
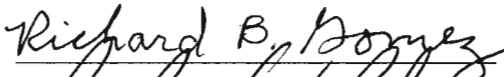
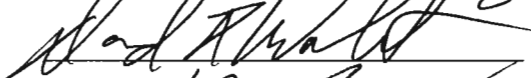
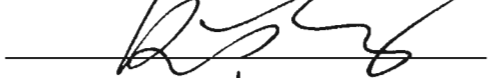
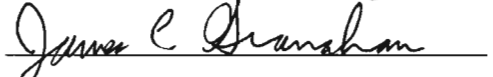
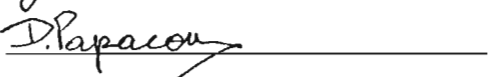




TRACE GAS DETECTION IN HYPERSPECTRAL IMAGERY USING THE
WAVELET PACKET SUBSPACE

by

Mark A. Z. Salvador
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Computational Sciences and Informatics

Committee:

	Dr. Ronald G. Resmini, Dissertation Co-Director
	Dr. Richard B. Gomez, Dissertation Co-Director
	Dr. David F. Walnut, Committee Member
	Dr. Ruixin Yang, Committee Member
	Dr. James C. Granahan, Committee Member
	Dr. Dimitrios A. Papaconstantopoulos, Department Chairperson
	Dr. Peter A. Becker, Associate Dean for Graduate Programs, College of Science
	Dr. Vikas Chandhoke, Dean, College of Science

Date: April 16, 2008

Spring Semester 2008
George Mason University
Fairfax, VA

Trace Gas Detection in Hyperspectral Imagery Using the Wavelet Packet Subspace

A dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy at George Mason University

By

Mark A. Z. Salvador
Master of Science
George Mason University, 2001
Bachelor of Science
Syracuse University, 1990

Co-Director: Ronald G. Resmini, Associate Professor
Co-Director: Richard B. Gomez, Research Professor
Computational and Data Sciences

Spring Semester 2008
George Mason University
Fairfax, VA

Copyright 2008 Mark A. Z. Salvador
All Rights Reserved

DEDICATION

This is dedicated to my loving wife Rebecca and to my mother and father, Concepcion Juan Zarate-Salvador MD, and Oscar Lapitan Salvador DMD.

“do something”

ACKNOWLEDGMENTS

I would like to thank my all of my committee members for their support and understanding throughout my research and dissertation work. I would also like to thank all of my colleagues from several organizations who have supported and inspired me over the years to complete the PhD program. They include: Dr. Stefanie Tompkins, Dr. Jessica Sunshine, Dr. Kerry Patterson, Mr. Gene Hart, Dr. Greg Hyslop, Dr. Greg Vernon, Dr. Lowell Wilkins, Dr. James Sweet, Dr. Julie Rosen, Dr. Guido Cervone, Dr. George Keramidas, CAPT. Robert N. Friedman, Mr. Nick Cassimatis.

Thanks also to Mr. David Williams at the EPA for providing the AHI data set used in this research.

Finally, a special note of thanks to Mr. Ambrose Lewis, who as my Deputy Division Manager and friend has continually supported my efforts.

TABLE OF CONTENTS

	Page
List of Tables.....	vii
List of Figures.....	viii
List of Acronyms.....	xii
Abstract.....	xiii
1 Introduction.....	1
1.1 Hyperspectral Imagery Data.....	6
1.2 The Atmosphere and Radiative Transfer Equation.....	11
1.3 Radiance Model for Gas Detection.....	21
1.4 Hyperspectral Detection Methods.....	23
1.5 Spectral Library.....	28
1.6 Hyperspectral Sensors.....	32
1.7 The Wavelet and Wavelet Packet Transforms.....	35
1.8 Computational Challenges.....	46
1.9 Research Goals.....	48
2 Methods.....	50
2.1 Data Collection.....	52
2.2 Descriptions of Simulated Data Cubes.....	54
2.3 Description of Airborne Hyperspectral Imager Data.....	57
2.4 Description of Advanced Infrared Sounder Data.....	60
2.5 Data Pre-Processing.....	63
2.6 Data Processing and Detection.....	67
2.7 Wavelet Selection.....	72
2.8 Cost Function for Best Basis.....	78
2.9 Products for Comparison and Analysis.....	80
2.10 Dissemination of AIRS Detections.....	82
2.11 Computational Environment.....	85
3 Results.....	86
3.1 Simulated Data Cube Results.....	86
3.2 Airborne Hyperspectral Imager Results.....	117
3.3 Advanced Infrared Sounder Data.....	130
4 Discussion.....	141
4.1 Construction of Simulated Data Cubes.....	141
4.2 Wavelet Selection Process.....	145
4.3 Observations from the Simulated Data Cubes.....	148

4.4	Comparison of WPS Approach to SAM and CMF	149
4.5	Analysis of the AHI Flightlines.....	152
4.6	Validation of Atmospheric SO ₂ Detection and Identification in AIRS Data..	162
4.7	Computational Improvements and Issues	169
4.8	Image Context and Visualization Aspects	171
5	Conclusions	172
5.1	Wavelet Packet Subspace for Spectral Trace Gas Detection	172
5.2	Computational Improvements	173
5.3	Context & Visualization.....	174
5.4	Future Research	174
	Appendix 1. Previous Wavelet Based Hyperspectral Analysis.....	176
	Appendix 2. Simulated Data Cube Images and Spectra	187
	Appendix 3. Additional Wavelet Selection Results	193
	Appendix 4. Additional AIRS Results.....	200
	Appendix 5. Detection and Regression Libraries.....	202

LIST OF TABLES

Table	Page
Table 1. Major Constituents of the Atmosphere	12
Table 2. Order of Magnitude Contributions to Total Radiance	19
Table 3. Gases Selected for Study.....	29
Table 4. Example Values for Dyadic Intervals.....	37
Table 5. Simulated Data Cube Characteristics.....	53
Table 6. Target Gases Used in Wavelet Selection	76
Table 7. Summary of Visualization and Comparison Methods	81
Table 8. Pragmatic Wavelet Selection Results	147
Table 9. SDC 4A Wavelet Selection Results.....	147

LIST OF FIGURES

Figure	Page
Figure 1. The Spaces of Hyperspectral Analysis	2
Figure 2. Black & White, Color, and Spectral Images	7
Figure 3. A Typical Image Collection Scenario	9
Figure 4. The Hyperspectral Data Cube	11
Figure 5. Simplified Model of the Atmosphere	13
Figure 6. Blackbody Radiation at Sun and Earth Temperatures	16
Figure 7. Solar Radiance Contribution to the Schott's Big Equation	17
Figure 8. Thermal Radiance Contribution to Schott's Big Equation	18
Figure 9. PNNL Target Spectra (AHI and AIRS Wavebands)	31
Figure 10. AHI Hyperspectral Image (Left) and Linescan Image (Right)	33
Figure 11. Illustration of Dyadic Intervals	37
Figure 12. Tree Diagram of the DWT	40
Figure 13. Tree Structure for the Discrete Wavelet Packet Transform	41
Figure 14. The Roadmap of the Data Analysis Process	51
Figure 15. Simulated Data Cube Construction Process	53
Figure 16. Simulated Data Cube 3	57
Figure 17. AHI Data Collection Process	58
Figure 18. AHI Flightlines from Plume Ex	59
Figure 19. AIRS Broadband Radiance Image, Data Granule 169, July 13, 2003	60
Figure 20. AIRS VNIR Processed Image	61
Figure 21. AIRS Manual Data Retrieval Process	62
Figure 22. AIRS Automated Data Retrieval Process	63
Figure 23. MODTRAN Generated Atmosphere Applied to Simulated Data Cubes	64
Figure 24. SDC and AHI Pre-Processing Steps	65
Figure 25. AIRS Pre-Processing Steps	66
Figure 26. Spectral Angle Mapper Process	68
Figure 27. Clutter Matched Filter Process	69
Figure 28. Wavelet Packet Algorithm Approach	70
Figure 29. Wavelet Coefficient Statistics for AHI Wavelengths	75
Figure 30. Wavelet Coefficient Mode Image for AHI Wavelengths	76
Figure 31. Wavelet Coefficient Statistics for AIRS Wavelengths	77
Figure 32. Wavelet Coefficient Mode Image for AIRS Wavelengths	78
Figure 33. Comparison of Spectral and WPS Coefficient Space	80
Figure 34. AIRS Automated Classification Map	82

Figure 35. AIRS Automated Mapping Process.....	83
Figure 36. AIRS Automated KML Process.....	84
Figure 37. AIRS SO ₂ , CH ₄ , NH ₃ CMF Detections in Google Earth.....	84
Figure 38. SO ₂ Detection Map for SDC 1	87
Figure 39. Histogram for SDC 1	88
Figure 40. Scatter Plot for SDC 1	88
Figure 41. WPS ROC Curve for SDC 1	89
Figure 42. Detection Map for SDC 2 (SO ₂).....	90
Figure 43. WPS Classification Map for SDC 2 (SO ₂).....	90
Figure 44. Histogram for SDC 2 (SO ₂)	91
Figure 45. Scatter Plot for SDC 2 (SO ₂).....	91
Figure 46. WPS ROC Curve for SDC 2 (SO ₂)	92
Figure 47. SAM ROC Curve for SDC 2 (SO ₂).....	92
Figure 48. CMF ROC Curve for SDC 2 (SO ₂)	93
Figure 49. Detection Map for SDC 2A (SO ₂).....	94
Figure 50. Classification Map for SDC 2A (SO ₂).....	94
Figure 51. Histogram for SDC 2A (SO ₂)	95
Figure 52. Scatterplot for SDC 2A (SO ₂).....	95
Figure 53. WPS ROC Curve for SDC 2A (SO ₂).....	96
Figure 54. Detection Map for SDC 3 (SO ₂).....	97
Figure 55. WPS Classification Map for SDC 3 (SO ₂).....	97
Figure 56. Histogram for SDC 3 (SO ₂)	98
Figure 57. Scatterplot for SDC 3 (SO ₂).....	98
Figure 58. WPS ROC Curve for SDC 3 (SO ₂)	99
Figure 59. Detection Map for SDC 3 (CH ₄)	99
Figure 60. WPS Classification Map for SDC 3 (CH ₄).....	100
Figure 61. Histogram for SDC 3 (CH ₄).....	100
Figure 62. Scatterplot for SDC 3 (CH ₄)	101
Figure 63. WPS ROC Curve for SDC 3 (CH ₄).....	101
Figure 64. SAM ROC Curve for SDC 3 (CH ₄).....	102
Figure 65. CMF ROC Curve for SDC 3 (CH ₄).....	102
Figure 66. Detection Map for SDC 4 (SO ₂).....	103
Figure 67. WPS Classification Map for SDC 4 (SO ₂).....	104
Figure 68. Histogram for SDC 4 (SO ₂)	104
Figure 69. Scatterplot for SDC 4 (SO ₂).....	105
Figure 70. WPS ROC Curve for SDC 4 (SO ₂)	105
Figure 71. Detection Map for SDC 4 (CH ₄)	106
Figure 72. Classification Map for SDC 4 (CH ₄).....	106
Figure 73. Histogram for SDC 4 (CH ₄).....	107
Figure 74. Scatterplot for SDC 4 (CH ₄)	107
Figure 75. WPS ROC Curve for SDC 4 (CH ₄).....	108
Figure 76. SAM ROC Curve for SDC 4 (CH ₄).....	108
Figure 77. CMF ROC Curve for SDC 4 (CH ₄).....	109
Figure 78. Detection Map for SDC 4A (SO ₂).....	110

Figure 79. WPS Classification Map for SDC 4A (SO ₂).....	110
Figure 80. Histogram for SDC 4A (SO ₂)	111
Figure 81. Scatterplot for SDC 4A (SO ₂).....	111
Figure 82. DWPT ROC Curve for SDC 4A (SO ₂).....	112
Figure 83. SAM ROC Curve for SDC 4A (SO ₂).....	112
Figure 84. CMF ROC Curve for SDC 4A (SO ₂)	113
Figure 85. Detection Map for SDC 4A (CH ₄)	113
Figure 86. WPS Classification Map for SDC 4A (CH ₄)	114
Figure 87. Histogram for SDC 4A (CH ₄).....	114
Figure 88. Scatterplot for SDC 4A (CH ₄).....	115
Figure 89. WPS ROC Curve for SDC 4A (CH ₄).....	115
Figure 90. SAM ROC Curve for SDC 4A (CH ₄).....	116
Figure 91. CMF ROC Curve for SDC 4A (CH ₄).....	116
Figure 92. SAM Detection Planes for AHI 121247.....	119
Figure 93. CMF Detection Planes for AHI 121247	119
Figure 94. WPS Detection Planes for AHI 121247	120
Figure 95. Selected Detection Planes for AHI 121247 (SO ₂).....	120
Figure 96. Scatterplot for AHI 121247.....	121
Figure 97. SAM Detection Planes for AHI 105948.....	123
Figure 98. CMF Detection Planes for AHI 105948.....	124
Figure 99. WPS Detection Planes for AHI 105948	125
Figure 100. Selected Detection Planes for AHI 105948 (Benzene).....	125
Figure 101. Scatterplot for AHI 105948.....	126
Figure 102. SAM Detection Planes for AHI 153924	127
Figure 103. CMF Detection Planes for AHI 153924	128
Figure 104. WPS Detection Planes for AHI 153924.....	128
Figure 105. Selected Detection Planes for AHI 153924 (Benzene).....	129
Figure 106. Scatterplot for AHI 153924.....	129
Figure 107. Detection map for AIRS Granule 169 (SO ₂).....	131
Figure 108. WPS Classification Map for AIRS Granule 169 (SO ₂).....	132
Figure 109. SAM Classification Map for AIRS Granule 169 (SO ₂).....	132
Figure 110. CMF Classification Map for AIRS Granule 169 (SO ₂).....	133
Figure 111. Scatterplot for AIRS Granule 169 (SO ₂)	133
Figure 112. Histogram for AIRS Granule 169 (SO ₂).....	134
Figure 113. SO ₂ Spectral Comparison for AIRS 169.....	134
Figure 114. Georegistered Image of AIRS 169	135
Figure 115. Detection map for AIRS Granule 102 (SO ₂).....	136
Figure 116. Detection map for AIRS Granule 107 (SO ₂).....	137
Figure 117. WPS Classification Map for AIRS Granule 102 (SO ₂).....	137
Figure 118. WPS Classification Map for AIRS Granule 107 (SO ₂).....	138
Figure 119. Scatterplot for AIRS Granule 102 (SO ₂)	138
Figure 120. Histogram for AIRS Granule 102 (SO ₂).....	139
Figure 121. Georegistered Image for AIRS 102	139
Figure 122. Georegistered Image for AIRS 107	140

Figure 123. Variation of Target Signature (Benzene) Caused by Thermal Contrast.....	144
Figure 124. Comparison of SO ₂ Detections FL121247.....	154
Figure 125. FL 121247 SO ₂ Plume Classification Map	155
Figure 126. FL 105948 WPS Plume Detection.....	157
Figure 127. Google Earth Image for FL 105948.....	157
Figure 128. FL 105948 Spectral Identification of CH ₄	158
Figure 129. FL 153924 WPS Plume Detection.....	160
Figure 130. FL 153924 Spectral Identification of Benzene	161
Figure 131. AIRS Data Granule 169 SO ₂ Detection Comparison	164
Figure 132. AIRS Data Granule 102 SO ₂ Detection Comparison	165
Figure 133. AIRS Data Granule 107 SO ₂ Detection Comparison	166
Figure 134. Pixel Spectra for AIRS SO ₂ Detections	168

LIST OF ACRONYMS

AHI	Airborne Hyperspectral Imager
AIRS	Advanced Infrared Sounder
C ₃ H ₈	Propane
C ₄ H ₁₀	Butane
C ₆ H ₆	Benzene
CH ₄	Methane
CAFO	Combined Animal Feeding Operation
cm ⁻¹	Inverse centimeter or Wavenumber
CMF	Clutter Matched Filter
DCT	Discrete Cosine Transform
DWT	Discrete Wavelet Transform
DWPT	Discrete Wavelet Packet Transform
EPP	Easily Parallelizable Problem
FFT	Fast Fourier Transform
FOV	Field of View
HDF	Hierarchical Data Format
KML	Keyhole Markup Language
MF	Matched Filter
μm	Micrometer or micron
HSI	Hyperspectral Imagery
LBLRTM	Line-by-Line Radiative Transfer Model
LWIR	Long-Wave Infrared
MODTRAN	Moderate Resolution Radiative Transfer Model
MWIR	Mid-Wave Infrared
NASA	National Aeronautics and Space Administration
NH ₃	Ammonia
PNNL	Pacific Northwest National Lab
SAM	Spectral Angle Mapper
S/C	Signal to Clutter Ratio
SDC	Simulated Data Cube
SMF	Spectral Matched Filter
SO ₂	Sulfur Dioxide
TOMS	Total Ozone Mapping Spectrometer
WPS	Wavelet Packet Subspace

ABSTRACT

TRACE GAS DETECTION IN HYPERSPECTRAL IMAGERY USING THE WAVELET PACKET SUBSPACE

Mark A. Z. Salvador, PhD.

George Mason University, 2008

Dissertation Co-Director: Dr. Ronald G. Resmini

Dissertation Co-Director: Dr. Richard B. Gomez

This dissertation describes research into a new remote sensing method to detect trace gases in hyperspectral and ultra-spectral data. This new method is based on the wavelet packet transform. It attempts to improve both the computational tractability and the detection of trace gases in airborne and spaceborne spectral imagery. Atmospheric trace gas research supports various Earth science disciplines to include climatology, vulcanology, pollution monitoring, natural disasters, and intelligence and military applications. Hyperspectral and ultra-spectral data significantly increases the data glut of existing Earth science data sets. Spaceborne spectral data in particular significantly increases spectral resolution while performing daily global collections of the earth. Application of the wavelet packet transform to the spectral space of hyperspectral and ultra-spectral imagery data potentially improves remote sensing detection algorithms. It

also facilitates the parallelization of these methods for high performance computing. This research seeks two science goals, 1) developing a new spectral imagery detection algorithm, and 2) facilitating the parallelization of trace gas detection in spectral imagery data.

1 Introduction

Traditional hyperspectral analysis has followed two paths, statistical and non-statistical. Statistical approaches use the population data of the hyperspectral cube through examination of the covariance matrix and eigenvalues. Once these statistics are known, a comparison is made to either other populations within the cube or a reference library. Non-statistical approaches examine individual pixels to identify their properties either in relation to other pixels within the hyperspectral cube or in relation to a reference library. This type of approach follows traditional spectroscopy where materials in an individual spectrum are identified and quantified. This information is used to identify and quantify materials in the scene. These approaches lie within three so-called spaces in which analysis is performed [1]. These spaces include Image Space, which represents the image plane or x and y spatial dimensions of the hyperspectral image, Spectral Space, which represents the wavelength or frequency dimension along the z-axis of the hyperspectral cube, or Feature Space, which represents the relationships between each or all of the wavelengths or bands within the z-axis of the hyperspectral cube. Image space contains literal information which is information that is visually interpreted by the user. These spaces are illustrated in Figure 1. Feature space and spectral space are considered non-literal information. Data in those spaces must be processed using various mathematical methods. These methods elicit additional information within the scene.

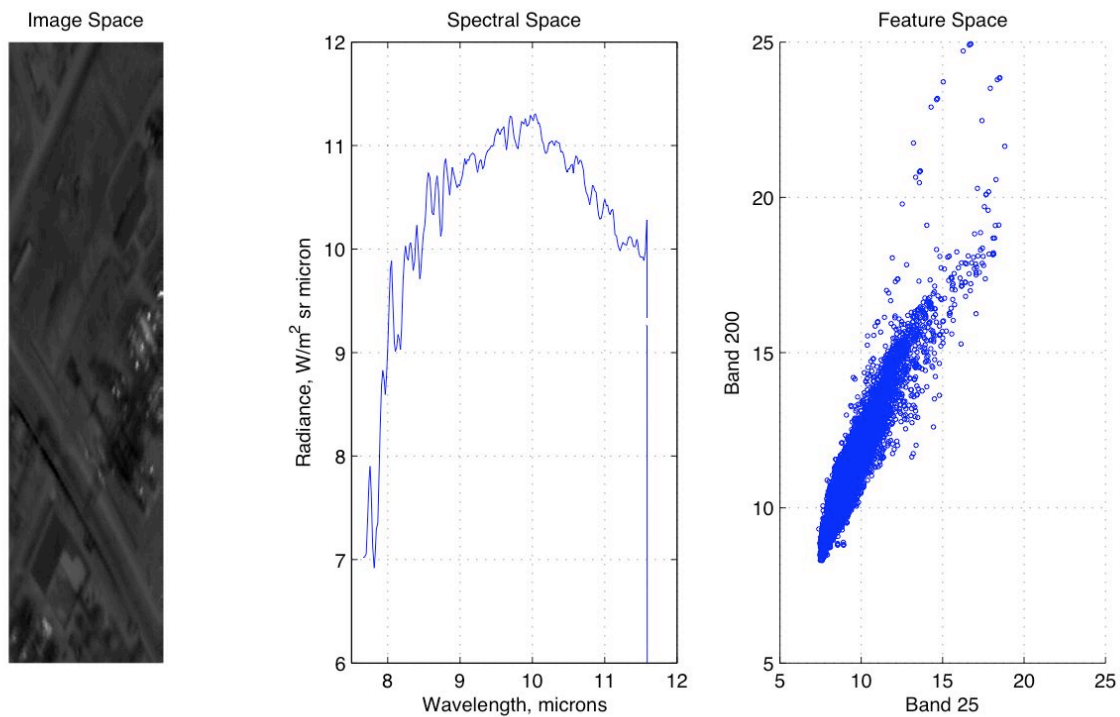


Figure 1. The Spaces of Hyperspectral Analysis

The general goals of hyperspectral analyses are relatively simple; 1) characterize or identify all of the materials either solid, liquid, or gas within the hyperspectral image, or 2) identify specific target materials of interest, which also maybe solid, liquid, or gas, within the hyperspectral image. The study of all the materials in the scene, or endmember analysis, is used in several scientific research areas; geology, geography, planetary science, earth science, atmospheric science, etc. The search for specific target materials may be thought of as a subset of this study, which has somewhat broader implications but a narrower focus in environmental monitoring, atmospheric science, planetary science, military applications, intelligence gathering, and many other research areas [2-5].

Several airborne and spaceborne platforms exist that collect spectral imagery data for both civil and military purposes. As more and more spectral imagers become available, data processing requirements shift from analysis of individual data collections to daily processing of global data. With the advent of NASA's Advanced Infrared Sounder (AIRS) on board the AQUA satellite, ultra-spectral data (2378 bands with $\lambda/\Delta\lambda \sim 1200$) increased by at least an order of magnitude the complexity and computational requirements on hyperspectral analysis [6]. This requires computationally tractable remote sensing methods to process data quickly for material detection and identification.

Typically hyperspectral analysis focuses on a single data set or a small number of data sets. This type of analysis seeks to find specific materials, properties, or characteristics, within the data sets of interest. Prior to Hyperion, on board NASA's EO-1 satellite, or Advanced Infrared Sounder (AIRS), on-board NASA's AQUA satellite, hyperspectral data were generally collected with an airborne platform. Airborne platforms have a much smaller data volume as their collections are limited to individual flights. AIRS covers the entire globe twice a day. With AIRS data, global atmospheric study is possible. AIRS data facilitates research in many areas providing information on both global and local scales. AIRS is as equally suited to measuring total atmospheric ozone across the globe as it is trace amounts of sulfur dioxide from regional volcanic events. It is easy to see the vast number of earth science applications for which AIRS data are suitable.

With this daily global data set comes an equally substantial data volume. This large data volume is on the order of 240 data sets/day or 30 GB/day of AIRS radiance data.

Sheer volume alone is not the only computational issue associated with AIRS data [7; 8]. The AIRS instrument has a spectral resolution of 1200 ($\lambda/\Delta\lambda$). This allows it to see both vibrational and rotational features of gas spectra [9]. In typical hyperspectral analysis ($\lambda/\Delta\lambda \sim 10$ -100), atmospheric compensation could be accomplished using a moderate resolution radiative transfer model such as MODTRAN. AIRS, with its high spectral resolution, requires radiative transfer models using the HITRAN database [10] and line-by-line radiative transfer models. This increases computational complexity of data.

If we agree that the computational science need for improved detection and identification for AIRS ultra-spectral data is clear, then we can focus on the earth science question of atmospheric trace gas detection and identification. Trace gases represent less than 1% of total content of the atmosphere [11] but are primarily responsible for such varied effects as climate change (CO_2), air pollution (chlorofluorocarbons, CFCs), protection from ultra-violet rays (O_3), and weather (H_2O). Detection, identification, and quantification of these trace gases are critical to the study of earth science [12; 13].

Most hyperspectral analysis can be divided between statistical methods in feature space and distance methods in spectral space. The current statistical methods do not lend themselves easily to parallelization, but spectral space methods are more accommodating. Unfortunately the statistical methods demonstrate better results in weak signal detection [14-17]. Thus methods applied to spectral space that provides performance as good as or better than statistical methods are much needed.

Several efforts have been made at applying the wavelet transform to hyperspectral analysis [18-24]. Applying a wavelet transform to spectral data follows the discussion of

[25] by applying a transformation to data to generate features for better pattern recognition. The fundamental problem of identifying a material (solid, liquid, or gas) within a spectral image is that of matching the pixel spectrum to the material signature from a reference library, ergo, pattern matching.

The wavelet transform has many properties that could be applied to hyperspectral data analysis. A fundamental property is the multiple resolution signal analysis that can be applied [26]. These multiple resolutions or scales separate the fine scale and large scale properties of the input signal. Thus, the wavelet transform used in spectral space is the basis for a new ultra-spectral remote sensing detection and identification method.

Preliminary results (see Appendix 1) for the development of a wavelet transform based detection and identification algorithm are promising. Initial applications of the Discrete Wavelet Transform (DWT) compare well with traditional hyperspectral methods. This is an expected result. Several researchers have already published their findings regarding classification improvements with the DWT [20; 27; 28].

The application of the DWT and the Wavelet Packet Transform (WPT) are somewhat analogous the principal component analysis (PCA or Karhunen-Loeve basis) approach to multispectral and hyperspectral feature detection. The most significant difference is the Karhunen-Loeve optimizes through finding the axes of most variance in the observed data, while a basis from a wavelet packet tree can be selected based on other cost functions [29].

