

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Elisa Anne Liszewski

Entitled INSTRUMENTAL AND STATISTICAL METHODS FOR THE COMPARISON OF
CLASS EVIDENCE

For the degree of Master of Science

Is approved by the final examining committee:

John Goodpaster

Chair

Jay Siegel

Sapna Deo

To the best of my knowledge and as understood by the student in the *Research Integrity and Copyright Disclaimer (Graduate School Form 20)*, this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Approved by Major Professor(s): John Goodpaster

Approved by: John Goodpaster

Head of the Graduate Program

07/09/10

Date

**PURDUE UNIVERSITY
GRADUATE SCHOOL**

Research Integrity and Copyright Disclaimer

Title of Thesis/Dissertation:

INSTRUMENTAL AND STATISTICAL METHODS FOR THE COMPARISON OF CLASS
EVIDENCE

For the degree of Master of Science

I certify that in the preparation of this thesis, I have observed the provisions of *Purdue University Teaching, Research, and Outreach Policy on Research Misconduct (VIII.3.1)*, October 1, 2008.*

Further, I certify that this work is free of plagiarism and all materials appearing in this thesis/dissertation have been properly quoted and attributed.

I certify that all copyrighted material incorporated into this thesis/dissertation is in compliance with the United States' copyright law and that I have received written permission from the copyright owners for my use of their work, which is beyond the scope of the law. I agree to indemnify and save harmless Purdue University from any and all claims that may be asserted or that may arise from any copyright violation.

Elisa Anne Liszewski

Printed Name and Signature of Candidate

07/09/10

Date (month/day/year)

*Located at http://www.purdue.edu/policies/pages/teach_res_outreach/viii_3_1.html

INSTRUMENTAL AND STATISTICAL METHODS FOR THE COMPARISON OF
CLASS EVIDENCE

A Thesis

Submitted to the Faculty

of

Purdue University

by

Elisa Anne Liszewski

In Partial Fulfillment of the
Requirements for the Degree
of
Master of Science

August 2010

Purdue University

Indianapolis, Indiana

For my loving and supportive family. Mom and Dad, you have encouraged me to pursue my dreams throughout my entire life, and I appreciate all you have done for me. My brother, Tony, for pushing me to reach my fullest potential. My fiancé Tom – you have been there for me through good times and bad, and I love you.

ACKNOWLEDGMENTS

I would like to thank Dr. John V. Goodpaster, my advisor and mentor, for assisting me throughout my graduate career. Your guidance and support have enabled me to successfully achieve my goals, and I am forever grateful for this experience. Sincere thanks also goes to Dr. Jay Siegel for giving me the opportunity to pursue forensic science and for all of your direction in my academic endeavors. In addition, the Microanalysis Unit at the Indiana State Police Laboratory has contributed greatly to my research, through my internship and the entirety of my studies. Special thanks also goes to Dr. Simon Lewis for all his help with the clear coat research. Also, I would like to thank the XLSTAT support team and Scott Ramos from Infometrix Inc. for assisting me with technical issues that arose during my research, as well as Tom Klaas from Testfabrics Inc. for providing us with dyed cotton exemplars. I am also grateful for the financial support provided by National Institute of Justice's Midwestern Forensics Resource Center (MFRC) grant. Lastly, I would like to extend my deepest appreciation to all those who have positively impacted my research.

TABLE OF CONTENTS

	Page
LIST OF TABLES	vi
LIST OF FIGURES	viii
LIST OF ABBREVIATIONS	xi
ABSTRACT	xiv
CHAPTER 1. INTRODUCTION	1
1.1. Chemometric Techniques and their Application to Forensic Science.....	2
1.1.1. Preprocessing Techniques	4
1.1.2. Agglomerative Hierarchical Clustering.....	6
1.1.3. Principal Components Analysis	10
1.1.4. Discriminant Analysis.....	11
CHAPTER 2. AUTOMOTIVE CLEAR COATS.....	14
2.1. Review of Analysis of Automotive Clear Coats.....	15
2.2. Materials and Methods	18
2.2.1. Instrumental Analysis.....	18
2.2.2. Data Analysis.....	21
2.3. Results and Discussion	22
2.3.1. Statistical Results	22
2.3.2. External Validation.....	32
2.3.3. Formation of Classes.....	34
2.3.4. Limitations to the Study	37
2.4. Conclusions	38
CHAPTER 3. FIBER DYE ANALYSIS	40
3.1. Review of Analysis of Dyed Textile Fibers.....	45
3.2. Materials and Methods	50
3.2.1. Instrumental Analysis.....	50
3.2.2. Data Analysis.....	53
3.3. Results and Discussion - Part I: Testfabrics Fibers Analyzed at IUPUI	54
3.3.1. Statistical Results	54
3.3.2. External Validation.....	65
3.3.3. Limitations to the Study	67
3.3.4. Conclusions	67
3.4. Results and Discussion - Part II: All Dyes Analyzed - IUPUI vs. ISP	69

	Page
3.4.1. Statistical Results of IUPUI Fiber Analysis.....	69
3.4.2. External Validation.....	84
3.4.3. Statistical Results of ISP Fiber Analysis	85
3.4.4. Limitations to the Study	97
3.4.5. Conclusions	97
CHAPTER 4. PLASTICS AND POLYMERS.....	104
4.1. Pyrolysis Gas Chromatography/Mass Spectrometry and its Use in Forensic Science	104
4.2. Materials and Methods	106
4.2.1. Instrumental Analysis.....	106
4.2.2. Data Analysis.....	109
4.3. Results and Discussion	120
4.3.1. Statistical Results	120
4.3.2. External Validation.....	130
4.4. Conclusions	131
CHAPTER 5. FUTURE DIRECTIONS.....	133
LIST OF REFERENCES	142
APPENDICES	
Appendix A. Clear Coats (Averaged Spectra).	154
Appendix B. Global Dye Averages	173
Appendix C. Polymer Standards.....	221

LIST OF TABLES

Table	Page
Table 2.1 Eigenvalues and Variability Associated with each Principal Component (PC).....	30
Table 2.2 Confusion Matrix for the Cross-Validation Results of DA.....	32
Table 2.3 Confusion Matrix for the External Validation Results of the Supplemental Data from DA	33
Table 3.1 The Various Types of Dye Classes.....	44
Table 3.2 List of the Dyed Exemplars from Testfabrics, Inc.	51
Table 3.3 List of the Dyed Exemplars from Dr. Stephen Morgan of the University of South Carolina.....	52
Table 3.4 Confusion Matrix for the Cross-Validation Results from DA (Three Classes)	63
Table 3.5 Confusion Matrix for the Cross-Validation Results from DA (Six Classes)	65
Table 3.6 Confusion Matrix for the External Validation Results of the Supplemental Data from DA using Three Classes	66
Table 3.7 Confusion Matrix for the External Validation Results of the Supplemental Data from DA using Six Classes	66
Table 3.8 Class Formation from the AHC Dendrogram of new IUPUI fibers using Three Classes.....	70
Table 3.9 Class Formation from the AHC Dendrogram of new IUPUI fibers using Seven Classes	71
Table 3.10 Confusion Matrix for the Cross-Validation Results from DA (Three Classes)	80
Table 3.11 Confusion Matrix for the Cross-Validation Results from DA (Seven Classes)	81
Table 3.12 Confusion Matrix for the Cross-Validation Results from DA (Twelve Classes).....	83
Table 3.13 Confusion Matrix for the External Validation Results of the Supplemental Data from DA using Three Classes	84
Table 3.14 Confusion Matrix for the External Validation Results of the Supplemental Data from DA using Seven Classes	85
Table 3.15 Class Formation from the AHC Dendrogram of ISP fibers using Three Classes.	86
Table 3.16 Confusion Matrix of the Cross-Validation Results from DA (Three AHC-designated Classes)	94

Table	Page
Table 3.17 Confusion Matrix for the Cross-Validation Results from DA (Twelve Classes)	96
Table 3.18 Dye Structures	101
Table 3.19 Dye Structures	102
Table 3.20 Dye Structures	103
Table 4.1 The Plastic Recycling Numbers and Polymer They Represent	107
Table 4.2 Specific Polymers Analyzed.	108
Table 4.3 Confusion Matrix for the Cross-Validation Results from DA	127
Table 4.4 Confusion Matrix for the Cross-Validation Results from DA using Nine Classes.	129
Table 4.5 Confusion Matrix for the External Validation Results of the Supplemental Data from DA	131

LIST OF FIGURES

Figure		Page
Figure 2.1	Absorbance Spectra of Two Central Objects with Lensbond as the Mounting Medium.....	19
Figure 2.2	Absorbance Spectra of the Same Two Central Objects Without Lensbond.....	19
Figure 2.3	Baseline Corrected, Normalized, and Offset Absorbance Spectra of Five Scans for a Clear Coat from a 1993 Chevy Lumina.....	21
Figure 2.4	Dendrogram from AHC of the Averages of each Clear Coat Sample.....	23
Figure 2.5	Central Objects of the Three Clusters from the Dendrogram.....	24
Figure 2.6	Observations Plot from PCA of Clear Coats.....	25
Figure 2.7	Factor Loadings Plot of the First Two PCs.....	26
Figure 2.8	Significant Factor Loadings Overlayed on the Central Objects Plot for Clear Coats.....	28
Figure 2.9	Observations Plot from DA.....	31
Figure 2.10	Percent Accuracy for each DA testing technique versus Varying Number of Classes.....	34
Figure 2.11	Samples of the Same Make and Model but Different Year Placed in Different Classes.....	35
Figure 2.12	Samples of the Same Make and Model but Different Year Placed in the Same Class.....	36
Figure 2.13	Samples of the Same Make, Model, and Year.....	36
Figure 3.1	Analytical Techniques Applied to the Analysis of Dyed Fibers.....	46
Figure 3.2	Background Subtracted, Normalized, and Offset Absorbance Spectra of Ten Scans for a Fiber Exemplar Dyed with Direct Red C-380.....	54
Figure 3.3	Dendrogram from AHC of Fibers A – F analyzed at IUPUI.....	56
Figure 3.4	Central Objects of the Three Clusters from the Dendrogram.....	57
Figure 3.5	Observations Plot from PCA.....	58
Figure 3.6	Factor Loadings Plot of the First Two PCs.....	59
Figure 3.7	Significant Factor Loadings Overlayed on the Central Objects Plot.....	61

Figure	Page
Figure 3.8 Observations Plot from DA using Three Classes.....	63
Figure 3.9 Observations Plot from DA using Six Classes	64
Figure 3.10 AHC Dendrogram of new IUPUI fibers (12) using Three Classes.....	70
Figure 3.11 AHC Dendrogram of new IUPUI fibers (12) using Seven Classes.....	71
Figure 3.12 Central Objects of the Three Clusters from the Dendrogram of IUPUI New Fibers	72
Figure 3.13 PCA Observations plot of IUPUI new data using Three Classes from AHC	74
Figure 3.14 PCA Observations plot of IUPUI new data using Seven Classes from AHC	74
Figure 3.15 Factor Loadings Plot of the First Two PCs	75
Figure 3.16 Significant Factor Loadings Overlaid on the Central Objects Plot	77
Figure 3.17 Observations Plot from DA using the Three Designated AHC Classes	79
Figure 3.18 Observations Plot from DA using the Seven Designated AHC Classes	81
Figure 3.19 Observations Plot from DA using 12 Classes (Each Dye is its Own Class)	82
Figure 3.20 AHC Dendrogram of all ISP Fibers.....	86
Figure 3.21 Central Objects of the Three Clusters from the Dendrogram of ISP Fibers	87
Figure 3.22 PCA Observations plot of ISP data using Three Classes from AHC	88
Figure 3.23 Factor Loadings Plot of the First Two PCs	89
Figure 3.24 Significant Factor Loadings Overlayed on the Central Objects Plot for ISP Data.....	91
Figure 3.25 Observations Plot from DA using the Three Designated AHC Classes	93
Figure 3.26 Observations Plot from DA Using Twelve Classes (Each Dye is its Own Class)	95
Figure 4.1 Chromatogram of HDPE	110
Figure 4.2 Chromatogram of LDPE.....	111
Figure 4.3 Chromatogram of HDPEblack.....	111
Figure 4.4 Chromatogram of Polypropylene	112
Figure 4.5 Chromatogram of PVC	112
Figure 4.6 Chromatogram of Other.....	113
Figure 4.7 Chromatogram of PETE.....	113
Figure 4.8 Chromatogram of Polystyrene	114
Figure 4.9 Chromatogram of PolystyreneRed.....	115
Figure 4.10 Raw Data of Three Replicates of LDPE	117

Figure	Page
Figure 4.11 Background corrected and Normalized Data for Three Replicates of LDPE.....	117
Figure 4.12 Aligned Data for Three Replicates of LDPE	118
Figure 4.13 Portion of a TIC of HDPE Showing All Time Points.....	119
Figure 4.14 Portion of a TIC of HDPE Showing Averaged Time Points.....	119
Figure 4.15 Dendrogram Produced Using AHC.....	121
Figure 4.16 Central Object of Class 1 (LDPE).....	121
Figure 4.17 Central Object of Class 2 (PVC).....	122
Figure 4.18 Central Object of Class 3 (PETE).....	122
Figure 4.19 Central Object of Class 4 (Polystyrene).....	123
Figure 4.20 Observations Plot from PCA of Plastics Expressed in Terms of the First Two PCs	124
Figure 4.21 PCA Observations Plot color-coded to show the four classes from AHC clustering.....	124
Figure 4.22 Observations Plot from DA	126
Figure 4.23 Observations Plot from DA Using Nine Classes in which Each Plastic is its Own Class	128

LIST OF ABBREVIATIONS

°	degree
x	magnification
µm	micrometer
3-D	three-dimensional
AHC	Agglomerative Hierarchical Clustering
C	Celsius
CCD	charge-coupled device
CE	capillary electrophoresis
CI	Colour Index
COW	correlation optimized warping
CV	canonical variate
DA	Discriminant Analysis
FTIR	fourier transform infrared spectroscopy
GC	gas chromatography
HDPE	high-density polyethylene
HPLC	high performance liquid chromatography
IR	infrared
ISP	Indiana State Police

LC	liquid chromatography
LDPE	low-density polyethylene
min	minute
mL	milliliter
mm	millimeter
MS	mass spectrometry
MSP	microspectrophotometry
NIR	near-infrared
NIST	National Institute of Standards and Technology
nm	nanometer
PC	principal component
PCA	Principal Components Analysis
PETE	polyethylene terephthalate
PLM	polarized light microscopy
PVC	polyvinyl chloride
Py	pyrolysis
SEM-EDS	scanning electron microscopy-energy dispersive
SWGMAAT	Scientific Working Group on Materials Analysis
TIC	Total Ion Chromatogram
TLC	thin layer chromatography
UV	ultraviolet
Vis	visible
VOC	volatile organic compound

XRD	x-ray diffraction
XRF	x-ray fluorescence

ABSTRACT

Liszewski, Elisa Anne. M.S., Purdue University, August, 2010. Instrumental and Statistical Methods for the Comparison of Class Evidence. Major Professor: John Goodpaster.

Trace evidence is a major field within forensic science. Association of trace evidence samples can be problematic due to sample heterogeneity and a lack of quantitative criteria for comparing spectra or chromatograms. The aim of this study is to evaluate different types of instrumentation for their ability to discriminate among samples of various types of trace evidence. Chemometric analysis, including techniques such as Agglomerative Hierarchical Clustering, Principal Components Analysis, and Discriminant Analysis, was employed to evaluate instrumental data. First, automotive clear coats were analyzed by using microspectrophotometry to collect UV absorption data. In total, 71 samples were analyzed with classification accuracy of 91.61%. An external validation was performed, resulting in a prediction accuracy of 81.11%. Next, fiber dyes were analyzed using UV-Visible microspectrophotometry. While several physical characteristics of cotton fiber can be identified and compared, fiber color is considered to be an excellent source of variation, and thus was examined in this study. Twelve dyes were employed, some being visually indistinguishable.

Several different analyses and comparisons were done, including an inter-laboratory comparison and external validations. Lastly, common plastic samples and other polymers were analyzed using pyrolysis-gas chromatography/mass spectrometry, and their pyrolysis products were then analyzed using multivariate statistics. The classification accuracy varied dependent upon the number of classes chosen, but the plastics were grouped based on composition. The polymers were used as an external validation and misclassifications occurred with chlorinated samples all being placed into the category containing PVC.

CHAPTER 1. INTRODUCTION

The aim of this study is to evaluate different types of instrumentation for their ability to discriminate among samples of various types of trace evidence. Chemometric analysis, including techniques such as Agglomerative Hierarchical Clustering, Principal Components Analysis, and Discriminant Analysis, was employed to evaluate instrumental data. Trace evidence is a major field within forensic science. Association of trace evidence samples can be problematic due to sample heterogeneity and a lack of quantitative criteria for comparing spectra or chromatograms. Therefore, in this project, automotive clear coats and red dyed cotton fibers were analyzed using microspectrophotometry (MSP) and their UV-visible spectra were evaluated using multivariate statistics. While several physical characteristics of cotton fibers can be identified and compared, fiber color is considered to be an excellent source of variation, and thus was examined in this study. Since automotive clear coats are colorless, their UV absorption characteristics were studied by microspectrophotometry. In addition, common plastic samples and other polymers were analyzed using pyrolysis-gas chromatography/mass spectrometry (Py-GC/MS), and their pyrograms were then analyzed using multivariate statistics. Overall, multivariate statistics can increase discrimination of samples as well as distinguishing between groups much more reliably than traditional visual examination of the data.

1.1. Chemometric Techniques and their Application to Forensic Science

The use of multivariate statistics has accelerated in recent years in forensic analyses. Forensic scientists are often tasked with identifying patterns in data as well as interpreting differences. Chemometrics has enabled this task to become more accurate and manageable. Its use in the trace evidence area is prominent and can be used with various types of evidence. For example, multivariate statistics has been applied to accelerants, document examination, inks, fibers, ammunition, glass, gunpowder, paint, nail polish, paint surfaces, and condom lubricants.¹ A complete review of chemometrics applied to trace evidence is beyond the scope of this thesis.

When trying to determine if known and unknown samples could have a common source, forensic chemists often rely only upon visual comparisons of complex chromatograms and other spectra. Because of this, examiners do not have any statistical basis for determining the value of the evidence in question. This is a concern for forensic laboratories regarding the reliability of comparisons such as dyed fibers, and the ability to compare samples in a quantitative way is desirable. Searching through databases, such as the Paint Data Query (PDQ), can assist in comparisons by having a large, readily available dataset with all factors about it known. Chemometrics could be readily applied to it. In addition, multivariate statistics could support issues raised in *Daubert v. Merrell Dow Pharmaceuticals* such as reliability and relevance of scientific evidence.¹ In addition, chemometrics can address some of the recommendations that were provided by the National Academy of Sciences (NAS) report on strengthening

forensic science. Specifically, chemometrics would address issues of accuracy and reliability in forensic science disciplines (Recommendation 3) as well as assist in research on human observer bias and sources of human error in forensic examinations (for example, visual analysis of data versus chemometric analysis of data) (Recommendation 5).²

The value of multivariate statistics has been recognized for many years. The fundamental idea of principal components analysis (PCA) was introduced by Pearson in 1901. Algorithms for computing principal components (PCs) were described by Hotelling in 1933. Mahalanobis established the multivariate distance that shares his name in 1936, and discriminant analysis (DA) first originated by Fisher in 1936.¹

In general, chemometrics is used for data reduction or structural simplification, sorting and grouping, investigation of the dependence among variables, prediction, or hypothesis construction and testing.³ Chemometrics can extract information from a large data set and thus reduce its complexity, and it can also assist in making accurate predictions about unknown samples. It uses all the information present in the data set but still maintains sensitivity to minor, significant features.¹ Further, chemometrics can be used to interpret results of forensic analysis particularly in areas where pattern recognition is involved. When working with chemometrics, replicate measurements of variables should be made as often as possible to increase the significance of differences found between samples as well as allow for experimental uncertainty.¹ Following pre-processing of data, three specific chemometrics techniques were utilized in this

study: Agglomerative Hierarchical Clustering (AHC), Principal Components Analysis (PCA), and Discriminant Analysis (DA).

1.1.1. Preprocessing Techniques

Preprocessing the data prior to performing multivariate statistical tests is typically required. Doing this can remove random noise and variation that might later affect interpretation. However, improper preprocessing may negatively influence the data; therefore, the techniques must be chosen carefully.

Smoothing the data can increase the signal-to-noise ratio if there is unnecessary noise. However, it can also cause distortions in peak height and width, as well as decrease resolutions. Therefore, smoothing must be done cautiously. Various smoothing techniques are possible. The running polynomial smooth fits a polynomial to points and replaces the center value with the predicted value of the model.¹ The Savitzky-Golay algorithm is the most common digital filtering method. Other smoothing filters include the mean smoother, running mean smoother, and a running median smoother.⁴

Background correction involves the manual subtraction of baselines between points chosen by the operator.¹ Another method of background correction could involve subtracting a fitted model for a trend present in the baseline. Lastly, sample vectors could be replaced by their first derivatives in order to correct for background noise. Baseline correction is a type of background correction, and occurs when a signal may contain a source of variation besides noise that is not significant. These baseline disturbances could

be detrimental to analysis if not removed, such as by causing chemometric classification of samples based on this baseline effect rather than main characteristic features. Baseline correction is a type of artificial removal and/or linearization, and it can typically improve accuracy and appearance.⁴

Normalization of the data can eliminate variability arising from sample amount, concentration, size, and instrument response. It is typically performed after smoothing and background correction. Normalization divides the values of the variables by a constant, and thus places them on the same scale.⁴ The values could be divided by the sum of absolute value of all intensities (normalizing to unit area) or it could be divided by the square root of the sum of squares of the values (normalizing to unit length).^{1, 4, 5}

Mean centering is carried out on one variable at a time. It involves calculating the mean of each variable and subtracting that value from the related elements of each sample vector.¹ It removes constant background without changing any differences in the variables.^{1, 5}

Autoscaling is a form of variable weighting that applies mean centering then followed by variance scaling. Variable weighting multiplies values in the variable by a set number. Variance scaling divides each value in the variable by the standard deviation of that variable.^{4, 5} This technique is recommended when variables have large differences in variance or are measured in different unit systems.¹

1.1.2. Agglomerative Hierarchical Clustering

Agglomerative Hierarchical Clustering (AHC) is a specific cluster analysis technique. Cluster analysis is used to classify individual samples into defined subgroups when no prior knowledge of groupings is known, thus it is referred to as an unsupervised technique. Clustering techniques can be utilized to perform several functions including data reduction, searching for natural groupings in data, generating hypotheses for future samples, evaluating dimensionality, and identifying outliers.^{3, 6} Determining the natural groupings of the variables is the basic objective and this is done on the foundation of similarities or dissimilarities (distances). Different approaches can be taken to measure similarity and dissimilarity. One such approach is using a Euclidean distance, which is a measure of dissimilarity. Euclidean distance, or true ruler distance, is the distance between two objects as if measured with a ruler. It is the simplest measure of proximity between patterns and is based on the Pythagorean theorem.¹ It can be tabulated using Equation 1.1, in which x and y are two points and $d_{(x,y)}$ is the ruler distance.^{3, 7}

Equation 1.1

$$d_{(x,y)} = [(x - y)'(x - y)]^{1/2}$$

Another approach is to use the standardized ruler distance in which all the variables are first standardized and then the Euclidean distance is calculated using their standardized Z-scores.⁷ The Mahalanobis distance is another measurement for similarity and dissimilarity. This method requires estimates of the within cluster variance-covariance matrices and then these matrices can be

pooled across the clusters. It is calculated from the centroid of a group of samples. It can be calculated following Equation 1.2.⁷

Equation 1.2

$$d_{(x,y)} = [(x - y)' \Sigma^{-1} (x - y)]^{1/2}$$

Similarity measures can also be determined by using sample correlation coefficients.³

Various cluster analysis techniques can be used to separate the data into these groups, or clusters. These can loosely consist of hierarchical techniques, optimization-partitioning techniques (clusters are formed by the optimization of a “clustering criterion”), density techniques (clusters are formed by searching for areas containing dense concentrations), and clumping techniques (classes can overlap).⁶ Focus will be placed on hierarchical techniques since AHC was employed in this research.

Hierarchical clustering groups data points into clusters either based on a series of successive fusions or successive partitions, as seen in the two main types of hierarchical clustering methods.³ Agglomerative hierarchical clustering (AHC) starts with every object being individual; therefore, there are as many clusters as there are objects. Objects are then grouped into subsets (“clusters”), such that those within each group are more closely related than other objects in different subsets. The most similar objects are grouped first and then these groups are merged according to their similarities. Eventually, as few clusters as possible will exist.³ Divisive hierarchical clustering initially has a single group

containing all objects and this large group is then divided into subgroups until every object has become its own group. The results of both of these methods, though, are a two-dimensional diagram called a dendrogram which illustrates the steps in the clustering process. The branches represent clusters and branches merge at nodes, in which these node positions indicate where the level of union occurred.^{3, 6}

Within hierarchical clustering, various linkage methods are possible. One such method is the nearest neighbor, or single linkage method. Initially, every group consists of one observation and these groups are then combined based on the distance between the nearest groups, with the smallest distance being joined first. The distance between groups is thus defined as the distance between their closest members (smallest distance, or largest similarity).^{3, 6} Observations will continue to be combined until only one large cluster of all observations remains.⁷

Furthest neighbor, or complete linkage, is another method used in hierarchical clustering, and it is the exact opposite of the nearest neighbor (single linkage) method. The distance between groups is defined as the distance between their furthest neighbors, or the most distant observations.^{6, 7}

The centroid method defines the distance between clusters as the distance between cluster means. Groups are replaced by their centroids, and groups are then joined according to the distance between the centroids, with the groups having the smallest distance being combined first.^{6, 7} The median method involves new groups forming between the two groups.⁶

The group average method the distance between groups is the average of the distances between all pairs of observations in the two groups.⁶ It takes the average of all the possible couplings of points so that one member of each pair is in each cluster.^{3, 7}

Ward's method requires the consideration of every possible pair of clusters being joined together at every step in the analysis. The two clusters whose union would result in the minimum increase in information loss (error sum of squares) are used.⁶ The error sum of squares is determined by measuring the total sum of squared deviations of every point from the mean of the cluster, or by squaring the distance between the cluster means and dividing that by the sum of the reciprocals of the number of points within each cluster.^{6, 7} Ward's method will minimize the heterogeneity within a group.

Overall, when trying to determine the number of clusters present, the hierarchical techniques will provide no indicators as to the number of groups. Therefore, examination of the dendrogram by the observer is needed.⁶ Because of this, AHC is a "pseudo-supervised" technique. The operator has the choice of linkage method and also the number of classes to select. An important point to remember, though, is that cluster analysis will be sensitive to outliers and noise.³ Clustering analysis has been applied to soil,^{8, 9} glass,¹⁰ and inks,¹¹ and specifically AHC has been applied to electrical tapes,^{12, 13} multi-layered paint chips,¹⁴ polymers,^{15, 16} and photocopy and printer toners.¹⁷

1.1.3. Principal Components Analysis

Principal Components Analysis (PCA) is a dimensionality reduction technique. Its main objectives are to reduce and interpret data, identify new variables, screen data, check assumptions, and verify clusters.⁷ It is an orthogonal linear transformation that transforms the data into a new coordinate system illustrating maximum variance.¹ A factor loadings plot illustrates the new coordinate system and data. A positive correlation exists between two factors that lie close to each other in the same plane, a negative correlation exists between two factors that lie 180° apart from each other, and no correlation exists between data points that are 90° away from each other.

The PCs are uncorrelated, and the greatest variance of the data lies on the first coordinate (line through the centroid of the data); the second greatest variance lies on the second coordinate (orthogonal to first PC). Each successive PC will account for less of the remaining variability. The possible number of PCs is the smaller number of samples or variables.^{1,7} Only PCs that capture a majority of the variance are retained and all other PCs that are associated with a small portion of the variance are ignored. Significant PCs will have eigenvalues that represent variation (signal) and not just noise.¹ Eigenvalues are the variances that are accounted for by the PCs and they sum the total variance.¹

A certain number of components are needed to obtain the total variance of a population. However, much of the variability can often be explained by a smaller number of the PCs. A smaller amount of PCs will describe most of the variance. Therefore, the original data set can be reduced to a smaller number of

PCs and still retain the same amount of measurement.³ The new PC axes compress the data by reducing the number of dimensions. All of this is done without much loss of information.

When determining the number of PCs to retain, a few options can be chosen. A percentage of total variability can be selected for how much variability is wanted. The more PCs that are required, the less useful each one becomes.⁷ Another option is to look at the plot of percent variance, often called a scree plot.¹ The scree plot displays eigenvalues against factor number. The location of a sudden break in the plot is an indicator of the number of significant principal components because the lower, leveled off eigenvalues are typically close to zero and thus not significant.⁷ For this research, in order to determine how many PCs were to be retained, the eigenvalues associated with each PC were observed. Enough PCs were retained in order to account for at least 95% of the total variance in the data set. PCA has been applied to glass,^{10, 18-20} paint,^{21, 22} polymers,^{15, 16, 23} inks,^{11, 24, 25} fibers,²⁶⁻²⁸ diesel fuels,²⁹ photocopy and printer toners,^{17, 30} lubricants,³¹ paper,³² nail polishes and paint surfaces,³³ steels,²⁰ accelerants,³⁴ gasoline,^{35, 36} electrical tapes,^{12, 13} and soil.^{8, 9}

1.1.4. Discriminant Analysis

Linear discriminant analysis (DA) reduces the dimensionality of the data in a way that best discriminates between groups. A new set of axes is created that divides the data into groups, and these are discriminant axes, or canonical variates (CVs), that are linear combinations of the original features.¹ The

classifications from PCA are changed by using a learning sample of known observations from PCA data in order to develop canonical variates. DA is a supervised technique, meaning that knowledge of group relationships for each sample is required. The number of samples must be larger than the number of variables, so PCA is commonly used to reduce the number of variables prior to DA. Theoretically, the DA clustering should be in agreement with the AHC patterns.

Data is classified into defined populations and in order to do this, a discriminant rule is created in which random samples are present in every classification group. DA then provides methods that allow the researcher to create rules to classify future samples into one of the groups.⁷ Four different ways can be used to generate a discrimination rule. These ways are the likelihood rule (multivariate normal probability density function), linear discriminant function rule, Mahalanobis distance rule, and the posterior probability rule.⁷

Once the discriminant rule has been created and new samples have been classified, the classification accuracy can be estimated using different methods. The resubstitution method applies the discriminant rule to the data set (making it the training set) and determines how often the rule correctly classifies these observations. Therefore, the classification procedure is based on the known class membership of each of the samples in the data set. A drawback to this method is that it often overestimates correct classifications.^{1, 7}

Another classification accuracy method is estimates from holdout data. The data is divided into two sections: a training set to create the classification model and a test set to predict future classifications. The test set (or “holdout data”) is the one where it is known where the observations should be classified. The discriminant rule is created using the training set and then applied to the test set. Though the estimate of error is unbiased, a drawback is that the paramount discriminant rule may not be created because all the data is not being used to create the rule.^{1, 7}

Lastly, the leave-one-out cross validation method could be applied. With this method, the first observation is removed from the data set and a discriminant rule is created based on the remaining data. The rule is then applied to the first observation and it is noted whether the classification was correct or incorrect. The first observation is then replaced in the data set and the second observation is removed. A discriminant rule is formed from the remaining data and the second observation is classified using the rule. This process continues through all the samples, treating each one as an unknown. This is the preferred method and it provides practically unbiased estimates of the true odds of correct and incorrect classifications.^{1, 7} DA has been applied to nylon,³⁷ fuel oils and asphalts,³⁸ paints,²¹ glass,^{18, 39} diesel fuels,²⁹ gasoline,^{35, 36} photocopy and printer toners,^{17, 30} lubricants,³¹ electrical tapes,^{12, 13} soil,⁸ and inks.^{11, 24}

CHAPTER 2. AUTOMOTIVE CLEAR COATS

Paint evidence is often found at the scenes of automobile crashes where one car hits another car, object, or pedestrian. In some cases, paint can be transferred between cars or from a car onto the clothing or body of the person. In most situations, paint cannot be associated to a specific source. As a result, forensic testing focuses on generating physical and chemical data on the paint in question and comparing it to a known sample of paint from the automobile. One aspect of automotive paint that has not been fully exploited for forensic purposes is the clear coat layer. By studying clear coats, more data can be generated about paint evidence and thus a better association can be drawn between the transferred and native paints.

Application of automotive paint usually consists in stages. The first stage is the application of the primer which is usually electrolyzed onto the body surface. Typically, above this layer is the top coat which consists of a color base coat followed by the clear coat. The clear coat is the top coat of paint and contains no pigmentation or color and protects the base coat from degradation.⁴⁰ Clear coats were first developed in the late 1970s when the topcoat paint system was split into two parts, a pigmented base coat and a colorless clear coat. By the 1980s, the clear coat system became popular, and in the 1990s, new types of paint binders were introduced as well as paint that contained lower amounts of

volatile organic compounds (VOCs) to meet new environmental standards.⁴¹

Automotive clear coats currently used are applied either by a liquid application (i.e. acrylic melamine and acrylic carboxy epoxy) or a powder coating (i.e. acrylic carboxy epoxy and acrylic urethane).⁴²

Almost all clear coats contain light stabilizers such as hindered amine light stabilizers or HALS. Clear coats also contain UV absorbers in order to protect the automobiles against UV light and weathering. UV absorbers must absorb light in the wavelength range of 290 to 350 nm because these photons can cause the photodegeneration of the polymers that make up the binders in the paint. UV absorbers found in automotive clearcoats are often benzotriazoles and triazines, but benzophenones and oxalanilides can also be used.^{41, 43} Clear coat binder chemistries are relatively simple, consisting of acrylics and polyurethanes, and are based on hydroxyl-functional polymers that react with cross linkers.⁴⁴ Both UV absorbers and the UV absorptions of the entire binder systems can contribute to the overall UV absorption of the clear coat layer.

2.1. Review of Analysis of Automotive Clear Coats

An established procedure exists for analyzing paint evidence in normal casework. The Scientific Working Group on Materials Analysis (SWGMA) created a guide for paint analysis and comparison of paints that involves a combination of several methods. An ASTM Standard E1610 (Standard Guide for Forensic Paint Analysis and Comparison) has since been established.⁴⁵ The first step of a typical paint exam involves preparation of cross sections of the paint

chips and then visual and microscopic examinations which assist in observing the different layers and layer sequences of the paint. Microscopic solvent tests can also be performed to determine solubility, texture change, and differences in resins. After the microscopic examinations, the chemistry of each layer is investigated. This is typically accomplished using Fourier transform infrared (FTIR) spectrophotometry and pyrolysis-gas chromatography/mass spectrometry (Py-GC/MS). Elemental composition can assist in identifying inorganic pigments and fillers and it can facilitate the determination of the type of paint in a single layer when used together with FTIR.^{46, 47} Other techniques that could be used to analyze paint are microspectrophotometry (MSP) for color analysis and x-ray diffraction (XRD) for identification of pigments and fillers.

FTIR spectrophotometry is the analytical tool that lies at the core of most paint examinations. FTIR allows information to be obtained about the binders, pigments, and additives in coatings and it can provide molecular structure information about organic and inorganic components.⁴⁸ This technique is useful for determining the general type of paint and identifying resins, pigments, and fillers. Extensive research has been conducted on the forensic applications of IR to paint samples^{14, 49, 50} and research has also been done specifically on the forensic examination of clear coats using IR. Both Edmonstone⁵¹ and Eyring⁵² used attenuated total reflectance (ATR) FTIR of the clear coat as a part of an overall analytical scheme to differentiate large collections of paint samples.

Py-GC/MS is a valuable, albeit destructive, technique in the forensic examination of paint, especially by providing discrimination of the polymeric

binders.⁵³ It is also useful in differentiating paint samples of very similar composition.^{22, 54, 55} More specifically, Py-GC/MS has been frequently touted for its ability to differentiate and identify paint samples that were indistinguishable by IR.⁵⁶⁻⁵⁹ Several authors have also used chemometric methods to analyze Py-GC/MS data.^{21, 22, 54} Finally, Py-GC/MS has been shown to discriminate different clear coats. Burns⁶⁰ and Plage⁶¹ demonstrated that this technique can classify clear coats through visual inspection and/or library searching of the chromatograms and mass spectra.

Microspectrophotometry (MSP) is a well established technique to distinguish dyes and pigments in paint since it allows discrimination of samples based on their interaction with light.⁶² For example, research on achromatic paints has shown that black and grey/silver topcoats produce spectral information in the visible region but undercoat samples do not.⁶³ Given that the UV-absorber concentrations in a clear coat can be measured at extremely low levels, automotive paint should be able to be differentiated based upon the clear coat, even if the same color coat is used. In particular, the findings of Stoecklein and Fujiwara are the most relevant to this research.⁴¹ In their paper, it was demonstrated that clear coats can be distinguished based on their UV absorbance spectra. While features in the UV spectra could arise from UV absorbers and/or binder systems, the spectral features of clear coats were largely attributed to the UV absorber itself. These UV absorbers each contained one of three structural cores: hydroxyphenylbenzotriazole, benzophenone and oxanilide.

2.2. Materials and Methods

2.2.1. Instrumental Analysis

Initially, samples of paint were collected from automobiles from junkyards and automobile body shops. A scalpel was used to scrape the paint chips down to the underlying metal to guarantee that all paint layers were present in the sample. The make, model, and year of each vehicle were noted for each sample. Samples were prepared by using a microtome to make a ten micron cross-section of each paint sample. These cross-sections were then mounted on a quartz slide with a quartz cover slip with Lensbond being chosen as the mounting medium. After analyzing this data, it was determined that the Lensbond was absorbing in the UV-range and thus obscuring the resulting spectra. The absorbance spectra had broad peaks and no clear definition, as seen in Figures 2.1 and 2.2.

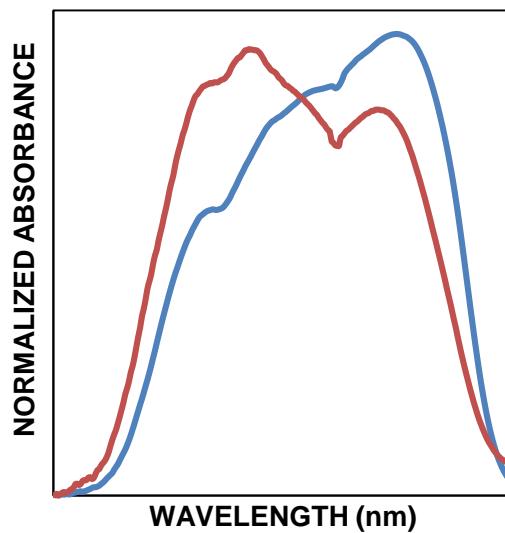


Figure 2.1 Absorbance Spectra of Two Central Objects with Lensbond Used as the Mounting Medium.

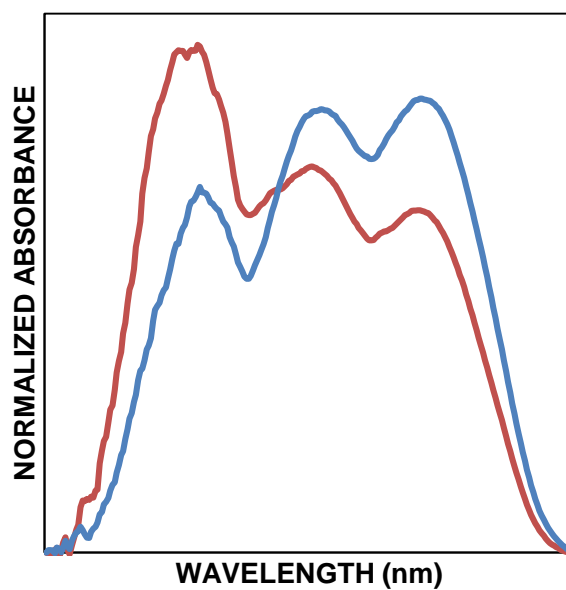


Figure 2.2 Absorbance Spectra of Same Two Central Objects Without Lensbond.

Because of the distortion caused by the mounting medium, this sample preparation method was abandoned and clear coats were instead peeled off and analyzed directly. Using a coverslip was determined to be unnecessary so the

peels were placed directly on a quartz slide without a coverslip. It is important to note that this form of sample preparation is not typically used for comparison of a known (K) and questioned (Q) paint chip, where both would be simultaneously mounted and sectioned. In contrast, this study is an attempt to use chemometrics to more quantitatively assess the diversity of clear coat samples, rather than conduct any so called “Q vs. K” comparisons. Additional samples collected from Australia were also included in this study. These were collected by taking a disc of metal from discarded panels at an automotive repair shop or by taking samples from car roofs removed at a car sun roof fitting business.

A CRAIC QDI 2000 microspectrophotometer (CRAIC Technologies, San Dimas, CA) was utilized with UV analysis being performed in transmitted light mode. Magnification was 35x. Prior to running samples, the microspectrophotometer was calibrated using NIST traceable standards. An autosect optimization, dark scan, and reference scan were run before each set of sample scans. Samples were taken as absorbance values, and five scans were taken at different locations for each paint sample. The five normalized scans taken from Sample 77 (1993 Chevy Lumina) are shown in Figure 2.3 to illustrate the reproducibility of this technique. The scans for each sample were then averaged together (for AHC analysis) or used directly for PCA and DA.

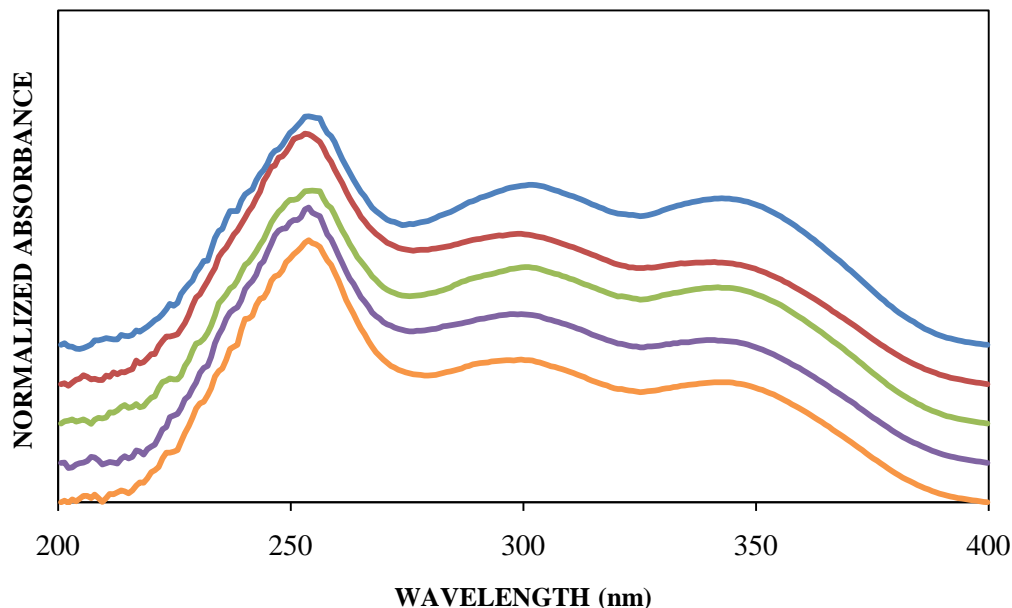


Figure 2.3 Baseline corrected, normalized, and offset absorbance spectra of five scans for a clear coat from a 1993 Chevy Lumina.

2.2.2. Data Analysis

One feature of multi-variate analysis is that relatively large sample populations are needed in order to confidently describe the organization of the data. In this paper, the sample collection consists of 355 spectra (71 total samples). The collection included a wide variety of cars; 18 different years are represented, 22 different makes, and 48 different models. Out of these spectra, nine samples with five scans each (total of 45 spectra) were replicates of other samples, so this data was set aside to be used as supplemental data in an external validation study. The entire dataset (wavelength range of 200 – 900nm) was first truncated to a range of 200 – 400nm (UV range), and this truncated data was baseline corrected by finding the slope and intercept of the raw data using the outermost points, calculating a new y-value by using the equation for a

straight line ($y = mx + b$), and then subtracting the calculated y-value from the raw data. Following this, the data set was normalized by dividing it by the square root of the sum of the squares of all absorbance values.^{4,5} This pre-treatment step eliminated variability in the data due to sample thickness. Statistical evaluation of the data was performed using Microsoft Excel and an add-in, XLSTAT2008 (Addinsoft, Paris, France).

2.3. Results and Discussion

2.3.1. Statistical Results

The AHC dendrogram for automotive clear coats is shown in Figure 2.4. AHC indicates that three distinct clusters are present based on the position of the truncation line, which was determined by a histogram of node positions. The histogram displays the number of nodes as a function of distance. In the dendrogram, small nodes exist (the leaves of the plot, starting with each sample as its own class), and the node positions eventually become more spread out as the number classes becomes smaller. Bifurcations that occur to the right of the truncation line are more significant in determining the number of classes. Averages for each automotive clear coat sample were used when performing AHC. By examination of the graphs of the central objects of each cluster (Figure 2.5), it can be seen that each group has a very distinct absorbance spectrum, and this is consistent with the node positions in the dendrogram showing that the

three classes separate at a relatively large distance as compared to other similar samples.

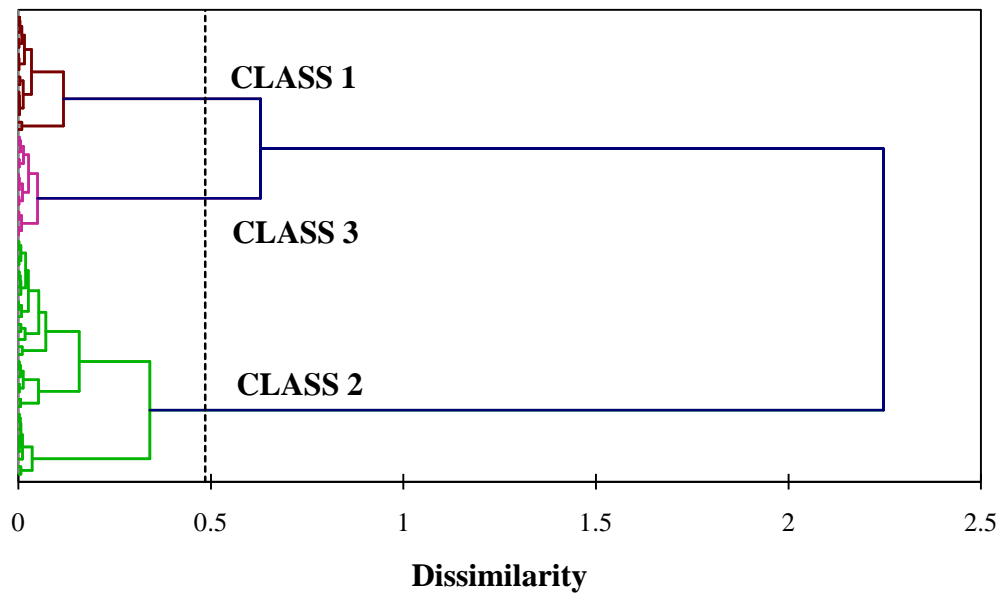


Figure 2.4 Dendrogram from AHC of the averages of each clear coat sample. Three distinct classes are formed.

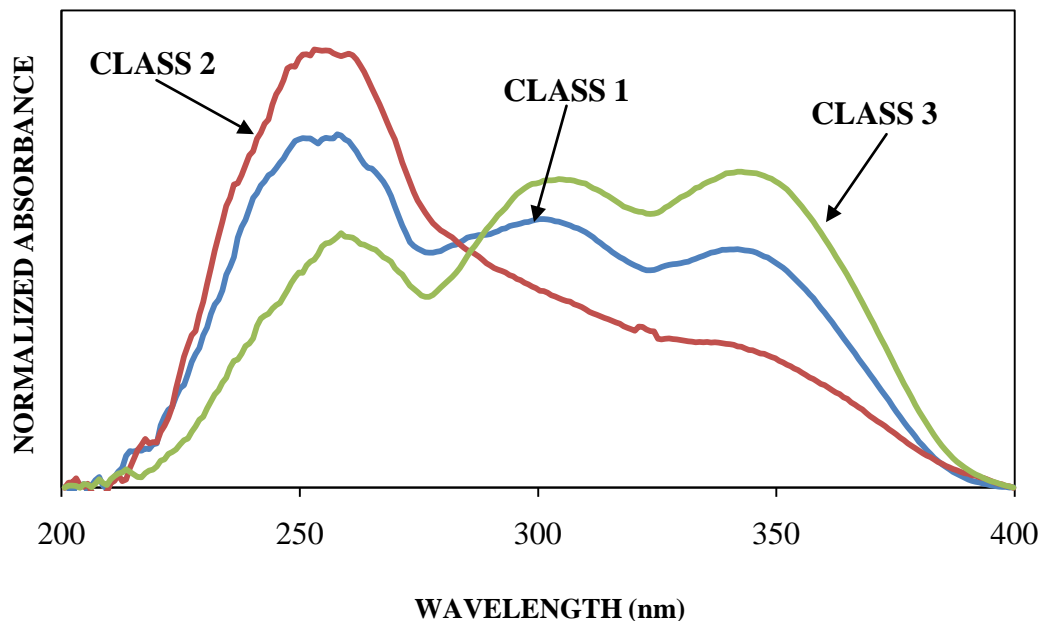


Figure 2.5 Central Objects of the Three Clusters from the Dendrogram.

For the purposes of PCA and DA, every scan was used rather than utilizing only the averages. Again, additional replicate samples were separated from the rest of the data set to be included as supplemental data. These replicates were not included in the PCA analysis but given a factor score for future use. The observations plot produced by PCA is illustrated in Figure 2.6. This plot displays the first two principal components, which captured 78.68% of the total variance of the data set. Though overlap is evident between the three classes, they still are clustering into the three classes. In this case, the total variance in two dimensions is still low; hence separation between samples could still exist in higher dimensions. Overall, three clusters with some overlap are present in this figure with the replicate scans for each sample near one another.

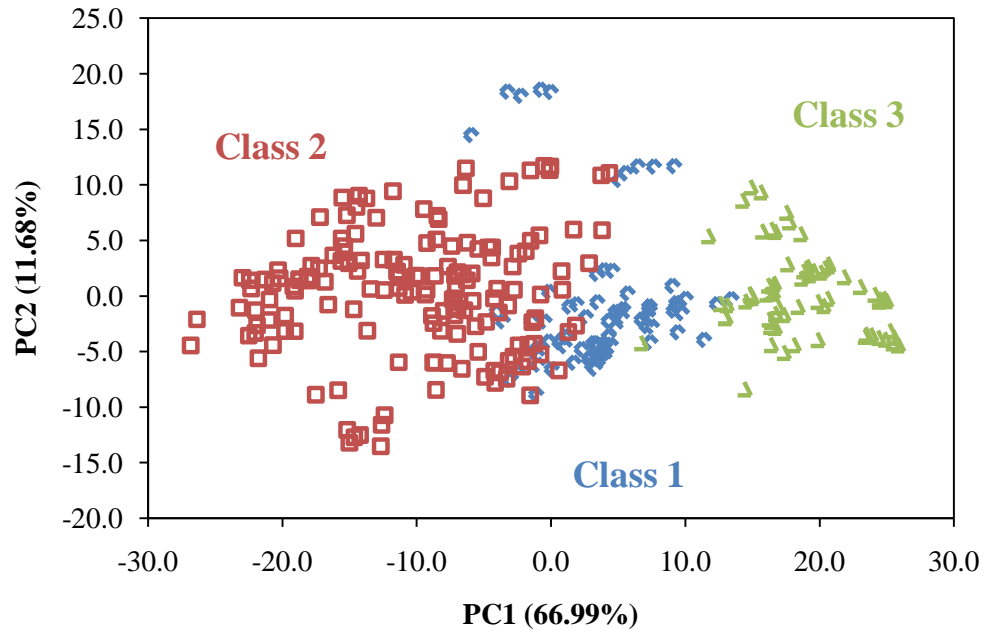


Figure 2.6 The Observations Plot from PCA of Clear Coats.

A factor loadings plot of the original variables can be used to illustrate the PCA coordinate system. In a traditional form of this plot, the variables are projected into the same PC space as the observations. A positive correlation exists between two variables that lie close to each other in the same plane, a negative correlation exists between two variables that lie 180° apart from each other, and no correlation exists between variables that are 90° away from each other. A different type of factor loadings plot was created for this data set (Figure 2.7) where the loadings in the form of the cosine of the angle between the principal component and each variable. Areas where the cosine is positive are areas of positive correlation, areas where the cosine is negative are areas of negative correlation, and areas where the cosine is close to zero have no

correlation. PC1 corresponds to the x-axis and PC2 corresponds to the y-axis of the observations plot (Figure 2.6).

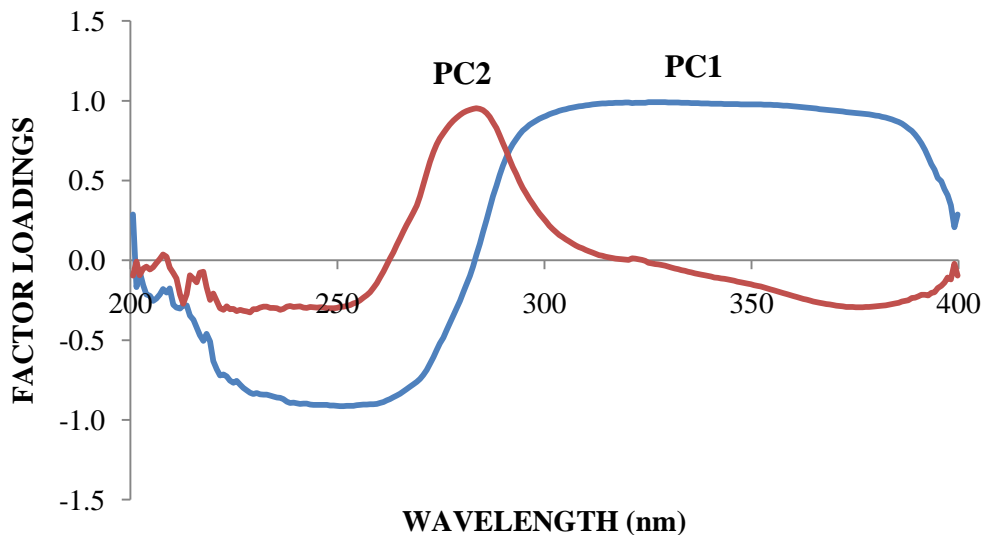


Figure 2.7 Factor Loadings Plot of the First Two PCs.

Focusing on PC1, the main areas of positive and negative correlation can be found, and the wavelengths corresponding to these areas can be highlighted on the central objects plot of all the clusters together. The purple bracketed areas in Figure 2.8 indicate these wavelength regions, and an inverse order is seen between the positive correlation wavelengths (right bracket) and negative correlation wavelengths (left bracket). The positive correlation wavelengths for PC1 range from 300 – 370nm, and the negative correlation wavelengths for PC1 range from 230 – 265nm. This could be due to the light absorbers found in clear coats, which typically absorb between 290 – 350nm.⁴¹ In an attempt to put more physical meaning behind the spectra, the ordering of the clusters within these brackets correlates with the ordering of the clusters in the PCA observations plot.

Samples with high absorbance in regions where there are high factor loadings on PC1 should be on the right of the PCA observations plot because those are the large and most positive values. Therefore, the observations plot should read (from right to left): Class 3, Class 1, and then Class 2. Referring back to Figure 2.6, this is the exact trend that is noticed in the observations plot.

The same physical meaning can be applied to the second principal component. One area of positive correlation is highlighted on the central objects plot (Figure 2.8) by an orange bracket. This area of positive correlation wavelengths ranges from 275 – 285nm. With PC2, samples with strong absorbance in the region of high factor loadings should be the highest on the PCA observations plot since PC2 is associated with the y-axis. Relating this to the PCA observations plot, it would be expected that all three classes will be located similarly in the y-axis direction, but Class 1 would have the highest points in the plot. Referring back again to Figure 2.6, this is the exact trend that is noticed.

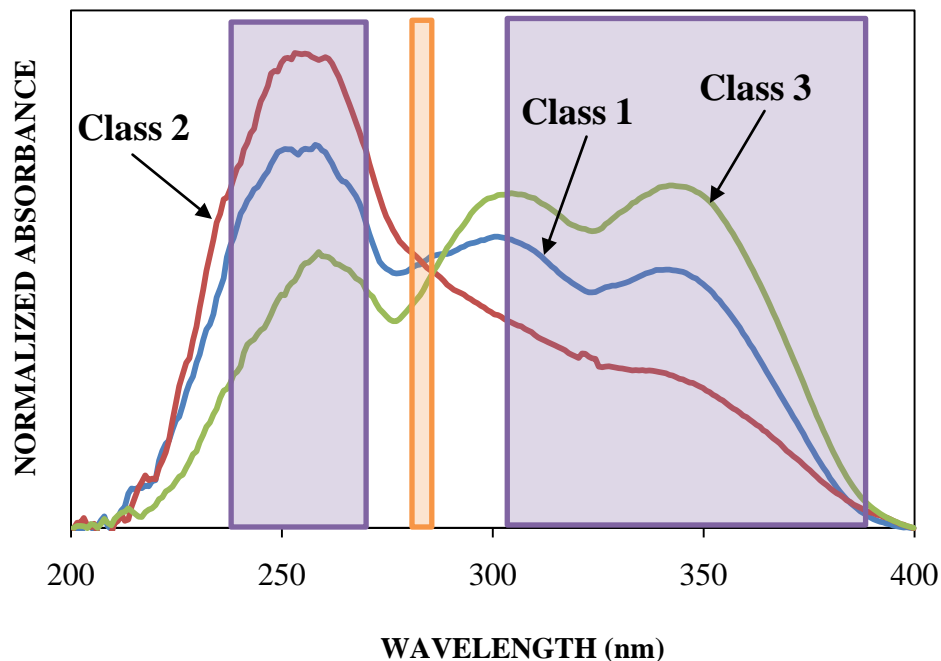


Figure 2.8 Significant Factor Loadings Overlaid on the Central Objects Plot for Clear Coats.

DA was then performed using the data from PCA. Every PC has an associated eigenvalue with it that describes the variance, as well as the percent variance and a cumulative variance for all the principal components. The specific values pertaining to this research are found in Table 2.1. Several methods exist that assist in determining the correct number of principal components to use. One method involves setting a targeted cumulative percentage of variance and using the number of principal components associated with that number. This method was used in this research, and 95% was the selected cumulative variance. Using this variance, the first nine principal components would be selected, as shown in Table 2.1. Another method involves using a scree plot, which displays eigenvalues against factor number. The location of a sudden

break in the plot is an indicator of the number of significant principal components. Based on the scree plot for this data set, four principal components would be used. Lastly, the Kaiser criterion can be used to determine the significant amount of principal components to utilize. Using this criterion, all eigenvalues greater than one would be considered significant, and based on Table 2.1, this would apply to the first twelve principal components.

Table 2.1 Eigenvalues and variability associated with each principal component (PC).

Principal Component (PC)	Eigenvalue	Variability (%)	Cumulative %
PC1	167.49	66.992	66.992
PC2	29.211	11.685	78.676
PC3	13.837	5.535	84.211
PC4	12.003	4.801	89.012
PC5	7.205	2.882	91.894
PC6	2.917	1.167	93.061
PC7	2.450	0.980	94.041
PC8	2.134	0.854	94.895
PC9	1.684	0.674	95.568
PC10	1.503	0.601	96.170
PC11	1.369	0.548	96.718
PC12	1.185	0.474	97.192
PC13	0.953	0.381	97.573

The results of DA using the first nine principal components are shown in Figure 2.9, with 100% of the variance accounted for in two dimensions. Three distinct groupings were used to be consistent with AHC. Overlap can be seen between the three groups, which affects the cross-validation results as seen in the confusion matrix results in Table 2.2. Samples located along the diagonal represent those that were correctly classified, while samples in bold outside of

this diagonal are incorrect classifications. Overall, 91.61% of the samples were correctly classified, which is considered to be an excellent result.

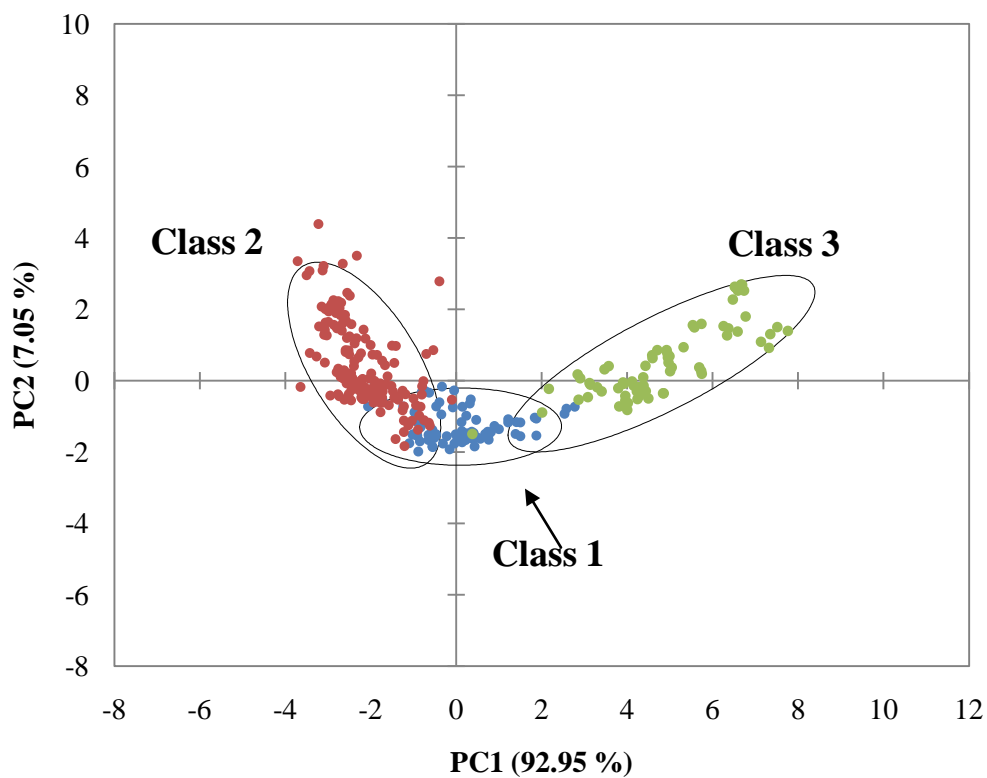


Figure 2.9 Observations plot from DA.

Table 2.2 Confusion Matrix for the Cross-Validation Results from DA.

<i>From/To</i>	<i>Class 1</i>	<i>Class 2</i>	<i>Class 3</i>	<i>Total</i>	<i>% Correct</i>
<i>Class 1</i>	72	6	2	80	90.00
<i>Class 2</i>	15	145	0	160	90.63
<i>Class 3</i>	3	0	67	70	95.71
<i>Total</i>	90	151	69	310	91.61

2.3.2. External Validation

The additional replicates were used as a form of external validation and the results are shown in Table 2.3. DA predicted which class the supplemental samples should be placed in and the correct placement was determined by where the original (non-replicate) sample was placed. The green, bolded diagonal numbers are samples that were placed correctly in the proper class. The red, bolded numbers outside of the diagonal indicate samples that were incorrectly classified. Overall, the performance of the classification model was good, with 81.11% of samples correctly assigned. In this external validation, the samples that were misclassified had absorbance spectra that more similarly resembled the absorbance spectra of the central object of a different class.

Table 2.3 Confusion Matrix for the External Validation Results of the Supplemental Data from DA.

<i>From/To</i>	<i>Class 1</i>	<i>Class 2</i>	<i>Class 3</i>	<i>Total</i>	<i>% Correct</i>
<i>Class 1</i>	6	2	2	10	60.00
<i>Class 2</i>	5	25	0	30	83.33
<i>Class 3</i>	0	0	5	5	100.00
<i>Total</i>	11	27	7	45	81.11

Given that DA is a supervised technique, the number of classes and their composition can have a large effect on the results. Furthermore, the use of internal versus external validation can also result in overly optimistic estimates of the accuracy of the DA model. Therefore, an additional study where the number of classes was varied demonstrated the validity of external samples. As is evident in Figure 2.10, the external validation had very high accuracy for three classes (81.11%), whereas the accuracy was significantly lower and never reached over 80% for all other number of classes. The estimation and cross-validation accuracy did not vary significantly when class number was changed and offered a clearly inflated view of the classification accuracy. Based upon external validation, three classes can be claimed to reasonably exist whereas larger number of classes have too low of an accuracy to say for certain that they exist.

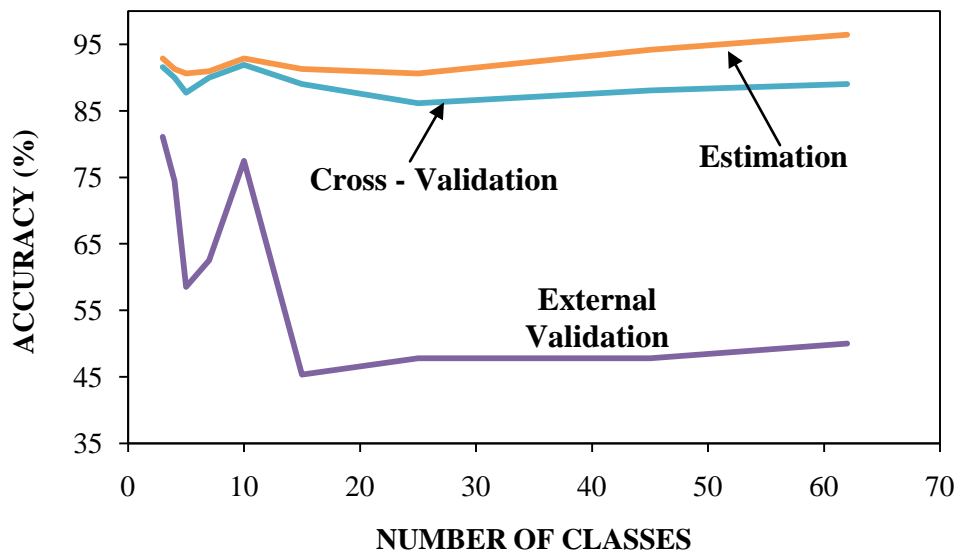


Figure 2.10 Percent accuracy for each DA testing technique versus varying number of classes.

2.3.3. Formation of Classes

Overall, chemometric analysis indicated three distinct groups that produced reliably discernible spectra. Using these results, the question of whether make, model, and year could be affecting the formation of these groups was addressed. Absorbance spectra of cars of the same make and model but different years were inspected to see if year could be affecting the cluster formation.

Based on Figures 2.11 and 2.12, cars of the same make and model but different year exhibited varying trends. Figure 2.11 shows cars of the same make and model but different year having visually dissimilar spectra and being placed in different classes (Year 2000 into Class 2 and Year 1998 into Class 3). However, Figure 2.12 shows cars of the same make and model but different year being placed in the same class (Class 2). Though the spectra in Figure 2.12 are

more similar to each other than spectra in Figure 2.11, it is still apparent that the spectra in Figure 2.12 are not the same even though they are being clustered into the same class. Therefore, year did not affect the groupings, and visual inspection can also assist in distinguishing clear coat spectra. Next, cars of the same make, model, and year were examined to see if make and model were an issue. Both 2008 Holden Captiva samples in Figure 2.13 had similar spectra and fell into Class 1. However, no others cars of the same make, model, and year were present in this dataset to show if any trend truly existed.

