


Spring 2013

# Novel Algorithms and Instrumentation for Vibrational Spectroscopic Methods of Analysis

Mohamed F. Abdelkader  
*Old Dominion University*

Follow this and additional works at: [https://digitalcommons.odu.edu/chemistry\\_etds](https://digitalcommons.odu.edu/chemistry_etds)

 Part of the [Applied Mathematics Commons](#), [Mathematics Commons](#), and the [Physical Chemistry Commons](#)

---

## Recommended Citation

Abdelkader, Mohamed F.. "Novel Algorithms and Instrumentation for Vibrational Spectroscopic Methods of Analysis" (2013). Doctor of Philosophy (PhD), dissertation, Chemistry and Biochemistry, Old Dominion University, DOI: 10.25777/sjy0-xz25  
[https://digitalcommons.odu.edu/chemistry\\_etds/29](https://digitalcommons.odu.edu/chemistry_etds/29)

This Dissertation is brought to you for free and open access by the Chemistry & Biochemistry at ODU Digital Commons. It has been accepted for inclusion in Chemistry & Biochemistry Theses & Dissertations by an authorized administrator of ODU Digital Commons. For more information, please contact [digitalcommons@odu.edu](mailto:digitalcommons@odu.edu).

NOVEL ALGORITHMS AND INSTRUMENTATION FOR VIBRATIONAL  
SPECTROSCOPIC METHODS OF ANALYSIS

by

Mohamed F. Abdelkader  
B.S. May 2007, Alexandria University, Egypt

A Dissertation Submitted to the Faculty of  
Old Dominion University in Partial Fulfillment of the  
Requirements for the Degree of

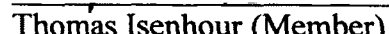
DOCTOR OF PHILOSOPHY

CHEMISTRY

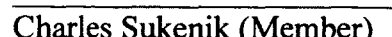
OLD DOMINION UNIVERSITY  
May 2013

Approved by:

  
John B. Cooper (Director)

  
Thomas Isenhour (Member)

  
Richard Gregory (Member)

  
Charles Sukenik (Member)

## **ABSTRACT**

### **NOVEL ALGORITHMS AND INSTRUMENTATION FOR VIBRATIONAL SPECTROSCOPIC METHODS OF ANALYSIS**

Mohamed F. Abdelkader  
Old Dominion University, 2013  
Director: Dr. John B. Cooper

Raman spectroscopy is a form of vibrational spectroscopy that has been increasingly applied to qualitative analysis of chemicals, explosives, pharmaceuticals, and fuels due to its non-invasive and non-destructive nature; its ease of sampling; and its high molecular specificity. These characteristics of Raman spectroscopy also facilitate its use for both in-line and at-line analysis. The principle limitation of Raman spectroscopy is optical interference arising from both analyte and non-analyte fluorescence. In this dissertation, a solution to this problem is presented in the form of a novel spectrometer design which operates in a sequentially shifted excitation mode to eliminate fluorescence backgrounds, fixed pattern noise, and room lights, while keeping the Raman data in true spectral space. The Raman data are extracted from the shifted excitation spectra using a novel algorithm which is three orders of magnitude faster than conventional iterative algorithms.

Near infrared spectroscopy is another form of vibrational spectroscopy which has also been increasingly used for the analysis of fuels in both at-line and in-line applications. Despite its popularity, near infrared spectroscopy lacks the spectral resolution of Raman spectroscopy and because of this, the transfer of quantitative calibration models is considered extremely difficult and expensive. This is considered one of the principal limitations in the use of near infrared spectroscopy for fuel analysis. In this dissertation, a solution to this problem is presented in the form of novel calibration transfer algorithms which allow the use of virtual fuels which are digitally synthesized from pure chemical standards and to transfer quantitative calibration models between different near infrared instruments. This solution eliminates the expense, time, and difficulty of traditional calibration transfer methods.

This dissertation is dedicated to my father who could not read or write but always motivated me to be educated and gave me all the support needed.

## ACKNOWLEDGMENTS

My first debt of gratitude goes to my advisor, Dr. John B. Cooper. He patiently provided the vision, encouragement and advice necessary for me to proceed through the program and complete my dissertation. I want to thank him for being a great advisor, a father, and a close friend. He has been a strong and supportive advisor to me throughout my graduate school career, but he has always given me great freedom to pursue independent work. Special thanks to my committee, Dr. Richard Gregory, Dr. Charles Sukenik, and Dr. Thomas Isenhour for their support, guidance, helpful suggestions, and revision to this final manuscript.

In addition, a thank you to Mrs. Alicia Herr, who always helped me with love and passion, gave me car rides, and walked with me around the campus when I first came.

A special thank you goes to Rohayma Rateb (EducationUSA, Alexandria Amideast) for her support, advising, and guidance. Indeed, without her help and guidance, I could not have come to the states.

## TABLE OF CONTENTS

	Page
LIST OF TABLES .....	viii
LIST OF FIGURES .....	x
 Chapter	
I. INTRODUCTION .....	1
RAMAN THEORY .....	2
SEQUENTIALLY SHIFTED EXCITATION RAMAN.....	7
NIR THEORY .....	9
NIR ABSORBANCES .....	14
OBJECTIVES .....	16
 II. NOVEL ALGORITHM AND INSTRUMENTATION FOR FLUORESCENCE- FREE RAMAN SPECTROSCOPY UTILIZING A SINGLE TUNABLE EXCITATION SOURCE.....	17
INTRODUCTION.....	17
EXPERIMENTAL SECTION.....	20
RESULTS AND DISCUSSION.....	36
INSTRUMENT CALIBRATION.....	56
 III. NOVEL ALGORITHM AND INSTRUMENTATION FOR FLUORESCENCE- FREE RAMAN SPECTROSCOPY UTILIZING DUAL-TUNABLE EXCITATION SOURCES.....	66
INTRODUCTION.....	66
EXPERIMENTAL SECTION.....	69
RESULTS AND DISCUSSION.....	84
INSTRUMENT CALIBRATION.....	100
 IV. CALIBRATION TRANSFER OF PLS MODELS USING VIRTUAL STANDARDS SLOPE AND BIAS (VSSB).....	109
INTRODUCTION.....	109
EXPERIMENTAL SECTION.....	112
VIRTUAL STANDARD CALIBRATION TRANSFER THEORY .....	115

RESULTS AND DISCUSSION.....	117
V. CALIBRATION TRANSFER OF PLS MODELS USING SEGMENTED VIRTUAL STANDARDS SLOPE AND BIAS (SVSSB).....	139
INTRODUCTION.....	139
EXPERIMENTAL SECTION.....	141
SEGMENTED VIRTUAL STANDARD CALIBRATION TRANSFER THEORY.....	144
RESULTS AND DISCUSSION.....	149
VI. CONCLUSIONS .....	165
ELIMINATING RAMAN OPTICAL INTERFERENCES VIRTUAL STANDARD CALIBRATION TRANSFER.....	166
REFERENCES .....	167
APPENDICES	
A. ABBREVIATIONS.....	176
B. NIR SPECTRA OF THE FIFTEEN SOLVENTS USED IN THE SEGMENTED VIRTUAL STANDARD CALIBRATION TRANSFER.....	177
VITA.....	185

## LIST OF TABLES

Table	Page
1.	NIR absorbance for C-H groups of the second overtone region.....14
2.	The SSE pixel axis of the sixteen acetaminophen peaks used in SSE axis calibration along with their corresponding ASTM Raman shift in wavelength (nm).....58
3.	The standard deviation of Raman peaks of acetaminophen at different temperature of the 852 nm and 785 nm diode lasers.....85
4.	Acetaminophen SSE Raman peaks in pixels with their ASTM Raman shift in wavelength for both the 852 and 785 nm excitations lasers.....102
5.	ASTM Raman shift of acetaminophen in wavenumber units along with their ASTM and SSE standard deviation error.....107
6.	ASTM Raman shift of polystyrene peaks in wavenumber with their standard deviation error compared to their SSE Raman shift in wavenumber.....108
7.	Jet fuels properties, ASTM methods, units, ASTM range, training set range, and PER.....118
8.	Summary of the jet fuel PLS models and properties used in VSSB.....120
9.	The standard errors of prediction of properties in the test set samples relative to the actual values for Uncorrected, VSSB Corrected, and PDS Corrected Models.....121
10.	The percentages of the twelve solvents used to digitally synthesize target spectra .....131
11.	The standard errors of prediction of properties in the test set samples relative to the master prediction for uncorrected, VSSB corrected, and PDS corrected.....132



12. Summary of the jet fuel PLS models and properties used in SVSSB calibration transfer.....149
13. The standard error of prediction of the properties in the test set samples relative to the master instrument prediction for uncorrected, SVSSB corrected, BS corrected, and PDS corrected .....153
14. The standard error of prediction of the properties in the test set samples relative to the master instrument prediction for uncorrected and SVSSB corrected.....161

## LIST OF FIGURES

Figure	Page
1. Different types of light scattering; Rayleigh, Raman Stokes, and Raman anti-Stokes.....	4
2. Raman Spectrum of CCl <sub>4</sub> showing Stokes and anti-Stokes lines acquired utilizing 448nm excitation.....	5
3. A virtual energy state overlay with an excited electronic state. This leads to fluorescence emission which plagues Raman signals.....	6
4. The effect of changing the excitation source wavelength on the Raman and fluorescence background.....	8
5. The harmonic model for potential energy of a diatomic molecule.....	10
6. Side-by-side comparison of the harmonic and anharmonic model for potential energy of a diatomic molecule. ....	11
7. The NIR spectra of a) toluene b) isooctane and c) nonane.....	15
8. The single-laser SSE Raman spectrometer design and components.....	21
9. The 785nm distributed Bragg reflector GaAs diode laser diagram within a custom TO-8 package.....	22
10. The sequences of operation of the single-laser SSE Raman.....	24
11. The overall method of processing the sequentially shifted excitation Raman data to obtain a background-free Raman spectrum .....	32
12. A detailed explanation of the SSE algorithm in matrix notation.....	33

13. A numerical example of extracting a back-ground free Raman spectrum from the shifted excitation Raman spectra with graphical depictions of initial, intermediate, and final values.....34
14. The plot between acetaminophen Raman shift and the diode laser temperature.....37
15. Dimethyle glyoxime Raman spectra acquired on the SSE Raman before and after applying SNV.....38
16. Raman spectra of dimethyl glyoxime acquired at four laser temperatures.....39
17. Raman spectra of dimethyl glyoxime extracted from shifted excitation Raman spectra using various numbers of iterations according to the previously described SSE algorithm.....41
18. Comparison of the FT-Raman spectrum to the SSE Raman spectrum of dimethyl glyoxime.....42
19. Comparison of the SSE Raman spectrum, the FT-Raman spectrum, and the dispersive Raman spectrum of catechol.....43
20. Comparison of the SSE Raman spectrum, the FT-Raman spectrum, and the dispersive Raman spectrum of magnesium carbonate.....44
21. The Raman spectra of the raw unprocessed and the SSE Raman spectra of a diesel fuel .....46
22. Comparison of the SSE Raman spectrum, the FT-Raman spectrum, and the dispersive Raman spectrum of ceric ammonium sulfate.....48
23. Comparison of the dispersive Raman spectroscopy, SERDS, and the FT-Raman spectroscopy to the SSE Raman of acenaphthylene.....50
24. Comparison of the dispersive Raman spectroscopy, SERDS, and the FT-Raman spectroscopy to the SSE Raman of 4-Bromo-N, N-dimethyl-aniline .....52
25. Performance of the SSE and the PCA algorithms at two different temperature

profiles of the diode laser.....	53
26. Overlay of acetaminophen Raman spectra acquired on the single-laser SSE before axis calibration and on the FT-spectrometer .....	56
27. The sixteen peaks of acetaminophen used in the SSE axis calibration.....	57
28. The third order polynomial used to generated the single-laser SSE axis.....	59
29. Overlay of acetaminophen Raman spectra acquired on the single-laser SSE after axis calibration and on the FT-spectrometer.....	60
30. The standard luminescence spectrum of the SRM generated using the five polynomial coefficients provided by NIST.....	62
31. The luminescence spect rum of the SRM 2241 acquired on the single-laser SSE Raman.....	63
32. Overlay of acetaminophen Raman spectra acquired on the single-laser SSE before and after intensity calibration.....	65
33. The dual-laser SSE Raman spectrometer design and components.....	70
34. The 852 nm distributed Bragg reflector (DBR) GaAs diode laser displayed with SEM micrographs .....	71
35. The sequences of operations of the dual-laser SSE Raman.....	75
36. The overall method of processing the shifted excitation Raman spectra acquired on the dual-laser SSE Raman to obtain a background-free single-merged Raman spectrum.....	79
37. The 785 and 852 nm excitation sources and their corresponding Ram an signal regions shown in an absolute wavelength axis.....	80
38. The 785 and 852 nm excitation sources and their corresponding Ram an spectra of acetaminophen shown in an absolute wavelength axis.....	81

39.	The 785 and 852 nm excitation sources and their corresponding Raman spectra of acetaminophen shown in an absolute wavenumber ( $\text{cm}^{-1}$ ) axis.....	82
40.	The 785 and 852nm Raman spectra of acetaminophen shown in a Raman shift wavenumber ( $\text{cm}^{-1}$ ) units.....	83
41.	The 785 and 852 merged acetaminophen Raman spectrum.....	83
42.	The plot between acetaminophen Raman shift and the dual-diode lasers temperature.....	85
43.	Raman spectra of 2,3-naphthalenediamine acquired at four laser temperatures for both the 785 and 852nm lasers.....	87
44.	The raw Raman spectrum and the SSE Raman spectrum of 2,3-naphthalenediamine acquired on the single-laser SSE .....	89
45.	The raw and the SSE Raman spectra of xanthan gum acquired on both the single-laser and the dual-laser SSE spectrometer.....	91
46.	The raw and the SSE Raman spectra of sodium carboxy methyl cellulose acquired on both the single-laser and the dual-laser SSE spectrometer.....	93
47.	The raw and the SSE Raman spectra of niacinamide acquired on both the single-laser and the dual-laser SSE spectrometer.....	95
48.	The raw Raman spectra of microcrystal cellulose acquired on the dual-laser SSE utilizing 785 nm excitation source and 852 nm excitation source.....	97
49.	The SSE Raman spectra of microcrystal cellulose acquired on the dual-laser SSE at six different temperature profiles.....	100
50.	The acetaminophen peaks used in the dual-laser axis calibration .....	101
51.	The two third order polynomials used to generate the dual-laser SSE axes.....	104

52. The 785nm and the 852nm acetaminophen spectra acquired on the dual-laser SSE Raman after axis calibration and both are overlaid with the FT-acetaminophen spectrum .....106
53. The 785nm and the 852nm polystyrene Raman spectra acquired on the dual-laser SSE Raman after axis calibration.....108
54. Overlay of three NIR spectra of jet fuel with maximum, median, and minimum cetane index value before and after applying a seven-points Savitzky-Golay first derivative .....117
55. Overlay of the first derivative SNV NIR spectra of the twelve solvents acquired on the master and the slave instruments .....122
56. NIR absorbance spectrum of a digitally synthesized virtual standard overlaid with its target fuel spectrum.....123
57. API gravity predictions of the test set fuels acquired on the FUElex slave instrument for both uncorrected and VSSB corrected.....125
58. Density, percent aromatic, cetane index, and percent distillation recovery predictions of the test set fuels acquired on the FUElex slave instrument for both uncorrected and VSSB corrected.....126
59. Flash point, 20, 50, and 90 percent distillation recovery predictions of the test set fuels acquired on the FUElex slave instrument for both uncorrected and VSSB corrected.....127
60. Freeze point, percent saturates, percent hydrogen content, and viscosity predictions of the test set fuels acquired on the FUElex slave instrument for both uncorrected and VSSB corrected.....128
61. The NIR spectra of a JP-8 fuel acquired on both the MPA and the FUElex NIR instruments.....129
62. Overlay of NIR absorbance spectra of digitally synthesized MPA virtual standard with a digitally synthesized FUElex virtual standard and their target NIR

spectra.....	131
63. API gravity, percent aromatic, and cetane index predictions of the test set fuels acquired on the FUElex slave instrument for both uncorrected, VSSB corrected, and PDS corrected relative to the MPA instrument.....	134
64. Density, 10, and 20 percent distillation recovery predictions of the test set fuels acquired on the FUElex slave instrument for both uncorrected, VSSB corrected, and PDS corrected relative to the MPA instrument.....	135
65. Flash point, freeze point, and 50 percent distillation recovery predictions of the test set fuels acquired on the FUElex slave instrument for both uncorrected, VSSB corrected, and PDS corrected relative to the MPA instrument.....	136
66. Viscosity, percent hydrogen content, and percent saturates predictions of the test set fuels acquired on the FUElex slave instrument for both uncorrected, VSSB corrected, and PDS corrected relative to the MPA instrument.....	137
67. A schematic of the segmented virtual standard calibration transfer .....	147
68. Overlay of the first derivative SNV NIR spectra of the fifteen molecular chemicals acquired on the master and the slave instruments.....	151
69. The NIR absorbance spectrum of a digitally synthesized virtual standard using segment target overlaid with digitally synthesized virtual standard using full target and their target fuel spectrum.....	152
70. API gravity, percent aromatic, and cetane index predictions of the test set fuels acquired on the FUElex slave instrument for both uncorrected, SVSSB corrected, and PDS corrected.....	155
71. Density, 10, and 20 percent distillation recovery predictions of the test set fuels acquired on the FUElex slave instrument for both uncorrected, SVSSB corrected, and PDS corrected. ....	156
72. Flash point, freeze point, and 90 percent distillation recovery predictions of the test set fuels acquired on the FUElex slave instrument for both uncorrected, SVSSB corrected, and PDS corrected.....	157

73. Viscosity, percent hydrogen content, and percent saturates predictions of the test set fuels acquired on the FUElex slave instrument for both uncorrected, SVSSB corrected, and PDS corrected.....158
74. The NIR spectra of a JP-8 fuel acquired on both the MPA and the FUElex NIR instruments.....159
75. NIR absorbance spectra of a digitally synthesized MPA secondary instrument virtual standard overlaid with a digitally synthesized FUELex virtual standard...160
76. API gravity predictions of the test set fuels acquired on the MPA secondary instrument for uncorrected, SVSSB corrected and BS corrected.....162
77. Cetane index predictions of the test set fuels acquired on the MPA secondary instrument for uncorrected, SVSSB corrected and BS corrected .....163
78. Percent hydrogen content predictions of the test set fuels acquired on the MPA secondary instrument for uncorrected, SVSSB corrected and BS corrected.....164



## **CHAPTER I**

### **INTRODUCTION**

Raman spectroscopy is a form of vibrational spectroscopy that has been increasingly applied to qualitative analysis of chemicals, explosives, pharmaceuticals, and fuels. The facts that Raman spectroscopy 1) can be sampled through a glass or plastic container, 2) provides specific molecular information, and 3) requires no sample preparations make it a popular technique for both in-line and at-line analysis. However, the huge optical interferences due to both analyte and non-analyte fluorescence limit the practical use of Raman spectroscopy. In this dissertation, a solution to this problem is presented. Basic understanding of Raman spectroscopy and Raman theory is presented in this chapter in order to better understand the solution presented in this dissertation.

Near infrared (NIR) spectroscopy is another form of vibrational spectroscopy which has also been increasingly used for the analysis of fuels in both at-line and in-line applications. Despite its popularity, NIR spectroscopy lacks the spectral resolution of Raman spectroscopy and because of this, the transfer of quantitative calibration models is considered extremely difficult and expensive. For fuel analysis, calibration transfer is a particularly difficult problem since maintaining the compositional integrity of fuel standards which often contain thousands of volatile chemicals is difficult at best. This problem is compounded by the need to both store these standards and to transport them to geographically diverse locations. To avoid these issues, the most common solution is to acquire a new spectral database of available fuels on a new instrument and build a new multivariate model specific for that instrument. In this dissertation a novel calibration transfer algorithm is presented which allows transfer of calibration models between different NIR instruments using only pure chemicals which can be obtained universally at high purity levels. To better understand this algorithmic approach, an overview of NIR spectroscopy and vibrational theory are presented in this chapter.

## RAMAN THEORY

When a sample is irradiated by an intense laser beam, that has frequency  $\nu_0$ , the radiation interacts with the sample's electrons and the radiation is scattered in all directions. The scattered light consists of two types: one is called Rayleigh scattering, and the other is called Raman scattering. Rayleigh scattering has the same frequency as the incident beam ( $\nu_0$ ), while Raman scattering has frequencies  $\nu_0 \pm \nu_{\text{vib}}$ , where  $\nu_{\text{vib}}$  is a vibrational frequency of a molecule. The  $\nu_0 - \nu_{\text{vib}}$  and  $\nu_0 + \nu_{\text{vib}}$  frequencies are called Stokes and anti-Stokes Raman, respectively. Thus, in Raman spectroscopy, we measure the vibrational frequency ( $\nu_{\text{vib}}$ ) as a shift from the incident beam frequency ( $\nu_0$ ). The intensity of Raman scattering is around six orders of magnitude less than the intensity of the incident beam due to the low probability of Raman scattering (the Raman cross-section). For photons which interact with a molecule, there is typically one Raman scattered photon for each million Rayleigh scattered photons. The interaction of a laser beam with a molecule polarizes its electrons to a virtual state, as shown in Figure 1. This can be understood by applying classic wave theory to a diatomic molecule. The electrical field strength ( $E$ ) of the laser beam (electromagnetic wave) fluctuates with time ( $t$ ) as given in Eqn. 1:

$$E = E_0 \cos 2\pi\nu_0 t \quad (1)$$

where  $\nu_0$  is the frequency of the laser beam and  $E_0$  is the amplitude of the laser beam. When a diatomic molecule is irradiated by the laser beam, an electrical dipole moment is induced due to the interaction of laser beam with the bonding electrons of the molecule. The induced electrical dipole moment ( $P$ ) is given by Eqn. 2:

$$P = \alpha E = \alpha E_0 \cos 2\pi\nu_0 t \quad (2)$$

where  $\alpha$  is the polarizability of the bonding electrons. When the diatomic molecule is vibrating with a frequency  $\nu_{\text{vib}}$ , its nuclear displacement ( $x$ ) is given by Eqn. 3:

$$x = x_0 \cos 2\pi\nu_{\text{vib}} t \quad (3)$$

where  $x_0$  is the equilibrium bond length of the diatomic molecule undergoing vibration. For all different vibrational amplitudes,  $\alpha$  can be represented as a linear function of  $x$  using Eqn. 4:

$$\alpha = \alpha_0 + \left(\frac{\partial\alpha}{\partial x}\right)_0 x + \dots \quad (4)$$

where  $\alpha_0$  is the polarizability at the equilibrium position, and  $(\partial\alpha/\partial x)_0$  is the rate of the change of  $\alpha$  with respect to the change in  $x$ . By substituting Eqn. 4 for the value of  $\alpha$  in Eqn. 2,  $P$  can be expressed as:

$$P = \alpha_0 E_0 \cos 2\pi\nu_0 t + \left(\frac{\partial\alpha}{\partial x}\right)_0 x E_0 \cos 2\pi\nu_0 t ,$$

and subsequent substitution of  $x$  using Eqn. 3 leads to:

$$P = \alpha_0 E_0 \cos 2\pi\nu_0 t + \left(\frac{\partial\alpha}{\partial x}\right)_0 x_0 E_0 \cos 2\pi\nu_0 t \cos 2\pi\nu_{vib} t ,$$

which can be simplified using the trigonometric identity,

$\cos(a)\cos(b) = 1/2 \{ \cos(a+b) + \cos(a-b) \}$ , to give:

$$P = \alpha_0 E_0 \cos 2\pi\nu_0 t + \frac{1}{2} \left(\frac{\partial\alpha}{\partial x}\right)_0 x_0 E_0 [ \cos\{2\pi(\nu_0 + \nu_{vib})t\} + \cos\{2\pi(\nu_0 - \nu_{vib})t\} ] \quad (5)$$

according to Eqn. 5, the first term represents a vibrating dipole which radiates light of frequency  $\nu_0$  which gives Rayleigh scattering. The second term corresponds to the scattering of frequency  $\nu_0 + \nu_{vib}$  which gives Raman Stokes scattering, and the third term ( $\nu_0 - \nu_{vib}$ ) gives the anti-Stokes Raman scattering. For Raman vibrations to be active the term  $(\partial\alpha/\partial x)_0$  has to be greater than zero. Practically, this means that the vibration must give rise to a change in polarizability. In practice this usually means that bonds which are not already highly polarized will give rise to strong Raman vibrations since the electromagnetic wave can easily polarize them giving rise to a change in polarizability. As an example, the Raman vibrations of alkanes (a main constituent in fuels) are relatively intense since the bonded carbon atoms in the chain have the same electronegativities. For this reason, the C-C bonds are not polarized in the ground state and are therefore easy to polarize by the electromagnetic radiation in a virtual state.

Figure 1 illustrates the different types of scattering: Rayleigh, Stokes Raman, and anti-Stokes Raman, where the interaction of laser beam with the diatomic molecule's electrons leads to the polarization of the electrons to one of many possible virtual energy states. This state is referred to as a virtual state since the distortion of the electrons by the electromagnetic field is not a natural excited state of the molecule. The polarized electrons relax releasing the same frequency  $\nu_0$ , by returning to the same vibrational level from which they were distorted to give Rayleigh scattering. Alternatively they can return to a different vibrational level releasing frequencies  $\nu_0 + \nu_{\text{vib}}$  and  $\nu_0 - \nu_{\text{vib}}$  which give Stokes Raman, and anti-Stokes Raman, respectively. Note that for Stokes Raman to occur, the molecule must originate from a ground vibrational level and return to an excited vibrational level. Conversely, for anti-Stokes Raman to occur, the molecule must originate in an excited vibrational level and return to the ground vibrational level. Since anti-Stokes Raman requires the molecule to initially be in an excited vibrational state, the intensity of the anti-Stokes Raman spectrum is governed not only by the Raman cross-section, but also by the Boltzmann distribution for the vibrational energy levels. For this reason, the anti-Stokes Raman is temperature dependent while the Stokes Raman is temperature independent.

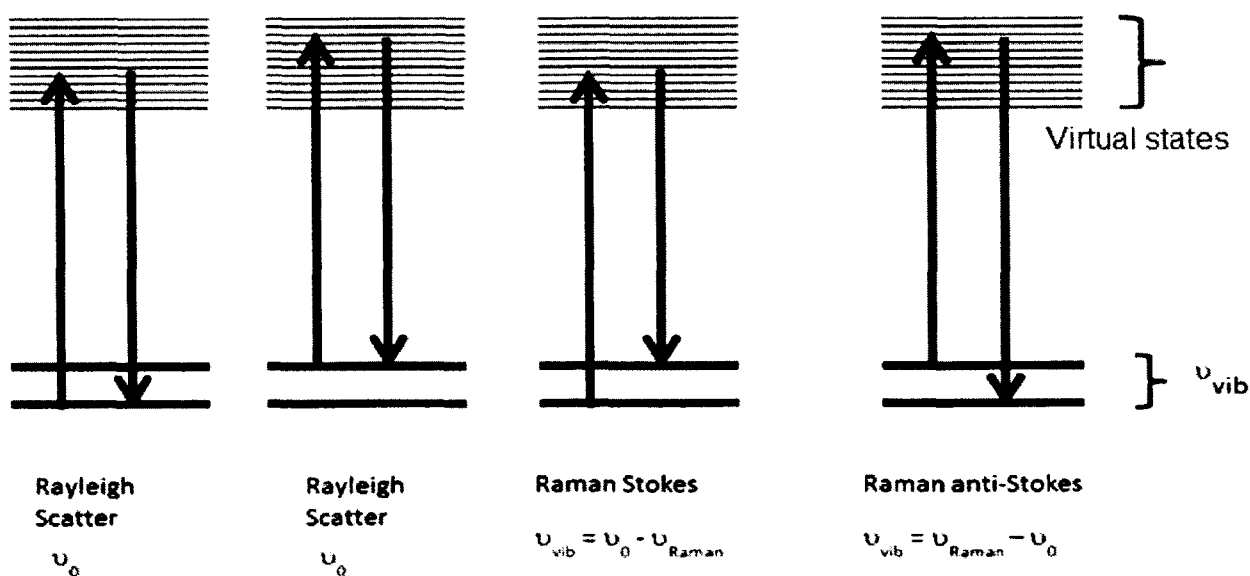


FIG. 1. Different types of light scattering; Rayleigh, Raman Stokes, and Raman anti-Stokes.

Figure 2 shows the Raman spectrum of  $\text{CCl}_4$ . As can be noticed, the Stokes lines have higher intensity than anti-Stokes, this is because the population of molecules at  $\nu = 0$  is much higher than at  $\nu = 1$  at room temperature. Since Stokes and anti-Stokes lines give the same vibrational information and Stokes lines are much stronger and temperature independent, only the Stokes side of the spectrum is typically measured in Raman Spectroscopy.

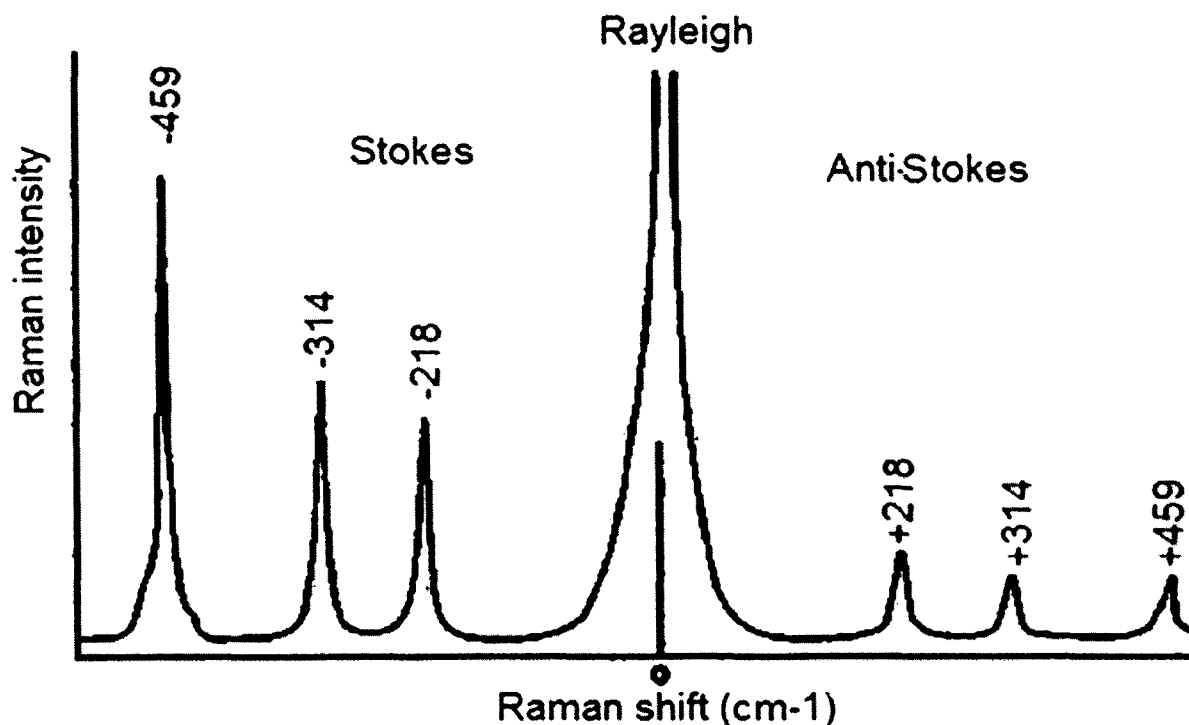


FIG. 2. Raman Spectrum of  $\text{CCl}_4$  showing Stokes and anti-Stokes lines. The spectrum was simulated based on 480nm excitation source  $\text{CCl}_4$  Raman spectrum.

The location of the virtual energy state depends on the wavelength of the excitation beam. The longer the wavelengths of the excitation source beam, the lower the position of the virtual energy state. When using a shorter wavelength excitation source (high energy lasers, such as 488 or 514 nm Argon ion laser lines), the position of the virtual energy state is high and usually the virtual energy state overlays with one of the higher excited electronic states of the molecule. In the case when the virtual energy state overlays with an excited electronic state, the electrons in the higher vibrational energy

level of the excited electronic state vibrationally relax to the lowest vibrational energy level of the excited electronic state. Finally, it is possible for the electrons to drop from the first vibrational energy level of the excited electronic state to the ground state releasing radiation known as fluorescence. The intensity of fluorescence is much more intense than the Raman signal and often obscures the Raman signals. To overcome this problem, Raman spectrometers usually utilize longer wavelength excitation sources, such as the 1064 nm line of the Nd:YAG laser. This leads to a large reduction in the fluorescence background. However, utilizing longer wavelength excitation sources has two problems: first, the intensity of the Raman scattering signal is proportional to  $1/\lambda^4$ . Second, the use of 1064 nm radiation prohibits the use of conventional silicon CCD detection (since the Raman photon energy is less than the silicon band gap), and hence FT-Raman spectrometers are used. These spectrometers typically contain Michelson interferometers. The combination of these two problems results in a large instrument with moving parts and a weak Raman signal requiring a long integration time. Thus they are difficult to deploy for in-line or at-line analysis.

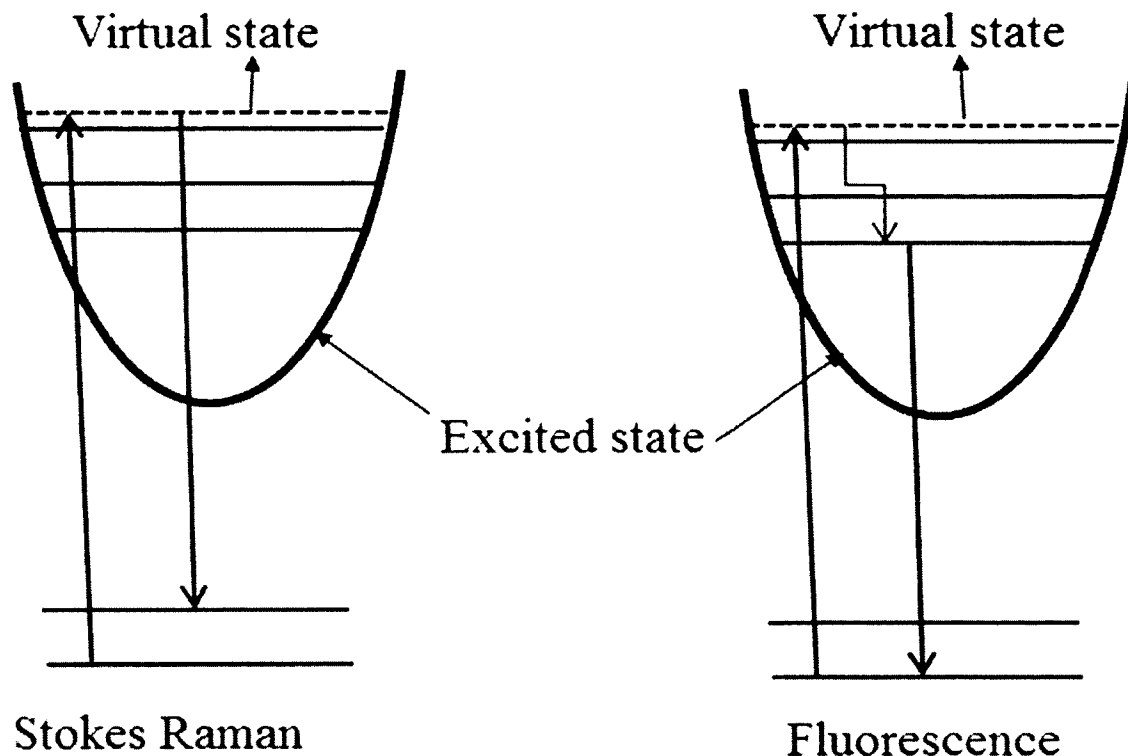


FIG. 3. A virtual energy state overlays with an excited electronic state. This leads to fluorescence emission which plagues Raman signals.

### SEQUENTIALLY SHIFTED EXCITATION RAMAN

Sequentially shifted excitation Raman spectroscopy is based on the hypothesis that a small shift in the excitation source wavelength leads to a shift in the wavelength of Raman signals while the fluorescence background does not shift. This is visually depicted in Figure 4, where increasing the excitation source energy leads to an increase in the position of the virtual energy state. The change in the virtual energy state position results in a shift in Raman signal, while fluorescence background frequency remains the same. Raman Stokes and anti-Stokes frequencies ( $\nu_0 \pm \nu_{\text{vib}}$ ) are dependent on the excitation source frequency for the same vibrational information. This means changing the excitation source frequency will alter the frequency of Raman signals by the same amount. However, in the case of fluorescence, electrons relax from the higher vibrational energy levels to the lowest vibrational energy level of the excited electronic state before they finally drop back to the ground electronic state. Thus, fluorescence frequency

depends only on the spacing between the electronic state and the ground state (specific to a molecule) and will not change due to a small change in the excitation source energy.

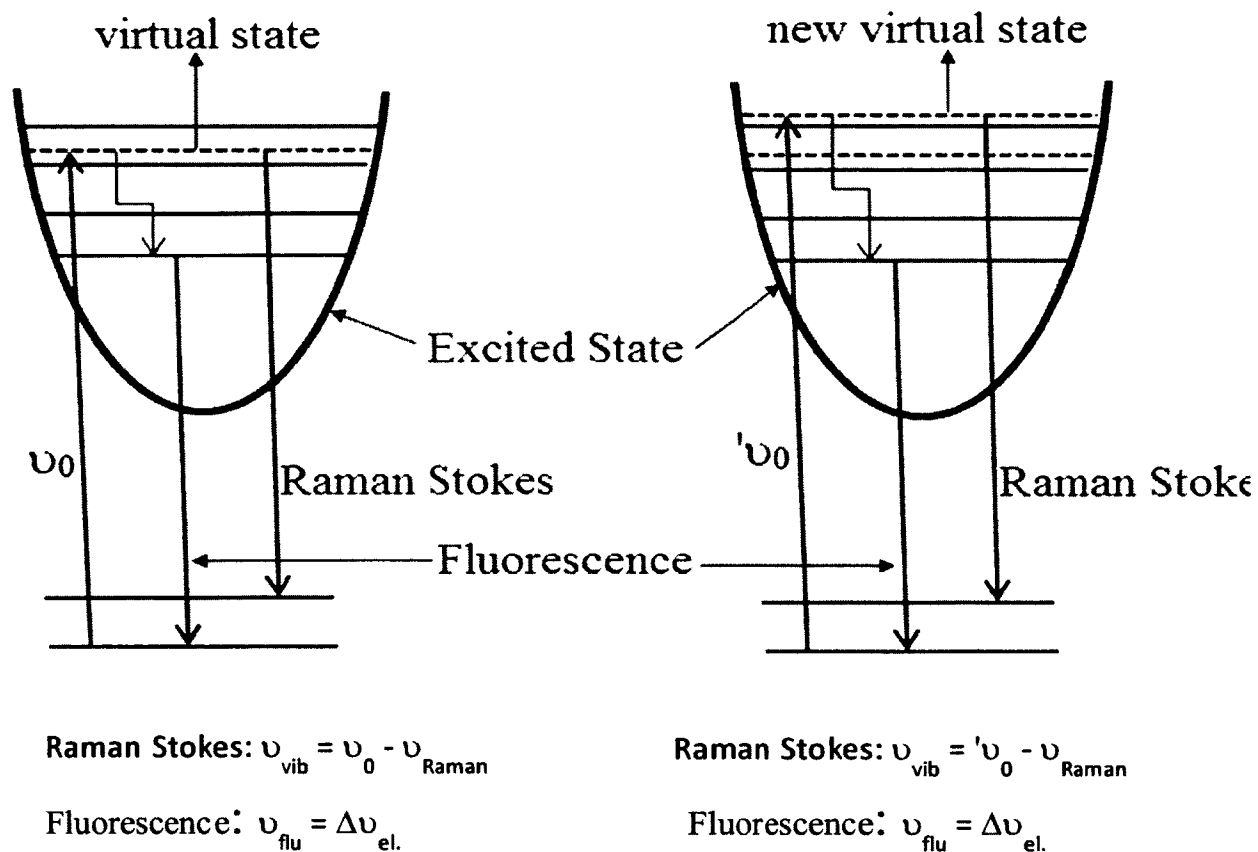


FIG. 4. The effect of changing the excitation source wavelength on the Raman and fluorescence background. Changing the excitation source energy leads to changing the virtual energy state which leads to changing Raman band frequency while fluorescence does not change.

In this dissertation, a novel design for a Raman spectrometer is presented. The spectrometer utilizes a temperature-tunable diode laser which allows the instrument to operate in a sequentially shifted excitation mode to eliminate fluorescence backgrounds. The Raman data is extracted from the shifted excitation spectra using a novel algorithm which separates constant fluorescence background out of shifted Raman signals. The instrument design as well as the algorithm allow for a real time solution.



## NIR THEORY

The near infrared region (NIR) spans the wavelength range of 750-2500 nm. The absorbance bands in this region mainly correspond to overtones and combinations of fundamental vibrations and are due to the interaction of NIR light waves with the vibrational characteristics of a molecule. A basic understanding of these vibrational characteristics is essential to understand the nature of NIR spectra. Vibrational characteristics of a diatomic molecule are the simplest example that can be used to understand the nature of the interaction between NIR waves and matter.

A diatomic molecule can be classically described as two balls, with masses  $m_1$  and  $m_2$ , connected by a spring (representing the chemical bond between the two atoms) that has a force constant  $k$ . The two balls will oscillate about an equilibrium position. The magnitude of the oscillation displacement is directly proportional to the force constant of the spring and inversely proportional to the masses of the two balls. Thus, the energy ( $E$ ) of this system, which is proportional to the vibrational frequency, can be described using a form of Hooke's law:

$$E = h \frac{1}{2\pi} (k/\mu)^{1/2} \quad (6)$$

where  $\mu$  is the reduced mass of the two atoms ( $\mu = \frac{m_1 * m_2}{m_1 + m_2}$ ),  $k$  is the force constant which is a characteristic of a system (bond strength), and  $h$  is Planck's constant ( $6.626 \times 10^{-34}$  J s). At any given displacement, the potential energy is given by Eqn. 7:

$$U = \frac{1}{2} k (x - x_0)^2 \quad (7)$$

where  $U$  is the potential energy at a displacement distance  $x$  from the equilibrium position,  $x_0$ , that represents the diatomic molecule bond length. This potential energy for a diatomic molecule is plotted in Figure 5 where the x-axis reflects the displacement from equilibrium. This depiction represents the harmonic oscillator model for a diatomic molecule. Since the vibrational levels of the molecule are quantized, the vibrational energies are restricted to discrete energies that are given by the following Eqn. 8:

$$E_v = (v + \frac{1}{2}) h \nu \quad (8)$$

where  $v$  is the vibrational quantum number, and  $E_v$  is the vibrational energy associated with  $v^{\text{th}}$  vibrational level, and  $\nu$  is the fundamental vibrational frequency,  $\nu = \frac{1}{2\pi} (k/\mu)^{1/2}$ . The difference in energy between any two adjacent energy levels is:

$$\Delta E_{vib} = E_{v2} - E_{v1} = \Delta v h \nu \quad (9)$$

Thus by absorbing a photon of energy  $h\nu$ , the existence of a vibration can be detected using absorbance spectroscopy. This is the same vibration detected using Raman Stokes spectroscopy ( $\nu_{\text{vib}} = \nu_0 - \nu_{\text{Raman}}$ ). Based on Eqn. 9, the only transitions that are possible between consecutive energy levels must have  $\Delta v = \pm 1$ . This means that the separation between the adjacent vibrations is always equidistant in energy and is equal to  $h\nu$ . From this, one would predict that for a NIR transition such as the transition from  $v_0$  to  $v_3$ , the energy difference would be equal to  $3 h\nu$ . However, this is not what is observed and thus a refinement to the harmonic model is required.

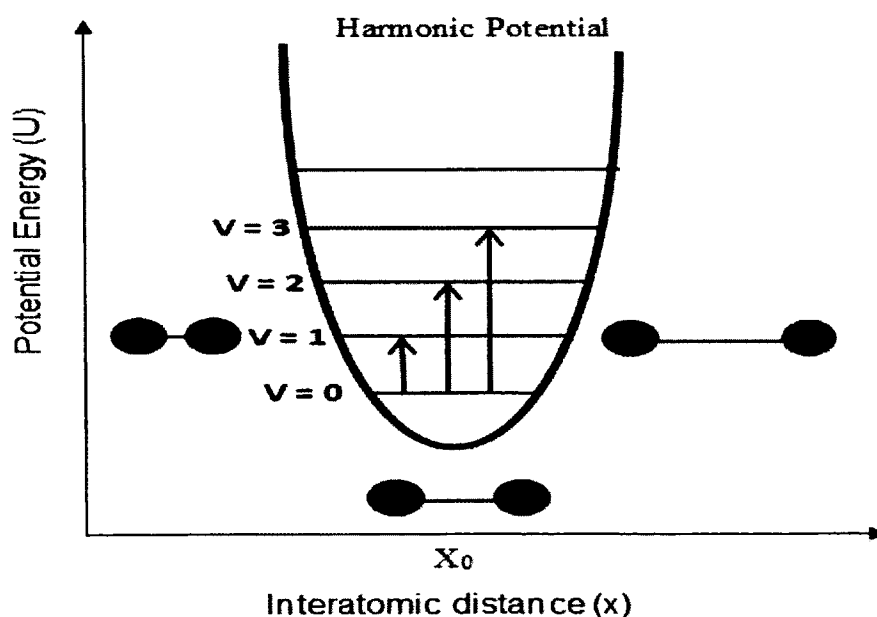


FIG. 5. The harmonic model for potential energy of a diatomic molecule.

The harmonic oscillator model cannot explain the NIR behavior in the actual molecules because it does not account for coulombic repulsion between atoms or the dissociation of bonds. In addition, for the quantized harmonic oscillator model, there is a requirement for  $\Delta v$  to be equal to +1 or -1, so that overtones are formally forbidden. However, the NIR behavior of molecules can be approximated by an anharmonic oscillator in which energy levels are not equally spaced which relaxes the selection rules for overtone vibrations. An overlay comparison of the harmonic model with the anharmonic model, which offers a correction for the harmonic “failures”, is shown in Figure 6. As can be seen, the anharmonic model accounts for repulsion between atomic nuclei at low displacements and bond dissociation at high displacements.

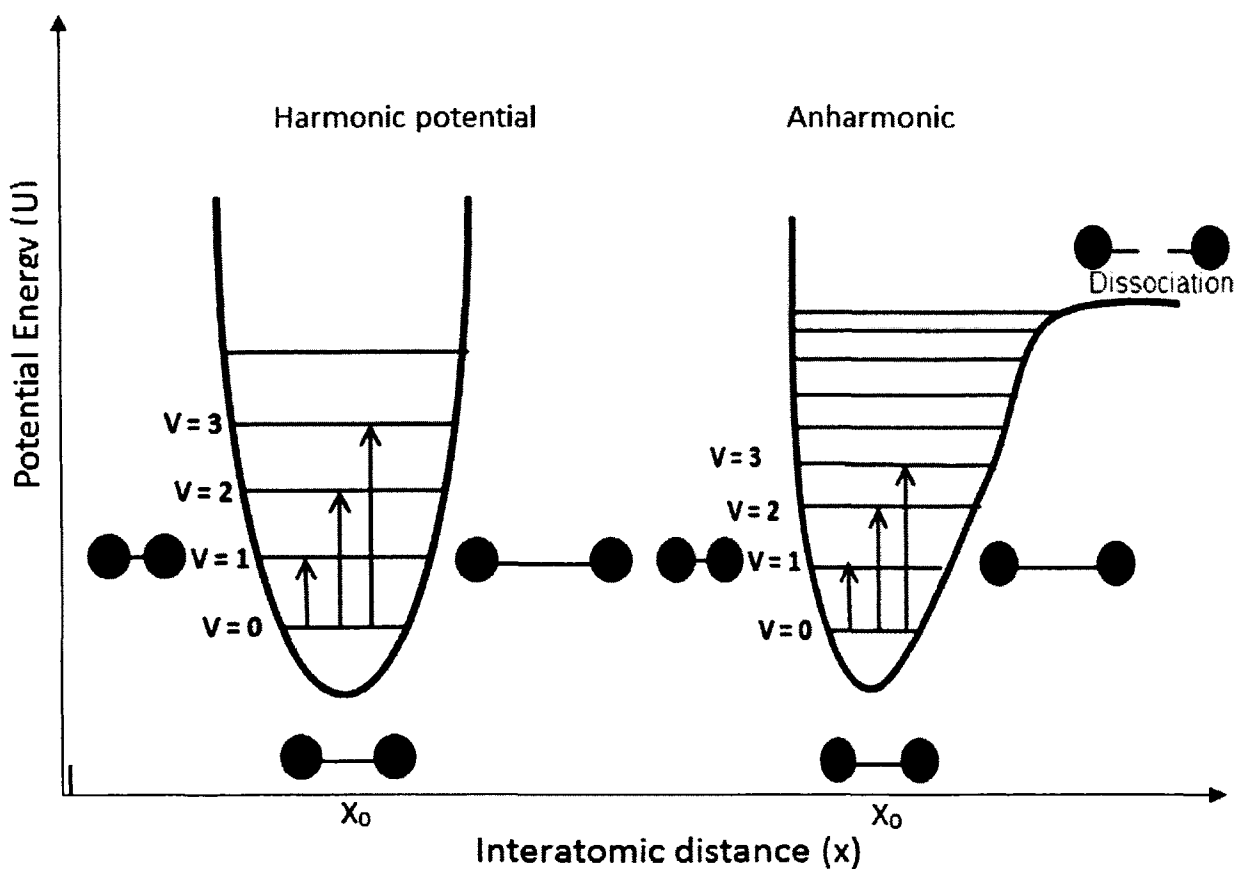


FIG. 6. Side-by-side comparison of the harmonic and anharmonic model for potential energy of a diatomic molecule.

The potential energy of the anharmonic model, shown in Figure 6, is given by the Morse relation which accounts for the anharmonic behavior:

$$U = D_e (1 - e^{-a(x-x_0)})^2 \quad (10)$$

where  $D_e$  is the bond dissociation energy, and  $a$  is a constant which is specific for a given molecule reflecting the degree of anharmonicity. Based on Eqn. 10, when  $x$  is large (the case of large displacement), the potential energy of the system approaches the dissociation energy. While, when  $x$  is equal to  $x_0$  (the displacement is zero) and the potential energy of the system is equal to the ground vibrational energy level. By applying the Schrödinger equation to the anharmonic diatomic model, the energy difference between vibrational states that describes NIR behavior of the actual diatomic molecule can be calculated as follow:

$$\Delta E_{\text{vib}} = h\nu \left( v + \frac{1}{2} \right) - ah\nu \left( v + \frac{1}{2} \right)^2 \quad (11)$$

where  $a$  is the anharmonicity constant. Based on Eqn. 11, transitions between vibrational states when  $\Delta v = \pm 2, \pm 3 \dots$  are more accurately calculated. These transitions yield NIR absorbance bands know as first, second, third overtones, etc. Although selection rules are relaxed under the anharmonic model, overtone transitions are much weaker than the corresponding fundamental vibrations. The first overtone band is generally 10-100 times weaker than its corresponding fundamental band. Overtone bands appear between 780

nm and 2000 nm, depending on the bond strength and nature, as well as the overtone type (e.g.  $\Delta v$ ). In polyatomic molecules, more than one NIR vibrational mode can interact and cause a simultaneous energy change that gives rise to an absorption band known as combination band. NIR combination bands generally appear between 1900 to 2500 nm.

NIR band intensity depends on the change in the dipole moment (as opposed to Raman intensity which depends on a change in the polarizability). The larger the change in the dipole moment during NIR absorbance, the more intense the NIR band will be. Within a molecule, stretching vibrations involving displacement of a hydrogen atom exhibit a large deviation from harmonic oscillator behavior because hydrogen atom is the lightest atom and the fundamental stretching vibration occurs at a very high energy (see Eqn. 6). In fact, the main bands observed in the NIR overtone absorbance regions typically correspond to bonds which contain light atoms such as C-H, N-H, O-H, and S-H. On the other hand, NIR bands corresponding to heavy atoms, such as C=O, C-C, and C-Cl, are much weaker or absent. Thus, the presence of intense overtone and combination bands are relatively prominent in the NIR spectra of molecules containing atoms that are bound to hydrogen (e.g. C-H, O-H, N-H), with particularly good resolution in the 2<sup>nd</sup> and 3<sup>rd</sup> overtone regions which is the case of fuel. The result is that fuels have absorption in the NIR region. In addition, the NIR spectral range is capable of being transmitted through fiber optics, which facilitates use of deployable portable fuel analyzers based on NIR absorbance.

## NIR ABSORBANCES

Fuels are hydrocarbon based and the only NIR absorbances of greatest significance are the C-H overtone and combination bands. There are three predominant C-H groups that yield NIR vibrational spectra: aromatic, methyl, methylene. (to a lesser extent methine and tertiary C-H stretching is important in gasoline fuels which are highly branched). A summary of the important NIR absorbances is presented in Table 1.

TABLE 1. Characteristic NIR absorbances for C-H groups of the 3rd overtone region.

Functional Group	Vibration	wavelength Range (nm)
Aromatic	C-H stretching	1100 - 1144
Methyl	C-H stretching	1190 - 1195
Methylene	C-H stretching	1135 - 1215
Methyl	C-H stretch and bending combination	1375 - 1400
Methylene	C-H stretch and bending combination	1410 - 1440

Figure 7 shows an overlay of the NIR spectra of three neat solvents: toluene, isooctane, and nonane. Each solvent was chosen based on molecular composition in order to better illustrate the three main overtone regions. Toluene is an aromatic compound, it therefore has a main absorbance in the aromatics region, and also a slight absorbance in the methyl region. Isooctane is comprised of mostly methyl groups and thus has a large absorption peak corresponding to methyl C-H stretching in the 1190-1195 region. Also, a slight shoulder as the peak tails off is indicative of minimal methylene content. Nonane, a long straight chain hydrocarbon, is comprised mostly of methylene groups which is reflected in its absorbance spectrum, where the methylene C-H stretching peak from roughly 1135 - 1215 is the most prominent feature.

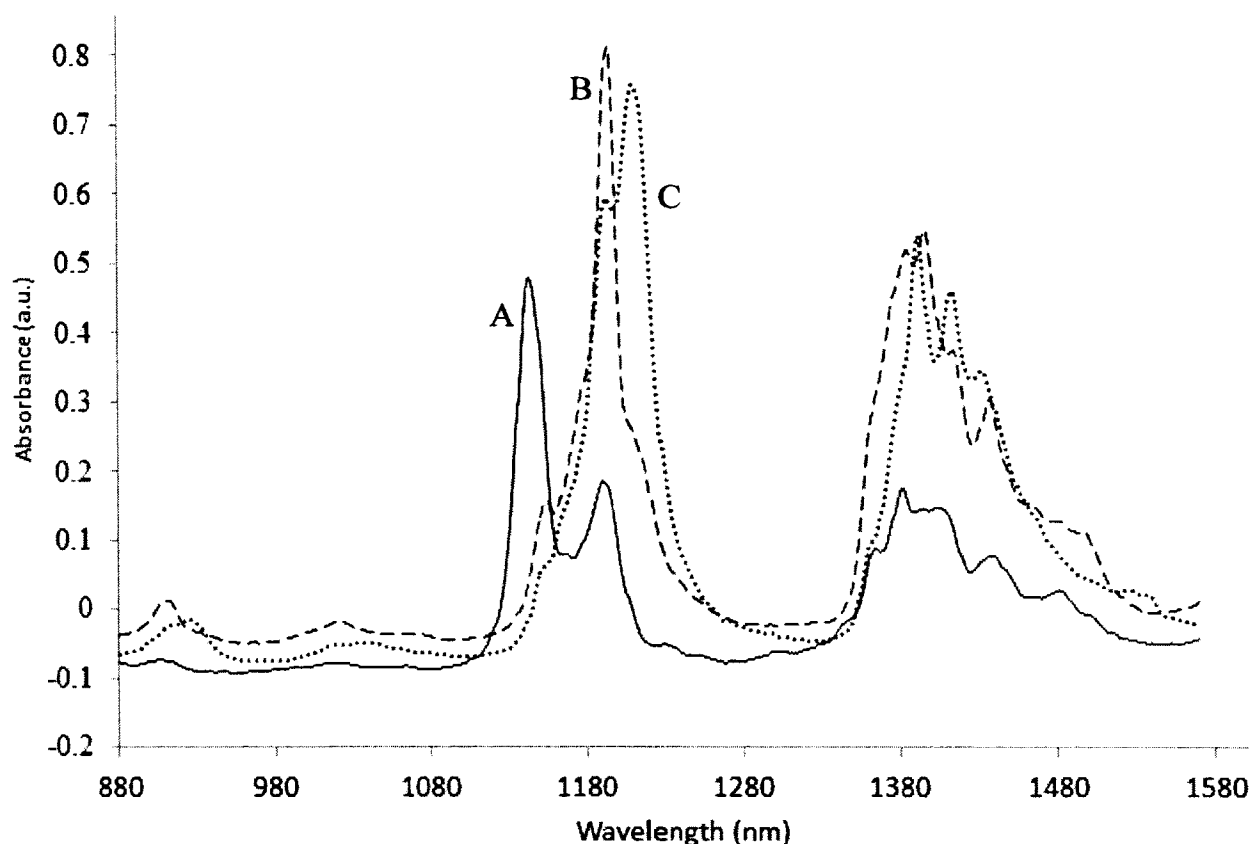


FIG. 7. The NIR spectra of a) toluene b) isooctane and c) nonane.

NIR spectra in conjunction with multivariate analysis have been used to quantitatively predict fuel properties. However, the lack of resolution in the NIR absorption bands (as shown in Figure 7) makes calibration transfer between different NIR instruments extremely difficult and expensive. For fuel analysis, calibration transfer is a particularly difficult problem because maintaining the compositional integrity of fuel standards which often contain thousands of volatile chemicals is difficult at best. This problem is compounded by the need to both store these standards and to transport them to geographically diverse locations. To avoid these issues, the most common solution is to acquire a new spectral database of available fuels on a new instrument and build a new multivariate model specific for that instrument. In this dissertation a novel calibration transfer algorithm is presented which allows transfer of calibration models between different NIR instruments using only pure chemicals which can be obtained universally at high purity levels.

## OBJECTIVES

The presented research in this dissertation is aimed to demonstrate that;

- 1) Optical interferences associated with Raman spectroscopy due to analyte and non-analyte fluorescence, fixed noise pattern, and room light can be eliminated by;
  - I. Designing a Raman spectrometer that utilizes a temperature-controlled distributed-Bragg-reflector diode laser which allows the instrument to operate in a sequentially shifted excitation mode.
  - II. Algorithmically removing the pure-Raman spectrum out of fluorescence background using a novel algorithm that discriminates between the shifted Raman signals and the unshifted background in the sequentially shifted excitation Raman spectra data.
  
- 2) Partial least square (PLS) models of fuel properties built on a NIR instrument can be transferred to different NIR instruments. The research in this dissertation will demonstrate the following;
  - I. NIR spectra of jet fuels in conjunction with multivariate analysis, PLS, can be used to quantitatively predict fuels properties.
  
  - II. Partial least square models of fuels built on a NIR-instrument can be transferred to different instruments using a novel algorithm that utilizes virtual standards digitally generated from the NIR spectra of 15 neat chemical solvents



## CHAPTER II

### NOVEL ALGORITHM AND INSTRUMENTATION FOR FLUORESCENCE-FREE RAMAN SPECTROSCOPY UTILIZING A SINGLE TUNABLE EXCITATION SOURCE

#### INTRODUCTION

Although Raman spectroscopy is a powerful analytical method for molecular analysis, Raman spectra are often plagued with intense fluorescence backgrounds resulting from impurities or from the population of a sample's excited state(s). The use of long wavelength lasers, such as the 1064 nm Nd:YAG laser commonly used for FT-Raman spectroscopy, results in a significant reduction in fluorescence backgrounds. However, since the Raman scattering is inversely proportional to  $\lambda^4$  for the excitation laser, it also results in less Raman signal and thus often requires the use of longer acquisition times and higher laser powers which can often lead to sample burning. In addition, FT-Raman instruments are typically large and expensive, containing integral interferometers which are sensitive to mechanical vibration. For these reasons FT-Raman spectroscopy does not easily lend itself to applications involving process control or remote deployment. Alternatively, dispersive Raman instruments using CCD detection with solid state laser excitation provide a robust option with no moving parts; however, the use of shorter wavelength excitation that the CCD systems require results in significantly more fluorescence. For this reason, several methods have been developed to extract the Raman information from interfering backgrounds. Overall, these methods of removing fluorescence can be separated into four categories: algorithm based baseline correction methods; methods and devices using specialized sampling optics; time gating methods; and shifted excitation methods.

Algorithm-based methods attempt to mathematically estimate a baseline and then subtract the estimated baseline from the Raman spectrum to give a fluorescence-free spectrum. Algorithm-based methods suffer from the requirement that they cannot be universally applied to all types of fluorescence without either significant degradation in performance or significant user intervention to adjust sensitive algorithm parameters. One reason for this disadvantage is the varied nature of fluorescence backgrounds. Another

reason is due to the complexity of the fitting algorithms. This prevents these methods from being used routinely by non-technical personnel. Methods using algorithm based baseline corrections are described in detail in the peer-reviewed scientific literature<sup>1-6</sup>.

Methods using unique sampling optics and geometries have the disadvantage of only mitigating the impact of fluorescence instead of completely eliminating it from the Raman spectrum. These methods also have the disadvantage of only being applicable to certain sample types. For instance, some of these methods can only be used on solids or only be used on liquids. Peer-reviewed scientific literature describes methods employing unique sampling optics and geometries in detail<sup>7-9</sup>.

Time gating methods involve using time to discriminate between the Raman signal (which occurs on a fast time scale) and the fluorescence (which generally occurs on a slower timescale). Since the time scale discrimination required for these methods is on the order of picoseconds to nanoseconds, these methods involve the use of either: 1) complex, bulky, and expensive instrumentation or 2) instrumentation which requires excessively long periods of time to collect an entire Raman spectrum. Because of the varied nature of fluorescence decay times, these methods are not universally successful. These methods also have the disadvantage of not being able to remove spectral backgrounds which arise from optical processes that occur on the same timescale as Raman. Such processes include stray light scatter, fixed pattern noise arising from a detector array, room lights, cosmic rays, and other extraneous sources of optical interference. These methods also have the disadvantage of requiring the Raman to be generated using a pulsed laser source. This results in a large number of Joules being delivered to the sample in a short period of time and often results in sample burning. Examples of time gating methods are described in detail in the peer-reviewed scientific literature<sup>10-17</sup>.

All shifted excitation methods involve changing the excitation laser during spectral acquisition. These methods rely on a common concept: the location of Raman intensities in spectral space changes with excitation while unwanted spectral intensities corresponding to fluorescence, stray light, fixed pattern detector noise etc., remain unchanged in spectral space. The difference in the various methods to date occurs in how the spectra are acquired and how the Raman data is extracted. The simplest extraction

method is taking a difference between two sets of excitation data; this is referred to as Shifted Excitation Raman Difference Spectroscopy (SERDS) and only requires excitation at two distinct wavelengths. This type of method has two distinct disadvantages: 1) random noise in two measurements is increased during subtraction; 2) the result is a derivative spectrum instead of a true Raman spectrum in spectral space. Since a true Raman spectrum is not generated, the data is not easily interpreted. Another disadvantage of SERDS is that there is significant difficulty in reconstructing the Raman spectrum from the derivative data. The reconstruction requires the use of advanced algorithms which have to be applied appropriately for specific samples since one set of algorithm parameters will not work universally for all samples. This method has been in use for over thirty years, so are many examples of its application in the peer-reviewed scientific literature<sup>18-26</sup>. Perhaps more intriguing is the recent work of Willet and co-workers in which a series of diode lasers at closely spaced wavelengths are used for the acquisition of shifted Raman spectra. The setup allows the researchers to extract the spectral data in true Raman space using an iterative expectation-maximization (EM) algorithm (Lucy-Richardson algorithm) based on a shift-matrix operator. This is the first example of the extraction of Raman data regardless of the complexity or nature of the fluorescence background (i.e., a universal approach). Despite the promising theory of the method, the instrumentation remains complex and it can be anticipated that reproducing the series of selected wavelength lasers would not be possible for multiple instruments. In addition, the processing of the data using the shift matrix operator requires tens of billions of double precision calculations, which results in samples with high background taking several minutes to process. In this chapter, I present a new algorithm and new instrumentation which allow the acquisition of fluorescence-free Raman spectra. I demonstrate that the practical nature of this instrumentation allows for implementation of shifted excitation Raman, even in a low-cost hand-held design. I also present a new algorithm, which allows for true Raman signal extraction that is identical in spectral performance to existing EM algorithms, but is three orders of magnitude faster. This method relies on sequentially shifted excitation (SSE) using a single temperature-controlled distributed Bragg reflector diode laser.

## EXPERIMENTAL SECTION

### *INSTRUMENT DESIGN*

The design of the SSE Raman instrument is shown in Figure 8. The excitation source (Fig.8A) is a distributed Bragg reflector (DBR) GaAs diode laser mounted on a Peltier cooler and contained within a custom TO-8 package. The DBR laser mounted on the thermo-electric controller (TEC) in the TO-8 package are shown in Figure 9. The DBR laser emits single-mode 785 nm radiation at 25 °C. The Bragg grating of the DBR laser is a periodic structure etched into the GaAs substrate at the rear facet of the diode cavity. Hence the cavity and the grating consist of a monolithic structure. The emitting wavelength of the laser at a given temperature is equal to:

$$\lambda = 785 \text{ nm} + 0.074 \text{ nm/}^\circ\text{C} (\Delta T) \quad (12)$$

where  $\Delta T$  is the set-point temperature of the TEC minus 25 °C. The feedback from the Bragg grating stabilizes the modal structure of the laser so that when the laser is operated at a constant current, a range of  $\Delta T$  exceeding 15 °C is obtainable without a change in the single-mode modal structure of the laser output (i.e., no mode hops occur). Once a desired optical power is chosen, its value can be varied typically by less than 15% to obtain a current which provides a mode-hop free range over a large temperature change of the laser. We have evaluated 5 different DBR lasers, all of which have demonstrated this characteristic (consistent with the manufacturer supplied current-voltage curves for each device). For the current work, a sample-incident optical power output of 50 mW was selected and the laser was run in constant current mode (100 mA) by using a fixed resistance on the TEC board (Fig 8B). The TEC board is capable of 0.01 °C precision. This provided a stable mode-hop free temperature range from 16-34°C.

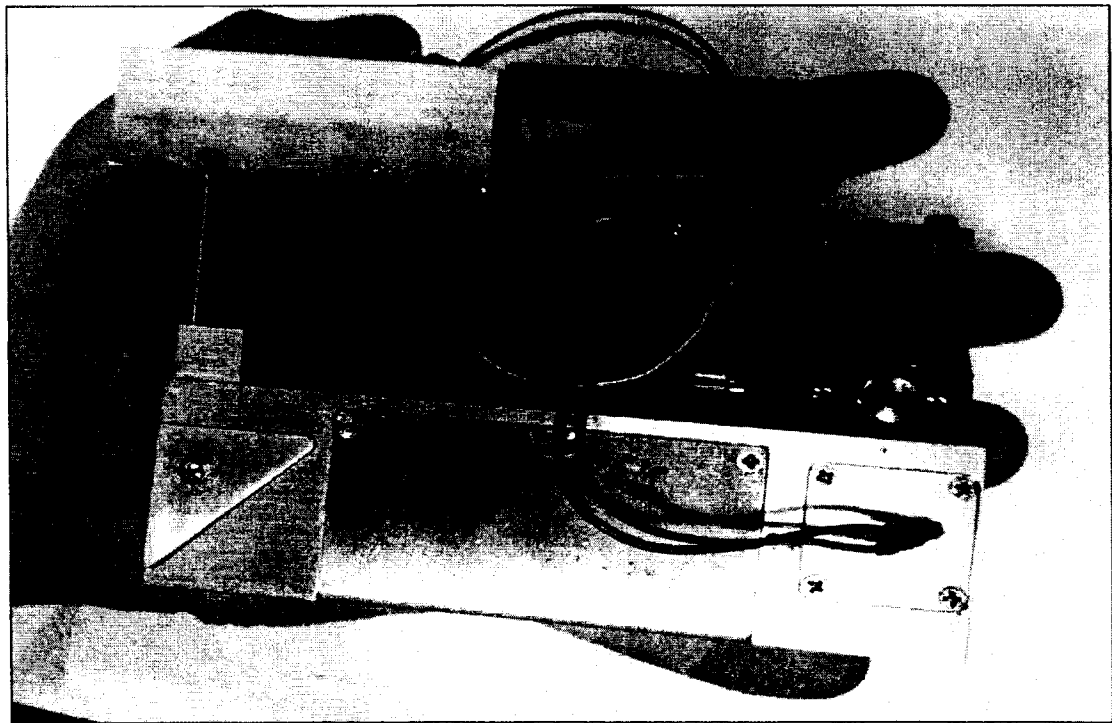
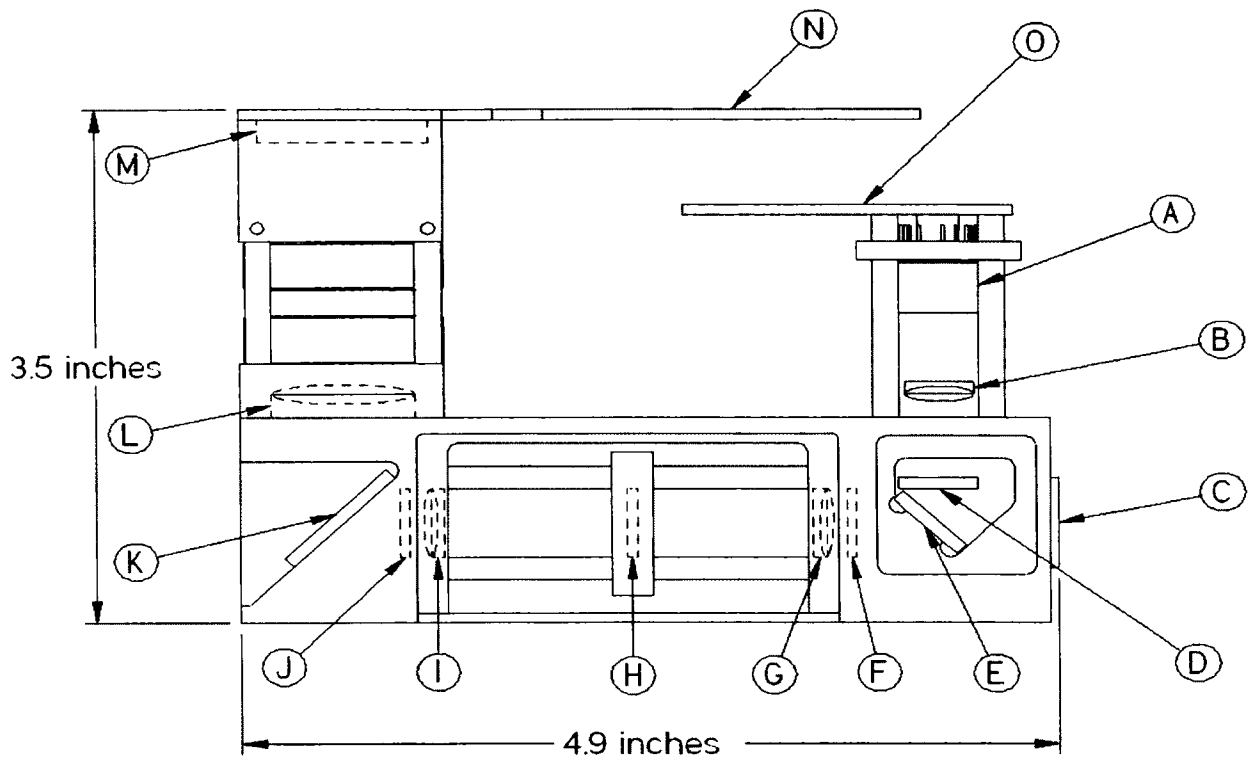


FIG. 8. The SSE Raman Spectrometer is composed of: A) a 785nm DBR GaAs diode laser, B) a collimating asphere, C) a focusing/collimating asphere, D) a laser bandpass filter, E) a dichroic beamsplitter, F) a long-pass filter, G) a focusing doublet achromat, H) a 40 micron rectangular slit, I) a collimating doublet achromat, J) a long-pass filter, K) a 1500 groove/mm volume holographic transmission grating, L) a focusing doublet achromat, M) a CCD detector, N) a microprocessor control board, and O) a thermoelectric control/laser diode driver control board.

