ALGORITHM DEVELOPMENT FOR COLUMN WATER VAPOR RETRIEVAL USING THE SUN AND AUREOLE MEASUREMENT (SAM) SENSOR

by

Joshua B. Williams

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Electrical Engineering

Approved:

Dr. Doran Baker Major Professor	Dr. Jacob Gunther Committee Member
Dr. Kay Baker Committee Member	Dr. Byron R. Burnham Dean of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2008

Abstract

Algorithm Development for Column Water Vapor Retrieval Using the Sun and Aureole Measurement (SAM) Sensor

by

Joshua B. Williams, Master of Science

Utah State University, 2008

Major Professor: Dr. Doran Baker
Department: Electrical and Computer Engineering

To understand and model the energetics of the Sun-Earth connection, measurements of specific atmospheric molecules are beneficial. The objective is to formulate an algorithm to derive temporally varying atmospheric water vapor concentrations as functions of altitude, latitude, and longitude from solar irradiance absorption measurements. The Visidyne SAM (Sun and Aureole Measurement) instrument, which studies the size and distribution of cloud particles, was used to obtain the experimental data. By introducing a spectrometer to the SAM instrument, column water vapor is produced as part of the data product. A new model optimized algorithm is developed and tested versus existing algorithms. Through a least-squares analysis, the new algorithm showed an improvement of a factor of 23 over the industry standard. A test was also conducted to determine which water absorption bandpass produces the smallest error. Through these tests a model optimized algorithm has been produced.

(78 pages)

Acknowledgments

I would like to thank Dr. Doran Baker for the many hours dedicated to helping me finish this project. I would like to thank Rocky Mountain Space Grant and Visidyne for their financial support of this project. I would also like to thank Dr. A. T. Stair, Jr. and Dr. John DeVore for mentioning this project. Last, but not least, I would like to thank my family for helping me get to where I am today. Without their upbringing I would not have been able to accomplish this important work.

Joshua B. Williams

Contents

List of Tables

List of Figures

Chapter 1

Overview

Finding new ways to study and characterize the Earth's atmosphere is being given increased emphasis due to climate change issues. To understand and model the energetics of the Sun-Earth connection, additional measurements of specific atmospheric molecules are required. The scope of this thesis is the needed algorithmic development regarding temporal atmospheric water vapor measurements as functions of latitude and longitude.

1.1 SAM Description

SAM (Sun and Aureole Measurement) is a new ground-based instrument designed and produced by Visidyne in Burlington, MA. Figure 1.1 shows the setup of a SAM sensor. SAM is mounted to a solar-tracking mount which accurately provides sensor aspect to within a tenth of a degree. It is designed for mobile field experiments as well as designated observation locations. It has been deployed to numerous locations in the United States and is planned to be used in a global network of sensors.

SAM is designed to measure the optical depth and forward scattering of clouds. These can be measured by observing the aureole of the Sun through the atmosphere. The aureole is the illuminated region surrounding the solar disk. To measure the aureole of the solar disk, SAM is mounted to a solar tracker. SAM has a 8^o field of view which is centered on the Sun. The direct sunlight from the solar disk is inputed to a blackened cavity. The radiance surrounding the solar disk is reflected into a camera enables the measurement of the aureole of the Sun out to 8^o. Radiance profiles are produced from the aureoles, which are angular slices of the image of the aureole. Figure 1.2 [1] is an illustration of a radiance profile. The radiance profiles are analyzed to ascertain the size distribution and scattering of the cloud particles. SAM sensors help provide ground-truth data to support the validation of satellite

cloud algorithms.

Along with contributing to the study of cloud particles and distributions, the radiance profile is also a measure of aerosols. Radiance profiles provide information concerning the angular Mie and Rayleigh scattering of sunlight due to aerosols and other particles. A radiance profile is a portrayal of the angular portion of this scattering. Measurements of angular scattering combined with the measurements of the extinction of direct sunlight can provide the necessary data to more accurately study aerosol size and distribution. SAM is equipped with a spectrometer at the rear of the instrument to acquire the direct sunlight scattering and absorption. The incident sunlight is transmitted through a fiber optic cable from the rear of the blackened cavity. Figure 1.3 is an illustration of the coupling device used to transmit the sunlight to the spectrometer. This device is also seen in fig. 1.1 at the rear of the instrument.

Another byproduct which can be measured by the incident sunlight is column water vapor. By introducing a spectrometer to the SAM instrument, column water vapor can be included with the other data products produced. As a large contributor to climate change, there is an interest to understanding the dynamics of water vapor. To accurately model the climate change frequent measurements are required to determine the changing distribution of water vapor.

1.2 Water Vapor

Column water vapor is the measurement of the concentration of water in the atmosphere. It represents:

> The total amount of water vapor contained in a vertical column of unit cross section area, expressed as the depth (in cm of precipitable water) of the liquid equivalent of the vapor in the column if all that water vapor were condensed and collected [2].

These values average between 0.5 and 5.0 cm [3]. Water vapor absorbs at multiple wavelengths of light in the visible and near-infrared spectral region. These reactions occur at

Fig. 1.1: SAM instrument setup.

specific wavelengths, and by using these H_2O reactions an algorithm can be developed to measure column water vapor.

1.3 Objectives

This thesis addresses the development of a model optimized algorithm to accurately ascertain column water vapor. To accomplish this development the following requirements need to be met:

1. Characterization and calibration of the SAM (Sun and Aureole Measurement) spectrometer.
2. Research existing methods and techniques of column water retrieval.
3. Development of an model optimized algorithm for the SAM (Sun and Aureole Measurement) instrument.
4. Validation of the measurements by minimum mean squared error tests.
5. Formulate, test, and document a model optimized algorithm.

Through these steps, using experimental data, a new algorithm will be tested against existing algorithms, and these tests will demonstrate the magnitude of improvement obtained through using the new approach. Upon accomplishing this, the new advanced algorithm will be tested to discover under what conditions this algorithm produces the least error. Both of these tests will be conducted using a least-squares approach. By minimizing the error a model optimized algorithm under optimal conditions will be formulated. This will provide a reliable data set to ascertain the distribution and dynamics of column water vapor, and in turn further the understanding of climate change trends.

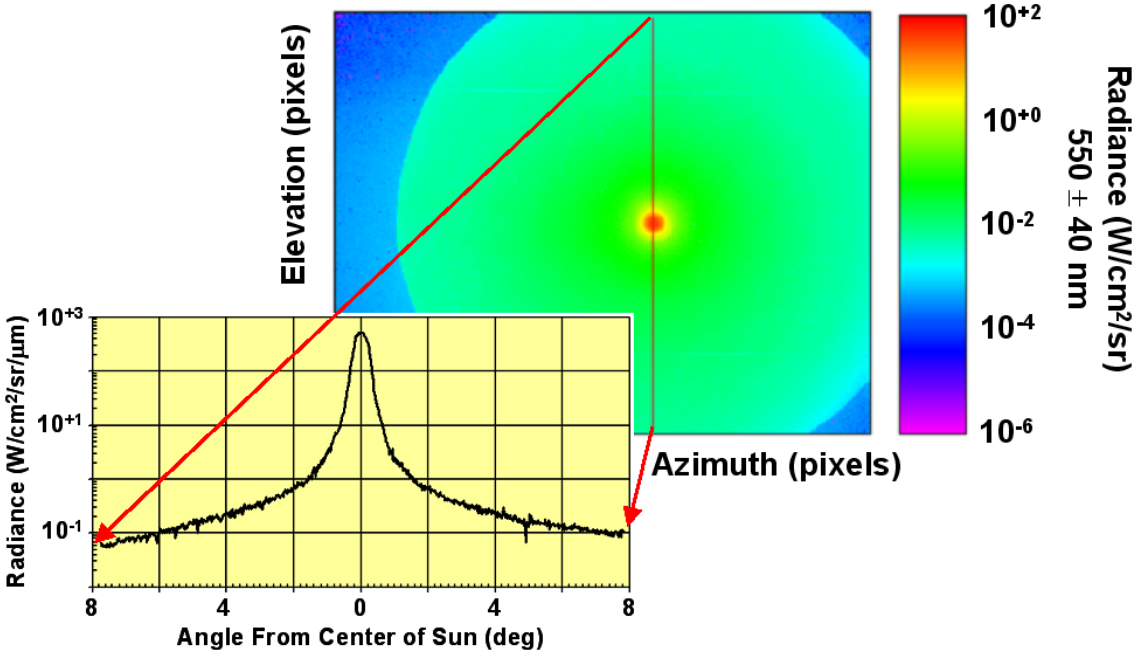


Fig. 1.2: Example of radiance profile.

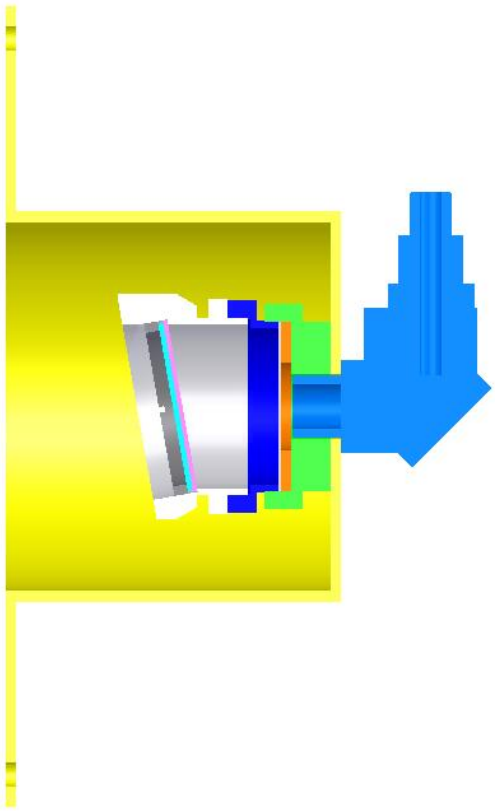


Fig. 1.3: Cutout of spectrometer.

Chapter 2

System Development

The Visidyne SAM instrumentation system is designed to measure the absorption of sunlight attributable to H_2O in the Earth's atmosphere. For SAM to measure the aureoles of the Sun, the instrument employs a sun tracker platform so that the direct sunlight is directed into a blackened cavity located at the rear of the instrument. This is called the beam dump. Located at the rear of the beam dump is a fiber-optic cable which transfers the direct sunlight into a spectrometer. This spectrometer was chosen for two types of scattering-absorption measurements, namely aerosols and water vapor. For the aerosol measurements the spectrometer needs high sensitivity in the visible range. For the water vapor a standard wavelength for the measurements is 940 nm. These factors are constraints upon the type of spectrometer to be employed.

2.1 Optical Spectrometer

To obtain the spectral absorption required for column water vapor retrieval, a grating spectrometer was acquired. The spectrometer selected was a Photon Control model number SPM-002-C with a silicon detector which has a high sensitivity in the visible range. The free spectral range of this spectrometer appeared to best fit the measurement needs in both the visible and the near infrared regions of the spectrum. This spectrometer uses a Toshiba TCD 1304 CCD linear image sensor [4]. The relative spectral response of this sensor is illustrated in fig. 2.1 [5].

It was known that at 940 nm the detector had a low sensitivity, but it was believed that this spectrometer would fit the requirements. If the requirements were not satisfied the selected spectrometer would only be used for testing purposes. This spectrometer was mounted onto the SAM mount and used in three field tests consisting of the following:

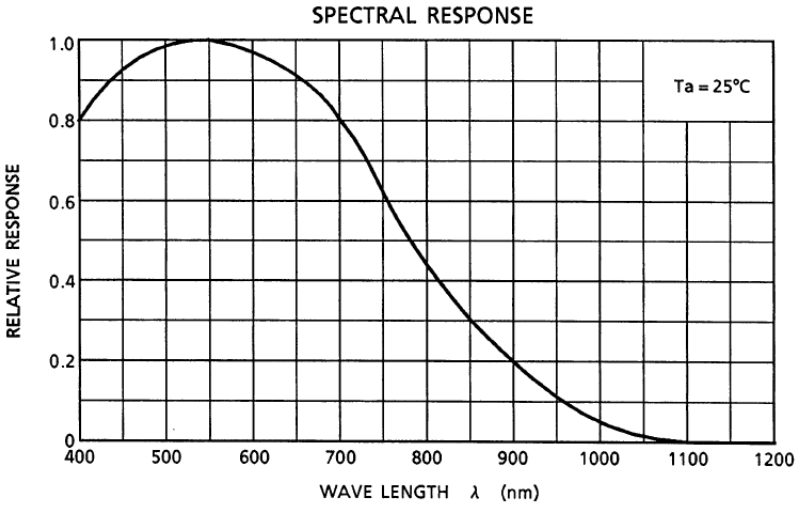


Fig. 2.1: Spectrometer sensitivity.

1. CHAPS, Cumulus Humilis Aerosol Processing Study [6], at the Atmospheric Radiation Measurement Southern Great Plains site located near Ponca City, Oklahoma, providing ground measurements during June, 2007. This was a joint collaboration of scientists who combined resources to study interactions between clouds and aerosols.

2. CATZ-C in Ridgely, Maryland, providing ground-based calibration exercises for the CALIPSO, Cloud-Aerosol Lidar and the Infrared Pathfinder Satellite Observation, satellite in July of 2007.

3. CATZ-E at NASA Goddard Space Flight Center(GSFC) and at Tuckahoe State Park, Maryland, providing ground-based calibration exercises for the CALIPSO satellite in August of 2007.

While the SAM sensor was in support for each of the above campaigns an AERONET, AErosol RObotic NETwork [7], sensor, the standard for water vapor measurements, was present providing a cross-calibration data set. During these campaigns, preliminary tests of these spectrometers were conducted which facilitated the formulation of column water vapor algorithms.

2.2 Preliminary Results

The preliminary approach to analyze data taken on these campaigns was to formulate a best fit for the SAM data set to that of AERONET. This was accomplished using an existing algorithm currently used by AERONET. In this algorithm there are two parameters, a and b, which approximate the water vapor absorption at 940 nm. Parameters a and b characterize water vapor absorption at specific wavelengths of light. These parameters remain constant for measurements a fixed measurement bandpass. The derivation of these parameters is found in sec. 4.3.2. By using AERONET as training data, parameters a and b were found using a least-squares comparison to create the best fit of the SAM data to the AERONET data during the CATZ-C studey. Using this method the results in figs. 2.2 and 2.3 demonstrated promising results.

At first glance the SAM values may appear to agree well with those seen with AERONET. However, on detailed examination these correlations are not sufficiently accurate. Once the parameters were calculated by using a least-squares fit, a separate data set was processed to test the accuracy and precision the parameters a and b. The second processed data set in comparison with the corresponding AERONET data set demonstrated unsatisfactory similarities. By using the same bandpasses as AERONET, parameters a and b should have been similar in value. This proved to be false. A least-squares analysis was then conducted on each of the SAM data sets which were taken next to an AERONET sensor. The parameters a and b demonstrated a large variation between each of the available data sets. Parameters a and b should remain constant for each specific measurement instrument. This test showed that additional considerations needed to be taken into account.

2.3 Problems / Observations

Sensor calibration is the process of determining both a senors accuracy and precision. It is the process of characterizing an instrument's output. The instrument characterization is used to stabilize measurements such that parameters a and b are constant values producing the desired response. The difficulty of this process depends on how external stimuli and instrument error affect the sensor's stability. A few of these problems which can affect a

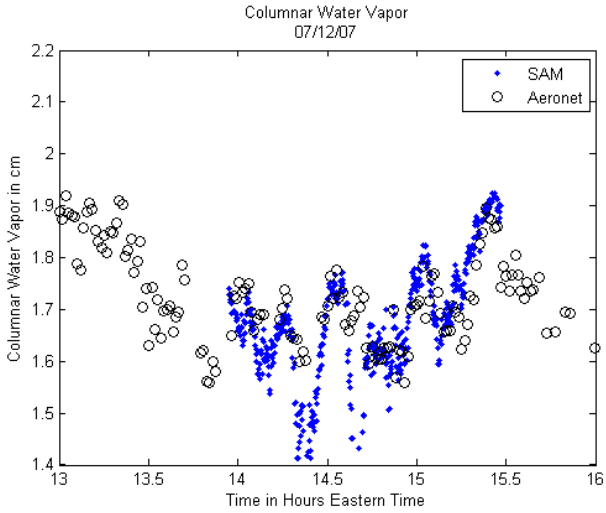


Fig. 2.2: CATZ-C column water vapor results.

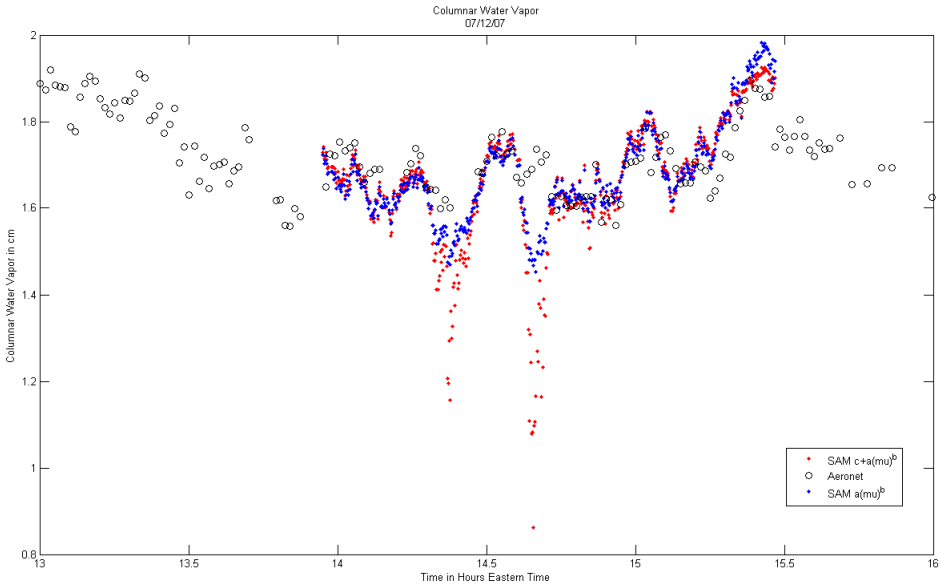


Fig. 2.3: CATZ-C column water vapor comparison.

sensors stability are:

1. responsivity,
2. temperature,
3. dark current.

