Summer 2012

Terahertz frequency analysis of gaseous and solid samples using terahertz time-domain spectroscopy

Ryan Michael Smith
University of Iowa

TERAHERTZ FREQUENCY ANALYSIS OF GASEOUS AND SOLID SAMPLES USING TERAHERTZ TIME-DOMAIN SPECTROSCOPY

by

Ryan Michael Smith

An Abstract

Of a thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Chemistry
in the Graduate College of
The University of Iowa

July 2012

Thesis Supervisor: Professor Mark A. Arnold

ABSTRACT

Developments in semiconductor and laser technologies have facilitated development of terahertz (THz)-frequency (~2-200 cm^{-1}) technologies. Results published in the literature as far back as the early 20th century demonstrate the utility of this frequency range for myriad applications, but the improved performance of modern THz technologies has renewed interest in THz-frequency analysis. Material presented in this dissertation focuses on three applications of terahertz time-domain spectroscopy (THz-TDS): quantitation of gas-phase molecular species, analysis of polymeric materials, and investigation of dental tissue/composite structures.

Gas phase species were quantified individually at concentrations ranging from several parts per million to several parts per thousand using various chemometric methods. Quantitative model robustness was evaluated by comparison of model precision, and partial least-squares (PLS) regression methods provided the highest precision. Species were quantified in mixtures using PLS with errors of prediction below the permissible exposure limits (PELs) set by the Occupational Safety and Health Administration. The effect of spectral broadening as a result of overall sample pressure was investigated, and species were analyzed in mixtures at various overall pressures. Errors of prediction were again near or below the PELs, demonstrating the utility of this method over the range of pressures common for atmospheric analyses. Chemical selectivity available in THz spectral features was evaluated and compared to selectivity over infrared frequencies.

Spectral parameters measured in the THz frequency range also provide insight into structural properties of polymeric materials. In some cases, spectral peaks may be used to identify the temperature at which phase changes occur within these materials. THz refractive index spectra were found to be a sensitive and non-destructive tool for identification of phase transition temperatures.

The time-resolved measurement of THz-TDS makes it particularly useful for rapid, non-destructive analysis of layered structures. Ordinarily, the strength of bonds between dental tissues and composite materials are evaluated in the laboratory using destructive failure analyses. Transparency of dental tissues and composite materials used for restorative procedures to THz pulses allows investigation of interfaces between these materials. Refractive index spectra indicate locations in which delamination has occurred between bonded layers. These results provide an overview of unique capabilities of the THz-TDS method in real-life spectral analyses.

Abstract Approved: _____
Thesis Supervisor

_____
Title and Department

_____
Date

TERAHERTZ FREQUENCY ANALYSIS OF GASEOUS AND SOLID SAMPLES USING TERAHERTZ TIME-DOMAIN SPECTROSCOPY

by

Ryan Michael Smith

A thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Chemistry
in the Graduate College of
The University of Iowa

July 2012

Thesis Supervisor: Professor Mark A. Arnold

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

_____

PH.D. THESIS

_____

This is to certify that the Ph.D. thesis of

Ryan Michael Smith

has been approved by the Examining Committee
for the thesis requirement for the Doctor of Philosophy
degree in Chemistry at the July 2012 graduation.

Thesis Committee: _____
Mark A. Arnold, Thesis Supervisor

_____
Thomas F. Boggess

_____
Amanda J. Haes

_____
Leonard R. MacGillivray

_____
Gary W. Small

To my wife Audrey, the love of my life. You bring happiness to me every day.

ACKNOWLEDGMENTS

First and foremost, I owe everything I have to God. Every day I live and every breath I take is a precious gift from Him. He has given me strength to overcome obstacles I never thought surmountable and has blessed me in ways I could never count.

I would like to thank Dr. Arnold for his support and guidance throughout my tenure at the University of Iowa. My experience in your research group has been wonderful and has prepared me to tackle analytical problems in the real world. Your work ethic and personal drive are inspiring, and I continually aspire to be a scientist of your caliber.

My wife and family are owed a great deal of thanks for their love and support throughout my life and graduate career. Thank you for continually encouraging me in tough times and always pushing me to excel in what I do. I will always love you all!

I would also like to thank the staff of the Department of Chemistry and Department of Physics and Astronomy technical shops. Your work in fabrication of custom equipment has enabled me to perform complex and specialized experiments with ease, and the time you all devoted to helping me understand the importance of technical design nuances is greatly appreciated!

Credit is also due to the personnel of TeraView Ltd. of Cambridge, UK, particularly Rebecca Goodall and Philip Taday. Their efforts to restore the TPS Spectra 1000D following the flood of 2008 are greatly appreciated, and are in part responsible for my timely completion of graduation requirements. Over the course of my graduate studies, these two have also helped me to understand the theory behind, and use of, terahertz spectroscopic instrumentation.

I would like to thank the faculty of the Department of Chemistry and the members of the Arnold Research Group. You are a collection of truly remarkable people with great talents and vision. I wish you all the best in your research and careers.

ABSTRACT

Developments in semiconductor and laser technologies have facilitated development of terahertz (THz)-frequency (~2-200 cm^{-1}) technologies. Results published in the literature as far back as the early 20^{th} century demonstrate the utility of this frequency range for myriad applications, but the improved performance of modern THz technologies has renewed interest in THz-frequency analysis. Material presented in this dissertation focuses on three applications of terahertz time-domain spectroscopy (THz-TDS): quantitation of gas-phase molecular species, analysis of polymeric materials, and investigation of dental tissue/composite structures.

Gas phase species were quantified individually at concentrations ranging from several parts per million to several parts per thousand using various chemometric methods. Quantitative model robustness was evaluated by comparison of model precision, and partial least-squares (PLS) regression methods provided the highest precision. Species were quantified in mixtures using PLS with errors of prediction below the permissible exposure limits (PELs) set by the Occupational Safety and Health Administration. The effect of spectral broadening as a result of overall sample pressure was investigated, and species were analyzed in mixtures at various overall pressures. Errors of prediction were again near or below the PELs, demonstrating the utility of this method over the range of pressures common for atmospheric analyses. Chemical selectivity available in THz spectral features was evaluated and compared to selectivity over infrared frequencies.

Spectral parameters measured in the THz frequency range also provide insight into structural properties of polymeric materials. In some cases, spectral peaks may be used to identify the temperature at which phase changes occur within these materials. THz refractive index spectra were found to be a sensitive and non-destructive tool for identification of phase transition temperatures.

The time-resolved measurement of THz-TDS makes it particularly useful for rapid, non-destructive analysis of layered structures. Ordinarily, the strength of bonds between dental tissues and composite materials are evaluated in the laboratory using destructive failure analyses. Transparency of dental tissues and composite materials used for restorative procedures to THz pulses allows investigation of interfaces between these materials. Refractive index spectra indicate locations in which delamination has occurred between bonded layers. These results provide an overview of unique capabilities of the THz-TDS method in real-life spectral analyses.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

LIST OF ABBREVIATIONS

A.U.	Absorbance unit
	Arbitrary unit
ATR	Attenuated total reflectance
BCC	Basal cell carcinoma
BWO	Backward-wave oscillator
CLS	Classical least-squares
CVSEP	Cross-validation standard error of prediction
CW	Continuous-wave
DFF	Digital Fourier filtering
DOAS	Differential optical absorption spectroscopy
EO	Electro-optic
EPA	Environmental Protection Agency
eV	Electron-volts
FET	Field-effect transistor
FFT	Fast Fourier transform
FID	free-induction decay
F-P	Fabry-Perot
fs	femtosecond(s)
FT	Fourier Transform
FTIR	Fourier Transform infrared
FTMW	Fourier Transform Microwave
FWHM	Full-width at half-maximum
GHz	Gigahertz
GVD	Group velocity dispersion
HWHM	Half-width at half-maximum

IR	Infrared
JPL	Jet Propulsion Laboratories
KLM	Kerr Lens Modelocking
LBO	Lithium borate frequency-doubling crystal
LED	Light emitting diode
LMA	Levenberg-Marquardt algorithm
LOD	Limit of detection
LPC	Linear predictive coding
LV	Latent variable
MMS	Mohs' micrographic surgery
NASA	National Aeronautics and Space Administration
NAS	Net analyte signal
NDM	Negative dispersion mirror
NIR	Near infrared
NMR	Nuclear magnetic resonance
NMS	Nelder-Mead simplex algorithm
OSHA	Occupational Safety and Health Administration
PC	Principal component or photoconductive
PCA	Principal component analysis
PCSA	Pure component selectivity analysis
PEL	Permissible exposure level
PLS	Partial least-squares
PNNL	Pacific Northwest National Laboratory
ppm	Parts-per-million
ppt	Parts-per-thousand
POM	Polyoxymethylene
QCL	Quantum cascade laser

RD-SOS	Radiation-damaged silicon-on-sapphire
RMS	Root mean square
SCA	Shift-correction algorithm
SEL	Multivariate selectivity
SNR	Signal-to-noise ratio
SVD	Singular value decomposition
THz	Terahertz
Ti:sapphire	Titanium-doped sapphire
TDS	Time-domain signal
	Time-domain spectroscopy
TMR	Transmission microradiography
TPI	Terahertz pulse imaging
UHMW	Ultra-high molecular weight
UV/Vis	Ultra-violet/visible
VPUF	Verdi-pumped ultrafast laser

CHAPTER I

INTRODUCTION

Background

Advances in the methods used to generate and detect terahertz (THz) frequency radiation have fueled the expansion of THz spectroscopy in recent years.[14] While there is agreement that the THz spectral region refers to frequencies between the infrared and microwave (MW) regions of the electromagnetic spectrum, the exact frequencies of this spectral range are not rigorously defined. Different ranges are specified in both the peer-reviewed literature[15-17] and documentation from instrument vendors.[18, 19] In general, THz frequencies are considered to extend from several hundred gigahertz (GHz) up to several THz, and these frequencies provide information related to crystalline vibrations in solids, intermolecular interactions in liquids, and rotational transitions in gases. Various characteristics of the spectral features and properties of the measurement have made this method particularly attractive for analysis of both gaseous and solid samples, each of which will be described in more detail in their respective chapters. For example, the rotational transitions probed with THz frequencies provide information related to the composition and molecular structure of absorbing species, as opposed to the functional group-related information obtained with vibrational spectroscopic methods. THz radiation can also penetrate materials that are opaque to other spectral frequencies such as polymers,[20] paper,[21, 22] cloth,[23] leather,[24] and dry food stuffs,[25] opening a window for noninvasive analysis and detection through these materials. These unique transmission properties as well as the non-ionizing nature of THz radiation make THz technology ideally suited for applications in various fields such as polymer manufacturing, security and screening, and quality assurance.[20]

Since the development of interferometric spectroscopy in the 1950's[26] and before the advent of modern THz sources and detectors in the 2000's,[27] broadband THz spectra

were often collected by using either dispersive or Fourier-transform (FT) spectrometers equipped with a mercury arc lamp in combination with a bolometric detector. The low radiant powers of the mercury arc lamps coupled with the poor sensitivity of the bolometric detectors over THz frequencies resulted in poor signal-to-noise ratios and low quality spectra. Modern THz spectrometers take advantage of novel semiconductor sources and detectors that provide superior performance. A direct comparison of THz sources reported by Zhang and co-workers reveals a peak power of only 0.0001 mW with a dynamic range of 300 for a mercury arc lamp compared to a typical semiconductor THz source that provides 1 mW of peak power with a dynamic range greater than 10^8.[28] Likewise, modern semiconductor THz detectors are more than 1,000 times more sensitive than the liquid-helium-cooled bolometers used for FT measurements.[29] Elimination of cryogenic cooling is another benefit of the modern semiconductor THz technology. The greater radiant powers and higher detector sensitivity of semiconductor-based THz spectrometers provide superior spectral results, particularly over the spectral range of 0.3 to 4 THz (10-133 cm^{-1}). The exact frequency range available for spectral measurements depends on key instrumental parameters, such as duration and power of the femtosecond laser pulse, composition of the semiconductor material, and design of the emission and detection antennae. A more thorough history of the genesis of THz spectroscopy has been published elsewhere.[30]

Spectroscopic and Imaging Analysis

THz instrumentation and associated signal processing are under development for both spectroscopic and imaging applications. Spectroscopic applications involve the use of broadband sources to provide a measure of chemical information on the basis of position and width of frequency-selective absorption bands. Imaging applications, on the other hand, typically involve an analysis of radiation transmitted or reflected at interfaces

using radiation composed of narrow THz spectral bandwidths, or through analysis of the time-domain signal (TDS).

For spectroscopic applications, a broad band of THz frequencies is needed, particularly for the analysis of solid samples where the absorption bands can be broad, having half-widths at half-maximum (HWHM) in the range of 2-11.6 cm^{-1} at room temperature.[31] In modern instruments, a broadband pulse of THz radiation is produced from a semiconductor source, and this pulse propagates through the sample cell before detection. The electric field vector of the received radiation is monitored as a function of time to produce a time domain signal (TDS) that is analogous to a free induction decay signal in NMR spectroscopy or an interferogram in FT-IR spectroscopy. A Fourier transformation is then used to convert this TDS to a frequency domain spectrum. The corresponding instrumentation is commonly referred to as a terahertz time-domain spectrometer (THz-TDS). In the literature, and hence in this review chapter, the THz-TDS abbreviation can also refer to a terahertz time-domain spectrum or a terahertz time-domain signal depending on context.

Instrumentation used for imaging can be simple or complex depending on the application. Efforts to develop methodology for the detection of concealed objects or the assessment of the physical structure of an interface typically use a relatively narrow band of THz radiation and the corresponding instrument is designed to monitor this band of radiation as it is transmitted through, or reflected from, surfaces of different refractive indexes.[5]

More complex instrumentation is required for chemical imaging. Mainly, a broadband THz spectrometer is used in conjunction with a translation stage or an elaborate robotic arm designed to orient the sample in three dimensions. A full THz spectrum is collected at each pixel location across either a two-dimensional or three-dimensional surface of the object under investigation. The combination of chemical

selectivity from a broadband THz-TDS instrument and spatial resolution from an imaging system provide a powerful tool for many practical applications.

Instrumentation

Principal Elements and Concepts of Operation

Modern instrumentation for collecting a THz-TDS is analogous to a single-beam FT instrument in the sense that in both apparatus, the intensity of transmitted or reflected radiation is measured as a function of the position of a moving optical element. A Michelson interferometer is generally used in FTIR spectroscopy to modulate the detected radiation to produce the time domain spectrum, whereas an optical delay line is used to create the TDS for THz-TDS. The THz electromagnetic radiation is produced by an emitter and detected by a receiver. The resulting THz-TDS measurement corresponds to the real-time detection of the electric-field vector of the THz electromagnetic wave after it propagates through the sample compartment.

Both the generation and detection of THz radiation originates from an ultrafast, femtosecond (fs) laser pulse at a near-infrared (NIR) wavelength. This NIR laser pulse is split into a pump beam that is used to emit a band of THz radiation and a probe beam that is used for detection of the THz radiation. Either the pump or probe beam passes through a delay line that increases the optical path of this radiation. Various delay line configurations are reported, each with its respective relative merits.[32] Overall, changes in the position of the delay line system allow sampling of the time-domain signal in a discrete, time-gated manner. Lock-in amplification is used to enhance the signal-to-noise ratio of the measurement.

Improvements in both the delay line and femtosecond laser are of interest, as these units present bottlenecks in the maximum attainable scan rate, scan repeatability, signal-to-noise ratio, and instrument footprint. In both photoconductive (PC) and electro-optic (EO) detection schema, the portion of the terahertz pulse which arrives at the

detection unit congruently with the probe laser pulse is sampled. The sampling window duration is defined in part by the NIR laser pulse width, as detection occurs only when birefringence modulation (EO) or current flow (PC) are induced.

General Measurement Configuration

A generic diagram of the TDS apparatus is shown in Figure I-1. Here, the ultrafast NIR laser pulse is split into the pump and probe beams. The pump beam strikes a biased semiconductor, thereby emitting a broadband pulse of THz radiation. This radiation is guided to the sample compartment where it interacts with the sample. Subsequently, the transmitted THz radiation is guided to the detection module where the time-delayed probe beam permits gated detection. A Ti:Sapphire laser is commonly used to produce an 800 nm optical pulse over a period of less than 100 fs and a power on the order of ten to several hundred milliwatts. Various semiconductor materials, such as GaAs, InAs, radiation-damaged Si-on-sapphire (RD-SOS), amorphous Si, InAlAs, ion-implanted Ge, and ZnTe, are available for emitting and receiving the THz radiation, and each type of semiconductor exhibits a different frequency response.[33] Components fabricated with the same semiconductor material and produced by the same manufacturer can have slightly different emission or detection characteristics, likely the result of subtle differences in the fabrication process.

THz Pulse Generation

While generation of THz pulses from unbiased semiconductor materials has been demonstrated,[34] most THz-TDS measurements utilize an emitter device produced by microfabrication of antenna structures on the surface of a thin slice of semiconductor material.[35] An electrical bias is applied across the antenna structure, and the fs NIR laser strikes the material between the two antenna electrodes, as shown schematically in Figure I-2.

As the NIR laser pulse strikes the biased semiconductor material, a set of charged carriers (excited electrons and holes) are generated and these charged carriers are instantly accelerated within the biased electric field. The resulting acceleration of flow of electrical current induces the formation of a coherent band of electromagnetic radiation at THz frequencies across the poles of the emitter antenna.[36] Current flow across these emitters is exponentially proportional to the applied bias for low voltages, and the intensity of the emitted THz radiation follows a similar trend.[35] This THz pulse is typically several hundred femtoseconds in duration, and propagates through the sample and toward the detection system.

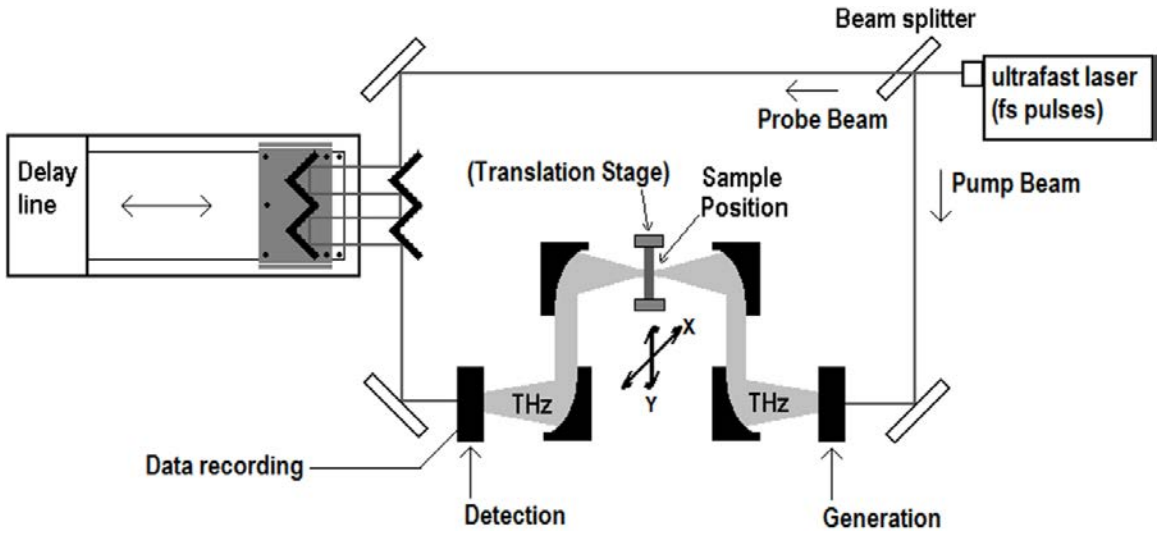


Figure I-1. Generic configuration of a THz-TDS apparatus for transmission measurements. A translation stage is depicted as an option that can be used for transmission-based imaging. The sample would otherwise be positioned in a stationary sample holder.

THz Detection

The two common modes of detection are the photo-conductive (PC) and electro-optic (EO) configurations. In the PC configuration, a semiconductor receiver antenna is

simultaneously illuminated by both the NIR probe beam and the THz radiation. As in the PC emitter device, each NIR laser pulse generates charge carriers between the antenna electrodes. The antenna structure is unbiased and current flow occurs via interaction of the THz electric field with the photo-generated charge carriers created by the NIR probe pulse. Current flow is measured across the antenna leads and is proportional to the intensity of the THz electric field in the sampled time window. The PC receiver device configuration is shown schematically in Figure I-3.

In the EO configuration, the THz radiation and polarized NIR probe pulse are simultaneously directed through an electro-optic crystal, and the transmitted NIR pulse travels through a quarter-wave plate and Wollaston polarizer, as illustrated in Figure I-4. The THz radiation alters the birefringence of the crystal, thereby altering the phase characteristics of the co-propagating NIR probe beam, and the degree of phase retardation imposed upon the transmitted NIR laser pulse is proportional to the intensity of the coincident THz pulse. Once the probe beam has propagated through the EO crystal, it passes through a quarter-wave plate, which converts the linearly-polarized NIR probe beam to circularly polarized radiation. Circularly polarized radiation can be thought of as a linear combination of two orthogonal linearly-polarized waves in which the relative amplitudes of the two components vary in a sinusoidal manner with time, where one of the two waves is 90°, ¼ λ, or ½ π radians out of phase with respect to the other polarization. The Wollaston prism separates the two orthogonal components of the polarized NIR probe radiation and diverts each component to one element of a pair of balanced photodiodes.[37] As a phase delay is created from modulation of the EO crystal by the THz radiation, the relative amplitudes of the two orthogonal polarization elements of the NIR probe pulse are altered, which serves as a measure of the magnitude of the THz radiant power.

The relative merits of the PC and EO detection configurations are summarized in the literature.[33] Basically, EO systems promise wider detection bandwidth, as the

bandwidth of the PC antenna is limited by the duration of the ultrafast laser pulse as well as by the lifetime of the photo-generated charge carriers within the semiconductor matrix.[38] EO systems are also simpler in implementation than their PC counterparts.[33] PC-based detection offers decreased sensitivity to laser power fluctuation since current flow in both emitter and receiver are proportional to the laser fluence. In practice, however, the impact of laser power fluctuations is mitigated for EO systems by modulation of the NIR laser with an acousto-optic modulator.[37] Instruments that use an EO detection scheme are more susceptible to interferences related to acoustic and mechanical noise because detection involves the discrimination of minute differences in polarization of the NIR probe beam.[37] More details for both detection modes are available elsewhere.[33]

Continuous Wave Components

Continuous wave (CW) THz radiation can also be used for both imaging and spectroscopic applications. CW measurements tend to have lower signal-to-noise ratios, however, and are most appropriate for imaging where a narrow band of radiation is acceptable. The simplicity and potential low cost of CW sources can be attractive for certain applications. Methods of CW generation include the use of molecular gas (such as CO_2), semiconductors, p-doped germanium, lead salt (such as PbSe), quantum cascade lasers (QCL's), diode-based MW devices, backward-wave oscillators (BWO's), synchrotrons, free-electron lasers, and photomixing devices.[25] CW detectors include photo diode-based systems, bolometers, Golay cells, and pyrometers.[25] A variety of instrumental configurations and measurement modes (TDS or CW) are described in the literature to meet a variety of novel analytical challenges.[25]

Signals in the Time and Frequency Domains

The THz-TDS is collected as a measure of the magnitude of the electric-vector of the broadband THz pulse as a function of time. The time delay of the measured signal

corresponds to the rate at which the THz radiation travels through the sample cell and reaches the detector optics and electronics. The timing of the measurement is controlled by the arrival of the NIR probe laser pulse as determined by the position of the delay line.

The measured TDS is characterized by a large intensity burst, referred to as the main peak, followed by decay in the amplitude of the detected signal. Prior to the main peak, the optical path of the probe beam is shorter than that of the pump beam because the velocity of the THz pulse is reduced by the refractive index of the material located within the sample compartment. The pre-main peak signal provides a measure of the background noise, as the THz pulse has not yet reached the detector. As the delay line is scanned, the path of the probe beam increases and the detected signal corresponds to the intensity of the THz radiation arriving at the detector at subsequent times.

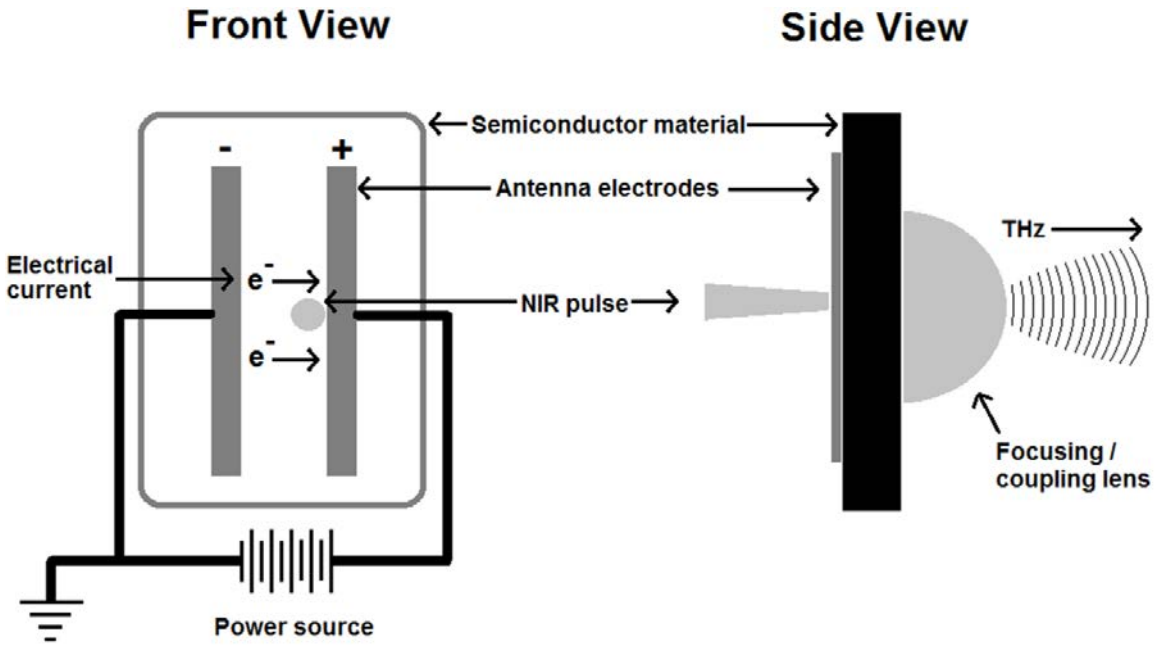


Figure I-2. Schematic diagram of a typical biased semiconductor THz emitter.

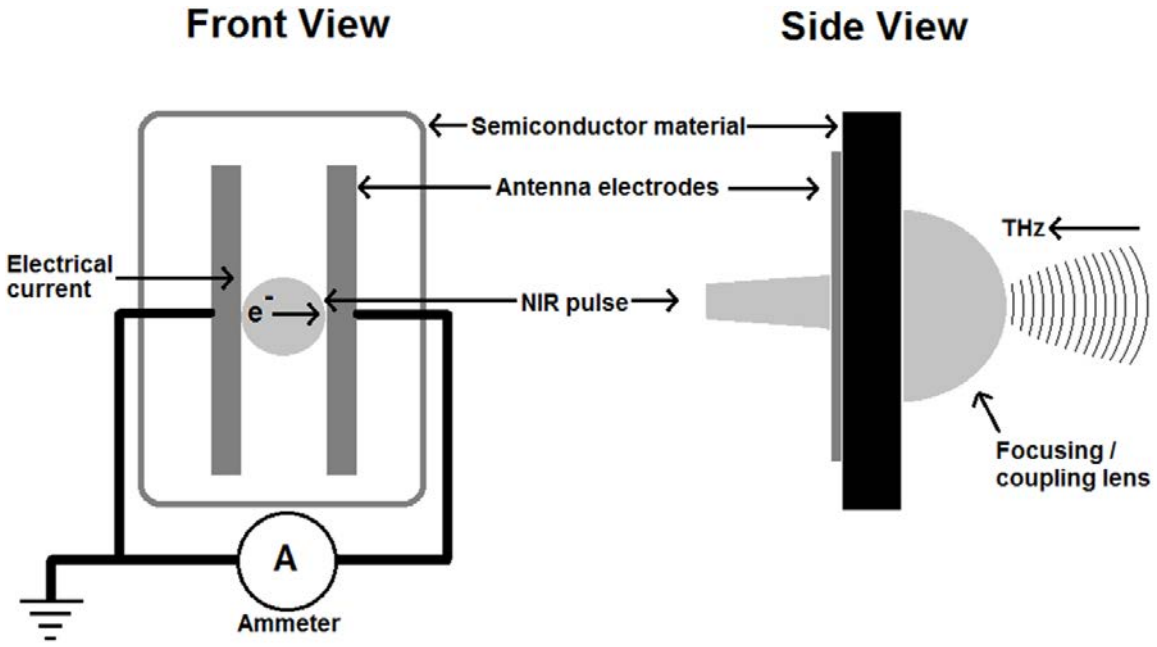


Figure I-3. Schematic representation of a PC antenna receiver.

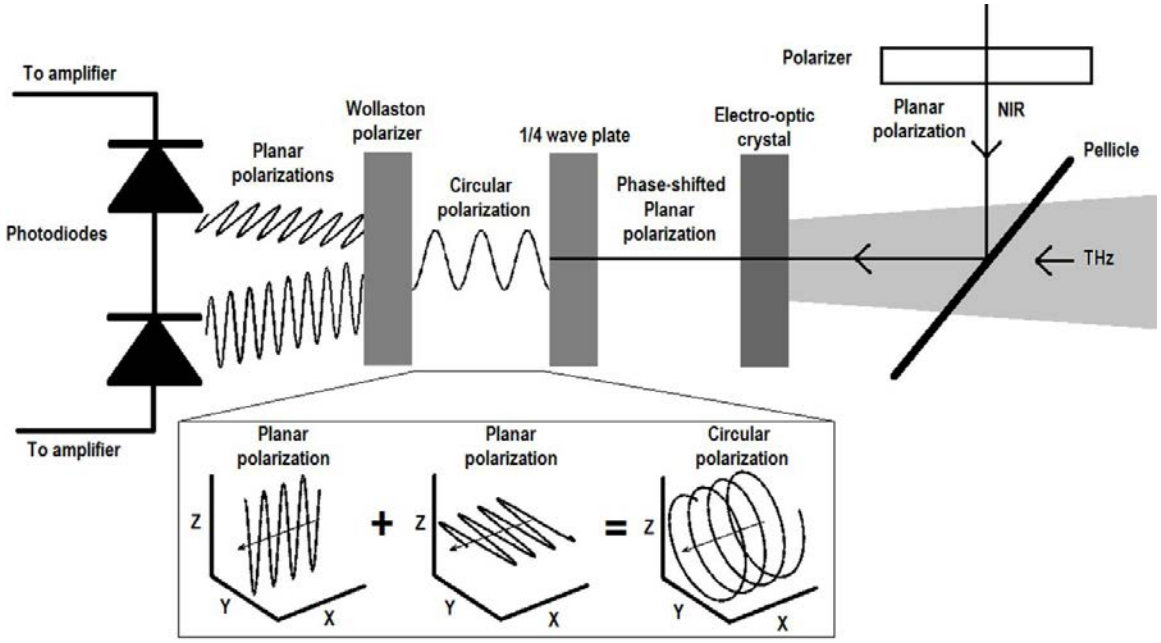


Figure I-4. Configuration of optical components for EO detection.

Examples of typical THz-TDS are provided in Figure I-5. As indicated in this figure, the THz-TDS is similar to an asymmetric interferogram from an FT-IR experiment or to the free-induction decay (FID) signal obtained during operation of an FT-NMR instrument. The THz-TDS plots in this figure illustrate typical spectra collected for solid samples when transmitting the incident radiation through a sample composed of the analyte dispersed within a pressed pellet of polyethylene (PE). The figure shows the reference and sample time-domain signals as well as the difference between the reference and sample waveforms. For the data in this figure, the reference signal was collected with a 200 mg pellet prepared with ultra-high molecular weight (UHMW) PE with a particle size range of 53-75 microns. The sample signal was collected with a 195 mg pellet comprised of the aforementioned PE powder mixed with crystalline β-lactose with a particle size less than 50 microns. The concentration of lactose in the mixture was 5.037 wt-%, and the pellets were prepared by compression in a 13 mm diameter die using 5 long tons (3,827 $kg \cdot cm^{-2}$) of pressure. The difference between the two signals represents the portion of the original signal absorbed by the crystalline lactose.

The THz-TDS measurement is based on pulses of coherent broadband radiation, so the signal can be thought of as a collection of cosine waves of a distribution of frequencies that have been superimposed, with constructive interference occurring at the main peak. As such, the reference and sample time-domain signals can be Fourier transformed (via the fast Fourier transform, or FFT, algorithm) to convert the data from the time domain to the frequency domain. The frequency-domain plot is referred to as a single-beam spectrum, since the reference and sample are measured independently, rather than simultaneously, as they could be measured in a dual beam spectrometer. The single-beam spectra for the reference and sample TDS's presented in Figure I-5 are shown in Figure I-6. The absorption spectrum of these signals is shown in Figure I-7. The final step in calculating the absorption spectrum ($A(\nu)$) is performed by taking the base-10

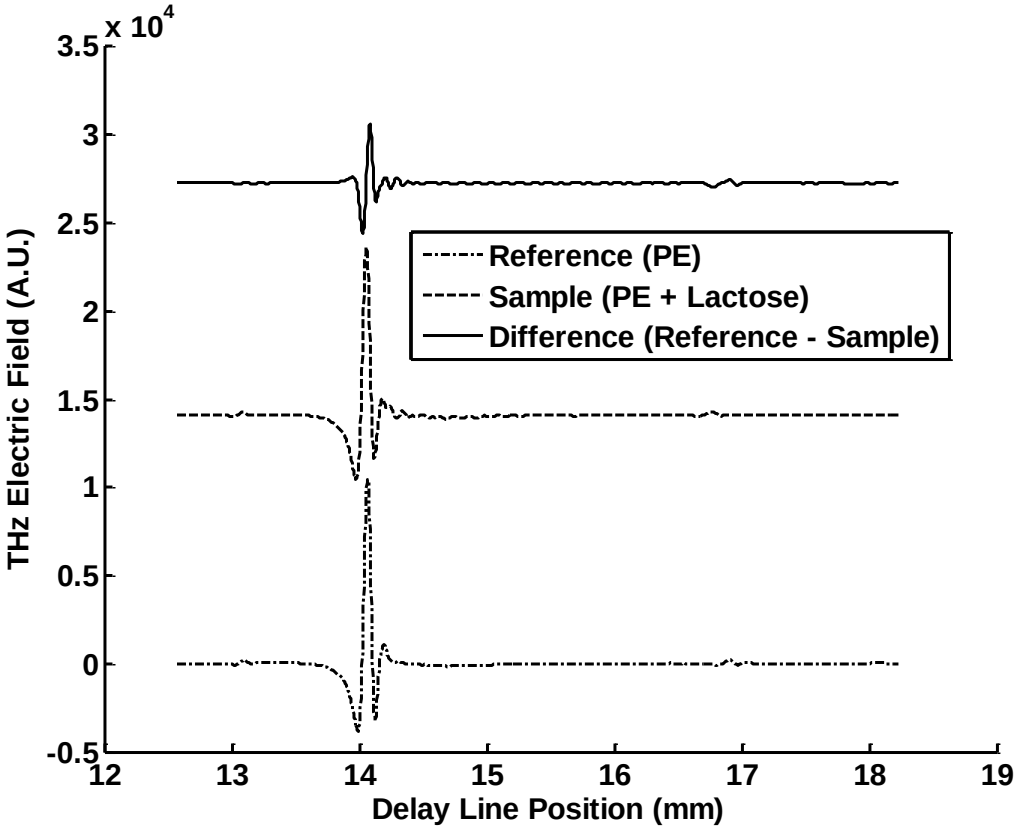


Figure I-5. Example of time domain signals (TDS) from THz-TDS measurements as well as the difference between the reference and sample TDS's. The reference signal was obtained from a 200 mg pellet of PE, the sample spectrum corresponds to a sample of 5.037 wt-% of β-lactose in PE, and the difference spectrum was calculated by subtracting the sample TDS from the reference TDS. Plots are offset vertically for comparison.

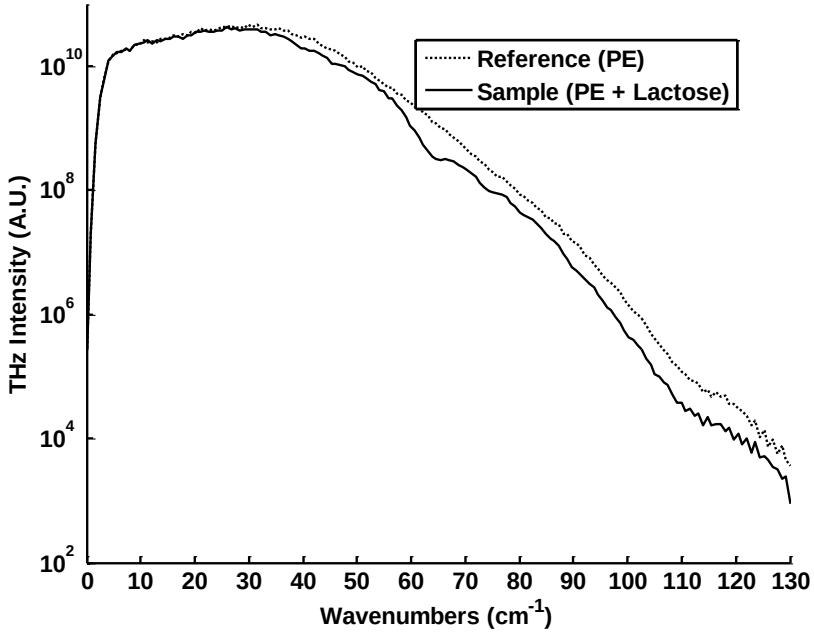


Figure I-6. Single-beam spectra of the PE reference (dashed) and PE/lactose sample (solid) described in the caption for Figure I-5.

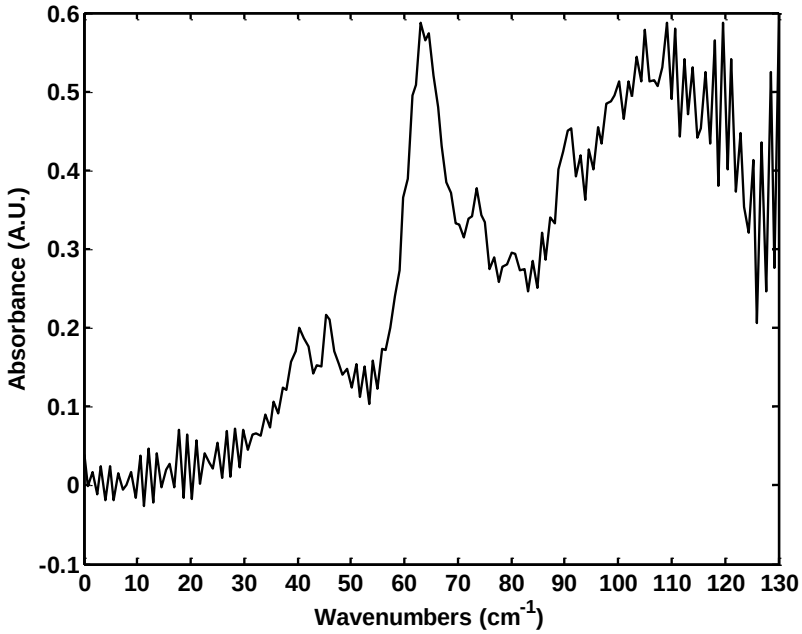


Figure I-7. Absorption spectrum of β-lactose generated from the single-beam transmission spectra presented in Figure I-6.

logarithm of the ratio of the sample single-beam spectrum ($I(\nu)_{samp}$) divided by the reference single-beam spectrum ($I(\nu)_{ref}$), as denoted in Equation 1:

$$A(\nu) = -\log_{10}\left[I(\nu)_{samp} \Big/ I(\nu)_{ref}\right] \qquad \text{Equation I-1}$$

The absorption spectrum calculated from the single-beam spectra presented in Figure I-6 is shown below. The observed absorption peaks correspond to crystal phonon vibrations of β-lactose. As with other FT spectroscopic methods, signal averaging can be performed with THz-TDS to enhance the measurement signal-to-noise ratio by decreasing random noise.

Signal Pre-Processing

Raw TDS's are typically processed prior to the FFT calculation in order to enhance the quality of the final spectrum. Two common pre-processing methods are apodization and zero-filling. An apodization step can be carried out with the time-domain signal immediately before transformation to the frequency domain. This step can be described as a weighting of the time-domain signal, and is performed using one of a variety of apodization or "windowing" functions. Application of the apodization function can improve spectral quality by excluding artifacts and reducing noise. Care is warranted, however, because extreme apodization can lead to distortion of the spectral features and loss of spectral information, especially when the instrumental resolution approaches the width of the analyte spectral features.

A variety of apodization functions are available to improve the final frequency domain spectrum. The simplest apodization function is the boxcar function where the raw data is essentially multiplied by unity so the FFT is performed on the raw, un-weighted time-domain signal. Other functions have a different multiplication pattern so different regions of the TDS are enhanced while other regions are attenuated in magnitude. These functions often have a value of 1 near the main peak and decrease

toward zero nearer the extremities of the sampled time window. Several examples of apodization functions include: triangular,[39] Happ-Genzel,[39] Hamming,[40] Norton-Beer weak medium and strong,[41] and Blackman-Harris 3-term and 4-term.[40] The "severity" of the apodization function can be described by the rapidity with which the function approaches its minimum value. Functions that approach zero nearer the main peak provide a smoother frequency-domain spectrum, but such functions have the propensity to diminish and broaden narrow absorption bands or high-frequency chemical information, the oscillations of which tend to be prominent at greater distances from the main peak in the TDS. Features of the spectrum are smoothed via this process, though the width of absorption peaks can also be increased. Asymmetry characteristic of the typical TDS affords another parameter for optimizing the apodization function. Functions can be applied in an asymmetric fashion, with each end reaching the same value, or symmetrically, where the function is calculated as though the pre-peak portion of the signal was the same length as the post-peak portion. Several sample asymmetric apodization functions are shown below in Figure I-8. The TDS in this image has been normalized to provide a consistent y-axis scale and allow visualization of the function features relative to features of the TDS. The order of strength of these apodization functions illustrated in Figure I-8 is Blackman-Harris 4-term > Blackman-Harris 3-term > boxcar.

The impact of the apodization functions can be seen in Figure I-9 which presented the resulting absorption spectra computed by applying each of the apodization functions presented in Figure I-8 to the TDS signals presented in Figure I-5. A comparison of the resulting absorption spectra highlights the reduction in high-frequency noise by applying a stronger apodization function to the TDS data.

A comparison of several apodization functions (including an optimized function) and application to THz-TDS has been reported by Dudley and co-workers.[42] In this paper, a process is described in which functions are applied to the time-domain signal, and the

resulting signal and frequency-domain spectra are compared to the original (unaltered) data sets. The performance of each function is evaluated on the basis of the root-mean-square (RMS) percent difference between the apodized and unapodized data. It is shown that the optimized function outperforms others, especially when the original data have been degraded significantly. In Dudley's work, the Mertz and Hamming apodization functions also enhance the THz spectral quality compared to a simple boxcar function (also known as a rectangular or Dirichlet apodization function), but to a lesser extent than their optimized function.

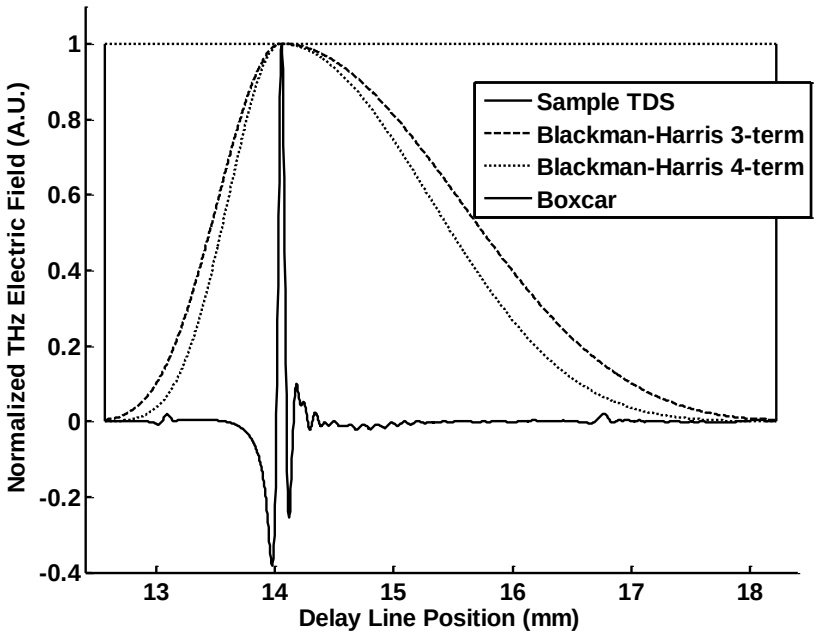


Figure I-8. The PE/lactose sample TDS is shown with several examples of apodization functions. Areas of the TDS falling under portions of the apodization function with high magnitude will be represented to a greater extent in the final absorption spectrum compared to those nearer the ends of the TDS where the apodization function is of lesser magnitude. Oscillations visible in the TDS from the main peak out to approximately 15.5 mm contain most of the spectral information of crystalline lactose. Note the appearance of a reflection artifact near 16.7 mm. An etalon "echo" is responsible for the appearance of this artifact, as described in the section titled "Analytical challenges for quantitative measurement of solids."

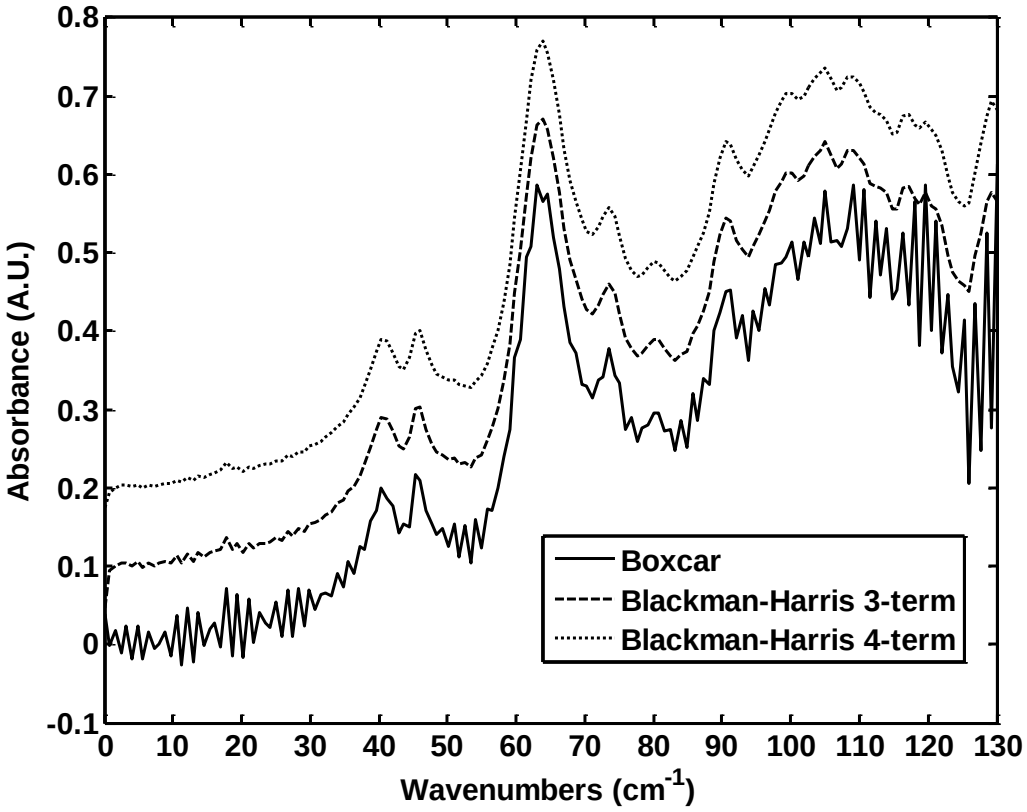


Figure I-9. Absorption spectra from the same PE/lactose sample calculated using several different apodization functions. Note the attenuation of baseline oscillations with an increase of the apodization strength.

Zero-filling, also known as "zero-padding", "Fourier interpolation" or "zooming," is a method that is commonly used to provide more spectral data points within a given spectral range. The additional data points do not contain chemically-specific information, but represent interpolated points between the actual, chemically-relevant points. The result is a smoothing of the spectral features within a spectrum. Zero-filling involves adding zeros to the end of the TDS prior to implementing the FFT calculation. The spectra presented in Figure I-10 show the impact of different levels of zero-filling on the principal absorption features in the absorption spectrum of β-lactose. The smoothing effect is evident as the number of points increases and the point spacing decreases. Still,

these additional points do not enhance the spectral resolution, as the amended portion of the signal is artificial and contains no true chemical information. Zero-filling can improve peak fitting by clarifying the shape of the spectral features that can be captured with the nominal resolution of the spectrometer.

The list below provides the experimental procedure used for the sample preparation, data collection, and data processing used to produce the signals and spectra above. Standard experimental parameters for each step are indicated in parenthesis.

(1) Sample preparation
   a. PE reference pellet (mass, compression pressure)
   b. PE/β-lactose sample pellet (mass, analyte concentration, compression pressure)

(2) Reference spectrum acquisition (resolution, signal averaging)

(3) Sample spectrum acquisition (resolution, signal averaging)

(4) Apodization (function selection)

(5) Zero-filling (levels of filling)

(6) Single-beam/absorption spectrum calculation

Measurement Modes

Transmission Spectroscopy

In many cases, spectroscopic characterization of a sample must precede imaging-based investigations, in order to establish the optical and spectroscopic properties of the material in question. Besides fundamental optical properties of the material, quantitative aspects of the sample are commonly important. Such quantitative information can be obtained from transmission spectroscopy where the concentrations of individual components are derived from an analysis based on the Beer-Lambert relationship presented in Equation I-2.

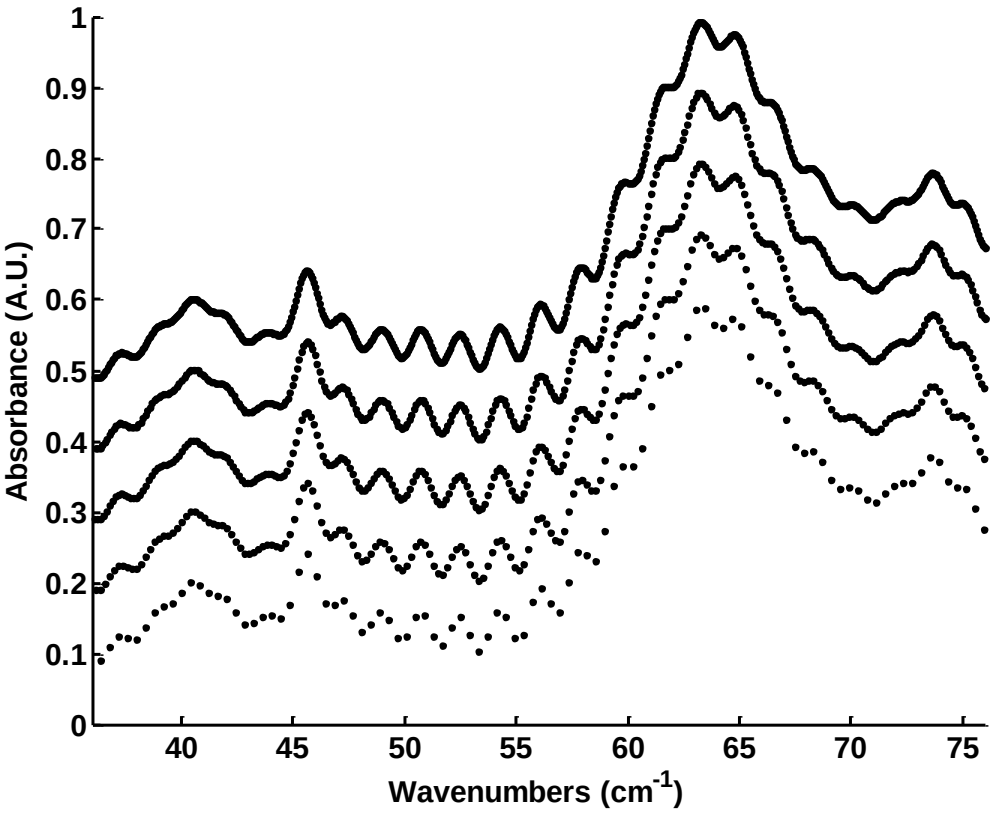


Figure I-10. Zero-filled absorption spectra from a sample of 5.037 wt-% β-lactose dispersed within a PE pellet. Spectra obtained from the analysis of zero-filled TDS extended to the nearest power of 2, where "N" is the next power of 2 from the length of the raw TDS, and higher orders of zero-filling are integer multiples of this value. The spectrum calculated with N data points is shown starting at ≈0 A.U., and 2×N, 4×N, 8×N, and 16×N are shown with vertical offset of 0.1 A.U. from this spectrum. As more levels of extension are added, more data points are present in the spectrum, providing a smoother appearance.

$$A(\nu) = \varepsilon(\nu)bc \qquad \text{Equation I-2}$$

where $A(\nu)$ is the absorption of radiation at a certain frequency ν, b is the optical path length, c is the concentration of the absorber, and $\varepsilon(\nu)$ is the molar absorptivity, in units of inverse path length and molar concentration (such as $cm^{-1} \cdot M^{-1}$). A linear relationship is observed experimentally between absorbance and concentration, but this linear relationship is ultimately limited at both low and high analyte concentrations. Below the limit of detection, absorption from the sample cannot be observed when the amount of

absorbed radiation is not significantly different compared to the background noise on the optical signal. The limit of detection depends on the sensitivity of the absorption signal (*e.g.*, $\varepsilon(\nu)b$ in Equation I-2), as well as the instrumental signal-to-noise ratio at the frequency of the spectroscopic measurement. Sensitivity is a crucial characteristic of analytical measurement which is determined by the change in measured signal as a response to changes in the analyte concentration. High sensitivity is obtained when the analytical signal changes greatly in response to a minute change in analyte concentration. Non-linearity can be observed at high analyte concentrations based upon inter-molecular interactions within the sample as well as by non-linearity in the detector response as transmitted radiation levels drop below the detectivity of the detector optics and electronics.

Equation I-2 is only appropriate when the analyte of interest is the sole absorber of the incident radiation. For cases where multiple absorbers are present, Equation I-2 must be expanded to include separate absorbance values from the individual components within the sample matrix. For example, if one were to simulate a pharmaceutical product containing lactose, glucose, and acetaminophen, the absorption would be expressed as presented below as Equation I-3:

$$A(\nu) = [\varepsilon(\nu)_{Ac}c_{Ac} + \varepsilon(\nu)_{La}c_{La} + \varepsilon(\nu)_{Gl}c_{Gl}]b \qquad \text{Equation I-3}$$

where $\varepsilon(\nu)_{Ac}$ and c_{Ac} are the absorptivity and concentration of acetaminophen, $\varepsilon(\nu)_{La}$ and c_{La} are the absorptivity and concentration of lactose, and $\varepsilon(\nu)_{Gl}$ and c_{Gl} are the absorptivity and concentration of glucose, respectively. In this case, b would represent the thickness of the tablet or pellet. The resulting absorption measured at any frequency is a sum of contributions from all species and this linear combination of absorbances assumes no chemical interactions between the sample components.

Chemical selectivity is required for accurate quantification within mixtures composed of multiple absorbing species. Chemical selectivity is a measure of the ability

to distinguish between two species based on differences in their spectral features. Selectivity thus relies upon the uniqueness of the position and shape of absorption features for each sample component. In simple cases where absorption bands are fully separated for each sample component, the degree of selectivity can be estimated by visual inspection of the corresponding frequency-domain spectra. When spectral features overlap, however, the degree of spectral uniqueness can be determined as the "net analyte signal" (NAS), which is a spectral vector that represents that portion of the analyte absorption spectrum that is orthogonal to all sources of spectral absorption within the sample matrix and background signal.[43] Equation I-4 provides the calculation of the NAS:

$$NAS_{k,un} = (I - R_{-k}R_{-k}^{+})r_k \qquad \text{Equation I-4}$$

where $NAS_{k,un}$ is the net analyte signal for species *k*, I is the identity matrix, R_{-k} is the response matrix (absorption spectra) without the spectrum of species *k*, r_k is the spectrum of species *k*, and the superscript "+" represents the Moore-Penrose pseudoinverse operation.

One measure of the selectivity of a measurement is to compare the length of the NAS spectral vector of the analyte to the length of the original spectral vector for the analyte of interest,[44] as indicated in Equation I-5:

$$SEL_{an} = \|NAS_{an}\|/\|r_{an}\| \qquad \text{Equation I-5}$$

where the Euclidian norm (indicated by the vertical bars) of the NAS vector (NAS_{an}) of the principal analyte species in the sample matrix (denoted with the subscript "an") is divided by the Euclidean norm of the pure analyte spectrum (r_{an}). In one extreme, the matrix contains no interfering species so there will be no spectral overlap with the pure component spectrum of the principal analyte. In this case, the NAS spectrum will be identical to the pure component spectrum, the length of the two vectors will be equivalent, and the ratio (SEL_{an}) will be unity. On the other extreme, there is 100%

overlap between the pure component spectrum of the analyte and components within the matrix. In this second case, no NAS spectrum exists, the length of the NAS vector is zero, and the ratio (SEL_{an}) is zero. For all situations in between, the selectivity ratio (SEL_{an}) ranges from 1.0 to 0. The theoretical basis and analytical utility of the NAS will be described in greater detail in Chapter III.

The THz region of the electromagnetic spectrum offers the potential for selectivity based on unique absorption features associated with inter-molecular interactions. As such, spectral features are a product of the entirety of the crystalline structure of the compound and the unique vibrational characteristics of this structure. For comparison, analytical measurements with vibrational spectroscopy over mid-infrared frequencies are based on absorptions associated with specific chemical bonds and molecular functional groups that can be common between sample constituents, thereby producing spectral overlap and adversely impacting measurement selectivity.

Transmission measurements from THz-TDS measurements can also be used to determine the refractive index of a sample. Such direct refractive index values are derived from the timing of the propagated THz radiation in combination with the coherency of the incident radiation.[45] Refractive index measurements are useful in characterizing materials and to quantify the heterogeneity of a given sample, assuming the species in the sample matrix have different THz refractive indices.[46] Moreover, the refractive index can be used to assess the dielectric constant for sample materials by squaring the frequency-dependent complex index of refraction.[47]

Attenuated Total Reflectance (ATR) Spectroscopy

Reflectance spectroscopic measurements are required in circumstances where the sample is strongly absorbing, highly scattering, thick, or present in limited quantities. In such cases, a reflectance measurement through an attenuated total reflection (ATR) module may be the best way to collect the desired spectral information. In an ATR

measurement, the sample is placed in contact with a crystal and the incident radiation is introduced into the sample through this crystal. Generally, the crystal is a high THz refractive index material, such as silicon or germanium, and the spectrum is collected through the evanescent field established at the interface between the sample and the crystal. A schematic of the instrumental configuration for such a measurement is provided in Figure I-11.

An important parameter for any ATR measurement is the penetration depth, which corresponds to the effective distance the radiation propagates into the sample material. This penetration depth defines the optical path length for analytical determinations, thereby impacting the overall sensitivity of the measurement. The penetration depth is given by Equation I-6:

$$d_{pen} = \lambda / \left[2\pi \left(n_{crystal}^2 \sin^2 \theta - n_{sample}^2 \right)^{1/2} \right] \qquad \text{Equation I-6}$$

where d_{pen} is the penetration depth in meters, λ is the wavelength in meters, $n_{crystal}^2$ is the square of the refractive index of the ATR crystal, θ is the angle of incidence (in degrees) of the THz radiation at the crystal/sample interface, and n_{sample}^2 is the square of the refractive index of the sample. For a measurement, the incident radiation is reflected at the sample-crystal interface and a portion of the incident wave propagates a short distance into the sample. The penetration depth depends on several parameters as noted in Equation I-6, including the wavelength of the propagating radiation and the angle of incidence at the sample-crystal interface. Figure I-12 shows this dependency for germanium and silicon crystals over the THz frequencies for different angles of incidence. The strong dependency on frequency of the THz radiation is evident by this plot with penetration depths of nearly 1 mm at the low frequencies and only 10 μm at high frequencies. Snell's law (Equation I-7) provides a guideline for samples that can be analyzed with an ATR geometry.

$$\sin \theta_2 = (n_1/n_2) \sin \theta_1 \qquad \text{Equation I-7}$$

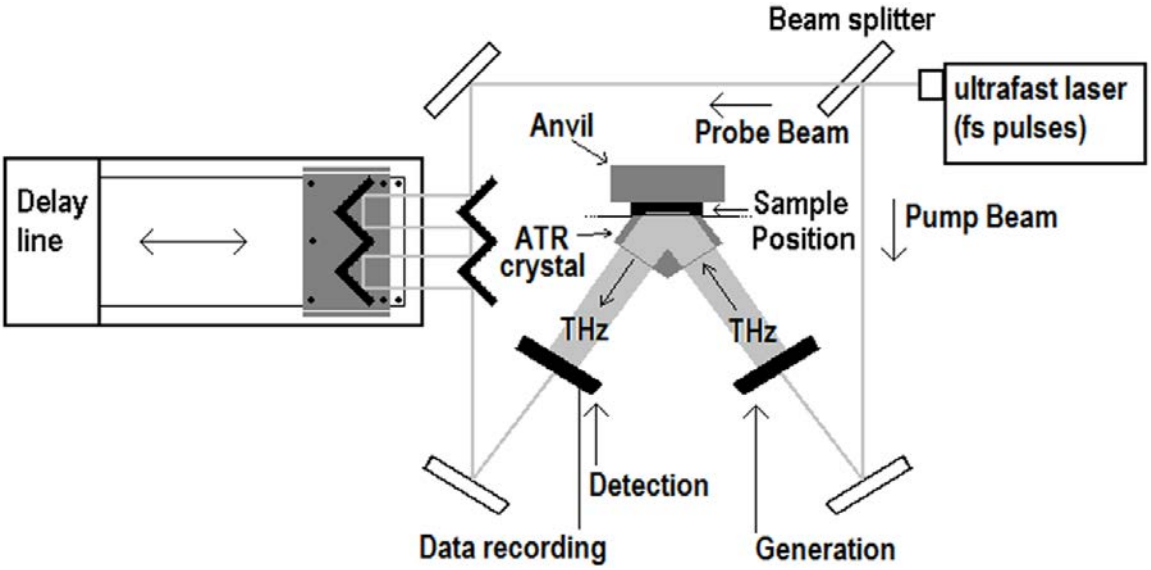


Figure I-11. Schematic diagram of a typical THz-TDS measurement with an ATR module.

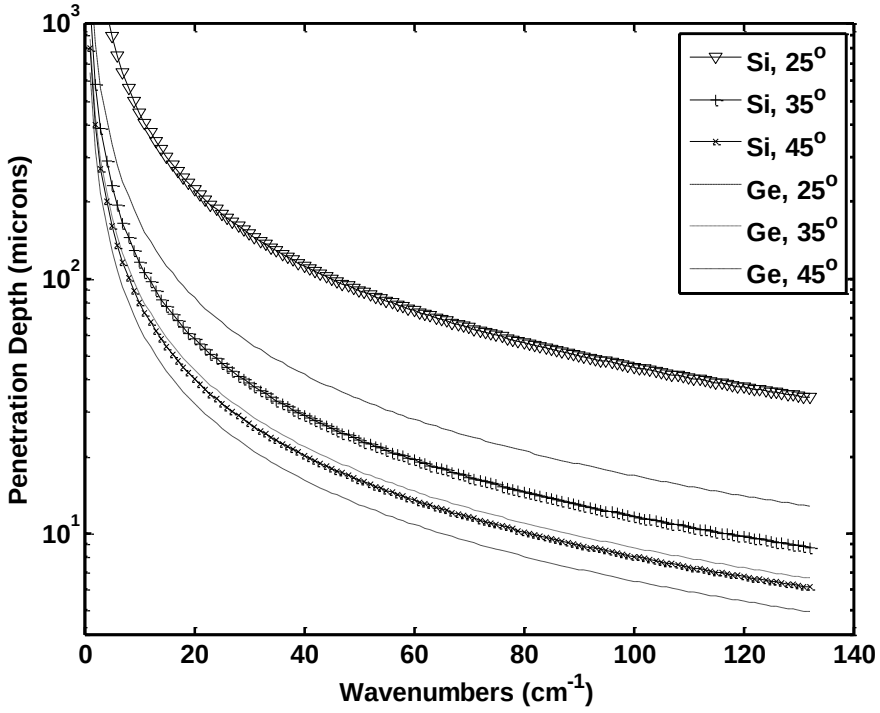


Figure I-12. Penetration depth in microns calculated for various ATR geometries with silicon and germanium crystals. A refractive index of 1.4 was used for the sample.

where θ_1 is the angle of incidence, n_1 is the refractive index of the ATR crystal, n_2 is the refractive index of the sample, and θ_2 is the angle of the refracted radiation. Equation I-7 illustrates the dependence of the critical angle of reflection on the refractive index of the ATR crystal and the sample, a condition which is met when θ_2 is equal to 90°. If the difference between the refractive indexes is not sufficiently large, reflection will not occur at the crystal/sample interface and the incident radiation will be lost in the sample and none of the radiation reaches the detection optics. As a result, manufacturers of ATR modules often provide optical configurations having different angles of incidence or continuously variable angles of incidence (varying θ_1), as well as a variety of ATR crystal materials. ATR-based experimental designs must account for both the absorptivity of the species as well as its refractive index.

A hand-tightened anvil is sometimes used to bring the sample into intimate contact with the surface of the ATR crystal and drive air from between analyte particles. While an anvil is not required for THz-TDS ATR measurements, improved performance is obtained when it is included. Results indicate, however, that absorption spectral features will change depending on the applied pressure.[48] In general, as the pressure of the anvil is increased, the sample is brought closer to the ATR crystal and the protruding evanescent wave. Air is displaced from this region as it is replaced by the sample, effectively increasing the degree of interaction between the radiation and sample and increasing the sensitivity of the measurement.[48]

The versatility and ease of use associated with the ATR geometry have made it appealing for qualitative work. Very small sample quantities (down to sub-milligram masses) can be analyzed rapidly and non-destructively without sample preparation.[48] Quantitative work with such a module would be far from trivial however, as the attenuation of reflected radiation is not only a function of the substance's absorptivity, but also a function of the sample's refractive index. Were two sample powders to be analyzed in a mixture, the ratio of the two species would need to be established prior to

analysis, as the penetration depth, and thus degree of attenuation of THz radiation in each powder, would vary. Sample mixture heterogeneity would also significantly impact the penetration depth. Refractive index of the sample represents another constraint on such measurements, as a sample with a high refractive index will allow radiation to couple from the ATR crystal to the sample with lower reflection according to Snell's law.

Reflection: Air Propagation

Reflection measurements can also be carried out by propagating the THz pulse through air and allowing the radiation to interact with the sample at the sample/air interface. This measurement mode has an advantage over ATR-based reflection measurements in that there are no reflective losses of radiation associated with the coupling of the THz pulses into a crystal of high refractive index. Measurements can also be performed with samples too harsh for an ATR crystal, such as hard samples that may abrade the ATR crystal surface or those which may chemically attack the crystal material. Samples with curved or irregular surfaces may require characterization with this measurement mode, as interrogation with an ATR crystal requires intimate contact. As can be seen in Figure I-12, surfaces at a distance greater than approximately 100 μm from the ATR crystal will go unprobed, as the evanescent field does not extend this far into the sample at the wavelengths accessible with most THz-TDS instruments.

For diffuse reflectance measurements, timing of the pump and probe beams must be adjusted to account for the thickness of the sample and the corresponding transit of the THz pulse within an irregularly shaped sample. Such a match can represent a practical limitation for diffuse reflectance-based measurements in that the effective thickness of the sample must be known and the distance through the corresponding optical delay of the probe beam must be adjustable to match the distance through the sample while maintaining a stable alignment. Figure I-13 shows a schematic of a THz-TDS in a reflection configuration.

Selected Gas Phase Literature

Spectroscopic investigation of gaseous species using THz frequencies is by no means a recent development. Scientific literature covering this topic dates back at least to 1934 when Cleeton and Williams published results from their experiments with gaseous ammonia using wavelengths around 1 cm,[49] followed by reports of the water vapor absorption spectrum in 1935 and 1937 by Barnes and colleagues[50] and Randall and colleagues,[51] respectively. These milestones in the THz literature were succeeded by Allied involvement in World War II, which spurred the development of various microwave technologies which would eventually be utilized for scientific endeavors.[26] A brief timeline of scientific milestones related to rotational spectroscopic analysis of gaseous species as well as the development of THz-TDS technology is shown below.

Timeline

*Stars denote advances integral to the operation of the THz-TDS apparatus.

- 1880 – Langley invents the bolometer.[52]
- 1896 – J. C. Bose uses electromagnetic waves to perform various tasks from a distance, including the detonation of gunpowder and ringing a bell.[53]
- 1897 – J. C. Bose develops components for the generation and detection of millimeter waves[54]
- 1897 – Rubens and Nichols describe production, manipulation, and detection of "rays of great wavelength".[55]
- *1904 – Hagen and Rubens examine the reflectivity of various metals in the THz range and discover very low losses with many metals.[56]
- 1911 – Rubens and Baeyer demonstrate the utility of a quartz-enveloped mercury arc lamp for generation of far-infrared radiation.[57]

- 1923 – Nichols and James Tear join the electrical and optical regions using a Hertzian oscillator to generate wavelengths as short as 220 μm (45.45 cm^{-1} or 1.36 THz).[58]
- 1924 – Glagolewa Arkadiewa uses brass filings immersed in oil to generate wavelengths around 90 μm (111.11 cm^{-1} or 3.33 THz).[59]
- 1933 – Cleeton and Williams report gas-phase spectroscopy of ammonia at 1 cm wavelength.[49]
- 1935 – Barnes, Benedict, and Lewis publish a complete absorption spectrum of water vapor to 170 μm (58.82 cm^{-1} or 1.76 THz).[50]
- 1937 – Randall and colleagues utilize a grating-based spectrometer to distinguish between absorption lines with a resolution of 0.5 cm^{-1} at 75 μm (133.33 cm^{-1} or 4.00 THz).[51]
- 1940's (early) – Microwave technology expands greatly with the Second World War. KRS-5, CsBr, and CsI optics extend useable spectral ranges to 35 μm (KRS-5, 285.71 cm^{-1} or 8.57 THz) and ~60 μm (CsBr and CsI, 166.67 cm^{-1} or 5.00 THz), extending the spectral range of dispersive instrumentation. The cooled bolometer is applied to the detection of far infrared (FIR) radiation.[26]
- 1946 – First mm-wave klystron described.[26]
- 1947 – Marcel Golay invents the pneumatic infrared detector, later referred to as the Golay cell.[60]
- *1950's – Computers become more common-place, enabling elaborate spectroscopic analyses, especially those involving interferometry.[26]
- 1953 – The backward-wave oscillator (BWO) is developed as a source of far-infrared radiation.[61]
- 1954 – King and Gordy extend the range of microwave sources to frequencies above 300 GHz.[62]

- 1956 – Gebbie and Vanasse report the spectroscopic application of a reflection interferometer to the analysis of radiation transmitted through water vapor.[63]
- 1957 – Strong describes developments in interferometry and applications to far-IR spectroscopy.[64]
- 1959 – Boyle and Rogers report a new and convenient carbon bolometer.[65]
- *1959 – Fray and colleagues describe the operation of a photoconductive detector for far-infrared radiation.[66]
- *1960's – The development of laser technology provides various intense sources in the far-IR region.[26]
- 1960 – Putley reports the development of the n-type indium antimonide bolometer.[67]
- 1961 – Low describes a cooled bolometer constructed with germanium.[68]
- 1966 – Pyroelectric detection is reported by Putley, providing enhanced access to far-IR frequencies.[69]
- 1967 – Ulrich describes efficient low-pass filters for the far-IR region.[70]
- 1970 – Optically pumped far-IR lasers (OPFIRL) are developed, providing an intense source of far-IR radiation.[71]
- 1975 – The first free electron laser (FEL) is brought into operation.[26]
- *1980's – Improvements in lithographic techniques yield smaller generative structures and thus extend the emitted frequency range of THz sources.[72]
- 1980's (late) – Advent of THz-TDS.[73]
- *1981 – Mourou et al. describe the generation of ps-duration microwave pulses using Nd: YAG laser-pumped Ge photoconductive switches.[74]
- 1981 – Smith and colleagues develop an improved photoconductive film consisting of radiation-damaged silicon-on-sapphire (RD-SOS).[75]

- *1984 – Auston et al. describe the generation, transmission, and detection of ps-duration electromagnetic pulses using excited photoconductor antennas, without intermediate excitation steps.[76]
- *1987 – DeFonzo and Lutz emit far-IR electromagnetic pulses and allow them to propagate in air, performing detection with identical photoconductive antennas.[77]
- *1988 – Fattinger and Grischkowsky demonstrate enhanced coupling of THz radiation via Si lenses.[78]
- *1991 – Femtosecond pulses are generated using a self-mode-locked Ti:sapphire laser by Spence and colleagues.[79]

Selected Literature by Species

In this dissertation, acetaldehyde (CH_3CHO), acetonitrile (CH_3CN), ammonia (NH_3), ethanol (CH_3CH_2OH), methanol (CH_3OH), propionaldehyde (CH_3CH_2CHO), propionitrile (CH_3CH_2CN), and water (H_2O) vapors were selected for analysis with the THz-TDS apparatus for a variety of reasons such as structural, toxicological, environmental, and spectroscopic properties. Rotational spectroscopic constants and features have been published for each of these species in the peer-reviewed scientific literature for various reasons such as astronomical investigations, structure elucidation, and development of theoretical models. While the intended use of the respective constants or conclusions published in these works may not overlap with the purpose of the THz-TDS investigations described herein, this body of literature nonetheless provides evidence of spectral features observed in the THz range and in some cases supports the theoretical origins of such features. Often times, these publications will provide tabulated parameters, some of which will be discussed in greater detail in Chapter IV.

Spectroscopic investigations of these species are carried out not only with the most common of isotopes, but also with rare species such as singly or multiply-deuterated species. Despite their decreased relative abundance, the characterization of these species

is still significant, as their contribution to spectral data is proportional to their concentration (or abundance), as dictated by Beer's law (Equation I-2). As such, any atmospheric measurement of a species will contain the spectral signature of the common isotope superimposed on that of the rare species. Deuterated species also provide an experimental parameter which may provide additional insight into theoretical models, since the rotational spectrum of species (like the vibrational spectrum) is dependent in part by the mass of constituents. The phrase "common isotope" is used to denote the natural abundance distribution of isotopes for each molecule: ^{1}H ^{12}C ^{14}N, and ^{16}O. The composition of rare isotopes is described explicitly when necessary. It is also important to note that in some cases, published accounts will describe transitions for the same molecule and isotope, but work in the field continues as new transitions are discovered and the theory is refined. These discoveries are due in part to the use of novel, improved instrument configurations as well as enhanced experimental methods.

Acetaldehyde

Several publications have focused on the rotational spectral features and parameters of acetaldehyde in spectral regions bordering the high and low-frequency ends of the useable range of current THz-TDS instrumentation. Microwave spectral analysis of acetaldehyde has been reported, providing a set of spectral parameters and molecular constants up to a frequency of 16.68 cm^{-1}.[80-86] The impact of torsional motion on the microwave spectrum of acetaldehyde and the corresponding parameters have been characterized,[80, 83, 85, 86] along with the fundamental torsional transition, which may be observed in the far-IR frequencies around 150 cm^{-1}.[80, 84, 87, 88] Deuterium substitution has also been used to investigate spectra at microwave frequencies, including CH_2DCHO,[89, 90] CHD_2CHO,[89] CD_3CDO,[83] CH_3CDO,[91] CHD_2CHO,[89, 90] and CD_3CHO.[92] The rich rotational spectrum of acetaldehyde is evident from the details of these reports.

Acetonitrile

Rotational spectral features of acetonitrile vapor have also been characterized. Molecular parameters and peak positions have been reported for the common isotope ($^{12}CH_3{}^{12}C^{14}N$) over microwave frequencies,[93] THz frequencies,[94] and the far-IR spectral range.[95, 96] Spectral broadening of acetonitrile caused by the sample pressure was characterized by Schwaab and coworkers[97] and is described in greater detail in Chapter IV. Scientific literature has also been published in which the spectral features and molecular parameters of isotopically labeled compounds are reported. For example, values have been reported for completely-deuterated acetonitrile (CD_3CN),[98, 99] as well as those containing isotopes of carbon and nitrogen such as $^{13}CD_3CN$,[98, 99] $CD_3{}^{13}CN$,[99] $CD_3C^{15}N$,[99] $^{13}CH_3CN$,[93] $CH_3{}^{13}CN$,[93] and $CH_3C^{15}N$.[93] For some molecules, rotational transitions of molecules in an excited vibrational state will overlap with the ground state rotation spectrum, increasing the complexity of observed spectral features. Rotational transitions of acetonitrile in the ν_8 vibrational state were reported by Boucher[93] and Matsumura.[99]

Ammonia

Spectroscopic parameters of ammonia have been published across the THz spectral range both from theoretical analyses[100, 101] and experimental data.[49, 102-105] These efforts (and others) helped to establish a set of parameters that accurately predict the molecular rotational spectra of ammonia. The spectrum of ammonia was also reported by Kolbe and Leskovar in 1983 up to 46.7 cm^{-1}.[106] Similar efforts have been published for $^{15}ND_3$.[107] The effect of collisional broadening of ammonia spectral features has also been characterized by several researchers, covering a range of rotational transitions.[108-111] Like acetonitrile, absorption features of ammonia rotational transitions can be readily observed for molecules in excited vibrational states such as the ν_2 and $2\nu_2$ states, as described by Husson, Goldman, and Orton.[100] One particularly interesting spectroscopic feature that is

observed with ammonia vapor is a set of transitions related to a phenomenon known as "inversion". To understand this phenomenon, imagine that a plane is drawn through the nitrogen atom in the ammonia molecule which is parallel to a plane drawn through the three hydrogen atoms. With perturbation via electromagnetic radiation of certain frequencies, these hydrogen atoms will cross the plane drawn through nitrogen, producing the mirror image of the initial ammonia molecule. Inversion transitions are observed at low frequencies (0.043 to 0.077 cm^{-1}, or 1.3 to 2.3 GHz),[107] but the rotational spectrum of the species is further complicated by these transitions.[100, 107]

Ethanol

The rotational spectrum of ethanol bears little resemblance to spectra of the other compounds described thus far. Throughout the decades following World War II, several groups of researchers have studied the spectroscopic features of ethanol. Bak and coworkers provided a rough outline of microwave spectral features near 1.25 cm^{-1}.[112] Michielsen-Effinger provided a set of rotational constants and centrifugal distortion coefficients in 1969.[113] Transitions associated with trans and gauche conformations have been identified by Sasada[114]and Pearson,[115] respectively. Durig and Larsen described the rotational and torsional spectrum of ethanol from 50 to 370 cm^{-1}.[116] Both the common isotope (CH_3CH_2OH) and a singly-deuterated isotope (CH_3CH_2OD) were analyzed in this work, providing spectra of each species. While the spectra of ethanol may appear to consist of one broad featureless continuum, Durig and Larsen report a large degree of ro-vibrational fine structure in the far-IR and THz spectral ranges.[116] A thorough investigation of the 0.40 to 55.04 cm^{-1} range was reported by Pearson in 2008, identifying at least 10,000 transitions in these frequencies.[117] A collection of microwave-frequency studies of ethanol was compiled by F. J. Lovas in 1982.[118]

Methanol

Methanol has been studied extensively by the spectroscopic community, likely owing to its strong far-IR emission as well as its identification in astronomical spectral observations.[119] Methanol also serves as a sort of model molecule, being one of the simplest molecules that exhibits a barrier to internal rotation.[120] Microwave spectra and parameters of this species have been reported over the span of several decades as theoretical models and instrumental methods developed. The microwave spectrum of natural abundance isotope molecule of ammonia has been described by Herbst,[120] Xu,[121] Lees,[122] Sakai,[96] and Moruzzi,[123] amongst others, covering various transitions between 0.017 and 90 cm^{-1}. The rotational spectrum of the natural abundance isotope molecule has also been studied extensively from a theoretical standpoint.[124, 125] Peak positions and molecular constants have also been reported for isotopically labeled molecules such as $^{13}CH_3OH$,[121, 122] $CH_3{}^{18}OH$,[122, 126, 127] CH_2DOH,[128] and CD_3OD.[129] The torsion-rotation spectrum of methanol has also been widely reported over the far-IR frequency range. Literature accounts are available for the common isotope[130] as well as rare isotopes such as CH_2DOH,[131] CH_3OD,[132] CD_3OD,[128, 132] CD_3OH,[130, 133] $^{13}CH_3OH$,[119] and $CH_3{}^{17}OH$,[134] as examples.

Propionaldehyde

Like ethanol, propionaldehyde exhibits a rotational barrier which gives rise to unique spectral features associated with unique rotational conformers, referred to as the s-cis and gauche conformers. Much of the found scientific literature focuses on the investigation of propionaldehyde's internal rotation barriers, and most involved the investigation of the natural abundance isotope molecule. Rotational features have been reported over the microwave region by Pickett,[135] Butcher,[136] and Randell.[137] These accounts also provide a set of molecular parameters describing the structure and spectrum of propionaldehyde. The barrier to internal rotation has been studied and reported in

several accounts from Smeyers,[138] and Durig,[139, 140] focusing on transitions in the far-IR frequency range. One of Durig's analyses also examined isotopically labeled propionaldehyde such as CH_3CH_2CDO, CH_3CD_2CHO, and CD_3CD_2CHO.[139]

Propionitrile

Propionitrile has been studied by several groups, though attention has been focused primarily on the microwave frequencies. Peak parameters and molecular parameters for the natural isotope molecule have been reported by Laurie,[141] Lerner,[142] Fukuyama,[143] and Burie.[144] These experimental efforts were later supplemented by the theoretical work of Demaison and coworkers in which the energy of the torsional barrier and equilibrium structure were determined.[145] Spectroscopic characteristics of isotopically labeled compounds such as $CH_3CH_2{}^{13}CN$,[142] CD_3CHDCN,[142] and CD_3CD_2CN[146] have also been reported. F. J. Lovas published a review of MW spectroscopic investigation of propionitrile in 1982.[118]

Water

Of all the candidate molecules described in this overview, water vapor has received a majority share of attention, likely due to its abundance in the atmosphere. Characterization of water's rotational absorption spectrum is also of interest to those performing astronomical and atmospheric measurements, as its absorption features will be present in any THz-frequency measurement performed within the earth's atmosphere.[147] Absorption peak positions and molecular parameters have been reported on various occasions, covering several decades of investigation. Many studies have been carried out with the common isotope of water, $H_2{}^{16}O$. Peak positions, and in some cases molecular parameters, have been reported by various researchers in different frequency ranges. Spectral parameters and molecular constants have been reported for microwave - frequency transitions of the common isotope molecule in the ground vibrational state[104, 148-153] as well as for rotational transitions in excited vibrational states.[154] Broadening

parameters[148, 155] and pressure-induced frequency shifts[156] have also been characterized for the common isotope. The absorption spectrum of water vapor was also reported in the range of 3 to 90 cm^{-1} by Sakai and colleagues in 1978 using an FT spectrometer,[96] and later by Kolbe and Leskovar in 1983 up to 46.7 cm^{-1}[106]as collected with traditional microwave instrumentation. Isotopically labeled compounds have also been studied individually, such as HDO,[104, 150, 151, 157] D_2O,[104, 151, 153] T_2O (T = tritium, or ^3H),[151] $H_2{}^{17}O$,[150, 152, 158] $H_2{}^{18}O$,[104, 150, 152, 157, 158] $HD^{18}O$,[104] and $D_2{}^{18}O$.[104]

Despite its molecular simplicity, theoretical models developed over the years have failed to accurately predict the absorption spectrum of water vapor in the atmosphere. The greatest discrepancies in these models have resided in the space between absorption peaks. These regions, which are expected to contain minimal water-related absorption often exhibit higher absorption in experimental data than is predicted by theoretical models. The apparent broadband absorption underlying the most intense transitions is often referred to as the water vapor absorption "continuum", due to the absence of discrete peaks or fine structure in these under-predicted baseline regions. This phenomenon has been studied in varying frequency ranges by different researchers,[147, 159-163] though the source of this discrepancy does not appear to have been conclusively determined to date. Some have attributed the anomalous absorption to dimers and higher-ordered clusters of water vapor molecules as well as weak absorptions from rotational transitions in excited vibrational states.

Selected THz-TDS Literature

The THz-TDS instrumentation used for this dissertation research has also been applied to analysis of many different gaseous species. Of course, a vast array of the THz spectroscopy literature uses non-TDS instrumentation.[17, 164-176] Still, attention is given primarily to literature describing efforts with THz-TDS instrumentation due to the increased relevance of these accounts to the experimental work described in Chapter IV.

THz-TDS instrumentation has been used for analysis of a variety of gaseous species. Lin et al. reported an algorithm designed to characterize THz-frequency spectra based on a simplified set of parameters related to peak positions and intensities.[177] Decomposition of spectral data into reduced data sets can be advantageous in that it reduces processing time and simplifies the calculative procedure. A simple Euclidean distance determination is used to compare experimental spectra to the aforementioned simplified spectral data sets.[177] The library spectral set which provides the minimum Euclidean distance is considered a match, and the species is identified.[177] Uno and coworkers published an account of the use of THz-TDS to analyze combustion products *in situ,* as well as a variety of other spectroscopic and imaging measurements.[178] In this series of experiments, an electric furnace was used to combust woodchips and polystyrene beads. Combustion byproducts were passed through a gas cell with an 8 cm path length, and terahertz pulses were transmitted through this cell. Transmitted pulses were recorded and used to produce transmittance spectra.[178] Absorption peaks were identified as those of CO and H_2O, and the experimental and HITRAN database peak positions of CO were compared graphically, showing the wavelength accuracy of the instrument.[178] The authors also concluded that THz-TDS is well-suited for such measurements because particulate matter produced during combustion did not appear to have a significant impact on the transmittance of THz pulses.[178]

Cheville and colleagues published an account of the measurement of terahertz pulses transmitted through flames.[179] This work exemplifies one advantage inherent in THz measurements in that the frequencies measured do not overlap with the frequencies of radiation emitted by hot chemical species.[179] This conclusion is similar to one reached by Takeshi and colleagues, but was observed in more extreme conditions.[178] Were it not for this unique property of THz radiation, the measurement would be confounded by detection of incoming radiation that was not generated by the spectrometer source. The authors also note a change in the arrival time (optical delay) of the measured pulse which

is modulated by the flame.[179] A shorter delay time was observed when the flame was "on" relative to times when it was "off". Spectral features were identified by comparison with calculated spectra, illustrating the presence of water vapor rotational transitions in both the ground and excited v_2 vibrational state.[179]

Yang et al. describe a more recent analysis of the aforementioned "continuum absorbance" observed when THz pulses propagate through water vapor.[180] Spectra are reported as collected in a variety of meteorological conditions such as high humidity, fog, low temperatures, high dust levels, and rain. Measurements were performed with a 6.18 m long-path cell, providing sensitive measurement of the water vapor spectrum. The increased absorption between spectral peaks as well as the discrepancy between prior published results are attributed to attenuation of THz radiation by water molecule dimers and polymers or clusters as well as self-broadening from the most intense rotational transitions.[180]

Foreign-gas broadening of water vapor spectral features has also been characterized with THz-TDS measurements. Seta and Hoshina have reported pressure broadening parameters for various rotational transitions across the water vapor absorption spectrum as broadened by N_2 and O_2 gases.[181, 182] Spectra of water vapor were collected in the presence of varying amounts of N_2 or O_2. Spectral features were then fit with a Lorentzian function, which provided the width of the absorption features. Once the widths had been calculated for samples at all pressures, linear regressions were performed with the pressure values and widths to obtain the respective broadening parameters (the slope of the best-fit line), in units of frequency per unit of pressure (MHz/Torr or MHz/kPa).[181, 182] The importance of these parameters was also discussed as they relate to the estimation of atmospheric water vapor concentrations by the Odin/SMR satellite. In previous atmospheric measurements, a broadening parameter of 4.37 MHz/Torr had been used, but it was found that this parameter was actually that of the same rotational transition of water vapor molecules in an excited vibrational state.[182]

Inclusion of the new parameter (4.012 MHz/Torr) provided a more accurate estimated water concentration for altitudes between 40 and 85 km.

Sagawa and coworkers reported broadening of water vapor features by CO_2. While CO_2 is not a majority constituent of Earth's atmosphere, it comprises approximately 96.5% of Venus' atmosphere.[183] Water vapor is a relatively minor component of Venus' atmosphere, being found at a concentration of less than 100 parts-per-million (ppm).[183] Though present in low concentrations, it plays a significant role in the climate and atmospheric chemistry of Venus. As such, a set of accurate broadening parameters for water vapor as broadened by CO_2 facilitates accurate estimation of water vapor in Venus' atmosphere and contributes to the understanding of its climate. The experiment was conducted as described in experimentation with nitrogen and oxygen, and the calculated broadening parameters were compared to theoretically-obtained values.[183] Ordinarily, the broadening parameters associated with foreign-gas broadening are estimated for species like CO_2 by multiplying the parameters associated with well-characterized species by a scaling factor.[183] While this may sometimes produce reasonable values for approximation, estimated concentrations can vary as much as ~75% when the broadening parameter is assumed to be equal to that of air.[183]

Deuterium oxide vapors have also been studied with THz-TDS systems. Yu and colleagues reported the THz spectrum of D_2O in 2006 covering 0-2 THz, reporting peak positions and widths of the 26 most apparent transitions.[184] The experimentally-determined parameters are compared with values from the scientific literature, showing agreement between the two sets of values to within several GHz. The impact of temperature on the observed peak widths is also reported.[184] At increased temperatures, broader spectral peak widths are observed as a result of increased collision frequency, similar to the effect of foreign-gas broadening for samples analyzed at elevated pressures.[184]

Harde and coworkers have reported analyses of THz spectral line shapes and broadening on several occasions, through the study of methyl chloride[185, 186] and ammonia vapors.[187] In Harde's first report of the spectral features of methyl chloride, much attention was devoted to the unique appearance of the TDS for these samples. Periodic features were observed in the signal which were attributed to the presence of methyl chloride vapor.[185] The experimental signal was compared to a calculated signal, showing good agreement between the two data sets. Data were collected at multiple pressures, and the effect of self-broadening was analyzed for a wide range of rotational transitions. Spectral absorption features are also shown to be more precisely fit by an assumed van-Vleck Weisskopf line profile, a line shape commonly observed in microwave-frequency investigations, rather than the Lorentzian profile, commonly observed in vibrational spectra.[185]

In a subsequent publication, Harde and coworkers advanced their work with methyl chloride with further studies of this compound as well as newer results from methyl fluoride and methyl bromide.[186] These molecules produced spectral features with a similar overall appearance, though the spacing of peaks and center of the grouping of peaks were changed by the different inertial moments for these molecules. Greater spacing was observed with methyl fluoride, which decreases as the fluorine was substituted with chlorine and ultimately with bromine. The center of the collection of spectral features also shifted from a high frequency (methyl fluoride) to a lower frequency (methyl bromide).[186] Spectral broadening was also characterized for each species as a function of the rotational quantum number, with the greatest broadening observed near J values of 10-15 and tailing off at higher and lower quantum numbers.[186] The authors also reported a new theoretical model for absorption features observed in this spectral range. Experimental spectral features were not so accurately modeled when using either the Lorentzian or van-Vleck Weisskopf line profiles. The use of a Lorentzian profile produced a calculated fit of lower magnitude than the experimental

data at the high frequency end of the absorption spectrum of methyl chloride, whereas the van Vleck-Weisskopf profile produced a fit of greater magnitude than the experimental data.[186] Harde and colleagues reported a modified theoretical model which utilizes a combination of the two line profiles, where the relative contribution of each line shape to the overall fit was altered via a set of switching functions.[186] This new model produced a more accurate fit to the experimental data, illustrating an overlap of the two line shape theories in this spectral region.

The molecular response theory developed by Harde et al. has also been applied to the analysis of the spectral features of ammonia vapor.[187] The model was found to fit the observed spectral features to a high level of accuracy with two exceptions. At low frequencies, the calculated absorption is too large which is attributed to the rotation-inversion spectrum of ammonia,[187] mentioned previously in the overview of microwave analysis of the species. Features associated with the inversion of the ammonia molecule were found to fit well with the molecular response theory.[187] The remainder of the observed absorption was attributed to the rotational spectrum of the vibrationally-excited ammonia molecule (in the ν_2 vibrational state), as well as absorption by the ammonia dimer. Dimerization of gaseous species and the analysis of these clusters will be described in greater detail in Chapter IV. These extra absorption features confound measurements of the line-shape parameter.[187]

Measurements of gaseous samples by THz-TDS can also be used in security-related applications. Foltynowicz and colleagues reported the gas-phase spectrum of 2,4-dinitrotoluene in 2006.[188] Measurements were performed with a 34 cm gas cell coupled to a temperature-controlled sample chamber. Mass spectra were collected with vapors from the sample chamber before and after sample scans to verify the presence of the analyte. Experimental THz spectra show a dense collection of absorption peaks from at least 50 GHz to 2.67 THz, providing a set of features that could potentially be used to identify this compound when present as a vapor.[188]

Published Quantitative Efforts

Reported efforts to quantify gaseous species on the basis of rotational spectroscopy are relatively rare. As such, this section provides an overview of analytical measurements executed with rotational spectroscopy, regardless of instrument configuration or useable spectral range.

Microwave Instrumentation

Lawrence Hrubesh demonstrated what was at the time rapid (requiring < 7 minutes) and automated identification and quantitation of several gaseous species using a computer-controlled microwave spectrometer.[189] A calibration plot was provided for acetone, ethanol, and ethylene oxide, showing linearity across the 1 – 1000 micron pressure range (~1.3×10^{-6} – 1.3×10^{-3} atm).[189] The experimental procedure described by Hrubesh involved sequential analysis of frequencies specific to the compounds contained in the system's spectral library. To perform qualitative analysis, the system would perform a measurement at the frequency of the strongest absorption feature for the first molecule in the spectral library. After signal averaging, the experimental data is analyzed to determine whether the measured signal-to-noise ratio (SNR) is greater than or equal to 3:1 at the frequency of the absorption peak. If the SNR exceeded the 3:1 threshold, a second measurement was then performed at a frequency equal to that of the second-strongest absorption peak of the species. If the SNR did not exceed the threshold, the automated system would move to the next compound in the spectral library and repeat the previous steps. Two frequencies were stored for each compound at which the highest absorption would be observed. Hrubesh found that species in sample mixtures could typically be identified in less than 3 minutes, though quantitative analysis required additional steps wherein the microwave power was sequentially increased at the frequency of the strongest absorption feature.[189] The power at which maximum absorption is achieved is then compared to experimentally-determined maxima.[189] The

system was also applied to measurement of toxic gaseous species in oil shale, and a table was provided showing the qualitative and quantitative results of the analysis.[189]

Quantitative efforts were reported by Andresen et al. in 1994 as performed with a Fourier-Transform Microwave (FTMW) spectrometer. This apparatus applies a brief microwave pulse (~0.5 μs) to the sample contained in the resonant cavity (analogous to gas cells typically used in optical spectroscopy). The sample gas is excited and polarized by the pulse, resulting in the emission of a coherent microwave signature that is specific to the molecule in the cavity.[190] Emission is recorded as a function of time, and the magnitude of digitized signal values are summed, producing a scalar value which is proportional to the concentration of the gas. This method provides a more rapid and computationally simple alternative to Fourier transformation of the time-domain decay signal.[190] However the authors point out that the Fourier transformation (which provides a frequency-domain emission spectrum) is preferable in situations where multiple absorbing (or emitting) species are present in the resonant cavity, as the emission signals of the two (or more) species will then be convoluted.[190] Detection limits associated with the method are listed for various species such as ammonia, chlorotrifluoromethane, 2,6-dimethylpyridine, methanol, toluene, cyclohexanone, m-xylene, and sulfur dioxide, with values ranging from 0.002 to <60 ppm.[190] Also included is a calibration plot based on methanol samples, illustrating the linearity of the instrument response to this species. Andresen and colleagues analyzed the impact of overall pressure on the method and found that the response of the system to analyte molecules was diminished at pressures exceeding 7×10^{-3} Pa.[190] This drop-off in instrument sensitivity was attributed to collisional transfer of excitation energy rather than emission of absorbed energy. No mention is made of analyses with sample mixtures.

Microwave spectroscopy has also been reported as a tool for quality control of natural gas. Leontakianakos published an article in which water vapor content is determined in air and in liquefied natural gas.[191] The sensitivity to water vapor was

increased by the inclusion of a Stark field modulation system, a component often used in microwave-frequency experiments. This system consists of an electrode that is placed within the resonator cavity which is insulated from the cavity. Gordy and Cook explained the effect as a modulation of microwave energy near the resonant frequency of an absorption line.[13] If measurement is performed with a narrowband detector tuned to the frequency of the applied field, greater sensitivity is achieved.[13] When an even harmonic of the Stark field is applied across the wire grid in the resonator cavity, the absorption of gaseous species can be enhanced, as demonstrated by Hughes and Wilson in 1947,[192] and by Hershberger in 1948.[193] As such, the length of the resonant cavity or waveguide may be reduced to provide a more compact system without compromising sensitivity. The decreased size of the resonant cavity also reduces the pumping time required to achieve the low pressures required for spectral analysis. Leontakianakos reports the use of the fourth harmonic in the detection of water vapor over a 300 to 9400 ppm concentration range.[191] Two quantitative methods were evaluated: one in which the absorption peak associated with water vapor was measured from the peak value to an assumed baseline (a), and one in which the height of the peak was measured from the peak value to the depth of adjacent absorption minima (b). Both calibration methods appear to produce linear calibration plots at low concentrations, but method (a) produced a non-linear response at high water concentrations.[191] Quantitative analysis of water vapor was demonstrated with a water vapor sensitivity of approximately 300 ppm.[191] Again, no mention is made of analyses performed with mixtures which include potential interferents, or the selectivity of the proposed method.

Rouleau and colleagues reported the use of a differential cavity microwave spectrometer for the determination of water vapor concentrations in 1999.[194] This apparatus is somewhat analogous to a dual-beam spectrophotometer in that the reference and sample are analyzed concurrently. Electromagnetic radiation generated by a microwave-frequency synthesizer is split equally between two beam paths and propagates

towards the reference and sample cavities. Radiation reflected back toward the synthesizer from the sample cavity is attenuated, but radiation propagating toward the detection system passes through a junction which imparts a 180° phase delay. When the reference and sample beams are recombined, the measured signal represents the difference between the beams reflected from the reference and sample cavities. Rouleau and colleagues provide a complete explanation of how this measured signal voltage is proportional to the difference in the dielectric constants of each beam path.[194] This method is reported to provide an improvement in sensitivity on the order 10^3.[194] Water vapor was added to argon and helium gas, and the change in measured voltage was determined as a function of water vapor concentration for each mixture, and the detection limit of the method was estimated at 6 ppm.[194]

Zhu and colleagues reported quantitative analysis of ethylene oxide in various matrices, such as nitrogen, wet air, and wet nitrogen.[195] The apparatus provided similar response across the three matrices when performing analysis with samples containing low levels of ethylene oxide, but the calibration plots became divergent and in some cases nonlinear at higher concentrations. Measurements were also carried out in a manner which simulated a sterilization procedure. In these experiments, nitrogen and water vapor were added to the chamber, followed by induction of ethylene oxide.[195] The gas was allowed to set for some time, representing the time of sterilization. This was followed by 5 air purging cycles to clear ethylene oxide from the system. The mean concentration of ethylene oxide across the resultant concentration-time profiles were near 50 %, and standard deviations close to 3.6 % were reported based upon the deviation of predicted concentrations from the reference concentrations.[195] These results suggest the applicability microwave spectroscopy to real-time monitoring of sterilization processes.

Rezgui and colleagues reported the use of a diode-based microwave spectrometer designed to generate and measure frequencies above those typically accessible with older microwave spectrometers. The apparatus was used to analyze various gases at ppm

concentrations such as HDO (9 ppm), OCS (2.7 ppm), $HCH^{18}O$ (2000 ppm), $HCH^{17}O$ (373 ppm), and CH_2CHCN (360 ppm).[196] The SNR of these spectra were low, and many spectra were reported as smoothed via Gaussian correlation. Absorption features were identified for each species despite the low concentrations and relatively high noise levels.[196]

The Rezgui apparatus was modified slightly and used again for measurements reported by Wilks et al.[197] In this work, a calibration plot was provided for sulfur dioxide at various volume fractions in N_2 based on the maximum absorbance values of rotational transitions.[197] Experimental absorption values were also compared with calculated values as well as those found in the literature.

Another paper published nearly concurrently with the previous report described microwave spectroscopic quantitation of oxygen. Ordinarily, gaseous species must possess a permanent dipole to exhibit microwave or THz-frequency absorbance. Oxygen lacks a permanent dipole, though it does absorb some microwave frequencies as a result of spin reversal of its unpaired electrons.[198] Absorption coefficients for these transitions were compared to literature values and used to construct a calibration plot for various mixtures of oxygen and nitrogen, ranging from 0 to 100 % oxygen by volume, showing the linearity of the series of measurements.[198] Measurements were performed with peak values from transitions at 58.324 and 60 GHz, illustrating the ability of the apparatus to determine the concentration of oxygen.[198]

Dumesh and colleagues described the quantitation of methanol using a microwave -frequency device known as an "orotron" oscillator as a tunable source.[199] This source contains elements similar to those of the more common backward-wave oscillator, but offers the advantage of providing highly frequency-stable output, as well as greater sensitivity than other microwave sources.[199] These characteristics are of primary importance when conducting high-resolution microwave spectroscopy with gaseous species. This apparatus was used to quantify methanol from the headspace of

methanol/water solutions. Samples were prepared by mixing methanol and water and cooling the solution to the temperature of liquid nitrogen. Once this temperature was reached, the absorption cell and transfer lines were evacuated with a vacuum pump to a pressure of 0.1 mTorr. The solution was then allowed to warm to room temperature and small quantities of the vapor from this compartment were released into the absorption cell for analysis. Response of the instrument to methanol vapors was shown to be linear across 4 orders of magnitude of solution volume concentration.[199] Though not displayed graphically in this work, Dumesh et al. performed analytical measurements of carbon monoxide gas, obtaining a peak absorption coefficient of approximately 3×10^{-4} and detection limit ca. 1 ppm.[199]

CW-THz

Bigourd and coworkers published an account of their work with a CW THz system in which they quantified hydrogen cyanide (HCN) and carbon monoxide (CO) in cigarette smoke.[15, 17, 200] These species were quantified by determining the area of the absorption peak produced by the associated rotational transition, generating a predicted concentration of 210.3 ppm for HCN and 1.7 ± 0.3% for CO.[15, 17, 200] Formaldehyde and water vapor were also identified in the smoke stream based on their distinct spectral features.[15, 17, 200]

THz-TDS

Liu and coworkers published an account describing their work with ammonia vapors in 2004, using an EO sampling configuration. The emitter and detector elements were contained in an airtight chamber with four off-axis parabolic mirrors, which were used to direct and focus the THz beam.[201] Pump and probe laser beams were directed into the sampling chamber to generate and detect the THz pulses. Samples were prepared by evacuating the chamber with a vacuum pump, bringing the pressure in the chamber down to less than 20 mTorr. Ammonia gas was inserted in the chamber and the pressure

reading on a manometer was recorded, which served as a reference concentration for the following quantitative analysis. In some cases, humid air or nitrogen was added to the ammonia vapor to adjust the overall pressure of the sample. This process was repeated with varying pressures of ammonia to construct a calibration plot. Quantitative analysis was performed using the peak absorption from a rotational transition at 0.572 THz (19.08 cm^{-1}) both for pure ammonia and ammonia in a mixture with 100 Torr of humid air.[201] Relative error in the predicted pressures ranged from 5 to 30%.[201] Analysis of ammonia in humid air provided a simple mixture, as water vapor in the air exhibited overlapping spectral features near atmospheric pressure; however this overlap was reduced by performing analysis at relatively low overall pressure.[201] Analysis at decreased pressures should provide improved selectivity as a result of decreased spectral overlap, but this may not be practical in real-life atmospheric measurements as these pressures would only be observed at higher altitudes.

The impact of band broadening was also examined by adding varying amounts of air to 20 Torr of ammonia. At low pressures, the spectral peak widths were relatively narrow, but at higher pressures the ammonia absorption peak at 0.572 THz started to overlap with a nearby water vapor absorption peak at ~0.557 THz (18.58 cm^{-1}).[201] This spectral broadening can increase error in quantitative measurements if appropriate measures are not taken to characterize the species individually, or if sufficiently low pressures are not employed in sample preparation. For comparison, quantitative analysis was also performed with various pressures of ammonia at 590 Torr overall pressure, this time provided by dry N_2 rather than air. The absorption of ammonia appears to increase linearly in this case, though the absorption peak at 0.572 THz (19.08 cm^{-1}) appears very broad in the published spectra.[201]

Another set of quantitative analyses was performed using a method known as differential optical absorption spectroscopy, or DOAS. This method has been used widely for atmospheric measurements of gaseous species,[202-205] and can be used when an

adequate reference spectrum is not available. In this method, a single beam spectrum of the sample was acquired, and the reference spectrum was calculated by interpolating between assumed baseline regions. Absorbance values measured at known absorption peaks are referenced to this line, ideally providing suitable data for quantitative analysis.[201] The authors report similar prediction errors when using DOAS when compared to the previous quantitative efforts.

Jacobsen, Mittleman and Nuss have published accounts describing quantitative measurements utilizing THz-TDS.[206, 207] In these articles, the authors describe the use of an algorithm known as linear predictive coding (LPC).[208] This algorithm is somewhat reminiscent of the algorithm implemented by Andresen et al. in that it is used to condense spectral information with the ultimate purpose of simplifying and enhancing quantitative analysis. Time-variant digital signals may be extrapolated using LPC, utilizing a weighted sum of the past M data points, as shown in Equation I-8:

$$s(n) = \sum_{k=1}^{M} \alpha_k s(n-k) + e(n) \qquad \text{Equation I-8}$$

where s is the signal, $s(n)$ is the extrapolated portion of the signal, M is the set of previous values, $e(n)$ is the residual, and α_k is the set of parameters used for linear prediction, which describes the spectral content of the waveform.[206] Jacobsen et al. report inclusion of a transform in the coding algorithm, as shown in Equations I-9 and I-10:

$$S(z) = \sum_{k=1}^{M} \alpha_k z^{-k} S(z) + E(z) \qquad \text{Equation I-9}$$

where

$$z = exp\,(i2\pi f \Delta) \qquad \text{Equation I-10}$$

S is the signal, E is the error, i is the imaginary unit, α_k is the set of parameters used for linear prediction, and Δ^{-1} is the signal sampling frequency.[206] The variable f is not

defined in the text, but appears to represent frequency, given the context. The authors cite two advantages of this algorithm: (1) the algorithm can accurately represent data sets in which the spectrum is comprised of sharp features and (2) the same number of coefficients (M) can be used to represent the spectral data no matter how many resonances, or spectral peaks are present in the frequency-domain spectrum.[206] The effect of LPC with TDSs on the subsequent frequency-domain spectrum was illustrated with water vapor spectral features[206] and HCl spectral features.[206, 207] Optimal coefficients can be obtained by least-squares, more specifically through minimization of the sum squared residual shown in Equation I-11:

$$E = \sum e(n)^2 \qquad \text{Equation I-11}$$

where E represents the sum of squared residuals. Parameters used for LPC can be treated as a vector in M-dimensional space. A library is assembled by measurement and training with the species of interest, encoding a maximum of M-1 species' spectral features. Species recognition may be performed by calculating the Euclidean distance between the vector of an unknown species and those stored in the library. The vector which provides the minimum Euclidean distance is considered to be the best match.[206]

Mixtures may also be analyzed via LPC. The vector calculated from an unknown mixture is comprised of the sum of the vectors of the component gases, with the length of each component vector representing the respective fractions of each species.[206] An orthogonal basis is assembled and used to calculate the length of each component vector, which represents the fraction of each component. The authors applied this method to quantitative analysis of a binary mixture of water and ammonia vapors, but found that LPC provides similar performance to Fourier analysis.[206] Samples consisted of mixtures of the two components with pressures of ammonia ranging from about 0.3 to 12.9 kPa and pressures of water vapor ranging from 0 to 2.25 kPa.[206] The models appear to provide a high degree of correlation between the reference and predicted pressures.

Jacobsen et al. also reported the use of wavelet denoising to improve the SNR of experimental data.[206, 207] This pre-processing procedure can provide benefits in various applications. Wavelet denoising could be used to reduce the level of noise in signals collected with low levels of averaging, thus simulating the quality of spectral data that may otherwise only be obtained through lengthy periods of signal-averaging. This would permit the acquisition of TDSs of greater data point length (to provide frequency-domain spectra with narrow data point spacing) with lower signal averaging, thus shortening acquisition times. Wavelet pre-processing may also be particularly useful when employing LPC in that the original TDS waveform is minimally scarified but would contain lower levels of instrumental variance following wavelet processing, which would ideally reduce the variability in LPC-predicted concentrations.

Selected Solid Phase Spectroscopic and Imaging Applications

Interest in the THz frequency investigation of solid samples has expanded greatly in recent years as instrumentation has become available commercially. The broadband response of the instrument has prompted various studies that exploit interesting phenomena related to the physical and chemical properties of different sample types. As was previously mentioned, THz-TDS instrumentation is often combined with translation stage modules, further broadening the potential for scientific inquiry. The relatively low maintenance requirements and room-temperature operation of the instrumentation adds to the popularity of the method.

A description of all THz-TDS measurements of solid samples is beyond the scope of this introduction. Instead, a brief overview is provided for several popular solid sample applications based on the recent peer-reviewed literature. Emphasis is given to reports that describe analytical applications, though qualitative results are also included to illustrate the breadth of the method's capabilities and opportunities for future quantitative

investigations. THz-TDS has been widely applied to solid-phase samples. The advent of THz-frequency imaging using a time-domain system has revolutionized quality control in several industries, as well as security-related activities. Imaging systems that utilize THz radiation have been developed in both continuous wave (CW) and pulsed (time-domain) operational modes. Each system has advantages and disadvantages, mainly in the size of the system, the amount of information provided by the measurements, and the processing time for collected data.[209]

Explosives and Security

THz spectroscopic characterization of explosive materials has been a major theme of research to provide the analytical basis for imaging and noninvasive detection of explosives for security purposes. An overview of established methods used for spectral identification of explosives has been published.[210] Various explosives have been examined, such as 2,4,6- trinitro-1,3,5-trihydroxybenzene (TNPG),[2] 1,3,5,7-tetranitro-1,3,5,7-tretraazacyclooctane (HMX),[211] 1,3,5-trinitro-s-triazine (RDX),[24] 2,4,6-trinitrotoluene (TNT),[3] decomposition products of TNT 2,4-dinitrotoluene (2,4-DNT)[3]and 2,6-dinitrotoluene (2,6-DNT),[212] pentaerythritol tetranitrate (PETN),[213] hexanitrohexaazaisowurtzitane (HNIW),[3] Semtex H,[4] SX2,[4] Metabel,[4] PBX 9501 and PBX 9502,[214] and some common components of typical "homemade" explosives.[215] Spectra of several of these substances are shown below in Figures I-14 through I-17 to illustrate the appearance of spectral features associated with crystalline phonon vibrations of these materials.

THz pulse imaging has been demonstrated as an effective security tool through detection of concealed contraband.[210] Controlled substances concealed in a container such as a polyethylene (PE) bag can be identified by using a transmission-based measurement through the bag since PE is largely transparent at THz wavelengths.[216] The spectral features of the substance within are superimposed with the broad, low-level

absorption of the bag, making chemical identification fairly straightforward. Yamamoto et al. demonstrated the feasibility of noninvasive THz-TDS spectroscopic detection of explosives concealed within envelopes.[21, 22]

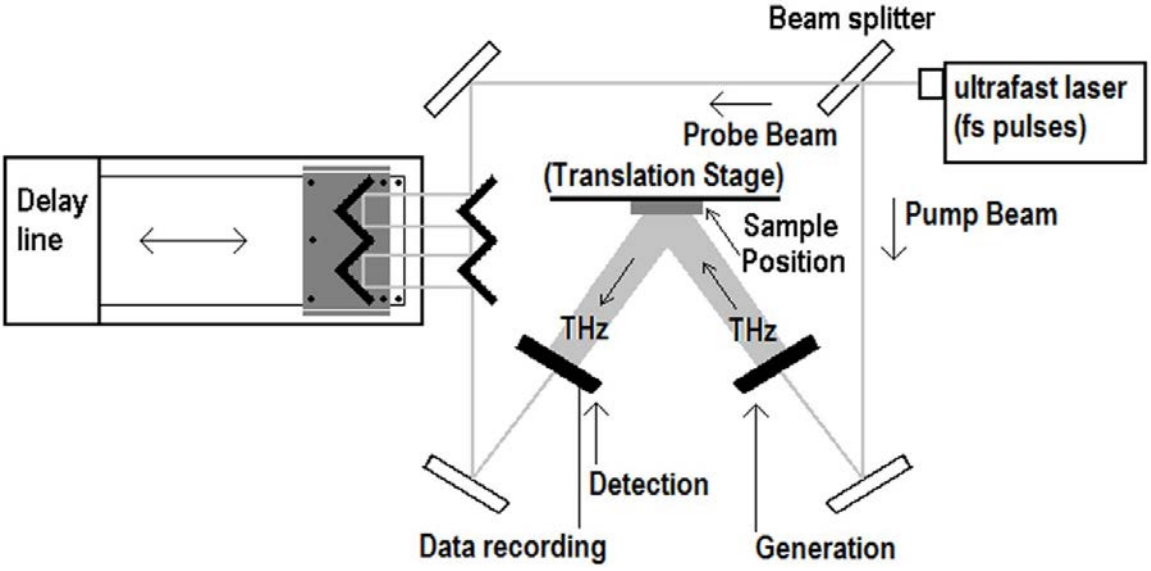


Figure I-13. Schematic diagram of the apparatus configuration for THz-TDS measurements in a reflectance geometry. In cases where a large sample is imaged, the emitter and detector may be mounted on a mobile gantry and coupled via fiber optic cables to a femtosecond laser.[217]

Imaging has also been performed with samples of uncontrolled substances, such as lactose and sucrose, compared with a sample of explosive material (RDX sheet).[218] In this work, powdered RDX was mixed in a PE powder matrix and compressed into a pellet. The PE powder acts as a binder,[218] providing structural integrity when sample powders are not sufficiently cohesive. The PE powder is a popular substrate for such investigations due to its low cost and high THz-frequency transparency, though alternative materials such as polytetrafluoroethylene (PTFE) powder may also be used. Initially, the RDX/PE pellet was analyzed spectroscopically with both THz-TDS and FT-IR spectroscopy to identify useful spectral features and compare them to previously

published results. Several pellets were prepared using lactose and sucrose powders, which were analyzed with the THz pulsed imaging system, along with a small piece of RDX sheet. While the raw TDS provided information describing the distribution of chemical components (as was determined by the magnitude of the main peak), the identification of species was performed through the calculation of absorption spectrum for each pixel in the THz image analyzer. By comparing the spectra of the materials to the image spectra, species in each sample were correctly identified, with clear distinction between RDX and the two sugars. Studies such as this ascertain the possibility for using THz spectroscopy to discern common materials from dangerous substances that might be encountered at a security check point, thereby establishing the utility of this approach.

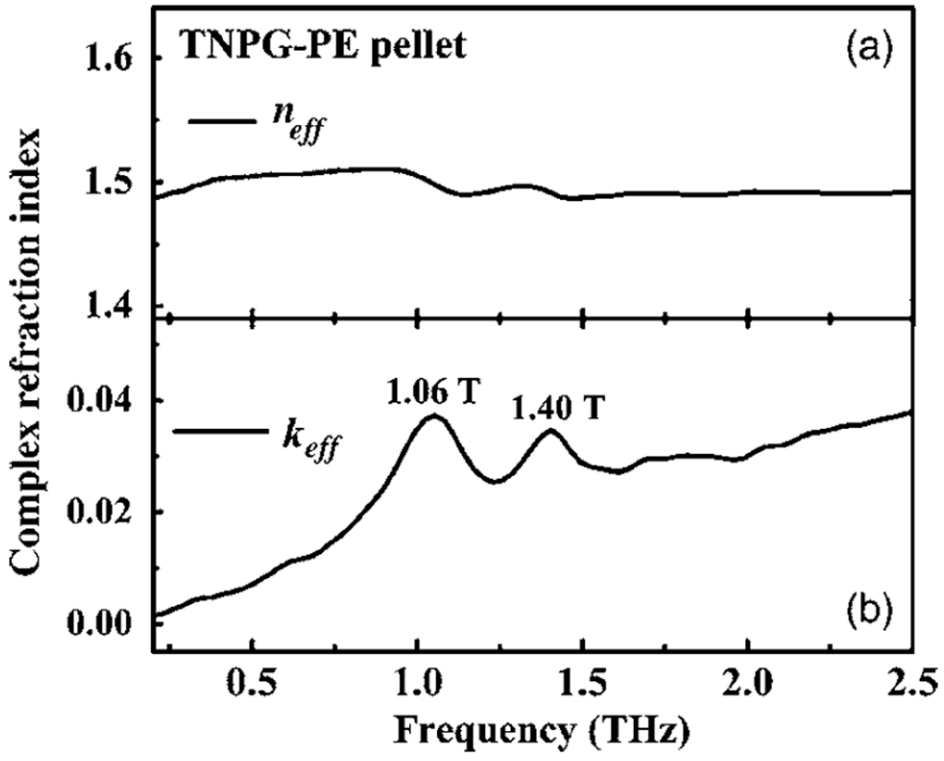


Figure I-14. THz refractive index and absorption spectrum of solid TNPG diluted in a PE pellet. Taken from[2] with permission.

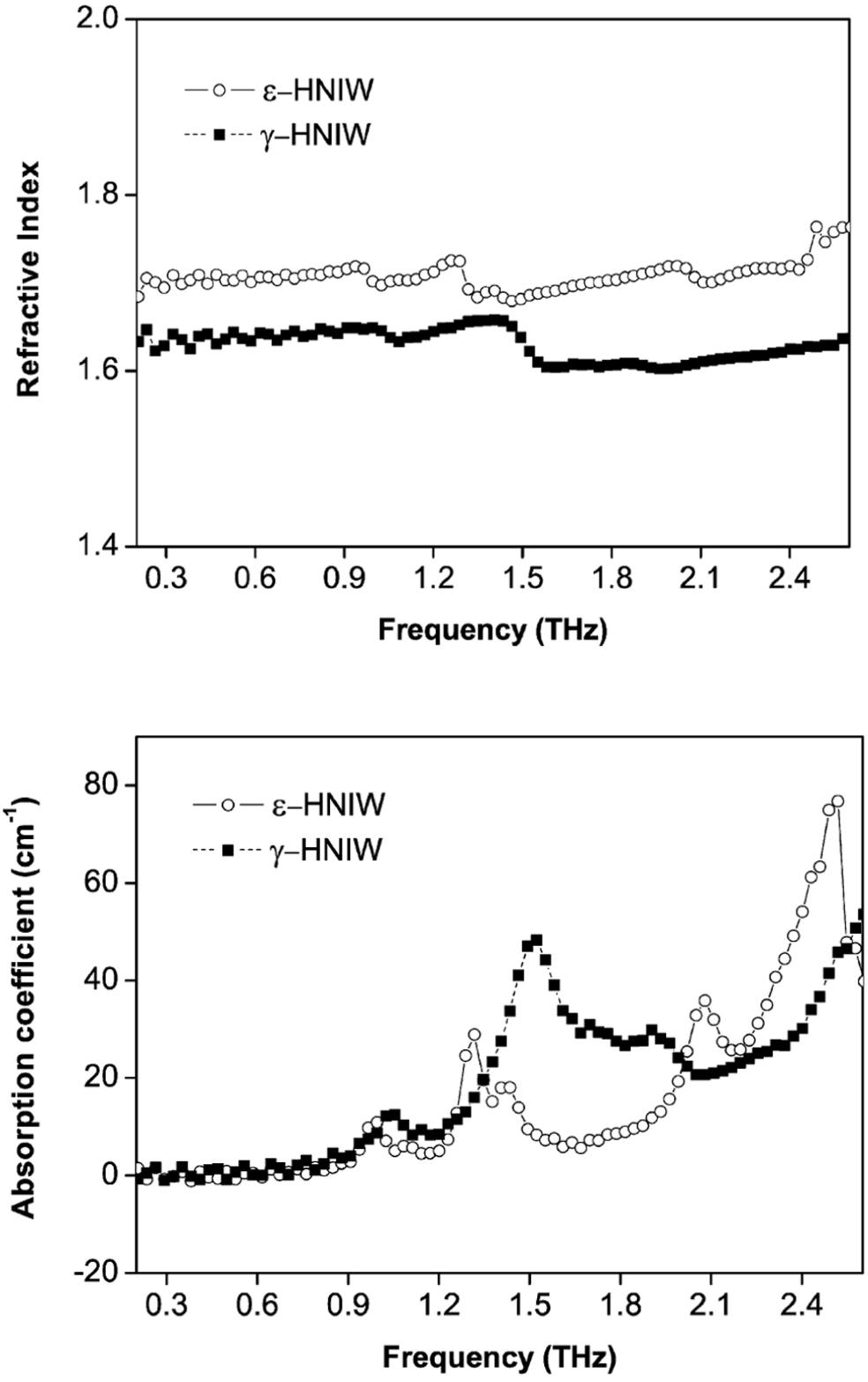


Figure I-15. THz refractive index (top) and absorptivity (bottom) spectra of solid ε- and γ-HNIW compressed to form pellet samples. Taken from[3] with permission.

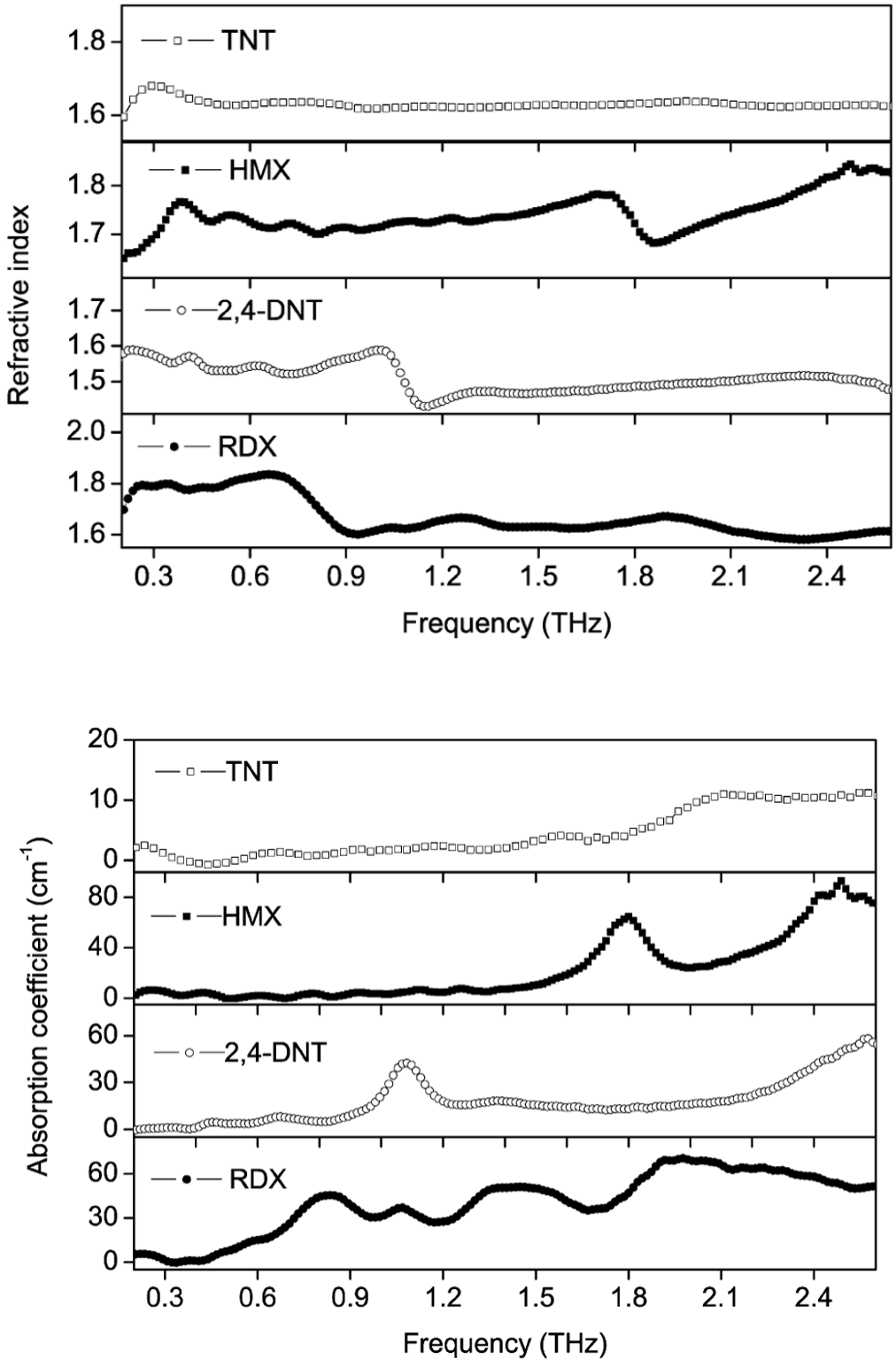


Figure I-16. THz refractive index (top) and absorptivity (bottom) spectra of solid TNT, HMX, 2,4-DNT, and RDX compressed to form pellet samples. Taken from[3] with permission.

