ESTIMATING POSTMORTEM INTERVAL USING VNIR SPECTROSCOPY ON HUMAN CORTICAL BONE

John A. Servello, M.S.

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

May 2018

APPROVED:

H. Gill-King, Committee Chair
Samuel Atkinson, Committee Chair
Rebecca Dickstein, Committee Member
Claire Kirchhoff, Committee Member
Elizabeth Peacock, Committee Member
Art Goven, Chair of the Department of Biological Sciences
Su Gao, Dean of the College of Science
Victor Prybutok, Dean of the Toulouse Graduate School

Servello, John A. *Estimating Postmortem Interval Using VNIR Spectroscopy on Human Cortical Bone*. Doctor of Philosophy (Biology), May 2018, 187 pp., 3 tables, 41 figures, references, 106 titles.

Postmortem interval (PMI) estimation is a necessary but often difficult task that must completed during a death investigation. The level of difficulty rises as time since death increases, especially with the case of skeletonized remains (long PMI). While challenging, a reliable PMI estimate may be of great importance for investigative direction and cost-savings (e.g. suspect identification, tailoring missing persons searches, non-forensic remains exclusion). Long PMI can be estimated by assessing changes in the organic content of bone (i.e. collagen), which degrades and is lost as the PMI lengthens. Visible-near infrared (VNIR) spectroscopy is one method that can be used for analyzing organic constituents, including proteins, in solid specimens. A 2013 preliminary investigation using a limited number of human cortical bone samples suggested that VNIR spectroscopy could provide a fast, reliable technique for assessing PMI in human skeletal remains. Clear separation was noted between "forensic" and "archaeological" specimen spectra within the near-infrared (NIR) bands. The goal of this research was to develop reliable multivariate classification models that could assign skeletal remains to appropriate PMI classes (e.g. "forensic" and "non-forensic"), based on NIR spectra collected from human cortical bone. Working with a large set of cortical samples (n=341), absorbance spectra were collected with an ASD/PANalytical LabSpec® 4 full range spectrometer. Sample spectra were then randomly assigned to training and test sets, where training set spectra were used to build internally cross-validated models in Camo Unscrambler® X 10.4; external validations of the models were then performed on test set spectra. Selected

model algorithms included soft independent modeling of class analogy (SIMCA), linear discriminant analysis on principal components (LDA-PCA), and partial least squares discriminant analysis (PLSDA); an application of support vector machines on principal components (SVM-PCA) was attempted as well. Multivariate classification models were built using both raw and transformed spectra (standard normal variate, Savitzky-Golay) that were collected from the longitudinally cut cortical surfaces (Set A models) and the superficial cortical surface following light grinding (Set B models). SIMCA models were consistently the poorest performers, as were many of the SVM-PCA models; LDA-PCA models were generally the best performers for these data. Transformed-spectra model classification accuracies were generally the same or lower than corresponding raw spectral models. Set A models out-performed Set B counterparts in most cases; Set B models often yielded lower classification accuracy for older forensic and non-forensic spectra. A limited number of Set B transformed-spectra models out-performed the raw model counterparts, suggesting that these transformations may be removing scattering-related noise, leading to improvements in model accuracy. This study suggests that NIR spectroscopy may represent a reliable technique for assessing the PMI of unknown human skeletal remains. Future work will require identifying new sources of remains with established extended PMI values. Broadening the number of spectra collected from older forensic samples would allow for the determination of how many narrower potential PMI classes can be discriminated within the forensic time-frame.

Copyright 2018

By

John A. Servello

ACKNOWLEDGEMENTS

Thanks to the 2014 ASD Alexander Goetz Instrument Support Program committee for providing the spectrometer used for this research. Thanks as well for the generous financial support provided by forensic anthropology organizations MD&C and SOFA.

To Dr. H-Gill-King, your mentorship throughout my professional development cannot be understated, and accepting me into your lab has been a great honor for me; I am grateful for your years of guidance and will continue to seek it for many to come. To my committee: Dr. Dickstein, Dr. Atkinson, Dr. Kirchhoff, and Dr. Peacock, page restrictions limit me, but know that you have my deepest respect and appreciation for all of your direction and support.

I would like to acknowledge the assistance of Donald Campbell (ASD-PANalytical) and John Chartier (Camo, Inc.), who were invaluable to my research. Additionally, my thanks to the following people who provided access to archived remains and relevant demographic information: Dr. Wendy Potter McQuade, Mark Ingraham, Dixie Peters, and Linda Larose (UNTCHI); Dr. Heather Edgars, Kate Rusk, and Anna Medendorp (Maxell Museum); and Dr. Ross Zumwalt (formerly, NMOMI). To my colleagues and friends Wendy and Mark, thanks as well for your continual structural and moral support.

To my “silent consultant,” Dr. A. Frank Servello, thank you for your keen eyes, ears, interpretations, and motivational speeches. To the remainder of my family and friends who gave relentless encouragement, you have my profound thanks.

Lastly, to Karla Garcia, my beautiful and talented wife, and our amazing daughter Frances Caroline, thank you both for your love and patience, and for providing me the drive to get through this – I love you both more than you can ever know.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

2Cl_FN	Two-class forensic/non-forensic model
2Cl_F_1F_2	Two-class forensic-only model
3Cl_F_1F_2N	Three-class model
CHI	Center for Human Identification
FOV	Field-of-view
GLY	Glycine
HYL	Hydroxylysine
HYP	Hydroxyproline
IndicoPro	Indico® Pro Spectral Acquisition Software package version 6.3
IR	Infrared
LDA	Linear discriminant analysis
LDA-PCA	Linear discriminant analysis on principal components
LFA	Laboratory of Forensic Anthropology
LMI	Laboratory of Molecular Identification
LV	Latent variable
LYS	Lysine
MMA	Maxwell Museum of Anthropology
MIR	Mid-infrared
MSExcel	Microsoft Excel spreadsheet
NIPALS	Nonlinear iterative partial least squares algorithm
NIR	Near-infrared
NMOMI	New Mexico Office of the Medical Investigator
PC	Principal component
PCA	Principal component analysis
PLSDA	Projection of latent squares discriminant analysis (partial least squares discriminant analysis)
PLSDA-MC	PLSDA on mean centered data
PLSDA-CO	PLSDA on corrected data
PLSR	Projection of latent squares regression (partial least squares regression)
PMI	Postmortem interval

PRO	Proline
RMSECV	Root mean-squared error of cross-validation
SG	Savitzky-Golay transformation
SG1	Savitzky-Golay first derivative
SG2	Savitzky-Golay second derivative
SIMCA	Soft independent modeling of class analogy
SNV	Standard normal variate
SVM	Support vector machine
SVM-PCA	Support vector machine on principal components
SWIR	Short-wave infrared
TGA	Thermogravimetric analysis
Unscrambler	Camo The Unscrambler® X version 10.4
UV	Ultraviolet
ViewSpec	ASD Viewspec Pro™ version 6.2
VNIR	Visible and near-infrared

*Further elucidation of data transformations and statistical techniques is presented in the text body and within Appendix H.

CHAPTER 1

INTRODUCTION

Medicolegal examiners are tasked with the responsibility of investigating the deaths of individuals in order to ultimately certify a cause and manner of death. One of the more complicated aspects of a death investigation is the determination of the postmortem interval (PMI), or the length of time for which an individual has been deceased, based on the degree of decompositional changes observed on the decedent (Ubelaker 1997). Estimating the PMI is often a difficult proposition, especially as the length of time since death increases, because the rate of decomposition is dependent upon many variables. These include factors specific to the decedent (e.g. body mass, health status at time of death, presence or absence of clothing) as well as the recovery context (e.g. temperature, microenvironment, level of exposure, presence of insects, scavenger access). The resultant taphonomic profile for a decedent is thus the sum of all physical and chemical changes that have occurred during the PMI as a result of both intrinsic and extrinsic sources; because there are many taphonomic factors, the rate of decomposition is highly variable.

Although a challenge, providing a reliable PMI estimate is important for investigative agencies. An accurate estimate of PMI may have civil implications, including the settling of estates and the resolution of questions on statute of limitations. For criminal investigators, a decedent's time of death may lead to the inclusion or exclusion of suspects. Additionally, a reliable PMI estimate can be used to pare down a large pool of missing persons to a smaller, more manageable number of potential matches. Likewise, remains that are of historical or archaeological origin must be identified and excluded from further forensic investigation so

that limited resources (e.g. time, labor, money) are not expended on cases in which all related parties are long-deceased.

PMI estimation of skeletonized remains, or those remains which no longer have associated soft tissue, is often done in a *qualitative* manner. Living bone is largely a mineralized extracellular matrix of primarily collagen, a flexible fibrillar protein that is embedded with rigid and compression-resistant calcium hydroxyapatite crystals (Shipman et al. 1985; Hall 2005). Over a lengthening PMI, the properties of bone will change as the native organic and inorganic components degrade (Hedges et al. 1995, Collins et al. 2002, Hedges 2002, Trueman and Martill 2002); such changes may be used to estimate the PMI, provided they are both predictable and quantifiable.

An estimate of relative PMI may be assigned based on textural and olfactory information, or on observed gross and microscopic morphological changes (Behrensmeyer 1978, Yoshino et al. 1991, Hedges et al. 1995, Bell et al. 1996, Collins et al. 2002, Trueman and Martill 2002, Nashelsky and McFeeley 2003, Swift 2006). Results from simple tests, such as the measurement of specific gravity, the degree of fluorescence on sectioned cortex, and the presence of blood residues may be deployed as well (Berg 1963, Knight and Lauder 1969, Introna et al. 1999, Ramsthaler et al. 2009, Ramsthaler et al. 2011). Each of these scenarios is rooted in bone compositional changes: specific gravity may decrease from loss of hydroxyapatite and replacement with soil carbonate; diminished fluorescence of cortex is attributed to collagen loss; and, the presence of decomposed blood products (heme) should indicate potentially “modern” origins.

Unfortunately, descriptive techniques are subjective and highly dependent on the experience of the examiner (Swift 2006), and are therefore not reproducible. Likewise, common simple tests (e.g. cortical fluorescence, specific gravity) as noted above are variable in outcome and are not validated. This presents a problem in the medicolegal setting, especially when cases are adjudicated, because non-validated techniques will likely be found inadmissible in court. The development of reliable, validated methods that can be used to provide a less subjective PMI range for skeletonized material based on the chemical and physical properties of recovered bone is therefore paramount.

Of immediate importance is determining whether skeletonized remains are of modern origin and forensically significant or of non-forensically significant provenance (i.e. historical, archeological, or medical origins). The defined distinction between "modern" and "historical" is somewhat variable in the literature, often ranging from 50-plus to 75 years, to less than 100 years (Knight and Lauder 1967, Knight and Lauder 1969, Swift 2006). In application, remains assigned to one side of an arbitrary cut-off (e.g. less than 75 years) are forensically significant, while those above are not. For forensic investigators, a major determination for significance of older remains is whether the relevant actors, such as criminal suspects or decedents' next-of-kin, are still alive. Equally important is the condition of the recovered skeletal remains and how useful the material will be for DNA extraction. In light of this, identifying a reliable, reproducible, and consistent set of criteria for the distinction between significant and non-significant remains is of major importance. The processing of non-significant historical or archaeological remains in a forensic investigation represents wasted analytical time and cost because 1) the bone is often so badly degraded that the DNA yield is poor, and perhaps more

importantly, 2) there are likely no DNA reference samples available for comparison, all related parties likely being deceased.

The goal of this research is the development of a robust classification model which can be used to reliably assign unknown skeletonized remains to a descriptive PMI class. Training sets, comprised of spectral absorbance data that were collected from cortical surfaces of bone specimens with known or estimated PMI, have been used for model construction. Similar reference (test) sets have been applied to the constructed models to perform an external validation.

This project focused on comparing various classification methods in an effort to see which could be best used to build classification models that could accurately separate older skeletal material of historical or archaeological origins from forensically significant remains (i.e. 2-class models). Selected classification techniques included Soft Independent Modeling of Class Analogies (SIMCA), Linear Discriminant Analysis on Principal Components (LDA-PCA), Support Vector Machine Classification on Principal Components (SVM-PCA), and Partial Least Squares (a.k.a. Projection of Latent Squares) Discriminant Analysis (PLSDA); these techniques are discussed further in Chapter 3. In addition to 2-class models used to assign remains to "non-forensic" or "forensic" classes, smaller subsets with *known* PMI were also used to develop models for the discrimination of 3-PMI classes ("newer" forensic cases, "older" forensic cases, and non-forensic remains) and 2-forensic classes (newer and older "forensic" remains, with non-forensic remains excluded). This was done in order to determine if absorbance spectra collected from bone are sensitive enough to reliably discriminate PMI range within the forensic

time frame. Ultimately, the development of a validated PMI technique would provide significant time and cost-savings for all parties involved in a death investigation.

This document has been organized into 6 primary chapters followed by 8 appendices and references. Chapter 1 serves as an introduction to the PMI, the importance of PMI within the medicolegal context, and the lack of robust, validated techniques available for PMI assessment, especially when dealing with skeletonized remains. Chapter 2 comprises a literature review and is organized in two primary sections. The first section presents the stages of the PMI as characterized by observed decompositional change. The second section of this chapter presents multiple PMI estimation techniques that have been attempted; an introduction to vibrational spectroscopy (VR) and a review of relevant VR applications using bone and cartilage are presented in this section as well.

Chapter 3 presents the results of preliminary research and the methodology developed for the current work. Equipment specifications and software selection are presented. Bone sample selection and preparation is described, followed by spectral data collection protocol. The final section of this chapter discusses the processes followed for the construction and validation of classification models, by algorithm.

Results of the research are presented in Chapter 4. The chapter begins with a justification for the selection of the spectral data sets used for model building, sample exclusion criteria, and applied data preparation techniques. The next section of this chapter presents the results of select classification models with associated external validation values. The final portion of this chapter reviews the results of tests conducted to determine if bone collagen in

extracted cortical samples is subject to oxidative change. A discussion of the results is presented in Chapter 5.

Chapter 6 concludes the text. The first section of this chapter presents the positive impacts that this technique would have if it can be successfully applied within both the medicolegal context and the examination and identification of remains from mass death scenarios such as contemporary and historic conflicts. Future avenues of research are presented in the closing section of this chapter.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Postmortem changes have been described from the gross anatomical to the subcellular and chemical levels, and can be broadly grouped by relative length of time into short PMI, intermediate PMI, and long PMI (Nashelsky and McFeeley 2003). This chapter will be comprised of two major sections. The first section will review the decompositional that often occur during each of these PMI stages and the limitations associated with PMI assessment. The second portion of the chapter will focus on a review of the techniques that have either been attempted or are in active use for the estimation of PMI, with an emphasis on the long PMI period.

2.2 Stages of the PMI

2.2.1 Short PMI

Short-term PMI changes (hours to days) include phenomena that begin to occur immediately following death (Nashelsky and McFeeley 2003). These processes include the onset of and disappearance of rigor mortis (stiffening of the musculature), settling of the blood to dependent portions of the body (livor mortis), and cooling of the body following cessation of metabolic activity (algor mortis). These early indicators usually occur during the first 1-2 days of the PMI (Nashelsky and McFeeley 2003, Henßge and Madea 2004, Perper 2006). Each may be taken in to account when assessing PMI, but none can be used definitively because they are highly dependent on corporeal and environmental factors. Other short-term biological indicators include changes in vitreous humor composition, such as increasing [K+] and

[hypoxanthine] (Nashelsky and McFeeley 2003, Henßge and Madea 2004, Perper 2006), although these may be problematic. Swift (2006) argues that the typical range of vitreous [K+] in living individuals may not be well-defined owing to sampling difficulty; thus, an elevated postmortem [K+] may be called into question. Additionally, an assessment of gut contents (i.e. "last meal") is regularly made in order to provide a rough estimate of PMI because the stomach is often assumed to empty at a predictable rate; however, this process is highly variable by individual (Nashelsky and McFeeley 2003, Perper 2006, Swift 2006).

2.2.2 Intermediate PMI

Decomposition processes begin during the intermediate-term (days to weeks) and are the result of both autolytic activities, which initiate at the onset of death, and putrefactive changes, which are driven by *in situ* and transmigrating gut bacteria (Nehring et al. 1972, Nashelsky and McFeeley 2003, Perper 2006). During this period, the body is often characterized by discoloration and bloating as a result of bacterial proliferation and gas release (e.g. methane, hydrogen sulfide; Gill-King 1997, Nashelsky and McFeeley 2003, Perper 2006). The superficial blood vessels may take on a dark purple stain as bacteria and decomposition gases infiltrate the vascular network and erythrocyte components are broken down, giving the body a "marbled" appearance (Gill-King 1997, Nashelsky and McFeeley 2003, Perper 2006). Additional autolytic changes include sloughing of the epidermis from the deeper dermal layer (skin slippage) and purging of fluids from bodily openings as visceral tissues break down and body cavity pressure increases from gas release. These decompositional changes may begin within the first days of

PMI and extend for a few weeks, depending upon the environment (e.g. ambient temperature, location).

Anamnestic evidence may also be useful for determining the PMI in the short- and intermediate-terms, when the life habits of the decedent are known (Nashelsky and McFeeley 2003). Such evidence focuses on knowledge of the decedent's normal life routines (e.g. retrieving the morning paper, arrival and departure times from work, daily trip to a local eatery). Identifying a window when an individual's normal habits have been interrupted may provide investigators a better indication of when that person died, assuming the death disrupts a timeline composed of known, repeated events.

Necrophagous insects, species that subsist on decomposing tissues, may also provide a robust method for determining PMI in the short- and intermediate-terms (Nashelsky and McFeeley 2003, Perper 2006). When allowed access, flies (Dipterans) are known to oviposit on exposed points of entry to the body (e.g. orifices, eyes, wound tracks) minutes following death (Goff 2009). Likewise, Coleopterids (beetles) will colonize remains later in the PMI (Goff 2009). Forensic entomologists may be consulted to identify the species of insects present, assess the developmental stages of the larvae, and determine the number of generations that have developed on the remains in order to provide a relatively narrow PMI window.

2.2.3 Long PMI

Given sufficient time and the right environment, the body may progress through the above series of soft tissue decompositional changes to essentially complete skeletonization (Clark et al. 1997, Nashelsky and McFeeley 2003, Perper 2006). Generally, following the

decomposition of organ tissues, the integument and connective tissues (periosteum, articular cartilages, tendons and ligaments) will decompose until they are eventually lost, and the bone itself may begin to break down from environmental exposure and the action of microorganisms (Behrensmeyer 1978, Gill-King 1997, Nashelsky and McFeeley 2003, Swift 2006). PMI estimation is difficult in the absence of some of the earlier rate-limited changes which characterize the short-term PMI (e.g. fixation of lividity, loss of rigor, cooling of the body to ambient temperature) and the benchmarks characterizing the intermediate-PMI (e.g. bloat, skin slippage, marbling), and it becomes increasingly difficult as death becomes more remote (Gill-King 1997, Nashelsky and McFeeley 2003, Swift 2006).

Additionally, microenvironment will directly influence long-term decomposition (Ubelaker 1997). It is reported that during the long-term PMI, remains may become skeletonized after many months (Ubelaker 1997); however, the process can occur rapidly (weeks) if the decedent is located in a humid environment. Alternatively, remains found in arid climates may rapidly become mummified due to fluid loss (Nashelsky and McFeeley 2003). The lower moisture content in turn retards bacterial decomposition and insect activity (Goff 2009); thus, some desiccated soft tissue may persist for many months to potentially years before remains become skeletonized (Galloway 1997).

Another long-term postmortem effect is the formation of an adipocere coating on the remains. Commonly referred to as “grave wax,” adipocere is formed from the hydrolysis of fatty acids (Nashelsky and McFeeley 2003, Perper 2006). When encasing remains, adipocere may greatly retard the rate of decomposition (Micozzi 1991). The tendency to saponify (form adipocere) is directly related to body composition (i.e. total adipose content). Common

adipocere locations are found in areas of high lipid content, such as brain and viscera (Perper 2006). Although saponification generally occurs when a body is sequestered in a wet environment, such as a body of water or a burial in water-rich soils, it may also develop on remains wrapped in protective layers of natural fabric or in loose, silty soils that are found in arid environments. Acting as a wick, these draw out tissue fluid and drive surface condensation of adipocere; likewise, the loss of moisture retards bacterial-driven decomposition, allowing soft tissues to persist. Thus, partial skeletonization, mummification, and adipocere potentially may all be found on one decedent (Nashelsky and McFeeley 2003).

Given access, scavengers may also reduce a decedent to skeletal elements in a short period of time (Ubelaker 1997). It should also be noted that scavengers often relocate portions of the remains from the original body deposition site. Thus, portions of the same decedent may be spread out in varying microenvironments (e.g. shaded area, in the sun, partially buried, in water), resulting in differential decompositional changes on the recovered material. In short, as PMI lengthens, it becomes increasingly difficult to assess in the extended time frame relative to the somewhat predictable postmortem benchmarks characterizing the short- and intermediate-terms.

Long-term taphonomic changes in bone, including diagenesis, are well-documented if not completely understood (Hedges et al. 1995). Diagenesis is as an exchange of individual ions or charged polyatomic complexes with different complexes. Acting through a watery interstitial environment, exchanges between soil and bone entail the postmortem chemical and physical changes in the organic (primarily collagen) and mineral (apatite) phases of osseous tissue that occur over many decades to millennia. These may include bone surface erosion, histological

degradation, changes in bone porosity, infiltration (soil, water, fungi, bacteria, and rootlets), and mineralization (Hedges et al. 1995). The extent and types of change are highly dependent upon the deposition environment (Hedges et al. 1995). Often, bone apatite is lost, and the subsequent infiltration of water, soil minerals, and microbes further drives the loss of both organic and mineral phases through replacement reactions. Along with time-dependent decomposition of collagen, the addition of soil fungi and bacteria accelerate the degradation of collagen. Bone mineral is replaced by passive means (i.e. concentration gradients) as minerals are incorporated from the surrounding soils, ultimately leading to recrystalization (Collins et al. 2002, Hedges 2002, Trueman and Martill 2002).

The PMI may extend into antiquity; however, forensic investigators are generally interested in "modern" remains with a PMI typically less than 50-75 years (Knight and Lauder 1969, Swift 2006). While investigators are usually consulted on medicolegal cases, remains of historical and archaeological origin are frequently recovered. Forensic anthropologists must determine whether or not such remains are of forensic interest. Although the remains of a 500 year deceased Native American may be of scientific interest, such remains are almost never forensically significant.

It is evident that in the absence of most or all of the soft tissue, gross and microscopic changes in bone will characterize the long PMI (Behrensmeyer 1978, Yoshino et al. 1991, Hedges et al. 1995, Bell et al. 1996, Gill-King 1997, Collins et al. 2002, Trueman and Martill 2002, Nashelsky and McFeeley 2003, Swift 2006). A review of normal bone composition is presented in Appendix A. Emphasis is placed on bone phases (mineral, organic), the general organization of compact bone, and the hierarchical organization of bone collagen. Techniques

used to assess these changes in an attempt to estimate PMI in the long-term are presented in section 2.2.2 below.

2.3 PMI Estimation Methods

The preceding pages introduce the various PMI stages in which a decedent may be discovered, as determined by the degree of observed decompositional change. The so-called short PMI ends fairly quickly, as each of the associated changes is rate-limited; likewise, a transition from the intermediate to the long PMI often occurs when most or all of the soft tissue "substrate" has been consumed. However, it has been noted that the long PMI is highly variable, and therefore the most difficult phase to assess. Because of this, numerous techniques have been attempted to solve the long PMI estimation problem. A review of some of the relevant PMI assessment techniques for each of these PMI stages is presented below.

2.3.1 Methods for Determining Early to Intermediate PMI

Much of the current anthropological research focuses on decompositional changes to soft tissues observed in the short- and intermediate-term PMI. Such studies include variability in decomposition patterns based on different environments, e.g. temperate versus desert (Bass 1997, Galloway 1997, Parks 2011), as well as the effects of deposition, including exposed, buried, or submerged remains (Rodriguez 1997). Documentation of scavenger activity, including species type, activity times, and scattering patterns have been analyzed (Haglund 1997, Reeves 2009), as have the effects of necrophagous insect colonization and species progression (Haskell et al. 1997, Goff 2009). The release of early- and intermediate-PMI biochemical decomposition

products into the soil from the breakdown of organ tissues has also been examined (Vass et al. 2002). Another avenue of interest for predicting PMI is the examination of microorganism species progression on and in decomposing remains, referred to as a "microbial clock" (Metcalf et al. 2016). Other strategies involve quantifying decompositional changes and associating these with relevant environmental variables (e.g. temperature, humidity, soil pH) in an attempt to build predictive models that can provide reliable estimated PMI ranges (Megyesi et al. 2005, Vass 2011). Such PMI formulae are summarized in Appendix B, Section 1. Note that these approaches may prove useful for narrowing the PMI estimate during the earlier stages of decomposition; however, they are of diminishing use as skeletonization occurs, and as the bone continues to break down with time.

2.3.2 Methods for Determining Long PMI

2.3.2.1 Qualitative Morphological Assessment

Long-term PMI estimation of skeletal remains is often a "best guess" based on the experience of the anthropologist or pathologist, a familiarity with the recovery context, and the results of the forensic investigation (i.e. potential concurrence information). Swift (2006) notes that long term PMI is often judged qualitatively by a visual morphological assessment. Relatively recent skeletonized remains may be recognized by the presence of soft tissue remnants such as integument and connective tissues, exuded grease, the expression of decompositional odor (byproducts of putrefaction, including putrescine and cadaverine), as well as associated live insects (Gill-King 1997, Nashelsky and McFeeley 2003, Swift 2006, Goff 2009). Bones at this stage are also often characterized as being relatively heavy and firm.

Lengthened PMI may lead to the loss of grease, soft tissue remnants, and periosteum, as well as diminishing odor (Swift 2006). Over time, the bone may take on a weathered appearance characterized by superficial exfoliation and splintering (Behrensmeyer 1978, Nashelsky and McFeeley 2003, Swift 2006). When buried, the invasion of bone by plant microrootlets provides access for microorganisms to the organic and inorganic components of bone, which facilitates degradation via secreted collagenases and phosphatases (Child 1995, Mullen 1998). This ultimately results in bone becoming increasingly "chalky" when sectioned, relatively light-weight, and increasingly friable, as PMI lengthens.

Sectioned bone can also be examined microscopically to assess PMI. Long-term damage to bone may be characterized by focal points of destruction, tunneling by microorganisms, and loss of normal bone histology (Hedges et al. 1995, Collins et al. 2002, Trueman and Martill 2002). Berg (1963) noted changes in optical activity with respect to PMI when bone is examined under polarized light. A decrease in collagen periodicity was observed using electron microscopy, with a shrinkage approximately 640Å (64nm) in relatively fresh bone to as low as 500Å (50nm) in Miocene fossil material (Shackleford and Wyckoff 1964, Shackleford 1966, Race et al. 1968), although the pattern of change was not sufficiently accurate to estimate the age fossil material (Shackleford 1966). Later electron microscopy studies attempted to describe the effects of the environment on both the rates and patterns of postmortem histological damage (Yoshino et al. 1991, Bell et al. 1996). See Appendix B, Section 2 for additional information about microscopic inspection.

Although gross examination, odor, and microscopic analyses of bone provide the means for roughly estimating PMI, these techniques are highly variable and non-quantifiable.

Additionally, interpretation of material conditions is highly dependent upon the experience of the individual, thus raising the question of interobserver variability. Lastly, the general condition of remains is not only the result of the amount of time that has passed, it is greatly affected by the physical and biological characteristics of the environment from which the remains were recovered.

2.3.2.2 Qualitative Tests

Other qualitative methods commonly used to assign remains to the "non-forensically significant" category often consist in assessments of lost collagen and/or the infiltration of sediment resulting in bone mineralization. These include numerous methods reported by Berg (1963), including the following: soil carbonate infiltration as indicated by increased reactivity of cortex with hydrochloric acid; collagen loss, demonstrated by diminished fluorescence of sectioned cortex under UV light and decreasing cortical specific gravity; and, staining with Nile Blue (mineral affinity) and Indophenol (affinity for proteins and vitamin C associated with collagen).

Nitrogen and amino acid yields from bones of different ages have also been examined (Knight and Lauder 1969) as well as attempts using benzidine staining to identify blood residues in bone (Knight and Lauder 1969, Facchini and Pettener 1977). Amino acid racemization may also be used for aging bone (Bada and Helfman 1975), although it should be noted that this method is only potentially applicable when the PMI is in the 1000's of years or greater. These techniques are summarized in Appendix B, Section 3.

2.3.2.3 Quantitative Methods

2.3.2.3.1 Luminol Reaction

Luminol can also be used to identify traces of blood. Luminol is oxidized when exposed to heme, an iron-containing porphyrin in blood; the catalyzed reaction drives the release of light (chemiluminescence) which can be viewed under an alternative light source such as ultraviolet light (Quickenden and Creamer 2001). Assuming that blood remnants in bone should decrease with increased PMI, multiple researchers have attempted to expose bone to luminol and quantify the reaction intensity (Introna et al. 1999, Ramsthaler et al. 2009, Ramsthaler et al. 2011). Although a stronger reaction is often observed in newer remains, older but still forensically significant remains can show little to no reaction, while some non-significant remains have been found to be mildly reactive. This suggests that PMI assessment cannot be based on the presence or absence of luminol reaction alone. These studies are summarized below in Appendix B, Section 4.

2.3.2.3.2 Autofluorescence

Fluorescence of sectioned bone under ultraviolet light has been previously described (Berg 1963). This phenomenon is the result of autofluorescence, which occurs when endogenous fluorophores (e.g. the cross-links and aromatic amino acids found in collagen) absorb higher energy photons (UV light) and become excited; upon return to a relaxed state, lower energy photons are emitted or "fluoresced" (Monici 2005). Although recent reexaminations of autofluorescence have noted a negative correlation between the degree of fluorescence and PMI as well as change in fluorescent color over the PMI (Ramsthaler et al.

2011, Hoke et al. 2013), autofluorescence is best used as a coarse indicator of PMI and must be considered in conjunction with the recovery context and the condition of the remains. See Appendix B, Section 5 for further detail.

2.3.2.3.3 Thermogravimetric Analysis

Thermogravimetric analysis (TGA) involves incremental heating of a material within a furnace, followed by measurement of the associated sample weight change as burned sample components are evolved (Coats and Redfern 1963). Attempted use of TGA to examine PMI is reviewed in Appendix B, Section 6 (Villanueva et al. 1976, Raja et al. 2009). While the technique may work well for estimating PMI for newer remains, both reported studies suggest that it may be limited for estimating long-term PMI. Additionally, samples are completely destroyed by TGA.

2.3.2.3.4 Citrate Concentration

Citrate (i.e. citric acid) comprises 1.5-2.0% of weight in living bone and is thought to play a role in inhibiting excess hydroxyapatite deposition during new bone ossification. Schwarcz and colleagues (2010) proposed using bone citrate content as a PMI indicator. Their rationale was based on the following observations: 1) citrate has a relatively uniform distribution throughout bone in life, 2) the concentration is independent of both sex and age, and 3) previous research (Gibbs 1991) demonstrated that citrate is lost gradually during the PMI, with remains greater than 100 years of age typically yielding less than 1.0% of the original citrate content in living bone. Results presented by Schwarcz and others indicated that citrate loss may

be a good PMI indicator. However, a later attempt testing the technique on unembalmed remains from exhumed cemetery burials was not as successful (Kanz et al. 2014). See Appendix B, Section 7 for additional details.

2.3.2.3.5 Radionuclide Techniques

Because radioactive isotopes decay at predictable rates irrespective of environment, quantifying the levels of radionuclides in mineralized tissues (e.g. teeth and bone) could theoretically provide a robust method for determining PMI. Natural isotopes (e.g. ^{210}Pb, ^{210}Po) and nuclear fission products (e.g. excess ^{14}C, ^{90}Sr) have been considered for dating forensically significant material (Swift 1998, Neis et al. 1999, Swift et al. 2001, Schrag et al. 2012, Ubelaker 2014, Speller et al. 2012, Schrag et al. 2014). Applications involving fission products are based on the presence of artificial isotopes released in the fallout from above-ground nuclear bomb testing conducted in the 1950s and early 1960s. Such isotopes were distributed worldwide and ultimately integrated into human tissues; thus, their presence could potentially be used to differentiate non-forensically significant remains from modern materials. A summary of radionuclide techniques are presented in Appendix B, Section 8.

Although attractive, radionuclide methods have significant practical limitations. Useful, naturally occurring radioisotopes are not uniformly distributed, so geographically specific curves must be developed (Swift 2006). Such techniques also assume that a decedent incorporates radioactive isotopes from the environment at a constant rate, which is not necessarily true when considering migration. Regarding artificial radionuclides, the global distribution is also non-uniform because much of the testing occurred in the northern

hemisphere. Of equal concern is the fact that atmospheric levels of artificial radionuclides generated from 1950s and 1960s era above-ground testing are continually decreasing. Because the time interval between testing and the present continues to increase, these methods will only be potentially useful for discriminating forensically significant from historical remains materials for, at most, another 20 years (Buchan and Anderson 2001). Lastly, radionuclide methods are time-consuming and destructive, the equipment is prohibitively expensive, and extensive, specialized knowledge and training are required for operation and interpretation. Such limitations make these techniques impractical for most investigative agencies.

2.3.3 Vibrational Spectroscopy Techniques

Each of the previously reported quantitative techniques illustrates the need for developing reliable methods for estimating PMI in the long-term. One analytical area not fully explored is vibrational spectroscopy. Although frequently used in various forms within industrial and medical fields (Murray et al. 2001, Boskey and Camancho 2007, Spahn et al. 2007, Spahn et al. 2008, Nagy et al. 2008, Afara et al. 2013), these methods thus far have been applied in a limited fashion to assessing PMI within the forensic context (Dogra 2009, Howes et al. 2012, Patonai et al. 2013). The following section provides an introduction infrared (IR) and near-infrared (NIR) spectroscopy, and relevant applications using cartilage and mineralized tissues.

2.3.3.1 Overview of Vibrational Spectroscopy

Vibrational spectroscopy involves the study of material characteristics via the absorption of light (Seisler et al. 2002), taking advantage of covalent bond vibration transitions

between relaxed (lower energy) and excited (higher energy) states. When a bond is exposed to broadband light, most wavelengths will be transmitted; however, a relaxed bond may absorb a photon of light if the energy of the wavelength equals the bond energy vibration, resulting in excited bond vibrations, such as symmetric or asymmetric stretches, bends, and scissoring (Seisler et al. 2002). Upon emission of the photon, the bond will return to a relaxed state. If absorbance of a photon produces a change in the dipole moment (e.g. an asymmetric stretch in CO_2), there is an associated absorbance peak in the infrared (IR) spectrum. These same stretches occur in the functional groups of large molecules such as proteins (e.g. C=O, NH, CH, OH); accordingly, IR spectra are useful for analyzing these proteins and associated functional groups (Seisler et al. 2002).

2.3.3.2 MIR Spectroscopy and Applications with Mineralized Tissues

Alternate techniques are used to examine different parts of the infrared spectrum. One method, IR spectroscopy, focuses on the mid-infrared (MIR). Since most of the absorbing bonds begin in a relaxed state (energy level n=0), photon absorption will elevate the majority to the first excited state (n=1) prior to reemission (Seisler et al. 2002). Thus, MIR spectra are characterized by well-defined “fundamental” peaks for functional groups of interest (Seisler et al. 2002).

MIR methods have been used to assess bone in medical applications to differentiate between natural and engineered bone (Boskey and Camancho 2007), examine cortical and trabecular bone (Paschalis et al. 1997), and for comparisons of pathological to healthy bone in human skeletal remains (Nagy et al. 2008). MIR methods have also been used to examine PMI-

related changes in bone organic and carbonate content as well as the degree of crystallinity (Howes et al. 2012, Patonai et al. 2013). These studies are reviewed in Appendix B, Section 9. MIR studies of bone have obvious strengths, most notably spectra composed of well-defined fundamental peaks that can be readily interpreted. Furthermore, these techniques can be used to study both the organic and inorganic bone phases. However, time-consuming sample preparation is required, potentially including dehydration, degreasing, and pulverizing. Additionally, the sample must generally be mixed with a reference material (e.g. KBr) and compressed into a pellet before analysis. Lastly, samples under study must be relatively thin in order to allow for the transmission of light as the longer, lower energy MIR wavelengths cannot penetrate the sample as deeply as shorter wavelength light.

2.3.3.3 NIR Spectroscopy and Applications with Mineralized Tissues

Although the bulk of absorbance events result in elevation from ground to the first excited state, higher levels (e.g. n=2, n=3) also exist. Elevation to one of these excited levels or a position between these levels produces overtone and combination bands. Because such excitations are relatively small in number with respect to the fundamentals, the corresponding overtones that appear in MIR spectra are generally reduced or overwhelmed by fundamental peaks. However, these bands represent the dominant features in near-infrared (NIR) spectra and are the focus of another method, NIR spectroscopy (Seisler et al. 2002, Manley 2014).

NIR spectra are composed of overtone bands (e.g., n=1 to n=2) associated with an absorbed wavelength of light which causes a change in dipole moment (e.g., asymmetric stretch) and/or combination bands, which consist in overlapping transitions between higher

energy levels. Displayed NIR spectra consist of broad, overlapping peaks and are difficult to interpret directly; this is a primary limitation of NIR spectroscopy (Seisler et al. 2002, Manley 2014). However, NIR spectroscopy is appealing because the shorter wavelength light can penetrate deeper into the sample than other vibrational techniques, allowing for interaction with the sample constituents (Seisler et al. 2002, Manley 2014). Additional strengths include minimal to no sample preparation, rapid data collection, and simple use with minimal prior training. To compensate for the less-defined spectra, NIR spectroscopic data are typically incorporated into chemometric models (Naes et al. 2002, Seisler et al. 2002). Such models may be predictive and constructed from a subset of known samples (e.g. NIR spectra and associated material characteristics such as pH and composition); spectra collected from unknown samples can be run through the model in order to characterize the sample. Alternatively, classification models can be developed to assign samples to specific classes based on spectral attributes.

NIR spectroscopy has been used to analyze a non-mineralized connective tissue, articular cartilage, for the relationship between water content and cartilage health (Spahn et al. 2007, Spahn et al. 2008), the prediction of collagen and proteoglycan content (Baykel et al. 2010), and the prediction of cartilage thickness (Afara et al. 2013). Published NIR applications with mineralized elements (e.g. bone and teeth) are limited. NIR spectroscopy has been used to identify the presence of bone and meat meals artificially mixed with fish meal (Murray et al. 2001). A related technique, NIR imaging spectroscopy, was used to map carious lesions on teeth (Zakian et al. 2009). Thomas and colleagues (2011) used a NIR spectrometer to discriminate a small sample of elk horn cores originating from two archaeological sites separated both geographically and temporally. Similarly, Linderholm and others (2013) used NIR imaging

spectroscopy to differentiate small, fragmented bone and tooth materials from other debris collected from a Bronze Age archaeological fire pit. Citing a need for a rapid method to identify ancient samples with adequately preserved organic content (collagen), Vincke and colleagues (2014) built predictive models using NIR imaging data and associated archaeological data for classifying zooarchaeological materials into two groups based on the presence of absence of organic content. The application of NIR to questions of PMI is currently limited to a dissertation (Dogra 2009) in which the researcher used NIR spectroscopy with chemometric modeling on a selection of defleshed porcine rib samples. This study is summarized in Appendix B, Section 10.

As noted, there is a paucity of NIR-related applications available in the literature that focus on bone characterization. However, the studies cited here collectively demonstrate that NIR techniques can be used to analyze bone and related mineralized tissues in general, and collagen content within bone specifically. Thus, NIR spectroscopy presents an appealing analytical technique for assessing the PMI of skeletonized human remains.

CHAPTER 3

RESEARCH DESIGN

3.1 Introduction

Living bone is a predictable composite of organic and inorganic components; it is therefore reasonable to assume that bone sampled from a recently deceased individual would display similar characteristics. Following death, the properties of bone change due to degradation as the postmortem interval lengthens, including the loss of organic content (primarily collagen) and the addition of environmental artifacts (e.g. soil carbonate, minerals) via diagenesis. Such changes should be accompanied by corresponding changes in the NIR spectra recorded from cortical bone samples representing a wide range of postmortem intervals.

3.2 Hypotheses

NIR spectroscopy can be used to indirectly analyze functional groups of organic compounds (e.g. O-H, SH, NH, amides). Because organic content is lost over an extended PMI, the following hypotheses are made:

- NIR spectra should change with PMI: Older, non-forensic remains will be spectrally distinct from more recent materials
- NIR spectra collected from cortical bone samples with a known or estimated PMI can be used to construct models that will accurately classify skeletal remains where PMI is not known
    - Developed models will accurately classify remains into 2-possible classes (Forensic, Non-forensic)
    - Developed models will accurately classify remains into 3-possible classes (Newer forensic, Older forensic, Non-forensic)

- Developed models will accurately classify forensic remains into 2-possible classes (Newer forensic, Older forensic)

3.3 Preliminary Work

A preliminary study was conducted in May 2013 at PANalytical in Boulder, Colorado, using a 350-2500nm LabSpec® 4 spectrometer on 20 samples of sectioned femoral cortical bone. Femoral samples examined in ASD Viewspec Pro™ 6.2 ("ViewSpec." ASD/PANalytical, Boulder, Colorado, USA. www.asdi.com) visualization software covered a wide range of PMIs broadly described as "recent" (≤ 10 years), "mid-range" (25 – 50 years), and "historic/archaeological." Examined as absorbance spectra, (i.e. pseudoabsorbance, log [1/Reflectance]) newer material tended to separate out first, followed by mid-range PMIs; the historical and archeological samples consistently fell out last. This pattern suggested an association between absorbance and PMI. Best separation was observed primarily within the NIR bands from roughly 1400 – 2500nm (Figure 1). Due to the low sample sizes, no statistical analyses were performed.

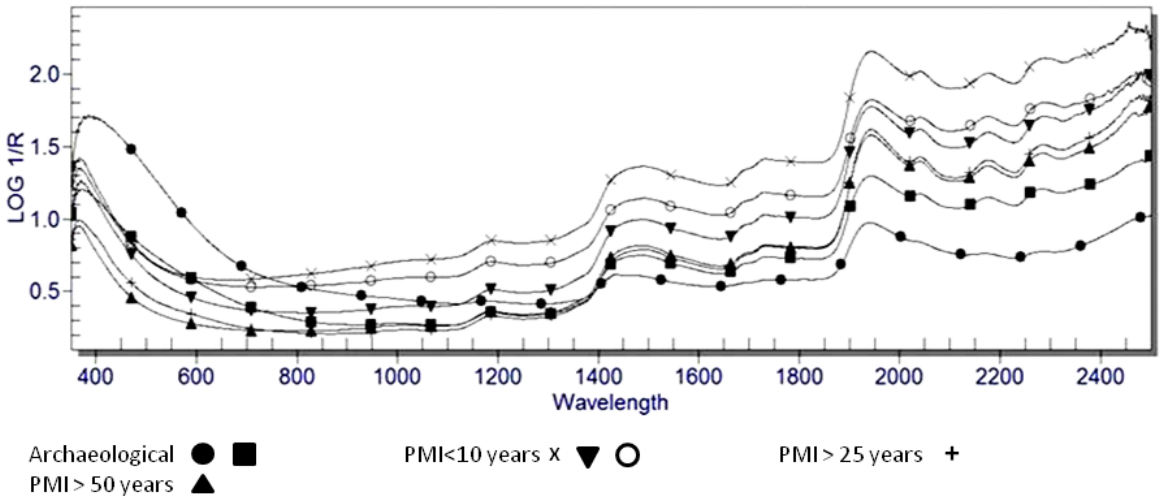


Figure 1. Selected absorbance spectra from the preliminary study (n=7). Relative PMI values by sample are presented in the legend. Curves separate best in the NIR bands (1400-2500nm) based on relative PMI. Diminishing absorbance is noted as PMI decreases.

To better examine this potential association, a subgroup of spectra from the current dataset (n=68) was plotted in Unscrambler® X version 10.4 ("Unscrambler." Camo Software, Oslo, Norway. www.camo.com). Selected spectra were those collected from the longitudinally cut surfaces of femora that is, cut surfaces running parallel to the long axis of the bone. These 68 samples included specimens with known PMI. Raw spectra were examined for areas of change and separation, and five peaks of interest were noted at around 1490nm (N-H 1st overtone region), 1730nm (S-H and C-H 1st overtone region), 1940nm (amide C=O 2nd overtone), 2040nm (O-H and amide combination bands), and 2160nm (N-H combination band). Spectra separated into two groupings: a narrow, low absorbing group comprising archaeological samples (PMI = 500-900 years), and a broader, higher absorbing group of spectra representing forensic cases with PMI ranging from days to 71 years (Figure 2 and Figure 3). Noted peaks of interest at 1490nm, 1730nm, 1940nm, 2040nm, and 2160nm were relatively flat in the archaeological spectra, which suggested that the relevant functional groups (NH, CH, amide carbonyl, and OH, respectively) were present in diminished amounts.

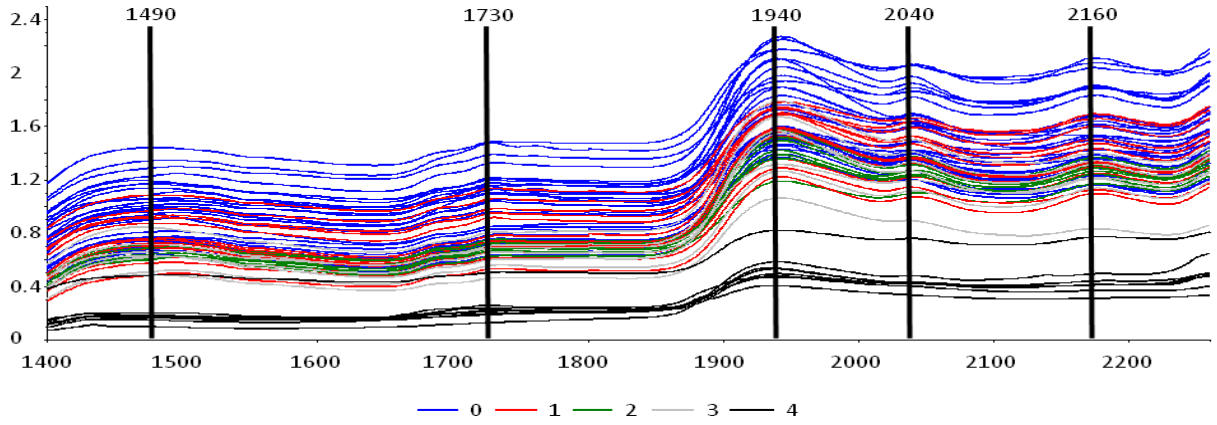


Figure 2. Absorbance spectra (1400-2260nm) collected from cases with known PMI, five classes presented (n=68). Classes: 0 (days – 2 years), 1 (2-10 years), 2 (10-20 years), 3 (20-71 years), and 4 (archaeological). Class 4 forms a low-absorbance thin band (500-900 years PMI). Peaks of interest (bars): 1490nm (N-H 1st overtone); 1730nm (S-H and C-H 1st overtone); 1940nm (amide C=O 2nd overtone); 2040nm (O-H and amide combination bands); and, 2160nm (N-H combination band). Peaks are relatively flattened in the archaeological samples (functional group loss).

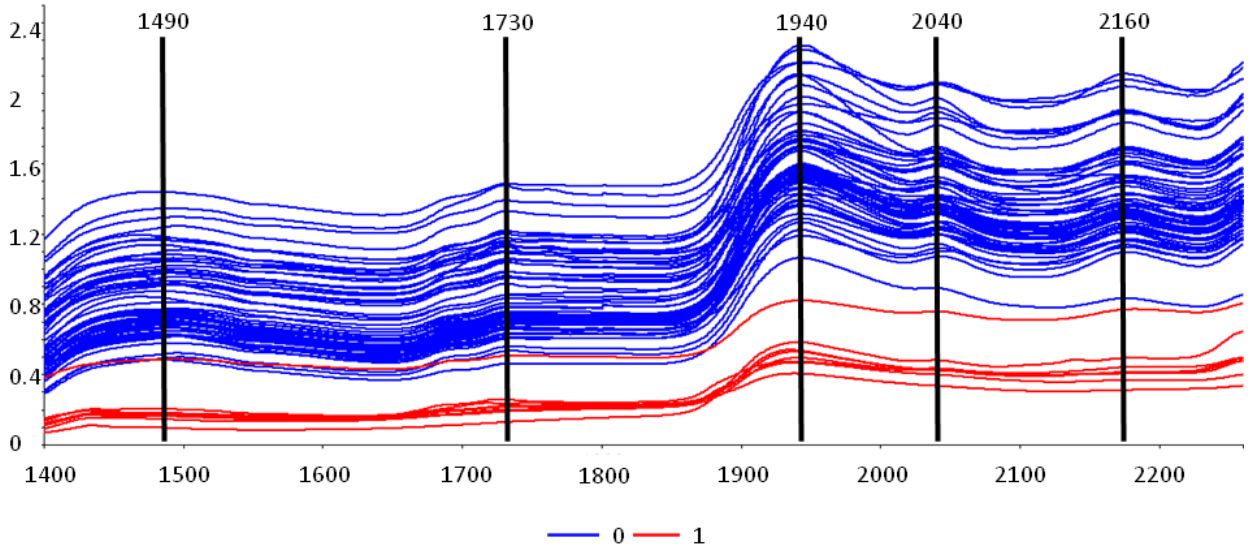


Figure 3. Absorbance spectra (1400-2260nm) collected from cases with known PMI, two classes presented (n=68). Peaks of interest are noted at 1490nm, 1730nm, 1940nm, 2040nm, and 2160nm. Classes include 0 ("forensic cases") and 1 ("archaeological").

Figure 4 presents a smaller collection of selected spectra from the larger group of 68, including 7 forensic cases with PMI ranging from "days" to 71 years, and 4 archaeological cases with established dates of 500 and 900 years, respectively. An additional spectrum collected from cremated bone is included. The abscissa (absorbance) appears to be correlated with time, with a notable decrease in absorbance and peak flattening as PMI increases. The cremains are added for comparison; note that in the cremated bone, all of the organic material has been burned off during the cremation process, resulting in a specimen of bone that appears spectrally "older" than the archaeological material.

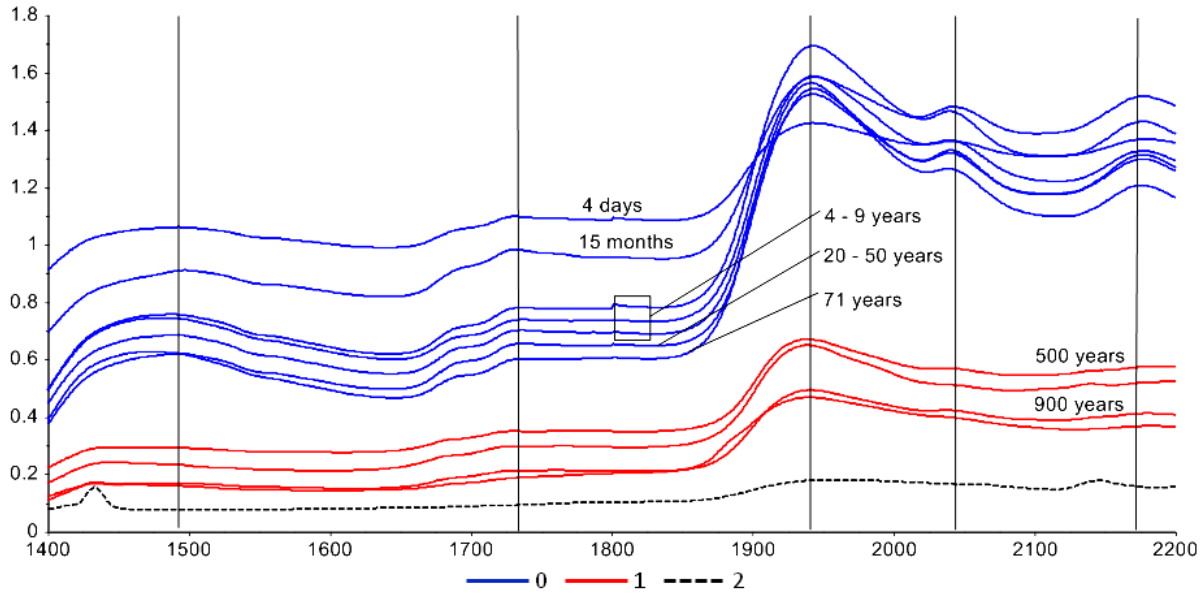


Figure 4. Raw spectra (1100-2300nm) collected from the longitudinally cut cortical surfaces of selected forensic (blue), archaeological (red), and cremated remains (dashes). PMI for the forensic and archeological samples is included. Peaks of interest are highlighted with dashed vertical lines. Note the diminishing peak heights with respect to PMI, indicating time-dependent loss of relevant functional groups. The same peaks are essentially flat on the cremains spectrum.

Principal components analysis (PCA) was performed on the larger group of 68 spectra. Using a narrower bandwidth (1400-2260nm), only two PCs were needed to account for 99.7% of the total variation in the dataset. Archaeological samples formed a cluster that was distinct from the remaining forensic samples. Although no clear separation between "forensically significant" PMI classes (i.e. PMI ≤ 71 years) was noted, the right side of the "forensic" cluster was dominated by remains with a shorter PMI (≤ 5 years) while older samples were found toward the middle, suggesting that time influences PC1 (Figure 5 and Figure 6). It should be noted that Unscrambler identified many of archeological samples as outliers. This suggested that relative to shorter-term PMI material, the properties of many forensically non-significant samples are so spectrally different that, statistically speaking, they may belong to a different population, and thus their presence may affect shorter-PMI class separation.

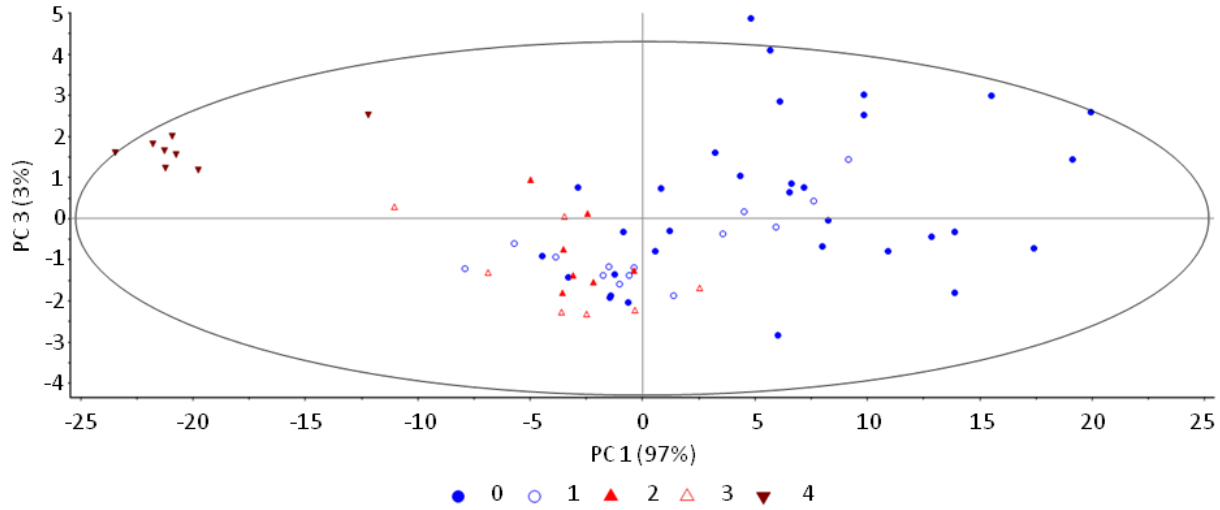


Figure 5. PCA on spectra (1400-2200nm) collected from femora with known PMI, four classes presented (n=68). Classes: 0 (days-2 years), 1 (2-10 years), 2 (10-20 years), 3 (20-71 years), and 4 (archaeological). Component 1 accounts for 97% of the total variation and appears to be influenced by time.

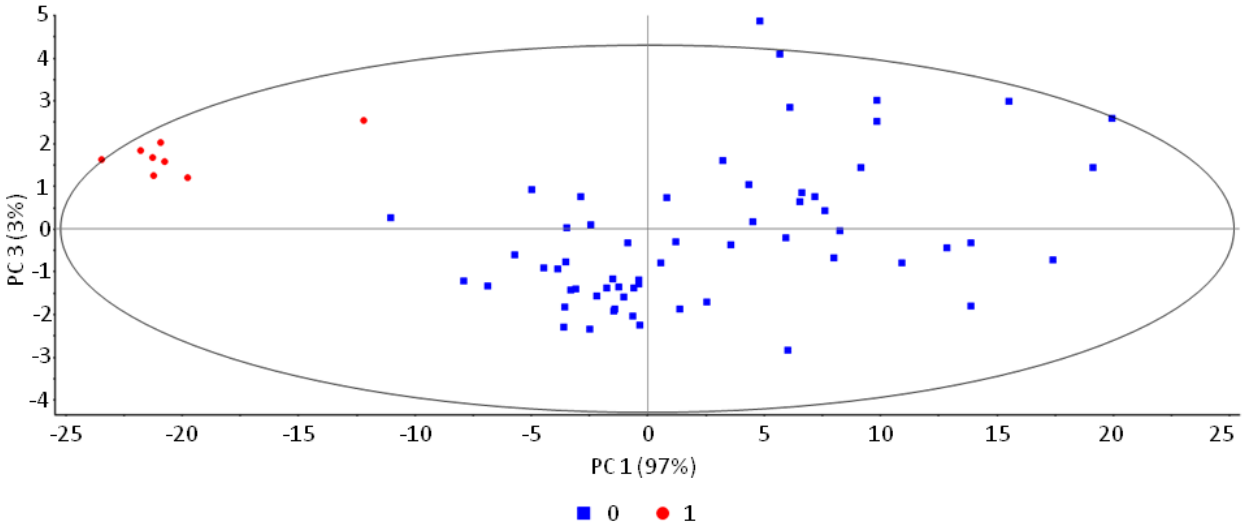


Figure 6. PCA on spectra (1400-2200nm) collected from femora with known PMI (n=68), two classes presented. Classes: 0 (“forensic cases,” 0-71 years) and 1 (“archaeological”).