The goal of this research is to develop a method for trace gas detection in spectral data using a wavelet packet best basis. Hereinafter referred to as the “Wavelet Packet

Subspace” (WPS) approach. By applying the wavelet packet transform in spectral space, the statistics of the entire data set need not be calculated, and the WPS approach would become an easily parallelizable problem (EPP). In this application, we seek to identify the most significant contributions of the target signal in wavelet coefficient space. Although we must also bear in mind the issues of using too few or too many bands in attempting to separate the target and the background.

The presented treatise is organized as follows. Following this introduction, which includes background information on hyperspectral analysis, a description of the atmosphere, various spectral detection methods, spectral libraries, and a discussion on wavelet and wavelet packet transforms. It then describes the methods used, which explain the approach to the data, wavelet packet selection, and the methods used to compare and visualize the results. The results are presented. Then, results for all of the data sets are analyzed. A discussion section follows, which provides key observations of the results and their implications to both the algorithm and its computational tractability. Finally, the conclusion discusses the results in regard to the original intent of the research and future research.

1.1 Hyperspectral Imagery Data

Spectral imaging for remote sensing arose as an alternative to high resolution, large aperture satellite systems [3]. Spectral imaging is an extension of typical visual imagery, like black and white or color imagery, into a larger portion of the electro-magnetic

spectrum. It uses the visual through long wave infrared (LWIR) portions of the electromagnetic (E-M) spectrum and measures several to hundreds of bands throughout the spectrum.

As a color image is a measure of radiance in three bands (red, green, and blue), a spectral image may contain several to hundreds of contiguous bands within the red, green, and blue portions of the spectrum. The spectral image may also extend through the visual spectrum into the mid-wave infrared (MWIR) and into the long-wave infrared (LWIR). Figure 2 illustrates this concept.

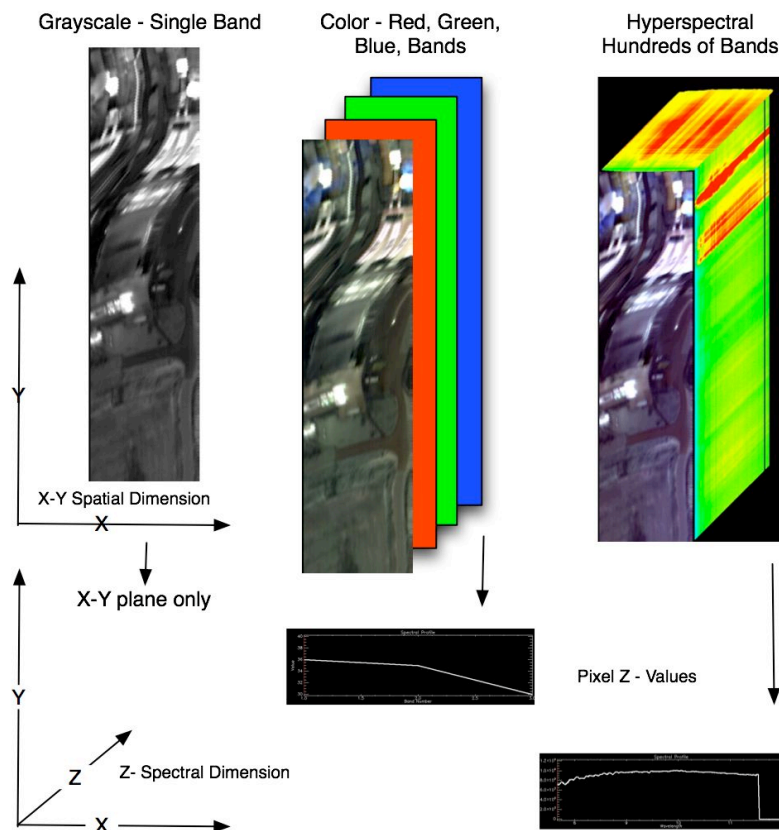


Figure 2. Black & White, Color, and Spectral Images

Hyperspectral data analysis is generally broken into three major segments of the electro-magnetic spectrum, the VNIR/SWIR (also called solar reflective) region $\sim 0.4 \mu\text{m}$ to $\sim 2.5 \mu\text{m}$, the mid-wave region $\sim 2.5 \mu\text{m}$ to $\sim 7.0 \mu\text{m}$, and long-wave (also called thermal) region $\sim 7.0 \mu\text{m}$ to $\sim 15.0 \mu\text{m}$. Within the reflective region the primary illuminating source of photons is the Sun. This is the visual portion of the E-M spectrum and extends into the near infrared. In the long-wave region, illumination from the Sun is negligible and the primary source of photons is self-emission from objects above zero Kelvin. Across the mid-wave region, both the Sun and self-emission contribute to the collected radiance and thus both sources of photons must be dealt with during analysis of the data. In addition to the particular phenomenologies of the source radiance, sensor design is based on materials that are sensitive to different parts of the E-M spectrum, e.g. Mercury-Cadmium-Telluride (HgCdTe). HgCdTe is a material used in sensor focal plane arrays. It is sensitive to long-wave radiation but significantly less sensitive to photons in the $\sim 0.4 \mu\text{m}$ to $\sim 2.5 \mu\text{m}$ region [30].

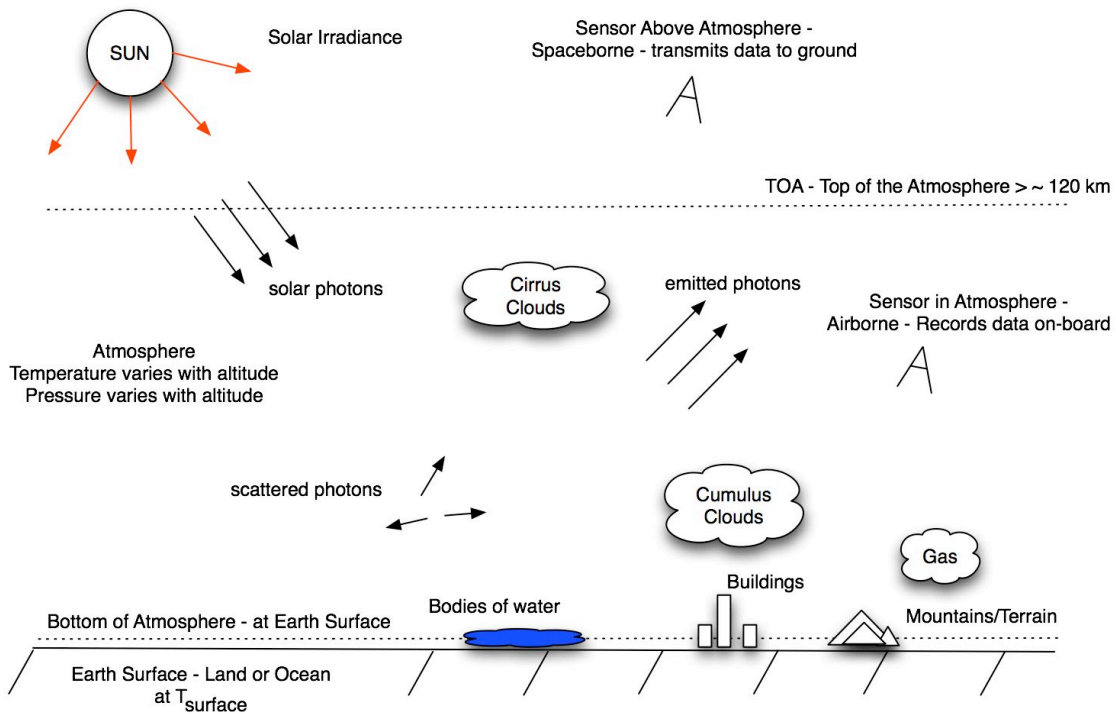


Figure 3. A Typical Image Collection Scenario

The model for a typical hyperspectral remote sensing collection scenario follows many other imaging systems as illustrated in Figure 3. There is a source, which provides the illuminating photons. Generally the Sun is the source for reflective hyperspectral collection. Radiant energy travels through the atmosphere, is reflected off of some object on the ground, and again travels through the atmosphere, and is then collected by the sensor. This radiance model will be discussed in more detail later. The sensor collects the received photons in both the spatial and spectral domains and records them.

Hyperspectral data are generally thought of as a data cube, with spatial information across the x-y plane and spectral information across the z-axis. Thus any plane in x-y corresponds to a wavelength along the z-axis. Each pixel within the x-y plane has a

corresponding spectrum along the z-axis (Figure 4). This brings us to the concept of literal and non-literal information [31]. In the x-y plane of the data cube we have information, which to any individual conveys visual information, an image. This image is literal information that conveys some context to us, i.e. the image of an Abrams M-1 main battle tank. The image is interpretable as a tank in a field or in a parking lot, etc. Along the z-axis of the data we have individual pixel spectra. These spectra are considered non-literal information as the data are not necessarily visually interpretable but must be processed before information can be interpreted.

This research focuses on exploiting non-literal information within the long-wave region of the E-M spectrum. Thus here, the sensors, radiance model, and data cubes used will correspond to the $\sim 7.0 \mu\text{m}$ to $\sim 12.0 \mu\text{m}$ region of the spectrum.

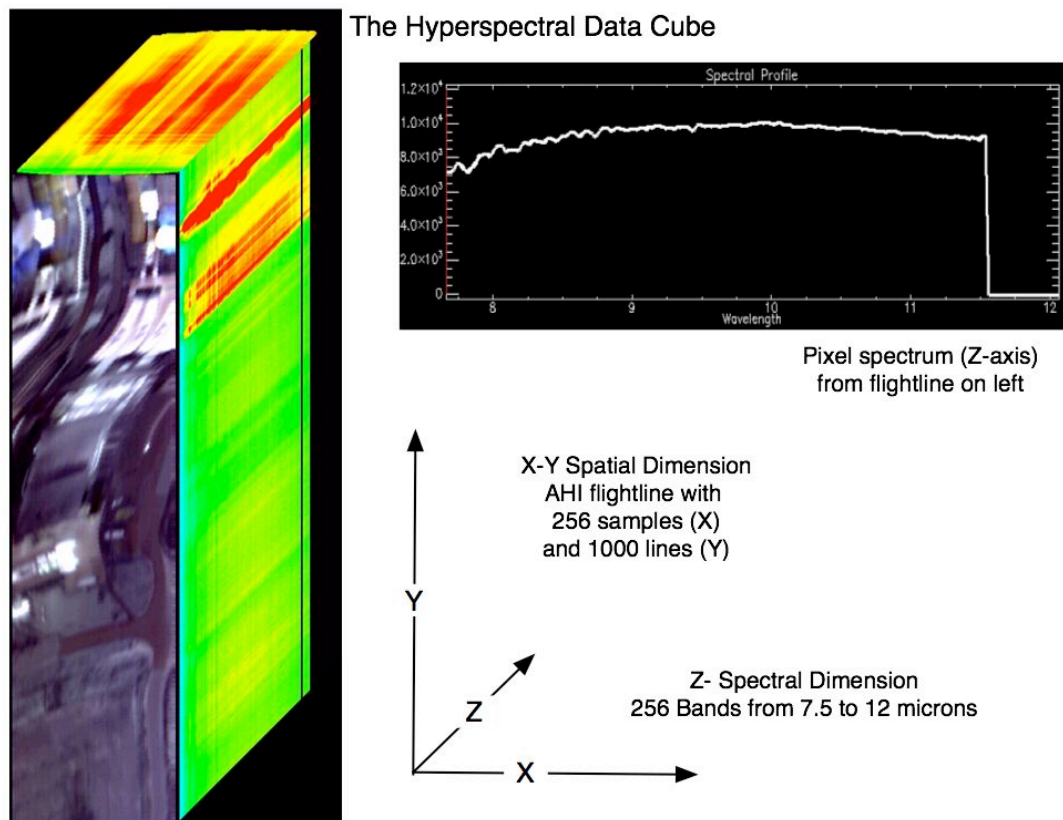


Figure 4. The Hyperspectral Data Cube

1.2 The Atmosphere and Radiative Transfer Equation

The atmosphere plays a significant role in the hyperspectral detection on any material whether the data collected is in the reflective or LWIR regions. The atmosphere is the medium that photons from the source travel to the sensor. The photons collected by the sensor have been attenuated and scattered by the atmosphere. The atmosphere itself is a mixture of ideal gases in which nitrogen and oxygen predominate by volume. Photons from the Sun in the visible, infrared, and ultra-violet wavelengths heat the atmosphere and are either absorbed by the atmosphere or are reflected back into space. The structure

of the atmosphere is stably stratified but varies in both pressure and temperature as altitude increases [11]. Table 1 lists major constituents of the atmosphere.

Table 1. Major Constituents of the Atmosphere¹

<i>Gas</i>	<i>Abundance</i>	<i>Infrared Active</i>
Nitrogen	78%	no
Oxygen	21%	yes
H ₂ O	< 4%	yes
CO ₂	350 ppm	yes
Ar	9340 ppm	no
Ne	18 ppm	no
He	5.2 ppm	no
CH ₄	1.7 ppm	yes
N ₂ O	320 ppb	yes
CO	125 ppb	yes
O ₃	10-100 ppb	yes

Given the structure of the atmosphere, the amount of absorption and transmission through that portion or layer of the atmosphere can vary significantly. There are many texts on the physics of the atmosphere [5; 11; 12; 32]. For our purposes a simplified model of the atmosphere is as shown in Figure 5. This helps to illustrate its impacts on hyperspectral data collection and analysis.

¹ Adapted from Yung & Demore (1999), pgs 318-319

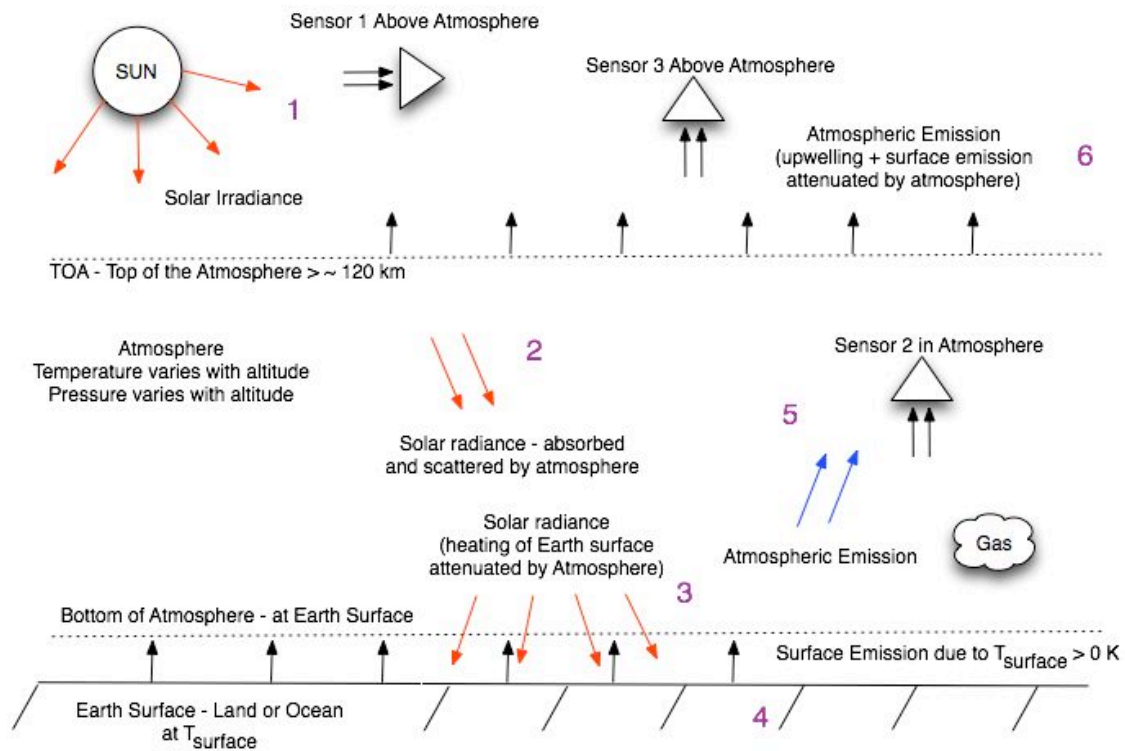


Figure 5. Simplified Model of the Atmosphere

Following Figure 5, at (1), the Sun's irradiance at 6000 K is close to a blackbody at the same temperature, (see Figure 6). The solar irradiance is incident on the Earth's atmosphere. At (2), its photons are absorbed and scattered by the atmosphere subsequently heating the atmosphere. Bear in mind since the atmosphere varies in pressure with altitude, the amount of absorption and scattering varies, and the subsequent heating of the atmosphere also varies.

Eventually the solar radiance (3), which has been attenuated by the atmosphere, is incident on the Earth's surface. It reflects off the surface and is also absorbed by the surface, thus, heating the surface. The Earth (4), is heated by the Sun. It also has internal heat generated by the geophysics of the planet. The surface then emits radiation back into

the atmosphere where it is either absorbed or transmitted through. The atmosphere (5) is emitting due to its temperature. It also contains the surface radiance and reflected solar radiance as attenuated by the atmosphere. Finally, the radiance at the top of the atmosphere (6) contains, reflected solar radiance, emission from the atmosphere and emission from the surface, all of which have been attenuated.

We will return to this model later when we discuss trace gas detection. It is important to note from Table 1 that not all gases are infrared active. That is, they do not absorb nor emit photons in the infrared region of the spectrum and thus do not contribute to the emission from the atmosphere.

Quantifying the radiance from Figure 5 brings us to the radiance equation used for hyperspectral detection. We will describe the Big Equation [33] with some modifications in notation. Starting with the solar irradiance we describe both the visible/near infrared components of the radiance as well as the long wave components. The total radiance at the sensor is

$$L_{total} = L_{reflective} + L_{thermal} \quad (1)$$

where L_{total} is total radiance at the sensor; $L_{reflective}$ is the solar irradiance contribution; and $L_{thermal}$ is the emissive radiance contribution.

The reflective contribution is the Solar irradiance reflected off objects into the sensor field of view (FOV) which occurs in the $\sim 0.4 \mu\text{m}$ to $\sim 5.0 \mu\text{m}$ range of the electromagnetic spectrum. The thermal contribution results from self-emission from the surface and the atmosphere within the sensor FOV. Self-emission is electromagnetic radiation

emitted by a body with a temperature greater than absolute zero. The spectral distribution of the radiation is governed by Planck's Law.

$$S(\lambda) = \frac{2\pi hc^2}{\lambda^5} \frac{1}{e^{\frac{hc}{\lambda kT}} - 1} \quad (2)$$

$S(\lambda)$	=	spectral radiant emittance
λ	=	wavelength
h	=	Planck's Constant
T	=	absolute temperature
c	=	velocity of light
k	=	Boltzman's constant

From Planck's Law, Wien's law can be derived which gives the wavelength of maximum emittance [34].

$$\lambda_{\max} = \frac{a}{T} \quad (3)$$

where $a = 2898 \mu\text{mK}$

For the surface of the Earth with a temperature ~ 300 K the peak emittance occurs at $\lambda_{\max} = 9.66 \mu\text{m}$ (LWIR). For the Sun, ~ 6000 K, peak emittance occurs at $\lambda_{\max} = 0.48 \mu\text{m}$, which is in the visible spectrum. Figure 6 illustrates the total radiance as a function of wavelength by Planck's Law for objects at 6000 K (center) and 300 K (right), also plotted is the Sun's irradiance at the top of the atmosphere (TOA) as calculated by MODTRAN (left).

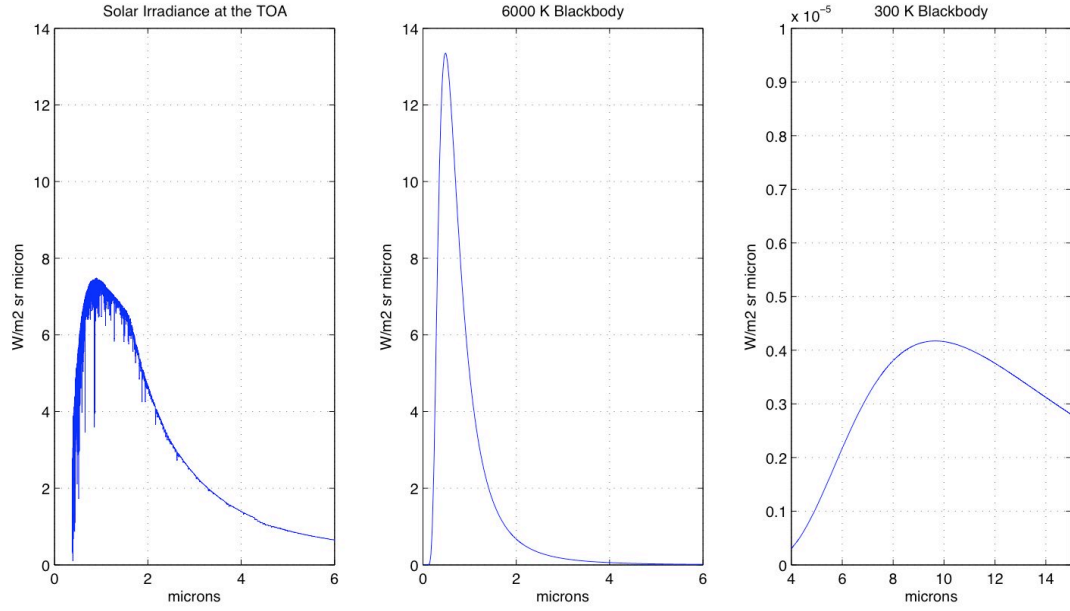


Figure 6. Blackbody Radiation at Sun and Earth Temperatures
(MODTRAN - left, 6000 K - center, 300 K - right)

The reflective contributions of total radiance is illustrated in Figure 7, and is described by equation (4).

$$L_{reflective} = L_1 + L_2 + L_3 + L_4 + L_5 \quad (4)$$

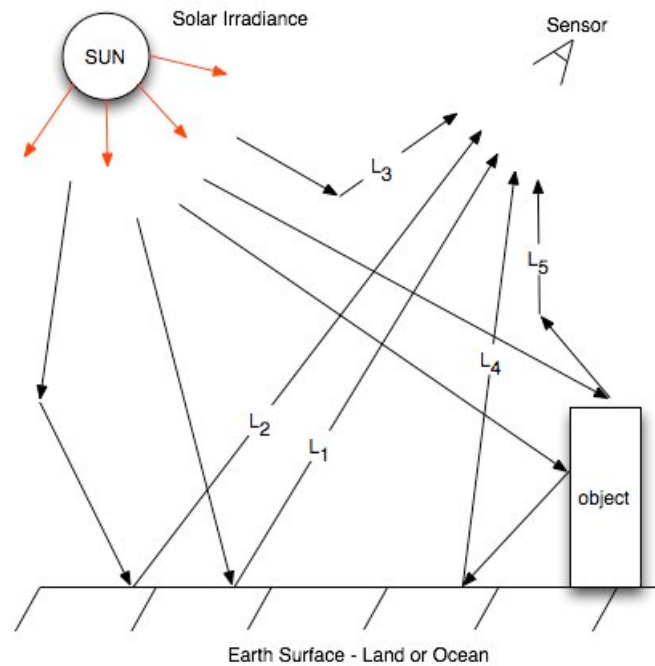


Figure 7. Solar Radiance Contribution to the Schott's Big Equation

L_1 represents photons from the Sun passing through the atmosphere and being reflected into the sensor. This is the most direct path between the Sun and the sensor. L_2 represents photons emitted from the Sun, scattered in the atmosphere and reflected off of the surface. They then travel through the atmosphere into the sensor. These photons are commonly referred to as sky shine. L_3 photons start at the sun and are directly reflected off the atmosphere. This is called upwelled radiance. L_4 and L_5 photons represent multiple bounce photons from background objects or scattered through the atmosphere and the surface.

Thermal energy radiance paths are illustrated in Figure 8. As evidenced in Figure 6, direct solar contribution to the LWIR is negligible. However, the Sun is primarily responsible for heating the atmosphere. L_8 in Figure 8, is this component.

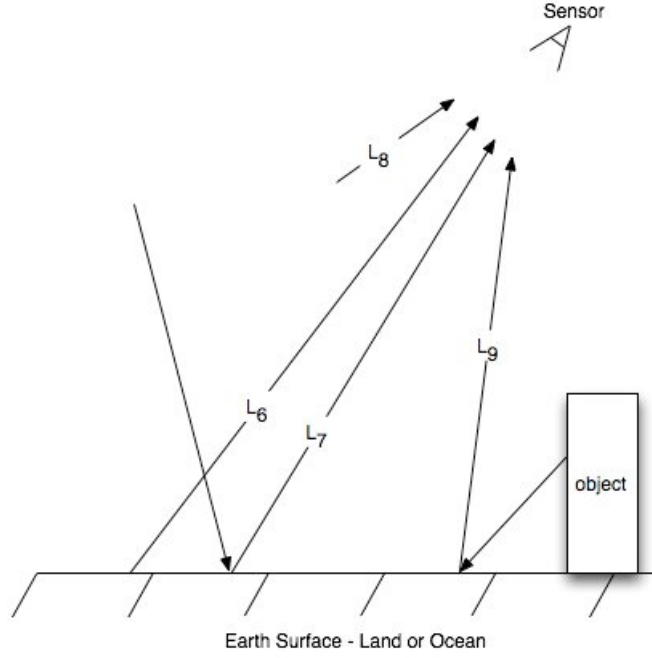


Figure 8. Thermal Radiance Contribution to Schott's Big Equation

$$L_{thermal} = L_6 + L_7 + L_8 + L_9 \quad (5)$$

L_6 photons are emitted directly from the surface and travel through the atmosphere into the sensor. L_7 photons are generated in the atmosphere and are reflected into the sensor off of the surface. This is called downwelled radiance. L_8 photons are emission from the atmosphere sent directly into the sensor. This is also upwelled radiance. Finally L_9 photons are emitted from bodies and scattered off multiple bodies prior to reaching the sensor. In the analysis of the Big Equation [33] a sensitivity study of the total radiance contributions for the visible through LWIR portions of the spectrum is performed. The order of magnitude results are summarized in Table 2.