Radiometers with solid-state sensors tend to be very stable and straight-forward to calibrate. Once this calibration procedure is conducted, that radiometer's measurements tend to remain precise and accurate. A spectrometer is affected by other stimuli. These stimuli can cause the spectrometer's measurements to become unstable. This instability affects the precision of the data. From the preliminary results of the field experiments, temperature change, responsivity, and dark current were found to cause variations in the data with the purchased spectrometer. Once discovered, methods of minimizing the effects of these problems on the accuracy and precision of the sensor will ensure an accurate data set.

The normalized responsivity of the spectrometer sensor is shown in fig. 2.1. Further examination of this figure at a wavelength of 940-nm shows that the relative response of the spectrometer used is approximately 0.1. As a result the signal-to-noise ratio (SNR) is relatively small. Unless the intensity of the incident light at 940 nm is high, the SNR ratio will not be sufficiently large to obtain a reliable data set. This problem can be minimized by either using a spectrometer whose photoelectric material has a higher relative response at 940 nm, such as indium gallium arsenide, or by using a different wavelength for which water vapor absorbs light.

The responsivity of the spectrometer detector was also found to change with temperature. Figure 2.1 shows the responsivity of the sensor at selected wavelengths, but is valid only for a specific temperature, namely 25^{o}C as indicated on the figure. The sensor temperature dependance is shown in fig. 2.4 [5]. This indicates that a small change in temperature leads to a significant measurement change. This temperature dependance of the spectrometer sensor creates significant issues if it is not constrained to a constant temperature. The

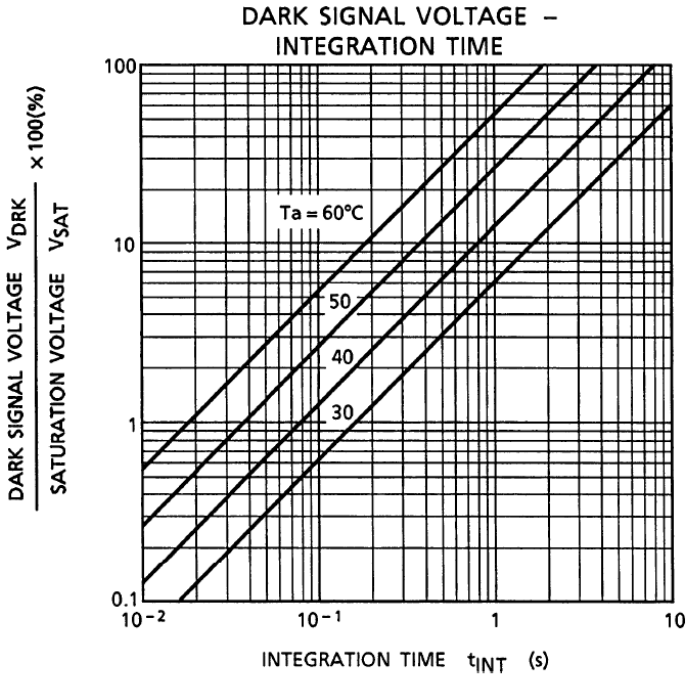


Fig. 2.4: Sensitivity (responsivity) temperature response of CCD array.

spectrometer used in these field experiments was not temperature stabilized, and therefore the produced data set exhibited temperature dependance. This is one of the major contributors to why, for each least-squares tests in which the parameters a and b were estimated, the SAM data sets did not statistically coincide with the AERONET data set.

In the absence of light, a small current propagates through photoelectric materials called dark current which is one the sources of noise in a photo-detection system. To minimize this additive noise, it can be measured and subtracted from the measurements taken which minimizes the effects of dark current. This is only valid if it can assumed that the dark current is constant at every wavelength measured due to temperature and time. This assumption is too broad. The magnitude of dark current increases with temperature and therefore can vary over time. To eliminate this problem the dark current would need to be measured before each measurement and subtracted from the output values. The spectrometer used in these field experiments did not have this capability.

These spectrometer calibration issues of the initial spectrometer were documented, and led to a search for a new spectrometer that minimized the temperature, responsivity, and dark current effects. Imposing these requirements a new spectrometer was selected that is temperature stabilized, automatically subtracts the dark current, and is more responsive in the near-infrared portion of the spectrum. This spectrometer is currently being tested and calibrated to produce more accurate, precise, and stable measurements.

These calibration issues motivated further study into the column water vapor retrieval algorithms, which, in turn, led to an in-depth study, reported in Chapters 3 and 4, which improved system performance.

Chapter 3

Algorithm Development

Research to determine more accurate methods to measure atmospheric column water vapor has been conducted over the last few decades [8]. Innovative studies have helped to significantly improve the validity of these measurements. These improvements have come both in the form of technology and in the algorithms used to ascertain the column water vapor. The more accurate these data sets are, the better the effects of the column water vapor upon the atmospheric transmittance and upon the climate trends can be modeled.

3.1 Existing Methods

3.1.1 Background

Column water vapor (CWV), otherwise known as precipitable water, has been studied using water absorption measurements at specific wavelengths. Water vapor absorption bands were discovered by Fowle in 1912-1915 using test tubes containing known amounts of water vapor. His findings are the foundation of methods of measuring CWV. Figure 3.1 [9] shows the atmospheric absorption of solar radiation by the Earth's atmosphere.

Einstein's research in the early 1900's concerning the quantum states of photon emissions from molecules explains this effect. He found that the energy of the incident photons are absorbed at specific wavelengths [10]. Figure 3.2 is a simplified energy level illustration of this absorption. Einstein and Fowle's findings help explain why specific wavelengths of light from the sun are absorbed by water vapor in the atmosphere before reaching ground level. Knowledge of the irradiance of the light outside the Earth's atmosphere in comparison with the irradiance at Earth's surface can therefore help to understand the concentration of water vapor in the atmosphere. This absorption is the basis of the column water re-

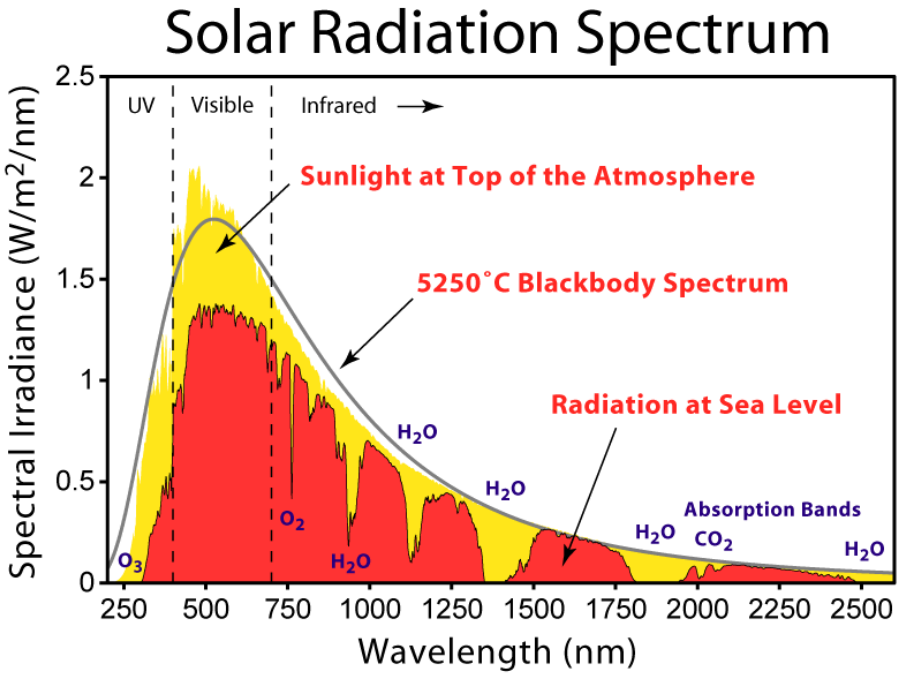


Fig. 3.1: Solar irradiance spectrum.

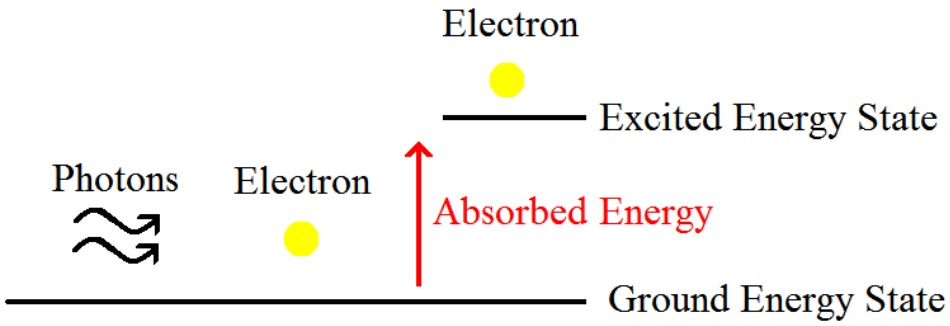
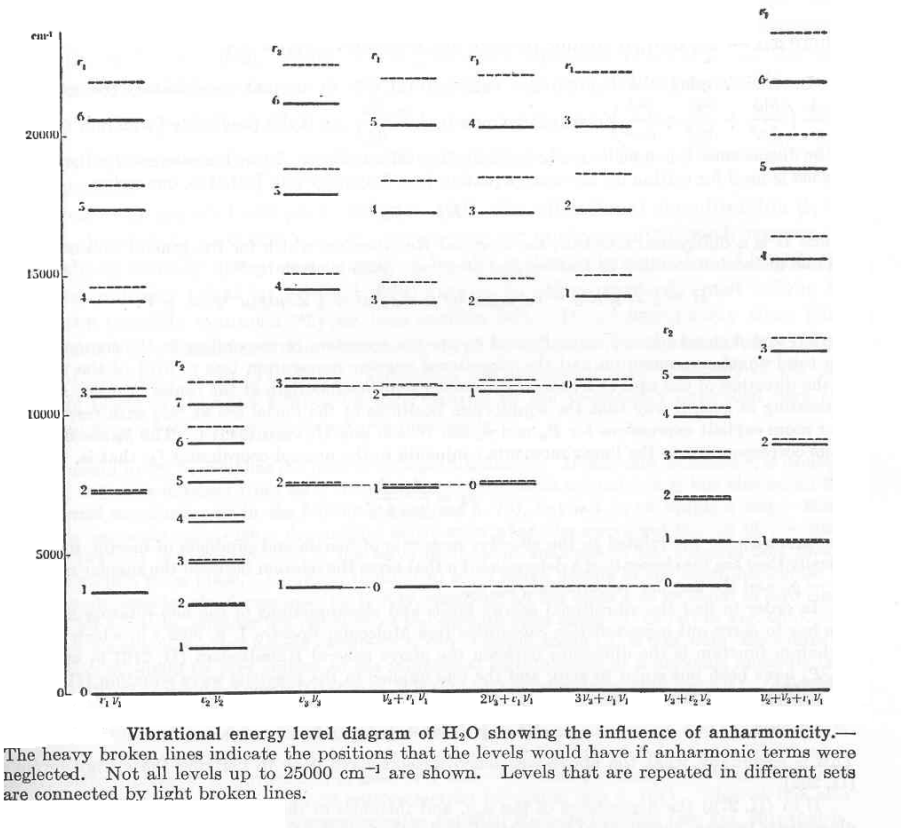


Fig. 3.2: Molecular excitation effect.

trieval method. Figure 3.1 demonstrates the irradiance of solar spectrum outside and inside Earth's atmosphere. Some of the most pronounced absorption bands attributable to the H_2O molecule are given in table 3.1 [11] and shown in fig. 3.4. Figure 3.3 [12] gives the energy transition levels of H_2O.



Vibrational energy level diagram of H_2O showing the influence of anharmonicity.—The heavy broken lines indicate the positions that the levels would have if anharmonic terms were neglected. Not all levels up to 25000 cm^{-1} are shown. Levels that are repeated in different sets are connected by light broken lines.

Fig. 3.3: Energy transition levels of H_2O.

Table 3.1: Water vapor absorption center wavelengths.

Wavelengths (nm)	**Wave Number**	**Spectrum**
514	19460	Visible
606	16500	Visible
660	15150	Visible
720	13530	Visible
820	11960	Infrared
940	10310	Infrared
1200	8330	Infrared
1465	6800	Infrared
1900	5260	Infrared

Research to develop an optical technique to measure atmospheric column water vapor began at the University of Arizona [13]. Algorithms were developed using two separate wavelength bandpasses at 820-nm and 940-nm. The center of one of these wavelength bandpasses was chosen such that it is located to cover a water absorption band, otherwise

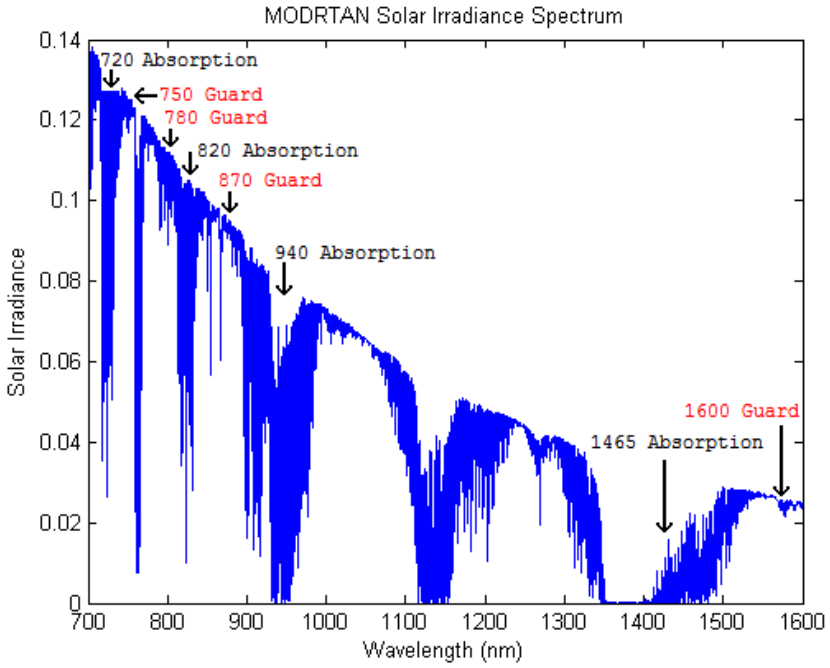


Fig. 3.4: Water absorption bands.

called as the water band. The center of the second bandpass is positioned close to, but not overlapping, the water band. This channel is known as the guard band. The derived algorithm [13] uses a ratio of these two bands to approximate the concentration of column water vapor in the atmosphere.

3.1.2 Symbol Identifications

The following symbols will be used throughout the derivation of the column water vapor algorithm:

CWV	Column water vapor
W	Center wavelength of water band
G	Center wavelength of guard band
$R(\lambda)$	Irradiance measured
$R(W)$	Irradiance at water band
$R(G)$	Irradiance at guard band
$R_O(\lambda)$	Solar spectral irradiance outside atmosphere
$R_O(W)$	Solar spectral irradiance outside atmosphere at water band
$R_O(G)$	Solar spectral irradiance outside atmosphere at guard band
O_R	Outside atmosphere constant $\frac{R_O(W)}{R_O(G)}$
$\tau_a(\lambda)$	Spectral optical depth for aerosol particles
$\tau_R(\lambda)$	Spectral optical depth for Rayleigh scattering
$\tau_3(\lambda)$	Spectral optical depth for O_3
τ_R	Combination of Rayleigh scattering $\tau_R(G) - \tau_R(W)$
d	Earth-to-Sun distance (in astronomical units)
u	Column water-vapor (cm)
m	Relative air-mass
a	Multiplicative parameter
b	Exponential parameter
c	Additive parameter

3.1.3 Previous Methods

The University of Arizona's column water vapor algorithm of the 1970's research has continued to be explored to formulate the optimal method to retrieve CWV. Research has continued since the conception of the algorithm to improve upon existing methods. Today, there are two different algorithms used to retrieve CWV. These are derived by using Beer's law [14] in conjunction with an empirical model of water vapor absorption. The two algorithms use two different empirical models to represent the absorption of water at specific absorption wavelengths. These empirical models are shown in eqs. 3.1 and 3.2.

$$\ln(R(\lambda)) + m(\tau_a(\lambda) + \tau_R(\lambda) + \tau_3(\lambda)) = \ln(R_0(\lambda)d^{-2}) + a(mu)^b \tag{3.1}$$

$$\ln(R(\lambda)) + m(\tau_a(\lambda) + \tau_R(\lambda) + \tau_3(\lambda)) = \ln(R_O(\lambda) * d^{-2}) + c + (mu)^b \tag{3.2}$$

Each represents a different method to model water vapor absorption using Beer's law [13]. The derivation for each the additive and multiplicative model are the same, therefore only the multiplicative model derivation will be shown.