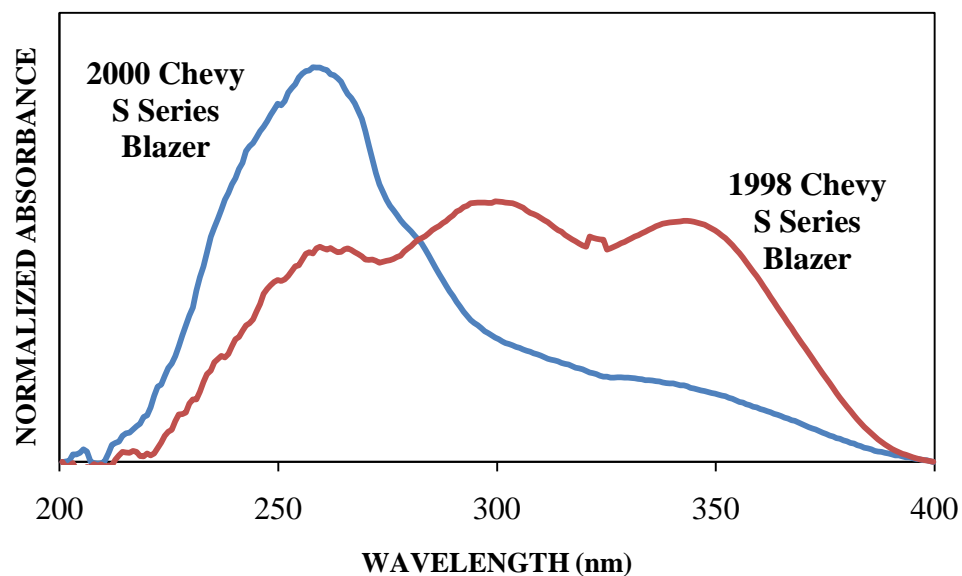


Figure 2.11 Samples of the Same Make and Model but Different Year Placed in Different Classes.

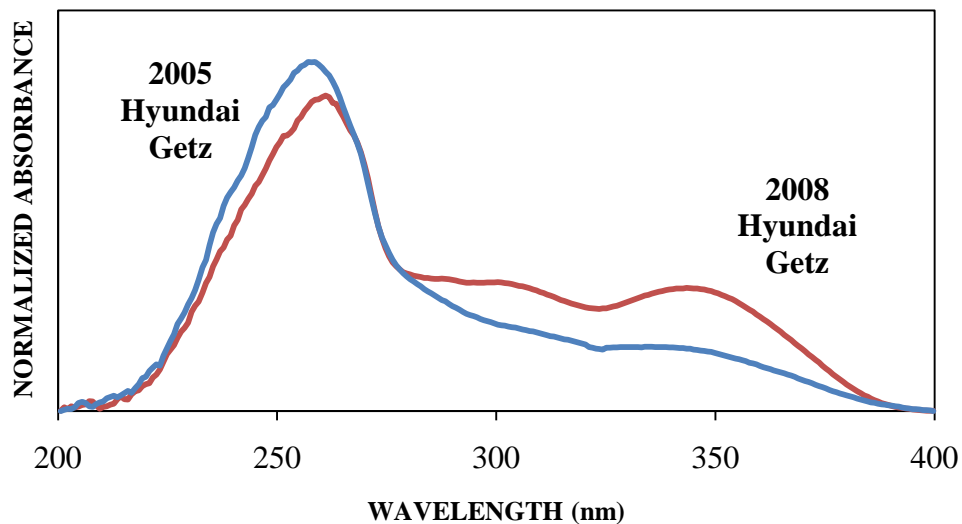


Figure 2.12 Samples of the Same Make and Model but Different Year Placed in the Same Class.

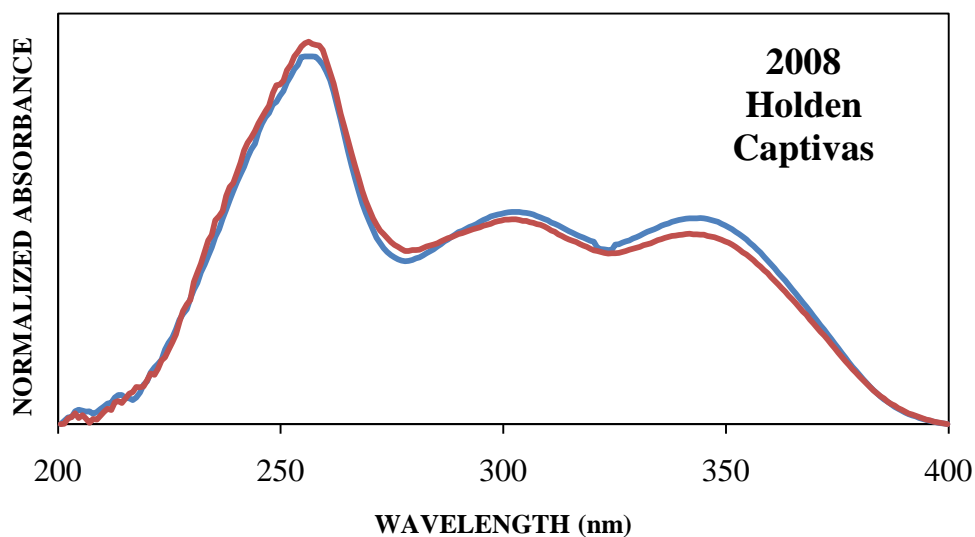


Figure 2.13 Samples of the Same Make, Model, and Year.

Another way to see that year is not affecting the class formation is by looking through the data set. Class 1 has automobiles ranging from 1993 to 2009, Class 2 has automobiles from 1990 to 2008, and Class 3 ranges from 1995 to 2009. Furthermore, the make and model of automobiles varies between

the classes. For example, automobiles of the same make and model are found in different classes such as a Chevrolet Lumina in Class 1 and Class 2. Certain makes of automobiles are found in all the classes as well, such as Toyota. Overall, these results indicate that there is not a correlation between the make, model, and year of the automobiles.

2.3.4. Limitations to the Study

It is important to note some significant limitations in this type of study. First, the initial conditions of the clear coats are unknown and environmental factors are likely to affect the clear coat. In this study, however, environmental exposure was not controlled as samples were gathered “as is” from various vehicles. For example, suppose clear coat samples were collected from two cars of the same make and model that were manufactured in the same year. If one was stored in a garage and the other was kept in the sun, the clear coat UV absorbance spectra would most likely be affected differently. In most cases, it is uncertain as to the history of the known and unknown during the interval prior to their becoming evidence.

Another possible limitation is that there can be differences in the finishes that are used on different parts of the automobile body, such as plastic and metal parts. Since a majority, if not all, of our samples were removed from the same part of a car, this limitation is avoided but it could affect casework. Lastly, it is possible that more than one manufacturer could supply the finish for a given vehicle. Therefore, a given make, model, and year could easily have two or

more types of finishes. Overall, these limitations could explain why a correlation is not seen between year, make, or model in the collection of clear coats analyzed here.

In addition, MSP does not provide a way for determining which UV absorbers were present in each clear coat. Knowing the UV absorbers could assist in determining if a correlation exists between the type of UV absorber used in a clear coat and the class into which a given sample was placed. Method development involving different instrumental techniques such as pyrolysis or extracting the components of the clear coat using LC/MS would elicit enough information to assign chemical meaning.

2.4. Conclusions

Based on the results discussed above, MSP is an appropriate and efficient step in the analysis of automotive clear coats, even though it is not as frequently used as other techniques. Proper sample preparation has been shown to be vital in obtaining accurate spectral data related to these automotive clear coats.

Overall, the comparison of automotive clear coats using AHC, PCA, and DA resulted in several findings. Three distinct groups formed within the collection of automotive clear coats. The central objects of the AHC dendrogram (Figure 2.5) illustrate that the three classes differ in their relative absorptions at approximately 250 nm, 300 nm and 350 nm. The spectral regions that correlated most strongly with PC1, and hence represent areas of high variability between samples, were 230 – 265 nm and 300 – 370 nm. Overall, the three classes were

differentiated with high accuracy using both cross-validation and external validation samples. Clear coats also varied widely by model and year, but no clear relationships were seen between these qualitative variables and MSP data.

CHAPTER 3. FIBER DYE ANALYSIS

Textile fibers are arguably one of the most important forms of trace evidence given that they have many classifications and subtypes, are physically and chemically differentiable, have various processing procedures, and are transferred easily. Various characteristics are evident in fibers, and certain ones play significant roles in fiber analysis. One of the most important characteristics for fiber comparisons is color, which reflects the dyes and pigments that were used on the fabric. In fact, the only characteristic of many fibers, such as cotton, that can be reliably used for discrimination of samples is its color.

Textile fibers are the basic unit of matter that form the components of fabrics and textiles. Although a multitude of classifications and subtypes exist, fibers can be broadly classified as either natural or man-made.⁶⁴ Natural fibers are further sub-divided according to their source (animal, vegetable or mineral). Man-made fibers are sub-divided according to their base material (synthetic polymer, natural polymer or other). There are over 1,500 manufactured fiber plants worldwide, according to 2007 data.⁶⁵ A majority of these plants are synthetic fiber producers, followed by cellulosic and glass fiber manufacturers. Regionally, Asia has the most manufactured fiber plants. Within the United States, 145 total fiber manufacturing plants exist, the vast majority of which are synthetic manufacturing plants.

Fibers can be transferred from one object to another either directly (a primary transfer) or through one or more intermediates (a secondary, tertiary, etc. transfer).^{66, 67} Forensic science laboratories can then examine these transferred fibers and compare them to a known fiber in order to discover possible common origins. However, a significant factor in fiber examination is the passage of time. For example, studies of fiber persistence have demonstrated that up to 80% of transferred fibers are lost after the first four hours.⁶⁶ Therefore, it is important for investigators to collect garments and any possible fibers quickly. Factors that can influence how fibers adhere to materials include the type of fiber transferred, the type of receiving material, and the extent to which the receiving material is used after transfer.⁶⁶

A dye is a colored substance that is able to absorb and reflect certain visible wavelengths of light. Dyes must have affinity for the substrate on which they are being applied. In contrast, a pigment has no affinity for the substrate and is instead incorporated into the fiber during production or bonded to the surface. An important feature of fiber dyes is that very few textiles are colored with only one dye, and therefore many dye combinations exist. Furthermore, thousands of dyes are produced worldwide. Important reference works such as the *Colour Index* published by the American Association of Textile Chemists and Colorists allow forensic chemists to keep abreast of the chemical dyes that are available. A three-color dye system is what is most frequently used to give a textile its color, and how color is applied and absorbed on the fiber is an important characteristic.

Fiber dyes are classified in various ways, including their method of application, chemical class, or the type of fiber to which they are applied. The main classification scheme typically used by forensic chemists is based on the method of dye application. The major dye classes in this scheme are: acid dyes, basic dyes, azoic dyes, direct dyes, disperse dyes, metallized dyes, reactive dyes, sulfur dyes, and vat dyes. While readers should consult Wiggins⁶⁸ for a more detailed discussion of fiber dyes, the characteristics of the major dye classes are summarized in Table 1 and can be described as follows:

- Acid dyes are applied under acidic conditions so that basic functional groups (e.g., amino) on the substrate are protonated and positively charged. These groups form ionic bonds with deprotonated functional groups (e.g., sulfonate) on the dye molecule.
- Basic dyes are also applied under acidic conditions, but in this case protonated/positively charged functional groups (e.g., ammonium) on the dye form ionic bonds with negatively charged functional groups on the substrate.
- Azoic dyes form a colored product via coupling between a diazo salt and a coupling component such as a naphthol.
- Direct dyes are directly incorporated into cellulosic fibers in the presence of heat and an electrolyte.
- Disperse dyes are also directly incorporated into synthetic fibers, associating with the substrate through weak Van der Waals forces and some hydrogen bonding.

- Metallized dyes form colored metal complexes within the fiber through the reaction of a mordant (binding agent) such as chrome with a separate dye molecule.
- Reactive dyes form covalent bonds with functional groups on the fiber. These dyes are growing in popularity in part because they are less likely to be removed by washing.
- Sulfur dyes require a reducing agent to make them soluble, where it is applied to the fiber, then oxidized within the fiber back to its original insoluble form.
- Vat dyes also require a reducing agent to make them soluble and undergo oxidation within the fiber to create an insoluble dye.

Table 3.1 The Various Types of Dye Classes.

Dye Class	Description	Fibers Typically Applied to
Acid	water-soluble anionic compounds; ionic bond between dye molecule and polymer	Wool, silk, polyamide, polypropylene
Basic	water-soluble, applied in weakly acidic dyebaths; very bright dyes; negatively charged fiber draws the dye cation	Polyacrylonitrile, acrylic, occasionally polyester and polypropylene
Direct	water-soluble, anionic compounds; applied directly to fiber from aqueous medium that has an electrolyte; positively charged ion is attracted to negatively charged fiber surface and the dye is able to enter the fiber	Cotton, rayon, other cellulosics
Disperse	not water-soluble; aqueous dispersion; hydrogen bonds and weak van der Waals' holds dye molecule in fiber	Polyester, acetate
Reactive	water-soluble; form covalent bond with functional groups of fiber; similar in structure to acid dyes and application similar to direct dyes	Cotton, wool, other cellulosics
Sulphur	organic compounds containing sulfur or sodium sulfide; reduced using sodium sulphide or sodium hydrosulphite; dye enters fiber and oxidized to original form	cellulosics
Vat	oldest dyes; more chemically complex; water-insoluble; good color fastness	cellulosics

3.1. Review of the Analysis of Dyed Textile Fibers

The primary methods for the identification and comparison of textile fibers according to their type (e.g., synthetic polymer) and sub-type (e.g., nylon) rely upon the principles of microscopy, spectroscopy, chromatography and mass spectrometry.

Microscopy is the most important tool for fiber examination.⁶⁹ For example, a stereomicroscope can record fiber characteristics such as size, crimp, color, and luster. Polarized light microscopy (PLM) is especially helpful with manufactured and synthetic dyed fibers because it reveals the polymer class based on the rotation of incident polarized light by the fiber. The use of infrared microspectroscopy is widely used in forensic laboratories to identify and compare single fibers.⁷⁰ Specific spectral features are used to discriminate between classes and sub-classes of fiber. Given the polymeric nature of synthetic and natural fibers, pyrolysis coupled to gas chromatography (as well as mass spectrometry) is an informative and minimally destructive technique given the small sample sizes required.⁷¹ Finally, scanning electron microscopy coupled with energy dispersive x-ray analysis (SEM-EDX) can be used primarily to examine the elemental content of fibers, with inorganic materials arising from the residues of the manufacturing process, additives, or environmental contaminants.⁷²

Approaches for comparing fiber color and identifying fiber dyes have been developed alongside those for identifying fiber materials. Figure 3.1 illustrates analytical techniques that have been applied to the analysis of dyed fibers. Older

methods of comparing fiber color involved comparing unknown shades to known hues and finding the best color match. Now, however, more reliable techniques are used to compare colors and dyes.

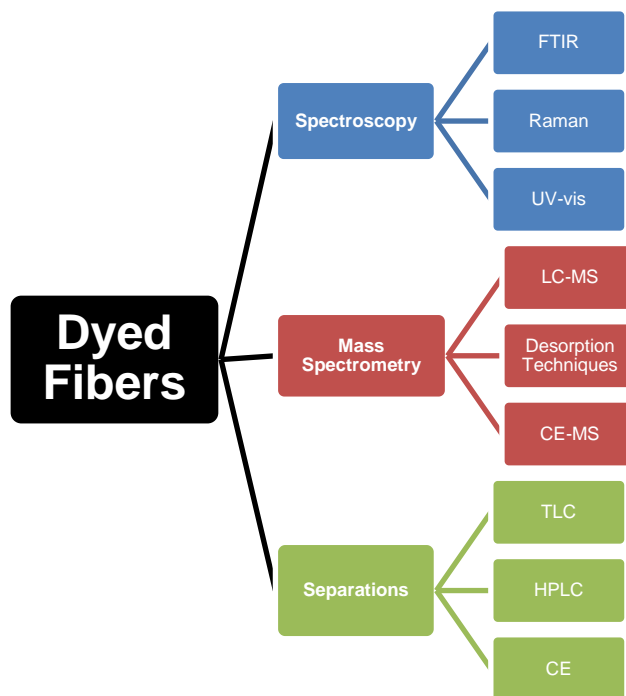


Figure 3.1 Analytical techniques applied to the analysis of dyed fibers.

Microscopic exams remain a key tool for color comparisons.⁶⁹ For example, simple observation of fiber color under the microscope can immediately eliminate two fibers as having come from the same source based on clear visible differences in hue. However, if fibers appear to be the same color that does not mean they have the same composition. Fibers that show the same hue but are dyed with different dyes or mixtures are called metamerism. A way to detect metamerism is by using different light sources to illuminate the fibers. Beyond microscopy, however, are several techniques that have found broad

usage and remain active areas of research. In some cases, extracting the dye from the fiber is crucial to the success of these techniques. The best known methods for dye identification are based on determining the application class, which is the type of dye that was used (acid, direct, etc.) and the generic structure. The current status of these methods will be discussed below in the approximate order in which they would be applied to an unknown sample.

UV-visible Microspectrophotometry uses a microscope to gather light from the sample and transmit it to a UV-visible spectrophotometer. This technique is ideal for objective observations of the color of many forms of trace evidence,⁷³ and the application of MSP to fiber analysis has been discussed in detail.⁷⁴ The use of an instrumental technique such as MSP is significant because it is repeatable, non-destructive, and requires little sample preparation, unlike other methods that require extraction of the dye. MSP can distinguish between colored fibers that may appear visually similar because absorbance in the visible region is dependent on the molecular structure of the chromophore as well as the environment in which the chromophore is found. It is important to note that MSP is not capable of identifying particular dyes or mixtures of dyes, but rather identifying the spectral characteristics of a sample for the purposes of comparison.

Typical MSP protocols call for several fibers to be examined per item (e.g., five for man-made fibers and ten for natural fibers). Upon examination using MSP, man-made fibers are usually found to have a homogenous chemical composition as the dye is bonded to a relatively constant chemical environment.

Cotton and other natural fibers are composed of many different chemical components which are not homogeneously distributed throughout the fiber.⁷⁴ Hence, variation within a control fiber sample may occur due to uneven dye uptake. MSP is currently the pivotal method in color comparisons of individual fibers. Modern UV-visible-NIR MSP instruments can measure transmittance, absorbance, reflectance, and fluorescence, as well as the polarization and luminescence of microscopic samples.

While the use of infrared microscopy is quite common for determining the composition of fibers,^{70, 75} its use for identifying fiber dyes is not. This stems from the inherently low level of dye that is found in most textiles and the lack of sensitivity of infrared absorbance to components that represent less than ~ 5% of a sample. Raman spectroscopy has steadily become established as a technique for fiber analysis^{70, 76} and it has been identified as a priority research area by the European Fibres Group.⁷⁷ While techniques such as MSP provide nondestructive ways to measure color, little information is gained about the specific dye. Raman is well suited for the analysis of colored fibers because this technique requires no sample preparation, is nondestructive, and requires only a small sample. Therefore, multiple tests can be performed on the same fiber.

Thin Layer Chromatography (TLC) is a long-standing tool for the separation of common classes of fiber dyes, with silica gel being the most typically used stationary phase.⁶⁸ Selecting an appropriate extraction and developing solvent can be difficult, however, and numerous solvent mixtures have been tabulated for these purposes.⁶⁸ Specific schemes for dye extraction

and TLC have been published for dyes on cotton,^{78, 79} wool,⁸⁰⁻⁸³ and various synthetics.^{78, 84, 85} TLC has the advantage of being able to identify small amounts of shading colors or intermediates that can often be found in extracted dye mixtures. Most textile fibers contain multiple dyes to obtain the desired shade, so TLC of a few fibers may produce very complex dye patterns. However, no single solvent system exists that can separate all classes of dyes, so some amount of pre-screening is necessary.

As discussed by Griffin and Spears, HPLC has many advantages over TLC and MSP for the analysis of fiber dyes.⁸⁶ For example, MSP is limited when attempting to analyze highly-absorbing (dark color) fibers. Furthermore, TLC requires relatively large quantities of dye and different eluent systems for various dye classes. HPLC, however, exhibits better chromatographic resolution, greater sensitivity and it can be used for quantitation. Capillary Electrophoresis (CE) and its related techniques offer even greater separation efficiency than traditional chromatographic techniques and hold great promise for analytes ranging from small ionizable compounds to large non-polar species. Coupling the separation efficiency of HPLC and CE with the sensitivity and specificity of mass spectrometry (MS) is a natural combination of instrumental methods.

3.2. Materials and Methods

3.2.1. Instrumental Analysis

In this research, we focused solely on red dyed cotton fibers, and therefore sought to obtain as many possible red dyes on cotton. Samples were provided from Testfabrics, Inc (West Pittston, PA), and consisted of six dyed cotton fabrics, all various shades of red, and three undyed substrates consistent with the substrates of the dyed exemplars. The specific samples are listed in Table 3.2. STC EMPA is a catalog designation. STC stands for “soil test cloth,” and EMPA is a Swiss Federal Testing facility in St. Gall, Switzerland that produces some of the dyed fabrics that were used in this research. Six additional samples were provided by Dr. Stephen Morgan at the University of South Carolina, and these specific fibers are listed in Table 3.3. The fiber ID numbers for Dr. Morgan’s samples were kept the same as the numbering he had previously assigned them, and he obtained his samples from different textile companies.

Table 3.2 List of the Dyed Exemplars from Testfabrics, Inc.

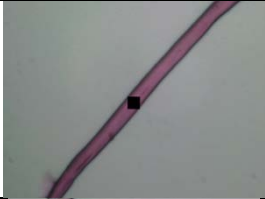




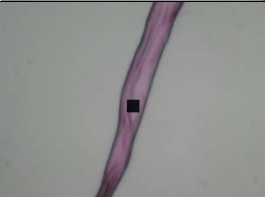
Fiber ID	Dyed Exemplars	Image of Fiber from MSP
A	Style 419, Bleached Mercerized Cotton Blend cloth dyed with Direct Red C-380	
B	Style 460, Bleached Cotton Interlock dyed with Fiber Reactive Red 120	
C	Style 460, Bleached Cotton Interlock dyed with Fiber Reactive Red 123	
D	STC EMPA 475 dyed with Reactive Red 195	
E	STC EMPA 495 dyed with Reactive Red 2	
F	STC EMPA 498 dyed with Reactive Red 228	

Table 3.3 Dyed Fiber Exemplars from Dr. Stephen Morgan of the University of South Carolina.

Fiber ID	CI Name
713	Direct Red 84
685	Reactive Red 180
686	Reactive Red 198
695	Reactive Red 239/241
721	Vat Red 10
722	Vat Red 15

Following standard MSP protocols, ten fibers were removed from each exemplar. These fibers were mounted on a glass slide with a glass coverslip with glycerin used as the mounting medium. This preparation technique was modeled after the technique the Indiana State Police laboratory uses for fiber analysis. In addition, glycerin is an appropriate mounting medium, as stated by the Scientific Working Group on Materials Analysis (SWGMA).⁸⁷ At IUPUI, a CRAIC QDI 2000 microspectrophotometer (CRAIC Technologies, San Dimas, CA) was utilized with Visible analysis being performed in transmitted light mode. At the Indiana State Police (ISP) laboratory, a CRAIC QDI 1000 microspectrophotometer was used. Magnification was 35x for both instruments. Prior to running samples, the microspectrophotometers were calibrated using NIST traceable standards. An autose optimization, dark scan, and reference scan were run before each set of sample scans. Samples were taken as absorbance values, and ten scans were taken at different locations for each fiber sample. The scans for each sample were then averaged together (to be used

only with Agglomerative Hierarchical Clustering) and the absorption data was used in data analysis.

3.2.2. Data Analysis

Three different data sets were utilized in this research. First, the Testfabrics, Inc. fibers were analyzed by this researcher and another individual ran two fibers from these exemplars as well. The additional two fibers per sample were used as an external validation. When analyzing these fibers, an attempt was made to analyze the varying shades of the dye along the fiber (dark, medium, light).⁸⁸⁻⁹¹ Next, new fibers were selected from all twelve exemplars and these were analyzed at the Indiana State Police (ISP) Laboratory. Following this, new fibers were selected from the twelve exemplars and these were then analyzed at IUPUI in order to do an inter-laboratory comparison. With the IUPUI-run fibers, extra care was now given to analyze a similar dye shade on the fibers. In terms of the dye analyzed, the most prevalent shade was selected, which typically was a medium-dyed shade. A similar attempt was made with the ISP fibers. An external validation was also performed on the new twelve fibers at IUPUI.

All data sets consisted of the truncated wavelength range of 350 – 800 nm to ensure that all the data consisted of the same wavelength ranges. They were then background subtracted by subtracting the minimum of each sample. Following this, the data sets were normalized by dividing each absorbance value by the square root of the sum of the squares of all absorbance values.^{28, 29} This

eliminated variability in the data due to sample thickness and dye uptake. The ten normalized scans taken from Sample A1 (Direct Red C-380) at IUPUI are shown in Figure 3.2 to illustrate the reproducibility of this technique. Statistical evaluation of the data was performed using Microsoft Excel XLSTAT2009.

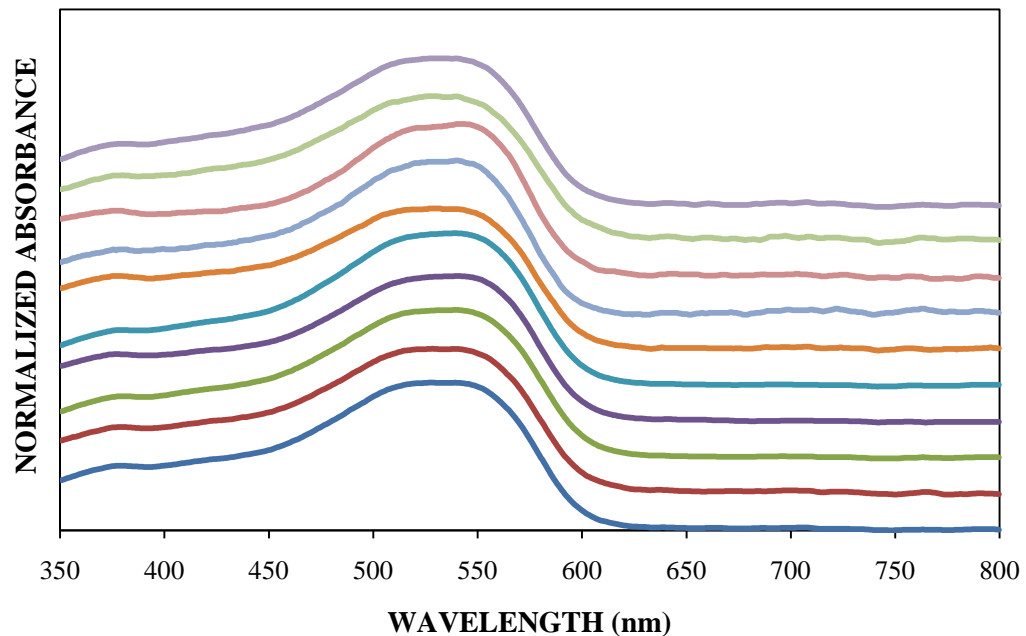


Figure 3.2 Background subtracted, normalized, and offset absorbance spectra of ten scans for a fiber from an exemplar dyed with Direct Red C-380.

3.3. Results and Discussion – Part I: Testfabrics Fibers Analyzed at IUPUI

3.3.1. Statistical Results

The AHC dendrogram for fibers A – F analyzed at IUPUI is shown in Figure 3.3, with three distinct classes being formed based on the truncation line. Bifurcations to the right of the truncation line are more significant in determining

the number of classes. In addition, the averages for each of the fiber dyes were used for AHC. Three classes were chosen to remain consistent with all other data sets because in the ISP data set, when more than three classes was tried, the replicates began to be split amongst the classes. In this IUPUI data set, though, three classes results in the separation of replicates amongst different classes. However, when a larger number of classes is chosen, some of the exemplars began to be increasingly split amongst the various classes. Figure 3.4 illustrates the central objects of each of the three classes, with each having distinct absorbance spectra. These distinctions are consistent with the node position from the dendrogram, showing that Class 3 is the most distinct compared to the other two classes.

When looking more closely at the AHC dendrogram, it is apparent that Fibers A, C, D, and F cluster well. This is also obvious when viewing these fibers' levels of dissimilarity. All four have very low levels of dissimilarity (A = 0.02; C = 0.01; D = 0.01; F = 0.03), thus proving they are very similar replicates. Fiber C is also so well-clustered that it is placed as its own class (Class 3). Fibers B and E, however, do not cluster well and have replicates that are placed in different classes.

In addition, outliers can be noted when viewing the dendrogram. The replicates that are more dissimilar than other replicates when viewing the dendrogram often have spectra that vary compared to the other replicates. For example, even though Fiber A replicates all remained clustered near each other, A3 and A4 are clustered together. When looking at the spectra for Fiber A,

replicates 3 and 4 are more intense compared to the other replicates. With Fiber E, replicates 6 and 10 are clustered into an entirely different class and those spectra do look slightly different compared to the other replicates' spectra. Also, E3 has a higher intensity and is separated from other replicates on the dendrogram. Lastly, replicate 8 on Fiber B is very intense in comparison to the other replicates of Fiber B and it is separated on the dendrogram. Though the other spectra for Fiber B look similar, the split between them (replicates 6, 10, 9, 4, and 5 versus 2, 1, 7, and 3) is clear.

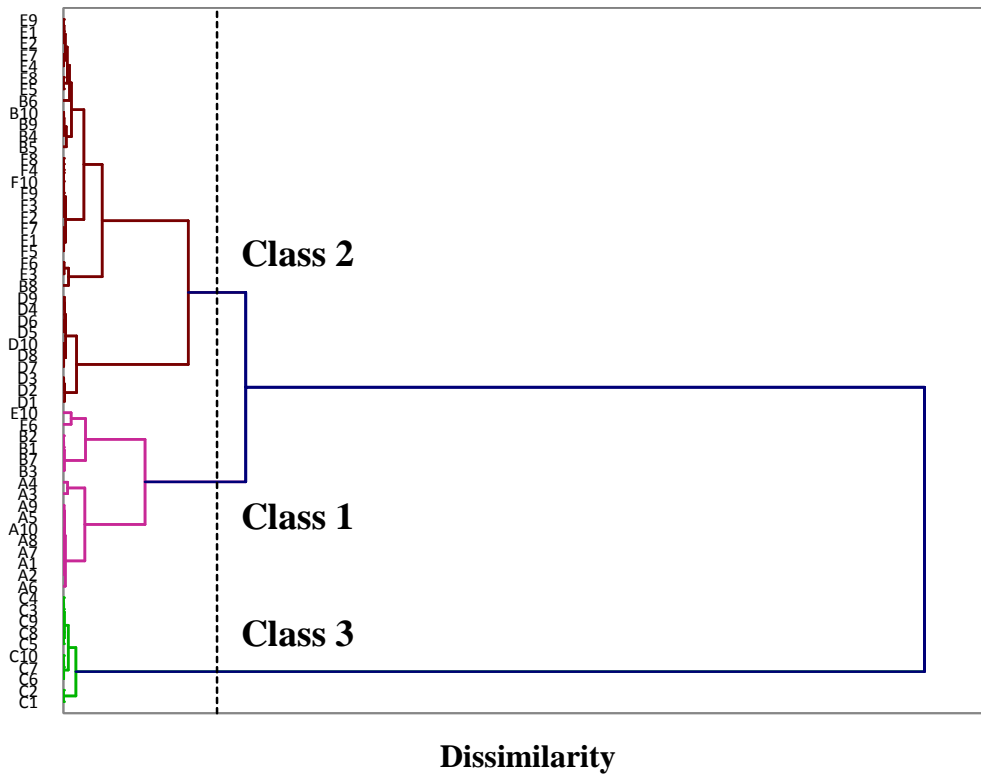


Figure 3.3 Dendrogram from AHC of Fibers A – F analyzed at IUPUI.

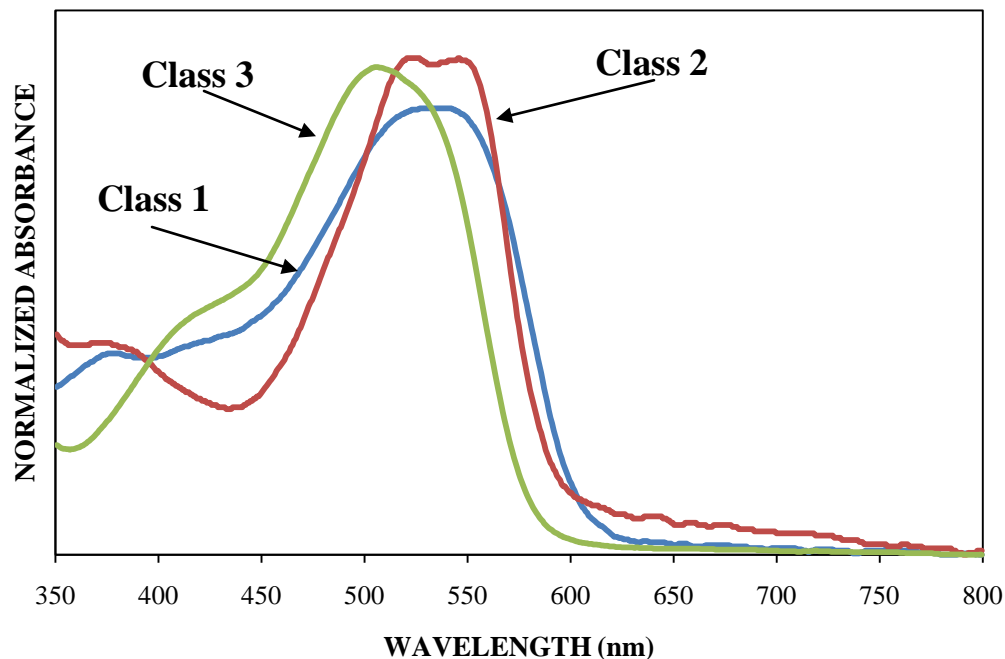


Figure 3.4 Central Objects of the Three Clusters from the Dendrogram.

For the purposes of PCA and DA, every scan was used instead of utilizing only the averages. The supplemental data, which was the additional two fibers analyzed per dye, was included in order for it to receive a factor score for future use, but it was not incorporated into the PCA analysis. The observations plot from PCA is shown in Figure 3.5. The plot displays the first two principal components, which captured 66.57% of the total variance of the data set. The same three classes from AHC were color-coded to be distinguishable. Through the three classes are distinct, some overlap is present. Similar to AHC, Class 3 is more distinct than the other two classes. The total variance in two dimensions is low, meaning more separation could be occurring at higher dimensions. All the replicates were placed into one class for this observation plot, as well as for

future DA analysis. Whichever class had the majority of the replicates was the class chosen to contain all of the replicates.

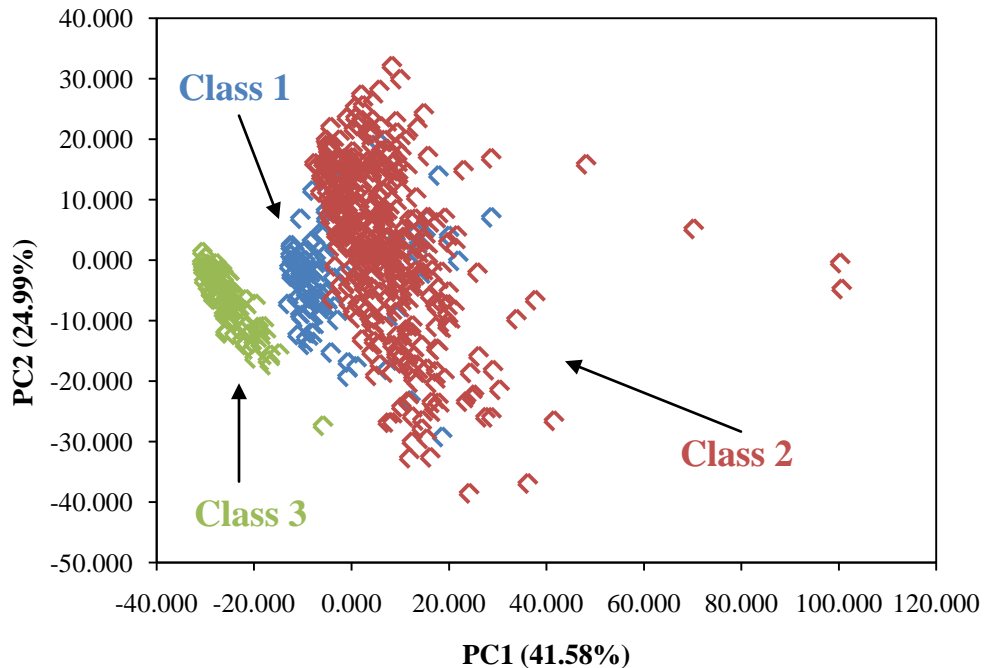


Figure 3.5 Observations Plot from PCA.

A factor loadings plot of the original variables can be used to illustrate the PCA coordinate system. A different type of factor loadings plot was created for this data set (Figure 3.6) where the loadings in the form of the cosine of the angle between the principal component and each variable. Areas where the cosine is positive are areas of positive correlation, areas where the cosine is negative are areas of negative correlation, and areas where the cosine is close to zero have no correlation. PC1 corresponds to the x-axis and PC2 corresponds to the y-axis of the observations plot (Figure 3.5).

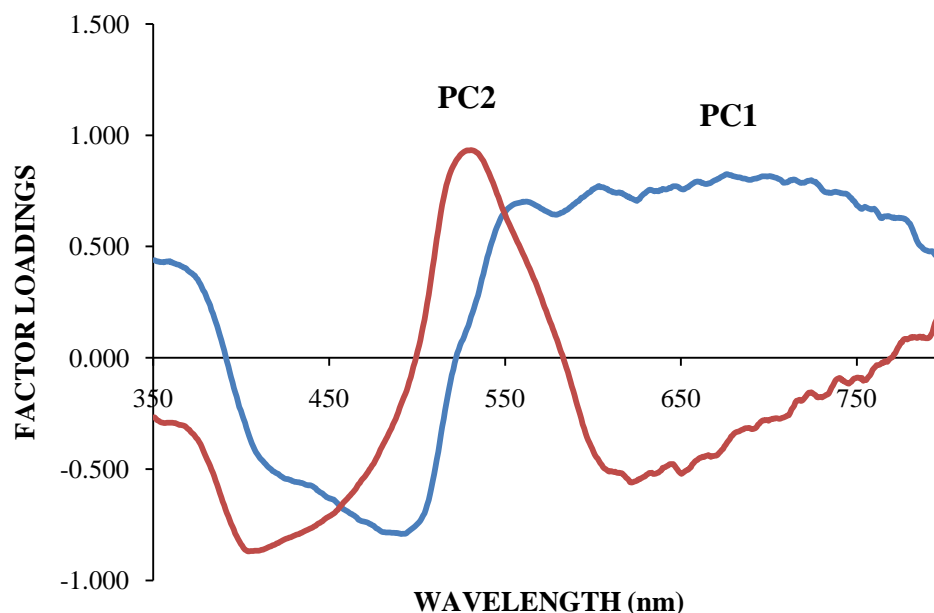


Figure 3.6 Factor Loadings Plot of the First Two PCs.

Focusing on PC1, the main areas of correlation can be found, and the wavelengths corresponding to this area can be highlighted on the central objects plot of all the clusters together. The purple bracketed areas in Figure 3.7 indicate these wavelength regions, where there is a negative correlation at approximately 460 – 520 nm and a positive correlation at 550 – 750 nm. In an attempt to put more physical meaning behind the spectra, the ordering of the clusters within the bracket correlates with the ordering of the clusters in the PCA observations plot. Samples with high absorbance in regions where there are high factor loadings on PC1 should be on the right of the PCA observations plot because those are the large and most positive values. The inverse is true as well; low factor loadings on PC1 should be on the left of the PCA observations plot because those are the smallest and least positive values. Therefore, using both significant areas, the

observations plot should read (from left to right): Class 3, Class 1, and then Class 2. Referring back to Figure 3.5, this is the exact trend that is noticed in the observations plot. In addition, both significant areas illustrate Class 3 as being the most distinct, and when looking at the observations plot, this distinction is also noticeable.

The same physical meaning can be applied to the second principal component. Two areas of significant correlation are highlighted on the central objects plot (Figure 3.7) by orange brackets. The area of positive correlation wavelengths ranges from 525 – 555 nm, and the area of negative correlation is around 400 nm. The negative correlation from PC1 and the positive correlation from PC2 could be attributed to the absorption of the red dye on the fibers. With PC2, samples with strong absorbance in the region of high factor loadings should be the highest on the PCA observations plot since PC2 is associated with the y-axis, and the inverse is true as well. Relating this to the PCA observations plot, it would be expected that Class 2 would be the highest and Classes 1 and 3 would be more similarly located. Referring back again to Figure 3.5, this is the exact trend that is noticed.

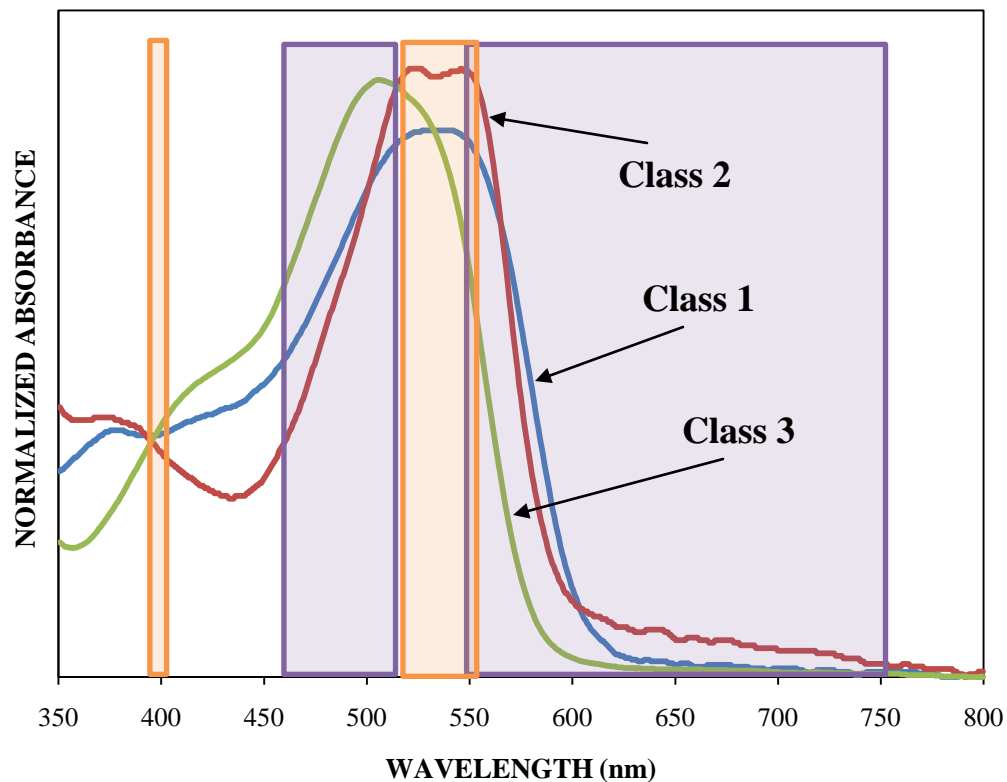


Figure 3.7 Significant Factor Loadings Overlaid on the Central Objects Plot.

DA was then performed using the data from PCA. Several methods exist that assist in determining the correct number of principal components to use. One method involves setting a targeted cumulative percentage of variance and using the number of principal components associated with that number. This method was used in this research, and 95% was the selected cumulative variance. Using this amount of cumulative variance as the criterion, the first five principal components would be selected. Another method involves using a scree plot, which displays eigenvalues against factor number. The location of a sudden break in the plot is an indicator of the number of significant principal components. Based on the scree plot for this data set, four principal components would be

used. Lastly, the Kaiser criterion can be used to determine the significant amount of principal components to utilize. Using this criterion, all eigenvalues greater than one would be considered significant, and this research this would apply to the first twelve principal components.

The results of DA using the first five principal components are shown in the observations plot in Figure 3.8. Three distinct groupings were used to be consistent with AHC. Minimal overlap can be seen between the three groups, but this will affect the cross-validation results as seen in the confusion matrix results in Table 3.4. Samples located along the diagonal in green represent those that were correctly classified, while samples in bolded red outside of this diagonal are incorrect classifications. Overall, 99.17% of the samples were correctly classified, which is considered to be an excellent result.

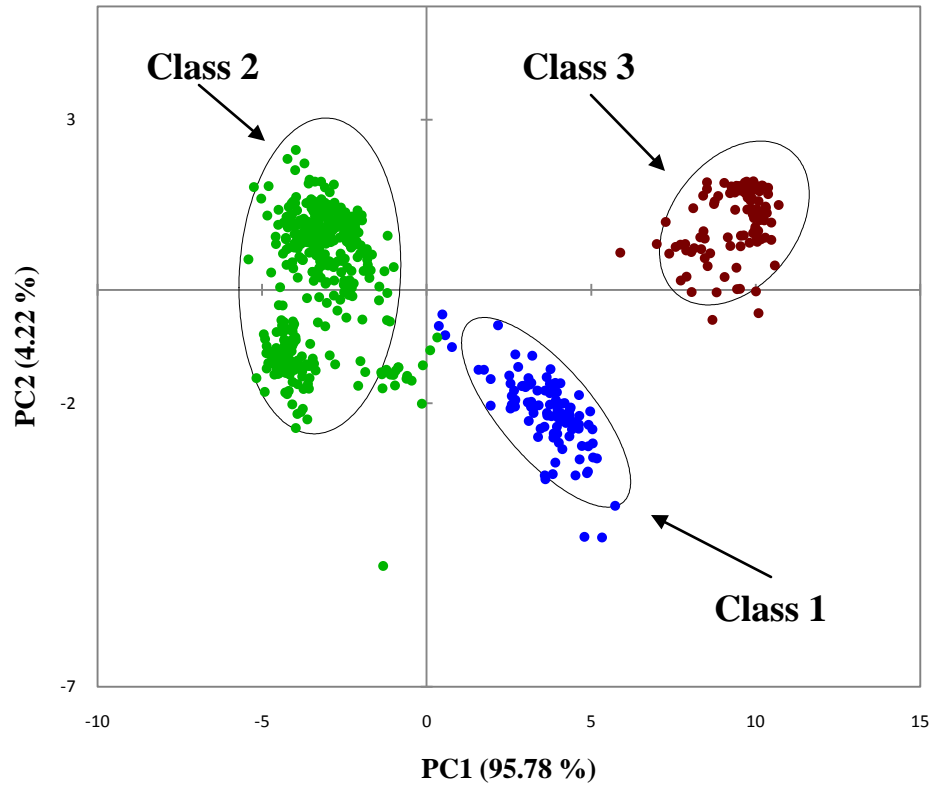


Figure 3.8 Observations Plot from DA using Three Classes.

Table 3.4 Confusion Matrix for the Cross-Validation Results from DA (Three Classes).

<i>From/To</i>	<i>Class 1</i>	<i>Class 2</i>	<i>Class 3</i>	<i>Total</i>	<i>% Correct</i>
<i>Class 1</i>	98	2	0	100	98.00
<i>Class 2</i>	2	398	0	400	99.50
<i>Class 3</i>	1	0	99	100	99.00
<i>Total</i>	108	392	100	600	99.17

DA was then performed using six classes, in which case each fiber dye was considered its own class. The observations plot for this data is shown in Figure 3.9. Using these six classes resulted in overlap between several of the classes (namely Fibers B, E, and F), whereas some fibers were able to be distinguished from the others using this plot. This overlap, though, affected the cross-validation results as seen in Table 3.5. Overall, 85.17% of the samples were correctly classified.

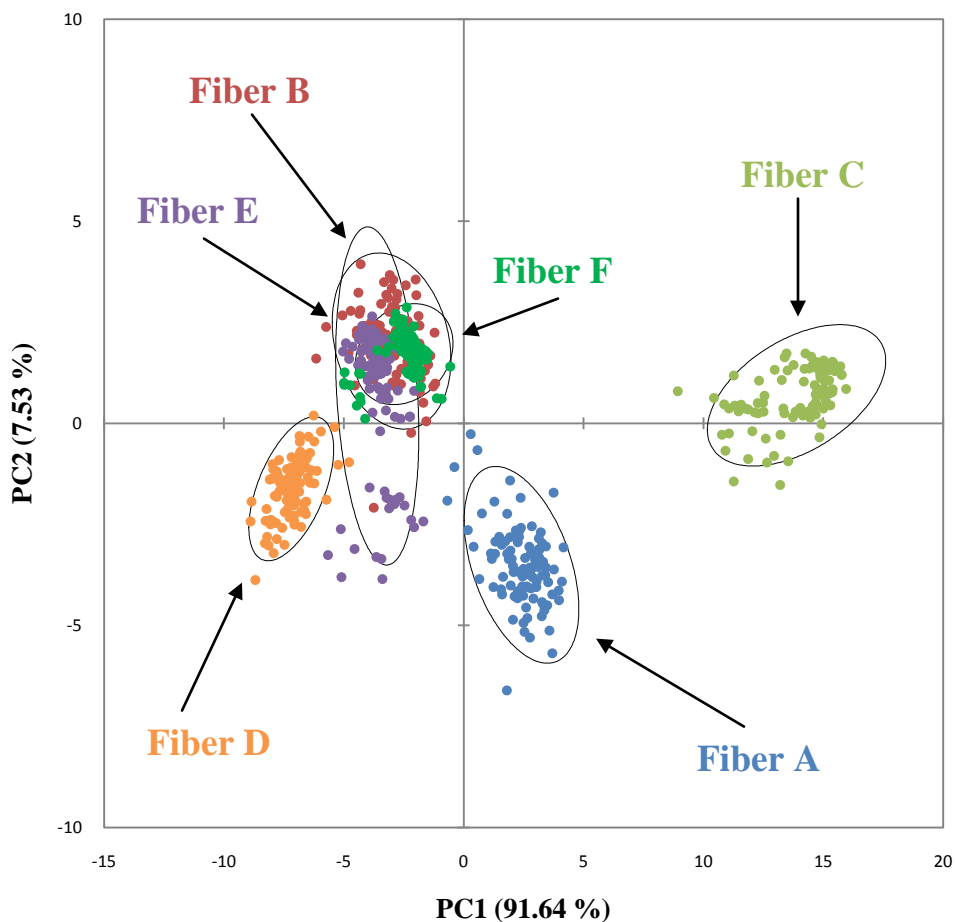


Figure 3.9 Observations Plot from DA using Six Classes.

Table 3.5 Confusion Matrix for the Cross-Validation Results from DA (Six Classes).

<i>From/To</i>	<i>Fiber A</i>	<i>Fiber B</i>	<i>Fiber C</i>	<i>Fiber D</i>	<i>Fiber E</i>	<i>Fiber F</i>	<i>TOTAL</i>	<i>% Correct</i>
<i>Fiber A</i>	98	1	0	0	1	0	100	98.00
<i>Fiber B</i>	0	74	0	0	6	20	100	74.00
<i>Fiber C</i>	0	0	100	0	0	0	100	100.00
<i>Fiber D</i>	0	0	0	98	2	0	100	98.00
<i>Fiber E</i>	0	23	0	6	56	15	100	56.00
<i>Fiber F</i>	0	4	0	0	11	85	100	85.00
<i>TOTAL</i>	98	102	100	104	76	120	600	85.17

3.3.2. External Validation

The supplemental data analyzed by another person was used as a form of external validation and the results are shown in Table 3.6, in which three classes were chosen. DA predicted which class the supplemental samples should be placed in and the correct placement was determined by where the original (non-replicate) sample was placed. The green, bolded diagonal numbers are samples that were placed correctly in the proper class. The red, bolded numbers outside of the diagonal indicate samples that were incorrectly classified. Overall, the performance of the classification model was excellent, with 98.33% of samples correctly assigned. When six classes was selected for an external validation, the classification accuracy was 87.50% (Table 3.7). When viewing this table, it is

evident that most errors were made between Classes E, F, and B, and this coincides with the observations plot in Figure 3.9.

Table 3.6 Confusion Matrix for the External Validation Results of the Supplemental Data from DA using Three Classes.

<i>From/To</i>	<i>Class 1</i>	<i>Class 2</i>	<i>Class 3</i>	<i>Total</i>	<i>% Correct</i>
<i>Class 1</i>	19	1	0	20	95.00
<i>Class 2</i>	0	80	0	80	100.00
<i>Class 3</i>	0	0	20	20	100.00
<i>Total</i>	20	80	20	120	98.33

Table 3.7 Confusion Matrix for the External Validation Results of the Supplemental Data from DA using Six Classes.

<i>From/To</i>	<i>Fiber A</i>	<i>Fiber B</i>	<i>Fiber C</i>	<i>Fiber D</i>	<i>Fiber E</i>	<i>Fiber F</i>	<i>TOTAL</i>	<i>% Correct</i>
<i>Fiber A</i>	19	0	0	0	1	0	20	95.00
<i>Fiber B</i>	0	18	0	0	2	0	20	90.00
<i>Fiber C</i>	0	0	20	0	0	0	20	100.00
<i>Fiber D</i>	0	0	0	20	0	0	20	100.00
<i>Fiber E</i>	0	4	0	5	11	0	20	55.00
<i>Fiber F</i>	0	2	0	0	1	17	20	85.00
<i>TOTAL</i>	19	24	20	25	15	17	120	87.50

3.3.3. Limitations to the Study

As mentioned before, when analyzing these fibers, an attempt was made to analyze the varying shades of the dye along the fiber (dark, medium, light). This method was decided upon after reading several articles that performed their research similarly.⁸⁸⁻⁹¹ However, it was noticed that in some cases chemometric analysis does not classify the replicate scans from the same sample into the same class. Even though the general shapes of the spectra are similar, the normalized intensities vary and thus can result in misclassification. After referring to SWGMAT's fiber analysis guidelines, it was determined for future research to analyze areas on the fiber that are similarly dyed. Therefore, the chemometric analysis will be truly determining if fiber dyes can be classified and how this compares amongst various dyes.

In addition, when these mounted fibers were revisited months after their initial sample preparation, many of the fibers had lost their dye (they were now clear). However, the glycerin in which the fibers were placed in was not red, pink, or any other shade of color. This issue will be explored further in Chapter 5 (Future Directions). Therefore, it was determined for future research to prepare the samples and analyze them as soon as possible to avoid a dye loss issue.

3.3.4. Conclusions

Based on the results discussed above, MSP is an appropriate and efficient step in the analysis of red fiber dyes. Proper sample preparation has been shown to be vital in obtaining accurate spectral data, such as exposure to time in

glycerin, as well as which shades of the fiber dye to analyze and what ranges of the dye. In addition, promising results were found when more than one person analyzed fibers, as evident from the external validation. Even though it was determined that sampling various shades of the dye was not beneficial for chemometric analysis, when both people analyzing fibers did the same routine analysis, the results were similar. Based on the external validation, almost 100% of the supplemental data was correctly classified with the original data set.

Overall, the comparison of fiber dyes using AHC, PCA, and DA resulted in several findings. Three distinct groups formed within the collection of fiber dyes. The central objects of the AHC dendrogram (Figure 3.4) illustrate that the three classes differ in their relative absorptions at approximately 500 nm, 525 nm and 540 nm. Overall, the three classes were differentiated with high accuracy using both cross-validation and external validation samples when three classes determined by AHC were used. However, when using seven classes to represent each dye differently, the classes were not differentiated well for several of the dyes, specifically Fibers E and F (Reactive Red 2 and Reactive Red 228). Other dyes, such as Fibers C (Reactive Red 123), A (Direct Red C-380), and D (Reactive Red 195), were differentiated very well.

3.4. Results and Discussion – Part II: All Dyes Analyzed – IUPUI vs. ISP

3.4.1. Statistical Results of IUPUI Fiber Analysis

The AHC dendrogram for all twelve fibers analyzed at IUPUI is shown in Figure 3.10, with three distinct classes being formed based on the truncation line. Table 3.8 lists which fibers were placed into specific classes. Bifurcations to the right to the immediate right of the truncation line are more significant in determining the number of classes. In addition, the averages for each of the fiber dyes were used for AHC. Three classes were selected to remain similar to the previous IUPUI fiber study, as well as the future ISP study. When more than three classes were tried, the replicates remained clustered together through seven classes. This allowed for five fibers to be completely distinguished from the others without any replicates being separated. The AHC dendrogram for seven classes is shown in Figure 3.11, and Table 3.9 shows the specific class formation. Since three classes will be used for the comparative study, Figure 3.12 illustrates the central objects of each of the three classes, with each having distinct absorbance spectra.

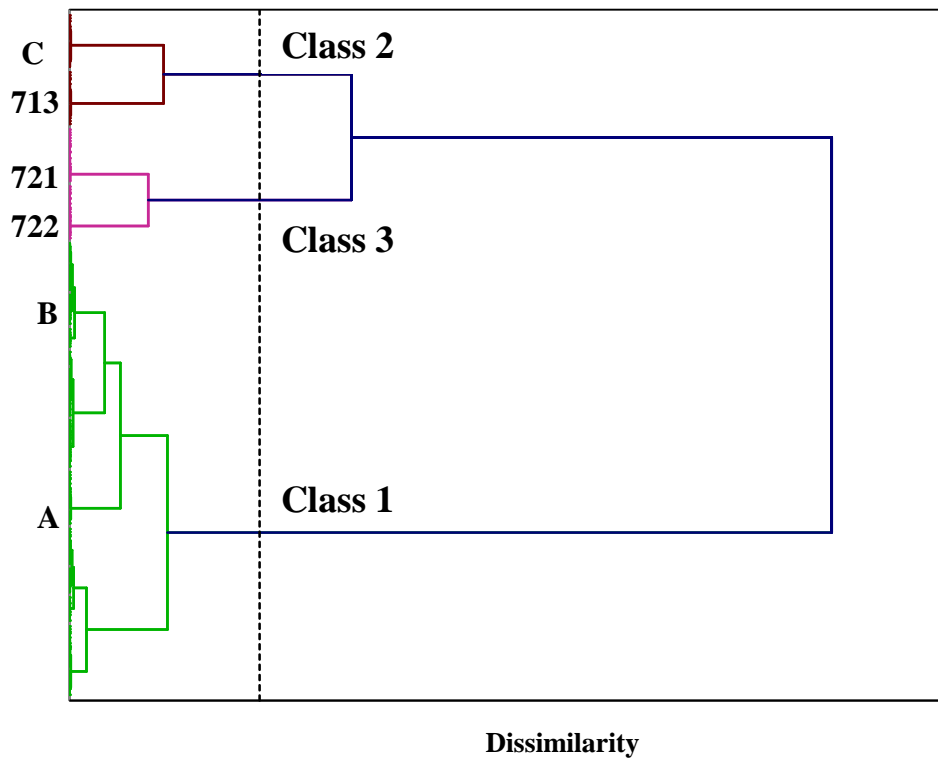


Figure 3.10 AHC Dendrogram of new IUPUI fibers (12) using Three Classes.

Table 3.8 Class Formation from the AHC Dendrogram of new IUPUI fibers using Three Classes.

CLASS 1	CLASS 2	CLASS 3
A	C	721
B	713	722
D		
E		
F		
685		
686		
695		

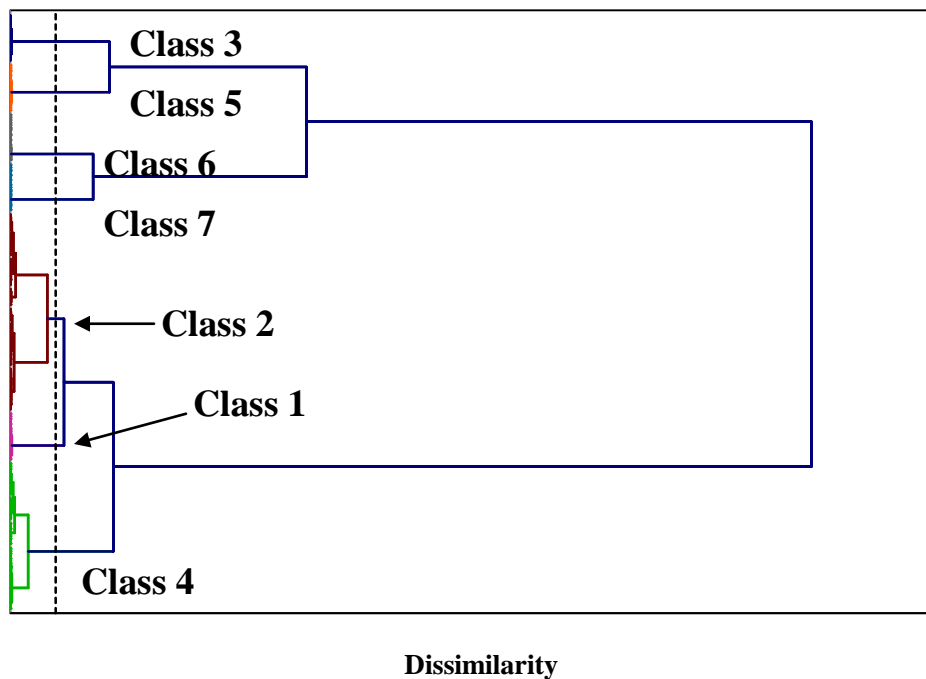


Figure 3.11 AHC Dendrogram of new IUPUI fibers (12) using Seven Classes.

Table 3.9 Class Formation from the AHC Dendrogram of new IUPUI fibers using Seven Classes.

Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7
A	B	C	D	713	721	722
	E		685			
	F		695			
	686					

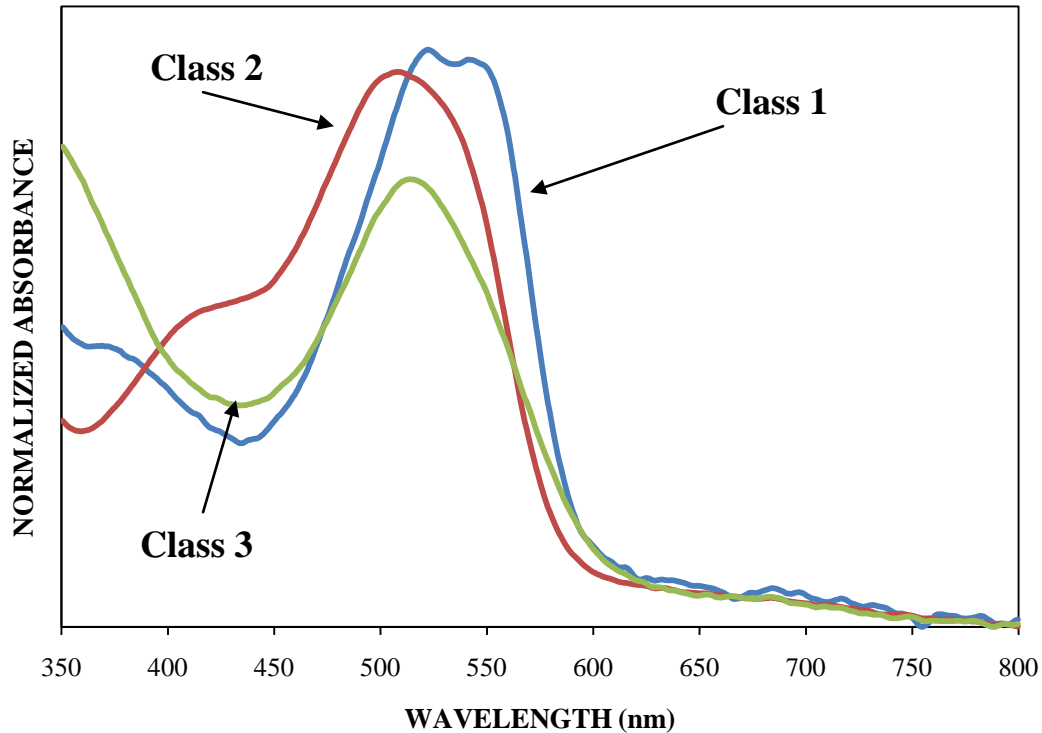


Figure 3.12 Central Objects of the Three Clusters from the Dendrogram of IUPUI new fibers.

When looking more closely at the dendrogram (Figure 3.10), Fibers A, B, C, 713, 721, and 722 are well-behaved and have replicates clustering together. Fiber 695 has Fiber D replicates mixed within its clustering. Noticeable outliers can also be viewed. For example, replicates 7, 8, and 9 of Fiber E are located below the other E replicates on the dendrogram. Those three separated replicates have a slightly increased intensity compared to the other replicates. In addition, replicate 1 of Fiber 686 looks different (lower intensity) and is separated on the dendrogram. Fiber F replicates all look similar with the exception of replicates 10 and 8 having a slightly lower intensity. However, the dendrogram split up several of the replicates. Therefore, the AHC dendrogram can identify outliers that have slight differences that are not outwardly visible.

For the purposes of PCA and DA, every scan was used instead of utilizing only the averages. The supplemental data to be used in an external validation was included in order for it to receive a factor score for future use, but it was not incorporated into the PCA analysis. The observations plot from PCA is shown in Figure 3.13. The plot displays the first two principal components, which captured 68.93% of the total variance of the data set. The same three classes from AHC were color-coded to be distinguishable. Though the three classes are distinct, some overlap is present. Class 2 can even be seen to separate within itself. The total variance in two dimensions is low, meaning more separation could be occurring at higher dimensions. In addition, the PCA observations plot for seven classes is displayed in Figure 3.14. More overlap can be seen, but certain fibers are removed from the larger group of data points. Again, since the total variance is low, several of these groups might be separated at higher dimensions.

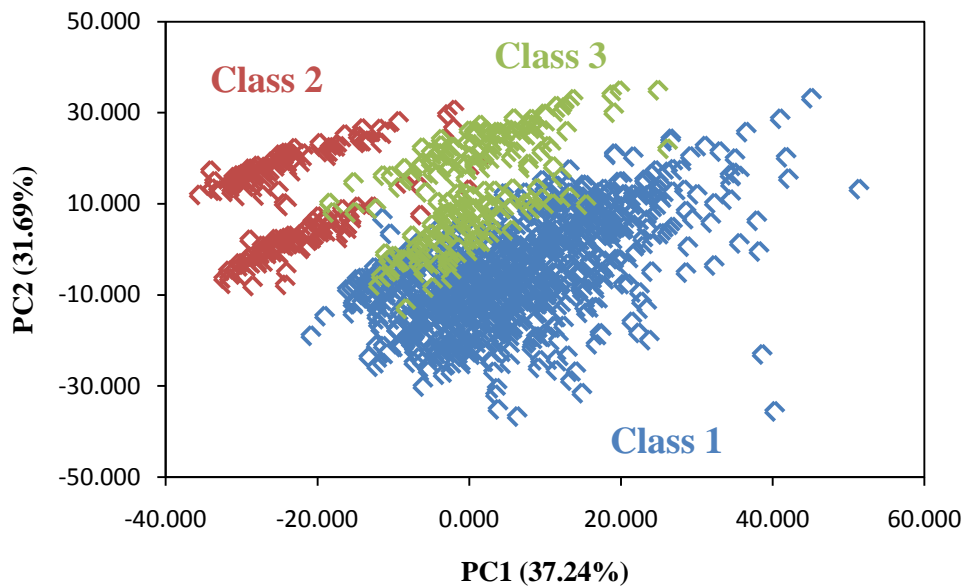


Figure 3.13 PCA Observations plot of IUPUI new data using Three Classes from AHC.