Table 2. Order of Magnitude Contributions to Total Radiance²

<i>Wavelength Range</i>	<i>L₁</i>	<i>L₂</i>	<i>L₃</i>	<i>L₄</i>	<i>L₆</i>	<i>L₇</i>	<i>L₈</i>	<i>L₉</i>	<i>L_{total}</i>
0.4 – 0.7 μm	10 ⁻³	10 ⁻⁴	10 ⁻³	10 ⁻⁵	10 ⁻²⁷	10 ⁻³¹	10 ⁻²⁹	10 ⁻²⁹	10 ⁻³
3 – 5 μm	10 ⁻⁵	10 ⁻⁷	10 ⁻⁷	10 ⁻⁷	10 ⁻⁴	10 ⁻⁶	10 ⁻²⁹	10 ⁻²⁹	10 ⁻⁴
8 – 14 μm	10 ⁻⁶	10 ⁻⁸	10 ⁻⁸	10 ⁻⁸	10 ⁻³	10 ⁻⁵	10 ⁻³	10 ⁻⁴	10 ⁻³

As we will see shortly the hyperspectral data used for our trace gas detection spans ~7.5 μm to ~12 μm. This allows us to make some simplifying assumptions to the radiance equation based on the orders of magnitude. Within the 8 μm to 14 μm range, the $L_{reflective}$ components are at least three orders of magnitude less than the total radiance. Within the $L_{thermal}$ components, L_7 is at least an order of magnitude less than the total radiance. L_9 is only an order of magnitude less than the total radiance, but very difficult to estimate since it is based on multiple scattering from objects in the scene. For our purposes, this is also set to zero. We can then set $L_1 = L_2 = L_3 = L_4 = L_6 = L_7 = L_9 = 0$, and our total radiance equation becomes,

$$L_{total} = L_6 + L_8 \quad (6)$$

where,

$$L_6 = \varepsilon(\lambda)L_{T\lambda}\tau(\lambda) \quad (7)$$

$\varepsilon(\lambda)$ = wavelength dependent emissivity of the body
 $L_{T\lambda}$ = blackbody radiation from the object from Planck's Law
 $\tau(\lambda)$ = wavelength dependent transmission through the atmosphere

² Adapted from Schott (1997), pg. 117

$$L_8 = L_{u\lambda} \quad (8)$$

$L_{u\lambda}$ = wavelength dependent atmospheric upwelling radiance

Then equation (6) becomes

$$L_{total} = \varepsilon(\lambda)L_{T\lambda}\tau(\lambda) + L_{u\lambda} \quad (9)$$

In solving equation (9), $L_{T\lambda}$ is easily calculated using Planck's Law while $\varepsilon(\lambda)$ is estimated as an average emissivity across the wavelengths. The emissivity is dependent on the background of the spectral scene and is selected accordingly. The other two components, atmospheric upwelling and atmospheric transmission are modeled using a radiative transfer or atmospheric model.

$$L_{total} = \varepsilon(\lambda)L_{T\lambda}\tau(\lambda) + L_{u\lambda} + \varepsilon_{noise} \quad (10)$$

For completeness we add ε_{noise} to equation (9), which represents the noise or error associated with the collected radiance data at the sensor. Consideration of this residual error in the inversion process is a fundamental computational issue [35; 36]. This is especially true in solving remote sensing problems. Solution to the remote sensing inversion problem requires solving a Fredholm integral equation of the first kind [35]. It is a computational issue many hyperspectral analysts overlook.

Two different atmospheric models are used for our trace gas detection. For simulated data cubes and AHI data, the Moderate Resolution Transmission code (MODTRAN) is used. MODTRAN is a broadband atmospheric model that has become

the de-facto standard for hyperspectral analysis [37]. At a spectral resolution of 0.1 to 1 cm^{-1} , this model is more than adequate for the collected data.

For the AIRS data, the spectral resolution of the AIRS sensor is significantly greater than AHI and MODTRAN ($\lambda/\delta\lambda > 1200$ or better than 10^{-3} cm^{-1}). A line-by-line radiative transfer model (LBLRTM) from Atmospheric and Environmental Research Inc. is used. LBLRTM is the atmospheric model that other models and sensors are compared to [38]. LBLRTM is typically the model used for atmospheric studies.

1.3 Radiance Model for Gas Detection

As a hyperspectral sensor collects data over large areas, it is obvious that gas/plume targets of interest are sparsely represented in the data. This is also true in the case of gas detection. Although, in applying the matched filter there is an assumption that the spectral data mean and covariance matrix are not significantly influenced by the presence of the target [15]. If the target signature is spread across the scene such that the covariance matrix is significantly influenced by the target signature, poor results can occur. This is actually the case in some of the AHI data presented in Chapter 3.

The fundamental problem in detecting weak signatures in hyperspectral data is separating the background from the target. The background has usually been differentiated into general categories such as urban, forest, and desert areas [39]. The background itself has been and continues to be a fundamental issue in the design of

hyperspectral collections and experiments. As research has shown, backgrounds cannot be easily characterized and generalized [39; 40].

In this model, the background represents not only the physical background in the scene, but all of the radiance contributions that are not the target signature. Equation (10) is essentially the background radiance model. For a gas in the target scene, the model requires the addition of the target signature. This radiance model follows Young's discussion in his report on detection of gases from industrial stack plumes [41].

The description of the background, equation (10) is equivalent to Young's equation describing radiance without a plume present in the scene. Both models are in agreement with Schott's radiative transfer expression. Young follows by describing the addition of the plume to the radiance equation. Added to the background model are:

1. Surface radiance attenuated by the plume, and
2. Radiance contribution of the plume itself.

Young makes assumptions and simplifications to his model concluding with his equation for the radiance in the presence of a plume.

$$L_{plume} = \sum_{n=1}^N C_n(\lambda) \tau(\lambda) k_n(\lambda) + L_0(\lambda) \quad (11)$$

where,

$C_n(\lambda)$	=	Thermal contrast spectrum for the n th Gas
$\tau(\lambda)$	=	wavelength dependent transmission through the atmosphere
$k_n(\lambda)$	=	Spectral absorption coefficient for the n th gas
$L_0(\lambda)$	=	L_{total} from equation (10)

The thermal contrast spectrum is defined as

$$C_n(\lambda) \approx C_n = c_{nP} [B_P(\bar{\lambda}_n) - \varepsilon_s(\bar{\lambda}_n) B_S(\bar{\lambda}_n)] + c_{nA} [B_P(\bar{\lambda}_n) - B_A(\bar{\lambda}_n)] \quad (12)$$

where,

c_n	=	denotes plume and ambient column density
B	=	Planck law radiance for the plume (P), surface (S), ambient (A)
ε_s	=	surface emissivity
$\bar{\lambda}_n$	=	denotes the average quantity for the broadband LWIR spectrum

Equation (11) is used to generate all the simulated data cubes and is the basis for trace gas detection in Chapter 2.

1.4 Hyperspectral Detection Methods

There are numerous works, which describe hyperspectral data analysis and signature matching methods in great detail [14; 42-44]. Provided here is an overview of the analysis process and description of the detection and matching methods used. The hyperspectral analysis process is broken into 5 stages.

1. Collection – physical data collections, sometimes the actual flight campaign
2. Pre-Processing – preparing the data for analysis, e.g. calibration
3. Processing to include material, object, event detection and identification – the actual analysis step

4. Product Generation – taking the results of analysis and making them suitable for external users
5. Dissemination of products – physical or electronic distribution of products and data to other users

Most of the steps listed above are intuitive perhaps to the experienced spectral data analyst but a short summary of each step is provided. The process begins with collection. This is the actual airborne flight or spacecraft overflight that activates the hyperspectral sensor and collects radiance data from the scene. The data is stored on-board for either data transmission to a ground station or physical transfer to the ground processing site. Depending on the sensor and system design, the collected radiance data may be radiometrically, spectrally, and/or spatially calibrated during the collection step. These steps may also be conducted on the ground as part of the pre-processing step.

Preprocessing generally encompasses any preparation of the data, calibration, splitting of flight lines, quality control, and so on. It may also include atmospheric compensation of the data. Atmospheric compensation includes modeling of the atmosphere and removal of the atmospheric effects from the data so that matching of spectral features in the processing step can be accomplished. Processing includes any analysis of the data that transforms the at-sensor radiance data into a new set of information. This may include detection and classification maps, vegetation analysis, anomaly detections, terrain categorization, and background characterization. The WPS algorithm is focused on the processing phase of hyperspectral analysis.

The last two steps are straightforward. Product generation is the process to take the newly generated information and to create a standard product. This may be formation of another data set such as NASA's Level 4 and Level 5 products, formal briefing slides, a community standard product format etc. The last step, dissemination is the not-so-simple task of transmitting these products to end users. This may sound unimportant, but given the volume of hyperspectral data, and the potential volume of the products it may generate, this task is anything but trivial.

Within the processing step, detection and identification are sometimes treated synonymously. Detection and identification are merely the confident identification of a material (solid, liquid or gas) in the spectral data. Most often, detection is performed via visual selection of areas of interest. This is followed by a comparison of spectral features of the pixels in the area of interest, to a reference library. The spectral comparison method for gas detection is usually a linear least squares fit (linear regression) of the pixel spectra against the target reference library [41].

The more difficult step that of absolute quantification of a specific material in the scene is often treated separately. In many cases, it is sufficient to detect the materials confidently in the data. The focus of this research is successful detection. The quantification process is a well-researched area based on Beer's Law. Quantification is fundamental to classical spectroscopy [45; 46].

The basis for comparison to the wavelet packet subspace approach is dual pronged. Two often used matching methods are used to assess WPS performance. One method, the spectral angle mapper or SAM, is a whole-pixel based approach while the second

method, the spectral matched filter (MF) is a statistical method. The matched filter examines the 2nd order statistics (i.e., data covariance) of the data population in order to determine the optimal filter. The spectral angle is simply the angle between 2 vectors in n-dimensional space. It calculates the spectral similarity between a reference spectrum and a pixel spectrum [4].

$$\Theta = \cos^{-1} \left(\frac{\sum_{i=1}^n t_i r_i}{\sqrt{\sum_{i=1}^n t_i^2 \sum_{i=1}^n r_i^2}} \right) \quad (13)$$

t = in-scene pixel spectrum
 r = reference spectrum

Spectral angle will be used as baseline both in spectral space and in wavelet coefficient space. It measures the angle between the target (gas) spectrum and the in-scene data. SAM is chosen because it is insensitive to the brightness difference between two spectra [31].

The second matching method chosen is the matched filter. The matched filter is commonly used for weak target detection in hyperspectral data especially gas detection [14; 15; 17; 47]. The first matched filter was developed over fifty years ago and was applied to radar signal processing [48]. The simple matched filter is defined as

$$r = q^T x \quad (14)$$

where,

$$q = K^{-1}t \quad (15)$$

$$K = \left(\frac{1}{N} \sum_{i=1}^N (x_i - m)(x_i - m)^T \right) \quad (16)$$

x	=	in-scene pixel vector (spectra)
K	=	covariance matrix
t	=	target spectra (library reference spectra)
m	=	spectral data mean

K is derived from the data through equation (16), but the best estimate of the covariance matrix, is not necessarily the best estimate of the inverse covariance matrix [15]. Although K is an n -by- n matrix it is not necessarily easily invertible [49]. There are several methods to calculate matrix inverse or a pseudoinverse, such as the Moore-Penrose matrix inverse [50]. It is important to examine the matrix rank and singularity in determining the appropriate inverse method. It should be noted however that this is a significant computational issue as a direct matrix inverse as calculated by MatLab or IDL may lead to erroneous results. This is due to both machine precision and the matrix inverse algorithm.

The simple matched filter is adapted into the clutter matched filter (CMF) as defined by Funk & Theiler [15]. The CMF takes into account correlations between spectral channels and modifies equation (15).

$$q = \frac{K^{-1}t}{\sqrt{t^T K^{-1}t}} \quad (17)$$

To summarize, SAM will be used as pixel-by-pixel matching method in both spectral space and wavelet coefficient space. CMF will be used as the statistical method that operates in feature space.

1.5 Spectral Library

The spectral library used for this research is the Pacific Northwest National Labs (PNNL) infrared reference database. The database was chosen for three reasons. First, the database is designed for hyperspectral imaging and detection. Second, it is well known in the hyperspectral community and third, it is publicly available.

The database is a collection of nearly 500 gas phase species designed for atmospheric monitoring, remote sensing, and hyperspectral imaging. The spectra are collected at 5° C, 25° C and 50° C using Fourier Transform spectrometry. The database has specific requirements for spectral resolution, species purity, accuracy and precision. The database is funded by the United States Government and publicly available at <http://nwir.pnl.gov> [51].

Only four gases are chosen for study. This limits the scope of the simulated data sets and the analysis to be performed. The gases chosen were selected based on four criteria:

1. The gas must have some earth science significance as detected by the AIRS instrument.

2. The gas must be potentially present in the AHI data set.
3. The gas must be present in the publicly available Pacific Northwest National Labs Infrared Spectral Library.
4. The gas must have significant spectral features in both the AIRS and AHI wavebands.

The selection criteria are shown in Table 3. A short description of each gas follows. Each gas is plotted in Figure 9.

Table 3. Gases Selected for Study

Name	Chemical Composition	In PNNL	Location in AHI data	Earth Science Significance	Spectral Features
Sulfur Dioxide	SO ₂	Yes	Potentially in Fairfield – Power Plant	Significant in Volcanic Eruptions and Ash Plumes	For AHI feature at 8-9 μm while larger feature at 7.5 μm in AIRS
Methane	CH ₄	Yes	N/A	Significant greenhouse gas	Large features 7-8 μm only partially in AHI
Ammonia	NH ₃	Yes	Potentially in CAFO site	Air pollutant, precursor to ozone	Strong periodic features 8.5-12.5 μm
Benzene	C ₆ H ₆	Yes	Potentially in refinery site	Known carcinogen, source burning coal and oil	Strong feature 14-15 μm in AIRS, smaller feature at 9.5 μm in AHI

Sulfur Dioxide is the third most abundant gas released from volcanic eruptions (H_2O and CO_2 are the first and second). Rates of emission from an active volcano can range from 10-20 million metric tons per day [52]. Significant man-made emission of sulfur dioxide released into the atmosphere is due to coal burning electrical utilities and is on the order of 13 million tons per year [53]. Its chemical composition is SO_2 . The infrared signature is due to sulfur and oxygen stretching vibration. This provides strong infrared bands between 1450 cm^{-1} and 1150 cm^{-1} ($6.8\text{ }\mu\text{m}$ – $8.7\text{ }\mu\text{m}$) [54]. In the AHI wavelengths detection will be made using the smaller feature near $8.7\text{ }\mu\text{m}$.

Methane is emitted from both natural sources (~ 190 million tons per year) and man-made sources (~ 540 million tons per year – US only). According to the EPA, 60% of the methane emissions are due to anthropogenic activities. These include, landfills, natural gas and petroleum activities, coal mining, livestock emission, and manure management. Natural sources include wetlands, termites, oceans, and methane hydrate. Of interesting note, is an estimated 13% of global methane emission is due to ocean and methane hydrates. There is much ongoing research determining exact sources and rates [55]. Methane has a chemical composition of CH_4 and is a hydrocarbon. Methane has strong absorption features in the $7\text{ }\mu\text{m}$ to $8\text{ }\mu\text{m}$ range, which are useful for both AIRS and AHI wavebands.

Ammonia is widely and commonly used in household and industrial cleaners. It is also present in the manufacture of synthetics, plastics, explosives, and has numerous other uses. It has both man made and natural sources and is tracked by the EPA as a

precursor/promoter of air pollutants. Ammonia's chemical composition is NH_3 and it has strong, narrow absorption features from 9 μm to 13 μm .

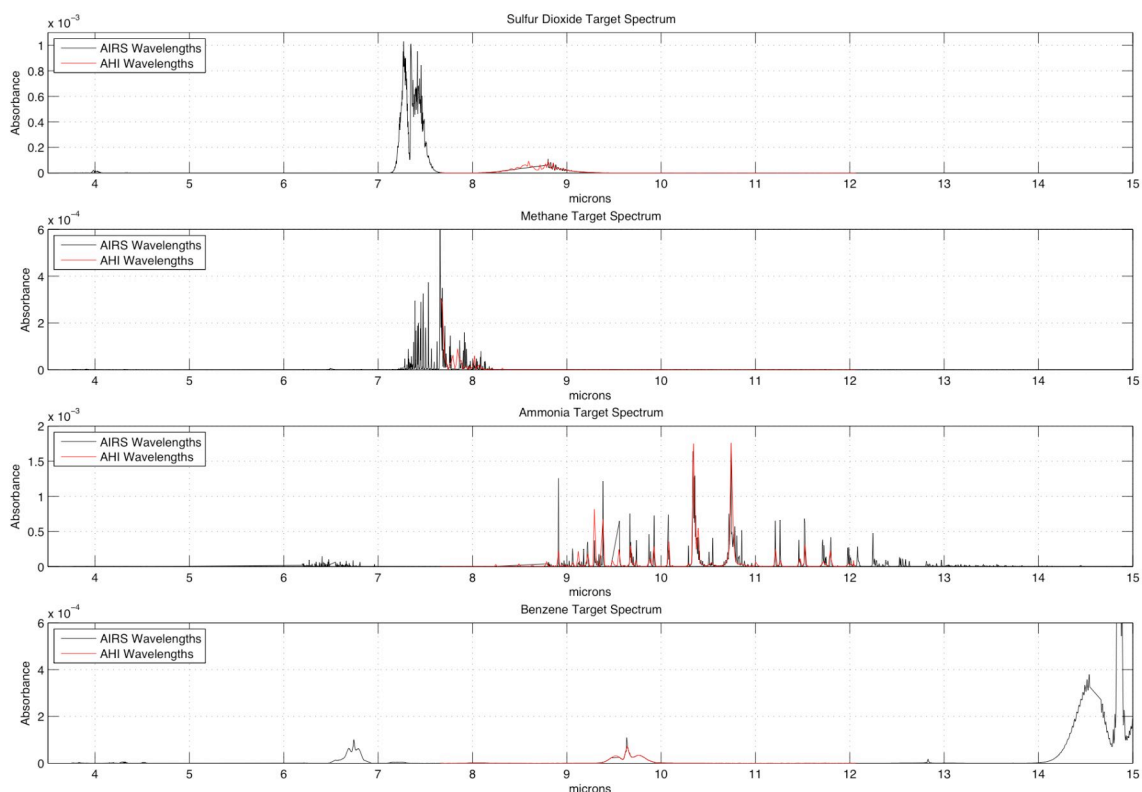


Figure 9. PNNL Target Spectra (AHI and AIRS Wavebands)

Benzene is used primarily in motor fuels and as a solvent. It is also used in the manufacture of detergents, explosives, and pharmaceuticals. It is found in tobacco smoke, emissions from burning coal and oil, motor vehicle exhaust, and evaporation from gasoline service stations [56]. Benzene has a chemical composition of C_6H_6 and is a hydrocarbon. Its peak absorbance occurs in the 3100 cm^{-1} ($3.2\text{ }\mu\text{m}$) range and has weaker

bands in the 2000 cm^{-1} to 1700 cm^{-1} ($5\text{ }\mu\text{m}$ to $6\text{ }\mu\text{m}$) range. Of most interest in AHI wavelengths is C-H bending between 1200 cm^{-1} to 1000 cm^{-1} ($8\text{ }\mu\text{m}$ to $10\mu\text{m}$) [54].

1.6 Hyperspectral Sensors

The Airborne Hyperspectral Imager (AHI) was developed by the Defense Advanced Research Project Agency (DARPA) as part of their Hyperspectral Mine Detection Program. It was built to exploit features in the $8.0\text{ }\mu\text{m}$ to $11.0\text{ }\mu\text{m}$ long-wave infrared range with a spatial and spectral resolution designed to detect buried mines. [57] AHI has demonstrated its ability to perform landmine detection when the appropriate processing algorithms are applied [58]. It has also demonstrated its ability to detect volcanic SO_2 [59], and toxic air pollutants [60].

The AHI system consists of a LWIR hyperspectral sensor with a boresighted 3-color LCD linescanning camera. The system performs on-board radiance calibration and has 256 spectral bands between $7.5\text{ }\mu\text{m}$ and $12\text{ }\mu\text{m}$. The AHI data used was provided by the Environmental Protection Agency from their 2004 Plume Ex experiment [60]. Figure 10 is an AHI hyperspectral image and linescan image of the Dow Chemical Plant in Freeport, Texas.

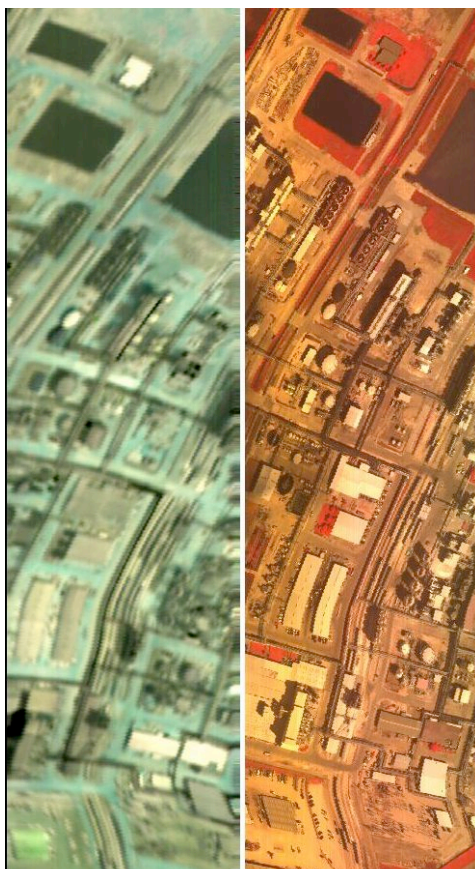


Figure 10. AHI Hyperspectral Image (Left) and Linescan Image (Right)
(Images not Geo-Rectified)

The AHI data was collected over 2 days in April 2004 and consists of over three dozen flightlines of data. Flightlines are as long as 7000 lines and have been manually split into 750 line segments for processing. The data was collected over several chemical plants, refineries, power generation facilities, and animal feeding operations.

The Advanced Infrared Sounder (AIRS) was launched aboard NASA's Aqua satellite in May 2002. It is a high spectral resolution infrared sounder that covers 3.74 μm to 15.4 μm across 2378 channels. It has 13.5 km spatial resolution with better than 1 K

temperature retrieval accuracy. Earth science products generated by the AIRS ground processing system include land and sea surface temperature, atmospheric temperature profiles, cloud cover, cloud-top height, cloud-top temperature, and surface emissivity. The AIRS instrument also carries a Vis/NIR multispectral sensor with 2.7 km spatial resolution. The multispectral sensor is intended to identify low level cloud, different terrain, and surface covers [6; 61-64].

AIRS is the first NASA instrument to measure global climate change parameters. The system has demonstrated its ability to perform high accuracy retrieval of global atmospheric and surface parameters. Its demonstrated temperature retrieval accuracy is better than 1 K [65]. Research and products developed with its data include cloud characteristics and parameters, volcanic sulfur dioxide, global methane, carbon dioxide, carbon monoxide and numerical weather prediction [65-70]. From published research, AIRS data and products support a rich area of research and study within the earth science and atmospheric science communities.

The AIRS data selected for this study include three separate data sets. The first data set is of the Montserrat volcanic eruption in July 2003. This data set is known to have SO₂ within its volcanic ash cloud. The second and third data sets are man-made emissions of SO₂. This occurred in June and July 2003 outside of Al-Mishraq, Iraq. This event was caused by a fire at a sulfur plant that burned for almost a month.

1.7 The Wavelet and Wavelet Packet Transforms

Before we begin a description of the wavelet and wavelet packet transforms, we must first make some observations regarding the signal that we are applying it to. The fundamental intent of remote sensing and specifically hyperspectral analysis is to collect electromagnetic information across multiple frequencies and separate the signal of interest from the noise. That noise may be background emissions, sensor artifacts, sensor noise, atmospheric contributions, contributions from adjacent pixels as well as sub-pixel contributions to the signal. For the hyperspectral analysis the target signal representation in spectral space is known. It is a gas spectrum from the reference library.

In the hyperspectral data cube, each pixel's z-vector is the signal under study. That signal is treated as a linear combination of the fundamental target signature being searched for and all of the background contributions. Even for the gas detection problem, the background is a linear addition to the weak target signal as described in the radiance, model equation (11).

The idea of applying a linear transformation to that signal or that population of signals is not new [31; 42; 43]. Its purpose is representing the signal in a new way and gaining new information [71]. The WPS approach applies the 1-D wavelet packet transform to each individual pixel spectrum in an attempt to separate the target signal from the background.

The wavelet ψ is a function of zero-average, dilated with a scale parameter s , and translated by u . [29],

$$\int_{-\infty}^{+\infty} \psi(x) dx = 0 \quad (18)$$

$$\psi_{u,s}(x) = \frac{1}{\sqrt{s}} \psi\left(\frac{x-u}{s}\right) \quad (19)$$

The wavelet transform of a function f at scale s and position u is

$$Wf(u,s) = \int_{-\infty}^{+\infty} f(x) \frac{1}{\sqrt{s}} \psi^*\left(\frac{x-u}{s}\right) dx \quad (20)$$

Traversing scale and location for s and u are accomplished through dyadic intervals. A change in notation is made here, moving from scale s , indicated now by j , and location u , indicated now by k . The intervals are defined as

$$I_{j,k} = [2^{-j}k, 2^{-j}(k+1)] \quad (21)$$

Stepping through scale j and location k , a simple example for a dyadic interval spanning 0 to 1 is illustrated in Table 4.

Table 4. Example Values for Dyadic Intervals

j	k	$2^{-j}k$	$2^{-j}(k+1)$
0	0	0	1
1	0	0	0.5
1	1	0.5	1
2	0	0	0.25
2	1	0.25	0.5
2	2	0.5	0.75
2	3	0.75	1
3	0	0	0.125
3	1	0.125	0.25
3	2	0.25	0.375
3	3	0.375	0.5
3	4	0.5	0.625
3	5	0.625	0.75
3	6	0.75	0.875
3	7	0.875	1

$j=0$	0.0-1.0 $k=0$							
$j=1$	0-0.5 $k=0$				0.5-1 $k=1$			
$j=2$	0-.25 $k=0$.25-.5 $k=1$.5-.75 $k=2$.75-1 $k=3$	
$j=3$	0-.125 $k=0$.125-.25 $k=1$.25-.375 $k=2$.375-.5 $k=3$.5-.625 $k=4$.625-.75 $k=5$.75-.875 $k=6$.875-1 $k=7$

Figure 11. Illustration of Dyadic Intervals

The maximum scale j and locations k are driven by the length of the input vector. The input signal to the wavelet transform must be of length 2^n or adjustments to the input vector must be made. (See [26] pgs 231-232 for additional details.) When dealing with the integral in equation (18) and its application to a discrete input vector, the integral can be converted to a summation over the appropriate intervals.