3.1.4 Derivation

The representations defined by eqs. 3.1 and 3.2 must be compared with a guard band. A guard band is chosen near an appropriate spectral water band feature. This is done to ensure similarity between the two bands. A small wavelength separation allows the reflection of two spectral bandpasses to be approximately the same. The guard band is selected such that the relative difference between the two bandpasses is only the water vapor absorption. Equation 3.3 gives the representation of the guard band.

$$\ln(R(\lambda)) + m(\tau_a(\lambda) + \tau_R(\lambda) + \tau_3(\lambda)) = \ln(R_O(\lambda)d^{-2}) \tag{3.3}$$

Comparing eqs. 3.1 and 3.3 demonstrates the difference between the water and guard bands. With only a small wavelength separation, parameters $\tau_a(\lambda)$, $\tau_R(\lambda)$, and $\tau_3(\lambda)$ are nearly the same values for both the water and guard bands. This observation is the basis of the algorithm. In order to solve for the parameters a, b, and c, which represent the water vapor, eqs. 3.2 and 3.3 are combined to form eq. 3.4.

$$\ln(\frac{R(G)}{R(W)}) + m(\tau_a(G) + \tau_R(G) + \tau_3(G)) - \tau_a(W) - \tau_R(W) - \tau_3(W)) = \ln(\frac{R_O(G)}{R_O(W)}) + a(m * u)^b \tag{3.4}$$

The center wavelength selection for the guard band near that of the water band facilitates some key assumptions. When emissions at two similar wavelengths are compared, the different reflections and refractions due to $\tau_a(\lambda)$, $\tau_R(\lambda)$ and $\tau_3(\lambda)$ are the approximately equal in value. This assumption is implicit in the simplification of eq. 3.4 to yield eq. 3.5.

$$\ln(\frac{R(G)}{R(W)}) = \ln(\frac{R_O(G)}{R_O(W)}) + a(m * u)^b \tag{3.5}$$

The result of solving for the column water vapor u is eq. 3.6.

$$u = \frac{1}{m}(\frac{\ln(O_R * R(G)/R(W))}{a})^{\frac{1}{b}} \tag{3.6}$$

This equation is the representation for the multiplicative model given in eq. 3.1. By a similar derivation, eq. 3.7 results from the additive model given in eq. 3.2.

$$u = \frac{1}{m}(\ln(O_R * R(G)/R(W)) - c)^{\frac{1}{b}} \tag{3.7}$$

3.1.5 Current Systems

The two eqs. 3.6 and 3.7 are the models for deriving CWV from the measurements using various instrumentation configurations including radiometers and spectrometers. Important factors in the validity and accuracy of the CWV determination are the bandpasses of the measurement devices, water absorption wavelengths, and guard band wavelengths.

A system currently in use, AERONET [7], is a globally distributed geographic network of sensors and are considered to establish a standard for CWV. The measurement data are processed using eq. 3.6. In most photometric instruments, the exponent b is assumed to be $\frac{1}{2}$ which introduces a square-root dependence [15]. This assumption is based upon the absorption properties of water vapor. Parameter a is then computed by using a least squares analysis. See sec. 4.3 for more detail on this process and how instrumental characteristics are taken into account as part of the calibration of each instrument, each with slightly varying instrumental specifications.

To ensure that AERONET produces the most accurate measurements possible, one instrument was taken to mount Mauna Loa, Hawaii, to "fine tune" these parameters. After this reference instrument is calibrated, the others are adjusted to agree, thereby setting the calibration standard. This is indicative of how sensitive these instruments are to slight changes in the calculated constants [16].

The parameters a and b or b and c are the key to being able to retrieve an accurate CWV data set. They have been studied in detail and are unique to each particular photometric instrument [15].

The formulation used by AERONET is adopted as the standard, which is represented in eq. 3.6. Papers reporting the comparisons of these techniques and measurement devices such as MODIS and AERONET are available in the references [15]. The measurement approach and technique to model the water absorption is the focus of the present thesis. As will be seen later, the current approach reliability approximates the water vapor absorption, but through a combination of techniques and a further study into measurement instrumentation a new technique will provide more accurate results.

3.2 New Three-Parameter Algorithm

3.2.1 Conceptual Development

The approach to finding an optimal model and method of approximating CWV is finding an optimal representation of water vapor absorption using Beer's law. By using an optimal water vapor absorption model, the error will be minimized between the calculated and actual measurements. Equations 3.1 and 3.2 are two different representations of water absorption used in combination with Beer's law. Upon a review of these two equations, tradeoffs were found approximating the water vapor absorption. This eventually led to the development of a new three-parameter approach. With the combination of the two two-parameter systems, a third degree of freedom is introduced into the system. This third parameter enables the algorithm to better approximate the water absorption. The third parameter adds some complexity to the overall system, but each parameter has to be

calculated only once for each instrument. See sec. 4.4 for comparison of the two- versus three-parameter approaches.

3.2.2 Derivation of the New Three-Parameter Algorithm

The derivation of the three-parameter approach is similar to that for each of the two-parameter approaches. The differences include:

1. It does not depend upon the square-root dependence of parameter b.

2. Its not assumed that the Rayleigh scattering in the water and guard band are the same. Existing systems assume the Rayleigh scattering to be the same for the water and guard bands [13].

3. All three parameters a, b, and c are included to approximate water absorption. Equation 3.8 will be now used to represent the water band, and eq. 3.3 is used to represent the guard band.

$$\ln(R(\lambda)) + m(\tau_a(\lambda)) + \tau_R(\lambda) + \tau_3(\lambda) = \ln(R_O(\lambda) * d^{-2}) + c + a(mu)^b \tag{3.8}$$

The Rayleigh scattering is similar in the water and guard bands, but due to different wavelength locations of these bands the Rayleigh scattering values can be slightly different. At this step in the derivation it is assumed, as before, that the optical scattering due to O_3 and aerosols, namely $\tau_a(\lambda)$ and $\tau_3(\lambda)$, are the same for both bands. In future work, if these scattering effects are measured and not assumed to be approximately the same, a more accurate measure of CWV will result. This will be pursued in future versions of the algorithm. Using these assumptions, eq. 3.9 results from the combination of eqs. 3.3 and 3.8.

$$\ln(\frac{R(G)}{R(W)}) + m(\tau_R(G) - \tau_R(W)) = \ln(\frac{R_O(G)}{R_O(W)}) + c + a(m * u)^b \tag{3.9}$$

Through simplification eq. 3.9 becomes eq. 3.10.

$$\ln(\frac{O_R * R(G)}{R(W)}) + m(\tau_R(G) - \tau_R(W)) = c + a(m * u)^b \tag{3.10}$$

3.2.3 Final Column Water Vapor Equation

Equations 3.9 and 3.10 produced by instituting the new three-parameter approach are more complex than the current two-parameter approach, but through the extensive testing, reported in Chapter 4, a more optimal solution is formulated. Equation 3.11 gives the resulting CWV equation model for this approach.

$$u = \frac{1}{m}(\frac{\ln(O_R * R(G)/R(W)) + m * \tau_R - c}{a})^{\frac{1}{b}} \tag{3.11}$$

Chapter 4

Algorithm Validation and Comparison

After conceptual development of the algorithm was developed, the next step was testing and implementation. For both the two- and three-parameter approaches (eqs. 3.6, 3.7, and 3.11), a method was formulated to solve for each of the parameters (a, b, and c). Once formulated, a comparison was conducted to ascertain which is the most optimal algorithm. These tests showed under what circumstances which algorithm performs the best, i.e., at what wavelengths and spectral bandwidths to use.

4.1 Problem Overview

The accuracy of each CWV algorithm is, in part, determined by how the parameters a, b, and c are formulated to model the water vapor absorption. The process [15] is sensitive to parameter values, and these parameters exhibit a large CWV variation with even slight alterations in each method of retrieval. Consequently, a method needed to be formulated to calculate these constants for each different retrieval technique. Parameters such as optimal bandwidth, center wavelength, and characterization of measurement instruments were taken into account.

4.2 Comparison of Retrieval Techniques

In practice, most instruments used to determine CWV employ radiometric measurements [15]. A radiometer measures a fixed bandpass of the electromagnetic spectrum which is selected by a fixed optical filter. The most important radiometric characteristics for CWV are spectral bandpass and sensitivity because they characterize that portion of the spectrum to be measured. Optical bandwidths are usually selected to be 10 nm or larger, depending upon the detector selected. The radiometric power detected over the bandpass

limits of this filter constitutes the radiometer measurement. Appropriate materials and methods are used to fabricate detectors to measure the flux of incident light. Advantages of using solid-state radiometers include reduced stability calibration issues. A disadvantage is that the bandwidth is a fixed value.

An instrument device that can be used to measure the incident light as a function of wavelength is a spectrometer. A spectrometer provided light spectrum over a wide free spectral range, whereas a radiometer only measures irradiance within a fixed bandpass. For a spectrometer, important characteristics are throughput grating, responsivity, and resolving power. The spectral resolution determines how much separation there is between each sequential measurement of light. This will determine how precise measurements can be taken.

The responsivity indicates what range of the incident light intensity will be detected. The free spectral range gives the maximum and minimum wavelengths over which the instrument can provide a spectrum. These numbers will vary depending upon the selection of detector. For instance, a silicon detector is best over the visible range, but in the near-infrared region indium-gallium-arsenide has the best detectivity. An advantage of using spectrometers is that they have a much larger measurement spectral range in which measurements can be obtained more precisely within a given water-band. A disadvantage is that they are more complicated and may not be temperature stable.

4.3 Least-Squares and MMSE Derivation

Independent of the sensor used to measure the guard and water bands, a method needed to be chosen to determine the values of a, b, and c which result in the smallest error. This error is defined as the correct value, produced by the atmospheric model, minus the calculated value, produced by the algorithm. This error determines which parameters best fit the actual values of CWV. It is noted that there will always be an error unless a perfect model exists of the system. Since the model used to approximate the CWV is not perfect, under all realistic conditions an error is expected. The goal of the present analysis is to minimize this error for the existing model. To find this minimum, the commonly used

least-squares error analysis was utilized.

4.3.1 Derivation

In the formulation of the CWV algorithm there are the three parameters a, b, and c. A least-squares analysis could be conducted for each of the three parameters, but prior knowledge simplifies the process and helps constrain the results to stay within physical bounds. When conducting a least-squares analysis a standard training data set is used in the calculation. Since an infinite set of data is not possible, the least-squares analysis can merely approximate the parameters which will result in the smallest error. However the parameter values that determine this error can lie outside the set of real possibilities.

In this study fewer than a hundred points of training data were used. So few points could cause the parameter values which result in the smallest error to lie outside of their possible range. From the understanding of the optical model used to approximate the water vapor absorption, it is known that the exponent b must physically lie between zero and two. This prior knowledge helps ensure that the calculated parameters are constrained to the appropriate range.

To initiate the least-squares analysis, the error is squared at each data point and is defined in eq. 4.1.

$$Error^2 = [c + a(m * u)^b - (ln(\frac{O_R * R(G)}{R(W)}) + m * \tau_R)]^2 \tag{4.1}$$

This is a representation of the error squared at one instance of time. The squared error was used in this analysis because minimizing the error is the equivalent of minimizing the error squared. Accordingly, the mean squared error is subsequently referred to as "the error." To calculate the mean squared error (MSE), eq. 4.1 is summed over the number of measurements, and divided by the number of measurements. The MSE is represented by eq. 4.2.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}[c + a(m_i * u_i)^b - (ln(\frac{O_R * R(G)_i}{R(W)_i}) + m_i * \tau_R)]^2 \tag{4.2}$$

The MSE shown in eq. 4.2 is the difference between the value of the actual water vapor absorption and the estimated value. It is noted that this is a different value than the MSE of the CWV. By minimizing the absorption MSE, in turn, the CWV MSE is minimized. These values will be different, but represent the same error. The magnitude of MSE introduced by the estimated absorption is proportional to the CWV MSE. For the derivation of the minimum mean squared error (MMSE), absorption MSE is used as shown in eq. 4.2. After the derivation the two are distinguished by the titles absorption and CWV MSE.

To minimize the MSE given by eq. 4.2, derivatives are taken with respect to parameters a and c. As stated earlier, the range of b is defined to lie between zero and two. For this part of the analysis, it is assumed that the value of b is known. The solution for the correct value of b is given later in the analysis. Taking the derivatives of the MSE with respect to a and c yields eqs. 4.3 and 4.4.

$$\frac{\partial MSE}{\partial a} = \frac{2}{n}\sum_{i=1}^{n}([c + a(m_i * u_i)^b - (ln(\frac{O_R * R_i(G)}{R_i(W)}) + m_i * \tau_R)]^2 * (m_i * u_i)^b) \tag{4.3}$$

$$\frac{\partial MSE}{\partial c} = \frac{2}{n}\sum_{i=1}^{n}[c + a(m_i * u_i)^b - (ln(\frac{O_R * R_i(G)}{R_i(W)}) + m_i * \tau_R)]^2 \tag{4.4}$$

To determine the MMSE with respect to a and c, both eqs. 4.3 and 4.4 are set to zero. This results in two equations and two unknowns. Equation 4.5 shows a simplification of eq. 4.3, and eq. 4.6 shows the simplification of eq. 4.4.

$$c\sum_{i=1}^{n}(m_i * u_i)^b + a\sum_{i=1}^{n}(m_i * u_i)^{2b} = \sum_{i=1}^{n}([ln(\frac{O_R * R_i(G)}{R_i(W)}) + m_i * \tau_R] * (m_i * u_i)^b) \tag{4.5}$$

$$c\sum_{i=1}^{n}1 + a\sum_{i=1}^{n}(m_i * u_i)^b = \sum_{i=1}^{n}[\ln(\frac{O_R * R_i(G)}{R_i(W)}) + m_i * \tau_R] \tag{4.6}$$

At this point, with two equations and two unknowns, it is possible to solve for both a

and c. In order to solve this set of equations, they are placed into matrix form as given in eq. 4.7.

$$\begin{bmatrix} \sum_{i=1}^{n}(m_i * u_i)^b & \sum_{i=1}^{n}(m_i * u_i)^{2b} \\ \sum_{i=1}^{n} 1 & \sum_{i=1}^{n}(m_i * u_i)^b \end{bmatrix} * \begin{bmatrix} c \\ a \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n}[\ln(\frac{O_R * R_i(G)}{R_i(W)}) + m_i * \tau_R](m_i * u_i)^b \\ \sum_{i=1}^{n}[\ln(\frac{O_R * R_i(G)}{R_i(W)}) + m_i * \tau_R] \end{bmatrix} \tag{4.7}$$

In order to solve for the parameters a and c, eq. 4.7 can be written as eq. 4.8 in a symbolic form.

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} * \begin{bmatrix} c \\ a \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \tag{4.8}$$

This produces a symbolic representation for both a and c. The symbolic representation is the least-squares solution for these parameters. This solution can be found by taking the inverse of the matrix A on both sides of the equation. Equation 4.9 shows the end result for a solution to explicitly yield parameters a and c.

$$\begin{bmatrix} c \\ a \end{bmatrix} = \begin{bmatrix} \frac{A_{22} * B_1 - A_{12} * B_2}{A_{11} * A22 - A12 * A21} \\ \frac{A_{11} * B_2 - A_{21} * B_1}{A_{11} * A22 - A12 * A21} \end{bmatrix} \tag{4.9}$$

Equation 4.9 is the general solution to the three-parameter approach. Having obtained a least-squares solution for both a and c, we can return back to the earlier assumption of knowing the value of b. Understanding the nature of the exponential parameter in this function demonstrates acceptable ranges of values for b. The range in this case is limited to between zero and two. To find the best method of solving for this parameter, a training data set is needed contains the following:

1. Water band measured value,
2. Guard band measured value,
3. Corresponding CWV value.

Using this training data set, a comparison between the actual CWV and the calculated value can be produced. The method used in this paper is the MMSE. This is accomplished by setting b to a fixed value, solving for both a and c in eq. 4.9, and then calculating the MSE for all values of u given by the training data set. This was done for incremental values of b between zero and two. The MMSE's is taken as the value of b which best fits the data. This method also produces the corresponding values for a and c. This process is shown in eq. 4.10.

$$MMSE = \min_{b} \sum_{i=1}^{n} (u_i(calculated) - u_i(measured))^2, 0 < b <= 2 \tag{4.10}$$

All derivations shown in this section have been formulated for the three-parameter approach. This same method is used to solve for the each of the two-parameter approaches. However, there is a slight variation in which there is only one parameter for which the least squares analysis is conducted. Equations 4.11 and 4.12 are similar to eq. 4.7 for the two-parameter versions. Using the MMSE to solve for b is used in all three cases.

$$a \sum_{i=1}^{n} (m_i * u_i)^{2b} = \sum_{i=1}^{n} [\ln(\frac{O_R * R_i(G)}{R_i(W)}) + m_i * \tau_R](m_i * u_i)^b \tag{4.11}$$

$$c \sum_{i=1}^{n} 1 = \sum_{i=1}^{n} [\ln(\frac{O_R * R_i(G)}{R_i(W)}) + m_i * \tau_R - (m_i * u_i)^b] \tag{4.12}$$

4.3.2 MODTRAN

Training data are needed to solve for the parameters using the least squares analysis derived in sec. 4.3. To provide the best approximation for the parameters a, b, and c these training data need to cover various conditions. In the case of CWV these conditions can vary from location to location, but can also vary due to the time of the year at one specific location. Due to this variability of CWV, these parameters need to provide the best results over a wide range of conditions. Another requirement is that the data need to cover the corresponding light spectrum with an optical resolution in the nm. For a valid comparison between radiometric and spectrometric data, the training data spectrum

needs to be sampled at a high enough resolution such that either narrow bandpass or wide bandpass are available as needed.

With these requirements, a valid atmospheric model with sufficiently high resolution could be used to produce these training data. The MODTRAN, MODerate spectral resolution atmospheric TRANSmittance, algorithm fits these requirements. MODTRAN was developed by the U.S. Air Force Research Laboratory to model the atmospheric radiance and transmittance [17]. LOWTRAN, LOW resolution TRANsmission, model was the predecessor to MODTRAN which was first developed to model the atmosphere. MODTRAN was developed to produce this model in higher resolution than was the case for LOWTRAN. Subsequently, an atmospheric emission-absorption model was developed called HITRAN, HIgh-resolution TRANsmission. The optical spectral resolution varies by wavelength. The spectral resolution in the visible to infrared regions is approximately 0.08 nm which is sufficient for the measurement resolution in nm, and covers the corresponding appropriate spectral range. MODTRAN also produces data over various conditions as would be seen from data at different locations. MODTRAN fills the necessary requirements for these training data, and therefore was used for the comparisons which have been drawn. Upon completion of this algorithm using MODTRAN data, the near optimal type of instrument and spectral range can be chosen.