3.4 Methodology

3.4.1 Instrumentation

An ASD/PANalytical LabSpec® 4 full range bench spectrometer was used to collect all spectral data[1]. The Indico® Pro Spectral Acquisition Software package version 6.3 ("IndicoPro." ASD/PANalytical, Boulder, Colorado USA. www.asdi.com) was used to control the instrument. The spectrometer collected spectral data in the 350-2500nm range. Instrument sampling intervals were 1.4nm and 2nm for the 350-1000nm and 1001-2500nm bands, respectively, with 10 spectra collected per second. The spectrometer was equipped with three detectors: a 512 pixel silicon photodiode array for the VNIR region (350-1000nm); a cooled indium-gallium-arsenide (InGaAs) photodiode for the SWIR1 region (1001-1800nm), and; an additional InGaAs photodiode for SWIR2 (1801-2500nm). Reported spectral resolutions were 3nm at 700nm and 10nm at both 1400nm and 2100nm. Spectral data were recorded with a bifurcated reflectance fiberoptic contact probe with a 4mm field of view (FOV); note that 4mm was the smallest FOV probe available through ASD/PANalytical, and for all but a limited number of cases was small enough to collect bone spectra. The probe was energized by an external light source (ASD Fiberoptic Illuminator) and was attached via two inputs to the light source and the spectrometer. This analysis equipment is depicted in Figure 7.

[1] Instrument was awarded through the 2014 Alexander Goetz Instrument Support Program and was made available from 4/18/2014 to 7/7/2014.

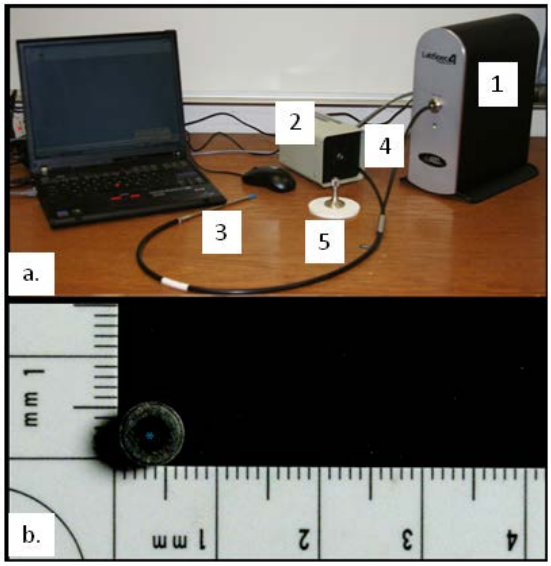


Figure 7. ASD/PANalytical LabSpec® 4 full range bench spectrometer and associated instrumentation. a) Typical instrument arrangement for data collection, labeled items include the following: 1. Spectrometer; 2. External light source; 3. Fiber optic contact probe; 4. Bifurcated fiber optic inputs for spectrometer (right) and light source (left); and 5. White reference. b) Collection end of the fiber optic contact probe with a 4mm field of view (blue asterisk).

3.4.2 Bone Sample Selection

The following skeletal collections were used as sources for bone samples: 1) new case submissions to the Center for Human Identification's (CHI) Laboratory of Forensic Anthropology (LFA); 2) previously sampled archived cases at LFA; 3) archived cortical samples at the CHI Laboratory for Molecular Identification (LMI); 4) previously sampled cold cases located at the New Mexico Office of the Medical Investigator (NMOMI) and the University of New Mexico Maxwell Museum of Anthropology Osteology Laboratory (MMA) in Albuquerque, New Mexico; and 5) fragmentary historical and archaeological long bones stored at the MMA with exposed broken edges (non-sectioned). A majority of the samples originated from various jurisdictions in Texas (n=150), New Mexico (n=105), and Nevada (n=32). To control for microenvironmental effects on decomposition, provenience data were collected, when available. It is noteworthy

that many of the samples were collected from remains found in the southwestern region of the United States.

A series of 341 human long bone diaphysis cortices representing a minimum of 290 individuals were spectrally sampled. Samples were collected from cases that were of forensic (n=278) and historical or archaeological origin (n=63). Those bones that were selected for sampling were visibly healthy with no evidence of pathology. Bone samples consisted of three types: 1) sections of cortex removed from a long bone via autopsy saw; 2) bones that had been previously sampled in similar fashion with exposed cortical windows; and 3) broken edges found on historical and archaeological material (Figure 8). Note that although the removal of a window of bone with a saw is a destructive process, the sample removed does not have to be processed further.

Cortex was selected over trabecular bone primarily because of its long-term durability, surface regularity, and decreased susceptibility to diagenesis (Haglund 1997). Typical trabecular bone, including diploë, vertebral bodies, and the epiphyses of long bones, is much more delicate than cortex. It is thus often fragmentary or absent in older sets of remains due to postmortem erosion and/or destruction by animal scavengers attempting to access marrow (Gill-King 1997, Haglund 1997).

Although some humeri (14) and tibiae (48) were sampled, femoral diaphyses (272) were selected as the primary cortical sample sites. Due to its relatively large size and dense construction, femora are frequently recovered; even a heavily-scavenged assemblage of remains will usually yield recognizable femoral diaphyseal fragments. Additionally, because the femur is a primary weight-bearing bone, the relatively large, dense, and thick femoral diaphysis

affords an extensive sampling area that is adequate for NIR spectral collection on both the superficial and sectioned surfaces when using a 4mm FOV contact probe.

The majority of modern, forensically significant cases examined were originally recovered from a surface microenvironment (n=184), although samples were also collected from remains originating in clandestine burials (n=47) and exhumed pauper's burials (n=5). When PMI was established for identified remains, the known PMI was recorded; otherwise, when available, an estimated PMI as reported in associated anthropology or pathology reports was included. Cortices sampled covered a broad range of PMI's, from relatively fresh remains in an early state of decomposition, to dated archaeological materials from New Mexico sites for Pottery Mound (1350-1500 C.E.) and Mimbres (1100-1150 C.E.) which represent PMI from approximately 500 to 900 old (Anyon et al. 1981; Franklin 2008). It should be noted that almost half of the sampled forensic cases have a known or estimated PMI of 2 years or less (n=132). All samples were coded with unique identifiers (001-344) in order to maintain the confidentiality of case-related information, including investigative agency, associated agency case number, and decedent identification, when applicable.

Figure 8. Representative examples of cortical samples. Samples consisted of three types, including: a) sections of cortex removed with an autopsy saw; b) previously sampled bones with exposed cortical windows; and c) broken edges on historical and archaeological material.

3.4.3 Acquisition of Spectra

Prior to each data collection session, the spectrometer and light source were allowed to warm up for 15 minutes, after which a standard baseline was established using a Spectralon® white optical standard disk. The optical probe was oriented at 90° to samples, which allowed the 4mm FOV to completely cover the surface of interest (Figure 9). A new baseline was taken every 10 minutes during each session. All bone samples were analyzed using a spectrum count of 20, meaning that the machine calculated a reflectance spectrum from 20 separate spectra. A subset of five reflectance spectra, each with a spectrum count of 20, was collected from a particular bone surface; the subset was then used to calculate a final average spectrum for output. Sampled bone was lightly wiped with a dry paper towel in order to remove adhered bone dust and dirt. Reflectance spectra were then collected from the following surfaces:

- Superficial bone surface
- Cross-sectional (transverse) cut surface of the cortex
- Longitudinal cut surface of the cortex

The superficial cortices for samples housed at LFA were then lightly ground with a stone wheel Dremmel® attachment in order to remove surface debris and weathered cortex. Following a wipe-down with dry fabric, a second spectral sample was collected from the ground surface (Figure 10).

It should be noted that not all surfaces could be sampled for every bone. Bones examined at the MMA and NMOMI could not be altered by grinding, and only a limited number provided adequately ground surfaces from prior molecular sampling. Additionally, many of the historical and archaeological samples at the MMA were coated superficially with preservative

sealant. Lastly, some of the samples had relatively narrow transverse sections; in such cases, the cross-sectional surface area was inadequate for covering the 4mm probe FOV, resulting in unusable spectra that were "saturated," where portions of the spectrum were outside of range (i.e. greater than 100% reflectance).

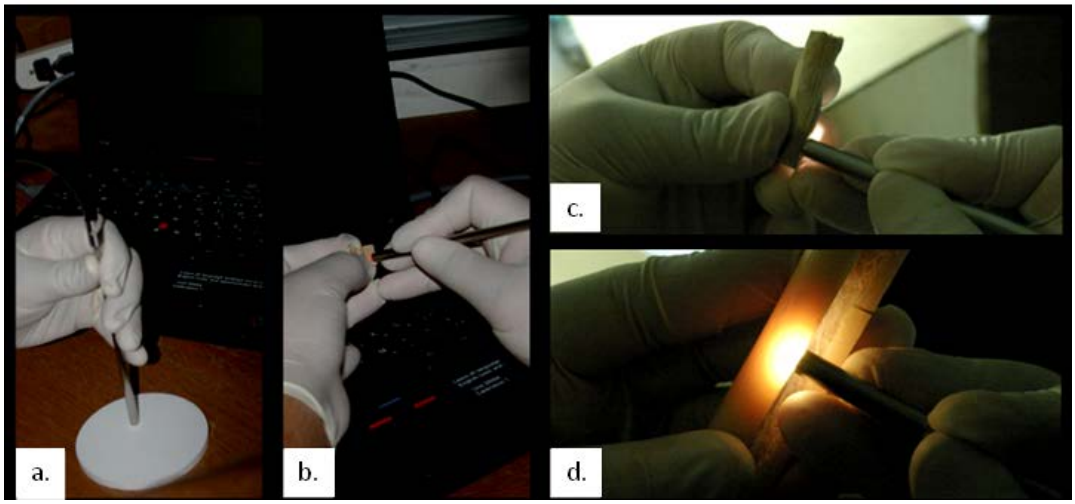


Figure 9. Fiber optic contact probe oriented 90° to samples. Presented images include a) baseline collection from a Spectralon® white optical standard disk and b) spectrum collection from the longitudinally cut surface of a cortical sample, as well as close-up images of spectral collection from c) a cortical sample and d) the exposed window in a previously sampled femur.

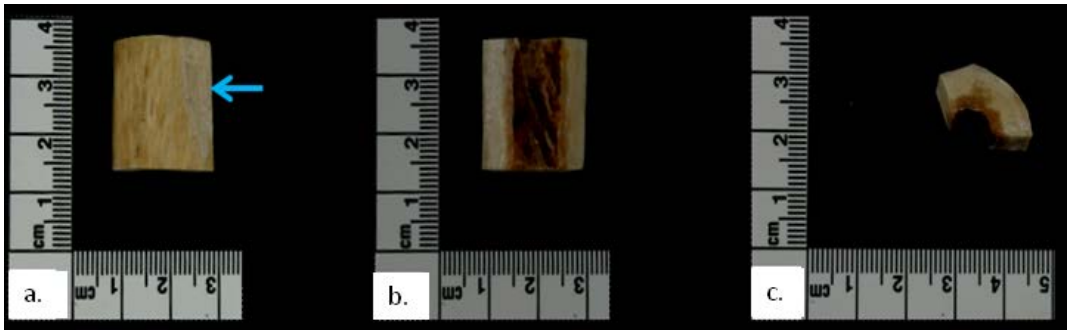


Figure 10. Cortical surfaces selected for spectral sampling. Spectra were collected from the a) superficial cortical surface, b) longitudinally-cut cortex, and c) transversely-cut cortex. Note that a portion of the superficial surface was subsequently ground down for an additional spectral sample (blue arrow).

3.4.4 Spectral Data Preparation

Spectral files (.asd) were imported in to Viewspec for visualization, data transformation, and export into text files (.txt). All spectra were exported as log-transformed

(pseudoabsorbance) spectra. Text files were then loaded into MSExcel, and relevant variable columns (sex, age, element, element side, deposition, geographic location, maceration, PMI/PMI estimate, and PMI Class) were added and populated for each sample.

3.4.5 Data Analysis

3.4.5.1 Software Selection

Due to a large number of highly intercorrelated variables, spectral data are best-analyzed using multivariate statistical approaches. Unscrambler, a platform designed to manipulate large datasets, was selected for this project. A full version of Unscrambler was purchased for a 1-year license (July 2015 - July 2016); a 1-year extension of the license was later purchased as well (March 2017 – March 2018). The primary strengths of Unscrambler include ease of user interaction, rapid processing time, and superior graphical output. Data were imported directly from MSExcel files into Unscrambler, where they could be subjected to a selection of frequently used preprocessing and statistical modeling methods (see following sections, this Chapter). Unscrambler is not a coding-based platform; however, selected procedures were customizable (e.g. sample and variable selection, sample exclusion with reanalysis, definition and analysis of subcategories within the dataset).

3.4.5.2 Exploratory Data Analyses

3.4.5.2.1 Spectral Selection

Raw spectra were plotted by surface in Unscrambler for visualization. Line plots of spectra were checked for specific samples that exhibited abnormal spectral morphology

relative to other samples. Although spectra with atypical shape may represent unique signatures of interest, they could signify poor spectral acquisition as well (i.e. bad samples). Such spectra would decrease the robusticity of any models for which they are randomly selected as training samples; likewise, if used as test samples, improper classification would artificially lower the model accuracy. Relevant information for samples with abnormal morphology, including original discovery context and condition of the remains when spectrally sampled, was reviewed to ascertain inclusion or exclusion from further analyses.

3.4.5.2.2 Data Pretreatment and Band Selection

The line plots were examined for bandwidths that exhibited changes in amplitude (i.e. peaks and valleys). Such features represent overtone and combination molecular vibrations of relevant functional groups; please see the chart of NIR absorption bands presented in Appendix C. Bands that were characterized by both shape change and separation between sample spectra were further investigated with PCA to determine if these regions were useful for discrimination.

Scattering-effect plots, a visualization tool available with Unscrambler, were generated as well. Scattering occurs when the reflection of light is non-specular, that is, the angle of reflectance is different from the angle of incidence (Jenson 2005). NIR spectra are often affected by additive and multiplicative scattering (Naes et al. 2002). The scattering effects plots produced in Unscrambler allow the user to examine spectra for evidence of additive effects, which are the result of variations of path lengths of light into different samples, and/or multiplicative scattering, which is generated from surface variability, including differences in

particle or grain size and texture. Spectra collected from solid samples, such as bone cortex, are prone to error resulting from light scattering off of sample particles. When the primary study goal involves studying chemical composition and not physical attributes, such artifacts must be removed (Naes et al. 2002, Rinnan et al 2009). Unscrambler scatter effects plots are generated by plotting each spectrum against the mean absorbance value at each wavelength; observed slope differences are indicative of multiplicative scattering, while baseline offset differences between spectra signify additive scattering (Brooke 2015). The data presented in these plots can assist the analyst in selecting appropriate pre-processing techniques that may remove most or all of the scattering (Brooke 2015). An examination of scattering-effect plots generated from some of the study samples revealed evidence of both additive and multiplicative scattering. Standard Normal Variate (SNV) and Savitzky-Golay derivation were thus explored as pretreatment methods (Rinnan et al 2009).

SNV is a commonly used preprocessing technique that can decrease scattering effects (Rinnan et al 2009, Brooke 2015). Considering a group of objects under study with associated spectra, an average spectral value (i.e. mean absorption) and standard deviation are calculated for that object. SNV corrected spectra are then calculated by subtracting the mean spectral value at each wavelength and then dividing by the standard deviation. This process is repeated for each object. Thus, the resulting dataset has been corrected for scattering and normalized as well (Rinnan et al 2009, Brooke 2015).

Derivations are transformation techniques that can correct for baseline as well as additive and multiplicative scattering (Rinnan et al 2009). Additionally, these techniques may be helpful in revealing hidden peaks that are in the areas of overlap on raw spectra (Brooke 2015).

Savitzky-Golay (SG) derivatives are commonly applied to spectral data (Rinnan et al 2009, Brooke 2015). The SG process involves defining a symmetric smoothing window of points on either side of a central point. A derivative is calculated for the central point by applying a polynomial function determined by least squares fitting to the smoothing points. This process is repeated for all points of the spectrum, resulting in a smoothed, derivatized spectrum; dependent upon the size of the window, an equal number of points are lost at the ends of the spectrum (Rinnan et al 2009). For this project, 1st and 2nd SG derivatives (SG1 and SG2) were applied to raw data. Output SG1 spectra were calculated using a 2nd order polynomial with 25 bilateral smoothing points (51 points total); SG2 spectra were derived with a 25 point symmetric window and a 3rd order polynomial (Rinnan et al 2009, Brooke 2015).

3.4.5.2.3 Principal Components Analysis

Principal components analysis (PCA) is an unsupervised technique that is often a first-step analysis used to identify patterns within datasets (Wold et al. 1987). In general, PCA reduces large, complex systems comprised of many X-variables that are highly correlated into a limited number of uncorrelated principal components (PCs) that can best explain the greatest amount of variation within the data while limiting noise. Thus, this technique can be used to both reduce the data dimensionality and potentially reveal underlying patterns in the data. This is especially useful in large, multivariate scenarios because 1) it is difficult to identify patterns when looking at large data tables filled with numbers, and 2) it is not possible to plot more than three variables at a time.

For a series of objects (i.e. cortical samples) that each have multiple associated X-variables (i.e. absorption values), PCA identifies the variables or groups of variables that account for the greatest amount of variation within the dataset: these are then used to compute a new directional vector (eigenvector) which defines PC1. The second PC (PC2) will be defined by the eigenvector that describes the second greatest amount of variation in a direction that is orthogonal to PC1 in the newly defined component space. These resulting PCs are then used to define the axes for a plot within component space as visualized on a PC scores plot; for example, see Figure 5 and Figure 6. Additional PCs can be calculated (i.e. PC3 through PC_{n-1}, where n= the total number of variables per object); however, the goal is to identify the fewest number of components that can still adequately describe a defined minimum amount of the total variation in the dataset. The number of PCs selected will be determined by examination of Scree plots produced in Unscrambler (D'Agostino and Russell 2005).

Following initial examination of the sample spectra, PCA was applied to all spectra in order to examine potential group patterns and identify which portions of the spectrum provided the best discrimination based on PMI. For exploratory purposes, the technique was applied to all samples for the following spectral bandwidths: complete spectra (350-2500nm); near-complete spectra (460-2360nm); visible bands (400-700nm); NIR bands (800-2500nm); and, specific bands of interest. All PCAs were performed on mean-centered, non-weighted spectral data. The NIPALS algorithm, a technique which calculates PCs iteratively from largest to smallest, was selected with a defined maximum of 100 iterations. No rotations were applied. To save on processing time, the number of components calculated was limited to 7; it should be noted that in most cases, the first two PCs often described 95% or more of the total variation.

All PCAs were validated using random cross-validation and a segment count of 20, with the total number of samples being used per segment equal to the number of samples divided by 20 (Brooke 2015).

3.4.5.3 Modeling

3.4.5.3.1 Sample Stratification

Prior to modeling, samples were assigned to either training sets that were used to build the models, or test sets which were used for external validation. This process was performed in MS Excel. A new column (RANDOM) was added to the spreadsheets. Each empty cell was then populated with a random number using the MS Excel RAND function and the data table was then sorted on RANDOM and PMI Class. A portion of the sorted samples from each PMI class were selected to build the training sets, while the remainders were used for the test groups.

3.4.5.3.2 Training Set Exploration and Pretreatment

Each training set was examined using line plots, scatter effect plots, and PCA in order to determine whether the data showed patterns similar to those observed on the full dataset. Pre-treatments (SNV and SG) were applied to the raw training data and the resultant transformed data were plotted and examined as well. Note that the test sets were not pretreated.

3.4.5.3.3 Model Construction and External Validation

Models were constructed on raw data as well as pretreated data. Note that only training set samples were used for model construction. Model accuracy was calculated using test sets

with known class assignments (external validation). Selected classification models included the following: Soft Independent Modeling of Class Analogy (SIMCA); Linear Discriminant Analyses on Principal Components (LDA-PCA); Support Vector Machine Classification on Principal Components (SVM-PCA); and, Projection of Latent Squares Discriminant Analysis (PLSDA).

3.3.5.3.3.1 Soft Independent Modeling of Class Analogy

Soft Independent Modeling of Class Analogy (SIMCA) operates essentially as a nested set of PCAs. Using training samples, local PCA models are constructed for each class in order to best describe the variation within a given class. These disjointed class PCAs are then combined to build a global PCA (the SIMCA classifier) which can then be used to assign unknown samples to the most appropriate class, based on which local PCA best describes the sample (Wold 1975). Assuming a two-class global PCA, SIMCA will place the unknown test samples in to one of the two possible classes or potentially into both classes (Maesschalck et al. 1999). In the latter case, the sample shares features with both possible classes and thus cannot be properly assigned. Another possibility is that a sample is assigned to no class, which often suggests that it either belongs to another undefined class or it is an outlier (Maesschalck et al. 1999).

Selected PCA settings for SIMCA were the same as those described in Section 3.3.5.2.3. Samples in the training set were selected by class to create Local PCA models based on selected bands, using raw or pretreated data. The SIMCA classifier was then constructed from the local PCA models (one Global PCA) and applied to the test set. All test set data were mean-centered, and, if necessary, pretreatments were applied. The output classification table was then imported into MS Excel where classification accuracy was calculated.

3.3.5.3.3.2 Linear Discriminant Analysis on Principal Components

Linear Discriminant Analysis (LDA) is another technique that can be used to classify two or more descriptive classes. For each group comprising a number of samples that have multiple measurable variables, LDA maximizes the separation between group multivariate means while minimizing the variation, or spread, within each group (Fisher 1936). To work properly, it is assumed that the number of variables is less than the number of samples, and that the variables are not correlated with each other. Unfortunately, spectral datasets are often extremely large and the data are highly intercorrelated. Working with a genetic dataset comprised of a large number of correlated variables, Jombart and colleagues (2010) showed that such data could be reduced to a limited number of PCs, which satisfies the first condition; additionally, PCs by definition are uncorrelated. This same concept can be applied to spectral data.

Unscrambler allows for a direct option for Linear Discriminant Analysis on Principal Components (LDA-PCA). Prior to model construction, the training set was analyzed with PCA (see section 5.3.5.2.3 for settings). The output was examined in order to identify the total number of PCs required to describe a minimum of 95% of the total variation in the dataset. Subsequently, LDA-PCA models were constructed. Descriptors consisted of training samples and relevant non-weighted raw or pretreated bands; classification settings were based on PMI Class. The band data were then used to calculate PCA scores. The number of PCs used for model construction was limited to the minimum number required to reach the 95% of variation threshold.

Developed LDA-PCA models were subsequently used to classify the test sets. After selecting the appropriate model, pretreatments were applied to the raw test data, if necessary, and the band(s) of interest were selected. Output figures were saved, and the classification table was then used to calculate classification accuracy in MS Excel.

3.3.5.3.3.3 Support Vector Machines on Principal Components

Given a dataset with known class values (e.g. PMI Classes), a Support Vector Machines (SVM) algorithm constructs a classifier based on machine learning which maximizes the distance between two or more groups. These SVM Classifiers can subsequently be used to classify unknown samples, assuming that such samples 1) arise from the same population as the training samples and 2) sample preparations are the same. The groups of interest are separated by the hyperplane with maximum-margins, that is, the plane that is the greatest distance from the closest samples in opposing groups (support vectors), providing a boundary which minimizes the number of misclassified samples (Kecman 2005).

SVM can be used on individual variables or groups of variables. In cases where the variable set is extremely large, it has been shown that data can be reduced with PCA, and SVM can be used on the resultant PCs as well (Dong and Liu 2011). Because of the large number of available variables in this study, SVM classification models were constructed with PCs (reduced data). Unlike the direct options offered with LDA, Unscrambler does not provide an automatic option to use PCs in an SVM Classifier Model. To compensate for this, PCA was run on combined training and tests sets, both raw and pretreated, using the previously described settings (section 5.3.5.2.3). The number of PCs required to describe a minimum of 95% of the

total variation was recorded; all PCA scores were then copied into the data matrix. SVM Classification models were constructed using training set samples and selected PC scores as predictors, with PMI Class variables selected as the classifiers. The SVM classification algorithm available in Unscrambler was originally reported by Chang and Lin (2011). PC data were left non-weighted. Models were constructed using a radial base function kernel type and were cross-validated using 10 equal segments. Default settings were used for Gamma (0) and C (1).

The SVM-PCA models were then applied to the test sets. For a given model, the appropriate number of PCs was selected. Following classification, figures and tables were saved, and model classification accuracies were calculated in MS Excel.

3.3.5.3.3.4 Partial Least Squares Discriminant Analysis

Partial Least Squares Discriminant Analysis (PLSDA) is an application based on PLS-Regression (PLSR). This process is similar to PCA in that the dimensionality of many X-variables is reduced to a limited number of latent variables (LV, similar to PC). However, the LVs are regressed on the Y-variables in order to explain the maximum amount of variation in those Y-variables (Wold et al. 2001). PLSR cannot be applied to the current dataset because the Y-variables, descriptive PMI classes, are categorical (i.e. PMI classes). In PLS-DA, the Y-variables are replaced with categories ("dummy variables"), and the reduced X-variables (LVs) are used to drive greater separation between the Y-classes (Brereton and Lloyd 2014).

Prior to building PLSDA models, the PMI Class column, a categorical variable with at least two possible discrete values, was split in order to create the new Y-variables (DUMMY). PLSDA models were then constructed in two ways. The first group of models used Unscrambler

default settings. Training set samples and select bands were used as predictors, and the DUMMY columns were selected as response variables. Models were built using the Kernal PLS algorithm on mean-centered and non-weighted data, with a total of 7 latent variables calculated. All models were internally validated by random cross-validation using 20 segments, with the number of samples per segment equal to the total number of samples divided by 20.

The classes under study have unequal samples sizes, which may potentially decrease the robusticity of the model. Brereton and Lloyd (2014) reported that when one class is larger than another, the resultant PLS models tend to underperform when predicting class identity for members of the larger group. To account for this, the authors recommend a correction for unequal samples sizes. This is performed by 1) calculating an average absorption value for each group (PMI Class), 2) calculating the overall mean of the groups, 3) subtracting that overall mean from each original value, 4) and, build PLSDA models using the corrected data with no mean-centering. This correction was applied to the full dataset. Corrected training set data were then used to construct additional PLSDA models with no mean-centering. Note that all other settings remained the same as described above.

Prediction of test set data was performed for each generated PLSDA model. The number of LVs selected was based on the number recommended by Unscrambler. Appropriate band(s) of interest were selected to represent the X-matrix of variables, and the DUMMY columns were chosen for the Y-variables. When necessary, pretreatments were applied to the corrected test data. Relevant figures produced by Unscrambler were saved, and the output prediction tables were imported into MS Excel. For a given sample, predicted DUMMY values were continuous and ranged from 0 to 1. All samples with predicted values less than 0.5 were assigned to class 0,

while those above were named as class 1; in the event Y = 0.5, a sample was deemed to be not classified (Brooke 2015). Using these decision rules, accuracy was calculated for all PLSDA classifications.

CHAPTER 4

RESULTS

4.1 Introduction

This chapter presents the results of the analyses described previously. The chapter is broken down into five primary sections. The first two sections include a rationale for the specific data sets selected for model building (e.g. longitudinally-cut cortices and superficially treated cortical surfaces) and documentation of excluded samples. The third section describes the steps taken in data preparation, including spectral band selections, transformations, and sample stratification. The fourth section presents a series of models constructed on raw spectral training sample data (1400-2200nm), and the associated external model validation results for the test sets. The final section reports on tests for potential protein oxidation that were conducted after the broader research was completed.

4.2 Data Set Selection

Averaged spectra were collected from the following surfaces: the untreated superficial cortical surface (n=336); the superficial cortical surface treated lightly with a grinding wheel (n=223); the longitudinally cut cortical surface (n=333); and, the transversely cut cortical surface (n=281). Line plots presented for the two superficial series are described below (Figure 11 and Figure 12). To allow for a direct comparison, spectra presented in these plots were collected from samples for which both surfaces could be examined (n=220). Note that the separation between the “forensic” and “non-forensic” spectra is poor for the untreated set (Figure 11a), while better separation is observed with the treated surface spectra (Figure 12a).

A subsequent PCA of these two series demonstrated class overlap for the untreated materials and partial separation into two clusters for the treated sample spectra; PCA scores plots are presented in Figure 11b and Figure 12b.

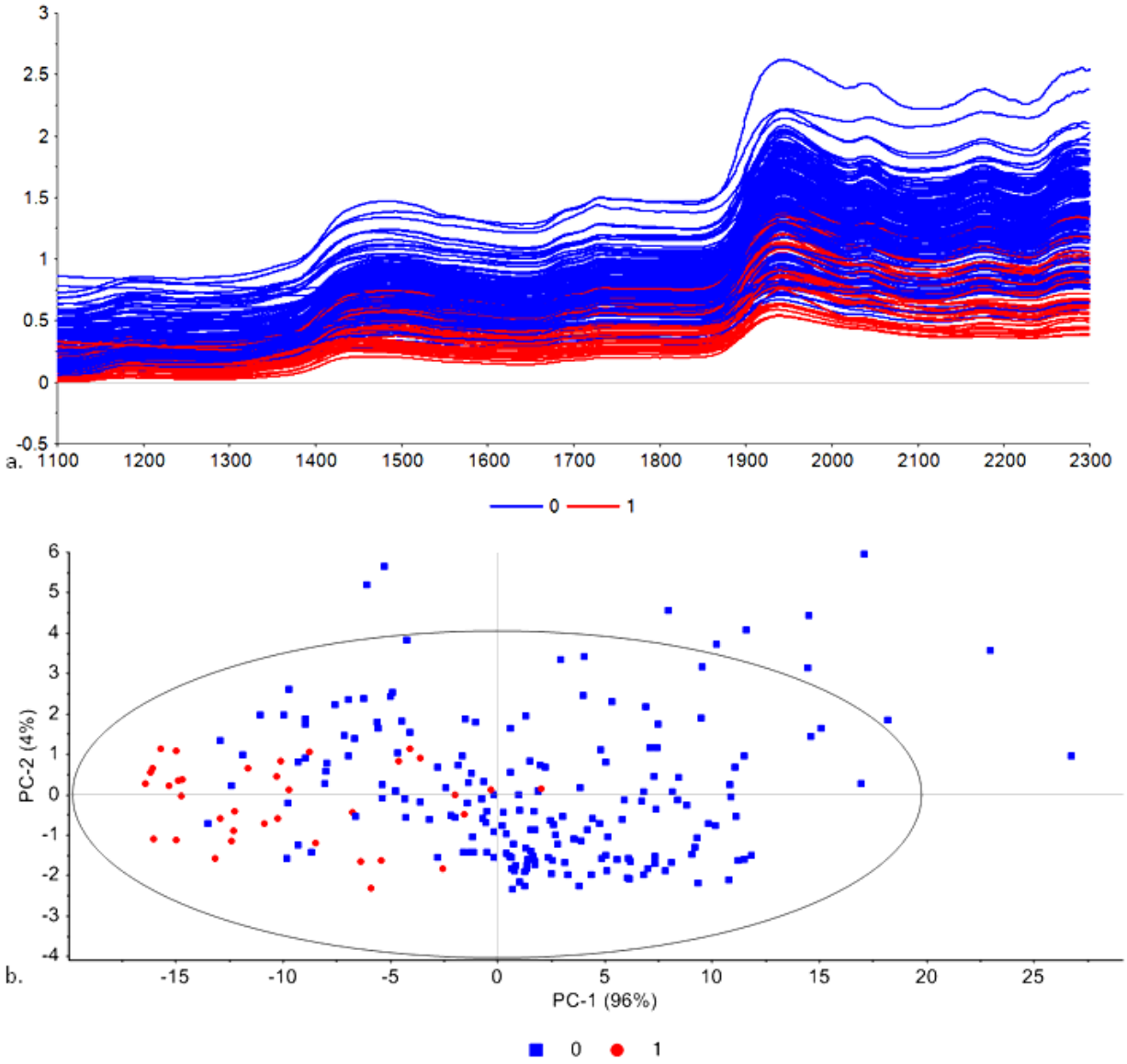


Figure 11. Line and score plots, superficial bone surfaces. a) Raw spectra (1100-2300nm) collected from the untreated superficial bone surface, n=220. The associated PCA scores plot (1400-2200nm) is presented in (b). Note the poor separation between the forensic and non-forensic samples.

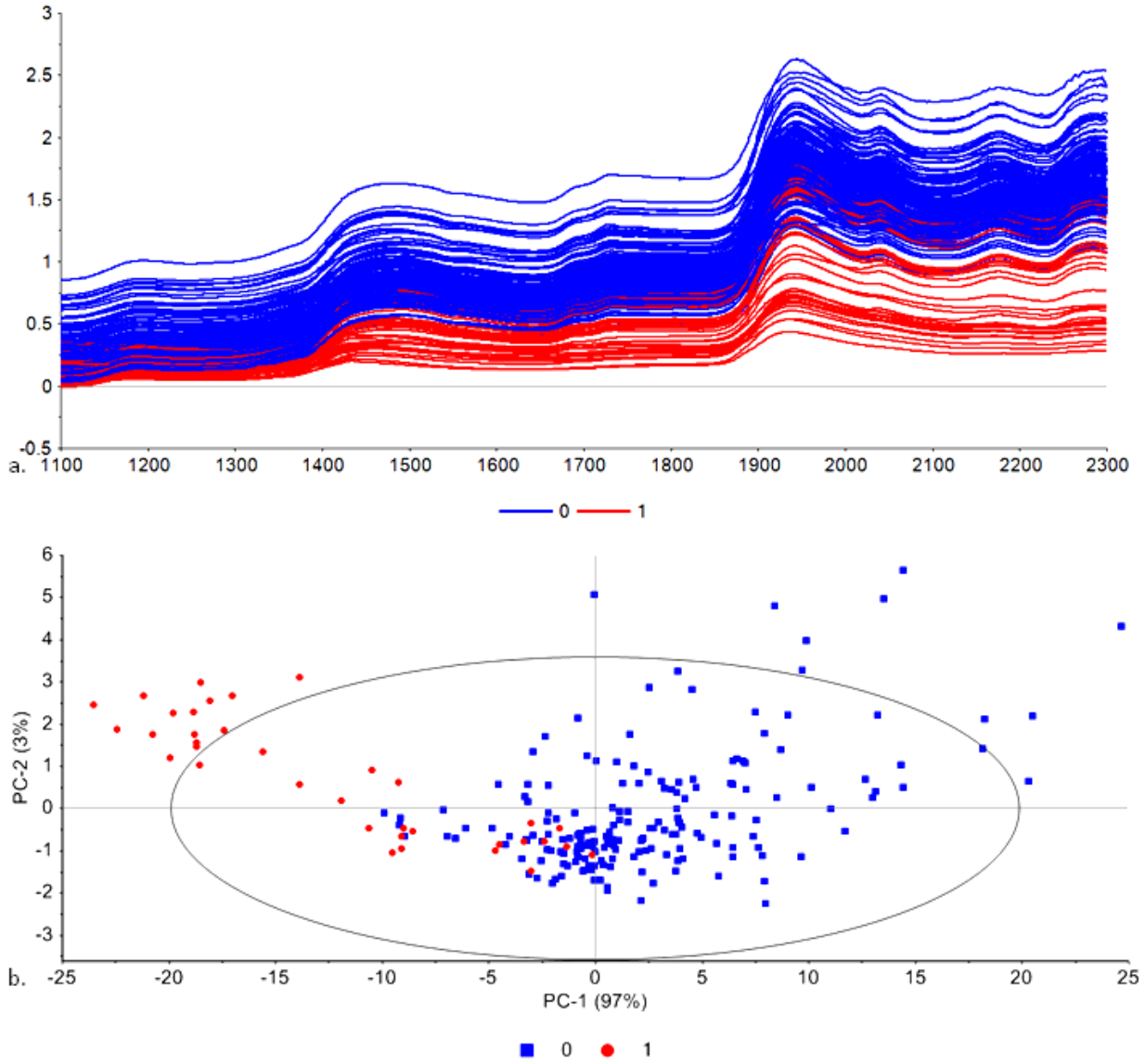


Figure 12. Line and score plots, superficial bone surfaces. a) Raw spectra (1100-2300nm) collected from the treated superficial bone surface, n=220. The associated PCA scores plot (1400-2200nm) is presented in (b). Note that separation between the forensic and non-forensic samples is more pronounced that what is observed on the untreated superficial surface.

Comparisons between spectra for transversely and longitudinally cut cortical surfaces were made as well (common samples n=159). Line plots for both sets of spectra demonstrate similar separation patterns between forensic and non-forensic spectra (Figure 13a and Figure 14a). Likewise, PCA scores plots reveal similar clustering patterns between the two classes (Figure 13b and Figure 14b).

Spectra collected from the longitudinally-cut surface and from the treated superficial surface were selected for future analyses, included model building. Spectra collected from the untreated superficial surface (i.e. no grinding) produced poor separation between forensic and non-forensic classes in both line and PCA score plots and were thus excluded. Conversely, the transverse and longitudinal cut cortical surfaces yielded spectral data that can be separated into classes, although the information is ultimately redundant because the same level of separation is observed. Because there are a greater number of spectra collected from the longitudinal cut surface, the transverse set was not included in further analyses.

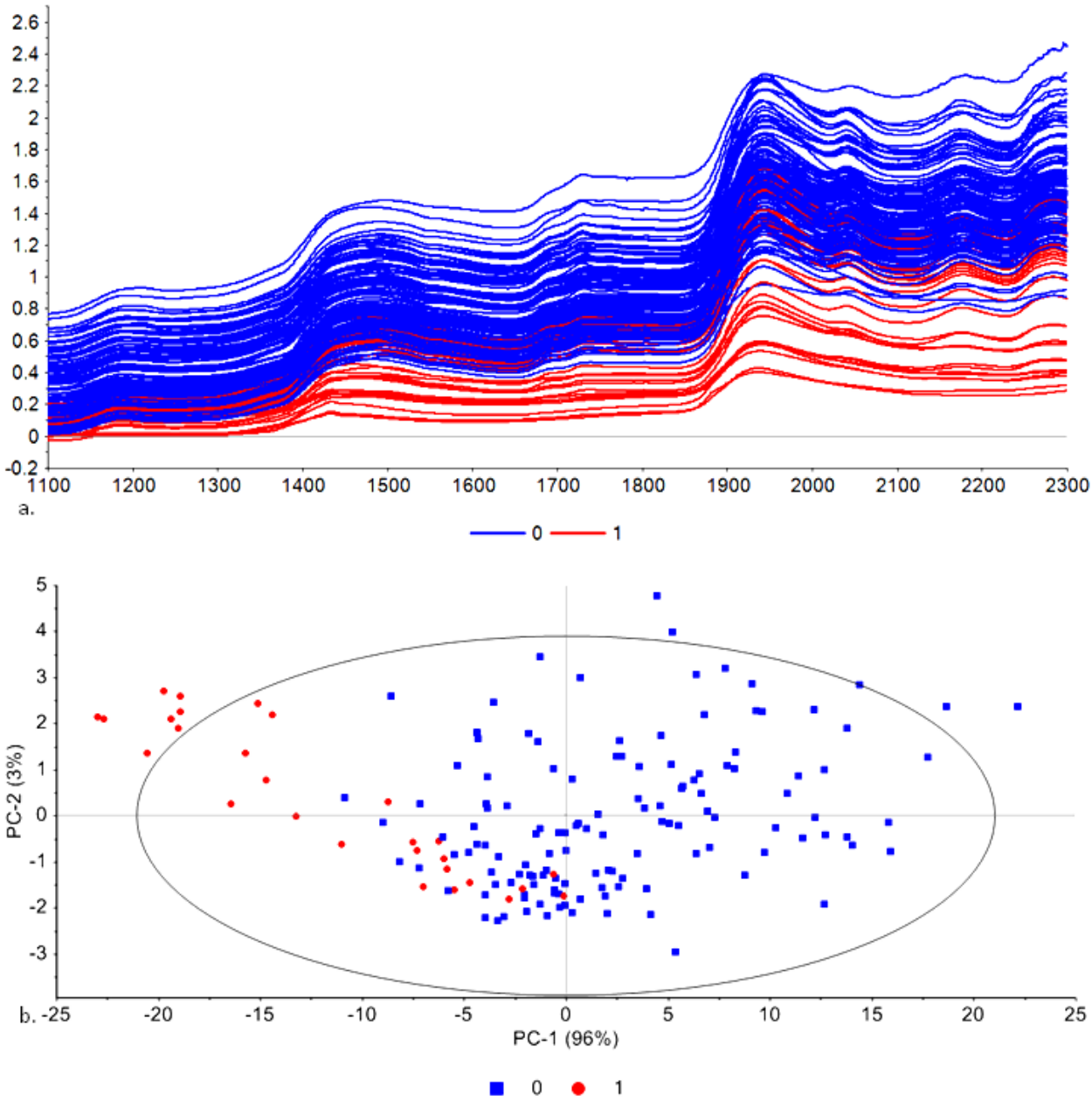


Figure 13. Line and score plots, sectioned bone surfaces. a) Raw spectra (1100-2300nm) collected from the longitudinally cut cortical surface, n=159. The associated PCA score plot (1400-2200nm) is presented in (b). Good separation is achieved between the forensic and non-forensic samples.

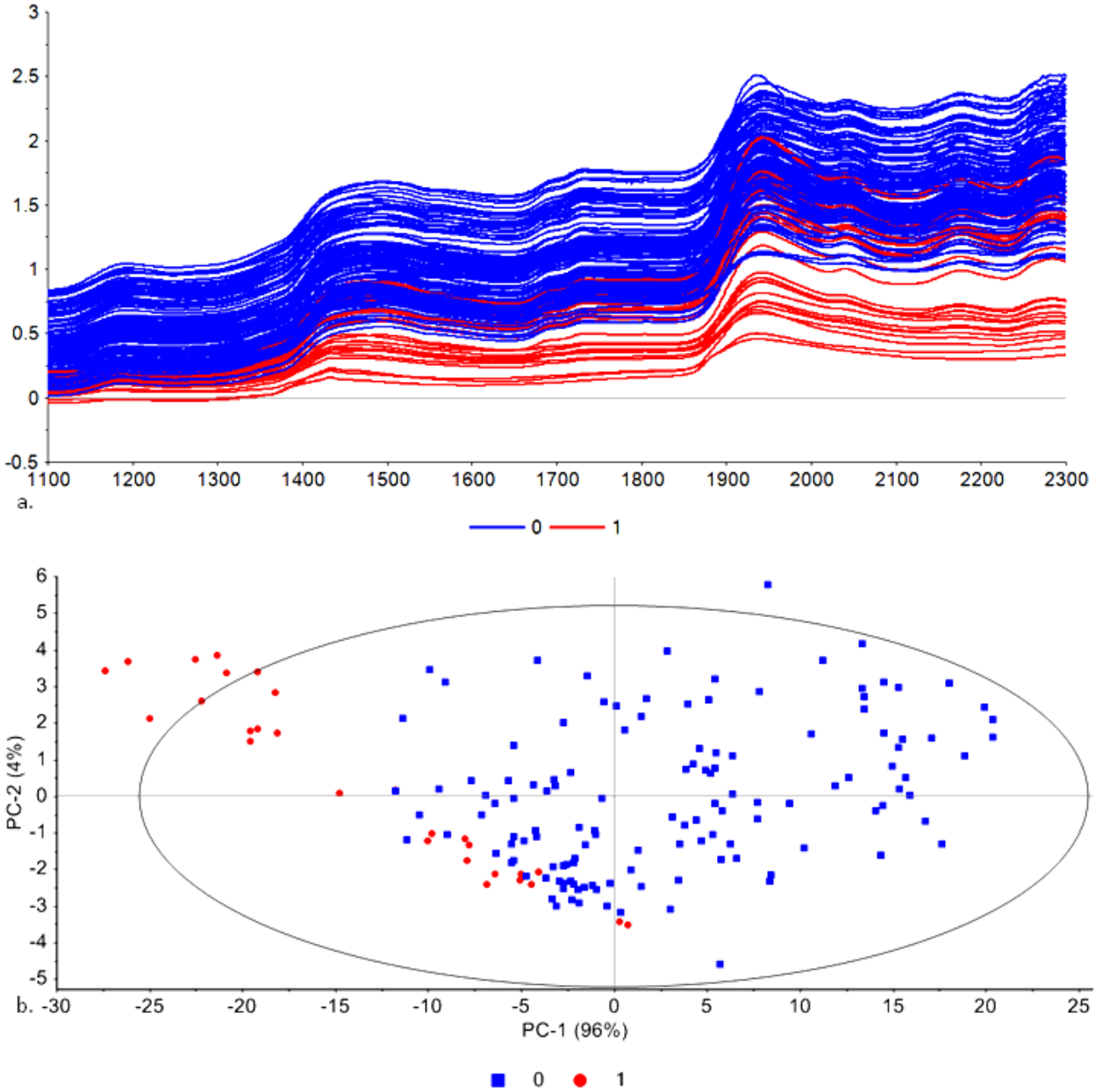


Figure 14. Line and score plots, sectioned bone surfaces. a) Raw spectra (1100-2300nm) collected from the transversely cut cortical surface, n=159. The associated PCA scores plot (1400-2200nm) is presented in (b). Good separation is achieved between the forensic and non-forensic samples.

4.3 Sample Exclusions

A total of 11 samples were excluded from the final study sets. Four of these samples came from remains that showed evidence of perimortem thermal damage. This included sample 35, an LFA femoral sample extracted from remains recovered from a structure fire, as well as samples 236, 239, and 243. The latter three specimens were scanned at the MMA, and

the line plots of these three spectra appeared within the band comprised primarily of historic and archeological materials even though these cases were of forensic origin (Figure 15). Although the spectra were collected from cortical regions that presented no gross evidence of thermal damage, it is likely that the heat disrupted the collagen even in remote areas of bone with no apparent scorching or visible heat-related damage; recall the cremains spectrum presented in Figure 4 (see Chapter 3), which can be thought of as an extreme example of this phenomenon. While sample 35 fell within the group of forensic spectra, spectra collected from all burned samples were excluded from further analysis for consistency.

Five additional forensic samples, including MMA specimens 241, 244, 265, and 268, as well as CHI archived sample 203, also produced spectra which grouped with the historic and archaeological materials (Figure 15). It was noted when the spectra were collected that each of these specimens was coated in a thin layer of adipocere, which likely resulted from incomplete drying following maceration. These five samples originated from the MMA and LMI and thus could not be cleaned prior to spectral collection. Because these spectra grouped inappropriately with the older, non-forensic material, and the commonality of adherent adipocere was observed, these samples were also excluded.

Sample 30 was a fragment of occipital bone originating from archaeological remains housed at LFA. Although the spectral shape was consistent with spectra collected from other archaeological materials, the sample was excluded because it did not originate from a long bone. Lastly, sample 308, a LMI specimen, was excluded because the associated case data were not recorded at the time of spectral collection.

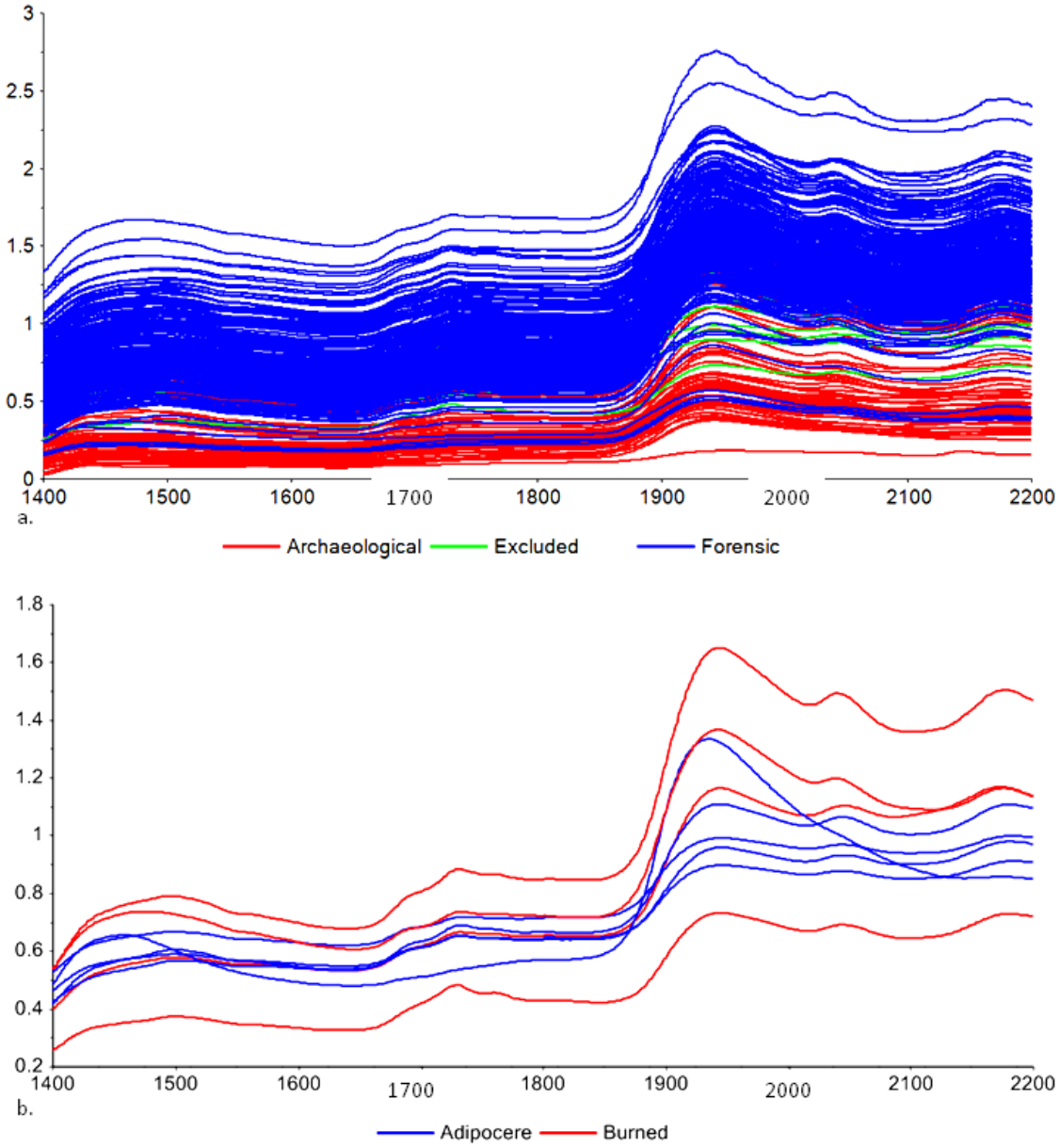


Figure 15. Raw absorbance spectra collected from the longitudinally-cut surface, 1400-2200nm band selected, included and excluded samples. Line plots present a) spectra selected for model building (n=314) and excluded samples (n=9); b) spectra for the 9 excluded samples.

4.4 Data Preparation

A closer examination of the line plots for both the superficial treated surface and the longitudinal cut surface revealed that the terminal ends of the spectra were noisy relative to the remaining central bands. Noisy bands are highlighted in Figure 16a below. These regions were cropped prior to further analysis, resulting in abridged spectra of 460-2360nm (Figure

16b). Note that separation between forensic and non-forensic spectral classes is relatively poor between 460-1200nm, while clearer separation is observed between 1400-2200nm.

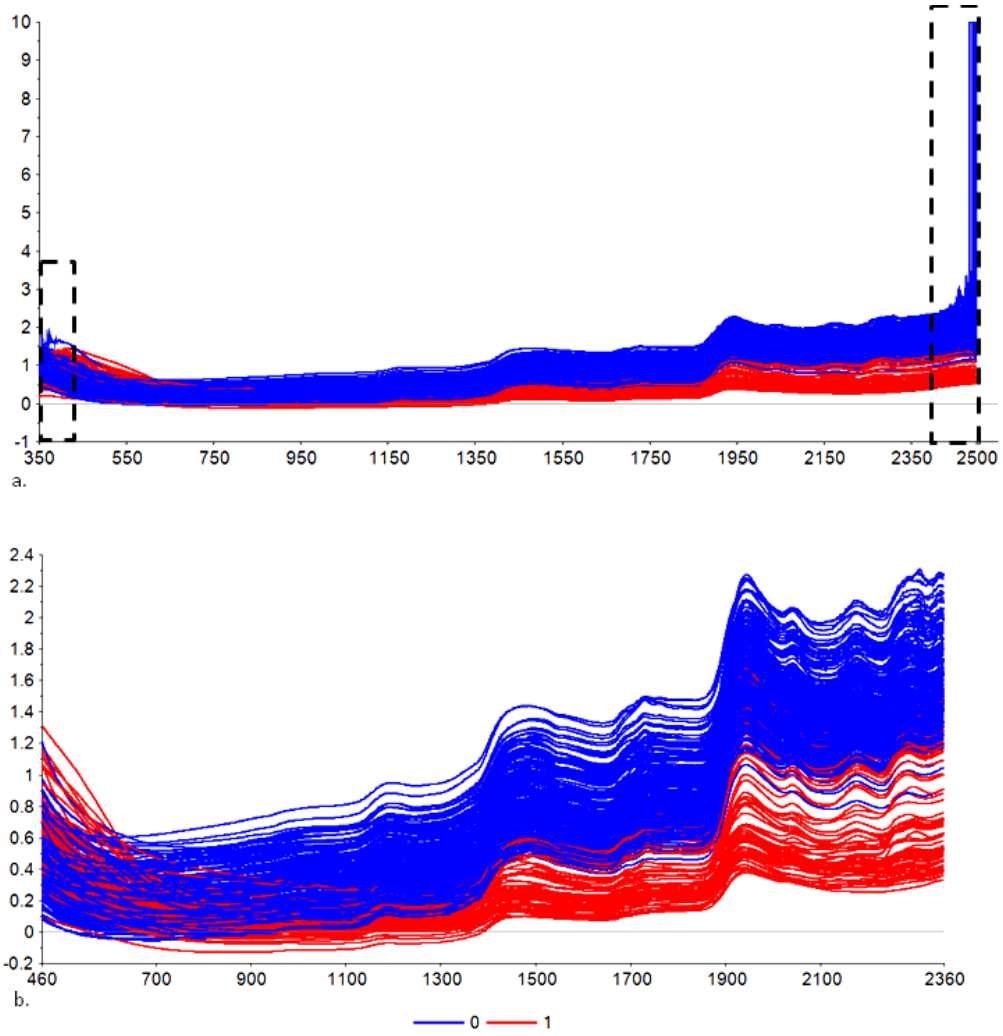


Figure 16. Line plots collected from the longitudinally cut cortical surface, n=314. a) The lower and upper ends (dashed boxes) of the raw absorbance spectra (350-2500nm) are noisy. b) Noisy areas have been cropped, resultant spectrum is 460-2360nm.

4.4.1 Data Transformations

Scatter effects plots generated in Unscrambler revealed evidence of both additive and multiplicative scattering in the spectral datasets; these plots are presented in Appendix D. The following pretreatments were applied (Rinnan et al. 2009, Brooke 2005):

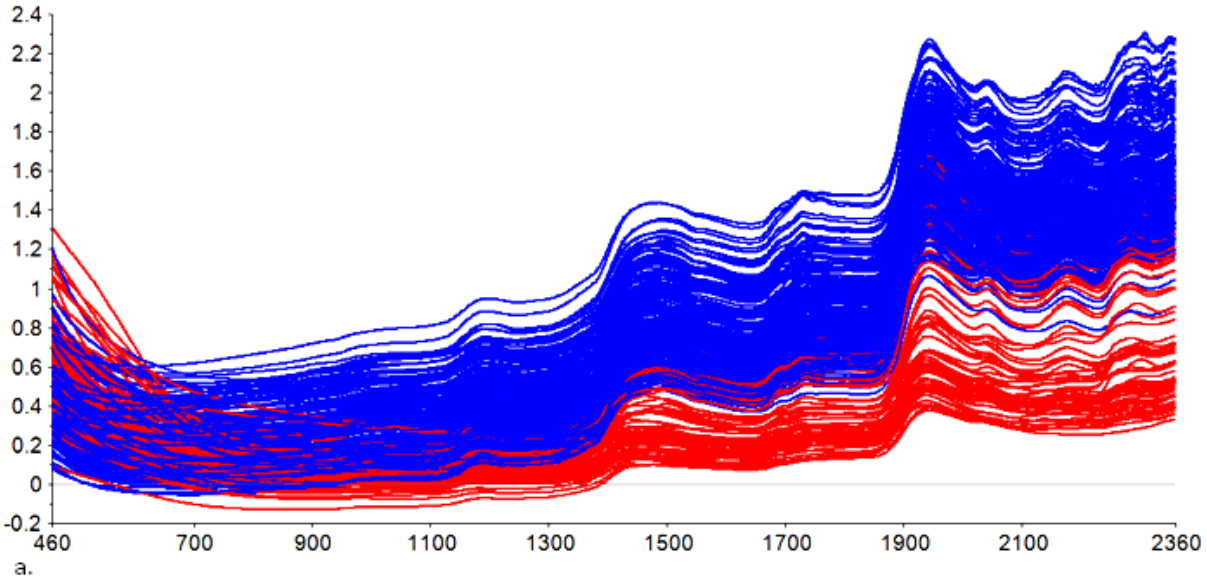
- Standard Normal Variate (SNV)
- Savitzky-Golay 1st derivative (SG1): 25 bilateral smoothing points with a 2nd order polynomial
- Savitzky-Golay 2nd derivative (SG1): 25 bilateral smoothing points with a 3rd order polynomial

No tandem transformations (i.e. SNV followed by SG1) were performed as previous attempts yielded transformed spectra which separated poorly with PCA.

