APPLICATION OF RAMAN AND INFRARED MICROSCOPY COUPLED WITH CHEMOMETRICS FOR THE FORENSIC EXAMINATION OF AUTOMOTIVE CLEAR COATS AND PAINT SMEARS

By

GEORGE PAA KWESI AFFADU-DANFUL

Bachelor of Science in Chemistry
University of Cape Coast
Cape Coast, Ghana
2015

Master of Science in Chemistry
East Tennessee State University
Johnson City, Tennessee
2018

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
July, 2022

APPLICATION OF RAMAN AND INFRARED

MICROSCOPY COUPLED WITH CHEMOMETRICS

FOR THE FORENSIC EXAMIINATION OF

AUTOMOTIVE CLEAR COATS AND PAINT SMEARS

Dissertation Approved:

Dr. Barry K. Lavine

Dissertation Adviser

Dr. Ziad El Rassi

Dr. Nicholas F. Materer

Dr. Richard Bunce

Dr. Albert T. Rosenberger

ACKNOWLEDGEMENTS

I am grateful to Dr. Barry Kenneth Lavine, my research advisor for his patience, direction, and inspiration. His immense knowledge contributed significantly to the success of my studies here at Oklahoma State University.

I would also like to thank my graduate advisory committee, Dr. Zaid El Rassi, Dr. Nicholas Materer, Dr. Richard Bunce, Dr. Albert T. Rosenberger, for their guidance, and insightful comments.

I thank Dr. Kaan Kalkan, OSU School of Mechanical and Aerospace Engineering for his guidance and granting me access to his Raman microscope and the Impact Tester. This was crucial to the success of the research projects described in this dissertation. I appreciate Dr. Linqi Zhang for his effective collaborative efforts during my time in the Kalkan research lab.

I also thank my lab members, Haoran, Isio, Tom, and Elizabeth for their support. Special thanks to former lab members, Dr. Francis Kwofie, Dr. Kaushalya Sharma, and Dr. Collin White for their helpful discussions and support.

Finally, I want to thank my entire family and friends especially my wife Mrs. Jane Affadu-Danful, my mother, Leticia Adjei, Mr. and Mrs. Affadu, Francis Baidoo, Gloria, Priscilla, Kwesi, Kuuku, Doxa and Evans Addo for their support and encouragement.

Name: GEORGE PAA KWESI AFFADU-DANFUL

Date of Degree: JULY, 2022

Title of Study: APPLICATION OF RAMAN AND INFRARED MICROSCOPY COUPLED WITH CHEMOMETRICS FOR THE FORENSIC EXAMINATION OF AUTOMOTIVE CLEAR COATS AND PAINT SMEARS

Major Field: CHEMISTRY

Abstract: Modern automotive paints typically use thinner undercoat and color coat layers protected by a thicker clear coat layer. All too often, a clear coat is the only layer of automotive paint left at the crime scene. Current approaches using automotive paint databases to identify clear coats have been unsuccessful because the FTIR spectra of clear coats are too similar to generate accurate hit lists by searching clear coat FTIR spectra alone. Recently published studies of pattern recognition methods applied to FTIR spectra of clear coats have shown that information about the line and model of a vehicle can be obtained from these spectra. To further enhance the general discrimination power of clear coats, Raman spectroscopy and pattern recognition techniques have been investigated as a better solution to the problem of extracting investigative lead information from automotive clear coats. The general discrimination power of Raman spectra for automotive paint comparisons involving 118 General Motors clear coat samples (6 assembly plants for the production years 2000-2006) have been compared to results previously obtained using FTIR spectroscopy. The results of the study show that Raman spectroscopy is a better solution to the problem of extracting investigative lead information from automotive clear coats than IR spectroscopy.

A procedure to simulate the type of paint smear generated in vehicle-vehicle collisions has also been developed. Currently, paint smears encountered in a vehicle collision cannot be created in a laboratory. To this end, testing accessories have developed and adapted to an impact tester which permitted the acquisition and analysis of a large number of paint smears on metal substrates. A machine learning approach to reconstruct IR spectra of the original layers from paint smears allowed searching of an in-house automotive paint library to seek a best match. The results of this study demonstrated that inter-comparisons of paint smears to OEM automotive paints using IR spectra alone to quantify discrimination power of OEM paint smears is feasible and the use of pattern recognition techniques can further efforts to communicate trace evidential significance to the courts.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

Figure Page

CHAPTER I

INTRODUCTION

Automotive paint is an important category of trace evidence as it can provide crucial links to the early phase of an investigation or evidence of association in criminal proceedings. Automotive paint in a form of paint chip or smear is often the only physical evidence recovered at the crime scene of a vehicle related fatality such as head-on collision or a hit-and-run, where injury or death to a pedestrian has occurred. Modern original equipment manufacturer (OEM) automotive paint is highly engineered, manufactured material applied to the frame of a vehicle to protect it from corrosion and photochemical degradation while providing the vehicle with the desired color and finish.

OEM automotive paints (see Figure 1.1) typically consist of four layers [1 - 3]. The primer (or the electro-coat) is first layer (approximately 20 μm thick) and is directly bonded to the automotive substrate. This layer protects the vehicle from corrosion, provides chip resistance and is the foundational layer on which the other layers of paint are applied. The next layer is the surfacer-primer (approximately 30 μm thick) which improves the surface uniformity of the paint and also provides chip resistance.

The color coat layer, which is also known as the basecoat (approximately 15 μm thick) is applied on top of the surfacer-primer layer and is responsible for the color of the vehicle. The top layer, the clear coat (approximately 50 μm thick) forms a glossy and transparent coating that is in direct contact with the environment. The clear coat layer protects the color coat layer from physical abrasions, stone chipping, and ultraviolet light. All four layers of an OEM paint contain binder (the polymer matrix that maintains the physical integrity of the paint and serves as a medium for pigments and additives suspended in it) and with the exception of the clear coat, each layer contains pigments (powdered compounds that impart color, luster, and opacity to the paint) and fillers (inorganic compounds such as kaolin, calcium carbonate, silicon dioxide and barium sulfate) that increase the bulk matter in the paint. A forensic chemist is able to individualize an OEM paint (i.e., determine the most likely make, model, and production year of the vehicle from an OEM paint sample recovered at a crime scene) because automotive manufacturers use different combinations of binders, fillers, and pigments in each layer for different assembly plants and production years.

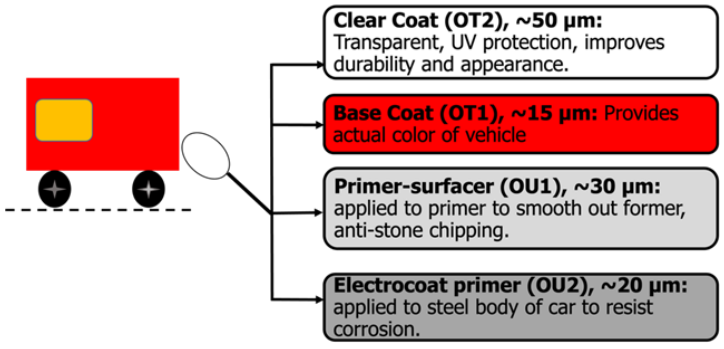

Figure 1.1 Modern automotive paint system containing the clear coat, the base coat, the primer-surfacer, and the electro-coat.

In the forensic examination of automotive paint, each layer of paint is visually and chemically analyzed. Forensic paint examiners have taken advantage of a variety of methods to analyze the chemical composition of automotive paint. Pyrolysis gas chromatography/mass spectrometry (PyMS) has been used by paint examiners to characterize the monomers comprising the binders in each layer [4, 5] after hand sectioning the paint chip. The drawback of PyMS is the length of time required for the analysis and the destruction of the sample. Furthermore, the mass of the paint sample must be greater than 10 milligrams. Many paint chips recovered from the crime scene of a vehicle related fatality are usually less than 10 milligrams. Other methods such as scanning electron microscopy coupled to energy dispersive X-ray spectroscopy [6, 7] and X-ray fluorescence spectroscopy are limited to elemental analysis of paint evidence [8, 9] and are complementary techniques. In North America, chemical analysis of automotive paints in forensic laboratories is typically done using Fourier transform infrared (FTIR) spectroscopy, which relies on the selective absorption of infrared (IR) radiation by the components in the paint. For FTIR analysis, each paint layer is separated and placed between two diamond anvil cells [10].

For automotive paint found at the crime scene of a vehicle related fatality such as a hit-and-run, each layer of the paint may be visually and chemically analyzed and compared to paint from a suspect vehicle. However, there are typically no witnesses to a hit-and-run, and police are often unable to develop a suspect. In these situations, the layer structure, color and composition of the automotive paint from the vehicle or clothing of the victim is characterized because it can serve as a fingerprint either linking or excluding suspect vehicles from association with the crime scene. This is the same approach that would be taken if a forensic paint examiner was presented with both a questioned (unknown) paint sample and a known

(control) sample from a suspect vehicle. In either scenario (with or without a control sample), an automotive paint database would be used to confirm the number of other possible vehicle makes/models/years that could have contributed the questioned paint sample, even if a known paint sample was obtained from a suspect vehicle.

Studies [11, 12] performed by the Royal Canadian Mounted Police (RCMP) over 45 years ago showed that automotive vehicles can be discriminated by directly comparing the color, layer sequence, and chemical composition of each individual layer in an OEM paint. To make these comparisons possible, a comprehensive automotive paint database was developed called the paint data query (PDQ) database. The PDQ concept [13, 14] is to narrow the list of potential vehicles to a reasonable number of suspects (approximately fifty to one hundred hits) that a forensic paint examiner can then work through to identify possible suspect vehicles, not to identify a single vehicle. Currently, PDQ contains over 22,000 samples (street samples and factory panels) that correspond to over 88,000 individual paint layers, representing the manufacturing paint systems used on most domestic and foreign vehicles marketed in North America. PDQ is the largest automotive paint database in existence and is being used by forensic scientists in Canada, USA, Australia, New Zealand, Singapore, Japan, South Africa, and many European countries within the EUCAP community. If the original layers are present in an unknown paint sample, PDQ can assist in identifying the make and model of the vehicle as well as the year range of potential vehicles from which an OEM paint sample may have originated.

Modern OEM automotive paints typically consist of a thin undercoat and color coat layer protected by a thicker clear coat layer. All too often, the clear coat is the only layer of paint left at the crime scene. In these situations, searches of the PDQ database will generate a

hit list consisting of thousands of samples. Modern automotive clear coats applied to any painted automotive substrate usually have only one of two possible formulations which are coded as acrylic melamine styrene or acrylic melamine styrene polyurethane in PDQ. Clear coat formulations are too similar for commercial library search algorithms to generate accurate hit lists by searching clear coat IR spectra alone, as there are no inorganic fillers or color with which to further discriminate a clear coat sample. Furthermore, commercial library search algorithms have not proven to be sufficiently sensitive at distinguishing subtle features in the data such as shoulders, unique shapes and patterns, and minor peaks. Band shifting is not handled well and bands of very low intensity, which may be highly informative, are often ignored [15]. The inability of FTIR spectroscopy and the PDQ database to identify a vehicle from the clear coat layer alone using the text-based searching system of PDQ or standard spectral search algorithms is a significant limitation in the use of FTIR spectroscopy and forensic automotive paint databases such as PDQ to obtain information about the make, model, and production year of a vehicle from an automotive paint sample.

Recently published studies [16 – 20] on pattern recognition methods applied to FTIR spectra of clear coats have shown that information about the make and model of the vehicle can be obtained from these spectra with a far greater likelihood of success than text based PDQ searches or direct searching of IR spectra by commercial library search algorithms. In a previously funded National Institute of Justice research project (2010-DN-BX-K17), library search prefilters were successfully developed from 478 clear coat IR spectra spanning 25 General Motors (GM) assembly plants as part of a prototype pattern recognition assisted IR library search system for clear coats, which consisted of search prefilters to reduce the size of the IR spectral library to a specific assembly plant or plants corresponding to the unknown

clear coat and a cross correlation library search algorithm to identify IR spectra most similar to the unknown in the subset of IR spectra identified by the search prefilters. Even in challenging trials where the samples evaluated were all of the same make (GM) within a limited production year range (2000-2006), the search prefilters were able to identify the specific assembly plant that manufactured the vehicle or to reduce the size of the IR spectral library to a few assembly plants (and hence specific lines and models) responsible for the automotive paint sample.

To further enhance the general discrimination power of clear coats, we propose to investigate Raman spectroscopy and pattern recognition techniques as a potentially better solution to the problem of extracting investigative lead information from clear coats. As noted above, the chemical formulation of modern automotive clear coats designated in PDQ is acrylic melamine styrene or acrylic melamine styrene polyurethane. Both Raman and IR spectra have distinct and recognizable peaks for melamine, styrene, and urethane. As for the acrylate component, there are eight distinct acrylate monomers that could be present in a typical clear coat formulation [21]. The key bands in the IR spectra of acrylates in clear coats are the carbonyl stretch and the C-O-C antisymmetric stretch. The latter band may be characteristic of specific acrylate monomers and is allowed in IR but is too weak to be observed unequivocally in the IR spectra of clear coats because of the presence of bands due to other components. Raman spectra contain acrylate bands, not only of the carbonyl stretching mode but also of the C-O-C symmetric stretch, which is characteristic of specific acrylate monomers, are less overlapped by neighboring bands in the Raman spectra of clear coats and are intense enough to be easily observed [22].

The complementary nature of Raman and IR spectroscopy and the importance of FTIR spectroscopy in forensic automotive paint analysis make it clear that Raman spectroscopy also has the potential to extract investigative lead information from automotive paints. Raman spectroscopy has advantages over IR spectroscopy. Raman bands generally do not overlap, whereas IR bands often are overlapped in spectra. Because of the different selection rules, IR bands, which may be too weak to be observed, can be sufficiently intense to be observed in Raman spectra. However, only a few studies have been reported in the literature on the use of Raman spectroscopy to characterize automotive clear coats. De Gelder [23] and He [24] demonstrated that reproducible Raman spectra of clear coats can be obtained from pearlescent automotive paints. Maric and coworkers [25] explored the potential of Fourier transform Raman spectroscopy and multivariate analysis to characterize and classify automotive clear coats by vehicle manufacturer and model. Unfortunately, the classification success rates reported by Maric were overly optimistic as replicate spectra of the same clear coat sample were included in both the training and validation sample sets. Furthermore, most of the automotive classes surveyed by the authors in this study were represented by only a few samples (less than four). Thus, most of the classes investigated failed to capture the variability of the chemical composition of this layer for these makes and models. The large number of vehicle manufacturers and models surveyed in this study and the small size of the data set collected by the authors makes the class membership problem that was investigated by the authors appear trivial as these automotive paint systems are from Japan, the United States, and several countries in Europe. Nevertheless, Maric's study raises some interesting possibilities about the potential of Raman spectroscopy to extract investigative lead information from automotive clear coats.

In the research described in this dissertation, the general discrimination power of Raman spectroscopy for automotive paint comparisons involving a library of 118 GM clear coat samples (spanning 6 assembly plants for the production years 2000-2006) will be compared to the results previously obtained using FTIR spectroscopy. Paint samples used in this study were obtained from the hood, roof, door, and trunk of the vehicle. The development of Raman search prefilters using the Raman spectra of 118 GM clear coats, which is a major thrust of this project, is needed to assess the general discrimination power of clear coats and the investigative lead information present in a clear coat paint sample. Information derived from these pattern recognition searches can also serve to quantify the general discrimination power of automotive paint comparisons using Raman versus IR spectroscopy and can further efforts to succinctly communicate the significance of evidence from vibrational spectra of clear coats to the courts. Addressing these concerns is a direct response to Recommendation 3 of the National Academies February 2009 report dealing with strengthening forensic science in the United States [26].

A second thrust of the research described in this dissertation is the development of a procedure to generate automotive paint smears (e.g., abraded or deformed clear coats, clear coat and color coat layers mixed together, or clear coat, color coat, primer, and e-coat layers mixed together) as a result of paint transfer from an automobile substrate to any surface of interest under conditions simulating those of a real collision. Realistic paint smears are difficult to create in a laboratory and have also proven challenging to analyze because of the mixing of the various automotive paint layers. To this end, testing accessories have been developed and adapted for a standard aggregate impact tester. This allows the capability to acquire and analyze a large number of paint smears on metal, fabric, and plastic surfaces by

attenuated total reflection (ATR) infrared imaging. The impact technique simulates an automobile collision with controllable collision speed and direction as well as momentum/energy transfer. Simulating the type of paint smear generated in a vehicle-vehicle or vehicle-pedestrian hit-and run-collision is a valuable asset to the forensic science community. One long-term goal of this paint transfer effort is the development of forensic standards for automotive paint transfer. Another is the development of new methods for paint smear analysis that can capture the vibrational spectra of individual layers of automotive paint. Currently, these standards and analytical methods are not available to forensic automotive paint examiners. By developing these standards and methods, the accuracy of paint smear analysis is enhanced. Furthermore, the paint transfer effort also makes possible the creation of realistic proficiency tests for forensic laboratories, and it can also be used for training of forensic scientists. This will result in a significant improvement in how the competency of laboratories conducting forensic automotive paint analysis is assessed.

References

1. H. J. Streitberger and K.F. Dossel Editor, "Automotive Paint and Coatings" Wiley-VCH, KGaA, Weinheim, 2008.

2. George Fettis, "Automotive Paints and Coatings" VCH, Weinheim, 1995.

3. R. Lambourne and T. A. Strivens, "Paint and Surface Coatings – Theory and Practice" Woodhead Publishing, New York, 1999.

4. A. R. Cassista and M. L. P. Sandercock, "Comparison and Identification of Automotive Topcoats: Microchemical Spot Tests, Microspectrophotometry, Pyrolysis Gas Chromatography, and Diamond Anvil Cell FTIR" *Can. Soc. Forensic Sci. J.* **1994,** *27,* 209-223.

5. Paul Kirkbride, "Paint and Coating Examinations" in J. Siegel, Forensic Chemistry: Fundamentals and Applications, Wiley Blackwell, NY, 2016, pp. 11-115.

6. M. L. Henson and T. A. Jergovich, "Scanning Electron Microscopy and Energy Dispersive X-ray Spectrometry (SEM/EDS) for the Forensic Examination of Paints and Coatings, In B. Caddy (Ed.) Forensic Examination of Glass and Paint: Analysis and Interpretation" Taylor & Frances, NY 2001, pp. 243-272.

7. A. Reynolds, K. Roberts, A. Thompson, and E. Runt, "Discrimination Power of Automotive Paint Comparisons using a Paint Analytical Scheme" *JASTEE,* **2018,** *8* (1), 4-1.

8. S. G. Ryland, T. A. Jergovich, and K. P. Kirkbride, "Current Trends in Forensic Paint Examination" *Forens. Sci. Rev*., **2006,** *18*, 97-117.

9. N. Yoshinori; W. Seiya, S. Osamu, S. Yasuhiro, N. Toshio, T. Yasuko, N. Toshio, and N. Izumi, "Trace Elemental Analysis of Titanium Dioxide Pigments and Automotive White Paint Fragments for Forensic Examination Using High-Energy Synchrotron Radiation X-Ray Fluorescence Spectrometry" *J. Forensic Sci*., **2009,** *54* (3), 564-570.

10. A. Beveridge, T. Fung, D. MacDougall, "Use of Infrared Spectroscopy for the Characterization of Paint Fragments" In: Forensic Examination of Glass and Paint: Analysis and Interpretation, B. Caddy, Ed., Taylor and Francis, NY, NY, 2001, pp. 220- 233.

11. P. G. Rodgers, R. Cameron, N. S. Cartwright, W. H. Clark, J. S. Deak, and E. W. Norman, "The Classification of Automobile Paint by Diamond Window Infrared Spectrophotometry by Binders and Pigments*" Can. Soc. Forens. Sci. J.* **1976,** *9* (1), 1-14

12. P. G. Rodgers, R. Cameron, N. S. Cartwright, W. H. Clark, J. S. Deak, and E. W. Norman, "The Classification of Automobile Paint by Diamond Window Infrared Spectrophotometry Part II Automotive Topcoats and Undercoats" *Can. Soc. Forens. Sci. J.* **1976,** *9* (2), 49-68.

13. J. L. Buckle, D. A. MacDougal, and R. R. Grant, "PDQ-Paint Data Queries: The History and Technology Behind the Development of the Royal Canadian Mounted Police Forensic Laboratory Services Automotive Paint Database" *Can. Soc. Forens. Sci. J.* **1997,** *30,* 199-212.

14. N. S. Cartwright, L. J. Cartwright, E. W. W. Norman, R. Cameron, W. H. Clark, D. A. MacDougal, "A Computerized System for the Identification of Suspect Vehicles Involved in hit-and -run Accidents" *Can. Soc. Forens. Sci. J*., **1982**, *15* (3/4), 105-115.

15. G. W. Small, "Automated Spectral Interpretation," *Anal. Chem*., **1987**, *59* (7), 535A-546A.

16. A. Fasasi, N. Mirjankar, R.-I. Stoian, C. White, M. Allen, M. P. Sandercock and B. K. Lavine, "Pattern Recognition Assisted Infrared Library Searching of Automotive Clear Coats" *Applied Spec*., **2015**, *69* (1) 84-94.

17. B. K. Lavine, A. Fasasi, N. Mirjankar, and M. Sandercock, "Search Prefilters to Assist in Library Searching of Infrared Spectra of Automotive Clear Coats" *Talanta,* **2015,** *120,* 182-190.

18. B. K. Lavine, A. Fasasi, N. Mirjankar, and M. Sandercock, "Development of Search Prefilters for Infrared Library Searching of Clear Coat Paint Smears" *Talanta*, **2014**, *119*, 331–340.

19. B. K. Lavine, A. Fasasi, N. Mirjankar, M. Sandercock, and S. D. Brown, "Search Prefilters for Mid-IR Spectra of Clear Coat Automotive Paint Smears Using Stacked and Linear Classifiers" *J. Chemom.,* **2014,** *28,* 385-394.

20. B. K. Lavine, A. Fasasi, N. Mirjankar, M. Sandercock, and S. D. Brown, "Search Prefilters for Mid-IR Spectra of Clear Coat Automotive Paint Smears Using Stacked and Linear Classifiers" *J. Chemom,* **2014,** *28,* 385-394.

21. S. G. Ryland and E. Suzucki, "Analysis of Paint Evidence" in Chemistry Forensic Handbook, John Wiley & Sons, NY, 2011, pp. 131-224.

22. H. A. Willis, V. I. J. Zichy, and P. J. Hendra, "The Laser-Raman and Infra-red Spectra of Poly Methyl Methacrylate Polymer" 1969, 10, 737-746.

23. J. De Gelder, P. Vandenabeele, F. Govaert, and L. Moens, "Forensic Analysis of Automotive Paints by Raman Spectroscopy" *J. Raman Spec*., **2005,** *36*, 1059-1067.

24. J. He, J. Lv, Y. Ji, J. Feng, and Y. Liu, "Multiple Characterizations of Automotive Coatings in Forensic Analysis" *Spec. Lett*., **2013,** *46*, 555-560

25. M. Maric, W. van Bronswijk, K. Pitts, and S. W. Lewis, "Characterization and Classification of Automotive Clear Coats with Raman Spectroscopy and Chemometrics for Forensic Purposes" *J. Raman Spec*., **2016,** *47* (8), 948-955.

26. "Strengthening Forensic Sciences in the United States: A Path Forward," National Research Council of the National Academies Press, Washington DC, February 2009.

CHAPTER II

METHODOLOGY

This section of the dissertation describes the methodology used to simulate automotive paint smears with an impact tester and the analysis of the paint smears using attenuated total reflection Fourier transform infrared (ATR-FTIR) microscopy and multivariate curve resolution. Raman analysis of automotive clear coats is also discussed with the emphasis on sample preparation and the collection of the Raman spectral data. Paint smears and clear coats represent OEM paint samples that are problematic in terms of identifying the make, model, and year of the vehicle from which the paint sample originated.

2.1 Generation of Automotive Paint Smears Using an Impact Tester.

Car paint evidence is often obtained from "head-on" collisions that result in physical smearing. Paint smears, unlike paint chips, are very difficult to analyze because of the mixing of the layers in the paint sample. To create paint smears mimicking those of real-life vehicle-vehicle collisions, a footprint collider for transferring paint in the form of a smear from an OEM automotive paint sample to a steel substrate under conditions emulating a vehicle-vehicle collision

was provided for this study by Dr. Kaan Kalkan (Department of Mechanical & Aerospace Engineering, Oklahoma State University-Stillwater). The design (see Figure 2.1) involves the collision between two wedge-shaped steel blocks. The automotive paint sample is mounted on the upper wedge, whereas the paint-transfer substrate is mounted on the lower wedge. The upper wedge is released from a specified height and accelerated by gravity towards the lower wedge. The upper wedge slides vertically on two guiding rods, but it transfers both vertical and horizontal momentum to the lower wedge upon collision, because the collision surfaces (i.e., wedge surfaces) are at 45^0 to the horizontal/vertical. Because of this design, the contact force, which is normal to these surfaces, has both vertical and horizontal components. Upon collision, energy and momentum is transferred to the lower wedge, as it is allowed to move vertically and horizontally on the rail guide. As a result, the lower wedge surface also slides against the upper wedge surface, creating friction. The frictional work is responsible for the transfer of the paint as a smear.

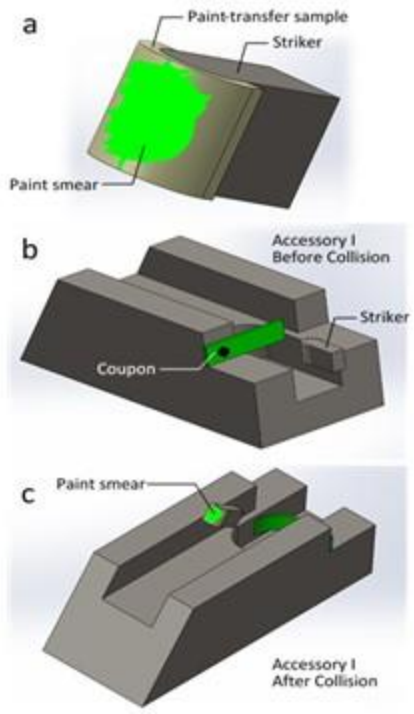

Figure 2.1. (a) Accessories to be adapted to the impact tester for producing paint smear samples. (b and c) Paint-transfer substrate clipped/attached to the striker. Illustration of the accessory concept based on bending of the automotive vehicle substrate and the 'coupon' upon collision.

To realistically emulate a vehicle collision, two tunable shock absorbers were employed (see Figure 2.2), which allowed adjustment of the effective force constant and damping coefficient associated with the collision (i.e., the mechanical coupling between the colliding bodies). Furthermore, the mass of each colliding body (coupon and wedge) can be increased by adding metal weights. (The coupon consists of the OEM paint sample and the automotive substrate to which it is directly bonded.) Collision speed is controlled by setting the height. The orientation of the sliding motion is altered by the surface angle of the lower wedge. Details about the collider design (e.g., geometry, materials, fabrication, dimensions, and masses) are discussed below.

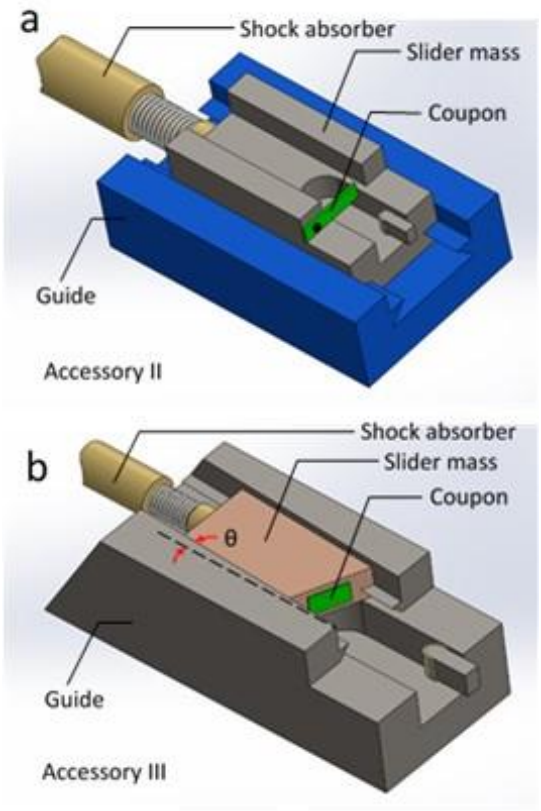

Figure 2.2 (a) Accessory shown in Figure 2.1b is upgraded with a shock absorber and inertial element. (b) Third accessory concept which emulates a collision with effective (adjustable) mass, force constant and damping coefficient.

The collider (see Figure 2.3), which can be operated in a laboratory environment, is approximately 1 m in height with a metal base of 22 kg. The base is cylindrical in shape, 30 cm in diameter and is supported by a solid wood block. Two vertical stainless-steel rod guides (which are set 16.5 cm apart and serve as the motion track for the upper collision wedge) are supported by the base. The stainless-steel rods on which the wedge is mounted guide a 14 kg metal slider. A lifting handle is provided to raise up the slider mass, and a bolt is used to lock it into a set position prior to release or when the system is off duty.

Figure 2.3 Complete assembly of the impact tester used to simulate automotive paint smears.

The lower wedge is mounted on a cross-shaped steel plate (1-inch-thick), which also slides vertically on the two guiding rods through holes bored on one axis of the cross-shaped steel plate. The cross-shaped plate is vertically coupled and supported by one of the two shock absorbers, which is mounted on the base. The horizontal motion of the lower wedge is guided by a slider-rail assembly along the other axis of the cross-shaped metal platform. The lower wedge is mounted directly on this slider-rail assembly, which is mounted on the cross platform. The horizontal motion of the lower wedge is coupled to the second shock absorber, which is also mounted on the cross-shaped steel plate. The design of the collider frame, bone, and upper slider are based on a well-known aggregate impact

tester (Ajanta). The vertical guides are tightened to the base by screw nuts to increase the stability and strength of the frame. The cross-shaped platform is machined (by milling) from a steel block. The sliding rail is cut to an appropriate length (25 cm) in order to be mounted on the platform. The two wedges are milled from low carbon steel blocks. Two mounting threaded bores are machined on at the base of each wedge. The inclined (collision) planes of the wedges are fine polished by 600-grit sandpaper.

The collider was inspected for vibration-free and smooth operation as well as paint smear formation. The collider was tested by releasing the upper colliding mass at various heights up to 1.0 m. The collider exhibited high mechanical stability and resilience. The shock absorbers can be tuned using their entire operational range, offering a broad span of damping coefficients. One collision parameter is the speed prior to the collision, which is controlled by the release height of the upper colliding mass. A release height of 1.0 meter was used in these studies. Another collision parameter is the contact area during collision. Using a smaller contact area, the pressure will be higher, and the smear heavier. The contact area is controlled by the size of the automotive paint sample on the coupon. For the simulation, the paint sample is mounted on the upper wedge. For rigid immobilization of the sample, the upper wedge was machined for a 15 mm wide slot across the wedge (63 mm wide). The slot was machined 0.5 mm deep, so a typical automotive paint sample including the coupon (1 mm thick) protrudes from the wedge surface and is exposed during the collision. Similarly, the lower wedge was machined for a groove of 20 x 50 x 0.5 mm^3 to hold the paint-transfer substrate. All paint-transfer substrates were sheared from a 0.9 mm thick SAE 304 austenitic stainless steel (polished finish) plate to exactly fit the groove. After preliminary testing, a 10 x 10 mm^2 paint substrate was used as the sample size. Smear

formation also depends on the masses of the colliding bodies and the direction of momentum transfer, which were maintained from the previous configuration. On the other hand, the range of force-constants (stiffnesses) and damping constants for the vertical and horizontal shock absorbers were varied for each simulation. Damping can be set to 8 different levels from 1 to 8 using the adjustment knob. The force-constants were set to 130 lb./inch and 15 lb./inch for the vertical and horizontal shock absorbers, respectively.

A necessary design iteration was an optional sample holder accessory, which was mounted on the upper wedge. This accessory will not only hold the auto paint sample securely during the collision, but it also imposes/defines a certain curvature for the sample surface. In a real collision, the contacting surfaces are generally not coplanar, because car bodies have curvature by design. The sample holder accessory consists of a rounded stainless-steel stage/base, on which the sample is wrapped/covered by bending. To install a paint sample, which is originally flat, one edge is first inserted into the machined crevice and the other edge is fixed by a screw-clamp.

The impactor was tested by generating paint smear samples with varying levels of the elastic force constant and the damping coefficient. These collision parameters are controlled using replaceable springs of two different stiffnesses and by changing the damping level of the shock absorbers. As expected, higher stiffness springs caused more severe collisions, with higher collisional forces and greater frictional work. By exploring the damping levels and the spring stiffness, various automotive paint smears of varying complexities were successfully simulated. Figure 2.4a shows an automotive paint sample received from the Royal Canadian Mounted Police (RCMP) which was then fixed on the upper wedge of the impactor. The paint transfer substrate was positioned into the groove

of the lower wedge (see Figure 2.4b). After the collision, the paint smear generated onto the transfer substrate is shown in Figure 2.4c. Figure 2.5 shows examples of automotive paint smears deposited onto the transfer substrate by the impact tester.

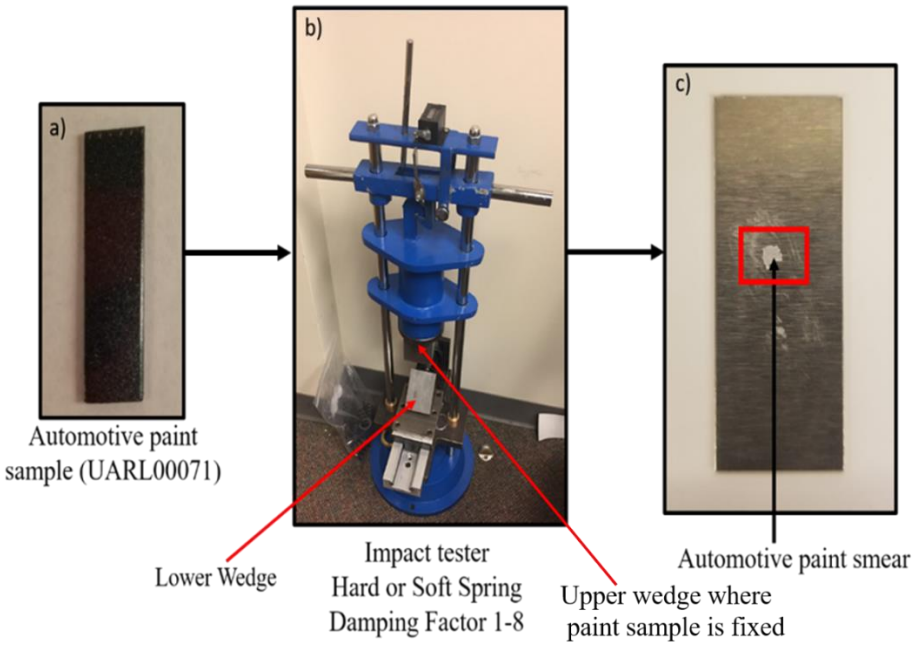

Figure 2.4 Generation of automotive paint smears. (a) automotive paint sample to be fixed on the upper wedge of the impactor (b) the impact tester (c) automotive paint smear transferred onto the transfer substrate after impact.

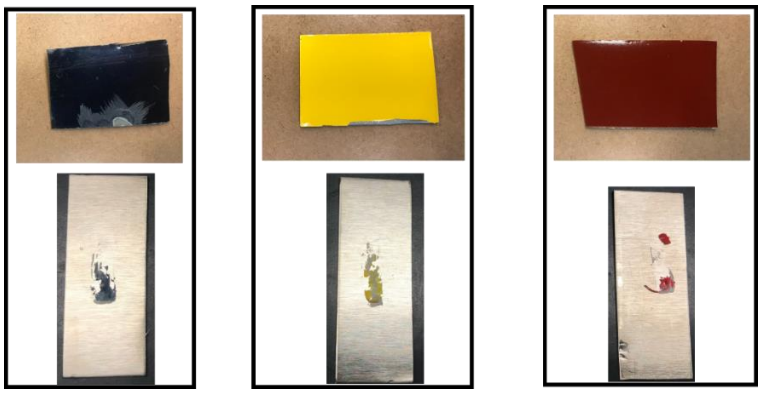

Figure 2.5 Images of automotive paint smears generated by the impactor.

2.2 Analysis of Automotive Paint Smears by Attenuated Total Reflection IR Microscopy.

FTIR spectroscopy is used to probe the composition of a sample based on its interactions with IR radiation. If IR absorbing constituents are present in the sample, the IR spectrum can be treated as a chemical fingerprint of the sample. IR radiation is divided into three spectral regions: near IR (13000 to 4000 cm^{-1}), mid IR (4000 to 400 cm^{-1}) and far IR (400 to 100 cm^{-1}). Most FTIR analyses focus on the mid IR region even though the near IR and far IR regions are also capable of providing useful information.

The absorption of mid-IR radiation by a sample can be performed experimentally using a variety of geometries. Transmission is the direct measurement of absorption of the IR radiation as it passes through the sample. Most solid, liquid, and gaseous samples are analyzed using transmission. Reflection is another widely used sampling geometry. This method is applied to samples that are difficult to analyze by transmission. There are two types of IR reflection experiments: internal reflection (e.g., attenuated total reflection spectroscopy) and external reflection (specular and diffuse reflection spectroscopy). FTIR is a mature technique that has been applied in a variety of disciplines including forensic chemistry [1], food science [2], pharmaceutics [3], agriculture [4], and medicine [5].

For the analysis of automotive paint smears, an attenuated total reflection (ATR) sampling geometry was employed. In ATR, the IR beam enters the internal reflecting element (IRE) (such as a germanium crystal) at an angle above the critical angle between the IRE and the paint smear sample. This causes internal reflection to occur, which generates an evanescent wave that extends beyond the IRE into the paint sample (see Figure 2.6). In the mid-IR region where the paint smear absorbs IR energy, the evanescent

wave from the IRE is attenuated. The reflected IR beam is the basis of the IR spectrum of the paint smear.

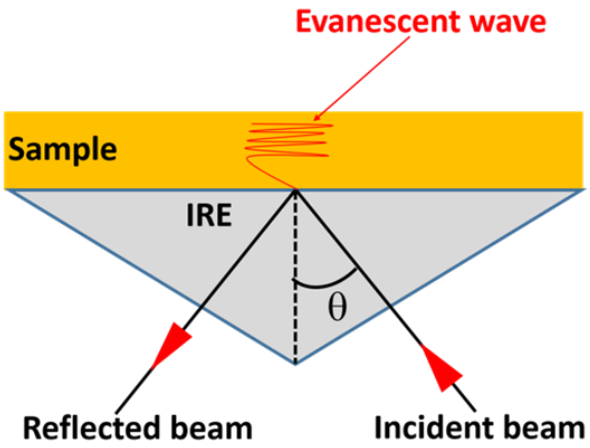

Figure 2.6 Attenuated total reflection infrared analysis

The extent to which the evanescent wave penetrates the sample is represented by its effective penetration depth given in Equation 2.1 [6 -8]

$$dp = \frac{\lambda}{2\pi n_p (sin^2\theta - n_{sp}^2)^{1/2}} \qquad (2.1)$$

where λ is the wavelength of the radiation, θ is the angle of incidence, n_p is the refractive index of the IRE and $n_{sp} = n_2/n_1$ is the ratio of the refractive indices of the sample (n_2) and the IRE (n_1). The penetration depth, which corresponds to the intensity of the evanescent wave decaying to 1/e of its original value, is the distance between the IRE and the sample boundary. The penetration depth of the evanescent wave can be tuned by varying the incidence angle (θ, see Figure 2.7) or the refractive index of the IRE (n_1).

ATR offers significant advantages in IR microscopy as it produces high quality spectra with enhanced spatial resolution. Spatial resolution is determined by the

wavelength and the numerical aperture of the objective of the microscope. Equations 2.2 and 2.3 summarize the spatial resolution achievable by an FTIR microscope where d is the spatial resolution, λ is the wavelength of light and NA is the numerical aperture of the microscope objective. θ is the acceptable half-angle of the IR radiation from the microscope objective illuminating the sample and n_1 is the refractive index of the imaging medium (see Equation 2.3).

$$d = \frac{0.61\lambda}{NA} \qquad (2.2)$$

$$NA = n_1 sin\theta \qquad (2.3)$$

In a transmission FTIR microscope, the spatial resolution (d) is equal to the wavelength of the IR radiation as the numerical aperture is equal to 0.6 (n_1 =1 and the half-angle of the light cone illuminating the sample is 36.87º). For an ATR imaging microscope using a germanium crystal, the numerical aperture is equal to 2.4 because the refractive index of germanium (n_1) equals 4. Thus, the spatial resolution of the ATR FTIR microscope is only one-fourth of the wavelength of the incident radiation. Since the penetration depth of the evanescent wave is lower due to the higher refractive index of germanium, analysis of paint smears transferred onto a substrate will be limited to the surface of the substrate. For this reason, a germanium tip-ATR accessory provided by Thermo-Nicolet for the iN10 FTIR microscope was used in this study (see Figure 2.7). The Tip-ATR accessory allows for collection of IR spectra quickly from small and isolated samples spaces on the substrate [6]. This is advantageous as the paint smears transferred onto the substrate are not evenly

distributed. Furthermore, the Tip-ATR accessory is rugged and is easy to clean between measurements. These factors make it possible to generate high quality spectra with enhanced spatial resolution when an ATR sampling geometry as implemented using the ATR tip is used to analyze automotive paint smears. By comparison, a transmission sampling geometry cannot be employed to analyze the simulated paint smears because the substrate on which the paint smears are deposited is opaque, preventing the IR beam from traversing through it.

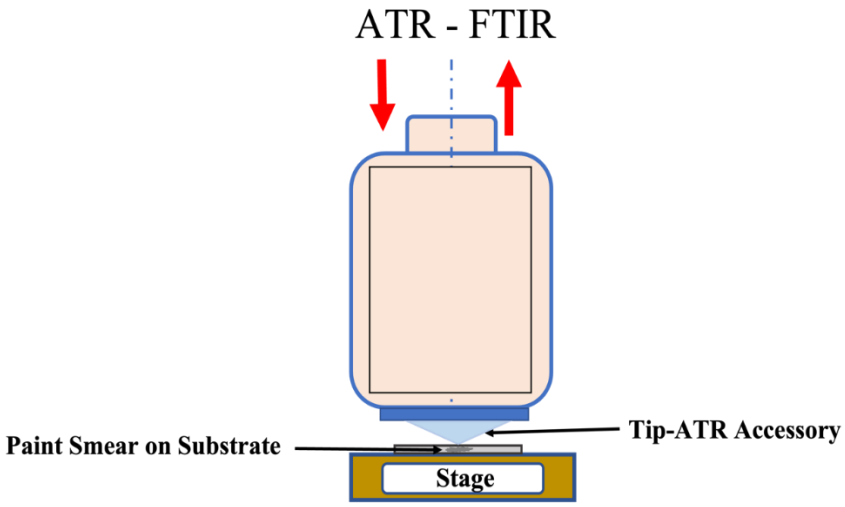

Figure 2.7 Attenuated total reflection infrared microscope

To analyze an automotive paint smear using the Tip-ATR accessory of the iN10-MX microscope, the transfer substrate containing the paint smear is positioned on the stage of the microscope. Using the digital image from the camera incorporated into the microscope as a guide, portions of the substrate containing the paint smear can be located and analyzed (see Figure 2.8).

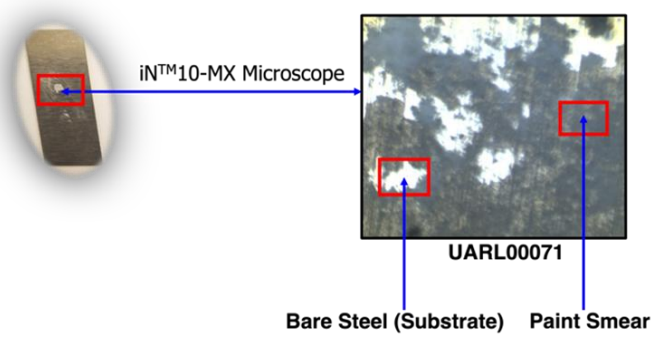

Figure 2.8 Paint sample showing bare substrate and the paint smear where IR spectra are collected.

The adjustable parameters used for this IR analysis include the aperture (150 x 150 with a step size of 5) and the number of scans (128 for these studies). The microscope is set to the reflection mode to allow a preview of the sample and the selection of the appropriate points to collect IR spectra using the ATR crystal (see Figure 2.9). This step is crucial in the analysis as it avoids collecting spectra from regions where the spectra appear to be noisy and also avoid regions on the substrate where the paint smears are not present. After selecting the desired points where the IR spectra will be collected, the microscope is switched to the ATR mode for spectral collection using a germanium crystal (embedded in the tip) as the IRE.

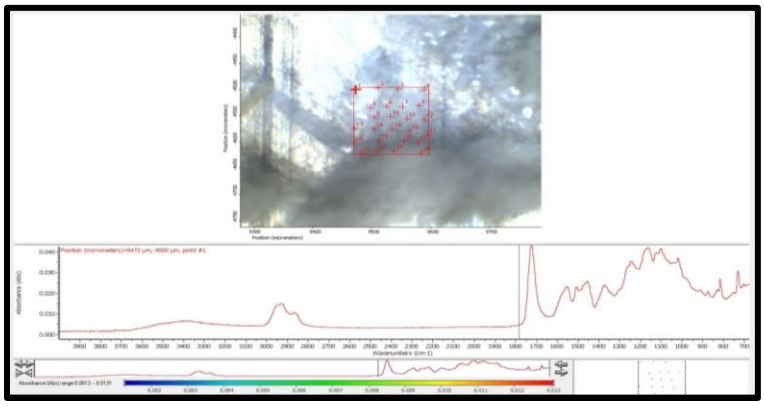

Figure 2.9 FTIR microscope set to the reflection mode to permit paint smear sample preview and selection of desired sample point for ATR analysis.

2.3 Sample Preparation for Raman Microscopy

The second part of the dissertation research focuses on the application of Raman spectroscopy and pattern recognition techniques to the problem of extracting investigative lead information from modern automotive clear coats. To obtain the clear coat layer from a paint chip for analysis, the surface of each paint sample is cleaned with methanol to remove particulates from the surface of the paint chip. Using a shark knife under a stereo microscope, the clear coat layer is gently scrapped off the paint sample and transferred onto a microscope glass slide (Globe Scientific Inc) covered with aluminum foil (see Figure 2.10). The aluminum foil offers a low and almost featureless background during Raman analysis. The sample is then taken to a Raman microscope for analysis.

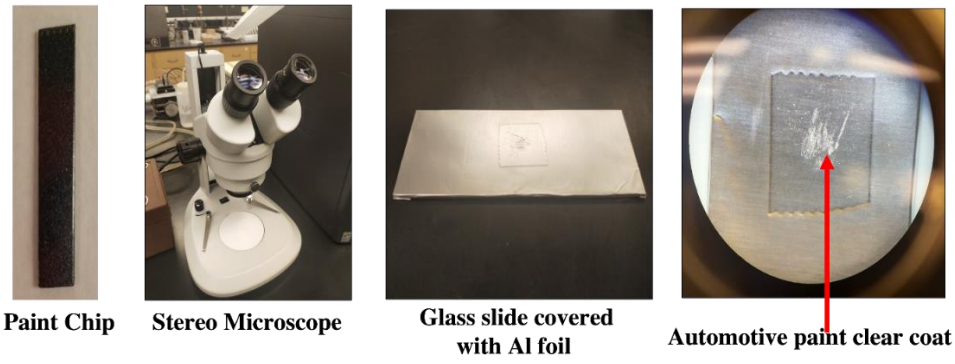

Figure 2.10 Collection of automotive clear coat sample for Raman analysis

2.4 Analysis of Automotive Clear Coats Using Raman Microscopy.

Raman spectra for the clear coats were measured using a WITEC Alpha 300 R confocal micro-Raman system equipped with a 532 nm neodymium-doped yttrium aluminum garnet (ND: YAG) laser and a charge-coupled detector (CCD) which is operated at -60^0 C. A silicon wafer, which exhibits a Raman peak at 520.5 cm^{-1} was used to calibrate the Raman spectrometer. The silicon wafer is positioned on the microscope stage using

tweezers. The Raman spectrum of the silicon wafer is then recorded using an oscilloscope. For a calibrated spectrometer, the expectation is that a Raman spectrum of the silicon wafer should exhibit a peak maximum at 520.5nm. If a different wavenumber is observed for the silicon peak, the laser wavelength is adjusted in small increments (e.g., 532.00 nm to 532.224 nm) until the center of the silicon wafer peak is 520.5 nm.

Each Raman spectrum is collected using an integration time of 10 seconds with 20 spectra collected per sample. If peak shifts are observed in the average sample spectrum, the laser power was lowered from 40 milli-watts to 10 milli-watts and the integration time was increased to 30 seconds with 30 spectra collected per sample. The magnification objectives can be varied from 4x to 100x magnification. In this study, a 20x objective lens (Nikon) with a numerical aperture (NA) of 0.4 for excitation and collection of the backscattered light is employed as it gave the highest S/N. The 20x lens collected the Raman-scattered radiation from a larger area approaching 10 μm diameter at a low confocal regime (i.e., 100 μm pinhole/fiber diameter). As a result, more spectra per sample could be collected with a higher level of statistical averaging. Alternatively, the laser spot can also be dispersed to a larger area minimizing the laser intensity for a given power. Spot size was increased for those samples where photo-degradation occurred. A spot size diameter of either 5 μm or 10 μm was adopted for these studies as a larger spot size resulted in a marked reduction in Raman counts. In this study, the laser from the Raman microscope was focused on the inner surface of each clear coat, not on the surface exposed to the environment.

An artifact-free Raman spectrum of a paint sample has minimal photo degradation, minimal fluorescence background and no photo-thermal effects which induce spectral peak

shifts. For the acquisition parameters selected for most of the samples in this study (40 milliwatts laser power, integration time of 10 seconds and accumulation of 20 spectra per sample), a weak fluorescence background was observed for most of the clear coat samples investigated which can probably be attributed to a small amount of the color coat layer dispersed in the clear coat paint sample collected from each paint chip. For these samples, the background fluorescence was found to gradually decrease during the 20 second signal acquisition with no changes observed in the Raman peaks (i.e., wavenumber, width, and relative intensity) as monitored up to a prolonged acquisition time of 800 seconds.

The spectral resolution of the Raman spectra collected was investigated using the 1800 lines/mm diffraction grating which is blazed at 520.5 nm. Although this grating provides a wavenumber range of only 1105 cm, it covers the fingerprint region which is sufficient for identifying the make and model of the automotive vehicle from which the clear coat sample was obtained. The spectra were acquired from 697 to 1802 cm^{-1} which corresponds to the same spectral region previously investigated in our studies using FTIR [9, 10]. Initial studies using this Raman microscope for clear coats revealed that the use of a lower resolution grating (600 lines/mm) prevented the alignment of spectra crucial for multivariate analysis. For this reason, the 1800 lines/mm grating was used in this study as it was able to generate spectra that are well resolved and properly aligned for multivariate analysis.

Although the collected Raman spectra were reproducible with high S/N, the excessive auto-fluorescence was of concern for some clear coat samples investigated. The fluorescence may be due to pigments transferred from the color coat layer when the clear coat layer, which was thin for these samples, was scrapped off the paint chip when isolating

the clear coat layer from the other layers of paint. Autofluorescence, which is the result of resonant or pre-resonant excitation of various moieties, such as carbonyls, may be initiated by the 532 nm laser (albeit weakly) of our WITec system. To address this concern, photo-bleaching was applied to these samples for a period of 5 minutes using a laser power of 40 mW to saturate the fluorophore so that is unable to fluoresce. A second problem encountered was peak shifting which occurred due to heating of the sample by the laser. Reducing the laser power to 10 mW but increasing the integration time for the detector from 10 seconds to 30 seconds usually addressed this problem. The Raman spectra were collected at five different sites on each clear coat (due to the small laser spot size used) to obtain a representative spectrum of each sample.

References

1. F. Kwofie, N. U. D. Perera, K. S. Dahal, G. P. Affadu-Danful, K. Nishikida, and B. K. Lavine, "Transmission Infrared Microscopy and Machine Learning Applied to the Forensic Examination of Original Automotive Paint" *Appl Spec*., **2022,** *76* (1) 118-113.

2. E. Li-Chan, J. M. Chalmers, and P.R. Griffiths, "Application of Vibrational Spectroscopy in Food Science" John Wiley and Sons., 2010, 2 Volume Set.

3. B. Igne and E. W Ciurczak, "Near-Infrared Spectroscopy in the Pharmaceutical Industry, in Near-Infrared Spectroscopy" Springer., 2021, 391-412.

4. D. Cozzolino, "Recent Trends on the use of Infrared Spectroscopy to Trace and Authenticate Natural and Agricultural Food Products" *Appl Spec. Reviews*., **2012,** *47* (7) 518-530.

5. R. Wang and Y Wang, "Fourier Transform Infrared Spectroscopy in Oral Cancer Diagnosis" *Int. J. Mol. Sci*., **2021**, *22* (3) 1206.

6. P. Larkin, "Instrumentation and Sampling Methods" Infrared and Raman Spectroscopy Principles and Spectral Interpretation. P. Larkin, Elsevier., 2011, 27-54.

7. K. L. A. Chan and S. G. Kazarian, "Attenuated Total Reflection Fourier Transform Infrared Imaging with Variable Angles of Incidence: A Three-Dimensional Profiling of Heterogenous Materials" *Appl Spec*., **2007,** *61* (1) 48-54.

8. F. M. Jr. Mirabella, "Strength of Interaction and Penetration of Infrared Radiation for Polymer Films in Internal Reflection Spectroscopy" *J. Polym. Sci*., **1983,** *21* (11) 2403-2417.

9. B. K. Lavine, A. Fasasi, N. Mirjankar, and M. Sandercock, "Search Prefilters to Assist in Library Searching of Infrared Spectra of Automotive Clear Coats" *Talanta*, **2015,** *120*, 182-190.

10. A. Fasasi, N. Mirjankar, R. I. Stoian, C. White, M. Allen, M. P. Sandercock and B. K. Lavine, "Pattern Recognition Assisted Infrared Library Searching of Automotive Clear Coats," *Appl Spec*., **2015,** 69 (1) 84-94.

CHAPTER III

CHEMOMETRICS

This chapter focuses on the multivariate analysis methods used in the investigation of automotive paint smears and clear coats. Specifically, principal component analysis (PCA) [1] and hierarchical clustering [2] were applied to the Raman spectra to identify the sample classes present in the data. Classification of these samples was undertaken using a genetic algorithm. To analyze the infrared images of the automotive paint smears generated using the impactor, multivariate curve resolution methods were applied to the spectral images. A feature common to all these studies is the use of high-level multivariate analysis techniques to extract relevant information embedded in the data, and to better understand how individual samples relate to the data cohort. Identifying similarities is a necessary first step to uncover the underlying structure (relationship between samples and their measurements) that exists in multivariate data. PCA and cluster analysis are two well-known pattern recognition methods used to identify relationships present in multivariate data [3, 4]. After identifying the presence of classes in the data, categorizing each sample as a member of a distinct sample class to predict a property is the next logical step.

A genetic algorithm for pattern recognition and feature selection [5], which identifies spectral features that optimizes the separation of the classes in a plot of the two or three largest principal components of the data, was employed in the studies discussed in this dissertation.

For infrared image analysis, the method of alternating least squares (ALS) [6] was used to deconstruct the spectral image into a score and loading matrix. Principal component analysis is an integral component of ALS as the number of distinct layers comprising the sample is determined by estimating the rank of the data matrix. However, determining the number of significant principal components is often a problem because of accidental correlations between signal and noise. Another problem is that principal component analysis constitutes a purely mathematical solution often devoid of physical or chemical meaning because there are more wavelengths than constituents (i.e., layers of automotive paint). The solution to both problems lies in the development of a suitable rotation, which is the thrust of the ALS method.

3.1 Principal Component Analysis

Principal component analysis is the most widely used multivariate analysis method in science and engineering [7]. PCA attempts to reduce the dimensionality of the data by finding a set of orthogonal axes that represent the directions of greatest variance in the data. The orthogonal axes are called principal components. Each principal component is a linear combination of the original measurement variables. The number of principal components that can be extracted from the data is the smaller of either the number of samples or number

of measurements in the data set as this number defines the largest number of independent axes in our data.

A measure of the amount of information conveyed by each principal component is its variance. (The variance is defined as the degree to which the data points are spread apart or scattered in the n-dimensional measurement space.) For this reason, the principal components are usually arranged in order of decreasing variance. Thus, the most informative principal component is the first and the least informative is the last. Typically, the first two principal components (those with the largest eigenvalues and hence the largest variance) are used to generate a plot representing the n-dimensional subspace which also summarize the information present in the data.

PCA is a powerful method for analyzing the structure of a data set. PCA is performed by a decomposition of the data matrix **X** (n x p) into a score matrix **T** (n x f), loading matrix **P** (f x p), and residual matrix **E** (n x p). (Note that f is smaller than p due to correlations among the measurement variables). The matrix equation for this decomposition is

$$\mathbf{X} = \mathbf{1}\mathbf{x}_{\text{mean}} + \mathbf{TP} + \mathbf{E} \tag{3.1}$$

where **1** is a column vector (n x 1) of ones and $\mathbf{x}_{\text{mean}}$ is a (1 x p) row vector that is the mean or centroid of the data. The score matrix defines the coordinates of the samples in the principal component space, and the loading matrix defines the relationship between the original measurement variables and the principal components. The score and the loading matrix describe the signal in the data, whereas the residual matrix describes the noise. Plotting the columns of the score matrix against each other can reveal the presence of

outliers as well as reveal similarities and differences among groups of samples. Clearly, PCA possesses two key attributes that can ensure a successful analysis of the data: (1) dimensionality reduction and hence simplification of multivariate data, and (2) separation of signal from noise in a data matrix.

Singular value decomposition [8] is the method often used to perform PCA on a data matrix. Usually, the data matrix, **X**, for singular value decomposition is mean centered. This is done by subtracting the mean of the variable from each entry in the column of the data matrix that corresponds to the variable. **X** can be autoscaled which involves adjusting the measurement such that each has a mean of zero and variance of one (see Equation 3.2)

$$x_{i,new} = \frac{x_{i,orig} - m_{i,orig}}{s_{i,orig}} \qquad (3.2)$$

where $m_{i,org}$ is the mean of the original component and $s_{i,org}$ is the standard deviation of the original component. Autoscaling removes inadvertent weighting of the variables that otherwise would occur due to differences in magnitude among the measurements. After autoscaling, all of the measurements have equal weight and, therefore, an equal effect in the analysis.

One would expect that only the first few principal components would convey information about the signal in the data, if the data were collected with due care as most of the information in the data should be about the effect that we seek to study. However, the situation is not always so straightforward. Each principal component describes some amount of signal and some amount of noise in the data because of accidental correlations between signal and noise. The larger principal components (i.e., those with the larger

eigenvalues) primarily describe signal, whereas the smaller principal components (i.e., those with smaller eigenvalues) essentially describe noise. When smaller the principal components are deleted, noise is discarded from the data, but so is a small amount of signal. However, the gain in signal to noise more than compensates for the biased representation of the data that results from discarding principal components that contain a small amount of signal but a large amount of noise. This approach to describing a data set in terms of important and unimportant variation is known as soft modeling in latent variables and is a major thrust of the methodology used for multivariate data analysis in this dissertation.

Principal component analysis takes advantage of the fact that a large amount of data is generated in a pattern recognition study. The data have a great deal of redundancy and therefore a great deal of collinearity. Because the measurement variables are correlated, 100-point spectra do not necessarily require 100 independent axes to define the position of the sample point in the measurement space. Utilizing PCA, the original measurement variables that constitute a correlated axis system can be converted to a new set of basis vectors that remove correlation by forcing the new axes to be independent and orthogonal, a requirement that greatly simplifies the data because the correlations present in the spectral data often allow us to use fewer axes to represent the sample points. In other words, the spectra for a set of automotive paint samples may reside in a subspace of the original 100-dimensional measurement space, and a plot of the two to three largest principal components of the data can help us to visualize the relative position of the samples in this subspace and the relationship among the samples to each other.

3.2 Cluster Analysis

Exploratory data analysis techniques are often helpful in understanding the complex nature of multivariate relationships. In the section above, the advantages of using PCA to understand the structure of a multivariate data set were presented. In this section, some additional techniques will be discussed which also give insight into the intrinsic structure of the data set. These methods are based on finding clusters of points in the data, hence the term cluster analysis.

Clustering methods [9] attempt to determine the structural characteristics of a data set by organizing the data into subgroups or clusters. These methods are based on the following principle: the distance between pairs of points (i.e., samples) in the measurement space is inversely related to their degree of similarity. Although several different types of clustering algorithms exist, by far the most popular is hierarchical clustering [10]. This particular algorithm works by computing the distances between all pairs of points in the data set, identifying the nearest pair, combining them into a new point which is located midway between the two original points and recalculating the distances from this new point to every other point in the data set. This process is continued until all points have been linked. The result of this procedure is a diagram called a dendrogram which can be analyzed for clusters using a variety of criteria. As these criteria can be subjective, the interpretation of the data structure on the basis of the dendrogram often depends upon the criteria used.

There are several approaches for computing distances between a data point and a cluster of points, and these define the different types of hierarchical clustering methods.

Nearest linkage calculates the similarity between a point and a cluster of points by computing the distance to the closest point in the cluster. Farthest linkage determines similarity by calculating the distance to the cluster's farthest point. For mean linkage, the distances between all point pairs (a point and each point in the cluster) are computed and the mean of these distances is used to define similarity.

3.3 Genetic Algorithm for Pattern Recognition and Variable Selection

So far, only unsupervised pattern recognition techniques have been discussed in this chapter. These methods analyze the structure of the data without directly incorporating *a priori* information about the class membership of the sample in the algorithm. However, the overall goal of a pattern recognition study is the development of a classification rule that can accurately predict the class membership of an unknown sample [11]. For the IR and Raman automotive paint studies discussed in this dissertation, discriminants for the classification of IR and Raman spectra were developed using a genetic algorithm (GA) for pattern recognition and feature selection [12-16], which selects features to optimize the separation of the classes in a plot of the two or three largest principal components of the data. Because the largest principal components capture the bulk of the variance in the data, the features chosen by the pattern recognition GA will usually contain information primarily about the differences between the classes in the data set. The principal component analysis routine embedded in the fitness function of the pattern recognition GA acts as an information filter, significantly reducing the size of the search space as it restricts the search to features whose principal component plots show clustering on the basis of class. The pattern recognition GA focuses on those classes and/or samples that are difficult to classify as it trains by boosting the weights associated with each sample and class in the

training set. Samples that consistently classify correctly are not as heavily weighted as samples that are more difficult to classify. Over time, the algorithm learns its optimal parameters in a manner similar to a neural network. The pattern recognition GA integrates aspects of artificial intelligence and evolutionary computations to yield a smart one pass procedure for feature selection and classification.

During each generation, the pattern recognition GA computes both the class and sample weights (which are an integral component of the fitness function) to facilitate the tracking and scoring of the principal component plots generated for each chromosome (i.e., feature subset) comprising the population, see Equations 3.3 and 3.4 where CW(c) is the weight of class c (with c varying from 1 to the total number of classes in the data set) and $SW_c(s)$ is the weight of sample s in class c. Class weights always sum to 100, and the sample weights comprising a particular class always sum to a value equal to the weight of the class in question.

$$\mathrm{CW(c)}=100\frac{\mathrm{CW(c)}}{\sum_{\mathrm{c}}\mathrm{CW\ (c)}} \tag{3.3}$$

$$SW(s) = CW(c)\frac{SW(s)}{\sum_{s\in c} SW(s)} \tag{3.4}$$

A principal component plot is scored using the k-nearest neighbor (K-NN) classification algorithm [17]. For a given point, Euclidean distances are computed between it and every other point in the plot. These distances are arranged from smallest to largest, and a poll is taken of the point's k nearest neighbors. For the most rigorous classification, k equals the number of samples in the class to which the point belongs. The sample hit count (SHC), or the number of like nearest neighbors is $0 \leq SHC(s) \leq K_c$ is calculated, and the fitness is computed using Equation 3.5.

$$F(d) = \sum_{c} \sum_{s \in c} \frac{1}{K_c} \times SHC(s) \times SW(s) \qquad (3.5)$$

To better understand how a principal component plot is scored (see Equation 3.5), consider a data set consisting of two classes. Class 1 has 50 samples, and class 2 has 10 samples. At generation 0, the samples in a given class have the same weight. Thus, each sample in class 1 has a sample weight of 1, whereas each sample in class 2 has a weight of 5. Suppose a sample from class 1 has as its nearest neighbors eight class 1 samples. Hence, SHC/K = 0.8, and (SHC/K) *SW = 0.8*5, which equals 4. By summing (SHC/K_c) *SW for each sample, each principal component plot is scored. One advantage of using this procedure to score the principal component plots is that a class with a large number of samples does not dominate the score because of the class weights.