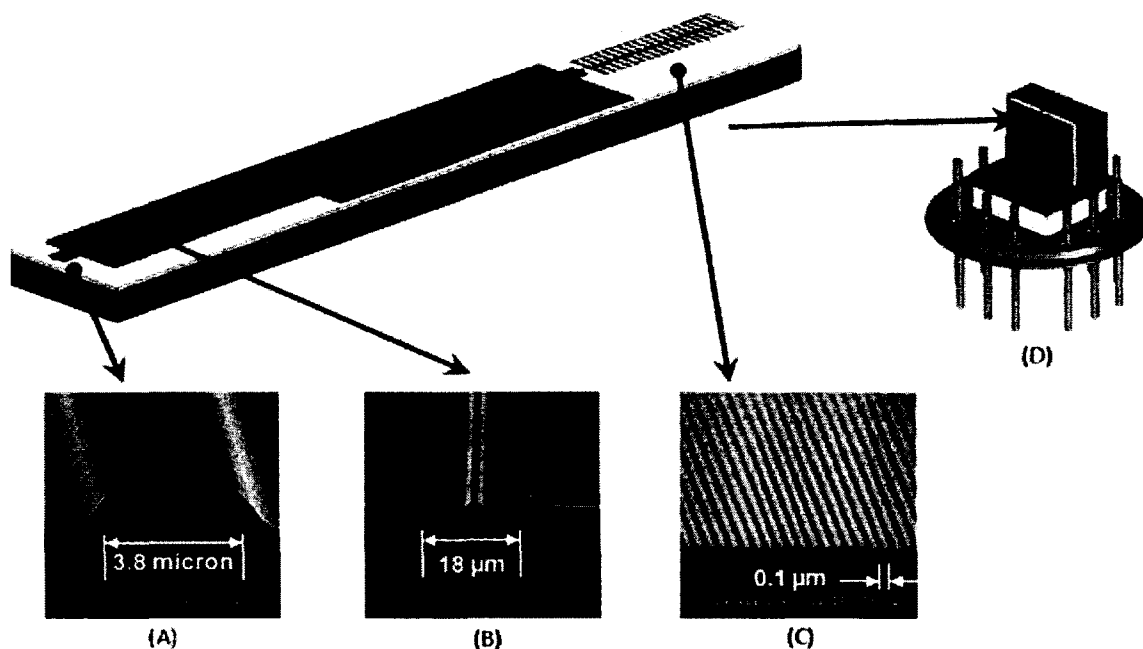


FIG. 9. A diagram of the 785 nm distributed Bragg reflector (DBR) GaAs diode laser is displayed with SEM micrographs showing A) the emitting facet of the laser cavity; B) the laser cavity; and C) the Bragg grating. The diode laser is also shown (D) within a custom TO-8 package mounted to a Peltier cooler with heat sink.

The laser beam is collimated with an asphere ( $NA = 0.55$ ,  $NA$  is the numerical aperture and it is equal to the diameter divided by two its focal length) and filtered with a short pass filter (Fig. 8D). A dichromic beam-splitter (Fig. 8E) directs the laser through a focusing asphere ( $NA = 0.50$ ) onto the sample. The asphere collected Raman scatter is passed through the beam-splitter where it is focused by a 12.7 mm diameter doublet achromat of focal length 32 mm (Fig. 8G) onto a 1mm x 50 micron slit (Fig.8H). An identical achromat is used to collimate the light. The collimated light is diffracted using a volume holographic transmission grating (1500 gr/mm) mounted at  $-45^\circ$  to the optical axis (Fig.8K). The diffracted light is focused onto the back-thinned CCD detector (Fig.8M) using a 45 mm focal length doublet achromat (Fig.8L, 25.4 mm diameter to account for the anamorphic magnification of the grating). The design provides a spectral resolution of approximately  $10 \text{ cm}^{-1}$ . A microcontroller board (Fig. 8N) is used to control both the TEC (Fig.8O) and the CCD data acquisition, and to upload collected raw spectra to a processing CPU over a USB 2.0 connection.

All lenses and opto-mechanical mounts were purchased from Thor Labs. All optical surfaces were anti-reflection coated for the spectral range used. The grating,

beam-splitter, and detector mounts were machined in house. The grating was purchased from Wasatch Photonics. The CCD detector was purchased from Hamamatsu (model S11500-1007) and was used at room temperature without any cooling. The slit with mounting was purchased from Thorlabs. The 785 nm DBR laser was purchased from Photodigm. The DBR laser was filtered with a 785nm laser-line band-pass filter from Chroma (5 nm FWHM). The collected Raman radiation was filtered using a 785nm Razor-edge long-pass filter from Semrock. The 785 di-chroicbeamsplitter was purchased from Semrock also. Custom electronics for the detector, laser driver, and TEC controller were developed in-house. CCD signal conversion was carried out using a 16 bit A/D converter. For typical experiments, the low integration times resulted in only 2000-5000 thermal counts, leaving over 55,000 counts of dynamic range for the spectral signals.

All control code was written in C and compiled using the GNU compiler build for the Freescale MCore. Spectra were uploaded to a computer with a Windows 7 operating system over USB using an USB 2.0 USB/serial converter (FTDI). All spectra were processed in real-time (<1s ) using a custom program written using Labview (National Instruments). All dispersive Raman spectra were calibrated to a wavenumber axis using the NIST standard acetaminophen.

All FT-Raman spectra were acquired using a Magna 950 FT-Raman (Thermo-Fischer Scientific). All FT-Raman spectra were acquired by averaging 100 scans at 4  $\text{cm}^{-1}$  resolution, with 800 mW laser power and Hap-Ganzelapodization. An InGaAs detector at room temperature was used for all acquisitions. Total acquisition time was 2 minutes for each spectrum.

### *INSTRUMENT OPERATION*

A typical measurement consists of collecting Raman spectra at DBR laser temperatures of 20, 23, 26, and 29 °C (i.e., four sequential excitations which are evenly spaced in temperature. This yields excitation wavelengths of 784.630, 784.852, 785.074, and 785.296 nm, respectively, and gives a constant excitation shift of 0.222 nm. When converted to  $\text{cm}^{-1}$ , this gives a separation of approximately 3.60  $\text{cm}^{-1}$  between the different excitations (approximate since the  $\text{cm}^{-1}$  scale is linear to 0.01  $\text{cm}^{-1}$  over such a

small change in wavelength). Due to the presence of the DBR grating and the use of fast optics with anti-reflection coatings, we have not been able to observe any changes in laser quality due to back-reflections into the diode cavity resulting from sample placement. For this reason an optical isolator in front of the diode laser was omitted in the current design.

Although the instrument can be operated in a variety of manners, the use of a microcontroller allows programming of a set sequence of events which minimizes acquisition time and electrical power consumption. For the current work these steps include: 1) turn on TEC; 2) cool laser to a starting temperature of 20 C; 3) turn on laser; 4) reset CCD and collect spectrum; 5) turn off laser; 6) heat laser to next temperature and upload data; 7) repeat steps 3-6 until final spectrum is collected; and 8) turn off TEC. For this sequence, the most time consuming steps are 2 and 4. A schematic of current operation is shown in Figure 10

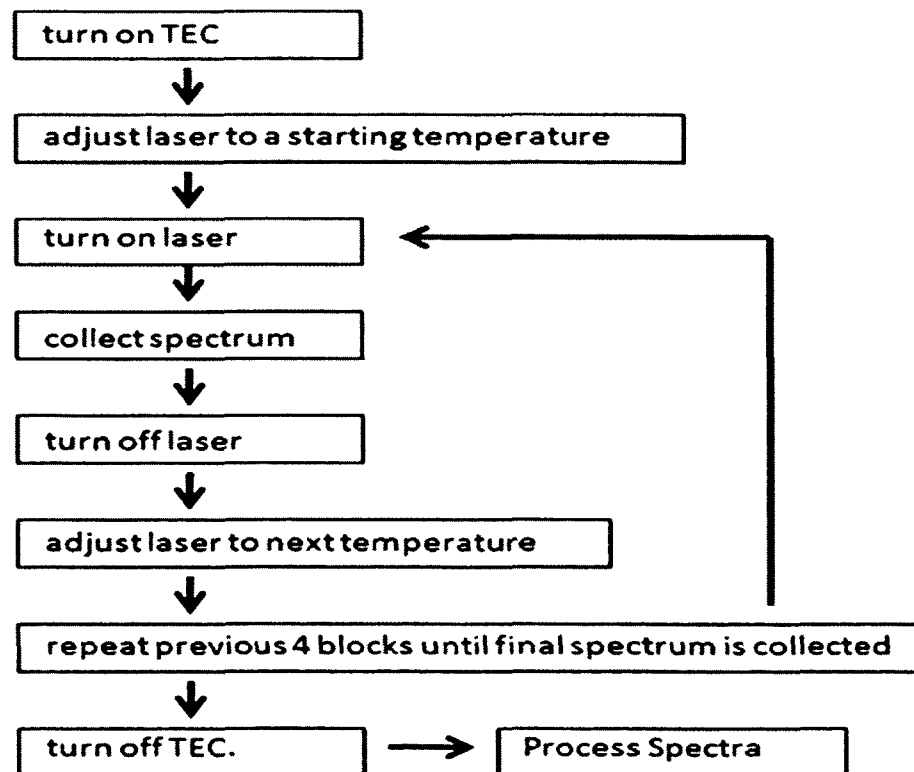


FIG. 10. This figure gives the current sequences of operation of SSE Raman. The most time consuming steps are heating and cooling the laser; heating laser to the next temperature requires 300ms while cooling the laser requires less than a second for temperature steps of 6 °C or less. The total acquisition time equals to the integration time plus less than two seconds.



Due to the small mass of the diode laser, the time to cool the laser (when room temperature is  $< 30$  °C) is less than 1s. The time required to heat the laser to the next temperature is 300 ms for temperature steps of 6 °C or less. The time to upload the data is 28 ms, and thus has no time penalty since it occurs during the TEC heating period. The time required to turn the laser on and off is less than 1 microsecond. Thus the total acquisition time for a four temperature experiment is equal to the total integration times of the spectra plus less than two seconds.

All chemicals were purchased from Aldrich. The chemicals were placed in borosilicate NMR tubes for spectral sampling.

The x-axis of all acquired Raman spectra were linearized to a calibrated x-axis with the intervals of the actual Raman shift ( $3.66 \text{ cm}^{-1}$  in current experiment). The all acquired Raman spectra were pretreated by the standard normal variate (SNV) to remove the difference in Raman intensities between the sequentially shifted excitation Raman spectra (due to the small variability of laser beam output). SNV is shown in Eqn. 13:

$$SNV_i = \frac{R_i - \hat{R}_i}{\sigma_i} \quad (13)$$

where  $i$  is the spectral index of the raw Raman spectra (i.e.,  $i$  is equal to the number of excitation used),  $R_i$  is a vector contains a Raman spectrum,  $\hat{R}_i$  is the mean of the  $R_i$  spectrum,  $\sigma_i$  is the standard deviation of the  $R_i$  spectrum.

### *RAMAN DATA PROCESSING*

Raman spectroscopy fluorescence-removal techniques which involve changing the excitation laser during spectral acquisition all rely on a common concept: the location of Raman intensities in spectral space changes with excitation while unwanted spectral intensities corresponding to fluorescence, stray light, fixed pattern detector noise etc., remain unchanged in spectral space. The difference in the various methods to date occurs in how the spectra are acquired and how the Raman data is extracted. The simplest extraction method is taking the difference between two sets of excitation data and is referred to as Shifted Excitation Raman Difference Spectroscopy (SERDS) and only

requires excitation at two distinct wavelengths. This type of method has two distinct disadvantages: 1) random noise in two measurements is increased during subtraction; 2) the result is a derivative spectrum instead of a true Raman spectrum in spectral space. With the addition of more than two excitation wavelengths, however, multivariate solutions can be used to extract the Raman intensities while minimizing noise. In such cases, the Raman spectra acquired at each unique laser excitation can be described by a matrix  $\mathbf{R}$ :

$$\mathbf{R} = \begin{pmatrix} r_{0,0} & \dots & r_{0,n} \\ \vdots & \ddots & \vdots \\ r_{k,0} & \dots & r_{k,n} \end{pmatrix} \quad (14)$$

where each row corresponds to the Raman spectrum acquired with  $k$  laser excitation (total of  $K$  excitations) and  $n$  is the spectral index (typically in  $\text{cm}^{-1}$ ) and each spectrum has a spectral length of  $N$ . Using conventional matrix decomposition such as singular value decomposition (SVD),  $\mathbf{R}$  can be related to a loadings matrix  $\mathbf{L}$ :

$$\mathbf{R} = \mathbf{T}\mathbf{L}^T \quad (15)$$

where  $\mathbf{L}^T$  is the transpose of the loadings matrix and where there are  $K$  rows in  $\mathbf{L}$  and each row corresponds to a principal component (loading vector) which describes orthogonal variance in  $\mathbf{R}$ . The scores matrix  $\mathbf{T}$  relates how much of each principal component is required to reconstruct each spectrum in  $\mathbf{R}$ . A characteristic of the loading matrix is that the principal components are sorted in descending order of variance. Since the principal variation of  $\mathbf{R}$  is the change in excitation lasers when collecting the data, and since this principally only affects the Raman signal, the first principal component describes the variation in the Raman signal as the excitation lasers are changed and is observed as a pseudo-derivative Raman spectrum. The spectral components which are independent of small changes in excitation wavelength (e.g., fluorescence) as well as a portion of the random noise are described by the higher principal components. This method is similar to SERDS, but has the added advantage of reducing spectral noise as opposed to increasing it. For both methods, however, generating the original Raman spectrum without a background is often difficult or impossible for many samples.

An alternative method for solving for the Raman spectrum when using multiple excitations involves relating the desired outcome (separated Raman spectrum and fluorescence spectrum) to the collected data ( $\mathbf{R}$ ) using an operator matrix  $\mathbf{H}$ :

$$\mathbf{H} \mathbf{S} = \mathbf{R} \quad (16)$$

where  $\mathbf{S}$  is a  $1 \times 2N$  matrix containing a column vector of spectral events which are independent of small changes in excitation wavelength such as fluorescence ( $\mathbf{S}^F$ ) and a column vector corresponding to the true Raman signal ( $\mathbf{S}^R$ ):

$$\mathbf{S} = \begin{pmatrix} \mathbf{S}^F \\ \mathbf{S}^R \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} S_0^F \\ \vdots \\ S_n^F \end{pmatrix} \\ \begin{pmatrix} S_0^R \\ \vdots \\ S_n^R \end{pmatrix} \end{pmatrix} \quad (17)$$

where both the fluorescent spectrum ( $\mathbf{S}^F$ ) and the Raman spectrum ( $\mathbf{S}^R$ ) consists of  $N$  spectral positions. The spectral data matrix ( $\mathbf{R}$ ) is rewritten as a concatenated column vector of spectra:

$$\mathbf{R} = \begin{pmatrix} R_0 \\ \vdots \\ R_k \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} r_{0,0} \\ \vdots \\ r_{0,n} \end{pmatrix} \\ \vdots \\ \begin{pmatrix} r_{k,0} \\ \vdots \\ r_{k,n} \end{pmatrix} \end{pmatrix} \quad (18)$$

where each spectrum has  $N$  spectral positions of intensity  $r_{k,n}$  at each spectral position, and there are a total number of spectra corresponding to the number of excitations ( $K$ ). The operator matrix  $\mathbf{H}$  consists of two columns of  $N \times N$  square sub-matrices:

$$\mathbf{H} = \begin{pmatrix} H_0^F & H_0^R \\ H_1^F & H_1^R \\ H_2^F & H_2^R \\ \vdots & \vdots \\ H_k^F & H_k^R \end{pmatrix} \quad (19)$$

where the first column corresponds to sub-matrices which describe the spectral position of fluorescence intensity ( $H_k^F$ ) at each excitation ( $k$ ), and the second column corresponds to sub-matrices which describe the spectral position of Raman intensity at each excitation ( $H_k^R$ ). Due to the sparse nature of the operator matrix, it is not possible to explicitly solve Eqn. 16 for  $S$ . However, an iterative approximation algorithm such as a Lucy-Richardson formula has been invoked and only requires a reasonable initial estimate of  $S^R$  and  $S^F$  to be provided. For example, selecting the maximum spectral intensity from each  $k$  spectrum at each spectral position  $n$  will maximize Raman contributions yielding an initial Raman estimate ( $S_0^R$ ). Likewise, selecting the minimum spectral intensity from each  $k$  spectrum at each spectral position  $n$  will minimize Raman contributions yielding an initial fluorescence estimate ( $S_0^F$ ). This iterative approach using an operator matrix is shown in Eqn. 20:

$$S_{i+1} = S_i \times (H^T(R \div HS_i)) \quad (20)$$

where multiplication and division operators are carried out element-wise and  $i$  is the iteration number ( $H, S$ , and  $R$  remain defined as in the previous three Eqn.s). Using this approach, McCain and coworkers have reported resolution of background-free Raman spectra using 100 iterations or less when the method is combined with de-noising approaches<sup>27, 28</sup>. However, as can be observed, the large size of the operator matrix ( $2N \times kN$ ) results in an extensive number of calculations. We have experienced that using this approach on a conventional desktop computer typically requires over 2 minutes to carry out 500 iterations when  $N = 1024$  and  $k = 4$  (over 25 billion calculations each requiring double precision). Indeed, the numerically intensive nature of this approach is the reason behind the inclusion of de-noising steps to minimize the iteration number<sup>27</sup>. Without a de-noising step, convergence does not often occur before 1000 iterations.

Since each row of the sub-matrices of the operator matrix (Eqn. 19) consists of a single unity value with the remaining elements being zero, the operator matrix is sparse. As one would expect since the fluorescence spectrum is not expected to change with excitation, for the first column of sub-matrices, every sub-matrix is the  $N \times N$  identity matrix. Also, the first sub-matrix in the second column of sub-matrices is the  $N \times N$  identity matrix, as expected since  $k=0$  and the laser excitation has not changed yet. An

example of this is given in eqn. 21 for an experiment using 3 laser excitations and N spectral positions, the operator matrix is not only sparse, it also consists of multiple identity sub-matrices:

$$H_{K=3} = \left( \begin{array}{c} k=0 \\ k=1 \\ k=2 \end{array} \begin{array}{c} \text{Fluorescence} \\ \text{Fluorescence} \\ \text{Fluorescence} \end{array} \begin{array}{c} \text{Raman} \\ \text{Raman} \\ \text{Raman} \end{array} \right) \quad (21)$$

Consideration of the facts that multiplication by the identity does not cause change, and considering that anything times zero is equal to zero, leads to the conclusion that there exists a much simpler and more efficient method. Our approach is to use a weighting vector,  $\hat{w}_i$ , which is calculated with each iteration,  $i$  and carry on only the important calculations; no multiplication by one or zero is included:

$$\hat{w}_i = \sum_{k=0}^{K-1} R_{k,n} \div (S_{i,n}^F + S_{i,n-k}^R) \quad (22)$$

where the division and addition operators are carried out element-wise.  $S_{i,n}^F$  is an initial estimate of the background signal (e.g. fluorescence or any non-Raman signal) is obtained by taking the minimum of each column in matrix R:

$$S_0^F = (\min_k r_{k,0}, \min_k r_{k,1}, \dots, \min_k r_{k,N-1}) \quad (23)$$

where  $S_0^F$  represents the vector containing the initial fluorescence estimate (e.g. the estimate at the 0<sup>th</sup> iteration).  $S_{i,n-k}^R$  An initial estimate of the Raman spectrum is obtained by calculating the standard deviation of each column in R:

$$S_0^R = \langle \sigma(r_{0,0}, r_{1,0}, \dots, r_{K-1,0}), \sigma(r_{0,1}, r_{1,1}, \dots, r_{K-1,1}), \dots, \sigma(r_{0,N-1}, r_{1,N-1}, \dots, r_{K-1,N-1}) \rangle \quad (24)$$

where the  $r_{k,n}$  elements in parentheses indicate a particular column of matrix R and  $\sigma$  is the standard deviation of that column, and  $S_0^R$  represents the vector containing the initial Raman estimate (e.g. the estimate at the 0<sup>th</sup> iteration). Then for each iteration  $i$ , the fluorescence spectrum is then calculated as:

$$S_{i+1}^F = S_i^F \times \hat{w}_i \quad (25)$$

and the Raman spectrum is calculated as:

$$S_{i+1}^R = S_i^R \times \hat{w}_i \quad (26)$$

Significantly, the number of calculations using this approach is proportional to  $k \times N$  per iteration, while that of Eqn. 20 is proportional to  $k \times N^2$  per iteration. For a spectral acquisition with 1000 spectral points (typical of CCD detection), this results in a decrease by 3 orders of magnitude in the number of calculations, resulting in total processing times of a fraction of a second using a conventional desktop computer. When presented with identical spectra, both algorithms yield identical results for  $S$ . In addition, the simple nature of Eqn. 22-26 allows implementation with only a few lines of simple code in

common computer languages such as C as a result of the elimination of the operator matrix. Elimination of  $\mathbf{H}$  and its transpose also results in a massive reduction in memory requirements for the CPU. Significantly, a part of the simplicity of the approach lies in the use of a fixed change in wavelength (excitation shift) between the utilized excitations. Use of varying excitation shifts would require the incorporation of a vector which indexes each of the shifts for each  $k$  calculation.

As presented, this new approach (Eqn. 22-26) and the previous method (Eqn. 20) both require that a spectral shift corresponds to the minimum difference in wavenumbers between the spectral data points. With the current method, that is possible by selecting the appropriate laser temperatures. However, if a series of different lasers are used with non-constant wavelength spacing, then the data must be splined so that the operator matrix indices can accommodate the varying shifts. The resulting loss in symmetry of the operator matrix not only complicates both algorithms, but it also results in the requirement that the algorithm be specifically tailored for each instrument which is produced. In addition, splining the data to obtain the requisite shift resolution will result in an increase in  $N$ , which further increases the processing time.

Figure 11 describes the overall extraction process of the Raman signals from the sequentially shifted excitation spectra. Where (A) represents the actual SSE Raman spectra (at least three spectra acquired at three different excitations temperature) acquired on the SSE Raman analyzer. The B represents the linearization of the SSE Raman spectra to a calibrated axis of energy spacing in wavenumber units which is equal to the actual shift in the Raman peaks. Also, the B includes the SSE Raman spectra pretreatment by the SNV given in Eqn. 13. The C & D give the initial estimates of fluorescence and Raman calculated from the SSE Raman spectra using Eqn.s 23 & 24, respectively. The E, F, & J represent the SSE algorithm shown in details in the Eqn.s 22, 25, & 26. The number of iteration (I) is set at 2000 iterations.

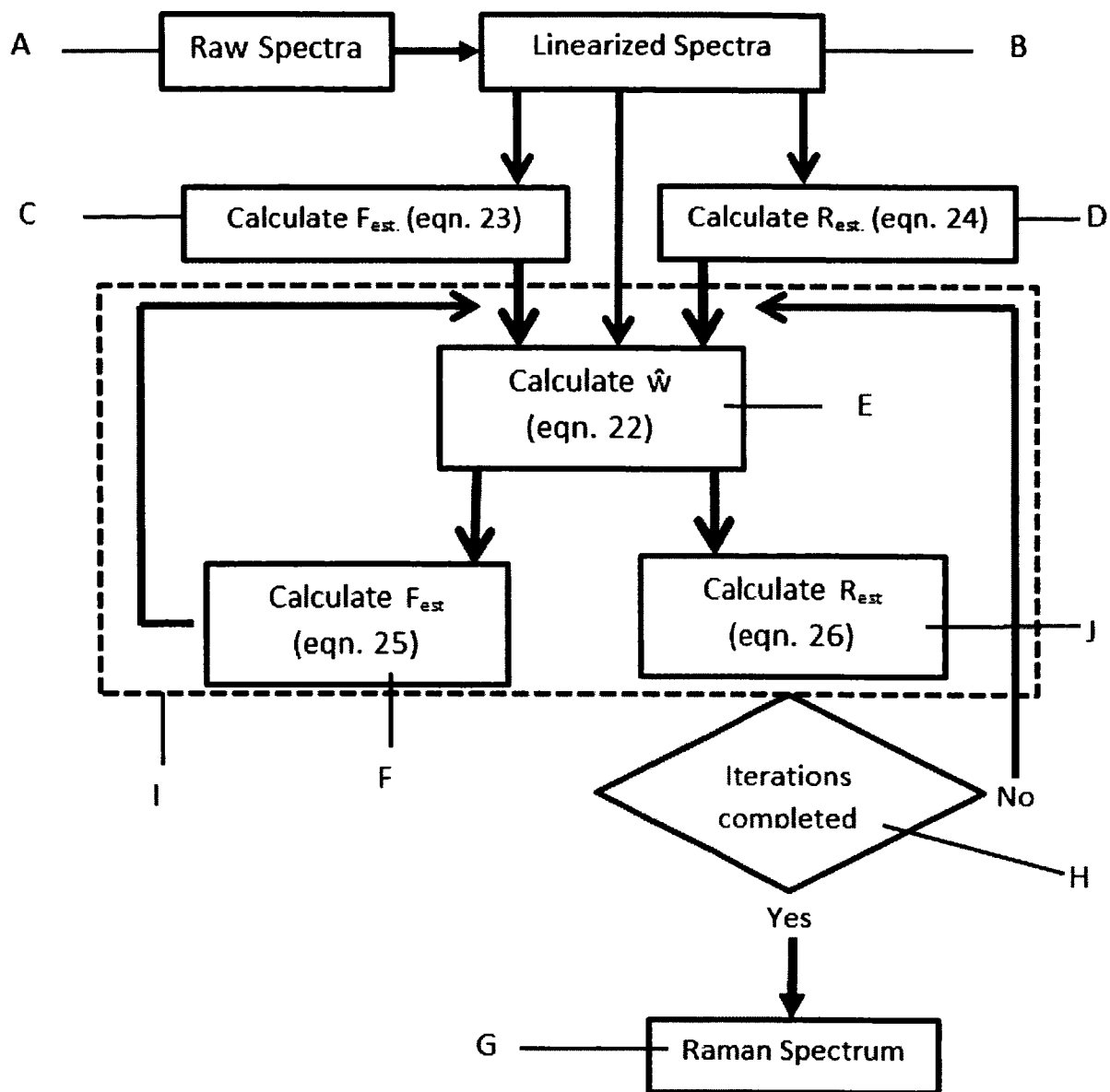


FIG. 11. This figure shows the overall method of processing the sequentially shifted excitation Raman spectra to obtain a background-free Raman spectrum. A) At least three SSE Raman spectrum acquired on the SSE Raman at different laser temperature. B) Preprocessing including axis linearization and applying SNV. The C and D) give the initial estimate of fluorescence and Raman, respectively. The E, F, and J) are iterated to a set number of iterations (2000 iteration). G) Gives the final pure Raman spectrum obtained from J after the completion of the 2000 iterations.

Since Eqn. 22 is based on element-wise math operations, it is instructive to the reader to display the element-wise operations in detail in another way. This is shown in Figure 12. As shown, the element-wise operations are indicated in matrix B for a shifted



excitation Raman spectra (R) consisting of N spectral positions and K excitations. Where N takes the values from 0 to n and K takes the values from 0 to k. The element wise addition of Eqn. 22 is depicted in matrix notation at C.

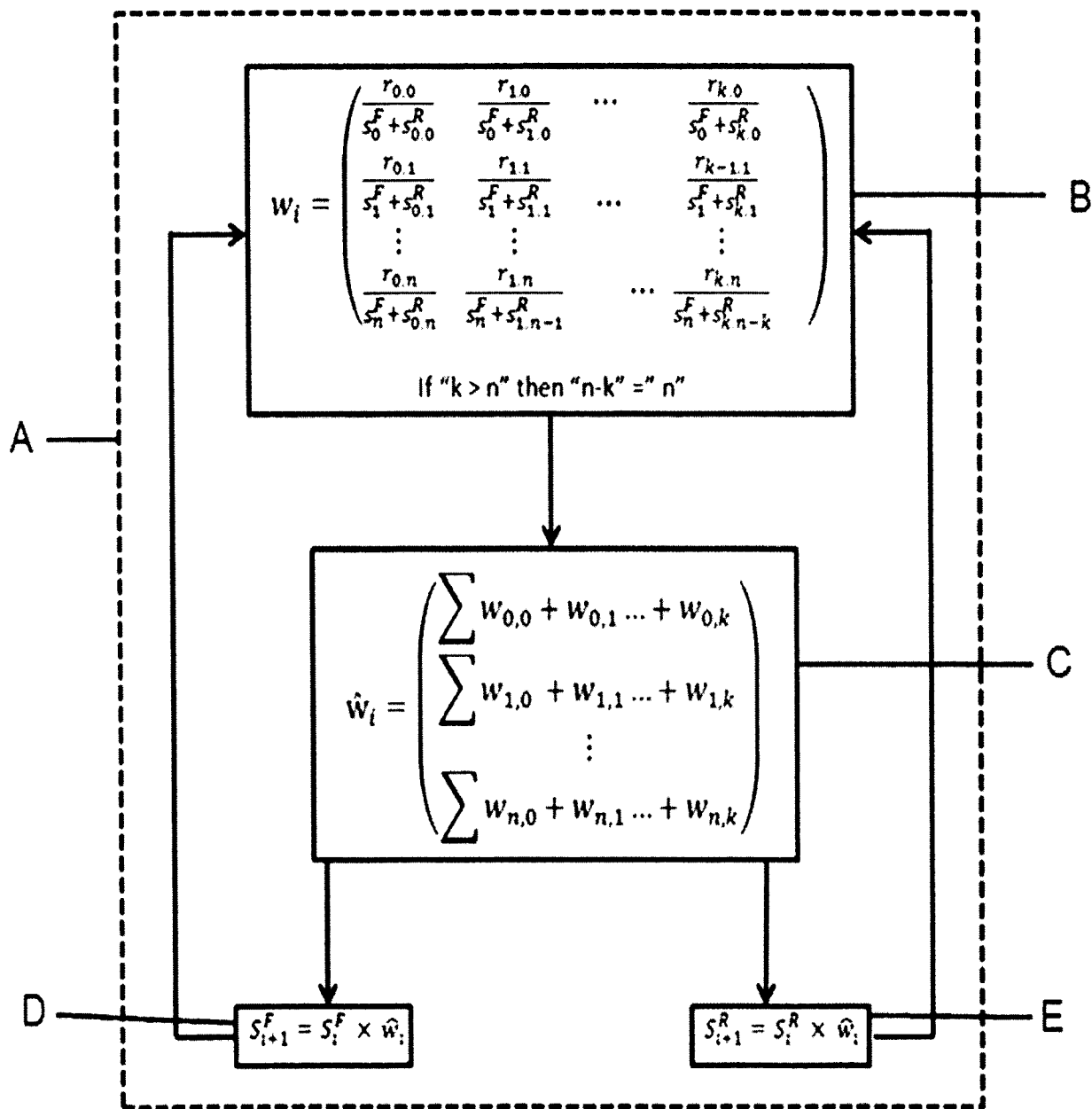


FIG. 12. This graph shows a detailed explanation of the SSE algorithm in matrix notation. B and C) give Eqn. 22 in matrix notation (calculations of the weighted vector). D and E) represent Eqn.s 25 and 26 and give the next fluorescence and Raman estimates, respectively. A) represents the number of iterations which is set to 2000.

As a further example, in Figure 13 a numerical example is given where the values at each spectral position for a three shifted excitation spectra are given (Fig.13A) and the spectra are graphed (Fig.13B). As can be seen, the SSE raw spectra have shifted Raman peaks in space while fluorescence background is not shifted. This represents the ideal scenario of the sequentially shifted excitation method; where Raman peak is shifted and fluorescence background is not at all. The ultimate goal of our algorithm is to differentiate between and separate the shifted Raman peaks from the unshifted fluorescence background.

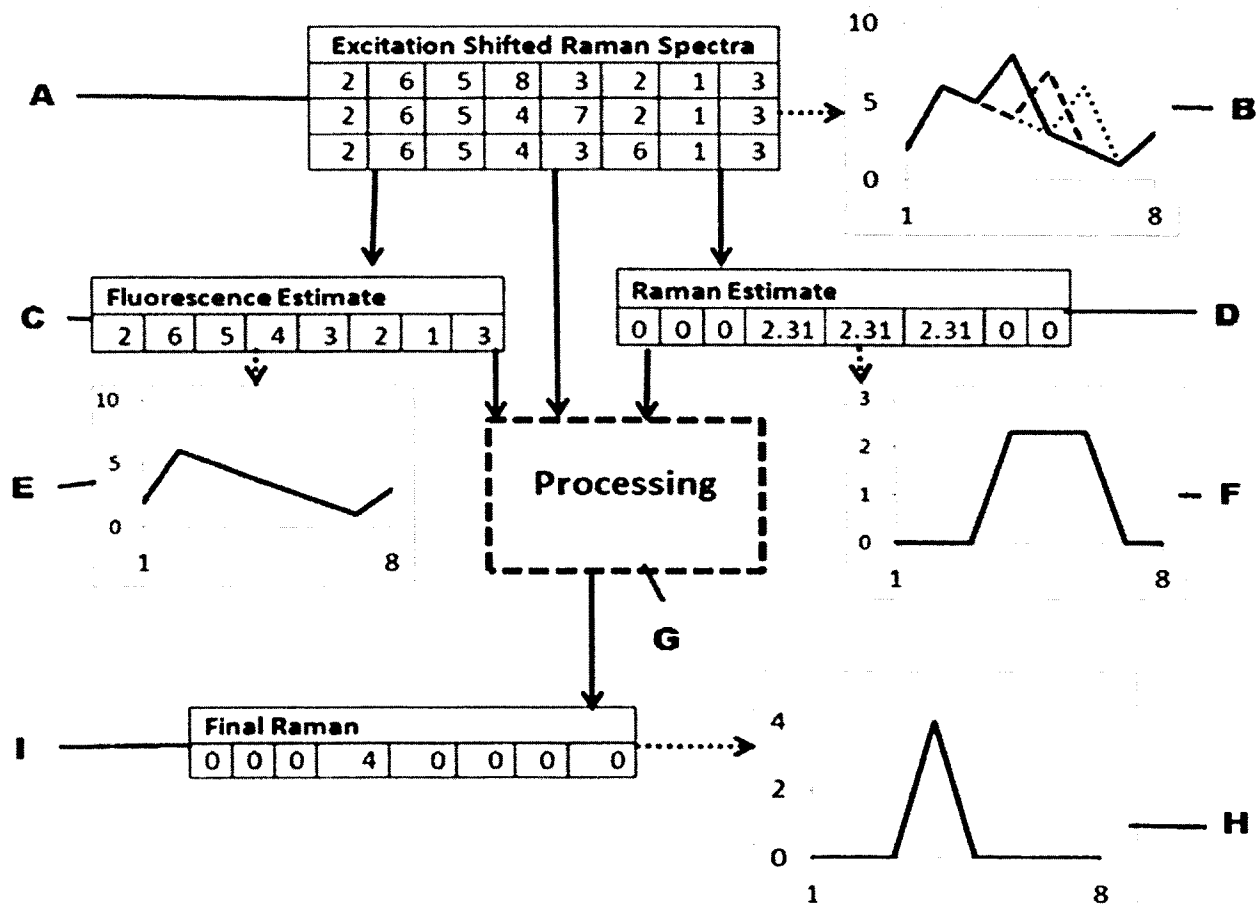


FIG. 13. This figure shows a numerical example of extracting a back-ground free Raman spectrum from the shifted excitation Raman spectra with graphical depictions of initial, intermediate, and final values.

As explained previously to get the SSE algorithm to start (Eqn.22) we have to have an initial estimate of Raman and fluorescence. The numerical values corresponding to the calculated initial fluorescence estimate (Fig.13C) using Eqn. 23 and the initial Raman estimate (Fig.13D) calculated using Eqn. 24 are given also and are plotted (Fig.13E and F), respectively. The numerical values for the final Raman spectrum (Fig.13I) are given and graphed (Fig.13H). In order to get the pure Raman spectrum shown in Fig. 13H, 100 iterations of the SSE algorithm, described in Figure 12 in details, were required.

The use of Eqn.s 22-26 or the use of Figure 11-13 to extract the Raman spectra from the fluorescence-background require that the spectra consist of uniformly spaced points which are proportional to energy (eg. wavenumbers). Methods to accomplish this requirement are well known and involve calibrating the spectra to an axis which is proportional to energy (e.g. wavenumbers). An additional requirement is that the differences in energy units between the various excitations (e.g. wavenumber difference) corresponds to the difference between consecutive spectral data points of the spectra or an integral multiple of the difference between consecutive spectral data points of the spectra. This requirement is accomplished by selecting the appropriate laser temperatures to obtain the desired excitation shifts. This is a significant advantage of the presented SSE design. For prior art which relies on the use of a series of different lasers to provide different excitations, it is not possible to accomplish this without great expense (either monetarily or computationally) since all lasers used must be manufactured to be integral values of the spectral spacing or the spectral data must be interpolated to a very small spacing so that each excitation shift can be represented by an integral number of spacings (difference in energy of consecutive spectral points).

## RESULTS AND DISCUSSION

### *LASER STABILITY*

To test the stability and the mode-hop of the 785 nm laser emissions, acetaminophen Raman spectra were acquired on the SSE Raman at different temperature. The acetaminophen spectra were acquired on the SSE at each temperature from 19 to 33 °C. The acquisition of the acetaminophen spectra were repeated thirty times. The acquisition time for each acquired acetaminophen spectra was set at 2 s. For each acquisition, the acetaminophen Raman peaks shift at each temperature was calculated relative to the lowest temperature acetaminophen Raman spectrum (19 °C). The stability of a five 785 nm lasers were tested. Figure 14 shows the plot between the acetaminophen Raman shift and the diode laser temperature for a two different lasers over the temperature range from 19 to 33 °C. The x-axis gives the change in the diode lasers temperature from 19 C° to 33 C° and the y-axis gives the Raman shift in units of wavenumbers. As shown in Figure 14, the Raman shift linearly increases as the diode laser one temperature increases. This indicates that there is no mode-hop in the case of laser one over the temperature range from 19-33 °C. However, Laser two exhibits a mode-hop (a jump change in the emitted wavelength) at temperature 28 °C. This indicates that laser two cannot be utilized in the sequentially shifted excitation mode at a temperature higher than or equal to 28 °C. Indeed, the suitable temperature range of operation can be picked from the current-voltage curve provided by the laser manufacturer.

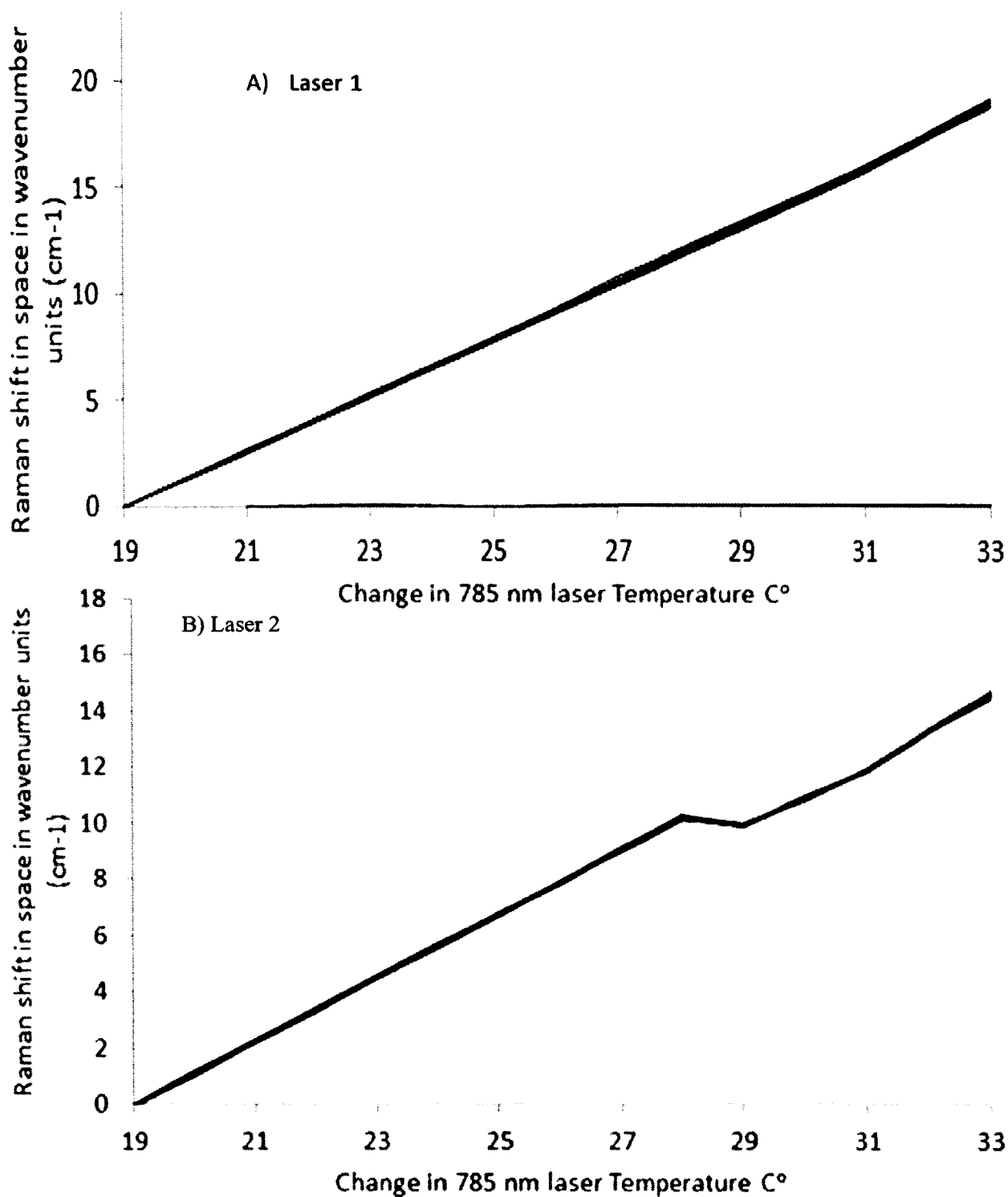


FIG. 14. This Figure shows the plot between acetaminophen Raman shift and the diode laser temperature for; (A) laser 1 which exhibits no mode-hop and (B) laser 2 which exhibits a mode-hop. The x-axis is expressed in Celsius scale from 19 to 33 °C and the y-axis is expressed in wavenumber units. The shift was calculated relative to the acetaminophen Raman spectrum acquired at the lowest temperature (19 °C). The acquisition time of each spectrum is 2 s.

*UTILITY OF THE SSE ALGORITHM IN REMOVING THE FLUORESCENCE BACKGROUND*

Figure 15 depicts the effect of the standard normal variate. The four dimethyl glyoxime Raman spectra acquired on the SSE Raman at the temperatures of 20, 23, 26, and 29 before and after applying SNV are shown in Figure 15A and 15B, respectively.

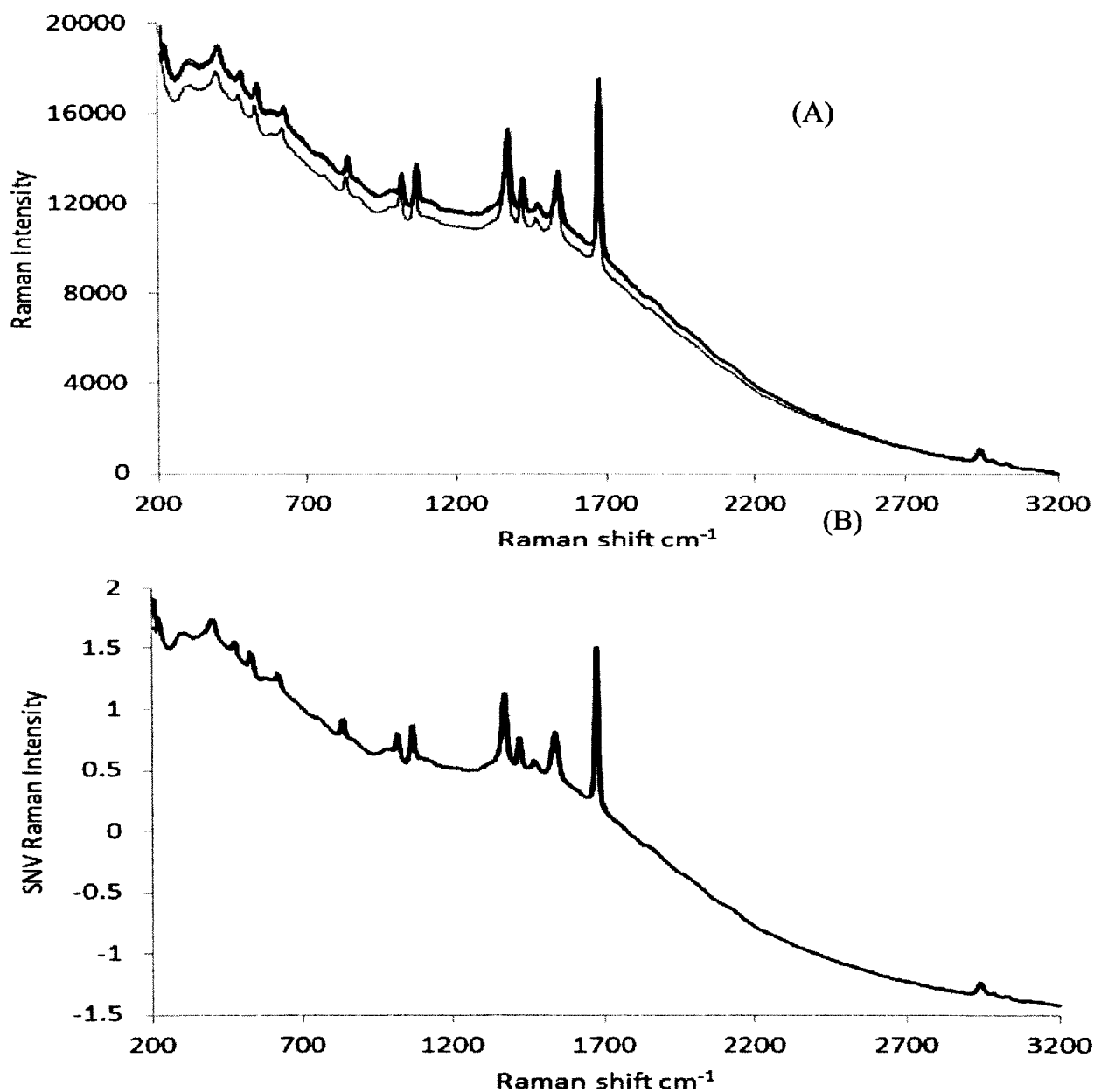


FIG. 15. Dimethyle glyoxime spectra acquired on the SSE Raman before applying (A) and after applying SNV (B). The x-axis is expressed in wavenumber Raman shift units.

As shown in Figure 15, performing SNV (Eqn. 13) removes the intensity difference between the SSE raw Raman spectra. However, SNV algorithm does not keep the original intensity of the SSE raw Raman spectra and produces negative intensity values. The later can be solved by subtracting the minimum from each intensity element.

Figure 16 shows the spectral data for the acquired Raman spectrum of dimethyl glyoxime acquired at four temperatures (20, 23, 26, and 29 °C) with a 785 nm DBR laser. Although it is difficult to resolve the spectra across the full spectral range, the change in laser temperature results in a shift of about  $3.5 \text{ cm}^{-1}$  for each of the spectra and the underlying background remains unchanged, as shown in the inset. Raman shifted obtained here is similar to the theoretical calculated one in the experimental section ( $3.6 \text{ cm}^{-1}$ ). SSE spectra shown in Figure 16 are pretreated by SNV.

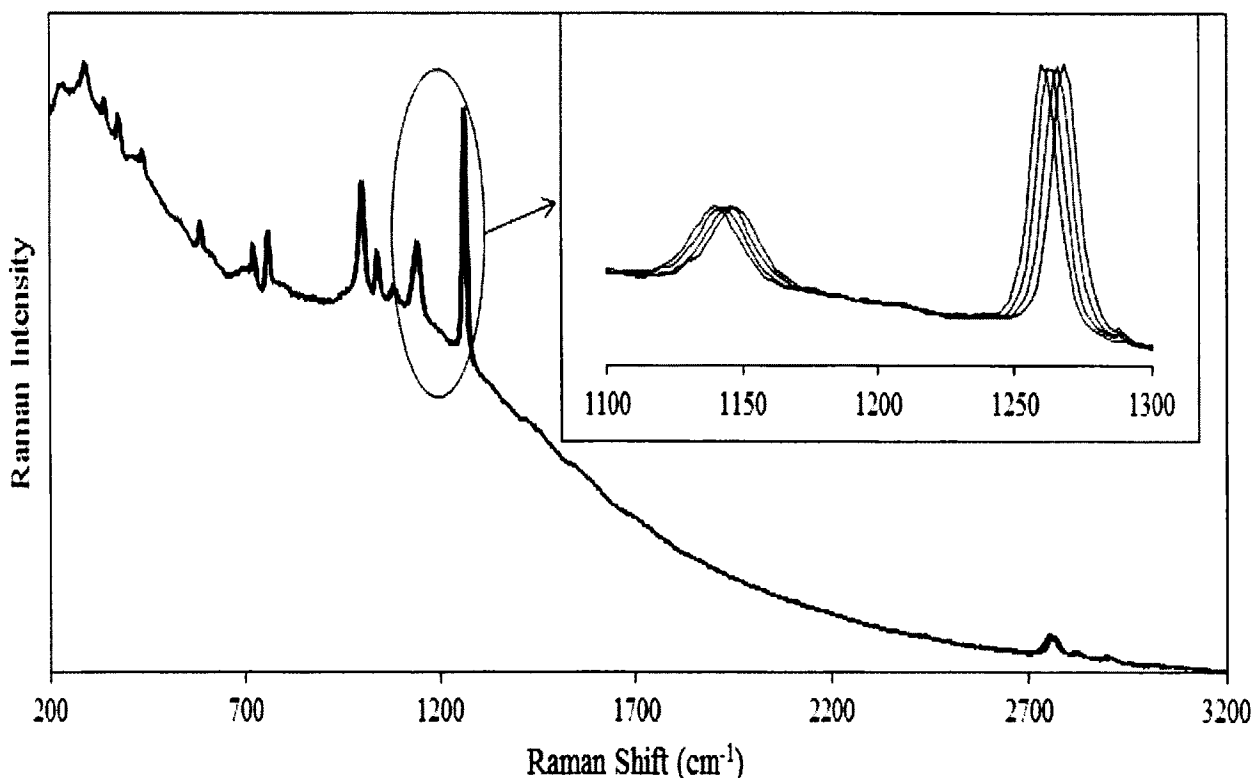


FIG. 16. Raman spectrum of dimethyl glyoxime acquired at four laser temperatures (20, 23, 26, and 29 °C) using a 785 nm DBR laser with an integration time of 0.8 s for each spectrum and a laser power of 50 mW. The change in laser temperature results in a shift of  $3.60 \text{ cm}^{-1}$  for each of the Raman spectra while the underlying background remains unchanged (as shown in the inset).

In Figure 17, the result of demodulating this data using Eqn. 22 and 26 is shown for various numbers of iterations. “A” gives the 0<sup>th</sup> iteration of Raman spectrum which corresponds to the standard deviation of the actual SSE Raman spectra at each point calculated using Eqn. 23. “B, C, D, and E” show the SSE Raman spectrum of dimethyl glyoxime at 10, 50, 100, and 1000 iterations, respectively.

In Figure 18, the result of demodulating this data using Eqn. 22 and 26 is shown along with the FT-Raman spectrum acquired using a 1064 nm laser. As can be observed, the two methods yield comparable results; both result in an absence of the fluorescence background. In the case of the FT-Raman spectrum, the absence of the fluorescence background is due to the long wavelength of the 1064 nm excitation laser. Although the results are comparable, the two experiments were carried out under dramatically different laser powers (800 mW for the FT-Raman and 50 mW for the SSE experiment). In addition, the total acquisition time for the FT-Raman spectrum was dramatically longer (120 seconds) compared to that of the SSE experiment (4 spectra each integrated for 800 ms). The slightly higher resolution of the FT-Raman spectrum is due to its acquisition at  $8\text{ cm}^{-1}$  resolution.



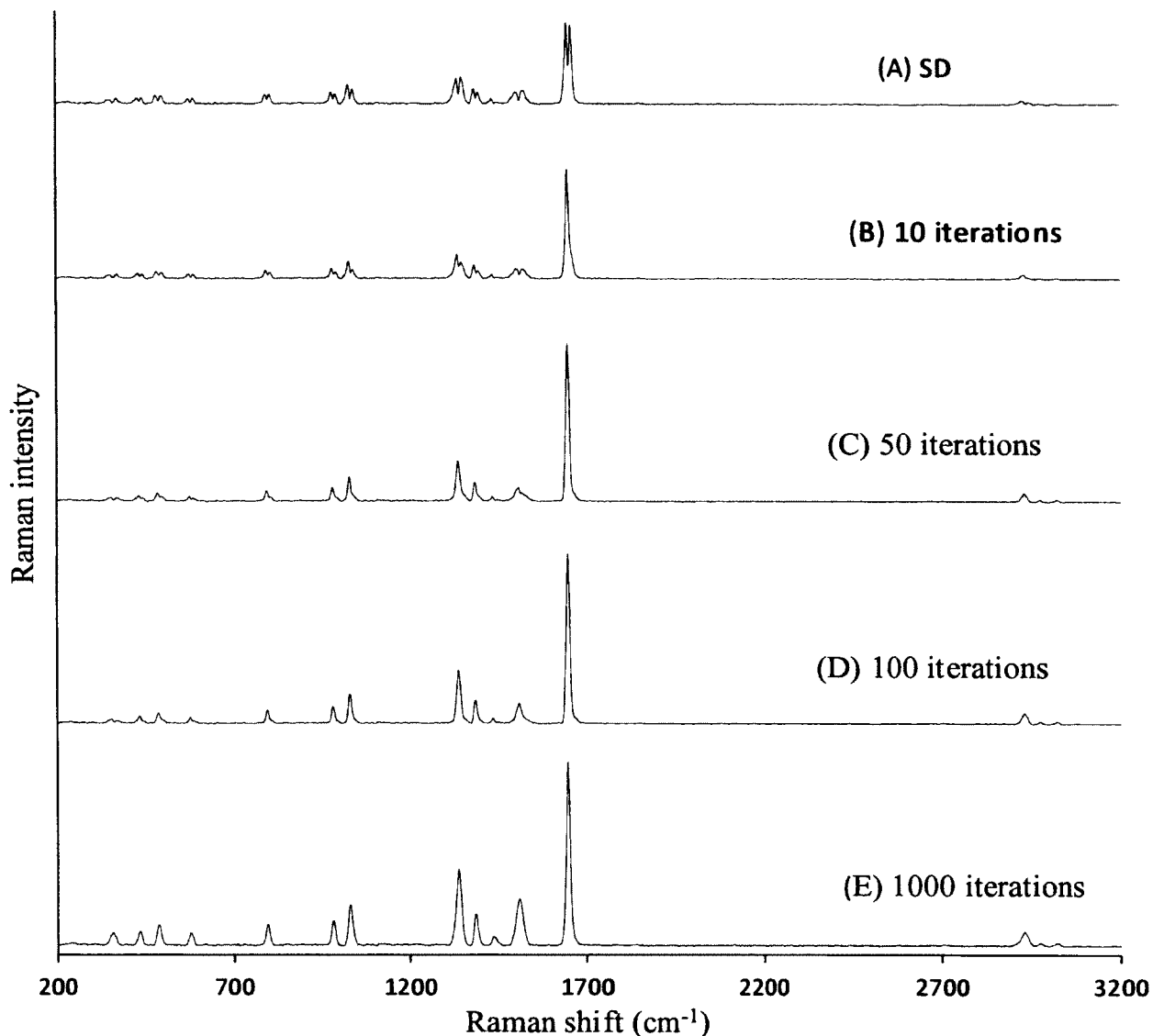


FIG. 17. This figure shows Raman spectra of dimethyl glyoxime extracted from shifted excitation Raman spectra using various numbers of iterations according to the previously described SSE algorithm; (A) standard deviation Raman spectrum (0 iterations), (B) 10 iterations, (C) 50 iterations, (D) 100 iterations, and (E) 1000 iterations.

Since neither instrument was corrected for spectral intensity throughput, the relative intensities of the peaks are not identical. This is particularly true when comparing the fingerprint region to the CH-stretching region. For the dispersive instrument, the silicon CCD detector has a minimum quantum efficiency for the CH-stretching spectral range (greater than 1020 nm), while for the FT-Raman spectrum, this range corresponds to the maximum quantum efficiency of the InGaAs detector used.

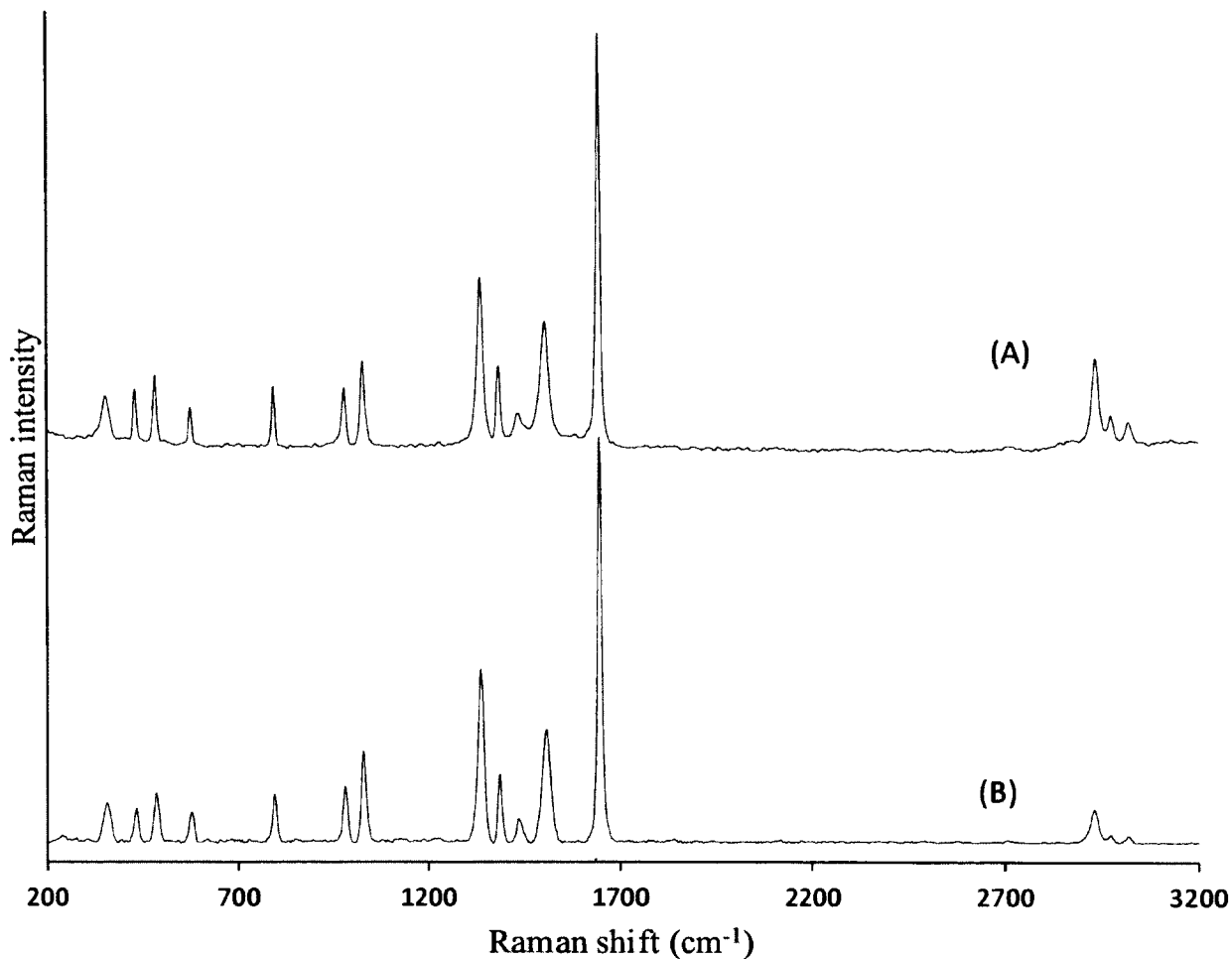


FIG. 18. This figure shows a graph comparing a FT-Raman spectrum to a Raman spectrum extracted from sequentially shifted excitation Raman spectra. (A) FT-Raman spectrum of dimethyl glyoxime acquired using a 1064 nm laser. The spectrum was acquired with an integration time of 120 s and a laser power of 800 mW. (B) Raman spectrum of dimethyl glyoxime which is obtained using the 4 sequentially shifted excitation Raman spectra of dimethyl glyoxime shown in Figure 16.

A second comparison is provided in Figure 19 with the spectra of catechol. Figure 19a corresponds to the FT-Raman spectrum of catechol (120 s acquisition and 800 mW); Figure 19b corresponds to the unprocessed dispersive Raman spectrum at 20 °C (2.5 s acquisition and 50 mW), and Figure 19c corresponds to the processed dispersive spectra (SSE) acquired at four laser temperatures (2.5 s x 4 acquisition and 50 mW). For the catechol sample, the fluorescence spectrum has a significantly different shape than that of the dimethyl glyoxime, but the dispersive Raman results are still comparable to the FT-Raman. It should however be noted that the FT-Raman spectrum has subtle background.

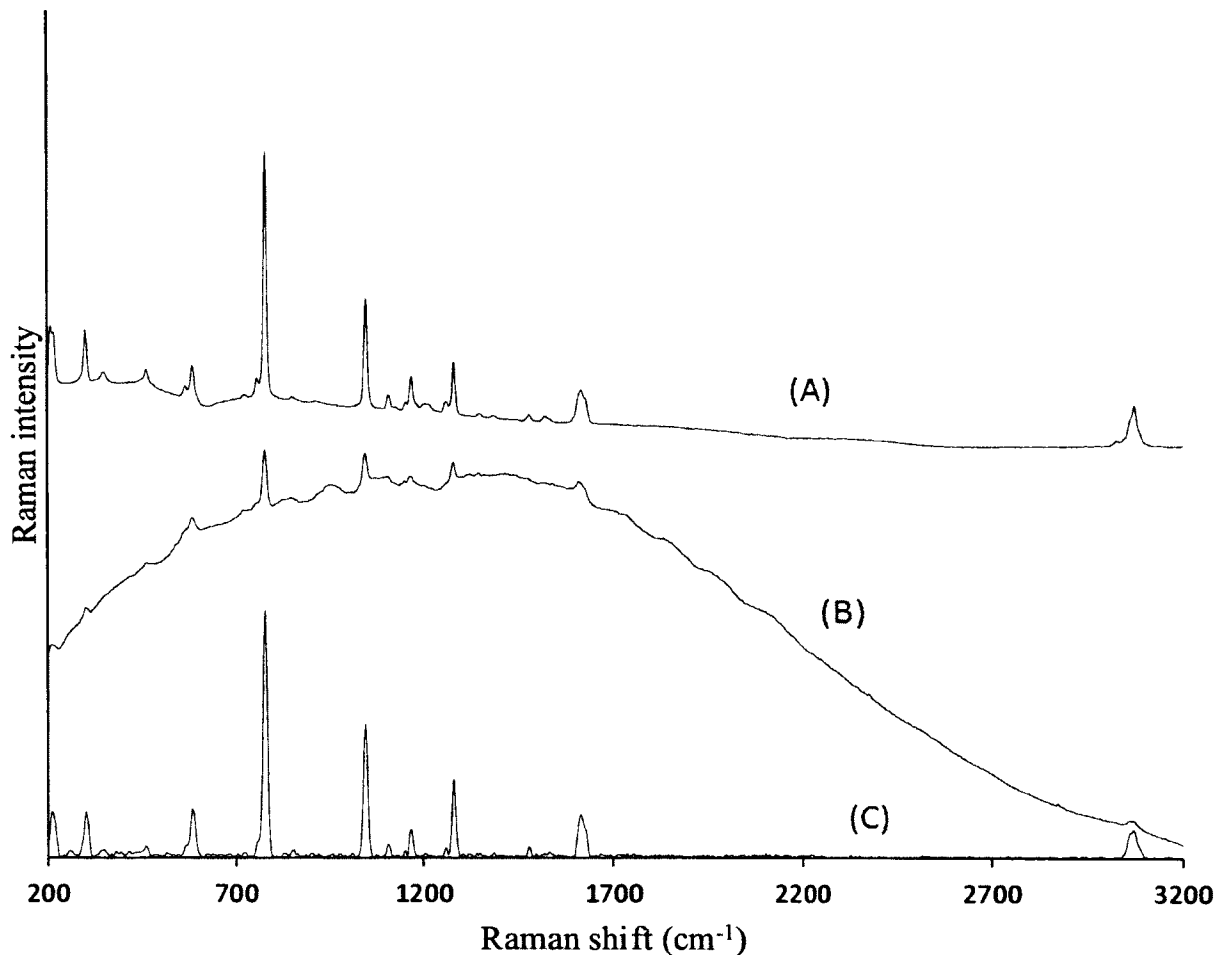


FIG. 19. This figure shows a graph comparing the SSE Raman spectrum of catechol to FT-Raman spectroscopy and dispersive Raman spectroscopy. A) FT-Raman spectrum of catechol acquired using 1064 nm laser (120 s acquisition and 800 mW). B) Dispersive Raman spectrum of Catechol acquired using 785 nm laser at 20 °C (2.5 s acquisition and 50 mW). C) Raman spectrum of Catechol obtained by processing sequentially shifted excitation Raman spectra acquired at three laser temperatures 23, 26, and 29 C (2.5 s x 3 acquisition and 50 mW).

A third comparison is given in Figure 20 for the spectra of magnesium carbonate. Due to the simplicity of the carbonate ion and its high symmetry, only a single Raman vibration is expected near 1180  $\text{cm}^{-1}$ . The high electronegativity of the oxygen atoms in the carbonate ion results in a large degree of polarization of the binding electrons which results in a very low Raman cross-section for the ion (i.e., small change in polarizability). This results in weak scattering and thus lower S/N for both the FT-Raman (Fig. 20a) and the SSE (Figure 20c) Raman spectra. For the raw dispersive

spectrum (Fig. 20b) the dominant spectral features are due to the fluorescence of the borosilicate NMR tube and weak emission lines due to the room lights. Since no effort was made to exclude room lights from the measurement, they were observed when weak signals were present. The SSE method results in the elimination of interference from both the fluorescence and from the room lights.