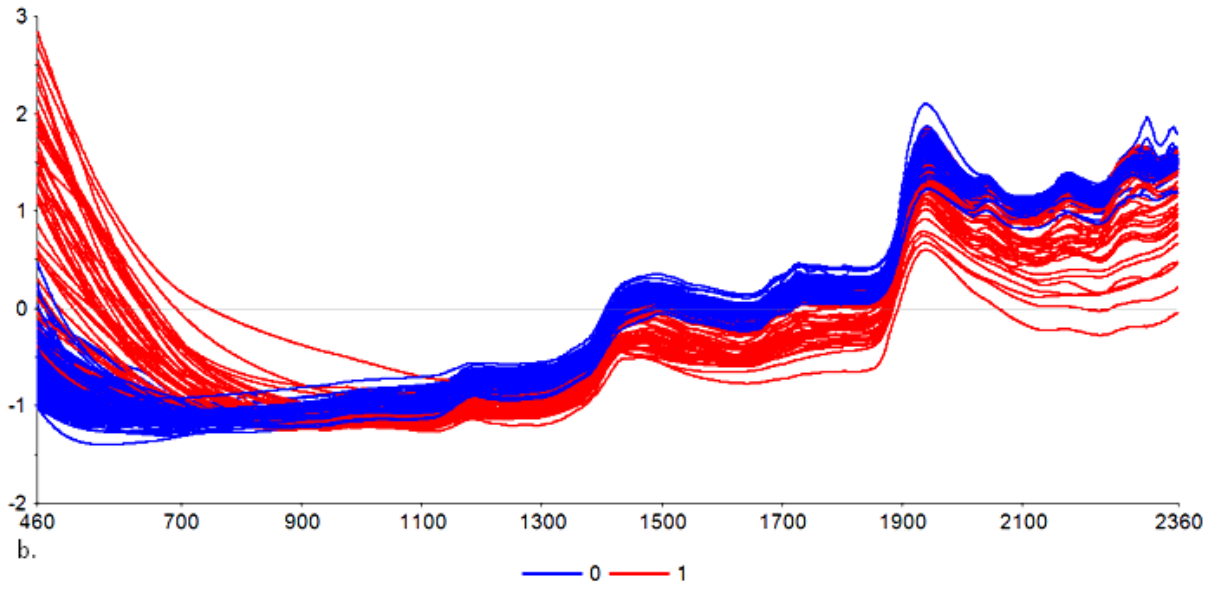
Following transformation, raw and transformed line plots were examined in order to identify areas in which best separation between forensic and non-forensic classes occurred. Raw data demonstrated best-separation from 1400-2200nm; SNV-transformed data also showed good differentiation between these classes for 1400-2200nm, especially within a subset band of 1425-1750nm. Line plots for SG1- and SG2-transformed data showed class separation in the 1400-2200nm band as well. In addition, the following SG subsets were identified:

- SG1 specific bands: 1120-1185nm and 1530-1720nm
- SG2 specific bands: 1235-1330nm and 1535-1585nm

Examples of raw and transformed spectral data are presented in Figure 17.


2.4
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
460
700
900
1100
1300
1500
1700
1900
2100
2360

a.


3
2
1
0
-1
-2
460
700
900
1100
1300
1500
1700
1900
2100
2360
0
1

b.

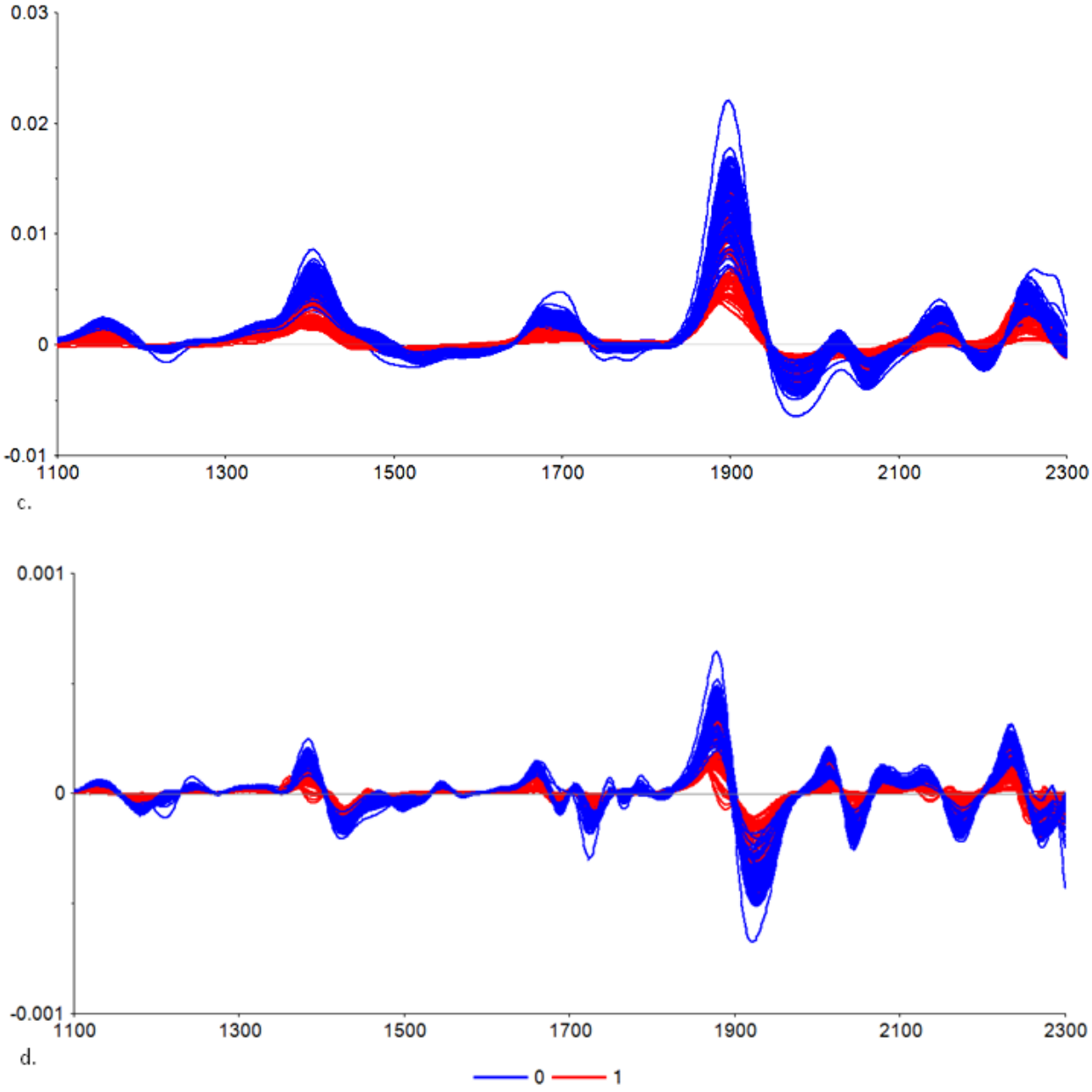


Figure 17. Line plots of raw and transformed spectral data collected from the longitudinally cut cortical surface, n=314. Presented plots include 460-2360nm (a) raw and (b) SNV-transformed absorbance spectra, as well as 1100-2300nm Savitzky-Golay (c) first and (d) second derivative absorbance spectra.

4.4.2 Variable Definitions

Prior to model building, spectral variable row sets (i.e. bands) and PMI class designations were defined. Spectral bands included the following: 460-2360nm; 1400-2200nm; 1425-1750nm; SG1 specific bands; and, SG2 specific bands. PMI class designations were as follows:

- PMI 2-class groupings: Forensic (class 0) and Non-forensic (class 1)
- PMI 3-class groupings: Forensic 1 (class 0), Forensic 2 (class 1), and Non-forensic (class 2)

"Forensic" samples in the 2-class column represent cases with a known or estimated PMI ranging from "days" to 71 years. "Non-forensic" samples are those estimated to be of historic or archeological origin, as well as the dated material examined at the MMA with a PMI 500 and 900 years, respectively. Specimens with values assigned in the 3-class column were those with known PMI only, with "Forensic 1" representing PMI of "days" to 4.9 years, "Forensic 2" comprising 5 to 71 years, and "Non-forensic" including the dated MMA specimens. Note that the "Forensic 2" class is broadly defined (5 – 71 years) owing to the limited of documented older forensic cases.

4.4.3 Sample Stratification

After spectral bands of interest and PMI class variables were defined, the longitudinal cut surface raw data set (Set A) and the treated superficial cortex raw set (Set B) were each split into training and test sets. For Set A, 70% of the samples were randomly assigned to the model

training sets, while the remaining 30% comprised the test sets. Resultant samples sizes used to train and test each group of Set A models are presented below:

- 2-class model (Forensic, Non-forensic): **2Cl_FN**
    - Train: Forensic = 175, Non-forensic = 45
    - Test: Forensic = 75, Non-forensic = 19

- 3-class model (Forensic 1, Forensic 2, Non-forensic): **3Cl_F_1F_2N**
    - Train: Forensic 1 = 29, Forensic 2 = 11, Non-forensic =4
    - Test: Forensic 1 = 20, Forensic 2 = 8, Non-forensic =3

- 2-class Forensic, Non-forensic samples excluded (Forensic 1, Forensic 2): **2Cl_F_1F_2**
    - Train: Forensic 1 = 29, Forensic 2 = 11
    - Test: Forensic 1 = 20, Forensic 2 = 8

Stratification of Set B was carried out in similar fashion; however, because the overall sample size was lower, it was decided that the training set would comprise 60% of the samples and the test set the remaining 40%. Sample sizes for Set B training and test sets are as follows:

- 2Cl_FN
    - Train: Forensic = 132, Non-forensic = 25
    - Test: Forensic = 56, Non-forensic = 10

- 3Cl_F_1F_2N
    - Train: Forensic 1 = 16, Forensic 2 = 5, Non-forensic =4
    - Test: Forensic 1 = 10, Forensic 2 = 3, Non-forensic =3

- 2Cl_F_1F_2
    - Train: Forensic 1 = 16, Forensic 2 = 5
    - Test: Forensic 1 = 10, Forensic 2 = 3

See Tables E.1 – E.6 in Appendix E for a list of sample numbers assigned by class for all training and test sets. Note that the sample sizes in the 3Cl_F_1F_2N and 2Cl_F_1F_2 models are low, especially with the Set B models. This is due to the fact that only cases with known PMI were included for these models. Although far from optimal, these models were constructed in order to determine if it is possible to discriminate within the forensic time frame. Following

stratification, selected pretreatments (SNV, SG1, and SG2) were applied to the raw training sets only.

4.5 Model Building and Validation

Training sets were used to construct the selected classification algorithms (SIMCA, LDA-PCA, SVM-PCA, and PLSDA). Alternate models were built using both raw and transformed data with the following spectral bands:

- Raw data models: 460-2360nm; 1400-2200nm
- SNV-transformed data models: 460-2360nm; 1400-2200nm; and, 1425-1750nm
- SG1-transformed data models: 1400-2200nm; 1120-1185 and 1530-1720nm
- SG2-transformed data models: 1400-2200nm; 1235-1330 and 1535-1585nm.

For brevity, the results reported in this section will be limited to the "raw 1400-2200nm" models constructed for Sets A and B (see Tables 1 – 3, below). The remainder are reported in Appendix F. Note that the cross-validated model accuracies and external validation rates for transformed data models were generally similar to or poorer than those seen with many of the raw data models presented below, although there are some better performing exceptions. Further discussion about the transformed models will be presented in the following chapter.

4.5.1 Two-Class Models (Classes: Forensic, Non-forensic)

4.5.1.1 SIMCA

The global SIMCA 2CL_FN classifier was built from two local PCA models describing Set A training set Forensic (Class 0, PMI = "days" to 71 years) and Non-forensic (Class 1, "historic and

archaeological" origins) samples; both local models required two components to describe a minimum of 95% of the group variation. Overall accuracy for this model as calculated from the test set classification was 67.0%; 64.0% of Class 0 samples were properly classified, while Class 1 accuracy was 79%. For the remaining samples, 28.7% were identified as belonging to both classes, and 3.2% remained unclassified. The results of this model are presented in a Coomans plot (Figure 18a), which plots orthogonal distances between a pair of disjointed PCA models. Properly classified test samples fall within the correct model boundaries, while those samples classified to both groups fall within the area of overlap (lower left quadrant); non-classified samples and outliers are assigned to the upper right quadrant of the plot.

The corresponding Set B SIMCA classifier was built from two component local PCA models for Group 0 and Group 1 as well. Model accuracy as calculated from test set classification was much lower (33.3%); 33.9% of Class 0 samples were properly classified, while 30.0% of Class 1 was correctly assigned. The remaining test samples were classified as belonging to both (54.6%) or neither class (6.1%). The output of this model is presented in (Figure 18b).

4.5.1.2 LDA-PCA

The Set A 2Cl_FN LDA-PCA model was constructed with two components. Cross-validated model accuracy was 95.6% (Figure 19a). Overall classification accuracy for the test set was 94.6%, with 96.0% of the Class 0 samples being properly classified. Class 1 accuracy was slightly lower at 82.2%.

Two components were used to build the Set B model as well, with a reported model accuracy of 94.3% (Figure 19b). The test set was correctly classified at a rate of 89.4%, with 98.2% of Class 0 samples being accurately classified. Class 1 accuracy was in contrast much lower than that observed in the Set A model (40%).

4.5.1.3 SVM-PCA

Two components were used to construct the 2Cl_FN SVM-PCA model for Set A. Model accuracy was reported at 95.9% (Figure 20a). External validation shows that the model properly classified 95.7% of the test set samples, including 98.7% accuracy for Class 0. Similar to the LDA-PCA model, Class 1 accuracy was somewhat lower (84.2%).

The Set B model was also built from two components; reported cross-validated accuracy was 94.3% (see Figure 20b). The model correctly classified all Class 0 test samples, but only 20% for Class 1. Overall accuracy for the test set was 87.9%.

4.5.1.4 PLSDA

PLSDA models were constructed on mean-centered spectral data (PLSDA-MC) as well as non-centered data that was corrected for unequal sample sizes (PLSDA-CO). Results for both attempted model types follow: note that because PLSDA is a modified regression technique, root mean square error of cross-validation (RMSECV) and R^2 are reported. Note that a relatively low RMSECV is desirable because it indicates a good fit of the training data to the model (i.e. lower residuals); higher R^2 values approaching 1.0 indicate a more accurate prediction (Naes et al. 2002).

The 2Cl_FN PLSDA-MC Set A model (4 factors) had a reported RMSECV of 0.2208 and an R^2 value of 0.7013 (Figure 21). This model properly predicted 94.7% of all test samples, with 98.7% accuracy for Class 0 and a relatively low prediction rate of 78.9% for Class 1. The PLSDA-CO model (4 factors) yielded a RMSECV of 0.2364 and an R^2 value of 0.9303 (Figure 22). Prediction accuracy for the test set was the same for this model (94.7%). Class 0 assignment was predicted accurately at 100.0% while Class 1 was prediction rate was also low (78.9%).

The 2Cl_FN Set B PLSDA-MC model required 5 factors; reported RMSECV and R^2 values are 0.2041 and 0.6888, respectively. Test set sample assignments were properly predicted at 90.9% accuracy. While 98.2% of Class 0 samples were correctly classified, Class 1 sample assignment was again low (50.0%). RMSECV and R^2 values reported for the alternate Set B PLSDA-CO model (4 factors) were 0.2252 and 0.6815. Overall, 90.8% of the test samples were properly classified; however, Class 1 prediction rates were again low (40.0%) relative to the higher accuracy observed for Class 0 (98.2%). See Figure 23 and Figure 24 for these two models.

Table 1. Model and test set classification accuracies for select two-class models (1400-2200nm). Class 0 represents forensic samples with known or estimated PMI of less than 71 years; Class 1 represents non-forensic specimens (historic and archaeological origins). Numbers in "( )" represent the number of components or factors used in a given model. Cells in white represent cross-validated model accuracies as a percentage (LDA-PCA and SVM-PCA) or as RMSECV and R^2 for PLSDA. Areas in grey represent accuracies (%) from model application to the test set.

Model	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
2-class (Set A)	95.9 (2)	95.9 (2)	RMSE = 0.2208 (4) R^2 = 0.7013	RMSE = 0.2360 (4) R^2 = 0.9303	Overall = 67.0 Class 0 = 64.0 Class 1 = 79.0 Both = 28.7 Neither = 3.2
	Overall = 94.6 Class 0 = 96.0 Class 1 = 82.2	Overall = 95.7 Class 0 = 98.7 Class 1 = 84.2	Overall = 94.7 Class 0 = 98.7 Class 1 = 78.9	Overall = 94.7 Class 0 = 100 Class 1 = 78.9	
2-class (Set B)	94.3 (2)	94.3(2)	RMSE = 0.2041 (5) R^2 = 0.6888	RMSE = 0.2252 (4) R^2 = 0.6815	Overall = 33.3 Class 0 = 33.9 Class 1 = 30.0 Both = 54.6 Neither = 6.1
	Overall = 89.4 Class 0 = 98.2 Class 1 = 40	Overall = 87.9 Class 0 = 100.0 Class 1 = 20.0	Overall = 90.9 Class 0 = 98.2 Class 1 = 50.0	Overall = 90.8 Class 0 = 98.2 Class 1 = 40.0	

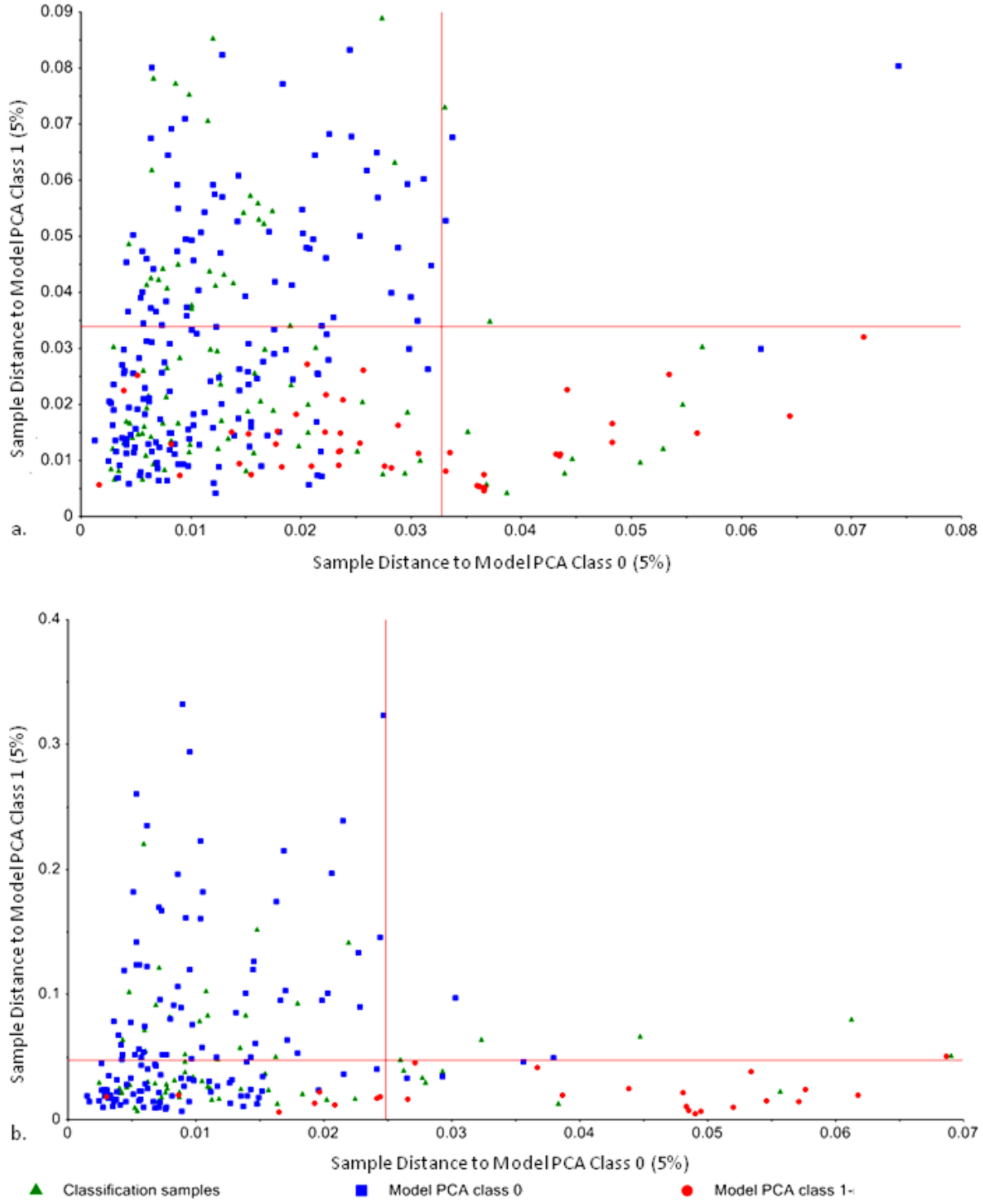


Figure 18. Coomans plots for two-class SIMCA models, raw 1400-2200nm. Models with classified samples (green) for Set A (a), n=220, and Set B (b), n=157, are presented. Classes include "Forensic" (Group 0: known or estimated PMI ≤ 71 years) and "Non-forensic" (Group 1: historical or archaeological).

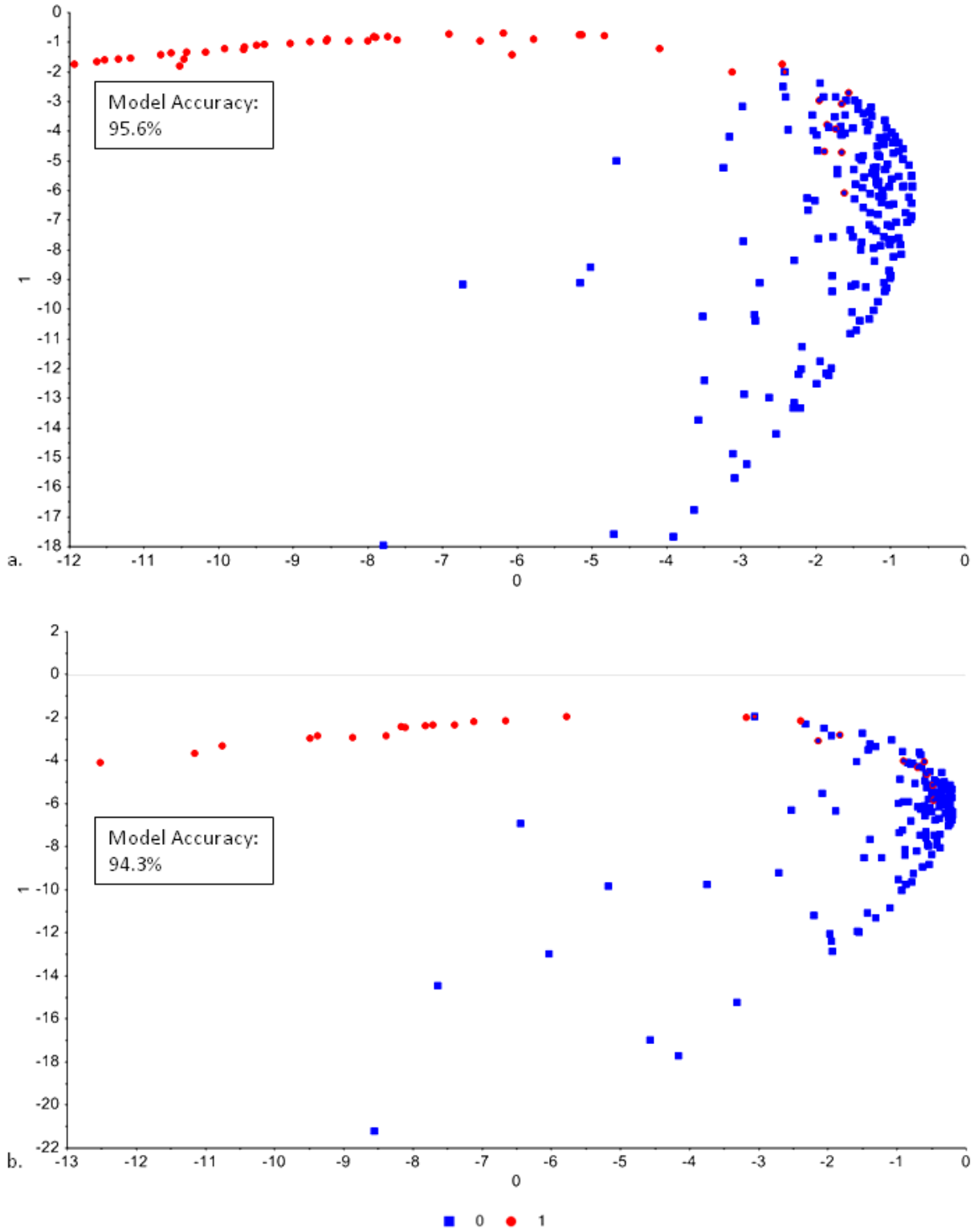


Figure 19. Two-class LDA-PCA models, raw 1400-2200nm. The Set A model (a) training set is n=220; Set B (b) is n=157 samples. Classes include “Forensic” (Group 0: known or estimated PMI ≤ 71 years) and “Non-forensic” (Group 1: historical or archaeological).

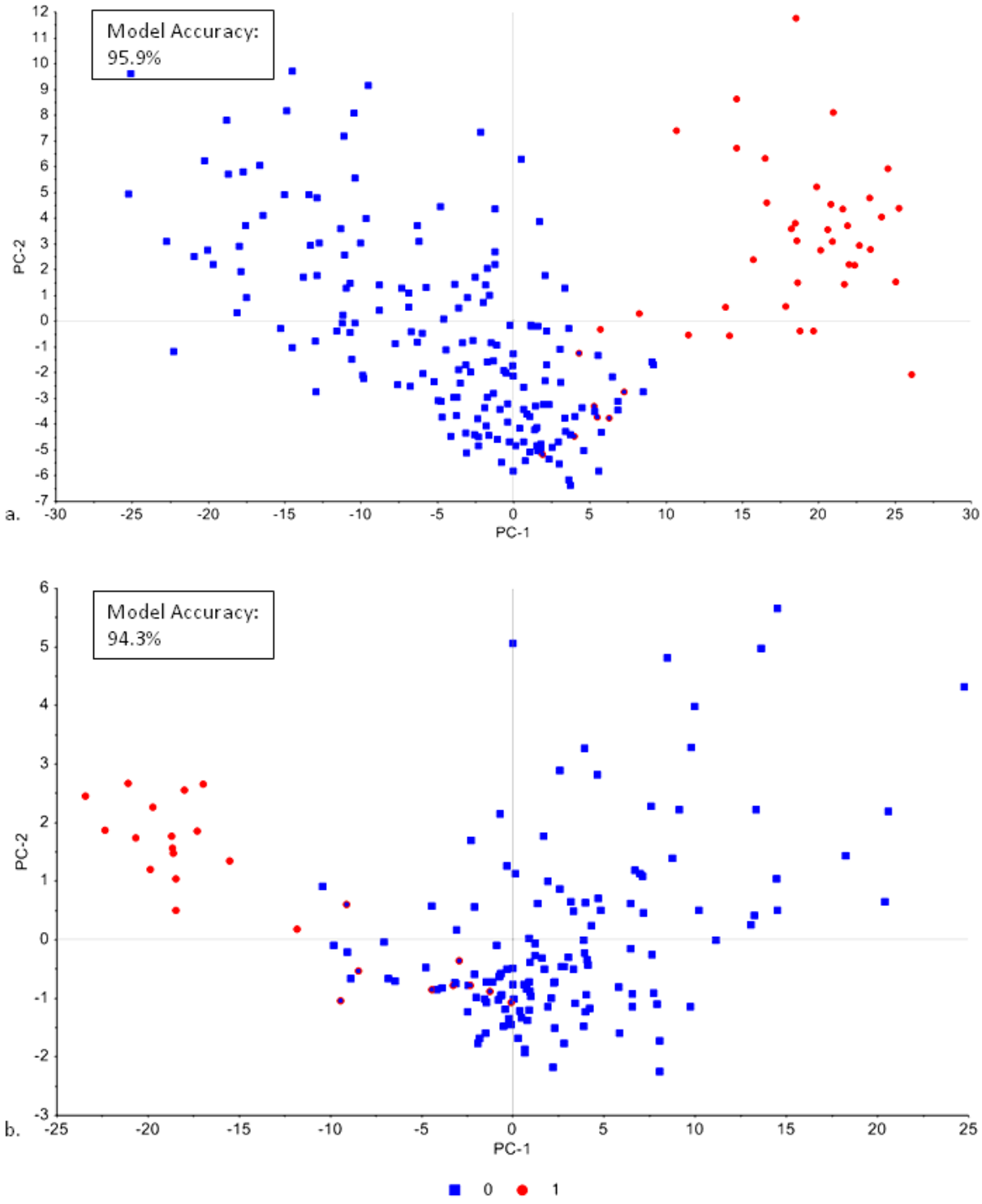


Figure 20. Two-class SVM-PCA models, raw 1400-2200nm. The Set A model (a) training set is n=220; Set B (b) is n=157 samples. Classes include “Forensic” (Group 0: known or estimated PMI ≤ 71 years) and “Non-forensic” (Group 1: historical or archaeological).

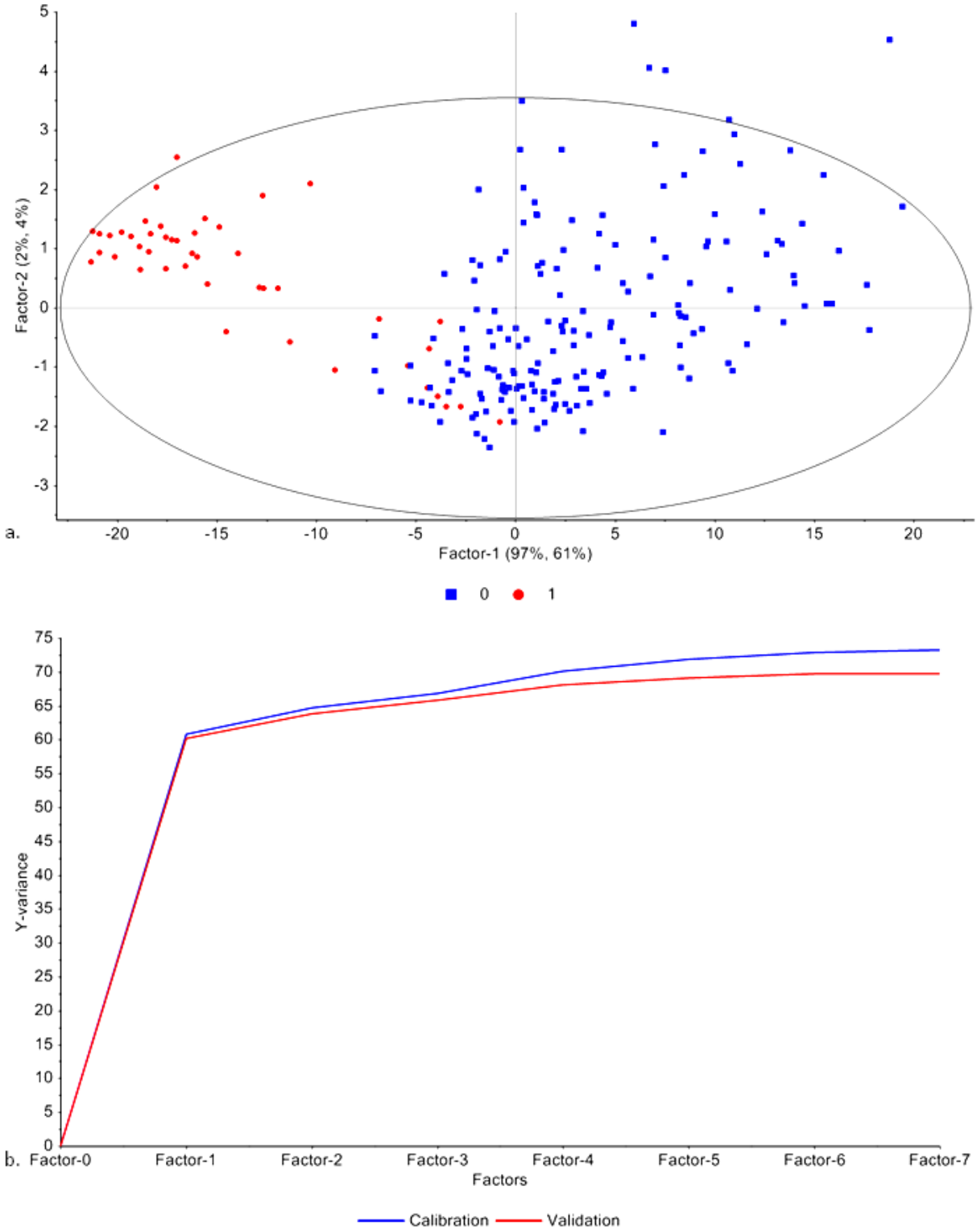


Figure 21. Two-class PLSDA models, a) scores and b) explained variance plots, raw 1400-2200nm. The presented model is built with mean-centered Set A training set sample spectra, n=220. Classes include "Forensic" (Group 0: known or estimated PMI ≤ 71 years) and "Non-forensic" (Group 1: historical or archaeological).

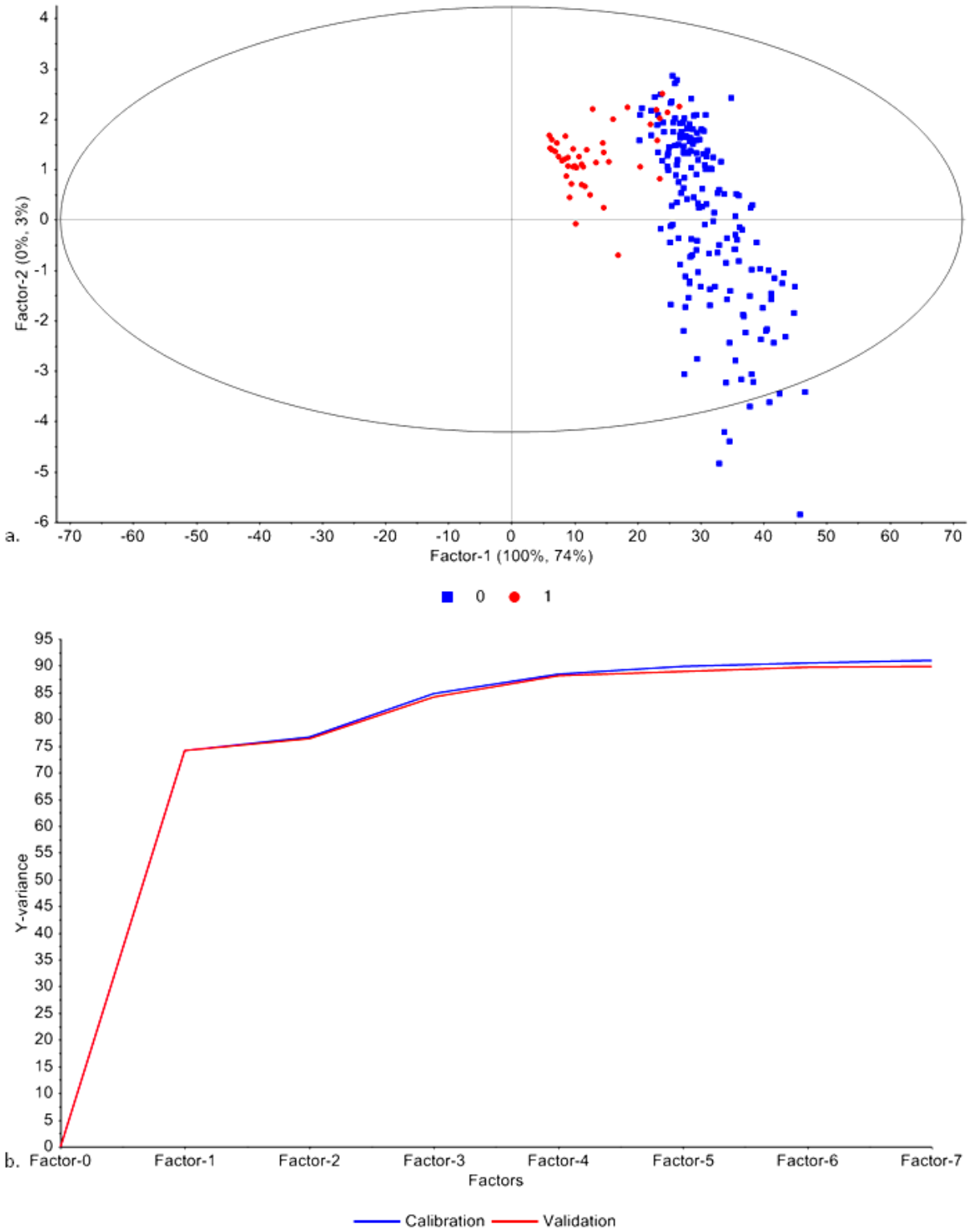


Figure 22. Two-class PLSDA models, a) scores and b) explained variance plots, raw 1400-2200nm. The presented alternate model is built with Set A spectral data corrected for sample size differences, n=220. Classes include "Forensic" (Group 0: known or estimated PMI ≤ 71 years) and "Non-forensic" (Group 1: historical or archaeological).

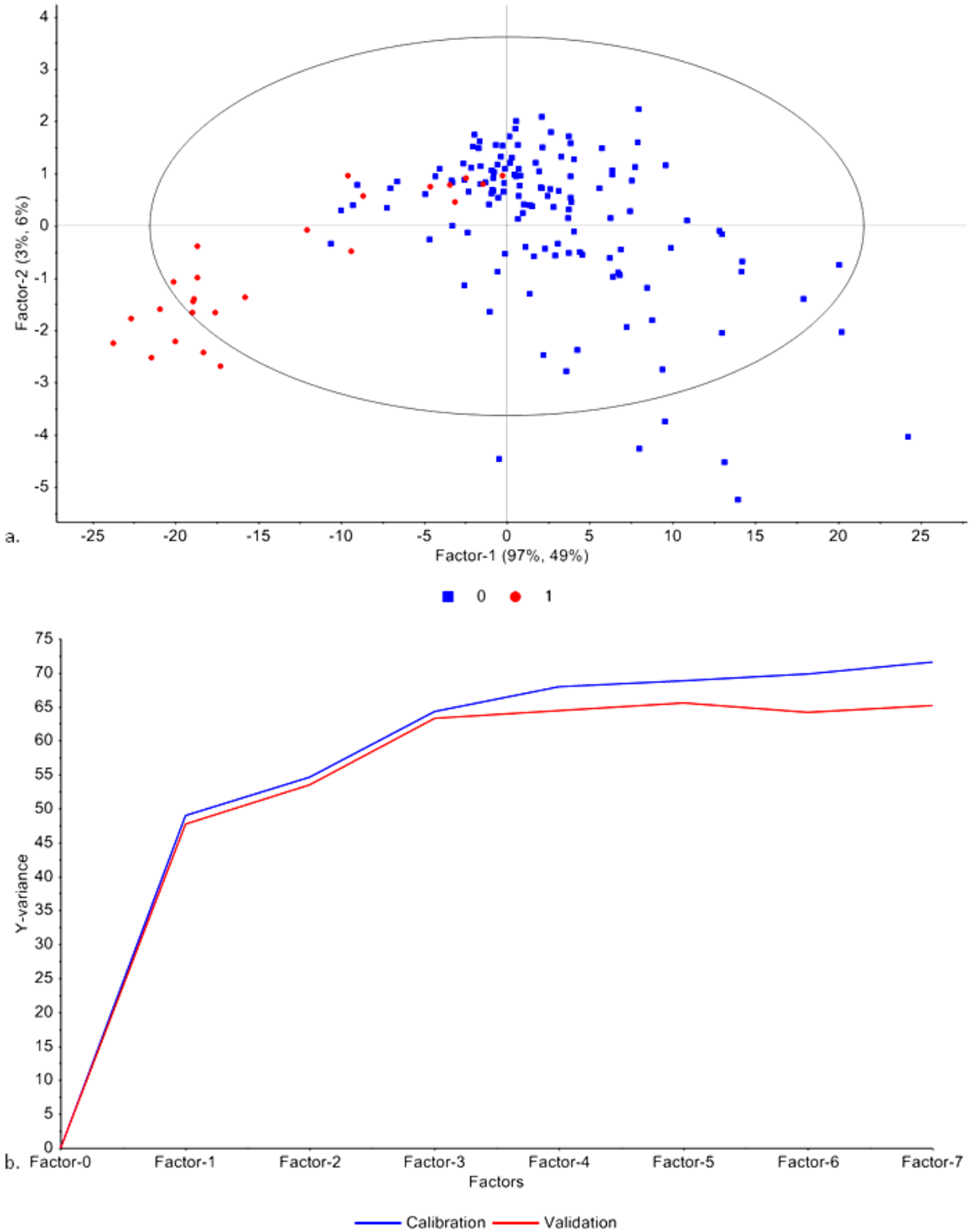


Figure 23. Two-class PLSDA models, a) scores and b) explained variance plots, raw 1400-2200nm. The presented model is built with mean-centered Set B training set sample spectra, n=157. Classes include "Forensic" (Group 0: known or estimated PMI ≤ 71 years) and "Non-forensic" (Group 1: historical or archaeological).

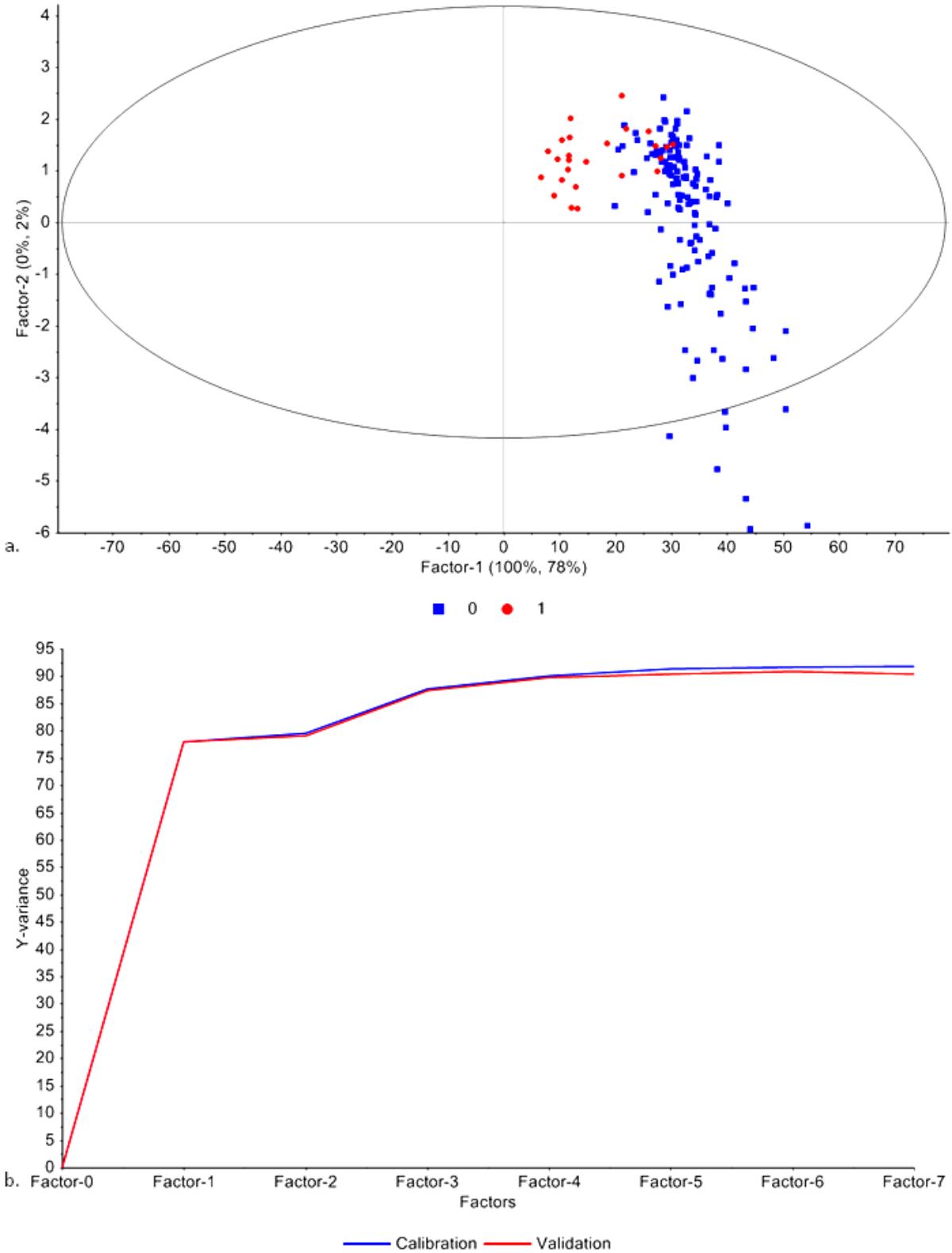


Figure 24. Two-class PLSDA models, a) scores and b) explained variance plots, raw 1400-2200nm. The presented alternate model is built with Set B spectral data corrected for sample size differences, n=157. Classes include "Forensic" (Group 0: known or estimated PMI ≤ 71 years) and "Non-forensic" (Group 1: historical or archaeological).

4.5.2 Three-Class Models (Classes: Forensic 1, Forensic 2, Non-forensic)

4.5.2.1 SIMCA

Three-class models were designed to see if bone cortex could be discriminated into one of three classes when PMI is known; classes included “Forensic 1” (Class 0: PMI = days-4.9 years), “Forensic 2” (Class 1: 5-71 years), and “Non-forensic” (Class 2: 500 or 900 years). Local PCAs used to build the Set A SIMCA model included a two component local PCA for Class 0 and one component PCAs for both Class 1 and Class 2. 48.4% of the test samples were correctly classified, including 65.0% of Class 0 samples and 67.0% of Class 2. However, no Class 1 samples were properly classified. 25.8% of samples were assigned to two classes, while 16.1% remained unassigned.

Results were similar for the Set B model. Components used for each class are as follows: Class 0 (2 PCs), Class 1 (2 PCs), and Class 2 (1 PC). 60.0% of the Class 0 and 33.3% of Class 2 test samples were correctly classified, while no Class 1 samples were properly assigned. Overall model accuracy was 43.8%, with 18.8% of samples being assigned to two or three classes and an equal amount remaining unassigned. Coomans plots for both three-class SIMCA models are presented in Figure 25 and Figure 26 below.

4.5.2.2 LDA-PCA

Two components were used for the Set A LDA-PCA 3Cl_F_1F_2N model. Cross-validated model accuracy is 88.6% (see Figure 27a). The model accurately classified 80.6% of the test set, including 80.0% of Class 0, 75.0% of Class 1, and all of the Class 2 samples. Reported accuracy for the Set B model (two components) is 88.0%; this model is presented in Figure 27b. Test set

classification accuracy for each class is identical to the Set A model (80.0, 75.0, and 100.0%, respectively). Overall test set classification accuracy using this model was 80.7%.

4.5.2.3 SVM-PCA

Two components were used for the three-class SVM-PCA model (Figure 28a). Internal cross-validated model accuracy (86.4%) was much higher than the actual model performance on the test set (71.0% accuracy). Although high classification accuracy was achieved by this model on samples belonging to Classes 0 (85.0%) and 2 (100.0), Class 1 accuracy was poor (25.0%).

Similarly, the two component Set B model validation is reported at 92.0% (Figure 28b). However, overall test set classification accuracy was much lower (68.8%). While 90.0% of Class 0 samples were properly identified, the model performed poorly at classifying samples in Classes 1 and 2 (33.3% for each).

4.5.2.4 PLSDA

Two-class PLSDA models perform by classifying samples into an “in” group (i.e. in Class 0) or an “out” group (i.e. not Class 0, then Class 1); an equivalent and redundant model for these two classes would assign a sample based on the rules of “in Class 1” or “not Class 1, then Class 0.” When dealing with more than two classes (e.g. Classes 0, 1 and 2), an individual model is used for each class, where:

- Model 1: Sample is in Class 0 or “not” Class 0 (in Class 1 or 2)
- Model 2: Sample is in Class 1 or “not” Class 1 (in Class 0 or 2)

- Model 3: Sample is in Class 2 or "not" Class 2 (in Class 0 or 1)

Note that these "submodels" will each be constructed from a recommended number of factors, and each will have associated RMSE and R^2 values as well.

The Set A three-class PLSDA-MC model consisted of the following three submodels: Class 0 (4 factors, RMSECV = 0.3344, R^2 = 0.5022), Class 1 (4 factors, RMSECV = 0.3498, R^2 = 0.6474), and Class 2 (4 factors, RMSECV = 0.1043, R^2 = 0.8683). See Figure 29. Overall prediction accuracy for the test set was 80.2%. The model performed well for test samples in Classes 0 and 2 (90.0% and 100.0% respectively); performance on Class 1 samples was much lower (50.0%). Relevant PLSDA-CO Set A model (Figure 30) statistics are as follows:

- Class 0: 3 factors, RMSECV = 0.3845, R^2 = 0.4086
- Class 1: 3 factors, RMSECV = 0.3424, R^2 = 0.8221
- Class 2: 3 factors, RMSECV = 0.1531, R^2 = 0.7442

The model performed similarly on test set samples for Class 0 (85.0%) and Class 2 (100.0%); Class 1 assignment was poor (12.5%). Overall test set classification accuracy was 67.7%.

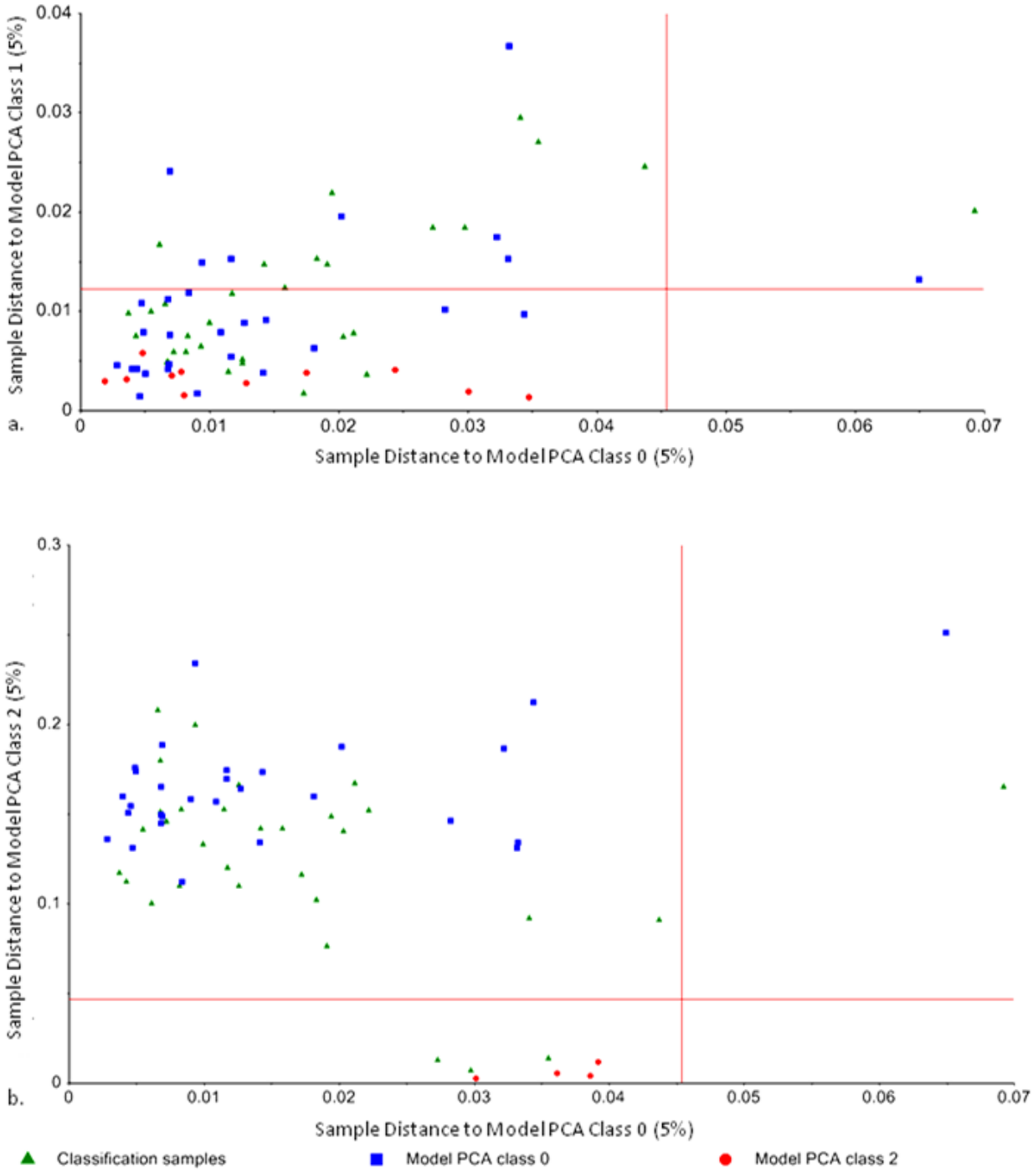
The Set B 3Cl_F_1F_2N PLSDA-MC model consisted of three submodels, each comprised of four factors with reported RMSECV and R^2 for each class as follows: Class 0 (0.3086, 0.5867), Class 1 (0.3122, 0.3910), and Class 2 (0.0837, 0.9479). This model is presented in Figure 31. 93.8% of the test samples were classified by the model, including 90.0% of Class 0 and all of Classes 1 and 2. The alternate Set B PLSDA-CO model consisted of three-factor submodels with reported RMSECV for Classes 0 through 2 being 0.3696, 0.3358, and 0.1810; associated R^2 values are 0.3171, 0.8238, and 0.7952, respectively (Figure 32). This model performed poorly relative to the mean-centered alternate, with only 68.8% of the test samples being accurately

classified. While Class 0 accuracy was high (90.0%), no Class 1 samples were identified, and

Class 2 accuracy was relatively low as well (66.7%).

Table 2. Model and test set classification accuracies for select three-class models (1400-2200nm). Classes 0 and 1 represent forensic samples with known PMI (days-4.9 years and 5-71 years, respectively); Class 2 samples are non-forensic specimens (PMI = 500 or 900 years). Numbers in "( )" represent the number of components or factors used in a given model. Cells in white represent cross-validated model accuracies as a percentage (LDA-PCA and SVM-PCA) or as RMSECV and R^2 for PLSDA. Areas in grey represent accuracies (%) from model application to the test set.