The wavelet transform in equation (20) convolves the wavelet ψ with some f at every scale j and location k forming an orthonormal basis (if the wavelet is proven to be orthonormal). A convolution of ψ with the input vector occurs at scale j . The convolution is then dilated from scale j to scale $j+1$. The reader is warned though that in practice the scales are inverted. The wavelet is applied to the input vector at the bottom of Figure 11 first and then applied to larger and larger scales.

It is the coefficients of this convolution operation at multiple scales that represent the new information being sought. Unlike the Fourier transform, which provides a convenient transformation between space or time and frequency, the wavelet transformation gives us information at multiple scales and locations.

The wavelet is developed from its scaling function $\varphi(x)$, which satisfies the two-scale dilation equation [26].

$$\varphi(x) = \sum_n h(n)\varphi_{1,n}(x) \quad (22)$$

h is the scaling filter and the wavelet is then,

$$\psi(x) = \sum_n g(n)\varphi_{1,n}(x) \quad (23)$$

where the wavelet filter is

$$g(n) = (-1)^n \overline{h(1-n)} \quad (24)$$

For j and $k \in \mathbb{Z}$, where \mathbb{Z} is the set of integers, the collection of wavelets, [26]

$\{\psi_{j,k}(x)\}_{j,k \in \mathbb{Z}}$ is an orthonormal basis if

$$\langle \psi_{j,k}, \psi_{j+1,k+1} \rangle = \delta_{j,k} \delta_{j+1,k+1} \quad (25)$$

and $f(x) \in L^2(\mathbb{R})$ can be written as

$$f(x) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} c_{j,k} \psi_{j,k}(x) \quad (26)$$

The scaling and wavelet filters are used in the construction of the approximation and detail operators H and G , defined as

$$(Hc)(k) = \sum_n c(n) h(n - 2k) \quad (27)$$

$$(Gc)(k) = \sum_n c(n) g(n - 2k) \quad (28)$$

The operators H and G are a convolution with the filters h and g followed by a downsampling. H and G are used in the orthogonal wavelet matrix W . The discrete wavelet transform can be thought of as linear transformation, taking the signal vector c of length 2^n , and transforming by W [26].

$$Wc = d \quad (29)$$

Similar to the approach that brings the Fourier integral equation into a linear transformation and matrix operation, the wavelet transform can also be proven to be a recursive matrix operation. The discrete wavelet transform is of $O(n)$ operations and is constructed with an approximation matrix H and a detail matrix G . These two matrices are of size $L/2 \times L$ where L is the length of the 2^n input vector. This is essentially a low pass and high pass filter applied to the input signal. The combination of the two matrices construct the wavelet matrix and are recursively applied. The tree structure in Figure 12 illustrates the process [26; 71].

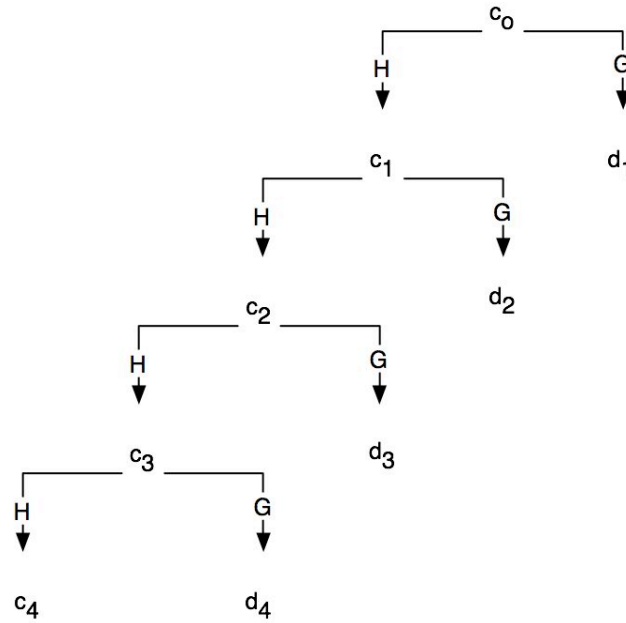


Figure 12. Tree Diagram of the DWT

C_0 is the input vector and c_1, d_1 are the approximation and detail coefficients, respectively. The wavelet matrix is initial(c_n) coefficients. The coefficients at c_4 and d_1

through d_4 make up the orthonormal wavelet basis, and their combined length is equivalent to the input vector.

The wavelet transform has properties that are applicable to detecting specific features. For example the Haar wavelet can capture jump discontinuities. Large wavelet coefficients across multiple scales indicates the presence of the features at specific locations across multiple scales. The Haar wavelet also has good localization around the feature. The coefficients quickly vanish outside of the interval of the feature. This ability to look at multiple scales simultaneously allows large spectral and small spectral features to be identified and filtered.

In reviewing the DWT tree structure of Figure 12 it seems obvious that the tree can be completed. This is accomplished by computing the DWT of the remaining nodes and creating a binary tree, (See Figure 13). This is the basis for the wavelet packet transform. The wavelet packets generated by the discrete wavelet packet transform (DWPT) form the discrete wavelet packet library.

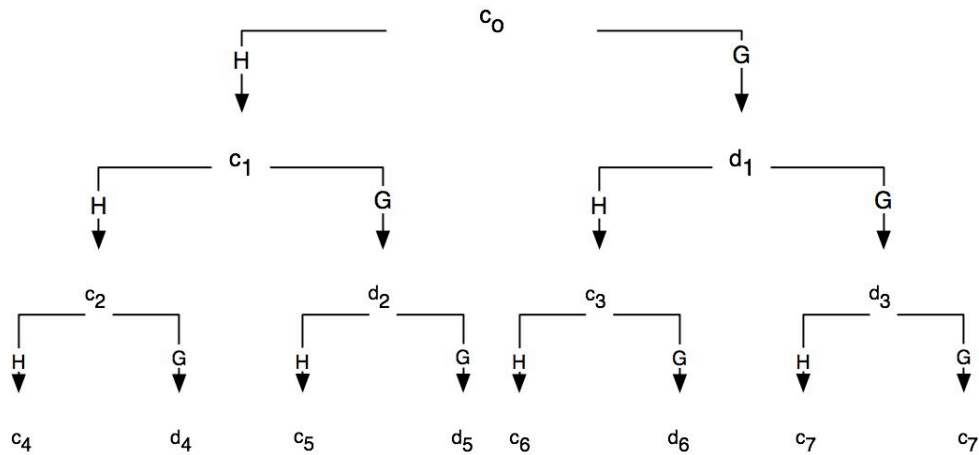


Figure 13. Tree Structure for the Discrete Wavelet Packet Transform

There are several reasons to apply the DWPT over the DWT. Primarily the orthonormal basis developed by the DWT represents a transform that is dependent on the construction of the wavelet and its properties. It is possible to find a best wavelet packet basis that optimizes the resulting coefficients for a specific measure of efficiency [26]. In other words, it is possible to find an orthonormal best basis that best represents the input signal according to some selected measure (the wavelet packet subspace). This parallels the principal components approach in generating an orthogonal basis. The difference is that the principal components approach selects a basis based on variance of the data. The DWPT basis is selected based on a cost function. This cost function may be entropy, energy, norm or other. For the trace gas detection problem, this means that an orthonormal basis can be selected from the wavelet packet library which best represents the gas spectrum. In fact there are $2^{n/2}$ (where n is the length of the input vector) discrete wavelet packet bases to choose from. One could also select a basis that best represents the background in the hyperspectral scene from the wavelet packet library. As discussed in Chapter 2, both are done as part of the WPS approach.

Finding this basis falls to a best basis algorithm. The algorithm calculates a cost function for each node of the DWPT tree relative to some measure. The smaller the cost function, the better that vector is represented by the basis. By traversing the tree from the bottom up, and requiring an additivity condition on the cost at each node, a best basis is quickly found. I refer to [26] pgs. 364-365 for a step-by-step description of the algorithm. There are numerous cost functions that have been used. Some better known examples are Shannon entropy, thresholding, and signal-to-noise ratio.

There is much precedent for application of the DWT to hyperspectral data. In multispectral and hyperspectral analysis, much work has been done to reduce the dimensionality of the data and make it more computationally tractable. Spectral linear unmixing has also been of significant interest as have specific feature extraction methods. There are numerous papers on the DWT approach to dimension reduction. These usually discuss the similarities and improvements of the DWT over the PCA approach. Several authors demonstrate improvements in both computational efficiency and classification results. Previously this type of work compared traditional dimension reduction and classification methods in spectral and feature space, to wavelet coefficient space. In these cases the 1-D DWT is applied and these coefficients are used. Some efforts applying the DWT to hyperspectral data are summarized below.

- In 1995, material identification based on multi-resolution wavelet techniques was proposed. Discrimination was based on wavelet transform decomposition using the fewest wavelet coefficients required. The research demonstrated dimension reduction based on a wavelet approach [72].
- In 2000, the use of wavelet based classifiers was proposed. Vegetation spectra and nearest-neighbor classification were used to demonstrate that wavelet based classifier results are comparable to traditional nearest-neighbor classification [23].
- In 2001, an adaptive multi-channel wavelet transform (AMDWT) was proposed. Sub-pixel targets were assumed to be Gaussian absorption bands. It was concluded the AMDWT was very promising for sub-pixel targets

[73]. Smoothing and derivative analysis using DWT was demonstrated in 2001. Derivative analysis and scale-space images was discussed. It was concluded that wavelet-based methods are feasible and practical for derivative analysis of hyperspectral images. In addition the use of the Mallot/Zhong wavelet to represent hyperspectral signatures across many scales was demonstrated. Also demonstrated were wavelet applications for: 1) smoothing pre-filters 2) generalized spectral decomposition 3) speed improvements [20].

- Wavelet Multi-Resolution Analysis incorporating an automated statistical classification system using Fisher's Linear Discriminant Analysis (LDA) for feature reduction, and Maximum Likelihood for classification was also demonstrated in 2001. It was found that the Continuous Wavelet Transform (CWT) performed more accurately than DWT, but the DWT (Haar) was computationally more efficient [27].
- Also in 2002, DWT dimensionality reduction and feature extraction with wavelet based linear-unmixing was performed. Using the Haar wavelet, wavelet coefficients at different composition levels are used as features. It was concluded that the linearity properties of the DWT are important for Linear Mixing Methods and that wavelet based features result in a promising improvement on abundance estimation [74].
- The DWT for feature extraction and dimension reduction was also proposed. The background of the wavelet transform and the concept Space-

Scale Tiling was discussed as well as wavelet based features and an automated classification system. Dimension reduction using PCA, feature selection using Fisher's LDA and then classify using Maximum Likelihood was performed. Comparing multiple wavelets, it was concluded in endmember applications DWT based systems provide over 95% classification accuracy compared to PCA. This is a significant increase in classification accuracy. The DWT outperformed both Fast Fourier Transform (FFT) and Discrete Cosine Transform (DCT) based feature extraction. For future work, they propose a more systematic choice of the mother wavelet and possible 2-D applications [75].

- In 2003, dimension reduction using several wavelets was proposed. Fisher's LDA and autocorrelation were applied to select features and Maximum Likelihood applied to perform classification. They concluded that it was possible to eliminate noise by discarding lower level DWT decomposition and that the Daubechies 6 DWT had higher classification accuracy than FFT or PCA methods [28].
- Also in 2003, a hybrid wavelet-based PCA reduction was proposed. The computational efficiency of the wavelet transform versus the PCA approach was discussed. Results show that applying the wavelet transform and then the principal components rotation had better classification accuracy than principal components alone. The computational cost of their

proposed hybrid method is higher than the wavelet method alone, but less than the PCA method [76].

The application of the DWPT to hyperspectral analysis, although proposed by some authors and proponents of the DWT, is not as well studied as the DWT applications have been.

- In 2002 a hyperspectral feature extraction methods using DWPT with the best basis cost function based on entropy or norm was proposed but not demonstrated. Accuracy of classification with improved computational efficiency was discussed [18]. Spectral un-mixing through an independent component analysis with wavelet packets was also proposed. This presented an automated endmember selection process [77].

As discussed previously, AIRS earth science research is a well published topic. AIRS data has also been a topic of wavelet based compression [78].

Given the large body of work, the general approach of applying a DWT or DWPT to hyperspectral data is accepted. But, this approach for trace gas detection in the thermal infrared using a DWPT is certainly a new area of research.

1.8 Computational Challenges

The computational challenges are somewhat straightforward. Solution of matched filter for trace gas detection requires $O(N^3)$ operations [29]. Matched filter operations require:

1. Calculation of the spectral data mean from the data set
2. Calculation of the covariance matrix from all pixels in all bands
3. Inversion of the covariance matrix using a pseudoinverse technique
4. Loop through all the pixels to calculate a matched filter score

To accomplish these four steps the following must be done:

1. The entire data cube must be accessible in memory (or significant time costs are incurred when accessing data in swap space)
2. The data must be initially looped through to calculate the data mean
3. Each pixel is looped through a third time to calculate the covariance matrix
4. The inverse (pseudoinverse) is calculated
5. The data must be looped through again to calculate the matched filter score

The wavelet packet approach overcomes these issues by performing pixel by pixel operations and has a lower computational cost, $O(N\log N)$. The wavelet packet approach applies the following steps:

1. Prior to processing data, select the best basis for target material and background from the DWPT library using the selected cost function
2. Loop through each individual pixel of data and apply the Wavelet Packet Transform, an $O(N\log N)$ operation

3. During this loop apply a matching method to calculate a match score

The wavelet packet approach eliminates the need for all of the data to be stored in memory. The order of operations is significantly less than the matched filter and the data is only looped through once. Finally, this approach is a pixel-by-pixel calculation, which easily lends itself to parallelization.

1.9 Research Goals

The goal of this research has two integrated components, 1) develop a wavelet packet hyperspectral detection method, and 2) develop an easily parallelizable method for trace gas detection in spectral data. By developing a wavelet packet trace gas detection method in spectral space, this becomes an easily parallelizable problem. Each individual pixel can be analyzed separately on any node of a cluster computer or on any processor of a multi-processor high performance computer architecture. This differs significantly from the current state-of-the art in trace gas detection. Current techniques are based on the matched filter. These require calculation of the inverse covariance matrix and subsequent eigenvalue decomposition. Thus, matched filter techniques or any other statistical techniques in feature space, are not easily parallelizable. In addition, a wavelet packet transform with a target specific best basis (the WPS approach), would in itself be a new method for spectral detection. This new method will be demonstrated using simulated

hyperspectral data cubes, data from the Airborne Hyperspectral Imager (AHI), and ultra-spectral data from the NASA Advanced Infrared Sounder (AIRS).

2 Methods

The methods section follows the hyperspectral analysis process as discussed in Chapter 1. There are two parts of the analysis process 1) individual data cube analysis, 2) automated data analysis. Algorithm and data analysis are performed on individual data cubes to measure performance and determine whether the wavelet packet approach compares well to other methods. Once the wavelet packet approach is successful, it is applied to an automated process. All three data types, simulated data cubes (SDC), AHI, and AIRS are examined in the individual cube analysis, while only AIRS data is automatically processed. The automated AIRS process is the end result of the WPS algorithm development.

The automated process downloads AIRS data weekly. The data are processed with CMF and the WPS approach. The products are then disseminated to a website. These detection products include detection maps and KML files, which allow users to view the detections in Google Earth. Results from the automated process are not discussed herein.

Flowcharts are used within this chapter to describe the key steps taken to process the data. Each section will discuss the details of data acquisition and manipulation, processing and analysis, comparison methods, and visualization methods as appropriate. Each flowchart is broken out by data type (SDC, AHI, AIRS). Figure 14 is a roadmap of the analysis steps performed.

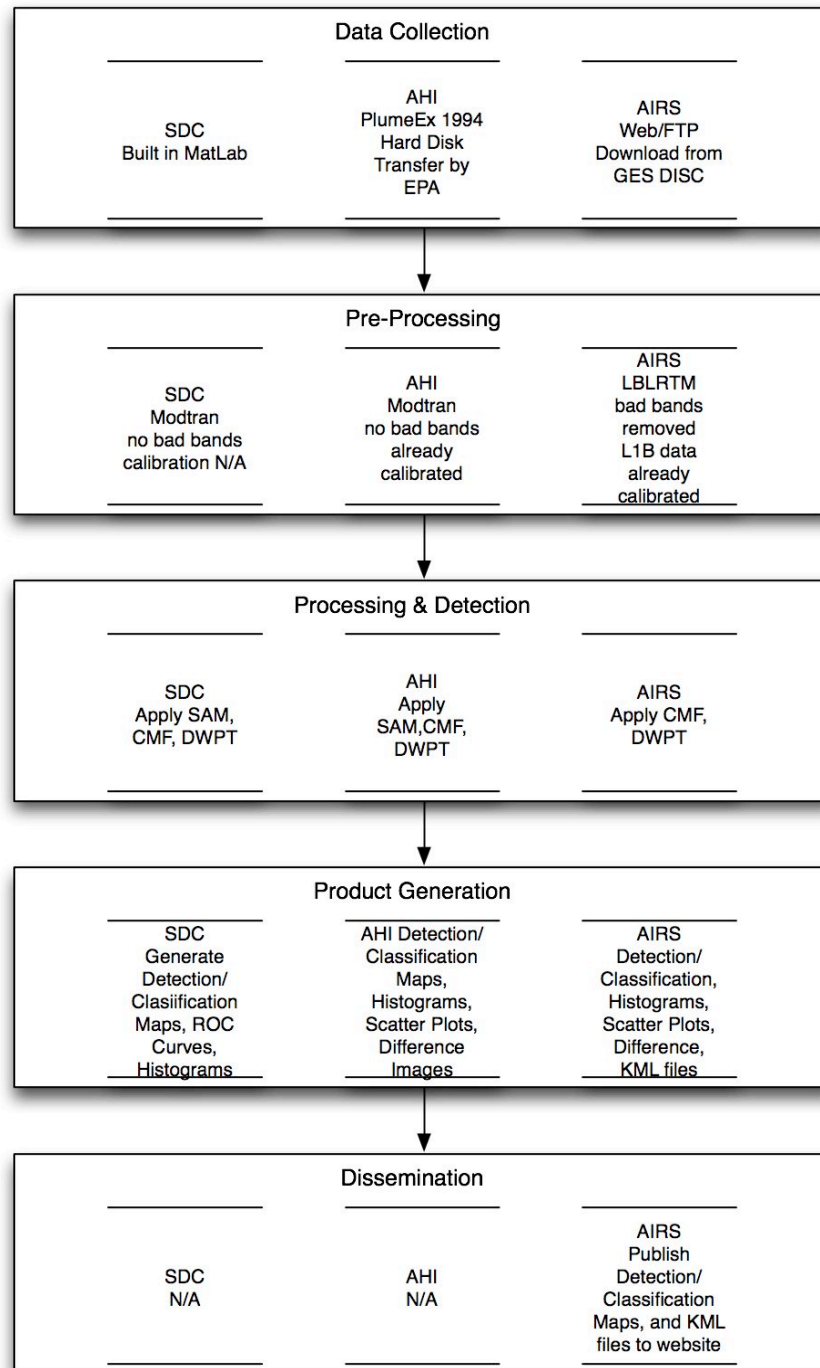


Figure 14. The Roadmap of the Data Analysis Process

2.1 Data Collection

Three types of data were chosen to demonstrate the performance of the wavelet packet approach against the matched filter technique in a logical manner. These types of data are, 1) simulated data cubes, 2) AHI data, and 3) AIRS data.

Simulated data cubes (SDC) provide a controlled data set that have complete ground truth. Plumes are inserted into a homogenous background in specific elliptical patterns. Their use provides complete knowledge of the data. Realistic ground truth for gas detection is excessively difficult. The use of simulated data cubes provides full knowledge of the scene so that algorithm performance metrics can be evaluated. The simulated data cubes contain SO_2 and CH_4 with known locations, known relative quantities, and known noise characteristics. From this information receiver operator characteristics (ROC) curves are constructed to compare multiple detection methods to each other. There are many instances of simulated hyperspectral data for algorithm development and comparison [40; 79; 80]. In this case, the constructed data is much less complex and contains a homogenous background with weak gas signatures.

Several cubes are constructed with various gases, locations, and quantities. The simulated cubes are constructed using the radiance model discussed in Chapter 1 for the AHI wavebands. Each cube is: 256 samples, by 500 lines, by 256 bands. Table 5 provides the characteristics of the simulated cubes and Figure 15 illustrates the process used to construct each data cube. SDC 1 is the simplest case and SDC 4 the most complex case.

Table 5. Simulated Data Cube Characteristics

<i>Name</i>	<i>Background</i>	<i>Gas 1</i>	<i>Gas 2</i>	<i>Location 1</i>	<i>Location 2</i>	<i>S/C</i>
SDC 1	Homogenous – surface emission and atmospheric upwelling	SO ₂	N/A	Quadrant 1	N/A	High
SDC 2	Homogenous + noise	SO ₂	N/A	Quadrant 1	N/A	Med
SDC 2A	Homogenous + noise	SO ₂	N/A	Quadrant 1	N/A	Low
SDC 3	Homogenous + noise	SO ₂	CH ₄	Quadrant 1	Quadrant 4	Med
SDC 4	Homogenous + noise	SO ₂	CH ₄	Quadrant 1	Quadrant 1	Med
SDC 4A	Homogenous + noise	SO ₂	CH ₄	Quadrant 1	Quadrant 1	Very Low

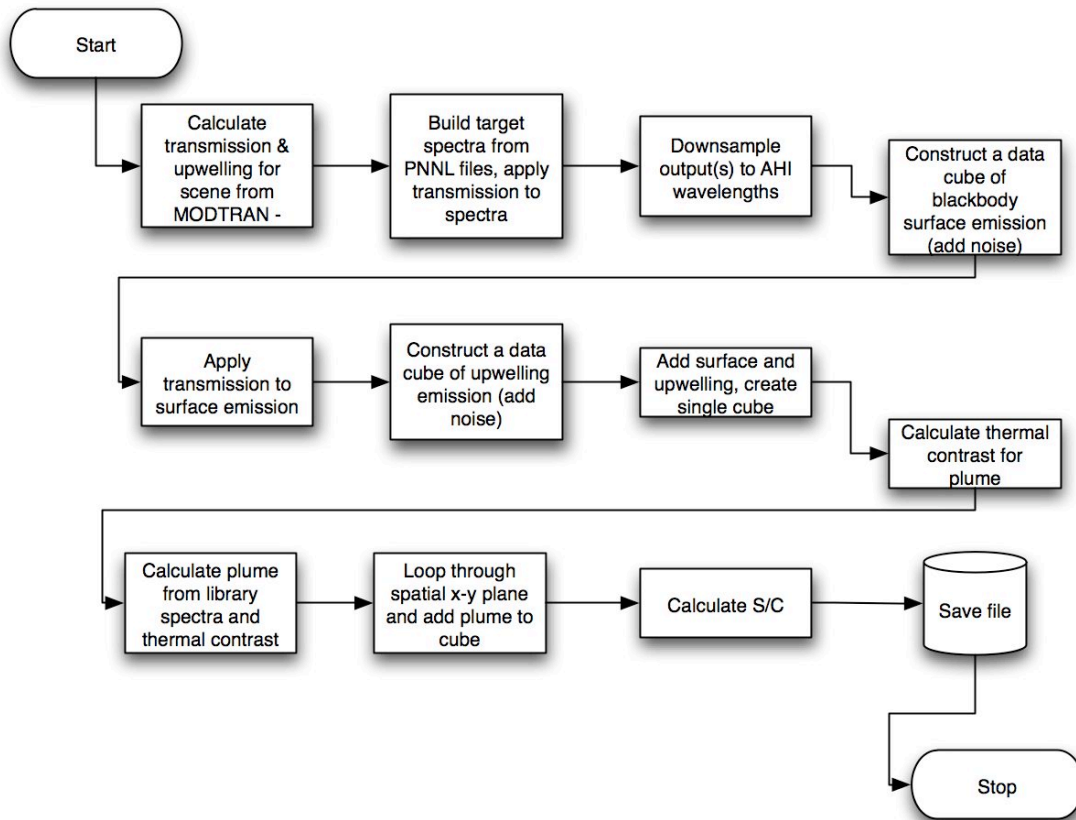


Figure 15. Simulated Data Cube Construction Process

Once the cubes are constructed, the clutter matched filter, wavelet packet subspace approach, and spectral angle mapper, are used against each simulated data cube. SDC 1 provides the opportunity to verify that all three algorithms are performing in an ideal noise-free environment. SDC 2, 3, and 4 have added random noise and a different gas to increase the complexity of the scene. SDC 4 also co-locates the gases. The gases overlap partially in Quadrant 1. (Quadrant 1 is upper left quadrant, Quadrant 2 is upper right, Quadrant 3 is lower left, and Quadrant 4 is lower right.)

2.2 Descriptions of Simulated Data Cubes

SDC 1 inserts a plume of SO_2 into the upper left quadrant of the scene. There are 183 pixels of SO_2 . The background is a linear combination of surface blackbody with atmospheric transmission applied and atmospheric upwelling. The sensor is positioned at a 2000 ft altitude to simulate an AHI flight collect. Gas contribution is approximately 0.02% of column density. The surface temperature is 290 K, the plume is 295 K, and the ambient atmospheric temperature is 292 K. A signal to clutter ratio (S/C) is calculated at greater than 100 dB. The signal to clutter is measured based on the covariance matrix and the target signature [15].

SDC 2 mimics SDC 1 and contains a plume of SO_2 in the upper left quadrant of the scene but has a small amount of random noise. There are 183 pixels of SO_2 . The homogenous background is composed of surface blackbody with atmospheric transmission applied and atmospheric upwelling. The sensor is again positioned at a 2000

ft altitude to simulate an AHI flight collection. Gas contribution (as in SDC 1) is approximately 0.02% of column density. The surface temperature as in SDC 1 is 290 K, the plume is 295 K, and the ambient atmospheric temperature is 292 K. Random noise is added to each pixel in the background. A signal to clutter ratio is calculated at approximately -23 dB.

SDC 2A increases the noise level from SDC 2. As with SDC 1, a plume of SO₂ is inserted into the upper left quadrant of the scene. There are 183 pixels of SO₂. The homogenous background is composed of surface blackbody with atmospheric transmission applied and atmospheric upwelling. The sensor is positioned at a 2000 ft altitude to simulate an AHI flight collection. Gas contribution is approximately 0.02% of column density. The surface temperature is 290 K, the plume is 295 K, and the ambient atmospheric temperature is 292 K. Random noise is added to each pixel in the background. A signal to clutter ratio is calculated at approximately -40 dB.