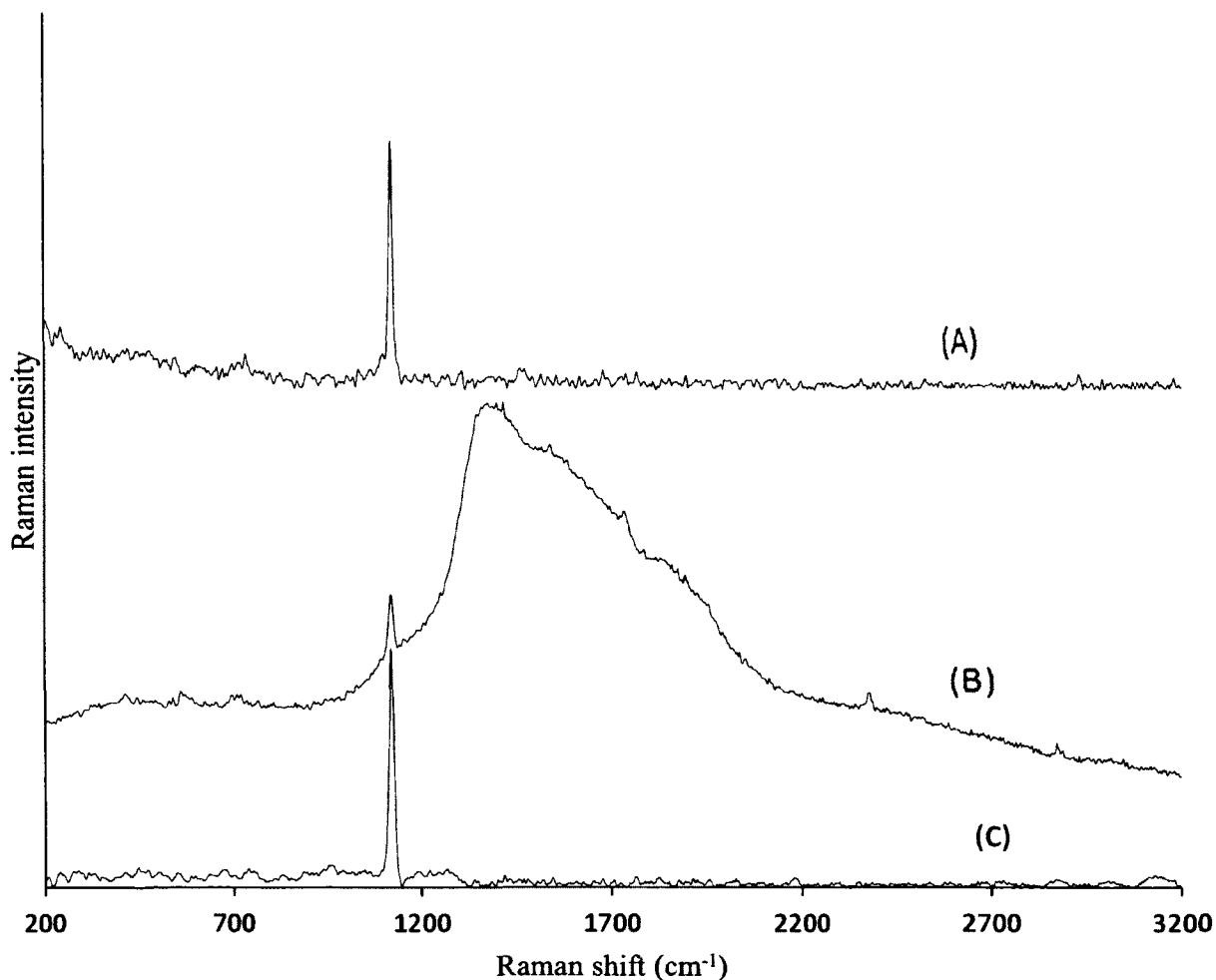


FIG. 20. This figure shows a graph comparing the SSE Raman spectrum of magnesium carbonate to FT-Raman spectroscopy and dispersive Raman spectroscopy. A) FT-Raman spectrum of magnesium carbonate acquired using 1064 nm laser (120s acquisition and 800 mW). B) Dispersive Raman spectrum of magnesium carbonate acquired using 785 nm laser at 20 °C (3 s acquisition and 50 mW). C) Raman spectrum of magnesium carbonate obtained by processing sequentially shifted excitation Raman spectra acquired at five laser temperatures: 20, 23, 26, 29, and 32 C (3 s x 5 acquisitions and 50 mW).

In Figure 21 the raw Raman spectra of a liquid diesel fuel sample are shown (Fig. 21A) along with the SSE Raman spectrum (Fig. 21B). The fluorescent background has a very high intensity at low wavenumbers, and drops off significantly at high wavenumbers. Also, the Raman spectrum has regions where the Raman spectrum is very weak ( $300\text{-}600\text{ cm}^{-1}$ ) as well as regions where there is significant S/N ( $1275\text{-}1675\text{ cm}^{-1}$ ). In Figure 21, an inset shows the  $300\text{-}600\text{ cm}^{-1}$  region with an expanded axis. It is easy to observe that the shifting Raman peaks have a very low intensity compared to the fluorescence background. In addition, the modulation of the fluorescence background in this region due to the response of the back-thinned CCD is quite large when compared to the weak Raman peaks. However, as shown in the SSE Raman spectrum, the algorithm is capable of extracting the peaks. The  $1275\text{-}1675\text{ cm}^{-1}$  region is also shown as an inset with an expanded axis and, as can be observed, the Raman peaks have significantly higher S/N. Although the Raman peaks can be observed to shift with changes in excitation wavelength, the broad underlying fluorescence remains essentially constant, giving rise to pseudo isosbestic points of constant intensity in the overlaid plots. In addition, by observing the strongest peak in this region, it can be observed that there is a component which shifts with changes to the excitation wavelength (Raman intensity) as well as a relatively sharp component which does not shift with changes to the excitation wavelength but rather decreases in intensity as the Raman peak shifts out from beneath it. This constant component is due to fluorescence intensity which is modulated by the back-thinned CCD. A comparison of this region with the SSE Raman spectrum clearly shows that only the shifted Raman peaks are retained. The superposition of the CCD response (particularly in the case of a back-thinned CCD) on an intense fluorescence signal would make it very difficult to remove the fluorescence background using just a background-removal algorithm on a single Raman spectrum since there are both broad components as well as sharp components to the fluorescence signal. In addition, using just 2 spectra (as with the case of SERDS) would also make it difficult to remove the sharp components since the intensity of the sharp fluorescence component is a function of the shifting Raman peak (this is not the case for a broad fluorescence background). The peak just below  $1475\text{ cm}^{-1}$  exhibits similar characteristics with a very weak Raman peak shifting

out from a much larger fluorescence peak. The corresponding region in the SSE Raman spectrum shows that only the small Raman peak is retained.

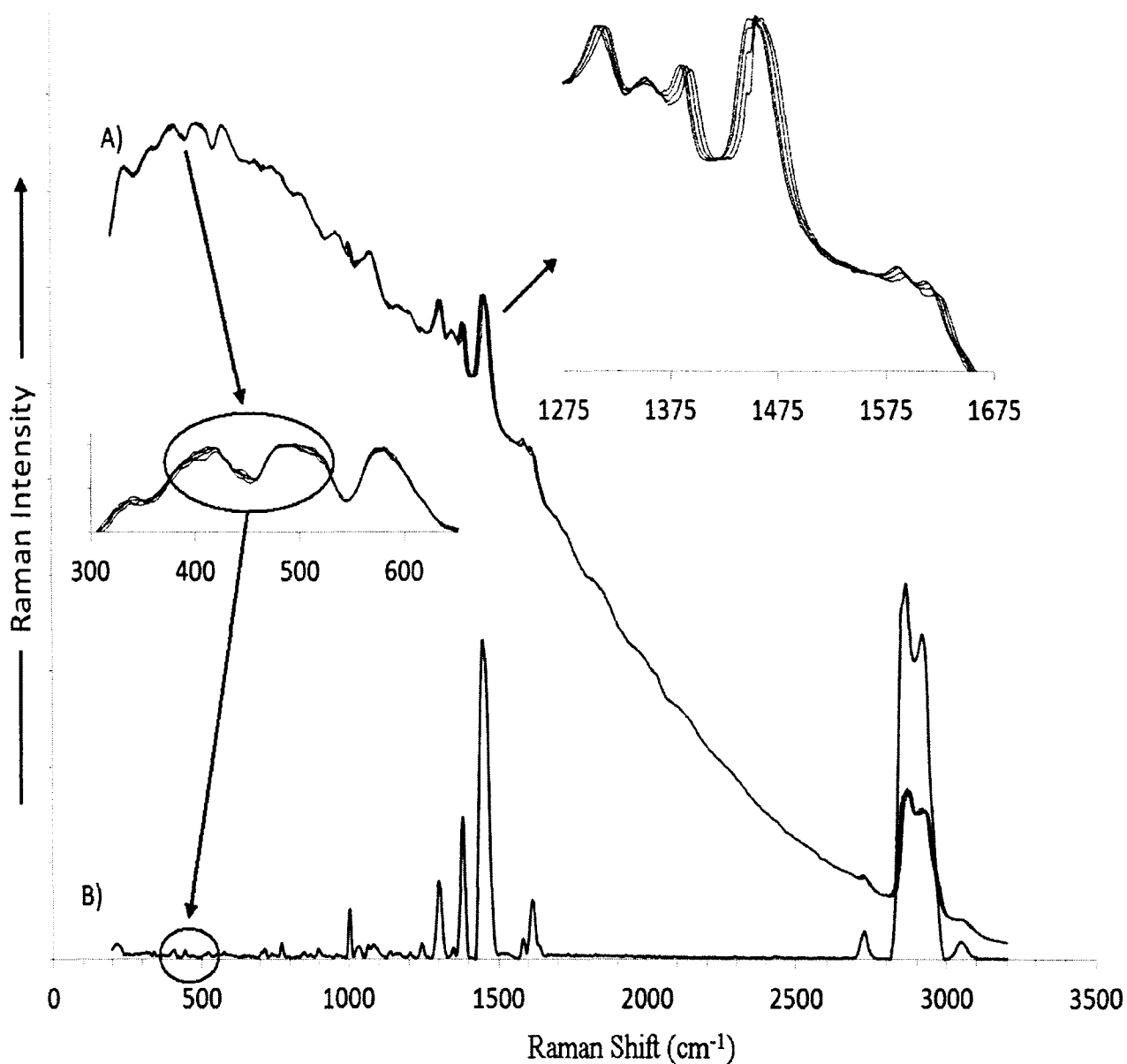


FIG. 21. The Raman spectra of a diesel fuel are shown: A) 4 overlaid raw unprocessed Raman spectra and B) the SSE Raman spectrum. Two expanded regions of the raw overlaid Raman spectra are shown in the insets. In one of the insets, a circle indicates the corresponding spectral region in the SSE Raman spectrum. The raw Raman spectra were acquired at four temperatures (22, 26, 30, and 34 °C) with an acquisition time of 15 s for each spectrum.

A more complex sample is given in Figure 22 for the Raman spectra of ceric ammonium sulfate. Figure 22a corresponds to the FT-Raman spectrum (120 s and 800

mW); Figure 22b corresponds to the unprocessed dispersive Raman spectrum at 20 °C (4 s and 50 mW), and Figure 22c corresponds to SSE Raman acquired at four laser temperatures (4 s x 4 acquisition and 50 mW). For Figure 21b, the 785 nm dispersive spectrum consists almost solely of fluorescence. In addition, the fluorescence of the f-block lanthanide ion is structured with peaks that have bandwidths which approximate those expected for Raman peaks for both the FT-Raman and dispersive Raman instruments. The inset in Figure 21b magnifies the lower part of the fingerprint region by a factor of 30 so that the Raman peaks can be observed. Although background effects are still present, a one to one correspondence of these peaks with those in Figure 22c can be observed (SSE Raman spectrum). For the FT-Raman, there is also some correspondence; the region close to  $200\text{ cm}^{-1}$  is very intense and broad relative to both the raw and SSE spectra however. In addition, the sulfate peak at  $\sim 1000\text{ cm}^{-1}$  is much broader than expected for the FT-Raman spectrum (see inset for Figure 22a) and must have a fluorescent contribution. As expected, the fluorescence of the cerium ion at 1064 nm is much different than that at 785 nm. It follows that the broad peak at  $1250\text{ cm}^{-1}$  and weak spectral features near  $1700\text{ cm}^{-1}$  observed for the FT-Raman spectrum correspond to fluorescence also. This is further justified by their absence in Figure 22c when using SSE Raman. Although typical broad fluorescent backgrounds are easily discerned by the eye, a significant advantage of shifted excitation methods of Raman over FT-Raman is the additional ability to discern sharp fluorescent features which have bandwidths which approach that of the Raman spectrum. Additionally, the fluorescence from such samples cannot be treated with baseline removal techniques such as polynomial fitting or derivatives.

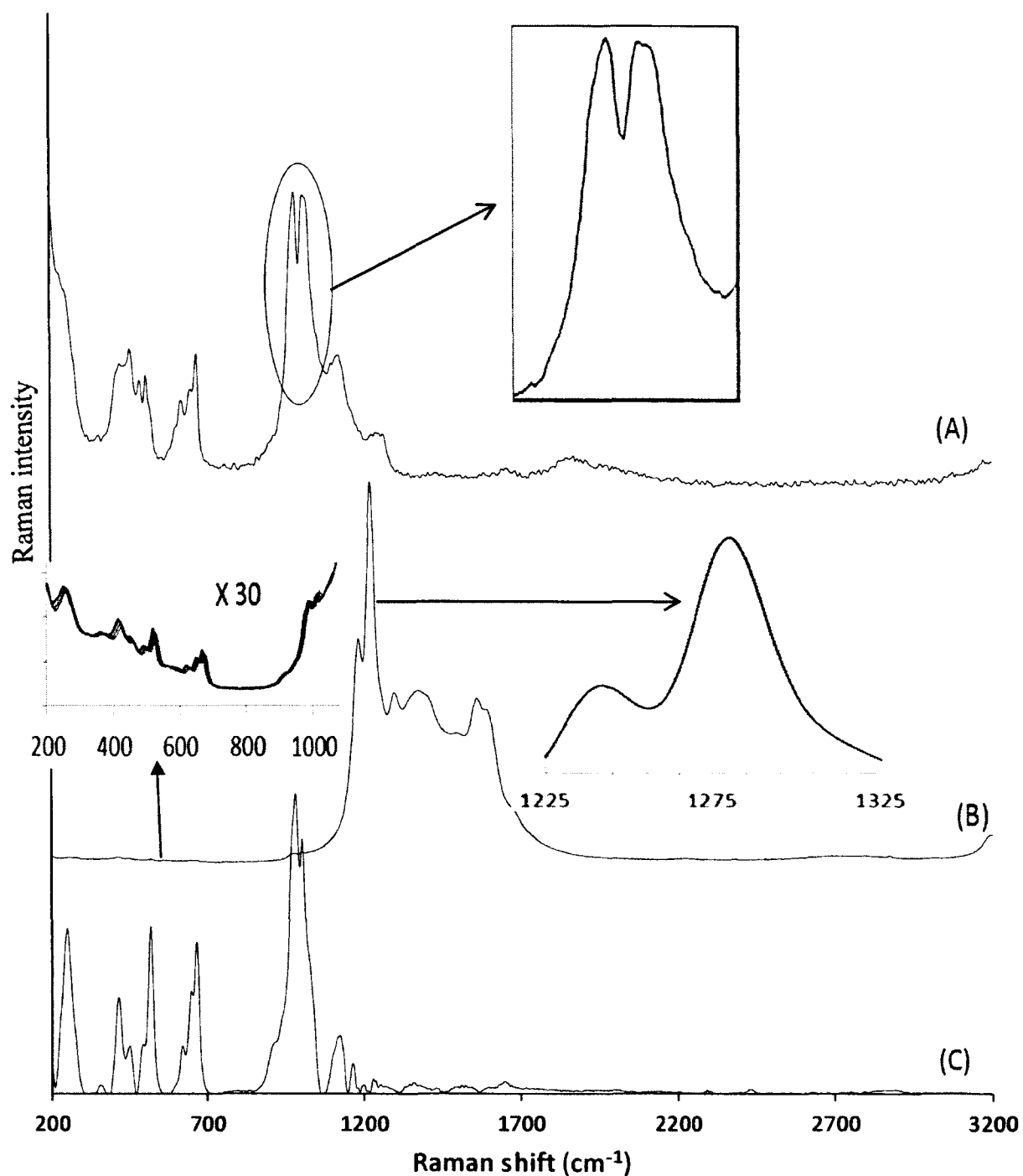


FIG. 22. This figure shows a graph comparing SSE ceric ammonium sulfate Raman spectrum to FT-Raman spectroscopy and dispersive Raman spectroscopy. A) FT-Raman spectrum of ceric ammonium sulfate acquired using 1064 nm laser (120s acquisition and 800 mW). B) Dispersive Raman spectrum of ceric ammonium Sulfate acquired using 785 nm laser at 20 °C (4 s acquisition and 50 mW). C) Raman spectrum of ceric ammonium Sulfate obtained using sequentially shifted excitation Raman spectra acquired at five laser temperatures: 17, 20, 23, 26, and 29 °C (4 s x 5 acquisitions and 50 mW).



The inability of FT-Raman to eliminate all of the fluorescence for some samples is exemplified in Figure 23a for the spectra of acenaphthylene (cyclopentanaphthalene). For the FT-Raman, there is a significant, broad, undulating background even with 1064 nm excitation due to the polycyclic aromatic nature of the molecule. For the unprocessed 785 nm dispersive Raman spectrum, the fluorescence is several orders of magnitude more intense than that of the Raman. Despite this intense fluorescence, the processed four-temperature SSE data effectively eliminates the fluorescence while providing comparable signal to noise when compared to the FT-Raman spectrum. The same observations can be made in Figure 24 for the spectra of 4-bromo-N,N-dimethylaniline. For these spectra, the signal to noise of the SSE Raman spectrum is slightly less than that of the FT-Raman spectrum especially in the CH stretching region. This is due to the much lower laser power and integration time for the SSE measurement and the quantum efficiency of the CCD detector. The FT-spectrometer optical power was 800 mW and the acquisition time is 120 s for both acenaphthylene and 4-bromo-N,N-dimethylaniline. While, the optical power of the 785 nm laser used in the SSE is 50 mW and the acquisition time is 4 s and 10 s for acenaphthylene and 4-bromo-N,N-dimethylaniline, respectively. The four unprocessed SSE Raman spectra of acenaphthylene and 4-bromo-N,N-dimethylaniline were acquired at the laser temperature of 20, 23, 26, and 19 °C. The three most successful alternative approaches are: standard deviation spectra, SERDs, and PCA extraction. As a comparison, these methods are also included in Fig. 23 and 24 using the same SSE data that was used for the SSE Raman spectra (Fig. 23C and 24C).

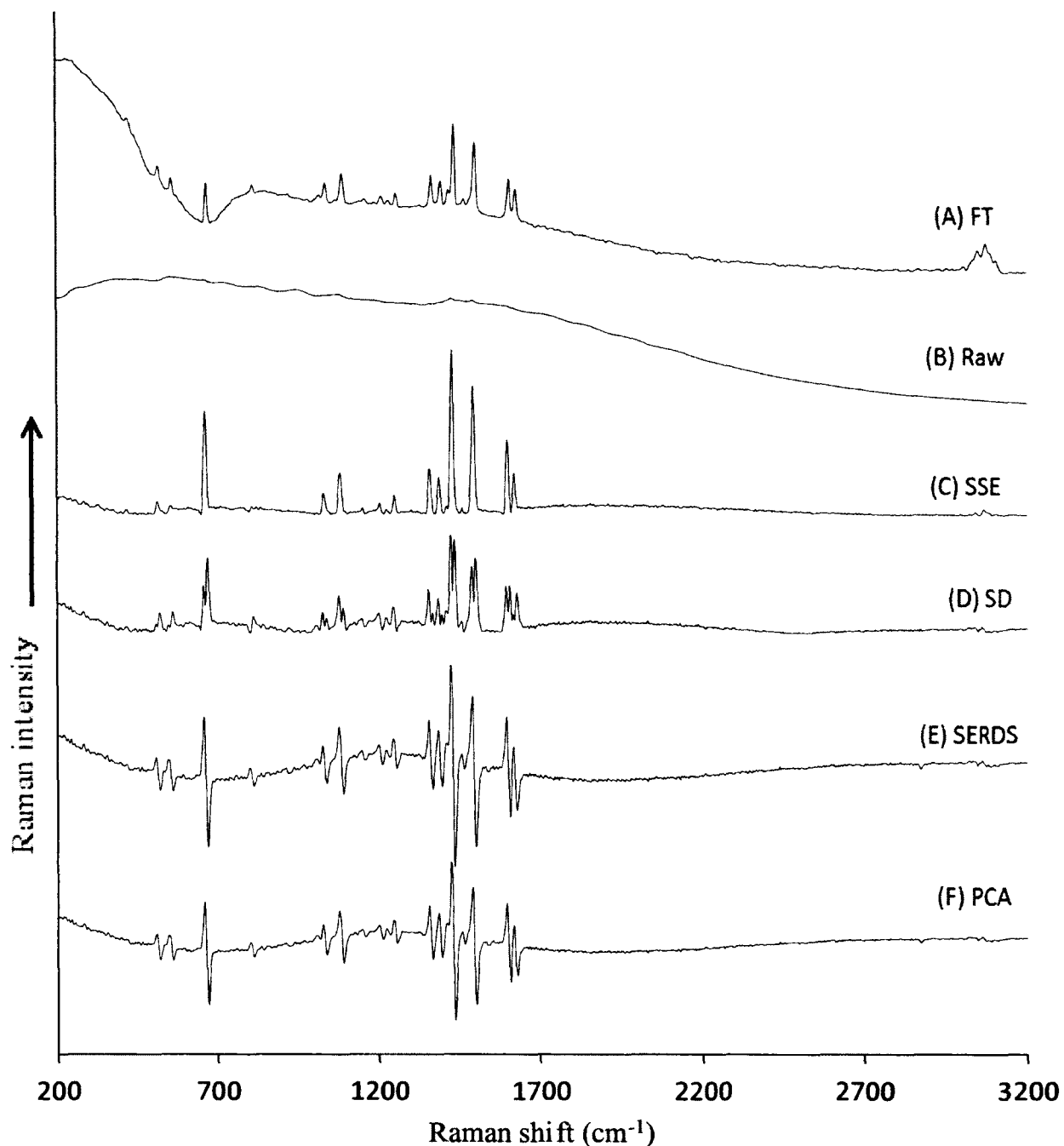


FIG. 23. This figure shows a graph comparing dispersive Raman spectroscopy, SERDS and FT-Raman spectroscopy to SSE Raman of acenaphthylene. (A) FT-Raman spectrum acquired using 1064 nm laser (120 s acquisition and 800 mW). (B) Unprocessed dispersive Raman acquired using 785 nm laser at 20 °C (10 s acquisition and 50 mW). (C) Processed dispersive Raman spectra (SSE). (D) Standard deviation Raman spectrum. (E) SERDS spectrum is obtained by acquiring only two spectra (20 and 29 °C and 10 s x 2 acquisition). (F) PCA Raman spectrum. All C, D, and F were acquired at four laser temperatures (10 s x 4 acquisition and 50 mW).

### *SSE ALGORITHM MERITS*

As described previously, many methods have been used to extract Raman spectra from spectral data acquired using multiple excitations. The three most successful alternative approaches are: standard deviation spectra, SERDs, and PCA extraction. As a comparison, these methods are also included in Fig. 23 and 24 using the same SSE data that was used for the SSE Raman spectra (Fig. 23C and 24C). For the SERDS calculation, only two spectra are required, so the best results are displayed (this always corresponded to the largest shift between spectra, ie, 20 C and 29 C). For acenaphthylene (Figure 23), all methods give comparable signal-to-noise, but the SSE Raman in Figure 23C results in the lowest background followed closely by the standard deviation spectrum. However, the standard deviation spectrum is difficult to interpret due to the doubling of peaks arising from the nature of the calculation. Both the SERDs and the PCA spectra result in similar yet higher levels of background. As mentioned, it is easier to observe the noise component in the dispersive spectra acquired for 4-bromo-N,N-dimethylaniline (Figure 24). The higher S/N for the SSE Raman can easily be seen relative to SD, SERDs, and PCA. In addition, the SERDs and PCA methods still suffer from the fact that the remaining background results in a baseline which makes it difficult to retrieve the true Raman spectrum from the processed spectral data without using additional algorithmic background filtering methods.

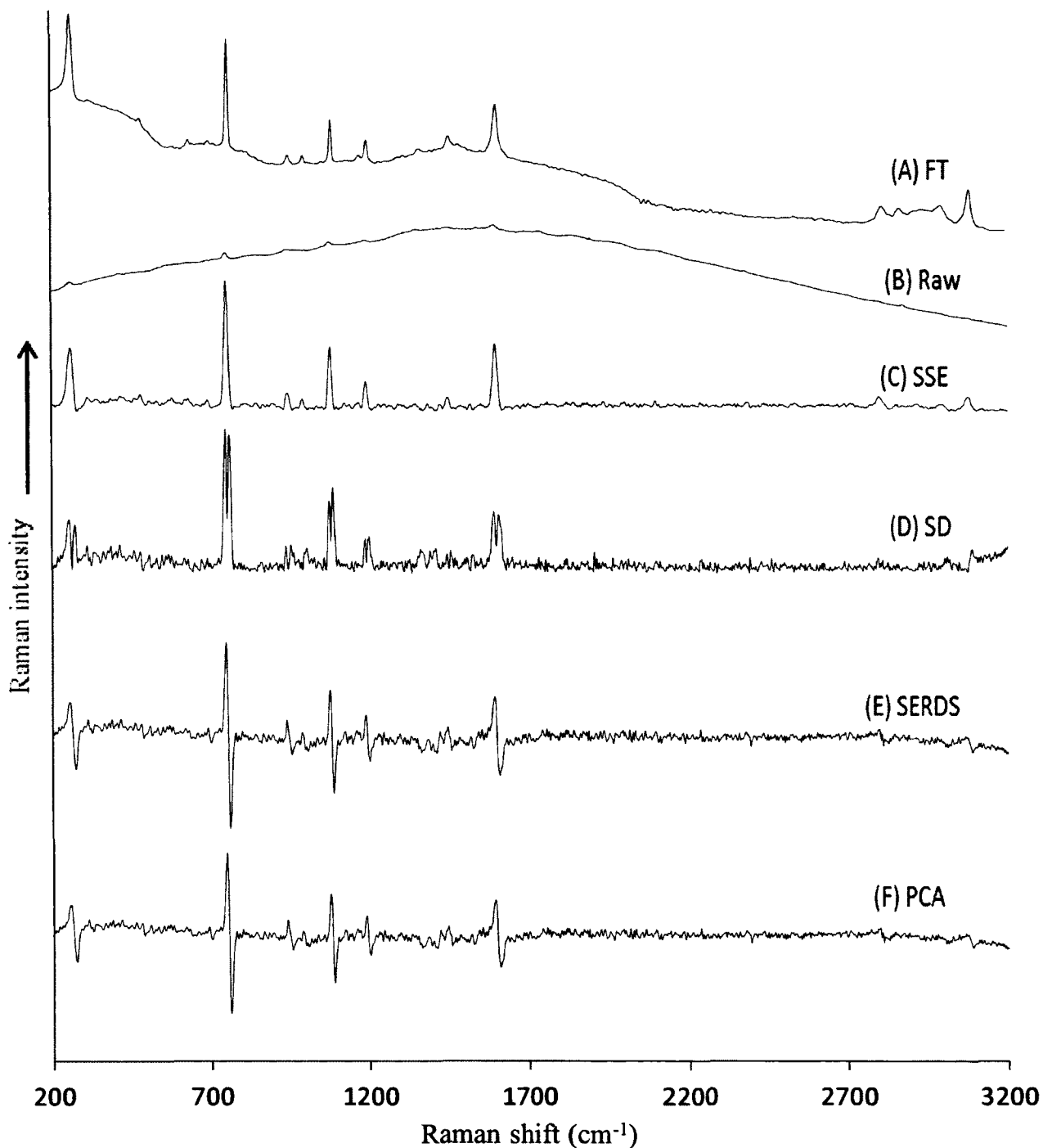


FIG. 24. This figure shows a graph comparing dispersive Raman spectroscopy, SERDS and FT-Raman spectroscopy to SSE Raman for 4-bromo-N, N-dimethyl-aniline (A) FT-Raman spectrum acquired using 1064 nm laser (120 s acquisition and 800 mW). (B) Unprocessed dispersive Raman acquired using 785 nm laser at 20 °C (5.2 s acquisition and 50 mW). (C) Processed dispersive Raman spectra (SSE). (D) Standard deviation Raman spectrum. (E) SERDS spectrum is obtained by acquiring only two spectra (20 and 29 °C and 5.2 s x 2 acquisition). (F) PCA Raman spectrum. All C,D, and F were acquired at four laser temperatures (5.2 s x 4 acquisition and 50 mW).

An additional advantage of SSE over the SERDS and PCA methods lies in the fact that the SSE experiment can be carried out under a wide range of excitation conditions without impacting the resolution of the Raman spectrum result. For both SERDS and PCA methods, however, the bandwidth of the resulting derivative peaks is a function of the excitation spacing. This is demonstrated in Figure 25 where two different excitation profiles are used. The first excitation profile consists of acquiring the spectra at 20, 23, and 26 °C; while the second excitation profile consists of acquiring spectra at 20, 26, and 32 °C. The change in the excitations for the experiment results in a change in peak bandwidth for the PCA data but not for the SSE data. Although not shown, the effect is even more dramatic for SERDS.

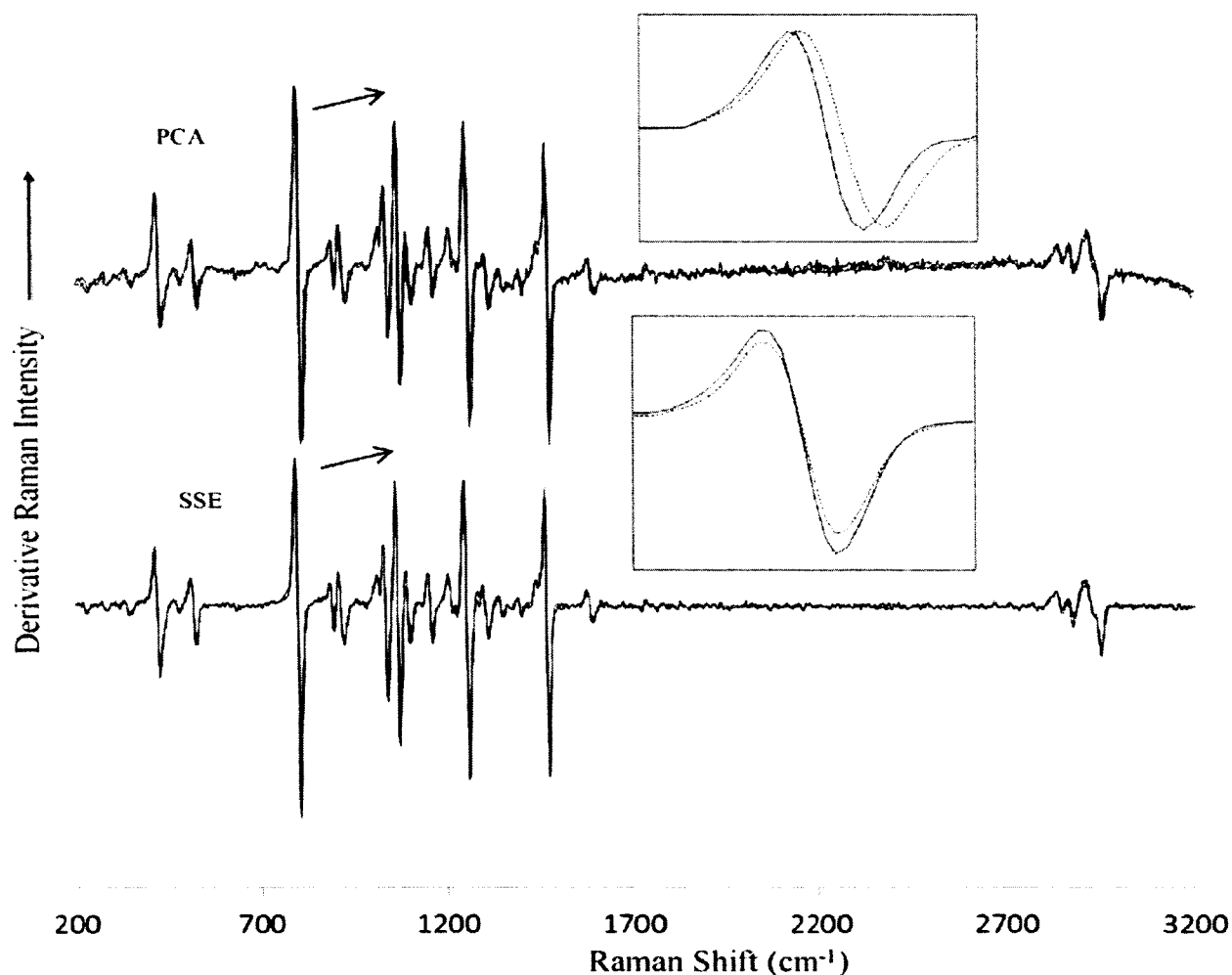


FIG. 25. Raman spectra of tris(hydroxymethyl)aminoethane acquired at two temperature profiles; 20, 23, and 26 °C (solid line), and 20, 26, 32 °C (dotted line). The acquisition time was 1 s x 3. A) PCA modulated Raman spectrum. B) Derivative of SSE Raman spectrum.

The primary advantage of the SSE algorithms (Eqn. 22-26) over Eqn. 20, is its execution speed since both methods utilize a Lucy-Richardson iterative method. For the current work, we have also found that using the standard deviation spectrum of the raw spectra as the initial estimate of  $S^R$  for the SSE method results in faster and more reliable convergence as opposed to using the maximum intensity of the shifted spectra at each spectral position. We have used the SSE method for over 500 chemicals and in each case convergence can be obtained using 2000 iterations. For samples with low levels of fluorescence or no fluorescence, this convergence is often reached with fewer iterations, although additional iterations offer no penalty in Raman quality. For samples with more dominating fluorescence, although 2000 iterations is sufficient, using iterations of 3000 or more often result in a further improvement in signal to noise and baseline removal. For the present work, all of the SSE data has been presented with 2000 iterations and a method for auto-detection of convergence has not been used.

In conclusion, the primary advantage of the SSE Raman method with the described instrumentation lies in its universal ability to extract Raman spectra from fluorescence interference using inexpensive and compact instrumentation in a time efficient manner. The SSE method offers superior S/N performance when compared to shifted excitation methods which do not use iterative algorithms. With respect to the previous reports using iterative algorithms and shifted-excitation, the SSE processing is 3 orders of magnitude faster due to its computational dependence on  $KN$  (compared to previous results which are dependent on  $KN^2$ ), resulting in sub-second processing times compared to processing times exceeding a minute. If derivative data is desired, it is possible to calculate the derivative using conventional algorithms and any changes in the resulting bandwidth are solely dependent on the algorithm used. This is not the case with SERDS or PCA methods whose derivative bandwidths are dependent on the experimental conditions (excitations chosen). In addition, SSE Raman is the only method which allows all of the excitation conditions to be easily varied in a predefined manner (number of excitations, separation of excitations, and integration time of each excitation) in order to obtain optimal results. We have begun to investigate the impact of these parameters in order to gain a better understanding of optimal conditions for varied samples. As with other

Raman extraction methods involving more than two excitations, even in the absence of a background, SSE Raman improves the S/N ratio by reducing random shot and thermal noise and by eliminating fixed pattern and random spike noise. In the current work, this allows an un-cooled CCD detector to be used. The use of a cooled CCD would allow much longer integration times to be used for weak Raman scatters without having the thermal noise impact the dynamic range of the 16 bit A/D converter.

Importantly, the ability to tune the DBR laser to set wavelengths allows the excitation shift to be set to a constant value. This simplifies implementation of the iteration algorithms since the shift index of either the operator matrix (Eqn. 20) or the weighting vector (Eqn. 22-26) can be set equal to the constant excitation shift (or an integral number multiplied times the excitation shift) and thus allow a single algorithm to be used even for different instruments. The only previous description of using shifted excitations with an iterative algorithm consisted of using eight single-mode excitation lasers with fixed wavelengths of 782.6 nm, 784.1 nm, 784.4 nm, 786.8 nm, 788.6 nm, 790.7 nm, 793.6 nm, and 794.3 nm with three or more of the lasers being used to collect the Raman spectra. As presented, the use of multiple excitations of varying separations requires that the data be splined (or fitted) so that the spectral spacing is equal to the lowest common denominator of all excitation shifts which are used if equivalent results to SSE are to be obtained. Such a procedure results in an increase in N and is particularly ineffective when using Eqn. 20 with calculation times which are proportional to  $KN^2$ .

## INSTRUMENT CALIBRATION

*AXIS CALIBRATION*

There are two ways to calibrate the x-axis of a dispersive Raman device. These are the use of the absolute wavenumber calibration using the emission line spectra of gases, such as neon and argon or a mixed of both, and the use of the Raman shift scatterers standards (e.g., cyclohexane, acetaminophen, acetonitrile, and benzonitrile). The choice of a standard depends on the number of well-established Raman peaks of the standard on the wavenumber range of interest. The Raman standards scatterers are a better choice since the excitation source wavelength does not have to be measured independently. Acetaminophen was chosen to calibrate the SSE Raman instrument since it has more than 20 Raman peaks that span the wavenumber range we are interested in (200 to 3200  $\text{cm}^{-1}$ ). The acetaminophen spectrum acquired on the SSE Raman (solid line) and the FT-Raman spectrometers (dotted line) are shown in Figure 26.

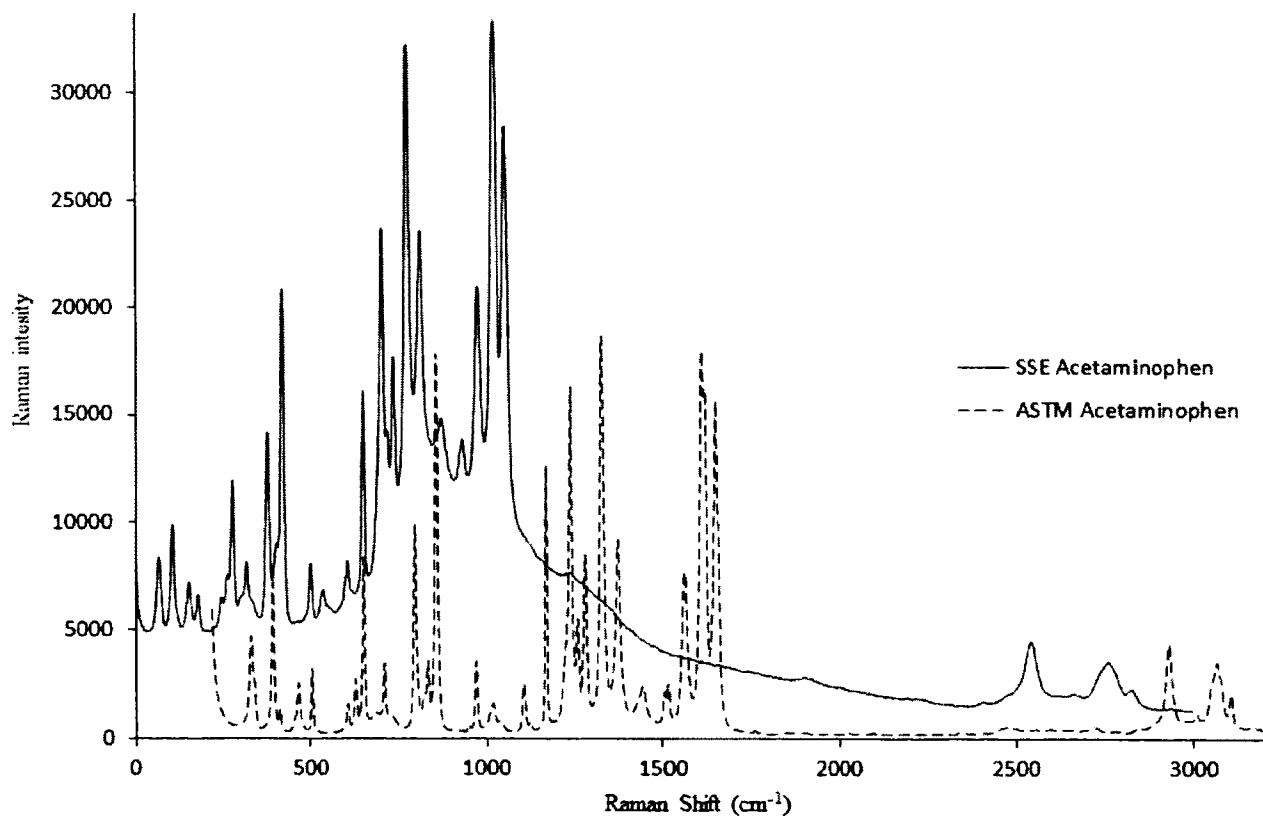


FIG. 26. Acetaminophen spectrum acquired on the SSE Raman before axis calibration (solidline) overlaid with acetaminophen spectrum acquired on the FT-Raman spectrometer (dotted line). The x-axis is expressed in wavenumber Raman shift units ( $\text{cm}^{-1}$ ).



The acquisition time of the SSE acetaminophen Raman spectrum is 5 s. The FT acetaminophen spectrum has a calibrated axis due to the internal interferometer calibration supplied by a built-in helium-neon laser. As shown in Figure 26, the acetaminophen spectrum acquired on the SSE Raman has different Raman shifts axis (in wavenumber units) than the calibrated FT-acetaminophen spectrum. It is obvious to notice the nonlinearity between the two axes where the difference is about  $200\text{ cm}^{-1}$  in CH-stretching region and about  $500\text{ cm}^{-1}$  in the fingerprint region.

Sixteen acetaminophen peaks were used to calibrate the x-axis of the SSE Raman. The acetaminophen peaks used in calibration are shown in Figure 27.

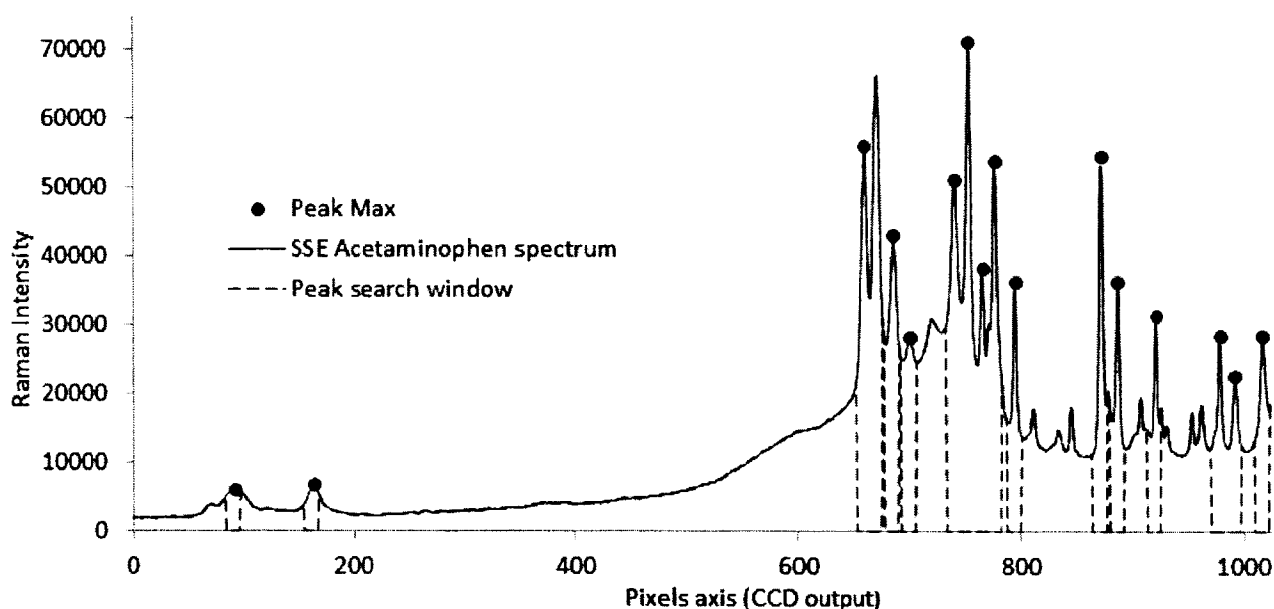


FIG. 27. This figure shows the sixteen peaks of acetaminophen used in the SSE axis calibration. A window of size 13 indices was assign around each peak location to automate its future picking. The x-axis is expressed in unit of pixels (CCD output).

The sixteen peaks were chosen so that they cover the whole spectral range we are interested in. The peaks were chosen based on their locations, shapes, consistency, and intensity. In order to automate the calibration procedure, software was written to automatically pick the peaks of interest. A window of 13 indices size was assign around each peak's location for its future picking. The 13 indices window of each peak is then spline a 100 times and the spline-window maximum is used as the peak maximum. As soon as a pick maximum is found its entire window is replaced by zeros in order to

prevent overlapping and picking the same peak twice.

The ASTM acetaminophen peaks in wavenumber ( $\text{cm}^{-1}$ ) units were changed to wavelength (nm) since there is a better linearity using a wavelength axis.

TABLE 2. The sixteen peaks of acetaminophen used in calibration of single-laser SSE in pixels along with their corresponding ASTM Raman shift in wavelength (nm) units.

SSE peaks locations (pixels)	ASTM peaks locations(nm)
92.21	1153.046
163.88	1135.562
660.88	991.262
686.73	982.676
700.5	978.314
740.58	964.771
753.33	960.322
765.96	956.151
776.93	952.378
794.9	946.268
872.39	919.24
886.8	914.129
920.76	902.119
978.31	881.428
991.51	876.644
1015.7	867.793

To calibrated the single-laser SSE axis to the FT-Spectrometer axis, a third order polynomial was generated between the acetaminophen peaks locations, acquired on the SSE Raman, and their corresponding ASTM Raman shift in absolute wavelength units. The wavelength of the sixteen SSE peaks in pixels, acquired on the SSE, along with their corresponding ASTM Raman shift in wavelength (nm) units are shown in table 2. The 3<sup>rd</sup> order polynomial is shown is Figure 28.

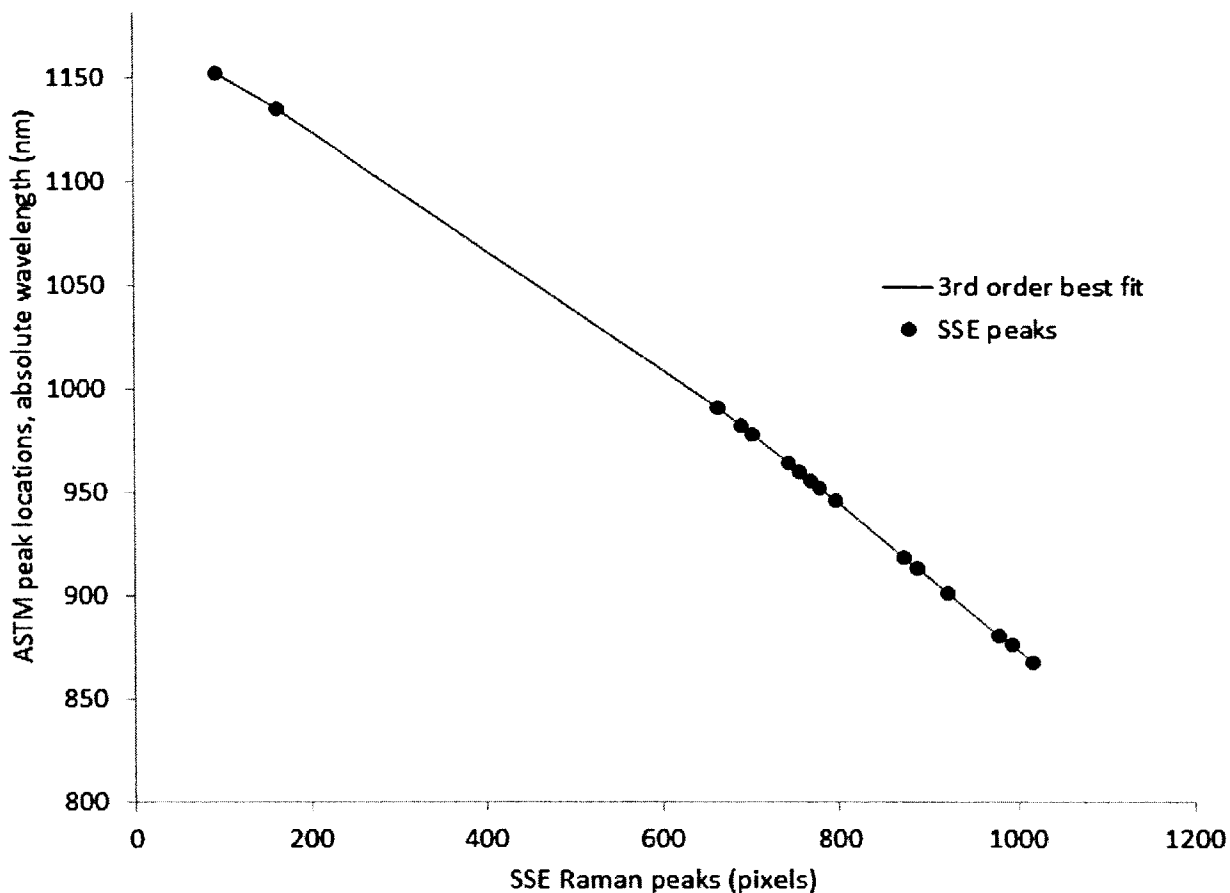


FIG. 28. This figure shows a 3<sup>rd</sup> order polynomial between the sixteen acetaminophen peaks locations in pixels acquired on the single-laser SSE and their ASTM certified wavelength Raman shift (solid line).

Using a third order polynomial produces the minimum error in the SSE axis. The third order polynomial coefficients were used to generate a calibrated x- axis ( $\lambda_{cal}$ ) for the SSE handheld Raman by using the following Eqn. 27;

$$\lambda_{cal}(i) = a_0 + a_1 x(i)^1 + a_2 x(i)^2 + a_3 x(i)^3 \quad (27)$$

where “i” gives the pixels index and it takes the values 0 to 1023 (CCD detector has 1024 pixels),  $a_n$ 's are the 3<sup>rd</sup> order polynomial coefficients, and  $\lambda_{cal}$  is the generated and calibrated wavelength Raman shift axis (nm) of the SSE. The SSE Raman shift in (wavelength nm unit) axis is then converted to wavenumber axis ( $\text{cm}^{-1}$ ).

An acetaminophen spectrum acquired on the SSE Raman after the x-axis calibration is plotted with the FT-acetaminophen spectrum in Figure 29. As shown in

Figure 29, the two spectra have the same x-axis in the units of wavenumber Raman shift ( $\text{cm}^{-1}$ ). The acetaminophen SSE spectrum has a little fluorescence background since it is shown before applying the fluorescence-removing algorithm.

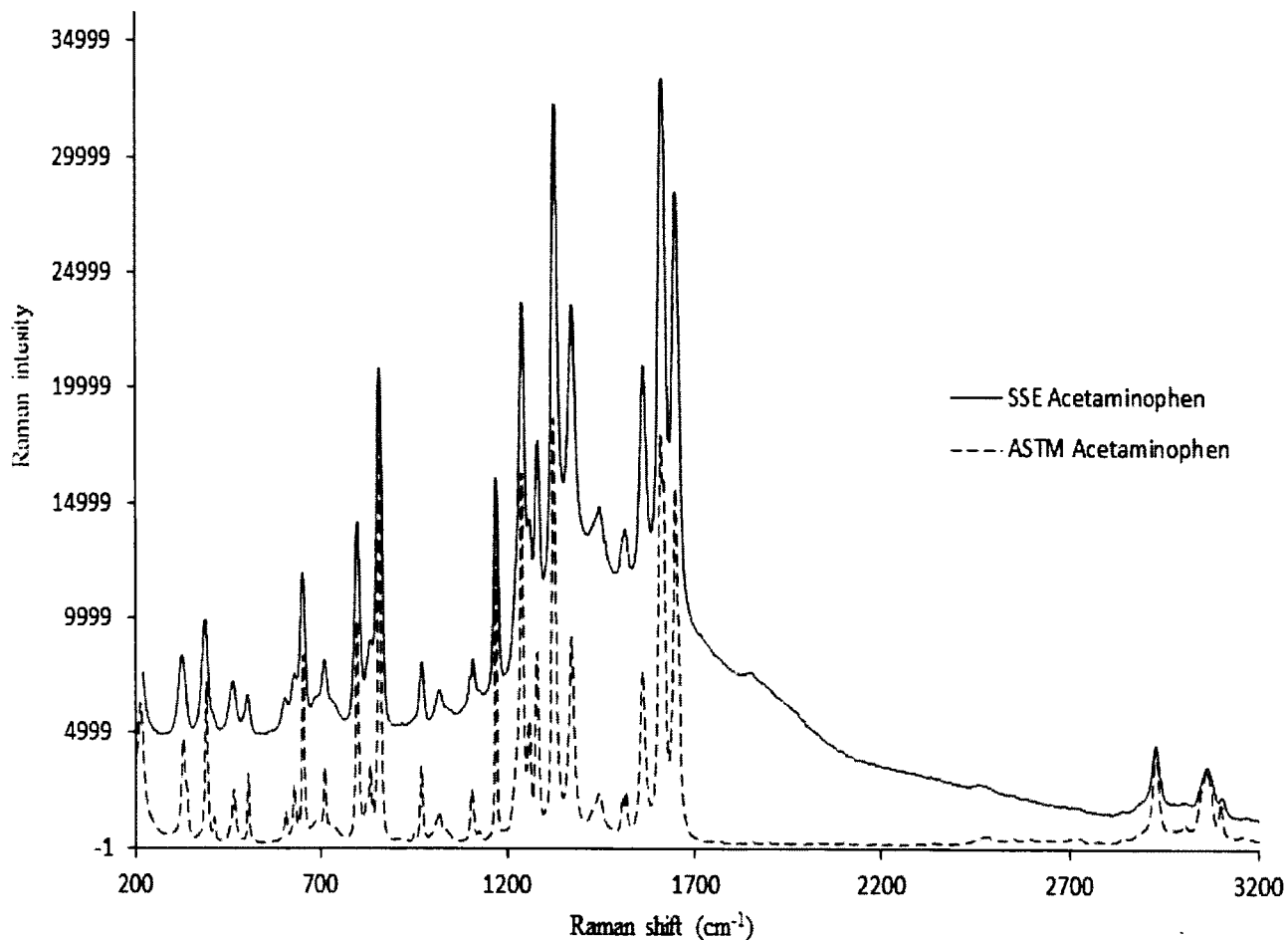


FIG. 29. Acetaminophen spectrum acquired on the SSE Raman after calibration (solid line) overlaid with the FT-acetaminophen spectrum (dotted line). The x-axis is expressed in unit of wavenumber Raman shift ( $\text{cm}^{-1}$ ).

### *INTENSITY CALIBRATION (Y-AXIS)*

Raman spectra acquired on an instrument are usually corrected in order to implement an existing Raman library database or in the case of inter-libraries transfer. Standard Reference Materials 2241 (SRM) is a certified spectroscopic standard for Raman spectra correction acquired on a Raman spectrometer utilizes a 785 nm excitation source. The SRM 2241 consists of an optical glass which emits a broadband luminescence spectrum when irradiated with a 785 nm excitation source. The standard SRM 2241 emission spectrum can be generated by using a 5<sup>th</sup> order polynomial coefficients (provided by NIST, national institute of standards and technology) and the Raman shift axis in wavenumber ( $\text{cm}^{-1}$ ) units of the instrument to be calibrated. The five coefficients of SRM 2241 provided by NIST are;  $a_0 = 9.71\ 937\ \text{E}-02$ ,  $a_1 = 2.28\ 325\ \text{E}-04$ ,  $a_2 = -5.86\ 762\ \text{E}-08$ ,  $a_3 = 2.16\ 023\ \text{E}-10$ ,  $a_4 = -9.77\ 171\ \text{E}-14$ , and  $a_5 = 1.15\ 596\ \text{E}-17$ . The standard SRM 2241 emission spectrum ( $Y_{\text{SRM}}$ ) can be computed using the following Eqn. 28:

$$Y_{\text{SRM}}(\Delta\nu) = a_0 + a_1 \times (\Delta\nu)^1 + a_2 \times (\Delta\nu)^2 + a_3 \times (\Delta\nu)^3 + a_4 \times (\Delta\nu)^4 + a_5 \times (\Delta\nu)^5 \quad (28)$$

where  $(\Delta\nu)$  is the wavenumber axis index in unit of Raman shift ( $\text{cm}^{-1}$ ) and  $a_n$ 's are the coefficients listed above. Eqn.28 is valid only in the spectral range from  $\Delta\nu = 200\ \text{cm}^{-1}$  to  $3500\ \text{cm}^{-1}$ . The standard spectrum of the SRM 2241 generated using the SSE wavenumber Raman shift axis is shown in Figure 30.

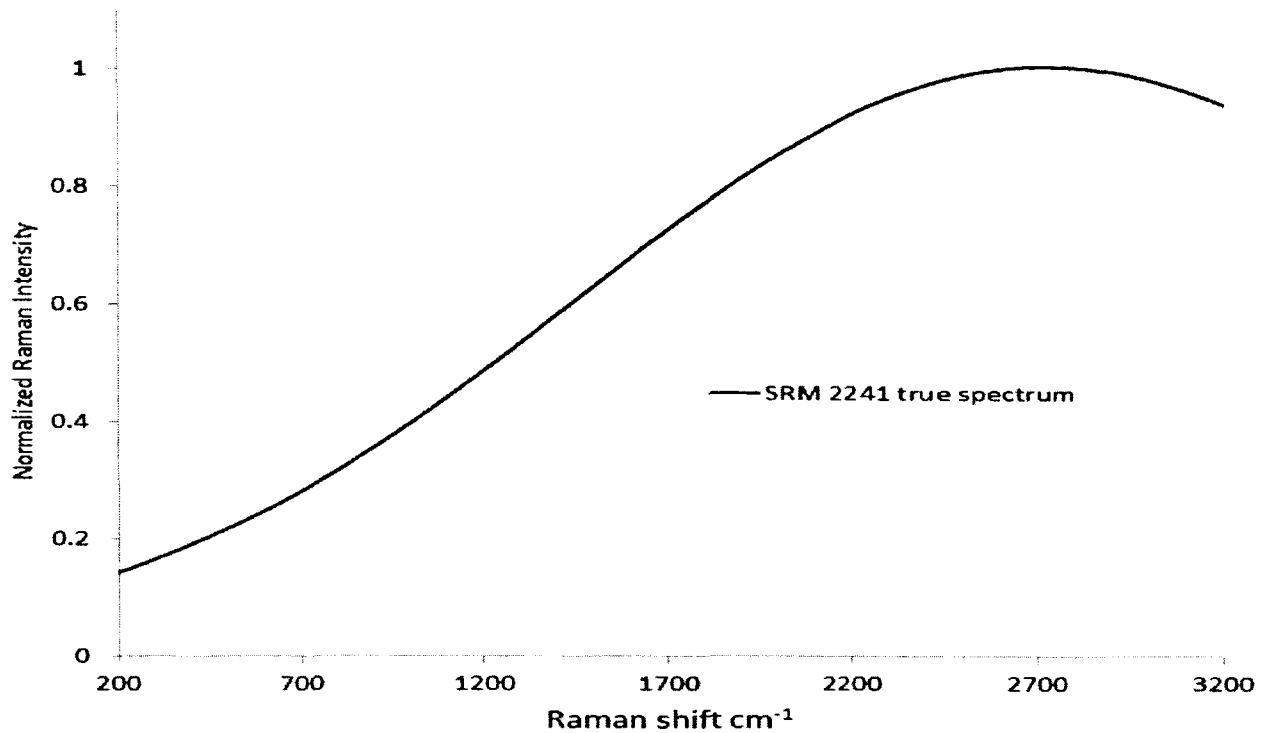


FIG. 30. The standard luminescence spectrum of the SRM 2241 generated using the five order polynomial coefficients provided by NIST and the SSE wavenumber axis. The x-axis is expressed in the units of Raman shift and the y-axis gives SRM 2241 luminescence intensity. The y-axis is normalized to the unity.

The standard SRM 2241 spectrum intensity is normalized to unity to universally fit any CCD detector. The SSE Raman spectral range is from  $200\text{ cm}^{-1}$  to  $3200\text{ cm}^{-1}$  and it is within the spectral range of the polynomial provided by NIST for SRM 2241.

The actual SRM 2241 luminescence spectrum ( $S_{\text{SRM}}$ ) was acquired on the SSE Raman. The integration time was increased so that we get a full count of the CCD detector (near to saturation, CCD detector currently implemented has 65,000 counts). The SRM 2241 luminescence spectrum acquired on the SSE Raman is shown in Figure 31. As shown in Figure 31, the  $S_{\text{SRM}}$  has different emission spectrum than the standard emission spectrum of the SRM 2241 provided by NIST.



FIG. 31. Luminescence spectrum of SRM 2241 acquired on the SSE with the 785 nm excitation source. The x-axis is expressed in wavenumber ( $\text{cm}^{-1}$ ) units and the y-axis gives the luminescence intensity of SRM 2241.

A correction vector ( $X_{\text{CORR}}$ ) is then calculated by dividing the standard luminescence spectrum over the acquired on the SSE ( $S_{\text{SRM}}$ ) spectrum of the SRMs 2241 as shown in Eqn. 29:

$$X_{\text{CORR}}(\Delta\nu) = Y_{\text{SRM}}(\Delta\nu) / S_{\text{SRM}}(\Delta\nu) \quad (29)$$

intensity-corrected Raman signals ( $I_{\text{CORR}}$ ) are obtained by the multiplication of the elements of the measured Raman spectrum of a sample ( $I_{\text{MEAS}}$ ) by the correction vector  $X_{\text{CORR}}$  as shown in Eqn. 30:

$$I_{\text{CORR}}(\Delta\nu) = I_{\text{MEAS}}(\Delta\nu) * X_{\text{CORR}}(\Delta\nu) \quad (30)$$

The correction vector  $X_{\text{CORR}}$  is determined one time at production and is saved for all future correction.

A Raman spectrum of acetaminophen acquired on the SSE Raman is shown in Figure 32 before intensity correction (solid line) and after applying intensity correction (dotted line). As it can be seen in Figure 32, the Raman intensity correction leads to an increase in the Raman intensity peaks at higher wavenumber Raman shifts (CH-stretching region) where the efficiency of the SSE CCD detector is low. This is particularly true when comparing the fingerprint region to the CH-stretching region. For the dispersive instrument, the silicon CCD detector has a minimum quantum efficiency for the CH-stretching spectral range (greater than 1020 nm), while for the FT-Raman spectrometer, this range corresponds to the maximum quantum efficiency of the InGaAs detector used. And hence, the intensity correction increases the intensity of the CH stretching region of the SSE Raman analyzer. In order to maintain the original intensity of the SSE Raman spectra after intensity correction, the corrected Raman spectrum were multiplied by the maximum intensity of the uncorrected Raman spectrum.



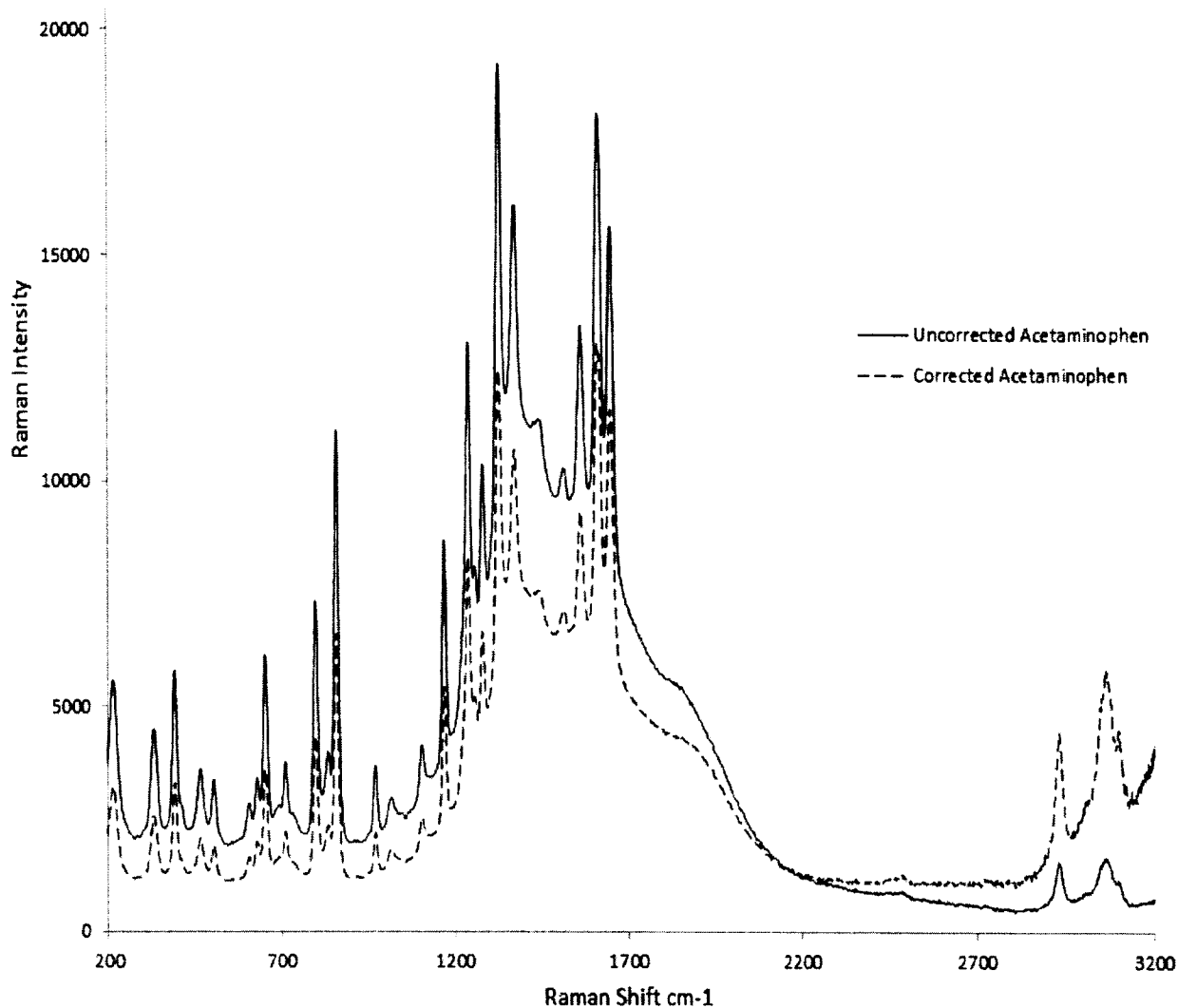


FIG. 32. Acetaminophen Raman Spectrum acquired on the SSE without Raman Intensity correction (solid line) and with Raman intensity correction (dotted line). The x-axis is expressed in Raman shift wavenumber units and the y-axis gives Raman intensity.

One has to consider two factors when using the SRM 2241 to calibrate Raman intensity; 1) the standard luminescence of the SRM 2241 is only certified for the use between  $200\text{ cm}^{-1}$  and  $3500\text{ cm}^{-1}$  Raman shift. The certified model cannot be used to extrapolate outside this wavenumber Raman shift range, 2) the standard luminescence spectrum and the measured luminescence spectrum of the SRM 2241 must have the same wavenumber Raman shift axis.

## **CHAPTER III**

### **NOVEL ALGORITHM AND INSTRUMENTATION FOR FLUORESCENCE-FREE RAMAN SPECTROSCOPY UTILIZING DUAL-TUNABLE EXCITATION SOURCES**

#### **INTRODUCTION**

Although the sequentially shifted excitation Raman analyzer, discussed in chapter II, is capable of removing a fluorescent background, there is still a requirement that Raman signal be present in addition to the fluorescence. Using 16 bit analog-to-digital conversion for the detector signal (typical for a high end CCD), this means that for every Raman photon there cannot be more than 65000 fluorescence photons or the detector will saturate. Of course this best case scenario assumes that there is no shot noise or electronic noise in the detector. However, for highly fluorescent compounds, even excitation with a 785 nm excitation source can give rise to fluorescence which is six orders of magnitude more intense than that of the Raman signal. However, utilizing a longer wavelength excitation source than 785 nm leads to decreasing fluorescence background. Unfortunately, dispersive Raman instruments utilizing CCD detectors and long wavelength excitation sources (significantly longer than 785 nm) cannot detect high energy scattered Raman such as CH-stretching due to the nature of silicon band gap used in the manufacturing of CCD detectors. For this reason, handheld Raman analyzers usually utilize a short wavelength excitation source such as 785 nm, along with a method of extracting Raman information out of the fluorescence background. Overall, these methods of removing fluorescence can be separated into four categories: algorithm based baseline correction methods; methods and devices using specialized sampling optics; time gating methods; and shifted excitation methods.

Algorithm-based methods attempt to mathematically estimate a baseline and then subtract the estimated baseline from the Raman spectrum to give a fluorescent free spectrum. Algorithm-based methods suffer from the requirement that they cannot be universally applied to all types of fluorescence without either significant degradation in performance or significant user intervention to adjust sensitive algorithm parameters. One reason for this disadvantage is the varied nature of fluorescence backgrounds. Another

reason is due to the complexity of the fitting algorithms. This prevents these methods from being used routinely by non-technical personnel. Methods using algorithm based baseline corrections are described in detail in the peer-reviewed scientific literature<sup>1-6</sup>

Methods using unique sampling optics and geometries have the disadvantage of only mitigating the impact of fluorescence instead of completely eliminating it from the Raman spectrum. These methods also suffer from the disadvantage of only being applicable to certain types of samples. For instance, some of these methods can only be used on solids or only be used on liquids. Methods using unique sampling optics and geometries are described in detail in the peer-reviewed scientific literature<sup>7-9</sup>.

Time gating methods involve using time to discriminate between the Raman signal (which occurs on a fast time scale) and the fluorescence (which generally occurs on a slower timescale). Since the time scale discrimination required for these methods is on the order of picoseconds to nanoseconds, the methods involve the use of either: 1) complex, bulky, and expensive instrumentation or 2) instrumentation which requires excessively long periods of time to collect an entire Raman spectrum. Also, these methods are not universally successful for all samples due to the varied nature of fluorescence decay times. These methods also have the disadvantage of not being able to remove spectral backgrounds which arise from optical processes which occur on the same timescale as Raman. Such processes include stray light scatter, fixed pattern noise arising from a detector array, room lights, cosmic rays, and other extraneous sources of optical interference. These methods also have the disadvantage of requiring the Raman to be generated using a pulsed laser source. This results in a large number of Joules being delivered to the sample in a short period of time and often results in sample burning. Examples of time gating methods are described in detail in the peer-reviewed scientific literature<sup>10-17</sup>.