Model	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
3-class (Set A)	88.6 (2)	86.4 (2)	Class 0: RMSE = 0.3344 R^2 = 0.5022 (4) Class 1: RMSE = 0.3498 R^2 = 0.3474 (4) Class 2: RMSE = 0.1043 R^2 = 0.8683 (4)	Class 0: RMSE = 0.3845 R^2 = 0.4086 (3) Class 1: RMSE = 0.3424 R^2 = 0.8221 (3) Class 2: RMSE = 0.1531 R^2 = 0.7442 (3)	Overall = 48.4 Class 0 = 65.0 Class 1 = 0.0 Class 2 = 66.7 Both = 25.8 Neither = 16.1
	Overall = 80.6 Class 0 = 80.0 Class 1 = 75.0 Class 2 = 100	Overall = 71.0 Class 0 = 85.0 Class 1 = 25.0 Class 2 = 100	Overall = 80.2 Class 0 = 90.0 Class 1 = 50.0 Class 2 = 100	Overall = 67.7 Class 0 = 85.0 Class 1 = 12.5 Class 2 = 100	
3-class (Set B)	88.0 (2)	92.0 (2)	Class 0: RMSE = 0.3086 R^2 = 0.5867 (4) Class 1: RMSE = 0.3122 R^2 = 0.3910 (4) Class 2: RMSE = 0.0837 R^2 = 0.9479 (4)	Class 0: RMSE = 0.3696 R^2 = 0.3171 (3) Class 1: RMSE = 0.3358 R^2 = 0.8238 (3) Class 2: RMSE = 0.1810 R^2 = 0.7952 (3)	Overall = 43.8 Class 0 = 60.0 Class 1 = 0.0 Class 2 = 33.3 Both = 18.8 Neither = 18.8
	Overall = 80.7 Class 0 = 80.0 Class 1 = 75.0 Class 2 = 100	Overall = 68.8 Class 0 = 90.0 Class 1 = 33.3 Class 2 = 33.3	Overall = 93.8 Class 0 = 90 Class 1 = 100 Class 2 = 100	Overall = 68.8 Class 0 = 90.0 Class 1 = 0.0 Class 2 = 66.7	


a.
Sample Distance to Model PCA Class 1 (5%)
0.04
0.03
0.02
0.01
0
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
Sample Distance to Model PCA Class 0 (5%)
b.
Sample Distance to Model PCA Class 2 (5%)
0.3
0.2
0.1
0
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
Sample Distance to Model PCA Class 0 (5%)
Classification samples
Model PCA class 0
Model PCA class 2

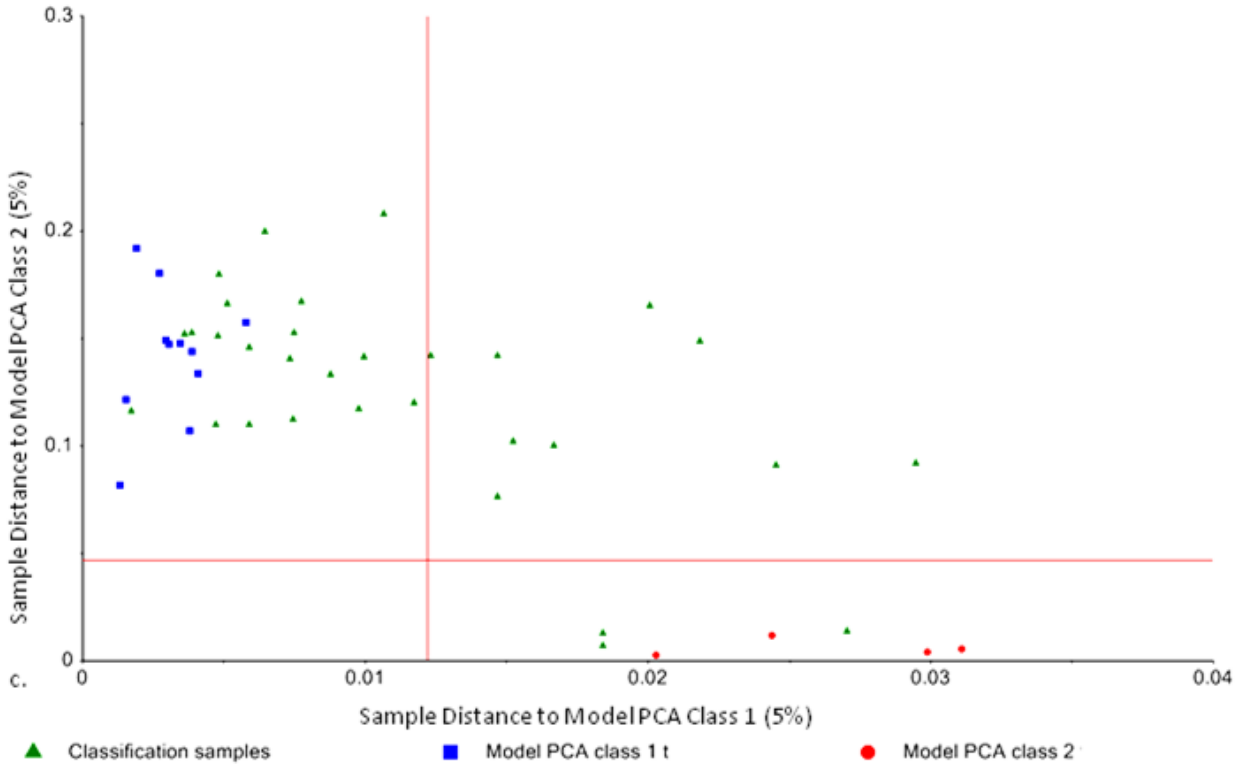
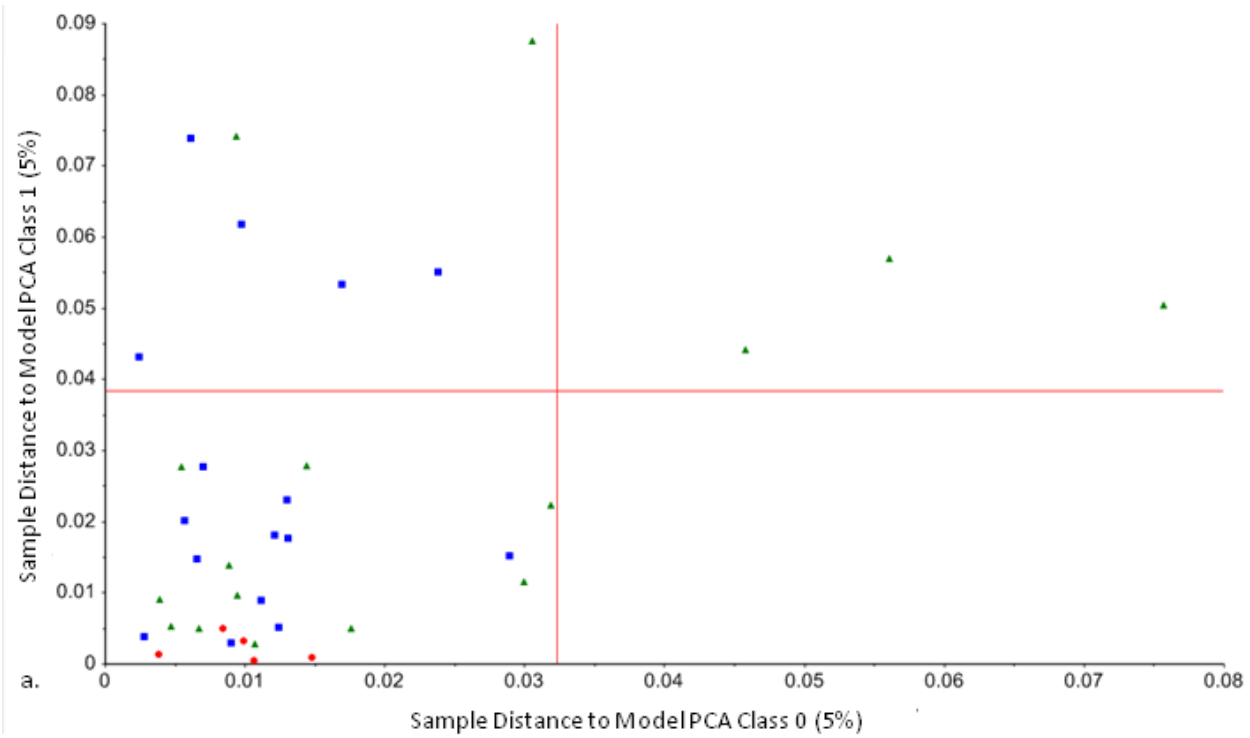
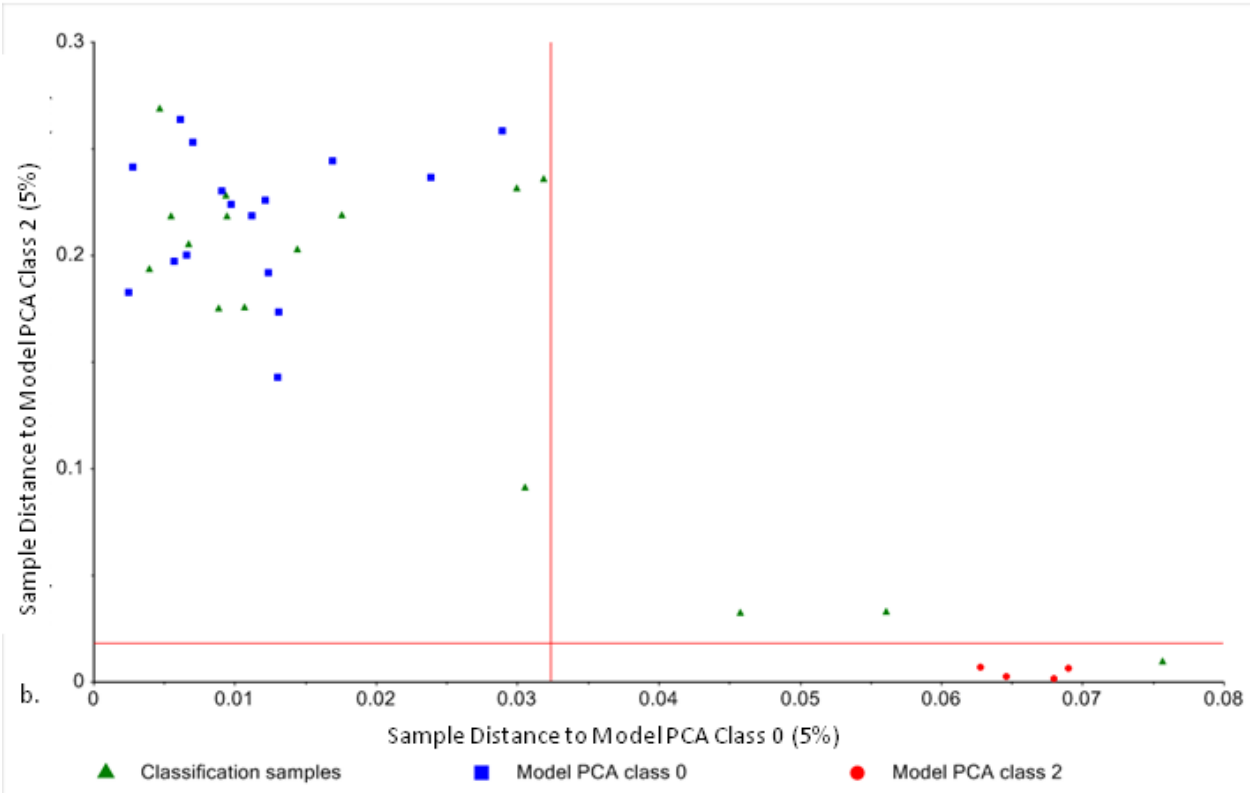


Figure 25. Coomans plots for three-class SIMCA models, raw 1400-2200nm. Models with classified samples (green) for Set A, n=44. Classes include “Forensic 1” (Group 0: PMI “days” - 4.9 years), “Forensic 2” (Group 1: PMI 5 - 71 years) and “Non-forensic” (Group 2: PMI 500-900 years). Presented plots show class assignment comparisons for groups (a) 0 or 1, (b) 0 or 2, and (c) 1 or 2, respectively.


0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
Sample Distance to Model PCA Class 1 (5%)
a.
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
Sample Distance to Model PCA Class 0 (5%)


0.3
0.2
0.1
0
Sample Distance to Model PCA Class 2 (5%)
b.
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
Sample Distance to Model PCA Class 0 (5%)
Classification samples
Model PCA class 0
Model PCA class 2

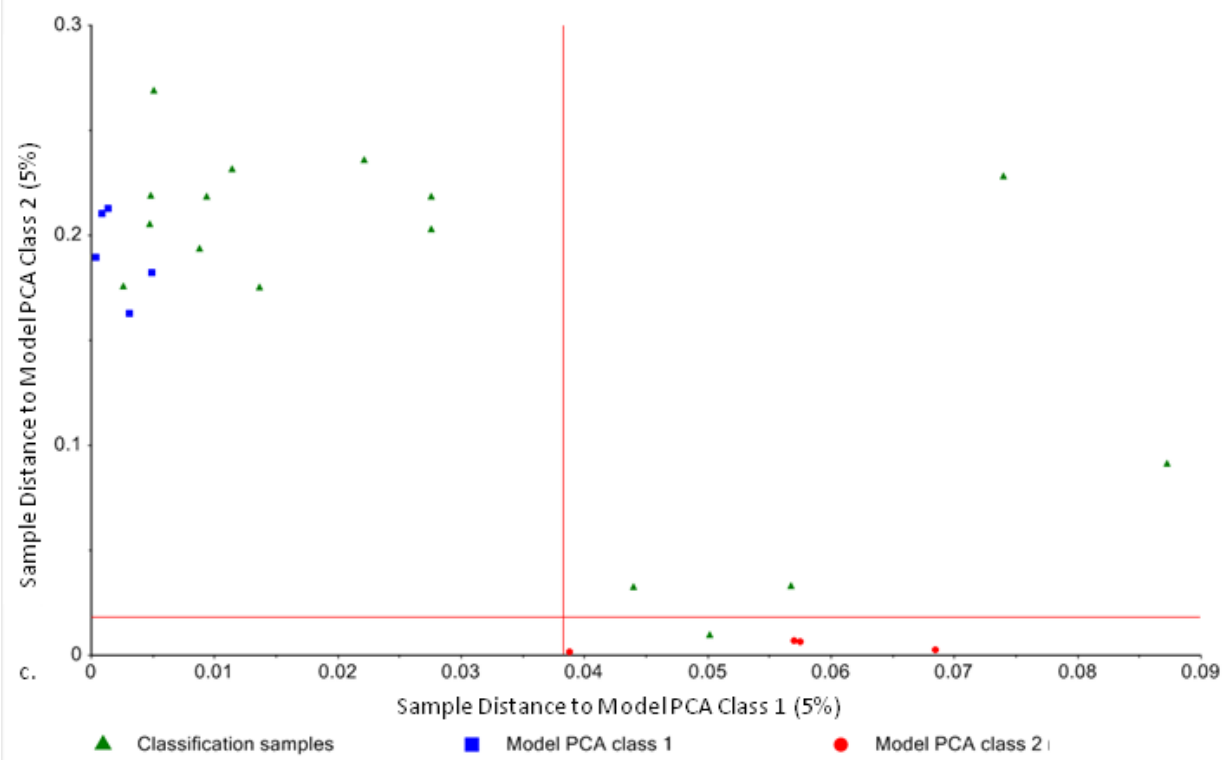


Figure 26. Coomans plots for three-class SIMCA models, raw 1400-2200nm. Models with classified samples (green) for Set B, n=25. Classes include “Forensic 1” (Group 0: PMI “days” - 4.9 years), “Forensic 2” (Group 1: PMI 5 - 71 years) and “Non-forensic” (Group 2: PMI 500-900 years). Presented plots show class assignment comparisons for groups (a) 0 or 1, (b) 0 or 2, and (c) 1 or 2, respectively.

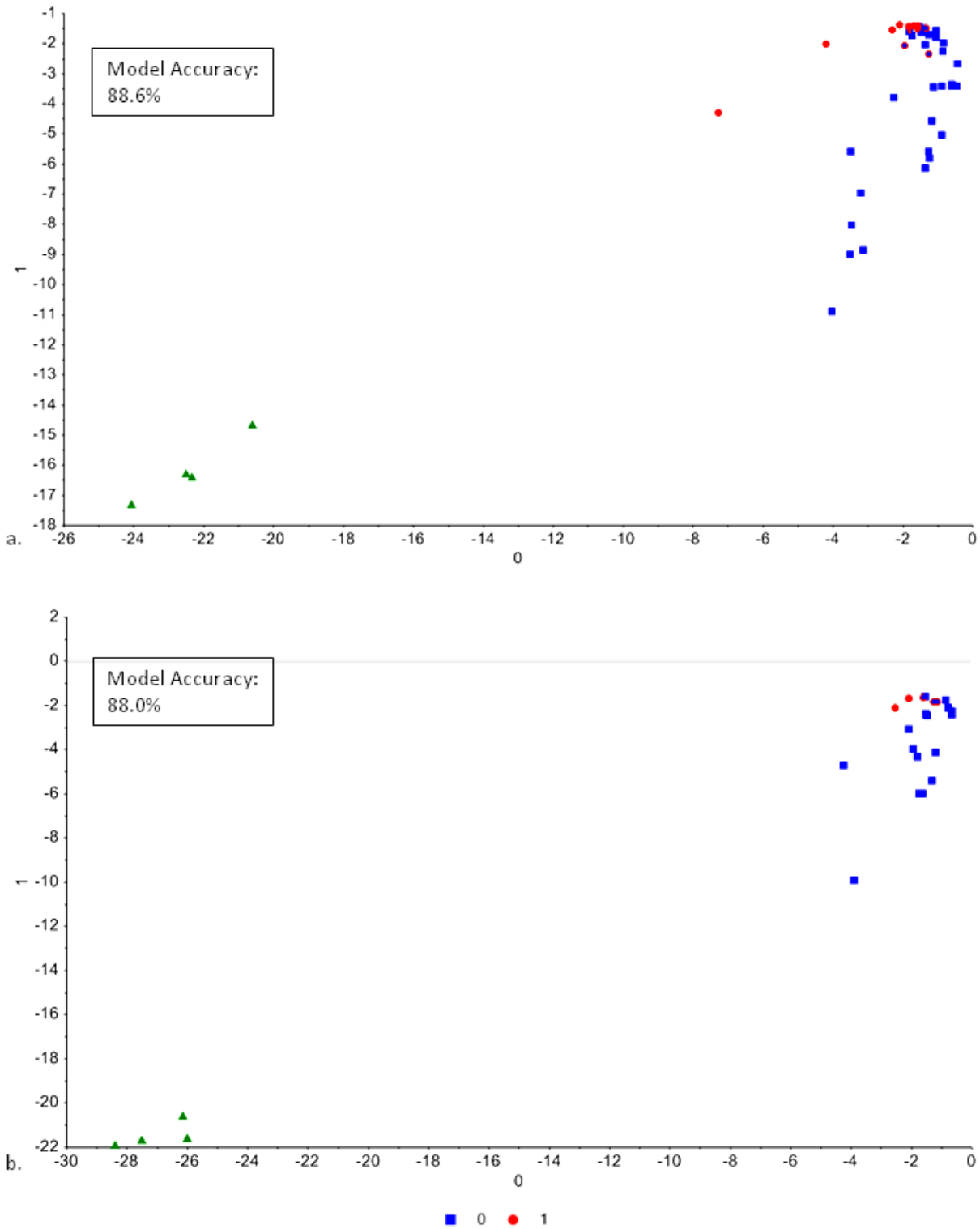


Figure 27. Three-class LDA-PCA models, raw 1400-2200nm. The Set A model (a) training set is n=44; Set B (b) is n=25 samples. Classes include “Forensic 1” (Group 0: PMI “days” - 4.9 years), “Forensic 2” (Group 1: PMI 5 - 71 years) and “Non-forensic” (Group 2: PMI 500-900 years).

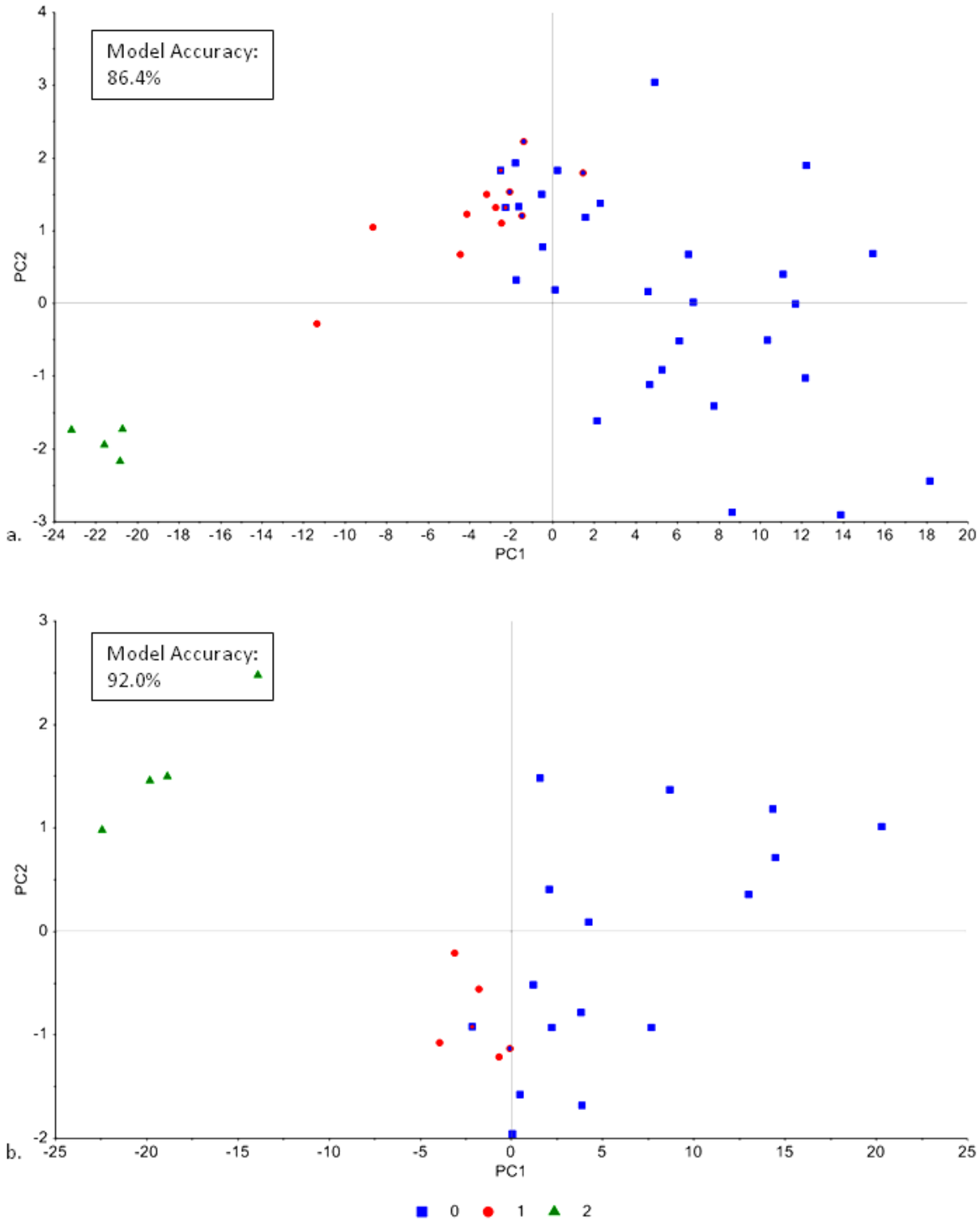


Figure 28. Three-class SVM-PCA models, raw 1400-2200nm. The Set A model (a) training set is n=44; Set B (b) is n=25 samples. Classes include “Forensic 1” (Group 0: PMI “days” - 4.9 years), “Forensic 2” (Group 1: PMI 5 - 71 years) and “Non-forensic” (Group 2: PMI 500-900 years).

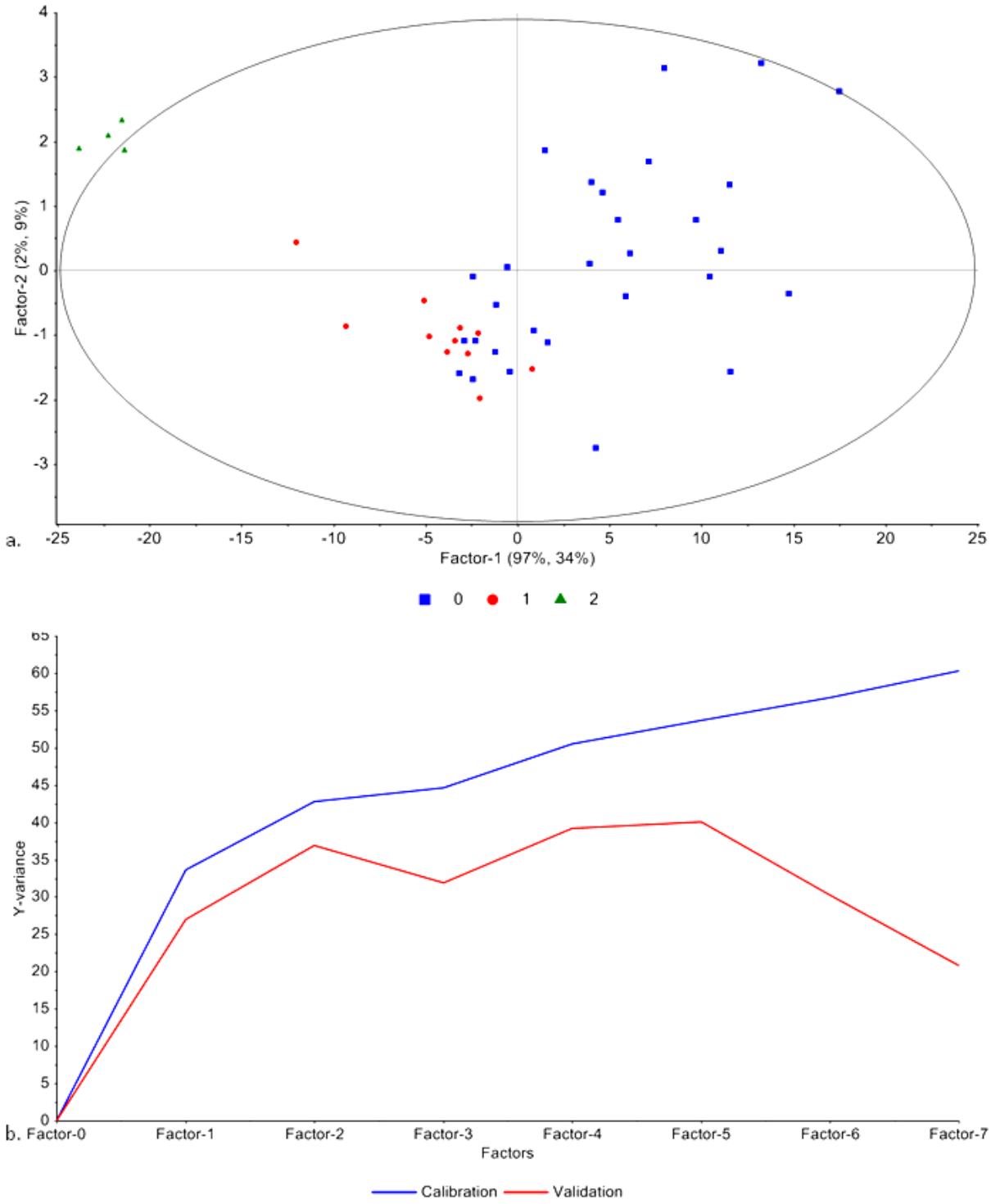


Figure 29. Three-class PLSDA models, a) scores and b) explained variance plots, raw 1400-2200nm. The presented model is built with mean-centered Set A training set sample spectra, n=44. Classes include “Forensic 1” (Group 0: PMI “days” - 4.9 years), “Forensic 2” (Group 1: PMI 5 - 71 years) and “Non-forensic” (Group 2: PMI 500-900 years).

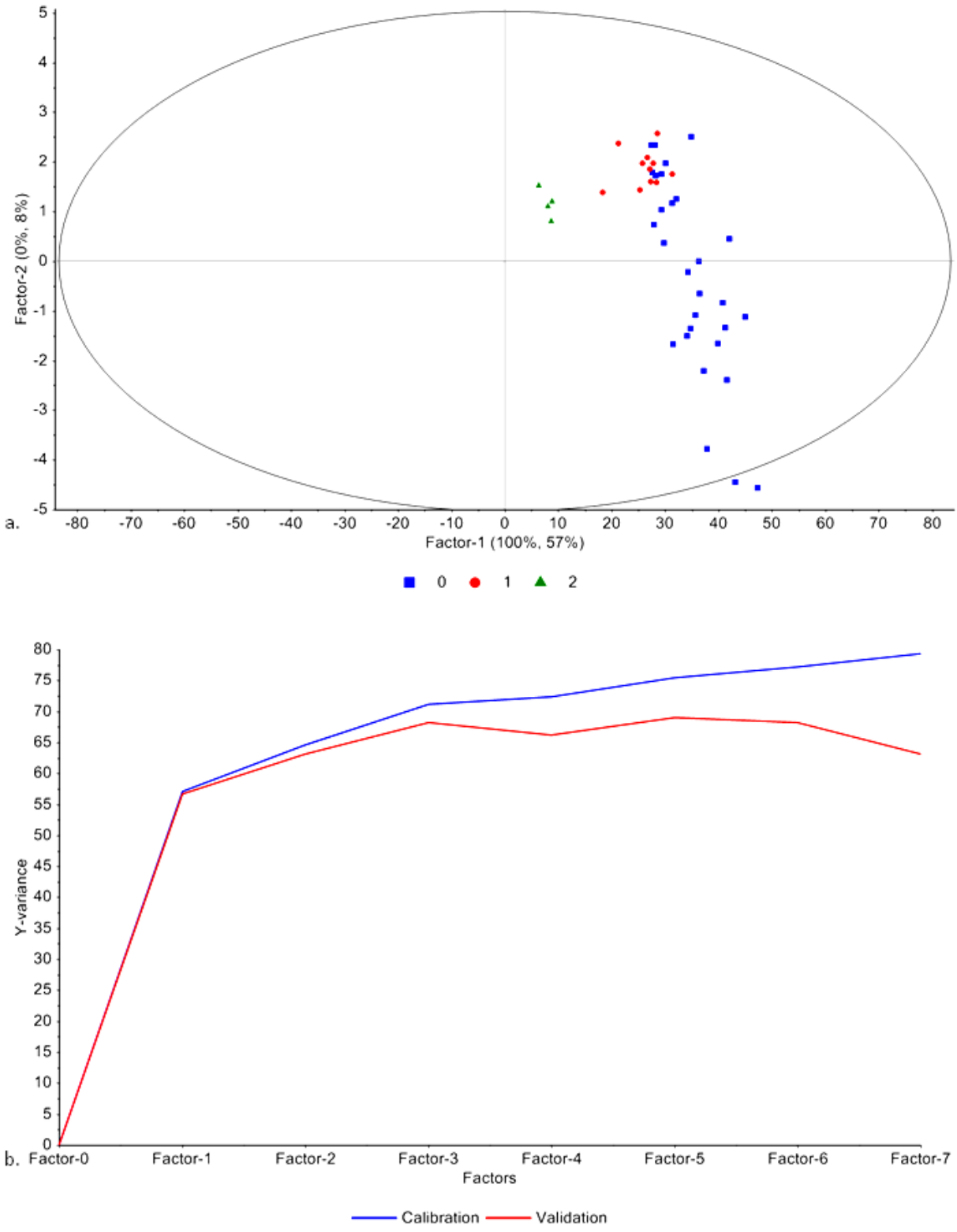


Figure 30. Three-class PLSDA models, a) scores and b) explained variance plots, raw 1400-2200nm. The presented alternate model is built with Set A spectral data corrected for sample size differences, n=44. Classes include "Forensic 1" (Group 0: PMI "days" - 4.9 years), "Forensic 2" (Group 1: PMI 5 - 71 years) and "Non-forensic" (Group 2: PMI 500-900 years).

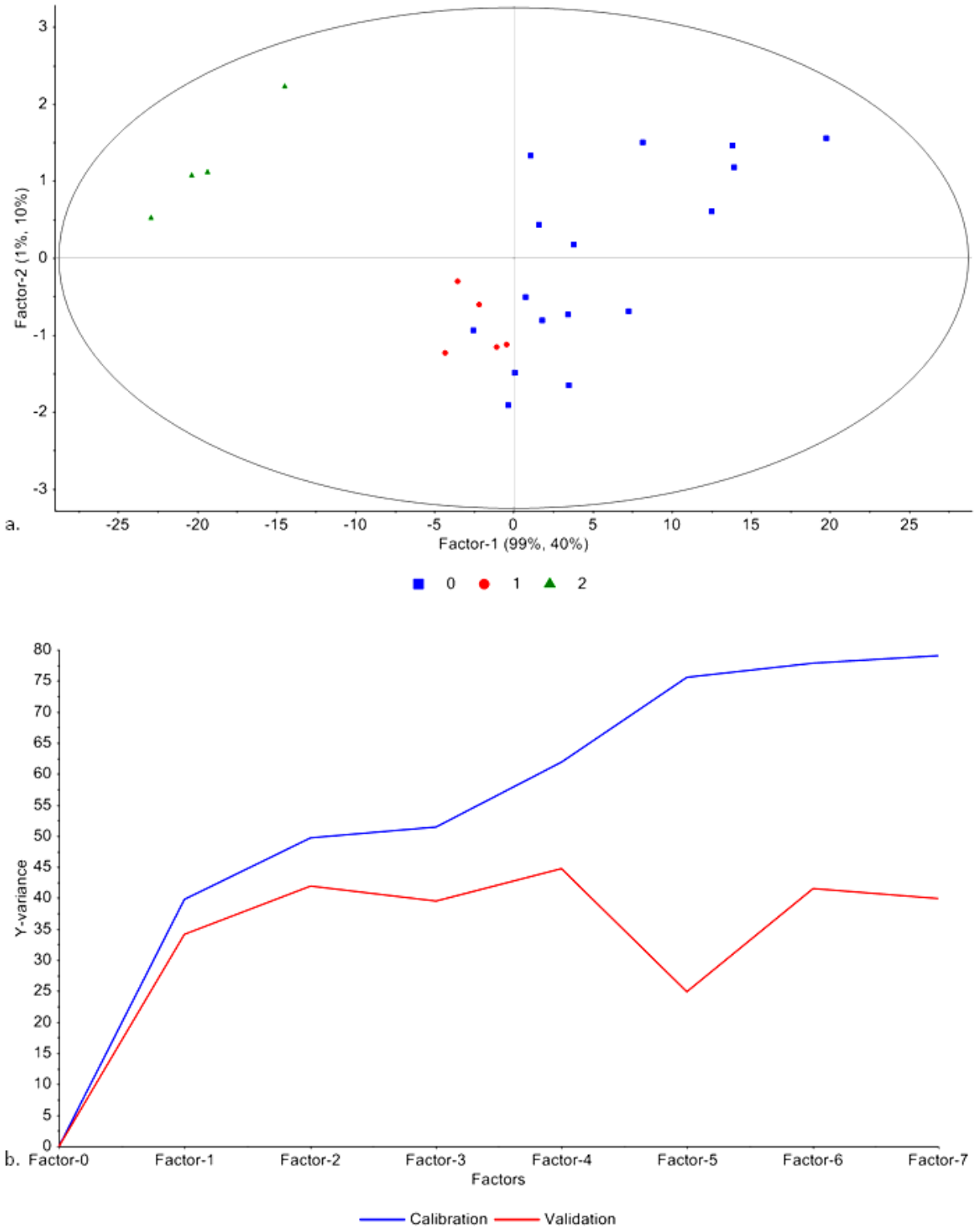


Figure 31. Three-class PLSDA models, a) scores and b) explained variance plots, raw 1400-2200nm. The presented model is built with mean-centered Set B training set sample spectra, n=25. Classes include “Forensic 1” (Group 0: PMI “days” - 4.9 years), “Forensic 2” (Group 1: PMI 5 - 71 years) and “Non-forensic” (Group 2: PMI 500-900 years).

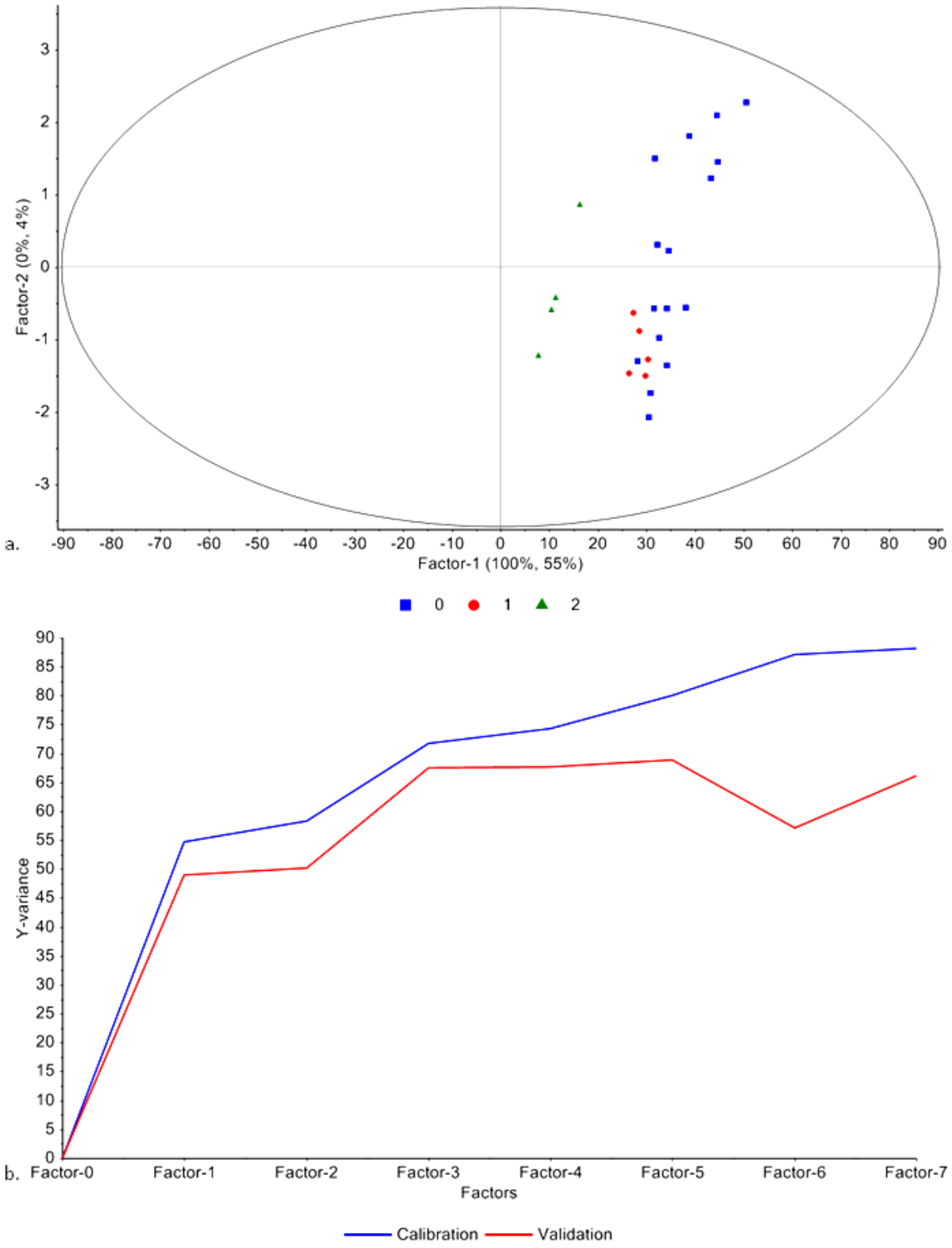


Figure 32. Three-class PLSDA models, a) scores and b) explained variance plots, raw 1400-2200nm. The presented alternate model is built with Set B spectral data corrected for sample size differences, n=25. Classes include "Forensic 1" (Group 0: PMI "days" - 4.9 years), "Forensic 2" (Group 1: PMI 5 - 71 years) and "Non-forensic" (Group 2: PMI 500-900 years).

4.5.3 Two-Class Forensic Models (Classes: Forensic 1, Forensic 2)

4.5.3.1 SIMCA

The last set of models was designed to determine if two classes of forensic samples could be accurately discriminated from each other. Forensic samples from the training and test sets used in the three-class models were applied for these models as well; the non-forensic samples were excluded. The Set A 2Cl_F_1F_2 SIMCA model was built with two local PCAs, a four component PCA for in Class 0 (Forensic 1), and a three component PCA for Class 1 (Forensic 2). Classification accuracy for Class 0 test samples was relatively low at 65.0%, while no Class 1 samples were properly classified. 28.6% of test samples were placed in both classes and the remainder were unclassified (10.7%). Overall classification accuracy was 46.4%.

Results were similar for the Set B model, which consisted of a four component Class 0 local PCA and three component Class 1 local model. Overall test set classification accuracy was 46.2%, with 60% of Class 0 samples being properly classified and no Class 1 samples with proper group assignment. The remaining samples were placed in either both (23.1%) or neither classes (7.7%). Coomans plots for the above models and resultant classified samples are presented in Figure 33.

4.5.3.2 LDA-PCA

The LDA-PCA 2Cl_F_1F_2 model on Set A samples used two components; model cross-validated accuracy is reported at 85.0% (Figure 34a). Overall test set classification accuracy was 78.6%. Class accuracies were 80.0% and 75.0% for Class 0 and Class 1, respectively.

Two components were used for the Set B model as well, with a resultant model accuracy of 90% (Figure 34b). Model performance on the test set was poorer than that observed in the Set A model; overall classification accuracy was 69.2%, with 80% of Class 0 samples being properly classified. Class 1 assignment was much lower (33.3%).

4.5.3.3 SVM-PCA

Cross-validated accuracy for the 2Cl_F_1F_2 SVM-PCA model on Set A samples was reported at 87.5% (Figure 35a). Two components were used to build the model. Test set classification accuracy was 60.7%, with Class 0 and Class 1 samples classified at rates of 75.0% and 25.0%.

Accuracy was much higher (100.0%) for the two component Set B model (Figure 35b). However, the model performed poorly on the test set; while 80.0% of Class 0 samples properly classified, only 33.3% of Class 1 samples were assigned to the correct class. Overall classification accuracy of the test set was 69.2%.

4.5.3.4 PLSDA

The Set A PLSDA-MC model required one factor; reported RMSECV and R^2 values are 0.3439 and 0.3719, respectively. 82.1% of test sample classes were properly predicted, including 85.0% of Class 0 samples and 75.0% of Class 1. For the alternate model on corrected data, two factors were used; RMSECV is reported at 0.3594, and the resultant R^2 was much higher (0.8218). However, overall prediction accuracy of the test set was much lower for this model (71.4%) rather than the mean-centered counterpart. Although 90.0% of Class 0 samples

were properly assigned, the model performed poorly on Class 1 samples (25.0%). See Figure 36 and Figure 37 for these models.

Similar results were observed in the Set B 2CL_F_1F_2 models. RMSECV and R^2 for the mean-centered model (one factor) were 0.3567 and 0.3215, respectively. The model accurately predicted only 76.9% of test samples, including 90.0% of Class 0 and 33.3% of Class 1. The one factor PLSDA-CO model reported a RMSECV of 0.3475 and a higher R^2 value of 0.8491. Overall test set prediction was the same with this model (76.9%). Although prediction accuracy for Class 0 was 100.0%, no Class 1 samples were properly identified. These models are presented in Figure 38 and Figure 39.

Table 3. Model and test set classification accuracies for select two-class models where non-forensic samples excluded (1400-2200nm). Class 0 represents forensic samples with known PMI of days to 4.9 years; Class 1 samples have known PMI of 5-71 years. Numbers in "( )" represent the number of components or factors used in a given model. Cells in white represent cross-validated model accuracies as a percentage (LDA-PCA and SVM-PCA) or as RMSECV and R^2 for PLSDA. Areas in grey represent accuracies (%) from model application to the test set.

Model	LDA-PCA	SVM-PCA	$PLSDA_{ctr}$	$PLSDA_{corr}$	SIMCA
2-class Forensic (Set A)	85.0 (2)	87.5 (2)	RMSE = 0.3439 (1) R^2 = 0.3719	RMSE = 0.3594 (2) R^2 = 0.8218	Overall = 46.4 Class 0 = 65.0 Class 1 = 0.0 Both = 28.6 Neither = 10.7
	Overall = 78.6 Class 0 = 80.0 Class 1 = 75.0	Overall = 60.7 Class 0 = 75.0 Class 1 = 25.0	Overall = 82.1 Class 0 = 85.0 Class 1 = 75.0	Overall = 71.4 Class 0 = 90.0 Class 1 = 25.0	
2-class Forensic (Set B)	90.0 (2)	100.0 (2)	RMSE = 0.3567 (1) R^2 = 0.3215	RMSE = 0.3475 (1) R^2 = 0.8491	Overall = 46.2 Class 0 = 60.0 Class 1 = 0.0 Both = 23.1 Neither = 7.7
	Overall = 69.2 Class 0 = 80.0 Class 1 = 33.3	Overall = 69.2 Class 0 = 80.0 Class 1 = 33.3	Overall = 76.9 Class 0 = 90.0 Class 1 = 33.3	Overall = 76.9 Class 0 = 100.0 Class 1 = 0.0	

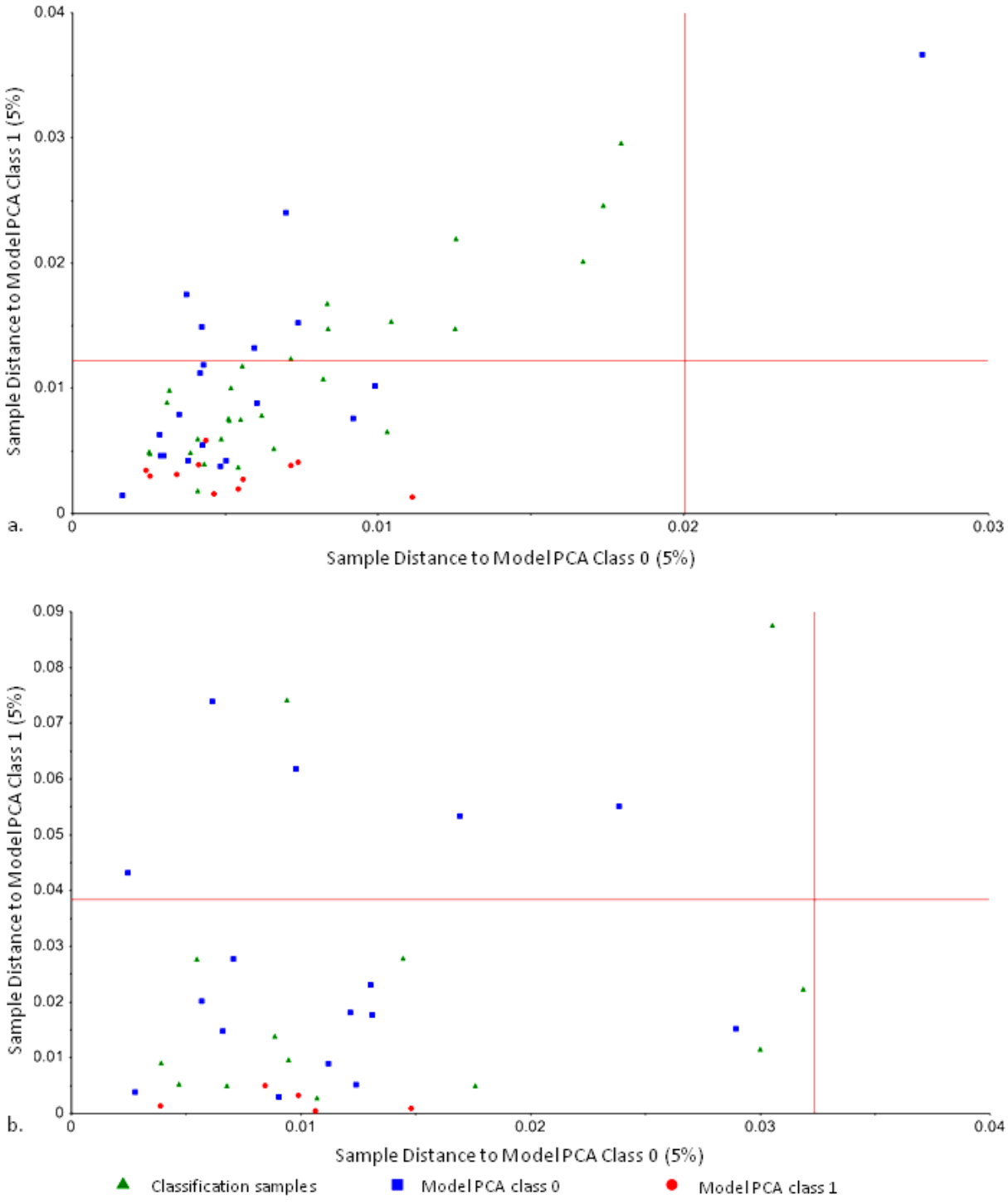


Figure 33. Coomans plots for two-class SIMCA models, raw 1400-2200nm. Models with classified samples (green) for Set A (a), n=40, and Set B (b), n=21, are presented. Classes include “Forensic 1” and (Group 0: PMI “days” - 4.9 years), “Forensic 2” (Group 1: PMI 5 - 71 years).

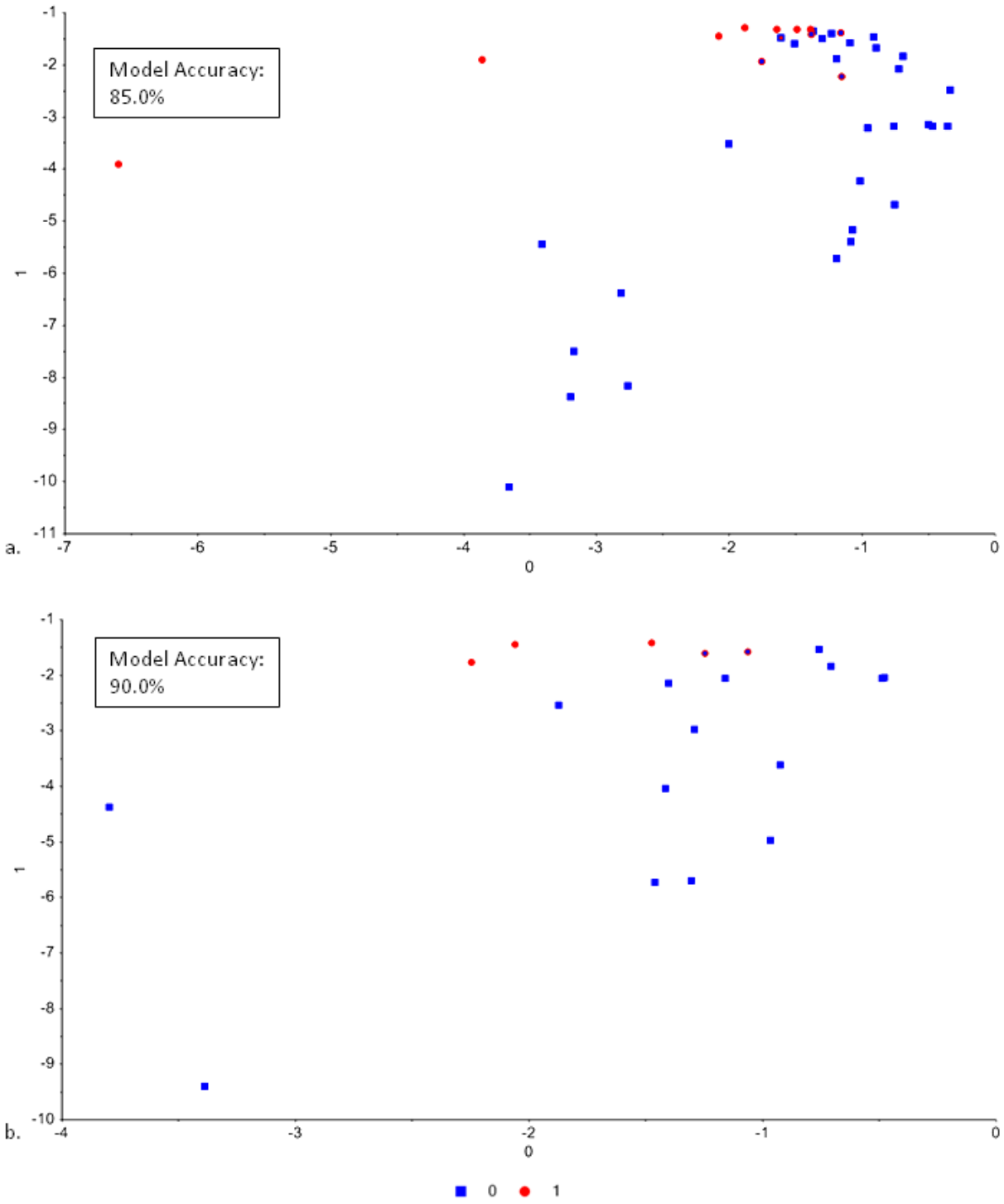


Figure 34. Two-class LDA-PCA models, raw 1400-2200nm. The Set A model (a) training set is n=40; Set B (b) is n=21 samples. Classes include “Forensic 1” and (Group 0: PMI “days” - 4.9 years), “Forensic 2” (Group 1: PMI 5 - 71 years).

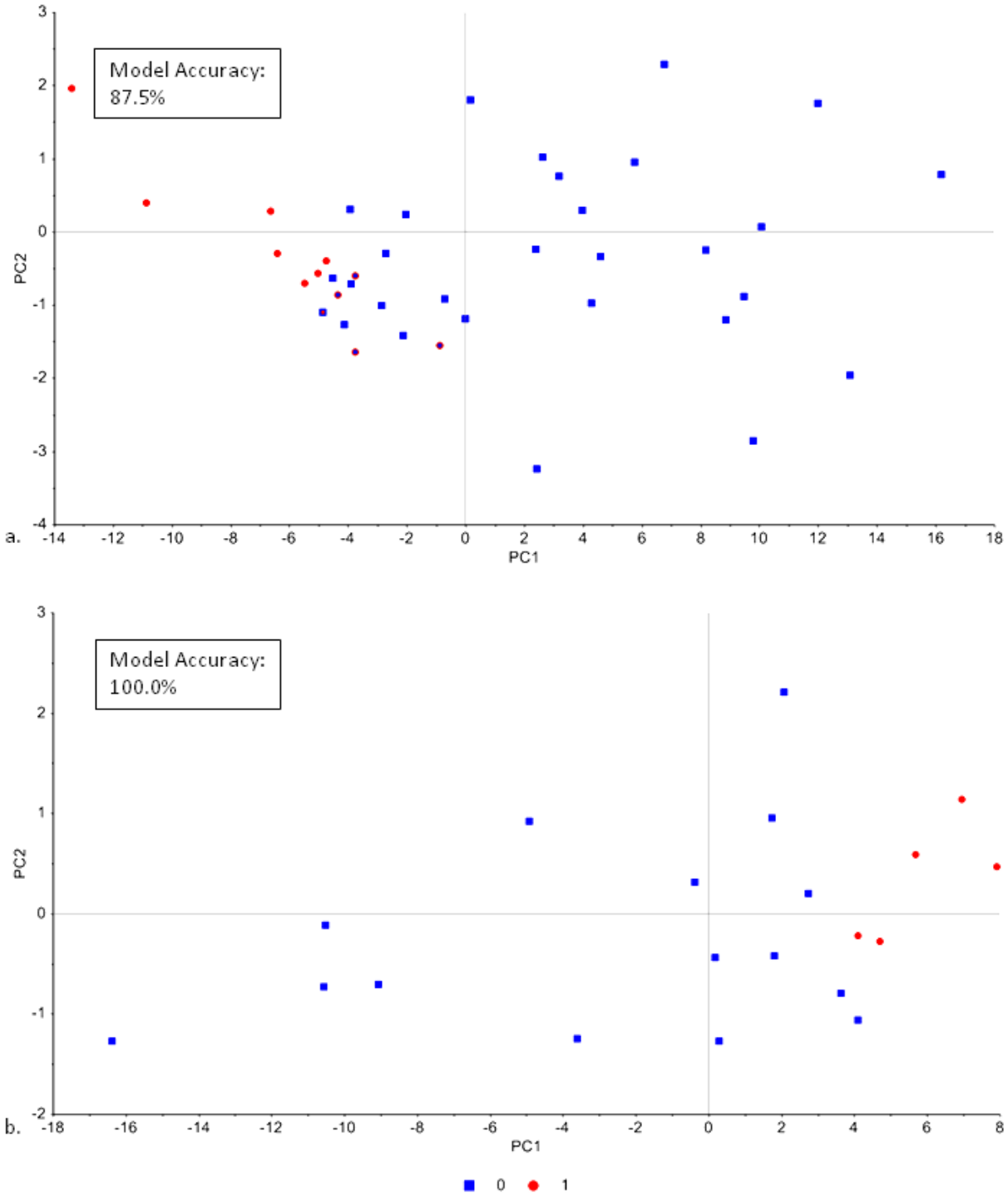


Figure 35. Two-class SVM-PCA models, raw 1400-2200nm. The Set A model (a) training set is n=40; Set B (b) is n=21 samples. Classes include “Forensic 1” and (Group 0: PMI “days” - 4.9 years), “Forensic 2” (Group 1: PMI 5 - 71 years).

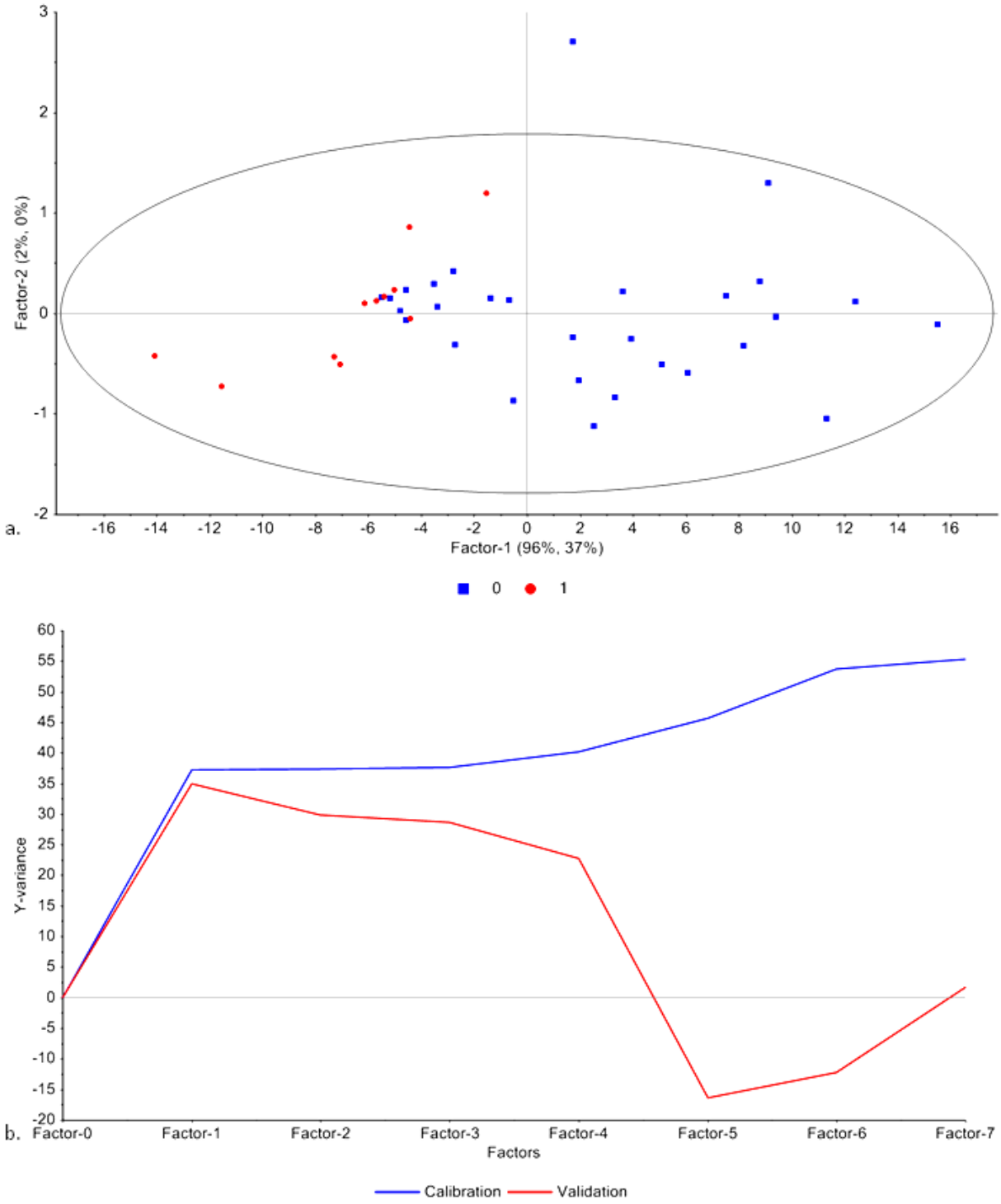


Figure 36. Two-class PLSDA models, a) scores and b) explained variance plots, raw 1400-2200nm. The presented model is built with mean-centered Set A training set sample spectra, n=40. Classes include “Forensic 1” (Group 0: PMI “days” - 4.9 years) and “Forensic 2” (Group 1: PMI 5 - 71 years).