The fitness function of the pattern recognition GA is able to focus on those samples and classes that are difficult to classify by boosting their class and sample weights over

successive generations. For boosting, it is necessary to calculate both the sample-hit rate (SHR), which is the mean value of SHC/K_c over all feature subsets (i.e., chromosomes) produced in a particular generation (see Equation 3.6), and the class-hit rate (CHR), which is the mean sample hit rate of all samples in a class (see Equation 3.7). ϕ in Equation 3.6 is the number of chromosomes in the population, and AVG in Equation 3.7 refers to the average or mean value. During each generation, class and sample weights are adjusted using a perceptron (see Equations 3.8 and 3.9) with the momentum, P, set by the user. (g + 1 is the current generation, whereas g is the previous generation.) Classes with a lower score are boosted more than those classes with a higher score. Boosting is crucial for the successful operation of the pattern recognition GA as the fitness function is modified by adjusting the values of the class and sample weights during the run. Boosting minimizes the problem of convergence to a local optimum. Hence, the fitness function of the pattern recognition GA changes as the population of chromosomes evolves towards a solution.

$$SHR(s) = \frac{1}{\emptyset}\sum_{i=1}^{\emptyset}\frac{SHC_i(s)}{K_c} \tag{3.6}$$

$$CHR_g(c) = AVG\left(SHR_g(s): \forall_{s\in c}\right) \tag{3.7}$$

$$CW_{g+1}(s) = CW_g(s) + P(1 - CHR_g(s))) \tag{3.8}$$

$$SW_{g+1}(s) = SW_g(s) + P(1 - CHR_g(s))) \tag{3.9}$$

3.4 Alternating Least Squares

ALS, which can resolve the IR spectra of individual components in a mixture that are spectral and/or spatial in nature, was used in the paint smear study to analyze the infrared images. Each IR spectrum in the image comprised a data vector, and the data vectors, in turn, were organized into a data matrix, X. ALS decomposes X into three matrices (see Equation 3.10), where C is the concentration profile of each component, S is the corresponding IR spectrum of the component, and E is the residual matrix. Equation 3.10 is solved iteratively using Equations 3.11 and 3.12. To perform ALS, the first step is to provide an initial estimate of C. Using this initial estimate, an estimate of S is computed (see Equation 3.11). Using the estimate of S, an improved estimate of C is computed (see Equation 3.12). From the product of C and S, an estimate of the principal component analysis (PCA) reproduced data matrix, X_{PCA}, is calculated. This process is repeated until the algorithm converges. The constraints assigned to ALS in the paint smear study to facilitate both convergence and accuracy were nonnegative concentration and nonnegative absorbance values for C and S. Furthermore, the concentration profile of each component was constrained to be unimodal. To obtain an initial estimate of C for ALS, a procedure known as the varimax extended rotation (VER) previously developed by our research group to identify the components of an oil in water emulsion from Raman imaging data and to resolve severely overlapped chromatographic peaks obtained using a liquid chromatograph equipped with a diode array detector was applied to the IR spectra comprising the line map [18, 19]. For ALS, using the entire spectral range (4000 cm^{-1} to 748 cm^{-1}) was more effective then selecting a wavelength region, e.g., the fingerprint region, as more information about how each layer changes as a function of the sample

position during scanning was obtained when the entire spectrum was subject to decatenation.

$$\mathbf{X} = \mathbf{C}\mathbf{S}^{T} + \mathbf{E} \tag{3.10}$$

$$\hat{S}^{T} = (\mathbf{C}^{T}\mathbf{C})^{-1}(\mathbf{C}^{T}\mathbf{X}_{PCA}) \tag{3.11}$$

$$\hat{C} = (\mathbf{X}_{PCA}\mathbf{S}^{T})(\mathbf{S}\mathbf{S}^{T})^{-1} \tag{3.12}$$

References

1. De Juan, A. E. Casassas, and R. Tauler. "Soft modeling of analytical data" in Encyclopedia of analytical chemistry: Applications, theory, and instrumentation. New York: Wiley., 2000 Vol. 11 9800-9837.

2. B. K. Lavine, and N. Mirjankar. “Clustering and classification of analytical data” in Encyclopedia of Analytical Chemistry, John Wiley, and Sons Ltd. 2012.

3. I. T. Jolliffee, “Principal Component Analysis” Springer-Verlag, New York, 1986.

4. D. L. Massart, and L. Kaufman, “The Interpretation of Analytical Chemical Data by the Use of Cluster Analysis” John Wiley & Sons, New York, 1983.

5. R. G. Brereton, “Multivariate Pattern Recognition in Chemometrics” Elsevier, Amsterdam, 1992.

6. S. C Rutan, A. D. Juan, and R. Tauler, “Introduction to Multivariate Curve Resolution” Brown SD, Tauler R, Walczak B, (Ed) Comprehensive Chemometrics, Elsevier., 2009 Vol. 2 249-260.

7. S. D. Brown, “Chemical Systems Under Indirect Observation: Latent Properties and Chemometrics” *Appl Spec*., **1995**, *49* 14A-31A.

8. G. Golub, and C. Van, “Matrix Computations” 1971, Baltimore: Johns Hopkins University Press.

9. P. Geladi, “Chemometrics in Spectroscopy Part 1. Classical Chemometrics” *Spectrochim. Acta Part B.,* **2003**, *58* (5), 767-782.

10. L. Kaufman and P. J. Rousseeuw, “Finding Groups in Data – An Introduction to Cluster Analysis” 1990, NY John Wiley & Sons.

11. J.H. Holland, “Adaptation in Natural and Artificial Systems” 6th edition, MIT Press, Cambridge, MA, 2001.

12. B. K. Lavine, C.G. White, T. Ding, M.M. Gaye, and D.E. Clemmer, "Wavelet Based classification of MALDI-IMS-MS Spectra of Serum N-Linked Glycans from Normal Controls and Patients Diagnosed with Barrett's Esophagus, High Grade Dysplasia, and Esophageal Adenocarcinoma" *Chemolab*., **2018**, *176*, 74-81.

13. M. M. Gaye, T. Ding, H. Shion, A. Hussein, Y. HU, S. Zhou, Z. T. Hammoud, B. K. Lavine, Y. Mechref, J. C. Gelbler, and D. E. Clemmer, "Delineation of Disease Phenotypes Associated with Esophageal Adencarcinoma by MALDI-IMS-MS Analysis of Serum N-linked Glycans" *Analyst*., **2017**, *142*, 1525-1535.

14. B. K. Lavine, A. Fasasi, N. Mirjankar, M. Sandercock, and S. D. Brown, "Search Prefilters for Mid-IR Spectra of Clear Coat Automotive Paint Smears using Stacked and Linear Classifiers" *J. Chemom*., **2014**, *28*, 385-39.

15. B. K. Lavine, K. Nuguru, N. Mirjankar, and J. Workman, "Pattern Recognition Assisted Infrared Library Searching," *Appl. Spec*., **2012**, *66* (8), 917-925.

16. B. K. Lavine, K. Nuguru, and N. Mirjankar, "One Stop Shopping - Feature Selection, Classification, and Prediction in a Single Step," *J. Chemom*., **2011**, *25*, 116-129.

17. A. Kataria, and M. D. Singh, "A review of data classification using k-nearest neighbour algorithm," *International Journal of Emerging Technology and Advanced Engineering*., **2013**, *3* (6), 354-360.

18. B.K. Lavine, C.E. Davidson, J.P. Ritter, D. Westover, and T. Hancewicz. "Varimax Extended Rotation Applied to Multivariate Spectroscopic Image Analysis". *Microchem. J*., **2004**, *76* (1): 173–180.

19. B.K. Lavine, J.P. Ritter and E. Voigtman. "Multivariate Curve Resolution in Liquid Chromatography: Resolving Two Way Multicomponent Data Using a Varimax Extended Rotation". *Microchem. J*., **2002**, *72* (2): 163–178.

CHAPTER IV

ANALYSIS OF AUTOMOTIVE PAINT SMEARS USING ATTENUATED TOTAL REFLECTION INFRARED MICROSCOPY

4.1 Introduction

In the forensic examination of automotive paint, each layer of the paint is visually and chemically analyzed. Forensic chemists have utilized a variety of methods to analyze the chemical composition of automotive paint. In North America, chemical analysis of automotive paint is typically done using FTIR spectroscopy, which relies on the selective absorption of infrared (IR) radiation by the constituents in each paint layer. For FTIR analysis, each paint layer is separated and placed between two diamond anvil cells [1]. The layer structure, color, and composition of the automotive paint recovered from the vehicle or clothing of the victim can serve as a fingerprint either linking or excluding suspect vehicles from association with the crime scene.

In many cases, automotive paint evidence obtained from vehicle-vehicle collisions or vehicle related fatalities such as a hit-and-run where injury or death to a pedestrian has occurred can result in the physical smearing of the car paint. Because it is not possible to physically separate the individual layers of the paint comprising the smear, the forensic chemist is not able to develop an investigative lead from the paint evidence recovered at the crime scene. The use of attenuated total reflection (ATR) infrared microscopy for the analysis of paints and the development of forensic standards for paint smear analysis using an impactor to simulate vehicle-vehicle collisions will significantly improve the accuracy of paint smear analysis and make possible the creation of realistic proficiency testing for forensic laboratories, which could also be used to train forensic scientists who are using forensic automotive paint databases such as PDQ.

4.2 Generation of Automotive Paint Smears

Using the impactor (discussed in Chapter II), the parameters controlling the impactor were systematically varied to simulate paint smears representative of a vehicle-vehicle collision. The tunable shock absorbers were adjusted using a soft spring (SS) or hard spring (HS) to control the force constant of the spring and the damping factor (DF) was varied from 1 to 8 (1 being the weakest and 8, the strongest) to control the overall force of the collision. With the appropriate parameters adjusted for each experiment, the automotive paint sample (including the automotive substrate to which the paint is bound) was cleaned with methanol to avoid interference from particulate matter and mounted on the upper wedge. The next step is releasing the upper wedge at a specified height, which then accelerates due to gravity colliding with the lower wedge to generate the paint smear,

which is deposited on the transfer substrate of the lower wedge. The transfer substrate is detached from the lower wedge and transferred to the microscope stage for analysis by an ATR-IR microscope. Each IR spectrum in the spectral map of the paint smear is baseline corrected by fitting a line between specified baseline points using routines in OMNIC. The baseline corrected IR spectra comprising the map are subject to multivariate curve resolution using Alternating Least Squares (ALS) (described in Chapter III) to obtain spectra representative of each layer recovered from the paint smear. Each ALS reconstructed IR spectrum was matched against an in-house IR spectral library using OMNIC. As the original in-house spectral library consisted of transmission IR spectra collected from a high-pressure diamond cell, it was necessary to convert these spectra into ATR spectra for library matching. The algorithm used to perform this conversion is described in detail in the next section.

Twenty-four OEM automotive paint samples were used to prepare thirty-one automotive paint smears. The OEM paint samples were from three automotive manufactures (General Motors, Chrysler, and Toyota) and ten assembly plants (Arlington, Belvidere, Flint, Fremont, Janesville, Orion, Linden, Lordstown, Silao, and Oshawa) (See Table 4.1). The number of layers recovered from the paint smears after analysis by ATR-IR microscopy and ALS varied from one layer (clear coat layer) to four layers (clear coat, color coat, surfacer-primer, and e-coat), see Table 4.2, depending upon the conditions (i.e., collisional parameters) used.

Table 4.1 OEM Automotive Paint Samples used in Smears Study

PDQ Number	Manufacturer	Assembly Plant	Make	Line	Year	Vehicle
UACD00166	General Motors	Oshawa	Chevrolet	Camaro	2010	Car
UARL00071	General Motors	Orion	Pontiac	PG6	2005	Car
UAZP00201	Chrysler	Belvidere	Dodge	Neon	2000	Car
UAZP00329	General Motors	Oshawa	GMC	Sierra	2001	Truck
UAZP00331	General Motors	Silao	Chevrolet	Suburban	2001	SUV
UAZP00332	General Motors	Silao	GMC	Yukon XL	2001	SUV
UAZP00334	General Motors	Flint	Chevrolet	Silverado	2001	Truck
UAZP00338	General Motors	Janesville	Chevrolet	CTA	2001	SUV
UAZP00390	General Motors	Linden	Chevrolet	Blazer	2001	SUV
UAZP00433	General Motors	Flint	Chevrolet	Silverado	2002	Truck
UAZP00434	General Motors	Janesville	Chevrolet	Suburban	2002	SUV
UAZP00494	General Motors	Silao	Chevrolet	Suburban	2003	SUV
UAZP00495	General Motors	Silao	Chevrolet	Suburban	2003	SUV
UAZP00496	General Motors	Silao	Chevrolet	Suburban	2002	SUV
UAZP00499	General Motors	Orion	Pontiac	Bonneville	2003	Car
UAZP00502	General Motors	Arlington	Chevrolet	CTA	2003	Truck
UAZP00561	Toyota	Fremont	Toyota	Tacoma	2005	Truck
UMOJ00105	General Motors	Oshawa	Chevrolet	Silverado	2003	Truck
UNJH00056	General Motors	Oshawa	Buick	Century	2001	Car
UNJH00057	General Motors	Oshawa	Chevrolet	Impala	2005	Car
UNVL00007	General Motors	Fairfax	Pontiac	Grand Am	2003	Car
UVAC00168	General Motors	Lordstown	Chevrolet	Cavalier	2002	Car
UWVC00112	General Motors	Linden	Chevrolet	S10	2001	Truck
UWVC00200	General Motors	Oshawa	Chevrolet	Impala	2002	Car

Table 4.2 Paint Smears Recovered from Infrared Analysis

Number of Paint Layer(s) in a Smear Sample	Number of Paint Smears Samples Generated
1 (clear coat)	10
2 (clear coat and surfacer-primer)	5
3 (clear coat, color coat, surfacer primer or e-coat)	4
4 (clear coat, color coat, surfacer primer, e-coat)	12
Total	31

4.3 ATR Correction Algorithm

The FTIR spectra collected from the iN10-MX microscope are ATR spectra. The optical configuration used to collect ATR spectra differs from transmission spectra in our in-house spectra library. Directly comparing ATR spectra to transmission spectra in a

library has often proved problematic because of differences in the band shapes, absorbances, and peak frequencies of ATR spectra. This problem has been addressed by converting the in-house transmission IR spectral library to an ATR spectra library using an ATR simulation algorithm developed by Lavine and coworkers [2].

The conversion is accomplished by considering the surface reflection processes that occur at the boundary between the internal reflecting element (IRE) and the sample. Fresnel's equations are used to describe the reflection of the incident beam from the IRE that interacts with the sample. The following parameters are required to generate the ATR spectra: n and k-indices of the sample (automotive paint), refractive index of the IRE, thickness of the sample, incident angle of the beam, and the number of internal reflections. Assigning values to these parameters and then inserting them into the Fresnel's equations, the reflectance due to s and p-polarized light is calculated. The average reflectance is then computed. To compute the reflectance of the s- and p- polarized light, the optical constants of the sample (i.e., the n- and k-indices) at each wavelength are computed from the transmission spectrum using Equations 4.1 and 4.2, where A(ν) is the absorbance value of the transmission spectrum as a function of wavenumber (ν), d is the sample thickness, P is the principal value of the integral and n (∞) is the refractive index at high wavenumber where there is no absorbance, which serves as an anchor value [3, 4]. The Kramers-Kronig equation (see Equation 4.2) conveys the complex refractive index (as it depends on both the real and imaginary components).

Using the complex refractive index and the computed values of the n- and k-indices, Equations 4.3 and 4.4 are applied to calculate the Fresnel's reflection coefficients [5] for the s- and p-polarized light at each wavelength, where n_0 and $\tilde{n}_1$ are the complex refractive

indices of the IRE and the sample, and θ_0 and θ_1 are incident angle of the beam in the IRE and in the sample. Finally, Equation 4.5 can be applied to calculate the reflectance at each wavelength in the ATR spectrum. To apply the ATR correction algorithm, the angle of incidence of the IR beam is also required. However, the angle of the incident beam of the Cassegrain reflectors and aperture size vary. Hence, the effective incident angle for each point in the spectrum is a range of values instead of single value. This range (47^0 - 49^0) was determined in a previous study [6]. The appropriate angle was determined by varying the minimum and maximum incident angles in the ATR correction algorithm until a representative spectrum identical to the reference clear coat spectra obtained on an iS50 FTIR spectrometer (Thermo-Nicolet) using a germanium ATR accessory was obtained.

$$k(\nu) = \frac{2.303A(\nu)\lambda}{4\pi d} \; or \; k(\nu) = \frac{2.303A(\nu)}{4\pi\nu d} \qquad (4.1)$$

$$n(\nu_a) = n(\infty) + \frac{2}{\pi} P \int_0^\infty \frac{\nu k(\nu)}{\nu^2 - \nu_a^2} d\nu \qquad (4.2)$$

$$r_s = \frac{n_0 cos(\theta_0) - \tilde{n}_1 cos(\theta_1)}{n_0 cos(\theta_0) + \tilde{n}_1 cos(\theta_1)} \qquad (4.3)$$

$$r_p = \frac{\tilde{n}_1 cos(\theta_0) - n_0 cos(\theta_1)}{\tilde{n}_1 cos(\theta_0) + n_0 cos(\theta_1)} \qquad (4.4)$$

$$R = \frac{(R_p + R_s)}{2} = (\,|r_p|^2 + |r_s|^2)/2 \qquad (4.5)$$

4.4. Analysis of Automotive Paint Smears from Spectral Maps

4.4.1 One Layer Recovered

For ten of the 31 automotive paint smears generated (see Table 4.3), only the clear coat layer was recovered from the smear by ALS. This suggests that only the clear coat layer was transferred from the OEM automotive paint sample mounted on the upper wedge to the transfer substrate. The ALS reconstructed IR spectra for these ten paint smear samples were searched against the in-house spectral library using OMNIC with search type set as correlation and apodization set to Happ-Genzel. Only the fingerprint region i.e., 1641 cm^{-1} to 860 cm^{-1} was used for the search as the region below 860 cm^{-1} is too noisy and the C=O stretch at 1730 cm^{-1} has been previously shown to be uninformative for discriminating among vehicle manufacturer and assembly plant [7, 8].

To ascertain whether a library search has been successful, the hit quality index (HQI) can be used as the tool for assessment. A library search was judged to be successful when the actual paint sample or a paint sample of the same make and model of the vehicle from which the paint sample originated is included in the top five hits of the hit-list generated for the search. Each search was restricted to a library that only contained IR spectra from a single layer (e.g., clear coat) and from the same manufacturer (e.g., General Motors) as the automotive paint sample within a narrow production year range (2000 – 2006). We selected the top five hits as our previous experience has shown that a search of the in-house spectral library has the potential to capture the correct line and model of the vehicle when the hit-list is limited to the top five hits. The ALS reconstructed IR spectrum of the clear coat layer from the paint smear compared well to the corresponding IR

spectrum of the same sample in the in-house General Motors library used to match the samples. Each match was a first, second or third hits with the hit quality index value greater than 95 percent for 9 of the 10 paint smears (See Table 4.4).

Figures 4.1 to 4.10 show plots of the ALS reconstructed spectra overlaid on the actual paint sample which corresponded to a top hit and are also present in the in-house spectral library, which consists of 643 clear coat spectra spanning 26 assembly plants for the years 2000 -2006. Only UAZP00329 has a percentage match less than 95 percent because the match between these two spectra in the 1300 – 900 cm^{-1} region is poor. Because the conditions used to generate the smear were extreme (hard spring, damping factor = 6), the paint chemistry of the sample may have been altered due to the heat generated in the simulated vehicle to vehicle collision.

Table 4.3 Paint Smear Samples with One Layer (Clear Coat)

PDQ Number	Manufacturer	Assembly Plant	Make	Line	Year	Vehicle
UARL00071	General Motors	Orion	Pontiac	PG6	2005	Car
UAZP00329	General Motors	Oshawa	GMC	Sierra	2001	Truck
UAZP00338	General Motors	Janesville	Chevrolet	CTA	2001	SUV
UAZP00390	General Motors	Linden	Chevrolet	Blazer	2001	SUV
UAZP00434	General Motors	Janesville	Chevrolet	Suburban	2002	SUV
UAZP00495	General Motors	Silao	Chevrolet	Suburban	2003	SUV
UAZP00499	General Motors	Orion	Pontiac	Bonneville	2003	Car
UAZP00502	General Motors	Arlington	Chevrolet	CTA	2003	SUV
UNJH00056	General Motors	Oshawa	Buick	Century	2001	Car
UNJH00057	General Motors	Oshawa	Chevrolet	Impala	2005	Car

Table 4.4 Collison Parameters and Library Matching Results for Paint Smear Samples with Only One Paint Layer (Clear Coat)

PDQ Number	Manufacturer	Collision Parameter	Smear Layer Recovered	HQI	% Match
UARL00071	General Motors	SS/DF-8	Clear Coat	1	97.33
UAZP00329	General Motors	HS/DF-6	Clear Coat	3	91.65
UAZP00338	General Motors	HS/DF-8	Clear Coat	2	98.53
UAZP00390	General Motors	HS/DF-7	Clear Coat	1	98.08
UAZP00434	General Motors	HS/DF-2	Clear Coat	1	96.85
UAZP00495	General Motors	HS/DF-1	Clear Coat	2	97.94
UAZP00499	General Motors	SS/DF-8	Clear Coat	1	97.45
UAZP00502	General Motors	SS/DF-7	Clear Coat	1	97.41
UNJH00056	General Motors	HS/DF-6	Clear Coat	3	95.68
UNJH00057	General Motors	SS/DF-1	Clear Coat	1	96.38

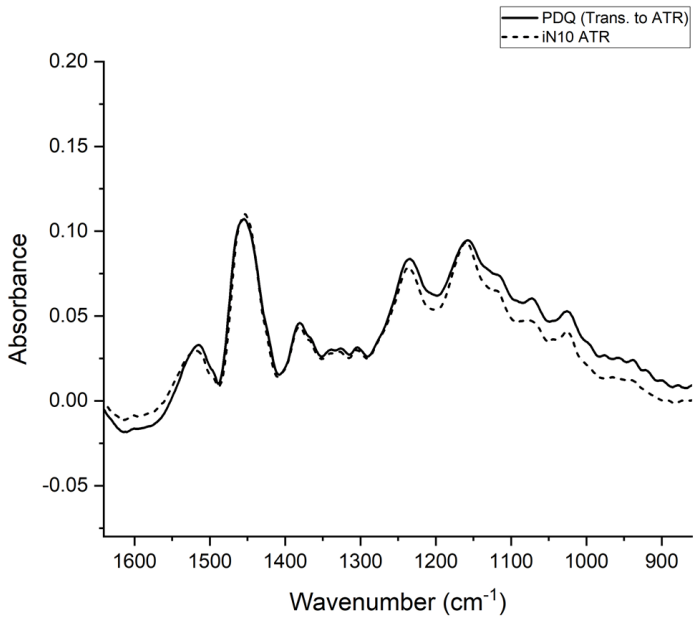

Figure 4.1 Comparison of ALS reconstructed clear coat ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UARL0071 -General Motors) in the PDQ spectral library.

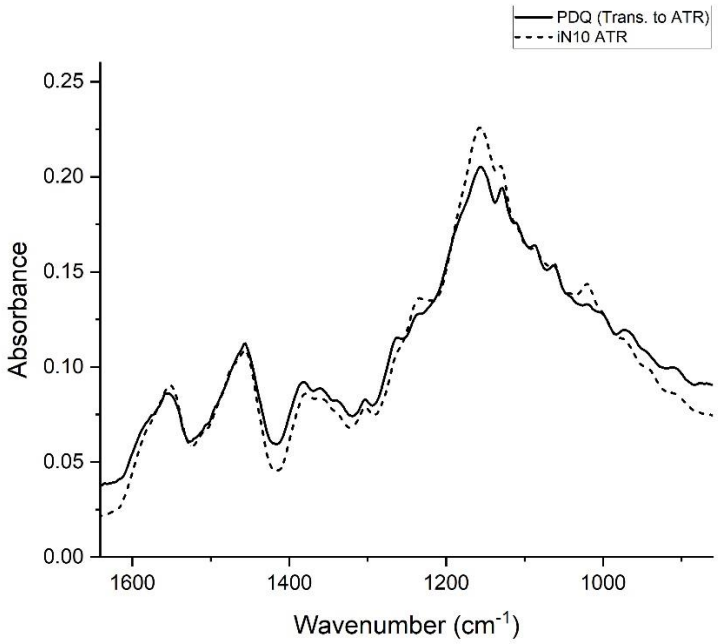

Figure 4.2 Comparison of ALS reconstructed clear coat ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UAZP00329 -General Motors) in the PDQ spectral library.

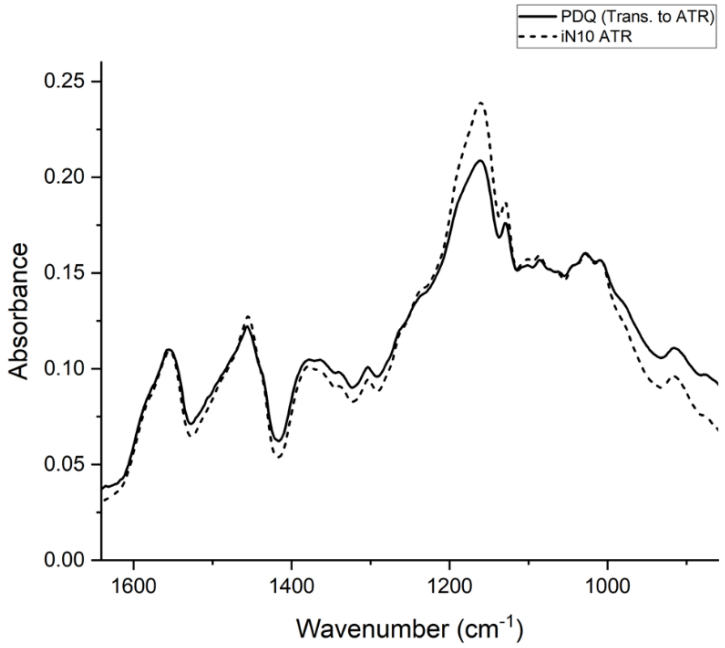

Figure 4.3 Comparison of ALS reconstructed clear coat ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UAZP00338 -General Motors) in the PDQ spectral library.

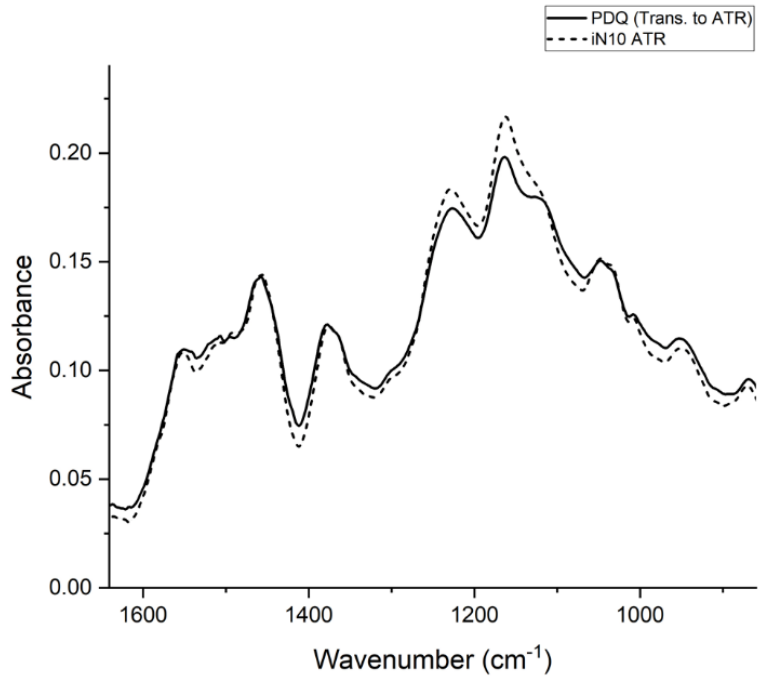

Figure 4.4 Comparison of ALS reconstructed clear coat ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UAZP00390 -General Motors) in the PDQ spectral library.

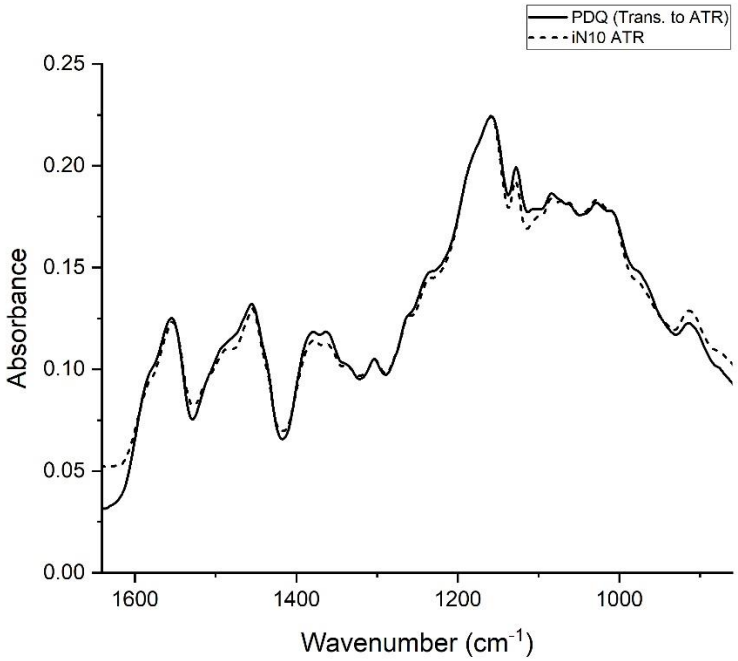

Figure 4.5 Comparison of ALS reconstructed clear coat ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UAZP00434 -General Motors) in the PDQ spectral library.

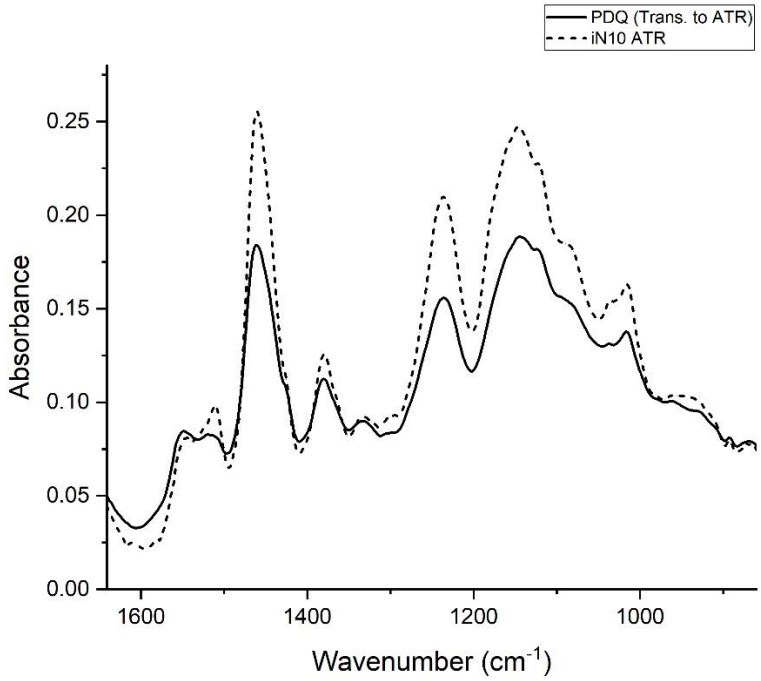

Figure 4.6 Comparison of ALS reconstructed clear coat ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UAZP00495 -General Motors) in the PDQ spectral library.

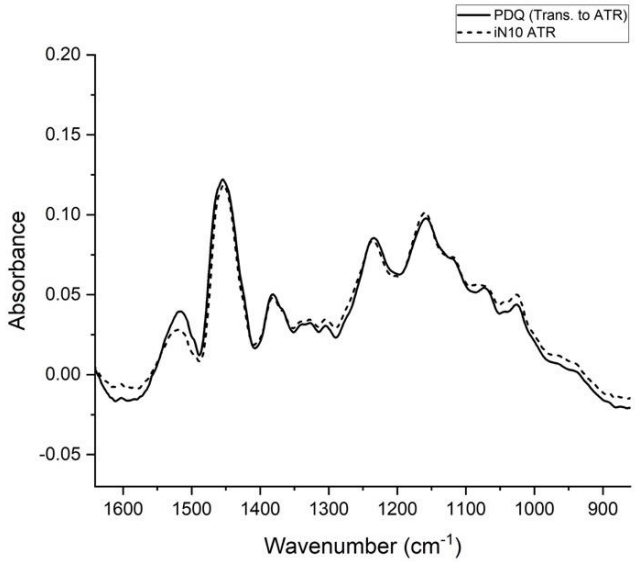

Figure 4.7 Comparison of ALS reconstructed clear coat ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UAZP00499 -General Motors) in the PDQ spectral library.

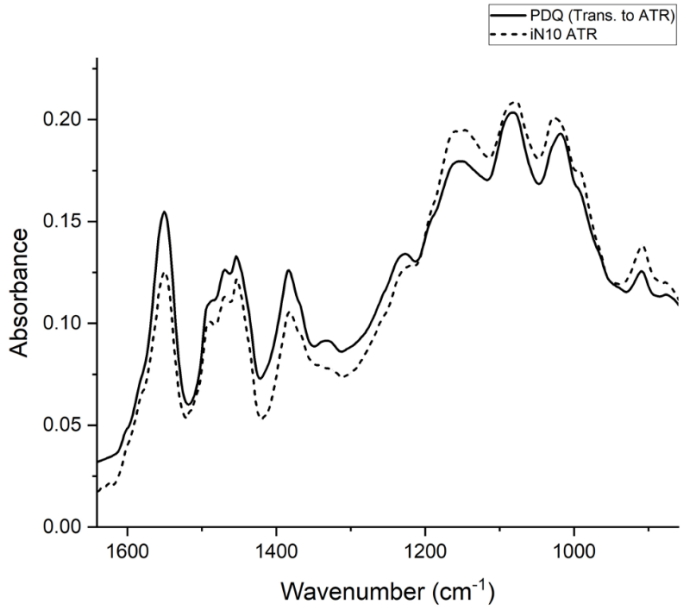

Figure 4.8 Comparison of ALS reconstructed clear coat ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UAZP00502 -General Motors) in the PDQ spectral library.

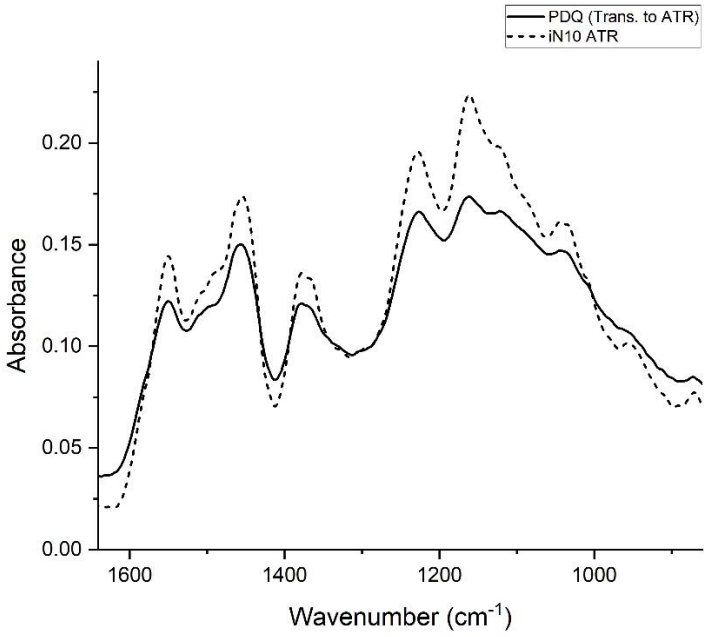

Figure 4.9 Comparison of ALS reconstructed clear coat ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UNJH00056 -General Motors) in the PDQ spectral library.

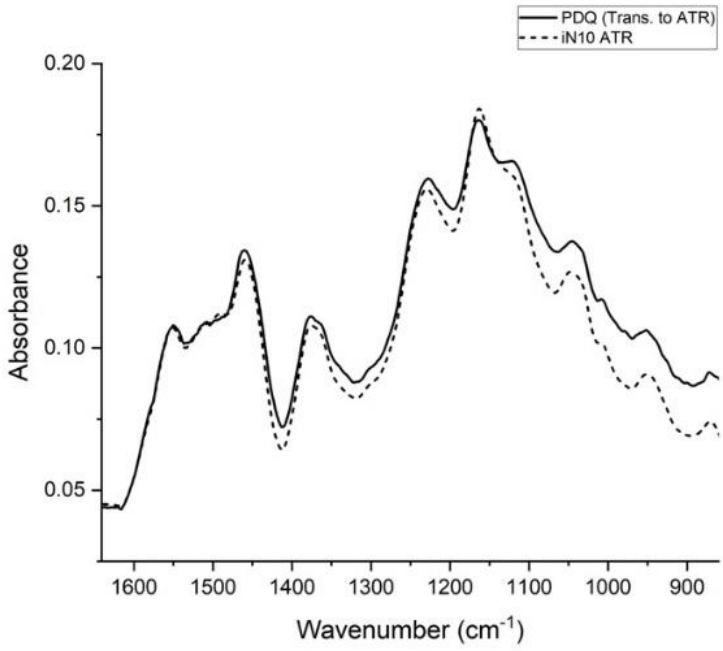

Figure 4.10 Comparison of ALS reconstructed clear coat ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UNJH00057 -General Motors) in the PDQ spectral library.

4.4.2 Two Layers Recovered.

For five of the 31 paint smears (see Table 4.5), two layers were recovered from each smear by ALS. The two layers recovered were the clear coat and the surfacer primer layer. The ALS reconstructed spectra for the clear coat layer and surfacer primer layer were matched against their respective in-house spectral library. (For both the General Motors clear coat and surfacer primer libraries, there were 643 spectra spanning twenty-six assembly plants for the years 2000-2006.) Table 4.6 summarizes the library search results for the paint smear samples that resulted in two paint layers. Figures 4.11 to 4.15 show plots of the ALS reconstructed IR spectra that are overlaid on the library sample of the same make and model as the smear and are also a top five hit. The HQI value of each match was greater than 95% with the exception of the surfacer-primer layer of UNJH00057 which was only 90.72 percent. An examination of the smear and the IR library spectrum corresponding to the third hit showed poor agreement between these two spectra in the region of 1100 – 950 cm^{-1}. As this spectral region corresponds to the problematic spectral region in UNJH00057, it is plausible that heat generated during the simulated collision is responsible for a change in the paint chemistry as reflected by the IR spectrum of this layer.

Table 4.5 Paint Smear Samples with Two Layers (Clear Coat and Surfacer Primer).

PDQ Number	Manufacturer	Assembly Plant	Make	Line	Year	Vehicle
UAZP00390	General Motors	Linden	Chevrolet	Blazer	2001	SUV
UAZP00433	General Motors	Flint	Chevrolet	Silverado	2002	Truck
UNJH00057	General Motors	Oshawa	Chevrolet	Impala	2005	Car
UWVC00112	General Motors	Linden	Chevrolet	S10	2001	Truck
UWVC00200	General Motors	Oshawa	Chevrolet	Impala	2002	Car

Table 4.6 Collison Parameters and Library Matching Results for Paint Smear Samples with Two Layers (Clear Coat and Surfacer Primer)

PDQ Number	Manufacturer	Collision Parameter	Clear Coat		Surfacer Primer	
			HQI	% Match	HQI	% Match
UAZP00433	General Motors	HS/DF-5	1	96.31	1	97.53
UNJH00057	General Motors	SS/DF-8	5	98.20	3	90.75
UWVC00112	General Motors	HS/DF-8	5	93.72	4	96.51
UWVC00200	General Motors	HS/DF-5	1	96.83	1	96.36

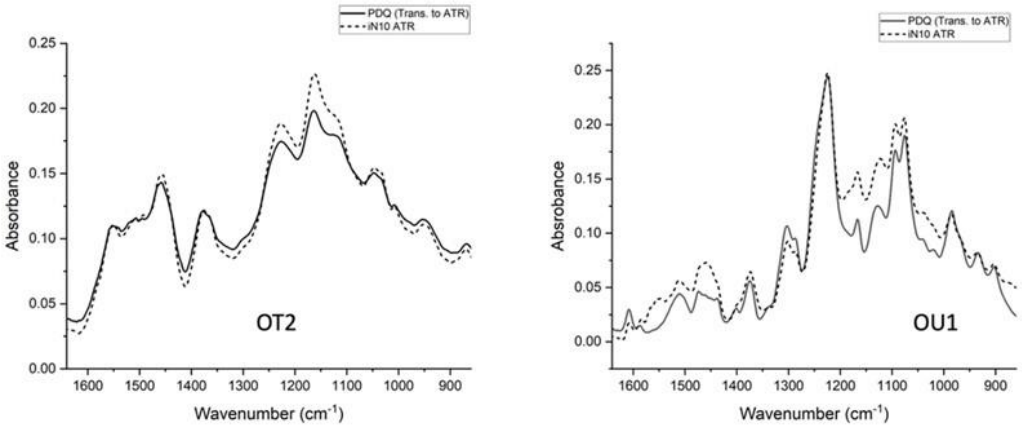

Figure 4.11 Comparison of ALS reconstructed ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UAZP00390 -General Motors) in the PDQ spectral library.

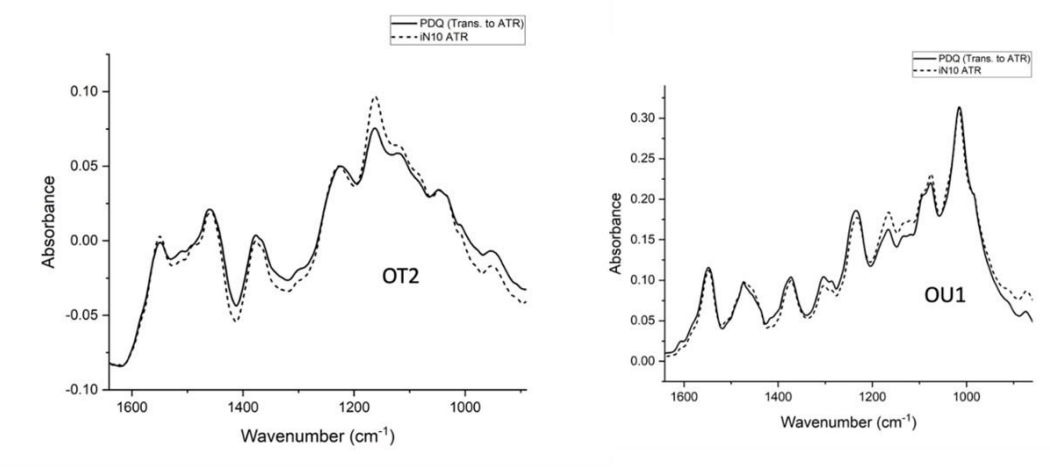

Figure 4.12 Comparison of ALS reconstructed ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UAZP00433 -General Motors) in the PDQ spectral library.

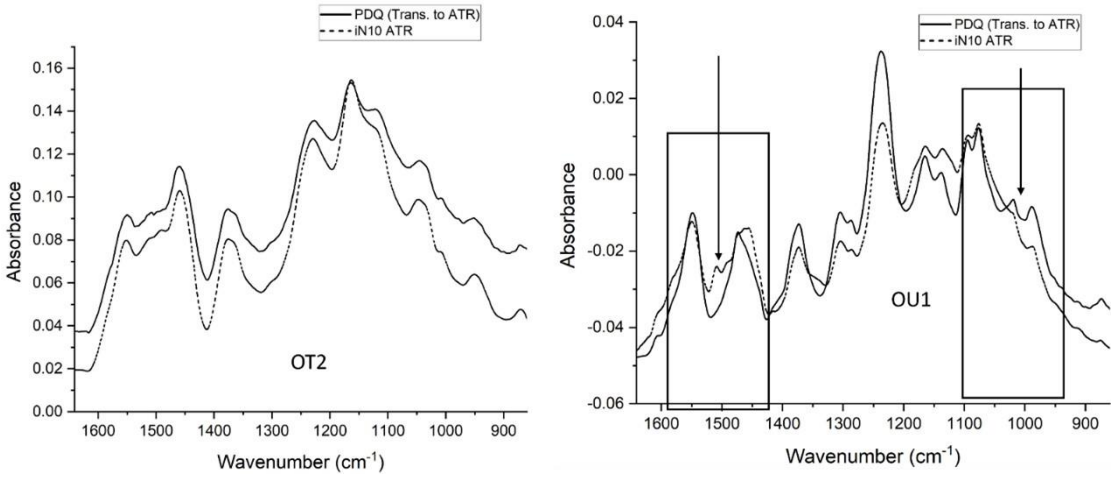

Figure 4.13 Comparison of ALS reconstructed ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UNJH00057 -General Motors) in the PDQ spectral library (Variations in OU1 spectra are highlighted).

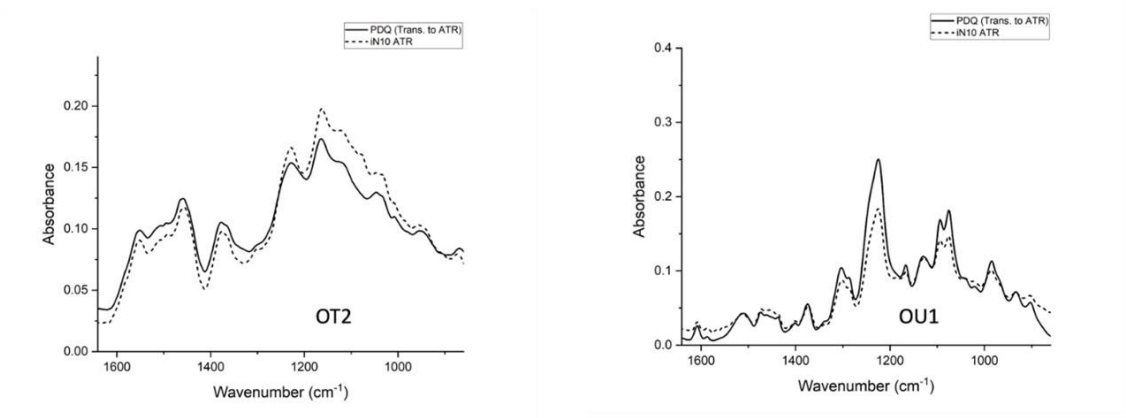

Figure 4.14 Comparison of ALS reconstructed ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UWVC00112 -General Motors) in the PDQ spectral library.

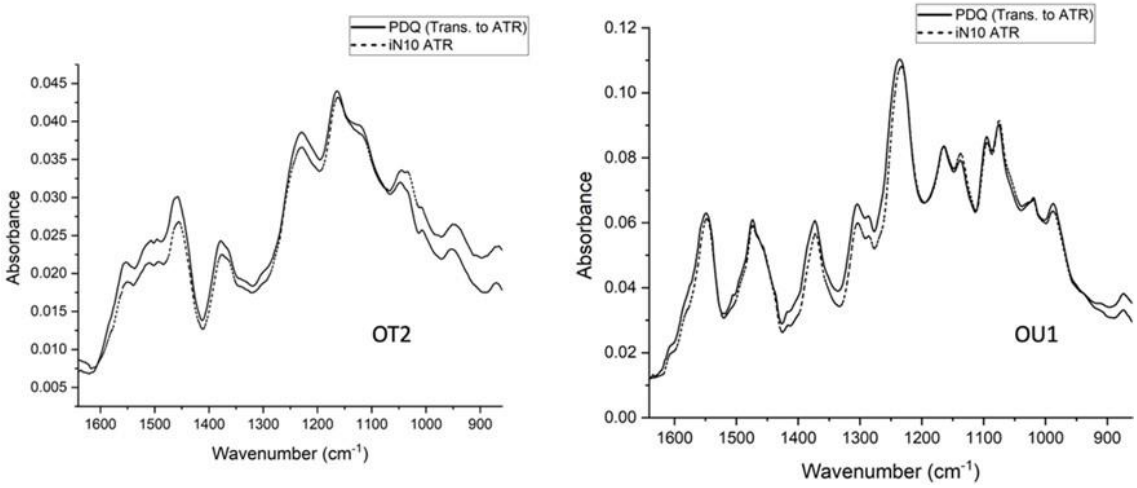

Figure 4.15 Comparison of ALS reconstructed ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UWVC00200 -General Motors) in the PDQ spectral library.

4.4.3 Three Layers Recovered.

For four of the thirty-one paint smears (see Table 4.7), three layers were recovered from each smear by ALS. The three layers recovered were the clear coat, color coat and the surfacer-primer or e-coat layer. The ALS reconstructed IR spectra of each of these layers were matched against their respective in-house spectral library. (For the General Motors clear coat, color coat, surfacer-primer and e-coat libraries, there were 643 spectra spanning twenty-six assembly plants for the years 2000 -2006. Table 4.8 summarizes the library search results for the paint smear samples that resulted in three layers.

Table 4.7 Paint Smear Samples with Only Three Layers (Clear Coat, Color Coat, Surfacer Primer, or E-coat)

PDQ Number	Manufacturer	Assembly Plant	Make	Line	Year	Vehicle
UAZP00433	General Motors	Flint	Chevrolet	Silverado	2002	Truck
UNJH00057	General Motors	Oshawa	Chevrolet	Impala	2005	Car
UNVL00007	General Motors	Fairfax	Pontiac	Grand Am	2003	Car
UVAC00168	General Motors	Lordstown	Chevrolet	Cavalier	2002	Car

Table 4.8 Collison Parameters and Library Matching Results for Paint Smear Samples with Only Three Paint Layers (Clear Coat, Color Coat, Surfacer Primer, or E-coat)

PDQ Number	**Impactor Parameter**	**Clear Coat**		**Color Coat**		**Surfacer Primer**		**E-coat**	
		HQI[1]	**% Match**	**HQI[2]**	**% Match**	**HQI**	**% Match**	**HQI**	**% Match**
UAZP00433	HS/DF-7	1	96.67	N/A	N/A	2	97.22	3	91.31
UNJH00057	HS/DF-6	3	93.28	-	79.78	N/A	N/A	5	97.74
UNVL00007	HS/DF-5	-	93.86	2	90.16	1	98.15	N/A	N/A
UVAC00168	HS/DF-7	1	93.37	N/A	N/A	1	92.02	2	96.71

[1]If the make and model of the vehicle corresponding to the paint smear was not present in the top five hits, the designation "-" was used.

[2]If the layer was not present in the smear, the designation "NA" was used.

Figures 4.16 to 4.19 show plots of the ALS reconstructed IR spectra that are overlaid on the library sample of the same make and model as the smear. The HQI value of each match was greater than 90% for those smears which registered a top-five hit. The clear coat, surfacer-primer, and e-coat layers generally compared well to the same sample in the in-house General Motors IR spectra library which also corresponded to a top-five hit for each paint smear. For UAZP00433 and UVAC00168, the region 1100 – 900 cm^{-1} for the ALS reconstructed IR spectra of the e-coat layer did not compare well with the corresponding IR spectra in the in-house spectral library.

The color coat layer gave poorer library matching results than the other layers as the IR spectra of color coats will usually be obscured by scattered light due to metal and pearlescent effect flakes present in this layer. Not obtaining an accurate IR library match for the color coat layer does not adversely impact on the investigative lead information that we hope to obtain from this data as the color coat layer is typically not indicative of a specific vehicle model because the same formulation is often used for the color coat throughout an entire vehicle line.

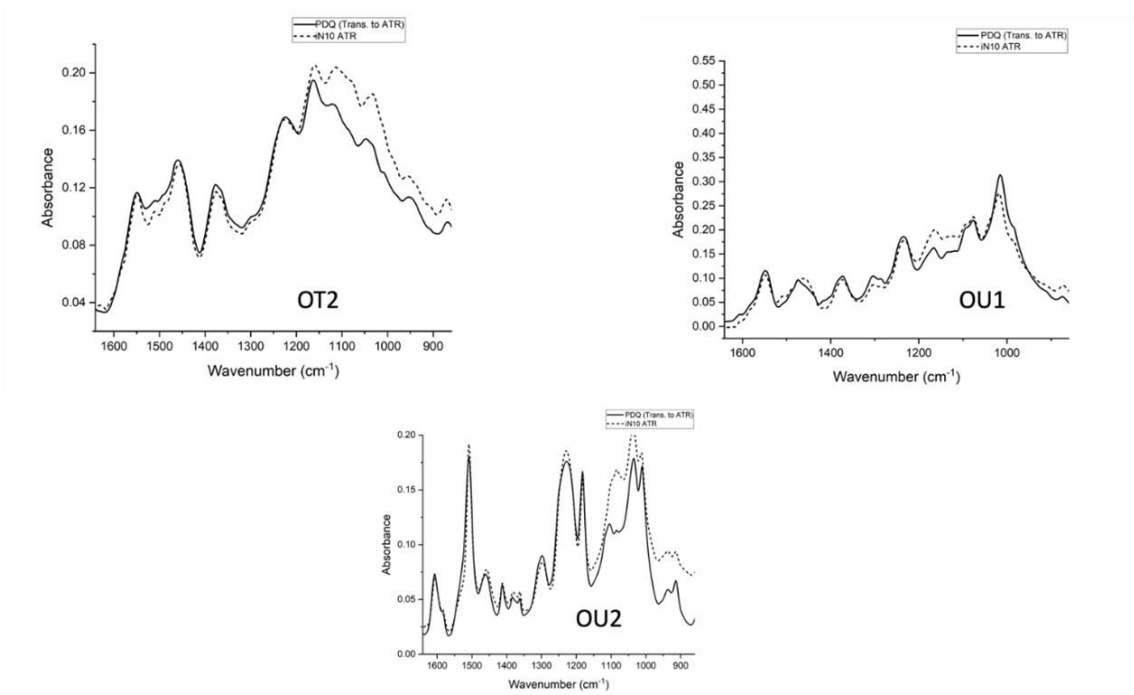

Figure 4.16 Comparison of ALS reconstructed ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UAZP00433 -General Motors) in the PDQ spectral library.

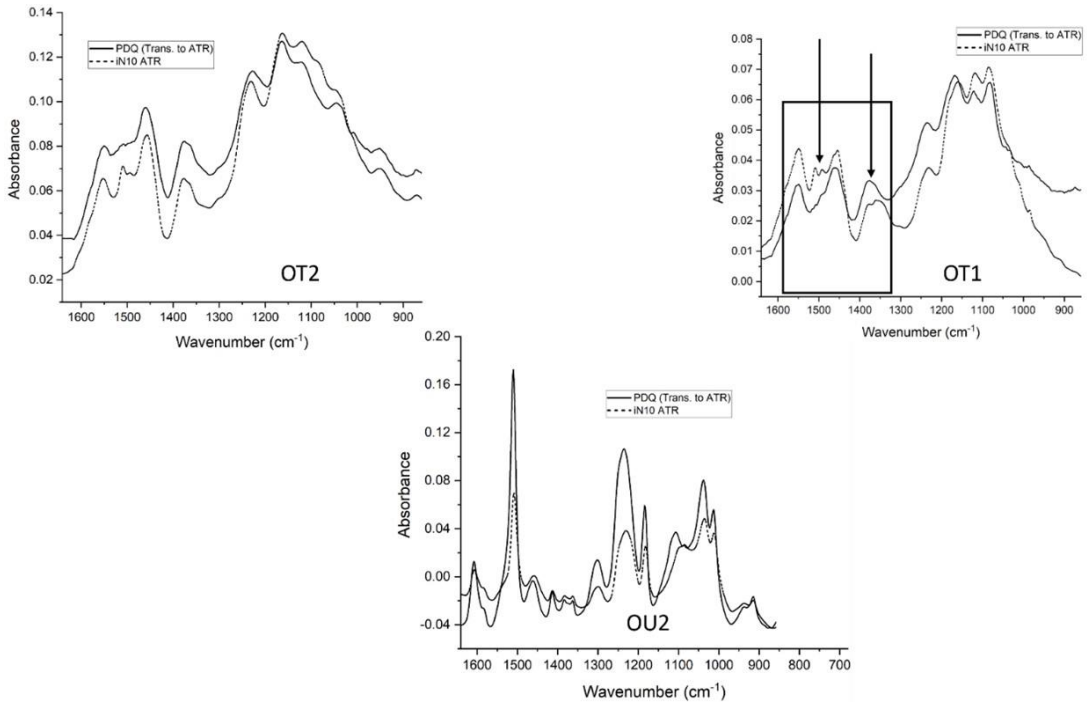

Figure 4.17 Comparison of ALS reconstructed ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UNJH00057 -General Motors) in the PDQ spectral library.

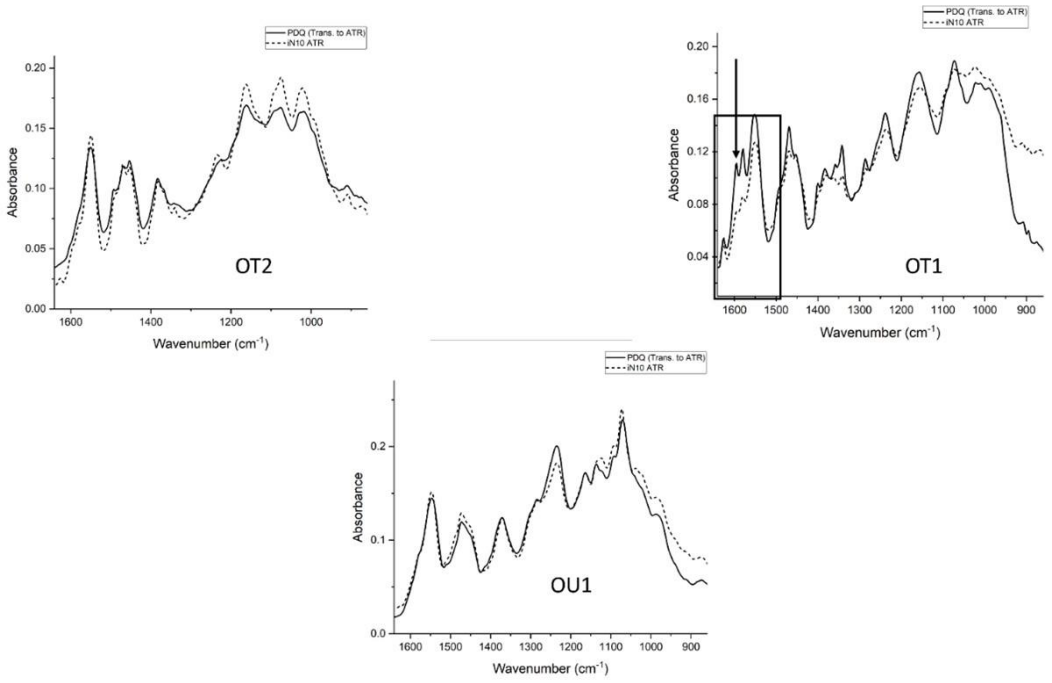

Figure 4.18 Comparison of ALS reconstructed ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UNVL0007 -General Motors) in the PDQ spectral library.

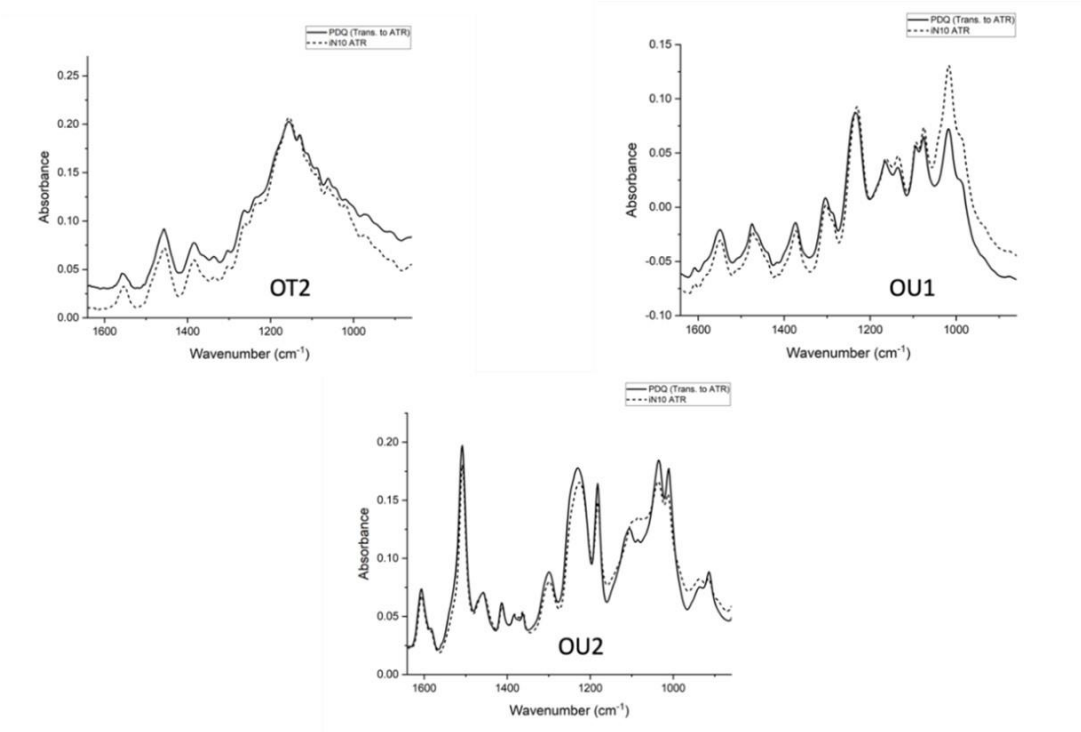

Figure 4.19 Comparison of ALS reconstructed ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UVAC00168 -General Motors) in the PDQ spectral library.

4.4.4 Four Layers Recovered.

For twelve of the thirty-one paint smears (see Table 4.9), four layers (clear coat, color coat, surfacer-primer, and e-coat) were recovered from each smear by ALS. The ALS reconstructed IR spectra of each layer were matched against their respective in-house spectral library. For the Chrysler spectral libraries, there were 465 IR spectra for the clear coats, color coats and surfacer primer layers and 446 spectra for the e-coat layer spanning fourteen assembly plants for the years 2000-2006. As for Toyota libraries, there were 264 IR spectra for the clear coat, 310 IR spectra for the color coat and surfacer primer layers, and 307 IR spectra for the e-coat encompassing twelve assembly plants for the years 2000-2006. Table 4.10 summarizes the library search results for these twelve paint smears. Figures 4.20 to 4.31 show plots of the ALS reconstructed IR spectra that were overlaid on the library sample of the same make and model as the smear and was also present in the hit-list for the paint smear. The IR spectra of the reconstructed layers for the clear coat and surfacer primer layers based on the HQI value of the match compared well to the corresponding IR spectrum of the same sample in the in-house General Motors, Chrysler, and Toyota libraries.

Table 4.9 Paint Smear Samples with Four Layers

PDQ Number	Manufacturer	Assembly Plant	Make	Line	Year	Vehicle
UACD00166	General Motors	Oshawa	Chevrolet	Camaro	2010	Car
UAZP00201	Chrysler	Belvidere	Dodge	Neon	2000	Car
UAZP00331	General Motors	Silao	Chevrolet	Suburban	2001	SUV
UAZP00332	General Motors	Silao	GMC	Yukon XL	2001	SUV
UAZP00334	General Motors	Flint	Chevrolet	Silverado	2001	Truck
UAZP00433	General Motors	Flint	Chevrolet	Silverado	2002	Truck
UAZP00494	General Motors	Silao	Chevrolet	Suburban	2003	SUV
UAZP00495	General Motors	Silao	Chevrolet	Suburban	2003	SUV
UAZP00496	General Motors	Silao	Chevrolet	Suburban	2002	SUV
UAZP00561	Toyota	Fremont	Toyota	Tacoma	2005	Truck
UMOJ00105	General Motors	Oshawa	Chevrolet	Silverado	2003	Truck
UNJH00057	General Motors	Oshawa	Chevrolet	Impala	2005	Car

Table 4.10 Collison Parameters and Library Matching Results for Paint Smear Samples with Four Layers

PDQ Number	**Impactor Parameter**	**Clear Coat**		**Color Coat**		**Surfacer Primer**		**E-coat**	
		HQI	**% Match**	**HQI[1]**	**% Match**	**HQI**	**% Match**	**HQI [1]**	**% Match**
UAZP00201	HS/DF-4	1	98.63	6	67.89	2	97.03	2	97.63
UAZP00329	HS/DF-8	1	96.28	3	83.81	1	93.73	1	97.93
UAZP00331	HS/DF-3	1	97.76	1	89.72	2	93.81	-	97.85
UAZP00332	HS/DF-7	1	98.06	-	78.21	4	94.76	-	97.57
UAZP00334	HS/DF-8	2	98.36	10	77.68	1	98.13	3	91.06
UAZP00433	HS/DF-5	1	92.74	2	82.12	3	90.61	5	97.75
UAZP00494	SS/DF-8	1	95.41	1	88.04	1	94.19	1	92.22
UAZP00494	HS/DF-1	1	97.73	1	76.43	2	90.11	-	90.52
UAZP00494	HS/DF-4	2	94.47	9	87.25	4	93.68	-	95.11
UAZP00494	HS/DF-6	2	97.13	9	87.25	1	95.55	-	90.06
UAZP00496	HS/DF-4	1	90.13	1	86.13	2	96.35	-	95.27
UAZP00561	HS/DF-4	1	97.32	4	79.98	1	98.38	5	97.70
UCAD00166	HS/DF-8	1	97.66	1	93.61	2	90.76	-	85.05
UMOJ00105	HS/DF-7	2	96.58	2	85.16	1	93.48	1	96.85
UMOJ00106	HS/DF-5	1	92.42	1	91.17	1	94.62	-	91.94

[1]If the make and model of the vehicle corresponding to the paint smear was not present in the top five hits, the designation "-" was used.

As for the e-coat layer, matching the IR spectra of the paint smear to the in-house spectral library proved problematic for ten of the twelve smears due to the lack of agreement in the spectral region 1100 – 900 cm^{-1} for the General Motors samples. This problem has been previously encountered and was attributed to an interaction between the e-coat layer of General Motors OEM automotive paints (2000 – 2006) and the germanium ATR tip [6].

For UAZP00494, the collision parameters used to generate an automotive paint smear always produced four distinct layers that could be recovered by ALS. This was true whether mild conditions (e.g., SS and damping factor 8) or more extreme conditions (HS and damping factor 8) were used. Five different sets of conditions were investigated for this OEM paint: SS/DF-8, HS/DF-1, HS/DF-4, and HS/DF-8 (see Table 4.10). These results suggest that UAZP00494 is "soft", and all layers are present in the paint smear formed.