All shifted excitation methods involve changing the excitation laser during spectral acquisition. These methods rely on a common concept: the location of Raman intensities in spectral space changes with excitation while unwanted spectral intensities corresponding to fluorescence, stray light, fixed pattern detector noise etc., remain unchanged in spectral space. The difference in the various methods to date occurs in how the spectra are acquired and how the Raman data is extracted. The simplest extraction

method is taking a difference between two sets of excitation data and is referred to as Shifted Excitation Raman Difference Spectroscopy (SERDS) and only requires excitation at two distinct wavelengths. This type of method has two distinct disadvantages: 1) random noise in two measurements is increased during subtraction; 2) the result is a derivative spectrum instead of a true Raman spectrum in spectral space. Since a true Raman spectrum is not generated, the data is not easily interpreted. Another disadvantage of SERDS is that there is significant difficulty in reconstructing the Raman spectrum from the derivative data. The reconstruction requires the use of advanced algorithms which have to be applied appropriately for specific samples since one set of algorithm parameters will not work universally for all samples. Since this method has been in use for over thirty years, there are many examples of its application in the peer-reviewed scientific literature<sup>18-26</sup>. Perhaps more intriguing is the recent work of Willet and co-workers in which a series of diode lasers at closely spaced wavelengths are used for the acquisition of shifted Raman spectra. The setup allows the researchers to extract the spectral data in true Raman space using an iterative expectation-maximization (EM) algorithm (Lucy-Richardson algorithm) based on a shift-matrix operator. This is the first example of the extraction of Raman data regardless of the complexity or nature of the fluorescence background (i.e., a universal approach). Despite the promise of the method, the instrumentation remains complex and it can be anticipated that reproducing the series of selected wavelength lasers would not be possible for multiple instruments. In addition, the processing of the data using the shift matrix operator requires tens of billions of double precision calculations so that processing time often requires several minutes for high background spectra. This problem was solved by application of our SSE algorithm, described in chapter II, which is practical and requires no user intervention.

In this chapter we present a new instrument which utilizes *two* tunable excitation sources and allows for the acquisition of fluorescence-free Raman spectra. The advantage of this dual-laser approach is that an 852 nm laser is used to collect the fingerprint region of the Raman spectrum where fluorescence is quite intense, while a 785 nm laser is used to collect the CH stretching region of the Raman spectrum where fluorescence is minimal. By using a 785 nm laser to collect the CH stretching region, a silicon based CCD detector can still be used to collect the signal. By using an 852 nm

laser to collect the fingerprint region, the fluorescence is significantly decreased, thus preventing saturation of the CCD before a Raman photon is detected.

## EXPERIMENTAL SECTION

### *INSTRUMENT DESIGN*

The design of dual-laser SSE Raman instrument is shown in Figure 33. The excitation source (Fig.33A) is two distributed Bragg reflector (DBR) GaAs diode lasers mounted on a Peltier cooler and contained within a custom TO-8 electronic package. The two lasers are spaced by 0.5 mm. The 852nm DBR laser is shown in Figure 34. One of the two DBR lasers emits single-mode 785 nm radiation and the other emits single-mode 852 nm at 25 °C. The Bragg grating of the DBR laser is a periodic structure etched into the GaAs substrate at the rear facet of the diode cavity. Hence the cavity and the grating consist of a monolithic structure. The emitting wavelength of the lasers at a given temperature is equal to:

$$\lambda = CW + (K \times \Delta T) \quad (31)$$

where CW is the center wavelength of the DBR laser at a specific temperature, K is the change in a laser wavelength per degree Celsius, and  $\Delta T$  is the set-point temperature of the thermo-electrical controller (TEC) minus the temperature where the laser has a wavelength equal to CW which is 25 °C for the two lasers used in the current SSE analyzer.

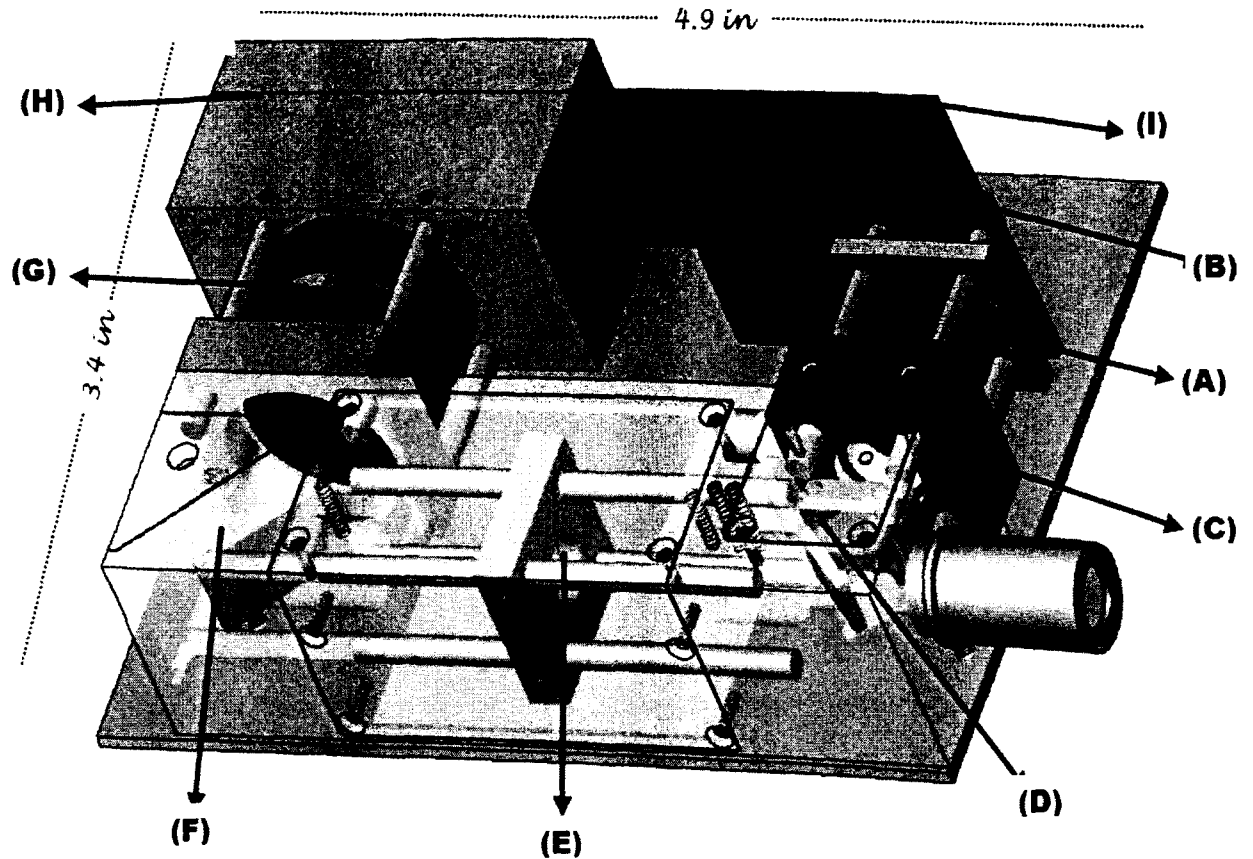


FIG. 33. The dual-laser SSE Raman instrument is composed of (A) two distributed Bragg reflector (DBR) GaAs diode lasers. The lasers emit 785 and 852 nm wavelengths at room temperature (25 C°) and are separated by 0.5 mm. ; (B) a thermoelectric controller (TEC) to control the lasers temperature; (C) a band pass lasers filter; (D) a dichromic beam-splitter; (E) a 50 micron rectangular slit; (F) a volume holographic transmission grating; (G) a doublet achromatic lens; (H) a 1024 x 256 element CCD detector; and (I) a microcontroller board to control both TEC and CCD.

The change in wavelength per degree Celsius was found to be 0.074 nm and 0.064 nm for 785 nm and 852 nm excitation sources, respectively. This means Eqn. 31 can be rewritten as;

$$\lambda = 785 \text{ nm} + 0.074 \text{ nm/}^{\circ}\text{C} (\Delta T) \quad (32)$$

$$\lambda = 852 \text{ nm} + 0.064 \text{ nm/}^{\circ}\text{C} (\Delta T) \quad (33)$$

The feedback from the Bragg grating stabilizes the modal structure of the laser so that when the laser is operated at a constant current, a range of  $\Delta T$  exceeding 15 °C is obtainable without a change in the single-mode modal structure of the laser output (i.e., no mod-hop occurs). Even using a Bragg grating to stabilize the laser, it is possible to have a laser which at a certain temperature and drive current gives rise to a modal instability. However, even in these instances, once a desired optical power is chosen, its value can be varied typically by less than 5% to obtain a current which provides a mode-hop free range over a large temperature change of the laser.

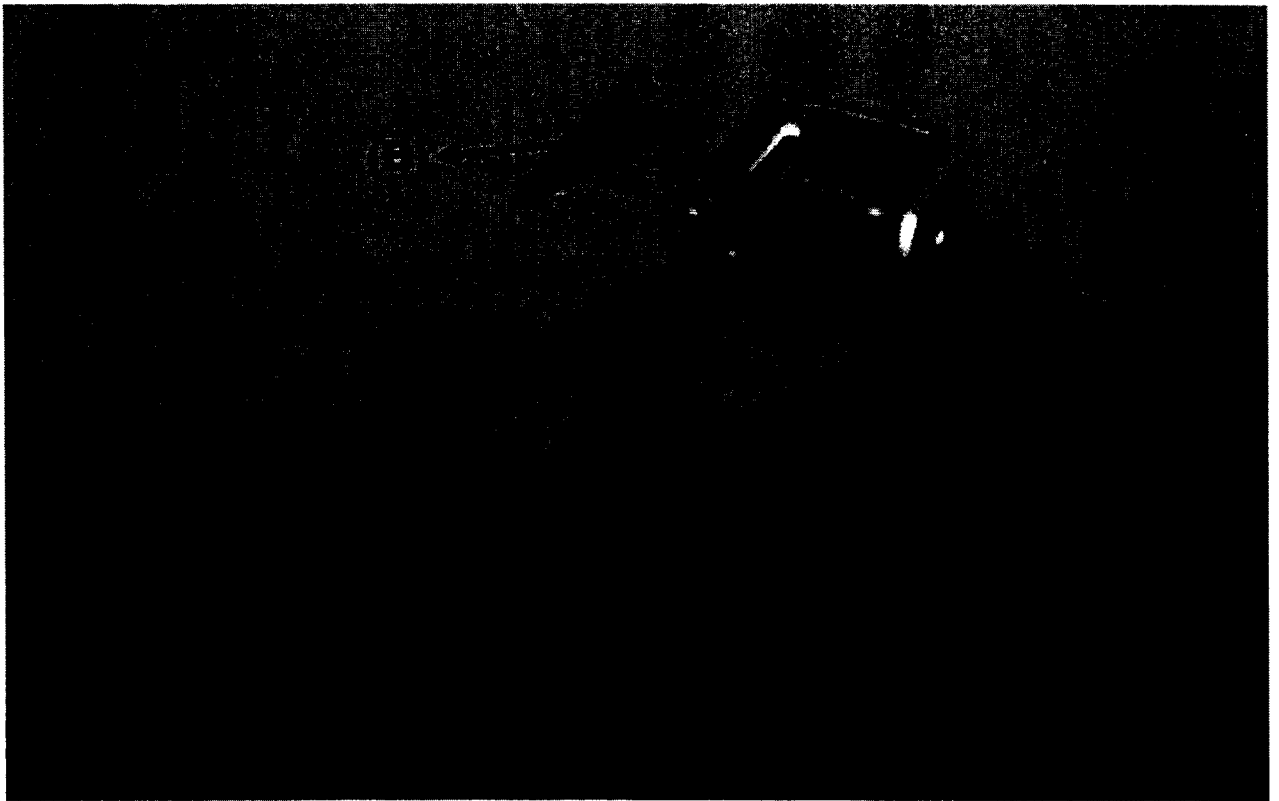


FIG. 34. A diagram of the 852 nm distributed Bragg reflector (DBR) GaAs diode laser is displayed with SEM micrographs showing A) the emitting facet of the laser cavity; B) a custom TO-8 package mounted to electronically control the laser.

We have evaluated 6 different DBR lasers, all of which have demonstrated this characteristic (consistent with the manufacturer supplied current-voltage curves for each device). For the current work, a sample-incident optical power output of 50 mW was selected and the lasers were run in constant current mode (100 mA) by using a fixed resistance on the thermoelectric controller (TEC) board (Fig 33B). The TEC controller

board is capable of 0.01 °C precision. This provided a stable mode-hop free temperature range from 16-34°C.

The laser beams are collimated with an asphere (NA = 0.55) and filtered with a short pass filter (Fig. 33C). A dichromic beam-splitter (Fig. 33D) directs the laser through a focusing asphere (NA = 0.50) onto the sample. The asphere collected Raman scatter is passed through the beam-splitter where it is focused by a 12.7 mm diameter doublet achromat of focal length 32 mm (Fig. 33E) onto a 1mm x 50 micron slit. An identical achromat is used to collimate the light. The collimated light is diffracted using a volume holographic transmission grating (1500 gr/mm) mounted at -45° to the optical axis (Fig.33F). The diffracted light is focused onto the back-thinned CCD detector (Fig.33H) using a 45 mm focal length doublet achromat (Fig.33G, 25.4 mm diameter to account for the anamorphic magnification of the grating). The design provides a spectral resolution of approximately 10 cm<sup>-1</sup>. A microcontroller board (Fig.33I) is used to control both the TEC and the CCD data acquisition and to upload the collected raw spectra to a processing CPU over a USB 2.0 connection.

All lenses and opto-mechanical mounts were purchased from Thor Labs. All optical surfaces were anti-reflection coated for the spectral range used. The grating, beam-splitter, and detector mounts were machined in house. The grating was purchased from Wasatch Photonics. The CCD detector was purchased from Hamamatsu (model S11500-1007) and was used at room temperature without any cooling. The slit with mounting was purchased from Thor Labs. The 785 nm and 852 nm DBR lasers were purchased from Photodigm. The DBR lasers were filtered with a dual-notch 785nm and 852nm laser-line band-passes filter from Chroma (5 nm FWHM). The collected Raman radiation was filtered using a 852nm Razor-edge long-pass filter from Semrock. The 860 di-chroic beamsplitter was purchased from Semrock also. Custom electronics for the detector, laser driver, and TEC controller were developed in-house. CCD signal conversion was carried out using a 16 bit A/D converter. For typical experiments, the low integration times resulted in only 2000-5000 thermal counts, leaving over 55,000 counts of dynamic range for the spectral signals.



All control code was written in C and compiled using the GNU compiler build for the Freescale MCore. Spectra were uploaded to a computer with a Windows 7 operating system over USB using an USB 2.0 USB/serial converter (FTDI). All spectra were processed in real-time (< 2s) using a custom program written using Labview (National Instruments). All dispersive Raman spectra were calibrated to a wavenumber axis using the acetaminophen NIST standard.

All FT-Raman spectra were acquired using a Magna 950 FT-Raman (Thermo-Fischer Scientific). All FT-Raman spectra were acquired by averaging 100 scans at 4  $\text{cm}^{-1}$  resolution, with 800 mW laser power and Hap-Ganzel apodization. An InGaAs detector at room temperature was used for all acquisitions. Total acquisition time was 2 minutes for each spectrum.

All the chemicals were purchased from Aldrich. The chemicals were placed in borosilicate NMR tubes for spectral sampling.

#### *INSTRUMENT OPERATION*

Since the DBR diode lasers are mounted to the same thermoelectric cooler, their temperatures are the same. A typical measurement consists of collecting Raman spectra at laser temperatures of 20, 23, 26, and 29 °C (i.e., four sequential excitations which are evenly spaced in temperature). This yields excitation wavelengths for laser one (785nm) of 784.630, 784.852, 785.074, and 785.296 nm, respectively. This gives a constant excitation shift of 0.222 nm. When converted to wavenumber ( $\text{cm}^{-1}$ ), an excitation-separation of approximately 3.60  $\text{cm}^{-1}$  is obtained (approximate since the  $\text{cm}^{-1}$  scale is linear to 0.01  $\text{cm}^{-1}$  over such a small change in wavelength). While laser two (852nm) excitation wavelengths are 851.680, 851.872, 852.064, and 852.256 nm, respectively, for the temperatures 20, 23, 26, and 29 °C. This gives a constant excitation shift of 0.192 nm. When converted to wavenumber ( $\text{cm}^{-1}$ ), a separation of 2.65  $\text{cm}^{-1}$  between the different excitations is obtained. Due to the presence of the DBR grating and the use of fast optics with anti-reflection coatings, we have not been able to observe any changes in laser quality due to back-reflections into the diode cavity resulting from sample placement. For this reason an optical isolator in front of the diode laser was omitted in the current design.

Although the instrument can be operated in a variety of manners, the use of a microcontroller allows programming of a set sequence of events which minimizes acquisition time and electrical power consumption. For the current work these steps included:

- 1) turn on TEC;
- 2) cool lasers to a starting temperature of 20 °C;
- 3) turn on laser 1;
- 4) reset CCD and collect spectrum;
- 5) turn laser 1 off and laser 2 on;
- 6) reset CCD and collect spectrum;
- 7) heat lasers to next temperature and upload data;
- 8) repeat steps 3-7 until final spectrum is collected; and
- 9) turn off TEC.

For this sequence, the most time consuming steps are 2 and 7. A schematic of the current operation is shown in Figure 35. Due to the small mass of the diode lasers, the time to cool the lasers (when room temperature is  $< 30$  °C) is less than 1s. The time required to heat the lasers to the next temperature is 300 ms for temperature steps of 6 °C or less. The time to upload the data is 28 ms, and thus has no time penalty since it occurs during the TEC heating period. The time required to turn a laser on and off is less than 1 microsecond. Thus the total acquisition time for a four temperatures experiment is equal to the total integration times of the spectra plus  $< 2$  s.

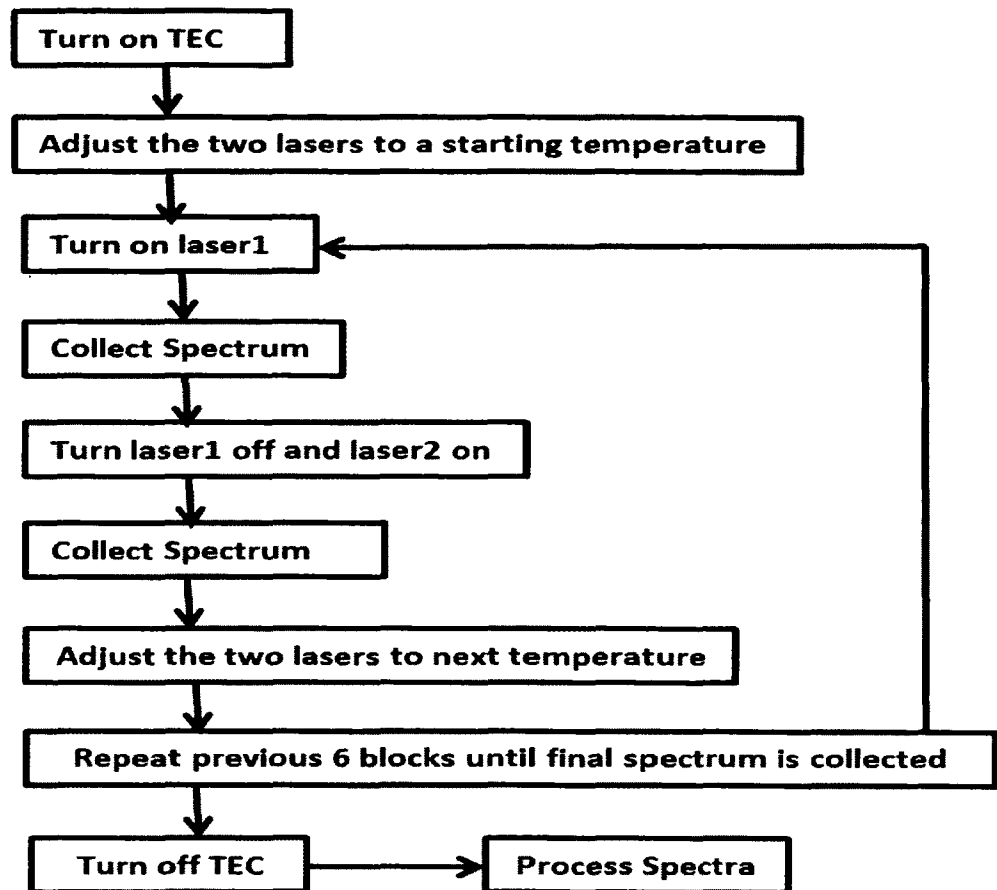


FIG. 35. This figure gives the current sequences of operation of the dual-laser SSE Raman. The most time consuming steps are heating and cooling the laser; heating laser to the next temperature requires 300ms while cooling the laser requires less than a second for temperature steps of 6 °C or less. The total acquisition time equals to the integration time plus less than two seconds.

The x-axis of the all acquired Raman spectra were linearized to a calibrated x-axis with the intervals of the actual Raman shift ( $3.66 \text{ cm}^{-1}$  for 785 nm Raman spectra and 2.65 nm for 852 nm Raman spectra for the temperatures 20, 23, 26, and 29 °C). Since the DBR lasers are run in a constant current mode, when the laser temperature is changed, there is a slight change in the optical power output. Across all temperatures used for both lasers, this change in optical power is less than 5%. A possible way to account for this small change in laser power is to use the standard normal variate scaling (SNV). Treating each Raman spectrum as a vector, the mean and standard deviations of the vector are calculated, and for each point in the vector, the mean is subtracted and the result is

divided by the standard deviation. Unfortunately, SNV eliminates the original intensity information. To retain this information, a modified version of SNV was used which we call scaled SNV. The calculation of scaled SNV is given in Eqn. 34;

$$SSNV_j = \left\{ \frac{R_j - \hat{R}_j}{\sigma_j} \right\} * \sigma_1 + \hat{R}_1 \quad (34)$$

where  $j$  is a spectral index referring to each of the shifted Raman spectra,  $R$  is the raw Raman spectrum (a vector of intensities) and  $\hat{R}_j$  and  $\sigma_j$  are the mean and the standard deviation of the  $R$  spectrum, respectively. The symbols  $\sigma_1$  and  $\hat{R}_1$  are the lowest temperature Raman spectrum standard deviation and mean, respectively. Effectively, SNV scaling is performed on each Raman spectrum, and then each Raman spectrum is multiplied by the standard deviation of the first Raman spectrum (unscaled) and added to its mean (unscaled).

#### *RAMAN DATA PROCESSING*

Raman spectroscopy fluorescence-removal techniques which involve changing the excitation laser during spectral acquisition all rely on a common concept: the location of Raman intensities in spectral space changes with excitation while unwanted spectral intensities corresponding to fluorescence, stray light, fixed pattern detector noise etc., remain unchanged in spectral space. The difference in the various methods to date occurs in how the spectra are acquired and how the Raman data is extracted.

The sequentially shifted excitation (SSE) algorithm, explained in details in Chapter II, is an iterative algorithm which relates the desired outcome (separated Raman spectrum and fluorescence spectrum) to the collected data ( $R$ ):

$$\hat{w}_i = \sum_{k=0}^{K-1} R_{k,n} \div (S_{i,n}^F + S_{i,n-k}^R) \quad (35)$$

where  $\hat{w}_i$  is a weighting vector which is calculated with each iteration  $i$ . The division and addition operators are carried out element-wise.  $\mathbf{R}$  is the spectral data matrix and is written as a concatenated column vector of spectra:

$$\mathbf{R} = \begin{pmatrix} R_0 \\ \vdots \\ R_k \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} r_{0,0} \\ \vdots \\ r_{0,n} \end{pmatrix} \\ \vdots \\ \begin{pmatrix} r_{k,0} \\ \vdots \\ r_{k,n} \end{pmatrix} \end{pmatrix} \quad (36)$$

where each spectrum has  $N$  spectral positions of intensity  $r_{k,n}$  at each spectral position, and there are a total number of spectra corresponding to the number of excitations ( $K$ ).  $S_{i,n}^F$  is an initial estimate of the background signal (e.g. fluorescence or any non-Raman signal) and is obtained by taking the minimum of each column in matrix  $\mathbf{R}$ :

$$S_0^F = (\min_k r_{k,0}, \min_k r_{k,1}, \dots, \min_k r_{k,N-1}) \quad (37)$$

where  $S_0^F$  represents the vector containing the initial fluorescence estimate (e.g. the estimate at the 0<sup>th</sup> iteration).  $S_{i,n-k}^R$  An initial estimate of the Raman spectrum is obtained by calculating the standard deviation of each column in  $\mathbf{R}$ :

$$S_0^R = \langle \sigma(r_{0,0}, r_{1,0}, \dots, r_{K-1,0}), \sigma(r_{0,1}, r_{1,1}, \dots, r_{K-1,1}), \dots, \sigma(r_{0,N-1}, r_{1,N-1}, \dots, r_{K-1,N-1}) \rangle \quad (38)$$

where the  $r_{k,n}$  elements in parentheses indicate a particular column of matrix and  $\sigma$  is the standard deviation of that column, and  $S_0^R$  represents the vector containing the initial Raman estimate (e.g. the estimate at the 0<sup>th</sup> iteration). For each iteration  $i$ , the fluorescence spectrum is calculated as:

$$S_{i+1}^F = S_i^F \times \hat{w}_i \quad (39)$$

and the Raman spectrum is calculated as:

$$S_{i+1}^R = S_i^R \times \hat{w}_i \quad (40)$$

The number of calculations using SSE is proportional to  $k \times N$  per iteration, this results in total processing times of a fraction of a second using a conventional dual-core 1.8 GHz Windows computer. Significantly, a part of the simplicity of the approach lies in the use of a fixed change in wavelength (excitation shift) between the utilized excitations. Use of varying excitation shifts would require the incorporation of a vector which indexes each of the shifts for each  $k$  calculation.

As presented, the SSE algorithm requires that a spectral shift which corresponds to the minimum difference in wavenumbers between the spectral data points. With the current method, that is possible by selecting the appropriate laser temperatures. However, if a series of different lasers is used with non-constant wavelength spacing, then the data must be splined so that the operator matrix indices can accommodate the varying shifts. The resulting loss in symmetry of the operator matrix not only complicates the algorithm, but it also results in the requirement that the algorithm be specifically tailored for each instrument which is produced. In addition, splining the data to obtain the requisite shift resolution will result in an increase in  $N$ , which further increases the processing time.

Figure 36 describes the overall method of extraction of pure Raman spectra from the sequentially shifted excitation Raman spectra. Where, the SSE Raman spectra acquired utilizing 852 nm and 785 excitation sources are processed separately and independently. In Figure 36, all the letters without prime and with prime represent 852 nm and 785nm SSE Raman spectra processing, respectively. (A) represents the actual SSE Raman spectra acquired on the SSE Raman analyzer utilizing the 852 nm excitation source and (A') gives the SSE Raman spectra acquired utilizing the 785nm.

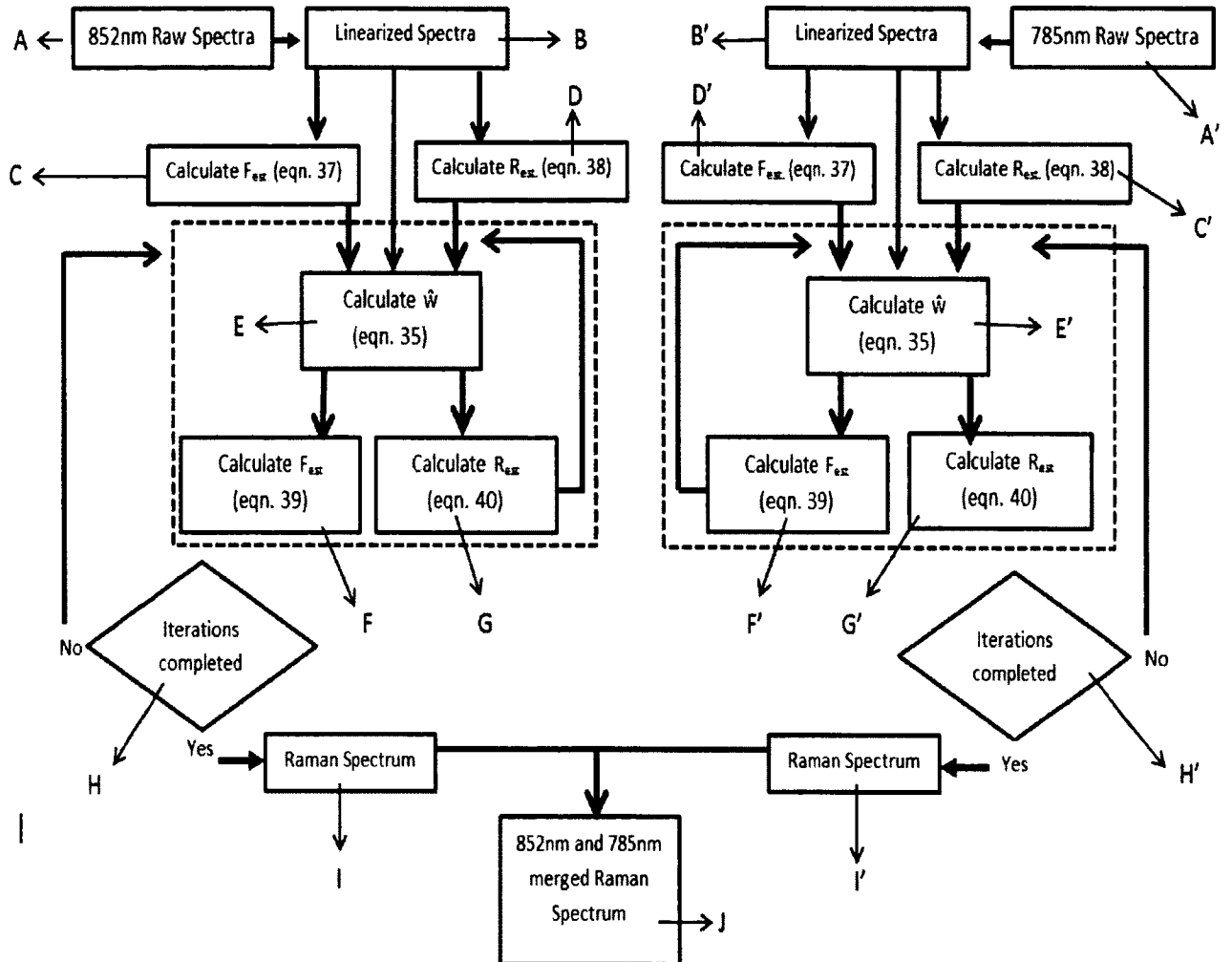


FIG. 36: This figure shows the overall method of processing the shifted excitation Raman spectra acquired on the dual-laser SSE Raman to obtain a background-free single-merged Raman spectrum.

The B represents the linearization of the 852 nm SSE Raman spectra to a calibrated axis of energy spacing in wavenumber equal to the actual shift in Raman peaks. Also, B includes the SSE Raman spectra pretreatment (scaled SNV) given in Eqn. 34. The B' is the same as B but for the 785 nm SSE Raman spectral data. The C & D give the initial estimates of fluorescence and Raman calculated using Eqn.s 37 & 38, respectively, for the Raman data acquired utilizing the 852nm. The C' & D' give the initial estimates of fluorescence and Raman calculated using Eqn.s 37 & 38, respectively, for the Raman data acquired utilizing the 785nm. E, F, G & E', F', G' represent the SSE algorithm steps

shown in details in Eqn.s 35, 39, & 40 for both the 852 and the 785 nm Raman data, respectively. The number of iterations (H and H') is set at 2000 iterations. The separated true Raman spectra of both the 852 and 785 nm are given at I and I'. The 852 nm Raman spectrum spans the spectral range from 300 to 2100  $\text{cm}^{-1}$  wavenumber Raman shift. While, the 785 nm Raman spectrum spans the spectral range from 1300 to 3200  $\text{cm}^{-1}$  wavenumber Raman shift. The two Raman spectra are then merged together at 2000  $\text{cm}^{-1}$ .

As a further clarification, a practical example is given in Figure 37-41. Figure 37 shows the two excitations sources and their corresponded Raman signals spectral range in absolute wavelength axis. As show in Figure 37, the 785 Raman signals spans the absolute wavelength spectral range from 881 nm to 1048 nm. While, the 852 Raman spectral range is from 874 to 1037 nm absolute wavelength.

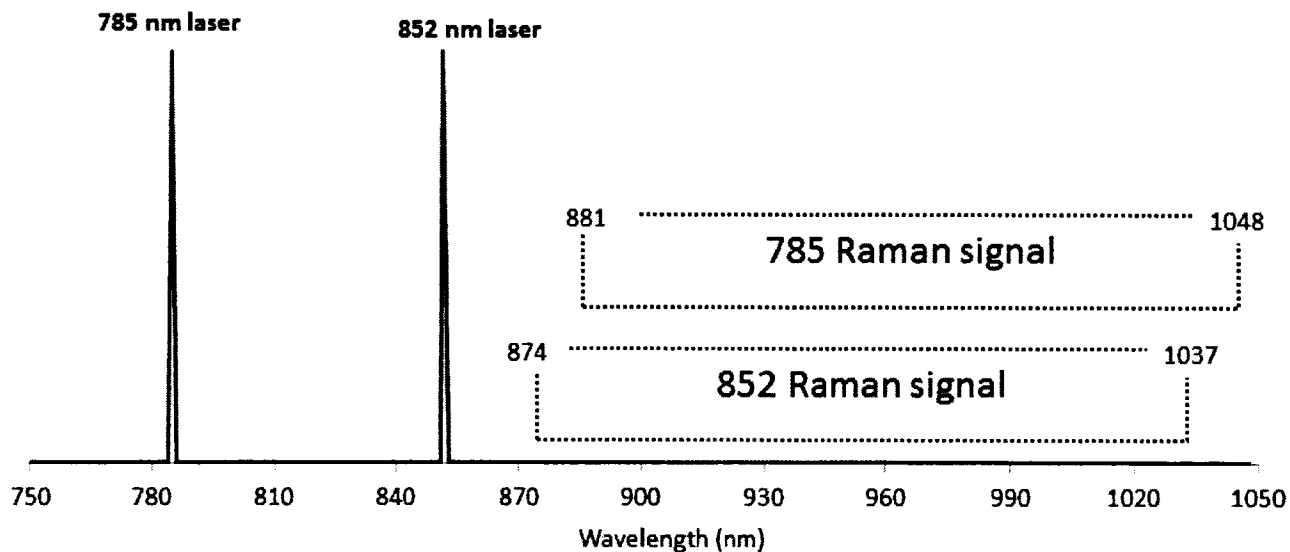


FIG. 37. The 785 and 852 nm excitation sources and their corresponding Raman signal regions shown in an absolute wavelength axis.

In Figure 38, the acetaminophen spectra acquired on the dual-laser SSE Raman analyzer are shown when both the 785 nm excitation source (top) and the 852 nm excitation source (bottom) are utilized. Although the two Raman spectra have almost



similar spectral range in absolute wavelength axis, they obviously show different Raman information. The absolute wavelength (nm) axis is converted to wavenumber ( $\text{cm}^{-1}$ ) axis and that is shown in Figure 39.

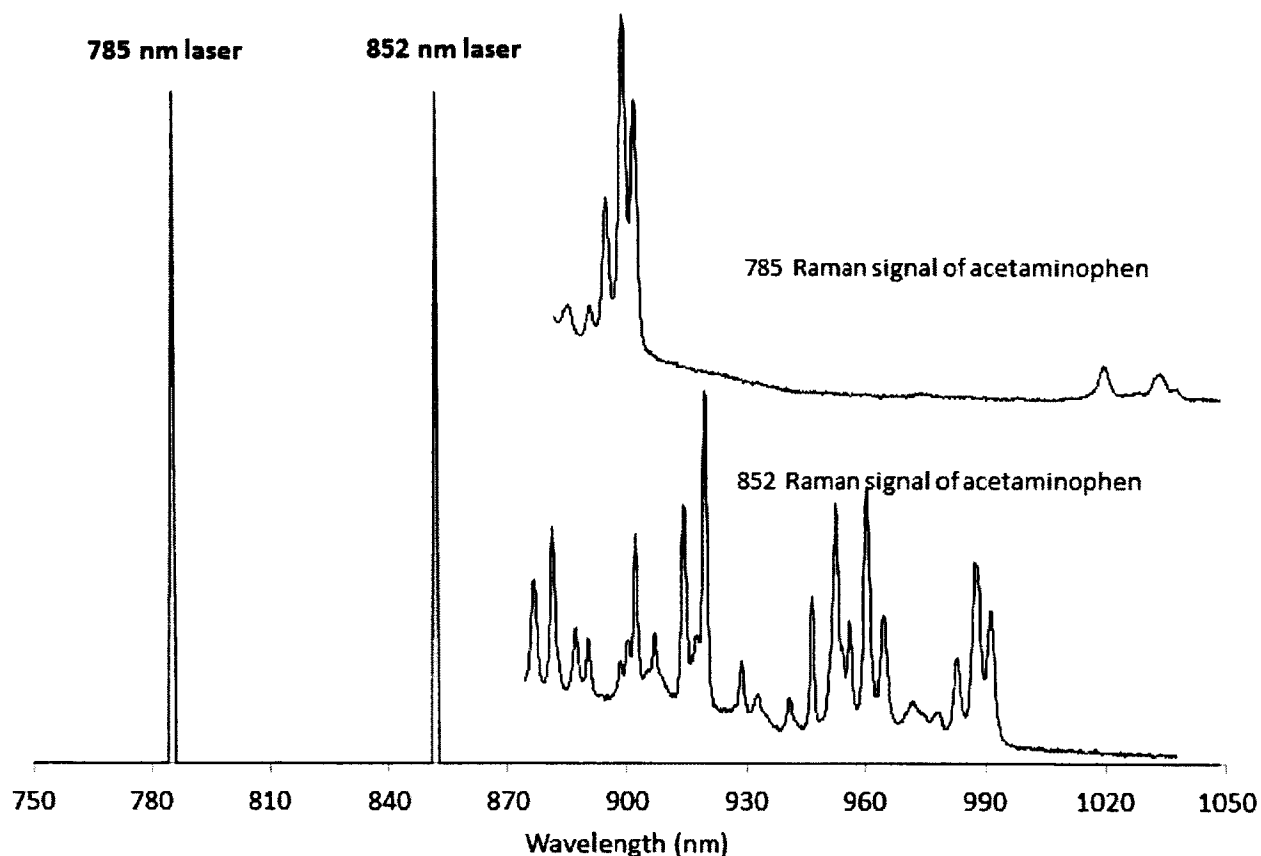


FIG. 38. The 785 and 852 nm excitation sources and their corresponding Raman spectra of acetaminophen shown in an absolute wavelength axis. Top and bottom are acetaminophen Raman spectra acquired utilizing 785 nm and 852 nm excitation sources, respectively.

The absolute wavenumber axis of the two acetaminophen Raman spectra shown in Figure 39, is converted to two Raman shift wavenumber axes by subtracting the corresponding excitation source in wavenumber units from the Raman signal (subtracting the 785 nm laser wavenumber from the absolute wavenumber axis gives a 785 axis in unit of wavenumber Raman shift. While, subtracting 852 nm laser wavenumber from the absolute wavenumber axis gives an 852 axis in units of wavenumber Raman shift). The

two spectra of acetaminophen are plotted in Figure 40, where the x-axis expresses Raman shift in wavenumber units. The two spectra have the same peaks over the overlap region (1300 to 2100  $\text{cm}^{-1}$  Raman shift). The disparity in photoresponse between the 785 and 852 Raman spectra of acetaminophen over the overlap region is due to the positional efficiency of the CCD detector. The two Raman spectra of the acetaminophen are finally merge at 2000  $\text{cm}^{-1}$  wavenumber Raman shift and the final merged Raman spectrum of acetaminophen is shown in Figure 41. The merged Raman spectrum spectral range is from 300 to 3200  $\text{cm}^{-1}$ , which includes both the fingerprint and the CH-stretching regions. The x-axis interval is 2  $\text{cm}^{-1}$  wavenumber which gives a 1450 data points in the merged Raman spectra of acetaminophen, and a resolution of 8  $\text{cm}^{-1}$ .

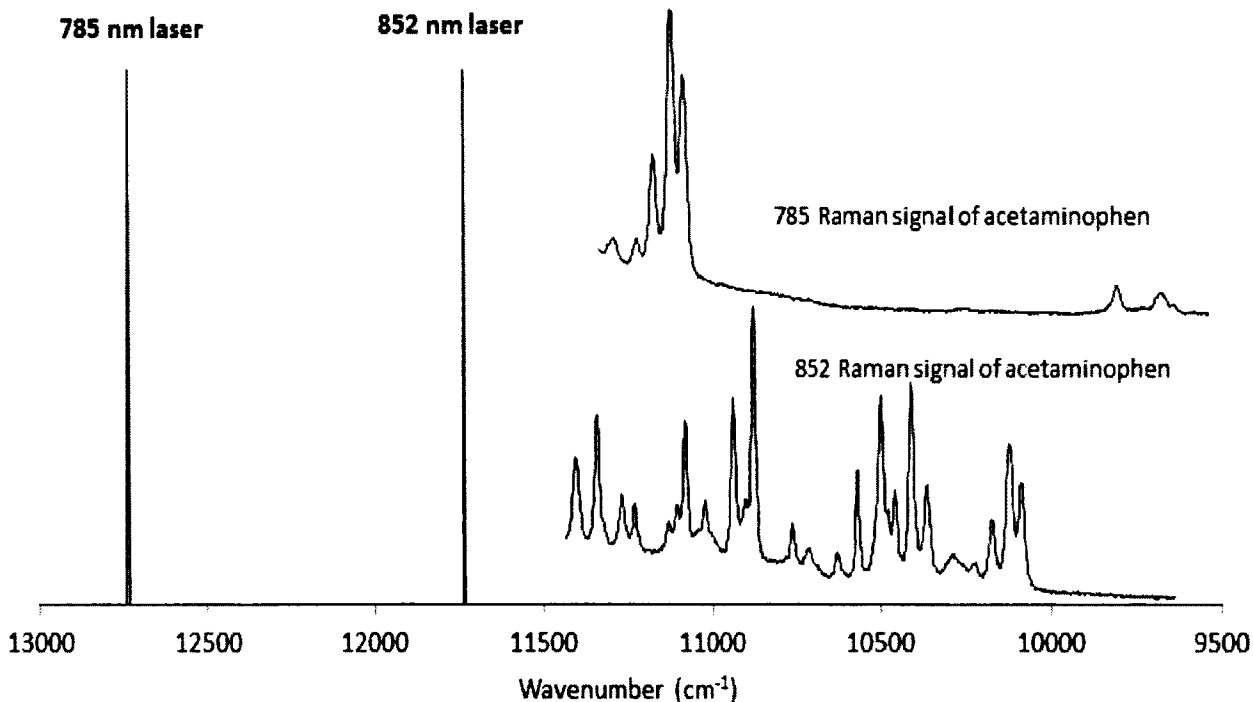


FIG. 39. The 785 and 852 nm excitation sources and their corresponding Raman spectra of acetaminophen shown in an absolute wavenumber ( $\text{cm}^{-1}$ ) axis. Top and bottom are acetaminophen Raman spectra acquired utilizing 785 nm and 852 nm excitation sources, respectively.

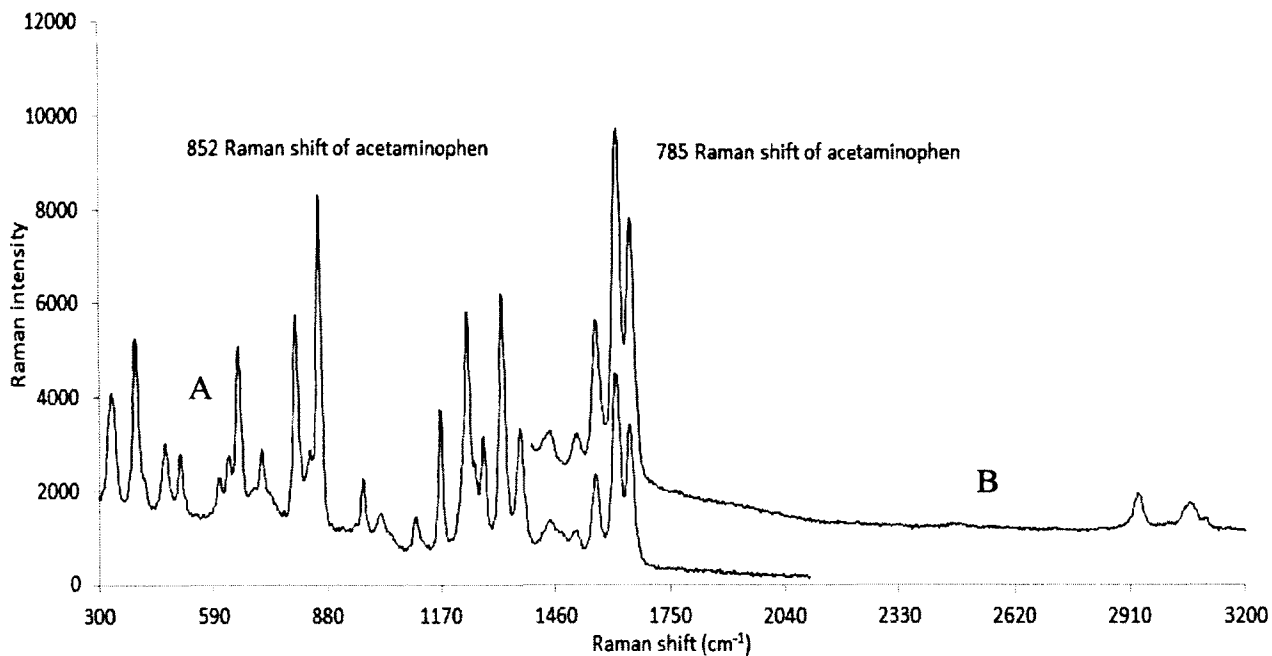


FIG 40. The 785 and 852nm Raman spectra of acetaminophen are plotted in a Raman shift wavenumber (cm<sup>-1</sup>) units. The A and B are acetaminophen Raman spectra acquired utilizing 785 nm and 852 nm excitation sources, respectively. The two spectra overlap over the spectral range from 1300 to 2100 cm<sup>-1</sup>.

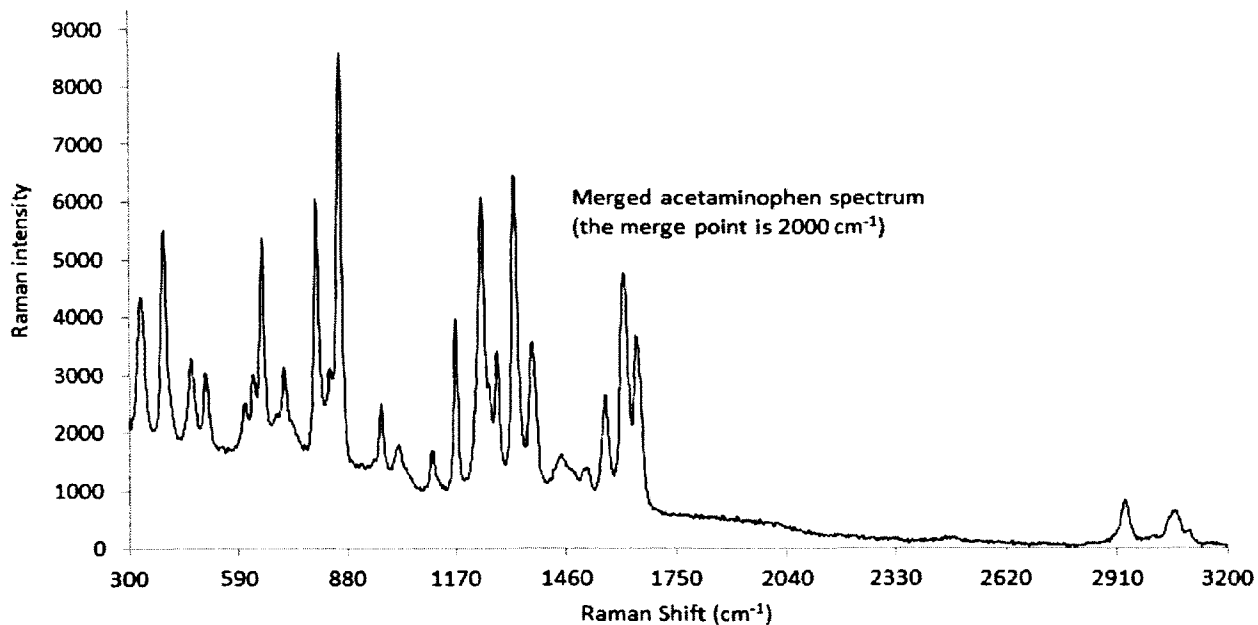


FIG. 41. The 785 and 852 merged acetaminophen Raman spectrum. The merge point is 2000 cm<sup>-1</sup>. The spectrum is shown as acquired without applying the SSE algorithm to remove the fluorescence-background.

## RESULTS AND DISCUSSION

### *LASER STABILITY*

To evaluate the stability of the 852 nm and the 785 nm laser emissions, thirty replicates of acetaminophen Raman spectra were acquired on the dual-laser SSE at different temperatures. The Raman spectra of acetaminophen we acquired on the dual-laser SSE at the temperature 19, 21, 23, 25, 27, 29, 31, and 33 °C. The acquisition of acetaminophen spectra were repeated thirty times. For each acquisition, the acetaminophen Raman peaks shift of the 785nm and the 852 nm excitations sources at each temperature were calculated relative to the lowest-temperature-acetaminophen Raman spectrum (19 °C) of each laser. The plot between the diode laser temperature and the Raman shift of acetaminophen Raman spectra acquired on the dual-laser SSE utilizing both the 852 and the 785nm excitation sources is shown in Figure 42A and 42B, respectively. Where the x-axis gives the change in the diode lasers temperature from 19 C to 33 °C and the y-axis gives the Raman shift in units of wavenumbers. As shown in Figure 42A &B, the Raman shift linearly increases as the diode laser temperature increases for the both lasers. This indicates that there is no mode-hop in the emission of the two lasers over the temperature range from 19-33 °C.

To calculate the stability of the two lasers emission at each temperature, the Raman shift standard deviation of the thirty replicates at each temperature is calculated. The standard deviations of Raman peaks shift at different temperature of the 785 and the 852 nm diode lasers are given is Table 3.

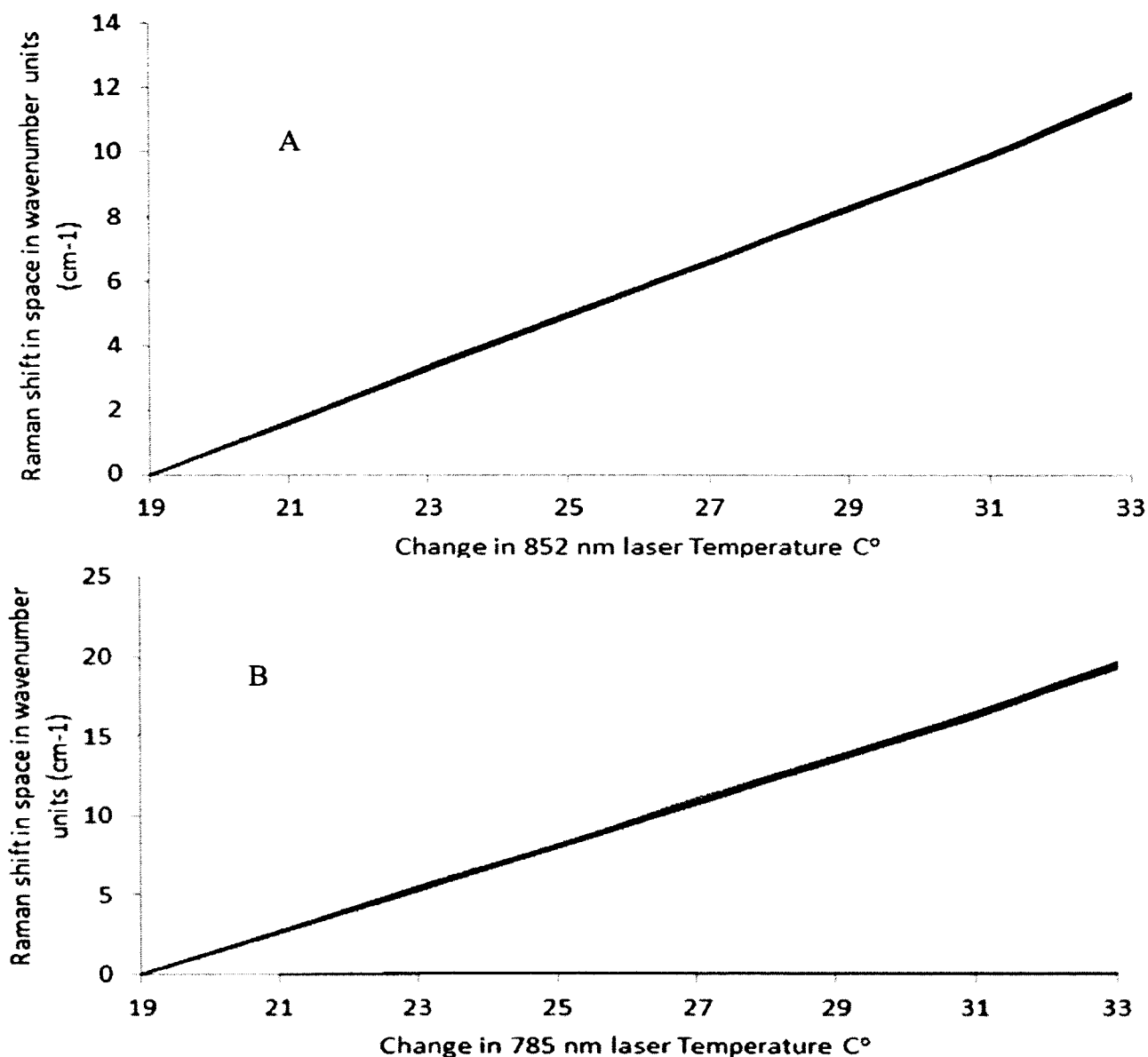


FIG. 42. This figure shows the plot between acetaminophen Raman shift and the diode laser temperature for the 852nm laser (A) and the 785nm laser (B). The x-axis is expressed in degree Celsius units.

TABLE 3. The standard deviation of Raman peaks of acetaminophen at different temperatures of the 852 and the 785nm diode laser.

laser temp °c	19	21	23	25	27	29	31	33
852nm stdv	0.0	0.029	0.040	0.034	0.034	0.039	0.034	0.048
785nm stdv	0.0	0.046	0.078	0.062	0.105	0.127	0.104	0.131

As shown in table 3, the 852 nm laser has lower standard deviations than the 785 nm. That is due to the fact that the 785 spectra are acquired first followed by the 852 spectra, which makes the temperature of the 852 diode laser more stable at acquisition time. A solution for that would be increasing the delay time after changing the temperature of the diode laser.

#### *UTILITY OF THE DUAL-LASER SSE IN REMOVING THE FLUORESCENCE BACKGROUND*

The Raman spectra of 2,3-naphthalene-diamine were acquired on the dual-laser SSE at three temperatures (22, 28, and 34 °C) utilizing the two lasers sequentially as described in the instrument operation section. Raman spectra of 2,3-naphthalene-diamine acquired on the dual-laser SSE utilizing the 785nm excitation source at the three different temperatures are shown in Figure 43A. Although it is difficult to resolve the spectra across the full spectral range, the change in laser temperature results in a shift of about  $7.2 \text{ cm}^{-1}$  for each of the spectra and the underlying background remains unchanged, as shown in the inset. Alternatively, Raman spectra of 2,3-naphthalene-diamine acquired on the dual-laser SSE utilizing the 852 nm laser at the three different temperatures are shown in Figure 43B. Although it is difficult to resolve the spectra across the full spectral range, the change in laser temperature results in a shift of about  $5.3 \text{ cm}^{-1}$  for each of the spectra and the underlying background remains unchanged, as shown in the inset. As shown in Figure 43, Raman spectra of 2,3-naphthalene diamine acquired utilizing the 785 nm laser have higher fluorescence background and show the CH-stretching region, while these acquired utilizing the 852 nm laser have less fluorescence background and show the fingerprint region. The SSE Raman spectra shown in Figure 43A and B are pretreated by the scaled SNV by applying Eqn. 43.

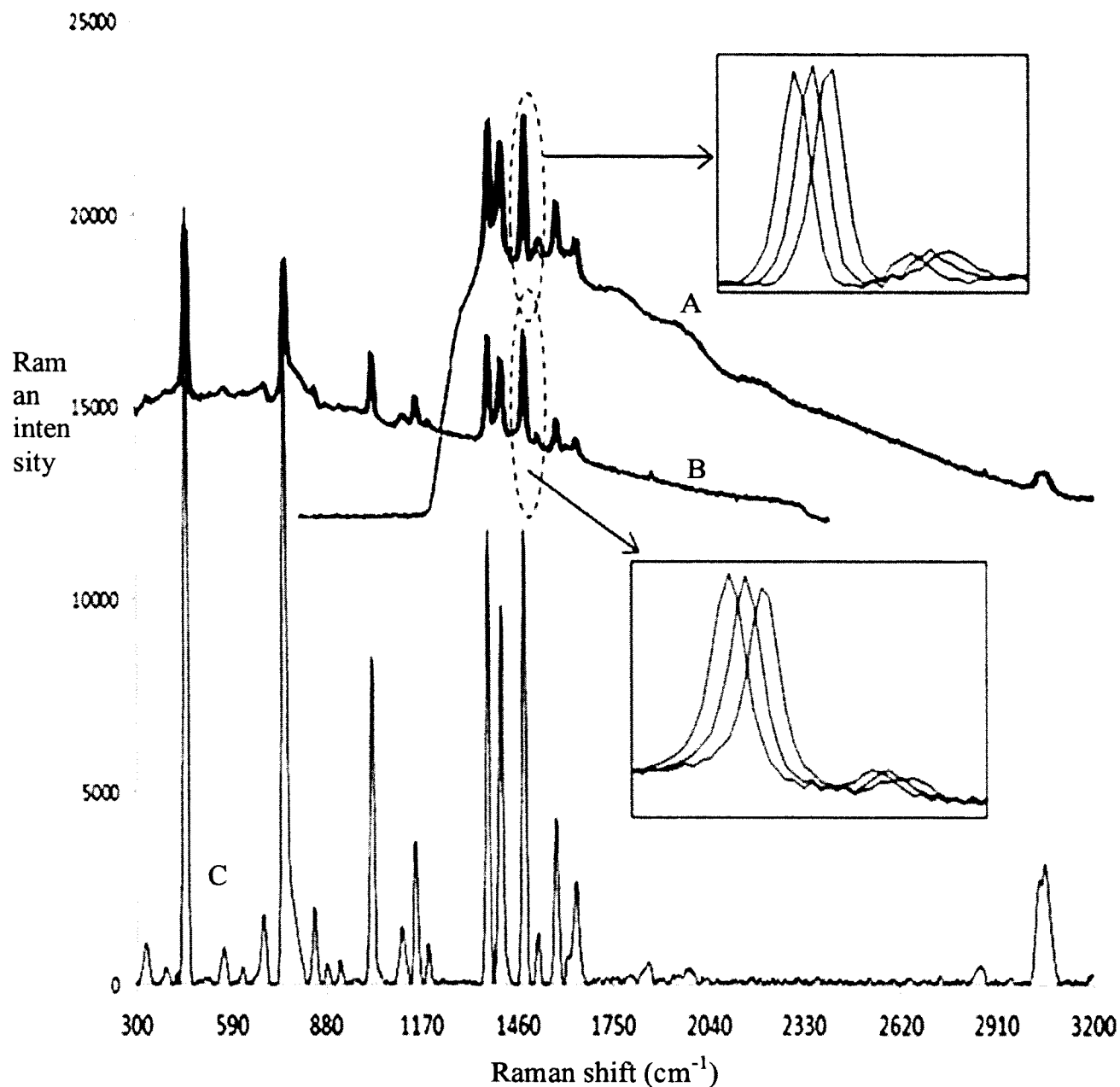


FIG. 43. The Raman spectra of 2,3-naphthalenediamine acquired at three laser temperatures (22, 28, and 34 °C) utilizing 785 nm DBR laser (A) and 852 nm DBR laser (B) with an integration time of 3 s for each spectrum and a laser power of 50 mW. The change in laser temperature results in a shift of 7.2 and 5.3  $\text{cm}^{-1}$  for each of the Raman spectra acquired utilizing 785 and 852 nm, respectively, while the underlying background remains unchanged (as shown in the insets). C is the demodulated and merged Raman spectrum of 2,3-naphthalenediamine.

The Fluorescent background is removed from the SSE Raman spectra of 2,3-

naphthalenediamine acquired utilizing the 785 and 852 nm excitation sources separately by applying the SSE algorithm (Eqn. 35-40). The 785 and 852 pure Raman spectra of 2,3-naphthalenediamine are then merged together at  $2000\text{ cm}^{-1}$ . The merged Raman spectrum of 2,3-naphthalenediamine is shown in Figure 43C. As shown, the merged Raman has no fluorescence background and spans the wavenumber range from  $300\text{ cm}^{-1}$  to  $3200\text{ cm}^{-1}$ , which include both the fingerprint as well as the CH-stretching region.

For comparison, Raman spectra of 2,3-naphthalenediamine were acquired on the single-laser SSE utilizing only the 785 nm laser (discussed in chapter II) at the same three temperatures (22, 28, and 34) and the spectra are shown in Figure 44A. The Pure Raman spectrum of 2,3-naphthalenediamine after removing the fluorescent background by the SSE algorithm is shown in Figure 44B. As can be observed, the two methods (using the single 785nm excitation source or dual-laser 785 and 852nm excitation sources) yield comparable results with both resulting in an absence of the fluorescent background. Thus, the performance of the single and the dual-laser SSE is the same in this example due to the low fluorescence profile of 2,3-naphthalenediamine. Since neither instrument was corrected for spectral intensity throughput, the relative intensities of the peaks are not identical. This is particularly true when comparing the fingerprint region to the CH-stretching region.



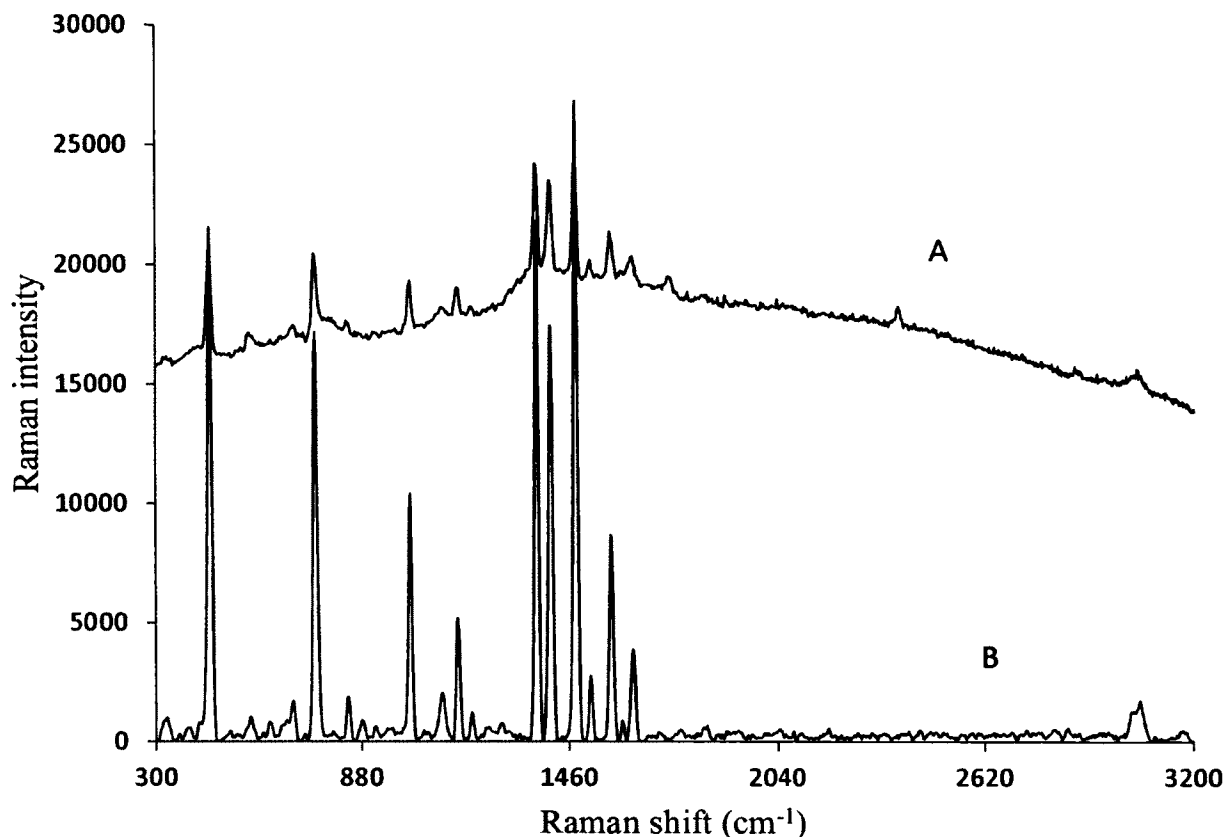


FIG. 44. This figure shows Raman spectrum of 2,3-naphthalenediamine acquired on the single-laser (785nm excitation source) SSE. (A) Raw Raman spectrum of 2,3-naphthalenediamine with an integration time of 3 s and a laser power of 50 mW. (B) Raman spectrum of 2,3-naphthalenediamine which is obtained using the 3 sequentially shifted excitation Raman spectra of 2,3-naphthalenediamine shown in A.

#### *MERITS OF THE DUAL-LASER AND THE SINGLE-LASER SSE*

Xanthan gum is one of the pharmaceuticals, which has very high fluorescence background, and it's a challenge to detect it using a dispersive handheld Raman analyzer. The xanthan gum Raman spectra were acquired on the two SSE devices (the single-laser and the dual-laser SSE Raman Analyzer). Figure 45I and II shows the Raman spectra of xanthan gum acquired on the dual-laser SSE and single-laser SSE, respectively. In Figure 45I, A gives the raw Raman spectrum of xanthan gum acquired on the dual-laser SSE utilizing the 785 nm as excitation source and B gives its spectrum when utilizing the 852 nm as excitation source. The integration time for each spectrum is 2 s and the average is 2. Three Raman spectra were acquired at the temperatures of 22, 28, and 34 °C for each laser. The total acquisition time is 26s (6 spectra x 2 s each x 2 averager + 2 s for

operation).

As can be observed, the Raman spectrum of xanthan gum acquired utilizing the 852 nm excitation source has much lower fluorescence background than that acquired when the 785 nm excitation source is utilized. The higher fluorescence background in the case of utilizing the 785 nm excitation source is due to the high energy of the 785nm laser which populates higher energy states of xanthan gum. The SSE algorithm is then applied separately on the raw Raman spectra of xanthan gum acquired utilizing the 852 and the 785 nm lasers (Figure 45I A and B) to remove the fluorescence background. The two pure Raman spectra are then merged at  $2000\text{ cm}^{-1}$  and the final spectrum is shown in Figure 45C.

For comparison, xanthan gum Raman spectra were also acquired on the single-laser SSE at the same conditions. Figure 45II shows the Raman spectrum of xanthan gum acquired on the single 785 nm SSE analyzer of; (D) raw Raman spectrum and (E) pure Raman spectrum of xanthan gum after removing fluorescence background. The integration time was kept the same as the dual-laser system and it was 2 s for each spectrum which makes the total acquisition time 14 s in the case of the single-laser SSE. As can be observed in Figure 45C and E, the pure Raman spectrum of xanthan gum acquired on the single-laser SSE has poor signal to noise ratio. In addition, some of the Raman peaks are missing due to the presence of high fluorescence background in the case of utilizing the 785 nm excitation source. Alternatively, the Raman spectrum of xanthan gum has good signal to noise ratio with all Raman peaks resolved when using the dual-laser SSE due to the low energy of the 852nm excitation source while we still getting the CH-stretching peaks due to the presence of the 785nm laser where the fluorescence is minimal.