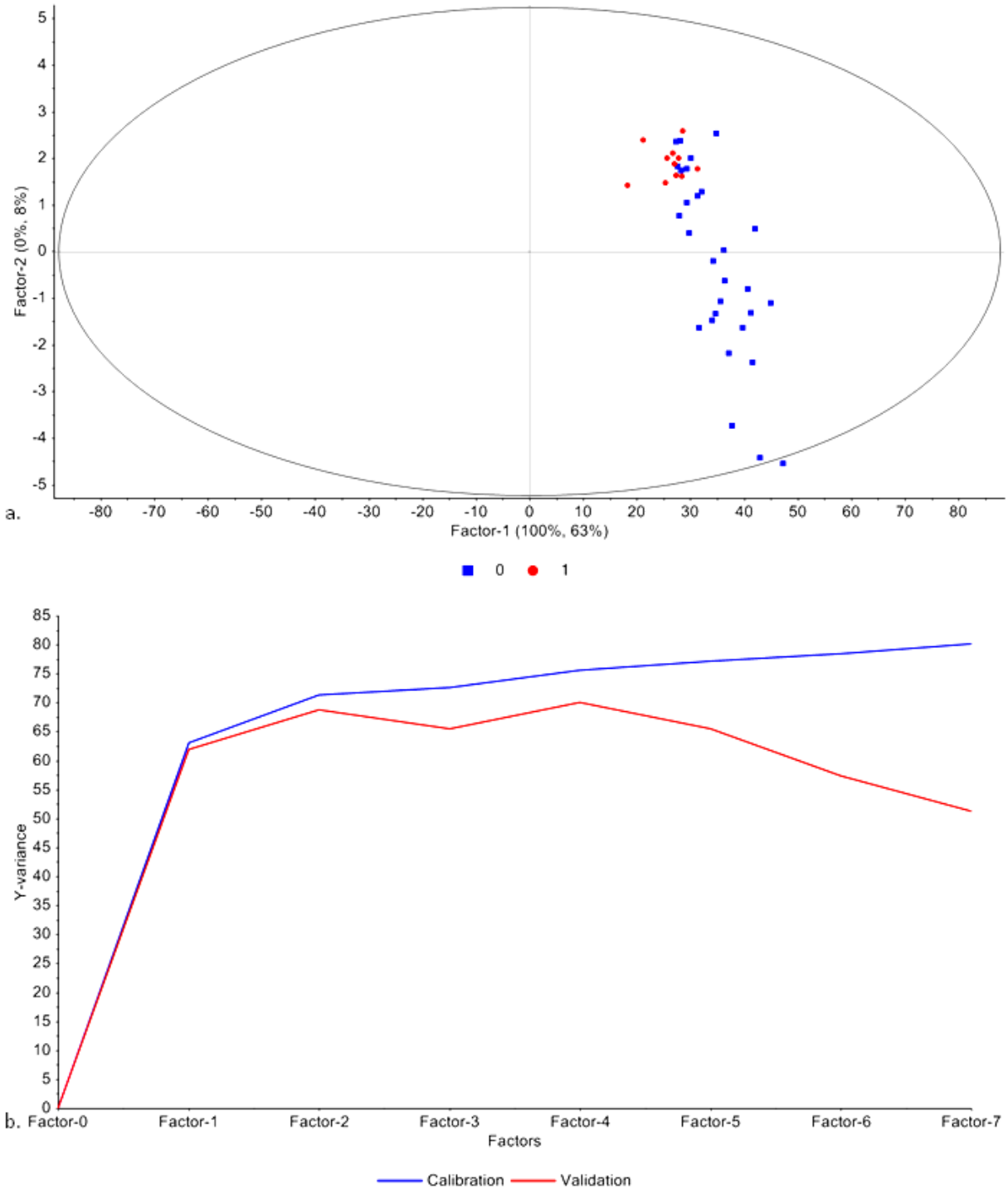


Figure 37. Two-class PLSDA models, a) scores and b) explained variance plots, raw 1400-2200nm. The presented alternate model is built with Set A spectral data corrected for sample size differences, n=40. Classes include "Forensic 1" (Group 0: PMI "days" - 4.9 years) and "Forensic 2" (Group 1: PMI 5 - 71 years).

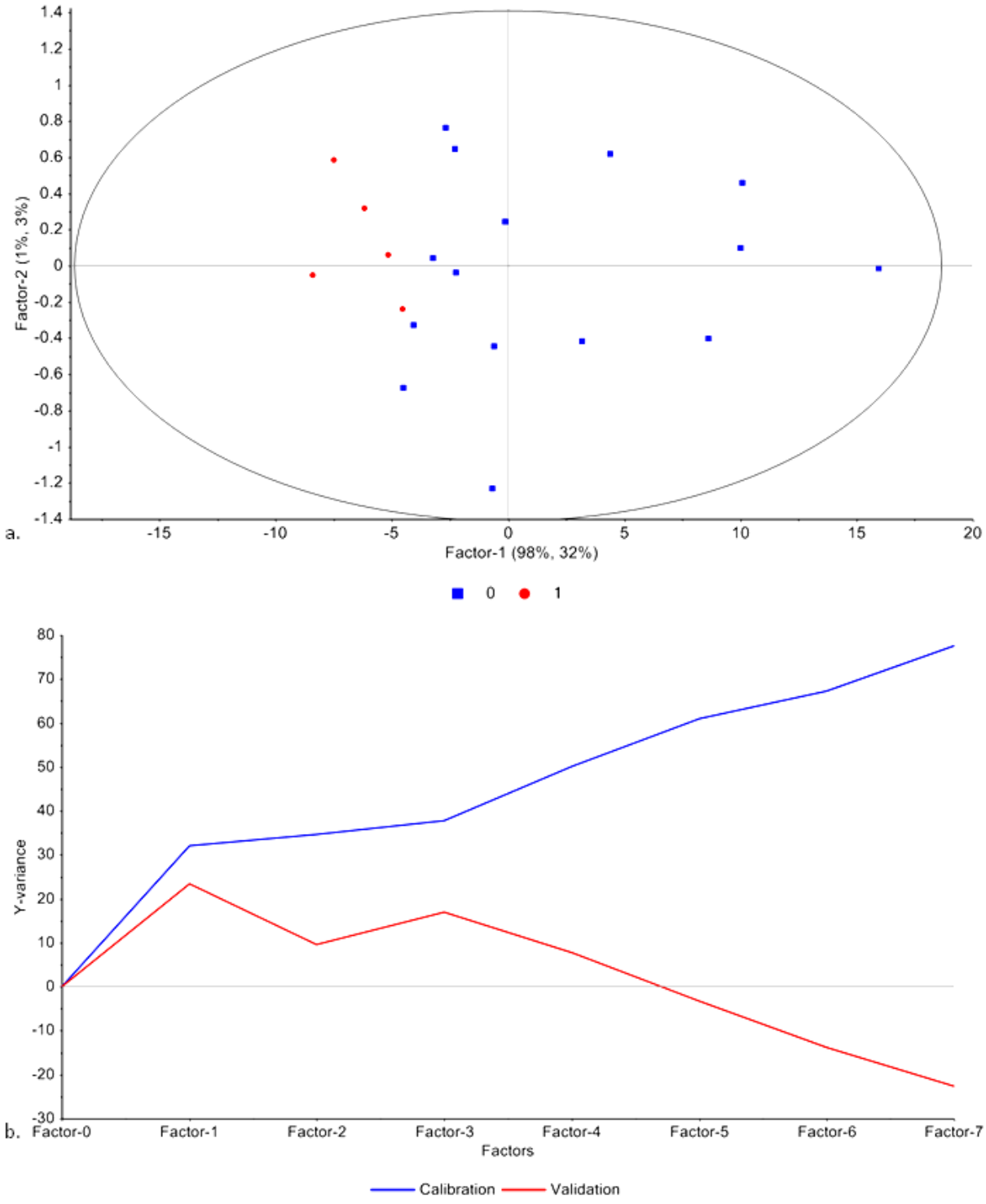


Figure 38. Two-class PLSDA models, a) scores and b) explained variance plots, raw 1400-2200nm. The presented model is built with mean-centered Set B training set sample spectra, n=21. Classes include “Forensic 1” (Group 0: PMI “days” - 4.9 years) and “Forensic 2” (Group 1: PMI 5 - 71 years).

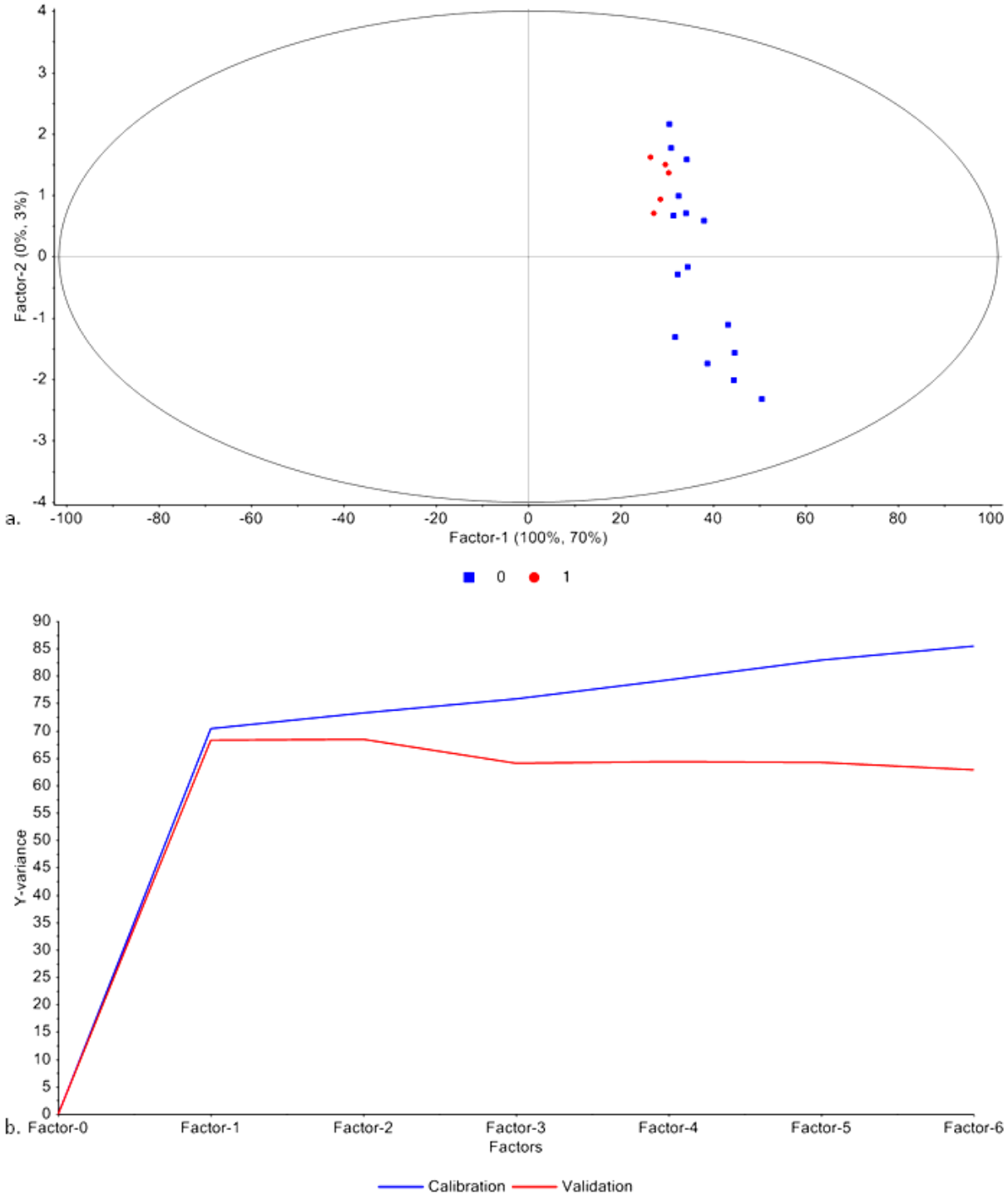


Figure 39. Two-class PLSDA models, a) scores and b) explained variance plots, raw 1400-2200nm. The presented alternate model is built with Set B spectral data corrected for sample size differences, n=21. Classes include "Forensic 1" (Group 0: PMI "days" - 4.9 years) and "Forensic 2" (Group 1: PMI 5 - 71 years).

4.6 Tests for Protein Oxidation

Following completion of the analyses reported above, a secondary study was conducted on a limited number of samples in order to determine if the protein content of extracted bone samples oxidizes over time. Experiments were conducted in February 2017 at ASD-PANalytical in Boulder, Colorado, and involved the collection of spectra from two groups of cortical samples, including the following: 16 samples used in the May 2013 preliminary study ("Resampled Set"); and, six samples of femoral cortex that were treated with a hydrogen peroxide H_2O_2 dilution series for a fixed period of time ("Dilution Set"). Resampled Set samples "Time 2" spectra were compared with "Time 1" spectra from 2013 (approximate time interval 4 years) in order to 1) determine if the spectra change with time and, if so, 2) identify the changes and tie them to potential chemical oxidative changes. Spectra collected from Dilution Set samples were analyzed to see if there were changes in absorbance spectra with respect to H_2O_2 concentration vs. control samples stored in water and air.

Minor peak changes were observed on some of the Resample Set spectra (SNV-transformed) at positions 1490nm, 1730nm, and 1940nm. Similar changes were observed in the Dilution Set series. These observed changes may be related to the presence of elevated carbonyl (1490nm, 1940nm) and methyl-related oxidative products; however, this currently cannot be determined without performing composition analyses, which were not a part of the broader study. It should be noted that the observed changes were relatively small and would likely not have a negative effect on the classification models reported above and in Appendix F. A more in depth discussion, including background, procedures, interpretation, and relevant figures, is presented in Appendix G.

CHAPTER 5

DISCUSSION

5.1 Introduction

This chapter discusses the results previously reported. The first sections of the chapter present explanations regarding why 1) treating the superficial surface with a grinding wheel prior to spectral collection is optimal and 2) why the spectra collected from the longitudinally and transversely cut cortical surfaces are essentially the same. This section is followed by a discussion of the models presented in Chapter 4. The next portion of this chapter will describe the use of transformed-data models and when data transformation may be desirable. The final segment will discuss the limitations regarding the tests for protein oxidation that were conducted after the broader study was completed.

5.2 Superficial Surface Comparison: Untreated versus Treated

It has been noted that spectra collected from the untreated cortical surface yielded poor class separation, as observed on the spectral line plots, in comparison to spectra collected from the same surface following treatment with a grinding wheel. This is most likely due to the presence of postmortem artifacts that artificially affect the spectral signature, e.g. the presence of soil staining, plant debris, and rootlets. Additionally, weathering and exfoliation of the outer cortex tend to make the spectrum appear more consistent with those of older specimens (Figure 40). Remnants of detergents used for maceration, or, accumulations of adipocere also frequently resulted in somewhat abnormally shaped spectra. The removal of superficial cortex by light grinding allows for direct scanning of the bone itself absent soil stains, plant material,

flaking cortex, and other non-bone materials. This is most pronounced when the bone is weathered in appearance but the PMI is relatively short: relatively new bone looks much more "fresh" when the flaky outer cortex is removed. In many ways, the treated surface approximates what is observed on sectioned longitudinal and transverse cortical edges, and better class separation is achieved both grossly in the line plots and statistically via PCA with spectra that are collected from the treated surface.

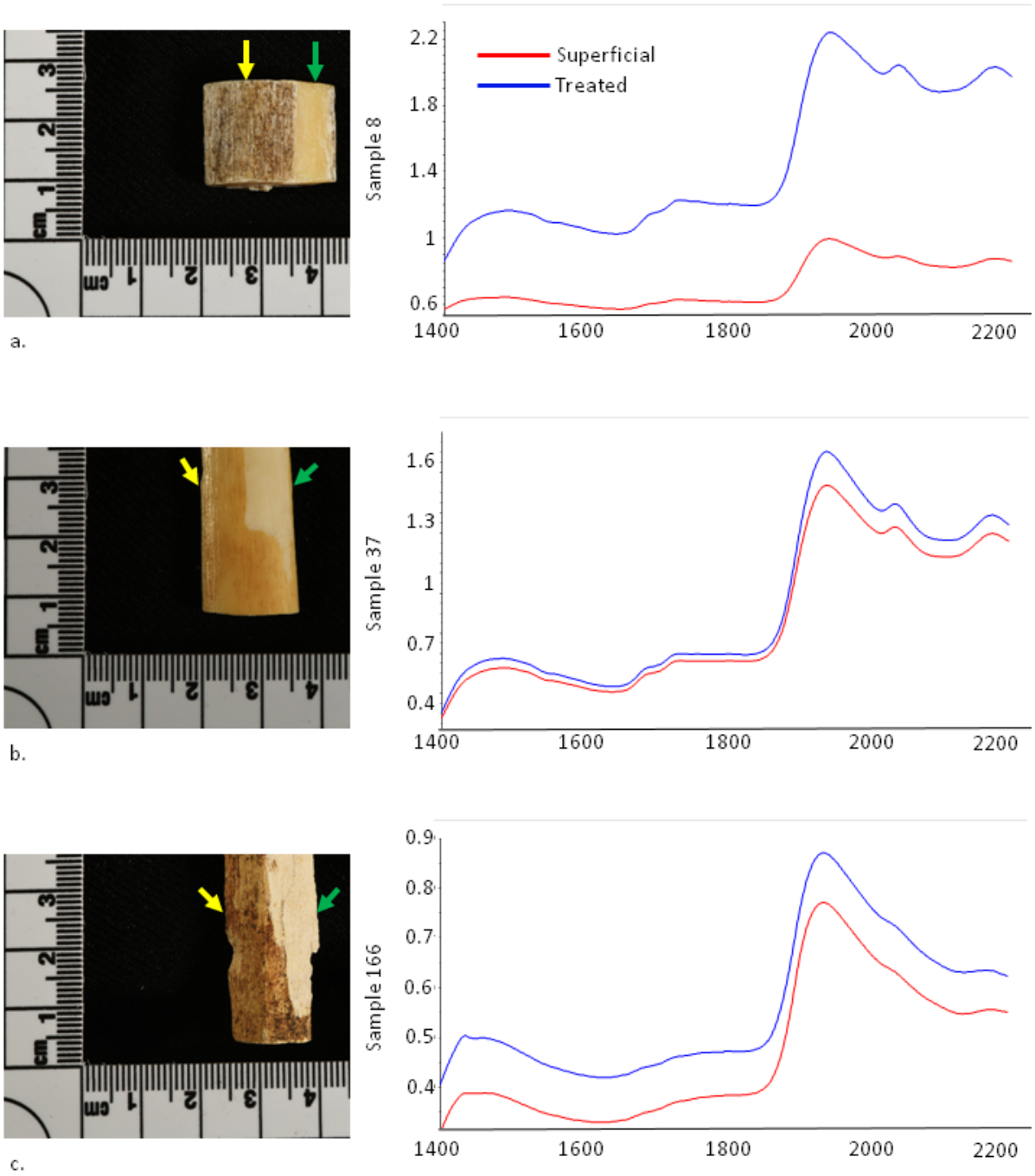


Figure 40. Comparison of raw spectra collected from the superficial (yellow arrow) and treated superficial (green arrow) surfaces of 3 cortical samples, 1400-2200nm selected. Images of the cortical samples (left) and associated spectra (right) are presented. Samples include a) a weathered femur (PMI<1 year), b) a non-weathered femur (PMI estimated 20-50 years), and c) an archaeological specimen.

5.3 Sectioned Surface Comparisons: Transverse versus Longitudinal

Bone is a biphasic composite of primarily collagen and mineral apatite. In compact bone, the mineralized collagen fibrils are organized into osteons which are arranged such that the bone can compensate for suite of applied forces. Osteons, and thus the collagen fibrils that are a major component of them, can be organized in longitudinal, transverse, and oblique orientations, and the ratio of these orientations will ultimately be dependent on both the bone in question and the position on the a given bone.

The diaphysis of a long bone, such as the femur, will contain many osteons (and fibers) that are longitudinally oriented in order to compensate for compressive forces, although other orientations are present as well. Thus, spectra collected from the longitudinal sectioned surface are the result of fibers that are oriented primarily parallel to the long axis, while those spectra collected from the transversely sectioned cortex result from viewing the fiber termini. Slight differences between the longitudinal and transverse section spectra are expected owing to fiber orientation and slight variations in their anisotropic properties. However amino acid composition along a collagen fibril is highly consistent; thus, longitudinal and transverse spectra are essentially the same morphologically and would be expected to provide redundant information (Figure 41). From a practical standpoint, sampling the “long cut” on the cortex is preferable because there is a greater amount of sampling surface available, and the anterior diaphyseal curvature can be avoided.

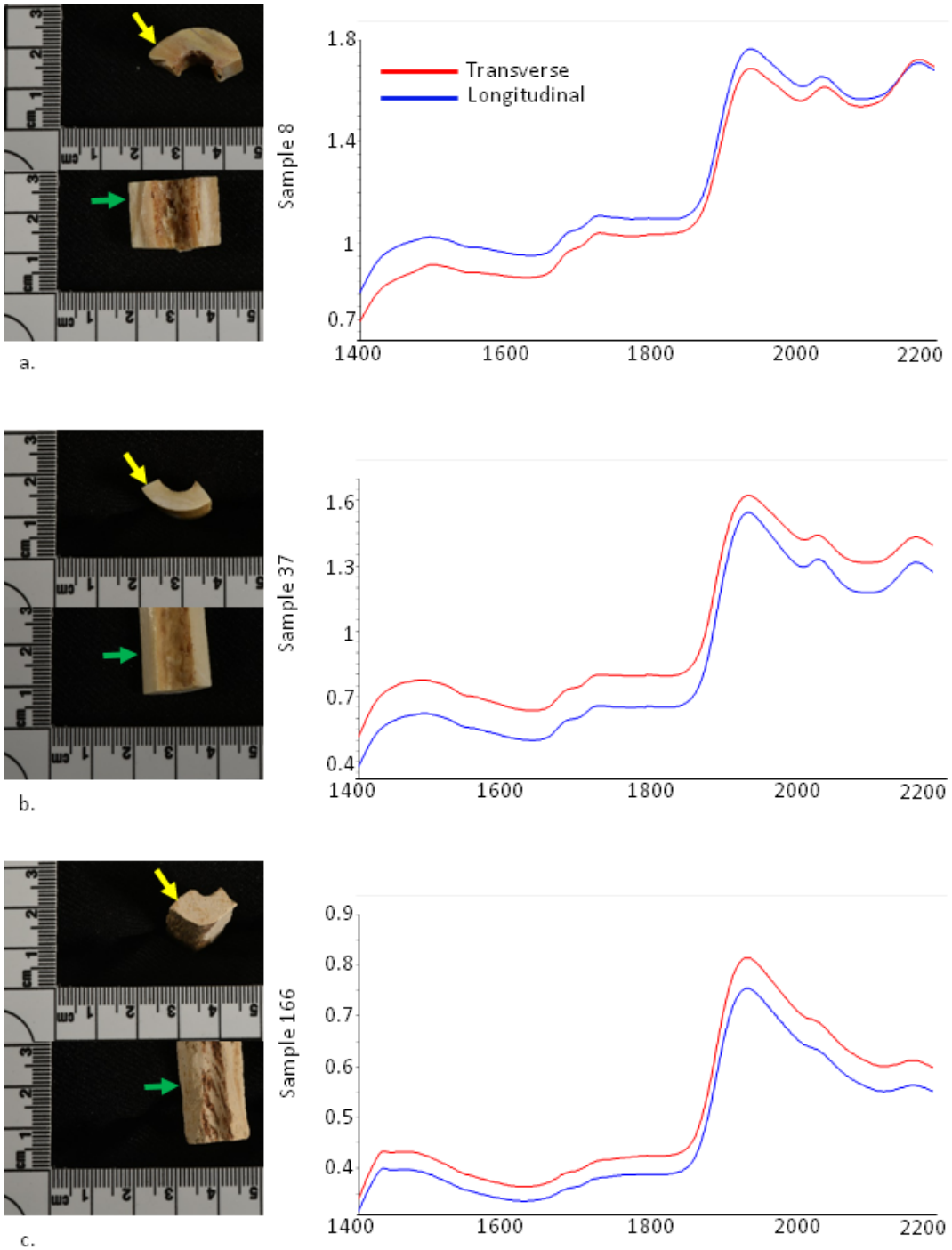


Figure 41. Comparison of raw spectra collected from the transverse (yellow arrow) and longitudinal (green arrow) cut surfaces of 3 cortical samples, 1400-2200nm selected. Images of the cortical samples (left) and associated spectra (right) are presented. Samples include a) a weathered femur (PMI<1 year), b) a non-weathered femur (PMI estimated 20-50 years), and c) an archaeological specimen.

5.4 Discussion of the Models

5.4.1 Performance of the SIMCA Models

A review of all developed models revealed that of the selected algorithms, SIMCA consistently performed the poorest on both raw and transformed data. The primary issue was that many samples were ultimately assigned to multiple classes or were given no class assignment at all. The best SIMCA results were observed with 2Cl_FN, although classification accuracies for both classes were low. When non-forensic samples were excluded, older forensic test samples run through two-class forensic models often classified at rates approaching 0.0%. These older forensic cortices were poorly classified in the three-class models as well. The consistently poor results achieved with SIMCA suggest that this classification method is not robust enough to use in this PMI classification application. Accordingly, the remainder of the discussion will focus on models built with the other selected classification algorithms (LDA-PCA, SVM-PCA, and PLSDA).

5.4.2 Performance of the 1400-2200nm Band Raw Models

2Cl_FN, 2Cl_F_1F_2, and 3Cl_F_1F_2N models were built using raw spectral data (1400-2200nm) collected from the longitudinally-cut cortical surface (Set A) and the treated superficial surface (Set B) of training samples. In most cases, Set A models outperformed Set B counterparts when classifying the test sets. For the 2Cl_FN models, LDA-PCA and SVM-PCA performed nearly the same on the Set A data. Both PLSDA models (mean centered and corrected) performed well at classifying Class 0, but yielded lower classification rates for Class 1 with respect to the LDA-PCA and SVM-PCA models.

For the Set A 2Cl_F_1F_2 and 3Cl_F_1F_2N models, the LDA-PCA models performed well, although class accuracies were lower than those in the 2Cl_FN model. The corresponding SVM-PCA models performed poorly at classifying Class 1 (25% correct classification in both models). Likewise, the PLSDA-CO models poorly classified Class 1 in both cases, 25% and 12.5%, respectively. The Set A PLSDA-MC 2Cl_F_1F_2 model outperformed the corresponding LDA-PCA model. Similar results were not observed in the three-class PLSDA-MC model; although classification accuracy for Class 0 exceeded that seen with the LDA-PCA model, Class 1 accuracy dropped to 50%, vs. 75.0%. These results indicate that for the raw 1400-2200nm band, LDA-PCA models appear to provide the most consistent classification results for the Set A samples.

As noted, Set A models outperformed Set B counterparts in all but one case. The 2Cl_FN models poorly classified non-forensic Set B samples. Older forensic samples were generally classified at lower accuracy in the 2Cl_F_1F_2 and 3Cl_F_1F_2N models as well. The exception to this was the three-class Set B PLSDA-MC model, which provided high overall accuracy of 93.8%, with 90% of Class 0 properly classified 100% accuracy for the other two classes; however, the results of this external validation are suspect. Although the model accurately classified all of the older forensic test specimens, Class 1 performed poorly during the internal cross-validation phase (40% accuracy). This, along with the generally observed lower classification rates for “older” forensic specimens, suggests that this class is poorly defined. Please see section 7.3.3 for further discussion about this issue.

Alternate PLSDA models were built using mean centered (MC) spectral data or spectral data that had been corrected (CO) for different sample sizes. In all cases (Set A and Set B models), the PLSDA-MC models outperformed corresponding PLSDA-CO models for the raw

1400-2200nm spectral band. This suggests that a correction for differing sample sizes may not be necessary in the application under study.

Additional models constructed with a broader raw spectral band (460-2360nm) are reported in Appendix F. There is a slight increase in classification accuracy for all LDA-PCA models for both Sets A and B. The most notable increase was observed in the Set B 2Cl_F_1F_2 PLSDA-MC model, which classified 90% of Class 0 and all of Class 1 test samples, with an overall accuracy of 92.3%. The corresponding 1400-2200nm model identified only 33.3% of Class 1.

Overall, poorer performance by the Set B models can likely be attributed to inadequate grinding of the superficial cortical surface. Sectioning the long axis of the bone diaphysis reveals clean, unblemished cortex with minimal environmental artifacts, thus making the Set A spectra optimal. The purpose of grinding down the superficial surface was to expose a comparably clean, underlying cortical layer. Grinding removed much of the environmental debris (soil staining, weathered cortex, etc.); however, some adherent material likely remained. This may have lead to a potential increase in scatter-related noise in the raw spectra. This is particularly true of the older, more altered bone, and is likely a contributing factor in the poorer observed classification accuracy of older Set B samples. An improvement of the method might include identifying an optimal amount of bone to grind off prior to spectral sampling

A final review of the 2Cl_FN raw models revealed that the SVM-PCA models typically underperformed all others, and PLSDA appeared to classify adequately without a correction for sample size. LDA-PCA was, in general, the most consistent method when applied to these data.

It was also observed that Class 1 accuracies were often lower than desired, even when using Set A data. Correct class assignments generally ranged from approximately 79-84%. In

contrast, the forensic test sample classification accuracy is consistently high (roughly 95% or greater). A subsequent review of the misclassified Class 1 samples reveals that the wrongly assigned samples were likely "historic" burials, that is, remains that are outside of the "forensic" time-frame but not within the more "archaeological" group of samples. It should be noted that, from a medicolegal perspective, the misclassification of a non-forensic case is of lower importance. However, improper classification of a forensic case is clearly undesirable, as this would represent the exclusion of an actual missing person from further investigation.

5.4.3 Limitations of the 3Cl_F_1F_2N and Cl_F_1F_2 Models

Although the primary goal of this study was to develop classification models that could assign unknown remains into "Forensic" and "Non-forensic" classes, a secondary goal was to determine whether "forensic" cases could be accurately placed into different subclasses (e.g. F_1, F_2). As noted in section 6.3.3, sample sizes for the 3Cl_F_1F_2N and Cl_F_1F_2 models are small, especially for Class 1 ("older forensic cases," PMI 5 – 71 years). At issue was the limited number of available "older" forensic cases with an established PMI values. As observed in the results of these 3Cl_F_1F_2N and Cl_F_1F_2 models, classification accuracy for Class 1 test samples was relatively low; although none of these samples was misclassified as non-forensic (i.e. Class 2), many were misclassified as Class 0 ("newer forensic cases," PMI "days" – 4.9 years). This suggests that Class 1 is either poorly defined, or the samples used for model building and/or external testing are poor candidates; both are likely contributors. Some of the "older" samples used included exhumed paupers burials, which consistently classified as "newer," suggesting that these remains might be better preserved than an extended PMI would suggest. Likewise,

the Class 1 PMI distribution of 5 to 71 years is overly broad; it is likely the addition of more samples with established, extended PMI values will reveal that this broad class can be further subdivided.

5.5 Data Transformations

5.5.1 Standard Normal Variate (SNV) Transformed Data Models

Selected bands of raw data were subjected to three transformations, including SNV as well as first and second SG derivations. The resultant model classification accuracy tables for each of these models are presented in Appendix F. SNV transformed models were constructed using the following bands: 460-2360nm, 1400-2200nm, and 1425-1750nm. On the whole, 2Cl_FN models and associated test set classification accuracies were the same as those reported for the Set A and Set B raw data 1400-2200 models reported in the previous chapter. The exception to this was the Set B PLSDA-MC models using the SNV 1400-2200nm and 1425-1750nm bands, both of which underperformed the raw model for non-forensic test samples (Class 1 accuracy = 30.0%, Overall accuracy 87.9%).

Discriminating power also decreased for the Set B LDA-PCA and PLSDA-CO three-class models using both SNV bands reported above. However, an increase in classification accuracy over the raw data models was observed in some cases for all three SNV model types. In many cases, the same increase was observed when using either SNV 1400-2200nm or 1425-1750nm, although the latter band provided additional models that outperformed the former. The better performing SNV 1425-1750nm models are reported below.

The 2Cl_FN Set B SVM-PCA model yielded an overall accuracy 90.9%, with 96.4% of the forensic test samples being properly classified; non-forensic test sample classification was a relatively low 60.0%, although this is considerably higher than that observed with the comparable raw data model (20% accuracy, Overall accuracy 87.9%).

Improvements were also noted in the 2Cl_F_1F_2 models. The Set A SVM-PCA model properly classified all newer forensic samples but only 37.5% of the older (Class 1) forensic material (Overall 82.1%). Similar to the above, although Class 1 accuracy remains low, it is higher than the corresponding raw data model (Class 1 = 25%, Overall 60.7%). All Set B models improved, with the greatest improvement noted on the PLSDA-MC model. An overall classification rate 92.3% was achieved, with 90% of the newer forensic test samples being properly classified; all older forensic test samples were assigned to the correct class.

Improved accuracies were also noted with three-class SNV transformed Set B data for models using SVM-PCA and PLSDA-CO. The overall accuracy for the SVM-PCA model was 81.3% (Class 0 = 100%, Class 1 = 33.3%, Class 2 = 66.7%), while the PLSDA model correctly classified 87.5% of the Set B test samples (accuracies of 100%, 33.3%, and 100% by class, respectively). In both cases, Class 1 (older forensic samples) was poorly classified, as seen in the reported raw data models.

Full band models were attempted as well (SNV 460-2360nm). Improvements in overall and class accuracies were noted in models that correspond with the better SNV 1425-1750nm performers. However, increased accuracies achieved using full bands were not as pronounced as those models built from narrower band.

5.5.2 Savitzky-Golay (SG) Transformed Data Models

SG first and second derivatives were performed on selected raw spectral data bands. The following derivatized bands were used for model construction:

- SG1: 1400-2200nm; 1120-1185nm and 1530-1720nm (SG1 “specific bands”)
- SG2: 1400-2200nm; 1235-1330nm and 1535-1585nm (e.g. SG2 “specific bands”)

In all model types (2Cl_FN, 2Cl_F_1F_2, and 3Cl_F_1F_2N models), both Set A and Set B SVM-PCA models built from SG1 and SG2 derivatized data were the worst performing models. SVM-PCA models were only successful at classifying Class 0 samples; remaining classes consistently classified at 0%, with the same outcome noted on repeated attempts. These results suggest that applying SVM-PCA to principal components derived from SG transformed data is not appropriate.

In contrast, the other 2Cl_FN models (LDA-PCA, PLSDA-MC, and PLSDA-CO) constructed with SG1 and SG2 derivatized data were of comparable accuracy with the reported raw 1400-2200 data models. These models also performed the same on both the Set A and Set B test sets. The three-class SG1 and SG2 derivatized LDA-PCA models showed decreased discriminating power for both Sets A and B test sets, similar to what was observed with the SNV transformed models. As with the SNV transformed 3Cl_F_1F_2N models, the primary driver for decreased classification accuracy was poor assignment of Class 1 test samples (older forensic material) samples.

Although most models built with derivatized data yielded classification accuracy results similar to or worse than models constructed with raw data, some exceptions were noted with SG1 and SG2 1400-2200nm attempts. The Set B 2Cl_F_1F_2 models (LDA-PCA and both PLDSA

methods) each classified 90% of Class 0 and 100% of Class 1 samples (overall 92.3%). The three-class Set B PLSDA-CO was an improvement as well, classifying 90% of the newer forensic test samples and all of the older forensic material; non-forensic test samples (Class 2) classified at only 67% accuracy (Overall = 87.5%). While Class 2 accuracy is still the same as that seen in the raw model, accuracy for older forensic cases in that same model was 0.0%. Note that corresponding SG1 and SG2 "specific band" models also showed increases in classification accuracy, although the improvements were not as pronounced as the above models built with the broader SG1 or SG2 1400-2200nm band.

5.5.3 Data Transformation: Interpretations

The present study suggests that spectral data transformations may be of some use in limited situations. When SNV transformation is applied to the raw spectral data, it appears that transforming and building classification models with a narrow spectral band (e.g. 1425-1750nm) is more beneficial. In contrast, SG1 and SG2 derivatives should be applied to a somewhat a broader spectral band (1400-2200nm) to improve model accuracy.

One of the more intriguing outcomes of the transformed data models is that in some cases, these models appeared to be better at discriminating between samples that were within the forensically significant time-frame; that is, differentiating between newer and older forensic samples. This was especially the case for transformed Set B spectral data. Some of the SNV- and SG-transformed data models showed notable increases in classification accuracies with respect to the corresponding raw data classification models.

It was previously suggested that the relatively poor performances of the raw 1400-2200nm Set B models may have been the result of inadequate surface preparation (grinding) and subsequent light scattering resulting from surface particulates. As noted by Rinnan and colleagues (2009), transformations such as SNV and SG can correct for scattering. Thus, it is possible that the application of these transformations to the raw Set B spectra may be removing scattering-related noise from the spectral signal, leading to increased classification accuracy in the final model. In order to determine if this is the case, additional work must be done in the future with a larger number of "forensic" samples that have a longer PMI.

5.6 Protein Oxidation Tests: Interpretations and Limitations

A limited number of tests for the potential of protein (collagen) oxidation were conducted in February 2017 upon completion of the broader study. This was done in order to determine if the collagen in archived bone samples oxidizes in the postmortem interval. Oxidative changes would likely lead to changes in associated spectra collected from the bone surfaces, which would thus need to be accounted for if such samples are reused in future research.

A review of the inventory of cortical samples used in the 2013 preliminary study revealed 16 samples that were reused for the current project. During a return trip to the ASD-PANalytical laboratories in Boulder, Colorado, new average spectra were collected from these 16 samples (i.e. resampling) using the same model spectrometer as used in the preliminary work. Following SNV transformation, comparisons were made between the "Time 1" spectra (2013) and "Time 2" spectra, with minor changes spectral changes noted for some samples.

This suggests that little to oxidative changes may be occurring, at least over a relatively short term of approximately 4 years; however, the potential for oxidation and resultant spectral differences at longer intervals (e.g. 10 years, 20 years) cannot be excluded. Due to the currently limited availability of the necessary analytical equipment, as well as time limitations with respect to completing the current research, these are questions that must be answered with future study.

In addition to the above 16 resampled cortices, an alternate set of cortices were examined for oxidative changes. Six adjacent samples of femoral diaphysis cortex were removed from one common femur and were treated with varying concentrations of diluted hydrogen peroxide (H_2O_2), ddH_2O (control), or air (control), over the defined time period of 10 days. This was done in order to determine if H_2O_2, a strong oxidizer, could lead to discernible spectral differences between these 6 samples. As with the resampled set, only minor changes were noted in SNV transformed spectra. It should be noted that, due to time limitations, the samples were only kept in solution for 10 day period, and it is not currently known what the resultant spectra would have looked like if the treatment was extended and the H_2O_2 solutions were periodically replaced. Additionally, H_2O_2 concentrations in the tubes were not measured over the 10 day course, so diminished H_2O_2 efficacy cannot be ruled out. Also note that alternative oxidizers of proteins (e.g. ozone, superoxide, and oxygen gas) were not tested in this study, and thus the effects or lack thereof of these oxidizers on bone and its collagen phase cannot be determined at this time.

CHAPTER 6

CONCLUSIONS

6.1 Project Impact

There are "as many as 100,000 active missing persons cases in the United States" with "tens of thousands" more disappearing yearly (Ritter 2007, p. 2). Recovered remains must be examined for PMI to maximize the possibility of association with a missing person. Coupled with remains from both historical conflicts (e.g. World War II, Korea, Vietnam, Laos, Peru, Guatemala, El Salvador, Vukovar, Rwanda, Cyprus) and contemporary warfare (e.g. Libya, Syria) that fall within a "time interval of interest," there are potentially millions of cases pending worldwide in which an accurate PMI assessment could ultimately help lead to identification (Budowle, personal communication).

The ability to rapidly assess the PMI of skeletal remains using a reliable and validated technique is of the utmost importance. Because NIR spectroscopy requires minimal sample preparation, it provides an avenue for rapid data collection and analysis. Additionally, because the spectral data produced can be incorporated into multivariate statistical models that can be validated, the technique is clearly appealing from an admissibility standpoint.

This project applied NIR spectroscopy to address the question of PMI for skeletonized remains. Multiple classification algorithms were used to construct models that could assign skeletal remains to an appropriate PMI class. Of the algorithms tested, LDA-PCA, followed by PLSDA on mean-centered data, appears to be the best overall option. NIR spectra collected from samples with known or estimated PMI were used for model construction; the models were subsequently validated on separate "unknown" test samples. Resulting models appear to

be well-suited for differentiating a large sample set of forensic and non-forensic material. Working with a much smaller set, there is some indication that samples may be discriminated spectrally within the "forensic time frame," i.e., remains with a PMI of less than 50 to 70 years, as well. To further assess this, additional work must be done with a larger sample of older forensic material.

Two surfaces were ultimately explored for model building in this study, the longitudinally cut cortical surface of long bone diaphyses and the ground superficial diaphyseal surfaces. In most cases, classification models involving spectra collected from the longitudinal cut surfaces performed better than superficial surface counterparts, although some exceptions were noted, especially when using transformed data. While it may eventually be determined that sectioned cortex produces the most dependable models, additional research must be conducted using treated superficial bone. Relatively speaking, cutting and removing a sample of cortex is more destructive and is thus likely unappealing to institutions that house osteological collections. If equally accurate models can be constructed from spectra that are collected from cortical surfaces that are only lightly ground down, this increases potential access to a much greater number of specimens with known PMI. Ideal samples would include forensic cases with known PMI ranging between 10 and 70 years, remains originating from modern and historic armed conflicts ("timed events"), as well as historic remains with documented PMI.

Ultimately, the successful application of this technology to the estimation of PMI would produce substantial future cost savings. Accurate assessments of PMI would allow law enforcement investigators and triers of fact to include or exclude suspects in specific instances and thus redirect investigations, reducing costs in both manpower and laboratory analyses.

Agencies can also triage forensically significant material, thereby focusing greater resources on more contemporary cases for which reference samples may be available.

Because many unidentified skeletal remains cases have profiles uploaded to the FBI Combined DNA Index System (CODIS), significant savings can also be realized in relevant molecular laboratories as well. An accurate determination of the PMI offers a significant savings for CODIS analytical work in multiple ways. First, expensive laboratory manipulations on non-significant material can be prevented. For example, the LFA was able to exclude 190 non-significant cases between 2006 and 2016; at a minimum cost of $2,000.00 per case, these exclusions represented a savings of at least $380,000.00 over a decade for the CHI and the taxpayers who fund it (Larose, personal communication). Second, additional cost savings are realized when CODIS analysts can avoid filtering and excluding meaningless molecular associations between non-significant remains and missing person profiles. Lastly, PMI can be used as metadata to filter through spurious potential matches.

6.2 Future Research

Classification modeling should be continued and refined, dealing with both sectioned and ground superficial cortex. To do this, a much larger collection of cortices will be required spanning multiple PMI groupings, including older forensic samples, historic samples with known PMI, and additional archaeological samples with known PMI. Ideal sources for such samples would include remains recovered from “timed events,” such as the numerous conflicts listed above. It is suspected that as these sample sizes are increased and PMI classes are more narrowly defined, group patterns will be discernible using techniques such as PCA.

A controlled examination of environmental and deposition effects on bone NIR spectra should be performed as well. Multiple scenarios could be set up in a controlled laboratory environment, including exposure to artificial sunlight, burial of remains in different types of soil and at different depths, submersion in water, treatment with acid, or exposure to heat. Spectra could be collected at defined time intervals and compared with the spectra of untreated control samples to see how they vary from each other.

Another area of interest is the effect of intraindividual variation on bone absorption spectra. A series of skeletal remains cases that are relatively complete can be identified for the study. Bone samples can be collected from multiple sites of antemere elements from the same individual (e.g. right and left humeri, femora, and tibiae). Additionally, multiple locations on the same bone can be sampled and compared (e.g. proximal, middle, and distal aspects of the anterior and posterior femoral diaphysis). Comparisons should be made between trabecular and cortical bone as well.

Lastly, predictive modeling for bone composition should be explored. The current study has demonstrated that NIR spectroscopy is a viable technique for developing classification models from spectra collected from cortex; predicting the composition of a sample of bone is a logical next step. To do this, a large set of cortical specimens with known background (demographics, deposition, location, PMI, etc.) should be gathered and randomly sampled into training and test sets. Spectra would be collected from all samples. Subsequently, all samples would be processed and analyzed to determine actual composition, including water, lipid, collagen, non-collagenous proteins, and DNA content. The training samples can then be used to construct PLSR models, where the X-variables are composed of spectra, and the Y-variables

consist of the composition data. The resultant models can then be used to predict test set composition. Model validations can then be performed by comparing the predicted values with the known composition data.

An accurate predictive model would be especially useful for predicting DNA content in bone prior to the processing UHR samples that are to be uploaded into the FBI Combined DNA Index System (CODIS). Samples that are predicted to yield no recoverable DNA can be excluded from further processing, preventing the waste of expensive reagents. Likewise, analysts can better plan the required extraction technique for yielding partial or complete sets of loci from samples in which there are detectable amounts of DNA present. Both scenarios would provide significant cost savings for CODIS analytical work.

APPENDIX A

MATERIAL CHARACTERISTICS AND COMPOSITION OF BONE

Bone, or osseous tissue, is a mineralized, dense connective tissue. It is a biphasic material, composed of a large mineral phase (~65%) of calcium hydroxyapatite ($Ca_5(PO4)_3OH$) and an organic phase (~35%). The organic phase is primarily type I collagen (85-90%), a major structural protein, with the remaining component (5-10%) comprised largely of non-collagenous proteins (e.g. osteocalcin, osteopontin, proteoglycans). Bone rigidity is due to the mineral phase, while flexibility results from the large amount of collagen (Shipman et al. 1985, Hall 2005).

Mature bone is organized as either trabecular (spongy) bone or more densely organized compact (cortical) bone (Figure A.1). The basic unit of cortex, the osteon, is generally a cylindrically shaped structure of approximately 200μm diameter. This osteon is constructed of concentric lamellae, each of which is approximately 5μm thick and composed of mineralized collagen fibrils (Shipman et al. 1985, Hall 2005).

Collagen fibrils are built from heterotrimeric tropocollagen proteins, each approximately 300nm long, 1.5nm wide, and 290k Da in mass (Figure A.2). The right-handed triple helix of tropocollagen is composed of three left-handed helical chains (2-α1, 1-α2). Each of the α-chains is built from a repeated structural motif of $(Glycine\text{-}X\text{-}Y)_n$, where n is generally ≥ 300 residues, "X" is often a proline (PRO) followed by "any" amino acid, or "X" can be any amino acid followed by hydroxyproline (HYP) (Fratzl et al. 2004, Fratzl and Weinkamer 2007, Hulmes 2008).

The tropocollagen triple helix is organized so that the glycine (GLY) residues are directed inward toward the central axis: because the R-group of GLY is a single hydrogen atom, steric interference is minimized and the tropocollagen is more stable. All residues in the X and Y positions, largely PRO and HYP, as well as lysine (LYS), hydroxylysine (HYL), and others, are

directed outward. Residue abundance and percentage values are presented by α1 and α2 chains in Table A. 1. Hydrogen bonding between α-chains and cross-links between HYL and LYS stabilize the triple helix (Fratzl et al. 2004, Fratzl and Weinkamer 2007, Hulmes 2008).

Collagen fibrils form when tropocollagens align in a parallel but non-registered, staggered orientation (Figure A.2). Staggering of tropocollagens leads to 32nm areas of overlap and gaps that are 35nm wide, for a total collagen periodicity of 67nm. The fibrils are then stabilized by additional hydrogen bonding and cross link formation. Mineralization occurs within the gaps via approximately 3nm thick plates of mineral that likely form by nucleation (Shipman et al. 1985, Fratzl et al. 2004, Fratzl and Weinkamer 2007).

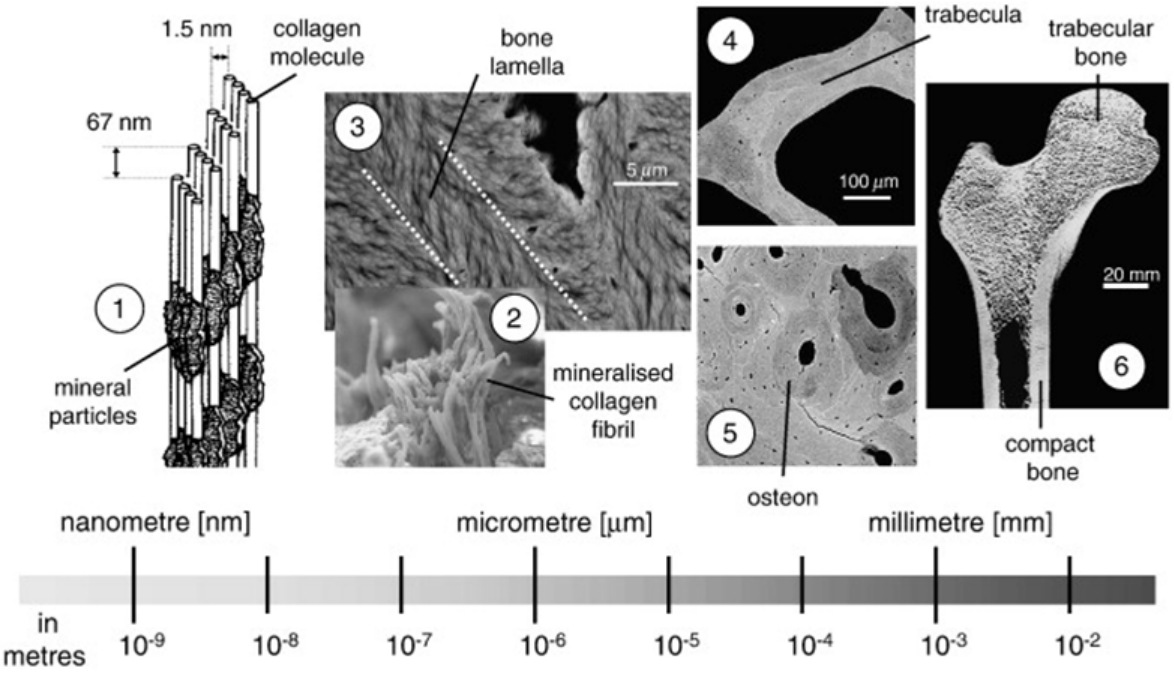


Figure A. 1. Hierarchical structure of bone. Reprinted from Science and Engineering C volume 31, R. Weinkamer and P. Fratzl. "Mechanical adaptation of biological materials – examples of bone and wood," 2011, pages 1164-73. With permission from Elsevier. http://dx.doi.org/10.1016/j.msec.2010.12.002.

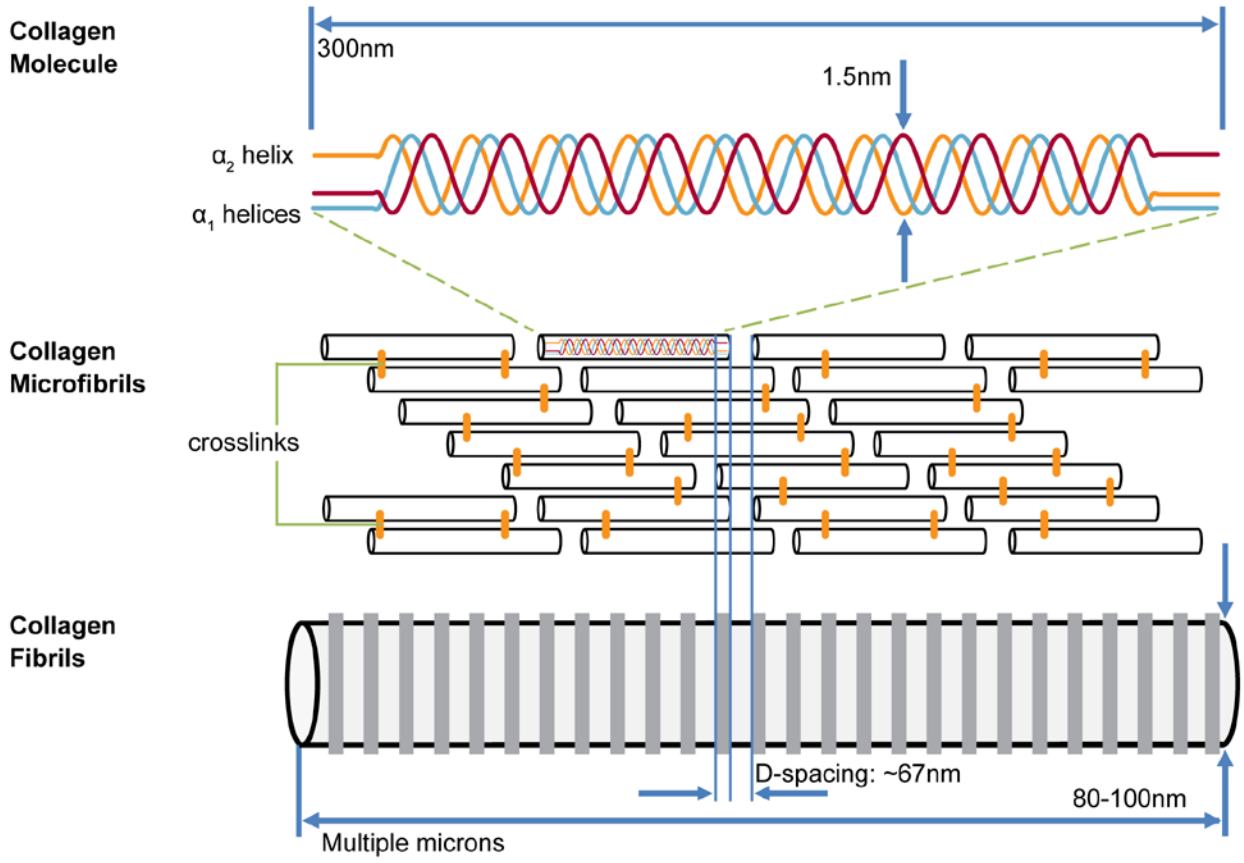


Figure A. 2. Collagen structure and organization. Collagen molecules self-assemble in a quarter-staggered array into microfibrils to form collagen fibrils with characteristic periodic D-spacing. Reprinted from PLOS One volume 11: e0166392, S. Canelón and J. Wallace, "β-Aminopropionitrile-Induced Reduction in Enzymatic Crosslinking Causes *In Vitro* Changes in Collagen Morphology and Molecular Composition," 2016. Reprinted under CC BY. https://doi.org/10.1371/journal.pone.0166392.

Table A. 2. Amino acid (AA) count per 1000 residues and calculated %AA composition in collagen (I) α1 and α2 chains. Reported values have been rounded. *Combined proline and hydroxyproline. **Combined lysine and hydroxylysine. Sequence data were acquired from the UniProt Consortium (Uniprot 2017). Data presented in the table below are modified from output generated at Bioinformatics.org (Stothard 2000).

