University of South Carolina
Scholar Commons

Summer 2019

Non-Destructive Determination of Magnetic Audio Tape Degradation for Various Tape Chemistries Using Spectroscopy and Chemometrics

Alyssa Marie Abraham

Follow this and additional works at: https://scholarcommons.sc.edu/etd

Part of the Chemistry Commons

www.manaraa.com

Non-Destructive Determination of Magnetic Audio Tape Degradation for Various
Tape Chemistries Using Spectroscopy and Chemometrics

by

Alyssa Marie Abraham

Bachelor of Science
University of New Haven, 2014

_____

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Chemistry

College of Arts and Sciences

University of South Carolina

2019

Accepted by:

Stephen L. Morgan, Major Professor

Susan D. Richardson, Committee Member

Michael L. Myrick, Committee Member

Claudia Benitez-Nelson, Committee Member

Cheryl L. Addy, Vice Provost and Dean of the Graduate School

DEDICATION

This dissertation is dedicated to my family and friends, without whom this would not be possible. Thank you for all of your support.

ACKNOWLEDGEMENTS

First I would like to thank my family, especially my parents, Edward and Wendy, for supporting me and giving me the opportunities that allowed me to achieve this accomplishment. Throughout my entire life you have been my biggest supporters. I could not have been successful without your loving support. A special thank you goes out to my siblings, Nicholas, Samantha, and Serena, who gave me a reprieve from the stressors of school. To Samantha, my twin, thank you for your support and help, even when I was stubborn while trying to learn calculus.

To my teachers and professors along the way, thank you for instilling in me a love of learning. From my high school teachers, especially Ms. Mo (Moriarty-Hawkins), to my college professors including my advisor, Dr. Nancy O. Savage, thank you for fueling my passion for learning. Thank you for believing in me. It is because of teachers like you that I had the confidence and skills to survive graduate school. I would also like to thank my advisor, Dr. Stephen Morgan for accepting me into his lab and providing the opportunity to work on such an amazing project. I have learned so much. My committee members, Dr. Susan Richardson, Dr. Michael Myrick, and Dr. Claudia Benitez-Nelson, also deserve recognition for their guidance and support.

Finally, I would like to thank my friends and peers, both from home and in graduate school. I could not have done this without your support. To Shannon Smith, Nicole Dib, Jess Cornell, and Delaney Sears, thank you for always making time for me, even with my crazy graduate school schedule. A special thanks to the Morgan lab: Brianna Cassidy, Zhenyu Lu, Nathan Fuenffinger, Kaylee McDonald, Molly Burnip, Nilmini Ratnasena, and Dayla Rich for your academic input and friendship along the way. I would also like to thank Amy King, Allison Rice, Meagan Smith, Samuel Putnam, and Hannah Liberatore for providing me a social life outside of school.

ABSTRACT

This dissertation focuses on creating a non-destructive method for the identification of degraded magnetic audio tapes using attenuated total reflectance Fourier transform infrared spectroscopy (ATR FT-IR) and chemometrics. The primary recording medium during the second half of the twentieth century was magnetic tapes. These tapes hold some of the world's modern cultural history. Unfortunately, the majority of these tapes were made with a polyester urethane (PEU) binder, which has shown to degrade through hydrolysis of the ester component, called sticky shed syndrome (SSS). As the name implies, degraded tapes will stick and shed onto player guides and heads during playback. This can irreversibly destroy the data, since the binder layer houses the magnetic particles responsible for the content. Therefore, libraries, institutions, and archives are digitizing their collections. The first half of this manuscript focuses on creating a model that can accurately determine degradation using a non-destructive approach for a mixture of tape manufacturers and models. Eleven different brand/model tapes were used, acquired from three different sources. Principal component analysis (PCA), followed by quadratic discriminant analysis (QDA) was used for classification. This was successfully performed for 95.19% of the calibration set consisting of 154 tapes and 97.79% of the test set, containing 52 tapes.

The second half of this manuscript focuses on applying this model to external tape test sets that were not used in the creation of the model. It was determined that the different binders and additives included during the manufacturing process would affect the model. Therefore, a new calibration model was created from the original 154-tape set. The new model was validated yielding a 94.64% accuracy. It was tested with the original 52-tape test set, correctly identifying the degradation status for 95.58% of those tapes. The model was then externally validated with tapes acquired from the Library of Congress (LC) and University of Maryland (UMD). The model successfully identified 90.79% of the LC tapes and 78.85% of the UMD tapes. The outcome of this work will greatly enhance preservationists' goal of triaging magnetic audio tape before digitization.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ATR...Attenuated Total Reflectance

FF ....Fast Forward

FT-IR.. Fourier Transform Infrared Spectroscopy

HCA .... Hierarchical Cluster Analysis

LC ....Library of Congress

MBRS ....Motion Picture Broadcasting and Recorded Sound Division

MC .... Mean Centering

MIRC... Moving Image Research Collection

PC...Principal Component

PCA .... Principal Component Analysis

PET.. Polyethylene Terephthalate

PEU .... Polyester Urethane

QDA... Quadratic Discriminant Analysis

RW..Rewind

SG...Savitzky-Golay

SM ....Smithsonian Institution

SNV ....Standard Normal Variate

SSS...Sticky Shed Syndrome

UMD..University of Maryland

USC ....University of South Carolina

WPT...Wisconsin Public Television

CHAPTER 1

A ROBUST MODEL FOR IDENTIFYING VARIOUS BRANDS OF
DEGRADED POLYESTER-URETHANE REEL-TO-REEL
MAGNETIC AUDIO TAPES USING ATTENUATED TOTAL
REFLECTANCE FOURIER TRANSFORM INFRARED
SPECTROSCOPY AND CHEMOMETRICS

1. ABSTRACT

Magnetic audio tapes were extensively used as the primary recording medium in the late twentieth century and remain an invaluable source of the world's modern cultural history. The majority of these tapes are composed of a polyester-urethane binder, placing the tape at risk of being degraded by hydrolysis. Many cultural heritage institutions are digitizing their tape collections, to preserve invaluable information for future generations. Although this process requires visually examining magnetic tapes to identify degraded tapes before playback, the accepted test for degradation has been playability testing performed by an audio engineer. However, this technique has the potential to damage tape surfaces and/or the recorder heads since degraded tapes stick and shed onto playback equipment.[1, 2]

Combining attenuated total reflectance Fourier transform infrared spectroscopy (ATR FT-IR) with multivariate statistics provides a fast, objective, and non-destructive method to assess a tape's playability/degradation status. This model classified magnetic audio tapes from the Library of Congress (LC) with 92-94% accuracy, but showed poor separation with tapes from sources other than the Library of Congress.[3] Different manufacturing brands/models of tapes (*e.g.*, Ampex 406, or 3M 808) have unique chemical formulations, which potentially influence assessment of degradation. The present research strives to create a robust model for the playability status of eleven diverse brands and models of tapes. ATR FT-IR spectroscopy spectra were acquired and compared to playability results. Principal component analysis (PCA) and quadratic discriminant analysis (QDA) were used to reduce the dimensionality of the projected spectral data to maximize the explained variance in an easily visualized space. As spectra from tapes of different brands and models were included, separation due to playability became apparent. Modeling outcomes from the current study performed with 95-97% accuracy for prediction of tape playability, despite containing a heterogeneous collection of tapes acquired from different collections and tape manufacturers.

2. INTRODUCTION

Audio recordings are one of the primary media for preserving the world's recent modern history. In the United States, it is estimated that there are over 21.5 million magnetic audio tapes in archival storage.[4] Cultural heritage institutions house the majority of these tapes and digitize their collections to

www.manaraa.com

preserve information for future generations. Digitization requires playing the tape to retrieve the analog signal and transfer audio content to a new media.[2] Unfortunately, a significant portion of magnetic audio tapes are at risk of being destroyed by ongoing chemical and physical degradation, known as sticky shed syndrome (SSS), which threatens the playability and integrity of the recordings. As the term SSS implies, degraded tapes have a tendency to stick and shed onto playback equipment causing irreversible data loss and downtime for cleaning and repair of playback equipment. Such degraded tapes are therefore at risk of being destroyed during any playing or digitization process.

Tapes are composed of two or three main layers: a magnetic layer, a substrate, and an optional back coating.[5, 6] The magnetic layer is the top most layer, and is comprised of magnetic particles, usually Fe_2O_3 or CrO_2, embedded in a binder that is 2-4 μm thick.[1, 7] This layer physically contacts player heads onto which sticky shed may accumulate for tapes exhibiting extensive damage. The magnetic layer adheres to the tape via a binder. Most magnetic audio tapes from the 1960s to 1980s were made with a polyester urethane (PEU) binder.[1, 7] Although tapes are composed of many different chemicals and layers, the degradation of PEU binders by hydrolysis is widely acknowledged as a significant cause of sticky shed syndrome.[2] When esters, such as PEU, contact water, hydrolysis causes the transformation into carboxylic acid and alcohol as seen in Figure 1.1.

Besides the magnetic particles, the binder may also contain additives such as lubricants for friction reduction, dispersants for better metal-particle

3

distribution, and abrasives to keep player heads clean.[1, 2, 5] The magnetic layer is the least stable layer since it is exposed to the most mechanical and environmental stress, making it exceptionally prone to degradation.[8, 2] The magnetic layer also contains the data, recorded in the magnetic particles, making it exceptionally important to not play a degraded tape so that the content can be digitized and preserved. To retain the physical integrity of the tape, typically a polyethylene terephthalate (PET), 35 µm thick substrate is added during manufacturing.[2, 9, 6] Some tapes consist of a third layer of back-coating,1-2 µm thick, to dissipate static charge during playing.[1, 6]

The standard technique for identifying degraded tapes has been a visual inspection followed by playability testing on a high-quality, vintage tape player, a process which involves an audio engineer playing the tape to determine the condition. Unless a tape is extremely degraded, visual inspection usually does not yield probative information on the tape's degradation status. Playability testing unfortunately can cause permanent damage to the tape and may not be practical for tapes suspected of being degraded due to poor storage conditions, *e.g.*, if the tape has been exposed to intense humidity and heat.

A degraded tape can be temporarily restored through baking, a process which requires placing the tape in a low temperature oven, about 50°C, for over eight hours.[9, 10] However, this process can damage a tape if not done correctly, and is only temporary since degradation can recur within weeks of treatment.[9, 11] Due to concerns with the integrity of tapes and time consumed during baking, it is

4

not feasible to bake all 21.5 million tapes in the U.S. Therefore, only those that have started to degrade should be put through this process.

Because magnetic audio tapes may contain valuable historical information, a non-destructive technique is needed to triage the tapes before digitization. Hobaica, *et al.* detected magnetic tape degradation using attenuated total reflectance Fourier transform infrared spectroscopy (ATR FT-IR) and achieved a ~70% classification accuracy by visual identification of degradation products in the spectra.[9] Hobaica, *et al.* isolated features indicative of degradation at 1730 cm^{-1}, 1364 cm^{-1}, 1252 cm^{-1}, and 1138 cm^{-1} which correspond to the C=O carbonyl stretching vibration, CH$_2$ wag, and C-O stretch from the carboxylic acid and alcohol, respectively. It was also determined that a shift occurs from 1693 cm^{-1} to 1730 cm^{-1} for the C=O carbonyl stretching vibration due to degraded tapes having more free carbonyl groups, while playable tapes have hydrogen bonded carbonyls.[9] While carbonyls were indicated as having the greatest influence in determining degraded from non-degraded tapes, IR spectra typically consist of thousands of wavenumbers making it possible that other features are also involved in differentiating the two groups. Visual inspection is time consuming and may miss these subtle, but significant, degradation markers.

ATR FT-IR with chemometrics can reveal the differences between playable and non-playable tape spectra that visual inspection might miss. Our previous paper demonstrated that ATR FT-IR and multivariate statistics yielded reliable classification accuracies of 91.30% and 93.78%, respectively for playable

5

and non-playable tapes in the calibration set, and classification accuracies for playable and non-playable test sets were 82.00% and 92.31%, respectively.[3] Although the former model worked well for predicting the playability of the LC test tapes, when analyzing tapes from the University of Maryland (UMD) (College Park, MD), the model did not enable acceptable playability classification for these new tapes, as evident in the PCA plot in Figure 1.2. Unlike the original test set used by Cassidy, *et al.*, there is no clear separation between the UMD playable and UMD non-playable tapes.

We hypothesize that this discrepancy is due to different chemical formulations between the UMD tapes and tapes used to create the previous model (LC). The tape sets could have been composed of different tape brands and/or formulations.[3, 12] Changes in process optimization could also cause frequent tape formulation changes within the same model, as well as between the same batch of tapes.[6] This may have affected the degradation assessment as well, since the model was based on chemical differences between the previous playable and non-playable groups in the LC tape set.[3] Many companies have protected their product formulation information.[6] Even with this information, manufacturer labels containing the brand and model are usually only found on the box and not the reel/hub or tape itself. Common rehousing practices of placing the tape on any open reel and in any box would also yield the information unreliable. This makes formulation information difficult to obtain and brand/model information can only be assured with unopened boxes of tape.[3]

6

For the present study, sources of variability between known, unopened, tape brands and models were evaluated using hierarchical cluster analysis (HCA) with Ward's method.[12, 13] The total variance within each cluster was based on spectra grouping by minimized squared Euclidean distance in a dendrogram. The higher the similarity between tape spectra, the closer they cluster together, as shown in Figure 1.3. Dissimilar tapes are located further from one another, and connected by longer branches. Tapes in the two main groups (labeled 1 and 2) were closely examined in an attempt to determine the primary cause of separation. When each group's spectra were averaged and compared, it was apparent that the lower cluster in the dendrogram resembled the LC classification tapes, made with a PEU binder.[12] The first cluster consisted of tapes made in the early 1960s, when manufacturers used acetate instead of PEU binders.

The results also showed a similarity based on manufacturer, seen by the labeled brand and model information, which may be due to slight formulation changes in the binder, as well as additives used by different manufacturers that can be detected in the IR region studied.[12] We conclude that the tapes in the LC collection were of similar formulation, since the test set behaved similarly to the calibration set, with variation due to playability being observed. However, the University of Maryland (UMD) tapes may belong to a cluster separated by formulation differences from the LC tapes, since the correct degradation status was not obtained with the LC model. Thus, an unknown tape would have to first be assigned a formulation group based on its brand/model before playability status is assigned.

The objective of this study was to create a robust model that could be employed to triage reel-to-reel magnetic audio tapes by using ATR FT-IR spectroscopy with multivariate statistics to (1) determine if subgroups of various tape brands/models are first needed by grouping chemistries or formulations of either individual tape brands/models or groups of similar tapes, then (2) determine the playability or degradation classification for each type or group of tape.

3. EXPERIMENTAL

3.1 Sample Selection

Due to the common rehousing practice of tapes being placed on the wrong reel or in the wrong box, a call for unopened tapes was sent to audio library conservators, list serves, and organizations in the tape/archive community. Since these tapes were never opened, they did not contain any recorded information. Over 700 tapes were collected from the Library of Congress (LC), Wisconsin Public Television (WPT), and the Smithsonian Institution (SM). The current study focuses on various brand and model tapes consisting of polyester polyurethane (PEU) that were deemed at risk for degradation due to sticky shed syndrome. All tapes were stored and analyzed at the Moving Image Research Collection (MIRC) at the University of South Carolina (USC), in a temperature and humidity-controlled room.

3.2 Playability Testing

The standard method for determining if a tape is degraded is to subject it to playability testing on vintage tape players. This method is subjective since an engineer has to manually play the tape and categorize a tape based on a list of characteristics. Playability testing of our tapes was performed using a vintage Scully 280 tape player (Scully Recording Instruments, Bridgeport, CT). Each tape went through a playback process consisting of alternating between playing the tape for ten seconds, two seconds of rewinding (RW) and fast forwarding (FF) for three iterations, then toggling between FF and RW to slowly advance the tape a third of the way to the end, before playing the tape for five seconds. This was repeated until the tape fully transferred from one reel to another over six stationary guides, with the read and recording heads removed. If at any point the tape exhibited one of the following characteristics, it was considered non-playable, (a) friction between the tape and player guides caused the tape to slow or stop; (b) the tape squealed; (c) the tape exhibited slow recovery between FF and RW transitions; (d) or significant tape material sloughed onto the player guides.[3] A small amount of shed tape was accepted without classifying the tape as non-playable. A tape was considered playable if it exhibited none of the aforementioned issues and played in a quiet, smooth manner.[3] Unfortunately, there is no empirical guidelines to determine the level of friction, slow recovery, or amount of sloughed tape material to deem a tape non-playable. Therefore, it is up to the audio engineer to make that determination, making the technique subjective and less reproducible.

3.3 Infrared Analysis

ATR FT-IR analysis was performed using a Nexus 670 FT-IR with Omnic® version 8.2 (Thermo-Nicolet, Madison, WI) equipped with a deuterated triglycine sulfate detector (DTGS) detector and a Thunderdome™ ATR (Thermo Spectra Tech, Inc., Shelton, CT), with a germanium crystal and an incident angle of 45°. The setup, as described in our previous paper, consisted of a Mylar™ film, ~0.1 mm thick, placed over the tape for protection followed by a steel plate, ~1 mm thick, to disperse the pressure.[3] Replicate spectra were obtained by collecting ten spectra of the magnetic side of the tape at ~50 cm and ten spectra at ~100 cm from the beginning of the tape for a total of 20 replicate spectra per tape. The spectral range consisted of 4000 cm^{-1} to 600 cm^{-1}, with 32 scans at 4 cm^{-1} resolution.

3.4 Data Analysis

All preprocessing and multivariate modeling was performed using MATLAB version 9.5, R2018b, (The MathWorks, Natick, MA). Spectra were truncated from 1750 cm^{-1} to 950 cm^{-1} (a total of 416 wavenumbers) to correspond with the spectral region descriptive of the degradation status for magnetic audio tapes.[9, 3] Preprocessing consisted of smoothing with a 13-point Savitzky-Golay (SG) fourth order polynomial.[14] The standard normal variate transform (SNV) was applied to remove multiplicative interferences caused by scatter and particle size. The SNV transform involves subtracting the mean absorbance from each spectral intensity, then dividing by the standard deviation of the absorbance for each individual spectrum as done previously.[15, 3] The data

10

were randomly split into a calibration set containing 154 tapes (3,080 spectra) and a test set containing 52 tapes (1,040 spectra). Finally, mean centering was performed by subtracting the average absorbance of a single feature or wavenumber (calculated from the 3,080 spectra calibration set) from all observations of that feature in both the calibration and test set.[12, 3]

Principal component analysis (PCA) was used to reduce the dimensionality of the tape degradation data, with the first principal component (PC) capturing the variables comprising the highest amount of variation in the data, followed by the second PC and so on.[16, 17, 18] The first five PCs were used to build the model, encompassing 95.8% variance and capturing the significant variation of the data set.