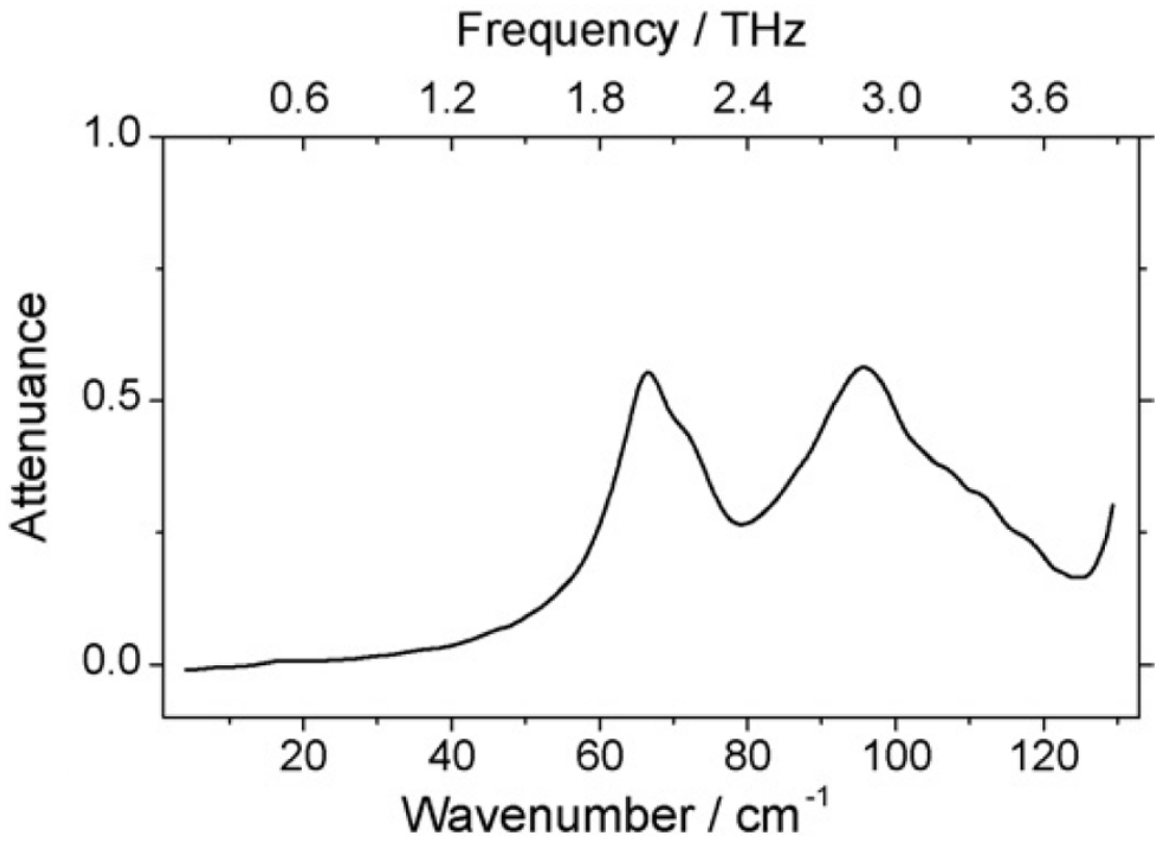


Figure I-17. THz absorption spectrum of solid PETN diluted within a PE pellet. Taken from[4] with permission.

The utility of THz imaging has been demonstrated for detection of concealed threats. Items contained either within luggage or underneath common clothing can be revealed by THz imaging. THz-TDS images taken through a briefcase and a leather shoe have been published,[5] and images of weapons and other contraband hidden within luggage and clothing has also been reported.[1] Such measurements are not exclusively limited to the detection of concealed metallic objects. Data produced by THz-TDS instrumentation is by nature sensitive to refractive index because of its time-resolved measurements. Objects fabricated with a high refractive index material will induce greater delay in transmitted radiation or reflect a greater amount of radiation in a reflection measurement. As such, concealed items made of glass or ceramics can also be

detected and imaged using this approach. Even items like plastic toy guns have been shown to be detectable with THz-TDS imaging.[219] Threats can be identified on the basis of spectral features as well as the shape of the imaged object. Instrumentation of this sort is attractive for security-based screening because of its sensitivity to materials such as these. Items like ceramic weapons that are undetectable to metal detection systems would be immediately visible with security systems armed with THz imaging instrumentation. Figures I-18 through I-21 illustrate THz-frequency images for the detection of concealed threats.

Spectroscopic Analysis of Pharmaceutical Solids

The promise of using THz-TDS spectroscopy for characterization of materials within the pharmaceutical industry has been demonstrated. The efficacy of pharmaceutical products can be impacted by the extent of hydration of the active ingredient.[6] Besides changes in mass upon adsorption of water, the degree of hydration of components within a tablet can also be assessed from changes in the THz spectrum. The sensitivity of THz-TDS spectroscopy to crystal lattice vibrations and intermolecular interactions make such measurements appealing for monitoring hydration within pharmaceutical preparations. Research has been published in which THz-TDS is used for the analysis of several pharmaceutical compounds and their hydrates.[6] The THz-TDS measurement was deemed useful not only for analyzing the substances during dehydration, but also for detection of water that had evaporated from the solid and the subsequent phase change in the crystalline structure of the drug at different temperatures. Figure I-22 shows the absorption spectra of carbamazepine (CBZ), lactose, theophylline (TP), and piroxicam (PXM) hydrates. Figure I-23 shows the change in the THz absorption spectrum of TP during heating, illustrating the effect of dehydration on spectral features, as well as the spectrum of water vapor leaving the substance.

Analytical information ascertained from THz-TDS spectroscopy was compared to

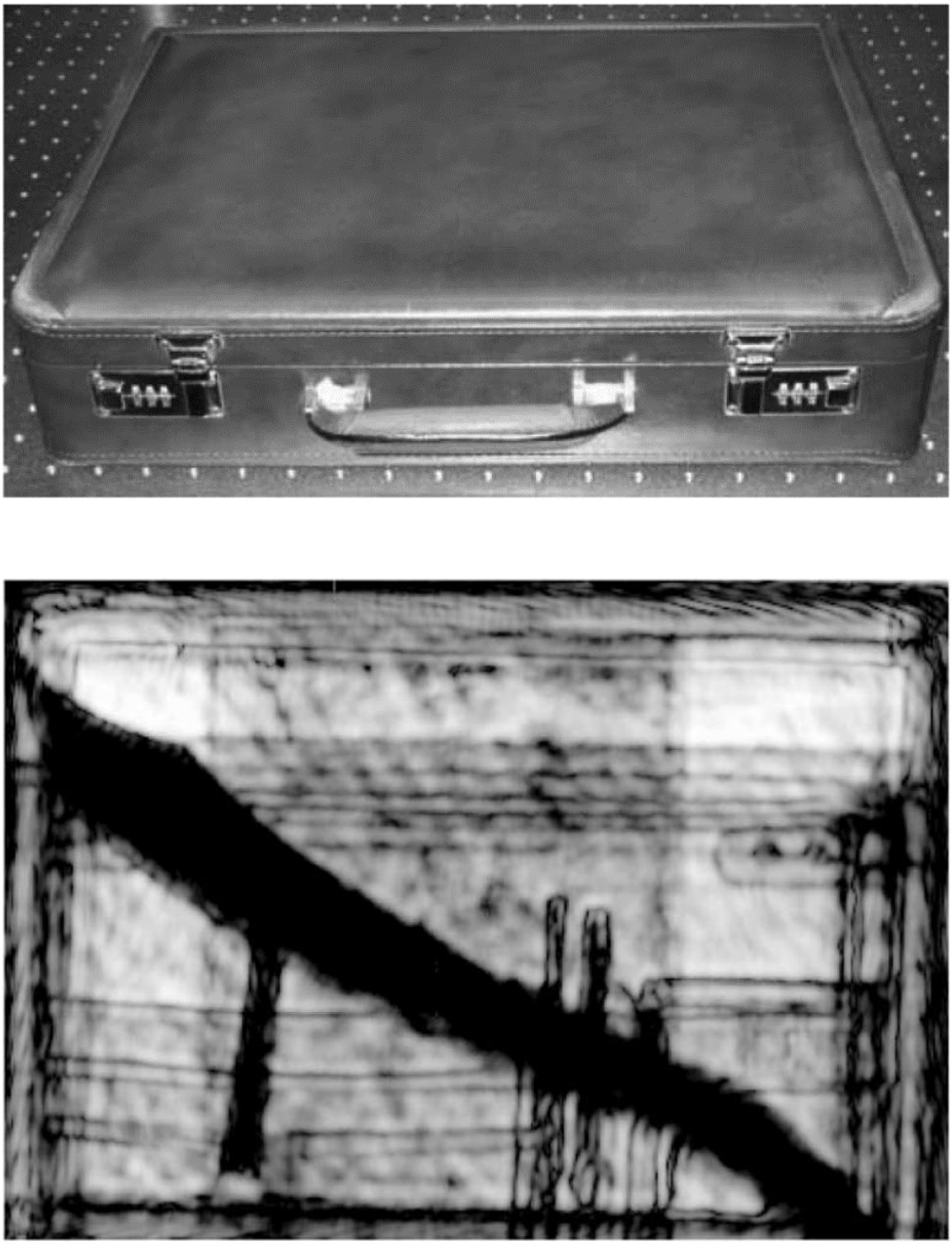


Figure I-18. Visible (top) and THz-frequency (bottom) images of a leather briefcase. A knife, magazine, and pens are visible in the 0.2 THz monochromatic image collected through the briefcase. Taken from[5] with permission (©2007 IEEE).

Figure I-19. Visible (top) and THz-frequency (bottom) images of a man's shoe. The interior structure of the shoe is visible in the 0.2 THz monochromatic image, highlighting stitching and the cushioned heel. Taken from[5] with permission (©2007 IEEE).

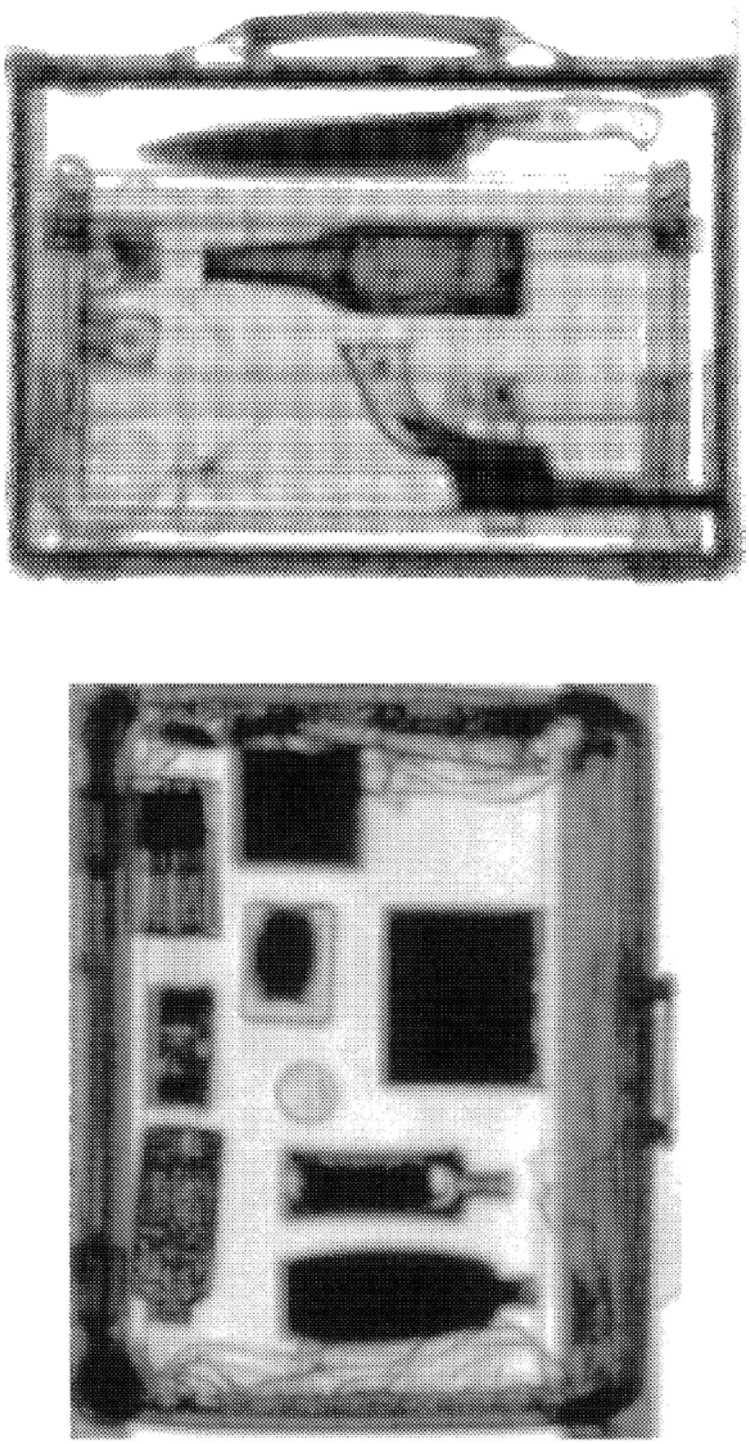


Figure I-20. THz-frequency images of an attaché case (top) and suitcase (bottom) containing various concealed threats. Taken from[1] with permission (Time domain terahertz imaging of threats in luggage and personnel, Zimdars, D., White, J., Stuk, G., Fichter, G., and Williamson, S. L., Int. J. High Speed Electron. Syst., ©2007 World Scientific Publishing Company, and Picometrix LLC).

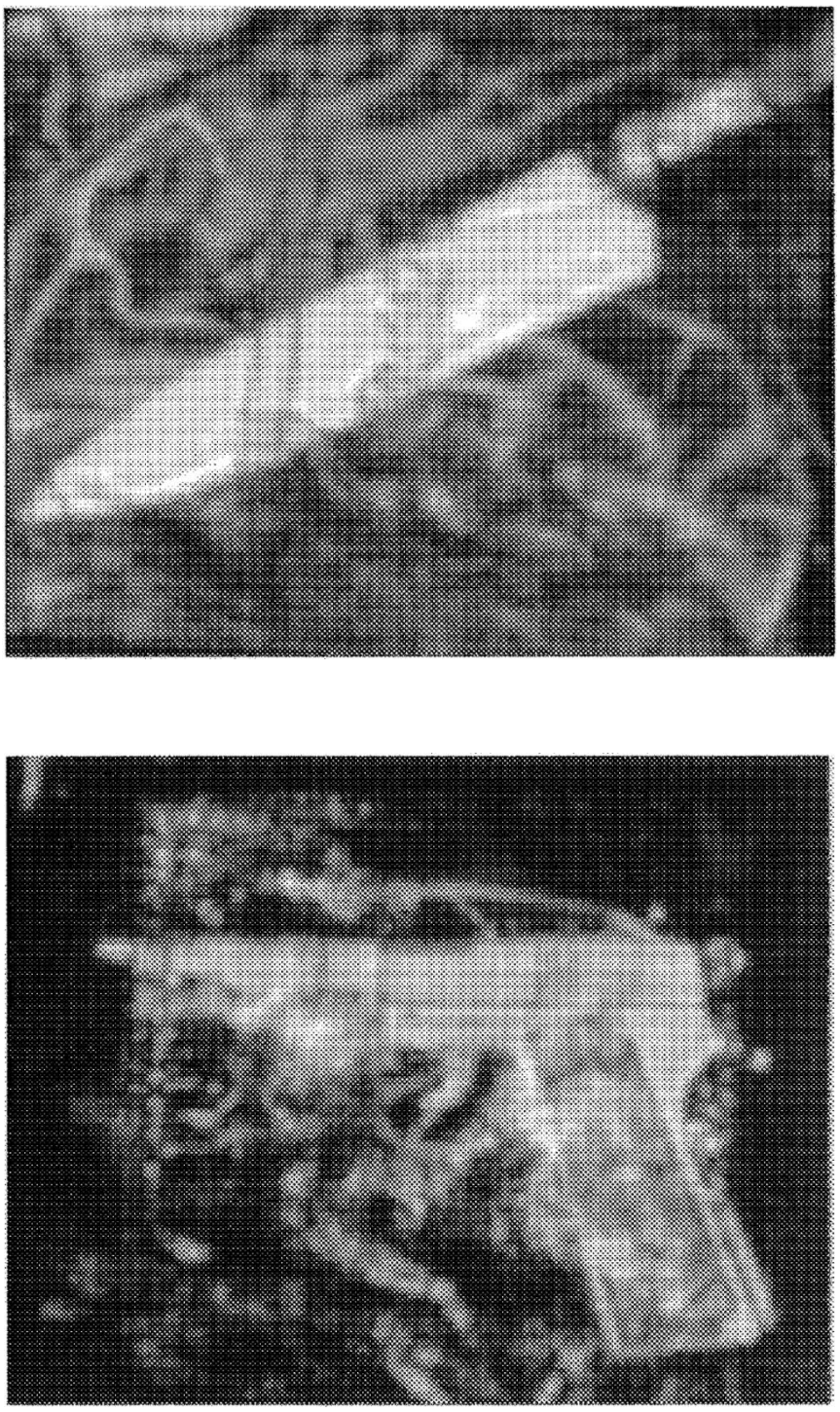


Figure I-21. THz-frequency images of a knife (top) and handgun (bottom) concealed within the pocket of a jacket. Taken from[1] with permission (Time domain terahertz imaging of threats in luggage and personnel, Zimdars, D., White, J., Stuk, G., Fichter, G., and Williamson, S. L., Int. J. High Speed Electron. Syst., ©2007 World Scientific Publishing Company, ©2007 World Scientific Publishing Company, and Picometrix LLC).

conventional methods of analysis in the pharmaceutical industry.[220] In comparison to variable-temperature x-ray powder diffractometry, NIR spectroscopy, Raman scattering spectroscopy, and thermogravimetric analysis, the THz-TDS method was deemed superior for the characterization of piroxicam. For these THz measurements, the piroxicam (anhydrous and hydrate forms) were mixed with PTFE and compressed into pellets for spectral analysis.[220] Several pellets were prepared in which the distribution of constituents was varied; one pellet was prepared in which the components were mixed, another was prepared with a high concentration of piroxicam at the surface, and another with piroxicam concentrated in the center (by depth) of the pellet. These pellets were analyzed to investigate the impact of sample heterogeneity on the appearance and quality of spectral features. Data collected with these pellets also provided some indication of the sampling depth of each method, as spectra collected with low penetration into the sample material would contain diminished (if any) absorption features when analyzing the analyte-centered pellet.

In the experiment described above, the THz-TDS spectral results were used to characterize the hydration state of the compound, and subsequently monitor the effect of dehydration on the THz signatures of PXM by acquiring spectra at temperatures ranging from 303 to 423 K.[220] Raman and NIR spectra were also collected with these samples at various temperatures (303-456 K) for comparison. The reported findings indicate that Raman and NIR spectroscopic measurements were limited by low sampling volumes inherent in the respective measurements and these conventional methods were hindered by interaction of the radiation with the polymer material used to dilute the drug sample.[220] X-ray powder diffraction measurements are commonly used to establish the structure of crystalline pharmaceuticals, but THz-TDS measurements have the advantage of being safer, faster, and non-destructive. Information obtained from the THz spectra could be used to follow dehydration as effectively as thermogravimetric analysis, while also providing structural information about the solid-state properties of the sample, which

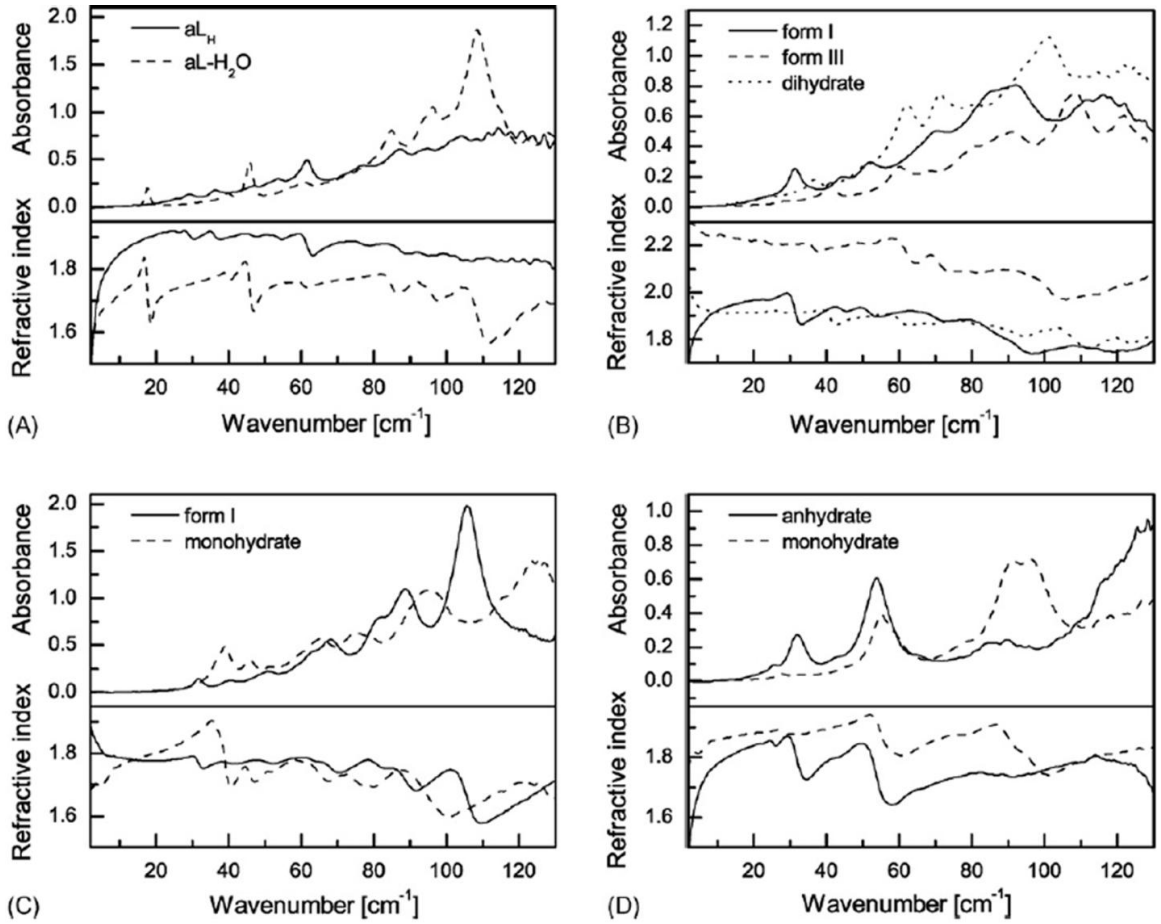


Figure I-22. THz absorption spectra of α-lactose (A), CBZ (B), PXM (C), and TP (D) collected at 293 K. Hydrates and other crystalline forms are denoted via differing line styles. Taken from[6] with permission.

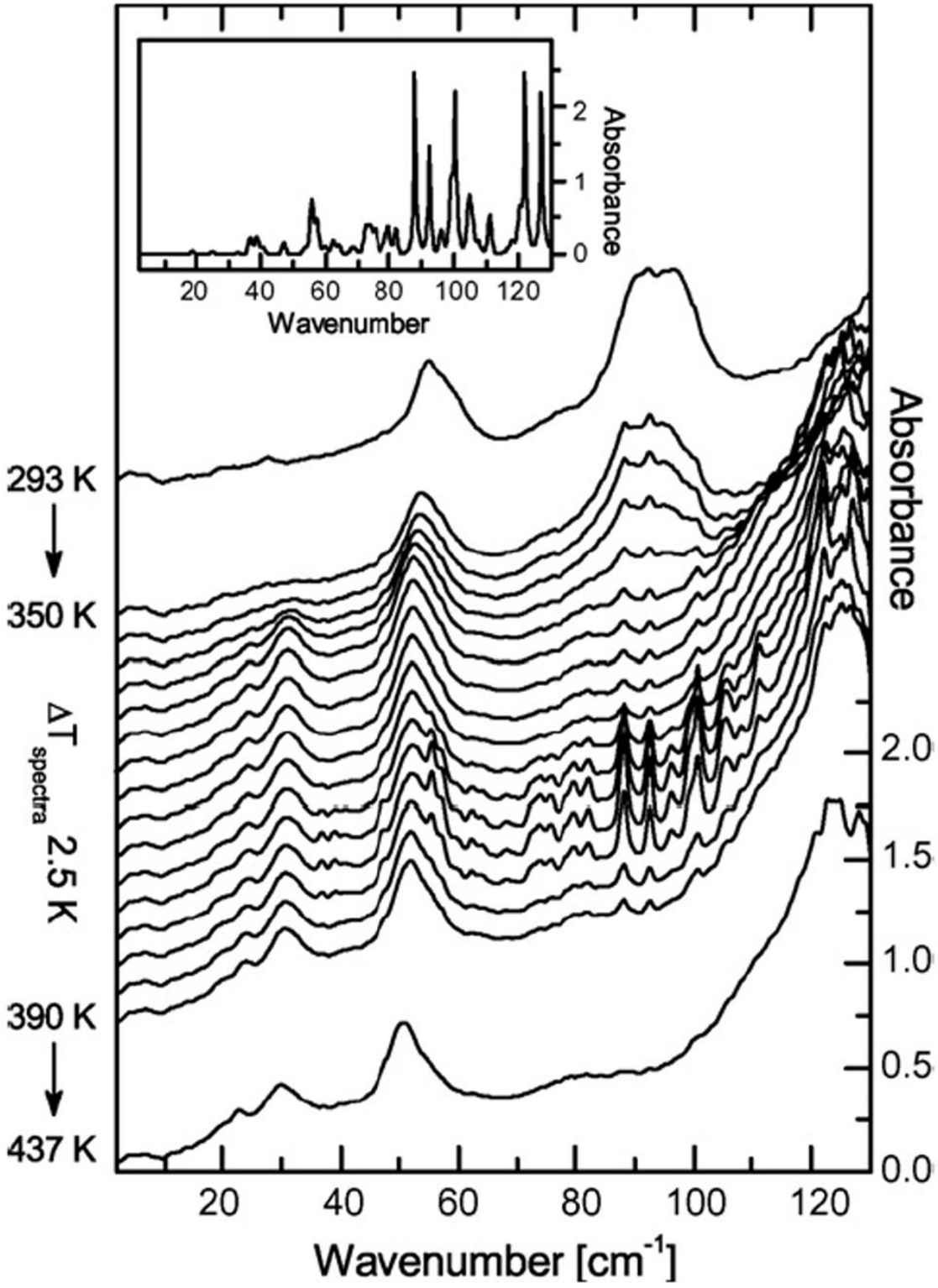


Figure I-23. THz absorption spectra of TP collected at different temperatures, ranging from 293 to 437 K. A THz absorption spectrum of water vapor is provided in the inset plot for comparison. Taken from[6] with permission.

cannot be obtained from thermogravimetric measurements.

THz-TDS has also been applied to identify different polymorphs of pharmaceutical crystalline materials. Polymorphism refers to the property of a compound to form different molecular crystalline conformations. The lowest energy crystalline conformation for a given pharmaceutical agent should provide the longest and most effective lifetime.[221] In addition, different polymorphs for a given material can have different dissolution properties, thereby affecting the rate of dosage delivery upon ingestion. Both active ingredients and excipients must be considered in such experiments. Polymorphs contain different geometric arrangements of the molecular constituents of the compound, thereby altering the vibrational properties of the crystal lattice. Such differences in the vibrational properties of the crystal lattice appear as different spectral features within the THz-TDS spectral data. Simple and fast THz spectra can be used to monitor changes in the crystalline structure of pharmaceutical compounds and to establish a fingerprint spectrum for polymorph structures. Such results have been demonstrated with pharmaceutical materials such as enalapril maleate,[222] fenoprofen calcium dehydrate,[222] indomethacin,[222] carbamazepine,[222] and theophylline.[223] As dissolution properties associated with various polymorphs are characterized, the method can be used to non-destructively assess the efficacy of pharmaceutical products by analysis of these fingerprint spectra.

THz-TDS has also been proposed to characterize carbohydrates and polysaccharides. For example, the structural conformation of β-glucan (an anticancer agent) was analyzed by comparing the THz spectra of the two major structural conformations, a single-stranded helix and a triple-stranded helix. Each conformation was found to display different optical properties of refractive index and dielectric constant over the investigated THz frequencies.[224] THz-TDS has also been demonstrated to provide selective identification and quantitation of multiple carbohydrate components in pharmaceutical products. For example, lactose monohydrate, starch, microcrystalline

cellulose, and theophylline were identified and quantified in tablets with the assistance of chemometric tools.[225]

Imaging Analysis of Pharmaceutical Solids

The sensitivity of the THz-TDS measurement to refractive index has been exploited within the pharmaceutical industry to image the physical structure of individual tablets. For these measurements, a narrow band of THz radiation is incident upon the surface of the tablet. A portion of this incident radiation is reflected at the air/tablet interface and detected. Some of the incident radiation passes through the outer coating layer of the tablet and interacts with a second interface that exists between the tablet coating and the core material within the tablet. A fraction of the radiation that hits this interface is reflected and detected. Figure I-24 illustrates the type of THz-TDS obtained for this type of measurement. Inspection of the time-domain signal reveals the initial radiation pulse peak corresponding to the reflection at the first air/coating interface as well as a second, much smaller, pulse that corresponds to the reflection at the second coating/core interface. The time delay between these two pulses is related to the thickness of the coating and the refractive index of the coating material. Thus, for a given coating material (constant refractive index), a thicker tablet coating will result in a larger time delay between the initial pulse and the portion reflected from the sub-surface interface.

Ho et al. have published several articles describing the application of THz imaging based on time-domain measurements.[7] Using a commercially available THz-TDS scanning reflectance instrument,[7] data were collected across the surface of the tablet, after which the data were assembled to produce a 3-dimensional image that represents the coating thickness across the tablet surface.[7] It is desirable to know the thickness of the tablet coating, as this is directly related to the time required for dissolution when the tablet is ingested.[226] Characterization of the coating thickness and

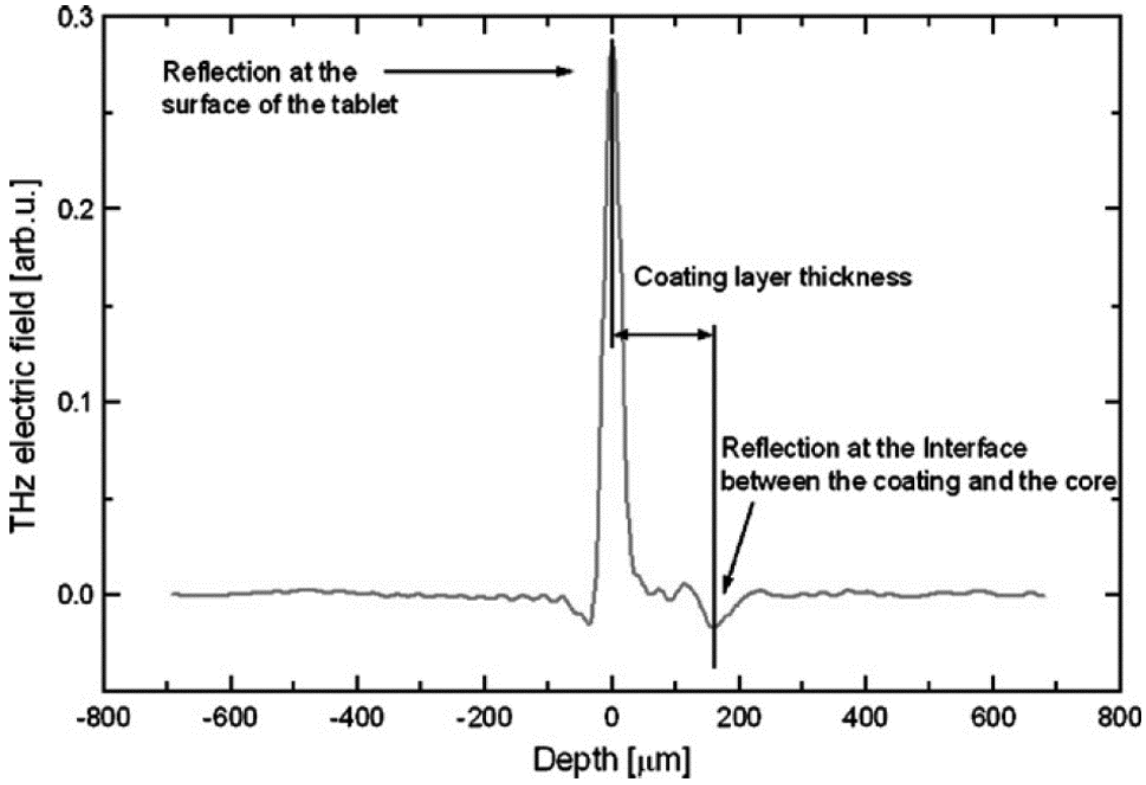


Figure I-24. Illustration of a time-domain signal collected as reflected from the surface of a pharmaceutical tablet. The largest feature centered at a depth of 0 microns is the portion of the THz pulse reflected from the outermost surface of the tablet. A significant dip can be seen near 160 microns which is the portion of the THz pulse reflected from the interior surface of the coating, at the boundary between the coating and the filler/excipient core. Taken from[7] with permission.

its uniformity can provide valuable diagnostics for enhanced quality control.

An interesting comparison between THz-TDS and FT-NIR tablet imaging has been published.[227] The NIR reflectance imaging data were collected with a focal plane array instrument designed to produce spatially resolved NIR spectra across the tablet surface. The THz-TDS image was shown to be more practical, because the coating thickness could be determined more directly and rapidly.[227] The NIR images provided some idea of coating thickness, as absorption features related to tablet core constituents became more prominent as coating thickness decreased. Overall, THz-TDS imaging

provides a powerful method for coating characterization compared to traditional methods of tablet analysis. The THz-TDS method is non-destructive and fast, whereas the more traditional methods require dissolution or dissection of the tablet, or provide indirect measurements of coating thickness.[228]

A review published in 2007 describes additional applications of THz-TDS measurements in the pharmaceutical industry,[229] with a focus on tablet layer uniformity and composition. Because THz radiation can penetrate deep into tablets, structures concealed by outer layers can be distinguished.[229] Another feature of using THz-TDS measurements to characterize tablets is the ability to analyze pharmaceutical products through packaging material for quality assurance. This last feature takes advantage of the high transparency of many polymers to THz radiation.[229] A more recent review further embellishes on developments in the characterization of crystalline polymorphs, quantitative efforts, hydration, and sample preparation.[46] Applications of THz imaging to tablet coating integrity, tablet hardness, component distribution are also reviewed.[46] Additional sources of related information are available.[230]

Dental Tissues

Research findings demonstrate the utility of THz spectroscopy and imaging in identifying abnormalities in biological tissues in general, and dental tissue in particular. Initial publications describe the proof-of-concept for dental THz measurements.[231] This initial feasibility work illustrates the potential utility of both transmission and reflection measurement geometries in identifying carious regions of the tooth structure. Reported results indicate that carious enamel produced ~35% higher absorptivities relative to non-carious enamel measured by transmission spectroscopy over the THz frequencies. In comparison, there is only a 2% difference in absorptivities between carious and normal enamel observed with clinical dental radiographs.[231] This preliminary work was expanded to a 3-dimensional reflection-based imaging for determination of enamel

thickness.[232] In addition, THz-TDS imaging is able to detect the dentine-enamel interface on the interior of the tooth, with a success rate of 91%.[232] Through the detection of this interface, the thickness of the enamel could be determined to within 10 μm of the expected value.[232]

Other dental applications involve direct comparison of THz pulse imaging (TPI) to transmission microradiography (TMR).[233] Results from this study reveal a good correlation ($r^2 = 0.995$) between TMR and TPI data, though the TPI measurements systematically indicated lesion depths shallower than those indicated by TMR.[233]

THz examination of dental tissue physiology has also been investigated. THz analysis of tooth structure was accomplished by transmitting a broadband pulse of THz radiation through an aperture of 0.3 mm to provide spatial resolution at specified locations on the tooth surface.[234] The tooth was placed against this aperture to produce region-specific spectra.[234] It was found through experimentation that different structures on the interior of the tooth produced different spectral characteristics in the 0 – 83.4 cm^{-1} range (0 – 2.5 THz).[234] Furthermore, it was found that the orientation of these internal tooth structures relative to the direction of the propagating THz radiation had an effect on the collected spectral features.

The peer-reviewed literature demonstrates the utility of THz spectroscopy and imaging for dental research and possibly clinical practice. Still, much research and development is needed to realize the potential of THz-TDS for practical *in vivo* dental applications.

Biological Tissues

Various applications of THz spectroscopy and imaging analyses have been examined. Examples presented below include distinction of different layers of skin, distinction of different types of tissues, and the distinction of cancerous versus

noncancerous tissue. In all cases, the use of THz-TDS methods for such analyses is in its infancy and these early measurements provide a proof of concept.

A major factor associated with THz measurements in biological samples is the strong absorption of THz radiation by water in its liquid state. Water drastically limits the penetration depth of THz radiation into the matrix of most biological samples. The THz spectrum of water is a broad, featureless strong absorption band that spans the entire range of THz frequencies.[235] The limited penetration depths of biological samples greatly favor reflectance measurements over transmission measurements. Even in a reflectance geometry, however, THz measurements are generally limited to only one millimeter into the biological matrix.

THz spectroscopic measurements of skin are reported in both transmission and reflectance modes[236] with measurements on skin originating from humans,[237] pigs,[238] and rats.[8] For experiments with human skin, *in vivo* skin measurements are sensitivity to variations in skin hydration, sun exposure, irritation, and density of hair, all of which have been shown to alter THz spectra collected from the same subject.[237] The best reproducibility for within subject spectra was achieved by collecting *in vivo* spectra from the volar forearm, as this area receives little exposure to sunlight and is covered with less hair than most other regions of the body.[237] Spectral differences are also significant between skin measurements from different human subjects, depending on the thickness and composition of the epidermis and dermis layers.[237]

THz imaging is compared directly to infrared imaging for biological samples.[238] For this comparison, pulsed THz images are collected and compared to images collected using a coherent optical time-gated technique similar to that described in.[239] The same ultrafast laser system was used for THz generation and acquisition of the near-IR image. Overall, the THz images were deemed superior on the basis of lower influence of scattering by the tissue matrix.[238] In addition, a 2 x 2 mm^2 Z-shaped metal structure (0.3 mm thick) embedded in the pork fat tissue was also discernible with the TPI system,[238]

indicating the possibility of using THz imaging techniques to locate and monitor superficial prostheses.

Research has shown that THz imaging can also distinguish between structures and layers within biological specimen. Work published by Hoshina and co-workers illustrates the ability to discern adipose tissue from muscle tissue with TPI measurements.[240] The difference in refractive index and absorptivity makes the two tissue types readily distinguishable without the use of complicated signal processing software.[240] Measurements performed with freshly harvested rat tissues and organs also produced similar results. Major features of the TDS, such as amplitude and width of the main peak could be used to discern the different organs and tissue types, with the exception of heart and pancreas tissues.[8] Frequency-domain values for absorptivity and refractive index were deemed to be statistically different for the various tissue types, suggesting the possibility of using THz spectra for qualitative identification of unknown tissue samples.[8] Figure I-25 illustrates THz reflectance data reported for a sample of heart tissue and Figure I-26 shows THz-TDS results reported for different types of biological tissues. Table 1 lists the TDS main peak amplitude and full-width at half-maximum (FWHM) and corresponding standard deviations for signals shown in Figure I-26.[8]

The ability of THz imaging to distinguish cancerous from noncancerous tissue has been demonstrated in both *in vivo* and *ex vivo* experiments.[241] Efforts to diagnose basal cell carcinoma (BCC) are particularly noteworthy, where BCC is a type of skin cancer that affects approximately 1 million people in America annually.[242] Currently, a method known as Mohs' Micrographic Surgery (MMS) is used to locate the tumor and determine the extent of its growth.[243] This method involves the analysis of excised samples during the removal procedure. As such, this method can be costly, time consuming, and uncomfortable. THz imagine promises to be a rapid (< 5 minutes for a 25 x 25 mm^2 scan[241]), non-invasive alternative to MMS, with a few limitations. The THz approach

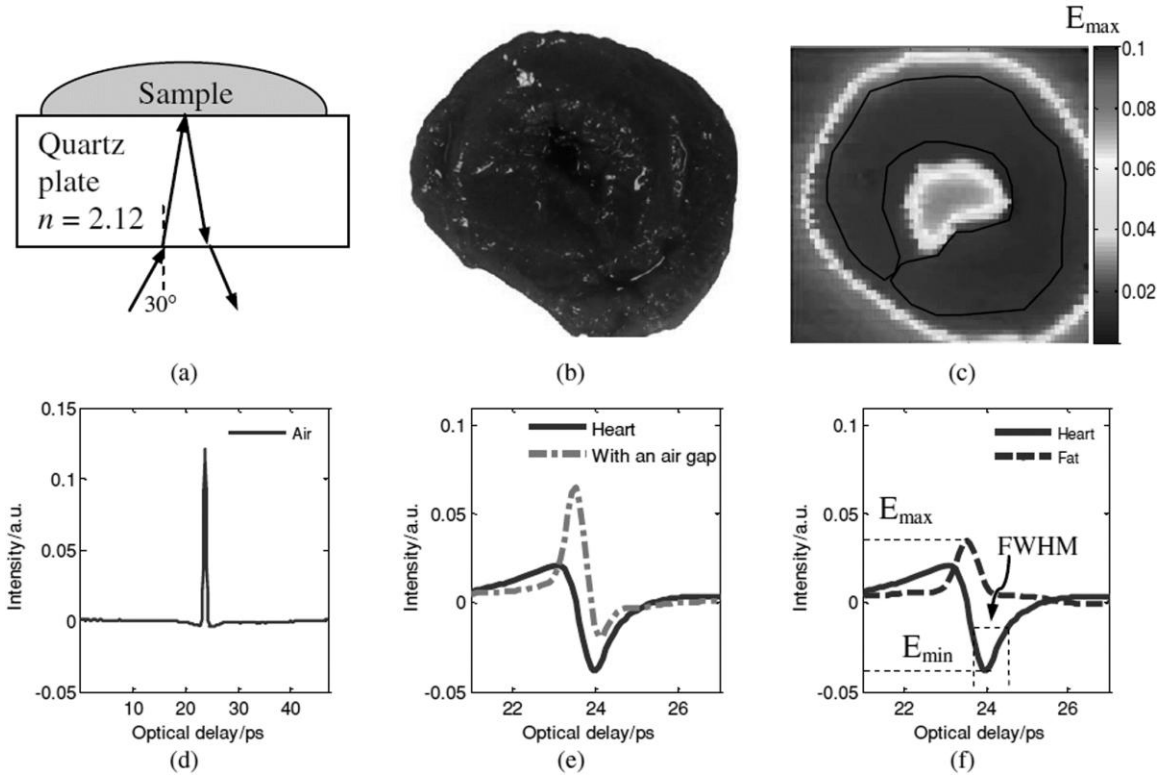


Figure I-25. Measurement geometry used for the THz analysis of rat tissues (top left). A visible image of the rat heart tissue sample is shown at top center, and the THz-frequency image of the same sample is illustrated at top right. Intensity in this image is determined by the measured value of the TDS peak (maximum) after reflection off the tissue. TDS data are shown as collected with air (bottom left), heart and fat tissues (bottom center). The plot at bottom right illustrates the procedure used to characterize the different TDS data. Values obtained from analysis of the TDS are available in Table 1. Taken from[8] with permission (© 2009 IOPP).

is insensitive to tumors growing beneath the skin owing to water absorption and the corresponding 1 mm penetration depth of THz radiation into skin.[241, 243, 244] In addition, necrotic cells can confound demarcation of the cancerous tissues.[245]

TPI has also been proven effective for the identification of breast tumors.[246] Excised tissue samples were sliced to expose suspected tumorous regions and were placed on a quartz imaging plate. Slight pressure was applied across the tissue samples to promote intimate contact between the imaging plate and tissues, thereby enhancing the

interrogation of tissues by THz pulses. Images were acquired by transmission of THz pulses through the quartz window and subsequent reflection from the tissue/quartz interface. Other methods such as confocal microscopy, optical coherence tomography, and high-frequency ultrasound are sometimes used for similar measurements of tumorous tissues, though these methods are often limited in application by shallow penetration depths or the degradation of spatial resolution deep in the tissue. At the very least, TPI instrumentation is a rapid, non-destructive supplement to other methods used to identify suspected cancer tissues.

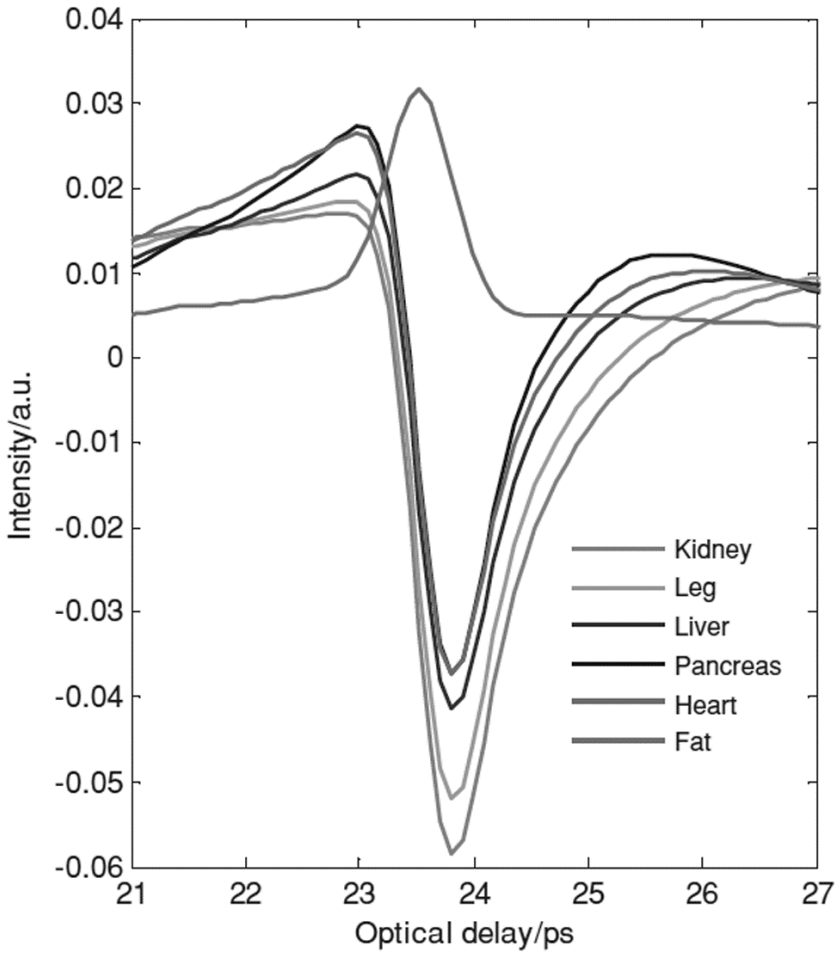


Figure I-26. Time-domain signals recorded for different biological tissues. Taken from[8] with permission (© 2009 IOPP).

Table I-1. Characteristics of time-domain signals recorded for different tissues.

Tissue	Amplitude (A.U.)	FWHM (ps)
Kidney (E_{min})	-0.0584 ± 0.0021	0.7508 ± 0.0382
Leg (E_{min})	-0.0520 ± 0.0055	0.7185 ± 0.0408
Liver (E_{min})	-0.0413 ± 0.0048	0.6431 ± 0.0597
Pancreas (E_{min})	-0.0372 ± 0.0037	0.5455 ± 0.0697
Heart (E_{min})	-0.0372 ± 0.0049	0.5879 ± 0.0724
Abdominal fat (E_{max})	0.0329 ± 0.0080	0.8274 ± 0.3781

Adapted from Huang, S. Y.; Wang, Y. X. J.; Yeung, D. K. W.; Ahuja, A. T.; and Zhang, Y.-T.; Pickwell-Macpherson, E. *Phys. Med. Biol.* **2009,** 54, 149-160. with permission.

Polymers

THz-TDS has also been used for a variety of applications related to polymeric compounds. Research has been conducted indicating the utility of THz spectroscopy for quality control of polymeric substances. An example is the detection of photo-oxidative damage to polymer materials, where the THz absorptivity of ultra-high molecular weight polyethylene was found to increase as the polymer was exposed to gamma-radiation.[247] This information was used to confirm that vitamin E decreases the extent of polymer damage caused by the gamma-radiation.

THz imaging has been demonstrated useful in imaging polymer-based compounds, such as low-density polyethylene (LDPE) with Ag-TiO_2 filler material, a glass fiber laminate material, and the cover from an automotive airbag.[248] The sensitivity of the instrument to difference in refractive index is clearly demonstrated, allowing for the identification of thickness variations in these samples, which can indicate areas of potential weakness within the structure. Unfortunately, the experiment with the additive-doped LDPE was somewhat inconclusive, since the sample thickness was small, and the additive content was low.[248] These experiments illustrate how refractive index and

absorption spectra can be confounded with artifacts associated with scattering and multiple reflections as the THz radiation propagates through the material of interest. In this published report,[248] the degree of absorption and dispersion was found to vary across the material's length, but it was unclear whether the variation resulted from the heterogeneous dispersion of filler within the material or the agglomeration of additive particles.

The application of THz-TDS to polymer manufacturing has also been studied. Wietzke et al. studied the ability of the THz-TDS system to determine the amount of $Mg(OH)_2$ dispersed within preparations of linear low-density polyethylene, $CaCO_3$ in polypropylene, glass fibers and other additives in polyamide 66, and Si or SiO_2 in polypropylene. Quantitation of the additive was primarily established by measuring the refractive index of the additive modified polymer.[249] The refractive index increases linearly as a function of the amount of additive.[249] Subsequently, these THz measurements were demonstrated for in-line monitoring capable of real-time measurements of the molten polymer as it was formed.[250] Again, refractive index was indicative of the amount of additive within the polymer, although in this case the refractive index was lower than its solid counterpart because of the inherent difference in density.[249] This set of measurements emphasizes the attractive ability of THz-TDS measurements to rapidly acquire absorptivity and refractive index data from polymer samples.

The transparency of many polymers renders these polymers well-suited to use as optical windows and sample matrix elements for THz measurements. Examples include polymethylpentene, Zeonor (a cyclo-olefin polymer), and polydimethylsiloxane, which is a material commonly used in the construction of microfluidic devices.[251] The THz properties of polystyrene foam were examined,[252] indicating that polystyrene may also be a suitable transparent material for THz spectroscopy.

Examples are also prominent in the literature illustrating the potential of THz methods to characterize physical and chemical properties of polymers, such as solvent diffusion[253] and hygroscopicity.[254] For example, reflection measurements were used to track the movement of acetone through polycarbonate and polyvinylchloride,[252] the results of which were in agreement with data from nuclear magnetic resonance (NMR) measurements. Hygroscopicity of various polymer materials was examined using THz-TDS, where cyclic-olefin copolymers COC 6013 and COC 5013 were recommended as THz optical windows based on low hygroscopicity and high optical transparency.[253] In this same study, high-density polyethylene and poly-tetrafluoroethylene (Teflon) were also deemed suitable on the basis of low hygroscopicity.[253]

Electrically conductive polymers have also been analyzed by using THz-TDS technology. Materials such as polypyrrole,[255] poly-3-methylthiophene,[255] poly-3-hexylthiophene,[256] and poly(2,5-bis(3-tetradecylthiophen-2yl)thieno(3,2-b)thiophene)[257] have been studied using THz-TDS. Characterization of these materials is of particular interest because they have performed well as alternative materials for the construction of semiconductor technology, such as field-effect transistors (FET's), light emitting diodes (LED's), solar cells, or a variety of other uses. Inorganic materials have already been used for these devices, but the flexibility of these polymer semiconductors would potentially enable a wide variety of portable applications. Determination of the charge carrier mobility can guide improvements in chemical structure. For example, solar cells constructed using materials with high carrier mobility will outperform cells constructed with materials having lower carrier mobilities. Carrier mobility (μ) is proportional to the relative change in THz transmission ($\Delta T/T_o$) divided by the efficiency (η) of carrier photogeneration, shown in Equation 8:

$$\mu = \frac{(1+N)h\nu\left[\left(\frac{\Delta T}{T_0}\right)/\eta\right]}{eF(1-R)(1-exp\{-\alpha d\})Z_0} \qquad \text{Equation I-12}$$

where ν is the pump laser beam frequency, N is the film refractive index, F is the incident fluence, α is the polymer film absorption coefficient, R is the polymer film reflection coefficient, e is electric charge, h is Planck's constant, and Z_0 is the impedance of free space.[258] THz-TDS is appealing for this set of measurements because information can be gained on the properties of the polymer without physical contact. This feature reduces the number of experimental variables that could affect the collected data. Traditional measurements of this type involve the use of electrical contacts on the surface of the semiconductor material, but the resistance at these junctions will often impact the observed mobility as much as a change in material composition.[256]

Polymers have also been of interest to the THz community as potential antenna substrates for the production and detection of THz radiation. Several publications indicate that certain polymers could be used in place of the biased group III-V semiconductor antennas described above.[259] Polymer materials offer several advantages and disadvantages over their inorganic counterparts. Polymer-based emitters and detectors can be manufactured with broadened frequency response, up to 12 THz in band width.[260] Internal reflective losses are less because of lower refractive index of the polymer compared to inorganic crystal.[260] The polymer-based components are easy to fabricate, flexible, less expensive, and easier to handle.[261] Less dispersion is also observed with polymer-based emitter devices.[262] Major drawbacks of these polymer materials include low radiant powers and photodegradation over time.[259] Further material development could simultaneously decrease the cost and improve the portability of THz-TDS instrumentation.

Analytical Challenges for Quantitative Measurements of Solids

The quantitative relationship between magnitude of THz radiation absorbed and concentration of the absorbing species should follow the Beer-Lambert law, as noted

above in Equation 2. Many papers in the THz literature explore the quantitative nature of THz-TDS measurements for solid samples, in particular samples where the targeted analyte is dispersed within a pellet of PE. Unfortunately, quantitative relationships are not well defined in the vast majority of these reports and plots of absorption versus concentration are highly scattered and of low overall quality. Clearly, analytical measurements of solids using THz-TDS methods are not straightforward.

A number of factors serve to complicate quantitative THz-TDS spectroscopy. In some cases, the targeted analyte is too highly absorbing to permit high quality spectra, particularly when spectra are collected with a transmission geometry. In such cases, reflectance measurements either in an ATR or diffuse reflectance format might be helpful.

Spectral features within THz spectra of solid samples tend to be relatively broad compared to those obtained from liquid and gaseous samples. Narrower spectral features enhance chemical selectivity by reducing spectral overlap and providing greater distinction between molecular species. Inhomogeneous and homogeneous broadening are two primary sources of peak broadening for these samples. Inhomogeneous broadening results from a large number of analyte particles with non-identical molecular conformations. The randomness of such granule morphologies creates a distribution of sizes, shapes, and crystal orientations that induce broadening of the crystalline phonon vibrations. The impact of inhomogeneous broadening can be reduced by using a single crystal,[263] though synthesis of single crystals can be challenging considering the size necessary for high quality spectra.

Alternatively, samples may be ordered through the use of a waveguide geometry, as reported by Laman and co-workers.[263] Waveguides consist of two parallel polished metal plates (such as Cu or Al) stacked vertically, leaving a gap between the two plates.[31] High throughput can often be achieved with these devices through careful alignment and use of optical components that focus radiation into the waveguide (between the parallel

plates), in spite of the sub-wavelength spacing of the plates.[263] Loss is also minimized by selecting waveguide materials having high THz reflectivity and polishing the plate surfaces. Samples can be drop-cast within the waveguide by distributing a sample-miscible solvent on the plate prior to addition of the sample.

Using the waveguide technique, Laman et. al.[263] obtained spectra from samples of benzoic acid, salicylic acid, 3-hydroxy-benzoic acid, 4-hydroxy-benzoic acid, and aspirin. These spectra revealed absorption peak full-widths at half-maximum (FWHM) are significantly reduced, to as little as 20 GHz (0.667 cm^{-1}). In addition to enhanced resolution achieved through narrowing of spectral features, Laman et. al. report an approximately 100-fold improvement in sensitivity afforded by the waveguide, obtained through a higher sample filling factor.[263] It was noted by the authors that crystalline structures formed within the confines of the waveguide exhibited an increased level of order, as crystals were found to be largely oriented in a planar fashion, which is thought to be the mechanism by which the effect of inhomogeneous broadening in the acquired absorption spectra is reduced.[31]

Homogeneous broadening is caused by the distribution of phononic vibration energy levels present in a sample at room temperature. Homogenous broadening can be reduced by lowering the temperature of the sample. Fewer vibrational states are occupied at lower temperatures, thereby narrowing the band of frequencies absorbed by the sample. Spectra were collected at reduced temperatures (77 K) as well as room temperature (295 K) to investigate the narrowing of absorption features. Such improvements are especially desirable when characterizing materials. The actual number of phononic transitions can be more closely determined if superimposed spectral features can be resolved, allowing visualization of weak transitions that are often lost in the shoulders of neighboring peaks. Transition center frequencies can be more easily estimated as spectral overlap is reduced. A slight trade-off is encountered when using a waveguide, as the sharpening of spectral features will require the acquisition of lengthier

TDSs to attain the necessary resolution. The increased signal length will require increased collection times, assuming an equivalent amount of signal averaging or integration time.

The granular nature of either the sample itself or the matrix in which the sample is dispersed can adversely affect the quality of THz spectra and, hence, the ability to perform quantitative measurements. Reports illustrate that scattering of the incident THz radiation depends on the size of the particles used to create the sample.[264] Spectroscopically, this scattering produces an exponential increase in the baseline from low to high THz frequencies and this sloping baseline can result in significant attenuation of radiation, particularly at higher frequencies. Figure I-27 shows an example of absorption spectra collected from a series of PE pellets prepared from stock PE materials with different particle sizes.

Large particles (150-200 μm) exhibited a higher degree of scattering when compared with samples containing smaller particles.[9] Milling granulated samples and/or sieving the sample particles prior to forming the pellet can reduce the slope of the baseline, but increases preparation time. The resulting baselines can be modeled by polynomial functions of different orders according to particle size. In general, the polynomial order which best fits a baseline was found to increase as the particle size increased.[9] Compensation for a sloping baseline could be achieved by using a pre-determined polynomial function to fit and remove the baseline similar to methods traditionally used in processing spectra collected from Raman scattering experiments.[265]

Other heterogeneous properties of the samples will contribute to degradation of the quality and reproducibility of THz spectra. Spatial distribution of components will affect spectroscopic investigation as the diameter of the THz beam is frequency dependent when generated in such a manner.[35] The beam diameter is smaller at high frequencies and increases at low frequencies. If, for example, analyte particles are located preferentially towards the periphery of a pellet or tablet, then these molecules will

not be probed equally at all frequencies, thereby distorting the spectrum. Sample pellets should be prepared from particles that have been thoroughly mixed to provide reproducible spectra. THz imaging might be useful to characterize analyte distribution.

Another source of spectral variation is a frequency-dependent attenuation of radiation transmitted or reflected at multiple interfaces within the sample.[234] Reports on this phenomenon illustrate its significance across the THz spectrum. An example is a set of parallel air/sample and sample/air interfaces for a sheet of PE with uniform thickness,

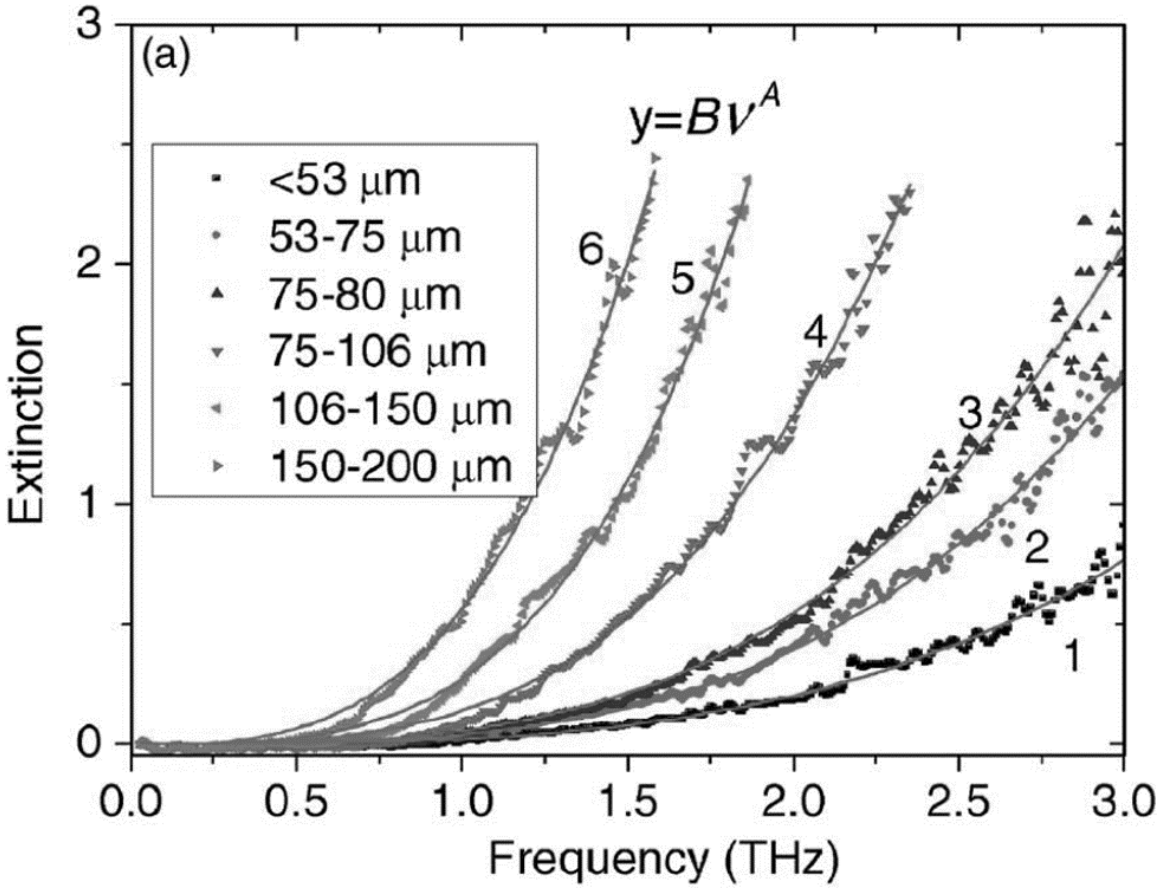


Figure I-27. Absorbance spectra of PE powders of various particle sizes. PE is a non-absorbing compound, but can still scatter THz radiation. Note that the increase in attenuation as particle size increases. Taken from[9] with permission.

as illustrated schematically in Figure I-28. Reflectance occurs when the THz radiation travels from a region of low refractive index (air) to a region of high refractive index (sample). The fraction of radiation lost by reflection is dependent on the difference in refractive index for the two materials, as described by Fresnel's equations. The resulting spectral impact will be a broadband attenuation in the transmitted signal. For many substances, the refractive index is wavelength dependent across the THz spectrum, thereby creating a matrix dependent source of spectral variation. The broadband attenuation can be removed by baseline fitting as described above or possibly by referencing the absorption spectrum to a suitably matched reference material. Still, the loss of transmitted signal diminishes the overall signal-to-noise ratio, which adversely impacts the quality of the spectra.

Fabry-Perot interference fringes can be an additional source of confounding spectral variation within THz spectra. As illustrated schematically in Figure I-28, the transmission measurement is essentially a Fabry-Perot cavity that results in multiple passes of the incident radiation through the sample. An etalon pattern is observed in the frequency-domain spectrum, depending on the thickness of the sample and the differences in refractive index at the interfaces. This etalon pattern corresponds to a periodic fluctuation in the baseline signal and the superposition of such periodic fluctuations and sample-dependent spectral features can confound qualitative and quantitative analyses. Although, in principle, the Fabry-Perot effect can occur beyond two reflections within the sample, higher order reflections are generally insignificant. An example of the Fabry-Perot phenomenon is illustrated in Figures I-29 and I-30, which correspond to the time and frequency domain spectra collected for a series of polyethylene sheets. The bottom signal shows results from dried air and no polymer. The remaining three spectra correspond to one, two, and three (moving upward) 0.81 mm thick sheets of high density polyethylene in the optical path, respectively. These spectral data were collected with a spectral resolution of 1.067 cm^{-1} and averaging 1800 scans.

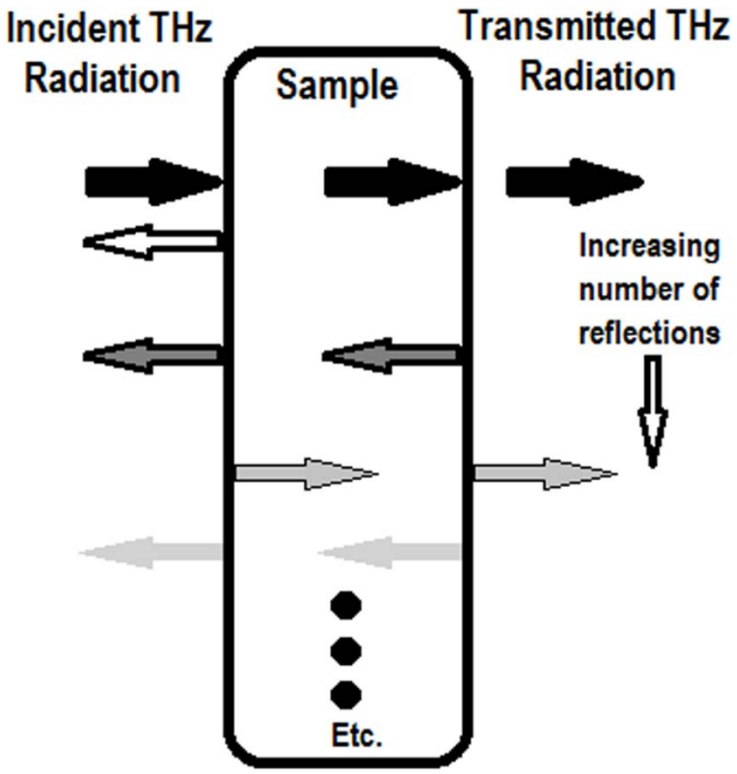


Figure I-28. Illustration of the Fabry-Perot effect showing reflection and transmission at the sample interfaces.

The Fabry-Perot phenomenon can be viewed directly in the time-domain spectrum and can provide information about the thickness of the sample. Indeed, as noted above for measurements of pharmaceutical tablets, relative positions of the main peak and the Fabry-Perot echo in the time-domain spectrum provides a measure of the depth and/or thickness of the various features within the tablet. Also, the number of interfaces can be determined directly from the time-domain spectrum. It should be noted that the reflection artifacts visible in the TDSs shown below are more prominent than those expected with a pellet, as multiple interfaces give rise to the intense second and third artifacts. The intensity of each echo artifact from ensuing reflections can be estimated using Fresnel's equation for reflection, shown in Equation I-13.

$$R = [(n_1 - n_2)/(n_1 + n_2)]^2 \qquad \text{Equation I-13}$$

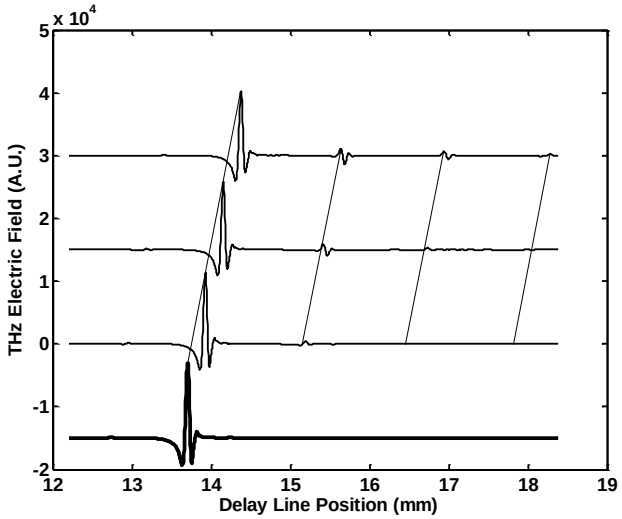


Figure I-29. Time-domain signals illustrating the etalon effect. Signals are offset vertically for comparison and lines are provided as a visual guide for position of the main peaks and the Fabry-Perot echoes. The bottom signal is for dried air and the three additional spectra are for one, two, and three (moving upward) sheets of HDPE in the beam path.

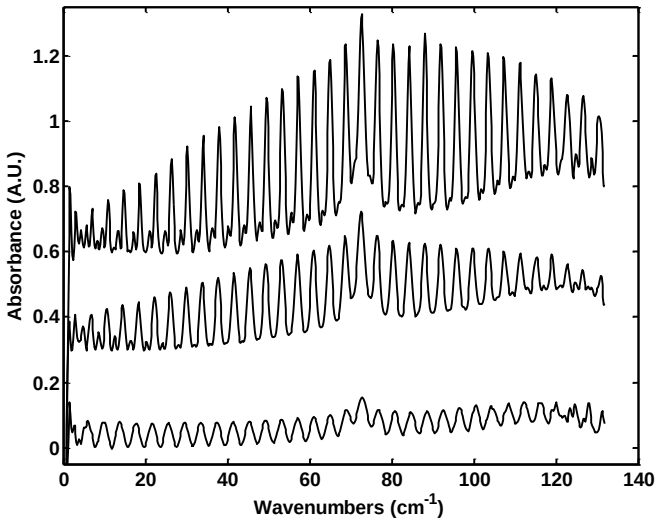


Figure I-30. Absorbance spectra calculated from the time-domain signals in Figure I-29. From bottom to top, the spectra represent one, two, and three sheets of HDPE placed in the beam path. The time-domain signals were extended with a linear function to 4×N data points, where N is the nearest power of 2 from the original TDS length. These signals were Fourier transformed after boxcar apodization. A spectrum collected with dried air was used as the reference for these spectra.

where R is the reflection coefficient, n_1 is the refractive index of one material (air, for example) and n_2 is the refractive index of the other material (PE in this example). Given an air/PE interface, one would expect the coefficient of reflection to be approximately 0.045, or 4.5% per interface. Because of the exponential decay in the intensity of the etalon echo, the second and third echoes are often obscured by instrumental noise or the waveform of the sample.

Although the number of interfaces can be determined by counting the number of Fabry-Perot echoes in the time-domain spectrum, it is important to understand that the scan length of the delay line must be sufficiently long to view echoes for all interfaces. If the reflected radiation were to travel through more material (of a greater thickness) or a material of sufficiently high refractive index, the time delay of the "echo" signal may be outside the region scanned with the delay line.

The corresponding frequency-domain spectra presented in Figure I-30 illustrate the complex fringe pattern that can result from multiple interfaces within the sample. The presence of such fringe patterns mixed with sample absorption features can certainly complicate quantitative measurements and degrade analytical performance. Various methods are proposed to reduce the impact of these Fabry-Perot patterns. One approach is to mathematically alter the time domain signal to remove the fringe pattern[266] and another, which has proven useful when applied in other frequency ranges, is to use windows and samples with non-parallel faces (wedge-shaped), so that the reflected radiation would no longer propagate toward the detector.[267]

As noted above for biological samples, the strong absorption properties of water presents a major challenge in collecting high quality THz-TDS spectra from samples with appreciable amounts of water. Water can interact with some compounds and affect the THz spectrum directly.[6] Mostly, however, liquid water attenuates the transmitted radiation, thereby reducing the signal-to-noise ratio of the measurement. A common practice is to dehydrate such samples if possible without damaging or altering the sample.

Alternatively, samples can be frozen, thereby taking advantage of the sharper spectral bands of ice compared to liquid water.[268]

Conclusions

THz-TDS, though in its infancy when compared to other spectroscopic methods, has been applied to a vast array of qualitative as well as quantitative applications. A variety of detection schemes and optical configurations have been developed for this method, providing greater flexibility which in turn broadens the applicability of the method. The time-resolved nature of the method provides unique information inaccessible to CW instrumentation which is often used for materials analysis, characterization, and quality control. Data processing methods can be applied prior to and following the Fourier transformation of the TDS, providing enhanced SNRs and allowing more precise determination of spectral peak parameters.

Through the use of both spectroscopic and imaging apparatus, THz frequencies have been used to meet a broad range of analytical challenges. Security-related objects such as concealed contraband have been imaged through non-metal barriers, and explosives have been identified based on spectral features. Tablet coating analysis and imaging of ingredient distribution have been performed using these frequencies, presenting an analytical method almost ideally-suited for the pharmaceutical industry. A variety of biological tissues have been examined using these frequencies, including dental, organ and dermal tissues, with promising results in the areas of caries and cancer detection. Analysis of polymeric compounds has illustrated the potential of this method in analyzing polymer additives in on-line measurements. Other polymers have been found to have potential in THz frequency optics, such as window materials or even antenna substrates.

Like any analytical method, hurdles must be overcome when analyzing solid samples. Water is the THz spectroscopist's nemesis, as it exhibits intense, broad

absorption across this frequency range. As such, investigation is restricted to relatively dry samples when using this method. This challenge may be overcome as methods are developed in which higher radiant powers are attained. Spectral features of some solid compounds have also been shown to vary depending on the extent of hydration. Scattering and reflection produce spectral artifacts, though these may be reduced or eliminated using special sample preparation techniques or spectral processing methods. Improvements to the conventional design of the apparatus are currently under investigation. Polymeric compounds have been shown to provide an inexpensive and relatively broadband antenna substrate alternative, though state-of-the-art devices are susceptible to photodegradation and produce low radiant powers. Advances in delay line technology have been reported, improving the scan rate of the apparatus. Improvement of femtosecond laser technology could provide reduction in the size of the apparatus, and lasers capable of producing NIR pulses of shorter duration would ultimately provide increased spectral bandwidth. Development of anti-reflective (AR) coatings suitable for THz frequencies could result in more efficient transmission of terahertz pulses, allowing a higher level of throughput and improved SNRs. As the method is developed and performance is enhanced, the number of potential analytical applications will undoubtedly increase. The THz frequency range has been shown to provide unique information for a variety of solid samples of scientific and engineering interest.

CHAPTER II

INSTRUMENT CONFIGURATION AND CHARACTERIZATION

The TeraView Ltd. TPS Spectra 1000 D

Measurements were performed with a TeraView TPS Spectra 1000D, a THz-TDS system that generates and detects terahertz radiation via a photoconductive (PC) antenna. Though designated as the model "1000D", this instrument is identical in design and function to the Spectra 2000. The Spectra 1000D was the second generation instrument, capable of operating both in transmission and reflection (ATR) modes, hence the "D" designation (meaning dual mode). Shortly after manufacture of the Spectra 1000D however, the model name was changed to Spectra 2000 to reflect the second generation technology and design. Given the unique characteristics of the instrument's operation and the relative infancy of the technology, descriptions of its configuration and principles of operation are warranted. The primary components of the apparatus are similar to those illustrated in Figure I-1, with the exception of the translation stage mounted in the beam path. Emitter and receiver antennae are similar to those illustrated in Figures I-2 and I-3.

The Ultrafast Laser (Coherent Vitesse-2)

As mentioned in Chapter I, the generation and detection of broadband THz pulses are dependent in part on ultrafast laser pulses. A Coherent Vitesse-2 titanium-doped sapphire (Ti:sapphire) ultrafast laser is used to generate pulses of NIR radiation in the TPS 1000D. Specifications of the laser unit are displayed in Table II-1. The laser consists of several small subunits: the pump laser diode, Verdi laser head, Verdi-Pumped Ultra-Fast (VPUF) laser head, and PowerTrack mirror, depicted schematically in Figure II-1. A complete description of the laser components and their functions may be found in the Coherent Vitesse User's Manual.[10] The PowerTrack mirror is an automatic alignment system designed by Coherent Inc. to reduce maintenance while providing optimal power output.

Table II-1. Selected specifications for the Coherent Vitesse 800-2 Ti:sapphire laser.

Property	Value	Range	Unit
Average Power	250	greater than	mW
Bandwidth	10	greater than	nm
Pulse width	100	less than	fs
Repetition Rate	80	± 1	MHz
Wavelength	800	± 1	nm
Power Stability	± 1	N/A	%

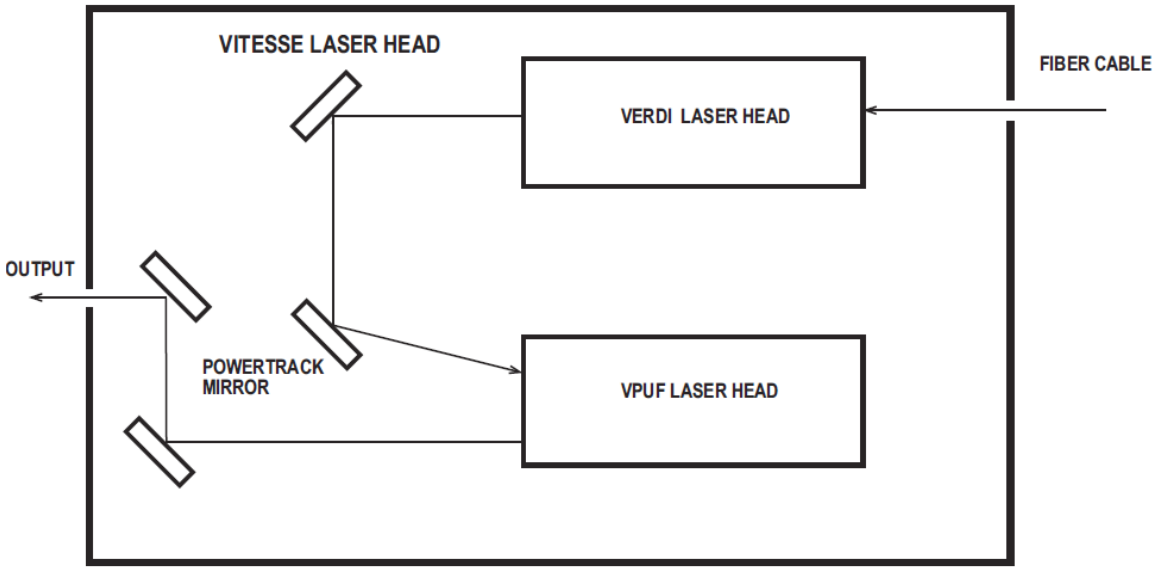


Figure II-1. A diagram of the subunits within the Vitesse 800-2. The fiber cable is used to couple in light from a solid-state diode laser. Adapted with permission.[10]

Generation of the ultrafast NIR pulses begins with a solid state CW diode laser, which produces NIR radiation at a wavelength of 808 nm. The output of this diode laser is coupled into the Vitesse unit via a fiber optic cable. Figure II-2 depicts an energy level diagram for this semiconductor laser. Once the light has entered the Verdi laser head, it is used to "pump" or excite a neodymium yttrium orthovanadate (Nd^{3+}:YVO_4, sometimes referred to as "vanadate") crystal, which in turn emits radiation at a wavelength of 1064 nm. Vanadate crystals provide several benefits in this application when compared to

common alternatives such as Nd:YAG (yttrium aluminum garnet host crystal) or Nd:YLF (yttrium lithium fluoride host crystal). Neodymium-doped vanadate has higher absorbance than Nd:YAG or Nd:YLF, which provides more efficient conversion and allows the use of smaller crystals. Vanadate's stimulated emission cross-section is almost 4 times greater than that of the YAG matrix, and the emitted radiation is both stable and of relatively narrow bandwidth centered near 1064 nm.[10] Light emitted from the vanadate crystal is also polarized.

Vanadate is also an efficient lasing medium due to the 4-level scheme involved in the lasing process, illustrated in Figure II-4. Four-level gain media have several advantages over 3-level media such as ruby lasers. The energy level depicted in Figure II-4 as E_y is not favored, so self-absorption losses (a transition from E_0 to E_y) are less significant, which provides greater efficiency of the conversion of pump energy to emitted laser light. This light then passes through the frequency-doubling lithium borate (LBO) crystal, producing light with a wavelength of 532 nm, which is green in color. An optical diode is included in the lasing cavity to induce unidirectional oscillation, which is the propagation of light in one direction about the beam path. When also combined with the temperature-stabilized etalon, frequency selectivity is improved.[10] An astigmatic compensator removes distortion caused by non-normal reflection of the laser beam from the surface of the curved mirrors.[10]

At this point, the light is coupled out of the Verdi laser head, reflected off the PowerTrack Mirror system, and enters the VPUF unit, which houses the Ti:sapphire crystal, from which the laser derives its name. A schematic of this unit is provided in Figure II-5. It is in this unit that the green light from the Verdi laser (532 nm) is converted to NIR light with a wavelength of 800 nm. Wavelength conversion is accompanied by the pulse-shaping process which is ultimately responsible for production of the required fs pulses.

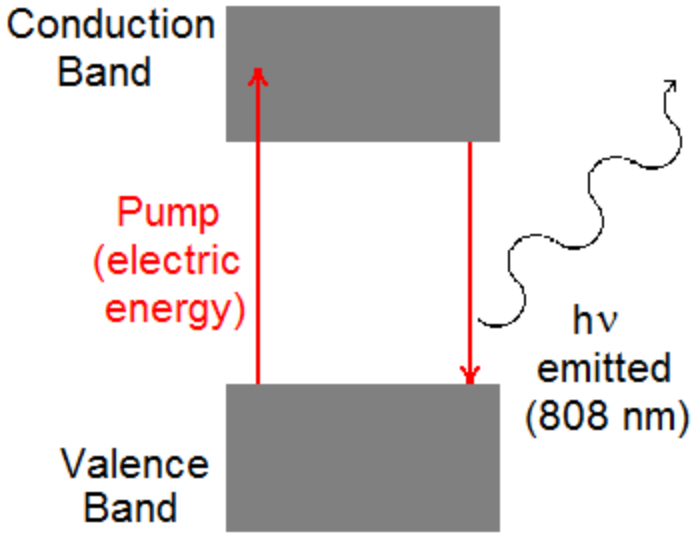


Figure II-2. The 2-level energy diagram depicting the electronic transitions occurring in the solid-state diode pump laser during emission.

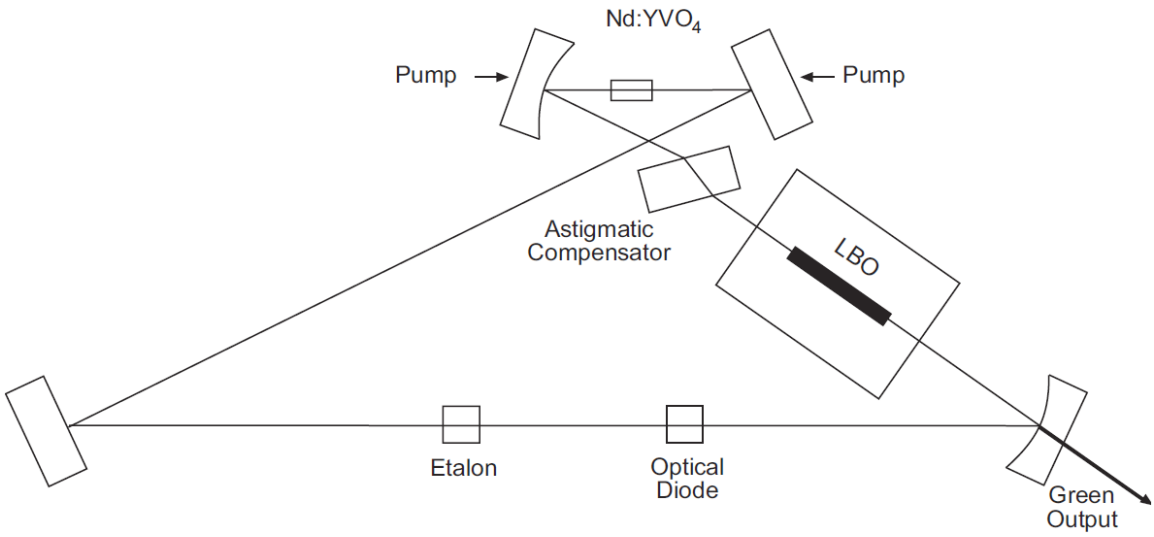


Figure II-3. Schematic of optical components within the Verdi laser head. Image reproduced with permission.[10]

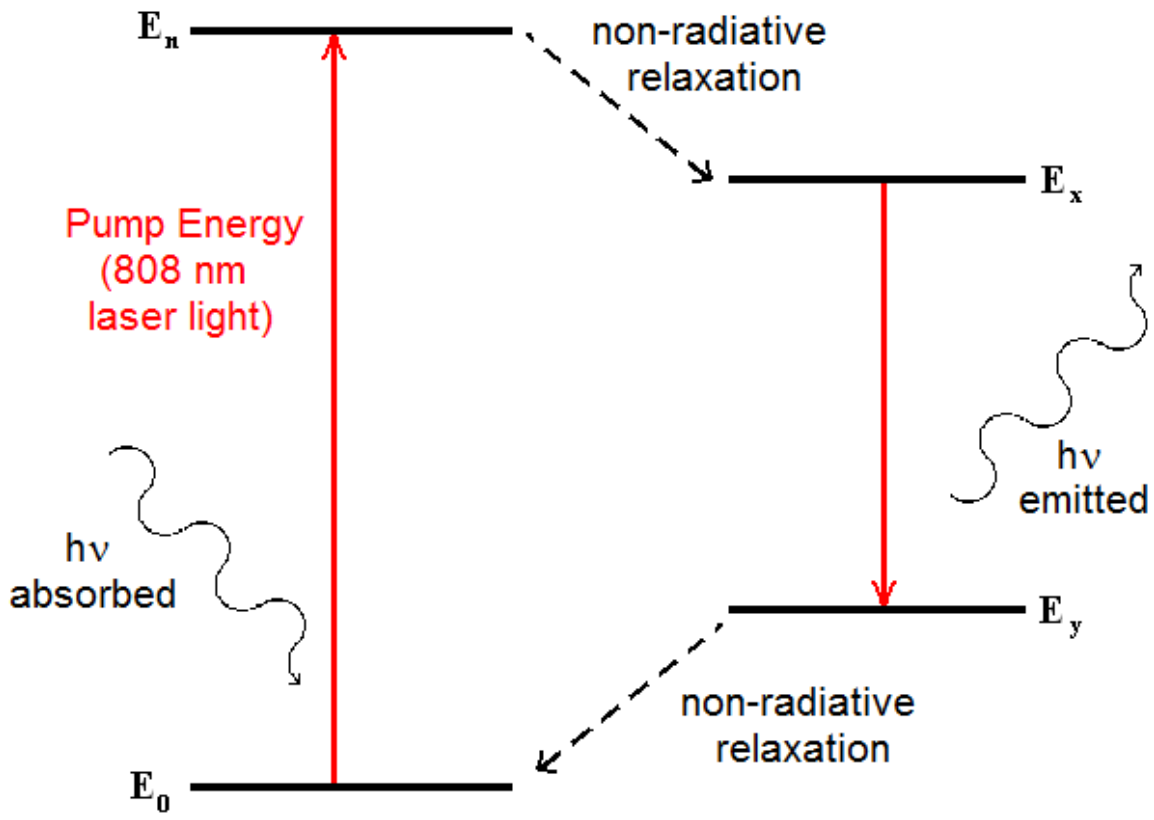


Figure II-4. Energy diagram depicting the electronic transitions occurring in the Nd^{3+}:YVO_4 laser gain medium.

Solid-state ultrafast lasers like the Vitesse 800 rely on a technique known as Kerr lens modelocking (KLM) in combination with the optical properties of the gain medium to generate fs light pulses.[10] KLM is initiated via an automatic starter unit and a slit. Prior to initiation of KLM, the laser operates in CW mode. When operating in CW mode, the laser beam profile is broader than during modelocked operation.[10] The slit's size is adjusted to induce appreciable attenuation when the laser operates in CW mode, but losses are decreased when the laser operates in a modelocked fashion. Kerr lens modelocking relies on a phenomenon known as the "optical Kerr effect" in this configuration, through which the beam size is modulated without inclusion of additional optical components in the laser cavity.

The optical Kerr effect occurs when bursts of sufficiently-intense light travel through a medium, altering the material's refractive index, and thus changing the size of the beam as it propagates through the medium.[10] Regions in which this occurs are referred to as Kerr lenses, and are critical in the KLM process. Ordinarily, Kerr lenses will not form while the laser is operating in CW mode,[10] but can be formed with assistance from the starter. The starter device consists of a high-reflectivity mirror mounted on a solenoid-actuated spring. When a potential is applied to the solenoid, the mirror will move rapidly, which in turn changes the length of the laser cavity. Rapid changes in the length of the laser cavity can induce temporary spikes in the laser power as a result of constructive interference between passing photons. In turn, the spikes can produce Kerr lenses in the laser cavity, which reduces the beam diameter and decreases attenuation of the laser beam by the slit. Decreased attenuation produces greater spikes in intensity when the cavity length is modulated by the starter, and the process is repeated. The starter is only required in this stage, after which the laser becomes "self-modelocked" and requires no further input to maintain pulsed operation.

Ultrafast laser pulse duration is a product of the emission characteristics of the gain medium. The uncertainty principle dictates that a short pulse in the time domain will have a correspondingly broad spectrum in the frequency domain.[11] This principle is illustrated graphically in Figure II-6. One can see from this figure that the pulse duration is inversely proportional to the bandwidth of the frequency domain emission spectrum and the laser repetition rate is proportional to the length of the laser cavity. The apparatus would not be complete without inclusion of the negative dispersion mirrors (NDM).

Kerr lens formation relies on a change in refractive index in the laser gain medium, but the altered refractive index can affect the pulse shape and duration through a phenomenon known as group velocity dispersion (GVD).[10] As discussed previously, generation of short laser pulses depends on high spectral bandwidth, but many materials

exhibit dispersion, which is characterized by frequency-dependent variation of the refractive index spectrum. As such, the refractive index for one frequency may be higher than that of another, causing it to be delayed with respect to the other frequency component. This results in broadening of the laser pulse in the time domain, which must be corrected via reflection from the NDMs.[10] Negative dispersion mirrors are optical components constructed from a thin, transparent material coated with a highly reflective surface on the back side which act like an F-P interferometer.

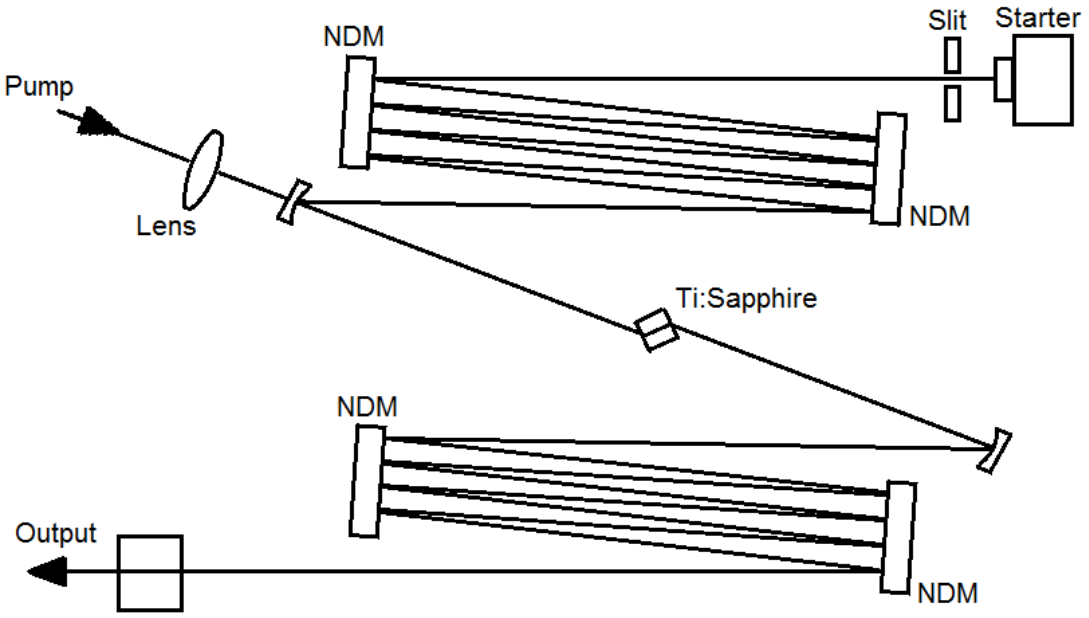


Figure II-5. Schematic of the VPUF laser head configuration. Adapted with permission.[10]

The thickness of the mirror is small enough that the difference in phase between photons reflected at the first surface and the reflective surface at the rear becomes negligible. As a result, the phase alteration induced by the NDM becomes a function of frequency, as shown in Equation II-1.[10]

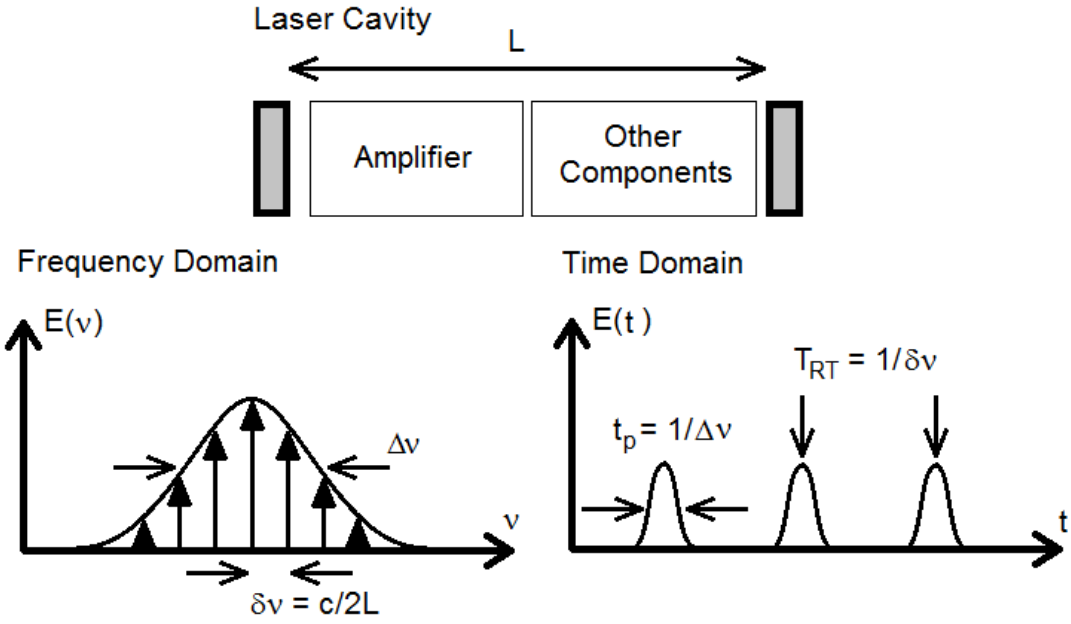


Figure II-6. Relationship of cavity length (top) and emission profile (lower left) to pulse duration and repetition rate (lower right). Adapted with permission.[11]

$$\tan\phi(\omega) = \frac{-2 * t_0^2 * (1 - r^2) * r * \sin[\omega * t_0]}{(1 + r^2 - 2 * r * \cos[\omega * t_0])^2} \qquad \text{Equation II-1}$$

where $\tan\phi(\omega)$ is the phase, r is the reflectivity of the front surface of the etalon, ω is the angular frequency of the light, and t_0 is the transit time of the light through the etalon. Four NDMs are used to reduce the overall GVD of the laser cavity to a near-zero level. The effect of uncorrected GVD on the laser pulse is illustrated in Figure II-7.

Complete descriptions of the generative and modelocking processes can be found in the literature.[10, 11] At this point, the laser produces ultrafast laser pulses which are coupled out of the laser unit via fiber optic cables, and are used to generate and detect THz pulses via the mechanism described in Chapter I. Further explanation of the principles governing the operation of ultrafast laser systems is available elsewhere.[269]

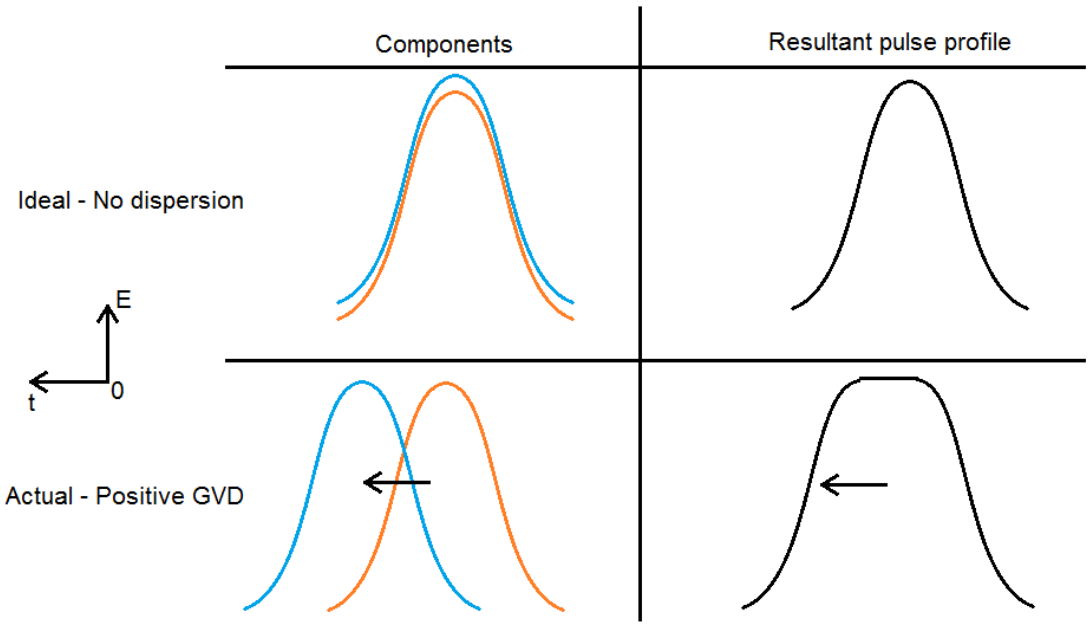


Figure II-7. Diagram of GVD phenomenon, illustrating frequency components and resultant pulse profiles. High-frequency components are represented by the blue trace, and low-frequency components are represented by the red trace. Note that the x-axis of these plots is in time, and the axis is reversed such that t=0 resides toward the right side of the pulse. The figure illustrates positive GVD, which is a case where the refractive index is greater for high frequencies, so high frequency constituents of the pulse are delayed relative to low frequencies, and are observed or measured at a later time. It is important to note that this effect is exaggerated for illustration and that the observed phase delay would be more continuous than shown here.

Instrument Characterization

Quantitative analysis of weak absorbers or low sample concentrations tends to challenge the capabilities of instrumentation, as mentioned in Chapter I. As such, it is wise to examine the properties of analytical instrumentation prior to sample analysis. A key instrumental figure of merit is the SNR, which describes the ratio of the magnitude of instrumental variance relative to a measured signal. The SNR is related to the spectral RMS noise, which may be determined using Equation II-2.[270]

$$RMS = \sqrt{\frac{\sum_i (A_i - \bar{A})^2}{n}}$$ Equation II-2

To calculate the RMS noise (RMS), the summed squared deviation of each spectrum (A_i) from the mean spectrum ($\bar{A}$) is normalized by the number of averaged points (n), either a range of resolution elements or multiple averaged spectra at a single wavelength. The square root of this value provides the RMS noise, which is related to the SNR via Equation II-3.

$$SNR = \frac{1}{2.303RMS}$$ Equation II-3

The SNR (SNR) is equal to the reciprocal of 2.303 times the RMS noise (RMS). RMS noise may be reduced by averaging replicate spectra, which is carried out by software installed on the TPS Spectra 1000D spectrometer, allowing rapid acquisition of low-noise spectral data. Thus, the RMS noise will decrease in a manner which is proportional to the reciprocal of the square root of the number of averaged scans, which in turn will produce a corresponding increase in the SNR.

Spectral noise is known to be proportional to the spectral resolution, amongst other instrumental characteristics as described by Griffiths and de Haseth.[12] As such, it is worthwhile to characterize the noise as a function of resolution to determine a practical value for measurement. The relationship between spectral resolution and the observed peak width is also of interest,[12] though the resolution selected for the reported gas phase spectra is still above that of the expected spectral peak widths.[181] Lower resolution (narrower data point spacing) is achieved by increasing the length of the time window sampled after the TDS main peak; however the decreased spectral resolution is obtained at the expense of the SNR as well as data acquisition time. Furthermore, published accounts have indicated that robust qualitative and quantitative models may be developed even when the instrument resolution is greater than the expected spectral peak widths of the analyte.[271-273] In some cases, the resolution may need to be optimized on a per-analyte basis.[271] It should be noted that while the SNR is higher when spectra are acquired at higher resolution, spectral peaks which are narrower than the instrumental

resolution will be broadened and their widths will approach that of the spectral resolution.[181] As the peaks are broadened, they will tend to overlap, which can have an adverse effect on the selectivity of the method.

To evaluate the effect of spectral resolution on the SNR, 3 TDSs were acquired at a resolution of 0.225 cm^{-1} as the average of 1800 scans each. An average of 1800 scans was suggested by TeraView representatives, but the effect of averaging was also examined. For this estimation, three single beam spectra were acquired with an open beam path after an extended (>24 hr) period of dry air purging and were collected manually in rapid succession. It should be noted that this set of data was measured shortly after alignment of the instrument's optical components, so the values presented represent near-optimal signal levels. Signal levels recorded during gas sample analysis were lower due to the extended period between the previous service appointment and the time of spectral acquisition.

While the TDSs were acquired at a single resolution (0.225 cm^{-1}), the spectral resolution may be increased by truncating the TDS and extending the signal with a linear function to provide uniform data point spacing (as described in Chapter I). The resolution of a TDS may be determined using Equation II-4.

$$Res = {}^{1}/_{2\Delta} \qquad \text{Equation II-4}$$

The resolution (Res, in cm^{-1}) is a function of the reciprocal of 2 times the post-peak optical delay (Δ, in cm) after truncation (if applicable) but before extrapolation of the signal. Single beam spectra were calculated from the TDSs, which were then used to calculate hundred percent lines in three combinations, shown in Figure II-8 and Figure II-9. RMS noise was calculated by fitting the spectrum with a quadratic function across the 20 – 60 cm^{-1} spectral range. This procedure can be represented by Equation II-2 in that n represents the number of resolution elements included in the fitted spectral region and $\bar{A}$ represents the quadratic fitted value at a particular resolution element. Measured RMS

noise and SNR values are plotted as a function of spectral resolution in Figures II-10 and II-11, respectively. A spectral resolution of 0.225 cm^{-1} was selected for sample spectra, as this provided a reasonable SNR, and the resolution could be altered via truncation of the TDS at a later time. Spectral resolution below 0.225 cm^{-1} requires acquisition of a fifth TDS segment (explained further in the next section), which increases collection time when averaging 1800 scans.

Replicate TDSs were also collected to evaluate the effect of signal averaging on the calculated RMS noise. One-hundred sixty-two signals were collected with the minimum level of averaging (30 scans), and various numbers of single beam spectra were averaged to mimic different levels of signal averaging. RMS noise was calculated across the 5-125 cm^{-1} range for each of the averaged hundred percent lines. Noise levels were found to change drastically at low levels of averaging, but the decrease in RMS noise is minimal at high levels of averaging, as shown in Figure II-12. When the same RMS noise values are plotted as a function of the reciprocal of the square root of the number of averaged scans, the plot becomes linear (as shown in Figure II-13), in accordance with Equation II-2. Eighteen hundred averaged scans were selected for experimental data, as this provides a low level of RMS noise without requiring prohibitively long scan times.

The triplicate single beam spectra used for analysis of the spectral resolution were also used to estimate the SNR in various spectral regions at the desired spectral resolution. The mean and standard deviation of the spectra are displayed in Figure II-14. Spectral SNR may be represented by a ratio of the mean and standard deviation, shown in Figure II-15, though this method tends to underestimate the SNR available with simple baseline correction steps. Calculated RMS noise values are tabulated in Table II-2 as calculated by least-squares fits of the spectral baseline. Note the correlation between the SNR values and the mean single beam spectrum magnitude shown in Figure II-14.

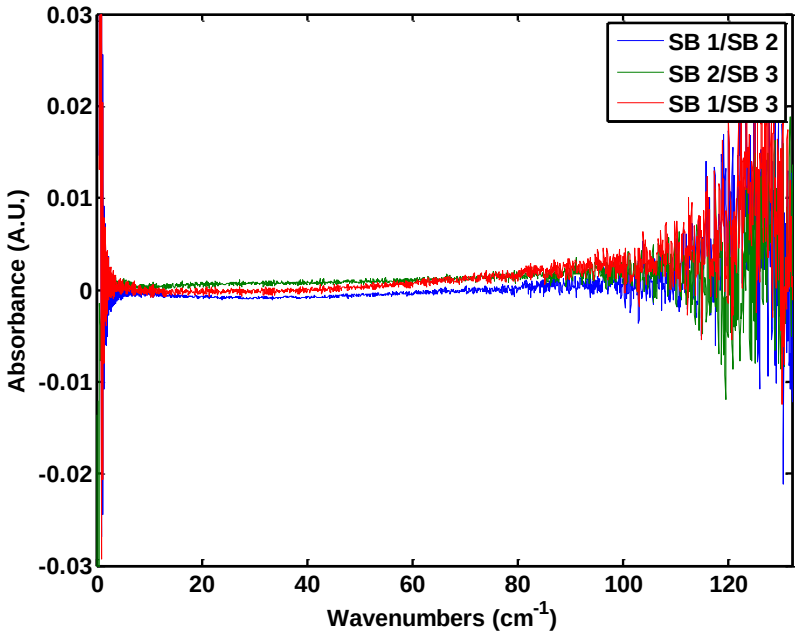


Figure II-8. Hundred percent lines calculated from three single beam spectra. Note the high variability at the extremes of the plotted spectral range.

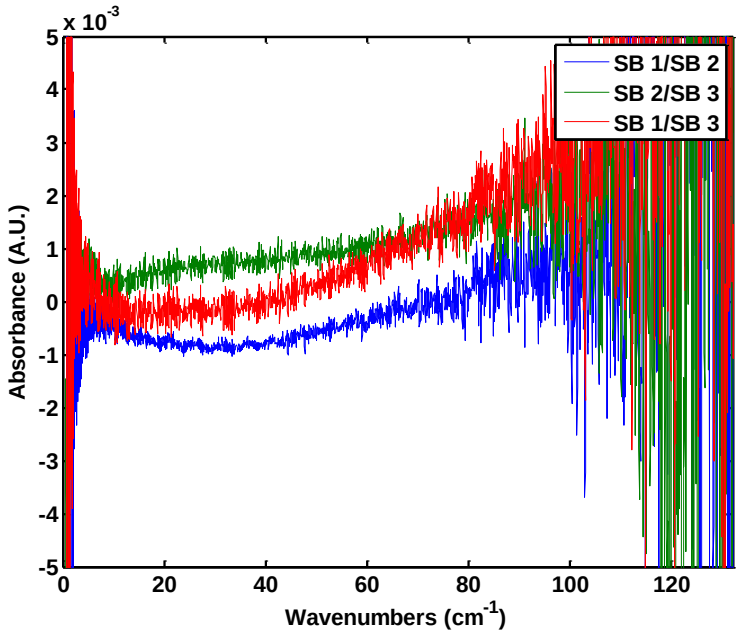


Figure II-9. Expanded view of hundred percent lines plotted in Figure II-8. Some high-frequency artifacts in the spectrum are non-random. The source of this variability will be discussed in the TDS shift-correction algorithm section.

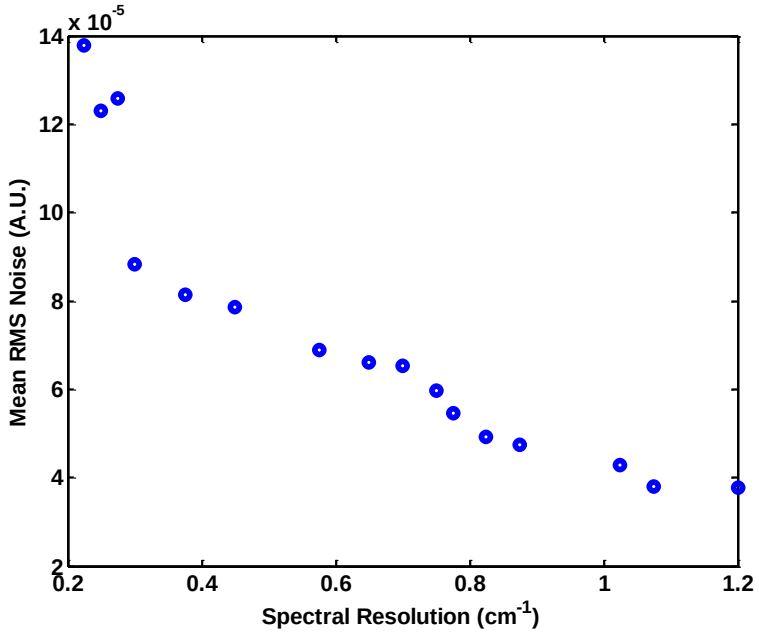


Figure II-10. Mean RMS noise from the 3 hundred percent lines plotted as a function of spectral resolution.

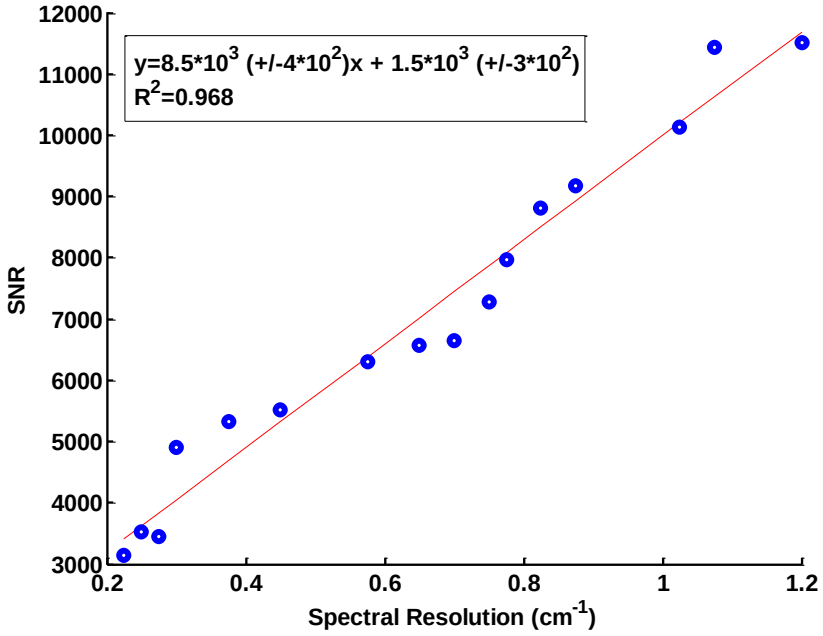


Figure II-11. SNR plotted as a function of resolution, demonstrating the linear relationship between SNR and resolution described by Griffiths and de Haseth.[12] The best fit line is shown in red.

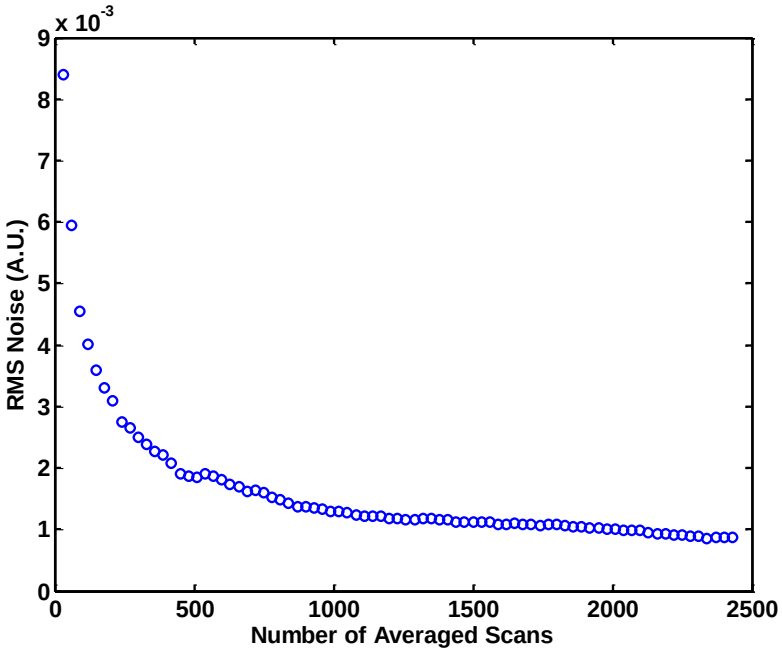


Figure II-12. The RMS noise from the 5-125 cm^{-1} spectral range plotted as a function of the number of averaged scans.

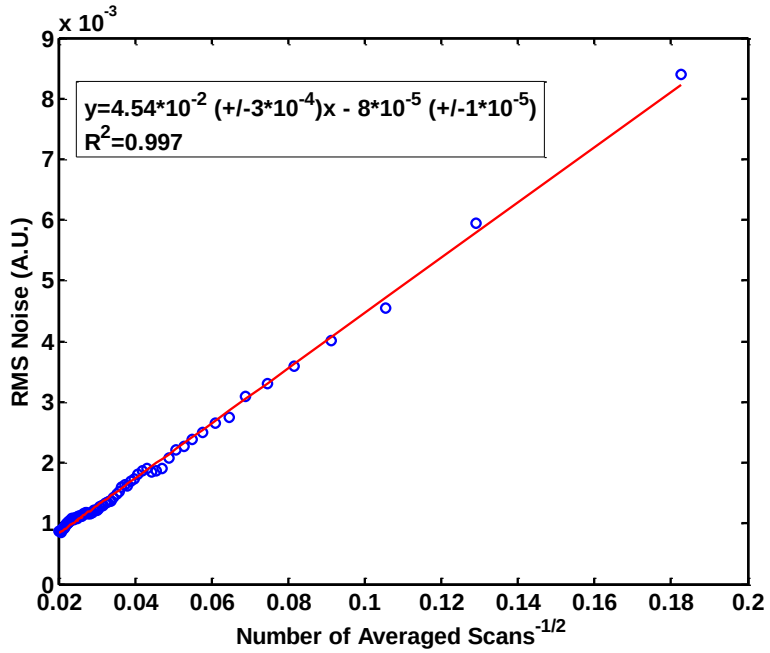


Figure II-13. RMS noise plotted as a function of the reciprocal of the square root of the number of averaged scans.

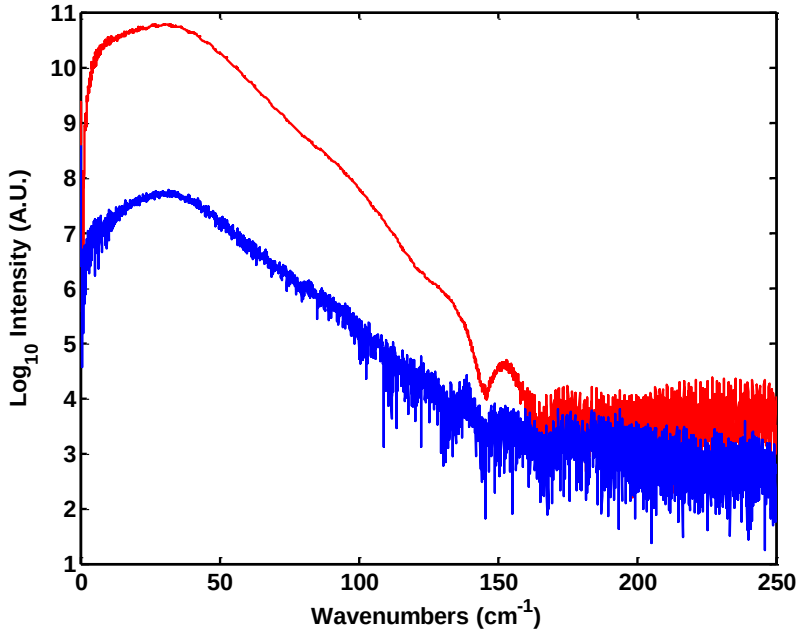


Figure II-14. Single beam spectrum magnitude (red) and standard deviation (blue) from three replicate spectra collected at a resolution of 0.225 cm^{-1} as the average of 1800 scans per spectrum.

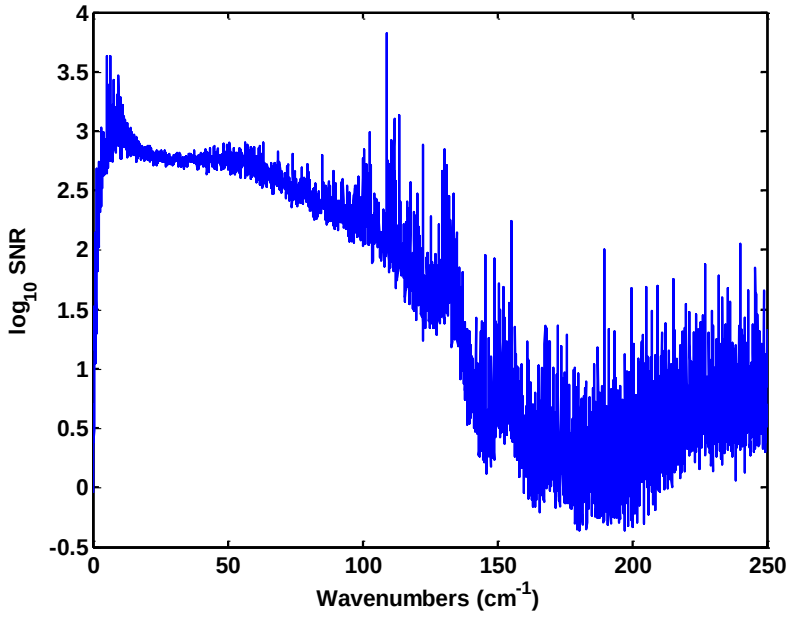


Figure II-15. Base-10 logarithm of the SNR calculated from the ratio of the mean single beam magnitude and standard deviation. Note the steep drop in SNR around 130 cm^{-1}.

Table II-2. RMS noise calculated from three 100% lines. Mean, standard deviation and SNR are provided for several spectral regions.

Spectral Region (cm^{-1})	RMS noise (mA.U.)					SNR
	SB 1/SB 2[1]	SB 2/SB 3[1]	SB 1/SB 3[1]	Mean	Std. dev.	
0 - 2.5	100	26	76	67	38	6
2.5 - 10	0.6	0.3	0.6	0.5	0.2	895
10 - 20	0.12	0.17	0.24	0.18	0.06	2479
20 - 30	0.08	0.14	0.19	0.13	0.06	3218
30 - 40	0.07	0.14	0.18	0.13	0.06	3357
40 - 50	0.09	0.13	0.17	0.13	0.04	3343
50 - 60	0.10	0.14	0.19	0.14	0.04	3038
60 - 70	0.14	0.17	0.21	0.17	0.04	2497
70 - 80	0.24	0.24	0.28	0.25	0.02	1728
80 - 90	0.44	0.41	0.48	0.45	0.03	976
90 - 100	0.66	0.68	0.70	0.68	0.02	643
100 - 110	1.51	1.39	1.54	1.48	0.08	294
110 - 120	4.2	3.3	3.6	3.7	0.4	117
120 - 130	6.7	5.5	6.5	6.2	0.6	70

1. Single beam ratios.

TDS Shift-Correction Algorithm (SCA)

All experimental spectra presented to this point have been produced using a program referred to as the shift-correction algorithm, or SCA. Signals acquired with the TeraView TPS 1000D are typically collected in what is known as “hybrid scan mode”. This means that the signal is collected in several discrete segments (2048 data points in length) during data acquisition. Upon completion of acquisition, the segments are concatenated, producing a single TDS. Segments are joined with overlapping sections, which are averaged at the junction to provide continuity between segments. However, a series of experiments revealed irreproducibility in the placement of these segments. When replicate TDSs were collected back-to-back, some signals contained segments which were shifted relative to the same segments of other signals, while other segments appeared to be correctly positioned.

The movement of these segments was found to correlate with the position of an unexplained artifact present in the TDS. This "stutter" was first identified when spectra were acquired while the beam path was blocked, similar to the signals displayed in Figures II-16 and II-17. One stutter is located in each segment, and the position of the artifact was found to vary only by one data point, and only in segments following the first. Figures II-18 and II-19 provide expanded views of the stutter in TDSs measured with an open beam path. When signals are acquired without complete absorption of the THz pulse, the stutter artifact is superimposed with the true THz pulse waveform. Figure II-19 portrays a case in which the shift has occurred. There is no longer good agreement between the two signals, and the stutter artifacts are no longer superimposed. The impact of the shift is clearer when inspecting the difference between signals, as shown in Figures II-20 and II-21. Note the increased magnitude of the difference in Figure II-21, as well as the difference caused by a shift in the stutter artifact. The difference between signals is also magnified when the shift occurs in a portion of the signal which contains a portion of the waveform with high magnitude.

A large set of signals was collected to determine the typical shape and position of the stutter artifacts. Signals containing stutters in identical positions were selected for use, and the regions containing the stutter were averaged to produce a reference set for the shape of the stutters. Baseline correction was performed with the extracted stutters in order to eliminate the shape of the waveform on which each stutter was superimposed. Data were also collected with interruptions: scans were cancelled during acquisition at various stages so the end points of each segment could be identified. Once the end points were determined, the SCA was designed to move the signal between these points to compensate for the shift. Segment positions were identified by calculating the correlation between the reference stutter sets and baseline-corrected portions of experimental signals which are known to contain the respective stutter artifacts.

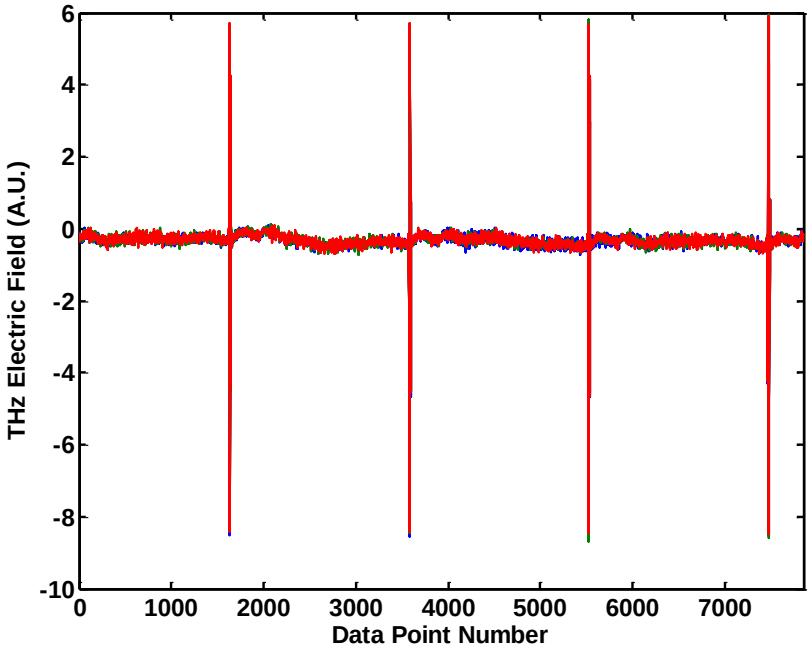


Figure II-16. TDSs collected with a blocked beam path. Signals were collected with a resolution of 0.225 cm^{-1}, which requires acquisition of four segments.

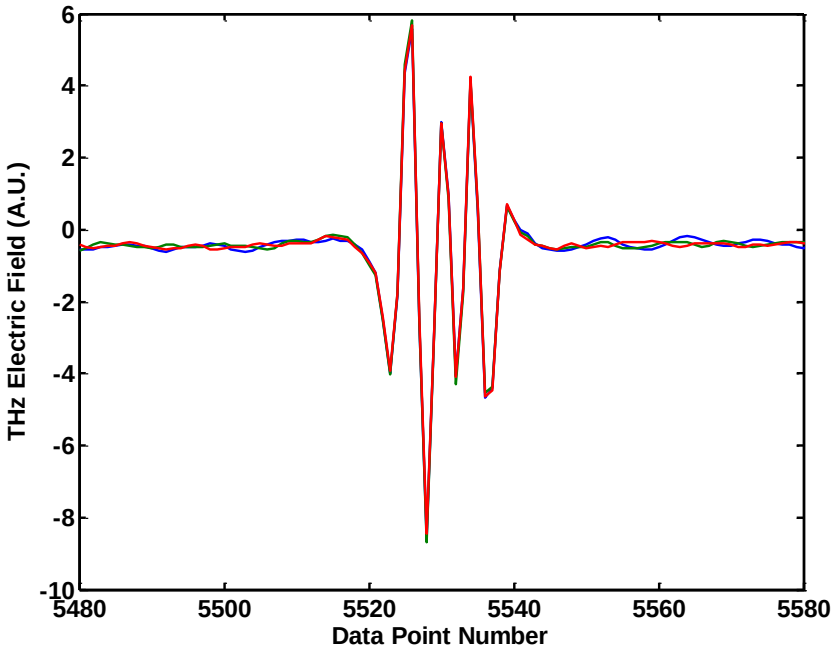


Figure II-17. Expanded view of one stutter artifact. Three replicate TDSs are shown in the plot.