4.4 MMSE Comparison of the Two- and Three-Parameter Approaches

The MMSE test was formulated in sec. 4.3.1. There are many possible different values of instrument parameters which need to be considered such as bandwidth and wavelength. Each of these parameter values has its own MMSE. Using this MMSE formulation along with the MODTRAN training data, these equations can be used to ascertain under which conditions the MMSE is the smallest. Therefore, by a simple comparison of these MMSE's the model optimized algorithm can be selected.

Each of the two- and three-parameter approaches has been derived in detail. These derivations provide the means by which a comparison can be made to ascertain which provides the most accurate representation of water vapor absorption. This comparison

was conducted over multiple center wavelengths and multiple bandwidths. It was done to show the validity of the results and to assess the magnitude of improvement of one approach over the other. When AERONET was formulated for an absorption technique the differences between the different two-parameter models were considered. It was found that the multiplicative model produced the least MSE, and therefore this model was used [15]. For this reason, only the multiplicative version of the two-parameter approach will be compared with the three-parameter approach herein.

The analysis was made by using the MODTRAN data, as discussed in sec. 4.3.2, and the MMSE derivation in sec. 4.3.1. To show the difference between the calculated and the actual CWV, a line with a slope of unity was plotted over the full range of values of CWV. The values of measured CWV were then plotted versus the actual values of CWV. The farther each point varies from the line, the larger the magnitude of the MSE.

To show the proper comparison with the technique used by AERONET, figs. 4.1 and 4.2 illustrate the magnitude of the improvement of the three-parameter approach. Figures 4.1-4.4 demonstrate which of the approaches produce the smallest MSE with respect to CWV. In each of these figures four separate plots are given to help with understanding the comparison. The top left panel contains the MMSE solution to the three-parameter approach. The top right panel contains a magnified view is presented to further demonstrate the magnitude of the MSE. The bottom left panel contains the MMSE solution to the two-parameter approach, and once again the bottom right panel gives a magnified view of the two-parameter approach.

These results are summarized in table 4.1. The magnitude of improvement from implementing the three-parameter approach is a function of wavelength; however, a significant improvement was found at each wavelength. Only a portion of the figures are shown in this section, but table 4.1 includes the results over a wide range of wavelengths. See Appendix A for a complete set of figures which show the comparisons of the two- versus three-parameter approaches.

Table 4.1 demonstrates the magnitude of improvement for each of the considered water

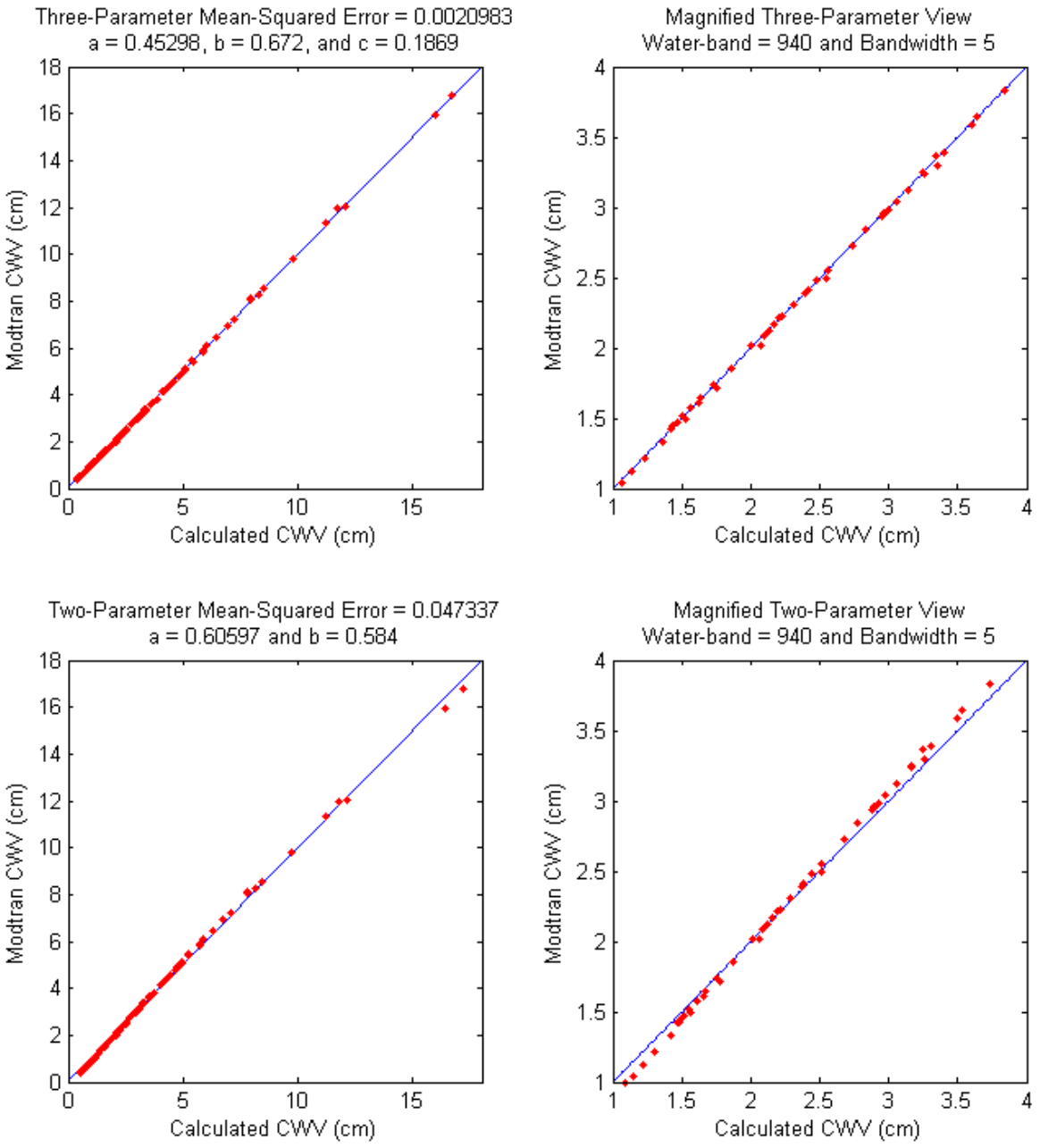


Fig. 4.1: Two- versus three-parameter comparison at 940 nm with 5-nm bandwidth. Magnified view on right.

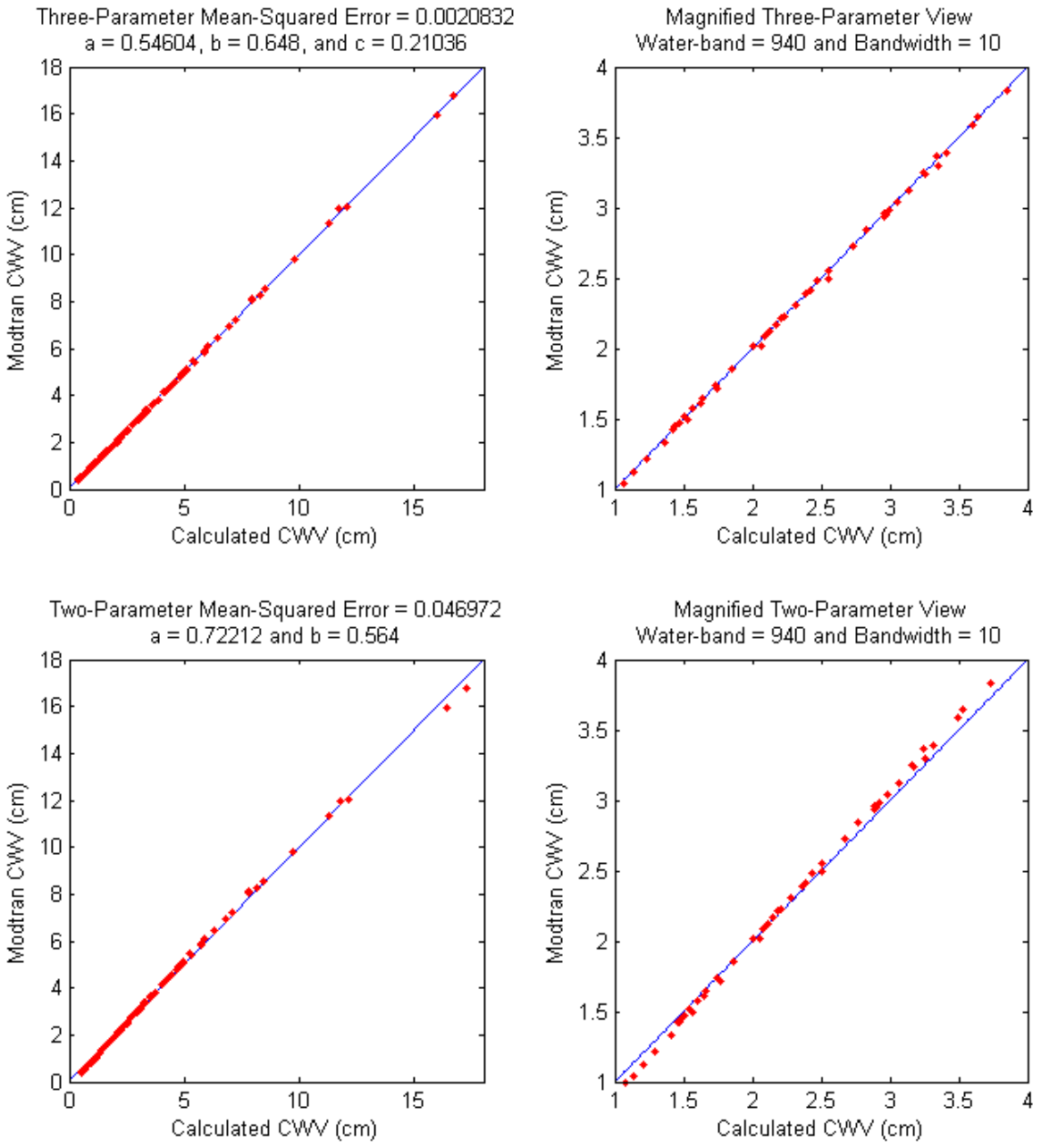


Fig. 4.2: Two- versus three-parameter comparison at 940 nm with 10-nm bandwidth. Magnified view on right.

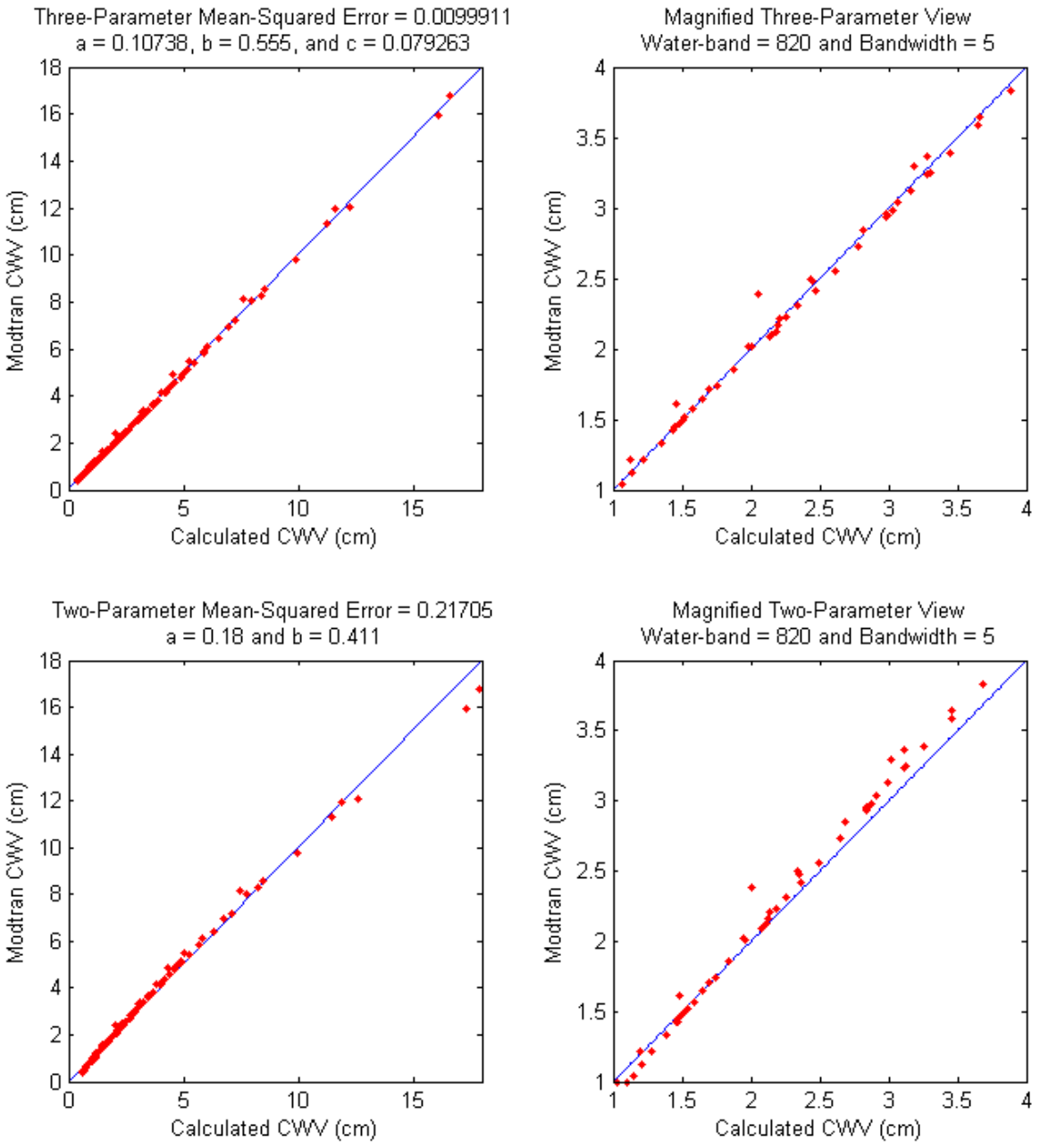


Fig. 4.3: Two- versus three-parameter comparison at 820 nm with 5-nm bandwidth. Magnified view on right.

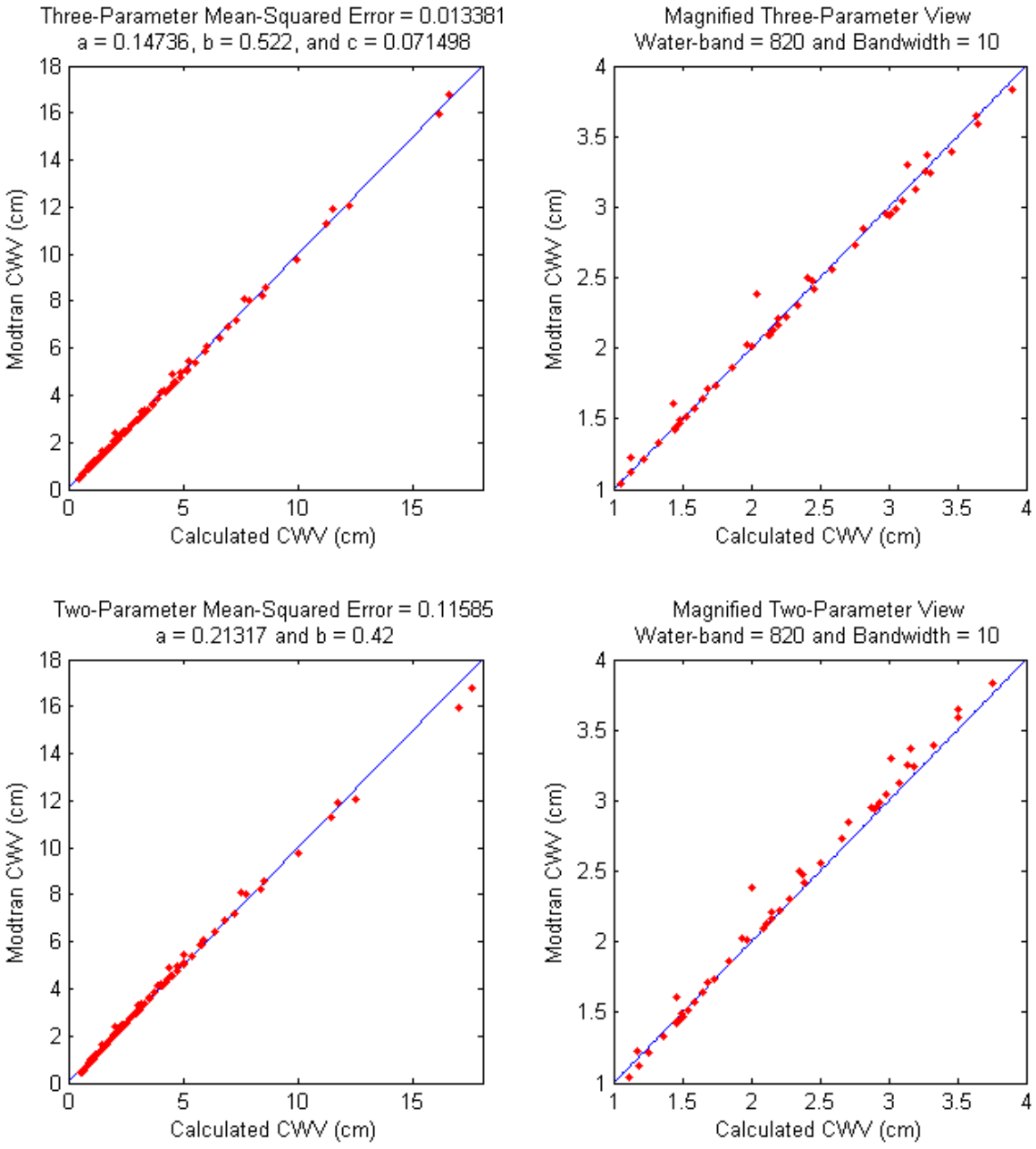


Fig. 4.4: Two- versus three-parameter comparison at 820 nm with 10-nm bandwidth. Magnified view on right.

Table 4.1: Two- versus three-parameter comparison with the magnitude of change corresponding water bands.