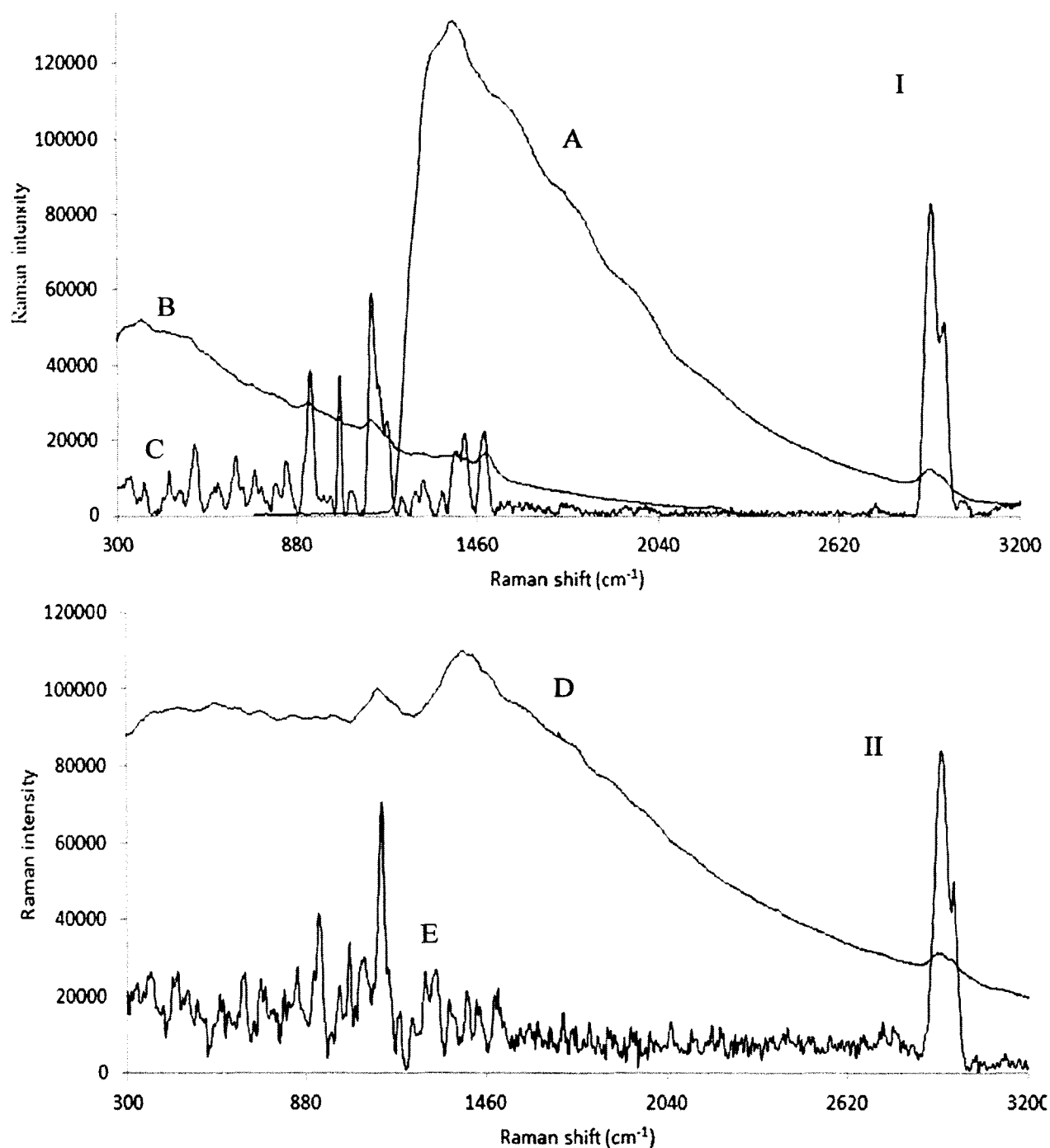


FIG. 45. The Raman spectra of xanthan gum acquired on the dual-laser SSE (I) and on the single-laser SSE (II). A and D give the raw xanthan gum spectra acquired on the dual and the single-laser SSE utilizing 785nm excitation source, respectively. B gives the raw xanthan gum spectrum acquired on the dual-laser SSE utilizing 852nm excitation source. C and E give the pure Raman spectra of xanthan gum after removing fluorescence-background of the raw spectra acquired on the dual-laser and the single-laser SSE, respectively.

For further comparison of the two devices, sodium carboxy methyl cellulose Raman spectra, a challenging sample, were acquired on the single-laser and dual-laser SSE. Figure 46I and II shows the Raman spectra of sodium carboxy methyl cellulose acquired on the dual-laser SSE and the single-laser SSE, respectively. In Figure 46I, A gives the Raman spectrum of sodium carboxy methyl cellulose acquired on the dual-laser SSE utilizing the 785 nm excitation source and B gives its spectrum when utilizing the 852 nm excitation source. The integration time for each spectrum is 3 s and the average is 2. Three Raman spectra were acquired at the following temperatures (22, 28, and 34) for each laser. The total acquisition time is 38 s (6 spectra x 3 sec each x 2 averager + 2 s for operation). As can be observed, the Raman spectrum of sodium carboxy methyl cellulose acquired utilizing the 852 nm excitation source has much lower fluorescence background than that acquired when the 785 nm excitation source is used. The high fluorescence background in the case of utilizing the 785 nm excitation source is due to the higher energy of 785nm laser which populates higher energy states of the sodium carboxy methyl cellulose. The SSE algorithm is then applied independently on the raw Raman spectra of sodium carboxy methyl cellulose (Figure 46I A and B) to remove the fluorescence background. The two pure Raman spectra are then merged at  $2000\text{ cm}^{-1}$  as shown in Figure 46C.

Figure 46II shows the Raman spectrum of sodium carboxy methyl cellulose acquired on the single-laser (785 nm) SSE analyzer for; (D) the raw Raman spectrum and (E) the pure Raman spectrum of sodium carboxy methyl cellulose after removing the fluorescence background. The integration time was kept the same as the dual-system and it was 3s for each spectrum with average of 2 which makes the total acquisition time 20 s. As can be observed in Figure 46C and E, the pure Raman spectrum of sodium carboxy methyl cellulose acquired on the single-laser SSE has poor signal to noise ratio. In addition, some of the Raman peaks are missing due to the presence of high fluorescence background because of utilizing the 785 nm laser as excitation source. However, the Raman spectrum of sodium carboxy methyl cellulose has good signal to noise ratio with all Raman peaks resolved and present when using the dual-laser SSE due to the low energy of the 852nm excitation source while we still getting the CH-stretching peaks due

to the presence of 785nm laser.

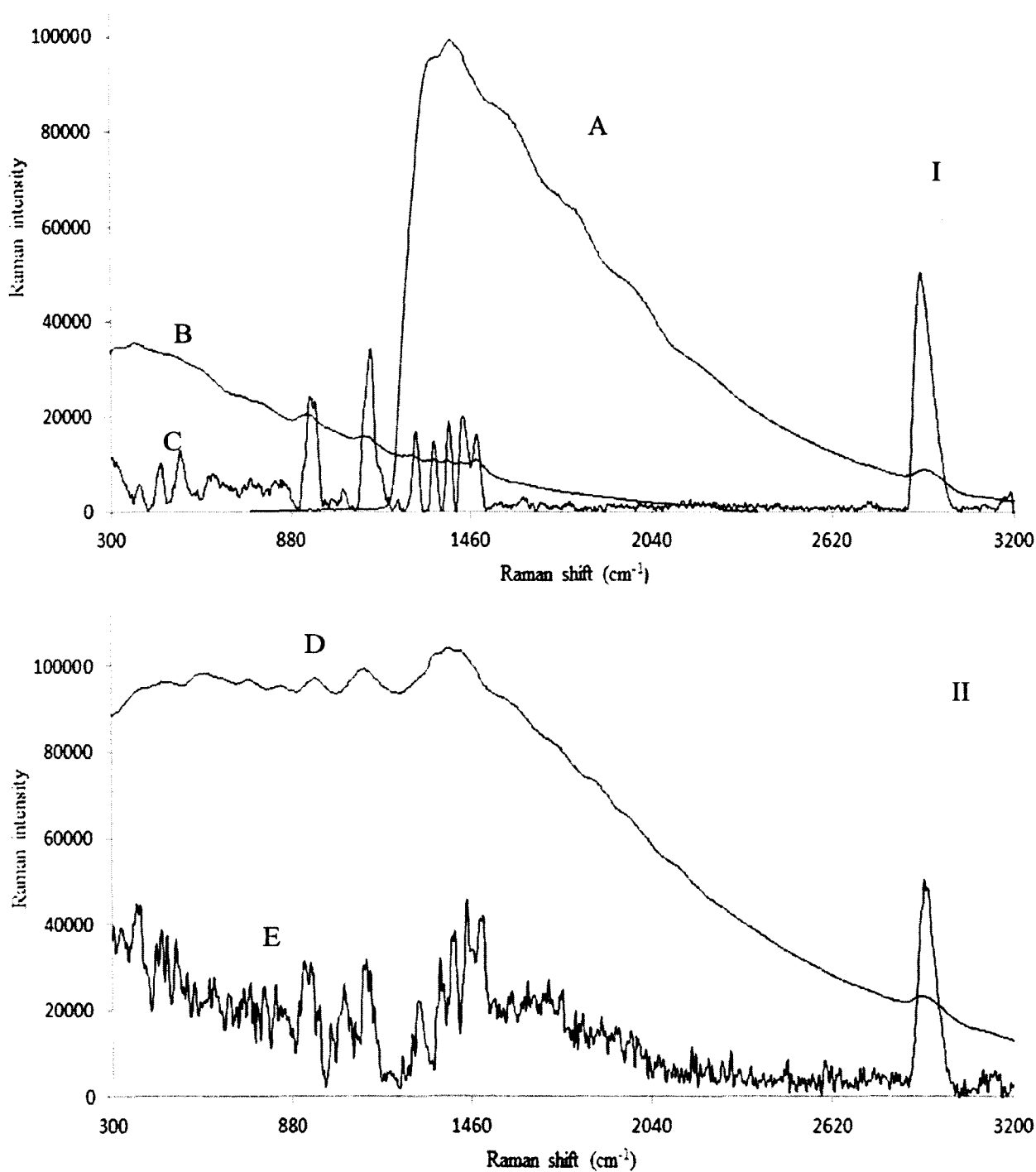


FIG. 46. The Raman spectra of sodium carboxy methyl cellulose acquired on the dual-laser SSE (I) and on the single-laser (785nm laser) SSE (II). The A and D give the raw sodium carboxy methyl cellulose spectra acquired on the dual and the single-laser SSE utilizing 785nm excitation source, respectively. B gives the raw sodium carboxy methyl cellulose spectrum acquired on dual SSE utilizing 852nm excitation source. C and E give the pure Raman spectra of sodium carboxy methyl cellulose after removing fluorescence-background of the raw spectra acquired on the dual and single-laser SSE, respectively.

The two previous examples clearly prove the superiority of using the dual-laser SSE in the case of very high fluorescent samples. However, this is not the case when samples have no or low fluorescence background. In order to show the performance of the two SSE devices in the case of low fluorescence background, niacinamide (nicotinic acid amide) Raman spectra were acquired on the two SSE devices. Figure 47I and II shows Raman spectra of niacinamide acquired on the dual-laser SSE and the single-laser SSE, respectively. In Figure 47I, A gives the Raman spectrum of niacinamide acquired on the dual-laser SSE utilizing the 785 nm as excitation source and B gives its spectrum when utilizing the 852 nm excitation source. The integration time for each spectrum is 1 s. Three Raman spectra were acquired at the temperatures of (22, 28, and 34) for each laser. The total acquisition time is 8 s (6 spectra x 1 sec each+ 2 s for operation). As can be observed, the niacinamide Raman spectra acquired utilizing the dual-laser have low fluorescence background. The SSE algorithm is then applied on the raw Raman spectra of niacinamide (Figure 47I A and B) to remove the fluorescence background. The two pure Raman spectra are then merged at  $2000\text{ cm}^{-1}$  and the final spectrum is shown in Figure 47C.

Figure 47II shows the Raman spectrum of niacinamide acquired on the single-laser (785 nm) SSE analyzer for; (D) raw Raman spectrum, (E) and pure Raman spectrum of niacinamide after removing the fluorescence background. The integration time was kept the same as the dual-system, and it was 1 s for each spectrum. This makes the total acquisition time 5 s in the case of using the single-laser SSE. As can be observed in Figure 47C and E, both the Raman spectrum of niacinamide acquired on the dual-laser and the single-laser SSE have good signal to noise ratio and all the fluorescence background is removed thus the single-laser system performs as well as the dual-laser system in the case of low fluorescent recipients while the single-laser system utilizes only one laser and has lower acquisition time.

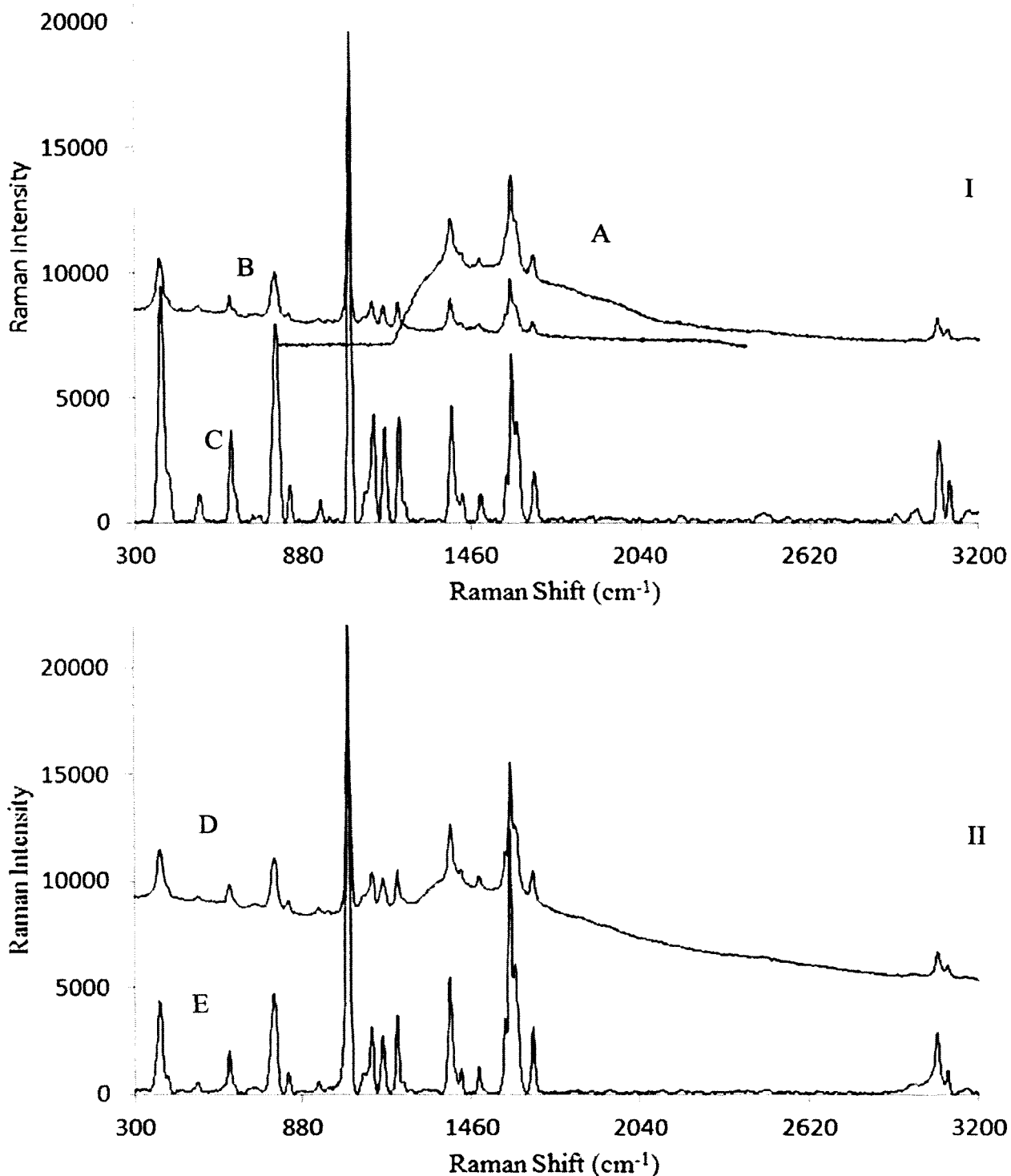


FIG. 47. The Raman spectra of niacinamide acquired on the dual-laser SSE (I) and on the single-laser (785nm laser) SSE (II). A and D give the raw niacinamide spectra acquired on the dual and the single-laser SSE utilizing 785nm excitation source, respectively. B gives the raw xanthan gum spectrum acquired on the dual-laser SSE utilizing 852nm excitation source. C and E give the pure Raman spectra of niacinamide after removing the fluorescence-background of the raw spectra acquired on the dual and single-laser SSE, respectively.

As can be observed, the primary usefulness of the SSE Raman method with the described instrumentation lies in its universal ability to extract Raman spectra from fluorescence interference using inexpensive and compact instrumentation in a time efficient manner. Using the dual-laser SSE allows extracting Raman spectra from high fluorescence samples while having a full Raman spectrum that includes both the fingerprint as well as CH-stretching region. In addition, the SSE Raman is the only method which allows all of the excitation conditions to be easily varied in a predefined manner (number of excitations, separation of excitations, and integration time of each excitation) in order to obtain optimal results. Even in the absence of a background, SSE Raman improves the S/N ratio by reducing random shot and thermal noise and by eliminating fixed pattern and random spike noise. In the current work, an un-cooled CCD detector is used. The use of a cooled CCD would allow much longer integration times to be used for weak Raman scatters without having the thermal noise impact the dynamic range of the 16 bit A/D converter.

Importantly, the ability to tune the DBR lasers to set wavelengths allows the excitation shift to be set to a constant value. This simplifies implementation of the iteration algorithms since the shift index of the weighting vector (Eqn. 35) can be set equal to the constant excitation shift (or an integral number multiplied times the excitation shift) and thus allows a single algorithm to be used even for different instruments. The Dual-laser SSE shows superior performance in the case of highly fluorescence samples, while both perform the same in the case of low or no fluorescence background.

### *LASER SHIFT EFFECT*

The magnitude of Raman shift in space between the acquired Raman spectra at different temperatures was evaluated to determine the proper shift to be used in the SSE dual-laser analyzer. In order to test the effect of Raman peak shift in space on the SSE algorithm efficiency and eventually on the quality of the separated pure Raman spectra, SSE Raman spectra of microcrystal cellulose were acquired at different temperatures profile. Microcrystal cellulose was chosen and shown here since it has high fluorescence background making it tedious on the SSE algorithm. Figure 48A gives the Raman



spectrum of microcrystal cellulose acquired on the dual-laser SSE utilizing the 785 nm excitation source and B gives the microcrystal cellulose Raman spectrum when utilizing the 852nm excitation source. As shown in Figure 48, using the 785 nm excitation source produces higher fluorescence background as expected.

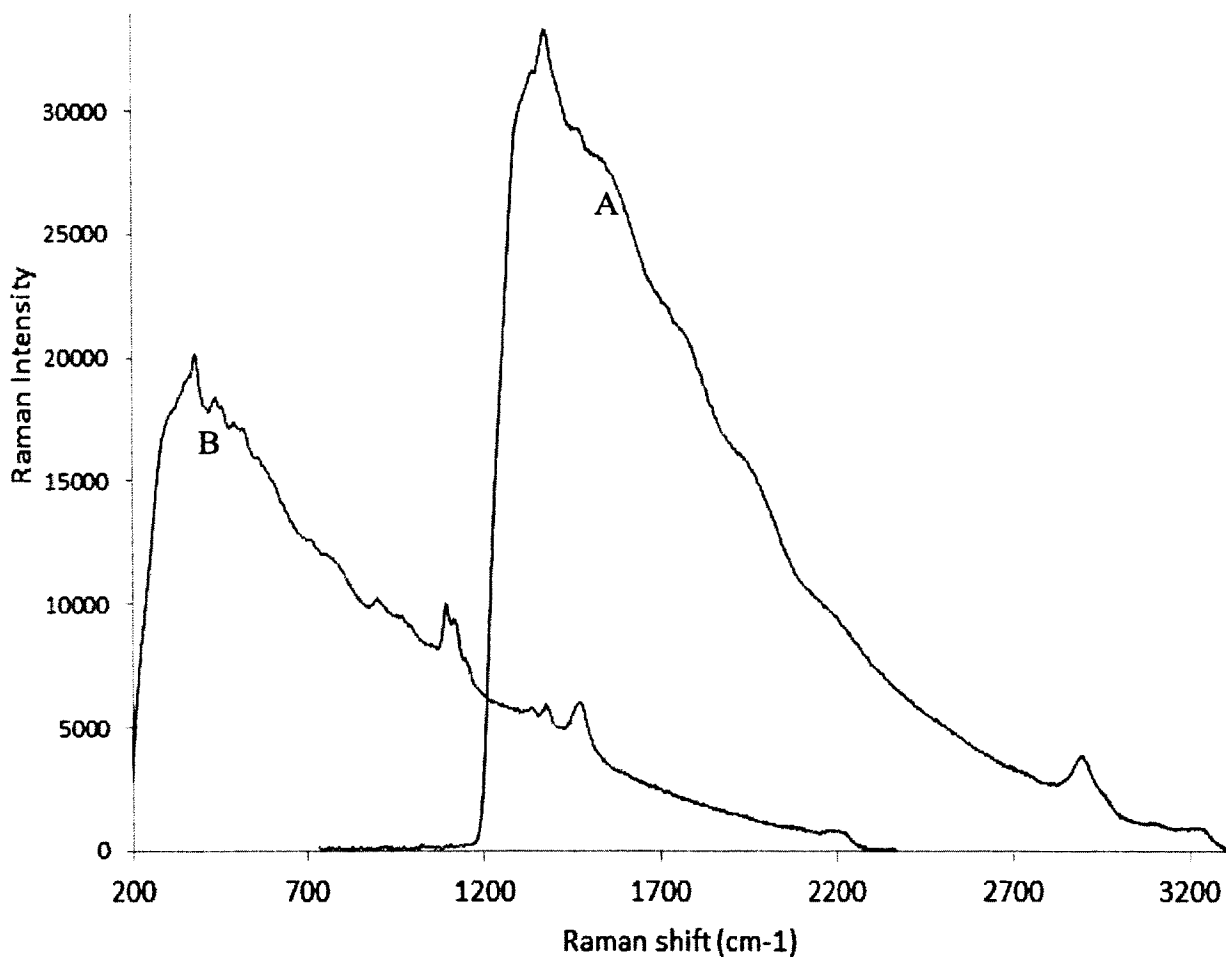


FIG. 48. The raw Raman spectra of microcrystal cellulose acquired on the dual-laser SSE utilizing 785 nm excitation source (A) and 852 nm excitation source (B). The integration time of each spectrum is 2 s. The x-axis is expressed in Raman shift wavenumber units.

The Raman spectra of microcrystal cellulose were acquired at six different experiments of different temperature profiles. The temperatures intervals of the six experiments are 2, 3, 4, 5, 6 and 7 °C for experiment one to six, respectively. The first temperature of all the six experiments was 20 °C. All experiments were acquired using three temperatures separated by the previous mentioned temperatures intervals. All the

experiments were acquired with the integration time of 2 s for each spectrum. The SSE pure Raman spectra of microcrystal cellulose, after removing the fluorescence background by applying the SSE algorithm, are shown in Figure 49 A, B, C, D, E, and F for the experiments one to six, respectively.

Using two degree Celsius separation between the three temperatures of the diode lasers (20, 22, and 24 °C, experiment one, Figure 49A) leads to low SNR of the microcrystal cellulose SSE Raman spectrum. A small shift in Raman signal results in increasing the number of data points of the SSE raw Raman spectra in the linearization step, and that leads to poor SNR since the SSE algorithm consider the random noise as Raman information at very small shift. As the laser temperature separation increases, Raman shift in space increases, and the signal to noise ratio of the separated Raman spectrum increases if there is no change in the optical interferences intensities. The optimal results are expected when the shift is close to the half Raman peaks width. The highest signal to noise ratio is obtained when using 6 degree Celsius separation (20, 26, and 32 °C), as shown in Figure 49E. Increasing the Raman shift in space more than the optimal shift leads to decreasing the SNR due to two reasons; 1) at high temperature separation, the optical interferences intensity starts to change, 2) a high Raman shift in space results in low number of data points of the raw Raman spectra which leads to a low resolution and broadening.

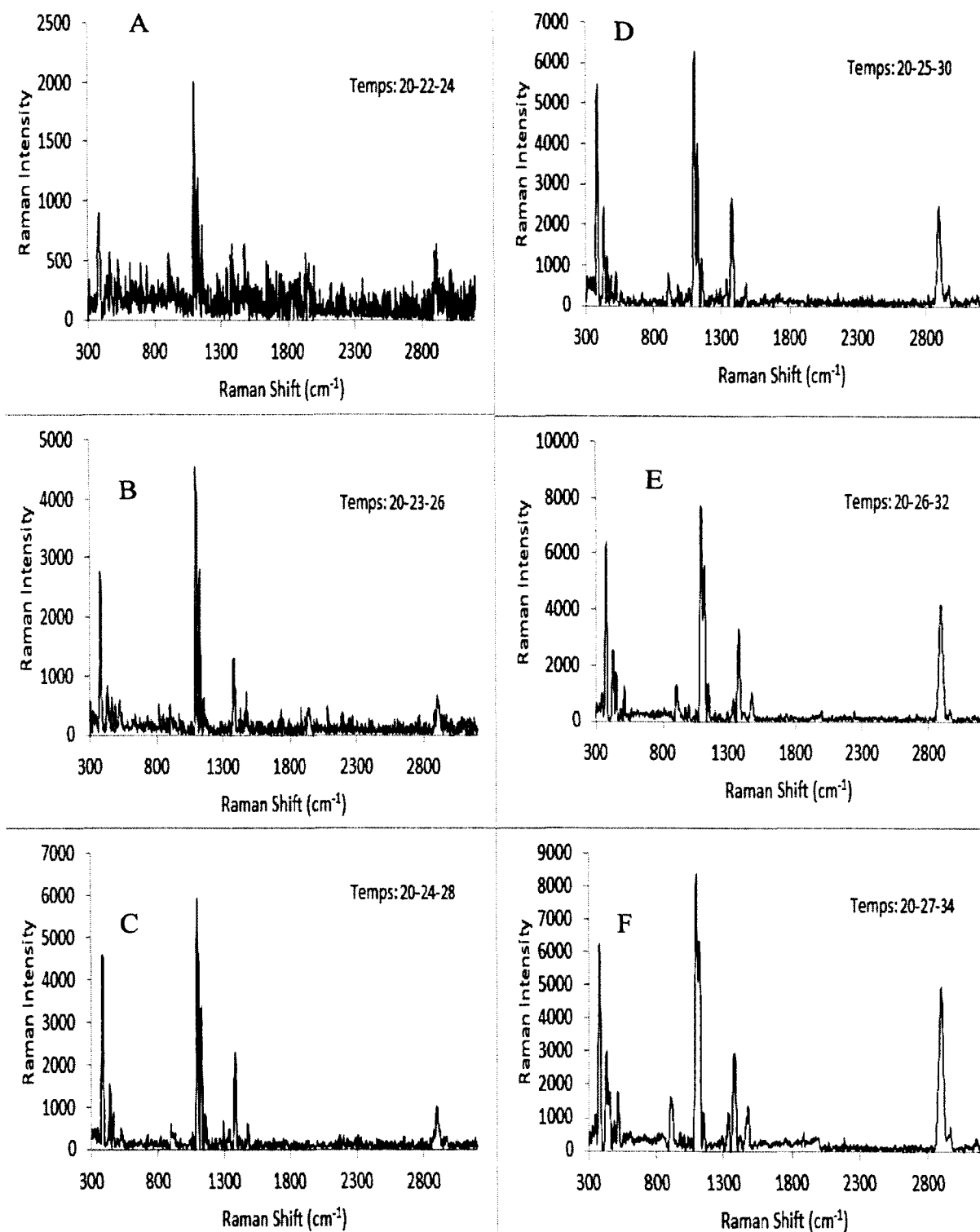


FIG. 49. The SSE Raman spectra of microcrystalline cellulose acquired on the dual-laser at six different temperature profiles. In all experiments three temperatures are acquired and the first temperature is always 20 °C. The temperatures spacing of the six experiments are 2, 3, 4, 5, 6, and 7 for the experiments 1 - 6 (given by A to F), respectively. The integration time for each spectrum was 2 s which makes the total acquisition time of each experiment 14 s.

## INSTRUMENT CALIBRATION

### *AXIS CALIBRATION*

There are two ways to calibrate the x-axis of a dispersive Raman device. These are the use of the absolute wavenumber calibration using the emission line spectra of gases, such as neon and argon or a mixed of both, and the use of the Raman shift scatterers standards (e.g., cyclohexane, acetaminophen, acetonitrile, and benzonitrile). The choice of a standard depends on the number of well-established Raman peaks of the standard on the wavenumber range of interest. The Raman standards scatterers are a better choice since the excitation source wavelength does not have to be measured independently. Acetaminophen was chosen to calibrate the dual-laser SSE Raman instrument since it has more than 20 Raman shift peaks which span the wavenumber range we are interested in (200 to 3200  $\text{cm}^{-1}$ ). Each laser was calibrated separately and independently using the acetaminophen to generate a calibrated axis for 785 and 852 Raman spectra.

Eleven acetaminophen peaks were used to calibrate the 785nm x-axis of the dual-laser SSE Raman. The eleven 785 peaks of acetaminophen used in the calibration are shown in Figure 50A. Among the eleven peaks, there are three peaks located in the CH-stretching region and nine peaks in the finger print region.

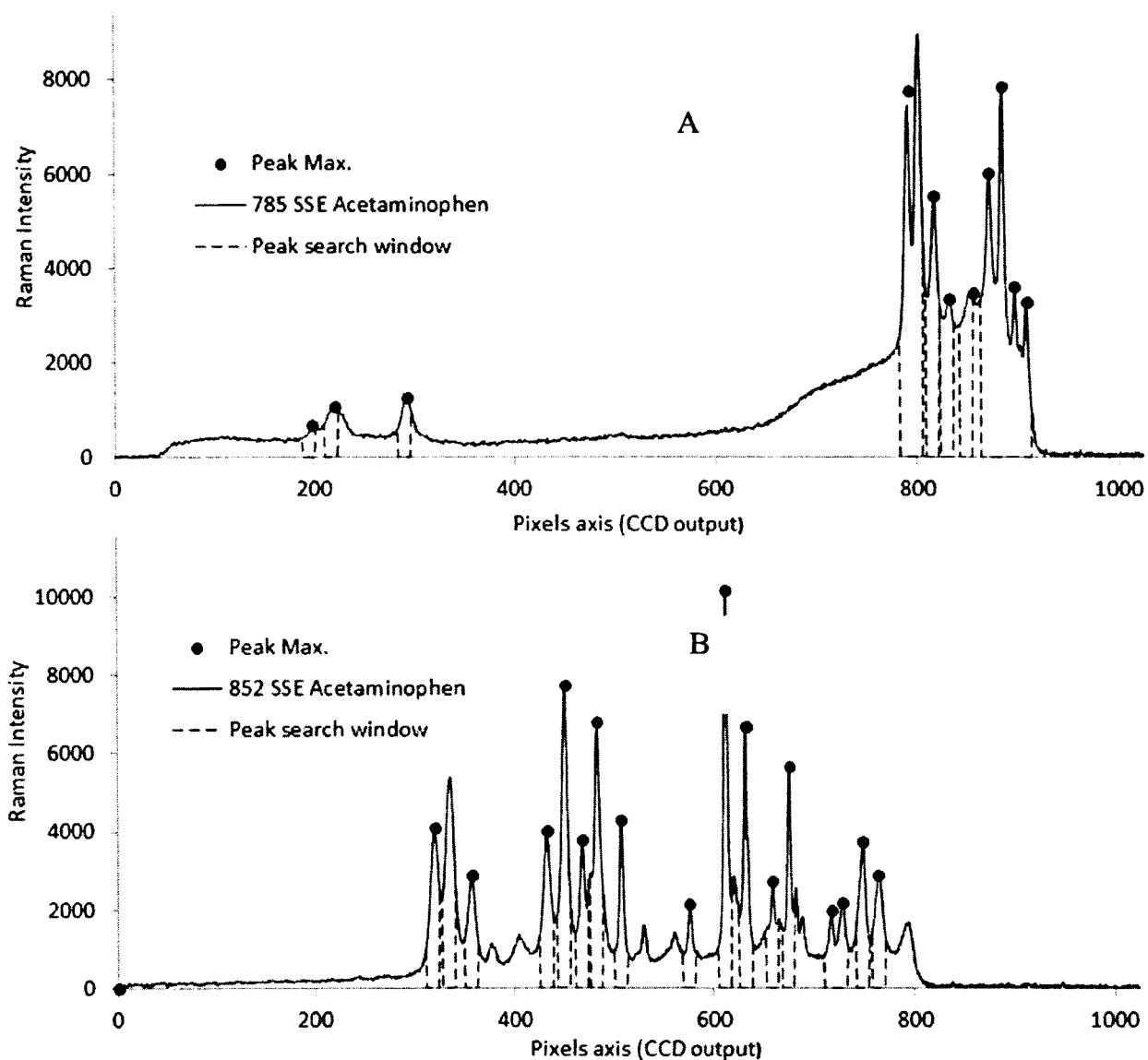


FIG. 50. This figure shows the 11 peaks of acetaminophen used in the 785nm laser axis calibration (A) and the 16 peaks used in the 852nm laser axis calibration (B). A window of size 13 indices was assign around each peak maximum to automate its future picking. The x-axis is expressed in unit of pixels.

The eleven peaks were chosen so that they cover the whole spectral range we are interested in. The Peaks were chosen based on their locations, shapes, consistency, and intensity. In order to automate the calibration procedure, software was written to automatically pick the peaks of interest. A window of a thirteen indices size was assigned around a peak's location for its future picking. A peak window is then spline 100X times and the maximum of the splined-window is used as the peak location. As soon as a peak maximum is found its entire window is replaced by zeros in order to prevent overlapping and picking the same peak twice.

To calibrate the 852 laser axis sixteen peaks were used that span the entire spectral range of interest. The peaks were automatically picked by assigning a peak window that has thirteen points window size (the same as the case of the 785nm axis) around each peak. A peak window is then spline 100X times and the maximum of the spline is then chosen as the peak maximum (i.e., peak location). The sixteen peaks of the acetaminophen used in calibration of the 852 axis are shown in Figure 50B.

The dual-laser SSE 785 and 852nm acetaminophen Raman peaks used in the axes calibration are shown in table 4 (with pixels units) along with their corresponding ASTM certified acetaminophen peaks in absolute wavelength (nm) unit.

TABLE 4. Acetaminophen SSE Raman peaks in pixels units with their ASTM certified Raman shift in wavelength units for both 852 and 785 nm excitations lasers.

852		785	
SSE peaks (pixels)	ASTM peaks (absolute wavenumber nm)	SSE peaks (pixels)	ASTM peaks (absolute wavenumber nm)
319.91	991.182	197.98	1158.056
356.39	982.725	220.06	1153.046
432.34	964.7	293.04	1135.562
450.19	960.373	791.49	991.262
467.7	956.091	816.59	982.676
482.81	952.366	831.74	978.314
507.39	946.243	854.96	971.554
575.78	928.771	870.92	964.771
612.03	919.227	883.84	960.322
630.78	914.208	896.34	956.151
657.37	906.991	907.45	952.378
675.08	902.115		
717.6	890.18		
728.07	887.19		
748.4	881.323		
764.06	876.748		

The ASTM Raman shift of acetaminophen in wavenumber units were changed to wavelength (nm) since there is a better linearity using a wavelength axis. A third order polynomial was generated between the acetaminophen peaks locations in pixel units, acquired on the dual-laser SSE Raman, and their corresponding ASTM peaks values in wavelength Raman shift units for each laser. The 3<sup>rd</sup> order polynomial was found to have the minimum error. Figure 51A shows the 3<sup>rd</sup> order polynomial generated between the eleven peaks of acetaminophen in pixels (SSE ccd detector throughput) and their ASTM Raman shift in absolute wavelength units (given in table 4). As can be observed in Figure 51A, a 3<sup>rd</sup> order polynomial linearly fits the pixel axis to the ASTM wavelength axis in the case of utilizing the 785nm as excitation source.

Another 3<sup>rd</sup> order polynomial was generated between 852 acetaminophen peaks in pixels units and their corresponding ASTM Raman shift in absolute wavelength (nm) unit, the 3<sup>rd</sup> order polynomial is shown in Figure 51B.

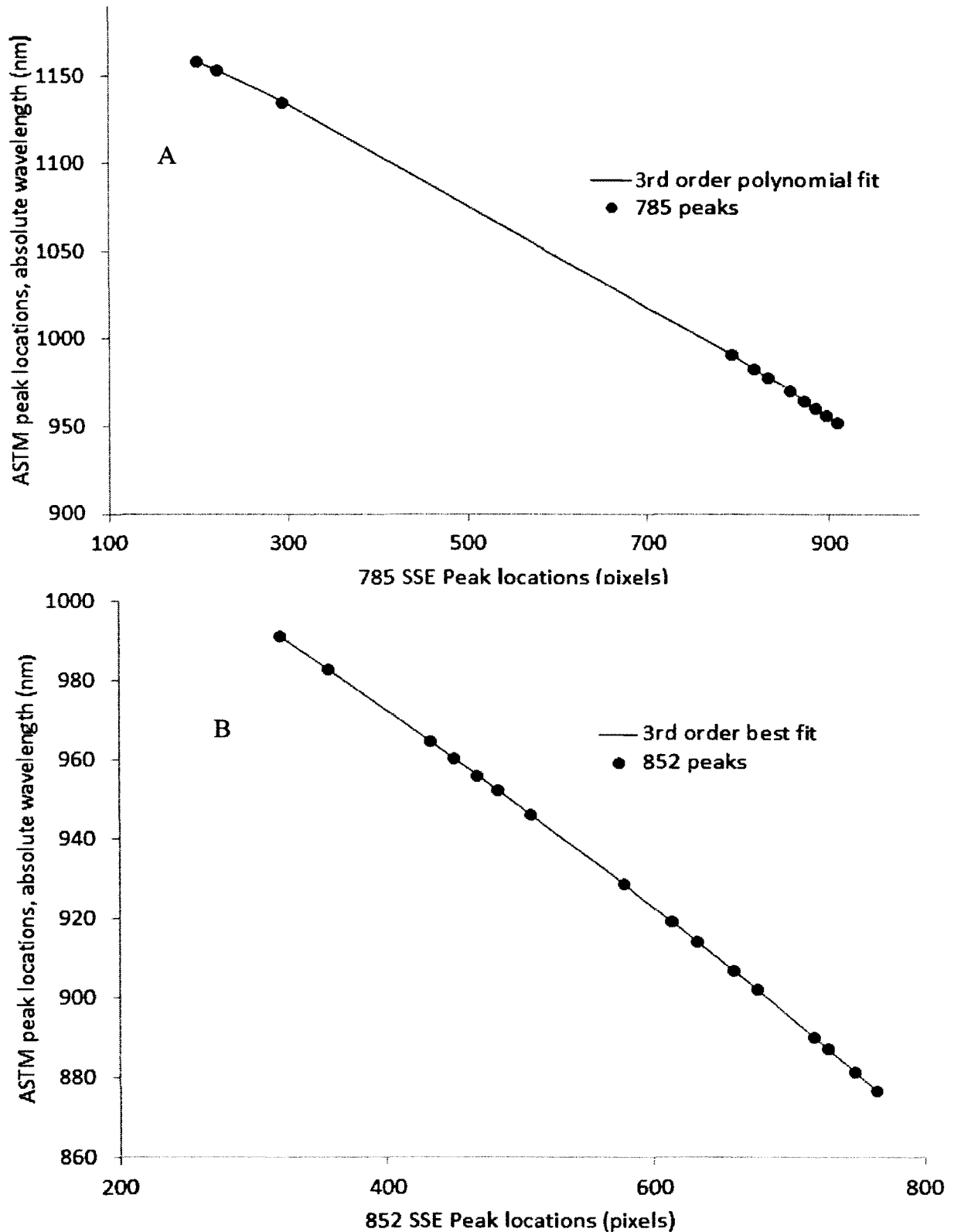


FIG. 51. This Figure shows the two 3<sup>rd</sup> order polynomials; A) between the 11 acetaminophen peaks of SSE (in pixels units) and their ASTM certified wavelength Raman shift (solid line), and B) between the 16 acetaminophen peaks of SSE (in pixels units) and their ASTM certified wavelength Raman shift (solid line).



The polynomial coefficients of the 785 and the 852 lasers were used to generate a calibrated x-axis for the 785 and the 852 Raman data according to the following Eqn. 41:

$$\lambda_{\text{cal}}(i) = a_0 + a_1 x(i)^1 + a_2 x(i)^2 + a_3 x(i)^3 \quad (41)$$

where “i” gives the pixels index and it takes the values 0 to 1023 (CCD detector has 1024 pixels) for each laser and  $a_n$ 's are the polynomials coefficients. The Raman shift axis in wavelength (nm) unit, generated using equation 41, is then converted to wavenumber ( $\text{cm}^{-1}$ ).

Acetaminophen spectrum acquired on the dual-laser SSE utilizing the 785nm laser with a calibrated x-axis is plotted with the FT-acetaminophen spectrum in Figure 52A. As shown in Figure 52A, the two spectra have the same x-axis in units of wavenumber Raman shift ( $\text{cm}^{-1}$ ). In addition, Figure 52B overlays the SSE acetaminophen spectrum acquired utilizing the 852nm excitation source with a calibrated axis with the FT-acetaminophen spectrum and as can be observed the acetaminophen peaks of the two axes perfectly overlay. The acetaminophen SSE spectrum has a little fluorescence background since it is shown before applying the fluorescence removing algorithm. The two axes (785 and 852) are then merged at the wavenumber 2000 ( $\text{cm}^{-1}$ ) to form one axis that can be used in the merged final spectrum.

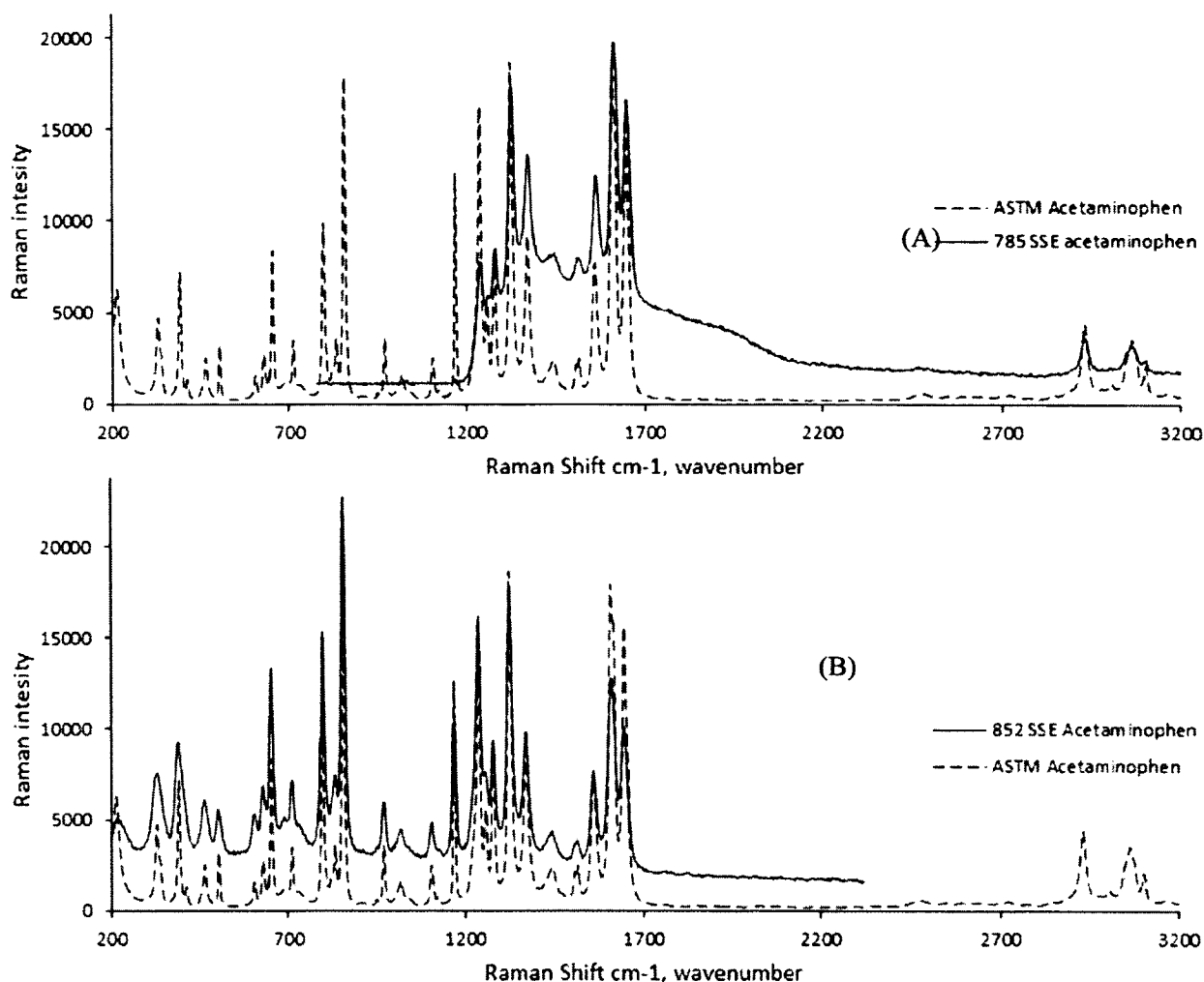


FIG. 52. The acetaminophen spectra acquired on the dual-laser SSE Raman after the axis calibration (solid line) overlaid with the FT-acetaminophen spectrum (dotted line). A and B give the acetaminophen Raman spectra acquired utilizing 785 nm and 852 nm laser, respectively. The x-axis is expressed in unit of wavenumber Raman shift ( $\text{cm}^{-1}$ ).

### AXIS VALIDATION

To validate the axis accuracy, acetaminophen spectra were acquired 30 times on the dual-laser SSE Raman. The 30 spectra were collected in two days early at morning and before leaving at evening. The sample position was changed before each acquisition. Ten peaks of acetaminophen were chosen to carry out the standard deviation calculations. The Raman shift in wavenumbers units of the 10 acetaminophen peaks used in this validation are shown in table 5. The row one in table 5 gives the 10 peaks ASTM Raman shift in wavenumbers units, while row two gives their corresponding ASTM standard deviation (reproducibility error associated with each peak), and the third row gives the

standard deviation (reproducibility error) at each peak acquired 30 times on the dual-laser SSE. As can be observed, all dual-laser SSE standard deviations lie within ASTM error.

TABLE 5. ASTM peaks of acetaminophen in Raman shift wavenumber units along with their ASTM and SSE standard deviation error.

ASTM peaks frequency ( $\text{cm}^{-1}$ )	329.2	465.1	651.6	857.9	968.7	1236.8	1323.9	1648.4	2931.1	3102.4
ASTM standard deviation	0.52	0.3	0.5	0.5	0.6	0.46	0.46	0.5	0.63	0.95
SSE standard deviation	0.24	0.22	0.18	0.13	0.37	0.08	0.16	0.38	0.35	0.88

In order to further validate the accuracy of the two lasers axes, polystyrene spectra were acquired on the dual-laser SSE after axes calibration. Polystyrene spectra acquired on the dual-laser SSE utilizing the 785 nm excitation source is shown in Figure 53A, while its Raman spectrum when utilizing the 852 nm excitation source is shown in Figure 53B. Both the polystyrene spectra, shown in Figure 53, have calibrated x-axes and as can be observed, the overlap peaks are at the same Raman shift in wavenumber units. The SSE Raman peaks of polystyrene in wavenumber units were compared to their corresponding ASTM Raman shift values. The SSE Raman peaks in wavenumber units along with their corresponding ASTM Raman shift in wavenumber, and the standard deviation at each peak are shown in Table 6. As can be observed, the all polystyrene peaks were found to be within the ASTM standard deviation error.

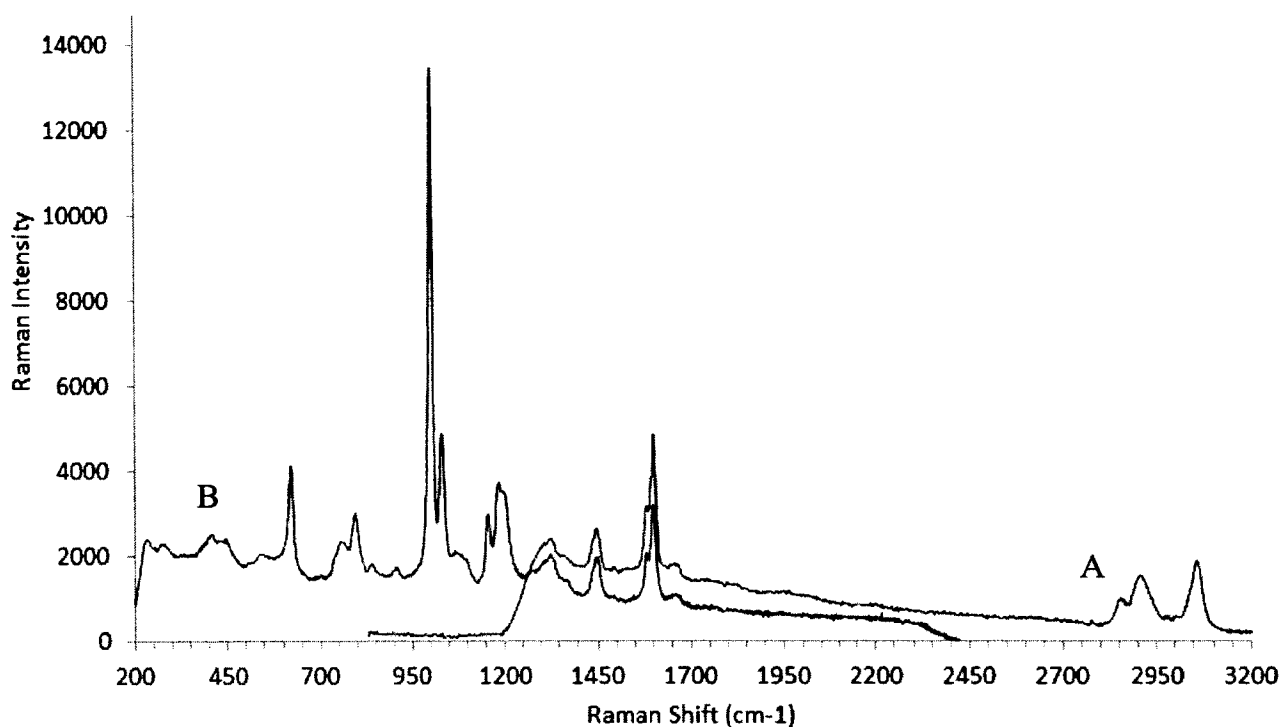


FIG. 53. The Polystyrene Raman spectra acquired on the SSE dual-laser utilizing the 785 nm excitation source (A) and the 852 nm excitation source (B). The x- axis is calibrated and expressed in wavenumber Raman shift. The acquisition time of each spectrum is 2 s.

TABLE 6. ASTM Raman shift of polystyrene peaks in wavenumber with their standard deviation error compared to their SSE Raman shift in wavenumber.

ASTM mean frequency (cm <sup>-1</sup> )	Standard Deviation	SSE mean frequency (cm <sup>-1</sup> )
620.9	0.69	620.81
795.8	0.78	795.43
1001.4	0.54	1001.8
1031.8	0.43	1031.8
1155.3	0.56	1155.36
1450.5	0.56	1448.63
1583.1	0.86	1583.41
1602.3	0.73	1602.13
2852.4	0.89	2852.34
2904.5	1.22	2905.28
3054.3	1.36	3055.12

## CHAPTER IV

### CALIBRATION TRANSFER OF PLS MODELS USING VIRTUAL STANDARDS

#### SLOPE AND BIAS

#### INTRODUCTION

Due to the correlation of infrared overtone bands with the chemical and physical properties of fuels, near-IR spectroscopy has gained wide-spread acceptance as a secondary analytical method in fuel analysis.<sup>29-44</sup> Typically using the partial least squares method, NIR spectra of fuels are correlated with their known ASTM determined properties. Near-IR analysis of fuels provides QA/QC feedback in real-time for fuel properties which traditionally require 1-2 days of laboratory analysis time per sample. Although PLS models can be built in a laboratory environment using a spectrometer and a set of calibration fuels, when the models are transferred to different instruments or used in a more demanding environment, the predictive ability of the models suffers greatly. This is a widely recognized phenomenon and can be attributed to subtle changes in the instrumentation during manufacture, transport, use, and storage.<sup>45, 46</sup>

Since full calibration of an instrument can often require the measurement of hundreds of samples and since the samples used for calibration may change with time, it is desirable to avoid performing a full calibration on each additional instrument due to the prohibitive associated costs. There are three common approaches to facilitate predictions on secondary instruments. The simplest and most common is the use of a spectral transform or combination of transforms. Common transforms include sample mean-centering, derivatives<sup>47</sup>, multiplicative scatter correction<sup>48</sup>, standard normal variate<sup>49</sup> and filtering algorithms.<sup>50-52</sup> Using transforms enables correction for changes in baselines and signal-to-noise ratios which invariably occur, however, they cannot correct for changes in an instrument's bandwidth or dispersion, or a detector's response linearity. These are more typically corrected using more sophisticated calibration transfer methods. These calibration transfers fall into two broad categories. spectral transfer and model transfer.

For spectral transfer, the spectra on the new instrument or spectra acquired in the new environment (secondary instrument) are made to resemble the spectra acquired on

the instrument used for the original calibration (master instrument). These methods include direct standardization<sup>53</sup>, piecewise direct standardization<sup>54, 55, 45, 56, 57, 53</sup>, orthogonal signal correction<sup>58-62</sup>, as well as other variants.<sup>63-74</sup> In all cases, there is a requisite to acquire standard spectra on the secondary instruments before calibration transfer can be performed. In addition, spectral transfer only works if the spectral standards resemble the spectra of the sample being analyzed, i.e., the standards at a minimum must have absorbance in the same spectral regions.<sup>75</sup> Also, the standards must include enough variability to account for the variations between the instruments; therefore a single sample cannot be used.

The alternative to spectral transfer is model transfer. In this category, the original master calibrated model is adjusted for the secondary instrument so that the predicted values closely match the values given by the master instrument. Methods for implementing this correction range from a simple slope bias correction<sup>54, 55, 76, 77</sup>, to model updating,<sup>78-82</sup> and transfer of the regression Eqn.<sup>83</sup>. However, regardless of the method used to implement the model transfer, there is still a requisite for a number of samples which closely resemble the samples being used to be measured on both the master and secondary instruments prior to transfer. Small and coworkers have introduced a calibration transfer method which requires no samples to be run on the instrument to be corrected, instead new background information from the instrument is used with component spectra from the master to simulate sample spectra.<sup>84, 85</sup> However, simulation of the spectra is only possible using this method if all of the components in the sample matrix are known.

For fuel analysis, the above methods present a dilemma since the fuels which the calibration samples must mimic consist of thousands of chemicals. Ideally, a range of standard fuels could be used, however, maintaining the integrity of these standards over time and during use and transport is a challenge due to contamination and compositional changes resulting from component volatility. Even if the integrity of the fuel standards is maintained, there is the dilemma of replicating standards of such complexity at a later date. To overcome this, the use of synthetic representative fuels is often used.<sup>77</sup> These standards are generated by titrating several pure chemical components (i.e., toluene, dodecane, etc.) to generate mixtures which are spectrally representative of the fuels

being analyzed. Although this overcomes the dilemma of standard replication, it does not solve the problem of maintaining sample integrity unless fresh standards are used. This adds a tremendous work load to the process. The use of synthetic representative fuels also introduces a new problem: the variability in the titration process itself. To overcome all of these limitations, a new model transfer method is described in this chapter, involving the use of digitally synthesized fuels generated from neat solvent spectra to form virtual standards for calibration transfer. The method requires no knowledge of the fuel composition and can be carried out with as few as twelve neat solvents. In this chapter, neat solvents refer to distilled liquid compounds which are commercially available with a known chemical purity.

## EXPERIMENTAL SECTION

### *INSTRUMENTATION*

Two FUELex NIR spectrometers (BrukerOptics, The Woodlands, TX, [www.brukeroptics.com/fuelex](http://www.brukeroptics.com/fuelex)) were used to collect the NIR spectra of fuels and chemicals. The FUELex consists of an  $f\#/2$  spectrograph with a 512 element InGaAs array detector and 3 nm resolution. The spectral wavelength range is 880 - 1570 nm. Each spectral acquisition consisted of 30 averaged spectra each with an integration time between 8 to 16 ms. The integration time for each instrument was automatically selected so that the signal intensity of an air reference yielded between 32,000 and 42,000 counts on the 16 bit A/D converter. This ensured that sufficient signal was present while preventing the A/D from saturating when the probe was inserted into the fuel. Total acquisition time for each spectrum was less than one second. All fuel spectra exhibited absorbance ranges within 0-0.8 absorbance units.

A third MPA NIR FT instrument (BrukerOptics, The Woodlands, TX, [www.brukeroptics.com/MPA](http://www.brukeroptics.com/MPA)) was used to collect the NIR spectra of fuels and chemicals. The spectral wavenumber range is 4000 - 12000  $\text{cm}^{-1}$ .

Each instrument was equipped with a transfectance dip probe with a fixed total pathlength of 1 cm. All spectra were collected at room temperature (22-25  $\text{C}^{\circ}$ ) by immersing a clean probe several inches into the sample and then swirling the probe to ensure no air bubbles were trapped. No effort was made to maintain a set fiber-optic geometry. All of the fuel and chemical spectra were preprocessed by Savitsky-Golay 7-point first derivative followed by a standard normal variate (SNV) transform to remove spectral off-set and baseline changes due to the variability of fiber-optic geometry in both Fuelex and MPA instruments. All chemicals were analytical grade (Aldrich).



## *BUILDING PLS MODELS*

NIR spectra of 420 JP-8 jet fuels were acquired using the primary instrument. The fuel samples were obtained from the Naval Research Laboratory in Washington D.C and from the Army TARDEC facility in Warren Michigan. The initial origin of the fuels ranged geographically on a sample by sample basis throughout the North American, European, African, Australian and Asian continents. The world-wide representation of fuel samples ensured variance in the training set used for PLS model building. ASTM reference values for each fuel set were provided; yet, most fuels lacked reference values for all properties listed in Table 7. ChemWave software (Bruker Optics, Woodlands, TX, USA) was used to construct and validate all PLS models of jet fuel properties from the spectra acquired on the primary instrument. Constructing a PLS model involves the following steps;

- 1) Selecting fuel spectra whose ASTM number for the property of interest are known, and combining these spectra with their corresponding property values into a training set,
- 2) Constructing a PLS model which correlates spectral intensity of the property of interest,
- 3) Selecting the number of principal components (latent variable) to use in the model,
- 4) Identifying and removing any outlier values,
- 5) Validating PLS model using leave-one-out cross validation,
- 6) Validating PLS model with an independent test set.

The number of latent variables was determined using the IND function and an F-test<sup>53</sup> with (n-k) degree of freedom and a probability level of 95%. For all property models shown in this chapter, five latent variables were used. Five latent variables were appropriate and showed the lowest SECV (standard error of cross-validation) with the minimum over-fitting of noise. Samples were considered outliers if both studentized residuals and leverage exceeded 98% probability of training set inclusion. Sample leverage is calculated using Eqn. 42:

$$\text{Sample leverage} = [\text{scores} (\mathbf{SCORES} \times \mathbf{SCORES}^T) \text{scores}^T]/n \quad (42)$$

where **SCORES** is a matrix of the entire training set where each row is a sample spectrum and numbers of columns give the latent variables used and **SCORES**<sup>T</sup> is the transposed matrix, and scores is a row vector containing the amount of each latent variable, and scores<sup>T</sup> is the transposed vector, and n is the number of samples in the training set.

In step 5, a particular model was validated by leave-one-out cross validation where one sample is removed from the training set, and then a PLS model was built using the remaining samples and finally this model is used to predict the property value of the excluded sample. This process is repeated until all the samples in the training set are excluded, each one once. The standard error of cross validation (SECV) is then calculated as in Eqn. 43:

$$\text{SECV} = [(\Sigma(\text{error})^2)/(n-k)]^{1/2} \quad (43)$$

where error is the difference between the true value (ASTM provided value) and the predicted value for all leave-one-out iterations, *n* is the total number of samples used to build the model, and *k* is the number of latent variable used to generate a model.

NIR spectra of an additional 180 JP-8 jet fuels (TARDEC, Warren, MI, USA) were acquired on both the primary and secondary instruments to form an independent test set (step 6). The independent test acquired on the primary instrument was used to validate the PLS models. The standard errors of prediction (SEP) are given in Table 8.

## VIRTUAL STANDARD CALIBRATION TRANSFER THEORY

To carry out the digital synthesis of the standards, 170 fuel spectra from the calibration set obtained on the Master were selected and used as target fuel spectra ( $f$ ) for the digital synthesis. The spectra were selected so that they reasonably represented the diversity of fuels used (i.e., inclusion of Jet-A, JP-8, JP-5) and also so that they reasonably spanned the range of PLS predicted property values. The virtual standards were first synthesized using the spectra of twelve neat solvents which were acquired on the master instrument. The solvents included: toluene, para-xylene, ortho-xylene, n-octane, meta-xylene, mineral oil, iso-octane, n-hexane, n-heptane, ethylbenzene, n-dodecane, and n-decane (Figure 55). All solvents were analytical grade (Aldrich) except mineral oil (Mallinkrodt white light paraffin oil). The virtual standards were constructed so that they approximated the fuels in the training set files (the target fuels). First, a multiple linear regression was set up:

$$f = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots \beta_i x_i + \varepsilon, \quad (44)$$

where  $x_i$  is a solvent spectrum, and  $\beta_0$  is an offset and  $\beta_i$  is the amount of solvent ( $x_i$ ) which is required to make each fuel target spectrum ( $f$ ). Also,  $\varepsilon$  is the spectral difference between the synthetic fuel and the target fuel. By setting up this relationship for every target fuel spectrum to be digitally synthesized, the above expression can be re-written as:

$$\mathbf{F} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (45)$$

where,  $\mathbf{F}$  is a matrix where each column is a fuel target spectrum,  $\mathbf{X}$  is a matrix where each column is a solvent spectrum, and  $\boldsymbol{\beta}$  is a matrix relating how much of each solvent is required to generate the target (number of rows equal to the number of solvents and number of columns equal to the number of targets). By setting the spectral difference matrix ( $\boldsymbol{\varepsilon}$ ) to zero, the matrix  $\boldsymbol{\beta}$  is solved as:

$$\boldsymbol{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{F}, \quad (46)$$

using singular value decomposition (*SVD*). For the current work, the *SVD* was performed at full rank (12), thus effectively the same as a multiple linear regression. The virtual standards were then generated as:

$$\mathbf{S} = \mathbf{X}\boldsymbol{\beta} \quad (47)$$

where  $\mathbf{S}$  is the virtual standard matrix with each column corresponding to a virtual standard and the number of virtual standards being equal to the number of fuel targets ( $f$ ). An example of a virtual standard overlaid with a target fuel is given in Figure 56. Once  $\boldsymbol{\beta}$  was determined, then solvent spectra acquired on the slave instruments, toluene, *p*-xylene, etc., were used to generate the virtual standards for the slave:

$$S_{slave} = X_{slave} \boldsymbol{\beta} \quad (48)$$

Where  $S_{slave}$  is the matrix of slave virtual standards,  $X_{slave}$  is the matrix of solvent spectra taken on a slave instrument and  $\boldsymbol{\beta}$  is from Eqn. (46). The virtual standards of the master instrument and the slave instrument were then predicted in a PLS model. Subsequently, a slope and bias correction for each slave PLS prediction was generated by regressing the master standard predictions against the slave standard predictions. For example, when correcting the master-generated PLS model for API gravity, the Master virtual standards were predicted using the model to yield  $P_{master}$  and the Slave virtual standards were predicted using the model to yield  $P_{slave}$ . The slope and bias correction for PLS future API gravity predictions on the slave was then given by:

$$P_{master} = m * P_{slave} + b \quad (49)$$

Where  $m$  is the value of the slope and  $b$  is the value of the bias.

All that is required for future transfers are the spectra of the solvents on the new instrument and the  $\boldsymbol{\beta}$  matrix. Once  $\boldsymbol{\beta}$  is determined, then solvent spectra acquired on the slave instruments are used to generate the virtual standards for each slave.

## RESULTS AND DISCUSSION

*VSSB CALIBRATION TRANSFER BETWEEN SIMILAR NIR-INSTRUMENTS*

Typical jet fuel NIR spectra are given in Figure 54. The spectra correspond to the jet fuels with the maximum, minimum, and median cetane indices in the training set. As can be seen, absorbances often occur with negative values due to the high refractive index of the fuel relative to that of air (used for reference).

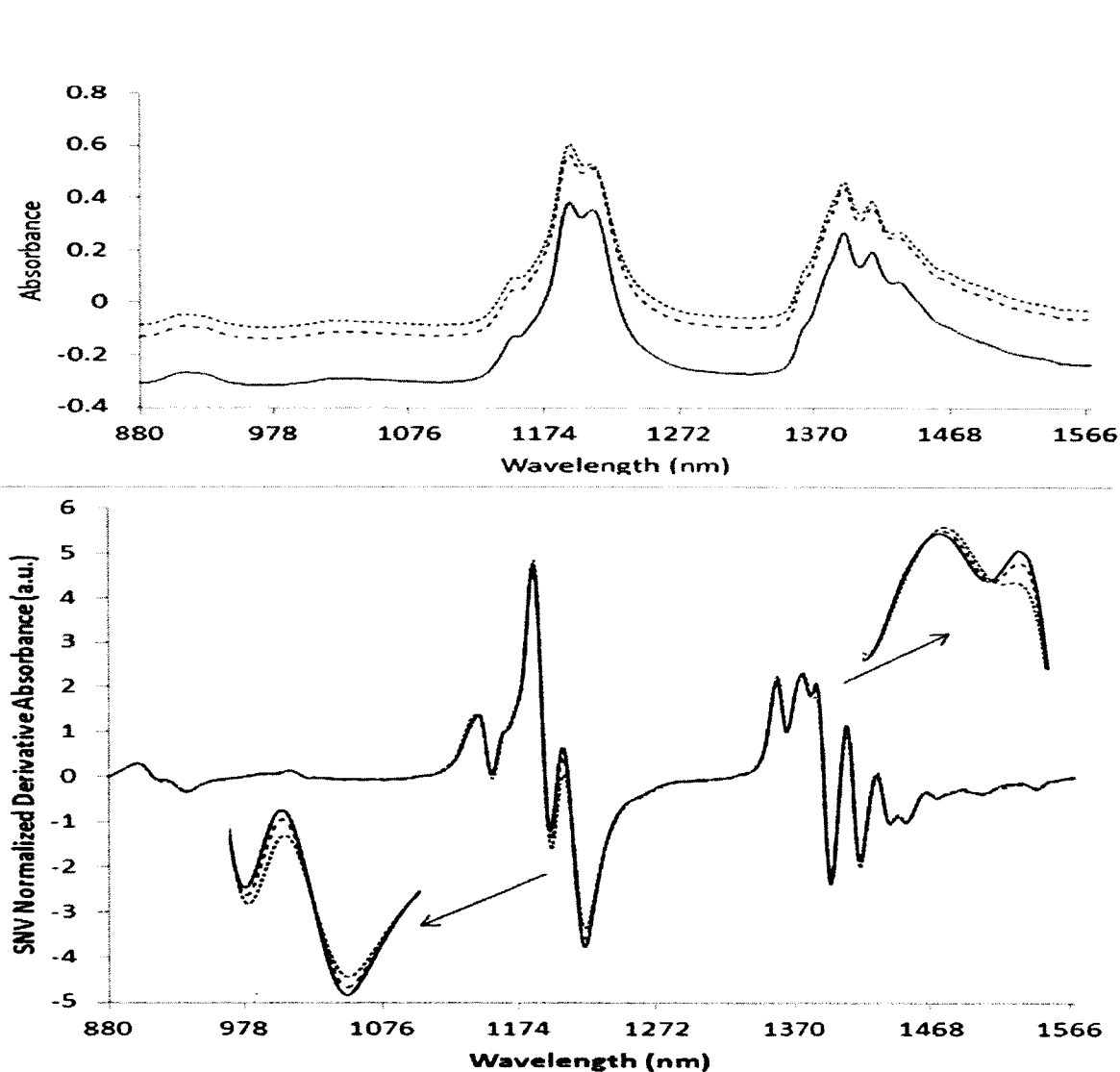


FIG. 54. (Top) Overlay of three spectra of jet fuel; spectrum of fuel with maximum cetane index value (solid line) overlaid with spectrum of median value (broken line) and minimum value (dashed line) over the wavelength range 880–1570 nm ( $11\,364\text{--}6369\text{ cm}^{-1}$ ). (Bottom) The same spectra after applying a 7-point Savitzky–Golay first derivative and SNV spectral transform.

Offsets in spectra were also observed to vary (to a lesser extent) with the degree of bending of the fiber-optic probe. In addition, across the range of fuels, baseline slope differences also occurred. For these reasons, spectra were transformed prior to analysis by application of a 7-point Savitzky–Golay first derivative followed by standard normal variate (SNV) normalization. This treatment (Fig. 54, bottom panel) eliminated baseline slope and offset differences. Within the spectral range investigated (880–1570 nm), the second overtone region accounts for the majority of the absorbance. Because jet fuels are hydrocarbon based, the C–H overtone and combination bands provide the bulk of the spectral features. There are three main C–H groups that yield vibrational spectra: aromatics (1100–1144 nm), methyl (1190–1195 nm and 1375–1400 nm), and methylene (1135–1215 nm and 1410–1440 nm). The modeled properties of the jet fuels are given in Table 7, along with the primary ASTM method for determining the property value, the range of values for that property according to the ASTM method, and the range of values determined for the jet fuels which were measured.

TABLE 7. Jet fuels properties, ASTM methods, units, ASTM range, training set range, and PER.

Property	ASTM method	Units	ASTM range		Training set range		RER
			Min	Max	Min	Max	
API gravity	D1298	°API	37	51	39.1	47.3	27
Aromatics	D1319	% volume	12.7	25	11.8	24.5	5
Cetane index	D976	index #	40	51	38.7	50.7	6
Density	D1298	g/mL	0.775	0.84	0.79	0.823	66
Distill. 10%	D86	°C	132	205	156	201	(12–15)
Distill. 20%	D86	°C	163	212	163	207	N/A
Distill. 50%	D86	°C	178	246	180	223	14
Distill. 90%	D86	°C	206	280	216	266	(13–15)
Flashpoint	D93	°C	38	74	38	68	(6–11)
Freeze point	D5972	°C	–63	–47	–62	–40	28
Hydrogen content	D3343	% mass	13.4	14.5	13.5	14.4	9
Saturates	D1319	% volume	69	89	69.7	89	5

<sup>a</sup> The range-error ratio (RER) is the range of values divided by the method reproducibility error. ASTM methods and ranges from ASTM International ([www.astm.org](http://www.astm.org)).