Five-fold cross-validation was used to estimate the accuracy of the model. The calibration set was randomly divided into five different groups. The first group was designated as the test set, and the other four groups were used to train the model, after which the model predictions were evaluated with the test set. This procedure was repeated, cycling through with each of the five groups acting as a test set, with the validation accuracy computed by the number of spectra misclassified.

Quadratic discriminant analysis (QDA), a supervised classification technique was performed, using the playability test outcomes as responses to train the model. The first five PCs determined from the ATR FT-IR spectra for the 154-tape calibration set were used to find the best projection to separate the classes using a quadratic decision boundary. QDA assumes unequal covariance

11

matrices and that the data is normally distributed for playable and non-playable classes. The Mahalanobis distance from each sample to the class centroids for the playable and non-playable groups was calculated. Samples were then assigned to whichever group's centroid was closest [16, 19] The preprocessed 1,080 spectra from the 52-tape test set were then projected into the same five-dimensional space as the calibration set. The spectra were classified as belonging to closest calibration group using the minimum Mahalanobis distance classifier to that group's centroid.

4. RESULTS AND DISCUSSION

4.1 Playability Testing

Playability testing was performed on all 154 tapes used to train the model, yielding 85 playable tapes and 69 non-playable tapes. Of the 52 tapes used in the test set, 27 were determined to be playable while the remaining 25 were non-playable.

4.2 Infrared Analysis

ATR FT-IR spectra were obtained from the 154 training set tapes by acquiring ten replicate spectra at both 50 cm and 100 cm from the beginning of the tapes. Figure 1.4 (top) shows the averaged spectra of the baseline corrected ATR FT-IR data for playable (denoted blue) and non-playable (denoted red) tapes assigned by playability testing. The bottom of Figure 1.4 shows the same averaged spectra after preprocessing by SG smoothing, SNV transform, and mean centering. While differences between the spectra can be seen in the raw

12

data, preprocessing enhances the differences and reveals additional effects. The biggest difference between playable and non-playable spectra is seen in the data around 1170 cm^{-1}, which corresponds to the C-O stretch of the alcohol peak due to the degradation of the ester from hydrolysis. Subtle differences, such as the peak intensity and shoulder found in degraded tapes at 1730 cm^{-1} and 1257 cm^{-1} correspond to the C=O stretch and C-O stretch of the carboxylic acid. These peaks are more pronounced after preprocessing, which enhances the ability of machine learning compared to visual inspection alone.[3, 9, 12]

4.3 Principal Component Analysis

PCA was useful for exploring the variability in the 154 tape, 3,080 spectral calibration set. Due to the previously described problems of introducing UMD tapes to the LC model, the present calibration set was built with eleven different tape brands/models acquired from three different sources, encompassing both playable and non-playable tapes as seen in Table 1.1. Due to the previous exploratory study using HCA described earlier, we expected to see groupings based on the brand and model of the tapes, *e.g.*, such that all Emtec tapes would be in one group, all BASF tapes in another, with Ampex 406 separated from those, and so on. Instead, the model is seen to classify based on playability status**,** as well as manufacturer information**.** Figure 1.5 shows the PCA projections for the calibration set, with each tape brand/model and institution as a different marker/color. Markers outlined in black indicate non-playable data points. Separation based on the acquisition source of the tape is not seen in the Figure. While storage conditions play a role in the progress of tape degradation,

13

whether storage conditions (*e.g.*, humidity and temperature) add another factor to the separation is not apparent.

The first PC accounts for 51.35% of the variability in the data set. Although assigning a definitive source for this variability is speculative, a combination of causes from varying additives in the binder layer used by various manufacturers, such as lubricants, plasticizers, and chemicals to reduce static charge and clean tape heads may be present.[8] The second PC accounts for 28.58% variability in the spectra and is attributable to chemical differences between playable and non-playable tapes. This is evident in the PCA plot in Figure 1.6: red circles indicate non-playable spectra, and blue circles represent playable spectra. As seen in the plot, there are some playable tapes in the positive PC 2 direction where the non-playable tapes reside, as well as some non-playable tapes in the negative PC 2 direction with the playable spectra. However, this does not definitively indicate that these spectra were misclassified since Figure 1.6 is 2-dimensional, but five PCs were used in the classification model.

4.4 Principal Component Loading Analysis

To determine chemical differences attributing to the separation between playable and non-playable spectra in PC 2, the eigenvector components for the second PC loading were examined (Figure 1.7). Non-playable spectra were found in the positive region, where eigenvector components correlate to wavenumbers found in non-playable spectra. The most influential peaks, indicated by the arrows pointing to the positive peaks in Figure 1.7, correspond to functional groups found in the PEU hydrolysis degradation products, including

14

the peak at 1730 cm^{-1} of the C=O ester stretch. This peak was found in both playable and non-playable tape spectra but was more intense in the non-playable spectra. Hobaica, *et al.* reported that the carbonyl in playable tapes centered closer to 1700 cm^{-1}.[9] The present results show a greater intensity of the carbonyl peak in the non-playable tape spectra due to the carbonyl from the carboxylic acid and ester. This effect was due to more free carbonyl groups present in sticky tapes than their playable counterparts.[9] The peak at 1257 cm^{-1} represents the C-O stretch from the carboxylic acid, while the peaks at 1170 cm^{-1} and 1140 cm^{-1} signify the C-O-H tertiary alcohol stretch.

The negative loading direction of PC 2 was due to eigenvector components indicative of wavenumbers found in playable spectra. This corresponds to the projected playable data points in the PCA plot in Figure 1.6, which are found in the negative PC 2 region. The peak at 1530 cm^{-1} was due to the N-H amide bend from the polyurethane segment of PEU, while the 1225 cm^{-1} and 1111 cm^{-1} peaks were the C-C-O and O-C-C ester components, respectively. These peaks indicate components that are found in the non-degraded PEU binder.

4.5 Quadratic Discriminant Analysis

A calibration model was created using QDA with the first five PCs and validated with five-fold cross-validation and the Mahalanobis distance classifier. A visual representation of the model can be seen in Figure 1.8, which shows the correctly classified non-playable and playable tapes as red and blue dots,

15

respectively. The purple and cyan data points correspond to the falsely predicted non-playable and playable tapes, respectively.

The confusion matrix in Table 1.2 shows that 2,931 of the total 3,080 spectra (95.19%) in the calibration set were correctly classified. Of the 1,380 spectra from the 69 tapes assigned as non-playable through playability testing, 1,373 (99.49%) were correctly classified as degraded. The seven misclassifications arose from seven different tapes. Two SM 3M 996 tapes, and one SM Ampex 478 and WPT Ampex 406 tapes all had their first replicate, sampled ~50 cm from the beginning of the tape misclassified. The remaining three incorrectly predicted playable spectra were the last replicate sampled ~100 cm from the beginning of a LC Ampex 456, SM 3M 996, and SM 3M 808 tape.

Of the 1,700 spectra from the 85 tapes assigned as playable, 1,559 (91.71%) were correctly classified with 141 spectra from 9 different tapes being misclassified as not degraded. Six tapes were fully misclassified with 20 replicates each being identified as non-playable through the QDA model. Another tape had nineteen spectra identified as degraded, while the remaining two misclassified spectra were from two additional tapes. Overall, the most often misclassified tape was 3M 996, acquired from the Smithsonian Institute. These tapes were housed on 10-inch metal reels, consisting of 2,400 feet of tape. As seen in Figure 1.9, playability testing was performed on a vertical tape player. The heavy 10-inch reels caused the tapes to slow and, in some cases, even stop during playback. It was difficult to determine if the tape slowed due to being degraded, or due to the weight transfer from the original reel to the uptake reel.

16

Thus, a tape was classified as playable if the entire tape made it through to the uptake reel without stopping. If the tape stopped, or there was squealing from the tape resisting being pulled onto the next reel, it was classified as non-playable. Any 3M 996 tape assigned as playable in our validation set had the majority of the replicates classified as degraded by the 5-fold cross-validation process in the QDA model. This led to all 20 spectra from five 3M 996 playable tapes being misclassified as non-playable, as well as 19 out of 20 spectra from a sixth 3M 996 tape being misclassified.

A full list of the misclassified tapes from the calibration set is shown in Table 1.3. It is preferred that misclassifications of playable tapes as non-playable occur, instead of non-playable classified as playable. If a playable tape is found to be non-playable, the tape would be routed through the baking process previously described. While this can impart some damage to the tape, it is preferred over playing a degraded tape. If a degraded tape is played during the digitization process, it is more probable that it may be destroyed as the binder or data layer sheds onto the player guides and heads, also causing downtime to the player for cleaning and maintenance.

The QDA model was then used to determine the degradation status of the 52-tape test set. The test set was first transformed into the same PC space as the training set to obtain the first five PC scores for the test set. These scores were then submitted to the model to predict which spectra were degraded and which were not degraded, portrayed in Figure 1.10. The confusion matrix in Table 1.4 shows the breakdown of the resulting output for the 1,040 test spectra,

with an overall classification accuracy of 97.79%. Of the 500 spectra from the 25 tapes determined to be non-playable during playability testing, 497 spectra were correctly classified at 99.40%. The three misclassifications were from a SM 3M 808, SM Quantegy 478, and WPT Ampex 406 tape with one spectrum misclassified as playable from each. There were 27 tapes, 540 spectra assigned as playable from the playability testing. The QDA model classified 520 or 96.30% of those as playable. One tape was fully misclassified as playable by the QDA model but classified non-playable by playability testing. This tape was a Library of Congress Scotch 3M 209 tape. The misclassifications for the test set did not yield a pattern as the first spectra from two tapes, the last from another, and all of the fourth were incorrect indicating degradation can occur randomly throughout the tape.

While these misclassifications could be due to the model not performing optimally, more likely the discrepancies arise from using playability testing to train the model. Playability testing can be very subjective, because an audio engineer determines at what level friction, slow recovery, squealing, or sloughed tape indicates a degraded tape. Some tapes can also cause additional problems such as the 3M 996 10-inch metal reel tapes previously described. Playability testing also can only classify the whole tape as being playable or degraded. With ATR FT-IR, the degradation status can be seen at any point sampled along the length of the tape. Therefore, while one location may not show degradation markers, if the other 19 do, the tape should be sent off to be restored through baking. If most of the sample locations have the chemical make-up of a playable tape, but one or

18

two show degradation markers, it is up to the conservationist as to whether to risk baking or playing the tape. Tapes do not necessarily degrade homogenously along the length of the tape, and both degraded and non-degraded regions may be revealed by infrared sampling. Tapes found to be degraded by the QDA model, but determined playable by playability testing, could be due to the level of degradation. ATR FT-IR spectroscopy can identify subtle features indicative of a degraded tape before the tape shows signs, such as friction, during playability testing. A tape could also be misclassified due to additives, such as lubricants, causing the tape to play without issues during playability testing, but exhibit markers of degradation in the IR spectrum.

5. CONCLUSION

Magnetic audio tape was the primary recording media of the late twentieth century and contain much of the world's modern history. Unfortunately, most of the reel-to-reel magnetic audio tape is composed of a polyester-polyurethane binder. This substrate is at risk of degradation through hydrolysis, and when played, a degraded audio tape can cause the data-containing binder to be destroyed along with the recorded information. Libraries and institutions such as the Library of Congress are continually taking preservation steps to save as much of this history as possible by digitizing their collections. Unfortunately, the digitization process involves playing the tape. It is vital then that a non-destructive, reliable, method be developed to analyze these tapes. Previous work by Cassidy *et al.* created a model to triage a subset of tapes obtained from the Library of Congress, laying the groundwork for this research.[3] The original model

19

unfortunately was not applicable to a wide variety of pro-audio tapes. Therefore, it is imperative to create a more robust model that can correctly classify a wider range of magnetic audio tapes.

The QDA model created in this research yielded an overall classification validation accuracy of 95.19% and an accuracy of 97.79% for the test set. The model also misclassified less non-playable tapes as playable, preventing tapes from incorrectly assumed to be playable and being sent to be digitized where a degraded tape could be destroyed. Depending on the importance of a tape, more sampling sites along the length of the tape can be sampled to determine the degradation at various points. This could locate areas of degradation that would not originally show during sampling.

The method shown here has the potential to triage a variety of tapes using one statistical model compared to our previous model. Archivists and cultural heritage institutions can obtain ATR FT-IR spectra for their tapes and analyze them through our model to determine whether baking is necessary before the digitization process takes place. The preservation strategies employed in this research can contribute to the important effort to preserve vital historical recordings.

ACKNOWLEDGEMENTS

Undergraduate Research Conference in Columbia South Carolina in 2017, at the Symposium on Research Computing in Columbia South Carolina in 2017, at the Great Scientific Exchange (SciX) Conference in Atlanta Georgia in 2018, and at the 2018 Southeastern Regional Meeting of the American Chemical Society in 2018. The authors are grateful for the support from Heather M. Heckman, Lydia C. Pappas, and Gregory J. Wilsbacher from the Moving Image Research Collections at the University of South Carolina, and Fennella France, Eric Breitung, and Eric Monroe from the Library of Congress. The authors also recognize the contributions made by University of South Carolina undergraduates and summer researchers, Tret Burdette, Jessica Rotheiser, Aleah Singleton, and Larissa Cunha.

REFERENCES

1. Thiébaut, B.; Vilmont, L.; Lavedrine, B. Report on Video and Audio Tape Deterioration Mechanisms and Considerations about Implementation of a Collection Condition Assessment Method; Report Deliverable D6.1; PrestoSpace: Paris, 2006.

2. Gómez-Sánchez, E.; Simon, S.; Koch, L-C.; Wiedmann, A.; Weber, T.; Mengel, M. "ATR-FTIR Spectroscopy for the Characterization of Magnetic Tape Materials," *E-Preservation Sci.* **2011**, *8*, 2-9.

3. Cassidy, B. M. et al "Rapid and Non-destructive Identification of Degraded Polyester-urethane Magnetic Tape using ATR FT-IR Spectroscopy and Multivariate Statistics," *Analytical Chemistry*, **2015**, 87(18), 9265-9272.

4. *A Public Trust at Risk: The Heritage Health Index Report on the State of America's Collections*, Heritage Preservation, Inc., Washington, D.C., 2005.

5. Edge, M.; Allen, N. S.; Hayes, M.; Jewitt, T. S. "Degradation of magnetic tape: Support and binder stability," *Polym. Degradation and Stability* **1993**, *39,* 207-214.

6. Hess R. L. "Tape Degradation Factors and Challenges in Predicting Tape Life," Association for Recorded Sound Collections Journal **2008**, 34, 240-274.

7. Thiébaut, B.; Vilmont, L.; Lavédrine, B. "Characterization of U-matic videotape deterioration by size exclusion chromatography and pyrolysis gas chromatography/mass spectrometry and the role of adipic acid," *Cultural Heritage* **2009**, *10*, 183-197.

8. Zelst, L. *Sticky Shed Syndrome: Characterization, Diagnosis, and Treatment,* Internal Library of Congress Report, March, 2008.

9. Hobaica, S. "Analysis of audio magnetic tapes with sticky shed syndrome by ATR-FTIR," *J. Appl. Polym. Sci.*, **2012**, DOI: 10.1002/APP.38364

10. Richardson C. A. Process for Restoring Magnetic Recording Tape Damaged by Sticky Shed Syndrome. U.S. Patent 6,797,072, Sep. 28, 2004.

11. Cuddihy, E. F. Hygroscopic Properties of Magnetic Recording Tape. *IEEE Trans. Magn.,* 1976, 12(2), 126-135.

12. Fuenffinger, N. C. Optical Spectroscopy and Chemometrics for Discriminations of Dyed Textile Fibers and Magnetic Audio Tapes. Ph.D. Dissertation, University of South Carolina, Columbia, SC, 2015.

13. Ward, J. Hierarchical Grouping to Optimize an Objective Function, *J. Amer. Statist. Assoc.* **1963**, 58, 236-244.

14. Savitzky, A.; Golay, M. J. E. "Smoothing and differentiation of data by simplified least squares procedures," *Anal. Chem*. **1964**, 36, 1627–39.

15. Barnes, R. J.; Dhanoa, M. S.; Lister, S. J. "Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra," *Appl. Spectrosc*. **1989**, *43*, 772-777.

16. Krzanowski, W. J. *Principles of Multivariate Analysis: A User's Perspective*, Rev. Ed.; Oxford University Press: New York, 2000.

17. Rencher, A. C. *Methods of Multivariate Analysis,* 2nd Ed.; John Wiley & Sons: New York, NY, 2002.

18. Morgan, S. L.; Bartick, E. G. Discrimination of Forensic Analytical Chemical Data Using Multivariate Statistics. *Forensic Analysis on the Cutting Edge: New Methods for Trace Evidence Analysis*, Robert D. Blackledge Ed.; John Wiley & Sons, New York, 2007, 331-372.

19. Mahalanobis, P. C. On the Generalized Distance in Statistics. *Proc. Natl. Inst. Sci. India* **1936**, *2*, 49−55.

Table 1.1. The playability status of the eleven different tape manufactures/brands acquired from three different sources.

Tape Source	Brand/Model	Playability Status
Library of Congress	Ampex 456	Both
Library of Congress	Quantegy 407	Playable
Library of Congress	Scotch 3M 209	Playable
Library of Congress, Wisconsin Public Television, Smithsonian Institute	Ampex 406	LC: Playable WPT: Both SM: Non-playable
Library of Congress, Wisconsin Public Television	Quantegy 406	LC: Playable WPT: Playable
Smithsonian Institute	Emtec	Playable
Smithsonian Institute	BASF	Both
Smithsonian Institute	Quantegy 478	Non-playable
Smithsonian Institute	Ampex 478	Non-playable
Smithsonian Institute	3M 808	Non-playable
Smithsonian Institute	3M 996	Both

Table 1.2. Confusion matrix of the QDA results compared to the playability results for the 154 tape, 3,080 spectra calibration set.

QDA Calibration Model (3080 Spectra)	QDA Classification 2932/3080 = **95.19%**	
Playability Classification	Non-Playable	Playable
Non-Playable	1373/1380 = **99.49%**	7/1380 = 0.51%
Playable	141/1700 = 8.29%	1559/1700 = **91.71%**

Table 1.3. A list of misclassified tape spectra in the calibration set.