Water (nm)	Guard (nm)	Bandwidth (nm)	Two- MMSE	Three- MMSE	Improvement Factor
1465	1600	1	0.0784	0.0313	2.5
1465	1600	5	0.0158	0.0157	1.0
1465	1600	10	0.0258	0.0149	1.7
940	870	1	0.0939	0.0033	28.5
940	870	5	0.0473	0.0021	22.5
940	870	10	0.0470	0.0021	22.4
820	780	1	0.0180	0.0017	10.6
820	780	5	0.2171	0.0010	217.1
820	780	10	0.1158	0.0134	8.6
720	750	1	0.0535	0.0175	3.1
720	750	5	0.2930	0.0227	12.9
720	750	10	0.3698	0.0253	15.7

vapor bands. This led to the following conclusions with respect to the two- versus three-parameter approaches:

1. For all water-bands the performance of the three-parameter approach is equal to or exceeds that of the two-parameter.

2. Both approaches produce the smallest MMSE's at water absorption bands of 820-nm and 940-nm.

3. At these wavelengths with a bandwidth less than 10-nm the three- outperforms the two-parameter by at least a factor of 10.

4. The three-parameter approach performs better over the wide range of possible values of CWV.

5. At 940-nm with a bandpass of 10-nm, AERONET's bandpass, the three-parameter approach demonstrated an improvement by a factor of 22.4.

From these observations the conclusion can be drawn that, for all wavelengths, the three-parameter water vapor absorption model is the best choice. Therefore from this point on, the comparison will focus only on the three-parameter approach.

Chapter 5

Optimal MMSE Comparison

Once the new three-parameter algorithm was selected, tests were performed to determine the optimal bandpass for each water vapor absorption band. Once found, each bandpass was compared to ascertain the optimal bandpass and water vapor absorption band. The optimal conditions found will determine the corresponding measurement technique and instrument.

5.1 Variable Center Wavelengths / and Bandwidths

In sec. 4.4 some observations can also be made concerning the location of the center wavelengths. By changing from one water vapor band to another the MMSE changes. These changes lead to the observation that one of these water bands produces the smallest MMSE. There are two variables over which the MMSE can change, water-band wavelength and bandwidth. This is seen in table 4.1 by examining the three-parameter MMSE column. As the wavelength and bandwidth change so does the magnitude of the MMSE.

To find the optimal bandwidth and water band, the MMSE was calculated over all four water bands and over various bandwidths. As seen in sec. 4.4 the two water bands which produce the smallest MMSE are 820 and 940 nm. These observations lead to a further examination of these two water bands. Figures 5.1-5.4 visually demonstrate which water bands, 820 nm and 940 nm, produce the smallest MMSE. This is shown in further detail in table 5.1.

5.2 Synopsis of Results

Wavelengths 820-nm and 940-nm have proven to be the best for minimizing the MMSE. To ascertain what measurement instrument characteristics are necessary to minimize the

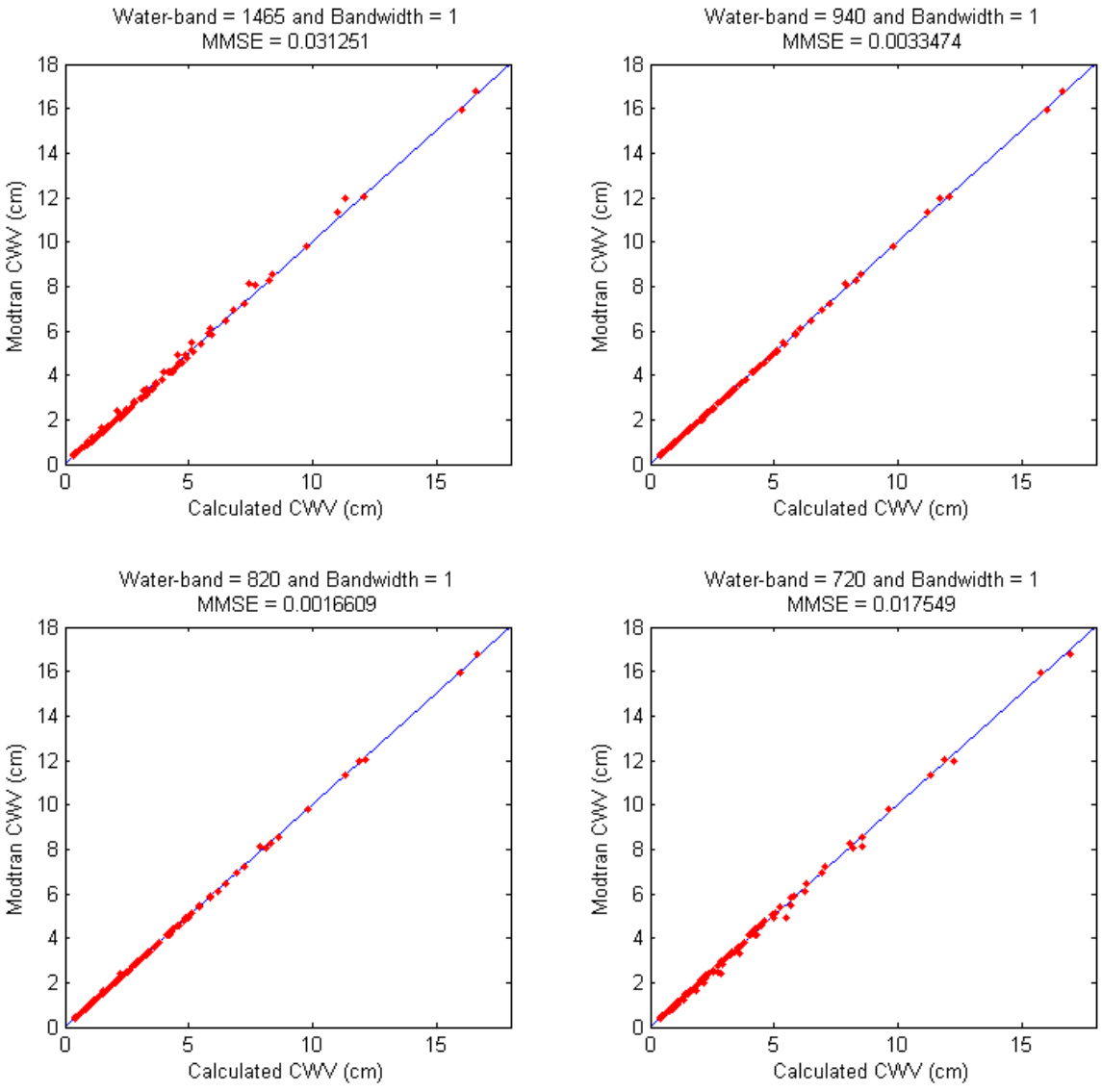


Fig. 5.1: Water band comparison with a bandwidth of 1 nm.

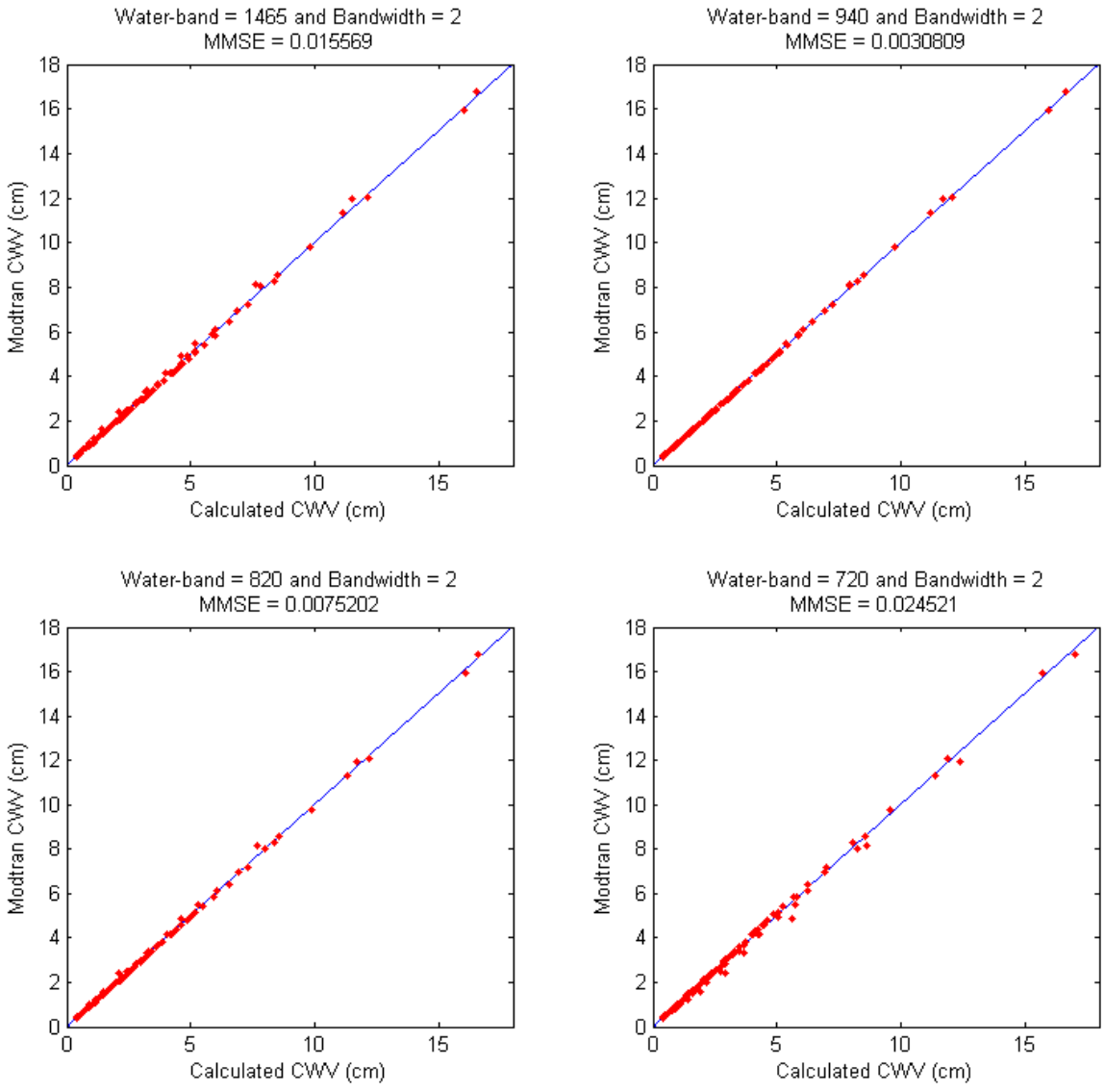


Fig. 5.2: Water band comparison with a bandwidth of 2 nm.

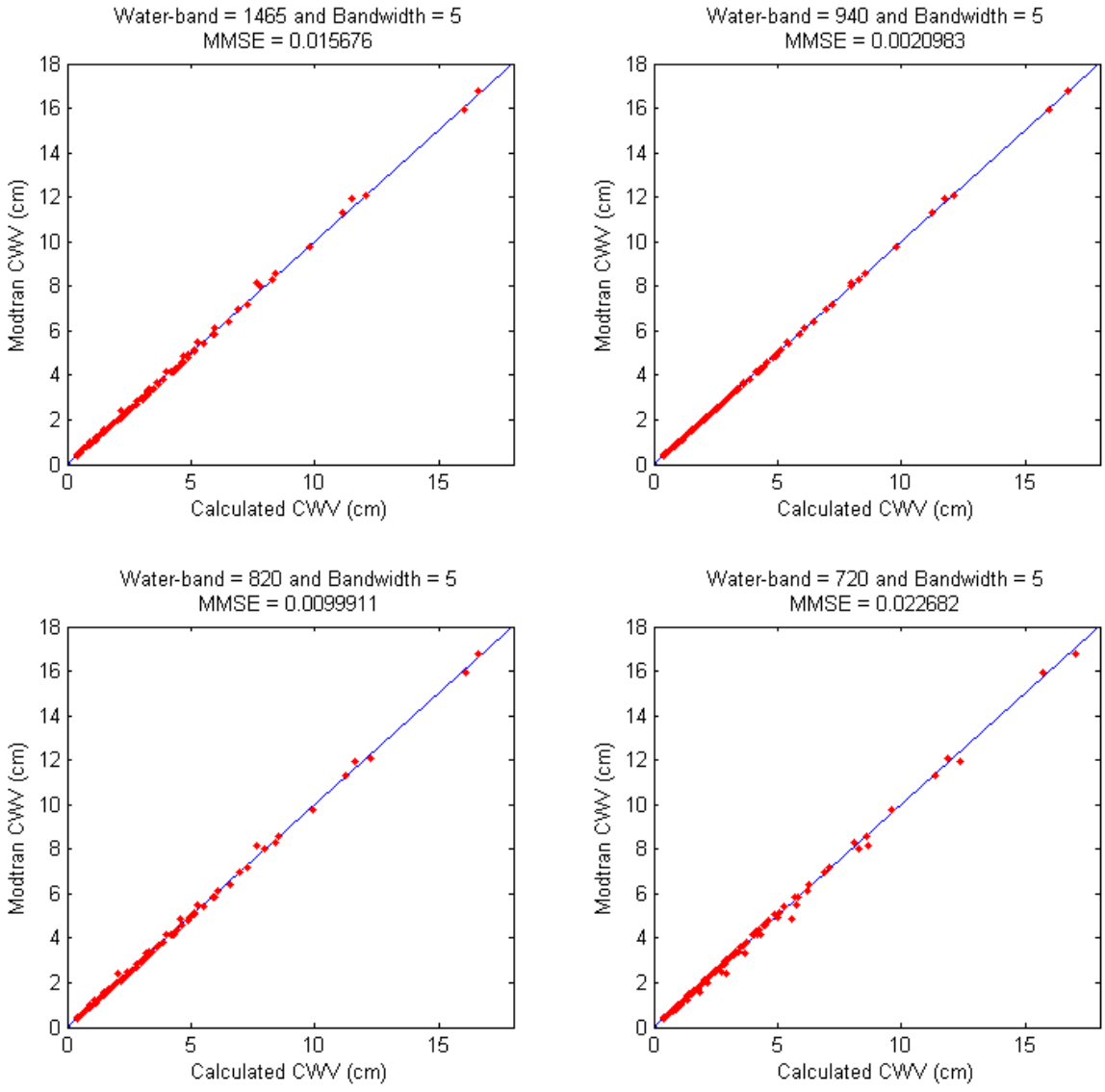


Fig. 5.3: Water band comparison with a bandwidth of 5 nm.

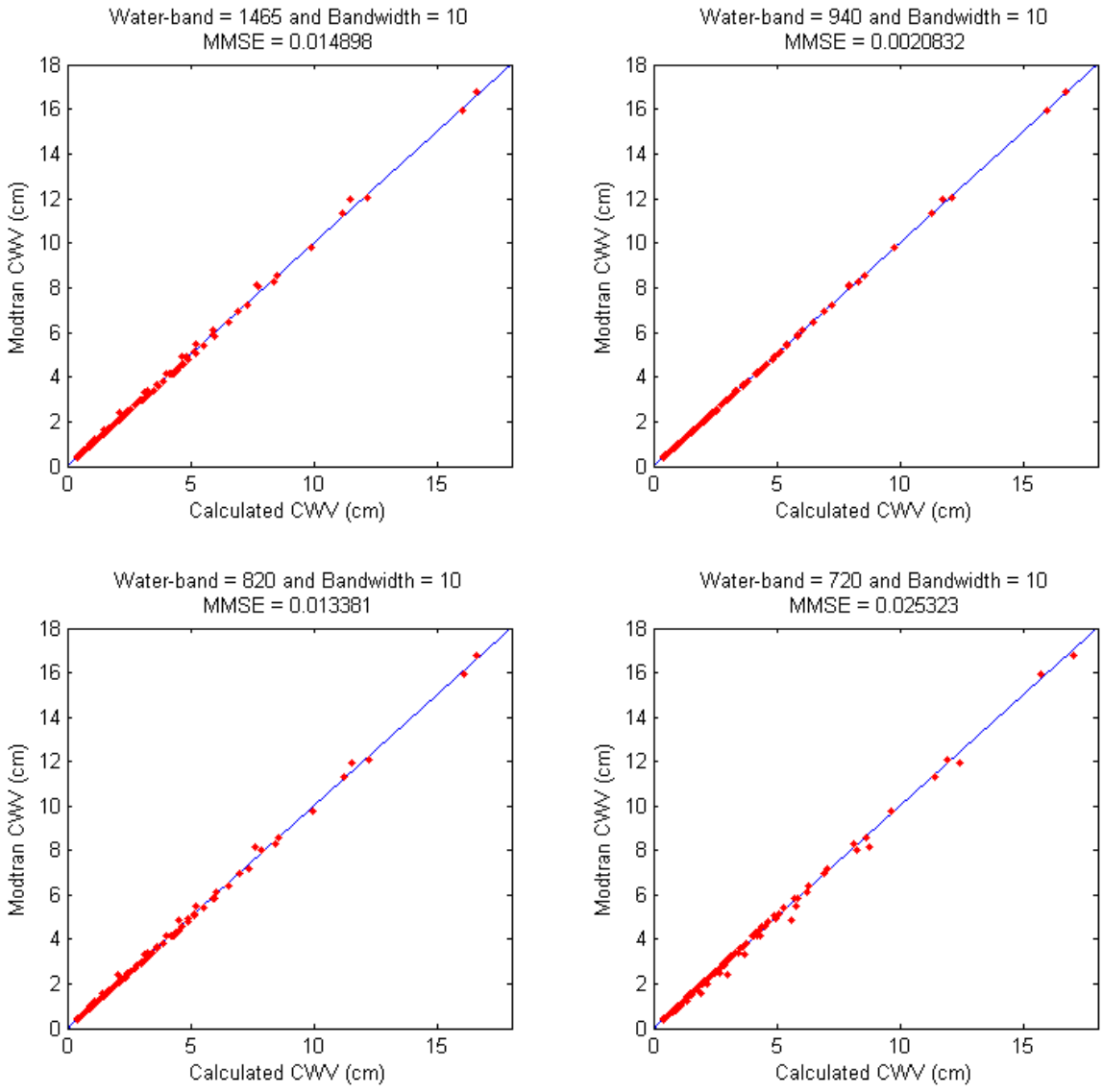


Fig. 5.4: Water band comparison with a bandwidth of 10 nm.