Although the ranges of the primary ASTM methods and the jet fuel training set are not identical, the training set does closely approximate the ranges used in the primary methods. Table 7 also includes the range error ratio (RER), which is simply the ASTM property range divided by the reproducibility error of the ASTM primary method. The RER for a given property gives a clue to how well that property can be modeled using a secondary correlation method since the correlation method cannot overcome the limitations of the primary method error unless multiple determinations are made on each sample. Typically a value of 10 or higher suggests a good modeling candidate because the range being sampled is much larger than the error in the measurement. The fuel properties being modeled are given in Table 8 along with the ASTM primary test method, the ASTM reproducibility, the standard error using leave-one-out validation, and the correlation coefficients for the predicted vs. actual results. The ASTM reproducibility reflects the maximum expected inter-lab measurement error for the primary ASTM test method at a 95% confidence level. As can be seen from the correlation coefficients and the standard errors, in all cases, the NIR PLS models built on the master instrument yielded reasonable estimates of the primary methods. For all PLS models of Jet fuel 5 principal components were used. Using a few principal components lead to throwing important spectral information, while using too many leads to over fitting the noise. The number of the principal components was determined by the F-test. The F-test is a statistical test that is always used to compare the least squares models and it is the ratio of two scaled sums of squares of two different PLS models.

TABLE 8. Summary of the jet fuel PLS models and properties.

Jet fuel property	Number of samples	ASTM Reproducibility	SECV <sub>M</sub>	Mean	Std. Dev.	r <sup>2</sup>
API Gravity	291	0.3	0.265	43.94	3.38	0.98
Aromatics	156	3	0.47	17.89	2.00	0.94
Cetane Index	138	2	0.309	45.83	1.72	0.97
Density	270	0.0005	0.001	0.80	0.02	0.96
Distillation 10% Recovered	372	3.43-4.42	2.65	176.97	19.61	0.92
Distillation 20% Recovered	397	N/A	2.57	185.58	11.09	0.92
Distillation 50% Recovered	381	2.97	2.09	204.09	11.24	0.92
Distillation 90% Recovered	340	3.24-3.99	1.99	240.42	5.92	0.77
Flash Point	314	2.7-4.83	2.36	50.92	6.69	0.86
Freeze Point	408	1.81	1.43	-49.91	3.11	0.52
Hydrogen Content	114	0.04	0.019	13.82	0.13	0.97
Saturates	120		1.08	78.67	2.71	0.81
Viscosity@ -20 °C	135	0.0800	0.07	4.29	0.23	0.90

\* <http://www.astm.org/>

SECV<sub>M</sub>: Standard error of validation of master property model using leave-one-out validation; Mean: mean of the training set values; Std.Dev.: standard deviation of the training set values; r<sup>2</sup>: correlation coefficient for predicted vs. actual values in leave-one-out validation plots. See Table 7 for units of measurement.

The results for the prediction of the independent test sets are given in Table 9. For the master instrument, the standard errors of prediction were higher for all properties except freeze point (see Tables 8 and 9). This is to be expected for such large training sets since the exclusion of a single spectrum from the models during a leave-one-out procedure has little leverage relative to the large number of samples remaining. Since the primary tests for each fuel were conducted at the geographically diverse points of origin of the fuels, the ASTM reproducibility represents a reasonable limit on the accuracy of the blind model predictions.



TABLE 9. Standard errors of prediction of properties in the test set samples relative to ASTM values for uncorrected, VSSB corrected, and PDS corrected models.

Jet fuel property	# of samples	Master SEP	Slave			$r^2$
			SEP <sub>U</sub>	SEP <sub>C</sub>	PDS	
API Gravity	125	0.35	1.33	0.28	0.33	0.93
Aromatics	107	1.41	0.86	1.37	1.36	0.96
Cetane Index	130	0.86	1.23	0.83	0.79	0.98
Density	124	0.0018	0.0059	0.0020	0.0016	0.92
Distil. 10%	128	3.34	13.11	3.47	3.56	0.74
Distil. 20%	127	3.43	16.1	3.17	3.41	0.72
Distil. 50%	127	2.81	10.81	2.62	2.66	0.71
Distil. 90%	126	2.86	2.81	2.99	3.04	0.97
Flash Point	103	2.74	8.82	2.82	2.81	0.65
Freeze Point	115	1.09	2.73	1.25	1.17	0.89
%Hydrogen	100	0.061	0.029	0.061	0.059	0.98
%Saturates	109	1.77	2.72	2.41	1.74	0.98
Viscosity @ -20 °C	112	0.105	0.44	0.11	0.11	0.99

Considering the global nature of the data set along with the test method reproducibility, the master prediction errors for the independent test set were consistent with expectations. Also shown in Table 9, are the prediction errors for the slave instruments using the master generated PLS models without any correction (SEP<sub>U</sub>). The slave standard errors were more than twice that of the master for API gravity, density, 10% distillation, 20% distillation, 50% distillation, flashpoint, freeze point, and viscosity.

In order to perform virtual standards calibration transfer, NIR spectra of the twelve standards was acquired on both master and slave instruments. Figure 55 displays the solvent spectra (1<sup>st</sup> derivative SNV) acquired on the master instrument overlaid with the solvent spectra acquired on the slave instrument.

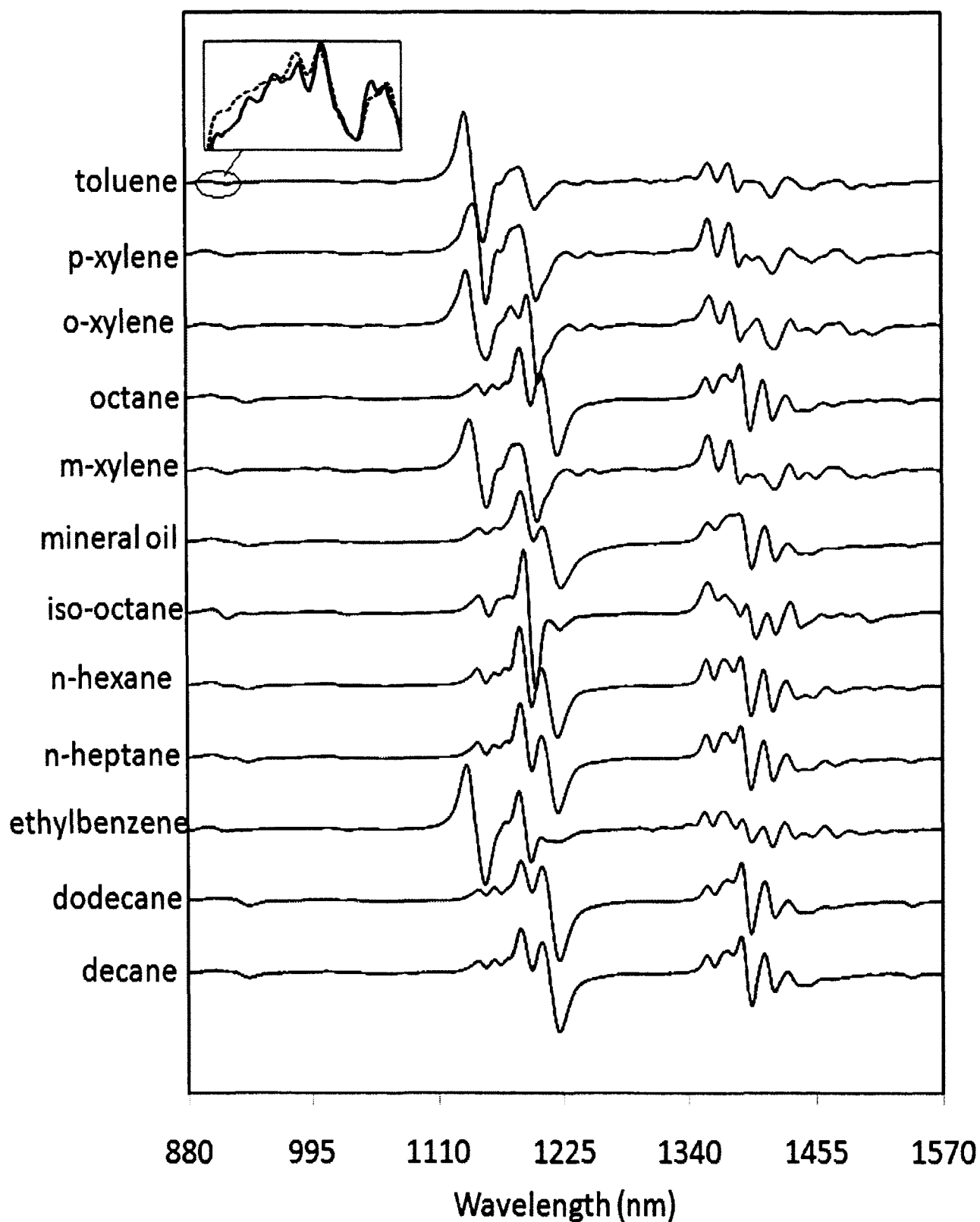


FIG. 55. NIR absorbance of 1<sup>st</sup> derivative SNV spectra of solvents acquired on the master instrument (Solid line) overlaid with the slave standards (broken line). The spectra cover the range of 6,269 - 11,364  $\text{cm}^{-1}$ . Seven point Savitsky-Golay 1<sup>st</sup> derivative and SNV spectral transforms have been applied. Spectra offset by 6 a. u.

Although the instruments yielded similar spectra, as observed in the inset, there were slight differences between the spectra. The spectra of the twelve standards were used with the spectra of the training set to generate virtual standards for both master and slave instruments according to the method described in the experimental section. Figure 56 displays the transformed spectra of one of the target jet fuels (taken from a training set) and the virtual fuel constructed from the master solvent spectra. Although not identical to the target fuel, the virtual fuel provided essentially the same absorbance in the same spectral regions as the target fuel. Similar results were obtained for all of the virtual fuels, thus ensuring a spectral variance for the virtual standards which was similar to that of the training sets.

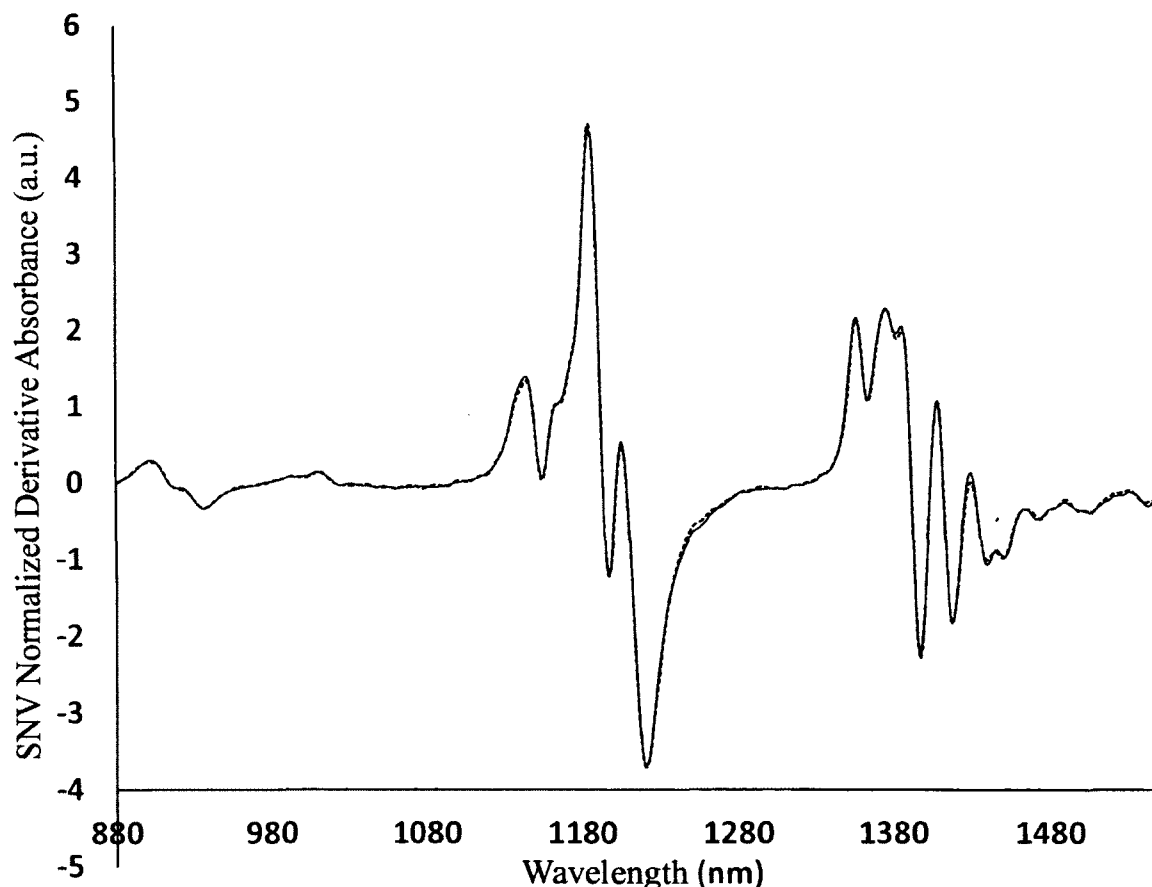


FIG. 56. NIR absorbance spectrum of a digitally synthesized virtual standard (dashed line) overlaid with its target fuel spectrum (solid line). Seven point Savitsky-Golay 1<sup>st</sup> derivative and SNV spectral transforms have been applied.

Table 9 also shows the standard error of prediction for the slave instruments after correction ( $SEP_c$ ) using the virtual standards and a slope and bias correction (VSSB correction). The VSSB correction resulted in a dramatic reduction in the slave standard error. In all instances where  $SEP_u$  was high, the error decreased to values that were less than, or similar to, that of the master SEP. However, even when the VSSB resulted in slightly higher SEP values, the slave  $SEP_c$  values were still lower than the master SEP. The only exceptions to this was for 90% distillation ( $SEP_c=2.99$ ,  $SEP_u=2.81$ ,  $SEP_{master}=2.86$ )

To further validate the performance of the VSSB correction, PDS correction, a very common used spectral correction method, was performed using the same training set used to generate the virtual standards. The standard error of prediction after applying PDS correction is shown in table 9. VSSB correction performs as well as PDS for most of the jet fuel properties; %aromatics, 10%, 50%, and 90% distillation, flash point, and viscosity @ - 20 C°. While VSSB correction shows better performance for API and 20% distillation.

Figure 57 graphically depicts the effects of the VSSB correction for the slave instruments for the API Gravity PLS model. Plot A shows the effect of the correction relative to the master predictions (solid line) and plot B shows the effect of the correction relative to the reported API gravity values (open circles) for each sample in the test set.

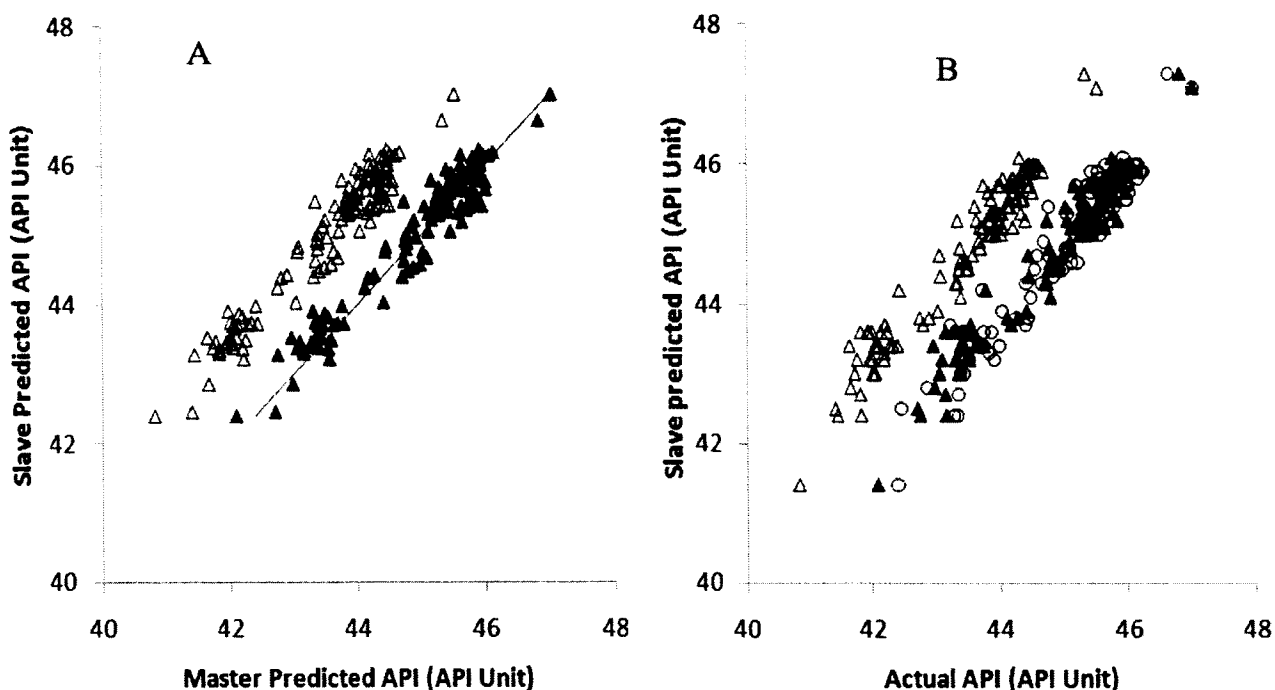


FIG. 57. API gravity predictions of the test set fuels acquired on the slave instrument for both uncorrected (open triangles) and VSSB corrected (filled triangles) data: A) slave predictions are plotted relative to the master instrument predictions and the solid line is the master predicted vs master predicted (ie., ideal line); B) predictions are plotted relative to the actual (ASTM D1298) values and open circles are the master predicted values.

Figures 58-60 show plots for the VSSB corrected predictions for slave (solid triangles) and the uncorrected predictions for slave (open triangles) relative to the master predictions for all of the properties modeled (with the exception of API gravity which is given in Figure 57).

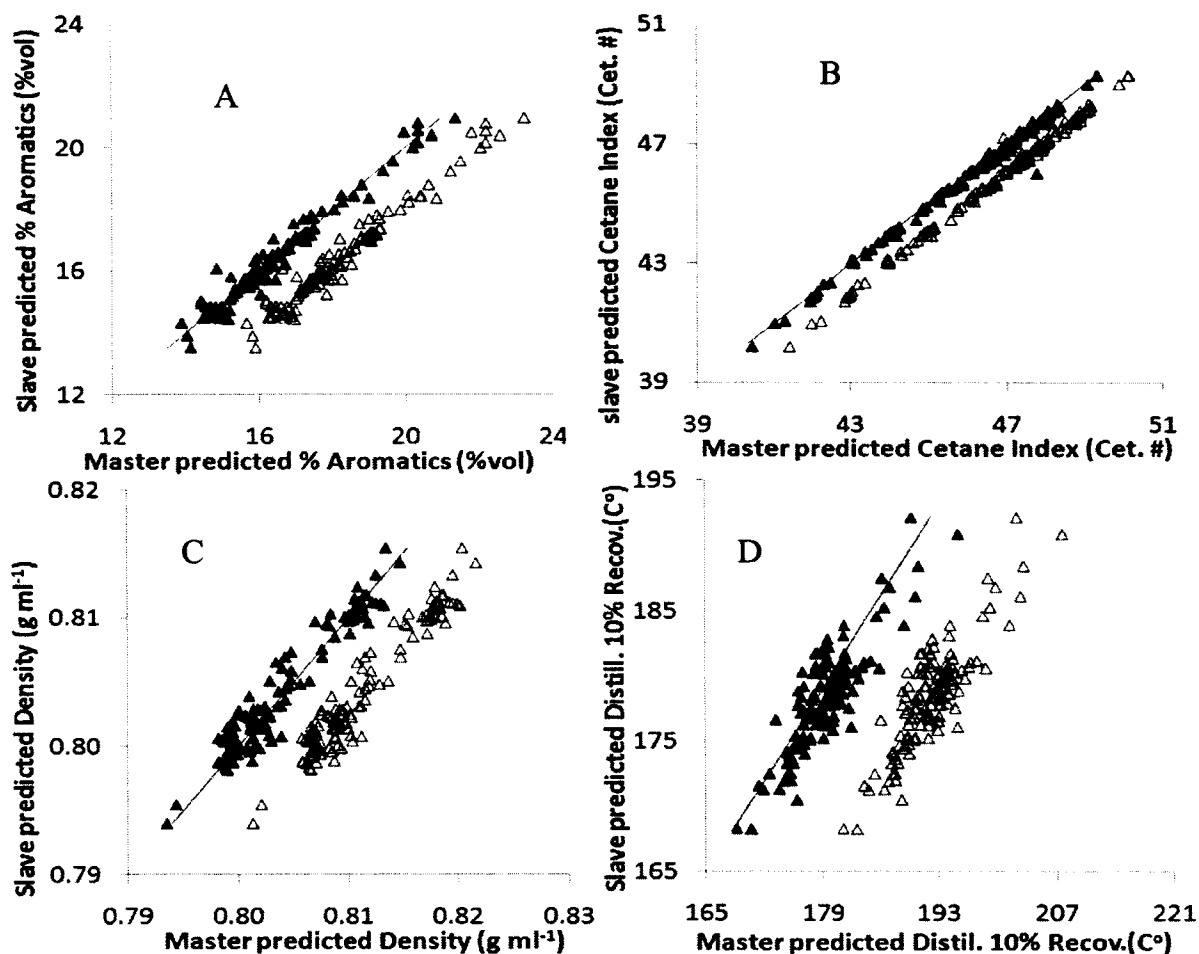


FIG. 58. Fuel property predictions of the test set fuels acquired on the slave instrument for both uncorrected (open triangles) and VSSB corrected (filled triangles) data. A) Aromatics; B) Cetane Index; C) Density; and D) Distillation Temperature at 10% volume recovered. All predictions are plotted relative to the master instrument predictions and the solid line is master predicted vs master predicted (ideal line).

This series of plots is perhaps the most telling indication of the utility of the VSSB method. For all of the property models, the VSSB correction resulted in prediction of values which more closely matched those of the master predictions. This puts the SEPC values in Table 9 into perspective. For example the %aromatics VSSB corrected

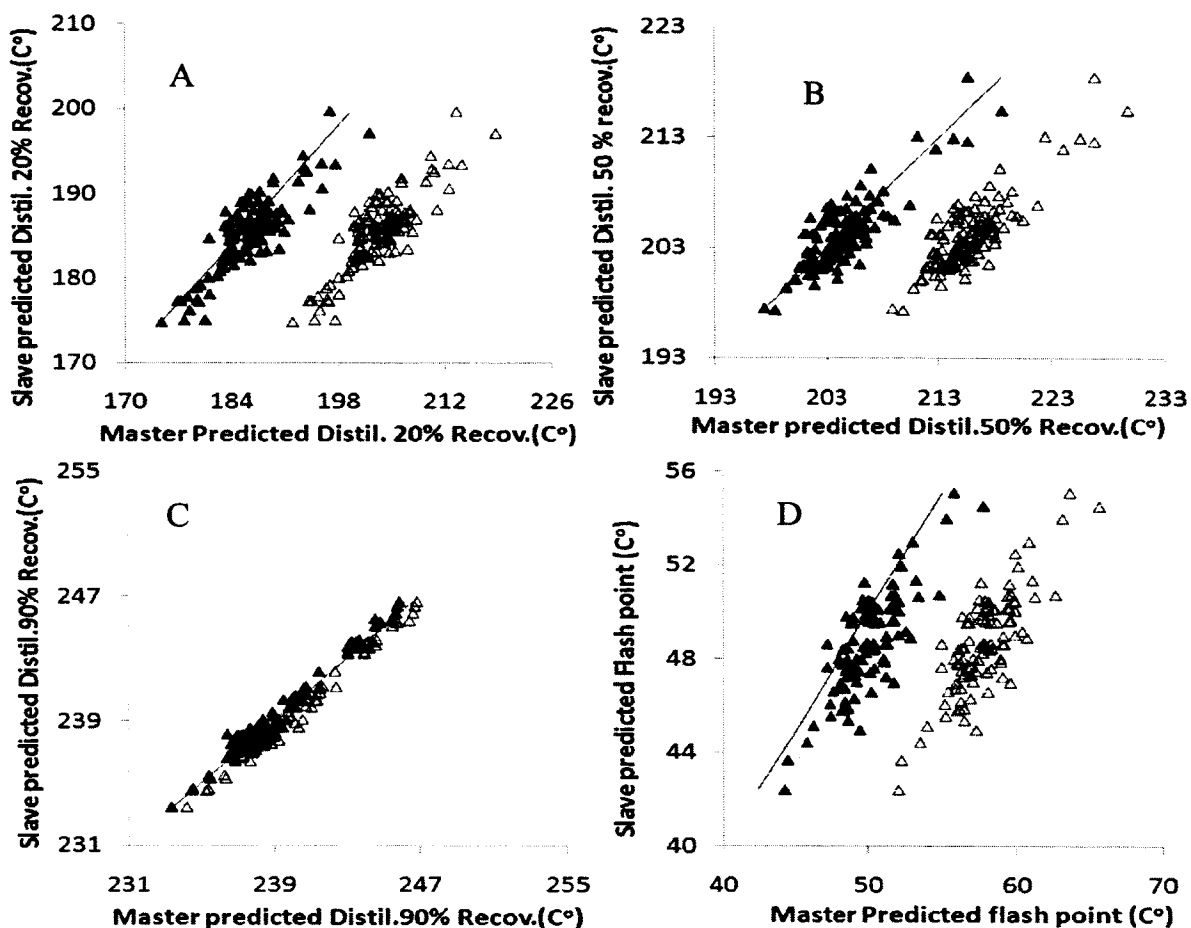


FIG. 59. Fuel property predictions of the test set fuels acquired on the slave instrument for both uncorrected (open triangles) and VSSB corrected (filled triangles) data. A) Distillation Temperature at 20% volume recovered; B) Distillation Temperature at 50% volume recovered; C) Distillation Temperature at 90% volume recovered; and D) Flashpoint. All predictions are plotted relative to the master instrument predictions and the solid line is master predicted vs master predicted (ideal line).

predictions for the slave showed an increase in error relative to the uncorrected predictions (Table 9), however, the plot in Figure 58a clearly shows that the VSSB correction resulted in predictions which much more closely matched to the master instrument predictions. The same was true for the 90% distillation and %hydrogen predictions for slave (Figures 59c and 60b, respectively) where the VSSB corrections were closer to the master predictions relative to the uncorrected predictions even though the standard error of prediction relative to the documented values increased with the VSSB correction.

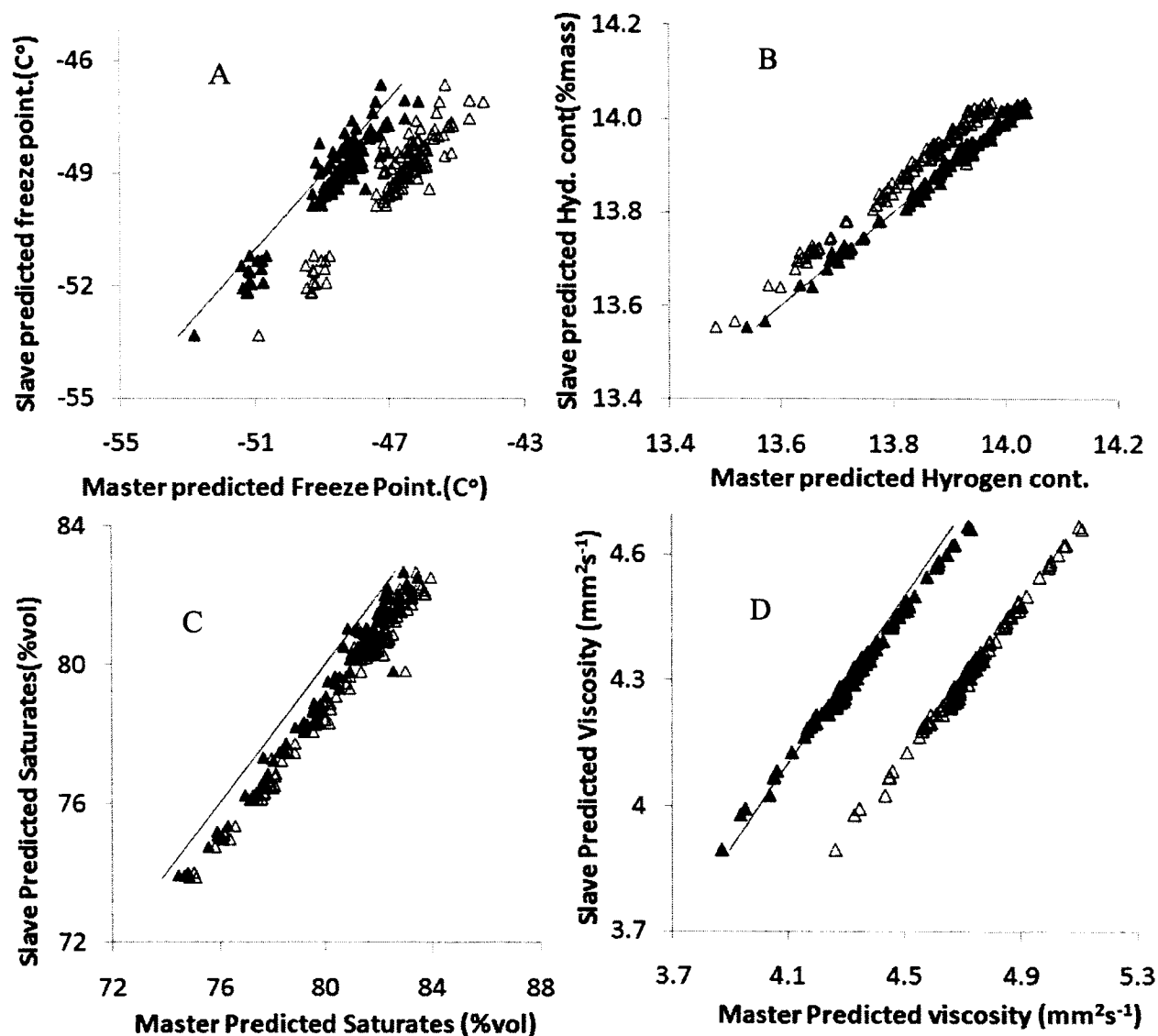


FIG. 60. Fuel property predictions of the test set fuels acquired on the slave instrument for both uncorrected (open triangles) and VSSB corrected (filled triangles) data. A) Freeze Point; B) % Hydrogen Content; C) % Saturates; and D) Viscosity at -20°C. All predictions are plotted relative to the master instrument predictions and the solid line is master predicted vs master predicted (ideal line)



*SVSSB CALIBRATION TRANSFER BETWEEN DIFFERENT NIR-INSTRUMENTS*

In this section the utility of the virtual standards slope and bias correction will be tested between two different NIR-instruments one of them is a dispersive handheld (FUElex) and the other is an FT-NIR desktop (MPA) instruments. NIR-spectra of 500 jet fuels (TARDEC, Warren, MI, USA) were acquired on both the FUElex NIR dispersive and the MPA FT-NIR instruments to form an independent test set. A raw JP-8 fuel spectra acquired on both the FUElex and the MPA are shown in Figure 61, A shows a MPA jet fuel raw spectrum and B shows a FUElex jet fuel raw spectrum. As shown in Figure 61, the two spectra have different ranges and resolution. The MPA instrument was used as the primary instrument and FUElex is the secondary instrument.

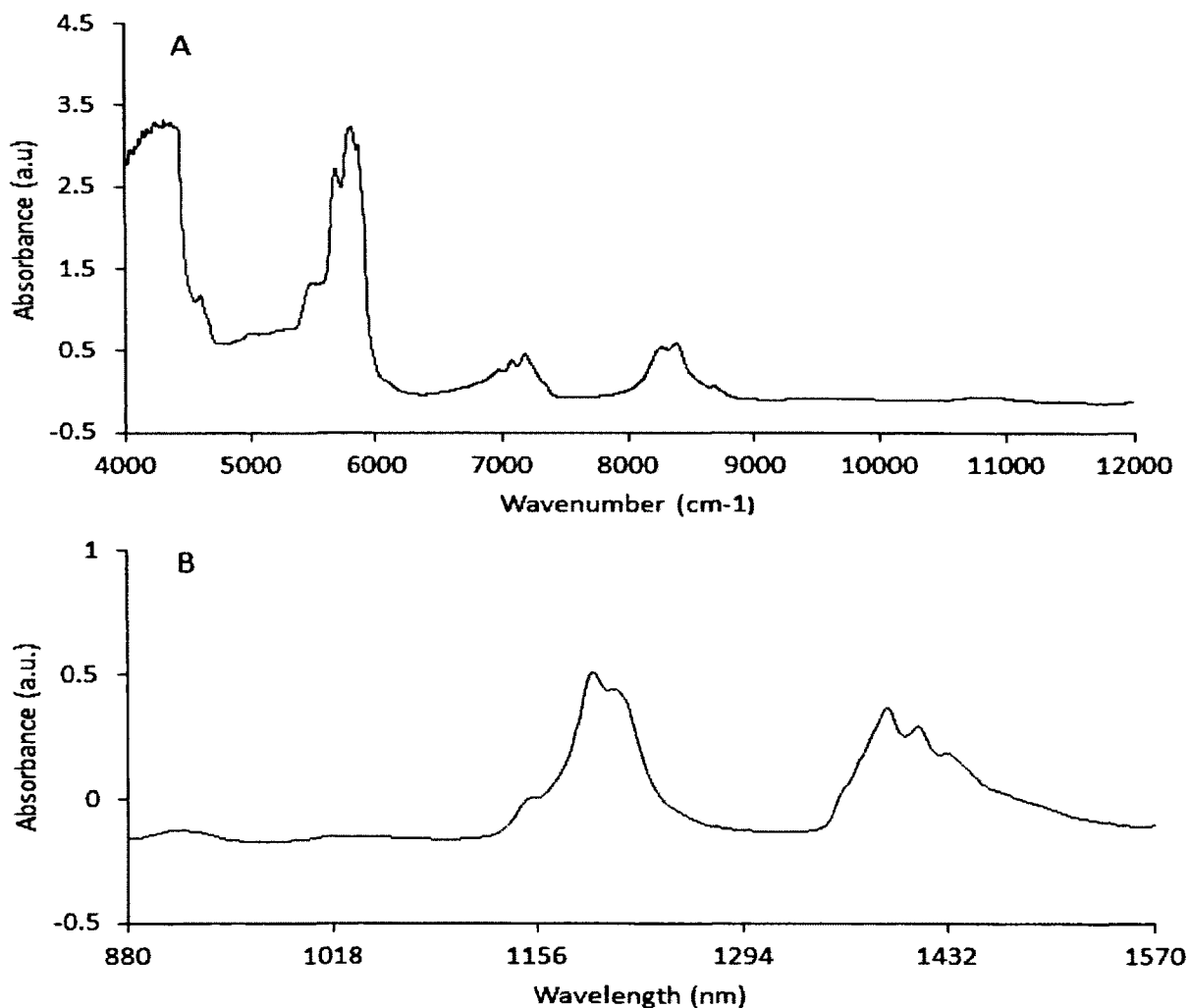


FIG. 61. (A) gives the NIR absorbance spectrum of a JP-8 fuel acquired on the MPA FT-NIR instrument. (B) gives the NIR absorbance spectrum of the JP-8 fuel acquired on the FUElex, a dispersive NIR-instrument.

The MPA spectra were truncated to span the same wavelength range as the FUELex (880-1570nm). The spectra were then pretreated by 7-points Savitsky-Golay 1<sup>st</sup> derivative followed by SNV. A pretreated FUELex spectrum is overlaid with a pretreated MPA spectrum in Figure 62A, the solid line is the FUELex spectrum and the dashed line is the MPA pretreated spectrum.

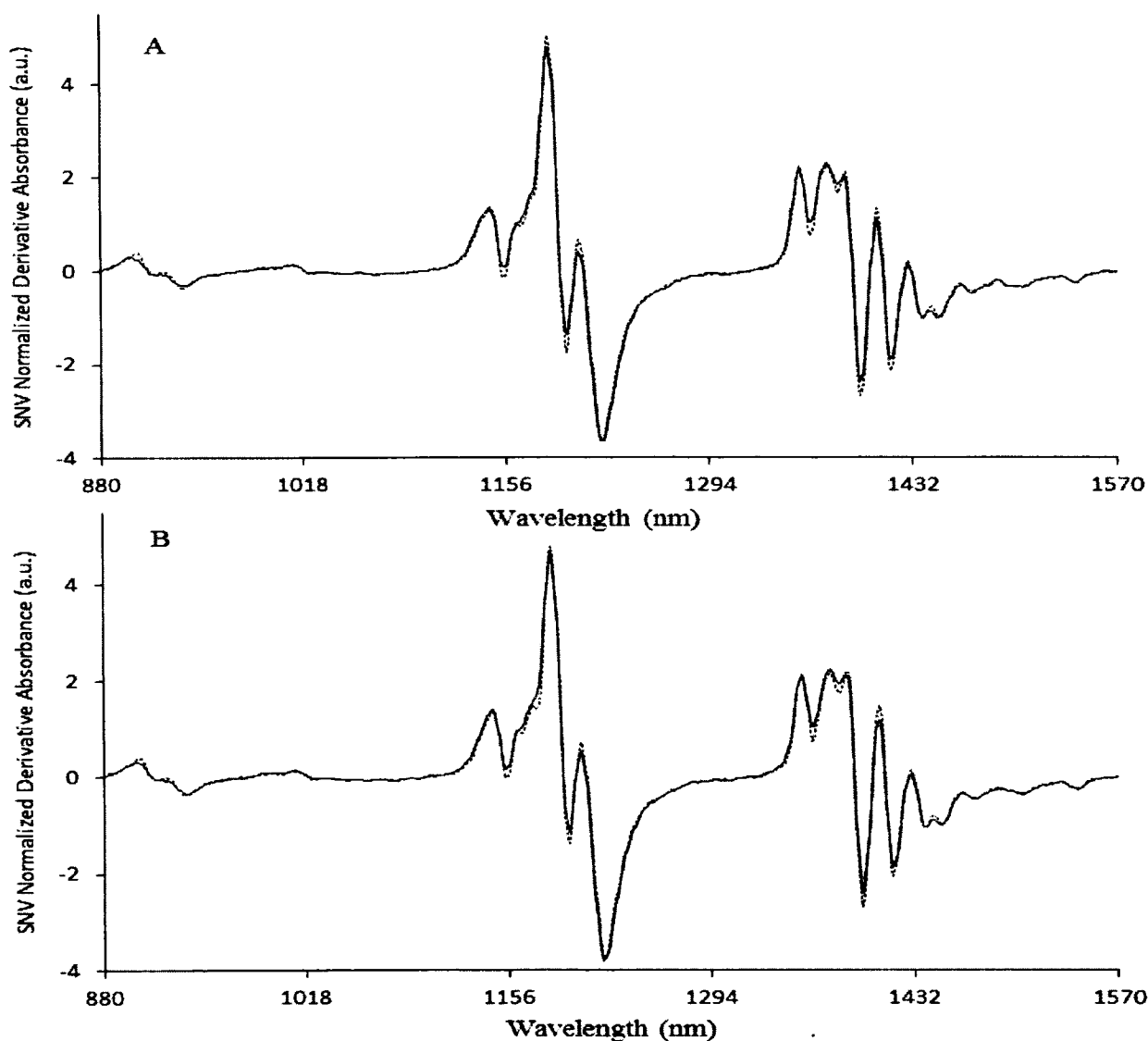


FIG. 62. NIR absorbance spectra of a digitally synthesized MPA secondary instrument virtual standard (dotted line) overlaid with a digitally synthesized FUELex virtual standard (solid line) are given in (B) and their target fuel spectra of MPA (dotted line) and FUELex (solid line) are given in (A). Seven point Savitsky-Golay 1<sup>st</sup> derivative and SNV spectral transforms have been applied.

In order to transfer the PLS models of jet fuel properties between the MPA master and the FUELex slave, the NIR-spectra of the twelve standards were acquired on both the master and slave instruments. The NIR-spectra of the standards acquired on the MPA are then truncated and pretreated by 7-points Savitsky-Golay 1<sup>st</sup> derivative and SNV. The standard spectra along with the training set acquired on the MPA instrument (used to build the PLS jet fuel properties models) were used to generate virtual standards for both the master (MPA) and the slave (FUELex) instruments according to the method described in the experimental section. A virtual standard spectrum of the MPA is overlaid with a virtual standard spectrum of the FUELex in Figure 62B. The two virtual standards spectra exactly reflect the difference in resolution and intensity of their targets in Figure 62A. Table 10 shows the percentage of each standard required to generate virtual standards of five randomly picked spectra of JP-8, JP-5, and JPA. As in table 10, the percentage of mineral oils (not pure solvents and its composition depends on the manufacturer) is always higher than 50%. This makes it hard to replace the mineral oils standard and puts a limitation of this method since the composition of the mineral oils is not always the same.

TABLE 10. The percentages of the twelve solvents used to digitally synthesize five target spectra.

	Fuel1	Fuel2	Fuel3	Fuel4	Fuel5
Decane	46.917	59.862	57.795	37.986	59.464
Dodecane	-44.216	-51.725	-49.465	-48.692	-53.018
Ebenzene	4.411	5.132	6.078	5.848	5.948
Heptane	6.025	7.245	14.07	2.416	12.029
Hexane	12.927	10.598	7.643	12.93	7.354
Isooctane	6.447	8.569	8.275	7.129	8.615
Mineral oils	64.291	58.962	60.564	58.989	60.529
Mxylene	6.709	7.084	8.967	11.593	9.443
Octane	5.896	4.262	-1.29	25.817	2.403
Oxylene	2.662	5.27	4.291	0.293	4.265
Pxylene	1.734	1.727	0.728	0.422	0.617
Toluene	-5.803	-6.914	-8.501	-5.697	-8.237

The five fuels were selected randomly from the training set of jet fuels (Jet-A, JP-5 and JP-8)

The results for the prediction of the independent test sets are given in Table 11. For the master instrument, the standard errors of prediction are within or close to the ASTM reproducibility. This is to be expected for such large training sets since the exclusion of a single spectrum from the models during a leave-one-out procedure has little leverage relative to the large number of samples remaining. Since the primary tests for each fuel were conducted at the geographically diverse points of origin of the fuels, the ASTM reproducibility represents a reasonable limit on the accuracy of the blind model predictions. Considering the global nature of the data set along with the test method reproducibility, the master prediction errors for the independent test set were consistent with expectations. Also shown in Table 11, are the prediction errors for the slave instruments using the master generated PLS models without any correction ( $SEP_u$ ). The slave standard errors were more than one order of magnitude higher than that of the master for API gravity, density, 10% distillation, 20% distillation, 50% distillation, flash point, and freeze point.

TABLE 11. The standard errors of prediction of properties in the test set samples relative to the master prediction for uncorrected, VSSB corrected, and PDS corrected models.

Jet fuel property	Samples Number	MPA Master		FuElex slave			
		SEP	$r^2$	$SEP_u$	$SEP_c$	PDS	$r^2$
API Gravity	287	0.43	0.99	6.36	0.33	0.37	0.98
Aromatics	171	0.51	0.92	3.33	0.26	0.56	0.98
Cetane Index	153	0.356	0.97	7.81	0.25	0.32	0.99
Density	287	0.0018	0.97	0.0034	0.0017	0.0017	0.94
Distill. 10%	427	3.32	0.92	62.17	4.11	2.29	0.86
Distil.20%	397	2.57	0.92	55.43	3.71	2.22	0.79
Distill.50%	425	3.33	0.88	46.26	7.35	2.21	0.71
Distill. 90%	422	3.21	0.49	5.16	2.1	0.96	0.81
Flash Point	328	3.19	0.78	24.35	3.56	2.11	0.71
Freeze Point	435	1.79	0.68	6.65	2.64	1.07	0.68
Hydrogen Cont.	115	0.022	0.98	0.41	0.022	0.011	0.99
Saturates	141	0.99	0.82	0.55	0.44	0.53	0.97
Visc@ -20 °C	148	0.088	0.81	1.35	0.046	0.058	0.95

Table 11 also shows the standard errors of prediction for the secondary instrument after the correction using the virtual standards and a slope bias correction method ( $SEP_c$ ). The VSSB correction results in a dramatic reduction in standard error of prediction. In all instances  $SE_{Pu}$  decreases to values that are similar or closed to that of the primary instrument SEP, as shown in table 11.

For further comparison PDS was carried out using the same training set target used to generate the virtual standards. The standard error of prediction of PDS correction is shown in table 11. PDS correction results in reduction of the prediction error of secondary instrument to values similar or close to the VSSB correction. the VSSB correction shows better correction than PDS for API gravity, cetane index, saturates, and viscosity, while the PDS correction is a little better or similar to VSSB correction for 10%, 20%, 50%, 90% distillation recovery, flash point, freeze point and hydrogen content (table 11). Although, the PDS error is closed to the VSSB error, there is a small slope after applying PDS correction, as can be seen in case of API and cetane index (Figure 63D)

The effect of the SVSSB as well as the PDS correction is graphically depicted in Figures 63 to 66 for all properties of jet fuel of the secondary instrument. In all Figures, A, B, and C display the VSSB correction, while D, E, and F display the PDS correction. Where open squares represent the uncorrected secondary instrument prediction jet fuel properties relative to primary instrument prediction (solid line). Solid squares show secondary instrument predicted jet fuel properties after correction using VSSB (A, B, C) and PDS (D, E, and F). For all of the property models, the VSSB correction results in predicted values which more closely match that of the primary instrument predictions values for all jet fuels properties being modeled, as shown on the plots (Figures 63-66).

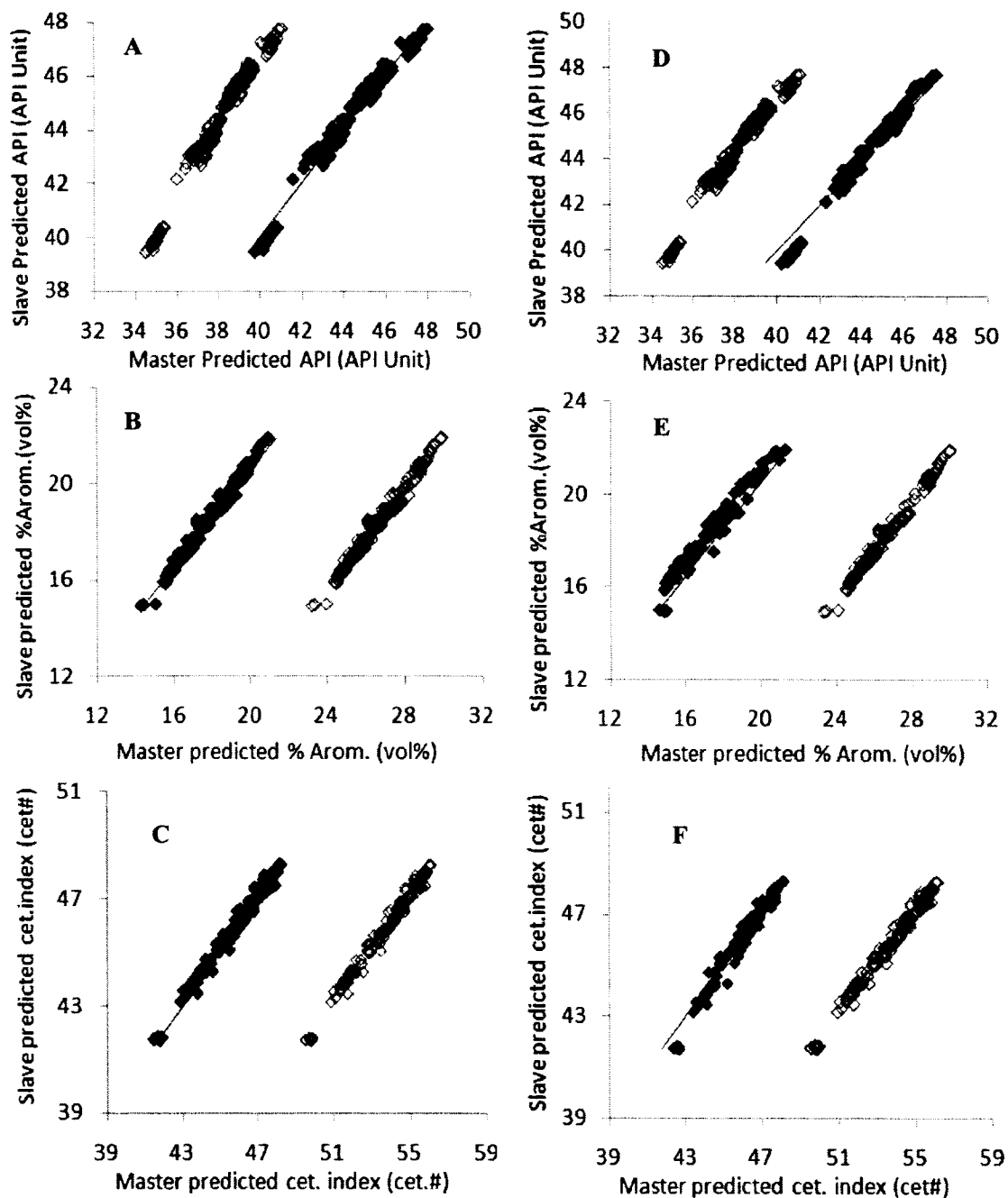


FIG. 63. Fuel property predictions of the test set fuels acquired on the secondary instrument for uncorrected (open squares), VSSB corrected and PDS corrected (filled squares) data. A, B, and C show the correction of API, %aromatic, and cetane index using VSSB correction respectively. D, E, and F show the correction of API, %aromatic, and cetane index using PDS correction respectively. All predictions are plotted relative to the primary instrument predictions and the solid line is primary instrument predicted vs primary predicted (ideal line).

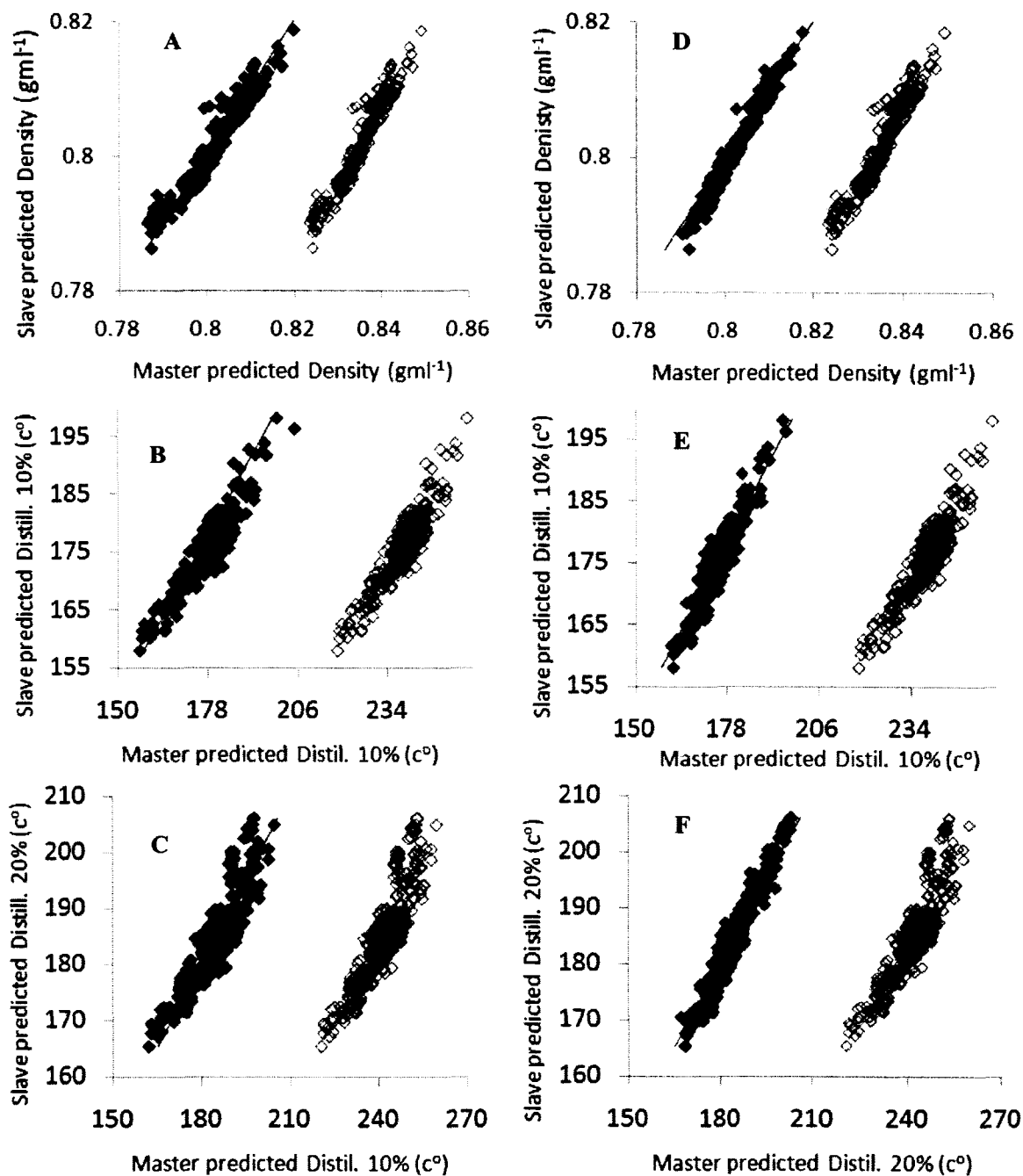


FIG. 64. Fuel property predictions of the test set fuels acquired on the secondary instrument for uncorrected (open squares), VSSB corrected and PDS corrected (filled squares) data. A, B, and C show the correction of density, 10% distill, and 20% distill using VSSB correction respectively. D, E, and F show the correction of density, 10% distill, and 20% distill using PDS correction respectively. All predictions are plotted relative to the primary instrument predictions and the solid line is primary instrument predicted vs primary predicted (ideal line).

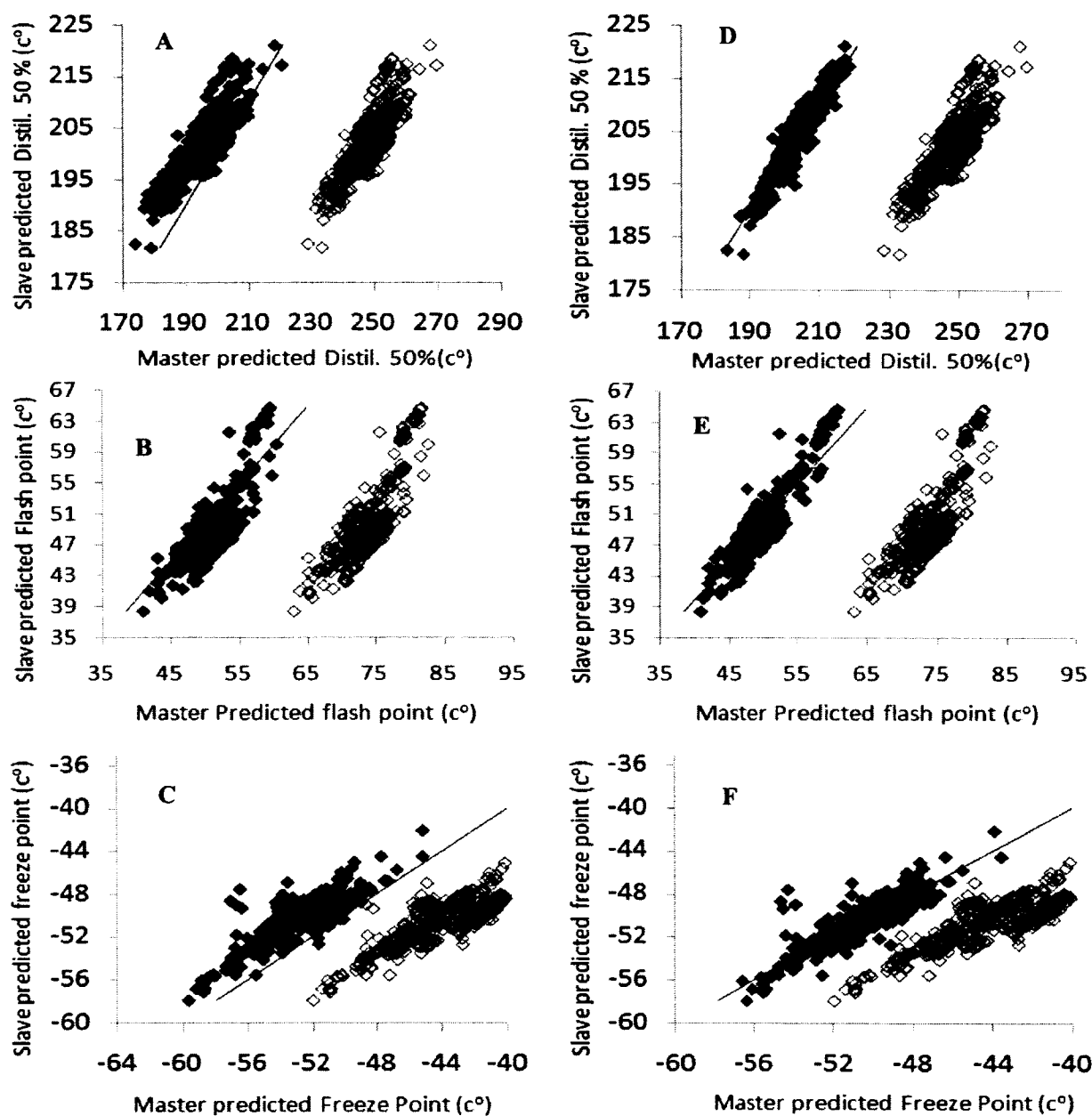


FIG. 65. Fuel property predictions of test set fuels acquired on the secondary instrument for uncorrected (open squares), VSSB corrected and PDS corrected (filled squares) data. A, B, and C show the correction of 50% distill, flash point, and freeze point using VSSB correction respectively. D, E, and F show the correction of 50% distill, flash point, and freeze point using PDS correction respectively. All predictions are plotted relative to the primary instrument predictions and the solid line is primary instrument predicted vs primary predicted (ideal line).



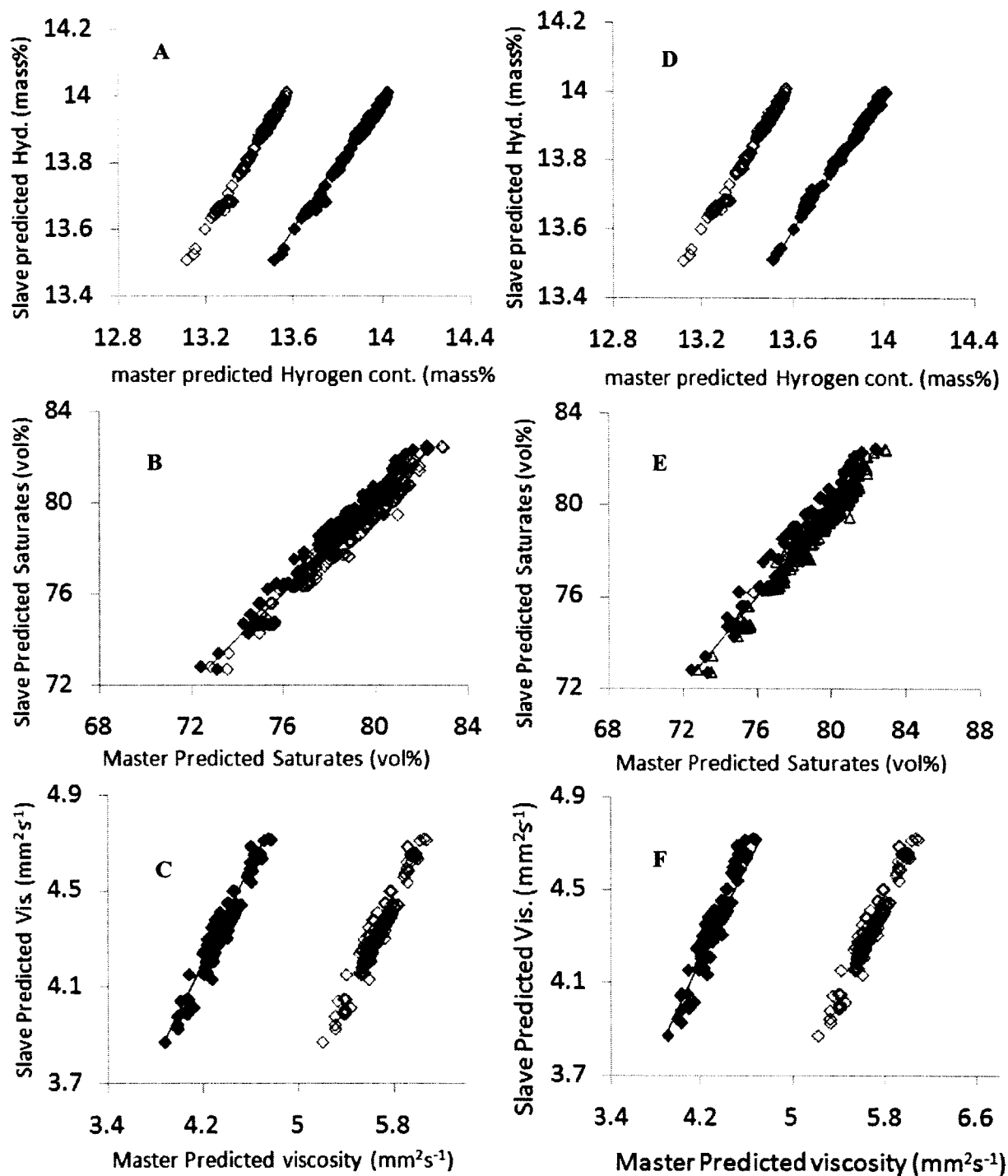


FIG. 66. Fuel property predictions of the test set fuels acquired on the secondary instrument for uncorrected (open squares), VSSB corrected and PDS corrected (filled squares) data. A, B, and C show the correction of hydrogen content, saturates, and viscosity @  $-20\text{ C}^\circ$  using VSSB correction respectively. D, E, and F show the correction of hydrogen content, saturates, and viscosity @  $-20\text{ C}^\circ$  using PDS correction respectively. All predictions are plotted relative to the primary instrument predictions and the solid line is primary instrument predicted vs primary predicted (ideal line).

In summary, the use of the VSSB correction performed as well as the conventional PDS correction when transferring the jet fuel property models listed in Table 7. However, the VSSB method does not require the use of jet fuels to transfer the models from the master instrument to the slave instrument. The use of a virtual standard, spectra from solvents, to transfer the PLS models eliminated the need to maintain fuel standards which are prone to compositional changes. Eventually, it eliminates the need to maintain a master instrument once the VSSB method is established for the models. The use of pure chemical components to transfer the models using the VSSB method is an ideal case. In this chapter, although neat solvents were used, one of the solvents (mineral oil) is not a pure chemical. This brings up two points: 1) can the same mineral oil be obtained at a later date and 2) is the mineral oil immune to compositional change over time. Both of these points are only valid for pure chemicals. Although the use of mineral oil in this chapter was seen as a useful for proof of concept of the VSSB method, a solution of this problem will be discussed in details in chapter V where only pure chemicals will be used for virtual standards calibration.

## CHAPTER V

### CALIBRATION TRANSFER OF PLS MODELS USING SEGMENTED VIRTUAL STANDARDS

#### INTRODUCTION

Multivariate models have gained widespread acceptance as a secondary analytical method in fuel analysis, pharmaceutical industry, and food industry.<sup>86, 87, 30-33, 35</sup> These methods are typically based on a partial least squares (PLS) regression correlating a property of interest with near infrared (NIR) or Raman spectra. Although PLS models can be built in a laboratory environment using a spectrometer and a set of calibration samples, development and validation of a robust multivariate model often involves hundreds of samples and the use of elaborate instrumentation to determine reference values. This leads to a large investment in time and money. Therefore, it is beneficial to be able to use a validated PLS model built on a primary instrument on additional instruments or on the same instrument as changes occur to the instrument over time. However, when measuring a sample on another instrument or the same instrument after some time, changes in the instrumental response lead to erroneous predictions by PLS models. One way to solve this problem is to re-measure every sample and reconstruct the PLS model. Yet, this is an impractical solution since it is time consuming, expensive, and requires a significant effort. For this reason, there has been extensive work to develop multivariate models calibration transfer methods between instruments. Where the term “Calibration” here means making an old instrument or a newly-built instrument give the same reading as the instrument in which the PLS model was built. These calibration transfer methods can be generally characterized as either spectral transfer or model transfer methods<sup>88, 89</sup>.

In the case of spectral transfer, the response of a sample measured on the secondary instrument is corrected to the response obtained on the primary instrument so that the spectra of standards collected on the master and secondary instruments match each other. Spectral transfer methods include, but are not limited to, direct standardization<sup>53</sup>,

piecewise direct standardization<sup>55</sup>, and orthogonal signal correction<sup>62</sup>. On the other hand, in the case of model transfer, the PLS model built and validated on the primary instrument is adjusted in a way that makes the predicted values for the secondary instrument match the predicted values of the primary instrument. Examples of model transfer methods are simple slope and bias correction<sup>45</sup>, model updating<sup>90</sup>, and regression equation transfer<sup>80</sup>. However, for both spectral and model transfer, there is a requisite to measure a number of samples (the transfer set) on both the primary and a secondary instrument. The transfer set spectra should closely match the spectra of the samples used to build the PLS model and must exhibit sufficient variability to account for the difference between the two instruments. Maintaining transfer set integrity is a challenge since contamination and compositional change can happen during transportation or storage over time due to component volatility and reactivity. This is particularly true with fuels which are composed of thousands of chemicals. If a viable transfer set cannot be maintained over time, then the master instrument must be maintained in pristine condition: an equally challenging task. In order to overcome these limitations, we have described virtual standard (VSSB) calibration transfer in chapter IV, in which the transfer set can be digitally synthesized using the NIR spectra of a twelve neat solvents. Although, VSSB eliminates the need to maintain a master instrument once the transfer set is acquired on the master instrument, one of the twelve solvents used (mineral oil) is not a pure chemical. This brings up two points: 1) can the same mineral oil be obtained at a later date and 2) is the mineral oil immune to compositional change over time. Both of these points are only valid for pure chemicals. To overcome the limitation of VSSB, a new model transfer method is described herein, involving the use of spectra of pure molecular chemicals to digitally generate segmented spectra of virtual standards that mimic the training set spectra used to build the PLS models. Virtual standard spectra from both the primary and a secondary instrument are then predicted on the PLS models and a simple slope and bias correction can be applied to the secondary instrument.

## EXPERIMENTAL SECTION

*INSTRUMENTATION*

Two FUELex NIR spectrometers (BrukerOptics, The Woodlands, TX, [www.brukeroptics.com/fuelex](http://www.brukeroptics.com/fuelex)) were used to collect the NIR spectra of fuels and chemicals. The FUELex consists of an  $f\#/2$  spectrograph with a 512 element InGaAs array detector and 3 nm resolution. The spectral wavelength range is 880 - 1570 nm. Each spectral acquisition consisted of 30 averaged spectra each with an integration time between 8 to 16 ms. The integration time for each instrument was automatically selected so that the signal intensity of an air reference yielded between 32,000 and 42,000 counts on the 16 bit A/D converter. This ensured that sufficient signal was present while preventing the A/D from saturating when the probe was inserted into the fuel. Total acquisition time for each spectrum was less than one second. All fuel spectra exhibited absorbance ranges within 0-0.8 absorbance units.

A third MPA NIR FT instrument (BrukerOptics, The Woodlands, TX, [www.brukeroptics.com/MPA](http://www.brukeroptics.com/MPA)) was used to collect the NIR spectra of fuels and chemicals. The spectral wavenumber range is 4000 - 12000  $\text{cm}^{-1}$ .

Each instrument was equipped with a transreflectance dip probe with a fixed total pathlength of 1 cm. All spectra were collected at room temperature (22-25 Co) by immersing a clean probe several inches into the sample and then swirling the probe to ensure no air bubbles were trapped. No effort was made to maintain a set fiber-optic geometry. All of the fuel and chemical spectra were preprocessed by Savitsky-Golay 7-point first derivative followed by a standard normal variate (SNV) transform to remove spectral off-set and baseline changes due to the variability of fiber-optic geometry in both Fuelex and MPA instruments. All chemicals were analytical grade (Aldrich).

## *BUILDING PLS MODELS*

NIR spectra of 420 JP-8 jet fuels were acquired using the primary instrument. The fuel samples were obtained from the Naval Research Laboratory in Washington D.C and from the Army TARDEC facility in Warren Michigan. The initial origin of the fuels ranged geographically on a sample by sample basis throughout the North American, European, African, Australian and Asian continents. World-wide representation of fuel samples ensured variance in the training set used for PLS model building. ASTM reference values for each fuel set were provided; yet, most fuels lacked reference values for all the three properties listed in Table 12. ChemWave software (Bruker Optics, Woodlands, TX, USA) was used to construct and validate all PLS models of jet fuel properties from the spectra acquired on the primary instrument. Constructing a PLS model involves the following steps;

- 1) Selecting fuel spectra whose ASTM number for the property of interest are known, and combining these spectra with their corresponding property values into a training set,
- 2) Constructing a PLS model which correlates spectral intensity of the property of interest,
- 3) Selecting the number of principal components (latent variable) to use in the model,
- 4) Identifying and removing any outlier values,
- 5) Validating PLS model using leave-one-out cross validation,
- 6) Validating PLS model with an independent test set.

The number of latent variables was determined using the IND function and an F-test<sup>53</sup> with (n-k) degree of freedom and a probability level of 95%. For all the three property models shown in this chapter, five latent variables were used. Five latent variables were appropriate and showed the lowest SECV (standard error of cross-validation) with the minimum over-fitting of noise. Samples were considered outliers if both studentized residuals and leverage exceeded 98% probability of training set inclusion. Sample leverage is calculated using Eqn. 50:

$$\text{Sample leverage} = [\text{scores} (\mathbf{SCORES} \times \mathbf{SCORES}^T) \text{scores}^T]/n \quad (50)$$

where **SCORES** is a matrix of the entire training set where each row is a sample spectrum and numbers of columns give the latent variables used and **SCORES<sup>T</sup>** is the transposed matrix, and where scores is a row vector containing the amount of each latent variable, and scores<sup>T</sup> is the transposed vector, and *n* is the number of samples in the training set.