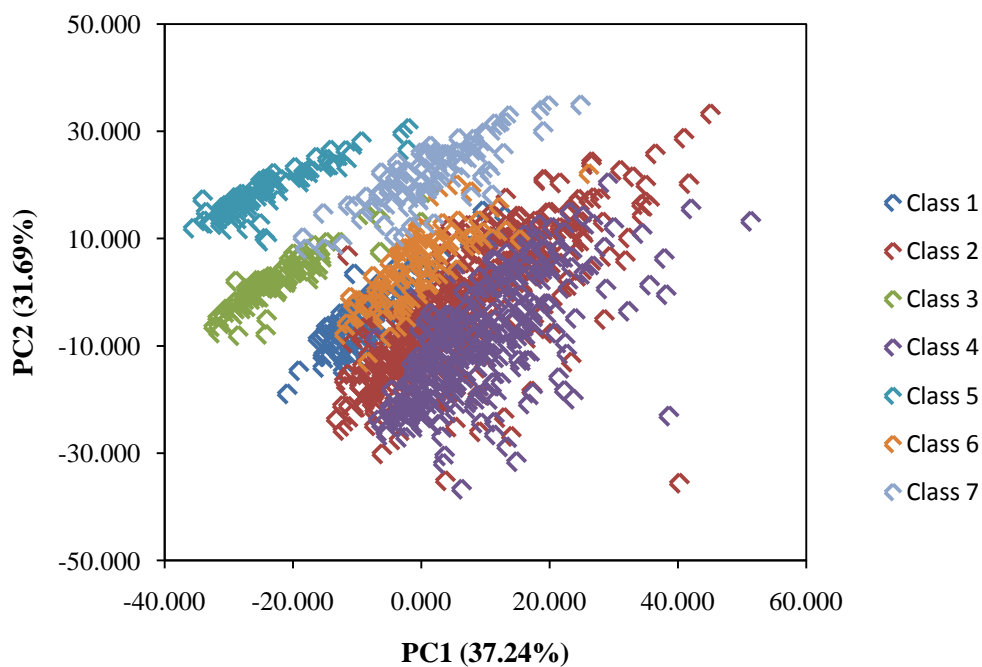


Figure 3.14 PCA Observations plot of IUPUI new data using Seven Classes from AHC.

A factor loadings plot of the original variables can be used to illustrate the PCA coordinate system. Similar to before, a different type of factor loadings plot was created for this data set (Figure 3.15) where the loadings in the form of the cosine of the angle between the principal component and each variable. Areas where the cosine is positive are areas of positive correlation, areas where the cosine is negative are areas of negative correlation, and areas where the cosine is close to zero have no correlation. PC1 corresponds to the x-axis and PC2 corresponds to the y-axis of the observations plot (Figure 3.13). Only the factor loadings for the AHC three clusters will be explored.

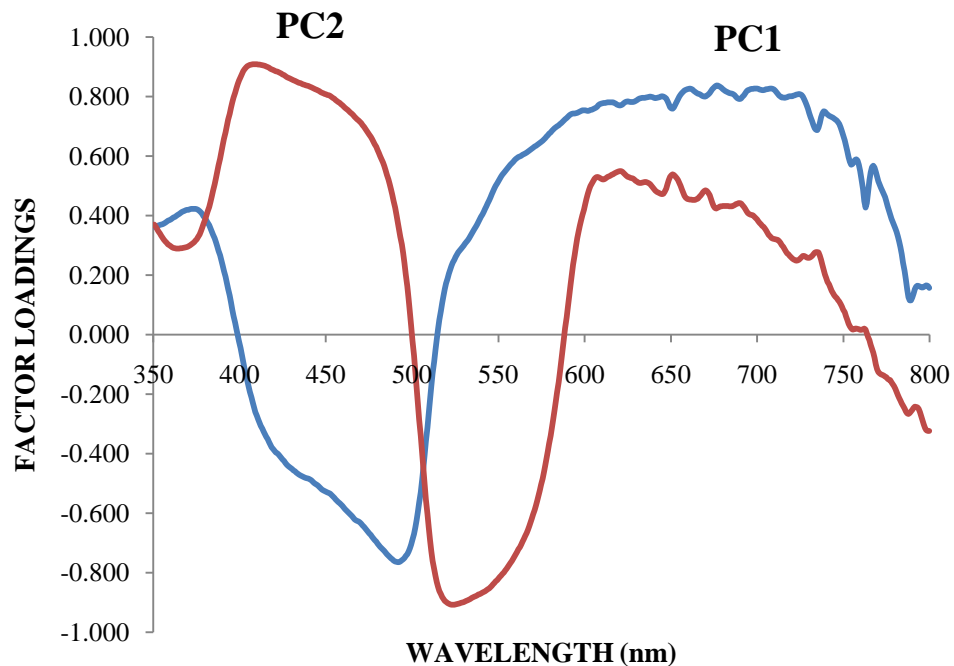


Figure 3.15 Factor Loadings Plot of the First Two PCs.

Focusing on PC1, the main areas of positive and negative correlation can be found, and the wavelengths corresponding to this area can be highlighted on the central objects plot of all the clusters together. The purple bracketed areas in Figure 3.16 indicates this wavelength regions, which is approximately 460 – 505 nm for the negative correlation and 550 – 750 nm for the positive correlation. In an attempt to put more physical meaning behind the spectra, the ordering of the clusters within the bracket correlates with the ordering of the clusters in the PCA observations plot. Samples with high absorbance in regions where there are high factor loadings on PC1 should be on the right of the PCA observations plot because those are the large and most positive values. The inverse is true as well; low factor loadings on PC1 should be on the left of the PCA observations plot because those are the smallest and least positive values. Therefore, the observations plot should have Class 2 furthest to the left, and Classes 1 and 3 similarly located along the x-axis to the right of Class 2 when using the negative correlation wavelengths. Using the positive correlation wavelengths, Class 1 should be the furthest to the right, and Classes 2 and 3 should be similarly located to the left of Class 1. Referring back to Figure 3.13, this is the exact trend that is noticed in the observations plot.

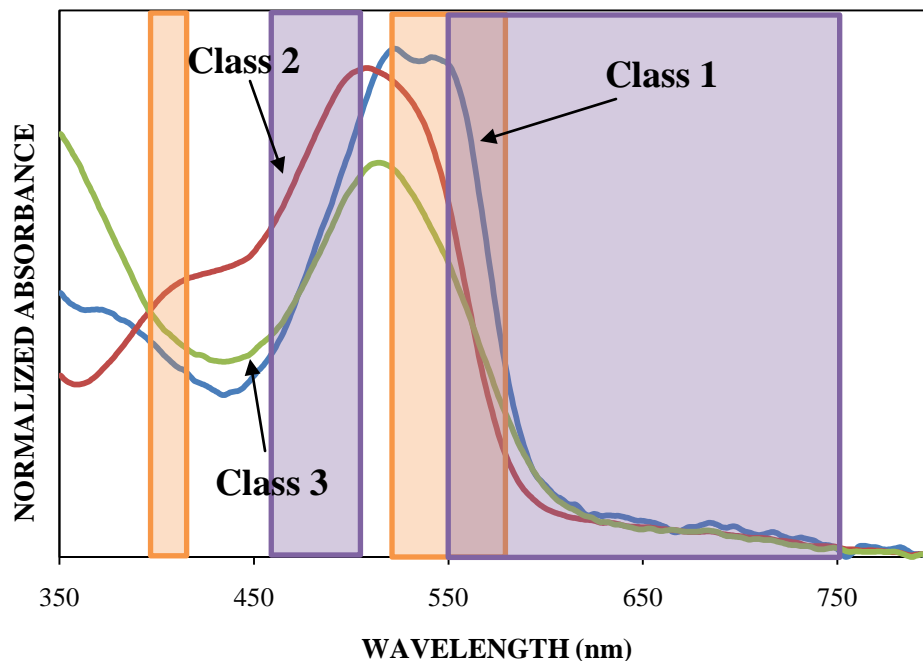


Figure 3.16 Significant Factor Loadings Overlaid on the Central Objects Plot.

The same physical meaning can be applied to the second principal component. Areas of negative and positive correlation are highlighted on the central objects plot (Figure 3.16) by orange brackets. The area of negative correlation wavelengths ranges from 525 – 575nm, and the area of positive correlation wavelengths is around 400 – 415 nm. This negative correlation area, as well as PC1's significant negative correlation area, could be attributed to the absorption of the red dye on the fibers. With PC2, samples with strong absorbance in the region of high factor loadings should be the highest on the PCA observations plot since PC2 is associated with the y-axis. The inverse is also true, such that areas of low factor loadings should be the lowest on the PCA observations plot. Relating this to the PCA observations plot, it would be expected that Class 1 would be the lowest and Classes 2 and 3 would be more

similarly located above Class 1. Referring back again to Figure 3.13, this is the exact trend that is noticed.

DA was then performed using the data from PCA. Several methods exist that assist in determining the correct number of principal components to use. One method involves setting a targeted cumulative percentage of variance and using the number of principal components associated with that number. This method was used in this research, and 95% was the selected cumulative variance. Using this variance, the first eight principal components would be selected. Another method involves using a scree plot, which displays eigenvalues against factor number. The location of a sudden break in the plot is an indicator of the number of significant principal components. Based on the scree plot for this data set, four principal components would be used. Lastly, the Kaiser criterion can be used to determine the significant amount of principal components to utilize. Using this criterion, all eigenvalues greater than one would be considered significant, and this research this would apply to the first fifteen principal components.

The results of DA using the first eight principal components are shown in the observations plot in Figure 3.17. Overall, DA accounted for 100% of the variance in two dimensions. Three distinct groupings were used to be consistent with AHC. No overlap can be seen between the three groups, and this is reflected in the cross-validation results as seen in the confusion matrix results in Table 3.10. Samples located along the diagonal in green represent those that

were correctly classified. Overall, 100.00% of the samples were correctly classified, which is considered to be an excellent result.

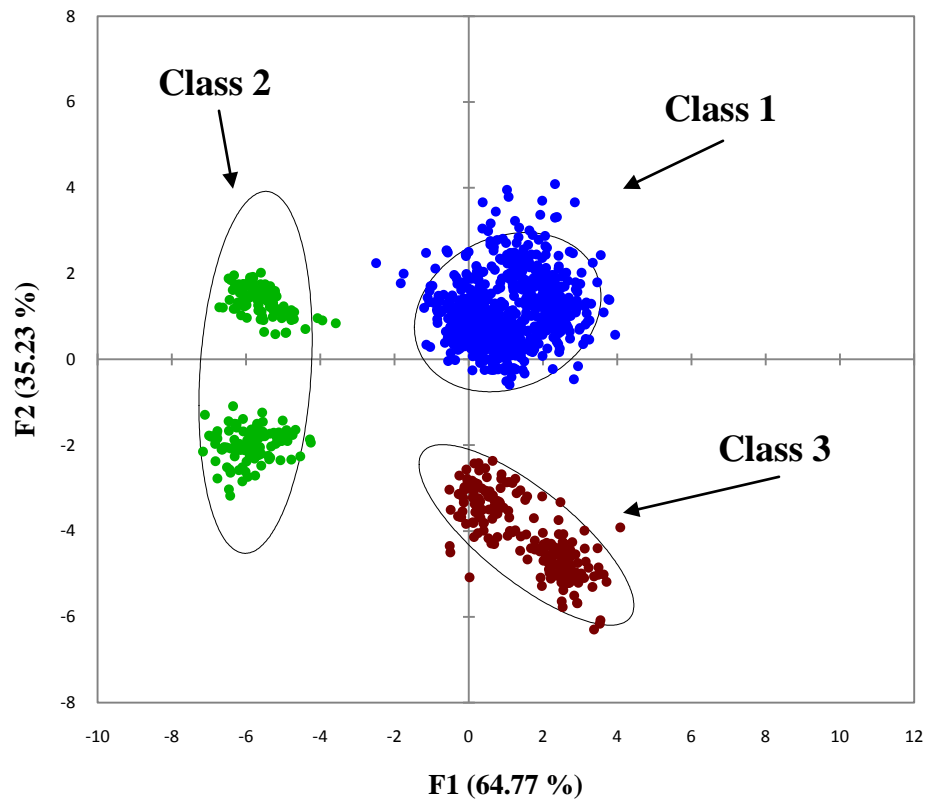


Figure 3.17 Observations Plot from DA using the Three AHC Designated Classes.

Table 3.10 Confusion Matrix for the Cross-Validation Results from DA (Three Classes).

<i>From/To</i>	<i>Class 1</i>	<i>Class 2</i>	<i>Class 3</i>	<i>Total</i>	<i>% Correct</i>
<i>Class 1</i>	800	0	0	800	100.00
<i>Class 2</i>	0	200	0	200	100.00
<i>Class 3</i>	0	0	200	200	100.00
<i>Total</i>	800	200	200	1200	100.00

DA was then performed using the seven classes designated from AHC.

The observations plot for this data is shown in Figure 3.18, in which 86.62% of the variance was accounted for in two dimensions. Using these seven classes resulted in no overlap amongst the classes, and this is also evident in the confusion matrix in Table 3.11. Overall, 100.00% of the samples were correctly classified. Lastly, twelve classes were used in DA, in which each dye was its own class. The observations plot showing this data is found in Figure 3.19, where overlap can be seen between a few classes and 87.14% of the total variance is accounted for in two dimensions. Five fibers, however, can be completely distinguished from all others and these are fibers A, C, 713, 721, and 722. The other fibers are in classes that overlap, and this is shown in the confusion matrix found in Table 3.12. Overall, though, 85.00% of the samples were correctly classified which is a good result.

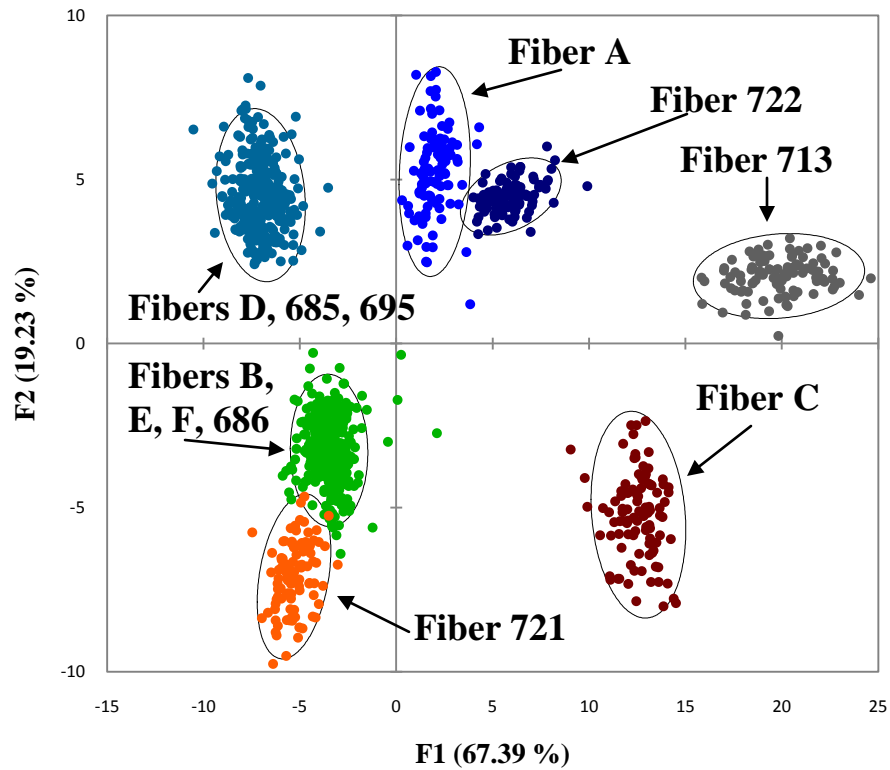


Figure 3.18 Observations Plot from DA using the Seven AHC Designated Classes.

Table 3.11 Confusion Matrix for the Cross-Validation Results from DA (Seven Classes).

<i>From/To</i>	<i>A</i>	<i>B, E, F, 686</i>	<i>C</i>	<i>D, 685, 695</i>	<i>713</i>	<i>721</i>	<i>722</i>	<i>TOTAL</i>	<i>% Correct</i>
<i>A</i>	100	0	0	0	0	0	0	100	100.00
<i>B, E, F, 686</i>	0	400	0	0	0	0	0	400	100.00
<i>C</i>	0	0	100	0	0	0	0	100	100.00
<i>D, 685, 695</i>	0	0	0	300	0	0	0	300	100.00
<i>713</i>	0	0	0	0	100	0	0	100	100.00
<i>721</i>	0	0	0	0	0	100	0	100	100.00
<i>722</i>	0	0	0	0	0	0	100	100	100.00
<i>TOTAL</i>	100	400	100	300	100	100	100	1200	100.00

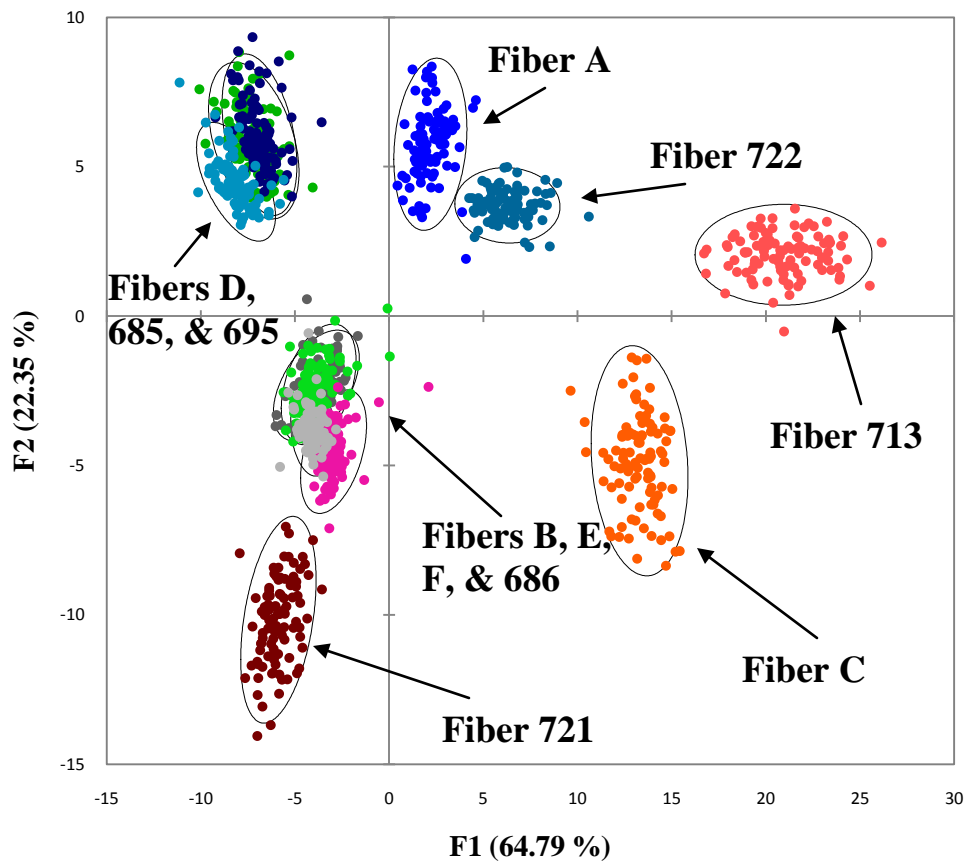


Figure 3.19 Observations Plot from DA using 12 Classes (Each Dye is its Own Class).

Table 3.12 Confusion Matrix for the Cross-Validation Results from DA (Twelve Classes).

<i>From/To</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>713</i>	<i>685</i>	<i>686</i>	<i>695</i>	<i>721</i>	<i>722</i>	<i>TOTAL</i>	<i>% Correct</i>
<i>A</i>	100	0	0	0	0	0	0	0	0	0	0	0	100	100
<i>B</i>	0	67	0	0	31	0	0	0	2	0	0	0	100	67.00
<i>C</i>	0	0	100	0	0	0	0	0	0	0	0	0	100	100
<i>D</i>	0	0	0	72	0	0	0	1	0	27	0	0	100	72.00
<i>E</i>	0	39	0	0	50	1	0	0	10	0	0	0	100	50.00
<i>F</i>	0	9	0	0	0	83	0	0	8	0	0	0	100	83.00
<i>713</i>	0	0	0	0	0	0	100	0	0	0	0	0	100	100
<i>685</i>	0	0	0	5	0	0	0	89	0	6	0	0	100	89.00
<i>686</i>	0	1	0	0	2	5	0	0	92	0	0	0	100	92.00
<i>695</i>	0	0	0	22	0	0	0	11	0	67	0	0	100	67.00
<i>721</i>	0	0	0	0	0	0	0	0	0	0	100	0	100	100
<i>722</i>	0	0	0	0	0	0	0	0	0	0	0	100	100	100
<i>TOTAL</i>	100	116	100	99	83	89	100	101	112	100	100	100	1200	85.00

3.4.2. External Validation

An external validation was performed since this data set was separating well. The supplemental data was collected several weeks after the initial data set was collected. DA predicted which class the supplemental samples should be placed in and the correct placement was determined by where the original (non-replicate) sample was placed. Table 3.13 illustrates the results of this external validation using the three AHC designated classes. The green, bolded diagonal numbers are samples that were placed correctly in the proper class. Overall, the performance of the classification model was excellent, with 100% of samples correctly assigned. Table 3.14 also illustrates the results of this external validation; however, seven AHC distinguished classes are used. Using the seven classes, 100% of the samples were correctly classified.

Table 3.13 Confusion Matrix for the External Validation Results of the Supplemental Data from DA using Three Classes.

<i>From/To</i>	<i>Class 1</i>	<i>Class 2</i>	<i>Class 3</i>	<i>Total</i>	<i>% Correct</i>
<i>Class 1</i>	160	0	0	160	100.00
<i>Class 2</i>	0	40	0	40	100.00
<i>Class 3</i>	0	0	40	40	100.00
<i>Total</i>	160	40	40	240	100.00

Table 3.14 Confusion Matrix for the External Validation Results of the Supplemental Data from DA using Seven Classes.

<i>From/To</i>	<i>A</i>	<i>B, E, F, 686</i>	<i>C</i>	<i>D, 685, 695</i>	<i>713</i>	<i>721</i>	<i>722</i>	<i>TOTAL</i>	<i>% Correct</i>
<i>A</i>	20	0	0	0	0	0	0	20	100.00
<i>B, E, F, 686</i>	0	80	0	0	0	0	0	80	100.00
<i>C</i>	0	0	20	0	0	0	0	20	100.00
<i>D, 685, 695</i>	0	0	0	60	0	0	0	60	100.00
<i>713</i>	0	0	0	0	20	0	0	20	100.00
<i>721</i>	0	0	0	0	0	20	0	20	100.00
<i>722</i>	0	0	0	0	0	0	20	20	100.00
<i>TOTAL</i>	20	80	20	60	20	20	20	240	100.00

3.4.3. Statistical Results of ISP Fiber Analysis

The AHC dendrogram for all twelve fibers analyzed at ISP is shown in Figure 3.20, with three distinct classes being formed based on the truncation line. Table 3.15 lists which fibers were placed into specific classes. Bifurcations to the right to the immediate right of the truncation line are more significant in determining the number of classes. In addition, the averages for each of the fiber dyes were used for AHC. Three classes were selected to remain similar to both IUPUI fiber studies. In addition, more than three classes results in the separating of replicates. Figure 3.21 illustrates the central objects of each of the three classes, with each having distinct absorbance spectra.

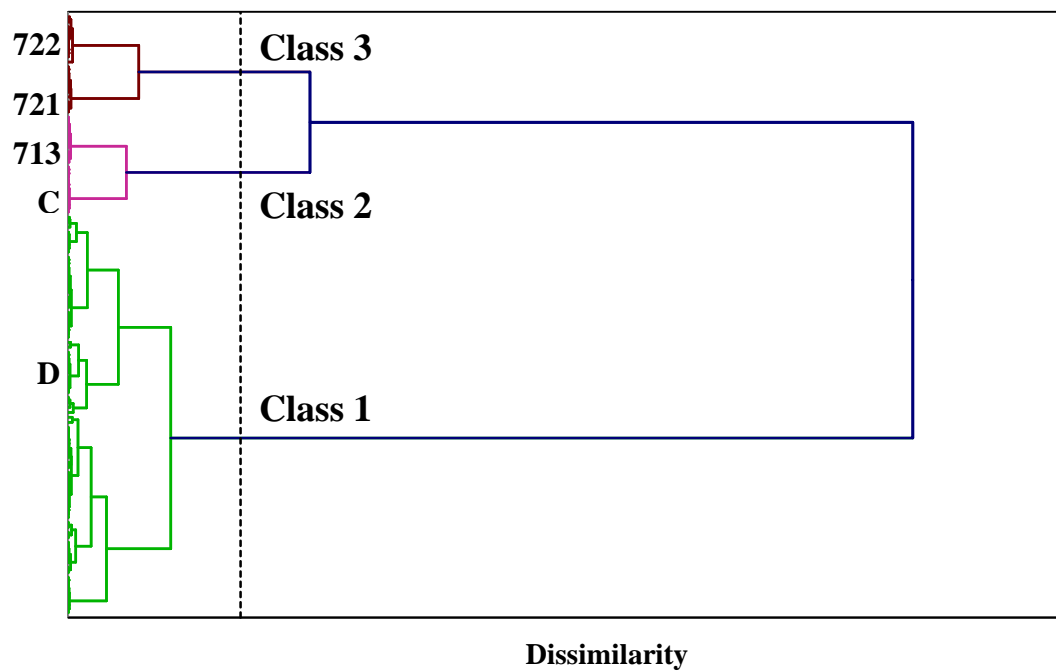


Figure 3.20 AHC Dendrogram of all ISP fibers.

Table 3.15 Class Formation from the AHC Dendrogram of ISP fibers using Three Classes.

Class 1	Class 2	Class 3
A	C	721
B	713	722
D		
E		
F		
685		
686		
695		

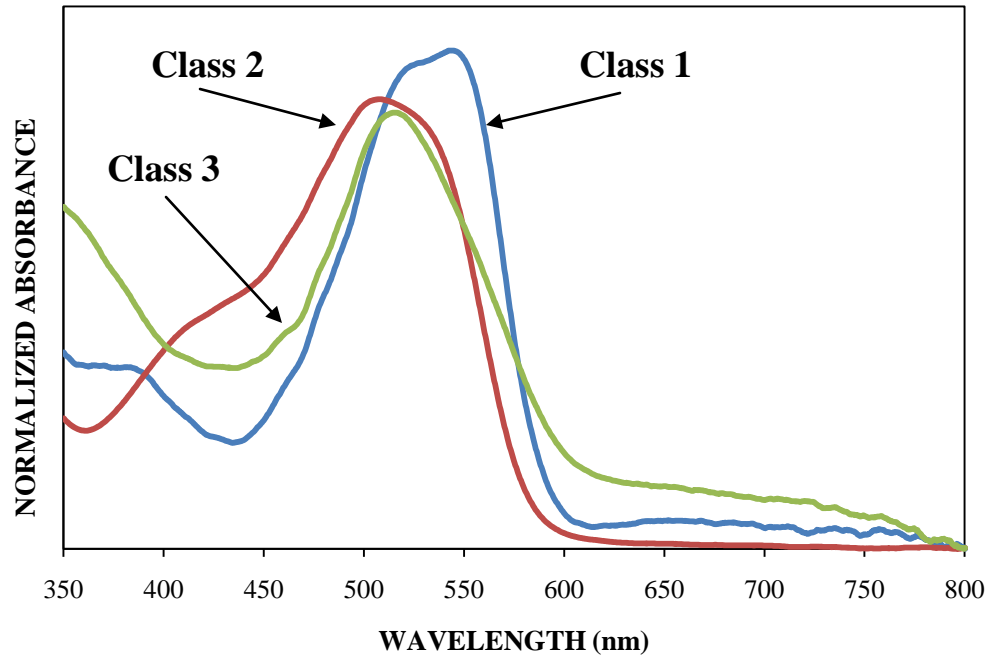


Figure 3.21 Central Objects of the Three Clusters from the Dendrogram of ISP fibers.

The dendrogram can reveal further information. Fibers that are well-behaved and cluster their replicates together include C, D, 713, 721, and 722. Fiber A has two replicates (6 and 10) that are separated from the rest of the replicates. When looking at the spectra for those two replicates, they look different than the others and thus proving why they are divided. Fiber B has many of its replicates split into smaller sections, and each of these sections have spectra that look different compared to other sections. More specifically, replicates 2, 3, and 9 have the lowest intensity and they are clustered together, and replicate 4 is the most intense and separated from the other groupings. Replicate E1 on Fiber E has a lower intensity than most of the replicates and thus is separated in the dendrogram. Overall, the separations in the dendrogram appear to be based on the intensity of the main absorbance peak. However,

Fiber 695 is the main exception that was found. Several of the replicates were individually placed throughout the dendrogram. Upon further investigation, these individual replicates had higher baselines compared to the other replicates that were clustered together.

For the purposes of PCA and DA, every scan was used instead of utilizing only the averages. The observations plot from PCA is shown in Figure 3.22.

The plot displays the first two principal components, which captured 72.12% of the total variance of the data set. The same three classes from AHC were color-coded to be distinguishable. Overlap between the classes is apparent.

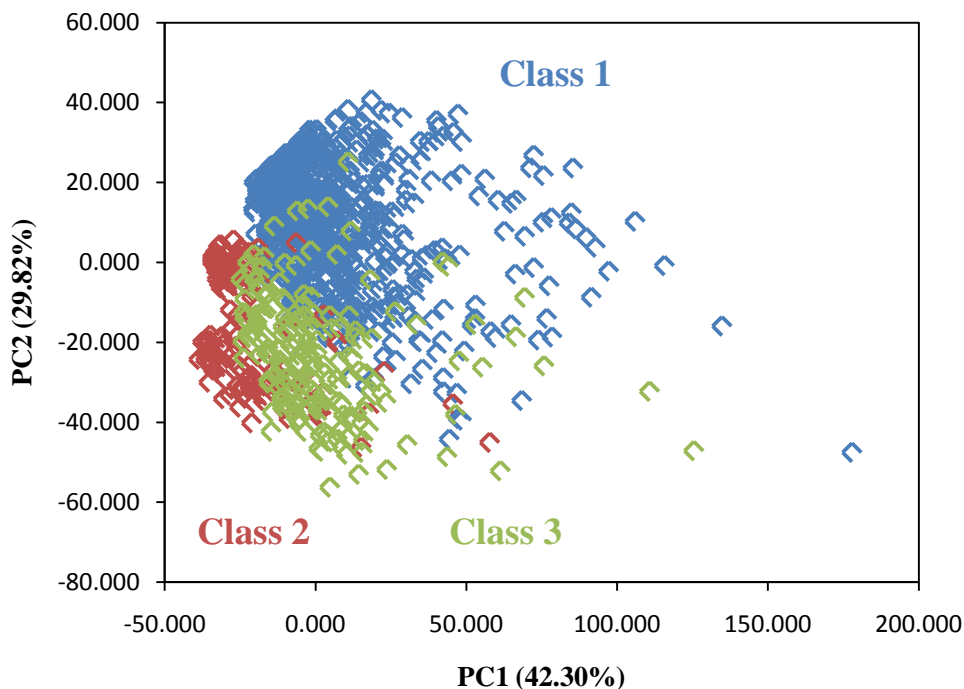


Figure 3.22 PCA Observations plot of ISP data using Three Classes from AHC.

A factor loadings plot of the original variables can be used to illustrate the PCA coordinate system. Similar to before, a different type of factor loadings plot

was created for this data set (Figure 3.23) where the loadings in the form of the cosine of the angle between the principal component and each variable. Areas where the cosine is positive are areas of positive correlation, areas where the cosine is negative are areas of negative correlation, and areas where the cosine is close to zero have no correlation. PC1 corresponds to the x-axis and PC2 corresponds to the y-axis of the observations plot (Figure 3.22).

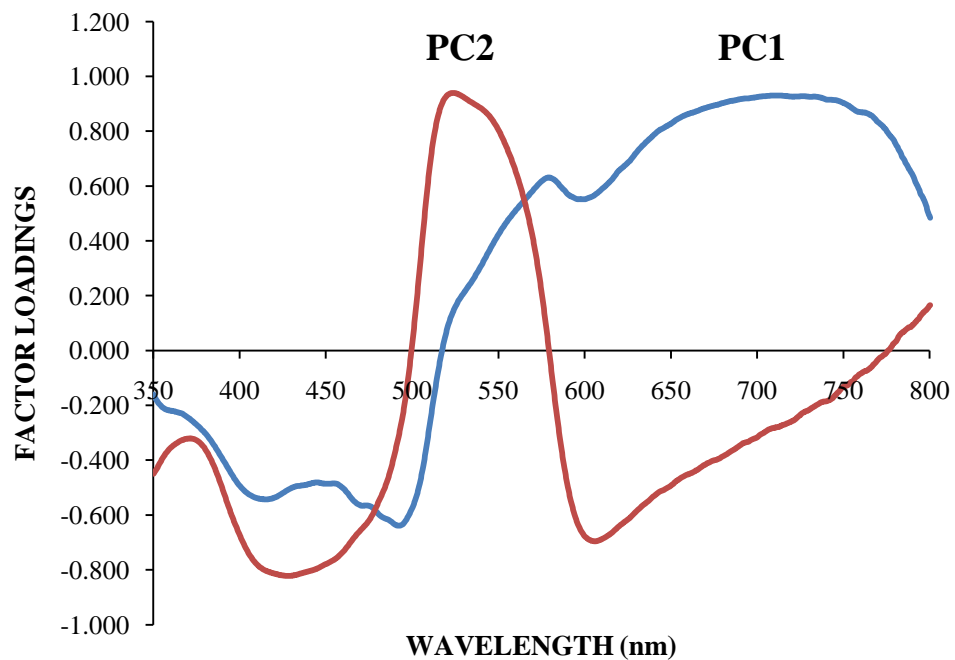


Figure 3.23 Factor Loadings Plot of the First Two PCs.

Focusing on PC1, the main areas of correlation can be found, and the wavelengths corresponding to these areas can be highlighted on the central objects plot of all the clusters together. The purple bracketed areas in Figure 3.24 indicate these wavelength regions, which are approximately 415 – 510 nm for the negative correlation, and 650 – 750 nm for the positive correlation. In an

attempt to put more physical meaning behind the spectra, the ordering of the clusters within the bracket correlates with the ordering of the clusters in the PCA observations plot. Samples with high absorbance in regions where there are high factor loadings on PC1 should be on the right of the PCA observations plot because those are the large and most positive values. The inverse is true as well; low factor loadings on PC1 should be on the left of the PCA observations plot because those are the smallest and least positive values. Therefore, using the significant negative correlation, the observations plot should read (from left to right): Class 2, Class 3, and then Class 1. The positive correlation reveals a slightly different result, but in both situations, Class 2 is the furthest to the left. Classes 1 and 3 are similarly located so variations between their significant regions are plausible. Referring back to Figure 3.22, this is the exact trend that is noticed in the observations plot.

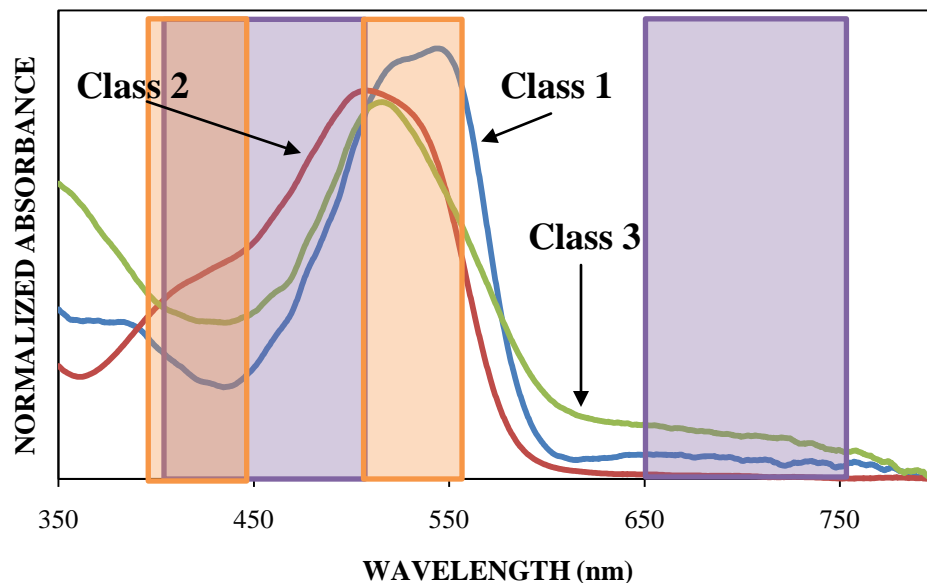


Figure 3.24 Significant Factor Loadings Overlaid on the Central Objects Plot for ISP data.

The same physical meaning can be applied to the second principal component. One area of positive correlation is highlighted on the central objects plot (Figure 3.24) by an orange bracket. This area of positive correlation wavelengths ranges from 510 – 555nm. This area, as well as PC1's significant area, could be attributed to the absorption of the red dye on the fibers. An area of negative correlation is also highlighted in orange in Figure 3.24, and these wavelengths range from 400 – 450 nm. With PC2, samples with strong absorbance in the region of high factor loadings should be the highest on the PCA observations plot since PC2 is associated with the y-axis. Relating this to the PCA observations plot, it would be expected that Class 1 would be the highest and Classes 2 and 3 would be more similarly located below Class 1. Referring back again to Figure 3.22, this is the exact trend that is noticed. The

negative correlation area reveals that Class 3 should be slightly above Class 2, and the few outliers above the main cluster from Class 3 fit this trend as well.

DA was then performed using the data from PCA. Several methods exist that assist in determining the correct number of principal components to use. One method involves setting a targeted cumulative percentage of variance and using the number of principal components associated with that number. This method was used in this research, and 95% was the selected cumulative variance. Using this variance, the first five principal components would be selected. Another method involves using a scree plot, which displays eigenvalues against factor number. The location of a sudden break in the plot is an indicator of the number of significant principal components. Based on the scree plot for this data set, the first eight principal components would be used. Lastly, the Kaiser criterion can be used to determine the significant amount of principal components to utilize. Using this criterion, all eigenvalues greater than one would be considered significant, and this research this would apply to the first twelve principal components.

The results of DA using the first five principal components are shown in the observations plot in Figure 3.25. Overall, DA accounted for 100% of the variance in two dimensions. Three distinct groupings were used to be consistent with AHC. Slight overlap can be seen between Classes 1 and 3, and this is reflected in the cross-validation results as seen in the confusion matrix results in Table 3.16. Samples located along the diagonal in green represent those that were correctly classified, and red bolded numbers outside of the diagonal

represent those that were incorrectly classified. Overall, 96.50% of the samples were correctly classified, which is considered to be an excellent result.

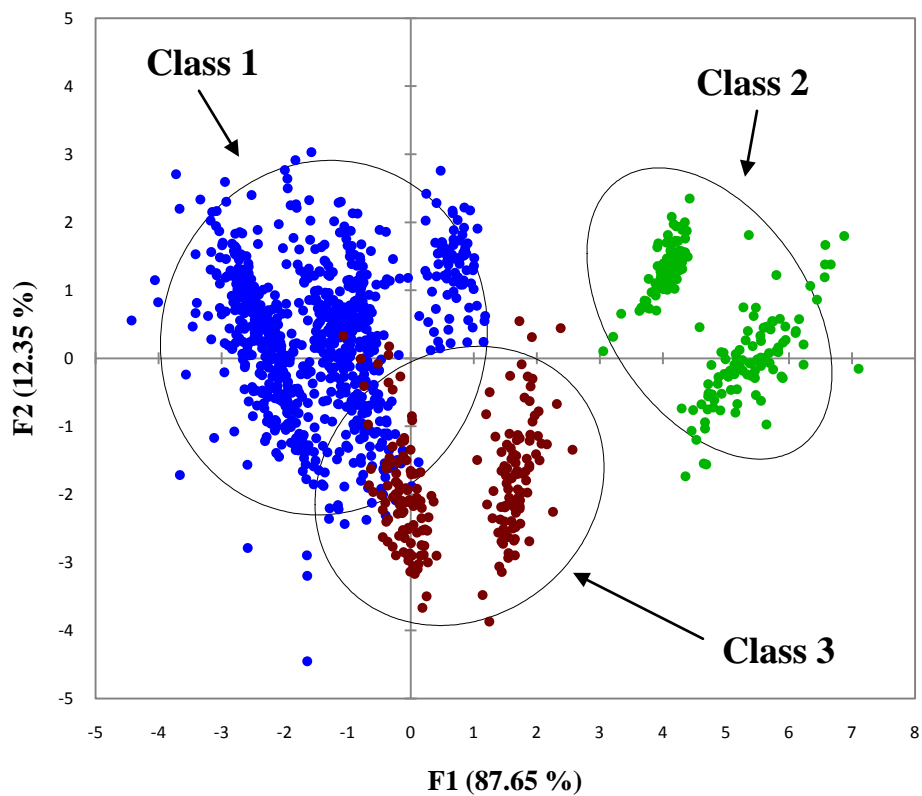


Figure 3.25 Observations Plot from DA using the Three AHC Designated Classes.

Table 3.16 Confusion Matrix of the Cross-Validation Results from DA (Three AHC-designated Classes).

<i>From/To</i>	<i>Class 1</i>	<i>Class 2</i>	<i>Class 3</i>	<i>Total</i>	<i>% Correct</i>
<i>Class 1</i>	778	0	22	800	97.25
<i>Class 2</i>	0	200	0	200	100.00
<i>Class 3</i>	19	1	180	200	90.00
<i>Total</i>	797	201	202	1200	96.50

DA was then performed using twelve classes, in which each dyes was its own class. The observations plot showing this data is found in Figure 3.26, where overlap can be seen between several classes and 96.53% of the total variance is accounted for in two dimensions. This overlap is seen in the confusion matrix found in Table 3.17. Overall, though, 81.33% of the samples were correctly classified which is a good result.

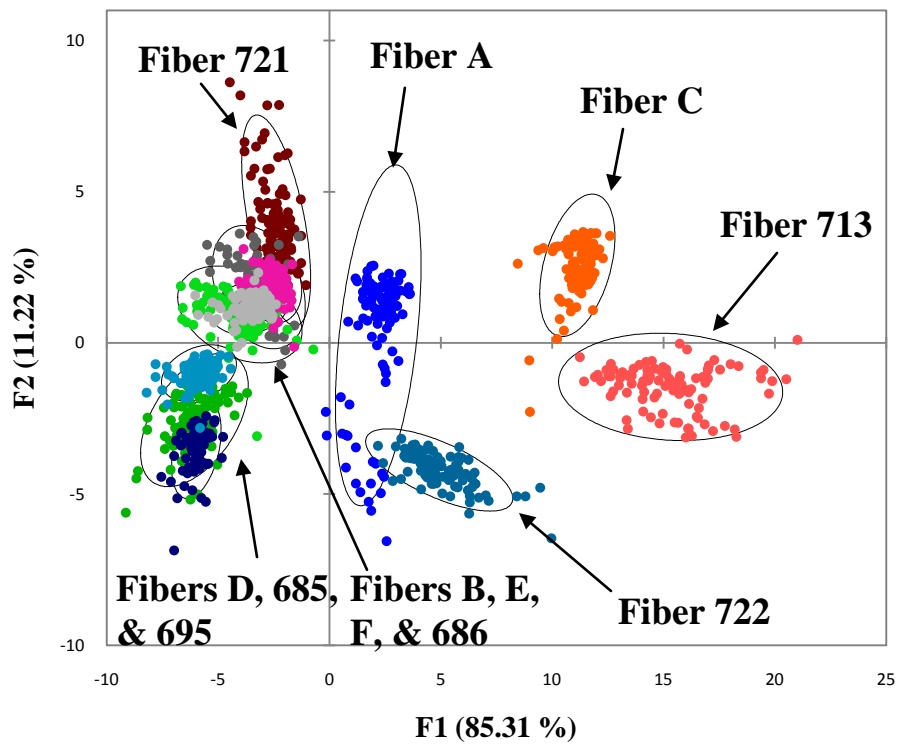


Figure 3.26 Observations Plot from DA using twelve classes (each dye is its own class).

Table 3.17 Confusion Matrix for the Cross-Validation Results from DA (Twelve Classes).

<i>From/To</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>713</i>	<i>685</i>	<i>686</i>	<i>695</i>	<i>721</i>	<i>722</i>	<i>TOTAL</i>	<i>% Correct</i>
<i>A</i>	92	0	0	0	0	0	0	0	0	0	0	8	100	92.00
<i>B</i>	0	62	0	0	14	10	0	0	10	0	4	0	100	62.00
<i>C</i>	0	0	99	0	0	0	0	0	0	0	0	1	100	99.00
<i>D</i>	0	0	0	93	0	0	0	0	0	7	0	0	100	93.00
<i>E</i>	0	22	0	0	39	0	0	3	35	1	0	0	100	39.00
<i>F</i>	0	16	0	0	2	77	0	0	5	0	0	0	100	77.00
<i>713</i>	0	0	1	0	0	0	99	0	0	0	0	0	100	99.00
<i>685</i>	0	0	0	3	1	0	0	94	0	2	0	0	100	94.00
<i>686</i>	0	7	0	0	15	8	0	0	70	0	0	0	100	70.00
<i>695</i>	0	0	0	21	0	0	0	27	0	52	0	0	100	52.00
<i>721</i>	0	1	0	0	0	0	0	0	0	0	99	0	100	99.00
<i>722</i>	0	0	0	0	0	0	0	0	0	0	0	100	100	100
<i>TOTAL</i>	92	108	100	117	71	95	99	124	120	62	103	109	1200	81.33

3.4.4. Limitations to the Study

Time played a major factor in this research. Samples were prepped and taken to ISP a one to two weeks before they were actually analyzed. Once the analysis began, a few of the fibers had already lost their dye content and therefore had to be remade and analyzed the following day. After examining the results, however, these samples remained clustered with the other replicates despite being made at a different time and analyzed on a different day. This could be attributed to analyzing the same shade along the fiber rather than recording a range of shades. Based on these outcomes, though, a few changes were made to the procedure. First, brand new samples were prepped for analysis at IUPUI even though several of the previously mounted samples still maintained their original color. Next, samples were prepped and analyzed immediately to avoid the loss of dye issue.

3.4.5. Conclusions

Based on the results discussed above, MSP is further shown to be an appropriate and efficient step in the analysis of red fiber dyes. Proper sample preparation has been shown to be vital in obtaining accurate spectral data, such as exposure to time in glycerin, as well as which shades of the fiber dye to analyze and what ranges of the dyes. After various analyses, it was determined that in order to achieve the best chemometric results possible, a consistent shade of the dye should be analyzed throughout the fibers.

The inter-laboratory comparison resulted in several findings. First, when comparing the AHC results from both IUPUI and ISP, the outcomes were exact. Three classes resulted in the same fibers placement amongst the classes. Both of the Class 1's contained fibers A, B, D, E, F, 685, 686, and 695. Class 2 contained fibers C and 713, and Class 3 contained fibers 721 and 722 for both laboratories' results. In addition, the central objects of these classes were similar between both laboratories. Different fibers were well-behaved when it came to the clustering of replicates. Fibers C, 713, 721, and 722 clustered well in both sets. However, A and B were clustered well in the IUPUI set whereas fiber D was clustered well in the ISP set. However, beyond this point, variations occurred between the data sets.

The IUPUI data can be divided into more classes without separating the replicates, whereas the ISP data cannot be separated past three classes before divisions among replicates occur. Because of this, the PCA and DA results varied. The PCA observations plot for the IUPUI data had more defined clusters but the ISP PCA observations plot did not. The same trend was seen in the DA observations plot. The IUPUI data resulted in 100% cross-validation accuracy but the ISP data had 96.50% cross-validation accuracy. Though the ISP result was excellent, it still varied from the IUPUI result.

Overall, many similarities existed between the two data sets from the different laboratories. However, discrepancies did occur. These deviations between the laboratories could be attributed to several factors. First, the fibers analyzed at ISP were not prepped and analyzed as quickly as the new fibers at

IUPUI. Though the results did not show any divergences due to this, it may have played a role in the chemometric results. For example, a slight variation may have been noticed by the statistical program that was not as noticeable to the observer. Next, more variables are present in the ISP data. The spacing between wavelength scans is shorter at ISP compared to IUPUI. Specifically, the IUPUI data set had 582 variables whereas the ISP data set has 3972 variables. Though both data sets analyzed the same wavelength range (approximately 350 – 800nm), more variables could mean more variations in the data and different features existing in the absorbance spectra. Lastly, another variation between the data sets is the number of significant principal components used. The IUPUI set utilized the first eight principal components and the ISP set used the first five principal components. The more PCs that are required, the less useful each one becomes. However, when determining the significant amount of principal components to use, the amount chosen had to account for at least 95% of the total variation.

Overall, the comparison of fiber dyes using AHC, PCA, and DA resulted in several findings. Three distinct groups formed within the collection of fiber dyes from both IUPUI and ISP. The central objects of the AHC dendrograms (Figures 3.12 and 3.21) illustrate that the three classes differ in their relative absorptions at approximately 500 nm, 520 nm and 540 nm. Overall, the three classes were differentiated with high accuracy using both cross-validation and external validation samples with the three classes determined by AHC were used. Further, the fiber dyes at IUPUI could be distinguished to seven classes,

meaning five fiber dyes were unique compared to the entire set. Looking at the structures of the dyes (Tables 3.18 – 3.20), it is evident that the classes containing several fibers (Class 2 containing fibers B, E, F, and 686; Class 4 containing D, 685, and 695) all have dyes with similar structures. The dyes that were unique (in a class of their own) had distinctive dye structures that thus set them apart. However, two of the dye structures (Fiber A [Direct Red C-380] and Fiber F [Reactive Red 228]) have not been found so no concrete conclusions can be based from the structures at this time.

Table 3.18 Dye Structures

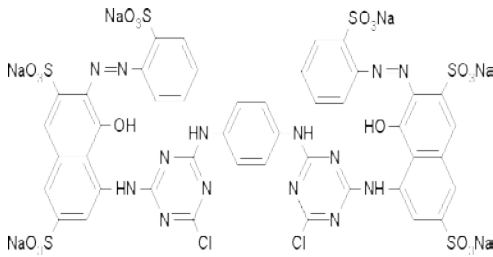
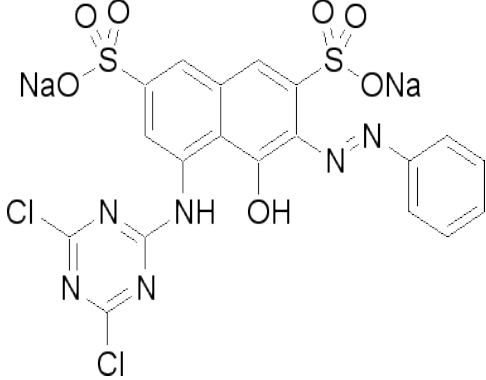
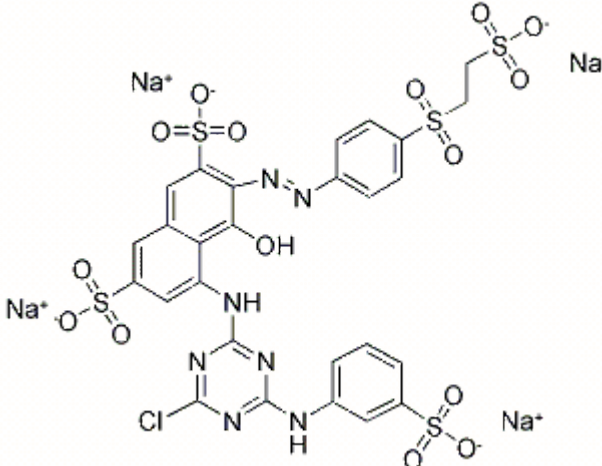
Fiber B – Reactive Red 120	Fiber E – Reactive Red 2	Fiber 686 – Reactive Red 198
 <p>The structure of Reactive Red 120 consists of two naphthalene rings connected by a central benzene ring. Each naphthalene ring has a hydroxyl group (-OH) and a chlorine atom (-Cl) at the 1-position, and a sodium sulfonate group (-SO₃Na) at the 2-position. The central benzene ring is substituted with a diazo group (-N=N-) at the 1-position and a chlorine atom (-Cl) at the 4-position.</p>	 <p>The structure of Reactive Red 2 features a central naphthalene ring with a hydroxyl group (-OH) at the 1-position and a chlorine atom (-Cl) at the 2-position. It is substituted with sodium sulfonate groups (-SO₃Na) at the 4 and 8 positions. A diazo group (-N=N-) is attached to the 5-position, which is further linked to a phenyl ring.</p>	 <p>The structure of Reactive Red 198 is a complex molecule with a central naphthalene ring. It has a hydroxyl group (-OH) at the 1-position and a chlorine atom (-Cl) at the 2-position. The 4-position is substituted with a diazo group (-N=N-) linked to a phenyl ring, which is further substituted with a sodium sulfonate group (-SO₃Na). The 5-position is substituted with a sodium sulfonate group (-SO₃Na). The 6-position is substituted with a chlorine atom (-Cl). The 7-position is substituted with a sodium sulfonate group (-SO₃Na). The 8-position is substituted with a sodium sulfonate group (-SO₃Na).</p>

Table 3.19 Dye Structures

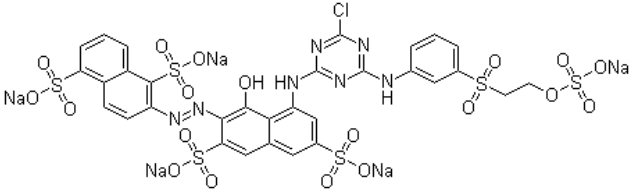
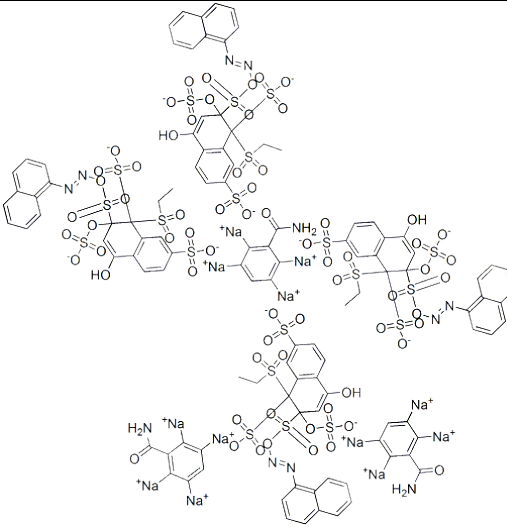
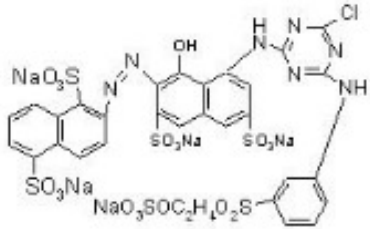
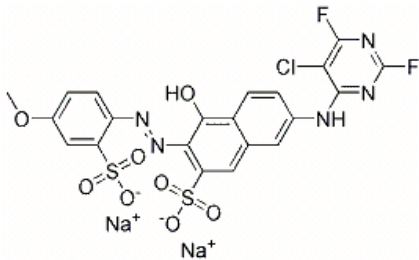
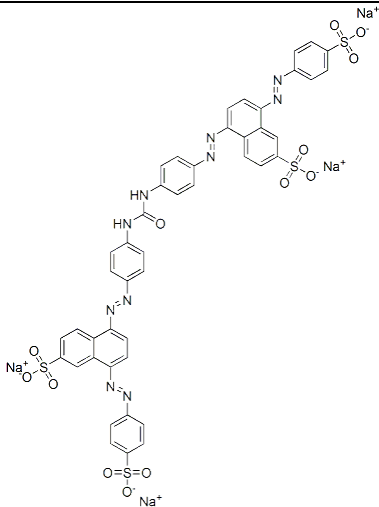
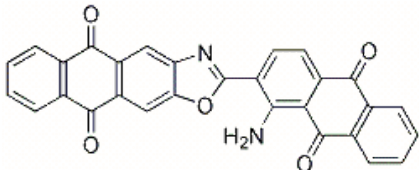
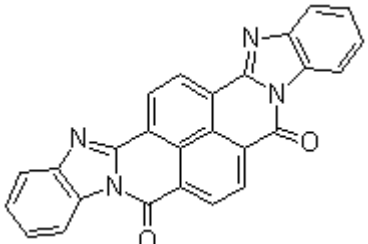
Fiber D – Reactive Red 195	Fiber 685 – Reactive Red 180	Fiber 695 – Reactive Red 239
 <p>The structure of Reactive Red 195 is a complex dye molecule. It features a central naphthalene ring system. Attached to this system are several functional groups: a chlorine atom on a pyrimidine ring, a hydroxyl group, and a sulfonate group (-SO₃Na). There are also two additional sulfonate groups (-SO₃Na) attached to the naphthalene ring. The molecule is highly polar due to the presence of these ionic groups.</p>	 <p>The structure of Reactive Red 180 is a very complex and large dye molecule. It consists of multiple fused and linked aromatic rings, including naphthalene and benzene rings. It features a variety of functional groups such as hydroxyl groups, sulfonate groups (-SO₃Na), and amine groups. The molecule is highly polar and has a complex, multi-ring architecture.</p>	 <p>The structure of Reactive Red 239 is a complex dye molecule. It features a central naphthalene ring system. Attached to this system are several functional groups: a chlorine atom on a pyrimidine ring, a hydroxyl group, and a sulfonate group (-SO₃Na). There are also two additional sulfonate groups (-SO₃Na) attached to the naphthalene ring. The molecule is highly polar due to the presence of these ionic groups.</p>

Table 3.20 Dye Structures

Fiber C – Reactive Red 123	Fiber 713 – Direct Red 84	Fiber 721 – Vat Red 10	Fiber 722 – Vat Red 15
 <p>The structure of Reactive Red 123 features a central benzene ring with a hydroxyl group (-OH) at the 1-position and a methoxy group (-OCH₃) at the 4-position. This central ring is connected via nitrogen atoms to two other benzene rings. One of these outer rings has a chlorine atom (-Cl) at the 1-position and a 2,6-difluoropyrimidin-4-ylamino group (-NH-C₄H₂N₂-F₂) at the 4-position. The other outer ring has a sodium sulfonate group (-SO₃⁻Na⁺) at the 1-position and a sodium sulfonate group (-SO₃⁻Na⁺) at the 4-position.</p>	 <p>The structure of Direct Red 84 consists of a central benzene ring with a sodium sulfonate group (-SO₃⁻Na⁺) at the 1-position and a 4-(4-sulfamoylphenyl)amino group (-NH-C₆H₄-NH₂-SO₂⁻Na⁺) at the 4-position. This central ring is connected via nitrogen atoms to two naphthalene rings. One naphthalene ring has a sodium sulfonate group (-SO₃⁻Na⁺) at the 1-position and a 4-(4-sulfamoylphenyl)amino group (-NH-C₆H₄-NH₂-SO₃⁻Na⁺) at the 4-position. The other naphthalene ring has a sodium sulfonate group (-SO₃⁻Na⁺) at the 1-position and a 4-(4-sulfamoylphenyl)amino group (-NH-C₆H₄-NH₂-SO₃⁻Na⁺) at the 4-position.</p>	 <p>The structure of Vat Red 10 is a complex polycyclic system consisting of two naphthalene rings connected by a central benzene ring. The central benzene ring has a primary amine group (-NH₂) at the 1-position and a 4-(4-sulfamoylphenyl)amino group (-NH-C₆H₄-NH₂-SO₂⁻Na⁺) at the 4-position. The naphthalene rings are substituted with carbonyl groups (-C=O) at various positions.</p>	 <p>The structure of Vat Red 15 is a complex polycyclic system consisting of two naphthalene rings connected by a central benzene ring. The central benzene ring has a primary amine group (-NH₂) at the 1-position and a 4-(4-sulfamoylphenyl)amino group (-NH-C₆H₄-NH₂-SO₂⁻Na⁺) at the 4-position. The naphthalene rings are substituted with carbonyl groups (-C=O) at various positions.</p>

CHAPTER 4. PLASTICS AND POLYMERS

Trace evidence includes a variety of samples such as paint, plastics, fibers, adhesives, and waxes. Because of this assortment, the analytical technique chosen must be sensitive and have discriminatory power for that specific material.⁹² Over the last 20-30 years, pyrolysis has been a primary technique for analyzing polymeric materials. Utilization of plastic products has increased in the past decades due to their resourcefulness and low cost,⁹³ and the technique of pyrolysis has proven to be beneficial with this type of sample.

4.1. Pyrolysis Gas Chromatography/Mass Spectrometry and its Use in Forensic Science

Pyrolysis was first used in 1860 to determine the structure of natural rubber,⁹⁴ and has become well-established in the forensic community within the past few decades. This technique involves chemical changes that take place in a substance as it is exposed to high temperatures in the absence of oxygen. These changes include molecular degradation as well as reactions between degradation products.^{94, 95} Pyrolysis is regularly used to characterize synthetic polymers, but in order to efficiently perform this task, pyrolysis must be coupled

with other analytical techniques, such as gas chromatography and mass spectrometry (GC/MS). Much research has been conducted on pyrolysis-GC and pyrolysis GC/MS.⁹³⁻⁹⁹ Pyrolysis can allow for GC analysis to occur on samples that the GC cannot separate, such as polymers due to their high molecular masses and low volatility.^{93, 99}

Advantages of pyrolysis GC/MS are that it is a quick, convenient, and powerful tool for a wide application to polymers. Polymers can be analyzed, characterized, and have their degradation studied.⁹²⁻⁹⁶ In addition, pyrolysis has a high heating speed, temperature reproducibility, and a wide temperatures range.⁹⁸ Furthermore, sample preparation is simple because pre-treatment is not required, small samples are ideal, and qualitative and quantitative results are possible. Disadvantages of pyrolysis, though, include destructiveness, lack of reproducibility, and sensitivity to sample morphology.^{96, 100}

Pyrolysis has been applied to forensic applications such as the analysis of paint,^{21, 55, 92, 101, 102} adhesives and tapes,^{92, 101, 103, 104} plastics and polymers,^{92, 93, 95, 96, 98, 101, 105} textile fibers,^{92, 101, 102, 106} environmental samples such as soil and wood,^{92, 95, 106} condom lubricants,¹⁰⁷ hair,⁹² and documents (photocopies and photocopy toners).^{17, 95, 101, 102}

Chemometrics has become increasingly important in scientific analyses, and specifically, research has been conducted on chemometrics applied to pyrolysis. In order to analyze pyrograms by multivariate statistical techniques, various ways of preprocessing and treating the data exist. An algorithm can be used to create a data matrix of peaks whose retention times and mass spectra

match throughout all samples, and thus pattern recognition is applied to relative peak areas.^{17, 105, 108} Data can also be preprocessed by using correlation optimized warping (COW) to align chromatograms and remove any baseline shift.^{109, 110} Other methods of data interpretation can include identifying compounds by their relative retention time and using the integrated areas of each peak for use in chemometric techniques.¹¹¹ In this study, which was in collaboration with the Indiana State Police laboratory, the feasibility of analyzing common plastics using pyrolysis-GC/MS was demonstrated. After a polymer library was established at their laboratory, a comparison was done using the plastic samples and certain polymer standards with chemometric analysis. Chemometric analysis of plastic samples using Py-GC/MS has not been documented and thus will assist in future research in this area.

4.2. Materials and Methods

4.2.1. Instrumental Analysis

The plastic samples were chosen by the Indiana State Police Microanalysis Unit so as to encompass the different types of common plastics that are currently readily available. The plastic recycling numbers (symbols) were used as a guideline to make sure a diverse sample collection was found. The numbers as well as the polymer they represent can be found in Table 4.1. The recycling numbers, also known as the Resin Identification Code, was created by the Society of the Plastics Industry in 1988. The code was brought

about to assist in separating similar-appearing plastics when they are recycled. The numbering in the Resin Code has no meaning besides identifying specific plastic. Using this code, at least one sample from each recycling number was chosen. The plastic “Other” is either a combination of the other plastics in the table or a different plastic altogether. In total, nine samples were collected in this research; one plastic from each category was used, and additional samples of HDPE and polystyrene were used with these additional samples being colored samples (a black HDPE sample and a red polystyrene sample). Several polymers were also analyzed and used in an external validation. These polymers can be found in Table 4.2.

Table 4.1 The Plastic Recycling Numbers and Polymer They Represent.

RECYCLING NUMBER	POLYMER	NUMBER OF PLASTIC SAMPLES USED
1	Poly(ethylene terephthalate) (PETE)	1
2	High-density Polyethylene (HDPE)	2
3	Poly(vinyl chloride) (V)	1
4	Low-density Polyethylene (LDPE)	1
5	Polypropylene (PP)	1
6	Polystyrene (PS)	2
7	Other	1

Table 4.2 Specific Polymers Analyzed.

Polyethylene (48% chlorinated)
Polyethylene (25% chlorinated)
Polyethylene (high density)
Polypropylene (chlorinated)
Polypropylene (isotactic)
Polystyrene
PVC (carboxylated 1.8%)
PVC (viscosity 1.26)

An Agilent® 6890 Gas Chromatograph/5975 Mass Spectrometer (Agilent Technologies, Santa Clara, CA) located in the Microanalysis Unit of the Indiana State Police Laboratory in Indianapolis, Indiana, was utilized in this research. The parameters for plastic research that is done at their laboratory were employed in this research. An injector temperature of 280°C and interface temperature of 280°C were used with a HP-5MS Agilent capillary column. The column was bonded 5% phenyl, 30 meters long, 0.25mm ID, and 0.25µm df. The initial temperature was 40°C and the initial time was 2.5 minutes. A program rate of 10°C/min was used with the final temperature reaching 300°C and a final time of 23 minutes. In the GC, a constant flow rate of 1mL/minute of helium was used with a split ratio of 20:1. A CDS 5100 Pyroprobe attachment (CDS Analytical, Oxford, PA) was used to pyrolyze the samples prior to introduction into the GC/MS. When the pyroprobe was triggered, the temperature progressed to 800°C for 20 seconds. The pyrolysis unit temperatures are 310°C for the valve

oven and the transfer line. The MS was equipped with a scan range of 35 – 350 amu and a scan speed of 4.45 scans/second. The source temperature was 230°C and the quad temperature was 150°C. There was no solvent delay, the threshold was 150, and the sampling rate was 2².

Small segments of the plastic were cut and inserted into the probe (approximately 1mm x 1mm). Polymer samples that were in powder form had to first be compacted into a solid disc using a press in the Indiana State Police laboratory. A small segment of this disc was then inserted into the probe. The sample was allowed to equilibrate for one minute before being pyrolyzed. The probe remained in the unit for all 50 minutes of the analysis. Three replicates were performed on each plastic sample to demonstrate reproducibility and be used for future chemometric analysis. Two replicates of each polymer sample were obtained but only one pyrogram from each sample was used for the external validation. In between each sample, a 30 minute blank was run in which the probe was first cleaned by pressing the “clean” button on the pyrolysis unit after each sample and then the empty probe was inserted into the unit and the blank was run.