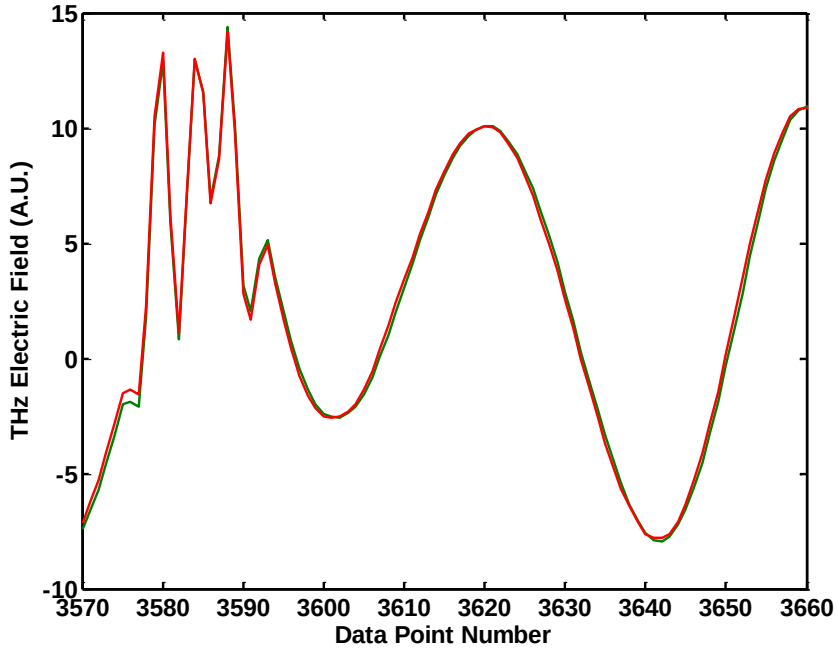


Figure II-18. Expanded view of the second stutter artifact in signals 2 and 3 from the RMS noise investigation.

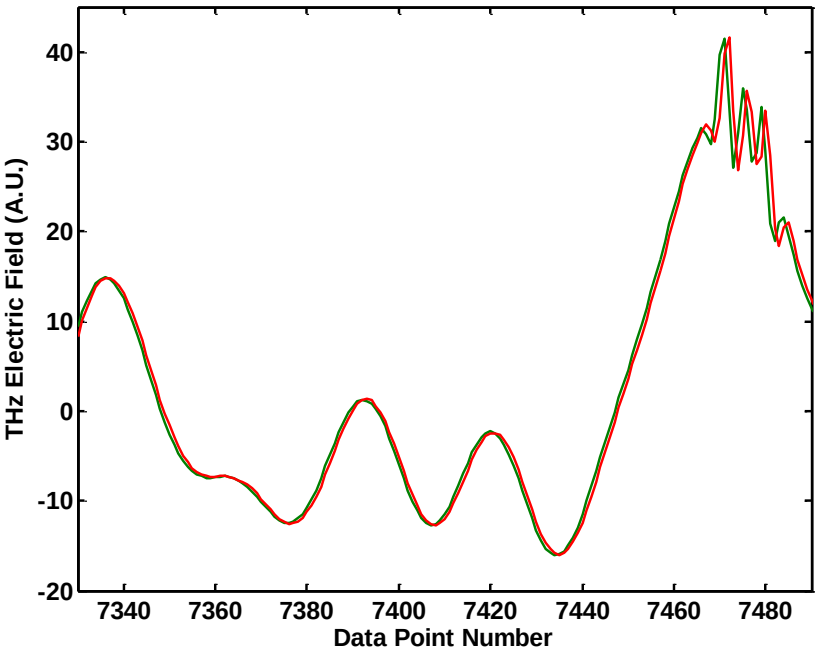


Figure II-19. Expanded view of the fourth stutter artifact in signals 2 and 3 from the RMS noise investigation.

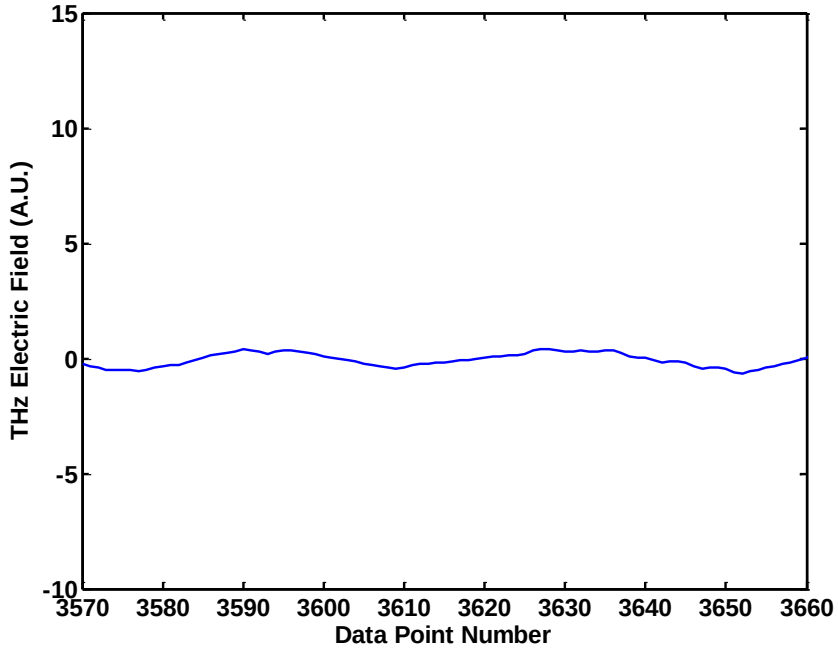


Figure II-20. Difference between signals 2 and 3, showing the region of the second stutter artifact.

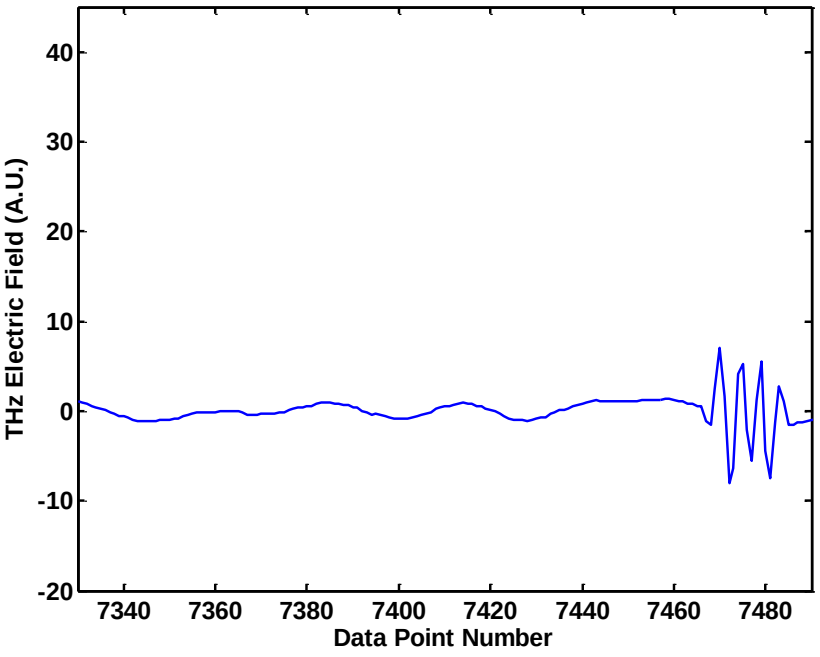


Figure II-21. Difference between signals 2 and 3, showing the region of the fourth stutter artifact.

If the calculated correlation was higher between a shifted reference stutter and the experimental signal, the segment number was recorded for shift correction. In cases where the correlation was higher between the original (un-shifted) reference stutter artifact and the experimental signal, the segment was left as-collected. Steps executed in shift correction are provided in the following outline.

1. Extract portions of experimental signals containing stutter artifacts (Figure II-22 and II-23)
2. Perform baseline correction with extracted portions to remove waveform superimposed with stutter artifacts (Figure II-24)
3. Calculate correlation coefficient between reference stutters and extracted experimental stutters (Table II-3)
4. Shift signal if correlation is higher between experimental stutter and shifted reference stutter

These steps are shown graphically in Figures II-22 through II-27 and Table II-3. Figure II-22 illustrates the reference stutter artifacts. Portions of the signal extracted for comparison of stutter artifacts are denoted by dotted lines in Figure II-23. Extracted experimental stutter artifacts are shown before (Figure II-24) and after (Figure II-25) baseline correction. Baseline-corrected stutter artifacts are overlaid with the reference stutters for comparison in Figure II-25. Boundaries for shift correction are illustrated in Figure II-26. In cases where two adjacent segments must be corrected, the signal is corrected from the lower boundary of the lower segment to the upper boundary of the upper segment in one step to avoid scarification of the averaged portion of the signal between adjacent segments. Differences between the second and third TDSs are shown with and without shift correction in Figure II-27. The standard deviations from data point number 6300 to 7400 are 0.63 and 1.10 for the shift-corrected and uncorrected signals, respectively. This comparison excludes the deviation attributed to the stutter artifact itself. Non-zero differences are still evident in Figure II-27, likely a result of sub-data

point shifts or signal drift. The impact of shift correction on the resulting frequency-domain spectra is discussed later.

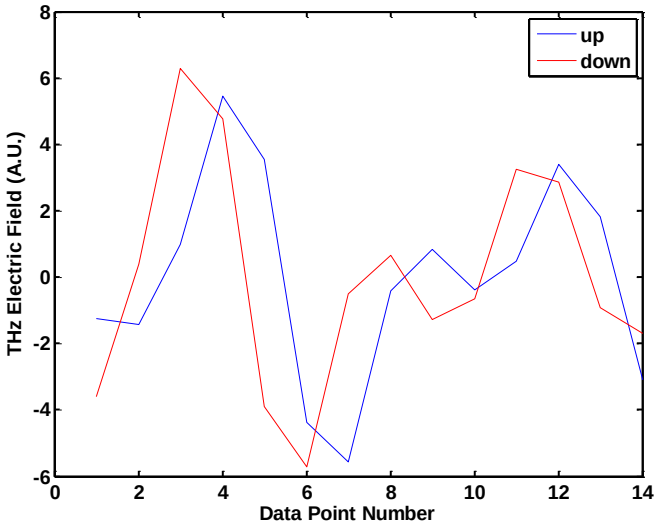


Figure II-22. Reference stutter artifacts. The red trace is shifted one data point "down", or toward shorter optical delay, relative to the blue trace. Signal correction was performed to match experimental data with the "up" position.

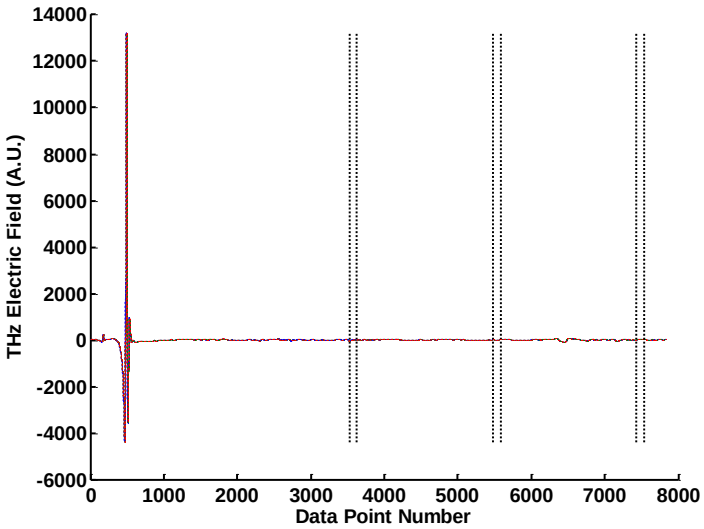


Figure II-23. Plot of experimental TDSs with dotted lines indicating extracted regions containing the stutter artifacts.

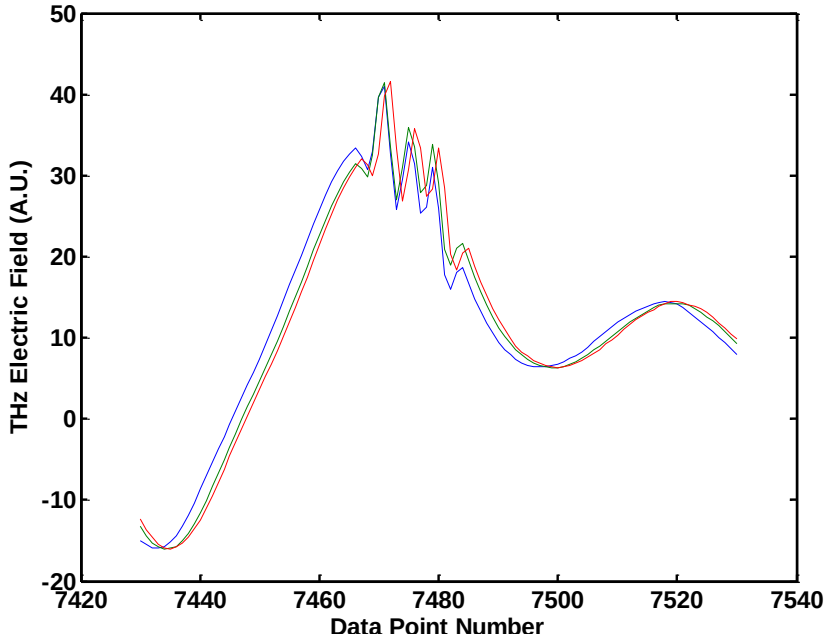


Figure II-24. Extracted stutter artifacts from the third TDS segment. No offset or other processing was applied to the signal to this point.

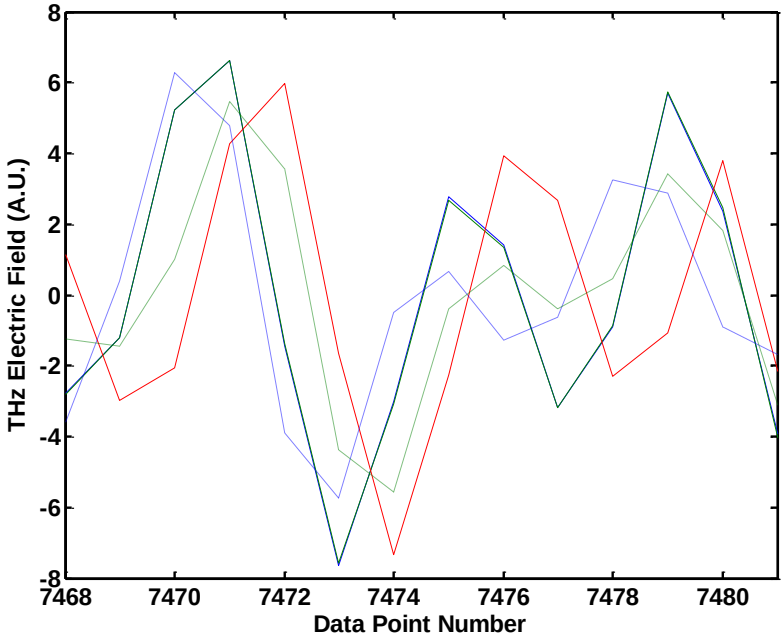


Figure II-25. Extracted stutter artifacts after baseline correction. Reference stutters, shown as dotted lines, are overlaid for comparison. The first TDS is shown in blue, the second in green, and the third in red.

Table II-3. Correlation coefficients (r) calculated between experimental TDS stutter artifacts and reference stutter artifacts.

	Down	Up	Shift
TDS 1	0.7943	0.7746	Yes
TDS 2	0.7932	0.7796	Yes
TDS 3	-0.1624	0.7124	No

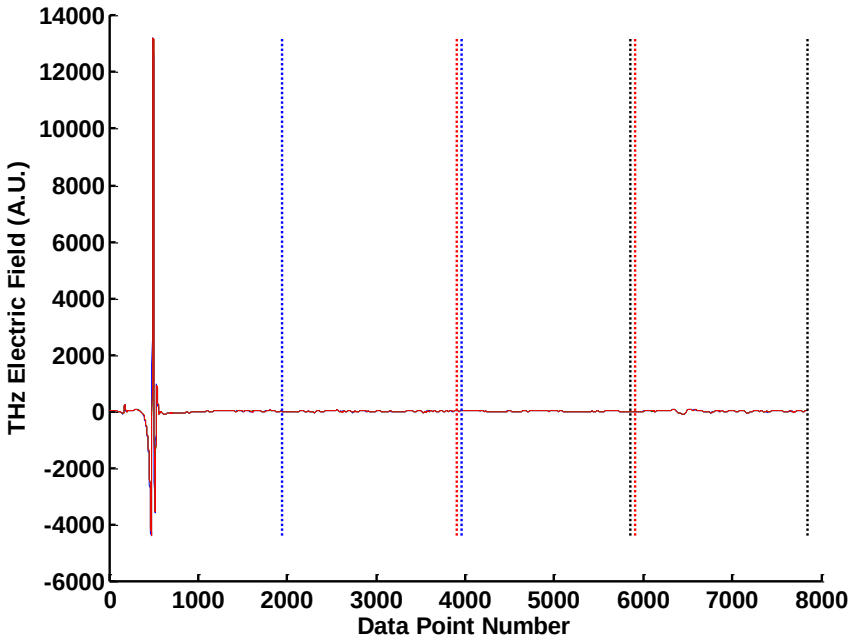


Figure II-26. Boundaries used for shift correction for each segment. The second segment is contained within the blue dotted lines, the third within the red dotted lines, and the fourth within the black dotted lines.

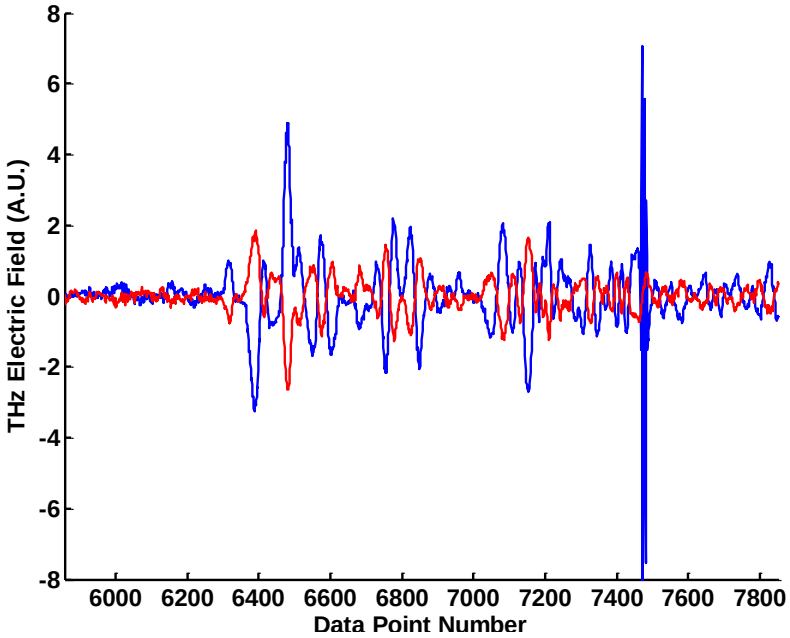


Figure II-27. Difference between TDS 2 and TDS 3 with (red) and without (blue) shift correction.

Correction of segment shifts provides significant enhancement in the calculated RMS noise while simultaneously improving the overall appearance of the absorption spectrum. Figure II-28 contains hundred percent lines calculated with and without shift-correction. Table II-4 provides the mean RMS noise and SNR calculated from spectra determined with each method. RMS noise values are greater by a factor of 2.3-9.5 when the SCA is not employed. Spectral peaks are also more reproducible when shift-correction is employed, as shown in Figures II-29 and II-30. Note the change not only in the position, but also in the shape of the spectral peaks when shift-correction is not performed, as illustrated by Figure II-30. Gas-phase spectral peaks are known to shift when the analyte is observed at various pressures, but the shift is expected to be miniscule (0.330 MHz/Torr,[156] or 1×10^{-5} cm^{-1}/Torr). Furthermore, the overall pressures at which these samples were prepared varied only by 0.5 Torr, a change which is not sufficient to induce a visible shift. Figure II-31 illustrates the difference in peak position

between water vapor absorption spectra calculated with the SCA method and those calculated without the SCA method. In some cases, the peak position is unchanged because the positions of the TDS stutters (and thus the segments) match those of the reference set, and no correction is applied.

While the SCA method provides significant improvement in the observed spectral noise, the effect of the shift cannot be eliminated completely. As mentioned previously, the junctions between the segments are averaged, producing a region which cannot be shift-corrected. The scarifying effect of the shift is still visible in corrected spectra, manifested as small periodic clusters of high-frequency artifacts. These features could not be attributed to absorbing species, as their widths are much narrower than those of the true observed spectral peaks. Identification of shifts by calculation of correlation coefficients provides a robust and hands-off method for signal correction.

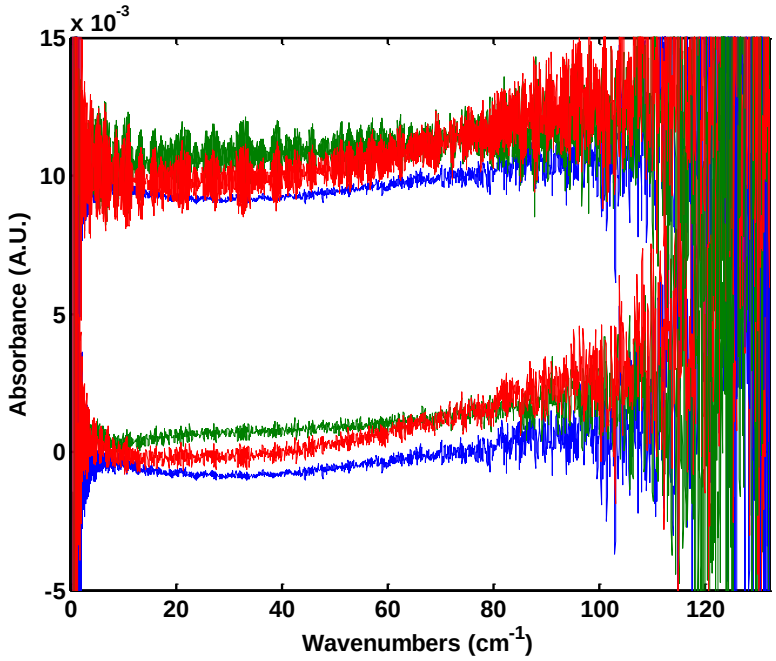


Figure II-28. Hundred percent lines calculated with (bottom) and without (top) shift-correction. Spectra calculated without shift correction are offset vertically by 0.01 A.U. for comparison.

Table II-4. RMS noise and SNR calculated from shift-corrected and uncorrected hundred percent lines.

	Shift-corrected			No shift correction		
	RMS noise (mA.U.)			RMS noise (mA.U.)		
Spectral Region (cm^{-1})	Mean	Std. dev.	SNR	Mean	Std. dev.	SNR
0 - 2.5	67	38	6	127	87	3
2.5 - 10	0.5	0.2	895	0.9	0.4	459
10 - 20	0.18	0.06	2479	0.5	0.4	842
20 - 30	0.13	0.06	3218	0.4	0.3	1081
30 - 40	0.13	0.06	3357	0.4	0.3	1045
40 - 50	0.13	0.04	3343	0.3	0.2	1501
50 - 60	0.14	0.04	3038	0.4	0.2	1151
60 - 70	0.17	0.04	2497	0.4	0.2	1198
70 - 80	0.25	0.02	1728	0.4	0.2	1006
80 - 90	0.45	0.03	976	0.8	0.3	554
90 - 100	0.68	0.02	643	1.0	0.3	439
100 - 110	1.48	0.08	294	1.6	0.1	269
110 - 120	3.7	0.4	117	3.8	0.2	113
120 - 130	6.2	0.6	70	6.3	0.3	69

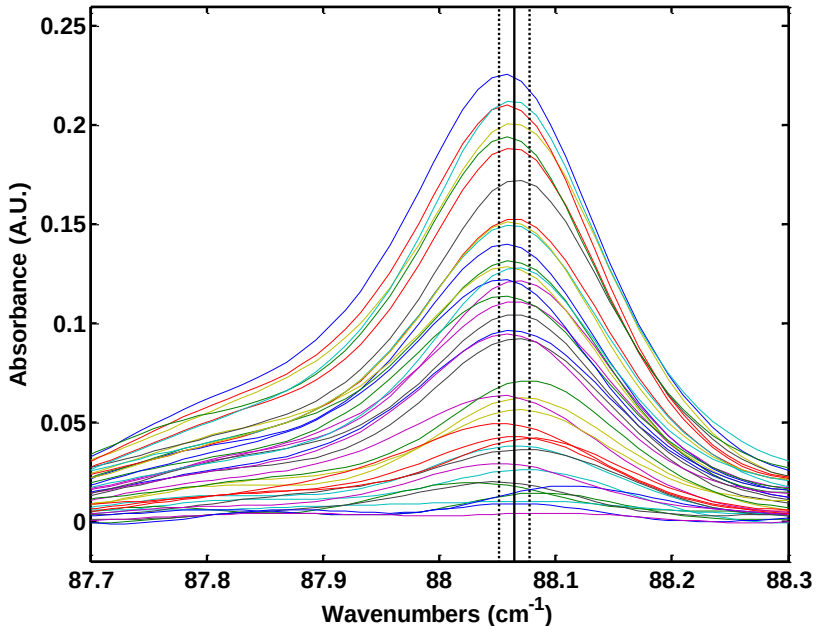


Figure II-29. Water vapor spectral peak as observed with shift-correction. The mean peak position is indicated with a vertical solid line, and the standard deviation of the observed positions is indicated by the two vertical dotted lines.

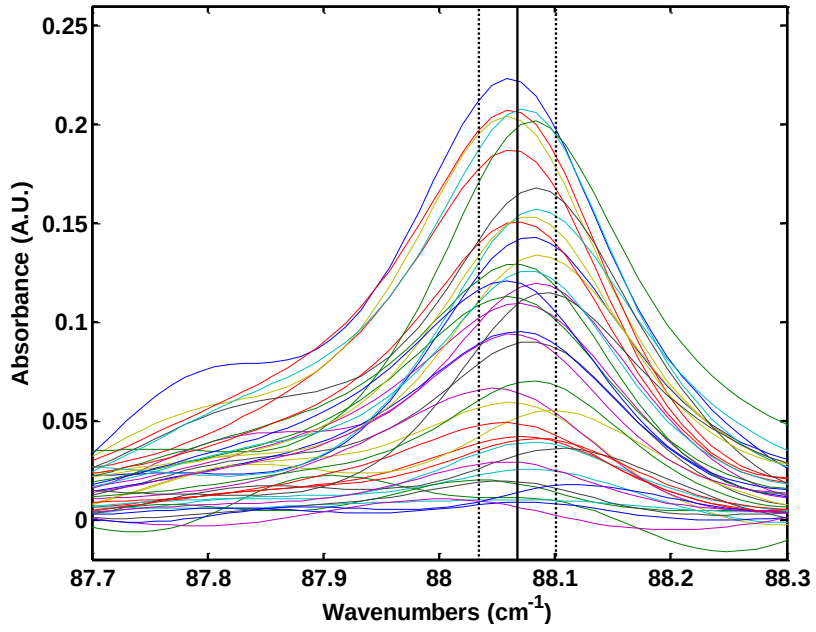


Figure II-30. Water vapor spectral peak as observed without shift-correction. The mean peak position is indicated with a vertical solid line, and the standard deviation of the observed positions is indicated by the two vertical dotted lines.

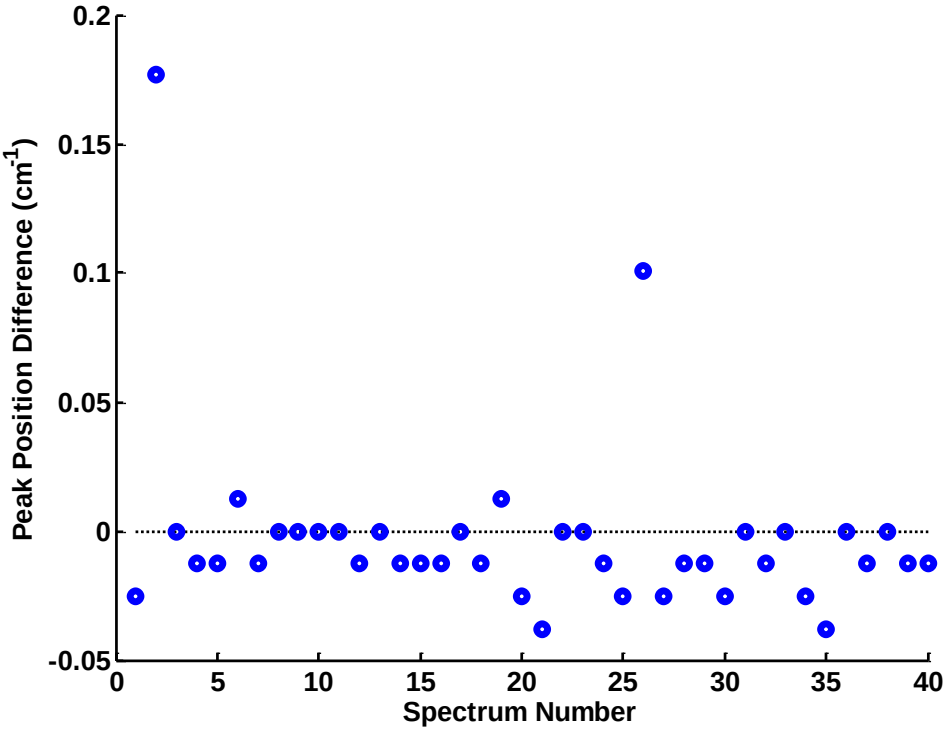


Figure II-31. Differences in water vapor peak position between spectral calculated with and without the SCA method (ν_{SCA}-ν_{noSCA}). The dotted line is plotted at $d\nu$=0 cm^{-1}.

CHAPTER III

CHEMOMETRIC METHODS

Introduction

A wide range of algorithms are available to analytical chemists for both quantitative and qualitative analysis. Complexity varies greatly from one to the next; where some methods require only a few commands or lines of code (such as the linear regression), others may require the use of computational clusters to achieve the desired result (e.g. genetic algorithms). Robust quantitative models will require at least two generic steps in implementation: (1) calibration and (2) validation, or prediction. Calibration is the step in which a predictive model is formed. A set of experimental data is used to establish a model which correlates reference concentrations to the observed instrumental response. The robustness of the model should be subsequently evaluated via the validation step. A validation set would contain experimental data similar to that of the calibration set, i.e. within the concentration range of the calibration samples, present in a similar or identical sample matrix, and analyzed under similar conditions (temperature, pressure, etc.). Further analyses may be performed, and are often warranted, to further characterize the robustness of the quantitative model.

A procedure known as leave-one-out cross-validation was used for all validation steps reported herein, unless otherwise noted. The resultant value, which is representative of the precision of the predictive model, is referred to as the cross-validation standard error of prediction, or CVSEP. It is obtained through a repetitive process by which experimental data are swapped step-by-step such that all spectra are used both in calibration and prediction steps by the end of the process. Figure III-1 depicts the calibration and prediction steps involved in the cross-validation process, and the CVSEP is calculated using Equation III-1. Low CVSEP values indicate a high level of precision in the quantitative model.

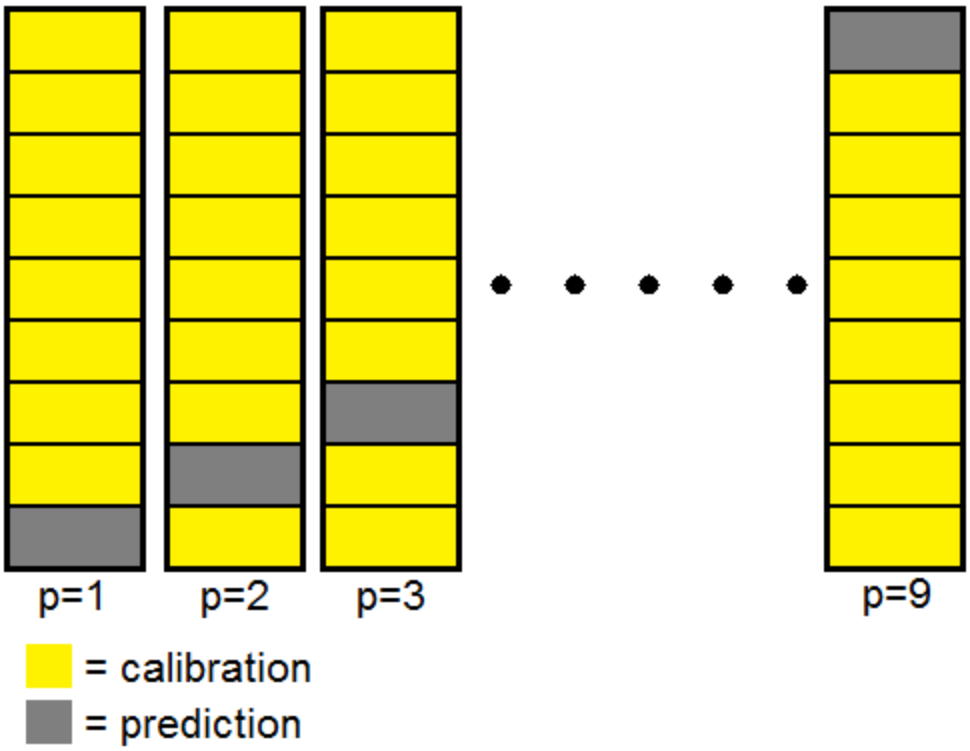


Figure III-1. Illustrated representation of the cross-validation procedure. Yellow squares represent calibration data and gray squares represent prediction data. Calculation of the CVSEP is shown in Equation III-1.

$$CVSEP = \sqrt{\frac{\sum_{p=1}^{n}(C_{ref,p} - C_{pred,p})^2}{n}} \qquad \text{Equation III-1}$$

where p is the prediction step, n is the total number of steps, which is equal to the number of samples in the leave-one-out case, $C_{ref,p}$ is the reference concentration at prediction step p, and $C_{pred,p}$ is the concentration predicted by the quantitative model at step p. It should be noted that data vector/matrix dimensions are provided in this context, where n-1 samples are used in calibration and 1 sample is used in prediction.

Linear Regression

The linear regression may be the simplest of tools available to the chemometrician. Its use is justified in spectroscopic applications where the Beer-

Lambert relationship holds, as this model assumes a linear relationship between the independent and dependent variables. The form of the assumed model is given in Equation III-2:

$$y = \beta_0 + \beta_1 x + \varepsilon \qquad \text{Equation III-2}$$

where y is the dependent variable (A.U.), β_0 represents the y-intercept (A.U.) which may result from the absence of a blank measurement or poor baseline correction, x represents the independent variable (analyte concentration, e.g. ppm), β_1 represents the slope, a measure of sensitivity (A.U./ppm), and ε represents the residuals (A.U.), which are often the product of instrumental noise or interfering species.

Calibration

The algorithm may be implemented by choosing a point on the analyte absorption spectrum, preferably one where the absorbance of the analyte is high but still within the linear range, in a spectral region with a high SNR, and a point with minimal or no overlap with interfering species. This would provide an (n-1) ×1 vector of absorption values, where n is the total number of samples. Likewise, the variable x would be an (n-1) × 1 vector of concentration values, except that it is augmented with a column of ones (n-1 × 1 in size) to provide an estimated y-intercept. If the x variable is not augmented with this column, the best fit line will have a y-intercept of zero, and only the slope term parameter would be estimated by the regression. Calibration is performed as shown in Equation III-3.

$$\beta = (x^t x)^{-1} x^t y \qquad \text{Equation III-3}$$

The 2 × 1 vector β contains the regression parameters. It is important to note that the order of the parameters in this vector depends on the order of the columns in the x variable. The superscript "t" represents transposition of the affected variable. Estimated absorption values can be obtained by multiplying the regression coefficients by the x matrix as shown in Equation III-4.

$$\hat{y} = x\beta$$ Equation III-4

Predicted absorption values are contained in the variable $\hat{y}$, an (n-1) ×1 vector. The "hat" symbol (ˆ) is typically used to represent predicted values. Residuals may be calculated via Equation III-5.

$$\varepsilon = y - \hat{y}$$ Equation III-5

Prediction

Once calibration has been performed, prediction is executed via what is known as the "inverse regression", whereby the regression parameters and experimental y-values from an unknown are used to predict the corresponding x value, or concentration. Simple algebraic rearrangement of Equation III-2 provides the inverse regression,

$$\hat{x} = \frac{y - \beta_0}{\beta_1}$$ Equation III-6

where $\hat{x}$ is the set of predicted concentrations and the residual in the measured value y is assumed to be zero. Whenever linear regression results are reported in this thesis, the best fit line is reported as calculated via Equation III-4 including all samples, but the CVSEP is calculated by calculating the best fit line from all but one sample, and the concentration of the sample left out is calculated as described in Equation III-6.

Diagnostics

Several methods may be used to evaluate the quality of a linear regression model. Non-linear trends in the data may be observed by plotting the residuals (ε) as a function of concentration (x). Non-linearity will be expressed as curvature of this plot, a phenomenon which would involve significant deviation of several data points from the best-fit line. Outliers may be identified through analysis of the "hat" matrix, which may be calculated using Equation III-7,

$$H = x(x^t x)^{-1} x^t$$ Equation III-7

which is synonymous with Equation III-8:

$$H = xx^{+} \quad \text{Equation III-8}$$

where H is the hat matrix and the superscript + denotes a Moore-Penrose pseudoinverse operation. This holds true because the pseudoinverse of x is equal to $(x^t x)^{-1}x^t$.

Diagonal values (referred to as "leverages") in the hat matrix describe the weight of each point in the calculation of predicted y-values. Thresholds for weighting may be calculated using Equations III-9 and III-10:

$$p = \sum_{i=1}^{n} h_{i,i} \quad \text{Equation III-9}$$

$$t_H = 2p/n \quad \text{Equation III-10}$$

where p is the sum of leverages, n is the number of points in the best-fit line, t_H is the threshold, and $h_{i,i}$ represents a diagonal element of the hat matrix. Equation III-10 may in some cases include a factor of 3 rather than 2, as no standard has been established for the use of this method. Elements from the diagonal of H which exceed t_H are considered outliers.

Outliers may also be identified through the exclusion of the sample in question. If a point is suspected to be an outlier, the CVSEP may be calculated for a series of linear regression models which do not include the point. The CVSEP values calculated with and without the point may then be compared using the F-test, a statistical method commonly used in the comparison of variance. F-values are calculated as shown in Equation III-11.

$$F_{calc} = \frac{CVSEP_m}{CVSEP_{-m}} \quad \text{Equation III-11}$$

In this equation, $CVSEP_m$ represents the error of prediction calculated with the point in question, and $CVSEP_{-m}$ represents the error of prediction calculated without the questionable point. The error of prediction without the questionable point is assumed to

be smaller than that produced by the model including the point. The variable F_{calc} represents the calculated F-value, which is then compared to a tabulated F-value selected with the appropriate degrees of freedom at the desired confidence level (95% confidence used throughout) appropriate for a 1-tailed test. If F_{calc} exceeds F_{table}, then the variance of the model with point m is considered to be significantly greater than that of the model excluding this point, and point m may be considered an outlier. This method is also used to compare the variance of other quantitative models.

Classical Least-Squares (CLS)

As mentioned previously, linear regression uses a single wavelength point for quantitation. While this provides straightforward interpretation and evaluation of model linearity, it is more sensitive to instrumental variance or interfering species than algorithms which utilize broader spectral ranges. The CLS algorithm is similar to linear regression in the mathematical process, while providing the benefit of multiple wavelength points, allowing the production of multivariate quantitative models.[274] Calibration spectra are the best fit of the points in the selected spectral range, which can average out some of the adverse impact of instrumental noise. A similar benefit is achieved in the prediction steps, making this algorithm less sensitive to random variance. Non-random variance such as baseline curvature or other artifacts should be characterized and included in the calibration steps if at all possible, as described below. Wavelength selection may also improve the predictive performance of this model by selecting spectral regions with a low SNR or regions in which there is minimal overlap with non-random artifacts or absorption from other species in the sample matrix.

Calibration

Just as the β vector of parameters in linear regression must be identified for calibration of the model, so must a matrix of parameters be identified for the CLS model. All spectra used for quantitative analyses reported in this thesis were pre-treated with a

simple polynomial (1-term or mean) baseline correction method to correct for broadband signal fluctuations prior to quantitative analysis. As such, only one set of parameters was identified for each compound, representing the absorptivity of the analyte. This is analogous to a linear regression in which only one parameter (the slope or sensitivity) is estimated. With this in mind, the vector of coefficients for each species was calculated as the sensitivity vector, shown in Equation III-12.

$$S = \frac{\bar{A}}{\bar{C}} \qquad \text{Equation III-12}$$

Sensitivity (S, in units of A.U./ppm) is represented by the n × 1 vector, where n is the number of wavelength points included in the model. The mean analyte absorption spectrum is represented by $\bar{A}$, and the mean concentration is represented by the variable $\bar{C}$. This method is suitable when the number of constituents within the sample is known and the absorptivity of each is readily calculable. Calibration may also be executed by performing a linear regression with the spectral data and concentration values, as described elsewhere.[275]

Prediction

Given the calibration procedure mentioned above, prediction of concentrations in unknown samples may be performed using a least-squares regression. The regression equation is shown in Equation III-13.

$$\hat{C} = (S^t S)^{-1} S^t A_{pred} \qquad \text{Equation III-13}$$

In Equation III-13, S is the sensitivity, or absorptivity vector or matrix produced by Equation III-12. The variable A_{pred} represents the spectrum for which the analyte concentration $\hat{C}$ is predicted. For convenience, the S matrix may be supplemented with additional vectors which encode baseline variation, such as a vector of ones for a baseline offset, or a linear function for a sloped baseline, or even higher-order polynomial functions.[275]

Diagnostics

Many of the diagnostic tools mentioned in the discussion of linear regression are also useful in evaluation of CLS models. Outliers may be identified either on the basis of the change in the CVSEP with inclusion/exclusion of a suspect data set, or by performing a linear regression of the concentration correlation (predicted vs. reference concentrations). If a linear regression is performed using the reference sample concentrations as the x data and the predicted sample concentrations as the y data, the slope would ideally be unity, and the y-intercept zero. Once a linear regression is performed, the hat matrix can be used to identify outliers based on leverages. Spectral residuals may also be obtained by calculating the difference between the predicted spectrum and the product of the concentration and sensitivity matrix, as shown in Equation III-14.

$$E_{spec} = A_{pred} - (\hat{C}S) \qquad \text{Equation III-14}$$

The spectral residuals (E_{spec}) should ideally be predominantly comprised of instrumental noise.

One benefit of CLS is that it allows simultaneous calibration and prediction of many components, though the algorithm's design requires the user to "know" the components in the sample under study. Species which contribute to the measured spectrum which are not included in the calibration and prediction steps will introduce errors, as its features will be fit with the shapes and spectra supplied in calibration. The magnitude of these errors depends on the magnitude of the spectral features that are not described by the shapes and spectra in the sensitivity matrix, as well as the spectral overlap this species has with the baseline shapes and analyte spectra.

Partial Least-Squares (PLS)

This method stands apart from the previous methods for a number of reasons in both design and performance. Its applications are rather diverse, ranging from

chemometric[276, 277] to marketing,[277] business,[277] and social science applications.[277] Several benefits may be gained from PLS implementation, such as reduction or elimination of random instrumental variance, decreased data collinearity, and a reduction in the number of independent variables.[278] The algorithm does not require the user to have comprehensive knowledge of background contributors for the generation of a robust predictive model, though correlation between the concentration of the analyte and the features of interfering species or the background can adversely impact the predictive ability of the model. Collinearity may be found in matrices where there is a strong relationship between variables, and can result in calculation of imprecise regression coefficients. Several variants of the algorithm have been reviewed in the peer-reviewed literature.[279] The variant described here and used in the reported investigations is known as PLS1. Unlike CLS, PLS1 is only capable of predicting the concentration of one analyte at a time.

Calibration

Spectral data are first processed via a step known as "mean centering" in which the mean of the calibration absorption spectra is subtracted from each calibration spectrum. This vector is also retained for the prediction steps. A similar step is executed for the vector of concentration values. The covariance matrix is then computed by multiplying the transpose of the mean-centered spectra by the concentration vector, as shown in Equation III-15.

$$cov = R^t c \qquad \text{Equation III-15}$$

Mean-centered concentrations are contained in the variable c, while the mean-centered spectral data is contained in the variable R. The covariance matrix, represented by the variable cov is then used to calculate the loading weights, as shown in Equation III-16.

$$w = cov/\|cov\| \qquad \text{Equation III-16}$$

Loading weights (w) may be calculated by dividing the covariance matrix by its Euclidean norm, as indicated by the pairs of vertical bars.

Scores are then calculated by multiplying the weights by the mean-centered spectra, as indicated by Equation III-17.

$$t = Rw \qquad \text{Equation III-17}$$

The scores (t) provide a means by which the spectral loadings may be calculated, as shown in Equation III-18. Note that this equation bears strong resemblance to the regressions of the prior methods, but in this case involves the scores, making it less sensitive to collinearity.

$$p^t = (t^t t)^{-1} t^t R \qquad \text{Equation III-18}$$

Spectral loadings (p) may be obtained by performing a linear regression with the scores and mean-centered spectral data. Spectral information not described by this set of scores is then extracted as shown in Equation III-19.

$$E = R - tp^t \qquad \text{Equation III-19}$$

The spectral information remaining (E) is analogous to the residuals from a linear regression. Concentration loadings are also calculated as shown in Equation III-20.

$$q = (t^t t)^{-1} t^t c \qquad \text{Equation III-20}$$

The concentration loadings (q) represent the concentration information encoded by the scores. Residuals may be calculated as with the spectral residuals E, as shown in Equation III-21.

$$e = c - tq \qquad \text{Equation III-21}$$

The concentration residuals (e) and spectral residuals (E) obtained to this point are treated as the concentration vector (c) and mean-centered spectral data (R), and the process is repeated for as many latent variables are desired for the model. A vector of

regression coefficients referred to as the "β -vector" (sometimes b-vector) is required for prediction, and is calculated from the calibration variables as shown by Equation III-22.

$$\beta = w(p^t w)^{-1} q^t \qquad \text{Equation III-22}$$

The regression coefficients (β) are analogous to those calculated in the linear regression method.

Prediction

Prediction using the PLS1 algorithm requires the mean calibration spectrum, regression coefficients, weights and spectral loadings calculated in the calibration steps. The spectra for prediction are mean-centered using the mean spectrum from calibration, and are then used to calculate scores, as shown in Equation III-23.

$$t_{pred} = r_{pred} w \qquad \text{Equation III-23}$$

The mean-centered spectra for prediction (r_{pred}) are multiplied by the weights from calibration to obtain the scores for prediction (t_{pred}). If more than one latent variable was used in calibration, this step is performed once for each latent variable, and at each step the spectral residuals are calculated once again, as shown in Equation III-24.

$$E_{pred} = r_{pred} - t_{pred} p \qquad \text{Equation III-24}$$

The spectral residuals (E_{pred}) are also calculated for each latent variable, and each time the spectral residuals are re-used for calculation of a new set of scores. Spectral residuals remaining after scores have been calculated for each latent variable represent the spectral information that will go un-described by the model. Prediction of the concentration of the analyte in the prediction spectrum is performed using Equation III-25.

$$\hat{c} = r_{pred} \beta \qquad \text{Equation III-25}$$

Multiplication of the regression coefficients by the prediction spectrum produces the predicted concentration ($\hat{c}$). Further explanation of the algorithm is available elsewhere.[276, 279, 280]

Diagnostics

One key concern when using PLS is that of "overfitting" in which an inappropriate number of latent variables is used. This phenomenon may be avoided by a combination of steps. One valuable "reality check" involves visual analysis of the spectral loadings included in the model. Each of the selected spectral loadings should bear some resemblance to the analyte or interfering species. If the loading is dominated by noise, the number of latent variables should be reduced. The significance of the latent variables may also be evaluated via the F-test. The error of prediction should be evaluated with varying numbers of latent variables. If the latent variable in question does not provide a significant decrease in the error of prediction upon inclusion in the quantitative model, its inclusion is often not warranted. As with CLS, the spectral residuals will provide useful diagnostic information.

When evaluating the predictive model for a mixture of samples, a particularly useful diagnostic tool known as "pure component selectivity analysis", or PCSA, should be used.[281-285] To perform this analysis, calibration models are developed using the spectra of the sample mixtures, with one model for each species. Prediction is then executed for the pure component spectrum of each component in the mixture. This procedure is carried out in a matrix-like fashion such that each pure component is involved in calibration (for prediction of all other species) and prediction (from calibration models developed with interfering species). A linear regression of the resultant concentrations provides a measure of selectivity. When an analyte is used in calibration, the prediction of the corresponding pure component spectra of the analyte should produce a slope of unity for a selective model. When the same calibration data are used in prediction of interfering species, the predicted concentrations will ideally be equal to zero, and the slope will likewise be zero. Selectivity of these models has been demonstrated to be sensitive to correlation between compound concentrations as well as spectral noise.[284]

Net Analyte Signal (NAS)

While the NAS algorithm can be utilized for quantitative efforts,[43, 286] the algorithm was utilized in the reported results for determination of multivariate analytical figures of merit. Analytical figures of merit provide practical measures of the quality of a quantitative model, though most analytical texts describe the calculation of these measures in a univariate context. The NAS provides a convenient method for estimation of these values. Implementation of the algorithm was mentioned briefly in Chapter I and the calculation of the NAS is described in Equation I-4. Selectivity may be evaluated quantitatively via the SEL coefficient or qualitatively via the appearance of the NAS regression coefficient vector, which may be calculated using Equation III-26.

$$\beta_{NAS} = NAS_{k,un} / \left\| NAS_{k,un} \right\|^{\mathbf{2}} \qquad \text{Equation III-26}$$

When a selective model has been constructed with PLS, the β vector of the PLS model will bear strong resemblance to the NAS regression coefficient (β) vector. Vertical bars in Equation III-26 denote the Euclidean norm. The relation of the NAS to the measured signal may be better understood via the graphical representation provided in Figure III-2.

Equation I-5 provides the SEL ratio as proposed by Lorber et al.,[43] but it may also be represented by the sine of the angle θ shown in Figure III-2, as shown in Equation III-27.

$$SEL_{an} = \sin\left(\frac{NAS_{an} \cdot r_{an}}{\|NAS_{an}\| \|r_{an}\|}\right) \qquad \text{Equation III-27}$$

In Equation III-27, the net analyte signal is represented by NAS_{an}, the analyte response by r_{an}, and the selectivity by SEL_{an}. Like the selectivity ratio described by Equation I-5, this coefficient ranges from zero to unity, with unity representing complete selectivity, and zero being representative of no selectivity. Other multivariate figures of merit can be calculated using methods proposed by Lorber.[43, 287] The multivariate error may be calculated by Equation III-28.

$$\varepsilon_k = \sqrt{\frac{r_k{}^t(I - R_{-k}R_{-k}^{+})r_k}{m-n}}$$ Equation III-28

Equation III-28 is used to estimate the error in the projection of the NAS onto the analyte response, as calculated in the numerator of the fraction. The error of projection is normalized by the number of degrees of freedom, which is the number of wavelength elements (m) less the number of interfering species (n) in the R_{-k} matrix.[287] The error ε_k for analyte k is thus not only dependent on the number of wavelength points in the spectral region of interest, but also on the number of interfering species and the magnitude of their spectral features.[287]

Once the error has been calculated, the multivariate SNR may be calculated using Equation III-29.[287]

$$SNR_k = \frac{r_k^t NAS_k}{\varepsilon_k \|NAS_{an}\|}$$ Equation III-29

Like the error, the SNR (SNR_k) of species k is dependent on the composition of the background. The limit of detection (LOD) may also be calculated using Equation III-30.[288]

$$LOD_k = 3\varepsilon_k \|\beta_k\|$$ Equation III-30

The multivariate LOD (LOD_k) is equal to three times the multivariate error for analyte k multiplied by the norm of the regression vector.

Digital Fourier Filtering (DFF)

Many pre-processing methods may be used to enhance spectral features of an analyte relative to background contributors such as noise or interfering species. As mentioned in Chapter I, apodization may be used to weight different frequencies prior to the FFT. These functions will only act as low-pass filters, removing the high-frequency components which are assumed to consist mostly of noise. However, when working with gas-phase samples, the time-domain signature of the analyte often extends near the end or

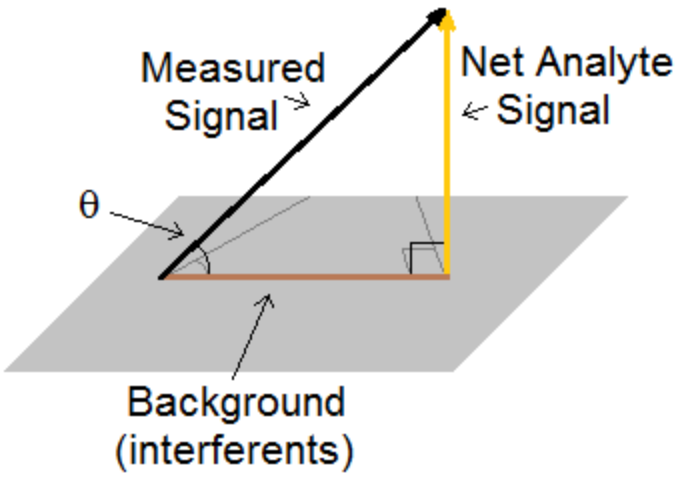


Figure III-2. Graphical representation of the NAS. The brown line represents the background contributors and spectral signatures of interfering species, which are represented by the matrix R_{-k} in Equation I-4. The black line labeled "measured signal" is the recorded analyte spectrum, represented by the vector r_k in Equation I-4. The NAS, which is the portion of the spectrum that is orthogonal to the background contributors and interfering species is represented by the yellow line and the vector $NAS_{k,un}$ in Equation I-4.

beyond the end of the sampled time window, so apodization can result in profound scarification of spectral features.

A particularly attractive alternative to this type of pre-processing is DFF,[289, 290] which allows enhancement of discrete spectral frequency components with an infinite number of possible filter shapes. Digital filtering may be executed starting with Equation III-31.

$$DF = FFT(A) \qquad \text{Equation III-31}$$

In Equation III-31, the experimental absorption spectrum is represented by the variable A, the FFT algorithm by FFT, and the digital frequency spectrum by DF. In this context, the absorption spectrum is treated in a manner analogous to the handling of the TDS as described in Chapter I. Once the digital frequency spectrum is obtained, it is filtered by convolution with the filter shape as shown in Equation III-32.

$$DF_{filt} = fDF \qquad \text{Equation III-32}$$

The digital filter (f) is multiplied by the digital frequency spectrum to obtain the filtered digital frequency spectrum (DF_{filt}). Following convolution, the filtered spectrum may be restored to the frequency domain via the inverse FFT, as shown in Equation III-33.

$$A_{filt} = iFFT\left(DF_{filt}\right) \qquad \text{Equation III-33}$$

The inverse FFT ($iFFT$) of the filtered digital frequency spectrum produces the filtered frequency-domain spectrum (A_{filt}).

Band-pass filters were applied for all filtering procedures reported in Chapter IV. This filter type allows a discrete range of frequencies to pass while discriminating against frequencies outside the band. During optimization, under certain circumstances a set of parameters would be suggested in which either the low-frequency or high-frequency cutoffs exceeded the range of possible digital frequencies. In these cases, the filter shape was converted to a low-pass, or high pass filter, respectively. A Gaussian shape was included at the cutoff frequency to prevent sharp discontinuity in the digital filter, which can introduce spurious features in the filtered spectra. The described filter function may be represented by a series of equations: Equation III-34a through Equation III-34c.

$$dv \leq \nu_{0l} \qquad f(dv) = exp\left(\frac{[dv - \nu_{0l}]^2}{2{\nu_{0l}}^2}\right) \qquad \text{Equation III-34a}$$

$$\nu_{0l} < dv < \nu_{0h} \qquad f(dv) = 1 \qquad \text{Equation III-34b}$$

$$\nu_{0h} \leq dv \qquad f(dv) = exp\left(\frac{[dv - \nu_{0h}]^2}{2{\nu_{0h}}^2}\right) \qquad \text{Equation III-34c}$$

Filter cutoffs at the high-frequency and low-frequency ends of the band are given the symbols ν_{0h} and ν_{0l}, respectively, and digital frequency is represented by the symbol dv. Each equation holds under certain conditions, provided to the left of the respective equation. The filtering process is shown in Figure III-3 through Figure III-7.

Figure III-3 illustrates the THz-frequency absorption spectra of propionaldehyde vapor at different ppm concentrations with 1 atm overall pressure. The spectra which appear to be composed of a single broad peak centered near 17 cm^{-1} with a tail extending out to ~65 cm^{-1} are in fact a dense collection of rotational transitions. Comparatively sparse features spread across the high-frequency end of the spectrum are the result of rotational transitions of water vapor, present in minute amounts due to leakage and hygroscopicity of the aldehyde. The spectrum is predominantly noise at frequencies above 100 cm^{-1}. Digital frequency spectra of the absorption spectra in Figure III-3 are displayed in Figure III-4, illustrating the rich frequency content of the absorption spectra.

The band-pass filter is shown in Figure III-5, which is applied to the digital frequency spectrum in Figure III-4 to produce the filtered digital frequency spectrum shown in Figure III-6. Note the near-zero magnitude of digital frequencies beyond the high-frequency cutoff (ν_{0h}). The inverse FFT of the filtered digital frequency spectrum produces the filtered frequency-domain absorption spectrum in Figure III-7. High frequency components of the spectra have been eliminated, as evidenced by removal of the water vapor absorption features, as well as some degree of de-noising. In some cases, depending on the digital frequency range passed by the filter, the filtered spectrum may contain shapes which were not present in the unfiltered spectrum, such as the troughs which are now present on each side of the strongest absorption bands of propionaldehyde near 0 – 7 cm^{-1} and 20 – 40 cm^{-1}.

Optimized filter conditions may sometimes induce changes of this sort in the spectrum, or remove low-frequency content which is known to be related to the analyte. Whatever the case, these filter conditions are selected because they improve the predictive performance of the respective quantitative model. The ultimate benefit of this pre-processing method is that it provides a method by which spectral variance not associated with the analyte may be selectively removed. In every instance, the filter was implemented with a standard deviation of 3 data points, or a digital frequency of

0.0007324 at the cutoff frequency. This value is small enough to provide efficient digital frequency discrimination without being so narrow that spurious features are introduced in the filtered spectrum.

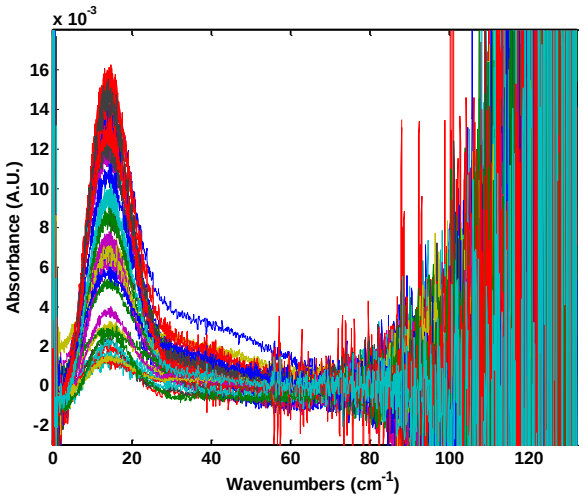


Figure III-3. The absorption spectra of propionaldehyde samples at ppm concentrations at one atmosphere overall pressure.

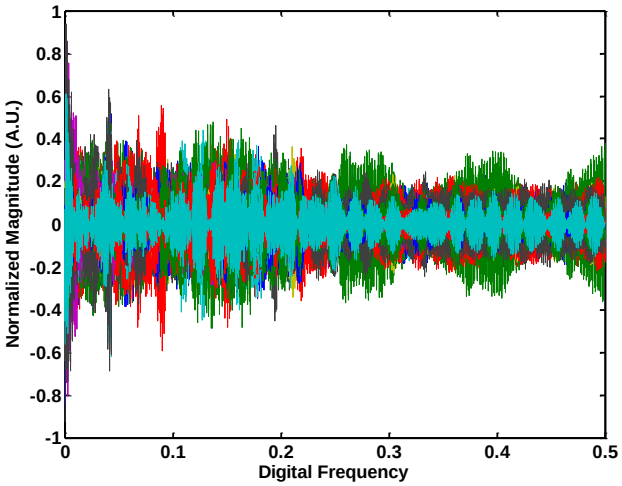


Figure III-4. The normalized digital frequency spectra of the propionaldehyde spectra shown in Figure III-3.

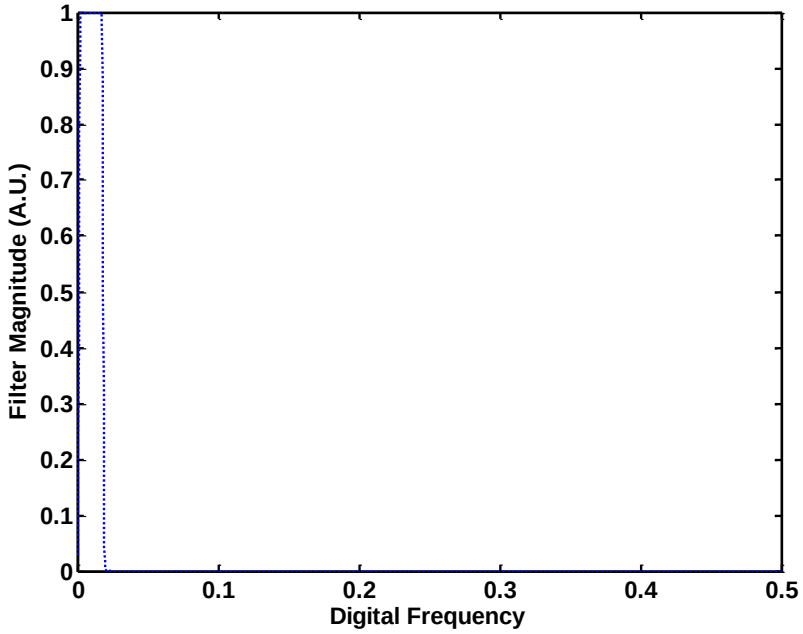


Figure III-5. The digital filter function. Low-frequency constituents will be retained, whereas digital frequencies above ~0.015 will be removed.

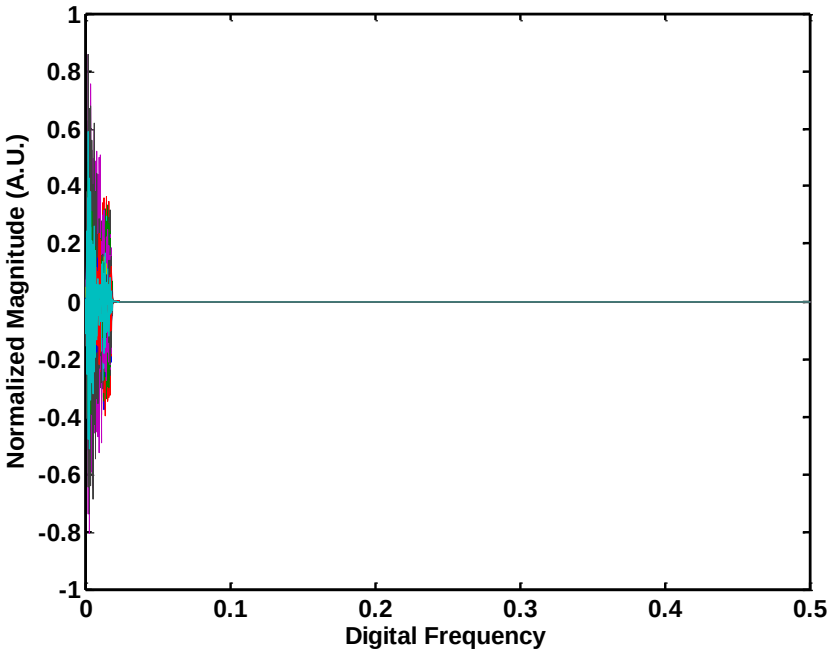
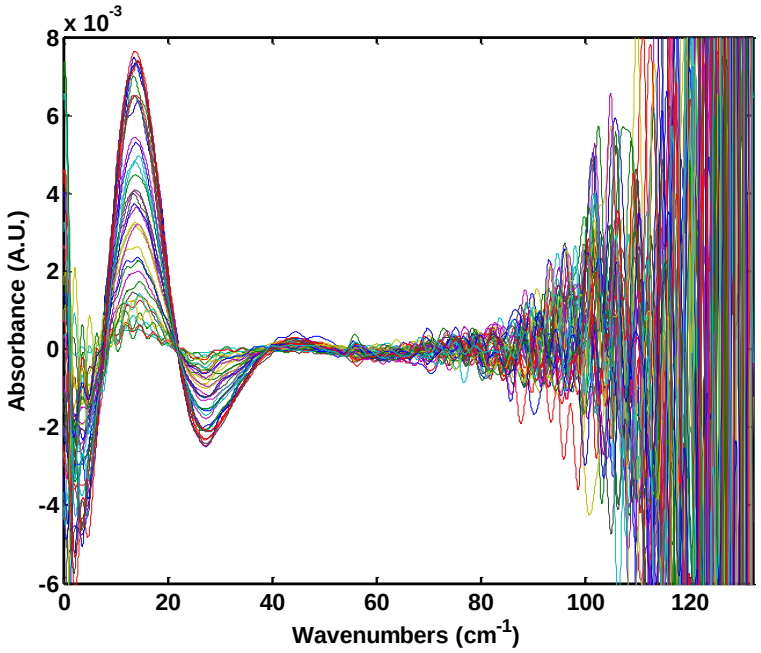


Figure III-6. The digital frequency spectrum following convolution with the filter function.



FigureIII-7. The filtered absorption spectra obtained from the inverse FFT of the filtered digital frequency spectra shown in Figure III-6.

Nelder-Mead Simplex (NMS)

When implementing quantitative models, pre-processing methods and wavelength selection (for multivariate methods) will often provide improvement in the predictive performance of the model, typically through the elimination or avoidance of non-analyte variance (as mentioned previously) such as noise or interfering species. Optimization of parameters used in pre-processing (such as filter cutoff frequencies) as well as those used for the quantitative model (such as the wavelength range or number of latent variables used) is thus warranted.

The effect of the parameters may be comprehensively investigated using a method known as the "grid search" method in which all combinations of parameters are examined and those which produce the error of prediction nearest the minimum are selected. Such a method becomes computationally prohibitive as the number of parameters is increased. For example, consider the samples discussed in Chapter IV

which represent mixtures of 8 components at various overall pressures. A total of 100 samples were prepared and analyzed. If 40 wavenumber points may be selected for the lower and upper bounds of the spectral region used for quantitation and digital filtering is used, introducing another 10 points for each of the digital filter cutoff frequencies (low and high), a total of 40 × 40 × 10 × 10 or 160,000 cross-validation procedures must be carried out. Remember also that the leave-one-out cross-validation process means that PLS calibration and prediction will occur 100 times for each set of parameters (given 100 samples), and the appropriate number of latent variables must also be determined (up to 20 maximum). Given these additional factors, the total number of PLS calibration and prediction events would be 100 × 20 × 160,000 or 320 million calibration/prediction steps. The need for alternative optimization procedures is thus obviated.

Enter the Nelder-Mead simplex algorithm, an optimization program reported by J. A. Nelder and R. Mead in 1965.[291] The program itself is mathematically simple: all steps at most require basic math operations and determination of a mean value. Optimization begins with the evaluation of a user-defined function at n+1 (or more) initial guesses, where n is the number of parameters being optimized. For example, in the case of wavenumber selection two parameters are optimized, so a minimum of three initial guesses must be provided. In the context of the reported optimizations, the function evaluated at each step is the PLS cross-validation of the gas-phase sample spectra, so the function value returned is the CVSEP. The simplex algorithm is thus designed to seek a minimum CVSEP, providing the highest accuracy quantitative model. Once the function has been evaluated at the conditions provided in the initial guesses, the resulting function values are ranked.

From this point on, five types of steps may be used to move the tested conditions towards those which provide the optimized value, referred to as reflection, expansion, outside contraction, inside contraction, and shrink. The size of each step is decided by the user, and is input to the simplex algorithm in the form of four coefficients: the

reflection coefficient (ρ, greater than 0 and less than the expansion coefficient), expansion coefficient (χ greater than 1), contraction coefficient (γ, 0 to 1), and shrinkage coefficient (σ, 0 to 1). A coefficient of shrinkage is not included in the algorithm described by Nelder and Mead, but provides an additional parameter by which the simplex algorithm's search may be controlled. The steps taken by the simplex algorithm are described in the following list.

1. Evaluate function at (at least) n + 1 initial guesses
2. Sort resulting values
3. Calculate mean of n + 1 best conditions ($\bar{x}$)
4. Calculate reflection conditions (x_r, Equation III-35)
    a. evaluate function at x_r (f_r)
    b. if f_r is greater than the best value (f_{best}), but less than the n + 1th best value (f_{n+1}), x_r is accepted, return to step 2
    c. if $f_r < f_{best}$, proceed to step 5
    d. if $f_{n+1} > f_r \geq f_n$, proceed to step 6
    e. if $f_r \geq f_{n+1}$, proceed to step 7
5. Calculate expansion conditions (x_e, Equation III-36)
    a. evaluate function at x_e (f_e)
    b. if $f_e < f_r$, accept x_e
    c. if $f_e > f_r$, accept x_r
6. Calculate outside contraction conditions (x_{co}, Equation III-37)
    a. evaluate function at x_{co} (f_{co})
    b. if $f_{co} \leq f_r$, accept x_{co} and return to step 2
    c. if $f_{co} > f_r$, proceed to step 8
7. Calculate inside contraction conditions (x_{ci}, Equation III-38)
    a. evaluate function at x_{ci} (f_{ci})
    b. if $f_{ci} \leq f_{n+1}$, accept x_{ci} and return to step 2

c. if $f_{ci} > f_{n+1}$, proceed to step 8

8. Calculate shrunk conditions ($x_{s,a}$, Equation III-39), where a is equal to the 1 through n+1 best points
    a. evaluate function at $x_{s,a}$ ($f_{s,a}$)
    b. return to step 2

These steps are also available in a concise format provided by Lagarias and colleagues.[292] The equations used to determine the conditions for each step are provided in Equation III-35 through Equation III-39.

Reflection	$x_r = \bar{x} + \rho(\bar{x} - x_{n+1})$	Equation III-35
Expansion	$x_e = \bar{x} + \chi(x_r - \bar{x})$	Equation III-36
Contraction (outside)	$x_{co} = \bar{x} + \gamma(x_r - \bar{x})$	Equation III-37
Contraction (inside)	$x_{ci} = \bar{x} - \gamma(\bar{x} - x_{n+1})$	Equation III-38
Shrinkage	$x_{s,a} = x_1 + \sigma(x_a - x_1)$	Equation III-39

The algorithm used in optimization of the reported quantitative models varies only slightly from that reported by Lagarias in that the values at the shrunk conditions were sorted with all other function values, whereas Lagarias prescribes the use of the shrunk conditions as an exclusive new set, which are not sorted with previous function values.[292] This variation is reflected in the list provided above. The simplex was also implemented with constraints that would prevent evaluation of the function value outside possible parameter values such as unrealistic wavenumber bounds or negative digital filter frequencies.

Optimization was also performed without a set of termination criteria. Nelder and Mead report the use of criteria which, if achieved during optimization, will result in termination of the algorithm.[291] Instead, a set number of iterations were allowed for optimization. Given the previous example of the number of calibration/prediction pairs,

consider a case where the NMS is used in optimization with 100 iterations. At most, given 100 iterations, 1 million calibration/prediction steps would be taken (100 iterations × 20 latent variables × 100 samples × 5 simplex steps). If the simplex does not proceed to the shrinkage step with each iteration, the total number of calibration/prediction pairs could be far less. This is a prime benefit of optimization algorithms, explaining why the NMS is reported for a variety of optimization tasks.[293-299] It is easy to imagine confinement of the simplex within a "local minimum", which is only an optimal value with respect to the nearest points. The goal of the NMS is to reach a "global minimum", which represents the optimal set of conditions of all possible combinations. In the reported work, efforts were made to provide a broad set of initial guesses to prevent local confinement.

Levenberg-Marquardt Algorithm (LMA)

Reported peak fitting results were generated via non-linear least-squares regression using the LMA, which is a modified form of the Gauss-Newton search method. The Gauss-Newton method[300, 301] (considered a Taylor series method[302]) has a tendency to take large steps in its approach to suitable solutions, and so the algorithms often fail to converge.[302] An alternative to Taylor series methods is the steepest-descent method, but these algorithms tend to be plagued by the opposite problem: the sizes of steps towards the solution tend to be small, causing slow convergence. Donald Marquardt reported an improved variant of Levenberg's algorithm[303] which is able to simultaneously evaluate the appropriate step direction and step size during operation.[302] Madsen et al. provide an accessible description of the LMA and its place amongst alternative non-linear least-squares methods.[304]

The LMA includes some elements which are familiar to analytical chemists. A least-squares regression is incorporated in the algorithm, though it is utilized in an iterative fashion, as the algorithm is designed for problems which are not linear with

respect to parameters, a condition which is met by the problem of peak fitting. Initially, several starting parameters must be selected by the user to enable the LMA to rapidly approach the solution. Since the LMA was used for peak fitting, initial guesses consisted of peak positions (center frequencies), peak heights, and peak widths. A parameter known as the "lambda multiplier" (ν) was implemented for adjustment of the step size, as well as a variable known as the Levenberg parameter (τ) which scales the initial step size.[304] The algorithm proceeds with calculation of the Jacobian matrix, as shown in Equation III-40.

$$J_{i,j} = \frac{f(\mathrm{x}) - f(\mathrm{x} + \mathrm{p})}{p} \qquad \text{Equation III-40}$$

The Jacobian matrix J is calculated by evaluating the function f at a set of parameters x for all observations (resolution elements) i for each parameter. In the case of peak fitting, the function is that of the expected line shape (to be described in Chapter IV). The parameter x_j is perturbed by amount p, which results in a numerator which describes the change in the function as a result of p, and is ultimately normalized by the value p. All parameters other than j are not perturbed. In terms familiar to analytical chemists, the Jacobian matrix resembles something like a collection of sensitivity spectra for various analytes.

At the start of the LMA, step size is calculated by determination of the Hessian matrix, which may be determined using the Jacobian calculated in Equation III-40. Equation III-41 describes the calculation of the Hessian matrix.

$$H = J^t J \qquad \text{Equation III-41}$$

The Hessian matrix (H) is a square matrix with of size m × m, where m is the total number of parameters. Once again, the superscript t represents transposition. Initial step size may be determined using Equation III-42.

$$\mu = \tau \max(\mathrm{h_{j,j}}) \qquad \text{Equation III-42}$$

The step size μ (often represented by the variable λ[302, 305]) is the product of the Levenberg parameter τ and the maximum diagonal element ($h_{j,j}$) of the Hessian matrix.

Once the step size has been determined, a set of perturbations (d) is determined using a least-squares regression, as shown by Equation III-43.

$$d = -\left(J^t J + \sqrt{\mu I .* J^t J}\right)^{-1} J^t \left(A_{exp} - f(x)\right) \qquad \text{Equation III-43}$$

Though it may not be apparent at first glance, the portion $(J^t J)^{-1} J^t$ of Equation III-43 has the same structure as the least-squares regressions of Equation III-3, Equation III-13, Equation III-18, and Equation III-20. The $\sqrt{\mu I .* J^t J}$ term weights the regression, providing control of the step size, representing the square root of the step size parameter multiplied by the identity matrix I, multiplied in an element-by-element fashion by $J^t J$. In this manner, the "sensitivity spectra" of the Jacobian matrix J are iteratively fit to the residuals between the experimental spectrum A_{exp} and the modeled spectrum $f(x)$. Published accounts often include the identity matrix in place of the Hessian in the weighting term $\sqrt{\mu I .* J^t J}$,[302, 304, 305] but the Hessian should provide parameter-specific weighting, as different parameters will produce changes of differing magnitudes. Moré also indicated that the matrix used in the weighting term may be any non-singular matrix.[305]

Given the steps above, the LMA will converge well, but a metric must be implemented to help speed or stop the progression of the algorithm. The gain ratio (ρ) meets this need and may be calculated with Equation III-44.

$$\rho = \frac{\sqrt{\sum\left(f(x+d) - f(x)\right)^2}}{\frac{1}{2} d^{\mathrm{t}}(\mu d - \mathrm{g})} \qquad \text{Equation III-44}$$

Perturbations calculated via Equation III-43 are added to the original parameters x and used to provide a new modeled spectrum $f(x+d)$. The gradient (g) is calculated via Equation III-45.

$$g = J^t \left(A_{exp} - f(x + d) \right)$$ Equation III-45

If the gain ratio (ρ) is greater than 0, the step size parameter (μ) is decreased either by a factor of one third, or by $1 - (2\rho - 1)^3$, whichever is the larger of the two values. This means that the LMA is approaching the solution, so the step size is decreased. However, if the gain ratio is less than one, the step size parameter is multiplied by the lambda multiplier, the lambda multiplier is increased by a factor of two, and the fitting process is repeated. This process provides larger step sizes with the goal of moving towards the solution at a faster pace. At any point during the operation of the LMA, its progression may be halted based on different criteria such as thresholds for the minimum perturbation, minimum standard deviation between the modeled and experimental spectra, or any other number of conditions. A clear outline of steps in the LMA was provided by Madsen.[304]

CHAPTER IV
THEORETICAL BASIS FOR THZ SPECTRAL FEATURES OF GASEOUS SAMPLES

Background

As mentioned in Chapter 1, the absorption of THz-frequency radiation by gaseous species results in promotion of electrons to an excited molecular rotational state. Table IV-1 provides a comparison of several spectral ranges with the associated frequencies, wavelengths, photon energy, and transitions. The energy associated with rotational transitions lies below those of vibrational or electronic transitions, but above electron and nuclear spin transitions. A simplified energy diagram is provided in Figure IV-1 illustrating the relation of the transitions listed in Table IV-1. Electron and nuclear spin transitions are excluded from this diagram to improve readability. In Figure IV-1, upward movement corresponds to an increase in energy or absorption event, while downward movement represents a decrease in energy, emission, or non-radiative relaxation event. Various molecular properties affect the spacing and population of each of these states, as will be discussed in the following sections. The theoretical basis for these features has been described in great depth over the years by the likes of Townes and Schawlow,[306] Gordy and Cook,[13] Sugden and Kenney,[307] and Wollrab.[308] Kroto published an overview of the theory,[309] from which much of the cited material is derived.

Molecular Structure

Spectral features observed in the MW and THz regions are heavily influenced by molecular composition and structure because these characteristics impact molecular rotation. As such, it is important that several parameters related to molecular structure are discussed before delving into related theory. A key parameter which affects the peak positions related to each transition is the molecular moment of inertia. The moment of inertia is a numeric representation of the distribution of massive bodies at some distance

from the molecule's center of mass, shown for the case of a simple diatomic molecule in Equation IV-1.

$$r_1 m_1 = r_2 m_2 \qquad \text{Equation IV-1}$$

Distances from the center of mass are represented by r_1 and r_2, and the masses of the atoms are represented by m_1 and m_2. If the total distance between the two bodies is described by Equation IV-2,

$$r = r_1 + r_2 \qquad \text{Equation IV-2}$$

and by solving for r_1 and substituting Equation IV-1 into Equation IV-2, the result is Equation IV-3.

$$r = r_2 \left(\frac{m_2}{m_1} + 1\right) \qquad \text{Equation IV-3}$$

Table IV-1. Comparison of approximate frequency, wavelength, energy and transitions associated with various spectral ranges.

Spectral Range	Frequency (Hz)	Wavelength (m)	Energy (kJ/mol)	Transition Type
γ - rays	$10^{20.5}$-$10^{19.5}$	10^{-11}-10^{-12}	10^{7}-10^{8}	ionization
x - rays	$10^{16.5}$-$10^{19.5}$	10^{-8}-10^{-11}	10^{4}-10^{7}	ionization
ultraviolet	10^{15}-$10^{16.5}$	$10^{-6.5}$-10^{-8}	10^{2}-10^{4}	electronic excitation
visible	$10^{14.5}$-10^{15}	10^{-6}-$10^{-6.5}$	10^{2}-10^{2}	electronic excitation
NIR	10^{14}-$10^{14.5}$	$10^{-5.5}$-10^{-6}	$10^{1.5}$-10^{2}	bond vibration (combination and overtone)
Mid - IR	10^{13}-10^{14}	10^{-5}-$10^{-5.5}$	10^{1}-$10^{1.5}$	bond vibration (fundamental)
Far - IR	10^{12}-10^{14}	10^{-4}-10^{-5}	10^{0}-10^{1}	bond torsion, molecular rotation
THz	10^{9}-10^{12}	10^{-1}-10^{-4}	10^{-1}-10^{0}	rotation, inversion
MW	$10^{7.5}$-10^{10}	10^{-1}-$10^{-2.5}$	10^{-3}-10^{-1}	rotation, inversion
Radio	$<10^{7.5}$	$>10^{-1}$	$<10^{-3}$	electron, nuclear motions

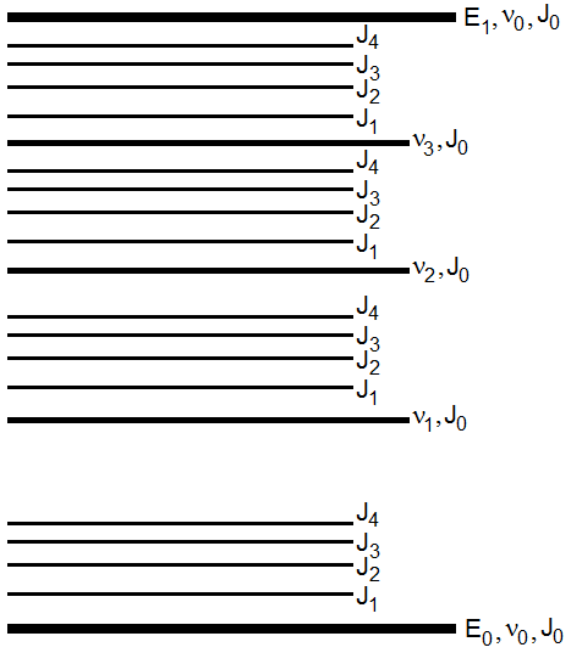


Figure IV-1. Simplified molecular energy diagram. Electronic states are indicated with the letter E, vibrational states are represented by ν and rotational states with the letter J. Other vibrational and rotational states can exist above E_1 and rotational states of a lower vibrational state may surpass the next excited vibrational state, but these are excluded to avoid clutter.

Equation IV-3 can be rearranged to solve for the distance from the center of mass to atom number 2, as shown in Equation IV-4.

$$r_2 = \left(\frac{m_1}{m_1 + m_2}\right) r \qquad \text{Equation IV-4}$$

If a similar substitution and rearrangement is performed, the distance to atom number 1 can also be determined. The moment of inertia for a rigid rotor is described by Equation IV-5.[310]

$$I = \sum mr^2 \qquad \text{Equation IV-5}$$

The product of the mass m and distance r squared is summed for each atom in the molecule to obtain the moment of inertia, I. In the case of a diatomic molecule, calculation of this value is trivial, as shown by Equation IV-6.

$$I = \left(\frac{m_1 m_2}{m_1 + m_2}\right) r^2$$ Equation IV-6

As the molecule becomes more complex, this equation is expanded to include the other masses and distances. Molecules which are not totally symmetric in two of the dimensions will also have moments of inertia in those other dimensions. Moments of inertia are typically given designations based on the magnitude of the moment: I_A is the smallest, I_B the intermediate, and I_C the largest.[307]

Classes of Molecules

As mentioned previously, molecular structure plays a significant role in the appearance of MW and THz gas-phase spectra. Molecules undergoing rotation, often referred to as "rotors" fall into four general categories of geometry[309]:

1) Spherical tops - $I_A = I_B = I_C$
2) Linear molecules - $I_A = 0$, $I_B = I_C$
3) Symmetric tops
    a. prolate - $I_A < I_B = I_C$
    b. oblate - $I_A = I_B < I_C$
4) Asymmetric rotor - $I_A < I_B < I_C$

Acetonitrile, when un-deuterated or triply-deuterated is classified as a prolate rotor, under the same constraints, ammonia is considered an oblate rotor, and the remaining six molecules of interest are classified as asymmetric rotors, as well as the singly or doubly-deuterated isotopes of ammonia and acetonitrile. The A, B, and C designations also represent molecular rotational constants, shown in Equation IV-7 through Equation IV-9.[311]

$$A = \frac{h^2}{8\pi^2 I_A}$$ Equation IV-7

$$B = \frac{h^2}{8\pi^2 I_B}$$ Equation IV-8

$$C = \frac{h^2}{8\pi^2 I_C}$$ Equation IV-9

In these equations, h represents Planck's constant, and I represents the moment of inertia for the respective axis. The molecular constants are typically reported in terms of energy or frequency (GHz, cm^{-1}, etc.). Most authors avoid quantum mechanical descriptions of the rotational motion of asymmetric rotors due to the complexity of the theoretical models. Their description is attempted here due to their prevalence both in nature and in the species of interest.

De Lucia has illustrated the effect of these parameters on the spectral features observed for symmetric tops.[312] In general, the center of the distribution of absorption peaks moves toward lower frequencies as these parameters increase, the width of the group of peaks decreases, the number of observed rotational transitions also increases, and the absorptivity is diminished as the molecular parameters decrease in magnitude.[312] This trend corresponds with an increase in the moment of inertia, i.e. an increase in mass or molecular size. Similar trends were observed by Harde[186] using a THz-TDS to study rotational spectra of methyl halides, as mentioned in Chapter I.

A set of three quantum numbers is traditionally used to describe rotational energy levels, indicated by the letters J, K, and M. Unless otherwise noted, the values of J can range from zero to positive integers, and K and M can be any integer value between ±J, including zero. Rotational quantum number J represents the angular momentum of the molecule, K represents the projection of the angular momentum on the internal figure axis, and M represents the projection of the angular momentum on a laboratory-fixed z axis. As moments of inertia change from one molecule to another, this changes not only the classification of the molecule but also the projection of the angular momentum on the two axes, thus changing the energy levels accessible by each molecule. Figure IV-2 illustrates the relationship of the vector representation of these quantum numbers in an arbitrary molecule.

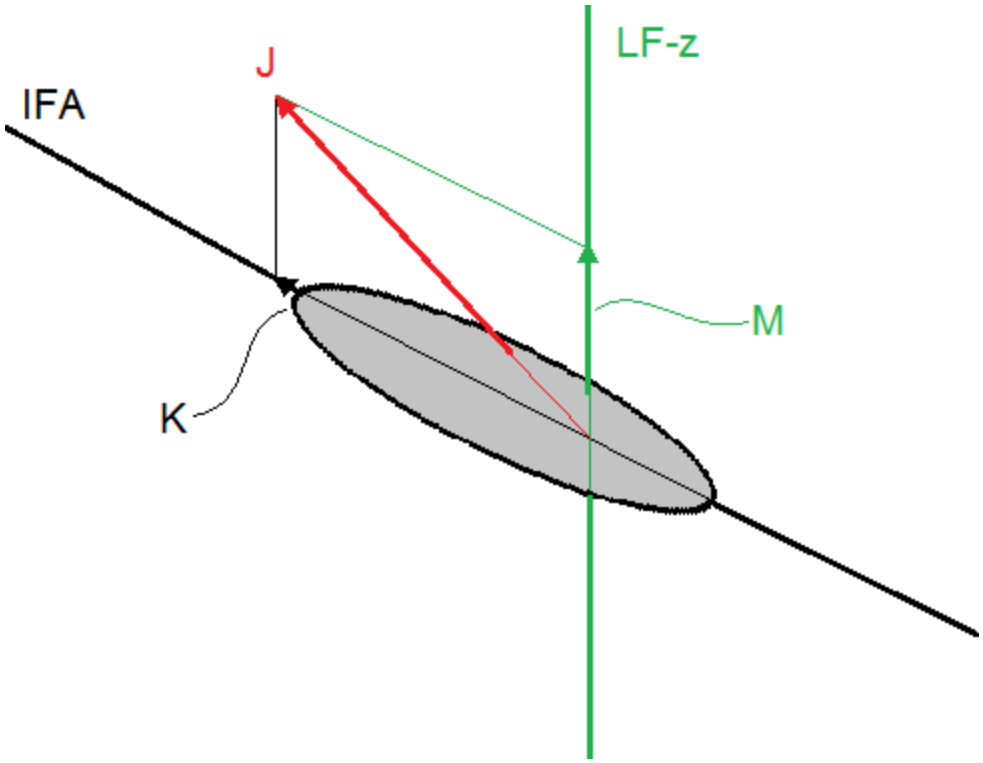


Figure IV-2. Relationship of rotational quantum numbers K and M to angular momentum quantum number J. The internal figure axis is indicated by the acronym IFA, and the laboratory-fixed z axis is represented by the acronym LF-z. As the orientation and magnitude of the angular momentum quantum number J changes, so do the values of the K and M quantum numbers.

Peak Positions

Various forces acting both within and without the molecule can affect the position and number of absorption features observed for a given species. Molecular structure plays the most significant role in the energy levels of rotational transitions. Other less-pronounced forces can act on the molecule to split energy levels into two or more closely-spaced transitions. As such, expected peak positions will be described in the context of the symmetry classifications described previously, followed by a brief overview of the less-significant contributors.

Spherical Top

Ordinarily, spherical tops do not exhibit rotational spectral features, owing to their symmetry. Symmetry of this sort prevents the molecule from possessing a permanent dipole moment, as the individual dipoles are equal and in opposing directions. The same is observed with the moments of inertia in such molecules. Under certain conditions, rotational spectral features have been observed for molecules of this category as a result of distortion of the molecular structure.[313] Rotational energy levels accessible by this category of molecule are described by Equation IV-10[309]:

$$E(J\ K) = BJ(J+1) \qquad \text{Equation IV-10}$$

where E is the energy of the state associated with the rotational quantum numbers J and K, and B is the molecular constant described previously. E is expressed in the same units of energy assigned to the molecular constant B. Each of these energy levels is 2J+1-fold degenerate for both K and M states.[309] The equations for the energy levels are derived from the Hamiltonian of the molecule, but the construction of this matrix is beyond the scope of this chapter. Please refer to one of the sources cited in the introductory paragraph for a complete description of the determination of components of the Hamiltonian.

Linear Rigid Rotor

Energy levels for rotational transitions of linear rotors may be calculated using an equation similar to that of the spherical top, shown in Equation IV-11.[309]

$$E(J) = BJ(J+1) \qquad \text{Equation IV-11}$$

The selection rules for linear rotors are $\Delta J=\pm1$, $\Delta K=0$, and $\Delta M=0, \pm1$.[308] Because $\Delta K=0$, the calculated energy levels are non-degenerate with respect to K states, but remain 2J+1-fold degenerate with respect to M.[309] Under ordinary conditions (excluding vibrational distortion and other splitting mechanisms), the rigid linear rotor absorbs frequencies with

very regular and predictable spacing, as illustrated by Kroto. If other distorting effects are neglected, the spacing is equal to even multiples of the B molecular parameter.[309]

Symmetric Tops

The rotational spectra of symmetric tops are more complex than those of linear molecules. Energy levels associated with these transitions are described by Equation IV-12 (for the prolate rotor) and Equation IV-13 (for the oblate rotor).

$$E(J,K) = BJ(J+1) + (A-B)K^2 \qquad \text{Equation IV-12}$$

$$E(J,K) = BJ(J+1) + (C-B)K^2 \qquad \text{Equation IV-13}$$

One can see from these equations that for both classes of molecules, the calculated energy level is doubly degenerate with respect to K when $K \neq 0$. When K is equal to zero, the resultant energy levels are non-degenerate. Each energy level is also 2J+1 fold degenerate in M quantum numbers. Due to the ranking of the respective moments of inertia, the energy of the rotational transitions decreases with an increasing absolute value of K for the prolate rotor, whereas the spacing increases with increasing absolute value of K for the oblate rotor. The presence of non-degenerate energy levels provides what appears to be a more complex spectrum, as different values of the K quantum number now produces a distinct transition, even when the J quantum number remains unchanged. Kroto provides a clear explanation for the degeneracy of the $K \neq 0$ states in that a change in the sign for the K value represents a change in the direction of spin,[309] but this does not affect the energy of the associated transition. Selection rules for symmetric rotors are $\Delta J = 0, \pm 1$ and $\Delta K = 0$.[309] A generic diagram of the relative spacing of such transitions is provided in Figure IV-3.

Asymmetric Rotor

Determination of energy levels for rotational transitions of asymmetric rotors is far more complex than for spherical, linear, or symmetric rotors. Ray's asymmetry parameter must first be calculated using Equation IV-14.

$$\kappa = \frac{2B - A - C}{A - C} \qquad \text{Equation IV-14}$$

In this equation A, B, and C are the molecular constants and κ is the asymmetry parameter. Values near -1 indicate a near-prolate symmetry, while values near 0 indicate a high level of asymmetry, and values near 1 represent near-oblate symmetry. The energy levels can be estimated based on the expectation values, shown in Equation IV-15 through Equation IV-18.

$$E = A\langle J_A^2\rangle + B\langle J_B^2\rangle + C\langle J_C^2\rangle \qquad \text{Equation IV-15}$$

where

$$\langle J_A^2\rangle = \frac{1}{2}\left[J(J+1) + E_{J\tau}(\kappa) - (\kappa+1)\frac{\partial E_{J\tau}(\kappa)}{\partial \kappa}\right] \qquad \text{Equation IV-16}$$

$$\langle J_B^2\rangle = \left[\frac{\partial E_{J\tau}(\kappa)}{\partial \kappa}\right] \qquad \text{Equation IV-17}$$

$$\langle J_C^2\rangle = \frac{1}{2}\left[J(J+1) - E_{J\tau}(\kappa) + (\kappa-1)\frac{\partial E_{J\tau}(\kappa)}{\partial \kappa}\right] \qquad \text{Equation IV-18}$$

The expectation values are averaged across the eigenfunctions for asymmetric rotors and provide a set of values useful for the determination of other molecular parameters such as intensities, centrifugal distortion, and quadrupoles.[309] Tabulated values reported in the scientific literature for the matrix $E_{J\tau}(\kappa)$[314, 315] are often used, as mentioned by Kroto,[309] Sugden,[307] Townes,[306] and Gordy.[13] The variables A, B, and C are the molecular constants, and E contains the energy levels of the asymmetric rotor. Other methods may be used to calculate the expected transitions for these molecules, though it is important to note that some models are only accurate for low values of J.[309]

Selection rules are dependent on the molecular structure and dipole moment vectors of asymmetric rotors, and may be determined using a procedure described by Kroto.[309]

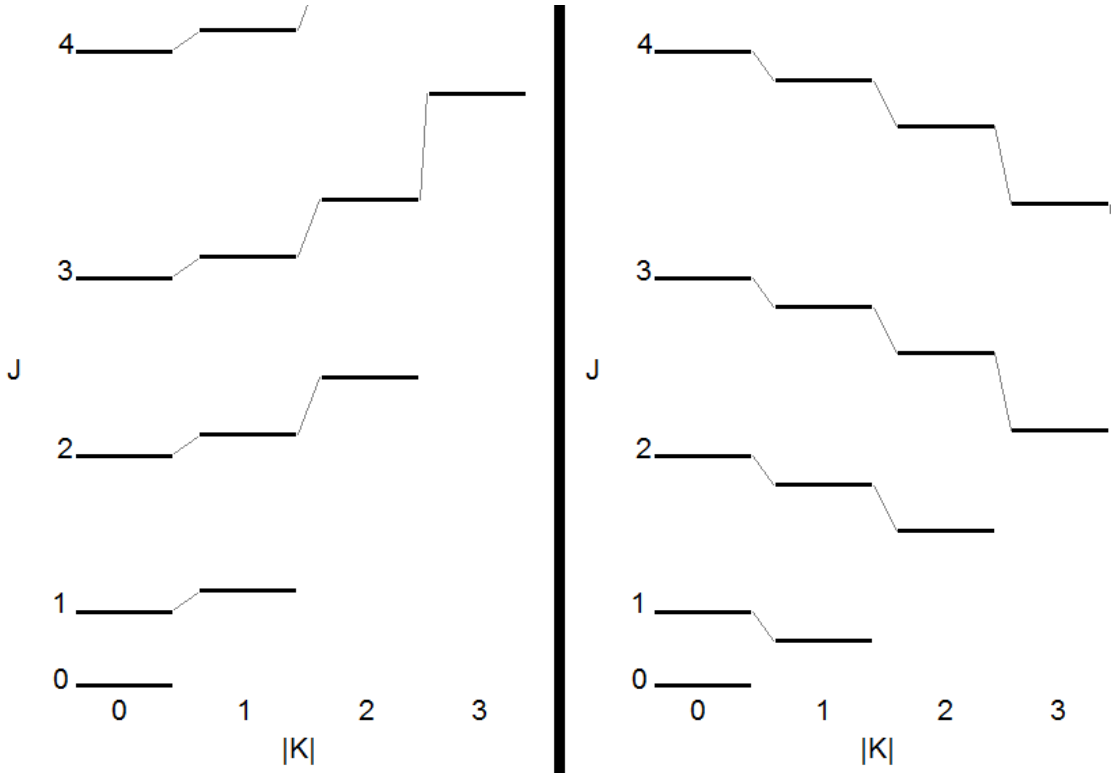


Figure IV-3. Examples of rotational transition energy diagrams for prolate symmetric rotors (left) and oblate symmetric rotors (right) for different J and K quantum numbers. Note that the x-axis is the absolute value of the K quantum number. The energy levels of a linear rotor may be represented by the levels of the K = 0 transitions.

In this class of molecules, a transition of $\Delta J = -1$ can still result in the absorption of electromagnetic radiation, as some (J+1) states may be of lower energy than some J states.[307]

Other Contributors

As mentioned in the literature overview in Chapter I and illustrated in Figure IV-1, the rotational spectra of species may also contain rotational absorption features of

molecules in excited vibrational states. These vibrational states can be characterized by various motions such as stretching, bending, scissoring, or torsion. The first three of these motions tend to be observed in the near and mid-IR regions, whereas the latter is often observed in the far-IR region. A ratio of the expected relative intensities may be determined by estimating the number of molecules in the excited vibrational state(s) vs. the number of molecules in the ground vibrational state using the Boltzmann distribution, shown in Equation IV-19.

$$\frac{N_e}{\mathrm{N_0}} = \frac{\mathrm{g_e}}{\mathrm{g_0}} \exp\left(-\frac{\Delta \mathrm{E}}{\mathrm{kT}}\right) \qquad \text{Equation IV-19}$$

The difference in energy between the excited (subscript e) and ground (subscript 0) is represented by ΔE, Boltzmann's constant is represented by k, the variable T represents the temperature in Kelvin, g represents the degeneracy of each state, and N represents the number of molecules occupying each state. At elevated temperatures, a greater number of species will be found in an excited vibrational state, according to Equation IV-19.

Several other characteristics of the molecule will affect the position and number of observed spectral features. A phenomenon known as "centrifugal distortion" is known to affect peak positions as a result of altered bond length. To this point in these explanations, the rotors have been treated as rigid bodies, simplifying the description of the observed features. In reality, when rotational transitions can occur when the bonds of a molecule are not at their equilibrium lengths, which results in slight changes to the moments of inertia. The distortion constant is the molecular parameter which describes this effect, often represented by the capital letter D. Equation IV-20 provides an example of the calculation for the change in energy levels associated with rotational transitions of a linear rotor.[311]

$$\Delta E(J) = 2B_v(J+1) - 4D_v(J+1)^3 \qquad \text{Equation IV-20}$$

The distortion coefficient is represented by the variable D_v, which is representative of the bond strength. For small values of D, the deviation from the rigid

rotor approximation is small, indicating a strong bond. When values of D are large, the deviation from the rigid rotor approximation is great, indicating a very weak bond. For rigid bonds, it is difficult to experimentally determine the distortion coefficient(s) using lower frequencies because the deviation from the theoretical position is miniscule.[311] Strandberg has also described the phenomenon and characterization of its effect on the spectral features of various gaseous species.[316]

Rotational spectra are said to contain "fine structure" as well as "hyperfine structure". Molecular rotational transitions are split into multiple closely-spaced transitions by electronic and nucleonic spin.[311] Kroto describes four distinct splitting phenomena related to electronic motion and spin: spin-orbit coupling, spin-rotation interaction, spin-spin interaction, and Λ-doubling. The energy of spin-orbit coupling is a product of the electronic angular momentum and the electron spin angular momentum, as shown in Equation IV-21.

$$E_{s-o} = AL_cS_c \qquad \text{Equation IV-21}$$

The electronic angular momentum is represented by L_c, and the electron spin angular momentum is S_c. The variable A is the spin-orbit coupling constant, a quantity both calculable and amenable to experimental determination.[309] Spin-orbit coupling energy is represented by the variable $E_{s\text{-}o}$. Spin-rotation interaction is the result of electrons orbiting nuclei and molecules at a distance, which forms weak magnetic fields.[309] These interactions tend to be most important at high rotational levels.[309] Spin-spin interactions are more significant at low rotational levels, and tend to be observed in molecules where electronic spin is aligned along the internuclear axis.[13] Λ-doubling is caused by coupling of electronic angular momentum to the molecular axis, but is relatively rare as there are few stable molecules which have ground states with electronic moments.[13]

Peak Magnitude

The magnitude of absorption features, often referred to as "intensities"[13, 308, 309] can also be estimated with knowledge of the molecule's structure and composition. The term "intensities" is often used to represent the magnitude of spectral features as a result of absorption or emission, though features reported in this thesis are only the result of absorption events. As mentioned previously, spherical tops only undergo rotational transitions when the molecule is distorted, though once distorted the molecule can then be classified as a symmetric top so explicit definition of absorptivities for these molecules is unnecessary.

In general, the absorption coefficient may be calculated for rotational transitions using Equation IV-22.

$$\varepsilon(\omega) = \frac{8\pi^3}{3hc}\omega\left(\frac{N_m}{g_m} - \frac{N_n}{g_n}\right)|\langle n|\mu|m\rangle|^2 S(\omega, \omega_0) \qquad \text{Equation IV-22}$$

This equation illustrates the dependence of the rotational transition on the change in the dipole moment of the molecule between the n (excited) and m (lower) states, represented by $|\langle n|\mu|m\rangle|$, referred to as the dipole moment matrix.[309] As with Equation IV-19, the variable N represents the population of each state, and g represents the degeneracy of each state. The frequency of the radiation is represented by the variable ω, the speed of light by c, Planck's constant by h, and the absorptivity by ε. The line shape function is represented by $S(\omega,\omega_0)$, which is a function of the frequency of the electromagnetic radiation as well as the energy of the transition in terms of frequency, represented by ω_0. This is often referred to as the "resonant frequency" of the transition, and represents the frequency at which the maximum absorption is observed for a given transition. Different line shapes will be described in greater detail. The absorptivity may also be determined as a function of the rotational quantum numbers, as described in the following paragraphs.

For linear molecules, the absorptivity of a given transition may be represented by Equation IV-23.

$$\varepsilon_{lin}(\omega_0) = \frac{8\pi^2}{3ck}\left[\frac{h}{k}\right] af_v \frac{N}{\Delta\omega_0}\frac{\omega_0^2 B}{T^2}\frac{g_I(J)}{G_I}\mu^2(J+1) \qquad \text{Equation IV-23}$$

Here, the absorptivity, $\varepsilon_{lin}(\omega_0)$ is determined at the resonant frequency, as this corresponds to the discrete energy spacing of the transitions described by the K quantum numbers. N represents the concentration of the absorbing species in units of molecules per cm^3. Isotopic abundance (as mentioned in Chapter I) is represented by the variable a, representing a fraction between 0 and 1. The fraction of molecules in the particular vibration state is represented by f_v, as the possible rotational transitions are specific to each vibrational state (also mentioned in Chapter I). The dipole moment is represented by μ, in units of Debye. G_I represents the degeneracy of states resulting from spin ($g_I(K)$), and $g_I(J)$ represents the degeneracy of the rotational state. The absorptivity is normalized by the peak width ($\Delta\omega_0$), as the intensity at the resonant peak is diminished by spectral broadening. B is the molecular constant, T is the temperature (in K), k is Boltzmann's constant, and J is the angular momentum quantum number.

Symmetric tops exhibit non-degenerate transitions of K quantum numbers, which complicates the calculation of the respective absorption coefficient, shown in Equation IV-24.

$$\varepsilon_{sym}(\omega_0) = \frac{8\pi^2}{3ck}\left[\frac{h^3}{k^3\pi}\right]^{1/2} af_v \frac{N}{\Delta\omega_0}\omega_0^2\frac{BA^{1/2}}{T^{5/2}}\frac{g_I(K)}{G_I}\mu^2\frac{(J+1)^2-K^2}{J+1} \qquad \text{Equation IV-24}$$

A is the molecular constant derived from the respective moment of inertia, and K is the internal figure axis quantum number. Various substitutions may be made to simplify Equation IV-24 as described by Kroto,[309] but it is provided in this form for comparison with Equation IV-23.

Absorptivities of asymmetric rotors may be determined using Equation IV-25.

$$\varepsilon_{asym}(\omega_0) = (3.8 \cdot 10^{-14}) a f_\nu \frac{\omega_0^2}{\Delta\omega_0^s} \frac{(ABC)^{1/2}}{T^{5/2}} \frac{g_I(JK_AK_B)}{G_I} \mu_\alpha^2 \lambda_\alpha(J'\tau', J''\tau'') \quad \text{Equation IV-25}$$

The constants at the start of this equation are condensed for simplicity. The molecular constant C is used in calculation of the absorptivity, as well as tabulated values for λ_α, values which were made available in a series of publications.[317-321] Comprehensive derivations of these equations can be found elsewhere.[309]

Spectral Broadening

Spectral absorption features are broadened by several contributors. The first and smallest of the contributors is what is known as the "natural line width". All rotational absorption peaks will contain broadening of this type, which is the result of finite lifetimes of the excited states.[13] Natural line widths may be estimated using Equation IV-26.

$$\Delta\nu \approx \frac{32\pi^3}{3hc^3} \nu_{mn}^3 |(m|\mu|n)|^2 \quad \text{Equation IV-26}$$

Line width (in Hz) is represented by the variable $\Delta\nu$, and the frequency of the transition by ν_{mn}. The dipole matrix element is represented by the variable (m|μ|n). Spread of energy for the state in question is also dependent on the number of states below the excited state to which the molecule may relax. In this case, the natural line width may be represented by a sum of dipole moments and frequencies for each relaxation event, as shown in Equation IV-27.[306, 308, 322]

$$\Delta\nu \approx \frac{32\pi^3}{3hc^3} \sum_i \nu_{im}^3 |(m|\mu|i)|^2 \quad \text{Equation IV-27}$$

The summation is performed across all i states to which the molecule might relax from state m. Natural line widths are expected to be on the order of 10^{-8}[13] to 10^{-4}[308, 322] Hz, widths which would require unprecedented spectral resolution to determine experimentally.

Another contributor to line broadening is a phenomenon known as Doppler broadening, which is typically observed under low pressures. This is observed as a result of most absorbing molecules having some component of their velocity which is parallel to the direction of propagation of the electromagnetic radiation produced by the instrument. For example, when a molecule is moving away from the source, the light reaching the molecule seems to have a lower frequency than that which is required to induce a rotational transition. This molecule will absorb a photon of slightly higher frequency than expected. Conversely, a molecule traveling toward the source will encounter light which seems to have a higher frequency than that necessary to induce the same transition. In this case, the molecule will absorb a photon of lower frequency than expected. The HWHM of the transition may be calculated using Equation IV-28.[13, 306, 308, 322]

$$\Delta\nu = \frac{\nu}{c}\left(\frac{2\mathrm{N}kT}{M}\ln 2\right)^{1/2} \qquad \text{Equation IV-28}$$

The HWHM ($\Delta\nu$) is proportional to the transition frequency ν. Spectral features are also broader at higher temperatures (T). In Equation IV-28, the variable c represents the speed of light, and k is the Boltzmann constant, N is Avogadro's number, and M is the molecular weight of the species. Under these conditions, the observed spectral peak will take the shape of a Gaussian distribution, with a HWHM around 40 kHz.[308]

Parsons and Roberts examined the contribution of Doppler broadening to microwave spectral features, and found that failure to account for Doppler broadening could produce errors in self-broadening parameters of as much as 1% and around 8% for calculated collision cross-sections obtained from the analysis of the spectrum of ammonia.[323] Gordy suggests a measurement geometry in which the sample vapor is sprayed at an angle which is orthogonal to the direction of microwave radiation propagation.[13] In doing so, the velocity components of the molecules under analysis

which are parallel to the beam path will be reduced, producing spectra with decreased or eliminated Doppler broadening.[13]

The last of the broadening phenomena most pertinent in the experiments described in this work is collisional broadening. This phenomenon includes three distinct subgroups, though not always described in this fashion in scientific literature. Weakest of the three is "wall collision broadening", which is observed as a result of the absorbing molecules colliding with the walls of the sample cell used for analyses at very low pressures (~1×10^{-6} atm[322]). The HWHM as a result of this phenomenon may be estimated using Equation IV-29.

$$\Delta\nu = \frac{A}{2\pi V}\left(\frac{RT}{2\pi M}\right)^{1/2} \qquad \text{Equation IV-29}$$

The surface area of the cell is represented by the variable A, while the volume is represented by V, and R is the gas constant. This phenomenon is expected to provide a spectral peak width of approximately 10 to 15 kHz.[308] Though samples analyzed in this work were not analyzed under such low pressures, this effect is still worth mentioning, as it falls in the same category as the most significant broadening effects.

The remaining two effects are known as self-broadening and foreign-gas broadening. Self-broadening is the process by which spectral features associated with a particular analyte are broadened by collisions between analyte molecules. Foreign-gas broadening is the process by which analyte spectral features are broadened by collision with other gas molecules, such as N_2, O_2, or CO_2 (as mentioned in Chapter I). Contribution of each effect to the width of observed spectral features is proportional to the pressure of each perturber and the broadening coefficient of each species, given in units of frequency per unit pressure (i.e. MHz/Torr). The theory behind each of these effects has been explained by Lorentz,[324] Debye,[325] and later by van Vleck and Weisskopf.[326] Lorentz's approach assumed a case in which the orientation of molecules after a collision had no relation to the pre-collision orientation.[306] Debye explained

molecular collisions in the case of polar molecules, which produced a theoretical model in which the post-collision orientations of polar molecules are not entirely random, but instead have an orientation related to the electric field at the time of collision. Re-orientation of the molecule with the electric field at the moment of impact causes absorption, even at points in the spectrum which are relatively distant from strong absorption peaks.[306] As a result, the absorption at a distance from the resonant frequency of a rotational transition will be related to collision frequency, or the mean time between collisions, which is often represented with τ. The effect of this parameter on spectral broadening is described by Equation IV-30.[13, 327]

$$\Delta\nu = \frac{1}{2\pi\tau} \qquad \text{Equation IV-30}$$

The width of the peak is inversely proportional to the time between collisions, τ. As the sample pressure increases, more molecules are present in the sample cell, which results in an increase in the number of collisions, and a corresponding decrease in the time between collisions, which causes an increase in spectral peak width.

Self-broadening and foreign-gas broadening can produce line widths anywhere above 25 kHz.[308] Studies of this phenomenon have been reviewed by Birnbaum.[328] Tsao and Curnutte have also provided an in-depth analysis of these phenomena.[329] Broadening coefficients will vary from one species to another, both in terms of self-broadening and foreign gas broadening. The source of this variability is a parameter known as the collisional cross-section, which is related to the mean time between collisions as shown in Equation IV-31.[327]

$$\tau = 1/n\bar{v}\sigma \qquad \text{Equation IV-31}$$

In Equation IV-31, the collision cross-section is represented by σ, n is the number of molecules per unit volume, and the relative velocity of impact is represented by $\bar{v}$. The cross-section is not strictly determined by the physical size of the molecule, but also by the intermolecular forces the molecule is able to exert on neighboring species.[327]

The effect of self-broadening may be described in this context with Equation IV-32.[110, 330]

$$\Delta\nu = \frac{n_1 \nu_1 b_1^2}{\sqrt{2}} \qquad \text{Equation IV-32}$$

In Equation IV-32, n_1 represents the molecular concentration (in molecules per unit volume) of the analyte molecule, ν_1 represents the mean velocity of the analyte molecule, and b_1 represents the collision diameter of the analyte molecule. Foreign-gas broadening is described by Equation IV-33.[110, 330]

$$\Delta\nu = \frac{\sqrt{2} n_1 \nu_1 b_1^2 + n_2 \nu_{12} (b_{12})^2}{2} \qquad \text{Equation IV-33}$$

Equation IV-33 includes the contribution of self-broadening to the observed line width, as indicated by the variables given the subscript 1. The variable n_2 represents the number of foreign-gas molecules per unit of volume, while ν_{12} and b_{12} represent the root-mean-square velocity of the foreign gas relative to the analyte molecule, and the collision diameter of the foreign gas, respectively. As indicated by Equations IV-32 and IV-33, the change in the observed spectral line width is proportional to the square of the collisional diameter, thus greatly increasing the spectral line width of analyte molecules when observed in the presence of bulky or polar molecules.

Spectral Peak Shapes

Three unique absorption profiles are thus suggested for modeling of rotational spectral features, given certain guidelines in the application of each line shape. As mentioned previously, when samples are analyzed at very low pressures (up to ~1 mTorr), a Doppler-broadened spectral peak is observed. This line is best represented by a Gaussian profile, shown in Equation IV-34.

$$A_G(\nu) = V\varepsilon_{\nu_0} \exp\left(\frac{[\nu - \nu_0]^2}{2\Delta\nu^2}\right) \qquad \text{Equation IV-34}$$

The magnitude of the absorption peak is dictated by the absorptivity $\varepsilon_{\nu 0}$ (in A.U./cm^3), with the maximum observed at the resonant frequency ν_0, and decreasing at adjacent frequencies ν. The HWHM of the absorption peak is represented by $\Delta\nu$, and the absorbance as measured in volume V (in cm^3) is calculated as $A_G(\nu)$. Under these conditions, the shape of the peak is representative of the velocity distribution of the absorbing molecules, which follows the Maxwell-Boltzmann distribution.

As the pressure is increased to a range from ~1 mTorr to 1 Torr, spectral absorption peaks will start to take on the shape of the Lorentzian profile as a result of collisions between molecules.[13] Gordy points out the underlying complexity of the sample under these conditions: not only is energy transferred between colliding molecules, but the presence of polar molecules (presumably the analyte, but not excluding interfering absorbing species) in greater quantities gives rise to longer-range interactions.[13] Stark modulation (mentioned in Chapter I) of analyte energy levels may be induced by other molecules in close proximity.[13] Molecules with particularly strong dipoles may experience attractive or repulsive forces due to dipole-dipole interactions. While the extent of broadening of spectral features of different polar molecules may provide insight into the magnitude of the dipole moment, Stark modulation of energy levels provides a more reliable experimental method for dipole determination.[13] The Lorentzian line profile, provided in Equation IV-35 is recommended for modeling of spectral features observed under conditions such as those above, primarily for spectral peaks with HWHM much less than the resonant frequency.

$$A_L(\nu) = V\varepsilon_{\nu_0}\Delta\nu\left[\frac{\Delta\nu}{(\nu_0-\nu)^2+\Delta\nu^2}\right] \qquad \text{Equation IV-35}$$

The variables in Equation IV-35 are identical in meaning to those of Equation IV-34. It is important to note that Equation IV-35 is of a slightly different form than that provided by Gordy.[13] The profile provided by Equation IV-35 includes the width parameter outside the bracketed term in order to normalize the profile. Excluding the

absorptivity term, this provides a peak in which the maximum value is always 1, making this form more practical for peak-fitting applications.

If the pressure of the analyte is increased further, to a level near 10 Torr or greater, spectral absorption peaks will start to deviate from the Lorentzian profile.[13] This line profile gives way to another line shape known as the van Vleck-Weisskopf profile, which bears some resemblance to the Lorentzian profile near the resonant frequency, but exhibits higher absorption at frequencies above this. The profile is described by Equation IV-36.

$$A_L(\nu) = V\varepsilon_{\nu_0}\frac{\nu}{\nu_0}\left[\frac{\Delta\nu}{(\nu_0-\nu)^2+(\Delta\nu)^2}+\frac{\Delta\nu}{(\nu_0+\nu)^2+(\Delta\nu)^2}\right] \qquad \text{Equation IV-36}$$

All variables in Equation IV-36 have the same meaning as those in Equations IV-34 and IV-35. Note the similarity between Equation IV-35 and the first term in the brackets in Equation IV-36. A comparison of these three line profiles is provided in Figures IV-4 and IV-5. Figure IV-5 illustrates the effect of increasing the peak width (Δν), mimicking the effect of spectral broadening on the appearance of spectral absorption features.

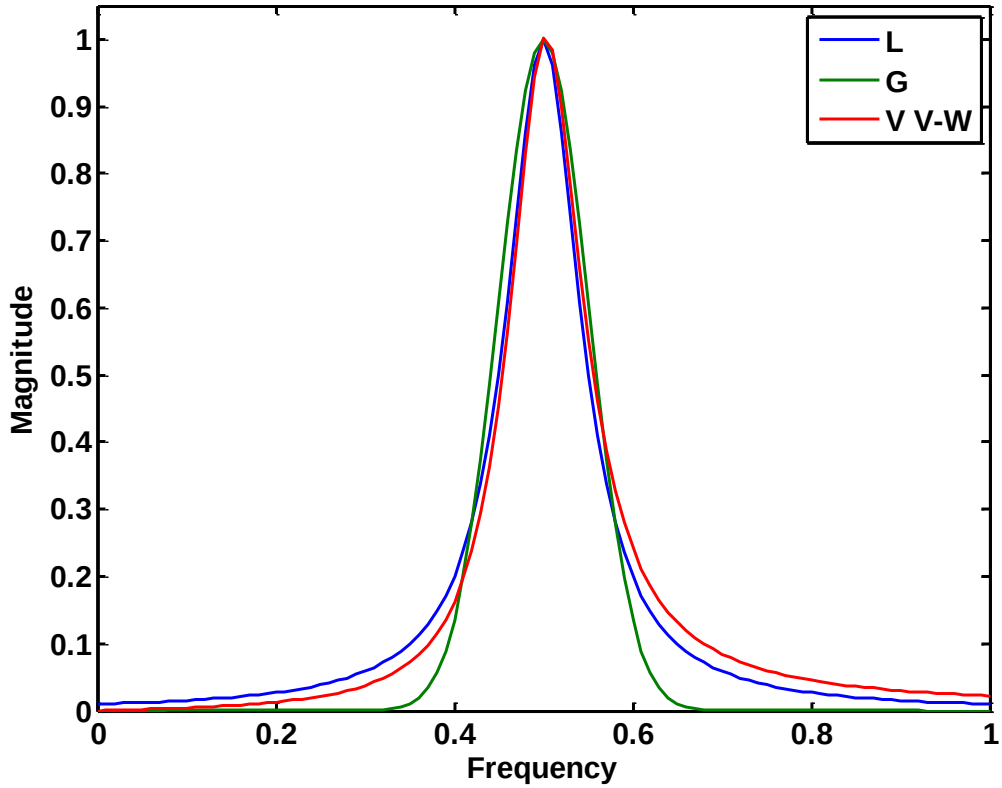


Figure IV-4. A comparison of the Lorentzian (L), Gaussian (G), and the van Vleck-Weisskopf (V V-W) line profiles. Note the difference between the Gaussian profile and others near the resonant frequency, where the Gaussian profile has a slightly increased width relative to the others. The Gaussian profile approaches zero much faster than the other profiles. Note also the difference in the magnitude between the Lorentzian and van Vleck-Weisskopf profiles at great distances from the resonant frequency. The Lorentzian profile has greater magnitude in the low-frequency tail, whereas the van Vleck-Weisskopf profile exhibits increased magnitude in the high-frequency tail. Figure IV-4 was produced using an x-axis point spacing of 0.01, peak center of 0.5, height of 1, and width of 0.05.

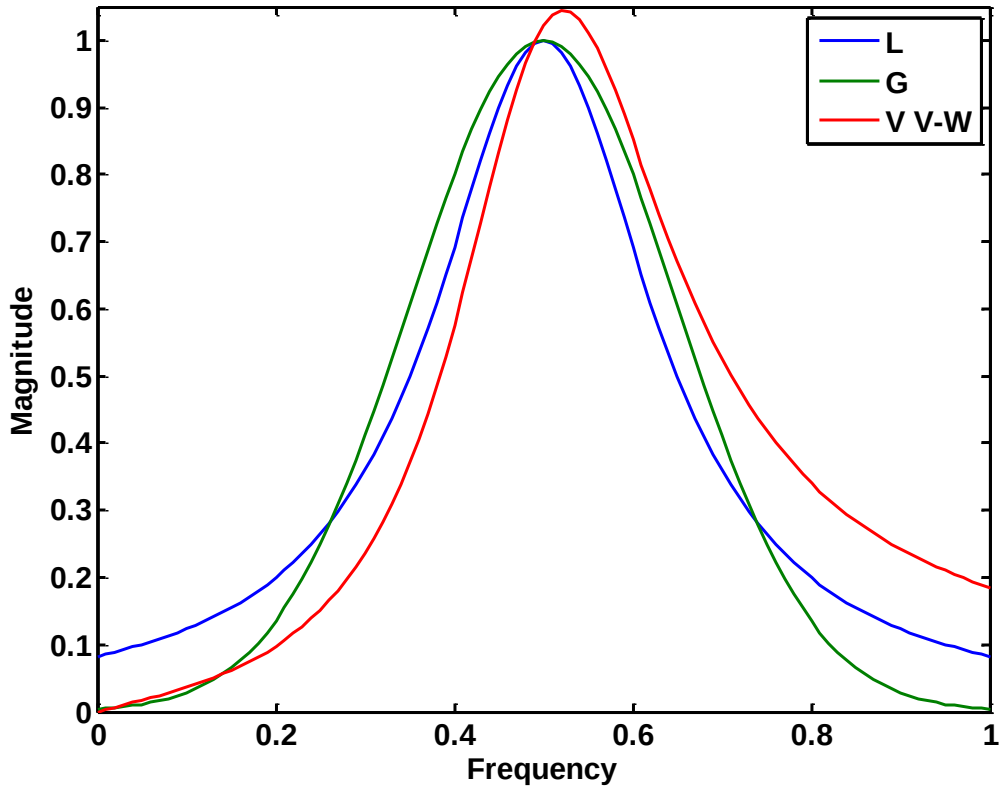


Figure IV-5. The three line profiles are plotted with a width of 0.15 and the same peak center and height as those used for Figure IV-4. Note the increased skew of the van Vleck-Weisskopf profile, similar to that illustrated by Gordy and Cook.[13] As the width increases, the height and apparent resonant frequency of the van Vleck-Weisskopf profile increase.

CHAPTER V

MOTIVATION FOR THZ FREQUENCY ANALYSIS OF GASEOUS SPECIES

Background

Precise measurement of gaseous species is of key importance in several fields: astronomical observation and study, atmospheric analysis, and industrial hygiene and health. As was mentioned briefly in some sources in Chapter I, microwave frequencies have been applied to the analysis of gaseous entities and the atmospheres of distant planets provides humankind with a deeper understanding of the Universe. Laboratory measurements are often carried out with various species to provide a reference set of spectral parameters for species of astrophysical interest. Armed with this knowledge, astronomers are then able to identify the species in spectral observations of extraterrestrial bodies.

The chemistry of the atmosphere is known to play a significant role in the Earth's climate. Atmospheric constituents, as well as pollutants exhibit unique absorptive and emissive properties, which affects the energy balance of the atmosphere.[331] Pollutants may also perturb natural chemical processes in the atmosphere, which in turn may affect the energy balance, or generate reaction products which can cause further harm to humans or the environment.[331] Furthermore, some species may reside in the atmosphere for long periods of time, increasing the long-term detrimental effects of the species.[331]

Industrial hygiene presents an opportunity for such measurements as well. Several of the compounds of interest are classified as hazardous air pollutants by the Occupational Safety and Health Administration (OSHA) and the Environmental Protection Agency (EPA). Permissible exposure levels (PELs) are set by OSHA which limit the concentrations to which workers may be exposed, based on an 8-hour time-weighted average.[332] Exposure to various species is restricted due to associated adverse

health effects, both short-term and long-term. Table V-1 provides a list of the species of interest, the respective chemical formula, CAS registry number, informal abbreviation, reference, and OSHA's PEL, if applicable.

Table V-1. Species of interest selected for THz-frequency analysis. Chemical names, formulas, informal abbreviations, CAS numbers, and PELs are provided with the respective references. Unregulated species are designated with "NA" in the PEL and Reference cells.