In step 5, a particular model was validated by leave-one-out cross validation where one sample is removed from the training set, and then a PLS model was built using the remaining samples and finally this model is used to predict the property value of the excluded sample. This process is repeated until all the samples in the training set are excluded, each one once. The standard error of cross validation (SECV) is then calculated as in Eqn. 51:

$$\text{SECV} = [(\Sigma(\text{error})^2)/(n-k)]^{1/2} \quad (51)$$

where error is the difference between the true value and the predicted value for all leave-one-out iterations, *n* is the total number of samples used to build the model, and *k* is the number of latent variable used to generate a model.

The ASTM reproducibility, the standard error of cross-validation, the mean and the standard deviation of actual values, and the correlation coefficient for the predicted vs. actual results are given in Table 12. ASTM reproducibility values reflect the maximum expected inter-lab measurements error for the primary ASTM test method at a 95% confidence level (2.6 times the standard deviation of the ASTM measured population). The standard error of cross validation (SECV<sub>M</sub>) for the PLS models of jet fuels properties are within the ASTM primary methods. The standard deviations of the jet fuel properties reflect the global nature of the jet fuel samples used in building these PLS models and the correlation coefficients values indicate that all the samples used are linearly correlated.

NIR spectra of an additional 180 JP-8 jet fuels (TARDEC, Warren, MI, USA) were acquired on both the primary and secondary instruments to form an independent test

set (step 6). The independent test acquired on the primary instrument was used to validate the PLS models. The standard errors of prediction (SEP) are given in Table 13.

## SEGMENTED VIRTUAL STANDARD CALIBRATION TRANSFER THEORY

Virtual standard that mimics the actual training set acquired on a master instrument, and that is used to build PLS models, can be digitally generated by using a few spectra of neat molecular chemicals acquired on the master instrument, the spectra of the training set, and a multiple linear regression algorithm. By acquiring the spectra of these neat molecular chemicals on a secondary instrument, virtual standards that represent the change on the secondary instrument can be digitally generated. Finally, a simple but effective slope and bias correction can be applied to the secondary instrument to correct PLS models predictions. A schematic summary of the segmented virtual standards slope-bias (SVSSB) method is given in Figure 69. To generate segmented virtual standards we used 15 neat molecular chemicals. The first derivative SNV spectra of the fifteen pure chemicals (butylcyclohexane, cis-decahydronaphtalene, dodecane, ethylebenzene, hexadecane, iso-octane, meta-xylene, nonane, ortho-xylene, pristane, squalene, tetradecane, toluene, tridecane, and undecane) are shown in Figure 68. As shown in Figure 67, the target spectra (the original training set fuel spectra acquired on the primary instrument) are divided into three segments. Segment one is from 880 to 1304 nm, segment two is from 1304 to 1432 nm and segment three is from 1432 to 1570 nm. This process is repeated for the 15 pure chemicals acquired on the primary instrument.

$$\mathbf{f}_{target} = [\mathbf{f}_1 \dots \mathbf{f}_j] \quad (52)$$

$$\mathbf{x}_{chemical} = [\mathbf{x}_1 \dots \mathbf{x}_j] \quad (53)$$

where  $\mathbf{f}_{target}$  is a fuel spectrum and  $\mathbf{f}_j$  is a segment of this spectrum and  $\mathbf{x}_{chemical}$  is a chemical spectrum (toluene, nonane,... etc) and  $\mathbf{x}_j$  is a segment of this spectrum. The brackets in Eqn.s 52 and 53 represent the vector concatenation operator. For each target fuel, the three segments are then approximated (one at a time) using a linear combination



of the corresponding segments of the 15 pure chemicals. This was accomplished by first setting up a multiple linear regression:

$$f_j = \beta_{1j}x_{1j} + \beta_{2j}x_{2j} + \dots + \beta_{ij}x_{ij} + \varepsilon \quad (54)$$

where  $f_j$  is a segment of a virtual fuel standard which approximates a particular target fuel;  $x_{ij}$  is a segment  $j$  of a chemical spectrum  $i$ ;  $i$  is the chemical spectrum index (1-15);  $j$  is the spectrum segment index (1-3); and  $\beta_{ij}$  is the amount of each chemical ( $x_{ij}$ ) required to make segment  $j$  of the virtual standard. The remaining term  $\varepsilon$  is the spectral difference between the virtual standard segment and the corresponding segment of the target fuel being approximated. By applying eqn. 54 to all three segments sequentially, and then concatenating the results, a virtual standard is formed which approximates one of the target fuels. Using matrix notation, eqn. 54 can be re-written to solve this relationship for every target fuel spectrum to be digitally synthesized:

$$F_j = X_j \beta_j + \varepsilon \quad (55)$$

where,  $F_j$  is a  $n \times m$  matrix where the number of rows ( $n$ ) corresponds to the number of wavelengths in the spectral segment  $j$ , and the number of columns ( $m$ ) corresponds to the number of target fuels being approximated. Thus each column of  $F_j$  is segment  $j$  of a virtual standard. The matrix  $X_j$  is an  $n \times i$  matrix where the number of rows ( $n$ ) also corresponds to the number of wavelengths in the spectral segment  $j$ , and the number of columns ( $i$ ) corresponds to the number of pure chemicals (15). Thus each column is segment  $j$  of chemical spectrum  $i$ .  $\beta_j$  is a matrix relating how much of each chemical is required to generate the targets for segment  $j$ . By setting the spectral difference matrix ( $\varepsilon$ ) to zero, the matrix  $\beta_j$  is solved as:

$$\beta_j = (X_j^T X_j)^{-1} X_j^T F_j \quad (56)$$

using singular value decomposition. In this work, the SVD was performed at full rank (15), thus conceptually the same as a multiple linear regression. The virtual standards for a particular segment  $j$  are then generated as:

$$S_j = X_j \beta_j \quad (57)$$

The complete set of virtual standards is then given by:

$$S = [S_1..S_j] \quad (58)$$

where the brackets indicate the matrix concatenation operator and each column of  $S$  is a virtual standard approximating a target fuel acquired on the primary instrument. For chemical spectra also acquired on the primary instrument,  $S$  becomes  $S_{primary}$  (the virtual standard spectra for the primary instrument).

An example of a virtual standard generated using this method is given in Figure 69, and is overlaid with both the target fuel spectrum and a virtual standard created without using the segmented approach. Once  $\beta_j$  is determined, then chemical spectra acquired on the secondary instruments are used to generate the virtual standards for each the secondary instrument:

$$S_{secondary} = [S_{1,secondary}..S_{j,secondary}] = [X_{1,secondary}\beta_1..X_{j,secondary}\beta_j] \quad (59)$$

where this Eqn. is the same as Eqn. 55, except the chemical spectra in the  $X_{j,secondary}$  matrix are acquired on the secondary instrument instead of the primary instrument.

The virtual standards of the primary instrument ( $S_{primary}$ ) and the secondary instrument ( $S_{secondary}$ ) are then predicted in a PLS model. Subsequently, a slope and bias correction for each secondary instrument PLS prediction is generated by regressing the predictions of the primary instrument virtual standards against those of the secondary instrument virtual standards.

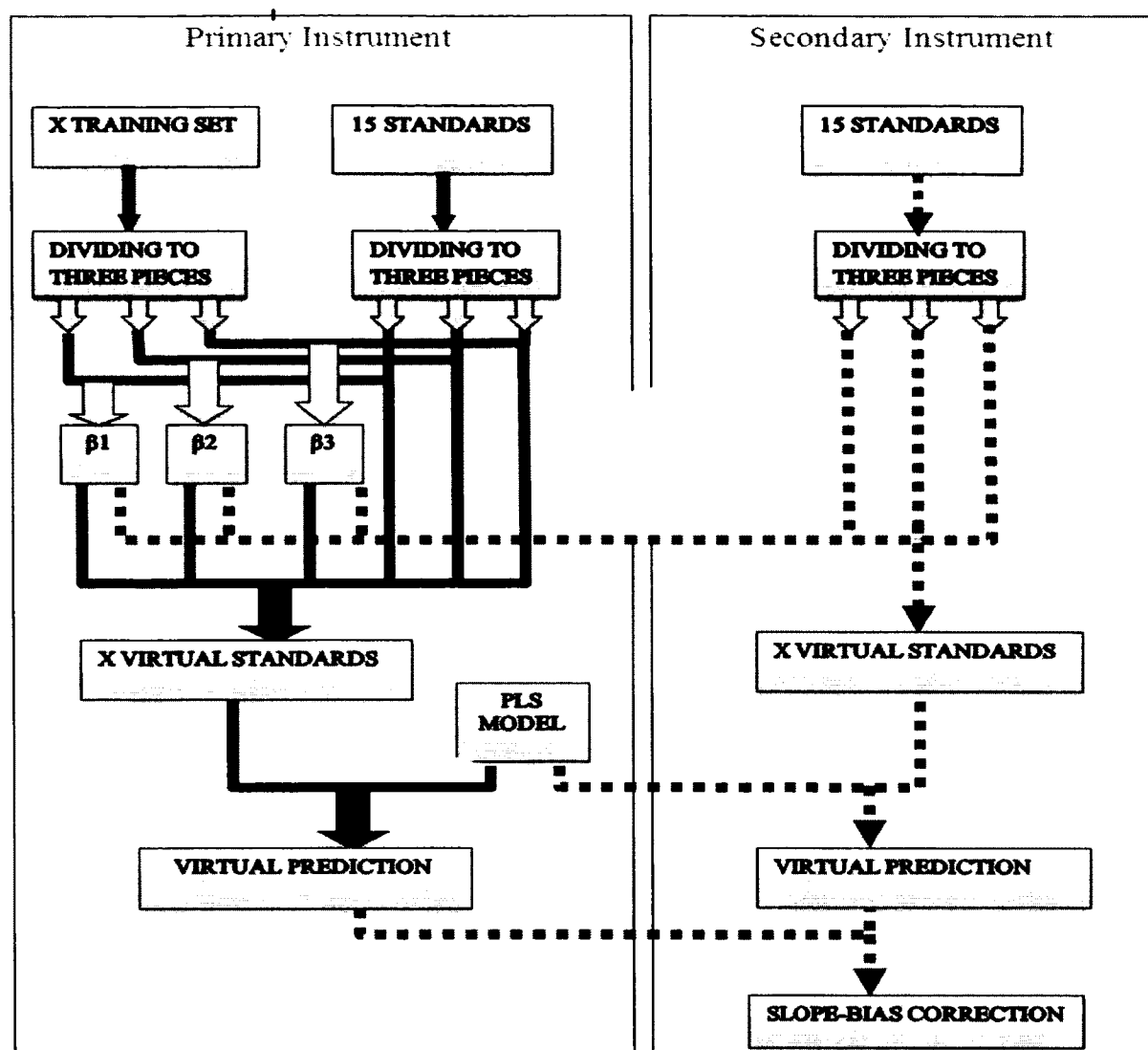


FIG. 67. Calibration transfer using segmented virtual standards which are digitally synthesized using 15 pure chemicals.

As an example, when correcting the primary instrument generated PLS model for API gravity, the primary instrument virtual standards ( $S_{primary}$ ) are predicted using the PLS model to yield  $P_{primary}$ , and the secondary instrument virtual standards ( $S_{secondary}$ ) are predicted using the PLS model to yield  $P_{secondary}$ . The slope-bias correction for future API gravity predictions on the secondary instrument is then given by solving:

$$P_{primary} = a * P_{secondary} + b \quad (60)$$

for the slope (a) and the bias (b). Future predictions on the secondary instrument are simply corrected as:

$$P_{corrected} = a * P_{secondary} + b \quad (61).$$

The essence of this method is that once the  $\beta_j$  values have been determined for the primary instrument, transfers can be carried out to an unlimited number of secondary instruments with only one requirement: the acquisition of the 15 pure chemicals on the secondary instrument. Since each chemical has a discrete molecular formula, they can always be obtained in the future in pristine condition. There is also no longer a need to maintain the primary instrument once the 15 chemical spectra have been acquired on it.

## RESULTS AND DISCUSSION

*SVSSB CALIBRATION TRANSFER BETWEEN SIMILAR NIR-INSTRUMENTS*

NIR spectra of 180 JP-8 jet fuels (TARDEC, Warren, MI, USA) were acquired on both the primary and secondary instruments to form an independent test set. The 13 jet fuel properties along with the ASTM primary test method, the ASTM reproducibility, the standard error using leave-one-out validation, the mean and the standard deviation of actual values, and the correlation coefficients for the predicted vs. actual results are shown in table 12.

TABLE 12. Summary of the jet fuel PLS models and properties.

Jet fuel property	ASTM method*	Units	Number of samples	ASTM Reproducibility	SECV <sub>M</sub>	Mean	Std. Dev.	r <sup>2</sup>
API Gravity	D1298	°API	291	0.3	0.265	43.94	3.38	0.98
Aromatics	D1319	% (vol.)	156	3	0.47	17.89	2.00	0.94
Cetane Index	D976	Cetan #	138	2	0.309	45.83	1.72	0.97
Density	D1319	Gml <sup>-1</sup>	270	0.0005	0.001	0.80	0.02	0.96
Distillation 10% Recovered			372	3.43-4.42	2.65	176.9	19.6	0.92
Distillation 20% Recovered			397	N/A	2.57	185.5	11.0	0.92
Distillation 50% Recovered	D86	°C	381	2.97	2.09	204.0	11.2	0.92
Distillation 90% Recovered			340	3.24-3.99	1.99	240.4	5.92	0.77
Flash Point	D93	°C	314	2.7-4.83	2.36	50.92	6.69	0.86
Freeze Point	D5972	°C	408	1.81	1.43	-49.91	3.11	0.52
Hydrogen Content	D3343	% (mass)	114	0.04	0.019	13.82	0.13	0.97
Saturates	D1319	% (vol.)	120		1.08	78.67	2.71	0.81
Viscosity@-20 °C	D445	mm <sup>2</sup> s <sup>-1</sup>	135	0.0800	0.07	4.29	0.23	0.90

\* <http://www.astm.org/>

SECV<sub>M</sub>: Standard error of validation of primary instrument property model using leave-one-out validation; Mean: mean of the training set values; Std.Dev.: standard deviation of the training set values; r<sup>2</sup>: correlation coefficient for predicted vs. actual values in leave-one-out validation plots.

The ASTM reproducibility reflects the maximum expected inter-lab measurement error for the primary ASTM test method at a 95% confidence level. The standard error of cross validation for PLS models of fuel properties are within the estimates of ASTM

primary methods. The correlation coefficients show that all the samples are linearly correlated. Thus, all NIR PLS models built on the primary instrument yield reasonable estimates of the primary methods. The results for the prediction of the independent test sets are given in table 13. For the primary instrument, the standard errors of prediction are higher for all properties except freeze point and flash point (see tables 12 and 13). This is to be expected for such large training sets since the exclusion of a single spectrum from the models during a leave-one-out procedure has little leverage relative to the large number of samples remaining. Since the primary instrument tests for each fuel were conducted at the geographically diverse points of origin of the fuels, the ASTM reproducibility represents a reasonable limit on the accuracy of the blind model predictions. Considering the global nature of the data set along with the test method reproducibility, the primary instrument prediction errors for the independent test set are consistent with expectations.

The 1<sup>st</sup> derivative SNV of the 15 pure chemicals (dodecane, pristane, etc) acquired on both primary and secondary instruments are overlaid in Figure 68. Although the two instruments yield similar spectra, there is difference between the spectra as observed on the inset. Where, the solid and the dotted line represent the 1<sup>st</sup> derivative SNV spectra acquired on the primary and the secondary instruments, respectively. We observed that dividing the spectra of training set and the 15 chemicals to three pieces result in identical fit between the training set and the segment virtual standards. As shown in Figure 69, NIR spectra of jet fuels exhibit major absorbance bands in three regions. Thus, dividing the spectra to three pieces, increase the degrees of freedom which lead to identical fit. A transformed target jet fuel spectra (taking from the training set used in PLS model building, solid line) is overlaid with its virtual standard (constructed using full target spectra and the 15 chemicals spectra, dashed line) and its segment virtual standard (constructed using segment target spectra and the segment of the 15 chemical spectra, dotted line) are displayed in Figure 69. The segment virtual standard is almost identical to the target spectrum, as observed on the inset. Although the virtual standard provides

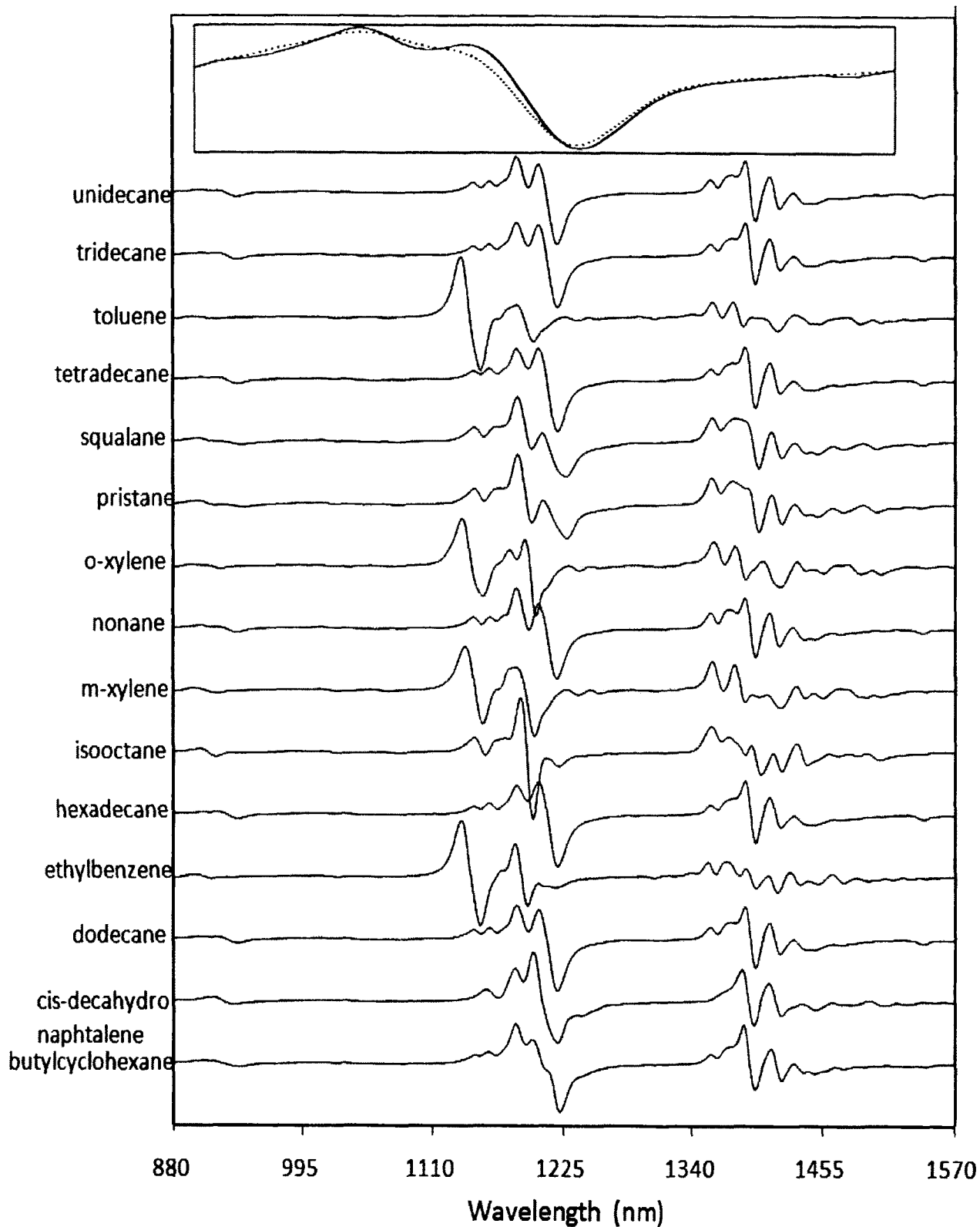


FIG. 68. The NIR absorbance of 1<sup>st</sup> derivative SNV spectra of the 15 chemicals acquired on the primary instrument (Solid line) overlaid with the secondary standards (broken line). Seven point Savitsky-Golay 1<sup>st</sup> derivative and SNV spectral transforms have been applied. Spectra offset by 6 a. u.

essentially the same absorbance in the same spectral regions as the target fuel, it is not identical to the target fuel while the segment virtual standard is. Similar results were obtained for all of the segment virtual fuels, thus ensuring a spectral variance for the segment virtual standards which is similar to that of the training sets.

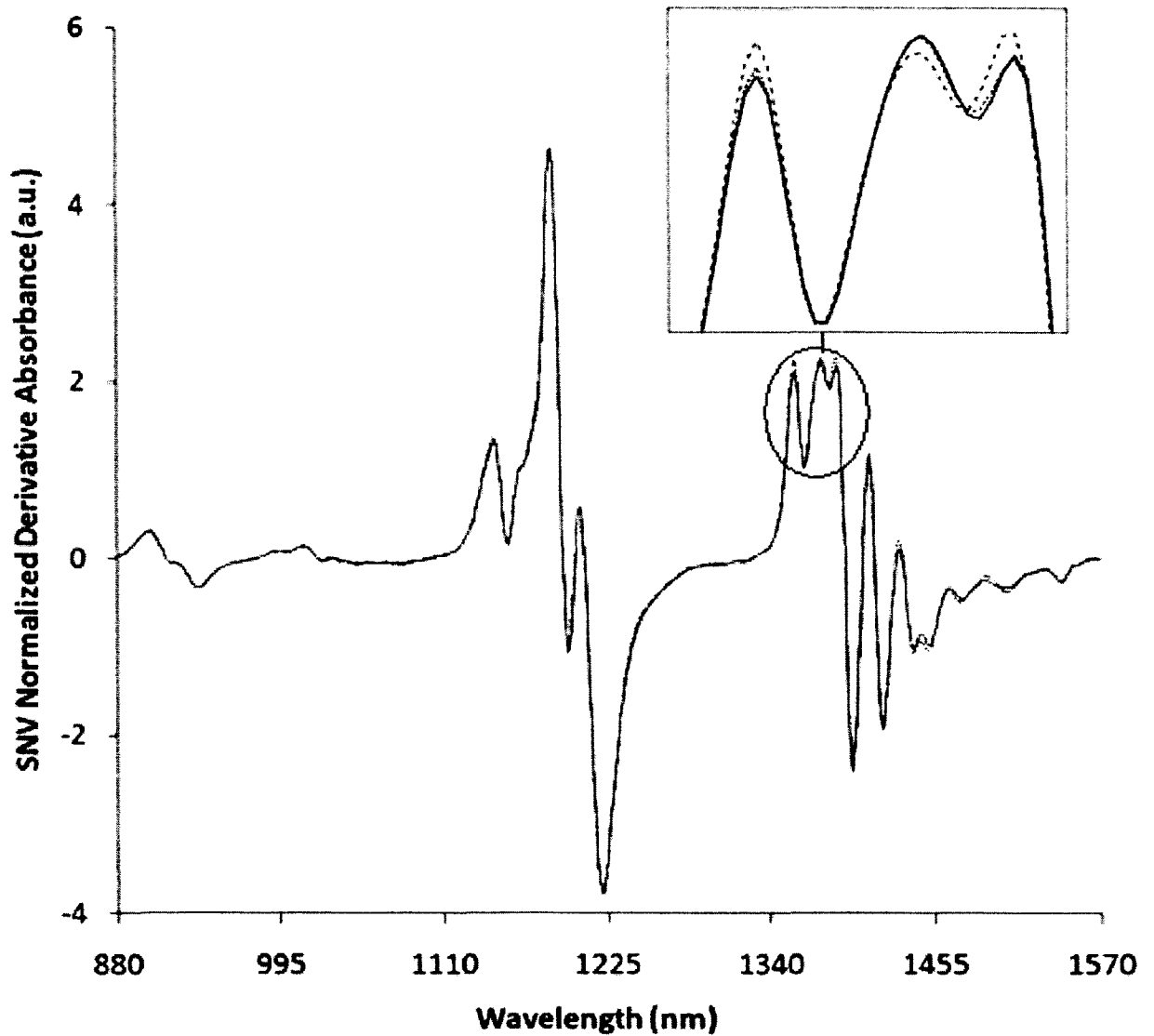


FIG. 69. The NIR absorbance spectrum of a digitally synthesized virtual standard using segment target (dotted line) overlaid with digitally synthesized virtual standard using full target (dashed line) and their target fuel spectrum (solid line). Seven point Savitsky-Golay 1st derivative and SNV spectral transforms have been applied.



The standard error of prediction without any correction ( $SEP_u$ ) for the secondary instruments using the PLS model generated on the primary instrument is shown in table 13. The secondary instrument shows more error of prediction than the primary instrument for a majority of property models. The  $SEP_u$  is four times higher than

TABLE 13. Standard error of prediction of the properties in the test set samples relative to the master instrument prediction for uncorrected, SVSSB corrected, BS corrected, and PDS corrected.

Jet fuel property	Samples Number	Primary Instrument SEP	Secondary Instrument				$r^2$
			$SEP_u$	$SEP_c$	SB	PDS	
API Gravity	180	0.32	1.46	0.18	0.18	0.24	0.93
Aromatics	180	1.23	1.97	0.32	0.33	0.35	0.96
Cetane Index	180	0.67	1.09	0.19	0.26	0.27	0.98
Density	180	0.0016	0.0058	0.00092	0.00084	0.00094	0.93
Distill. 10%	180	3.21	14.1	1.88	1.66	1.85	0.74
Distil.20%	180	3.18	15.1	2.13	1.81	1.99	0.72
Distill.50%	180	2.51	12.55	1.92	1.62	1.77	0.71
Distill. 90%	180	2.01	5.17	0.9	0.93	0.99	0.98
Flash Point	180	2.07	8.67	1.23	0.93	1.01	0.65
Freeze Point	180	0.85	3.46	0.92	0.61	0.66	0.89
Hydr. Cont.	180	0.05	0.057	0.015	0.014	0.014	0.98
Saturates	180	1.56	1.31	0.32	0.51	0.55	0.97
Vis @ -20 c	180	0.081	0.41	0.043	0.042	0.046	0.93

SEP. Standard error of prediction of primary instrument relative to the actual property values;  $SEP_u$ : Standard error of prediction of uncorrected secondary instrument relative to primary instrument prediction;  $SEP_c$ : Standard error of prediction of SVSSB corrected secondary instrument relative to primary instrument prediction; PDS: Standard error of prediction of PDS corrected secondary instrument relative to primary instrument prediction;  $r^2$ : correlation coefficient of primary prediction relative to secondary prediction.

that of the primary instrument for 10% distillation, 20% distillation, and 50% distillation and more than twice for API gravity, density, flashpoint, freeze point, and viscosity, as shown in table 13. For example, 20% distillation predicted values of secondary instrument are 15.1 C° higher than primary instrument prediction.

Table 13 also shows the standard error of prediction for the secondary instrument after correction using the segment virtual standards and a slope bias correction ( $SEP_c$ ), and normal fuel standards slope bias correction (SB). The SVSSB correction results in a

dramatic reduction in standard error of prediction. In all instances SE<sub>Pu</sub> decreases to values that are closed to that of the primary instrument SEP. For comparison, 36 jet fuel samples were taken from the target set (roughly 20% of the target set) to carry out fuel standards slope bias correction (SB). The samples were selected based on Kennard stone algorithms. Fuel standards slope bias correction leads to reduction in the standard error of prediction to be, in most cases, similar or closed to that of segment virtual standards slope bias correction SE<sub>Pc</sub>. For example, SE<sub>Pc</sub> for API gravity (0.18), % Aromatics (0.32), Hydrogen content (0.015), and viscosity at -20 C° (0.043) are similar to that of fuel standards slope and bias correction (SB) which are 0.18, 0.33, 0.014, and 0.042 respectively. While, 10%, 20%, and 50% distillation recovery, density and freeze point have standard error of prediction error of prediction in case of correction using SVSSB than correction using SB, and cetane index, 90% distillation recovery and saturates show less standard error of prediction when using SVSSB than SB, as shown in table 13.

For further comparison PDS was carried out using the same 36 samples used to carry out the fuel standard slope bias correction (BS). The standard error of prediction of PDS correction is shown in table 13. PDS correction result in reduction of prediction error of secondary instrument to values similar or close to SVSSB correction. SVSSB correction show better correction than PDS for API gravity, % Aromatic, cetane index, density, 90% distillation recovery, saturates, and viscosity, while PDS correction is a little better or similar to SVSSB correction for 10%, 20%, 50% distillation recovery, flash point, freeze point and hydrogen content (table 13). Although, the PDS error is closed to the SVSSB error, there is a small slope after applying PDS correction, as can be seen in case of 90% distillation recovery (Figure 72, D)

The effect of SVSSB as well as PDS correction is graphically depicted in Figures 70 to 73 for all properties of jet fuel of the secondary instrument. In all Figures, A, B, and C display the SVSSB correction, while D, E, and F display the PDS correction. Where open triangles represent the uncorrected secondary instrument prediction jet fuel properties relative to primary instrument prediction (solid line). Solid triangles and solid circles show secondary instrument predicted jet fuel properties after correction using SVSSB and PDS, respectively. For all of the property models, the SVSSB correction results in

predicted values which more closely match that of the primary instrument predictions values for all jet fuels properties being modeled, as shown on plots (Figures 70-73).

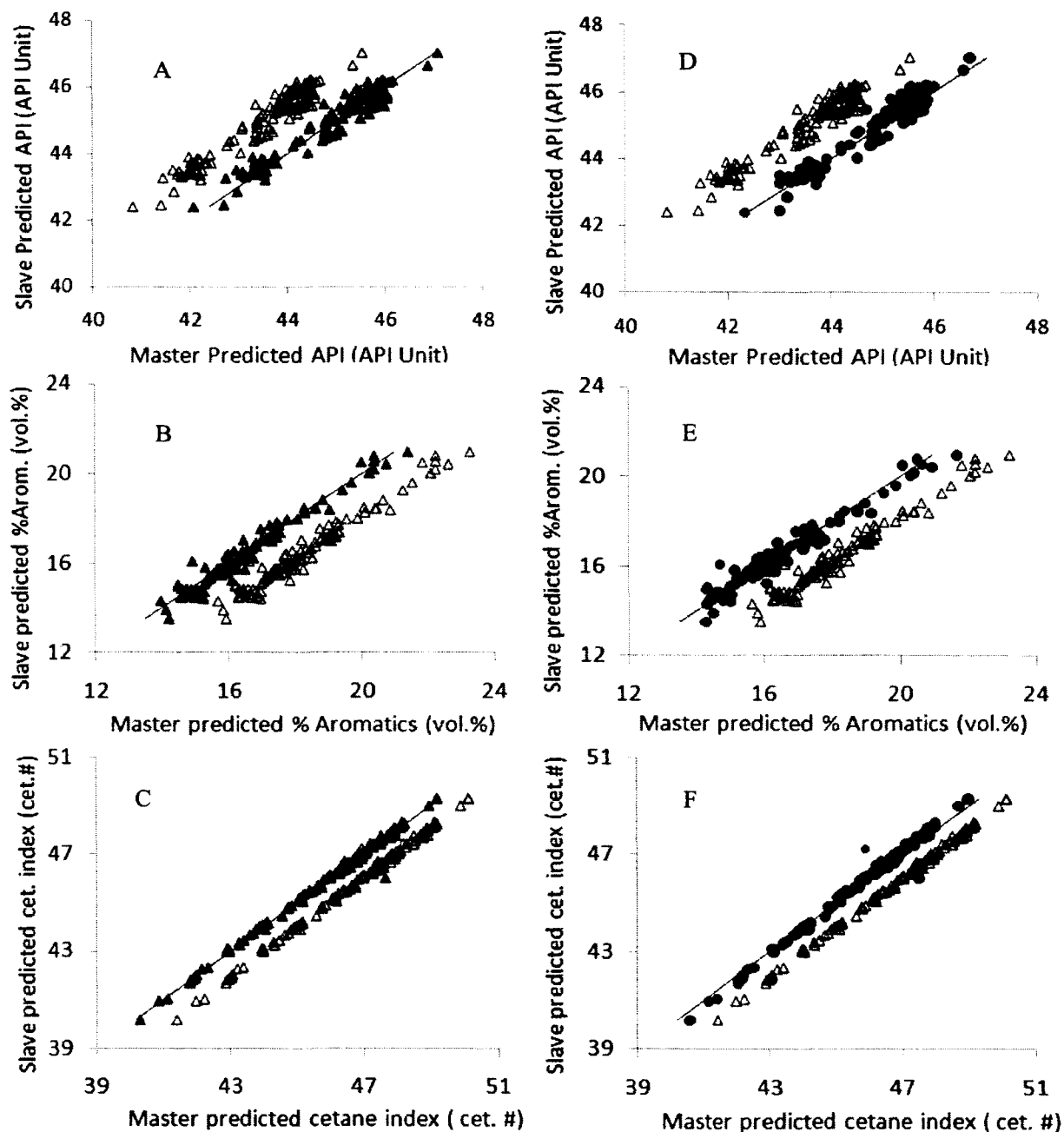


FIG. 70. Fuel property predictions of the test set fuels acquired on the secondary instrument for uncorrected (open triangles), SVSSB corrected (filled triangles), and PDS corrected (filled circles) data. A, B, and C give the API gravity, %aromatic, and cetane index SVSSB correction, respectively. D, E, and F give the API gravity, %aromatic, and cetane index PDS correction, respectively. All predictions are plotted relative to the primary instrument predictions and the solid line is primary instrument predicted vs primary predicted (ideal line).

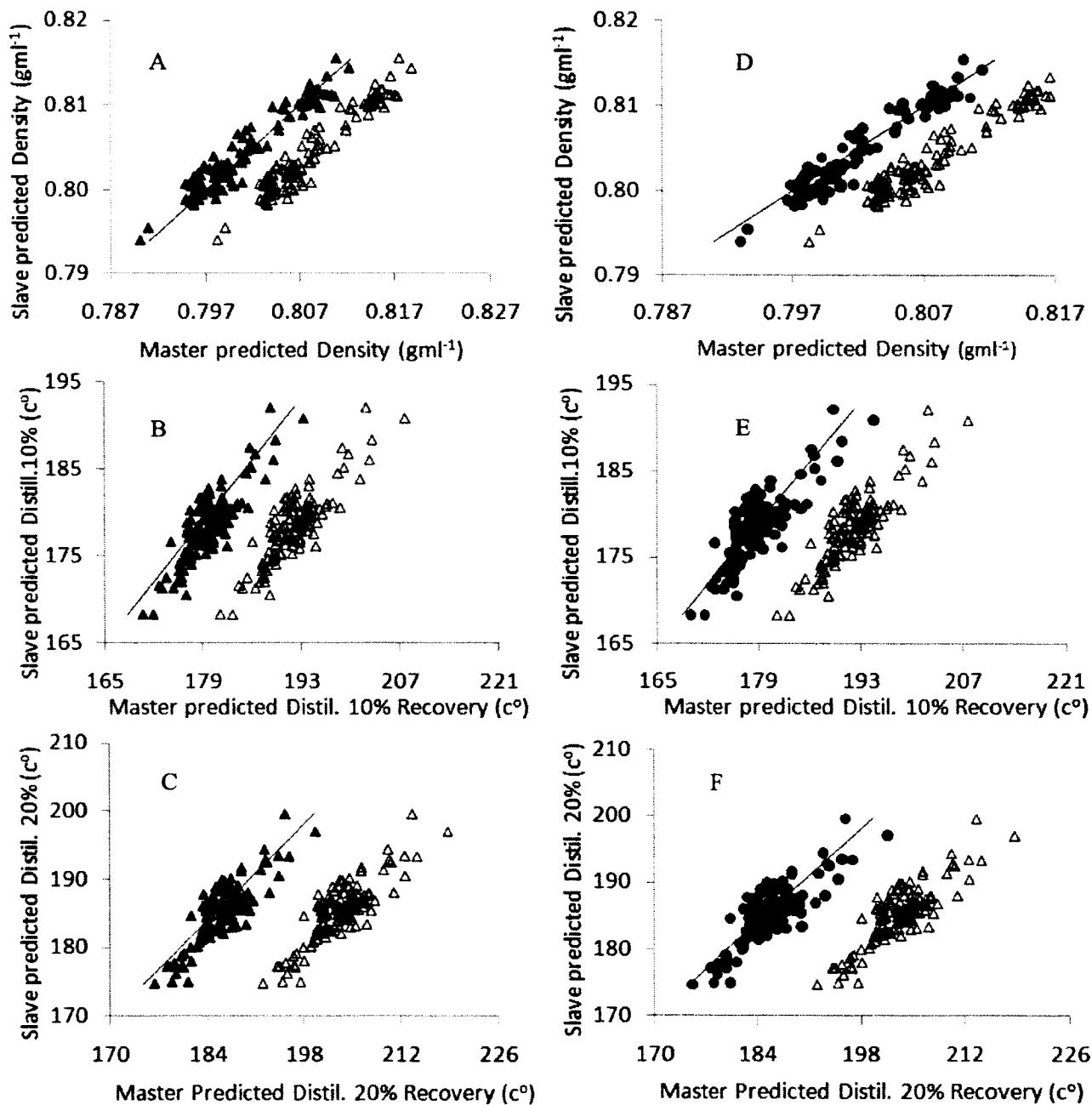


FIG. 71. Fuel property predictions of the test set fuels acquired on the secondary instrument for uncorrected (open triangles), SVSSB corrected (filled triangles), and PDS corrected (filled circles) data. A, B, and C give the density, 10%, and 20% distillation recovery SVSSB correction, respectively. D, E, and F give the density, 10%, and 20% distillation recovery PDS correction, respectively. All predictions are plotted relative to the primary instrument predictions and the solid line is primary instrument predicted vs primary instrument predicted (ideal line).

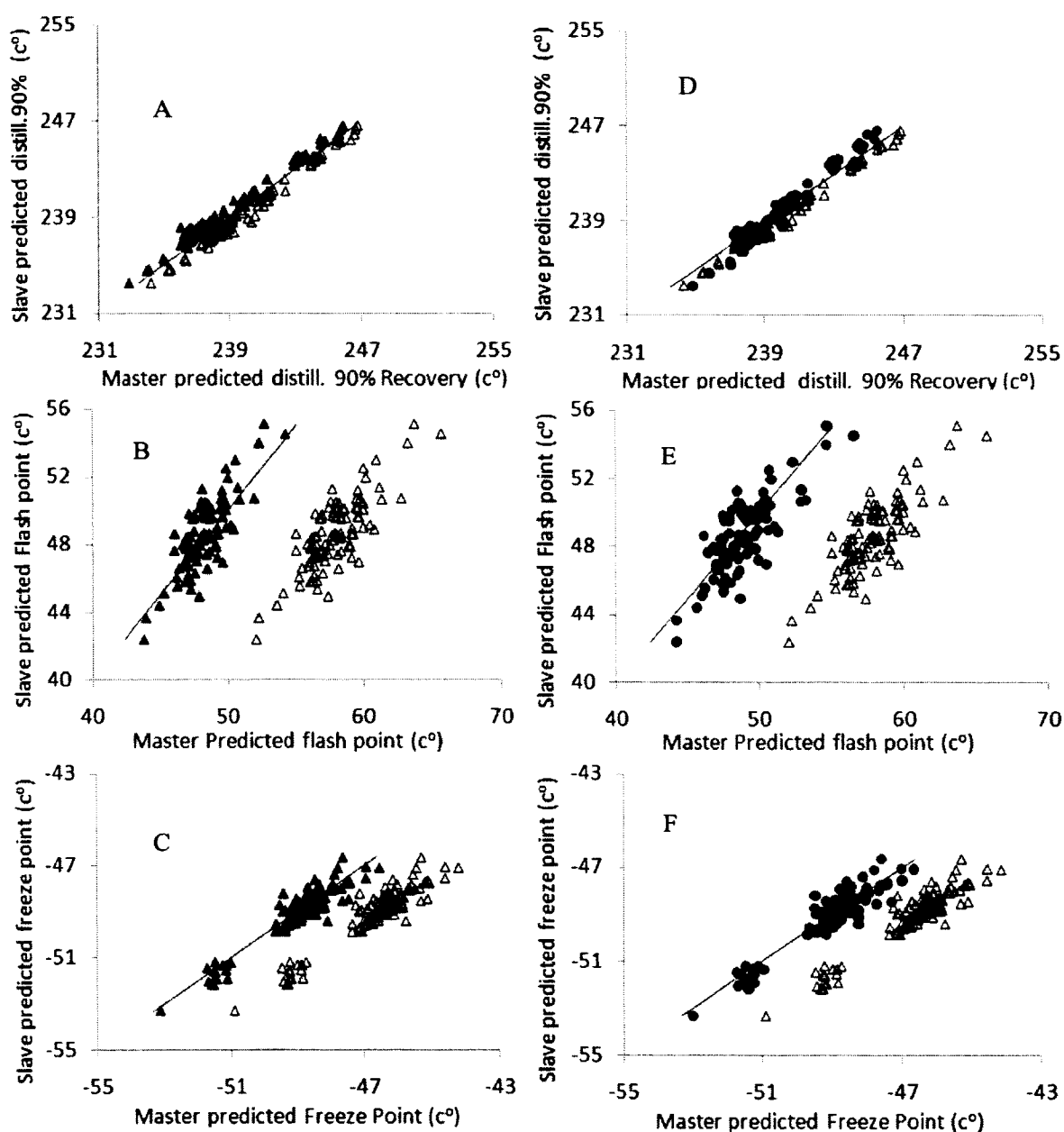


FIG. 72. Fuel property predictions of the test set fuels acquired on the secondary instrument for uncorrected (open triangles), SVSSB corrected (filled triangles) and PDS corrected (filled circles) data. A, B, and C give the 90% distillation recovery, flash point, and freeze point SVSSB correction, respectively. D, E, and F give the 90% distillation recovery, flash point, and freeze point PDS correction, respectively. All predictions are plotted relative to the primary instrument predictions and the solid line is primary instrument predicted vs primary instrument predicted (ideal line).

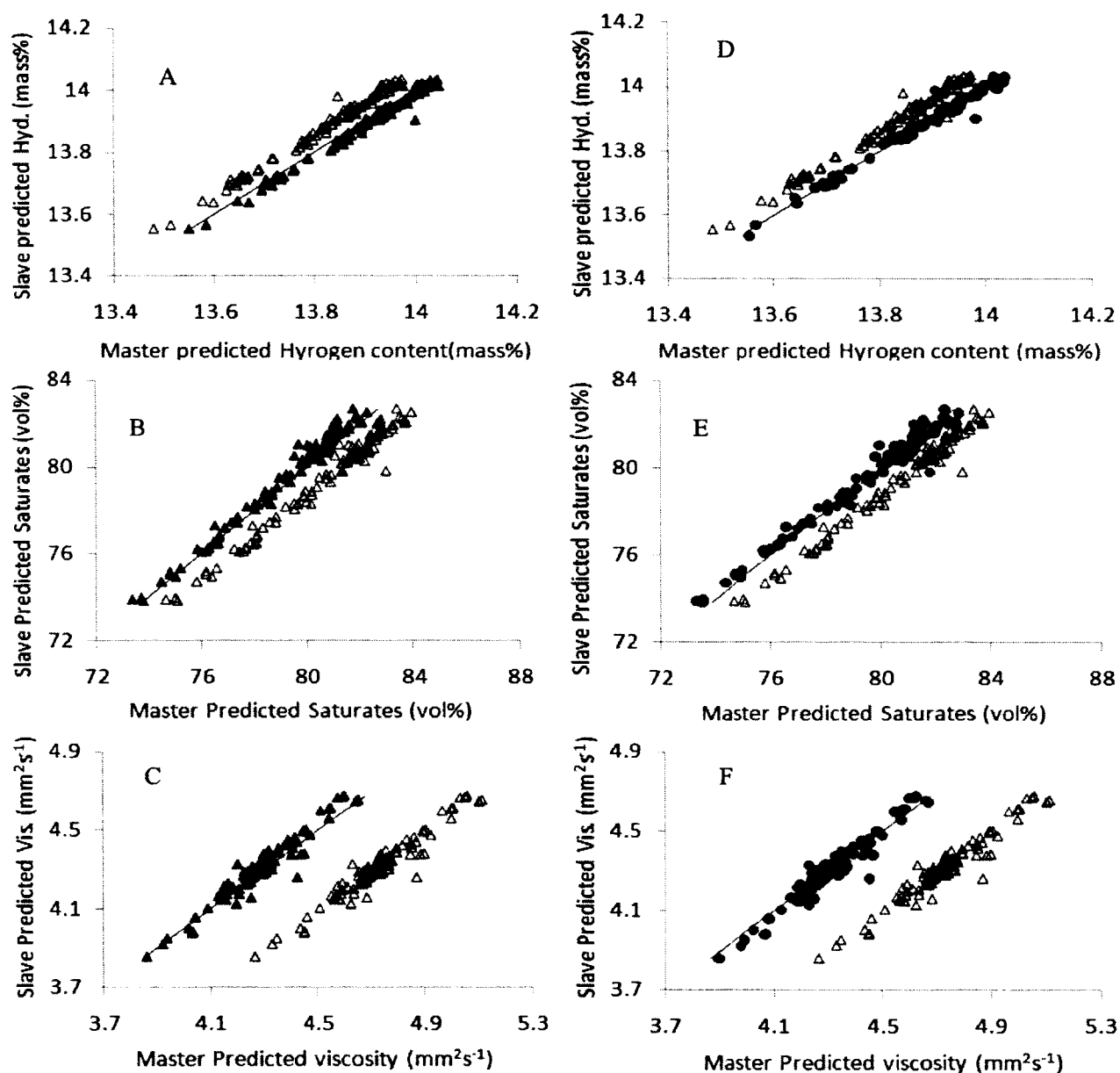


FIG. 73. Fuel property predictions of the test set fuels acquired on the secondary instrument for uncorrected (open triangles), SVSSB corrected (filled triangles) and PDS corrected (filled circles) data. A, B, and C give the % hydrogen content, % saturates, and viscosity @ -20 °C SVSSB correction, respectively. D, E, and F give % hydrogen content, % saturates, and viscosity @ -20 °C PDS correction, respectively. All predictions are plotted relative to the primary instrument predictions and the solid line is primary instrument predicted vs primary instrument predicted (ideal line).

*SVSSB CALIBRATION TRANSFER BETWEEN DIFFERENT NIR-INSTRUMENTS*

The segmented virtual standards calibration transfer was tested between two different NIR-instruments. The master instrument is a dispersive handheld NIR-instrument (FUELex) while the slave instrument is FT-NIR bench-top instrument. NIR spectra of 57 JP-8 jet fuels (TARDEC, Warren, MI, USA) were acquired on both FUELex NIR dispersive and MPA FT-NIR instruments to an independent test set. Raw JP-8 fuel spectra of acquired on both the FUELex and the MPA are shown in Figure 74, “A” shows the FUELex jet fuel raw spectrum and “B” shows the MPA jet fuel raw spectrum. As shown in Figure 74, the two spectra have different ranges and resolution. The FUELex instrument was used as the primary instrument since PLS models of the 3 shown jet fuel properties were built on it, and the MPA is the secondary instrument.

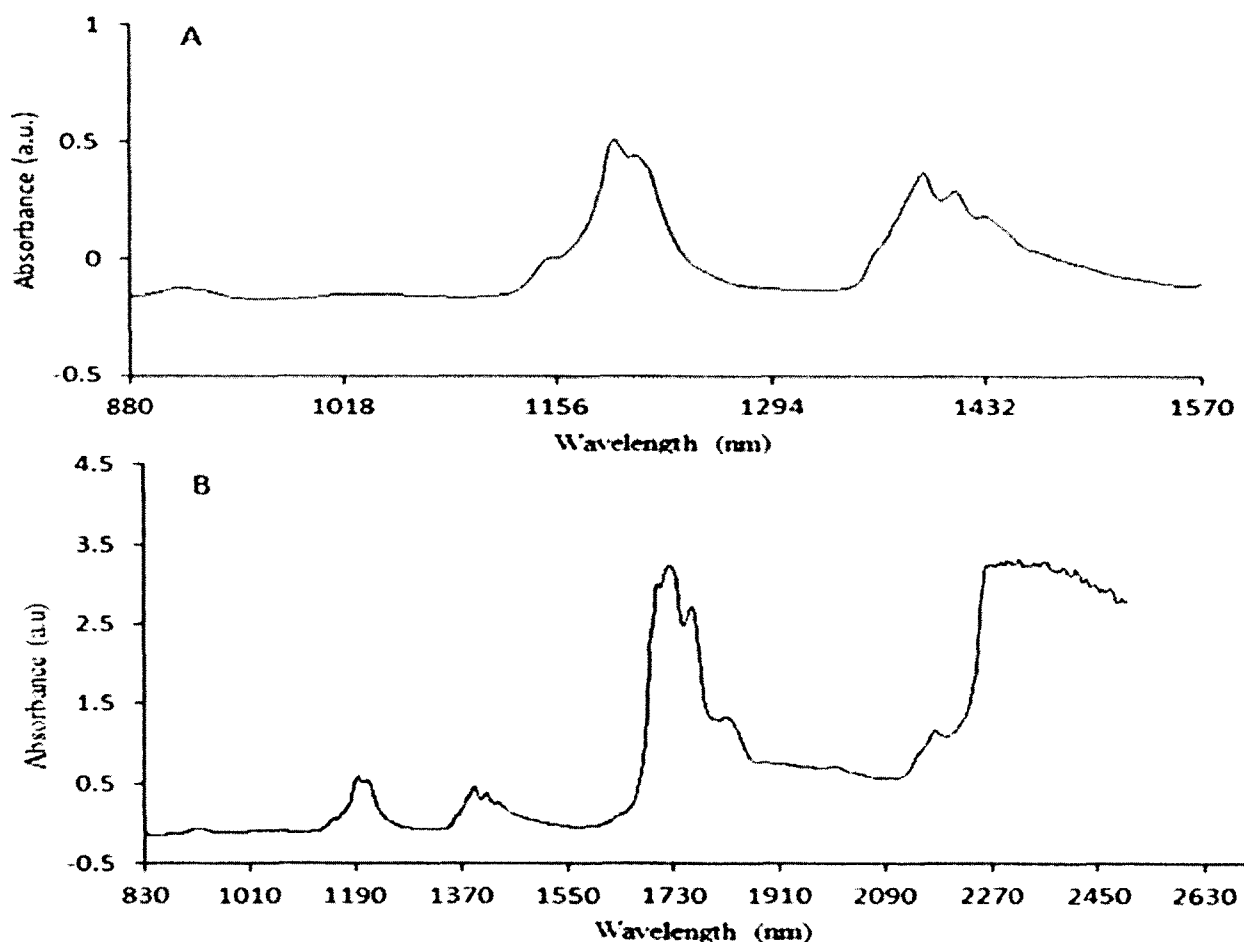


FIG. 74. (A) gives the NIR absorbance spectrum of a JP-8 fuel acquired on the FUELex dispersive NIR-instrument. (B) gives the NIR absorbance spectrum of the JP-8 fuel acquired on the MPA FT-NIR instrument.

The MPA spectra were truncated to span the same wavelength range as the FUELex (880-1570nm). The spectra were then pretreated by Savitsky 1<sup>st</sup> derivative followed by SNV. A pretreated FUELex spectrum is overlaid with a pretreated MPA spectrum in Figure 75A.

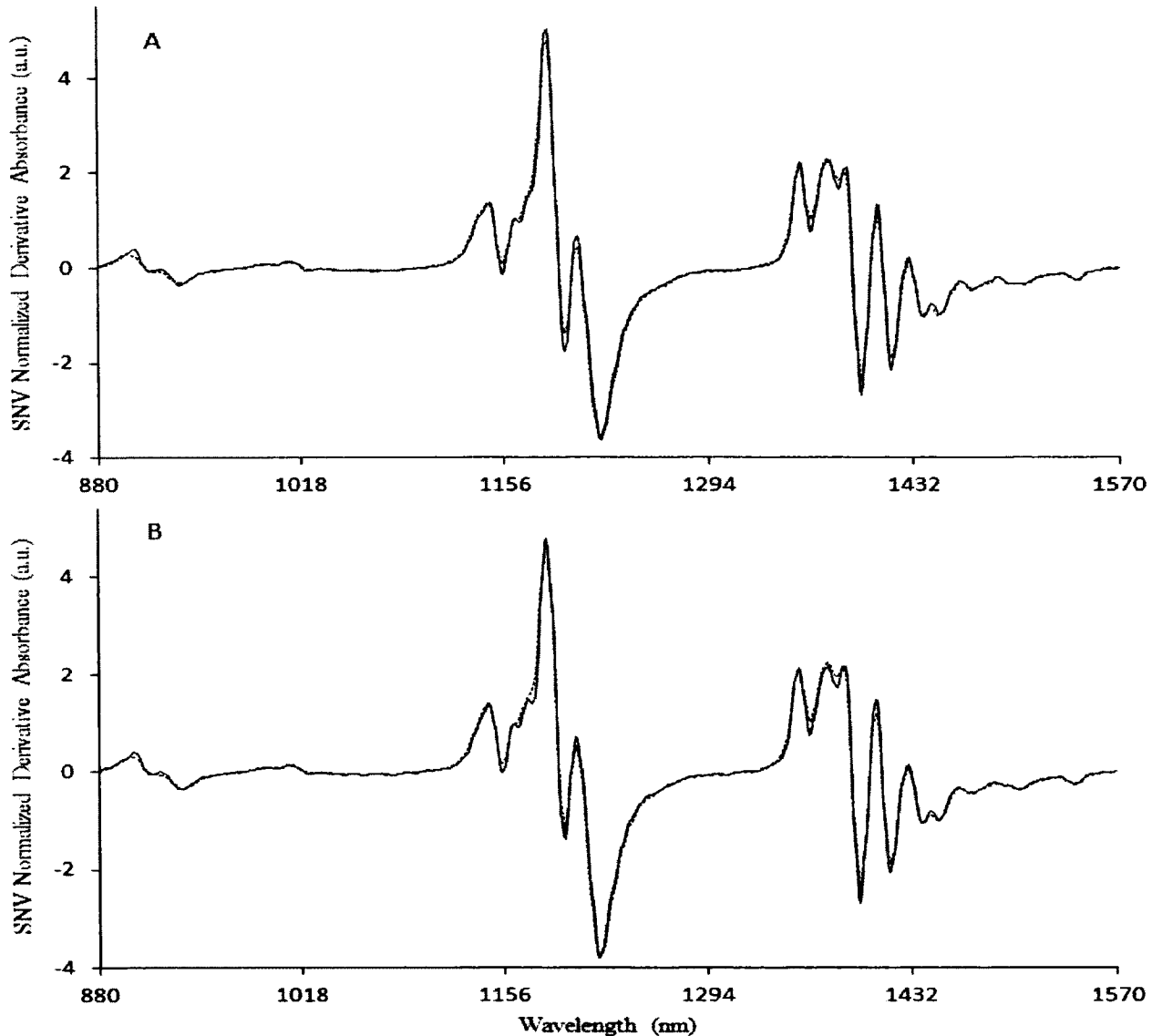


FIG. 75. NIR absorbance spectra of a digitally synthesized MPA secondary instrument virtual standard (dotted line) overlaid with a digitally synthesized FUELex virtual standard (solid line) are given in (B) and their target fuel spectra of MPA (dotted line) and FUELex (solid line) are given in (A). Seven point Savitsky-Golay 1<sup>st</sup> derivative and SNV spectral transforms have been applied.



The solid line is the MPA spectrum and the dashed line is the FUELex pretreated spectrum. The pretreated spectra were then predicted on the PLS models. The errors of prediction are tabulated in Table 14. SEP is the standard error of prediction of the primary FUELex instrument relative to the actual values. For the API gravity, cetane index, and % hydrogen content the SEP values are 0.27, 0.26, and 0.016, respectively, and all of the SEP for the primary instrument are acceptable and fall within the ASTM reproducibility values. The standard errors of prediction for the secondary MPA instrument ( $SEP_U$ ) relative to the master instrument are 5.53, 4.30, and 0.23 for API gravity, cetane index, and % hydrogen content, respectively, and all of the error are much higher than the accepted ASTM reproducibility given in Table 12.

TABLE 14. Standard error of prediction of the properties in the test set samples relative to the master instrument prediction for uncorrected and SVSSB corrected.

Jet fuel property	Samples Number	Secondary Instrument			$r^2$
		$SEP_U$	$SEP_C$	SB	
API Gravity	57	5.53	0.27	0.18	0.99
Cetane Index	57	4.30	0.26	0.25	0.99
Hydr. Cont.	57	0.23	0.016	0.035	0.99

$SEP_U$ : Standard error of prediction of uncorrected secondary instrument relative to primary instrument prediction;  $SEP_C$ : Standard error of prediction of SVSSB corrected secondary instrument relative to primary instrument prediction;  $r^2$ : correlation coefficient of primary prediction relative to secondary prediction.

To apply the SVSSB calibration transfer, the 15 neat chemicals were acquired on both the FUELex and the MPA instruments. The spectra of the chemicals were then pretreated and used to generate virtual standards as described previously. In Figure 75B, a virtual standard of the FUELex (dashed line) is overlaid with a MPA virtual standard (solid line). As can be observed, the difference between the two virtual standards (Figure 75B) is the same as the real fuel spectra (Figure 75A).

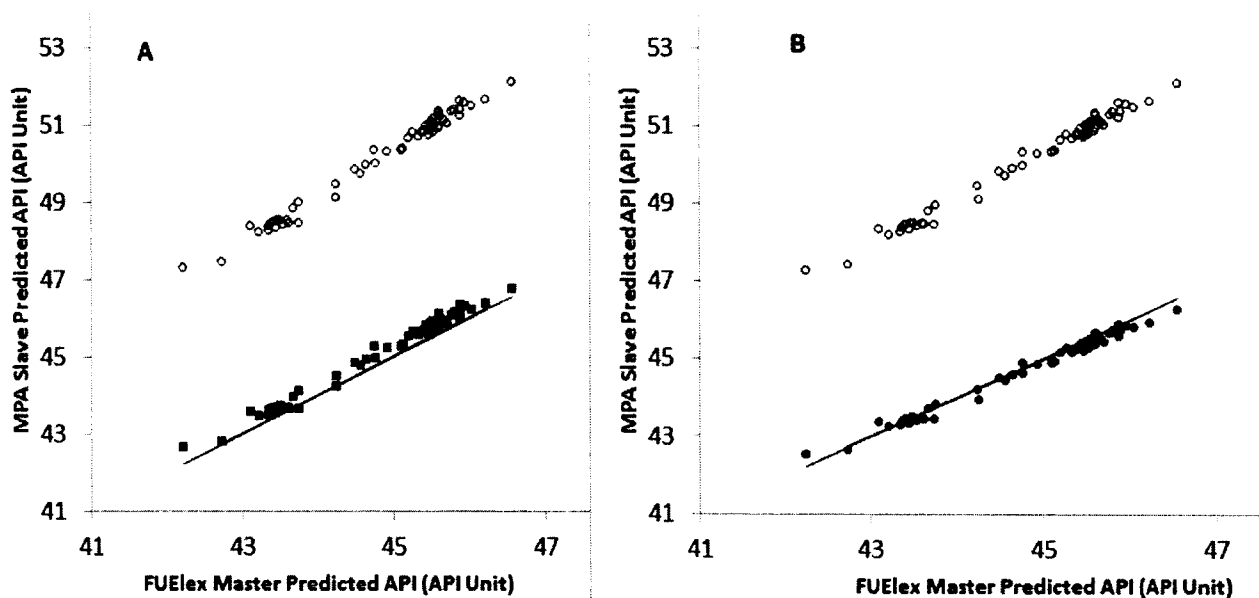


FIG. 76. API predictions of the test set fuels acquired on the MPA secondary instrument for uncorrected (open circles), SVSSB corrected (solid squares) and BS corrected (solid circles) data. (A) Shows SVSSB correction and (B) shows SB correction using the actual fuel standards. The solid line is an ideal line (slope=1 and intercept =0).

The SVSSB calibration transfer was applied to the MPA instrument. The standard errors of prediction for the MPA secondary instrument after correction ( $SEP_C$ ) is shown in Table 14 and are 0.27, 0.26, and 0.016 for API gravity, cetane index, and % hydrogen content, respectively. After applying SVSSB calibration transfer to the secondary instrument, the  $SEP_C$  is now within the ASTM reproducibility for all the three Jet fuel properties. In addition, the calibration transfer was applied to the secondary instrument using the real fuel spectra and a simple slope bias correction (SB) to test the reliability of SVSSB. The SB prediction errors are so close to the SVSSB prediction errors except in the case of hydrogen content where the SVSSB prediction is better than SB.

The effect of SVSSB as well as a simple slope and bias using actual fuel standards (SB) correction is graphically depicted in Figures 76-78 for the API, cetane index, and % hydrogen content respectively. In all Figures, “A” displays the SVSSB correction and “B” displays the SB correction. In addition, the open markers represent the uncorrected MPA secondary instrument prediction and the solid marker represents the corrected MPA secondary instrument prediction. Also, for all Figures, the solid line represents an ideal relation between the primary and secondary predictions (slope =1 and

bias = 0). As can be seen, for all of the property models, the SVSSB correction results in predicted values which more closely match that of the primary instrument's values for all jet fuels properties being modeled and in all cases the results are comparable to that of SB using actual fuel standards to transfer the calibration.

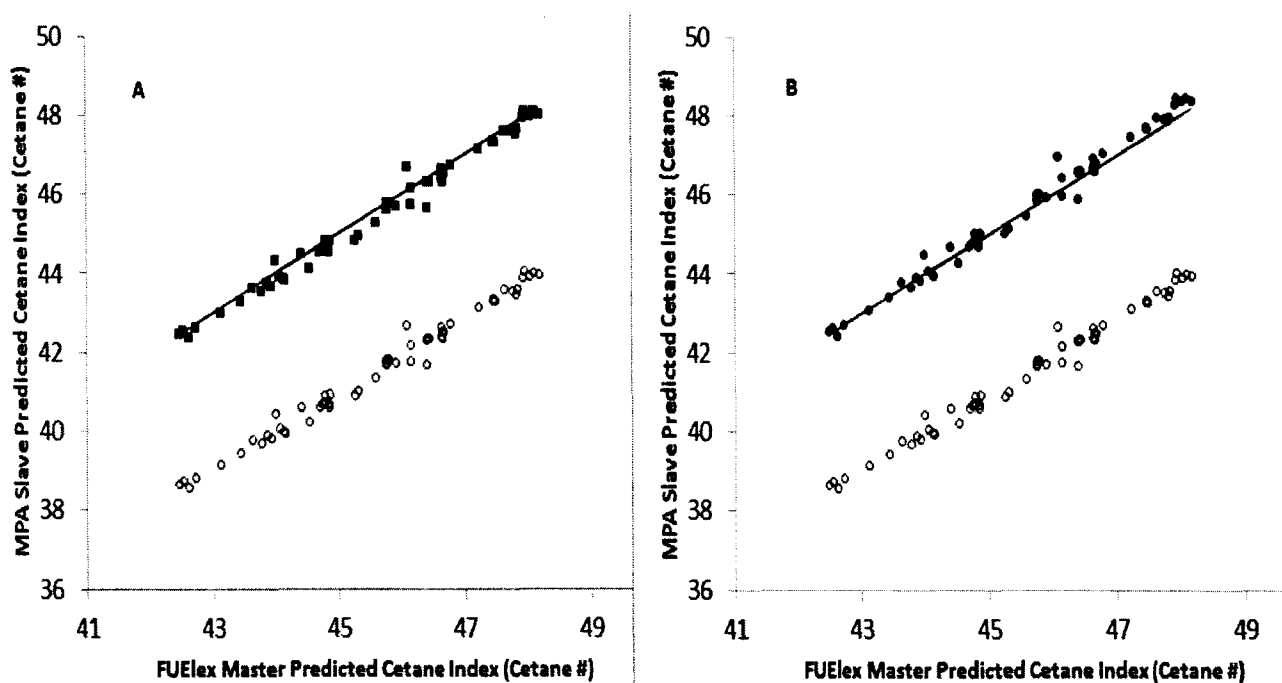


FIG. 77. Cetane index predictions of the test set fuels acquired on the MPA secondary instrument for uncorrected (open circles), SVSSB corrected (solid squares) and BS corrected (solid circles) data. (A) Shows SVSSB correction and (B) shows SB correction using the actual fuel standards. The solid line is an ideal line (slope=1 and intercept =0).

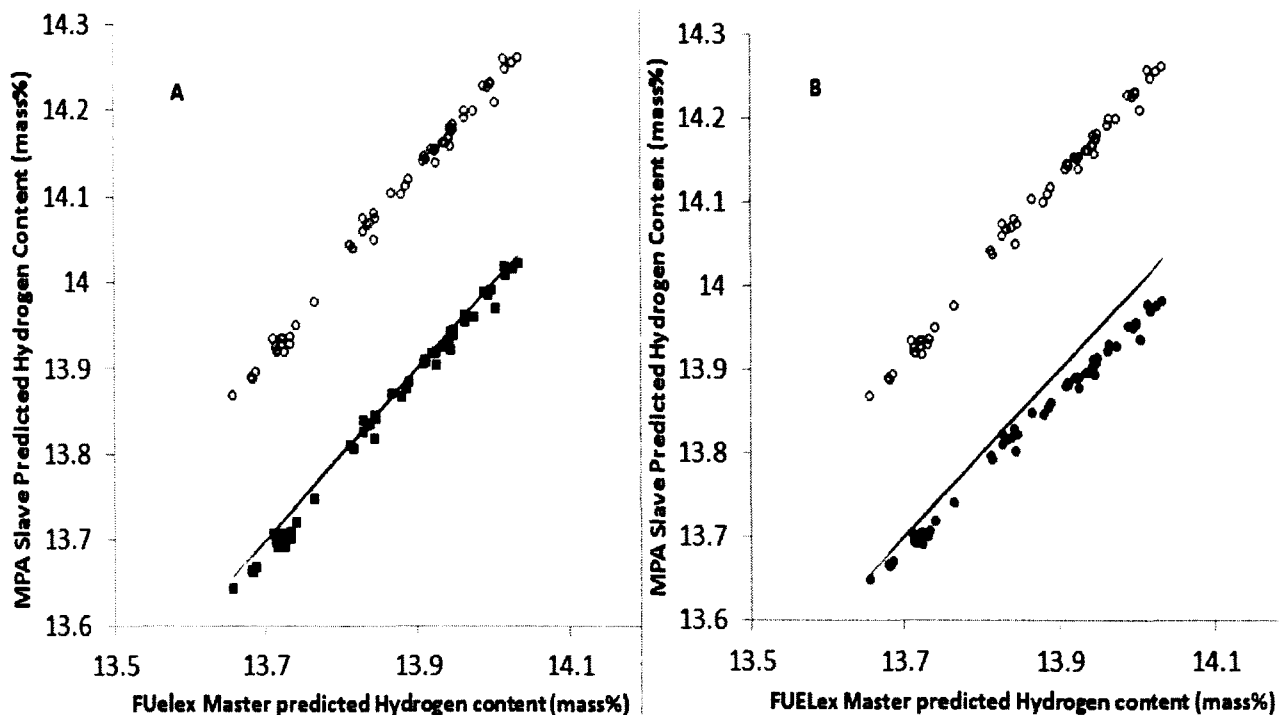


FIG. 78. % Hydrogen content predictions of the test set fuels acquired on the MPA secondary instrument for uncorrected (open circles), SVSSB corrected (solid squares) and BS corrected (solid circles) data. (A) Shows SVSSB correction and (B) shows SB correction using the actual fuel standards. The solid line is an ideal line (slope=1 and intercept =0).

As shown in this chapter, PLS models of jet fuels properties were correctly transferred using virtual standards constructed using a segmented approach from 15 pure chemicals standards acquired on both the primary and a secondary instrument. The SVSSB calibration transferred was carried out between two dispersive instruments and between a dispersive and an FT instruments. The standard error of prediction of SVSSB corrected secondary instrument is similar to the standard error of prediction for a slope and bias correction using real fuels. The use of the SVSSB approach eliminates the need to maintain fuel standards and a primary instrument, once the chemical spectra have been collected on the primary instrument. In addition, the acquisition of the spectra of 15 pure chemicals is all that is required to enable future calibration transfers to secondary instruments.

## CHAPTER VI

### CONCLUSIONS

#### ELIMINATING RAMAN OPTICAL INTERFERENCES

The primary usefulness of the SSE Raman method with the described instrumentation lies in its universal ability to extract Raman spectra from fluorescence interference using inexpensive and compact instrumentation in a time efficient manner. Using the dual-laser SSE allows extracting Raman spectra from high fluorescence samples while having a full Raman spectrum that includes both the fingerprint as well as the CH-stretching regions. The SSE method offers superior SNR performance when compared to shifted excitation methods which do not use iterative algorithms as well as sub-second processing times which allows real-time results. If derivative data is desired, it is possible to calculate the derivative using conventional algorithms and any changes in the resulting bandwidth are solely dependent on the algorithm used. This is not the case with SERDS or PCA methods whose derivative bandwidths are dependent on the experimental conditions (excitations chosen). In addition, SSE Raman is the only method which allows all of the excitation conditions to be easily varied in a predefined manner (number of excitations, separation of excitations, and integration time of each excitation) in order to obtain optimal results. Even in the absence of a background, SSE Raman improves the S/N ratio by reducing random shot and thermal noise and by eliminating fixed pattern and random spike noise. In this dissertation an un-cooled CCD detector was used. The use of a cooled CCD would allow much longer integration times to be used for weak Raman scatters without having the thermal noise impact the dynamic range of the 16 bit A/D converter.

The ability to tune the DBR lasers to set wavelengths allows the excitation shift to be set to a constant value. This simplifies implementation of the iteration algorithms since the shift index of the weighting vector can be set equal to the constant excitation shift (or an integral number multiplied times the excitation shift) and thus allows a single algorithm to be used even for different instruments. The dual-laser SSE shows superior performance in the case of highly fluorescent samples, while both the single and the dual-laser perform the same in the case of low or no fluorescence background.