4.2.2. Data Analysis

Overall, the sample collection consists of 27 pyrograms (9 total plastic samples). The eight polymer samples remained separate from this data set. The entire plastic dataset ranged from 0 – 50 minutes. Total Ion Chromatograms (TICs) of the raw data for each plastic can be seen in Figures 4.1 – 4.9, and

major peaks that were identified by searching a library of mass spectra (NIST Mass Spectral Database, Version 2.0) have been labeled.

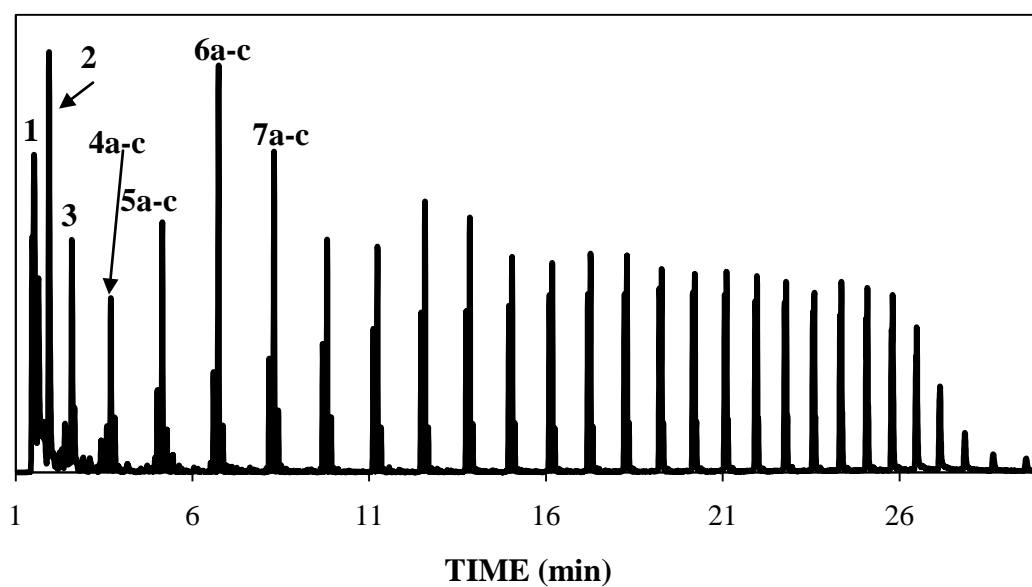


Figure 4.1 Pyrogram of HDPE; **(1)** 2-Butene, **(2)** 1-hexene, **(3)** 1-heptene, **(4)** *a*: 1,7-octadiene; *b*: 1-octene; *c*: octane, **(5)** *a*: 1,8-nonadiene; *b*: 1-nonene; *c*: nonane, **(6)** *a*: 1,9-decadiene; *b*: 1-decene; *c*: decane, **(7)** *a*: 1,10-undecadiene; *b*: 1-undecene; *c*: undecane.

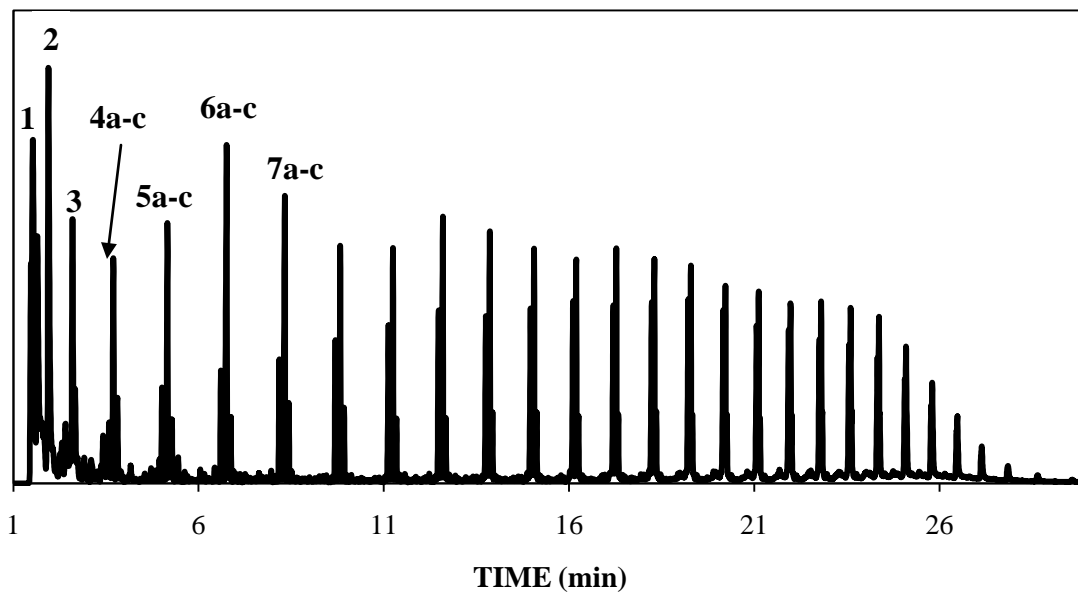


Figure 4.2 Pyrogram of LDPE; **(1)** 2-Butene, **(2)** 1-hexene, **(3)** 1-heptene, **(4)** *a*: 1,7-octadiene; *b*: 1-octene; *c*: octane, **(5)** *a*: 1,8-nonadiene; *b*: 1-nonene; *c*: nonane, **(6)** *a*: 1,9-decadiene; *b*: 1-decene; *c*: decane, **(7)** *a*: 1,10-undecadiene; *b*: 1-undecene; *c*: undecane.

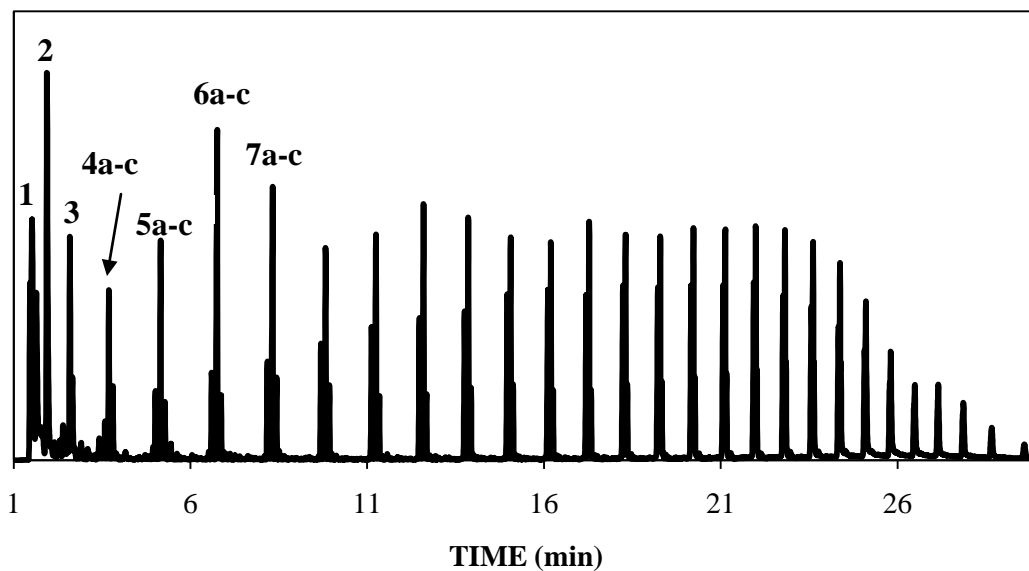


Figure 4.3 Pyrogram of HDPEblack; **(1)** 2-Butene, **(2)** 1-hexene, **(3)** 1-heptene, **(4)** *a*: 1,7-octadiene; *b*: 1-octene; *c*: octane, **(5)** *a*: 1,8-nonadiene; *b*: 1-nonene; *c*: nonane, **(6)** *a*: 1,9-decadiene; *b*: 1-decene; *c*: decane, **(7)** *a*: 1,10-undecadiene; *b*: 1-undecene; *c*: undecane.

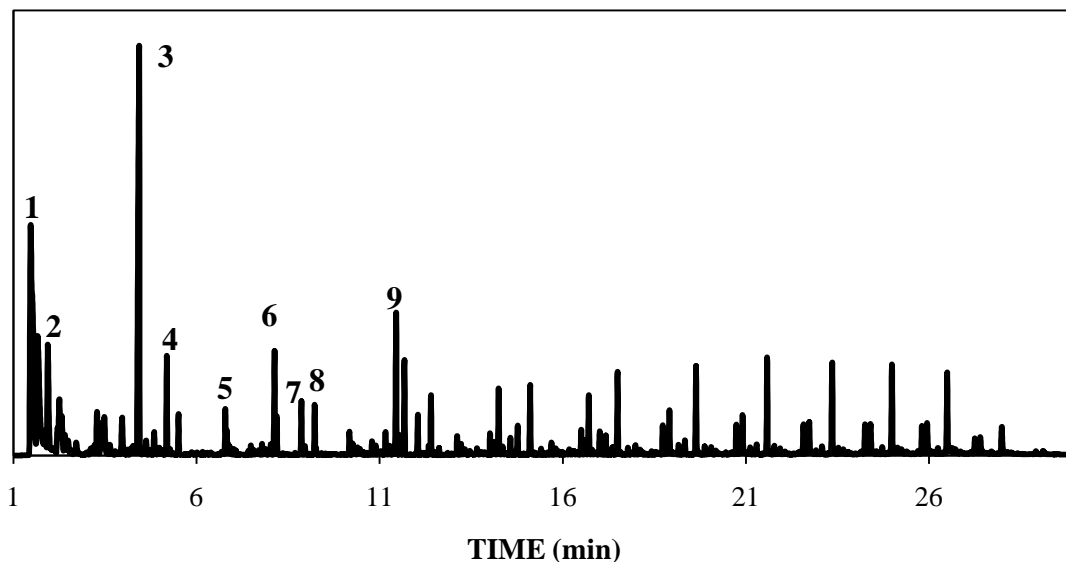


Figure 4.4 Pyrogram of Polypropylene; **(1)** cyclopropane, **(2)** 1-pentene,2-methyl, **(3)** 2,4-dimethyl-1-heptene, **(4)** 3-octene,2,2-dimethyl-, **(5)** nonane, 2-methyl-3-methylene-, **(6)** cyclooctane,1,4-dimethyl-,trans-, **(7)** 2-isopropyl-5-methyl-1-heptanol, **(8)** (2,4,6-trimethylcyclohexyl)methanol, **(9)** 9-Eicosene.

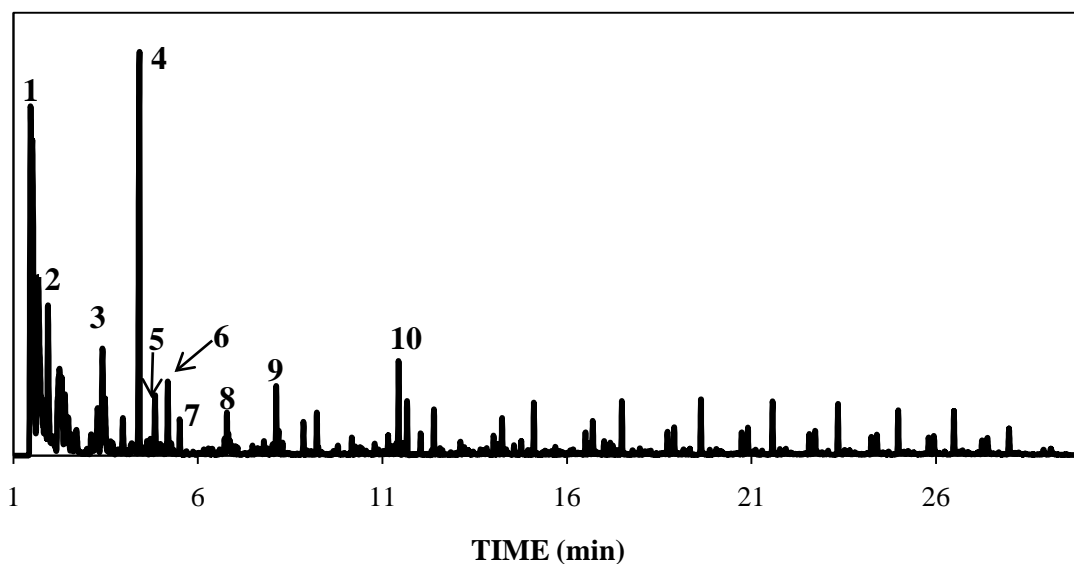


Figure 4.5 Pyrogram of PVC; **(1)** cyclopropane, **(2)** 1-pentene,2-methyl, **(3)** toluene, **(4)** 2,4-dimethyl-1-heptene, **(5)** o-xylene, **(6)** 3-octene, 2,2-dimethyl, **(7)** 1,6-octadiene,2,6-dimethyl-,(z)-, **(8)** nonane, **(9)** cyclooctane,1,4-dimethyl-,trans-, **(10)** 9-Eicosene.

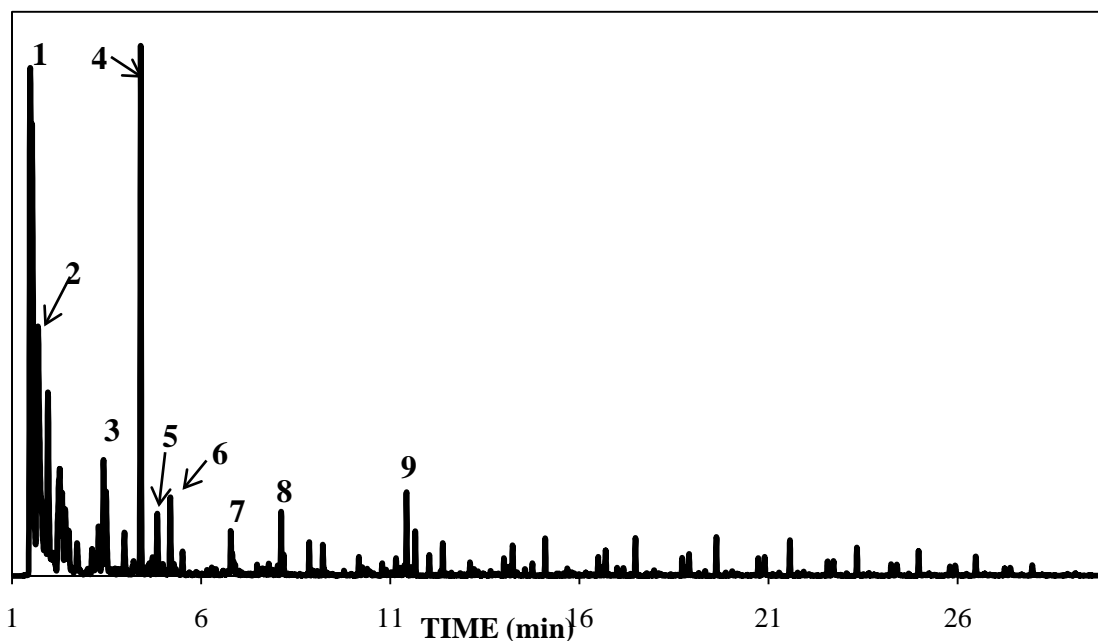


Figure 4.6 Pyrogram of Other; **(1)** cyclopropane, **(2)** 1-pentene,2-methyl, **(3)** toluene, **(4)** 2,4-dimethyl-1-heptene, **(5)** o-xylene, **(6)** 3-octene, 2,2-dimethyl, **(7)** nonane, **(8)** 2-undecanethiol,2-methyl, **(9)** 9-Eicosene.

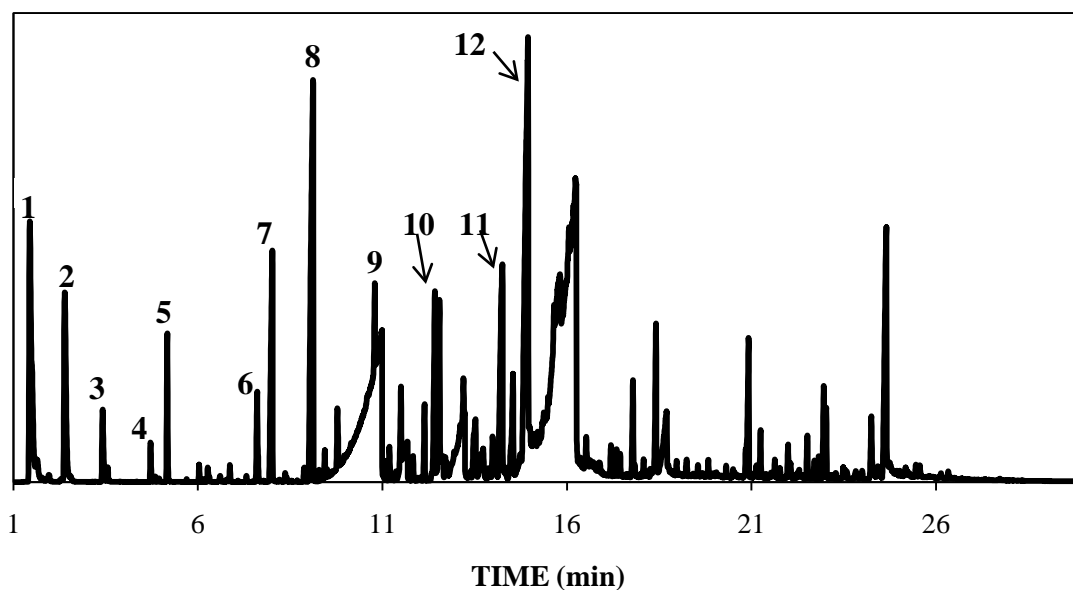


Figure 4.7 Pyrogram of PETE; **(1)** carbon dioxide, **(2)** benzene, **(3)** toluene, **(4)** ethylbenzene, **(5)** styrene, **(6)** benzeneacetaldehyde, **(7)** acetophenone, **(8)** vinyl benzoate, **(9)** benzoic acid, **(10)** biphenyl, **(11)** benzene, p-diacetyl-, **(12)** quinoline, 6-methoxy,1-oxide.

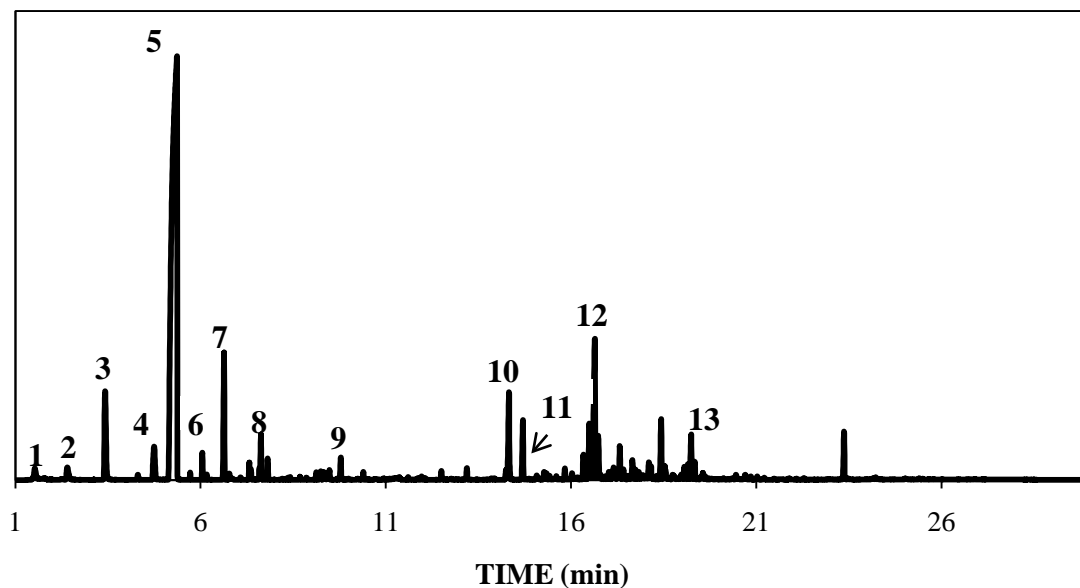


Figure 4.8 Pyrogram of Polystyrene; **(1)** methylenecyclopropane, **(2)** benzene, **(3)** toluene, **(4)** ethylbenzene, **(5)** styrene, **(6)** benzene,cyclopropyl-, **(7)** α -methylstyrene, **(8)** Indene, **(9)** naphthalene, **(10)** bibenzyl, **(11)** benzene,1,1'-(1-methyl-1,2-ethanediyl)bis-, **(12)** naphthalene,1,2,3,4-tetrahydro-2-phenyl, **(13)** 2-phenylnaphthalene.

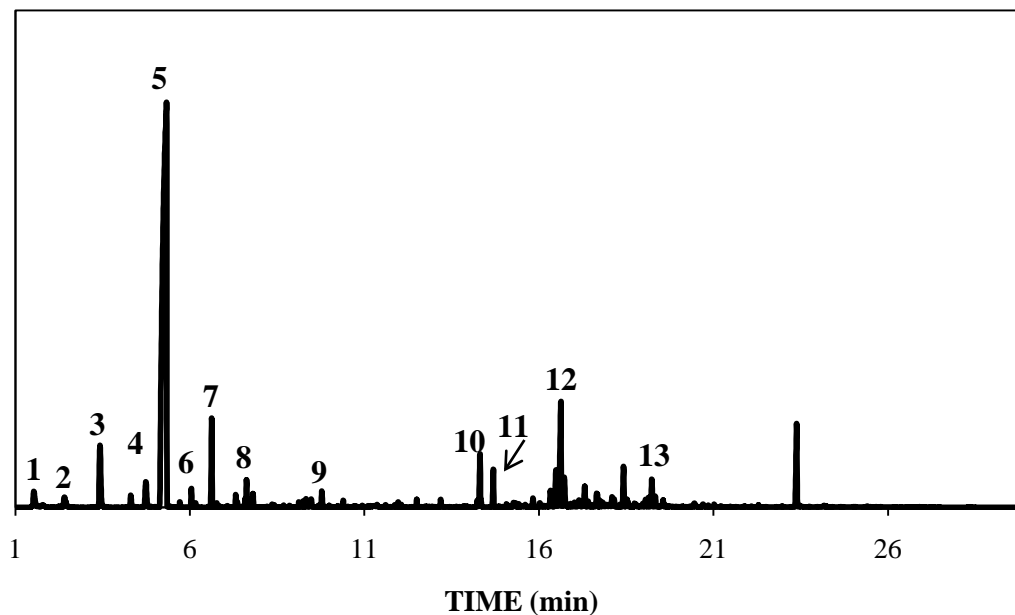


Figure 4.9 Pyrogram of PolystyreneRed; **(1)** methylenecyclopropane, **(2)** benzene, **(3)** toluene, **(4)** ethylbenzene, **(5)** styrene, **(6)** benzene, cyclopropyl-, **(7)** α -methylstyrene, **(8)** Indene, **(9)** naphthalene, **(10)** bibenzyl, **(11)** benzene, 1,1'-(1-methyl-1,2-ethanediyl)bis-, **(12)** naphthalene, 1,2,3,4-tetrahydro-2-phenyl, **(13)** 2-phenylnaphthalene.

A macro created by Infometrix, Inc. was used to convert the total ion chromatograms (TICs) into ASCII files and this involved using only the time frame of 1 – 30 minutes. Following this, the ASCII files were converted into Excel files through the computer program Piroutte (Infometrix, Inc., Bothell, WA). The Excel files were then pre-treated in Microsoft Excel by background correcting (subtracting the minimum value of the data file from each data point) and normalizing each data point by dividing its intensity by the square root of the sum of the squares of all abundance values.^{4, 5} This eliminated variability in the data due to sample amount. Next, the data was multiplied by 1,000,000 to guarantee that the y-axis values were large enough to interpret. Target alignment followed, which is an important step because the pyrograms need to be aligned properly in

order for chemometric analysis to be accurate. The target selection was performed by creating a new pyrogram which is the average of all pyrograms to be aligned. The Euclidean distance was then measured using Microsoft Excel and XLSTAT2009. The closest pyrogram to the averaged pyrogram was chosen as the target.

Following the alignment target selection, the pyrograms were converted to .cdf files using Piroutte, and these files were then used in the software program LineUp (Infometrix, Inc., Bothell, WA). LineUp is a chromatographic line-up tool that is used to correct for differences due to minor changes and shifts in chromatograms. The program uses the COW algorithm in which the chromatograms are aligned by piecewise linear stretching and compression of the time axis.¹¹⁰ Within the program, the slack parameter (warp) was set to 4 and the segment size was set to 20. Once the data was aligned, the files were converted back into .xls files and XLSTAT2009 was used to perform the chemometric analysis. Figures 4.10 – 4.12 illustrate these processes, showing raw data, pre-treated data, and aligned data for one specific plastic example (LDPE).

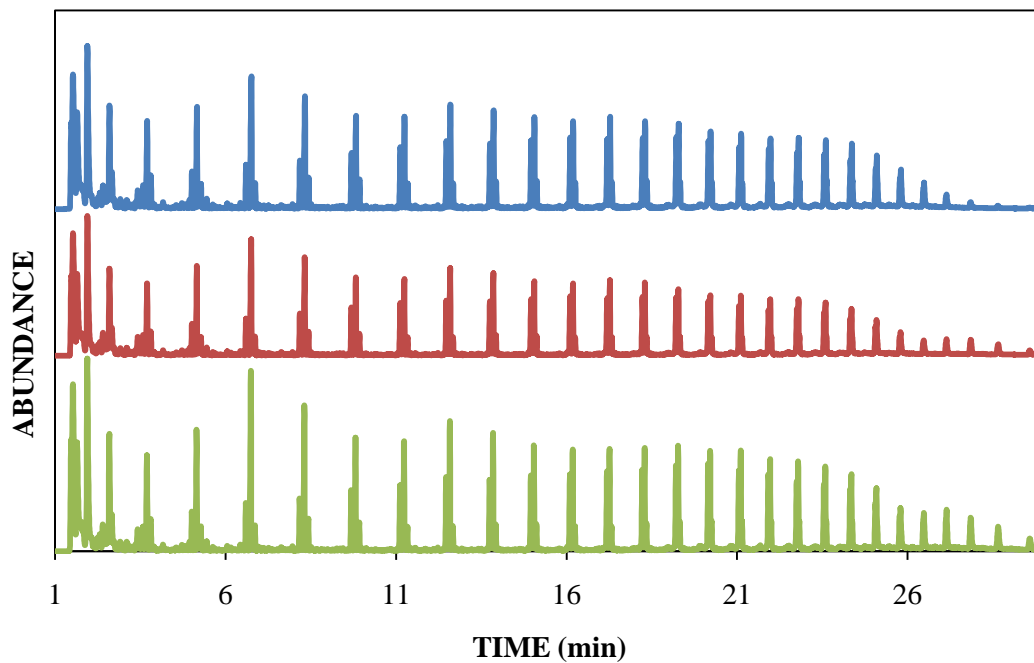


Figure 4.10 Raw data of three replicates of LDPE.

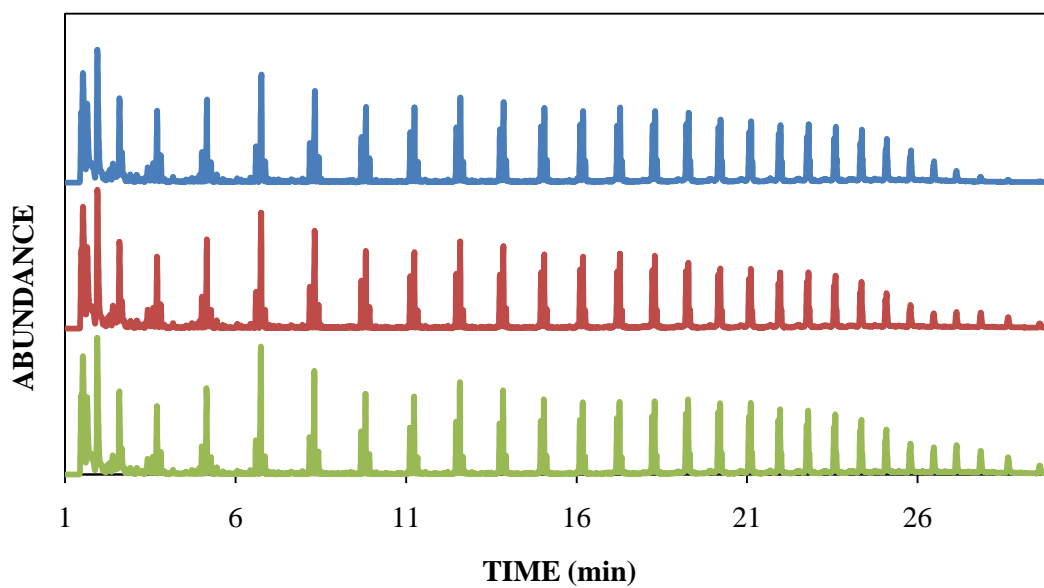


Figure 4.11 Background corrected and normalized data for three replicates of LDPE.

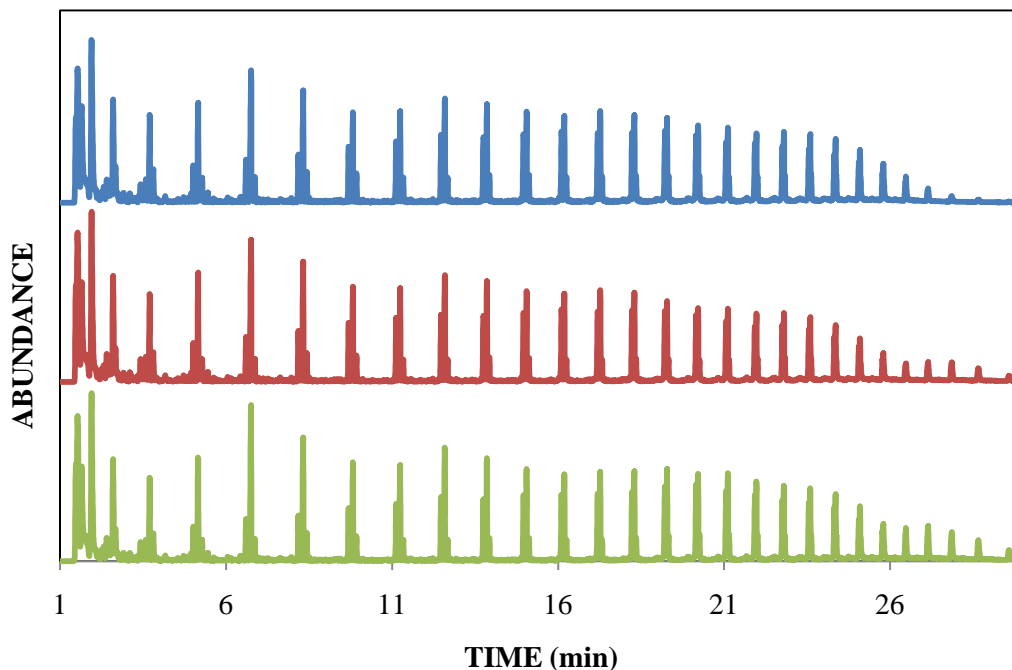


Figure 4.12 Aligned data for three replicates of LDPE.

Before chemometrics was applied, the dataset had to be reduced in order for the programs to analyze them because the dataset exceeded the limits on the number of data points that XLSTAT can handle. This was done by averaging every three time point scans together. Since there originally was 0.225 seconds between every scan, the reduction in the number of time points did not reduce the chromatographic resolution. This can be seen in Figures 4.13 and 4.14, which are zoomed in TICs of HDPE, with Figure 4.13 showing all time points and Figure 4.14 illustrating the averaged time points.

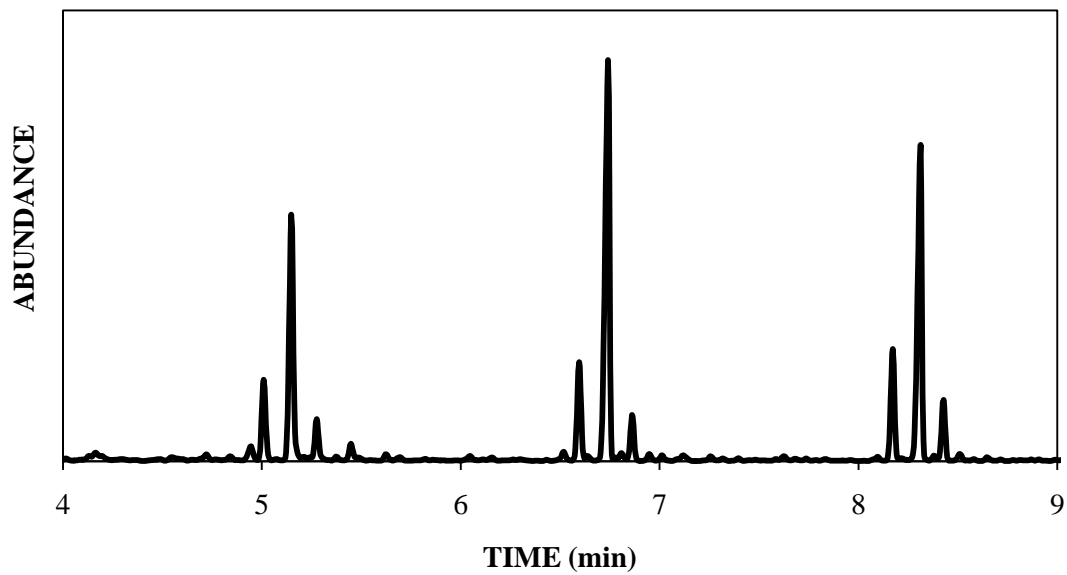


Figure 4.13 Portion of a TIC of HDPE showing all time points.

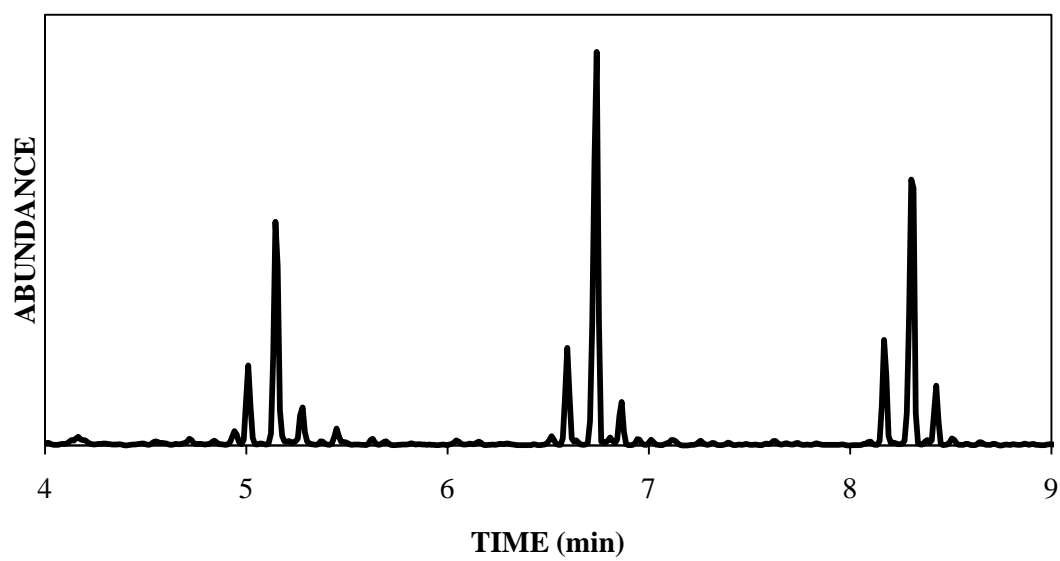


Figure 4.14 Portion of a TIC of HDPE showing averaged time points.

4.3. Results and Discussion

4.3.1. Statistical Results

The AHC dendrogram for plastics is shown in Figure 4.15, where four distinct clusters are present based on the position of the truncation line, which was determined by a histogram of node positions. Bifurcations to the right of the truncation line are considered to be meaningful in determining the number of classes. By examination of the graphs of the central objects of each cluster (Figures 4.16 – 4.19), it can be seen that each group has a very distinct chromatogram, and this is consistent with the node positions in the dendrogram showing that the four classes bifurcate at a relatively large distance as compared to other similar samples.

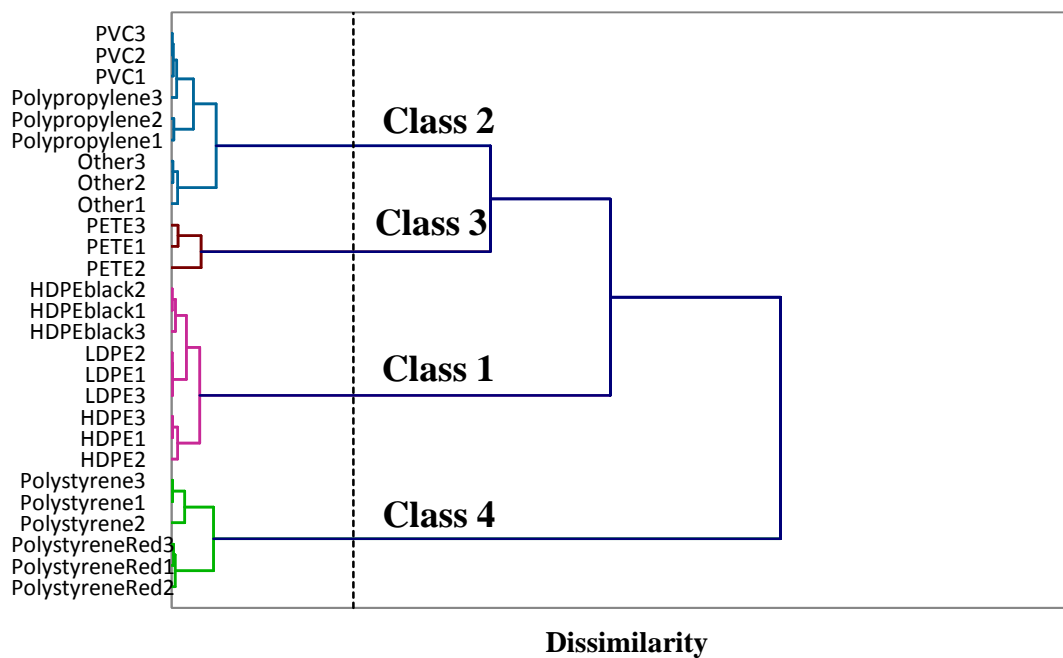


Figure 4.15 Dendrogram Produced Using AHC.

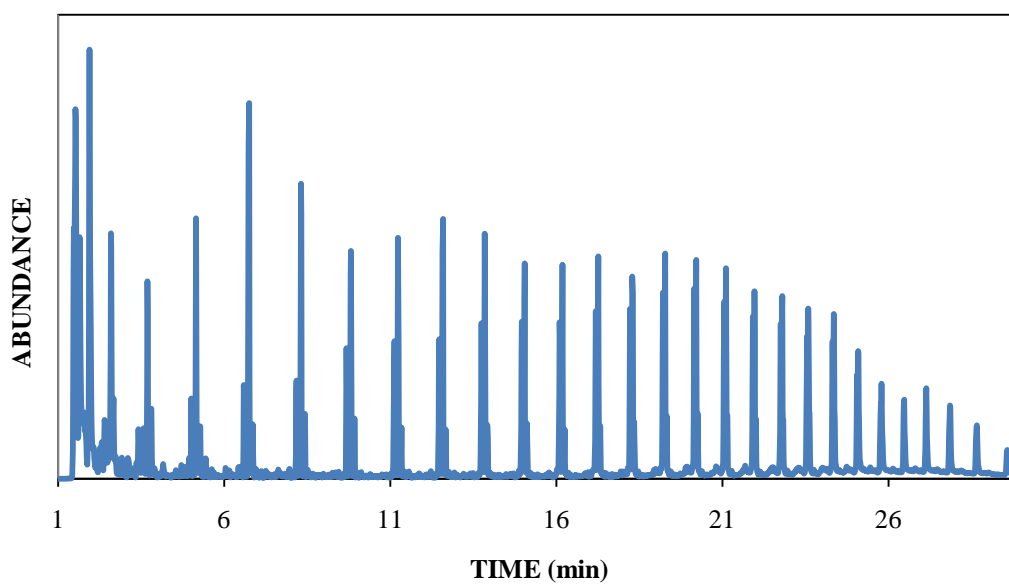


Figure 4.16 Central Object of Class 1 (LDPE).

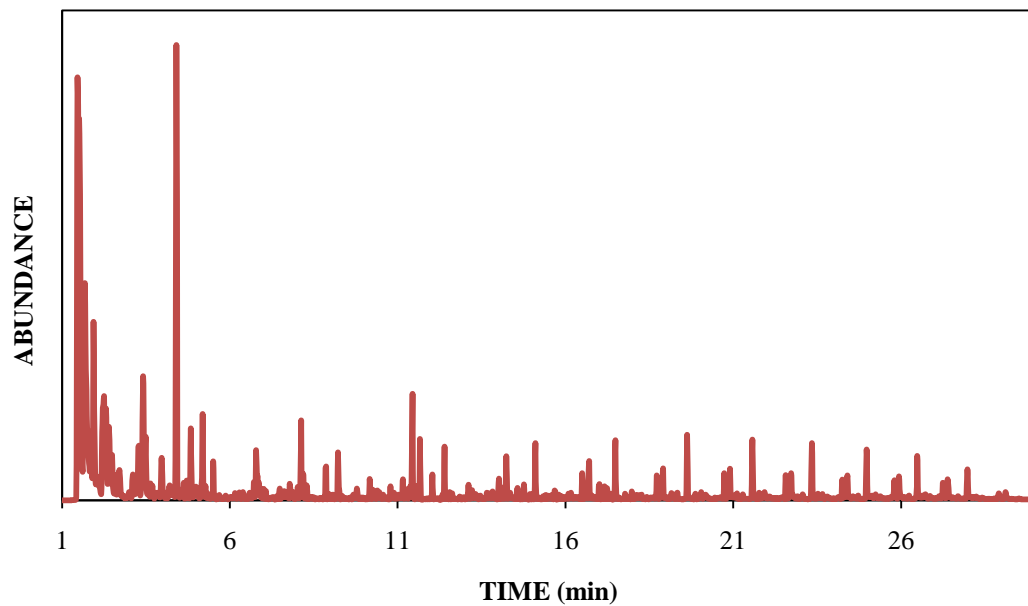


Figure 4.17 Central object of Class 2 (PVC).

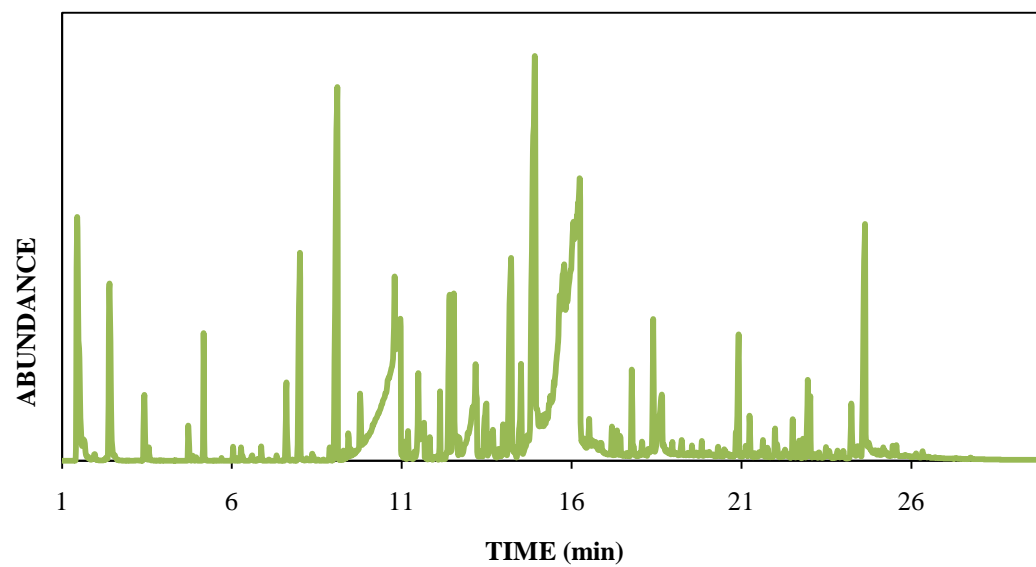


Figure 4.18 Central object of Class 3 (PETE).

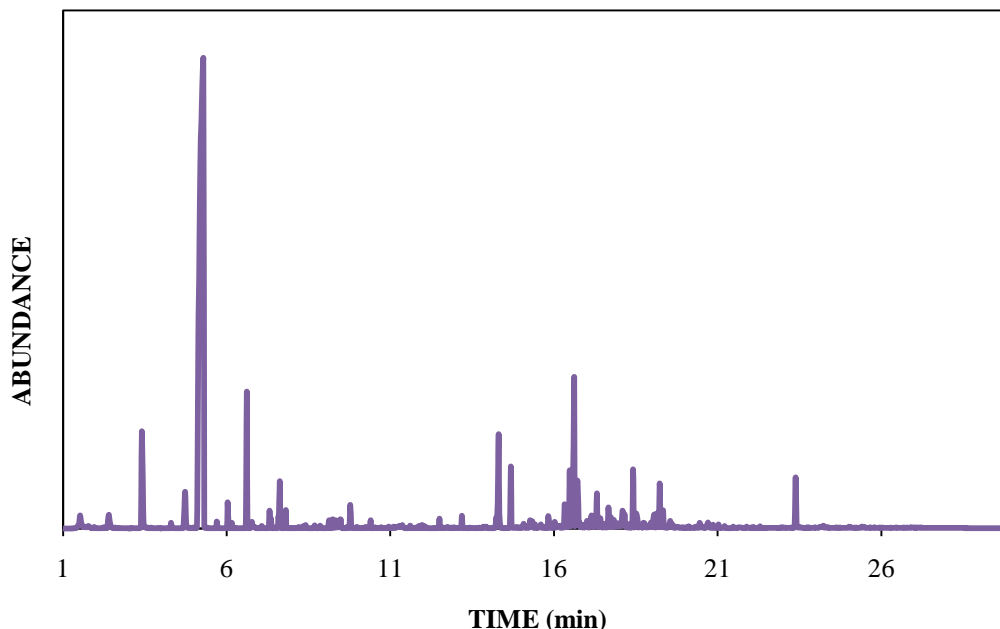


Figure 4.19 Central object of Class 4 (Polystyrene).

The observations plot produced by PCA is illustrated in Figure 4.20. The first two principal components captured 59.05% of the total variance of the population, with PC1 capturing 30.63% of the total variance and PC2 describing 28.42% of the total variance. Since this is a low amount of variance being described by the first two PCs, much of the information is present in higher PCs that these graphs are not illustrating. PCA condensed the components, and only the first two were used which represented the most variance in the population, and hence is able to be viewed in two dimensions. However, the total variance is still low. Four clusters with some overlap are present in Figure 4.21, which is the same PCA observations plot as Figure 4.20 except it is color-coded to illustrate the class formation. In addition, if the third PC was positive, a filled symbol was used; if the third PC was negative, an unfilled symbol was used. This was done

in order to illustrate a three-dimensional effect. Classes 1 and 2 appeared to be closely located, but using the third PC, they can be set apart.

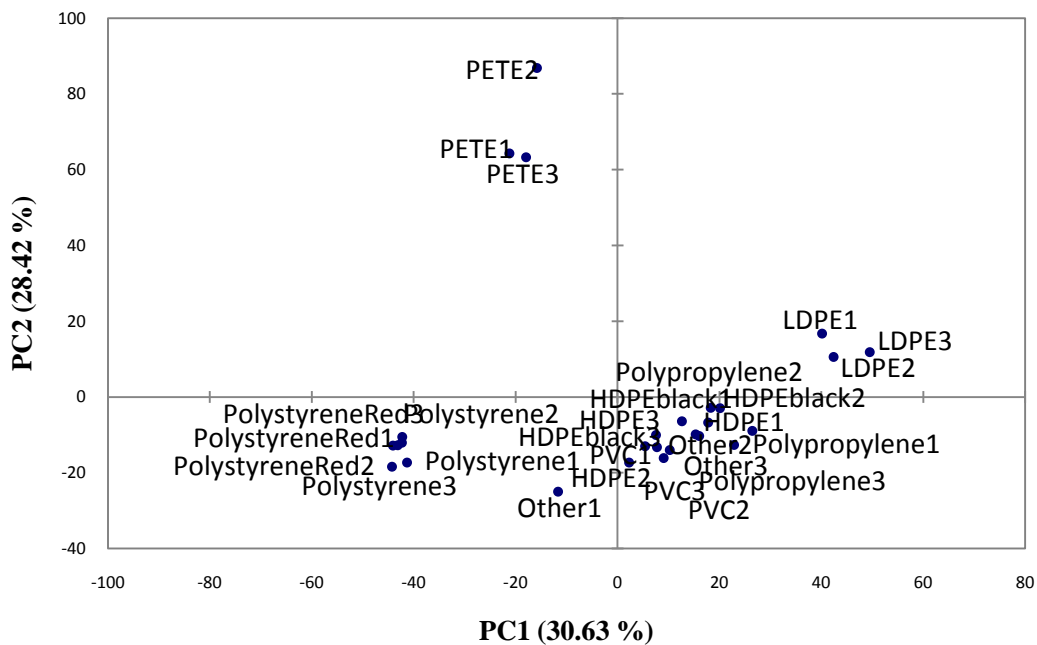


Figure 4.20 Observations Plot from PCA of Plastics Expressed in Terms of the First Two PCs.

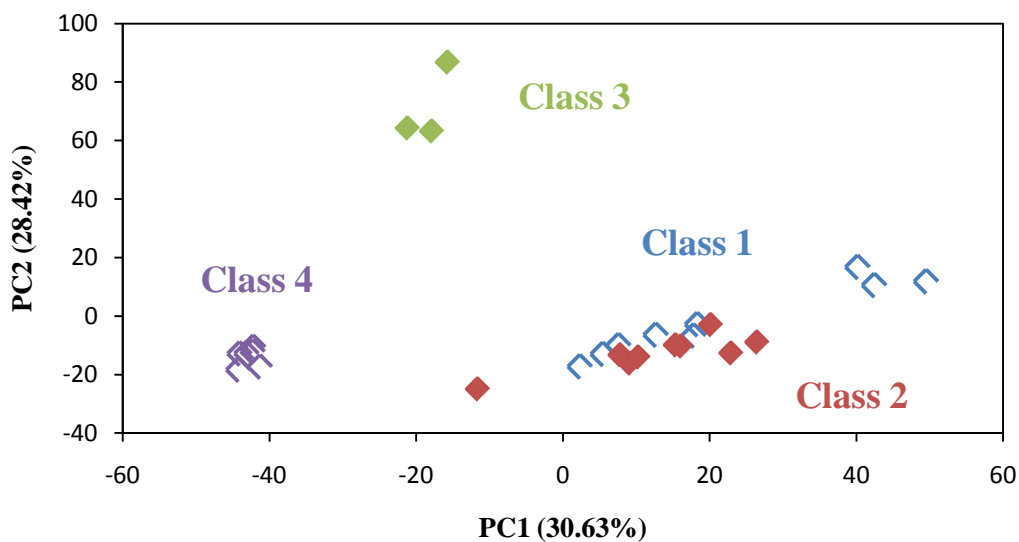


Figure 4.21 PCA observations plot color-coded to show the four classes from AHC clustering.

DA was then performed using the data from PCA; however, the Other plastic was removed. Every PC has an associated eigenvalue with it that describes the variance, as well as the percent variance and a cumulative variance for all the principal components. Several methods exist that assist in determining the significant number of principal components to use. One method involves setting a targeted cumulative percentage of variance and using the number of principal components associated with that number. This method was used in this research, and 95% was the selected cumulative variance. Using this variance, the first nine principal components would be selected, accounting for 95.83% of the total variance. Another method involves using a scree plot, which displays eigenvalues against factor number. A sudden drop or break in the plot is related to the number of significant principal components because after this drop, the plot remains consistent. Based on the scree plot, four principal components would be used. Lastly, the Kaiser criterion can be used to determine the significant amount of principal components to utilize. Using this criterion, all eigenvalues greater than one would be considered significant, and this would apply to the first 24 principal components.

The results of DA using the first nine principal components are shown in Figure 4.22, with 95.83% of the variance accounted for in two dimensions. The plastic Other was removed from DA and external validation because there is no polymer to match this plastic and the training set should consist of knowns. Four distinct groupings were used, and the assigned group membership was based upon the four AHC classes. No overlap between the clusters exists, and the

cross-validation results reflect this, as seen in the confusion matrix results in Table 4.3, which shows the classification of each sample into the four groups. Samples located along the diagonal represent those that were correctly classified. Overall, 100% of the samples were correctly classified.

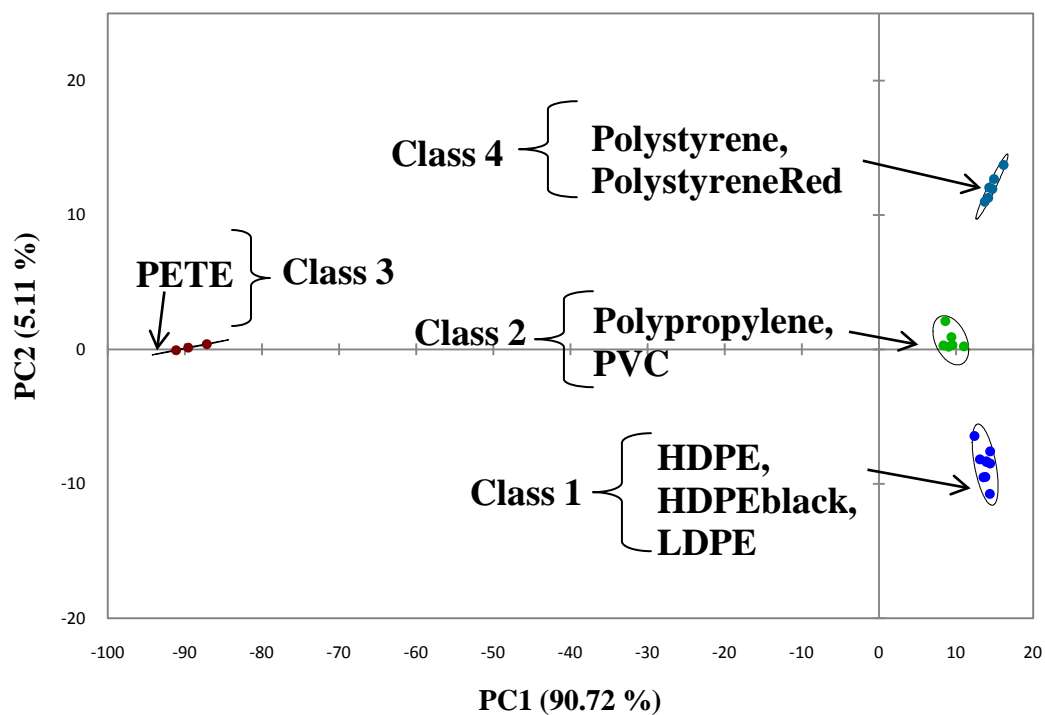


Figure 4.22 Observations Plot from DA.

Table 4.3 Confusion Matrix for the Cross-Validation Results from DA.

<i>From/To</i>	<i>Class 1</i>	<i>Class 2</i>	<i>Class 3</i>	<i>Class 4</i>	<i>TOTAL</i>	<i>% Correct</i>
<i>Class 1</i>	9	0	0	0	9	100.00
<i>Class 2</i>	0	6	0	0	9	100.00
<i>Class 3</i>	0	0	3	0	3	100.00
<i>Class 4</i>	0	0	0	6	6	100.00
<i>TOTAL</i>	9	6	3	6	24	100.00

DA was also performed using nine classes (instead of four), in which each type of plastic was classified into its own class. The first nine principal components were again used, and the DA observations plot can be seen in Figure 4.23, in which 90.24% of the total variance is accounted for in two dimensions. Some overlap can be seen with these nine clusters, and this is evident when looking at the cross validation results in Table 4.4. Samples located along the diagonal represent those that were correctly classified, while red samples outside of this diagonal are incorrect classifications. Overall, 91.67% of the samples were correctly classified, which is considered to be an excellent result.

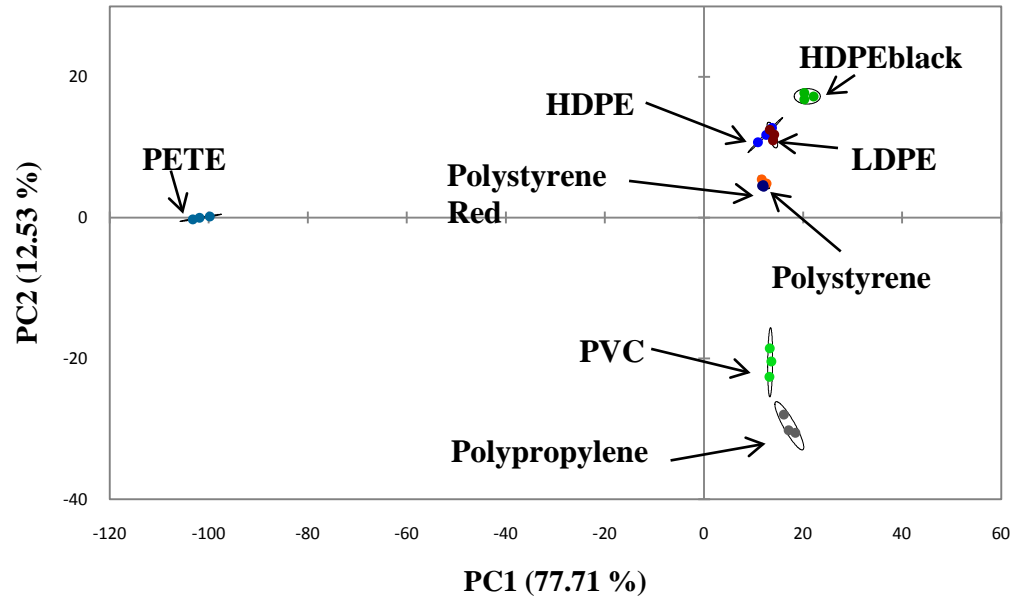


Figure 4.23 Observations plot from DA using nine classes, in which each plastic is its own class.

Table 4.4 Confusion Matrix for the Cross-Validation Results from DA Using Nine Classes.

<i>From/To</i>	<i>HDPE</i>	<i>HDPEblack</i>	<i>LDPE</i>	<i>PETE</i>	<i>PP</i>	<i>PS</i>	<i>PS Red</i>	<i>PVC</i>	<i>TOTAL</i>	<i>% Correct</i>
<i>HDPE</i>	3	0	0	0	0	0	0	0	3	100
<i>HDPEblack</i>	0	3	0	0	0	0	0	0	3	100
<i>LDPE</i>	0	0	3	0	0	0	0	0	3	100
<i>PETE</i>	0	0	0	3	0	0	0	0	3	100
<i>PP</i>	0	0	0	0	3	0	0	0	3	100
<i>PS</i>	0	0	0	0	0	2	1	0	3	66.67
<i>PS Red</i>	0	0	0	0	0	1	2	0	3	66.67
<i>PVC</i>	0	0	0	0	0	0	0	3	3	100
<i>TOTAL</i>	3	3	3	3	3	3	3	3	24	91.67

The misclassifications occurred between Polystyrene and PolystyreneRed. These misclassifications are reasonable because the only difference between these samples was the color; their chromatograms looked very similar and they were clustered originally into the same class. Also, when looking at the DA observations plot (Figure 4.23), those two classes are overlapping.

4.3.2. External Validation

The polymer samples were used as supplemental data in a form of external validation. The external validation results are shown in Table 4.5. DA predicted which class the supplemental samples should be placed. Correct placement was determined by seeing if the polymer was placed into the same class as the plastic containing that polymer (i.e if the PVC polymer was placed into the same class as the plastic PVC sample). The four classes from AHC were used and overall, the performance of the classification model was good, with 77.78% of samples correctly assigned. In this external validation, the samples that were misclassified were the two chlorinated polyethylene samples. Both of these were placed into the class that contained PVC. However, a chlorinated polyethylene sample would have a very similar structure to PVC and therefore under a similar pyrolysis breakdown (side group scission), so this placement of the polymer in the external validation is correct. PETE did not have a polymer standard to use in the external validation so Class 3 was not represented.

Table 4.5 Confusion Matrix for the External Validation Results of the Supplemental Data from DA.

<i>From/To</i>	<i>Class 1</i>	<i>Class 2</i>	<i>Class 4</i>	<i>TOTAL</i>	<i>% Correct</i>
<i>Class 1</i>	1	2	0	3	33.33
<i>Class 2</i>	0	4	0	4	100.00
<i>Class 4</i>	0	0	1	1	100.00
<i>TOTAL</i>	1	6	1	8	77.78

4.4. Conclusions

Based on the results discussed above, Py-GC/MS is an appropriate and efficient step in the analysis of plastics and polymers due to qualities such as its convenience and lack of physical pretreatment to the sample preparation. Pyrolysis allowed for the analysis of components that are otherwise unable to be analyzed by just GC/MS alone. It degraded molecules into smaller, more volatile compounds that could be analyzed using the GC/MS.^{94, 99}

Chemometrics was used as a guide in interpreting these pyrolysis results, and overall, the comparison of plastic and polymer samples using AHC, PCA, and DA resulted in several findings. It distinguished between a collection of plastics, as seen by the four distinct groups that formed in which also each plastic replicate remained clustered with its other replicates. In addition, color samples were placed into the same overall grouping as their non-colored

samples when using four classes but they were in separate groups within this large cluster. Therefore, even though the chromatograms of a colored versus non-colored sample of the same type of plastic appeared visually similar, the color must have some affect on the overall pyrolysis outcome.

CHAPTER 5. FUTURE DIRECTIONS

The work in this research has been a basis for future research and determining whether these applications are worth pursuing for potential implementation in routine analysis. Specifically, will chemometrics be able to assist forensic scientists in quantitatively analyzing various types of class evidence, and is MSP an appropriate technique for analyzing clear coats and fiber dyes. In order to determine this, other instrumentation should be used on these samples to see if the results are similar to the conclusions drawn in this research. Lastly, chemometrics is shown to cluster samples into groupings and if these groupings remain consistent using various instrumentation, then a plan to train scientists in chemometrics would be desirable.

When focusing on clear coats, several directions could be taken in the future. First, different sampling techniques could be tried to ensure that the chosen method of paint peels is the optimum method. Lensbond was shown to be a detriment to data analysis and it would be useful to determine if other mounting media would do the same. In addition, a larger sample population should be collected to try to encompass a variety of cars of different makes, models, years, and foreign or domestic to further prove that make, model, and year are not affecting class formation, as evident in this research. Specifically, the remaining samples in IUPUI's clear coat collection should be analyzed, and

samples of the same make, model, and year would be beneficial for further comparisons. Even if statistical results may not vary, it is ideal for comparison purposes to have a complete collection. More instrumental techniques should also be employed, such as FTIR and Raman spectroscopy, to see if the clusters remain the same.

Lastly, it is imperative to determine why the spectra are separating into the specific clusters and what factors are causing the formation of groups. Potential factors could include the UV absorber and/or binder system⁴¹ or the migration/weathering of the UV-absorbers over time.¹¹² When dealing with the UV absorber and binder system, a possible procedure could be to contact the manufacturer of the clear coat samples to determine information about the exemplars such as the UV absorbers or light stabilizers that were used. Once this information is known, it would be important to obtain the standards of these absorbers and stabilizers and then analyze them using MSP. This information may be difficult to come by; however, it is possible to obtain, as seen in Stoecklein's research involving clear coats.⁴¹ More likely, the development of analytical methods to correlate the absorbers present and the class the samples were placed in should occur. As previously mentioned, utilizing instrumentation such as pyrolysis GC/MS or extract the components of the clear coat using LC/MS would be imperative in assigning chemical meaning to the clear coats.

With fiber dyes, the first issue to consider is trying different preprocessing techniques, especially to correct the sloping baseline. A possible option is to fit the sloping baseline to a straight line in order to do background subtraction

instead of subtracting the minimum. By correcting the baseline, this will ensure that its features are not affecting the formation of clusters. The first derivative could also be used to see if that would remove any excess noise in the data. Further, when performing chemometrics, it would be beneficial to view the 3-D representation of the PCA observations plot. In these data sets, the two-dimensional plots often had a low amount of the variance accounted for and therefore, a 3-D plot might be able to illustrate more variation and show distinctions between classes that appear to overlap in two-dimensions.

More information about the dyes would be beneficial, such as finding all structures for the dyes. This could assist in determining the formation of clusters to see if they are based on structure or perhaps just visual color. Further, extracting the dyes from the fiber would provide more details about the dyes. TLC can be used in addition to MSP because MSP is generally held to be a complementary technique to TLC and running both techniques is a long-standing practice.^{78-80, 113, 114} TLC and MSP have been used to compare dye batches and this would be a beneficial aspect to explore in this research.^{83, 91} HPLC is another possible technique that could be used to analyze the dye after it has been extracted. Considerations to regard when using this technique are the chemical nature of the dye, possible extraction solvents, and possible degradation of the sample. The results of the dye extraction could then be analyzed with chemometrics and a comparison of the clusters with the MSP clusters could be performed to assess the validity of the techniques. Another technique that could be used for fiber analysis is Raman spectroscopy, which

has steadily been established as a technique in this area.^{70, 76} Analyzing the fiber dye by itself as well as on the fiber would be beneficial to note if any changes in the spectra are present.

Just as with clear coats, a larger sample collection of red dyes on cotton would be advantageous in ascertaining if the sample population is adequately represented. Dyes such as Direct Red 1 would be beneficial in increasing the sample size and thus making more meaningful conclusions. With this, the creation of a sufficient training set is necessary for classification to accurately occur. Additional dyes should be added to the collection as well as the same dyes but from different companies. This would determine if dyeing processes could affect class formation. Class formation could also be affected by environmental factors such as light. Over time, dye loss from the fiber was experienced so determining the cause of this would be beneficial for future research. Possible causes of this could be light, air exposure, or the dye being in glycerin (the mounting medium). An experiment can be designed to test these possible causes once more samples are collected. First, samples can be prepared in glycerin and microscopically examined at various time intervals (days) to determine which fibers experience loss of the dye and how long it takes for this to happen. Next, different mounting mediums can be tested in a similar fashion to see if they act similarly to the glycerin. In addition, storage of the samples can be experimented with by preparing slides and leaving the slide holders in the open room or in a dark drawer to see if light quickens the dye leeching process. If the glycerin is discovered to be the reason why the dye is

diffusing from the fiber, experiments can be designed to test if a dye can be extracted by soaking it in glycerin.