Species	Formula	CAS number	PEL	Reference
Acetaldehyde	CH_3CHO	75-07-0	200	333
Acetonitrile	CH_3CN	75-05-8	40	334
Ammonia	NH_3	7664-41-7	50	335
Ethyl Alcohol	CH_3CH_2OH	64-17-5	1000	336
Methyl Alcohol	CH_3OH	67-56-1	200	337
Water	H_2O	7732-18-5	NA	NA
Propionitrile	CH_3CH_2CN	107-12-0	NA	NA
Propionaldehyde	CH_3CH_2CHO	123-38-6	NA	NA

Alternative Measurement Configurations

A broad range of instrumentation is available for analysis of gaseous species in the atmosphere. Microwave-frequency instrumentation is currently in use in a variety of measurement configurations, including ground-based, airborne and satellite based units.[338] Infrared spectroscopic instrumentation has also been applied to measurements of atmospheric species in active and passive modes.[338] Active measurement configurations are those in which the analytical system incorporates both a source and detector, whereas passive instrumentation relies on an external source, such as the sun or infrared radiation

emitted by the Earth's surface.[338] Ultraviolet/visible (UV/Vis) spectroscopy has also been applied in active and passive modes for the measurement of minor stratospheric pollutants.[338] Other methods such as chemoluminescence,[339] gas chromatography,[339, 340] liquid chromatography,[339] laser-induced fluorescence[341] or breakdown,[342] mass spectrometry,[339] photoacoustic detection,[339] and many others. Many other techniques are available for such measurements. Each method has inherent advantages and disadvantages, which is likely the cause for such variety in the available apparatus. Merits and shortcomings of generalized classes of analytical instrumentation will be discussed to address areas in which THz-TDS measurement may be advantageous. These are primarily discussed with respect to the sensitivity, selectivity, and practical limitations of each method. Generic challenges met in most atmospheric measurements have been reported.[343]

Portable Chemical Sensors

Portable chemical sensors are becoming increasingly commonplace, as the electrical components for these devices become cheaper and more compact. Solid-state sensors also provide a durable measurement platform, as they are far simpler than laboratory bench-top instrumentation. This simplicity offers a benefit in the stability of the system: it will tend to be less sensitive to shock and vibration, and may be deployed in environments otherwise deemed unsuitable for an analytical instrument. Simplicity also seems to be the downfall of this class of apparatus, as they often exhibit low selectivity and sensitivity.[344] As mentioned in Chapter I, selectivity is the ability to distinguish between chemical species, and sensitivity is the ability to detect species which incite a miniscule instrumental response. Some devices require elevated temperatures for effective measurements, which may also make them unsuitable for certain applications.[344] Wilson points out a particularly challenging aspect of sensor development: selectivity and sensitivity are often traded off.[344]

Chromatographic Methods

Chromatographic apparatus often provides increased chemical selectivity owing to the separation process(es) preceding detection. Separation modules can also be combined with detection methods which further improve the selectivity of the method, such as a diode array for broadband spectroscopic detection or mass spectrometer for mass/charge-based detection. Selectivity and/or sensitivity are achieved at the expense of measurement time. Separation methods provide chemical selectivity via chromatographic resolution, which has a distinct lower limit. If the sample is passed through the separation component at too high a rate, separation will not be achieved.[345] As a result, measurements may require extended periods of time to achieve the desired resolution. Furthermore, remote deployment would require a re-stocking regimen to replenish carrier gases or solvents used in the separation. Chromatographic methods also require the use of some sort of inlet system which will transfer the sample from the atmosphere to the environment in which the measurement is performed. Reactivity of species with the inlet material is a potential concern when performing analyses with this apparatus, though some designs have circumvented this challenge.[343]

Spectroscopic (Narrowband)

Narrowband optical devices often employ lasers as sources due to the high radiant powers possible with modern units. Spectroscopy may be performed with these devices in either absorption or emission modes. Those apparatus which perform absorption measurements, such as cavity ringdown spectrometers, benefit from the high radiant power in that the laser may be propagated through a multipass cell many times without significant loss of the analytical signal. The long path lengths attainable with these devices provide excellent sensitivity, but this benefit is gained at the expense of selectivity in many cases. When the measurement is performed with a very narrow spectral range, spectral overlap from interfering species cannot be characterized, resulting

in potential error in predicted concentrations. In some cases, it is preferable for the apparatus to perform measurements in narrow bands as the measurement will otherwise take on "multiexponential" character.[346] This means that the measured decay signal will no longer be well fit by the single decay expression often used in this type of measurement.[346] Ball and Jones reviewed reports of (relatively) broadband cavity ringdown systems and their application to measurement of different atmospheric trace species.[346] Broader spectral output is often achieved by incorporation of dye gain media or xenon flash lamps in the apparatus,[346] but the useable bandwidth is still well below that of traditional spectroscopic methods. Species with widely-spaced absorption features are well-suited for measurements with this method, such as di- or triatomic species as well as simple polyatomic molecules.[272]

Laser-based emission measurements can provide enhanced selectivity by exciting species to a point at which they will emit broad spectral signatures. Further enhancements in sensitivity and selectivity may be gained by incorporating resonant pulses to maximize excitation of the analyte atom or molecule.[341] Laser-induced breakdown spectroscopy is a method which utilizes laser pulses to atomize molecules in the sample. These excited atoms then emit radiation at wavelengths specific to the atoms in the molecule. As such, the method provides spectral data representative of the analyte's elemental composition, though this information is not useful when analyzing compounds comprised of the same elements, such as those studied in the following experiments. Emission spectral features of the analyte may also be superimposed on emission features of other species in the sample matrix.[341] Spectral peak width can be reduced by analyzing samples at decreased pressures, providing some improvement in the selectivity of this method.

Spectroscopic (Broadband)

Broadband spectroscopic methods provide greater chemical selectivity relative to their narrowband counterparts, but there is some tradeoff in sensitivity since the radiant power produced by these sources is spread across a wide range of wavelengths. Even so, these devices may still be coupled to a multipass cell to increase sensitivity. Measurements have also been performed in open path configurations to achieve high sensitivity with relatively simple setups.[272] Successful measurement has been demonstrated for a wide range of analyte molecules.[343] Open path measurements provide the benefit of measuring the compound *in situ*, avoiding the problem of reactivity with instrument surfaces. Greater attention will be devoted to this category in a discussion of the place of THz-TDS amongst current spectroscopic methods.

Mass Spectrometry

Mass spectrometry is another popular method for atmospheric analyses. These devices first ionize incoming sample molecules and then separate them by their mass-to-charge (or m/z) ratios.[343] The selectivity of this method depends heavily on the instrument configuration: the ionization method and mass spectral resolution will determine the extent of molecular fragmentation and resolving power available for qualitative analysis, respectively.[343] Harsh ionization methods such as electron-impact ionization will fragment incoming species to a high degree, making qualitative analysis unreliable. Selectivity can be restored by coupling the device to a separation method such as a gas chromatograph or high-performance liquid chromatograph.[272]

Mass spectrometers utilizing multi-pole (quadrupole or octopole) separation regions can provide very high-resolution m/z spectra, but measurements may only be performed for one m/z ratio at a time.[343] As such, detection of multiple species requires m/z scans, lengthening the time required for sample analysis. Recent advances in the design of time-of-flight mass spectrometers have facilitated multi-component analysis, as

these instruments are capable of measuring all m/z ratios simultaneously. Time-of-flight mass spectrometers are starting to replace their multi-pole counterparts in atmospheric analyses.[347] A concise summary of instrumental configurations and ionization techniques was reported by Farmer and Jimenez.[343]

THz Spectral Features of the Compounds of Interest

Rotational spectral features have been identified for the species of interest, both theoretically and experimentally, as described in Chapter I. Many of these transitions fall within the spectral range accessible by the TPS Spectra 1000D. It is worthwhile to compare features observed in experiments to those reported in the literature to ensure that features assigned to each species are genuine and are not the result of instrumental artifacts. Rotational spectral peak positions and relative peak heights are available on the National Aeronautics and Space Administration's (NASA) Jet Propulsion Laboratories (JPL) molecular spectroscopy site.[348] Revisions to experimental and calculated peak parameters are reported frequently in the Journal of Quantitative Spectroscopy and Radiative Transfer.[349]

The following figures provide a comparison between the experimental absorptivity spectra (A.U./ppm) measured in this work, shown as solid traces, and the JPL reference parameters, shown as vertical bars and scaled for comparison. While the JPL database provides spectral parameters for many of the species of interest, no entries were found for propionaldehyde or for isotopologues of some other species of interest. The Cologne Database for Molecular Spectroscopy[350] and HITRAN[351] databases were not used to fill the gaps.

Table V-2 lists the isotopologues and excited vibrational states available for each species of interest. Vibrational states are denoted by the symbol ν. Each species is given a species identifier number which indicates the molecular weight of the species (the first two digits) and the sequence of its listing (the last three digits). Trans and gauche

designations are provided for singly-deuterated propionitrile, and a unique parameter set is available for each conformation.[349] Figures V-1 through V-21 contain overlays of the experimental spectra and available JPL reference parameters for 7 of the 8 species of interest. Isotopologues and rotational transitions of vibrationally-excited species are included in the spectra with scales appropriate for respective relative abundances. No baseline is plotted in these spectra. In some cases, such as the spectrum of water vapor, it may appear as though the reference parameters form a baseline near zero, but these are simply dense collections of weak rotational transitions. It is this collection of weak transitions that has been cited as the source of the water vapor absorption continuum, as described in Chapter I.

Absorptivity spectra shown in Figures V-1 through V-21 represent the average of ≥30 samples analyzed at concentrations ranging from 26 to 710 ppm. Spectra were collected at a resolution of 0.225 cm^{-1} as the average of 1800 scans. In each spectrum, the solid blue line represents the experimental absorptivity spectrum with 16×N levels of TDS extension, providing clearer peak shapes with a greater number of data points in the resulting spectrum. Rotational transitions contained in the JPL spectral database are represented by red vertical lines, each of which is illustrated with a small red dot at half-height to facilitate visual discrimination of closely-spaced transitions. Rotational transitions end abruptly for some species towards the high-frequency end of the spectrum, while significant spectral features may be observed in the experimental data at these frequencies. Parameters may not have been determined for these frequencies, or the discrepant experimental data represents rotational transitions of vibrationally-excited species, isotopologues, or other spectral contributors.

The Case for THz-Frequency Spectroscopic Analysis

Spectroscopic analysis of gaseous pollutants may be performed with a variety of analytical methods, as described previously, but THz frequency radiation has several

Table V-2. JPL rotational transition parameter sets and species identifiers for each compound of interest.

Species	Formula	Species Identifier
Acetaldehyde	CH_3CHO	44003
Acetonitrile	CH_3CN	41001
Acetonitrile	CH_3CN - ν_8	41010
Acetonitrile	$^{13}CH_3{}^{13}CN$	43005
Acetonitrile	$^{13}CH_3CN$	42006
Acetonitrile	$CH_3{}^{13}CN$	42007
Acetonitrile	CH_2DCN	42008
Acetonitrile	$CH_3C^{15}N$	42001
Ammonia	NH_3	17002
Ammonia	NH_3 - ν_2	17004
Ammonia	NH_2D	18004
Ammonia	$^{15}NH_3$	18002
Ethanol	CH_3CH_2OH	46004
Methanol	CH_3OH	32003
Water	H_2O	18003
Water	H_2O - ν_1 ν_2 ν_3 $2\nu_2$	18005
Water	HDO	19002
Water	D_2O	20001
Water	$H_2{}^{17}O$	19003
Water	$H_2{}^{18}O$	20003
Water	$HD^{18}O$	21001
Propionitrile	CH_3CH_2CN	55001
Propionitrile	$^{13}CH_3CH_2CN$	56003
Propionitrile	$CH_3{}^{13}CH_2CN$	56002
Propionitrile	$CH_3CH_2{}^{13}CN$	56001
Propionitrile	CH_2DCH_2CN-trans	56005
Propionitrile	CH_2DCH_2CN-gauche	56006
Propionitrile	$CH_3CH_2C^{15}N$	56004

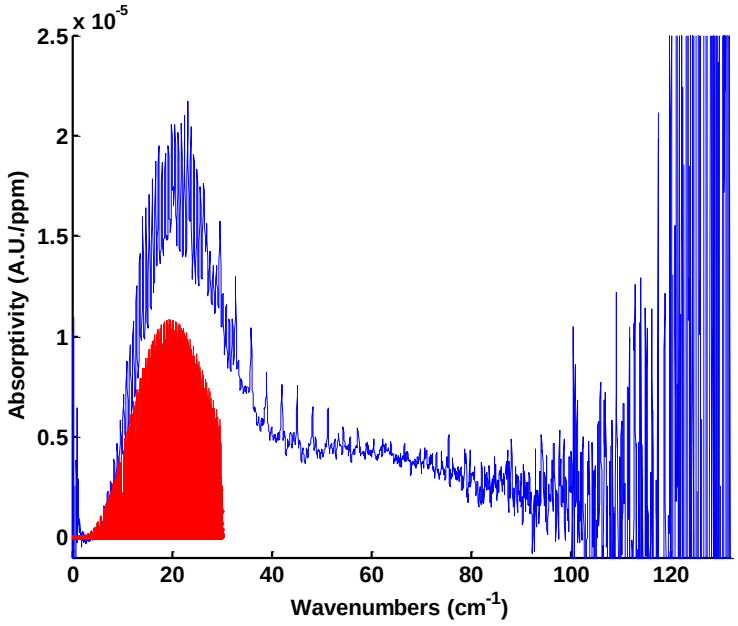


Figure V-1. THz-frequency absorptivity spectrum of acetaldehyde (blue) and corresponding JPL parameters (red). Note the abrupt end of parameters near 30 cm^{-1}.

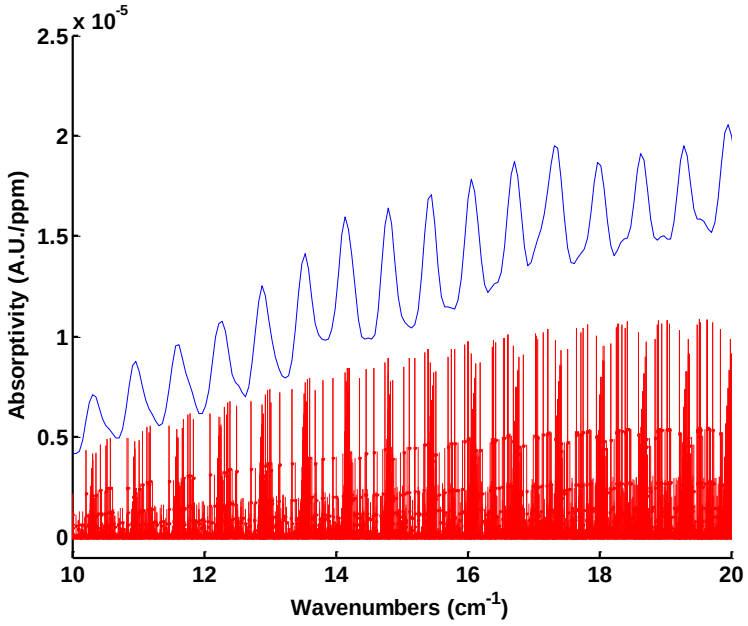


Figure V-2. Expanded view of the absorptivity spectrum (blue) shown in Figure V-1 with corresponding JPL parameters (red). Note the complexity of the underlying features which comprise the THz spectrum.

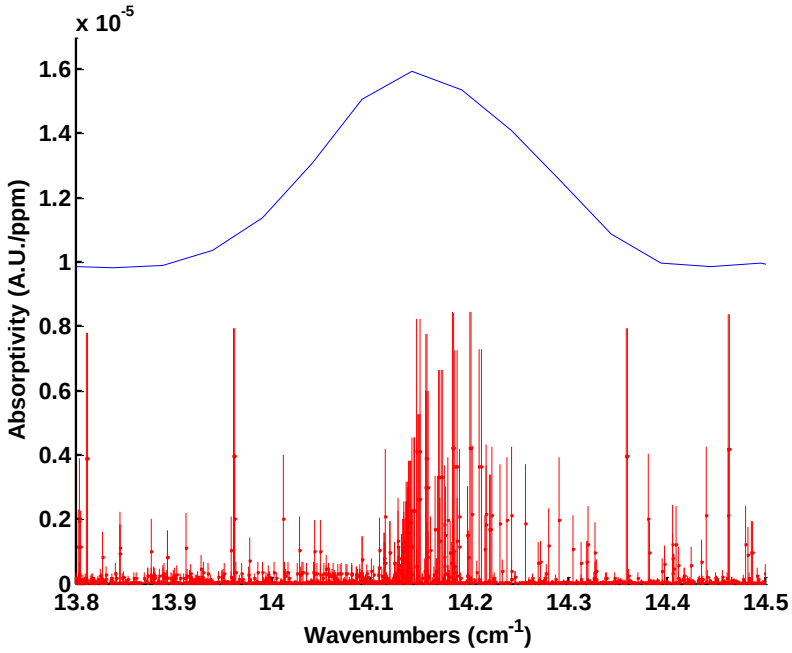


Figure V-3. Expanded view of the absorptivity spectrum (blue) shown in Figure V-2 with the corresponding JPL parameters (red).

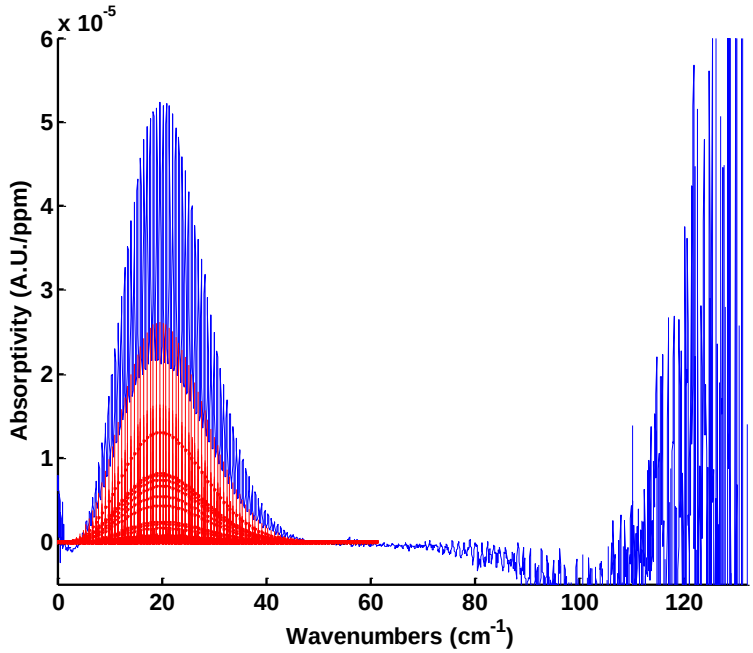


Figure V-4. THz-frequency absorptivity spectrum of acetonitrile (blue) and corresponding JPL parameters (red).

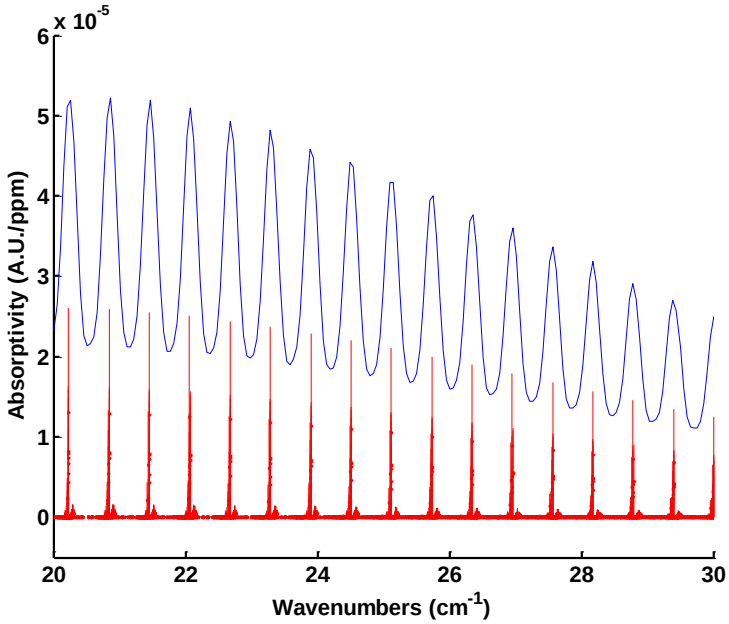


Figure V-5. Expanded view of the absorptivity spectrum (blue) shown in Figure V-4 with the corresponding JPL reference parameters (red). Note the complexity of the underlying features which comprise the THz spectrum.

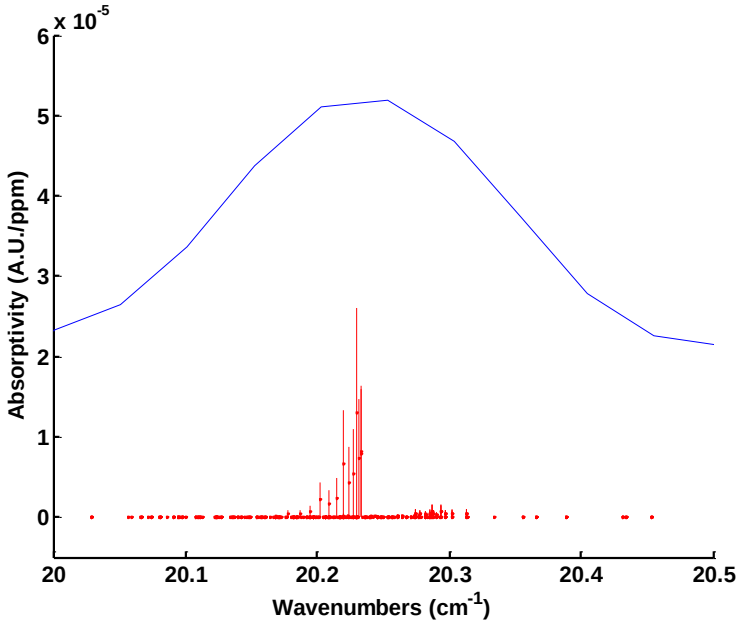


Figure V-6. Expanded view of the absorptivity spectrum (blue) shown in Figure V-5 with the corresponding JPL parameters (red).

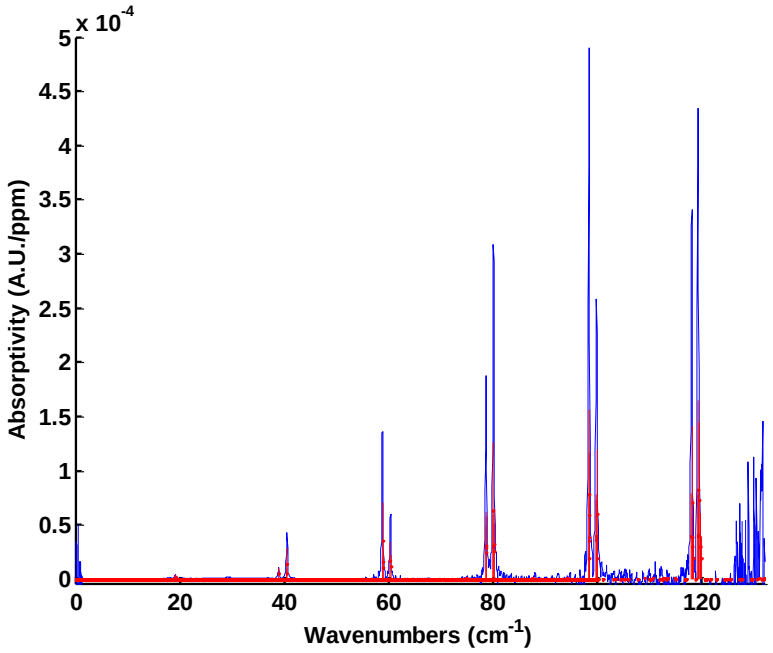


Figure V-7. THz-frequency absorptivity spectrum of ammonia (blue) and corresponding JPL parameters (red).

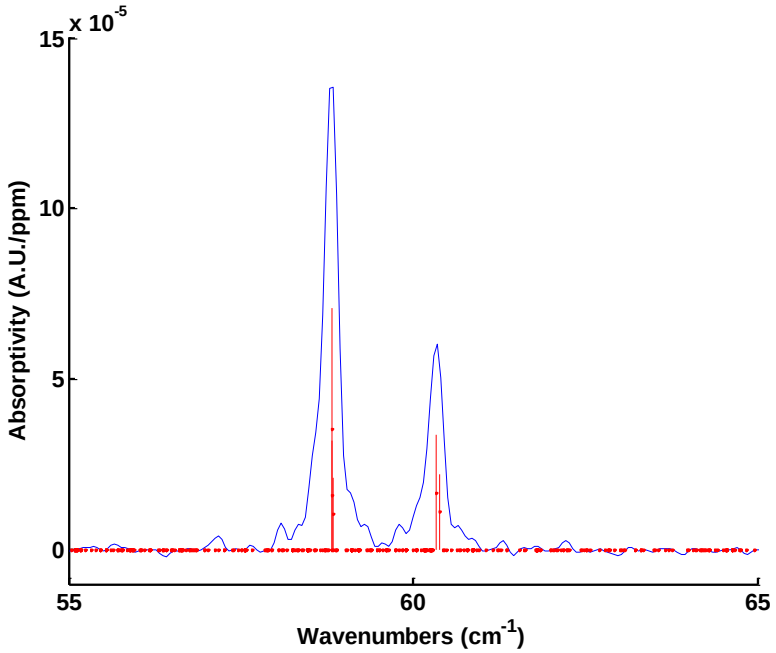


Figure V-8. Expanded view of the absorptivity spectrum (blue) shown in Figure V-7 with the corresponding JPL parameters (red).

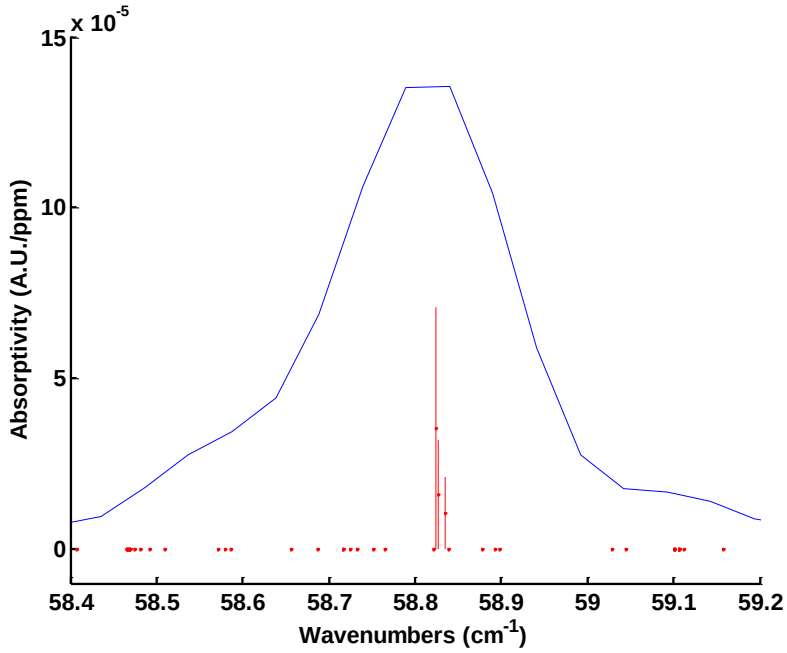


Figure V-9. Expanded view of the absorptivity spectrum (blue) shown in Figure V-8 and the corresponding JPL parameters (red).

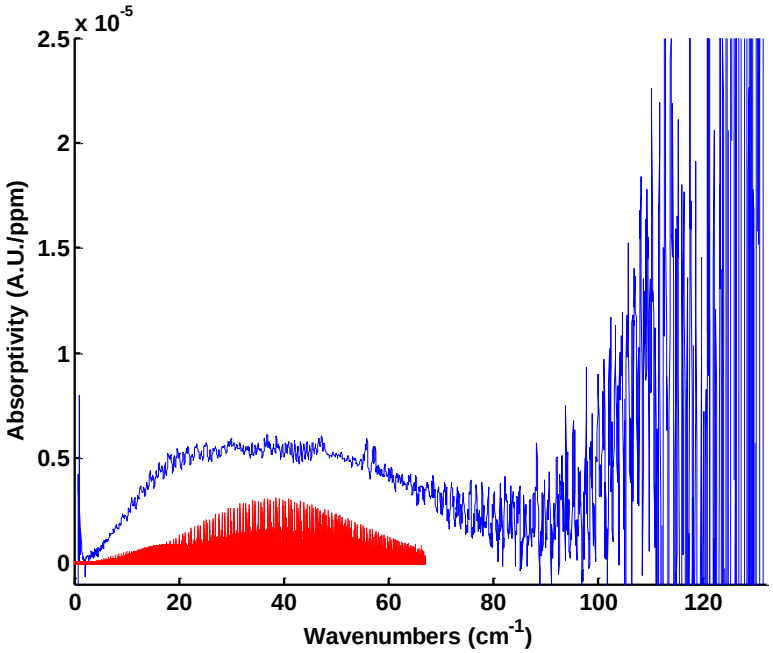


Figure V-10. THz-frequency absorptivity spectrum of ethanol (blue) and corresponding JPL parameters (red). Note the abrupt end to features near 66 cm^{-1}.

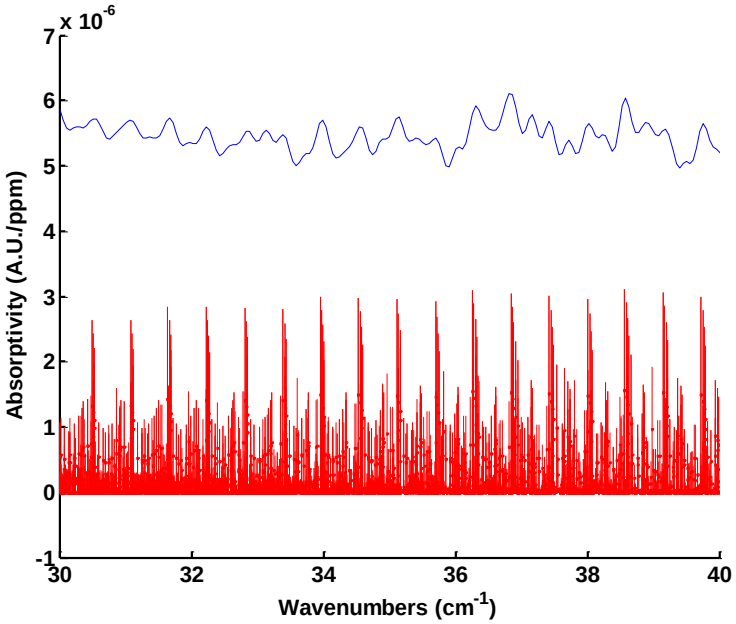


Figure V-11. Expanded view of the absorptivity spectrum (blue) shown in Figure V-10 with the corresponding JPL parameters (red). Note the complexity of the underlying features which comprise the THz spectrum.

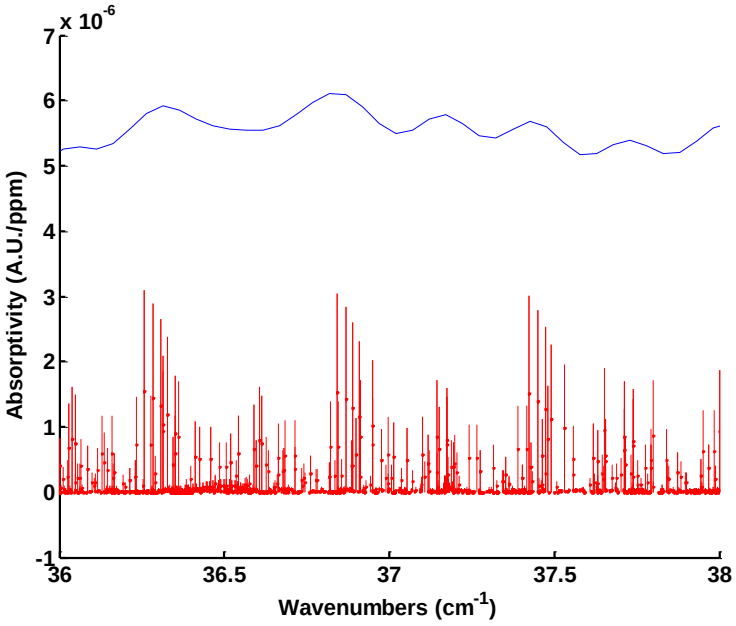


Figure V-12. Expanded view of the absorptivity spectrum (blue) shown in Figure V-11 and the corresponding JPL parameters (red).

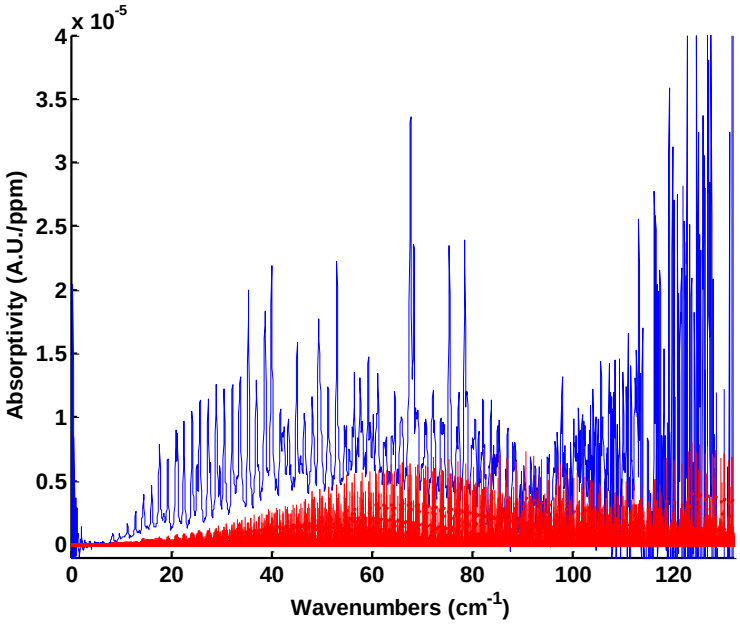


Figure V-13. THz-frequency absorptivity spectrum of methanol (blue) and corresponding JPL parameters (red).

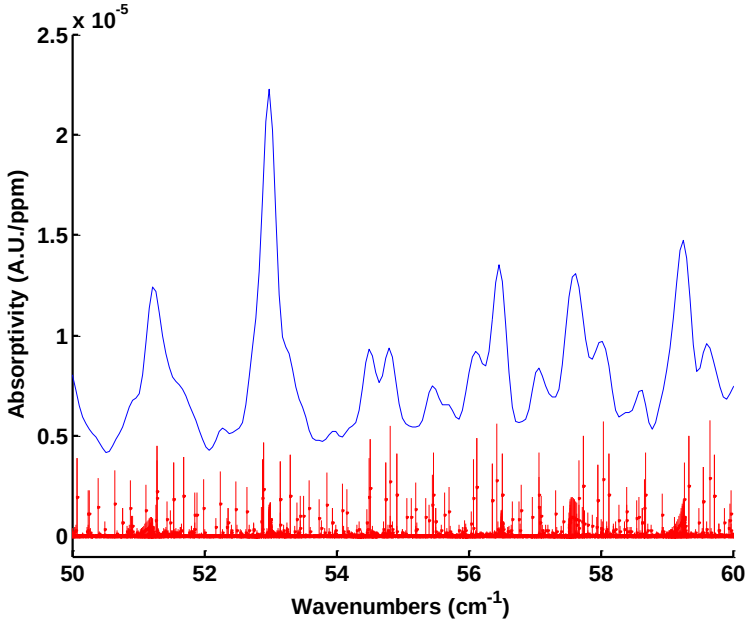


Figure V-14. Expanded view of the absorptivity spectrum (blue) shown in Figure V-13 and the corresponding JPL parameters (red). Note the complexity of the underlying features which comprise the THz spectrum.

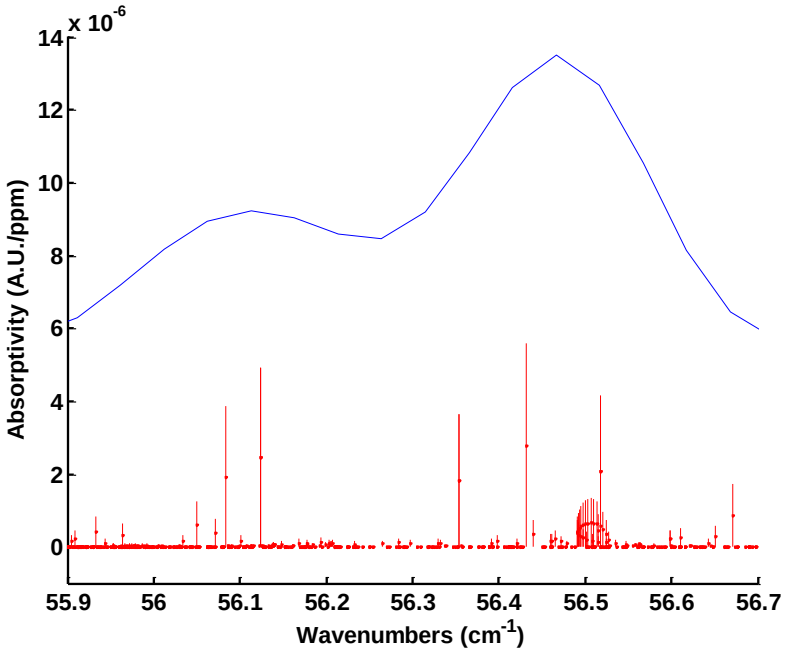


Figure V-15. Expanded view of the absorptivity spectrum (blue) shown in Figure V-14 and corresponding JPL parameters (red).

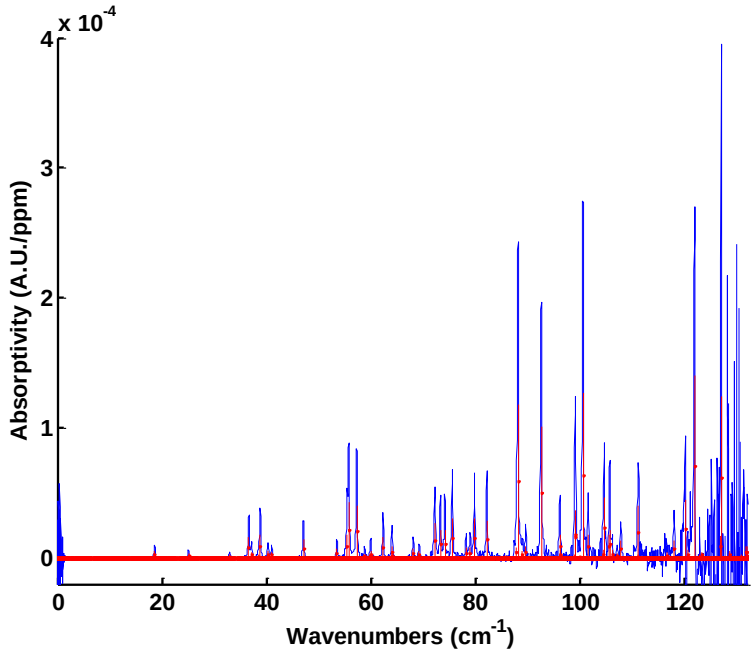


Figure V-16. THz-frequency absorptivity spectrum of water vapor (blue) and corresponding JPL parameters (red).

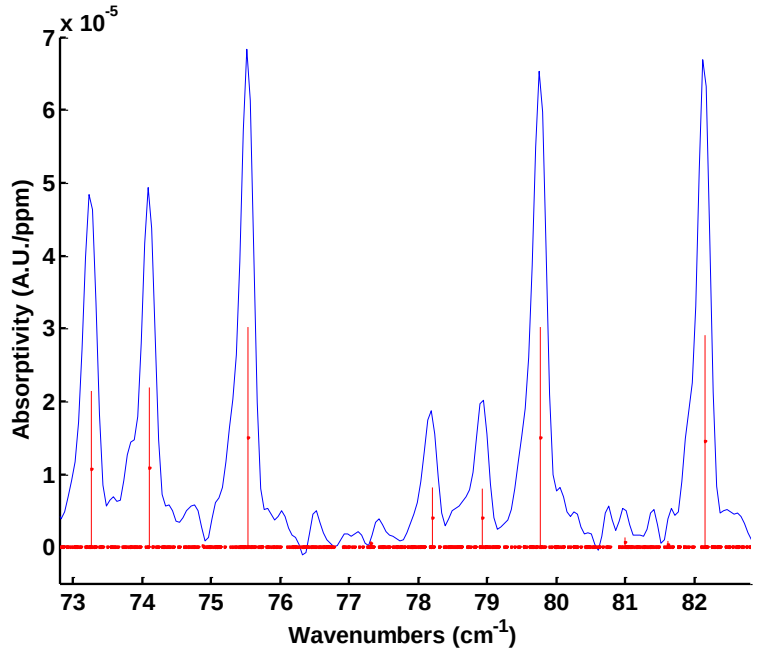


Figure V-17. Expanded view of the absorptivity spectrum (blue) shown in Figure V-16 with the corresponding JPL parameters (red).

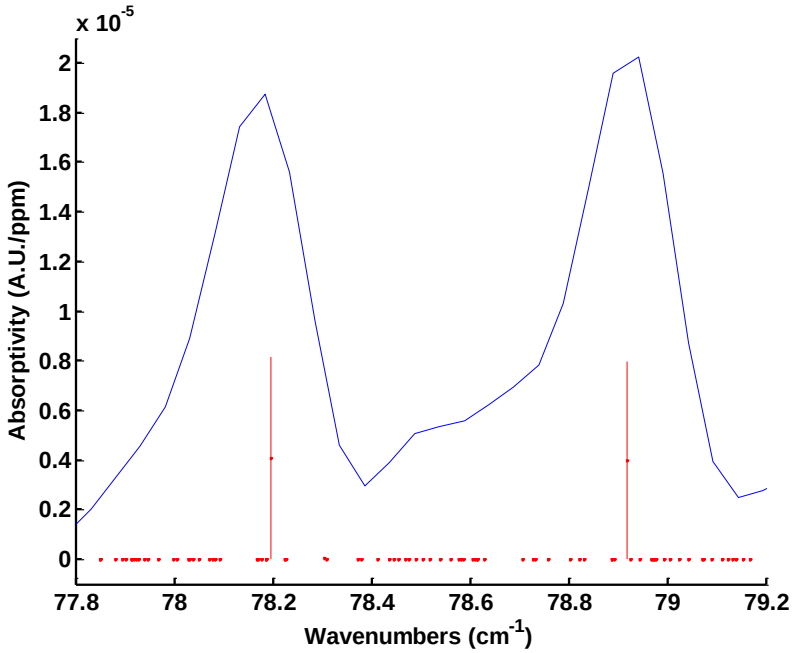


Figure V-18. Expanded view of the absorptivity spectrum (blue) shown in Figure V-17 and the corresponding JPL parameters (red).

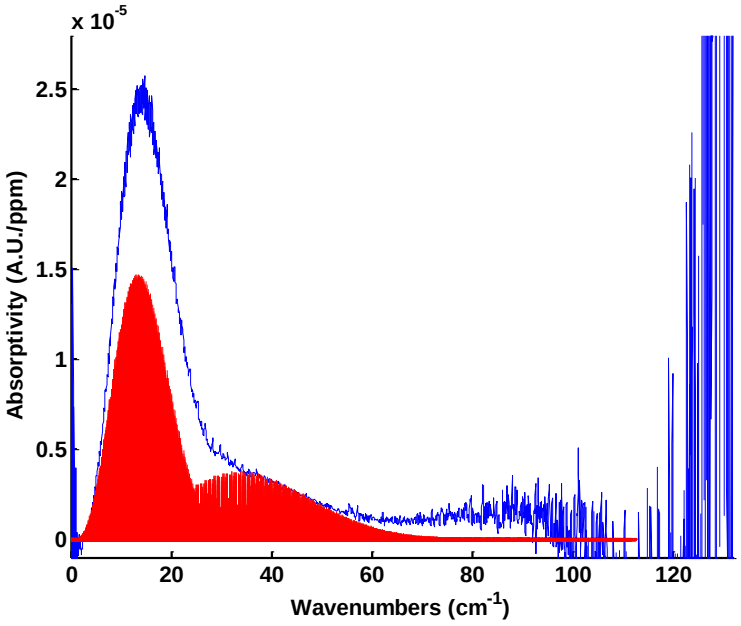


Figure V-19. THz-frequency absorptivity spectrum of propionitrile (blue) and corresponding JPL parameters (red).

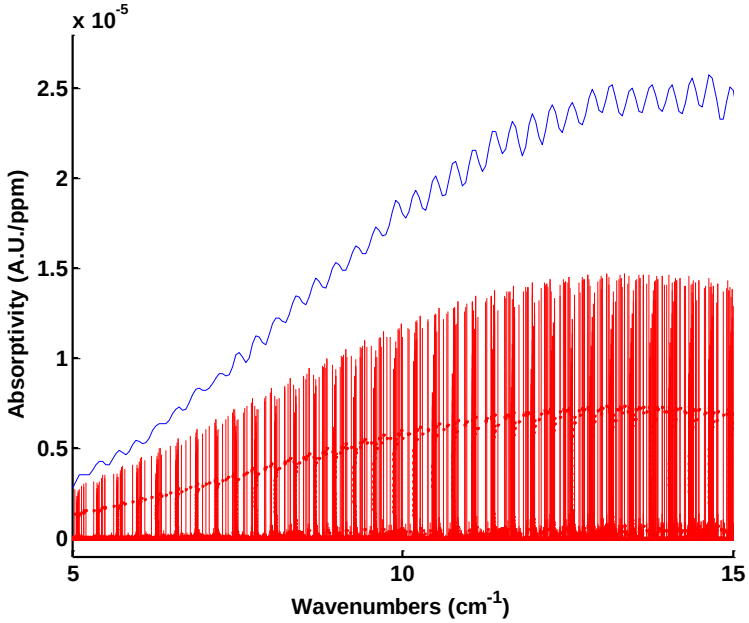


Figure V-20. Expanded view of the absorptivity spectrum (blue) shown in Figure V-19 and the corresponding JPL parameters (red). Note the complexity of the underlying features which comprise the THz spectrum.

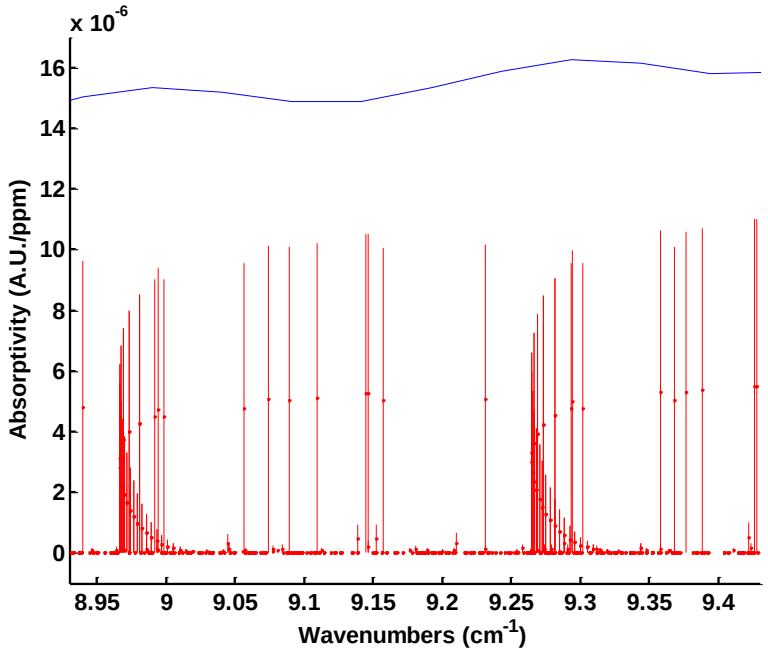


Figure V-21. Expanded view of the absorptivity spectrum (blue) shown in Figure V-20 and the corresponding JPL parameters (red).

properties which set it apart from other regions of the electromagnetic spectrum.

As mentioned in Chapter I, thermally-excited chemical species do not emit electromagnetic radiation near the sampled THz-frequency range.[179] Thus, species may be analyzed at various temperatures without observation of a combination of absorption and emission processes. As described in Chapter IV, the temperature does affect the height of observed absorption peaks as a result of a shift in the Boltzmann distribution, but this phenomenon affects all spectroscopic measurements.

A second argument may be made for this method based on selectivity. As mentioned in Chapter I and described theoretically in greater detail in Chapter IV, THz-frequency spectral features are comprised in part of pure rotational transitions of the analyte molecule. Because absorption is a function of the motion of the entire molecule, these frequencies should provide greater chemical selectivity relative to functional group analysis in infrared or Raman methods. This argument may be supported by analysis of

absorption spectra of these species in different spectral ranges. Library spectra of the species of interest were obtained from the Pacific Northwest National Laboratory (PNNL) as collected with FTIR instrumentation with path lengths ranging from 19.94 – 815.76 cm.[352] Overlays of the THz and IR spectra for the species of interest are shown in Figures V-22 through V-23 and Figures V-24 through V-29, respectively.

These spectra cover the 600-6500 cm^{-1} spectral range, and are provided with a data point spacing of 0.06 cm^{-1}. Chemical selectivity available with THz frequencies and IR frequencies was estimated using the NAS multivariate selectivity measures described in Chapters I (Equation I-5, shown in Table V-3) and III (Equation III-27, shown in Table V-4). Data point spacing in the PNNL library spectra was increased to 0.96 cm^{-1} to enable calculation of the selectivity coefficient. The SEL coefficient was calculated for various IR spectral regions with both SEL measures, as shown in Tables V-3 and V-4. For each species, the background matrix (R_{-k}) was constructed with the spectra of the remaining 7 species of interest. No spectra were included for either spectral range which described instrumental variance, as none were available for the PNNL spectral data. The 2000 – 2600 cm^{-1} spectral range was excluded from the SEL comparison due to the miniscule magnitude of the spectral features therein.

The chemical selectivity available with rotational spectral frequencies is evident from the comparison of the selectivity values listed in Tables V-3 and V-4. Similar trends are seen in the SEL values obtained with each method, though values calculated with the projection angle method tend to be lower in magnitude. Infrared absorption features associated with acetonitrile and water vapor provide greater multivariate selectivity than those in the THz spectral range, but the THz spectral features of the remaining species provide the greatest chemical selectivity, and THz SEL values for water and acetonitrile are similar in magnitude to those calculated for the IR spectral ranges.

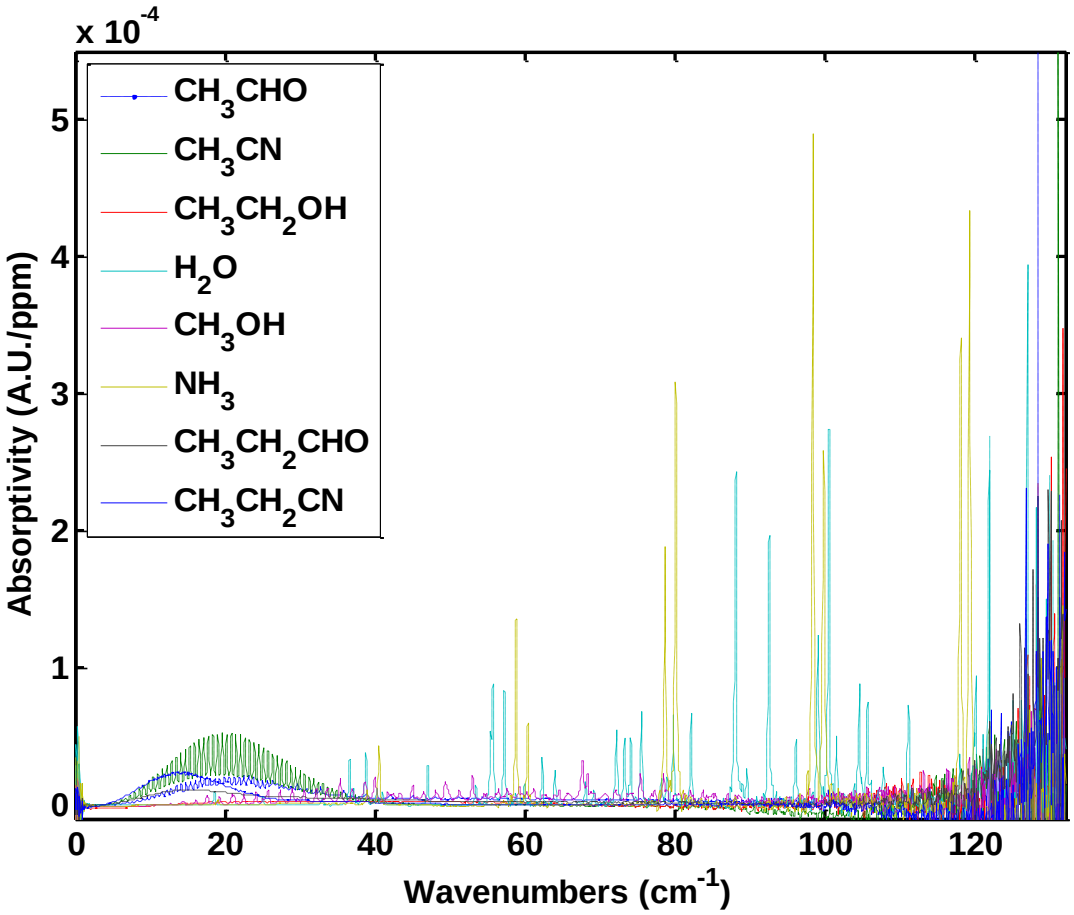


Figure V-22. THz absorptivity spectra of the compounds of interest.

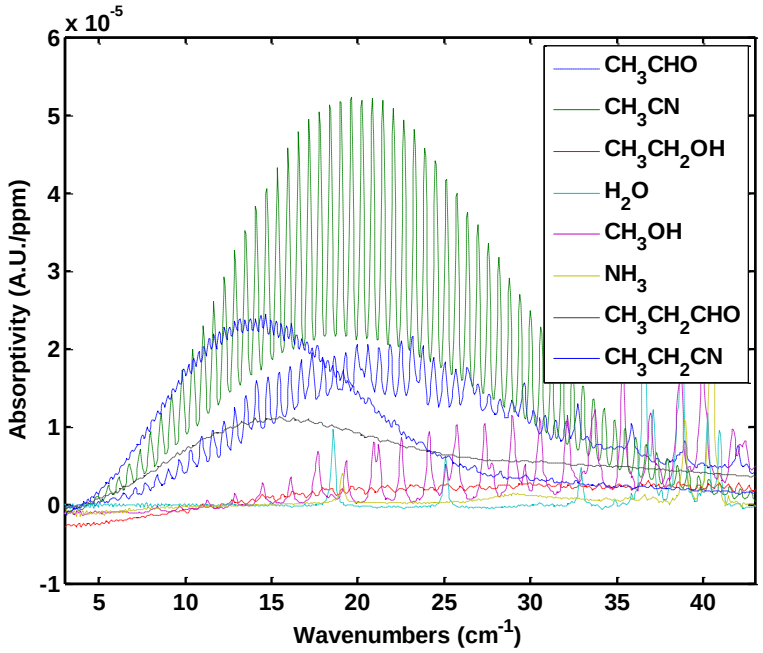


Figure V-23. Expanded view of Figure V-22 highlighting spectral features of rotors with increased molecular weight.

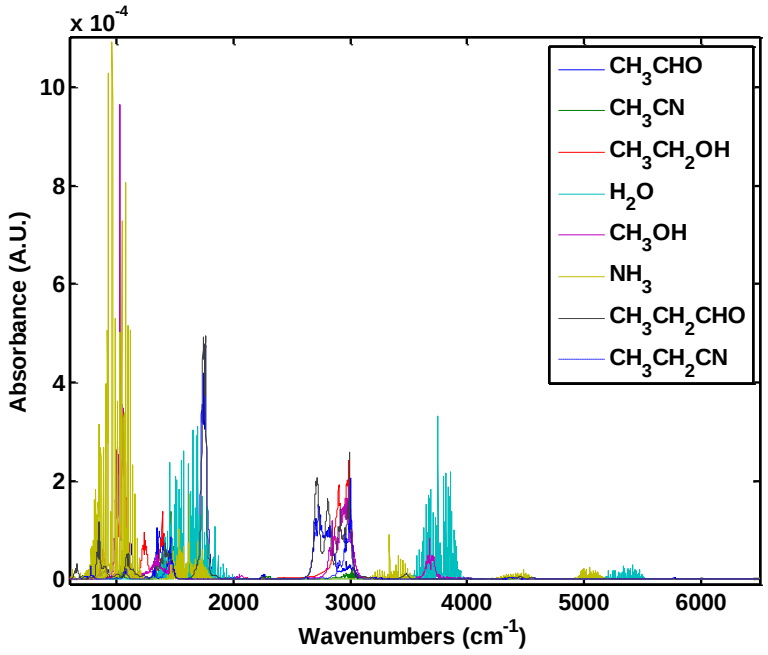


Figure V-24. IR absorption spectra of the compounds of interest. Note the decrease in absorption with increasing frequency.

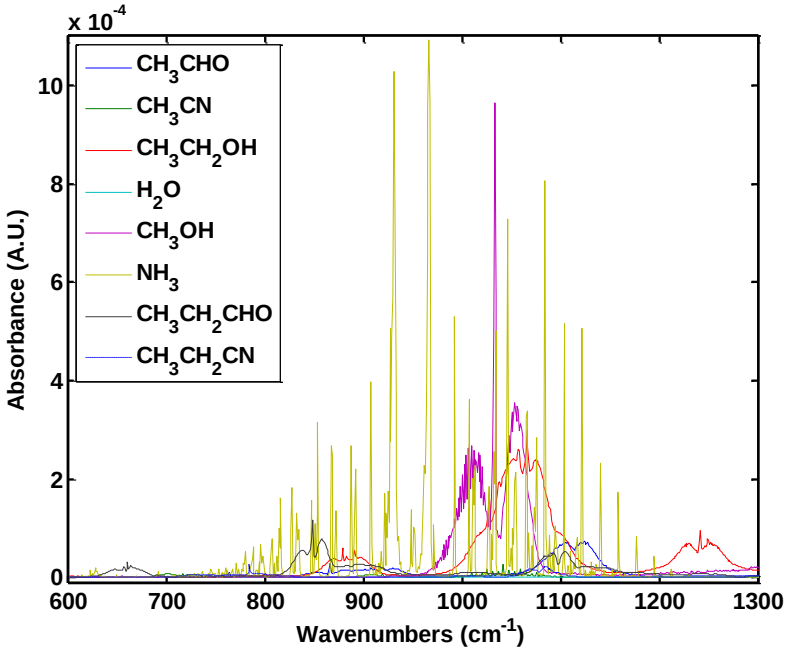


Figure V-25. IR absorption spectra of the compounds of interest over the 600 – 1300 cm^{-1} range.

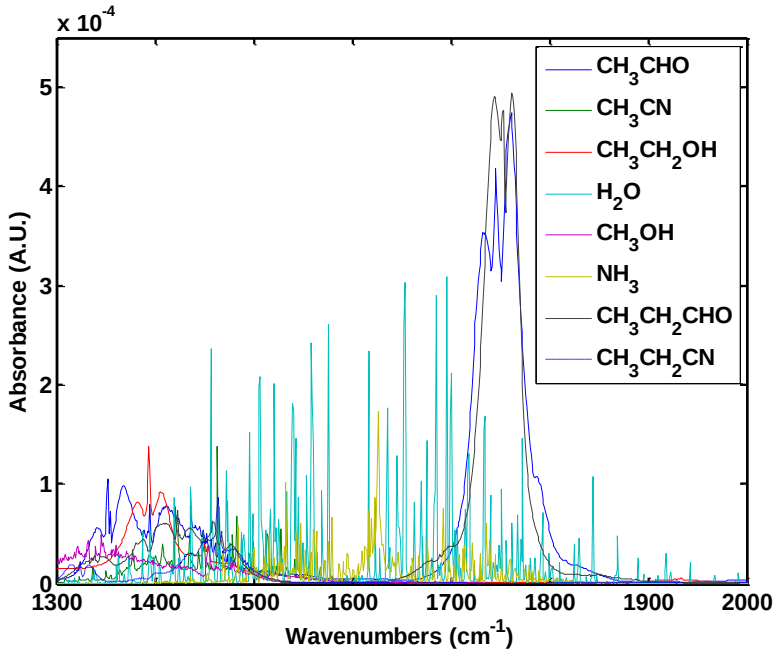


Figure V-26. IR absorption spectra of the compounds of interest over the 1300 – 2000 cm^{-1} range.

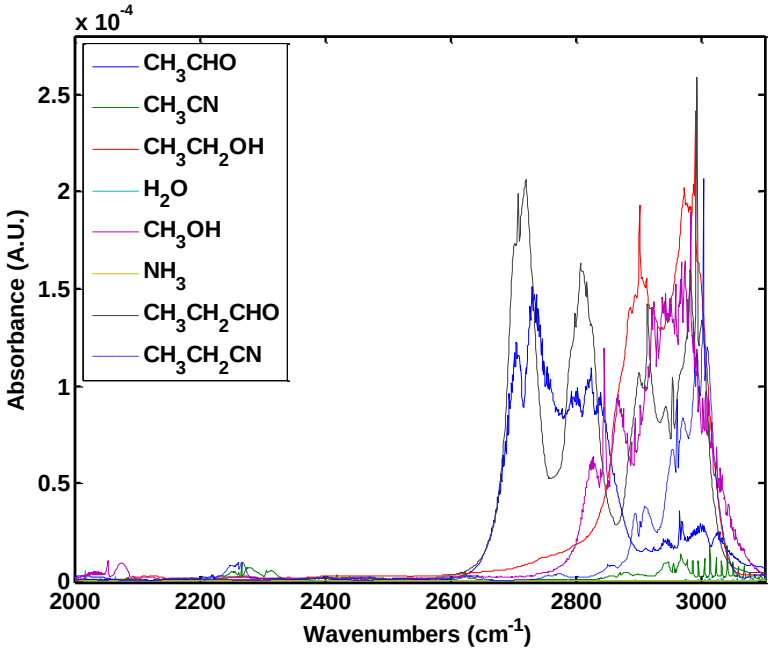


Figure V-27. IR absorption spectra of the compounds of interest over the 2000 – 3100 cm^{-1} range.

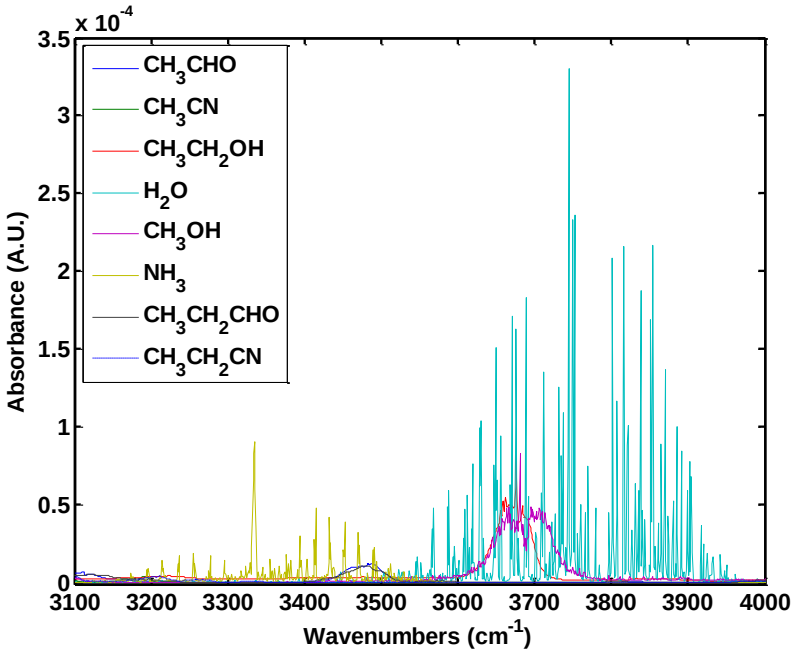


Figure V-28. IR absorption spectra of the compounds of interest over the 3100 – 4000 cm^{-1} range.

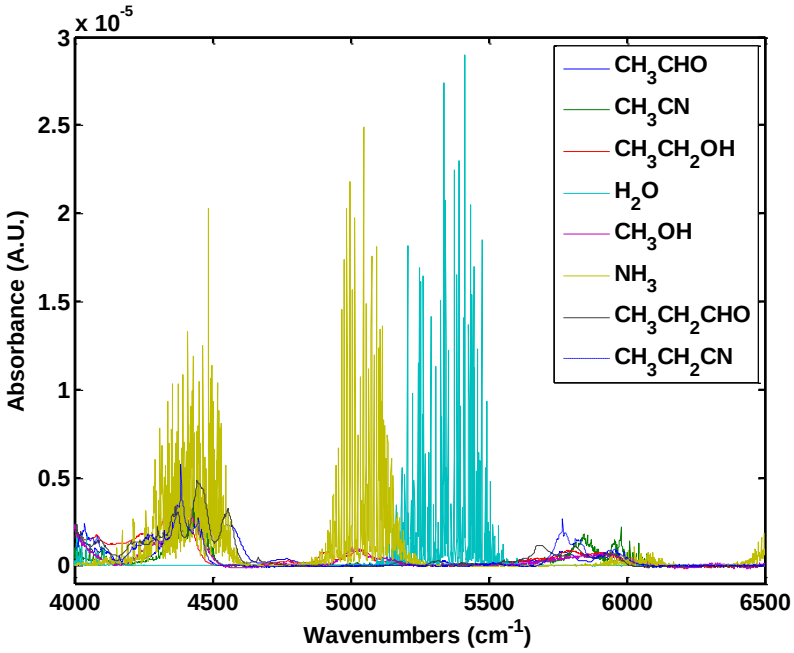


Figure V-29. IR absorption spectra of the compounds of interest over the 4000 – 6500 cm^{-1} range.

Table V-3. The SEL coefficient calculated from the ratio of the lengths of the NAS and analyte spectrum vectors. The highest SEL value for each species is highlighted in yellow.

Species	Spectral Region (cm^{-1})					
	THz	Mid-IR				Near-IR
	2-125	600-1300	1300-2000	2600-3100	3100-4000	4000-6500
CH_3CHO	0.723	0.714	0.259	0.432	0.205	0.343
CH_3CN	0.695	0.759	0.586	0.540	0.700	0.460
CH_3CH_2OH	0.968	0.512	0.574	0.251	0.533	0.398
H_2O	0.950	0.995	0.940	0.900	0.955	0.999
CH_3OH	0.869	0.662	0.583	0.272	0.542	0.340
NH_3	0.970	0.940	0.934	0.920	0.961	0.797
CH_3CH_2CHO	0.759	0.735	0.264	0.344	0.198	0.301
CH_3CH_2CN	0.779	0.722	0.632	0.510	0.715	0.379

Table V-4. The SEL coefficient calculated via the angle between vectors. The highest SEL value for each species is highlighted in yellow.

Species	Spectral Region (cm^{-1})					
	THz	Mid-IR				Near-IR
	2-125	600-1300	1300-2000	2600-3100	3100-4000	4000-6500
CH_3CHO	0.661	0.655	0.256	0.418	0.203	0.337
CH_3CN	0.641	0.688	0.553	0.515	0.644	0.444
CH_3CH_2OH	0.823	0.490	0.543	0.248	0.508	0.387
H_2O	0.813	0.839	0.808	0.783	0.816	0.841
CH_3OH	0.764	0.614	0.550	0.269	0.515	0.333
NH_3	0.824	0.807	0.804	0.795	0.820	0.716
CH_3CH_2CHO	0.688	0.671	0.261	0.338	0.197	0.297
CH_3CH_2CN	0.703	0.661	0.591	0.488	0.656	0.370

CHAPTER VI

QUANTITATIVE ANALYSIS OF INDIVIDUAL GASEOUS SPECIES

Experimental Configuration

Gas samples were prepared and analyzed using a variety of components shown schematically in Figure VI-1. Liquids were injected into the mixing manifold (3) via the liquid injection port (13), which relies on evaporation of the compound in the sample loop. As such, the system is most useful for compounds which have appreciable vapor pressures. Sample vapors flow from the mixing manifold (3) to the stainless steel gas cell (8) via the stainless steel transfer tube (7). THz spectral analysis was performed with the TPS Spectra 1000D (9), which was purged with dried air (15). The temperature of the gas(es) in the cell was (were) measured by a thermocouple (Omega model TJC36, 10) inserted in the cell, which provided a reading on the temperature readout unit (Omega model DP 462, 11). Transmission of THz pulses through the cell was made possible by two Teflon windows. The path length of the assembled cell was 7.780 ± 0.004 cm. Pressure was measured by a heated capacitance manometer (MKS Instruments, model 627D, 5), which displayed a reading on the gauge readout unit (MKS Instruments, model 660B, 6). Ammonia (Specialty Gases of America, 99.99% anhydrous, 12) and nitrogen (Airgas, ultra-pure carrier grade, 14) gases were inducted into the manifold as gases, requiring semi-permanent connection to the manifold. All chemicals were used without additional purification or treatment. Nitrogen gas was used in place of pressurized air to prevent oxidation of analyte species by oxygen, which is relatively abundant (~27%) in air. Oxidation products would exhibit different spectral features, as the moment of inertia would change as molecules become oxidized.

Dilution flasks (4) were used to store mixtures of compounds of interest with nitrogen during analysis of samples at ppm concentrations. The dilution procedure and its benefits will be described later. Once analysis of a gas sample was complete, it was

withdrawn from the manifold, cell, transfer lines, and flask through a liquid nitrogen trap (2). This component is used to remove dangerous vapors from the gases before removal with the vacuum pump (Welch Ultratorr model 8907, 1). Were it not for the trap, the vacuum pump oil would become contaminated with the species under analysis, which would affect pumping performance.

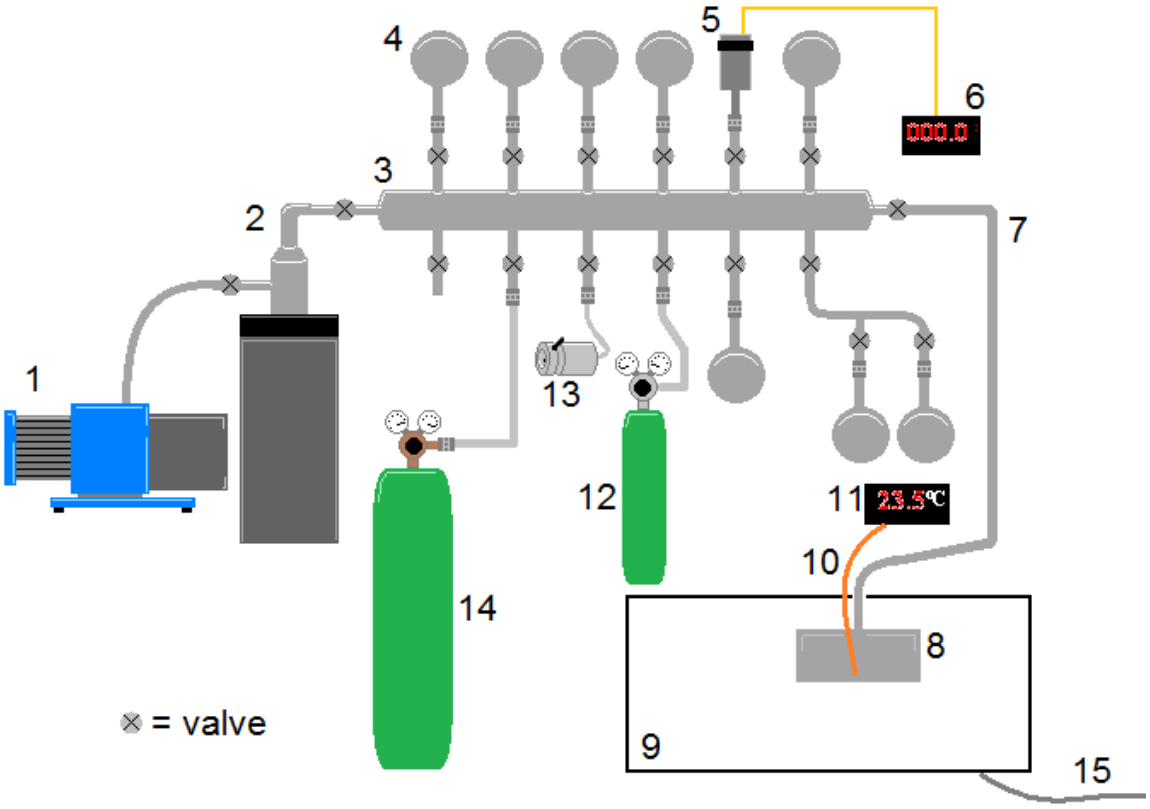


Figure VI-1. Schematic representation of equipment used for gas sample preparation and analysis. The identity and function of each numbered component is described in the text.

Pure Components at Part-Per-Million Concentrations

In this work, each species of interest was analyzed as a pure component (mixed only with nitrogen) and in mixtures (mixed with nitrogen and other species of interest).

Pure components were analyzed at concentrations ranging from 26 ppm to 66.5 parts-per-thousand, and this section is devoted to concentrations ranging from 26 to 710 ppm. The ppm concentrations are similar to those expected for atmospheric measurements. Quantitative performance of linear regression, CLS, and PLS models was examined for each pure component to establish a baseline for analytical performance with mixtures, which more closely mimic real-life atmospheric measurements.

Quantitative analysis was performed with linear regressions, CLS, and PLS. Predictive models were prepared with unprocessed as well as DFF-processed spectra, and multivariate methods were also optimized with wavenumber selection. Filter parameters and wavenumber bounds were optimized as described in Chapter III using the NMS algorithm, where the CVSEP served as the minimized function value. Baseline correction was performed for samples in which baseline was visible in high SNR regions by subtracting the mean of a small blank spectral range from the entire spectrum. Errors reported from linear regressions represent the mean of the error from 5 regressions at various frequencies covering the analyte spectrum.

Experimental

Samples were prepared with a low propagated reference error using a dilution process. This reference error is many times lower than that attained by a single measurement. Preparation is described by the following outline.

1. Stock mixture
    a. evacuate system (0.0 Torr)
    b. insert analyte (various pressures, $P_{analyte}$)
    c. add N_2 (~760.0 Torr, P_{mix})
    d. retain mixture in flask
2. Reference
    a. evacuate system (0.0 Torr)
    b. add N_2 (760.0 Torr)
    c. collect reference spectrum
3. Sample
    a. evacuate system (0.0 Torr)

b. add stock mixture (various pressures, P_{stock})
c. add N_2 (~760.0 Torr, P_{total})
d. adjust volume (760.0 Torr, P_{final})
e. collect sample spectrum

Reference and sample spectra were collected alternately so each sample has a unique reference spectrum.

Given the steps in the outline above, the analyte concentration may be calculated using Equation VI-1.

$$C_{ppm} = 1 \times 10^6 \left(\frac{P_{analyte}}{P_{mix}}\right)\left(\frac{\mathrm{P_{stock}}}{\mathrm{P_{total}}}\right)\left(\frac{\mathrm{P_{final}}}{\mathrm{P_{total}}}\right) \qquad \text{Equation VI-1}$$

Concentration (C_{ppm}, in ppm) is the product of the ratios of the pressures for each step. The first term enclosed in parentheses represents the pressure fraction of analyte in the stock mixture, and the second represents the dilution factor obtained when the mixture is released into the manifold and N_2 is added. The last term in parentheses is the result of volume adjustment. Despite the use of metering valves, it is difficult to add N_2 to a pressure equal to 760.0 Torr. After addition of the nitrogen, metering valves near the pressure transducer and steel transfer lines may be opened further to increase the internal volume, which decreases the pressure of the sample according to the ideal gas law. In the end, the final step enables preparation of samples with tighter control on the overall pressure of the sample.

Preparation of samples by dilution significantly reduces the propagated reference concentration error for ppm samples. For example, if a sample must be prepared with a concentration of 260 ppm, and the error of the pressure measurement is assumed to be equal to the smallest increment on the pressure display, a stock mixture of 12.3 Torr of analyte would be prepared with a mixture pressure of 760.0 Torr. A 12.2 Torr portion would be released and diluted to 760 Torr with N_2 gas, which would produce the propagated error shown in Equation VI-2 (assuming no volume adjustment is performed).

$$\left(\sqrt{\left(\frac{0.1}{12.3}\right)^2+\left(\frac{0.1}{760.0}\right)^2+\left(\frac{0.1}{12.2}\right)^2+\left(\frac{0.1}{760.0}\right)^2}\right)*260\ ppm=\ 2.1\ ppm \quad \text{Equation VI-2}$$

On the other hand, if this sample is prepared using the single-measurement method, a pressure of 0.2 Torr of analyte must be diluted in 760.0 Torr of N_2 gas, as shown in Equation VI-3.

$$\left(\sqrt{\left(\frac{0.1}{0.2}\right)^2+\left(\frac{0.1}{760.0}\right)^2}\right)*260\ ppm=\ 130\ ppm \quad \text{Equation VI-3}$$

The propagated reference error thus prohibits preparation of samples in the ppm range with precisely-known concentrations if the dilution process is not performed. Error in the reference measurement may be reduced by preparing stock mixtures at elevated overall pressures, but this is practically limited to pressures near 760.0 Torr, as the integrity of the glass and sealing components may be compromised at elevated pressures.

Propagated error is calculated for the ppm-range samples using Equation VI-2 with a slight modification. The numerator of each relative error term shown in Equations VI-2 and VI-3 would be replaced with a value which would represent the error as a result of variability in the gauge measurement and readout stability, which are defined by the manufacturer as 0.12% of the reading and 0.01% of the reading ±1digit. These sources of error are assumed to be additive, so the numerator may be calculated using Equation VI-4.

$$e_{num}=\sqrt{(1.2\times10^{-3}r)^2+(1.0\times10^{-4}r)^2+2(0.1)^2} \quad \text{Equation VI-4}$$

The magnitude of the numerator error term (e_{num}) is determined by the pressure reading (r) and the error introduced by the gauge and readout devices. The third term in Equation VI-4 represents the precision of the value shown on the readout, as 0.1 is the smallest increment of the measurement. It is counted twice because each reading is provided relative to a zero-measurement similar to a mass measurement with an analytical balance.

All spectra were collected at a spectral resolution of 0.225 cm^{-1} as the average of 1800 scans, as mentioned in Chapter II. The SCA was used in calculation of all spectra, and the TDS of each reference and sample was extended such that the total signal length was 4×N, where N is the next highest power of 2 from the number of data points in the original signal. Table VI-1 provides statistics for each sample set including number of samples, minimum, maximum, and mean concentration, correlation in the collection sequence, and mean and standard deviation of temperature and final pressure values. Samples were prepared in a manner which reduced correlation between the collection sequence and analyte concentration. This procedure decreases the likelihood that systematic instrumental variance will correlate with the magnitude of analyte spectral features.

Results

Several symbols will be used to represent quantitative data obtained using different procedures. Results which are reported with the symbol "Raw" represent results obtained with spectra taken after baseline correction and across the full spectral range (2-125 cm^{-1}, when using multivariate methods). Values provided with the symbol "WN" represent those obtained with wavenumber window optimization, and those reported with the symbol "DFF" were calculated with optimized wavenumber bounds and digital filter parameters. Table VI-2 lists the wavenumber points used for linear regressions, and Table VI-3 contains the DFF parameters used for linear regression models for each analyte of interest. Digital filter means and widths are provided in units of digital frequency, and digital filter widths represent the half-width of the passed digital frequency band.

Filter parameters tend to follow the frequency content of the respective analyte spectrum. For example, as seen in Figure V-10, the spectrum of ethanol is comprised of a dense collection of peaks which makes the spectrum appear very broad. As a result,

much of ethanol's chemical information in the digital frequency domain is focused on the low-frequency end, and the optimized filter is located at a correspondingly low digital frequency. Conversely, the spectra of methanol or acetonitrile consist of well-separated narrow peaks compared to those of ethanol. As a result, the chemical information for methanol and acetonitrile is focused in higher digital frequencies, and the optimized filter parameters reflect this difference.

Tables VI-4 and VI-5 list the optimized parameters selected for CLS and PLS regression models, respectively. Table VI-6 provides the CVSEP obtained for each species of interest with each chemometric method using the different sets of parameters. The mean reference error and PEL are grouped for comparison in Table VI-7. Table VI-8 provides calculated F values which represent the statistical significance of processing methods used for each algorithm. This table provides a rapid evaluation of the benefit of different quantitative methods. For example, the cell under PLS, Raw-DFF for CH_3CN is highlighted yellow. This means that the error of prediction obtained when using the full spectral range in the PLS model for acetonitrile was significantly greater than the error of prediction obtained when the optimized DFF pre-processing and wavenumber optimization method was used. Concentration correlation plots are shown for each quantitative model in Figures VI-2 through VI-9. The precision of each model is reflected graphically by the respective concentration correlation plots in that imprecise quantitative models exhibit high scattering about the ideal line. The need for optimization of the quantitative models is apparent, despite the simplicity of these samples.

In most cases, digital filtering provided improved predictive ability of the linear regression models. The precision of quantitative models constructed with the un-filtered spectra generally follow the relative absorptivities for each species. Digital filtering is not expected to consistently provide improved predictive performance with linear regression models when implemented in a multi-wavelength fashion with linear

Table VI-1. Part-per-million sample set conditions and statistics. The minimum concentration (C_{min}), maximum concentration (C_{max}), and mean concentration (C_{mean}) are provided for each species, along with the collection sequence correlation, average temperature and average overall pressure across samples.

Species	No. Samples	C_{min} (ppm)	C_{max} (ppm)	C_{mean} (ppm)	R^2_C	T (°C)	P_{final} (Torr)
CH_3CHO	39	27	644	294	6.2E-04	23.7 ± 0.2	760.2 ± 0.4
CH_3CN	39	26	650	291	4.7E-03	24.0 ± 0.6	761 ± 1
NH_3	40	50	656	384	6.6E-05	25 ± 2.3	760 ± 1
CH_3CH_2OH	30	33	699	281	2.9E-01	23.7 ± 0.2	761 ± 1
CH_3OH	40	44	693	362	1.3E-03	28.9 ± 0.2	760.3 ± 0.7
CH_3CH_2CHO	40	37	625	327	2.2E-04	23.8 ± 0.2	760.3 ± 0.7
CH_3CH_2CN	39	38	647	332	3.9E-04	23.6 ± 0.2	760.3 ± 0.8
H_2O	40	43	710	379	1.2E-06	23.6 ± 0.2	760.2 ± 0.5

Table VI-2. Wavenumber points used for linear regressions for each species of interest.

Species	Wavenumbers (cm^{-1})				
CH_3CHO	12.27	17.32	23.18	29.65	32.78
CH_3CN	8.59	13.49	19.65	26.36	31.21
NH_3	19.09	40.51	58.79	78.64	98.34
CH_3CH_2OH	16.67	23.13	31.01	40.35	52.58
CH_3OH	14.50	28.94	40.00	52.98	67.68
CH_3CH_2CHO	8.44	12.42	15.56	25.96	45.25
CH_3CH_2CN	6.87	9.24	14.65	20.91	27.68
H_2O	18.59	38.79	55.71	75.51	100.50

Table VI-3. Parameters (in digital frequency units) used for digital filtered linear regression models constructed with ppm-range sample spectra.

Species	DFF	
	Filter μ	Filter width
CH_3CHO	0.1149	0.0426
CH_3CN	0.0912	0.0283
NH_3	0.0650	0.0504
CH_3CH_2OH	0.0000	0.0023
CH_3OH	0.0296	0.0142
CH_3CH_2CHO	0.0007	0.0011
CH_3CH_2CN	0.0049	0.0166
H_2O	0.0732	0.0421

Table VI-4. Parameters used for CLS quantitative models constructed with ppm-range sample spectra.

Species	WN		DFF			
	WN_{low}[1]	WN_{high}[1]	Filter μ[2]	Filter width[2]	WN_{low}[1]	WN_{high}[1]
CH_3CHO	7.0	27.0	0.0677	0.0298	13.3	66.0
CH_3CN	7.7	27.7	0.1330	0.0347	6.4	41.9
NH_3	3.0	100.0	0.0119	0.0055	15.2	105.7
CH_3CH_2OH	0.6	88.3	0.1276	0.0494	14.3	47.9
CH_3OH	20.2	40.2	0.0366	0.0244	5.0	70.0
CH_3CH_2CHO	1.2	21.2	0.0082	0.0052	16.0	43.2
CH_3CH_2CN	9.3	29.3	0.0078	0.0057	15.8	42.6
H_2O	3.8	87.8	0.0049	0.0049	7.8	91.7

1: Wavenumber units (cm^{-1})

2: Digital frequency units

Table VI-5. Parameters used for PLS quantitative models constructed with ppm-range sample spectra.

Species	Raw	WN			DFF				
	# LV[1]	WN_{low}[2]	WN_{high}[2]	# LV[1]	Filter μ[3]	Filter width[3]	WN_{low}[2]	WN_{high}[2]	# LV[1]
CH_3CHO	1	10.0	30.0	1	0.0737	0.0080	10.5	31.0	1
CH_3CN	2	15.4	45.6	2	0.1109	0.0474	23.2	44.8	1
NH_3	2	3.0	100.0	1	0.0133	0.0079	18.5	99.5	1
CH_3CH_2OH	1	9.5	29.5	1	0.1207	0.0423	11.7	35.7	1
CH_3OH	1	16.8	36.8	2	0.0363	0.0237	9.4	68.0	1
CH_3CH_2CHO	1	16.3	39.0	2	0.0076	0.0050	15.8	42.5	1
CH_3CH_2CN	2	16.0	36.0	2	0.0100	0.0077	16.8	46.3	1
H_2O	1	6.2	92.1	1	0.0468	0.0185	10.2	42.9	1

1: Number of latent variables

2: Wavenumber units (cm^{-1})

3: Digital frequency units

Table VI-6. CVSEP values (in ppm) obtained for each quantitative model for each species of interest.

Species	Linear Regression		CLS			PLS		
	Raw (mean)	DFF (mean)	Raw	WN	DFF	Raw	WN	DFF
CH_3CHO	42	24	211	39	16	182	22	16
CH_3CN	32	8	89	23	4	77	5	4
NH_3	27	23	50	33	29	26	19	16
CH_3CH_2OH	142	138	301	129	54	199	108	55
CH_3OH	35	23	164	44	25	174	17	16
CH_3CH_2CHO	123	117	312	62	28	194	28	30
CH_3CH_2CN	99	101	137	42	11	114	14	9
H_2O	49	48	98	92	88	59	45	45

Table VI-7. Mean reference error, PEL, and best CVSEP for each analyte in ppm. The quantitative method which produces the best result is indicated in the last column.

Species	Mean ref. error	PEL	Best CVSEP	Best Method
CH_3CHO	5.4	200	16	PLS, DFF[1]
CH_3CN	5	40	4	PLS, DFF[1]
NH_3	6	50	16	PLS, DFF
CH_3CH_2OH	4.5	1000	54	CLS, DFF[1]
CH_3OH	5.8	200	16	PLS, DFF
CH_3CH_2CHO	5.6	X	28	CLS, DFF[1]
CH_3CH_2CN	5.6	X	9	PLS, DFF[1]
H_2O	6	X	45	PLS, WN/DFF

1: CLS and PLS results are not significantly different

Table VI-8. Calculated F values comparing the errors of prediction obtained with different processing methods. Significant differences are indicated with yellow highlighted values. Models constructed with the un-processed spectra are denoted with the word Raw; those constructed with wavenumber optimization are indicated with the symbol WN, and those constructed with digital filtering and wavenumber optimization are represented by the DFF abbreviation.

Species	F_{crit}	F_{calc}						
		Lin. Reg.	CLS			PLS		
		Raw-DFF	Raw-WN	Raw-DFF	WN-DFF	Raw-WN	Raw-DFF	WN-DFF
CH_3CHO	1.72	3.07	30.05	177.46	5.90	70.53	126.91	1.80
CH_3CN	1.72	17.13	15.56	444.56	28.57	199.94	452.31	2.26
NH_3	1.70	1.37	2.37	2.89	1.22	1.87	2.44	1.31
CH_3CH_2OH	1.86	1.81	5.39	31.38	5.82	3.38	13.30	3.93
CH_3OH	1.70	2.30	13.69	43.52	3.18	107.62	115.47	1.07
CH_3CH_2CHO	1.70	1.11	25.42	124.59	4.90	48.83	42.83	1.14
CH_3CH_2CN	1.72	1.03	10.50	165.06	15.71	62.57	163.89	2.62
H_2O	1.70	1.03	1.13	1.22	1.08	1.70	1.70	1.00

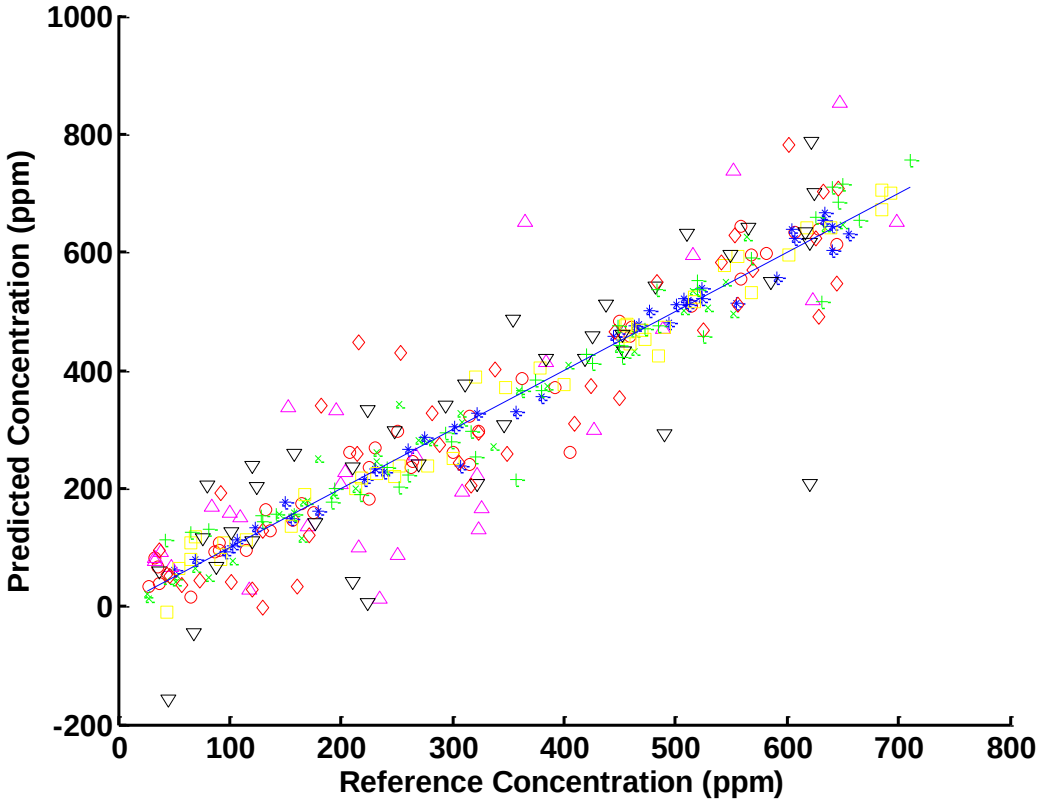


Figure VI-2. Concentration correlation from linear regression models constructed with un-processed spectra. Acetaldehyde is represented by red circles, acetonitrile by green x's, ammonia by blue asterisks, ethanol by magenta triangles, methanol by yellow squares, propionaldehyde by inverted black triangles, propionitrile by red diamonds, and water vapor by green +'s. The ideal line is shown as a solid blue line with a slope of 1.

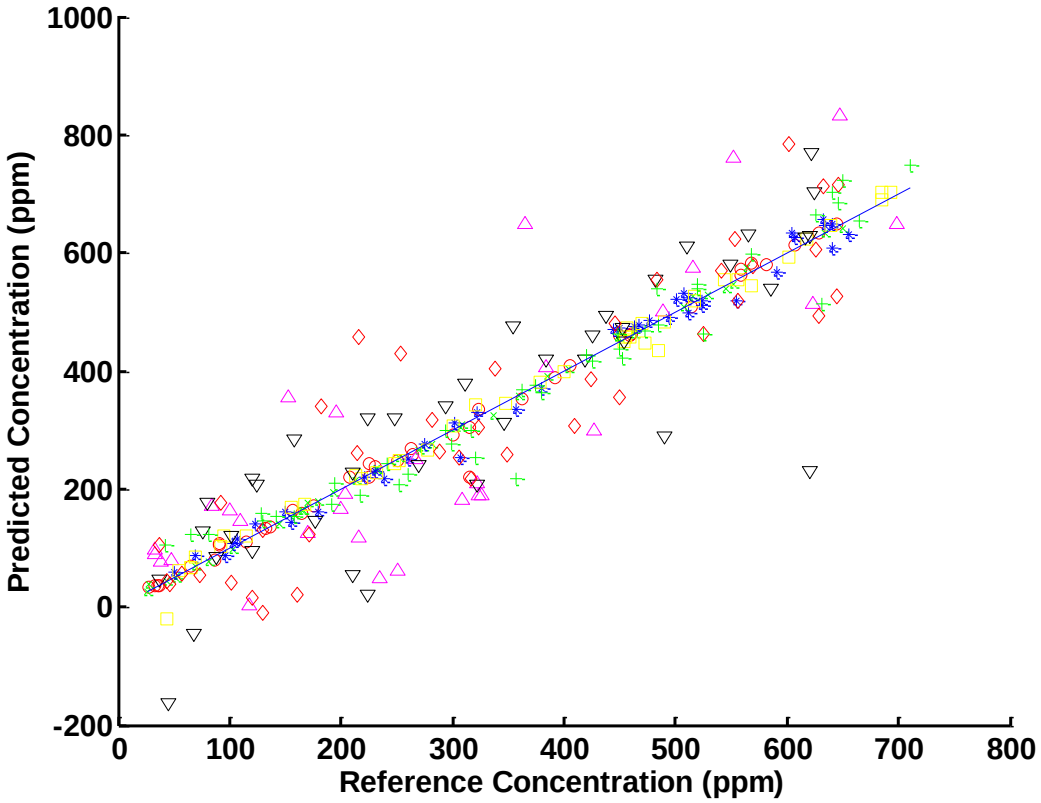


Figure VI-3. Concentration correlation from linear regression models constructed with DFF-processed spectra. Acetaldehyde is represented by red circles, acetonitrile by green x's, ammonia by blue asterisks, ethanol by magenta triangles, methanol by yellow squares, propionaldehyde by inverted black triangles, propionitrile by red diamonds, and water vapor by green +'s. The ideal line is shown as a solid blue line with a slope of 1.

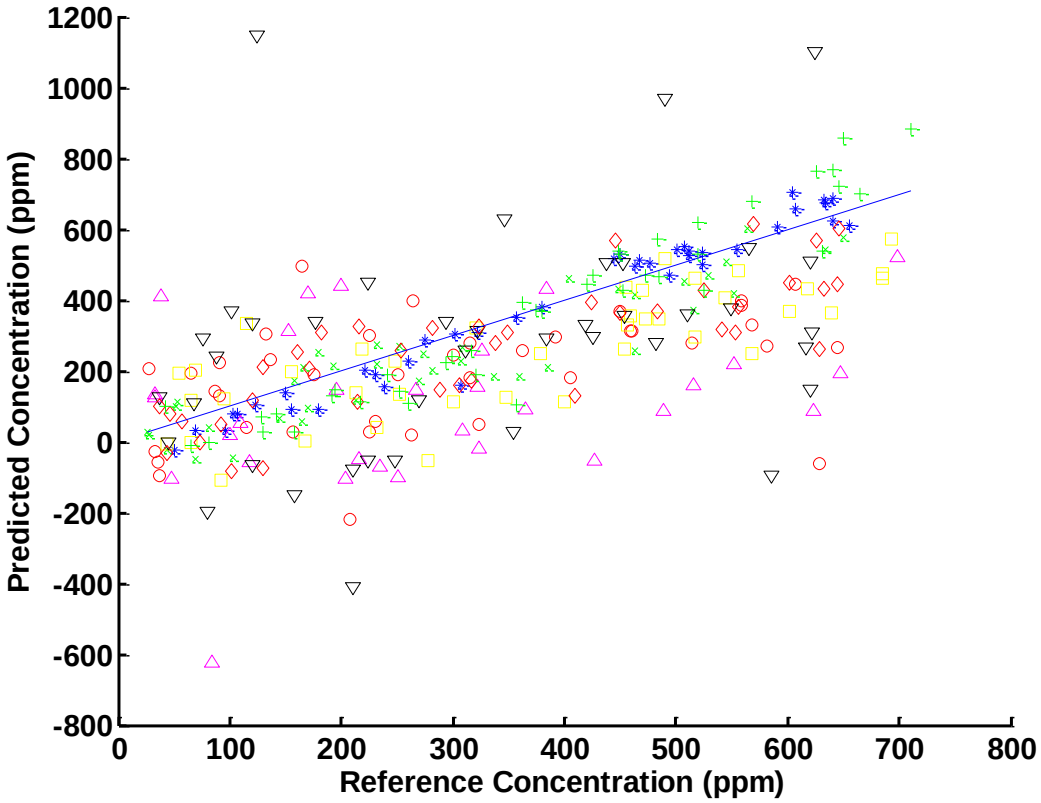


Figure VI-4. Concentration correlation from CLS regression models constructed with the full spectral range. Acetaldehyde is represented by red circles, acetonitrile by green x's, ammonia by blue asterisks, ethanol by magenta triangles, methanol by yellow squares, propionaldehyde by inverted black triangles, propionitrile by red diamonds, and water vapor by green +'s. The ideal line is shown as a solid blue line with a slope of 1.

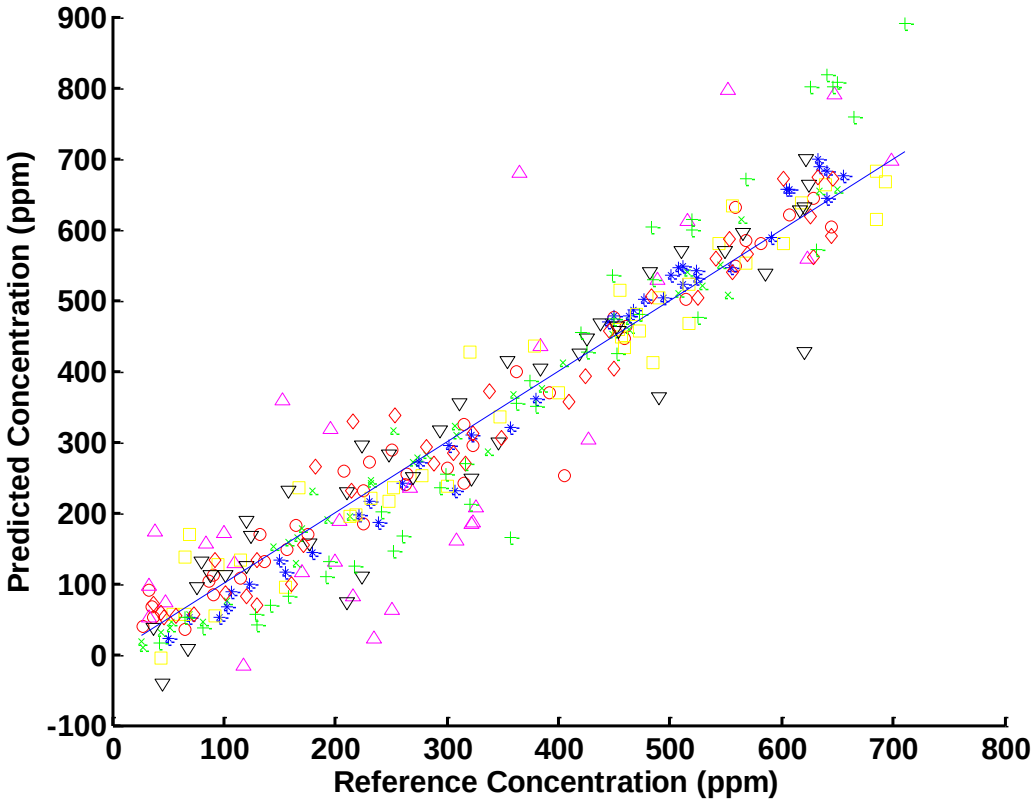


Figure VI-5. Concentration correlation for CLS models constructed using the optimized spectral ranges. Acetaldehyde is represented by red circles, acetonitrile by green x's, ammonia by blue asterisks, ethanol by magenta triangles, methanol by yellow squares, propionaldehyde by inverted black triangles, propionitrile by red diamonds, and water vapor by green +'s. The ideal line is shown as a solid blue line with a slope of 1.

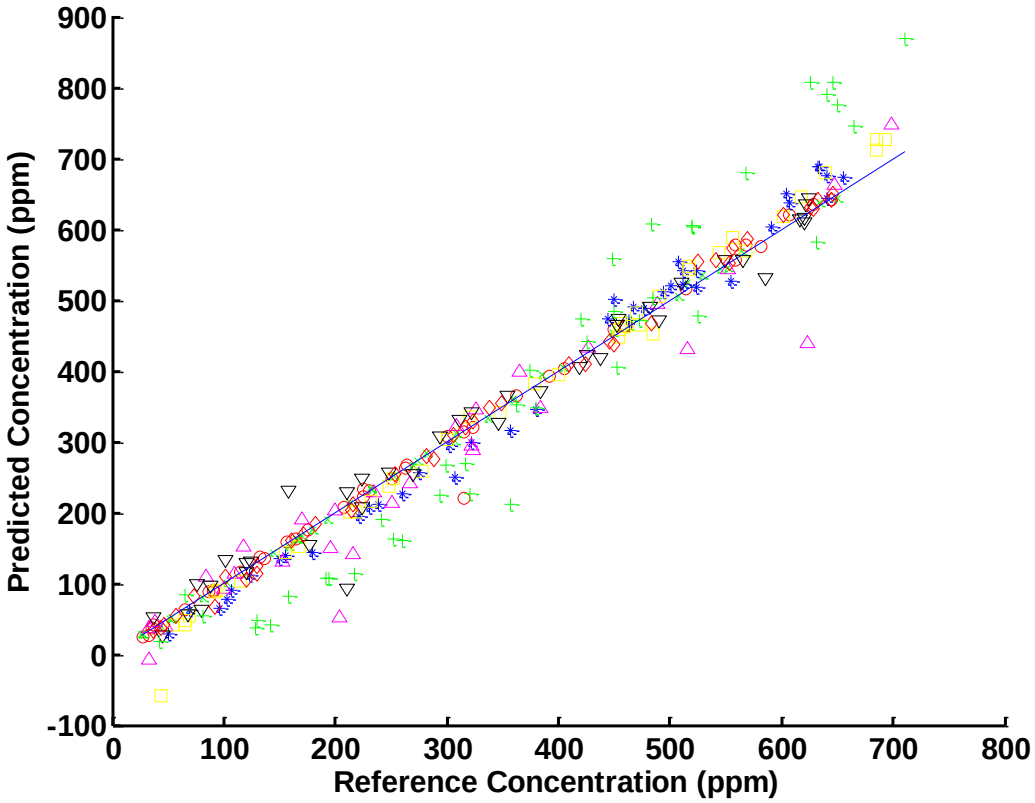


Figure VI-6. Concentration correlation for CLS models using the optimized spectral range and DFF pre-processing. Acetaldehyde is represented by red circles, acetonitrile by green x's, ammonia by blue asterisks, ethanol by magenta triangles, methanol by yellow squares, propionaldehyde by inverted black triangles, propionitrile by red diamonds, and water vapor by green +'s. The ideal line is shown as a solid blue line with a slope of 1.

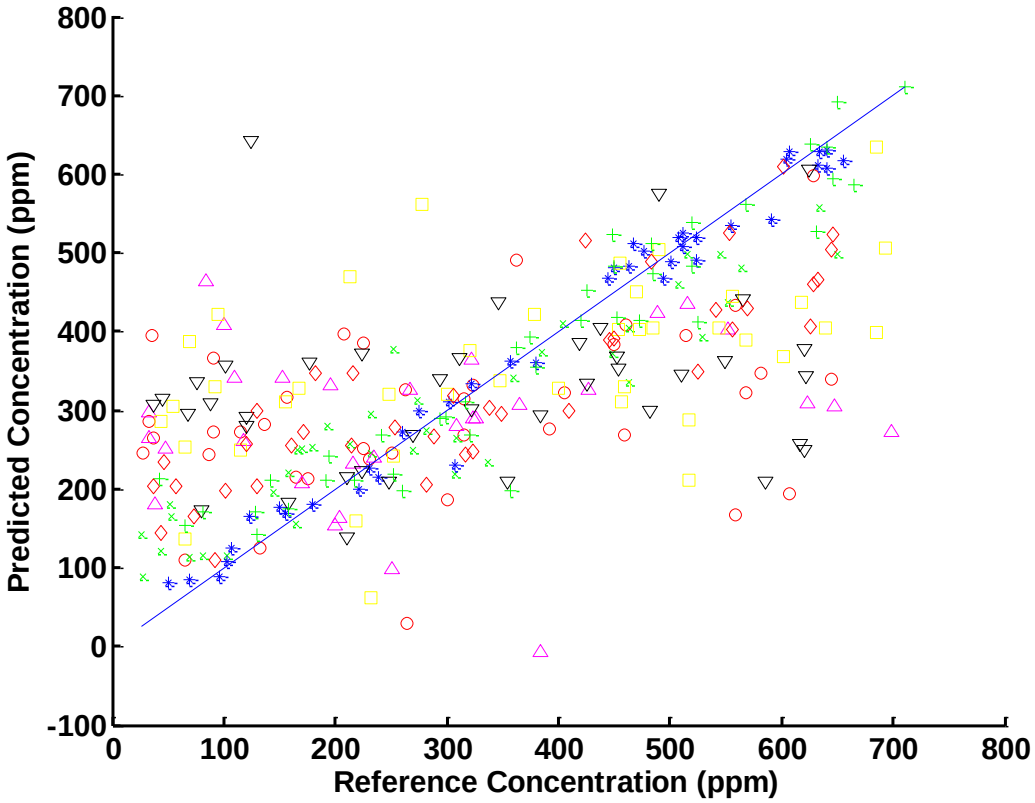


Figure VI-7. Concentration correlation for PLS regression models constructed with the full spectral range. Acetaldehyde is represented by red circles, acetonitrile by green x's, ammonia by blue asterisks, ethanol by magenta triangles, methanol by yellow squares, propionaldehyde by inverted black triangles, propionitrile by red diamonds, and water vapor by green +'s. The ideal line is shown as a solid blue line with a slope of 1.

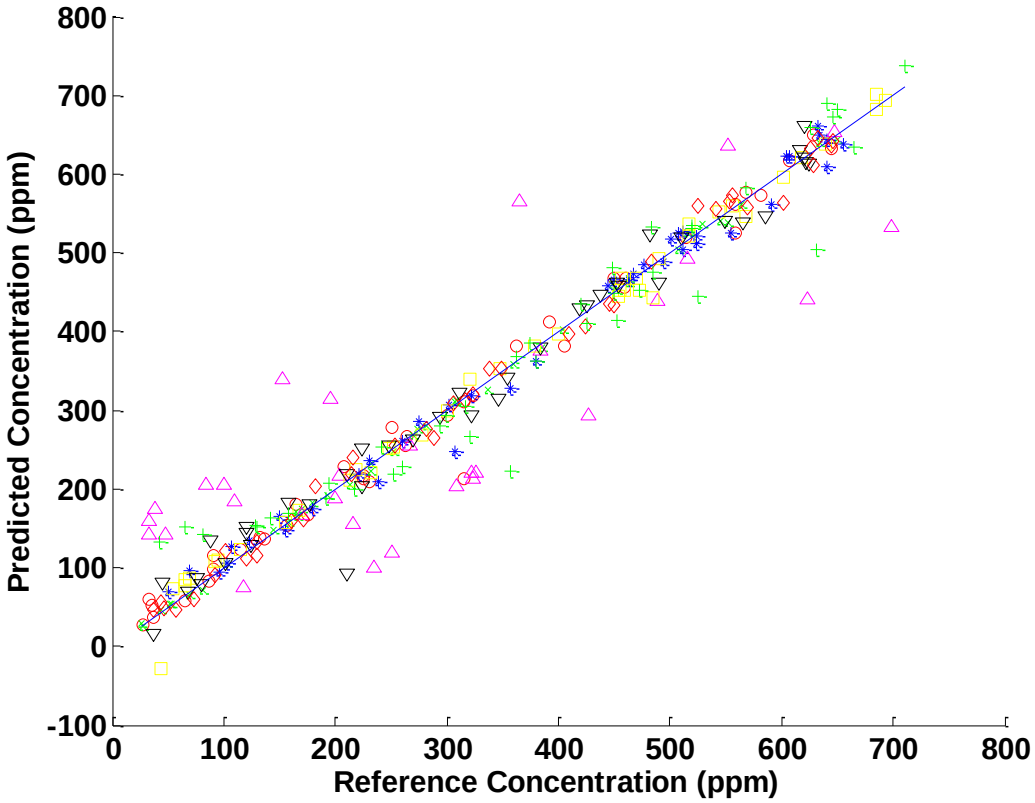


Figure VI-8. Concentration correlation for PLS regression models constructed with the optimized spectral range. Acetaldehyde is represented by red circles, acetonitrile by green x's, ammonia by blue asterisks, ethanol by magenta triangles, methanol by yellow squares, propionaldehyde by inverted black triangles, propionitrile by red diamonds, and water vapor by green +'s. The ideal line is shown as a solid blue line with a slope of 1.

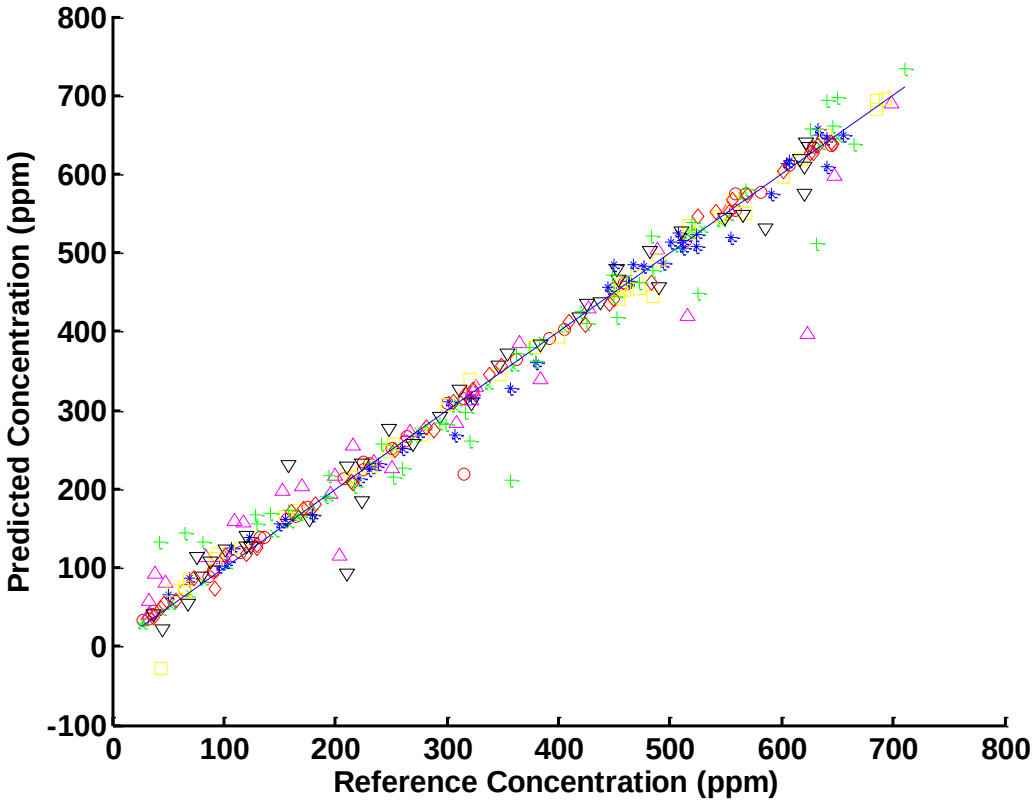


Figure VI-9. Concentration correlation for PLS models constructed with the optimized spectral range and DFF parameters. Acetaldehyde is represented by red circles, acetonitrile by green x's, ammonia by blue asterisks, ethanol by magenta triangles, methanol by yellow squares, propionaldehyde by inverted black triangles, propionitrile by red diamonds, and water vapor by green +'s. The ideal line is shown as a solid blue line with a slope of 1.

regressions, as improvements will only be observed if a given filter enhances the SNR at the wavelength points selected for regression. This explains the mixed results when comparing the linear regression models constructed with the unprocessed spectra to those of the regression models constructed with DFF pre-processed spectra.

The possibility of improvement of linear regression model precision exists, as evidenced by some values in the first two columns of Table VI-6 and the first column of Table VI-8, though some results suggest the possible need for simultaneous filter parameter and regression wavelength point optimization. Were the frequency points selected for regression and DFF parameters optimized concurrently, wavelength points might be selected at which higher SNRs are available. However, due to inherent limitations of the linear regression algorithm, no further efforts were expended towards improvement of the linear regression models.

Errors of prediction were below the PELs for each species when using the raw spectra or filtered spectra, and are improved significantly for almost all species when using wavenumber selection or DFF/wavenumber selection with the multivariate methods. Insignificant improvement is seen in the analyses of water vapor absorption spectra when different processing methods are compared, as shown in Table VI-8. This may be the result of variability in the concentration of water vapor present in the gas cell as a result of small leaks or adsorption, making spectral variance a less-significant contributor to the total error in the measurement.

Similar trends may be found amongst the errors of prediction obtained with CLS and PLS models. Predictive performance improved significantly when optimized wavenumber ranges were used for calibration and prediction, and the optimized ranges are primarily focused in the spectral regions in which a high SNR is available for the respective analyte. Further improvement was observed for most species when DFF was used for pre-processing, though the improvement was not significant in all cases as indicated by values in Table VI-8. These spectra were also used to calculate the

multivariate SNR and LOD as described in Chapter III. The background matrix for the figures of merit consisted of vectors representing instrumental variance, which were the first five principal components of the hundred percent lines collected at the start of each day of data acquisition. Significance of each principal component was evaluated via the F-test. Figures of merit were calculated based on the parameters used for PLS cross-validation and are tabulated in Tables VI-9 and VI-10.

The calculated figures of merit generally follow the trends in the errors of prediction from one species to another. When a high CVSEP was observed, a similarly poor LOD was calculated, and a correspondingly-low SNR resulted. The LOD calculated for water vapor is relatively poor despite its high absorptivity in this frequency range, which results from leaks in the sample handling system. As mentioned previously, the background matrix is constructed with principal components from hundred percent lines. The spectra were collected with the cell and sampling system under vacuum to remove species from the system, but the decreased pressure causes water vapor features to become sharpened, and the sensitivity is increased. Overlap of the analyte features, i.e. the water vapor added during sample analysis with those of the background, i.e. water vapor leaking into the system, results in a poorer multivariate LOD due to the decreased magnitude of the β_{NAS}.

As mentioned in Chapter III, the efficacy of PLS regression models may be evaluated by comparing the PLS regression vector with that of the NAS algorithm. Representative vectors are displayed in Figures VI- 10 through VI-15 for water and methanol pure component samples. In general, the regression vectors constructed with the full spectral range exhibited great similarity in shape, but not magnitude. PLS models for all compounds produced regression coefficients lower than those of the NAS models when using the full spectral range. Vectors calculated using the optimized spectral ranges show similar spectral shapes and differing magnitudes as with the full spectral range. However, vectors calculated using the DFF pre-processed spectra are very

Table VI-9. The multivariate SNR calculated for each species of interest using the wavenumber bounds and filter parameters obtained from optimization of PLS regression models.

Species	Raw	WN	DFF
CH_3CHO	13.6	19.7	19.9
CH_3CN	18.1	24.2	20.4
NH_3	35.5	42.1	39.1
CH_3CH_2OH	11.2	23.6	13.0
CH_3OH	14.3	18.6	31.2
CH_3CH_2CHO	14.9	18.9	20.5
CH_3CH_2CN	16.3	19.5	23.7
H_2O	28.3	38.6	24.4

Table VI-10. The multivariate LOD (in ppm) calculated for each species of interest using parameters obtained from optimization of PLS regression models.

Species	Raw	WN	DFF
CH_3CHO	63.9	44.6	44.0
CH_3CN	46.5	35.9	42.5
NH_3	29.1	26.6	29.0
CH_3CH_2OH	73.0	32.0	57.7
CH_3OH	75.1	55.8	32.6
CH_3CH_2CHO	64.5	47.6	44.9
CH_3CH_2CN	59.6	50.7	41.5
H_2O	35.7	28.0	45.1

different from those of the other two models. When comparing vectors calculated for the strongest absorbers (water and ammonia gases), the shape and magnitude of the regression vectors exhibit less agreement than for the weaker absorbers, which may

explain the miniscule change in the CVSEP between PLS models constructed using the optimized wavenumber range and DFF pre-processed spectra for the strong absorbers. Excellent agreement is observed between the regression vectors calculated for the weak absorbers when DFF pre-processing is utilized. Fair agreement is seen in the comparisons of the regression vectors, with some spectral noise superimposed on the analyte features.

For some species (acetonitrile, ammonia, ethanol, and water vapors), the multivariate LOD is higher for digital filtered spectra relative to the wavenumber-optimized models, despite diminished errors of prediction. The cause of this change becomes apparent when comparing absorption spectra of the analyte and background before and after DFF pre-processing, as presented for ethanol in Figures VI-16 and VI-17. Digital filtering with the optimized values produces a spectrum in which the magnitude of the analyte features is not much greater than the magnitude of the background variance, whereas the difference in magnitude is much greater in the unprocessed spectra. As a result, the multivariate LOD is worse. This is thought to be purely an artifact of the LOD calculation method, as the regression coefficient vectors of the corresponding PLS and NAS models appear to be similar in shape and magnitude, as depicted in Figure VI-18. It is worth noting that the magnitude and to a lesser extent the shape of the regression vectors will not be identical between the two regression vectors in practice because the assumed background contributors and actual background contributors are never exactly the same between data sets.

Conclusions

Errors of prediction calculated with various chemometric algorithms suggest the utility of the THz-TDS for atmospheric measurements. Despite the short path length (7.78 cm), appreciable sensitivity is achieved and each analyte molecule is quantified to ppm precision. Several CVSEP values are also well below the PELs defined by OSHA.

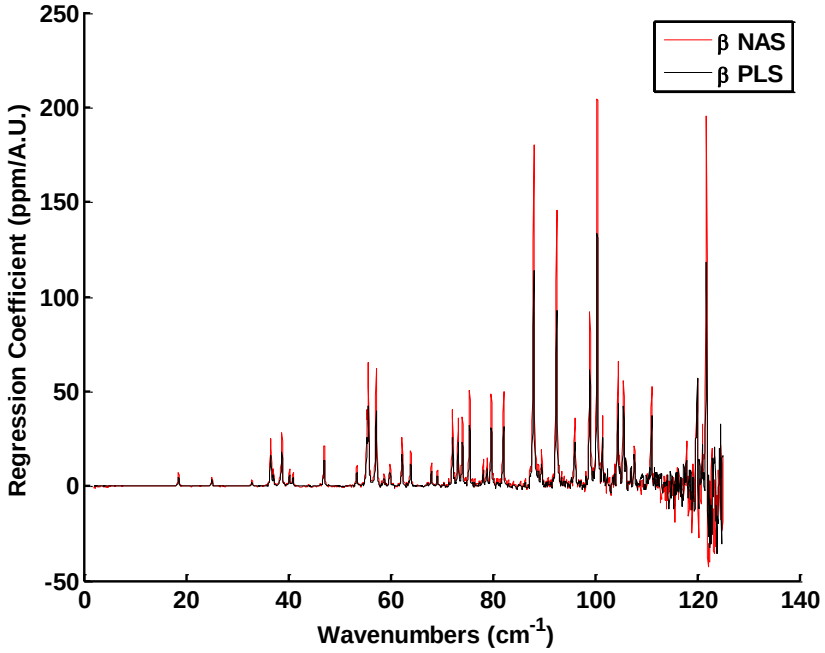


Figure VI-10. NAS and PLS regression vectors constructed with the full spectral range from water vapor spectra.

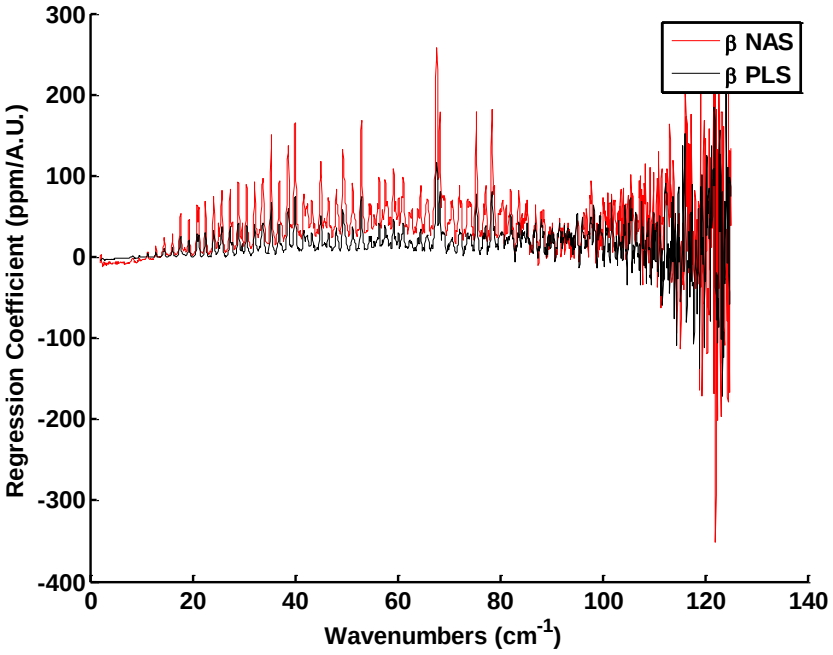


Figure VI-11. NAS and PLS regression vectors constructed with the full spectral range from methanol vapor spectra.

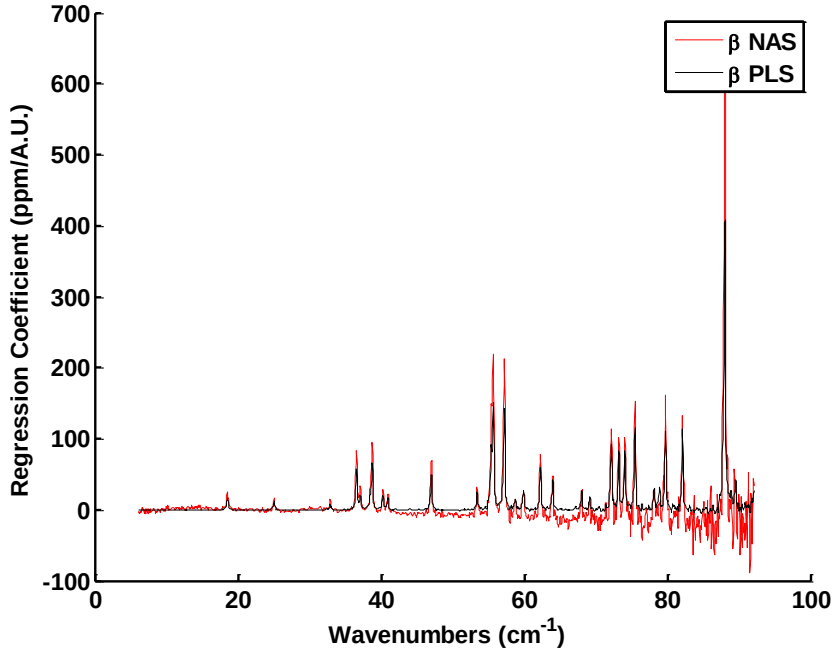


Figure VI-12. NAS and PLS regression vectors constructed with the optimized spectral range from water vapor spectra.

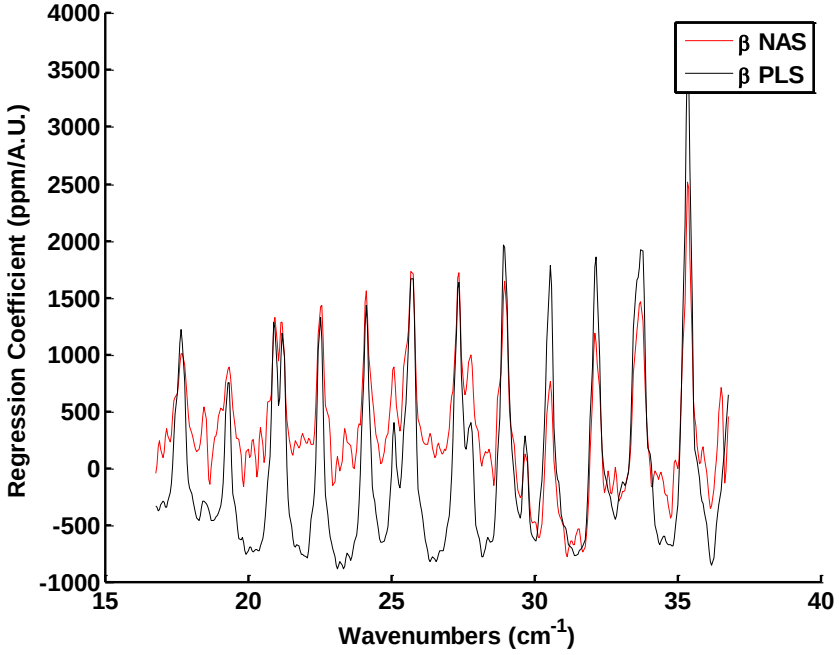


Figure VI-13. NAS and PLS regression vectors constructed with the optimized spectral range from methanol vapor spectra.

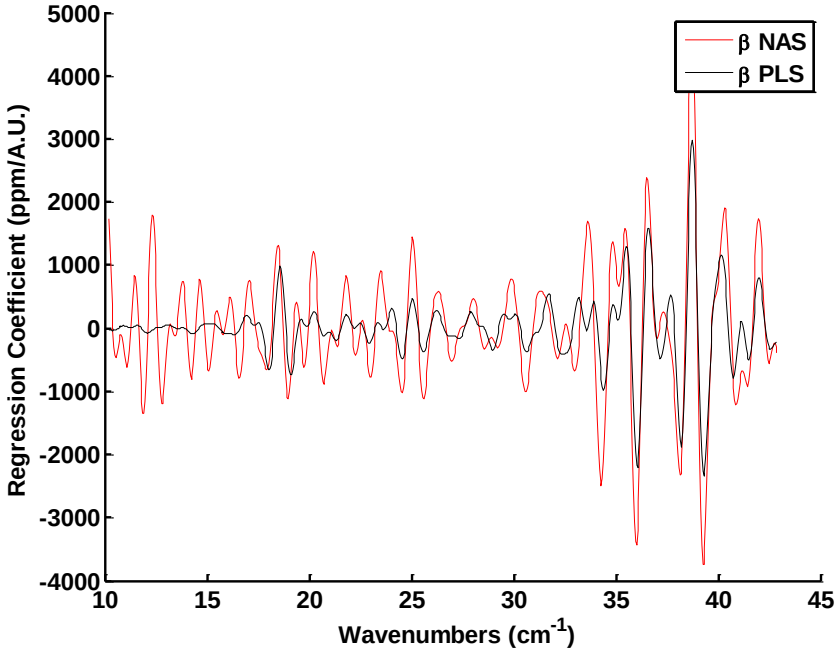


Figure VI-14. NAS and PLS regression vectors constructed with the DFF pre-processed water vapor spectra.