## VIRTUAL STANDARD CALIBRATION TRANSFER

The use of the VSSB correction performed as well as the conventional PDS correction when transferring the jet fuel property models. However, the VSSB method does not require the use of jet fuels to transfer the models from the master instrument to the slave instrument. The use of a virtual standard, spectra from solvents, to transfer the PLS models eliminated the need to maintain fuel standards which are prone to compositional changes. In VSSB, although neat solvents were used, one of the solvents (mineral oil) is not a pure chemical. This brings up two points: 1) can the same mineral oil be obtained at a later date and 2) is the mineral oil immune to compositional change over time. Both of these points are only valid for pure chemicals. Although the use of mineral oil was seen as a useful for proof of the concept of the VSSB method, the use of SVSSB approach solve this problem. The use of the SVSSB approach eliminates the need to maintain fuel standards and a primary instrument, once the chemical spectra have been collected on the primary instrument. In addition, the acquisition of the spectra of 15 pure chemicals is all that is required to enable future calibration transfers to secondary instruments.

## REFERENCES

1. T. Hasegawa, J. Nishijo, and J. Umemura. "Separation of Raman spectra from fluorescence emission background by principal component analysis". *Chem. Phys. Lett.* 2000. 317(6): 642.
2. Z. Jianhua, L. Harvey, I. M. David, and Z. Haishan. "Automated Autofluorescence Background Subtraction Algorithm for Biomedical Raman Spectroscopy". *Appl. Spectrosc.* 2007. 61(11): 1225.
3. Z. Jianhua, H. Lui, D. I. McLean, and Z. Haishan. "Integrated real-time Raman system for clinical in vivo skin analysis". *Skin Res. Technol.* 2008. 14(4): 484.
4. C. A. Lieber and A. Mahadevan-Jansen. "Automated Method for Subtraction of Fluorescence from Biological Raman Spectra". *Appl. Spectrosc.* 2003. 57(11): 1363.
5. P. A. Mosier-Boss, S. H. Lieberman, and R. Newbery. "Fluorescence Rejection in Raman Spectroscopy by Shifted-Spectra, Edge Detection, and FFT Filtering Techniques". *Appl. Spectrosc.* 1995. 49(5): 630.
6. D. Zhang and D. Ben-Amotz. "Enhanced Chemical Classification of Raman Images in the Presence of Strong Fluorescence Interference". *Appl. Spectrosc.* 2000. 54(9): 1379.
7. I. G. Cormack, M. Mazilu, K. Dholakia, and C. S. Herrington. "Fluorescence suppression within Raman spectroscopy using annular beam excitation". *Appl. Phys. Lett.* 2007. 91(2): 023903.
8. E. C. Le Ru, L. C. Schroeter, and P. G. Etchegoin. "Direct Measurement of Resonance Raman Spectra and Cross Sections by a Polarization Difference Technique". *Anal. Chem.* 2012. 84(11): 5074.
9. M. V. Schulmerich, R. Reddy, A. K. Kodali, L. J. Elgass, K. Tangella, and R. Bhargava. "Dark Field Raman Microscopy". *Anal. Chem.* 2010. 82(14): 6273.
10. Benniston, C. A, Matousek, P, McCulloch, E. I, Parker, W. A, Towrie, and M. "Detailed Picosecond Kerr-Gated Time-Resolved Resonance Raman Spectroscopy and Time-Resolved Emission Studies of Merocyanine 540 in Various Solvents". *J. Phys. Chem. A.* 2003. 107(22): 4347.

11. E. V. Efremov, J. B. Buijs, C. Gooijer, and F. Ariese. "Fluorescence Rejection in Resonance Raman Spectroscopy Using a Picosecond-Gated Intensified Charge-Coupled Device Camera". *Appl. Spectrosc.* 2007. 61(6): 571.
12. N. Overall, T. Hahn, P. Matousek, A. W. Parker, and M. Towrie. "Picosecond Time-Resolved Raman Spectroscopy of Solids: Capabilities and Limitations for Fluorescence Rejection and the Influence of Diffuse Reflectance". *Appl. Spectrosc.* 2001. 55(12): 1701.
13. V. E. Evtim, B. B. Joost, G. Cees, and A. Freek. "Fluorescence Rejection in Resonance Raman Spectroscopy Using a Picosecond-Gated Intensified Charge-Coupled Device Camera". *Appl. Spectrosc.* 2007. 61(6): 571.
14. A. Lakshmana, B. Mallick, and S. Umamathy. "Ultrafast Raman loss spectroscopy: a new approach to vibrational structure determination". *Curr. Sci* (00113891). 2009. 97(2): 210.
15. D. Mandal, M. Mizuno, and T. Tahara. "Temporal fluorescence rejection in Raman spectroscopy using femtosecond up-conversion with single- and multi-channel detection". *J. Mol. Struct.* 2005. 735-736: 189.
16. D. V. Martyshkin, R. C. Ahuja, A. Kudriavtsev, and S. B. Mirov. "Effective suppression of fluorescence light in Raman measurements using ultra fast time gated charge coupled device camera". *Rev. Scientific Instruments.* 2004. 75(3): 630.
17. P. Matousek, M. Towrie, and A. W. Parker. "Fluorescence background suppression in Raman spectroscopy using combined Kerr gated and shifted excitation Raman difference techniques". *J. Raman Spectrosc.* 2002. 33(Copyright (C) 2011 American Chemical Society (ACS): 238.
18. M. Maiwald, G. Erbert, A. Klehr, H. D. Kronfeldt, H. Schmidt, B. Sumpf, and G. Traenkle. "Rapid shifted excitation Raman difference spectroscopy with a distributed feedback diode laser emitting at 785 nm". *Appl. Phys. B: Lasers Opt.* 2006. 85(Copyright (C) 2011 American Chemical Society (ACS): 509.
19. M. Maiwald, H. Schmidt, and B. Sumpf. "Microsystem 671 nm light source for shifted excitation Raman difference spectroscopy". *Applied Optics.* 2009. 48(15): 2789.



20. P. Matousek, M. Towrie, and A. W. Parker. "Simple Reconstruction Algorithm for Shifted Excitation Raman Difference Spectroscopy". *Appl. Spectrosc.* 2005. 59(6): 848.
21. K. H. Michaelian, H. Yuan, R. H. Hall, and J. T. Bulmer. "Fluorescence rejection in Raman spectra of Syncrude Sweet Blend distillation fractions". *Spectrochimica Acta Part A: Mol. Biomol. Spectrosc.* 2005. 62(1-3): 582.
22. I. Osticioli, A. Zoppi, and E. M. Castellucci. "Shift-Excitation Raman Difference Spectroscopy Difference Deconvolution Method for the Luminescence Background Rejection from Raman Spectra of Solid Samples". *Appl. Spectrosc.* 2007. 61(8): 839.
23. S. Wolf and H. Doring. "Application of cheap lasers in shifted excitation Raman difference spectroscopy". J. Popp, W. Drexler, V. V. Tuchin and D. L. Matthews, Eds. (SPIE, Brussels, Belgium, 2012), p. 84271A.
24. C. Xie and Y.-q. Li. "Confocal micro-Raman spectroscopy of single biological cells using optical trapping and shifted excitation difference techniques". *J. of Appl. Phys.* 2003. 93(5): 2982.
25. J. Zhao, M. M. Carrabba, and F. S. Allen. "Automated Fluorescence Rejection Using Shifted Excitation Raman Difference Spectroscopy". *Appl. Spectrosc.* 2002. 56(7): 834.
26. W. Zou, Z. Cai, and J. Wu. "Fluorescence rejection by shifted excitation Raman difference spectroscopy". *Proc. SPIE.* 2010. 7855(Copyright (C) 2011 American Chemical Society (ACS). All Rights Reserved. ): 78551M/1.
27. S. T. McCain, R. M. Willett, and D. J. Brady. "Multi-excitation Raman spectroscopy technique for fluorescence rejection". *Opt. Express.* 2008. 16(15): 10975.
28. R. Willett. "Multiscale Reconstruction for Photon-Limited Shifted Excitation Raman Spectroscopy". in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on* (2007), p. III.
29. J. Cooper, C. Larkin, and M. Abdelkader. "Virtual standard slope and bias calibration transfer of partial least squares jet fuel property models to multiple near infrared spectroscopy instruments". *J. Near Infrared Spectrosc.* 2011. 19(2): 139.

30. J. B. Cooper, C. M. Larkin, and M. F. Abdelkader. "Calibration transfer of near-IR partial least squares property models of fuels using virtual standards". *J. Chemom.* 2011. 25(9): 496.
31. J. B. Cooper, C. M. Larkin, J. Schmitigal, R. E. Morris, and M. F. Abdelkader. "Rapid Analysis of Jet Fuel Using a Handheld Near-Infrared (NIR) Analyzer". *Appl. Spectrosc.* 2011. 65(2): 187.
32. J. B. Cooper, K. L. Wise, W. T. Welch, R. R. Bledsoe, and M. B. Sumner, "Determination of Weight Percent Oxygen in Commercial Gasoline: A Comparison between FT-Raman, FT-IR, and Dispersive Near-IR Spectroscopies". *Appl. Spectrosc.* 1996. 50(7): 917.
33. J. B. Cooper, K. L. Wise, W. T. Welch, M. B. Sumner, B. K. Wilt, and R. R. Bledsoe. "Comparison of Near-IR, Raman, and Mid-IR Spectroscopies for the Determination of BTEX in Petroleum Fuels". *Appl. Spectrosc.* 1997. 51(11): 1613.
34. J. A. Cramer, K. E. Kramer, K. J. Johnson, R. E. Morris, and S. L. Rose-Pehrsson. "Automated wavelength selection for spectroscopic fuel models by symmetrically contracting repeated unmoving window partial least squares". *Chemometr. Intell. Lab. Sys.* 2008. 92(1): 13.
35. J. A. Cramer, R. E. Morris, B. Giordano, and S. L. Rose-Pehrsson. "Partial Least-Squares Predictions of Nonpetroleum-Derived Fuel Content and Resultant Properties When Blended with Petroleum-Derived Fuels". *Energy Fuels.* 2009. 23(2): 894.
36. J. Jaumot, J. C. Menezes, and R. Tauler. "Quality assessment of the results obtained by multivariate curve resolution analysis of multiple runs of gasoline blending processes". *J. Chemom.* 2006. 20(1-2): 54.
37. K. E. Kramer, R. E. Morris, S. L. Rose-Pehrsson, J. Cramer, and K. J. Johnson. "Statistical Significance Testing as a Guide to Partial Least-Squares (PLS) Modeling of Nonideal Data Sets for Fuel Property Predictions". *Energy Fuels.* 2007. 22(1): 523.
38. T. r. A. Lestander and R. Samuelsson. "Prediction of Resin and Fatty Acid Content of Biorefinery Feedstock by On-line Near-Infrared (NIR) Spectroscopy". *Energy Fuels.* 2010. 1.
39. S. M. Maggard, U.S. Patent 5,362,965.1994.
40. S. M. Maggard, U.S. Patent 5,349,188.1994.

41. S. M. Maggard and W. T. Welch, U.S. Patent 5,370,790. 1994.
42. R. E. Morris, M. H. Hammond, J. A. Cramer, K. J. Johnson, B. C. Giordano, K. E. Kramer, and S. L. Rose-Pehrsson. "Rapid Fuel Quality Surveillance through Chemometric Modeling of Near-Infrared Spectra". *Energy Fuels*. 2009. 23(3): 1610.
43. W. T. Welch, M. B. Sumner, B. K. Wilt, R. R. Bledsoe, and S. M. Maggard, U.S. Patent 5,712,481. 1998.
44. Z.-n. Xing, J.-x. Wang, Y. Ye, and G. Shen. "Rapid Quantification of Kinematical Viscosity in Aviation Kerosene by Near-Infrared Spectroscopy". *Energy Fuels*. 2006. 20(6): 2486.
45. E. Bouveresse and D. L. Massart. "Standardisation of near-infrared spectrometric instruments: A review". *Vib. Spectrosc.* 1996. 11(1): 3.
46. U. Horrmann and N. Zanier-Szdlowski. "Portability of near infrared spectroscopic calibrations for petrochemical parameters". *J. Near Infrared Spectrosc.* 1999. 7(1): 33.
47. P. Geladi, D. MacDougall, and H. Martens. "Linearization and Scatter-Correction for Near-Infrared Reflectance Spectra of Meat". *Appl. Spectrosc.* 1985. 39(3): 491.
48. K. E. Kramer, R. E. Morris, and S. L. Rose-Pehrsson. "Comparison of two multiplicative signal correction strategies for calibration transfer without standards". *Chemometr. Intell. Lab. Sys.* 2008. 92(1): 33.
49. R. J. Barnes, M. S. Dhanoa, and S. J. Lister. "Standard Normal Variate Transformation and De-trending of Near-Infrared Diffuse Reflectance Spectra". *Appl. Spectrosc.* 1989. 43(5): 772.
50. T. B. Blank, S. T. Sum, S. D. Brown, and S. L. Monfre. "Transfer of Near-Infrared Multivariate Calibrations without Standards". *Anal. Chem.* 1996. 68(17): 2987.
51. Koehler, G. W. Small, R. J. Combs, R. B. Knapp, and R. T. Kroutil. "Calibration Transfer Algorithm for Automated Qualitative Analysis by Passive Fourier Transform Infrared Spectrometry". *Anal. Chem.* 2000. 72(7): 1690.
52. B. Walczak, E. Bouveresse, and D. L. Massart. "Standardization of near-infrared spectra in the wavelet domain". *Chemometr. Intell. Lab. Sys.* 1997. 36(1): 41.
53. Y. Wang, D. J. Veltkamp, and B. R. Kowalski. "Multivariate instrument standardization". *Anal. Chem.* 1991. 63(23): 2750.

54. E. Bouveresse, C. Hartmann, D. L. Massart, I. R. Last, and K. A. Prebble. "Standardization of Near-Infrared Spectrometric Instruments". *Anal. Chem.* 1996. 68(6): 982.
55. E. Bouveresse and D. L. Massart. "Improvement of the piecewise direct standardisation procedure for the transfer of NIR spectra for multivariate calibration". *Chemometr. Intell. Lab. Sys.* 1996. 32(2): 201.
56. E. Bouveresse, D. L. Massart, and P. Dardenne. "Modified Algorithm for Standardization of Near-Infrared Spectrometric Instruments". *Anal. Chem.* 1995. 67(8): 1381.
57. P. J. Gemperline, J. Cho, P. K. Aldridge, and S. S. Sekulic. "Appearance of Discontinuities in Spectra Transformed by the Piecewise Direct Instrument Standardization Procedure". *Anal. Chem.* 1996. 68(17): 2913.
58. T. Fearn. "On orthogonal signal correction". *Chemometr. Intell. Lab. Sys.* 2000. 50(1): 47.
59. R. N. Feudale, Y. Liu, N. A. Woody, H. Tan, and S. D. Brown. "Wavelet orthogonal signal correction". *J. Chemom.* 2005. 19(1): 55.
60. J. Sjöblom, O. Svensson, M. Josefson, H. Kullberg, and S. Wold. "An evaluation of orthogonal signal correction applied to calibration transfer of near infrared spectra". *Chemometr. Intell. Lab. Sys.* 1998. 44(1-2): 229.
61. N. A. Woody, R. N. Feudale, A. J. Myles, and S. D. Brown. "Transfer of Multivariate Calibrations between Four Near-Infrared Spectrometers Using Orthogonal Signal Correction". *Anal. Chem.* 2004. 76(9): 2595.
62. N. G. Yee and G. G. Coghill. "Factor selection strategies for orthogonal signal correction applied to calibration of near-infrared spectra". *Chemometr. Intell. Lab. Sys.* 2003. 67(2): 145.
63. F. Despagne, B. Walczak, and D.-L. Massart, "Transfer of Calibrations of Near-Infrared Spectra Using Neural Networks" *Appl. Spectrosc.* 1998. 52(5): 732.
64. R. N. Feudale, H. Tan, and S. D. Brown. "Piecewise orthogonal signal correction". *Chemometr. Intell. Lab. Sys.* 2002. 63(2): 129.

65. R. N. Feudale, N. A. Woody, H. Tan, A. J. Myles, S. D. Brown, and J. Ferré. "Transfer of multivariate calibration models: a review". *Chemometr. Intell. Lab. Sys.* 2002. 64(2): 181.
66. P. Geladi. "Some recent trends in the calibration literature". *Chemometr. Intell. Lab. Sys.* 2002. 60(1-2): 211.
67. C. V. Greensill, P. J. Wolfs, C. H. Spiegelman, and K. B. Walsh. "Calibration Transfer between PDA-Based NIR Spectrometers in the NIR Assessment of Melon Soluble Solids Content". *Appl. Spectrosc.* 2001. 55(5): 647.
68. F. A. Honorato, R. K. H. Galvão, M. F. Pimentel, B. de Barros Neto, M. C. U. Araújo, and F. R. de Carvalho. "Robust modeling for multivariate calibration transfer by the successive projections algorithm". *Chemometr. Intell. Lab. Sys.* 2005. 76(1): 65.
69. B. Igne and C. R. Hurburgh. "Local chemometrics for samples and variables: optimizing calibration and standardization processes". *J. Chemom.* 2010. 24(2): 75.
70. H. B. Paul Geladi, E. D. rring, and J. T. bakk, Henrik Antti, Svante Wold, Bo Karlberg. "Calibration transfers for predictig lake-water pH from near infrared spectra of lake sediments". *J. Near Infrared Spectrosc.* 1999. 7(4): 251.
71. D. G. Robert, M. W. Christine, J. P. Randy, and M. H. David. "Importance of Prediction Outlier Diagnostics in Determining a Successful Inter-vendor Multivariate Calibration Model Transfer". *Appl. Spectrosc.* 2007. 61(7): 747.
72. H. W. Tan and S. D. Brown. "Wavelet hybrid direct standardization of near-infrared multivariate calibrations". *J. Chemom.* 2001. 15(8): 647.
73. C. M. Wehlburg, D. M. Haaland, and D. K. Melgaard. "New Hybrid Algorithm for Transferring Multivariate Quantitative Calibrations of Intra-vendor Near-Infrared Spectrometers". *Appl. Spectrosc.* 2002. 56(7): 877.
74. J. Yoon, B. Lee, and C. Han. "Calibration transfer of near-infrared spectra based on compression of wavelet coefficients". *Chemometr. Intell. Lab. Sys.* 2002. 64(1): 1.
75. R. E. Wrolstad, E. A. Decker, S. J. Schwartz, and C. Ebooks. "Handbook of Food Anal. Chem". (John Wiley & Sons, Inc., Hoboken, 2004).
76. J. A. Jones, I. R. Last, B. F. MacDonald, and K. A. Prebble. "Development and transferability of near-infrared methods for determination of moisture in a freeze-dried injection product". *J. Pham. and Biomed. Anal.* 1993. 11(11-12): 1227.

77. S. M. Maggard, U.S. Patent 5,243,546.1993.
78. A. Candolfi and D. L. Massart. "Model Updating for the Identification of NIR Spectra from a Pharmaceutical Excipient". *Appl. Spectrosc.* 2000. 54(1): 48.
79. K. E. Kramer and G. W. Small. "Digital Filtering and Model Updating Methods for Improving the Robustness of Near-Infrared Multivariate Calibrations". *Appl. Spectrosc.* 2009. 63(2): 246.
80. M. R. Kunz, J. Ottaway, J. H. Kalivas, and E. Andries. "Impact of standardization sample design on Tikhonov regularization variants for spectroscopic calibration maintenance and transfer". *J. Chemom.* 2010. 24(3-4): 218.
81. A. J. Myles, T. A. Zimmerman, and S. D. Brown. "Transfer of Multivariate Classification Models Between Laboratory and Process Near-Infrared Spectrometers for the Discrimination of Green Arabica and Robusta Coffee Beans". *Appl. Spectrosc.* 2006. 60(10): 1198.
82. C. M. Wehlburg, D. M. Haaland, D. K. Melgaard, and L. E. Martin. "New Hybrid Algorithm for Maintaining Multivariate Quantitative Calibrations of a Near-Infrared Spectrometer". *Appl. Spectrosc.* 2002. 56(5): 605.
83. M. Forina, G. Drava, C. Armanino, R. Boggia, S. Lanteri, R. Leardi, P. Corti, P. Conti, R. Giangiacomo, C. Galliena, R. Bigoni, I. Quartari, C. Serra, D. Ferri, O. Leoni, and L. Lazzeri. "Transfer of calibration function in near-infrared spectroscopy". *Chemometr. Intell. Lab. Sys.* 1995. 27(2): 189.
84. Y. Sulub and G. W. Small. "Spectral Simulation Methodology for Calibration Transfer of Near-Infrared Spectra". *Appl. Spectrosc.* 2007. 61(4): 406.
85. L. Zhang, G. W. Small, and M. A. Arnold. "Multivariate Calibration Standardization across Instruments for the Determination of Glucose by Fourier Transform Near-Infrared Spectrometry". *Anal. Chem.* 2003. 75(21): 5905.
86. M. F. Abdelkader, J. B. Cooper, and C. M. Larkin. "Calibration transfer of partial least squares jet fuel property models using a segmented virtual standards slope-bias correction method". *Chemometr. Intell. Lab. Sys.* 2012. 110(1): 64.
87. A. Andrew and T. Fearn. "Transfer by orthogonal projection: making near-infrared calibrations robust to between-instrument variation". *Chemometr. Intell. Lab. Sys.* 2004. 72(1): 51.

88. E. Fernández-Ahumada, A. Garrido-Varo, J. E. Guerrero, D. Pérez-Marín, and T. Fearn. "Taking NIR Calibrations of Feed Compounds from the Laboratory to the Process: Calibration Transfer between Predispersive and Postdispersive Instruments". *J. Agricultural and Food Chemistry*. 2008. 56(21): 10135.
89. J. A. F. Pierna, P. Vermeulen, B. Lecler, V. Baeten, and P. Dardenne. "Calibration Transfer from Dispersive Instruments to Handheld Spectrometers". *Appl. Spectrosc.* 2010. 64(6): 644.
90. K. Kramer, G. Small. "Digital Filtering and Model Updating Methods for Improving the Robustness of Near-Infrared Multivariate Calibrations". *Appl. Spectrosc.* 2009. 63: 246.

## APPENDIX A

### ABBREVIATIONS

SSE: Sequentially Shifted Excitation  
FT-Raman: Fourier transform Raman  
Nd:YAG: Neodymium-doped yttrium aluminum garnet; Nd:Y<sub>3</sub>Al<sub>5</sub>O<sub>12</sub>  
CCD detector: Charge-coupled device detector  
SERDS: Shifted Excitation Raman Difference Spectroscopy  
DBR: Distributed Bragg Reflector  
GaAs laser: Gallium Arsenide laser  
TEC: Thermoelectric controller  
NA: Numerical Aperture  
VSSB: Virtual Standard Slope and Bias  
SVSSB: Segmented Virtual Standard Slope and Bias  
API: American Petroleum Institute  
ASTM: American Society for Testing Materials  
LOO-CV: Leave one out - cross validation  
MLR: Multiple linear regression  
NIR: Near-infrared  
NRL: Navy Research Laboratory  
PC: Principal component  
PCA: Principal component analysis  
PCR: Principal component regression  
PLS: Partial least squares  
RER: Range error ratio  
SEP: Standard error of prediction  
SEV: Standard error of validation  
SG: Savitsky-Golay  
SNV: Standard Normal Variate  
SPE: Specification  
TARDEC: Tank-Automotive Research, Development and Engineering Center



**APPENDIX B**  
**NIR Spectra of the Fifteen Solvents used in the Segmented Virtual Standard Calibration Transfer**

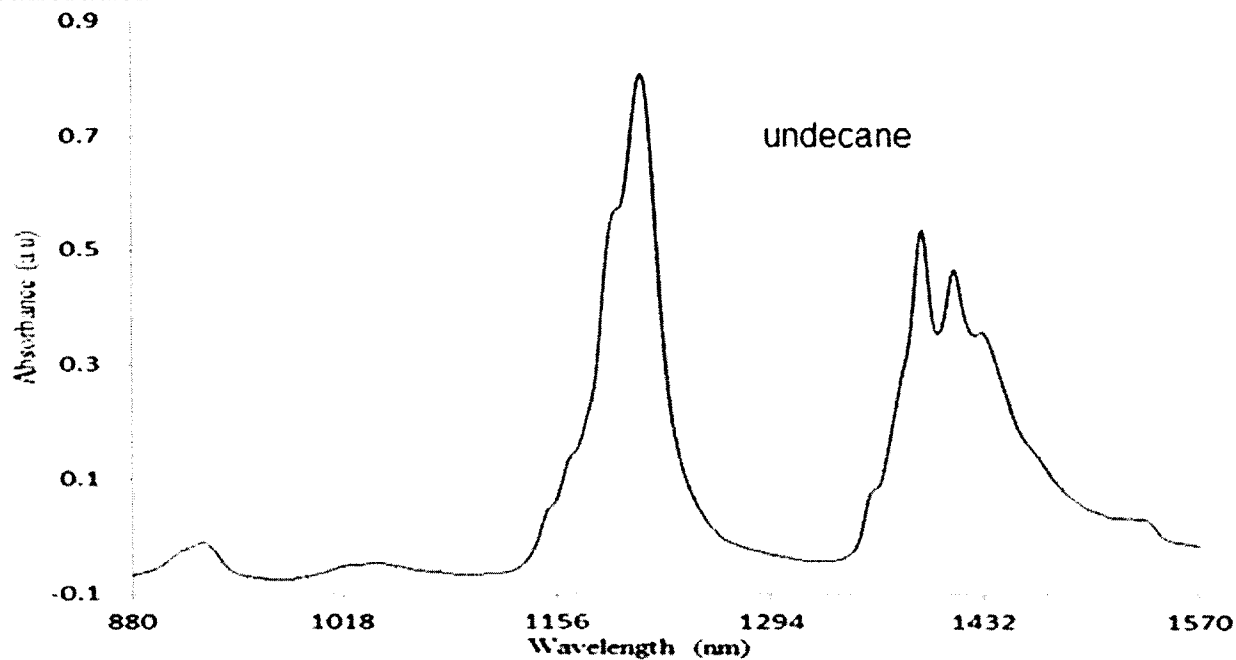


FIG. 79B. NIR spectrum of undecane ( $C_{11}H_{24}$ ), one of the fifteen molecular chemicals used in the PLS models calibration transfer of jet fuels properties.

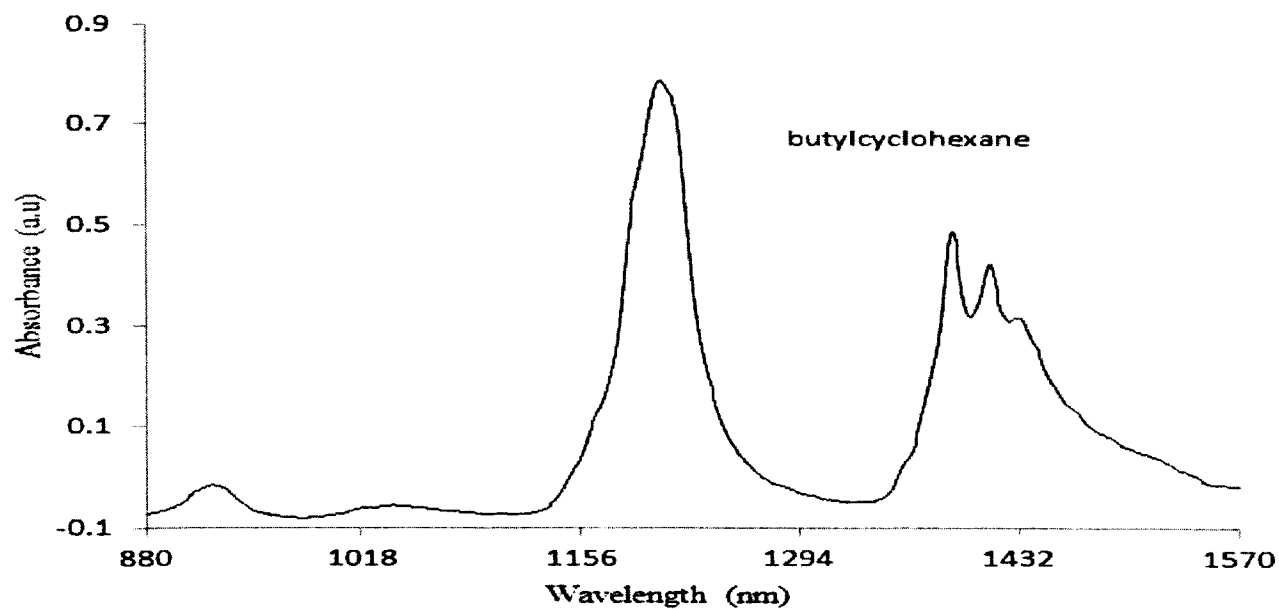


FIG. 80B. NIR spectrum of butylcyclohexane ( $C_{10}H_{20}$ ), one of the fifteen molecular chemicals used in the PLS models calibration transfer of jet fuels properties.

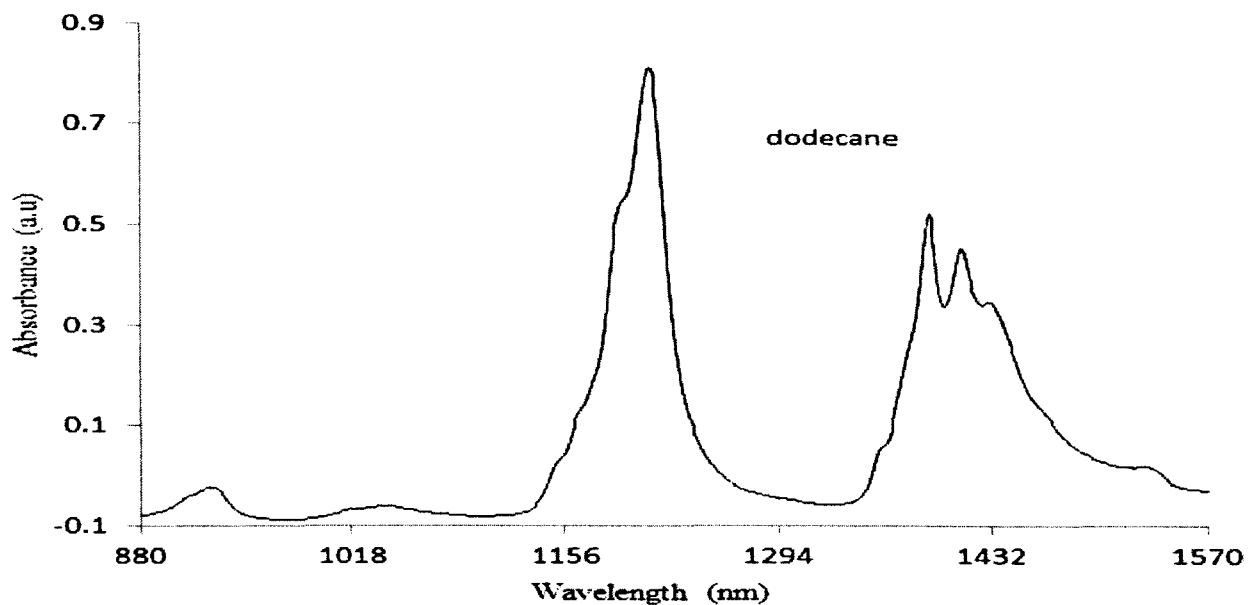


FIG. 81B. NIR spectrum of dodecane ( $C_{12}H_{26}$ ), one of the fifteen molecular chemicals used in the PLS models calibration transfer of jet fuels properties.

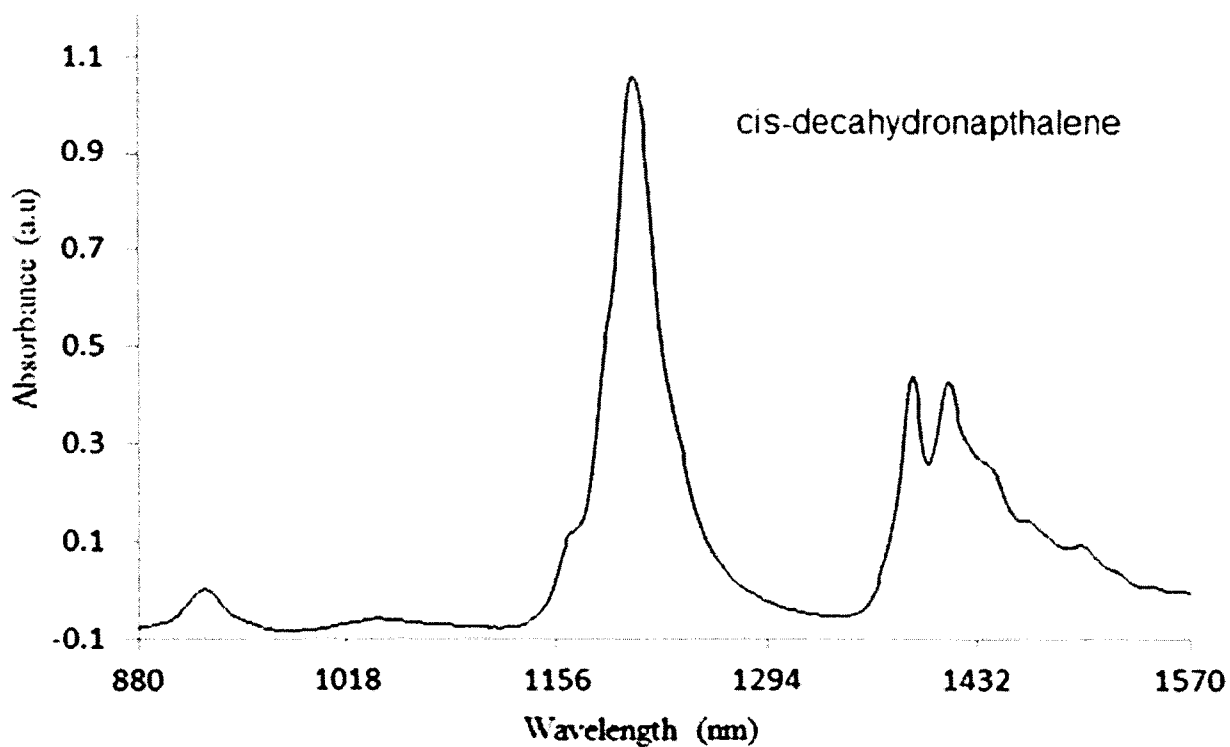


FIG. 82. NIR spectrum of cis-decahydronaphthalene ( $C_{10}H_{18}$ ), one of the fifteen molecular chemicals used in the PLS models calibration transfer of jet fuels properties.

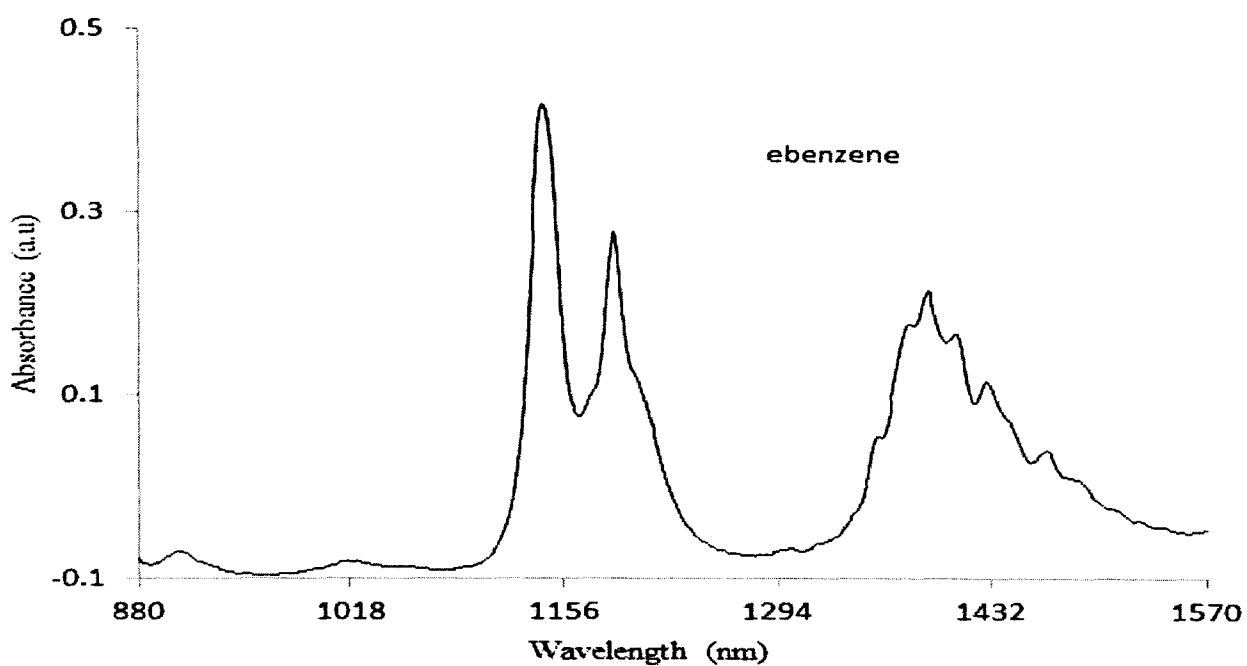


FIG. 83B. NIR spectrum of ethylbenzene ( $C_8H_{10}$ ), one of the fifteen molecular chemicals used in the PLS models calibration transfer of jet fuels properties.

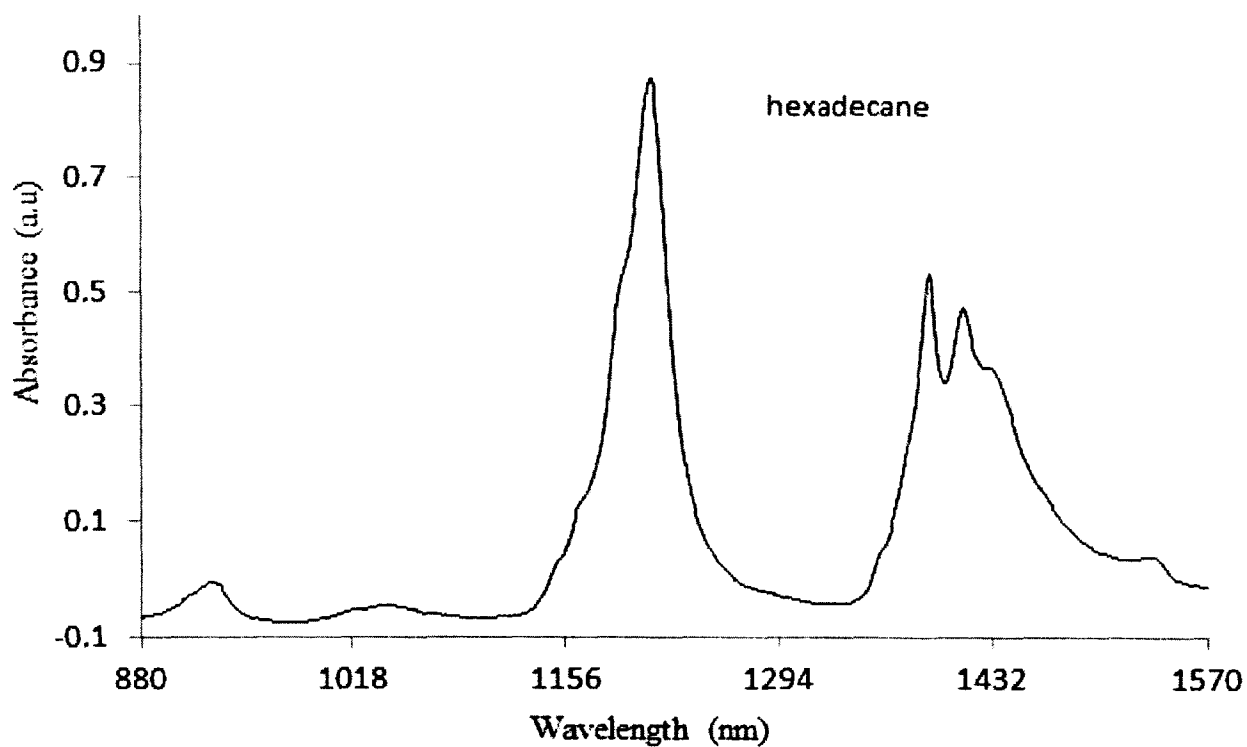


FIG. 84B. NIR spectrum of hexadecane ( $C_{16}H_{34}$ ), one of the fifteen molecular chemicals used in the PLS models calibration transfer of jet fuels properties.

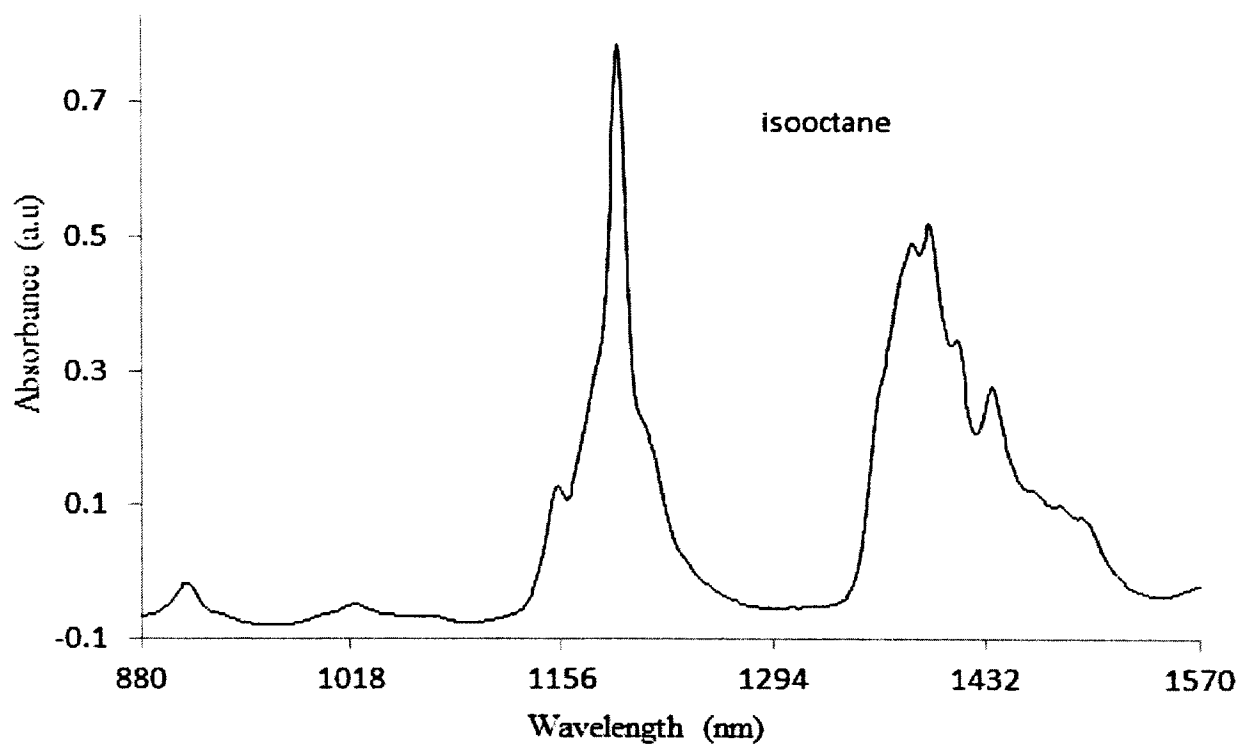


FIG. 85B. NIR spectrum of isooctane ( $C_8H_{18}$ ), one of the fifteen molecular chemicals used in the PLS models calibration transfer of jet fuels properties.

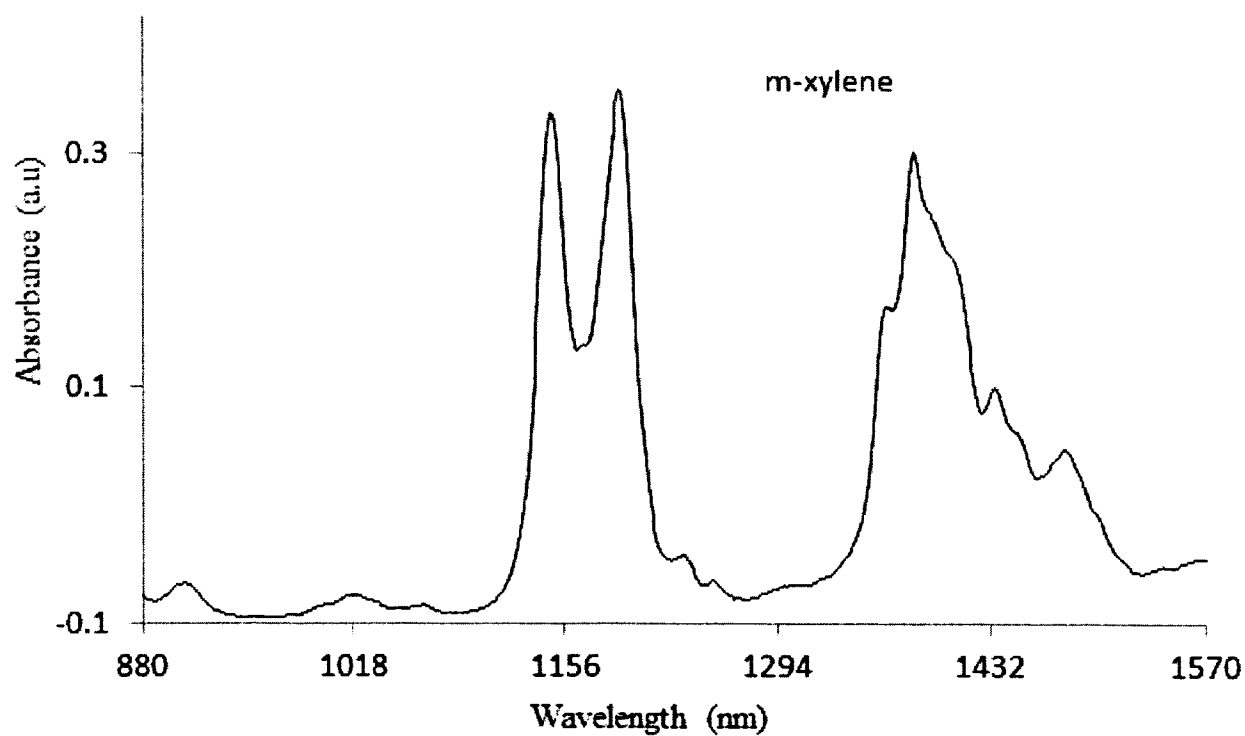


FIG. 86B. NIR spectrum of meta-xylene ( $C_8H_{10}$ ), one of the fifteen molecular chemicals used in the PLS models calibration transfer of jet fuels properties.

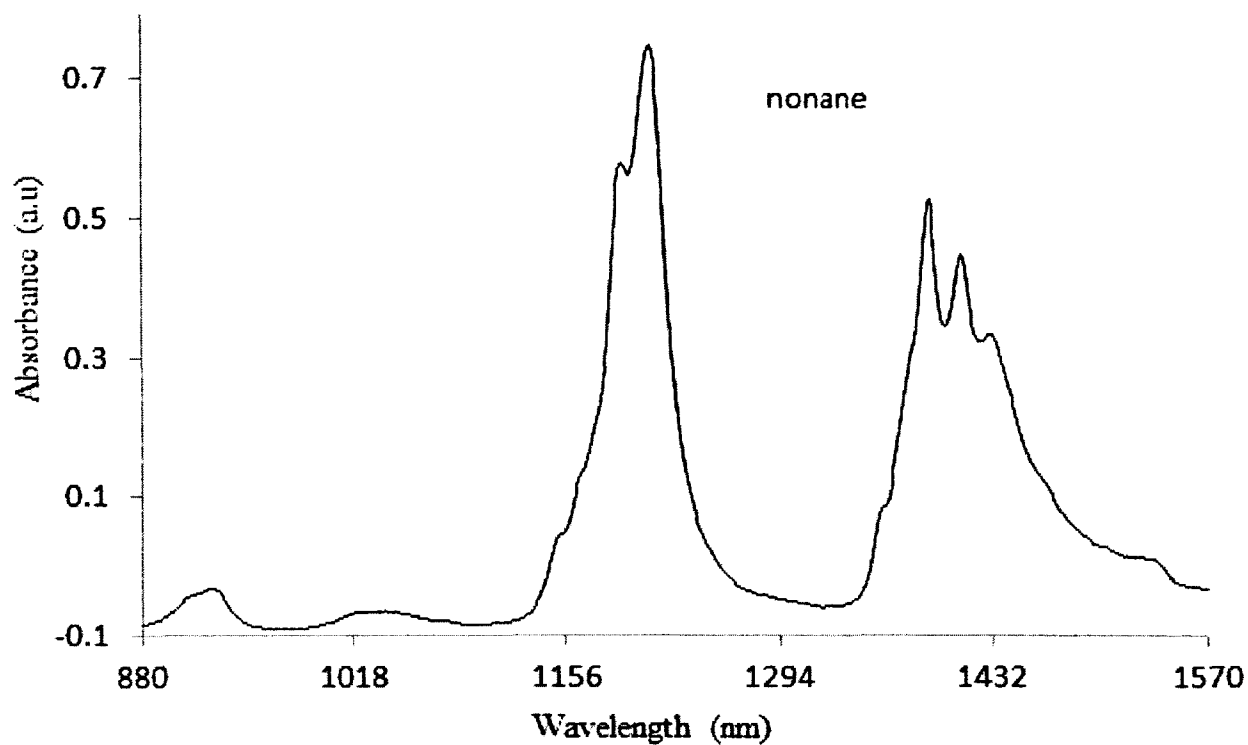


FIG. 87B. NIR spectrum of n-nonane ( $C_9H_{20}$ ), one of the fifteen molecular chemicals used in the PLS models calibration transfer of jet fuels properties.

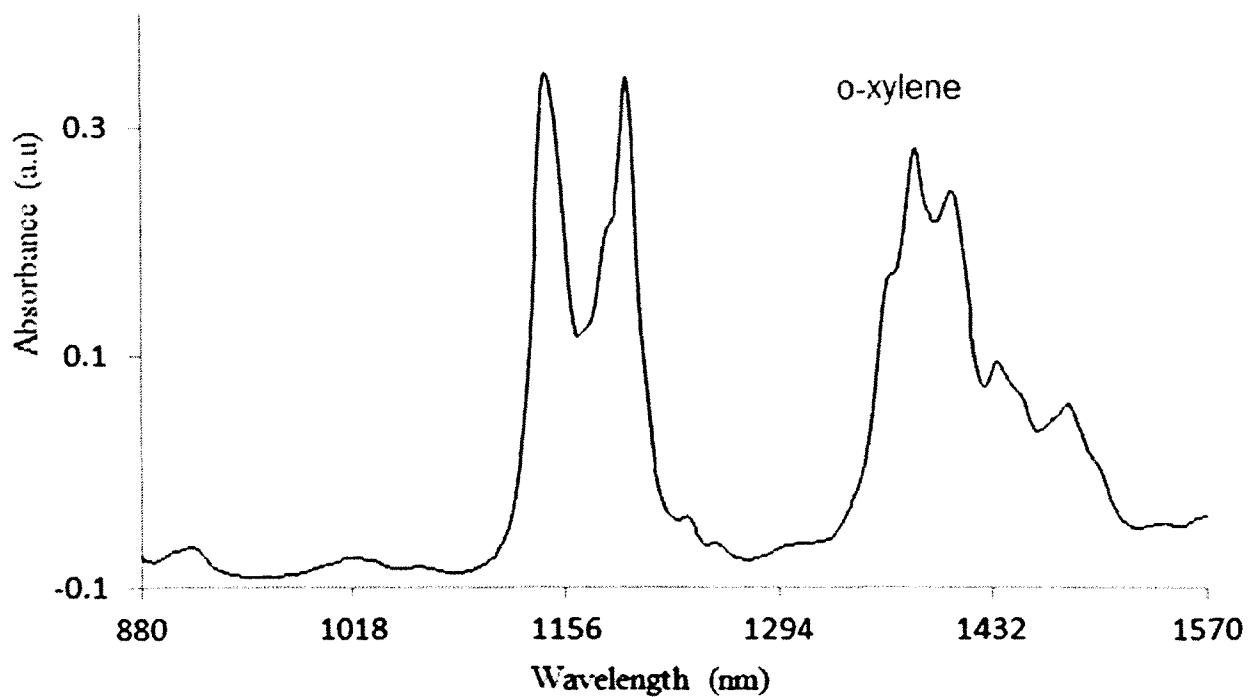


FIG. 88B. NIR spectrum of ortho-xylene ( $C_8H_{10}$ ), one of the fifteen molecular chemicals used in the PLS models calibration transfer of jet fuels properties.

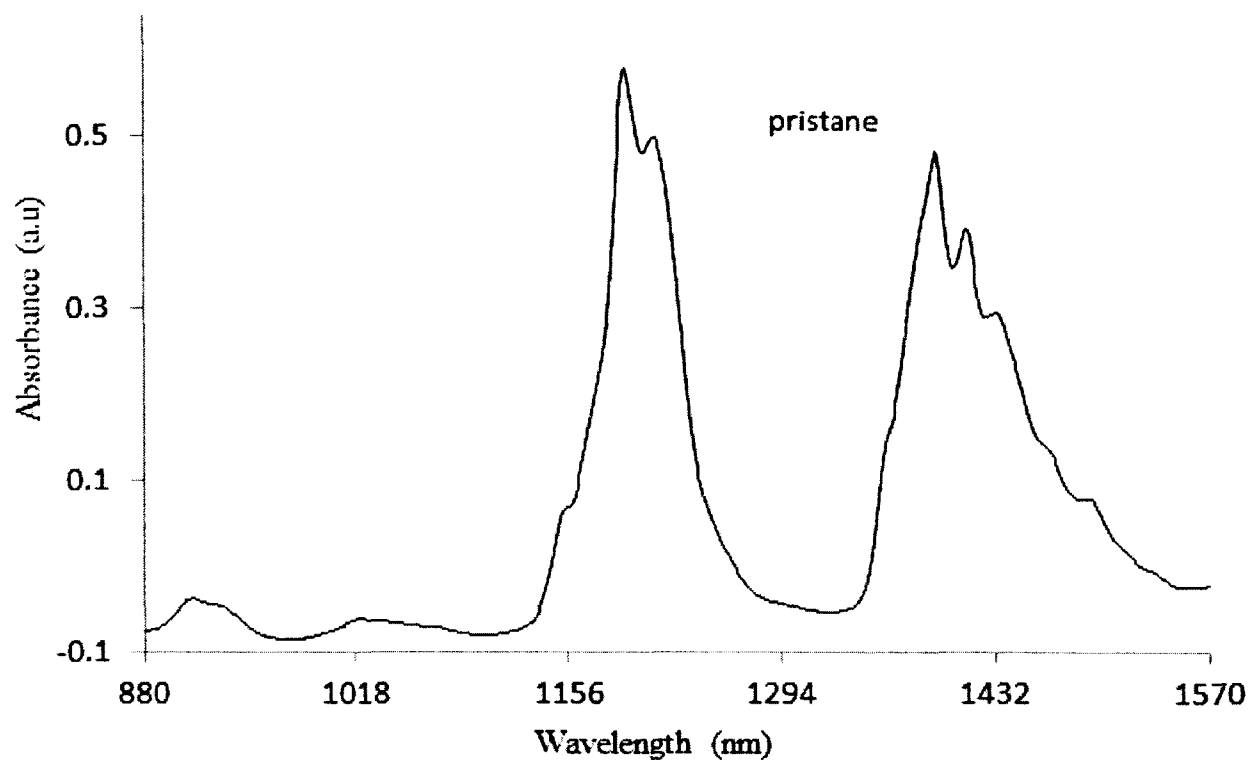


FIG. 89B. NIR spectrum of pristane (C<sub>19</sub>H<sub>40</sub>), one of the fifteen molecular chemicals used in the PLS models calibration transfer of jet fuels properties.

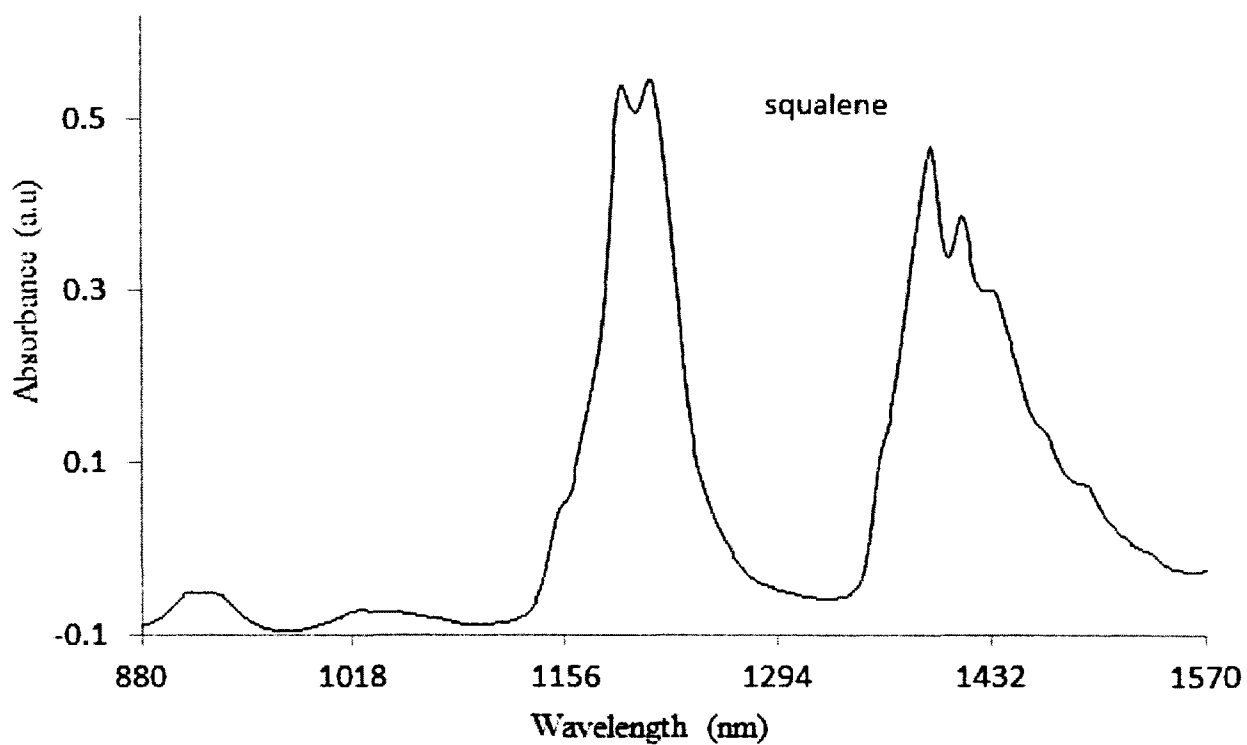


FIG. 90B. NIR spectrum of squalene (C<sub>30</sub>H<sub>62</sub>), one of the fifteen molecular chemicals used in the PLS models calibration transfer of jet fuels properties.

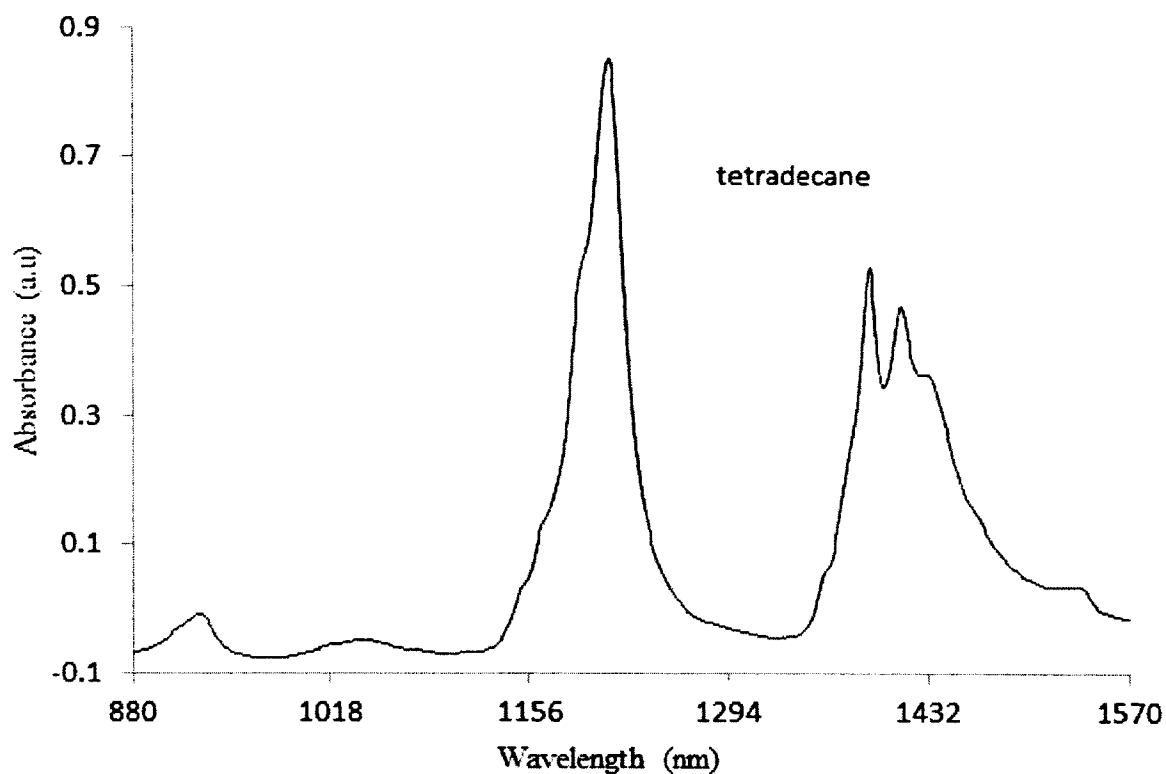


FIG. 91B. NIR spectrum of tetradecane ( $C_{14}H_{30}$ ), one of the fifteen molecular chemicals used in the PLS models calibration transfer of jet fuels properties.

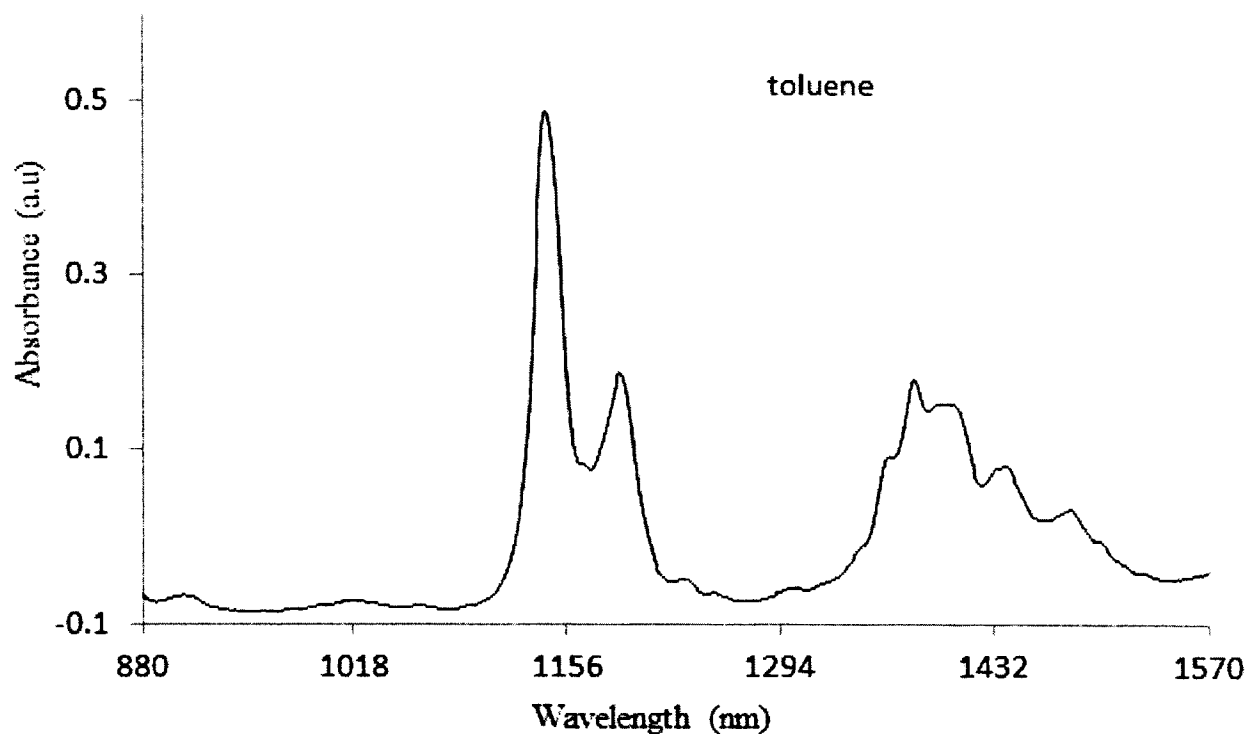


FIG. 92B. NIR spectrum of toluene ( $C_7H_8$ ), one of the fifteen molecular chemicals used in the PLS models calibration transfer of jet fuels properties.

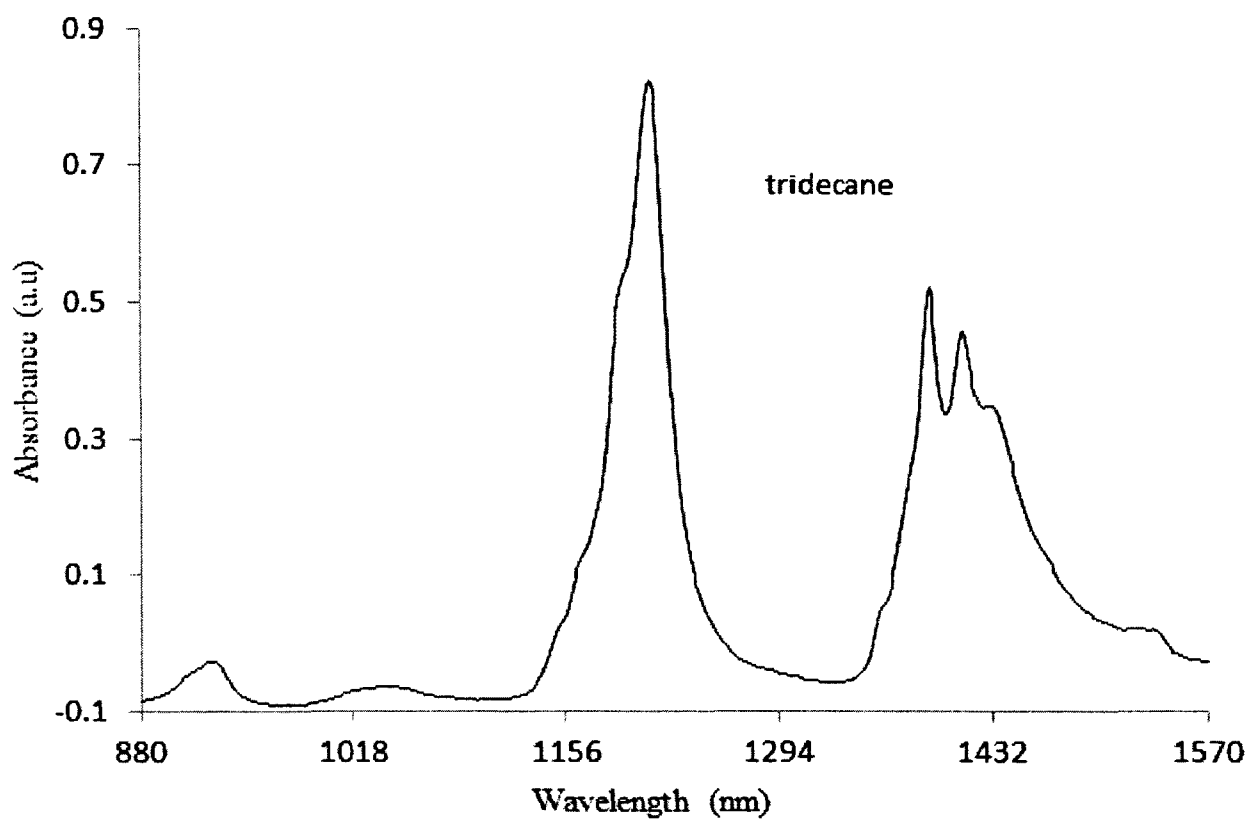


FIG. 93B. NIR spectrum of tridecane ( $C_{13}H_{28}$ ), one of the fifteen molecular chemicals used in the PLS models calibration transfer of jet fuels properties using SVSSB.



**VITA****Mohamed F. Abdelkader**

*Department of Chemistry and Biochemistry  
Old Dominion University  
Norfolk, VA 23529  
Email: mhdifarag7@gmail.com*

**Education**

OLD DOMINION UNIVERSITY, NORFOLK, VA  
PhD in chemistry, May 2013

ALEXANDRIA UNIVERSITY, ALEXANDRIA, EGYPT  
Bachelor of Science Degree in chemistry, 2007

**Publications**

- 1 Cooper, J. B., Abdelkader, M. F., Wise, K. "method and apparatus for acquiring Raman spectra without background interferences". U.S patent
- 2 Cooper, J. B., Abdelkader, M. F., Wise, K. "Sequentially shifted excitation Raman spectroscopy: novel algorithm and instrumentation for fluorescence-free Raman spectroscopy in spectral space" (ms no. 12-06852). J of applied spec. (in press)
- 3 Abdelkader, M. F., Cooper, J. B. & Larkin, C. M. "Calibration Transfer; theory and application". Chapter 7. Calibration transfer using virtual standard. Nova publication 2013. Editor: Ikumastu Fujimoto
- 4 Abdelkader, M. F., Cooper, J. B. & Larkin, C. M. Calibration transfer of partial least squares jet fuel property models using a segmented virtual standards slope-bias correction method. Chemometr. Intell. Lab. Sys 110, 64-73, doi:10.1016/j.chemolab.2011.09.014 (2012).
- 5 Cooper, J. B., Larkin, C. M. & Abdelkader, M. F. Calibration transfer of near-IR partial least squares property models of fuels using virtual standards. J. Chemom. 25, 496-505, doi:10.1002/cem.1395 (2011).
- 6 Cooper, J. B., Larkin, C. M. & Abdelkader, M. F. Virtual standard slope and bias calibration transfer of partial least squares jet fuel property models to multiple near infrared spectroscopy instruments. Journal Near Infrared Spectroscopy 19, 139-150, doi: 10.1255/jnirs.922 (2011).