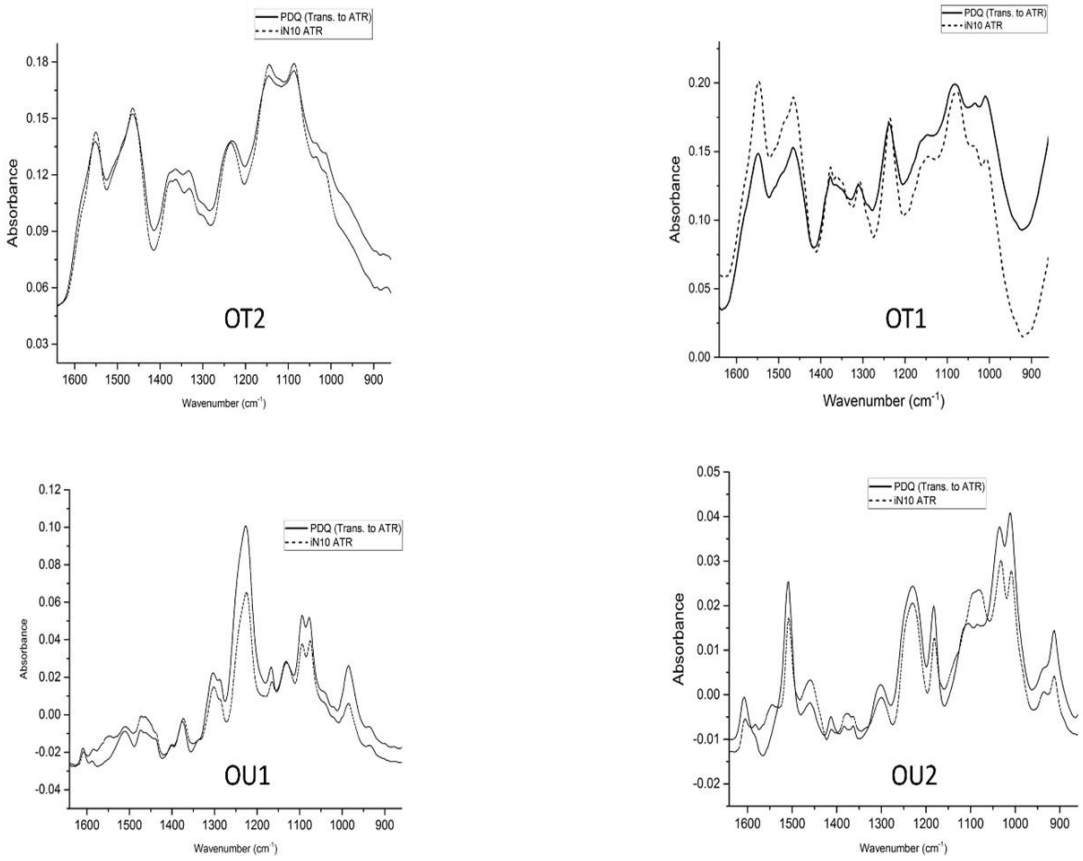

Figure 4.20 Comparison of ALS reconstructed ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UCAD00166 -General Motors) in the PDQ spectral library.

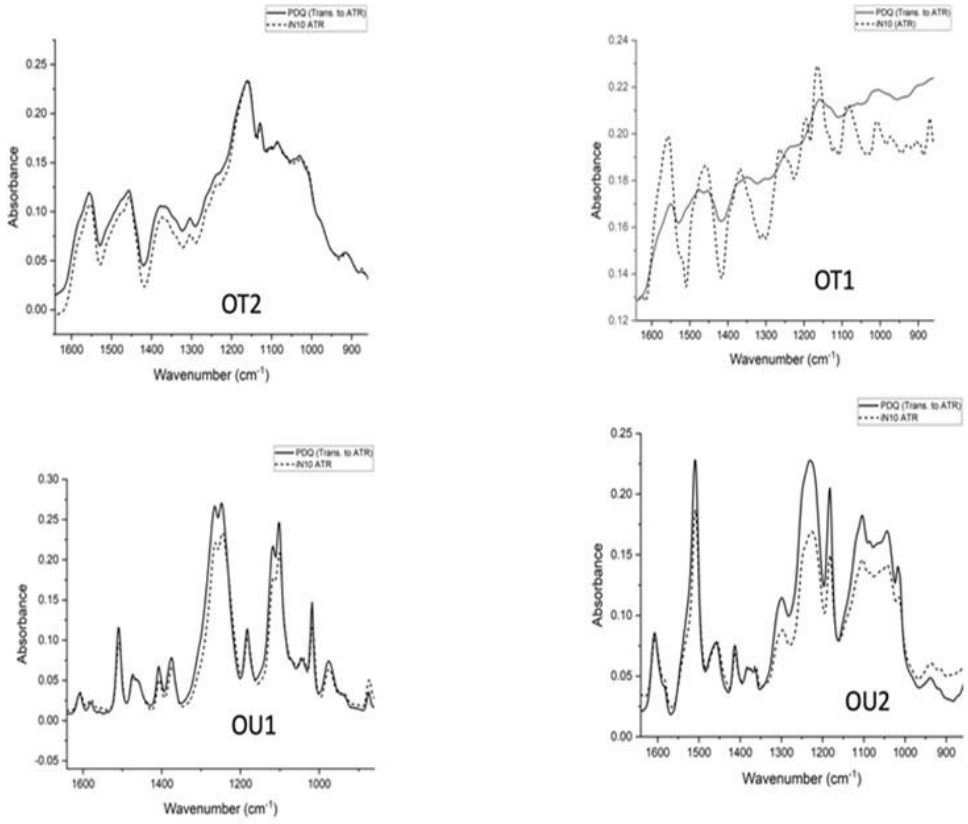

Figure 4.21 Comparison of ALS reconstructed ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UAZP00201 -Chrysler) in the PDQ spectral library.

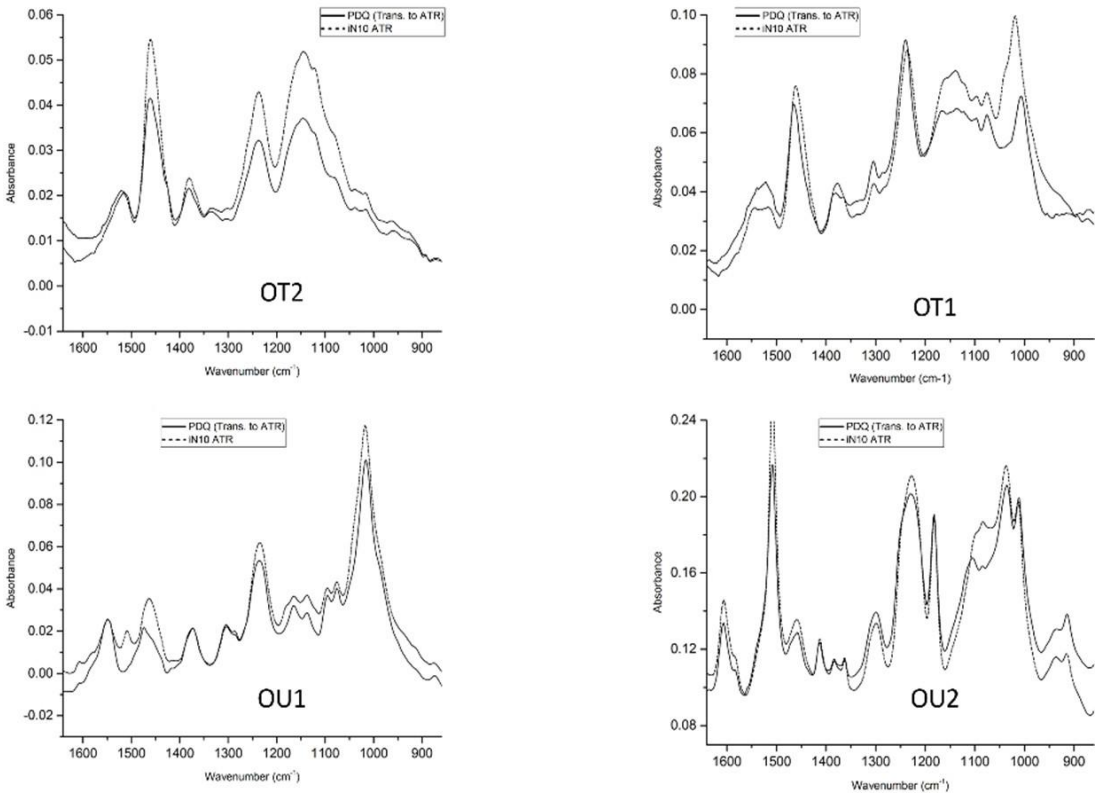

Figure 4.22 Comparison of ALS reconstructed ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UAZP00331 - General Motors) in the PDQ spectral library.

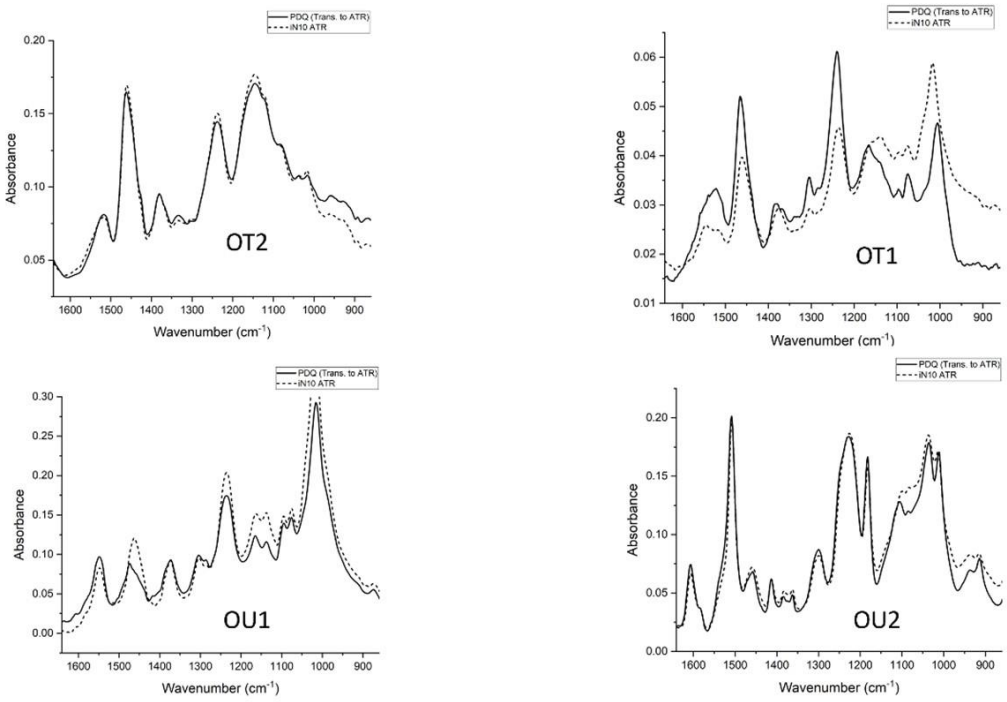

Figure 4.23 Comparison of ALS reconstructed ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UAZP00332 - General Motors) in the PDQ spectral library.

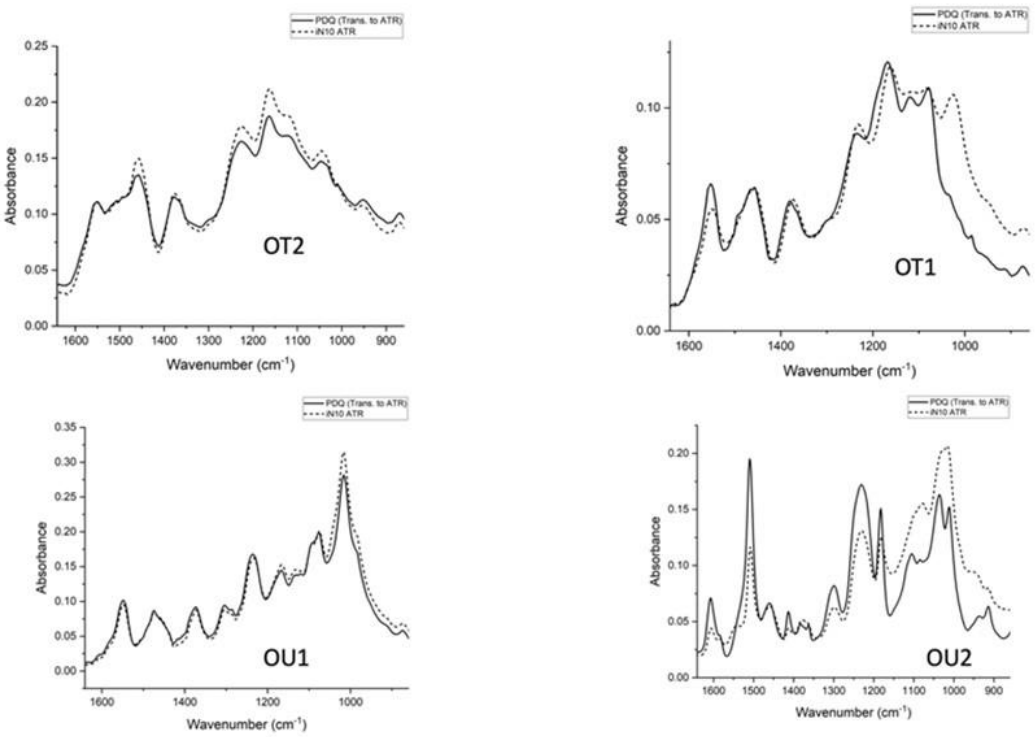

Figure 4.24 Comparison of ALS reconstructed ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UAZP00334 - General Motors) in the PDQ spectral library.

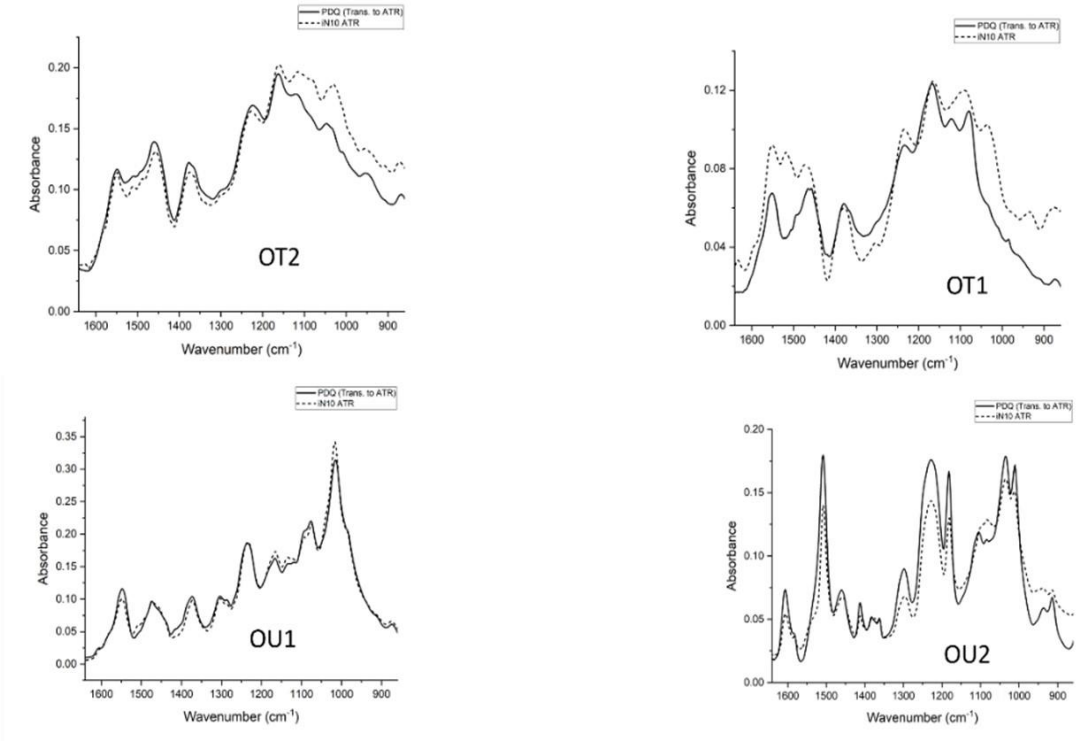

Figure 4.25 Comparison of ALS reconstructed ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UAZP00433 - General Motors) in the PDQ spectral library.

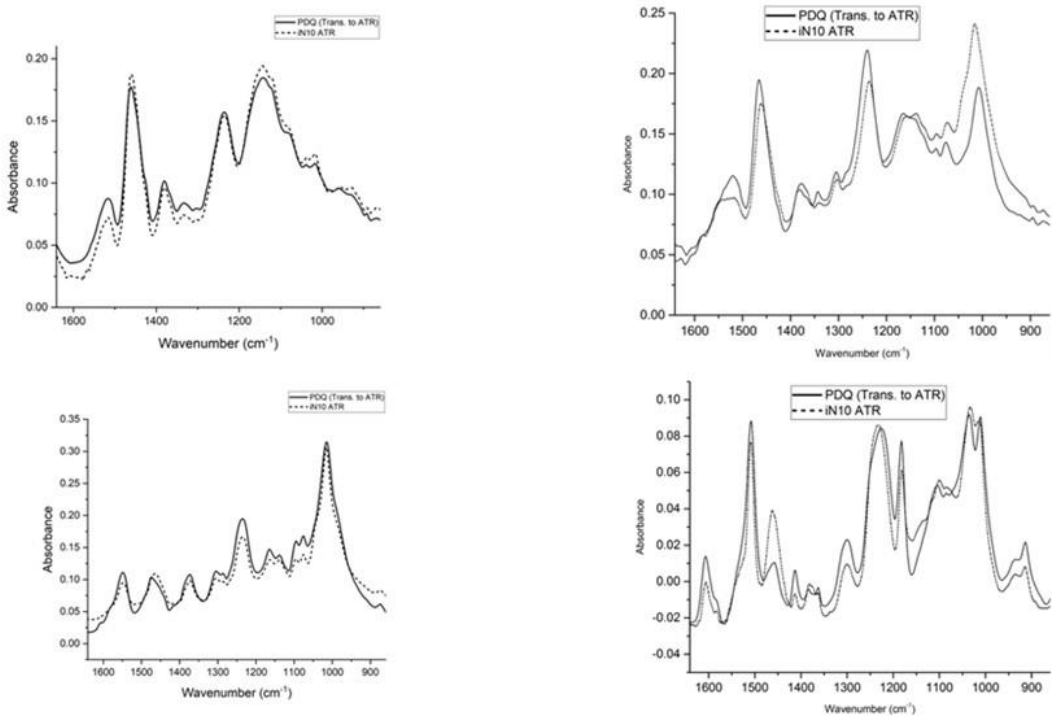

Figure 4.26 Comparison of ALS reconstructed ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UAZP00494 - General Motors) in the PDQ spectral library.

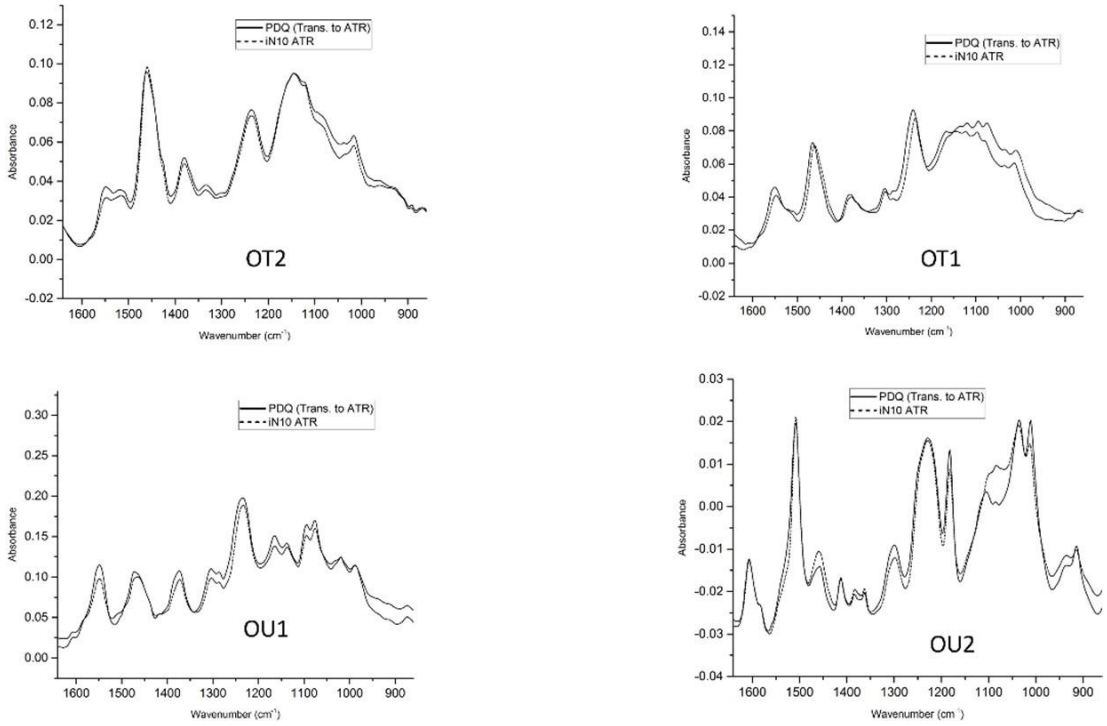

Figure 4.27 Comparison of ALS reconstructed ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UAZP00495 - General Motors) in the PDQ spectral library.

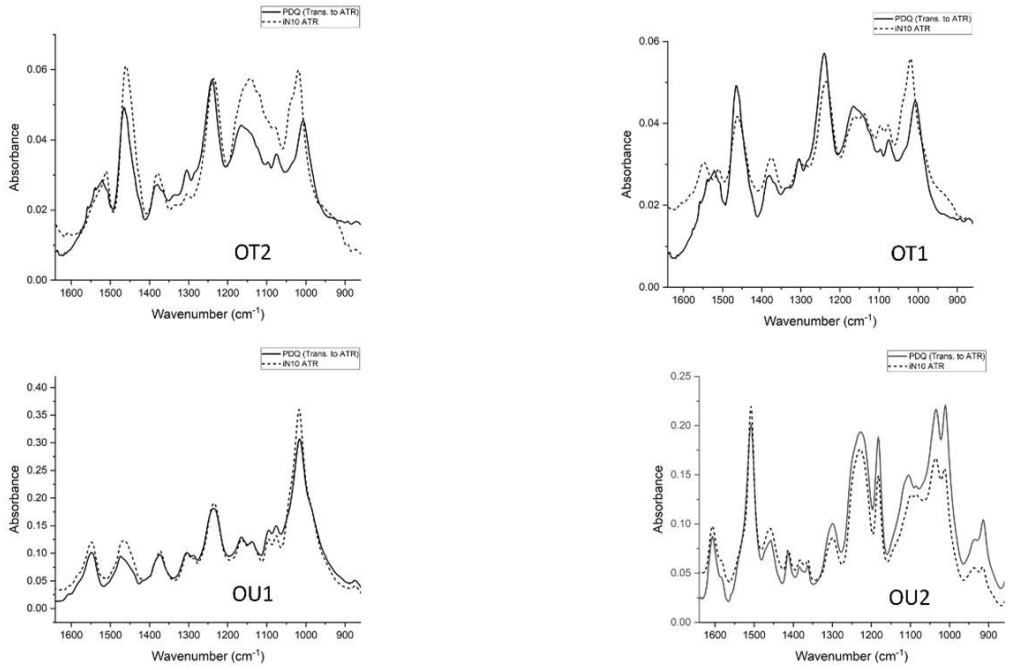

Figure 4.28 Comparison of ALS reconstructed ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UAZP00496 - General Motors) in the PDQ spectral library.

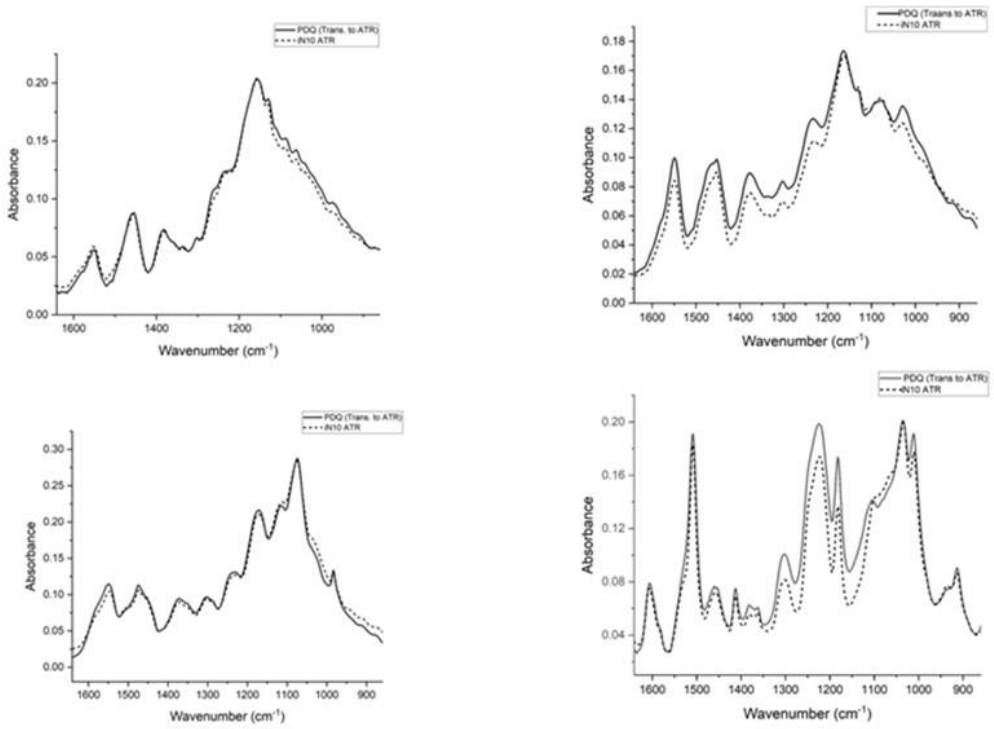

Figure 4.29 Comparison of ALS reconstructed ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UAZP00561 -Toyota) in the PDQ spectral library.

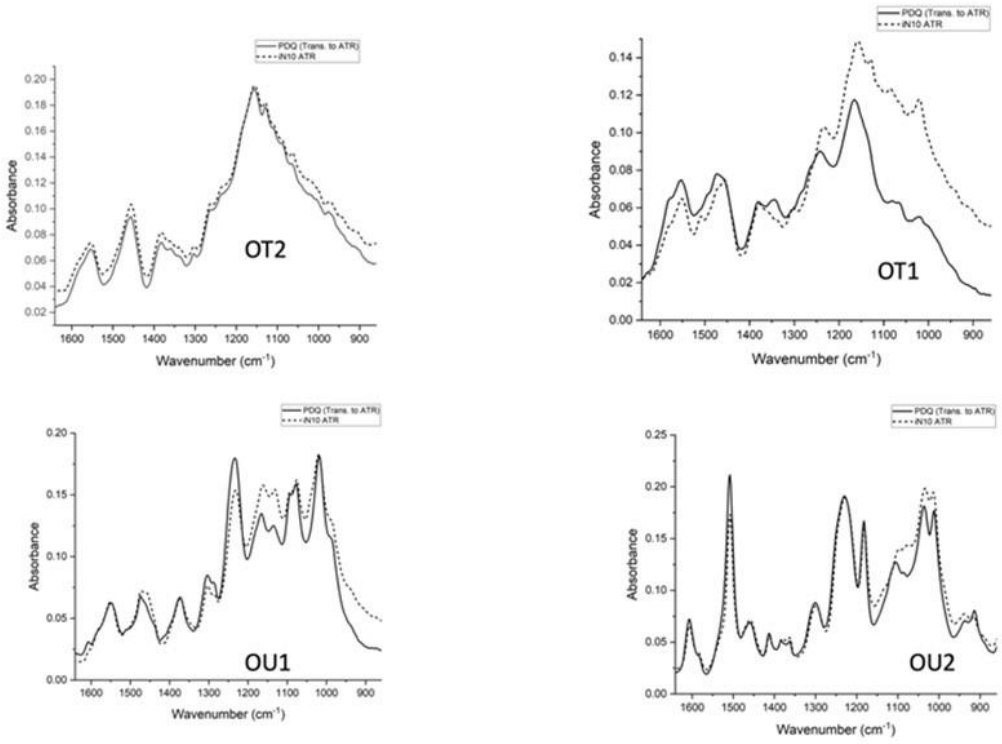

Figure 4.30 Comparison of ALS reconstructed ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UMOJ00105 -General Motors) in the PDQ spectral library.

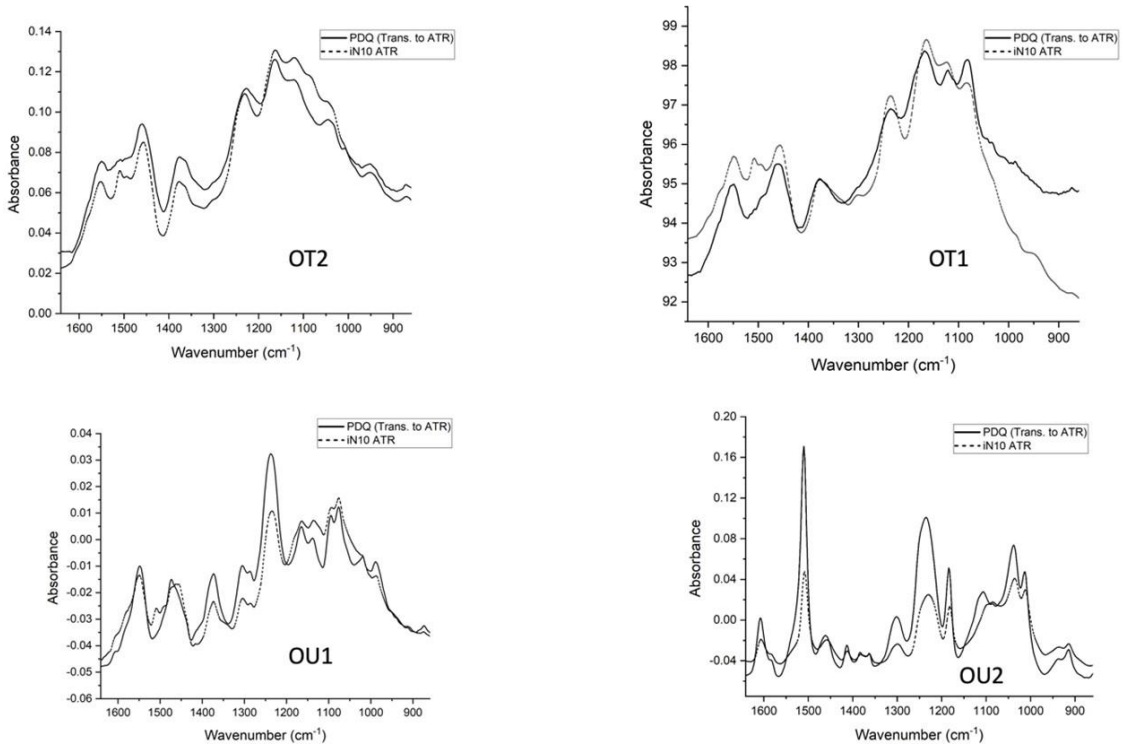

Figure 4.31 Comparison of ALS reconstructed ATR paint smear spectrum (dashed line) to IR spectrum of the actual paint sample (UNJH00057 -General Motors) in the PDQ spectral library.

4.5 Pattern Recognition Assisted Infrared Library Searching

The ALS reconstructed PDQ spectra of the OEM paint samples from which the clear coat, surfacer-primer, and e-coat layers were recovered were analyzed using a prototype pattern recognition search engine [9 - 12]. The IR spectra comprising the in-house spectral library used for the development of search prefilters in this study consisted of 4956 IR spectra of the clear coat, surfacer-primer, and e-coat layers from 1652 OEM automotive paint systems spanning six vehicle manufacturers and 62 assembly plants within a narrow production year range (2000 – 2006). Each spectrum in the library was normalized to the helium neon laser frequency of 15798.0 cm^{-1} using OMNIC (Thermo-Nicolet) to ensure proper spectral alignment along the wavelength axis for all spectra in the library.

Search prefilters were applied to the IR spectra of the fingerprint region for the ALS reconstructed clear coat, surfacer-primer, and e-coat layers to identify the manufacturer and the assembly plant of the vehicle from which each of the paint smears originated. To develop these search prefilters, IR spectra from 1652 OEM paint samples comprising the in-house IR spectral library were preprocessed using the discrete wavelet transform [13]. The sym6 mother wavelet (Symlet wavelet family, sixth smallest filter size, eighth level of decomposition) was applied to each spectrum in the region 1641 cm^{-1} to 860 cm^{-1}. The wavelet coefficients (both the approximation and detailed coefficients) for the IR spectra of the clear coat, surfacer-primer, and e-coat layers were concatenated for each OEM paint in the in-house library to form the 1652 data vectors used to develop the search prefilters. Prior to pattern recognition analysis, the data vectors were autoscaled to ensure that each wavelet coefficient had a mean of zero and standard deviation of one.

Wavelet coefficients characteristic of vehicle manufacturer or assembly plant for the search prefilters were selected for this study using a genetic algorithm (GA) for pattern recognition analysis, which identifies wavelet coefficients that optimize the separation of the classes (vehicle manufacturer or assembly plant) in a plot of the two or three largest principal components of the IR spectral data. Since principal components maximize variance, the information encoded by these wavelet coefficients is primarily about differences between the vehicle manufacturers and/or assembly plants as reflected in the set of 3825 IR spectra comprising our in-house spectra library. The pattern recognition GA was able to focus on specific vehicle manufacturers or assembly plants that were difficult to classify during training by adjusting the values of the weights for the classes (vehicle manufacturer or assembly plant) and/or samples (specific OEM paint samples) using a perceptron algorithm. Classes and/or samples that were classified correctly were not as heavily weighted as those classes and/or samples that were difficult to classify. Over time, the pattern recognition GA learned its optimal parameters in a manner similar to that of a neural network. Further details about the pattern recognition GA algorithm used to develop the search prefilters can be found in Chapter 3.

An overview of the manufacturer search prefilter system developed from the in-house IR spectral library is shown in Figure 4.32. A nine-tiered hierarchical classification scheme was used in this study to exploit the linear separability of the different assembly plants in the IR spectral data [14]. Prefilter 1 separates six Chrysler assembly plants from the remaining eight Chrysler assembly plants and the plants of the other five vehicle manufacturers. Prefilter 2 separates the two Chrysler and four General Motors assembly plants from the assembly plants of the other four vehicle manufacturers and the remaining

six Chrysler and thirteen General Motors assembly plants. Prefilter 3 separates the two Chrysler assembly plants from the four General Motors assembly plants. Prefilter 4 was developed to discriminate three of the six remaining Chrysler assembly plants from the other vehicle manufacturers including the three remaining Chrysler assembly plants. All remaining General Motors OEM paints are discriminated from the remaining data cohort using Prefilter 5. Prefilter 6 discriminates all Toyota assembly plants from those of Honda, Nissan, Ford, and the remaining three Chrysler assembly plants. Prefilter 7 discriminates the three remaining assembly plants of Chrysler and all those of Ford from Honda and Nissan. Finally, prefilters 8 and 9 were developed to discriminate Honda from Nissan and Ford from the three remaining Chrysler assembly plants.

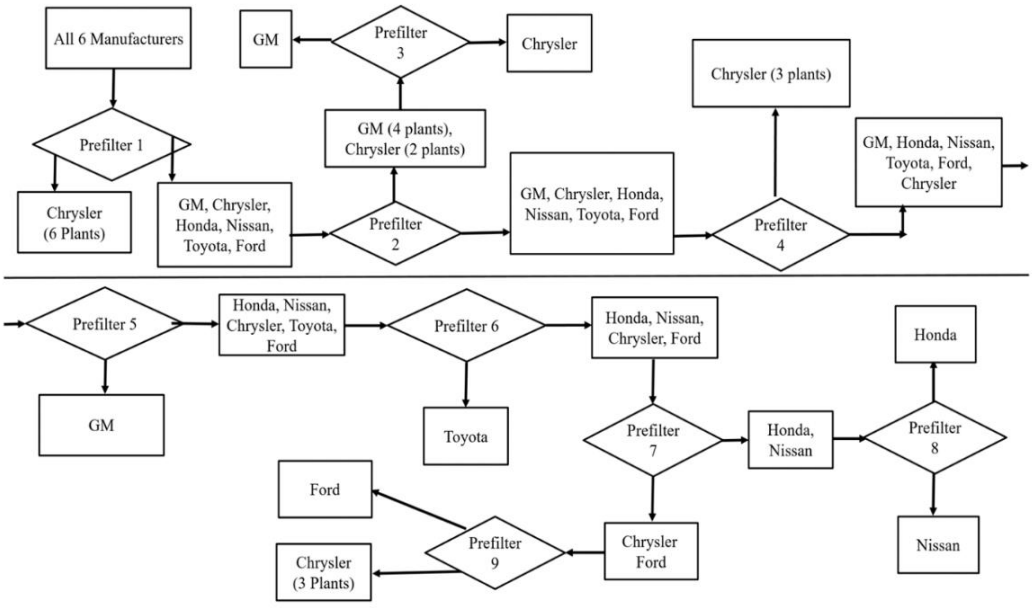

Figure 4.32 The manufacturer search prefilter system for six automotive paint manufacturers (General Motors, Chrysler, Nissan, Honda, Toyota, and Ford).

Of the fourteen paint smears that were passed through the vehicle search prefilter (two smears with three layers recovered and twelve smears with four layers recovered), only two (Toyota and Chrysler) were correctly classified as to vehicle manufacturer. The other twelve paint samples were from General Motors vehicles. For these paint samples, the interaction of the e-coat layer with the germanium ATR tip distorted the IR spectrum of the e-coat layer in the 1100 – 900 cm^{-1} spectral region adversely impacting the performance of the manufacturer search prefilter.

To demonstrate the performance of the vehicle manufacturer search prefilter system for the two OEM paint samples that are suitable for pattern recognition analysis, the ALS reconstructed IR spectra of the clear coat, surfacer-primer, and e-coat layers from UAZP00201 (Chrysler, Belvidere assembly plant) and UAZP00561 (Toyota, Fremont assembly plant) are passed through the manufacturer search prefilter. Prefilter 1 (see Figure 4.33) assigns UAZP00201 to the six Chrysler assembly plants. Therefore, the vehicle manufacturer for UAZP00201 has been identified.

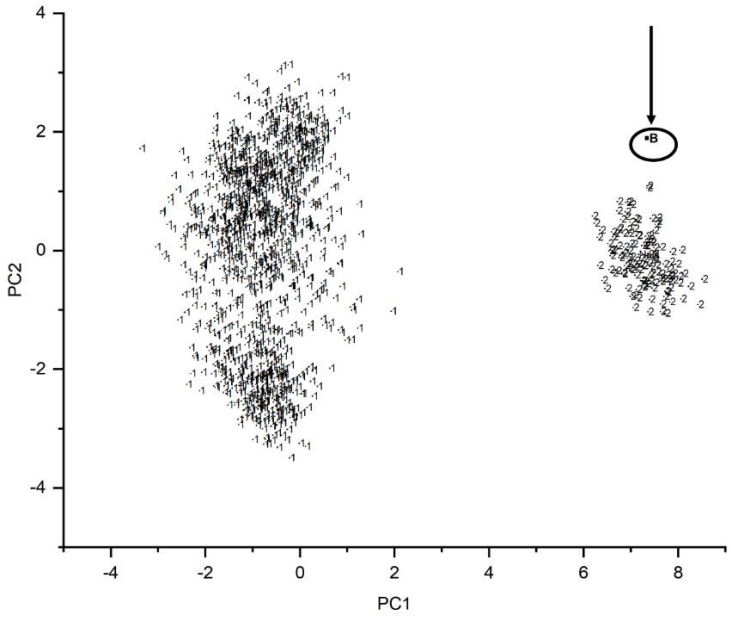

Figure 4.33 Projection of UAZP00201 onto the PC plot of Prefilter 1 by the pattern recognition GA. Training set: 1 = General Motors, Chrysler, Honda, Nissan, and Toyota; 2 = Chrysler (6 assembly plants). Validation set B = UAZP00201 (Manufacturer Chrysler).

UAZP00561 was also run through the manufacturer search prefilter. Prefilter 1 (see Figure 4.34) assigns UAZP00561 to six automotive manufacturers including nine Chrysler assembly plants that remain in the sample cohort. UAZP00561 was next run through prefilter 2 (see Figure 4.35) and was assigned to the cluster that contained samples from all the six vehicle manufacturers. Using prefilter 4 (see Figure 4.36), UAZP00561 is assigned to another sample cohort that included all six vehicle manufacturers, and prefilter 5 (see Figure 4.37) assigns UAZP00561 to the Honda, Nissan, Chrysler, Toyota, and Ford sample cohort. Finally, prefilter 6 (see Figure 4.38) assigns UAZP00561 to the Toyota cluster. Therefore, the vehicle manufacturer is Toyota.

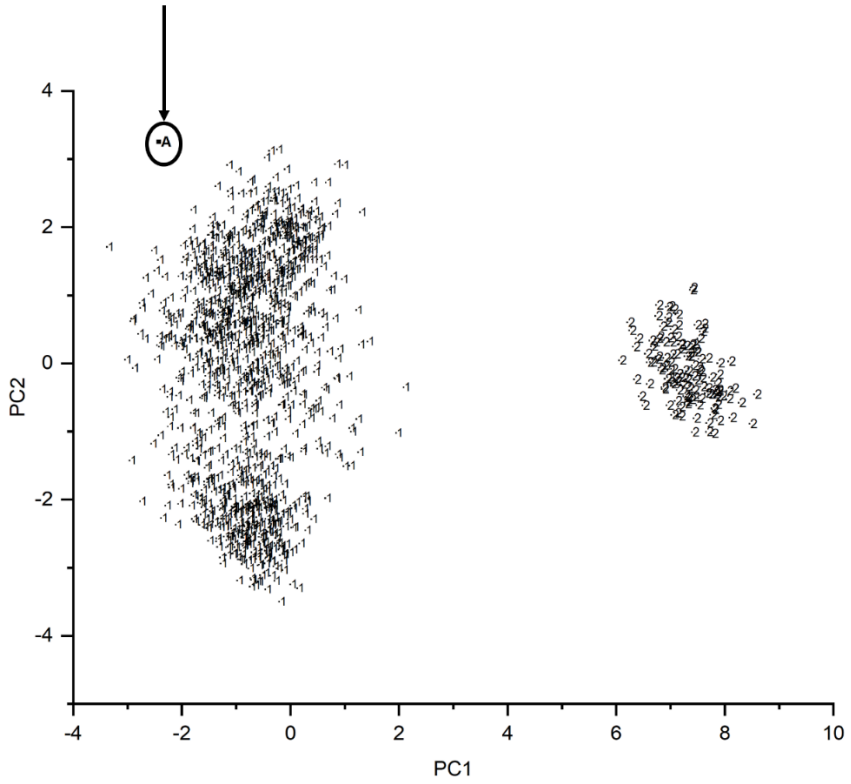

Figure 4.34 Projection of UAZP00561 onto the PC plot of Prefilter 1 by the pattern recognition GA. Training set: 1 = General Motors, Chrysler, Honda, Nissan, and Toyota; 2 = Chrysler (6 assembly plants). Validation set A = UAZP00561 (Manufacturer Toyota).

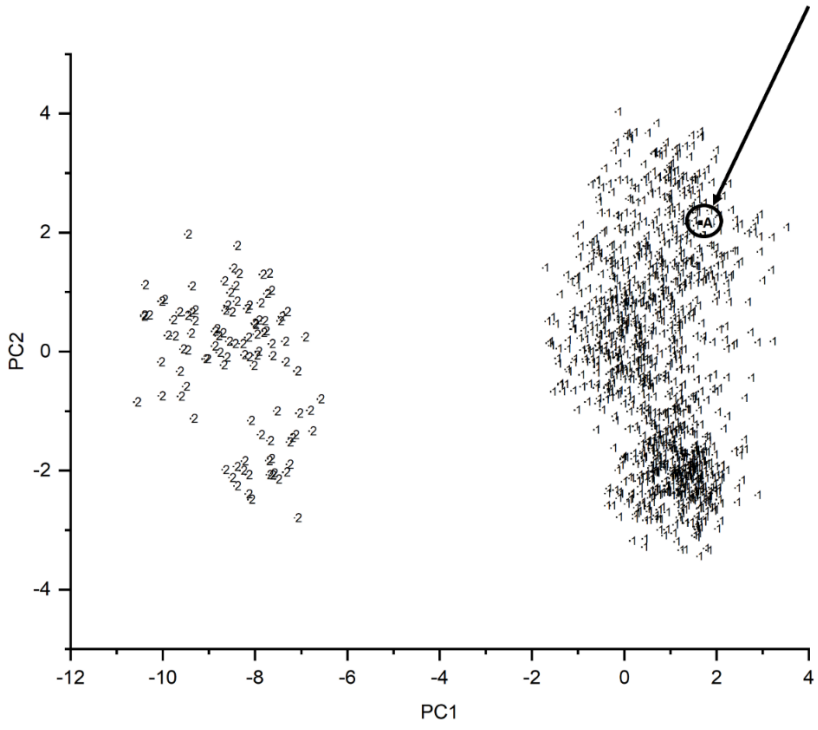

Figure 4.35. Projection of UAZP00561 onto the PC plot of Prefilter 2 by the pattern recognition GA. Training set: 1 = General Motors, Chrysler, Honda, Nissan, and Toyota; 2 = Chrysler (2 assembly plants) and General Motors (4 assembly plants). Validation set A: UAZP00562 (Manufacturer Toyota).

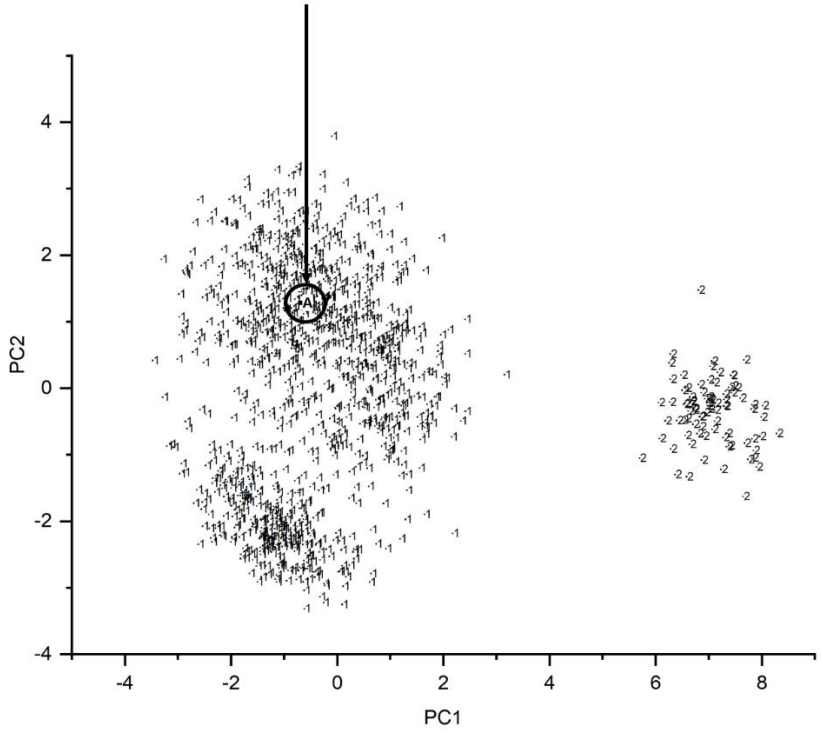

Figure 4.36. Projection of UAZP00561 onto the PC plot of Prefilter 4 by the pattern recognition GA. Training set: 1 = General Motors, Chrysler, Honda, Nissan, and Toyota; 2 = Chrysler (3 assembly plants). Validation set A = UAZP00561 (Manufacture Toyota).

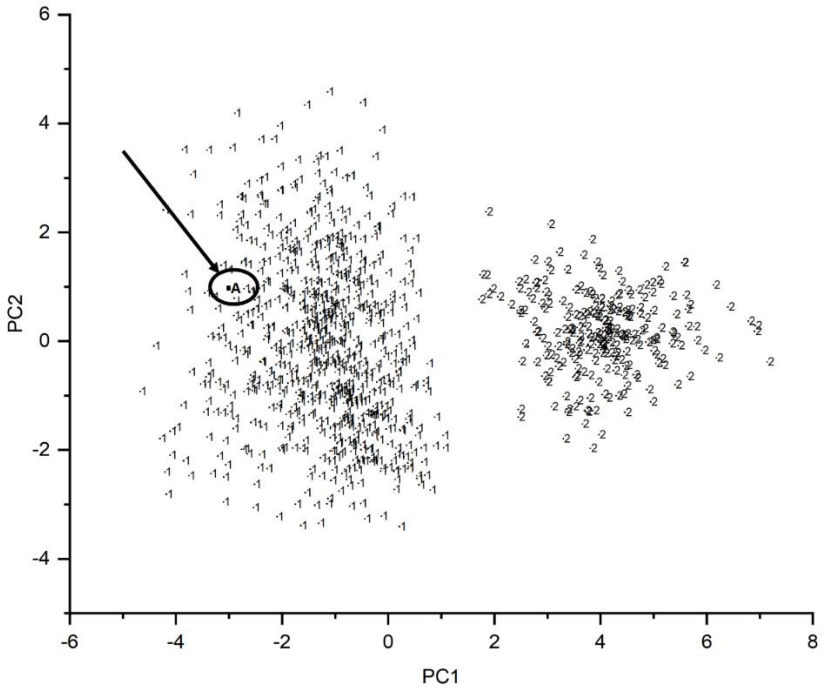

Figure 4.37. Projection of UAZP00561 onto the PC plot of Prefilter 5 defined by the 22 by the pattern recognition GA. Training set 1: Chrysler, Ford, Honda, Nissan, and Toyota; 2 = General Motors (all assembly plants). Validation set A = UAZP00561 (Manufacturer Toyota).

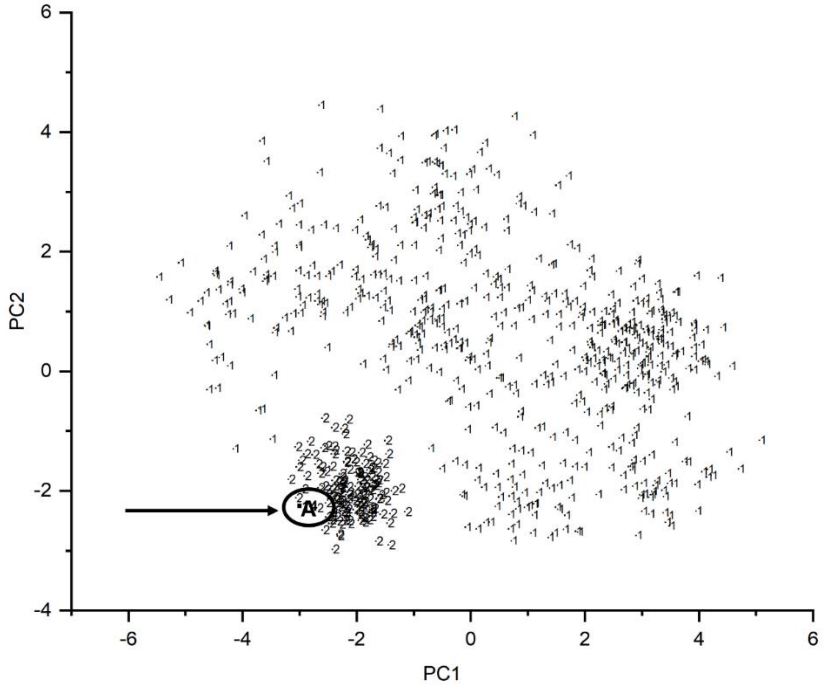

Figure 4.38. Projection of UAZP00561 onto the PC plot of Prefilter 6 by the pattern recognition GA. Training set: 1 = Chrysler, Ford, Honda, Nissan, and General Motors; 2 = Toyota. Validation set A = UAZP00561.

The assembly plant search prefilter for Chrysler, which employs a hierarchical classification of the spectral data, was applied to UAZP00201 to determine the plant group of the automotive paint smear sample and then the assembly plant. UAZP00201 was assigned to Plant Group 11 which consists of the Belvidere, Saltillo, Toluca, Toledo, Dodge Main, and St. Louis. (Figure 4.39). Finally, the assembly plant search prefilter assigns UAZP0001 to Belvidere (see Figure 4.40).

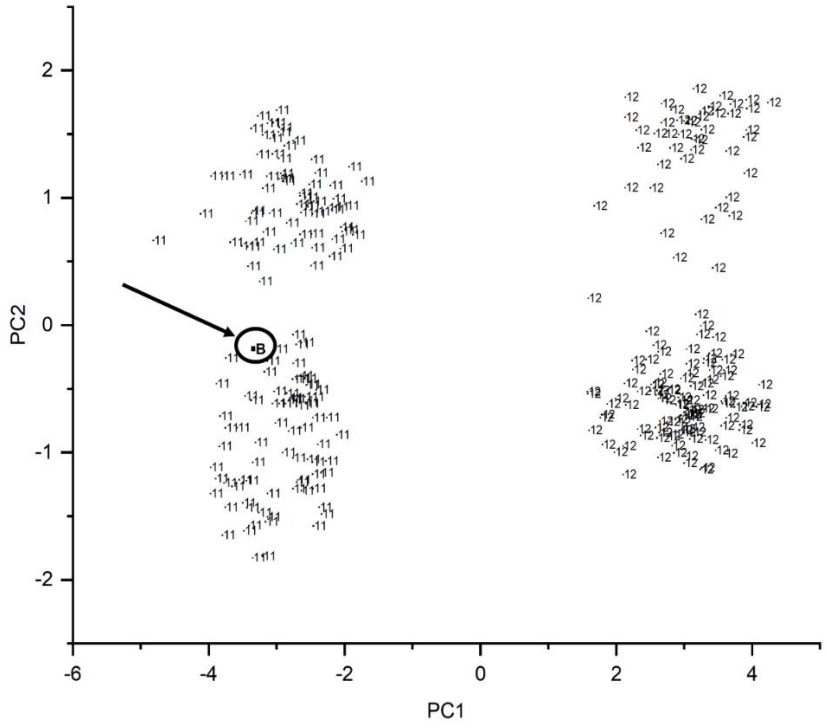

Figure 4.39 Projection of UAZP00201 onto the PC plot of Chrysler search prefilter for plant Group. UAZP00201 is assigned to Plant Group 11.

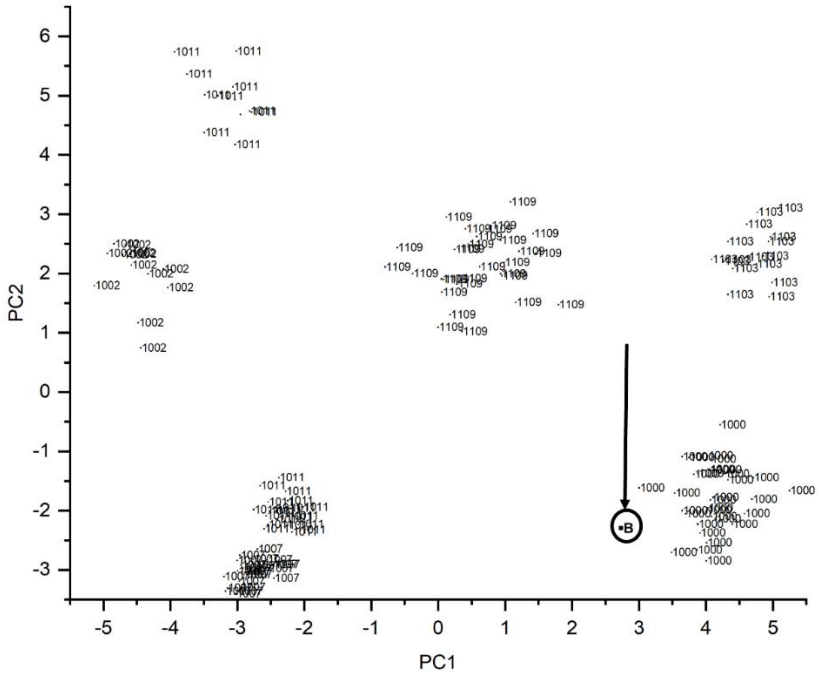

Figure 4.40. Projection of UAZP00201 onto the PC plot of the Chrysler search prefilter for assembly plant. UAZP00201 was obtained from a vehicle manufactured from the Belvidere assembly plant. B = UAZP00201, 1000 = Belvidere, 1002 = Bramalea/Brampton, 1011 = Saltillo and Toluca, 1007 = Toledo, 1103 = Dodge Main, 1109 = St. Louis.

UAZP00561 was also evaluated using the assembly plant search prefilter. As Toyota is only represented by two assembly plants (Freemont and Princeton) in our in-house spectral library, it is a simple matter to apply the Toyota search prefilter to UAZP00561 for identifying the assembly plant of the vehicle from which the paint sample originated. As shown in Figure 4.41, UAZP00561 is projected in a region of the PC plot of the Toyota search prefilter that does not contain paint samples from either assembly plant. ALS reconstructed IR spectra of UAZP00561 were also studied with discriminants to directly compare the IR spectra of the paint samples in each of the two assembly plants. In one study, UAZP00561 is merged with paint samples from Freemont, the other class

being Princeton. In a second study, UAZP00561 and Princeton are merged into a single class, the other class being Fremont. For both binary classification problems, the pattern recognition GA is able to identify a set of features that yields a separation between the two assembly plants that is comparable to the degree of separation for the same assembly plants shown in Figure 4.41. Furthermore, UAZP00561 lies in a region of the PC plot with paint samples from the class to which it was merged. Therefore, UAZP00561 cannot be assigned to either assembly plant.

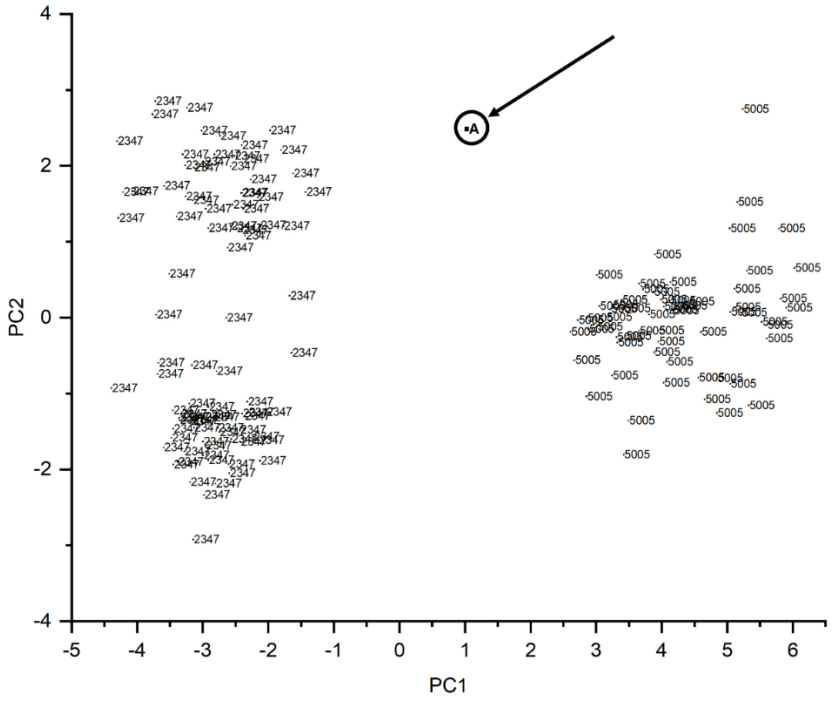

Figure 4.41 Projection of UAZP00561 onto the PC plot of the Toyota search prefilter for the assembly plant. UAZP00561 were obtained from vehicles manufactured in Fremont. A= UAZP00561, 2347 = Fremont and 5005 = Princeton

4.6 Conclusion

A machine learning approach using ALS to reconstruct IR spectra of the original layers from paint smears allows forensic paint databases to be searched to seek a best match. The results of this study also demonstrate the inter-comparison of paint smears to OEM automotive paint layer systems using IR spectra alone to quantify discrimination power of OEM paint smears encountered in actual casework and to further efforts for communicating trace evidential significance to the courts using pattern recognition.

References

1. B. Caddy, “Forensic Examination of Glass and Paint” Taylor & Francis, 2011, NY, pp. 194.

2. B. K. Lavine, A. Fasasi, N. Mirjankar, K. Nishikida, and J. Campbell, “Simulation of attenuated total reflection infrared absorbance spectra: applications to automotive clear coat forensic analysis” *Appl. Spec.,* **2014**, *68* (5), 608-615.

3. A. G. Marshall, “Biophysical Chemistry: Principles, Techniques, and Applications” Wiley New York, 1978.

4. R. M. A. Azzam, and N. M Bashara, "Ellipsometry and Polarized Light North-Holland Personal Library" Elsevier, Amsterdam, 1977, 19872, 417-86.

5. F. Wooten, "Dispersion Relations and Sum Rules: Optical Properties of Solids” F. Wooten (Ed), Academic Press, 1972, 173-85.

6. K. S. Dahal, “Infrared Microscopy and Liquid Chromatography Applied to Problems in Forensics and Bioanalytical Chemistry” PhD Dissertation, Oklahoma State University, Stillwater, OK. July 2019.

7. B. K. Lavine, C. G. White, T. Ding. “Library Search Prefilters for Vehicle Manufacturer to Assist in the Forensic Examination of Automotive Paints” *Appl. Spec*., **2018**, *72* (3), 476-488.

8. B. K. Lavine, C. G. White**,** M. D. Allen, A. Weakley. “Pattern Recognition Assisted Infrared Library Searching of the Paint Data Query Database to Improve Investigative Lead Information from Automotive Paint Trace Evidence” *Appl. Spec*., **2017**, *71* (3), 480-495.

9. U. D. N. Perera, K. Nishikida, and B. K. Lavine, “Development of infrared library search prefilters for automotive clear coats from simulated attenuated total reflection (ATR) spectra” Appl. Spec., 2018, 72(6), 886-895.

10. A. Fasasi, N. Mirjankar, R. I. Stoian, C. White, M. Allen, M. P. Sandercock, and B. K. Lavine, “Pattern recognition-assisted infrared library searching of automotive clear coats” *Appl. Spec*., **2015**, *69* (1), 84-94.

11. B. K Lavine, C. G. White, M. D. Allen, A. Fasasi, and A. Weakley, "Evidential significance of automotive paint trace evidence using a pattern recognition based infrared library search engine for the Paint Data Query Forensic Database" *Talanta*, **2016**, *159*, 317-329.

12. B. K. Lavine, C. White, and M. Allen, "Forensic analysis of automotive paints using a pattern recognition assisted infrared library searching system: Ford (2000–2006)" *Microchemical Journal*, **2016**, *129*, 173-183.

13. J. S. Walker, "A primer on wavelets and their scientific applications". Chapman and Hall/CRC, 2008.

14. F. Kwofie, U. D. N. Perera, K. S. Dahal, G. P. Affadu-Danful, K. Nishikida, and B. K. Lavine, "Transmission Infrared Microscopy and Machine Learning Applied to the Forensic Examination of Original Automotive Paint" *Appl. Spec*., **2022**, *76* (1), 118-131.

CHAPTER V

RAMAN SPECTROSCOPY TO ENHACE INVESTIGATIVE LEAD INFORMATION FOR AUTOMOTIVE CLEAR COATS

5.1 Introduction

Automotive paint consists of several layers: original and repaint layers, topcoats, and primers. Modern original equipment manufacturer (OEM) automotive paint systems have a typical layer sequence of substrate, primer (e-coat), surfacer-primer, color coat, and clear coat. As modern OEM paint systems typically have thinner undercoat and color layers protected by a thicker clear coat layer, a clear coat (all too often) is the only layer of automotive paint recovered from the crime scene of a vehicle related fatality such as a hit-and-run where injury or death to a pedestrian has occurred. In these situations, automotive paint databases such as PDQ, are unable to identify the make and model of the vehicle from which the paint sample originated. However, the clear coat layer, like the undercoat and color coat layers, exhibits features in its IR spectra that are characteristic of the assembly plant where the automotive paint was applied.

Hence, crucial investigative lead information potentially can be extracted from the IR spectra of clear coats. In recently published studies [1-7], Lavine and coworkers have demonstrated that search prefilters developed from clear coat IR spectra alone can reveal information about the make and model of a vehicle. Even in challenging trials where the samples evaluated were all of the same make and model within a limited production year range (2000-2006), the search prefilters were able to identify the specific vehicle assembly plant or potential assembly plants that manufactured the vehicle and hence the specific line and model of the vehicle that is responsible for the automotive paint sample.

In the present study, Raman spectroscopy and pattern recognition techniques have been investigated as an alternative solution to the problem of extracting investigative lead information from automotive clear coats. The complementary nature of Raman and infrared spectroscopy and the important role played by IR spectroscopy in forensic automotive paint analysis suggests that Raman spectroscopy also has the potential to extractive investigative lead information from automotive clear coats. Raman spectroscopy has several advantages over IR spectroscopy. Raman bands generally do not overlap, whereas IR bands often are overlapped in the spectra. IR bands that are too weak to be observed may be sufficiently intense to be observed in the corresponding Raman spectra. However, only a few studies [8-12] have been reported in the literature on the use of Raman spectroscopy to characterize automotive clear coats. The results reported by the authors in these studies are overly optimistic as most of the automotive clear coat classes surveyed by these authors contained only a few samples. Furthermore, the clear coats investigated in these studies were manufactured in Japan, the United States, and several countries in Europe. The results obtained by the authors are probably more indicative of

differences in the supply chain for these automotive paints. By comparison, the study presented in this chapter focuses on 118 clear coat IR spectra obtained from automotive vehicles manufactured in six General Motors assembly plants between 2000 and 2006. In a previous study [6], only one of the six assembly plants could be differentiated from the other plants using IR spectra of the clear coats. Hence, this problem was selected because it is challenging and provides a reasonable test bed to validate the advantages of using Raman spectroscopy for the forensic examination of automotive clear coats.