Tape Identifier	Replicate Number	Misclassified Prediction
LC Ampex 456	20	Playable
SM 3M 996	20	Playable
SM 3M 996	1	Playable
SM 3M 996	1	Playable
SM 3M 808	20	Playable
SM Ampex 478	1	Playable
WPT Ampex 406	1	Playable
LC Quantegy 407	1	Non-playable
SM BASF	1	Non-playable
SM 3M 996	2-20	Non-playable
SM 3M 996	1-20	Non-playable
SM 3M 996	1-20	Non-playable
SM 3M 996	1-20	Non-playable
SM 3M 996	1-20	Non-playable
SM 3M 996	1-20	Non-playable
LC Scotch 3M 209	1-20	Non-playable

Table 1.4. Confusion matrix of the QDA results compared to the playability results for the 52 tape, 1040 spectra test set.

QDA Test Set (1040 Spectra)	QDA Classification 1017/1040 = **97.79%**	
Playability Classification	Non-Playable	Playable
Non-Playable	497/500 = **99.40%**	3/500 = 0.60%
Playable	20/540 = 3.70%	520/540 = **96.30%**
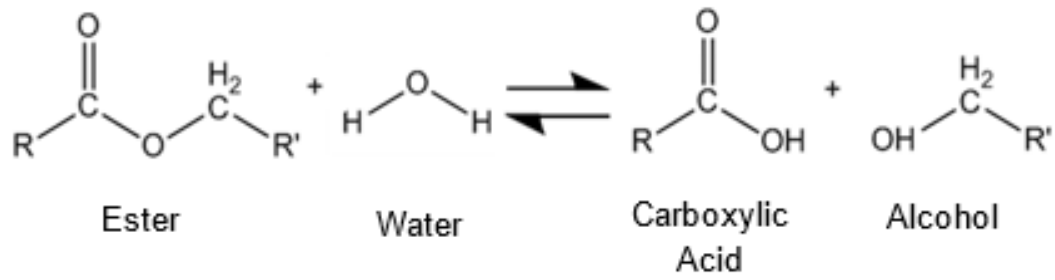
Figure 1.1. Hydrolysis degradation of the polyester binder when in contact with water yielding carboxylic acid and alcohol.
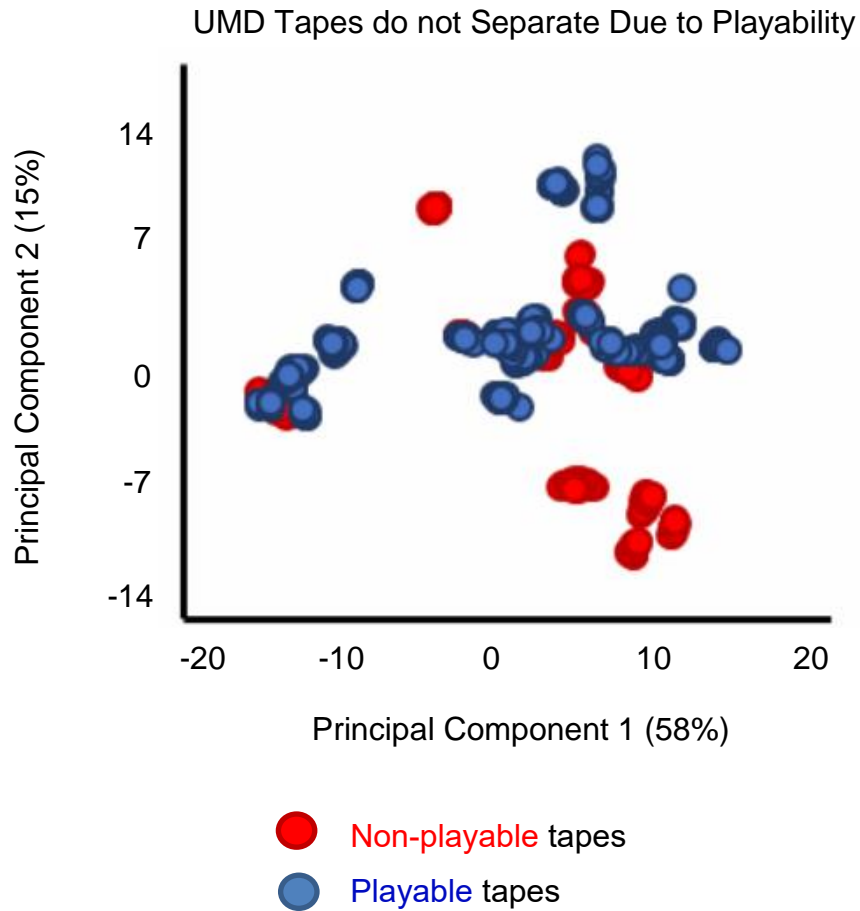
Figure 1.2. Projection of the UMD tapes using the original statistical model from the LC tapes. The blue points represent playable tape spectra while the red points represent non-playable tape spectra.
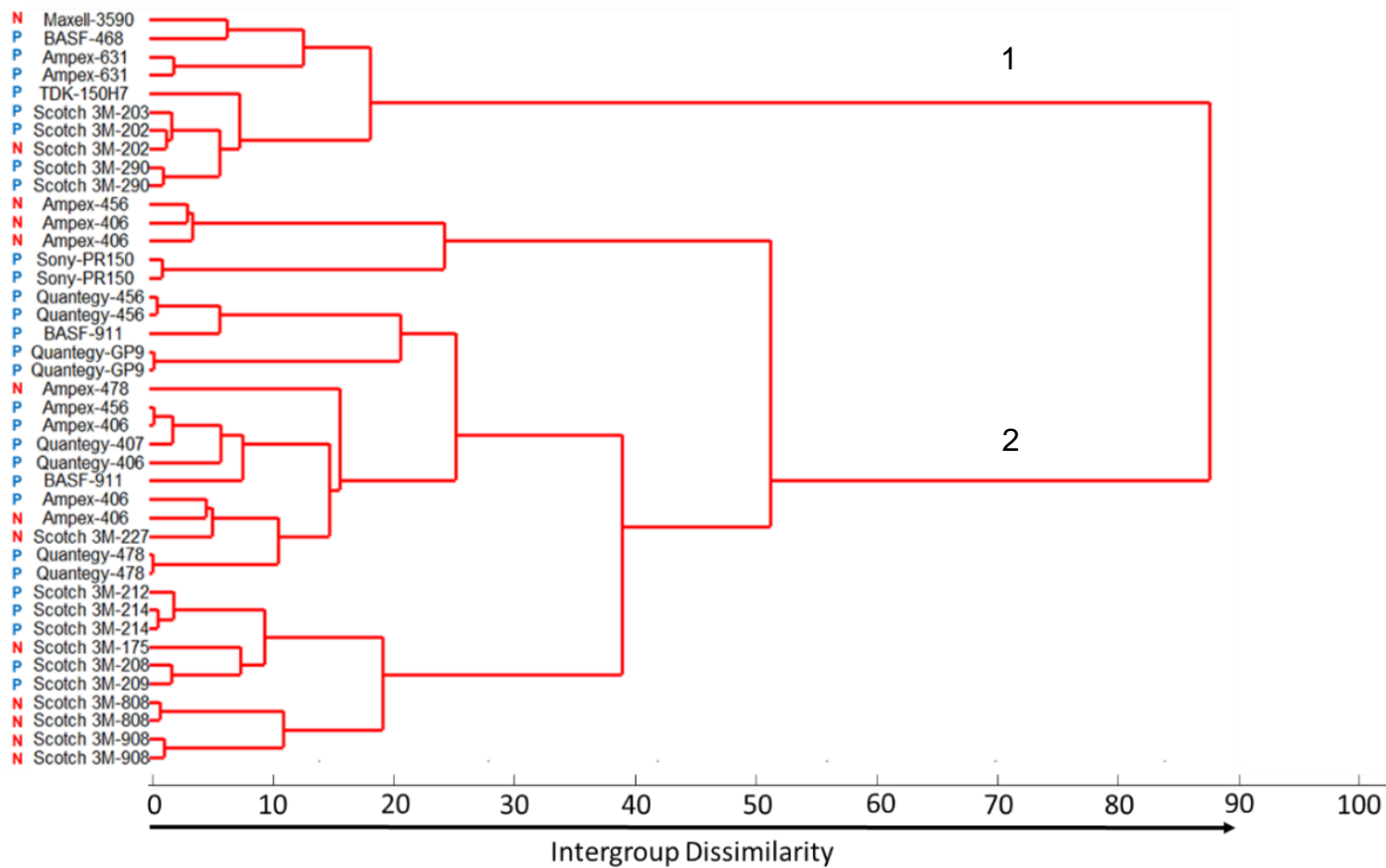
Figure 1.3. HCA describing the degree of dissimilarity between various tape brands and models suggesting playability status is not the main source of differentiation. The blue P represents playable tape spectra, while the red N represents non-playable spectra. Group 1 consists on non-polyester urethane tapes, while group 2 contains tapes manufactured with a polyester urethane binder.
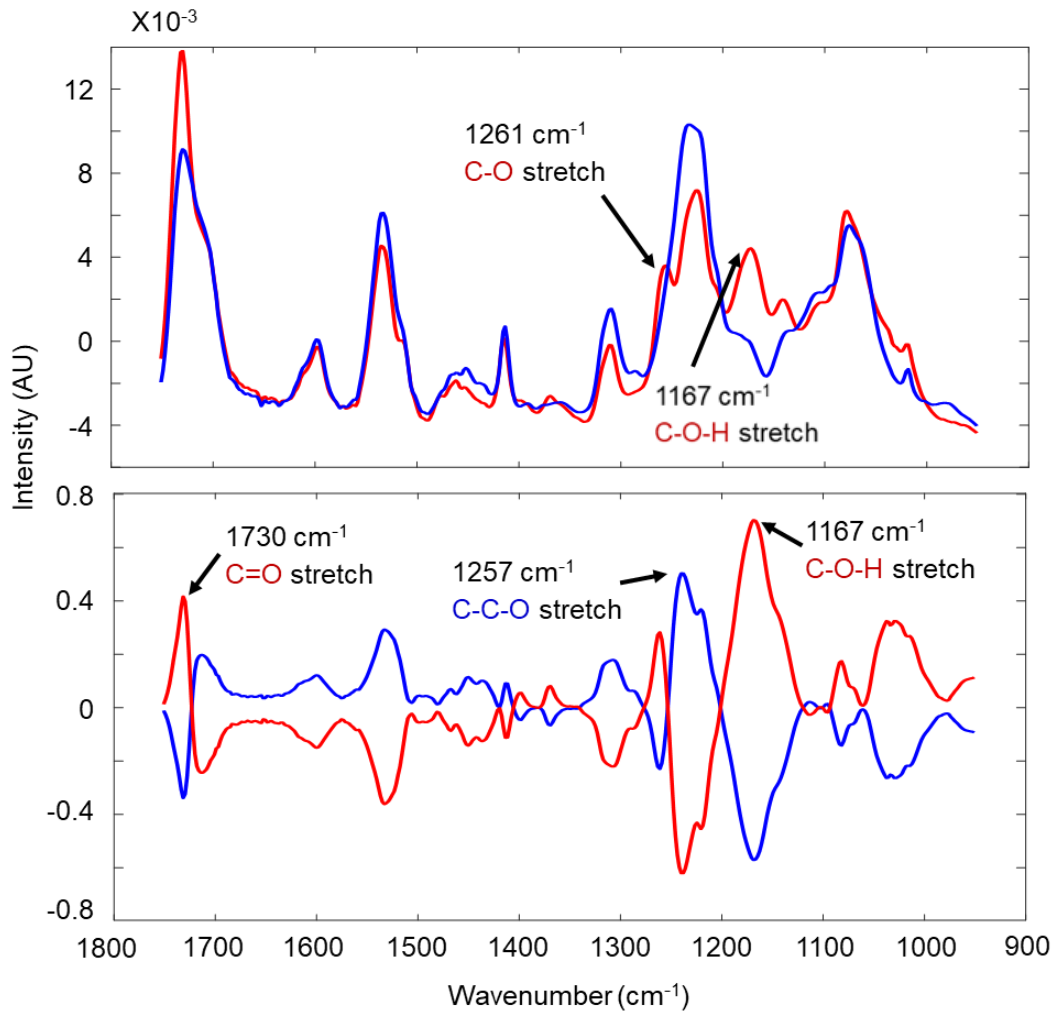
Figure 1.4. Absorbance spectra from the magnetic audio tape samples in the calibration set showing (top) background corrected spectra for playable (blue) and non-playable (red) tapes and (bottom) preprocessed Savitzky-Golay smoothing, standard normal variate transform, and mean centering spectra.
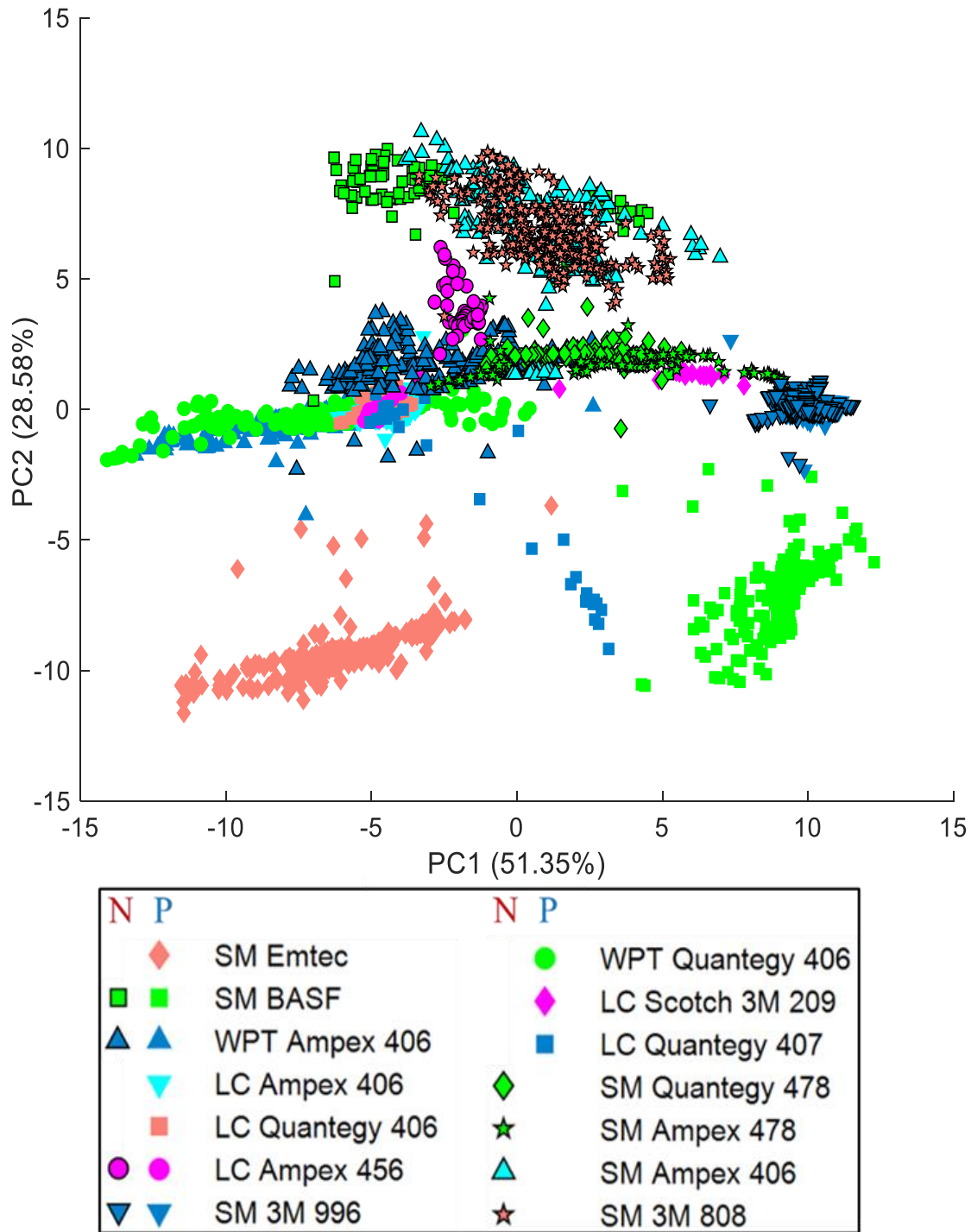
Figure 1.5. Two dimensional PCA plot showing the projection of the 3,080-calibration set spectra into the space containing the first two PCs. PC 1 accounts for 51.35% variability, while PC 2 encompasses 28.58%. Each type of tape is denoted by a different color/marker combination with the non-playable data points outlined in black.
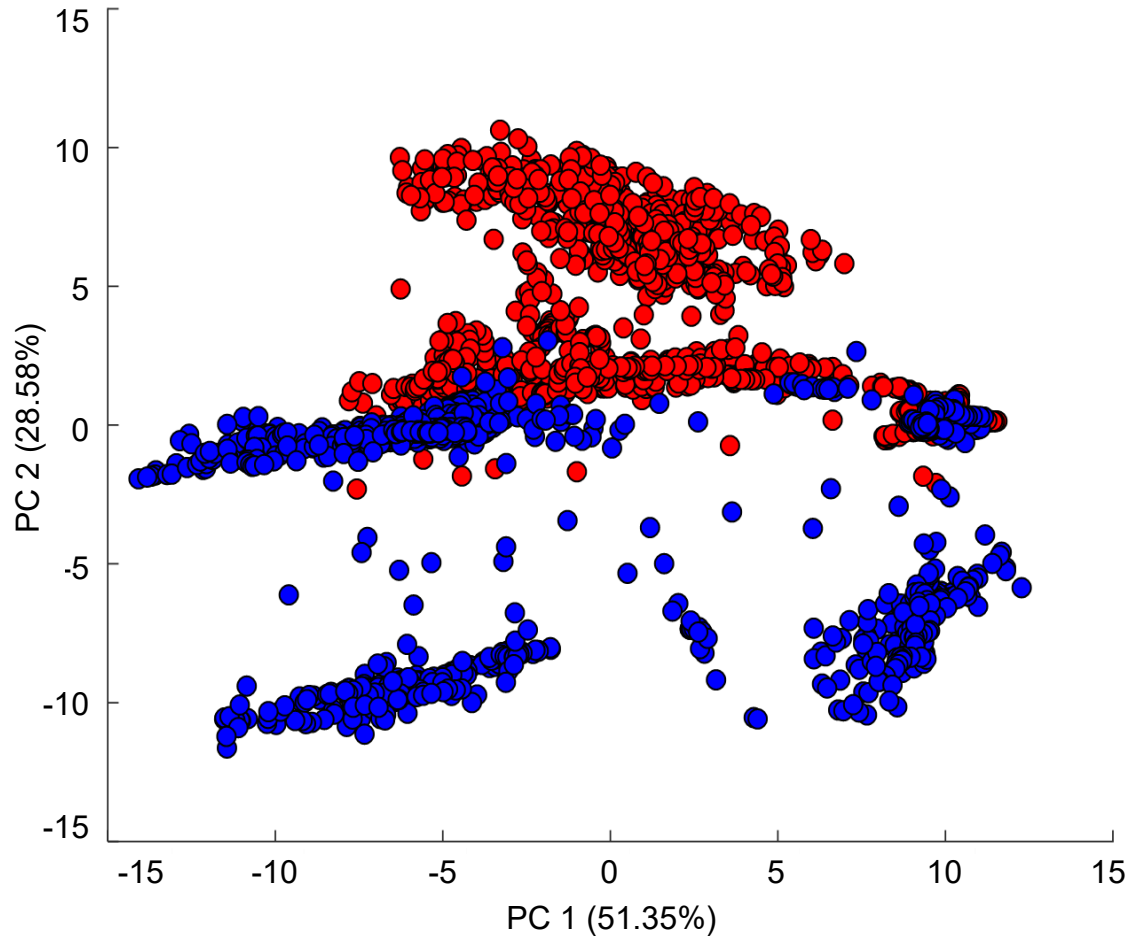
Figure 1.6. Projection of the 3,080-calibration set spectra into the PC space. PC 1 accounts for 51.35% variability, while PC 2 encompasses 28.58% of the variability. The 1700 blue data points represent determined playable through playback. The 1380 red data points represent spectra from tapes considered non-playable during playability testing.
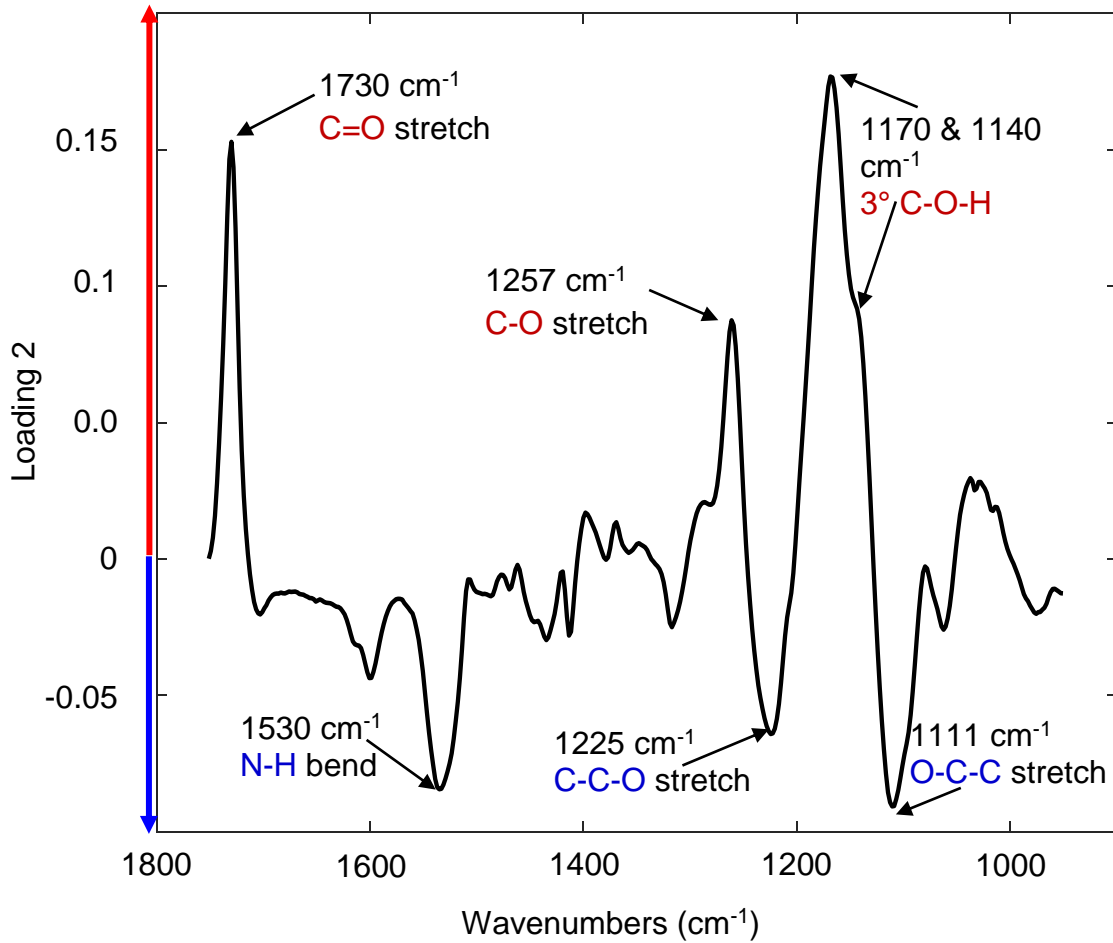
Figure 1.7. The loading for principal component 2 shows the chemical differences between the spectra projected in the positive region and those in the negative PC space. These wavenumbers correspond to the degradation products found in non-playable tapes.
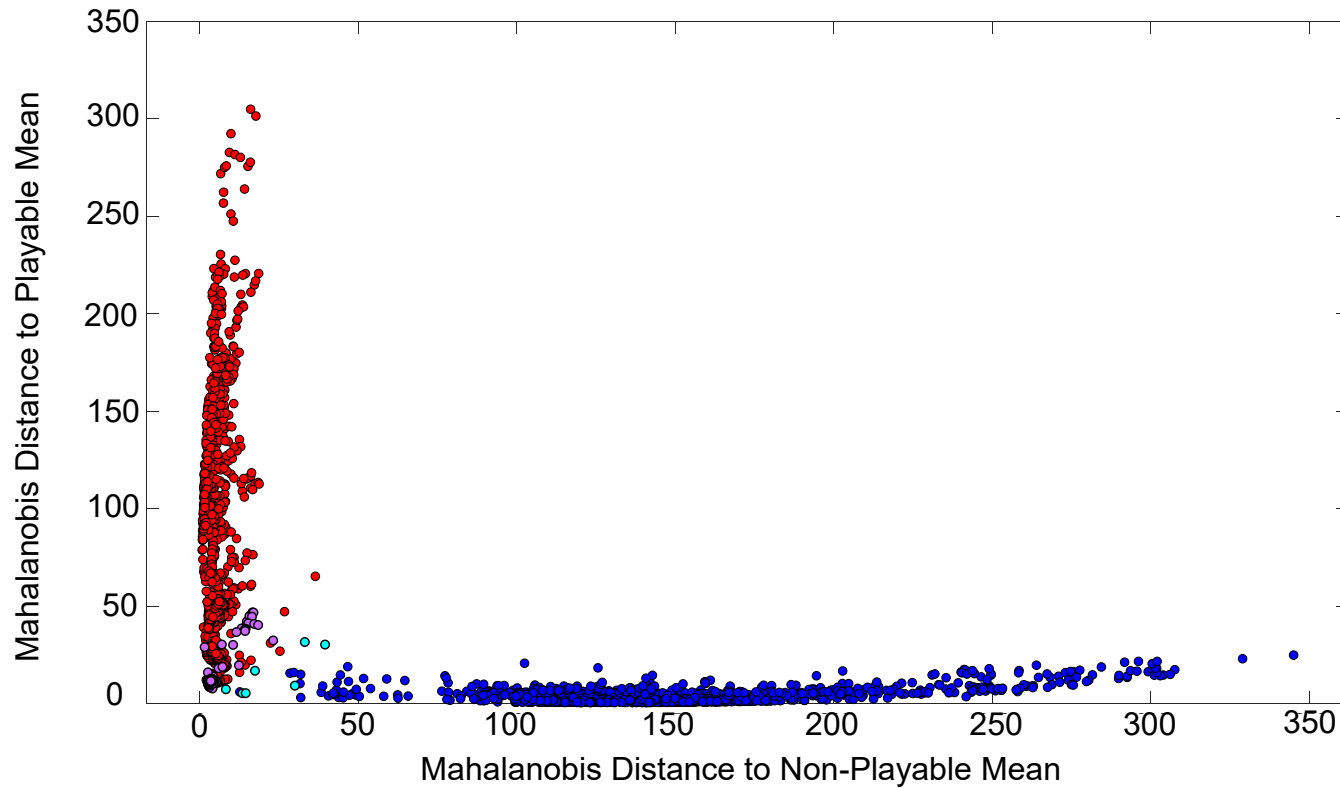
Figure 1.8. QDA response for the playability validation of the calibration model. The Mahalanobis distance from the 3,080 spectra to the playable and non-playable class centroids (means) were calculated. The 1,559 blue data points represent playable spectra correctly classified as playable while the 141 purple data points represent playable spectra classified as non-playable through the validated QDA model. The 1,373 red points were correctly classified as non-playable but the 7 cyan were identified as playable by the QDA model.