Amino Acid (AA)	Symbol	α1 (per 1000 residues)	α2 (per 1000 residues)
Glycine	G	328 (33%)	337 (34%)
Proline	P	227* (23%)	196* (20%)
Alanine	A	114 (11%)	104 (10%)
Arginine	R	50 (5%)	53 (5%)
Glutamic acid	E	46 (5%)	43 (4%)
Serine	S	37 (4%)	31 (3%)
Lysine	K	36** (4%)	30** (3%)
Aspartic acid	D	32 (3%)	23 (2%)
Glutamine	Q	28 (3%)	21 (2%)
Leucine	L	20 (2%)	33 (3%)
Valine	V	20 (2%)	39 (4%)
Threonine	T	17 (2%)	18 (2%)
Phenylalanine	F	14 (1%)	12 (1%)
Asparagine	N	10 (1%)	23 (2%)
Isoleucine	I	7 (1%)	17 (2%)
Methionine	M	7 (1%)	5 (<1%)
Tyrosine	Y	4 (<1%)	5 (<1%)
Histidine	H	3 (<1%)	12 (1%)
Cysteine	C	0	0
Tryptophan	W	0	0
AA Class (totals)			
Aliphatic	G,A,V,L,I	488 (49%)	529 (53%)
Aromatic	F,W,Y	18 (2%)	16 (2%)
Sulphur	C,M	7 (1%)	5 (<1%)
Basic	K,R,H	89 (9%)	94 (9%)
Acidic	B,D,E,N,Q,Z	117 (12%)	111 (11%)
Aliphatic hydroxyl	S,T	54 (5%)	49 (5%)

APPENDIX B

QUALITATIVE AND QUANTITATIVE METHODS USED FOR THE ASSESSMENT OF THE POSTMORTEM INTERVAL (PMI)

Presentation of multiple qualitative and quantitative methods (Sections 1 – 10) for the assessment of the postmortem interval (PMI). Included abbreviations: Sect (Section), Microscop (Microscopy), Qual. (Qualitative), Autofl. (Autofluorescence), TGA (thermogravimetric analysis), Radionucl. (radionuclide), MIR (mid-infrared), NIR (near infrared), ARF (University of Tennessee Anthropological Research Facility), Arch (archaeological), Hx (historical).

Sect.	Reference	Technique	Samples and Methodology	Results
.1	Megyesi et al. 2005	PMI Formula	Retrospective study of forensic cases (n=68), known PMI (0-200days). Used written and photodocumentation. All cases surface finds. Quantified decomposition changes by body region (head, trunk, limbs) by developing point system, summed points for Total Body Score (TBS). Calculated Accumulated Degree Days (ADD), summed mean daily temperatures for PMI. Examine ADD-TBS relationship	Well-fit loglinear relationship between PMI and TBS, ADD and TBS. Noted that most cases originated from Illinois/Indiana
	Vass 2011		Presented formulae developed from decomposition studies conducted at ARF. Equations for surface remains (humidity, scored decomposition, temperature) and buried remains (scored decomposition, temperature, soil moisture, adipocere)	Both equations worked well when applied to actual forensic cases originating from environments similar to those near ARF
.2	Berg 1963	Microscop.	Technique review. Author recommendation to examine decalcified bone under polarized light	Diminished optical activity in older remains due to loss of organic content and infiltration of new mineral from soil. Remnants of fat within Haversian systems may indicate PMI≤50yr.
	Shackleford & Wyckoff 1964		Examined thin sections of bone and dentine with electron microscopy. Samples were collected from nonhuman bone dated to the Miocene, Pleistocene, and Pliocene, as well as modern human dentine and modern bovine long bone	Note a collagen periodicity of 640Å in fresh material, with diminishing periodicity as PMI increased. Periodicity around 600Å by 25kya-1mya (Pleistocene) and 500Å by 13-25mya (Miocene). Technique is best suited for ancient bone.
	Race et al. 1968		Examined thin sections of bone collected from 14kya human remains (Nubian desert) with electron microscopy	Observed similar decrease in collagen periodicity from fresh (640Å) to archaeological (600Å).
	Yoshino et al. 1991		Human humeri (n=51) placed in three environments: surface (33), buried (15), at sea (4). Placed in environments for 15yr, sampled accordingly. Examined with SEM (SE mode)	Superficial damage and shallow bacterial tunneling surface sample by 15yr. Extensive superficial and endosteal damage by 5yr on buried samples, with increased tunneling to mid-cortex with longer PMI. Superficial damage on sea samples by 4-5yrs.

Sect.	Reference	Technique	Samples and Methodology	Results

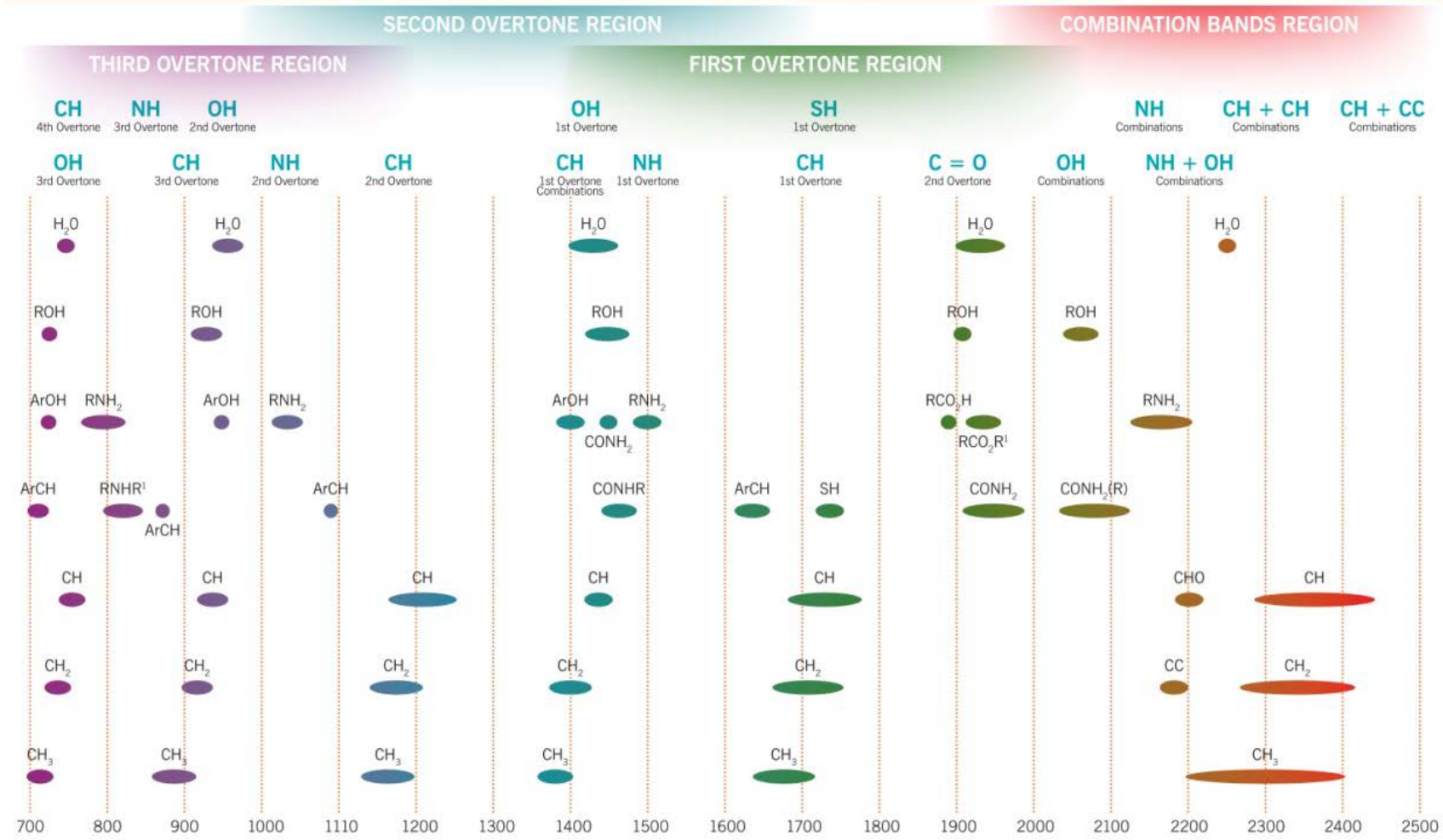
.2	Bell et al. 1996		Mixed human samples (n=11) from forensic cases; long bone fragment, ribs, teeth, PMI 3mo-83yr. Cases originated from terrestrial, intertidal, and lacustrine environs. Interested in how quickly bone damage may occurring PMI. Examiner with SEM (BSE)	General agreement with Yoshini et al. 1991. Noted focal demineralization on longbone fragment (terrestrial, PMI=3mos) and a rib (intertidal, PMI=15mos). Longbone fragment found in scat
.3	Berg 1963	Qual. Methods	Review chapter of previously attempted techniques, includes reactivity of cortex with 3M HCl (test for soil carbonate infiltration), cortical fluorescence under UV light, cortical specific gravity (SPGR), and cortical staining with Nile Blue and Indophenol	Longterm burial (arch) in carbonate-heavy soils yield high reactivity; diminished fluorescence under UV with increased PMI (collagen loss), continuous blue-white fluorescence on newer materials; SPGR ≤ 1.7g/ml; indicates arch; diminished affinity for Nile Blue with greater PMI (decreased organics) increased affinity for Indophenol (relative increase of mineral content as organics decrease and mineral infiltrates)
	Knight & Lauder 1967		Primarily human long bone cortex (n=68), PMI 1-3000yr. Includes cortical staining with Nile Blue and Indophenol, reaction with HCL, nitrogen content in bone powder, amino acid (AA) content in bone powder (paper chromatography), reaction with benzidine (surfaces and powder) UV cortical fluorescence	Staining with Nile Blue/Indophenol and reaction with HCl are poor discriminators; diminished nitrogen content with PMI (>3.5gm% indicates <50yr); diminished AA content with increased PMI (>7 AA's eluted indicates <100yr); positive benzidine reaction indicates ≤ 100-150yr; diminished fluorescence with PMI (continuous blue fluorescence generally indicates recent)
	Bada & Helfman 1975		Examination of racemization (conversion) of amino acids from the typical L-form to the R-form as PMI increases	Notes increased ratio of R:L form at longer PMI. Racemization highly dependent upon recovery environment (e.g. pH, moisture, temperature). Can potentially calibrate with radionuclide data. Technique is only potentially of use when dealing with remains that are 1000s of years old
	Facchini & Pettener 1977		Human femoral shaft cortex (n=71), classes include: 1935-45 (10), 1600-1800 (11), 700-1000 CE (10), 100-300 CE (11), 650-450 BCE (20), 1700-1500 BCE (10). Tests include: application of benzidine (detect blood) to bone powder and whole cortex, cortical fluorescence under UV (366nm), and specific gravity (SPGR)	Diminished benzidine reaction (powder) intensity as PMI increases, with some false positives in arch samples (iron oxide infiltration); diminished fluorescence with increased PMI by 200-350yr; decrease SPGR with PMI, 1.9g/ml indicates <350yr, 1.5g/ml indicates greater than 1000yr

Sect.	Reference	Technique	Samples and Methodology	Results
.4	Introna et al. 1998	Luminol	Powdered human femoral shaft cortex (n=80). 5 PMI Classes (1mo-3yr, 10-15yr, 25-35, 50-60, >80). Expose powder to luminol, record reaction to video under alternate light examine image pixel variation	No reaction >80, Strong reaction ≤ 3yr. Variable for classes between (some false negatives)
	Ramsthaler et al. 2009		Powdered human femoral shaft cortex (n=80). 4 classes (For 1995-2002; For 1965-1972; Hx cem 1878-1902; Arch burials 1st-3rd century CE). Expose powder to luminol, view reaction	Strong reaction recent cases. False positives in some hx and arch remains
	Ramsthaler et al. 2011		Powdered human femoral shaft cortex (n=39), PMI <10yr-1000+ yr. Expose powder to luminol, view reaction	Positive reaction <10yr; false negative reactions 11-100yr. False positives in 101-1000+ yr, likely due to Fe(III)oxide infiltration
.5	Ramsthaler et al. 2011	Autofl.	Human femoral shaft cortex (n=39). Expose sectioned cortex to UV light (254nm, 366nm)	Overall negative correlation between PMI and fluorescence. Newest bone diminished fluorescence due to grease
	Hoke et al. 2013		Long bone cortices (n=213), hx/arch (79 human, 76 nonhuman), PMI 90-1000+ yr, and degreased forensic cases (58) PMI 8-60yr. Developed index system to classify fluorescent color (Blue, Mixed, Yellow/Brown)	Forensic cases generally blue (newer) or mixed. PMI 90-150 cases typically mixed to yellow/brown. PMI 151-1000 increase in amount of Blue, disappears >1000yr
.6	Villanueva et al. 1976	TGA	Human long bone fragments (n=34), primarily modern cases (PMI 0-100yr), limited hx/arch. Placed in furnace and measured mass loss as function of temperature	Best at differentiating newest cases from ancient material
	Raja et al. 2009		Porcine ribs recovered from buried swine, PMI 3mos-7rs. Placed in furnace, measured loss of water (50-200°C), organics (600°C), and CO_2 evolved from apatite	Precipitous mass loss at associated peaks first 2yr of PMI. Gradual decrease in mass loss beyond 2yr
.7	Schwarcz et al. 2010	Citrate	Porcine ribs (7), forensic cases (6, mixed elements), ARF cadaver ribs (3). Porcine ribs buried (6), recovered at 1mo intervals. Assayed bone citrate in samples from procine ribs, forensic samples, ARF samples. Forensic cases recovered from surface	Loglinear decline of citrate vs. time for porcine bone. Developed a regression. Forensic cases lower citrate than porcine (extended PMI). Developed a regression to predict PMI from citrate content, suggested further research with much larger sample with broader PMI
	Kanz et al. 2014		Femoral shaft and temporal squamosa from 20 disinterred cemetery burials (n=40), PMI 27-52yr. All remains unembalmed, in collapsed caskets; half of remains also in body bags. Assayed bone citrate from selected samples, applied regression developed by Schwarcz et al. 2010	General overestimation of PMI on bagged remains. Considerable underestimation of PMI on non-bagged remains. Suggested further research (buried research (buried vs. surface)

Sect.	Reference	Technique	Samples and Methodology	Results
.8	Neis et al. 1999	Radionucl.	Occipital bone (n=9), unburied individuals, deaths 1931-1994. Measured levels of artificial isotope (Sr90) in samples	Sr90 ranged from high in recent samples to non-significant in oldest samples
	Swift et al. 2001		Femoral shaft samples (n=15) from exhumed cemetery remains that had been moved to ossuaries, PMI 15-77yrs. Examined amount of natural isotope Po210 (decay from Pb210), ratio of U234/U238, presence of artificial isotopes	Po210 decreases with PMI, inverse correlation for U234/U238 with PMI. Artificial isotopes (Pu239-240, Cs137) only present in those that died after beginning of nuclear age
	Schrag et al. 2012		Used vertebrae collected from autopsies (1960-2001) to build Sr90 calibration curve ("bomb curve"). Measured amount of Sr90 and Po210 in 30 skeletal cases recovered from 2001-2009. Used Po210 decrease (Swift et al 2001) to target position on Sr90 bomb curve.	Identified 10 cases as non-significant (e.g. hx/arch), sharpened ranges for other. Noted Sr90 in some arch remains (diagenesis) and leaching of Po210 from some forensic cases found on surface (leaching out)
	Speller et al. 2012		Multidisciplinary Case study, juvenile cranium recovered in 1968. Used C14 from tooth enamel to determine ranges for birth and death years. Radionuclide data coupled with mtDNA profile and anthropological profile (filtering)	C14 levels indicated decedent born 1958-1962, death 1963-1968; with bioprofile, search narrowed. ID made via mtDNA comparison, 4.6yr old, missing since 1965 (drowning)
.9	Howe et al. 2012	MIR	Porcine ribs recovered from buried carcasses, PMI 3-23mos. Examined with FTIR spectrometer, used spectra to examine relationship between PMI and organic content, carbonate content, and crystallinity index (CI - crystal ordering)	Observed negative logarithmic relationship between PMI and 1) organic phase, 2) inorganic phase. Weak positive relationship between PMI and CI
	Patonai et al. 2013		Human thoracic vertebrae (n=76). Modern specimens (36) from forensic and clinical cases. Arch. (40) with PMI approximately 1000-6000yr.Examined with FTIR spectrometer, used bands to calculate CI and carbonate/phosphate (C/P) index	Both indices could discriminate between the two groups. Higher CI index values observed in arch; higher C/P index values forensic and medical cases. Separate peak (francolite) indicates arch.
.10	Dogra 2009	NIR	Sectioned porcine ribs allowed to decay in lab setting. Collected daily spectra with NIR spectrometer for 90 days. Attempted to build predictive and classification models	Limited success for predictive model (days by spectrum); slightly better on classification model (month by spectrum).

APPENDIX C

ABSORPTION BANDS FOR RELEVANT FUNCTIONAL GROUPS IN NIR SPECTROSCOPY.



Appendix C. Absorption bands for relevant functional groups in NIR spectroscopy. Wavelength is in (nm). Image was acquired from ASD PANalytical (www.asdi.com).

APPENDIX D

SCATTER-EFFECTS PLOTS GENERATED IN UNSCRAMBLER X 10.4

Figures D. 1 – D. 10. Scatter effects plots for Set A and Set B samples. Each sample is plotted against an average sample. Slope differences are indicative of multiplicative scattering, and offset differences indicate additive effects.

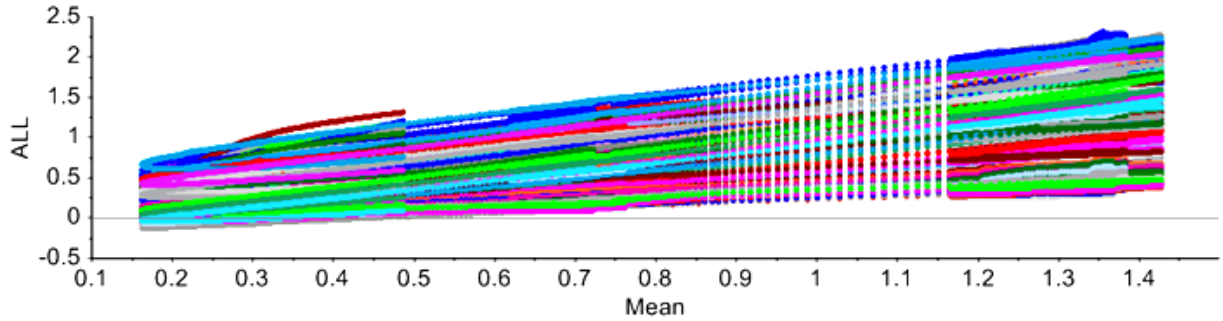


Figure D. 1. Scatter effects plot for the full Set A analysis set (n=314), raw 460-2360nm selected. Slope and offset differences indicate multiplicative and additive scattering.

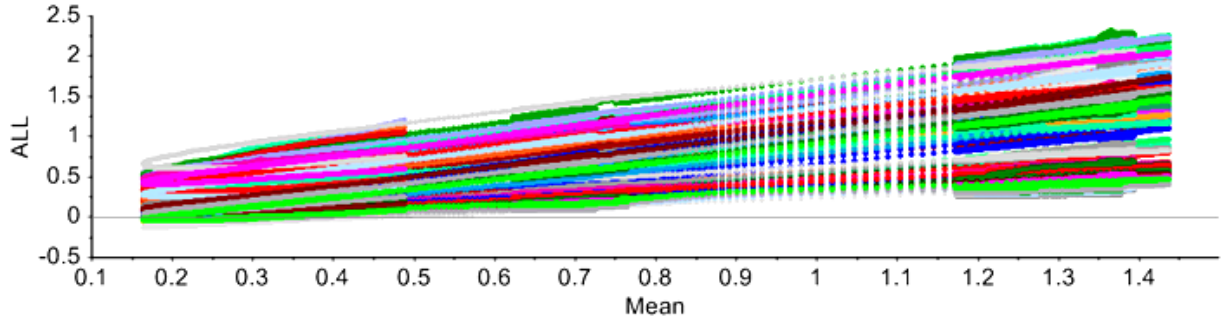


Figure D. 2. Scatter effects plot for the Set A two-class forensic/non-forensic model training set (n=220), raw 460-2360nm selected. Slope and offset differences indicate multiplicative and additive scattering.

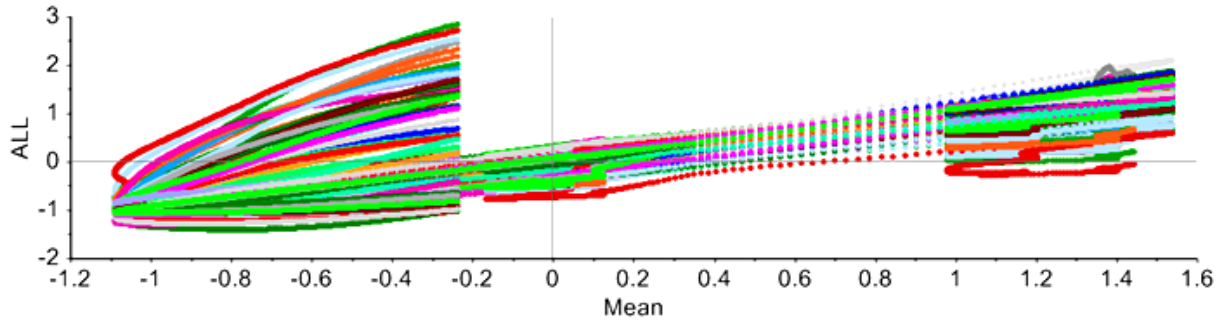


Figure D. 3. Scatter effects plot for the Set A two-class forensic/non-forensic model training set (n=220), SNV-transformed 460-2360nm selected. Slope and offset differences are partially corrected following transformation.

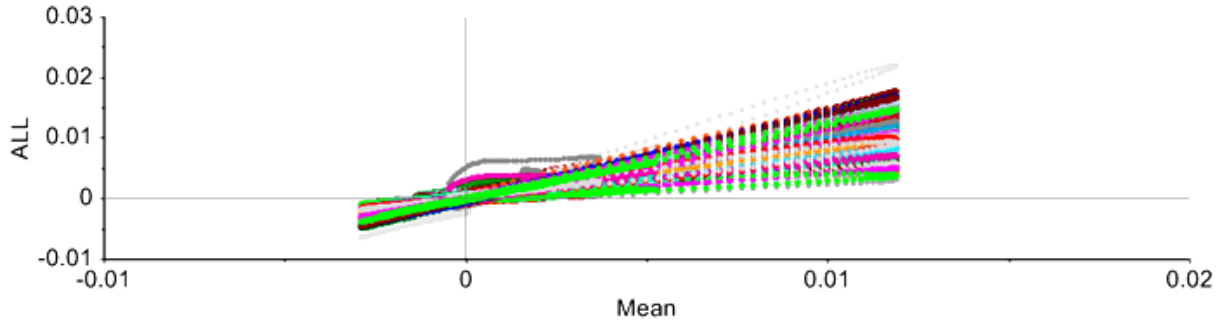


Figure D. 4. Scatter effects plot for the Set A two-class forensic/non-forensic model training set (n=220), SG1-transformed 1400-2200nm selected. Slope and offset differences are partially corrected following transformation.

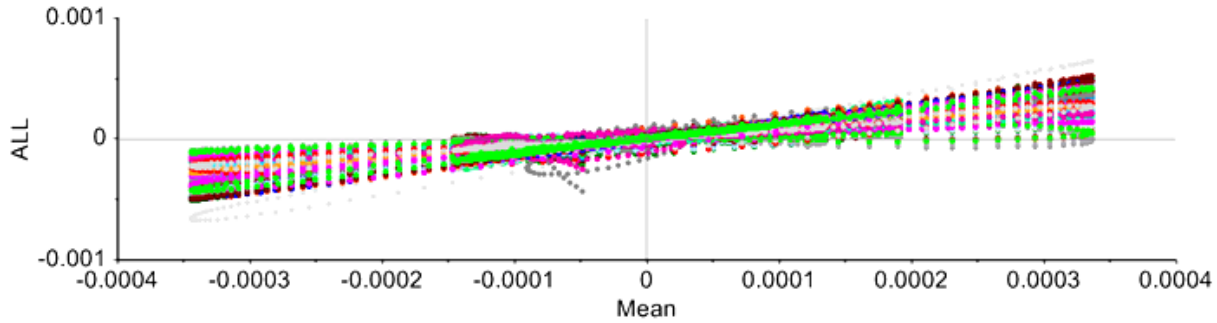


Figure D. 5. Scatter effects plot for the Set A two-class forensic/non-forensic model training set (n=220), SG2-transformed 1400-2200nm selected. Slope and offset differences are partially corrected following transformation.

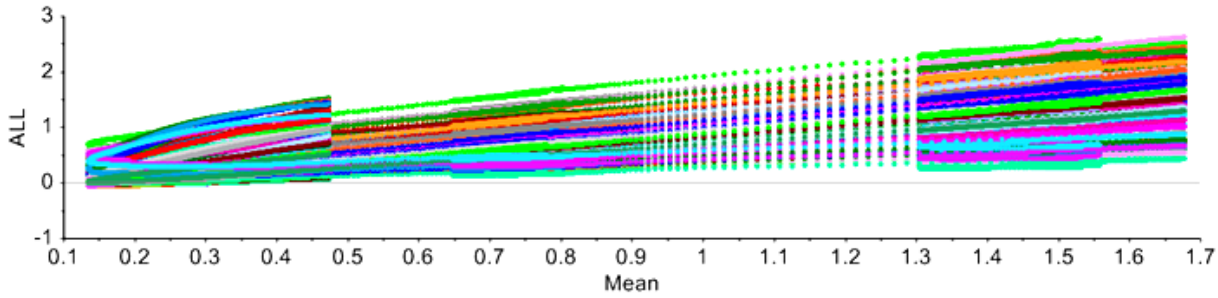


Figure D. 6. Scatter effects plot for the full Set B analysis set (n=223), raw 460-2360nm selected. Slope and offset differences indicate multiplicative and additive scattering.

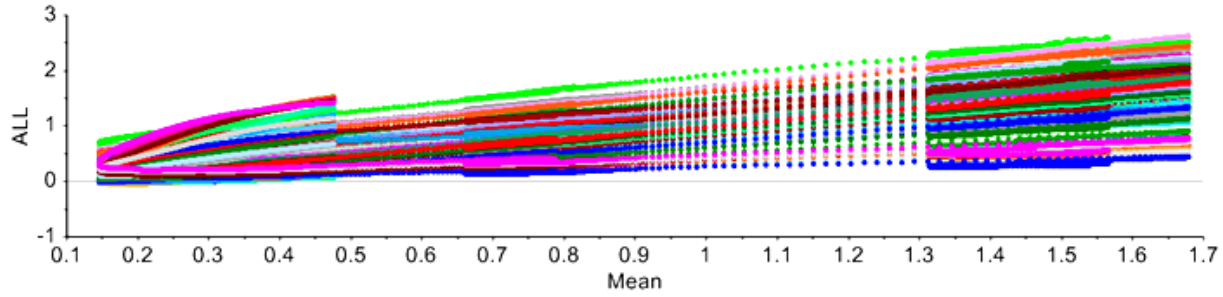


Figure D. 7. Scatter effects plot for the Set B two-class forensic/non-forensic model training set (n=157), raw 460-2360nm selected. Slope and offset differences indicate multiplicative and additive scattering.

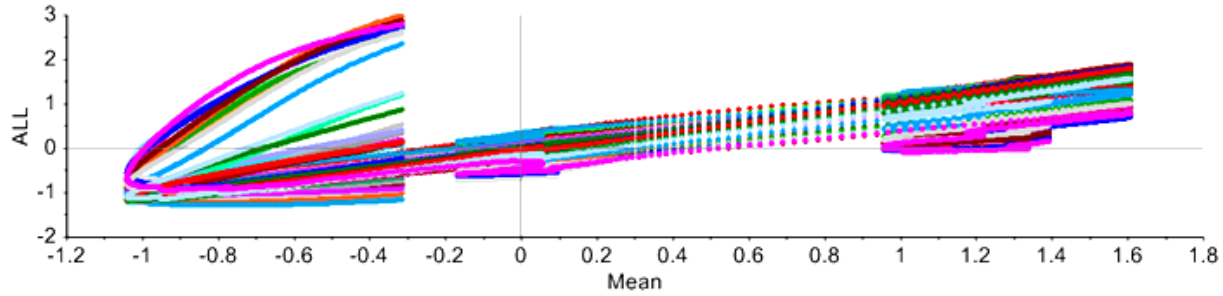


Figure D. 8. Scatter effects plot for the Set B two-class forensic/non-forensic model training set (n=157), SNV-transformed 460-2360nm selected. Slope and offset differences are partially corrected following transformation.

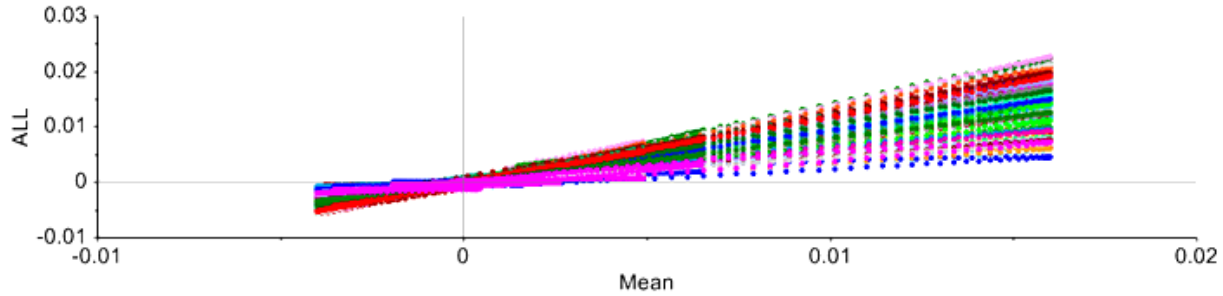


Figure D. 9. Scatter effects plot for the Set B two-class forensic/non-forensic model training set (n=157), SG1-transformed 1400-2200nm selected. Slope and offset differences are partially corrected following transformation.

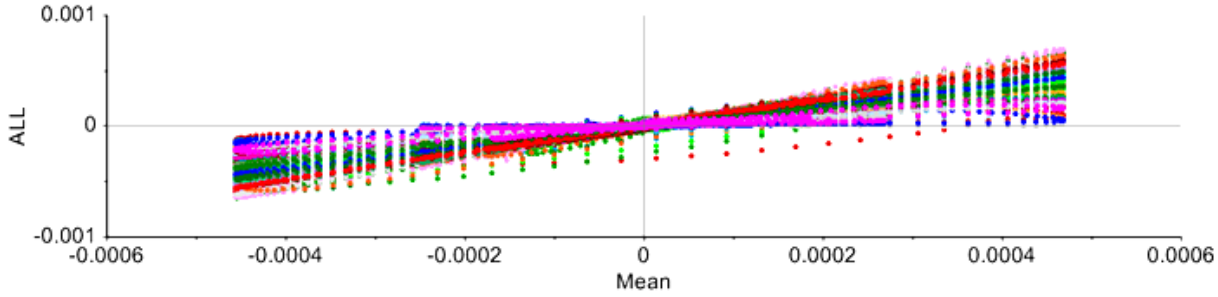


Figure D. 10. Scatter effects plot for the Set B two-class forensic/non-forensic model training set (n=157), SG2-transformed 1400-2200nm selected. Slope and offset differences are partially corrected following transformation.

APPENDIX E

TRAINING AND TEST SAMPLES USED FOR MODEL CONSTRUCTION AND EXTERNAL VALIDATION

Tables E. 1 – E. 6. Selected training and test set samples used for model building and external validation. Class assignments are as follows: 1) Two-class forensic/non-forensic models (Class 0 = forensic, Class 1 = non-forensic); 2) Three-class models (Class 0 = PMI days – 4.9 years, Class 1 = PMI 5 – 71 years, Class 2 = PMI 500 – 900 years); 3) Two-class forensic models (Class 0 = PMI days – 4.9 years, Class 1 = 5 – 71 years).

Table E. 1. Training and test set samples for two-class forensic samples models. Spectra were collected from the longitudinally-cut surface (Group A).

Training Set			
Class 0			**Class 1**
5	93	222	31
6	108	223	42
17	135	226	75
32	147	229	77
38	158	232	125
45	205	234	198
65	206	329	220
66	210		228
67	212		322
81	214		324
82	215		333
Test Set			
4	70	225	23
33	72	313	144
34	91	325	209
40	109	332	217
41	197		309
44	201		311
49	216		323
64	224		334

Table E. 2. Training and test set samples for three -class forensic models. Spectra were collected from the longitudinally-cut surface (Group A).

Training Set				
Class 0			**Class 1**	**Class 2**
5	93	222	31	288
6	108	223	42	301
17	135	226	75	302
32	147	229	77	303
38	158	232	125	
45	205	234	198	
65	206	329	220	
66	210		228	
67	212		322	
81	214		324	
82	215		333	
Test Set				
4	70	225	23	294
33	72	313	144	297
34	91	325	209	300
40	109	332	217	
41	197		309	
44	201		311	
49	216		323	
64	224		334	

Table E. 3. Training and test set samples for two-class forensic/non-forensic models. Spectra were collected from the longitudinally-cut surface (Group A).

Training Set								
Class 0							**Class 1**	
2	38	77	121	153	234	316	134	273
3	39	84	122	154	238	317	162	274
4	40	86	123	155	240	320	165	275
5	42	87	125	156	242	321	168	276
6	45	89	126	158	245	322	169	277
7	46	90	127	194	246	323	171	281
8	47	91	128	195	251	324	175	284
9	49	94	129	196	253	325	176	285
10	50	95	130	197	255	327	177	288
11	51	96	131	201	256	328	178	289
15	52	97	132	202	257	329	179	290
17	53	99	135	205	258	330	180	292
19	55	100	136	211	259	331	181	293
20	56	101	137	213	260	332	182	294
21	57	103	138	214	261	334	183	296
22	58	104	139	215	263	335	184	297
23	59	105	140	217	267	336	186	299
25	60	108	142	221	280	338	189	300
26	61	110	144	222	286	339	190	301
28	68	111	146	224	309	340	233	303
29	69	114	147	226	311	341	248	
32	71	115	148	227	312	342	249	
33	72	116	150	229	313	343	266	
34	73	117	151	230	314	344	270	
36	75	119	152	232	315		271	
Test Set								
1	48	78	106	157	210	252	133	278
13	54	79	107	159	212	254	166	279
14	62	80	109	160	216	262	167	282
18	63	81	112	161	218	264	170	283
24	64	82	118	198	220	287	187	291
27	65	85	120	200	223	310	188	295
31	66	88	124	204	225	318	191	298
37	67	92	141	206	228	319	219	302
41	70	93	143	207	231	333	250	
43	74	98	145	208	235		269	

44	76	102	149	209	237	272

Table E. 4. Training and test set samples for two-class forensic samples models. Spectra were collected from the longitudinally-cut surface (Group B).

Training Set		
Class 0		**Class 1**
5	67	23
6	70	31
12	82	77
17	91	125
33	93	144
41	135	
65	147	
66	158	
Test Set		
4	64	42
32	72	75
38	81	334
40	92	
49	325	

Table E. 5. Training and test set samples for three -class forensic models. Spectra were collected from the longitudinally-cut surface (Group B).

Training Set			
Class 0		**Class 1**	**Class 2**
5	67	23	166
6	70	31	171
12	82	77	185
17	91	125	219
33	93	144	
41	135		
65	147		
66	158		
Test Set			
4	64	42	186
32	72	75	188
38	81	334	281
40	92		
49	325		

Table E. 6. Training and test set samples for two-class forensic/non-forensic models. Spectra were collected from the longitudinally-cut surface (Group B).

Training Set								
Class 0							**Class 1**	
2	33	62	90	117	145	255	133	190
4	35	63	91	118	146	256	134	191
5	37	64	93	119	147	259	162	219
6	38	65	95	123	149	261	163	233
8	39	67	96	125	150	262	168	266
12	45	69	97	126	151	263	170	281
13	47	70	98	127	152	267	173	
14	48	71	101	128	153	286	174	
15	51	73	102	129	154	318	175	
17	52	74	104	130	156	325	177	
18	53	75	105	131	157	334	178	
19	54	76	107	132	158	336	180	
20	55	78	108	135	161	339	182	
21	56	79	110	136	200	340	183	
22	57	80	111	137	246	341	184	
23	58	81	112	140	247	342	185	
25	59	85	113	141	251	343	186	
27	60	86	115	142	253	344	188	
32	61	87	116	143	254		189	
Test Set								
1	28	44	82	103	138	245	164	181
3	29	46	84	106	139	252	165	187
7	31	49	88	109	144	257	166	
9	36	50	89	114	148	264	167	
10	40	66	92	120	155	265	169	
11	41	68	94	121	159	335	171	
24	42	72	99	122	160	337	176	
26	43	77	100	124	211	338	179	

APPENDIX F

CLASSIFICATION MODELS THAT WERE NOT REPORTED IN CHAPTER 5

Tables F. 1 – F. 8. Remaining models not reported in the Results Section, including models constructed on raw and transformed spectral data.

Table F. 1. Tables are separated by model type, including a) two-class models (forensic/non-forensic), b) two-class models (forensic samples only), and c) three-class models. Numbers in "( )" represent the number of components or factors used in a given model. Cells in white represent cross-validated model accuracies as a percentage (LDA-PCA and SVM-PCA) or as RMSECV and R^2 for PLSDA. Areas in grey represent accuracies (%) from model application to the test set.

a. Two-class models (forensic, non-forensic). Class 0 represents forensic samples with known or estimated PMI of less than 71 years; Class 1 represents non-forensic specimens (historic and archaeological origins).

Set	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
A	Model: 95.8 (2)	Model: 96.8 (2)	RMSE = 0.2259 (2)	RMSE = 0.2434 (2)	Overall = 61.7
			R^2 = 0.6873	R^2 = 0.9258	Class 0 = 57.3
	(2) Overall = 95.7	(2) Overall = 90.4	(2) Overall = 95.7	(2) Overall = 92.5	Class 1 = 79.0
	Class 0 = 98.7	Class 0 = 98.6	Class 0 = 98.6	Class 0 = 97.3	Both = 35.1
	Class 1 = 84.2	Class 1 = 57.9	Class 1 = 57.9	Class 1 = 73.7	Neither = 2.1
B	Model: 95.5 (3)	Model: 96.8 (2)	RMSE = 0.1948 (3)	RMSE = 0.2092 (3)	Overall = 36.4
			R^2 = 0.7165	R^2 = 0.9479	Class 0 = 33.9
	(2) Overall = 90.9	(2) Overall = 83.3	(3) Overall = 86.4	(3) Overall = 92.3	Class 1 = 50.0
	Class 0 = 96.4	Class 0 = 96.4	Class 0 = 96.4	Class 0 = 98.2	Both = 53.0
	Class 1 = 60.0	Class 1 = 10.0	Class 1 = 30.0	Class 1 = 50.0	Neither = 4.6

b. Two-class models with no archaeological samples. Class 0 represents forensic samples with known PMI of days to 4.9 years; Class 1 samples have known PMI of 5-71 years.

Set	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
A	Model: 90.0 (2)	Model: 87.5 (2)	RMSE = 0.3484 (2)	RMSE = 0.3415 (3)	Overall = 50.0
			R^2 = 0.3912	R^2 = 0.8164	Class 0 = 65.0
	(2) Overall = 85.7	(2) Overall = 67.9	(1) Overall = 75.0	(3) Overall = 84.1	Class 1 = 12.5
	Class 0 = 90.0	Class 0 = 95.0	Class 0 = 80.0	Class 0 = 95.0	Both = 28.6
	Class 1 = 75.0	Class 1 = 0.0	Class 1 = 62.5	Class 1 = 50.0	Neither = 14.3
B	Model: 90.0 (2)	Model: 95.0 (3)	RMSE = 0.2943 (3)	RMSE = 0.3481 (1)	Overall = 38.5
			R^2 = 0.5382	R^2 = 0.8485	Class 0 = 40.0
	(2) Overall = 84.6	(2) Overall = 69.2	(3) Overall = 92.3	(1) Overall = 76.9	Class 1 = 33.3
	Class 0 = 90.0	Class 0 = 90.0	Class 0 = 90.0	Class 0 = 100.0	Both = 30.8
	Class 1 = 66.7	Class 1 = 0.0	Class 1 = 100.0	Class 1 = 0.0	Neither = 23.1

c. Three-class models. Classes 0 and 1 represent forensic samples with known PMI (days-4.9 years and 5-71 years, respectively); Class 2 samples are non-forensic specimens (PMI = 500 or 900 years).

Set	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
A	Model: 88.6 (3)	Model: 88.6 (2)	Cl 0: RMSE = 0.3203 (5) R^2 = 0.5435 Cl 1: RMSE = 0.3240 (5) R^2 = 0.4401 Cl 2: RMSE = 0.0999 (5) R^2 = 0.8062	Cl 0: RMSE = 0.3910 (3) R^2 = 0.3885 Cl 1: RMSE = 0.3556 (3) R^2 = 0.8082 Cl 2: RMSE = 0.1041 (3) R^2 = 0.8808	Overall = 54.8 Class 0 = 65.0 Class 1 = 12.5 Class 2 = 100.0 Both = 29.0 Neither = 9.7
	(2) Overall = 87.1 Class 0 = 90.0 Class 1 = 75.0 Class 2 = 100.0	(2) Overall = 67.4 Class 0 = 95.0 Class 1 = 0.0 Class 2 = 66.7	(5) Overall = 83.9 Class 0 = 95.0 Class 1 = 50.0 Class 2 = 100.0	(6) Overall = 80.7 Class 0 = 95.0 Class 1 = 37.5 Class 2 = 100.0	
B	Model: 88.0 (2)	Model: 96.0 (2)	Cl 0: RMSE = 0.3312 (2) R^2 = 0.5241 Cl 1: RMSE = 0.3689 (2) R^2 = 0.1493 Cl 2: RMSE = 0.0729 (2) R^2 = 0.9605	Cl 0: RMSE = 0.3699 (3) R^2 = 0.3160 Cl 1: RMSE = 0.3292 (3) R^2 = 0.8307 Cl 2: RMSE = 0.1124 (3) R^2 = 0.9210	Overall = 43.8 Class 0 = 40.0 Class 1 = 33.3 Class 2 = 66.7 Both = 25.0 Neither = 25.0
	(2) Overall = 87.1 Class 0 = 90.0 Class 1 = 75.0 Class 2 = 100.0	(2) Overall = 87.5 Class 0 = 90.0 Class 1 = 66.7 Class 2 = 100.0	(2) Overall = 81.3 Class 0 = 100.0 Class 1 = 0.0 Class 2 = 100.0	(3) Overall = 75.0 Class 0 = 100.0 Class 1 = 0.0 Class 2 = 66.7	

Table F. 2. Model and test set classification accuracies for SNV-transformed data (1400-2200nm) presented by Study Set (A and B). Tables are separated by model type, including a) two-class models (forensic/non-forensic), b) two-class models (forensic samples only), and c) three-class models. Numbers in "( )" represent the number of components or factors used in a given model. Cells in white represent cross-validated model accuracies as a percentage (LDA-PCA and SVM-PCA) or as RMSECV and R^2 for PLSDA. Areas in grey represent accuracies (%) from model application to the test set.

a. Two-class models (forensic, non-forensic). Class 0 represents forensic samples with known or estimated PMI of less than 71 years; Class 1 represents non-forensic specimens (historic and archaeological origins).

Set	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
A	Model: 94.5 (2)	Model: 95.9 (2)	RMSE = 0.2371 (2)	RMSE = 0.2284 (4)	Overall = 40.4
			R^2 = 0.6557	R^2 = 0.9347	Class 0 = 25.3
	(2) Overall = 95.7	(2) Overall = 95.7	(2) Overall = 96.8	(4) Overall = 94.7	Class 1 = 100
	Class 0 = 98.7	Class 0 = 98.7	Class 0 = 100	Class 0 = 100	Both = 53.2
	Class 1 = 84.2	Class 1 = 84.2	Class 1 = 84.2	Class 1 = 78.9	Neither = 1.1
B	Model: 94.3 (2)	Model: 96.2 (2)	RMSE = 0.2139 (2)	RMSE = 0.2379 (2)	Overall = 56.1
			R^2 = 0.6583	R^2 = 0.6444	Class 0 = 57.1
	(2) Overall = 89.4	(2) Overall = 89.4	(2) Overall = 87.9	(2) Overall = 92.3	Class 1 = 50.0
	Class 0 = 100.0	Class 0 = 94.6	Class 0 = 98.2	Class 0 = 100.0	Both = 39.4
	Class 1 = 30	Class 1 = 50.0	Class 1 = 30.0	Class 1 = 40.0	Neither = 3.0

b. Two-class models with no archaeological samples. Class 0 represents forensic samples with known PMI of days to 4.9 years; Class 1 samples have known PMI of 5-71 years.

Set	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
A	Model: 80.0 (3)	Model: 87.5 (2)	RMSE = 0.3712 (1)	RMSE = 0.3551 (4)	Overall = 32.1
			R^2 = 0.3879	R^2 = 0.8201	Class 0 = 45.0
	(2) Overall = 82.1	(2) Overall = 75.0	(1) Overall = 85.7	(4) Overall = 78.6	Class 1 = 0.0
	Class 0 = 100.0	Class 0 = 90.0	Class 0 = 100.0	Class 0 = 85.0	Both = 35.7
	Class 1 = 37.5	Class 1 = 37.5	Class 1 = 37.5	Class 1 = 62.5	Neither = 28.6
B	Model: 80.0 (3)	Model: 85.0 (5)	RMSE = 0.2951 (4)	RMSE = 0.4002 (1)	Overall = 15.4
			R^2 = 0.5355	R^2 = 0.7998	Class 0 = 10.0
	(2) Overall = 84.6	(2) Overall = 84.6	(4) Overall = 84.6	(1) Overall = 76.9	Class 1 = 33.3
	Class 0 = 90.0	Class 0 = 100.0	Class 0 = 80.0	Class 0 = 100.0	Both = 61.5
	Class 1 = 66.7	Class 1 = 33.3	Class 1 = 100.0	Class 1 = 0.0	Neither = 23.1

c. Three-class models. Classes 0 and 1 represent forensic samples with known PMI (days-4.9 years and 5-71 years, respectively); Class 2 samples are non-forensic specimens (PMI = 500 or 900 years).

Set	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
A	Model: 77.3(2)	Model: 84.1 (2)	Cl 0: RMSE = 0.3632 (2) R^2 = 0.4130 Cl 1: RMSE = 0.3843 (2) R^2 = 0.2122 Cl 2: RMSE = 0.1092 (2) R^2 = 0.8557	Cl 0: RMSE = 0.3272 (6) R^2 = 0.5717 Cl 1: RMSE = 0.3242 (6) R^2 = 0.8405 Cl 2: RMSE = 0.0411 (6) R^2 = 0.9814	Overall = 38.7 Class 0 = 45.0 Class 1 = 0.0 Class 2 = 100.0 Both = 53.2 Neither = 1.1
	(2) Overall = 83.9 Class 0 = 100.0 Class 1 = 37.5 Class 2 = 100.0	(2) Overall = 83.9 Class 0 = 100.0 Class 1 = 37.5 Class 2 = 100.0	(2) Overall = 77.4 Class 0 = 100.0 Class 1 = 12.5 Class 2 = 100.0	(6) Overall = 83.9 Class 0 = 95.0 Class 1 = 50.0 Class 2 = 100.0	
B	Model: 88.0 (2)	Model: 88.0 (2)	Cl 0: RMSE = 0.3362 (3) R^2 = 0.5094 Cl 1: RMSE = 0.3473 (3) R^2 = 0.2464 Cl 2: RMSE = 0.0686 (3) R^2 = 0.9469	Cl 0: RMSE = 0.3949 (2) R^2 = 0.2201 Cl 1: RMSE = 0.3798 (2) R^2 = 0.7746 Cl 2: RMSE = 0.0884 (2) R^2 = 0.9512	Overall = 18.75 Class 0 = 10.0 Class 1 = 0.0 Class 2 = 66.7 Both = 56.3 Neither = 25.0
	(2) Overall = 83.9 Class 0 = 100.0 Class 1 = 37.5 Class 2 = 100.0	(2) Overall = 81.3 Class 0 = 100.0 Class 1 = 33.3 Class 2 = 66.7	(3) Overall = 81.3 Class 0 = 90.0 Class 1 = 66.7 Class 2 = 66.7	(2) Overall = 75.0 Class 0 = 100.0 Class 1 = 0.0 Class 2 = 66.7	

Table F. 3. Model and test set classification accuracies for SNV-transformed data (1425-1750nm) presented by Study Set (A and B). Tables are separated by model type, including a) two-class models (forensic/non-forensic), b) two-class models (forensic samples only), and c) three-class models. Numbers in "( )" represent the number of components or factors used in a given model. Cells in white represent cross-validated model accuracies as a percentage (LDA-PCA and SVM-PCA) or as RMSECV and R^2 for PLSDA. Areas in grey represent accuracies (%) from model application to the test set.

a. Two-class models (forensic, non-forensic). Class 0 represents forensic samples with known or estimated PMI of less than 71 years; Class 1 represents non-forensic specimens (historic and archaeological origins).

Set	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
A	Model: 95.0 (2)	Model: 94.5 (2)	RMSE = 0.2361 (1) R^2 = 0.6586	RMSE = 0.2453 (2) R^2 = 0.9247	Overall = 24.5 Class 0 = 6.7
	(2) Overall = 95.7 Class 0 = 98.7 Class 1 = 84.2	(2) Overall = 95.7 Class 0 = 98.7 Class 1 = 84.2	(1) Overall = 95.8 Class 0 = 98.7 Class 1 = 84.2	(2) Overall = 94.7 Class 0 = 100 Class 1 = 78.9	Class 1 = 94.7 Both = 70.2 Neither = 1.1
B	Model: 94.3 (2)	Model: 95.5 (2)	RMSE = 0.2144 (1) R^2 = 0.6567	RMSE = 0.2129 (3) R^2 = 0.9461	Overall = 22.7 Class 0 = 16.1
	(2) Overall = 90.9 Class 0 = 98.2 Class 1 = 50.0	(2) Overall = 90.9 Class 0 = 96.4 Class 1 = 60.0	(2) Overall = 87.9 Class 0 = 98.2 Class 1 = 30.0	(3) Overall = 92.3 Class 0 = 100.0 Class 1 = 40.0	Class 1 = 60.0 Both = 69.7 Neither = 3.0

b. Two-class models with no archaeological samples. Class 0 represents forensic samples with known PMI of days to 4.9 years; Class 1 samples have known PMI of 5-71 years.

Set	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
A	Model: 80.0 (2)	Model: 82.5 (2)	RMSE = 0.3809 (1) R^2 = 0.2725	RMSE = 0.3295 (5) R^2 = 0.8503	Overall = 35.7 Class 0 = 50.0
	(2) Overall = 82.1 Class 0 = 100.0 Class 1 = 37.5	(2) Overall = 82.1 Class 0 = 100.0 Class 1 = 37.5	(1) Overall = 82.1 Class 0 = 100.0 Class 1 = 37.5	(5) Overall = 82.1 Class 0 = 95.0 Class 1 = 50.0	Class 1 = 0.0 Both = 50.0 Neither = 10.7
B	Model: 85.0 (2)	Model: 85.0 (2)	RMSE = 0.2852 (4) R^2 = 0.5662	RMSE = 0.3292 (3) R^2 = 0.8646	Overall = 7.7 Class 0 = 0.0
	(2) Overall = 84.6 Class 0 = 90.0 Class 1 = 66.7	(2) Overall = 84.6 Class 0 = 100.0 Class 1 = 33.3	(4) Overall = 92.3 Class 0 = 90.0 Class 1 = 100.0	(3) Overall = 76.9 Class 0 = 90.0 Class 1 = 33.3	Class 1 = 33.3 Both = 61.5 Neither = 30.8

c. Three-class models. Classes 0 and 1 represent forensic samples with known PMI (days-4.9 years and 5-71 years, respectively); Class 2 samples are non-forensic specimens (PMI = 500 or 900 years).

Set	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
A	Model: 81.8 (2)	Model: 84.1 (2)	Cl 0: RMSE = 0.3547 (3) R^2 = 0.4401 Cl 1: RMSE = 0.3567 (3) R^2 = 0.3214 Cl 2: RMSE = 0.0601 (3) R^2 = 0.9563	Cl 0: RMSE = 0.3288 (5) R^2 = 0.5677 Cl 1: RMSE = 0.3244 (5) R^2 = 0.8403 Cl 2: RMSE = 0.0487 (5) R^2 = 0.9739	Overall = 41.9 Class 0 = 50.0 Class 1 = 0.0 Class 2 = 100.0 Both = 45.2 Neither = 9.7
	(2) Overall = 83.9 Class 0 = 100.0 Class 1 = 37.5 Class 2 = 100.0	(2) Overall = 83.9 Class 0 = 100.0 Class 1 = 37.5 Class 2 = 100.0	(3) Overall = 83.9 Class 0 = 100.0 Class 1 = 37.5 Class 2 = 100.0	(5) Overall = 80.7 Class 0 = 100.0 Class 1 = 25.0 Class 2 = 100.0	
B	Model: 88.0 (2)	Model: 88.0 (2)	Cl 0: RMSE = 0.3162 (4) R^2 = 0.5659 Cl 1: RMSE = 0.3212 (4) R^2 = 0.3553 Cl 2: RMSE = 0.0611 (4) R^2 = 0.9722	Cl 0: RMSE = 0.3043 (5) R^2 = 0.5369 Cl 1: RMSE = 0.2895 (5) R^2 = 0.8691 Cl 2: RMSE = 0.0559 (5) R^2 = 0.9805	Overall = 18.5 Class 0 = 10.0 Class 1 = 0.0 Class 2 = 66.7 Both = 56.3 Neither = 25.0
	(2) Overall = 83.9 Class 0 = 100.0 Class 1 = 37.5 Class 2 = 100.0	(2) Overall = 81.3 Class 0 = 100.0 Class 1 = 33.3 Class 2 = 66.7	(4) Overall = 87.5 Class 0 = 90.0 Class 1 = 100.0 Class 2 = 66.7	(5) Overall = 87.5 Class 0 = 100.0 Class 1 = 33.3 Class 2 = 100.0	

Table F. 4. Model and test set classification accuracies for SNV-transformed data (460-2360nm) presented by Study Set (A and B). Tables are separated by model type, including a) two-class models (forensic/non-forensic), b) two-class models (forensic samples only), and c) three-class models. Numbers in "( )" represent the number of components or factors used in a given model. Cells in white represent cross-validated model accuracies as a percentage (LDA-PCA and SVM-PCA) or as RMSECV and R^2 for PLSDA. Areas in grey represent accuracies (%) from model application to the test set.

a. Two-class models (forensic, non-forensic). Class 0 represents forensic samples with known or estimated PMI of less than 71 years; Class 1 represents non-forensic specimens (historic and archaeological origins).

Set	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
A	Model: 94.5 (2)	Model: 95.0 (2)	RMSE = 0.2217 (2)	RMSE = 0.2542 (2)	Overall = 26.6
			R^2 = 0.6988	R^2 = 0.9191	Class 0 = 10.7
	(2) Overall = 95.8	(2) Overall = 95.4	(2) Overall = 95.7	(2) Overall = 94.7	Class 1 = 89.5
	Class 0 = 98.7	Class 0 = 98.7	Class 0 = 98.7	Class 0 = 100	Both = 86.7
	Class 1 = 84.2	Class 1 = 84.2	Class 1 = 84.2	Class 1 = 78.9	Neither = 0.0
B	Model: 95.5 (2)	Model: 96.8 (2)	RMSE = 0.1899 (2)	RMSE = 0.2013 (2)	Overall = 42.4
			R^2 = 0.7306	R^2 = 0.7456	Class 0 = 39.3
	(2) Overall = 90.9	(2) Overall = 89.4	(2) Overall = 89.4	(2) Overall = 90.8	Class 1 = 60.0
	Class 0 = 96.4	Class 0 = 94.6	Class 0 = 96.4	Class 0 = 100.0	Both = 45.5
	Class 1 = 60.0	Class 1 = 60.0	Class 1 = 50.0	Class 1 = 30.0	Neither = 3.0

b. Two-class models with no archaeological samples. Class 0 represents forensic samples with known PMI of days to 4.9 years; Class 1 samples have known PMI of 5-71 years.

Set	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
A	Model: 80.0 (4)	Model: 87.5 (2)	RMSE = 0.3712 (1)	RMSE = 0.3705 (2)	Overall = 10.7
			R^2 = 0.3089	R^2 = 0.8107	Class 0 = 10.0
	(2) Overall = 85.7	(2) Overall = 75.0	(1) Overall = 85.7	(2) Overall = 85.7	Class 1 = 12.5
	Class 0 = 100.0	Class 0 = 95.0	Class 0 = 95.0	Class 0 = 95.0	Both = 60.7
	Class 1 = 50.0	Class 1 = 25.0	Class 1 = 62.5	Class 1 = 62.5	Neither = 25.0
B	Model: 85.0 (2)	Model: 85.0 (5)	RMSE = 0.3666 (1)	RMSE = 0.3992 (1)	Overall = 15.4
			R^2 = 0.2833	R^2 = 0.8008	Class 0 = 0.0
	(2) Overall = 76.9	(2) Overall = 76.9	(1) Overall = 84.6	(1) Overall = 76.9	Class 1 = 66.7
	Class 0 = 80.0	Class 0 = 100.0	Class 0 = 90.0	Class 0 = 100.0	Both = 61.5
	Class 1 = 66.7	Class 1 = 0.0	Class 1 = 66.7	Class 1 = 0.0	Neither = 15.4

c. Three-class models. Classes 0 and 1 represent forensic samples with known PMI (days-4.9 years and 5-71 years, respectively); Class 2 samples are non-forensic specimens (PMI = 500 or 900 years).