Table 5.1: MMSE of three-parameter algorithm at various bandwidths and wavelengths.

Water (nm)	Width 1 nm	Width 2 nm	Width 5 nm	Width 10 nm
1465	0.0313	0.0156	0.0157	0.0149
940	0.0033	0.0031	0.0021	0.0021
820	0.0017	0.0075	0.0010	0.0134
720	0.0175	0.0245	0.0227	0.0253

error, an in-depth comparison was conducted to ascertain the model optimized solution. For this comparison, several bandwidths were used for both wavelengths. The results of this comparison are seen in table 5.2. The full comparison of all four wavelengths is given in Appendix B.

The results in table 5.2 demonstrate that both water bands produce a very small error. The following observations are made from analyzing these results:

1. A nominal change in the error with respect to bandwidth over the 940 band. Therefore changing the bandwidth from 10 to 1 nm did not significantly improve the accuracy of the algorithm as seen at other bandwidths.

2. The optimal bandwidths for 940 nm are 3 and 5 nm.

3. The improvement of the 3 to 5 nm bandwidth is nominal over the 940 nm band.

4. The 820 nm band shows improvement with using a smaller and smaller bandwidth.

5. The MMSE results by using the 820 nm band, with a bandwidth of 1-nm.

6. The 820 nm band shows an improvement factor of 14 by changing the bandwidth from 15 to 1 nm.

7. Nominal improvement is seen by using the optimal 820 nm band versus the 940 nm band.

5.3 Model Optimized Algorithm / Measurement Instrument

From the observations in secs. 4.4-5.2, the model optimized algorithm and instrument setup can be chosen. Through testing of the two- and three-parameter algorithms in sec.

4.4, it was found that the three-parameter demonstrates significant improvement over the existing two-parameter approach at all water bands. This demonstrates that by adding the third additive parameter the water vapor absorption can be better approximated. It was also found in this section that the 820 and 940 nm water bands produce significantly smaller errors than the 720 and 1465 nm bands.

These results changed the focus from which of approaches, two- versus three-parameter, is the model optimized algorithm to under what circumstances is the error minimized due to the choice of water vapor absorption band and of bandwidth. In sec. 5.2 it was found that the error is minimized under the circumstances that the water band is located at 820 nm and the bandwidth at 1-nm. It was also found that the algorithm shows little to no improvement of the 820 nm band at 1 nm compared with the 940 nm band at 3 to 5 nm. Therefore, the results show both the 820 and 940 nm bands create the optimal solutions. This facilitates the selection of which measurement instrument to use. If the 820 band were chosen, a spectrometer would be needed to produce measurements with such narrow bandwidths. If the 940 nm band were chosen, then either a spectrometer or a radiometer would produce the good results. This is due to the error not changing relative to bandwidth.

By the selection of water band and bandwidth, the optimal parameters for these solutions are available by the same methods. These can parameters have been calculated for water bands 820 and 940 nm and been placed in table 5.3. Refer to Appendix C for the code used to produce these results.

Table 5.2: 820 versus 940 nm comparison.

Bandwidth (nm)	820 nm MMSE	940 nm MMSE
1	0.0017	0.0033
2	0.0075	0.0031
3	0.0084	0.0021
4	0.0098	0.0040
5	0.0100	0.0021
6	0.0079	0.0023
7	0.0080	0.0023
8	0.0104	0.0024
9	0.0123	0.0031
10	0.0134	0.0021
11	0.0161	0.0032
12	0.0177	0.0032
13	0.0189	0.0032
14	0.0219	0.0031
15	0.0242	0.0033

Table 5.3: Calculated parameters a, b, and c.

Water Band (nm)	Bandwidth (nm)	a	b	c
820	1	0.1541	0.5160	0.0329
820	2	0.1026	0.5555	0.0487
820	3	0.0920	0.5780	0.0887
820	4	0.0993	0.5675	0.0904
820	5	0.1074	0.5550	0.0793
820	6	0.1271	0.5370	0.0757
820	7	0.1503	0.5180	0.0649
820	8	0.1506	0.5215	0.0606
820	9	0.1443	0.5300	0.0663
820	10	0.1474	0.5220	0.0715
940	1	0.2864	0.7430	0.1897
940	2	0.3140	0.7330	0.1228
940	3	0.3897	0.6945	0.1701
940	4	0.4255	0.6760	0.1933
940	5	0.4530	0.6720	0.1869
940	6	0.4837	0.6615	0.2080
940	7	0.5222	0.6480	0.1820
940	8	0.5015	0.6600	0.2146
940	9	0.5587	0.6340	0.2166
940	10	0.5460	0.6480	0.2104

Chapter 6

Conclusions / Proposed Future Work

6.1 Conclusions

The following are some of the conclusions which were formulated from the research reported in this thesis:

1. Through field tests using an optical spectrometer incorporated in the SAM (Sun and Aureole Measurement) instrument, it was ascertained that the spectrometer could not be sufficiently calibrated to produce a valid column water vapor data set. A replacement spectrometer was acquired based upon the guidelines formulated from the field tests including (a) temperature stabilization, (b) auto-dark current subtraction, and (c) response in the near-infrared spectral range. The replacement spectrometer is under calibration tests.

2. Research into already-available algorithms gave improved insight into approximating water vapor absorption. The combination of two separate water vapor absorption models demonstrated improvements by a factor of 22 over the available models' individual performance using a three fitting parameter approach. Through a minimum mean squared error test the improvement of a new algorithm over the original demonstrated for each of the 720, 820, 940, and 1465 nm water vapor absorption bands. Consequently, using the specifications of the national AERONET network sensor at 940 nm, a factor of improvement of 22 was demonstrated. Such a significant improvement demonstrated that the three-parameter algorithm to be a model optimized algorithm.

3. Minimum mean-squared error tests were conducted for the 720, 820, 940, and 1465 nm absorption bands. These results were tabulated to select the optimal water band to

employ in the SAM instrument. The results also demonstrated for each of the water vapor absorption bands, what corresponding spectral bandwidth would give the best results.

4. A model optimized algorithm was developed. The near optimal algorithm was found to produce the least MMSE by using either the 820 nm band with a 1-nm bandwidth or the 940 nm band with a 5-nm bandwidth.

5. Characterization parameters a, b, and c were calculated and tabulated for quick reference for various instrument setups. Each set of parameters provide the least minimum mean squared error for each corresponding water band and bandwidth.

6.2 Proposed Future Work

As ascertained in Chapter 4 the near-optimal algorithm uses the new three-parameter approach at either 820 or 940 nm. The results were formed by using the MODTRAN absorption model. The data produced by the MODTRAN model represent a portion of the solar spectrum as if there were no instrument measurement error.

In the case of the SAM instrument, the first spectrometer was tested in order to characterize its response. The responsivity was a function of wavelength. This function signifies the need to characterize the signal-to-noise ratio of the spectrometer dependant on wavelength. The variable responsivity, shown in fig. 2.1, is directly related to the signal-to-noise ratio.

A new spectrometer has been selected which addresses the problems of the varying responsivity due to temperature and the varying dark current. A Ocean Optics USB2000+ Miniature Fiber Optic Spectrometer was chosen to fit the SAM sensor specifications. The new spectrometer is temperature stabilized. This temperature stabilization needs to be quantified.

Tests to characterize the percent change in measurements as the spectrometer varies over this temperature change degree are under way. A constant light source is measured by the spectrometer which is tested at the maximum temperature range. The percent change

values from this test, will be produced according to wavelength. A smaller magnitude of change demonstrates greater validity at that specific wavelength.

Another test concerns the responsivity of the spectrometer selected. In Chapter 4 ideal conditions were assumed, and therefore the results for a near optimal algorithm at 820 and 940 nm could change with respect to the spectrometer selected. A test to find a transfer function between the signal-to-noise ratio responsivity of a spectrometer should be conducted. The transfer function derived could then be used in the minimum mean squared error derivation in Chapter 4. This new derivation would then represent the near optimal algorithm for the selected spectrometer. This algorithm would produce the parameters a, b, and c and the corresponding optimal bandpass as well.

References

[1] J. DeVore, "SAM Ground-Based Measurement of Cloud Optical Properties," [http://www.visidyne.com/sam/samhome.htm], Feb. 2008.

[2] E. J. Hopkins, "Radiosonde Observation (RAOB) Data Tabulation," [http://www.aos.wisc.edu/ hopkins/aos100/raobdoc.htm], Feb. 2008.

[3] N. Robinson, *Solar Radiation.* Elsevier Publishing Company, 1966.

[4] Ocean Optics, "Photon Control Products - Spectroscopy," [http://www.photon-control.com/spectroscopy.html], Jan. 2008.

[5] Toshiba, "TCD1304AP CCD," [http://www.zygo.btinternet.co.uk/tcd1304ap.pdf], Jan. 2008.

[6] Atmospheric Science Program (ASP), "Cumulus Humilis Aerosol Processing Study (CHAPS)," [http://asp.labworks.org/], Jan. 2008.

[7] NASA Goddard Space Center, "AERONET (AErosol RObotic NETwork) program," [http://aeronet.gsfc.nasa.gov/], Jan. 2008.

[8] J. Reagan, "A comparison of columnar water vapor retrievals obtained with near-ir solar radiometer and microwave radiometer measurements," *American Meteorological Society*, vol. 34, pp. 1384–1391, Jun. 1995.

[9] Wikipedia, "Solar Spectrum," [http://upload.wikimedia.org/wikipedia/commons/4/4c/Solar_Spectrum.png], Dec. 2007.

[10] A. Z. Jones, "Einstein's Wonderful Year," [http://physics.about.com/od /quantumphysics/a/photoelectric2.htm], Mar. 2008.

[11] M. Chaplin, "Water Absorption Spectrum," [http://www.lsbu.ac.uk/water/vibrat.html], Jan. 2008.

[12] G. Herzberg, *Molecular Specta and Molecular Structure.* Litton Educational Publishing, 1945.

[13] K. J. Thome, "Determination of precipitable water from solar transmission," *American Meteorological Society*, vol. 31, pp. 157–165, Feb. 1993.

[14] B. M. Tissue, "Beer-Lambert Law," [http://elchem.kaist.ac.kr/vt/chem-ed/spec/beerslaw.htm], Mar. 2008.

[15] B. Schmid, "Comparison of columnar water-vapor measurements from solar transmittance methods," *Applied Optics*, vol. 40, no. 12, pp. 1886–1896, Apr. 2001.

[16] R. N. Halthore, "Sun photometric measurements of atmospheric water vapor column abundance in the 940-nm band," *Journal of Geophysical Research*, vol. 102, no. D4, pp. 4343–4352, Feb. 1997.

[17] U.S. Air Force Research Lab, "Fact Sheet, Modtran 4 Software," [http://www.kirtland.af.mil/library/factsheets/factsheet.asp?id=7915], Jan. 2008.

Appendices

Appendix A

Additional Two- Versus Three-Parameter Figures

Table A.1: Two- versus three-parameter comparison.

Water (nm)	Guard (nm)	Bandwidth (nm)	Two- MMSE	Three- MMSE	Improvement
1465	1600	1	0.0784	0.0313	2.5
1465	1600	5	0.0158	0.0157	1.0
1465	1600	10	0.0258	0.0149	1.7
940	870	1	0.0939	0.0033	28.5
940	870	5	0.0473	0.0021	22.5
940	870	10	0.0470	0.0021	22.4
820	780	1	0.0180	0.0017	10.6
820	780	5	0.2171	0.0010	217.1
820	780	10	0.1158	0.0134	8.6
720	750	1	0.0535	0.0175	3.1
720	750	5	0.2930	0.0227	12.9
720	750	10	0.3698	0.0253	15.7

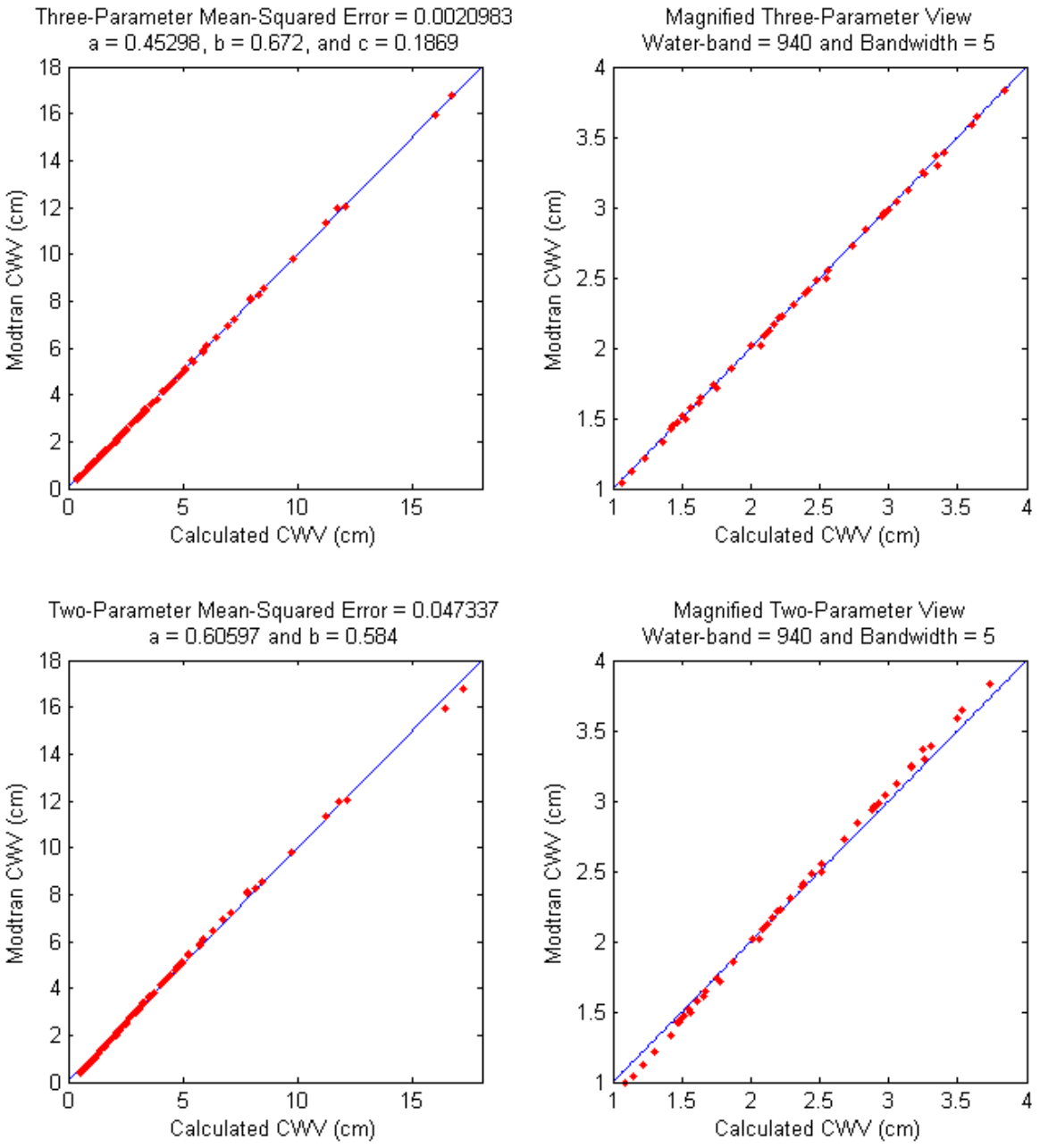


Fig. A.1: Two- versus three-parameter comparison at 940 nm with 5-nm bandwidth. Magnified view on right.

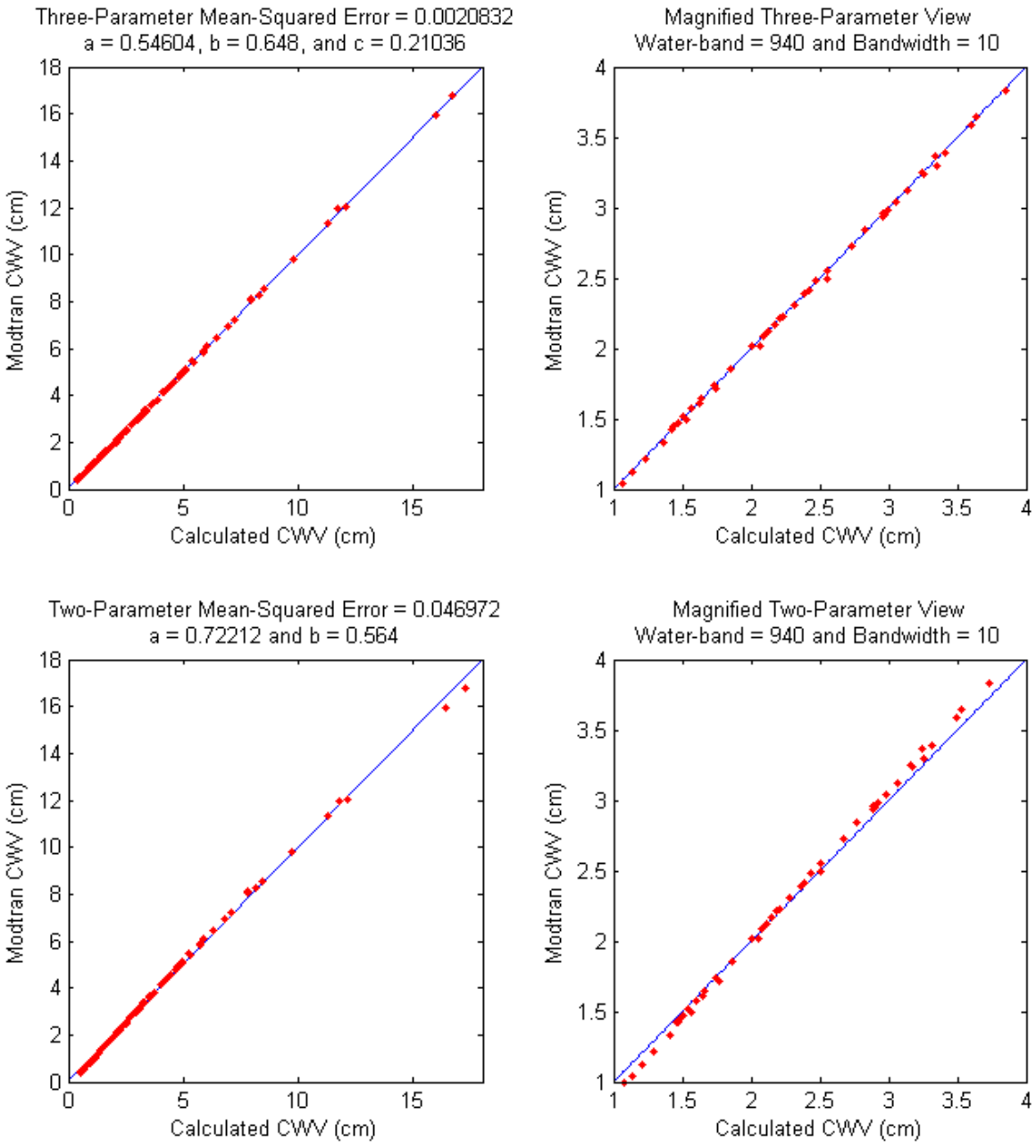


Fig. A.2: Two- versus three-parameter comparison at 940 nm with 10-nm bandwidth. Magnified view on right.