Figure 1.9. Scully 280 tape player used for playability testing. A 10-inch reel housing a 3M 996 tape is being transferred to a free uptake reel.
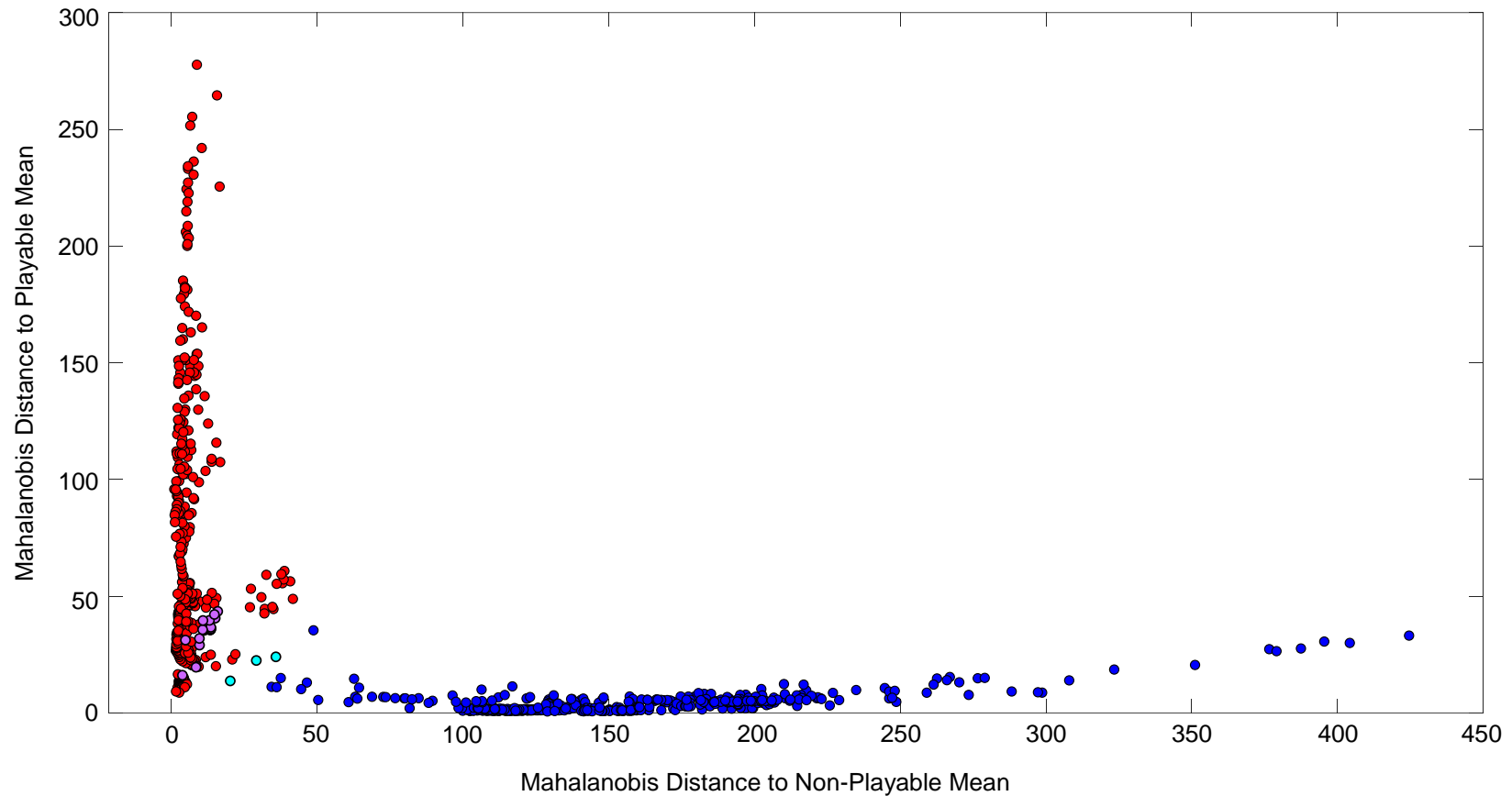
Figure 1.10. QDA results for the degradation of the test set using the calibration model. The 520 blue data points represent playable spectra correctly classified as playable while the 20 purple data points represent playable spectra classified as non-playable through the validated QDA model. The 497 red points were correctly classified as non-playable but the 3 cyan were identified as playable by the QDA model.

CHAPTER 2

EXTERNAL VALIDATION FOR THE TRIAGE OF MAGNETIC AUDIO TAPES USING SPECTROSCOPY AND MULTIVARIATE STATISTICS

1. ABSTRACT

There are over 4.8 billion artifacts housed in various U.S. archives, libraries, and scientific institutions, encompassing rare books, art, artifacts, and magnetic audio tapes. The preservation of magnetic audio tapes is of utmost importance to cultural heritage institutions, museums, and audiophiles alike. This form of media was primarily used in the twentieth century, but remains in use today. Unfortunately, this source of the world's modern cultural history is degrading. This is especially true for magnetic tapes manufactured with polyester-urethane (PEU) binders. While it is agreed that the source of the degradation is hydrolysis of the ester component, triaging these tapes has been difficult without the possibility of destroying them. We have previously shown that degradation of magnetic audio tapes can be determined using non-destructive attenuated total reflectance Fourier transform infrared spectroscopy (ATR FT-IR) coupled with multivariate modelling.[1] Chapter 1 in this manuscript elaborated on that model by using a more robust data set consisting of eleven

different types of tapes from various manufacturers housed in different institutional collections.

The focus of this present study is to create a robust model that can determine the degradation status of the type of tapes used in the calibration model and direct triage of tapes that may not be composed of those exact tape chemistries by validating with external test sets. The ATR FT-IR spectra of a collection of magnetic tapes were subjected to quadratic discriminant analysis, creating a predictive model for the non-destructive degradation determination of PEU based magnetic audio tape. A collection of 206 tapes was divided into a 154-tape calibration set, with an additional set consisting of 52 tapes held back as a test set. The resulting model produced a validation accuracy of 94.64% and a test set accuracy of 95.58%. The model was also validated with two external test sets, one obtained from the Library of Congress's (LC) Motion Picture Broadcasting and Recorded Sound Division (MBRS), and the other from the University of Maryland (UMD). The MBRS and UMD test sets performed with 90.79% and 78.85% accuracies, respectively.

2. INTRODUCTION

The history of magnetic audio tape can be traced back to Thomas Edison's invention of the phonograph in 1877, which could transcribe spoken vibrations into grooves using a recording needle and metal cylinder. The device would then play back the message by reading the indented pattern.[2] This advance was followed by Oberlin Smith, an American engineer, who in 1888

39

suggested using magnetized steel wire to record sound.[3] The first magnetic recorder was invented ten years later by Valdemar Poulsen, a Danish telephone engineer who used Smith's ideas in his Telegraphone based on a magnetized thin metal wire. The first use of magnetic tape, called the Magnetophone, was developed in 1928 by German inventor Fritz Pfleumer using magnetic oxide on a paper strip. The German company AEG furthered this technology by partnering with BASF and incorporating polymers into production to form the first practical means of tape recording. [4] Since their initial production, magnetic audio tape recordings have captured much of the world's twentieth century history.

Magnetic tape is a complex substance composed of a magnetic top layer, a substrate, and an optional back coating as shown in Figure 2.1.[5, 6] The top layer consists of magnetic particles, usually iron oxide (Fe_2O_3) or chromium dioxide (CrO_2), embedded in a binder.[6, 7] Additives found in the binder include lubricants (fatty acid esters) to reduce friction, dispersants (zinc naphthenate, dioctyl sodium sulfoccinate, or mono-, di-, and triesters of oleic acid, to name a few) used to distribute metal particles, as well as pigments, such as Al_2O_3 or SiO_2, that act as abrasives to clean tape player heads.[5, 8, 9] In addition, the binder contains a mixture of soft and hard segments, varying by brand, which may change during manufacturing.[8] The soft segment is formed from a high molecular weight polyester, giving the magnetic tape elastomeric properties. The soft segment is usually composed of poly(ethylene adipate) (PEA), poly(1,4-butylene adipate) (PBA), poly(caprolactone) PCL, or polycarbonate.[6, 10] The hard segments are usually a polyurethane, such as m-xylene diisocyanate, toluene

40

diisocyanate, 3,3'-tolidene-4-4'-diisocyanate (TODI), or 4,4'-methylene diphenyl diisocyanate (MDI). The hard segments are linked by small molecule chain extenders between the isocyanate molecules and are usually butanediol (BD) or hexane-1,6-diol (HD).[8, 10] Due to the combination of varying polyester soft segments and polyurethane hard segments that can be used, the binder is generally referred to as polyester-urethane (PEU). The substrate or base film is responsible for the mechanical properties and physical integrity of magnetic audio tapes.[6, 8, 11] The first tapes used paper backings, then cellulose acetate from 1935 to the 1960s.[9, 12] This was followed by polyvinyl chloride (PVC) in the 1940s to the 1970s, which was more popular in Germany than the United States (US), before manufacturers settled on a polyester backing, polyethylene terephthalate (PET), in the 1950s.[9, 12] An optional back-coating made from either polyester-urethane or polyether-urethane with carbon particles was used by some manufacturers to dissipate static charge while playing the tape.[6, 7]

Archives, institutions, and libraries, such as the Library of Congress, are responsible for preserving and providing a safe environment for historical collections. The number of magnetic tapes in US libraries is estimated to be more than 21.5 million, accounting for the highest media containing recorded sound. Unfortunately, magnetic tapes tend to degrade in storage, leading to the loss of the information contained.[13] The complex nature of magnetic tapes leads them to be susceptible to different forms of degradation: oxidation of the magnetic oxide particles, as well as pyrolysis due to excessive heat. However, the primary degradation mechanism is attributed to the hydrolysis of ester in the PEU

41

binder.[14, 15] Hydrolysis causes polyester in the binder to breakdown into smaller molecules, degrading the integrity of the binder. This causes the tapes to stick to the player guides and heads, as well as shed when played, which is known as known as sticky shed syndrome (SSS).[6] As previously described, magnetic audio tapes can be manufactured with an array of different chemicals and additives to impart properties such as stability, minimizing friction, and static dispersion. The type of polymers used in the binder, both for the soft and hard segments, can also vary, changing the stretch and polymeric properties of the tape. Although magnetic tapes can vary in composition, the general use of PEU type binders imbue most brands with susceptibility to degradation by hydrolysis mechanisms.