SDC 3 inserts both a plume of SO₂ into the upper left quadrant of the scene and a colder (less than ambient air temperature) methane (CH₄) plume into the lower right quadrant. There are 183 pixels of SO₂ and 629 pixels of CH₄. The homogenous background is composed of surface blackbody with atmospheric transmission applied and atmospheric upwelling. The sensor is positioned at a 2000 ft altitude to simulate an AHI flight collection. SO₂ contribution is approximately 0.02% of column density and CH₄ contribution is approximately 0.001% of column density. The surface temperature is 290 K, the SO₂ plume is 295 K and the CH₄ plume is 285 K, and the ambient atmospheric

temperature is 292 K. Random noise is added to each pixel in the background. A signal to clutter ratio is calculated at approximately -23 dB for both gases.

SDC 4 co-locates both the SO₂ and CH₄ into the upper left quadrant. There are 183 pixels of SO₂ and 629 pixels of CH₄. 167 pixels contain both SO₂ and CH₄. The homogenous background is composed of surface blackbody with atmospheric transmission applied and atmospheric upwelling. The sensor is positioned at a 2000 ft altitude to simulate an AHI flight collection. SO₂ contribution is approximately 0.02% of column density and CH₄ contribution is approximately 0.001% of column density. The surface temperature is 290 K, the SO₂ plume is 295 K and the CH₄ plume is 285 K, and the ambient atmospheric temperature is 292 K. Random noise is added to each pixel in the background. A signal to clutter ratio is calculated at approximately -23 dB for both gases.

SDC 4A emulates SDC 4 but changes the CH₄ to emission and significantly decreases S/C. There are 183 pixels of SO₂ and 629 pixels of CH₄. 167 pixels contain both SO₂ and CH₄. The homogenous background is composed of surface blackbody with atmospheric transmission applied and atmospheric upwelling. The sensor is positioned at a 2000 ft altitude to simulate an AHI flight collection. SO₂ contribution is approximately 0.02% of column density and CH₄ contribution is less than 0.001% of column density. The surface temperature is 290 K, the SO₂ plume is 295 K and the CH₄ plume is 293 K, and the ambient atmospheric temperature is 292 K. A signal to clutter ratio is calculated at approximately -40 dB and -44 dB for SO₂ and CH₄ respectively.

Figure 16 below illustrates the broadband image of the simulated cubes and spectra for the background and the target gases for SDC 3. Images for the other simulated data cubes are shown in Appendix 2.

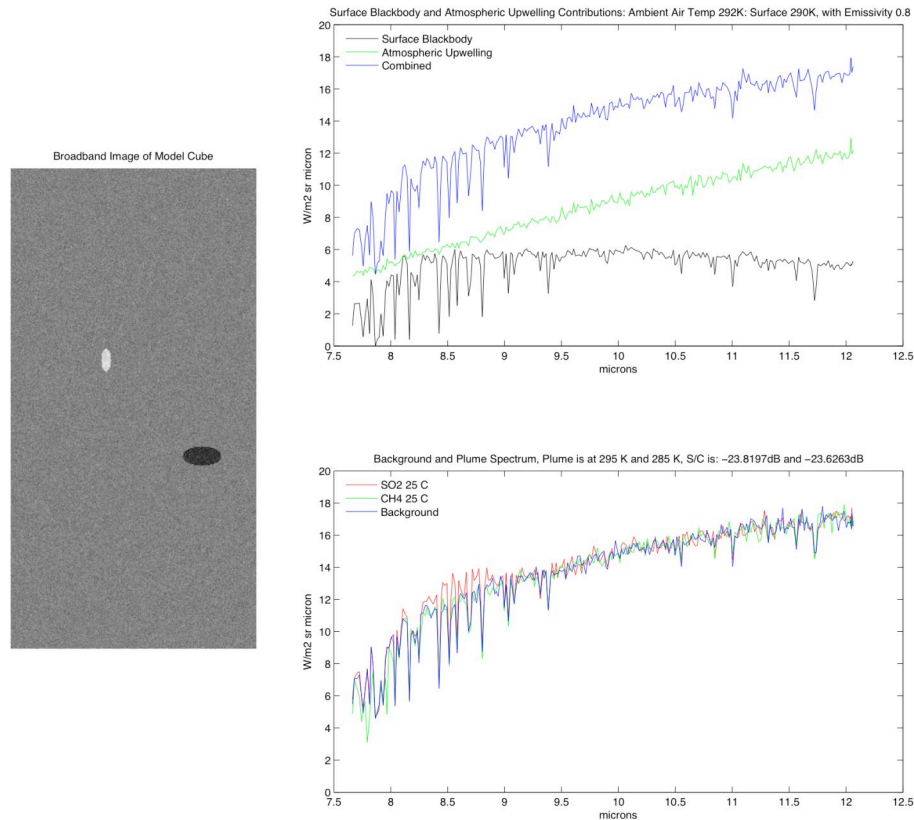


Figure 16. Simulated Data Cube 3

2.3 Description of Airborne Hyperspectral Imager Data

The AHI data used was part of the April 2004 Plume EX experiment conducted by the EPA [60]. Data provided and published papers [60] indicate the possible presence of sulfur dioxide, ammonia, and benzene. The collects occurred over a confined animal

feeding operation (CAFO), a petrochemical plant, and a refinery in southeast Texas. The AHI sensor was flown at 2000 ft and 5000 ft altitudes for the facilities. From this experiment, four flight lines were selected that had indications of these gases in the provided data or published data. Each flight line was 3000 to 7000 lines long and was split into smaller segments for ease of processing. AHI data was provided by Mr. David Williams of the EPA via external hard disk drive.

The AHI data selection process evaluated and selected flightlines based on the geospatial data and metadata included with each flightline. Google Earth was used to visualize the location of each flight. Figure 17 describes the process.

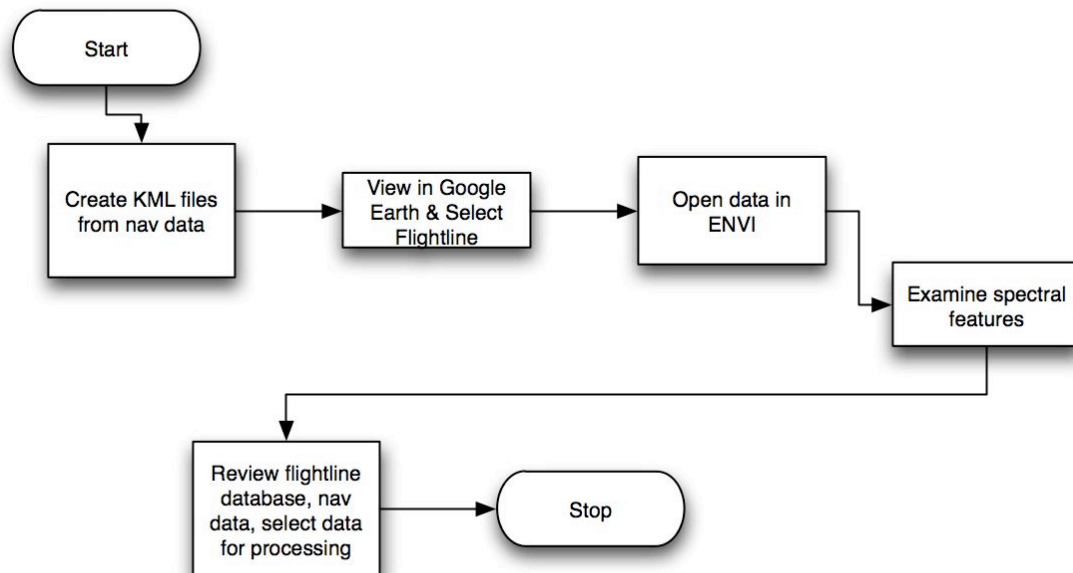


Figure 17. AHI Data Collection Process

The first selected flightline is 121247 collected on April 20, 2004 and was flown at an altitude of 2000 ft. This scene contains a possible SO₂ plume adjacent to power plant stacks. The second flightline, 150652, was flown on April 19, 2004 at an altitude of 5000 ft. In this flightline a suspected ammonia plume is present along the length of the lower buildings. Flightline 105948 is an industrial stack at a refinery. Benzene was reported within this scene. The last flightline, 153924, is of a chemical plant and contains multiple potential plumes. The metadata for 153924 includes facility information for potential chemical releases and their locations. Figure 18 below has the hyperspectral image on the left and the linescan false color image on the right.

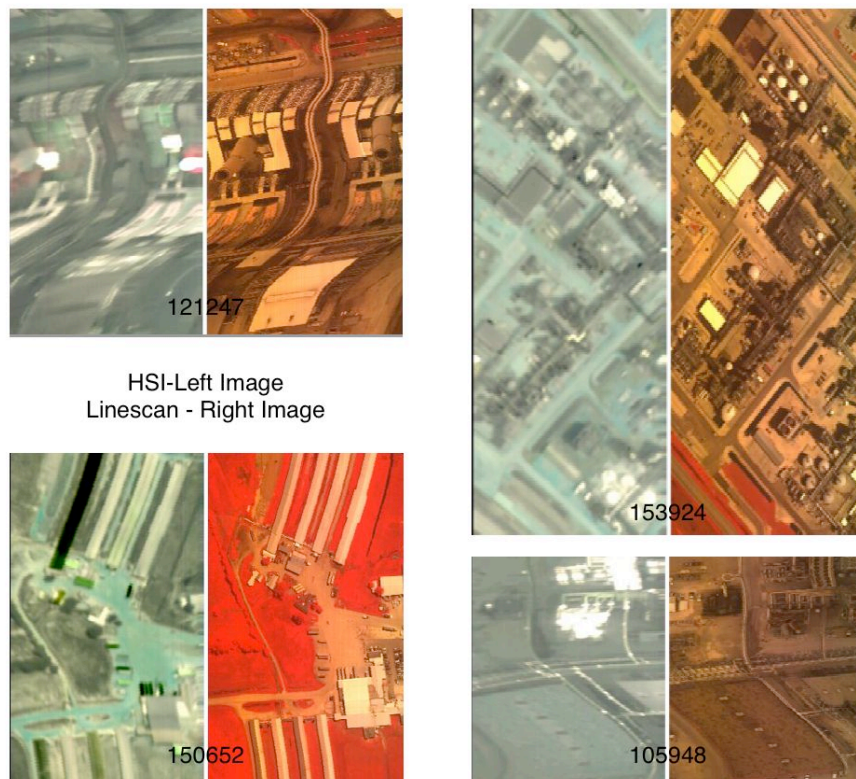


Figure 18. AHF Flightlines from Plume Ex

2.4 Description of Advanced Infrared Sounder Data

AIRS data selected for analysis include two sulfur dioxide emission events. The first is a natural event, the volcanic eruption at Soufriere Hills in Montserrat on July 13, 2003. This event was detected by the NASA's Total Ozone Mapping Spectrometer (TOMS). The event was captured in AIRS data granule 169 from the same day. A Level 1B (unprocessed) AIRS broadband radiance image is shown in Figure 19. An example of the processed AIRS visible/near-infrared high-resolution geo-registered image follows in Figure 20.



Figure 19. AIRS Broadband Radiance Image, Data Granule 169, July 13, 2003

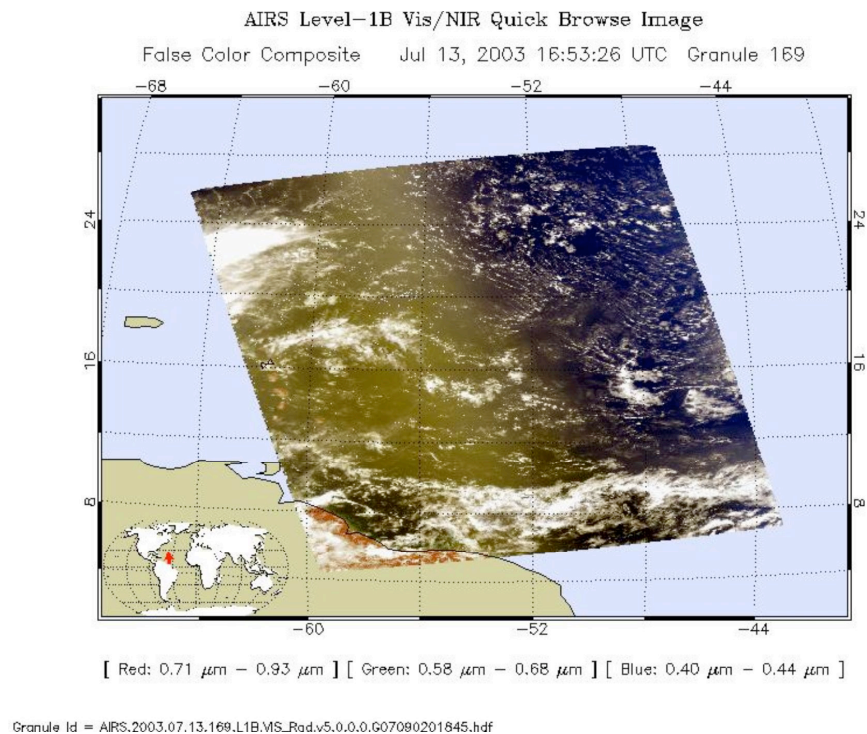


Figure 20. AIRS VNIR Processed Image³

The second and third AIRS data sets are a man-made SO_2 events caused by the fire at the Al-Mishraq Sulfur Plant near Mosul, Iraq, in June of 2003. The fire raged for almost a month. Several AIRS data granules captured the site during the event. The SO_2 plume was also detected in TOMS data and emission quantities were calculated [81]. AIRS data granule 102 from June 29 and data granule 107 from July 2, 2003 were selected for analysis. On June 29, 2003 approximately 100 KT were emitted from the plant. On July 3, 2003 approximately 36 KT was emitted from the plant. Both quantities were estimated with TOMS data. The range of daily quantities emitted from the plant spans from 1 KT to over 100 KT per day.

³ http://disc.sci.gsfc.nasa.gov/get_data.shtml

AIRS data was retrieved from NASA Goddard's Earth Sciences Data and Information Services Center at http://disc.sci.gsfc.nasa.gov/get_data.shtml. The GES DISC is an intuitive website which provides preview images as well as full data sets for immediate download. The DISC also provides an FTP data pool for automated retrieval of collected data. The manual and automated retrieval process are described in Figure 21 and Figure 22 respectively.

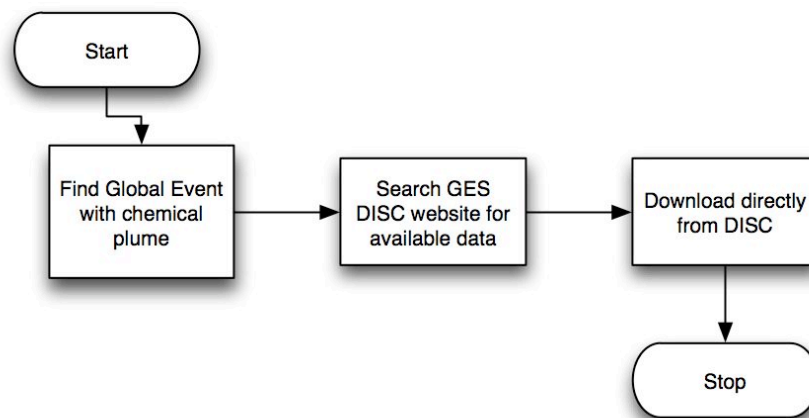


Figure 21. AIRS Manual Data Retrieval Process

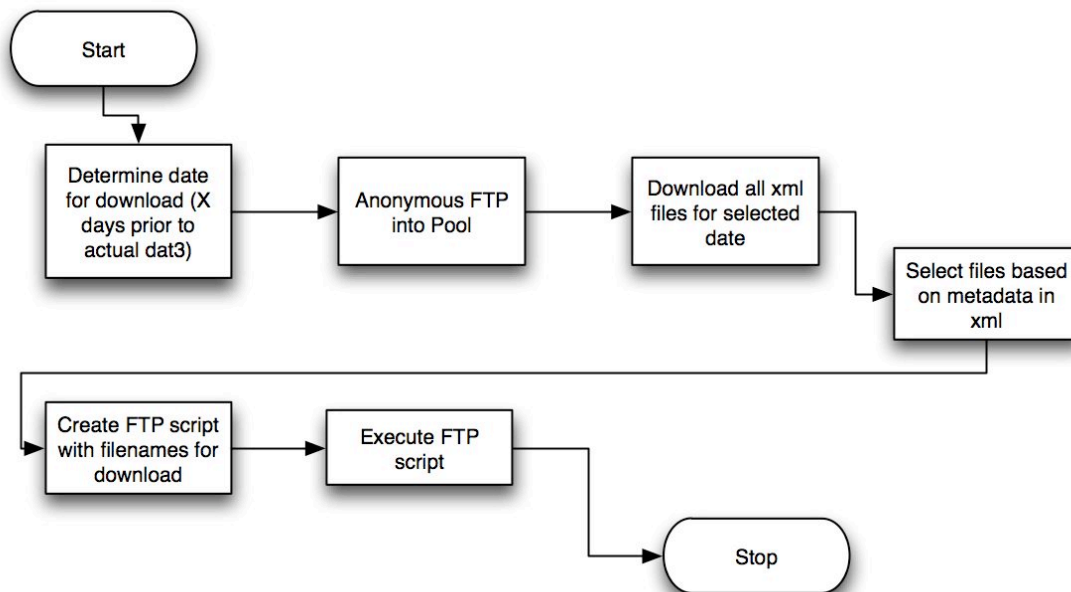


Figure 22. AIRS Automated Data Retrieval Process

2.5 Data Pre-Processing

Pre-processing of the data is dependent on the data type. The three types of data used required similar but different steps in pre-processing. The general pre-processing steps include spectral and radiometric calibration, bad band removal, atmospheric characterization, and target library compensation.

The simulated data cubes were constructed based on the radiance model discussed in Chapter 1 and were designed to emulate AHI data cubes. AHI wavelengths were used and no radiometric or spectral calibration was required for the data. No bad bands were modeled in the data. An atmospheric model was required to generate the atmospheric transmission and upwelling for a sensor at altitude. This was accomplished using

MODTRAN. A straightforward MODTRAN run was performed for a mid-latitude summer collection. The sensor is nadir looking at an altitude of 2000 ft. No aerosols or rain were modeled in the scene. Figure 23 illustrates the surface blackbody and upwelling radiance applied to the simulated data cubes as generated by MODTRAN.

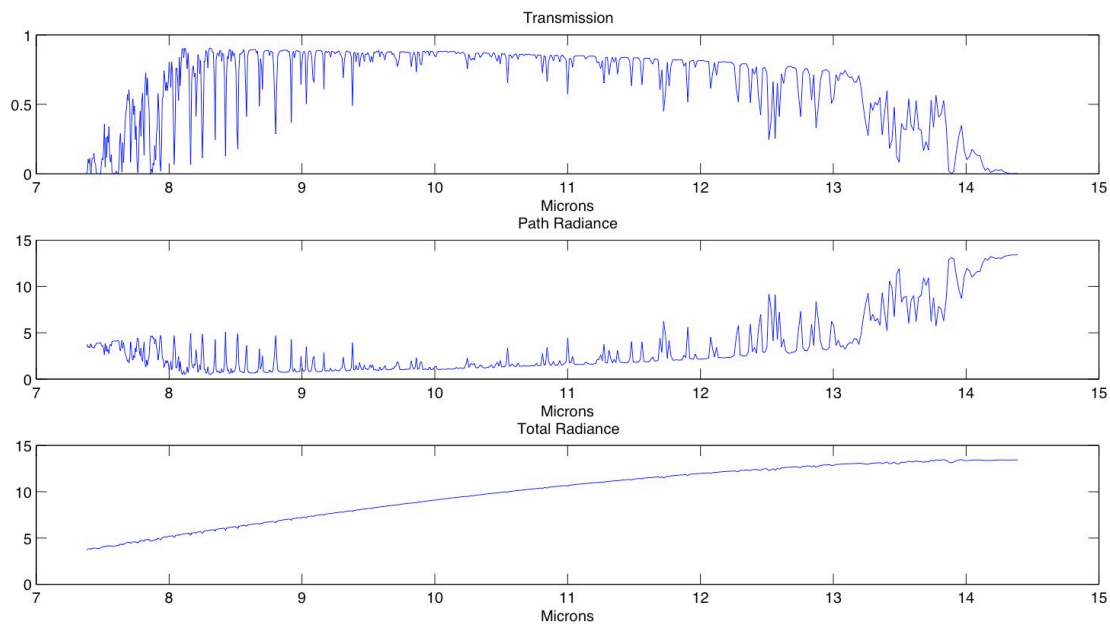


Figure 23. MODTRAN Generated Atmosphere Applied to Simulated Data Cubes

The AHI data provided by the EPA had been previously spectrally and radiometrically calibrated into micrometers and $1000 \cdot \text{W}/\text{sr}/\text{m}^2/\mu\text{m}$ respectively. Initial spectral calibration was slightly off as indicated by several readme files. A separate wavelength file was provided with the data. No bad bands were identified upon initial visual inspection of the data. MODTRAN was used to characterize the atmosphere based on the supplied aircraft flight data. The MODTRAN model was a nadir looking sensor at

the appropriate AHI altitude. No aerosols or rain were modeled. The provided data were binary files with ENVI header files. The binary data was opened in MatLab and spatially subset into files with 300-750 lines each.

Both the simulated data and AHI data were pre-processed in the same manner.

Figure 24 describes the steps.

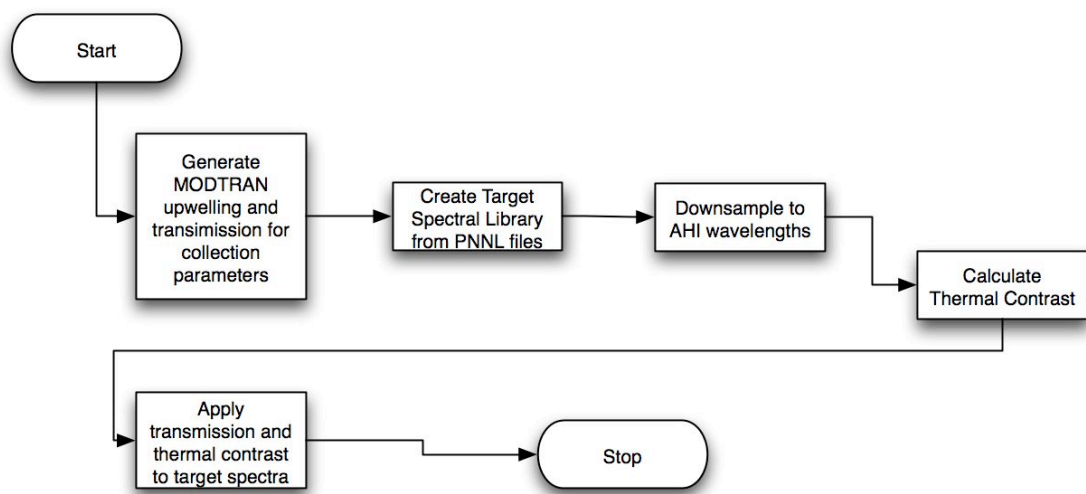


Figure 24. SDC and AHI Pre-Processing Steps

AIRS data retrieved from the GES DISC is provided in hierarchical data format (HDF). These HDF files include the raw data and several layers of metadata, which include all the information required for processing. The metadata includes satellite parameters that define sensor location, pointing angles, sun angles, sensor performance, sensor calibration, and geo-registration data.

The HDF files are opened in MatLab and the metadata examined. The data retrieved is Level 1B. Level 1B data is radiometrically and spectrally calibrated by NASA. Calibration and quality assurance flags indicate good and questionable spectral bands. Based on these flags, any questionable bands are removed from the data for processing. Sensor metadata is also used to drive the atmospheric model. For AIRS data, LBLRTM is used to characterize the atmosphere. Inputs from the metadata into LBLRTM include: sensor altitude, location, CO₂ concentration, and surface temperature. Figure 25 illustrates the AIRS pre-processing steps.

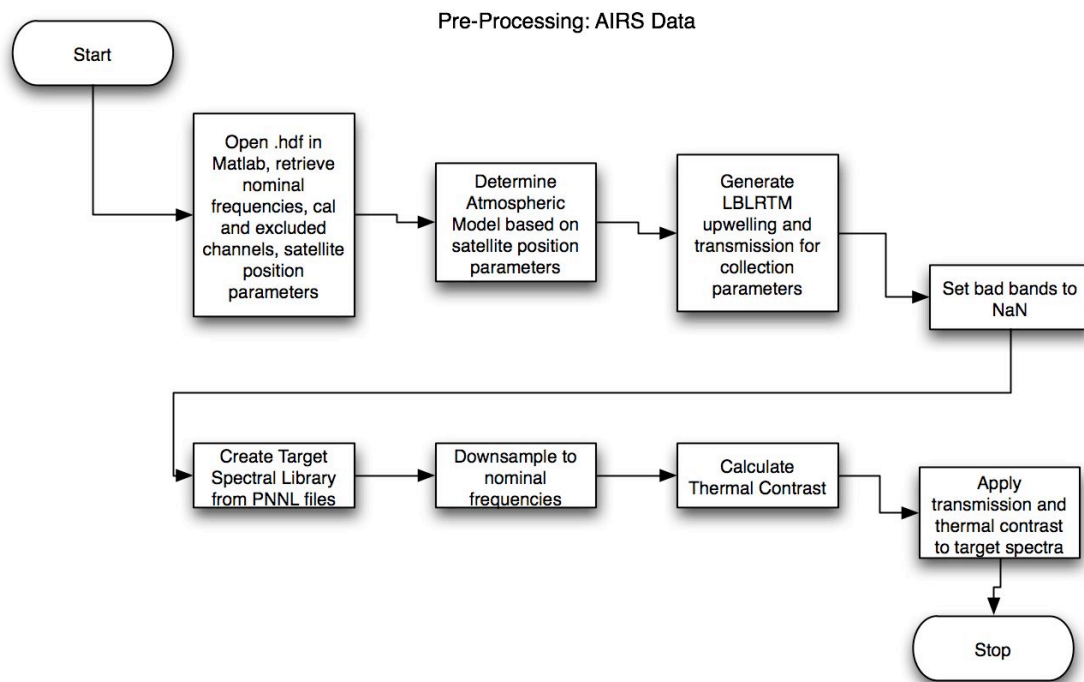


Figure 25. AIRS Pre-Processing Steps

For all three data types, once atmospheric transmission is calculated, that transmission is applied to the target or library spectra. This step follows the radiance model of equation (11). Application of the transmission to the library spectra completes the pre-processing step.