SWGMAAT fiber guidelines state that occasionally, an aromatic solvent reduced mounting medium such as XAM can adversely affect some fiber dyes by dissolving them and allowing them to diffuse from the fiber. However, this normally happens very quickly after mounting. The guidelines also state that a mounting medium, particularly solvent-based media, can have a detrimental effect on the fiber if mounted for a long time.⁸⁷ Based on these statements, testing glycerin's effect on the red dyes on cotton fibers would be beneficial, even if glycerin is an approved mounting medium for fibers. Specifically, fiber dyes Reactive Red 2 (Fiber E), Reactive Red 228 (Fiber F), Direct Red C-380 (Fiber A), Reactive Red 120 (Fiber B), and Direct Red 84 (Fiber 713) should be tested since those had noticeable dye loss during short and long-term storage of mounted samples. In short-term storage, dyes Reactive Red 198 (Fiber 686) and Reactive Red 239/241 (Fiber 695) had one of ten fibers experience the dissolving of the dye so those should be monitored closely as well. All other dyes did not experience dye loss.

Dye uptake clearly varied within certain fibers, as previously mentioned. When analyzing areas of different shades of the dye within a fiber, outliers occurred. Therefore, a way to control for the intra-fiber variability should be established. SWGMAAT guidelines suggest analyzing areas of a similar color but this leads to subjectivity due to the observer's eyesight and how they see the fiber colors.⁸⁷ A proposed method would be to follow the SWGMAAT guidelines of

analyzing areas of similar color, and the most prevalent color should be analyzed. External validations using other observers/researchers would be valuable in order to determine if the similar prevalent color is being selected. In addition, another intra-laboratory comparison should be performed but this time using the most frequent shade of the dye in the fiber (the medium shade). By doing this, the clusters may be more revealing.

The plastics research can be improved in several ways. First, an effort must be made to ensure that sample thickness is kept consistent. The solid polymers and plastics were shaved from the original sample and approximately the same size; however, variations most likely occurred. The powder samples were compacted together using a press and then a small section was removed from the larger piece. Again, the size was approximately the same for all samples, but variations occurred. An appropriate amount of sample is necessary to avoid peak broadening, and consistent sample sizes will assist in obtaining reproducibility. In real situations and cases, consistent sample size is needed for optimum comparative conclusions.

A much larger sample size will be necessary for future research and comparative studies. Specifically, polymer standards should be used that will match with the plastics being analyzed and therefore, an external validation study could be appropriately achieved. Further, a larger training set will be more useful in chemometric analyses to draw conclusions and make comparisons of groups. More replicates of the same sample would also assist in giving the model more data points since three or more is generally accepted. Besides a larger data set,

more similar types of plastics would be advantageous, as well as more samples that would be found at a crime scene. In addition, comparing methodologies of pyrolysis methods would be a valuable assessment. For example, the Pyroprobe method used in this research could be compared to a Shimadzu pyrolysis unit to see if the chemometric results are similar.

When performing chemometrics, a 3-D PCA observations plot would be beneficial since the amount of variability described in two dimensions was very low. Because of this, more variability could have been explained using more PCs, as was evident in Figure 4.21 when a color-coding scheme was used to distinguish clusters that appeared to overlap but were actually separated when using PC3.

Lastly, peak alignment was necessary to perform due to the shifts that occur in chromatograms. However, the peak alignment method does have disadvantages; namely, different parameters can be set during the process. No particular requirement or set standard is required to align the peaks and therefore, each operator is allowed to choose what they deem the best possible fit of the peaks. A consistent parameter was set for the slack and segment size but each chromatogram is different so the set parameter affected the chromatograms differently. Different parameters would have optimized the individual results for each chromatogram, and therefore, the percent match may have been higher. Choosing the parameters is also a difficult task. The segment size should be a small number to get better alignment, but too small of a number could result in over-fitting the data. In addition, a larger slack size provides more

flexibility and freedom to move features of the chromatogram, but it also allows strong deformation and could match features that the chromatogram did not originally resemble.¹¹⁰ Based on this, more time to investigate the various combinations of these parameters would result in a better understanding of the optimum alignment values.

LIST OF REFERENCES

LIST OF REFERENCES

1. Morgan, S. L.; Bartick, E. G., Discrimination of Forensic Analytical Chemical Data Using Multivariate Statistics. In *Forensic Analysis of the Cutting Edge: New Methods for Trace Evidence Analysis*, Blackledge, R. D., Ed. John Wiley & Sons: New York, 2007; pp 333-367.
2. *Strengthening Forensic Science in the United States - A Path Forward*. The National Academies Press: Washington, D.C., 2009; p 22; 24.
3. Johnson, R.; Wichern, D., *Applied Multivariate Statistical Analysis*. Prentice Hall, Inc: Englewood Cliffs, NJ, 1982.
4. Beebe, K.; Pell, R.; Seasholtz, M., *Chemometrics A Practical Guide*. John Wiley & Sons, Inc: New York, 1998.
5. Kramer, R., *Chemometric Techniques for Quantitative Analysis*. Marcel Dekker, Inc: New York, 1998.
6. Everitt, B., *Cluster Analysis*. Heinemann Educational Books Ltd: London, 1974.
7. Johnson, D., *Applied Multivariate Methods for Data Analysis*. Brooks/Cole Publishing Co.: Pacific Grove, CA, 1998.
8. Thanasoulis, N. C.; Piliouris, E. T.; Kotti, M. S. E.; Evmiridis, N. P., Application of multivariate chemometrics in forensic soil discrimination based on the UV-Vis spectrum of the acid fraction of humus. *Forensic Science International* **2002**, 130 (2-3), 73-82.
9. Chazottes, V.; Brocard, C.; Peyrot, B., Particle size analysis of soils under simulated scene of crime conditions: the interest of multivariate analyses. *Forensic Science International* **2004**, 140 (2-3), 159-166.
10. Koons, R. D.; Fiedler, C.; Rawalt, R. C., Classification and Discrimination of Sheet and Container Glasses by Inductively Coupled Plasma-Atomic Emission Spectrometry and Pattern Recognition. *Journal of Forensic Sciences* **1988**, 33 (1), 49-67.
11. Thanasoulis, N. C.; Parisi, N. A.; Evmiridis, N. P., Multivariate chemometrics for the forensic discrimination of blue ball-point pen inks based on their Vis spectra. *Forensic Science International* **2003**, 138 (1-3), 75-84.

12. John V. Goodpaster, A. B. S., Kristen L. Andrews, Leanora Brun-Conti, Identification and Comparison of Electrical Tapes Using Instrumental and Statistical Techniques: I. Microscopic Surface Texture and Elemental Composition. *Journal of Forensic Sciences* **2007**, 52 (3), 610-629.
13. Goodpaster, J. V.; Sturdevant, A. B.; Andrews, K. L.; Briley, E. M.; Brun-Conti, L., Identification and Comparison of Electrical Tapes Using Instrumental and Statistical Techniques: II. Organic Composition of the Tape Backing and Adhesive. *Journal of Forensic Sciences* **2009**, 54 (2), 328-338.
14. Flynn, K.; O'Leary, R.; Lennard, C.; Roux, C.; Reedy, B. J., Forensic applications of infrared chemical imaging: Multi-layered paint chips. *Journal of Forensic Sciences* **2005**, 50 (4), 832-841.
15. Chalmers, J. M.; Everall, N. J., FTIR, FT-Raman and chemometrics: applications to the analysis and characterisation of polymers. *Trends in Analytical Chemistry* **1996**, 15 (1), 16-25.
16. Lukasiak, B. M.; Faria, R.; Zomer, S.; Brereton, R. G.; Duncan, J. C., Pattern recognition for the analysis of polymeric materials. *Analyst* **2006**, 131, 73-80.
17. Egan, W. J.; Galipo, R. C.; Kochanowski, B. K.; Morgan, S. L.; Bartick, E. G.; Miller, M. L.; Ward, D. C.; Mothershead, R. F., Forensic discrimination of photocopy and printer toners. III. Multivariate statistics applied to scanning electron microscopy and pyrolysis gas chromatography/mass spectrometry. *Anal. Bioanal. Chem.* **2003**, 376 (8), 1286-1297.
18. May, C. D.; Watling, R. J., A comparison of the use of refractive index (RI) and laser ablation inductively coupled plasma mass spectrometry (LA-ICP-MS) for the provenance establishment of glass bottles. *Forensic Sci. Med. Pathol.* **2009**, 5 (2), 66-76.
19. Bajic, S. J.; Aeschliman, D. B.; Saetveit, N. J.; Baldwin, D. P.; Houk, R. S., Analysis of Glass Fragments by Laser-Ablation Inductively Coupled Plasma-Mass Spectrometry and Principal Component Analysis. *Journal of Forensic Sciences* **2005**, 50 (5), 1123-1127.
20. Aeschliman, D. B.; Bajic, S. J.; Baldwin, D. P.; Houk, R. S., Multivariate pattern matching of trace elements in solids by laser ablation inductively coupled plasma-mass spectrometry: Source attribution and preliminary diagnosis of fractionation. *Anal. Chem.* **2004**, 76 (11), 3119-3125.
21. Wilcken, H.; Schulten, H., Quality control of paints: Pyrolysis-mass spectrometry and chemometrics. *Analytica Chimica Acta* **1996**, 336 (1-3), 201-208.

22. Kochanowski, B.; Morgan, S., Forensic discrimination of automotive paint samples using pyrolysis-gas chromatography-mass spectrometry with multivariate statistics. *J. Chromatogr. Sci.* **2000**, *38* (3), 100-108.
23. Elomaa, M.; Lochmuller, C. H.; Kudrjashova, M.; Kaljurand, M., Classification of polymeric materials by evolving factor analysis and principal component analysis of thermochromatographic data. *Thermochim. Acta* **2000**, *362* (1-2), 137-144.
24. Kher, A.; Mulholland, M.; Green, E.; Reedy, B., Forensic classification of ballpoint pen inks using high performance liquid chromatography and infrared spectroscopy with principal components analysis and linear discriminant analysis. *Vib. Spectrosc.* **2006**, *40* (2), 270-277.
25. Kher, A. A.; Green, E. V.; Mulholland, M. I., Evaluation of principal components analysis with high-performance liquid chromatography and photodiode array detection for the forensic differentiation of ballpoint pen inks. *Journal of Forensic Sciences* **2001**, *46* (4), 878-883.
26. Kokot, S.; Tuan, N. A.; Rintoul, L., Discrimination of reactive dyes on cotton fabric by Raman spectroscopy and chemometrics. *Appl. Spectrosc.* **1997**, *51* (3), 387-395.
27. Kokot, S.; Crawford, K.; Rintoul, L.; Meyer, U., A DRIFTS study of reactive dye states on cotton fabric. *Vib. Spectrosc.* **1997**, *15* (1), 103-111.
28. Gilbert, C.; Kokot, S., Discrimination Of Cellulosic Fabrics By Diffuse-Reflectance Infrared Fourier-Transform Spectroscopy and Chemometrics. *Vib. Spectrosc.* **1995**, *9* (2), 161-167.
29. Marshall, L. J.; McIlroy, J. W.; McGuffin, V. L.; Smith, R. W., Association and discrimination of diesel fuels using chemometric procedures. *Anal. Bioanal. Chem.* **2009**, *394* (8), 2049-2059.
30. Egan, W. J.; Morgan, S. L.; Bartick, E. G.; Merrill, R. A.; Taylor, H. J., Forensic discrimination of photocopy and printer toners. II. Discriminant analysis applied to infrared reflection-absorption spectroscopy. *Anal. Bioanal. Chem.* **2003**, *376* (8), 1279-1285.
31. Burger, F.; Dawson, M.; Roux, C.; Maynard, P.; Doble, P.; Kirkbride, P., Forensic analysis of condom and personal lubricants by capillary electrophoresis. *Talanta* **2005**, *67* (2), 368-376.
32. Kher, A.; Stewart, S.; Mulholland, M., Forensic classification of paper with infrared spectroscopy and principal components analysis. *J. Near Infrared Spectrosc.* **2005**, *13* (4), 225-229.

33. Gresham, G. L.; Groenewold, G. S.; Bauer, W. F.; Ingram, J. C., Secondary ion mass spectrometric characterization of nail polishes and paint surfaces. *Journal of Forensic Sciences* **2000**, *45* (2), 310-323.
34. Tan, B. J.; Hardy, J. K.; Snavely, R. E., Accelerant classification by gas chromatography/mass spectrometry and multivariate pattern recognition. *Analytica Chimica Acta* **2000**, *422* (1), 37-46.
35. Doble, P.; Sandercock, M.; Du Pasquier, E.; Petocz, P.; Roux, C.; Dawson, M., Classification. of premium and regular gasoline by gas chromatography/mass spectrometry, principal component analysis and artificial neural networks. *Forensic Science International* **2003**, *132* (1), 26-39.
36. Sandercock, P. M. L.; Du Pasquier, E., Chemical fingerprinting of unevaporated automotive gasoline samples. *Forensic Science International* **2003**, *134* (1), 1-10.
37. Enlow, E. M.; Kennedy, J. L.; Nieuwland, A. A.; Hendrix, J. E.; Morgan, S. L., Discrimination of nylon polymers using attenuated total reflection mid-infrared spectra and multivariate statistical techniques. *Appl. Spectrosc.* **2005**, *59* (8), 986-992.
38. Kawahara, F.; Santner, J.; Julian, E., Characterization of Heavy Residual Fuel Oils and Asphalts by Infrared Spectrophotometry Using Statistical Discriminant Function Analysis. *Anal. Chem.* **1974**, *46* (2), 266-273.
39. Hicks, T.; Sermier, F. M.; Goldmann, T.; Brunelle, A.; Champod, C.; Margot, P., The classification and discrimination of glass fragments using non destructive energy dispersive X-ray mu fluorescence. *Forensic Science International* **2003**, *137* (2-3), 107-118.
40. Bentley, J., Composition, Manufacture, and Use of Paint. In *Forensic Examination of Glass and Paint - Analysis and Interpretation*, Caddy, B., Ed. Taylor & Francis: London, 2001.
41. Stoecklein, W.; Fujiwara, H., The examination of UV-absorbers in 2-coat metallic and non-metallic automotive paints. *Science & Justice* **1999**, *39* (3), 188-195.
42. Thornton, J. I., Forensic Paint Examination. In *Forensic Science Handbook*, 2nd ed.; Saferstein, R., Ed. Prentice Hall: Upper Saddle River, New Jersey, 2002; Vol. I.
43. Oberg, P., Effectiveness of UV absorbers in selected automotive topcoats. *Polymeric Materials Science and Engineering* **2000**, *83* (129).

44. J Huybrechts, A. V., K Dusek, M Duskova-Smrckova, R Barsotti, Scratch- and mar-resistant refinish two-pack clear coats - linear versus branched acrylics. *Surface Coatings International Part B: Coatings Transactions* **2006**, 89 (B4), 275-282.
45. SWGMAT, Forensic Paint Analysis and Comparison Guidelines. *Forensic Science Communications* **1999**, 1 (2).
46. Henson, M.; Jergovich, T., Scanning electron microscopy and energy dispersive X-ray spectrometry (SEM/EDS) for the forensic examination of paints and coatings. In *Forensic Examination of Glass and Paint*, Caddy, B., Ed. Taylor & Francis: London, 2001; pp 243-272.
47. Zieba-Palus, J.; Borusiewicz, R., Examination of multilayer paint coats by the use of infrared, Raman, and XRF spectroscopy for forensic purposes. *Journal of Molecular Structure* **2006**, 792-793, 286-292.
48. Beveridge, A.; Fung, T.; MacDougall, D., Use of infrared spectroscopy for the characterisation of paint fragments. In *Forensic Examination of Glass and Paint*, Caddy, B., Ed. Taylor & Francis: London, 2001; pp 183-272.
49. Zieba-Palus, J., Application of micro-Fourier transform infrared spectroscopy to the examination of paint samples. *Journal of Molecular Structure* **1999**, 511-512, 327-335.
50. Ryland, S.; Bishea, G.; Brun-Conti, L.; Eyring, M.; Flanagan, B.; Jergovich, T.; MacDougall, D.; Suzuki, E., Discrimination of 1990s original automotive paint systems: A collaborative study of black nonmetallic base coat/clear coat finishes using infrared spectroscopy. *Journal of Forensic Sciences* **2001**, 46 (1), 31-45.
51. Edmondstone, G.; Hellman, J.; Legate, K.; Vardy, G.; Lindsay, E., An Assessment of the Evidential Value of Automotive Paint Comparisons. *Canadian Society of Forensic Science Journal* **2004**, 37 (3), 147-153.
52. Eyring, M.; Lovelace, M.; Sy, D. In *A Study of the Discrimination of Some Automotive Paint Films Having Identical Color Codes*, NIJ Trace Evidence Symposium, Clearwater Beach, FL, Clearwater Beach, FL, 2007.
53. Challinor, J. M., Pyrolysis techniques for the characterisation and discrimination of paint. In *Forensic Examination of Glass and Paint*, Caddy, B., Ed. Taylor & Francis: London, 2001; pp 165-182.
54. Wilcken, H.; Schulten, H. R., Differentiation of resin-modified paints by pyrolysis-gas chromatography mass spectrometry and principal component analysis. *Fresenius J. Anal. Chem.* **1996**, 355 (2), 157-163.

55. Wampler, T.; Bishea, G.; Simonsick, W., Recent changes in automotive paint formulation using pyrolysis gas chromatography mass spectrometry for identification. *J. Anal. Appl. Pyrolysis* **1997**, 40-1, 79-89.
56. Burns, D.; Doolan, K., A comparison of pyrolysis-gas chromatography-mass spectrometry and fourier transform infrared spectroscopy for the characterisation of automotive paint samples. *Analytica Chimica Acta* **2005**, 539 (1-2), 145-155.
57. Zieba-Palus, J.; Milczarek, J.; Koscielniak, P., Application of Infrared Spectroscopy and Pyrolysis-Gas Chromatography-Mass Spectrometry to the Analysis of Automobile Paint Samples. *Chemia analityczna* **2007**, 52 (109), 109-121.
58. Zieba-Palus, J.; Zadora, G.; Milczarek, J.; Koscielniak, P., Pyrolysis-gas chromatography/mass spectrometry analysis as a useful tool in forensic examination of automotive paint traces. *Journal of Chromatography A* **2008**, 1179, 41-46.
59. Zieba-Palus, J.; Zadora, G.; Milczarek, J., Differentiation and evaluation of evidence value of styrene acrylic urethane topcoat car paints analyzed by pyrolysis-gas chromatography. *Journal of Chromatography A* **2008**, 1179, 47-58.
60. Burns, D.; Doolan, K., The discrimination of auto-motive clear coat paints indistinguishable by Fourier transform infrared spectroscopy via pyrolysis-gas chromatography-mass spectrometry. *Analytica Chimica Acta* **2005**, 539 (1-2), 157-164.
61. Plage, B.; Berg, A.; Luhn, S., The discrimination of automotive clear coats by pyrolysis-gas chromatography/mass spectrometry and comparison of samples by a chromatogram library software. *Forensic Science International* **2008**, 117, 146-152.
62. Cousins, D.; Platoni, C.; Russell, L., The use of microspectrophotometry for the identification of pigments in small paint samples. *Forensic Science International* **1984**, 24 (3), 183-196.
63. Kopchick, K.; Bommarito, C., Color Analysis of Apparently Achromatic Automotive Paints by Visible Microspectrophotometry. *Journal of Forensic Sciences* **2006**, 51 (2), 340-343.
64. David, S.; Pailthorpe, M., Classification of Textile Fibers: Production, Structure, and Properties. In *Forensic Examination of Fibres*, Robertson, J.; Grieve, M., Eds. CRC Press: New York, 1999.
65. Bureau, F. E., *2008 World Directory of Manufactured Fiber Producers*. Arlington, VA, 2007.

66. Robertson, J.; Roux, C., Transfer, Persistence, and Recovery of Fibres. In *Forensic Examination of Fibres*, Robertson, J.; Grieve, M., Eds. CRC Press: New York, 1999.
67. Deedrick, D. W., Hairs, Fibers, Crime, and Evidence; Part 2: Fiber Evidence. *Forensic Science Communications* **2000**, 2 (3).
68. Wiggins, K., Thin Layer Chromatographic Analysis for Fibre Dyes. In *Forensic Examination of Fibres*, Robertson, J.; Grieve, M., Eds. CRC Press: New York, 1999.
69. Palenik, S., Microscopical Examination of Fibers. In *Forensic Examination of Fibres*, Robertson, J.; Grieve, M., Eds. CRC Press: New York, 1999; pp 153-177.
70. Kirkbride, K.; Tungol, M., Infrared Microspectroscopy of Fibres. In *Forensic Examination of Fibers*, Robertson, J.; Grieve, M., Eds. CRC Press: New York, 1999; pp 179-222.
71. Challinor, J., Fibre Identification by Pyrolysis Techniques. In *Forensic Examination of Fibres*, Robertson, J.; Grieve, M., Eds. CRC Press: New York, 1999; pp 223-237.
72. Roux, C., Scanning Electron Microscopy and Elemental Analysis. In *Forensic Examination of Fibres*, Robertson, J.; Grieve, M., Eds. CRC Press: New York, 1999.
73. Eyring, M., Visible Microscopical Spectrophotometry in the Forensic Sciences. In *Forensic Science Handbook*, Saferstein, R., Ed. Regents/Prentice Hall: Upper Saddle River, NJ, 1993; pp 321-387.
74. Adolf, F.; Dunlop, J., Microspectrophotometry/Colour Measurement. In *Forensic Examination of Fibres* Robertson, J.; Grieve, M., Eds. CRC Press: New York, 1999; pp 251-289.
75. Bartick, E.; Tungol, M., Infrared Microscopy and its Forensic Applications. In *Forensic Science Handbook*, Saferstein, R., Ed. Regents/Prentice Hall: Upper Saddle River, NJ, 1993; pp 196-252.
76. White, P., Surface Enhanced Resonance Raman Scattering Spectroscopy. In *Forensic Examination of Fibres*, Robertson, J.; Grieve, M., Eds. CRC Press: New York, 1999; pp 337-342.
77. Wiggins, K. G., The European Fibres Group (EFG) 1993-2002: "Understanding and improving the evidential value of fibres". *Anal. Bioanal. Chem.* **2003**, 376 (8), 1172-1177.

78. Golding, G. M.; Kokot, S., The Selection Of Noncorrelated Thin-Layer Chromatographic (TLC) Solvent Systems For The Comparison Of Dyes Extracted From Transferred Fibers. *Journal of Forensic Sciences* **1989**, *34* (5), 1156-1165.
79. Rendle, D. F.; Crabtree, S. R.; Wiggins, K. G.; Salter, M. T., Cellulase Digestion Of Cotton Dyed With Reactive Dyes And Analysis Of The Products By Thin-Layer Chromatography. *J. Soc. Dye. Colour* **1994**, *110* (11), 338-341.
80. Wiggins, K.; Crabtree, S.; March, B., The importance of thin layer chromatography in the analysis of reactive dyes released from wool fibers. *Journal of Forensic Sciences* **1996**, *41* (6), 1042-1045.
81. Macrae, R.; Dudley, R.; Smalldon, K., The Extraction of Dyestuffs from Single Wool Fibers *Journal of Forensic Sciences* **1979**, *24* (1), 109-116.
82. Resua, R., A Semi-Micro Technique for the Extraction and Comparison of Dyes in Textile Fibers. *Journal of Forensic Sciences* **1980**, *25* (1), 168-173.
83. Wiggins, K.; Cook, R.; Turner, Y., Dye Batch Variation In Textile Fibers. *Journal of Forensic Sciences* **1988**, *33* (4), 998-1007.
84. Beattie, I.; Dudley, R.; Smalldon, K., The Extraction and Classification of Dyes on Single Nylon, Polyacrylonitrile, and Polyester Fibers. *J. Soc. Dye. Colour* **1979**, *95*, 295-302.
85. Resua, R.; DeForest, P.; Harris, H., The Evaluation and Selection of Uncorrelated Paired Solvent Systems for Use in the Comparison of Textile Dyes by Thin-Layer Chromatography. *Journal of Forensic Sciences* **1981**, *26* (3), 515-534.
86. Griffin, R.; Speers, J., Other Methods of Colour Analysis - High-Performance Liquid Chromatography In *Forensic Examination of Fibres*, Robertson, J.; Grieve, M., Eds. CRC Press: New York, 1999.
87. SWGMAT, Forensic Fiber Examination Guidelines. *Forensic Science Communications* **1999**, *1* (1).
88. Grieve, M. C.; Dunlop, J.; Haddock, P., An Assessment Of The Value Of Blue, Red, And Black Cotton Fibers As Target Fibers In Forensic-Science Investigations. *Journal of Forensic Sciences* **1988**, *33* (6), 1332-1344.
89. Grieve, M. C.; Dunlop, J.; Haddock, P., An Investigation Of Known Blue, Red, And Black Dyes Used In The Coloration Of Cotton Fibers. *Journal of Forensic Sciences* **1990**, *35* (2), 301-315.

90. Wiggins, K. G.; Holness, J. A.; March, B. M., The importance of thin layer chromatography and UV microspectrophotometry in the analysis of reactive dyes released from wool and cotton fibers. *Journal of Forensic Sciences* **2005**, *50* (2), 364-368.
91. Wiggins, K.; Holness, J., A further study of dye batch variation in textile and carpet fibres. *Science & Justice* **2005**, *45* (2), 93-96.
92. Wheals, B.; Noble, W., Forensic applications of pyrolysis gas chromatography. *Chromatographia* **1972**, *5* (9), 553-557.
93. Serrano, D. P.; Aguado, J.; Escola, J. M.; Rodriguez, J. M.; San Miguel, G., An investigation into the catalytic cracking of LDPE using Py-GC/MS. *J. Anal. Appl. Pyrolysis* **2005**, *74* (1-2), 370-378.
94. Moldoveanu, S. C., Pyrolysis GC/MS, present and future - (Recent past and present needs). *J. Microcolumn Sep.* **2001**, *13* (3), 102-125.
95. Sobeih, K. L.; Baron, M.; Gonzalez-Rodriguez, J., Recent trends and developments in pyrolysis-gas chromatography. *Journal of Chromatography A* **2008**, *1186* (1-2), 51-66.
96. Wang, F. C. Y., Polymer analysis by pyrolysis gas chromatography. *Journal of Chromatography A* **1999**, *843* (1-2), 413-423.
97. Smith, P. B.; Pasztor, A. J.; McKelvy, M. L.; Meunier, D. M.; Froelicher, S. W.; Wang, F. C. Y., Analysis of synthetic polymers and rubbers. *Anal. Chem.* **1999**, *71* (12), 61R-80R.
98. Rial-Otero, R.; Galesio, M.; Capelo, J. L.; Simal-Gandara, J., A Review of Synthetic Polymer Characterization by Pyrolysis-GC-MS. *Chromatographia* **2009**, *70* (3-4), 339-348.
99. Wampler, T. P., Introduction to pyrolysis-capillary gas chromatography. *Journal of Chromatography A* **1999**, *842* (1-2), 207-220.
100. Nordmark, U., Quantitative determination of resin mixtures in coating colours for paper by multivariate analysis of data from pyrolysis-gas chromatography. *J. Anal. Appl. Pyrolysis* **2000**, *55* (1), 93-103.
101. Wampler, T. P., Practical applications of analytical pyrolysis. *J. Anal. Appl. Pyrolysis* **2004**, *71* (1), 1-12.
102. Armitage, S.; Saywell, S.; Roux, C.; Lennard, C.; Greenwood, P., The analysis of forensic samples using laser micro-pyrolysis gas chromatography mass spectrometry. *Journal of Forensic Sciences* **2001**, *46* (5), 1043-1052.

103. Sakayanagi, M.; Konda, Y.; Watanabe, K.; Harigaya, Y., Identification of Pressure-Sensitive Adhesive Polypropylene Tape. *Journal of Forensic Sciences* **2003**, *48* (1), 68-76.
104. Williams, E., R.; Munson, T. O., The Comparison of Black Polyvinylchloride (PVC) Tapes by Pyrolysis Gas Chromatography. *Journal of Forensic Sciences* **1988**, *33* (5), 1163-1170.
105. Sellers, K. W.; Towns, C. M.; Mubarak, C. R.; Kloppenburg, L.; Bunz, U. H. F.; Morgan, S. L., Characterization of high molecular weight poly(p-phenylenethynylene)s by pyrolysis gas chromatography/mass spectrometry with multivariate data analysis. *J. Anal. Appl. Pyrolysis* **2002**, *64* (2), 313-326.
106. Kristensen, R.; Coulson, S.; Gordon, A., THM PyGC-MS of wood fragment and vegetable fibre forensic samples. *J. Anal. Appl. Pyrolysis* **2009**, *86* (1), 90-98.
107. Campbell, G. P.; Gordon, A. L., Analysis of condom lubricants for forensic casework. *Journal of Forensic Sciences* **2007**, *52* (3), 630-642.
108. Sahota, R. S.; Morgan, S. L., Vector Representation, Feature-Selection, And Fingerprinting - An Application Of Pattern-Recognition To Pyrolysis-Gas Chromatography Mass-Spectrometry Of Nucleosides. *Anal. Chem.* **1993**, *65* (1), 70-77.
109. Cheung, W.; Xu, Y.; Thomas, C. L. P.; Goodacre, R., Discrimination of bacteria using pyrolysis-gas chromatography-differential mobility spectrometry (Py-GC-DMS) and chemometrics. *Analyst* **2009**, *134*, 557-563.
110. Nielsen, N.-P.; Carstensen, J.; Smedsgaard, J., Aligning of single and multiple wavelength chromatographic profiles for chemometric data analysis using correlation optimised warping. *Journal of Chromatography A* **1998**, *805* (1-2), 17-35.
111. Gray, N. R.; Lancaster, C. J.; Gethner, J., Chemometric Analysis Of Pyrolysate Compositions - A Model For Predicting The Organic-Matter Type Of Source Rocks Using Pyrolysis-Gas Chromatography. *J. Anal. Appl. Pyrolysis* **1991**, *20*, 87-106.
112. Cliff, N.; Kanouni, M.; Peters, C.; Yaneff, P.; Adamsons, K., Use of Reactable Light Stabilizers to Prevent Migration and to Improve Durability of Coatings on Plastic Substrates. *Journal of Coatings Technology and Research* **2005**, *2* (5), 371-387.

113. Grieve, M.; Dunlop, J.; Haddock, P., An Investigation of Known Blue, Red, and Black Dyes Used in the Coloration of Cotton Fibers. *Journal of Forensic Sciences* **1990**, 35 (2), 301-315.
114. Wiggins, K., Dye Batch Variation in Textile Fibers. *Journal of Forensic Sciences* **1988**, 33 (4), 998-1007.

APPENDICES

Appendix A. Clear Coats (Averaged Spectra)