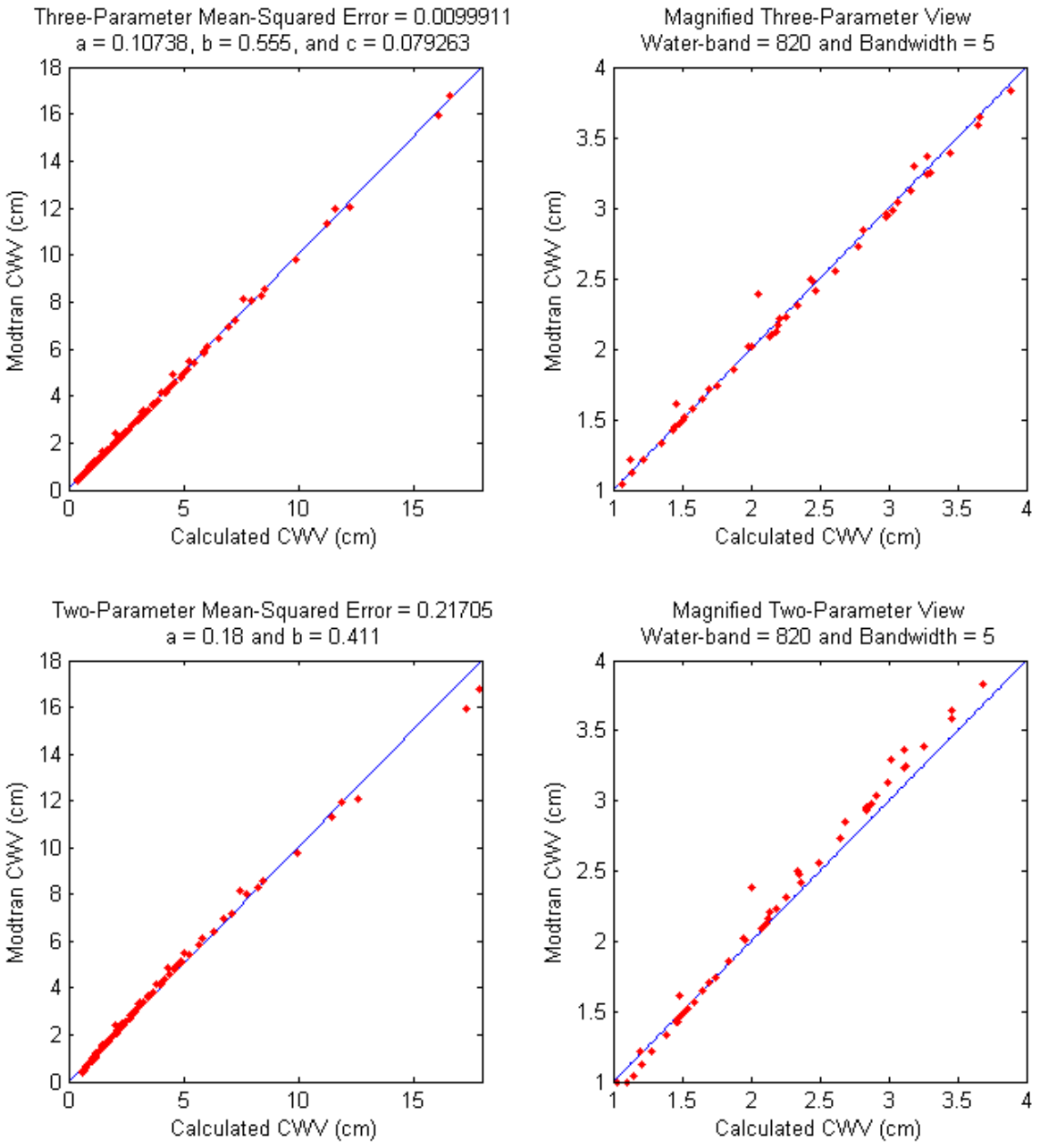


Fig. A.3: Two- versus three-parameter comparison at 820 nm with 5-nm bandwidth. Magnified view on right.

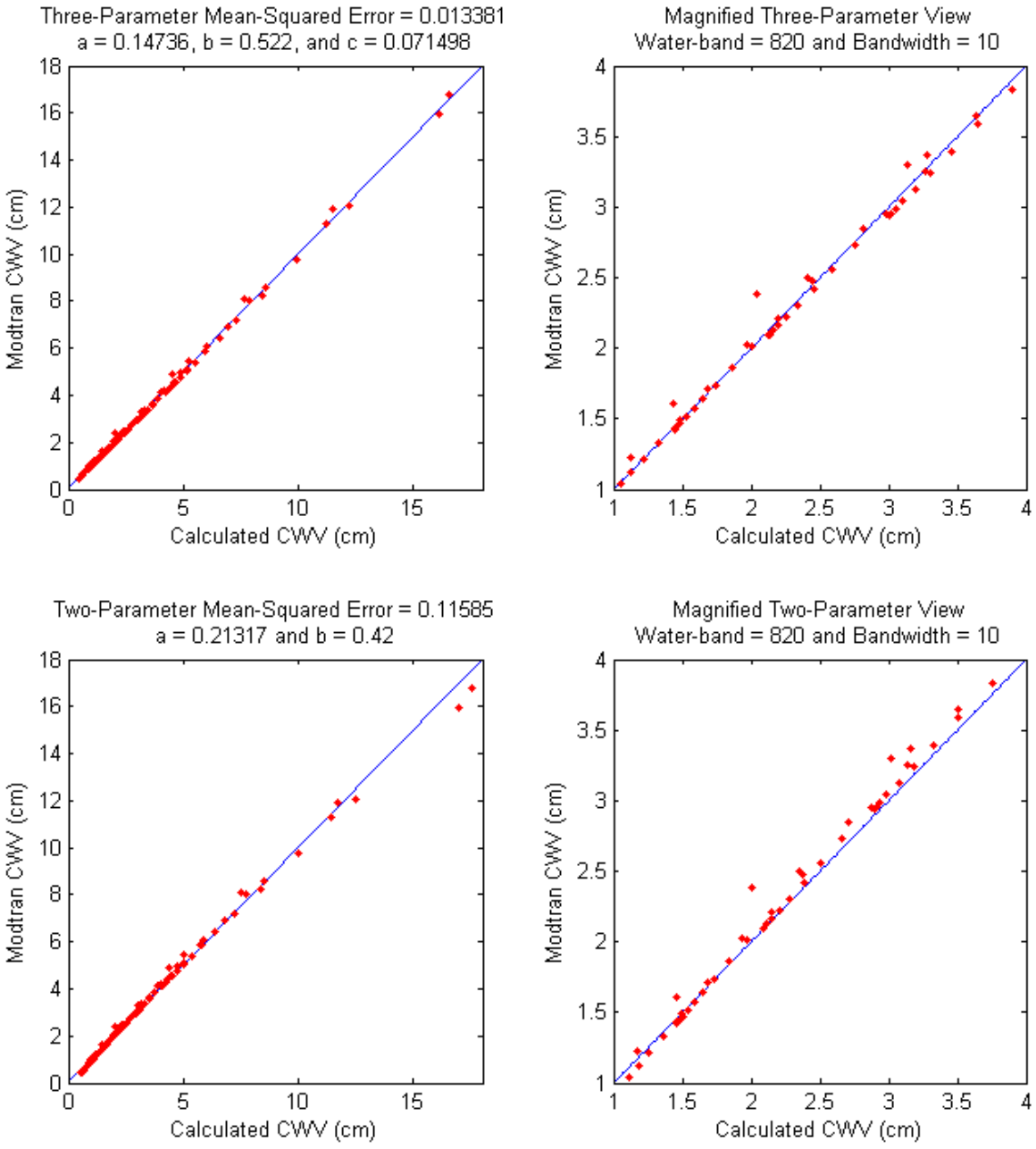


Fig. A.4: Two- versus three-parameter comparison at 820 nm with 10-nm bandwidth. Magnified view on right.

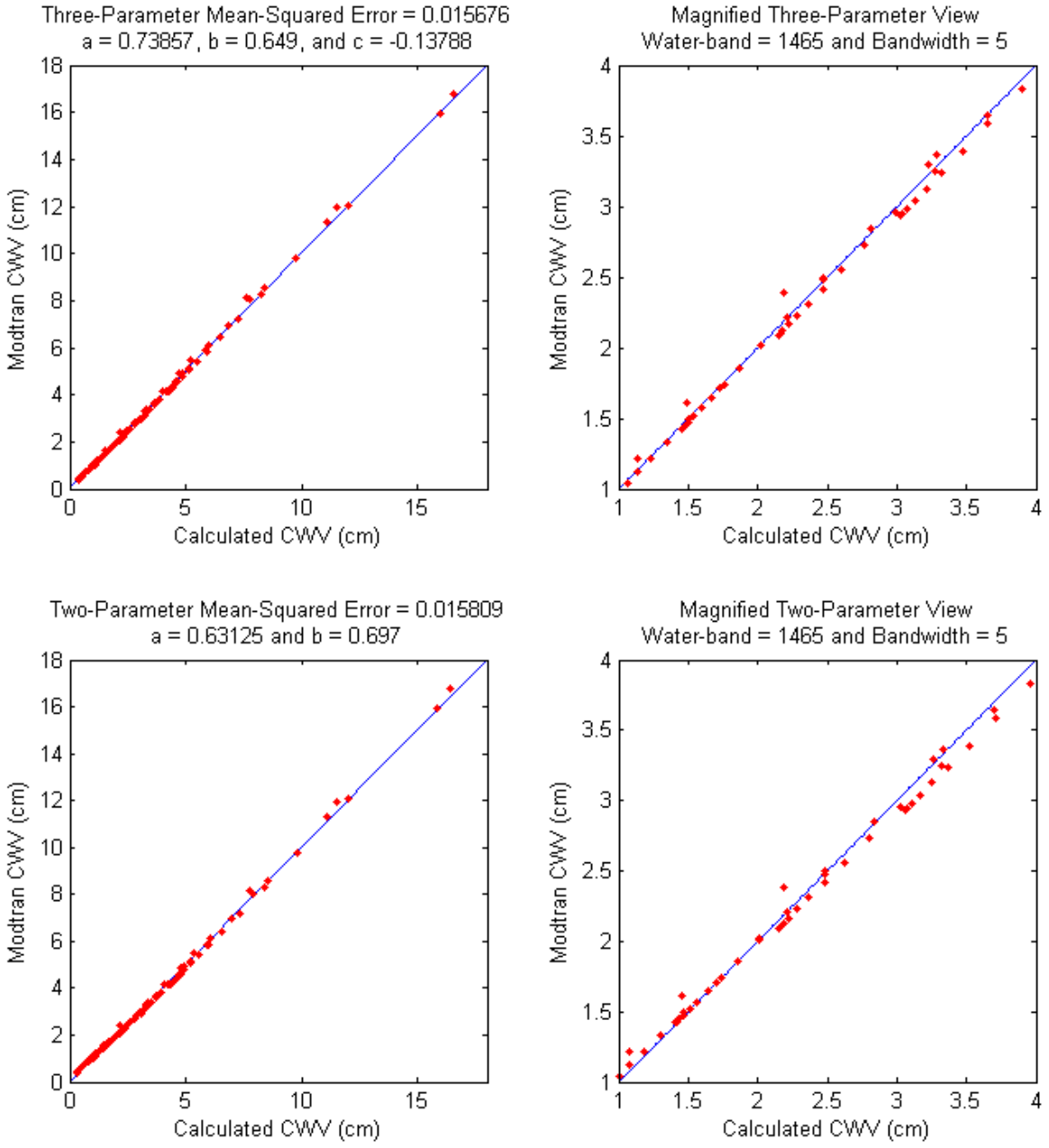


Fig. A.5: Two- versus three-parameter comparison at 1465 nm with 5-nm bandwidth. Magnified view on right.

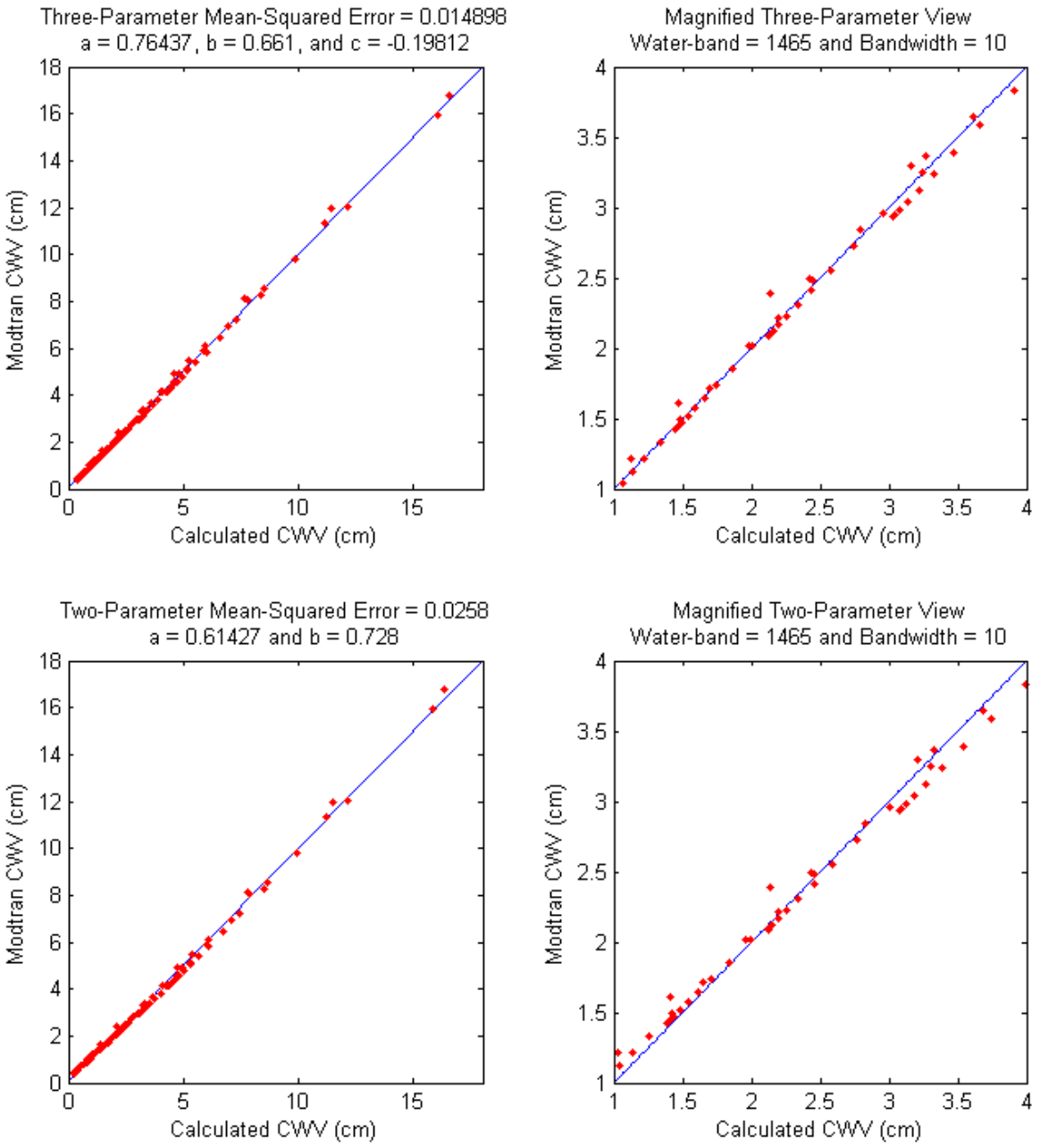


Fig. A.6: Two- versus three-parameter comparison at 1465 nm with 10-nm bandwidth. Magnified view on right.