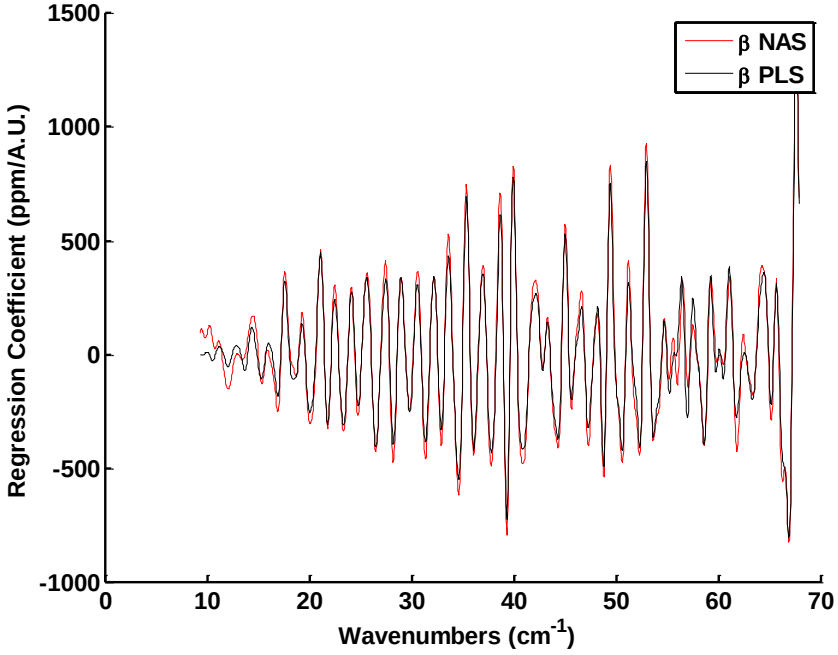


Figure VI-15. NAS and PLS regression vectors constructed with the DFF pre-processed methanol vapor spectra.

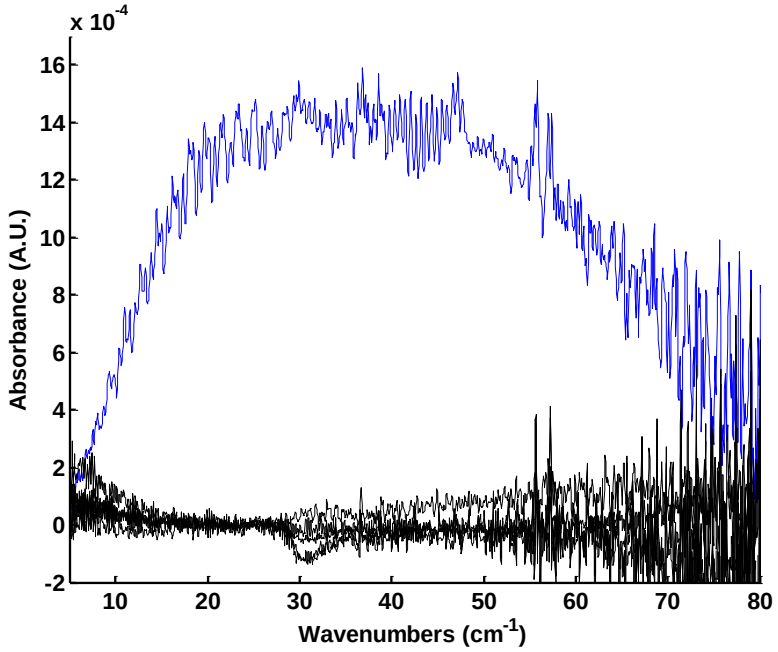


Figure VI-16. Mean absorption spectrum of ppm-range ethanol samples (blue) and spectra of background variance (black) without pre-processing.

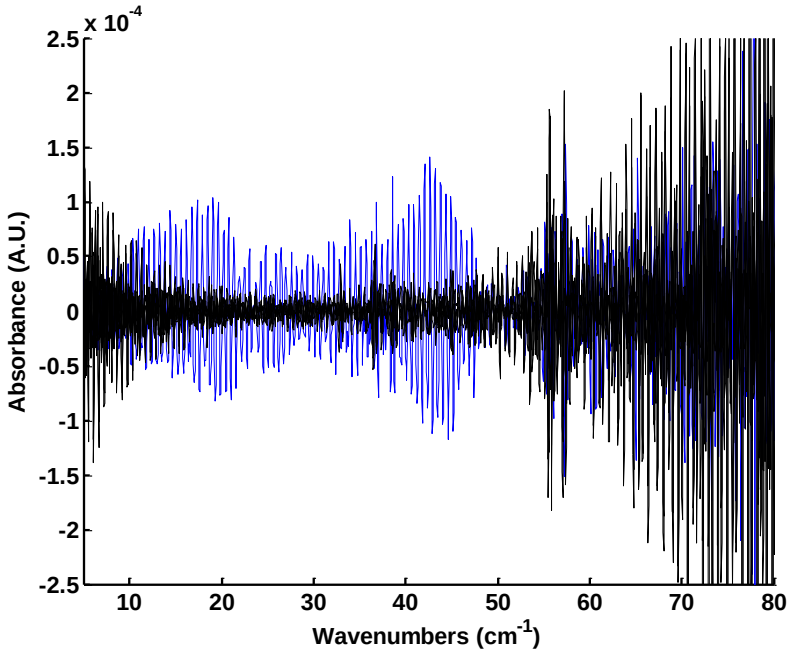


Figure VI-17. Mean absorption spectrum of ppm-range ethanol samples (blue) and spectra of background variance (black) with DFF pre-processing.

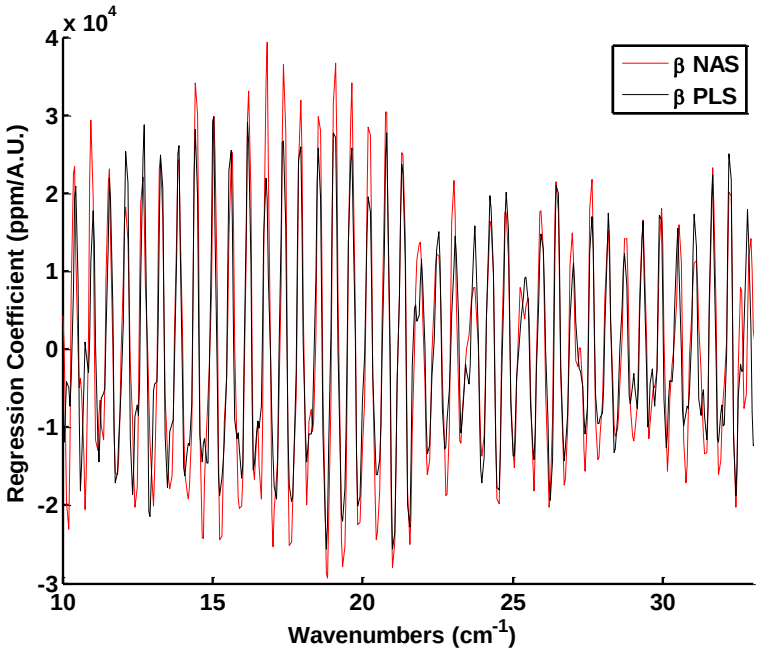


Figure VI-18. Regression coefficient vectors for PLS and NAS models constructed with DFF-processed ethanol absorption spectra.

Wavenumber selection and DFF pre-processing provided significant improvement in the quantitative performance of the predictive models for most analytes. The PLS algorithm typically provided the lowest error of prediction when quantitation was performed with the full spectral range and with the optimized wavenumber range, but CLS and PLS provide similar performance with weakly-absorbing species when DFF pre-processing is used in combination with wavenumber selection. The errors of prediction for acetonitrile with wavenumber optimization and digital filtering are comparable to the propagated reference error. Agreement was observed between regression vectors calculated for PLS and NAS quantitative models, except for models constructed with DFF pre-processing of the absorption spectra of water vapor and ammonia, indicating selectivity of these models against background variance. Wavenumber optimization is expected to be sufficient for improvement of the quantitative models for H_2O and NH_3.

Pure Components at Part-Per-Thousand Concentrations

Though often used to represent concentrations on the order of parts-per-trillion, the ppt designation is used here to represent concentrations on the order of parts-per-thousand. One would not expect to find polluting species in the atmosphere at this concentration except in the worst of scenarios, but quantitation of these samples demonstrates the linear range for each compound. Nearly all published accounts of quantitative efforts based on THz-TDS measurement describe work in this concentration range.

Experimental

Reference concentrations represent the fractional pressure of the analyte, as calculated from pressure readings using Equation VI-5.

$$C_{ppt} = 1 \times 10^3 \left(\frac{P_{analyte}}{P_{total}}\right)\left(\frac{P_{final}}{P_{total}}\right) \qquad \text{Equation VI-5}$$

The concentration (C_{ppt}) may be determined by dividing the measured pressure of the analyte ($P_{analyte}$) by the total pressure (P_{total}), which was near one atm. This is the single-measurement method mentioned in the previous section. Samples may not be prepared at the ppt concentration level using the dilution procedure, as liquid analytes injected into the manifold will condense as the saturation pressure is approached. The final pressure for each sample was also adjusted using the volume adjustment mentioned in the previous section. Propagated error may be calculated using Equation VI-3, where the numerator of each relative error term is defined by Equation VI-4. Samples were prepared such that correlation was minimized between analyte concentration and collection sequence so systematic instrumental variance would not correlate with the analyte. The following steps were followed for sample preparation and data acquisition:

1. Reference
    a. evacuate system (0.0 Torr)
    b. add N_2 gas (760.0 Torr)
    c. collect reference spectrum

2. Sample
    a. evacuate system (0.0 Torr)
    b. add analyte (various pressures, $P_{analyte}$)
    c. add N_2 gas (~ 760.0 Torr, P_{total})
    d. adjust overall pressure (760.0 Torr, P_{final})
    e. collect sample spectrum

Reference and sample spectra were collected in an alternating fashion such that each sample spectrum was referenced from a unique spectrum collected with 760 Torr of nitrogen. Spectral processing and quantitative analysis were performed in a manner consistent with the ppm-range samples. Parameters were optimized for the ppt-range samples, and quantitative analysis was also performed with the optimized parameters from the ppm-range samples, denoted in tables by the subscript "ppm". Samples were prepared in a sequence which reduced the correlation between the sample number and the analyte concentration as with ppm-range samples. Table VI-11 lists the conditions and statistics for samples prepared with each analyte.

Results

Optimized digital filter parameters for linear regressions are tabulated in Table VI-12. The parameters are different than those obtained for the ppm-range samples of most species. Such changes arise as a result of changing digital frequency content in the spectral data. As the analyte concentrations increase, self-broadening increases. This increasing of peak widths will be manifested as increased magnitude in the low-frequency end of the digital frequency spectrum. The change in digital filters from the ppm-range to ppt-range data sets is illustrated in Figure VI-19. Optimized filter parameters do remain fairly consistent in some for some species, but the change was fairly large for ethanol, methanol, and water vapor samples.

Trends in optimized parameters used for CLS and PLS regression models are similar to those of the ppm-range samples, as shown in Tables VI-13 and VI-14. Optimized spectral ranges are fairly different for the strongest absorbers when comparing

the two concentration ranges. This can be attributed to the high absorptivity of these compounds at the high-frequency end of the recorded THz spectra. At such high concentrations, spectral peaks in this end of the spectrum tend to exceed the linear range of the instrument, and the precision of quantitative models including these regions is diminished. As a result, the NMS moves the spectral range toward the low-frequency end where the absorptivity is lower and the precision is enhanced. It is this, in combination with changing digital frequency content which results in diminished predictive ability when ppt-range samples are processed and quantified using the optimized parameters developed with ppm-range spectral data. Predictive errors from models constructed with optimized parameters approach the propagated reference concentration errors listed in Table VI-18, indicating good linearity in the instrument response. Concentration correlation plots shown in Figures VI-20 through VI-27 illustrate the precision available with various quantitative models.

Diminished precision is observed in quantitative models developed using the full spectral range for many of the species. Scatter about the ideal line is not so remarkable for the ppt-range samples, especially when compared to the ppm-range quantitative models. Enhanced precision is still somewhat visible when comparing the concentration correlation plots from un-optimized models to those developed with the optimized spectral range and/or DFF parameters. Distinct non-linearity is visible for un-optimized models for ammonia and water vapor samples, likely a result of analyte-related absorption exceeding the linear range of the THz-TDS. Samples with higher concentrations than those shown on the concentration correlation plots were excluded because of non-linearity observed in the spectra prior to construction of the quantitative models. Further investigation of PLS models revealed some similarities with those of the ppm-range samples. Representative PLS regression vectors are overlaid with corresponding NAS regression vectors in Figures VI-28 through VI-33.

Table VI-11. Part-per-thousand sample set conditions and statistics. Correlation coefficients represent the correlation between sample concentrations and collection sequence.

Species	No. Samples	C_{min} (ppt)	C_{max} (ppt)	C_{mean} (ppt)	R^2_C [1]	T (°C) [2]	P_{final} (torr) [2]
CH_3CHO	13	1.2	48.9	15.0	2.4E-05	23.8 ± 0.2	760.4 ± 1.5
CH_3CN	15	1.0	37.7	14.8	8.1E-06	23.7 ± 0.2	761 ± 1
NH_3	15	1.3	33.8	13.9	2.2E-04	23.7 ± 0.2	761 ± 1
CH_3CH_2OH	13	0.8	23.5	10.3	9.8E-06	23.7 ± 0.3	760.3 ± 0.7
CH_3OH	12	4.0	66.5	32.2	1.4E-02	23.4 ± 0.2	761 ± 2.1
CH_3CH_2CHO	12	1.2	37.7	12.8	4.3E-03	23.8 ± 0.2	760.0 ± 0.1
CH_3CH_2CN	12	0.5	24.5	10.2	7.7E-04	23.8 ± 0.2	760.7 ± 1.3
H_2O	11	1.1	12.9	7.3	4.5E-03	23.6 ± 0.3	760.3 ± 0.7

1: Correlation between sample concentrations and collection sequence.

2: Mean ± standard deviation of measured sample temperatures and final pressures of all samples.

Table VI-12. Digital filter parameters used for linear regressions with ppt-range sample spectra.

Species	DFF	
	Filter μ[1]	Filter width[1]
CH_3CHO	0.1149	0.0426
CH_3CN	0.0039	0.0032
NH_3	0.0052	0.0105
CH_3CH_2OH	0.0037	0.0024
CH_3OH	0.0051	0.0049
CH_3CH_2CHO	0.0041	0.0032
CH_3CH_2CN	0.0057	0.0085
H_2O	0.0049	0.0049

1: Digital frequency units.

Table VI-13. Parameters used for CLS quantitative models constructed with ppt-range sample spectra.

Species	WN		DFF			
	WN_{low}[1]	WN_{high}[1]	Filter μ[2]	Filter width[2]	WN_{low}[1]	WN_{high}[1]
CH_3CHO	0.0	20.0	0.0677	0.0298	6.9	40.3
CH_3CN	31.6	60.5	0.1330	0.0347	44.4	63.7
NH_3	20.0	40.0	0.0071	0.0140	17.1	51.7
CH_3CH_2OH	17.4	37.4	0.0185	0.0145	9.6	31.7
CH_3OH	20.2	40.2	0.0049	0.0049	5.0	70.0
CH_3CH_2CHO	10.1	30.1	0.0063	0.0020	15.5	41.5
CH_3CH_2CN	12.7	41.4	0.0088	0.0072	18.6	38.9
H_2O	16.3	36.3	0.0049	0.0049	14.0	55.0

1: Wavenumber units (cm^{-1}).

2: Digital frequency units.

Table VI-14. Parameters used for PLS quantitative models constructed with ppt-range sample spectra.

Species	Raw	WN			DFF				
	# LV[1]	WN_{low}[2]	WN_{high}[2]	# LV[1]	Filter μ[3]	Filter width[3]	WN_{low}[2]	Wn_{high}[2]	#LV[1]
CH_3CHO	1	7.0	27.0	1	0.074	0.008	35.5	68.2	1
CH_3CN	1	30.3	51.4	1	0.111	0.047	0.0	116.8	1
NH_3	2	34.2	54.2	2	0.010	0.015	42.7	62.9	2
CH_3CH_2OH	1	1.0	21.0	2	0.001	0.001	5.3	23.3	2
CH_3OH	1	16.8	36.8	1	0.005	0.005	9.4	68.0	1
CH_3CH_2CHO	2	0.0	20.0	1	0.008	0.004	20.1	40.3	1
CH_3CH_2CN	1	10.7	30.7	1	0.010	0.005	12.2	46.1	2
H_2O	1	7.0	90.3	2	0.006	0.005	75.0	88.5	2

1: Number of latent variables.

2: Wavenumber units (cm^{-1}).

3: Digital frequency units.

Table VI-15. Error of prediction from linear regression models, given in units of ppt.

Species	Raw (mean)[1]	DFF (mean)[1]	DFF_{ppm} (mean)[2]
CH_3CHO	0.4	0.2	4.8
CH_3CN	1.0	0.3	5.2
NH_3	4.7	1.3	8.2
CH_3CH_2OH	0.8	0.5	0.5
CH_3OH	1.5	0.8	1.6
CH_3CH_2CHO	0.4	0.4	0.4
CH_3CH_2CN	0.7	0.7	0.7
H_2O	2.8	0.5	4.4

1: Results obtained with optimized parameters.

2: Results obtained using optimized parameters from ppm-range samples.

Table VI-16. Error of prediction from CLS models, given in units of ppt.

Species	Raw[1]	WN[1]	WN_{ppm}[2]	DFF[1]	DFF_{ppm}[2]
CH_3CHO	0.2	0.2	0.3	0.1	5.2
CH_3CN	0.5	0.2	0.5	0.6	6.4
NH_3	4.6	0.3	4.5	0.1	3.5
CH_3CH_2OH	0.6	0.5	0.6	0.4	0.7
CH_3OH	1	0.9	1	0.7	1.8
CH_3CH_2CHO	0.6	0.4	0.4	0.2	0.2
CH_3CH_2CN	0.7	0.5	0.6	0.5	0.5
H_2O	1.1	0.3	0.3	0.2	0.5

1: Results obtained with optimized parameters.

2: Results obtained using optimized parameters from ppm-range samples.

Table VI-17. Error of prediction from PLS models, given in units of ppt.

Species	Raw[1]	WN[1]	WN_{ppm}[2]	DFF[1]	DFF_{ppm}[2]
CH_3CHO	0.2	0.2	0.2	0.1	4.8
CH_3CN	0.4	0.2	0.3	0.2	2.6
NH_3	1.5	0.1	1.3	0.1	1.2
CH_3CH_2OH	0.5	0.2	0.2	0.2	0.4
CH_3OH	0.8	0.8	0.8	0.6	1.5
CH_3CH_2CHO	0.3	0.4	0.4	0.2	0.3
CH_3CH_2CN	0.7	0.6	0.6	0.3	0.3
H_2O	0.6	0.2	0.2	0.2	0.3

1: Results obtained with optimized parameters.

2: Results obtained using optimized parameters from ppm-range samples.

Table VI-18. Mean propagated reference concentration error (in ppt) for ppt samples.

Species	Mean ref. error
CH_3CHO	0.19
CH_3CN	0.19
NH_3	0.19
CH_3CH_2OH	0.19
CH_3OH	0.20
CH_3CH_2CHO	0.19
CH_3CH_2CN	0.19
H_2O	0.19

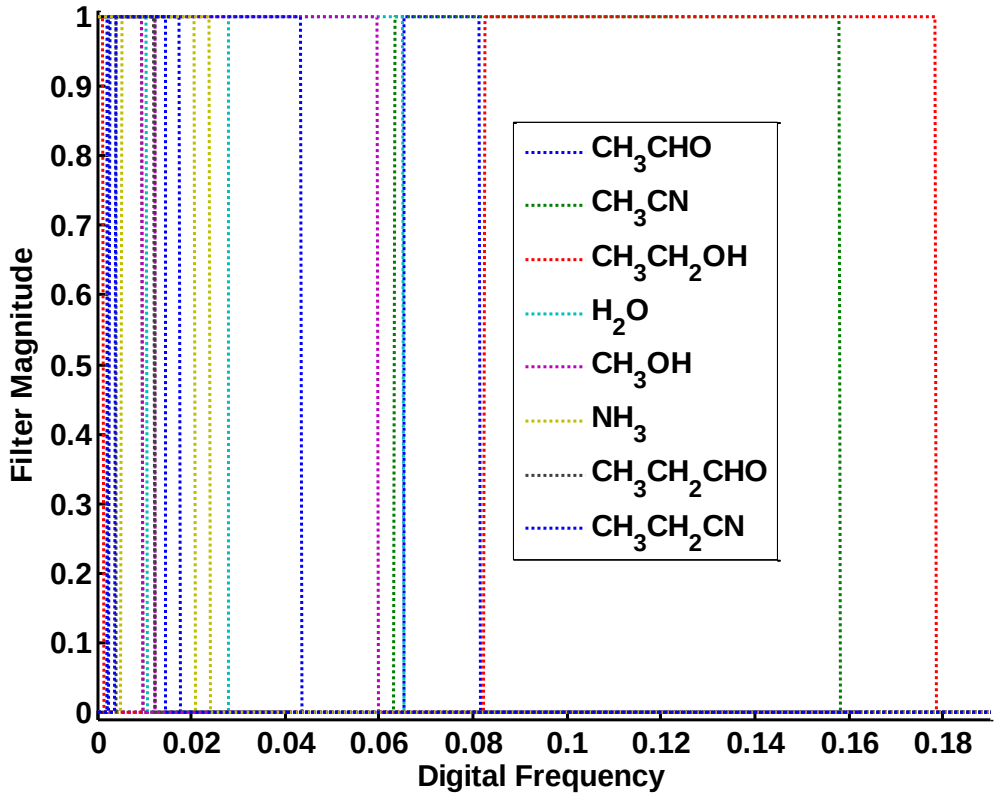


Figure VI-19. Optimized digital frequency filters for the species of interest. Filters generated for ppm-range samples are shown in dotted lines, and those developed for ppt-range samples are shown with solid lines. All filters have magnitudes very near zero at digital frequencies above 0.17.

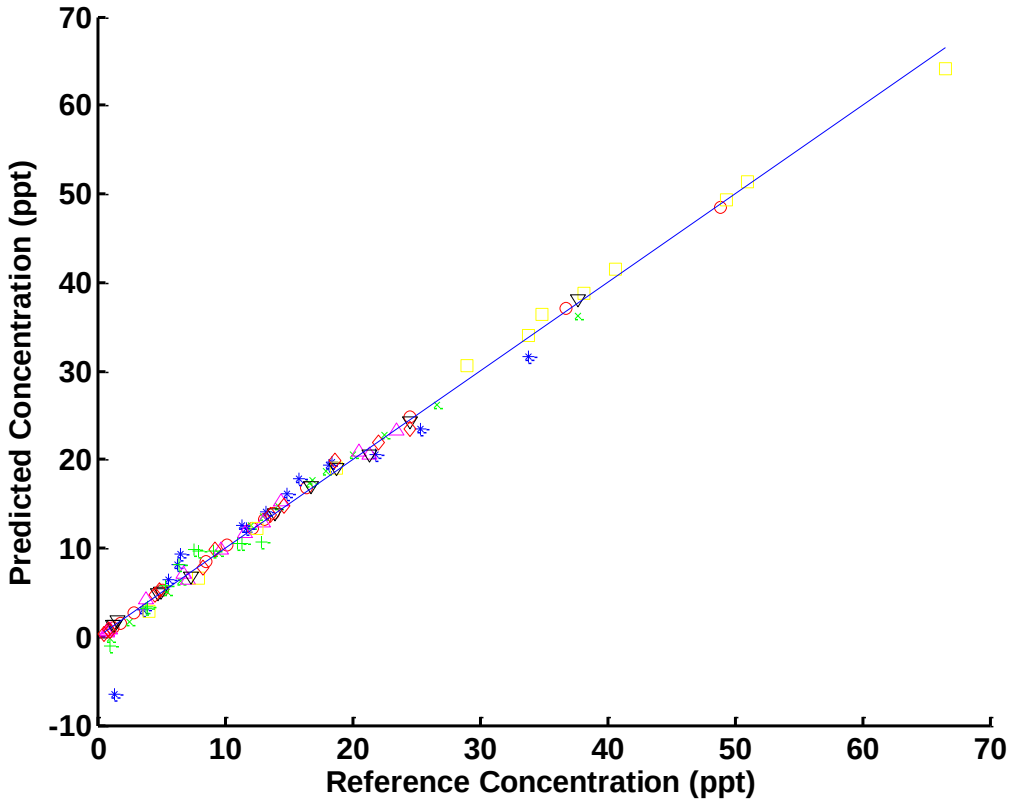


Figure VI-20. Concentration correlation plot from linear regression models using un-processed spectra. Acetaldehyde is represented by red circles, acetonitrile by green x's, ammonia by blue asterisks, ethanol by magenta triangles, methanol by yellow squares, propionaldehyde by inverted black triangles, propionitrile by red diamonds, and water vapor by green +'s. The ideal line is shown as a solid blue line with a slope of 1.

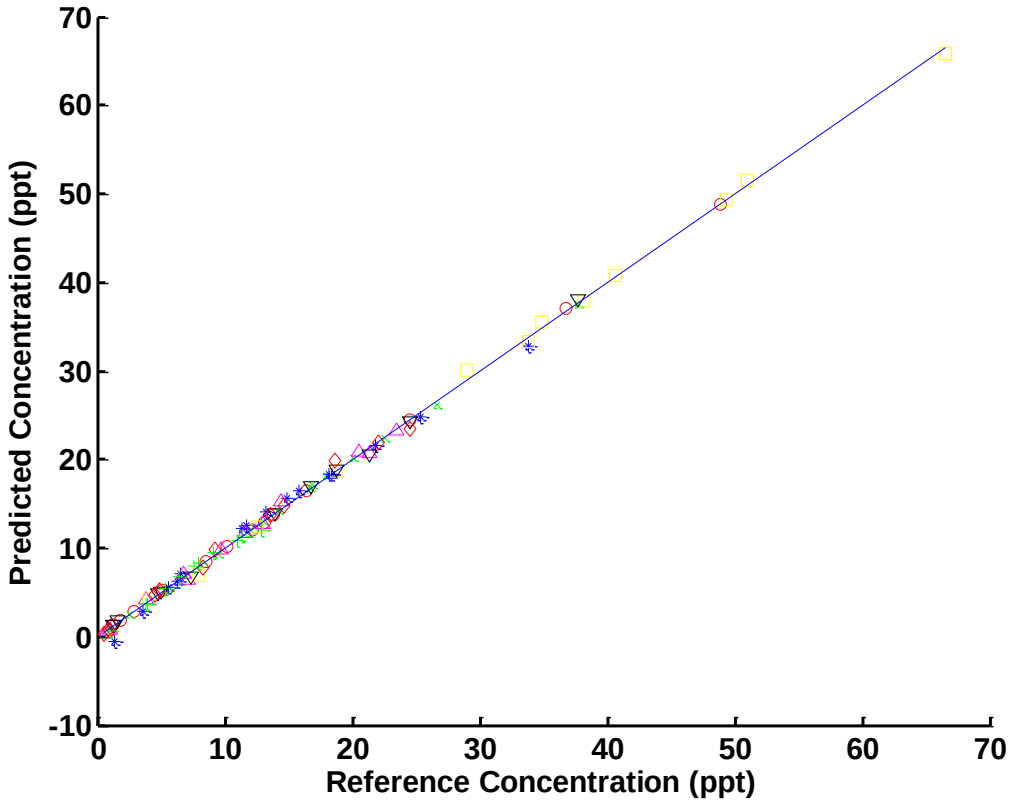


Figure VI-21. Concentration correlation plot from linear regression models using DFF pre-processing. Acetaldehyde is represented by red circles, acetonitrile by green x's, ammonia by blue asterisks, ethanol by magenta triangles, methanol by yellow squares, propionaldehyde by inverted black triangles, propionitrile by red diamonds, and water vapor by green +'s. The ideal line is shown as a solid blue line with a slope of 1.

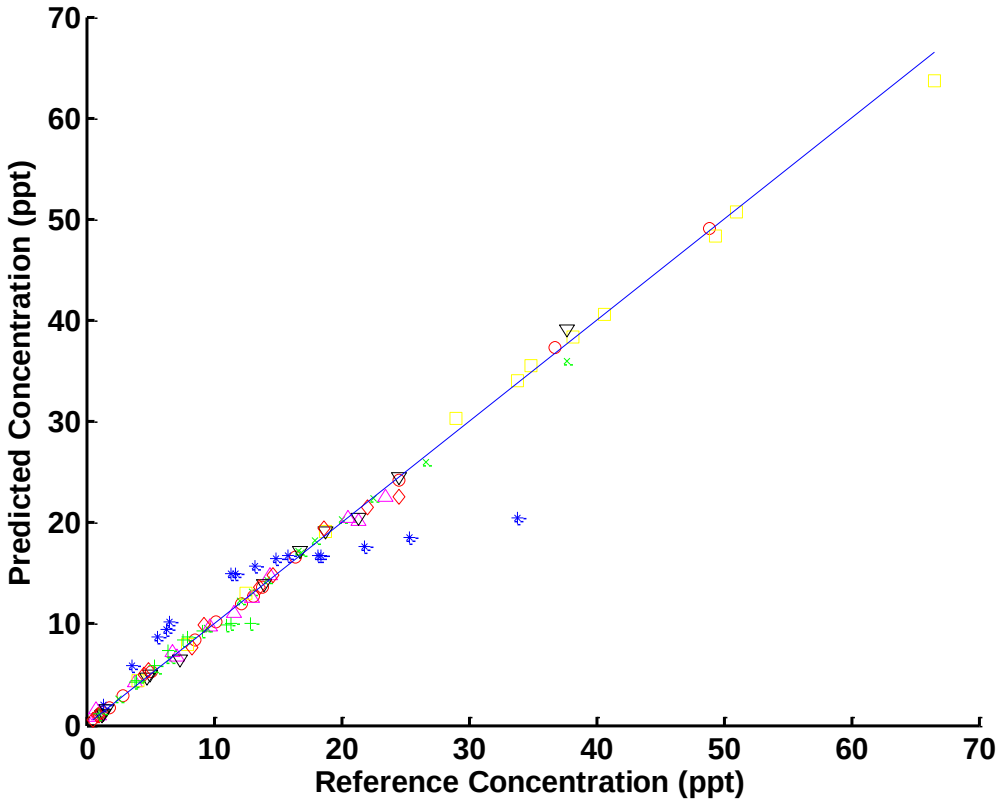


Figure VI-22. Concentration correlation plot from CLS regression models using the full spectral range. Acetaldehyde is represented by red circles, acetonitrile by green x's, ammonia by blue asterisks, ethanol by magenta triangles, methanol by yellow squares, propionaldehyde by inverted black triangles, propionitrile by red diamonds, and water vapor by green +'s. The ideal line is shown as a solid blue line with a slope of 1.

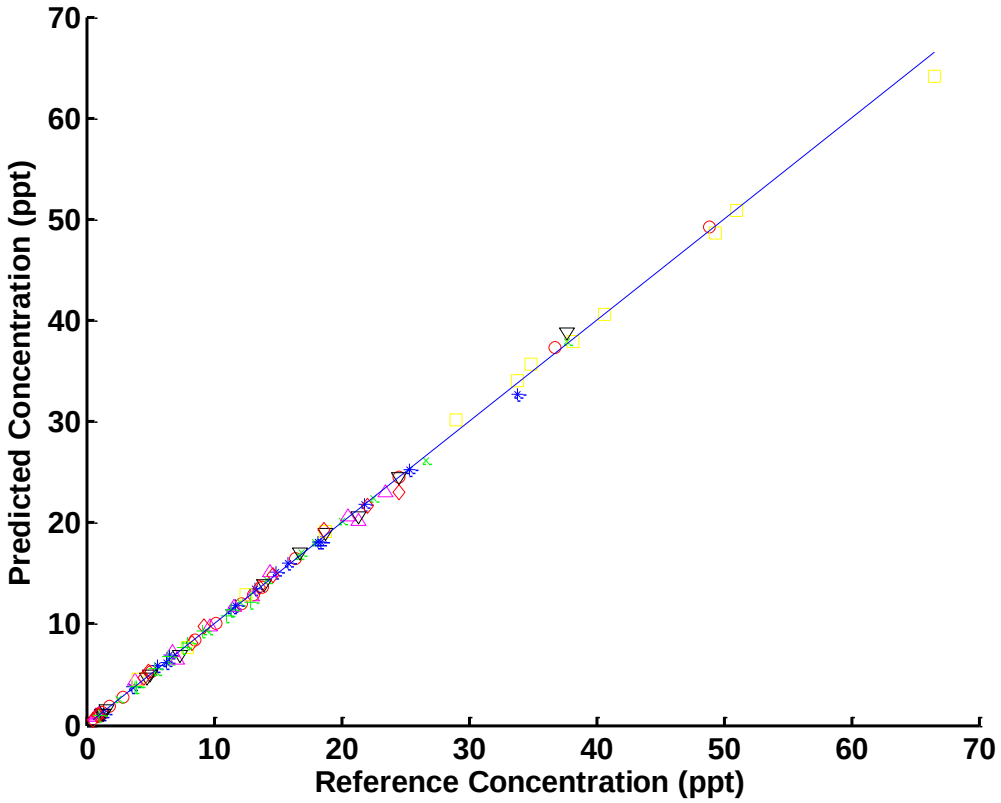


Figure VI-23. Concentration correlation plot from CLS regression models using the optimized spectral ranges. Acetaldehyde is represented by red circles, acetonitrile by green x's, ammonia by blue asterisks, ethanol by magenta triangles, methanol by yellow squares, propionaldehyde by inverted black triangles, propionitrile by red diamonds, and water vapor by green +'s. The ideal line is shown as a solid blue line with a slope of 1.

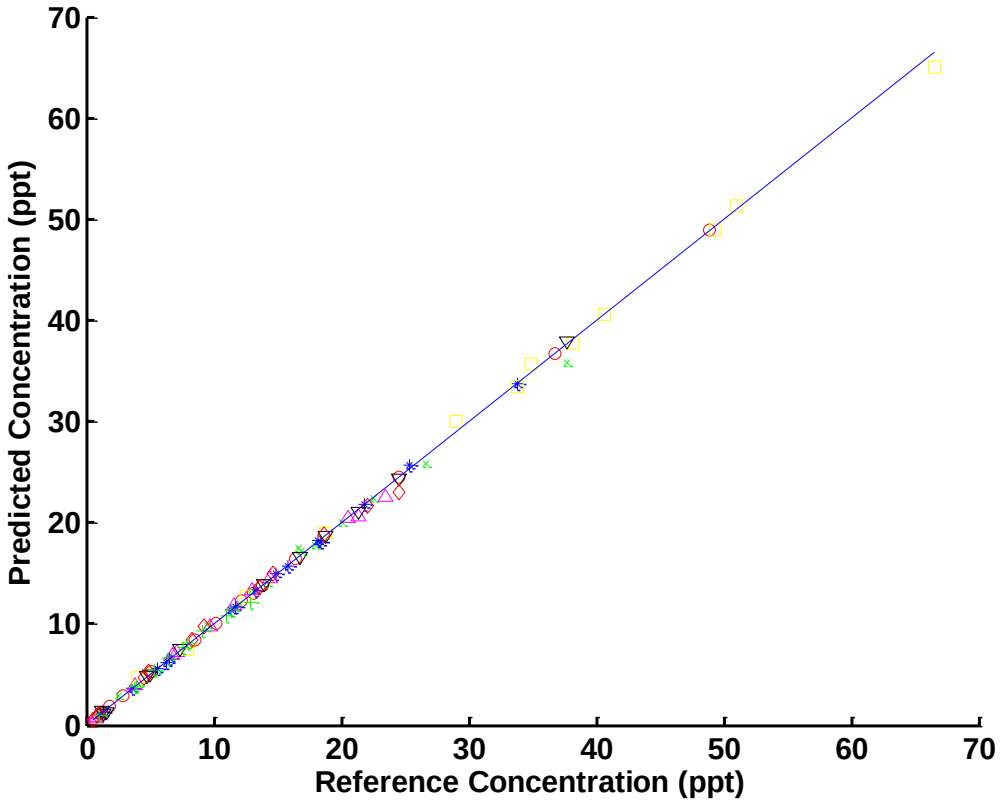


Figure VI-24. Concentration correlation plot from CLS regression models using DFF pre-processing with wavenumber optimization. Acetaldehyde is represented by red circles, acetonitrile by green x's, ammonia by blue asterisks, ethanol by magenta triangles, methanol by yellow squares, propionaldehyde by inverted black triangles, propionitrile by red diamonds, and water vapor by green +'s. The ideal line is shown as a solid blue line with a slope of 1.

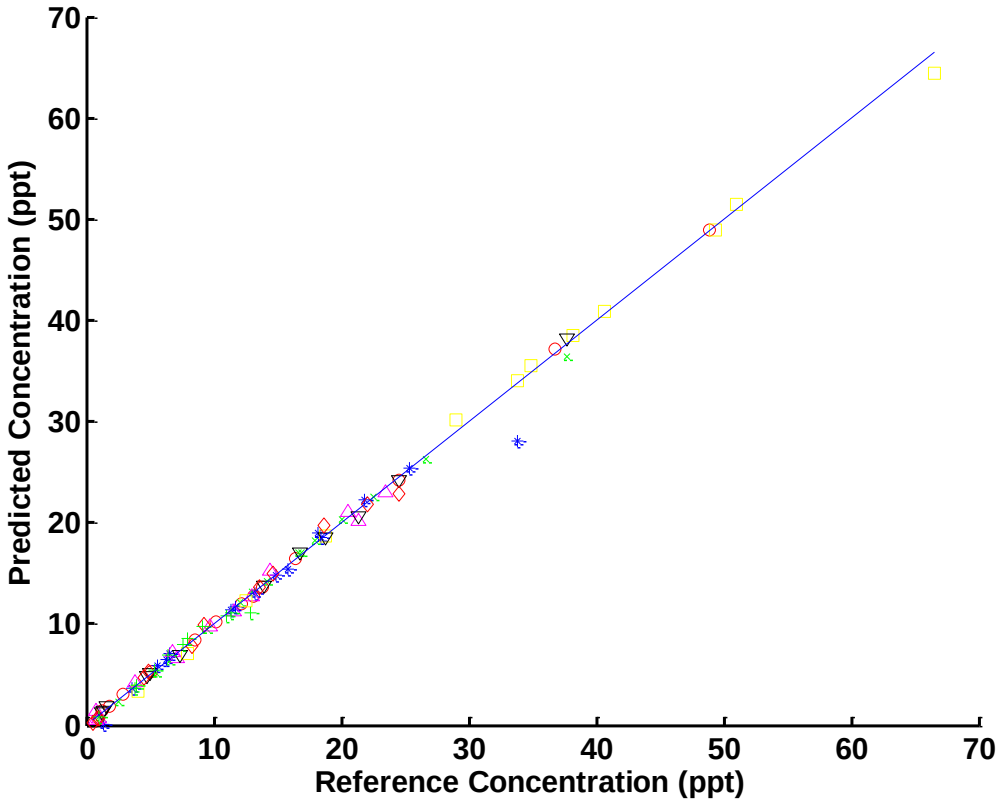


Figure VI-25. Concentration correlation plot from PLS regression models using the full spectral range. Acetaldehyde is represented by red circles, acetonitrile by green x's, ammonia by blue asterisks, ethanol by magenta triangles, methanol by yellow squares, propionaldehyde by inverted black triangles, propionitrile by red diamonds, and water vapor by green +'s. The ideal line is shown as a solid blue line with a slope of 1.

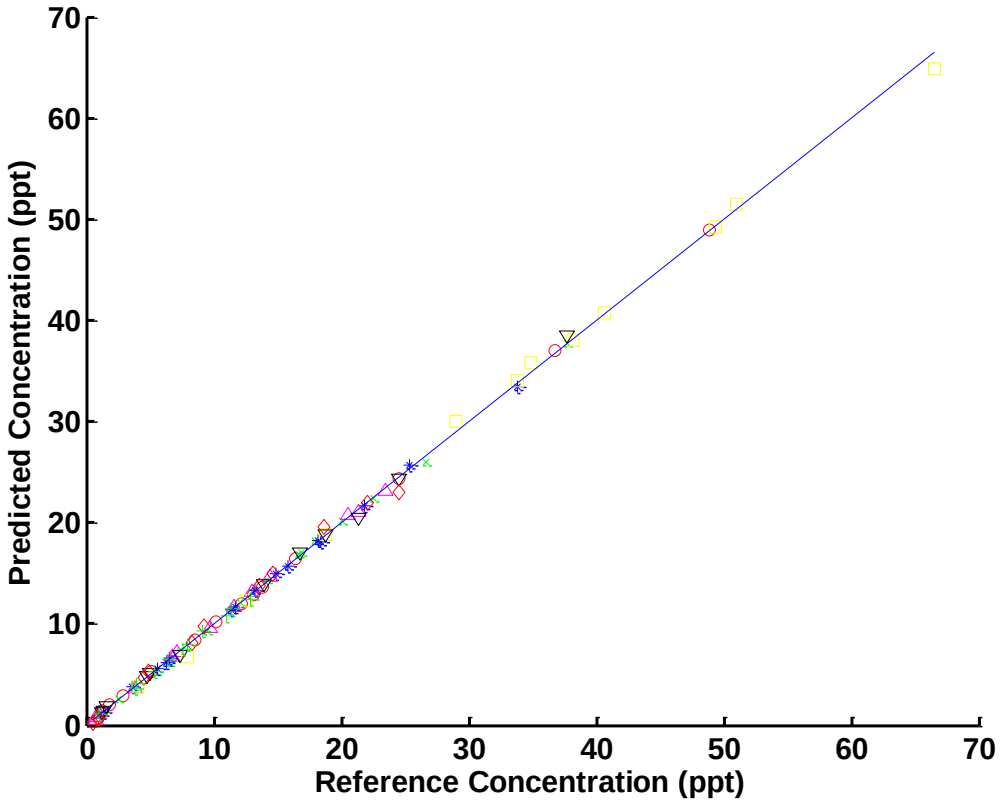


Figure VI-26. Concentration correlation plot from PLS regression models using the optimized spectral ranges. Acetaldehyde is represented by red circles, acetonitrile by green x's, ammonia by blue asterisks, ethanol by magenta triangles, methanol by yellow squares, propionaldehyde by inverted black triangles, propionitrile by red diamonds, and water vapor by green +'s. The ideal line is shown as a solid blue line with a slope of 1.

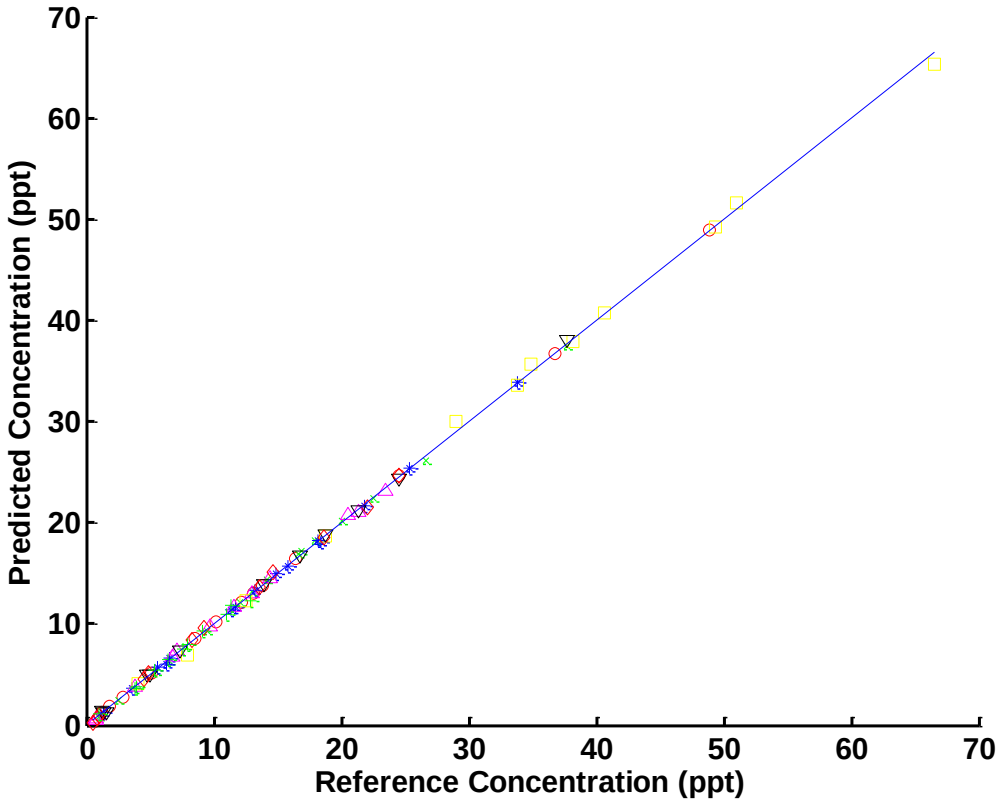


Figure VI-27. Concentration correlation plot from PLS regression models using DFF pre-processing with wavenumber optimization. Acetaldehyde is represented by red circles, acetonitrile by green x's, ammonia by blue asterisks, ethanol by magenta triangles, methanol by yellow squares, propionaldehyde by inverted black triangles, propionitrile by red diamonds, and water vapor by green +'s. The ideal line is shown as a solid blue line with a slope of 1.

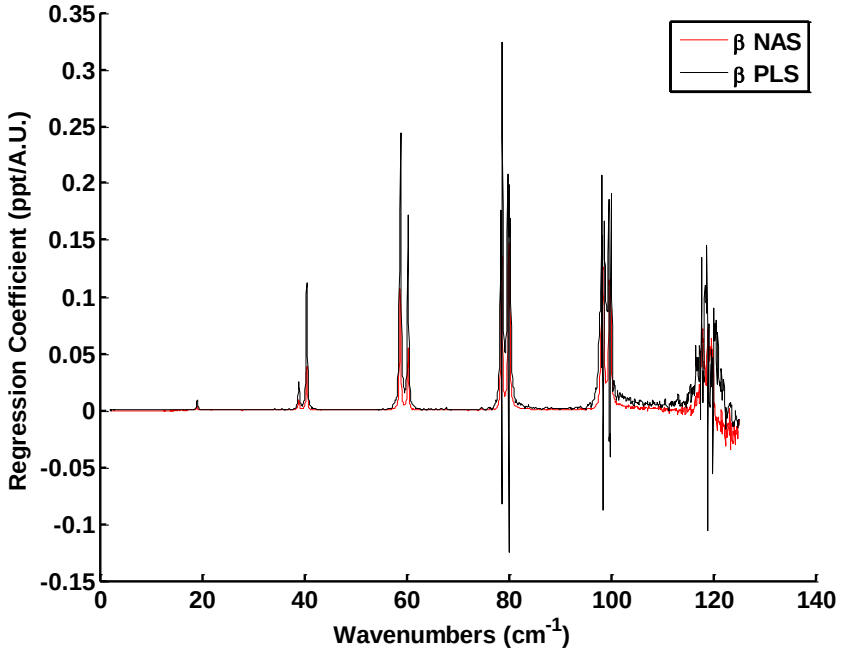


Figure VI-28. NAS and PLS regression vectors constructed with the full spectral range from ammonia vapor spectra.

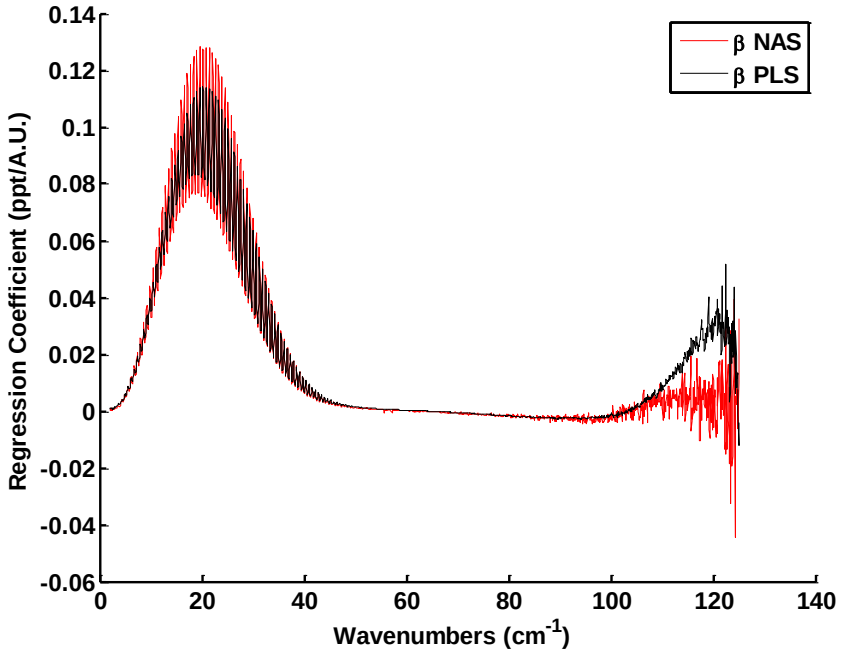


Figure VI-29. NAS and PLS regression vectors constructed with the full spectral range from acetonitrile vapor spectra.

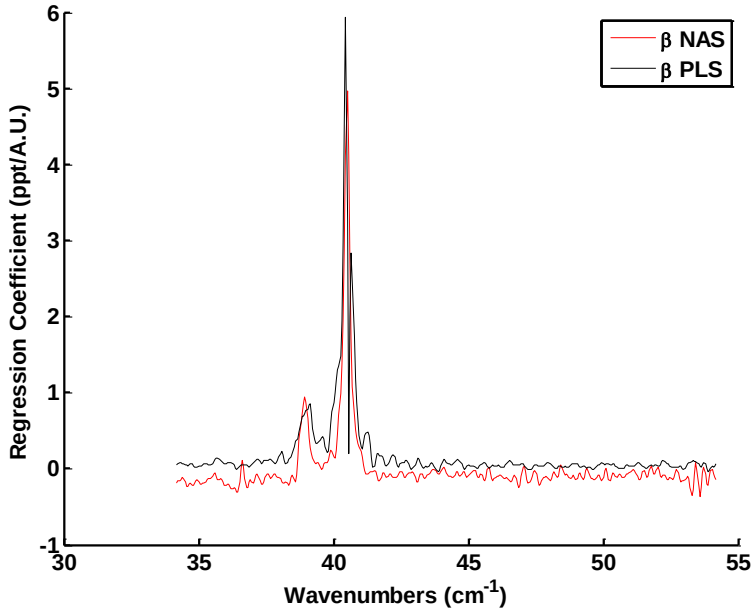


Figure VI-30. NAS and PLS regression vectors constructed with the optimized spectral range from ammonia vapor spectra.

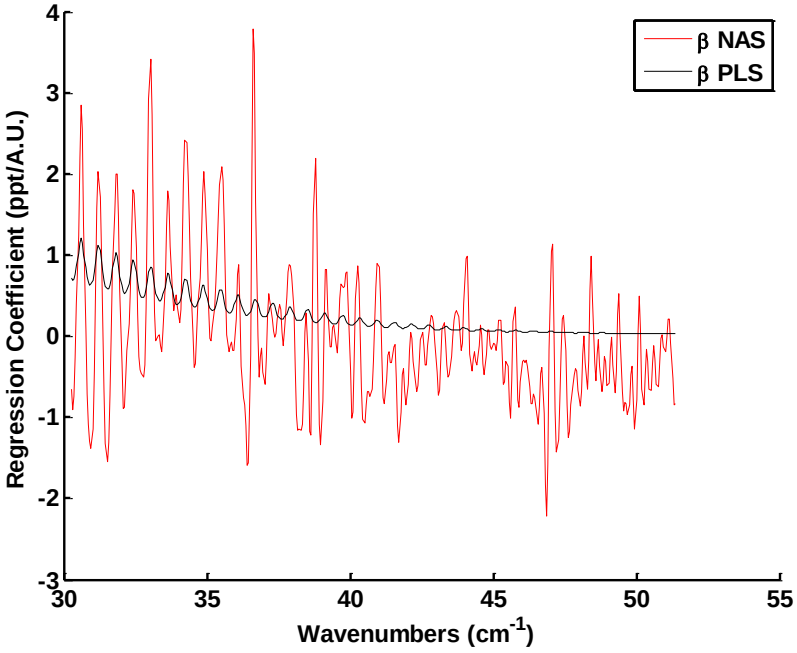


Figure VI-31. NAS and PLS regression vectors constructed with the optimized spectral range from acetonitrile vapor spectra.

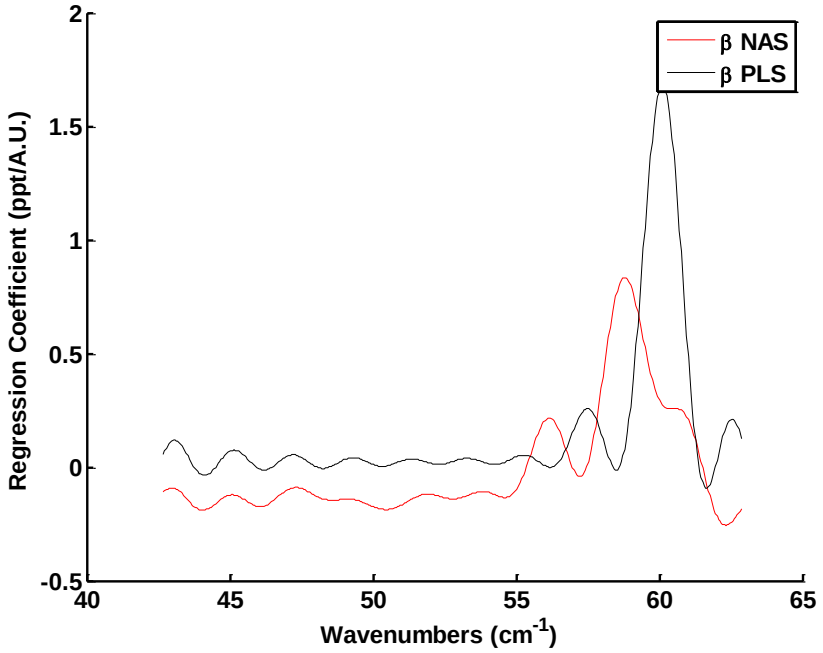


Figure VI-32. NAS and PLS regression vectors constructed with the DFF pre-processed ammonia vapor spectra.

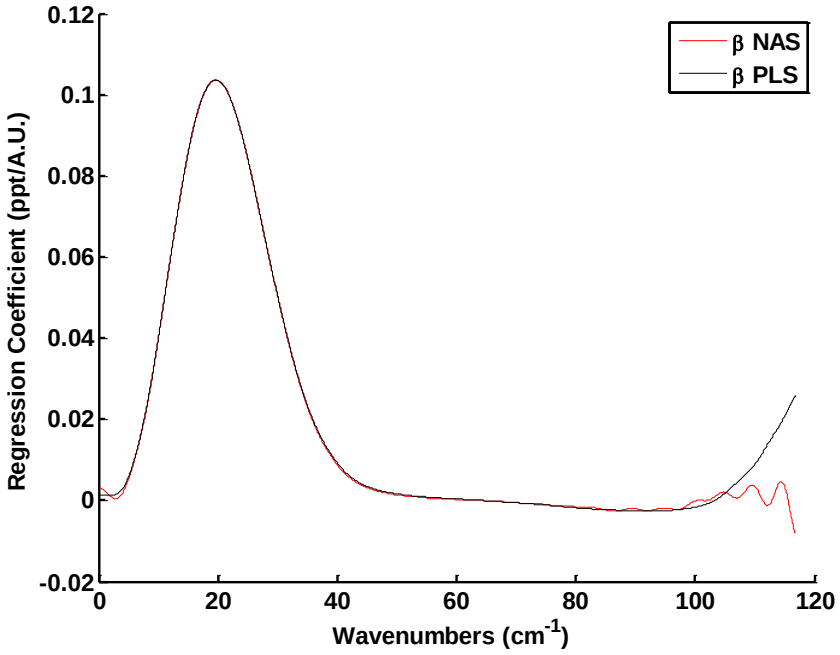


Figure VI-33. NAS and PLS regression vectors constructed with the DFF pre-processed acetonitrile vapor spectra.

Similar trends were observed in some cases when comparing the regression vectors calculated for ppt-range samples with those calculated for ppm-range samples. Regression vectors calculated from the full spectral range are of similar magnitude and shape. When comparing those calculated in the optimized wavenumber region, PLS regression vectors are of much lower magnitude than the corresponding NAS vector for weakly-absorbing species. Vectors calculated for strong absorbers using the optimized spectral range are closer in magnitude. Regression vectors calculated for DFF pre-processed spectra showed poor agreement for strong absorbers, as in the case of the ppm-range samples. Agreement between vectors for weak absorbers varied from one species to another, with some showing excellent agreement (acetonitrile, propionitrile, and acetaldehyde), fair agreement (propionaldehyde, ethanol), or in one case poor agreement (methanol). This is thought to be a result of the significant change in THz absorption from low to high ppt-level concentrations, as optimized filter parameters of some species changed between the ppm and ppt-range sample sets.

Conclusions

Quantitative procedures carried out with the ppt-range samples yield similar results to those analyzed in the ppm concentration range. Errors of prediction are near the propagated reference error for several species of interest, and wavenumber selection and digital filtering provide significant improvement in the predictive performance of quantitative models. Analysis of these samples reveals the breadth of the linear range available with this method: it covers approximately 3 orders of magnitude in concentration. For some samples, the upper limit in the ppt range is the result of physical phenomena rather than instrumental limitations as the highest concentrations are close to the saturation pressure in some cases. The breadth of the linear range is also dependent on the selected frequency range and filter parameters. Thus, for successful implementation of THz-TDS for atmospheric measurements, some general knowledge of

the atmospheric composition may be required such as the order of magnitude of the concentration of the species of interest. Measurement path length may also be altered to achieve a different linear range: short paths will provide good linearity to high concentrations and long path lengths will provide linearity to low concentrations, according to the Beer-Lambert law.

The change of optimal DFF parameters between concentration ranges as well as the disparity between NAS and PLS regression vectors calculated for models developed with DFF pre-processed spectra suggest the need for further examination and development of the filters used for pre-processing when species must be quantified across a large concentration range. Disparity between the vectors suggests a lack of orthogonality in the PLS model, which may have an adverse effect on the method's chemical selectivity. Improved filters may be developed by more thorough optimization procedures such as grid search. Furthermore, the complexity of analyte digital frequency spectra may warrant the use of more complex filters which are not limited to a single pass-band. According to the uncertainty principle illustrated in Chapter II, the frequency domain bandwidth of a short time-domain pulse is correspondingly broad. Thus, for species exhibiting sharp, sparse spectral features (such as ammonia or water), digital filtering of spectra collected with the analytes present at high concentrations may be improved by incorporation of a broad band-pass digital filter. In-depth analysis of the digital frequency spectra may reveal more useful regions which provide improved PLS model orthogonality.

CHAPTER VII

QUANTITATIVE ANALYSIS OF SAMPLE MIXTURES

Eight-Component Sample Mixtures at 1 atm Total Pressure

To this point, the compounds of interest were analyzed individually. While those experiments represent a good initial step in the determination of the feasibility of the proposed measurement, they are far from representative samples. A far more likely scenario is one in which multiple species are present in the sampled volume of gas. The next set of experiments involved the preparation and analysis of samples which consisted of up to all eight of the species of interest with an overall sample pressure of 1 atm. Robustness of the measurement and quantitative models may be seen in the level of precision achieved despite significant spectral overlap between many of the species.

Experimental

Stock mixtures of each analyte were prepared as described in the ppm-range pure component analyses using the apparatus described in Chapter VI. Mixtures of nitrogen and individual species of interest were prepared and retained in the flasks connected to the mixing manifold. Reference spectra were collected with nitrogen at 1 atm as before. During preparation of samples, a portion of each analyte was released into the manifold in a serial-addition process. The propagated error is calculated in a fashion similar to the ppm-range pure components, except that the error for each pressure reading after addition of stock mixture is determined in part by the pressure reading of the previous mixture (if applicable).

Samples were added to the mixture in a particular order. Stock gases were added from propionaldehyde to ammonia following the order of species on the left side of Table VII-1. As such, species with high molecular weights were added first, followed by those of decreasing weights. Addition was performed in this manner to reduce error which might have resulted from effusion of light species into the storage flasks containing

mixtures with species having higher molecular weights. Steps involved in sample preparation are described in the following list.

1. Stock mixtures
    a. evacuate system (0.0 Torr)
    b. insert analyte (various pressures)
    c. add N_2 (~760.0 Torr)
    d. retain mixture in flask
    e. repeat for each species of interest
2. Reference
    a. evacuate system (0.0 Torr)
    b. add N_2 (760.0 Torr)
    c. collect reference spectrum
3. Sample
    a. evacuate system (0.0 Torr)
    b. add stock mixtures (various pressures, repeat for each analyte)
    c. add N_2 (~760.0 Torr)
    d. adjust volume (760.0 Torr)
    e. collect sample spectrum

Reference and sample spectra were acquired in an alternating fashion such that each sample spectrum had its own unique reference spectrum. As with the pure component analyses, sample concentrations were prepared such that correlation with collection sequence was minimized. Furthermore, correlation between concentration of the species of interest was reduced so the features of interfering species would not correlate with the analyte. Table VII-1 provides the R^2 correlation coefficients, where the diagonal elements of the matrix (highlighted in gray) represent the correlation between analyte concentrations and collection sequence, whereas the off-diagonal elements represent the correlation between a given species and the respective interfering compound. Only the lower triangular portion is displayed in this table because the matrix is symmetrical about the diagonal elements and values in blacked-out cells would be identical to those shown in the lower-left half of the matrix. The small values indicate low levels of correlation. Correlation coefficients on the diagonal tend to be larger than those off the diagonal, as preparation of samples without correlation in all cases was

challenging. Though efforts were made to reduce all correlation, higher values were allowed for analyte-sequence correlation as this was expected to have a lesser impact on the predictive ability of the quantitative models.

Spectra were calculated with SCA-processed signals with 4×N signal length after extension. Baseline correction was carried out by subtracting the mean of the 2.72 to 4 cm^{-1} baseline region from the entire spectrum. While there is some absorption in this portion of the spectrum, it is relatively weak, making this region suitable for this purpose. Only the PLS algorithm was used for quantitative analysis, due to its superior performance with pure component samples and inherent analytical benefits. The three methods (full range, optimized range and optimized DFF and wavenumber range) were evaluated for PLS cross-validation with each compound of interest. A total of 55 samples were prepared and analyzed in the temperature range of 23.9 ± 0.2 °C at final pressures of 760.3 ± 1.1 Torr. Concentration ranges for each analyte are provided in Table VII-2.

Results

Spectral data acquired with these samples is displayed in Figure VII-1. Overlap between species prevents observation of distinct features attributable to most species. Features of water vapor, acetonitrile, and acetaldehyde are perhaps the most apparent. Spectral noise dominates the extremities of the plotted frequency range. Parameters selected for each PLS regression model are listed in Table VII-3, and the resultant CVSEP values are provided in Table VII-4. Optimized parameters vary from those of the pure components, but still center on spectral regions containing the strongest spectral features. The number of latent variables required for each regression model has increased as is expected for complex mixtures, and the number used for each species exhibits some correlation with the sensitivity for each analyte when considering the parameters chosen for wavenumber-optimized models. Weakly-absorbing species tend to require additional

Table VII-1. Correlation between reference concentrations for analyte-interferent pairs and analyte-sequence pairs. Values on the diagonal (highlighted in gray) represent analyte-sequence correlation, whereas values off the diagonal represent analyte-interference correlation.

R^2	CH_3CH_2CHO	CH_3CH_2CN	CH_3CH_2OH	CH_3CHO	CH_3CN	CH_3OH	H_2O	NH_3
CH_3CH_2CHO	1.8E-03							
CH_3CH_2CN	1.5E-03	1.9E-04						
CH_3CH_2OH	1.5E-02	6.4E-04	5.0E-04					
CH_3CHO	3.2E-03	7.7E-05	8.5E-05	2.1E-03				
CH_3CN	6.5E-05	5.0E-05	1.1E-04	1.1E-06	4.3E-02			
CH_3OH	8.1E-05	5.2E-06	8.3E-06	5.9E-05	1.2E-05	1.5E-01		
H_2O	9.1E-03	5.5E-05	4.4E-06	6.3E-06	6.2E-06	4.4E-05	1.2E-01	
NH_3	9.0E-04	1.6E-06	1.4E-04	7.9E-07	5.3E-05	3.7E-06	7.0E-05	7.8E-03

latent variables to produce improved quantitative models. This trend is no longer visible in the DFF-processed spectra, which is a result of the elimination of non-analyte variance by the filter.

The error of prediction was calculated for all PLS models constructed with varying numbers of latent variables, shown in Figures VII-2 through VII-4. Significant latent variables were determined using the F-test as described previously. These traces not only provide information related to the significance of each latent variable, but also provide a general idea of the relative magnitudes of non-analyte and analyte-related variance in the models. A slow decrease in the CVSEP is observed in instances where the analyte variance is relatively small compared to the non-analyte variance, a trend observed in Figure VII-2. A great deal of instrumental noise and features of interfering species are included in the PLS model when utilizing the spectral range. Conversely, the CVSEP approaches a minimum value with inclusion of fewer latent variables when the wavenumber range is optimized or when DFF pre-processing is implemented, since features of the analyte become more prominent compared to other sources of spectral variance.

Errors of prediction were below the PELs for all species, and wavenumber optimization consistently provided significant improvement in the precision of the quantitative model when compared to models constructed using the full spectral range. Likewise, DFF pre-processing (combined with wavenumber optimization) reduced the variability of the regressions when compared to models utilizing the entire spectrum. Pre-processing with DFF did not consistently provide significant improvement when compared to models using only wavenumber optimization; increased precision was only observed in quantitative models constructed for two of the weakest absorbers: ethanol and propionaldehyde. The errors of prediction obtained with PLS regression models using the full spectral range are nearly equal to the standard deviation of the reference concentrations, indicating minimal predictive ability.

Analyte concentrations predicted by the various PLS models are plotted as a function of the reference concentration in Figures VII-5 through VII-7. Each of the 55 samples is represented in these plots by 8 points: 1 for each analyte in the sample mixture. Scatter of the points about the ideal line is indicative of the precision of the regression model, where greater scatter represents poorer precision. The ideal line itself is a line with a slope of 1, representing a situation in which the predicted concentration of the analyte in a given sample is exactly equal to the reference concentration of the analyte. CVSEP values are related to this scatter; high CVSEP values will result from high scatter and vice versa. A high degree of scatter about the ideal line can be seen in the concentration correlation plot for regressions constructed using the full spectral range, as expected based on the CVSEP values listed in the first column of Table VII-4. Far less scatter is observed in the concentration correlation plots from PLS models constructed with optimized wavenumber bounds and DFF parameters, which is reflected by the correspondingly-low errors of prediction listed in the second and third columns of Table VII-4.

The performance of these quantitative models was also investigated by analysis of the residuals of prediction, shown in Figures VII-5 through VII-10. Residuals calculated from the PLS models using the full spectral range show a fairly significant linear trend. The magnitude of the residuals is very large, spanning a range of about -300 ppm to 400 ppm. Conversely, the prediction residuals calculated with PLS models utilizing the optimized spectral range and digital filtering show no significant trends. Moreover, the magnitude of these residuals is diminished, providing evidence of improved quantitative precision. Figures VII-8 through VII-10 were constructed with residuals of analyte species plotted as a function of the respective reference concentrations. Correlation between the residuals of the analyte and the concentration of an interfering species can provide further information about model selectivity or reactivity between species.

Correlation coefficients are shown for each of the three PLS models in Tables VII-6 through VII-8.

Table VII-2. Concentration statistics for each compound of interest in the 8-component sample mixtures.

Species	C_{mean} (ppm)	C_{min} (ppm)	C_{max} (ppm)
CH_3CH_2CHO	289 ± 183	0	760
CH_3CH_2CN	244 ± 145	0	545
CH_3CH_2OH	207 ±152	0	558
CH_3CHO	243 ± 135	0	485
CH_3CN	243 ± 142	0	498
CH_3OH	262 ± 167	0	580
H_2O	303 ± 160	49	659
NH_3	213 ± 135	0	536

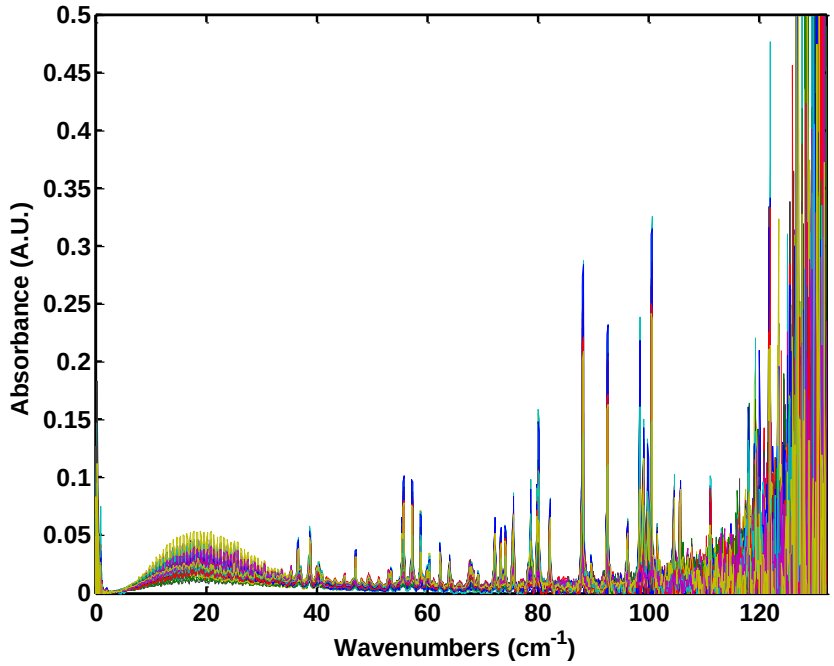


Figure VII-1. Spectra collected with ppm-range mixtures of the 8 species of interest.

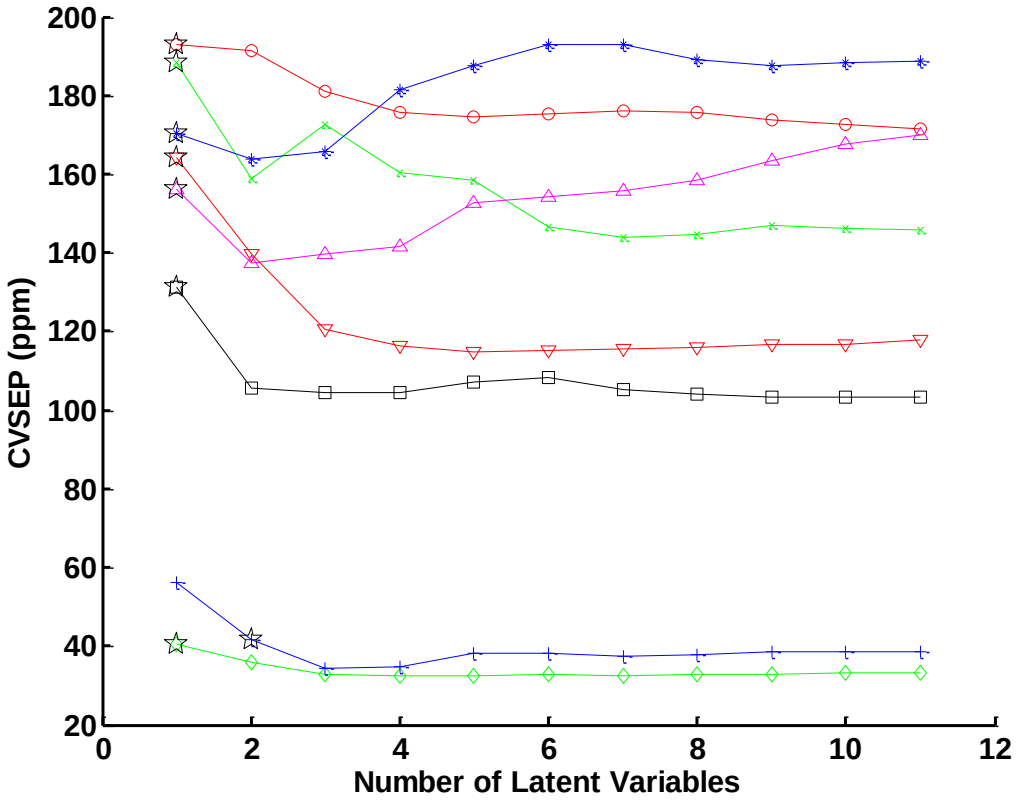


Figure VII-2. Errors of prediction calculated with PLS models constructed with different numbers of latent variables using the full spectral range. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, and ammonia by blue +'s. The appropriate number of latent variables is indicated by a star for each analyte.

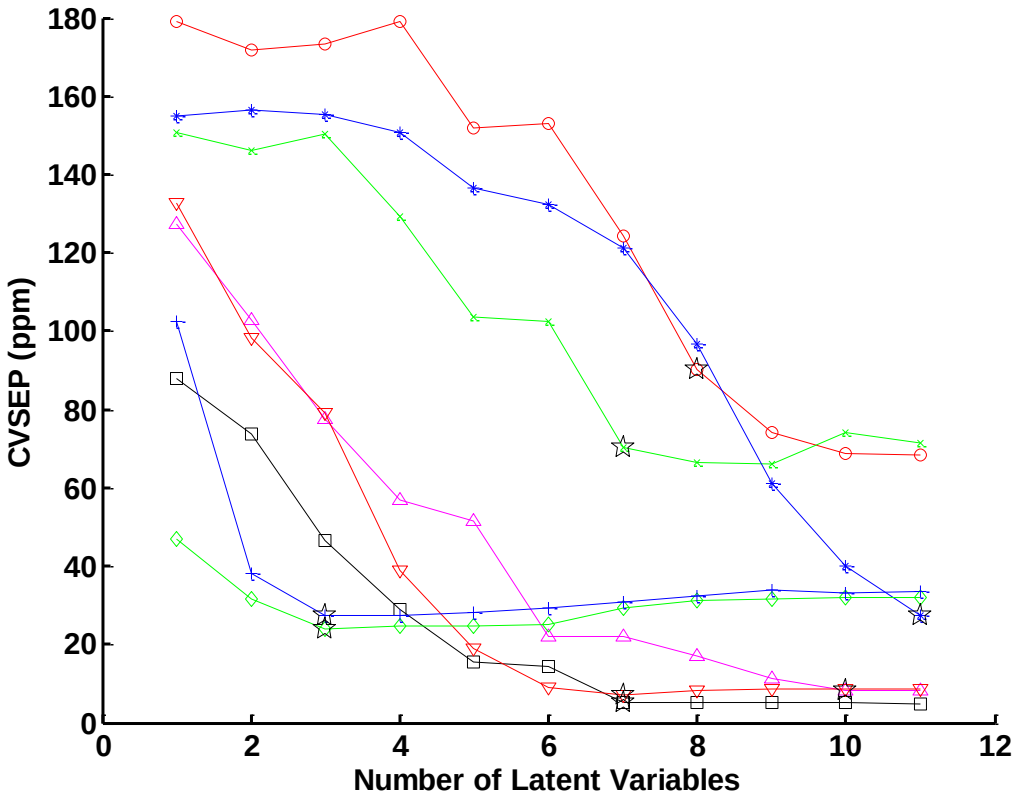


Figure VII-3. Errors of prediction calculated with PLS models constructed with different numbers of latent variables using the optimized spectral range. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, and ammonia by blue +'s. The appropriate number of latent variables is indicated by a star for each analyte.

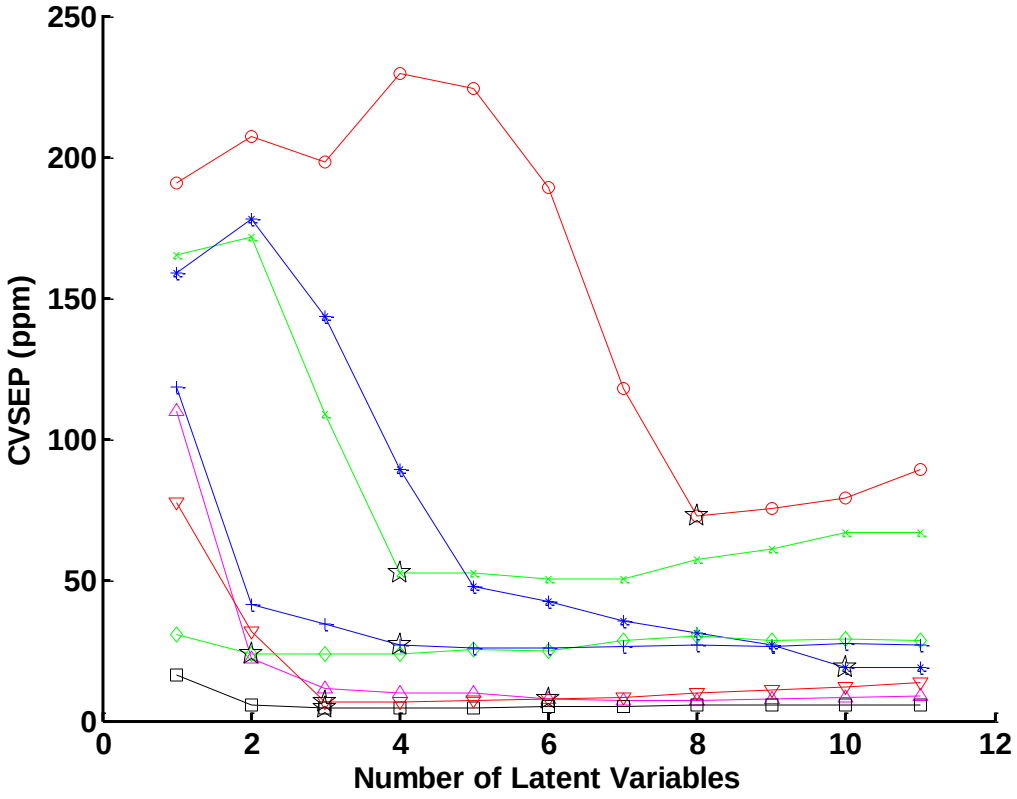


Figure VII-4. Errors of prediction calculated with PLS models constructed with different numbers of latent variables using the optimized spectral range and DFF parameters. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, and ammonia by blue +'s. The appropriate number of latent variables is indicated by a star for each analyte.

Table VII-3. Parameters used for PLS cross-validation with 8-component mixtures. Optimized values are provided for each calibration method.

Species	RAW	WN			WN, DFF				
	# LV[1]	WN_{low}[2]	WN_{high}[2]	# LV[1]	Filter μ[3]	Filter width[3]	WN_{low}[2]	WN_{high}[2]	# LV[1]
CH_3CH_2CHO	1	10.0	30.0	8	0.0749	0.0278	10.9	32.8	8
CH_3CH_2CN	1	20.4	75.1	7	0.1325	0.0496	10.8	42.4	4
CH_3CH_2OH	1	15.0	45.0	11	0.1007	0.0360	10.0	30.0	10
CH_3CHO	1	4.9	41.7	10	0.0931	0.0361	5.3	25.3	6
CH_3CN	1	5.0	70.0	7	0.2488	0.0933	9.6	58.6	3
CH_3OH	1	24.3	66.0	7	0.0506	0.0287	18.8	63.6	3
H_2O	1	22.7	61.8	3	0.0707	0.0549	18.6	60.3	2
NH_3	2	22.9	73.3	3	0.4474	0.2010	22.7	61.8	4

1: Number of latent variables.

2: Wavenumber units (cm^{-1}).

3: Digital frequency units.

Table VII-4. Errors of prediction (in ppm) from PLS regression models using different processing methods.

Species	Raw	WN	WN, DFF	PEL
CH_3CH_2CHO	193	90	73	NA
CH_3CH_2CN	188	70	52	NA
CH_3CH_2OH	170	27	19	1000
CH_3CHO	156	8	8	200
CH_3CN	131	5	4	40
CH_3OH	164	7	7	200
H_2O	41	24	24	NA
NH_3	42	27	27	50

Table VII-5. Calculated F-values representing the significance of the processing and optimized methods. The critical F value is 1.57, and values exceeding this threshold are highlighted in yellow.

Species	F_{calc} (RAW-WN)	F_{calc} (DFF-RAW)	F_{calc} (DFF-WN)
CH_3CH_2CHO	4.59	7.04	1.53
CH_3CH_2CN	7.15	12.99	1.82
CH_3CH_2OH	39.25	83.48	2.13
CH_3CHO	365.35	412.39	1.13
CH_3CN	683.31	907.06	1.33
CH_3OH	547.76	621.72	1.14
H_2O	2.86	2.98	1.04
NH_3	2.28	2.35	1.03

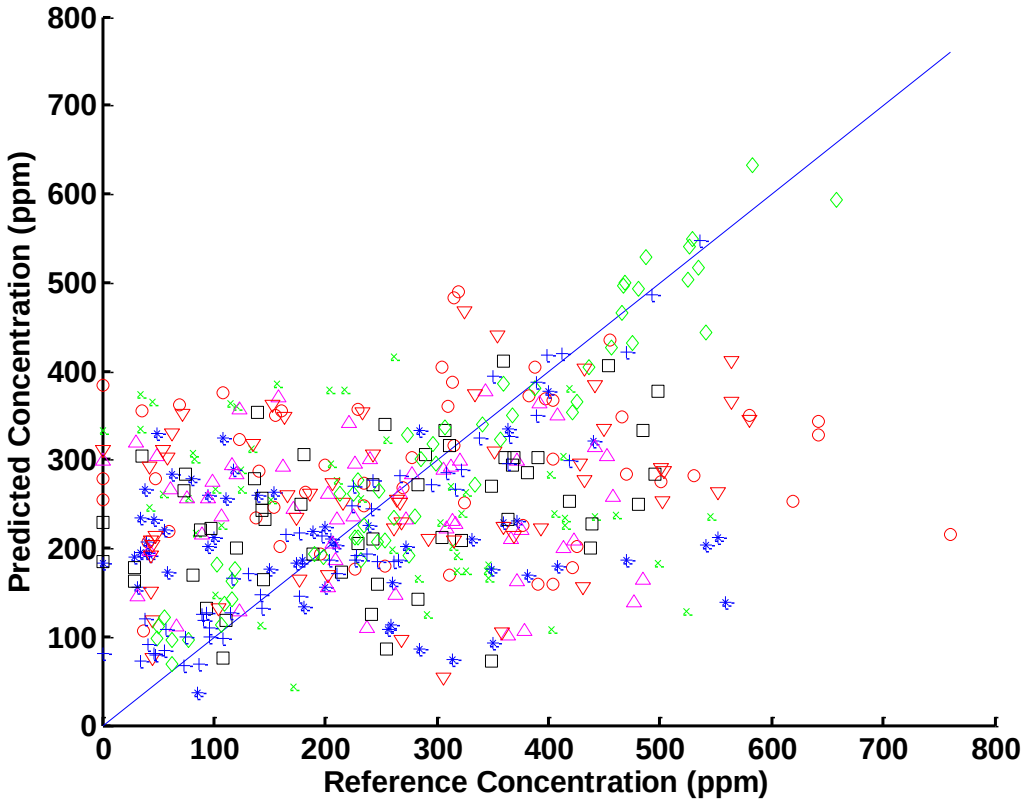


Figure VII-5. Concentration correlation plot from PLS prediction of species in 8-component mixtures at 1 atm overall pressure using optimized wavenumber bounds. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, and ammonia by blue +'s. Poor precision is evidenced by extreme scatter about the ideal line (shown as a solid blue line).

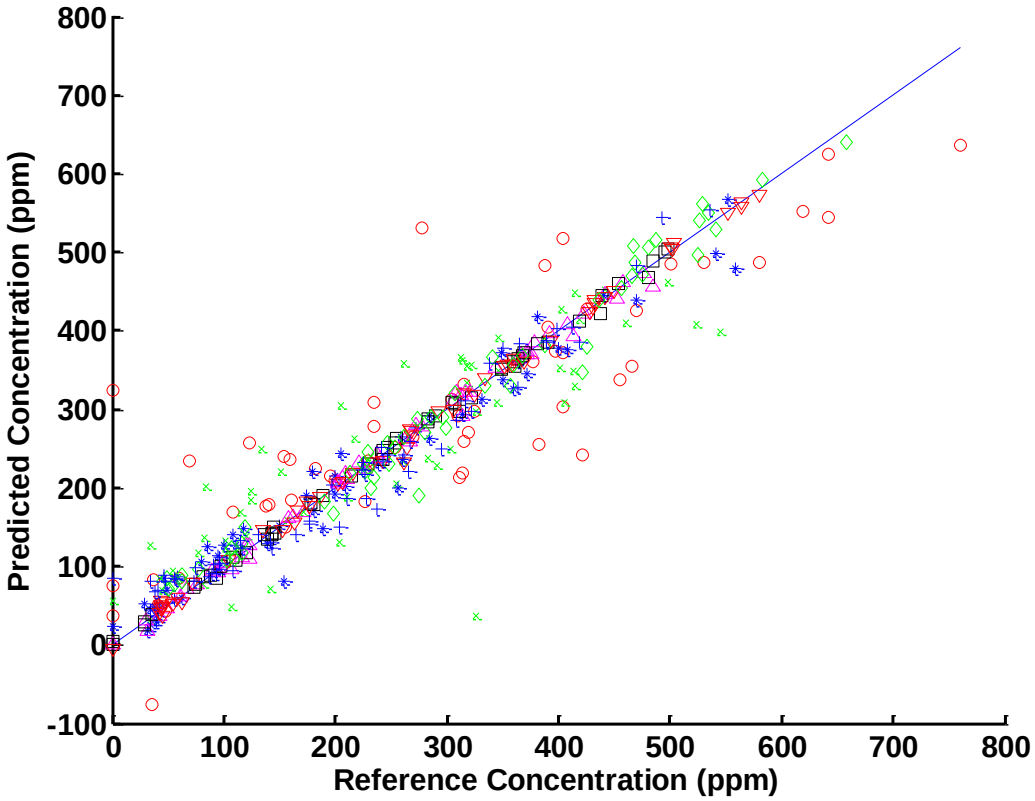


Figure VII-6. Concentration correlation plot from PLS prediction of species in 8-component mixtures at 1 atm overall pressure using optimized wavenumber bounds. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, and ammonia by blue +'s. Improved precision is evidenced by decreased scatter about the ideal line (shown as a solid blue line).

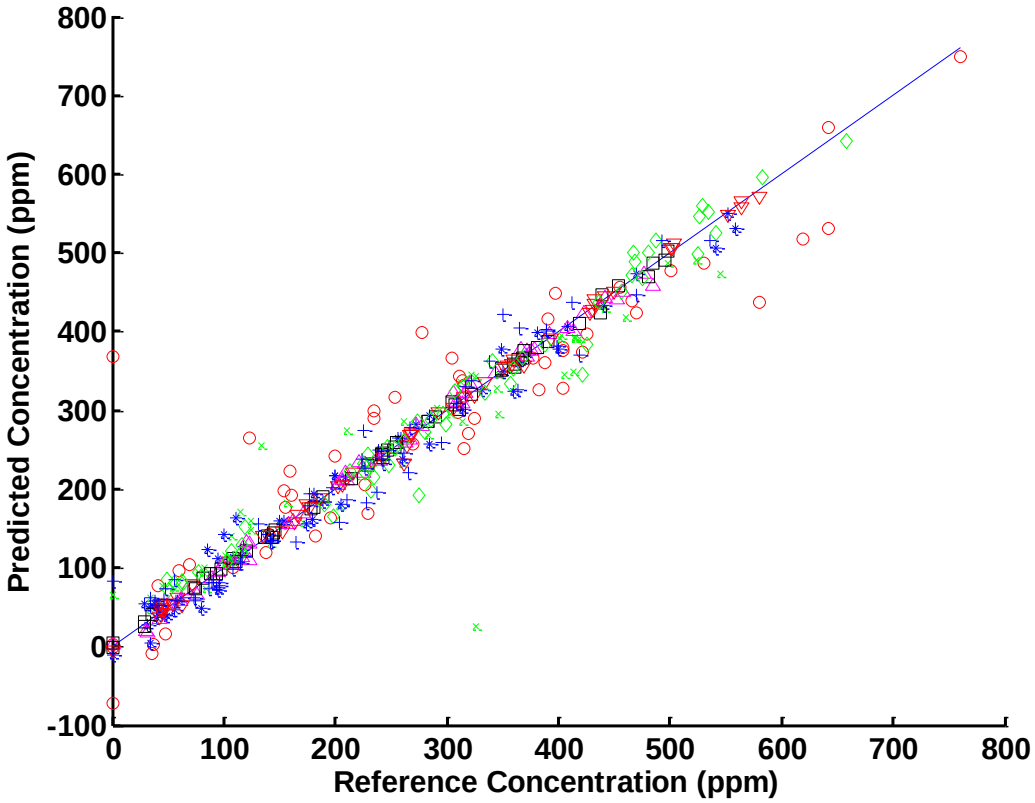


Figure VII-7. Concentration correlation plot from PLS prediction of species in 8-component mixtures at 1 atm overall pressure using optimized DFF parameters and wavenumber bounds. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, and ammonia by blue +'s. Improved precision is evidenced by decreased scatter about the ideal line (shown as a solid blue line).

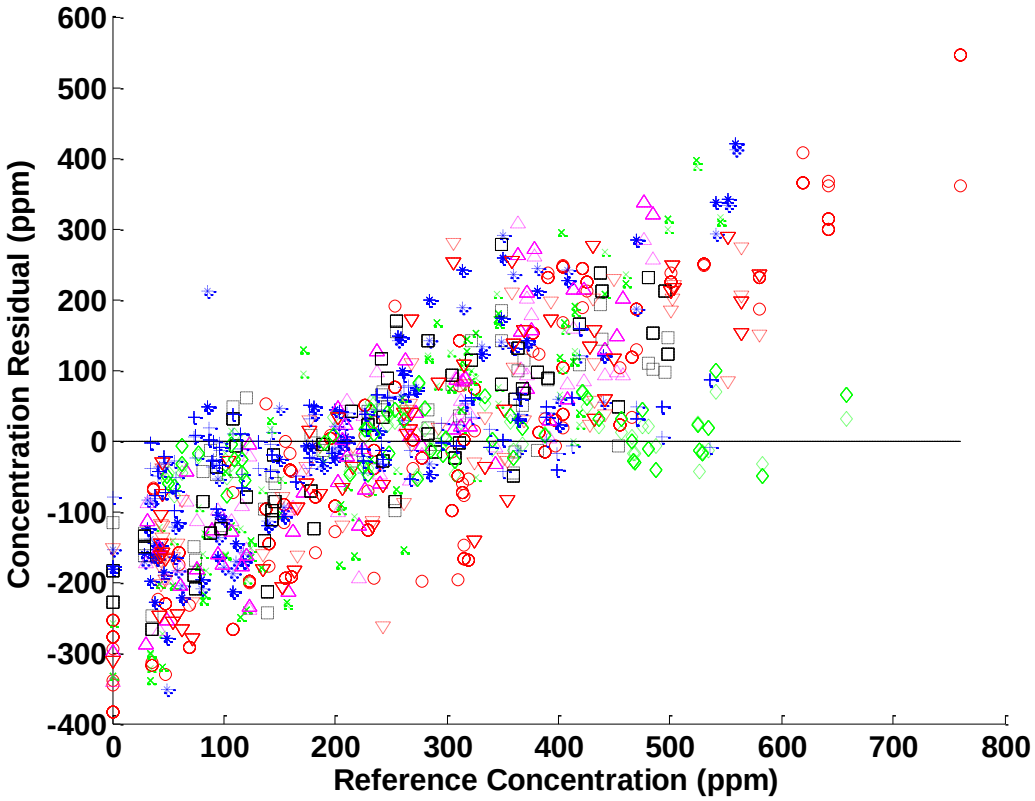


Figure VII-8. Residuals of predicted concentrations as a function of the reference concentration from PLS regression models constructed with the full spectral range. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, and ammonia by blue +'s. Zero is indicated by a black dotted line.

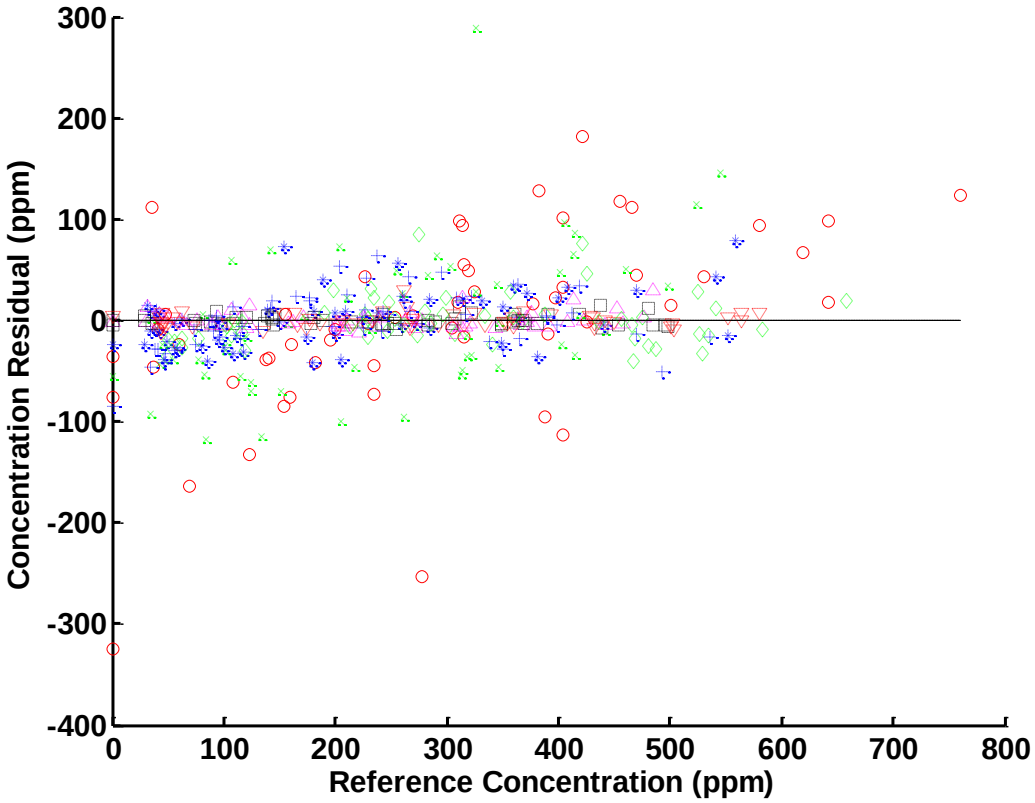


Figure VII-9. Residuals of predicted concentrations as a function of the reference concentration from PLS regression models constructed with the optimized spectral range. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, and ammonia by blue +'s. Zero is indicated by a black dotted line.

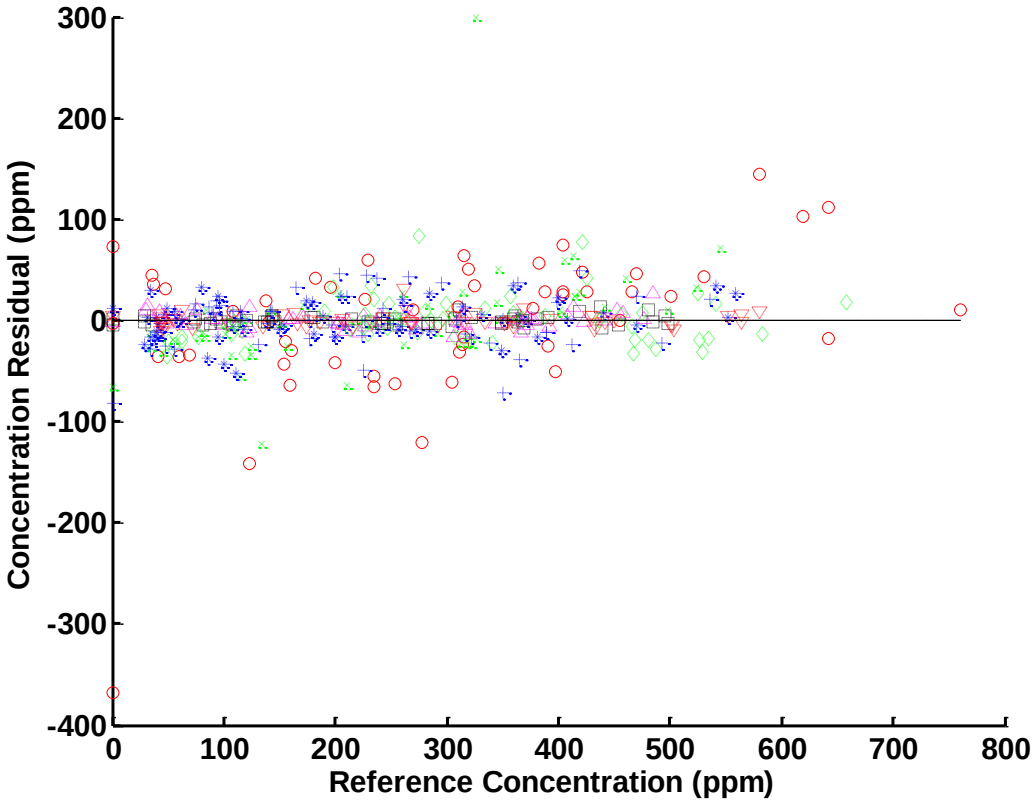


Figure VII-10. Residuals of predicted concentrations as a function of the reference concentration from PLS regression models constructed with the optimized spectral range and DFF pre-processing. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, and ammonia by blue +'s. Zero is indicated by a black dotted line.

Table VII-6. Correlation between analyte concentration and prediction residuals from PLS quantitative models utilizing the full spectral range.

		Residual Species							
	Species	CH_3CH_2CHO	CH_3CH_2CN	CH_3CH_2OH	CH_3CHO	CH_3CN	CH_3OH	H_2O	NH_3
Analyte Species	CH_3CH_2CHO	0.81	0.00	0.02	0.02	0.04	0.01	0.00	0.02
	CH_3CH_2CN	0.00	0.84	0.00	0.01	0.03	0.00	0.00	0.03
	CH_3CH_2OH	0.02	0.00	0.84	0.00	0.00	0.00	0.01	0.01
	CH_3CHO	0.01	0.01	0.00	0.79	0.04	0.01	0.01	0.09
	CH_3CN	0.02	0.05	0.00	0.04	0.70	0.00	0.02	0.05
	CH_3OH	0.00	0.00	0.00	0.01	0.00	0.72	0.12	0.02
	H_2O	0.07	0.00	0.02	0.01	0.03	0.19	0.18	0.00
	NH_3	0.00	0.05	0.00	0.06	0.05	0.00	0.13	0.30

Table VII-7. Correlation between analyte concentration and prediction residuals from PLS quantitative models utilizing the optimized spectral range.

		Residual Species							
	Species	CH_3CH_2CHO	CH_3CH_2CN	CH_3CH_2OH	CH_3CHO	CH_3CN	CH_3OH	H_2O	NH_3
Analyte Species	CH_3CH_2CHO	0.29	0.36	0.04	0.00	0.00	0.00	0.01	0.00
	CH_3CH_2CN	0.10	0.25	0.01	0.00	0.00	0.00	0.01	0.02
	CH_3CH_2OH	0.00	0.09	0.22	0.02	0.00	0.00	0.00	0.01
	CH_3CHO	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.02
	CH_3CN	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
	CH_3OH	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
	H_2O	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00
	NH_3	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.09

Table VII-8. Correlation between analyte concentration and prediction residuals from PLS quantitative models utilizing the optimized spectral range and DFF pre-processing.

		Residual Species							
	Species	CH_3CH_2CHO	CH_3CH_2CN	CH_3CH_2OH	CH_3CHO	CH_3CN	CH_3OH	H_2O	NH_3
Analyte Species	CH_3CH_2CHO	0.17	0.08	0.00	0.01	0.04	0.00	0.01	0.01
	CH_3CH_2CN	0.07	0.21	0.01	0.00	0.01	0.00	0.01	0.11
	CH_3CH_2OH	0.00	0.00	0.06	0.01	0.01	0.06	0.00	0.02
	CH_3CHO	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.05
	CH_3CN	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00
	CH_3OH	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.02
	H_2O	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00
	NH_3	0.00	0.02	0.00	0.03	0.00	0.00	0.00	0.01

Correlation coefficients calculated from PLS models constructed with the full spectral range reflect the strong linear trends observed in Figure VII-8, in that the values on the diagonal of the matrix have values close to 1. Correlation coefficients in off-diagonal positions are extremely small, indicating minimal correlation between the reference concentrations of the analytes and the residuals of prediction for the interfering species.

Coefficients on the diagonals of Tables VII-7 and VII-8 are much lower, providing evidence of the PLS regression models' precision and selectivity. Correlation is minimal between most interfering species in these models with the exception of propionaldehyde and propionitrile ($R^2 = 0.36$). These two species have similar spectral features, and both are very weak absorbers, likely causing the observed correlation. This relationship is diminished considerably by DFF pre-processing ($R^2 = 0.08$). Further diagnosis of the quality of the quantitative models may be performed by comparison of PLS and NAS regression vectors from models developed with each method. Representative regression vectors are shown in Figures VII-11 through VII-16.

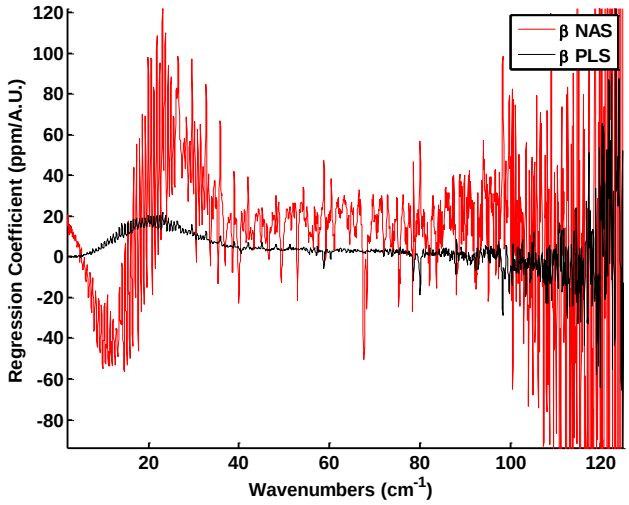


Figure VII-11. NAS and PLS regression vectors constructed using the full spectral range for acetaldehyde in 8-component sample mixtures at 1 atm overall pressure.

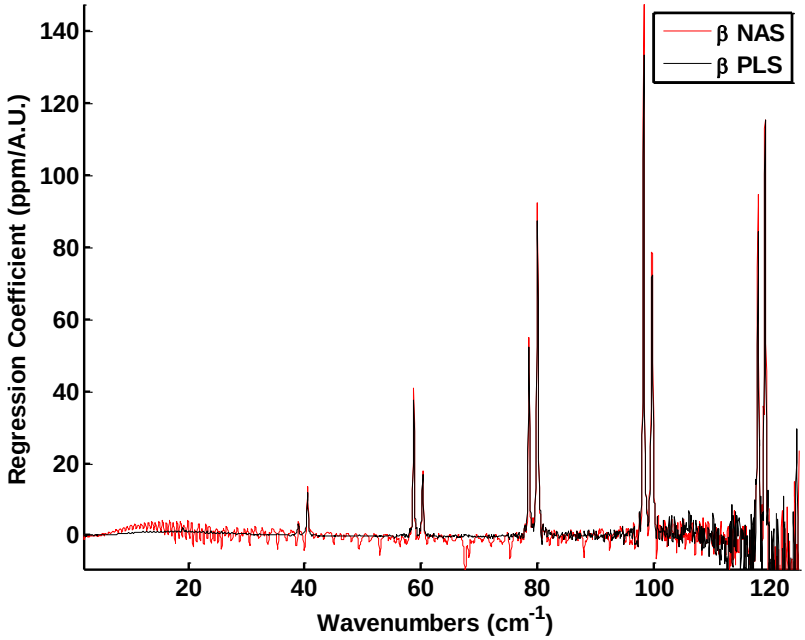


Figure VII-12. NAS and PLS regression vectors constructed using the full spectral range for ammonia in 8-component sample mixtures at 1 atm overall pressure.

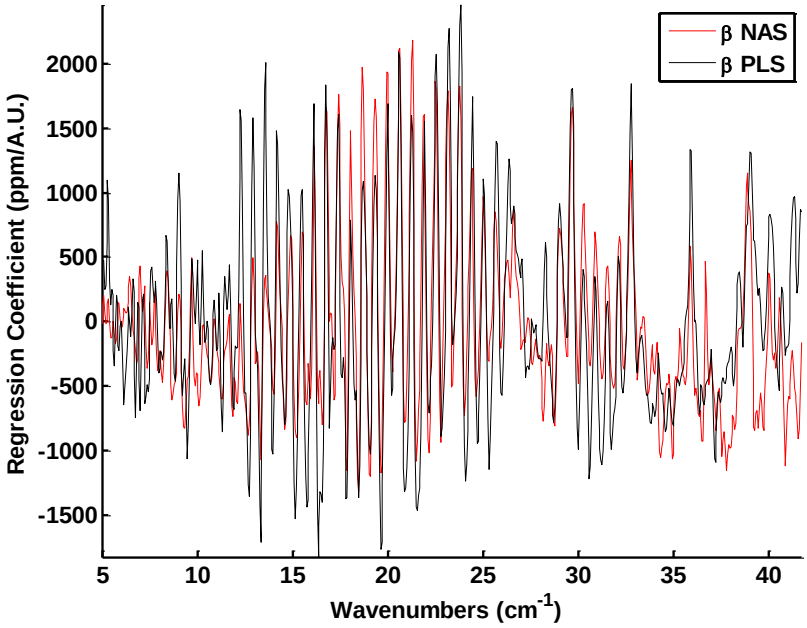


Figure VII-13. NAS and PLS regression vectors constructed using the optimized spectral range for acetaldehyde in 8-component sample mixtures at 1 atm overall pressure.

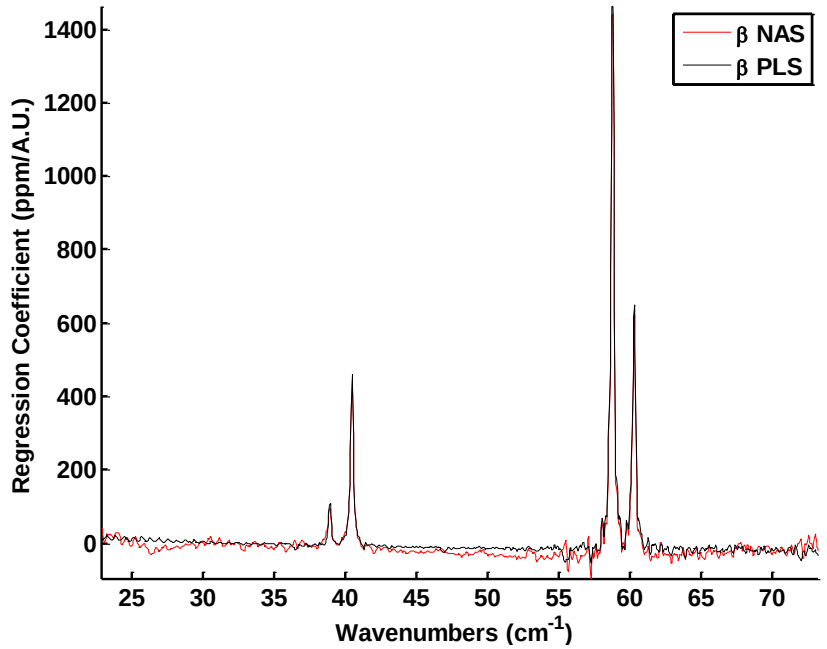


Figure VII-14. NAS and PLS regression vectors constructed using the optimized spectral range for ammonia in 8-component sample mixtures at 1 atm overall pressure.

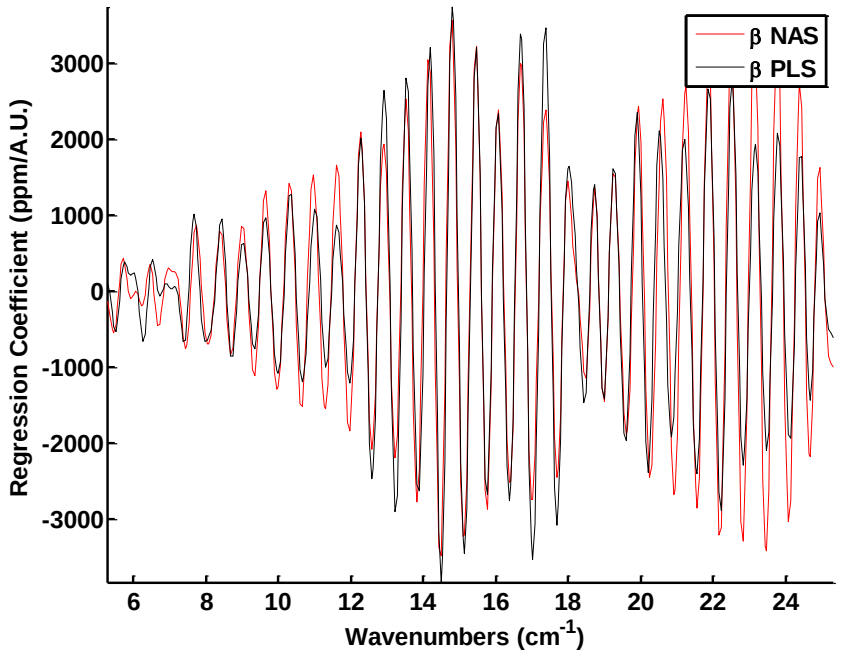


Figure VII-15. NAS and PLS regression vectors constructed using DFF pre-processing for acetaldehyde in 8-component sample mixtures at 1 atm overall pressure.

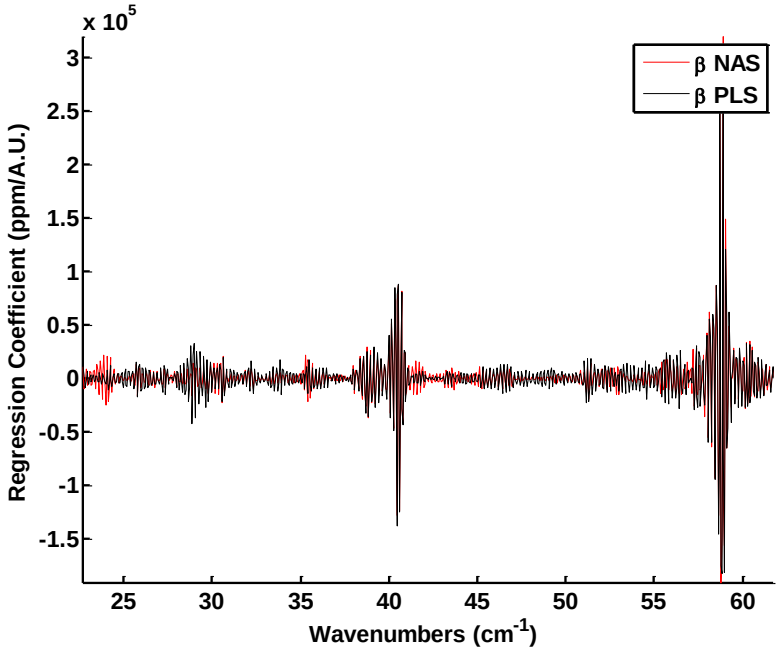


Figure VII-16. NAS and PLS regression vectors constructed using DFF pre-processing for ammonia in 8-component sample mixtures at 1 atm overall pressure.

Similarities are evident between the regression vectors of the pure components and the regression vectors calculated for the sample mixtures. Vectors calculated for weakly-absorbing species using the full spectral range tend to be remarkably different in magnitude. The regression vectors calculated from the sample mixtures with these conditions also display great differences in the shape of the vectors as a result of the spectral overlap between the analyte and interfering species. Vectors calculated for strong absorbers using the full spectral range show far better agreement as with the pure components, except that the vectors now include significant features related to interfering species, as shown in the low-frequency end of Figure VII-12 where the spectral features of the aldehydes and nitriles are discernible from the baseline.

Vectors calculated with the optimized spectral ranges bear much more similarity than those calculated for the full spectral range. In contrast to the regression vectors calculated for the pure components, these no longer bear such strong resemblance to the

absorption spectra of the pure analyte, as some portion of the regression vectors has changed due to spectral overlap with interfering species, a trend which is especially evident in the regression vectors of the weak absorbers, as shown in Figure VII-13. Regression vectors calculated for strong absorbers using the optimized spectral range still resemble the pure component absorption spectra, as illustrated by Figure VII-14.

Vectors calculated with the optimized DFF parameters and wavenumber bounds show excellent agreement in shape and magnitude, as was observed in comparison of regression vectors for the pure components. Selectivity of the quantitative models is evidenced by the high level of agreement between the regression vectors, explaining the significant decreases observed in the corresponding errors of prediction.

Conclusions

Robust quantitative models were constructed for each of the species of interest in the 8-component mixtures. As in the pure component analyses, wavenumber and DFF optimization provided significant improvement over models constructed using the full spectral range. Errors in prediction of each analyte's concentration were well below the PELs for each species, despite overlap between species' spectral features. These results demonstrate the sensitivity and selectivity of the THz-TDS measurement at a relatively consistent overall pressure. While this represents a significant advance towards the goal of the project, further studies were carried out to examine the impact of the overall sample pressure on the precision of quantitative models.

Foreign-Gas Broadening of Spectral Features

Overall pressure is known to affect the width of spectral features for gaseous absorbers, as described theoretically in the preceding section titled "Spectral broadening". Because atmospheric pressure varies as a function of altitude, and to a lesser extent weather conditions, this variability must be taken into account when exploring the feasibility of a proposed analytical method. As described previously, the extent of

spectral overlap will affect the chemical selectivity of the method, and thus varying spectral peak widths will affect selectivity.

Experimental

For this series of experiments, the compounds of interest were analyzed individually at various overall pressures. The procedure is outlined in the following list.

1. Reference
    a. evacuate system (0.0 Torr)
    b. collect reference spectrum
2. Sample
    a. add analyte (various pressures)
    b. collect sample spectrum
    c. add N_2 (various pressures)
    d. collect sample spectrum
    e. repeat steps c and d as necessary

Sample spectra for a given analyte were all referenced to one spectrum as collected with the cell under vacuum. As such, the sample was not removed from the sampling system until the final sample spectrum of the set had been recorded. Variability resulting from self-broadening was avoided by retaining the analyte between sample spectra. Samples were collected for each of the eight species of interest under the conditions listed in Table VII-9.

Spectra were processed with the SCA and extended to 16 × N, which provides greater certainty in the fit without requiring large amounts of time for attainment of a satisfactory fit. Reference spectral parameters were obtained from NASA's JPL,[349] as shown in the plots of rotational transitions in Figures V-1 through V-21. Parameters were condensed to reduce the time required for the fit and to retain a reasonable number of degrees of freedom. For example, at a spectral resolution of 0.225 cm^{-1} with 4 × N signal length after extension, the data point spacing in the resulting wavenumber axis is approximately 0.0505 cm^{-1}. Parameters were condensed such that transitions within a window of ±0.02525 cm^{-1} from each frequency axis data point were summed in

magnitude to produce a single transition with a center position determined by an intensity-weighted average of the resonant frequencies of the transitions.

With extension to a signal length of 16 × N, the number of data points in the spectrum is quadrupled, and the spacing is approximately 0.0126 cm^{-1}. Thus, even when dense rotational spectra are fitted with the Lorentzian function, 3 parameters are estimated and a minimum of 1 degree of freedom is available per fitted peak. To regain this degree of freedom without condensing spectral parameters, a threshold would have to be enforced such that a significant number of relatively strong rotational transitions would be altogether excluded from the fit, increasing the magnitude of the residuals between the experimental and fitted spectra. Parameter condensation is also justified given the extremely narrow spacing of the transitions (in some cases on the order of 1×10^{-3} cm^{-1} or less) relative to the spectral resolution available with the TPS Spectra 1000D. Furthermore, the widths of the observed spectral peaks far exceed the spacing of the condensed reference peak positions.

Spectra were collected with various pressures of nitrogen referenced to an evacuated cell to examine baseline variability which might arise from the serial collection of samples at various pressures. Slight changes were observed in the spectrum which followed the change in pressure, manifested as broad spectral absorption with low magnitude, which is thought to be the result of window flexion or seal compression at various pressures. This effect was removed from the experimental data described previously, and additional baseline correction was performed to compensate for broadband signal variation. Spectral peaks were fit with a Lorentzian function using the condensed parameters as starting guesses for peak positions and heights. Initial guesses for widths were determined by placing data cursors at various points across the absorption spectrum to estimate the HWHM of the spectral peaks in the experimental data. Fitting was achieved via 25 iterations of the LMA. Spectral regions included in the fit were determined by the availability of parameters, which were restricted with a

Table VII-9. Sample pressures from broadening experiments in units of Torr and sample temperatures. Minimum and maximum pressures (P_{min} and P_{max}, respectively) represent the total pressure, which is the sum of the analyte pressure and added nitrogen.

Species	$P_{analyte}$	P_{min}	P_{max}	T (°C)
CH_3CH_2CHO	27	253	807.9	23.7 ± 0.1
CH_3CH_2CN	13.5	280.8	831	23.7 ± 0.1
CH_3CH_2OH	11.8	255	800.7	23.5 ± 0.1
CH_3CHO	4.9	233.7	801.7	23.6 ± 0.1
CH_3CN	12.6	249.3	801.1	23.6 ± 0.1
CH_3OH	11.5	285.7	807.8	26.6 ± 0.1
H_2O	4.9	233.7	801.7	23.5 ± 0.1
NH_3	5.6	233.7	798.6	23.6 ± 0.1

threshold to exclude transitions which are approximately 1×10^{-2} times the strongest transition, as these parameters would primarily be fit to baseline features or noise.

Fitted parameters were used to estimate the pressure broadening coefficients of 5 transitions selected from various parts of the analyte spectra. Linear regressions of the peak widths versus the overall pressure of the sample produce a slope and intercept, where the slope represents the pressure broadening coefficient (in cm^{-1}/atm) and the y-intercept represents the self-broadened line width. Since the spectra were collected with the analyte at a single concentration, no effort is devoted to the interpretation of the y-intercept.

Results

Some representative spectra of acetaldehyde are presented below in Figures VII-17 through VII-19. In Figures VII-17 through VII-19, the effect of foreign-gas on the THz spectra of gaseous molecules is clear: as the overall pressure increases, the spectral peak widths of the analyte increase. The experimental spectrum and residuals of the fits

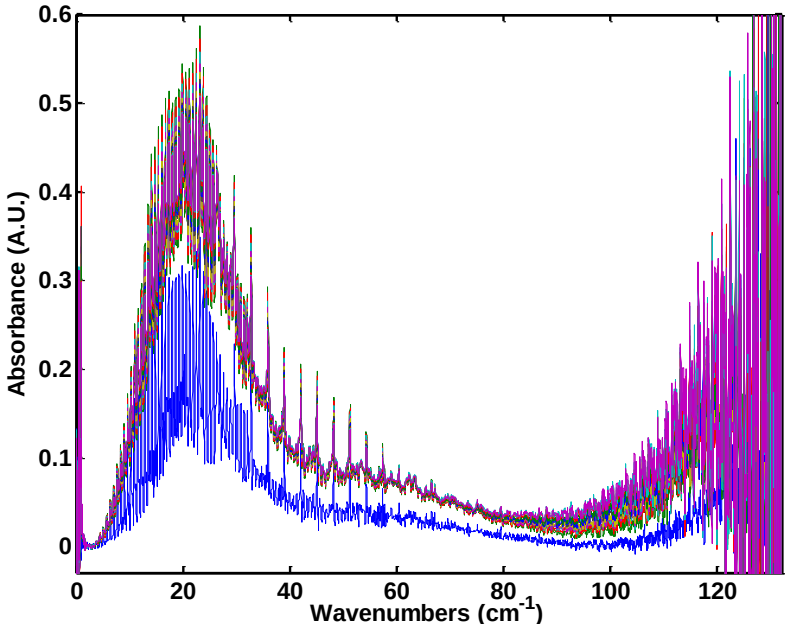


Figure VII-17. Absorption spectra of acetaldehyde vapor at various overall pressures. No offset is added to the spectra beyond baseline correction.

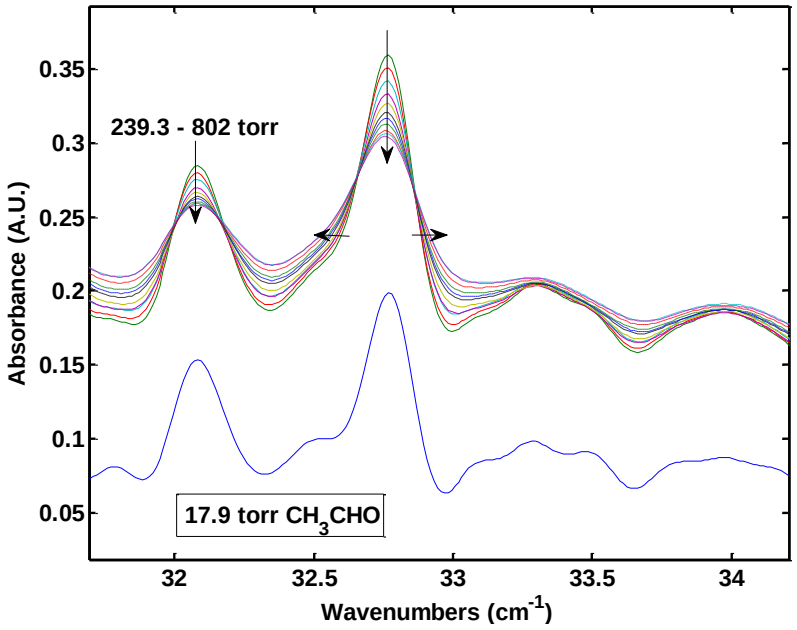


Figure VII-18. Expanded THz absorption spectra of acetaldehyde vapor at various overall pressures. The bottom (blue) trace is that of the pure vapor. No offset was added to the spectra beyond baseline correction.

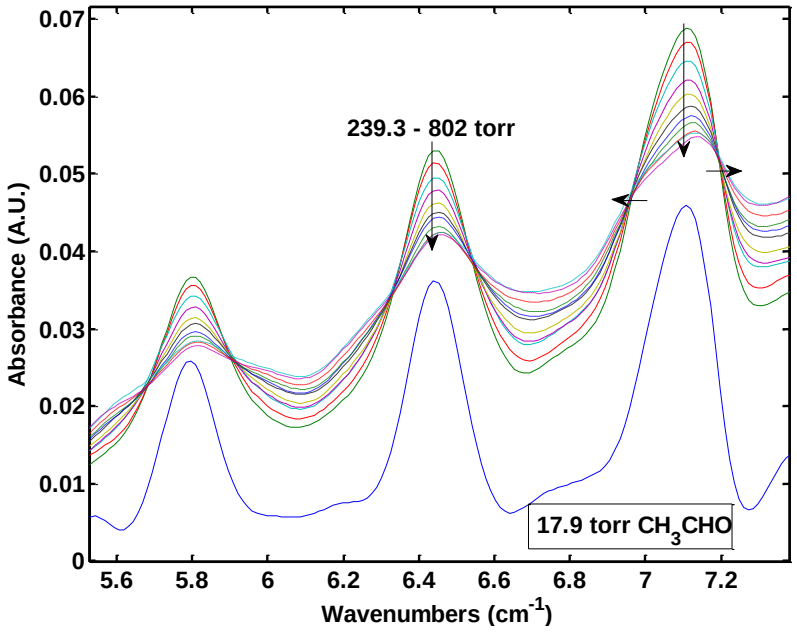


Figure VII-19. Expanded THz absorption spectra of acetaldehyde vapor at various overall pressures. The bottom (blue) trace is that of the pure vapor. No offset was added to the spectra beyond baseline correction.

of acetonitrile spectra are illustrated in Figure VII-20. In some cases, the residuals do not solely resemble spectral noise, as parameters used in the fit may only be those of the most common isotopologue(s) of the species. Also, as mentioned previously, rotational spectra of the analyte in an excited vibrational state may overlap with that of the analyte in the ground vibrational state. If parameters are not available for these species, significant residuals may appear. Observed peak widths are sufficiently high to obscure these extraneous transitions and thus prevent their isolation and exclusion from the fit. In other cases, the absorption features of species which are not described by the reference parameters may be weak enough that they are satisfactorily fit by the shoulder of a neighboring peak, which can result in calculation of peak widths which exceed the expected value.

Once spectral peaks were fitted to the reference parameters, the fitted peak width was plotted as a function of the overall pressure of the sample. As expected, the spectral

peak widths increased in a manner which was linearly proportional to the overall sample pressure, as shown in Figure VII-21. Scatter of the fitted widths about the best fit line is the result of several factors. Spectral noise affects the fitted parameters to a relatively low extent in these spectra, as the SNR is very high. The most significant contributors to the scatter about these plots are the isotopologues present in the sample which are not described by the reference parameters and baseline artifacts arising from window flexion or movement at various pressures. Scatter in the broadening plots is represented by the uncertainty in the fitted slope, as increased scatter will increase the uncertainty in the slope. Table VII-10 lists the wavenumber positions of select fitted peaks, and Table VII-11 lists the calculated broadening parameters.