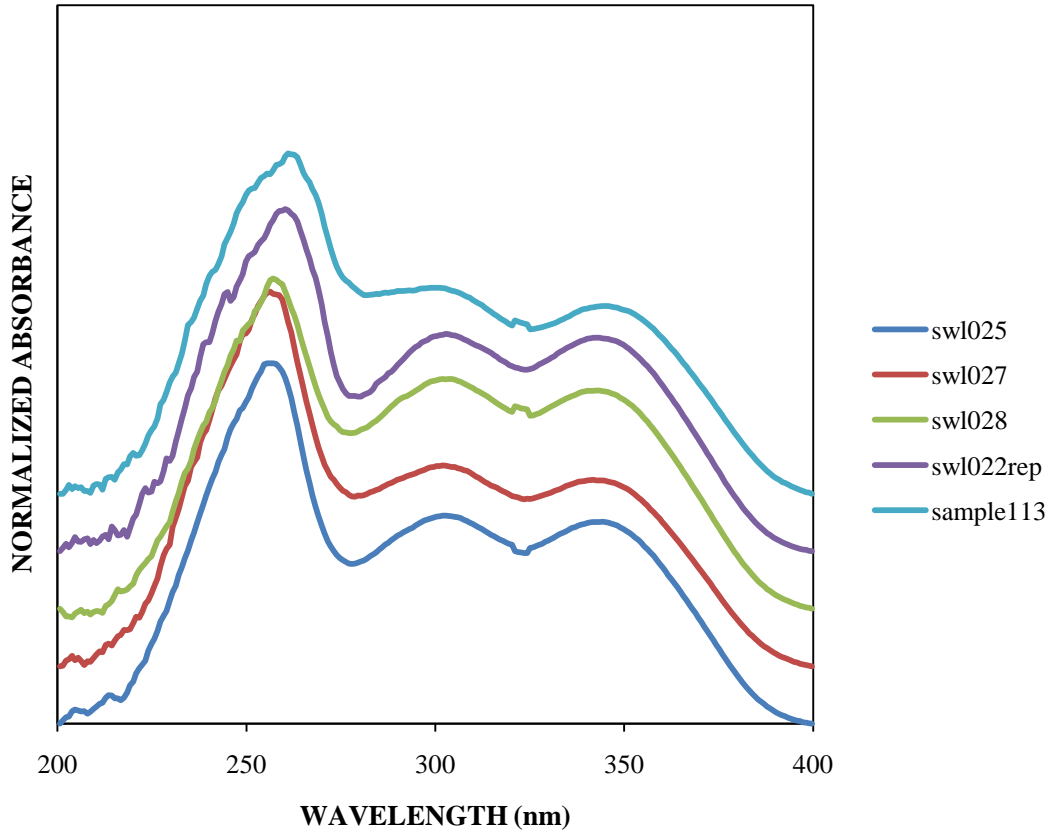


Figure A.1

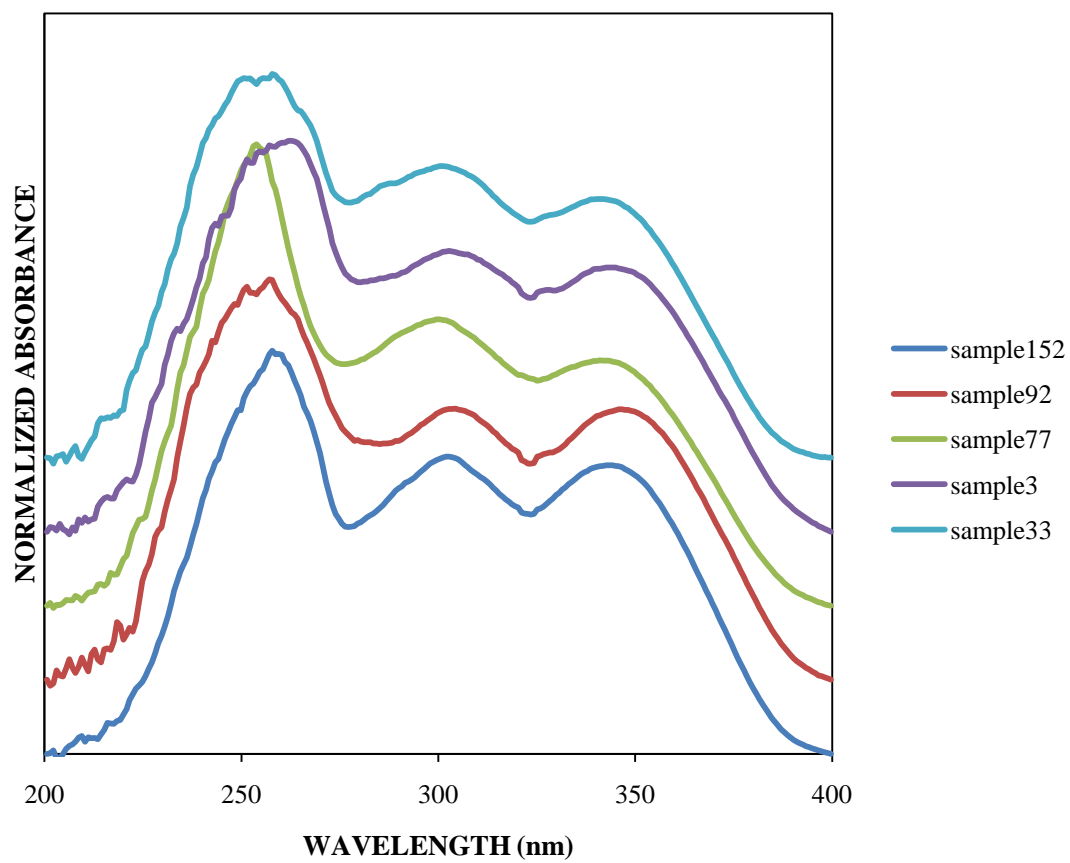


Figure A.2

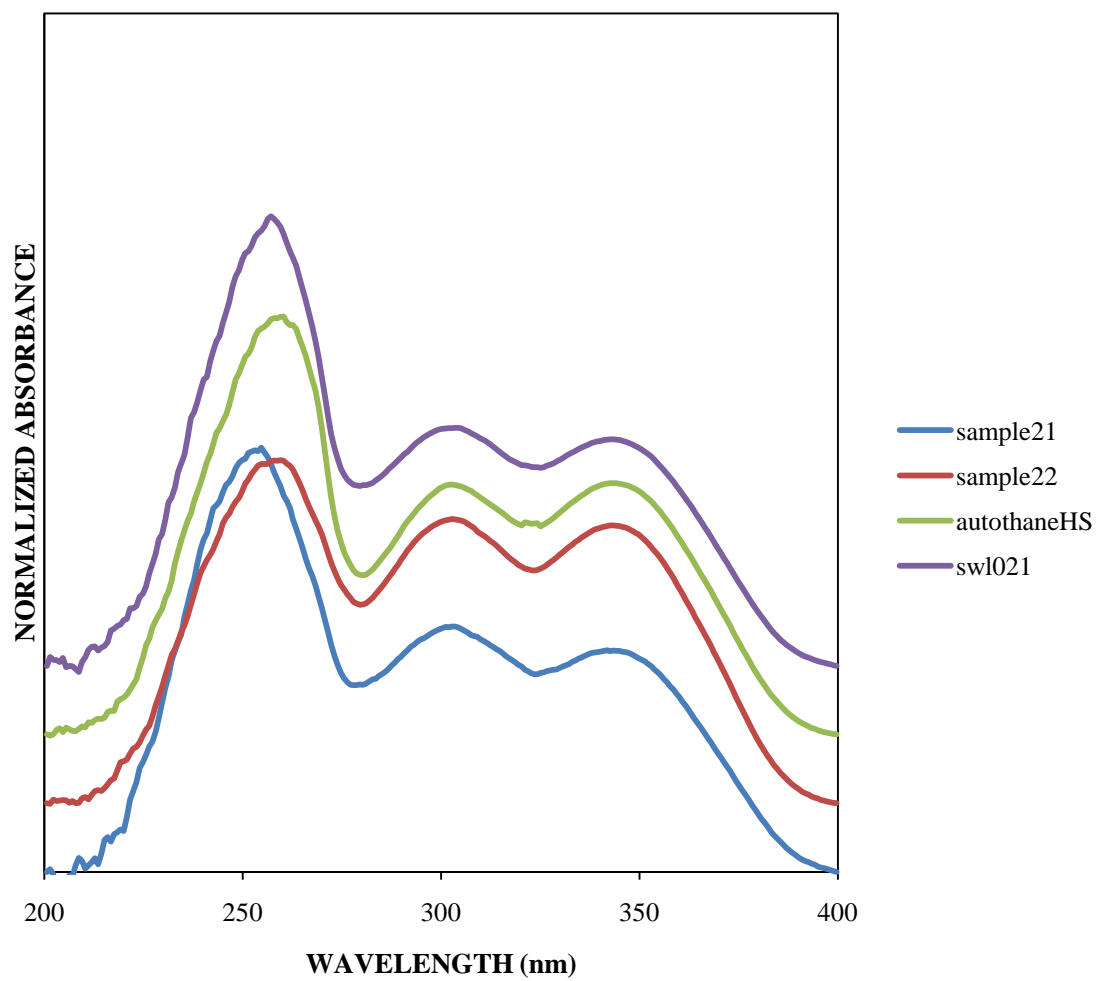


Figure A.3

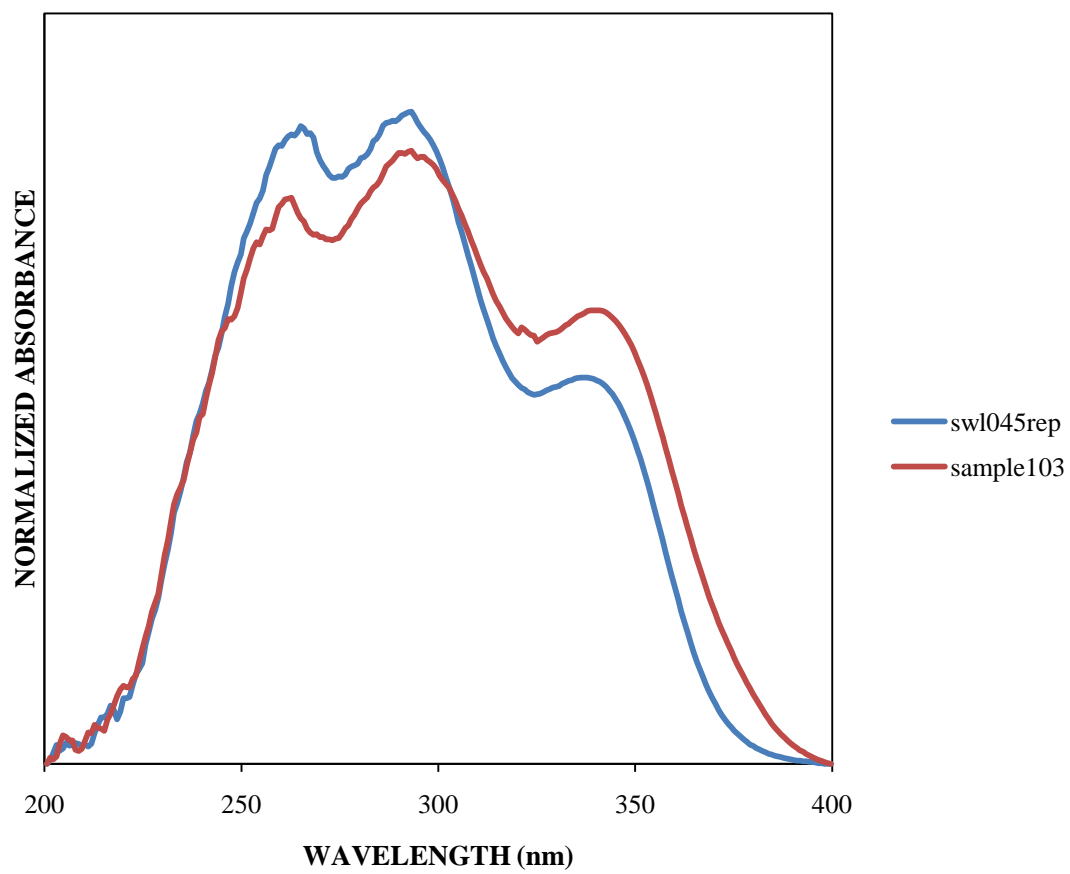


Figure A.4

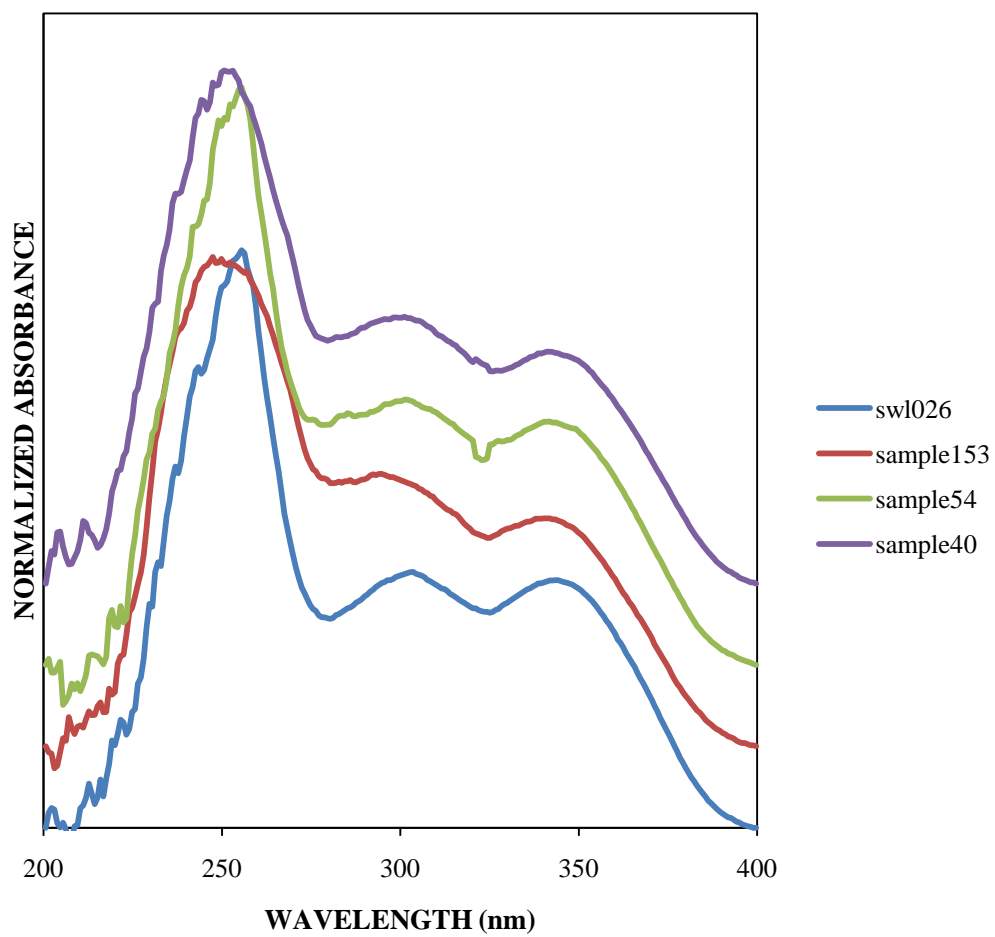


Figure A.5

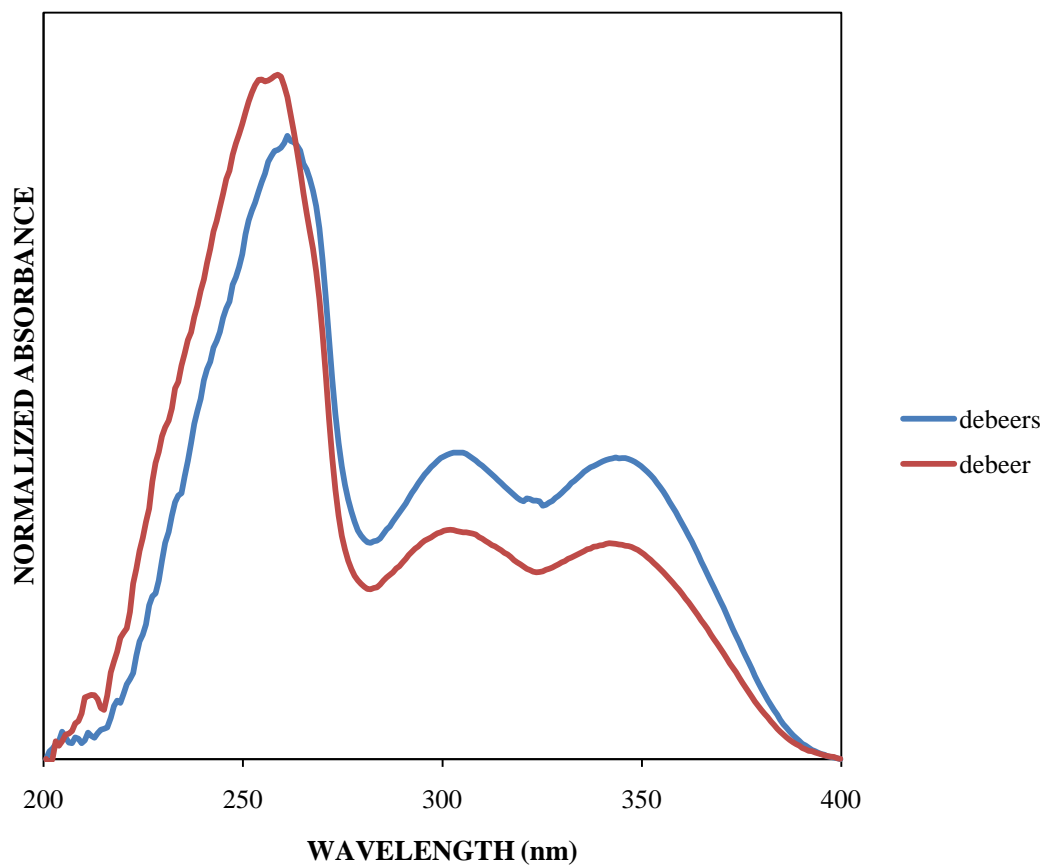


Figure A.6

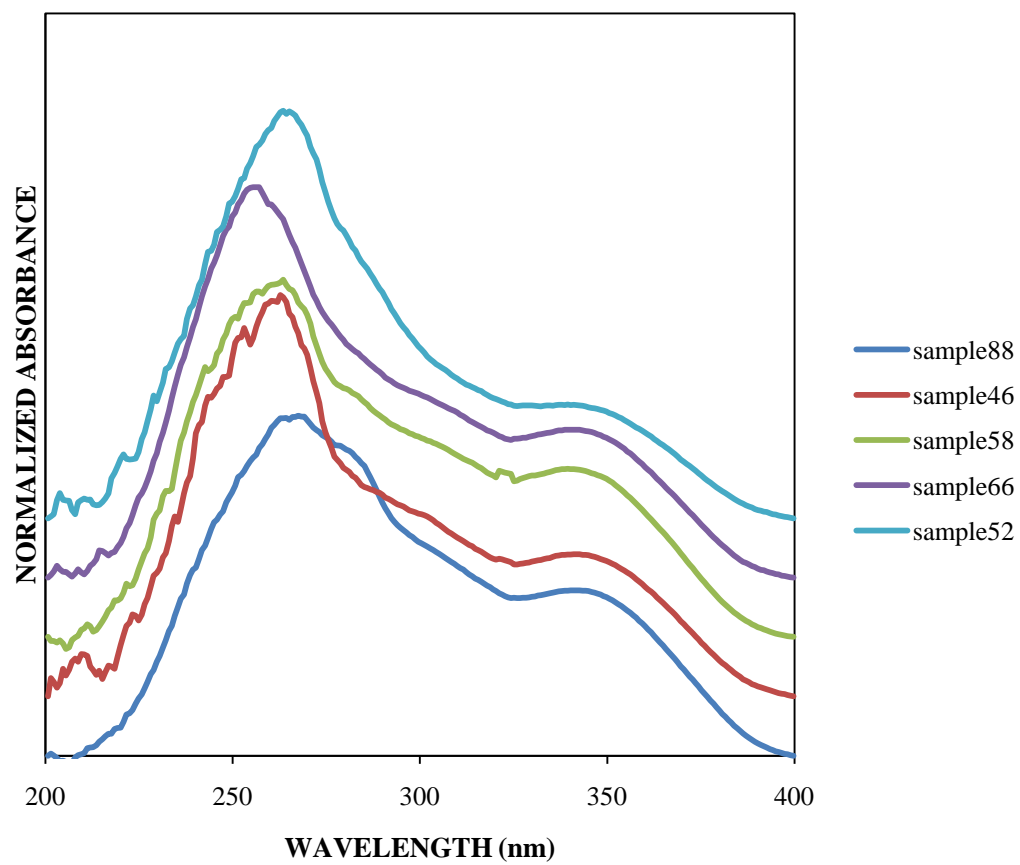


Figure A.7

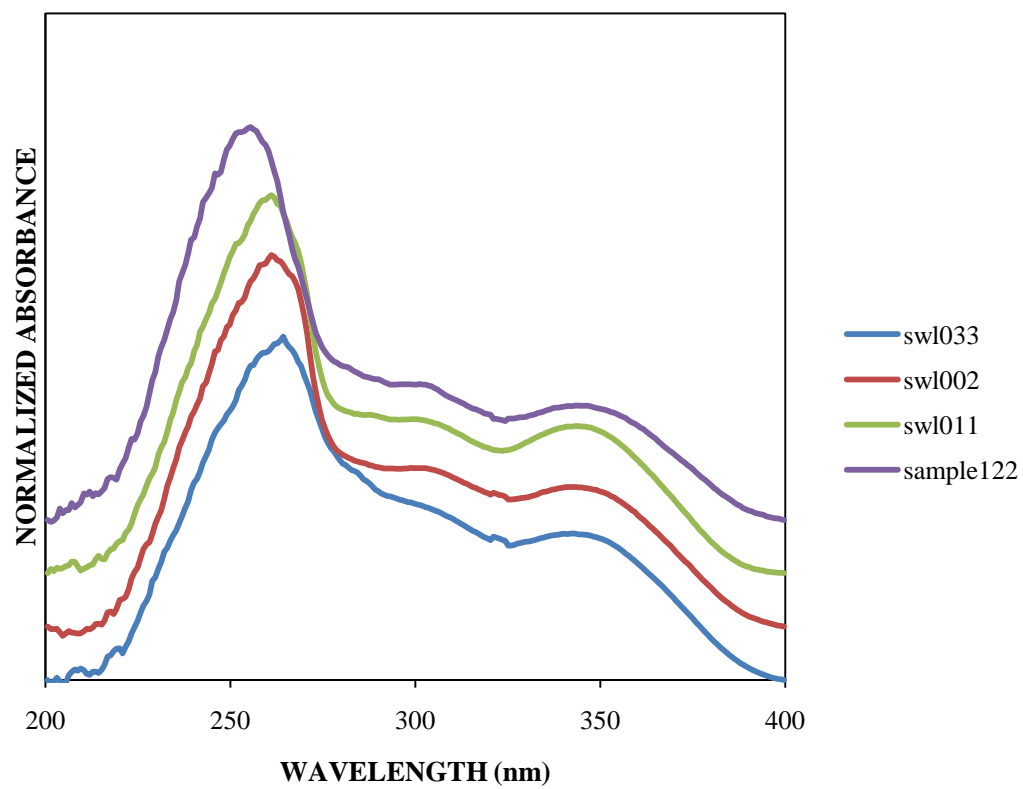


Figure A.8

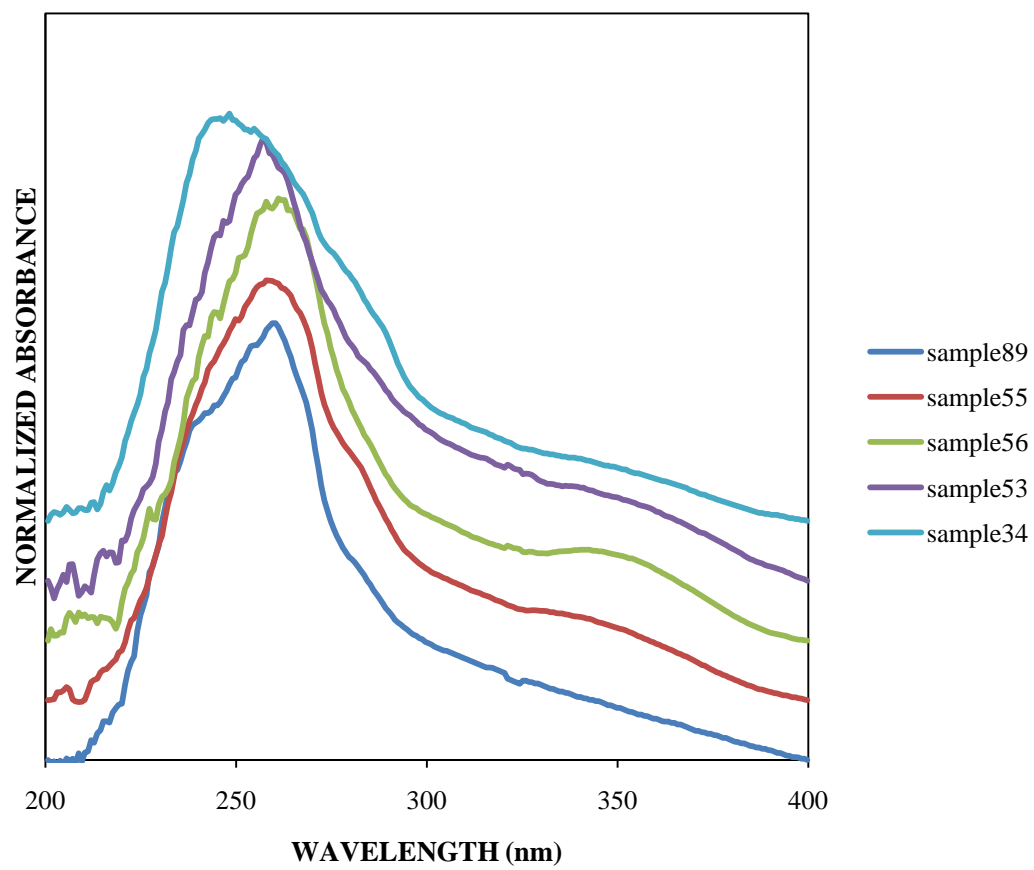


Figure A.9

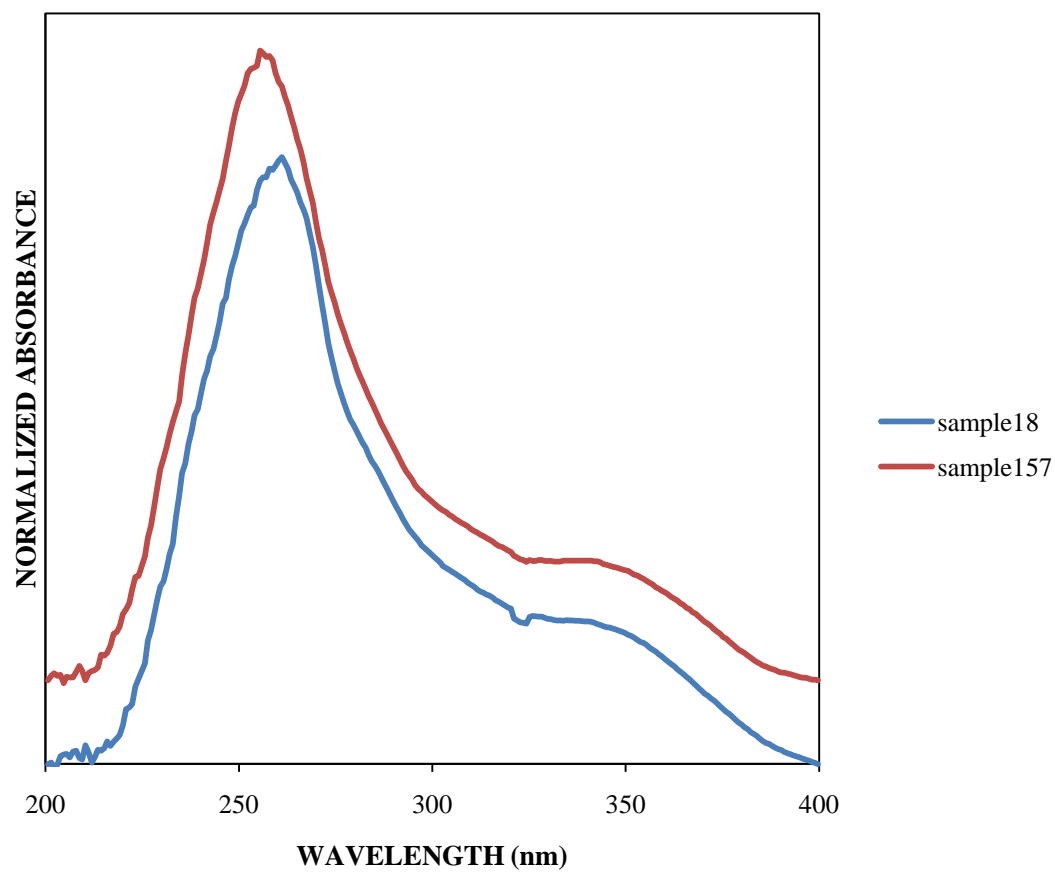


Figure A.10

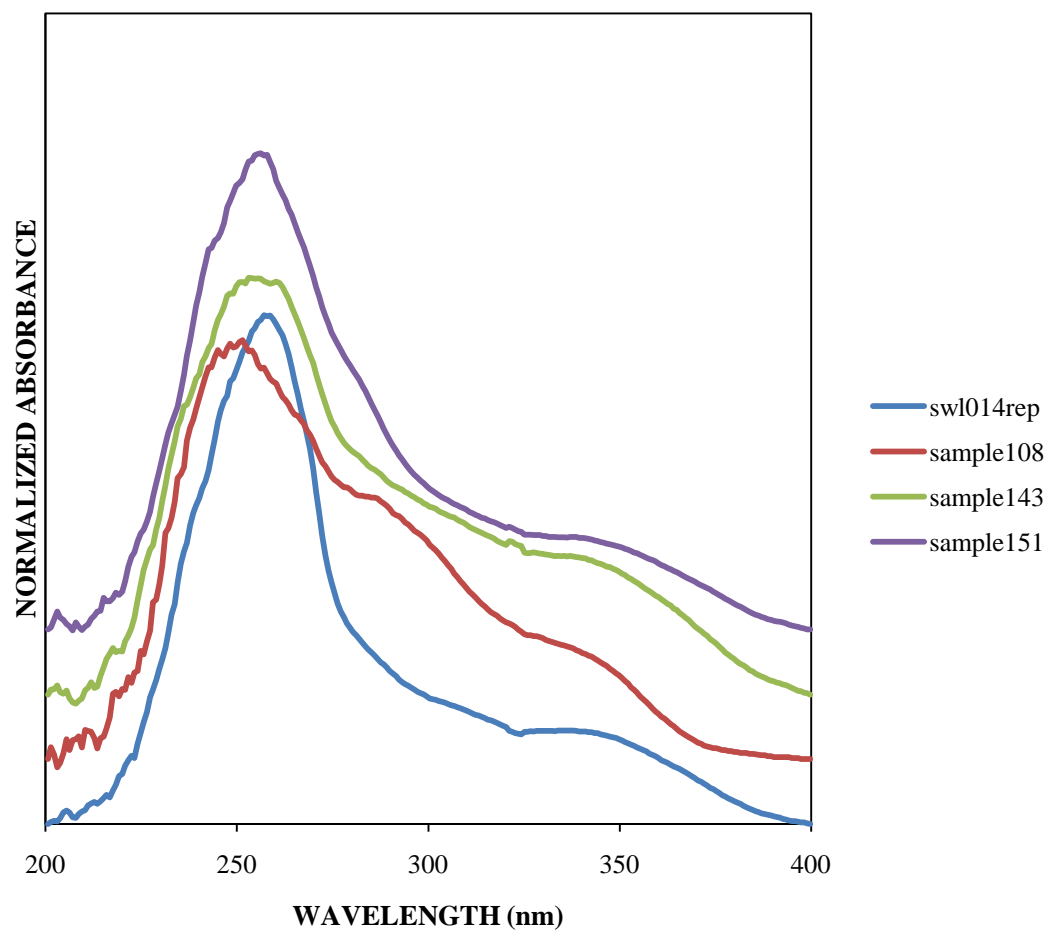


Figure A.11

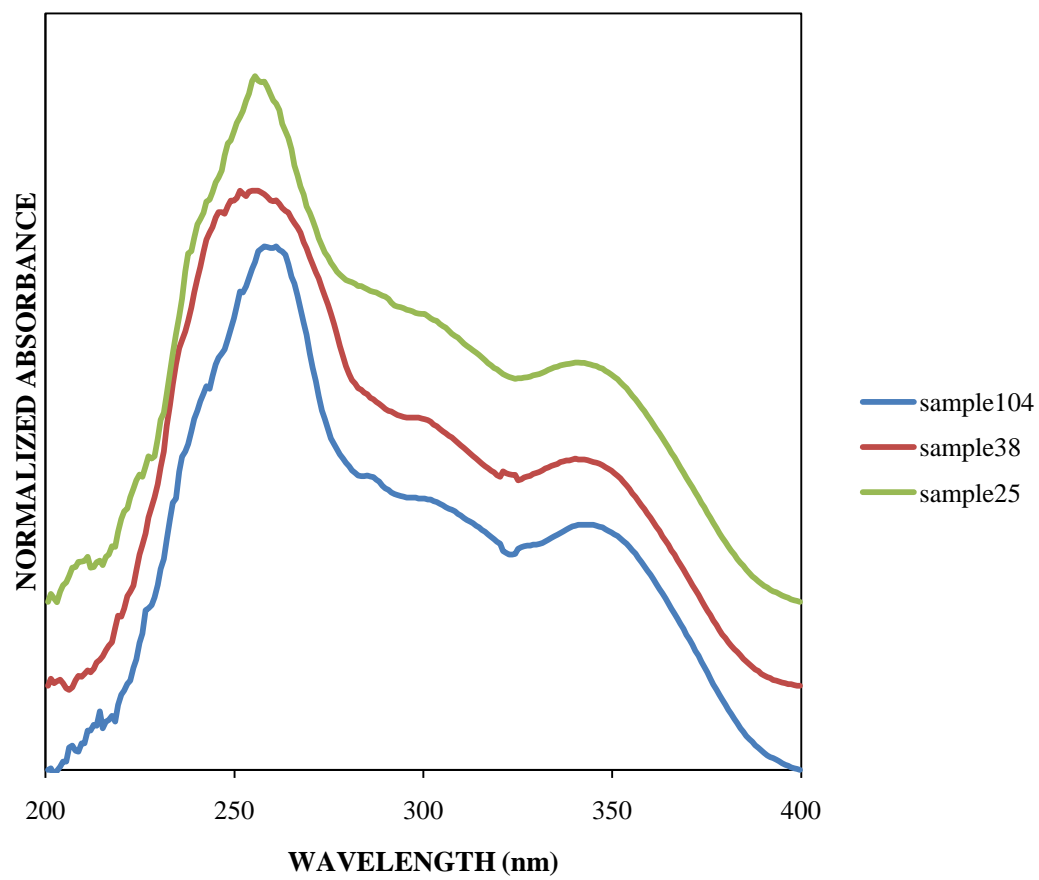


Figure A.12

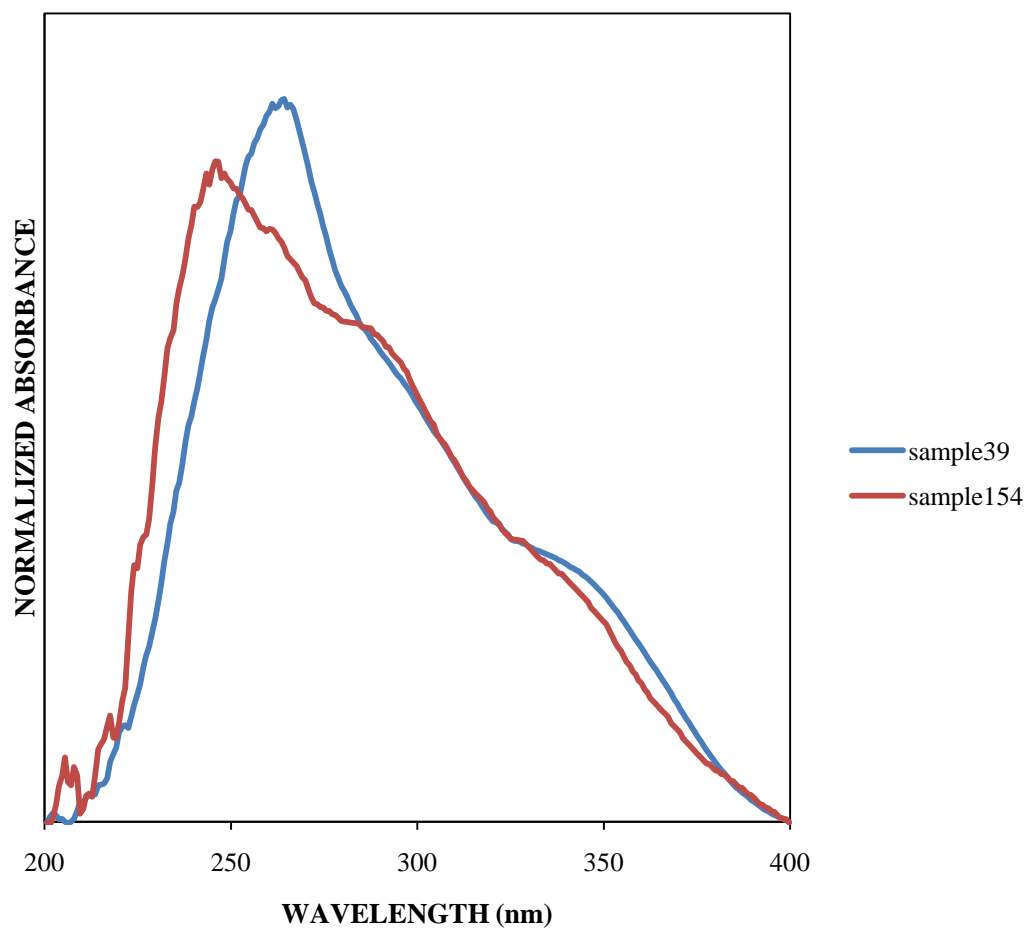


Figure A.13

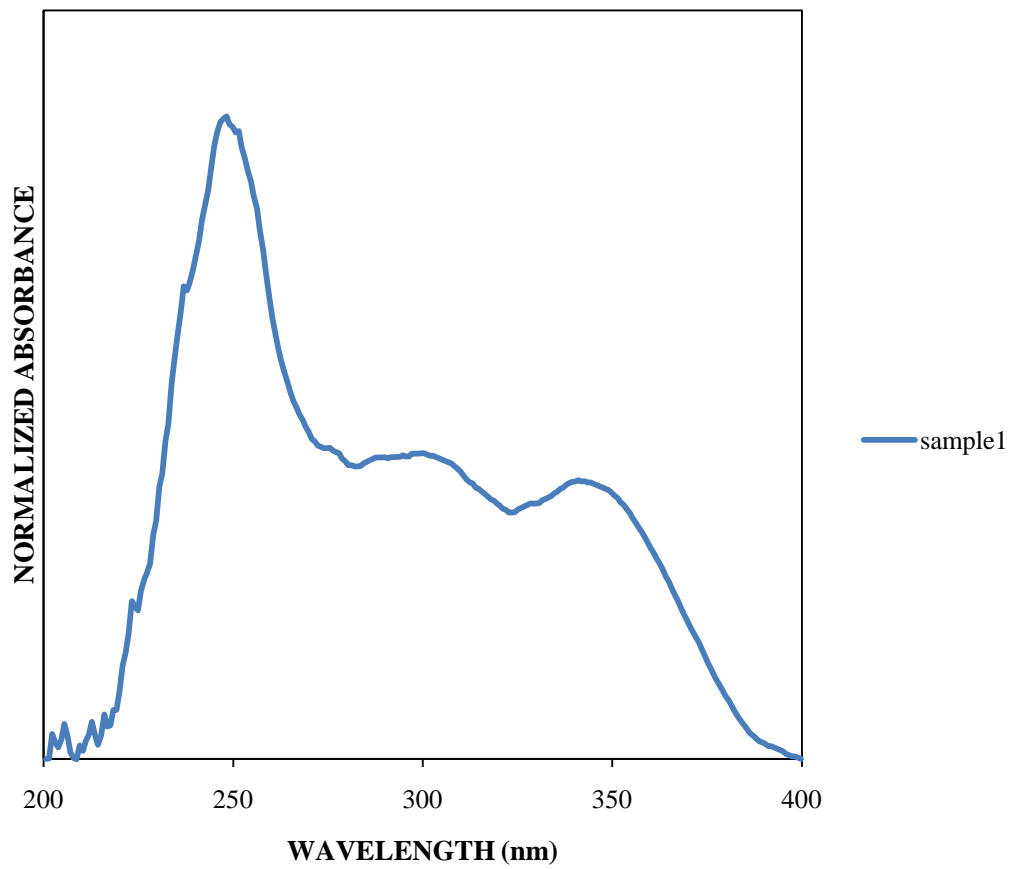


Figure A.14

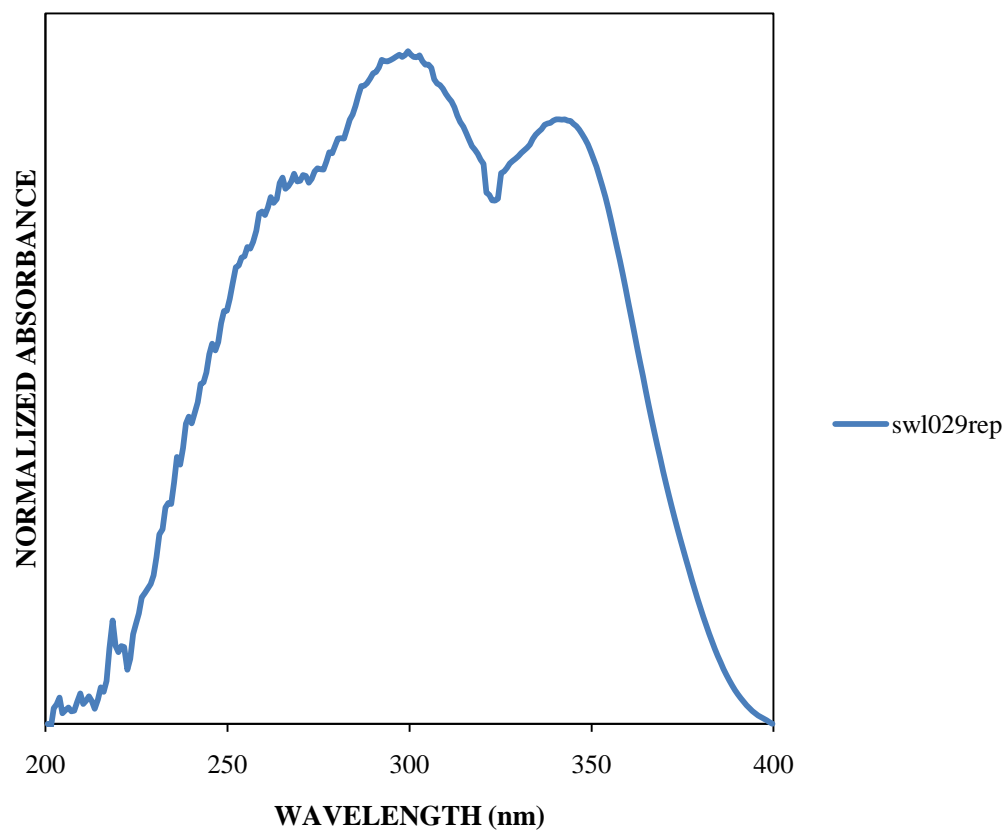


Figure A.15

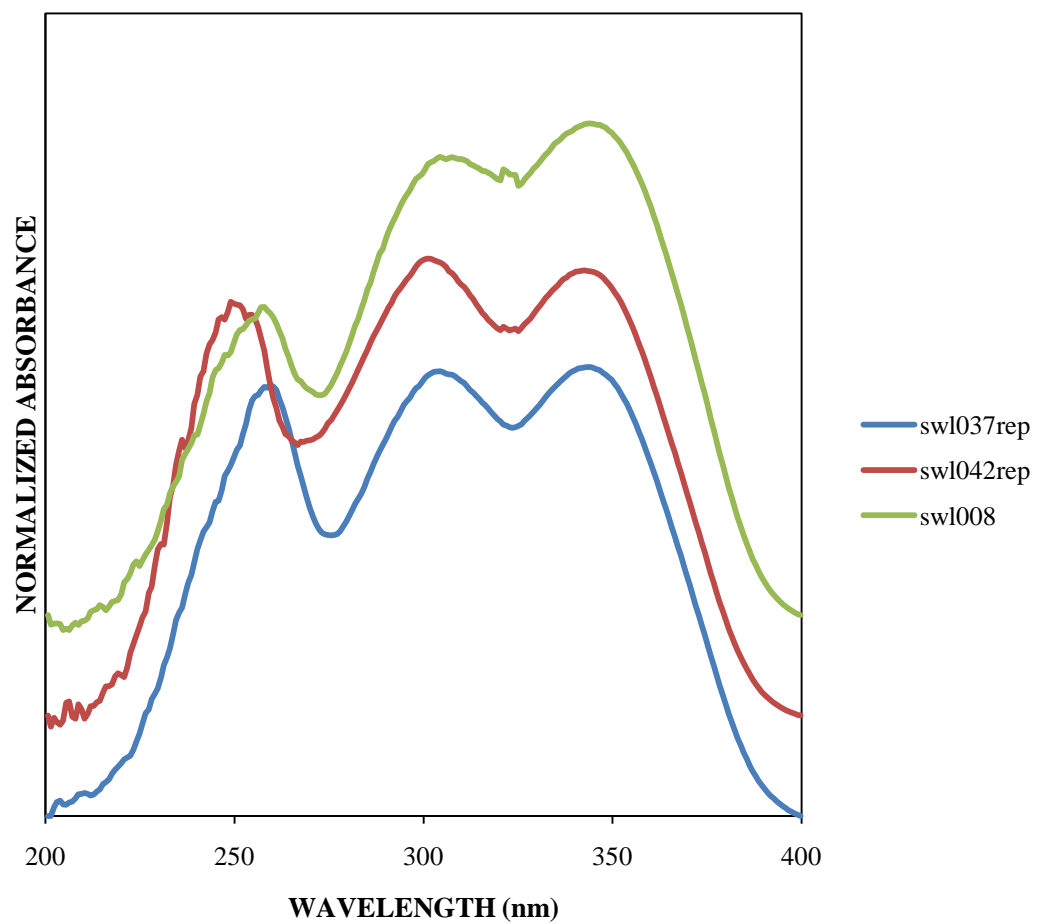


Figure A.16

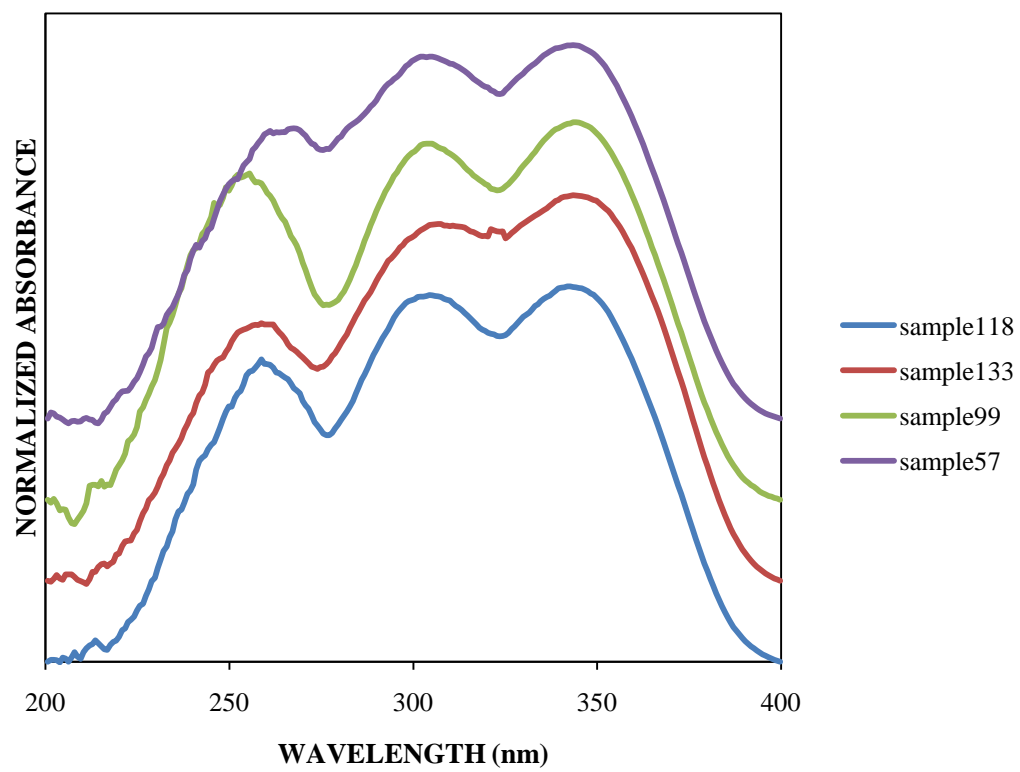


Figure A.17

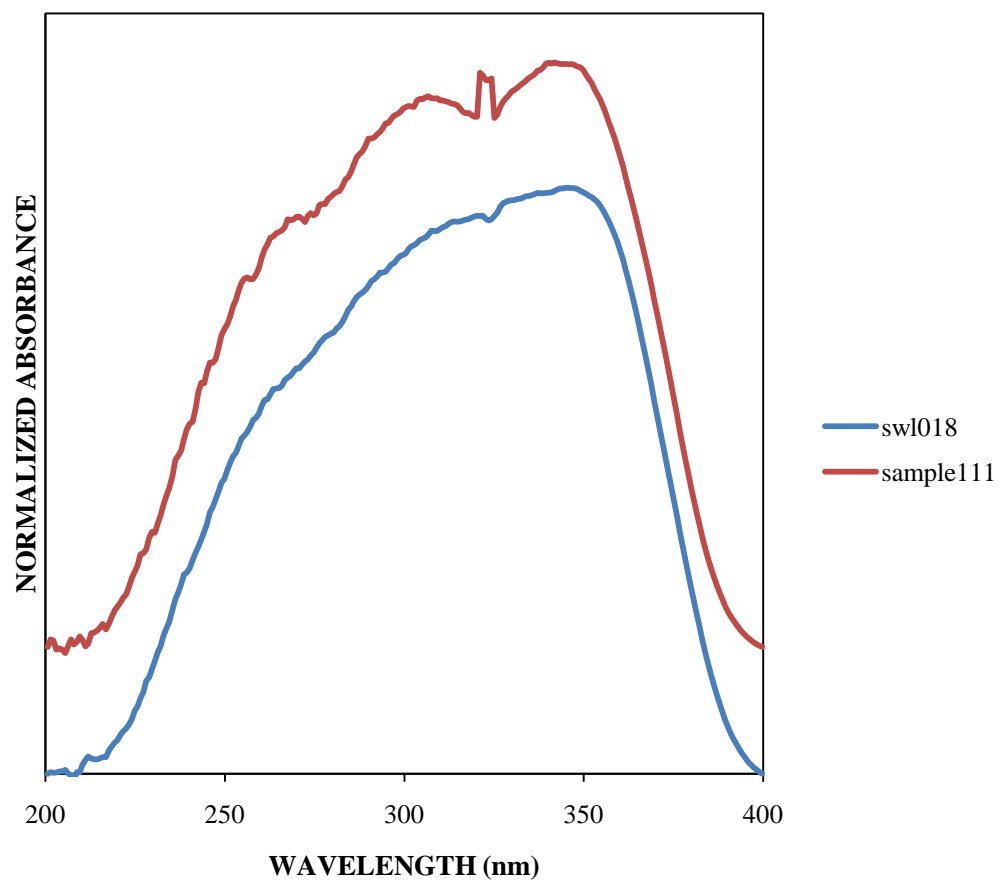


Figure A.18

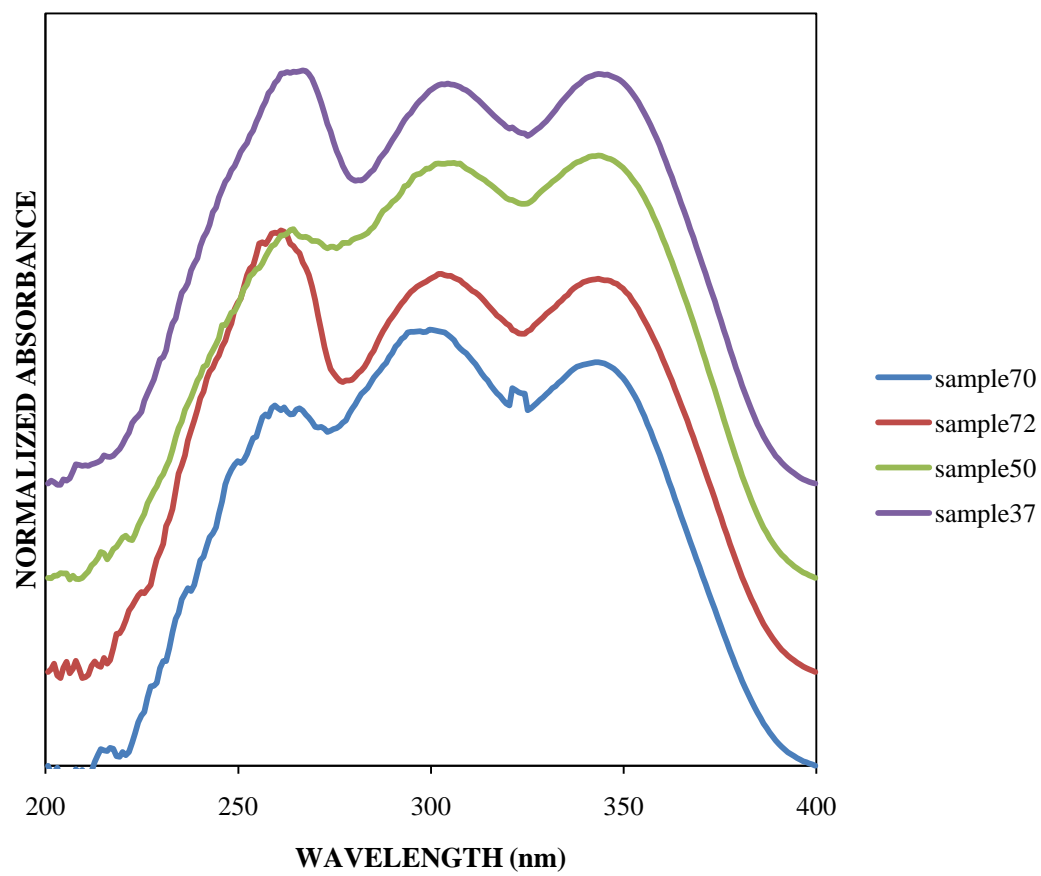


Figure A.19

Appendix B. Global Dye Averages

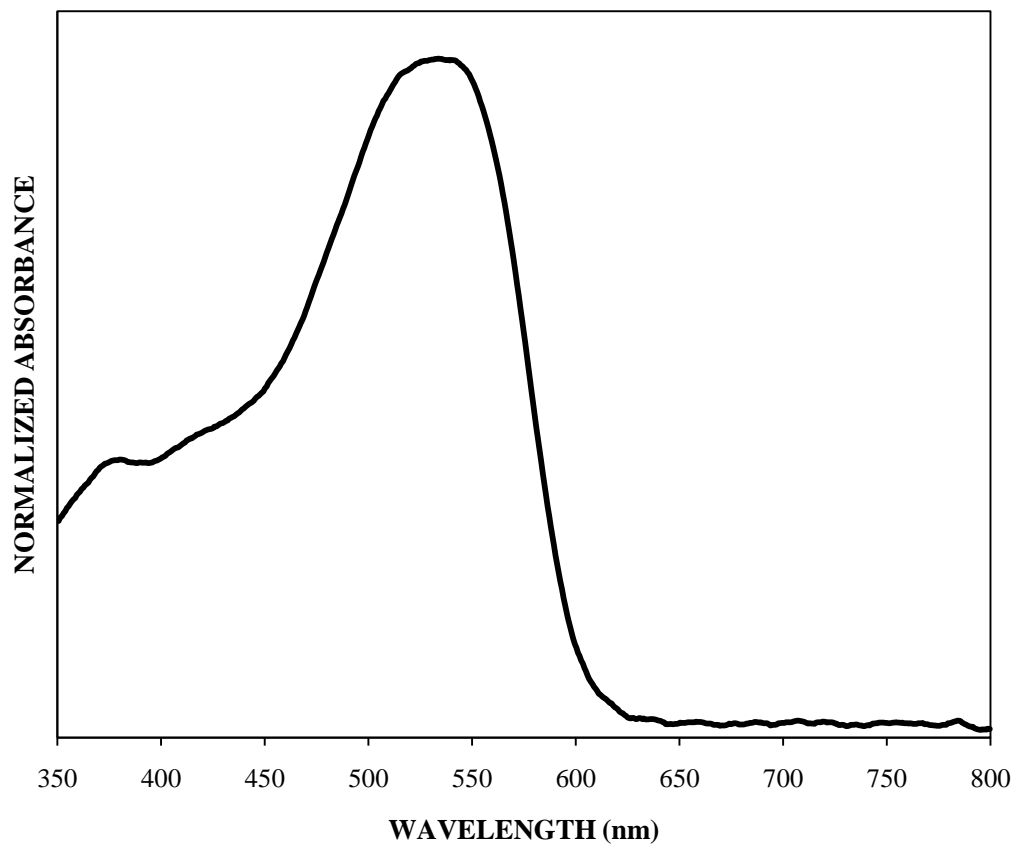


Figure B.1 Fiber A (IUPUI – first set)

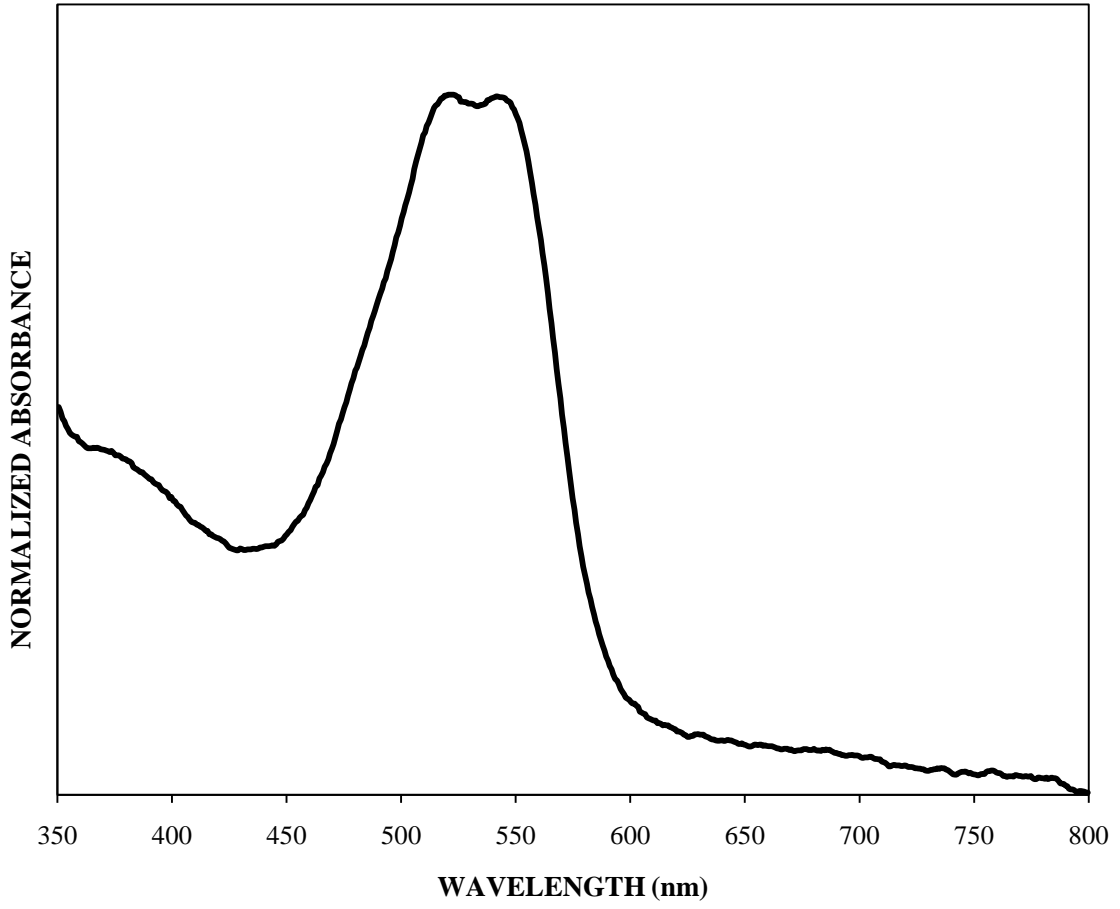


Figure B.2 Fiber B (IUPUI – first set)

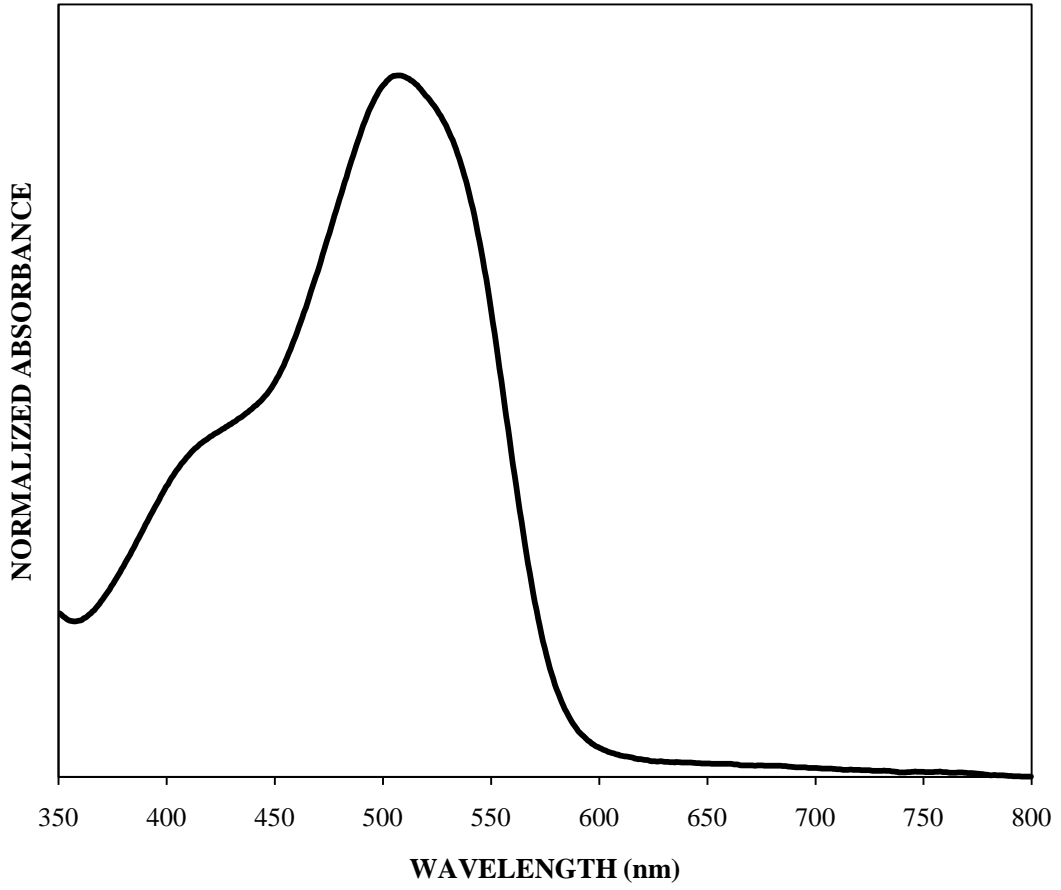


Figure B.3 Fiber C (IUPUI – first set)

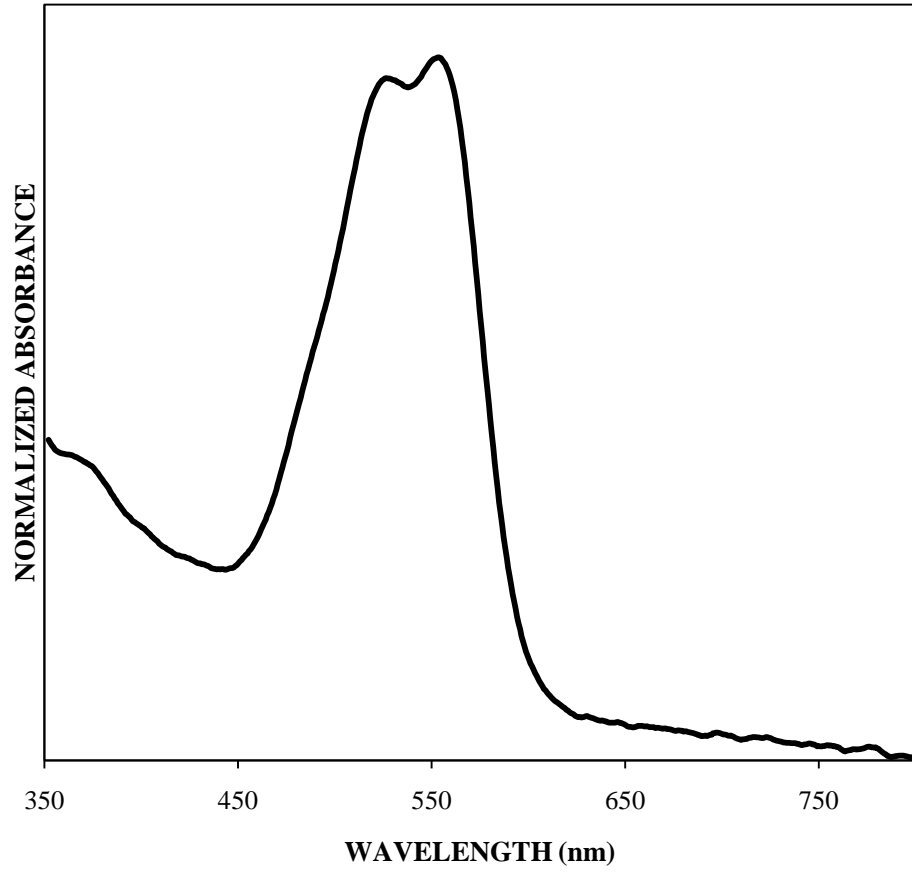


Figure B.4 Fiber D (IUPUI – first set)

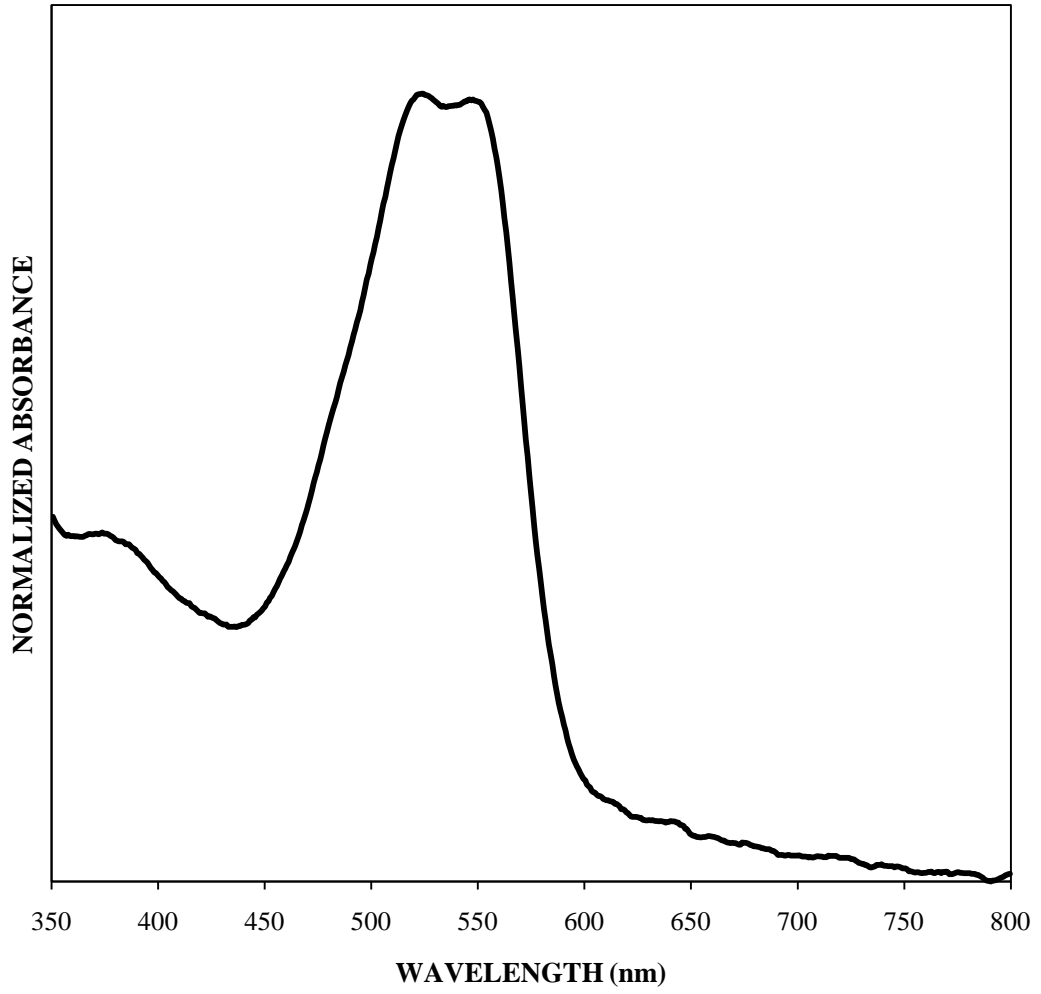


Figure B.5 Fiber E (IUPUI – first set)

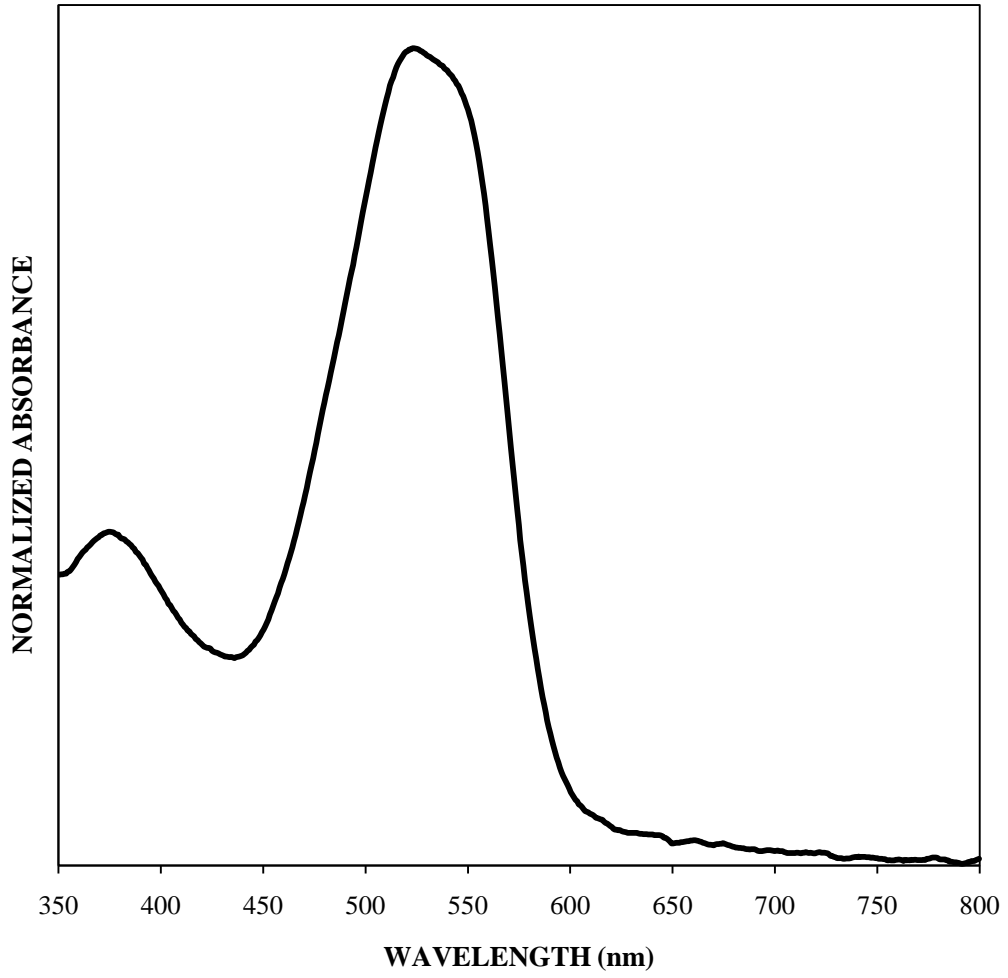


Figure B.6 Fiber F (IUPUI – first set)

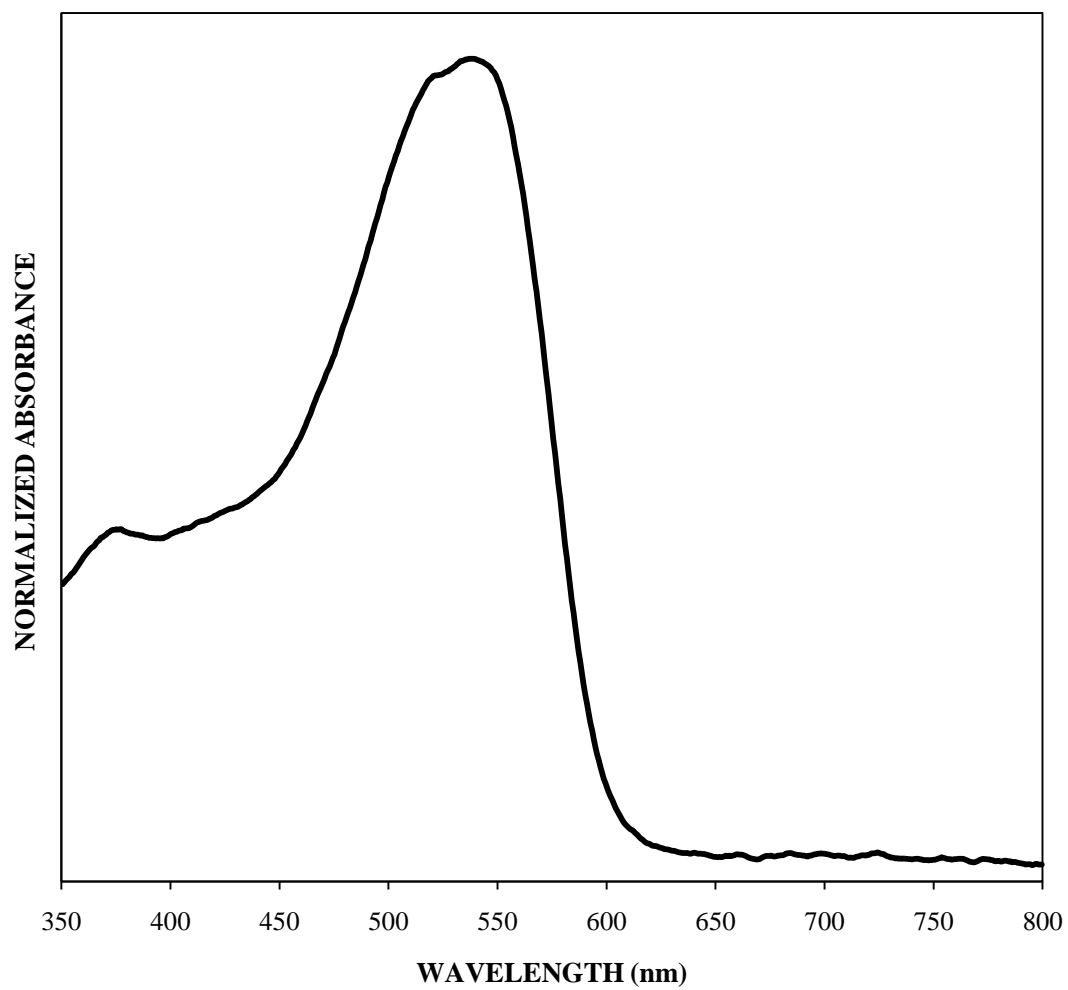


Figure B.7 Fiber A (IUPUI – first set; supplemental data)

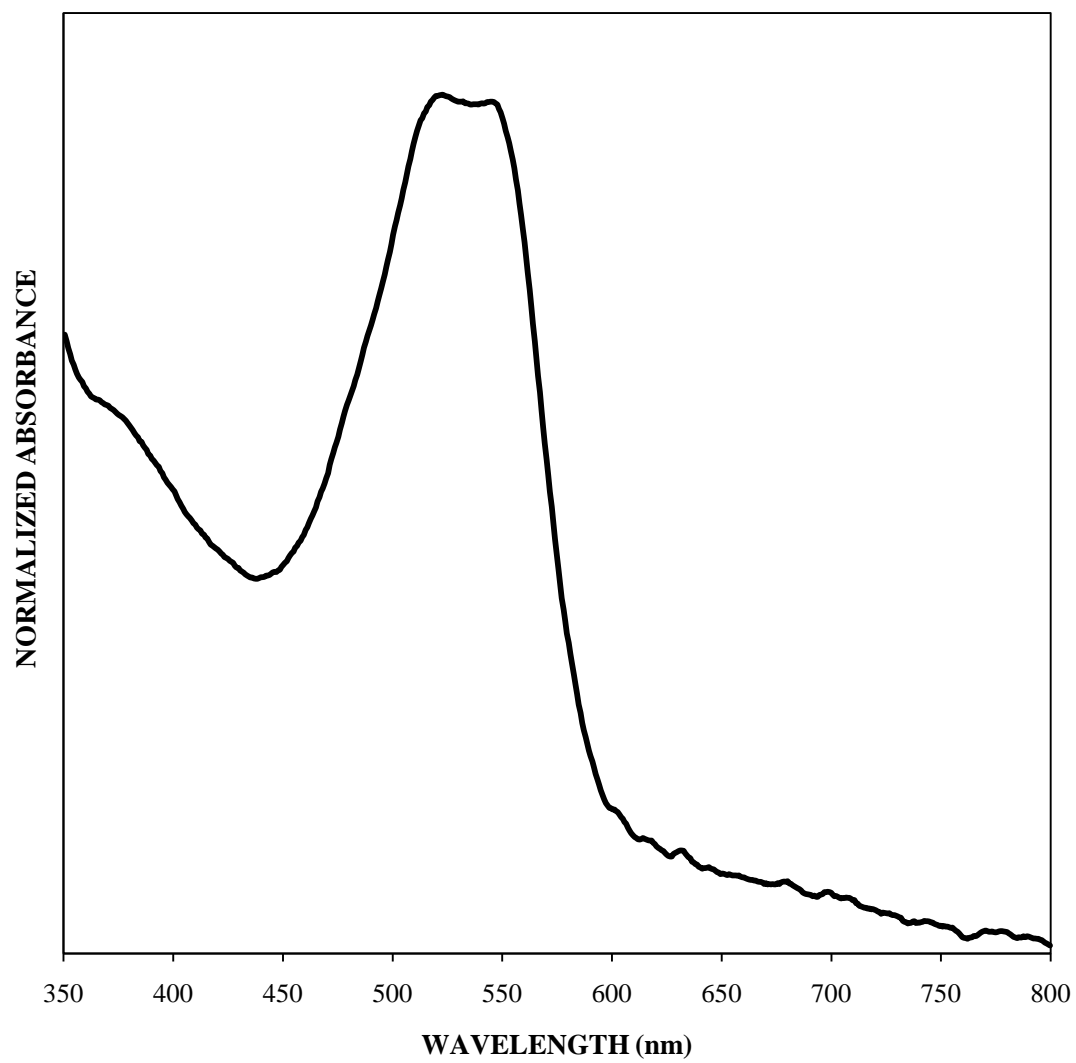


Figure B.8 Fiber B (IUPUI – first set; supplemental data)

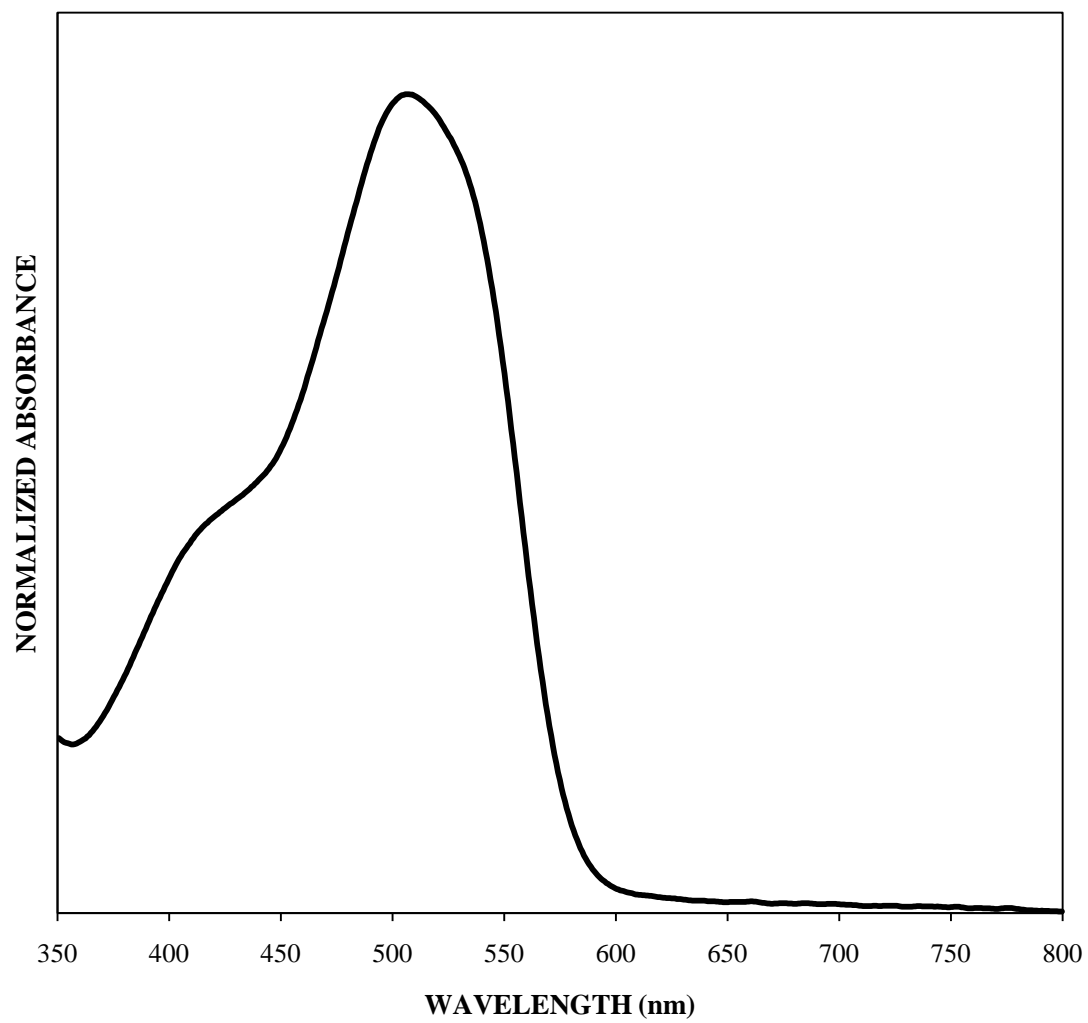


Figure B.9 Fiber C (IUPUI – first set; supplemental data)

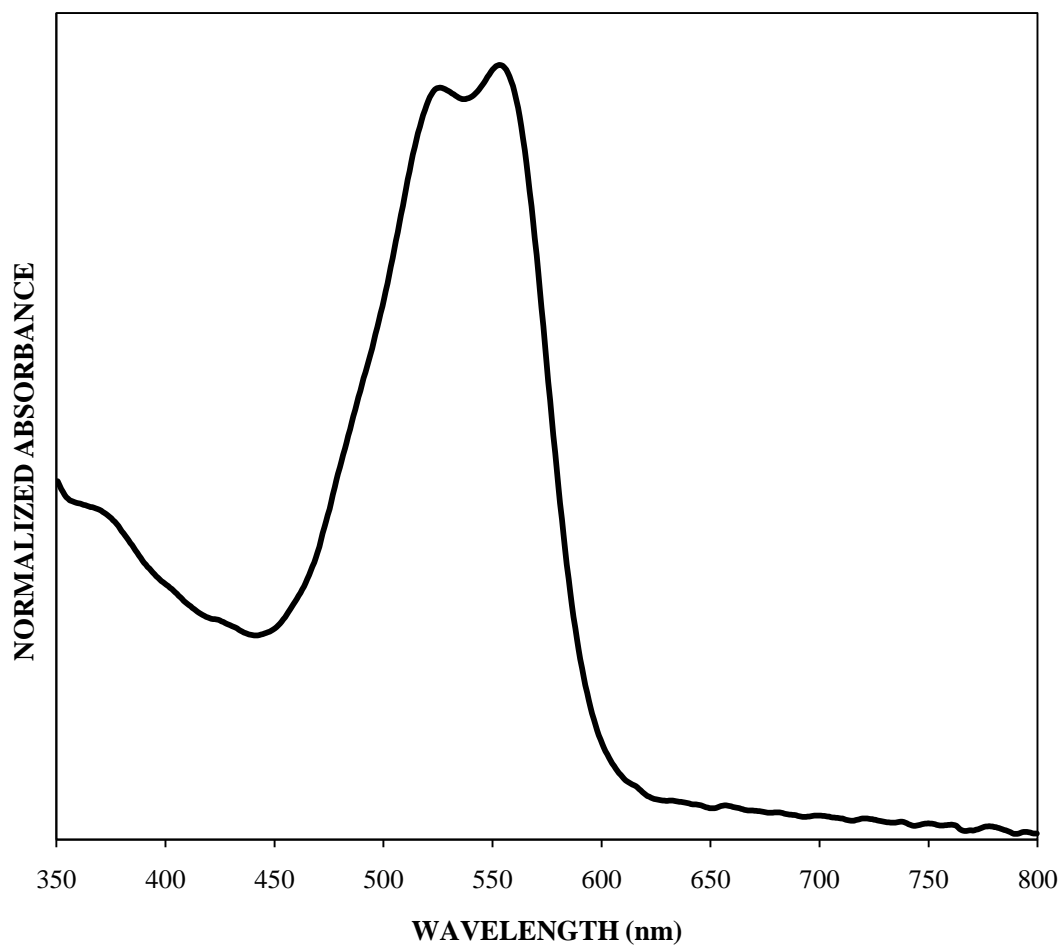


Figure B.10 Fiber D (IUPUI – first set; supplemental data)

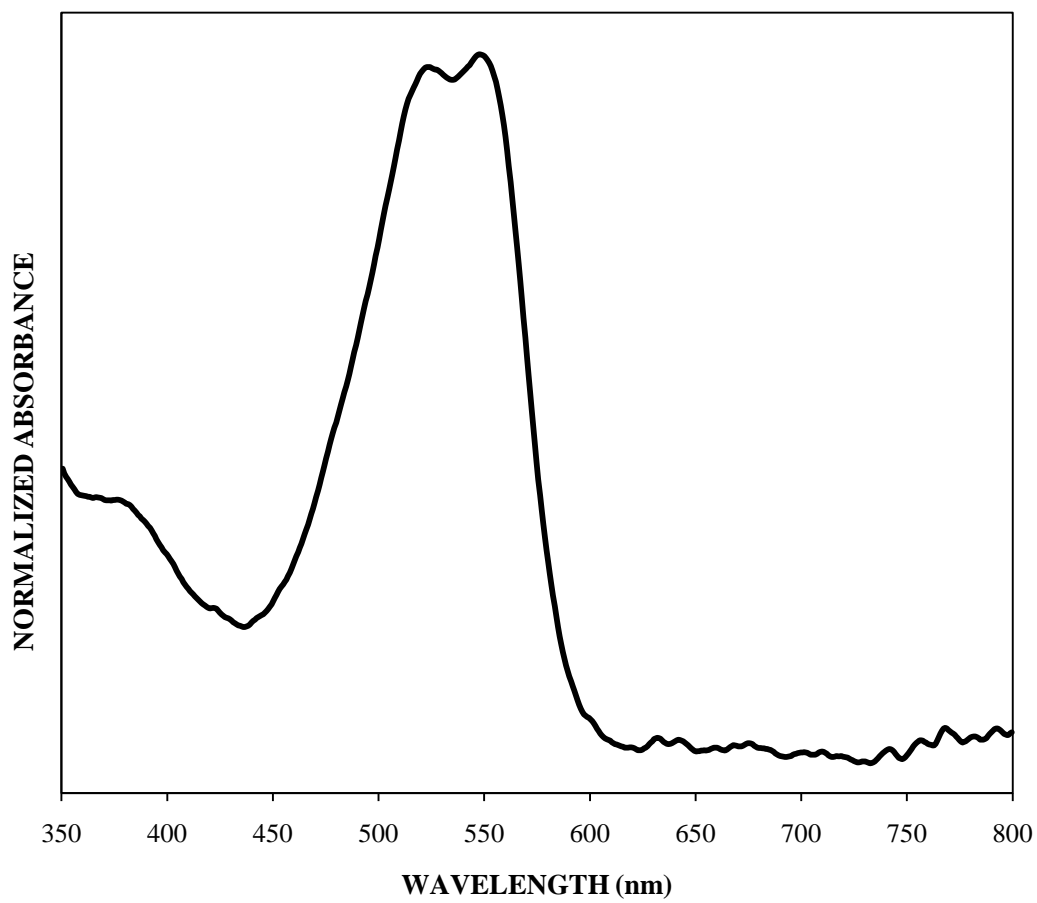


Figure B.11 Fiber E (IUPUI – first set; supplemental data)

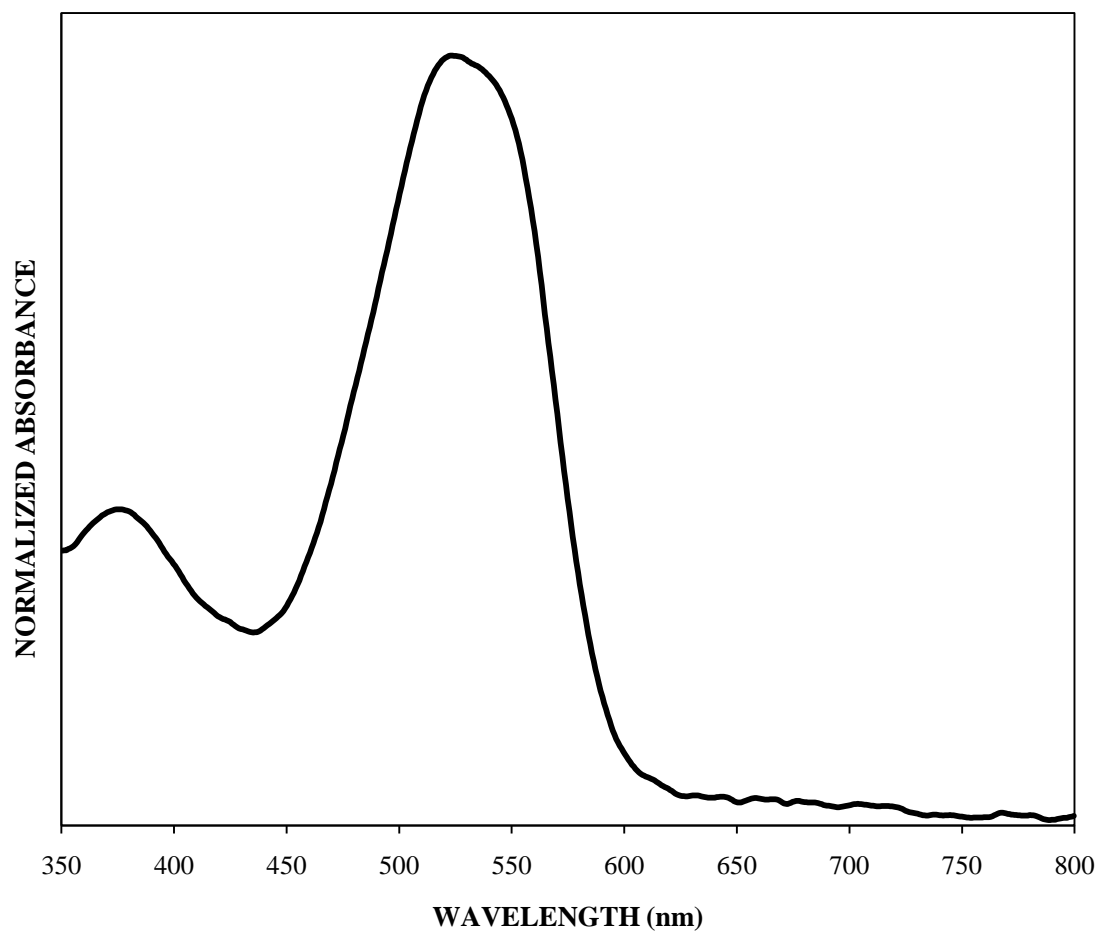


Figure B.12 Fiber F (IUPUI – first set; supplemental data)

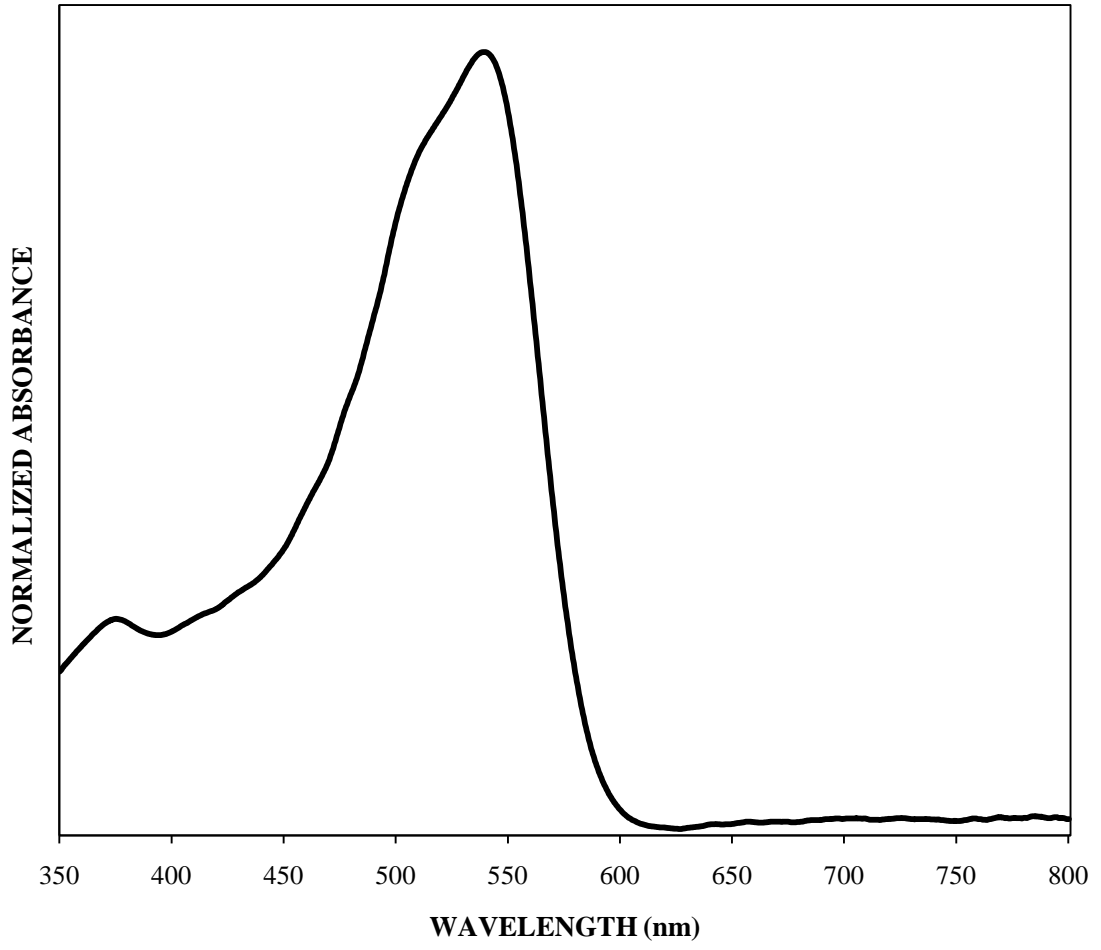


Figure B.13 Fiber A (ISP)

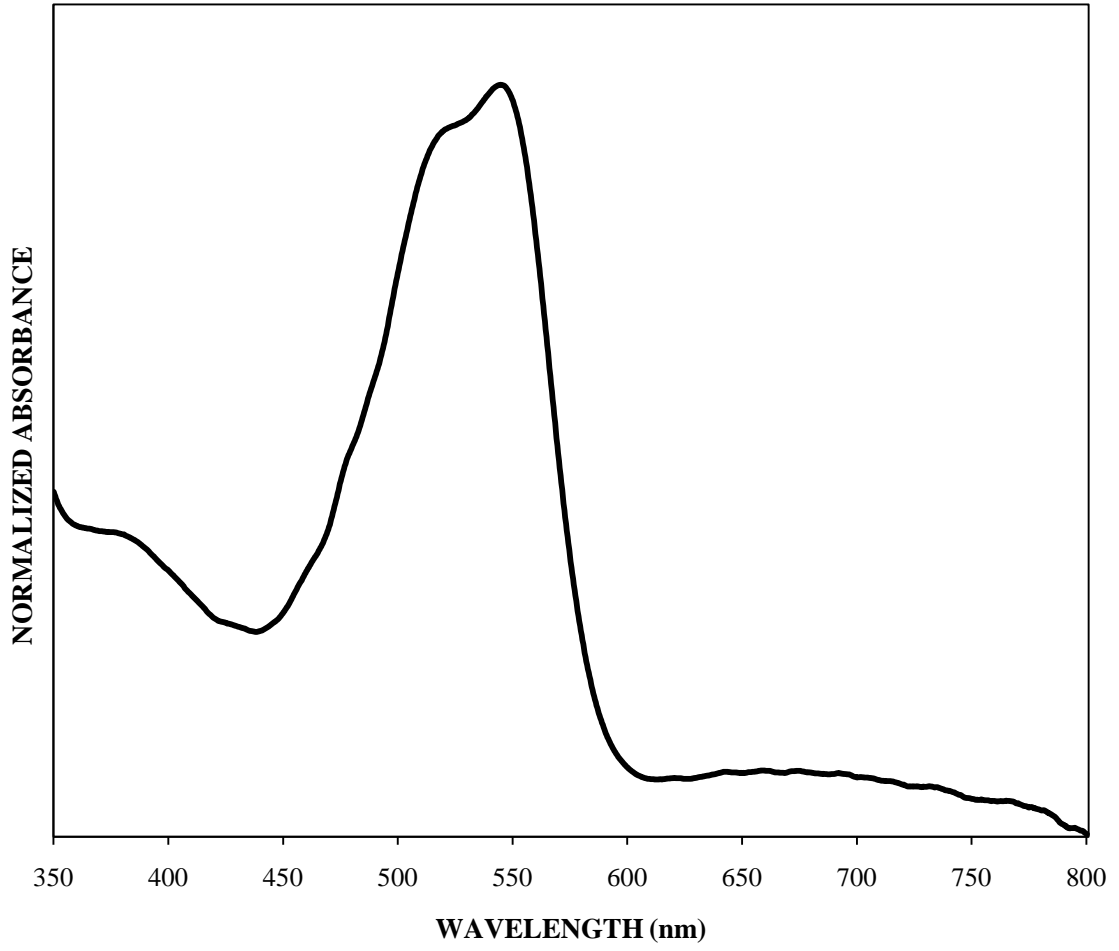


Figure B.14 Fiber B (ISP)

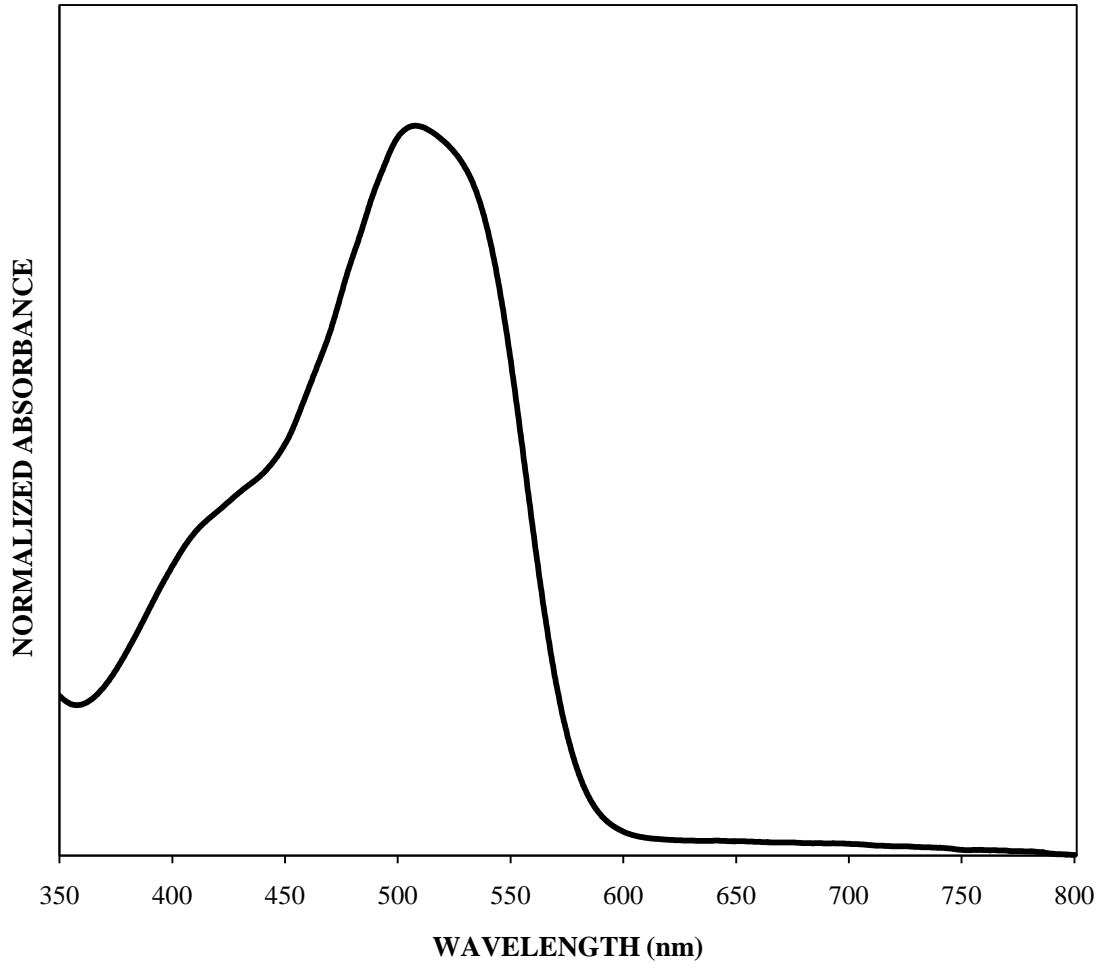


Figure B.15 Fiber C (ISP)