5.2 Materials

118 OEM automotive paint samples (see Table 5.1) were selected to investigate the discrimination power of Raman and infrared spectroscopy towards automotive clear coats. All OEM automotive paint samples are from General Motors and span six distinct assembly plants (Doraville, Arlington, Fairfax, Fort Wayne, Lansing, and Moraine, see Tables 5.2 to 5.7). Raman spectra were collected using a WITEC Alpha 300, 532 nm neodymium-doped yttrium aluminum garnet (ND: YAG) laser Raman microscope equipped with charge-coupled detector (CCD). FTIR spectra were obtained from an in-house IR spectral library provided by the RCMP. Each paint sample in the library was hand sectioned with a scalpel under a stereomicroscope. Because the clear coat layer is thick, the approach used was to first access the primer and color coat layers by removing one layer at a time starting with the e-coat layer and performing IR analysis of each layer as it is exposed and removed using a high-pressure diamond cell for compression of the thin peels to collect IR transmission spectra. Further details about the collection of the Raman and IR spectra discussed in this chapter are provided in Chapter II.

Table 5.1 Assembly Plants Used in the Raman and FTIR Studies

Assembly Plant	Plant ID Number	Number of Samples
Arlington	1	33
Doraville	4	5
Fairfax	5	20
Fort Wayne	8	16
Lansing	14	22
Moraine	18	22

Table 5.2 Doraville Assembly Plant

PDQ Number	Manufacturer	Assembly Plant	Make	Line	Year
CONT00475	General Motors	Doraville	Chevrolet	Venture	2001
CONT00699	General Motors	Doraville	Chevrolet	Venture	2002
CONT01501	General Motors	Doraville	Chevrolet	Venture	2004
CONT01504	General Motors	Doraville	Chevrolet	Uplander	2005
CONT01505	General Motors	Doraville	Pontiac	Montana	2006

Table 5.3 Arlington Assembly Plant

PDQ Number	Manufacturer	Assembly Plant	Make	Line	Year
CONT00902	General Motors	Arlington	Chevrolet	Suburban	2002
CONT01861	General Motors	Arlington	Cadillac	Escalade	2004
UARL00113	General Motors	Arlington	Chevrolet	CTA	2007
UAZP00242	General Motors	Arlington	GMC	Yukon	2012
UAZP00333	General Motors	Arlington	Chevrolet	CTA	2001
UAZP00391	General Motors	Arlington	Chevrolet	CTA	2001
UAZP00396	General Motors	Arlington	Chevrolet	CTA	2001
UAZP00435	General Motors	Arlington	Chevrolet	CTA	2001
UAZP00437	General Motors	Arlington	Chevrolet	CTA	2002
UAZP00502	General Motors	Arlington	Chevrolet	CTA	2003
UAZP00504	General Motors	Arlington	Chevrolet	CTA	2002
UAZP00796	General Motors	Arlington	Cadillac	Escalade	2011
UAZP00797	General Motors	Arlington	Cadillac	Escalade	2011
UAZP00798	General Motors	Arlington	Chevrolet	Tahoe	2010
UAZP00799	General Motors	Arlington	Chevrolet	Tahoe	2009
UAZP01122	General Motors	Arlington	Cadillac	Escalade	2010
UCAD00069	General Motors	Arlington	Chevrolet	Tahoe	2009
UCAD00141	General Motors	Arlington	GMC	SUB	2009
UCAD00147	General Motors	Arlington	Chevrolet	Tahoe	2010
UCAD00196	General Motors	Arlington	GMC	Yukon	2009
UCAD00229	General Motors	Arlington	Chevrolet	Tahoe	2012
UKYF00203	General Motors	Arlington	Chevrolet	Avalanche	2003
UMNP00135	General Motors	Arlington	Cadillac	Escalade	2000
UMNP00141	General Motors	Arlington	Cadillac	Escalade	2000
UMNP00180	General Motors	Arlington	GMC	Yukon	2000
UNVL00074	General Motors	Arlington	Chevrolet	CTA	2003
USCC00246	General Motors	Arlington	Chevrolet	Tahoe	2009
USCC00301	General Motors	Arlington	Chevrolet	Tahoe	2011
USCC00513	General Motors	Arlington	Chevrolet	Tahoe	2011
USCC00514	General Motors	Arlington	Chevrolet	Tahoe	2010
UTNN00054	General Motors	Arlington	Cadillac	Escalade	2002
UTXA00234	General Motors	Arlington	Chevrolet	Tahoe	2002
UWVC00116	General Motors	Arlington	Chevrolet	CTA	2000

Table 5.4 Fairfax Assembly Plant

PDQ Number	Manufacturer	Assembly Plant	Make	Line	Year
CONT01472	General Motors	Fairfax	Pontiac	Grand P	2002
CONT01475	General Motors	Fairfax	Oldsmobile	Intrigue	2002
CONT01545	General Motors	Fairfax	Pontiac	Grand P	2002
UARL00039	General Motors	Fairfax	Chevrolet	Malibu	2009
UAZP00801	General Motors	Fairfax	Chevrolet	Malibu	2012
UAZP00802	General Motors	Fairfax	Chevrolet	Malibu	2012
UAZP00998	General Motors	Fairfax	Chevrolet	Malibu	2012
UAZP00999	General Motors	Fairfax	Buick	Lacrosse	2012
UCAD00049	General Motors	Fairfax	Chevrolet	Malibu	2009
UCAD00067	General Motors	Fairfax	Chevrolet	Malibu	2009
UCAD00068	General Motors	Fairfax	Chevrolet	Malibu	2009
UCAD00071	General Motors	Fairfax	Chevrolet	Malibu	2008
UCAD00086	General Motors	Fairfax	Chevrolet	Malibu	2009
UCAD00089	General Motors	Fairfax	Buick	Lacrosse	2010
UCAD00179	General Motors	Fairfax	Chevrolet	Malibu	2011
UCAD00181	General Motors	Fairfax	Chevrolet	Malibu	2010
UNJH00101	General Motors	Fairfax	Chevrolet	Malibu	2011
UNVL00007	General Motors	Fairfax	Pontiac	Grand P	2003
UNVL00059	General Motors	Fairfax	Chevrolet	Malibu	2006
UNYV00138	General Motors	Fairfax	Chevrolet	Malibu	2004

Table 5.5 Fort Wayne Assembly Plant

PDQ Number	Manufacturer	Assembly Plant	Make	Line	Year
CONT01394	General Motors	Fort Wayne	Chevrolet	Silverado	2002
CONT01395	General Motors	Fort Wayne	GMC	Sierra	2002
CONT01396	General Motors	Fort Wayne	Chevrolet	Silverado	2000
CONT01397	General Motors	Fort Wayne	GMC	Sierra	2000
CONT01515	General Motors	Fort Wayne	Chevrolet	Silverado	2006
CONT01856	General Motors	Fort Wayne	Chevrolet	Silverado	2006
UARL00038	General Motors	Fort Wayne	Chevrolet	Silverado	2010
UARL00085	General Motors	Fort Wayne	GMC	Sierra	2003
UCAR00191	General Motors	Fort Wayne	Chevrolet	Silverado	2000
UKYF00200	General Motors	Fort Wayne	Chevrolet	Silverado	2003
UNVL00068	General Motors	Fort Wayne	Chevrolet	Silverado	2004
UOHL00234	General Motors	Fort Wayne	Chevrolet	Silverado	2004
UOKO00115	General Motors	Fort Wayne	Chevrolet	Silverado	2000
UOKO00153	General Motors	Fort Wayne	Chevrolet	Silverado	2002
UTXA00172	General Motors	Fort Wayne	Chevrolet	Silverado	2002
UWVC00140	General Motors	Fort Wayne	Chevrolet	Silverado	2000

Table 5.6 Lansing Assembly Plant

PDQ Number	Manufacturer	Assembly Plant	Make	Line	Year
CONT00984	General Motors	Lansing	MAL	CXX	2004
CONT00985	General Motors	Lansing	Oldsmobile	OXX	2000
CONT01042	General Motors	Lansing	Pontiac	Grand Am	2003
CONT01046	General Motors	Lansing	Oldsmobile	Alero	2003
CONT01047	General Motors	Lansing	Oldsmobile	Alero	2003
CONT01049	General Motors	Lansing	Chevrolet	Malibu	2001
CONT01474	General Motors	Lansing	Pontiac	Grand Am	2003
CONT01522	General Motors	Lansing	Pontiac	Grand Am	2002
CONT01550	General Motors	Lansing	Oldsmobile	Alero	2002
CQCM00137	General Motors	Lansing	Pontiac	Grand Am	2001
UAZP00567	General Motors	Lansing	Chevrolet	Malibu	2003
UCAR00180	General Motors	Lansing	Pontiac	Grand Am	2000
UMNP00151	General Motors	Lansing	Pontiac	Grand Am	2000
UMNP00186	General Motors	Lansing	Oldsmobile	Alero	2000
UNCR00360	General Motors	Lansing	Oldsmobile	Alero	2001
UOCN00015	General Motors	Lansing	Pontiac	Grand Am	2000
UOHL00050	General Motors	Lansing	Pontiac	Grand Am	2000
UOHL00139	General Motors	Lansing	Oldsmobile	Alero	2003
UTNN00056	General Motors	Lansing	Pontiac	Grand Am	2002
UTNN00057	General Motors	Lansing	Oldsmobile	Alero	2001
UVAC00134	General Motors	Lansing	Oldsmobile	Alero	2000
UVAC00134	General Motors	Lansing	Oldsmobile	Alero	2000

Table 5.7 Moraine Assembly Plant

PDQ Number	Manufacturer	Assembly Plant	Make	Line	Year
CONT00918	General Motors	Moraine	Chevrolet	Trail Blazer	2003
CONT00920	General Motors	Moraine	Chevrolet	Trail Blazer	2002
CONT00921	General Motors	Moraine	GMC	Jimmy	2001
CONT01483	General Motors	Moraine	GMC	Envoy	2003
CONT01583	General Motors	Moraine	GMC	Envoy	2003
UAZP00430	General Motors	Moraine	Chevrolet	Blazer	2002
UAZP00431	General Motors	Moraine	Chevrolet	Blazer	2002
UAZP00432	General Motors	Moraine	Chevrolet	Blazer	2002
UAZP00498	General Motors	Moraine	Chevrolet	Trail Blazer	2003
UCAD00033	General Motors	Moraine	GMC	Envoy	2006
UKYF00171	General Motors	Moraine	GMC	Jimmy	2000
UNCR00377	General Motors	Moraine	GMC	Sonoma	2000
UNJH00055	General Motors	Moraine	Chevrolet	Trail Blazer	2004
UOHL00116	General Motors	Moraine	GMC	Envoy	2003
UOHL00230	General Motors	Moraine	Chevrolet	Trail Blazer	2005
UOKO00214	General Motors	Moraine	Chevrolet	Trail Blazer	2003
USCC00008	General Motors	Moraine	Chevrolet	Blazer	2000
USCC00088	General Motors	Moraine	GMC	Envoy	2004
USCC00313	General Motors	Moraine	Chevrolet	Trail Blazer	2007
UTXA00206	General Motors	Moraine	GMC	Jimmy	2000
UVAC00143	General Motors	Moraine	Chevrolet	Blazer	2001
UWAS00001	General Motors	Moraine	Chevrolet	Blazer	2000

5.3 Methods

5.3.1 Data Preprocessing of Raman Spectra

All Raman spectra were preprocessed to obviate both noise and background using a three-step procedure performed using the PLS Toolbox (Eigenvector Technology). First, the Whitaker filter was applied to each spectrum for baseline correction. All baseline corrected spectra were smoothed using a Savitzky-Golay filter with a 5-point window and polynomial order 1. Each spectrum was then normalized to unit length. Figure 5.1 shows a Raman spectrum that has been preprocessed, i.e., baseline corrected, smoothed, and normalized to unit length. For some paint samples (e.g., UAZP00795), cosmic spikes were randomly generated by the CCD detector. To eliminate these cosmic spikes, they have been replaced with the average intensity of the wavelengths adjacent to the spike as shown in Figure 5.2. In addition, significant peak shifts (greater than 2 cm^{-1}) were observed for some OEM automotive paint samples (e.g., UAZP00796). This problem has been addressed by irradiating the sample using less laser power (10 mW instead of 40 mW), longer integration times (30 seconds instead of 20 seconds) and more spectral accumulations (30 instead of 20), see Figure 5.3. A final problem encountered was fluorescence. For the samples that exhibited fluorescence (e.g., CONT00928), photobleaching was used to address this problem. (Fluorescence probably occurred due to the color coat layer being mixed with the clear coat layer during hand sectioning of the paint sample.) For photobleaching, the sample was exposed to the laser source for about 5 minutes before the spectra were collected. Exposing the sample to the laser for this length of time saturated the fluorophore, thereby preventing the sample from fluorescing (see Figure 5.4).

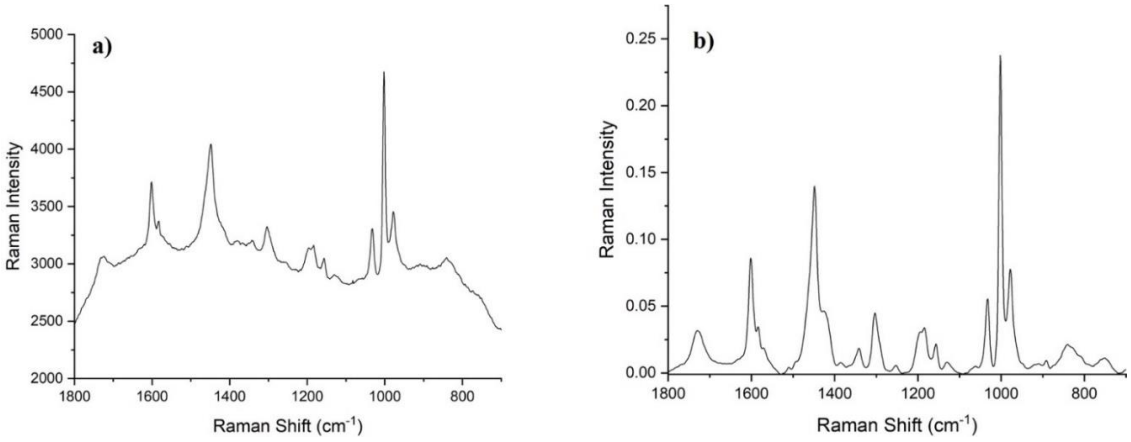

Figure 5.1 A representative Raman spectrum of automotive paint clear coat (UNVL00007): a) before baseline correction, smoothening and normalizing to unit length and b) after baseline correction and normalizing to unit length.

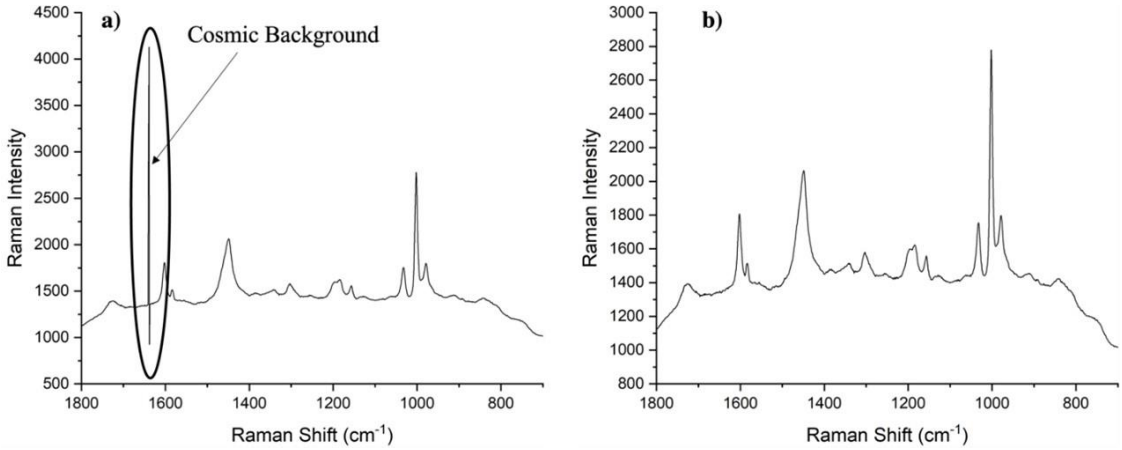

Figure 5.2 A representative Raman spectrum of automotive paint clear coat (UAZP00795): a) before cosmic background removal and b) after cosmic background removal.

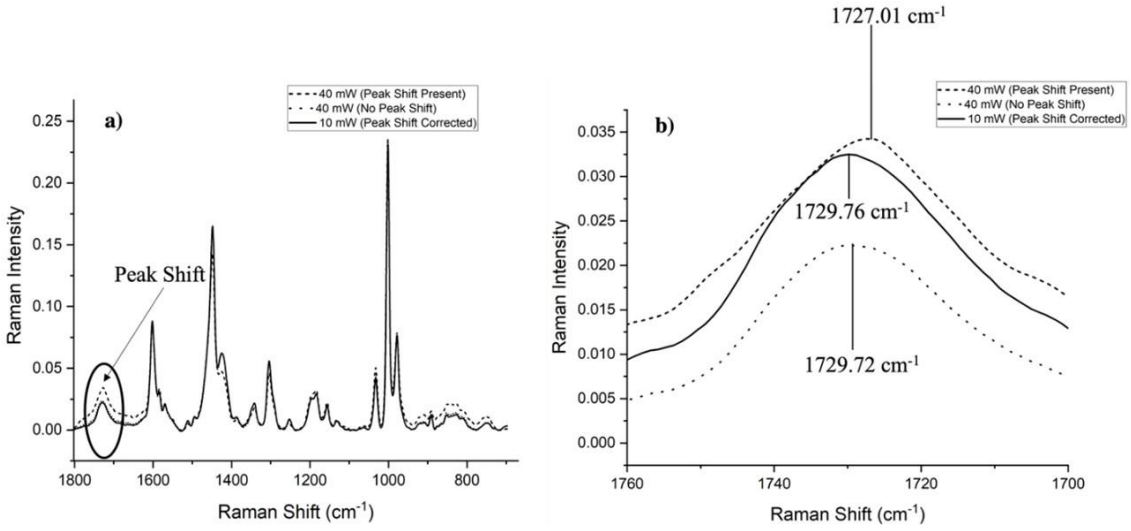

Figure 5.3 A representative Raman spectrum of automotive paint clear (UAZP00796): a) Raman spectra showing the entire region with peak shift region highlighted and b) peak shift correction using 10 mW lower laser power.

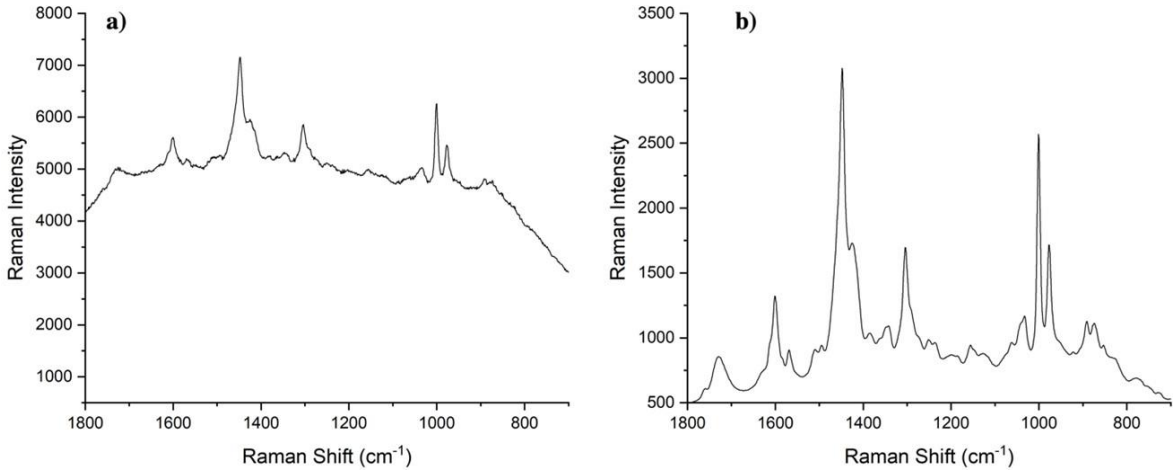

Figure 5.4 A representative Raman spectrum of automotive paint clear coat (CONT00928): a) before photobleaching and b) after photobleaching for 5 minutes.

5.3.2 Pattern Recognition Analysis.

The discrimination power of the average Raman clear coat spectrum and the FTIR spectrum for each sample was computed using pattern recognition methods. For this pattern recognition study, the same spectral range was used for both the Raman and IR spectra (1802 cm^{-1} to 697 cm^{-1}). Classifiers (i.e., search prefilters) were developed to differentiate between Raman and IR clear coat spectra and to identify a clear coat sample as to the assembly plant from which it originated. The Raman and IR spectra were preprocessed using wavelets to enhance subtle but important spectral features in the data. The mother wavelet used should be the one that best matches the shape of the bands present in the spectra. Selection of the appropriate mother wavelet is crucial to ensure that information about assembly plant of the vehicle is successfully extracted from the spectra using the discrete wavelet transform [13], each Raman and IR spectrum was passed through two scaling filters: a high pass filter and a low pass filter. The low-pass filter only allows the low frequency component of the signal to be passed as a set of wavelet coefficients called “approximation” coefficients, whereas the high-pass filter allows only the high frequency component of the signal to be passed as a set of wavelet coefficients set called “detail” coefficient. This process of decomposition is continued with different scales of the wavelet filter pair in a step-by-step manner until the necessary level of signal decomposition has been achieved. For this study, the Symlet 6 mother wavelet at the eighth level of decomposition was utilized for both the Raman and IR spectra of the clear coats.

Wavelet coefficients characteristic of the assembly plant of the vehicle for both the Raman and FTIR spectra were identified using the pattern recognition GA, which takes advantage of both supervised and unsupervised learning to identify wavelet coefficients that optimize the separation of the vibrational spectra by assembly plant in a plot of the two or three largest principal components of the data. Because principal components (PCs) maximize variance, the bulk of the information encoded by the wavelet coefficients selected by the pattern recognition GA is about the differences between assembly plants in the General Motors dataset. A principal component (PC) plot that shows separation of the data by assembly plot can only be generated using coefficients whose variance or information is primarily about differences between these assembly plants. The fitness function of the pattern recognition dramatically reduces the size of the search space as it limits the search to these types of wavelet coefficient subsets. In addition, the pattern recognition GA focuses on those classes and/or samples that are difficult to classify as it trains by boosting the relative importance of the classes and samples that are consistently misclassified. Over time, the algorithm learns its optimal parameters in a manner similar to a neural network. The pattern recognition GA integrates aspects of artificial intelligence and evolutionary computations to yield a "smart" one-pass procedure for wavelet coefficient selection and pattern classification.

Because of the similarity of the IR spectra, variable selection was the logical step as deletion of uninformative wavelet coefficients ensures that discriminatory information about the assembly plant of the vehicle is the major source of variation in the data. If the uninformative wavelet coefficients are retained, their presence is often detrimental as the informative wavelet coefficients are swamped out by the large amount of qualitative and

quantitative data due to other sources of variation in the spectra. The pattern recognition GA is able to identify informative wavelet coefficients by sampling key feature subsets, scoring their PC plots, and tracking those assembly plants and/or spectra that are difficult to classify. The boosting routine uses this information to steer the population to an optimal solution. After 200 generations, the pattern recognition GA is able to identify a set of wavelet coefficients that contain information about the pattern recognition problem of interest, which is the assembly plant of the vehicle from which each paint sample originated.

For this study, the fitness function of the pattern recognition GA was modified to allow for incorporation of model inference into the variable selection process. The goal is to identify variables that minimize the error across the entire model. This was accomplished by assessing the uncertainty of the sample scores in the principal component plot using the jack-knife [14] to generate estimates of dispersion. During each generation, the fitness function of the pattern recognition GA evaluates thousands of principal component plots, one for each feature subset (i.e., chromosome) in the population of solutions. For each principal component score plot, the corresponding training set samples are removed one at a time, and the score matrix and loading matrix for the resampled (i.e., jackknifed) training set is recomputed. (Due to the rotational ambiguities of PCA, the loading matrix for each resampled training set must be rotated using a Procrustean rotation [15] to match the loading matrix associated with the score plot containing all the samples.) For each training set sample, scores across all leave-one-out score plots are projected onto the original principal component plot of the feature subset which is then scored using the fitness function. Thus, information about the level of confidence in the classification of

each training set sample is directly incorporated into the variable selection process with the jack-knifed scores for each sample effectively comprising an error cloud to depict the uncertainty associated with each training set sample.

5.4 Results and Discussion

The data cohort (118 Raman and IR spectra) was divided into 16 training and prediction set pairs (see Table 5.8). Each sample was present in the prediction set only once. Six samples were flagged as outliers by the generalized distance test at the 0.01 level implemented using SCOUT [16]. Three were from Fairfax (CONT01472, CONT01475, and UNVL00007) and three were from Lansing (CONT00985, CONT01049 and UVAC00134). In addition, a visual examination of the Raman spectra of these six samples revealed small but noticeable differences compared to the other 112 spectra. Therefore, the six outliers were removed from the training.

Two studies were performed on the 112 Raman and IR spectra. In the first study, a hierarchical classification of the spectral data was undertaken to solve a six-way classification problem. In the second study, the linear separability of each class was assessed using the Raman and IR spectra from each assembly plant. These two studies are described in detail below.

5.4.1 Hierarchical Classification Study

In the first study, a six-way classification of the spectra was formulated using a two-step approach. In the first step, the Arlington assembly plant was discriminated from the other five assembly plants (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine)

using the pattern GA to identify informative features in the Raman spectra. The results for each Raman training set/prediction set pair (Arlington versus Doraville, Fairfax, Fort Wayne, Lansing, and Moraine), including, and excluding the sample outliers, are summarized in Figures 5.5 to 5.20. For each pair, Raman spectra comprising the prediction set were projected onto the PC plot of the training set samples and the wavelet coefficients identified by the pattern recognition GA. If the prediction set samples are correctly classified, the Raman spectra representing the clear coats from the Arlington assembly plant will lie in a region of the plot with other Arlington paint samples, and the Raman spectra representing the Doraville, Fairfax, Fort Wayne, Lansing, and Moraine assembly plants will lie in a region of the PC plot containing the samples from these five assembly plants

In the second step, the Raman spectra from the other five assembly plants (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) were compared using the pattern recognition GA to solve a five-way classification problem. For each training set/prediction pair, the actual prediction set samples are those in the first step that were projected onto the region of the PC plot encompassing the other five assembly plants. Prediction set samples from Doraville, Fairfax, Fort Wayne, Lansing, or Moraine classified as Arlington in the first step were not passed to the discriminant (i.e., PC score plot) developed for the simultaneous classification of Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. If a Doraville prediction set sample is correctly classified, for example, the sample will lie in a region of the plot with other Doraville paint samples. The results for the second step are summarized in Figures 5.21 to 5.36 for each training set/prediction set pair. The combined

results (Step 1 and Step 2) for the 16-training set/prediction set pairs are summarized in Table 5.9.

A hierarchical classification study of the 118 IR spectra was also undertaken using this same two-step process. The results for each IR training set/prediction set pair (Arlington versus Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) are summarized in Figures 5.37 to 5.52 and the results of the five-way classification problem (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) are summarized in Figures 5.53 to 5.68. The combined IR results for the 16-training set/prediction set pairs are summarized in Table 5.10.

For the Raman spectra, 105 of the 112 samples or 93.75% of the samples in the prediction sets were correctly classified. By comparison, 75 of the 112 IR spectra or 66.96% of the samples in the prediction set were correctly assigned to their respective class. These results demonstrate that information derived solely from the clear coats can categorize paint samples as to assembly plant. Furthermore, Raman spectroscopy and pattern recognition appear to be a better solution to the problem of extracting investigative lead information from automotive clear coats than IR spectroscopy and pattern recognition.

5.4.2 Two-way Classifications

In the second study, each assembly is compared to the other five assembly plants in a two-way classification. This is the simplest approach to take for comparing Raman and IR spectra from each assembly plant using pattern recognition techniques. The feature space is divided into two regions. Clear coat paint samples from a specific assembly plant will be found in one region of the space while those paint samples comprising the other

five classes will be found in a different region of the space. Any sample in the dataset can be classified into one of two categories by obtaining the coordinates (i.e., scores) of the sample in the PC plot developed from the training set samples and the wavelet coefficients identified by the pattern recognition GA.

For the two-way classification studies, both the Raman and IR spectral data were again divided into 16-training set prediction set pairs (see Table 5.8). Again, each sample was a member of the prediction set only once. The results for Arlington are summarized in Figures 5.5 to 5.20 and Figures 5.37 to 5.52. For the other five assembly plants, Doraville, Fairfax, Fort Wayne, Lansing, and Moraine, the same two-way classification studies were performed, and the results for the Raman and IR spectra are summarized in Figures 5.69 to 5.137 and Figures 5.138 to 5.206. The combined results (Arlington, Doraville, Fairfax, Fort Wayne, Lansing, and Moraine assembly plants) for the 16-training set/prediction set pairs are summarized in Tables 5.11 and 5.12.

For the Raman spectra, 97 of the 112 samples or 86.6% of the samples in the prediction sets were correctly classified. Three of the six assembly plants (Doraville, Fairfax, and Moraine) were linearly separable. By comparison, only 61 of the 112 IR spectra or 54.5% of the samples in the prediction set were correctly assigned to their respective class. Furthermore, no assembly plant was linearly separable from the other assembly plants. Although in a previous study [5] we reported that Moraine could be differentiated from the other five assembly plants using FTIR spectroscopy, only the apparent classification success rate of the clear coat paint samples was computed for each class. By comparison, cross validation (which was used in this study) is a more rigorous

procedure to assess differences between the Raman and IR spectra in each dataset and establish the quality of a model for classification [17].

The cross-validation procedure used in this study [18] differs from the procedure used in other published studies. In this study, the coefficients selected for each training set are different. In most cross-validation studies, the features used in each training set are the same. These features are identified using the entire data set prior to dividing the data into training set/prediction set pairs. For this reason, cross validation will often give overly optimistic estimates of the error rate. In this study, the prediction (validation) set samples do not influence the coefficients that are selected for each training set as the coefficients are chosen independent of the samples in each of the prediction sets by the pattern recognition GA. Hence, the error rate reported by the cross-validation procedure used in this study is less biased.

The chemical formulation of the automotive clear coats used in this study is acrylic melamine styrene or acrylic melamine styrene polyurethane. Both Raman and IR spectra have distinct and recognizable peaks for melamine, styrene, and urethane. As for the acrylate component, there are eight distinct acrylate monomers that could be present in a typical clear coat formulation [19]. The key bands in the IR spectra of acrylates in clear coats are the carbonyl stretch and the C─O─C antisymmetric stretch. The latter band may be characteristic of specific acrylate monomers and although allowed in IR, is too weak to be observed unequivocally in the IR spectra of clear coats because of the presence of bands due to other components. Raman spectra, on the other hand, contain acrylate bands not only of the carbonyl stretching mode but also of the C─O─C symmetric stretch, which is

characteristic of specific acrylate monomers, is less overlapped by neighboring bands in the Raman spectra of clear coats and is intense enough to be clearly observed [20].

Table 5.8 Prediction Set for Each Training Set/Prediction Set Pair for Raman and IR Data

Prediction Set 1		
Plant ID	Plant Name	Prediction Set Sample
1	Arlington	UTNN00054
4	Doraville	CONT01505
5	Fairfax	UAZP00998
8	Fort Wayne	UKYF00200
14	Lansing	CQCM00137
18	Moraine	CONT01483 and CONT00918
Prediction Set 2		
Plant ID	Plant Name	Prediction Set Sample
1	Arlington	UAZP00796 and UWVC00116
4	Doraville	CONT00475
5	Fairfax	UAZP00999
8	Fort Wayne	UNVL00068
14	Lansing	UAZP00567
18	Moraine	UAZP00430 and UTXA00206
Prediction Set 3		
Plant ID	Plant Name	Prediction Set Sample
1	Arlington	UAZP00797 and UARL00113
4	Doraville	CONT00699
5	Fairfax	UCAD00049
8	Fort Wayne	UOHL00234
14	Lansing	UNCR00360
18	Moraine	UAZP00431 and CONT01583
Prediction Set 4		
Plant ID	Plant Name	Prediction Set Sample
1	Arlington	UCAD00069 and UAZP00242
4	Doraville	CONT01501
5	Fairfax	UCAD00067
8	Fort Wayne	UOKO00115

14	Lansing	UOHL0050
18	Moraine	CONT00920 and UAZP00498
Prediction Set 5		
Plant ID	Plant Name	Prediction Set Sample
1	Arlington	UCAD00141 and UCAD00196
4	Doraville	CONT01504
5	Fairfax	UCAD00068
8	Fort Wayne	UOKO00153
14	Lansing	UOCN00015
18	Moraine	UCAD00033
Prediction Set 6		
Plant ID	Plant Name	Prediction Set Sample
1	Arlington	UCAD00229 and UAZP01122
4	Doraville	N/A
5	Fairfax	UCAD00071
8	Fort Wayne	UTXA00172
14	Lansing	UMNP00151
18	Moraine	UKYF00171
Prediction Set 7		
Plant ID	Plant Name	Prediction Set Sample
1	Arlington	UMNP00135 and UCAD00147
4	Doraville	N/A
5	Fairfax	UCAD00086
8	Fort Wayne	UWVC00140
14	Lansing	CONT01550
18	Moraine	UNCR00377 and UVAC00143
Prediction Set 8		
Plant ID	Plant Name	Prediction Set Sample
1	Arlington	UAZP00504 and USCC00246
4	Doraville	N/A
5	Fairfax	UCAD00089
8	Fort Wayne	CONT01394
14	Lansing	CONT01474
18	Moraine	UNJH00055
Prediction Set 9		
Plant ID	Plant Name	Prediction Set Sample
1	Arlington	UMNP00180 and UTXA00234
4	Doraville	N/A

5	Fairfax	UCAD00179
8	Fort Wayne	CONT01395
14	Lansing	UCAR00180
18	Moraine	UOHL00116
Prediction Set 10		
Plant ID	Plant Name	Prediction Set Sample
1	Arlington	UAZP00437 and UKYF00203
4	Doraville	N/A
5	Fairfax	UCAD00181
8	Fort Wayne	CONT01396
14	Lansing	UOHL00139
18	Moraine	UOKO00214
Prediction Set 11		
Plant ID	Plant Name	Prediction Set Sample
1	Arlington	UNVL00074 and UMNP00141
4	Doraville	N/A
5	Fairfax	UNJH00101
8	Fort Wayne	CONT01397
14	Lansing	UWVC00123
18	Moraine	USCC00088
Prediction Set 12		
Plant ID	Plant Name	Prediction Set Sample
1	Arlington	CONT00902, UAZP00437 and USCC00301
4	Doraville	N/A
5	Fairfax	UNVL00059
8	Fort Wayne	CONT01515
14	Lansing	UTNN0056
18	Moraine	USCC00313
Prediction Set 13		
Plant ID	Plant Name	Prediction Set Sample
1	Arlington	CONT01861 and UAZP00435
4	Doraville	N/A
5	Fairfax	UNYV00138
8	Fort Wayne	CONT01856
14	Lansing	UMNP00186
18	Moraine	UAZP00432
Prediction Set 14		
Plant ID	Plant Name	Prediction Set Sample
1	Arlington	UAZP00396 and UAZP00798
4	Doraville	N/A

5	Fairfax	CONT01545
8	Fort Wayne	UARL00038
14	Lansing	UTNN00057 and CONT01046
18	Moraine	UOHL00230
Prediction Set 15		
Plant ID	Plant Name	Prediction Set Sample
1	Arlington	UAZP00391 and UAZP00799
4	Doraville	N/A
5	Fairfax	UARL00039
8	Fort Wayne	UARL00085
14	Lansing	CONT01522 and CONT00984
18	Moraine	CONT00920 and USCC00008
Prediction Set 16		
Plant ID	Plant Name	Prediction Set Sample
1	Arlington	UAZP00333, USCC00513 and USCC00514
4	Doraville	N/A
5	Fairfax	UAZP00801 and UAZP00802
8	Fort Wayne	UCAR00191
14	Lansing	CONT01042 and CONT01047
18	Moraine	UWAS00001

Table 5.9 Summary of Raman Results for First Study

Prediction Set Raman	Raman (Outliers Removed)	
Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine	Correct	Wrong
Total (First Step)	27	6
Five-way classification (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine)	Correct	Wrong
Total (Second Step)	78	1
Total (First Step + Second Step)	105	7

Table 5.10 Summary of IR Results for First Study

Prediction Set IR	IR (Outliers Removed)	
Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine	Correct	Wrong
Total (First Step)	18	15
Five-way classification (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine)	Correct	Wrong
Total (Second Step)	57	22
Total (First Step + Second Step)	75	37

Table 5.11 Summary of Raman Results for Second Study

Prediction Set	Correct	Wrong	Total
Arlington versus Doraville, Fairfax, Fort Wayne, Lansing, and Moraine	27	6	33
Doraville versus Arlington, Fairfax, Fort Wayne, Lansing, and Moraine	5	0	5
Fairfax versus Arlington, Doraville, Fort Wayne, Lansing, and Moraine	17	0	17
Fort Wayne versus Arlington, Doraville, Fairfax, Lansing, and Moraine	14	2	16
Lansing versus Arlington, Doraville, Fairfax, Fort Wayne, and Moraine	12	7	19
Moraine versus Arlington, Doraville, Fairfax, Fort Wayne, Lansing, and Moraine	22	0	22
Total	97	15	112

Table 5.12 Summary of IR Results for Second Study

Prediction Set	Correct	Wrong	Total
Arlington versus Doraville, Fairfax, Fort Wayne, Lansing, and Moraine	18	15	33
Doraville versus Arlington, Fairfax, Fort Wayne, Lansing, and Moraine	2	3	5
Fairfax versus Arlington, Doraville, Fort Wayne, Lansing, and Moraine	5	12	17
Fort Wayne versus Arlington, Doraville, Fairfax, Lansing, and Moraine	10	6	16
Lansing versus Arlington, Doraville, Fairfax, Fort Wayne, and Moraine	6	13	19
Moraine versus Arlington, Doraville, Fairfax, Fort Wayne, Lansing, and Moraine	20	2	22
Total	61	51	112

First Study Raman: First Step (Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine)

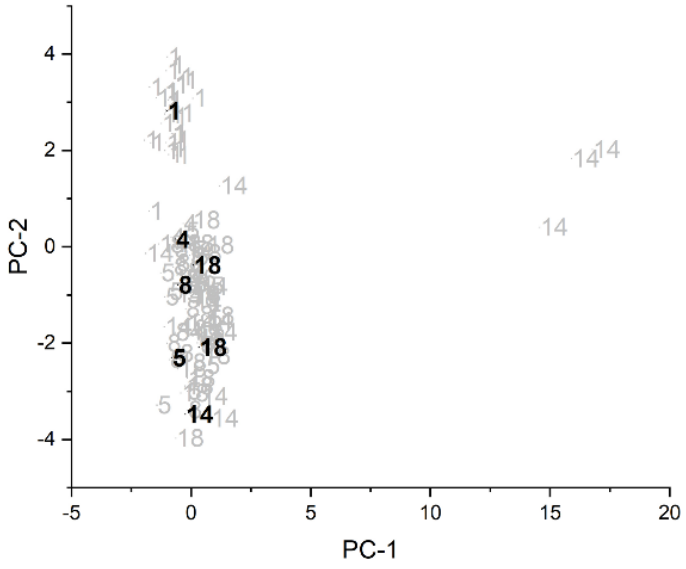

Figure 5.5 PC plot for the first training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

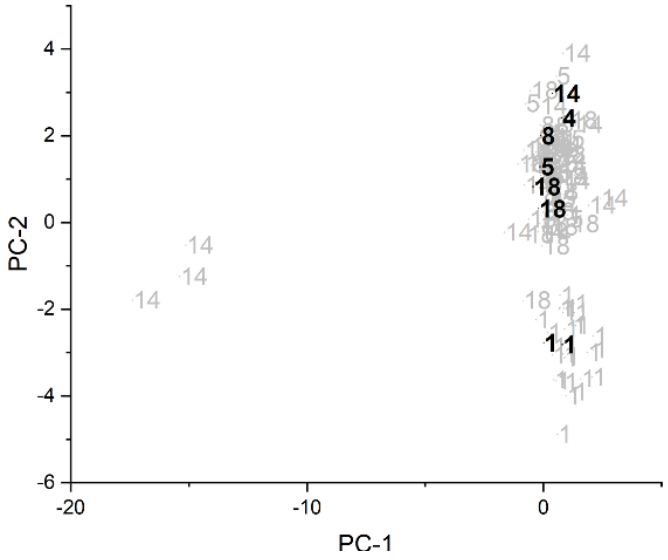

Figure 5.6 PC plot for the second training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

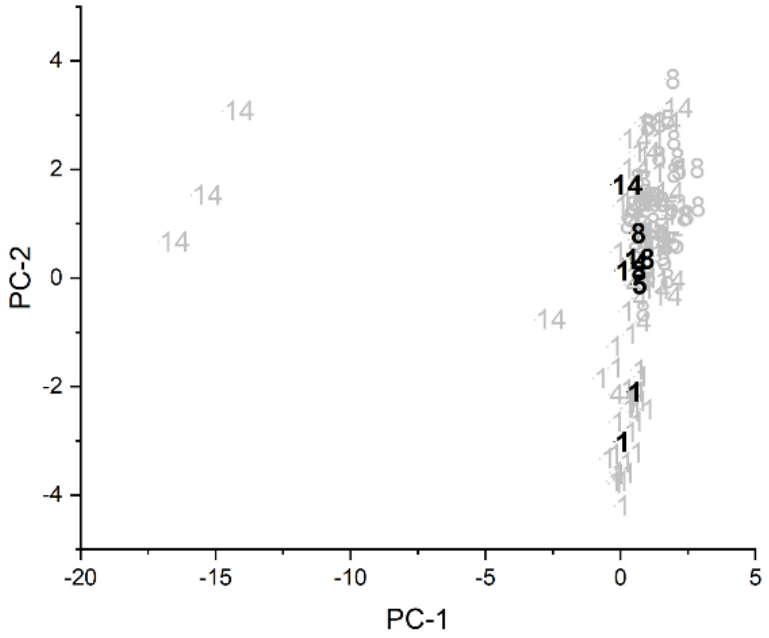

Figure 5.7 PC plot of the third training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

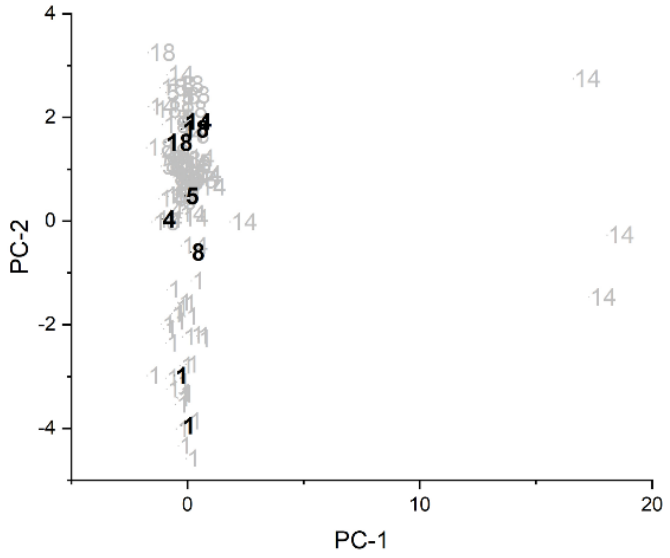

Figure 5.8 PC plot of the fourth training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

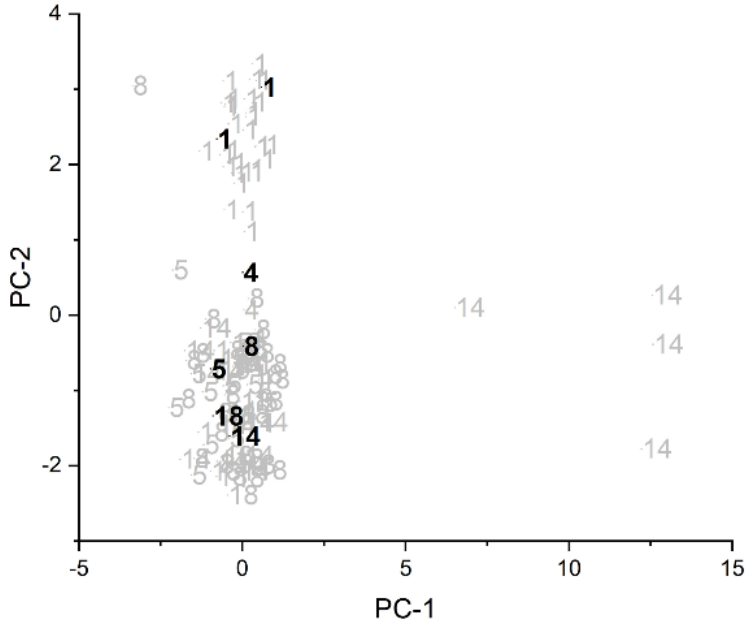

Figure 5.9 PC plot of the fifth training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

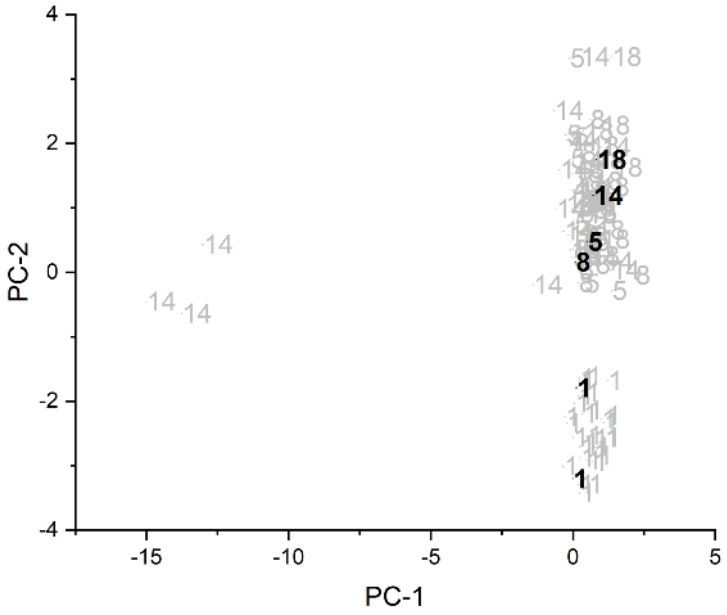

Figure 5.10 PC plot of the sixth training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

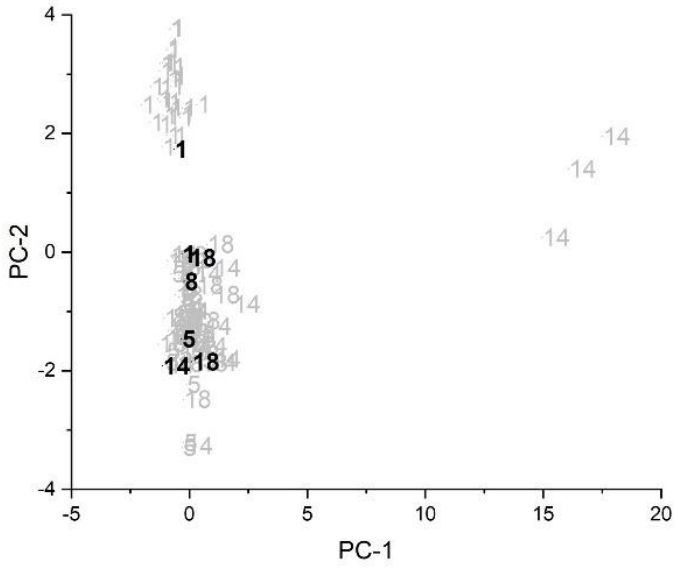

Figure 5.11 PC plot of the seventh training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

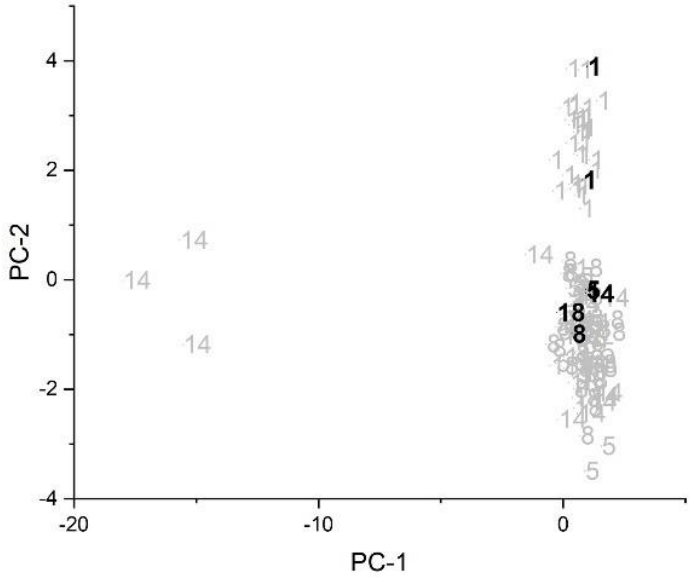

Figure 5.12 PC plot of the eighth training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

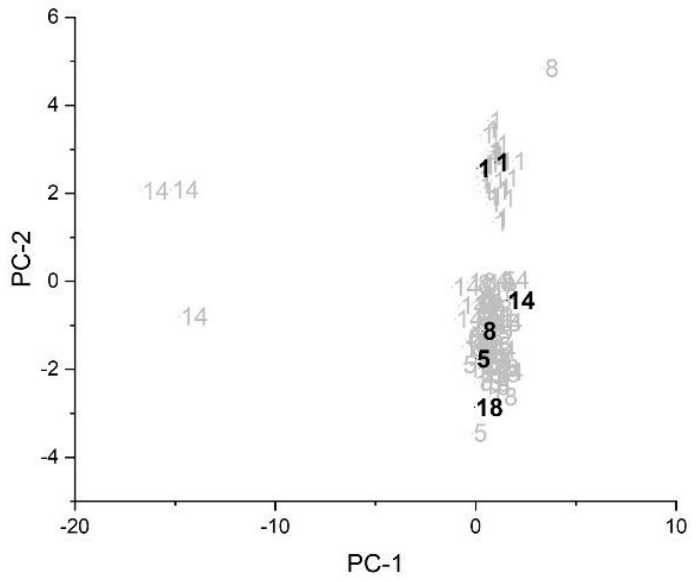

Figure 5.13 PC plot of the ninth training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

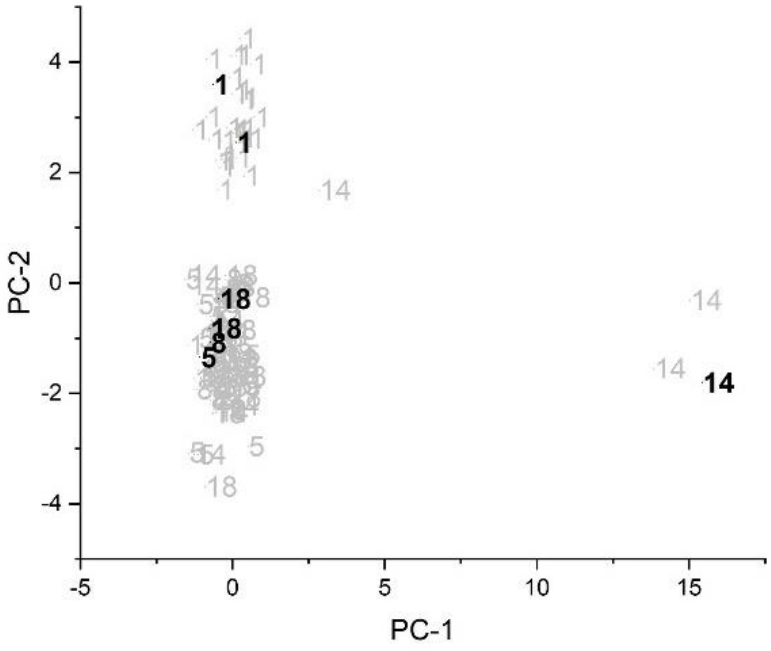

Figure 5.14 PC plot of the tenth training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

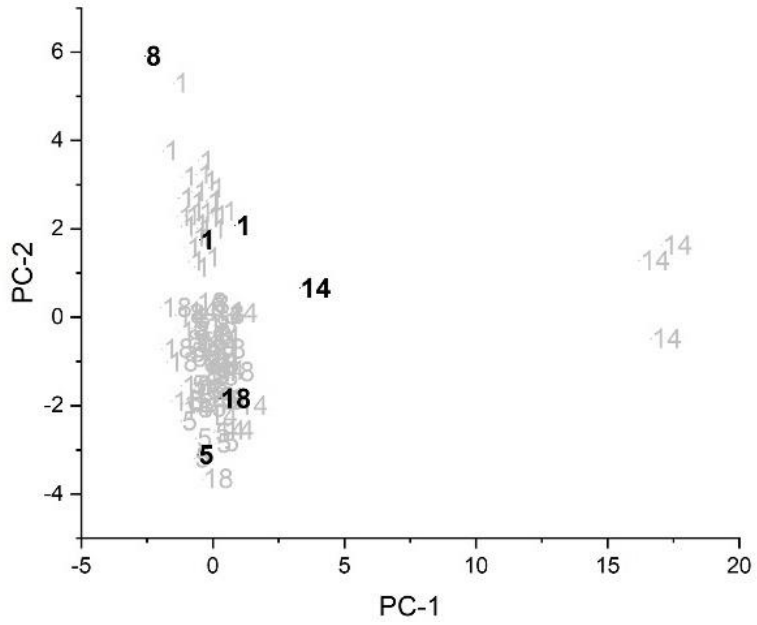

Figure 5.15 PC plot of the eleventh training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

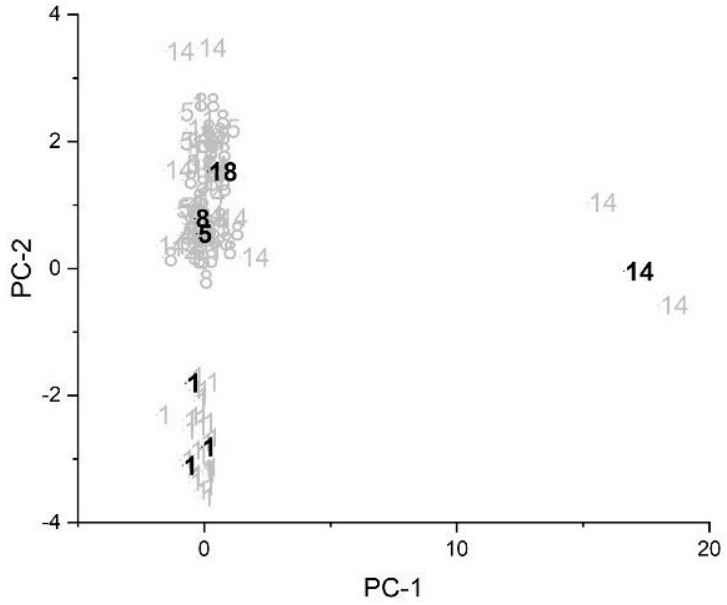

Figure 5.16 PC plot of the twelfth training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

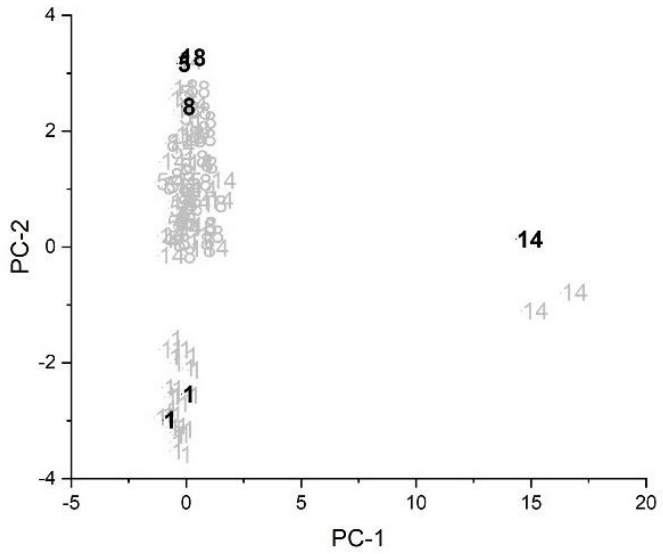

Figure 5.17 PC plot of the thirteenth training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

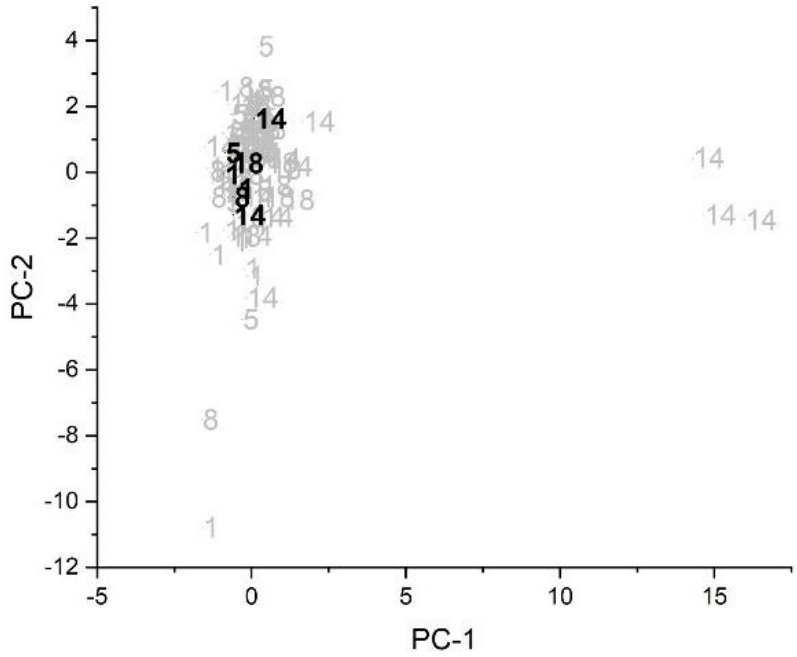

Figure 5.18 PC plot of the fourteenth training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

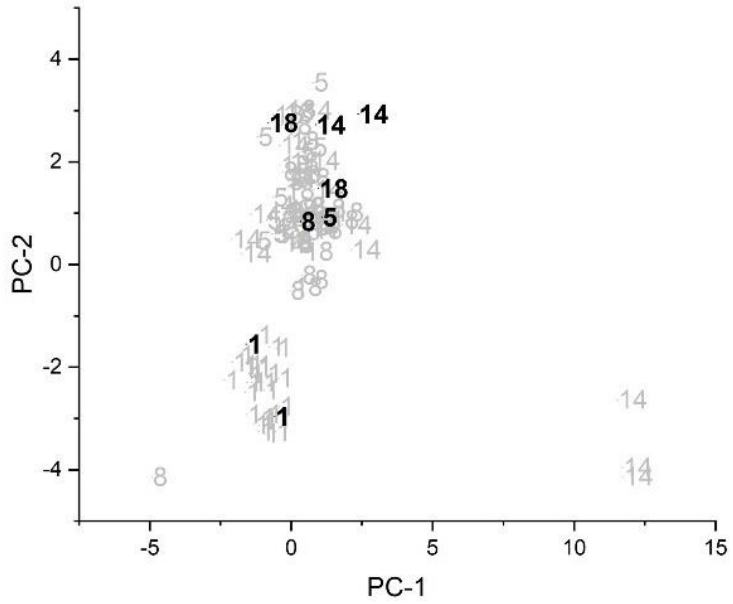

Figure 5.19 PC plot of the fifteenth training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

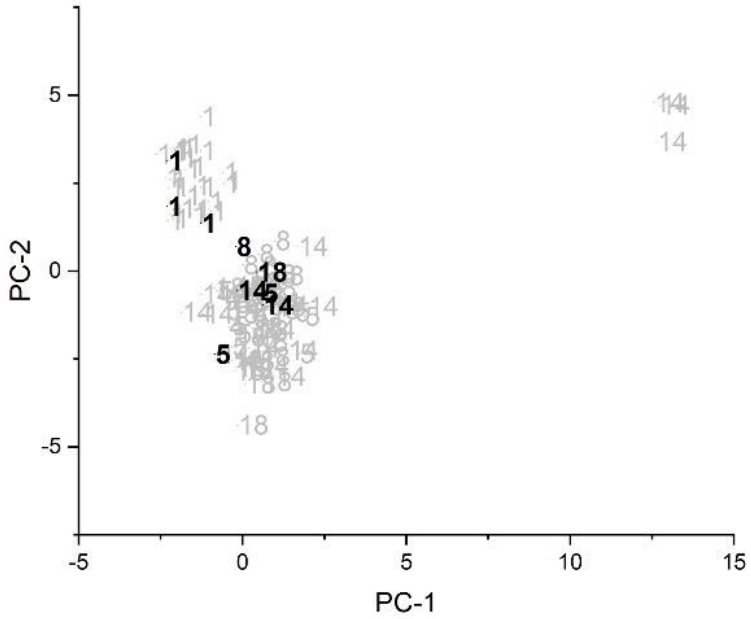

Figure 5.20 PC plot of the sixteenth training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

First Study Raman: Second Step, Five-way classification (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine)

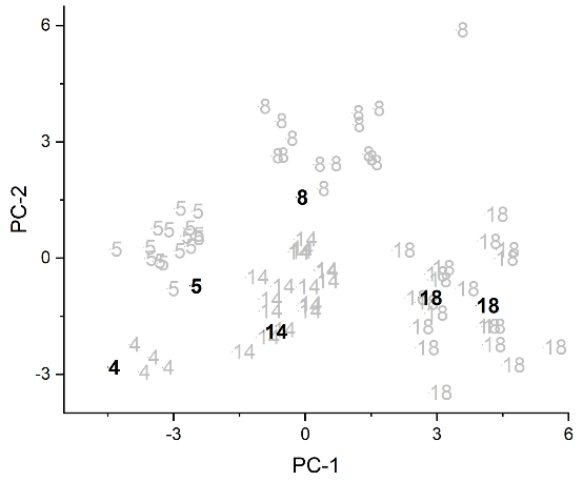

Figure 5.21 PC plot of the first training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

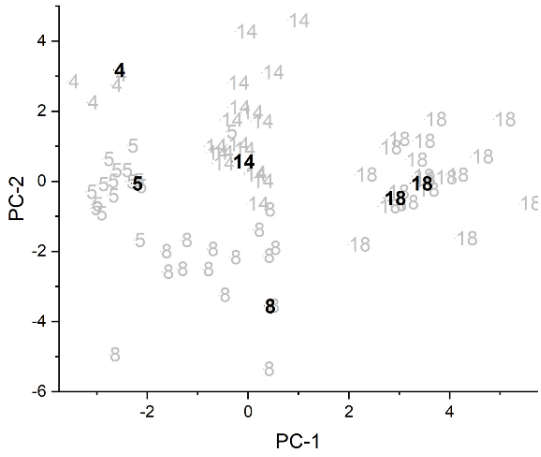

Figure 5.22 PC plot of the second training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

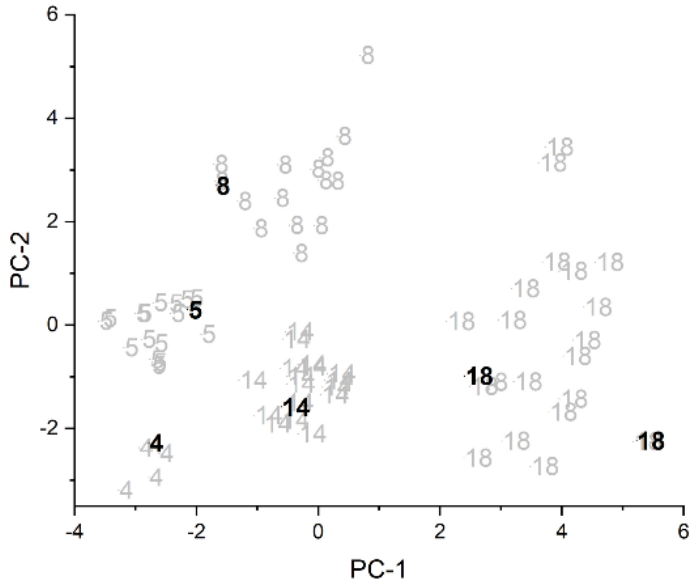

Figure 5.23 PC plot of the third training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

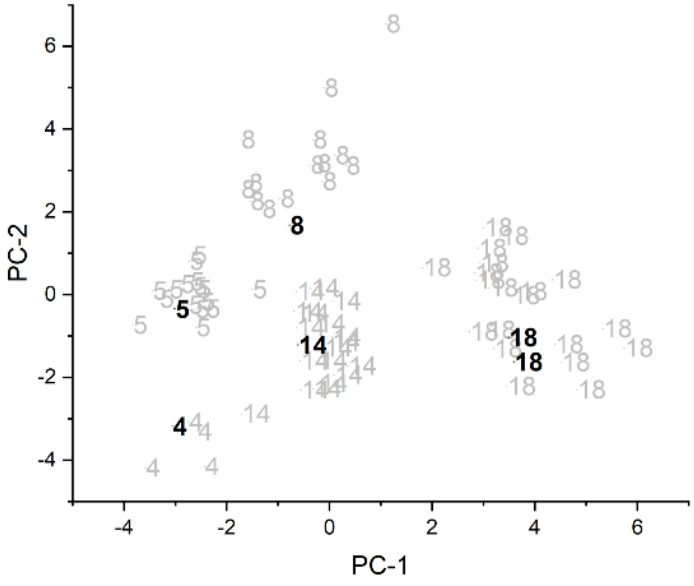

Figure 5.24 PC plot of the fourth training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

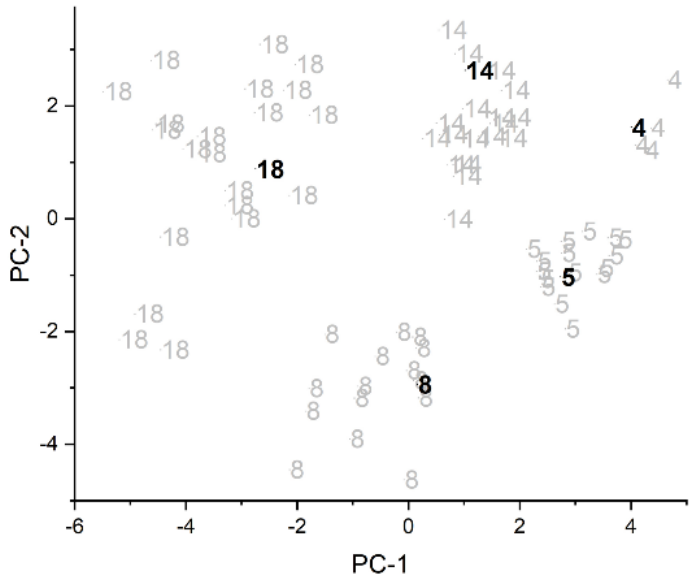

Figure 5.25 PC plot of the fifth training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