The calculated broadening coefficients range from 0.010 (CH_3CH_2OH) to 0.163 (CH_3CN) cm^{-1}/atm. Given the relatively low magnitude of these values when compared to the available spectral bandwidth, spectral broadening is not expected to limit the application of this method to atmospheric analyses. Experimental broadening coefficients could not be directly compared to literature values in most cases, as much of the published literature containing these values focuses on microwave frequencies which are below those accessible with the TPS Spectra 1000D. Furthermore, no foreign-gas broadening coefficients were found for propionitrile, ethanol, or methanol. Values located in the literature for comparison were reported by Pandey,[353] Fabian,[354] Hoshina,[181] and Lightman.[109]

Pandey's reported coefficients are larger than most of the parameters determined in this work. For example, the broadening coefficient for a N_2-broadened rotational transition near 0.8 cm^{-1} was determined to be ~0.101cm^{-1}/atm, a value which is nearly double the values obtained from fits near the low-frequency end of the spectrum of acetaldehyde.[353] It is difficult to say whether this is a result of poor quality of fits to the experimental data, or whether the extent of broadening is so different between transitions near 11 cm^{-1} and 0.8cm^{-1}. Broadening coefficients have been shown to vary from one

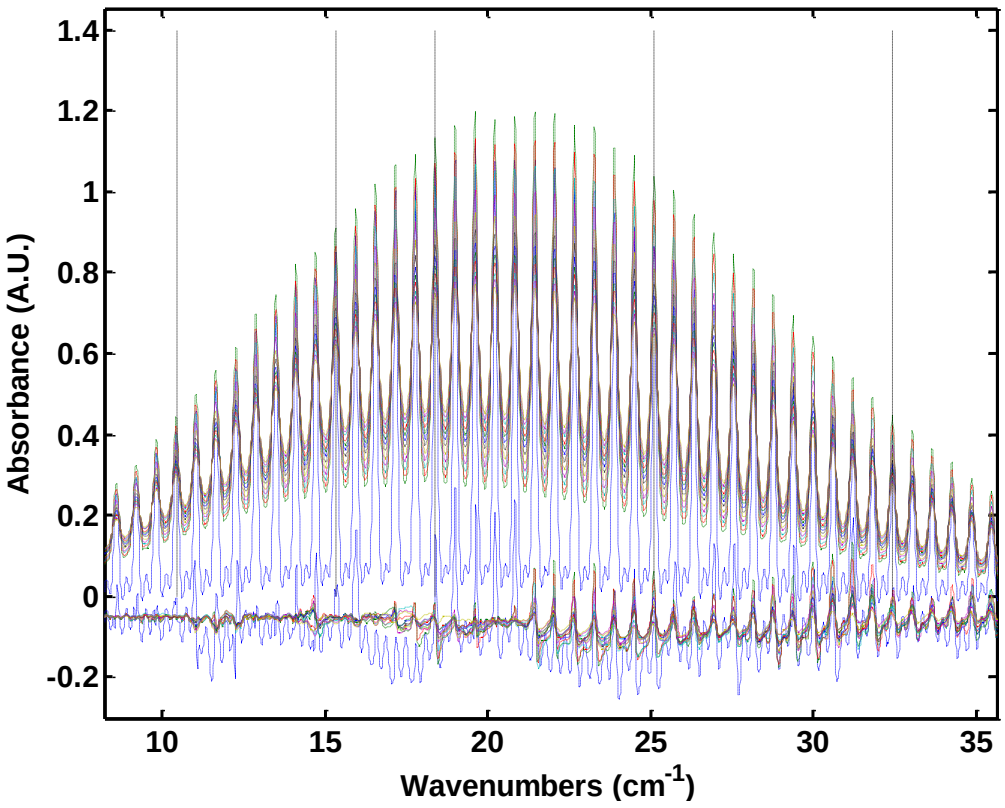


Figure VII-20. Broadened spectra of acetonitrile vapors. Vertical dashed lines represent the transitions selected for broadening analysis, solid lines represent the experimental data, and the dotted traces at the bottom of the spectrum represent the residuals between the experimental data and fitted spectra. Residuals are offset by -0.05 A.U. for comparison.

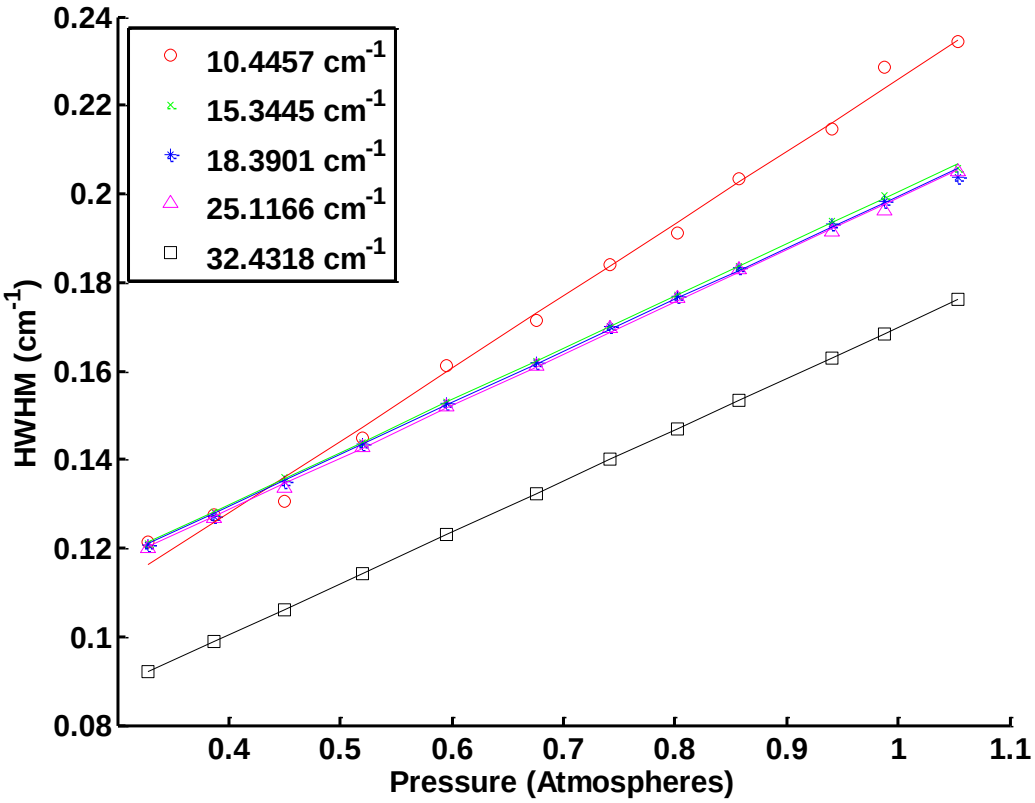


Figure VII-21. Peak widths and corresponding linear regressions. The slopes of the best fit lines represents the broadening coefficient of the respective rotational transition. The widths and broadening coefficients of the three center peaks (15 to 25 cm^{-1}) are nearly identical and are highly overlapped in this scale.

Table VII-10. Wavenumber positions (cm^{-1}) of select fitted peaks used for comparison of spectral broadening.

Species	Position 1	Position 2	Position 3	Position 4	Position 5
CH_3CH_2CN	14.11	15.81	19.51	20.07	23.30
CH_3CH_2OH	17.37	23.67	32.22	38.84	44.95
CH_3CHO	10.96	17.06	20.21	26.06	26.96
CH_3CN	10.45	15.34	18.39	25.12	32.43
CH_3OH	12.90	24.14	30.55	40.00	45.09
H_2O	18.57	25.08	47.04	69.17	73.24
NH_3	19.10	38.96	40.49	60.32	78.66

Table VII-11. Calculated spectral broadening coefficients in units of $cm^{-1}atm^{-1}$ and associated uncertainties from the selected fitted peaks.

Species	Position 1	Position 2	Position 3	Position 4	Position 5
CH_3CH_2CN	0.041 ± 0.005	0.032 ± 0.006	0.032 ± 0.005	0.046 ± 0.007	0.06 ± 0.01
CH_3CH_2OH	0.018 ± 0.004	0.063 ± 0.007	0.021 ± 0.001	0.069 ± 0.008	0.010 ± 0.001
CH_3CHO	0.046 ± 0.009	0.07 ± 0.01	0.051 ± 0.005	0.14 ± 0.02	0.020 ± 0.004
CH_3CN	0.163 ± 0.004	0.1175 ± 0.0005	0.117 ± 0.001	0.1172 ± 0.0008	0.1159 ± 0.0000
CH_3OH	0.07 ± 0.01	0.09 ± 0.02	0.09 ± 0.01	0.035 ± 0.007	0.034 ± 0.005
H_2O	0.094 ± 0.002	0.084 ± 0.002	0.0896 ± 0.0008	0.0805 ± 0.0008	0.0902 ± 0.0006
NH_3	0.076 ± 0.004	0.081 ± 0.005	0.09 ± 0.01	0.085 ± 0.009	0.08 ± 0.01

transition to another, so the fact that a difference is observed between calculated values is not in itself a cause for concern.

Fabian reported broadening coefficients for rotational transitions in acetonitrile vapor.[354] One such transition is located at 3.67 cm^{-1}, for which the N_2 broadening coefficient was reported as 0.160 cm^{-1}/atm.[354] This value is within one standard deviation of the experimentally-determined value of 0.163 ± 0.004 cm^{-1}/atm, calculated for the transition at 10.45 cm^{-1}. This is indicative of a relatively large collisional cross-section, as described earlier in this chapter.

Hoshina et al. published several broadening coefficients for water vapor, as mentioned in Chapter I.[181] A broadening coefficient of 0.113 cm^{-1}/atm was calculated for the transition near 18.57 cm^{-1} when the spectral features were broadened by nitrogen gas.[181] This value is significantly higher than the experimentally-determined value of 0.094 ± 0.002 cm^{-1}/atm, given the low standard deviation of the fitted parameter.

Lightman and BenReuven reported broadening parameters of ammonia in the presence of nitrogen, as described in Chapter I. Parameters were reported for several rotational transitions near the high-frequency end of the spectral range accessible with the TPS Spectra 1000D. A broadening coefficient of 0.115 ± 0.010 cm^{-1}/atm was published for the transition near 99 cm^{-1}, which is separated from the experimentally-determined value of 0.08 ± 0.01 cm^{-1}/atm (at 78.66 cm^{-1}) by about 1.5 standard deviations.

Discrepancies between values reported in the literature and those determined in these experiments may arise from a number of factors. As illustrated in Chapter II, implementation of the SCA provides greater reproducibility in the shape and position of the observed spectral peaks. Irreproducibility in the position of TDS segments could be a cause of the discrepancies between the reported values and observed spectral features. Furthermore, given the relatively short path length, higher quantities of analyte must be analyzed to obtain a useable SNR, and the increased concentration could conceivably

result in distortion of the observed spectral peaks as many of the analyte gases are strongly-polar molecules.

Conclusions

The effect of foreign-gas broadening by nitrogen gas was characterized and quantified by fitting Lorentzian functions to the observed spectral peaks, and calculated broadening parameters ranged from 0.010 (CH_3CH_2OH) to 0.163 (CH_3CN) cm^{-1}/atm. Spectral bandwidth provided by the TPS Spectra 1000D is more than sufficient given the comparatively-small extent of spectral broadening, but overlap between the features of the species of interest is still expected to vary appreciably with the overall pressure. The extent of this overlap will be much more pronounced in the spectra of the heavy species (all but ammonia and water vapor), as their features are in much closer proximity than those of the light species. Values reported herein provide a useful set for generalization of trends, but several are fairly distant from values reported in the literature. Ideally, this set of experiments would be repeated with the analyte present in a very low concentration at various low overall pressures, with analyses carried out in a long-pass cell to provide improved resolution of closely-spaced transitions. Additionally, the use of an apparatus with narrower spectral resolution would be desirable, as spectral peaks would be represented by more experimental data points and the observed peak widths would not be limited by the spectral resolution of the instrument.

Eight-Component Sample Mixtures at Various Overall Pressures

Sample mixtures were prepared in which the overall pressure was adjusted by adding varying amounts of nitrogen gas to examine the impact of varied spectral broadening on the precision of quantitative models. These samples represent another step toward experimental representation of realistic atmospheric analyses by incorporating the spectral variability which is introduced by varying the overall pressure of the sample. As

demonstrated in the previous section, the width of spectral peaks observed in this frequency range vary linearly with the overall sample pressure, which in turn affects the overlap of spectral features of the species of interest.

Experimental

Stock mixtures were prepared in the same manner as described for other ppm-range samples. The mixtures consisted of one analyte species mixed with nitrogen, and were stored in flasks attached to the sample handling manifold. Reference spectra were collected with the cell filled with nitrogen at the same pressure as the respective sample. By referencing a sample spectrum to these data, the effect of window flexion on the spectrum could be eliminated. Samples were prepared by sequentially adding the species to the manifold and cell, followed by addition of nitrogen to the desired overall pressure. The propagated error in the sample concentration was determined using the same method as was used for the 8-component mixtures at 1 atm overall pressure. Analyte concentrations were calculated using Equation VI-1. The overall pressure for this calculation was fixed at 760 Torr so the analyte concentrations would not be correlated to the overall pressure of the sample.

As with the 8-component mixtures prepared at 1 atm overall pressure, the concentrations of each species in each sample were selected to reduce correlation between the concentrations of the species and the correlation between species concentrations and collection sequence. In addition to these considerations, samples were added in order from the highest molecular weight to the lowest molecular weight in an effort to reduce the effect of sample contamination which may result from effusion of interfering species into the sample flasks. The overall pressure of the sample mixtures was treated as an analyte, meaning pressures were selected such that the correlation between species concentrations and the overall pressure was reduced. Correlation coefficients from these samples are provided in Table VII-12.

1. Stock mixtures
    a. evacuate system (0.0 Torr)
    b. insert analyte (various pressures)
    c. add N_2 (~760.0 Torr)
    d. retain mixture in flask
    e. repeat for each species of interest
2. Reference
    a. evacuate system (0.0 Torr)
    b. add N_2 (various pressures, ~233-799 Torr)
    c. collect reference spectrum
3. Sample
    a. evacuate system (0.0 Torr)
    b. add stock mixtures (various pressures, repeat for each analyte)
    c. add N_2 (various pressures, ~233-799 Torr)
    d. adjust volume (adjusts pressure to match target pressure)
    e. collect sample spectrum

Small values in Table VII-12 indicate low correlation between the concentrations of interfering species. Values on the diagonal tend to be higher than those off the diagonal, indicating higher correlation between the sample concentrations and the collection sequence, though these values are still very low. This experimental design increases the robustness of the PLS regression models, as spectral features of interfering species will not be attributed to the analyte.

Spectra were calculated with SCA-processed signals with 4×N signal length after extension. Baseline correction was carried out by subtracting the mean of the 2.72 to 4 cm^{-1} baseline region from the entire spectrum as with the 8-component samples analyzed at 1 atm overall pressure. Once again, only the PLS algorithm was used for quantitative analysis. The three model construction methods (full range, optimized range and optimized DFF and wavenumber range) were evaluated for PLS cross-validation with each compound of interest. A total of 100 samples were prepared and analyzed in the temperature range of 23.9 ± 0.1 °C at final pressures of 522.9 ± 164.9 Torr. Concentration ranges for each analyte are provided in Table VII-13. The overall pressure of the sample is denoted by the symbol P_{total}.

Table VII-12. Correlation between reference concentrations for analyte-interferent and analyte-sequence pairs. Values on the diagonal (highlighted in gray) represent analyte-sequence correlation, whereas values off the diagonal represent analyte-interference correlation.

R^2	CH_3CH_2CHO	CH_3CH_2CN	CH_3CH_2OH	CH_3CHO	CH_3CN	CH_3OH	H_2O	NH_3	P_{total}
CH_3CH_2CHO	4.6E-02								
CH_3CH_2CN	3.0E-06	1.7E-02							
CH_3CH_2OH	1.1E-06	2.2E-08	7.0E-02						
CH_3CHO	1.7E-06	7.8E-06	1.6E-07	2.5E-02					
CH_3CN	2.0E-07	1.9E-06	2.0E-06	3.5E-08	7.8E-06				
CH_3OH	2.0E-06	8.9E-08	7.5E-08	6.6E-06	2.0E-06	2.3E-02			
H_2O	2.8E-06	1.7E-05	3.3E-07	2.4E-06	5.3E-06	5.8E-06	5.1E-03		
NH_3	3.2E-07	5.8E-06	1.8E-09	8.4E-08	2.3E-06	1.4E-06	9.5E-06	3.1E-03	
P_{total}	3.0E-07	2.2E-06	2.6E-06	4.7E-07	4.1E-07	8.9E-07	2.9E-06	1.8E-06	8.8E-08

Results

Spectral data collected with these samples are illustrated in Figure VII-22. As with the samples analyzed at 1 atm overall pressure, the features of individual species are somewhat difficult to identify in the low-frequency end of the spectrum. Spectral noise was once again found to be a majority contributor to the observed spectrum in the extremities of the accessible spectral range. Optimized parameters for each regression model are provided in Table VII-14. Once again, the optimized wavenumber bounds and filter parameters are not identical with those of the pure components, but are selected to provide improved predictive performance. While the overall pressure of the sample is not a species of interest, PLS models were also constructed in an effort to use the observed spectral features to gauge the pressure in the cell.

Those species which are weak absorbers of these THz frequencies required an increased number of latent variables in construction of the respective PLS regression models, a trend which was also observed in quantitation of the species in the mixtures at 1 atm overall pressure. However, in all cases (except those of water and ammonia), more latent variables were required to produce robust quantitative models in samples analyzed at various overall pressures than for those analyzed at 1 atm overall pressure. Increased model complexity is likely required to account for the variation in spectral overlap, as will be discussed in greater detail with comparison of regression vectors. Traces of the error of prediction are plotted as a function of the number of latent variables in Figures VII-23 through VII-25. The shapes of these traces are similar to those observed with mixtures analyzed at 1 atm overall pressure. Predictive error decreases minimally as model complexity increases when utilizing the full spectral range, as shown in Figure VII-23. Conversely, the error of prediction drops fairly rapidly with increasing model complexity when using the optimized spectral range, as illustrated in Figure VII-24. When DFF pre-processing and wavenumber selection are combined, the error of

Table VII-13. Concentration statistics for each compound of interest in the 8-component sample mixtures analyzed at various overall pressures.

Species	C_{mean} (ppm)[1]	C_{min} (ppm)[2]	C_{max} (ppm)[3]	S_{conc} (ppm)[4]
CH_3CH_2CHO	315	0	664	183
CH_3CH_2CN	354	0	693	188
CH_3CH_2OH	320	0	706	182
CH_3CHO	353	0	690	182
CH_3CN	368	0	759	192
CH_3OH	363	0	791	198
H_2O	335	0	772	200
NH_3	355	0	787	202
P_{total}	524	235.8	797	165

1: Mean of analyte concentrations in prepared sample mixtures.

2: Minimum of analyte concentrations in prepared sample mixtures.

3: Maximum of analyte concentrations in prepared sample mixtures.

4: Standard deviation of analyte concentrations in prepared sample mixtures.

prediction decreases rapidly as more latent variables are included in the regression model, as shown in Figure VII-25.

Like the pure component and 8-component mixtures analyzed at 1 atm overall pressure, the errors of prediction obtained with the PLS regression models are greatest when using the full spectral range, as shown in Table VII-15. Predictive errors decrease when wavenumber selection is performed, and decrease further in most cases when DFF pre-processing is used. The significance of these changes was evaluated using the F-test, and the calculated F-values are provided in Table VII-16. Wavenumber optimization provided significant improvement for quantitative models developed for propionaldehyde, propionitrile, ethanol, acetaldehyde, acetonitrile, methanol, and the

overall sample pressure. Wavenumber and DFF parameter optimization provides a significant decrease in the error of prediction for all species, and provides significant improvement in the error of prediction for several species when compared to wavenumber optimization. In some cases, DFF pre-processing increased the error of prediction over that of the wavenumber-optimized models.

Concentration correlation plots reflect the precision of the respective models. Poor precision can be seen in PLS regression models constructed using the full spectral range, manifested as scatter about the ideal line as shown in Figure VII-26. Only the concentrations of the strongest absorbers, i.e. water and methanol, are somewhat determined by these regression models. When wavenumber optimization is used, the precision is drastically improved, as shown in Figure VII-27. The scatter about the ideal line is greatly reduced for almost all of the compounds of interest. Some improvements may be seen when DFF pre-processing is used, though the precision of some models are adversely impacted by implementation of DFF, as shown in Figure VII-28. Regression model precision is also evident in the plots of the concentration residuals, shown in Figures VII-29 through VII-31.

Correlation coefficients were calculated to examine correlation between reference and residual concentrations. Large R^2 values on the diagonal of Table VII-17 reflect the strong correlation between the concentration residuals of the analytes and their reference concentrations from regression models constructed with the full spectral range. This correlation is the result of poor model performance. Values off the diagonal are near zero, indicating some degree of chemical selectivity or a lack of reactivity between the compared species. Diagonal elements in Tables VII-18 and VII-19 are much closer to zero, reflecting the drastic improvement in the precision of the quantitative models, and values off the diagonal remain very close to zero, with the exception of some values in the rows for the analyte propionaldehyde.

In Table VII-18, some correlation is observed between the reference concentration of propionaldehyde and the residual concentrations for propionaldehyde and propionitrile. Correlation between propionaldehyde and propionaldehyde residuals is the result of one or two relatively large residuals. The correlation observed between the concentration of propionaldehyde and the residual of propionitrile is thought to be the result of the strong spectral overlap and weak absorptivity of these two species. Further investigation into the chemical selectivity of these models is discussed in greater detail in a following section. Similar correlation is present in Table VII-19, where an R^2 value of 0.22 is calculated for propionaldehyde reference concentrations and propionaldehyde concentration residuals. This correlation value is also thought to be the result of relatively large residuals.

Comparison of the regression vectors reveals slight changes in the shape and magnitude of the coefficients. Generally speaking, the PLS and NAS model regression coefficients calculated using the full spectral range are dissimilar in shape and magnitude. Vectors calculated from the optimized spectral ranges tend to show some similarity in magnitude, but the shapes of the features are different, likely a result of the changing spectral peak widths. Vectors calculated with wavenumber optimization and DFF pre-processing exhibit a high level of correlation in shape and magnitude. Figures VII-32 through VII-38 portray representative vectors for NAS and PLS models developed for propionitrile and ammonia.

Features in the regression vectors of ammonia (and water vapor) tend to match well between NAS and PLS models constructed with the full range, optimized ranges, and DFF pre-processed spectra. Similarly, little change is seen in the observed error of prediction. Regression coefficients for propionitrile (and the other weak absorbers) tend to exhibit little agreement when comparing NAS and PLS-calculated vectors produced by models using the full spectral range, exemplified by the traces in Figure VII-32. Vectors calculated with the optimized spectral ranges show improved agreement in magnitude,

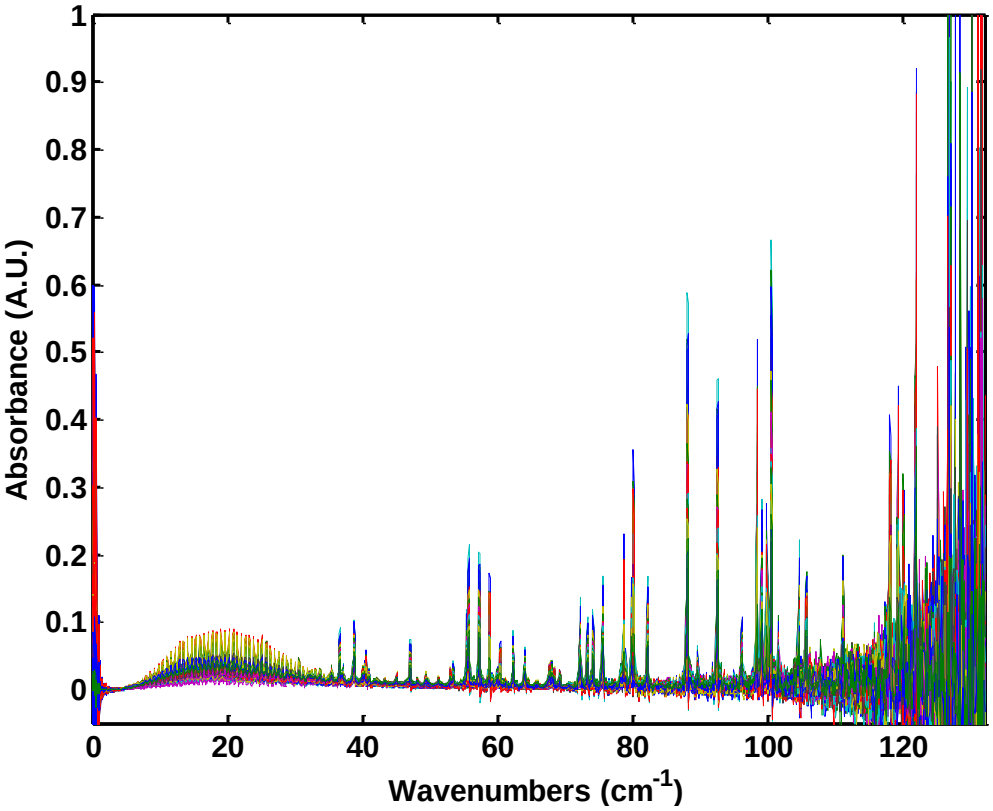


Figure VII-22. Spectra collected with analyte mixtures at various overall pressures.

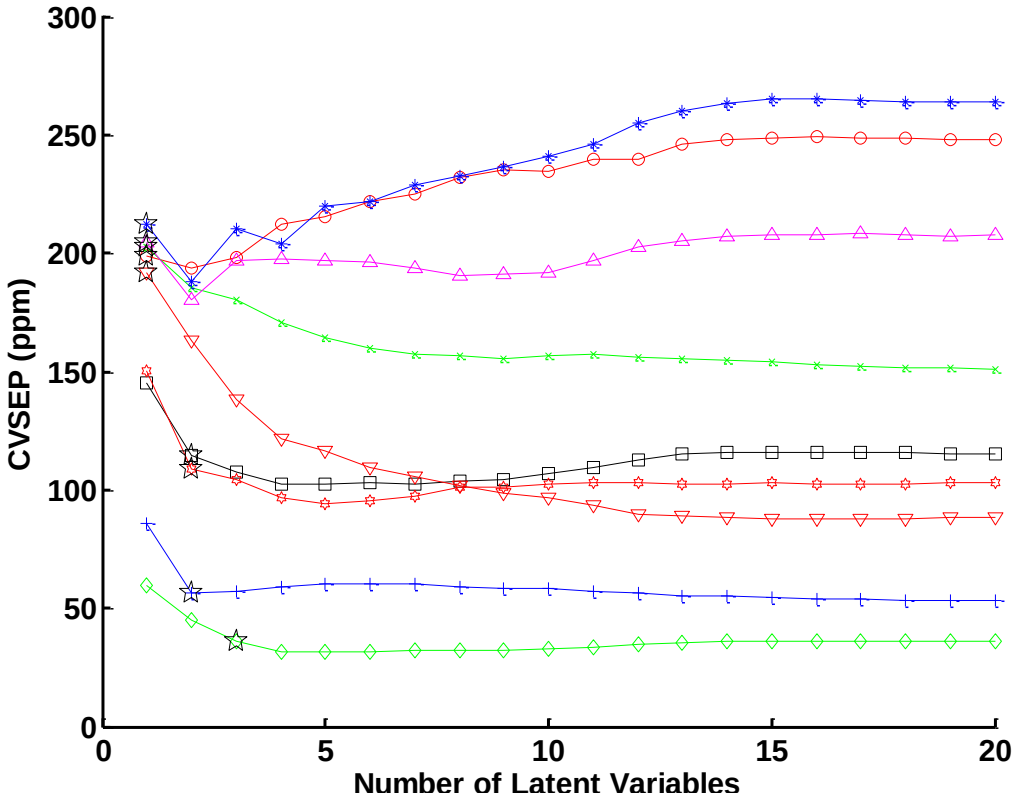


Figure VII-23. Errors of prediction calculated with PLS models constructed with different numbers of latent variables using the full spectral range. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, ammonia by blue +'s, and the overall pressure by red hexagrams. The appropriate number of latent variables is indicated by a star for each analyte.

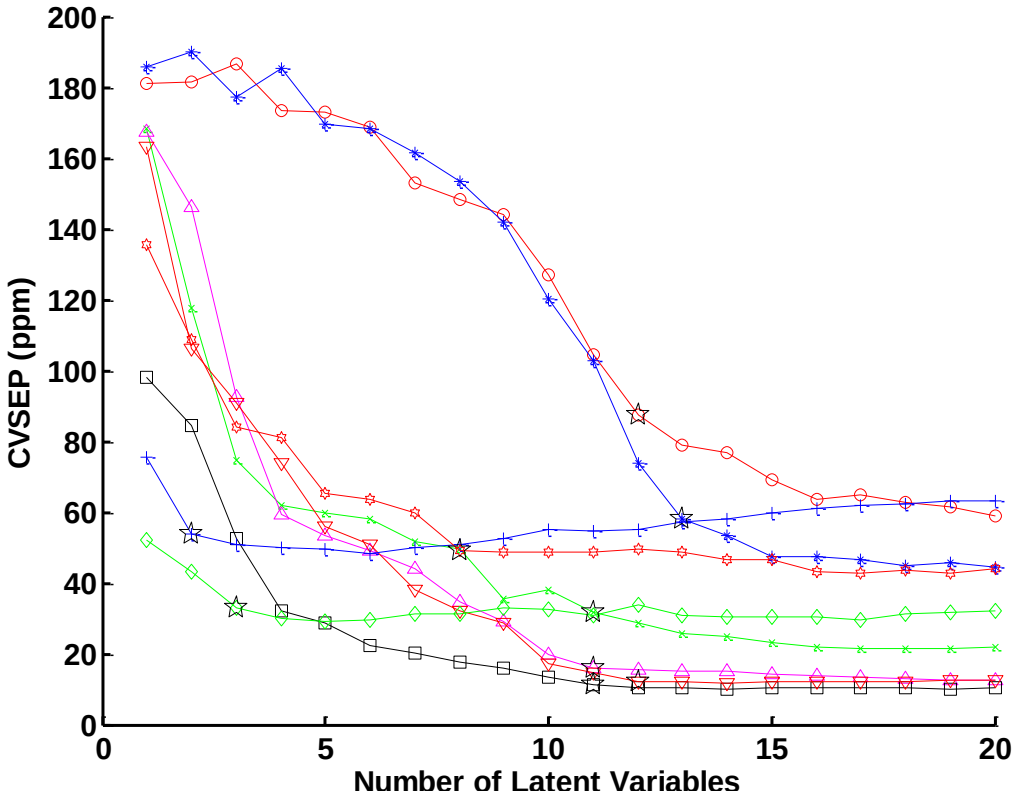


Figure VII-24. Errors of prediction calculated with PLS models constructed with different numbers of latent variables using the optimized spectral range. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, ammonia by blue +'s, and the overall pressure by red hexagrams. The appropriate number of latent variables is indicated by a star for each analyte.

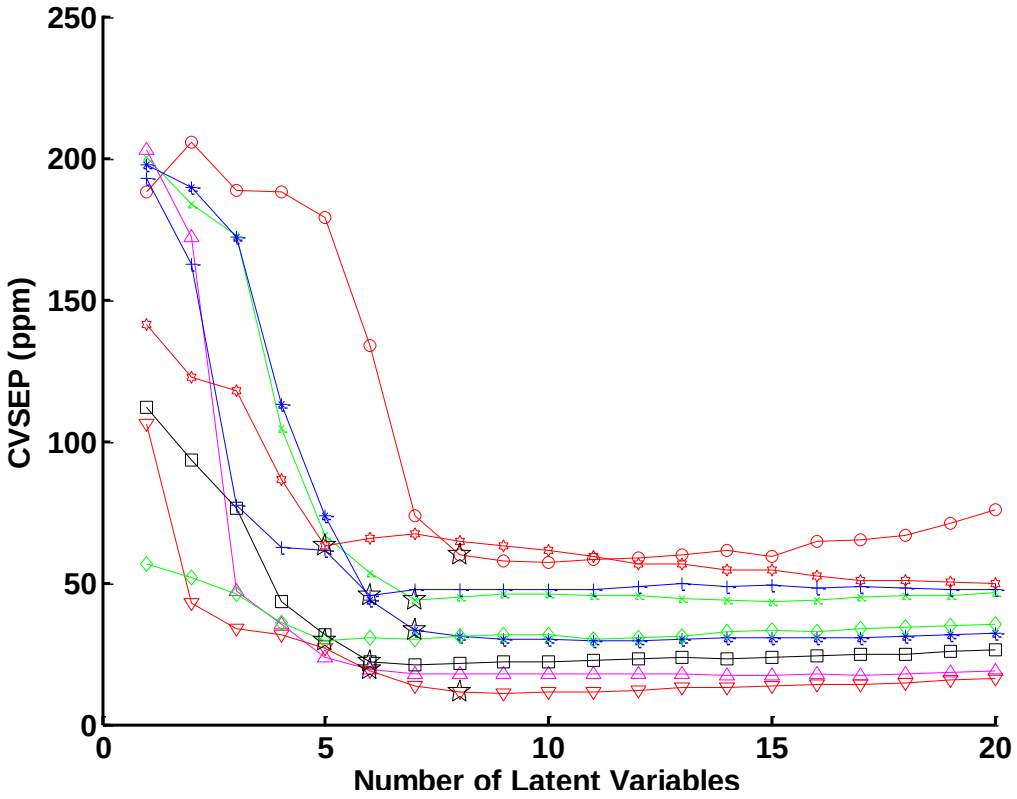


Figure VII-25. Errors of prediction calculated with PLS models constructed with different numbers of latent variables using the optimized spectral range and DFF pre-processing. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, ammonia by blue +'s, and the overall pressure by red hexagrams. The appropriate number of latent variables is indicated by a star for each analyte.

Table VII-14. Optimized parameters used for PLS cross-validation with 8-component sample mixtures analyzed at various overall pressures.

Species	RAW	WN			WN, DFF				
	# LV[1]	WN_{low} (cm^{-1})[2]	WN_{high} (cm^{-1})[2]	# LV[1]	Filter μ[3]	Filter width[3]	WN_{low} (cm^{-1})[2]	WN_{high} (cm^{-1})[2]	# LV[1]
CH_3CH_2CHO	1	9.0	49.4	12	0.0803	0.0353	9.4	30.4	8
CH_3CH_2CN	1	8.9	51.4	11	0.0495	0.0185	14.5	45.6	7
CH_3CH_2OH	1	9.1	61.2	13	0.0220	0.0122	15.0	45.0	7
CH_3CHO	1	15.3	35.3	11	0.0295	0.0185	12.2	58.1	6
CH_3CN	2	10.0	30.0	11	0.0986	0.0344	10.3	31.0	6
CH_3OH	1	9.1	61.2	12	0.0598	0.0519	32.9	58.4	8
H_2O	3	28.3	56.9	3	0.0481	0.0436	20.0	40.0	5
NH_3	2	3.0	100.0	2	0.0073	0.0055	15.0	45.0	6
P_{total}	2	20.5	40.7	8	0.0961	0.0438	11.9	36.9	5

1: Number of latent variables.

2: Wavenumber units (cm^{-1}).

3: Digital frequency units.

Table VII-15. Errors of prediction (in ppm) from PLS regression models utilizing various processing methods.

Species	RAW	WN	WN, DFF	PEL
CH_3CH_2CHO	199	87	60	NA
CH_3CH_2CN	202	32	44	NA
CH_3CH_2OH	212	58	33	1000
CH_3CHO	204	16	19	200
CH_3CN	115	11	22	40
CH_3OH	192	12	11	200
H_2O	36	33	30	NA
NH_3	56	54	45	50
P_{total}	108.6	49.3	63.1	NA

Table VII-16. Calculated F-values representing the significance of the optimized processing methods. The critical F value is 1.39, and values exceeding this threshold are denoted by yellow highlighting.

Species	F_{calc} (RAW-WN)	F_{calc} (DFF-RAW)	F_{calc} (DFF-WN)
CH_3CH_2CHO	5.21	11.02	2.11
CH_3CH_2CN	40.32	21.29	1.89
CH_3CH_2OH	13.31	41.36	3.11
CH_3CHO	164.77	112.26	1.47
CH_3CN	105.33	27.53	3.83
CH_3OH	245.84	285.66	1.16
H_2O	1.21	1.50	1.25
NH_3	1.10	1.55	1.41
P_{total}	4.85	2.96	1.64

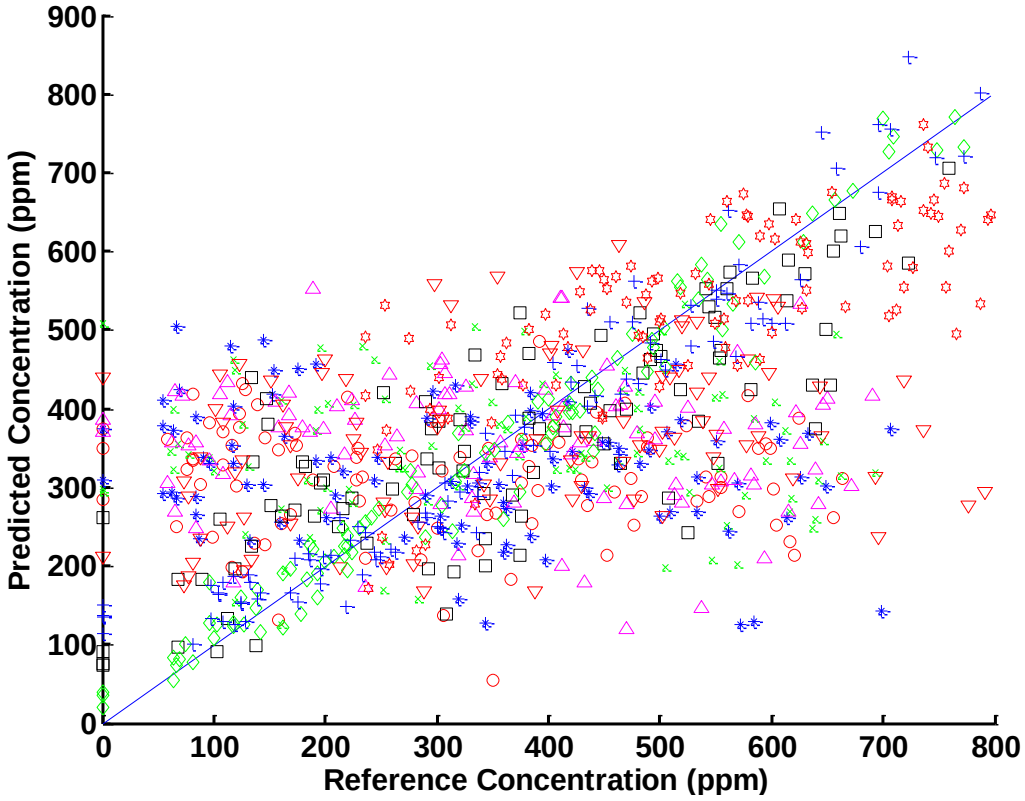


Figure VII-26. Concentration correlation plot from PLS prediction of species in 8-component mixtures at various overall pressures using the full spectral range. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, ammonia by blue +'s, and the overall sample pressure by red hexagrams. Poor precision is evidenced by extreme scatter about the ideal line (shown as a solid blue line).

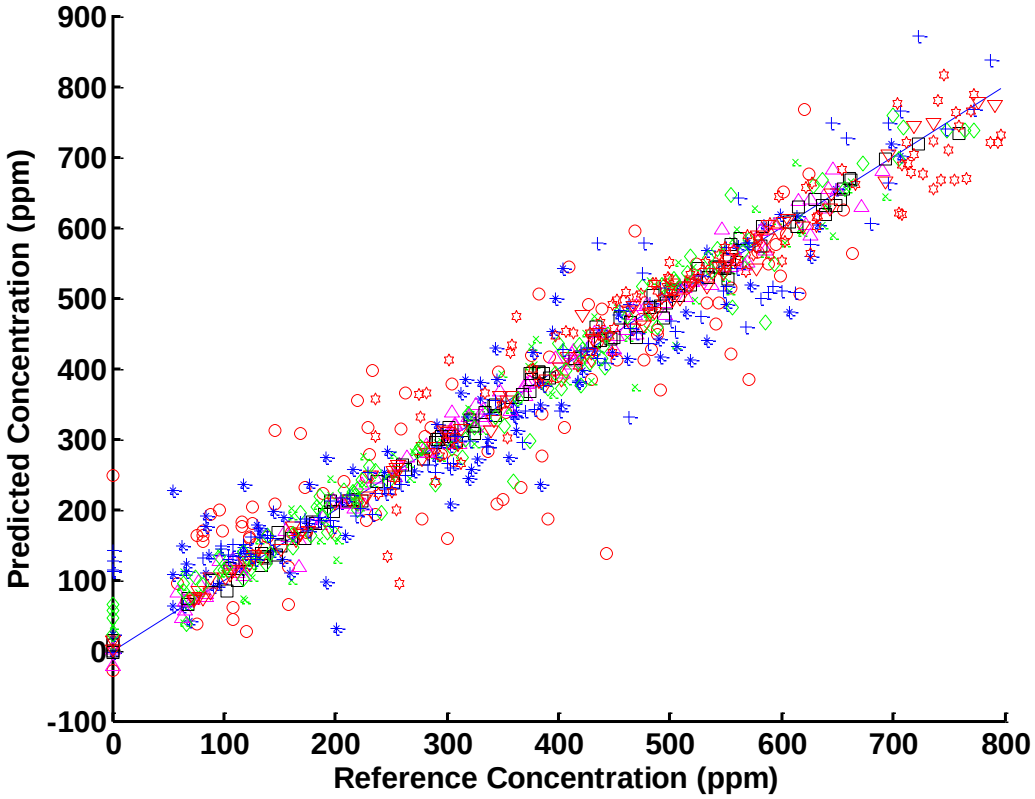


Figure VII-27. Concentration correlation plot from PLS prediction of species in 8-component mixtures at various overall pressures using the optimized spectral ranges. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, ammonia by blue +'s, and the overall sample pressure by red hexagrams. The ideal line is shown as a solid blue line with a slope equal to one.

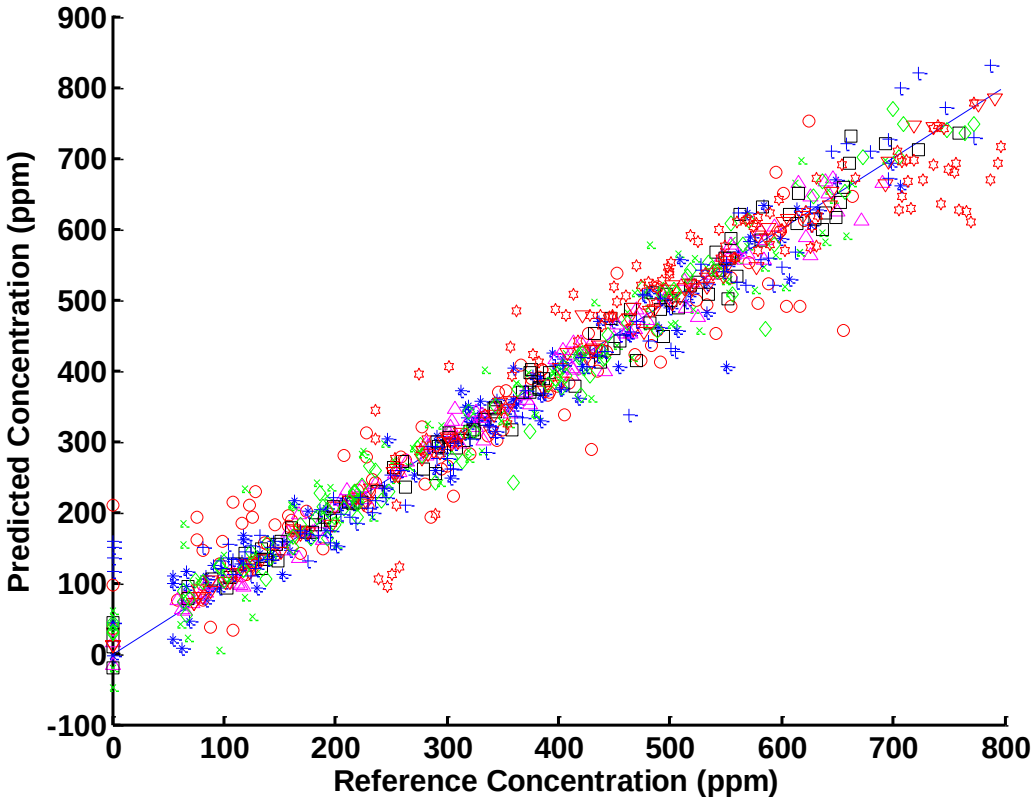


Figure VII-28. Concentration correlation plot from PLS prediction of species in 8-component mixtures at various overall pressures using the optimized spectral ranges and DFF pre-processing. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, ammonia by blue +'s, and the overall sample pressure by red hexagrams. The ideal line is shown as a solid blue line with a slope equal to one.

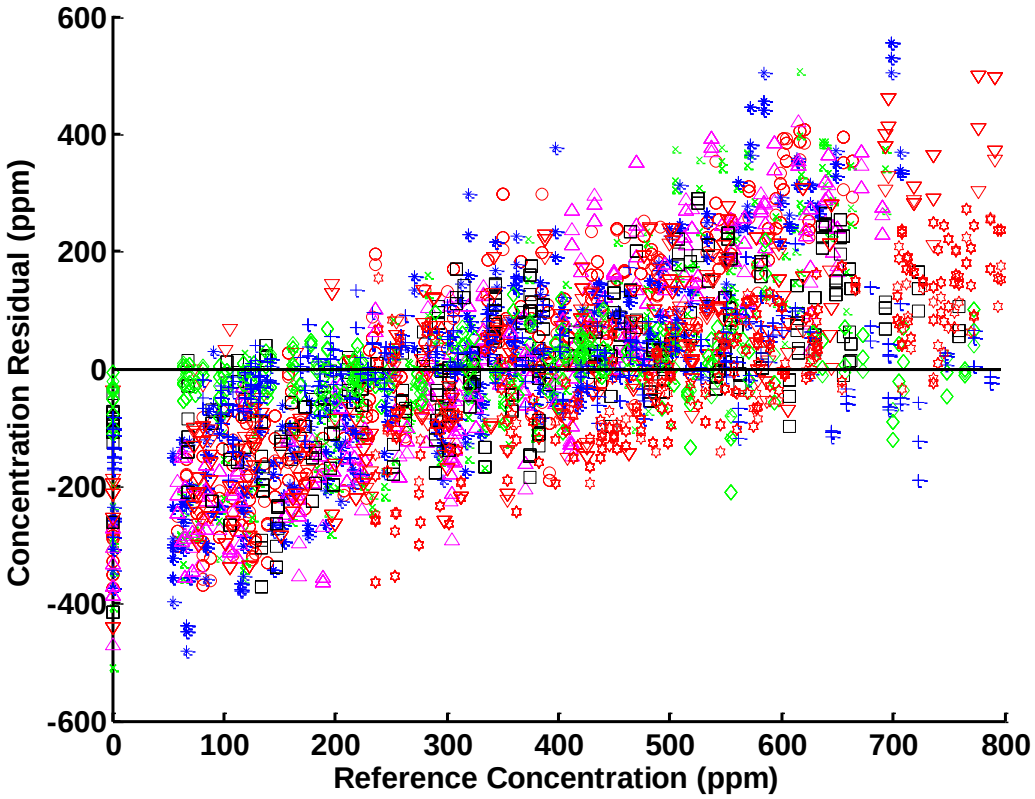


Figure VII-29. Residuals of predicted concentrations as a function of the reference concentration from PLS regression models constructed with the full spectral range. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, ammonia by blue +'s, and the overall sample pressure by red hexagrams. Zero is indicated by a solid black line.

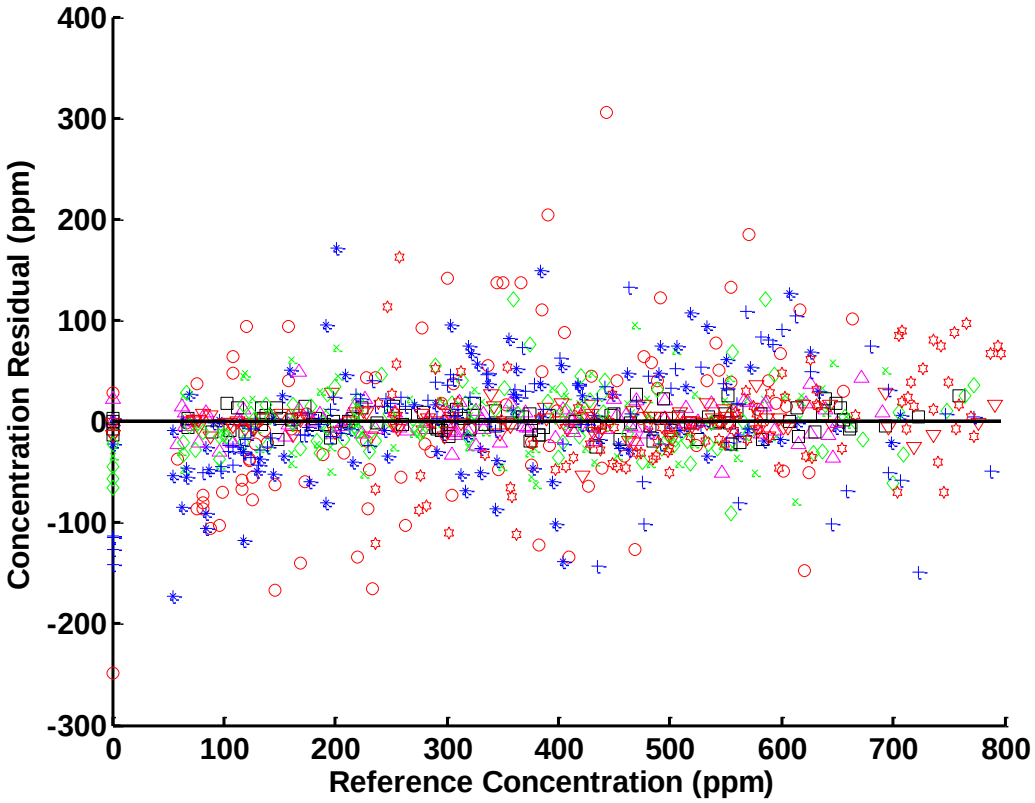


Figure VII-30. Residuals of predicted concentrations as a function of the reference concentration from PLS regression models constructed with the optimized spectral range. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, ammonia by blue +'s, and the overall sample pressure by red hexagrams. Zero is indicated by a solid black line.

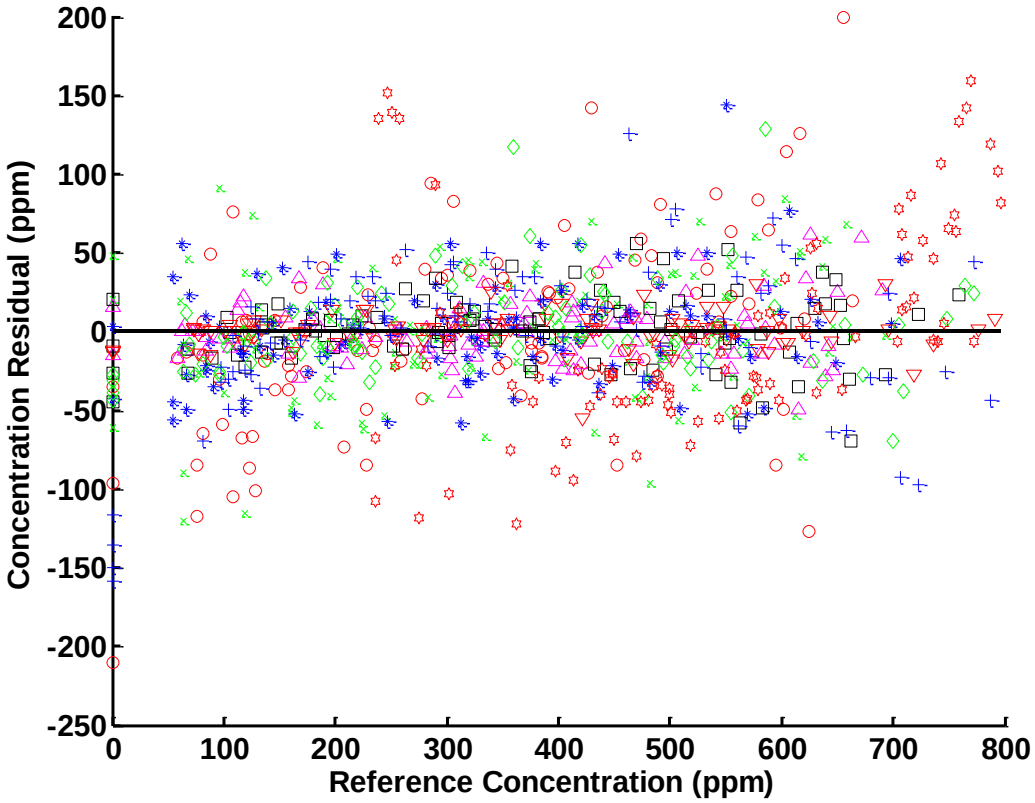


Figure VII-31. Residuals of predicted concentrations as a function of the reference concentration from PLS regression models constructed with the optimized spectral range and DFF pre-processing. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, ammonia by blue +'s, and the overall sample pressure by red hexagrams. Zero is indicated by a solid black line.

Table VII-17. Correlation between analyte concentration and prediction residuals from PLS regression models utilizing the full spectral range. Gray cells contain the correlation coefficient for the relationship between analyte reference and analyte residual concentrations. All other cells provide the correlation between the analyte reference and interferent residual concentrations.

		Residual Species								
	Species	CH_3CH_2CHO	CH_3CH_2CN	CH_3CH_2OH	CH_3CHO	CH_3CN	CH_3OH	H_2O	NH_3	P_{total}
Analyte Species	CH_3CH_2CHO	0.88	0.01	0.00	0.00	0.04	0.00	0.00	0.04	0.00
	CH_3CH_2CN	0.00	0.83	0.00	0.01	0.16	0.00	0.01	0.00	0.00
	CH_3CH_2OH	0.00	0.00	0.87	0.01	0.02	0.01	0.01	0.02	0.02
	CH_3CHO	0.00	0.02	0.01	0.84	0.21	0.00	0.00	0.03	0.00
	CH_3CN	0.01	0.07	0.01	0.06	0.44	0.00	0.04	0.01	0.04
	CH_3OH	0.00	0.00	0.01	0.01	0.00	0.74	0.01	0.00	0.01
	H_2O	0.01	0.03	0.00	0.00	0.00	0.17	0.04	0.00	0.00
	NH_3	0.05	0.00	0.04	0.04	0.00	0.01	0.00	0.12	0.02
	P_{total}	0.00	0.00	0.01	0.00	0.04	0.00	0.01	0.05	0.51

Table VII-18. Correlation between analyte concentration and prediction residuals from PLS regression models utilizing the optimized spectral ranges. Gray cells contain the correlation coefficient for the relationship between analyte reference and analyte residual concentrations. All other cells provide the correlation between the analyte reference and interferent residual concentrations.

		Residual Species								
	Species	CH_3CH_2CHO	CH_3CH_2CN	CH_3CH_2OH	CH_3CHO	CH_3CN	CH_3OH	H_2O	NH_3	P_{total}
Analyte Species	CH_3CH_2CHO	0.14	0.18	0.05	0.02	0.06	0.01	0.00	0.04	0.07
	CH_3CH_2CN	0.01	0.02	0.00	0.00	0.00	0.00	0.02	0.01	0.01
	CH_3CH_2OH	0.00	0.03	0.12	0.01	0.01	0.02	0.01	0.04	0.01
	CH_3CHO	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.03	0.00
	CH_3CN	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.02	0.01
	CH_3OH	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00
	H_2O	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00
	NH_3	0.00	0.00	0.00	0.00	0.01	0.00	0.09	0.07	0.00
	P_{total}	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.07	0.10

Table VII-19. Correlation between analyte concentration and prediction residuals from PLS regression models utilizing the optimized spectral ranges and DFF pre-processing. Gray cells contain the correlation coefficient for the relationship between analyte reference and analyte residual concentrations. All other cells provide the correlation between the analyte reference and interferent residual concentrations.

		Residual Species								
	Species	CH_3CH_2CHO	CH_3CH_2CN	CH_3CH_2OH	CH_3CHO	CH_3CN	CH_3OH	H_2O	NH_3	P_{total}
Analyte Species	CH_3CH_2CHO	0.22	0.00	0.00	0.06	0.00	0.07	0.00	0.00	0.01
	CH_3CH_2CN	0.00	0.06	0.00	0.01	0.00	0.00	0.02	0.00	0.00
	CH_3CH_2OH	0.01	0.00	0.06	0.00	0.02	0.02	0.01	0.02	0.00
	CH_3CHO	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.02
	CH_3CN	0.00	0.03	0.00	0.07	0.00	0.00	0.00	0.00	0.01
	CH_3OH	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
	H_2O	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00
	NH_3	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.05
	P_{total}	0.00	0.01	0.02	0.06	0.00	0.00	0.00	0.00	0.13

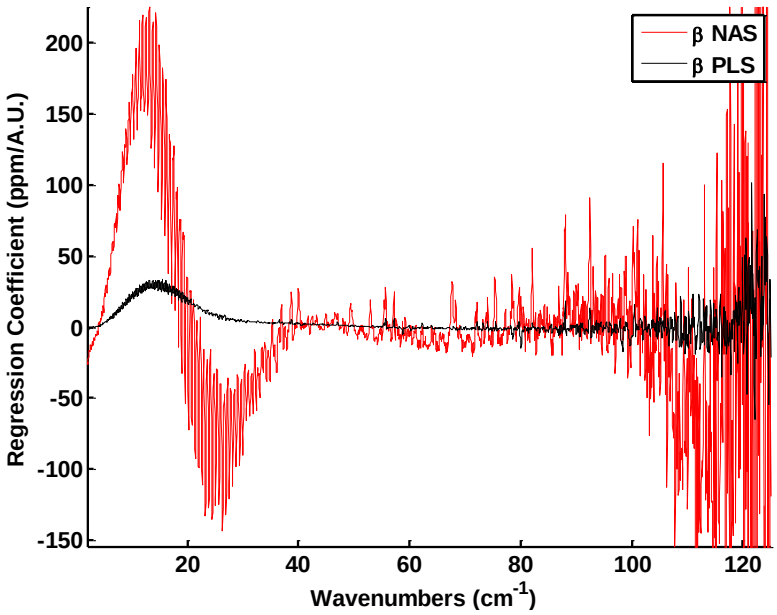


Figure VII-32. Regression coefficients from NAS and PLS models constructed for propionitrile using the full spectral range.

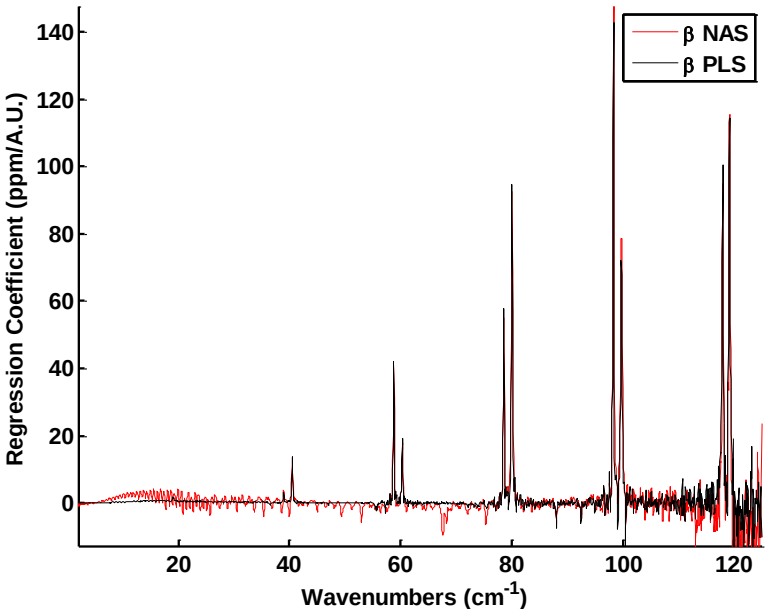


Figure VII-33. Regression coefficients from NAS and PLS models constructed for ammonia using the full spectral range.

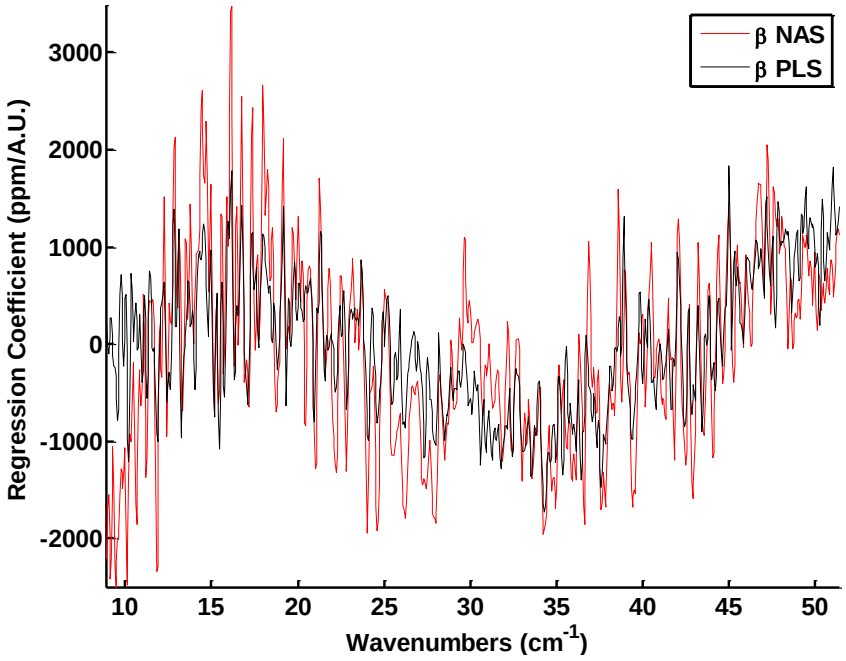


Figure VII-34. Regression coefficients from NAS and PLS models constructed for propionitrile using the optimized spectral range.

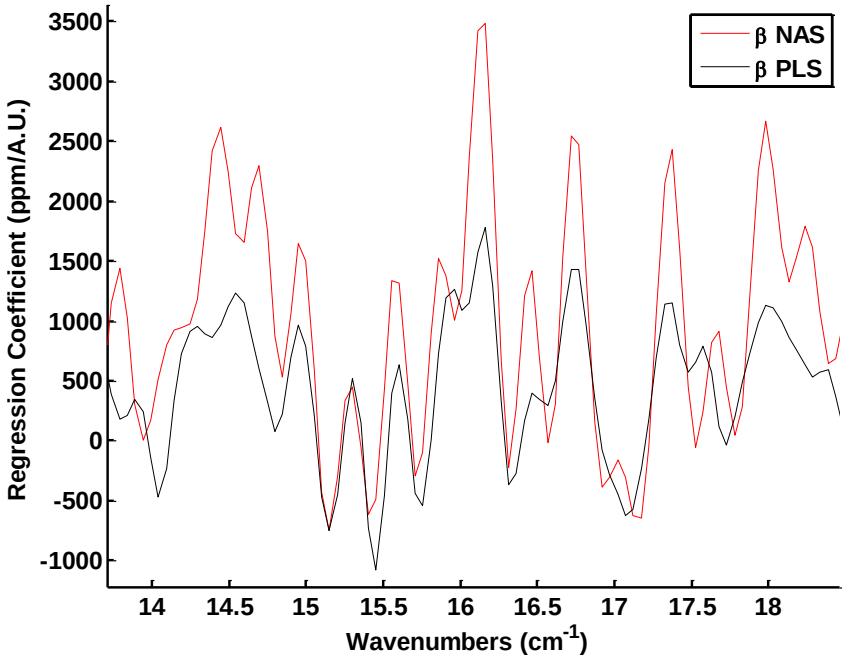


Figure VII-35. Expanded view of regression coefficients from NAS and PLS models constructed for propionitrile using the optimized spectral range.

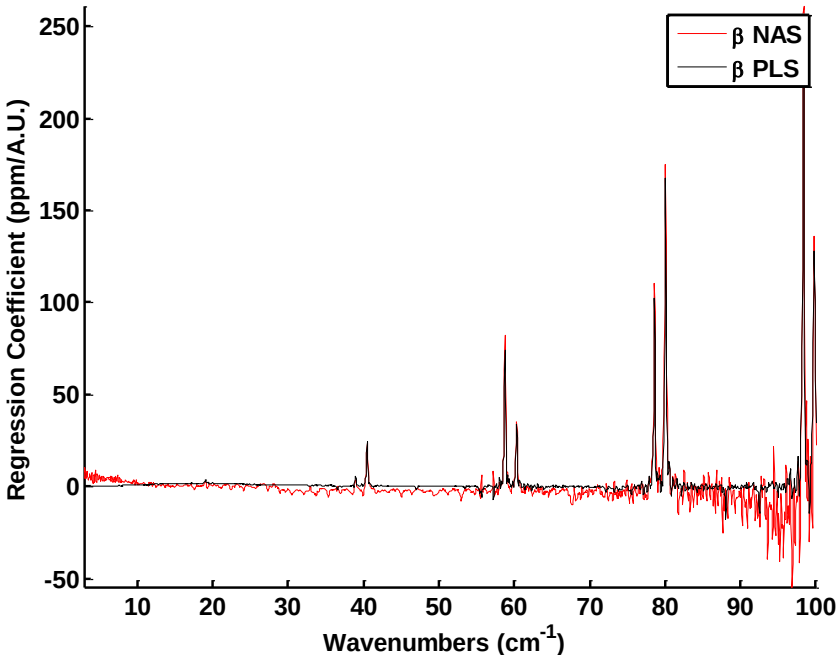


Figure VII-36. Regression coefficients from NAS and PLS models constructed for ammonia using the optimized spectral range.

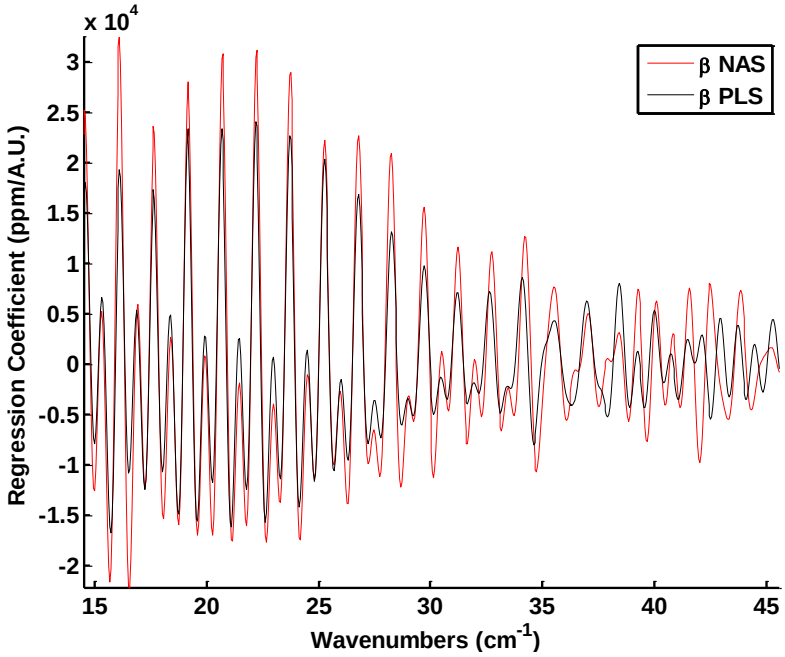


Figure VII-37. Regression coefficients from NAS and PLS models constructed for propionitrile using the optimized spectral range and DFF pre-processing.

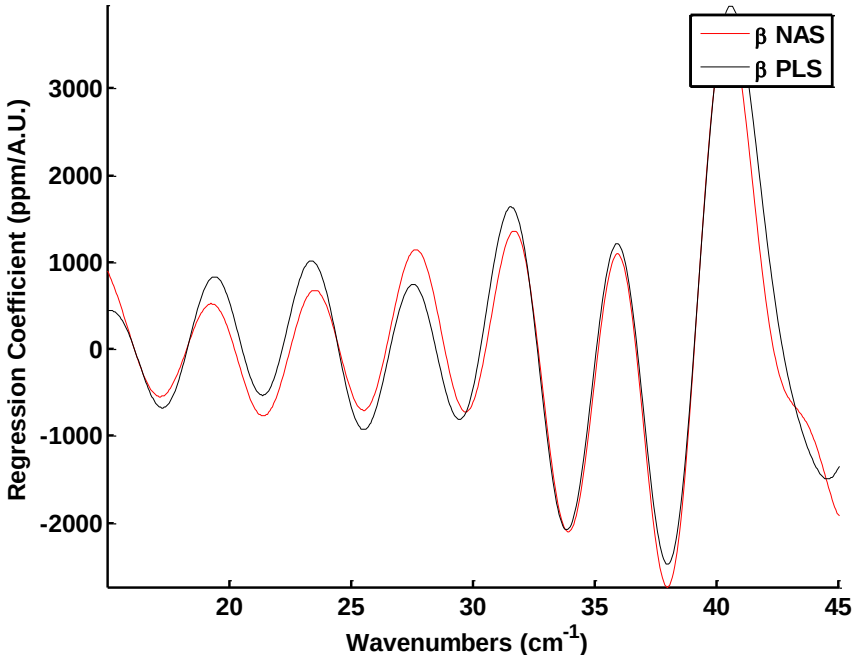


Figure VII-38. Regression coefficients from NAS and PLS models constructed for ammonia using the optimized spectral range and DFF pre-processing.

though the shapes in the vectors are not in complete agreement. This is thought to be a result of the changes in spectral overlap resulting from changes in the overall sample pressure. Similarity between the regression vectors calculated with DFF pre-processing suggests the DFF method is able to compensate for the effect of spectral broadening and changing peak widths. It is important to note that the NAS vectors were generated using spectra acquired at 1 atm overall pressure. Due to the nature of the NAS algorithm, regression vectors could not be produced which would simultaneously represent the sensitivity for analyte concentration and overall pressure.

The effect of spectral broadening on the precision of the regression models was further investigated by performing an F-test to compare the errors of prediction obtained with models developed for samples analyzed at 1 atm overall pressure and those analyzed in varying overall pressures. Calculated F values are shown in Table VII-20. For all species except propionitrile, the error of prediction increased in models developed with

varying overall pressure. Precision was improved for models constructed for propionitrile, and the improvement was found to be significant at the 95% confidence level as determined by the F test. Despite these increases, the errors of prediction for most of the PLS regression models were well below the PELs, except in the case of ammonia.

Table VII-20. Calculated *F* values for comparison of the errors of prediction obtained with various regression models developed for samples at 1 atm overall pressure and varying overall pressures. The critical F value is 1.51, and values exceeding this limit are highlighted in yellow.

Species	RAW	WN	DFF
CH_3CH_2CHO	1.06	1.07	1.48
CH_3CH_2CN	1.15	4.79	1.40
CH_3CH_2OH	1.56	4.61	3.02
CH_3CHO	1.71	4.00	5.64
CH_3CN	1.30	4.84	30.25
CH_3OH	1.37	2.94	2.47
H_2O	1.30	1.89	1.56
NH_3	1.78	4.00	2.78

Conclusions

Precision of the quantitative models constructed with the optimized wavenumber range, and in some cases with DFF pre-processing, was demonstrated despite varying spectral overlap. Most errors of prediction were well below the PELs for the respective species, reinforcing the utility of this method for atmospheric analyses. Pre-processing of spectral data with DFF increased the error of prediction for some species when compared to the error of prediction obtained with wavenumber-optimized models. Further study might include a more thorough optimization process such as a grid search or NMS

optimization with an increased number of starting guesses covering a broader range of wavenumber bounds and digital filter parameters. The change in the precision of these models did not appear to be related to trends in the relationship between NAS and PLS regression vectors. Regression vectors calculated using the NAS and PLS algorithms show some similarity to those calculated at 1 atm overall pressure, but shapes vary in some cases as a result of peak broadening and the related spectral overlap. These experiments represent a further step in evaluation of the feasibility of real-life atmospheric measurement using the THz-TDS apparatus.

CHAPTER VIII

IMPACT OF SPECTRAL OVERLAP ON PLS REGRESSION MODELS

Spectral Resolution and Precision of PLS Models

As mentioned in the comparison of THz instrumentation with alternative methods, the spectral resolution represents one of several experimental parameters which can be adjusted to improve the predictive performance of quantitative models. Experimental data from the sections describing analysis of sample mixtures (both at 1 atm and at various overall pressures) were used for this investigation. As mentioned in Chapter II, the resolution of spectra calculated via the FFT algorithm is dependent on the length of the time domain signal, more specifically the length of the post-peak portion of the signal in the case of the THz-TDS. Truncation of the TDS will thus produce a spectrum with poorer spectral resolution. Though spectral resolution was not considered in the optimization of quantitative models developed to this point, some attention was devoted to examine the possibility of further improvement of model performance.

Experimental

Spectral resolution values between 0.225 and 2.0 cm^{-1} were achieved by truncating the post-peak portion of the sample mixture TDSs to the appropriate lengths. Absorbance spectra were calculated from the truncated signals, and the extent of TDS extension was varied to provide consistent data point spacing between all spectra, regardless of spectral resolution. Baseline correction was carried out with these spectra as with the results reported for the mixtures when analyzed at 0.225 cm^{-1} spectral resolution. Quantitation was performed with PLS regressions using the optimized parameters for each respective data set (wavenumber region and DFF parameters, if applicable). Errors of prediction were used to calculate a CVSEP index, as shown in Equation VIII-1.

$$I_{CVSEP} = \frac{CVSEP_{res}}{CVSEP_{0.225}}$$ Equation VIII-1

In this equation, the CVSEP index (I_{CVSEP}) is calculated as the ratio of the CVSEP calculated at 0.225 cm^{-1} resolution ($CVSEP_{0.225}$) and the CVSEP calculated at the modified resolution ($CVSEP_{res}$). Thresholds of significance were determined by taking the square root of the critical F value as well as the reciprocal of the square root of the critical F value, representing a significant increase or a significant decrease in the error of prediction, respectively.

Results

The CVSEP index was calculated for the three processing methods (full range, optimized range, and digital filtered optimized range) for mixtures at 1 atm overall pressure and mixtures at various overall pressures. Traces of the CVSEP index as a function of the spectral resolution are provided in Figures VIII-1 through VIII-6. Slight changes in the error of prediction are observed when changing the resolution of spectra used for PLS regression across the full spectral range. No significant gains are made for samples analyzed at 1 atm overall pressure, and a significant improvement is only observed for quantitation of methanol at various overall pressures. Given the extreme difference in the errors of prediction of the full range and optimized range/filtered spectra, the change in resolution is not expected to provide any practical benefit when utilizing the entire spectral range.

Changes in the CVSEP index are greater when considering the results obtained with the optimized spectral range. At 1 atm overall pressure, the CVSEP increases significantly for several species at high spectral resolution values. These are species which exhibit high spectral overlap at increased resolution, but diminished overlap near 0.225 cm^{-1} resolution. For samples analyzed at various overall pressures, alteration of the spectral resolution provides significant gains in the predictive ability of the PLS models

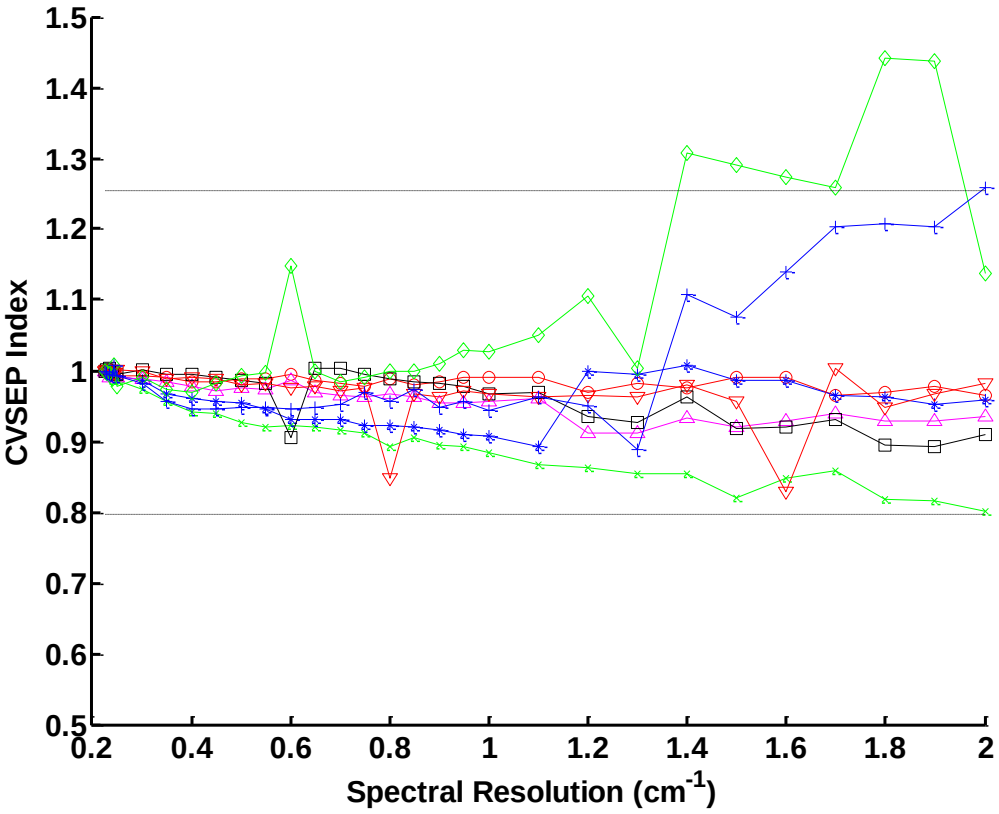


Figure VIII-1. CVSEP index calculated at various resolutions using the full spectral range with ppm-range sample mixtures at 1 atm overall pressure. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, and ammonia by blue +'s.

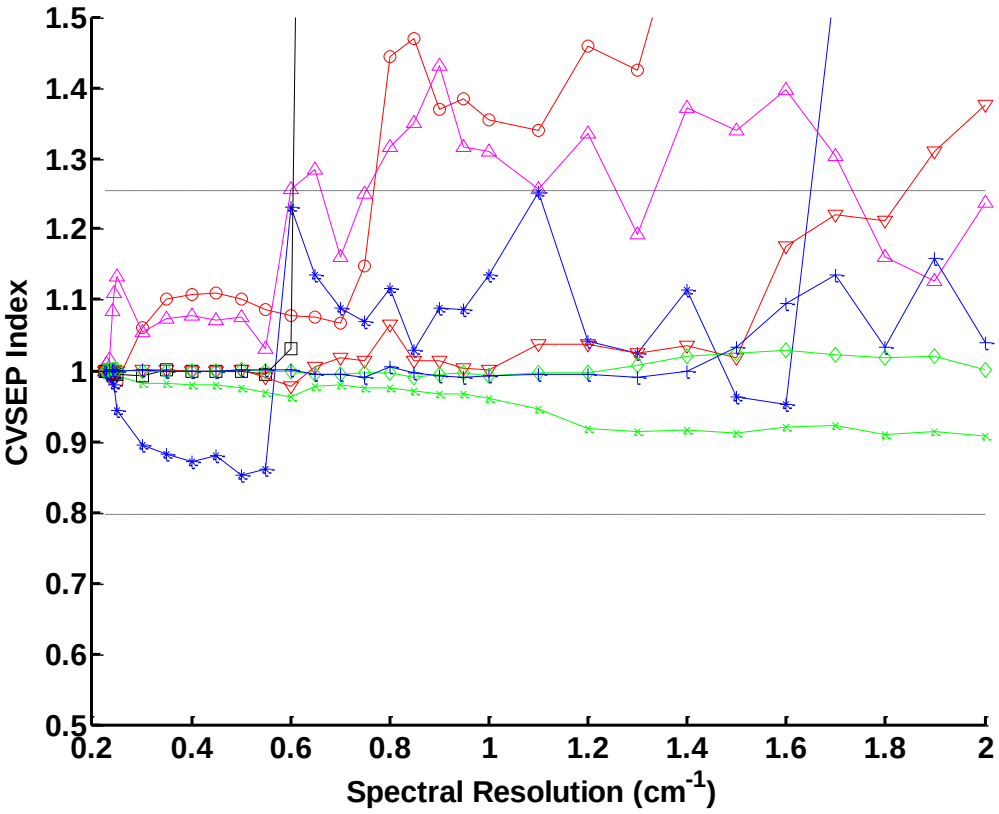


Figure VIII-2. CVSEP index calculated at various resolutions using optimized spectral ranges with ppm-range sample mixtures at 1 atm overall pressure. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, and ammonia by blue +'s.

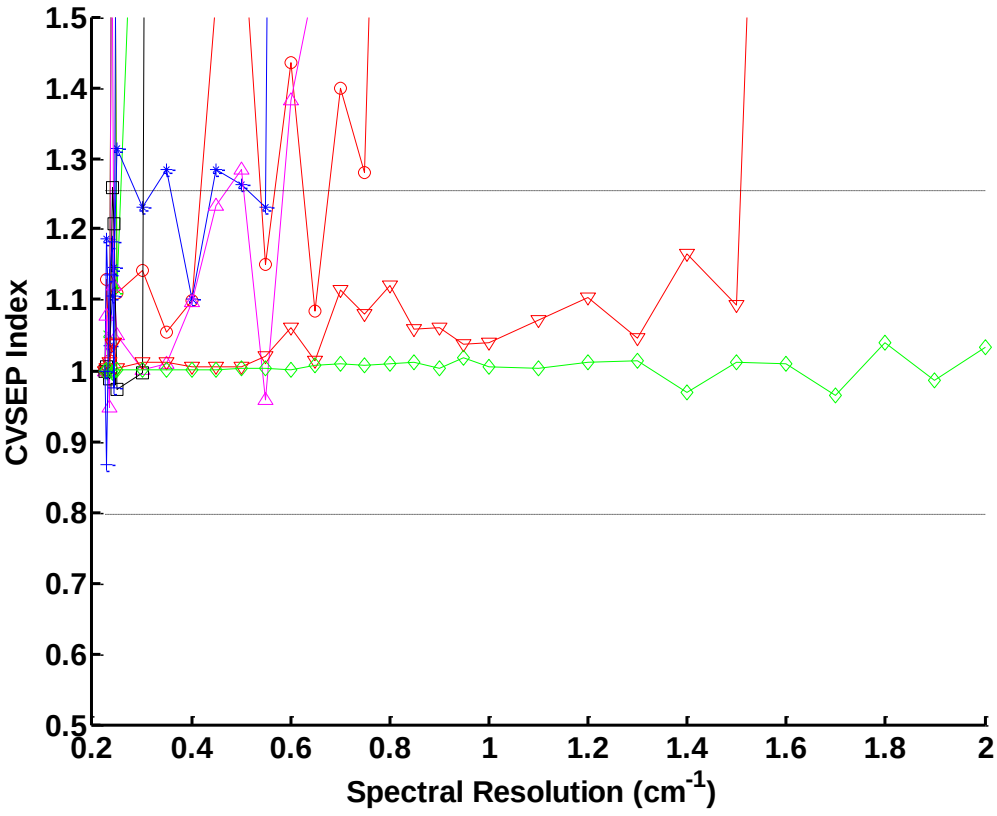


Figure VIII-3. CVSEP index calculated at various resolutions using optimized spectral ranges and DFF with ppm-range sample mixtures at 1 atm overall pressure. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, and ammonia by blue +'s.

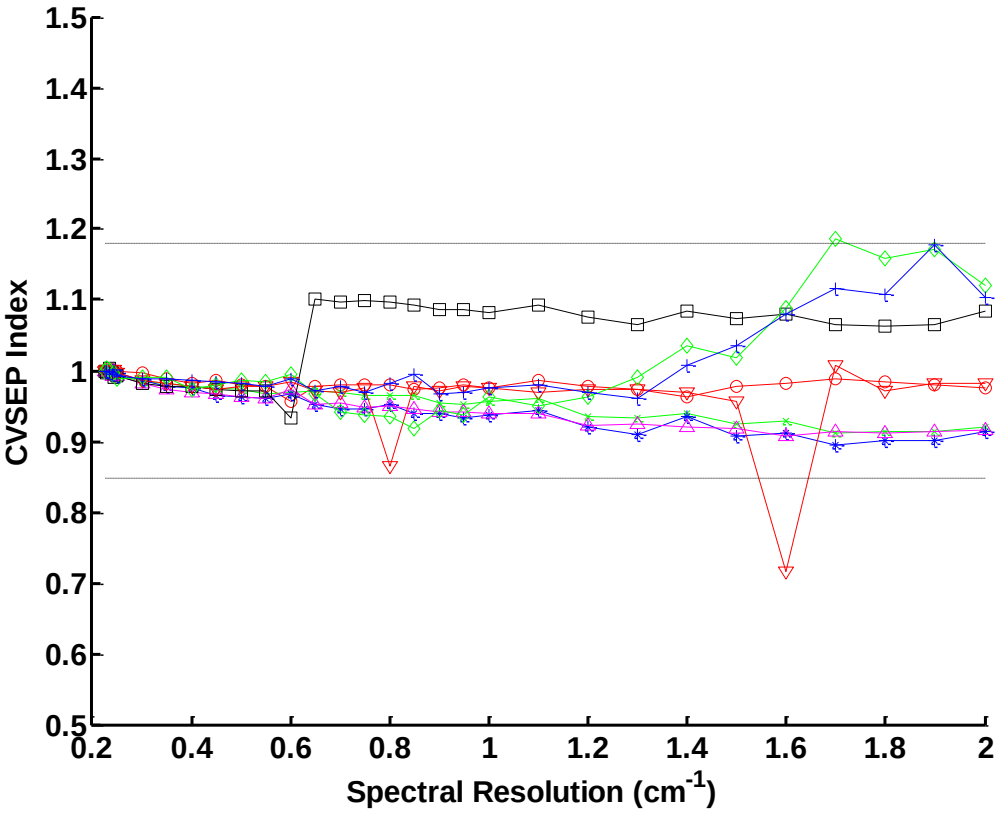


Figure VIII-4. CVSEP index calculated at various resolutions using the full spectral range with ppm-range sample mixtures at various overall pressures. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, and ammonia by blue +'s.

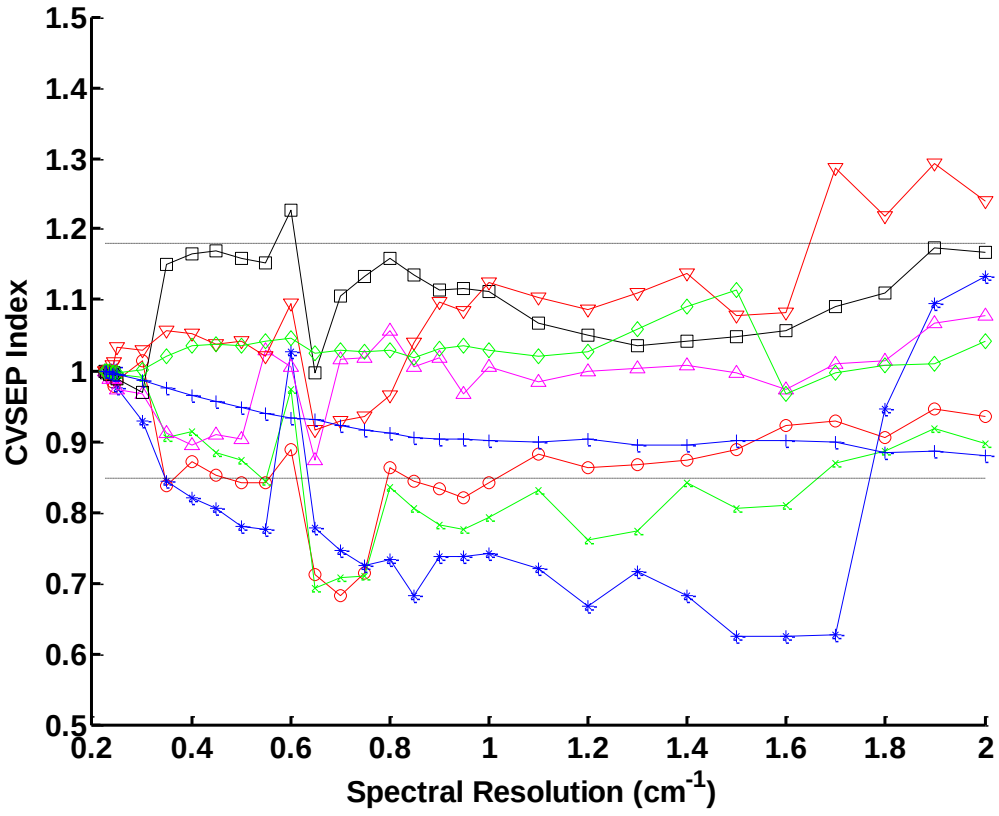


Figure VIII-5. CVSEP index calculated at various resolutions using optimized spectral ranges with ppm-range sample mixtures at various overall pressures. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, and ammonia by blue +'s.

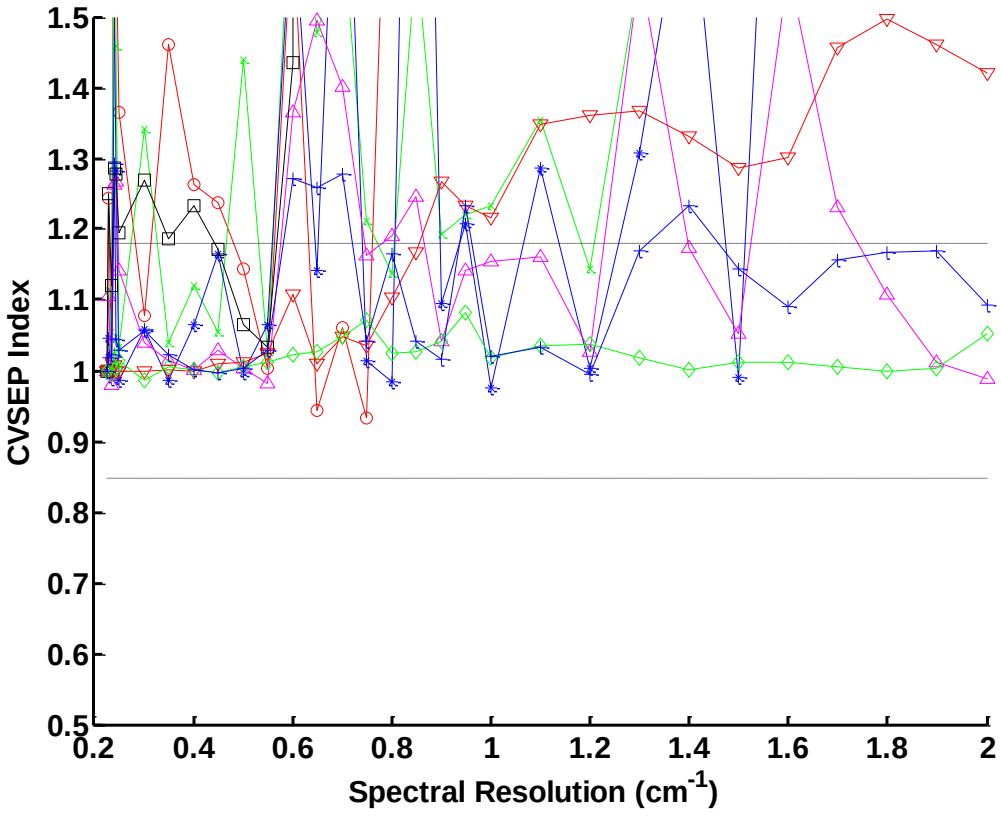


Figure VIII-6. CVSEP index calculated at various resolutions using optimized spectral ranges and DFF with ppm-range sample mixtures at various overall pressures. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, and ammonia by blue +'s.

for some species (propionaldehyde, propionitrile, and ethanol). Errors of prediction for other species undergo slight changes, but few significant increases in the CVSEP are observed, particularly at the lower resolution values.

Behavior of the models becomes somewhat unpredictable when spectral resolution is altered prior to digital filtering. This is a result of the change in digital frequency content for each analyte as the resolution is changed. In cases where the optimized digital filter centered on relatively high digital frequencies, the chemical content of this portion of the digital frequency spectrum is diminished or eliminated

altogether, as increased spectral resolution decreases spectral noise as well as the magnitude of narrow spectral features.

The effect of spectral overlap and peak width on the precision of the PLS model may also be assessed through visualization of concentration residuals for each of the models. Figures VIII-7 through VIII-9 display the residual concentrations as a function of overall pressure for the three methods.

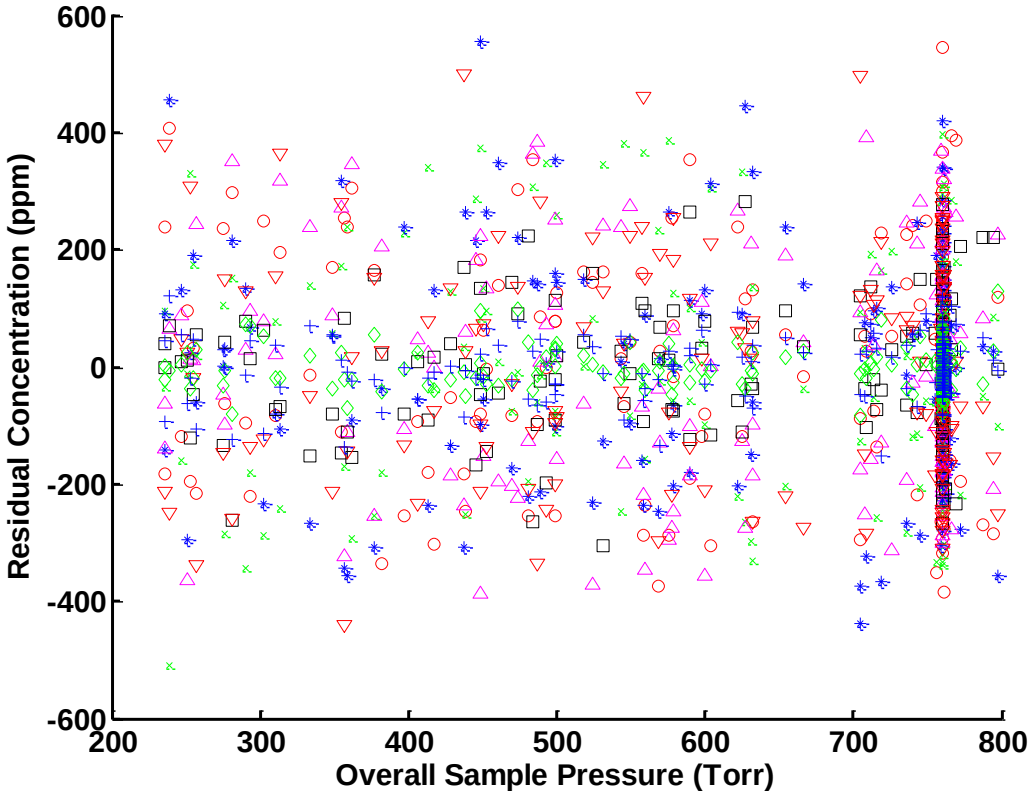


Figure VIII-7. Residual concentrations from PLS models constructed using the full spectral range plotted as a function of overall pressure. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, and ammonia by blue +'s. The tight band of points at the right side of the plot contains residuals from samples analyzed at one atmosphere overall pressure.

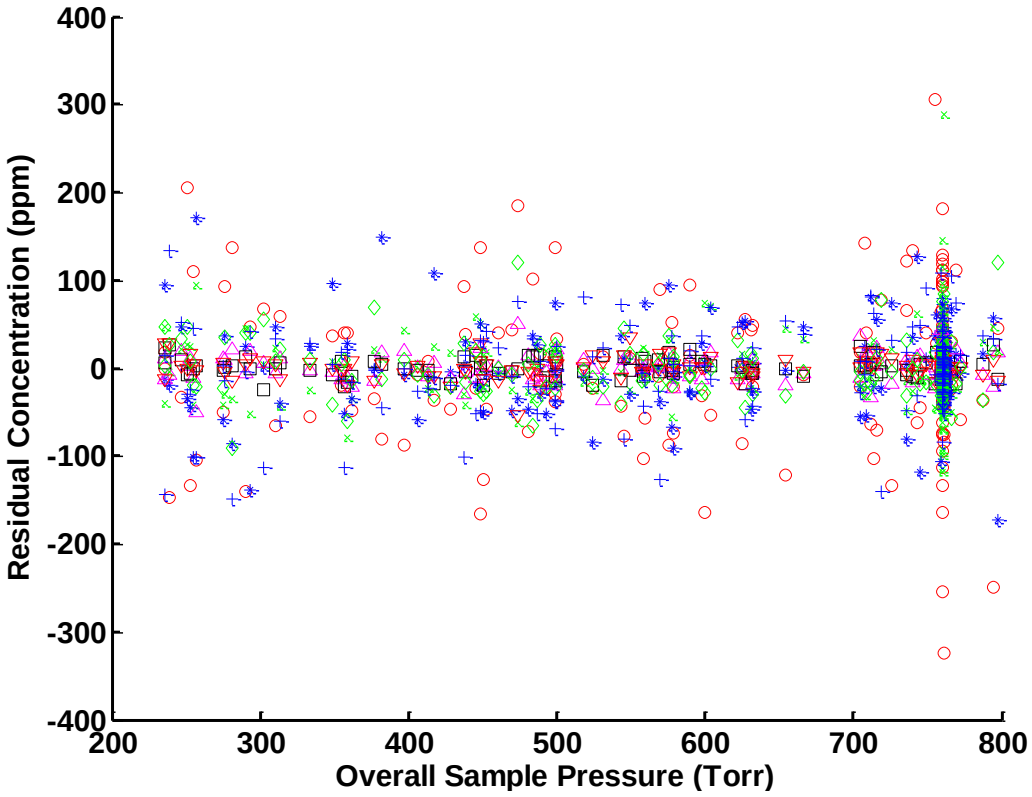


Figure VIII-8. Residual concentrations from PLS models constructed using the optimized spectral range plotted as a function of overall pressure. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, and ammonia by blue +'s. The tight band of points at the right side of the plot contains residuals from samples analyzed at one atmosphere overall pressure.

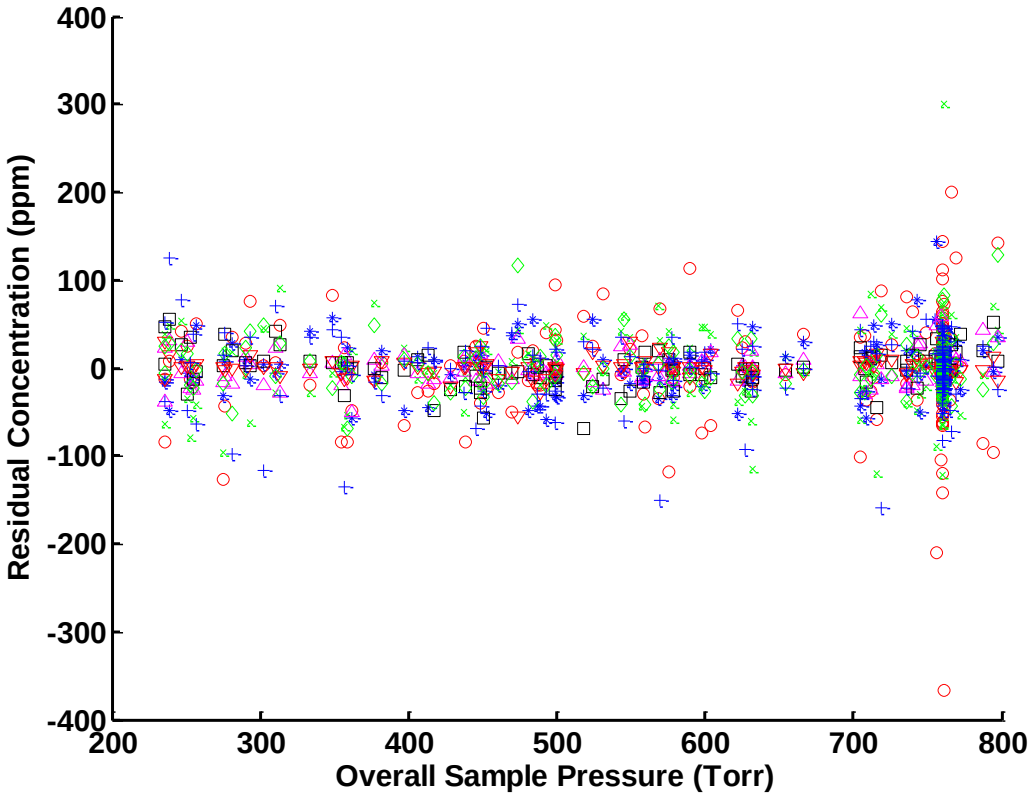


Figure VIII-9. Residual concentrations from PLS models constructed using the optimized spectral range and digital filter plotted as a function of overall pressure. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, and ammonia by blue +'s. The tight band of points at the right side of the plot contains residuals from samples analyzed at one atmosphere overall pressure.

As described in Chapter VII, the residuals decrease in magnitude when wavenumber optimization and digital filtering are employed. No strong relationship is apparent in these data. To simplify visualization, the absolute value of the residuals was obtained by taking the square root of the squared residual concentrations. The mean of residuals for all species of interest were averaged, producing one mean residual magnitude value for each sample. The mean residual magnitudes are plotted as functions of the overall pressure in Figures VIII-10 through VIII-12.

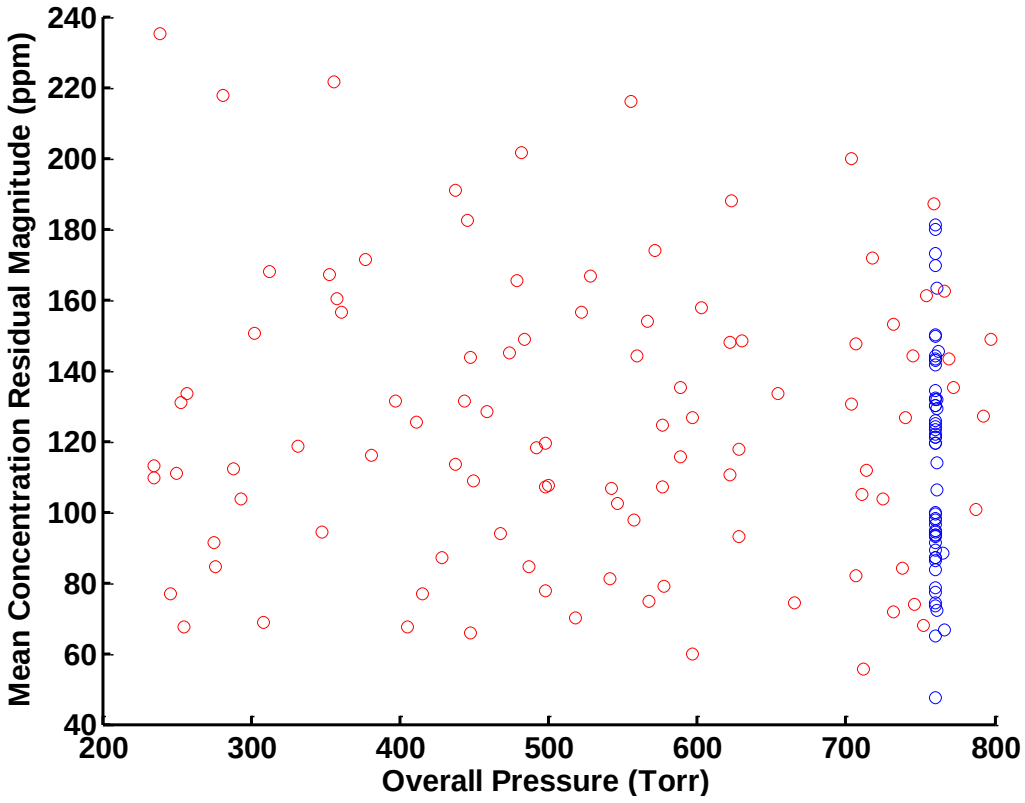


Figure VIII-10. Mean residual concentration magnitude from PLS regression models constructed using the full spectral range plotted as a function of overall sample pressure. Red circles are residuals from samples analyzed at various overall pressures and blue circles are residuals from samples analyzed at 1 atm overall pressure.

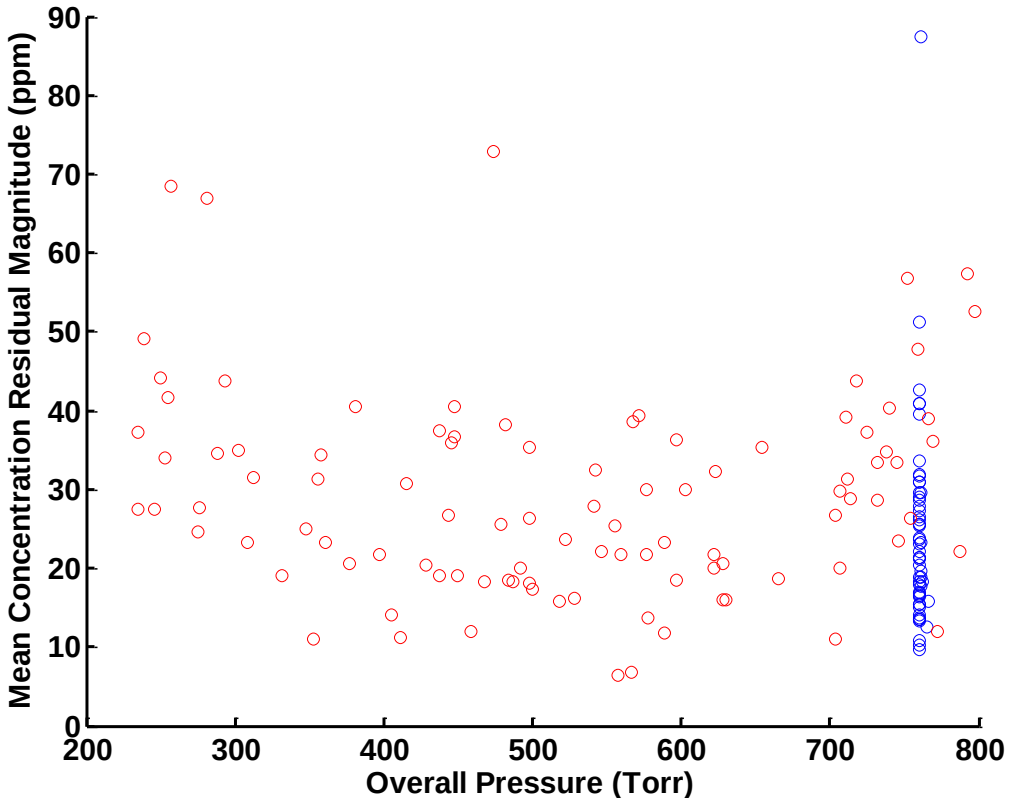


Figure VIII-11. Mean residual concentration magnitude from PLS regression models constructed using the optimized spectral range plotted as a function of overall sample pressure. Red circles are residuals from samples analyzed at various overall pressures and blue circles are residuals from samples analyzed at 1 atm overall pressure.

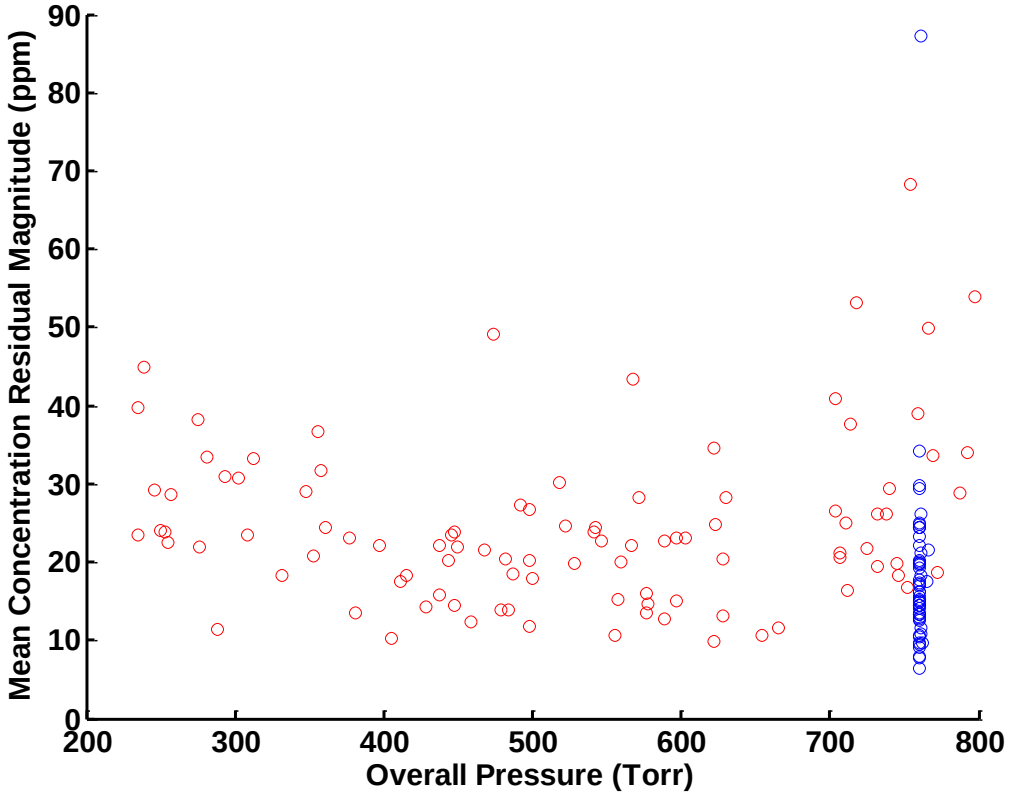


Figure VIII-12. Mean residual concentration magnitude from PLS regression models constructed using the optimized spectral range and digital filter plotted as a function of overall sample pressure. Red circles are residuals from samples analyzed at various overall pressures and blue circles are residuals from samples analyzed at 1 atm overall pressure.

No strong trends are visible in Figure VIII-10, though the mean and spread of concentration residuals for samples collected at 1 atm overall pressure appear smaller than those of samples collected at varying overall pressure. Interesting trends are visible in plots produced with concentration residuals from the optimized PLS models. The magnitude of residuals is greatest at the extremes of the examined pressure range and lowest towards the center. Averages of the concentration residual magnitudes appear lower for samples collected at 1 atm overall pressure. The significance of this change was evaluated using the *t*-test, and the means, standard deviations and Student's *t* values (95% confidence level) are provided in Table VIII-1.

Table VIII-1. Means (R_{mean}) and standard deviations (s_R) of concentration residual magnitudes from the PLS regression models produced for the 8-component mixtures.

PLS model	1 atm		Various Pressures			
	R_{mean}	s_R	R_{mean}	s_R	t_{crit}	t_{calc}
RAW	113.6	31.2	125.5	40.4	1.98	2.04
WN	23.0	9.0	29.1	12.8	1.98	3.40
DFF	16.9	6.1	24.6	10.6	1.98	5.71

For all PLS models, the mean magnitude of residuals is significantly higher for samples collected at varying overall pressures, as is the standard deviation of the residual magnitude. The slight U-shape in the residual magnitude may result from failure of the PLS model to adapt to spectral peaks of varying height:width ratios. Regression models for species at various overall pressures which are constructed with optimized parameters provide the best quantitative performance at the mean of the overall pressure of the samples, as evidenced by the approximate minimum in the concentration residual magnitudes near 524 Torr.

While it may seem that peak fitting would be a viable alternative given this limitation, the vast number of rotational transitions in any spectral window makes quantitative analysis via peak fitting a virtual impossibility. Furthermore, quantitative analysis via peak fitting would require a comprehensive parameter set for each of the species of interest (including all isotopologues and common vibrational excited states), data which are not currently available from any spectral database. The LMA is incapable of selectively discriminating against background contributors, which would likely result in poorer model precision than those reported for PLS models.

Further evaluation of the sensitivity of the PLS algorithm to the peak height:width ratio was performed by comparison of concentration residual magnitudes from spectra calculated with varied resolution. Concentration residuals were taken from the adjusted

resolution which provided the lowest CVSEP index as shown in Figure VIII-5. The magnitudes of these values are plotted in Figure VIII-13, along with quadratic fits of the data illustrating the curvature of the concentration residual magnitudes.

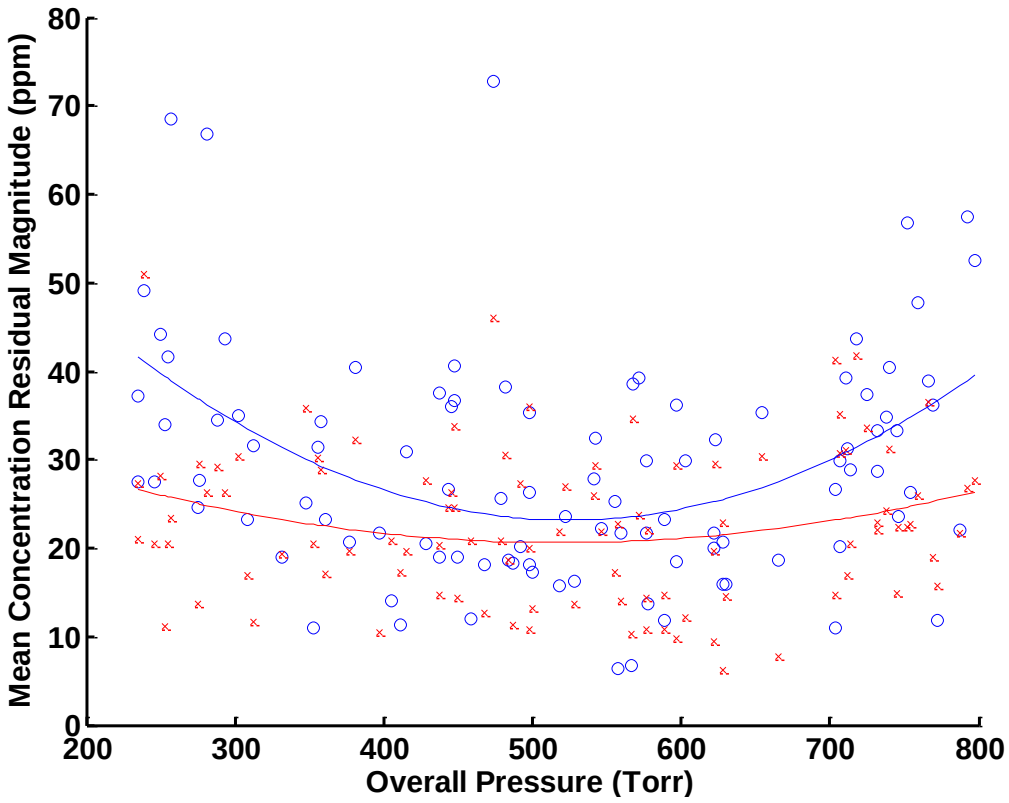


Figure VIII-13. Mean concentration residual magnitudes from PLS models developed with the optimized spectral range (blue o's), and the optimized spectral range with improved spectral resolution (red x's). Quadratic fits of the data are shown as solid lines in the respective colors.

Changes in the spectral resolution decrease the curvature of the mean residual magnitudes, an indicator of diminished sensitivity to peak height:width ratio variability. As with the data reported previously, the means of these residual magnitudes were compared using the *t*-test, the results of which are provided in Table VIII-2.

Means of the mean concentration residual magnitudes are significantly different,

Table VIII-2. Mean of mean concentration residual magnitudes (R_{mean}) and standard deviation of mean concentration residual magnitudes (s_R) from models constructed with the optimized spectral range as well as those constructed with the optimized spectral range with the improved spectral resolution. Student's *t* values (t_{crit} and t_{calc}) are provided in the two right-most columns.

	R_{mean}	s_R	t_{crit}	t_{calc}
WN	29.1	12.8	1.97	4.18
WN, res	22.6	8.7		

further reinforcing the benefit of spectral resolution optimization. Optimization of spectral resolution decreases the magnitude of concentration residuals while also decreasing the spread of these values, which is manifested as a U-shaped curvature illustrated in Figure VIII-13. These data demonstrate the benefit of spectral resolution optimization for atmospheric sensing applications which utilize PLS regressions.

Conclusions

Changing the spectral resolution can provide slight improvement for some species, though based on the lack of unified trends resolution should be optimized on an analyte-by-analyte basis. Greatest improvements are observed when the resolution is altered for PLS models constructed with optimized spectral ranges. Changes in the predictive performance of models using DFF in combination with wavenumber optimization exhibited unpredictable changes, and almost always resulted in significant increases in the CVSEP. Alteration of the spectral resolution has the potential to provide practical benefit both through the improvement of quantitative performance as well as reduction in the time of spectral acquisition. Changes in the height:width ratio of spectral peaks appear to negatively impact the precision of PLS regression models, as evidenced by analysis of concentration residual magnitudes. Despite this drawback, the algorithm is still expected to provide superior performance to peak fitting methods due to its ability to

discriminate against background contributors. Precision might be improved by generating quantitative models with an algorithm such as PLS-2, which has the ability to simultaneously model two components, in this case the analyte and the overall pressure.

Evaluation of Chemical Selectivity via PCSA

Efficacy of any quantitative measurement is dependent on chemical selectivity. In terms of spectroscopic methods, selectivity is demonstrated by effective discrimination between species based on spectral features. As mentioned in Chapter III, the PCSA procedure is useful for evaluation of chemical selectivity. This procedure also mimics a scenario in which a calibration set is constructed in the lab with a multitude of species which are expected to be present in the measurement at some point in time. The calibration set may be put to use in predictions where some, or only one of the species included in the calibration set is present at the time of the measurement.

Experimental

Calibration was carried out with the PLS algorithm using the three methods described previously in which the full spectral range, optimized spectral ranges, and optimized DFF pre-processing and wavenumber bounds were utilized. Spectra collected at 1 atm overall pressure and spectra collected at varying overall pressures were used separately to ascertain the effect of varying spectral overlap on the observed chemical selectivity. The concentrations of the pure component spectra were predicted for each analyte in all combinations, such that prediction was carried out with each of the eight sets of pure component spectra after calibration was performed for one analyte with mixture spectra. By repeating this process, a total of 64 calibration/prediction processes are carried out, with 8 of these representing analyte-analyte pairs, and 56 representing analyte-interferent pairs.

Linear regressions were constructed with the concentration correlation from each of the 64 sets, where a slope of 1 is expected for analyte-analyte pairs and a slope of 0 is

expected for analyte-interferent pairs where complete chemical selectivity is achieved. As chemical selectivity decreases, the slopes of analyte-interferent pairs will increase, and the slopes of analyte-analyte pairs will decrease, providing a quantitative metric of the chemical selectivity in realistic measurements. The slopes calculated for each data set (1 atm and various overall pressures) were compared using the *t*-test.

Results

A representative set of PCSA concentration correlation plots is provided in Figures VIII-14 through VIII-16. Slopes of the concentration correlation plots are listed for each data set in Tables VIII-3 through VIII-8, with the first three tables representing PCSA results using samples prepared at 1 atm overall pressure, and the latter three representing those prepared at various overall pressures. Chemical selectivity increases when wavenumber or DFF and wavenumber optimization are utilized in construction of the PLS regression models, as evidenced by the increased slopes of the analyte-analyte concentration correlations and the decreased slopes of the analyte-interferent concentration correlations. Many of the analyte-interferent slopes are near zero or are not significantly different from zero. As described earlier, it is highly desirable for these slopes to equal zero, as this is indicative of high chemical selectivity.

Slopes listed in Tables VIII-3 through VIII-8 were compared in two manners. Comparison was performed between the slopes calculated with different model construction methods (full range, optimized range, and optimized range/DFF parameters). Figures VIII-14 through VIII-16 demonstrate these changes graphically. Results of the *t*-test indicate significant differences in chemical selectivity with incorporation of wavenumber selection or DFF/wavenumber optimization. Significant differences in chemical selectivity are observed in comparison of models constructed with the optimized spectral range and those constructed with wavenumber and DFF optimization in several cases.

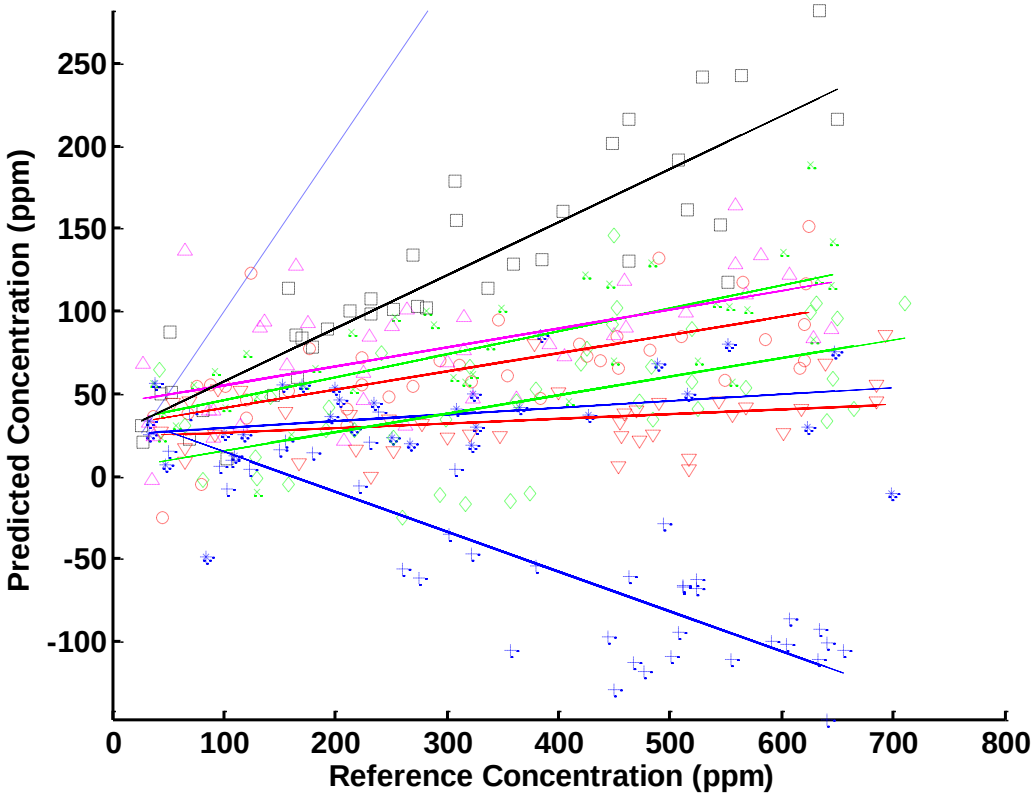


Figure VIII-14. PCSA concentration correlation plot of pure component concentrations predicted by a PLS regression model constructed for acetonitrile using the full spectral range. The overall pressure of the calibration data was 1 atm. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, and ammonia by blue +'s. The blue dotted line represents the ideal case where the slope is 1, and the y-intercept is 0.

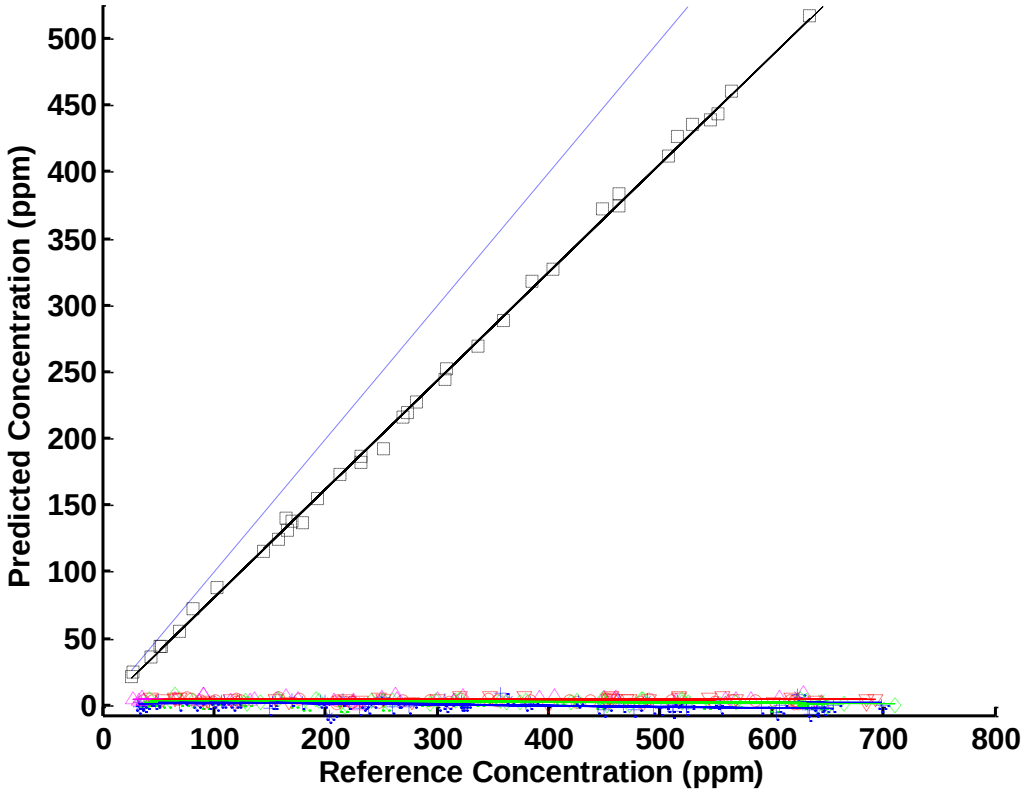


Figure VIII-15. PCSA concentration correlation plot of pure component concentrations predicted by a PLS regression model constructed for acetonitrile using the optimized spectral range. The overall pressure of the calibration data was 1 atm. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, and ammonia by blue +'s. The blue dotted line represents the ideal case where the slope is 1, and the y-intercept is 0.