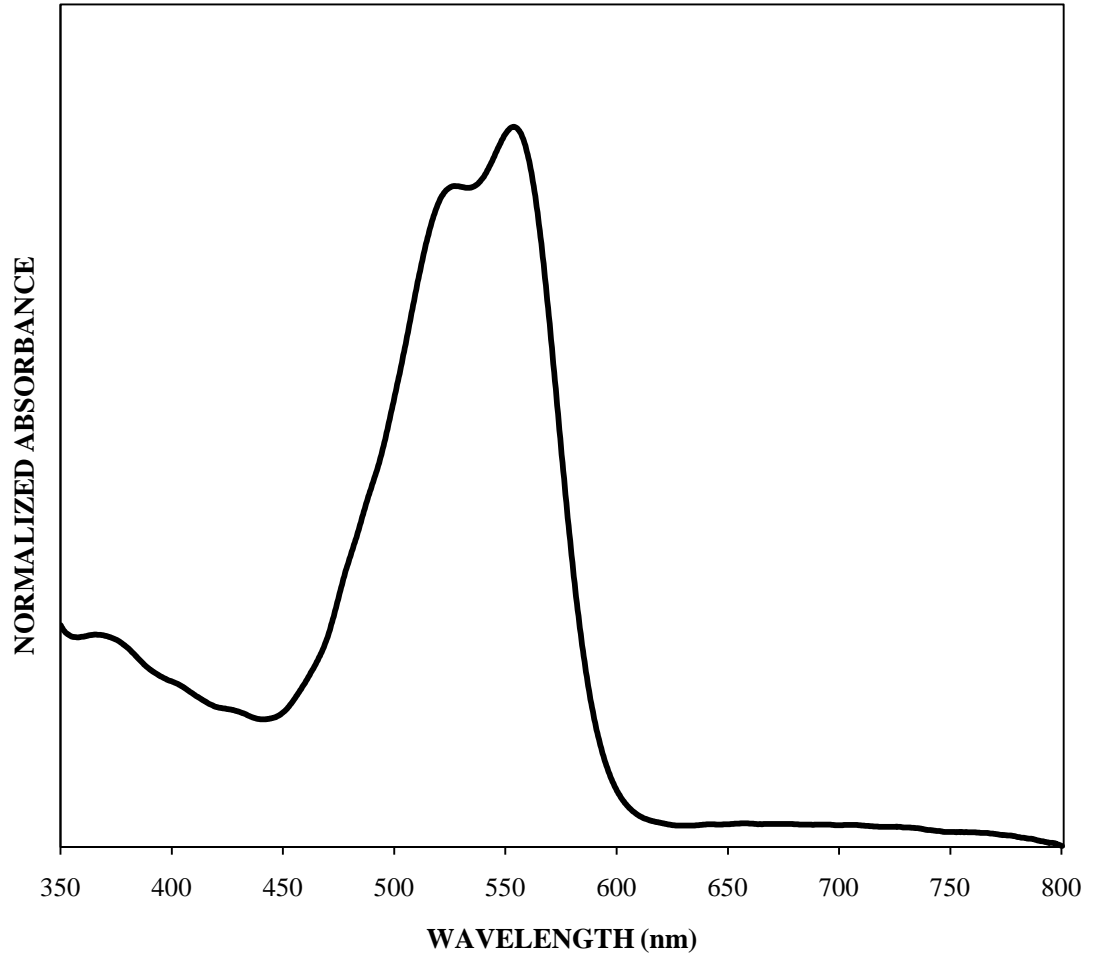


Figure B.16 Fiber D (ISP)

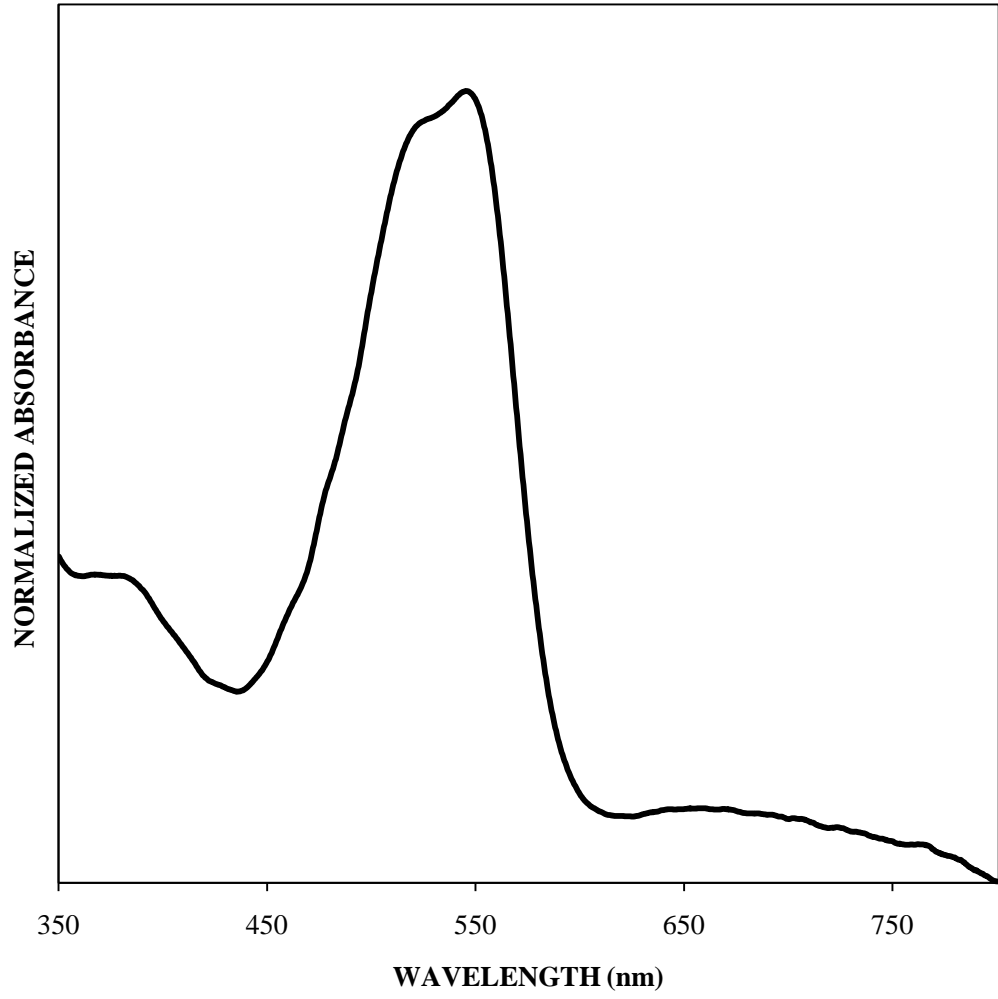


Figure B.17 Fiber E (ISP)

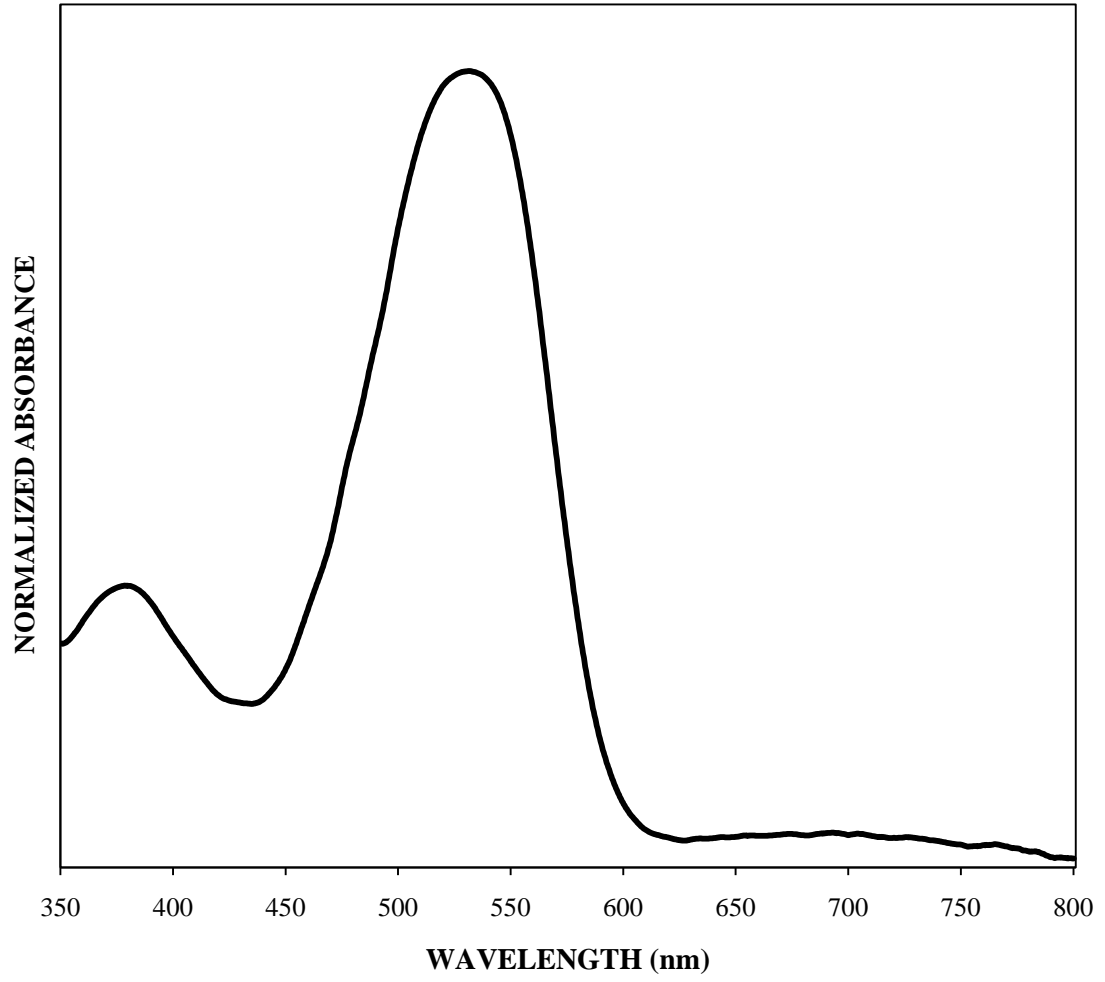


Figure B.18 Fiber F (ISP)

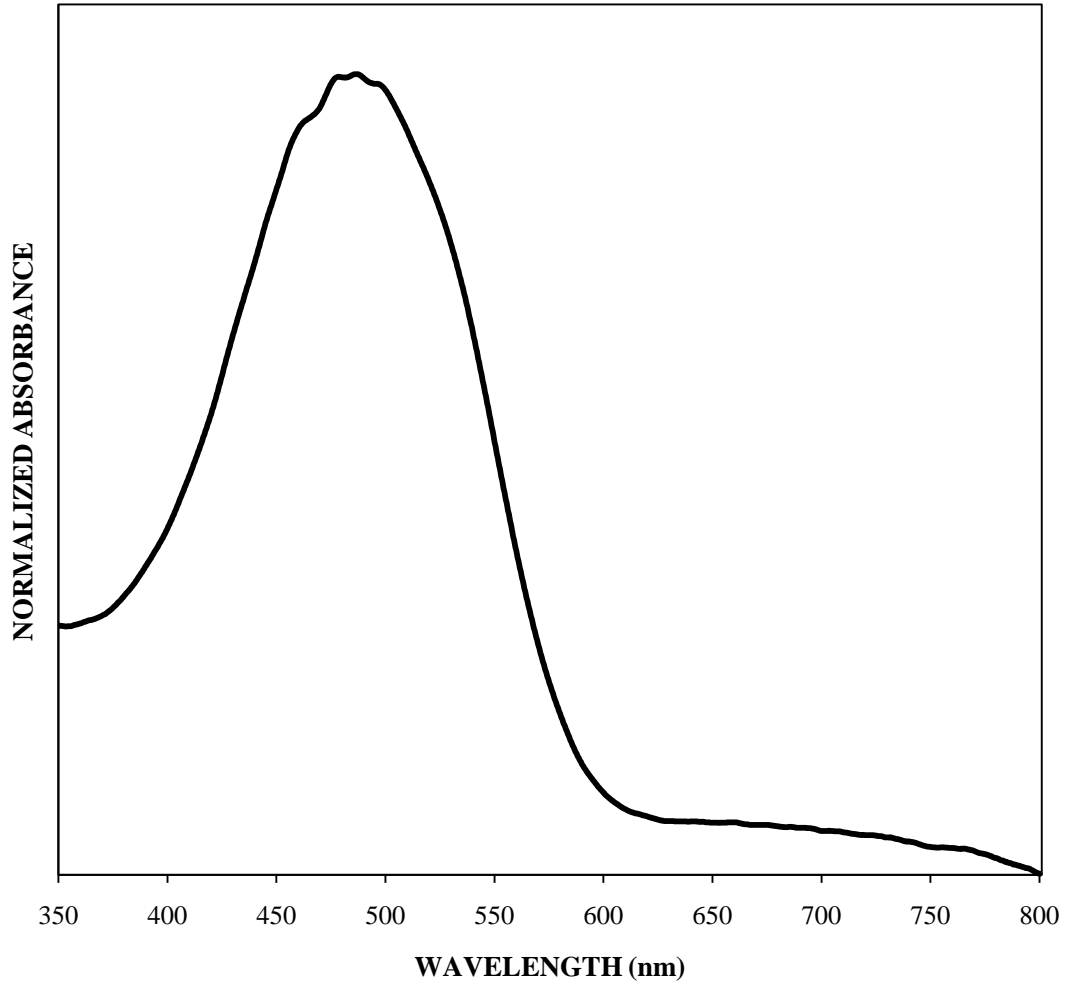


Figure B.19 Fiber 713 (ISP)

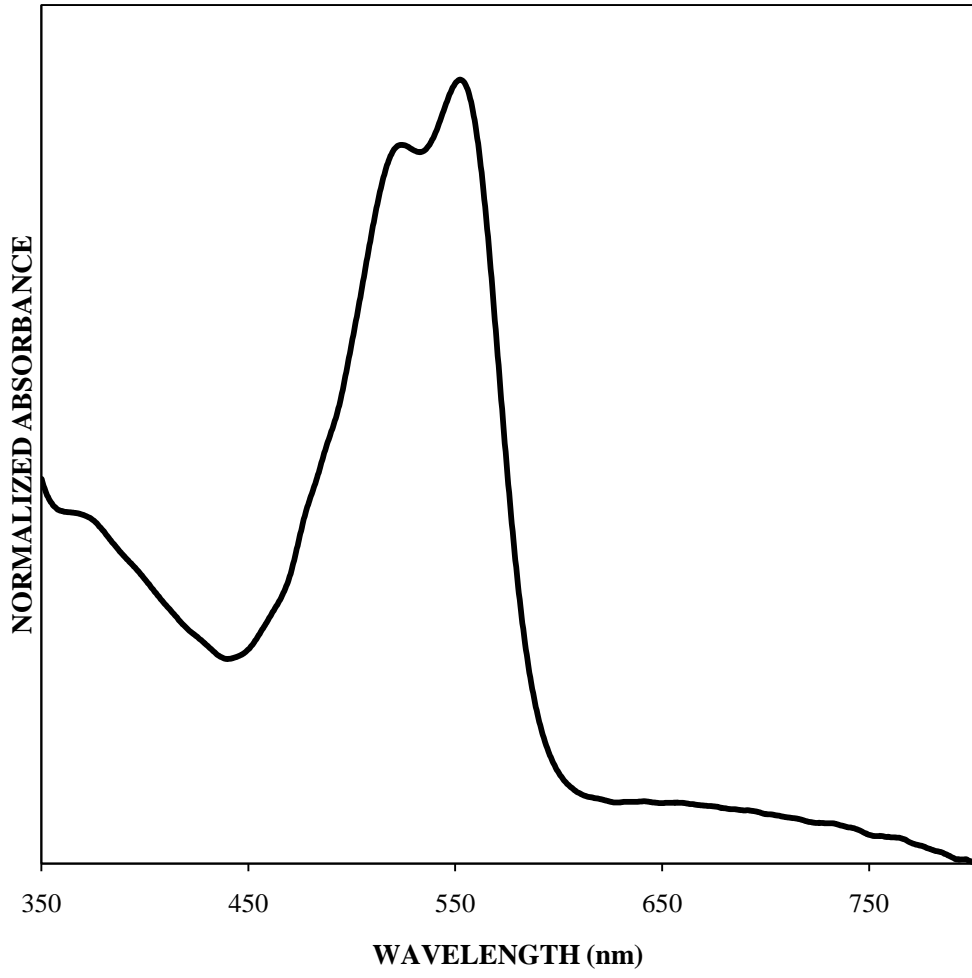


Figure B.20 Fiber 685 (ISP)

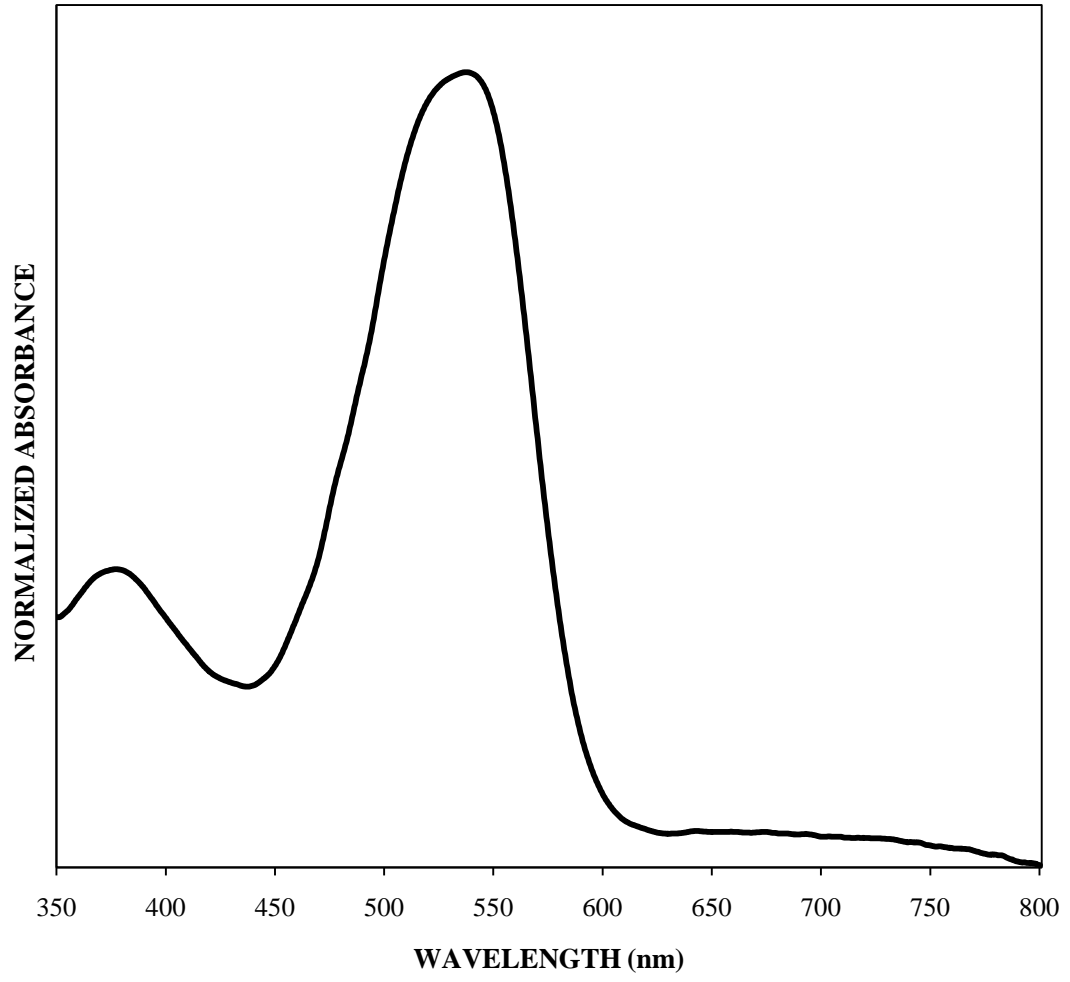


Figure B.21 Fiber 686 (ISP)

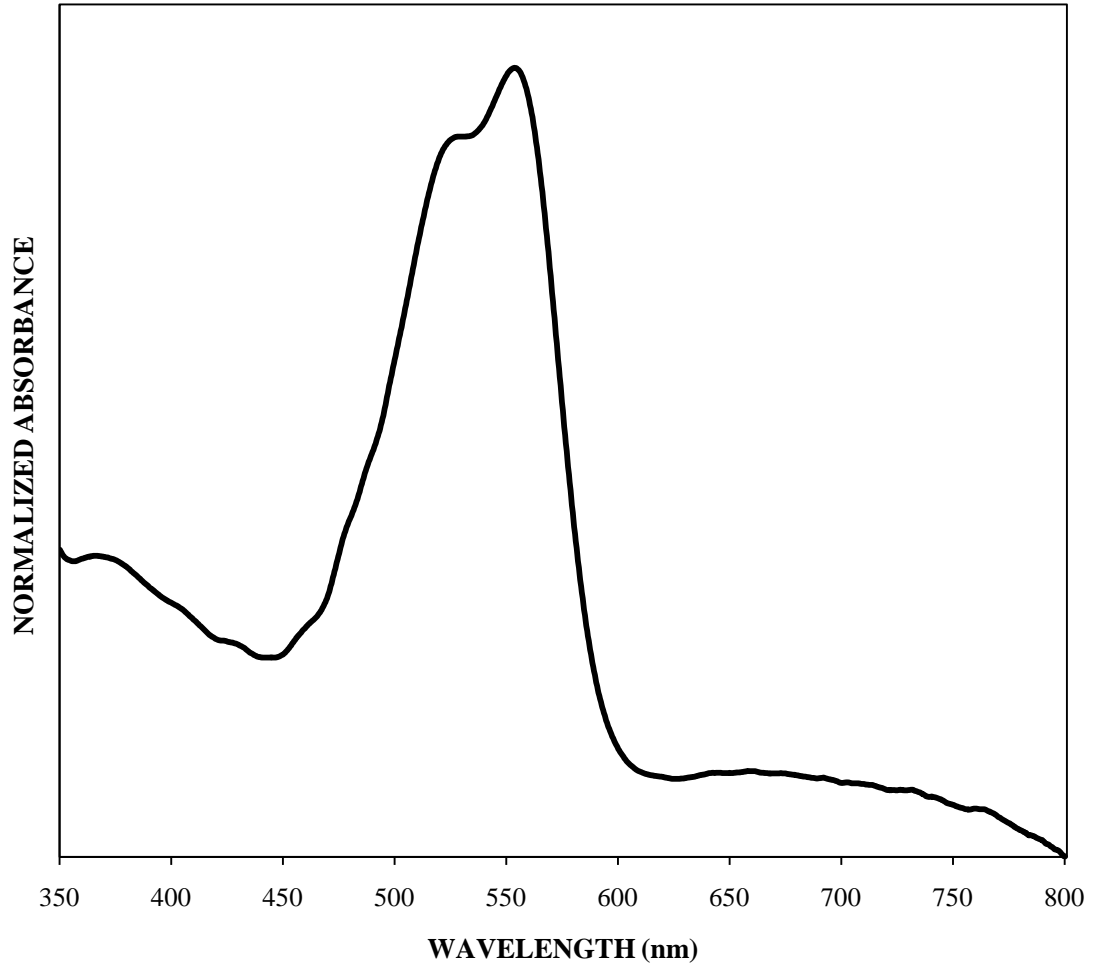


Figure B.22 Fiber 695 (ISP)

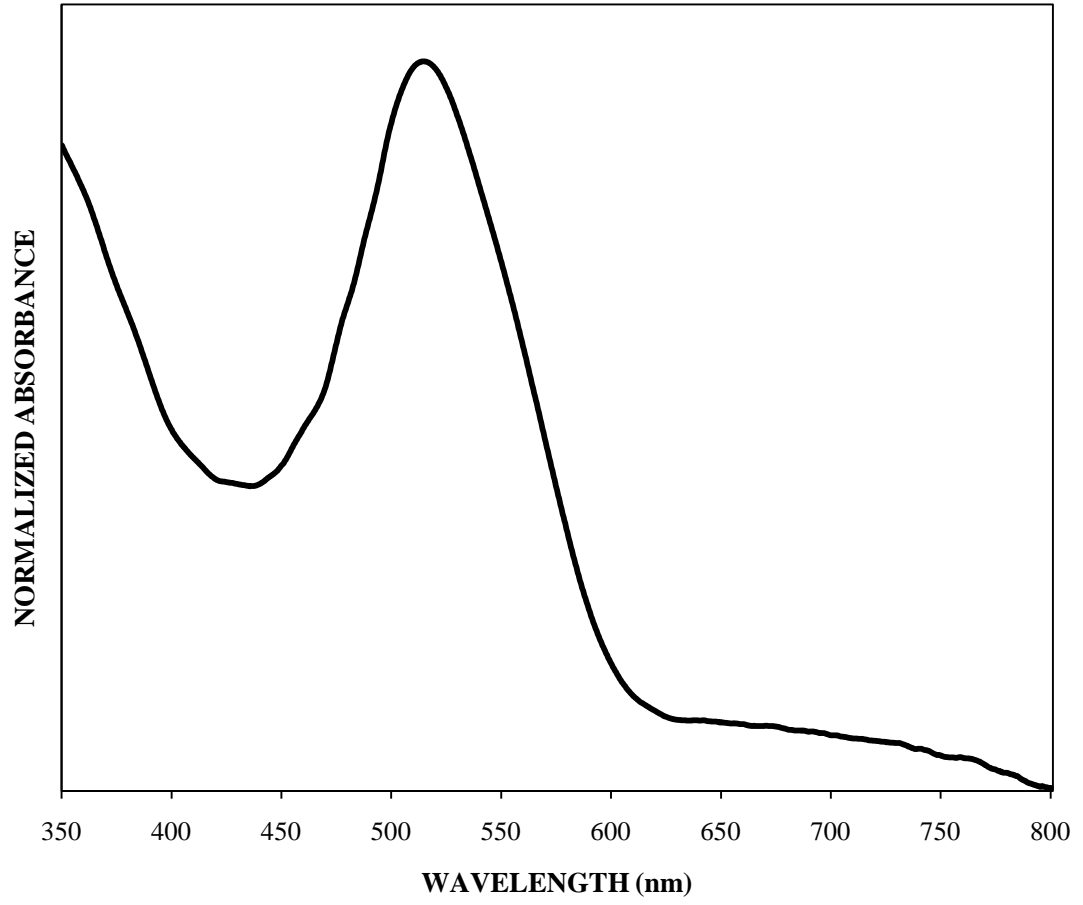


Figure B.23 Fiber 721 (ISP)

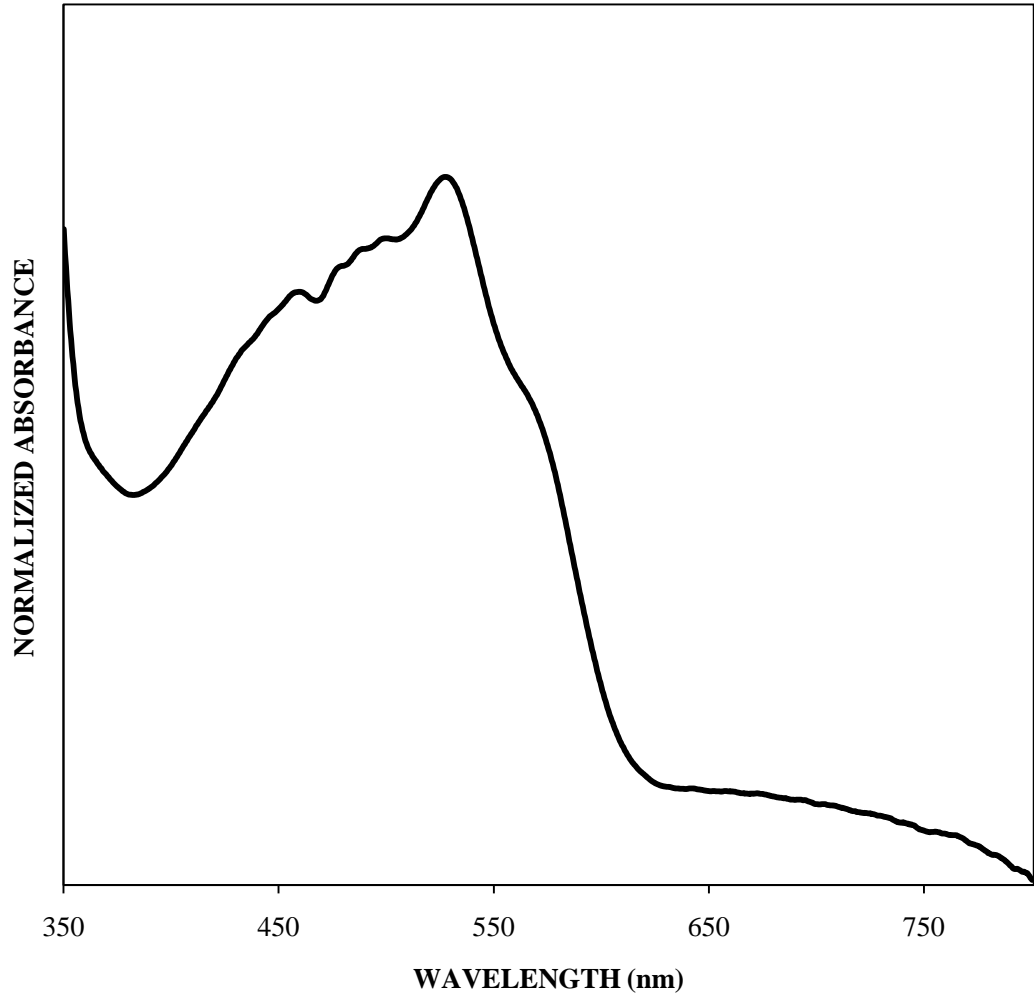


Figure B.24 Fiber 722 (ISP)

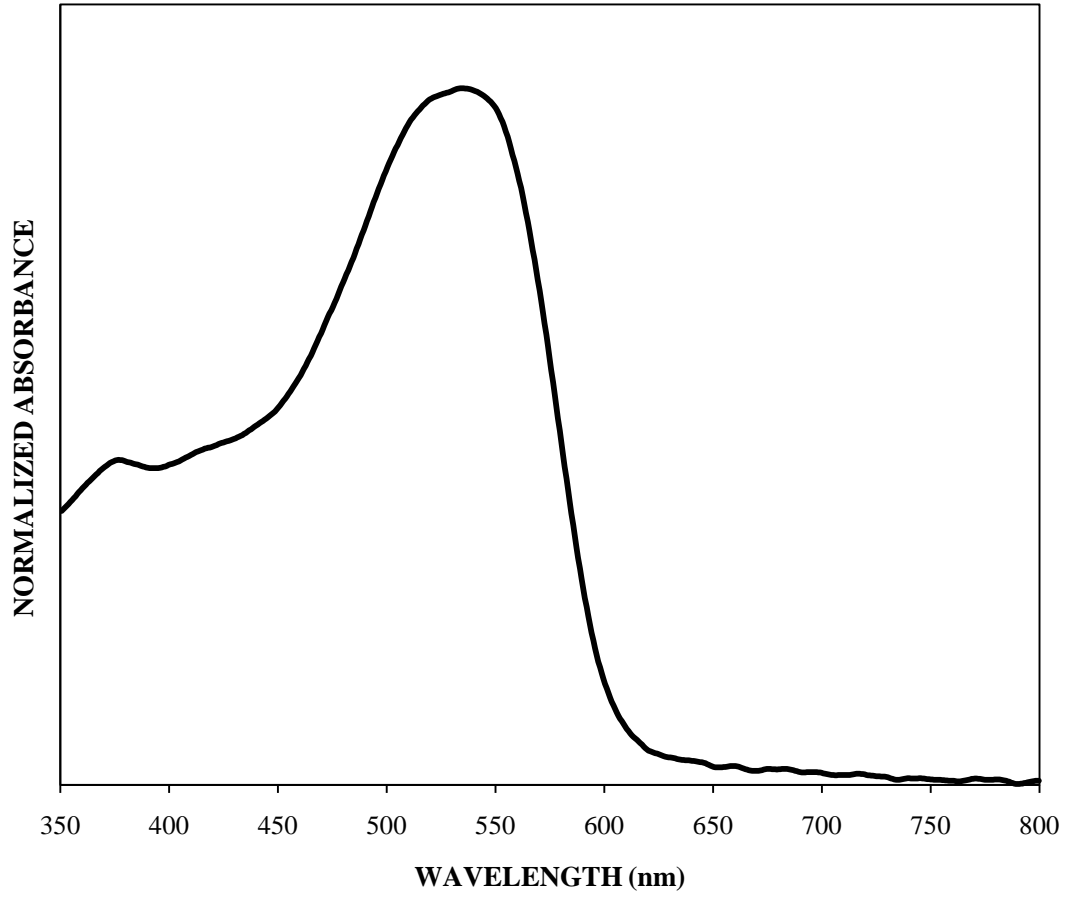


Figure B.25 Fiber A (IUPUI – newest set)

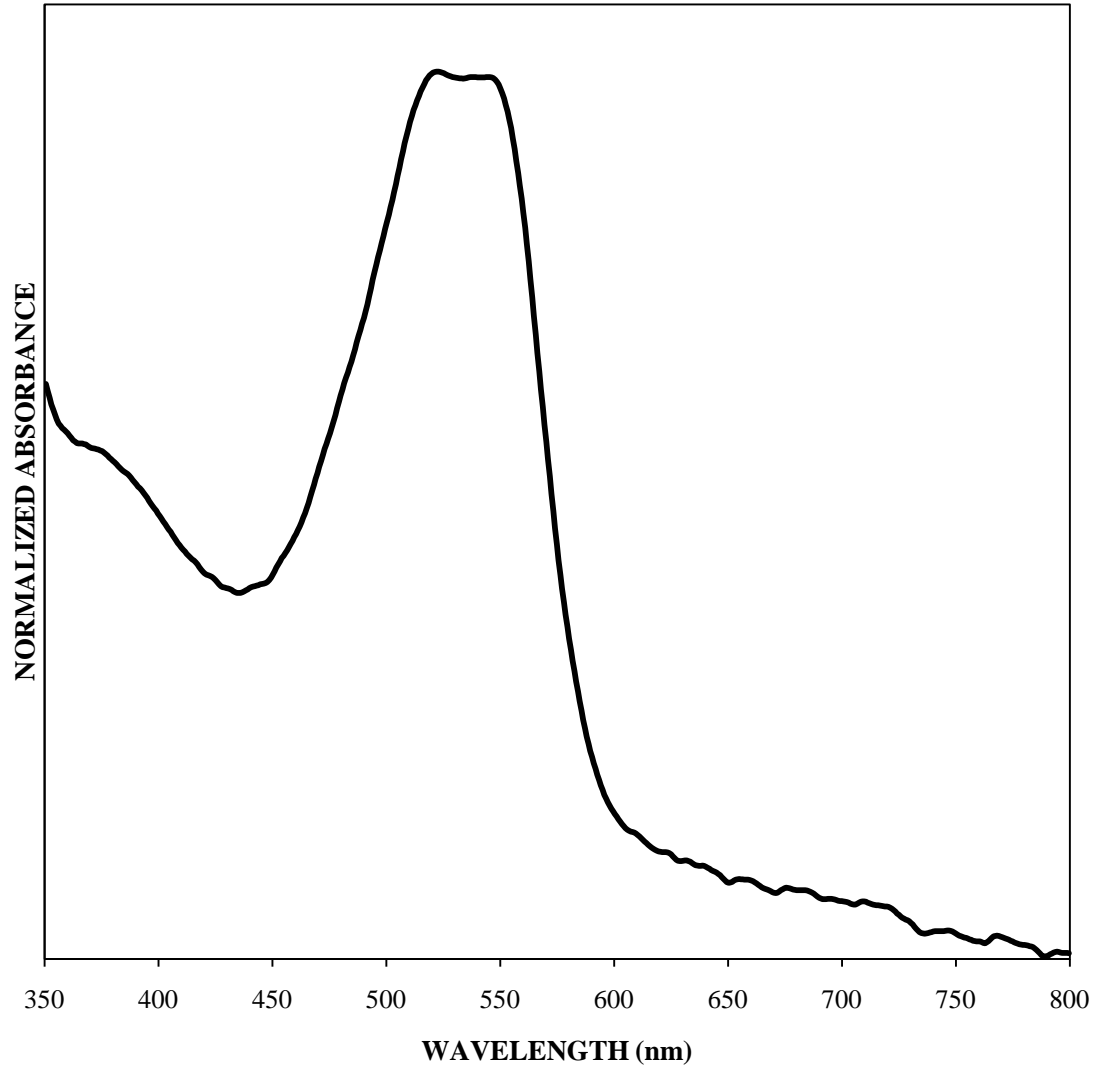


Figure B.26 Fiber B (IUPUI – newest set)

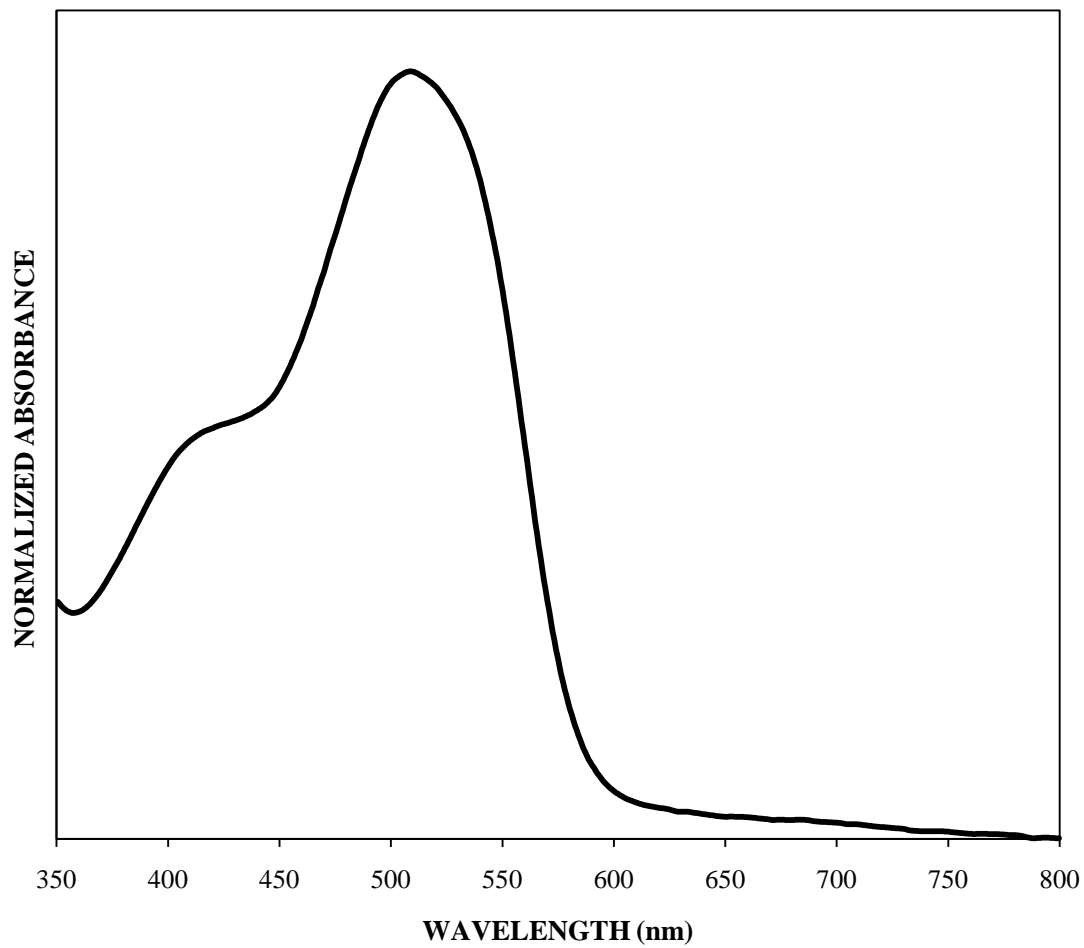


Figure B.27 Fiber C (IUPUI – newest set)

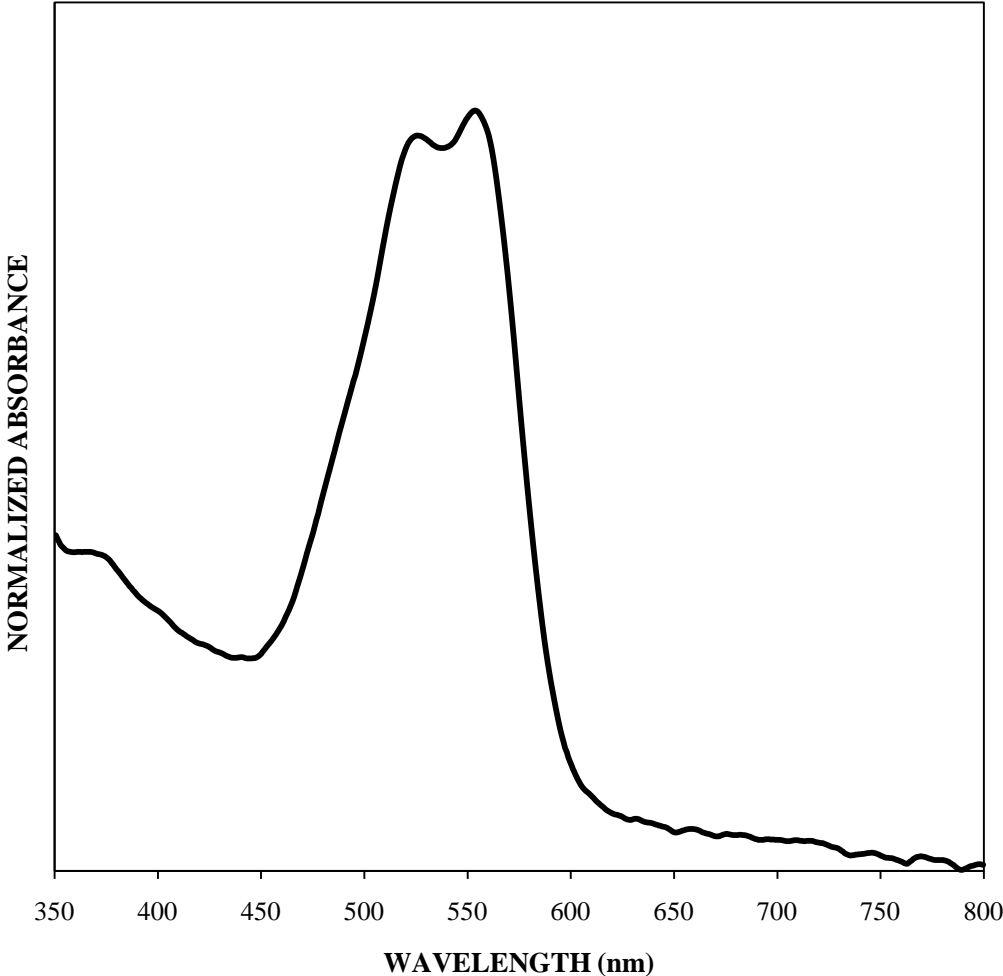


Figure B.28 Fiber D (IUPUI – newest set)

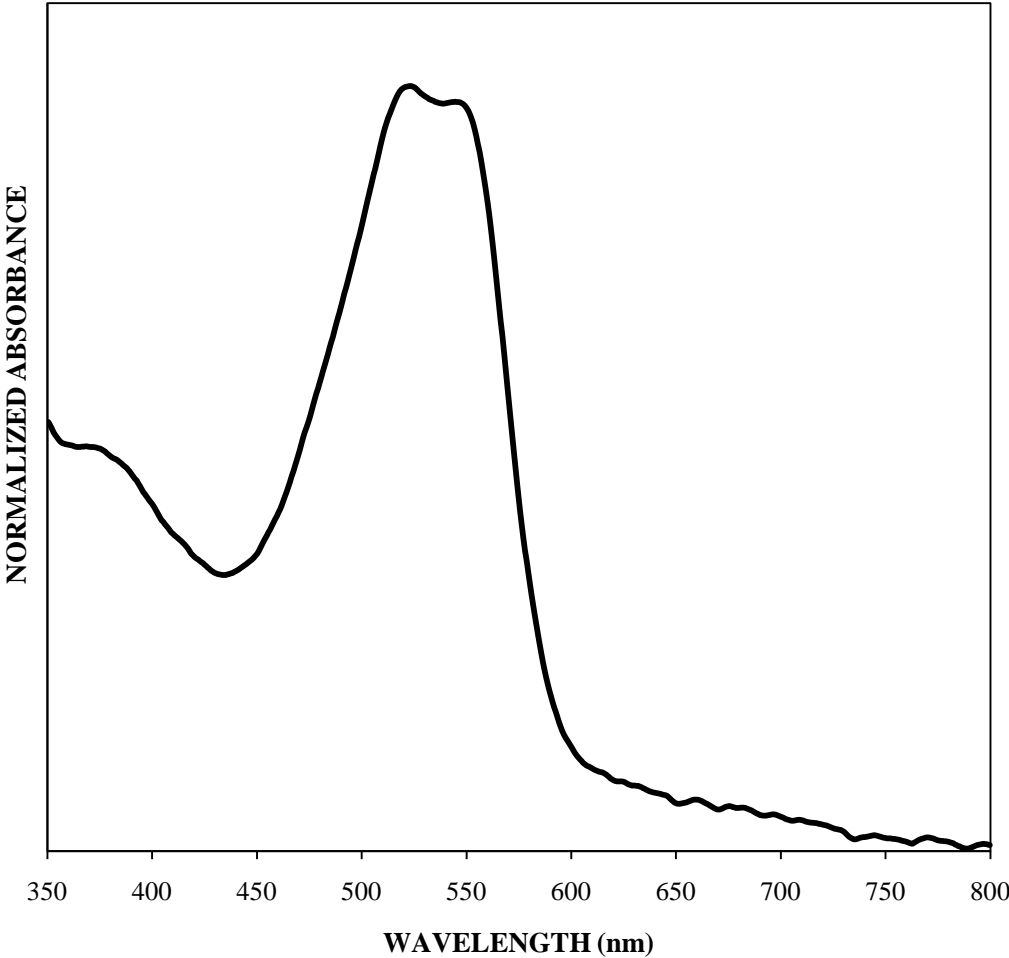


Figure B.29 Fiber E (IUPUI – newest set)

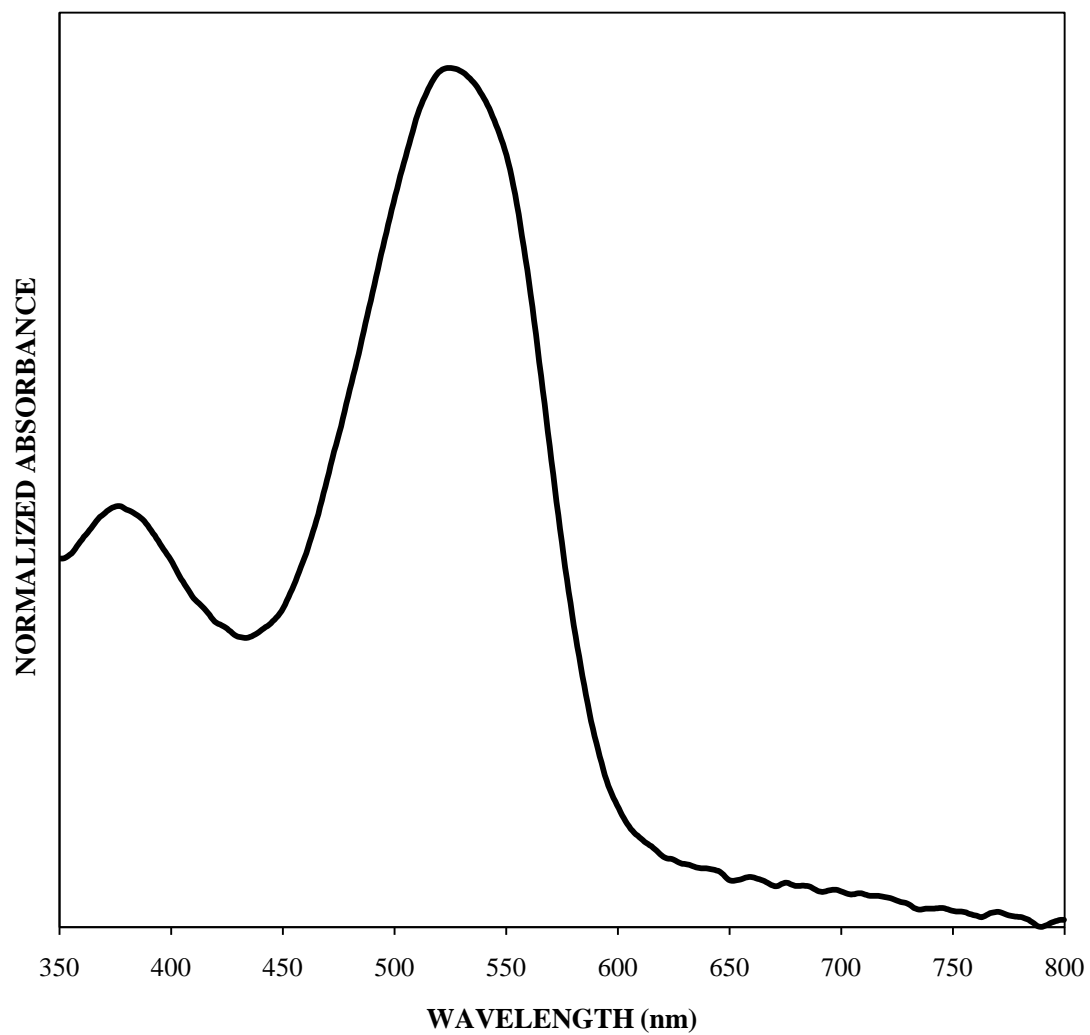


Figure B.30 Fiber F (IUPUI – newest set)

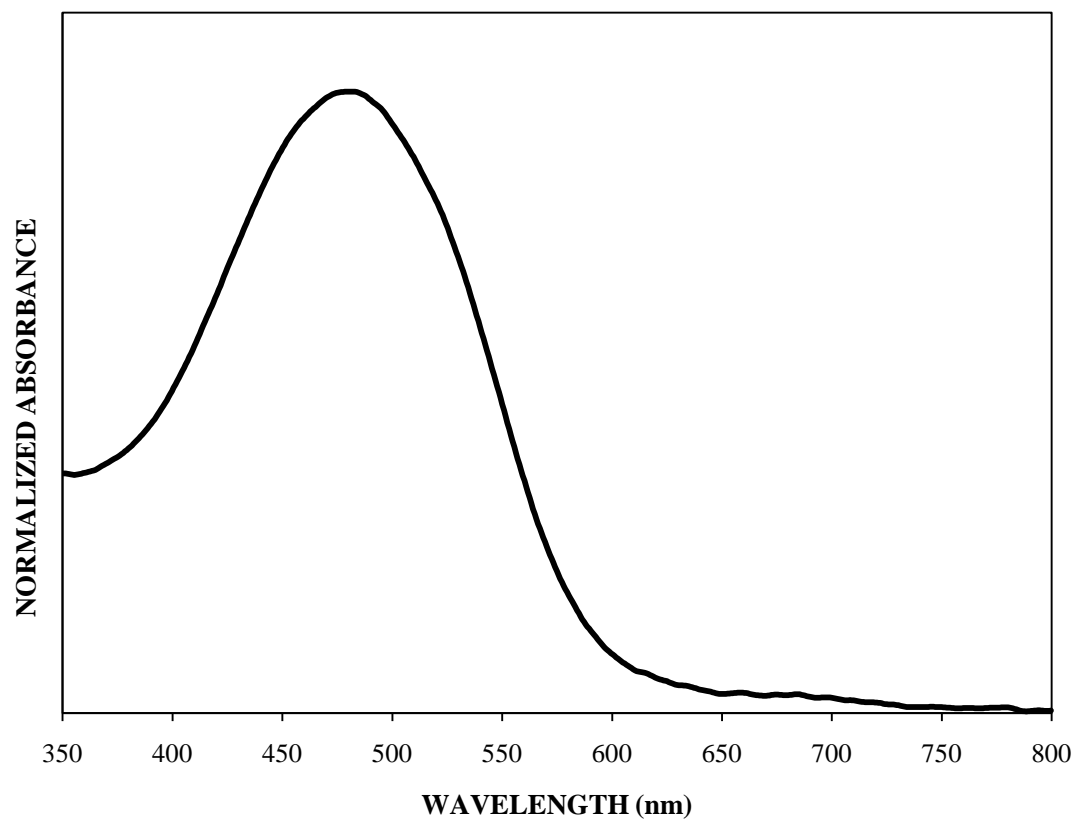


Figure B.31 Fiber 713 (IUPUI – newest set)

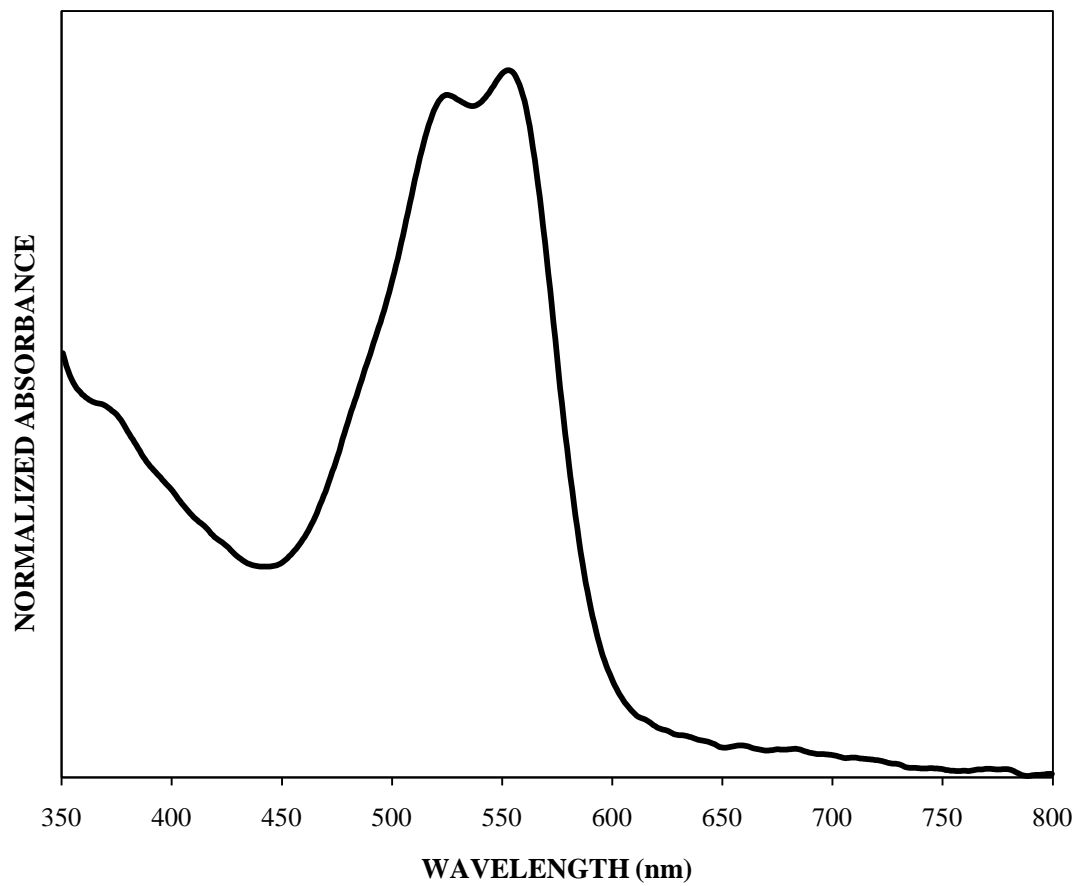


Figure B.32 Fiber 685 (IUPUI – newest set)

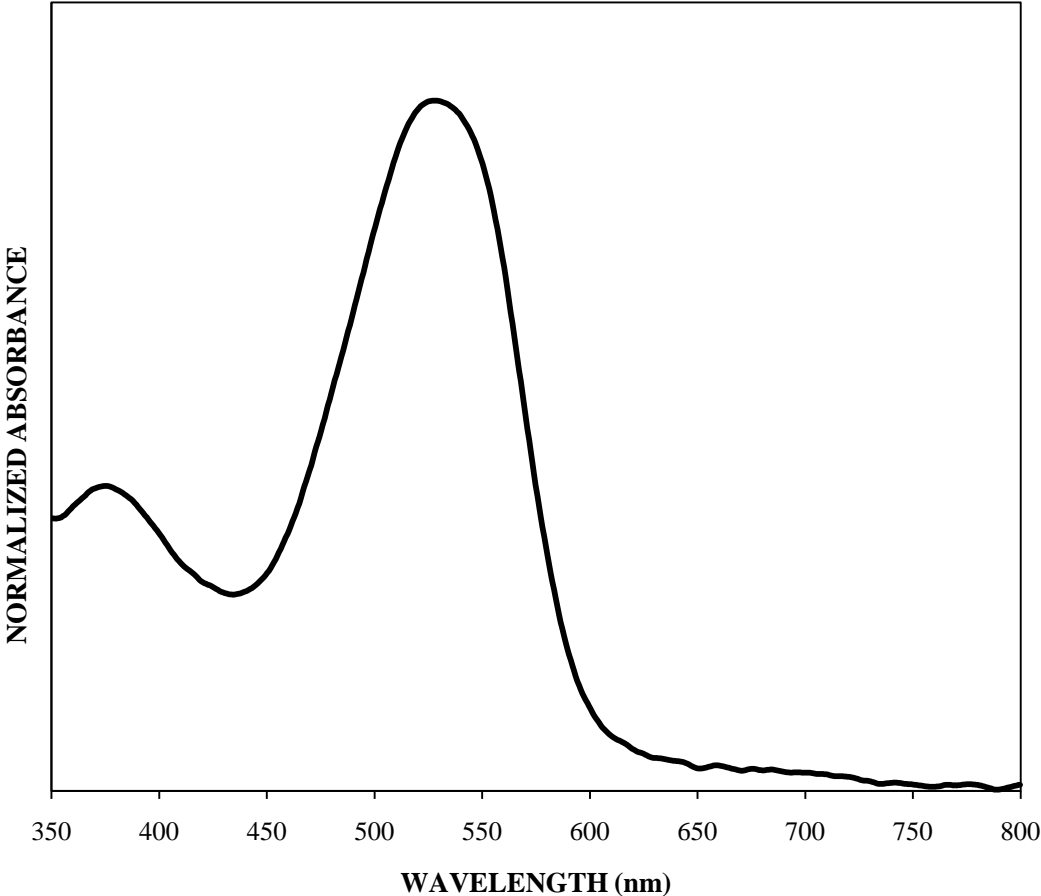


Figure B.33 Fiber 686 (IUPUI – newest set)

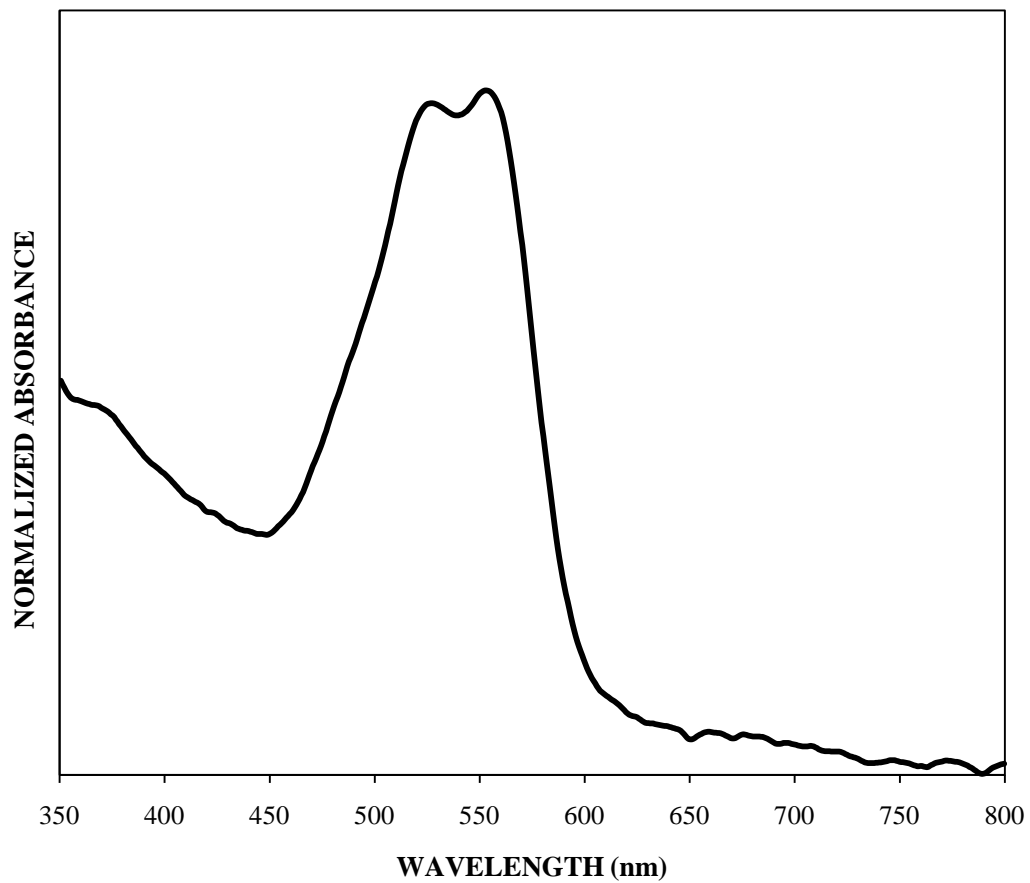


Figure B.34 Fiber 695 (IUPUI – newest set)

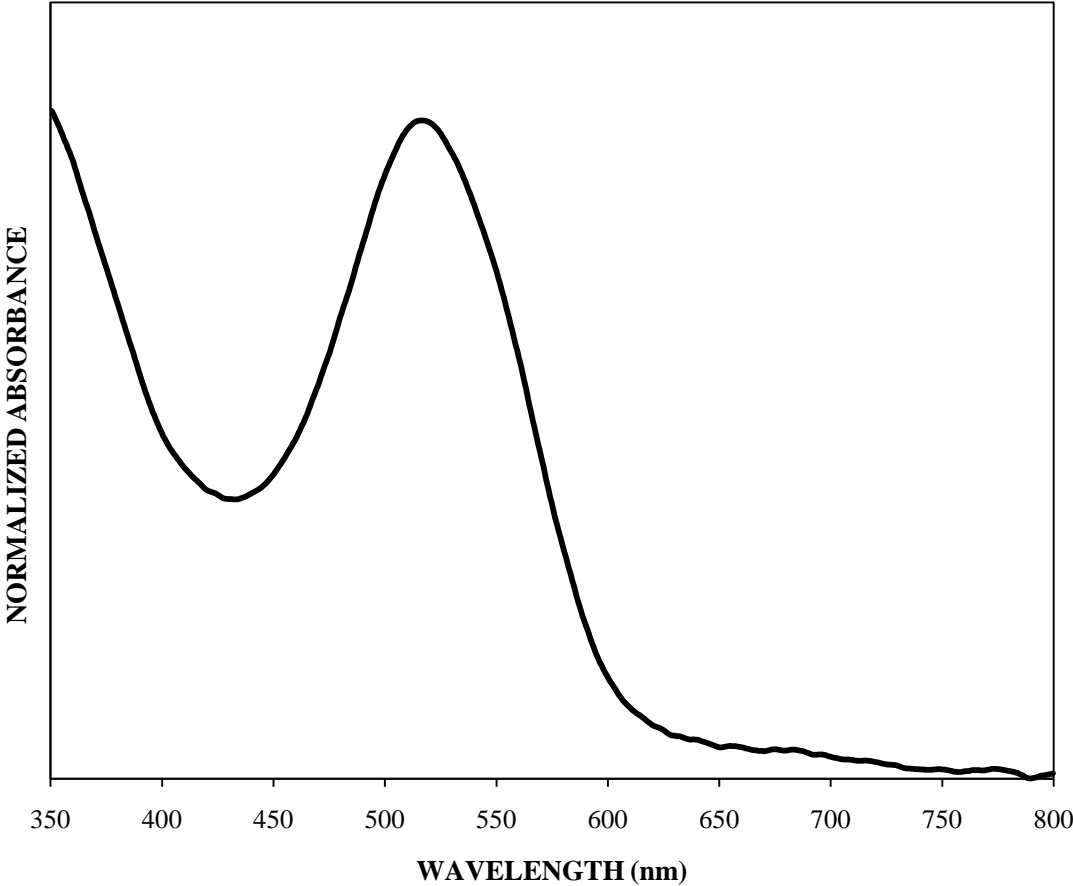


Figure B.35 Fiber 721 (IUPUI – newest set)

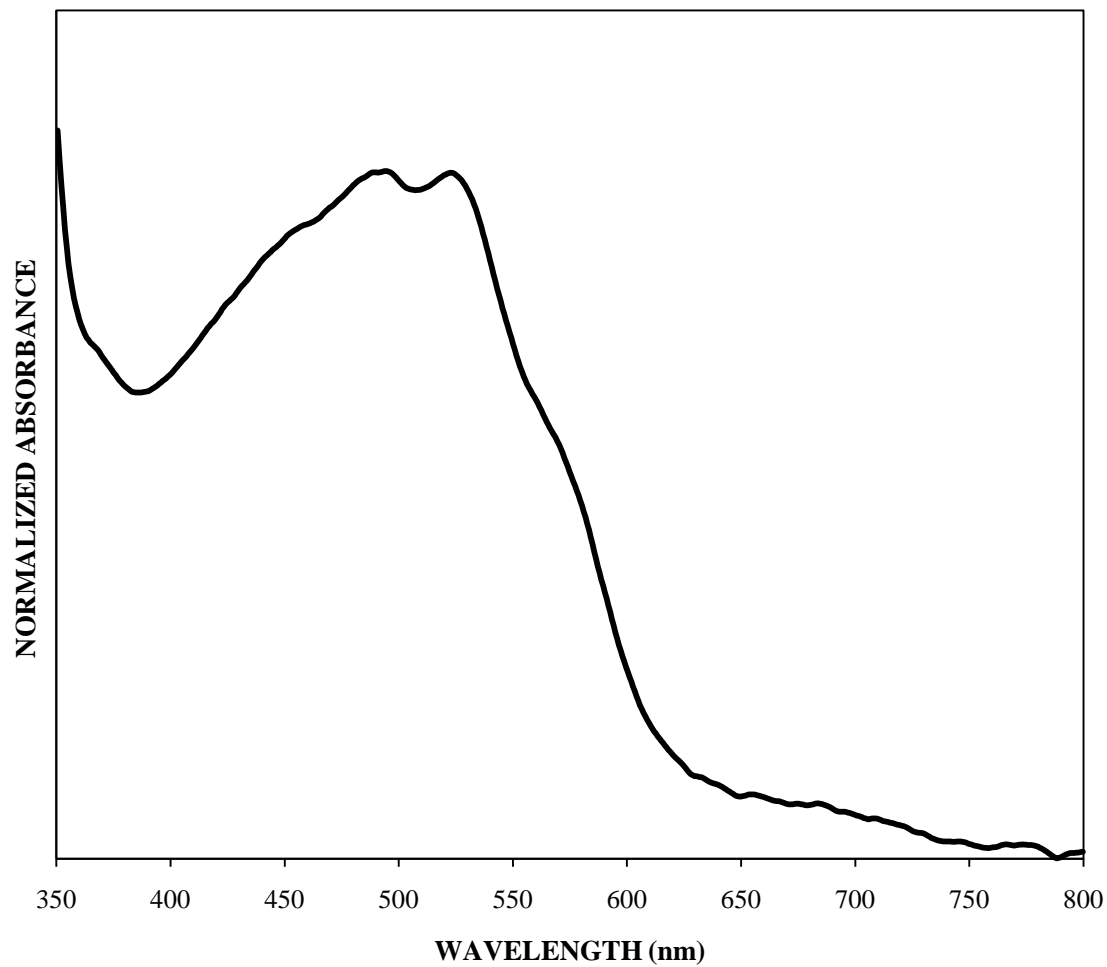


Figure B.36 Fiber 722 (IUPUI – newest set)