2.6 Data Processing and Detection

The processing and detection step involves application of the detection method to the data. All three detection methods are applied to all data types as part of the individual data cube analysis. Comparisons are then made between each detection method. Application of spectral angle and matched filter is straightforward. The equations presented in Chapter 1 are applied to the data in either a pixel-by-pixel method, or in the case of matched filter to the entire data set. A score value is calculated for each pixel in the scene with the score representing the strength of the match between the pixel vector and the target spectra being searched for. The resulting image or map, sometimes called a detection map or classification map, is the output of the detection method. Each target spectra generates a detection map with the score values in the spatial domain.

Once the detection maps are complete, the results are analyzed in terms of their statistics. As with an image, the score values can be plotted as a histogram illustrating the distribution. From the histogram, pixels are found which are potentially strong matches to the target spectra. A threshold can then be set indicating that values above or below the threshold represent the detected target signature with some confidence level.

Classification maps are generally created from the detection results once a threshold is chosen and applied to the detection image. The threshold is usually based on score values from a histogram or ROC curve that defines the acceptable level of detections versus false alarms. For this comparison, thresholds that represent multiples of the standard deviation from each detection map are used. Figures 26 and 27 describe the SAM and CMF processing steps.

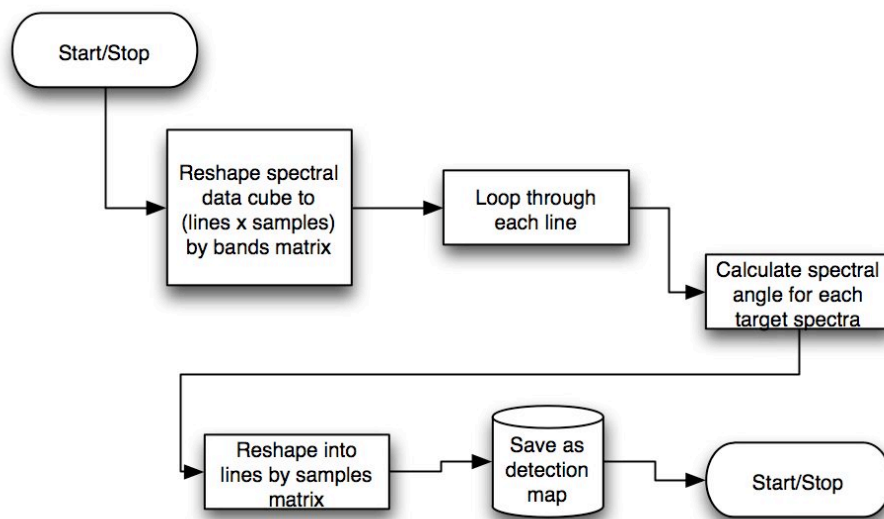


Figure 26. Spectral Angle Mapper Process

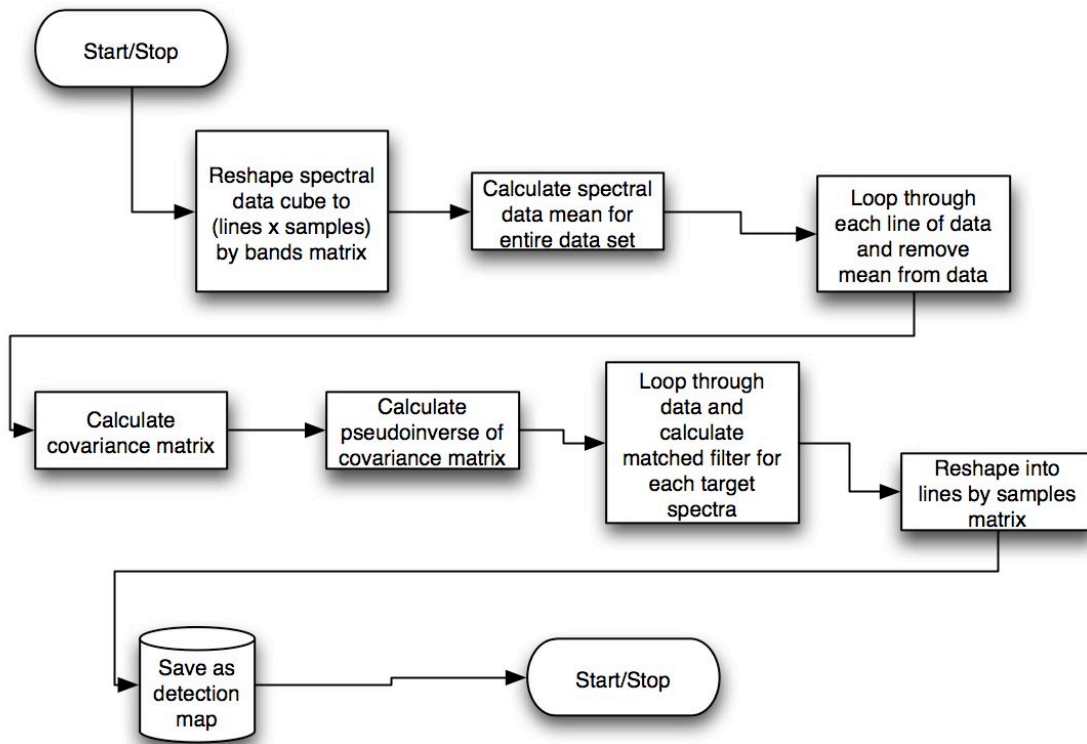


Figure 27. Clutter Matched Filter Process

The wavelet packet subspace approach, as with the other methods, also generates a detection map for each target spectrum. The wavelet packet subspace approach applies a wavelet packet decomposition to each pixel vector prior to generating the detection map.

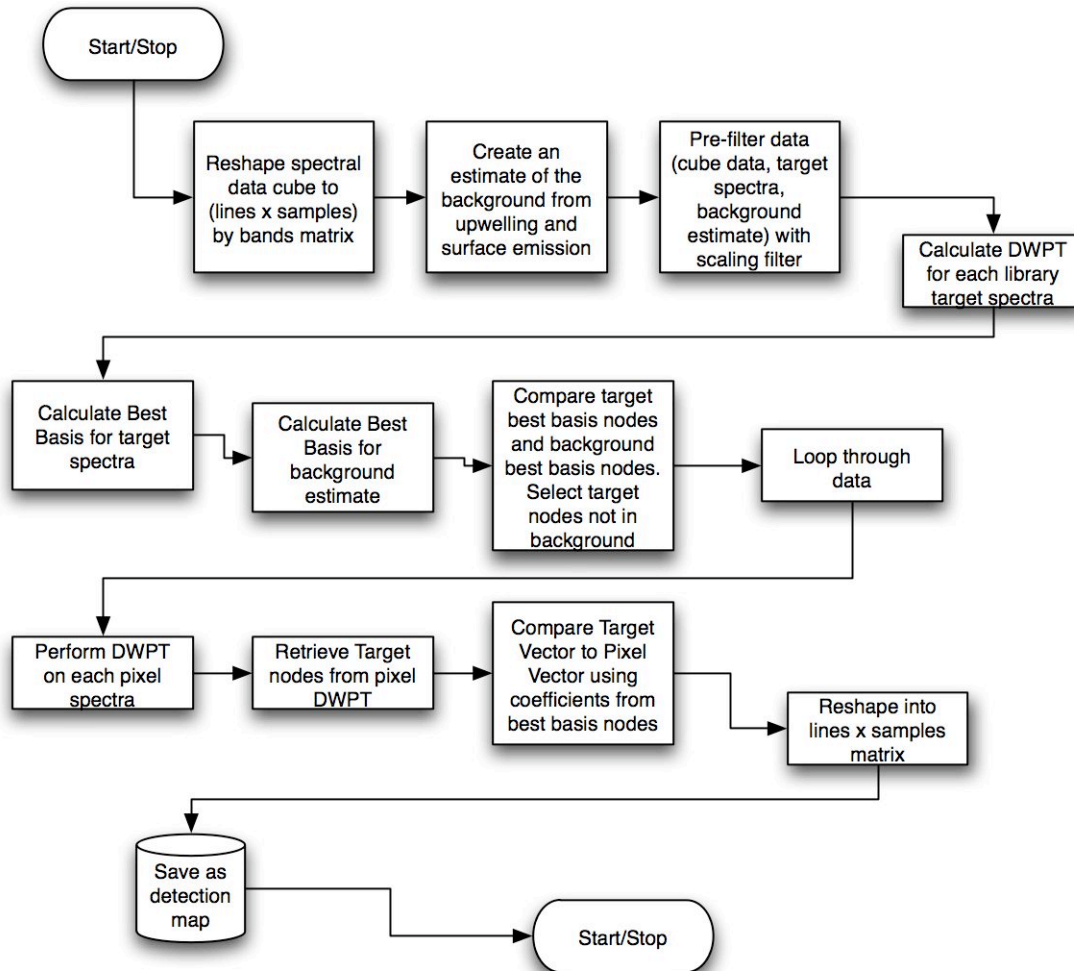


Figure 28. Wavelet Packet Algorithm Approach

Examining Figure 28, two significant operations are performed to transform the data from its original spectral space to wavelet coefficient space. The first is the application of the wavelet packet decomposition, and the second is the selection of the best basis for both the target signature and an estimate of the background. Prior to discussing those operations, we will discuss the overall progression of the process.

1. The data are first reshaped from an x by y by z cube into an x by y matrix. This is done to speed row major operations in MatLab. This step is also performed for CMF.
2. An estimate of the background is calculated. The background estimate is used to identify the contribution of the background to the overall pixel signature. This differs from the calculation of the spectral data mean in CMF, in that it is used to filter background contributions in the wavelet packet library. This estimate can be taken from the data itself. In this study, the estimate is created from the first 25 lines of the data.
3. The scaling filter of the selected mother wavelet is applied to the data (target, background, and data). This prevents any “criminal application” of the DWT or DWPT [71].
4. A wavelet packet decomposition is then applied to the filtered target spectra and a best basis is determined using a cost function. The cost function used is the simple L^1 norm. The depth and terminal nodes of the best basis are retrieved.
5. A wavelet packet decomposition is performed on the filtered background estimate and a best basis is determined using the same cost function. The terminal nodes of the best basis are retrieved.
6. The target spectra best basis nodes and the background best basis nodes are compared. Nodes that intersect both the target and background are eliminated. The remaining target best basis nodes are retained and their coefficients used for detection.

7. The data are then looped through. Each pixel vector has the scaling filter applied then undergoes a wavelet packet decomposition. The coefficients from the target best basis nodes of the data are retrieved and compared to the target spectra coefficients using a matching method. In this study spectral angle is used as the matching method.
8. The resulting score values are reshaped into a matrix representing the spatial domain and this is the detection map.

The steps above reveal that the wavelet packet subspace approach attempts to filter out the background contribution in wavelet coefficient space. The target spectrum is emphasized by selecting its best basis in coefficient space. This process is analogous to de-noising a signal by removing the noise contributions in wavelet coefficient space and then reconstructing the signal in its original space. This approach performs the detection in wavelet coefficient space negating the need to re-construct the original signal. Key to the performance of the approach is the selection of the both wavelet and the best basis.

2.7 Wavelet Selection

Many texts on wavelets discuss the details of theory and application. Selection of a wavelet is met with general recommendations and with significant dependence on the signal under analysis. For the hyperspectral analysis process there are several considerations when selecting a wavelet.

1. Vanishing Moments – The number of vanishing moments describes a wavelet's smoothness, approximation, and ability to reproduce polynomial functions. With increasing vanishing moments comes a greater ability for the wavelet to reproduce functions and polynomials with fewer coefficients.
2. Size of Support/Compact Support – If we are seeking to efficiently capture spectral features of a given size, the size of support by a selected wavelet must generate high amplitude coefficients near or at the same scale as the feature. This will then minimize high amplitude coefficients outside the scale/location of the feature.
3. Size of Filters – Longer filters increase computation time and can add artifacts during reconstruction. Smoother wavelets with higher vanishing moments require longer filters. Since the emphasis is not with reconstruction in the process, there exists a balance between the length of the filter and its computational impact on analysis.
4. Regularity/Symmetry – The regularity and symmetry of the wavelet influences the errors introduced in analysis and reconstruction. Regularity impacts reconstruction in image coding applications, while symmetry is of concern for edge effects in periodic functions. These factors are of less concern for the hyperspectral application.

General consensus for wavelet selection is to choose a wavelet that produces the most number of coefficients closest to zero [26; 29; 82]. Previous work with wavelet and hyperspectral data has made conclusions and recommendations on which wavelets outperform others. Unfortunately none of the research generalizes a best wavelet for hyperspectral data. Additionally, none of it addresses long wave infrared gas detection.

This approach emphasizes wavelet selection based on the target spectrum. It follows the general recommendation to choose a wavelet that produces the most number of coefficients near zero. Wavelet selection follows three steps: 1) numerical analysis of several orthogonal wavelets to determine best wavelet for target spectra, 2) qualitative analysis of numerical results relative to the wavelet selection criteria listed previously, and 3) pragmatic analysis in which the most promising wavelets from Step 2 are applied as the wavelet packet subspace approach and are directly compared to spectral angle and matched filter.

Starting with the target spectra under study (sulfur dioxide, methane, ammonia, benzene), all temperatures available within the PNNL library are selected and downsampled to AHI (or AIRS) wavelengths. Applying a single-level wavelet decomposition to each spectrum and examining the statistics of the coefficients is next. Figure 29 and Figure 30 illustrate the results for AHI wavelengths and Figure 31 shows results for AIRS wavelengths.

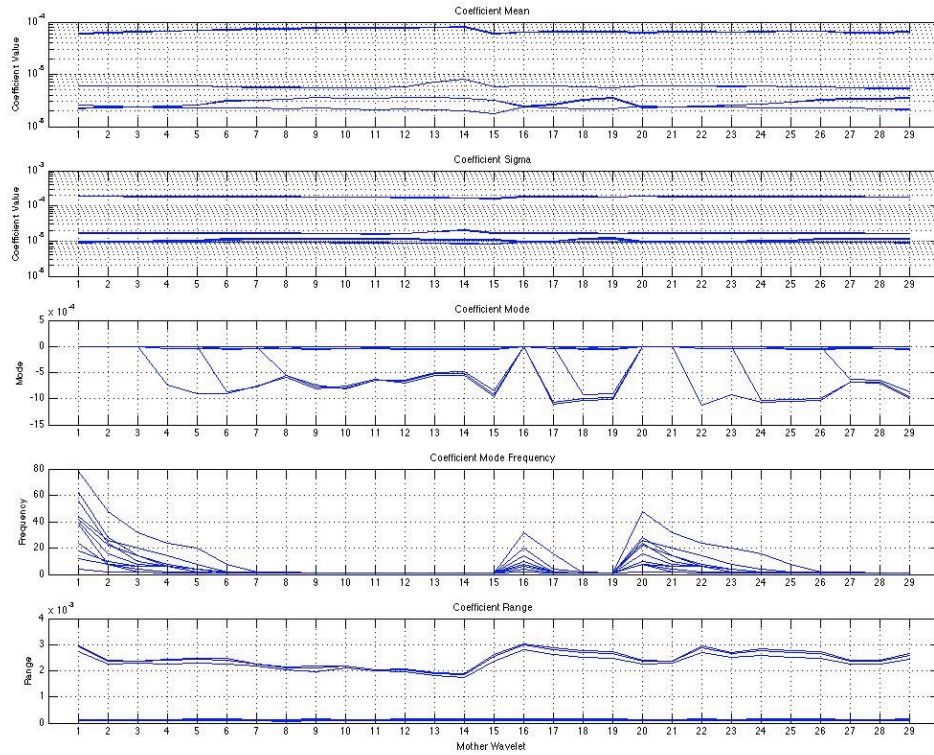


Figure 29. Wavelet Coefficient Statistics for AHI Wavelengths

The x-axis in the figures indicates the mother wavelet applied. The y-axis represents the statistic score. Twelve target spectra are analyzed numbering 1 to 12. There are three temperatures (5°C , 25°C , 50°C) for each of the 4 gases. The list of gases follows in Table 6. In the mode image figures, darker regions represent fewer zero coefficients and whiter regions represent more zero coefficients. The mathematical mode indicates the numbers that occur most frequently.

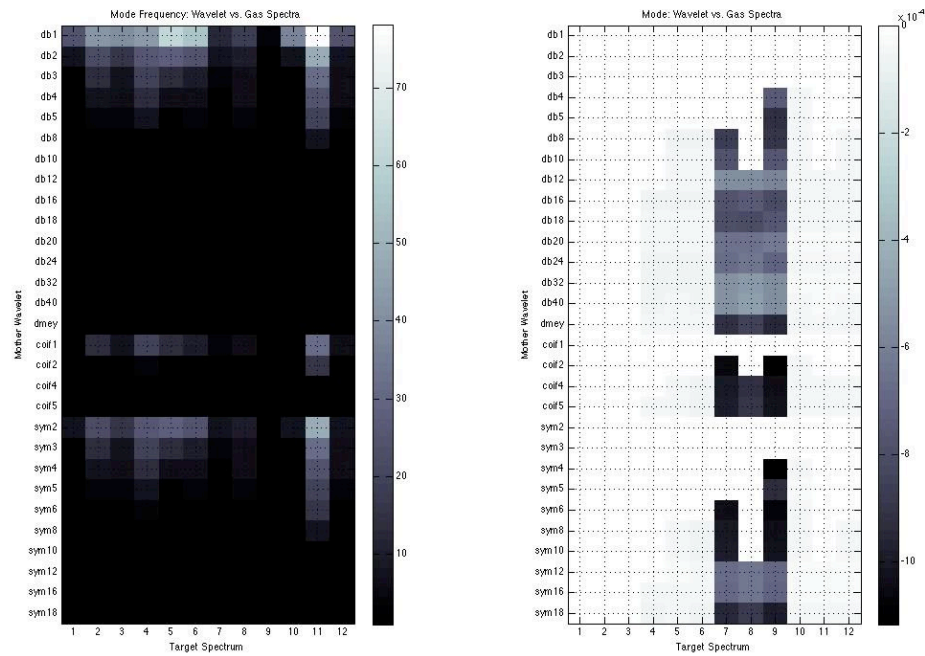


Figure 30. Wavelet Coefficient Mode Image for AHI Wavelengths

Table 6. Target Gases Used in Wavelet Selection
(Number indicates the x-axis target spectrum in Figure 30)

<i>Gas</i>	<i>Temperature</i>		
	<i>5° C</i>	<i>25° C</i>	<i>50° C</i>
<i>Methane</i>	1	2	3
<i>Sulfur Dioxide</i>	4	5	6
<i>Benzene</i>	7	8	9
<i>Ammonia</i>	10	11	12

The means and standard deviations are uninteresting for the spectra at AHI wavelengths. The modes of all wavelets remain close to zero with increasing frequency of modes occurring for the shorter wavelets (Haar, Coif2, Sym2, etc.).

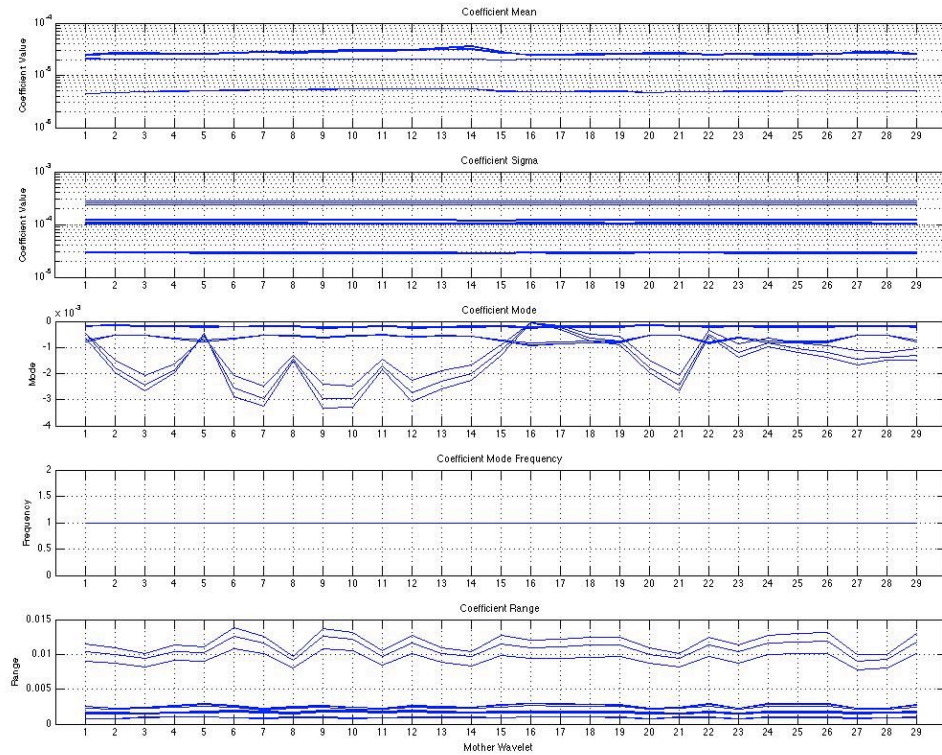


Figure 31. Wavelet Coefficient Statistics for AIRS Wavelengths

Statistics for the AIRS wavelengths are all generally uninteresting with means and standard deviations remaining flat for all wavelengths. Modes are misleading in that the frequency of the mode is one. An in depth review of the score indicates many coefficients near zero, but no wavelet significantly outscores the others. The mode image for AIRS wavelengths is equally uninteresting.

This process is also applied to all spectra in the PNNL library. Results of this are provided in Appendix 2. Results for biorthogonal wavelets are also included.

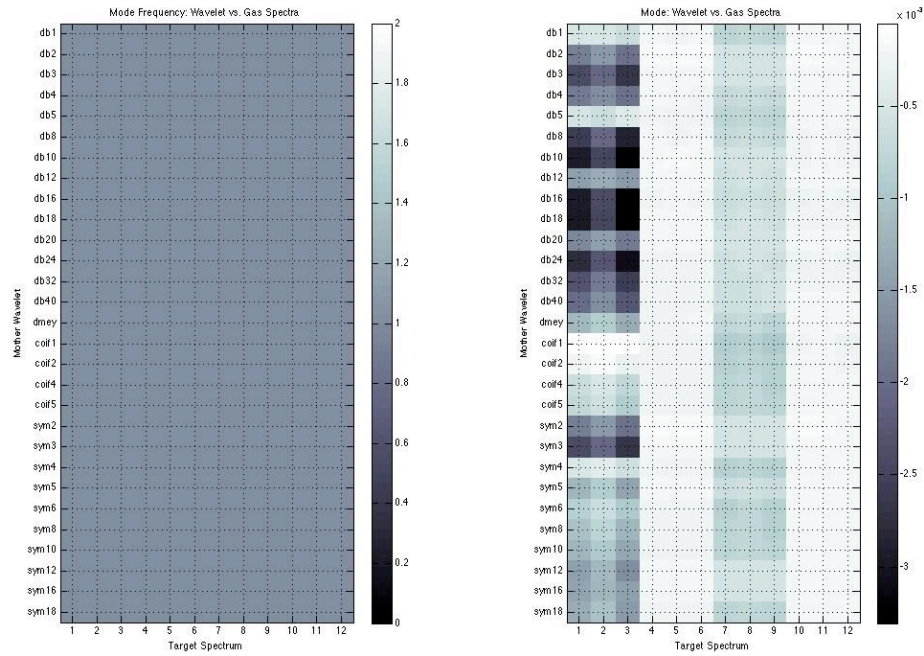


Figure 32. Wavelet Coefficient Mode Image for AIRS Wavelengths

Given the numerical results, there is no clear best choice of wavelet. Taking into account vanishing moments, length, size of support, the most promising wavelets appear to be the shorter Daubechies, Coifman, Symlet, and Biorthogonal wavelets. Results from the pragmatic case are discussed in Chapter 4.

2.8 Cost Function for Best Basis

Previous work in dimension reduction for hyperspectral data analysis has focused primarily on information content with respect to the variance in the data. The wavelet packet library provides the opportunity to select a basis on a multitude of cost functions. There are many texts in information theory that describe entropy and information content

applicable to our best basis selection and cost function. Given the many small variables that drive hyperspectral analysis, the cost function chosen can have a significant impact to the performance of the algorithm. It is however chosen as a simple geometric approach to the information.

The cost function chosen is the L^1 norm. The L^1 norm is a distance measure, similar to the Euclidean distance in n-dimensional space. Euclidean distance has been used as a classification tool in hyperspectral analysis. The L^1 norm chosen based on simplicity as well as similarity to Euclidean Distance. The cost function for the best basis is a distance measure of the target spectrum from a reference. In this case a zero vector. This is analogous to a principal components analysis where the components with the largest variance are chosen as the best representation of the input signal. Figure 33 illustrates the target spectrum as it progress from original spectrum (top), is then filtered using the mother wavelet (middle), and is then represented in wavelet coefficient space (bottom) after best basis selection.