As discussed in Chapter 1, playability testing is the main technique for determining if a tape was degraded. However, if a tape that is experiencing sticky shed syndrome (SSS) is played, the tape could be destroyed. Hobaica, *et al.* showed that attenuated total reflectance Fourier transform infrared spectroscopy (ATR FT-IR) could correctly classify a tape 70% of the time through a simple, visual comparison of the spectra.[8] This work was followed by Cassidy, *et al.* by applying multivariate statistics to analyze spectra for differences between playable and non-playable (degraded) tapes, which enabled reliable classification accuracies in the low 90% range.[1] Since magnetic tapes are complex, it is necessary to obtain a diverse data set consisting of different tape models from different manufacturers. Only then could the degradation status of tape health be investigated by creating a robust model to predict the degradation status, and enable the triage of tapes held in various collections. Chapter 1 exhibited the

success of that approach with a training set of 11 different brands/models acquired from three different sources, producing classification accuracies of 97.87% and 97.79% for the resulting validation and test sets. While this model correctly classified the original test set, it could not determine the degradation status of an external test set consisting of the 95-tape calibration set from Cassidy, *et al.*, which yielded an accuracy of only 55.26%, as seen in the confusion matrix in Table 2.1. Figure 2.2 shows the quadratic discriminant analysis model response using the Mahalanobis distance classifier for this data set. The red, blue, purple, and cyan data points represent correctly classified non-playable and playable spectra, and the misclassified non-playable and playable spectra, respectively.

A peak located between 1196 cm^{-1} and 1151 cm^{-1} was found only in the Chapter 1 model's degraded tapes, used to create the calibration model, and which was associated with the C-OH alcohol stretch from the ester hydrolysis degradation. This feature was found in both playable and non-playable spectra in the 95-tape MBRS tape set used by Cassidy, *et al.* and is thought to be due to the type of ester used in the binder manufacturing. Most Scotch 3M tapes exhibited this peak, regardless of degradation. We believe the 95-tape set is comprised of Scotch 3M tapes since they contain the 1196 cm^{-1} and 1151 cm^{-1} peak regardless of degradation status. Most of the tapes were also found in boxes labeled Scotch 3M, although rehousing practices make this an assumption since the box label does not guarantee the correct manufacturer information for the tape inside the box. Chemical formulations for magnetic audio tape are

43

proprietary, so determining the exact components is not possible. It is believed that Scotch 3M may have used a specific polymer in their binder called Estane®, a type of polyester polyurethane that uses poly(caprolactone) as the polyester precursor versus the poly(1,4-butylene adipate) polyester precursor used in Ampex tapes.[15] Schoonover, *et al.* compared Estane® to poly(butylene adipate) and found that Estane® contained a peak at 1176 cm^{-1} related to the C-O-C ester stretch.[16] It is believed that the original 95-tape MBRS tape set from Cassidy, *et al.* did not classify correctly due to the precursor used in the polyester segment of the polyester polyurethane binder. Therefore, a new model must be made to correct for this feature. This peak is seen in Figure 2.3, where the top panel shows the averaged baseline corrected playable (blue) and non-playable (red) Ampex 406 spectra. The bottom panel of Figure 2.3 shows the averaged baseline corrected playable (blue) and non-playable (red) spectra for Scotch 3M 175 tapes. The arrow in the top panel of Figure 2.3 points to the feature at 1175 cm^{-1} corresponding to the C-OH degradation peak in the Ampex tapes. The arrow in the bottom panel of Figure 2.3 identifies the C-O-C ester stretch in the Scotch 3M tapes. Since this peak cannot be contributed solely to the degradation of the tapes, and therefore causes misclassifications in tapes such as Scotch 3M, it must be removed to create a robust tape triage model.

The objective of this study is to create a robust tape model to predict the playability status of external test sets consisting of unknown brand tapes using ATR FT-IR spectroscopy and chemometrics. This is accomplished by creating a new model with the same eleven different brands/models tapes from the three

different institutions used in Chapter 1, but with the aforementioned feature omitted. The model will be evaluated using the original 52-tape test set. Two unknown brand/model external tape sets will then be analyzed using the model and classified using the chemical markers indicative of playable and non-playable tapes. The first external test set consists of the 95-tape calibration set used in the Cassidy, *et al.* paper, further referred to as the MBRS tape test set. The second external test set will use the University of Maryland (UMD) tapes described in Chapter 1 that did not classify correctly with the Cassidy model. All test sets will have the features from 1196 cm^{-1} to 1151 cm^{-1} omitted to correct for the differing polyester precursor used in the binder formulations between manufacturers.

3. EXPERIMENTAL

3.1 Sample Selection

This study focuses on quarter-inch reel-to-reel PEU pro-audio tapes that were determined to suffer from SSS. The tapes used in the calibration model were unopened tapes collected from the Library of Congress (LC), Wisconsin Public Television (WPT), and the Smithsonian Institution (SM). The first external test set consisted of 95 tapes selected from the Library of Congress's Motion Picture Broadcasting and Recorded Sound Division (MBRS) (Culpeper, VA) and were the tapes used in the Cassidy, *et* al. calibration set. The brand and model information are believed to be primarily Scotch 3M, but due to rehousing practices, the information found on the tape hub or box is unreliable. The second

external test set consists of thirteen tapes donated from UMD and will be denoted as the UMD test set for the remainder of this paper. The brand/model information are not known for these tapes as well.

3.2 Playability Testing

Playability testing is a subjective method to determine degraded tapes. An audio engineer plays a tape using a vintage tape player and classifies it as degraded if the tape sheds onto the player, squeals, has slow recovery between forwarding and reversing, or stops all together. Playability testing for this study was performed on a Scully 280 tape player (Scully Recording Instruments, Bridgeport, CT). The playability testing for the 154-tape calibration set and the 52-tape test sets was performed by the author of this paper. The playability testing for the MBRS and UMD tapes were performed by a sound engineer at the MBRS at the LC with a separate Scully 280 tape player. For each set of tapes, playability testing involved repeating playing the tape for five to ten seconds, rewinding then fast forwarding for two seconds each, repeated three times, before toggling between fast forward and rewind to advance the tape a third of the way through. This was repeated three times until the tape was fully transferred to the uptake reel.

3.3 Infrared Analysis

ATR FT-IR spectra were obtained using a Nexus 670 FT-IR with Omnic® version 8.2 (Thermo-Nicolet, Madison, WI) equipped with a deuterated triglycine

sulfate (DTGS) detector and a Thunderdome™ ATR (Thermo Spectra Tech, Inc.,

Shelton, CT), with a germanium crystal and an incident angle of 45°. The setup

described in our previous paper by Cassidy, *et al.* was followed to protect the

tape. This entailed orienting the tape magnetic side down against the ATR

crystal, followed by a Mylar™ film to protect the tape, and a ~1 mm thick steel

plate to disperse the pressure from the ATR pressure tower evenly.[1] This

arrangement minimized imparting any pressure points or damage to the tapes.

Ten replicate spectra were acquired at ~50 cm, and an additional ten spectra at

~100 cm from the beginning of the tape with a spectral range from 4000 cm^{-1} to

600 cm^{-1}, using 32 scans at 4 cm^{-1} resolution.[1] The spectra for the 154-tape

calibration and 52-tape test sets were acquired by the authors of this paper, while

the MBRS and UMD spectra were taken from previous analysts in our laboratory

and at the LC.

3.4 Data Analysis

MATLAB version 9.5, R2018b, (The MathWorks, Natick, MA) was used for

all preprocessing and data analysis for both the calibration and test sets. The

'fitcdiscr' and 'classify' functions were used with the discriminant type 'quadratic',

and the 'mahalanobis' distance classifier.

3.4.1 Preprocessing

The region from 1750 cm^{-1} to 950 cm^{-1} was identified as containing the

degradation products for magnetic audio tapes.[8] The features from 1196 cm^{-1} to

47

1151 cm^{-1} were also removed to correct for the peak being present in most playable and non-playable MBRS tapes. This peak was only found in the degraded tapes from the calibration tape set. All tapes were subjected to preprocessing using a 13-point Savitzky-Golay (SG) fourth order polynomial for spectral smoothing.[17] This was followed by applying the standard normal variate transform (SNV) to remove scatter effects by subtracting the mean absorbance of the individual spectrum from each feature's intensity, then dividing by the standard deviation for the absorbance of each spectrum.[1, 18] Mean centering was then performed by calculating the average absorbance per wavenumber in the 154-tape calibration set and subtracting that average from each wavenumber in the 154-tape calibration set, 52-tape test set, MBRS, and UMD test sets.[1]

3.4.2 Principal Component Analysis

Principal component analysis (PCA) was performed as a dimensionality reduction technique with new variables created from correlated wavenumbers. The first principal component (PC) accounts for the most variability in the data set, followed by the second PC, and so on.[19] Unlike the original Chapter 1 model that used five PCs to account for 95% variability, when the 1196 to 1151 cm^{-1} features were removed, six PCs were needed, accounting for 95.12% variability.

3.4.3 Quadratic Discriminant Analysis

Quadratic discriminant analysis (QDA) was used to create the model from the first six PCs. As previously mentioned, QDA is a supervised classification

48

technique for which response assignments from playability testing were used to train the model. A quadratic decision boundary was used to determine playable and non-playable tapes. QDA assumes that the data is normally distributed with unequal covariance matrices for playable and non-playable classes. K-fold cross-validation was used to validate the model with k set to five groups, randomly split from the original 154-tape calibration set. One group was held out as a test set, while the remaining four were used as training data. A QDA model was created using the training data and validated using the test set. This was repeated until each group had been used as the test set. The Mahalanobis distance was then calculated from each sample to the class centroids for both playable and non-playable groups. A sample was assigned to the group with the closer class centroid.[19, 20] The 52-tape, 1,080 spectral test set from Chapter 1, the 95-tape, 1,900 spectral MBRS test set, and the 13-tape, 260 spectral UMD test set were then projected into the same six-dimensional PC space as the calibration set. Using the new QDA model and the minimum Mahalanobis distance classifier, the spectra from each test set were classified as degraded or non-degraded, depending on the closer class centroid.

4. RESULTS AND DISCUSSION

4.1 Playability Testing

Playability testing was performed on the 154-tape calibration set, 52-tape test set, 95-tape MBRS test set, and 13-tape UMD test set. Of the calibration set tapes, there were 85 tapes determined to be playable, with the remaining 69

determined to be non-playable. From the 52-tape test set, there were 27 test set tapes determined to be playable and 25 non-playable, 46 of the 95-MBRS test set tapes were playable with 49 being non-playable, and of the 13-UMD test set tapes, 6 were playable, leaving the remaining 7 as non-playable.

4.2 Infrared Analysis

ATR FT-IR spectra were acquired from the calibration set and the three test sets. Ten replicate spectra were obtained at ~50 cm, and an additional ten spectra at ~100 cm from the beginning of the tape. The averaged baseline corrected non-playable and playable spectra for the calibration set can be seen in the top of Figure 2.4, represented by the red and blue spectra, respectively. The bottom panel of Figure 2.4 depicts the preprocessed non-playable and playable spectra as red and blue respectively, after SG smoothing, SNV transform, and MC. As described in Chapter 1, and seen here in Figure 2.4, the raw spectra exhibit some differences in the peaks at 1146 cm^{-1} and 1261 cm^{-1} found in the non-playable spectrum. These differences, as well as additional subtle effects, are more prominent in the preprocessed spectra. As seen in the top of Figure 2.4, the peak and shoulder located at 1730 cm^{-1} to 1700 cm^{-1} from the C=O ester stretch is apparent in both the playable and non-playable spectra. After preprocessing, this feature differentiates the two spectra, as shown in the bottom of Figure 2.4.

4.3 Principal Component Analysis

PCA was used as a dimensional reduction technique and for exploring the variability in the 154-tape calibration set. The first six PCs were used since they encompassed over 95% of the variability with PC 1 to PC 6 being 42.14%, 28.60%, 11.48%, 7.60%, 3.86%, and 2.95% respectively. While a combination of these six PCs contribute to the overall separation of non-playable and playable tapes, PC 2 most closely resembles the chemical differences attributed to the degradation of the PEU binder, as seen in Figure 2.5, showing PC 1 and PC 2. In the figure legend each brand/model tape is indicated by a different symbol and color combination. Symbols with a black outline indicate a non-playable spectral data point. The PCA plot (Figure 2.5) shows separation due to the chemical differences between playable and non-playable tapes. Similar brand/model tapes group together, as in the case of the LC Ampex 209 spectra. It is also evident that the playable spectra are projected in a tighter cluster than their non-playable counterparts (see Ampex 406 data points). This is due to the differing level of degradation found in the non-playable tapes, as some have just started to degrade while others are more degraded. Since it is apparent that the brand and model information is not needed to determine playability status when a wide array of tape manufacturers are used, Figure 2.6 shows the PCA plot with only playability information. This figure only shows PC 1 and PC 2, causing some spectra to appear as if they may be misclassified. This may not be the case since the model contains the first six PCs for classification and four other PCs were considered in playability determination.

www.manaraa.com

4.4 Principal Component Loading Analysis

To validate PC 2 as encompassing the chemical markers attributed to the degradation of magnetic audio tape, the eigenvector components for the second PC loading were shown in Figure 2.7. The eigenvector components in the positive direction correlate to wavenumbers found in the non-playable tapes, since the non-playable tapes were projected in the positive region of the PC 2 space. The prominence of a peak dictates the influence a feature yields in determining playability status. The most influential features are those found at 1730 cm^{-1}, 1261 cm^{-1}, and 1146 cm^{-1}. These wavenumbers relate to functional groups found in the ester degradation via hydrolysis. The 1730 cm^{-1} C=O ester stretch peak is found in playable spectra but was more intense in the non-playable counterparts. While both degraded and non-degraded tapes contain esters, degraded tapes also contain a carbonyl stretch in the carboxylic acid product. The carbonyl stretch from the PEU carboxylic acid hydrolysis product was more intense than the carbonyl from the PEU ester.[1] It was also determined that a shift occurs from 1693 cm^{-1} to 1730 cm^{-1} from playable to non-playable tapes due to non-playable tapes having more free carbonyl groups while playable tapes have hydrogen bonded carbonyls.[8] Esters degrade via hydrolysis into carboxylic acids and alcohols. The C-O carboxylic acid stretch was found at 1261 cm^{-1} and the C-O-H alcohol stretch at 1146 cm^{-1}. These features were also seen in the positive region of the second PC loading, corroborating the belief that the variability found in PC 2 was due to the ester degradation products.

Eigenvector components related to playable tapes are found in the negative direction of the second PC loading with playable data points projected into the negative PC 2 space. Features found in playable spectra include the 1537 cm^{-1} N-H amide bend from the polyurethane hard segment of the PEU binder, the C-C-O ester stretch at 1225 cm^{-1}, and the O-C-C ester stretch at 1111 cm^{-1}.

4.5 Quadratic Discriminant Analysis

QDA was used to create a calibration model using the first six PCs determined from the 154-tape calibration set. Validation was performed using five-fold cross-validation with the Mahalanobis distance classifier. Figure 2.8 shows a visual representation of the validated QDA model. The correctly classified non-playable and playable spectra are depicted as red and blue data points, respectively, while the purple and cyan points represent spectra falsely classified as non-playable and playable, respectively.

Table 2.2 displays the confusion matrix summarizing the validation of the calibration model using playability information as "true" assignments. Of the 3,080 spectra in the calibration set, 2,915 were correctly classified, yielding a total validation accuracy of 94.64%. From the 1,380 spectra for the 69 tapes assigned non-playable through playability testing, 1,344 (97.39%) were determined degraded through the model. For the 85 tapes, 1,700 spectra assigned as playable, 1,571 (92.41%) were correctly classified. More playable spectra were misclassified than non-playable. This is a preferred outcome since

53

these tapes would be baked to temporarily reverse the degradation. A non-playable tape assigned playable may be destroyed during the digitization process as the binder sticks and sheds onto player heads and guides.

The only full tapes, with all twenty spectra to be misclassified, were from SM 3M 996 tapes. As described in Chapter 1, these tapes were particularly hard to evaluate during playability testing due to their large size and mass. Of the six SM 3M 996 tapes determined playable through playability testing in our calibration set, all but one spectrum was misclassified as non-playable. Two non-playable SM 3M 996 spectra from different non-playable tapes were misclassified as playable. The remaining misclassifications were scattered throughout the other brand/model tapes, most only containing one misclassified replicate. There was no pattern as to which replicate was misclassified, as some tapes had the first replicate, some the last, and other misclassifications were scattered throughout the twenty replicates. Given a mixed degradation prediction, a preservationist would have to determine whether to bake or digitize a tape based on the overall outcome and value of the tape. If most of a tape was deemed playable, it may be digitized, while if a tape was primarily non-playable it would be advantageous to bake it before digitization. The value of a tape would also need to be considered to prevent ruining a tape by prematurely digitizing it before restoration.

The degradation of the 52-tape test set, the MBRS test set, and the UMD test sets were determined using the QDA model. The first six PC scores for each

test set were obtained by projecting the data into the same PC space as the calibration set before being submitted to the model for prediction. The 52-tape test set output can be seen in Figure 2.9, with the same color combination as used in the calibration set. The confusion matrix in Table 2.3 summarizes the predictions for this test set. A total 994 of the 1,040 spectra were correctly predicted with an overall classification accuracy of 95.58%. Of the 500 spectra from the 25 tapes determined non-playable through playability testing, 457 were correctly assigned, achieving 91.40% correct classification. Misclassifications arose from seven different tapes. Four different Ampex 406 tapes from WPT were misclassified, of which the first spectrum from one, the twentieth from another, seventeen spectra from a third, and four spectra throughout the fourth were assigned as playable. There were two SM Quantegy 478 tapes misclassified with only one having the first spectra predicted playable, and the second contained the last ten replicates, or the ~100 cm location misidentified. The last non-playable tape with an incorrect prediction was a SM 3M 996 tape whose first ten replicates located at the ~50 cm location were misclassified. There were 537 spectra (99.44%) out of a total 540 spectra from 27 tapes accurately deemed playable. The three misclassified spectra as were from three different tapes: one replicate spectrum (replicate twenty) from a LC Ampex 406 tape, replicate twenty from a SM Emtec tape, and replicate one from a LC Scotch 3M tape. In Chapter 1 all the LC Scotch 3M replicates were misclassified. By omitting the region from 1196 cm^{-1} to 1151 cm^{-1}, all but one replicate was

correctly classified, improving upon the model's ability to correctly predict the degradation status for Scotch 3M tapes.