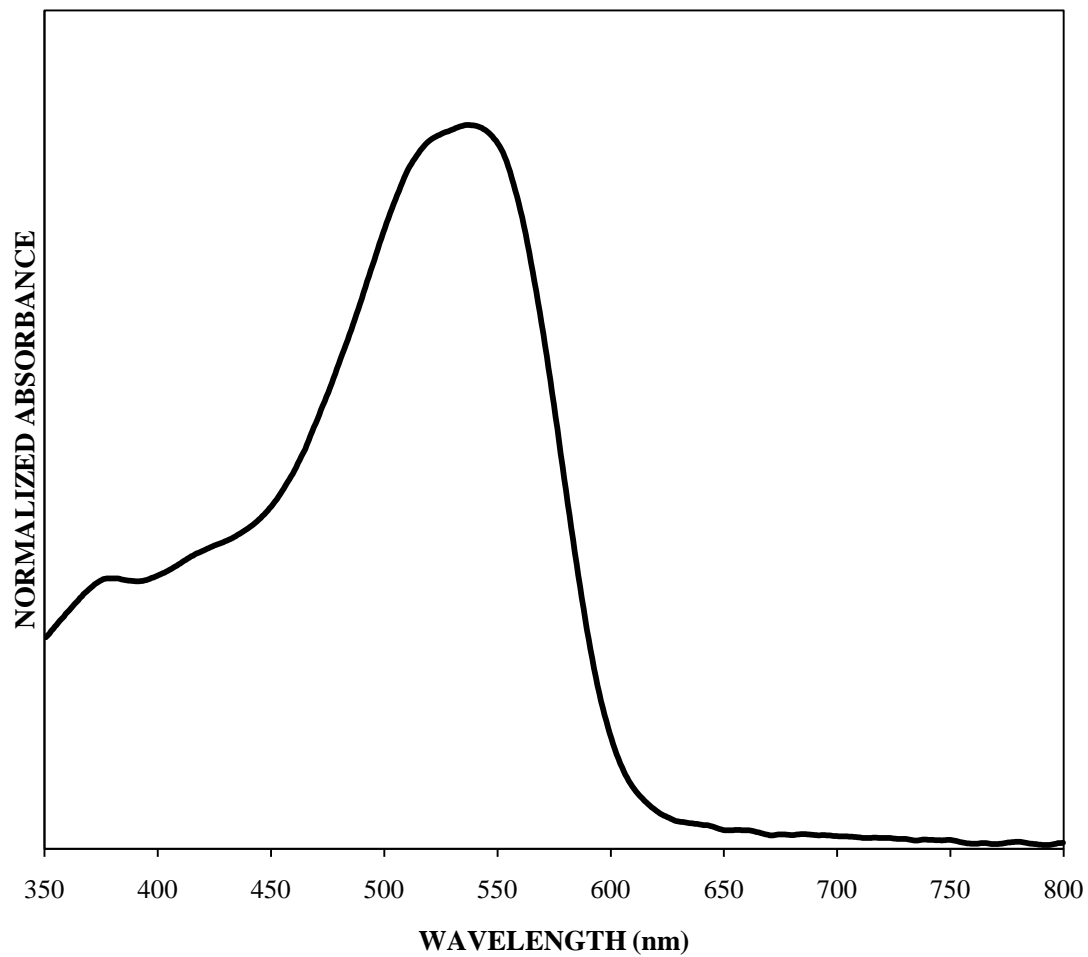


Figure B.37 Fiber A (IUPUI – newest set; supplemental data)

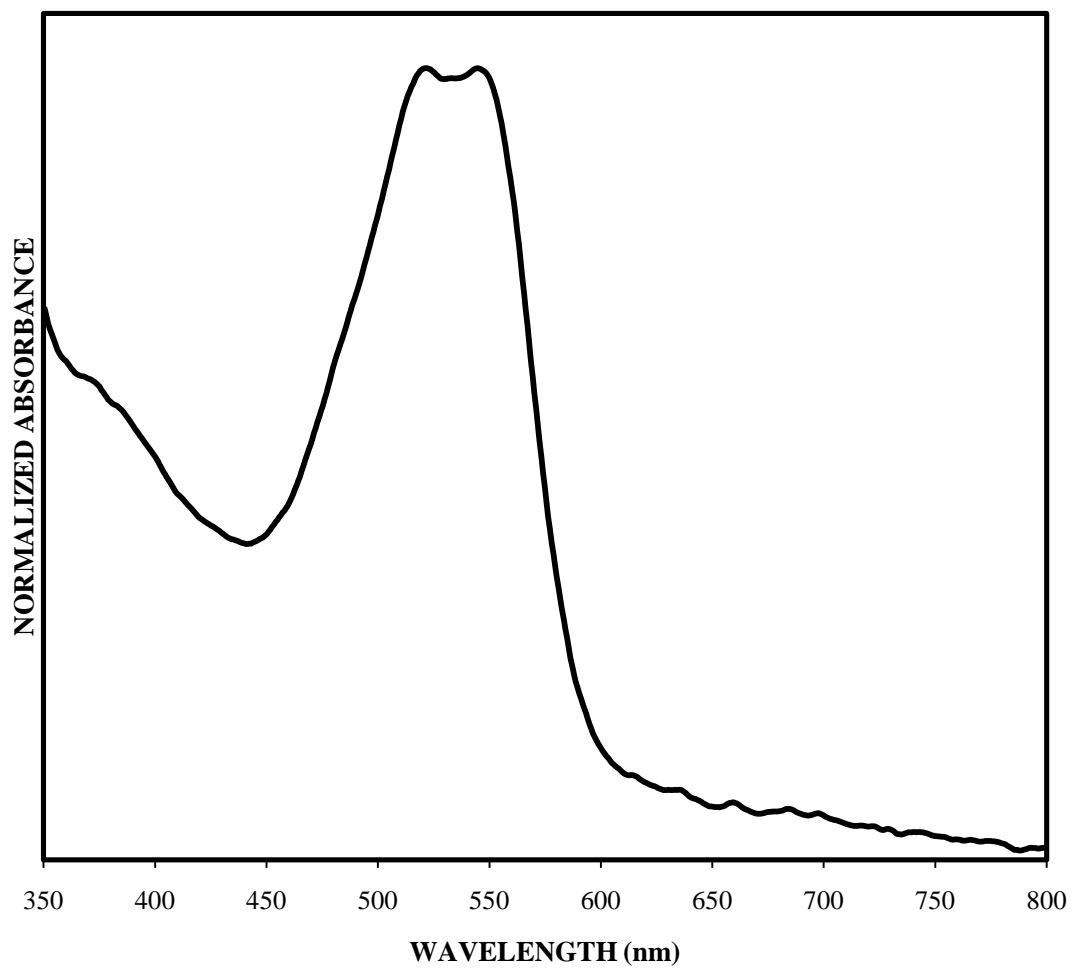


Figure B.38 Fiber B (IUPUI – newest set; supplemental data)

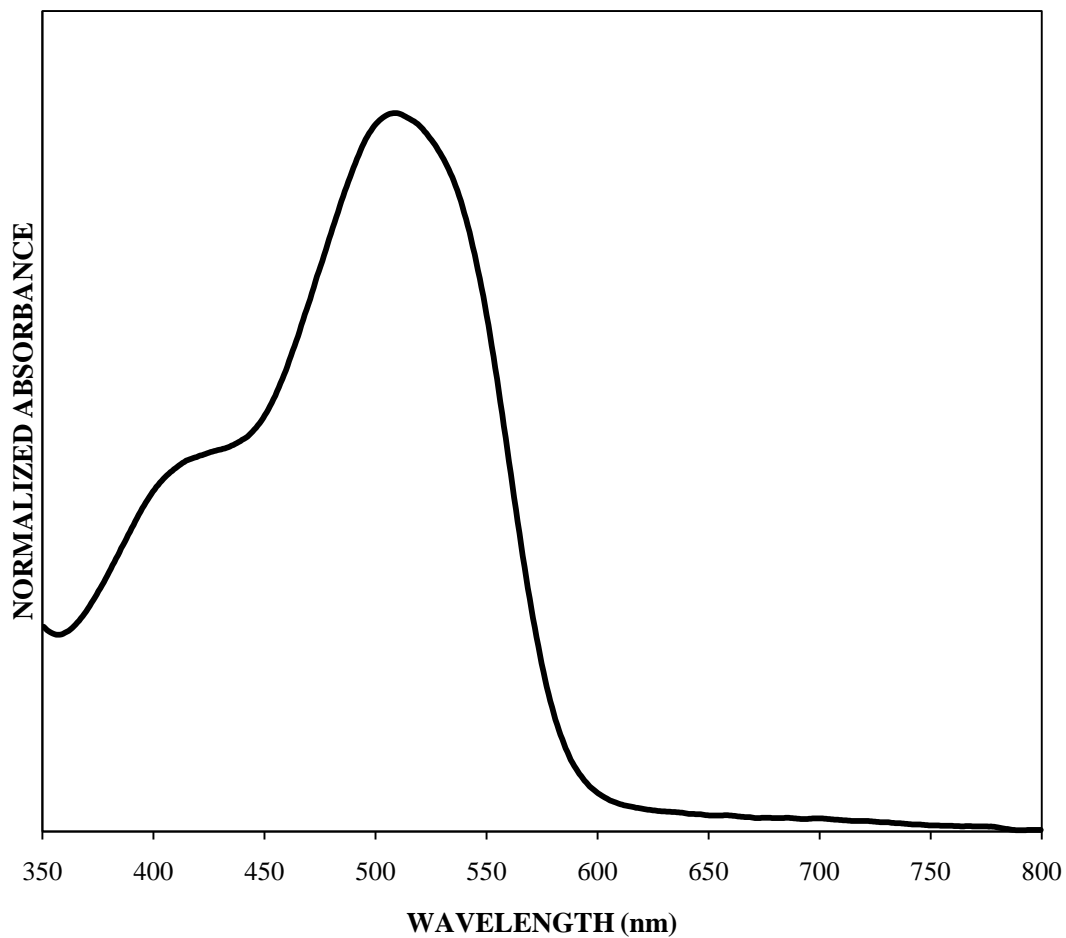


Figure B.39 Fiber C (IUPUI – newest set; supplemental data)

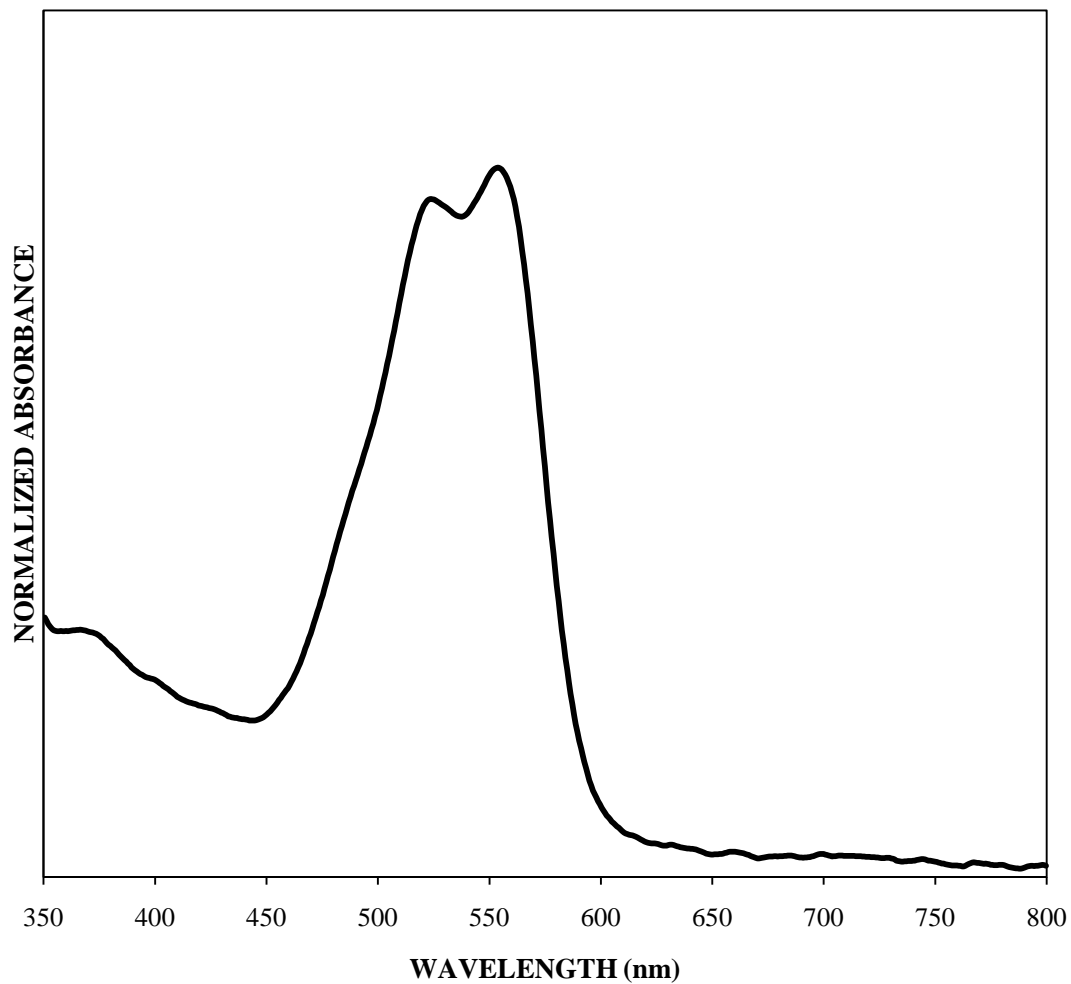


Figure B.40 Fiber D (IUPUI – newest set; supplemental data)

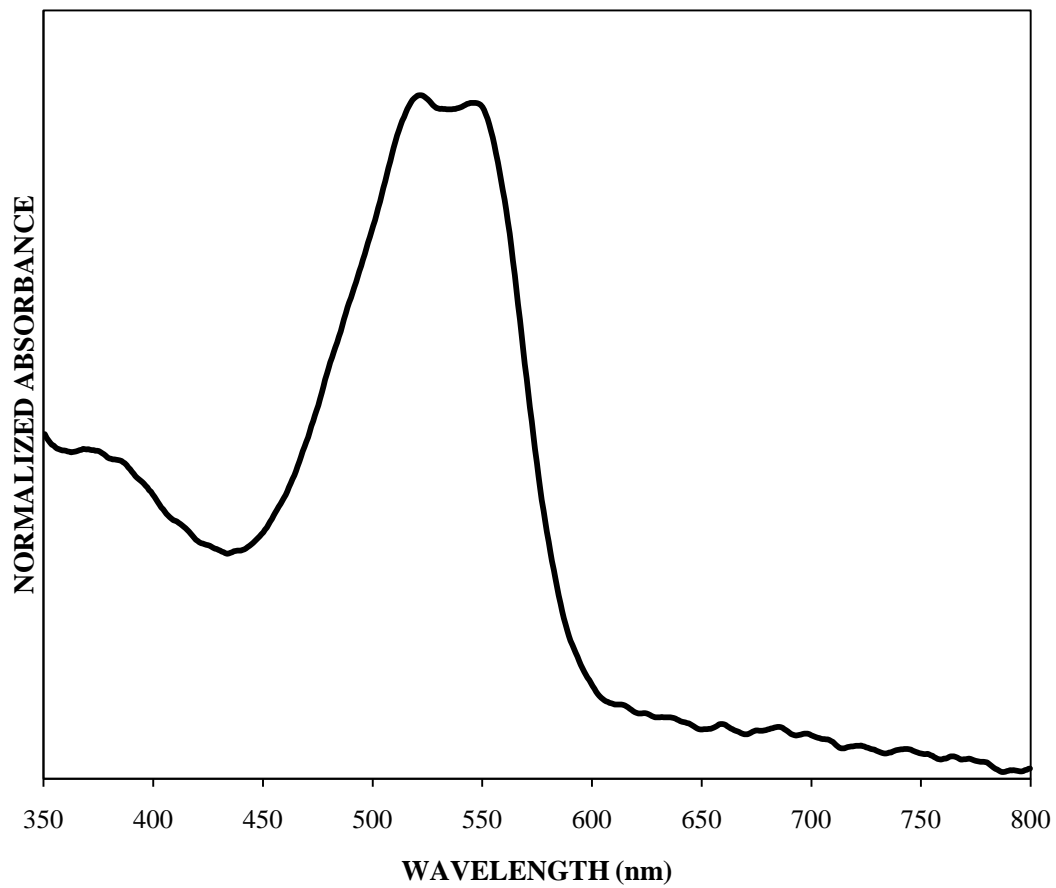


Figure B.41 Fiber E (IUPUI – newest set; supplemental data)

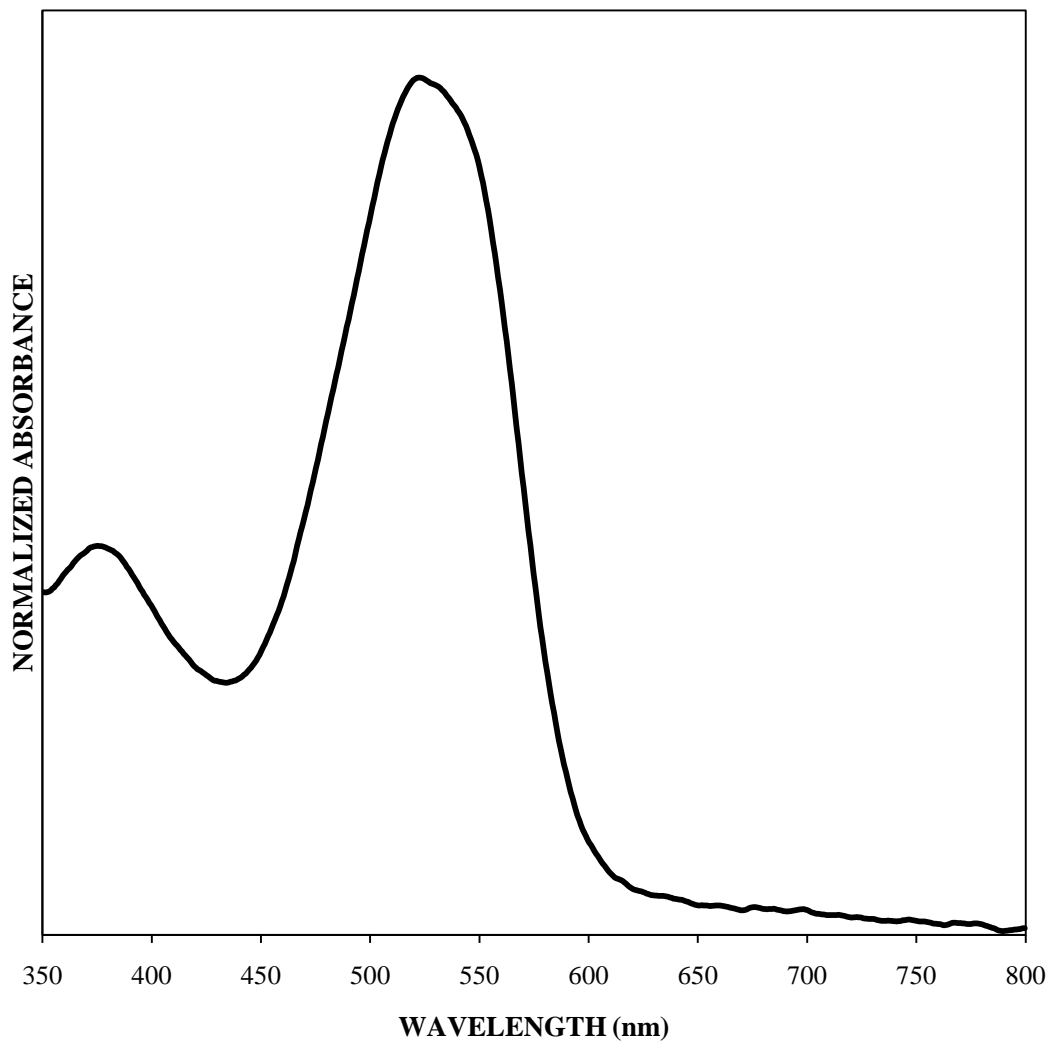


Figure B.42 Fiber F (IUPUI – newest set; supplemental data)

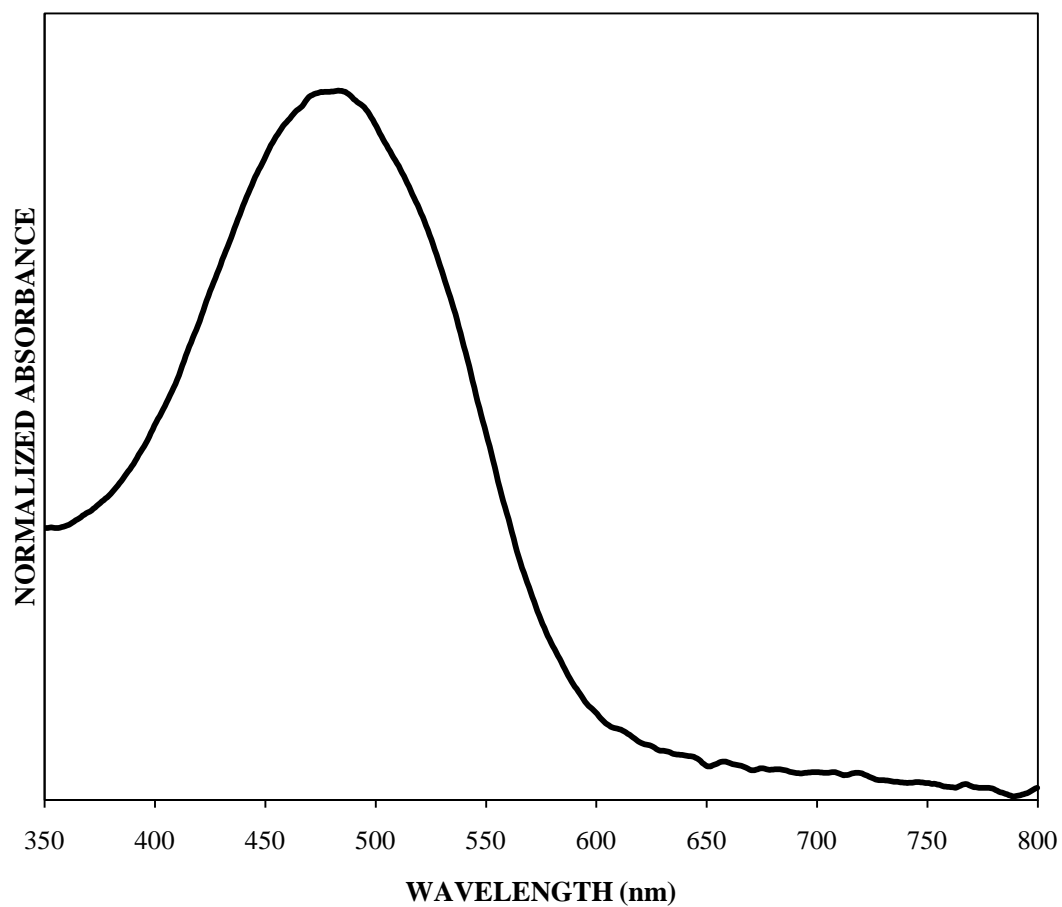


Figure B.43 Fiber 713 (IUPUI – newest set; supplemental data)

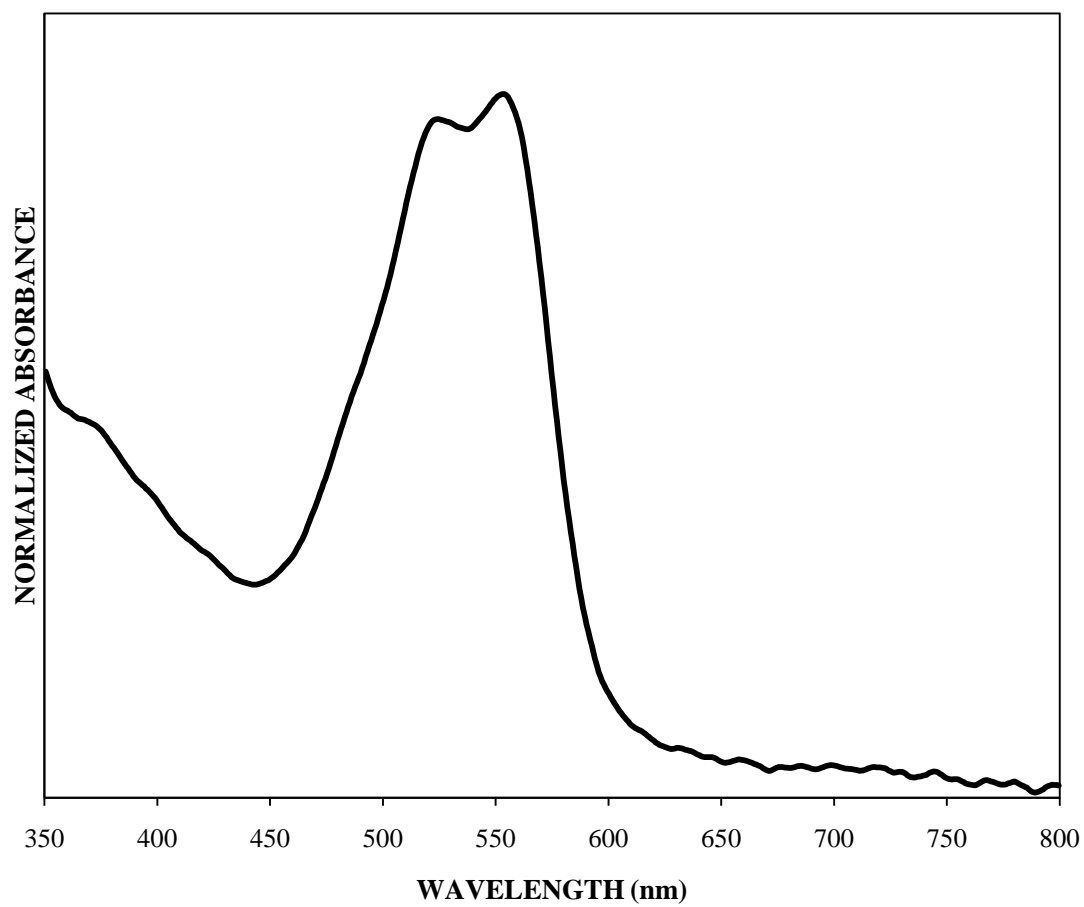


Figure B.44 Fiber 685 (IUPUI – newest set; supplemental data)

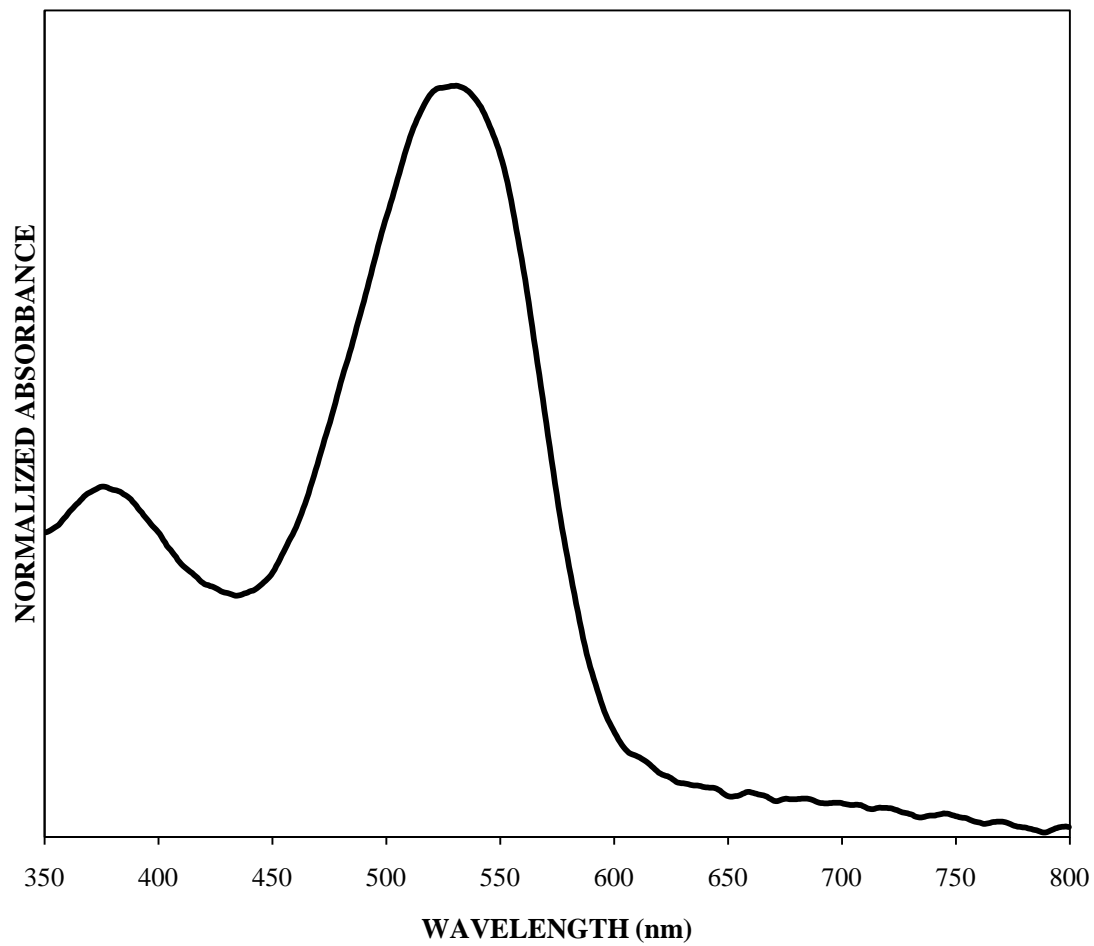


Figure B.45 Fiber 686 (IUPUI – newest set; supplemental data)

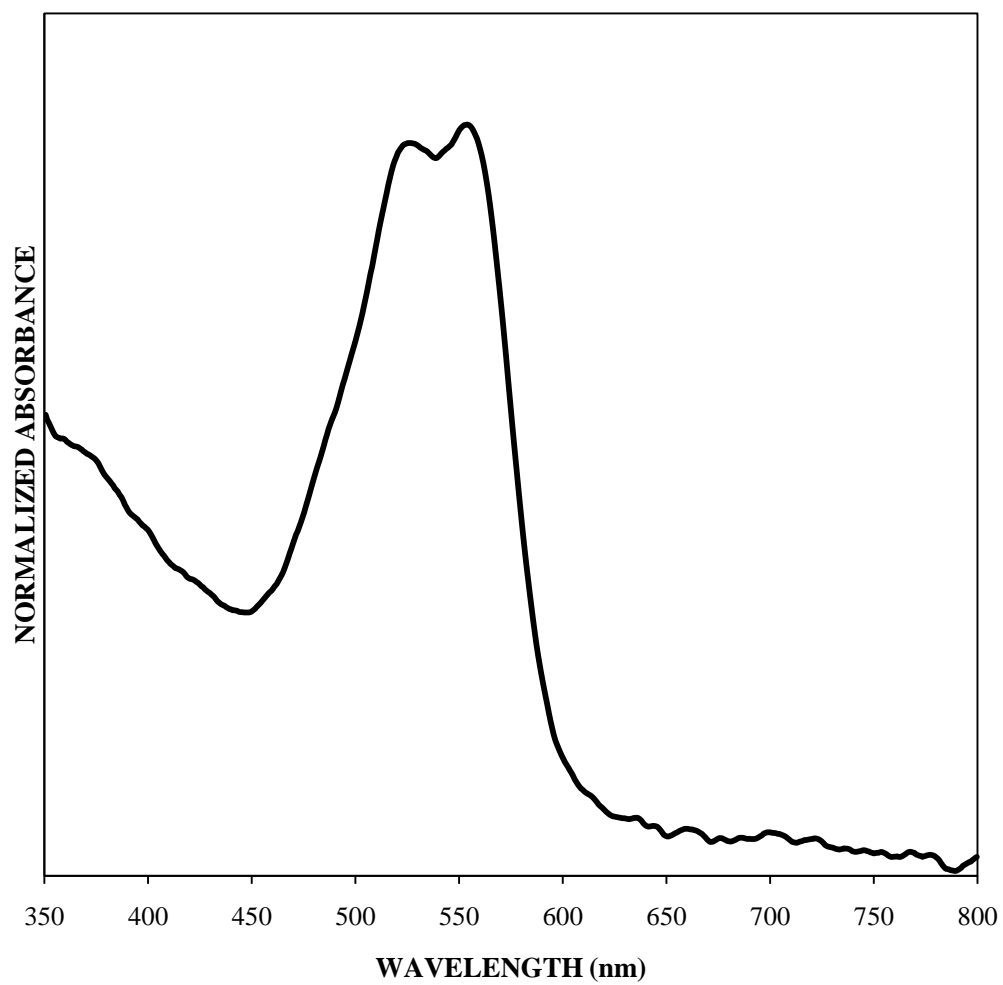


Figure B.46 Fiber 695 (IUPUI – newest set; supplemental data)

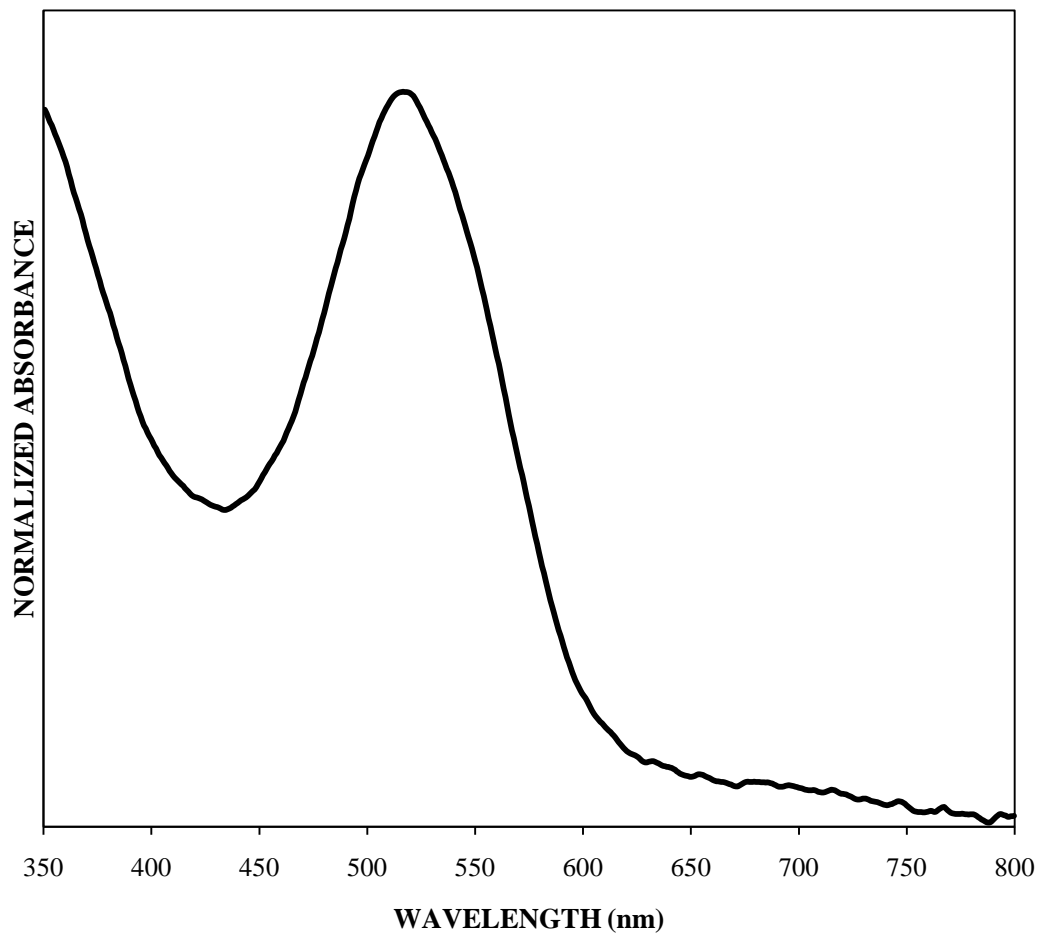


Figure B.47 Fiber 721 (IUPUI – newest set; supplemental data)

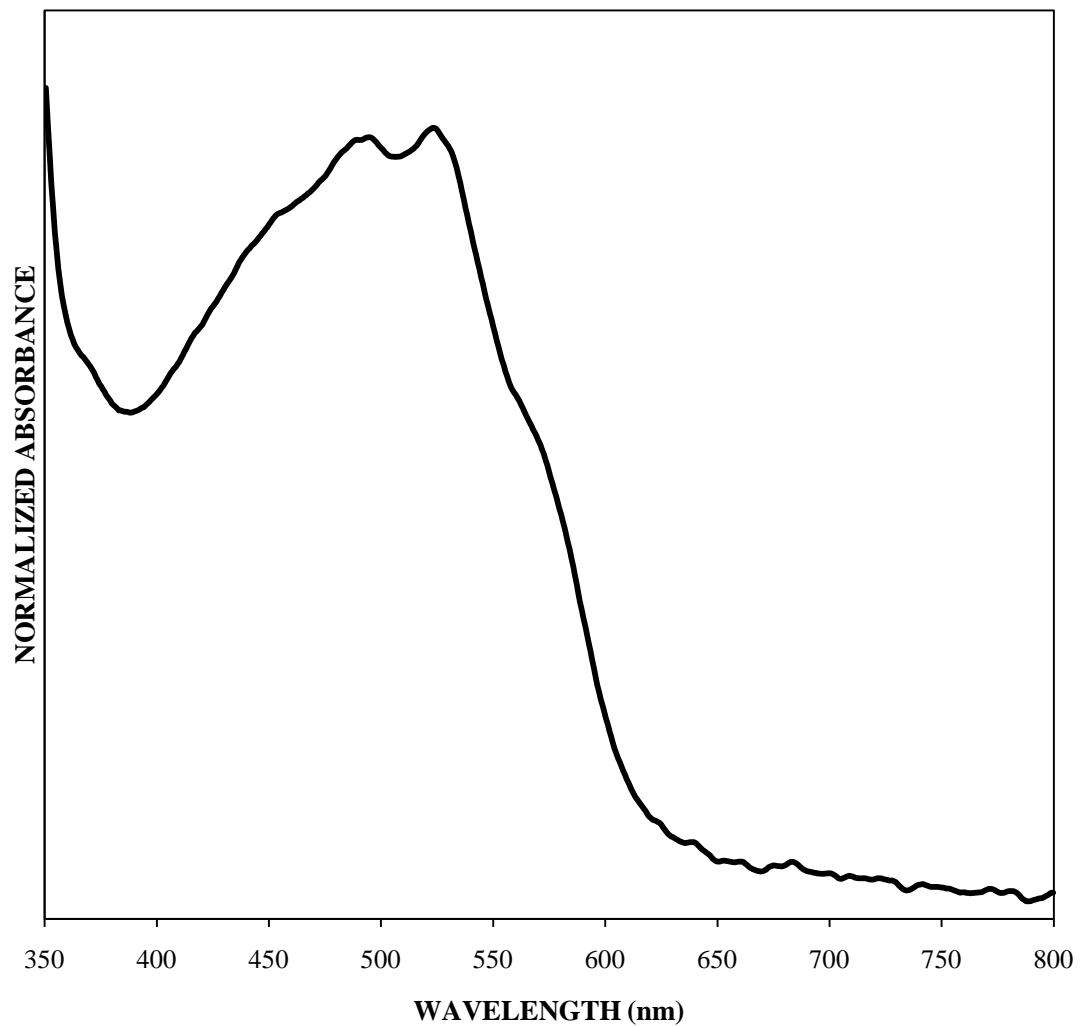


Figure B.48 Fiber 722 (IUPUI – newest set; supplemental data)

Appendix C. Polymer Standards

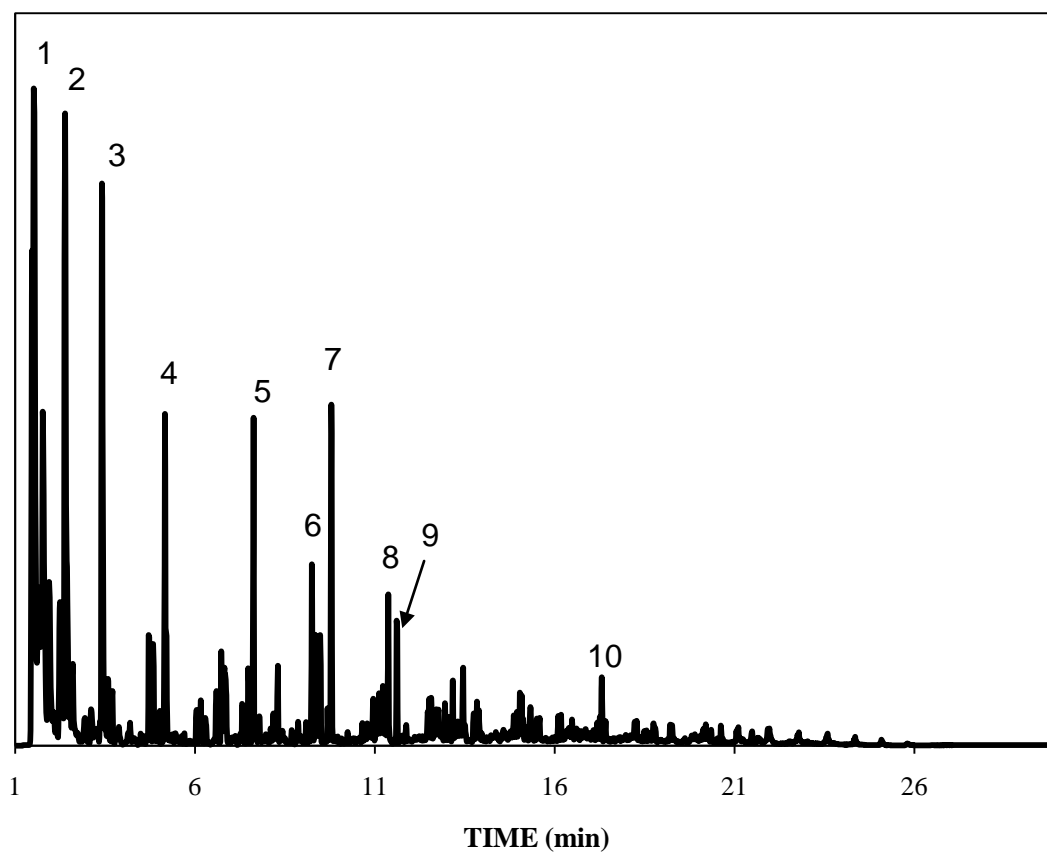


Figure C.1 Chromatogram of Polyethylene (25% chlorine); **(1)** methylenecyclopropane; **(2)** benzene; **(3)** toluene; **(4)** benzocyclobutene; **(5)** indene; **(6)** naphthalene; **(7)** 3-methyl-3-phenyl-cyclopropene; **(8)** naphthalene, 1-methyl-; **(9)** naphthalene, 1-methyl-; **(10)** anthracene

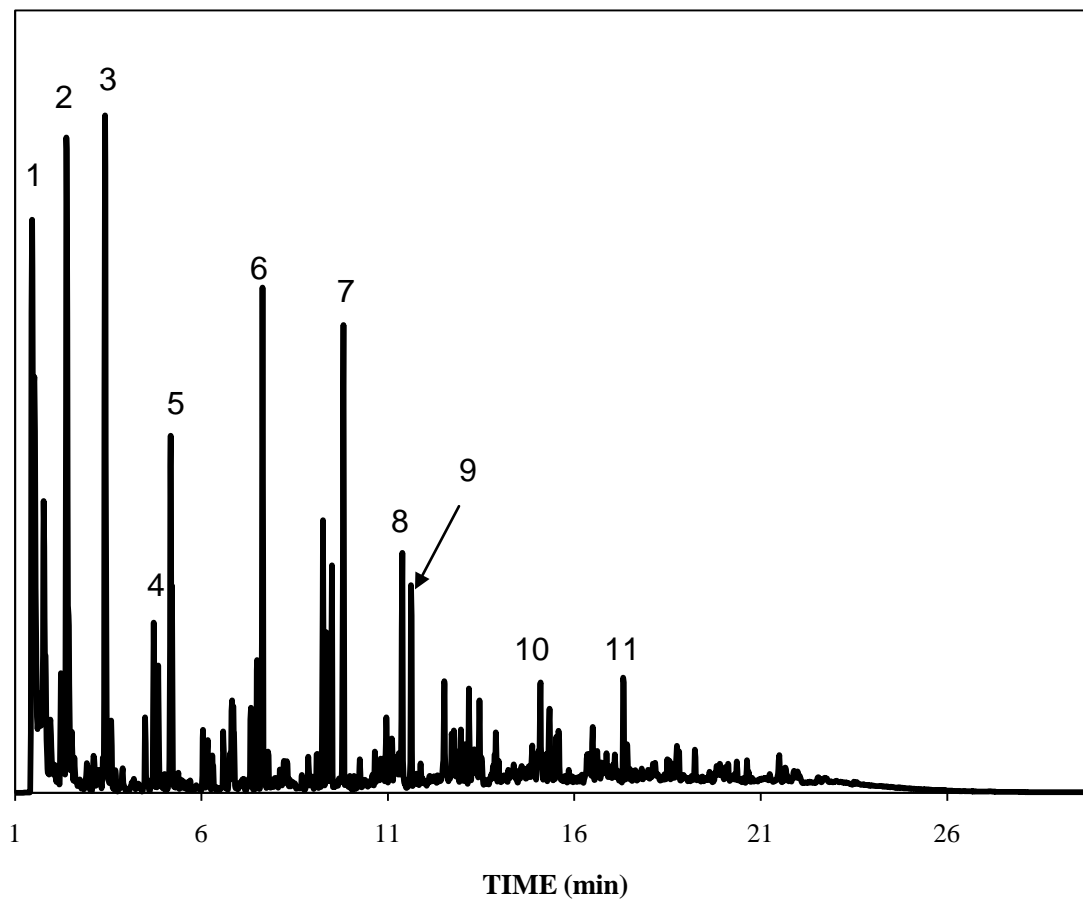


Figure C.2 Chromatogram of Polyethylene (48% chlorine); **(1)** hydrogen chloride; **(2)** benzene; **(3)** toluene; **(4)** ethylbenzene; **(5)** styrene; **(6)** indene; **(7)** naphthalene; **(8)** naphthalene, 1-methyl-; **(9)** naphthalene, 1-methyl-; **(10)** fluorene; **(11)** phenanthrene

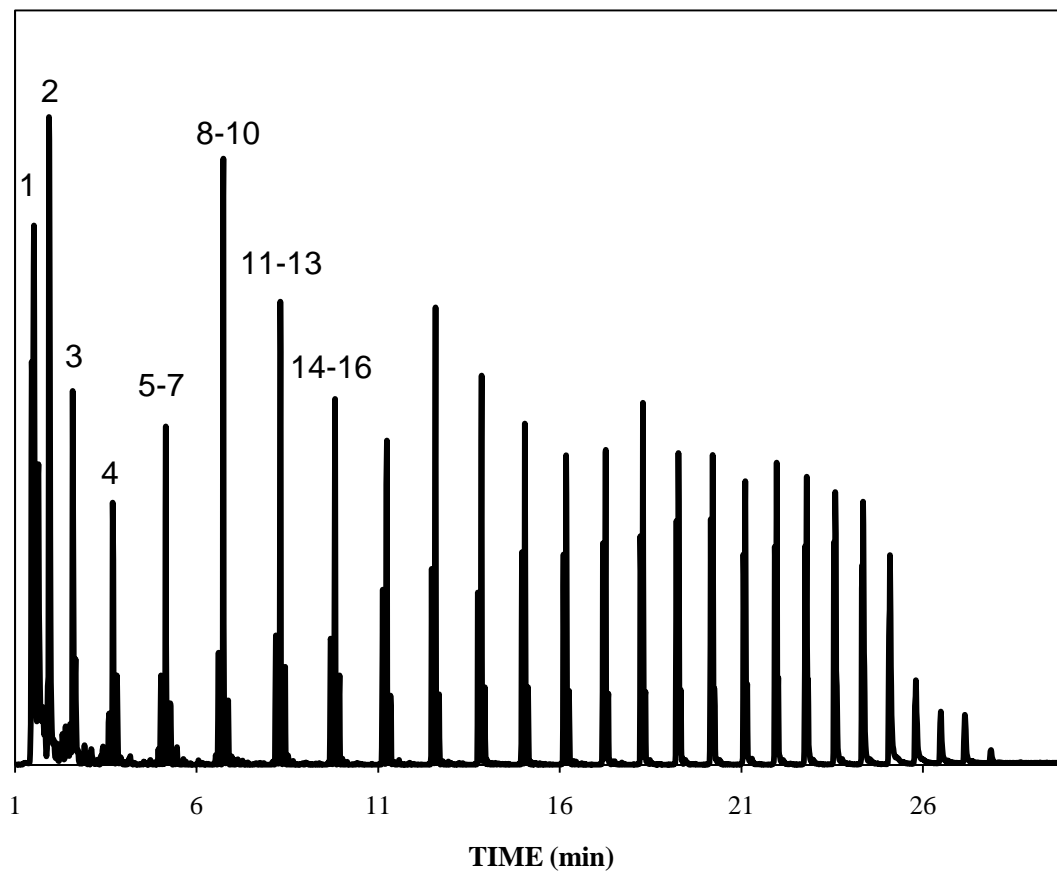


Figure C.3 Chromatogram of Polyethylene, high density; **(1)** 2-Butene, (E)-; **(2)** 1-hexene; **(3)** 1-heptene; **(4)** 1-octene; **(5)** 1,8-nonadiene; **(6)** 1-nonene; **(7)** nonane; **(8)** 1,9-decadiene; **(9)** 1-decene; **(10)** decane; **(11)** 1,10-undecadiene; **(12)** 1-undecene; **(13)** undecane; **(14)** 1,11-dodecadiene; **(15)** 1-dodecene; **(16)** dodecane

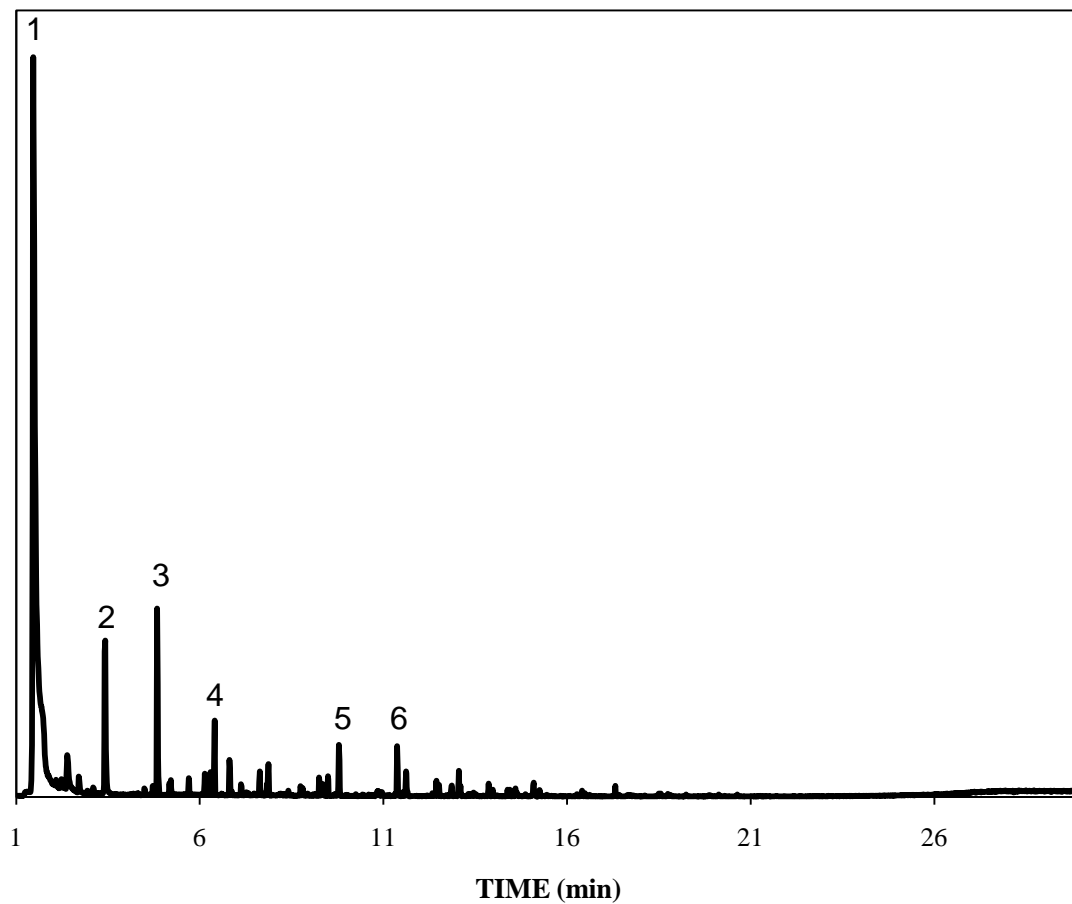


Figure C.4 Chromatogram of Polypropylene, chlorinated; **(1)** hydrogen chloride; **(2)** toluene; **(3)** benzene, 1,3-dimethyl-; **(4)** benzene, 1,3,5-trimethyl-; **(5)** 1H-indene,1-methylene-; **(6)** naphthalene, 1-methyl-

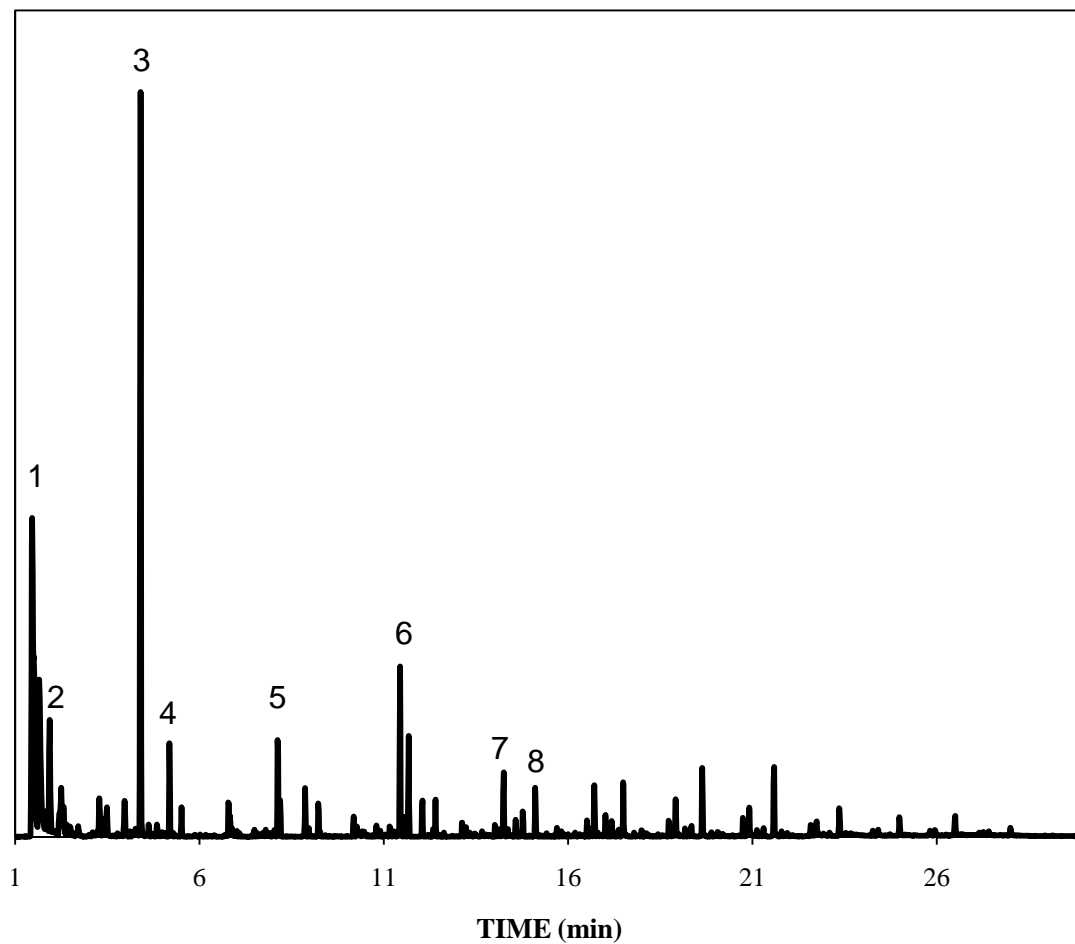


Figure C.5 Chromatogram of Polypropylene, isotactic; **(1)** cyclopropane; **(2)** 1-pentene, 2-methyl-; **(3)** 2,4-dimethyl-1-heptene; **(4)** cyclopentane, 1,2,3,4,5-pentamethyl-; **(5)** 1-decene, 2,4-dimethyl-; **(6)** 2-isopropyl-5-methyl-1-heptanol; **(7)** 3-Eicosene, (E)-; **(8)** 1,19-Eicosadiene

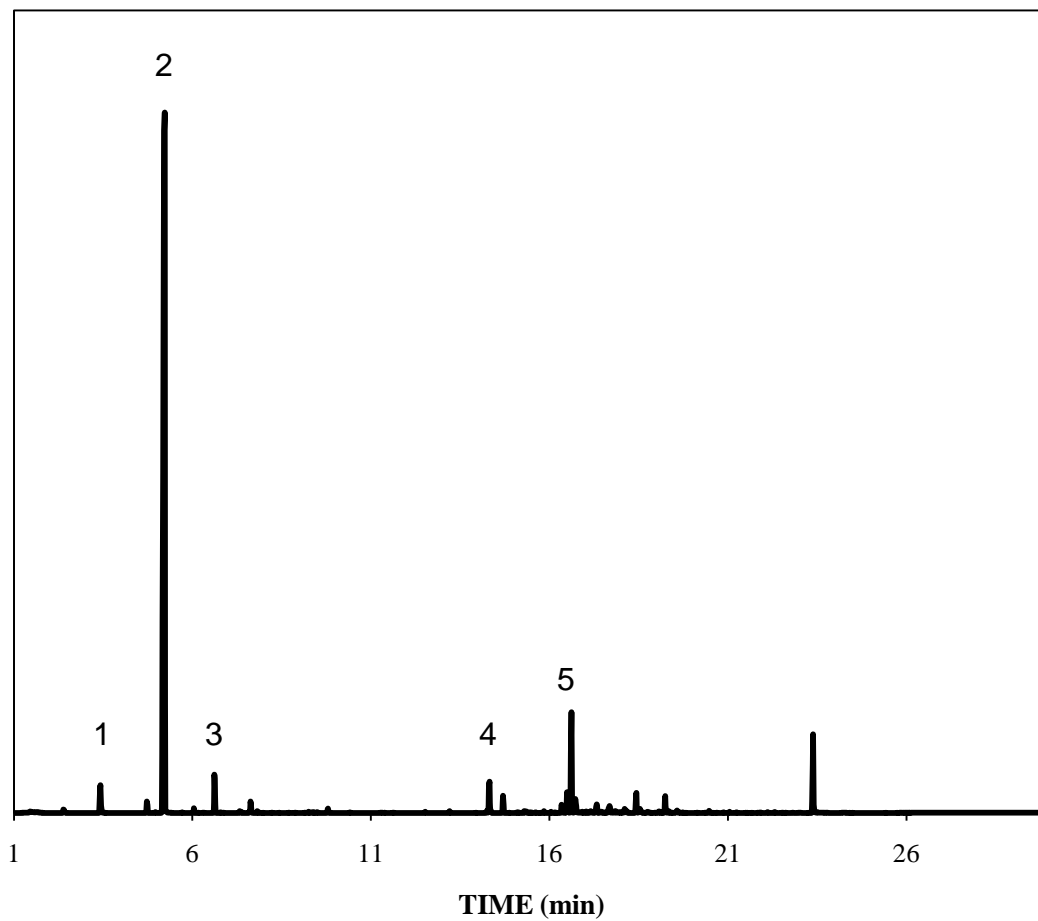


Figure C.6 Chromatogram of Polystyrene; **(1)** toluene; **(2)** styrene; **(3)** α -methylstyrene; **(4)** Dibenzyl; **(5)** naphthalene, 1,2,3,4-tetrahydro-2-phenyl

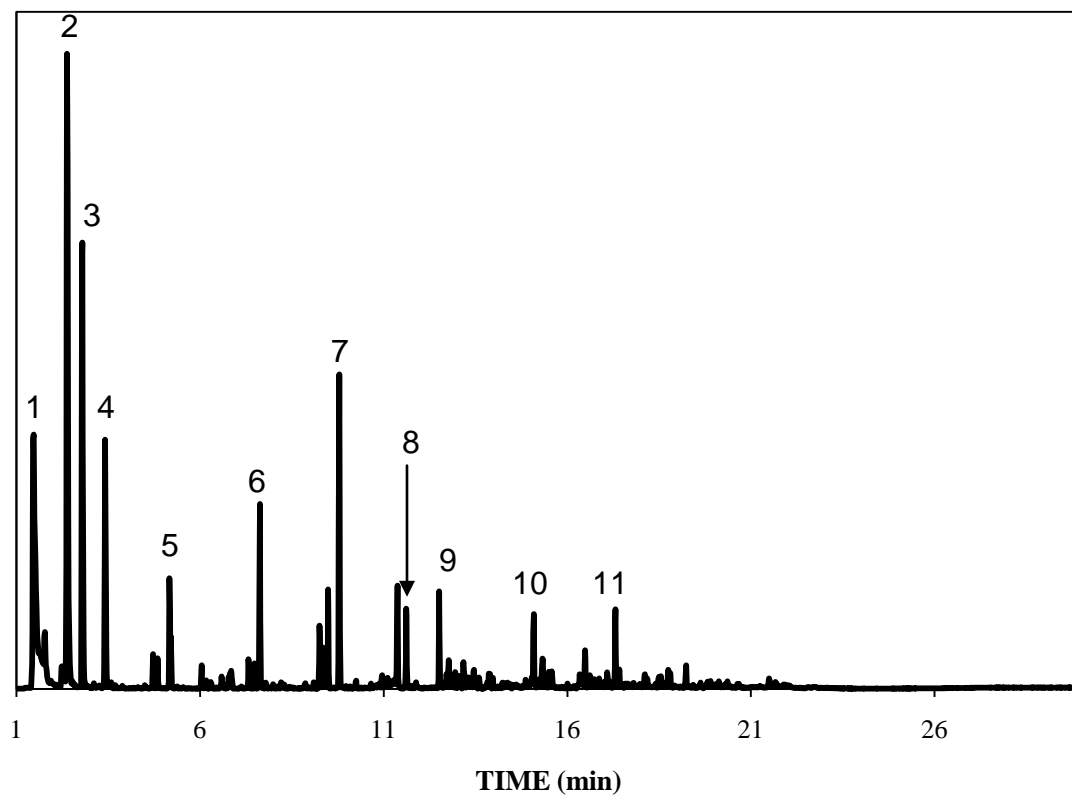


Figure C.7 Chromatogram of PVC, carboxylated 1.8%; **(1)** hydrogen chloride; **(2)** benzene; **(3)** 2-propenoic acid, 2-methyl-,methylester; **(4)** toluene; **(5)** styrene; **(6)** Indene; **(7)** naphthalene; **(8)** naphthalene, 2-methyl-; **(9)** biphenyl; **(10)** fluorene; **(11)** anthracene

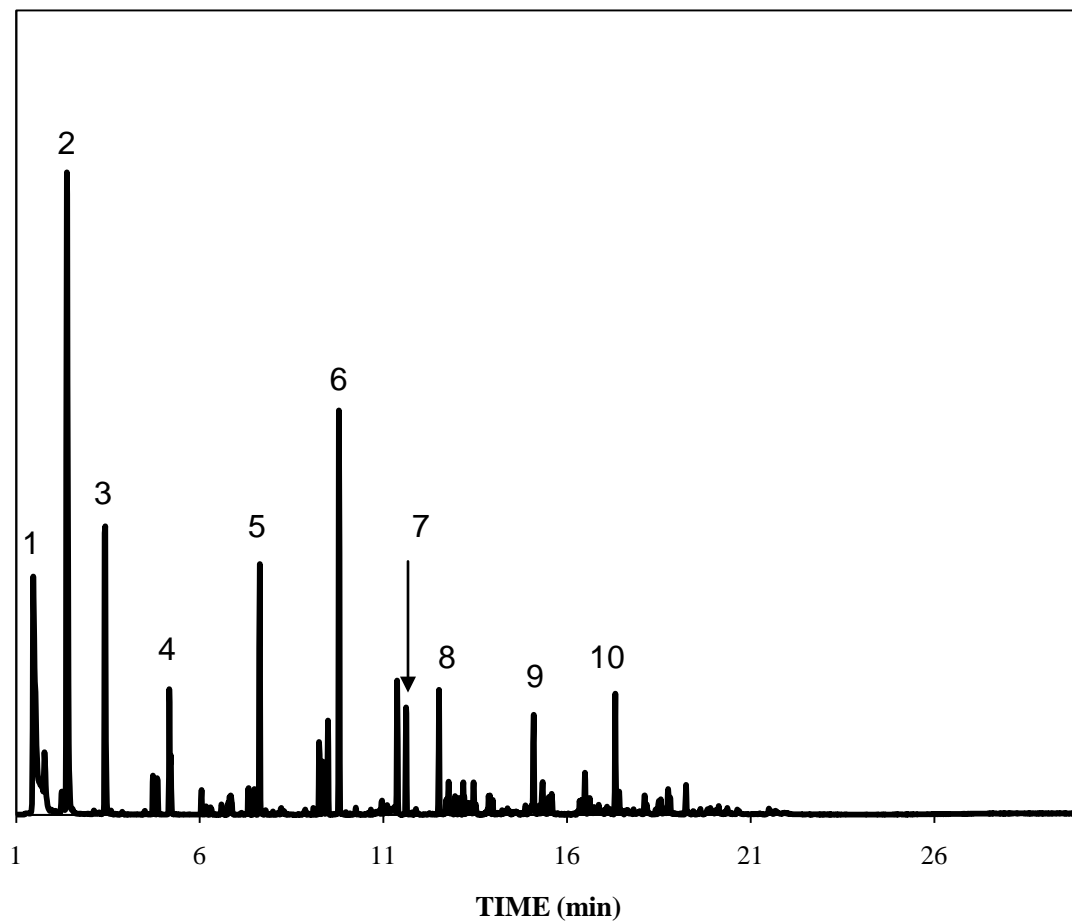


Figure C.8 Chromatogram of PVC, viscosity 1.26; **(1)** hydrogen chloride; **(2)** benzene; **(3)** toluene; **(4)** styrene; **(5)** Indene; **(6)** naphthalene; **(7)** naphthalene, 2-methyl-; **(8)** biphenyl; **(9)** fluorene; **(10)** anthracene