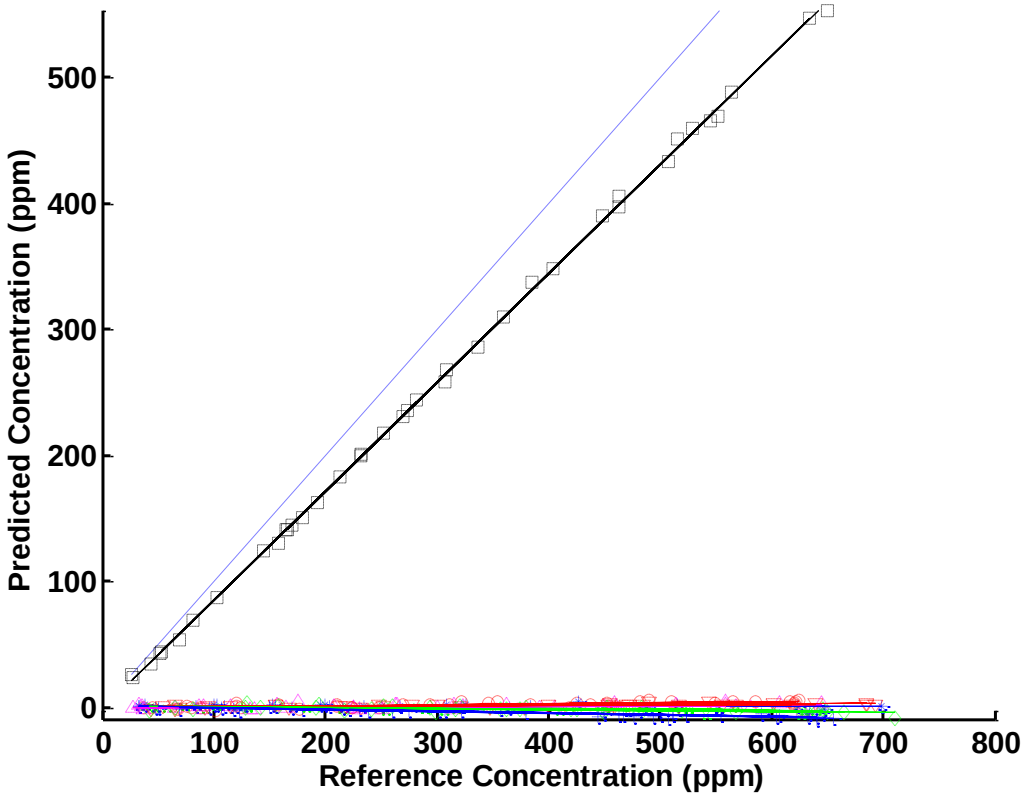


Figure VIII-16. PCSA concentration correlation plot of pure component concentrations predicted by a PLS regression model constructed for acetonitrile using the optimized spectral range and DFF pre-processing. The overall pressure of the calibration data was 1 atm. Propionaldehyde is represented by red circles, propionitrile by green x's, ethanol by blue stars, acetaldehyde by magenta triangles, acetonitrile by black squares, methanol by red inverted triangles, water by green diamonds, and ammonia by blue +'s. The blue dotted line represents the ideal case where the slope is 1, and the y-intercept is 0.

Table VIII-3. Slopes and uncertainties from linear regressions of PCSA concentration correlation data. Calibration and prediction were performed using the full spectral range. Spectra collected at 1 atm overall pressure were used for calibration.

		Pure analyte (Prediction)							
	Species	CH_3CH_2CHO	CH_3CH_2CN	CH_3CH_2OH	CH_3CHO	CH_3CN	CH_3OH	H_2O	NH_3
Analyte in Mixture (Calibration)	CH_3CH_2CHO	0.05 ±0.03	0.05 ±0.02	0.007 ±0.02	0.08 ±0.03	0.16 ±0.04	0.05 ±0.02	0.31 ±0.04	0.02 ±0.02
	CH_3CH_2CN	0.10 ±0.02	0.16 ±0.03	0.03 ±0.03	0.13 ±0.02	0.31 ±0.03	0.003 ±0.02	0.003 ±0.03	-0.33 ±0.03
	CH_3CH_2OH	-0.02 ±0.04	0.06 ±0.05	0.07 ±0.08	-0.04 ±0.07	-0.06 ±0.06	-0.004 ±0.04	0.02 ±0.05	-0.32 ±0.08
	CH_3CHO	0.10 ±0.02	0.05 ±0.02	0.04 ±0.03	0.08 ±0.03	0.22 ±0.04	0.04 ±0.03	0.006 ±0.03	-0.28 ±0.03
	CH_3CN	0.11 ±0.02	0.14 ±0.02	0.04 ±0.03	0.11 ±0.03	0.32 ±0.03	0.03 ±0.01	0.11 ±0.03	-0.24 ±0.02
	CH_3OH	0.03 ±0.03	-0.01 ±0.03	0.08 ±0.03	0.06 ±0.04	0.06 ±0.03	0.14 ±0.03	0.17 ±0.03	0 ±0.04
	H_2O	0.005 ±0.01	-0.007 ±0.01	0.04 ±0.02	0.01 ±0.02	0.006 ±0.03	0.06 ±0.01	0.63 ±0.03	0.09 ±0.02
	NH_3	-0.02 ±0.01	-0.02 ±0.01	0.02 ±0.02	0.02 ±0.02	0.007 ±0.02	0.02 ±0.01	-0.02 ±0.02	1.07 ±0.02

Table VIII-4. Slopes and uncertainties from linear regressions of PCSA concentration correlation data. Calibration and prediction were performed using the optimized spectral ranges. Spectra collected at 1 atm overall pressure were used for calibration.

		Pure analyte (Prediction)							
	Species	CH_3CH_2CHO	CH_3CH_2CN	CH_3CH_2OH	CH_3CHO	CH_3CN	CH_3OH	H_2O	NH_3
Analyte in Mixture (Calibration)	CH_3CH_2CHO	0.88 ±0.04	0.0002 ±0.04	0.08 ±0.04	0.02 ±0.03	0.27 ±0.05	-0.08 ±0.04	0.04 ±0.03	0.048 ±0.07
	CH_3CH_2CN	0.16 ±0.01	0.81 ±0.02	-0.04 ±0.03	0.03 ±0.02	0.07 ±0.06	0.03 ±0.01	-0.002 ±0.01	-0.03 ±0.09
	CH_3CH_2OH	-0.007 ±0.01	-0.07 ±0.01	0.74 ±0.03	-0.03 ±0.02	0.04 ±0.02	-0.06 ±0.01	-0.03 ±0.01	0.002 ±0.03
	CH_3CHO	-0.008 ±0.003	-0.013 ±0.003	-0.0009 ±0.004	0.90 ±0.01	0.037 ±0.009	0.024 ±0.003	-0.009 ±0.003	0.005 ±0.01
	CH_3CN	-0.0022 ±0.0009	-0.0007 ±0.002	0.002 ±0.003	0.0008 ±0.002	0.813 ±0.004	-0.0002 ±0.001	-0.004 ±0.002	-0.009 ±0.002
	CH_3OH	-0.001 ±0.001	-0.004 ±0.002	-0.009 ±0.004	0.007 ±0.002	0.003 ±0.002	0.80 ±0.01	-0.029 ±0.002	-0.030 ±0.003
	H_2O	-0.006 ±0.004	0.0001 ±0.003	0.011 ±0.008	-0.003 ±0.003	-0.03 ±0.01	-0.014 ±0.004	0.72 ±0.03	-0.135 ±0.007
	NH_3	0.023 ±0.006	0.001 ±0.004	0.040 ±0.005	0.045 ±0.004	0.0401 ±0.0009	0.033 ±0.004	-0.004 ±0.001	1.16 ±0.02

Table VIII-5. Slopes and uncertainties from linear regressions of PCSA concentration correlation data. Calibration and prediction were performed using the optimized spectral ranges and DFF parameters. Spectra collected at 1 atm overall pressure were used for calibration.

		Pure analyte (Prediction)							
	Species	CH_3CH_2CHO	CH_3CH_2CN	CH_3CH_2OH	CH_3CHO	CH_3CN	CH_3OH	H_2O	NH_3
Analyte in Mixture (Calibration)	CH_3CH_2CHO	0.86 ±0.05	0.03 ±0.04	0.08 ±0.05	0.10 ±0.04	0.08 ±0.05	0.07 ±0.03	0.08 ±0.06	-0.02 ±0.06
	CH_3CH_2CN	0.009 ±0.009	0.96 ±0.01	0.02 ±0.01	-0.0009 ±0.01	-0.02 ±0.02	0.06 ±0.01	0.06 ±0.01	0.005 ±0.01
	CH_3CH_2OH	-0.05 ±0.02	-0.04 ±0.03	0.73 ±0.03	-0.03 ±0.03	0.03 ±0.02	-0.002 ±0.01	-0.05 ±0.02	-0.007 ±0.02
	CH_3CHO	-0.004 ±0.003	-0.016 ±0.004	-0.004 ±0.002	0.89 ±0.01	0.019 ±0.003	0.010 ±0.003	-0.007 ±0.003	0.002 ±0.003
	CH_3CN	0.009 ±0.001	-0.005 ±0.001	-0.0008 ±0.001	0.003 ±0.001	0.863 ±0.003	0.0037 ±0.0009	-0.007 ±0.001	-0.014 ±0.001
	CH_3OH	0.001 ±0.001	-0.001 ±0.001	-0.005 ±0.003	0.004 ±0.001	0.001 ±0.001	0.80 ±0.01	-0.026 ±0.002	-0.032 ±0.003
	H_2O	-0.005 ±0.003	0.002 ±0.003	0.002 ±0.007	0.003 ±0.003	-0.03 ±0.01	-0.002 ±0.004	0.72 ±0.03	-0.134 ±0.006
	NH_3	0.007 ±0.003	0.008 ±0.004	-0.004 ±0.006	0.014 ±0.007	0.027 ±0.008	-0.008 ±0.004	-0.029 ±0.005	0.94 ±0.03

Table VIII-6. Slopes and uncertainties from linear regressions of PCSA concentration correlation data. Calibration and prediction were performed using the full spectral range. Spectra collected at various overall pressures were used for calibration.

		Pure analyte (Prediction)							
	Species	CH_3CH_2CHO	CH_3CH_2CN	CH_3CH_2OH	CH_3CHO	CH_3CN	CH_3OH	H_2O	NH_3
Analyte in Mixture (Calibration)	CH_3CH_2CHO	0.02 ±0.01	0.07 ±0.02	0.01 ±0.02	0.008 ±0.02	0.07 ±0.03	0.009 ±0.01	-0.04 ±0.02	-0.26 ±0.03
	CH_3CH_2CN	0.11 ±0.03	0.13 ±0.02	0.01 ±0.04	0.11 ±0.03	0.25 ±0.04	0.03 ±0.03	0.09 ±0.04	-0.11 ±0.04
	CH_3CH_2OH	0.05 ±0.03	-0.01 ±0.02	0.01 ±0.03	0.05 ±0.03	0.13 ±0.04	0.04 ±0.03	-0.04 ±0.03	-0.29 ±0.03
	CH_3CHO	0.07 ±0.02	0.09 ±0.01	0.02 ±0.02	0.13 ±0.02	0.20 ±0.02	0.04 ±0.01	-0.03 ±0.02	-0.27 ±0.01
	CH_3CN	0.15 ±0.02	0.17 ±0.02	0.04 ±0.02	0.24 ±0.03	0.54 ±0.03	0.03 ±0.02	-0.03 ±0.01	-0.04 ±0.02
	CH_3OH	0.05 ±0.03	0.02 ±0.02	0.05 ±0.03	0.07 ±0.03	0.06 ±0.03	0.13 ±0.02	0.23 ±0.03	-0.03 ±0.02
	H_2O	-0.014 ±0.008	-0.01 ±0.01	0.05 ±0.02	0.0006 ±0.02	-0.02 ±0.02	0.03 ±0.01	0.58 ±0.03	-0.07 ±0.02
	NH_3	0.002 ±0.01	-0.005 ±0.01	0.02 ±0.02	0.03 ±0.01	0.02 ±0.01	0.02 ±0.01	-0.005 ±0.02	1.04 ±0.02

Table VIII-7. Slopes and uncertainties from linear regressions of PCSA concentration correlation data. Calibration and prediction were performed using the optimized spectral ranges. Spectra collected at various overall pressures were used for calibration.

		Pure analyte (Prediction)							
	Species	CH_3CH_2CHO	CH_3CH_2CN	CH_3CH_2OH	CH_3CHO	CH_3CN	CH_3OH	H_2O	NH_3
Analyte in Mixture (Calibration)	CH_3CH_2CHO	0.79 ±0.04	0.008 ±0.05	0.16 ±0.09	-0.09 ±0.06	-0.07 ±0.2	-0.01 ±0.05	0.03 ±0.04	0.24 ±0.23
	CH_3CH_2CN	0.07 ±0.01	0.95 ±0.02	-0.02 ±0.03	0.006 ±0.02	0.06 ±0.07	0.07 ±0.02	0.02 ±0.02	-0.05 ±0.1
	CH_3CH_2OH	0.08 ±0.02	-0.14 ±0.03	0.60 ±0.04	-0.03 ±0.02	0.02 ±0.04	0.008 ±0.03	0.03 ±0.02	-0.11 ±0.05
	CH_3CHO	-0.0001 ±0.004	-0.024 ±0.003	-0.015 ±0.005	0.87 ±0.01	0.076 ±0.006	0.021 ±0.003	-0.011 ±0.004	0.004 ±0.01
	CH_3CN	0.017 ±0.004	-0.015 ±0.004	-0.017 ±0.005	0.005 ±0.005	0.84 ±0.02	-0.004 ±0.004	-0.003 ±0.003	-0.009 ±0.03
	CH_3OH	-0.007 ±0.003	0.003 ±0.002	-0.005 ±0.004	0.017 ±0.003	0.004 ±0.004	0.78 ±0.01	-0.022 ±0.002	-0.012 ±0.006
	H_2O	0.004 ±0.005	0.004 ±0.002	0.027 ±0.008	0.015 ±0.004	0.004 ±0.009	-0.012 ±0.004	0.64 ±0.03	-0.053 ±0.005
	NH_3	0.004 ±0.005	0.004 ±0.003	0.027 ±0.005	0.002 ±0.005	0.013 ±0.005	0.046 ±0.005	0.010 ±0.004	1.03 ±0.02

Table VIII-8. Slopes and uncertainties from linear regressions of PCSA concentration correlation data. Calibration and prediction were performed using the optimized spectral ranges and DFF parameters. Spectra collected at various overall pressures were used for calibration.

		Pure analyte (Prediction)							
	Species	CH_3CH_2CHO	CH_3CH_2CN	CH_3CH_2OH	CH_3CHO	CH_3CN	CH_3OH	H_2O	NH_3
Analyte in Mixture (Calibration)	CH_3CH_2CHO	0.65 ±0.06	0.05 ±0.03	0.03 ±0.08	0.08 ±0.02	0.19 ±0.04	0.007 ±0.02	0.03 ±0.06	-0.009 ±0.03
	CH_3CH_2CN	-0.02 ±0.02	0.69 ±0.02	0.04 ±0.02	0.01 ±0.02	-0.02 ±0.02	0.03 ±0.01	-0.03 ±0.02	0.05 ±0.03
	CH_3CH_2OH	-0.04 ±0.01	-0.07 ±0.01	0.65 ±0.03	-0.01 ±0.02	0.012 ±0.003	0.007 ±0.009	0.05 ±0.01	0.12 ±0.03
	CH_3CHO	-0.003 ±0.003	-0.006 ±0.004	-0.029 ±0.007	0.85 ±0.01	0.005 ±0.007	0.081 ±0.006	0.006 ±0.005	0 ±0.01
	CH_3CN	-0.009 ±0.005	-0.035 ±0.004	-0.003 ±0.008	0.008 ±0.005	0.679 ±0.005	0.002 ±0.005	-0.017 ±0.005	-0.016 ±0.004
	CH_3OH	0.002 ±0.002	0.006 ±0.003	-0.013 ±0.003	0.018 ±0.003	-0.004 ±0.003	0.80 ±0.01	-0.034 ±0.004	-0.029 ±0.003
	H_2O	-0.004 ±0.003	0.0008 ±0.003	0.013 ±0.008	-0.006 ±0.004	0.01 ±0.01	0.007 ±0.004	0.64 ±0.03	-0.113 ±0.009
	NH_3	0.050 ±0.009	0.018 ±0.008	0.06 ±0.01	0.065 ±0.008	0.047 ±0.008	0.034 ±0.006	0.013 ±0.007	1.16 ±0.02

The difference in chemical selectivity is significant in models constructed for propionitrile, acetaldehyde, acetonitrile, and ammonia for both pressure range data sets. Models for ethanol and methanol at various overall pressures also exhibit significant differences in chemical selectivity when comparing wavenumber optimized and DFF/wavenumber optimized models. While these trends reinforce the benefits of wavenumber selection and pre-processing, the changes between models using wavenumber optimization and those using optimized DFF and wavenumber ranges are sensitive to the parameters used in construction of the model and can vary greatly.

Slopes obtained with the different sample pressures (1 atm or various overall pressures) were also compared with the *t*-test. Chemical selectivity of models constructed with the full spectral range changes significantly when the overall pressure of the sample is varied. Differences between the slopes are still significant when comparing the data sets using optimized wavenumber ranges, but the magnitudes of almost all differences are greatly diminished. This trend is broken when comparing models using wavenumber and DFF optimization, as some changes are significant, and some are not. Models for propionaldehyde, propionitrile, acetaldehyde, acetonitrile, and ammonia undergo significant changes in chemical selectivity when the overall pressure is varied. Chemical selectivity of models constructed for ethanol, methanol, and water vapor are insensitive to the change in overall sample pressure. This may contraindicate the need for DFF pre-processing, as minimal change is observed in the chemical selectivity. Likewise, for some PLS regression models, the errors of prediction were not improved significantly (or at all) when DFF pre-processing was utilized.

Conclusions

Chemical selectivity of the quantitative models described in previous sections is demonstrated. Significant differences in chemical selectivity are observed when wavenumber optimization or DFF/wavenumber optimization are utilized in construction

of the regression models. Differences in selectivity between wavenumber-optimized and DFF/wavenumber optimized models were not consistent, perhaps demonstrating the sensitivity of selectivity to the selected digital filters for some species. The effect of the overall sample pressure on chemical selectivity was also verified by comparing the slopes from the PCSA. Slopes obtained from PCSA may provide an additional metric for evaluation of wavenumber bounds or pre-processing parameters during optimization procedures.

CHAPTER IX

SOURCES OF ERROR AND FUTURE STUDIES FOR GAS SAMPLE ANALYSIS

Estimation of Gas-Phase Dimerization

Molecules are known to form dimeric or higher-order polymeric structures while in a gaseous state.[355] Theoretical and experimental work has been devoted to the study of this phenomenon for complexes such as water vapor,[356] methanol,[357] ammonia-water,[358] ammonia,[359, 360] water-methanol,[361] HF-HCl,[362] and HCN-HF,[363] amongst others. Dimer or polymer formation is the result of intermolecular forces such as hydrogen-bonding or dipole-dipole interactions. Though molecules are of rather low density while in the gas-phase, passing molecules can associate with each other under the right conditions. Mechanisms for dimer formation have been reported, and studies suggest that formation only requires collision of two species, as opposed to the previously-held belief that formation was the product of collisions between three molecules.[364] Dimerization of gas-phase species is of interest because different spectral features for the dimeric species represents a potential source of error in the reported measurements.

Curtiss and Blander described several methods for analysis of gas-phase dimer formation such as “gas imperfection”, thermal conductivity, and spectroscopic methods.[355] Gas imperfection analyses involve the study of physical properties of the gaseous mixtures such as the speed of sound, vapor density, or pressure-volume-temperature behavior.[355] Spectroscopy, as Curtiss and Blander point out, should be the most accurate of these methods for determination of the thermodynamic properties of gas-phase dimers because it is a direct method.[355] Furthermore, spectroscopic methods are chemically selective, allowing discrimination between species, while the indirect methods do not provide any meaningful specificity. When working with vibrational spectroscopic methods, it can be difficult to distinguish between the features of the

monomeric and di/polymeric species. Thus, this is an application particularly suited to rotational spectroscopy, as the spectra of the monomeric and complexed species would be markedly different because of the large change in the inertial moment(s) and possible changes in symmetry (changes in the rotor classification).

Dimerization may be represented by the following chemical equation

$$M_A + M_B = D_{AB} \qquad \text{Equation IX-1}$$

where M_A and M_B are monomers, and D_{AB} is the dimer, similar to the form prescribed by Pine,[362] with the exception of allowance for mixed dimeric species, *i.e.* those comprised of two different molecules. Given this chemical equation, the concentration of the dimeric species may be determined by rearranging the equation for the equilibrium constant K_D,

$$K_D = \frac{[D_{AB}]}{[M_A][M_B]} \qquad \text{Equation IX-2}$$

yielding

$$[D_{AB}] = K_D[M_A][M_B] \qquad \text{Equation IX-3}$$

Since the concentrations of the monomeric species are known from the reference concentration measurements (provided by the manometer readout), the concentration of the dimeric species may be estimated to assess the contribution of dimerization to the observed errors of prediction.

Experimental

Equilibrium constants reported in sources cited by Curtiss and Blander were assembled and averaged, as shown in Table IX-1. It is worth noting that these values only represent the formation of dimers between two molecules of the same species. Equilibrium constants for mixed or hetero-dimers are somewhat rare. Those which were found represented dimers comprised of species of interest bound to haloacids. To

compensate for the lack of credible values, the concentrations of mixed dimers were calculated using the larger of the two homo-dimer constants for each of the species in the hetero-dimer. No equilibrium constants were found for propionitrile or propionaldehyde, so these values were assumed to be equal to the methyl analogs of the respective species.

Table IX-1. Reported equilibrium constants for dimer formation. Mean equilibrium constants are provided for each species.

Species	K_{lit1}	K_{lit2}	K_{lit3}	K_{lit4}	K_{lit5}	K_{lit6}	mean
CH_3CH_2OH	0.0231	0.0236	0.0391				0.0286
CH_3CHO	0.00579	0.0086					0.00719
CH_3CN	0.035	0.0612	0.0206				0.0389
CH_3OH	0.0191	0.0192	0.0177	0.0378	0.0177	0.0012	0.0188
H_2O	0.005	0.0193	0.0111				0.0118
NH_3	0.0005						0.0005

Results

Calculated concentrations of dimers containing each species for sample mixtures prepared at 1 atm overall pressure are illustrated in Figure IX-1. Concentrations of the dimer-bound species range from 0 to 0.0548 ppm. Though several assumptions were made in calculation of these values, the resulting values demonstrate the magnitude of error contributed by this phenomenon. In fact, the K values would need to increase by 2 to 3 orders of magnitude for this variation to be comparable to the observed errors of prediction. Were the K value to be increased by this amount, this would mean that more species would be present in a dimerized state than in a monomeric state, and the spectroscopic features of the dimeric species would become apparent contributors in the THz spectra of the sample mixtures. No features could be seen in the reported data

which would be attributed to absorption features of the dimeric species. Furthermore, HITRAN, the CDMS database, and the JPL spectral database do not currently include spectral parameters explicitly attributed to dimeric species.

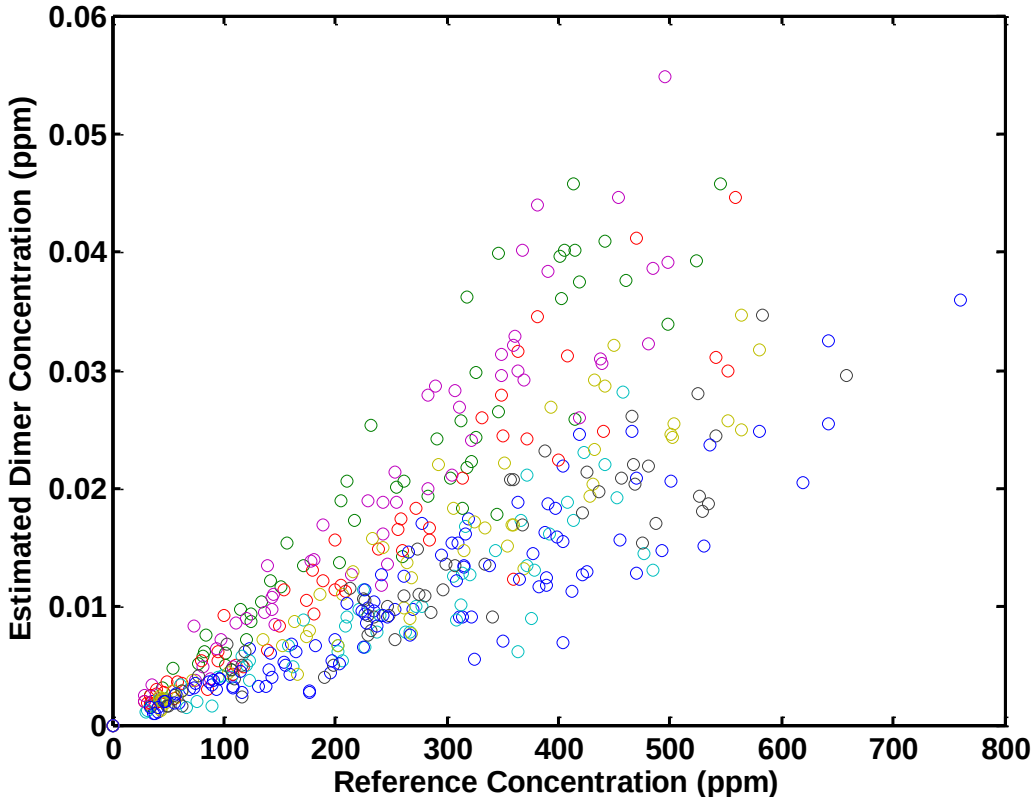


Figure IX-1. Estimated concentrations of dimers containing each species of interest.

Conclusions

While dimerization of gas-phase species presents a plausible source of error, it is negligible when compared to the errors of prediction from the PLS regression models. Dimerization of species in gas-phase samples such as these could represent ~ 0.1% of the total error, given the calculated concentrations. Further investigation of this phenomenon might include preparation of mixtures at elevated concentrations such that the spectral features of dimerized species could be distinguished. K values could be estimated for the

species of interest based on quantitative comparison of monomer spectral features with dimer spectral features.

Suggested Investigation/Future Studies

Terahertz radiation in the sample compartment of the TeraView TPS Spectra 1000D has a beam profile similar in shape to an hour glass turned on its side. This profile is a result of the configuration of optical components within the apparatus, which focuses the THz beam to a diffraction-limited waist near the center of the sample compartment, as shown in Figure IX-2. This helps to achieve high optical throughput with samples which are small by comparison with a collimated beam profile, but this configuration may also be a source of non-linearity in quantitative efforts. The diameter of the THz beam is also frequency dependent, which results in a narrowed hour-glass shape at short wavelengths and a broad shape at long wavelengths, as indicated in Figure IX-2.

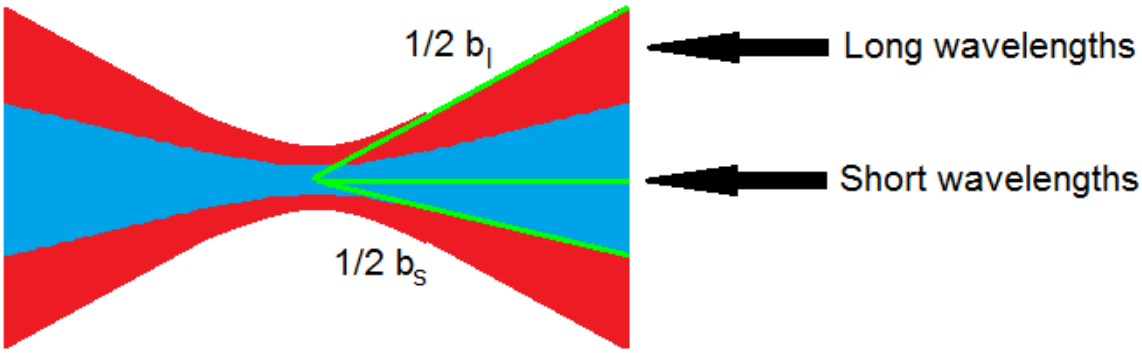


Figure IX-2. Beam profile in the TeraView TPS Spectra 1000D sample compartment. Half-path lengths are indicated by green lines for long and short wavelengths (1/2 b_l and 1/2 b_s, respectively).

Comparison of the two half-path lengths represented in Figure IX-2 reveals a longer path length at low frequencies at the edge of the beam profile, which can be confirmed using

the Pythagorean Theorem. This is thought to impart some level of non-linearity in the observed spectral features due to alteration of the beam profile by changing refractive indices. If the refractive index is increased by elevated concentrations within the gas cell, the focal point of the beam profile should move to the left, assuming photons travel from right to left in Figure IX-2. As a result, the path length would thus become not only a function of frequency, but also of sample concentration. This effect could be further complicated by dispersion, a phenomenon in which the refractive index becomes nonlinear, especially in frequency ranges bordering strong absorption features. Suggested studies would involve calculation of the refractive index for the investigated samples and simulation of the beam profile under different sample conditions to estimate the change in path length.

As described in the theoretical background of rotational transitions, the height of the observed spectral peaks is in part dependent on the temperature of the sample. While temperature control was used on the manifold and transfer line to reduce error in sample concentration measurements, the cell itself was not insulated or temperature controlled, and the measured gas temperatures fluctuated by a few degrees within and between some data sets. Fluctuation of temperature within a data set may account for some level of predictive error, though this is expected to be relatively miniscule, as the within-set standard deviation was typically close to 0.2 °C, or close to the precision of the temperature readout unit. However, ppm-range spectra of pure component samples differed from the sample mixture data sets by as much as 5°C, which may have some effect on the slopes calculated for the PCSA. Suggested studies would involve calculation of the spectra for the species of interest at various temperatures to ascertain the magnitude of the change in absorptivity with changing temperature.

Further studies may be warranted in evaluation of the impact of spectral resolution on the robustness of the method for this application. Results described for the study of resolution only evaluate the impact of spectral resolution on the precision of the

quantitative models, but less attention was given to the interplay between selectivity and resolution. Suggested studies would involve the calculation of SEL ratios from spectra calculated with varying spectral resolution as well as PCSA with such modified data. This would provide insight into the potential benefits (or negative impact) of spectral resolution optimization for each species.

CHAPTER X

THZ-TDS ANALYSIS OF POLYMER SAMPLES

Background

Polymeric materials have been used widely in THz applications, as described in Chapter I. Often times, powdered polymers are used as a support matrix for samples due to their transparency in the THz frequency range. These materials are also used as optical components due to the ease of working and low loss compared to high refractive index materials, such as silicon or quartz.[365, 366] Researchers have noted correlation between the THz spectral features and physical properties of polymers, such as the composition,[367, 368] thermal history,[368] and glass transition temperature.[369]

Spectral properties of these materials can change under certain conditions, confounding or at least complicating spectral interpretation and analysis when the polymer is used as a window or as a sample support matrix. The THz refractive index, absorption peak magnitude, and absorption peak position have been shown to change at varying temperatures for different polymers.[369-372] While the presence of these features does not preclude the use of these materials in THz applications, investigators should be mindful of this variability when designing experiments or selecting optical components. Changes in the refractive index (in regions distant from absorption peaks) are caused by thermally-induced changes in the specific gravity of the polymer.[369] As the temperature of the sample decreases, the refractive index is expected to increase as a result of an increase in the specific gravity.[369] In some polymers, a broadband increase in spectral absorption is also observed at elevated temperatures, arising from increased interstitial space, that is the distance between polymer chains.[369]

The relationship between temperature and free space within the polymer material has been used to explain the origin of glass transitions.[373] In semi-crystalline polymers such as PTFE or polyoxymethylene, some portion of the polymer exists in an ordered,

crystalline state, while the remaining portion is a highly-disordered amorphous state. It is this amorphous phase that exhibits a glass transition.[374] As mentioned previously, changes in temperature affect the density of the polymer material and in turn affect the interstitial space within the polymer. The glass transition is a temperature range in which the space between molecules reaches a critical threshold in which polymer chain motion is constrained or liberated, for cooling or heating processes, respectively.[369] As molecular motion is restricted, long-range order may arise in the polymer structure, permitting propagation of vibrational modes which were previously absent. Zerbi points out that the vibrational transitions observed in polymer materials do not arise from formation of a true crystal lattice as in many biomolecules, but rather from regular arrangements of polymer chains.[375] Morphological transitions studied in this chapter are not glass transitions, but are considered phase transitions which occur as a result of crystalline ordering and disordering processes.[376] The origin of these transitions is similar to that of glass transitions in that they arise from changes in the space within the polymer as well as the length and orientation of polymer chains.

Polytetrafluoroethylene (PTFE, Teflon®)

Spectral variability was noted by the author while collecting spectra with ammonia at ppm concentrations as mentioned in Chapter VI. Absorption spectra of these samples are shown in Figure X-1. Sharp, intense absorption peaks near 18, 38, and 40.5 cm^{-1} are rotational transitions of ammonia molecules. Baseline between these peaks is disrupted by a broad, low-magnitude feature that does not appear to have a consistent position and is not attributed to any of the species of interest. The changing sign of the magnitude of the absorption in this range is a result of changes in the peak magnitude between acquisition of the reference and sample spectra. When the absorption of greatest magnitude in this region was plotted as a function of the change in temperature between the reference and sample, a linear trend was observed, illustrated in Figure X-2. Potential

sources of variation were considered aside from gas-phase species due to the character of the observed features, and literature was found in which spectral features of PTFE (the gas cell window material) were characterized at varying temperatures.

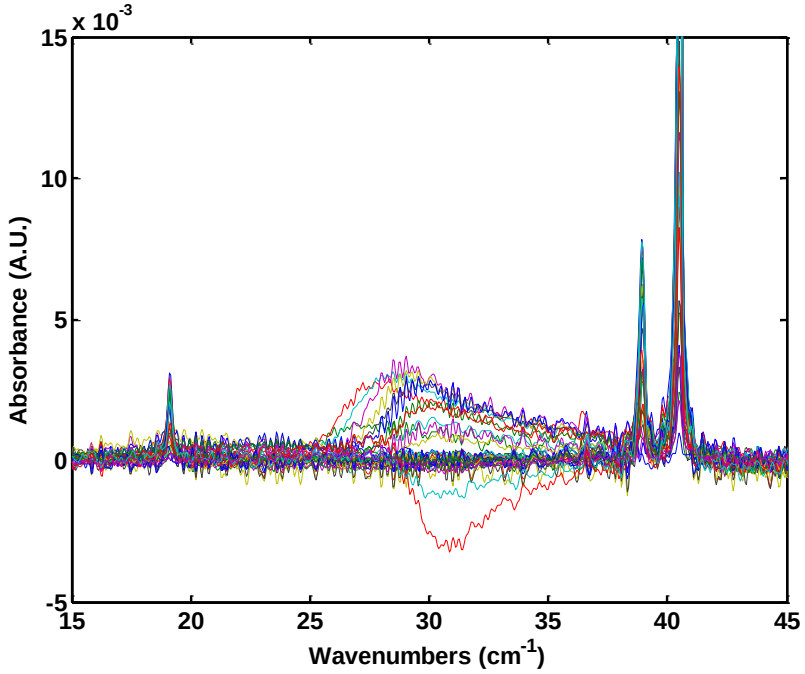


Figure X-1. Expanded view of absorption spectra collected with ammonia gas at ppm concentrations.

Teflon has several physical properties which have made the material attractive for various applications such as its low coefficient of friction, oleophobicity, and hydroscopicity. Of particular interest given the reported experimentation is the existence of a solid-solid phase transition which is known to occur in PTFE at temperatures near 19 °C at 1 atm pressure.[377-381] The glass transition temperature (T_g) of PTFE is far removed from the temperatures measured during gas sample analysis ($T_g \approx 160$ K, or -113 °C), so the observed features were to this point attributed to the phase transition. During this transition, the structure of the polymer changes from 15 CF_2 groups in 7 helical turns to

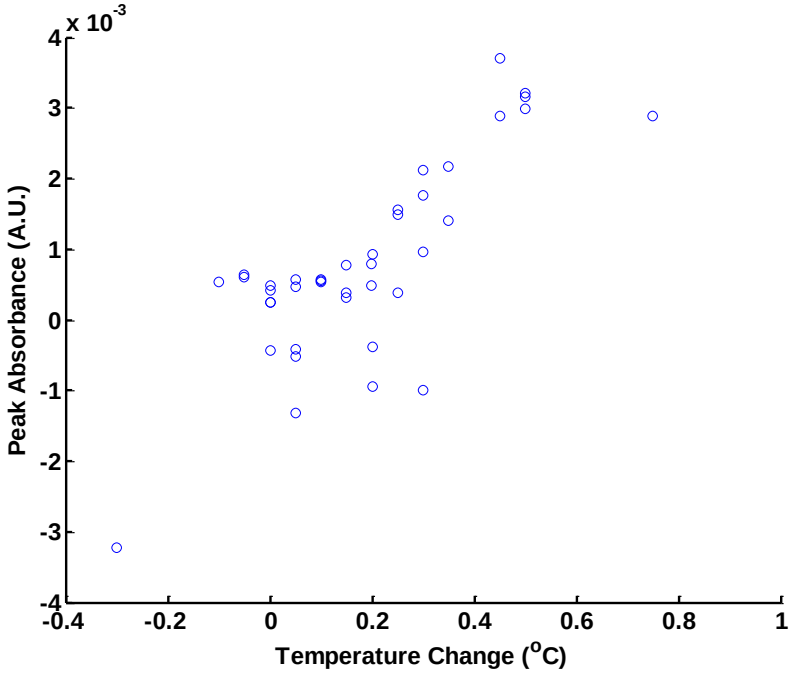


Figure X-2. Peak absorption plotted as a function of temperature change between reference and sample spectra.

13 CF_2 groups in 6 helical turns.[380]

Experimental

To investigate the cause of this baseline shape, several discs were cut from the same rod of virgin electrical grade PTFE from which the windows had been cut. Each side of these discs was polished to improve throughput of THz by reducing scattering. Data were collected from a disc which was 3.43 mm thick. Spectra were analyzed for each disc to ensure consistency between pieces prior to data acquisition. Discs were held in the beam path within the TPS Spectra 1000D by a custom sample holder, shown schematically in Figure X-3. Stable alignment was maintained with a sample retaining clip. Thermocouple probes were held in place by alligator clips against each side of the polymer discs to provide temperature readings during measurement, and the readings from each side were recorded at the start and end of spectral acquisition for each sample.

Readings were averaged to produce a single temperature for each sample spectrum. The temperature of the polymer sample was adjusted by blowing cooled or heated dry air on the sample. Spectra were collected as the average of 1800 scans at a spectral resolution of 1.2 cm^{-1}. A total of 20 spectra were collected from 8.1 to 35.9 °C, covering the temperature range in which the solid-solid phase transition is expected to occur.

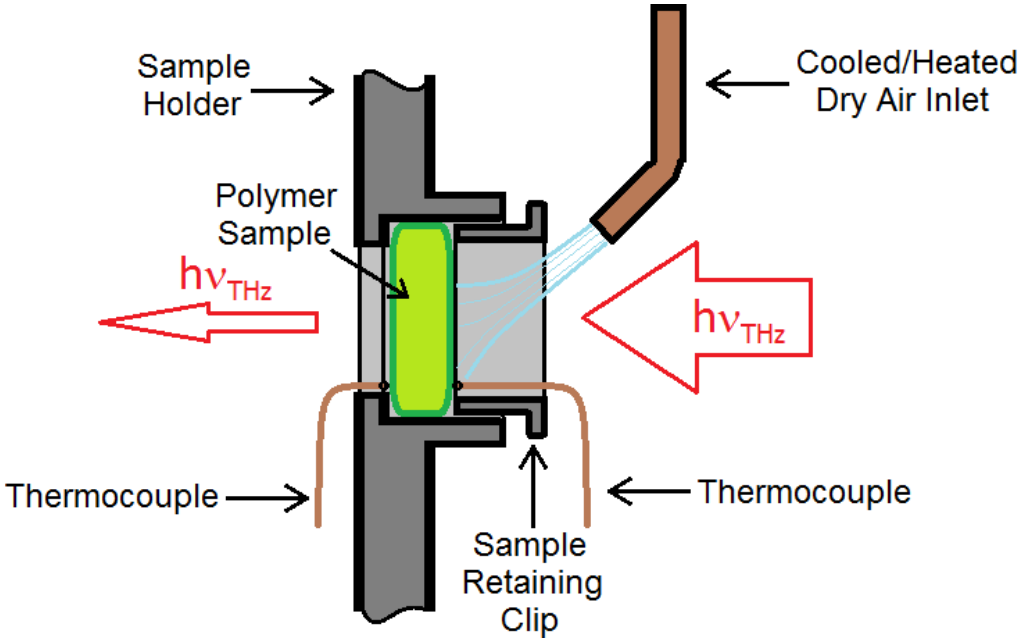


Figure X-3. Schematic representation of the configuration used for polymer sample analysis.

Results

Unprocessed experimental data are shown in Figures X-4 and X-5. Weak absorption peaks are visible in the THz spectrum near 31, 41, and 52 cm^{-1} at decreased temperatures. These features shift in position and decrease in magnitude as the temperature of the sample increases, until they are ultimately indistinguishable. Some seemingly non-random character is evident in the absorption spectra at the high frequencies, possibly attributable to a polymer chain extension vibration, according to

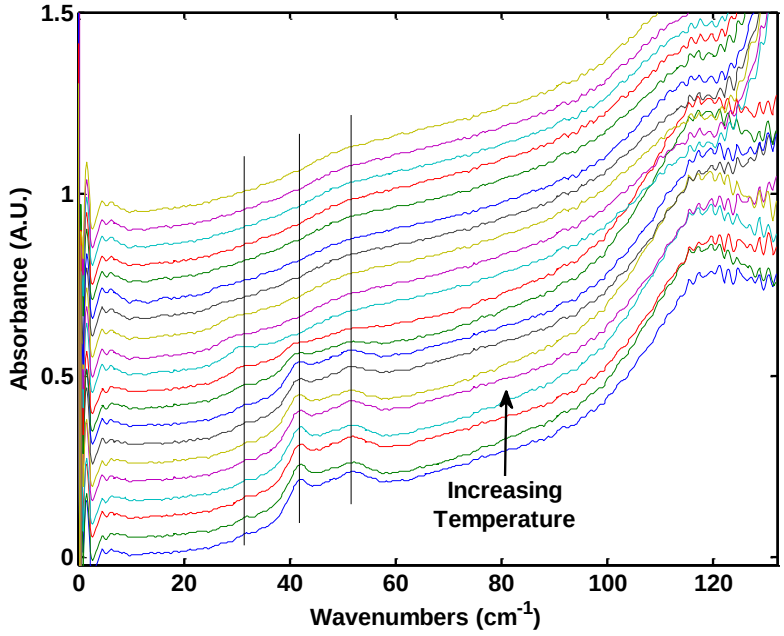


Figure X-4. Experimental THz absorption spectra of a PTFE disc at various temperatures. Spectra are offset vertically by 0.05 A.U. and sorted by temperature. Lines indicate observed peak positions.

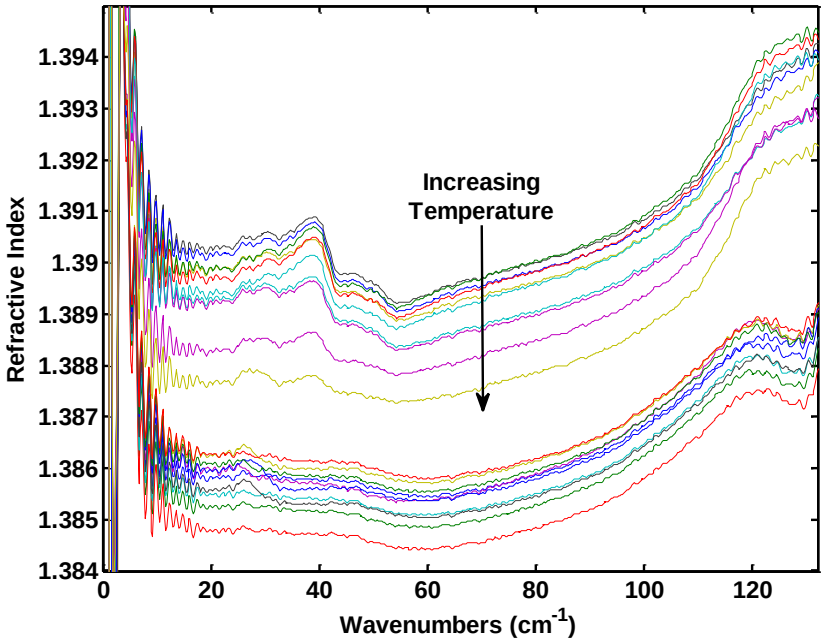


Figure X-5. Experimental THz refractive index spectra of a PTFE disc at various temperatures. No offset is added to these spectra.

Chantry and colleagues.[380] No attention is devoted to interpretation of the shape or magnitude of this portion of the spectrum, as the peak center frequency assigned to this transition (136 cm^{-1}[382] or 140 cm^{-1}[383]) is outside the useable bandwidth. The slope of the baseline has also been discussed before.[380, 381]

Some dispersion can be seen in the refractive index spectra, with refractive indices changing by about 0.003 across the useable bandwidth. The mean refractive index varies with temperature, as has been described by other researchers.[369, 370] Dispersion increases in the vicinity of the observed absorption features as the magnitude of absorption increases. The spectral baseline was linearized by fitting two spectral regions with a cubic polynomial function in the 10.71-26.87 and 62-80 cm^{-1} spectral regions. Difference spectra were calculated by subtracting the spectrum collected at 35.93 °C from the remaining spectra. This provides a much clearer picture of the changes in the measured spectrum, as illustrated in Figure X-6.

The standard deviation of 3 replicate spectra collected with the PTFE disc with minimal temperature variation is plotted in Figure X-6 as a bold dotted black line. These spectra were collected by removing and re-inserting the sample holder between times of spectral acquisition, a procedure which mimics that of the temperature-variant data. Comparison with the standard deviation reveals the significance of the observed spectral features. While the magnitude of features below 25 cm^{-1} appears significant when compared with the standard deviation, these features are known to arise from reflections within the pellet and should be disregarded. These would ordinarily be eliminated by calculation of the spectral differences, but the thickness of the pellet varies with the change in temperature, and thus the frequency of the observed etalon varies from one sample to another. The change in thickness across this temperature range is expected to be less than 0.015 mm, depending on the grade of Teflon.

The DFF pre-processing method utilized for quantitative analysis of the gas samples could be used for solid samples such as these, however given the width and

periodicity of the baseline artifacts, their removal with DFF pre-processing results in significant scarification and attenuation of the features of interest. Absorption spectra provided for the PTFE sample were calculated with TDS truncation to reduce the appearance of reflection-induced artifacts, but the presence of other periodic features in the TDS in closer proximity to the main peak prevents the total elimination of these artifacts in the resulting frequency-domain spectrum.

In general, the magnitude of the absorption peaks near 41 and 51 cm^{-1} increase with decreasing temperature, and the peak near 30 cm^{-1} is observed at temperatures near the phase transition temperature. Experimental spectra were fit with Gaussian profiles, as the quality of fit was consistently higher for this line shape compared to either Lorentzian or Voigt line shapes. Spectra were fit with 4 peaks near 30.8, 33.6, 41.6, and 50.2 cm^{-1}. Figure X-7 illustrates a representative experimental spectrum with the calculated fit and residuals of the fit. Figure X-8 displays the fitted peak height as a function of temperature. The height of the peaks at 41.6 and 50.2 decrease sharply as the temperature increases in the vicinity of the 19 °C solid-solid transition temperature. In addition to the steep decrease in peak magnitude, two weak transitions increase in magnitude near the phase transition temperature, and decrease above this temperature. Another popular way to visualize trends in spectral peaks is to plot the integrated peak area as a function of the independent variable, as shown in Figure X-9. Similar trends may be seen in the peak areas. The sharp decline in peak height and peak area near the transition temperature suggest the utility of THz spectroscopy for non-contact, non-destructive identification of solid-solid phase transition temperatures in polymers.

According to results reported in the scientific literature, the glass transition temperature of a polymer may be determined by measuring the THz refractive index of a polymer at various temperatures surrounding the glass transition temperature.[369, 370] Refractive index values determined for the PTFE discs was plotted as a function of temperature to evaluate the utility of this method for identification of other solid-solid

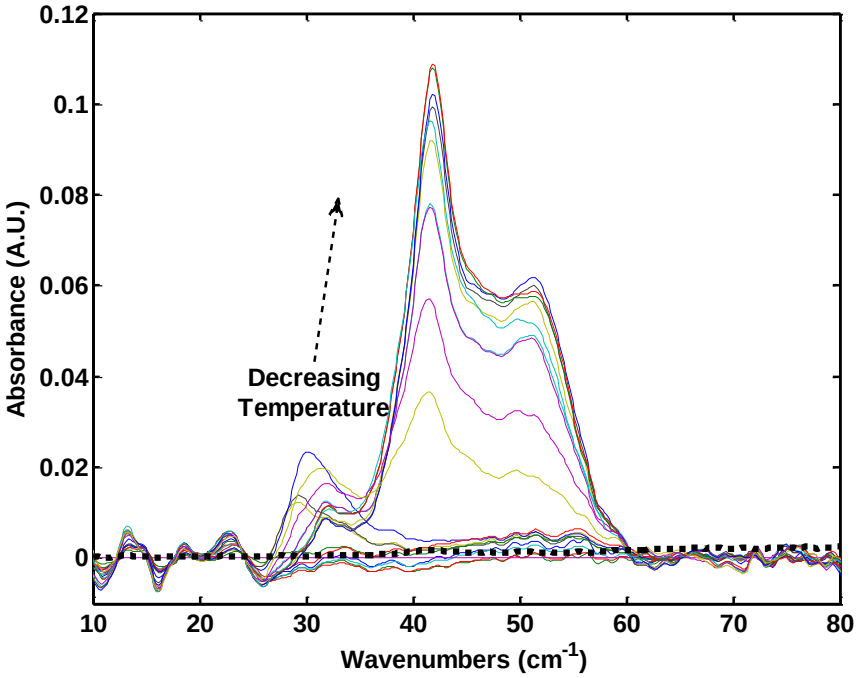


Figure X-6. Difference spectra from the PTFE sample.

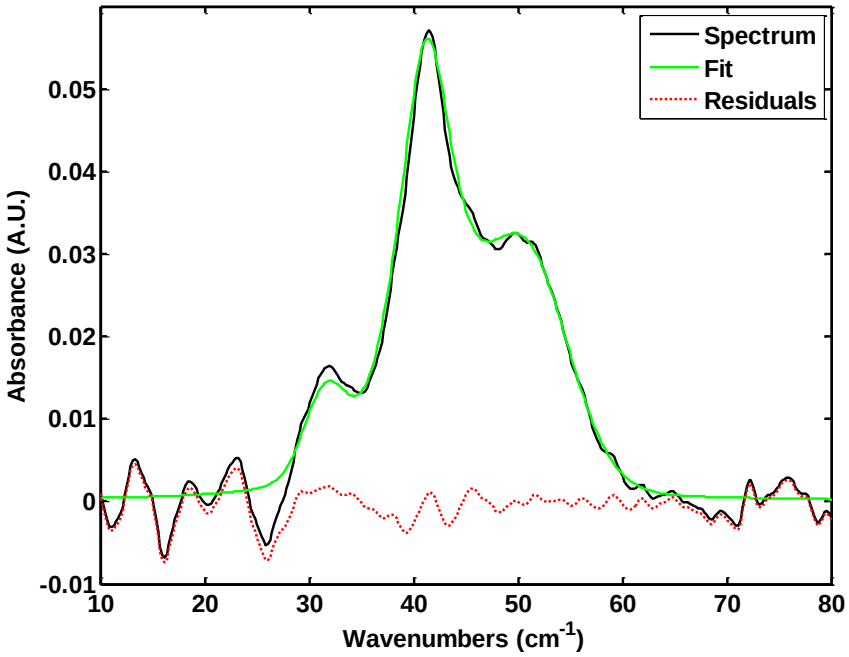


Figure X-7. Representative plot of experimental data (collected at 20.68 °C), fitted spectrum, and spectral residuals from a PTFE sample.

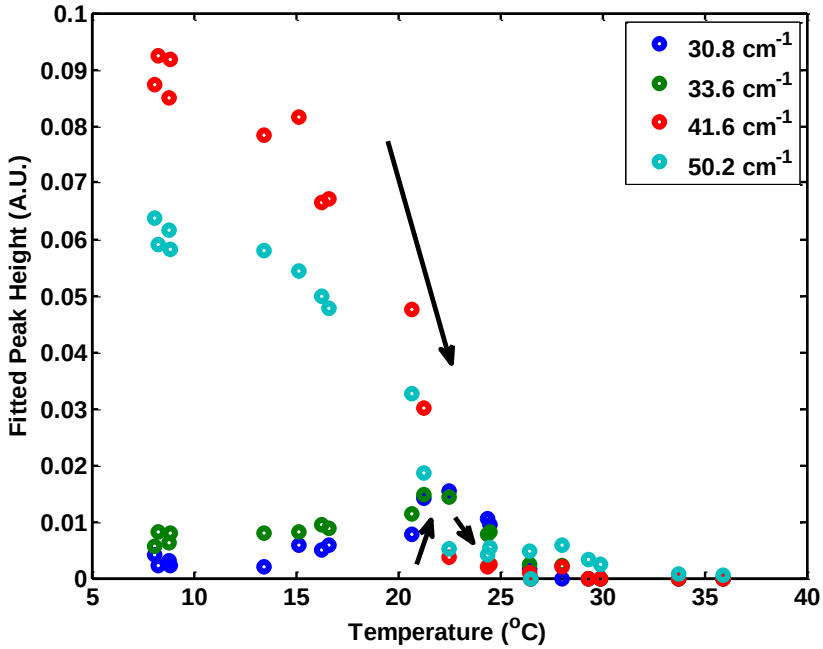


Figure X-8. Fitted peak heights of the four observed absorption peaks at various temperatures. Arrows show trends observed with increasing temperature.

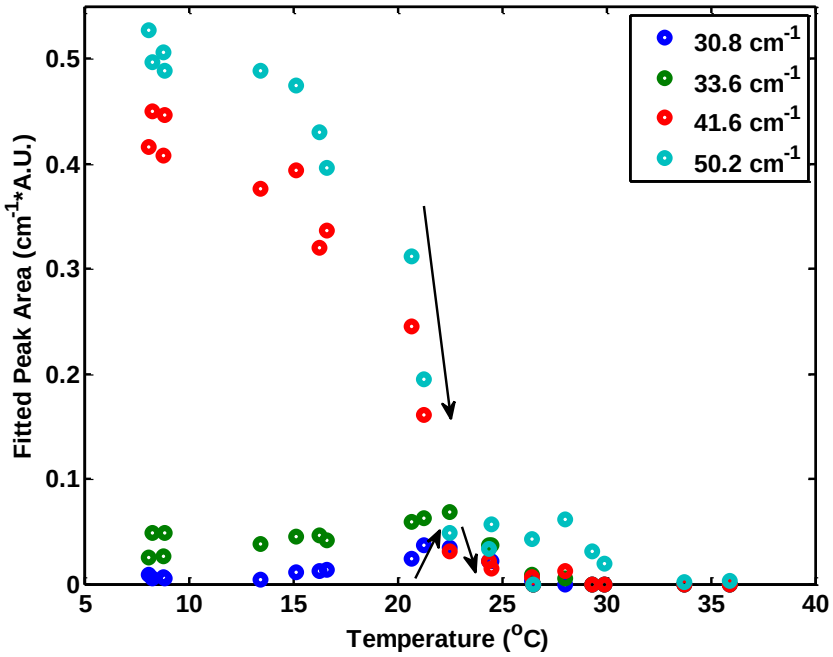


Figure X-9. Fitted spectral peak areas at various temperatures. Arrows indicate trends observed with increasing temperature.

transitions, as shown in Figure X-10. In this plot, best fit lines were calculated for refractive index values taken from various points on the refractive index spectrum displayed in Figure X-5. These points are not the same as the frequencies used for peak fitting, as the peak positions varied with temperature. Instead, four points were selected representing two regions of the refractive index spectra: 39.19 and 54.35 cm^{-1} are near areas exhibiting dispersion which is sensitive to temperature (close proximity to absorption bands), and 75.76 and 96.97 cm^{-1} are located in areas of minimal temperature-dependent dispersion. Fits were performed in two temperature regimes bordering the observed transition, excluding the measured values at 21.27 °C. The difference in the refractive index measured at this temperature from those measured at other temperatures is so great that inclusion of these points in regressions to either side would have produced large leverages.

According to Wietzke et al., transition temperatures are marked by a change in the slope of the refractive index as a function of temperature.[369] Slopes and y-intercepts of the best-fit lines (provided in Table X-1) were compared with a 1-tailed *t*-test (at 95% confidence) to evaluate the significance of the change in parameters. Calculated *t* values are provided in Table X-2. Both the slopes and intercepts changed significantly between the low- and high-temperature regimes. This supports the utility of Wietzke's method for identification of solid-solid phase transitions. While not observed clearly in the refractive index data, an additional phase transition is known to occur in PTFE near 30 °C.[376] The fitted peak area, and to some extent the fitted peak height of the spectral peak near 50.2 cm^{-1} exhibit a second dip in magnitude near 30 °C. These vibrational modes are weak or nonexistent beyond ~22.5 °C, so the authenticity of this second dip is uncertain.

Observed peak positions are not totally congruous with those reported in the literature, and those reported in the literature are often not in agreement. Two explanations for these discrepancies are proposed here. Given that the thermal history of polymers may affect the observed THz spectral response, the observed feature may be

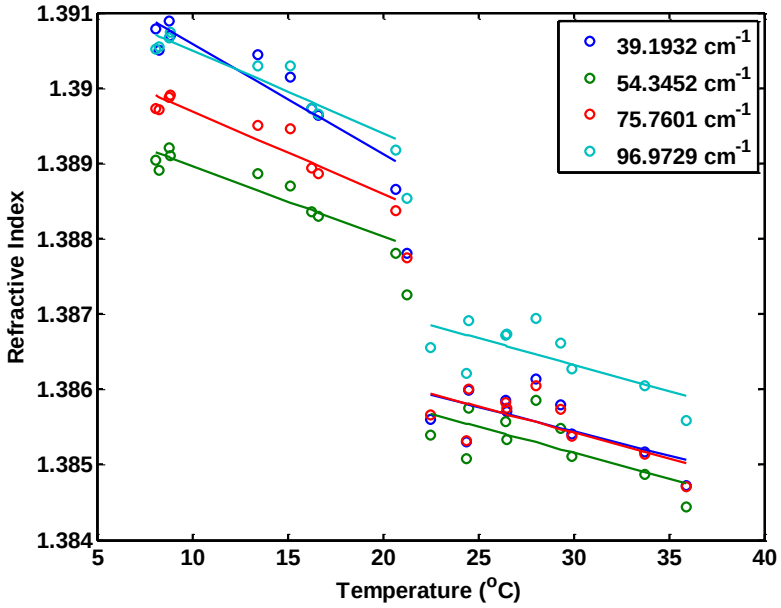


Figure X-10. Refractive index at various frequencies measured at different temperatures. Best-fit lines are provided for low- and high-temperature regimes, excluding the data points at 21.27 °C.

Table X-1. Equations for best-fit lines plotted in Figure X-10.

Low Temperature (< 21.27 °C)	
Wavenumber (cm^{-1})	Equation
39.19	N = -6E-5 (± 3E-5)T + 1.3874 (± 8E-4)
54.35	N = -7E-5 (± 3E-5)T + 1.3872 (± 7E-4)
75.76	N = -7E-5 (± 2E-5)T + 1.3875 (± 7E-4)
96.97	N = -7E-5 (± 3E-5)T + 1.3884 (± 7E-4)
High Temperature (> 21.27 °C)	
Wavenumber (cm^{-1})	Equation
39.19	N = -1.5E-4 (± 2E-5)T + 1.3921 (± 3E-4)
54.35	N = -9.4E-5 (± 1E-5)T + 1.3899 (± 2E-4)
75.76	N = -1.1E-4 (± 1E-5)T + 1.3908 (± 2E-4)
96.97	N = -1.1E-4 (± 2E-5)T + 1.3916 (± 2E-4)

Table X-2. Calculated *t*-values for comparison of regression parameters. The critical *t*-value is 2.1.

Wavenumber (cm^{-1})	Intercept	Slope
39.19	17.9	7.3
54.35	11.2	2.7
75.76	14.0	4.2
96.97	13.2	4.0

unique to these samples due to the preparation procedure. The composition of samples described in the scientific literature is not totally clear, as some blends of PTFE may contain polymers from previous batches or other polymer materials. Samples analyzed herein were virgin electrical grade PTFE, meaning that no additives were used in polymer fabrication.

Polyoxymethylene

Polyoxymethylene (POM) is a semi-crystalline thermoplastic produced by DuPont under the trade name Delrin® which offers a low coefficient of friction and high mechanical rigidity. This polymer exhibits two transitions: one weak transition observed around 0-15 °C and a strong transition at -80 °C.[384] While often molded into various shapes, POM may also be machined due to its high rigidity. The chemical formula of POM is illustrated in Figure X-11.

Experimental

To investigate the solid-solid phase change in POM, several discs were cut from a single piece of polymer. No further sample preparation was performed, as the polymer strip was provided with a smooth surface. Data were collected with a disc which was 1.05 mm thick. Spectra were analyzed with each disc to ensure consistency between

pieces prior to data acquisition. Discs were held in the beam path using the same sample holder shown in Figure X-3. Temperature was measured and mean temperatures were calculated using the same method as for analysis of the PTFE samples. Spectra were collected as the average of 1800 scans at a resolution of 1.2 cm^{-1}. A total of 20 spectra were collected from -52.0 to 33.2 °C, covering the temperature range in which the aforementioned weak solid-solid phase transition is expected to occur.

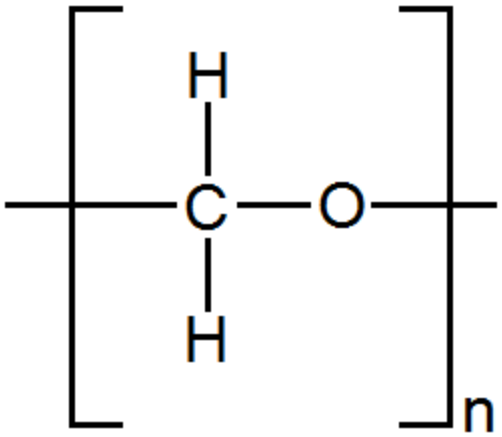


Figure X-11. Formula of POM polymer.

Results

Absorption spectra collected with the POM sample and calculated with TDS truncation are shown in Figure X-12. The shape of the baseline appears more linear than that of the PTFE samples. Two absorption peaks are visible in these spectra near 41.5 and 91.7 cm^{-1}. Overall, absorption of the THz frequencies is higher in POM than in PTFE samples. The shape of the spectrum does not appear to change significantly with the changing temperature – no distinct peaks appear in the spectrum as the temperature changes, unlike the spectra of PTFE. The position, height and width of the absorption peaks appear to change slightly with temperature, however.

The THz refractive index of POM is provided in Figure X-13. Like PTFE, POM exhibits some dispersion in the THz range, with slight changes in the shape and magnitude of the refractive index as the temperature varies. Strong oscillations are visible in the refractive index spectrum at low frequencies, which are the result of reflections within the POM sample. Dispersion is greatest in the vicinity of the 91.7 cm^{-1} transition. Weak dispersion is expected in the refractive index spectrum near the 41.5 cm^{-1} transition, but this obscured by the etalon artifact.

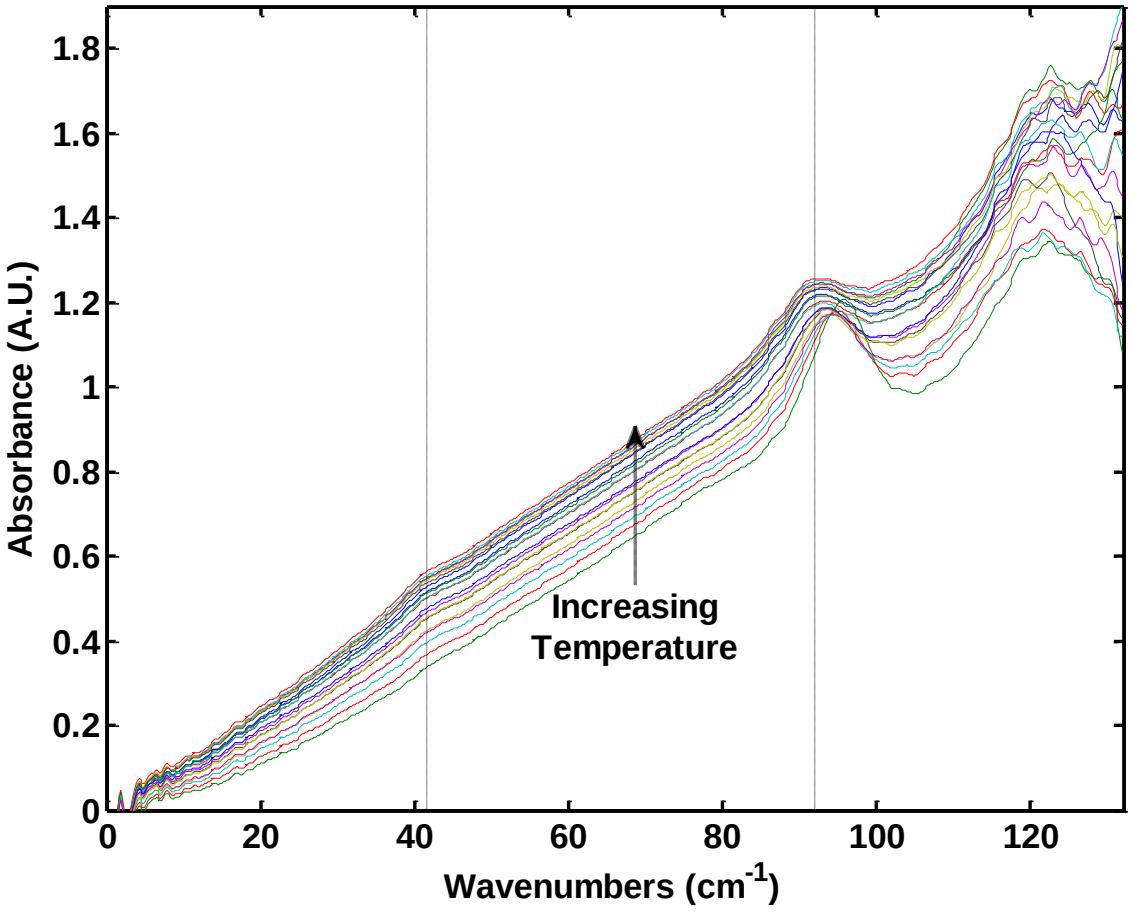


Figure X-12. Raw absorption spectra measured with a POM sample.

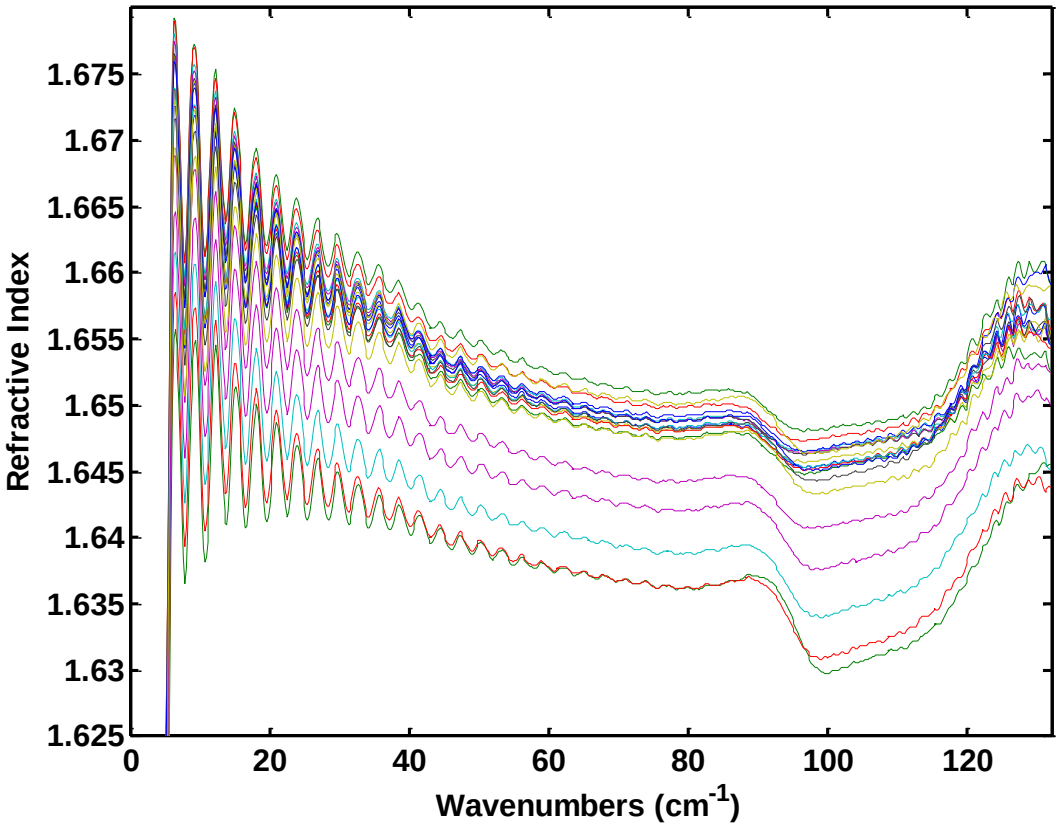


Figure X-13. Refractive index spectra measured with a POM sample.

Spectra were prepared for peak fitting using a cubic polynomial fit of the baseline regions at 11.5-33.94, 47-80, and 103-106 cm^{-1}, producing the spectra plotted in Figure X-14. DFF pre-processing was used to eliminate high-frequency noise and spectral artifacts, producing the absorption spectra plotted in Figure X-15. Difference spectra were not calculated from the POM data because all absorption spectra contained the spectral features of interest, and the baseline was adequately removed by subtracting the cubic function from the spectra. At the lowest temperature, a slight peak at 78 cm^{-1} appears, which may be associated with the glass transition at -80 °C, but this temperature could not be achieved with the described apparatus to verify the source of this feature. Figure X-16 illustrates a representative spectrum of the POM sample with the corresponding fit and residuals for each of the two spectral peaks.

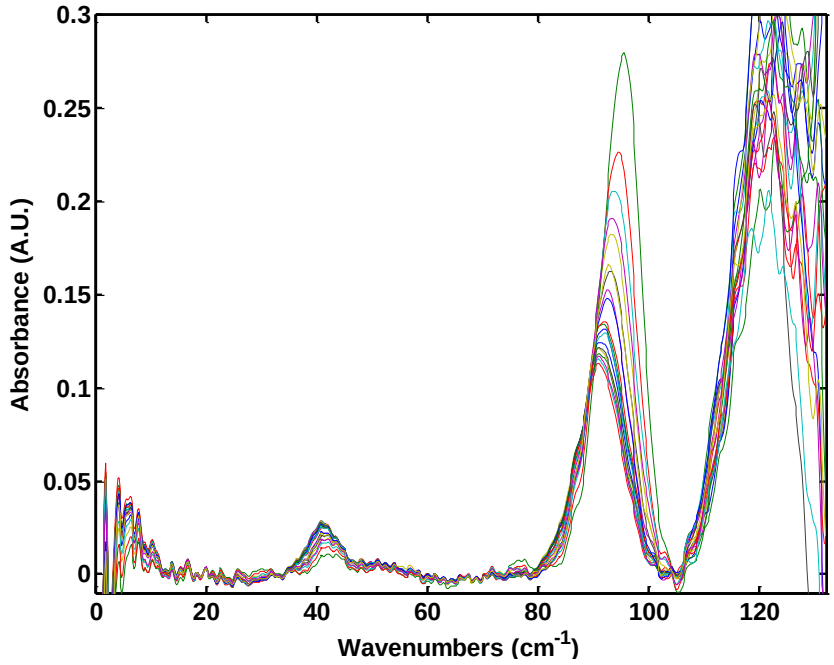


Figure X-14. Baseline-corrected absorption spectra of a POM polymer sample.

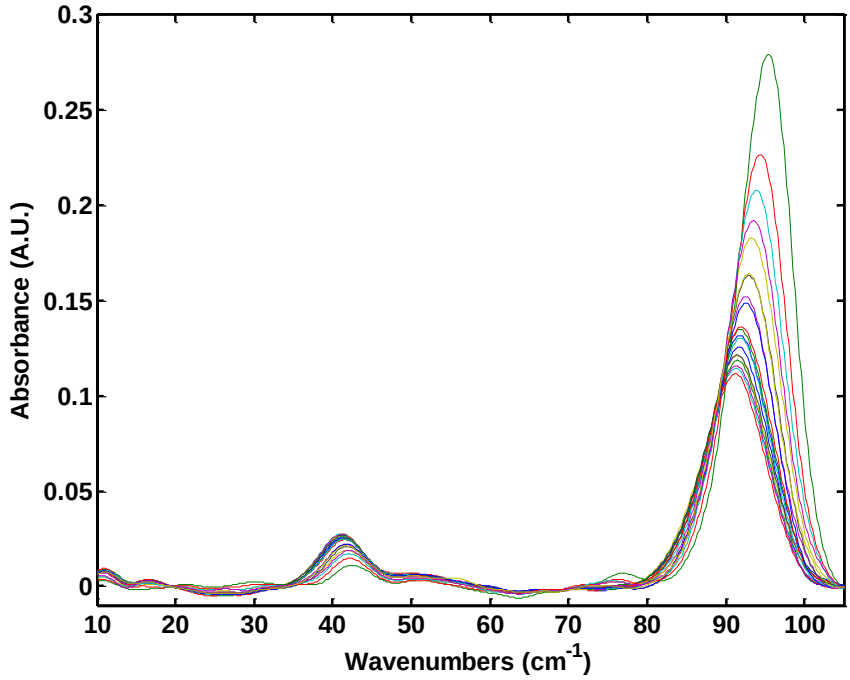


Figure X-15. Baseline-corrected, DFF pre-processed absorption spectra of a POM polymer sample.

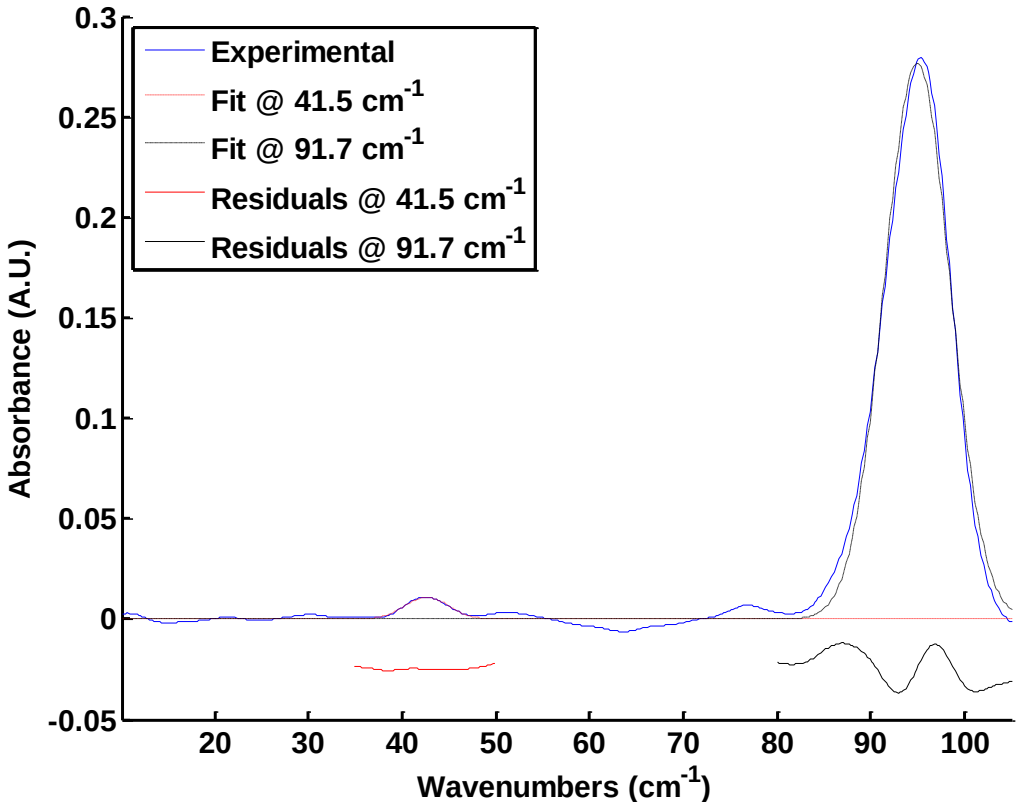


Figure X-16. Representative plot of experimental data (collected at -51.98 °C), fitted spectrum, and spectral residuals from a POM sample. Residuals are offset by -0.025 A.U. for comparison.

As with the PTFE sample, peaks in the absorption spectra were fit with Gaussian profiles. Peak parameters from the fits were plotted as a function of temperature in an attempt to identify the weak solid-solid transition near 0-15 °C. The peak height, change in peak position, width, and area were plotted as a function of temperature, shown in Figures X-17, X-18, X-19, and X-20, respectively. The relationship between the peak area and sample temperature was re-evaluated using values normalized by the areas calculated at 22.8 °C to improve visualization of the trend, plotted in Figure X-21. Best-fit line parameters for all linear regressions are provided in Table X-3, and Student's *t* values from comparison of the parameters are provided in Table X-4.

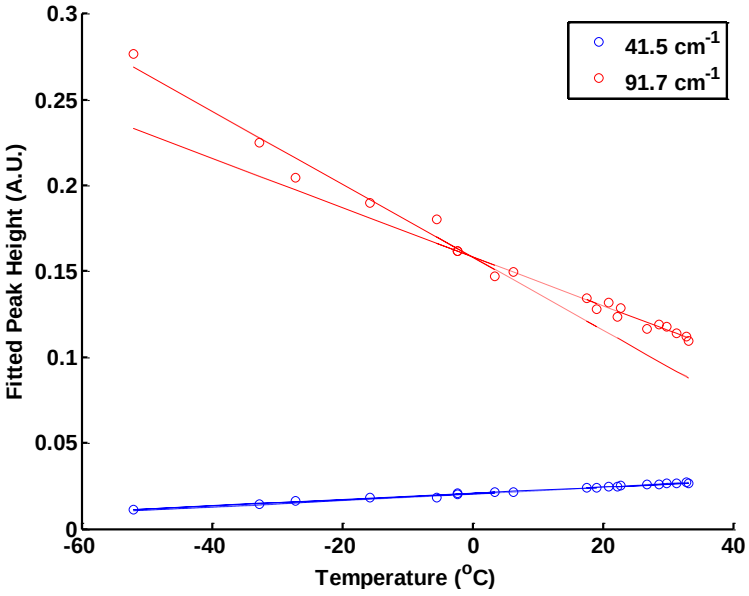


Figure X-17. Fitted peak height plotted as a function of the POM sample temperature. Samples collected at temperatures above the reported transition are fit with a dotted line and samples collected below the reported transition are fit with a solid line.

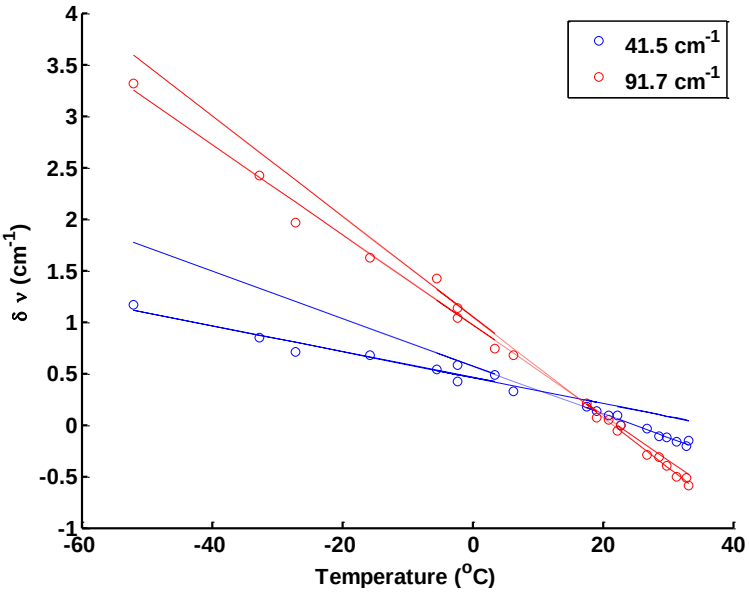


Figure X-18. Fitted peak position change (relative to the position determined at 22.8 °C) as a function of temperature. Samples collected at temperatures above the reported transition are fit with a dotted line and samples collected below the reported transition are fit with a solid line.

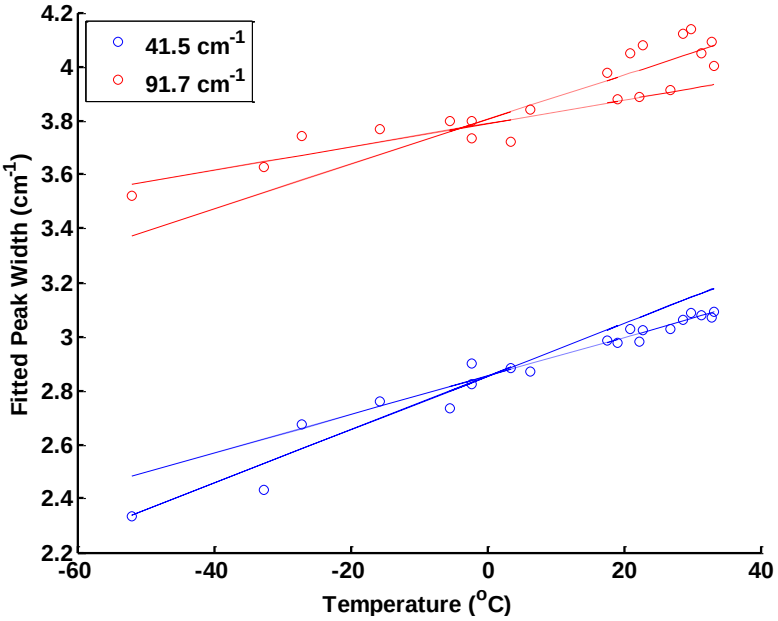


Figure X-19. Fitted peak width plotted as a function of temperature. Samples collected at temperatures above the reported transition are fit with a dotted line and samples collected below the reported transition are fit with a solid line.

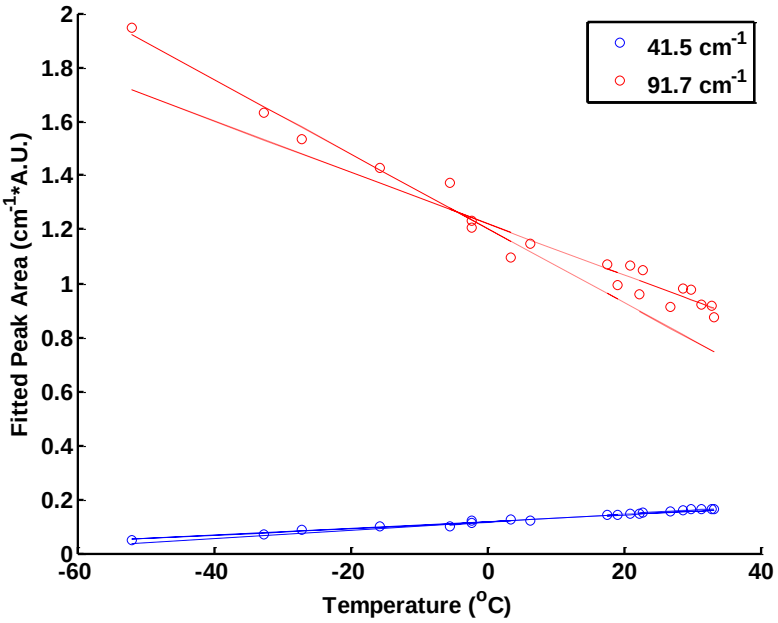


Figure X-20. Fitted peak area plotted as a function of temperature. Samples collected at temperatures above the reported transition are fit with a dotted line and samples collected below the reported transition are fit with a solid line.

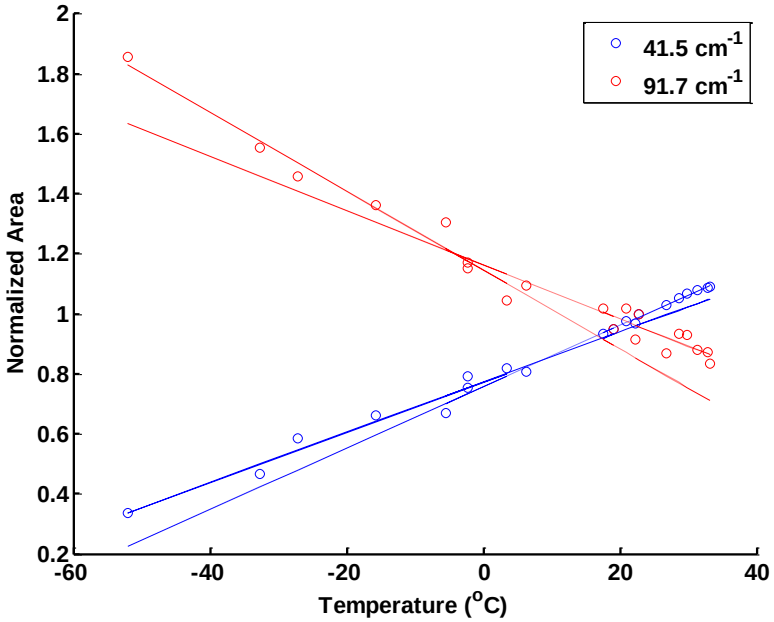


Figure X-21. Normalized fitted peak area plotted as a function of temperature. Areas were normalized by the value calculated at a temperature of 22.8 °C. Samples collected at temperatures above the reported transition are fit with a dotted line and samples collected below the reported transition are fit with a solid line.

Table X-3. Linear regression parameters from peak fit analysis regressions. Parameters are provided for temperatures above and below the glass transition temperature range.

	Peak 1 (41.5 cm^{-1})							
	T<0°C				T>0°C			
Trend	m	s_m	b	s_b	m	s_m	b	s_b
H vs. T	1.8E-04	1E-05	2.07E-02	3E-04	2.02E-04	5E-06	2.02E-02	1E-04
δv vs. T	-1.2E-02	1E-03	3.9E-01	3E-02	-2.0E-02	1E-03	4.1E-01	3E-02
W vs. T	1.0E-02	1E-03	2.86E+00	3E-02	7.9E-03	7E-04	2.83E+00	2E-02
A vs. T	1.3E-03	1E-04	1.17E-01	2E-03	1.55E-03	6E-05	1.15E-01	2E-03
A (norm) vs. T	8.4E-03	6E-04	7.7E-01	1E-02	1.02E-02	4E-04	7.6E-01	1E-02
	Peak 2 (91.7 cm^{-1})							
	T<0°C				T>0°C			
Trend	m	s_m	b	s_b	m	s_m	b	s_b
H vs. T	-2.1E-03	1E-04	1.58E-01	3E-03	-1.4E-03	1E-04	1.59E-01	4E-03
δv vs. T	-4.4E-02	2E-03	9.7E-01	5E-02	-4.9E-02	2E-03	1.05E+00	5E-02
W vs. T	4E-03	1E-03	3.79E+00	2E-02	8E-03	5E-03	3.8E+00	1E-01
A vs. T	-1.38E-02	9E-04	1.20E-00	2E-02	-1.0E-02	2E-03	1.22E+00	6E-02
A (norm) vs. T	-1.31E-02	9E-04	1.15E+00	2E-02	-9E-03	2E-03	1.16E+00	6E-02

Table X-4. Calculated *t* values from comparison of the regression parameters provided in Table X-3.

Peak 1 (41.5 cm^{-1})				
	Intercept		Slope	
Trend	t_{crit}	t_{calc}	t_{crit}	t_{calc}
H vs. T	2.10	1.94	2.20	3.54
δν vs. T	2.10	6.91	2.10	16.58
W vs. T	2.10	0.12	2.10	5.32
A vs. T	2.10	2.80	2.10	7.83
A (norm) vs. T	2.10	2.80	2.10	7.83
Peak 2 (91.7 cm^{-1})				
	Intercept		Slope	
Trend	t_{crit}	t_{calc}	t_{crit}	t_{calc}
H vs. T	2.10	0.11	2.10	10.81
δν vs. T	2.10	3.38	2.10	5.25
W vs. T	2.18	0.39	2.16	2.81
A vs. T	2.13	0.93	2.12	5.81
A (norm) vs. T	2.13	0.93	2.12	5.81

Strong linear trends were observed between all of the fitted spectral parameters and the measured temperature. The high-frequency vibrational transition peak decreased in magnitude as the temperature increased, while an opposite trend was observed in the peak toward the low-frequency end of the spectrum. Both peaks shift towards lower frequencies and become broader as the temperature increases. Trends in the fitted peak areas are similar to those of the peak height. The slope from the relationships between normalized peak area and temperature show similarity in the rate of change in the area, though no structure or chemical significance is inferred from this observation.

As with PTFE samples, the refractive index was plotted as a function of temperature to investigate the relationship between sample temperature and the measured refractive index. A plot of the refractive index at several frequencies is provided in Figure X-22. The refractive index measured at the lowest temperature is excluded from the linear fits because the error in the measured temperature is expected to be very large. For this particular sample, the temperature measured on each side indicated a large temperature gradient was established across the sample (-81.3 °C on the chilled side of the disc and -25.5 °C on the other side at the start of spectral acquisition). Furthermore, the measured values on the chilled side changed by several degrees during data collection. Both of these effects are explained by heat transfer between the sample and holder, as well as transfer between the sample and the instrument's supply of dried purging air. Regression parameters for the best fit lines are provided in Table X-5, and the calculated Student's *t* values from comparison of the regression parameters are shown in Table X-6.

As with the PTFE sample, the refractive index underwent a significant change both in the slope and y-intercept near the expected solid-solid transition temperature. The change in the measured refractive index of POM was opposite of the change observed in PTFE, with higher refractive indices measured at elevated temperatures. The change in refractive index observed for POM was also opposite that which was reported by Wietzke and colleagues,[369] but this difference in trends may be explained by the anisotropy of the thermal expansion of POM. The chain spacing in crystalline POM is known to decrease with increasing temperature in the (009) plane, whereas a remarkable increase in spacing is observed with increasing temperature in the (100) plane.[385] This may also explain the discrepancy between the observed spectral peak positions and those reported in the literature.[375] Amrhein and Heil reported a peak for POM near 90 cm^{-1}.[386]

No mention of the peak near 41.5 cm^{-1} was found in the THz or far-IR literature. While Wietzke's work covered this spectral range, no absorption spectra were provided

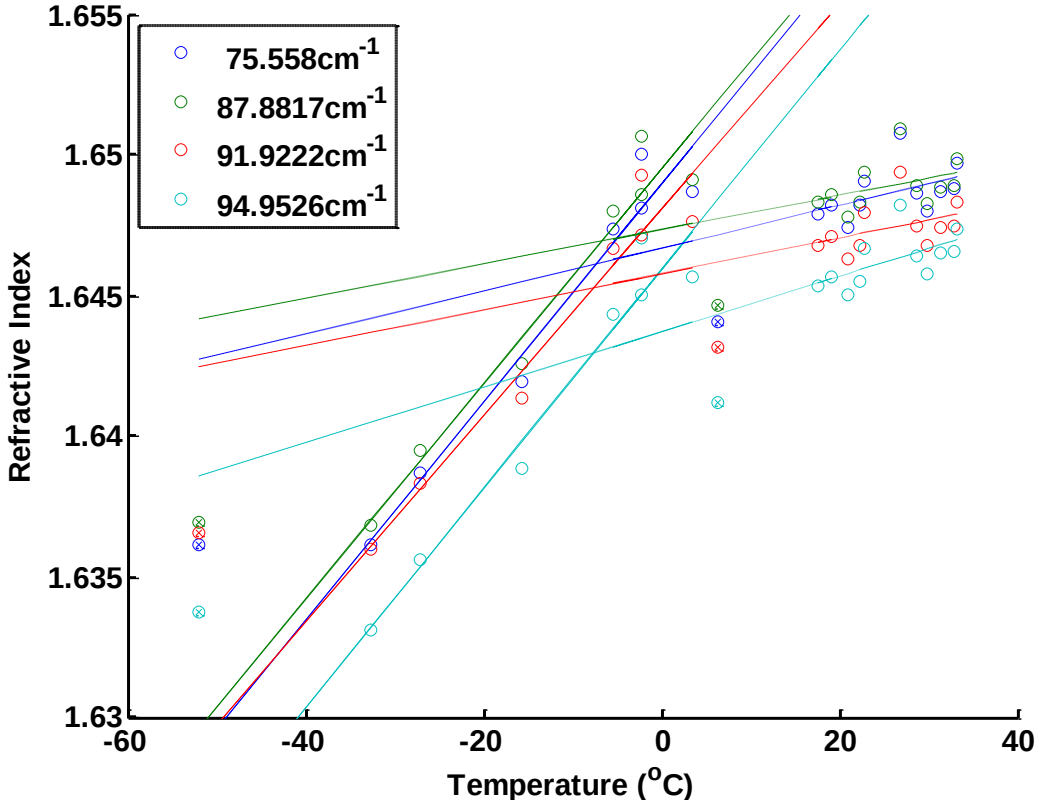


Figure X-22. The THz refractive index of POM at various temperatures. Best fit lines are provided for values measured at several frequencies. Outlying points are marked with x's and are not included in the regressions. Samples collected at high temperatures (above the reported transition temperature) are fit with a dotted line, while samples collected at low temperatures are fitted with a solid line.

Table X-5. Regression parameters from the relationship between sample temperature and measured refractive index.

Low Temperature (< 6.35 °C)	
Wavenumber (cm^{-1})	Equation
75.56	N = 3.9E-4 (± 4E-5)T + 1.6490 (± 6E-4)
87.88	N = 3.8E-4 (± 4E-5)T + 1.6495 (± 7E-4)
91.92	N = 3.7E-4 (± 4E-5)T + 1.6481 (± 7E-4)
94.95	N = 3.9E-4 (± 4E-5)T + 1.6459 (± 6E-4)
High Temperature (≥ 6.35 °C)	
Wavenumber (cm^{-1})	Equation
75.56	N = 8E-5 (± 5E-5)T + 1.6467 (± 1E-3)
87.88	N = 6E-5 (± 5E-5)T + 1.6473 (± 1E-3)
91.92	N = 6E-5 (± 5E-5)T + 1.6458 (± 1E-3)
94.95	N = 1.0E-4 (± 5E-5)T + 1.6437 (± 1E-3)

Table X-6. Calculated *t* values from comparison of the regression parameters provided in Table X-5. The critical *t* value is 2.1.

Wavenumber (cm^{-1})	Intercept	Slope
75.56	5.2	15.2
87.88	4.5	16.1
91.92	4.8	15.2
94.95	4.8	15.0

for comparison.[369] Furthermore, this vibrational transition increases in intensity with increasing temperature, which means it might be attributed to motions in the amorphous fraction of the polymer. This would in turn prevent its observation in experimentation with samples having high levels of crystallinity.

Conclusions

Spectroscopic investigation of polymers with the THz-TDS method has been demonstrated. Various spectral parameters provide insight into the properties of polymers, and the temperature at which solid-solid phase transitions occur may be gleaned from refractive index and/or absorption spectral features, depending on the particular polymer. A solid-solid phase change was observed in PTFE near 19 °C using both spectral features and the refractive index method described by Wietzke and colleagues.[369, 370] Absorption spectral features of POM did not provide strong evidence of a solid-solid phase transition, but analysis of the refractive index at several frequencies provided evidence of a transition between 3 and 6 °C. This method provides a non-contact and non-destructive alternative to differential scanning calorimetry for determination of phase transition temperatures, as was also concluded by Wietzke et al.[369]

Several modifications are suggested for further investigation of this topic. Incorporation of a temperature-controlled cryostat would improve control over the temperature of the sample and would reduce the magnitude of the temperature gradient across the polymer disc, due to insulation provided by evacuation of gases from the cryostat sample chamber. This would also enable analysis of a wider range of polymers, beyond those exhibiting phase transitions near room temperature. The experimental procedure could be changed to include reference spectral acquisition at the start and end of data collection to reduce spectral variability resulting from un-mounting and re-mounting of the sample holder within the beam path. Thinner samples should not be used for reduction of thermal heterogeneity, as the glass transition temperature is depressed in thin polymer layers.[387]

Several studies are suggested based on the reported results. The anisotropic nature of the thermal expansion of POM suggests further investigation into the THz spectral features of polymers and the potential directionality of said features. Given that toughness can be affected by the crystallinity of a polymer,[388] the THz-TDS may provide

a rapid and non-destructive method for the quantitation of crystallinity, as it is this fraction of the polymer material which exhibits distinct THz features. Further studies are warranted in the investigation of the phase transition in PTFE near 30 °C, as the observed change in absorption was relatively weak. More spectra could be collected around this temperature range with a thicker piece of PTFE to provide increased absorption, and thus greater certainty regarding the change in the spectral features. Analysis of POM at temperatures higher than those accessible with the current apparatus are also suggested for investigation of the vibrational transition near 41.5 cm^{-1}. Should this peak continue to increase in magnitude as the sample temperature nears the melting point, the source of this peak may be better ascertained. If POM samples are available with varying crystallinity, spectroscopic analysis of these samples should test this hypothesis.

CHAPTER XI

THZ-TDS ANALYSIS OF DENTAL TISSUES AND COMPOSITES

Background

As mentioned in Chapter I, the THz-TDS method has been applied for analysis of dental tissues in both healthy and damaged states. Given the utility of the method in these applications, a series of experiments were performed in collaboration with Dr. Stephen Armstrong of the University of Iowa's College of Dentistry to evaluate the utility of THz-TDS for non-contact and non-destructive investigation of bond quality between dental tissues and reconstructive composite materials and adhesives. Such an evaluation would prove useful in diagnostic settings where composite/tissue bond strength may not be evaluated by destructive tensile or shear strength mechanical testing. Even in controlled laboratory settings, mechanical failure testing has proven to be a complex undertaking, with many variables affecting the failure rate of such bonds.[389, 390]

Experimental

Several non-carious, un-restored human molars were acquired less than 6 months from the extraction date and were mounted in dental stone at the roots. The occlusal surface (the surface of the tooth used to chew food) was abraded to expose the dentin using a carbide bur (US No. 55: Brasseler, Savannah, GA, USA) under an air-water spray within a custom cutting device (Computer Numeric Controlled Specimen Former: University of Iowa). Adhesives were applied and cured according to the manufacturers' directions. Each exposed dentin surface was dried with oil- and water-free pressurized air, with the manufacturers' prescribed distance and angle. Composite brands and lots are listed in Table XI-1.

Photo-polymerization was carried out using a dental light curing source (Optilux 500, Kerr, Orange, CA, USA). Light output was calibrated with a 6" integrating sphere and a UBS2000 spectrometer (Medical College of Georgia, Augusta, GA, USA).

Illumination was performed at a distance of approximately 1 mm from the sample surface.

Table XI-1. Dental composite brands, manufacturers, and lot numbers.

Brand	Type	Manufacturer	Lot
Bond Force	SE 1 step 1 bottle	Tokuyama Dental	021E29A
Optibond FL	ER 3 step	Kerr	3471551
Estilite Sigma Quick	Supra-nanofill	Tokuyama Dental	E464

Resin-based composite was applied to the dentin surface and cured with light for 40 seconds. Consistent illumination was achieved with the help of a custom holder for the illumination device. Following curing, samples were stored in artificial saliva and later sectioned with a water-cooled low-speed diamond saw the following day.

Spectra were collected as the average of 1800 scans at a spectral resolution of 1.2 cm^{-1}, with the exception of the dentin spectrum, which was acquired at a spectral resolution of 0.25 cm^{-1} in a preliminary study. Samples analyzed using a custom sample holder with a small (~0.8 mm) aperture, which provided high spatial selectivity at the expense of THz radiation throughput. The sample holder is shown schematically in Figure XI-1. A similar configuration has been reported in the scientific literature.[391] Reference spectra were collected with the empty sample holder after 5 minutes of purging with dried air. Sample spectra were collected with the sample in the holder following a similar 5 minute purging period. Foam padding in the sample holder allowed movement of the piece under study for analysis of various points of interest on the tissue and composite samples. Samples of the three composites were analyzed, as well as

composite/tissue structures formed with Bond Force and dentin, and Optibond FL and dentin.

Samples were analyzed without further preparation such as grinding or polishing. The thickness of each sample was measured with a digital caliper, and the absorption spectra and refractive index for each sample were calculated with the TeraView software (TPS Spectra Analysis and Acquisition Software, Version 1.17.1 TeraView Ltd., Cambridge, UK) and analyzed without further processing. Photographs of the dental composite and tissue samples are provided in Figures XI-2 through XI-10. Regions of each sample through which the THz pulses were transmitted are indicated with circles (and numbers, when appropriate) in each figure. Circles are illustrated approximately to scale. In cases where THz pulses were transmitted through the samples at an angle perpendicular to the dentin-composite interface, samples were analyzed in one direction such that the THz pulses traveled from the composite to the dentin, mimicking the orientation of a clinical measurement. Photographs shown in Figures XI-6 and XI-10 are taken from this perspective. Figure XI-11 depicts the anatomy of a human tooth and Figure XI-12 illustrates the sections from which the dentin/composite structures were fabricated.

Results

Absorption and refractive index spectra of the individual tissues and composites are illustrated in Figures XI-13 and XI-14, respectively. Spectra are shown without pre-processing. Periodic artifacts in the absorption spectra caused by multiple reflections within the sample materials were not removed by TDS truncation due to their close proximity to the main peak. Figure XI-15 illustrates the absorption spectra after DFF pre-processing, which provided a significantly smoother appearance. While a feature is visible near 73 cm^{-1} in the absorption spectrum of the Optibond FL composite, the authenticity of this feature is questionable due to the extremely high absorption at this

frequency. All spectra exhibit a strongly non-linear baseline, which is thought to be the result of scattering at each interface due to surface roughness. Aside from these characteristics, the absorption spectra of the individual components lack significant features, i.e. distinct absorption peaks. Slight curvature in the spectra of Estilite Sigma Quick and Optibond FL composites (near 78.5 and 50 cm^{-1}, respectively) suggest the presence of weak absorption features, but they are far from apparent in these data. The downward slopes of the absorption spectra at the high frequency ends are a product of the dynamic range of the THz-TDS and the extreme attenuation of radiation at these frequencies.

DFF pre-processing was not applied for smoothing of the refractive index spectra due to the variability in the useable bandwidth and extreme discontinuity at high frequencies. Nevertheless, each material exhibits a relatively unique refractive index spectrum within the useable bandwidth. Replicate samples were not available for the evaluation of the utility of the refractive index for material/tissue identification, though these results suggest the feasibility of THz-TDS for this application.

Spectra collected with composite/dentin structures are provided in Figures XI-16 through XI-19 after DFF pre-processing using the same parameters as before. Numbers in the figure legends correspond to the numbers and sampled regions illustrated in the photos of the composite/tissue samples. Figure XI-16 illustrates spectra collected with the Bond Force/dentin sample with an orientation similar to Figure XI-5 where THz pulses propagate through the sample parallel to the interface between the composite and tissue. Figure XI-17 displays spectra collected with the Bond Force/dentin sample with an orientation similar to Figure XI-6, where the THz pulses were transmitted through the sample in a direction perpendicular to the composite/tissue interface, entering on the side of the composite. Spectra illustrated in Figure XI-18 are those collected with the Optibond FL/dentin sample at an orientation similar to Figure XI-8, where THz pulses were transmitted through the sample at an angle parallel to the composite/tissue interface.

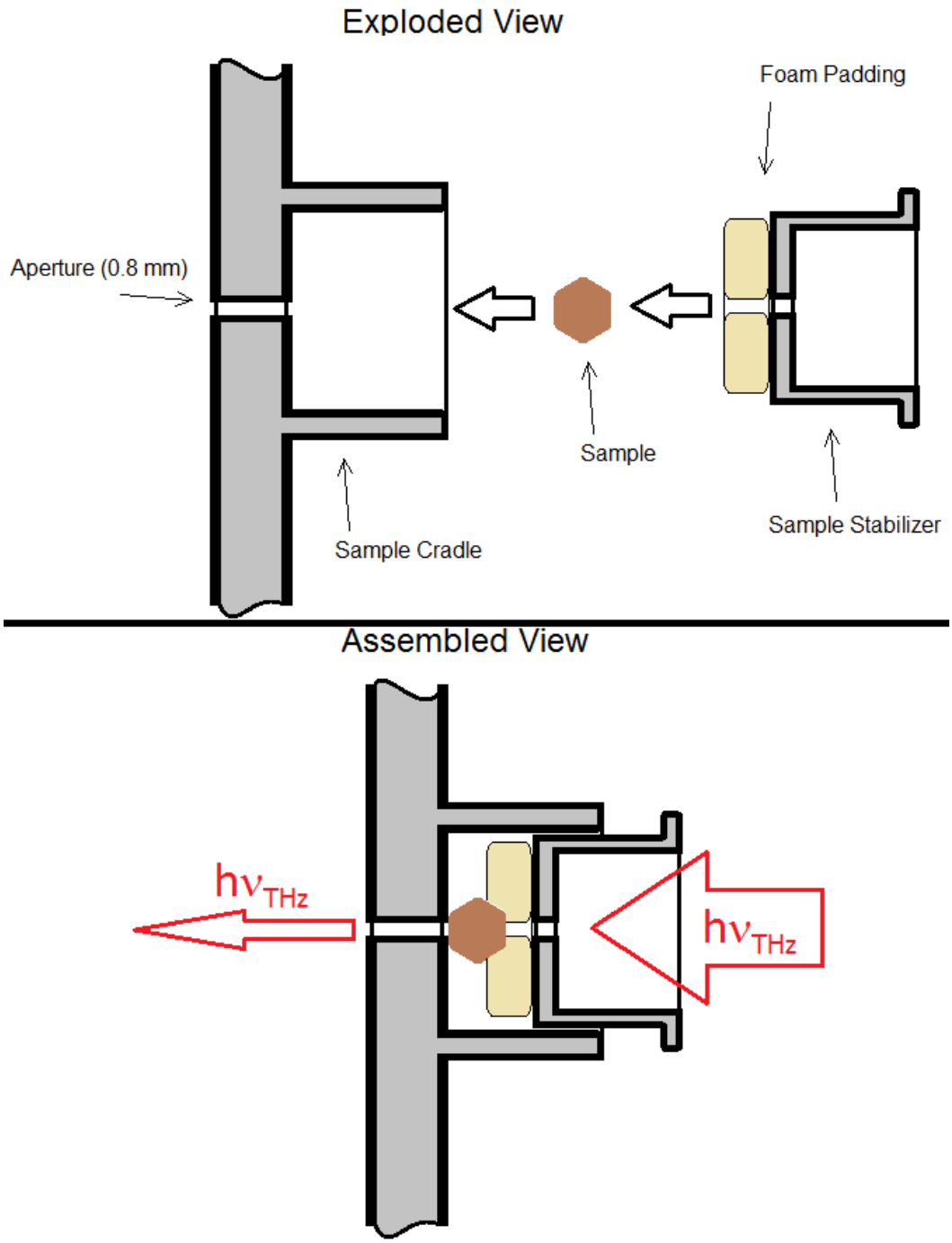


Figure XI-1. Schematic of the aperture sample holder used for dental tissue analysis.

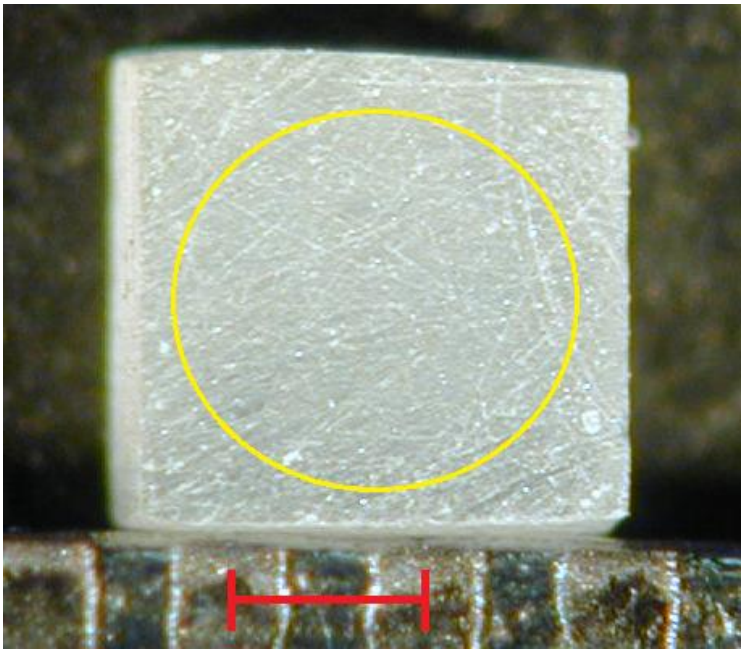


Figure XI-2. Bond Force composite sample. The red scale at the bottom of the image is 1/64th of one inch, or 0.397 mm.

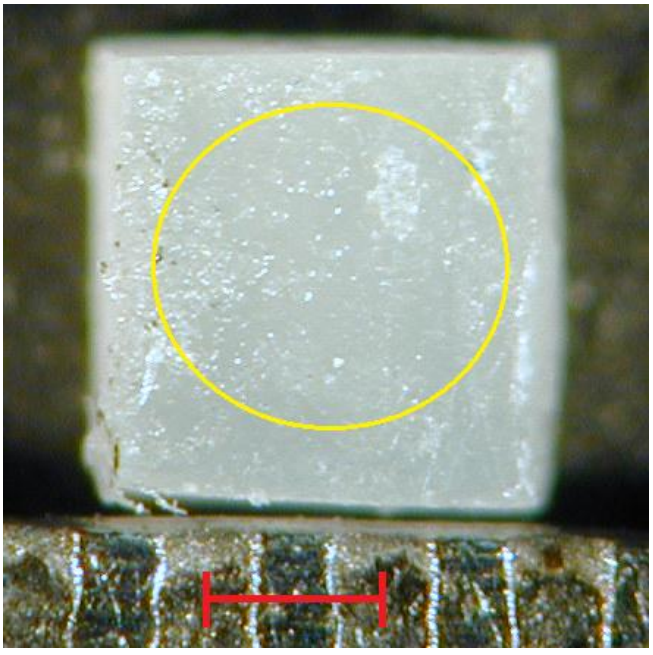


Figure XI-3. Estilite Sigma Quick composite sample. The red scale at the bottom of the image is 1/64th of one inch, or 0.397 mm.

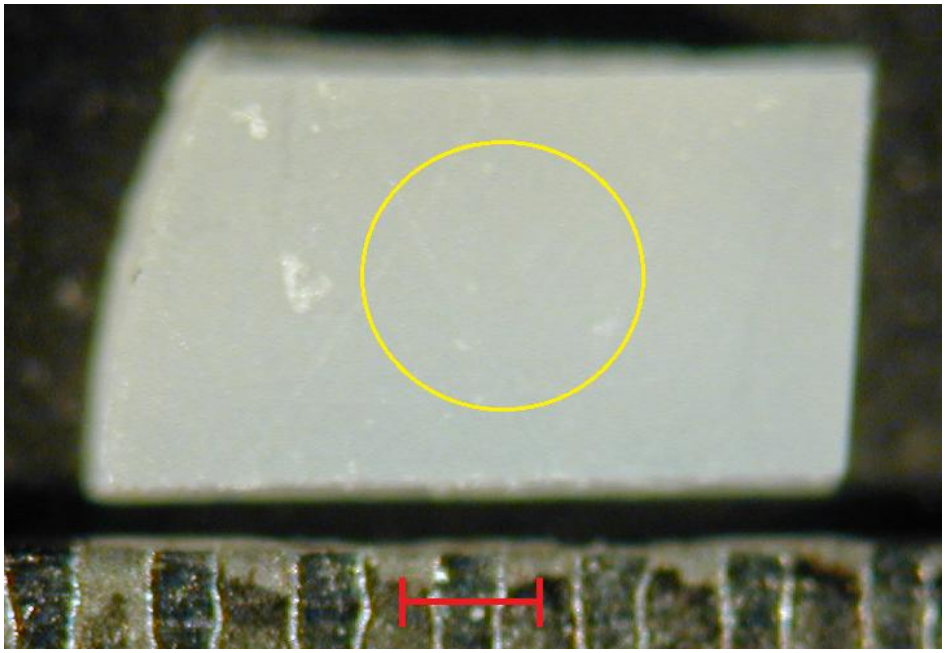


Figure XI-4. Optibond FL composite sample. The red scale at the bottom of the image is 1/64th of one inch, or 0.397 mm.

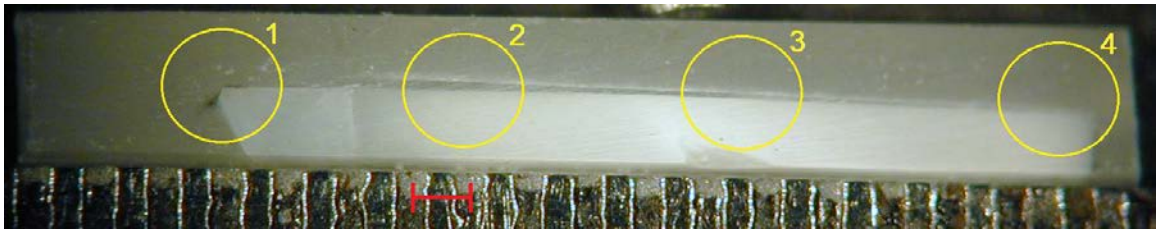


Figure XI-5. Bond Force composite bonded with dentin. The pale white portion of the sample is dentin, and the off-white portion is the composite material. The red scale at the bottom of the image is 1/64th of one inch, or 0.397 mm.

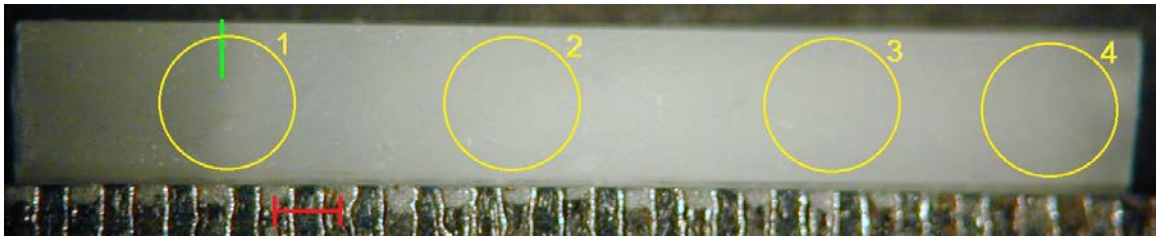


Figure XI-6. Bond Force composite bonded with dentin (same as Figure XI-5), shown from the Bond Force side of the sample. The boundary between the composite and tissue is faint, partially indicated by the bright green line near the top left corner of the image. The red scale at the bottom of the image is 1/64th of one inch, or 0.397 mm.

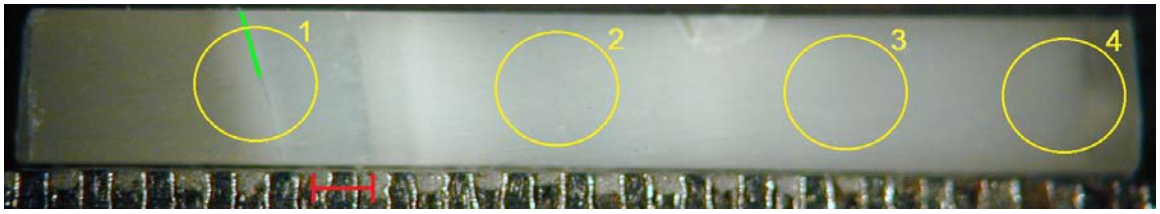


Figure XI-7. Bond Force composite bonded with dentin (same as Figures XI-5 and XI-6), shown from the dentin side of the sample. The boundary between the composite and tissue is readily visible, partially indicated by the bright green line near the top left corner of the image. The red scale at the bottom of the image is 1/64th of one inch, or 0.397 mm.

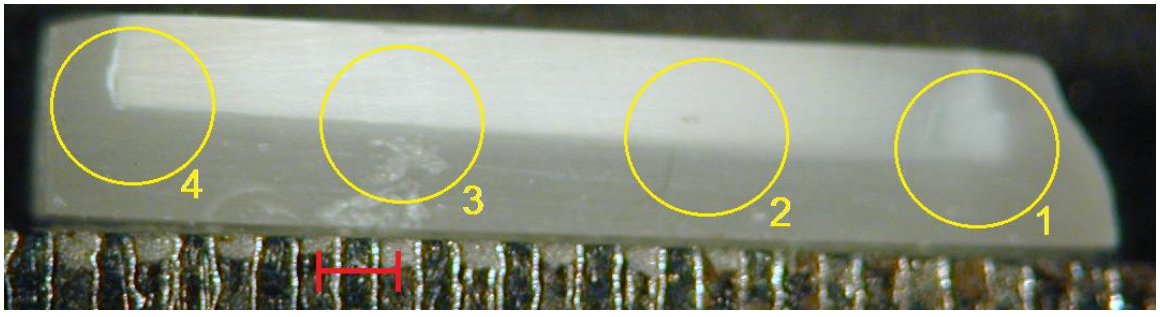


Figure XI-8. Optibond FL composite bonded with dentin. Dentin appears as the pale white portion, and the composite is the off-white portion. The red scale at the bottom of the image is 1/64th of one inch, or 0.397 mm.

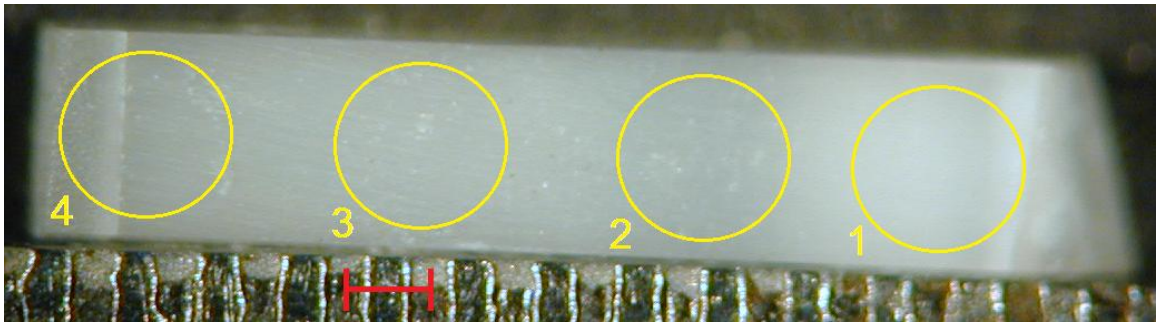


Figure XI-9. Optibond FL composite bonded with dentin (same as Figure XI-8), viewed from the side facing upward in Figure XI-8 (dentin). Dentin appears as the pale white portion, and the composite is the off-white portion. The red scale at the bottom of the image is 1/64th of one inch, or 0.397 mm.

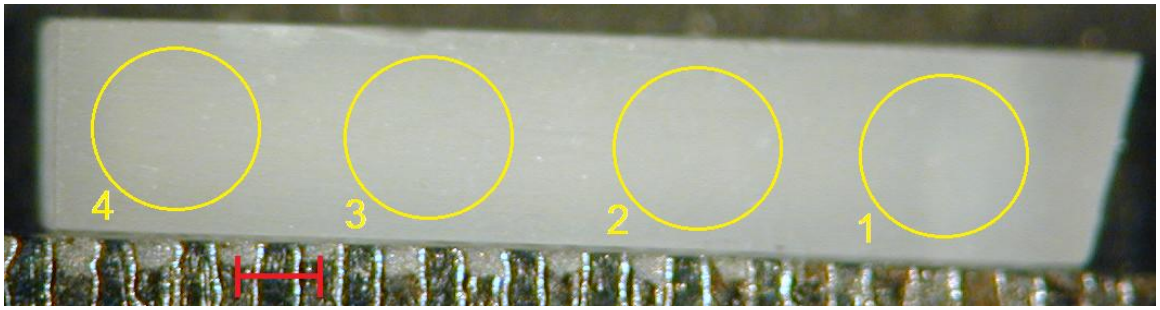


Figure XI-10. Optibond FL composite bonded with dentin (same as Figure XI-8), viewed from the side facing downward in Figure XI-8 (composite). Only the composite is visible from this perspective. The red scale at the bottom of the image is 1/64th of one inch, or 0.397 mm.

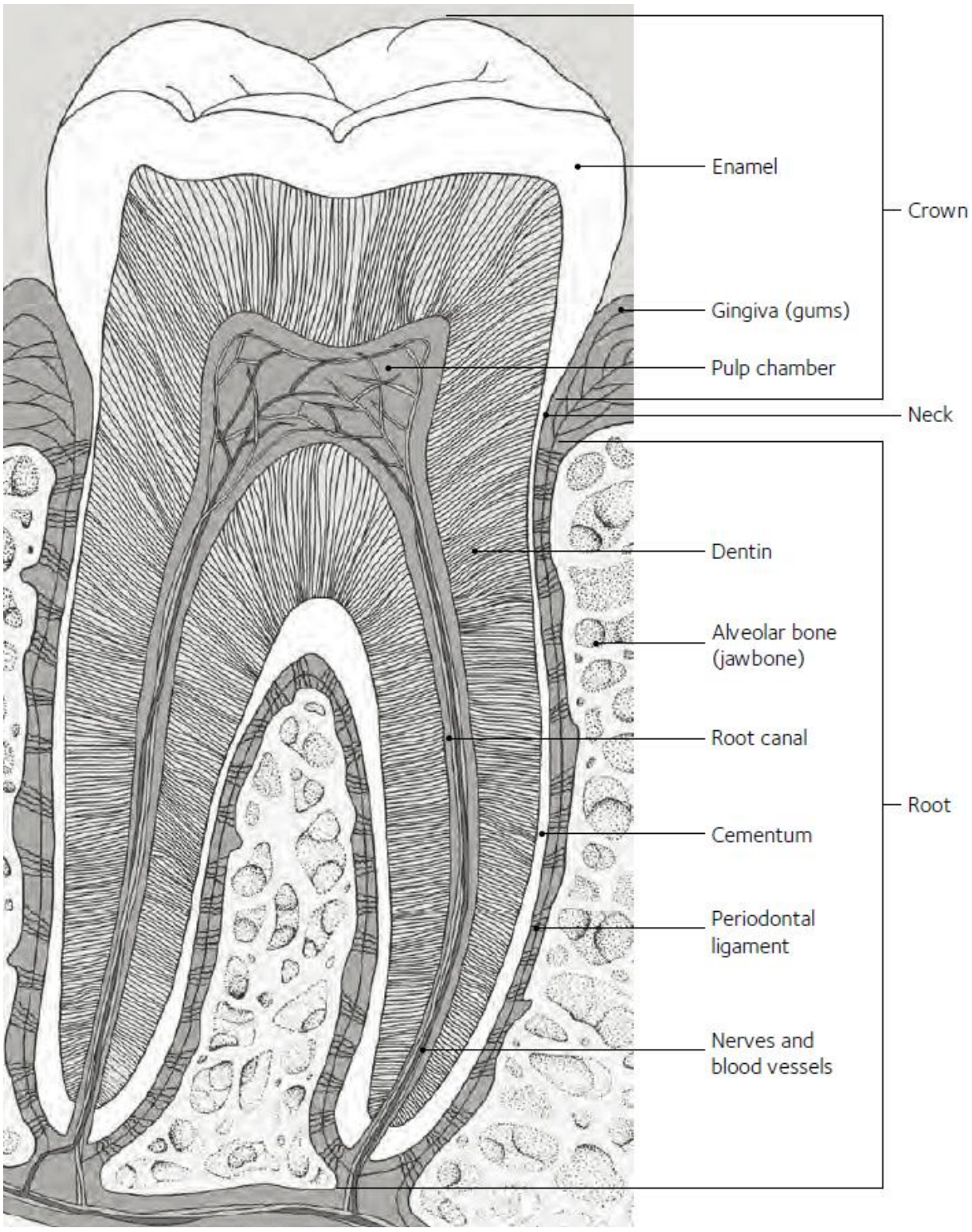


Figure XI-11. Anatomy of a human tooth. Note the tubular structures in the dentin and the variability in the orientation of these structures. Image courtesy of the American Dental Association.

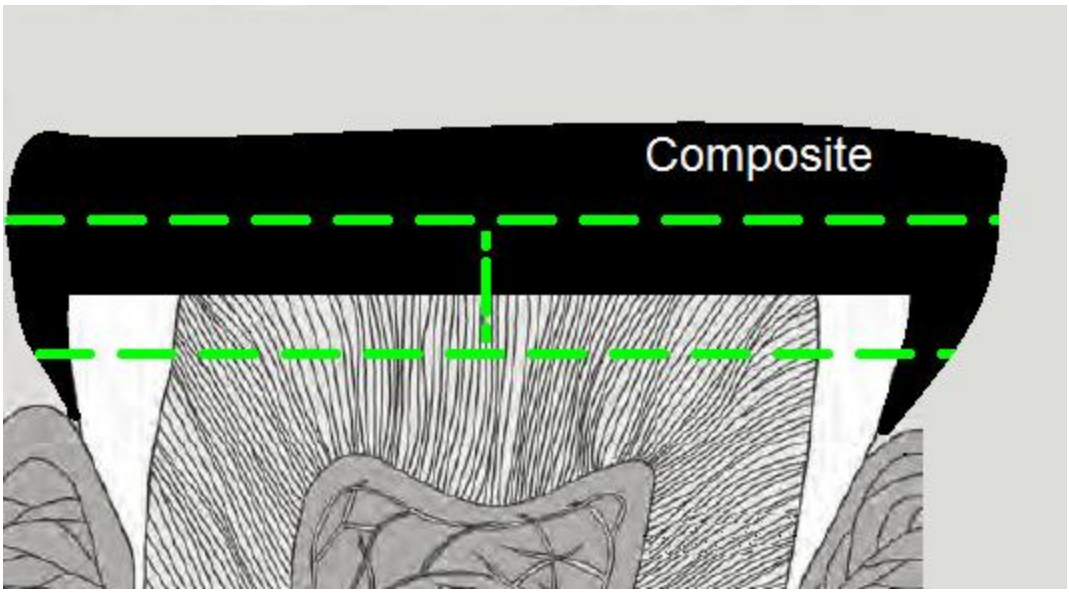


Figure XI-12. Human tooth with composite bonded to dentin. Green dashed lines represent sections cut for samples reported herein. Image adapted with permission from the American Dental Association.