To test the model with external test sets containing tapes that may not be in the calibration mode, two test sets were chosen. The first was the MBRS 95-tape calibration set from Cassidy, *et al.* and the second was a set of 13 tapes from UMD. The UMD tapes did not correctly classify when the MBRS 95-tape set was used to create a model, most likely due to the small variety of tapes in that calibration set. The degradation status of these test sets was predicted using the current model built from a robust 154-tape calibration set consisting of eleven different tape brands and models.

The MBRS tape set consisted of 1,900 spectra from 95 tapes. Using the current model, 1,725 spectra (90.79%) were correctly classified as seen in Table 2.4. Results show that 938 of the 980 (95.71%) non-playable spectra and 787 of the 920 (85.54%) playable spectra were correctly predicted. The Mahalanobis distance from each class centroid can be seen in Figure 2.10. There were three tapes whose non-playable spectra were misclassified as playable. The majority of spectra from these tapes were misclassified, with eleven from one, twelve from another, and nineteen of twenty from the last. As with the 52-tape test set, there was no pattern to the misclassified replicates, as there were misclassifications in both regions sampled. There were twelve tapes with playable spectra misclassified as non-playable. Four of the twelve tapes had all replicates incorrectly assigned, two only had the first sample location differing, and the

remaining five had varying numbers of misclassifications throughout the tape.

Again, there was no pattern to those misclassified.

The last tape set used to test the QDA model contained thirteen tapes, six non-playable and seven playable, from UMD. The results can be seen in Table 2.5 and Figure 2.11. Of the 260 spectra, 205 were correctly classified, yielding 78.85% classification accuracy. There were 113 of the 120 spectra (94.17%) determined through playability testing to be non-playable that were correctly classified. Seven non-playable spectra were misclassified as playable from two different tapes (five from one and two from another). In all, there were fewer correctly classified playable tapes with 92 of the 140 spectra (65.71%) predicted as playable. The 48 misclassified spectra were from four different tapes, two of which had all twenty spectra incorrectly predicted, one had five, and the last contained three false non-playable assignments. The overall accuracy predicted through the 154-tape QDA calibration model is greater than when the degradation of the UMD tapes were predicted with the 95-tape MBRS QDA calibration model (66.92%).

The validity of this QDA model was compared to playability testing, with playability testing responses taken as the true degradation status of the tapes. Unfortunately, playability testing can be subjective, lending the results unreliable. Playability testing also can only assign one response to the entire tape, causing it to be less sensitive than instrumental analysis where multiple sampling locations can occur. ATR FT-IR is more sensitive, identifying subtle differences between

57

degraded and non-degraded tapes. Therefore, it can identify tapes starting to degrade before signs are apparent during playback. Playability testing does not identify a tape as degraded if it can fully play without squealing, friction, or shedding onto the player guides. It is possible that additives, such as lubricants, will allow a degraded tape to play. It is also probable that tapes do not degrade homogeneously, nor with a pattern to the starting point, since there was no pattern to which number replicate was more likely to be misclassified with the QDA model. Based on this data, it appears that random spots along the tape can show signs of degradations via ATR FT-IR analysis. There is no evidence that one region of the tape, (*e.g.,* the beginning or outer most region) is more likely to start degrading before another region (*e.g.,* the end or inner most region when on the reel). Ester hydrolysis is autocatalytic, so once degradation occurs in a location, the rate will increase, causing surrounding areas to degrade rapidly.[5, 21] Ultimately, it would be up to the preservationist to determine whether a tape should be baked before digitization or not based on the results of the model and the importance of the tape.

5. CONCLUSION

Magnetic audio tapes are a vital part of our cultural heritage. Unfortunately, many tapes manufactured with a polyester-urethane binder are degrading. It is agreed that the main source of this degradation arises from the hydrolysis of the ester component with moisture found in the air. This is

58

especially prominent for tapes housed in poor storage conditions that lack control of temperature and humidity.

This research has demonstrated the importance of having a robust data set when creating a calibration model, with three of the four tape sets classifying over 90%. The validation accuracy for the calibration set was 94.64%, while the test set consisting of the same brand and model tapes were correctly determined with 95.58% accuracy. The two test sets from the Library of Congress's Motion Picture Broadcasting and Recorded Sound Division and the University of Maryland had 90.79% and 78.85% accurate degradation determinations, respectively.

The majority of the misclassifications arose from tapes identified as playable through playability testing but non-playable via our model. This is preferred since these tapes would be baked to remove the excess water. It is less risky for a true playable tape to be baked than for a true non-playable tape to be played during digitization, since these tapes have a higher possibility of being destroyed. Overall, the preservationist would decide whether the results from our model would warrant a tape be sent for baking or digitizing. If a tape is important, one may want to obtain more spectra along the length of the tape to obtain an overall status of the tape's degradation level to determine how likely it would be for the tape to be destroyed during digitization. This model has demonstrated the utility of spectroscopy combined with chemometrics to triage a variety of PEU based magnetic audio tapes for historical preservation.

ACKNOWLEDGEMENTS

REFERENCES

1. Cassidy, B. M., et al. "Rapid and Non-destructive Identification of Degraded Polyester-urethane Magnetic Tape using ATR FT-IR Spectroscopy and Multivariate Statistics," *Analytical Chemistry*, **2015**, 87(18), 9265-9272.

2. Gelatt, R.; *The Fabulous Phonograph: From Tin Foil to High Fidelity.* Phyladelphia: J. B. Lippincott Company, 1955.

3. Engel, F. K. Oberlin Smith and the Invention of Magnetic Sound Recording. https://museumofmagneticsoundrecording.org/images/R2R/EngelOberlinSmith2006.pdf (accessed May 3, 2019).

4. Kimizuka, M. *Historical Development of Magnetic Recording and Tape Recorder*; rep.; National Museum of Nature and Science, 2012.

5. Edge, M.; Allen, N. S.; Hayes, M.; Jewitt, T. S. "Degradation of magnetic tape: Support and binder stability," *Polym. Degradation and Stability* **1993**, *39,* 207-214.

6. Hess R. L. "Tape Degradation Factors and Challenges in Predicting Tape Life," Association for Recorded Sound Collections Journal **2008**, 34, 240-274.

7. Thiébaut, B.; Vilmont, L.; Lavedrine, B. Report on Video and Audio Tape Deterioration Mechanisms and Considerations about Implementation of a Collection Condition Assessment Method; Report Deliverable D6.1; PrestoSpace: Paris, 2006.

8. Hobaica, S. "Analysis of audio magnetic tapes with sticky shed syndrome by ATR-FTIR," *J. Appl. Polym. Sci.*, **2012**, DOI: 10.1002/APP.38364

9. Kreiselmaier, K. W. Pigmentation of Magnetic Tapes. In *Pigment Handbook*; Patton, T. C., Wiley: New York, 1973; 315-329.

10. Thiébaut, B.; Vilmont, L.; Lavédrine, B. *J.* "Characterization of U-matic videotape deterioration by size exclusion chromatography and pyrolysis gas chromatography/mass spectrometry and the role of adipic acid," *Cultural Heritage* **2009**, *10*, 183-197.

11. Gómez-Sánchez, E.; Simon, S.; Koch, L-C.; Wiedmann, A.; Weber, T.; Mengel, M. "ATR-FTIR Spectroscopy for the Characterization of Magnetic Tape Materials," *E-Preservation Sci.* **2011**, *8*, 2-9.

12. Hill, V. E. "The Preservation of Sound Recordings," *Music Reference Services Quarterly* **2012**, *15* (2), 88-98.

13. A Public Trust at Risk: The Heritage Health Index Report on the State of America's Collections, Heritage Preservation, Inc., Washington, D.C., 2005.

14. Cuddihy, E. F. Aging of Magnetic Recording Tape. *IEEE Trans. Magn*. **1980**, *16,* 558−568.

15. Zelst, L. *Sticky Shed Syndrome: Characterization, Diagnosis, and Treatment,* Internal Library of Congress Report, March, 2008.

16. Schoonover, J. R.; Dattelbaum, D. M.; Osborn, J. C.; Bridgewater, J. S.; Kenney III, J. W. "Pressure-dependent Fourier transform infrared spectroscopy of a poly(ester urethane)," *Spectrochem. Acta A*. **2003**, *59*, 309-19.

17. Savitzky, A.; Golay, M. J. E. "Smoothing and differentiation of data by simplified least squares procedures," *Anal. Chem*. **1964**, *36*, 1627−39.

18. Barnes, R. J.; Dhanoa, M. S.; Lister, S. J. "Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra," *Appl. Spectrosc*. **1989**, *43*, 772-777.

19. Krzanowski, W. J. *Principles of Multivariate Analysis: A User's Perspective*, Rev. Ed.; Oxford University Press: New York, 2000.

20. Mahalanobis, P. C. On the Generalized Distance in Statistics. *Proc. Natl. Inst. Sci. India* **1936**, *2*, 49−55.

21. Brown, D. W.; Lowry, R. E.; Smith, L. E. Prediction of the Long Term Stability of Polyester-Based Recording Media; Report Number NBSIR 84-2988; National Bureau of Standards: Washington, DC, 1984.

Table 2.1. Confusion matrix of the playability prediction of the MBRS 95-tape, 1,900 spectra test set using the Chapter 1 QDA model.

QDA MBRS Test Set (1900 Spectra)	QDA Classification 1050/1900 = **55.26%**	
Playability Classification	Non-Playable	Playable
Non-Playable	962/980 = **98.16%**	18/980 = 1.84%
Playable	832/920 = 90.43%	88/920 = **9.57%**

Table 2.2. Confusion matrix of the QDA results for the 154-tape, 3,080-spectra validation of the calibration set after omitting features between 1196 cm^{-1} and 1151 cm^{-1} when compared to playability results.

QDA Calibration Model (3080 Spectra)	QDA Classification 2915/3080 = **94.64%**	
Playability Classification	Non-Playable	Playable
Non-Playable	1344/1380 = **97.39%**	36/1380 = 2.61%
Playable	129/1700 = 7.59%	1571/1700 = **92.41%**

Table 2.3. Confusion matrix of the QDA results for the 52-tape, 1,040-spectra test set after omitting features between 1196 cm^{-1} and 1151 cm^{-1} when compared to playability results.

QDA Test Set (1040 Spectra)	QDA Classification 994/1040 = 95.58%	
Playability Classification	Non-Playable	Playable
Non-Playable	457/500 = 91.40%	43/500 = 8.60%
Playable	3/540 = 0.56%	537/540 = 99.44%

Table 2.4. Confusion matrix of the QDA results for the 95-tape, 1,900-spectra MBRS test set after omitting features between 1196 cm^{-1} and 1151 cm^{-1} when compared to playability results.

QDA MBRS Test Set (1900 Spectra)	QDA Classification 1725/1900 = **90.79%**	
Playability Classification	Non-Playable	Playable
Non-Playable	938/980 = **95.71%**	42/980 = 4.29%
Playable	133/920 = 14.46%	787/920 = **85.54%**

Table 2.5. Confusion matrix of the QDA results for the 13-tape, 260-spectra UMD test set after omitting features between 1196 cm^{-1} and 1151 cm^{-1} when compared to playability results.