Set	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
A	Model: 77.3(2)	Model: 93.2 (2)	Cl 0: RMSE = 0.3224 (4) R^2 = 0.5373 Cl 1: RMSE = 0.3329 (4) R^2 = 0.4088 Cl 2: RMSE = 0.0987 (4) R^2 = 0.8825	Cl 0: RMSE = 0.3265 (5) R^2 = 0.5737 Cl 1: RMSE = 0.3277 (5) R^2 = 0.8370 Cl 2: RMSE = 0.0.085 (5) R^2 = 0.9202	Overall = 19.3 Class 0 = 10.0 Class 1 = 12.5 Class 2 = 100.0 Both = 41.9 Neither = 12.9
	(2) Overall = 83.9 Class 0 = 95.0 Class 1 = 50.0 Class 2 = 100.0	(2) Overall = 74.2 Class 0 = 90.0 Class 1 = 37.5 Class 2 = 66.7	(4) Overall = 80.7 Class 0 = 100.0 Class 1 = 25.0 Class 2 = 100.0	(5) Overall = 80.7 Class 0 = 100.0 Class 1 = 25.0 Class 2 = 100.0	
B	Model: 88.0 (2)	Model: 88.0 (2)	Cl 0: RMSE = 0.3431 (2) R^2 = 0.4892 Cl 1: RMSE = 0.3602 (2) R^2 = 0.1889 Cl 2: RMSE = 0.0575 (2) R^2 = 0.9754	Cl 0: RMSE = 0.3920 (2) R^2 = 0.2316 Cl 1: RMSE = 0.3826 (2) R^2 = 0.7713 Cl 2: RMSE = 0.0600 (2) R^2 = 0.9775	Overall = 12.5 Class 0 = 0.0 Class 1 = 0.0 Class 2 = 66.7 Both = 62.5 Neither = 18.8
	(2) Overall = 83.9 Class 0 = 95.0 Class 1 = 50.0 Class 2 = 100.0	(2) Overall = 75 Class 0 = 100.0 Class 1 = 0.0 Class 2 = 66.7	(2) Overall = 87.5 Class 0 = 100.0 Class 1 = 66.7 Class 2 = 66.7	(2) Overall = 75.0 Class 0 = 100.0 Class 1 = 0.0 Class 2 = 66.7	

Table F. 5. Model and test set classification accuracies for Savitzky-Golay first derivative transformed data (1400-2200nm) presented by Study Set (A and B). Tables are separated by model type, including a) two-class models (forensic/non-forensic), b) two-class models (forensic samples only), and c) three-class models. Numbers in "( )" represent the number of components or factors used in a given model. Cells in white represent cross-validated model accuracies as a percentage (LDA-PCA and SVM-PCA) or as RMSECV and R^2 for PLSDA. Areas in grey represent accuracies (%) from model application to the test set.

a. Two-class models (forensic, non-forensic). Class 0 represents forensic samples with known or estimated PMI of less than 71 years; Class 1 represents non-forensic specimens (historic and archaeological origins).

Set	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
A	Model 91.3 (2)	Model: 79.5 (2)	RMSE = 0.2263 (3)	RMSE = 0.2149 (7)	Overall = 48.9
			R^2 = 0.6864	R^2 = 0.9422	Class 0 = 42.7
	(2) Overall = 91.5	(2) Overall = 79.8	(3) Overall = 95.8	(7) Overall = 95.7	Class 1 = 73.7
	Class 0 = 93.3	Class 0 = 100	Class 0 = 100	Class 0 = 98.6	Both = 43.6
	Class 1 = 84.2	Class 1 = 0	Class 1 = 78.9	Class 1 = 84.2	Neither = 6.4
B	Model 94.3 (2)	Model: 84.1 (2)	RMSE = 0.1838 (6)	RMSE = 0.1975 (5)	Overall = 21.2
			R^2 = 0.7478	R^2 = 0.9536	Class 0 = 17.9
	(2) Overall = 90.9	(2) Overall = 84.8	(3) Overall = 89.4	(5) Overall = 92.3	Class 1 = 40.0
	Class 0 = 100.0	Class 0 = 100.0	Class 0 = 98.2	Class 0 = 100.0	Both = 71.2
	Class 1 = 40.0	Class 1 = 0.0	Class 1 = 40.0	Class 1 = 40.0	Neither = 6.1

b. Two-class models with no archaeological samples. Class 0 represents forensic samples with known PMI of days to 4.9 years; Class 1 samples have known PMI of 5-71 years.

Set	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
A	Model 82.5 (3)	Model: 72.5 (2)	RMSE = 0.3655 (3)	RMSE = 0.3745 (3)	Overall = 28.6
			R^2 = 0.3300	R^2 = 0.8066	Class 0 = 35.0
	(2) Overall = 82.1	(2) Overall = 71.4	(3) Overall = 82.1	(3) Overall = 75.0	Class 1 = 12.5
	Class 0 = 90.0	Class 0 = 100.0	Class 0 = 90.0	Class 0 = 100.0	Both = 46.4
	Class 1 = 62.5	Class 1 = 0.0	Class 1 = 62.5	Class 1 = 12.5	Neither = 10.7
B	Model 95.0 (3)	Model: 75.0 (3)	RMSE = 0.3279 (2)	RMSE = 0.2785 (4)	Overall = 7.7
			R^2 = 0.4265	R^2 = 0.9031	Class 0 = 10.0
	(2) Overall = 92.3	(2) Overall = 76.9	(2) Overall = 92.3	(4) Overall = 92.3	Class 1 = 0.0
	Class 0 = 90.0	Class 0 = 100.0	Class 0 = 90.0	Class 0 = 90.0	Both = 61.5
	Class 1 = 100.0	Class 1 = 0.0	Class 1 = 100.0	Class 1 = 100.0	Neither = 30.8

c. Three-class models. Classes 0 and 1 represent forensic samples with known PMI (days-4.9 years and 5-71 years, respectively); Class 2 samples are non-forensic specimens (PMI = 500 or 900 years).

Set	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
A	Model 75.0 (2)	Model: 65.9 (2)	Cl 0: RMSE = 0.3555 (3) R^2 = 0.4376 Cl 1: RMSE = 0.3927 (3) R^2 = 0.1774 Cl 2: RMSE = 0.1260 (3) R^2 = 0.8079	Cl 0: RMSE = 0.3961 (4) R^2 = 0.3723 Cl 1: RMSE = 0.3560 (4) R^2 = 0.8078 Cl 2: RMSE = 0.1523 (4) R^2 = 0.7450	Overall = 41.9 Class 0 = 50.0 Class 1 = 12.5 Class 2 = 66.7 Both = 32.3 Neither = 12.9
	(2) Overall = 74.2 Class 0 = 90.0 Class 1 = 25.0 Class 2 = 100.0	(2) Overall = 64.5 Class 0 = 100.0 Class 1 = 0.0 Class 2 = 0.0	(3) Overall = 71.0 Class 0 = 90.0 Class 1 = 12.5 Class 2 = 100.0	(4) Overall = 64.5 Class 0 = 95.0 Class 1 = 12.5 Class 2 = 0.0	
B	Model 76.0 (2)	Model: 64.0 (2)	Cl 0: RMSE = 0.2872 (4) R^2 = 0.6420 Cl 1: RMSE = 0.2807 (4) R^2 = 0.5075 Cl 2: RMSE = 0.0684 (4) R^2 = 0.9652	Cl 0: RMSE = 0.2840 (4) R^2 = 0.5968 Cl 1: RMSE = 0.2887 (4) R^2 = 0.8698 Cl 2: RMSE = 0.1597 (4) R^2 = 0.8406	Overall = 12.5 Class 0 = 10.0 Class 1 = 0.0 Class 2 = 33.3 Both = 25.0 Neither = 25.0
	(2) Overall = 74.2 Class 0 = 90.0 Class 1 = 25.0 Class 2 = 100.0	(2) Overall = 62.5 Class 0 = 100.0 Class 1 = 0.0 Class 2 = 0.0	(4) Overall = 93.8 Class 0 = 90.0 Class 1 = 100.0 Class 2 = 100.0	(4) Overall = 87.5 Class 0 = 90.0 Class 1 = 100.0 Class 2 = 66.7	

Table F. 6. Model and test set classification accuracies for Savitzky-Golay first derivative transformed data (1120-1185, 1530-1720nm) presented by Study Set (A and B). Tables are separated by model type, including a) two-class models (forensic/non-forensic), b) two-class models (forensic samples only), and c) three-class models. Numbers in "( )" represent the number of components or factors used in a given model. Cells in white represent cross-validated model accuracies as a percentage (LDA-PCA and SVM-PCA) or as RMSECV and R^2 for PLSDA. Areas in grey represent accuracies (%) from model application to the test set.

a. Two-class models (forensic, non-forensic). Class 0 represents forensic samples with known or estimated PMI of less than 71 years; Class 1 represents non-forensic specimens (historic and archaeological origins).

Set	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
A	Model 95.4 (4)	Model: 79.5 (2)	RMSE = 0.2284 (2)	RMSE = 0.2213 (5)	Overall = 41.5
			R^2 = 0.6806	R^2 = 0.9387	Class 0 = 33.3
	(4) Overall = 97.9	(2) Overall = 79.8	(2) Overall = 95.8	(5) Overall = 94.7	Class 1 = 73.7
	Class 0 = 100	Class 0 = 100	Class 0 = 100	Class 0 = 100	Both = 51.1
	Class 1 = 89.5	Class 1 = 0	Class 1 = 78.9	Class 1 = 73.7	Neither = 6.4
B	Model 94.3 (2)	Model: 84.1 (2)	RMSE = 0.2259 (2)	RMSE = 0.2186 (3)	Overall = 48.5
			R^2 = 0.6189	R^2 = 0.9432	Class 0 = 51.8
	(2) Overall = 90.9	(2) Overall = 84.8	(3) Overall = 90.9	(3) Overall = 92.3	Class 1 = 30.0
	Class 0 = 98.2	Class 0 = 100.0	Class 0 = 98.2	Class 0 = 100.0	Both = 47.0
	Class 1 = 50.0	Class 1 = 0.0	Class 1 = 50.0	Class 1 = 40.0	Neither = 3.0

b. Two-class models with no archaeological samples. Class 0 represents forensic samples with known PMI of days to 4.9 years; Class 1 samples have known PMI of 5-71 years.

Set	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
A	Model 77.5 (3)	Model: 72.5 (2)	RMSE = 0.3984 (1)	RMSE = 0.3466 (5)	Overall = 42.9
			R^2 = 0.2038	R^2 = 0.8343	Class 0 = 60.0
	(2) Overall = 78.6	(2) Overall = 71.4	(1) Overall = 64.3	(5) Overall = 82.1	Class 1 = 0.0
	Class 0 = 85.0	Class 0 = 100.0	Class 0 = 90.0	Class 0 = 95.0	Both = 35.7
	Class 1 = 62.5	Class 1 = 0.0	Class 1 = 0.0	Class 1 = 50.0	Neither = 3.6
B	Model 90.0 (3)	Model: 75.0 (4)	RMSE = 0.2935 (2)	RMSE = 0.3618 (1)	Overall = 46.2
			R^2 = 0.5405	R^2 = 0.8361	Class 0 = 50.0
	(2) Overall = 84.6	(2) Overall = 76.9	(2) Overall = 92.3	(1) Overall = 76.9	Class 1 = 33.3
	Class 0 = 80.0	Class 0 = 100.0	Class 0 = 90.0	Class 0 = 100.0	Both = 46.2
	Class 1 = 100.0	Class 1 = 0.0	Class 1 = 100.0	Class 1 = 0.0	Neither = 7.7

c. Three-class models. Classes 0 and 1 represent forensic samples with known PMI (days-4.9 years and 5-71 years, respectively); Class 2 samples are non-forensic specimens (PMI = 500 or 900 years).

Set	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
A	Model 81.8 (2)	Model: 65.9 (2)	Cl 0: RMSE = 0.3805 (2) R^2 = 0.3556 Cl 1: RMSE = 0.4003 (2) R^2 = 0.1455 Cl 2: RMSE = 0.0911 (2) R^2 = 0.8996	Cl 0: RMSE = 0.3315 (6) R^2 = 0.5604 Cl 1: RMSE = 0.3243 (6) R^2 = 0.8405 Cl 2: RMSE = 0.1036 (6) R^2 = 0.8819	Overall = 41.9 Class 0 = 55.0 Class 1 = 0.0 Class 2 = 66.7 Both = 32.3 Neither = 19.4
	(2) Overall = 74.2 Class 0 = 80.0 Class 1 = 50.0 Class 2 = 100.0	(2) Overall = 64.5 Class 0 = 100.0 Class 1 = 0.0 Class 2 = 0.0	(2) Overall = 67.7 Class 0 = 90.0 Class 1 = 0.0 Class 2 = 100.0	(6) Overall = 83.9 Class 0 = 95.0 Class 1 = 50.0 Class 2 = 100.0	
B	Model 92.0 (2)	Model: 64.0 (2)	Cl 0: RMSE = 0.2790 (4) R^2 = 0.6622 Cl 1: RMSE = 0.3172 (4) R^2 = 0.3713 Cl 2: RMSE = 0.0788 (4) R^2 = 0.9538	Cl 0: RMSE = 0.316 (6) R^2 = 0.5176 Cl 1: RMSE = 0.2718 (6) R^2 = 0.8846 Cl 2: RMSE = 0.0946 (6) R^2 = 0.9441	Overall = 37.5 Class 0 = 50.0 Class 1 = 0.0 Class 2 = 33.3 Both = 37.5 Neither = 18.8
	(2) Overall = 74.2 Class 0 = 80.0 Class 1 = 50.0 Class 2 = 100.0	(2) Overall = 62.5 Class 0 = 100.0 Class 1 = 0.0 Class 2 = 0.0	(4) Overall = 81.3 Class 0 = 90.0 Class 1 = 33.3 Class 2 = 100.0	(6) Overall = 75.0 Class 0 = 90.0 Class 1 = 33.3 Class 2 = 66.7	

Table F. 7. Model and test set classification accuracies for Savitzky-Golay second derivative transformed data (1400-2200nm) presented by Study Set (A and B). Tables are separated by model type, including a) two-class models (forensic/non-forensic), b) two-class models (forensic samples only), and c) three-class models. Numbers in "( )" represent the number of components or factors used in a given model. Cells in white represent cross-validated model accuracies as a percentage (LDA-PCA and SVM-PCA) or as RMSECV and R^2 for PLSDA. Areas in grey represent accuracies (%) from model application to the test set.

a. Two-class models (forensic, non-forensic). Class 0 represents forensic samples with known or estimated PMI of less than 71 years; Class 1 represents non-forensic specimens (historic and archaeological origins).

Set	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
A	Model 93.2 (3)	Model: 79.5 (2)	RMSE = 0.2276 (4)	RMSE = 0.2329 (5)	Overall = 35.1
			R^2 = 0.6828	R^2 = 0.9321	Class 0 = 24.0
	(3) Overall = 93.6	(2) Overall = 79.8	(4) Overall = 95.8	(5) Overall = 94.7	Class 1 = 79.0
	Class 0 = 97.3	Class 0 = 100	Class 0 = 100	Class 0 = 100	Both = 60.6
	Class 1 = 78.9	Class 1 = 0	Class 1 = 78.9	Class 1 = 73.7	Neither = 2.1
B	Model 94.3 (3)	Model: 84.1 (2)	RMSE = 0.1992 (5)	RMSE = 0.2303 (4)	Overall = 13.6
			R^2 = 0.7037	R^2 = 0.6669	Class 0 = 8.9
	(2) Overall = 87.9	(2) Overall = 84.8	(3) Overall = 90.9	(4) Overall = 92.3	Class 1 = 40.0
	Class 0 = 96.4	Class 0 = 100.0	Class 0 = 98.2	Class 0 = 100.0	Both = 81.8
	Class 1 = 40.0	Class 1 = 0.0	Class 1 = 50.0	Class 1 = 40.0	Neither = 1.5

b. Two-class models with no archaeological samples. Class 0 represents forensic samples with known PMI of days to 4.9 years; Class 1 samples have known PMI of 5-71 years.

Set	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
A	Model 77.5 (5)	Model: 72.5 (2)	RMSE = 0.3726 (2)	RMSE = 0.3880 (2)	Overall = 39.3
			R^2 = 0.3037	R^2 = 0.7555	Class 0 = 55.0
	(2) Overall = 75.0	(2) Overall = 71.4	(2) Overall = 78.6	(2) Overall = 75.0	Class 1 = 0.0
	Class 0 = 80.0	Class 0 = 100.0	Class 0 = 85.0	Class 0 = 100.0	Both = 42.9
	Class 1 = 62.5	Class 1 = 0.0	Class 1 = 62.5	Class 1 = 12.5	Neither = 10.7
B	Model 90.0 (3)	Model: 75.0 (4)	RMSE = 0.2573 (4)	RMSE = 0.2496 (4)	Overall = 15.4
			R^2 = 0.6469	R^2 = 0.9222	Class 0 = 20.0
	(2) Overall = 84.6	(2) Overall = 76.9	(4) Overall = 92.3	(4) Overall = 92.3	Class 1 = 0.0
	Class 0 = 80.0	Class 0 = 100.0	Class 0 = 90.0	Class 0 = 90.0	Both = 61.5
	Class 1 = 100.0	Class 1 = 0.0	Class 1 = 100.0	Class 1 = 100.0	Neither = 23.1

c. Three-class models. Classes 0 and 1 represent forensic samples with known PMI (days-4.9 years and 5-71 years, respectively); Class 2 samples are non-forensic specimens (PMI = 500 or 900 years).

Set	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
A	Model 84.3 (3)	Model: 65.9 (4)	Cl 0: RMSE = 0.3677 (2) R^2 = 0.3984 Cl 1: RMSE = 0.3930 (2) R^2 = 0.1762 Cl 2: RMSE = 0.1471 (2) R^2 = 0.7382	Cl 0: RMSE = 0.3924 (4) R^2 = 0.3842 Cl 1: RMSE = 0.3634 (4) R^2 = 0.7996 Cl 2: RMSE = 0.1306 (4) R^2 = 0.8123	Overall = 38.7 Class 0 = 50.0 Class 1 = 0.0 Class 2 = 66.7 Both = 22.6 Neither = 12.9
	(3) Overall = 77.4 Class 0 = 90.0 Class 1 = 37.5 Class 2 = 100.0	(4) Overall = 64.5 Class 0 = 100.0 Class 1 = 0.0 Class 2 = 0.0	(2) Overall = 71.0 Class 0 = 90.0 Class 1 = 12.5 Class 2 = 100.0	(4) Overall = 64.5 Class 0 = 90.0 Class 1 = 12.5 Class 2 = 33.3	
B	Model 80.0 (2)	Model: 64.0 (2)	Cl 0: RMSE = 0.2528 (4) R^2 = 0.7225 Cl 1: RMSE = 0.2397 (4) R^2 = 0.6408 Cl 2: RMSE = 0.0729 (4) R^2 = 0.9605	Cl 0: RMSE = 0.2594 (4) R^2 = 0.6635 Cl 1: RMSE = 0.2746 (4) R^2 = 0.8822 Cl 2: RMSE = 0.1021 (4) R^2 = 0.9349	Overall = 18.75 Class 0 = 10.0 Class 1 = 0.0 Class 2 = 66.7 Both = 50.0 Neither = 31.25
	(3) Overall = 77.4 Class 0 = 90.0 Class 1 = 37.5 Class 2 = 100.0	(2) Overall = 62.5 Class 0 = 100.0 Class 1 = 0.0 Class 2 = 0.0	(4) Overall = 81.3 Class 0 = 90.0 Class 1 = 33.3 Class 2 = 100.0	(4) Overall = 87.5 Class 0 = 90.0 Class 1 = 100.0 Class 2 = 66.7	

Table F. 8. Model and test set classification accuracies for Savitzky-Golay second derivative transformed data (1235-1330, 1535-1585nm) presented by Study Set (A and B). Tables are separated by model type, including a) two-class models (forensic/non-forensic), b) two-class models (forensic samples only), and c) three-class models. Numbers in "( )" represent the number of components or factors used in a given model. Cells in white represent cross-validated model accuracies as a percentage (LDA-PCA and SVM-PCA) or as RMSECV and R^2 for PLSDA. Areas in grey represent accuracies (%) from model application to the test set.

a. Two-class models (forensic, non-forensic). Class 0 represents forensic samples with known or estimated PMI of less than 71 years; Class 1 represents non-forensic specimens (historic and archaeological origins).

Set	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
A	Model 94.1 (3)	Model: 79.5 (2)	RMSE = 0.2360 (3)	RMSE = 0.2408 (2)	Overall = 45.7
			R^2 = 0.6618	R^2 = 0.9274	Class 0 = 36.0
	(3) Overall = 94.7	(2) Overall = 79.8	(3) Overall = 95.8	(2) Overall = 94.7	Class 1 = 84.2
	Class 0 = 97.3	Class 0 = 100	Class 0 = 100	Class 0 = 100	Both = 47.9
	Class 1 = 84.2	Class 1 = 0	Class 1 = 78.9	Class 1 = 78.9	Neither = 4.3
B	Model 94.3 (3)	Model: 84.1 (2)	RMSE = 0.2406 (2)	RMSE = 0.2389 (2)	Overall = 27.3
			R^2 = 0.5677	R^2 = 0.9321	Class 0 = 25.0
	(2) Overall = 90.9	(2) Overall = 84.8	(2) Overall = 90.9	(2) Overall = 92.3	Class 1 = 40.0
	Class 0 = 100.0	Class 0 = 100.0	Class 0 = 100.0	Class 0 = 100.0	Both = 72.7
	Class 1 = 40.0	Class 1 = 0.0	Class 1 = 40.0	Class 1 = 40.0	Neither = 0.0

b. Two-class models with no archaeological samples. Class 0 represents forensic samples with known PMI of days to 4.9 years; Class 1 samples have known PMI of 5-71 years.

Set	LDA-PCA	SVM-PCA	$PLSDA_{ctr}$	$PLSDA_{corr}$	SIMCA
A	Model 82.5 (5)	Model: 72.5 (2)	RMSE = 0.3946 (1)	RMSE = 0.3629 (3)	Overall = 42.9
			R^2 = 0.2189	R^2 = 0.7860	Class 0 = 60.0
	(2) Overall = 75.0	(2) Overall = 71.4	(1) Overall = 75.0	(3) Overall = 67.9	Class 1 = 0.0
	Class 0 = 80.0	Class 0 = 100.0	Class 0 = 90.0	Class 0 = 75.0	Both = 39.3
	Class 1 = 62.5	Class 1 = 0.0	Class 1 = 37.5	Class 1 = 50.0	Neither = 14.3
B	Model 90.0 (4)	Model: 75.0 (4)	RMSE = 0.3414 (1)	RMSE = 0.3089 (2)	Overall = 15.4
			R^2 = 0.3784	R^2 = 0.8807	Class 0 = 20.0
	(2) Overall = 84.6	(2) Overall = 76.9	(1) Overall = 84.6	(2) Overall = 84.6	Class 1 = 0.0
	Class 0 = 80.0	Class 0 = 100.0	Class 0 = 80.0	Class 0 = 90.0	Both = 69.4
	Class 1 = 100.0	Class 1 = 0.0	Class 1 = 100.0	Class 1 = 66.7	Neither = 15.4

c. Three-class models. Classes 0 and 1 represent forensic samples with known PMI (days-4.9 years and 5-71 years, respectively); Class 2 samples are non-forensic specimens (PMI = 500 or 900 years).

Set	LDA-PCA	SVM-PCA	$PLSDA_{MC}$	$PLSDA_{CO}$	SIMCA
A	Model 77.3 (3)	Model: 65.9 (3)	Cl 0: RMSE = 0.3815 (1) R^2 = 0.3524 Cl 1: RMSE = 0.4316 (1) R^2 = 0.0063 Cl 2: RMSE = 0.1472 (2) R^2 = 0.7379	Cl 0: RMSE = 0.3910 (3) R^2 = 0.3885 Cl 1: RMSE = 0.3556 (3) (3) R^2 = 0.8082 Cl 2: RMSE = 0.1041 (3) R^2 = 0.8808	Overall = 32.3 Class 0 = 40.0 Class 1 = 0.0 Class 2 = 66.7 Both = 48.4 Neither = 16.1
	(3) Overall = 67.7 Class 0 = 65.0 Class 1 = 62.5 Class 2 = 100.0	(3) Overall = 64.5 Class 0 = 100.0 Class 1 = 0.0 Class 2 = 0.0	(1) Overall = 64.5 Class 0 = 85.0 Class 1 = 0.0 Class 2 = 100.0	(3) Overall = 74.2 Class 0 = 95.0 Class 1 = 12.5 Class 2 = 100.0	
B	Model 96.0 (2)	Model: 64.0 (2)	Cl 0: RMSE = 0.3446 (1) R^2 = 0.4845 Cl 1: RMSE = 0.3997 (1) R^2 = 0.0.0015 Cl 2: RMSE = 0.1104 (1) R^2 = 0.9093	Cl 0: RMSE = 0.3985 (2) R^2 = 0.2061 Cl 1: RMSE = 0.3466 (2) R^2 = 0.8123 Cl 2: RMSE = 0.1107 (2) R^2 = 0.9234	Overall = 25.0 Class 0 = 20.0 Class 1 = 0.0 Class 2 = 66.7 Both = 56.3 Neither = 18.8
	(3) Overall = 67.7 Class 0 = 65.0 Class 1 = 62.5 Class 2 = 100.0	(2) Overall = 62.5 Class 0 = 100.0 Class 1 = 0.0 Class 2 = 0.0	(1) Overall = 93.8 Class 0 = 90.0 Class 1 = 100.0 Class 2 = 100.0	(2) Overall = 81.3 Class 0 = 100.0 Class 1 = 0.0 Class 2 = 100.0	

APPENDIX G

TESTS FOR PROTEIN OXIDATION

G1.0 Examination of Potential Oxidative Changes on Cortical Samples

G1.1 Introduction

Recall that bone is a biphasic material comprised of approximately 30% organic content, the bulk of which is Type I collagen. Organic materials, including proteins such as collagen, are known to oxidize in biological systems in the presence of reactive species, including oxidized lipids, oxygen and oxygen radicals, and metal ions, to name a few (Berlett and Stadtman 1997, Hawkins and Davies 2001, Stadtman and Levine 2003, Davies 2005). Oxidative modifications of proteins may involve the protein backbone or individual amino acid side groups. Considerations must therefore be made about the potential for bone protein oxidation occurring, not only during the postmortem interval, but on the superficial surfaces and freshly exposed cortex of sectioned surfaces of sampled bone as well. If such changes occur, absorbance spectra collected from those surfaces may vary over time.

The primary question is whether or not spectra collected from sectioned bone remain acceptably constant. If not, what are the likely causes for observed absorbance changes, and more importantly, are the changes predictable? Sources for spectral change can be attributed to the following: loss of water content; loss of total organic content; potential oxidation effects, primarily on the collagen; and, selection of different spectral sampling points along the bone (Campbell, personal communication). When dealing with sampled cortex in a lab setting, the first two scenarios are unlikely to be sources for spectral change because 1) the bone is already "dry" and 2) the samples are stored in a protected environment where further collagen loss is minimized, i.e. no potential for additional diagenesis.

G1.2 Methodology

To assess the potential for time-related protein oxidation on extracted bone samples, a follow-up study was conducted in February 2017 at ASD-PANalytical in Boulder, Colorado. Two groups of cortical samples were analyzed. The first group consisted of 16 samples that were used in the preliminary study conducted in May 2013 ("Resampled Set"). This set was examined to determine if there was a change in the absorbance spectrum between Time 1 (May 2013) and Time 2 (February 2017), a roughly 4 year interval. Selected samples included the following: "Forensic" samples 1, 2, 4, 37, 46, 47, 48, 78, 85, 86, 87, and 88; and "Non-forensic" (historic/archaeological) samples 165, 166, 167, and 169.

An additional sampling group ("Dilution Set") was comprised of six new pieces of cortex, four of which were exposed to dilutions of hydrogen peroxide (H_2O_2). This was done in order to determine how bone collagen reacts when exposed to a strong oxidizer over a fixed period of time. To assess its potential oxidative effect on absorbance spectra, cortical samples were exposed to varying concentrations of H_2O_2. Six adjacent pieces of cortex were removed from the anterior diaphyseal surface of a right femur. The bone, made available from the LFA donated study collection, was best-described as untreated and slightly greasy to the touch with some associated fibrous tissue. Samples were removed on 1/31/2017 with an autopsy saw, each having approximate dimensions and weight of 1cm x 2cm and 2.25 grams, respectively. An H_2O_2 dilution series was prepared on 2/1/2017 and is presented in Table G.1 below. Samples 1 through 4 were placed in marked 75ml centrifuge tubes containing 15ml of diluted H_2O_2 (ml H_2O_2 in 100ml of ddH_2O, reported as percentage; 30%, 15%, 7%, and 3%, respectively). For control, sample 5 was placed in tube containing 15ml of ddH_2O. The final sample, number 6,

was placed in an empty “air-filled” tube with no added fluid. Each tube was closed, sealed in parafilm, placed in a tube rack, and stored in the dark under a fume hood at room temperature for 10 days. Samples were subsequently removed from the tubes on 2/10/2017, rinsed with ddH_20, and allowed to air dry under the hood for an additional 10 days. After drying, LFA staff members assigned random numbers to each tube and reassigned the randomized samples in new tubes marked A through F (Table G.2).

Resampled Set and Dilution Set spectra were collected on 2/24/2017 with a 350-2500 Hi-Res Lab LabSpec® 4 spectrometer at ASD PANalytical in Boulder, Colorado. The general protocol described in Section 5.3 was used for data collection. However, it should be noted that this instrument, unlike the spectrometer made available through the Goetz Program, had an internal light source, and was the same model as the one used in the 2013 preliminary study.

Resampled Set spectra (Time 2) were collected from the longitudinally cut surfaces of each cortical sample. The Time 2 spectra were later overlaid in Unscrambler with the corresponding Time 1 series. SNV-transformation was applied in order to remove baseline shift artifacts. This allowed for a direct comparison between spectra to readily identify potential peak changes (i.e. potential oxidative chemical change). Resultant SNV-transformed spectra are presented for each sample in Figures G.1 through G.16 below.

Dilution Set spectra were collected from the superficial, transversely cut, and longitudinally cut surfaces for each of the six cortical samples. All collected spectra were later loaded into Unscrambler and SNV-transformed. These spectra are presented by group and surface in Figures G.17 through G.19.

G1.3 Results

G1.3.1 Resample Set

Aside from fluctuations that can be attributed to probe positioning along the sample edge, Time 1 and Time 2 SNV-transformed spectra were the same for forensic samples 1, 37, 78, 85, as well as a sample 169, which was extracted from historic cemetery burial remains (Figures G.1, G.4, G.8, G.9, and G.16). Slight differences were noted on the remaining forensic samples. Numbers 47, 87, and 88 showed an elevation in the Time 2 spectra at the 1940 peak; as noted previously, this peak is associated with a carbonyl vibration (Figures G.6, G.11, and G.12). There was a minor separation between Time 1 and Time 2 spectra for samples 46 and 86 over a broad band of roughly 1400-1880nm; this band includes peaks at 1490nm and 1730nm, which are often associated with N-H 1^{st} overtones and the C-H and S-H 1^{st} overtones, respectively (Figures G.5 and G.10). Note that other vibrations observed at 1490nm are associated with water and alcohols, as well as carbonyl groups (e.g. aldehydes, ketones, and amides; see Appendix B). An additional three forensic samples (2, 4, and 48) showed separation over the 1400-1880nm band and at the 1940nm peak (Figures G.2, G.3, and G.7). Time 1 and Time 2 spectra were broadly separated from 1400-2200nm on non-forensic samples 165-167 (Figures G.13 – G.15).

G1.3.2 Dilution Set

Minor separation of the “air” sample spectrum (lower absorption) from the remaining group of spectra was noted at the 1490nm peaks for the SNV-transformed spectra collected from superficial or longitudinally cut surfaces (Figures G.17 and G.18, respectively). Slight

differences were also noted on the spectra collected from the transversely cut surfaces at 1730nm and 1940nm. The control sample stored in air showed an elevated absorbance at 1730nm, while the dilution series and water control sample were lower. The inverse relationship was observed at 1940nm (Figure G.19).

G1.4 Discussion

G1.4.1 Resample Set

Spectrum-wide differences were noted on three of the non-forensic samples (165 – 167). The primary issues with older bone are a combination of organic content loss over the extended PMI, and changes in mineral composition and organization related to diagenesis. This modified cortex is often characterized as relatively rough in texture with increased grain size, which results in uneven and grainy surfaces that produce slight changes in both probe orientation and incident light directed from the spectrometer probe. Coupled with variation in sampling point selections, separation between Time 1 and Time 2 is not unexpected.

The samples originating in the "forensic" time frame have relatively smoother sectioned cortex that can be more consistently sampled. Recall that four of the forensic samples (1, 37, 78, and 85) and historic sample 169 yielded no difference between Time 1 and Time 2; in the remaining eight forensic samples, only slight differences were observed. These samples were notable for relative high grease content (4, 88) and significant weathering and cortical exfoliation (2, 46, 47, 48, 86, and 87). Changes at the 1490nm and 1940nm peaks may indicate a relative increase in carbonyl compounds (ketones and aldehydes), which are primary breakdown products of amino acid oxidation via O_2 gas (Stadtman and Levine 2003, Davies

2005). As noted, absorption peaks at 1730nm generally represent SH and CH vibrations. Due to its absence in mature collagen, the oxidation of cysteine is unlikely (i.e. formation of disulfide bonds). However, aliphatic amino acid residues account for approximately 50% of the residues in mature collagen (Appendix A). Accordingly, changes at this peak may represent an elevated concentration of methyl groups.

G1.4.2 Dilution Set

Similar to some of the forensic samples, slight changes were observed at 1730nm and 1940nm for transverse surface spectra, and 1490nm for the superficial and longitudinal surface spectra. As previously noted, these peaks are most likely associated with CH and carbonyl vibrations. Because the six samples in the Dilution Set Differences were extracted from adjacent positions on a single femoral diaphysis, native composition should not change appreciably. The differences observed between the transverse group and the other two groups are therefore likely related to general osteon, and thus collagen fibril, orientation.

G1.5 Conclusions

Peak changes observed with some of the forensic samples (Resample Set), as well as the Dilution Set spectra, may be due to collagen oxidation; however, sampling error related to probe position cannot be ruled out (Campbell, personal communication). Final resolution of the oxidation question would require a determination of chemical composition, which was not a part of this study. It is important to note that although slight peak changes were present on

some of the Resample Set spectra, those changes were morphologically minimal and would likely have no impact on the classification methods developed in this study.

Table G. 1. Dilution Set. Concentration of hydrogen peroxide (H2O2) is reported as a percentage (ml H2O2 in 100ml ddH2O).

Sample #	$[H_2O_2]$	v(ml) 30% H_2O_2	v(ml) ddH_2O	Total volume (ml)
1	30%	15	0	15
2	15%	7.5	7.5	15
3	7%	3.5	11.5	15
4	3%	1.5	13.5	15
5	0%	0	15	15
6	Air	n/a	n/a	n/a

Table G. 2. Dilution Set. Blind sample assignments are reported.

Sample #	$[H_2O_2]$	Random	Blind Sample
1	30%	38	B
2	15%	93	F
3	7%	44	C
4	3%	2	A
5	0%	49	D
6	Air	55	E

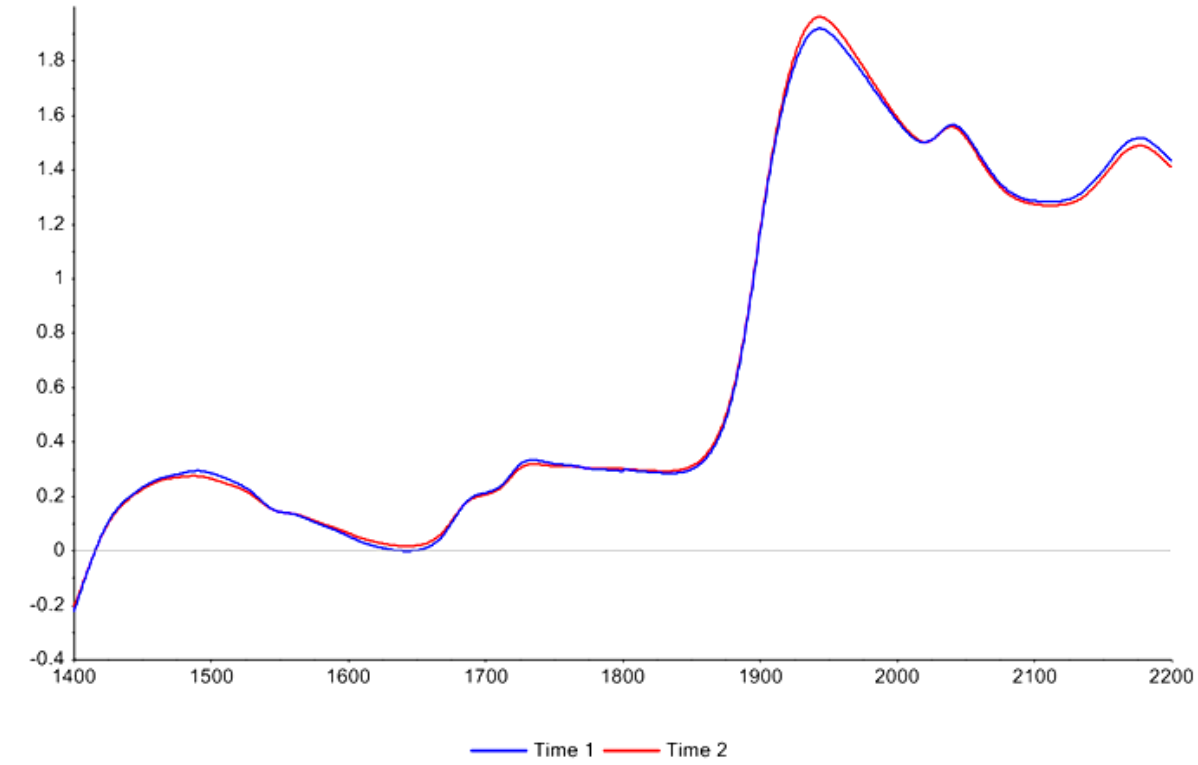


Figure G. 1. SNV-transformed Time 1 and Time 2 spectra for forensic sample 1.

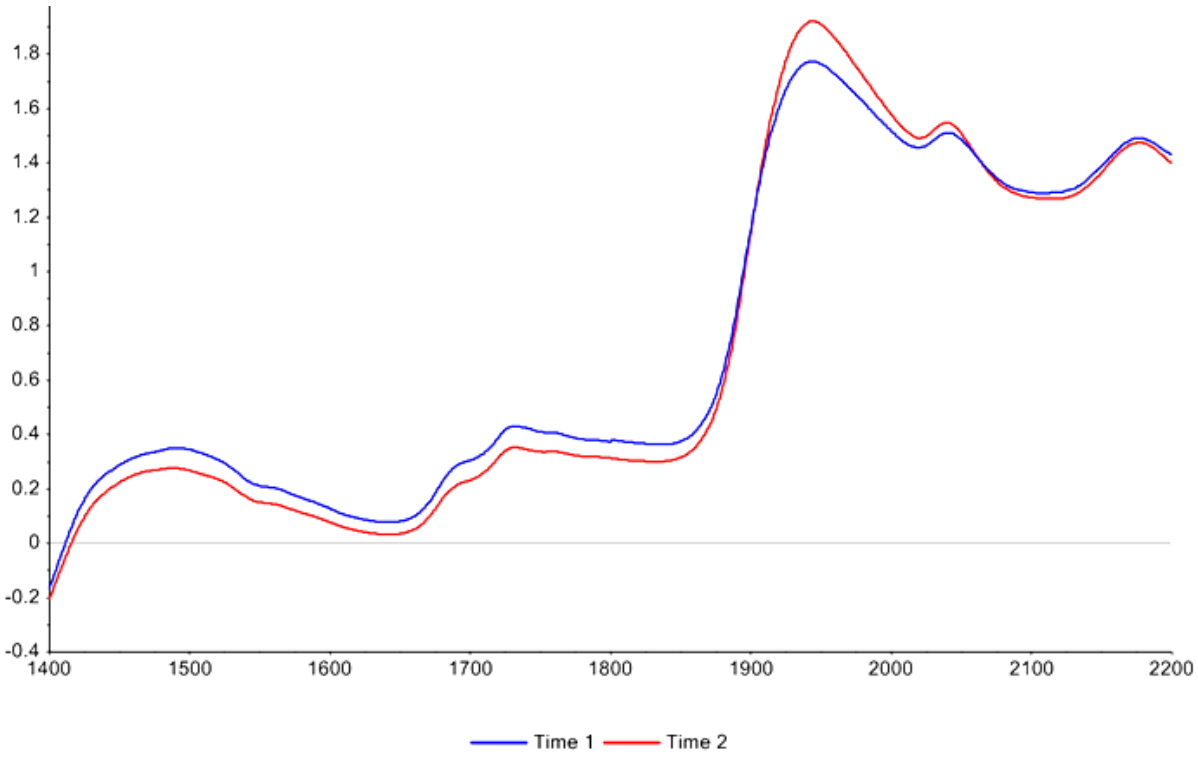


Figure G. 2. SNV-transformed Time 1 and Time 2 spectra for forensic sample 2. Slight separation is noted broadly from 1400-1880nm, as well as at the 1940nm and 2040nm peaks.

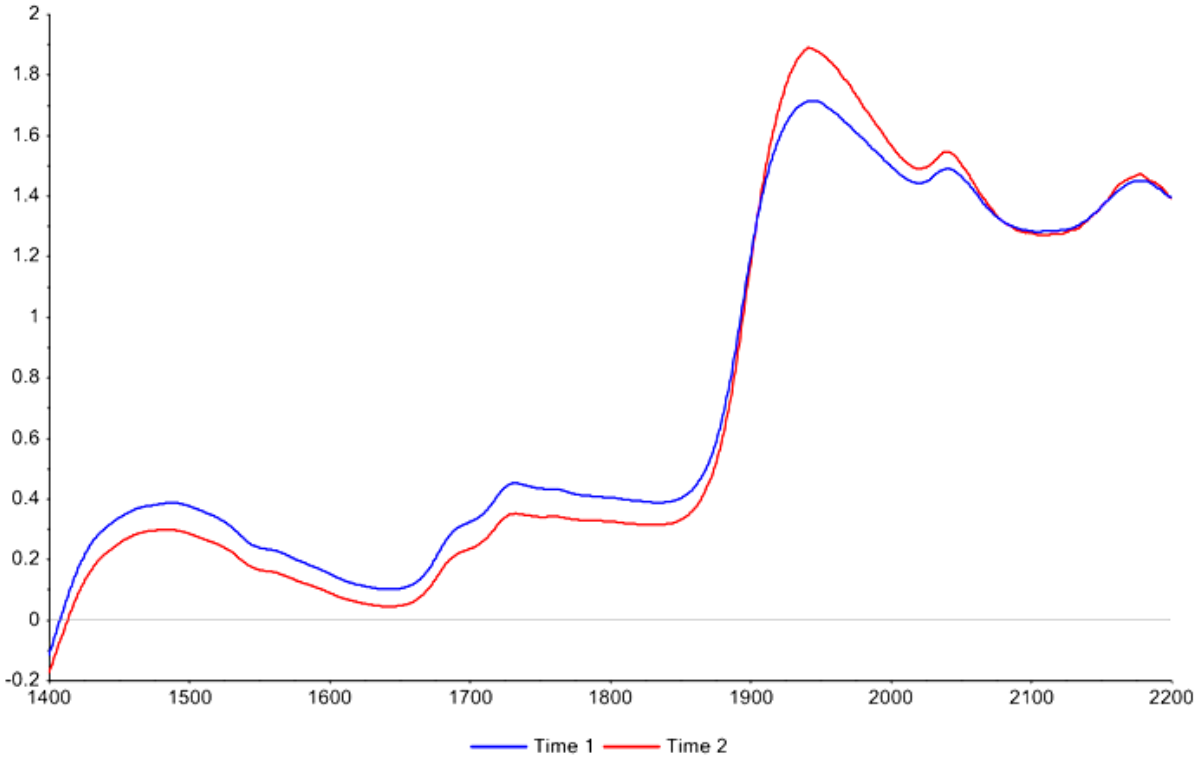


Figure G. 3. SNV-transformed Time 1 and Time 2 spectra for sample 4. Slight separation is noted broadly from 1400-1880nm, as well as at the 1940nm and 2040nm peaks.

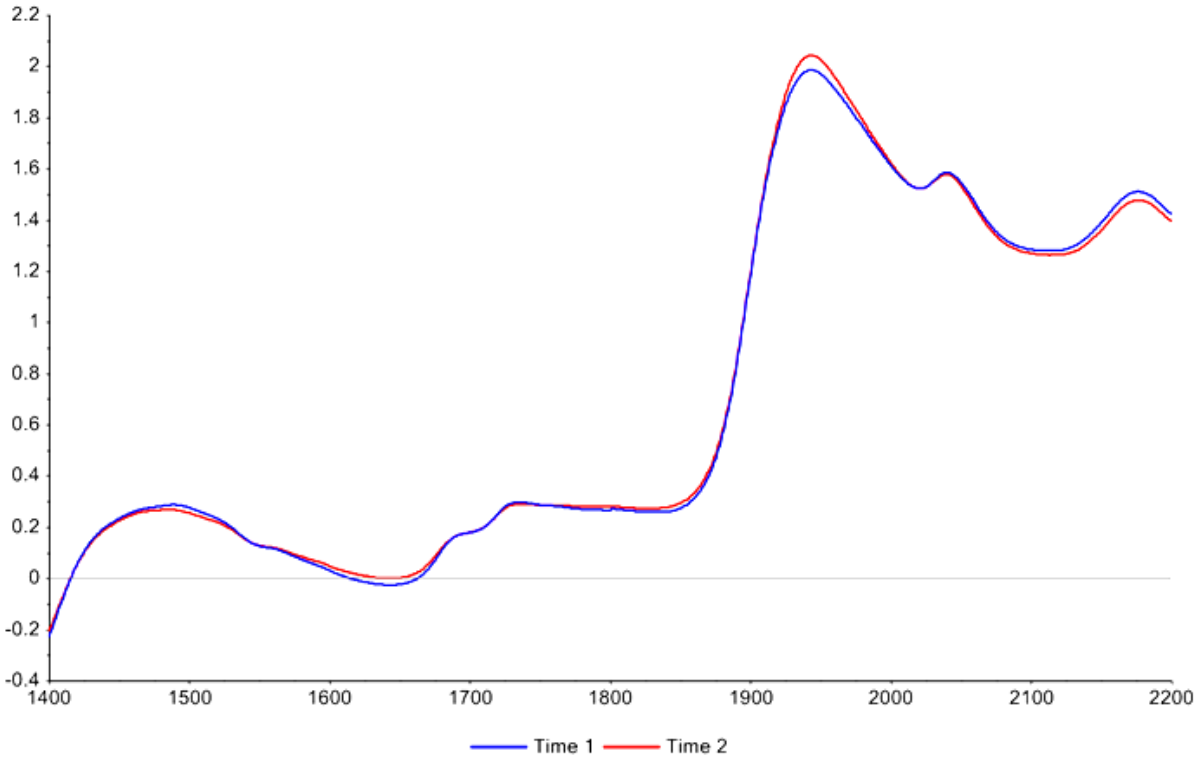


Figure G. 4. SNV-transformed Time 1 and Time 2 spectra for forensic sample 37.

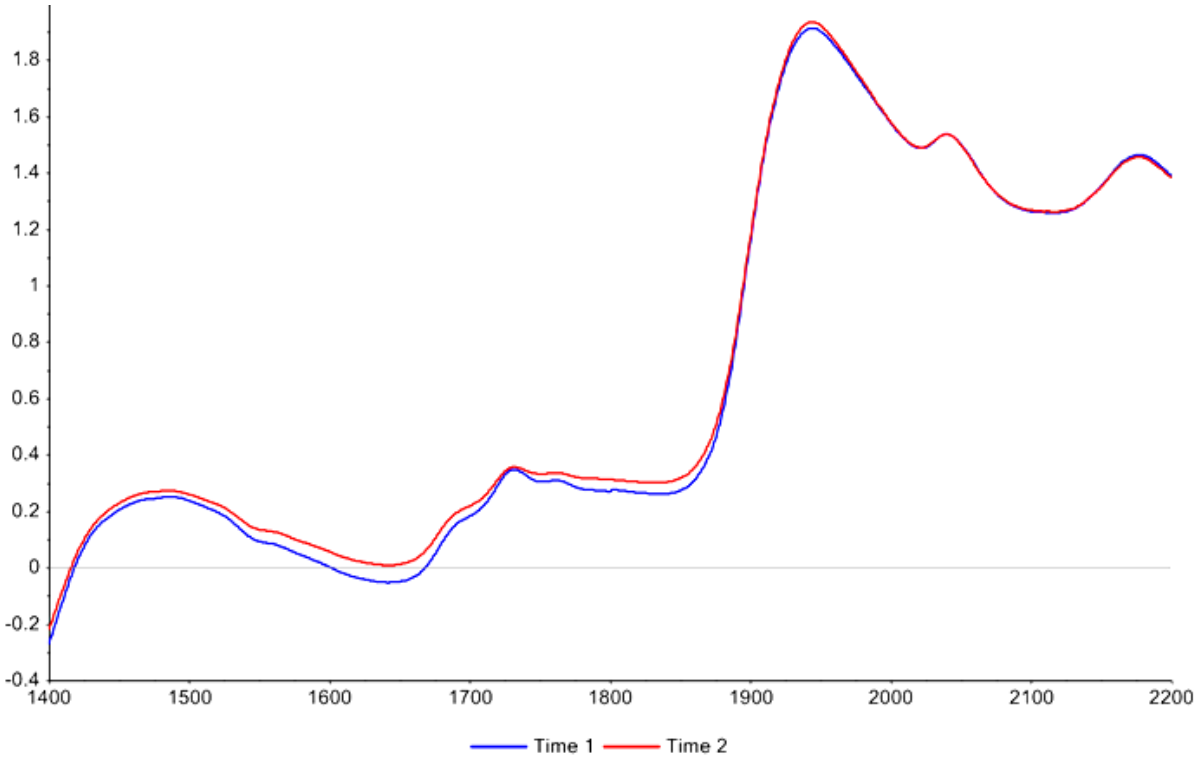


Figure G. 5. SNV-transformed Time 1 and Time 2 spectra for forensic sample 46.

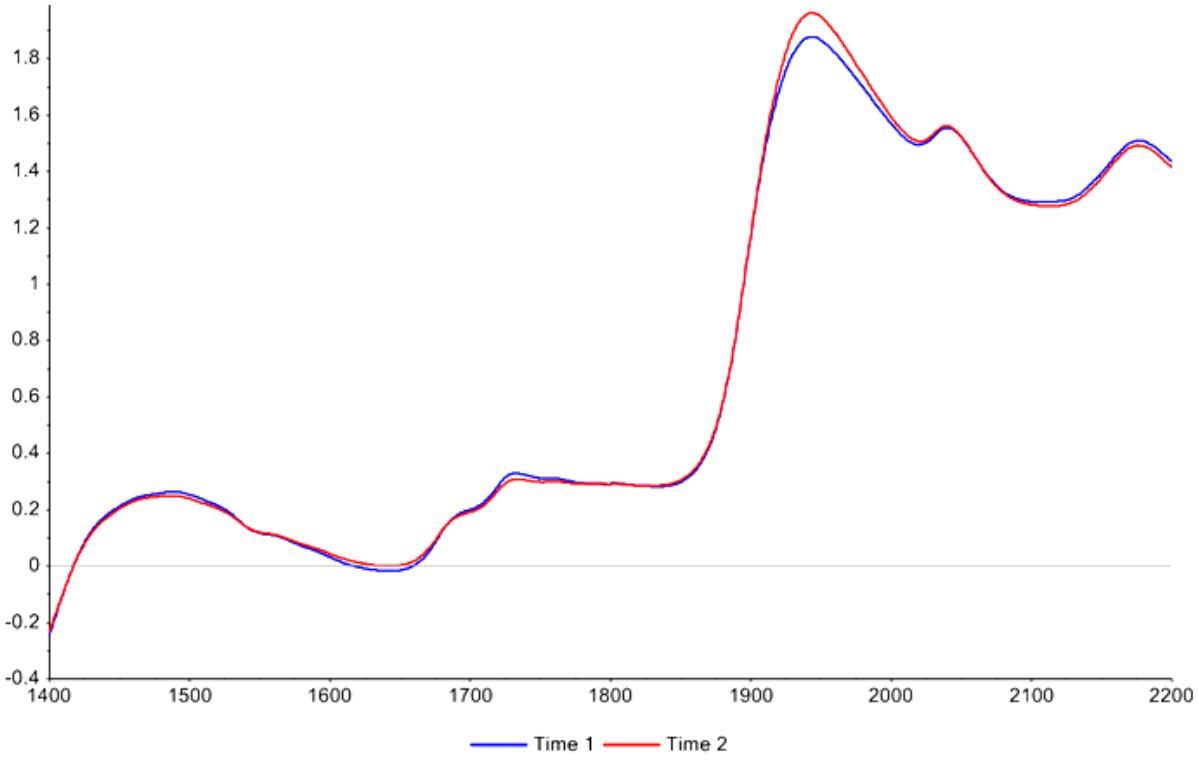


Figure G. 6. SNV-transformed Time 1 and Time 2 spectra for forensic sample 47. A slight elevation is noted at the 1940nm peak.

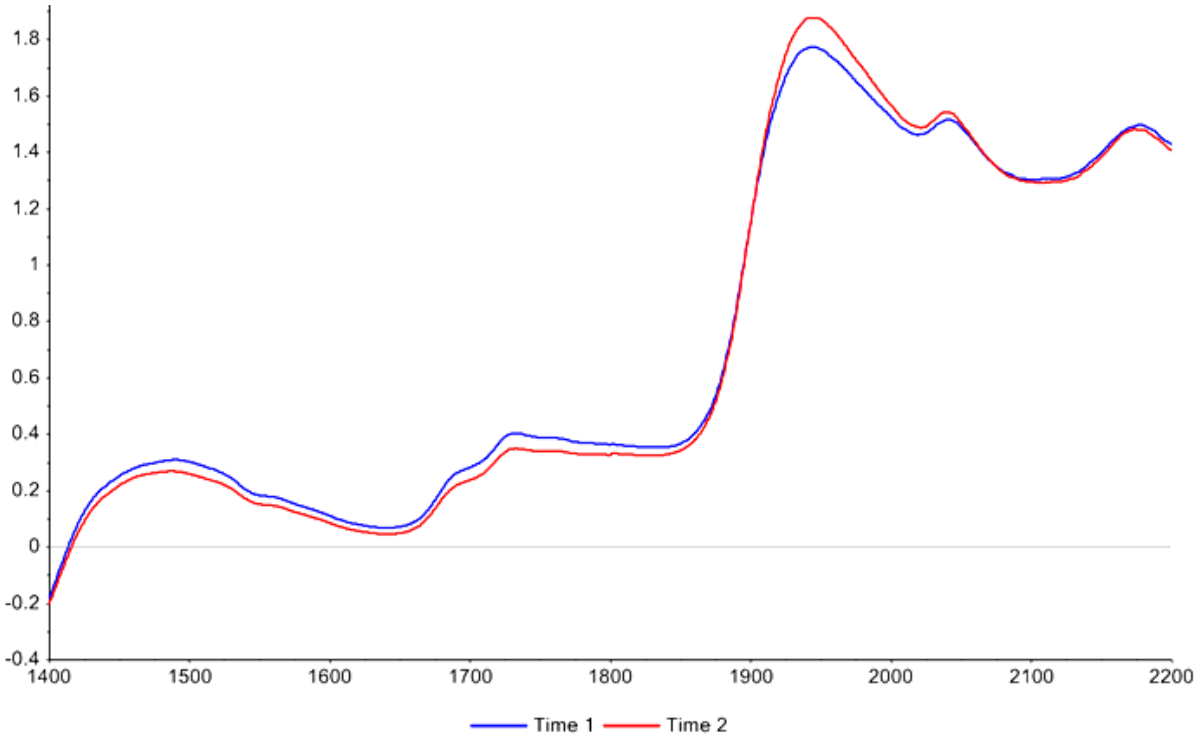


Figure G. 7. SNV-transformed Time 1 and Time 2 spectra for forensic sample 48. Slight separation is noted broadly from 1400-1880nm, as well as at the 1940nm and 2040nm peaks.

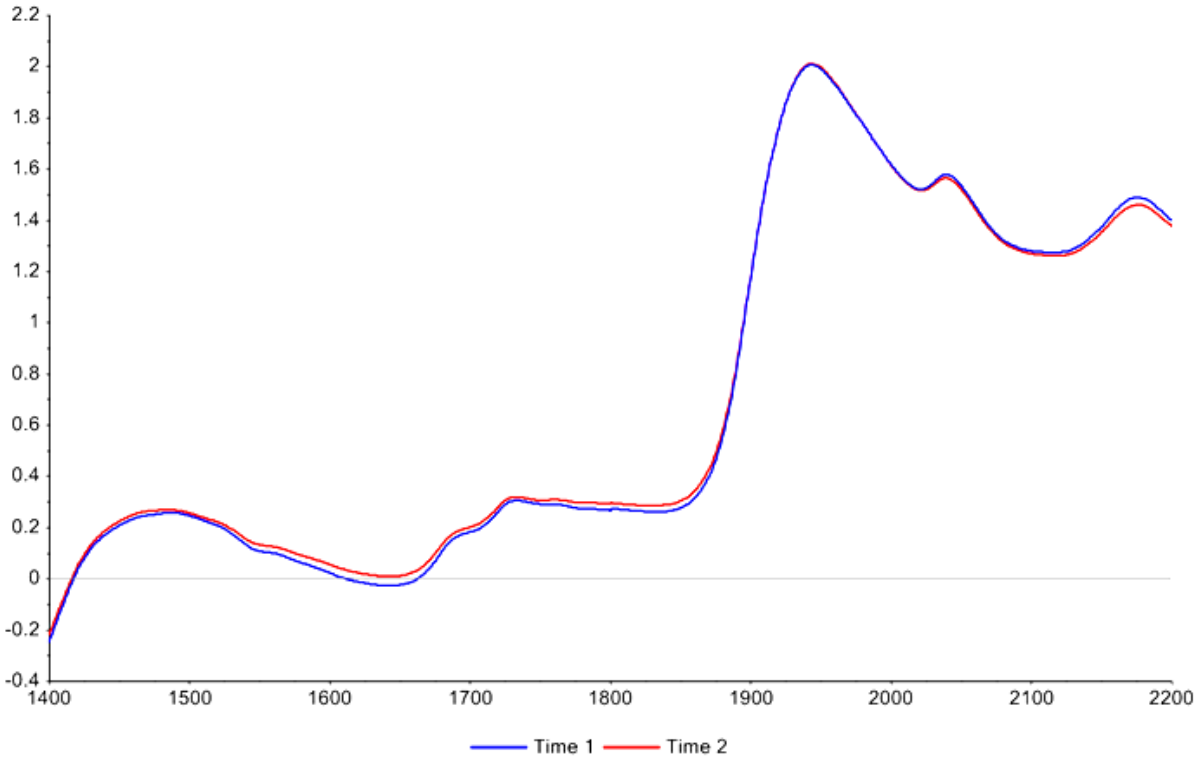


Figure G. 8. SNV-transformed Time 1 and Time 2 spectra for sample 78.