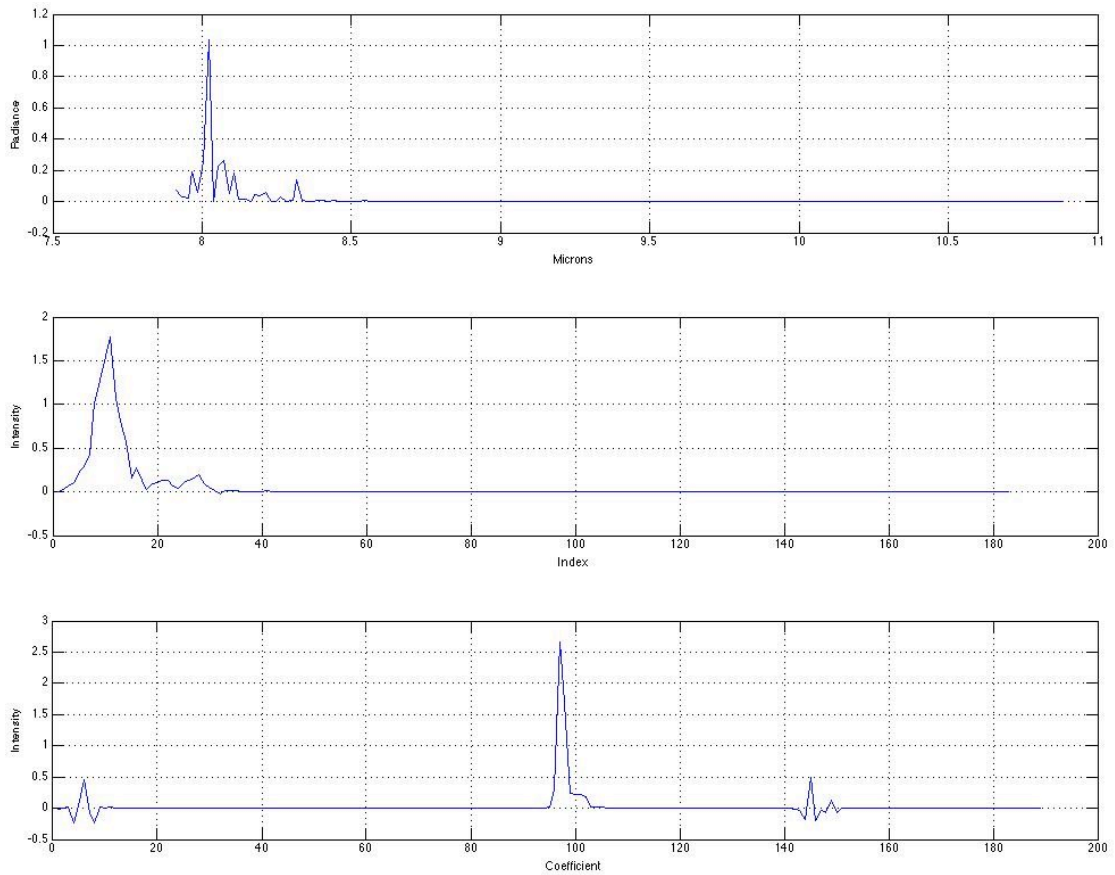


Figure 33. Comparison of Spectral and WPS Coefficient Space

2.9 Products for Comparison and Analysis

From the detection maps, several other plots and images are constructed to compare and contrast results.

1. Image Histograms of score values are generated from the detection maps. These histograms are used to measure and visualize the separation between detected gases and the background. Threshold values are determined from the histograms.

2. Classification Maps are thresholded detection maps. The thresholds from image histograms and from Receiver Operator Characteristics (ROC) curves are used to build classification maps.
3. Receiver Operator Characteristics (ROC) Curves are used to quantify algorithm performance. They require truth data and quantify probability of detection (P_d) versus probability of false alarm (P_{fa}).
4. Scatter Plots illustrate the correlation of an algorithm's score values against another algorithm's score values. Correlation between the sets of score values indicates the algorithms are in agreement with one another.
5. Keyhole Markup Language (KML) files used to visualize results in Google Earth. Results from AIRS data are visualized in Google Earth as part of the automated AIRS process. Georegistered images are used to validate AIRS detections.

The following table summarizes which products will be used for each data type.

Table 7. Summary of Visualization and Comparison Methods

	<i>SDC</i>	<i>AHI</i>	<i>AIRS</i>
<i>Histograms</i>	Yes	Yes	Yes
<i>Classification Maps</i>	Yes	Yes	Yes
<i>ROC Curves</i>	Yes	No	No
<i>Scatter Plots</i>	Yes	Yes	Yes

2.10 Dissemination of AIRS Detections

The automated AIRS process generates classification maps and KML files. The classification maps are thresholded detection maps that are geo-registered to specific map zones. Ten map zones are created across the globe. As AIRS data are processed, its geospatial information is used to register the results to one of the zones. Separate maps are created for each zone, each target spectrum detection, and detection method. Figure 34 is the matched filter methane classification map for the western United States. Figure 35 walks through the steps in the automated process.

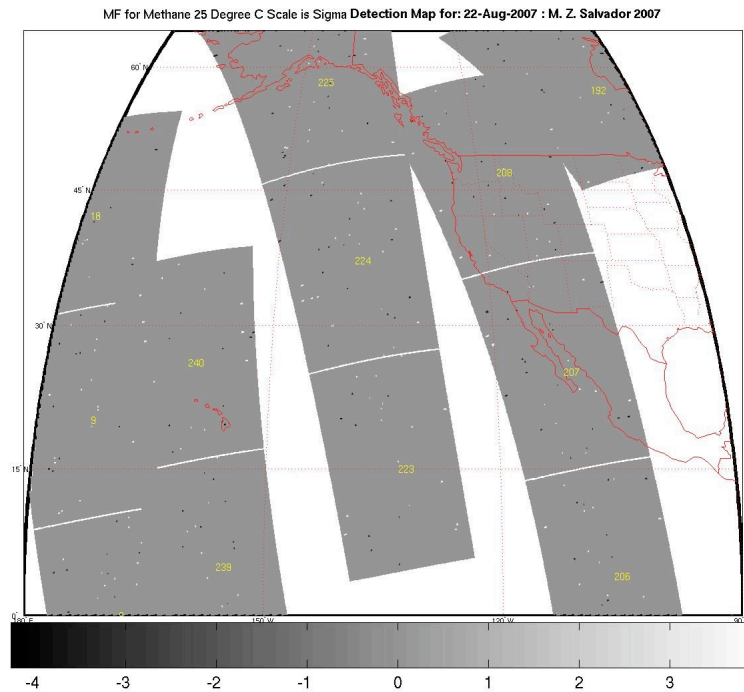


Figure 34. AIRS Automated Classification Map

Once the thresholded detections are geo-registered, each pixel containing a thresholded detection is used to create a polygon for use in Google Earth. The spatial coordinates of the polygon are written to a KML file which are also split up by map zone, gas, and detection method. Polygons are color coded to indicate positive or negative scores, with the height of each polygon representing the scaled score. Figure 36 illustrates the automated KML process. Updated KML files are available for download at http://www.earthintelligence.us/Trace_Gas_Detection/. An example is shown in Figure 37.

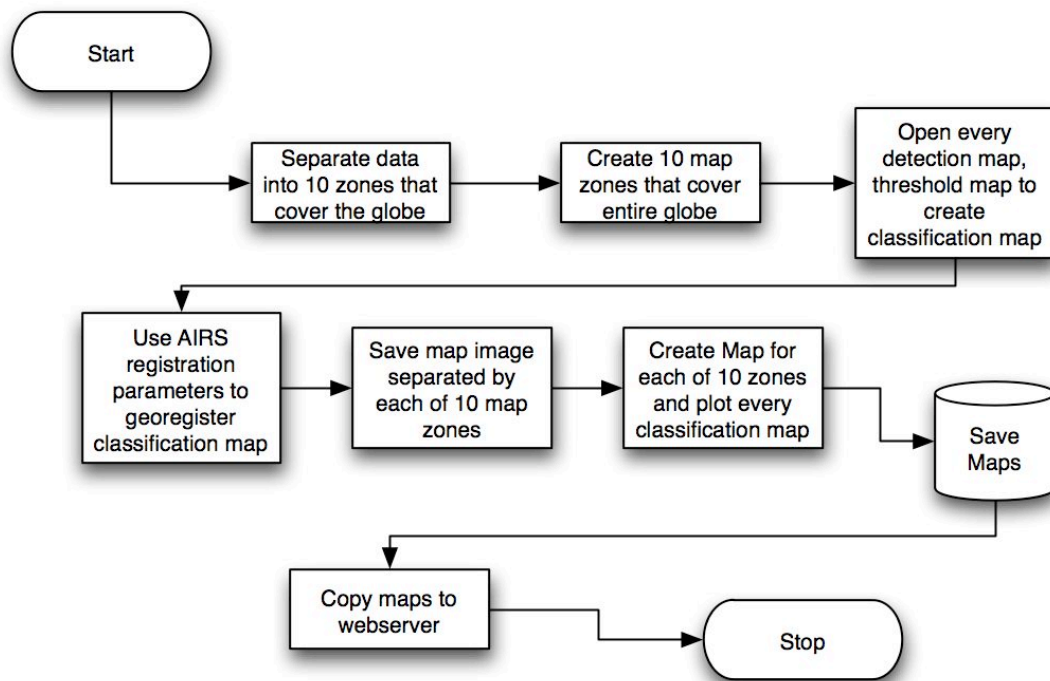


Figure 35. AIRS Automated Mapping Process

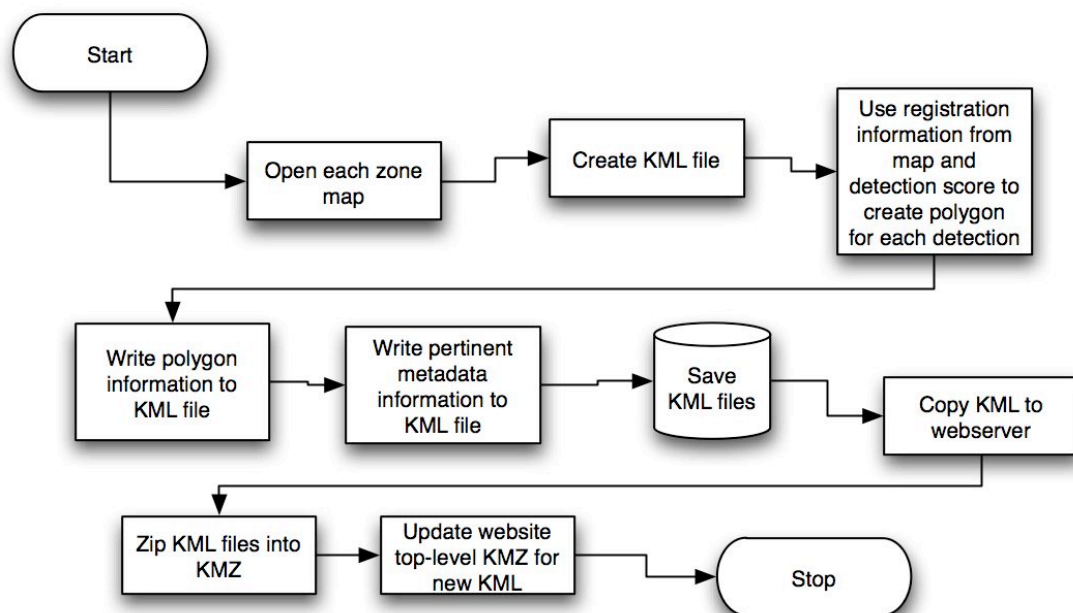


Figure 36. AIRS Automated KML Process



Figure 37. AIRS SO₂, CH₄, NH₃ CMF Detections in Google Earth

2.11 Computational Environment

The majority of algorithm performance evaluation, prototyping, and data processing is performed in MatLab on a Macintosh 2.5 Ghz Dual Processor G5 workstation. The workstation utilizes a 64-bit processor with 8 GB of RAM (4 GB available to each processor). An additional 32-bit dual-core Intel processor is also used to perform wavelet packet subspace analysis of the hyperspectral data.

Although one of the original intents of the proposed research was to port the algorithms to a computer cluster, the original cluster chosen is no longer available. At the successful completion of algorithm development, parallelization of the wavelet packet subspace approach onto a computer cluster is an easily parallelizable problem. Only a suitable hardware platform would be required to complete the parallelization.

Automated AIRS processing will be accomplished on the single Macintosh workstation. A combination of Applescript, shell scripting, and MatLab code is used to automatically download, process, and disseminate detection products via website. The automated process will perform weekly downloads and processing. This could be accomplished on a daily basis using a cluster computer with very high bandwidth internet connectivity.

3 Results

Results are presented in the following section. The results proceed from the simulated data into AHI and then AIRS data. Summary descriptions of the results and the process are discussed with detailed discussion left for Chapter 4.

3.1 Simulated Data Cube Results

The simulated data cubes are used to initially test the performance of all three algorithms against known gases and a homogenous background. They are also used to help characterize the behavior of the WPS approach. The results of the less complex data sets are uninteresting. Results from the more complex data sets guide wavelet selection and provide an initial performance assessment.

SDC 1 contains SO₂ with a simple background (known surface and atmospheric upwelling), and demonstrates that all algorithms easily separate the target signature from the background. SO₂ is easily observed visually in Figure 38. Reviewing the statistics of the detection images demonstrates a clear separation (greater than 3σ) between the background and the target pixels. Only two bars appear in the histogram, one representing the background and one the target pixels (barely visible), see Figure 39. Similarly, the scatter plot has only two points, see Figure 40. One point represents the

single score for all the target pixels and the other point is a single score for all background pixels. The WPS ROC curve is shown (the other two are similar) in Figure 41. P_d and P_{fa} quickly jump to 1 as the target is detected below 1.17 radians.

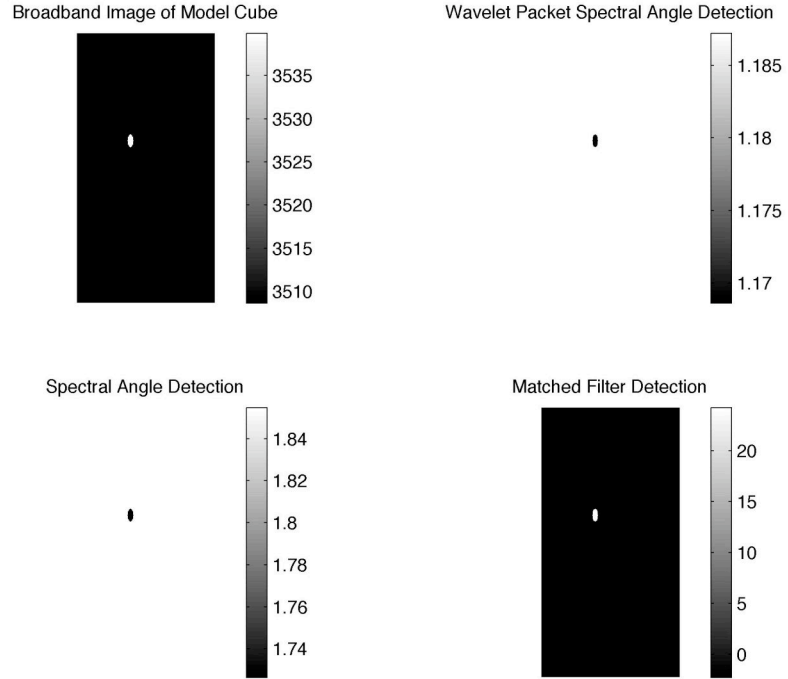


Figure 38. SO₂ Detection Map for SDC 1

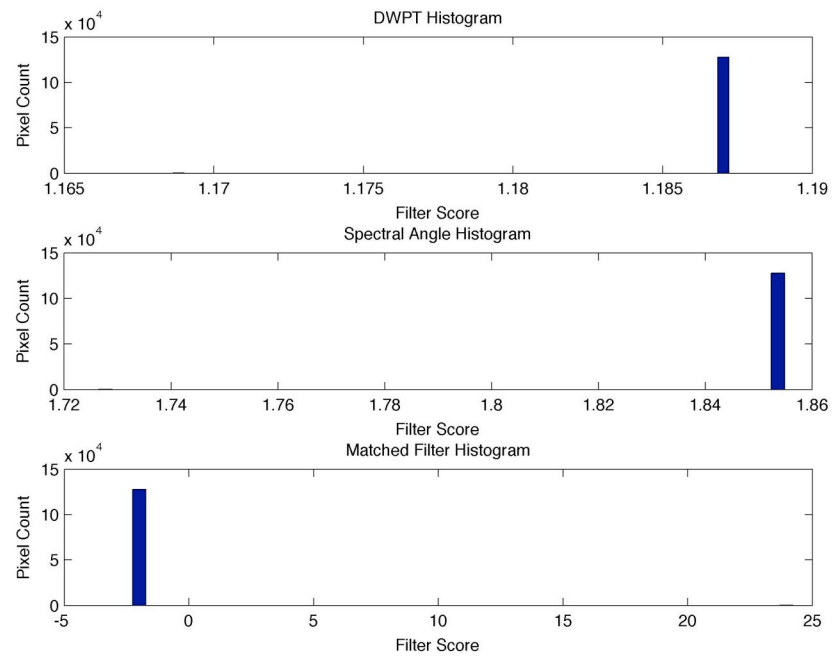


Figure 39. Histogram for SDC 1

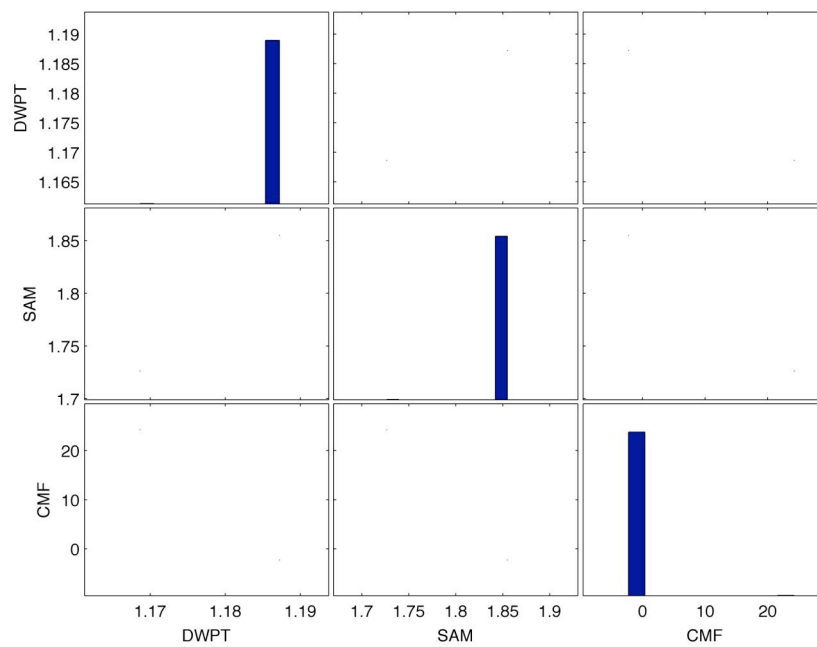


Figure 40. Scatter Plot for SDC 1

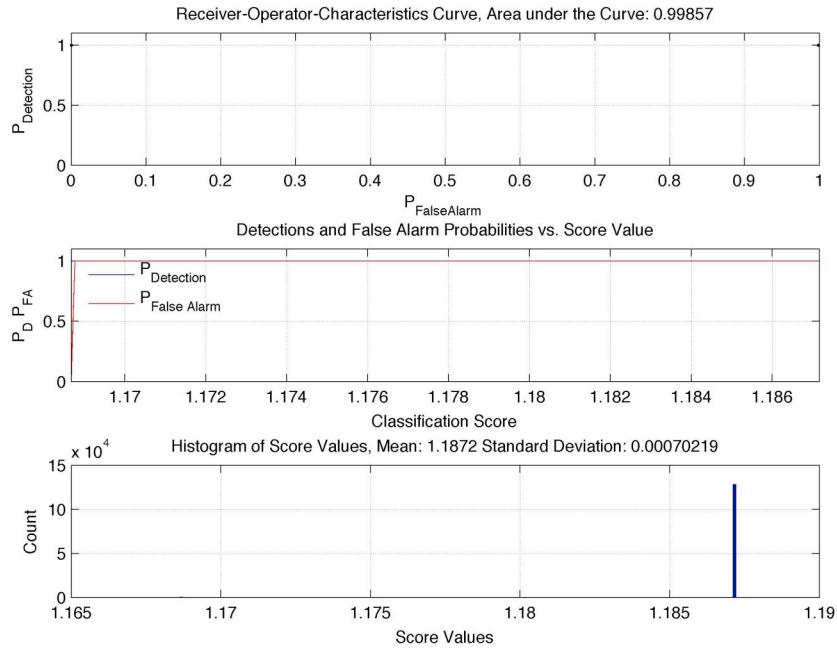


Figure 41. WPS ROC Curve for SDC 1

SDC 2, SO_2 with low noise background, has similar results to SDC 1. The target and background are visually detected and easily separated as seen in Figure 42. Normally distributed random noise is added to the background. Noise is added to all pixels prior to adding the target radiance. The results of the noise are clear in the data statistics. The histogram and scatter plot, (Figure 44 and Figure 45) now have a distribution of scores for both the target and background pixels. A classification map for the WPS is added to illustrate target separation at 1σ , 2σ , and 3σ in Figure 43. At 3σ the target is clearly separated from the background. The histogram and scatter plot also show clear separation between the target and background. All three ROC curves demonstrate the full detection of the target prior to any false alarms as illustrated by Figures 46-48.

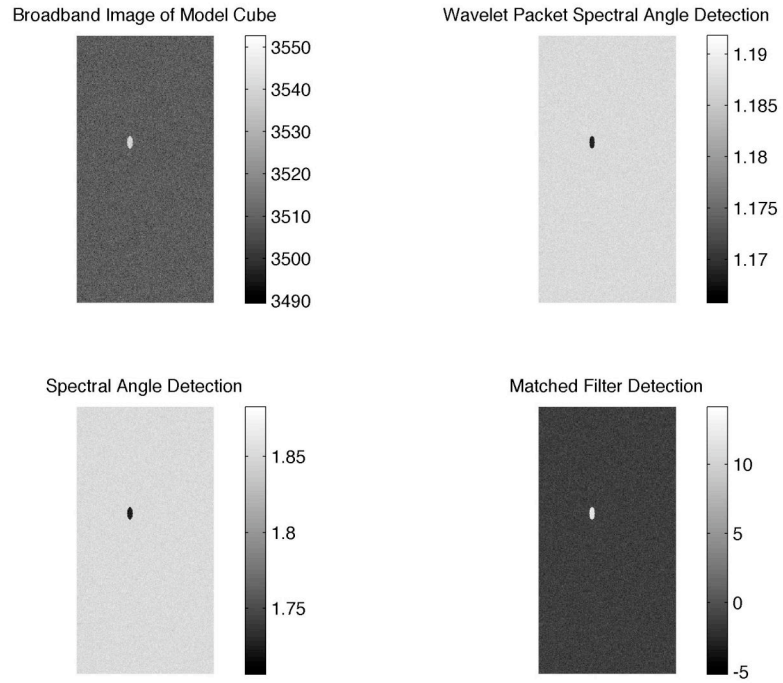


Figure 42. Detection Map for SDC 2 (SO_2)

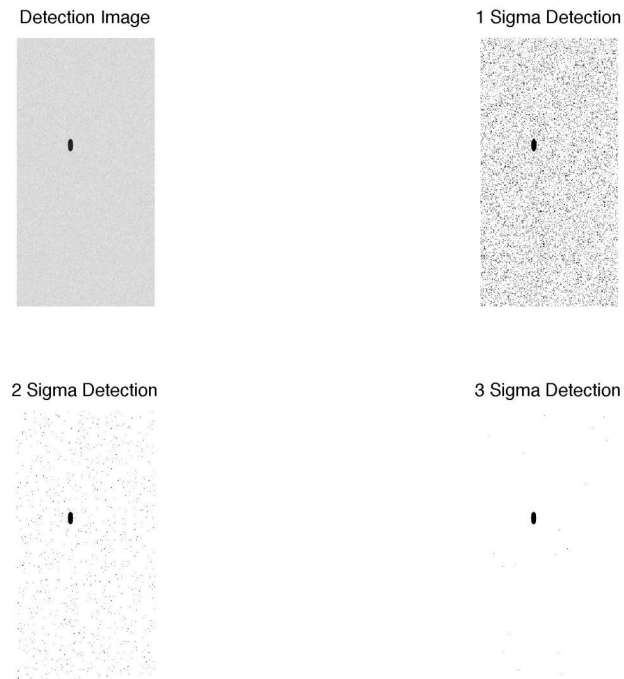


Figure 43. WPS Classification Map for SDC 2 (SO_2)

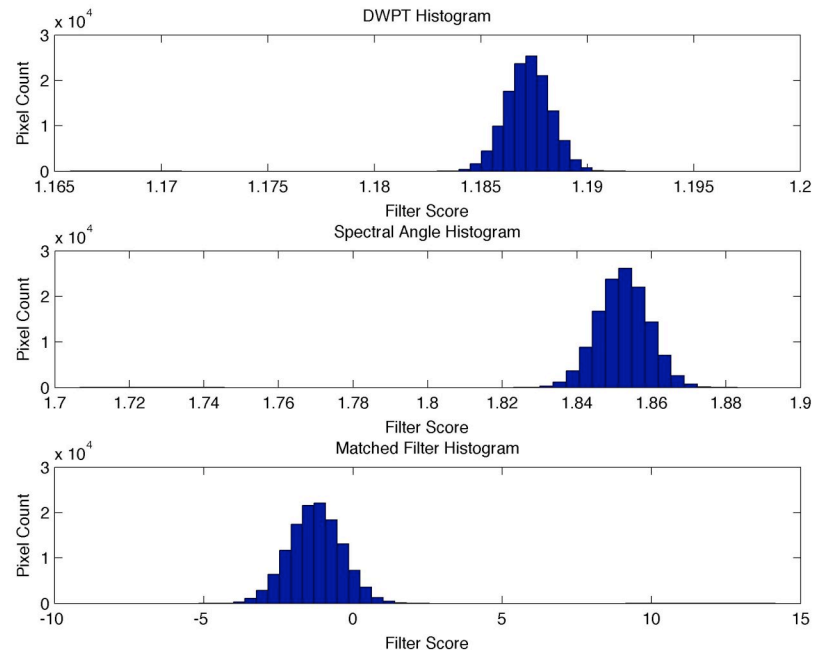


Figure 44. Histogram for SDC 2 (SO_2)

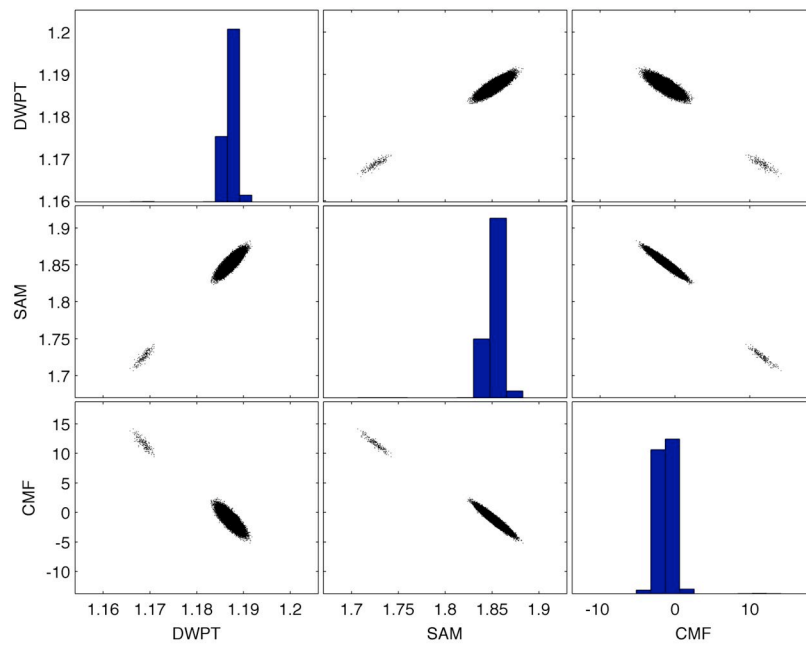


Figure 45. Scatter Plot for SDC 2 (SO_2)