QDA UMD Test Set (260 Spectra)	QDA Classification 205/260 = **78.85%**	
Playability Classification	Non-Playable	Playable
Non-Playable	113/120 = **94.17%**	7/120 = 5.83%
Playable	48/140 = 34.29%	92/140 = **65.71%**
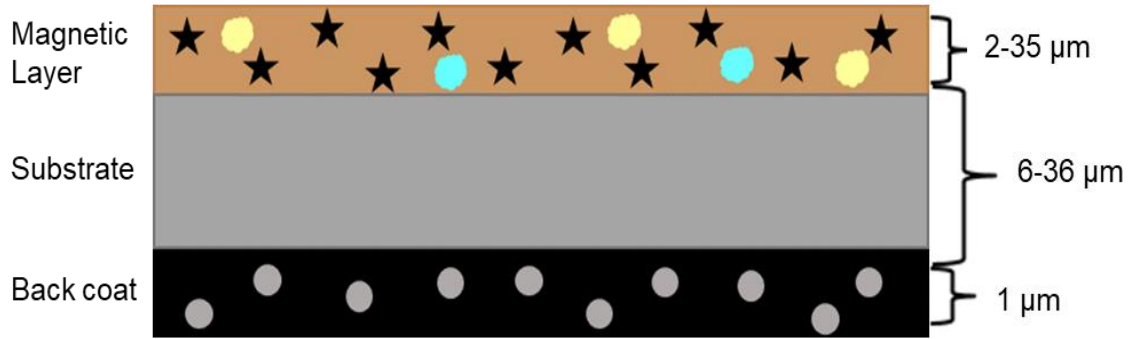
Figure 2.1. Cross-section of magnetic audio tape consisting from top down, a PEU binder layer, substrate, and optional back-coat. The binder contains magnetic particles (depicted by the black stars) and additives (represented by blue and yellow particles) such as lubricants and pigments. The substrate is usually PET, while the back-coat is carbon particles embedded in a polyester or polyether urethane.
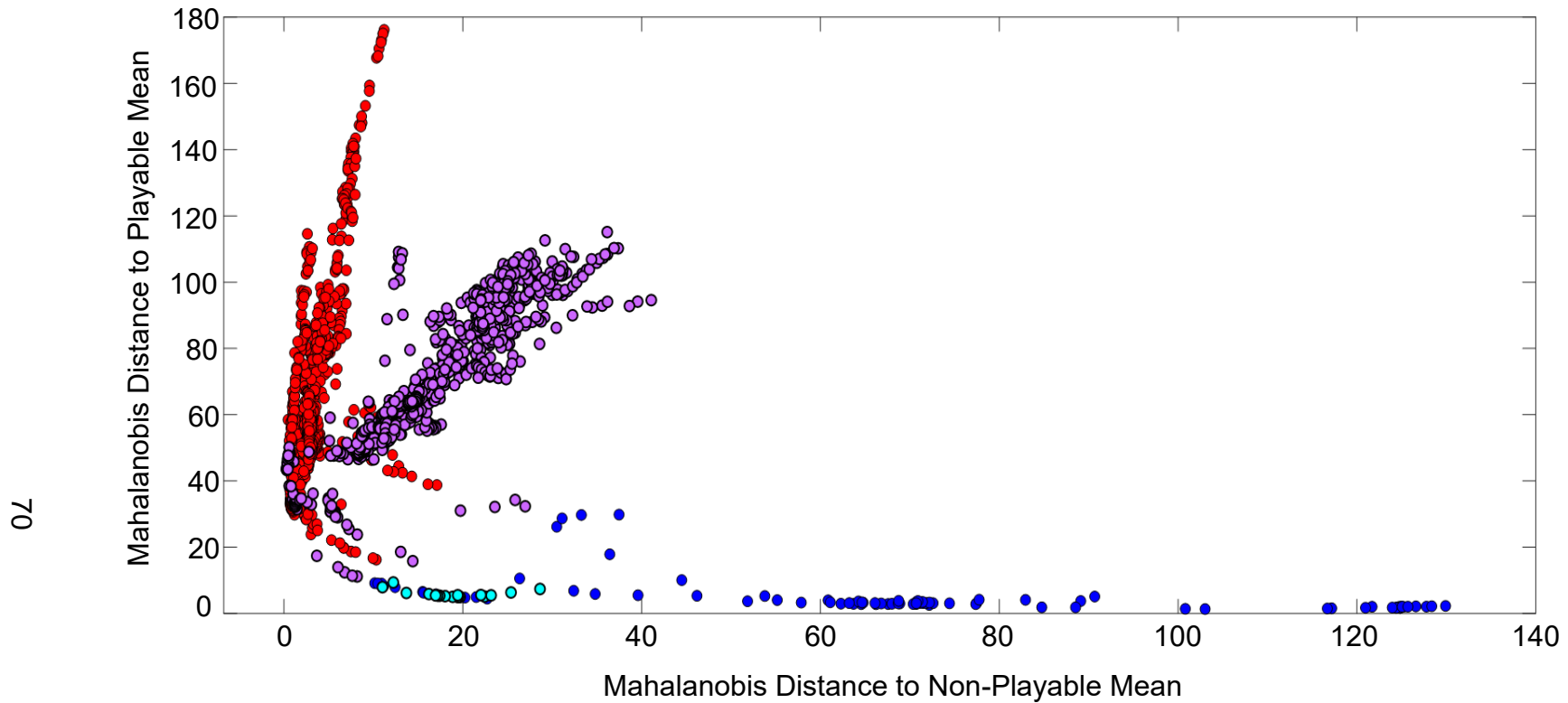
Figure 2.2. QDA response for the playability of the 1,900 spectra, 95-tape MBRS test set when the Chapter 1 model was used. The 962 red data points represent correctly classified non-playable tapes determined through playability testing, while the 88 blue points are correctly predicted playable tapes. The 18 cyan and 832 purple points are incorrectly predicted playable and non-playable spectra, respectively.
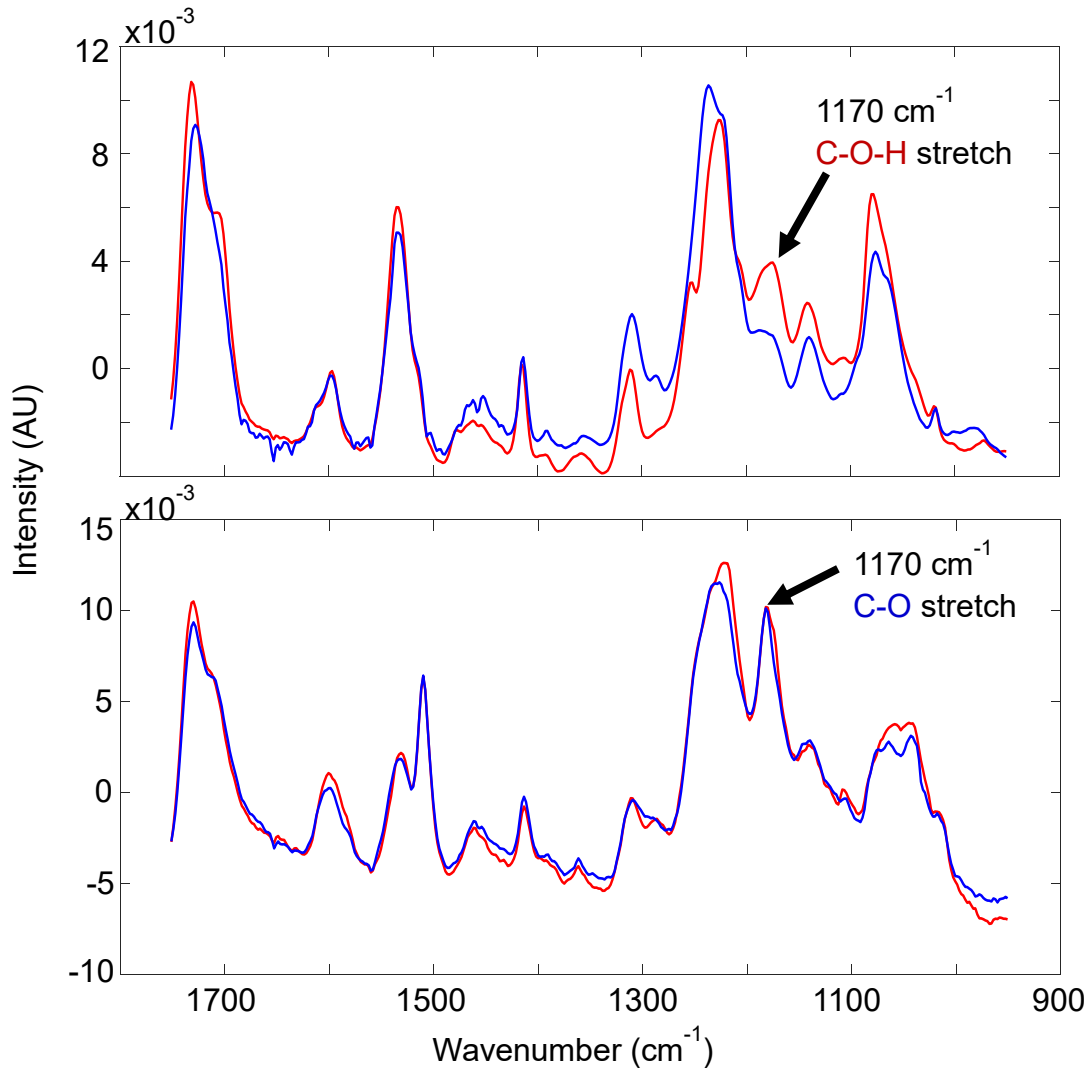
Figure 2.3. Top: Averaged playable (blue) and non-playable (red) WPT Ampex 406 spectra. Bottom: Averaged spectra of Scotch 3M 175 playable (blue) and non-playable (red) tapes. The arrows points to the feature around 1170 cm^{-1} found in only non-playable Ampex tapes but both playable and non-playable Scotch tape spectra.
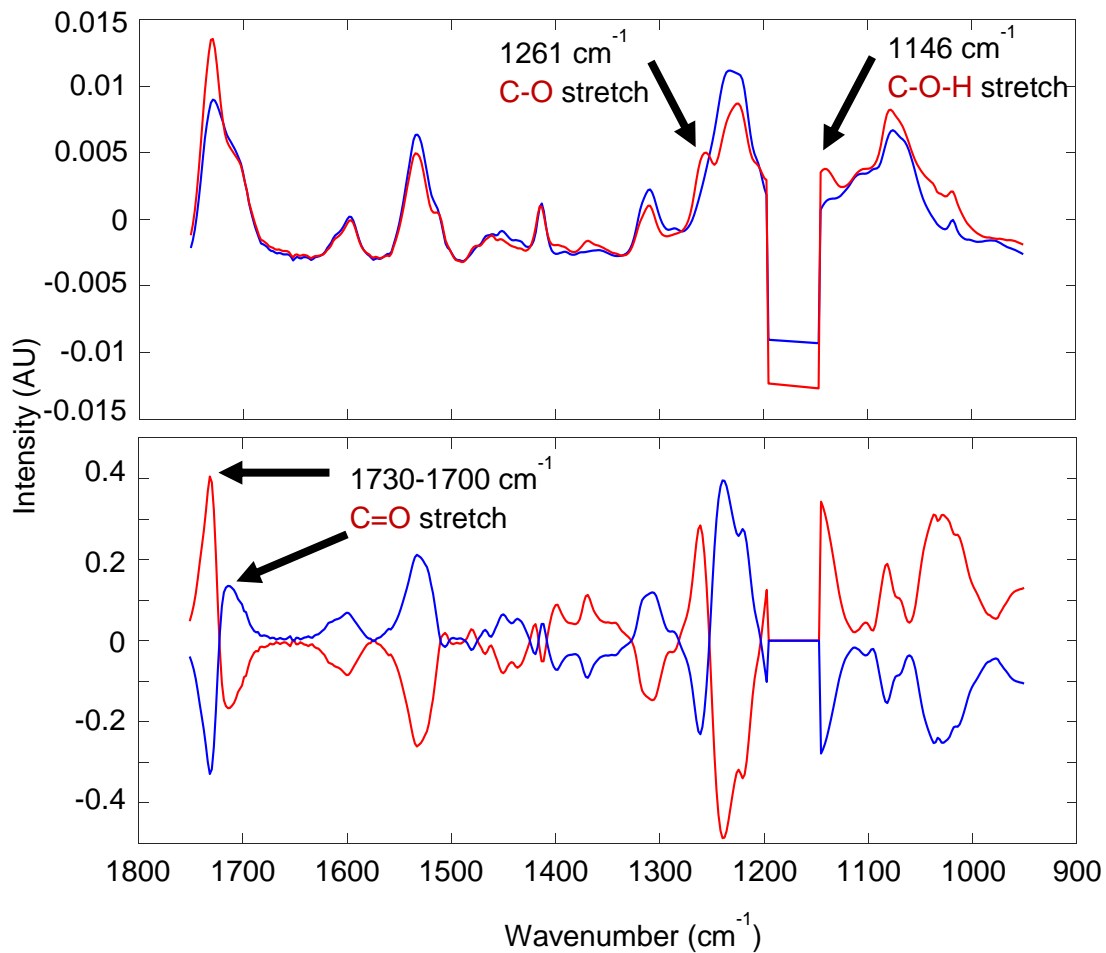
Figure 2.4. Absorbance spectra from the magnetic audio tape samples in the calibration set with the features from 1196 cm^{-1} to 1151 cm^{-1} omitted. The top shows the background corrected spectra for playable (blue) and non-playable (red) tapes, while the bottom demonstrates the enhanced difference after preprocessing with Savitzky-Golay smoothing, standard normal variate transform, and mean centering spectra. Arrows point to the 1261 cm^{-1} C-O stretch, 1146 cm^{-1} C-OH stretch, and 1730-1700 cm^{-1} C=O stretch.
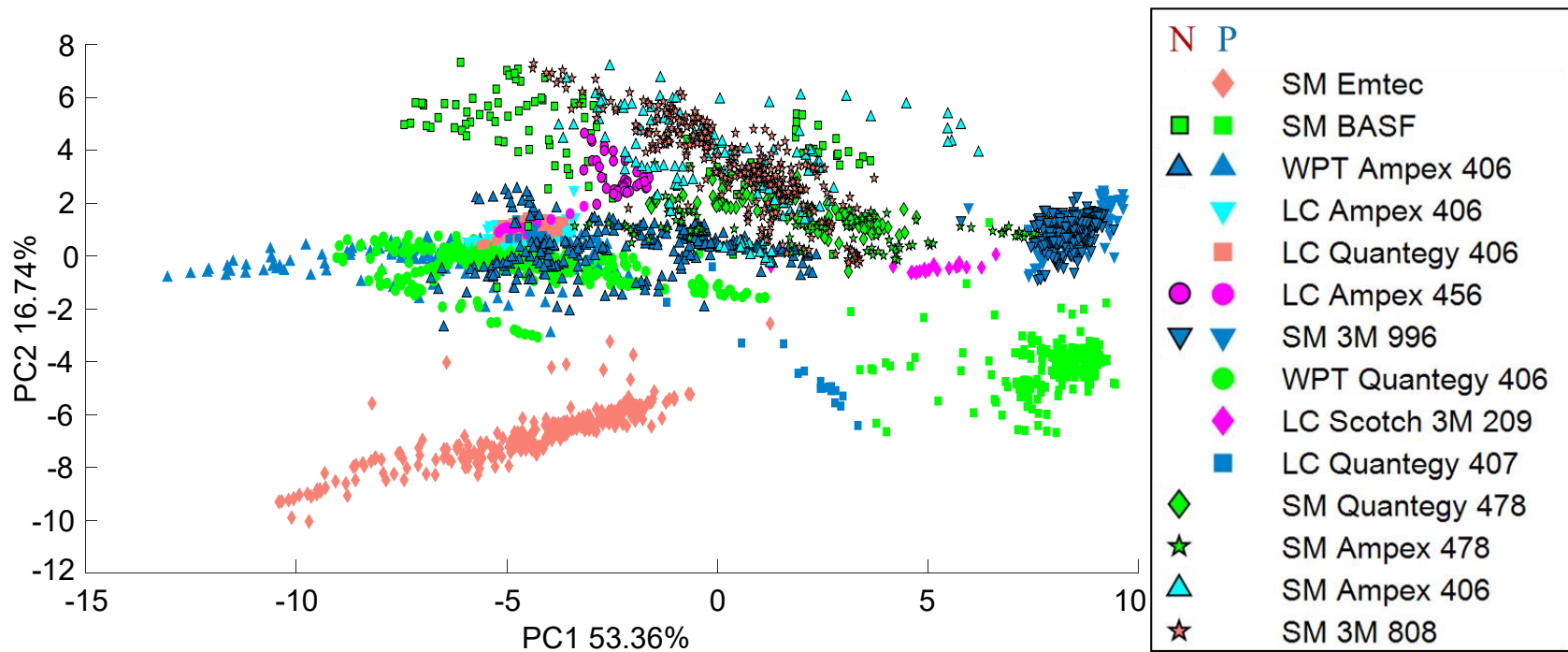
Figure 2.5. PCA projection of the preprocessed spectra in the calibration set with PC 1 accounting for 53.36% variability and PC 2 encompassing 16.74%. The brand and model information is depicted by the varying colored shapes, as indicated by the key. Playability information is denoted using a black outline on the data points belonging to a non-playable tape as determined during playability testing.
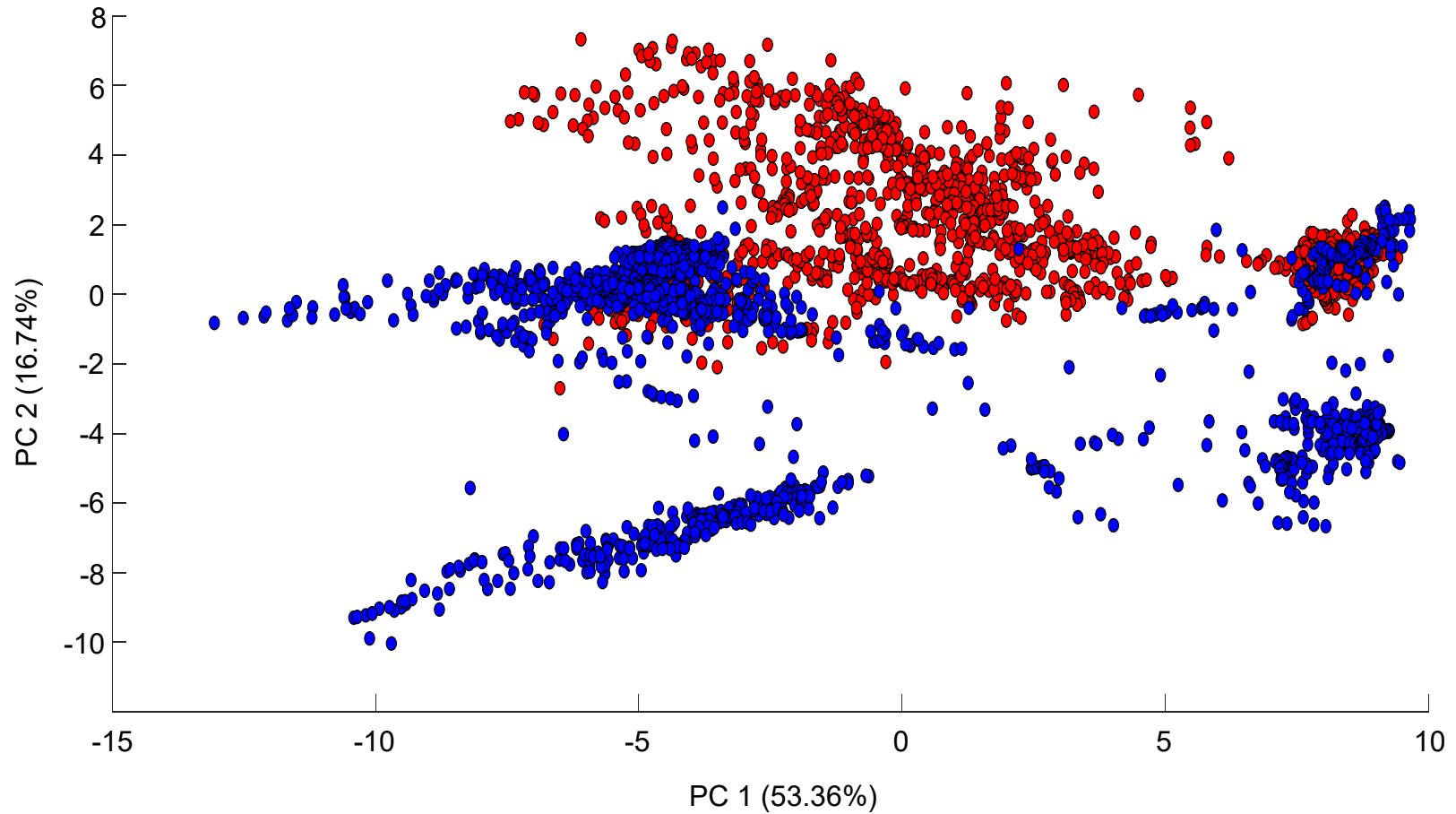
Figure 2.6. Projection of the truncated 3,080-calibration set spectra into the space containing PC 1 and PC 2. PC 1 accounts for 53.36% variability, while PC 2 encompasses 16.74%. The 1,700 playable spectra are represented by the blue data points and the 1,380 non-playable spectra are shown with red data points.
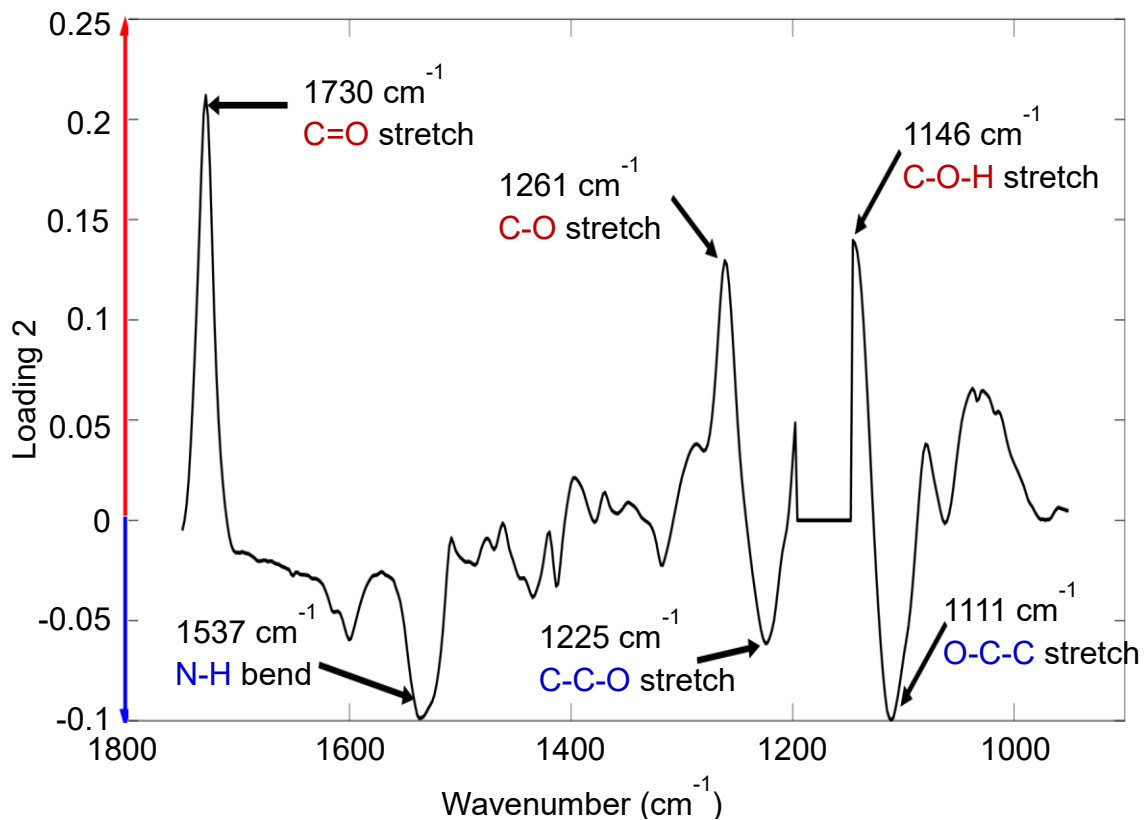
Figure 2.7. Principal component 2 loading showing that the projection along PC 2 was due to the indicated wavenumbers. Playable tapes were found in the negative region of the PCA plot due to exhibiting features found in the negative direction of the loading: 1537 cm[-1], 1225 cm[-1], and 1111 cm[-1]. Non-playable tapes were projected into the positive region of the PCA plot due to having the features found in the positive direction of the loading: 1730 cm[-1], 1261 cm[-1], and 1146 cm[-1].
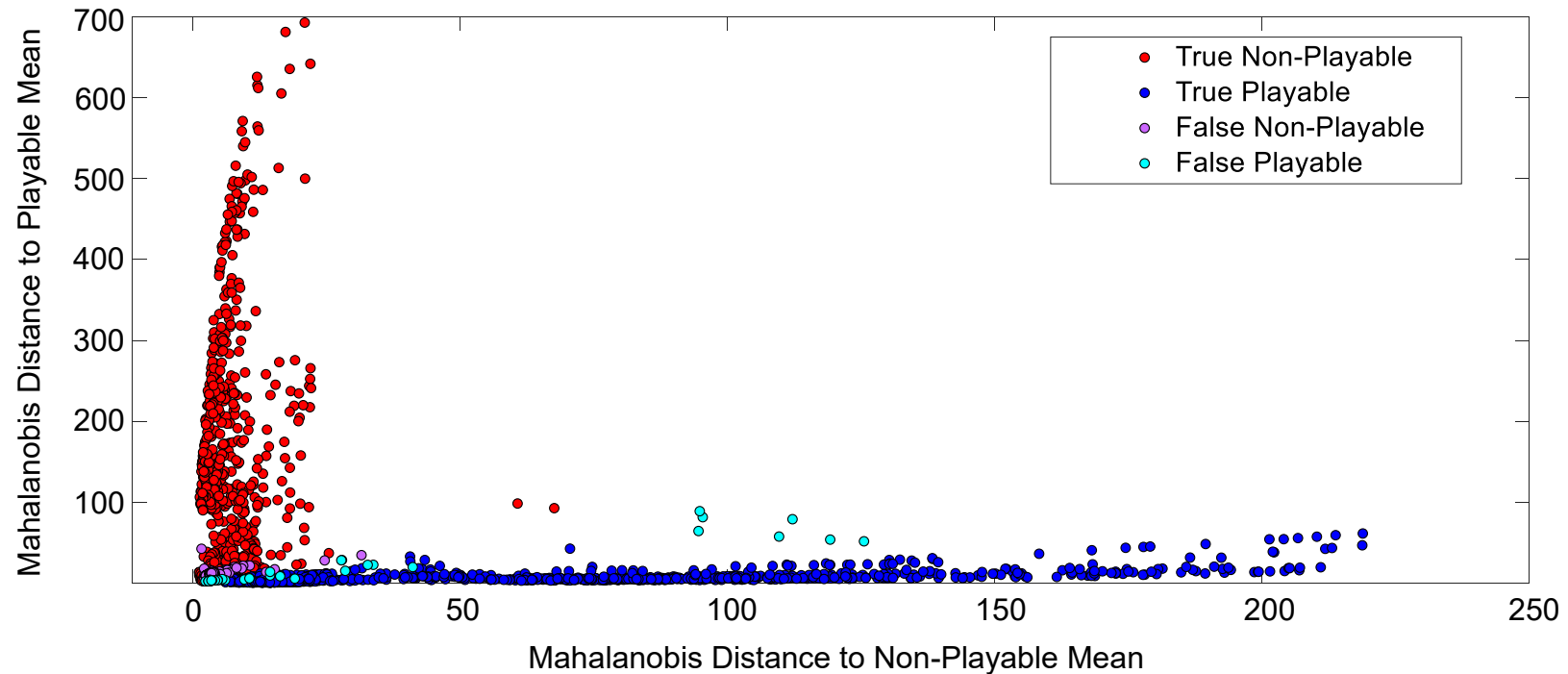
Figure 2.8. Mahalanobis distance from each spectra to the non-playable and playable class means for the 3,080 spectra calibration set. The 1,571 blue data points represent playable spectra correctly classified as playable while the 129 purple data points represent playable spectra classified as non-playable through the validated QDA model. The 1,344 red points were correctly classified as non-playable, while the 36 cyan were identified as playable by the QDA model but non-playable through playability testing.
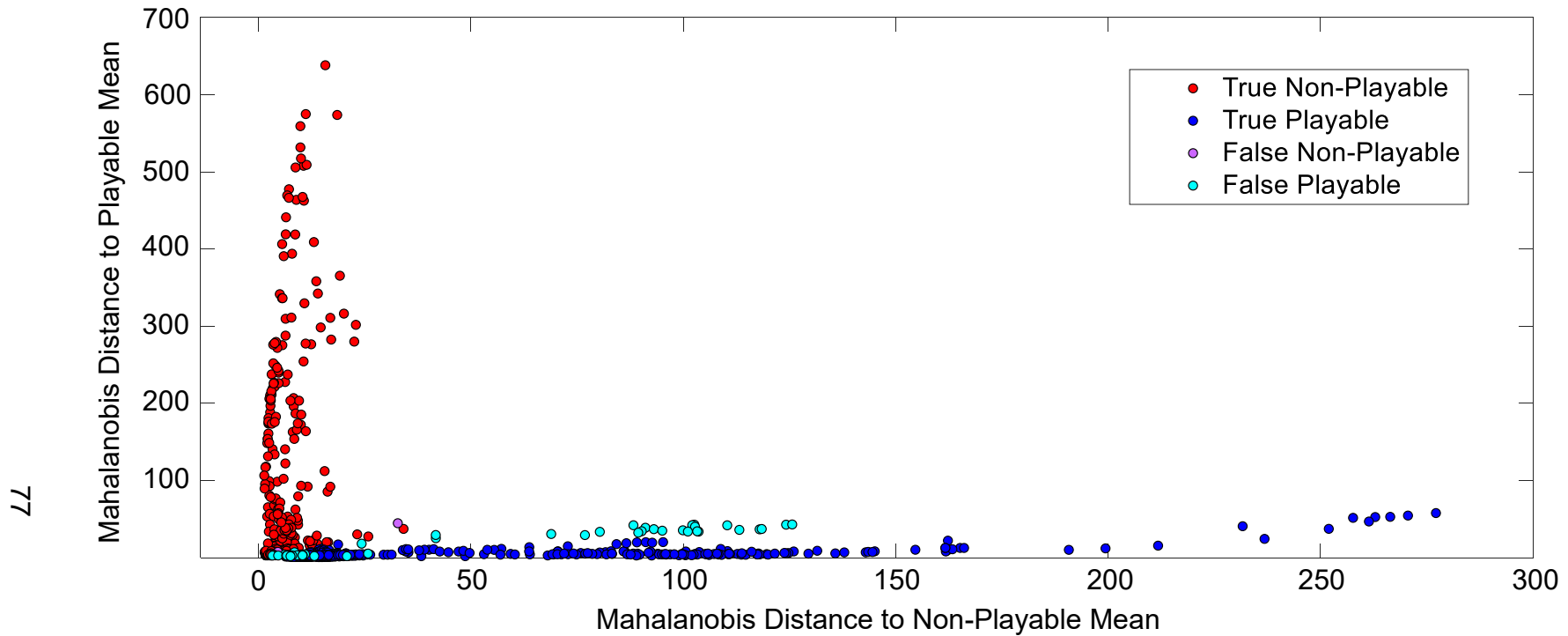
Figure 2.9. Results for the 52-tape, 1,040-spectra test set showing the Mahalanobis distance from each data point to the non-playable and playable class centroids. The 457 red data points indicate correctly classified non-playable spectra by the QDA model when compared to the playability testing results. The 43 cyan spectra were incorrectly classified as playable through the QDA model but non-playable by playability testing. The 537 blue and 3 purple points are the correctly classified playable and misclassified non-playable spectra using the QDA model.
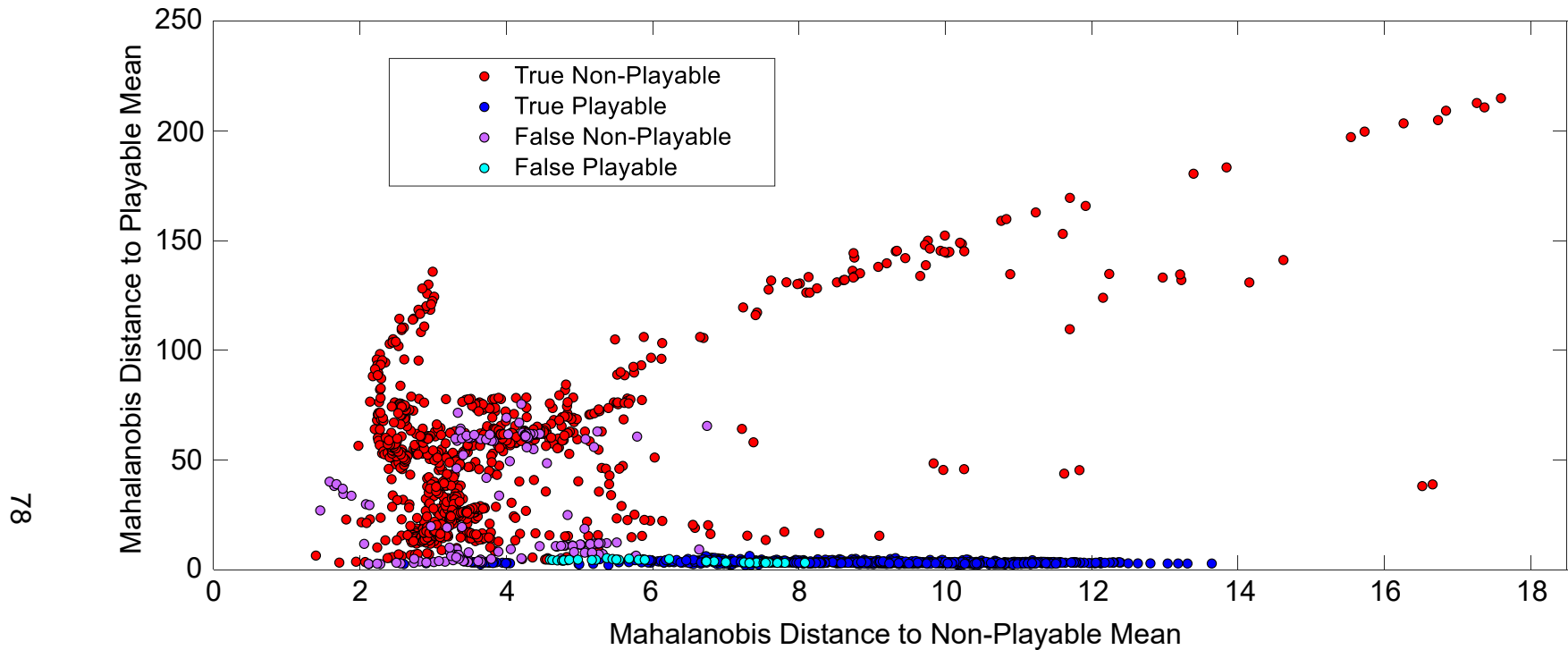
Figure 2.10. QDA predicted results for the 95-tape, 1,900-spectra MBRS test set. The 938 red data points represent non-playable spectra correctly classified, while the cyan data points show the 42 spectra misclassified as non-playable. The 787 blue points indicate the correctly classified playable spectra, with the 133 incorrectly classified non-playable spectra shown in purple.
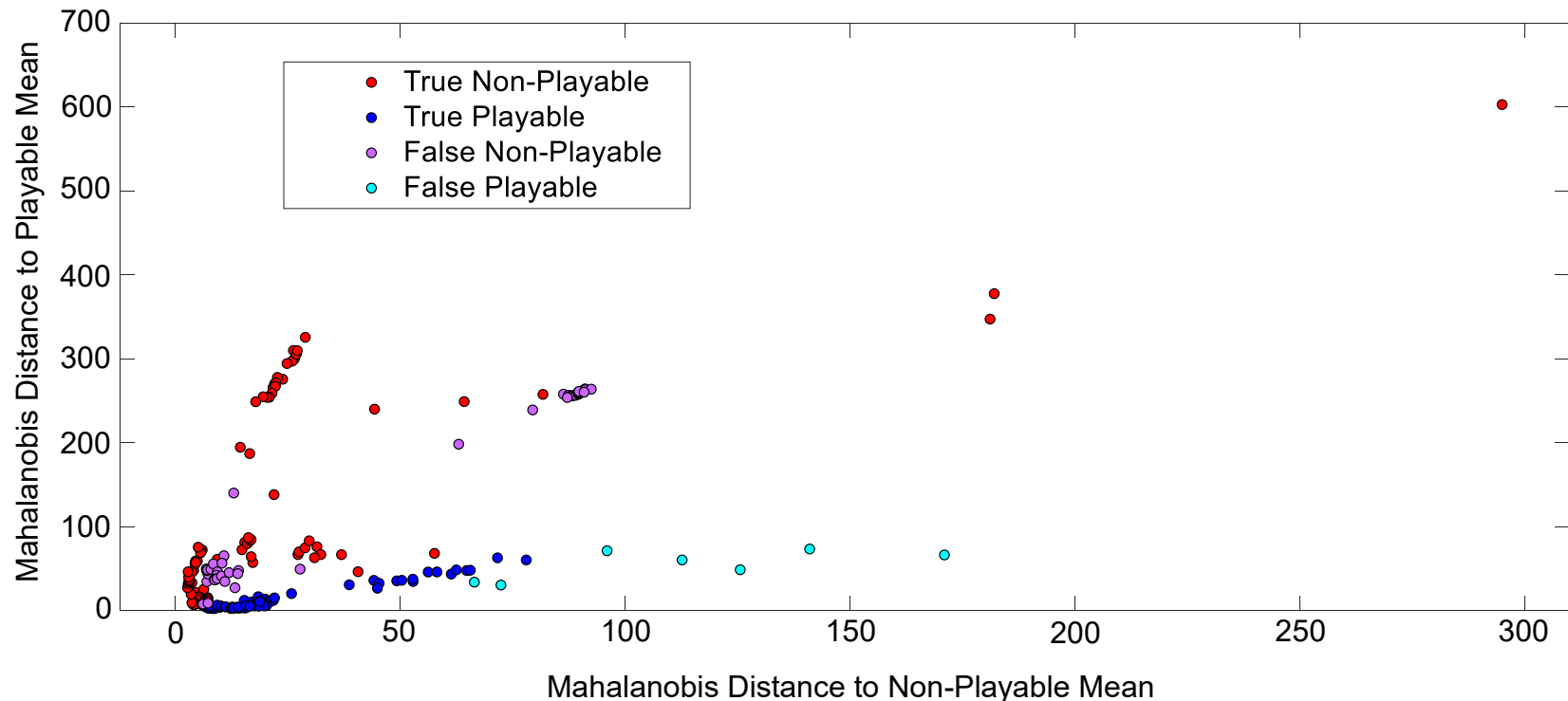
Figure 2.11. QDA response for the 260-spectra UMD test set. The Mahalanobis distance from each spectra to the playable and non-playable calibration set class means were calculated. The 92 blue data points represent playable spectra correctly classified as playable while the 48 purple data points represent playable spectra classified as non-playable through the validated QDA model. The 113 red points were correctly classified as non-playable but the 7 cyan were identified as playable by the QDA model.