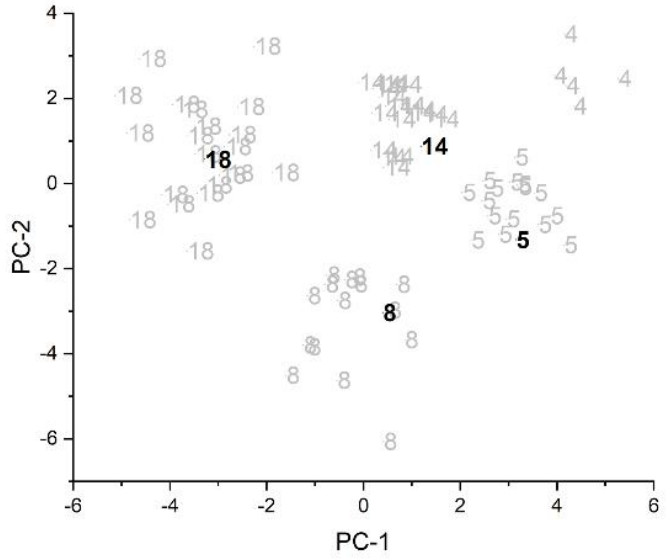

Figure 5.26 PC plot of the sixth training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

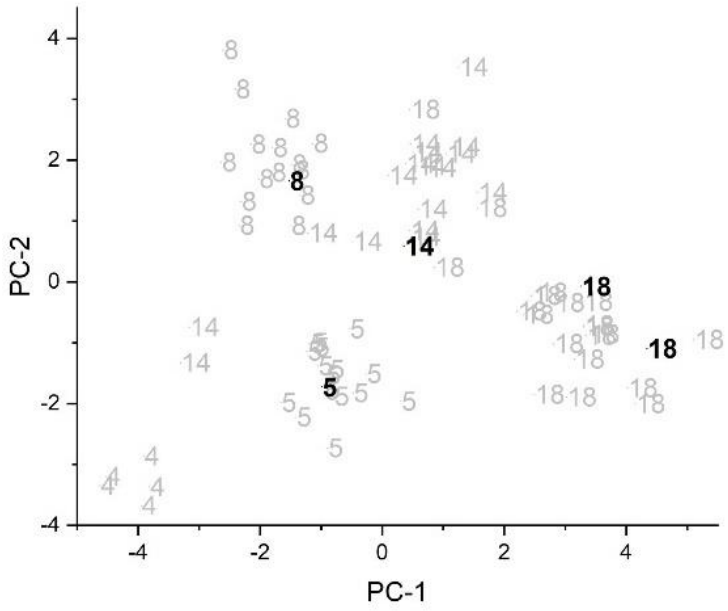

Figure 5.27 PC plot of the seventh training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

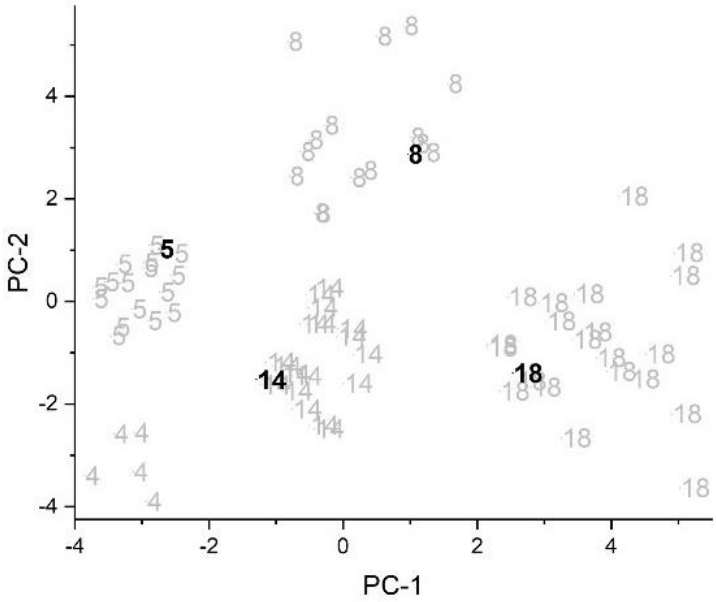

Figure 5.28 PC plot of the eighth training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

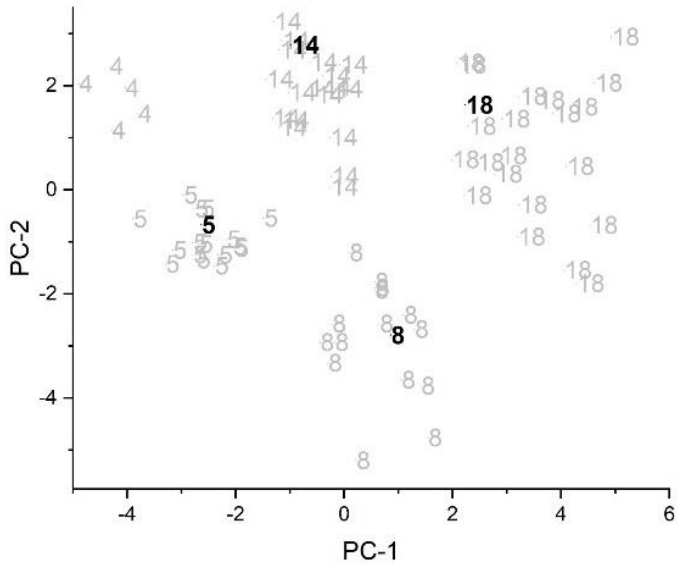

Figure 5.29 PC plot of the ninth training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

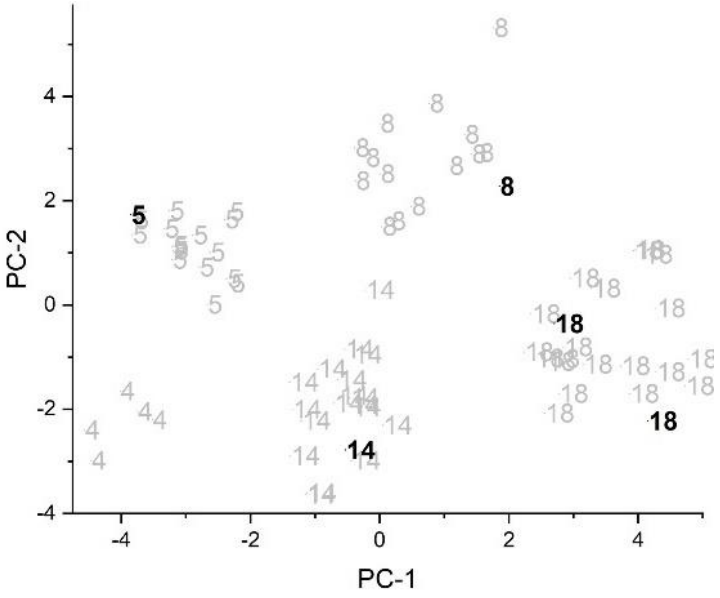

Figure 5.30 PC plot of the tenth training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

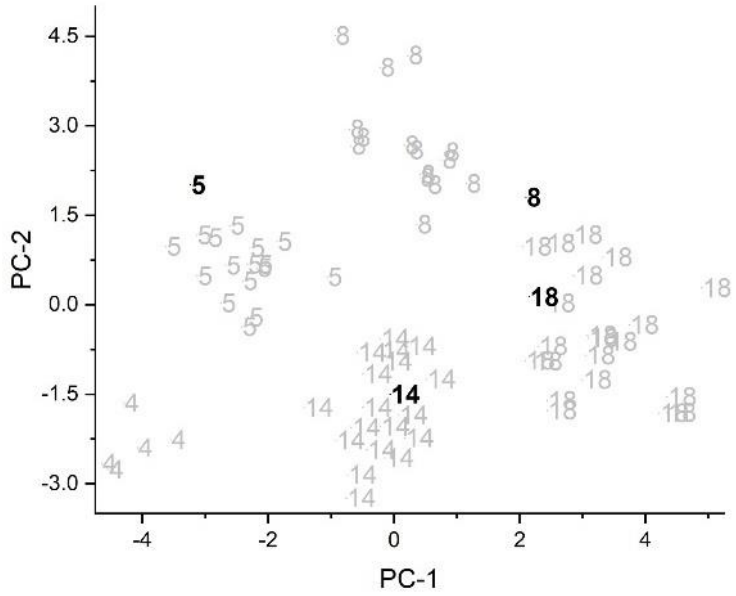

Figure 5.31 PC plot of the eleventh training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

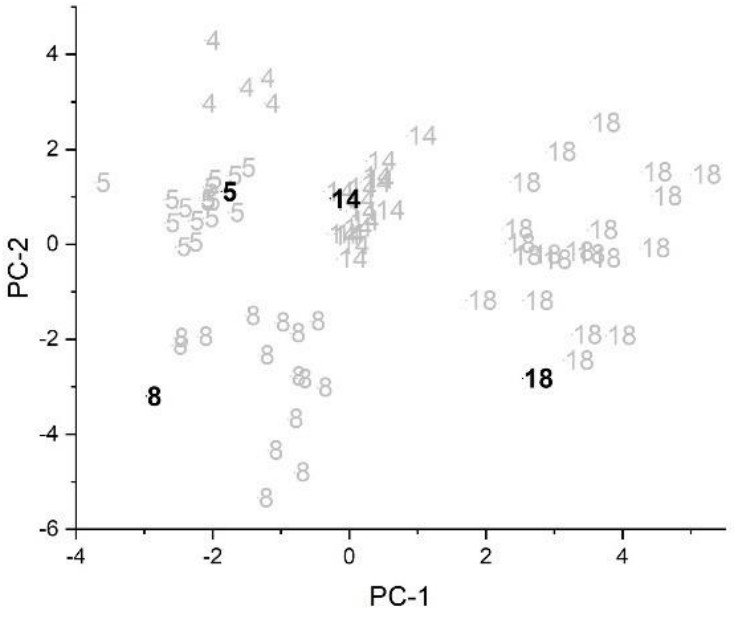

Figure 5.32 PC plot of the twelfth training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

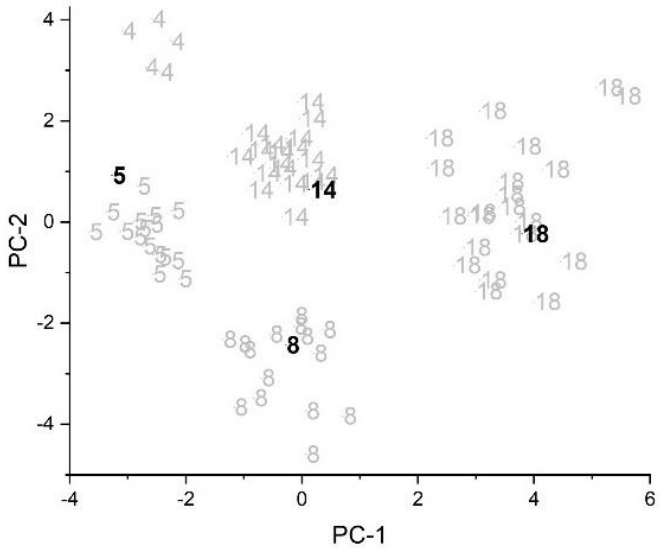

Figure 5.33 PC plot of the thirteenth training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

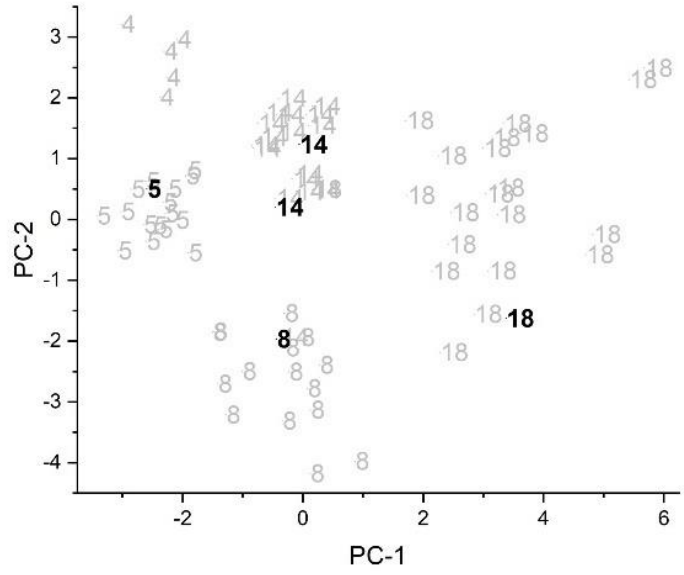

Figure 5.34 PC plot of the fourteenth training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

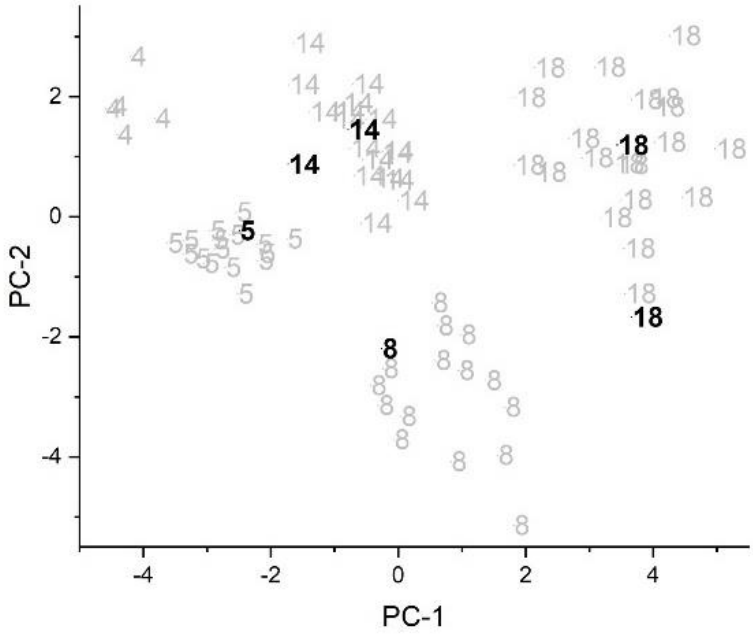

Figure 5.35 PC plot of the fifteenth training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

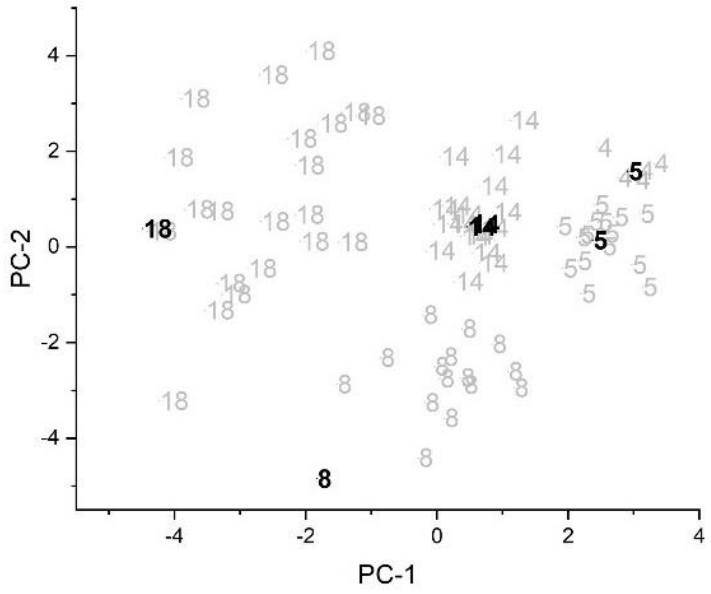

Figure 5.36 PC plot of the sixteenth training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

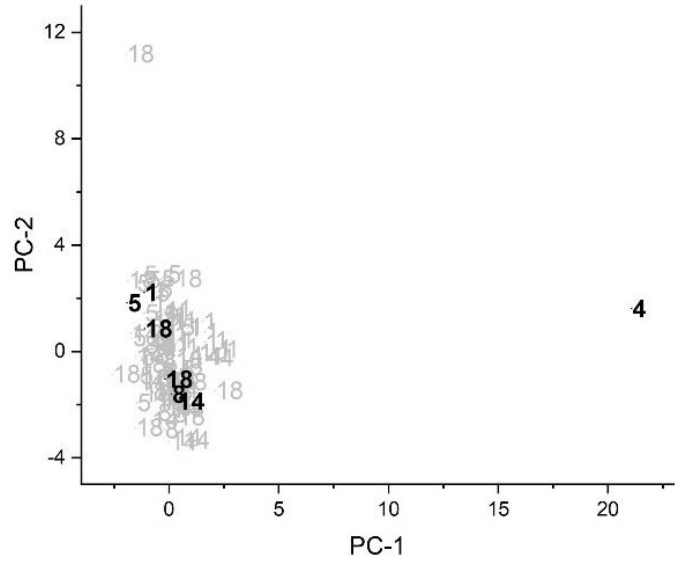

Figure 5.37 PC plot for the first training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

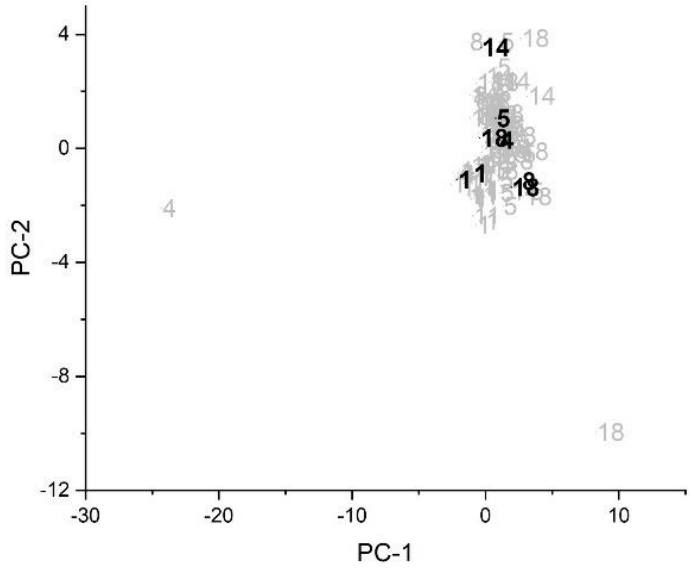

Figure 5.38 PC plot for the second training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

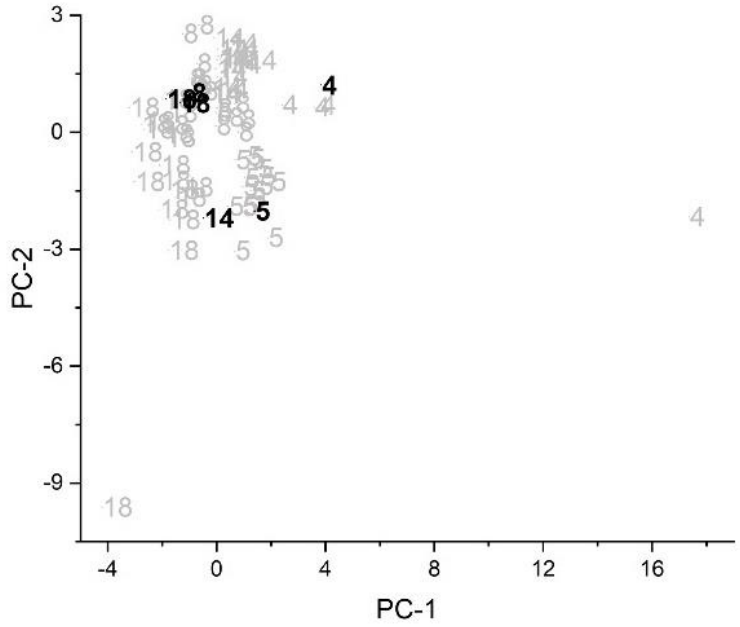

Figure 5.39 PC plot for the third training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

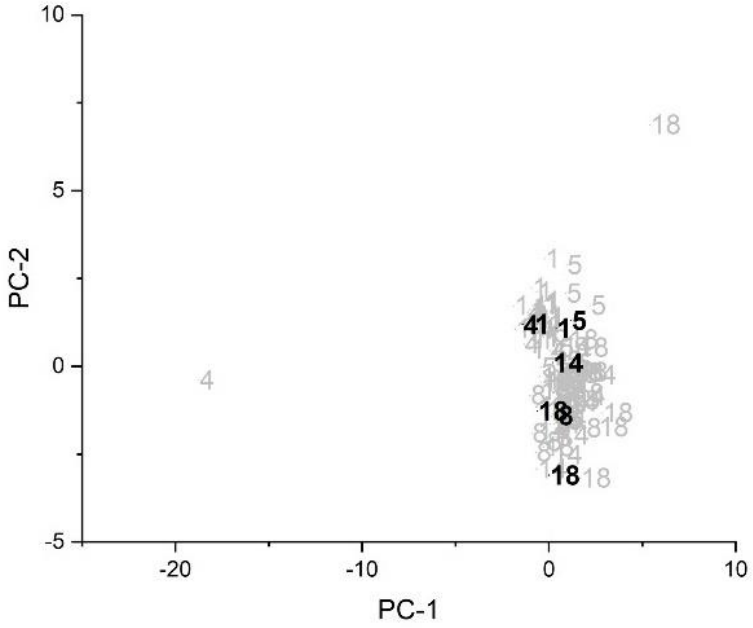

Figure 5.40 PC plot for the fourth training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

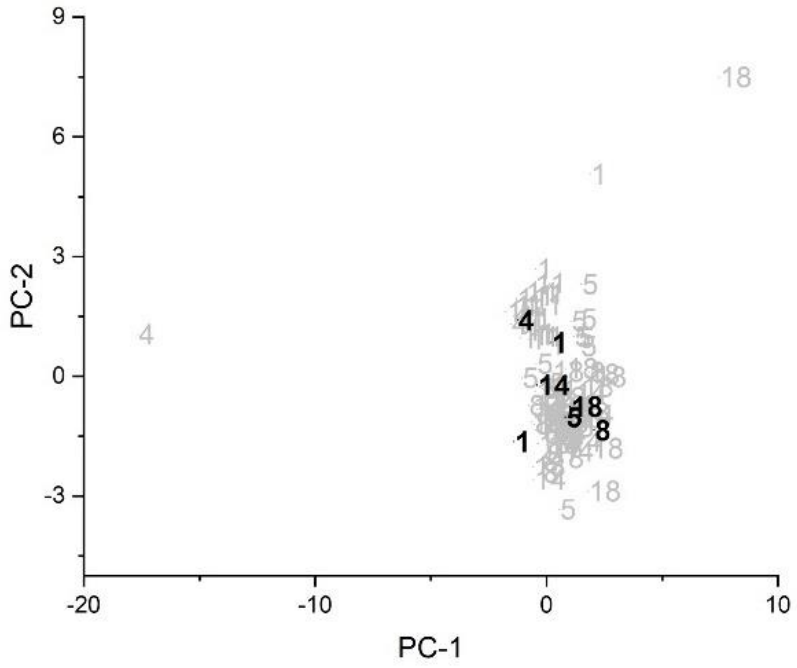

Figure 5.41 PC plot for the fifth training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

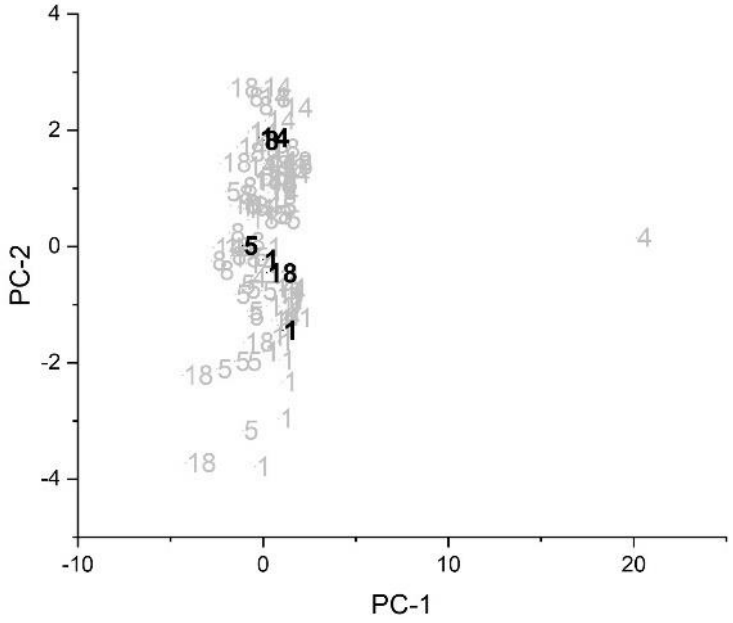

Figure 5.42 PC plot for the sixth training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

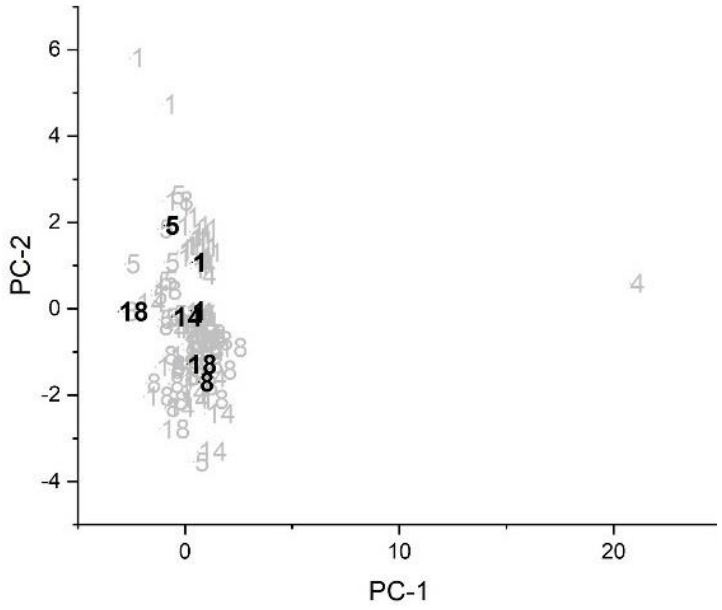

Figure 5.43 PC plot for the seventh training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

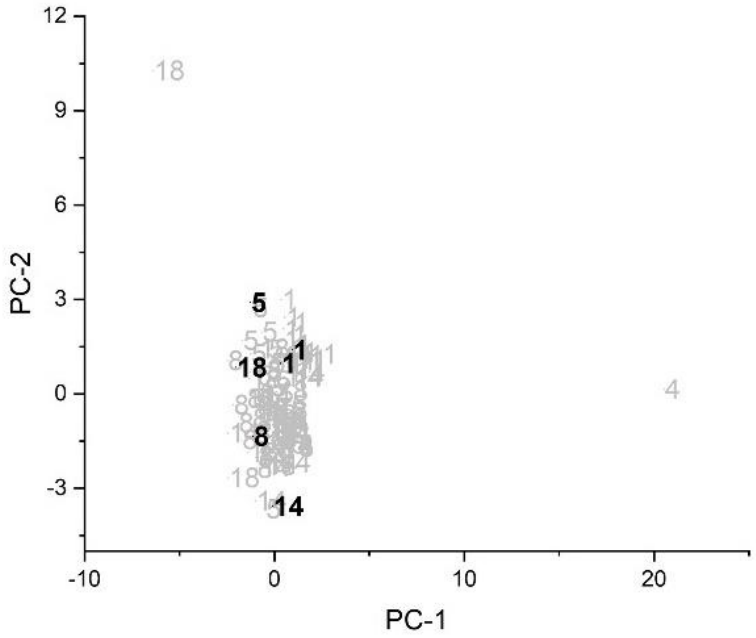

Figure 5.44 PC plot for the eighth training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

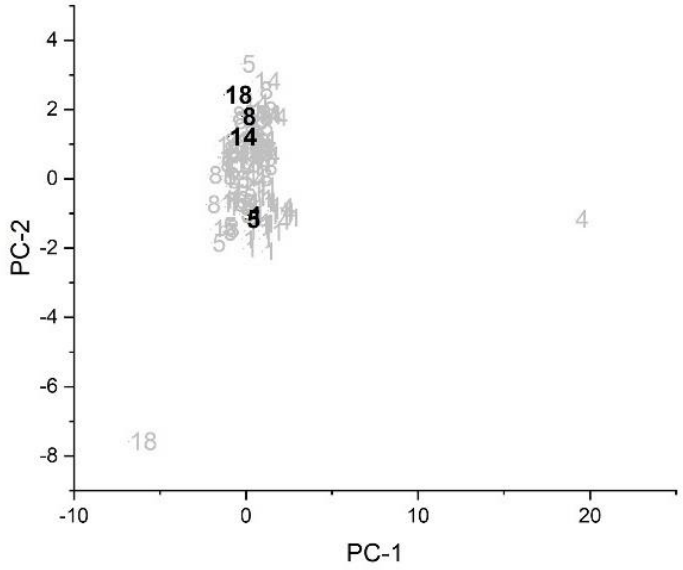

Figure 5.45 PC plot for the ninth training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

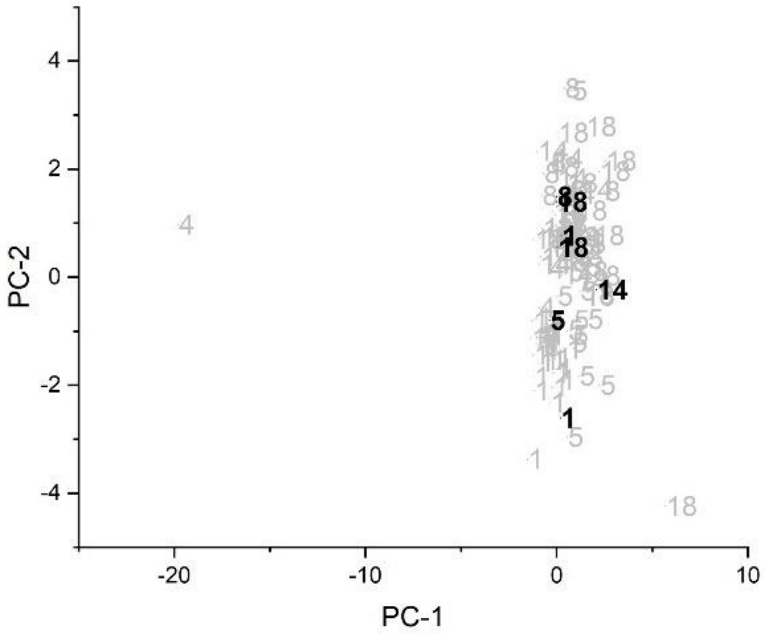

Figure 5.46 PC plot for the tenth training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

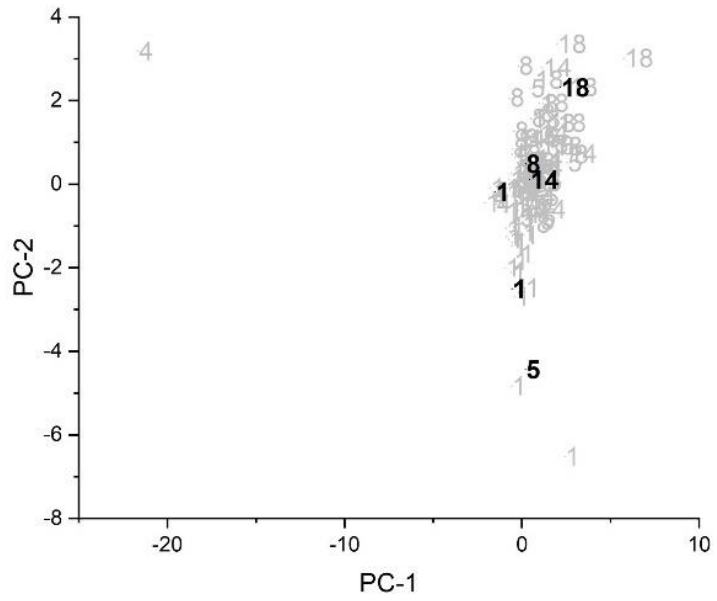

Figure 5.47 PC plot for the eleventh training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

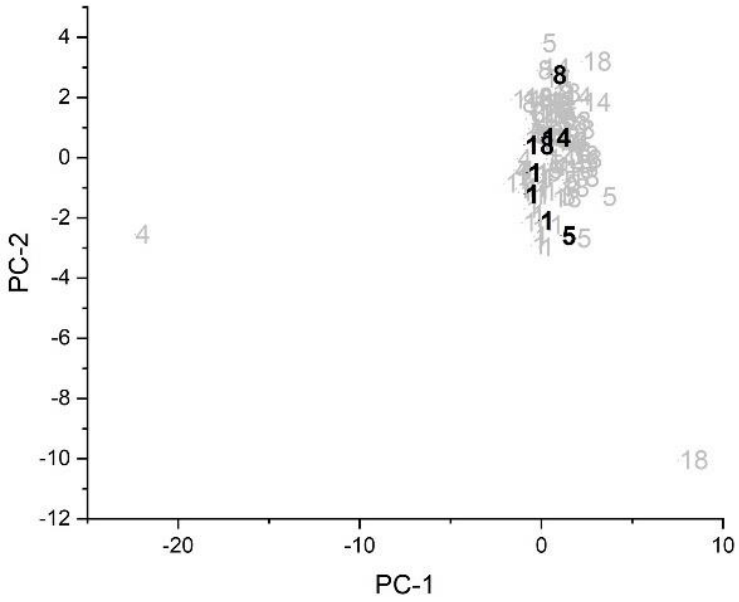

Figure 5.48 PC plot for the twelfth training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

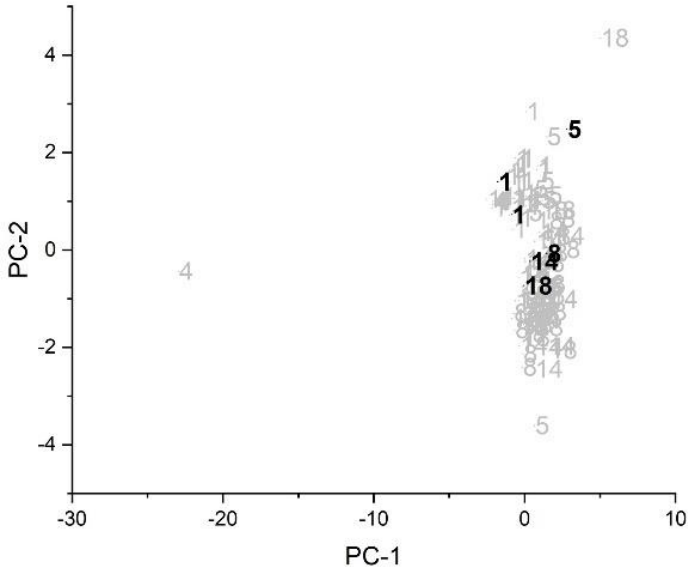

Figure 5.49 PC plot for the thirteenth training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

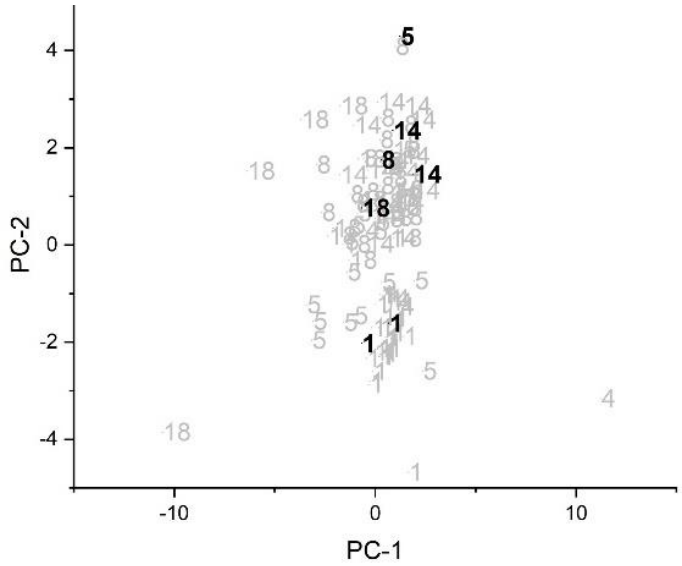

Figure 5.50 PC plot for the fourteenth training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

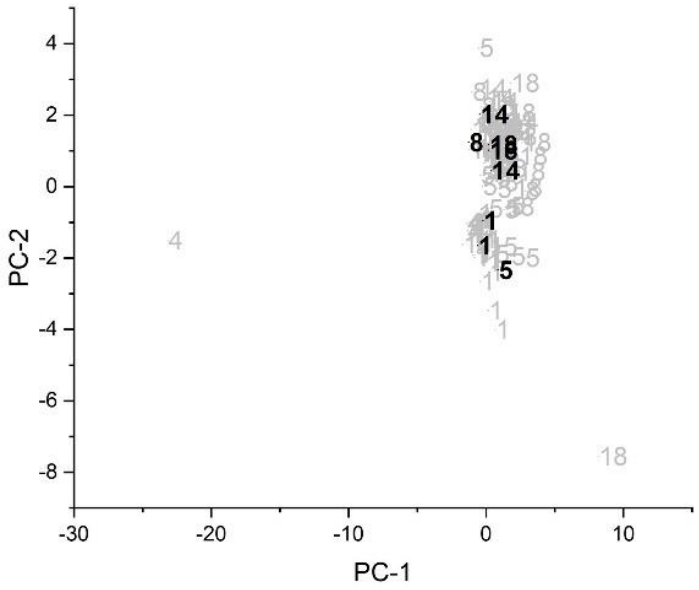

Figure 5.51 PC plot for the fifteenth training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

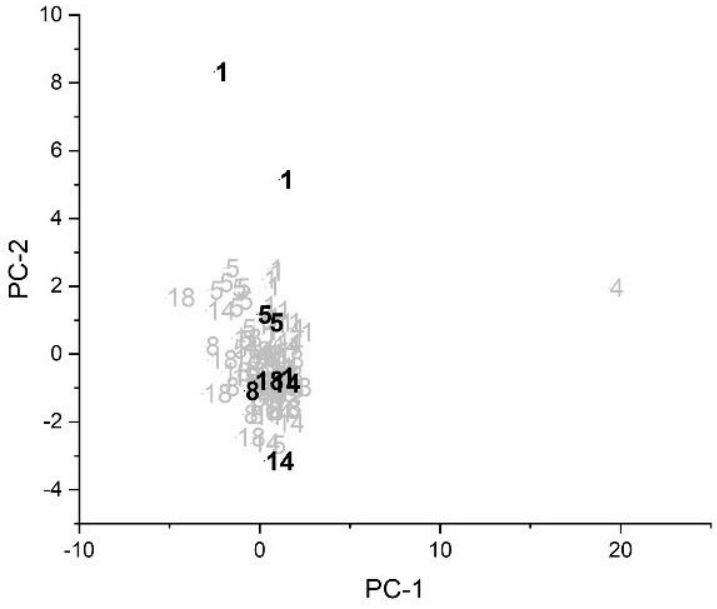

Figure 5.52 PC plot for the sixteenth training set/prediction set pair: Arlington vs Doraville, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

First Study Five-way classification FTIR (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine)

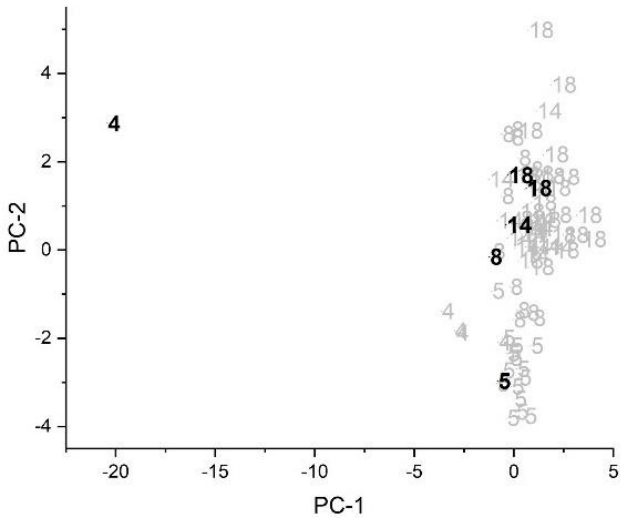

Figure 5.53 PC plot of the first training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

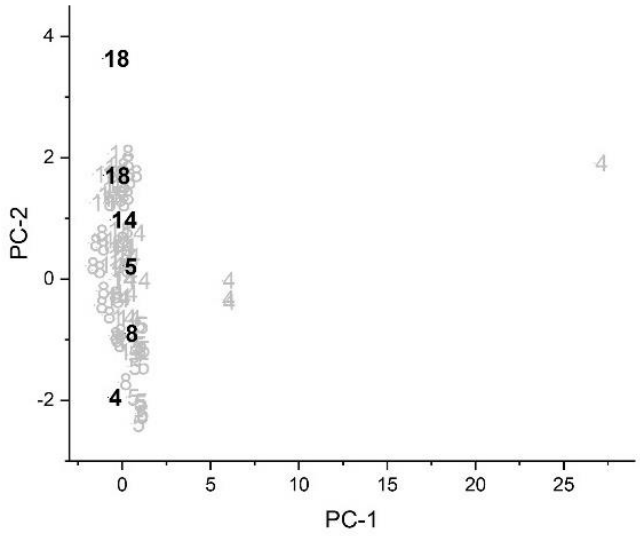

Figure 5.54 PC plot of the second training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

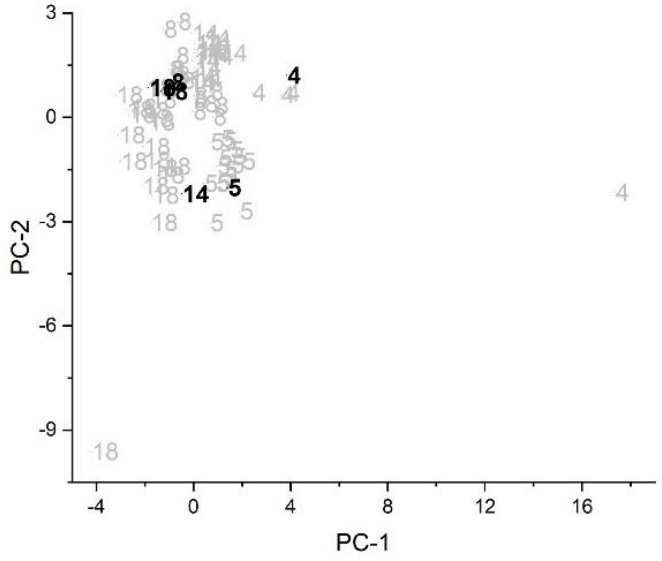

Figure 5.55 PC plot of the third training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

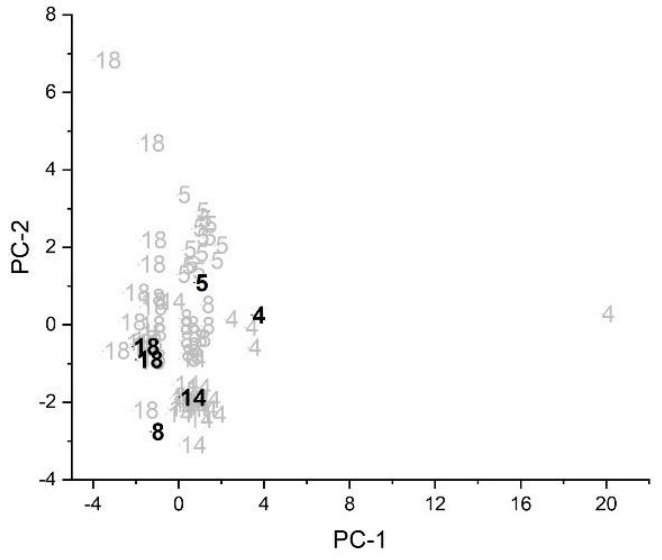

Figure 5.56 PC plot of the fourth training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

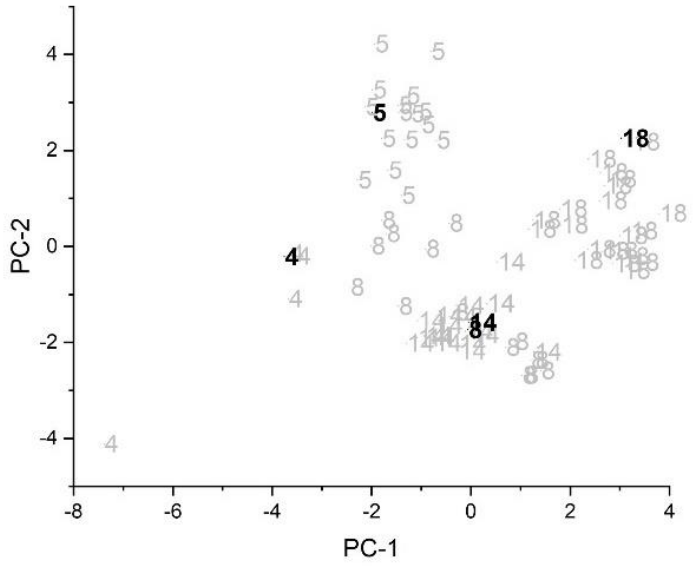

Figure 5.57 PC plot of the fifth training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

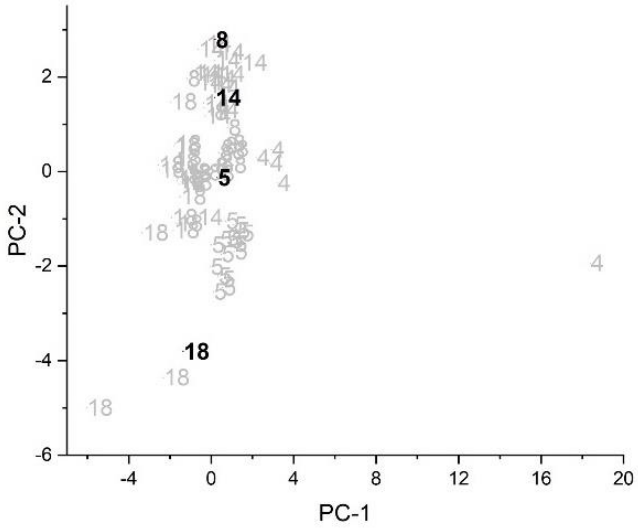

Figure 5.58 PC plot of the sixth training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

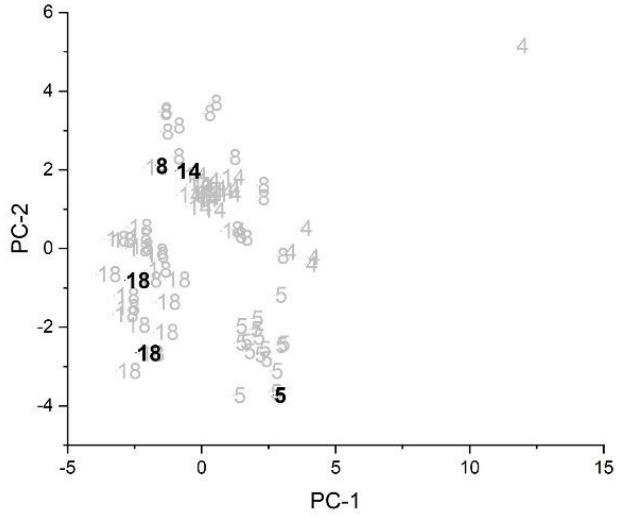

Figure 5.59 PC plot of the seventh training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

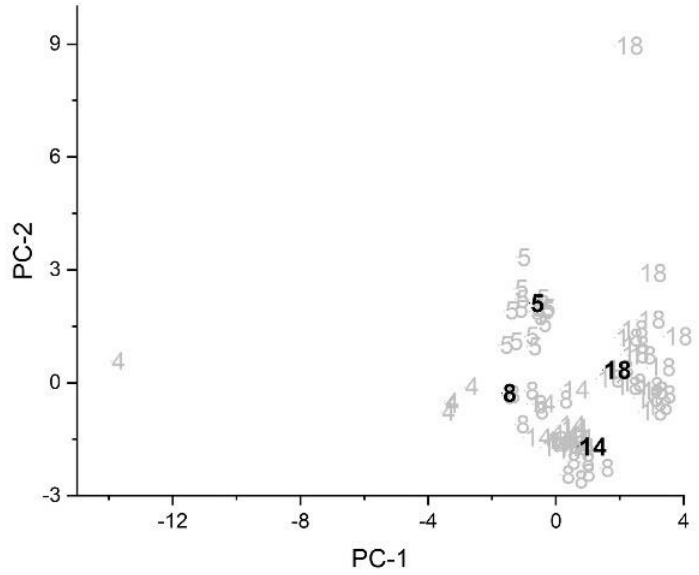

Figure 5.60 PC plot of the eighth training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

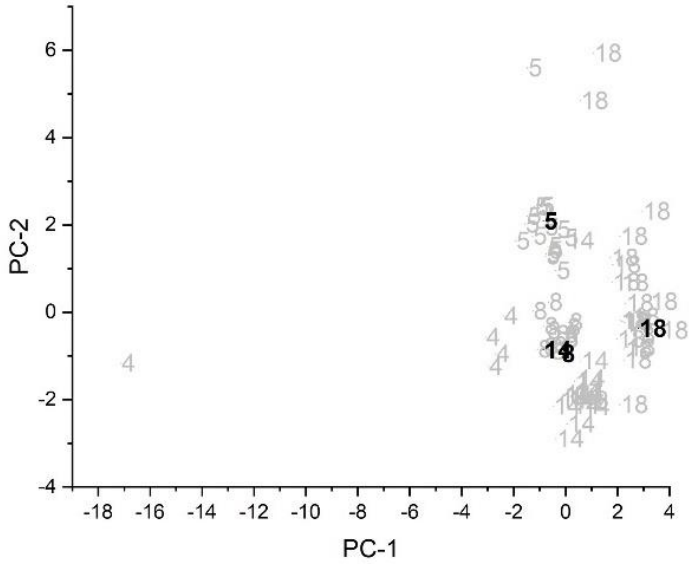

Figure 5.61 PC plot of the ninth training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

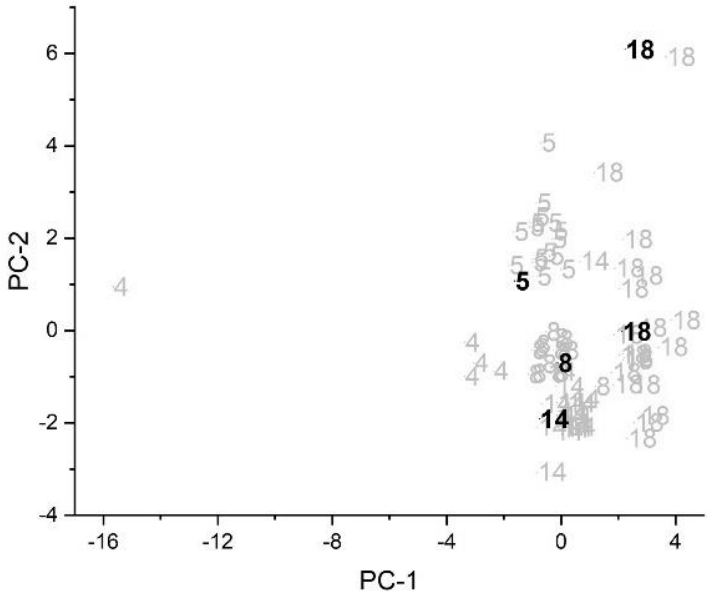

Figure 5.62 PC plot of the tenth training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

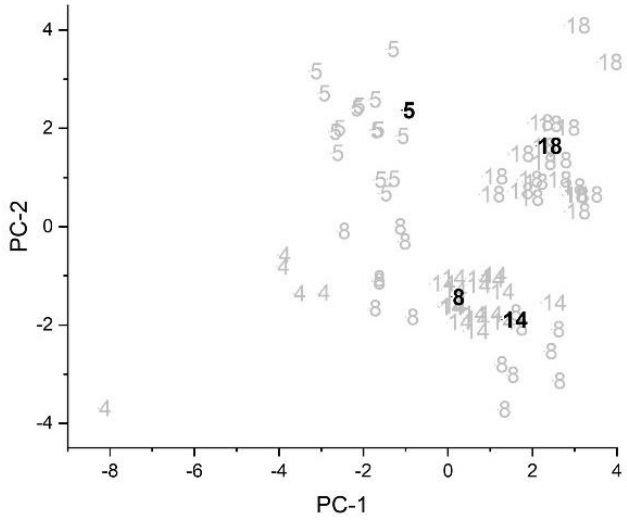

Figure 5.63 PC plot of the eleventh training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

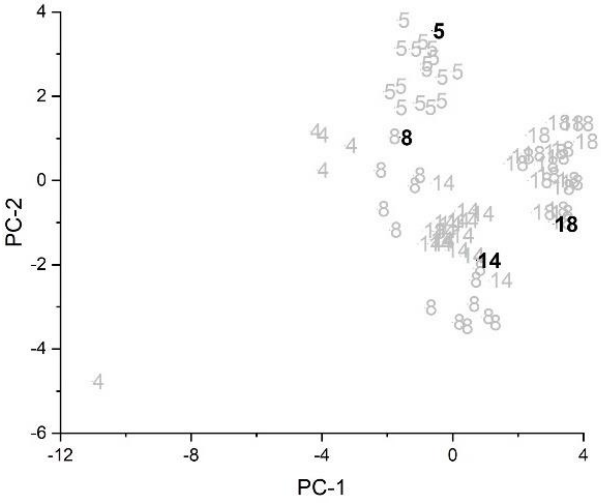

Figure 5.64 PC plot of the twelfth training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

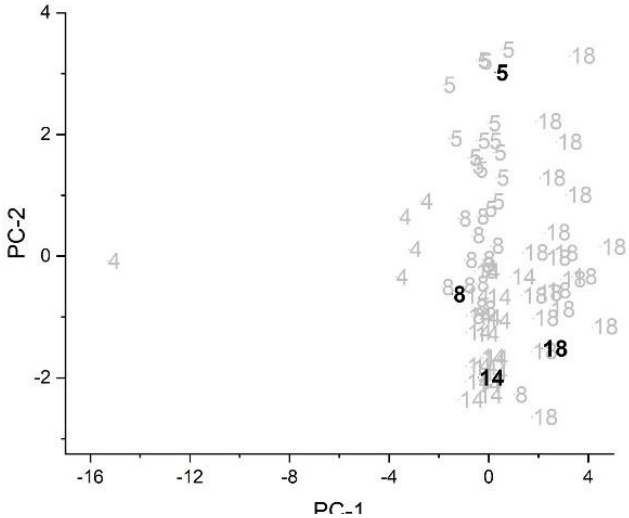

Figure 5.65 PC plot of the thirteenth training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

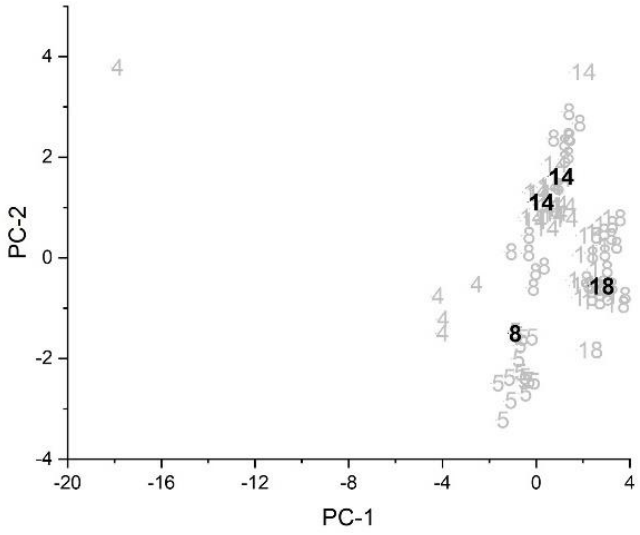

Figure 5.66 PC plot of the fourteenth training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

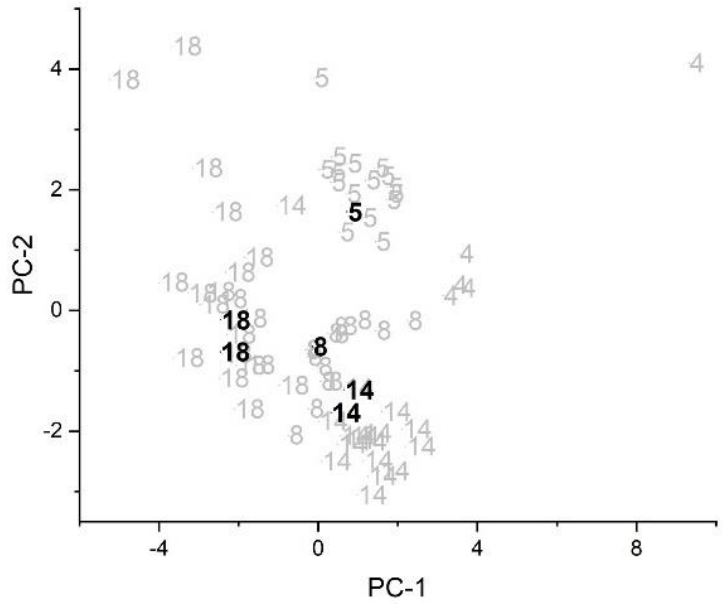

Figure 5.67 PC plot of the fifteenth training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

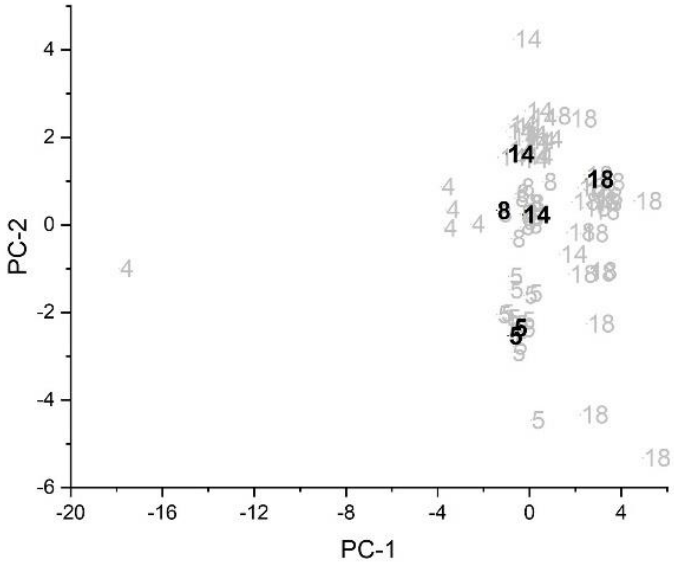

Figure 5.68 PC plot of the sixteenth training set/prediction set pair: Five-way classification study (Doraville, Fairfax, Fort Wayne, Lansing, and Moraine) for prediction set. Prediction set samples (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

Second Study Raman (Binary Classification)

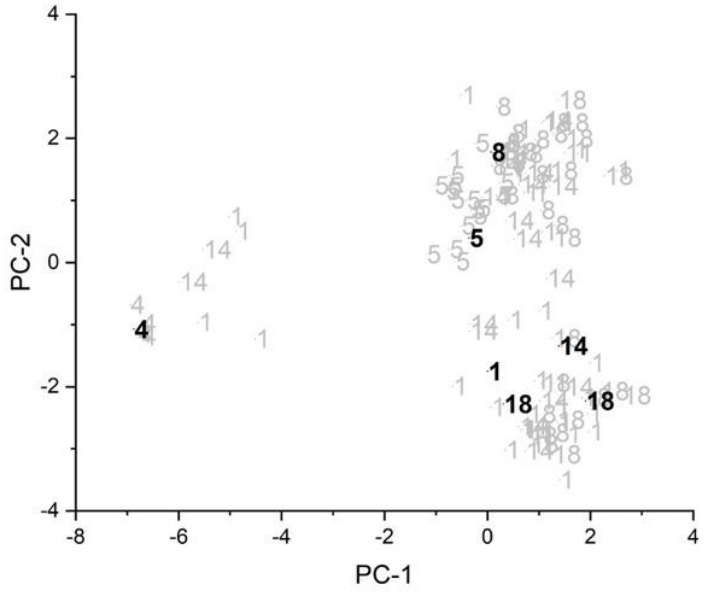

Figure 5.69 PC plot for the first training set/prediction set pair: Doraville vs Arlington, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

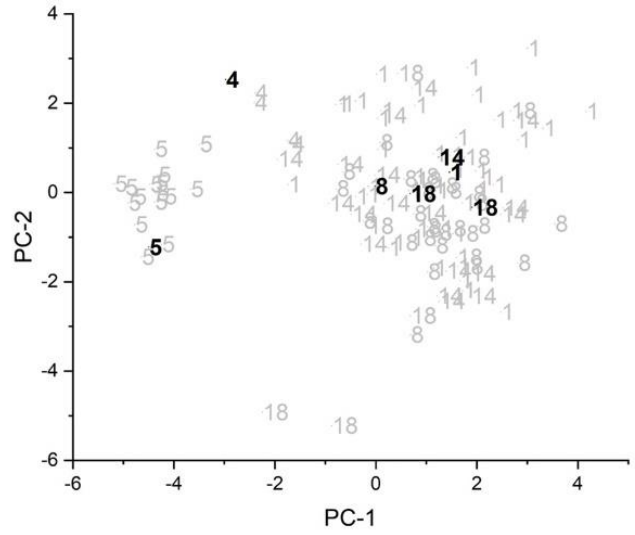

Figure 5.70 PC plot for the first training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

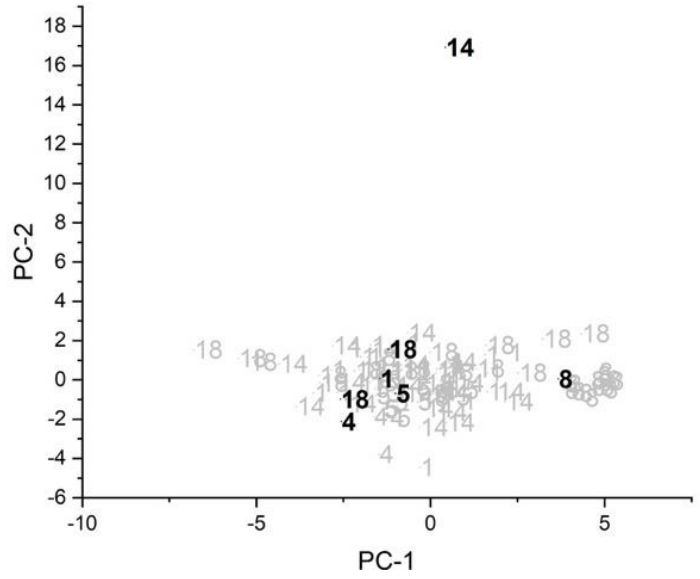

Figure 5.71 PC plot for the first training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

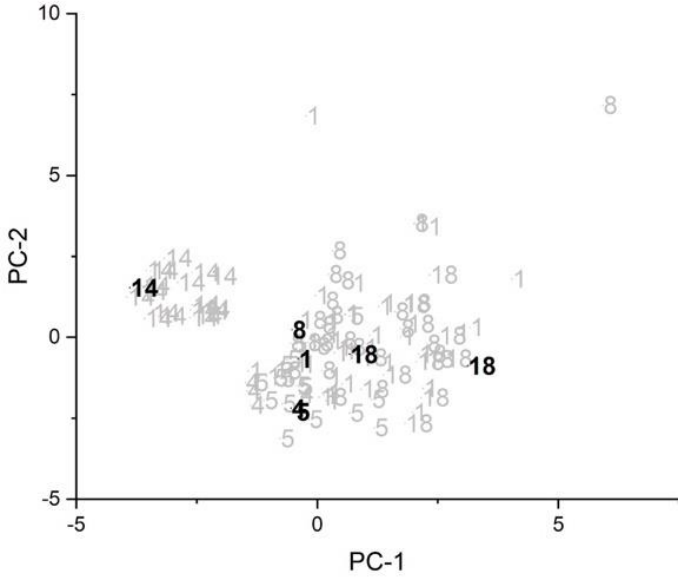

Figure 5.72 PC plot for the first training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

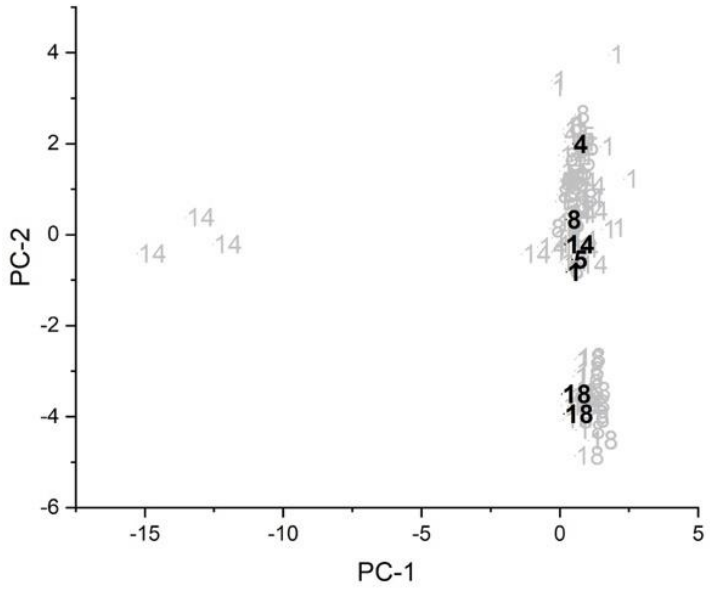

Figure 5.73 PC plot for the first training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