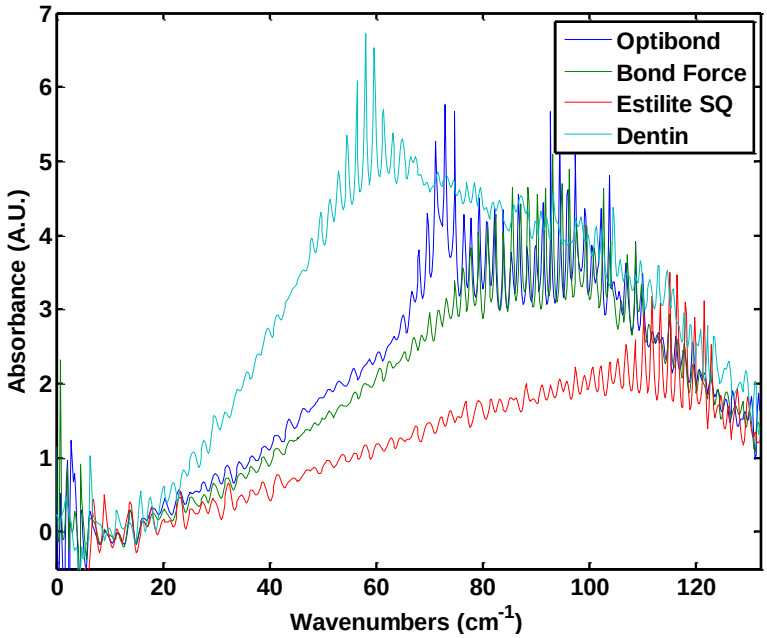


Figure XI-13. Absorption spectra of composite materials and dentin tissue (un-bonded).

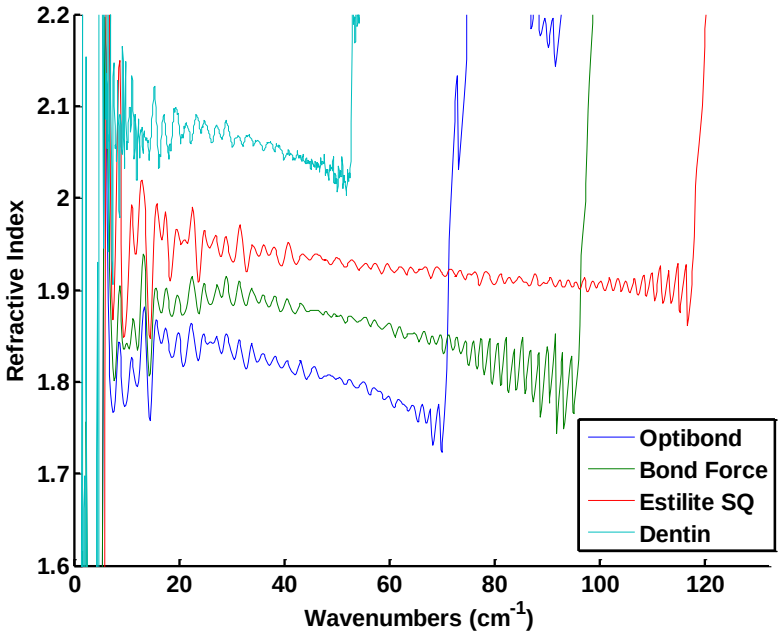


Figure XI-14. Refractive index spectra of un-bonded dentin and adhesive samples.

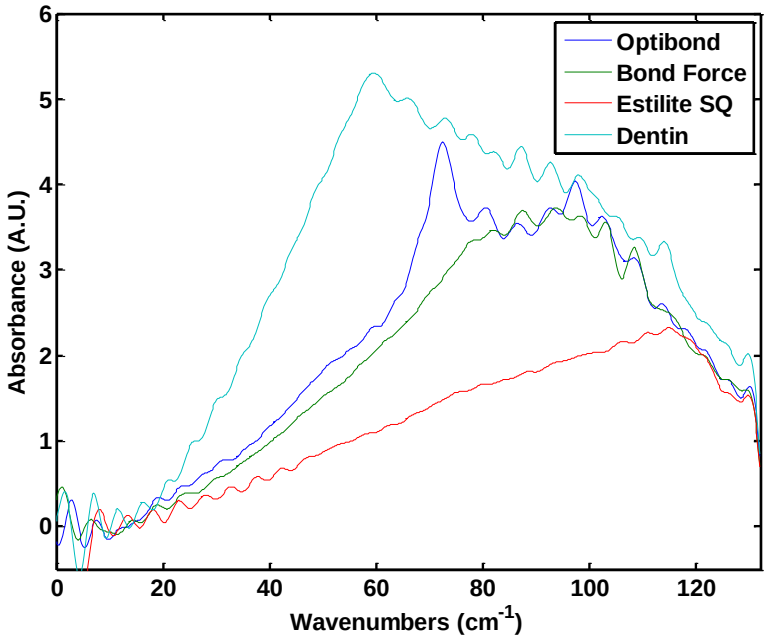


Figure XI-15. Absorption spectra of un-bonded dentin and adhesive samples after DFF pre-processing.

Data displayed in Figure XI-19 are those collected with the Optibond FL/dentin sample in a configuration similar to Figure XI-10 where THz pulses travelled through the sample at an angle perpendicular to the composite/tissue interface, incident on the composite side of the sample.

Many of the spectra exhibit baseline shapes similar to those of the pure materials where strong increases in attenuation are observed from the low to high frequency end of the absorption spectra. The slope of this baseline is of greater magnitude in samples where THz pulses are transmitted through the sample perpendicular to the composite/tissue interface. Generally, strong spectral features are observed in the absorption spectra collected with THz pulses travelling through the sample parallel to the composite/tissue interface. The authenticity of these features was verified by comparison with hundred percent lines calculated from reference spectra collected with the empty sample holder, shown in Figure XI-20. Spectral features appearing below ~6.5 cm^{-1} are primarily the result of variability cause by the sample holder. Thus, this spectral range is disregarded in the interpretation of spectra collected with the composite/tissue samples. Absorption spectra collected with THz pulses propagating through the sample orthogonal to the composite/dentin interface appear to exhibit increased scattering in regions where delamination is observed in corresponding sample photographs. Refractive index spectra of these samples are provided in Figures XI-21 through XI-24 without pre-processing.

Upon closer inspection of the Bond Force/dentin sample in Figure XI-5, one can see delamination of the composite from the tissue near sampled regions 2, 3, and 4 (to a lesser extent). While the absorption spectra of this sample do not appear to provide strong evidence of this trend (the spectrum of site 3 is very similar to that of site 1), the refractive index spectra from analysis perpendicular to the composite/tissue interface indicate a change between the two regions. The refractive index is lower (~2.93) at site 1, while it is higher (~3.1-3.2) in sites where delamination appears. The lack of this trend in the spectra collected with the interface parallel to incident THz radiation is attributed

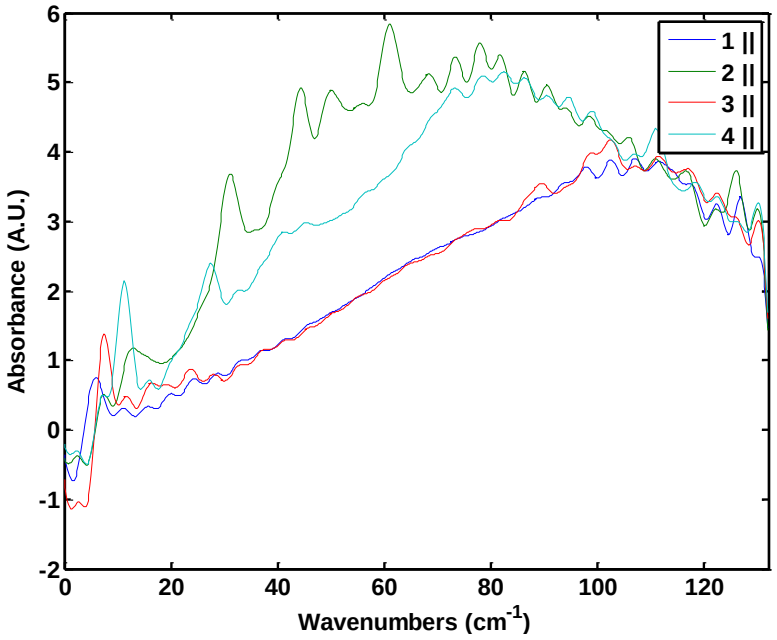


Figure XI-16. Absorption spectra collected with Bond Force/dentin structures with THz radiation transmitted parallel to the composite/tissue interface.

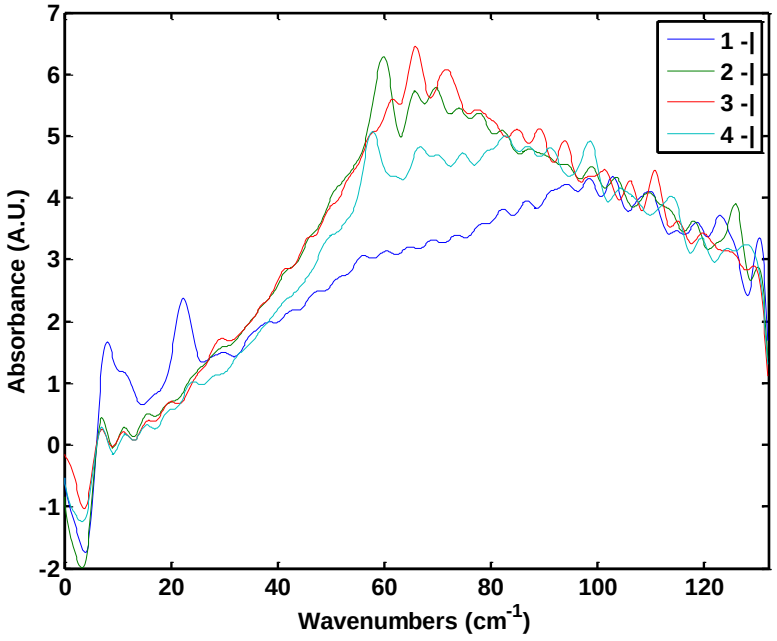


Figure XI-17. Absorption spectra collected with Bond Force/dentin structures with THz radiation transmitted perpendicular to the composite/tissue interface.

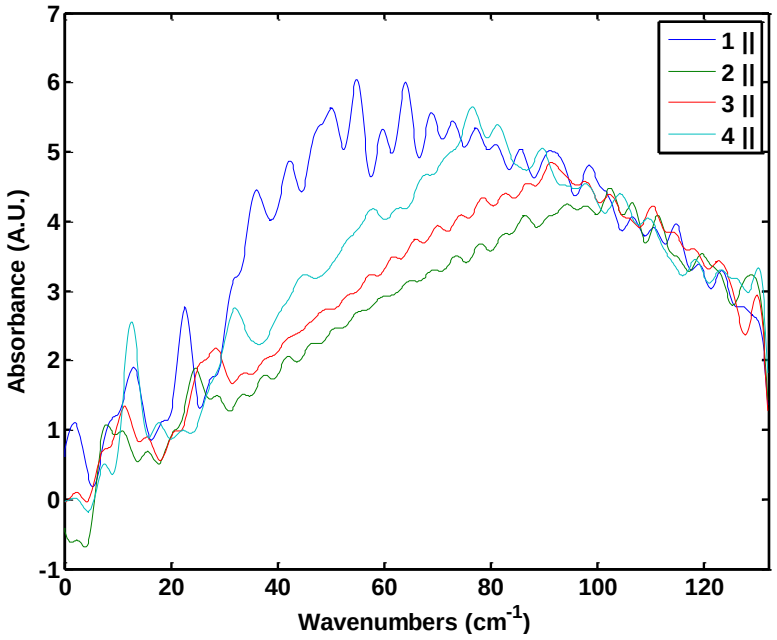


Figure XI-18. Absorption spectra collected with Optibond FL/dentin structures with THz radiation transmitted parallel to the composite/tissue interface.

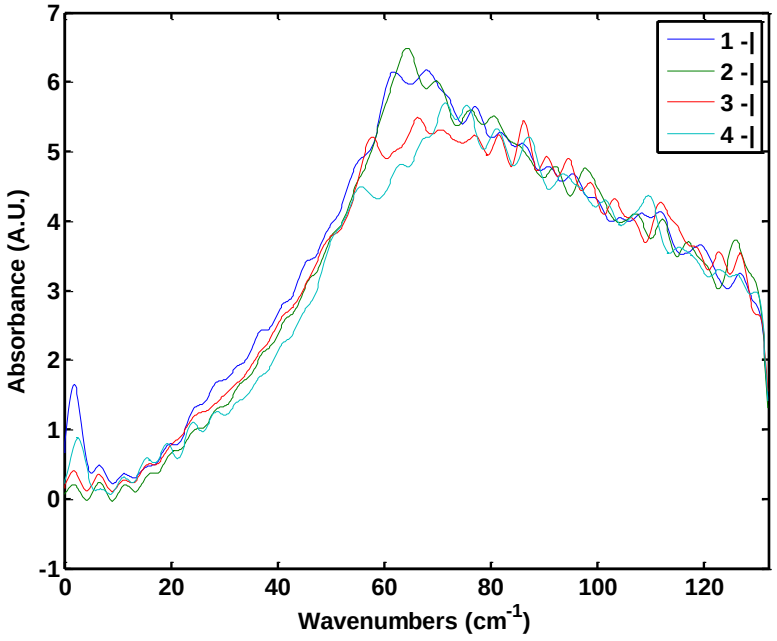


Figure XI-19. Absorption spectra collected with Optibond FL/dentin structures with THz radiation transmitted perpendicular to the composite/tissue interface.

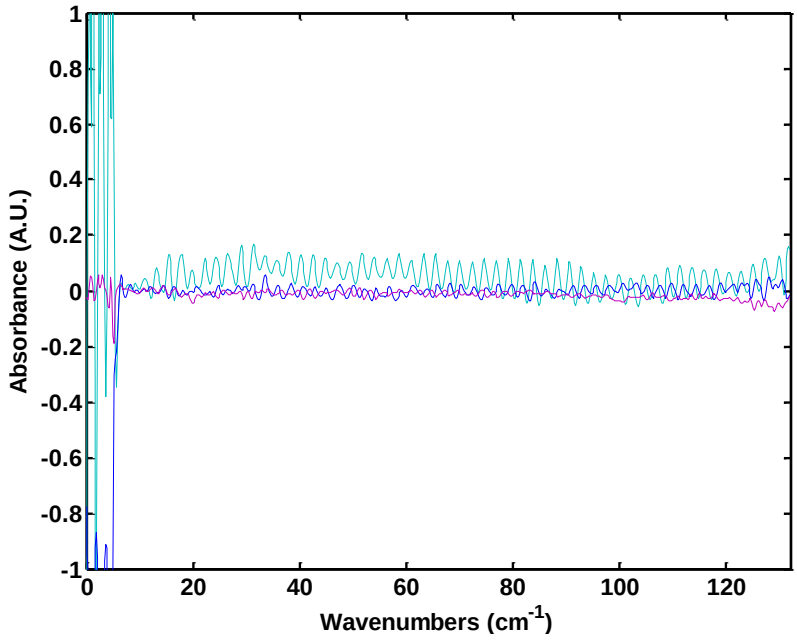


Figure XI-20. Hundred percent lines calculated from TDSs collected with the empty sample holder. Note the large magnitude of spectral features near the low-frequency end.

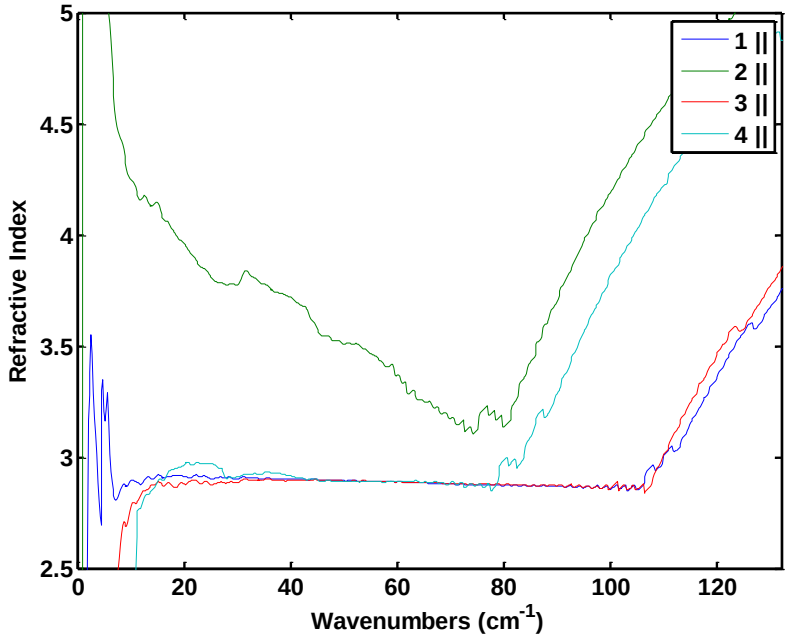


Figure XI-21. Refractive index spectra of Bond Force/dentin samples where THz pulses are transmitted through the sample parallel to the composite/tissue interface.

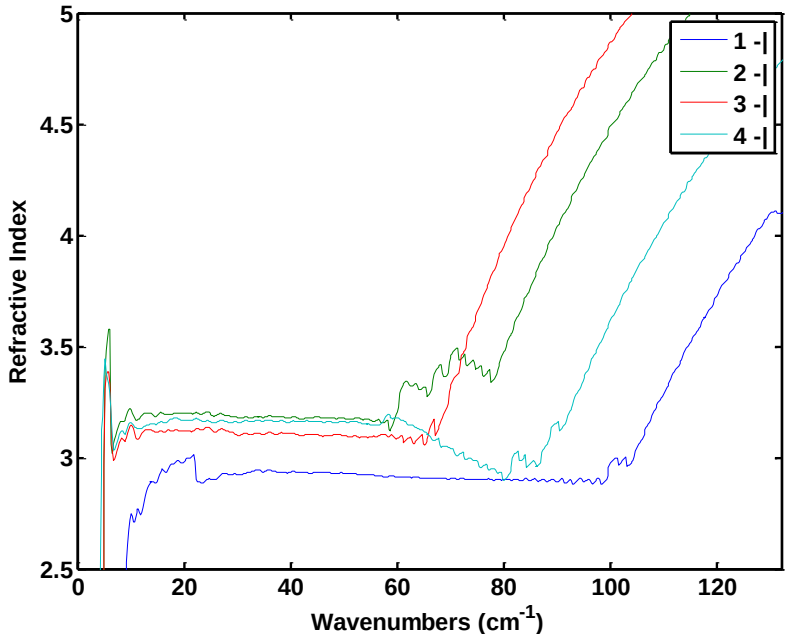


Figure XI-22. Refractive index spectra of Bond Force/dentin samples where THz pulses are transmitted through the sample perpendicular to the composite/tissue interface.

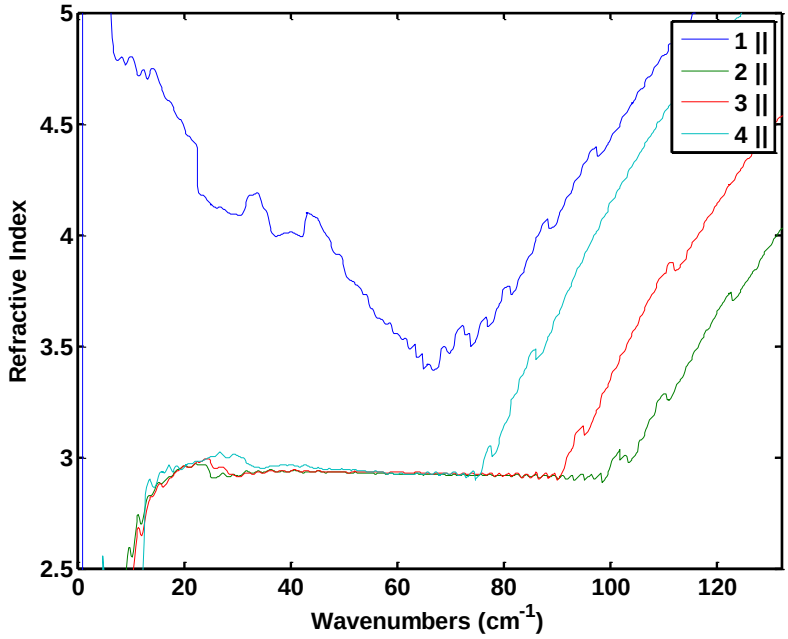


Figure XI-23. Refractive index spectra of Optibond FL/dentin samples where THz pulses are transmitted through the sample parallel to the composite/tissue interface.

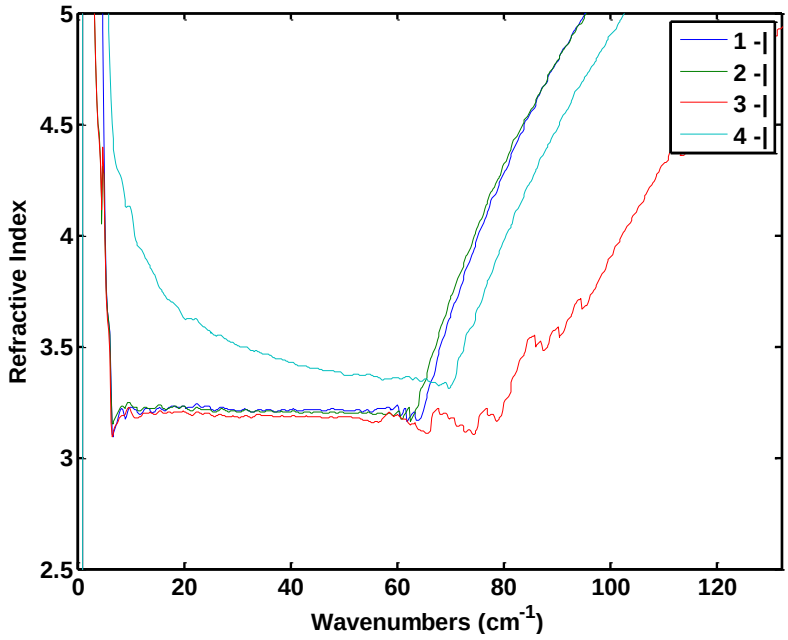


Figure XI-24. Refractive index spectra of Optibond FL/dentin samples where THz pulses are transmitted through the sample perpendicular to the composite/tissue interface.

to a minimal number of photons traversing the interface between each material, providing no spectral information related to the bond between the two materials.

The Optibond FL/dentin sample illustrated in Figure XI-8 appears to be free of delamination. Refractive index spectra collected with THz pulses travelling parallel to the interface between the two materials are quite similar with the exception of site 1, as shown in Figure XI-23. Site 1 was analyzed in close proximity to a jagged edge toward the end of the sample, and the sample may have shifted slightly after the sample stabilizer was secured, causing scattering of THz radiation on the rough surface. Refractive index spectra shown in Figure XI-24 are also very similar, except the spectrum collected in site 4. This site contains two interfaces, where one is parallel to the incident THz pulses, and the other is orthogonal to the radiation. The cause of non-linearity of the refractive index spectrum collected at this site is not yet known. Had the non-linearity been attributed to interference between photons experiencing differing optical delays, the absorption

spectrum would be expected to contain spectral features which would set it apart from the other three sites analyzed on the sample.

The THz refractive indices of the composite/dentin structures (~3.2) are higher than would be predicted based on the refractive indices of the individual components (~1.8-2.1). Spectra were collected with two different sample holders using a polymer sample to examine whether the design of the aperture had an effect on the measured refractive index. A piece of polyoxymethylene used for polymer characterization was analyzed in the sample holder and subsequently analyzed in the holder used for tooth samples. Absorption spectra and refractive index spectra are provided in Figures XI-25 and XI-26, respectively. Each plot includes the standard deviation of three replicate measurements performed with each holder removed and reinserted in the sample compartment.

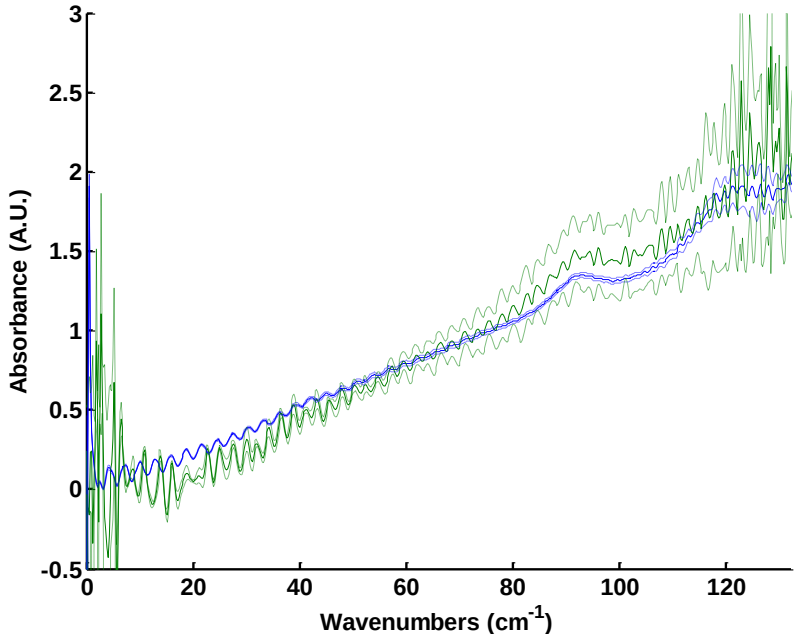


Figure XI-25. Absorption spectra of polyoxymethylene in the standard sample holder and tooth sample holder. The blue trace is the absorption spectrum collected with the standard holder, and the green trace was collected with the tooth sample holder. Dotted lines represent one standard deviation.

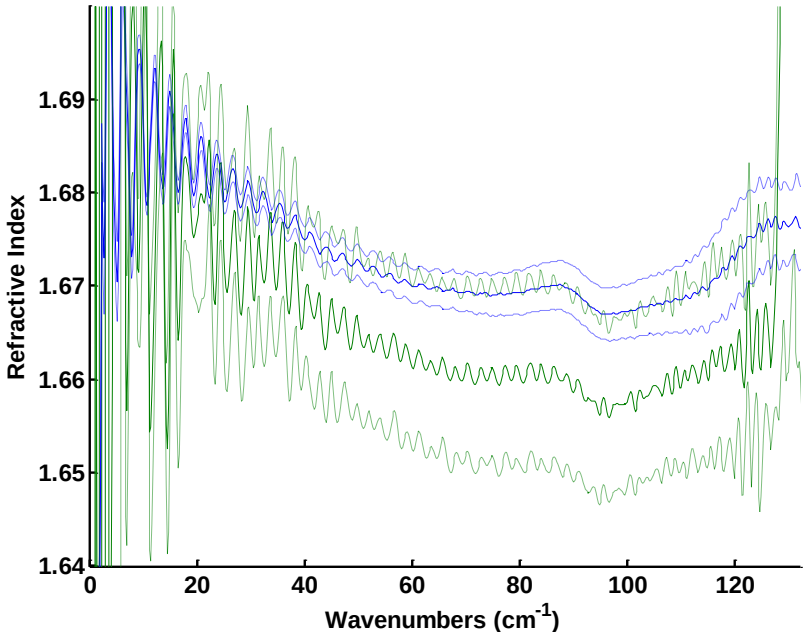


Figure XI-26. Refractive index spectra of polyoxymethylene in the standard sample holder and tooth sample holder. The blue trace is the spectrum collected with the standard holder, and the green trace was collected with the tooth sample holder. Dotted lines represent one standard deviation.

The tooth sample holder had a relatively insignificant impact on the measured spectral parameters, though the variability in these measurements was higher than that of the measurements performed with the standard holder. In the absorption spectrum, the measured values were slightly low in the low frequency range when using the tooth sample holder, but better agreement is seen in the high frequency end of the spectrum. The refractive index, when measured with the tooth sample holder, is lower than that of the standard holder, but the difference is not large enough to be the sole source of the change observed between the composite/tissue samples and those of the individual materials, as well as the change observed between regions exhibiting delamination and pristine interfaces.

Variability in the measured spectral parameters was also thought to be a product of interference between portions of the THz pulses transmitted though portions of the

sample with different refractive indices, or two regions having the same refractive index and differing thicknesses. This concept is illustrated in Figures XI-27 and XI-28.

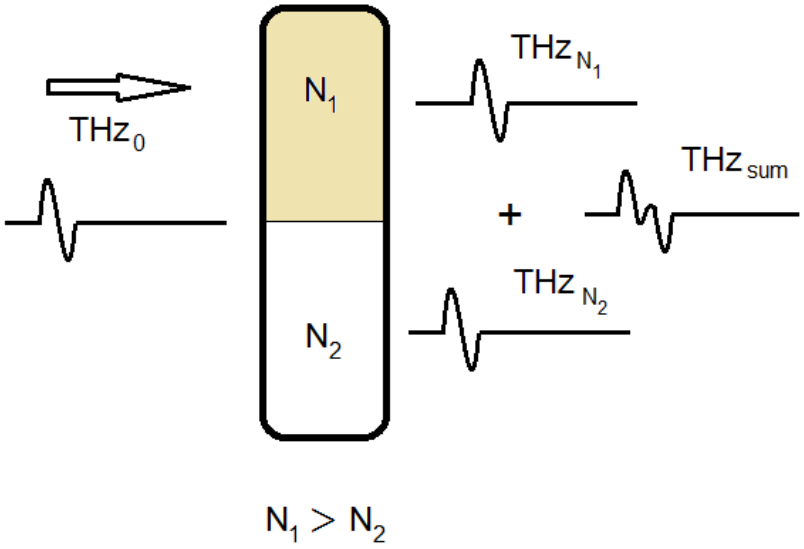


Figure XI-27. Illustration of interference between THz radiation transmitted through two regions of differing refractive indices and identical thicknesses.

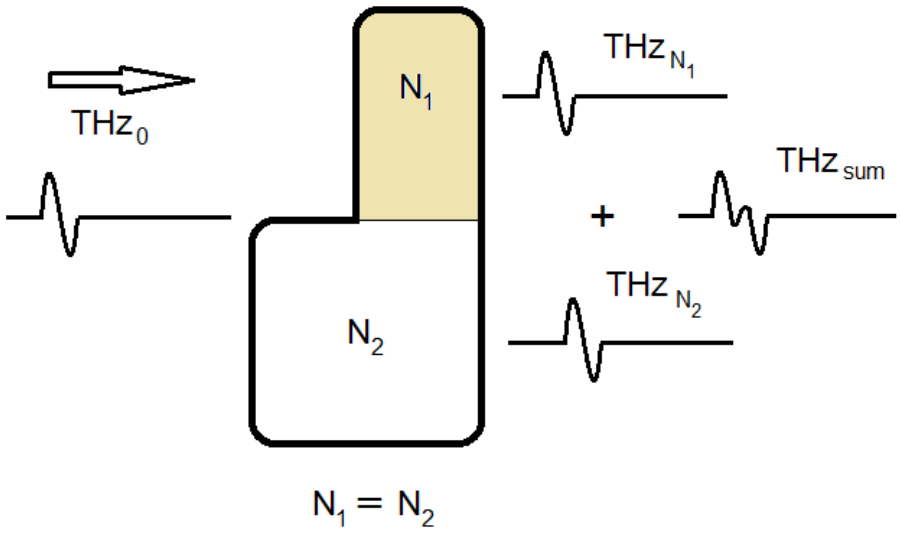


Figure XI-28. Illustration of interference between THz radiation transmitted through two regions of identical refractive indices and differing thicknesses.

This phenomenon was examined experimentally by cutting a disc of polyoxymethylene in half and performing spectral acquisition with the normal sample holder. Such a setup is analogous to Figure XI-27, where the polymer is material 1 is the polymer ($N_1 \approx 1.67$) and material 2 is the dried purging air ($N_2 = 1$). In this configuration, half of the THz beam travels through the polymer, and the remaining half propagates through air. The absorption spectrum and refractive index calculated for this sample are provided in Figures XI-29 and XI-30, respectively.

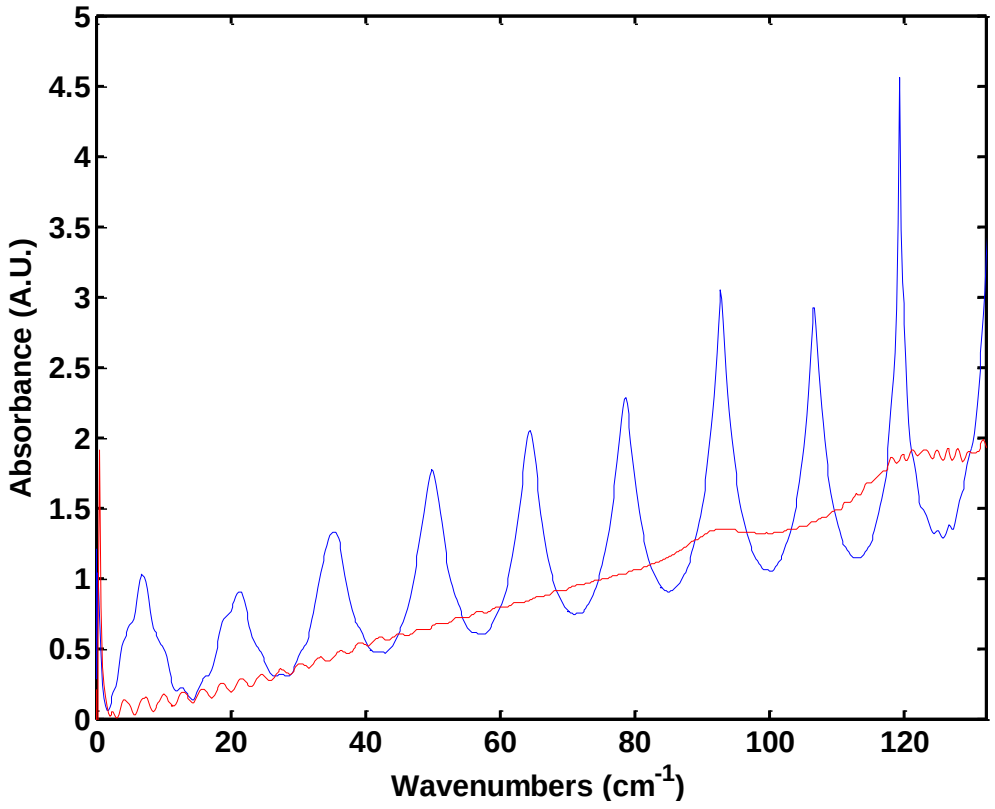


Figure XI-29. Absorption spectra of normal polyoxymethylene sample (red) and polyoxymethylene sample with TDS interference (blue).

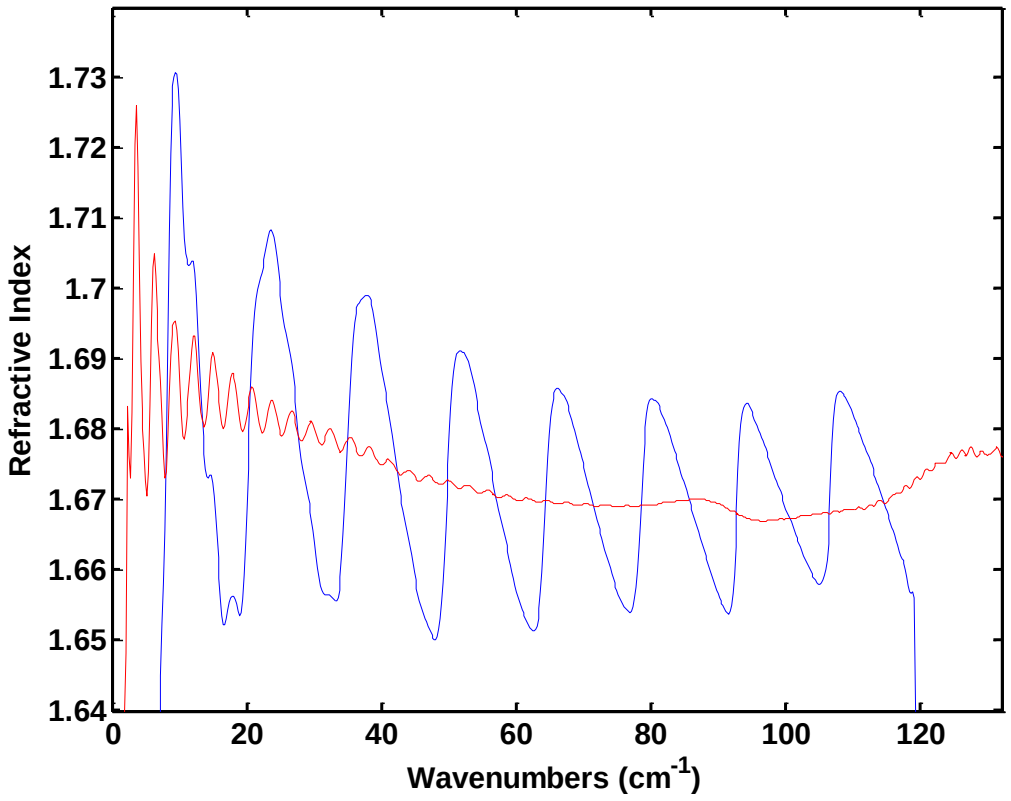


Figure XI-30. Refractive index spectra of normal polyoxymethylene sample (red) and polyoxymethylene sample with TDS interference (blue).

The absorption and refractive index spectrum collected with interference contain strong, periodic features. Absorption spectral peaks take on a nearly-triangular shape, and the magnitudes of the peaks are extremely high. Variability is increased in the refractive index spectrum, though the mean value remains unchanged. Changes in the measured spectral parameters for each composite/dentin sample at each location reflect differences in the sample morphology, but the source of these changes is difficult to pinpoint without further characterization of the sample using methods which may probe and image the interior of the samples.

Conclusions

Spectral parameters were measured for several common dental composite materials and dentin. No distinct reproducible spectral features were observed in the spectra from any of the samples. Analysis of composite/tissue structures revealed the utility of the THz-TDS method for investigation of delamination of composite materials from dental tissue. Absorption spectra of the samples appeared to show slightly increased scattering in regions where delamination occurred when THz pulses were transmitted through the sample at a right angle to the composite/dentin interface. For these same spectra, the refractive index appeared to provide a reliable indicator for separation of the materials, as the refractive index increased in regions where delamination was observed. The cause of this change is not yet understood, and further studies are warranted.

Suggested studies include imaging analyses for investigation of the interior of the sample/composite interfaces using either X-ray or THz-TDS images. The shape and angle of the interior features of these samples as well as the extent of delamination between the two materials can have a significant impact on the measured spectral parameters. Spectral features could then be correlated to the whole structure of the sample, rather than relying on visual analysis of exposed surfaces. Polished samples are expected to provide improved throughput of THz radiation, aiding interpretation of spectral data. Preparation of thicker composite materials would also allow evaluation of the authenticity of the weak absorption features observed with these samples.

REFERENCES

(1) Zimdars, D.; White, J.; Stuk, G.; Sucha, G.; Fichter, G.; Williamson, S. L. Time domain terahertz imaging of threats in luggage and personnel. *Int. J. High Speed Electron. Syst.* **2007***, 17*, 271-281.

(2) Zhou, Q.; Zhang, C.; Mu, K.; Jin, B.; Zhang, L.; Li, W.; Feng, R. Optical property and spectroscopy studies on the explosive 2,4,6-trinitro-1,3,5-trihydroxybenzene in the terahertz range. *Appl. Phys. Lett.* **2008***, 92*, 101106/1-101106/3.

(3) Hu, Y.; Huang, P.; Guo, L.; Wang, X.; Zhang, C. Terahertz spectroscopic investigations of explosives. *Phys. Lett. A* **2006***, 359*, 728-732.

(4) Lo, T.; Gregory, I. S.; Baker, C.; Taday, P. F.; Tribe, W. R.; Kemp, M. C. The very far-infrared spectra of energetic materials and possible confusion materials using terahertz pulsed spectroscopy. *Vib. Spectrosc.* **2006***, 42*, 243-248.

(5) Liu, H.; Zhong, H.; Karpowicz, N.; Chen, Y.; Zhang, X. Terahertz spectroscopy and imaging for defense and security applications. *Proc IEEE* **2007***, 95*, 1514-1527.

(6) Zeitler, J. A.; Kogermann, K.; Rantanen, J.; Rades, T.; Taday, P. F.; Pepper, M.; Aaltonen, J.; Strachan, C. J. Drug hydrate systems and dehydration processes studied by terahertz pulsed spectroscopy. *Int. J. Pharm.* **2007***, 334*, 78-84.

(7) Ho, L.; Mueller, R.; Roemer, M.; Gordon, K. C.; Heinaemaeki, J.; Kleinebudde, P.; Pepper, M.; Rades, T.; Shen, Y. C.; Strachan, C. J.; Taday, P. F.; Zeitler, J. A. Analysis of sustained-release tablet film coats using terahertz pulsed imaging. *J. Controlled Release* **2007***, 119*, 253-261.

(8) Huang, S. Y.; Wang, Y. X. J.; Yeung, D. K. W.; Ahuja, A. T.; Zhang, Y.; Pickwell-Macpherson, E. Tissue characterization using terahertz pulsed imaging in reflection geometry. *Phys. Med. Biol.* **2009***, 54*, 149-60.

(9) Shen, Y. C.; Taday, P. F.; Pepper, M. Elimination of scattering effects in spectral measurement of granulated materials using terahertz pulsed spectroscopy. *Appl. Phys. Lett.* **2008***, 92*, 051103/1-051103/3.

(10) Inc., C. Operator's manual, Vitesse diode-pumped modelocked Ti:sapphire laser. **2002**, 1-118.

(11) French, P. M. W. Ultrafast solid-state lasers. *Contemporary Physics* **1996***, 37*, 283-301.

(12) Griffiths, P. R.; de Haseth, J. A. Fourier transform infrared spectrometry. **2007**, 529.

(13) Gordy, W.; Cook, R. L. In *Techniques of chemistry, vol. 56: Microwave molecular spectra;* John Wiley and Sons: 1984; , pp 929.

(14) Redo-Sanchez, A.; Zhang, X. Terahertz science and technology trends. *IEEE J. Sel. Top. Quantum Electron.* **2008***, 14*, 260-269.

(15) Bigourd, D.; Cuisset, A.; Hindle, F.; Matton, S.; Fertein, E.; Bocquet, R.; Mouret, G. Detection and quantification of multiple molecular species in mainstream cigarette smoke by continuous-wave terahertz spectroscopy. *Opt. Lett.* **2006***, 31*, 2356-2358.

(16) Plusquellic, D. F.; Siegrist, K.; Heilweil, E. J.; Esenturk, O. Applications of terahertz spectroscopy in biosystems. *ChemPhysChem* **2007***, 8*, 2412-2431.

(17) Hindle, F.; Cuisset, A.; Bocquet, R.; Mouret, G. Continuous-wave terahertz by photomixing: applications to gas phase pollutant detection and quantification. *C. R. Phys.* **2008***, 9*, 262-275.

(18) TeraView Limited Terahertz radiation. *http://teraview.co.uk/terahertz/id/27* (accessed 5/14, 2009).

(19) Picometrix: An API Company What is terahertz (THz)? *http://www.picometrix.com/pico_products/terahertz_whatis.asp* (accessed 5/14, 2009).

(20) Ikeda, T.; Matsushita, A.; Tatsuno, M.; Minami, Y.; Yamaguchi, M.; Yamamoto, K.; Tani, M.; Hangyo, M. Investigation of inflammable liquids by terahertz spectroscopy. *Appl. Phys. Lett.* **2005***, 87*, 034105/1-034105/3.

(21) Yamamoto, K.; Yamaguchi, M.; Miyamaru, F.; Tani, M.; Hangyo, M.; Ikeda, T.; Matsushita, A.; Koide, K.; Tatsuno, M.; Minami, Y. Noninvasive inspection of C-4 explosive in mails by terahertz time-domain spectroscopy. *Jpn. J. Appl. Phys. , Part 2* **2004***, 43*, L414-L417.

(22) Hoshina, H.; Sasaki, Y.; Hayashi, A.; Otani, C.; Kawase, K. Noninvasive mail inspection system with terahertz radiation. *Appl. Spectrosc.* **2009***, 63*, 81-86.

(23) Shen, X.; Dietlein, C. R.; Grossman, E.; Popovic, Z.; Meyer, F. G. Detection and segmentation of concealed objects in terahertz images. *IEEE Trans. Image Process.* **2008***, 17*, 2465-75.

(24) Wang, Y.; Zhao, Z.; Chen, Z.; Kang, K.; Feng, B.; Zhang, Y. Terahertz absorbance spectrum fitting method for quantitative detection of concealed contraband. *J. Appl. Phys.* **2007***, 102*, 113108/1-113108/6.

(25) Koch, M. Terahertz technology: a land to be discovered. *Opt. Photonics News* **2007***, 18*, 20-25.

(26) Kimmitt, M. F. Restrahlen to T-rays - 100 years of terahertz radiation. *J. Biol. Phys.* **2003***, 29*, 77-85.

(27) Picometrix: An API Company Leading supplier of high-speed optical receivers and terahertz instrumentation. *http://www.picometrix.com* (accessed 6/20, 2011).

(28) Han, P. Y.; Tani, M.; Usami, M.; Kono, S.; Kersting, R.; Zhang, X. -. A direct comparison between terahertz time-domain spectroscopy and far-infrared Fourier transform spectroscopy. *J. Appl. Phys.* **2001***, 89*, 2357-2359.

(29) Grischkowsky, D.; Keiding, S.; Van, E., Martin; Fattinger, C. Far-infrared time-domain spectroscopy with terahertz beams of dielectrics and semiconductors. *J. Opt. Soc. Am. B: Opt. Phys.* **1990***, 7*, 2006-15.

(30) Beard, M. C.; Turner, G. M.; Schmuttenmaer, C. A. Terahertz spectroscopy. *J Phys Chem B* **2002***, 106*, 7146-7159.

(31) Melinger, J. S.; Laman, N.; Harsha, S. S.; Grischkowsky, D. Line narrowing of terahertz vibrational modes for organic thin polycrystalline films within a parallel plate waveguide. *Appl. Phys. Lett.* **2006***, 89*, 251110/1-251110/3.

(32) Baxter, J. B.; Guglietta, G. W. Terahertz spectroscopy. *Anal. Chem. (Washington, DC, U. S.)* **2011***, 83*, 4342-4368.

(33) Sakai, K.; Tani, M. Introduction to terahertz pulses. *Top. Appl. Phys.* **2005***, 97*, 1-30.

(34) Gu, P.; Tani, M. Terahertz radiation from semiconductor surfaces. *Top. Appl. Phys.* **2005***, 97*, 63-97.

(35) Jepsen, P. U.; Jacobsen, R. H.; Keiding, S. R. Generation and detection of terahertz pulses from biased semiconductor antennas. *J Opt Soc Am B* **1996***, 13*, 2424-2436.

(36) Cai, Y.; Brener, I.; Lopata, J.; Wynn, J.; Pfeiffer, L.; Stark, J. B.; Wu, Q.; Zhang, X. C.; Federici, J. F. Coherent terahertz radiation detection: Direct comparison between free-space electro-optic sampling and antenna detection. *Appl. Phys. Lett.* **1998***, 73*, 444-446.

(37) Duvillaret, L.; Garet, F.; Roux, J.; Coutaz, J. Analytical modeling and optimization of terahertz time-domain spectroscopy experiments using photoswitches as antennas. *IEEE J. Sel. Top. Quantum Electron.* **2001***, 7*, 615-623.

(38) Chen, Q.; Tani, M.; Jiang, Z.; Zhang, X. -C. Electro-optic transceivers for terahertz-wave applications. *J Opt Soc Am B* **2001***, 18*, 823-831.

(39) Rabolt, J. F.; Bellar, R. The nature of apodization in Fourier transform spectroscopy. *Appl. Spectrosc.* **1981***, 35*, 132-5.

(40) Harris, F. J. On the use of windows for harmonic analysis with the discrete Fourier transform. *Proceedings of the IEEE* **1978***, 66*, 51-83.

(41) Norton, R. H.; Beer, R. New apodizing functions for Fourier spectrometry. *J. Opt. Soc. Am.* **1976***, 66*, 259-264.

(42) Galvao, R. K. H.; Hadjiloucas, S.; Zafiropoulos, A.; Walker, G. C.; Bowen, J. W.; Dudley, R. Optimization of apodization functions in terahertz transient spectrometry. *Opt. Lett.* **2007***, 32*, 3008-3010.

(43) Lorber, A.; Faber, K.; Kowalski, B. R. Net analyte signal calculation in multivariate calibration. *Anal. Chem.* **1997***, 69*, 1620-1626.

(44) Olivieri, A. C.; Faber, N. M.; Ferre, J.; Boque, R.; Kalivas, J. H.; Mark, H. Uncertainty estimation and figures of merit for multivariate calibration. *Pure Appl. Chem.* **2006***, 78*, 633-661.

(45) Zeitler, J. A.; Taday, P. F.; Newnham, D. A.; Pepper, M.; Gordon, K. C.; Rades, T. Terahertz pulsed spectroscopy and imaging in the pharmaceutical setting - a review. *J. Pharm. Pharmacol.* **2007***, 59*, 209-223.

(46) Shen, Y. Terahertz pulsed spectroscopy and imaging for pharmaceutical applications: A review. *Int. J. Pharm.* **2011***, 417*, 48-60.

(47) Schall, M.; Helm, H.; Keiding, S. R. Far infrared properties of electro-optic crystals measured by THz time-domain spectroscopy. *Int. J. Infrared Millimeter Waves* **1999***, 20*, 595-604.

(48) Newnham, D. A.; Taday, P. F. Pulsed terahertz attenuated total reflection spectroscopy. *Appl. Spectrosc.* **2008***, 62*, 394-398.

(49) Cleeton, C. E.; Williams, N. H. Electromagnetic waves of 1.1 cm wave-length and the absorption spectrum of ammonia. *Phys.Rev.* **1934***, 45*, 234-237.

(50) Barnes, R. B.; Benedict, W. S.; Lewis, C. M. The far infrared spectrum of H_2O. *Physical Review* **1935***, 47*, 918-921.

(51) Randall, H. M.; Dennison, D. M.; Ginsburg, N.; Weber, L. R. The far infrared spectrum of water vapor. *Phys.Rev.* **1937***, 52*, 160-174.

(52) Langley, S. P. The bolometer. *Nature* **1881***, 25*, 14-15, 16.

(53) Emerson, D. T. The work of Jagadis Chandra Bose: 100 years of millimeter-wave research. *IEEE T. Microw. Theory* **1997***, 45*, 2267--2273.

(54) Bose, J. C. On the determination of the wave-length of electric radiation by diffraction grating. *Proceedings of the Royal Society of London* **1896***, 60*, pp. 167-178.

(55) Rubens, H.; Nichols, E. F. Heat rays of great wave length. *Phys.Rev.(Series I)* **1897***, 4*, 314-323.

(56) Hagen, E.; Rubens, H. XVI. On some relations between the optical and the electrical qualities of metals. *Philosophical Magazine Series 6* **1904***, 7*, 157-179.

(57) Rubens, H.; von Baeyer, O. LXXX. On extremely long waves, emitted by the quartz mercury lamp. *Philosophical Magazine Series 6* **1911***, 21*, 689-695.

(58) Nichols, E. F.; Tear, J. D. Short electric waves. *Phys.Rev.* **1923***, 21*, 587-610.

(59) Arkadiewa, A. G. Short electromagnetic waves of wave-length up to 92 μm. *Nature* **1924***, 113*, 640.

(60) Golay, M. J. E. A pneumatic infra-red detector. *Rev. Sci. Instrum.* **1947***, 18*, 357-62.

(61) Kompfner, R.; Williams, N. T. Backward-wave tubes. *Proc.IRE.* **1953***, 41*, 1602-1611.

(62) King, W. C.; Gordy, W. One-to-two millimeter wave spectroscopy. IV. Experimental methods and results for OCS, CH_3F, and H_2O. *Phys.Rev.* **1954***, 93*, 407-412.

(63) Gebbie, H. A.; Vanasse, G. A. Interferometric spectroscopy in the far infrared. *Nature* **1956***, 178*, 432.

(64) Strong, J. Interferometry for the far infrared. *J. Opt. Soc. Am.* **1957***, 47*, 354-357.

(65) Boyle, W. S.; Rodgers, K. F. Performance characteristics of a new low-temperature bolometer. *J. Opt. Soc. Am.* **1959***, 49*, 66-69.

(66) Fray, S. J.; Oliver, J. F. C. Photoconductive detector of radiation of wavelength greater than 50 μ. *J. Sci. Instrum.* **1959***, 36*, 195.

(67) Putley, E. H. Impurity photoconductivity in n-type InSb. *Proc. Phys. Soc. , London* **1960***, 76*, 802-5.

(68) Low, F. J. Low-temperature germanium bolometer. *J. Opt. Soc. Am.* **1961***, 51*, 1300-1304.

(69) Putley, E. H. Solid-state devices for infrared detection. *J. Sci. Instrum.* **1966***, 43*, 857-68.

(70) Ulrich, R. Effective low-pass filters for far infrared frequencies. *Infrared Physics* **1967***, 7*, 65-74.

(71) Chang, T. Y.; Bridges, T. J.; Burkhardt, E. G. CW submillimeter laser action in optically pumped methyl fluoride, methyl alcohol, and vinyl chloride gases. *Appl. Phys. Lett.* **1970***, 17*, 249-251.

(72) Nuss, M.; Orenstein, J. Terahertz time-domain spectroscopy. **1998***, 74*, 7-50.

(73) Han, P. Y.; Zhang, X. Free-space coherent broadband terahertz time-domain spectroscopy. *Measurement Science and Technology* **2001***, 12*, 1747.

(74) Mourou, G.; Stancampiano, C. V.; Blumenthal, D. Picosecond microwave pulse generation. *Appl. Phys. Lett.* **1981***, 38*, 470-2.

(75) Smith, P. R.; Auston, D. H.; Johnson, A. M.; Augustyniak, W. M. Picosecond photoconductivity in radiation-damaged silicon-on-sapphire films. *Appl. Phys. Lett.* **1981***, 38*, 47-50.

(76) Auston, D. H.; Cheung, K. P.; Smith, P. R. Picosecond photoconducting Hertzian dipoles. *Appl. Phys. Lett.* **1984***, 45*, 284-286.

(77) DeFonzo, A. P.; Lutz, C. R. Optoelectronic transmission and reception of ultrashort electrical pulses. *Appl. Phys. Lett.* **1987***, 51*, 212-214.

(78) Fattinger, C.; Grischkowsky, D. Point source terahertz optics. *Appl. Phys. Lett.* **1988***, 53*, 1480-1482.

(79) Spence, D. E.; Kean, P. N.; Sibbett, W. 60-Femtosecond pulse generation from a self-mode-locked titanium-doped sapphire laser. *Opt. Lett.* **1991***, 16*, 42-4.

(80) Hollenstein, H.; Winther, F. K-structure analysis of the A and E components of the torsional fundamental of acetaldehyde and investigation of torsional combination bands. *J. Mol. Spectrosc.* **1978***, 71*, 118-144.

(81) Bauder, A.; Lovas, F. J.; Johnson, D. R. Microwave spectra of molecules of astrophysical interest IX. Acetaldehyde. *Journal of Physical and Chemical Reference Data* **1976***, 5*, 53-78.

(82) Kleiner, I.; Lovas, F. J.; Godefroid, M. Microwave spectra of molecules of astrophysical interest. XXIII. Acetaldehyde. *Journal of Physical and Chemical Reference Data* **1996***, 25*, 1113-1210.

(83) Lin, C. C.; Kilb, R. W. Microwave spectrum and internal barrier of acetaldehyde. *J. Chem. Phys.* **1956***, 24*, 631-631.

(84) Kleiner, I.; Godefroid, M.; Herman, M.; McKellar, A. R. W. The fundamental torsion band in acetaldehyde. *J. Mol. Spectrosc.* **1990**, *142*, 238-253.

(85) Baker, J. G.; Petty, G. C. Torsional motion in acetaldehyde studied by microwave and double resonance spectroscopy. *J. Mol. Struct.* **1988**, *189*, 227-238.

(86) Liang, W.; Baker, J. G.; Herbst, E.; Booker, R. A.; De Lucia, F. C. The millimeter-wave spectrum of acetaldehyde in its two lowest torsional states. *J. Mol. Spectrosc.* **1986**, *120*, 298-310.

(87) O'Loane, J. K. Some absorption bands in the far infrared. *J. Chem. Phys.* **1953**, *21*, 669-674.

(88) Fateley, W. G.; Miller, F. A. Torsional frequencies in the far infrared—I: Molecules with a single methyl rotor. *Spectrochimica Acta* **1961**, *17*, 857-868.

(89) Turner, P. H.; Cox, A. P. Centrifugal distortion and internal rotation in the microwave spectrum of acetaldehyde. *Chemical Physics Letters* **1976**, *42*, 84-88.

(90) Turner, P. H.; Cox, A. P.; Hardy, J. A. Microwave spectrum of acetaldehyde with asymmetric CH_2D and CHD_2 internal rotors. *J.Chem.Soc. Faraday Trans.2* **1981**, *77*, 1217-1231.

(91) Martinache, L.; Bauder, A. Microwave spectrum of acetaldehyde-1-d_1: Deuterium quadrupole splittings and internal rotation analysis. *Chemical Physics Letters* **1989**, *164*, 657-663.

(92) Kleiner, I.; Lopez, J. C.; Blanco, S.; McKellar, A. R. W.; Moazzen-Ahmadi, N. The ground and first torsional states of CD_3CHO. *J. Mol. Spectrosc.* **1999**, *197*, 275-288.

(93) Boucher, D.; Burie, J.; Bauer, A.; Dubrulle, A.; Demaison, J. Microwave spectra of molecules of astrophysical interest. XIX. Methyl cyanide. *Journal of Physical and Chemical Reference Data* **1980**, *9*, 659-720.

(94) Carlotti, M.; Di Lonardo, G.; Fusina, L.; Carli, B. The far-infrared spectrum of methyl cyanide, CH_3CN. *J. Mol. Spectrosc.* **1988**, *129*, 314-325.

(95) Crowder, G. A.; Cook, B. R. Acetonitrile: far-infrared spectra and chemical thermodynamic properties. Discussion of an entropy discrepancy. *J. Phys. Chem.* **1967**, *71*, 914-16.

(96) Sakai, K.; Ichimura, K.; Masumoto, H.; Kitagawa, Y. High-resolution spectroscopy of some gaseous molecules with a submillimetre Fourier transform spectrometer. *Infrared Physics* **1978**, *18*, 577-583.

(97) Schwaab, G. W.; Evenson, K. M.; Zink, L. R. Far-infrared self-broadening and pressure shift measurements of methyl cyanide. *Int J Infrared Millim Waves* **1993**, *14*, 1643-1655.

(98) Cosleou, J.; López, J. C.; Alonso, J. L.; Wlodarczak, G.; Demaison, J. The rotational spectrum of methyl cyanide-d_3 analysis of the low-lying excited vibrational states. *J. Mol. Spectrosc.* **1991**, *149*, 242-251.

(99) Matsumura, C.; Hirota, E.; Oka, T.; Morino, Y. Microwave spectrum of acetonitrile-d_3, CD_3CN. *J. Mol. Spectrosc.* **1962**, *9*, 366-380.

(100) Husson, N.; Goldman, A.; Orton, G. Spectroscopic line parameters of NH_3 and PH_3 in the far infrared. *Journal of Quantitative Spectroscopy and Radiative Transfer* **1982**, *27*, 505-515.

(101) Hadley, L. N.; Dennison, D. M. The microwave spectrum of ammonia. *Phys.Rev.* **1946**, *70*, 780-781.

(102) Foley, H. M.; Randall, H. M. Fine structure in the far infra-red spectrum of NH_3. *Phys.Rev.* **1941**, *59*, 171-173.

(103) Gordy, W.; Kessler, M. Microwave spectra: The hyperfine structure of ammonia. *Phys.Rev.* **1947**, *71*, 640-640.

(104) Fleming, J. W.; Gibson, M. J. Far-infrared absorption spectra of water vapor $H_2{}^{16}O$ and isotopic modifications. *J. Mol. Spectrosc.* **1976**, *62*, 326-337.

(105) Jerome M., D. The rotation-inversion spectrum of ammonia. *J. Mol. Spectrosc.* **1968**, *27*, 527-538.

(106) Kolbe, W. F.; Leskovar, B. Millimeter and submillimeter wave absorption by atmospheric pollutants and constituents. *Journal of Quantitative Spectroscopy and Radiative Transfer* **1983**, *30*, 463.

(107) Fusina, L.; Carlotti, M.; Di Lonardo, G.; Murzin, S. N.; Stepanov, O. N. Pure inversion and inversion-rotation spectra of $^{15}ND_3$ in the ground state. *J. Mol. Spectrosc.* **1991**, *147*, 71-83.

(108) Fulford, J. A. Line breadths in the ammonia spectrum. *Nature* **1960**, *188*, 1097-1098.

(109) Lightman, A.; Ben-Reuven, A. Line mixing by collisions in the far-infrared spectrum of ammonia. *J. Chem. Phys.* **1969**, *50*, 351-353.

(110) Potter, C. A.; Bushkovitch, A. V.; Rouse, A. G. Pressure broadening in the microwave spectrum of ammonia. *Phys.Rev.* **1951**, *83*, 987-989.

(111) Yamada, C.; Shigemune, T.; Hirota, E. Pressure broadening spectroscopy of formaldehyde, ammonia, isocyanic acid, trimethylene sulfide, and methyl chloride using far-infrared lasers. *J. Mol. Spectrosc.* **1975***, 54*, 261-274.

(112) Bak, B.; Knudsen, E. S.; Madsen, E. Microwave absorption of some organic vapors. *Phys.Rev.* **1949***, 75*, 1622-1623.

(113) Michielsen-Effinger, J. Spectre de rotation en microondes de la molécule d'alcool éthylique CH_3CH_2OH. *J.Phys.France* **1969***, 30*, 336-340,.

(114) Sasada, Y. Excited states of trans ethyl alcohol by microwave spectroscopy. *J. Mol. Struct.* **1988***, 190*, 93.

(115) Pearson, J. C.; Sastry, K. V. L. N.; Herbst, E.; De Lucia, F. C. Gauche ethyl alcohol: Laboratory assignments and interstellar identification. *Astrophys. J.* **1997***, 480*, 420.

(116) Durig, J. R.; Larsen, R. A. Torsional vibrations and barriers to internal rotation for ethanol and 2,2,2-trifluoroethanol. *J. Mol. Struct.* **1990***, 238*, 195-222.

(117) Pearson, J. C.; Brauer, C. S.; Drouin, B. J. The asymmetric top - asymmetric frame internal rotation spectrum of ethyl alcohol. *J. Mol. Spectrosc.* **2008***, 251*, 394.

(118) Lovas, F. J. Microwave spectra of molecules of astrophysical interest. XXI. Ethanol (C_2H_5OH) and propionitrile (C_2H_5CN). *Journal of Physical and Chemical Reference Data* **1982***, 11*, 251-312.

(119) Mukhopadhyay, I.; Lees, R. M. Far-infrared spectrum of excited torsional states of C-13 methanol. *Int J Infrared Millim Waves* **1995***, 16*, 99-115.

(120) Herbst, E.; Messer, J. K.; De Lucia, F. C.; Helminger, P. A new analysis and additional measurements of the millimeter and submillimeter spectrum of methanol. *J. Mol. Spectrosc.* **1984***, 108*, 42-57.

(121) Xu, L.; Lovas, F. J. Microwave spectra of molecules of astrophysical interest. XXIV. Methanol (CH_3OH and $^{13}CH_3OH$). *Journal of Physical and Chemical Reference Data* **1997***, 26*, 17-156.

(122) Lees, R. M.; Lovas, F. J.; Kirchhoff, W. H.; Johnson, D. R. Microwave spectra of molecules of astrophysical interest: III. Methanol. *Journal of Physical and Chemical Reference Data* **1973***, 2*, 205-214.

(123) Moruzzi, G.; Prevedelli, M.; Evenson, K. M.; Jennings, D. A.; Vanek, M. D.; Inguscio, M. Ultrahigh resolution far-infrared spectroscopy of methanol. *Infrared Physics* **1989***, 29*, 541-549.

(124) De Lucia, F. C.; Herbst, E.; Anderson, T.; Helminger, P. The analysis of the rotational spectrum of methanol to microwave accuracy. *J. Mol. Spectrosc.* **1989**, *134*, 395-411.

(125) Pearson, J. C.; Drouin, B. J.; Yu, S.; Gupta, H. Microwave spectroscopy of methanol between 2.48 and 2.77 THz. *J Opt Soc Am B* **2011**, *28*, 2549-2577.

(126) Lees, R. M.; Goulding, R. R. J.; Zhao, S.; Mukhopadhyay, I.; Johns, J. W. C. Torsion-rotation far-infrared spectrum of O-18 methanol. *Int J Infrared Millim Waves* **1994**, *15*, 2021-2030.

(127) Gerry, M. C. L.; Lees, R. M.; Winnewisser, G. The torsion-rotation microwave spectrum of $^{12}CH_3{}^{18}OH$ and the structure of methanol. *J. Mol. Spectrosc.* **1976**, *61*, 231-242.

(128) Mukhopadhyay, I.; Duan, Y. B. THz spectrum of methanol-d_4 and microwave spectrum of methanol-d_1. *Int J Infrared Millim Waves* **2003**, *24*, 855-898.

(129) Lees, R. M. Torsion-vibration-rotation interactions in methanol. II. Microwave spectrum of CD_3OD. *J. Chem. Phys.* **1972**, *56*, 5887-5890.

(130) Stern, V.; Belorgeot, C.; Kachmarsky, J.; Möller, K. D. Far infrared spectrum of the internal rotation of CH_3OH and CD_3OH. *J. Mol. Spectrosc.* **1977**, *67*, 244-264.

(131) Hilali, A. E.; Coudert, L. H.; Konov, I.; Klee, S. Analysis of the torsional spectrum of monodeuterated methanol CH_2DOH. *J. Chem. Phys.* **2011**, *135*, 194309.

(132) Stern, V.; Goff, N.; Kachmarsky, J.; Möller, K. D. Far-infrared internal rotation spectrum of CH_3OD and CD_3OD. *J. Mol. Spectrosc.* **1980**, *79*, 345-362.

(133) Mollabashi, M.; Lees, R. M.; Johns, J. W. C. Far infrared spectrum of CD_3OH methanol from 40-220 cm^{-1}. *Int J Infrared Millim Waves* **1993**, *14*, 1727-1753.

(134) Moruzzi, G.; Murphy, R. J.; Lees, R. M.; Predoi-Cross, A.; Billinghurst, B. E. High-resolution synchrotron-based Fourier transform spectroscopy of $CH_3{}^{17}OH$ in the 120-350 cm^{-1} far-infrared region. *Mol. Phys.* **2010**, *108*, 2343-2353.

(135) Pickett, H. M.; Scroggin, D. G. Microwave spectrum and internal rotation potential of propanal. *J. Chem. Phys.* **1974**, *61*, 3954-3958.

(136) Butcher, S. S.; Wilson, E. B. Microwave spectrum of propionaldehyde. *J. Chem. Phys.* **1964**, *40*, 1671-1678.

(137) Randell, J.; Hardy, J. A.; Cox, A. P. The microwave spectrum and potential function of propanal. *J. Chem. Soc., Faraday Trans. 2* , 1199.

(138) Smeyers, Y. G.; Villa, M. A theoretical determination of the methyl and aldehydic torsion far-infrared spectrum of propanal-d_0 with the vibrational zero point correction. *J. Chem. Phys.* **2002***, 116*, 4087.

(139) Durig, J. R.; Guirgis, G. A.; Bell, S.; Brewer, W. E. Far-infrared spectrum, ab initio calculations, conformational energy differences, barriers to internal rotation, and r_0 structure of propanal. *The Journal of Physical Chemistry A* **1997***, 101*, 9240-9252.

(140) Durig, J. R.; Compton, D. A. C.; McArver, A. Q. Low frequency vibrational spectra, methyl torsional potential functions, and internal rotational potential of propanal. *J. Chem. Phys.* **1980***, 73*, 719-724.

(141) Laurie, V. W. Microwave spectrum and internal rotation of ethyl cyanide. *J. Chem. Phys.* **1959***, 31*, 1500-1505.

(142) Lerner, R. G.; Dailey, B. P. Microwave spectrum and structure of propionitrile. *J. Chem. Phys.* **1957***, 26*, 678-680.

(143) Fukuyama, Y.; Omori, K.; Odashima, H.; Takagi, K.; Tsunekawa, S. Analysis of rotational transitions in excited vibrational states of propionitrile (C_2H_5CN). *J. Mol. Spectrosc.* **1999***, 193*, 72-103.

(144) Burie, J.; Demaison, J.; Dubrulle, A.; Boucher, D. Microwave spectrum of propionitrile: Determination of the quartic and sextic centrifugal distortion constants. *J. Mol. Spectrosc.* **1978***, 72*, 275-281.

(145) Demaison, J.; Margulès, L.; Mäder, H.; Sheng, M.; Rudolph, H. D. Torsional barrier and equilibrium structure of ethyl cyanide. *J. Mol. Spectrosc.* **2008***, 252*, 169-175.

(146) Li, Y. S.; Durig, J. R. Nitrogen-14 quadrupole coupling constants and barrier to internal rotation of ethylcyanide-d_5. *J. Mol. Spectrosc.* **1975***, 54*, 296-302.

(147) Podobedov, V. B.; Plusquellic, D. F.; Siegrist, K. E.; Fraser, G. T.; Ma, Q.; Tipping, R. H. New measurements of the water vapor continuum in the region from 0.3 to 2.7 THz. *Journal of Quantitative Spectroscopy and Radiative Transfer* **2008***, 109*, 458-467.

(148) Emery, R. Atmospheric absorption measurements in the region of 1 mm wavelength. *Infrared Physics* **1972***, 12*, 65-79.

(149) De Lucia, F. C.; Helminger, P.; Cook, R. L.; Gordy, W. Submillimeter microwave spectrum of $H_2{}^{16}O$. *Phys. Rev. , A* **1972***, 5*, 487-490.

(150) DeLucia, F. C.; Helminger, P.; Kirchhoff, W. H. Microwave spectra of molecules of astrophysical interest V. Water vapor. *Journal of Physical and Chemical Reference Data* **1974**, *3*, 211-219.

(151) Nabiev, S.; Vaks, V.; Domracheva, E.; Palkina, L.; Pripolzin, S.; Sobakinskaya, E.; Chernyaeva, M. Express analysis of water isotopomers in the atmosphere with the use of nonstationary subterahertz and terahertz spectroscopy methods. *Atmospheric and Oceanic Optics* **2011**, *24*, 402-410.

(152) Partridge, R. H. Far-infrared absorption spectra of $H_2{}^{16}O$, $H_2{}^{17}O$, and $H_2{}^{18}O$. *J. Mol. Spectrosc.* **1981**, *87*, 429-437.

(153) Stephenson, D. A.; Strauch, R. G. Water vapor spectrum near 600 GHz. *J. Mol. Spectrosc.* **1970**, *35*, 494-495.

(154) Belov, S. P.; Kozin, I. N.; Polyansky, O. L.; Tret'yakov, M. Y.; Zobov, N. F. Rotational spectrum of the $H_2{}^{16}O$ molecule in the (010) excited vibrational state. *J. Mol. Spectrosc.* **1987**, *126*, 113-117.

(155) Podobedov, V. B.; Plusquellic, D. F.; Fraser, G. T. THz laser study of self-pressure and temperature broadening and shifts of water vapor lines for pressures up to 1.4 kPa. *Journal of Quantitative Spectroscopy and Radiative Transfer* **2004**, *87*, 377-385.

(156) Markov, V. N.; Krupnov, A. F. Measurements of the pressure shift of the 110-101 water line at 556 GHz produced by mixtures of gases. *J. Mol. Spectrosc.* **1995**, *172*, 211.

(157) Burroughs, W. J.; Harries, J. E.; Gebbie, H. A. Submillimetre absorption by isotopic and vibrationally excited forms of water vapour. *Nature* **1969**, *222*, 658-660.

(158) Winther, F. The rotational spectrum of water between 650 and 50 cm^{-1}: $H_2{}^{18}O$ and $H_2{}^{17}O$ in natural abundance. *J. Mol. Spectrosc.* **1977**, *65*, 405-419.

(159) Bohlander, R. A.; Gebbie, H. A.; Pardoe, G. W. F. Absorption spectrum of water vapour in the region of 23 cm^{-1} at low temperatures. *Nature* **1970**, *228*, 156-157.

(160) Podobedov, V. B.; Plusquellic, D. F.; Fraser, G. T. Investigation of the water-vapor continuum in the THz region using a multipass cell. *Journal of Quantitative Spectroscopy and Radiative Transfer* **2005**, *91*, 287-295.

(161) Yasmin, K.; Armstrong, R. L. Theoretical modeling of microwave absorption by water vapor. *Appl. Opt.* **1990**, *29*, 1979-1983.

(162) Davis, G. R. The far infrared continuum absorption of water vapour. *Journal of Quantitative Spectroscopy and Radiative Transfer* **1993***, 50*, 673-694.

(163) Thomas, M. E.; Nordstrom, R. J. The N_2-broadened water vapor absorption line shape and infrared continuum absorption—II. Implementation of the line shape. *Journal of Quantitative Spectroscopy and Radiative Transfer* **1982***, 28*, 103-112.

(164) Matton, S.; Rohart, F.; Bocquet, R.; Mouret, G.; Bigourd, D.; Cuisset, A.; Hindle, F. Terahertz spectroscopy applied to the measurement of strengths and self-broadening coefficients for high-J lines of OCS. *J. Mol. Spectrosc.* **2006***, 239*, 182.

(165) Baum, O.; Esser, S.; Gierse, N.; Brünken, S.; Lewen, F.; Hahn, J.; Gauss, J.; Schlemmer, S.; Giesen, T. F. Gas-phase detection of HSOD and empirical equilibrium structure of oxadisulfane. *J. Mol. Struct.* **2006***, 795*, 256-262.

(166) Almoayed, N. N.; Piyade, B. C.; Afsar, M. N. In *High-resolution absorption coefficient and refractive index spectra of common pollutant gases at millimeter and THz wavelengths;* Mehdi Anwar, Anthony J. DeMaria and Michael S. Shur, Eds.; SPIE: 2007; Vol. 6772, pp 67720G.

(167) Hübers, H. -W.; Pavlov, S. G.; Richter, H.; Semenov, A. D.; Mahler, L.; Tredicucci, A.; Beere, H. E.; Ritchie, D. A. High-resolution gas phase spectroscopy with a distributed feedback terahertz quantum cascade laser. *Appl. Phys. Lett.* **2006***, 89*, 061115.

(168) Medvedev, I. R.; Neese, C. F.; Plummer, G. M.; De Lucia, F. C. Impact of atmospheric clutter on Doppler-limited gas sensors in the submillimeter/terahertz. *Appl. Opt.* **2011***, 50*, 3028-3042.

(169) Hindle, F.; Yang, C.; Mouret, G.; Cuisset, A.; Bocquet, R.; Lampin, J.; Blary, K.; Peytavit, E.; Akalin, T.; Ducournau, G. Recent developments of an opto-electronic THz spectrometer for high-resolution spectroscopy. *Sensors* **2009***, 9*, 9039-9057.

(170) Müller, H. S. P.; Drouin, B. J.; Pearson, J. C. Rotational spectra of isotopic species of methyl cyanide, CH_3CN, in their ground vibrational states up to terahertz frequencies. *Astronomy & Astrophysics* **2009***, 506*, 1487-1499.

(171) Lu, X.; Karpowicz, N.; Chen, Y.; Zhang, X. -. Systematic study of broadband terahertz gas sensor. *Appl. Phys. Lett.* **2008***, 93*, 261106.

(172) De Lucia, F. C.; Petkie, D. T. In *In THz gas sensing with submillimeter techniques (Invited Paper);* R. Jennifer Hwu, Dwight L. Woolard and Mark J. Rosker, Eds.; SPIE: 2005; Vol. 5790, pp 44-53.

(173) Hongqian Sun; Ding, Y. J.; Zotova, I. B. THz spectroscopy by frequency-tuning monochromatic THz source: from single species to gas mixtures. *Sensors Journal, IEEE* **2010***, 10*, 621-629.

(174) Johnson, K.; Price-Gallagher, M.; Mamer, O.; Lesimple, A.; Fletcher, C.; Chen, Y.; Lu, X.; Yamaguchi, M.; Zhang, X. -C. Water vapor: An extraordinary terahertz wave source under optical excitation. *arXiv. org, e-Print Arch. , Phys.* **2009**, 1-9, arXiv:0902.2024v1 [physics.optics].

(175) Guo, R.; Akiyama, K.; Minamide, H.; Ito, H. Frequency-agile terahertz-wave spectrometer for high-resolution gas sensing. *Appl. Phys. Lett.* **2007***, 90*, 121127-1.

(176) Matsuura, S.; Tani, M.; Abe, H.; Sakai, K.; Ozeki, H.; Saito, S. High-resolution terahertz spectroscopy by a compact radiation source based on photomixing with diode lasers in a photoconductive antenna. *J. Mol. Spectrosc.* **1998***, 187*, 97.

(177) Lin, H.; Withayachumnankul, W.; Fischer, B. M.; Mickan, S. P.; Abbott, D. Gas recognition with terahertz time-domain spectroscopy and spectral catalog: a preliminary study. *Proc. SPIE* **2008***, 6840*, 68400X/1-68400X/9.

(178) Uno, T.; Tabata, H. *In situ*measurement of combustion gas using terahertz time domain spectroscopy setup for gas phase spectroscopy and measurement of solid sample. *Japanese Journal of Applied Physics* **2010***, 49*, 04D17.

(179) Cheville, R. A.; Grischkowsky, D. Observation of pure rotational absorption spectra in the ν_2 band of hot H_2O in flames. *Opt. Lett.* **1998***, 23*, 531-533.

(180) Yang, Y.; Shutler, A.; Grischkowsky, D. Measurement of the transmission of the atmosphere from 0.2 to 2 THz. *Opt.Express* **2011***, 19*, 8830-8838.

(181) Hoshina, H.; Seta, T.; Iwamoto, T.; Hosako, I.; Otani, C.; Kasai, Y. Precise measurement of pressure broadening parameters for water vapor with a terahertz time-domain spectrometer. *Journal of Quantitative Spectroscopy and Radiative Transfer* **2008***, 109*, 2303-2314.

(182) Seta, T.; Hoshina, H.; Kasai, Y.; Hosako, I.; Otani, C.; Loßow, S.; Urban, J.; Ekström, M.; Eriksson, P.; Murtagh, D. Pressure broadening coefficients of the water vapor lines at 556.936 and 752.033 GHz. *Journal of Quantitative Spectroscopy and Radiative Transfer* **2008***, 109*, 144-150.

(183) Sagawa, H.; Mendrok, J.; Seta, T.; Hoshina, H.; Baron, P.; Suzuki, K.; Hosako, I.; Otani, C.; Hartogh, P.; Kasai, Y. Pressure broadening coefficients of induced by for Venus atmosphere. *Journal of Quantitative Spectroscopy and Radiative Transfer* **2009***, 110*, 2027-2036.

(184) Yu, B. L.; Yang, Y.; Zeng, F.; Xin, X.; Alfano, R. R. Terahertz absorption spectrum of D_2O vapor. *Opt. Commun.* **2006**, *258*, 256-263.

(185) Harde, H.; Katzenellenbogen, N.; Grischkowsky, D. Terahertz coherent transients from methyl chloride vapor. *J Opt Soc Am B* **1994**, *11*, 1018-1030.

(186) Harde, H.; Cheville, R. A.; Grischkowsky, D. Terahertz studies of collision-broadened rotational lines. *The Journal of Physical Chemistry A* **1997**, *101*, 3646-3660.

(187) Harde, H.; Zhao, J.; Wolff, M.; Cheville, R. A.; Grischkowsky, D. THz time-domain spectroscopy on ammonia. *The Journal of Physical Chemistry A* **2001**, *105*, 6038-6047.

(188) Foltynowicz, R. J.; Allman, R. E.; Zuckerman, E. Terahertz absorption measurement for gas-phase 2,4-dinitrotoluene from 0.05 THz to 2.7 THz. *Chemical Physics Letters* **2006**, *431*, 34-38.

(189) Hrubesh, L. W. Gas analysis by computer-controlled microwave rotational spectrometry. *Appl. Spectrosc.* **1978**, *32*, 425-429.

(190) Andresen, U.; Dreizler, H.; Kretschmer, U.; Stahl, W.; Thomsen, C. A molecular beam fourier transform microwave spectrometer developed for analytical purposes. *Fresenius J. Anal. Chem.* **1994**, *349*, 272-276.

(191) Leontakianakos, A. N. Determination of water vapor by microwave spectroscopy with application to quality control of natural gas. *Instrumentation and Measurement, IEEE Transactions on* **1992**, *41*, 370-374.

(192) Hughes, R. H.; Wilson, E. B. A microwave spectrograph. *Phys.Rev.* **1947**, *71*, 562-563.

(193) Hershberger, W. D. Minimum detectable absorption in microwave spectroscopy and an analysis of the stark modulation method. *J. Appl. Phys.* **1948**, *19*, 411-419.

(194) Rouleau, J. F.; Goyette, J.; Bose, T. K.; Frechette, M. F. Investigation of a microwave differential cavity resonator device for the measurement of humidity in gases. *Rev. Sci. Instrum.* **1999**, *70*, 3590-3594.

(195) Zhu, Z.; Matthews, I. P.; Samuel, A. H. Quantitative measurement of analyte gases in a microwave spectrometer using a dynamic sampling method. *Rev. Sci. Instrum.* **1996**, *67*, 2496-2501.

(196) Rezgui, N. D.; Allen, J.; Baker, J. G.; Alder, J. F. Quantitative millimetre wave spectrometry. Part I: Design and implementation of a tracked millimetre wave confocal Fabry-Perot cavity spectrometer for gas analysis. *Anal. Chim. Acta* **1995***, 311*, 99-108.

(197) Wilks, A. T.; Baker, J. G.; Alder, J. F.; Thompson, M. Quantitative millimetre wavelength spectrometry at pressures approaching atmospheric: Part I. Absolute determination of transition line strengths and sample concentrations with a confocal Fabry–Perot cavity spectrometer. *Anal. Chim. Acta* **2002***, 466*, 57-67.

(198) Wilks, A. T.; Thompson, M.; Alder, J. F.; Baker, J. G. Quantitative millimetre wavelength spectrometry at pressures approaching atmospheric: II. Determination of oxygen at atmospheric pressure. *Anal. Chim. Acta* **2002***, 468*, 323-333.

(199) Dumesh, B. S.; Surin, L. A. Two highly sensitive microwave cavity spectrometers. *Rev. Sci. Instrum.* **1996***, 67*, 3458-3464.

(200) Bigourd, D.; Cuisset, A.; Hindle, F.; Matton, S.; Bocquet, R.; Mouret, G.; Cazier, F.; Dewaele, D.; Nouali, H. Multiple component analysis of cigarette smoke using THz spectroscopy, comparison with standard chemical analytical methods. *Appl. Phy. B* **2007***, 86*, 579-586.

(201) Liu, H.; Chen, Y.; Yuan, T.; Al-Douseri, F.; Xu, J.; Zhang, X. In *Quantitative analysis of ammonia by THz time-domain spectroscopy;* James O. Jensen, Jean-Marc Theriault, Eds.; SPIE: 2004; Vol. 5268, pp 43-52.

(202) Edner, H.; Ragnarson, P.; Spännare, S.; Svanberg, S. Differential optical absorption spectroscopy (DOAS) system for urban atmospheric pollution monitoring. *Appl. Opt.* **1993***, 32*, 327-333.

(203) Winer, A.; Biermann, H. Long pathlength differential optical absorption spectroscopy (DOAS) measurements of gaseous HONO, NO_2 and HCNO in the California South Coast Air Basin. *Research on Chemical Intermediates* **1994***, 20*, 423-445.

(204) Parsons, M. T.; Sydoryk, I.; Lim, A.; McIntyre, T. J.; Tulip, J.; J\ager, W.; McDonald, K. Real-time monitoring of benzene, toluene, and p-xylene in a photoreaction chamber with a tunable mid-infrared laser and ultraviolet differential optical absorption spectroscopy. *Appl. Opt.* **2011***, 50*, A90-A99.

(205) Kuriyama, K.; Kaba, Y.; Yoshii, Y.; Miyazawa, S.; Manago, N.; Harada, I.; Kuze, H. Pulsed differential optical absorption spectroscopy applied to air pollution measurement in urban troposphere. *Journal of Quantitative Spectroscopy and Radiative Transfer* **2011***, 112*, 277-284.

(206) Jacobsen, R. H.; Mittleman, D. M.; Nuss, M. C. Chemical recognition of gases and gas mixtures with terahertz waves. *Opt. Lett.* **1996***, 21*, 2011-2013.

(207) Mittleman, D. M.; Jacobsen, R. H.; Neelamani, R.; Baraniuk, R. G.; Nuss, M. C. Gas sensing using terahertz time-domain spectroscopy. *Applied Physics B: Lasers & Optics* **1998***, 67*, 379.

(208) O'Shaughnessy, D. Linear predictive coding. *Potentials, IEEE* **1988***, 7*, 29-32.

(209) Karpowicz, N.; Zhong, H.; Xu, J.; Lin, K.; Hwang, J.; Zhang, X. -C. Comparison between pulsed terahertz time-domain imaging and continuous wave terahertz imaging. *Semicond. Sci. Technol.* **2005***, 20*, S293-S299.

(210) Leahy-Hoppa, M. R.; Fitch, M. J.; Osiander, R. Terahertz spectroscopy techniques for explosives detection. *Anal. Bioanal. Chem.* **2009***, 395*, 247-257.

(211) Hooper, J.; Mitchell, E.; Konek, C.; Wilkinson, J. Terahertz optical properties of the high explosive beta-HMX. *Chem. Phys. Lett.* **2009***, 467*, 309-312.

(212) Laman, N.; Harsha, S. S.; Grischkowsky, D.; Melinger, J. S. 7 GHz resolution waveguide THz spectroscopy of explosives related solids showing new features. *Opt. Express* **2008***, 16*, 4094-4105.

(213) Barber, J.; Hooks, D. E.; Funk, D. J.; Averitt, R. D.; Taylor, A. J.; Babikov, D. Temperature-dependent far-infrared spectra of single crystals of high explosives using terahertz time-domain spectroscopy. *J Phys Chem A* **2005***, 109*, 3501-3505.

(214) Funk, D. J.; Calgaro, F.; Averitt, R. D.; Asaki, M. L. T.; Taylor, A. J. THz transmission spectroscopy and imaging: application to the energetic materials PBX 9501 and PBX 9502. *Appl. Spectrosc.* **2004***, 58*, 428-431.

(215) Chen, J.; Chen, Y.; Zhao, H.; Bastiaans, G. J.; Zhang, X. -C. Absorption coefficients of selected explosives and related compounds in the range of 0.1-2.8 THz. *Opt. Express* **2007***, 15*, 12060-12067.

(216) Fischer, B.; Hoffmann, M.; Helm, H.; Modjesch, G.; Jepsen, P. U. Chemical recognition in terahertz time-domain spectroscopy and imaging. *Semicond. Sci. Technol.* **2005***, 20*, S246-S253.

(217) Zimdars, D.; Fichter, G.; Chernovsky, A. In *Rapid time domain terahertz axial computed tomography for aerospace non-destructive evaluation;* Infrared, Millimeter and Terahertz Waves, 2008. IRMMW-THz 2008. 33rd International Conference on; 2008; pp 1-3.

(218) Shen, Y. C.; Lo, T.; Taday, P. F.; Cole, B. E.; Tribe, W. R.; Kemp, M. C. Detection and identification of explosives using terahertz pulsed spectroscopic imaging. *Appl. Phys. Lett.* **2005**, *86*, 241116/1-241116/3.

(219) Zhang, L.; Karpowicz, N.; Zhang, C.; Zhao, Y.; Zhang, X. -C. Real-time nondestructive imaging with THz waves. *Opt. Commun.* **2008**, *281*, 1473-1475.

(220) Kogermann, K.; Zeitler, J. A.; Rantanen, J.; Rades, T.; Taday, P. F.; Pepper, M.; Heinamaki, J.; Strachan, C. J. Investigating dehydration from compacts using terahertz pulsed, Raman, and near-infrared spectroscopy. *Appl. Spectrosc.* **2007**, *61*, 1265-1274.

(221) Singhal, D.; Curatolo, W. Drug polymorphism and dosage form design: a practical perspective. *Adv. Drug Delivery Rev.* **2004**, *56*, 335-347.

(222) Strachan, C. J.; Taday, P. F.; Newnham, D. A.; Gordon, K. C.; Zeitler, J. A.; Pepper, M.; Rades, T. Using terahertz pulsed spectroscopy to quantify pharmaceutical polymorphism and crystallinity. *J. Pharm. Sci.* **2005**, *94*, 837-846.

(223) Upadhya, P.; Nguyen, K.; Shen, Y.; Obradovic, J.; Fukushige, K.; Griffiths, R.; Gladden, L.; Davies, A.; Linfield, E. Characterization of crystalline phase-transformations in theophylline by time-domain terahertz spectroscopy. *Spectrosc. Lett.* **2006**, *39*, 215-224.

(224) Shin, H. J.; Oh, S. J.; Kim, S. I.; Kim, H. W.; Son, J. Conformational characteristics of beta-glucan in laminarin probed by terahertz spectroscopy. *Appl. Phys. Lett.* **2009**, *94*, 111911/1-111911/3.

(225) Palermo, R.; Cogdill, R. P.; Short, S. M.; Drennen, J. K.; Taday, P. F. Density mapping and chemical component calibration development of four-component compacts via terahertz pulsed imaging. *J. Pharm. Biomed. Anal.* **2008**, *46*, 36-44.

(226) Spencer, J. A.; Gao, Z.; Moore, T.; Buhse, L. F.; Taday, P. F.; Newnham, D. A.; Shen, Y.; Portieri, A.; Husain, A. Delayed release tablet dissolution related to coating thickness by terahertz pulsed image mapping. *J. Pharm. Sci.* **2008**, *97*, 1543-1550.

(227) Maurer, L.; Leuenberger, H. Terahertz pulsed imaging and near infrared imaging to monitor the coating process of pharmaceutical tablets. *Int. J. Pharm.* **2009**, *370*, 8-16.

(228) Taday, P. F. Terahertz pulsed imaging for nondestructive testing of pharmaceutical products. *Spectroscopy (Duluth, MN, U. S.)* **2009**, *24*, 28, 30-36.

(229) Zeitler, J. A.; Shen, Y.; Baker, C.; Taday, P. F.; Pepper, M.; Rades, T. Analysis of coating structures and interfaces in solid oral dosage forms by three dimensional terahertz pulsed imaging. *J. Pharm. Sci.* **2006***, 96*, 330-340.

(230) Fitzgerald, A. J.; Cole, B. E.; Taday, P. F. Nondestructive analysis of tablet coating thicknesses using terahertz pulsed imaging. *J. Pharm. Sci.* **2005***, 94*, 177-183.

(231) Crawley, D. A.; Longbottom, C.; Cole, B. E.; Ciesla, C. M.; Arnone, D.; Wallace, V. P.; Pepper, M. Terahertz pulse imaging: a pilot study of potential applications in dentistry. *Caries Res.* **2003***, 37*, 352-9.

(232) Crawley, D.; Longbottom, C.; Wallace, V. P.; Cole, B.; Arnone, D.; Pepper, M. Three-dimensional terahertz pulse imaging of dental tissue. *J. Biomed. Opt.* **2003***, 8*, 303-7.

(233) Pickwell, E.; Wallace, V. P.; Cole, B. E.; Ali, S.; Longbottom, C.; Lynch, R. J. M.; Pepper, M. A comparison of terahertz pulsed imaging with transmission microradiography for depth measurement of enamel demineralisation in vitro. *Caries Res.* **2007***, 41*, 49-55.

(234) Nazarov, M. M.; Shkurinov, A. P.; Tuchin, V. V. Tooth study by terahertz time-domain spectroscopy. *Proc. SPIE* **2008***, 6791*, 679109/1-679109/9.

(235) Thrane, L.; Jacobsen, R. H.; Uhd, J., P.; Keiding, S. R. THz reflection spectroscopy of liquid water. *Chem. Phys. Lett.* **1995***, 240*, 330-3.

(236) Png, G. M.; Choi, J. W.; Ng, B. W.; Mickan, S. P.; Abbott, D.; Zhang, X. The impact of hydration changes in fresh bio-tissue on THz spectroscopic measurements. *Phys. Med. Biol.* **2008***, 53*, 3501-17.

(237) Pickwell, E.; Cole, B. E.; Fitzgerald, A. J.; Pepper, M.; Wallace, V. P. In vivo study of human skin using pulsed terahertz radiation. *Phys. Med. Biol.* **2004***, 49*, 1595-607.

(238) Han, P. Y.; Cho, G. C.; Zhang, X. C. Time-domain transillumination of biological tissues with terahertz pulses. *Opt. Lett.* **2000***, 25*, 242-4.

(239) Yoo, K. M.; Xing, Q.; Alfano, R. R. Imaging objects hidden in highly scattering media using femtosecond second-harmonic-generation cross-correlation time gating. *Opt. Lett.* **1991***, 16*, 1019-21.

(240) Hoshina, H.; Hayashi, A.; Miyoshi, N.; Miyamaru, F.; Otani, C. Terahertz pulsed imaging of frozen biological tissues. *Appl. Phys. Lett.* **2009***, 94*, 123901/1-123901/3.

(241) Woodward, R. M.; Cole, B. E.; Wallace, V. P.; Pye, R. J.; Arnone, D. D.; Linfield, E. H.; Pepper, M. Terahertz pulse imaging in reflection geometry of human skin cancer and skin tissue. *Phys. Med. Biol.* **2002***, 47*, 3853-63.

(242) The Skin Cancer Foundation Basal cell carcinoma. *http://www.skincancer.org/basal-cell-carcinoma/* (accessed 5/28, 2009).

(243) Woodward, R. M.; Wallace, V. P.; Arnone, D. D.; Linfield, E. H.; Pepper, M. Terahertz pulsed imaging of skin cancer in the time and frequency domain. *J. Biol. Phys.* **2003***, 29*, 257-259.

(244) Wallace, V. P.; Fitzgerald, A. J.; Shankar, S.; Flanagan, N.; Pye, R.; Cluff, J.; Arnone, D. D. Terahertz pulsed imaging of basal cell carcinoma ex vivo and in vivo. *Br. J. Dermatol.* **2004***, 151*, 424-32.

(245) Nakajima, S.; Hoshina, H.; Yamashita, M.; Otani, C.; Miyoshi, N. Terahertz imaging diagnostics of cancer tissues with a chemometrics technique. *Appl. Phys. Lett.* **2007***, 90*, 041102/1-041102/3.

(246) Fitzgerald, A. J.; Wallace, V. P.; Jimenez-Linan, M.; Bobrow, L.; Pye, R. J.; Purushotham, A. D.; Arnone, D. D. Terahertz pulsed imaging of human breast tumors. *Radiology* **2006***, 239*, 533-40.

(247) Yamamoto, K.; Yamaguchi, M.; Tani, M.; Hangyo, M.; Teramura, S.; Isu, T.; Tomita, N. Degradation diagnosis of ultrahigh-molecular weight polyethylene with terahertz-time-domain spectroscopy. *Appl. Phys. Lett.* **2004***, 85*, 5194-5196.

(248) Rutz, F.; Koch, M.; Khare, S.; Moneke, M.; Richter, H.; Ewert, U. Terahertz quality control of polymeric products. *Int. J. Infrared Millimeter Waves* **2006***, 27*, 547-556.

(249) Wietzke, S.; Jansen, C.; Rutz, F.; Mittleman, D. M.; Koch, M. Determination of additive content in polymeric compounds with terahertz time-domain spectroscopy. *Polym. Test.* **2007***, 26*, 614-618.

(250) Krumbholz, N.; Hochrein, T.; Vieweg, N.; Hasek, T.; Kretschmer, K.; Bastian, M.; Mikulics, M.; Koch, M. Monitoring polymeric compounding processes inline with THz time-domain spectroscopy. *Polym. Test.* **2009***, 28*, 30-35.

(251) Podzorov, A.; Gallot, G. Low-loss polymers for terahertz applications. *Appl. Opt.* **2008***, 47*, 3254-3257.

(252) Zhao, G.; ter, M., Maarten; Wenckebach, T.; Planken, P. C. M. Terahertz dielectric properties of polystyrene foam. *J Opt Soc Am B* **2002***, 19*, 1476-1479.

(253) Obradovic, J.; Collins, J. H. P.; Hirsch, O.; Mantle, M. D.; Johns, M. L.; Gladden, L. F. The use of THz time-domain reflection measurements to investigate solvent diffusion in polymers. *Polymer* **2007***, 48*, 3494-3503.

(254) Balakrishnan, J.; Fischer, B. M.; Abbott, D. Sensing the hygroscopicity of polymer and copolymer materials using terahertz time-domain spectroscopy. *Appl. Opt.* **2009**, *48*, 2262-2266.

(255) Jeon, T.; Grischkowsky, D.; Mukherjee, A. K.; Menon, R. Electrical and optical characterization of conducting polymers by THz time-domain spectroscopy. *Synth. Met.* **2003**, *135-136*, 451-452.

(256) Esenturk, O.; Melinger, J. S.; Heilweil, E. J. Terahertz mobility measurements on poly-3-hexylthiophene films. Device comparison, molecular weight, and film processing effects. *J. Appl. Phys.* **2008**, *103*, 023102/1-023102/8.

(257) Esenturk, O.; Kline, R. J.; Delongchamp, D. M.; Heilweil, E. J. Conjugation effects on carrier mobilities of polythiophenes probed by time-resolved terahertz spectroscopy. *J. Phys. Chem. C* **2008**, *112*, 10587-10590.

(258) Ostroverkhova, O.; Shcherbyna, S.; Cooke, D. G.; Egerton, R. F.; Hegmann, F. A.; Tykwinski, R. R.; Parkin, S. R.; Anthony, J. E. Optical and transient photoconductive properties of pentacene and functionalized pentacene thin films: Dependence on film morphology. *J. Appl. Phys.* **2005**, *98*, 033701/1-033701/12.

(259) Nahata, A.; Auston, D. H.; Wu, C.; Yardley, J. T. Generation of terahertz radiation from a poled polymer. *Appl. Phys. Lett.* **1995**, *67*, 1358-60.

(260) Zheng, X.; Sinyukov, A.; Hayden, L. M. Broadband and gap-free response of a terahertz system based on a poled polymer emitter-sensor pair. *Appl. Phys. Lett.* **2005**, *87*, 081115/1-081115/3.

(261) Sinyukov, A. M.; Hayden, L. M. Generation and detection of terahertz radiation with multilayered electro-optic polymer films. *Opt. Lett.* **2002**, *27*, 55-57.

(262) Hayden, L. M.; Sinyukov, A. M.; Leahy, M. R.; French, J.; Lindahl, P.; Herman, W. N.; Twieg, R. J.; He, M. New materials for optical rectification and electrooptic sampling of ultrashort pulses in the terahertz regime. *J. Polym. Sci. , Part B: Polym. Phys.* **2003**, *41*, 2492-2500.

(263) Laman, N.; Harsha, S. S.; Grischkowsky, D. Narrow-line waveguide terahertz time-domain spectroscopy of aspirin and aspirin precursors. *Appl. Spectrosc.* **2008**, *62*, 319-326.

(264) Zurk, L. M.; Orlowski, B.; Winebrenner, D. P.; Thorsos, E. I.; Leahy-Hoppa, M.; Hayden, M. R. Terahertz scattering from granular material. *J Opt Soc Am B* **2007**, *24*, 2238-2243.

(265) Leger, M. N.; Ryder, A. G. Comparison of derivative preprocessing and automated polynomial baseline correction method for classification and quantification of narcotics in solid mixtures. *Appl. Spectrosc.* **2006***, 60*, 182-193.

(266) Naftaly, M.; Miles, R. E. A method for removing etalon oscillations from THz time-domain spectra. *Opt. Commun.* **2007***, 280*, 291-295.

(267) Mansfield, C. D.; Rutt, H. N. Evaluation of spurious results in the infrared measurement of CO_2 isotope ratios due to spectral effects: a computer simulation study. *Phys. Med. Biol.* **1999***, 44*, 1155-67.

(268) Rungsawang, R.; Ueno, Y.; Ajito, K. Detecting a sodium chloride ion pair in ice using terahertz time-domain spectroscopy. *Anal. Sci.* **2007***, 23*, 917-920.

(269) Wollenhaupt, M.; Assion, A.; Baumert, T. Femtosecond laser pulses: Linear properties, manipulation, generation and measurement. **2007**, 937-983.

(270) Harris, D. C. In *Quantitative chemical analysis;* W. H. Freeman and Company: New York, NY, 2003; pp 744.

(271) Bacsik, Z.; Mink, J.; Keresztury, G. FTIR spectroscopy of the atmosphere. I. Principles and methods. *Applied Spectroscopy Reviews* **2004***, 39*, 295-363.

(272) Anonymous In *Biogenic trace gases: Measuring emissions from soil and water;* Matson, P. A., Harriss, R. C., Eds.; Blackwell: 1995; pp 394.

(273) Bangalore, A. S.; Demirgian, J. C.; Boparai, A. S.; Small, G. W. Effect of spectral resolution on pattern recognition analysis using passive Fourier transform infrared sensor data. *Appl. Spectrosc.* **1999***, 53*, 1382-1391.

(274) Olivieri, A. C. Analytical advantages of multivariate data processing. One, two, three, infinity? *Anal. Chem.* **2008***, 80*, 5713-5720.

(275) *Classical least-squares;* Chemometric Techniques for Quantitative Analysis; CRC Press: 1998; pp 51-70.

(276) *Partial least-squares;* Chemometric Techniques for Quantitative Analysis; CRC Press: 1998; pp 131-142.

(277) Vinzi, V. E.; Chin, W. W.; Henseler, J.; Wang, H., Eds.; In *Handbook of partial least squares;* Gentle, J. E., Hardle, W. K. and Mori, Y., Eds.; Springer Handbooks of Computational Statistics; Springer Verlag: New York, 2010; pp 798.

(278) Abdi, H. Partial least squares regression and projection on latent structure regression (PLS regression). *Wiley Interdisciplinary Reviews: Computational Statistics* **2010***, 2*, 97-106.

(279) Rosipal, R.; Krämer, N. Overview and recent advances in partial least squares. **2006***, 3940*, 34-51.

(280) Haaland, D. M.; Thomas, E. V. Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. *Anal. Chem.* **1988***, 60*, 1193-1202.

(281) Liu, L.; Arnold, M. A. Selectivity for glucose, glucose-6-phosphate, and pyruvate in ternary mixtures from the multivariate analysis of near-infrared spectra. *Anal Bioanal Chem* **2009***, 393*, 669-77.

(282) Ren, M.; Arnold, M. A. Comparison of multivariate calibration models for glucose, urea, and lactate from near-infrared and Raman spectra. *Anal Bioanal Chem* **2007***, 387*, 879-88.

(283) Lee, Y.; Chung, H.; Arnold, M. A. Improving the robustness of a partial least squares (PLS) model based on pure component selectivity analysis and range optimization: Case study for the analysis of an etching solution containing hydrogen peroxide. *Anal. Chim. Acta* **2006***, 572*, 93-101.

(284) Xiang, D. Advances in near-infrared glucose monitoring using pure component selectivity analysis for model characterization and a novel digital micromirror array spectrometer. 2006.

(285) Arnold, M. A.; Small, G. W.; Xiang, D.; Qui, J.; Murhammer, D. W. Pure component selectivity analysis of multivariate calibration models from near-infrared spectra. *Anal. Chem.* **2004***, 76*, 2583-2590.

(286) Ferré, J.; Faber, N. (. M. Net analyte signal calculation for multivariate calibration. *Chemometrics Intellig. Lab. Syst.* **2003***, 69*, 123-136.

(287) Lorber, A. Error propagation and figures of merit for quantification by solving matrix equations. *Anal. Chem.* **1986***, 58*, 1167-1172.

(288) Boqué, R.; Rius, F. X. Multivariate detection limits estimators. *Chemometrics Intellig. Lab. Syst.* **1996***, 32*, 11-23.

(289) Betty, K. R.; Horlick, G. A simple and versatile Fourier domain digital filter. *Appl. Spectrosc.* **1976***, 30*, 23-7.

(290) Horlick, G. Digital data handling of spectra utilizing Fourier transformations. *Anal. Chem.* **1972***, 44*, 943-7.

(291) Nelder, J. A.; Mead, R. A simplex method for function minimization. *The Computer Journal* **1965***, 7*, 308-313.

(292) Lagarias, J. C.; Reeds, J. A.; Wright, M. H.; Wright, P. E. Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM Journal of Optimization* **1998***, 9*, 112-147.

(293) Oral, A.; Anlas, G.; Lambros, J. Determination of Gurson-Tvergaard-Needleman model parameters for failure of a polymeric material. *Int J Damage Mech* **2012***, 21*, 3-25.

(294) Yildiz, Z.; Sener, E.; Yuceer, M. Optimization of electromagnetic filtration variables using a Nelder-Mead simplex method. *J. Dispersion Sci. Technol.* **2011***, 32*, 1230-1234.

(295) Tse, C. H.; Tang, M.; Shum, P. P.; Wu, R. F. Nelder-Mead simplex method for modeling of cascaded continuous-wave multiple-stokes Raman fiber lasers. *Opt. Eng. (Bellingham, WA, U. S.)* **2010***, 49*, 091009/1-091009/6.

(296) Sacco, W. F.; Alves, F., Hermes; Henderson, N.; de, O., Cassiano R. E. A Metropolis algorithm combined with Nelder-Mead Simplex applied to nuclear reactor core design. *Ann. Nucl. Energy* **2008***, 35*, 861-867.

(297) Cernivec, G.; Krc, J.; Smole, F.; Topic, M. Band-gap engineering in CIGS solar cells using Nelder-Mead simplex optimization algorithm. *Thin Solid Films* **2006***, 511-512*, 60-65.

(298) Larsen, L. M. An evaluation of the use of the Nelder Mead modified simplex method to find the region of optimal dose therapy combinations subject to toxicity constraints in a phase I/II clinical trial, 1997.

(299) Carlson, G. M. Kinetic analysis of consecutive reactions using Nelder-Mead simplex minimization. *Polym. Mater. Sci. Eng.* **1985***, 52*, 555-9.

(300) Gill, P. E.; Murray, W. Algorithms for the solution of the nonlinear least-squares problem. *SIAM Journal on Numerical Analysis* **1978***, 15*, 977-992.

(301) Hartley, H. O. The modified Gauss-Newton method for the fitting of non-linear regression functions by least squares. *Technometrics* **1961***, 3*, pp. 269-280.

(302) Marquardt, D. W. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics* **1963***, 11*, 431-441.

(303) Levenberg, K. A method for the solution of certain non-linear problems in least squares. *Quart. Appl. Math.* **1944***, 2*, 164-168.

(304) Madsen, K.; Nielsen, H. B.; Tingleff, O. Methods for non-linear least squares problems (2^{nd} ed.). **2004**, 60.

(305) Moré, J. The Levenberg-Marquardt algorithm: Implementation and theory. **1978**, *630*, 105-116.

(306) Townes, C. H.; Schawlow, A. L. In *Microwave spectroscopy;* McGraw-Hill Book Co: 1955; pp 698.

(307) Sugden, T. M.; Kenney, C. N. In *Microwave spectroscopy of gases;* Van Nostrand: 1966; pp 332.

(308) Wollrab, J. E. In *Rotational spectra and molecular structure physical chemistry. Series of monographs;* Academic: 1967; Vol. 13, pp 468.

(309) Kroto, H. W. In *Molecular rotation spectra;* Wiley-Interscience: 1974; pp 324.

(310) Eisenberg, D.; Crothers, D. In *Physical chemistry, with applications to the life sciences;* Benjamin/Cummings Publishing Co: 1979; pp 868.

(311) Walker, N. R. New opportunities and emerging themes of research in microwave spectroscopy. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences}* **2007**, *365*, 2813-2828.

(312) De Lucia, F. C. The submillimeter: A spectroscopist's view. *J. Mol. Spectrosc.* **2010**, *261*, 1-17.

(313) Watson, J. K. G. Forbidden rotational spectra of polyatomic molecules. *J. Mol. Spectrosc.* **1971**, *40*, 536-44.

(314) Erlandsson, G. Microwave spectrum and molecular structure of fluorobenzene. *Ark. Fys.* **1953**, *7*, 189-92.

(315) Erlandsson, G. Microwave spectrum of fluorobenzene. *Ark. Fys.* **1953**, *6*, 477-8.

(316) Strandberg, M. W. P. Centrifugal distortion. *Ann. N. Y. Acad. Sci.* **1952**, *55*, 808-813.

(317) Wacker, P. F.; Mizushima, M.; Petersen, J. D.; Ballard, J. R. Microwave spectral tables. Diatomic molecules. *Natl. Bur. Stand. Circ. (U. S.)* **1964**, *70*, 144.

(318) Wacker, P. F.; Pratto, M. R. In *Microwave spectral tables. Vol. 2;* U.S. Govt. Printing Office: 1964; , pp 338.

(319) Cord, M. S.; Lojko, M. S.; Petersen, J. D. In *Microwave spectral tables, vol. 5: Spectral line listing (National Bureau of Standards, monograph 70, vol. 5);* U.S. Govt. Printing Office: 1968; , pp 533.

(320) Cord, M. S.; Petersen, J. D.; Lojko, M. S.; Haas, R. H. Microwave spectral tables. IV. Polyatomic molecules without internal rotation. *Nat. Bur. Stand. (US), Monogr.* **1968***, No. 70*, 418.

(321) Wacker, P. F.; Cord, M. S.; Burkhard, D. G.; Petersen, J. D.; Kukol, R. F. Microwave spectral tables. III. Polyatomic molecules with internal rotation. *Nat. Bur. Stand. (US), Monogr.* **1969**, 265.

(322) Geschwind, S. High-resolution microwave spectroscopy. *Ann. N. Y. Acad. Sci.* **1952***, 55*, 751-773.

(323) Parsons, R. W.; Roberts, J. A. The Doppler contribution to microwave line widths. *J. Mol. Spectrosc.* **1965***, 18*, 412-417.

(324) Lorentz, H. A. The absorption and emission lines of gaseous bodies. *Proc. K. Ned. Akad. Wet.* **1906***, 8*, 591-611.

(325) Debye, P. In *Polar molecules;* Chem. Catalog Co: , pp 172.

(326) Van Vleck, J. H.; Weisskopf, V. F. On the shape of collision-broadened lines. *Rev.Mod.Phys.* **1945***, 17*, 227-236.

(327) Smith, W. V.; Howard, R. Microwave collision diameters II. Theory and correlation with molecular quadrupole moments. *Phys.Rev.* **1950***, 79*, 132-136.

(328) Birnbaum, G. Microwave pressure broadening and its application to intermolecular forces. *Advan. Chem. Phys.* **1967***, 12*, 487-548.

(329) Tsao, C. J.; Curnutte, B. Line-widths of pressure-broadened spectral lines. *Journal of Quantitative Spectroscopy and Radiative Transfer* **1962***, 2*, 41.

(330) Howard, R. R.; Smith, W. V. Microwave collision diameters I. Experimental. *Phys.Rev.* **1950***, 79*, 128-131.

(331) Brimblecombe, P. In *Air composition and chemistry: Second edition;* Cambridge Univ Press: 1996; pp 253.

(332) Occupational Safety and Health Administration Safety and health topics | Permissible exposure limits (PELs). *http://www.osha.gov/SLTC/pel/* (accessed 3/23, 2012).

(333) Occupational Safety and Health Administration Chemical sampling information - Acetaldehyde. *http://www.osha.gov/dts/chemicalsampling/data/CH_216300.html* (accessed 3/23, 2012).

(334) Occupational Safety and Health Administration Chemical sampling information - acetonitrile. *http://www.osha.gov/dts/chemicalsampling/data/CH_216700.html* (accessed 3/23, 2012).

(335) Occupational Safety and Health Administration Safety and health topics: Ammonia. *http://www.osha.gov/dts/chemicalsampling/data/CH_218300.html* (accessed 3/23, 2012).

(336) Occupational Safety and Health Administration Chemical sampling information: Ethyl alcohol. *http://www.osha.gov/dts/chemicalsampling/data/CH_239700.html* (accessed 3/23, 2012).

(337) Occupational Safety and Health Administration Chemical sampling information - Methyl alcohol. *http://www.osha.gov/dts/chemicalsampling/data/CH_251600.html* (accessed 3/23, 2012).

(338) Payan, S.; de La Noë, J.; Hauchecorne, A.; Camy-Peyret, C. A review of remote sensing techniques and related spectroscopy problems. *Comptes Rendus Physique* **2005***, 6*, 825-835.

(339) Platt, U. Modern methods of the measurement of atmospheric trace gases. *Phys. Chem. Chem. Phys.* **1999***, 1*, 5409-5415.

(340) Jones, C. Chemistry in the atmosphere. *Chem. Rev. (Deddington, U. K.)* **2010***, 19*, 16-18.

(341) Fiddler, M. N.; Begashaw, I.; Mickens, M. A.; Collingwood, M. S.; Assefa, Z.; Bililign, S. Laser spectroscopy for atmospheric and environmental sensing. *Sensors* **2009***, 9*, 10447-10512.

(342) Xu, H. L.; Chin, S. L. Femtosecond laser filamentation for atmospheric sensing. *Sensors* **2010***, 11*, 32-53.

(343) Farmer, D. K.; Jimenez, J. L. Real-time atmospheric chemistry field instrumentation. *Anal. Chem. (Washington, DC, U. S.)* **2010***, 82*, 7879-7884.

(344) Wilson, D. M.; Hoyt, S.; Janata, J.; Booksh, K.; Obando, L. Chemical sensors for portable, handheld field instruments. *Sensors Journal, IEEE* **2001***, 1*, 256-274.

(345) *Chromatography: Concepts and contrasts, second edition;* Miller, J. M., Ed.; John Wiley & Sons: 2004; , pp No pp. given.

(346) Ball, S. M.; Jones, R. L. Broad-band cavity ring-down spectroscopy. *Chem. Rev.* **2003***, 103*, 5239-5262.

(347) Canagaratna, M. R.; Jayne, J. T.; Jimenez, J. L.; Allan, J. D.; Alfarra, M. R.; Zhang, Q.; Onasch, T. B.; Drewnick, F.; Coe, H.; Middlebrook, A.; Delia, A.; Williams, L. R.; Trimborn, A. M.; Northway, M. J.; DeCarlo, P. F.; Kolb, C. E.; Davidovits, P.; Worsnop, D. R. Chemical and microphysical characterization of ambient aerosols with the aerodyne aerosol mass spectrometer. *Mass Spectrom. Rev.* **2007***, 26*, 185-222.

(348) National Aeronautics and Space Administration JPL molecular spectroscopy. *http://spec.jpl.nasa.gov/* (accessed 4/3, 2012).

(349) Pickett, H. M.; Poynter, R. L.; Cohen, E. A.; Delitsky, M. L.; Pearson, J. C.; Muller, H. S. Submillimeter, millimeter, and microwave spectral line catalog. *J. Quant. Spectrosc. Radiat. Transfer* **1998***, 60*, 883-890.

(350) Mueller, H. S. P.; Schloeder, F.; Stutzki, J.; Winnewisser, G. The Cologne Database for Molecular Spectroscopy, CDMS: a useful tool for astronomers and spectroscopists. *J. Mol. Struct.* **2005***, 742*, 215-227.

(351) Rothman, L. S.; Gordon, I. E.; Barbe, A.; Benner, D. C.; Bernath, P. F.; Birk, M.; Boudon, V.; Brown, L. R.; Campargue, A.; Champion, J. -.; Chance, K.; Coudert, L. H.; Dana, V.; Devi, V. M.; Fally, S.; Flaud, J. -.; Gamache, R. R.; Goldman, A.; Jacquemart, D.; Kleiner, I.; Lacome, N.; Lafferty, W. J.; Mandin, J. -.; Massie, S. T.; Mikhailenko, S. N.; Miller, C. E.; Moazzen-Ahmadi, N.; Naumenko, O. V.; Nikitin, A. V.; Orphal, J.; Perevalov, V. I.; Perrin, A.; Predoi-Cross, A.; Rinsland, C. P.; Rotger, M.; Simeckova, M.; Smith, M. A. H.; Sung, K.; Tashkun, S. A.; Tennyson, J.; Toth, R. A.; Vandaele, A. C.; Vander, A., J. The HITRAN 2008 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transfer* **2009***, 110*, 533-572.

(352) Pacific Northwest National Laboratory NorthWest Infrared. *https://secure2.pnl.gov/nsd/nsd.nsf/Welcome* (accessed 4/3, 2012).

(353) Pandey, P. C.; Srivastava, S. L. Collision broadening of acetaldehyde lines. II. Broadened by quadrupolar molecules nitrogen and hydrogen. *J. Chem. Phys.* **1973***, 58*, 1630-5.

(354) Fabian, M.; Morino, I.; Yamada, K. M. T. Analysis of the line profiles of CH_3CN for the J = 5 ← 4 and J = 6← 5 rotational transitions. *J. Mol. Spectrosc.* **1998***, 190*, 232-239.

(355) Curtiss, L. A.; Blander, M. Thermodynamic properties of gas-phase hydrogen-bonded complexes. *Chem. Rev.* **1988***, 88*, 827-41.

(356) Braun, C.; Leidecker, H. Rotation and vibration spectra for the H_2O dimer: Theory and comparison with experimental data. *J. Chem. Phys.* **1974***, 61*, 3104-3113.

(357) Lovas, F. J.; Hartwig, H. The Microwave Spectrum of the Methanol Dimer for K = 0 and 1 States. *J. Mol. Spectrosc.* **1997***, 185*, 98-109.

(358) Stockman, P. A.; Bumgarner, R. E.; Suzuki, S.; Blake, G. A. Microwave and tunable far-infrared laser spectroscopy of the ammonia–water dimer. *J. Chem. Phys.* **1992***, 96*, 2496.

(359) Havenith, M.; Cohen, R. C.; Busarow, K. L.; Gwo, D.; Lee, Y. T.; Saykally, R. J. Measurement of the intermolecular vibration–rotation tunneling spectrum of the ammonia dimer by tunable far infrared laser spectroscopy. *J. Chem. Phys.* **1991***, 94*, 4776.

(360) Karyakin, E. N.; Fraser, G. T.; Loeser, J. G.; Saykally, R. J. Rotation-tunneling spectrum of the deuterated ammonia dimer. *J. Chem. Phys.* **1999***, 110*, 9555.

(361) Stockman, P. A.; Blake, G. A.; Lovas, F. J.; Suenram, R. D. Microwave rotation-tunneling spectroscopy of the water-methanol dimer: Direct structural proof.. *J. Chem. Phys.* **1997***, 107*, 3782.

(362) Pine, A. S.; Howard, B. J. Hydrogen bond energies of the HF and HCl dimers from absolute infrared intensities. *J. Chem. Phys.* **1986***, 84*, 590-596.

(363) Legon, A. C.; Millen, D. J. Gas-phase spectroscopy and the properties of hydrogen-bonded dimers. HCN···HF as the spectroscopic prototype. *Chem. Rev.* **1986***, 86*, 635-657.

(364) Eerkens, J. W. Equilibrium dimer concentrations in gases and gas mixtures. *Chem. Phys.* **2001***, 269*, 189-241.

(365) Scherger, B.; Wietzke, S.; Scheller, M.; Vieweg, N.; Wichmann, M.; Koch, M.; Wiesauer, K. Characterization of micro-powders for the fabrication of compression molded THz lenses. *J. Infrared, Millimeter, Terahertz Waves* **2011***, 32*, 943-951.

(366) Sengupta, A.; Bandyopadhyay, A.; Bowden, B. F.; Harrington, J. A.; Federici, J. F. Characterisation of olefin copolymers using terahertz spectroscopy. *Electronics Letters* **2006***, 42*, 1477-1479.

(367) Bershtein, V. A.; Ryzhov, V. A. Relationship between molecular characteristics of polymers and parameters of far-infrared spectra. *J. Macromol. Sci. , Phys.* **1984***, B23*, 271-89.

(368) Wietzke, S.; Reuter, M.; Nestle, N.; Klimov, E.; Zadok, U.; Fischer, B. M.; Koch, M. Analyzing morphology and thermal history of poly(butylene terephthalate) by THz time-domain spectroscopy. *J. Infrared, Millimeter, Terahertz Waves* **2011***, 32*, 952-959.

(369) Wietzke, S.; Jansen, C.; Jung, T.; Reuter, M.; Baudrit, B.; Bastian, M.; Chatterjee, S.; Koch, M. Terahertz time-domain spectroscopy as a tool to monitor the glass transition in polymers. *Opt.Express* **2009***, 17*, 19006-19014.

(370) Wietzke, S.; Jansen, C.; Reuter, M.; Jung, T.; Kraft, D.; Chatterjee, S.; Fischer, B. M.; Koch, M. Terahertz spectroscopy on polymers: A review of morphological studies. *J. Mol. Struct.* **2011***, 1006*, 41-51.

(371) Johnson, K. W.; Rabolt, J. F. Far infrared measurements on the 19 °C phase transition in poly (tetrafluoroethylene). *J. Chem. Phys.* **1973***, 58*, 4536-8.

(372) Wietzke, S.; Jansen, C.; Reuter, M.; Jung, T.; Hehl, J.; Kraft, D.; Chatterjee, S.; Greiner, A.; Koch, M. Thermomorphological study of the terahertz lattice modes in polyvinylidene fluoride and high-density polyethylene. *Appl. Phys. Lett.* **2010***, 97*, 022901.

(373) Cowie, J. M. G.; Arrighi, V. Polymers, chemistry and physics of modern materials. *Cellul. Chem. Technol.* **2008***, 42*, 140-141.

(374) Suzuki, H.; Grebowicz, J.; Wunderlich, B. Glass transition of poly(oxymethylene). *British Polymer Journal* **1985***, 17*, 1-3.

(375) Zerbi, G.; Masetti, G. Lattice vibrations of orthorhombic polyoxymethylene. *J. Mol. Spectrosc.* **1967***, 22*, 284-289.

(376) Sperati, C.; Starkweather, H. Fluorine-containing polymers. II. Polytetrafluoroethylene. **1961***, 2*, 465-495.

(377) Beecroft, R. I.; Swenson, C. A. Behavior of polytetrafluoroethylene (Teflon) under high pressures. *J. Appl. Phys.* **1959***, 30*, 1793-1798.

(378) Clements, B. E. Low-pressure equation of state of polymers. *AIP Conference Proceedings* **2006***, 845*, 57-60.

(379) Wu, C.; Nicol, M. Raman spectra of high pressure phase and phase transition of polytetrafluoroethylene (teflon). *Chemical Physics Letters* **1973***, 21*, 153-157.

(380) Chantry, G. W.; Fleming, J. W.; Nicol, E. A.; Willis, H. A.; Cudby, M. E. A.; Boerio, F. J. The far infra-red spectrum of crystalline polytetrafluoroethylene. *Polymer* **1974***, 15*, 69.

(381) Chantry, G. W.; Nicol, E. A.; Jones, R. G.; Willis, H. A.; Cudby, M. E. A. On the vibrational assignment problem for polytetrafluoroethylene: 1. The far infra-red spectrum. *Polymer* **1977***, 18*, 37.

(382) Piseri, L.; Powell, B. M.; Dolling, G. Lattice dynamics of poly(tetrafluoroethylene). *J. Chem. Phys.* **1973**, *58*, 158-71.

(383) Boerio, F. J.; Koenig, J. L. Refinement of a valence force field for polytetrafluoroethylene by a damped least-squares method. *J. Chem. Phys.* **1970**, *52*, 4826-9.

(384) E. I. du Pont de Nemours and Company. DuPont Delrin acetal resin molding guide. *http://www2.dupont.com/Plastics/en_US/assets/downloads/processing/DELRIN_Mold_Guide_11_06.pdf*.

(385) Choy, C. L.; Nakafuku, C. Thermal expansion of polyoxymethylene. *Journal of Polymer Science Part B: Polymer Physics* **1988**, *26*, 921-934.

(386) Amrhein, E. M.; Heil, H. Dielectric absorption of polymers from the millimeter to the far infrared region. *Journal of Physics and Chemistry of Solids* **1971**, *32*, 1925-1933.

(387) Forrest, J. A.; Dalnoki-Veress, K.; Stevens, J. R.; Dutcher, J. R. Effect of free surfaces on the glass transition temperature of thin polymer films. *Phys. Rev. Lett.* **1996**, *77*, 2002-2005.

(388) Brown, E. N.; Rae, P. J.; Orler, E. B.; III, G. T. G.; Dattelbaum, D. M. The effect of crystallinity on the fracture of polytetrafluoroethylene (PTFE). *Materials Science and Engineering: C* **2006**, *26*, 1338.

(389) Roeder, L.; Pereira, P. N. R.; Yamamoto, T.; Ilie, N.; Armstrong, S.; Ferracane, J. Spotlight on bond strength testing-Unraveling the complexities. *Dental Materials* **2011**, *27*, 1197.

(390) Armstrong, S.; Geraldeli, S.; Maia, R.; Raposo, L. H. A.; Soares, C. J.; Yamagawa, J. Adhesion to tooth structure: A critical review of "micro" bond strength test methods. *Dental Materials* **2010**, *26*, e50.

(391) Nazarov, M. M.; Shkurinov, A. P.; Tuchin, V. V. Tooth study by terahertz time-domain spectroscopy. *Proc. SPIE* **2008**, *6791*, 679109/1.