BIBLIOGRAPHY

A Public Trust at Risk: The Heritage Health Index Report on the State of America's Collections, Heritage Preservation, Inc., Washington, D.C., 2005.

Barnes, R. J.; Dhanoa, M. S.; Lister, S. J. "Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra," *Appl. Spectrosc*. **1989**, *43*, 772-777.

Brown, D. W.; Lowry, R. E.; Smith, L. E. Prediction of the Long Term Stability of Polyester-Based Recording Media; Report Number NBSIR 84-2988; National Bureau of Standards: Washington, DC, 1984.

Cassidy, B. M. et al "Rapid and Non-destructive Identification of Degraded Polyester-urethane Magnetic Tape using ATR FT-IR Spectroscopy and Multivariate Statistics," *Analytical Chemistry*, **2015**, 87(18), 9265-9272.

Cuddihy, E. F. Hygroscopic Properties of Magnetic Recording Tape. *IEEE Trans. Magn.,* 1976, 12(2), 126-135.

Edge, M.; Allen, N. S.; Hayes, M.; Jewitt, T. S. "Degradation of magnetic tape: Support and binder stability," *Polym. Degradation and Stability* **1993**, *39,* 207-214.

Engel, F. K. Oberlin Smith and the Invention of Magnetic Sound Recording. https://museumofmagneticsoundrecording.org/images/R2R/EngelOberlinSmith2006.pdf (accessed May 3, 2019).

Fuenffinger, N. C. Optical Spectroscopy and Chemometrics for Discriminations of Dyed Textile Fibers and Magnetic Audio Tapes. Ph.D. Dissertation, University of South Carolina, Columbia, SC, 2015.

Gelatt, R.; *The Fabulous Phonograph: From Tin Foil to High Fidelity.* Phyladelphia: J. B. Lippincott Company, 1955.

Gómez-Sánchez, E.; Simon, S.; Koch, L-C.; Wiedmann, A.; Weber, T.; Mengel, M. "ATR-FTIR Spectroscopy for the Characterization of Magnetic Tape Materials," *E-Preservation Sci.* **2011**, *8*, 2-9.

Hess R. L. "Tape Degradation Factors and Challenges in Predicting Tape Life," Association for Recorded Sound Collections Journal **2008**, 34, 240-274.

Hill, V. E. "The Preservation of Sound Recordings," *Music Reference Services Quarterly* **2012**, *15* (2), 88-98.

Hobaica, S. "Analysis of audio magnetic tapes with sticky shed syndrome by ATR-FTIR," *J. Appl. Polym. Sci.*, **2012**, DOI: 10.1002/APP.38364

Kimizuka, M. *Historical Development of Magnetic Recording and Tape Recorder*; rep.; National Museum of Nature and Science, 2012.

Kreiselmaier, K. W. Pigmentation of Magnetic Tapes. In *Pigment Handbook*; Patton, T. C., Wiley: New York, 1973; 315-329.

Krzanowski, W. J. *Principles of Multivariate Analysis: A User's Perspective*, Rev. Ed.; Oxford University Press: New York, 2000.

Mahalanobis, P. C. On the Generalized Distance in Statistics. *Proc. Natl. Inst. Sci. India* **1936**, *2*, 49−55.

Morgan, S. L.; Bartick, E. G. Discrimination of Forensic Analytical Chemical Data Using Multivariate Statistics. *Forensic Analysis on the Cutting Edge: New Methods for Trace Evidence Analysis*, Robert D. Blackledge Ed.; John Wiley & Sons, New York, 2007, 331-372.

Rencher, A. C. *Methods of Multivariate Analysis,* 2nd Ed.; John Wiley & Sons: New York, NY, 2002.

Richardson C. A. Process for Restoring Magnetic Recording Tape Damaged by Sticky Shed Syndrome. U.S. Patent 6,797,072, Sep. 28, 2004.

Savitzky, A.; Golay, M. J. E. "Smoothing and differentiation of data by simplified least squares procedures," *Anal. Chem*. **1964**, 36, 1627–39.

Schoonover, J. R.; Dattelbaum, D. M.; Osborn, J. C.; Bridgewater, J. S.; Kenney III, J. W. "Pressure-dependent Fourier transform infrared spectroscopy of a poly(ester urethane)," *Spectrochem. Acta A*. **2003**, *59*, 309-19.

Thiébaut, B.; Vilmont, L.; Lavédrine, B. "Characterization of U-matic videotape deterioration by size exclusion chromatography and pyrolysis gas chromatography/mass spectrometry and the role of adipic acid," *Cultural Heritage* **2009**, *10*, 183-197.

Thiébaut, B.; Vilmont, L.; Lavedrine, B. Report on Video and Audio Tape Deterioration Mechanisms and Considerations about Implementation of a Collection Condition Assessment Method; Report Deliverable D6.1; PrestoSpace: Paris, 2006.

Ward, J. Hierarchical Grouping to Optimize an Objective Function, *J. Amer. Statist. Assoc.* **1963**, 58, 236-244.

Zelst, L. *Sticky Shed Syndrome: Characterization, Diagnosis, and Treatment,* Internal Library of Congress Report, March, 2008.