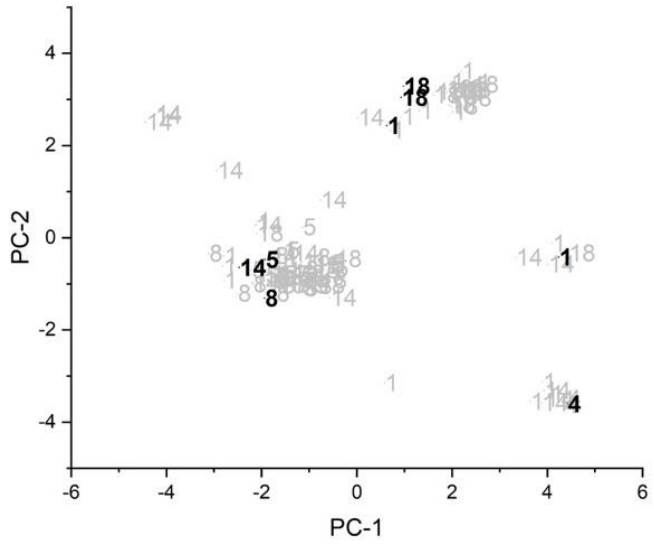

Figure 5.74 PC plot for the second training set/prediction set pair: Doraville vs Arlington, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

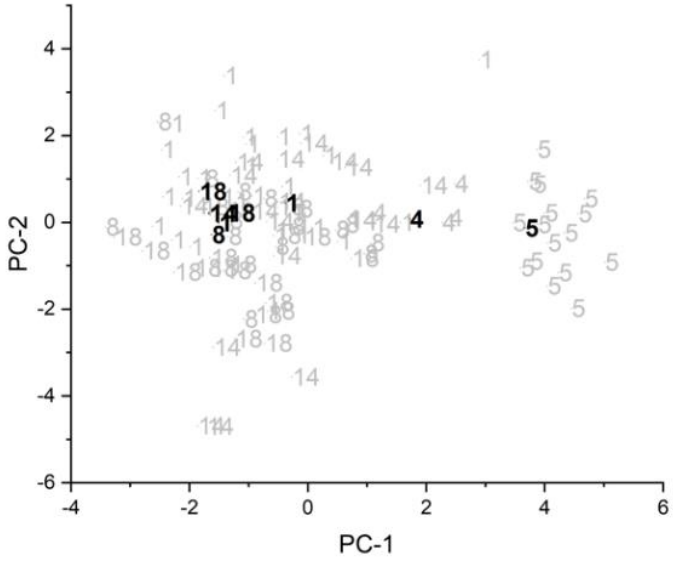

Figure 5.75 PC plot for the second training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

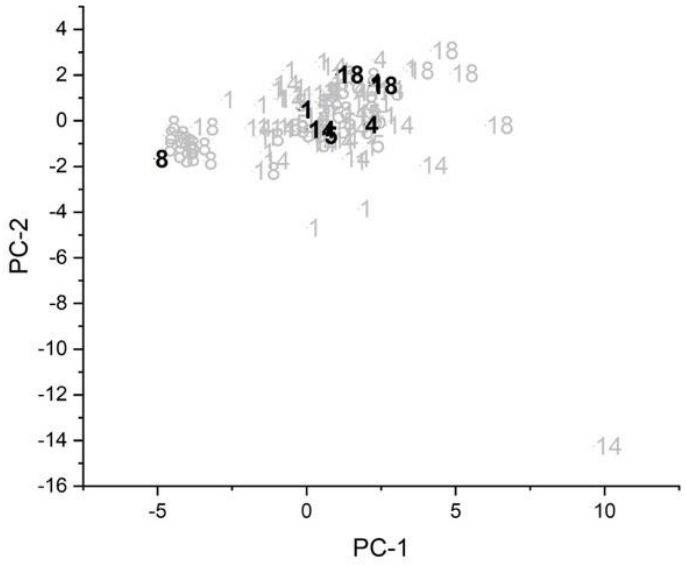

Figure 5.76 PC plot for the second training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

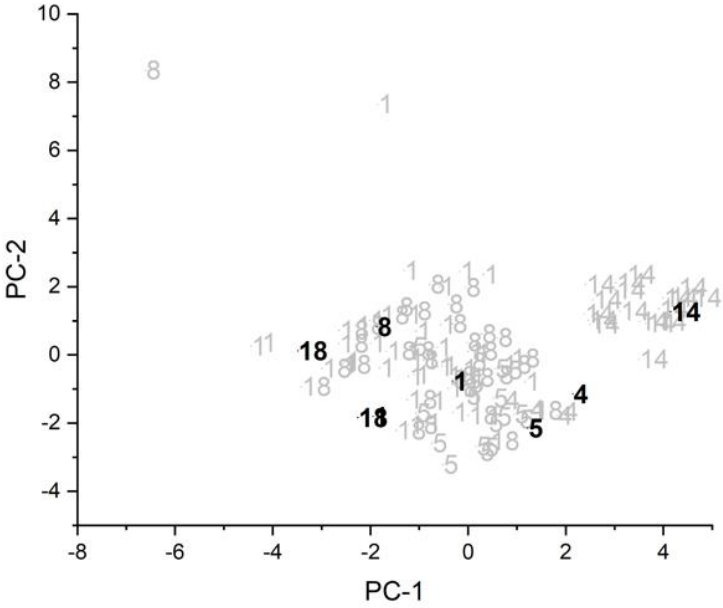

Figure 5.77 PC plot for the second training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

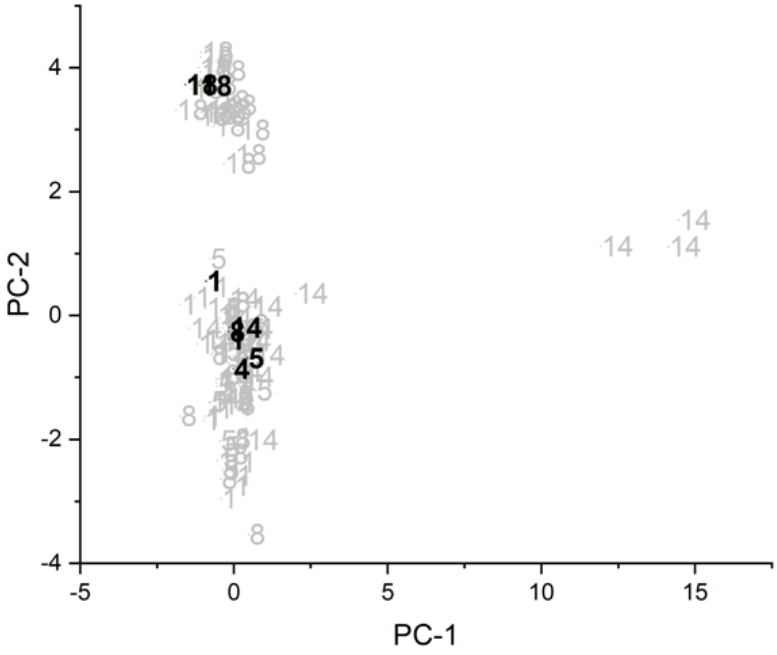

Figure 5.78 PC plot for the second training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

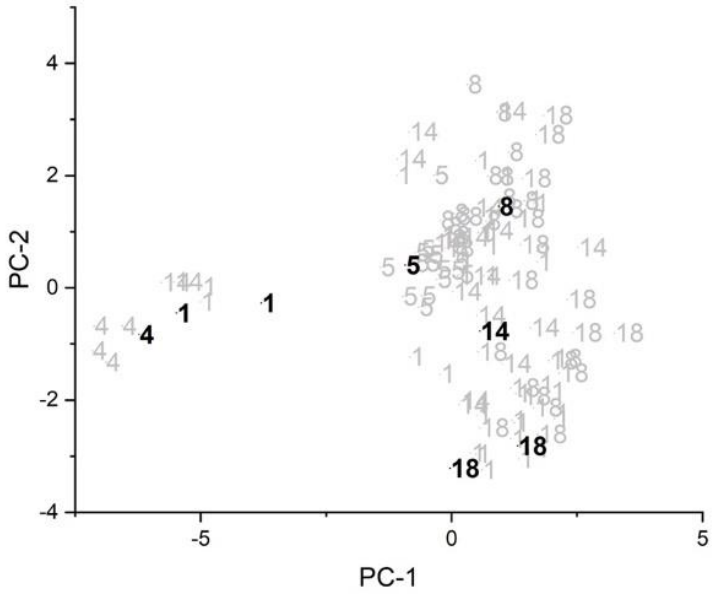

Figure 5.79 PC plot for the third training set/prediction set pair: Doraville vs Arlington, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

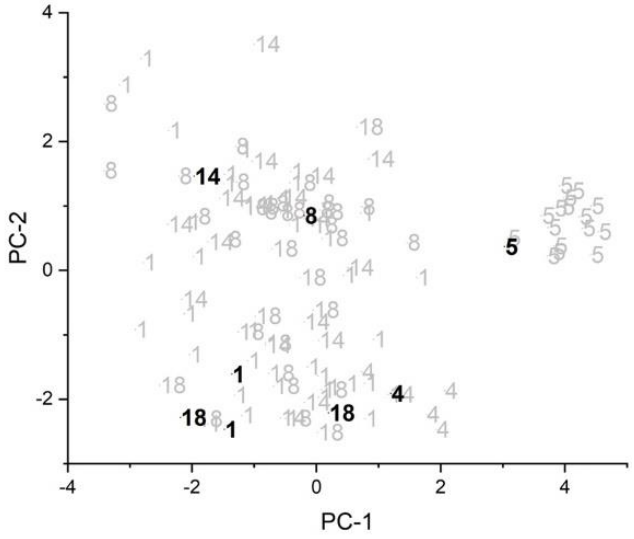

Figure 5.80 PC plot for the third training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

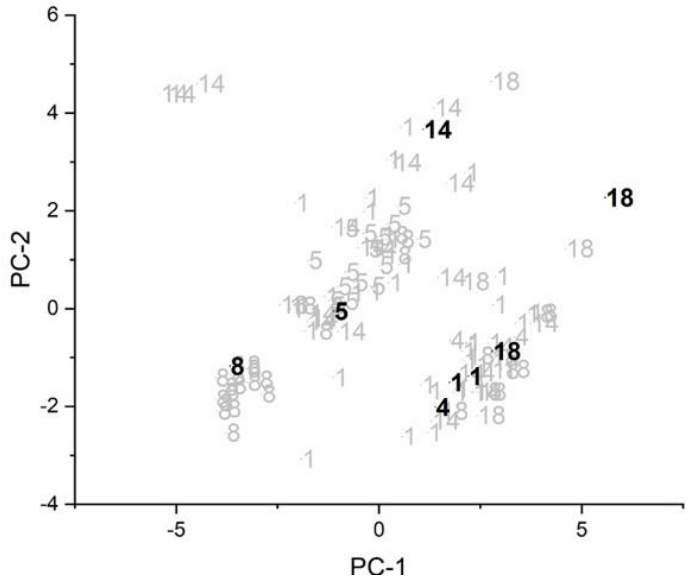

Figure 5.81 PC plot for the third training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

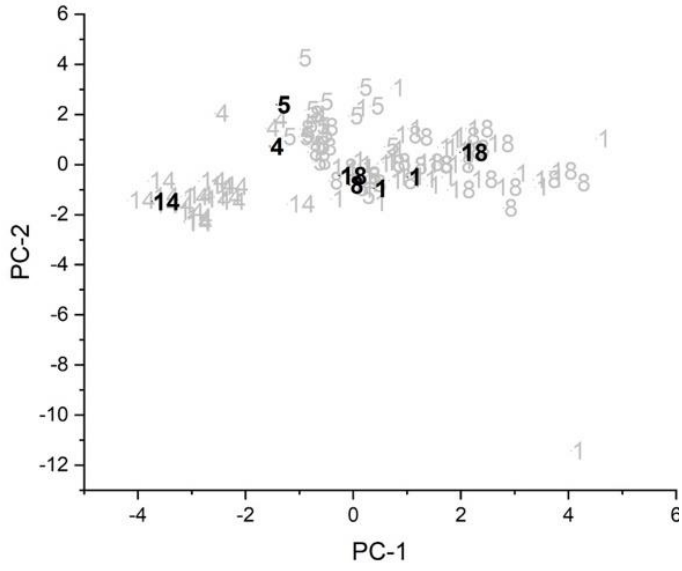

Figure 5.82 PC plot for the third training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

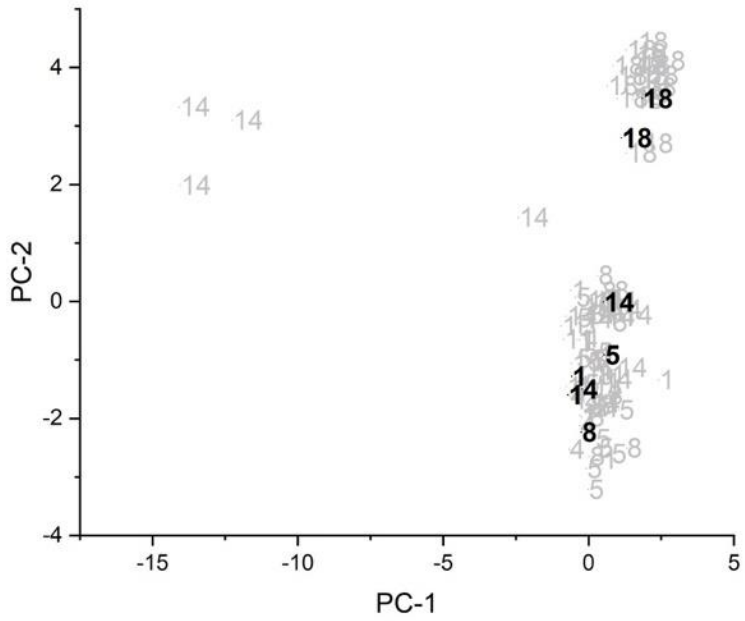

Figure 5.83 PC plot for the third training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

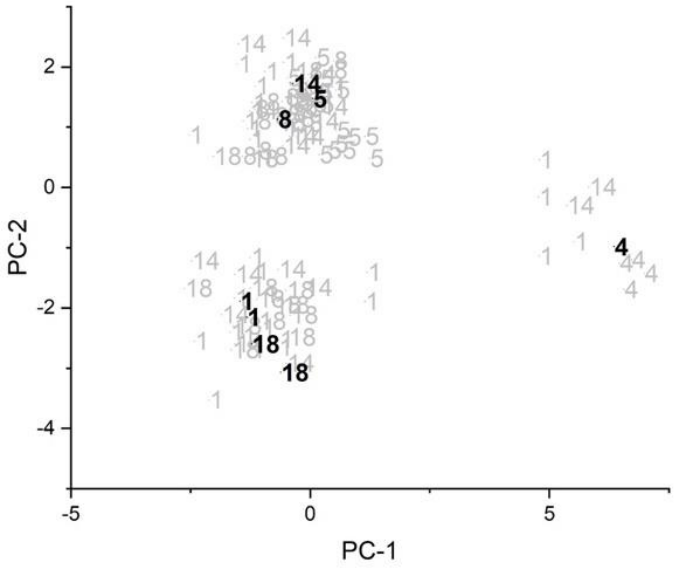

Figure 5.84 PC plot for the fourth training set/prediction set pair: Doraville vs Arlington, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

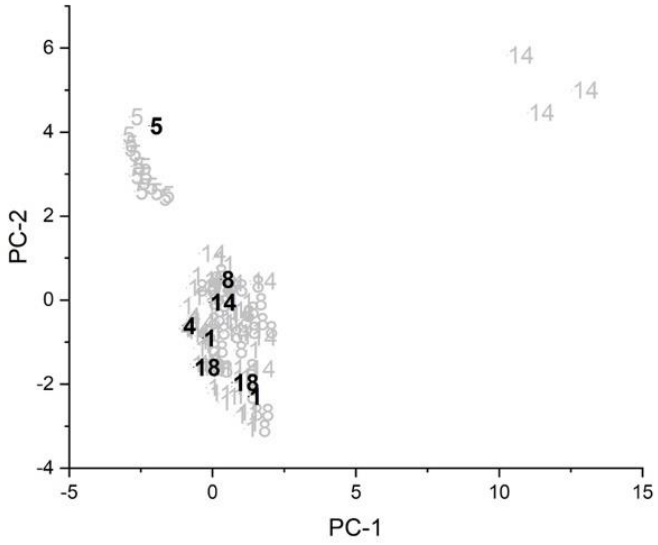

Figure 5.85 PC plot for the fourth training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

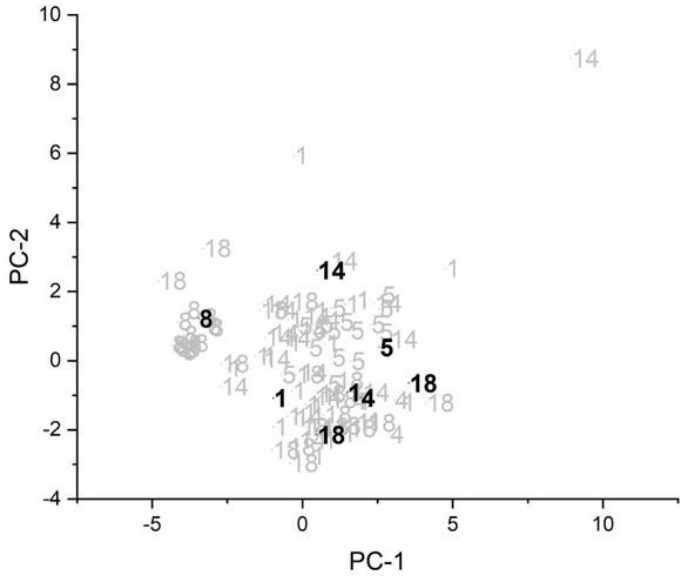

Figure 5.86 PC plot for the fourth training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

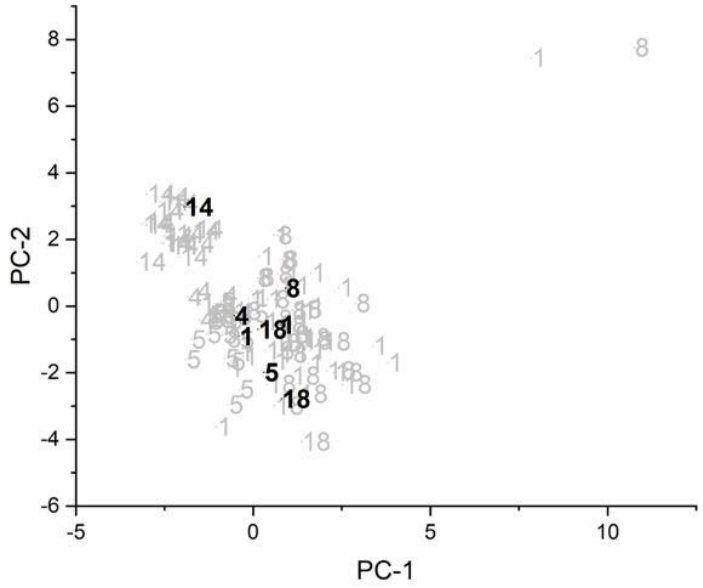

Figure 5.87 PC plot for the fourth training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

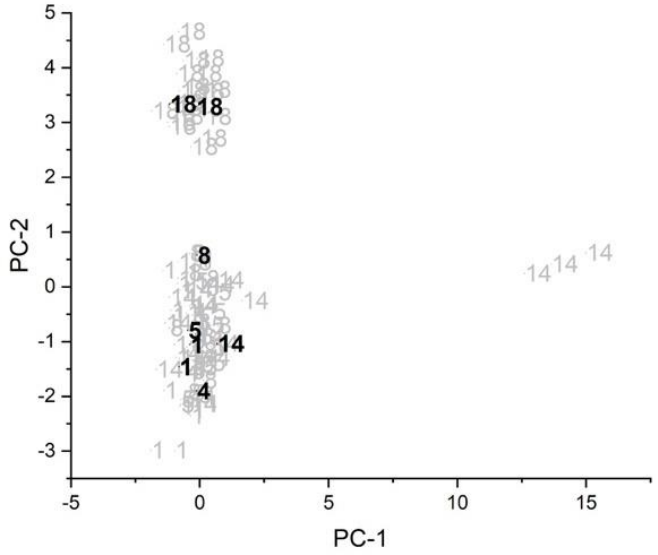

Figure 5.88 PC plot for the fourth training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

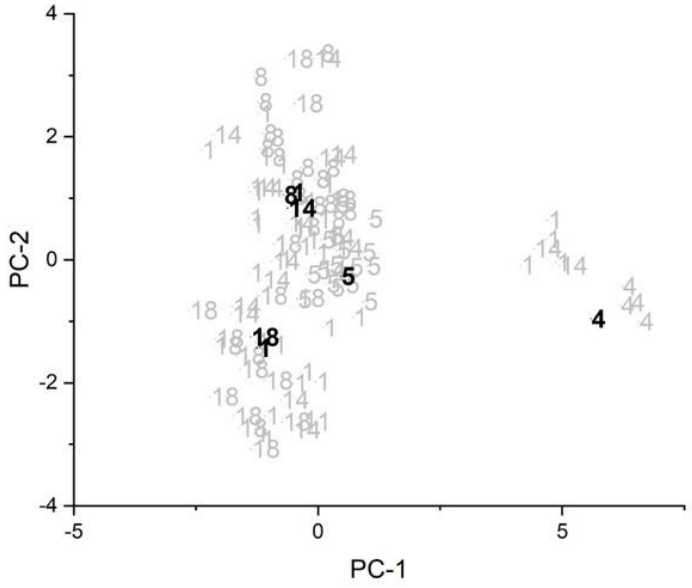

Figure 5.89 PC plot for the fifth training set/prediction set pair: Doraville vs Arlington, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

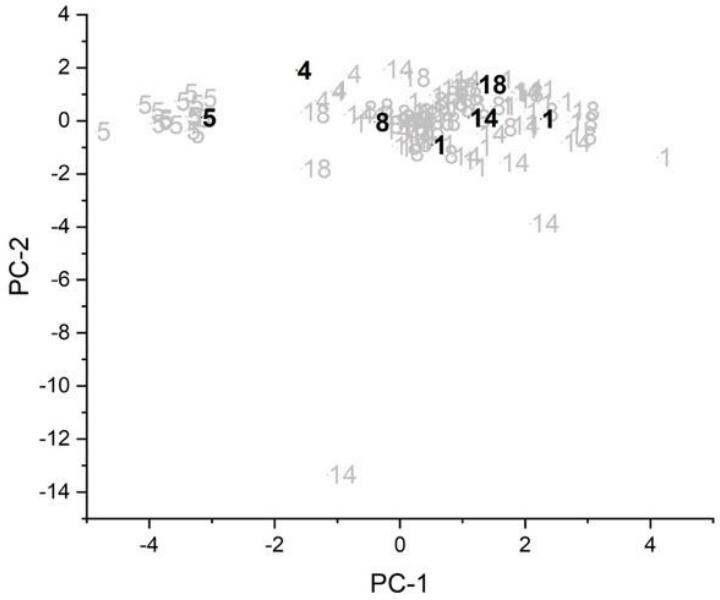

Figure 5.90 PC plot for the fifth training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

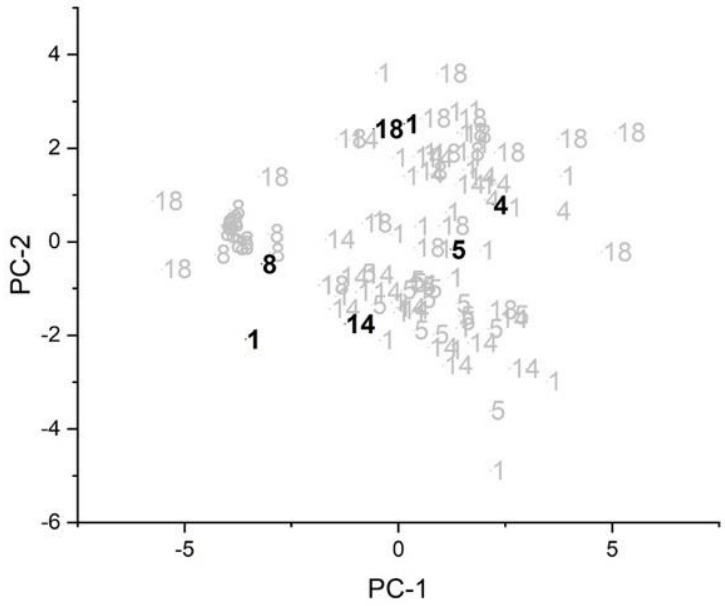

Figure 5.91 PC plot for the fifth training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

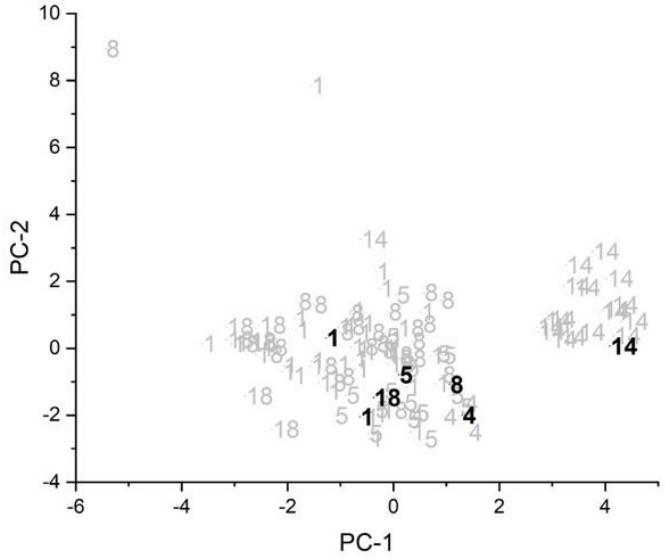

Figure 5.92 PC plot for the fifth training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

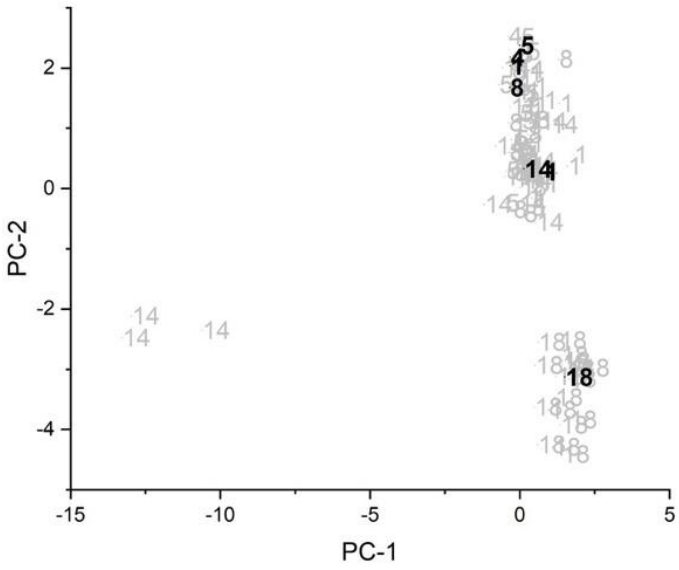

Figure 5.93 PC plot for the fifth training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

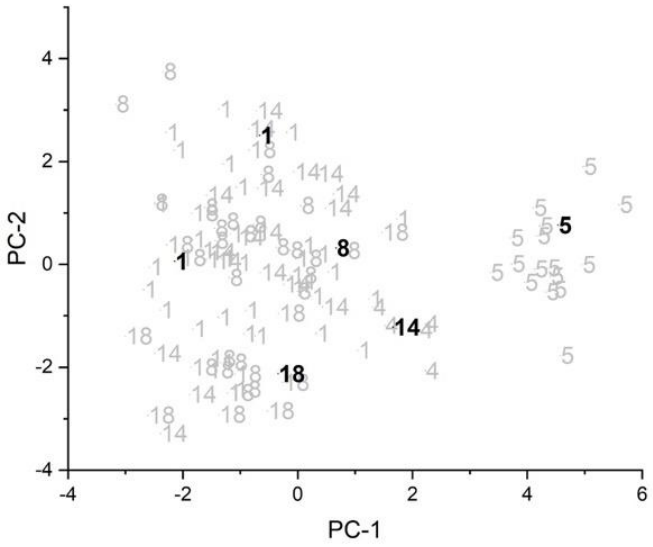

Figure 5.94 PC plot for the sixth training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

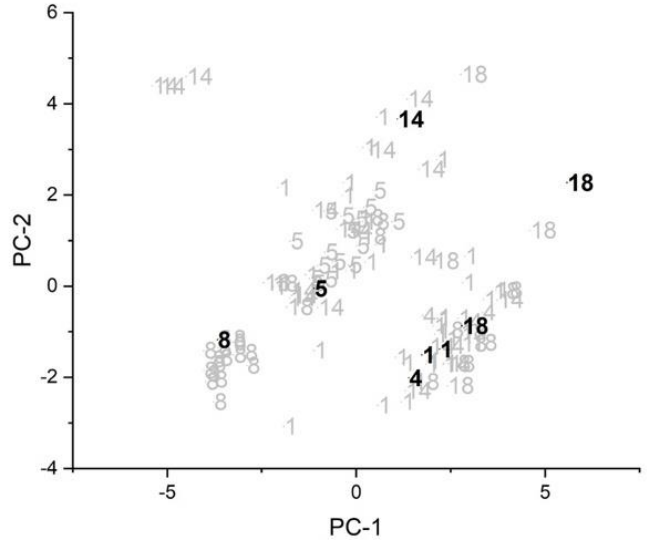

Figure 5.95 PC plot for the sixth training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

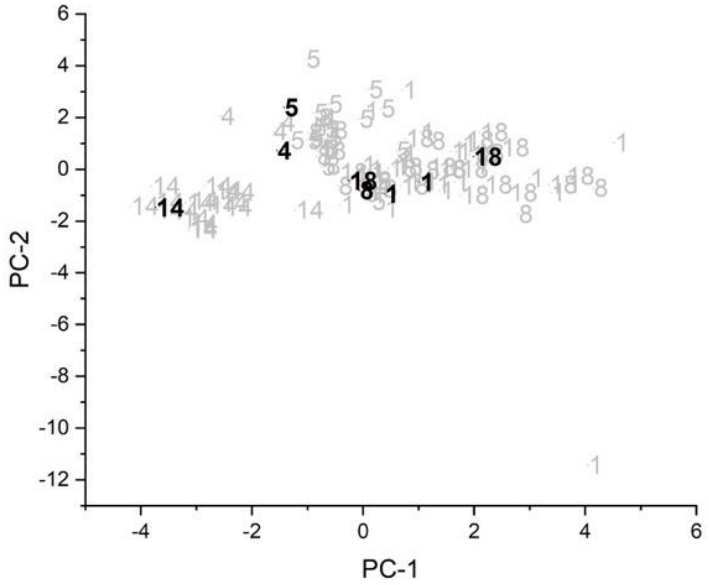

Figure 5.96 PC plot for the sixth training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

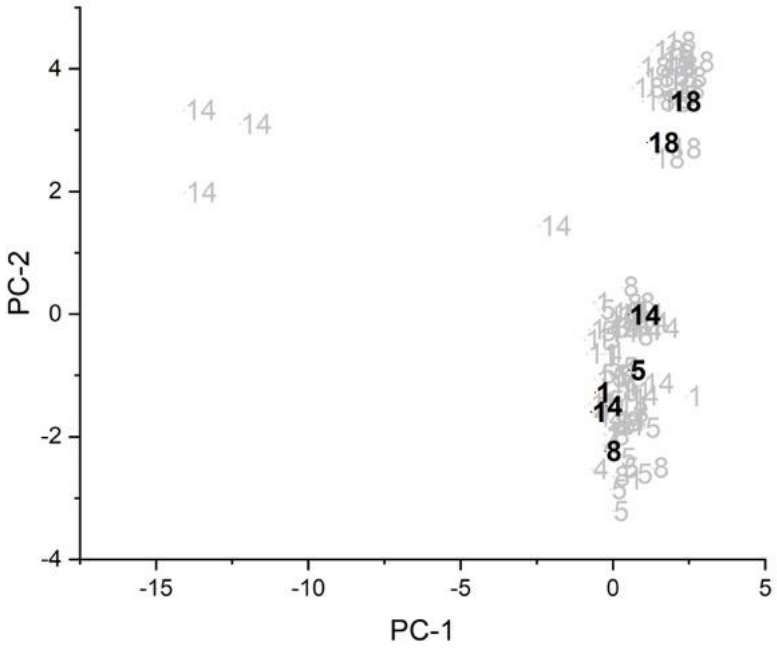

Figure 5.97 PC plot for the sixth training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

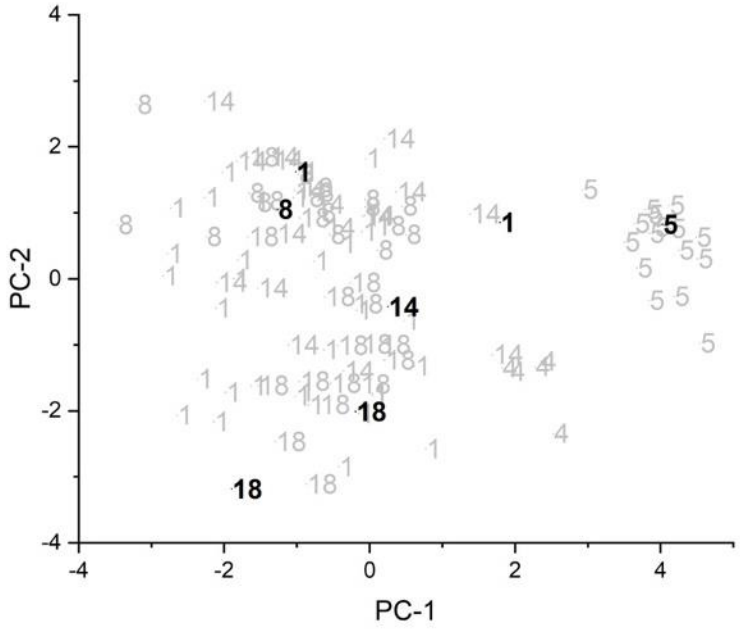

Figure 5.98 PC plot for the seventh training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

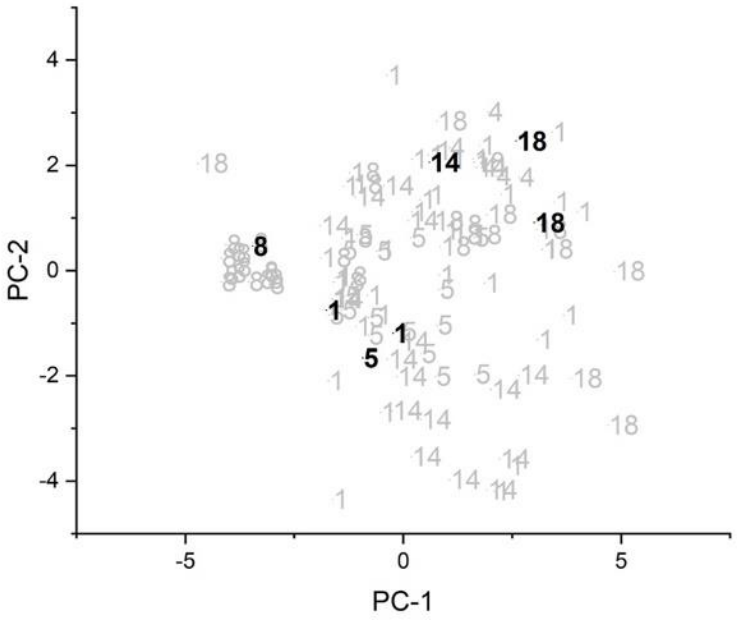

Figure 5.99 PC plot for the seventh training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

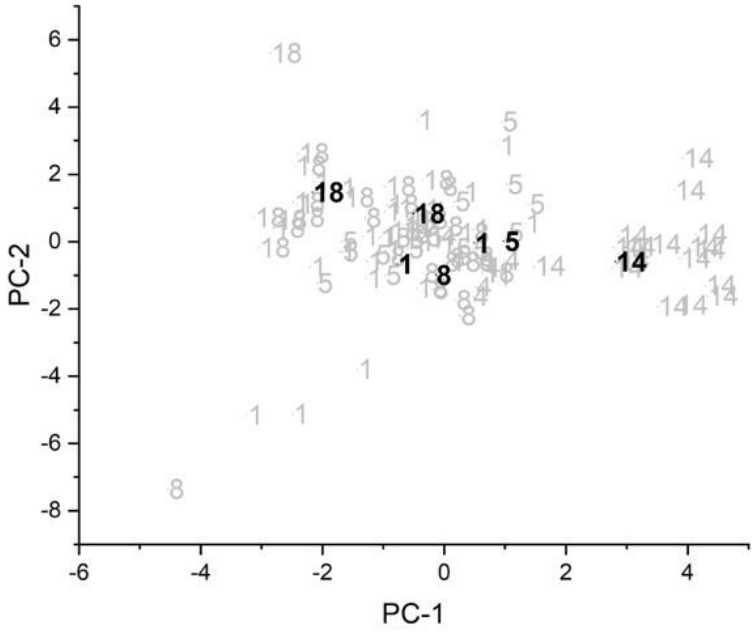

Figure 5.100 PC plot for the seventh training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

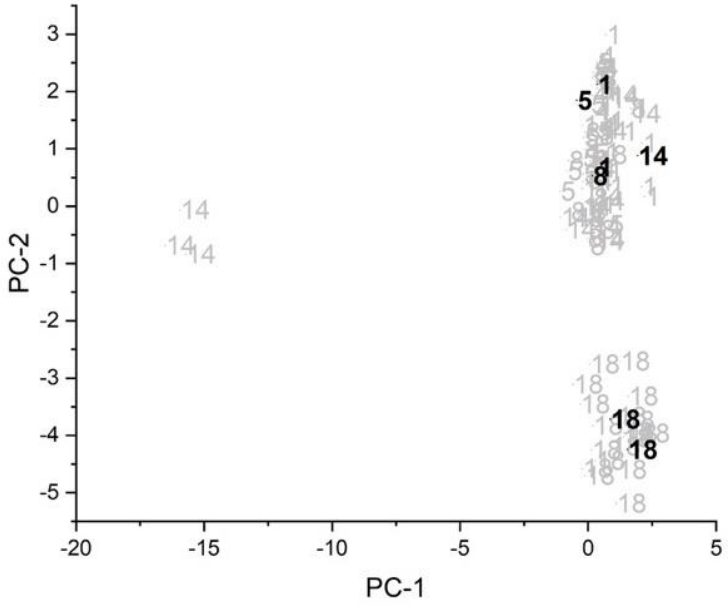

Figure 5.101 PC plot for the seventh training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

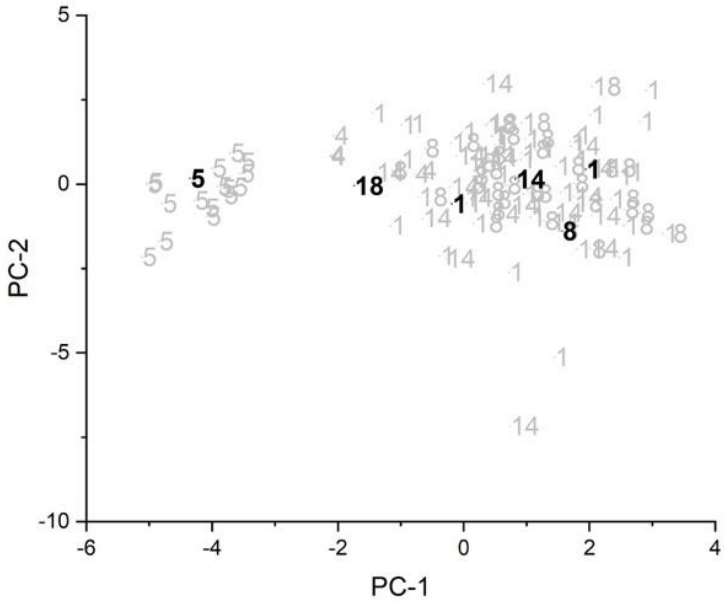

Figure 5.102 PC plot for the eighth training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

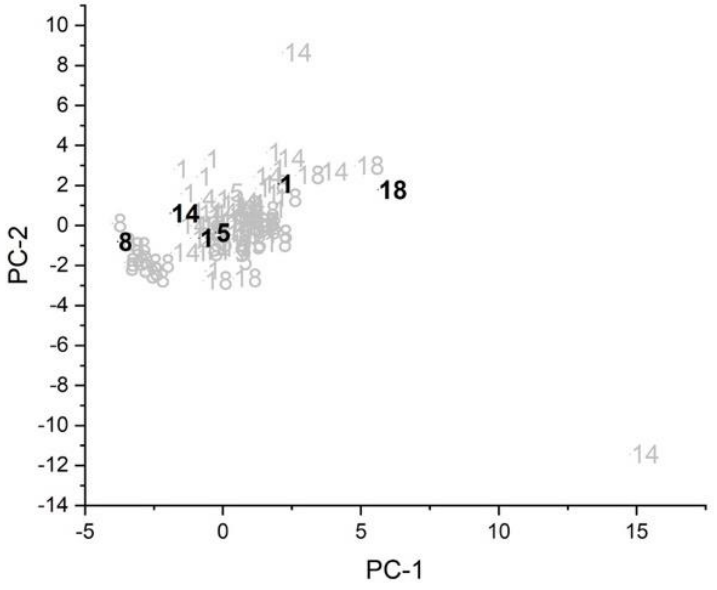

Figure 5.103 PC plot for the eighth training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

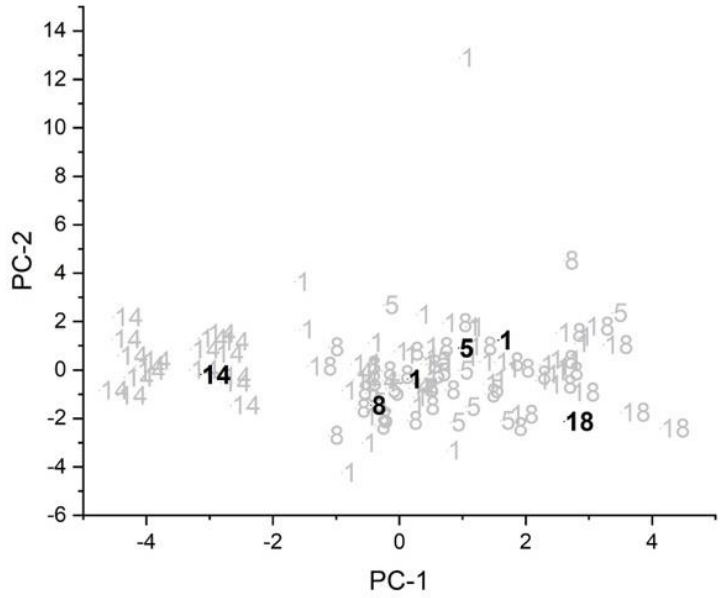

Figure 5.104 PC plot for the eighth training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

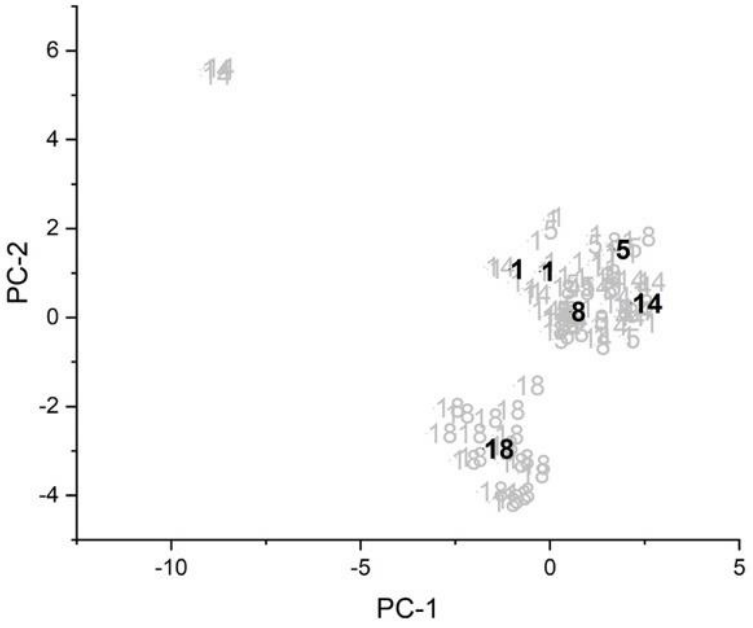

Figure 5.105 PC plot for the eighth training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

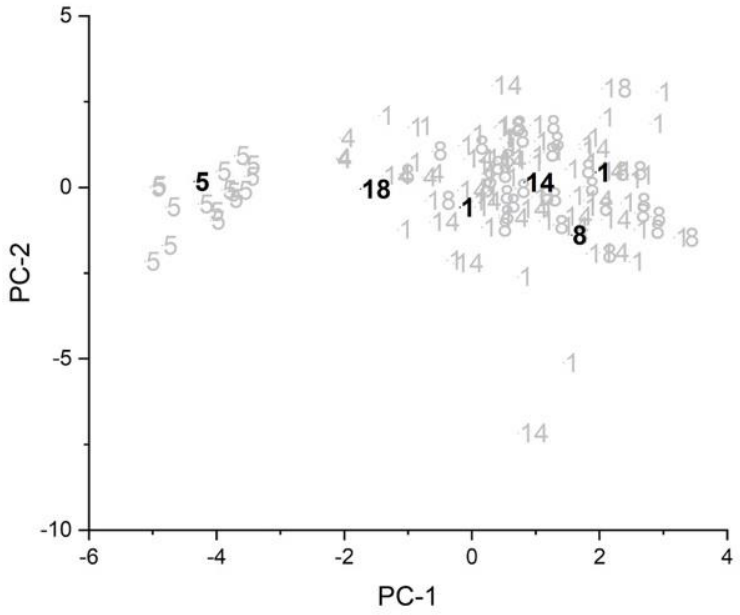

Figure 5.106 PC plot for the ninth training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

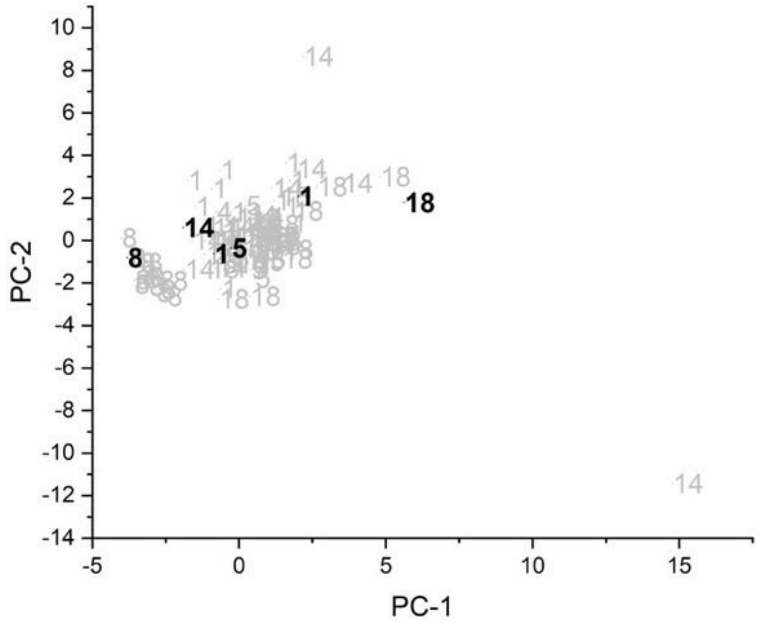

Figure 5.107 PC plot for the ninth training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

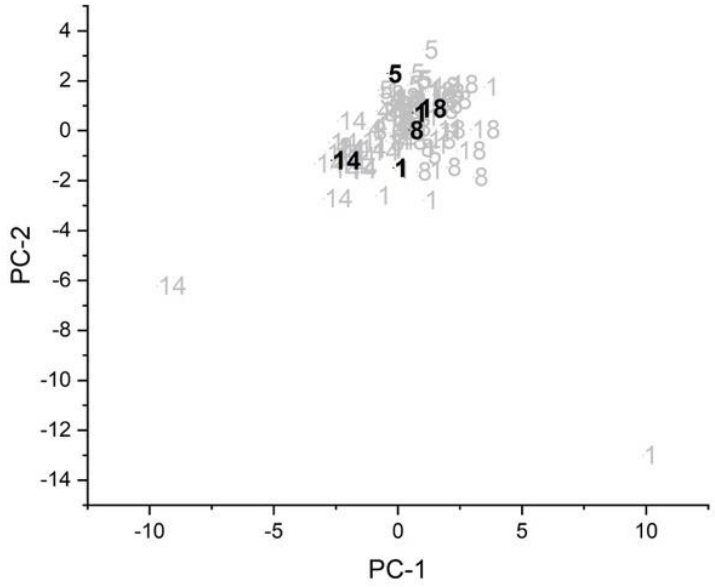

Figure 5.108 PC plot for the ninth training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

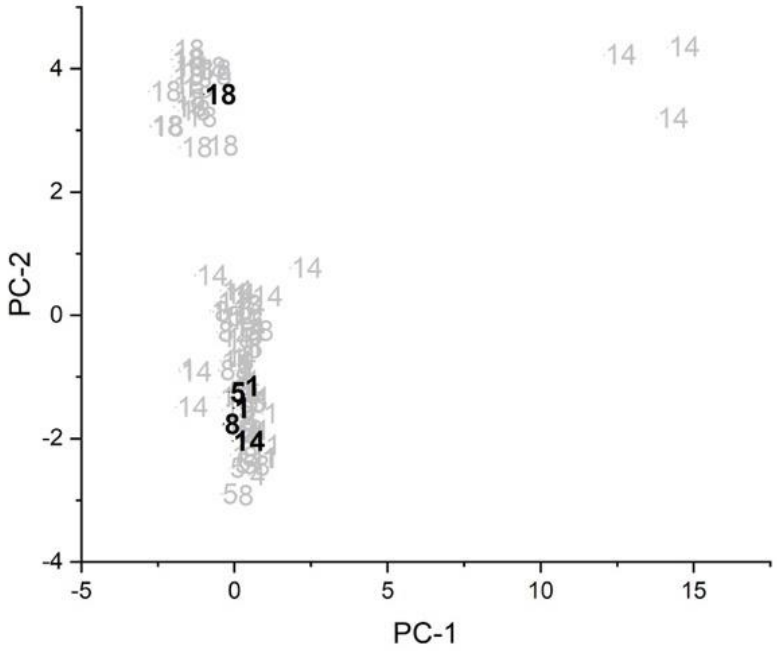

Figure 5.109 PC plot for the ninth training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

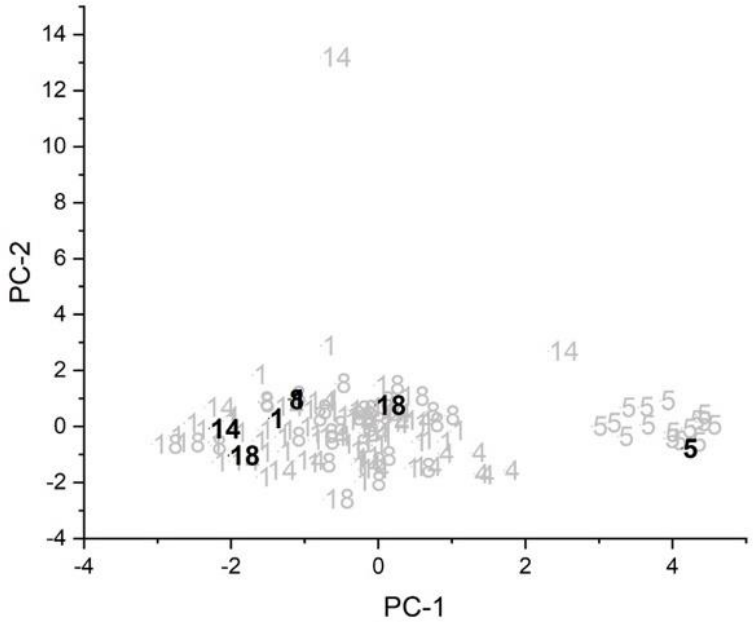

Figure 5.110 PC plot for the tenth training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

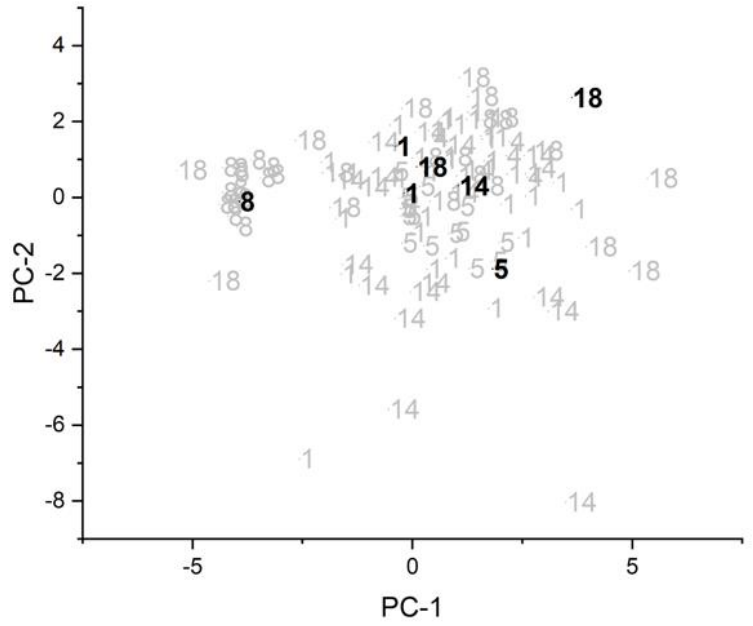

Figure 5.111 PC plot for the tenth training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

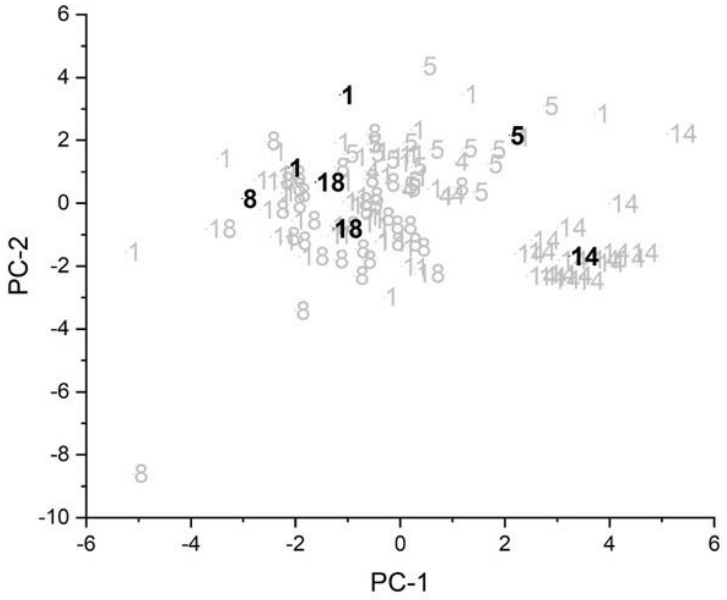

Figure 5.112 PC plot for the tenth training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

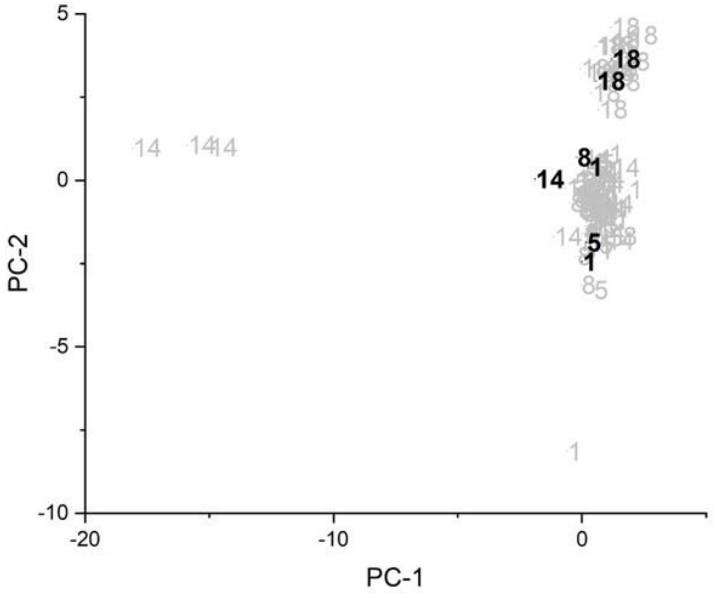

Figure 5.113 PC plots for the tenth training set/prediction set pair: Moraine vs Arlington, Fairfax, Fort Wayne, and Lansing for prediction set 10 Raman. Prediction set projected onto the PC plot developed from the training set and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Ft. Wayne, 14 = Lansing and 18 = Moraine.