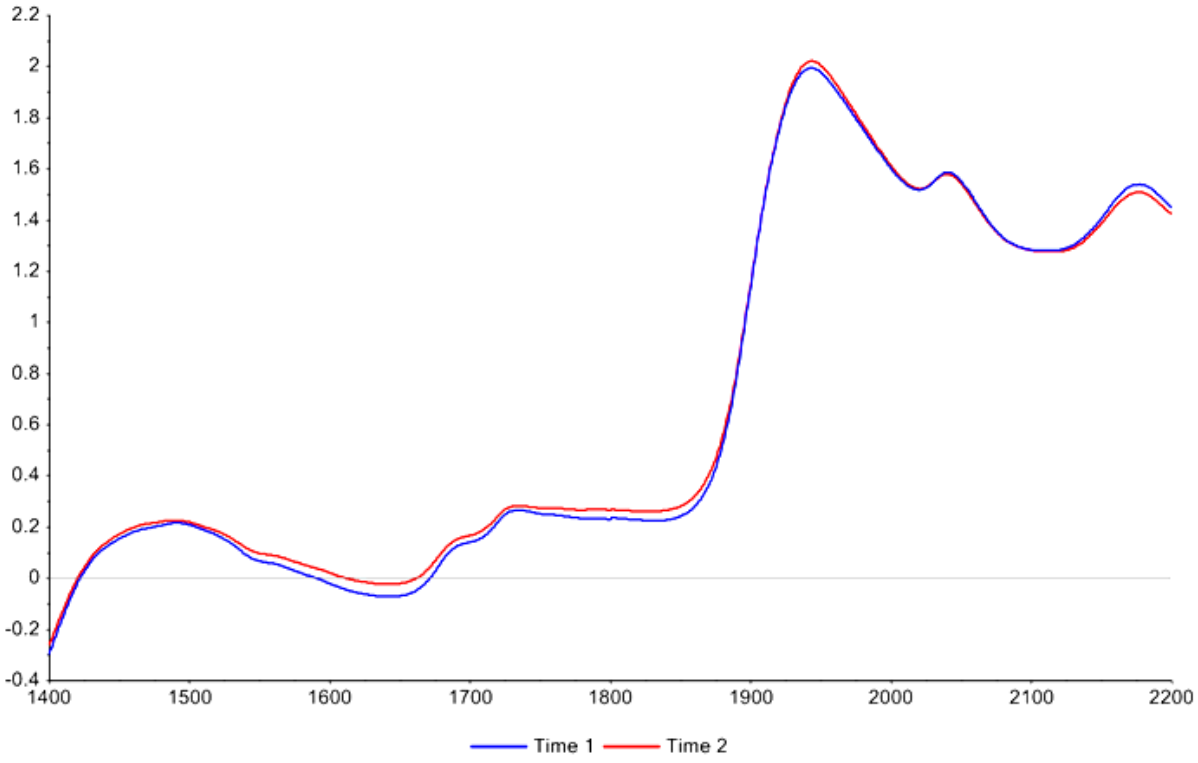


Figure G. 9. SNV-transformed Time 1 and Time 2 spectra for forensic sample 85.

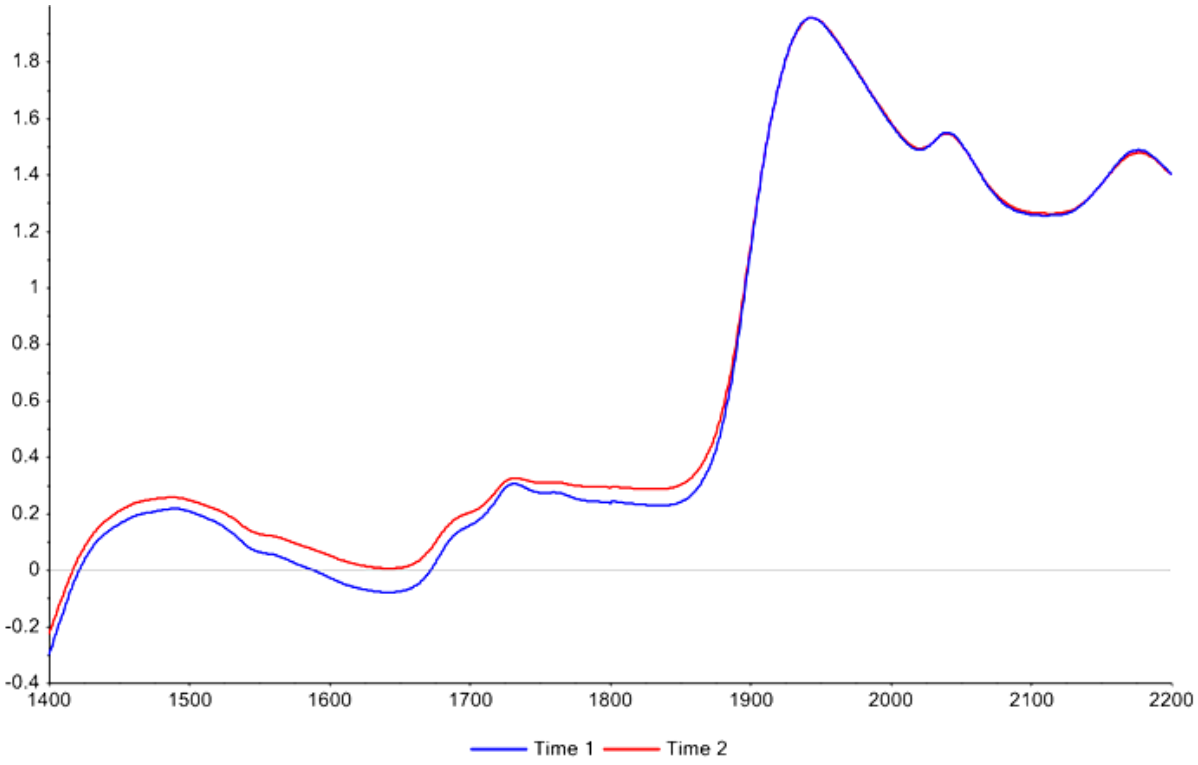


Figure G. 10. SNV-transformed Time 1 and Time 2 spectra for forensic sample 86. A slight separation is noted broadly from 1400-1880nm.

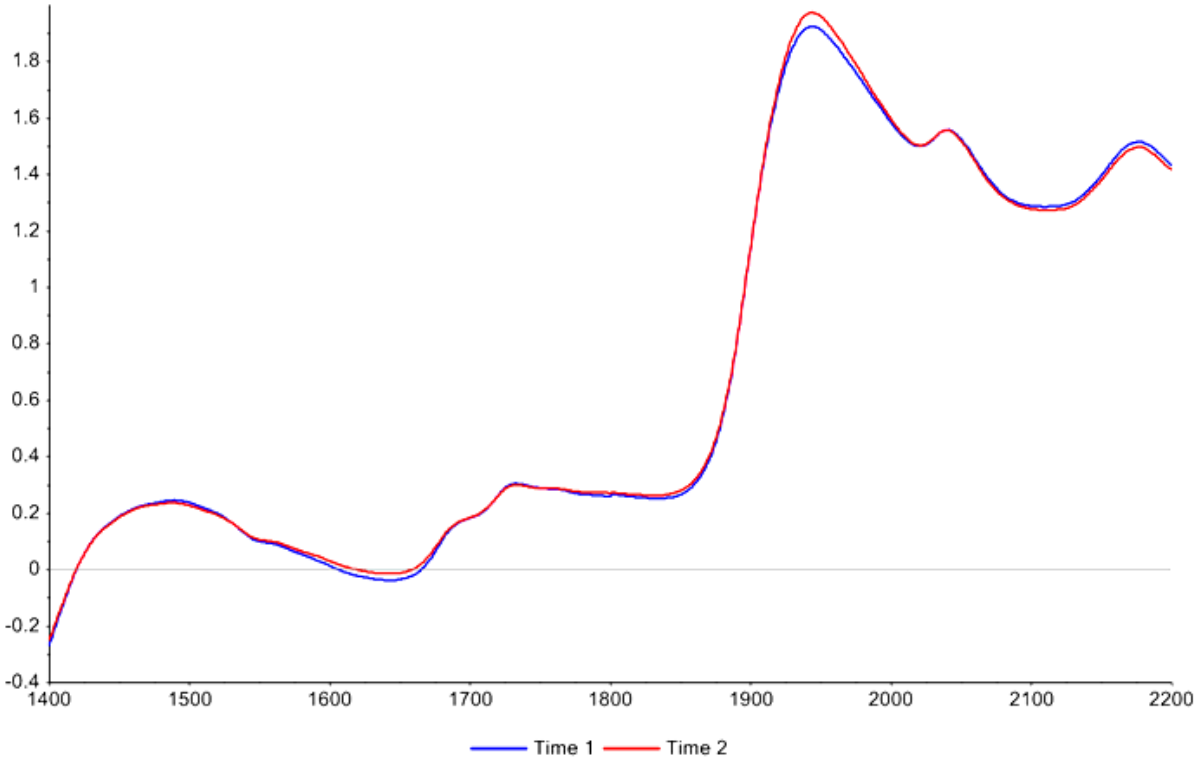


Figure G. 11. SNV-transformed Time 1 and Time 2 spectra for forensic sample 87. A slight elevation is noted at 1940nm.

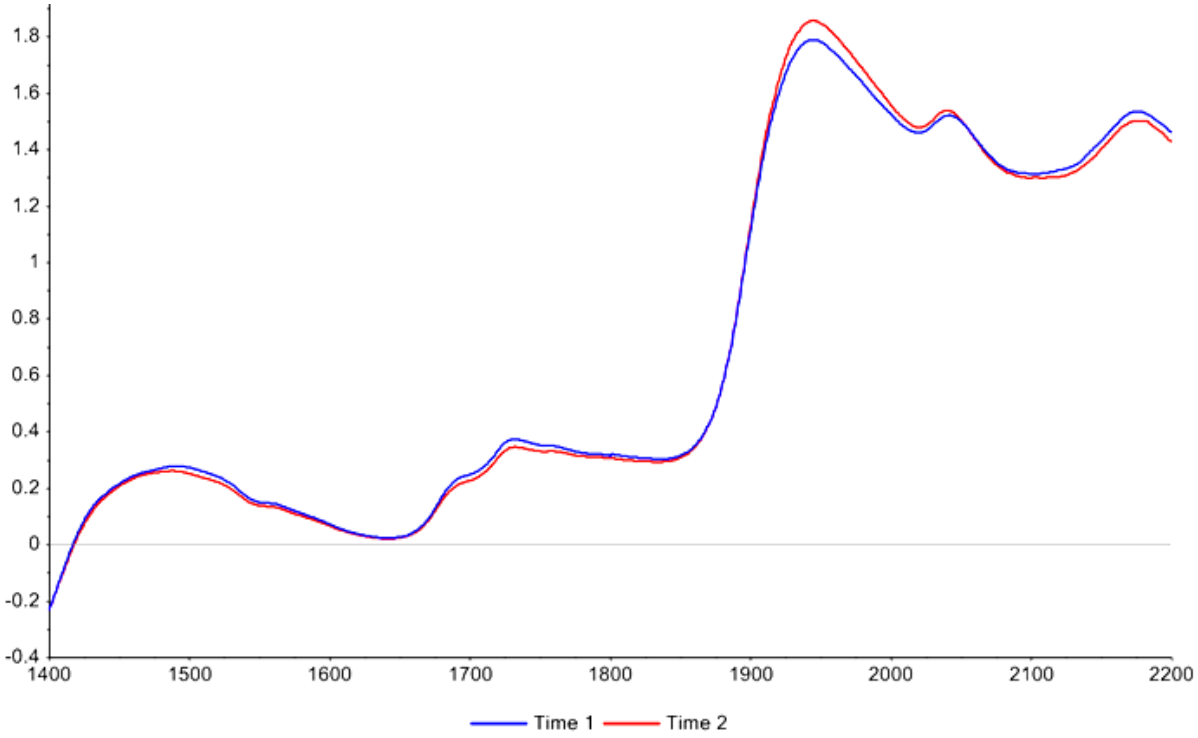


Figure G. 12. SNV-transformed Time 1 and Time 2 spectra for forensic sample 88. A slight elevation is noted at 1940nm.

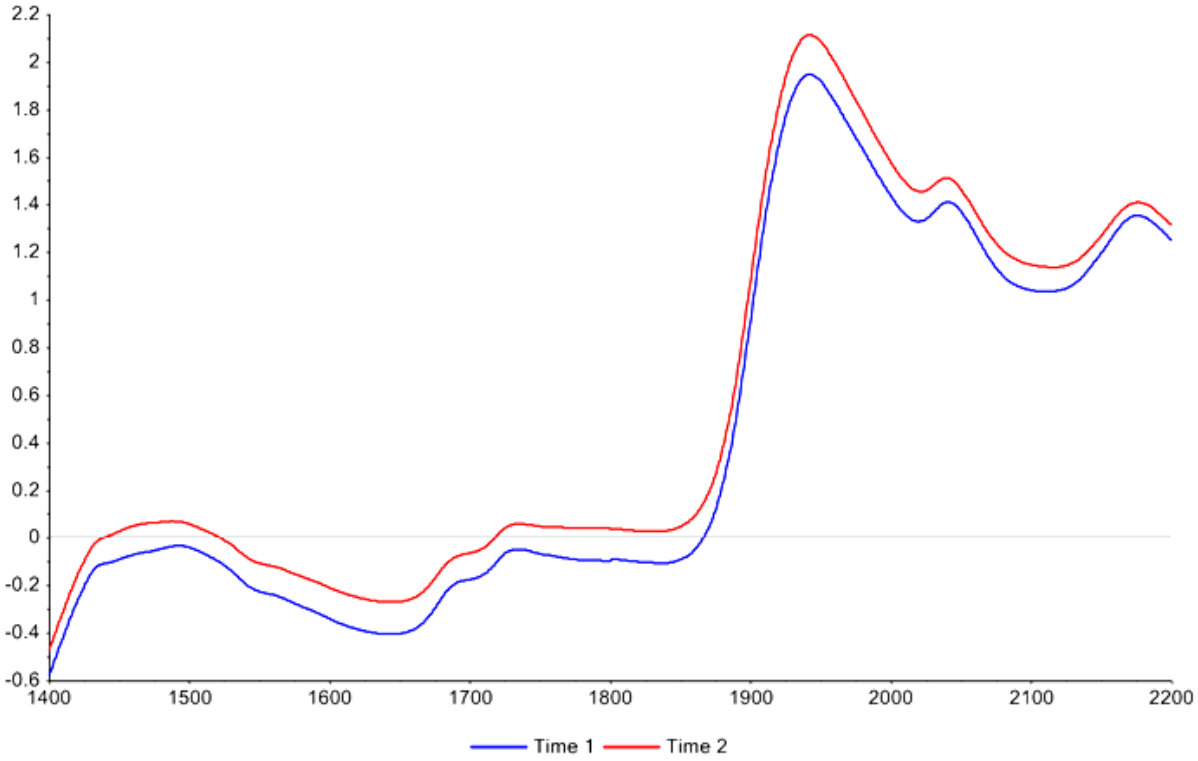


Figure G. 13. SNV-transformed Time 1 and Time 2 spectra for historic sample 165. Broad whole-spectra separation is noted.

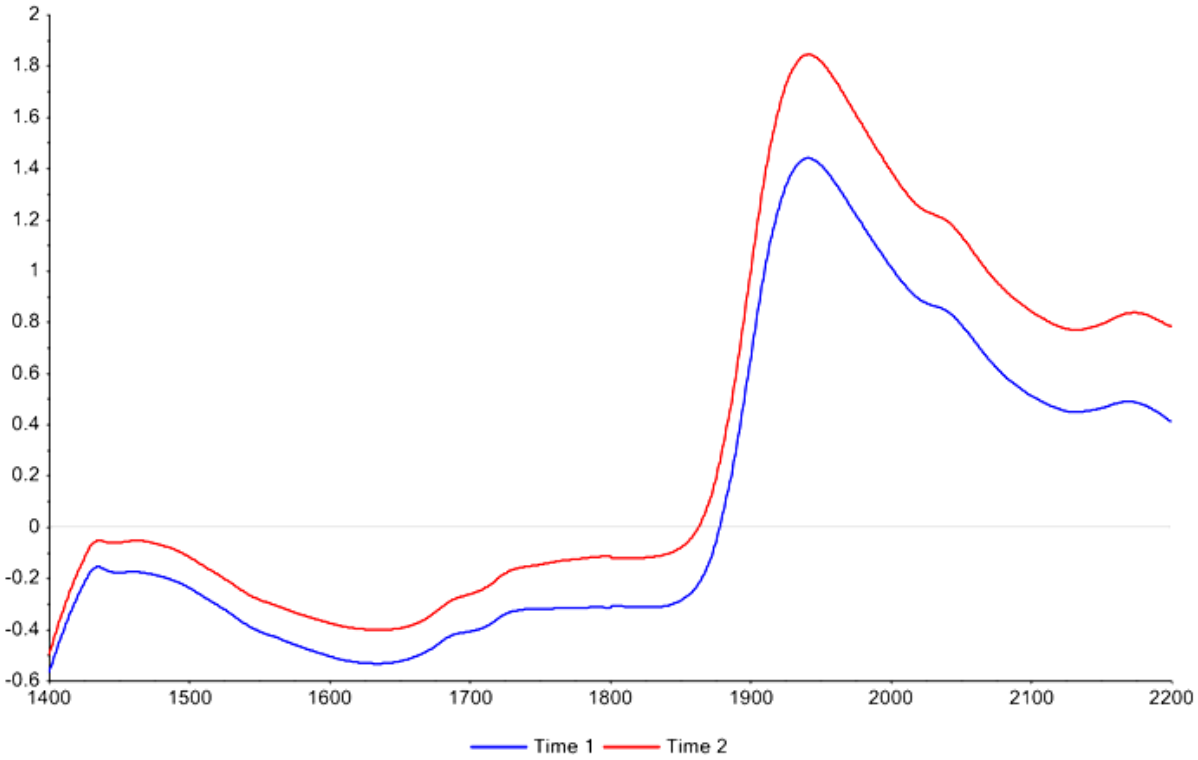


Figure G. 14. SNV-transformed Time 1 and Time 2 spectra for archaeological sample 166. Broad whole-spectra separation is noted.

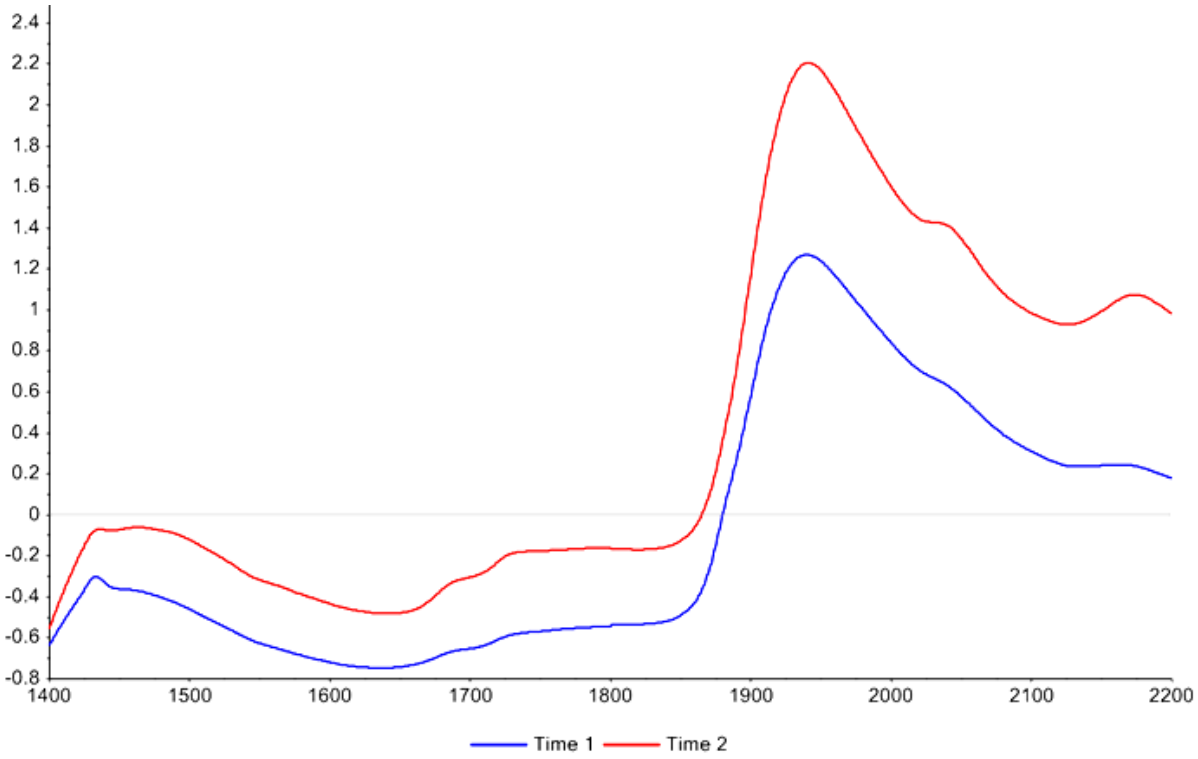


Figure G. 15. SNV-transformed Time 1 and Time 2 spectra for archaeological sample 167. Broad whole-spectra separation is noted.

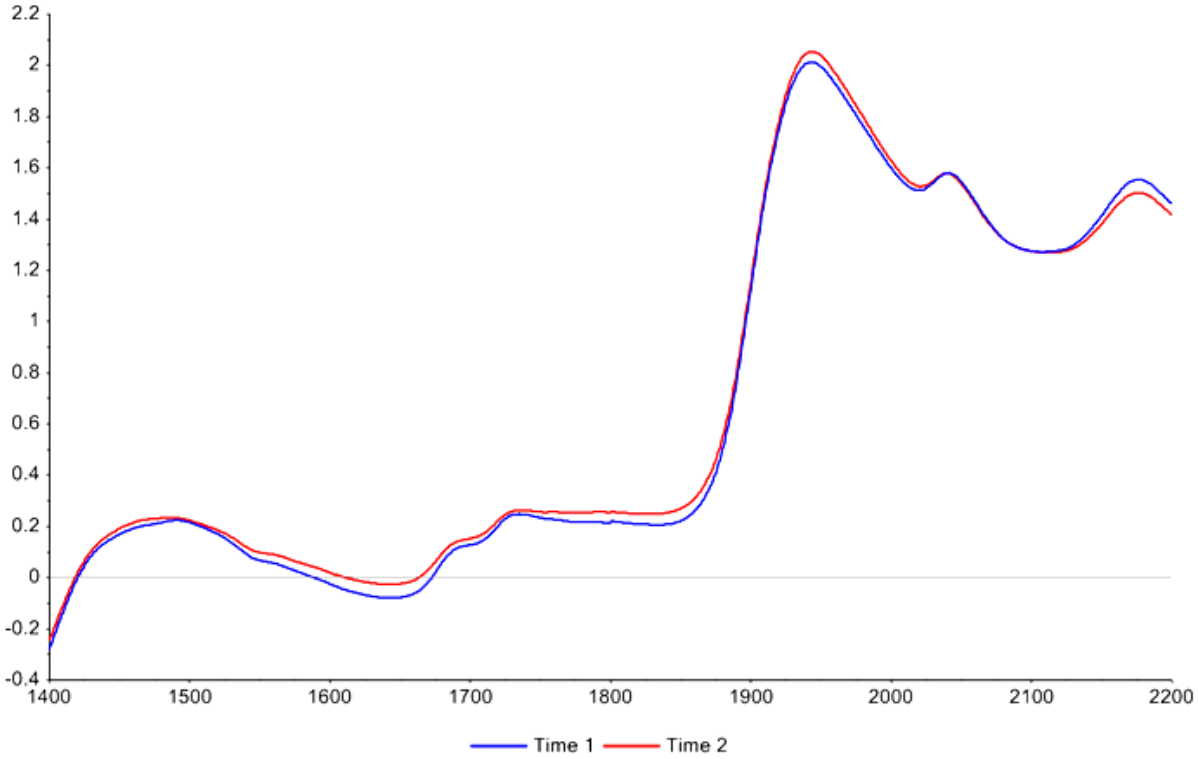


Figure G. 16. SNV-transformed Time 1 and Time 2 spectra for historic sample 169.

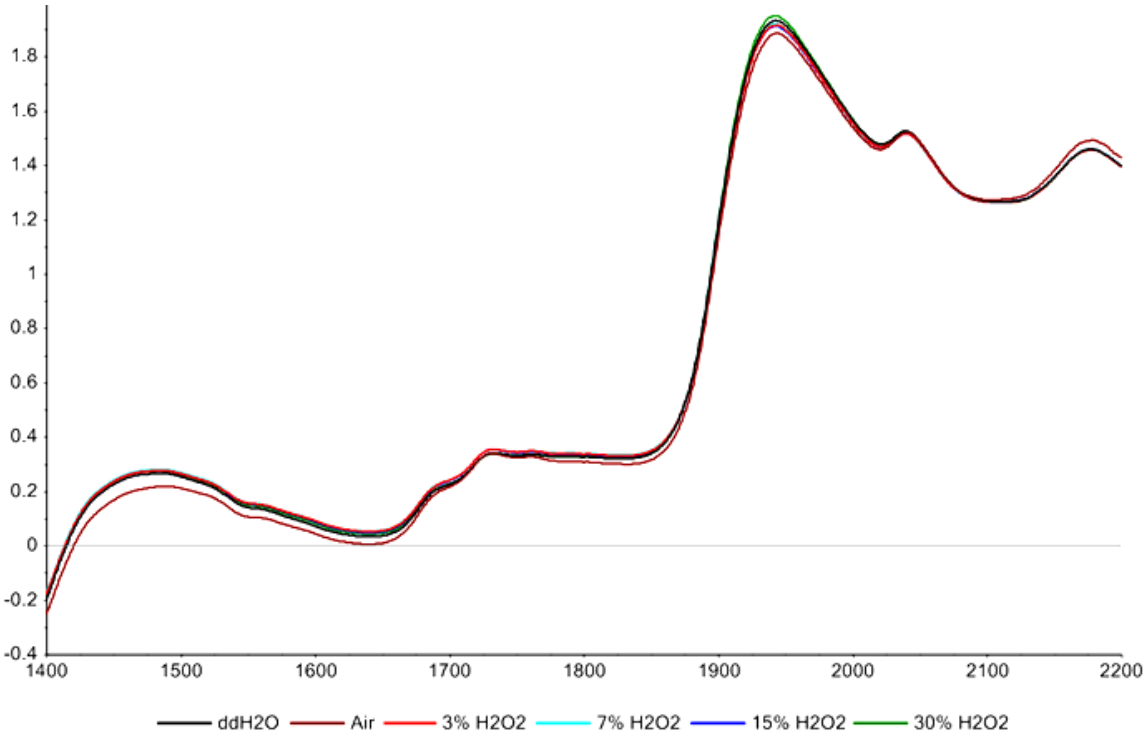


Figure G. 17. SNV-transformed spectra collected from superficial surfaces of the Dilution Set. A slight separation is noted at the 1490nm peak.

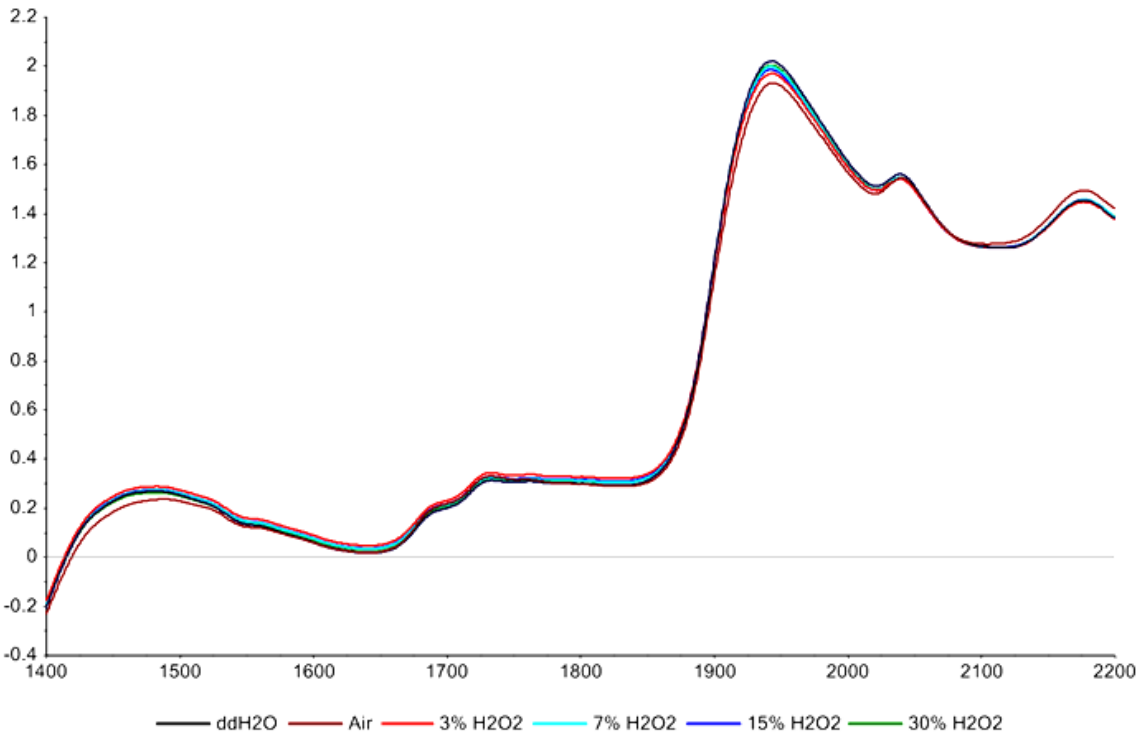


Figure G. 18. SNV-transformed spectra collected from longitudinally cut surfaces of the Dilution Set. A slight separation is noted at the 1490nm peak.

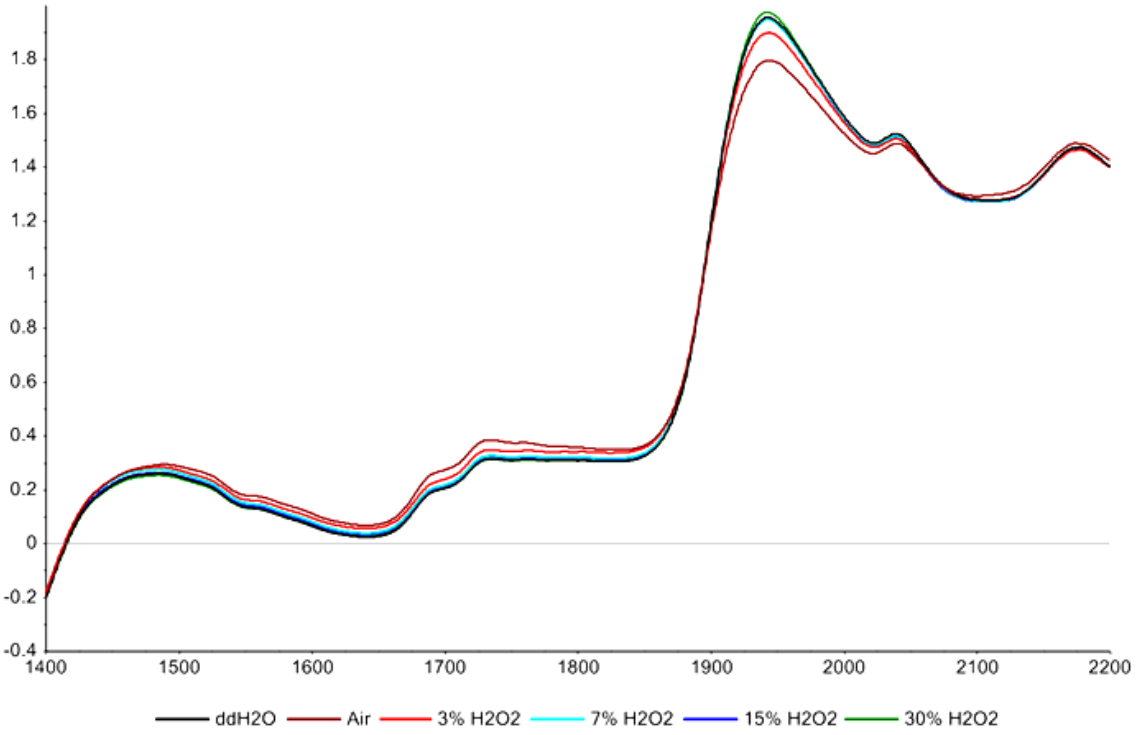


Figure G. 19. SNV-transformed spectra collected from transversely cut surfaces of the Dilution Set. A slight separation is noted at the 1730nm and 1940nm peaks.

APPENDIX H

GLOSSARY OF SELECT TERMS

ASD Viewspec Pro™ version 6.2 (Viewspec): A free software package developed by ASD Inc. for visualization, transformation, and export of spectral data.

Camo The Unscrambler® X version 10.4 (Unscrambler): A non-coding statistical package designed CAMO Software to manipulate large datasets. Unscrambler includes a selection of frequently used preprocessing and statistical modeling methods.

Indico® Pro Spectral Acquisition Software package version 6.3 (IndicoPro): A software package developed by ASD Inc. that is used for the settings and operation of the ASD/PANalytical LabSpec® 4 full range bench spectrometer.

Linear Discriminant Analysis (LDA): A statistical technique that can be used to classify two or more descriptive classes that are each defined by samples with multiple measurable variables which are not correlated. LDA maximizes the separation between group multivariate means while minimizing the variation within each group.

Linear Discriminant Analysis on Principal Components (LDA-PCA): A modification on LDA whereby PCA is used to reduce a large number of inter-correlated variables into a limited number of non-correlated principal components (PC), and LDA is subsequently applied to the PCs.

Principal Component Analysis (PCA): An unsupervised statistical technique that can be used to reduce large, complex systems of highly correlated variables into a limited number of uncorrelated Principal Components (PCs) that best explain the greatest amount of variation. PCs are used to define the axes for plots in a newly created component space. The greatest amount of variation is described by PC1, followed by the second highest in an orthogonal PC2, and so on.

Partial Least Squares Discriminant Analysis (PLSDA): Also known as Projection of Latent Squares Discriminant Analysis. A regression-based classification technique that reduces a large, complex system of highly correlated X-variables into a limited number of latent variables (LV). LVs are then regressed on Y-variables (defined categories) in order to explain the maximum amount of variation in Y.

Postmortem Interval (PMI): The length of time for which an individual has been deceased

Root Mean-Squared Error of Cross-Validation (RMSECV): A measure of total error for a calibration model. RMSECV for a predictive model is calculated from the difference between known training set values and predicted values,

where each individual sample is left out as RMSECV is iteratively calculated.

Savitzky-Golay transformation (SG): A preprocessing technique that can be used for spectral smoothing and removal of scattering effects. SG involves defining a symmetric smoothing window of points on either side of a central point. A derivative is then calculated for the central point by applying a polynomial function. The process is repeated for all points of the spectrum, resulting in a smoothed, derivatized spectrum

Scattering: Scattering occurs when the reflection of light is non-specular. NIR spectra are often affected by additive scattering, which results from variations of path lengths of light into different samples, and/or multiplicative scattering, which is generated from surface variability, including sample differences in particle or grain size and texture

Soft Independent Modeling of Class Analogy (SIMCA): A classification technique whereby a Global PCA classifier is constructed from disjointed local PCA models for each descriptive class. A sample applied to the SIMCA model is classified based on which local PCA best describes the sample.

Standard Normal Variate (SNV): A preprocessing technique that can decrease scattering effects in raw spectra. For a set of spectra, an average spectral value and standard deviation are calculated each, after which SNV-corrected spectra are calculated by subtracting the mean spectral value at each wavelength and then dividing by the standard deviation. The resulting spectra are corrected for scattering and normalized.

Support Vector Machine (SVM): A classification technique based on machine learning which maximizes the distance between two or more groups. Groups of interest are separated by the hyperplane with maximum-margins, or the plane that is the greatest distance from the closest samples in opposing groups (support vectors), providing a boundary which minimizes the number of misclassified samples.

Support Vector Machine on Principal Components (SVM-PCA): A modification on SVM whereby PCA is used to reduce a large number of inter-correlated variables into a limited number of non-correlated PCs, and an SVM classifier is subsequently constructed using the PCs.

REFERENCES

Afara, I., S. Singh, and A. Oloyede. Application of near-infrared (NIR) spectroscopy for determining the thickness of articular cartilage. Medical Engineering and Physics. 35(2013)88-95.

Anyon, R., P. Gilman, and S. LeBlanc. A reevaluation of the Mongollon-Mimbres archaeological sequence. Kiva. 46(1981):209-225.

Bada, J. and P. Helfman. Amino acid racemization dating of fossil bones. World Archaeology. 7(1975):160-173.

Bass, W. 1997. Outdoor decomposition rates in Tennessee. In: "Forensic Taphonomy: the postmortem fate of human remains," edited by W. Haglund and M. Sorg, pp. 181-186. CRC Press, Boca Raton.

Baykal, D., O. Irrechukwu, P. Lin, K. Fritton, R.G. Spencer, and N. Pleshko. Non-destructive assessment of engineered cartilage constructs using near-infrared spectroscopy. Applied Spectroscopy. 10(2010):1160-1166.

Bell, L.S., M.F. Skinner, S.J. Jones. The speed of post mortem change to the human skeleton and its taphonomic significance. Forensic Science International. 82(1996):129-140.

Behrensmeyer, A. Taphonomic and ecologic information from bone weathering. Paleobiology. 4(1978):150-162.

Berg, S. 1963. The determination of bone age. In Methods of forensic science volume II, edited by F. Lundquist, pp.231-252. John Wiley & Sons, London/New York.

Berlett, B. and E. Stadtman. Protein oxidation in aging, disease, and oxidative stress. The Journal of Biological Chemistry. 272(1997):20313-20316.

Boskey, A. and N.P. Comacho. FT-IR imaging of native and tissue-engineered bone and cartilage. Biomaterials. 28 (2007):2465-2478.

Brereton, R. and G. Lloyd. Partial least squares discriminant analysis: taking the magic away. Journal of Chemometrics. 28(2014):213-225.

Brooke, H. "Introduction to Multivariate Analysis: Three Day Theory and Practical Course." Camo Software. Orlando, Florida. November 2015.

Buchan, M.J., G.S. Anderson. Time since death: a review of the current status of methods used in the later postmortem interval. Canadian Society of Forensic Sciences Journal. 34(2001):1-22.

Chih-Chung Chang and Chih-Jen Lin, LIBSVM : a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1--27:27, 2011. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

Child, A. Towards an understanding of the microbial destruction of archaeological bone in the burial environment. Journal of Archaeological Science. 22(1995):165-174.

Clark, M., M. Worrell, and J. Pless. 1997. Postmortem changes in soft tissues. In: "Forensic Taphonomy: the postmortem fate of human remains," edited by W. Haglund and M. Sorg, pp. 151-164. CRC Press, Boca Raton.

Coats, A.W., J.P. Redfern. Thermogravimetric analysis: a review. Analyst. 88(1963):906-924.

Collins, M., C. Nielsen-Marsh, J. Hiller, C. Smith, J. Roberts, R. Prigodich, T. Wess, J. Csapò, A. Millard, and G. Turner-Walker. The survival of organic matter in bone: a review. Archaeometry. 3(2002):383-394.

D'Agostino, R. and H. Russell. 2005. Scree Test. "Encyclopedia of Biostatistics," vol. 7, edited by P. Armitage and T. Colton, pp. 4790-4793. Wiley.

Davies, M. The oxidative environment and protein damage. Biochimica et Biophysica Acta. 1703(2005):93-109.

Dogra, J. Multivariate analyses of near-infrared and UV spectral data. Ph.D. 2009. Baylor University.

Dong, P. and J. Liu. Hyperspectral image classification using support vector machines with an efficient principal component analysis scheme. Foundations of Intelligent Systems, pp131-40. Springer Berlin Heidelberg.

Facchini, F. and D. Pettener. Chemical and physical methods in dating human skeletal remains. American Journal of Physical Anthropology. 47(1977)65-70.

Fisher, A. The use of multiple measurements in taxonomic problems. Annuls of Eugenics. 7(1936):179-188.

Franklin, H. New Dates from Pottery Mound. Maxwell Museum Technical Series No. 7. 2008. Maxwell Museum of Anthropology, University of New Mexico.

Fratzl, P., H. Gupta, E. Paschalis, and P. Roschger. Structure and mechanical quality of the collagen–mineral nano-composite in bone. Journal of Materials Chemistry. 14(2004):2115-2123.

Fratzl, P. and R. Weinkamer. Nature's hierarchical materials. Progress in Materials Science. 52(2007):1263-1334.

Galloway, A. 1997. The process of decomposition: a model from the Arizona-Sonoran desert. In: "Forensic Taphonomy: the postmortem fate of human remains," edited by W. Haglund and M. Sorg, pp. 139-149. CRC Press, Boca Raton.

Gibbs, M.L. What's sex in the east is not necessary sex in the west: citrate, sex, and human skeletal remains. M.A. Thesis 1991. McMaster University.

Gill-King, H. 1997. Chemical and ultrastructural aspects of decomposition. In: "Forensic Taphonomy: the postmortem fate of human remains," edited by W. Haglund and M. Sorg, pp. 93-108. CRC Press, Boca Raton.

Goff, M.L. Early post-mortem changes and stages of decomposition in exposed cadavers. Experimental and Applied Acarology. 49(2009):21-36.

Haglund, W. 1997. Dogs and coyotes: postmortem involvement with human remains. In: "Forensic Taphonomy: the postmortem fate of human remains," edited by W. Haglund and M. Sorg, pp. 367-381. CRC Press, Boca Raton.

Hall, B. 2005. Bones and Cartilage: Developmental and Evolutionary Skeletal Biology. Elsevier Academic Press, San Diego/London.

Halliwell, B. and J. Gutteridge. Free Radicals in Biology and Medicine, 4th ed. 2007. Oxford University Press.

Haskell, N., R. Hall, V. Cervenka, and M. Clark. 1997. On the body: insects' life stage presence, their postmortem artifacts. In: "Forensic Taphonomy: the postmortem fate of human remains," edited by W. Haglund and M. Sorg, pp. 415-448. CRC Press, Boca Raton.

Hawkins, C. and M. Davies. Generation and propagation of radical reactions in proteins. 1504(2001):196-219.

Hedges, R., A. Millard, and A. Pike. Measurements and relationships of diagenetic alteration of bone from three archaeological sites. Journal of Archaeological Science 22(1995):201-209.

Hedges, R. 2002. Bone diagenesis: an overview of processes. Archaeometry 44(3)319-328.

Henßge, C. and B. Madea. Estimation of the time since death in the early post-mortem period. Forensic Science International. 114(2004):167-175.

Hoke, N., A. Grigat, G. Grupe, and M. Harbeck. Reconsideration of bone postmortem interval estimation by UV-induced autofluorescence. Forensic Science International. 228(2013):176e1-176e6.

Howes, J., B. Stuart, P. Thomas, S. Raja, and C. O'Brien. An investigation of model forensic bone in soil environments studied using infrared spectroscopy. Journal of Forensic Sciences. 57(2012):1161-1167.

Hulmes, D. 2008. Collagen diversity, synthesis and assembly. In: “Collagen Structure and Mechanics,” edited by P. Fratzl, pp15-47. Springer Science+Business Media, New York.

Introna, F., G. DiVella, and C.P. Campobasso. Determination of postmortem interval from old skeletal remains by image analysis of luminol test results. Journal of Forensic Sciences. 44(1999):535-538.

Jenson, J. 2005. Introductory Digital Image Processing: A Remote Sensing Perspective. 3rd edition. Prentice Hall, Upper Saddle River.

Jombart, J., S. Devillard, F. Balloux. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BioMed Central Genetics. 11(2010). www.biomedcentral.com/1471-2156/11/94. Accessed 7/15/2016.

Kanz, F., C. Reiter, D.U. Risser. Citrate content in bone for time since death estimation: results from burials with different physical characteristics and known PMI. Journal of Forensic Sciences. 59(2014):613-620.

Kecman, V. Support vector machines: theory and applications. Studies of Fuzziness and Soft Computing. 177(2005): 1-47.

Knight, B. and I. Lauder. Proceedings of the British Association in Forensic Medicine: Practical Methods of Dating Skeletal Remains: A Preliminary Study. Medicine, Science and the Law 7(1967): 205-208.

Knight, B. and I. Lauder. Methods of dating human skeletal remains. Human Biology. 41(1969):322-341.

Linderholm, J., J. Fernandez Pierna, D. Vincke, P. Dardenne, and V. Baeten. Identification of fragmented bones and their state of preservation using near infrared hyperspectral imaging. Journal of Near Infrared Spectroscopy. 21(2013):459-466.

De Maesschalck, R. A. Candolfi, D. Massart, and S. Heuerding. Decision criteria for soft modelling of class analogy applied to near infrared data. Chemometrics and Intelligent Laboratory Systems. 47(1999):65-77.

Manley, M. Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis of biological materials. Chemical Society Review. 43(2014):8200-8214.

Megyesi, M., S. Nawrocki, N. Haskell. Using accumulated degree-days to estimate the postmortem interval for decomposed human remains. Journal of Forensic Sciences. 50(2005):618-626.

Metcalf, J., Z. Xu, S. Weiss, S. Lax, W. Van Treuren, E. Hyde, S. Song, A. Amir, P. Larsen, N. Sangwan, D. Haarmann, G. Humphrey, G. Ackermann, L. Thompson, C. Lauber, A. Bibat, C. Nicholas, M. Gebert, J. Petrosino, S. Reed, J. Gilbert, A. Lynne, S. Bucheli, D. Carter, R. Knight.

Microbial community assembly and metabolic function during mammalian corpse decomposition. Science. 351(2016): 158-162.

Monici, M. Cell and tissue autofluorescence research and diagnostic applications. Biotechnology Annual Review. 11(2005):227-56.

Micozzi, M. 1991. Postmortem Changes in Human and Animal Remains. Charles C Thomas, Springfield, Il.

Mullen, M. 1998. Transformations of other elements. In Principles and Applications of Soil Microbiology, edited by D. Sylvia, J. Fuhrmann, P. Hartel, and D. Zuberer, pp 369-386. Prentice Hall. Upper Saddle River, NJ.

Murray, I., L.S. Aucott, and I.H. Pike. Use of discriminant analysis on visible and near infrared reflectance spectra to detect adulteration of fishmeal with meat and bone meal. Journal of Near Infrared Spectroscopy. 9(2001):297-311.

Naes, T., T. Isaksson, T. Fearn, and T. Davies. 2002. A User-Friendly Guide to Multivariate Calibration and Classification. NIR Publications. Chichester, UK.

Nagy, G., T. Lonrand, Z. Patonai, G. Montsko, I. Bajnoczky, A Marcsik, and L. Mark. Analysis of pathological and non-pathological human skeletal remains by FT-IR spectroscopy. Forensic Science International. 175 (2008):55-60.

Nashelsky, M., P. McFeeley. 2003. Time of death. In: "Handbook of Forensic Pathology," 2nd ed., edited by R.C. Froede, pp. 69-78. College of American Pathologists.

Neis, P., R. Hille, M. Paschke, G. Pilwat, A. Schnabel, C. Neiss, and H. Bratzke. Strontium90 for determination of time since death. Forensic Science International. 99(1999):47-51.

Nehring, J., M. Sheridan, W. Funk, and G. Anderson. The possibility of postmortem bacterial transmigration. Archives of Pathology. 93(1972)266-270.

Parks, C.L. A study of the human decomposition sequence in Central Texas. Journal of Forensic Science. 56(2011):19-22.

Paschalis, E.P., F. Betts, E. DiCarlo, R. Mendelsohn, and A.L. Boskey. FTIR microscopic analysis of normal human cortical and trabecular bone. Calcified Tissue International. 61 (1997):480-486.

Patonai, Z., G. Maasz, P. Avar, J. Schmidt, T. Lorand, I. Bajnoczky, and L. Mark. Novel dating method to distinguish between forensic and archaeological human skeletal remains by bone mineralization indexes. International Journal of Legal Medicine. 127(2013):529-533.

Perper, J. 2006. Chapter III. Time of death and changes after death. Part 1. Anatomical considerations. In: "Spitz and Fisher's Medicolegal Investigation of Death. Guidelines for the

Application of Pathology to Criminal Investigations," edited by W. Spitz and D. Spitz. 4th Ed. Charles C. Thomas, Springfield, IL.

Quickendon, T.I., J.I. Creamer. A study of common interference with the forensic luminol test. Luminescence. 16(2001):295-298.

Race, G., E. Fry, J. Matthews, M. Wagner, J. Martin, and J. Lynn. Ancient Nubian human bones: a chemical and ultrastructural characterization including collagen. American Journal of Physical Anthropology. 28(1968):157-162.

Raja, S., P.S. Thomas, B.H. Stuart, J.P. Guerbois, and C. O'Brien. The estimation of pig bone age for forensic application using thermogravimetric analysis. Journal of Thermal Analysis and Calorimetry. 98(2009):173-176.

Ramsthaler, F., K. Kreutz, K. Zipp, M.A. Verhoff. Dating skeletal remains with luminol-chemiluminescence. Validity, intra- and interobserver error. Forensic Science International. 187(2009):47-50.

Ramsthaler, F., S.C. Ebach, C.G. Birngruber, M.A. Verhoff. Postmortem interval of skeletal remains through the detection of interosseal hemin traces. A comparison of UV-fluorescence, luminol, Hexagon-OBTI®, and Combur® tests. Forensic Science International. 209(2011):59-63.

Reeves, N. Taphonomic effects of vulture scavenging. Journal of Forensic Sciences. 54(2009):523-528.

Rinnan, Å., F. van den Berg, and S. Engelsen. Review of the most common pre-processing techniques for near-infrared spectra. Trends in Analytical Chemistry. 28(2009):1201-1222.

Ritter N. Jan 2007. Missing persons and unidentified remains: the nation's silent mass disaster. NIJ Journal 256:1-7 [internet]. [accessed 9/25/2013]. Available from: http://www.ojp.usdoj.gov/nij/journals/256/missing-persons.html

Rodriguez, W. 1997. Decomposition of buried and submerged bodies. In: "Forensic Taphonomy: the postmortem fate of human remains," edited by W. Haglund and M. Sorg, pp. 459-467. CRC Press, Boca Raton.

Schrag, B., T. Uldin, P. Mangin, and P. Froidevaux. Dating human skeletal remains using a radiometric method: biogenic versus diagenic 90Sr and 210Pb in vertebrae. Forensic Science International. 220(2012):271-278.

Schrag, B., T. Uldin, P. Mangin, F. Bochad, and P. Froidevaux. Dating human skeletal remains using 90Sr and 210Pb: case studies. Forensic Science International. 234(2014):190e1-190e6.

Schwarcz, H.P., K. Agur, and L. Meadows Janz. A new method for determination of postmortem interval: citrate content of bone. Journal of Forensic Sciences. 55(2010):1516-1523.

Shackleford, J. and R. Wyckoff. Collagen in fossil teeth and bones. Journal of Ultrastructure Research. 11(1964):173-180.

J. Shackleford. The ultrastructure of Mississippian and Archaic Indian bones from various soil and drainage conditions. American Journal of Physical Anthropology. 25(1966):291-198.

Siesler, H.W., Y. Ozaki, S. Kawata, and H.M. Heis (editors). 2002. Near-infrared spectroscopy. Principles, instruments, applications. Wiley-VCH/Weinheim.

Shipman, P., A. Walker, and D. Bichell. 1985. The human skeleton. Harvard University Press, Cambridge MA.

Soloman S., and E. Hackett. 1996. Setting Boundaries between Science and Law: Lessons from Daubert v. Merrell Dow Pharmaceuticals, Inc. Science, Technology & Human Values. 21(2):131-156.

Spahn, G., H. Plettenberg, E. Kahl, H.M. Klinger, T. Muckley, and G.O. Hofmann. Near-infrared (NIR) spectroscopy. A new method for arthroscopic evaluation of low grade degenerated cartilage lesions. Results of a pilot study. BMC Musculoskeletal Disorders. 8(2007):47.

Spahn, G., H. Plettenberg, H. Nagel, E. Kahl, H.M. Klinger, T. Muckley, M. Gunther, G.O. Hofman, J.A. Mollenhauer. Evaluation of cartilage defects with near-infrared spectroscopy (NIR): an ex vivo study. Medical Engineering and Physics. 30(2008):285-292.

Speller, C.F., K.L. Spalding, B.A. Buchholz, D. Hildebrand, J. Moore, R. Mathewes, M.F. Skinner, and D.Y. Yang. Personal identification of cold case remains through combined contribution from anthropological, mtDNA, and bomb-pulse dating analyses. Journal of Forensic Sciences. 57(2012):1354-1360.

Stadtman, E., and R. Levine. Free radical-mediated oxidation of free amino acids and amino acid residues in proteins. Amino Acids. 25(2003):207-218.

Stothard, P. The Sequence Manipulation Suite: JavaScript programs for analyzing and formatting protein and DNA sequences. Biotechniques 28(2000):1102-1104.

Swift, B. Dating human skeletal remains: investigating the viability of measuring the equilibrium between 210Po and 210Pb as a means of estimating the post-mortem interval. Forensic Science International. 98(1998):119-126.

Swift, B, I. Lauder, S. Black, and J. Norris. An estimation of the post-mortem interval in human skeletal remains: a radionuclide and trace element approach. Forensic Science International. 117(2001):73-87.

Swift, B. 2006. The timing of death. In Essentials of Autopsy Practice: Current Methods and Modern Trends. Edited by G. Rutty, pp. 189-214. Springer-Verlag, London.

Thomas, D.B., C.M. McGoverin, A. Chinsamy, and M. Manley. Near infrared analysis of fossil bone from the Western Cape of South Africa. Journal of Near Infrared Spectroscopy. 19(2011):151-159.

Trueman, C., and D. Martill. 2002. The long-term survival of bone: the role of bioerosion. Archaeometry 44(3):371-382.

Ubelaker, D. 1997. Taphonomic applications in forensic anthropology. In: "Forensic Taphonomy: the postmortem fate of human remains," edited by W. Haglund and M. Sorg, pp. 77-90. CRC Press, Boca Raton.

Ubelaker, D.H. Radiocarbon analysis of human remains: a review of forensic applications. Journal of Forensic Sciences 59(2014):1466-72.

The UniProt Consortium. Nucleic Acids Res (2017) 45 (D1): D158-D169. DOI: https://doi.org/10.1093/nar/gkw1099.

Vass, A.A., S.A. Barshick, G. Sega, J. Caton, J.T. Skeen, J.C. Love, J.A. Synstelien. Decomposition chemistry of human remains: a new methodology for determining the postmortem interval. Journal of Forensic Sciences. 47(2002):542-553.

Vass, A.A. The elusive post-mortem interval formula. Forensic Science International. 204(2011)34-40.

Villanueva, P.E., F. Girela, and M. Castellanos. The application of differential thermal analysis and thermogravimetric analysis to dating bone remains. Journal of Forensic Sciences. 21(1976):822-830.

Vincke, I., R. Miller, E. Stassart, M. Otte, P. Dardenne, M. Collins, K. Wilkinson, J. Stewart, and V. Baeten. Analysis of collagen preservation in bones recovered in archaeological contexts using NIR hyperspectral imaging. Talanta. 125(2014):181-188.

Wold, S. Pattern recognition by means of disjoint principal components models. Pattern Recognition. 8(1975):127-139.

Wold, S., K. Esbensen, and P. Geladi. Principal component analysis. Chemometrics and Intelligent Laboratory Systems. 2(1987):37-52.

Wold, S., M. Sjöström, and L. Ericksson. PLS-regression: a basic tool of chemometrics. Chemometrics and Intelligent Laboratory Systems. 58(2001):109-130.

Yoshino, M., T. Kimijima, S. Miyasaka, H. Sato, S. Seta. Microscopical study on estimation of time since death in skeletal remains. Forensic Science International. 49(1991):143-158.

Zakian, C., I. Pretty, and R. Ellwood. Near-infrared hyperspectral imaging of teeth for dental caries detection. Journal of Biomedical Optics. 14.6(2009):064047-064047.