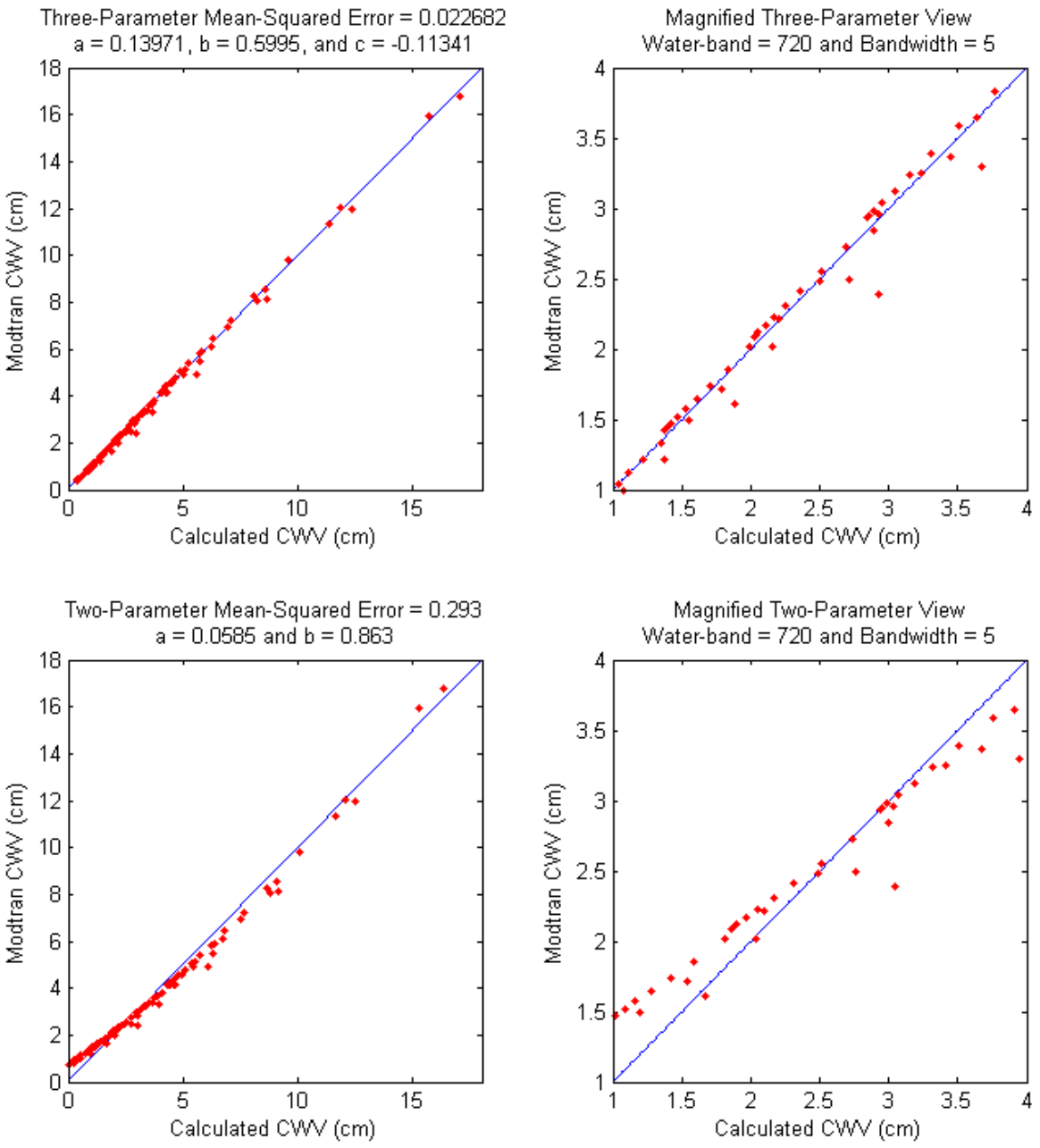


Fig. A.7: Two- versus three-parameter comparison at 720 nm with 5-nm bandwidth. Magnified view on right.

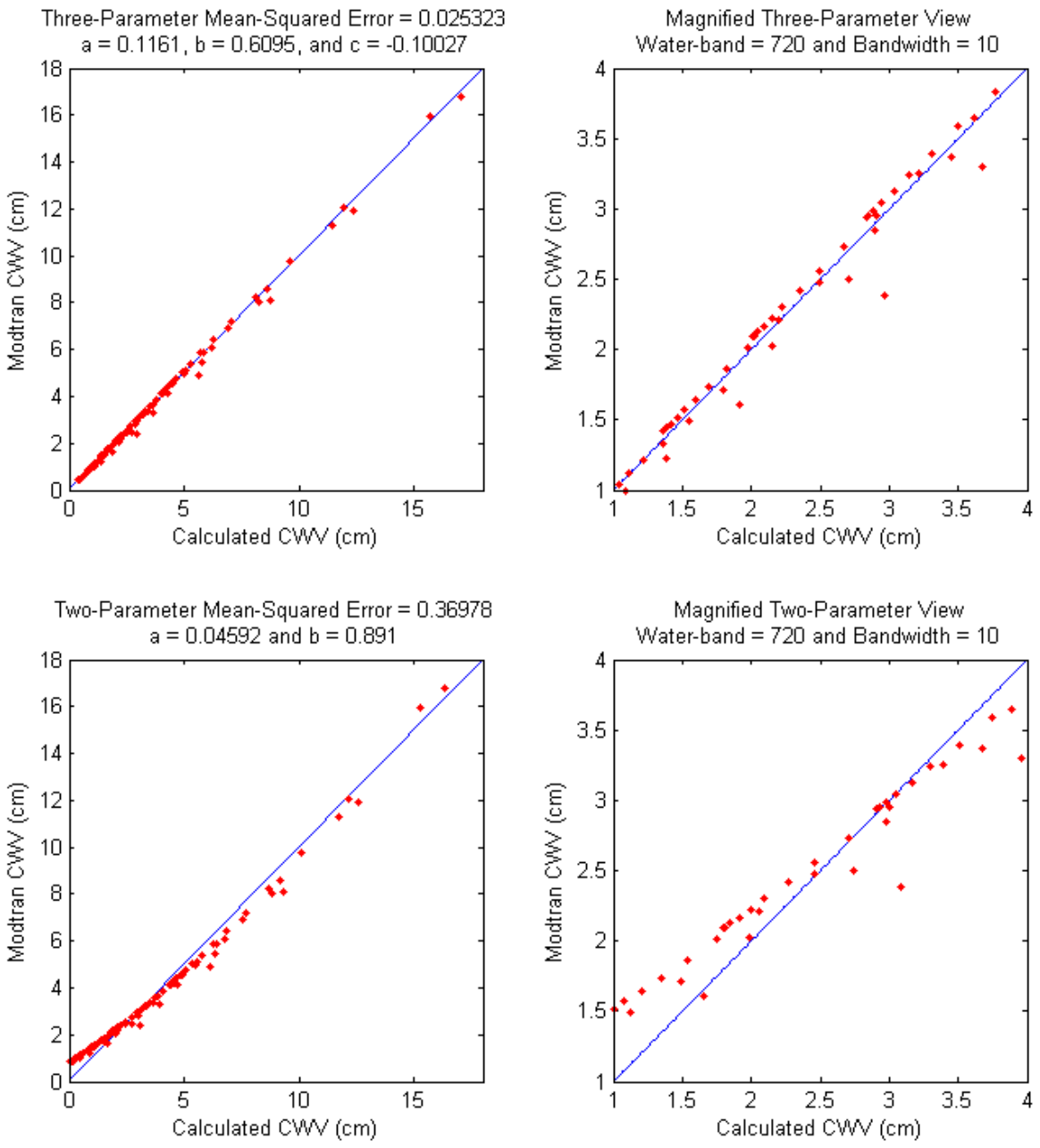


Fig. A.8: Two- versus three-parameter comparison at 720 nm with 10-nm bandwidth. Magnified view on right.

Appendix B

Full Table of Results

Table B.1: All water bands comparison.

Bandwidth (nm)	720 nm MMSE	820 nm MMSE	940 nm MMSE	1465 nm MMSE
1	0.0175	0.0017	0.0033	0.0313
2	0.0245	0.0075	0.0031	0.0156
3	0.0238	0.0084	0.0021	0.0168
4	0.0253	0.0098	0.0040	0.0253
5	0.0227	0.0100	0.0021	0.0157
6	0.0207	0.0079	0.0023	0.0219
7	0.0189	0.0080	0.0023	0.0217
8	0.0217	0.0104	0.0024	0.0204
9	0.0246	0.0123	0.0031	0.0164
10	0.0253	0.0134	0.0021	0.0149
11	0.0282	0.0161	0.0032	0.0131
12	0.0313	0.0177	0.0032	0.0142
13	0.0334	0.0189	0.0032	0.0150
14	0.0354	0.0219	0.0031	0.0162
15	0.0389	0.0242	0.0033	0.0150

Appendix C

Code

C.1 Least-Squares Code

```
    function [a_out b_out c_out MMSE] =
constants(width,water_band,
guard_band,plots)
%————————————————————————
% This function performs a least-squares brute force estimation of the
% constants a, b, and c for the column water vapor measurements. This code
% was written for the thesis of Josh Williams, but can be used to calculate
% the correct a, b, and c for any set of water and guard bands. This will
% also allow you to change the width of the filter used from the
% spectrometer. This code will conduct the analysis and plot the result for
% both the two and three parameter approaches. The error_squared and values
% of the constants is output in the title of each plot.
%
% Inputs: width = width of the filter in nm over both the guard and water
% bands.
% water_band = center of the water_band in nm
% guard_band = center of the guard_band in nm
% plots = 1 to plot, 0 to avoid plotting
%
% Outputs: a, b, c = The calculated values of each of the parameters in
% the three constant approach. Two constant can be
% output if wanted, but is not in this version of the
% code. This is produced in the plots.
% MMSE = The MMSE value of the error for the entry parameters
```

```
%
% See also: READ_MODTRAN_FILE.M, MMSE_VALUES.M
%-----------------------------------------------------------------------
if( plots)
closeall;
end
load('Modtran.mat');
%Load in the training data
Modtran.wavelength = 1000 * Modtran.wavelength;
%Change wavelength from um to nm
w = find(Modtran.wavelength < water_band + width/2&
Modtran.wavelength > water_band - width/2);
%Find the index of the correct wavelength band from input for water-band
g = find(Modtran.wavelength < guard_band + width/2&
Modtran.wavelength > guard_band - width/2);
%Find the index of the correct wavelength band from input for guard-band
u = Modtran.cwv;
%Assign u to the corresponding cwv values from training data
tau_z = 8.6 * 10^-3./(Modtran.wavelength.^(4 + .04)); %All Reyleigh scattering
tau_g = mean(tau_z(g)); %Solve for the Reyleigh scattering of guard band
tau_w = mean(tau_z(w)); %Solve for the Reyleigh scattering of water
tau_r = tau_g - tau_w; %Solve for tau_r for equation
%Assign Corresponding values over ranges specified by inputs
Rw = sum(Modtran.data(w,:)); Rg = sum(Modtran.data(g,:)); R0g =
mean(Modtran.exo_solar(g)); R0w = mean(Modtran.exo_solar(w));
%Calculate the Atmospheric constant value
R0 = R0w/R0g; Or = R0;
%Assume all m values equal zero, See Josh Williams thesis for explaination
m = ones(1, length(Rw));
b = .001 : .0005 : 2; %Possible range of values for b
%Intialize all vectors
a = zeros(1, length(b)); c = a; er = a; er2 = a;
```

```
%Start Analysis of Three-Parameter Approach over all values of 0¡b¡=2
for i = 1 : length(b)
m11 = 0;
m12 = 0;
m21 = length(Rw);
m22 = 0;
d1 = 0;
d2 = 0;
for j = 1 : length(Rw)
m11 = m11 + (u(j) * m(j))^b(i);
m12 = m12 + (u(j) * m(j))^(2 * b(i));
d1 = d1 + (log(Or * Rg(j)/Rw(j)) + m(j) * tau_r) * (m(j) * u(j))^b(i);
d2 = d2 + log(Or * Rg(j)/Rw(j)) + m(j) * tau_r;
end
m22 = m11;
M = [m11, m12; m21, m22];
temp = inv(M) * [d1; d2];
a(i) = temp(2);
c(i) = temp(1);
for j = 1 : length(Rw)
er2(i) = er2(i) + (c(i) + a(i) * (m(j) * u(j))^b(i) - log(Or * Rg(j)/Rw(j)))^2;
er(i) = er(i) +
(1/m(j) * ((log(Or * Rg(j)/Rw(j)) - c(i))/a(i)).^(1/b(i)) - u(j))^2;
end
er(i) = er(i)/length(Rw);
end
valindx = min(er2);
MSE = real(er(indx));
MMSE = MSE;
if(plots)
figure
subplot(221)
```

```
line = 0 : .01 : 18;
plot(line, line);
holdon
plot(1./m. * ((log(Or * Rg./Rw) - c(indx))./a(indx)).^(1/b(indx)), u,'r.');
ylabel('ModtranCWV(cm)')
xlabel('CalculatedCWV(cm)')
xlim([0, 18]);
ylim([0, 18]);
subplot(222)
plot(line, line);
holdon;
plot(1./m. * ((log(Or * Rg./Rw) - c(indx))./a(indx)).^(1/b(indx)), u,
'r.');
ylabel('ModtranCWV(cm)')
xlabel('CalculatedCWV(cm)')
xlim([1, 4]);
ylim([1, 4]);
end
%if(plots)
%plot(b, er2)
%xlabel('b')
%ylabel('ErrorSquared')
%figure
%end
a_out = a(indx); b_out = b(indx); c_out = c(indx);
% Start Analysis of Two-Parameter Approach over all values of 0¡b¡=2
b = .001 : .001 : 2; a = zeros(1, length(b)); er = a; er2 = a;
fori = 1 : length(b)
num = 0;
den = 0;
forj = 1 : length(Rw)
num = num + log(R0 * Rg(j)/Rw(j)) * (m(j) * u(j))^b(i);
```

$den = den + (m(j) * u(j))^(2 * b(i));$

end

$a(i) = num/den;$

$for j = 1 : length(Rw)$

$er2(i) = er2(i) + (a(i) * (m(j) * u(j))^b(i) - log(R0 * Rg(j)/Rw(j)))^2;$

$er(i) = er(i) +$

$(1/m(j) * ((log(Or * Rg(j)/Rw(j)))/a(i)).^(1/b(i)) - u(j))^2;$

end

$er(i) = er(i)/length(Rw);$

end

$valindx = min(er2); MSE = real(er(indx));$

$if(plots)$

$subplot(223)$

$line = 0 : .01 : 18;$

$plot(line, line);$

$holdon;$

$plot(1./m. * (log(R0 * Rg./Rw)./a(indx)).^(1/b(indx)), u, 'r.');$

$ylabel('ModtranCWV(cm)')$

$xlabel('CalculatedCWV(cm)')$

$xlim([0, 18]);$

$ylim([0, 18]);$

$subplot(224)$

$plot(line, line);$

$holdon$

$plot(1./m. * (log(R0 * Rg./Rw)./a(indx)).^(1/b(indx)), u, 'r.');$

$ylabel('ModtranCWV(cm)')$

$xlabel('CalculatedCWV(cm)')$

$xlim([1, 4]);$

$ylim([1, 4]);$

end

$\%if(plots)$

$\%figure$

$\%plot(b, er2)$

$\%xlabel('b')$

$\%ylabel('ErrorSquared')$

$\%end$

C.2 Wavelength Comparison

function [MMSE] = wavelength_comparison_plots()

%————————————————————————–

%This function is written to produce the comparison plots over water band

%wavelengths. This will produce one plot with four figures in each.

%Inputs: []

%Outputs: plots for each of the conditions.

%Example Code: two_three_comparison_plots;

%See also: CONSTANTS.M, READ_MODTRAN_FILE.M

%————————————————————————-

close all;

$water = [1465, 940, 820, 720];$

$guard = [1600, 870, 780, 750];$

$MMSE = zeros(5, 5);$

$MMSE(1, 2 : end) = [12510];$

$MMSE(2 : end, 1) = water';$

$count = 2; fori = [1, 2, 5, 10]$

$figure$

$forj = 1 : 4$

$subplot(2, 2, j);$

$MMSE(j + 1, count) = constants3(i, water(j), guard(j), 1);$

end

$count = count + 1;$

end

C.3 Read MODTRAN File

function [] = read_modtran_file()

%————————————————————

%This function reads in the MODTRAN values from the two files

%MODTRAN_CWV.xls and MODTRAN_data.xls. This was done because this eases

%the process of reading in the exel file produced by John Devore. This

%file does not output anything, but does save the structure to an .mat file

%named Modtran.mat. This file is to be loaded later by another program

%such as constants.m

%Inputs: []

%Outputs: []

%Example Code: read_modtran_file;

%See also: CONSTANTS.M

%————————————————————

$Modtran = struct('m', zeros(1,804),'cwv', zeros(1,804),'wavelength', ...$

$zeros(1,19002),'exo_solar', zeros(1,19002),'data', zeros(19002,804));$

$temp = xlsread('MODTRAN_data.xls');$

$Modtran.wavelength = temp(:,1);$

$Modtran.data = temp(:,3:end);$

$Modtran.exo_solar = temp(:,2);$

$temp = xlsread('MODTRAN_CWV.xls');$

$Modtran.m = temp(1,:);$

$Modtran.cwv = temp(2,:);$

$cleartemp$

$save('Modtran.mat')$

C.4 Two- and Three-Parameter Comparison Table

function out = two_three_comparison_table()

%————————————————————————————

%This function is written to produce the comparison table of the two versus

```
%three parameter approach. The output is a matrix of values
%Inputs: []
%Outputs: out = matrix of value in the following order, each representing a
% column. water, guard, width,two_error,three_error
%Example Code: two_three_comparison_table;
%See also: CONSTANTSM, READ_MODTRAN_FILEM
%————————————————————————————
water = [1465,940,820,720];
guard = [1600,870,780,750];
out = zeros(8,5);
count = 1;
for j = 1:4
for i = [1 5 10]
out(count,4)out(count,5) = constants2(i,water(j),guard(j),0);
out(count,1) = water(j);
out(count,2) = guard(j);
out(count,3) = i;
count = count + 1;
end
end
```

C.5 MODTRAN Plot

```
    function [] = Modtran_water_band_plot()
close all
clear all
clc
load('Modtran.mat');
Modtran.wavelength = Modtran.wavelength * 1000;
range = find(Modtran.wavelength ¡ 1600 & Modtran.wavelength ¿ 700);
plot(Modtran.wavelength(range),Modtran.data(range,10));
title('MODRTAN Solar Irradiance Spectrum')
```

xlabel('Wavelength (nm)')

ylabel('Solar Irradiance')