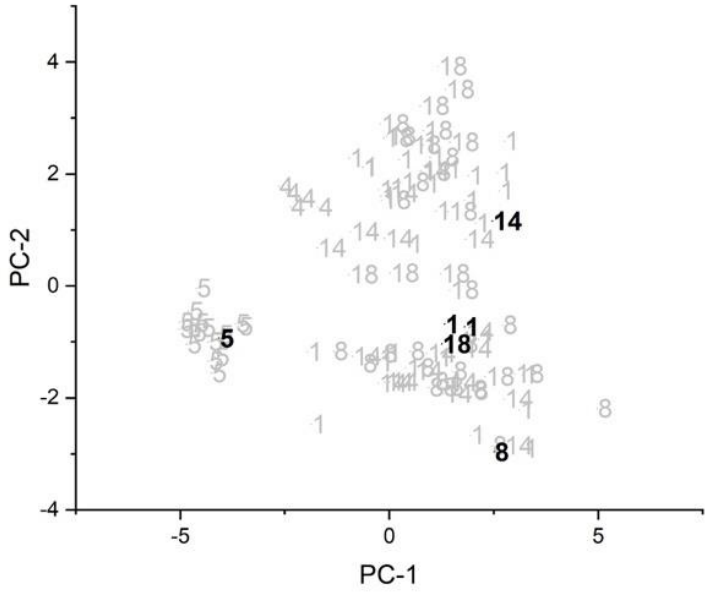

Figure 5.114 PC plot for the eleventh training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

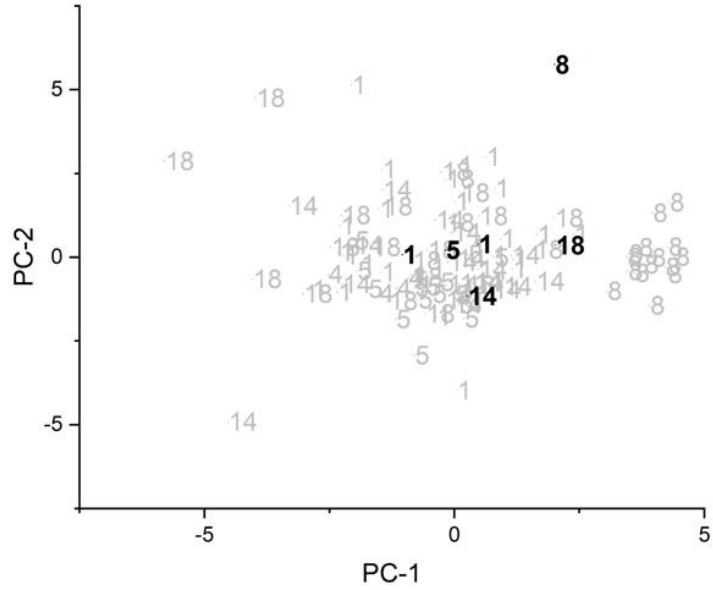

Figure 5.115 PC plot for the eleventh training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

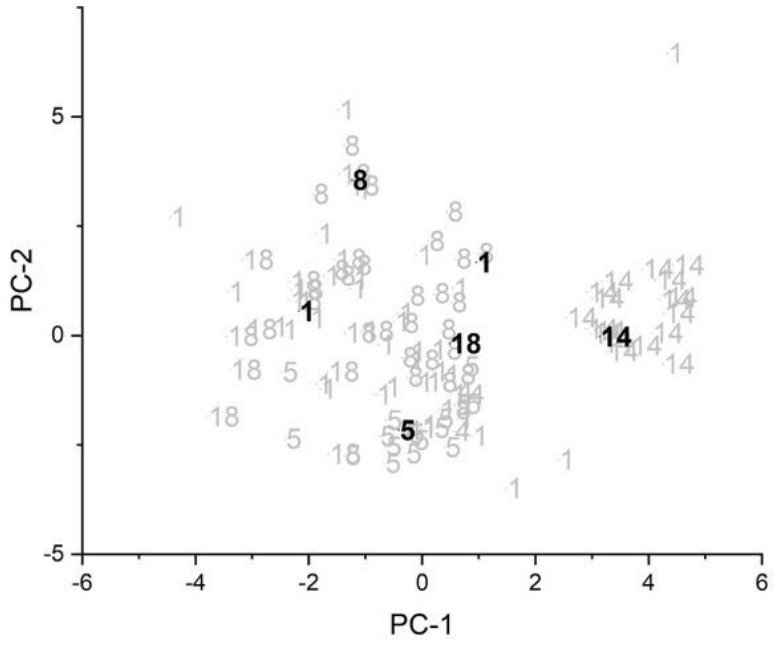

Figure 5.116 PC plot for the eleventh training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

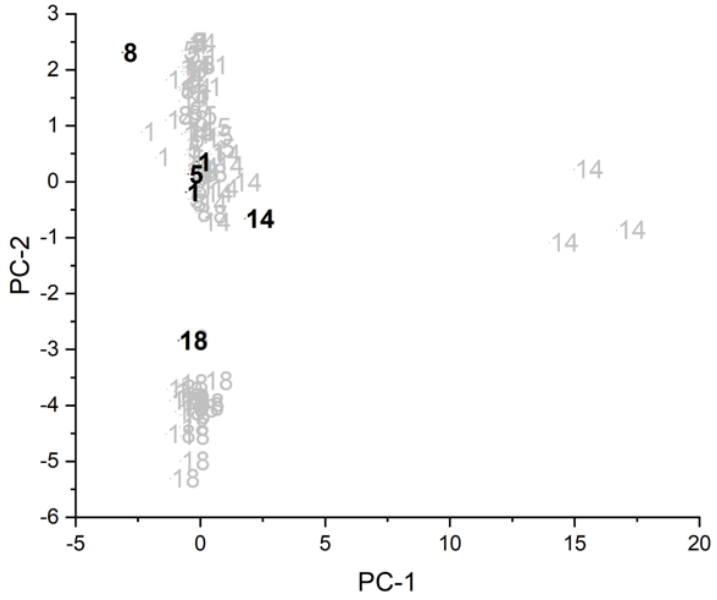

Figure 5.117 PC plot for the eleventh training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

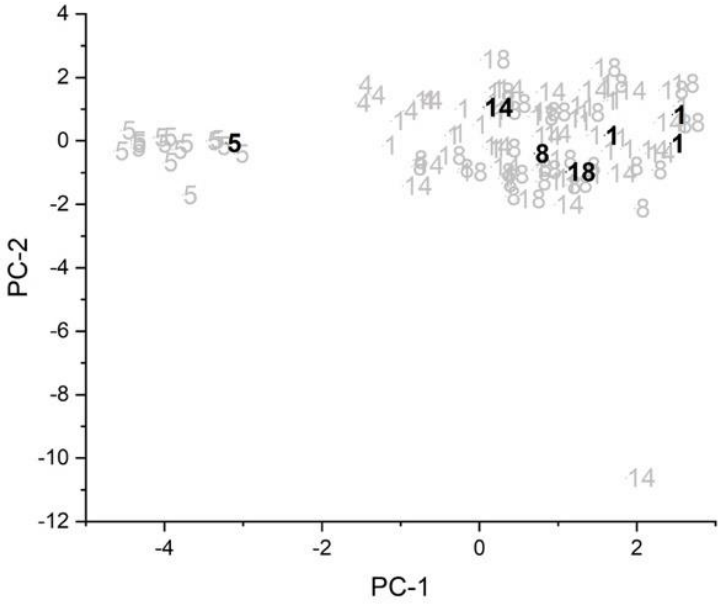

Figure 5.118 PC plot for the twelfth training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

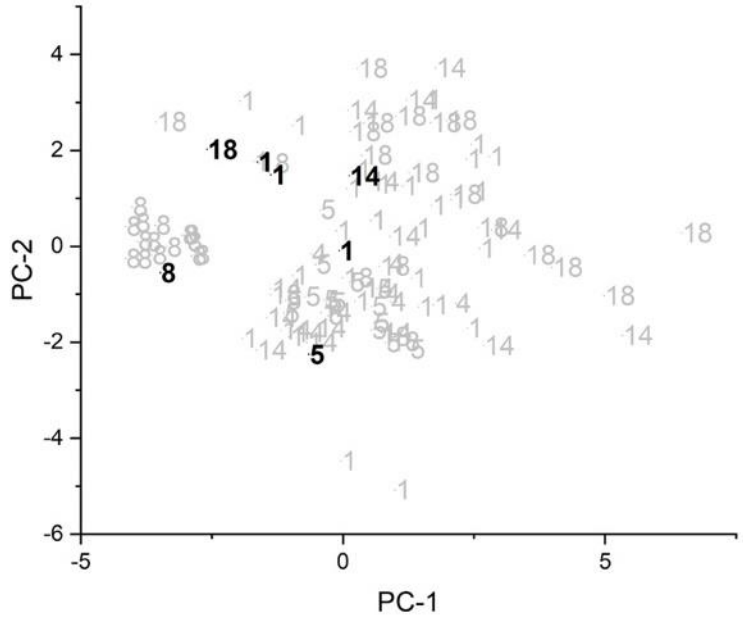

Figure 5.119 PC plot for the twelfth set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

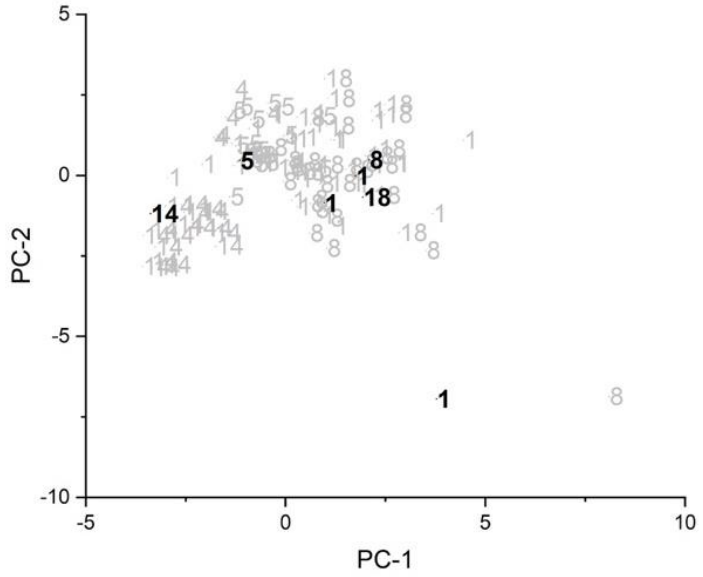

Figure 5.120 PC plot for the twelfth training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

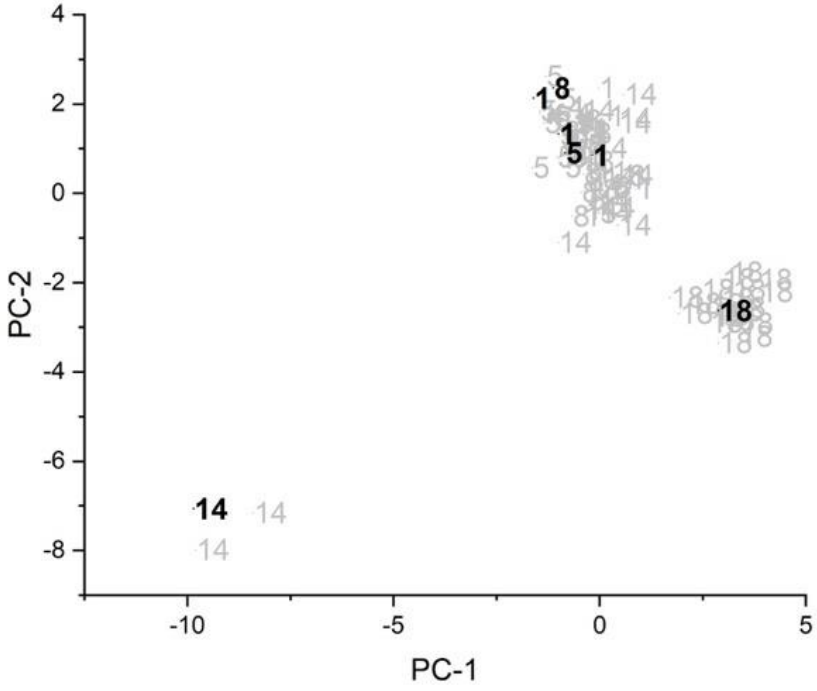

Figure 5.121 PC plot for the twelfth training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

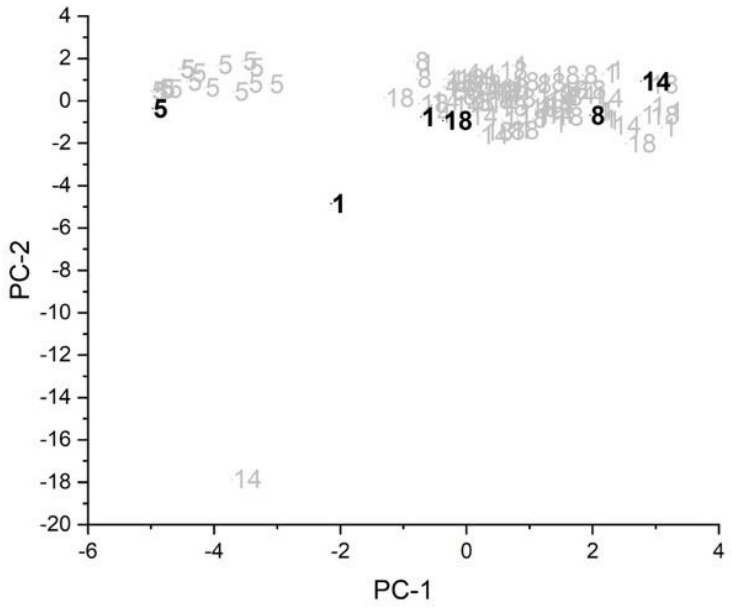

Figure 5.122 PC plot for the thirteenth training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

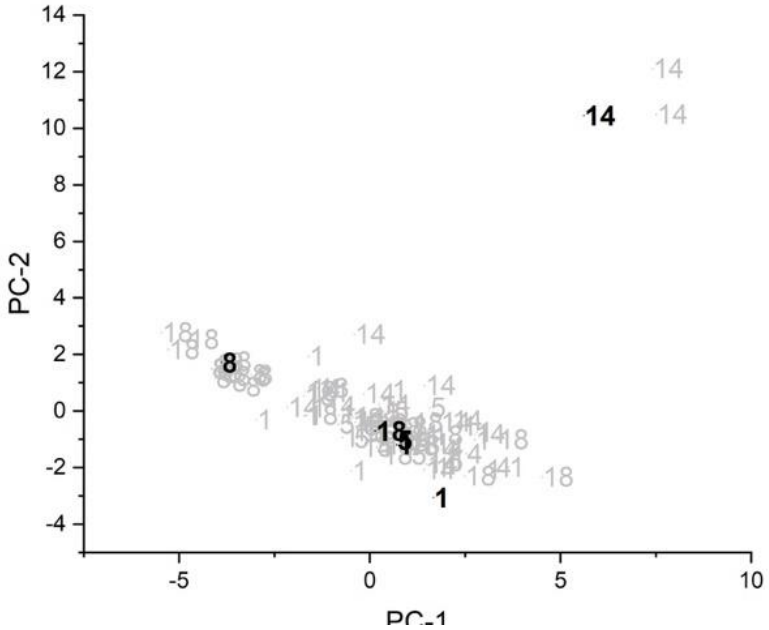

Figure 5.123 PC plot for the thirteenth training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

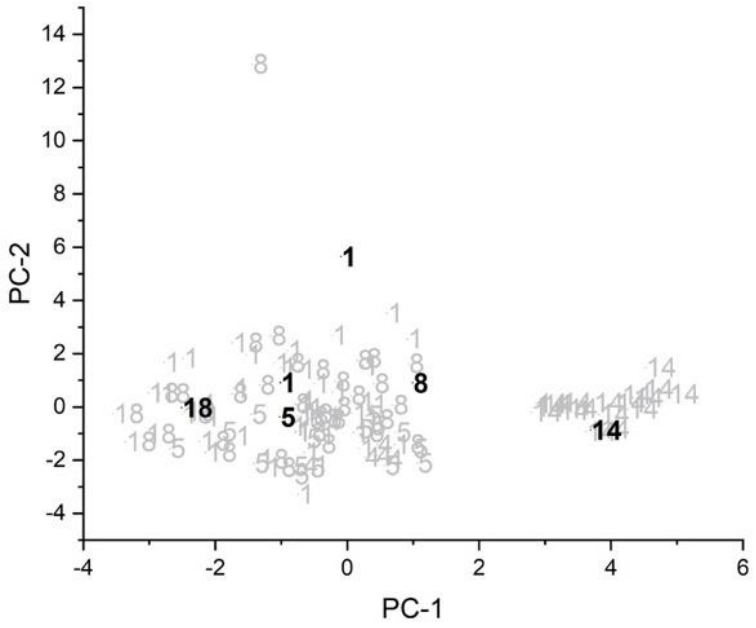

Figure 5.124 PC plot for the thirteenth training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

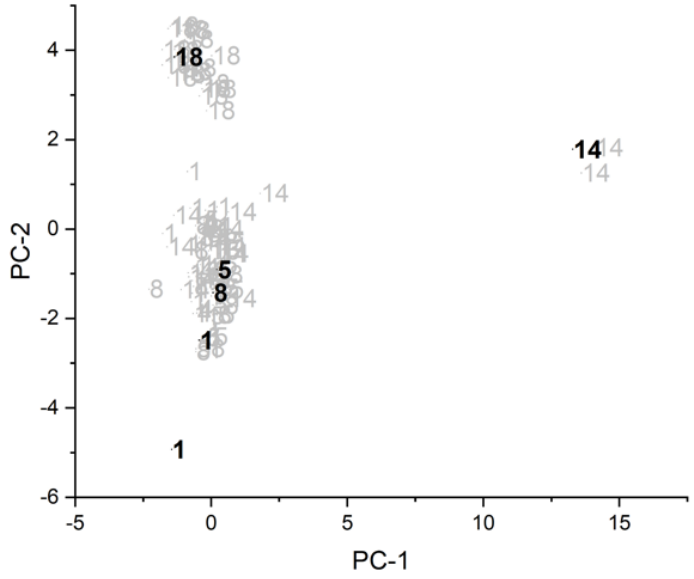

Figure 5.125 PC plot for the thirteenth training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

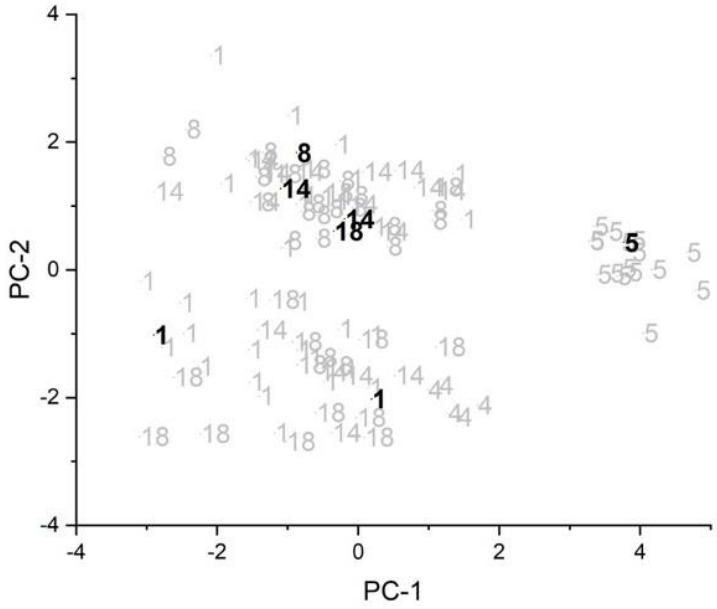

Figure 5.126 PC plot for the fourteenth training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

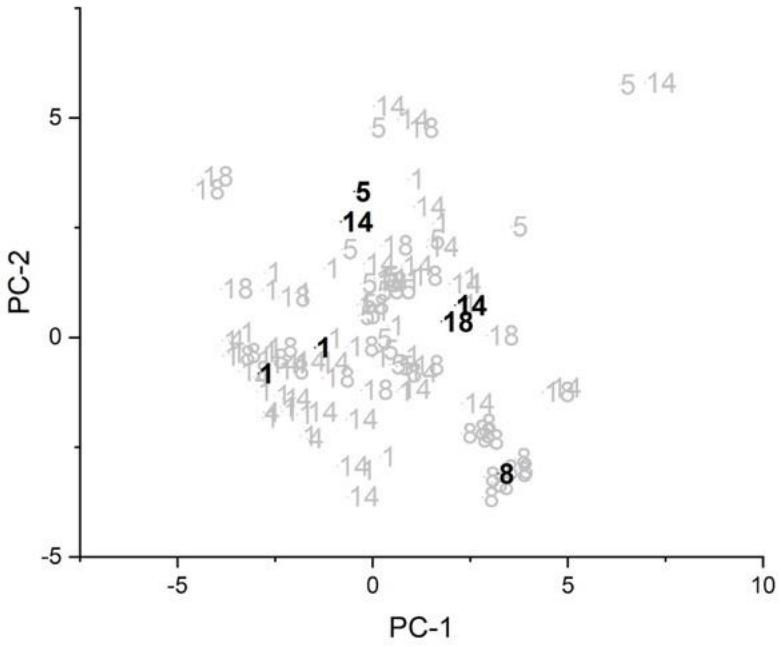

Figure 5.127 PC plot for the fourteenth training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

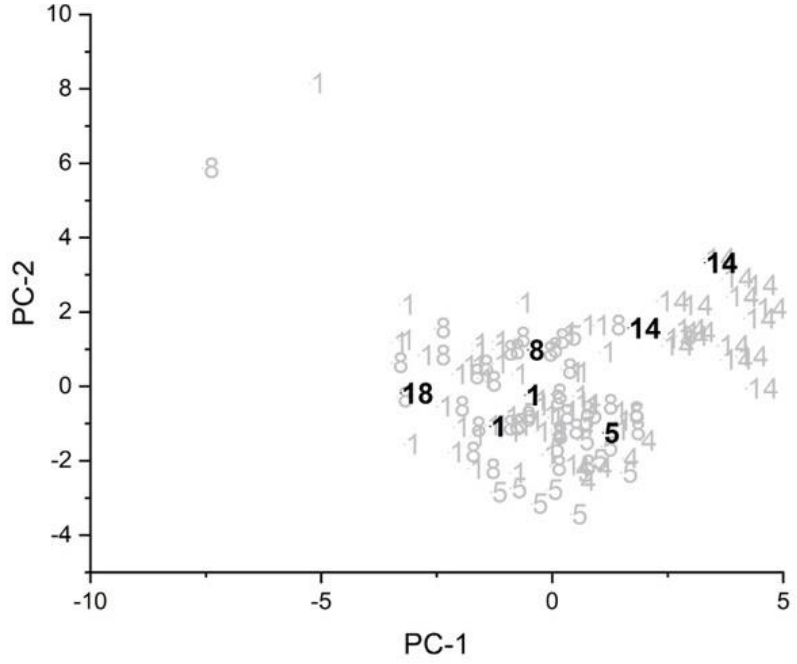

Figure 5.128 PC plot for the fourteenth training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

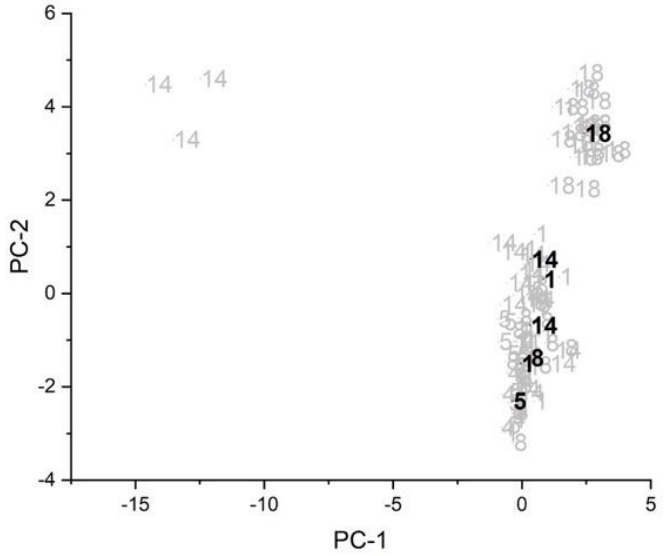

Figure 5.129 PC plot for the fourteenth training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

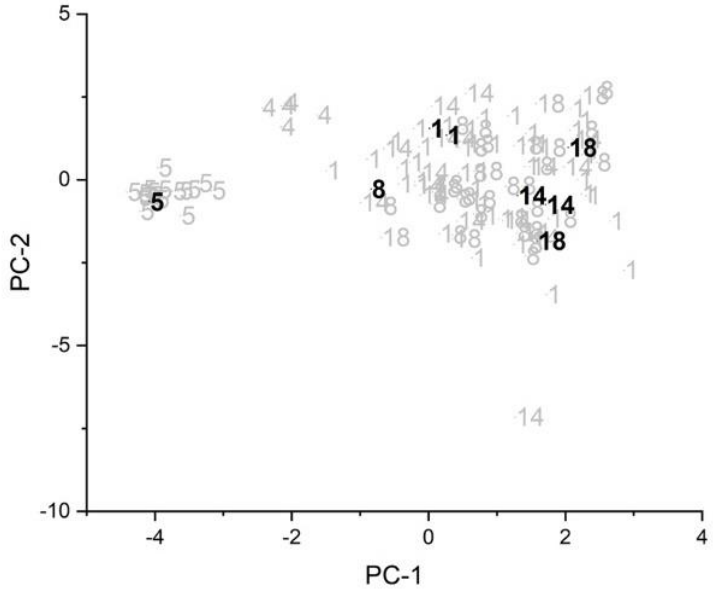

Figure 5.130 PC plot for the fifteenth training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

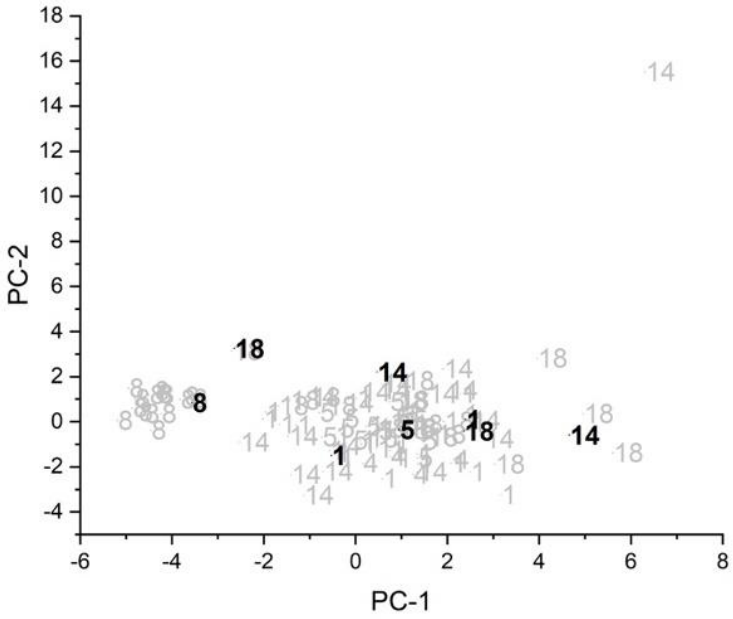

Figure 5.131 PC plot for the fifteenth training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

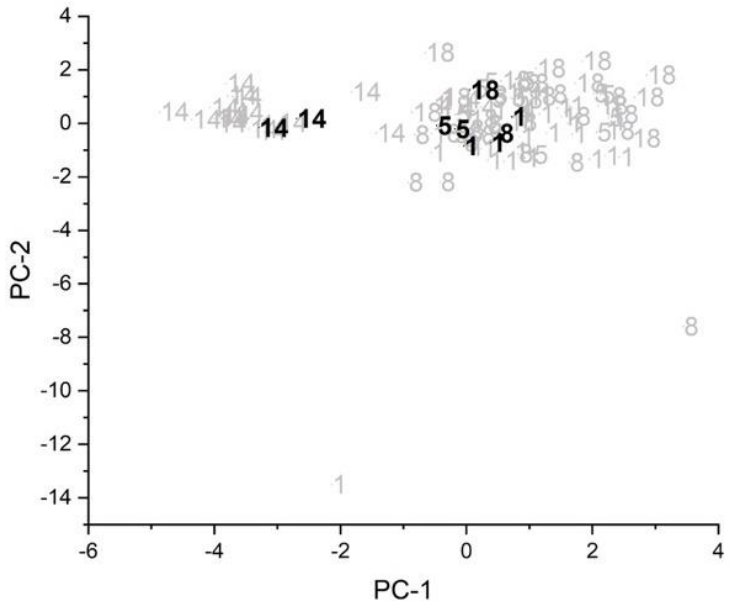

Figure 5.132 PC plot for the fifteenth training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

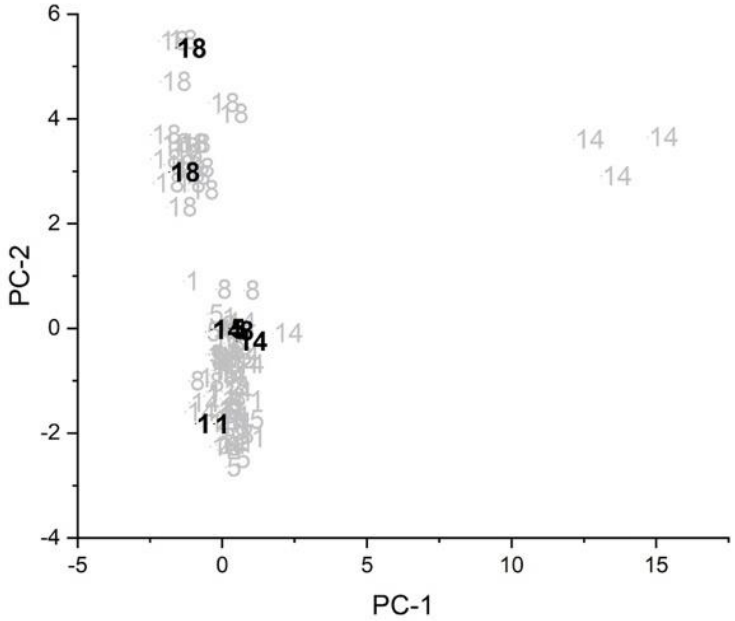

Figure 5.133 PC plot for the fifteenth training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

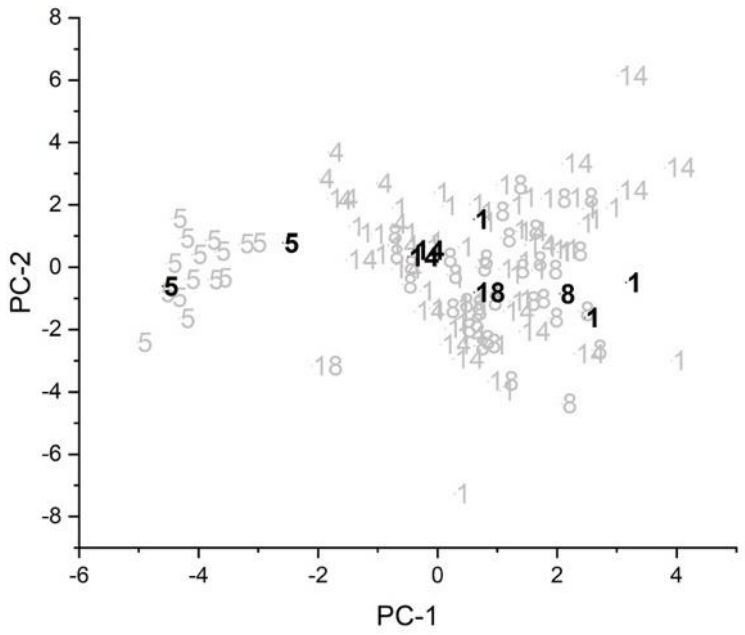

Figure 5.134 PC plot for the sixteenth training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

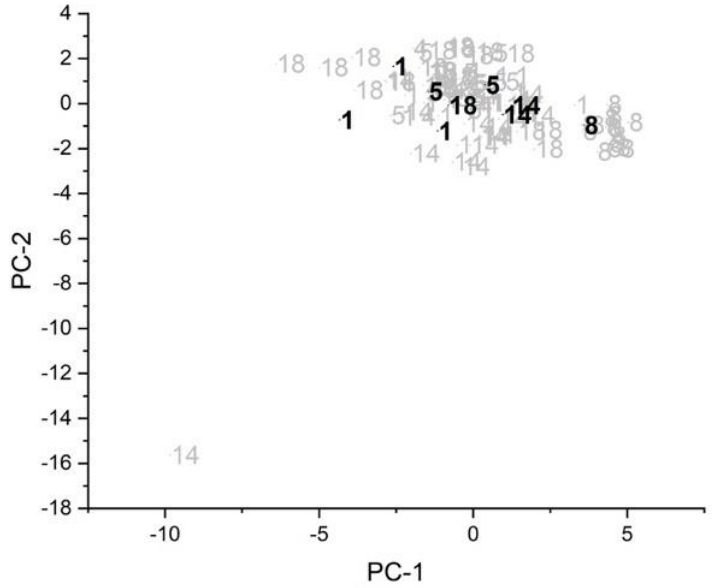

Figure 5.135 PC plot for the sixteenth training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

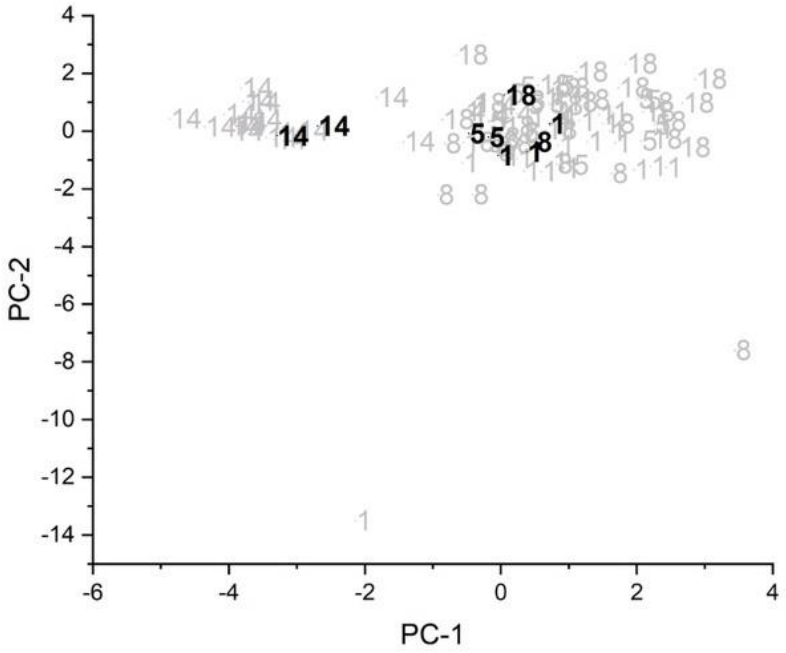

Figure 5.136 PC plot for the sixteenth training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

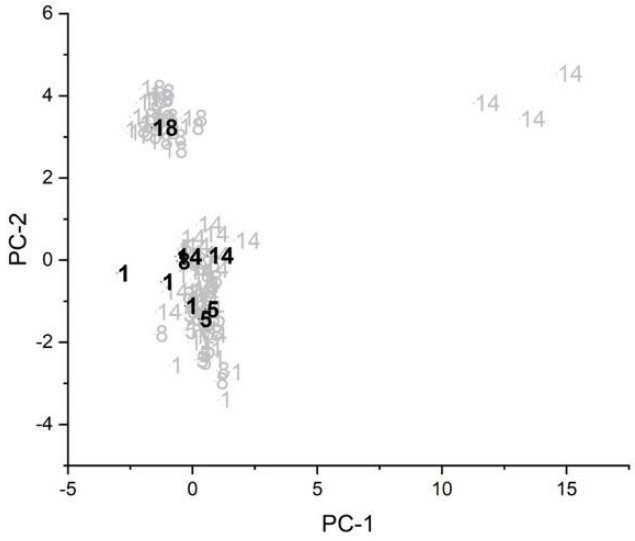

Figure 5.137 PC plot for the sixteenth training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

Second Study FTIR (Binary Classification)

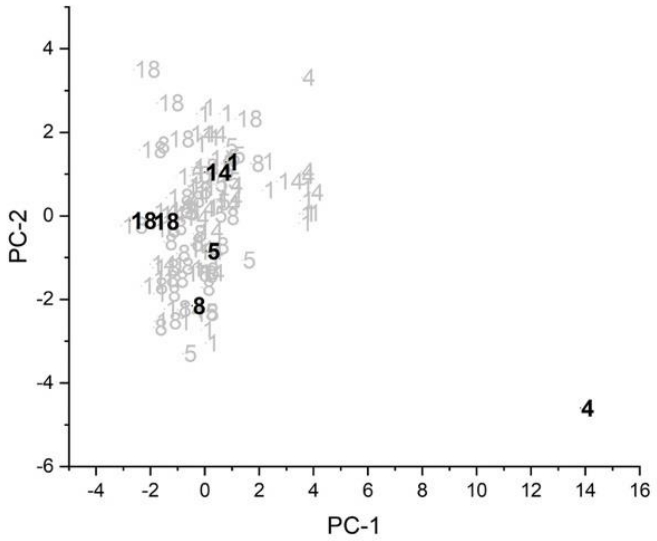

Figure 5.138 PC plot for the first training set/prediction set pair: Doraville vs Arlington, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

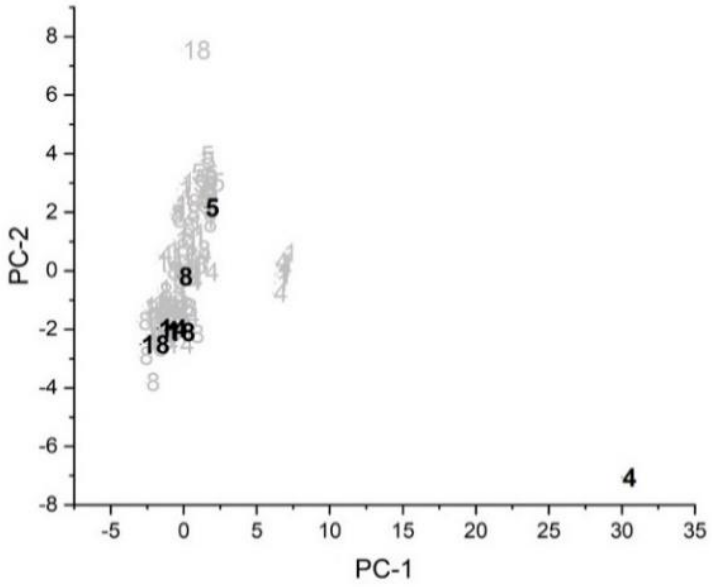

Figure 5.139 PC plot for the first training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

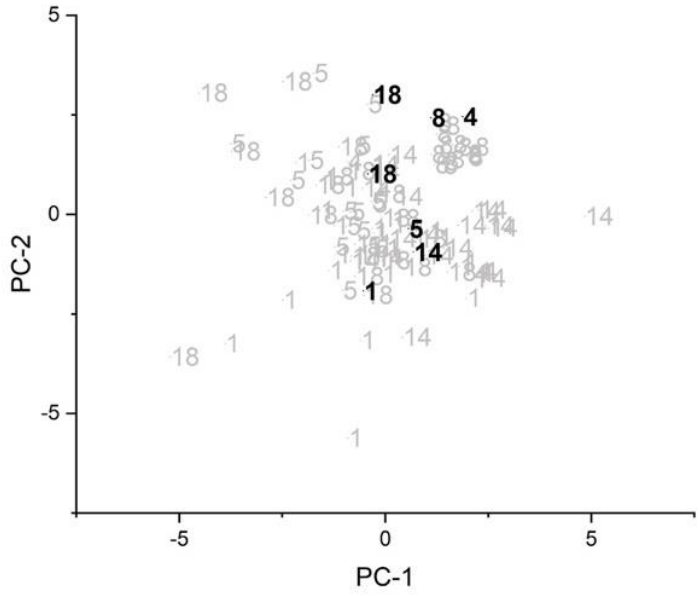

Figure 5.140 PC plot for the first training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

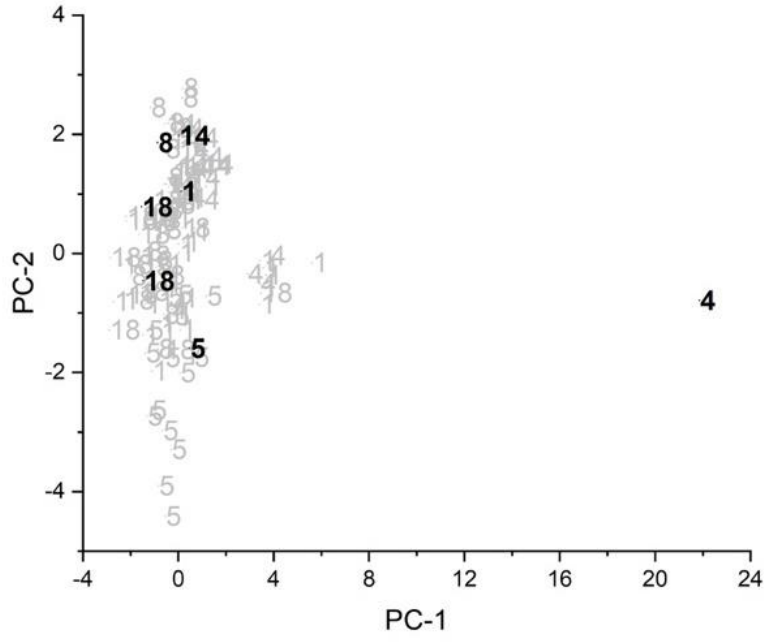

Figure 5.141 PC plot for the first training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

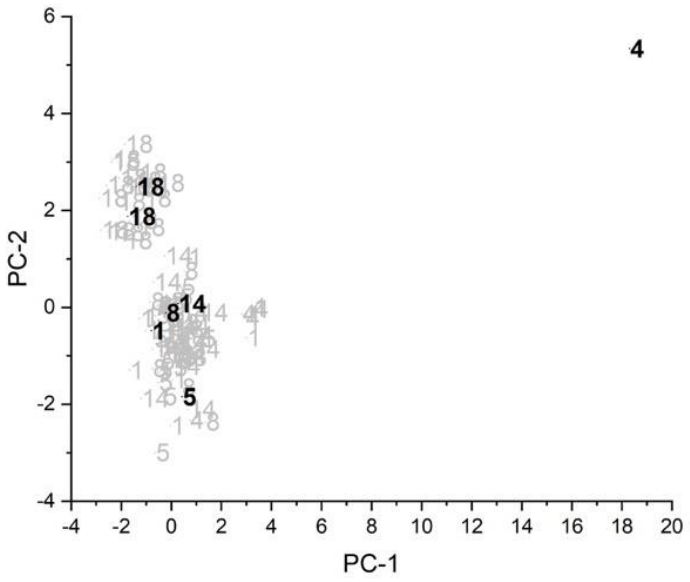

Figure 5.142 PC plot for the first training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

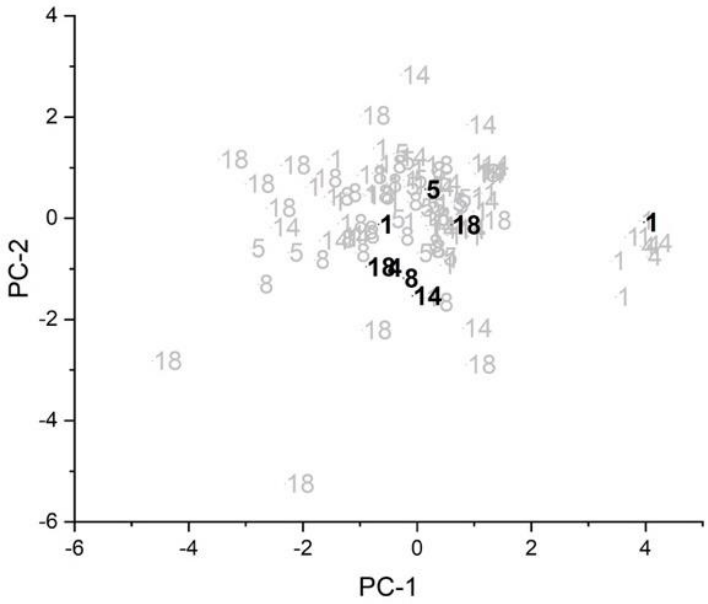

Figure 5.143 PC plot for the second training set/prediction set pair: Doraville vs Arlington, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

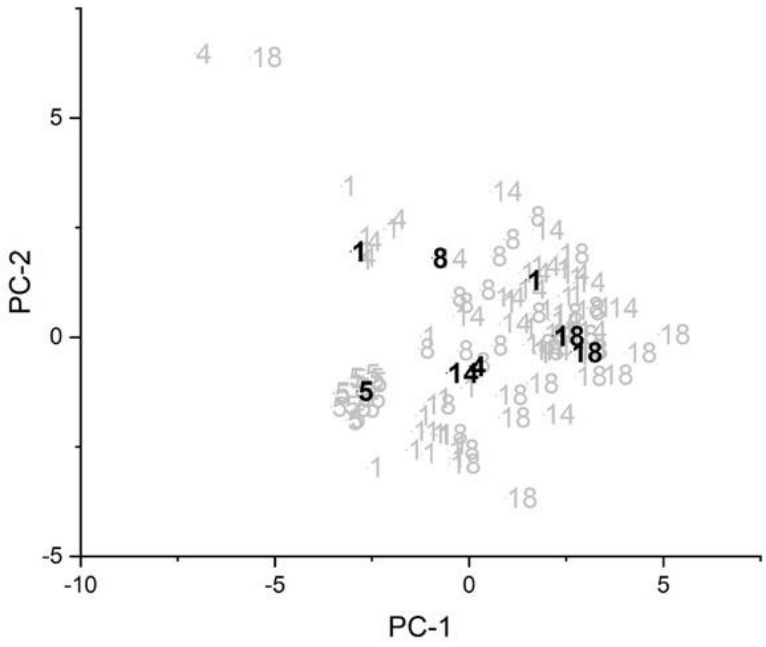

Figure 5.144 PC plot for the second training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

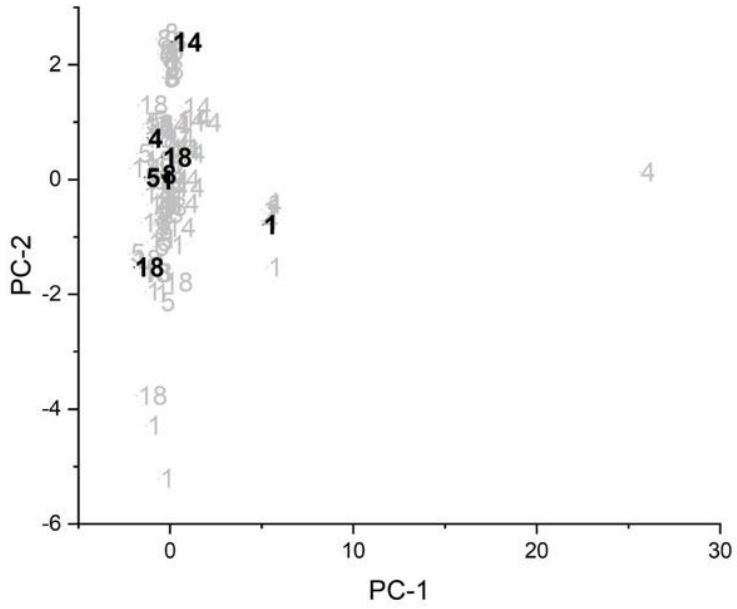

Figure 5.145 PC plot for the second training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

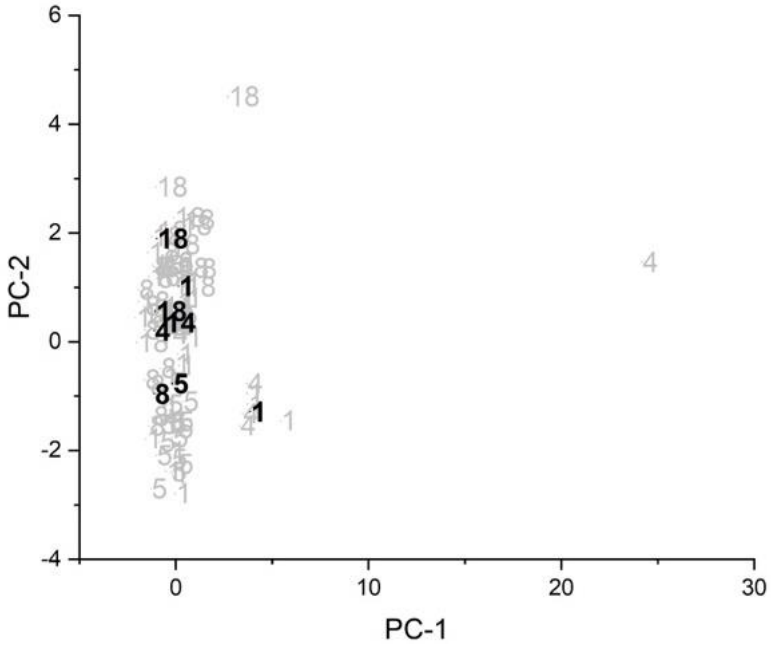

Figure 5.146 PC plot for the second training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

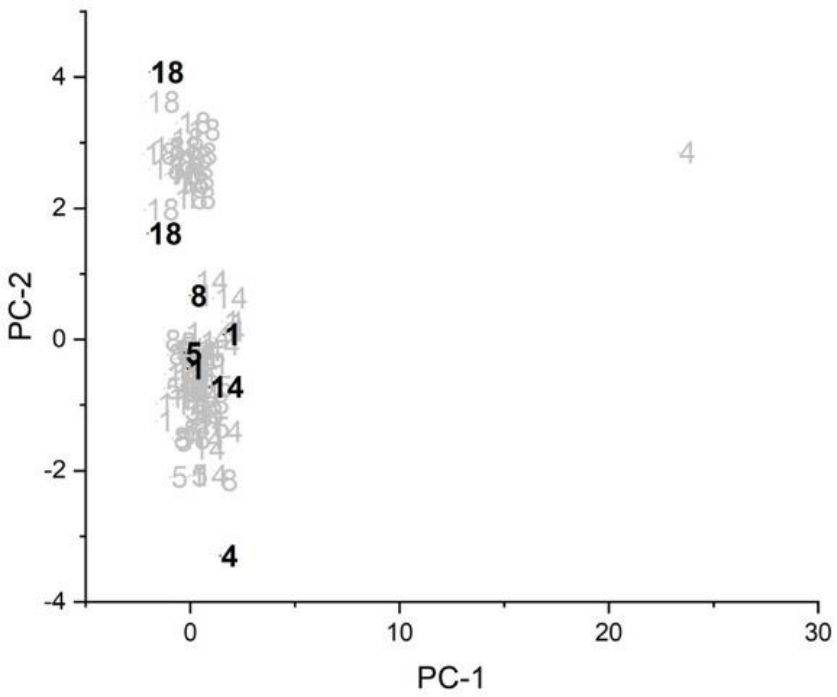

Figure 5.147 PC plot for the second training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.
.

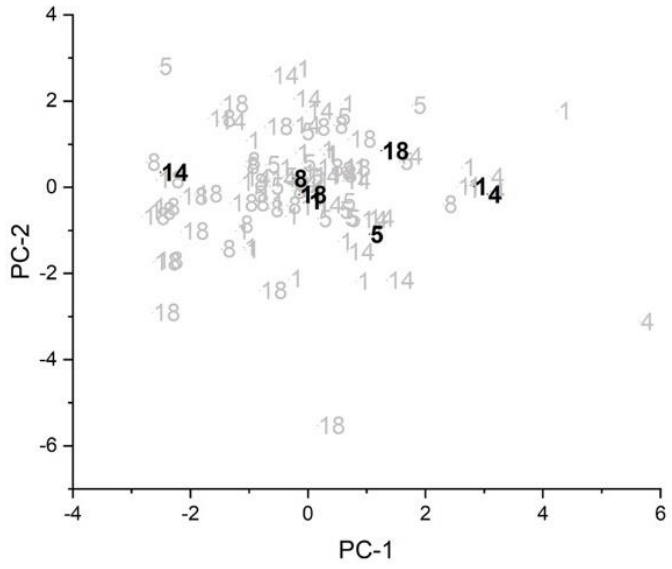

Figure 5.148 PC plot for the third training set/prediction set pair: Doraville vs Arlington, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

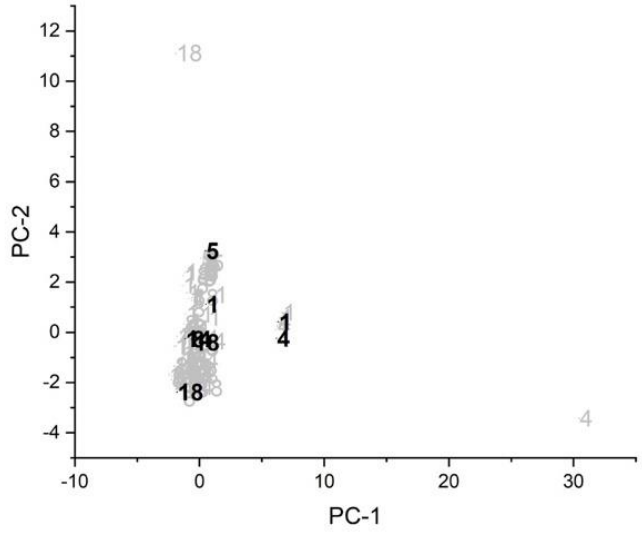

Figure 5.149 PC plot for the third training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

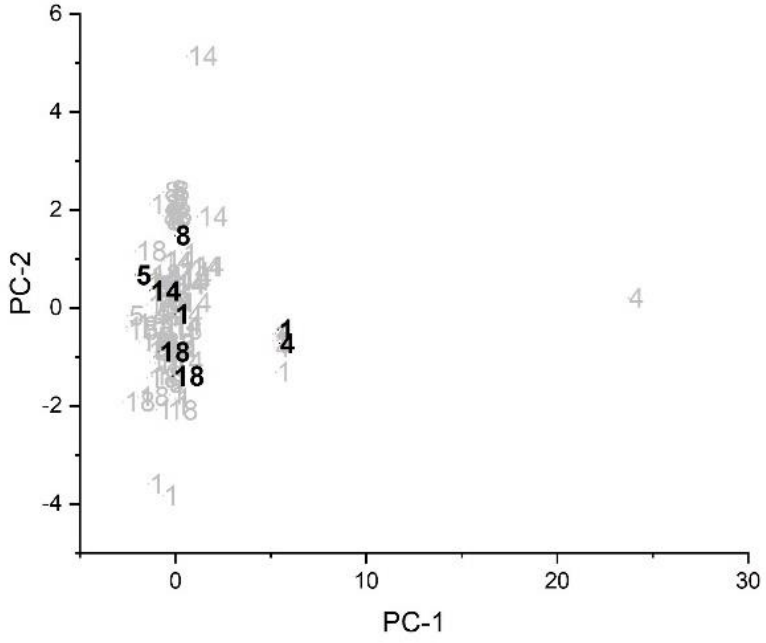

Figure 5.150 PC plot for the third training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

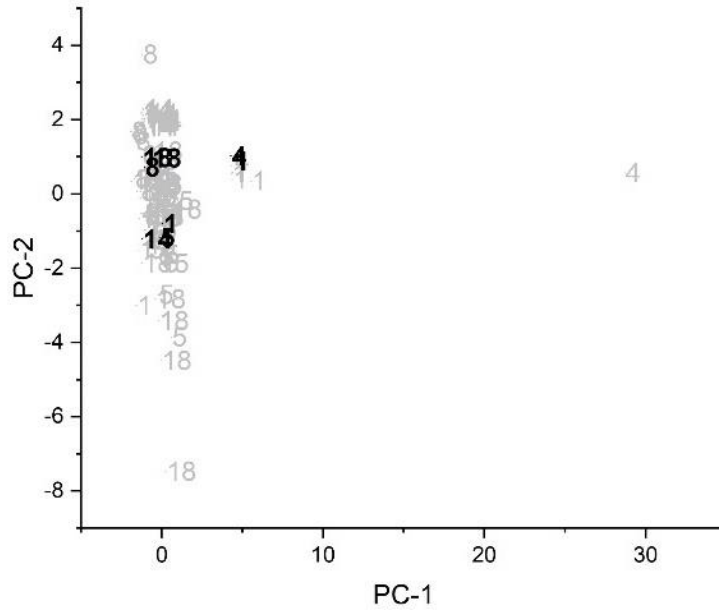

Figure 5.151 PC plot for the third training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

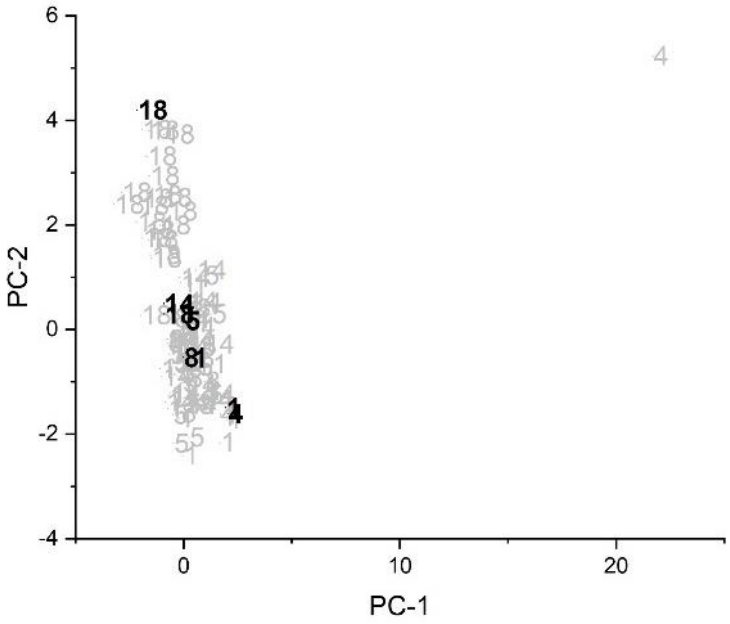

Figure 5.152 PC plot for the third training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

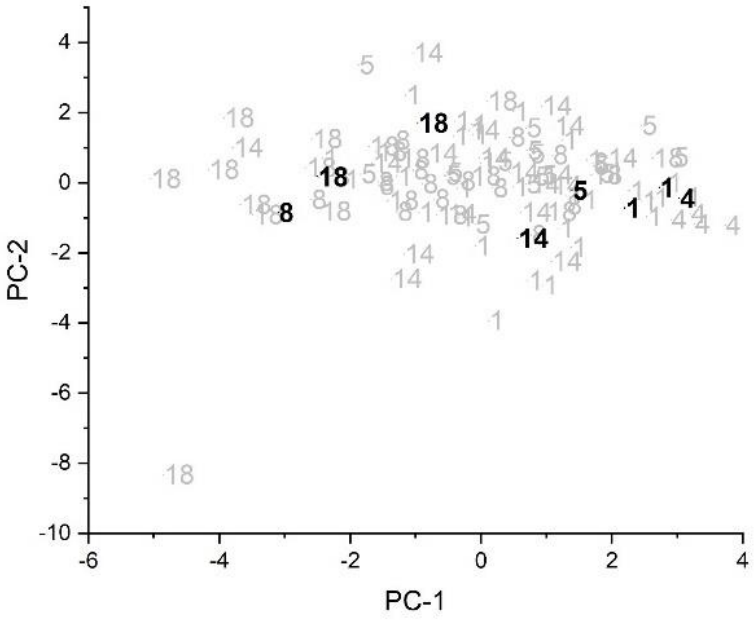

Figure 5.153 PC plot for the fourth training set/prediction set pair: Doraville vs Arlington, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

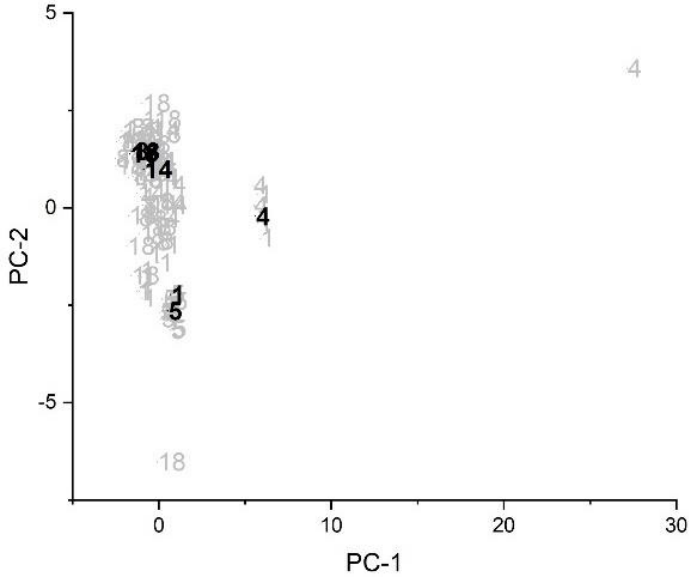

Figure 5.154 PC plot for the fourth training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

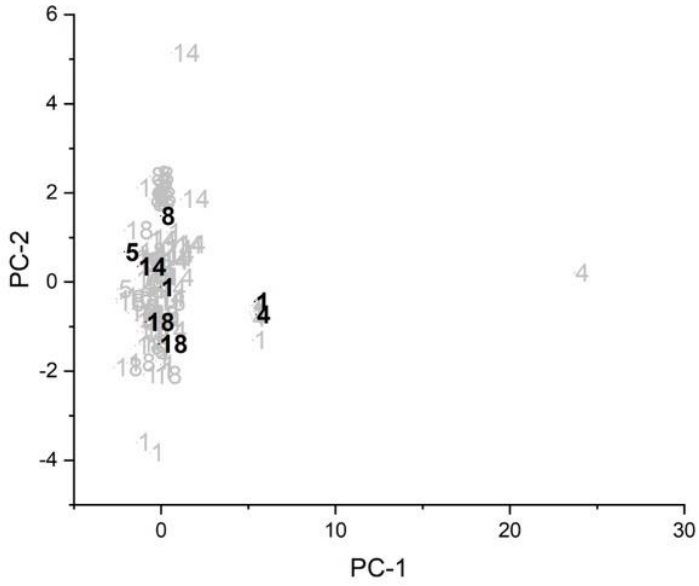

Figure 5.155 PC plot for the fourth training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

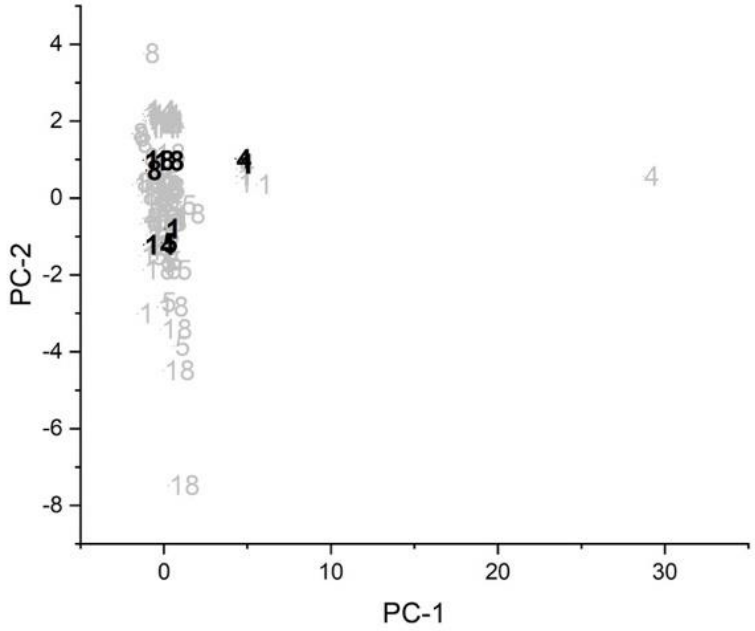

Figure 5.156 PC plot for the fourth training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

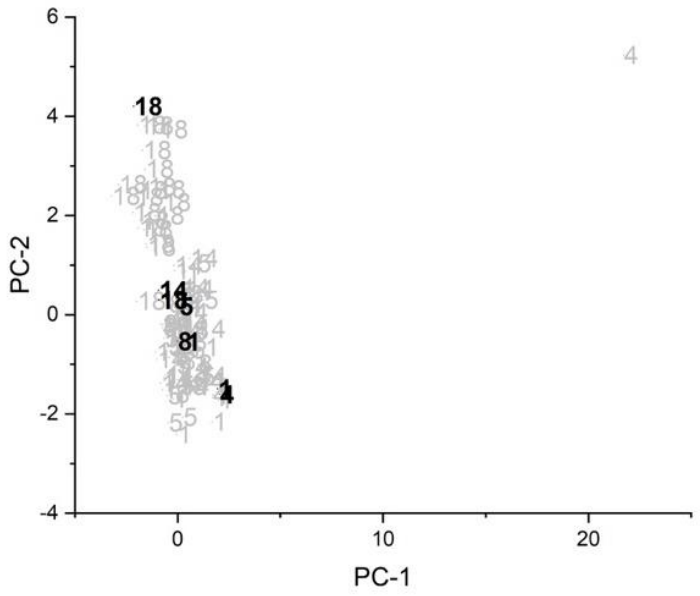

Figure 5.157 PC plot for the fourth training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

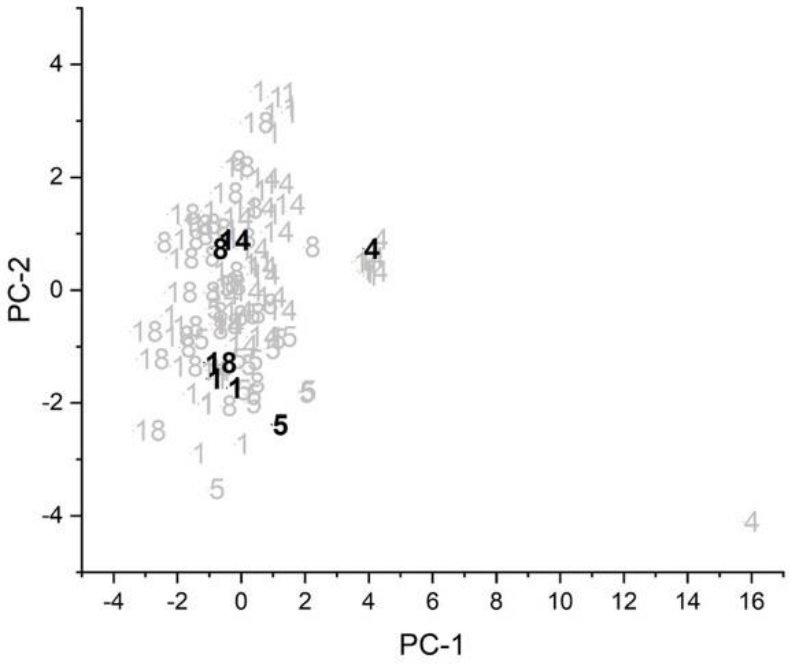

Figure 5.158 PC plot for the fifth training set/prediction set pair: Doraville vs Arlington, Fairfax, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

.

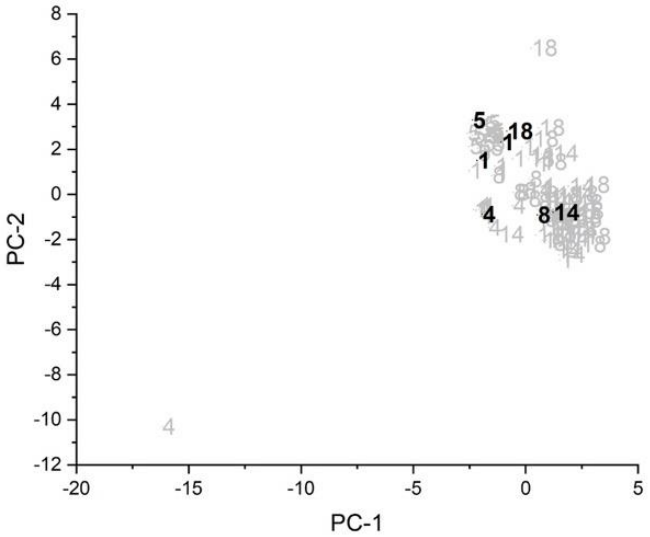

Figure 5.159 PC plot for the fifth training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

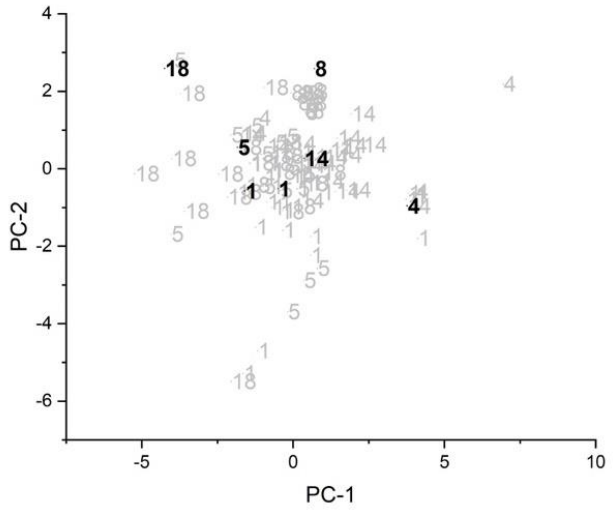

Figure 5.160 PC plot for the fifth training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

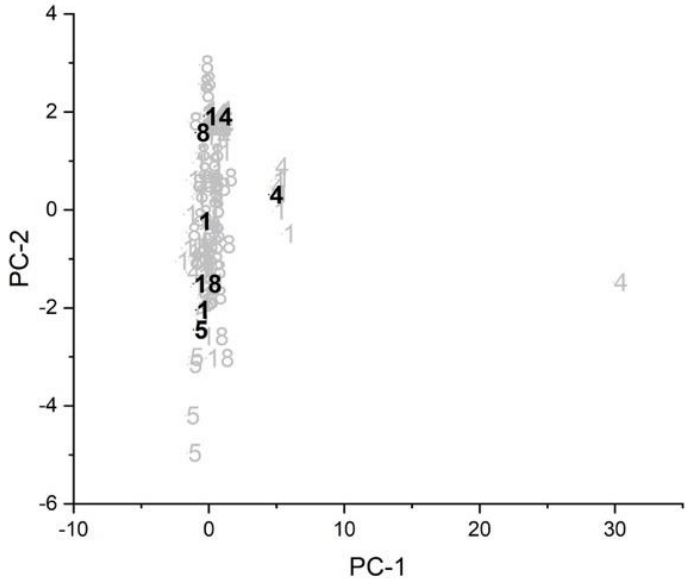

Figure 5.161 PC plot for the fifth training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

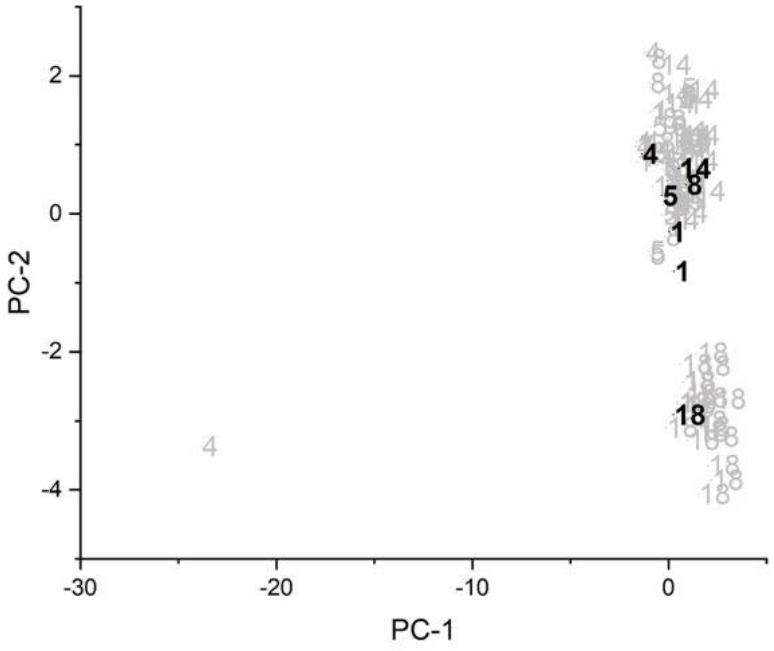

Figure 5.162 PC plot for the fifth training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

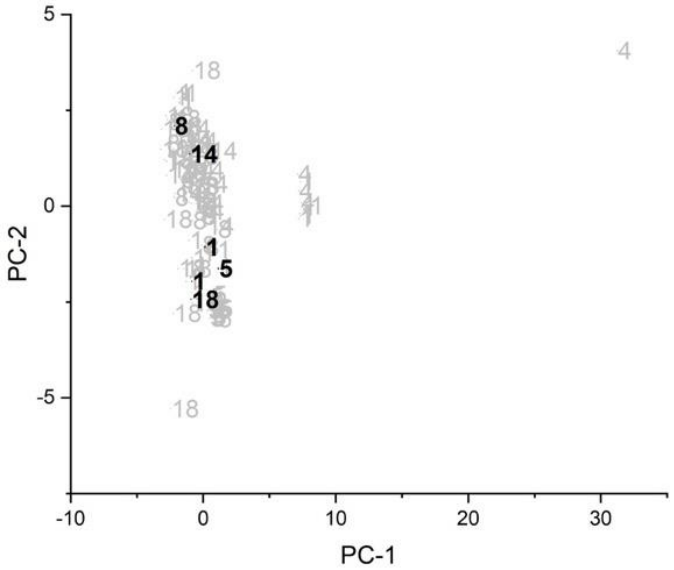

Figure 5.163 PC plot for the sixth training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

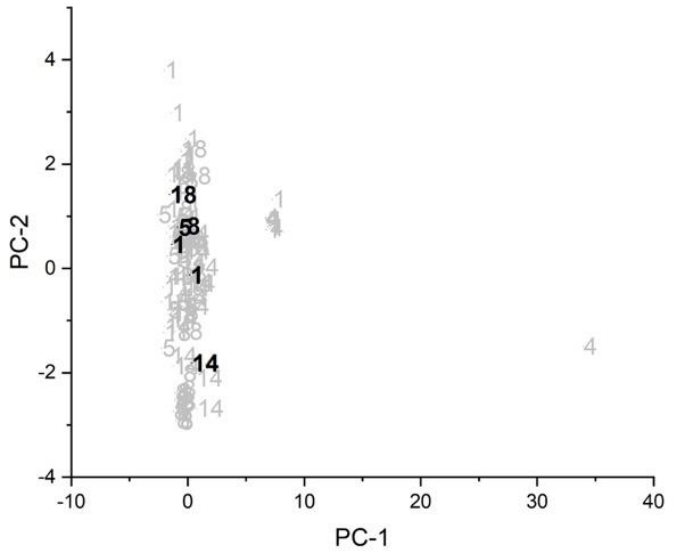

Figure 5.164 PC plot for the sixth training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

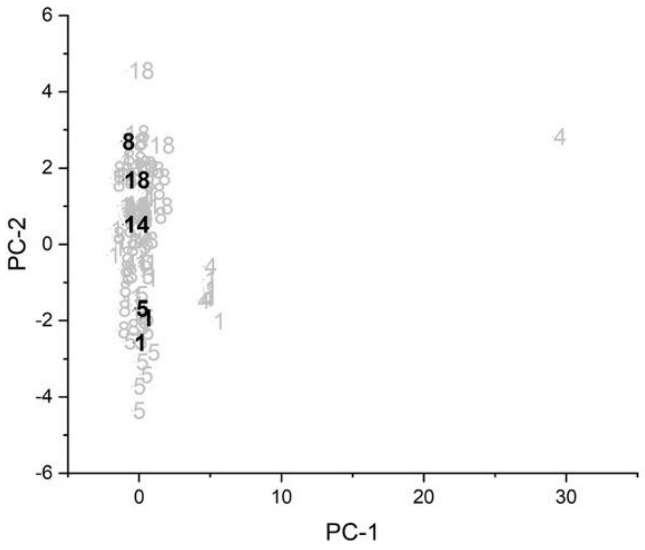

Figure 5.165 PC plot for the sixth training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

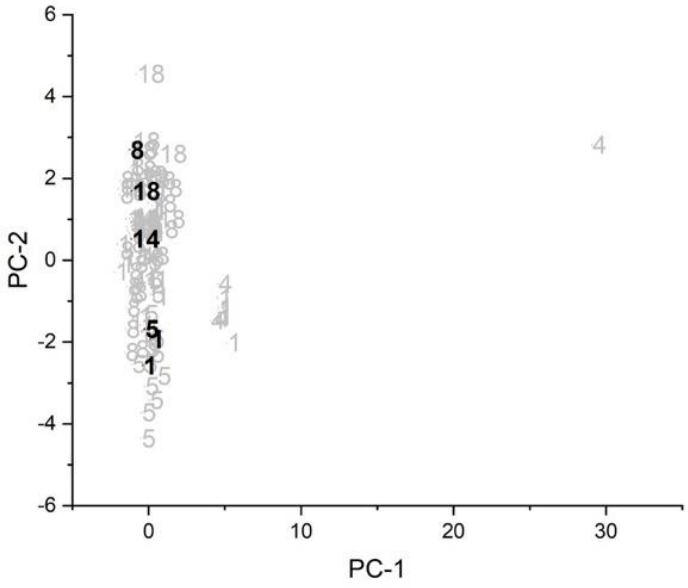

Figure 5.166 PC plot for the sixth training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

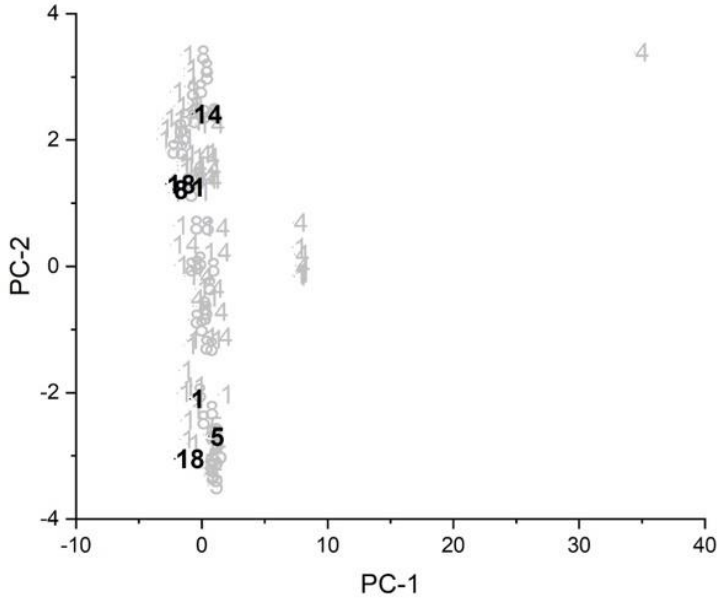

Figure 5.167 PC plot for the seventh training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

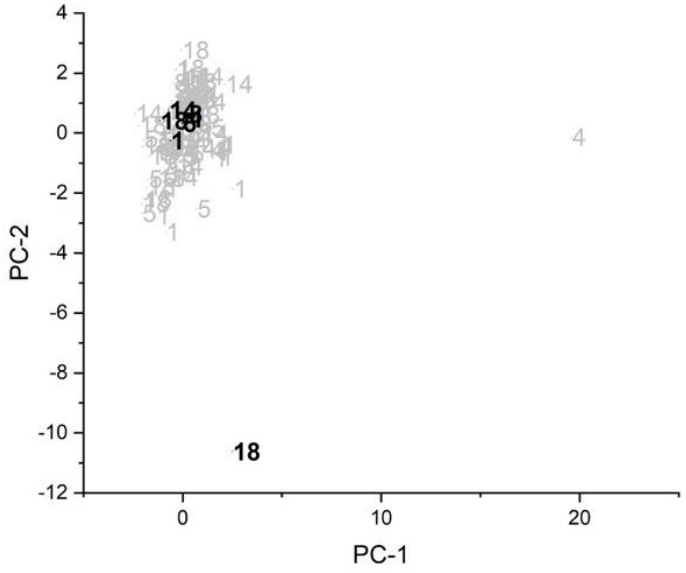

Figure 5.168 PC plot for the seventh training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

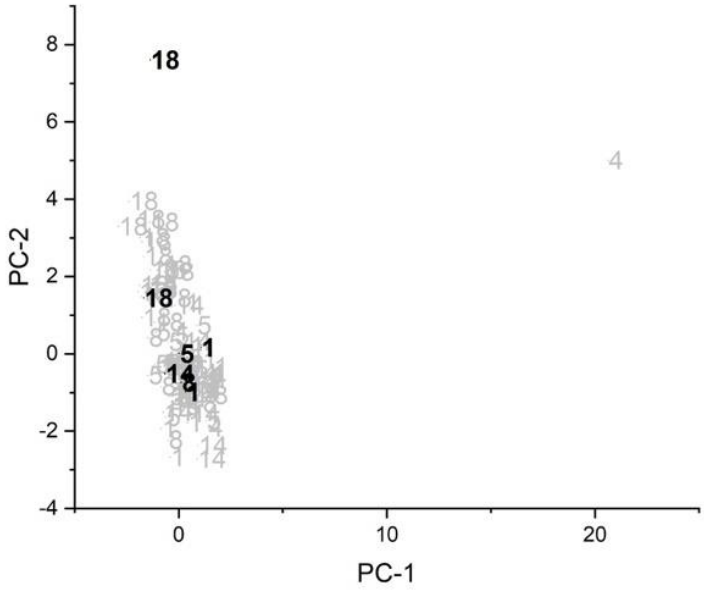

Figure 5.169 PC plot for the seventh training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

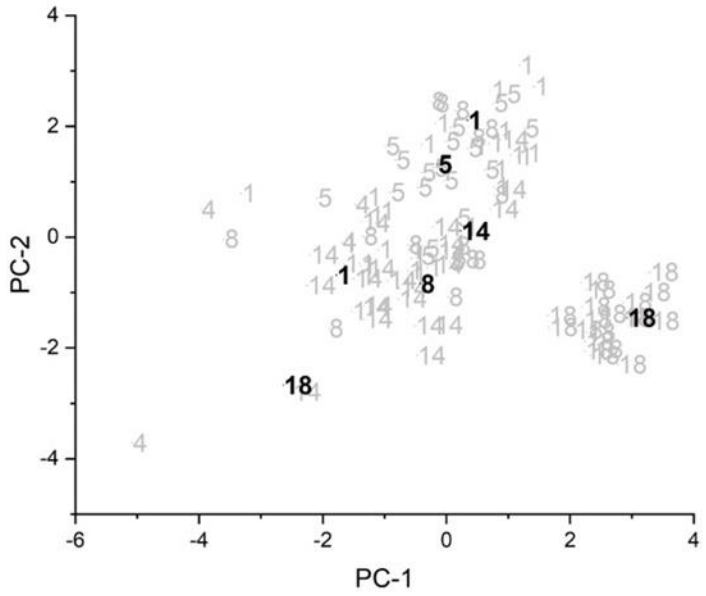

Figure 5.170 PC plot for the seventh training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

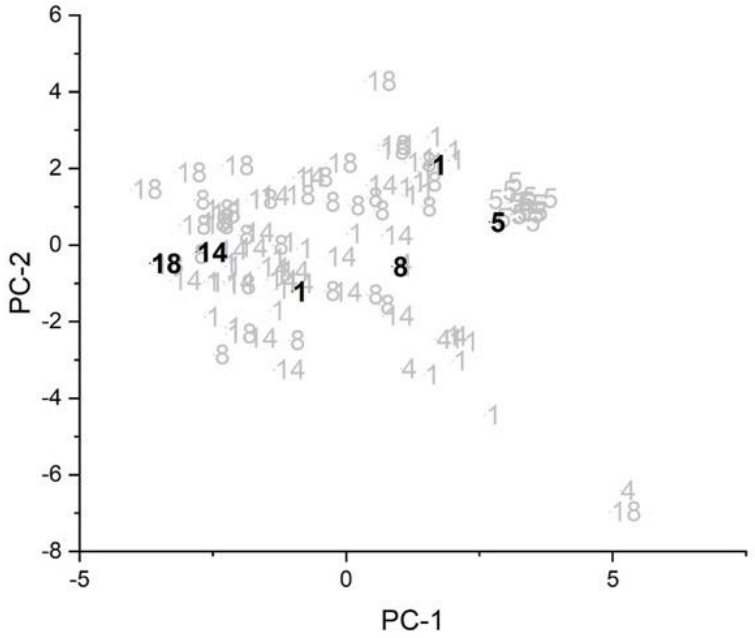

Figure 5.171 PC plot for the eighth training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

.

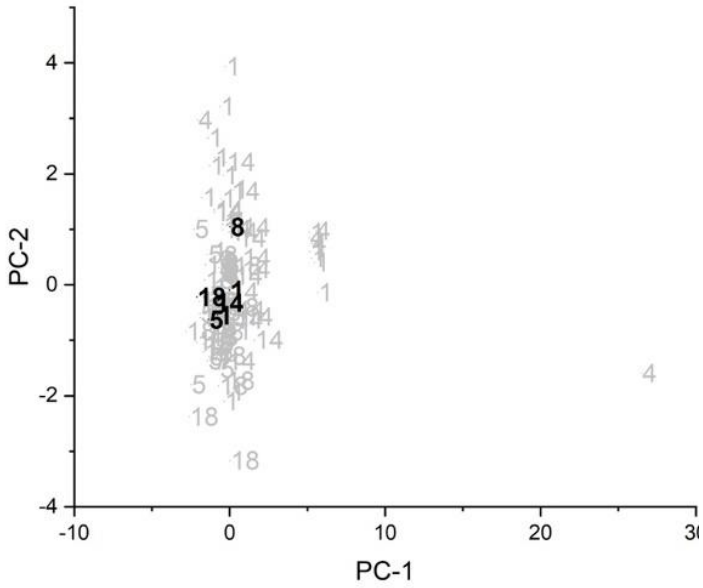

Figure 5.172 PC plot for the eighth training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

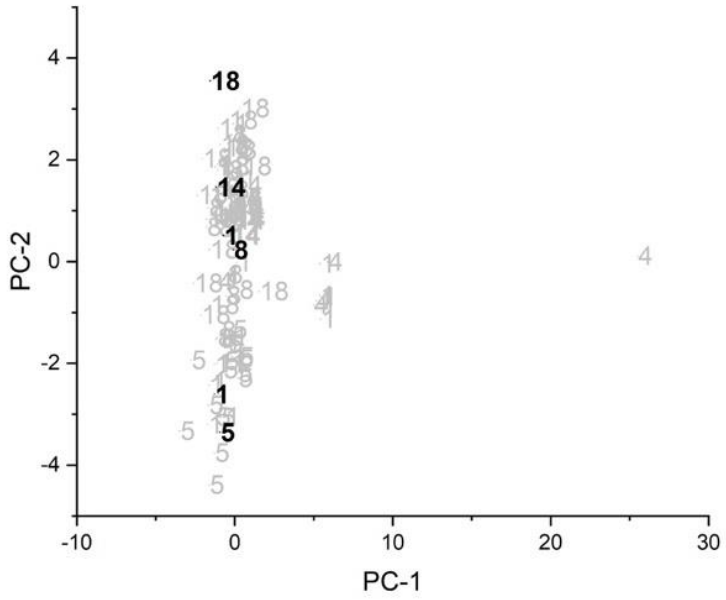

Figure 5.173 PC plot for the eighth training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

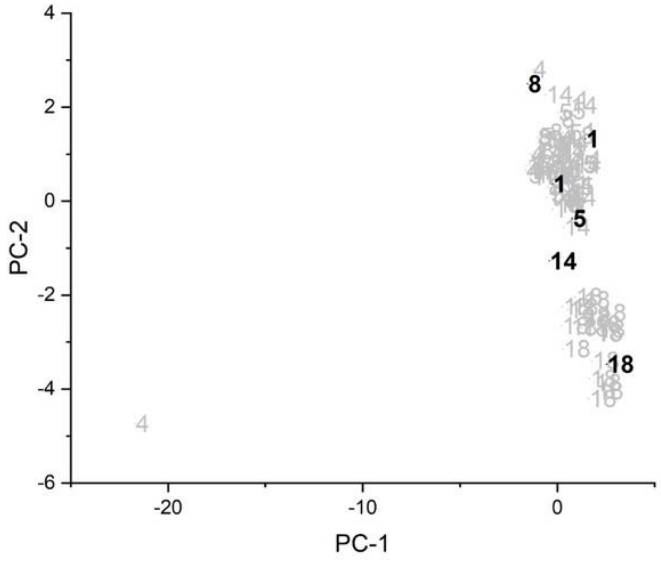

Figure 5.174 PC plot for the eighth training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

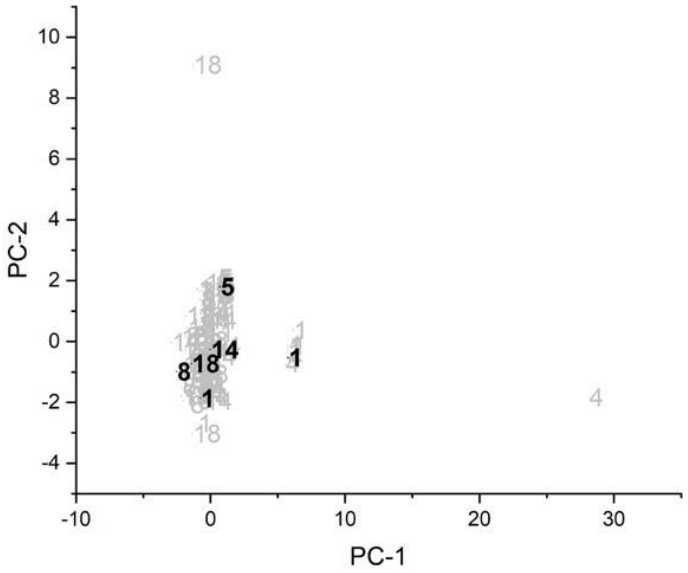

Figure 5.175 PC plot for the ninth training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

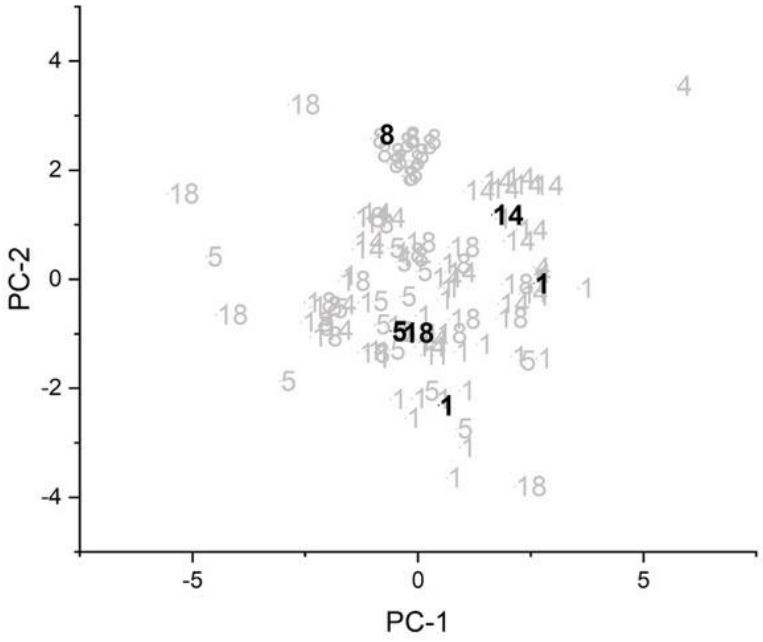

Figure 5.176 PC plot for the ninth training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

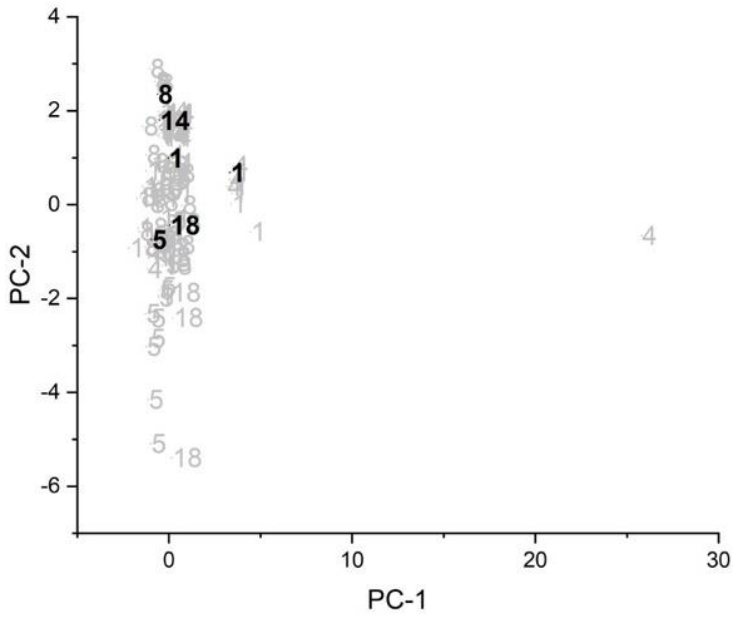

Figure 5.177 PC plot for the ninth training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

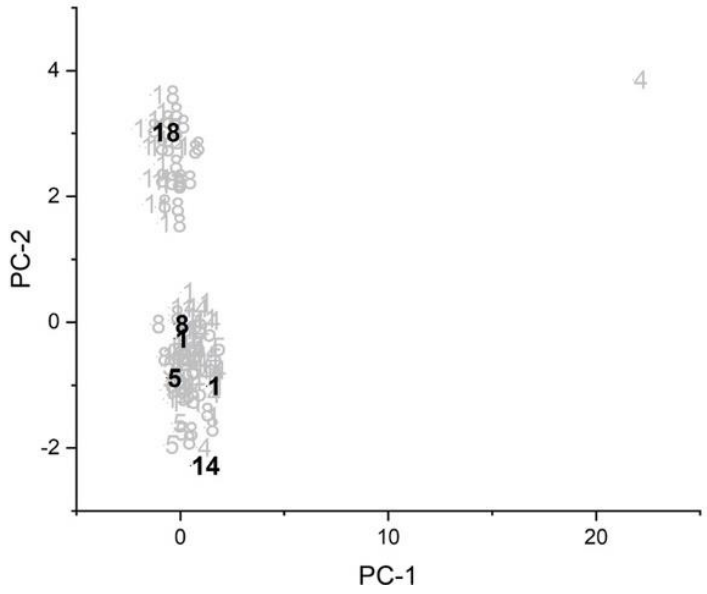

Figure 5.178 PC plot for the ninth training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

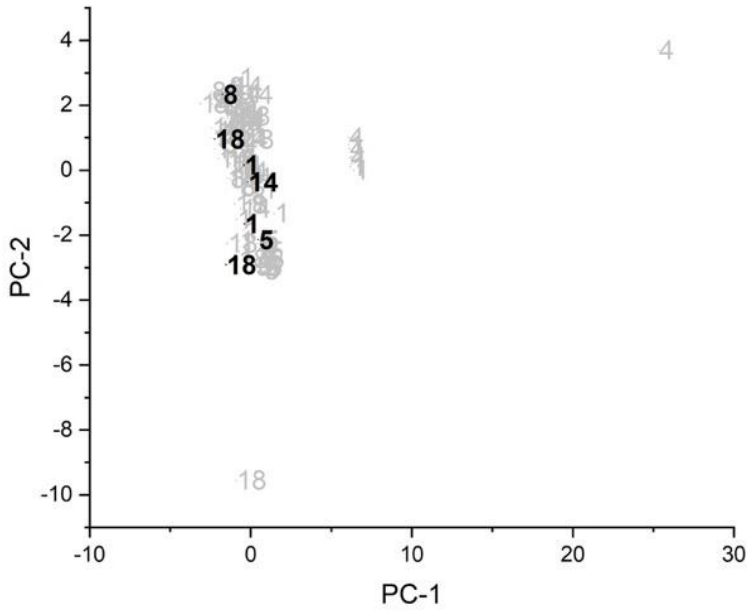

Figure 5.179 PC plot for the tenth training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

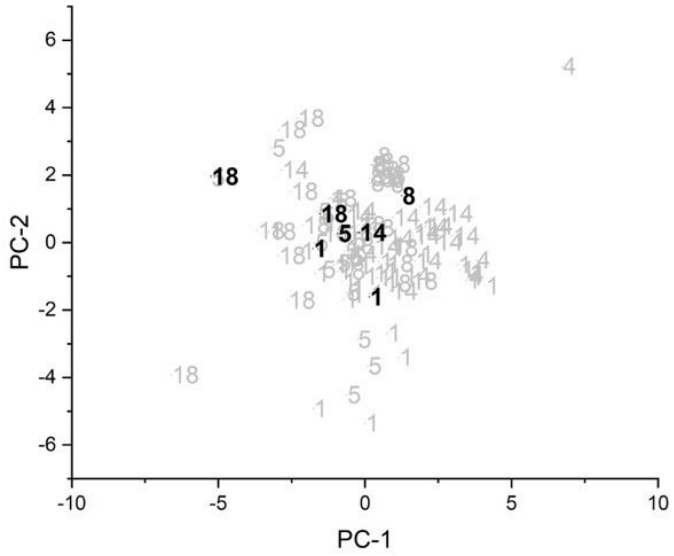

Figure 5.180 PC plot for the tenth training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

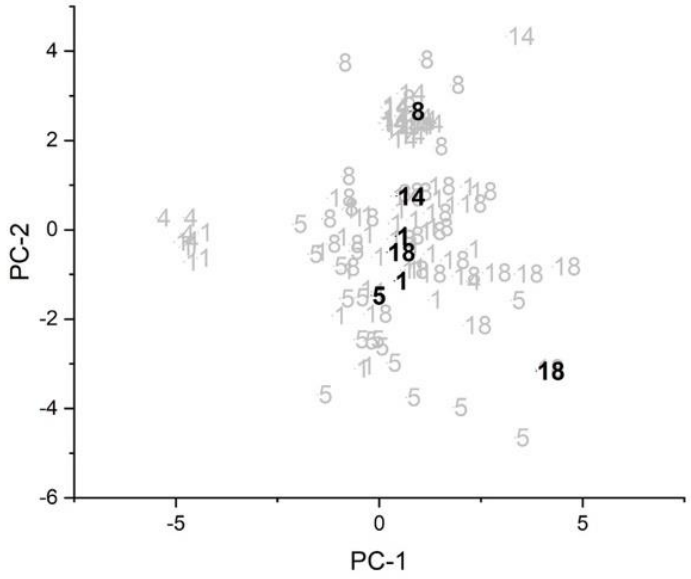

Figure 5.181 PC plot for the tenth training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

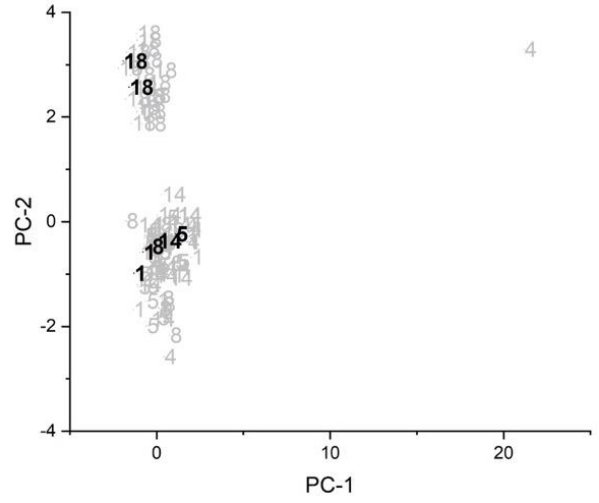

Figure 5.182 PC plot for the tenth training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

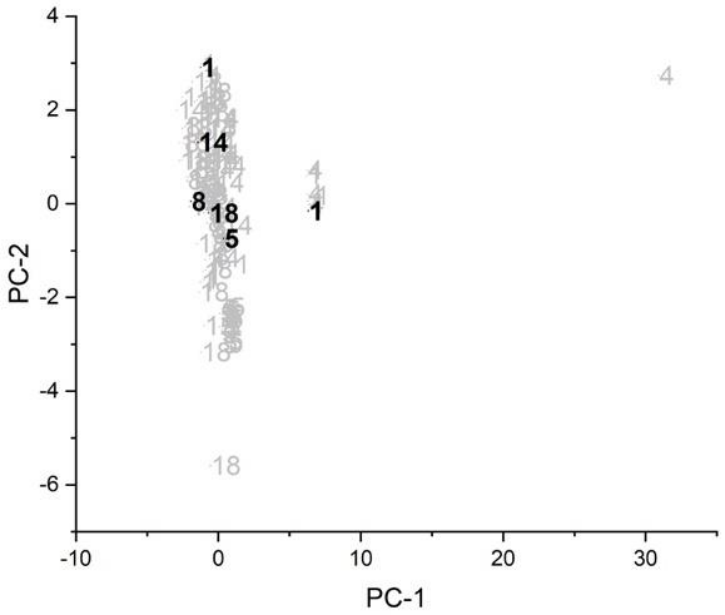

Figure 5.183 PC plot for the eleventh training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

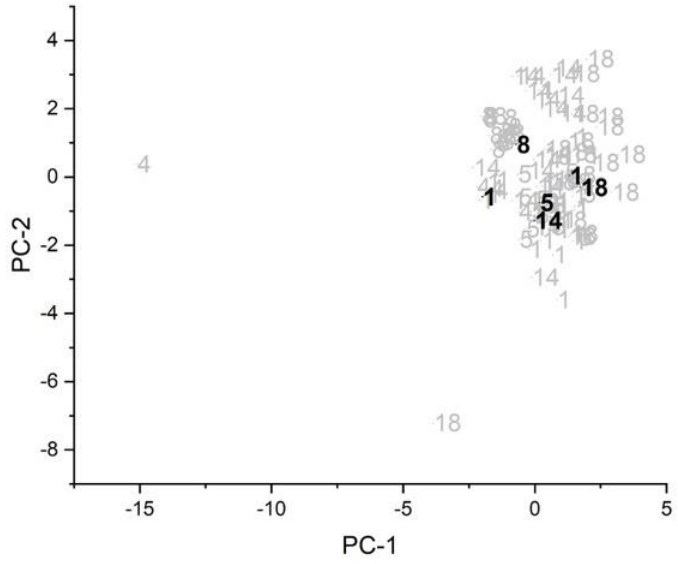

Figure 5.184 PC plot for the eleventh training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

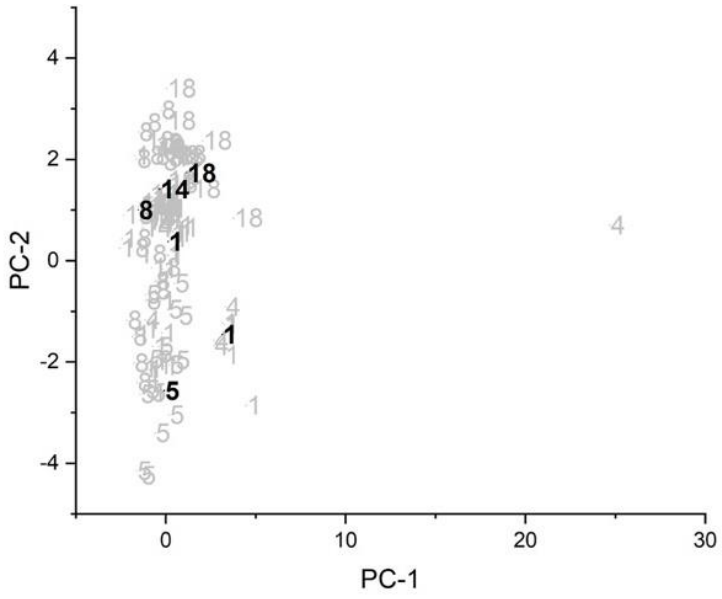

Figure 5.185 PC plot for the eleventh training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

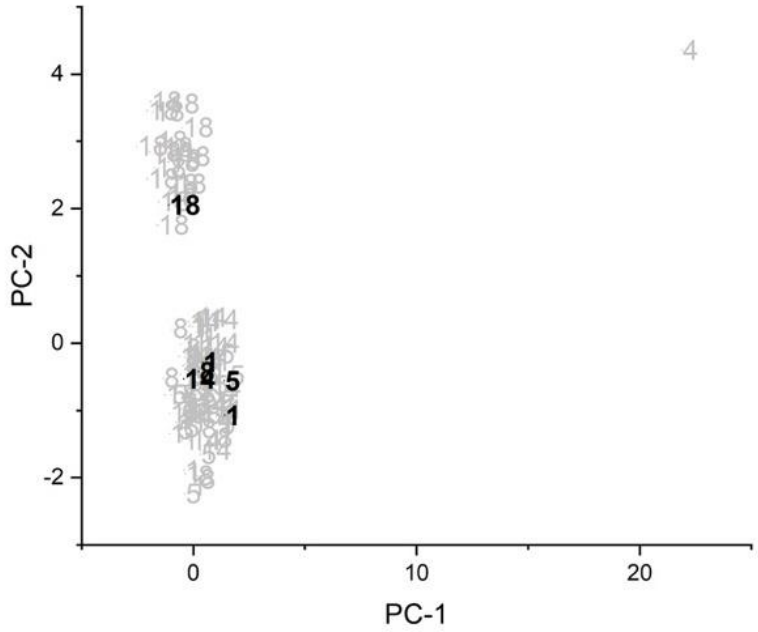

Figure 5.186 PC plot for the eleventh training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

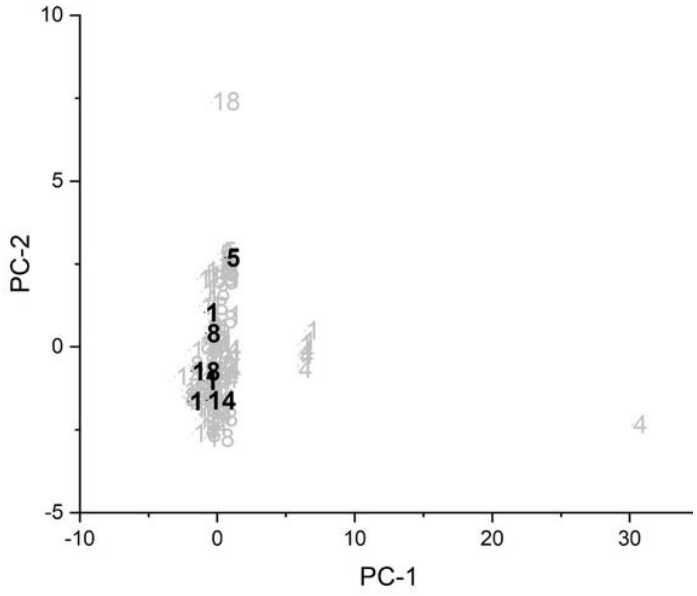

Figure 5.187 PC plot for the twelfth training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

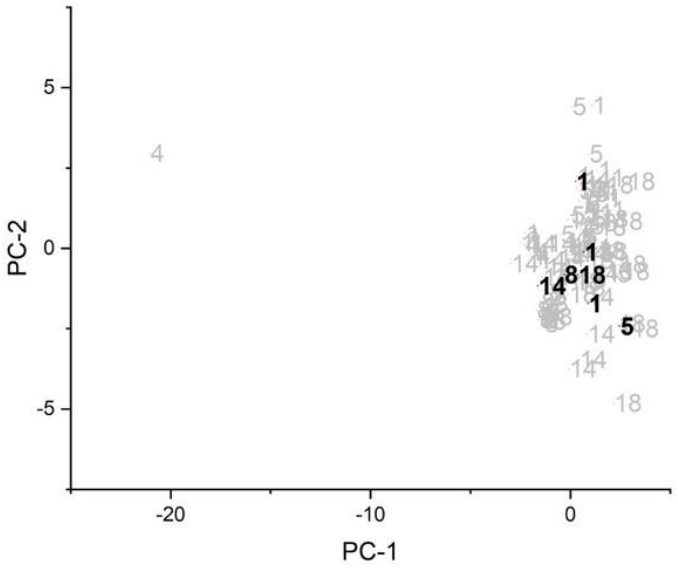

Figure 5.188 PC plot for the twelfth training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

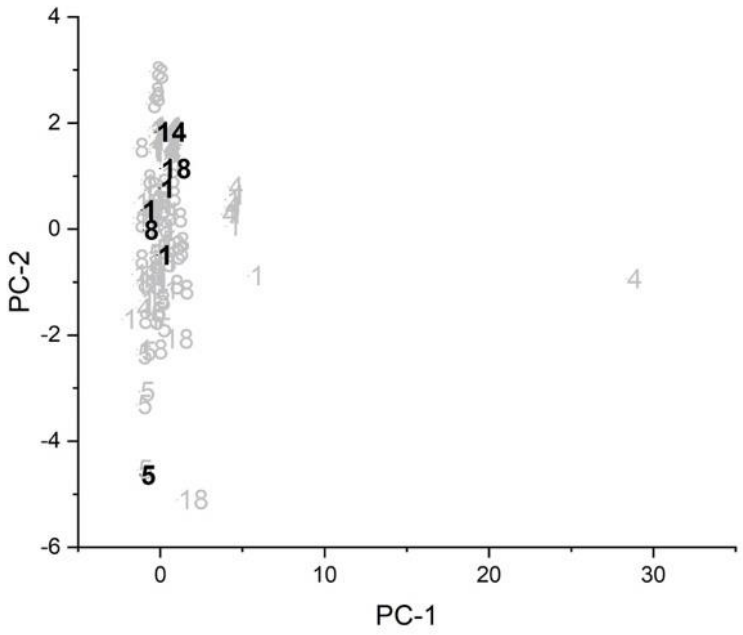

Figure 5.189 PC plot for the twelfth training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

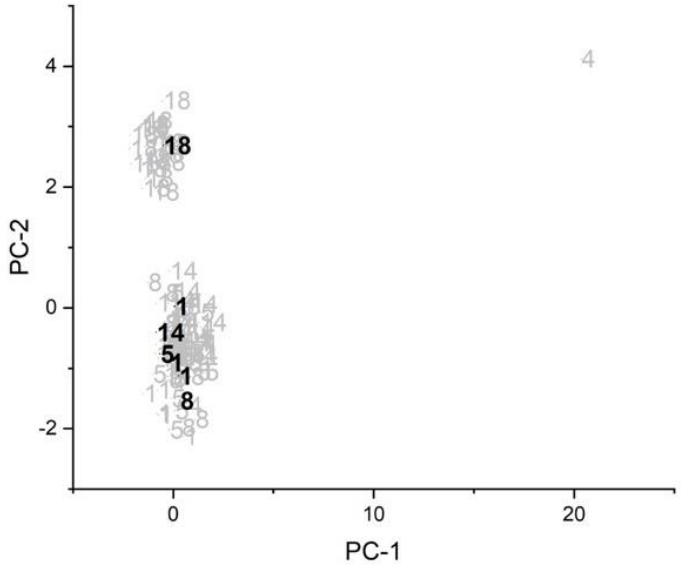

Figure 5.190 PC plot for the twelfth training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

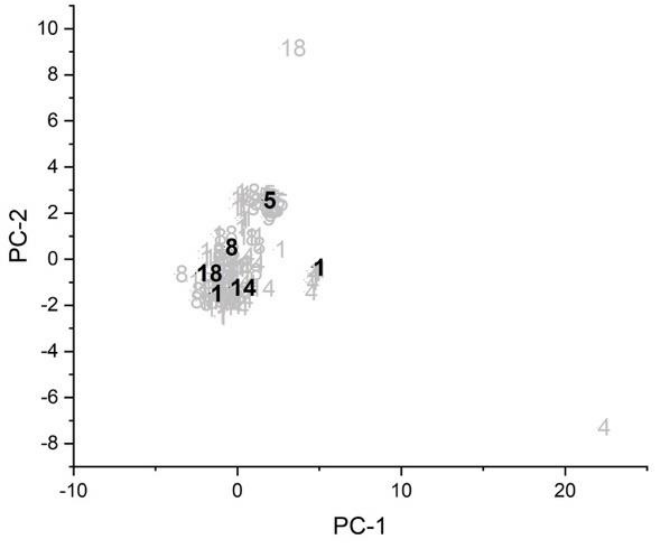

Figure 5.191 PC plot for the thirteenth training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

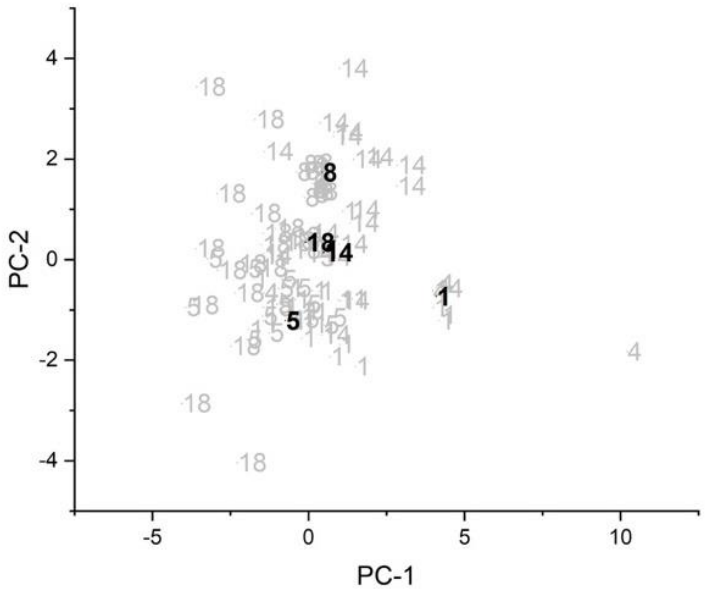

Figure 5.192 PC plot for the thirteenth training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

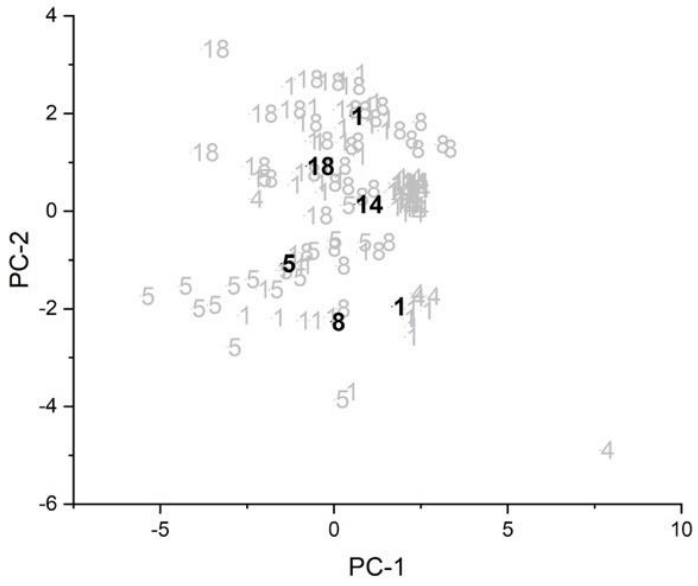

Figure 5.193 PC plot for the thirteenth training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

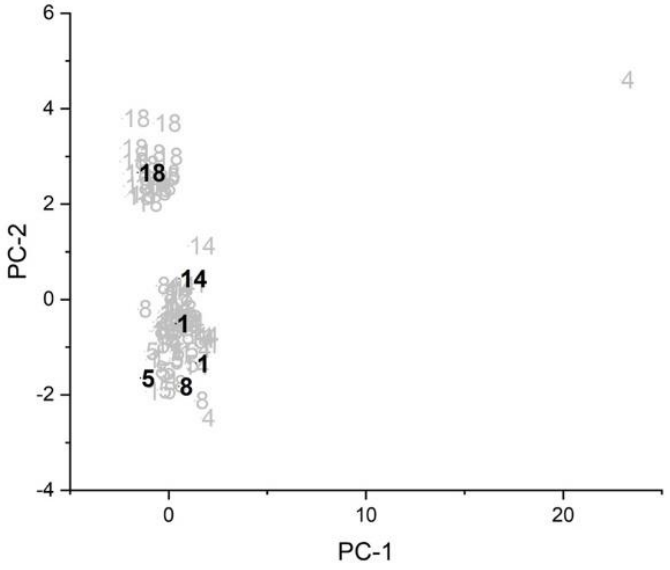

Figure 5.194 PC plot for the thirteenth training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

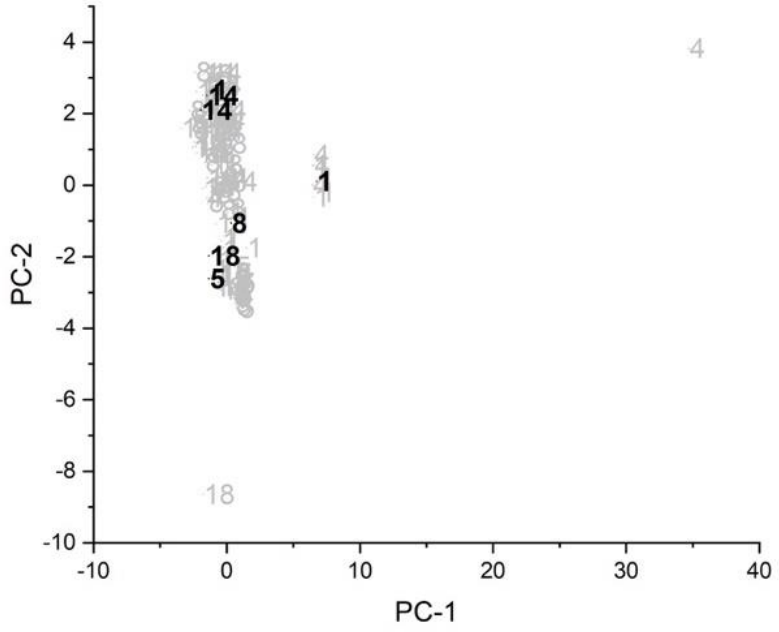

Figure 5.195 PC plot for the fourteenth training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

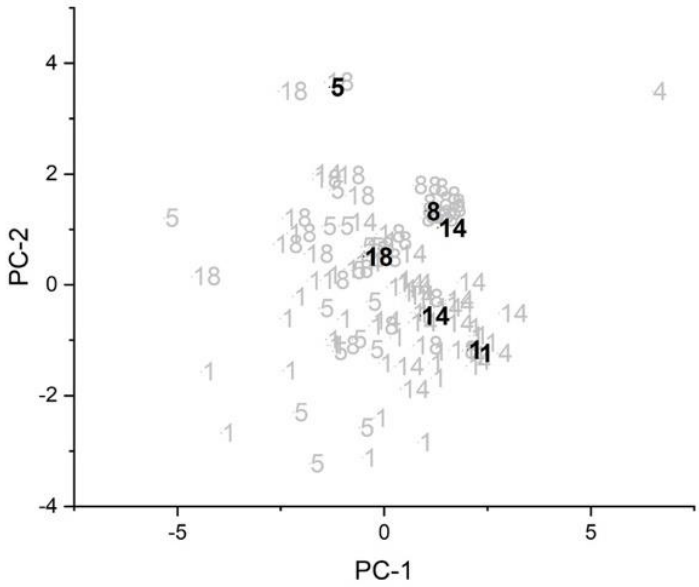

Figure 5.196 PC plot for the fourteenth training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

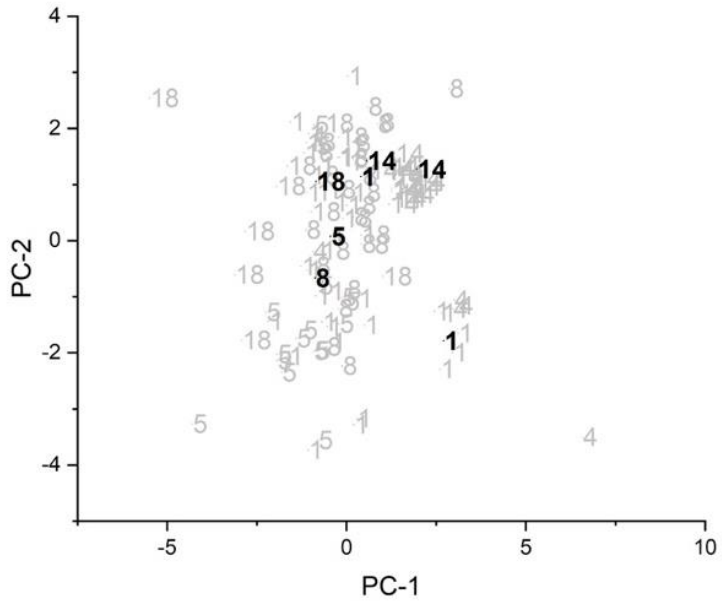

Figure 5.197 PC plot for the fourteenth training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

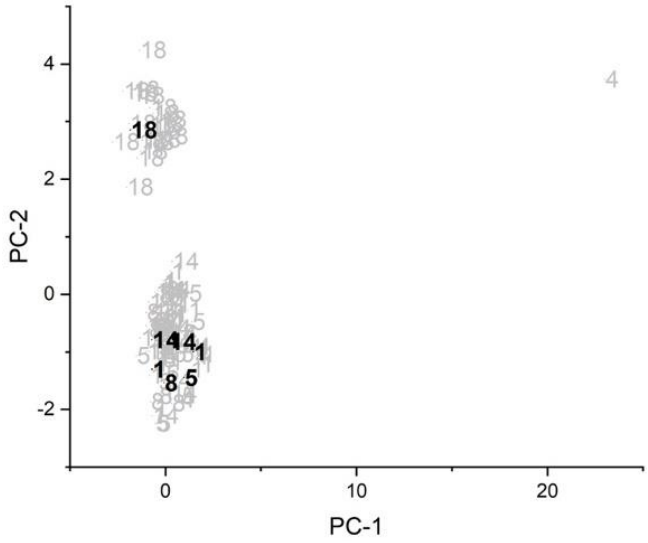

Figure 5.198 PC plot for the fourteenth training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

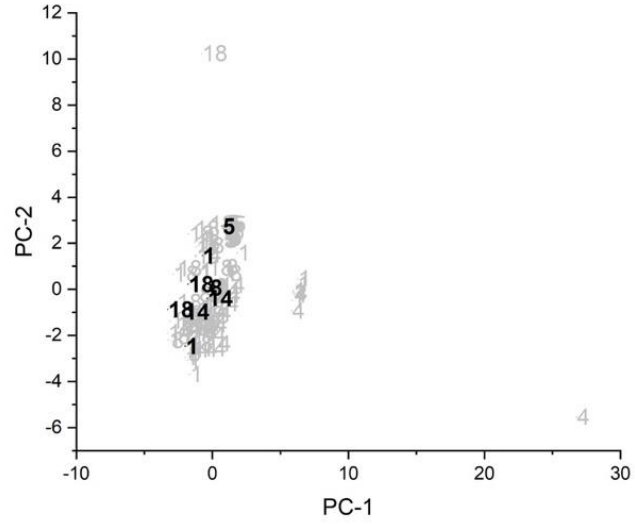

Figure 5.199 PC plot for the fifteenth training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

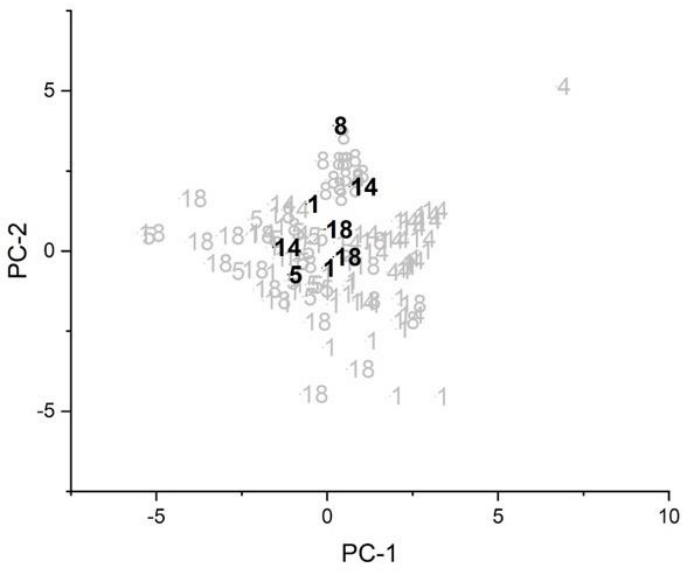

Figure 5.200 PC plot for the fifteenth training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

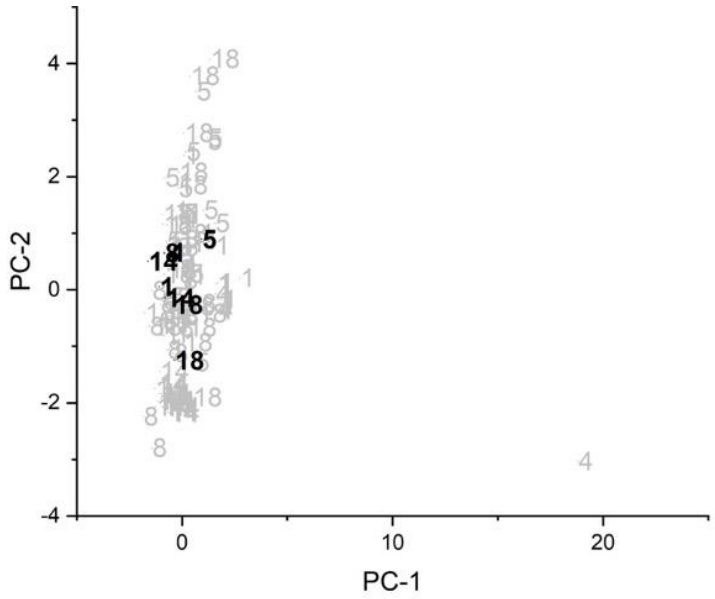

Figure 5.201 PC plot for the fifteenth training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

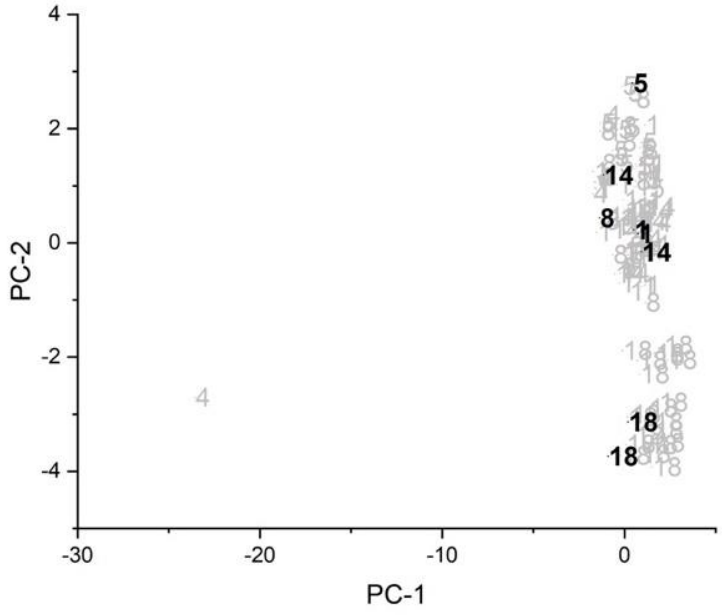

Figure 5.202 PC plot for the fifteenth training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

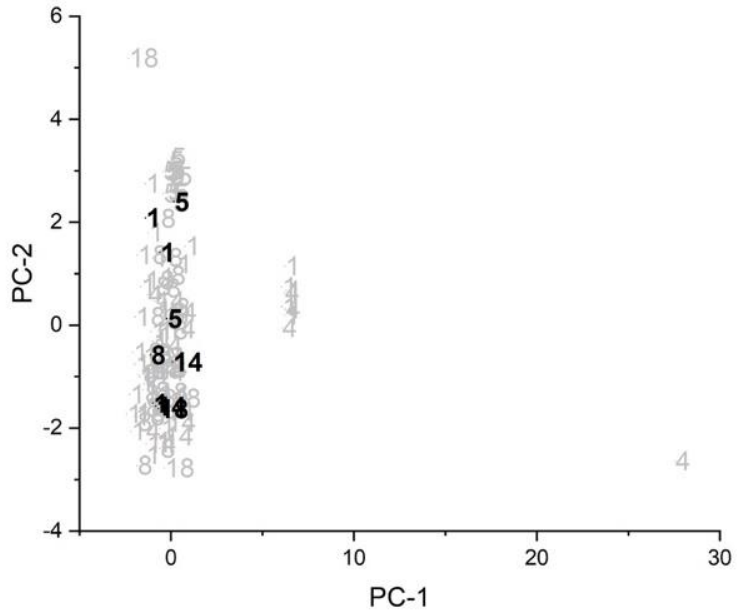

Figure 5.203 PC plot for the sixteenth training set/prediction set pair: Fairfax vs Arlington, Doraville, Fort Wayne, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

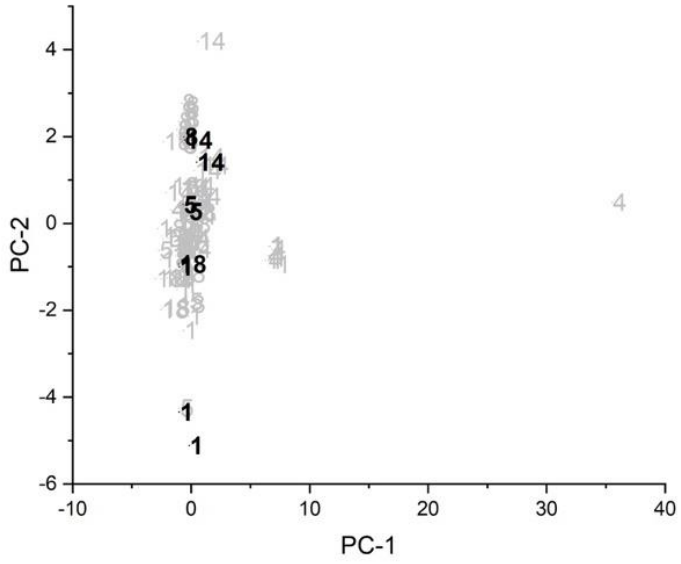

Figure 5.204 PC plot for the sixteenth training set/prediction set pair: Fort Wayne vs Arlington, Doraville, Fairfax, Lansing, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

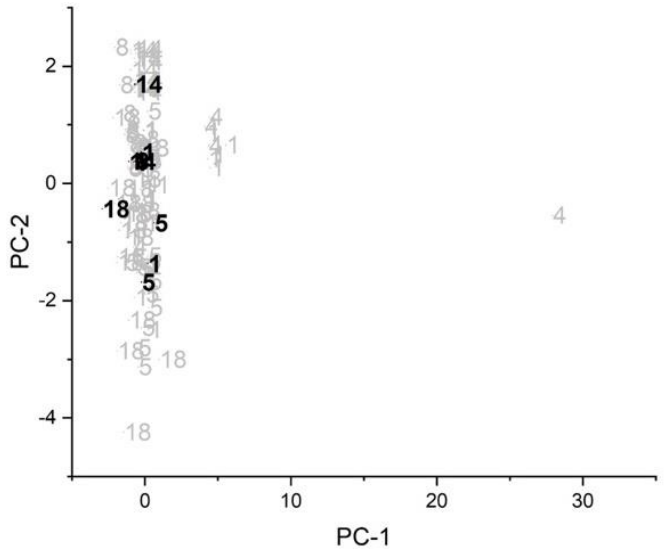

Figure 5.205 PC plot for the sixteenth training set/prediction set pair: Lansing vs Arlington, Doraville, Fairfax, Fort Wayne, and Moraine. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

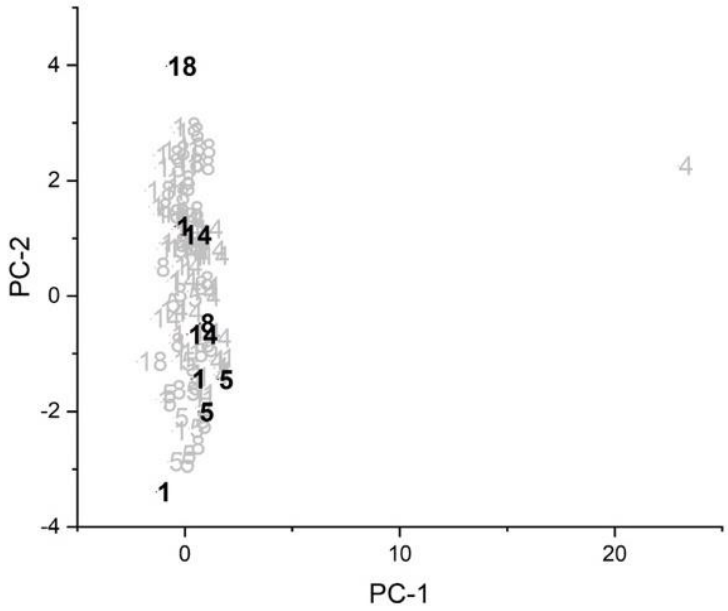

Figure 5.206 PC plot for the sixteenth training set/prediction set pair: Moraine vs Arlington, Doraville, Fairfax, Fort Wayne, and Lansing. Prediction set (black) projected onto the PC plot developed from the training set samples (grey) and the wavelet coefficients identified by the pattern recognition GA. 1 = Arlington, 4 = Doraville, 5 = Fairfax, 8 = Fort Wayne, 14 = Lansing and 18 = Moraine.

References

1. B. K. Lavine, N. Mirjankar, S., and M. Sandercock, "Wavelets and Genetic Algorithms Applied to Search Prefilters for Spectral Library Matching in Forensics" *Talanta*, **2011**, *87*, 46-52.

2. B. K. Lavine, A. Fasasi, N. Mirjankar, and M. Sandercock, "Development of Search Prefilters for Infrared Library Searching of Clear Coat Paint Smears" *Talanta*, **2014**, *119*, 331–340.

3. B. K. Lavine, A. Fasasi, N. Mirjankar, and C. White, "Search Prefilters for Library Matching of Infrared Spectra in the PDQ Database using the Autocorrelation Transformation" *Microchem. J*., **2014**, *113*, 30–35.

4. B. K. Lavine, A. Fasasi, N. Mirjankar, M. Sandercock, and S. D. Brown, "Search Prefilters for Mid-IR Spectra of Clear Coat Automotive Paint Smears Using Stacked and Linear Classifiers" *J. of Chemom*., **2014**, *28*, 385-394.

5. B. K. Lavine, A. Fasasi, N. Mirjankar, C. White, and M. Sandercock, "Search Prefilters to Assist in Library Searching of Infrared Spectra of Automotive Clear Coats" *Talanta*, **2015**, *132*, 182-190.

6. A. Fasasi, N. Mirjankar, R.-I. Stoian, C. White, M. Allen, M. P. Sandercock and B. K. Lavine, "Pattern Recognition Assisted Infrared Library Searching of Automotive Clear Coats" *Appl. Spec*., **2015**, *69* (1) 84-94.

7. U. D. N. Perera, K. Nishikida, and B. K. Lavine, "Development of Infrared Library Search Prefilters for Automotive Clear Coats from Simulated ATR Spectra" *Appl. Spec*., **2018**, *186*, 662-669.

8. J. De Gelder, P. Vandenabeele, F. Govaert, and L. Moens, "Forensic Analysis of Automotive Paints by Raman Spectroscopy" *J. Raman Spec*., **2005**, *36*, 1059-1067.

9. J. He, J. Lv, Y. Ji, J. Feng, and Y. Liu, "Multiple Characterizations of Automotive Coatings in Forensic Analysis," *Spec. Lett*., **2013**, *46*, 555-560.

10. M. Maric, W. van Bronswijk, K. Pitts, and S. W. Lewis, "Characterization and Classification of Automotive Clear Coats with Raman Spectroscopy and Chemometrics for Forensic Purposes" *J. Raman Spec*., **2016**, *47* (8), 948-955.

11. G. Sauzier, M. Maric, W. van Bronswijk and S. W. Lewis, "Preliminary studies into the effect of environmental degradation on the characterisation of automotive clear coats by attenuated total reflectance infrared spectroscopy" *Anal. Methods*., **2013**, 5, 4984 4990.

12. K. B. Ferreira, A. G. G. Oliveira and J. A. Gomes, "Raman spectroscopy of automotive paints: Forensic analysis of variability and spectral quality" *Spec. Lett*., **2017**, *50* (2), 102–110.

13. J. S. Walker, "Primer on Wavelets and Their Scientific Applications" Chapman & Hall/CRC Press, Boca Raton, FL 1999.

14. M. Meloun, J. Militky, and M. Forina, "Chemometrics for analytical chemistry. volume 1: PC aided statistical data analysis" Ellis Horwood, NY, 1992, pp. 138-139.

15. B.G.M. Vandeginste, D. L. Massart, L. M. C. Buydens, S. DeJong, P. J. Lewi, and J. Smeyers-Verbeke, "Handbook of Chemometrics and Qualimetrics" Elsevier, Amsterdam, 1998, pp. 310.

16. M. A. Stapanian, F. C. Garner, K. E. Fitzgerald, G. T. Flatman, and J. M. Nocerino, "Finding suspected causes of measurement error in multivariate environmental data" *J. Chemom*., **1993**, *7*, 165-176.

17. B. K. Lavine, K. Nuguru, and N. Mirjankar, "One Stop Shopping - Feature Selection, Classification, and Prediction in a Single Step" *J. Chemom*., **2011**, *25*, 116-129.

18. J. M. Stone, "Cross Validation Choice and Assessment of Statistical Predictions" *J. Royal Statist*. Soc., **1977**, *B39*, 44-47.

19. S. G. Ryland and E. Suzucki, "Analysis of Paint Evidence" in Chemistry Forensic Handbook, John Wiley & Sons, NY, 2011, 131-224.

20. H. A. Willis, V. I. J. Zichy, and P. J. Hendra, "The Laser-Raman and Infra-red Spectra of Poly (Methyl Methacrylate)" *Polymer*, **1969**, *10*, 737-746.

CHAPTER VI

CONCLUSION

The research described in this dissertation is directly targeted to develop new approaches for the forensic analysis of automotive clear coats and paint smears. Forensic automotive paint analysis is common in all crime laboratories. The research described in this dissertation will ensure that forensic automotive paint databases such as PDQ remain current with respect to modern automotive paints. A methodology to generate paint smears in a realistic manner is a useful addition to forensic laboratories because car paint evidence is often obtained from collisions that result in physical smearing. The development of forensic standards for automotive paint transfer and paint smear analysis can significantly improve the accuracy of paint smear analysis and make possible the creation of realistic proficiency testing for forensic laboratories, which could also be used to train forensic scientists who are using automotive paint databases such as PDQ.

The use of machine learning methods including multivariate curve resolution to assess paint smears will aid in evidential significance assessment, both at the investigative lead stage and at the courtroom testimony stage.

Direct impact on over 57 local, state, and federal forensic laboratories currently using forensic automotive paint database in the United States is anticipated. There will also be direct impact on international forensic laboratories including the Forensic Laboratory Services Division of the RCMP, the Centre of Forensic Sciences in Toronto, Canada, the ENFSI network of European forensic science institutes, the Australian Police Services, and the New Zealand Police Services. Since 2014, the Bundeskriminalamt (BKA) has started a European Automotive Paint Collection to analyze automotive clear coats by Raman spectroscopy. Currently, the BKA is systematically collecting Raman spectra of automotive clear coats. The BKA has not yet performed an evaluation of the forensic value of their Raman database. The Raman clear coat study described in this dissertation is timely and addresses important questions crucial for assessing the forensic value of Raman spectroscopy for the analysis of automotive paints.

The dissertation research is an international collaborative effort. Pioneering studies previously performed by Lavine and coworkers at Oklahoma State University and Mark Sandercock of the RCMP on Chrysler and General Motors clear coats spanning a limited production year range have demonstrated the advantages of using pattern recognition approaches to extract investigative lead information from IR spectra of clear coats, which all too often is the only layer of paint left at the crime scene. The proposed methodology is a significant improvement over the way clear coats and paint smears are currently handled by forensic paint examiners.

An added advantage of pattern recognition-based approaches to identify automotive paint samples such as those investigated in this dissertation research is an increase in accuracy because objective criteria are substituted for subjective human input

and interpretation in the extraction of investigative lead information from vibrational spectra. Information derived from these pattern recognition searches will allow forensic scientists to quantify the general discrimination power of original automotive paint comparisons encountered in casework. This, in turn, will improve the forensic scientist's ability to succinctly communicate trace evidential significance to the courts. Furthermore, the Raman studies of clear coats and the development of a procedure to investigate paint smears as a result of paint transfer from an automotive substrate to a variety of surfaces under conditions simulating those of a real collision will be of significant interest to the wider scientific community. For the forensic community, the proposed methodology for the development and analysis of paint smears will ensure that PDQ and other automotive paint databases will remain current for forensic science casework.

VITA

George Paa Kwesi Affadu-Danful

Candidate for the Degree of

Doctor of Philosophy

Dissertation: APPLICATION OF RAMAN AND INFRARED MICROSCOPY COUPLED WITH CHEMOMETRICS FOR THE FORENSIC EXAMINAITON OF AUTOMOTIVE CLEAR COATS AND PAINT SMEARS

Major Field: Chemistry

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Chemistry at Oklahoma State University, Stillwater, Oklahoma in July, 2022.

Completed the requirements for the Master of Science in your Chemistry at East Tennessee State University, Johnson City, Tennessee, USA in 2018.

Completed the requirements for the Bachelor of Science in Chemistry major at University of Cape Coast, Cape Coast, Ghana in 2015.

Experience:

Graduate Research Assistant and Graduate Teaching Assistant at Oklahoma State University from August 2018 to July 2022.

Graduate Research Assistant and Teaching Assistant at East Tennessee State University from August 2016 to July 2018.

Teaching Assistant at University of Cape Coast from September 2015 to August 2016.