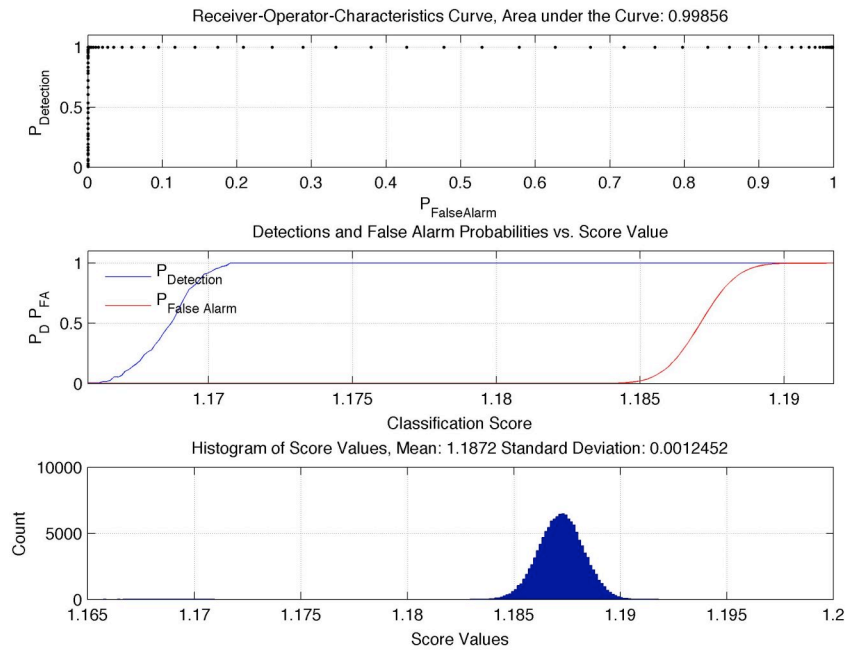


Figure 46. WPS ROC Curve for SDC 2 (SO_2)

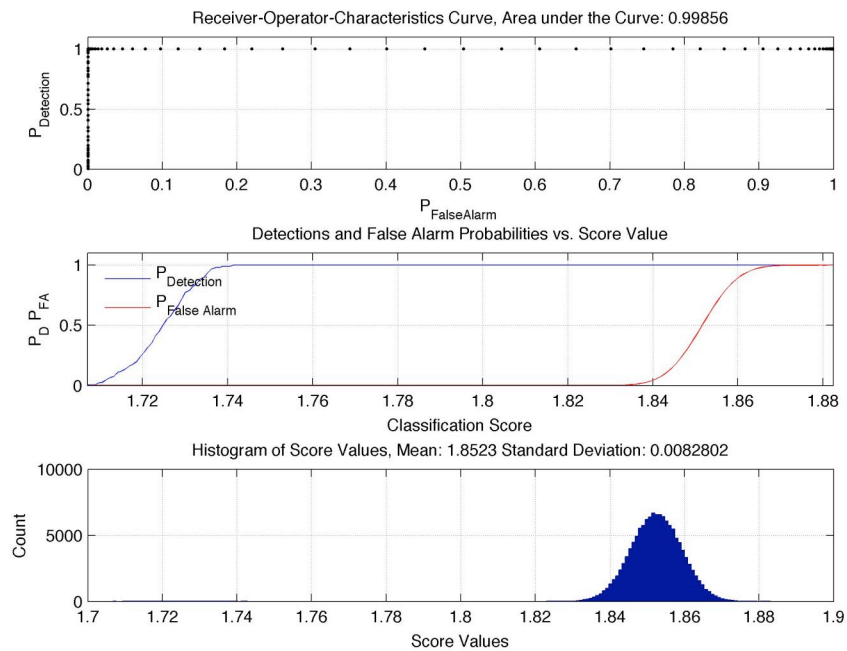


Figure 47. SAM ROC Curve for SDC 2 (SO_2)

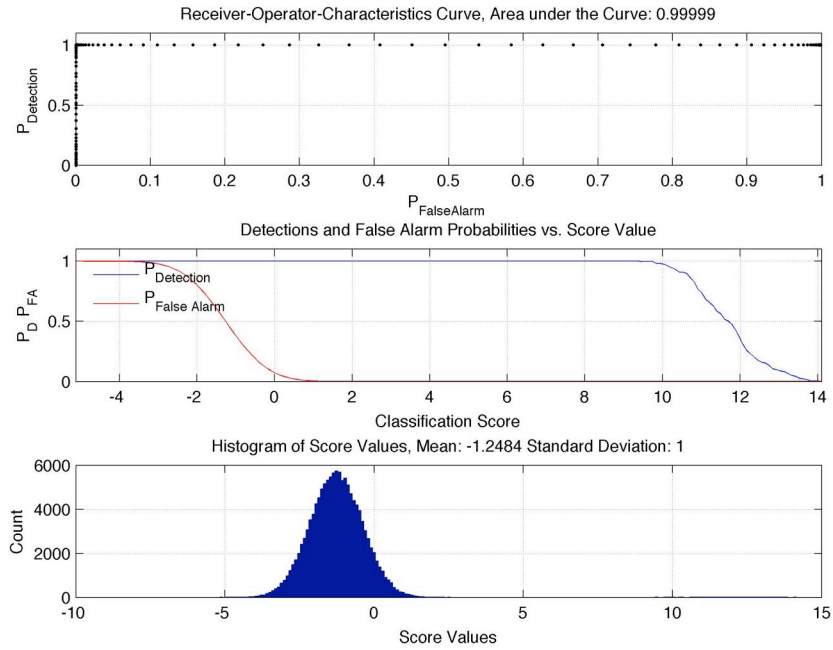


Figure 48. CMF ROC Curve for SDC 2 (SO_2)

SDC 2A, SO_2 with high noise background, begins to add enough noise to reduce the target and background separation. All three algorithms can still separate the target and background as evidenced by the detection image and data statistics in Figure 49, Figure 51 and Figure 52. In the classification map, Figure 50, erroneous pixels begin to appear at the 3σ level. Only the WPS ROC curve is shown as all three algorithms have good performance, see Figure 53. The SAM and CMF ROC curves are almost identical. All three algorithms can separate target from background with detections and false alarms clearly separated.

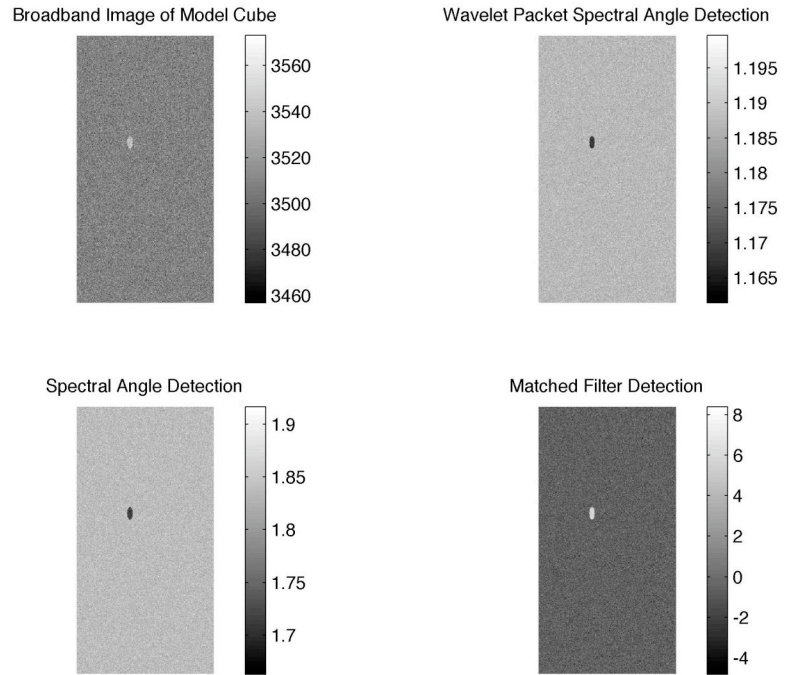


Figure 49. Detection Map for SDC 2A (SO_2)

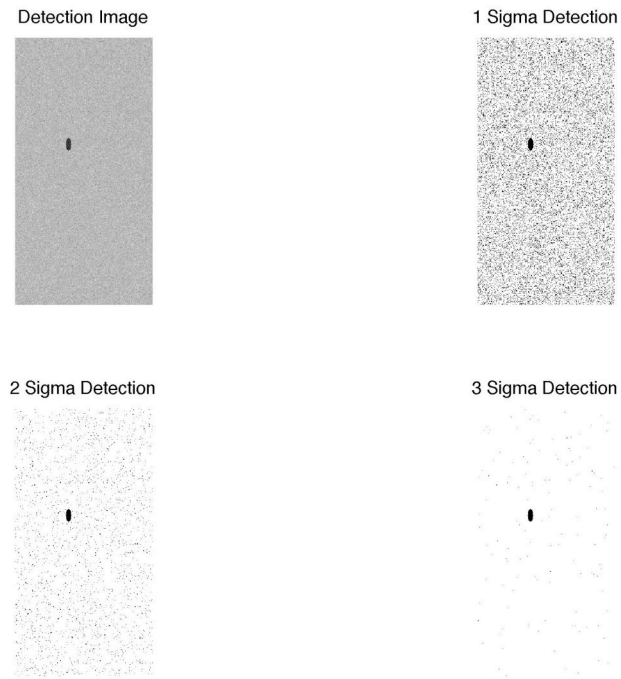


Figure 50. Classification Map for SDC 2A (SO_2)

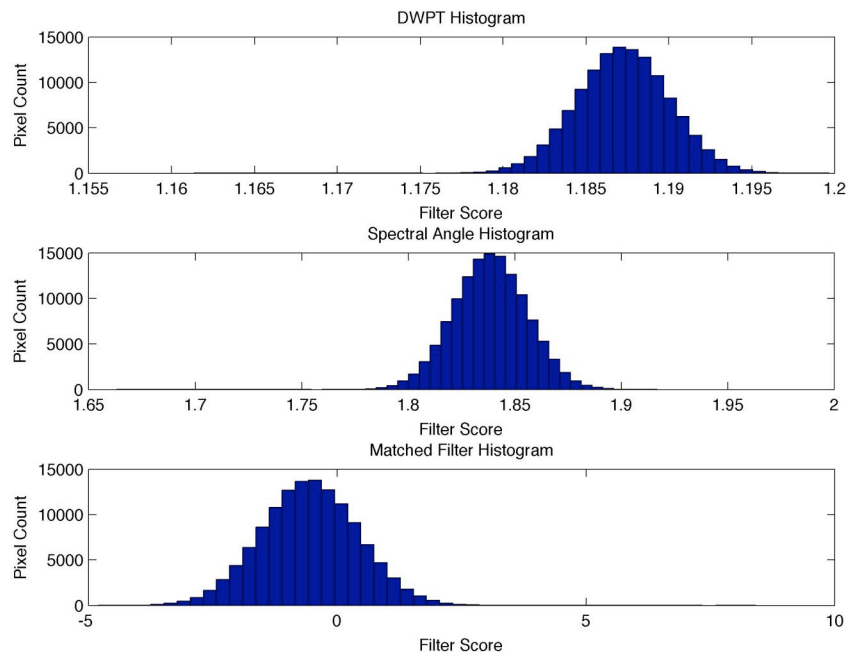


Figure 51. Histogram for SDC 2A (SO_2)

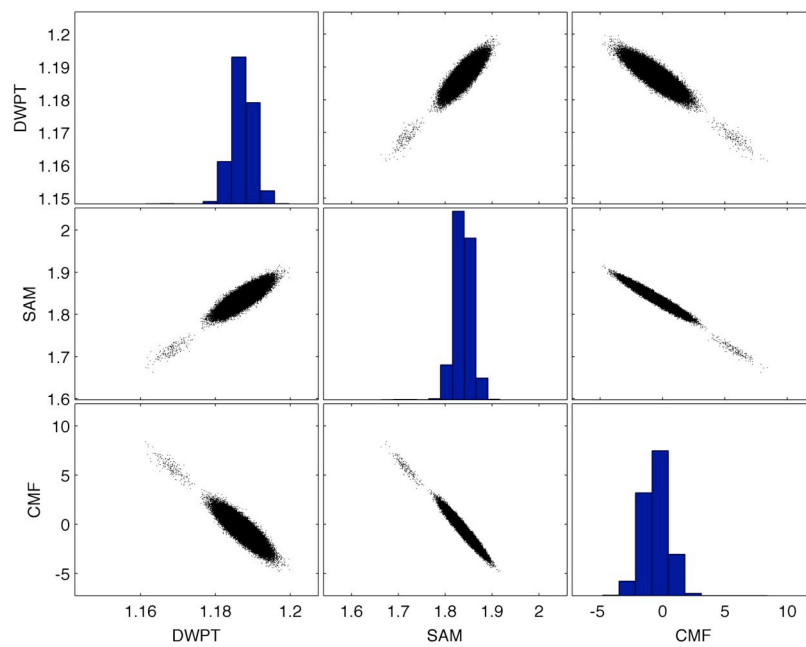


Figure 52. Scatterplot for SDC 2A (SO_2)

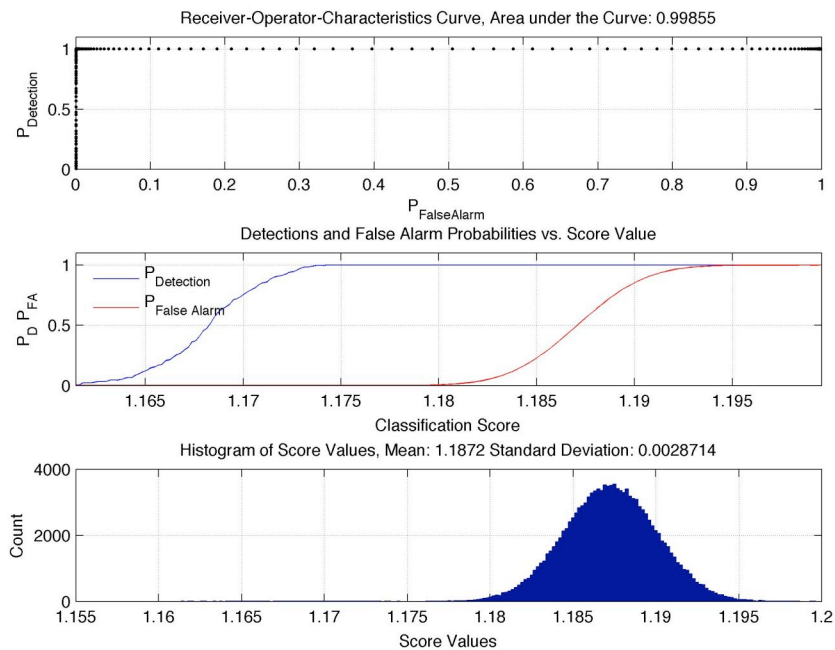


Figure 53. WPS ROC Curve for SDC 2A (SO_2)

SDC 3 adds a second target plume to the scene, (SO_2 and CH_4 with low noise homogenous background). All three algorithms do well against SO_2 but CMF begins to show significant degradation against CH_4 . The SO_2 target is separable both visually and statistically for all three (Figures 54-57), while the CH_4 target is inseparable for CMF, (Figures 59-62). WPS classification at 3σ separates SO_2 and CH_4 from the background and from each other. Below 3σ , some false alarms occur among the background and CH_4 . Only the WPS ROC curve is shown for SO_2 as SAM and CMF have similar performance, see Figure 58. For CH_4 all three ROC curves are shown in Figures 63-65.

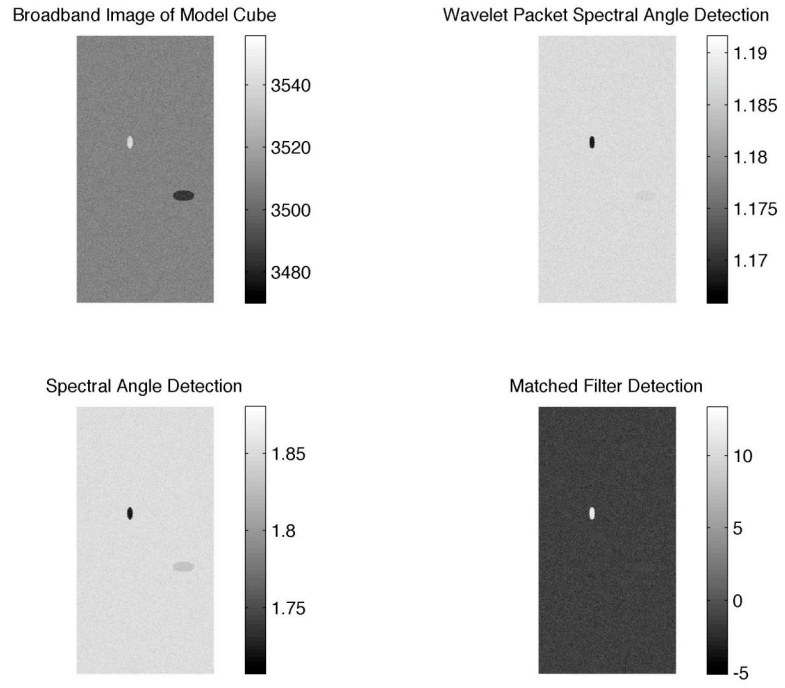


Figure 54. Detection Map for SDC 3 (SO_2)

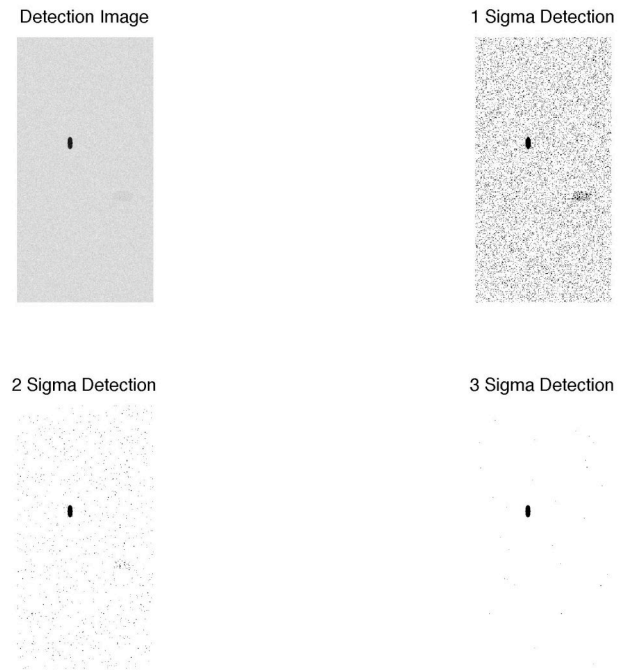


Figure 55. WPS Classification Map for SDC 3 (SO_2)

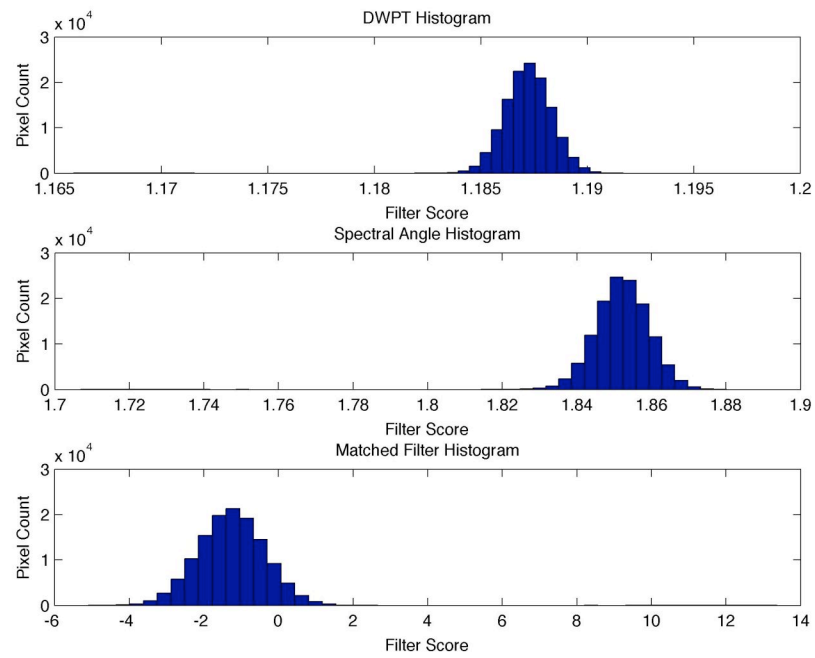


Figure 56. Histogram for SDC 3 (SO_2)

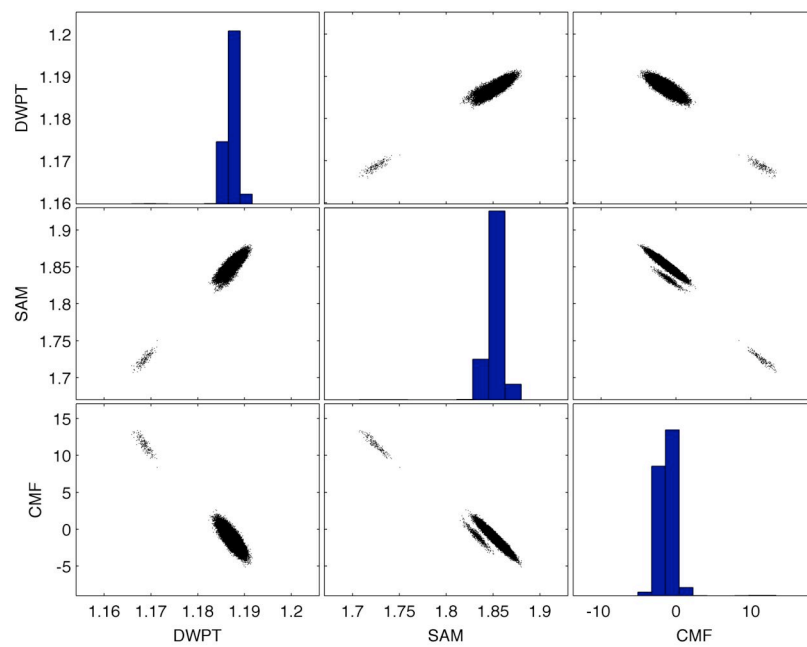


Figure 57. Scatterplot for SDC 3 (SO_2)

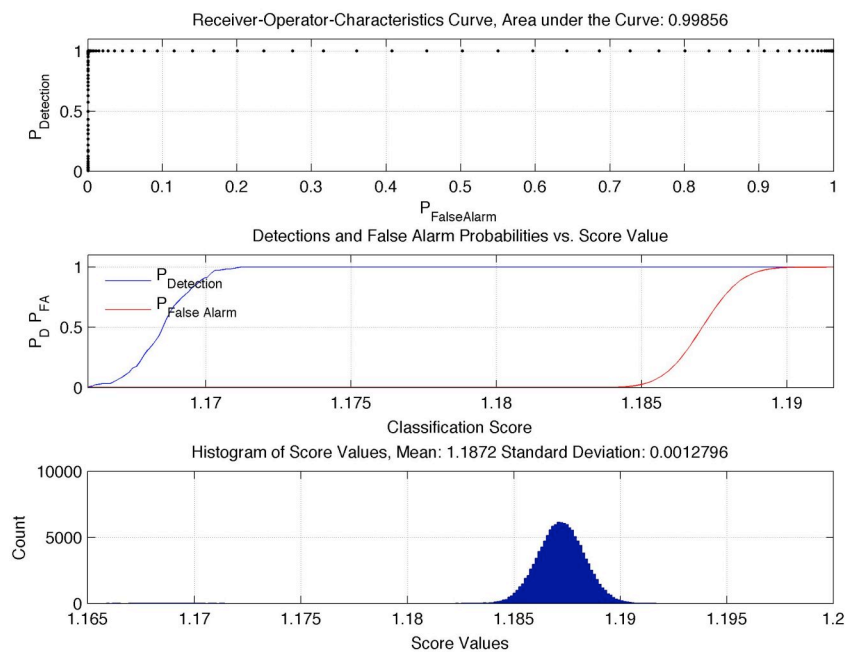


Figure 58. WPS ROC Curve for SDC 3 (SO₂)

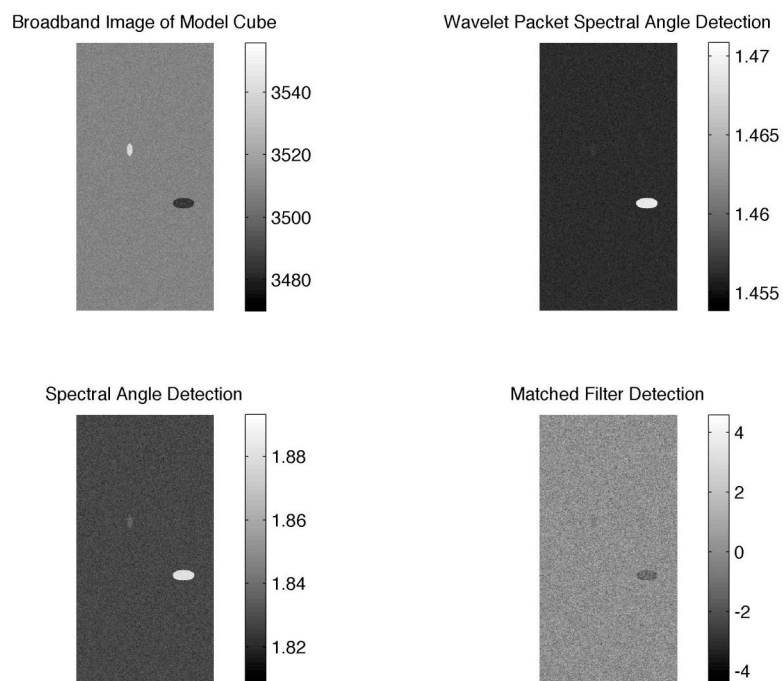


Figure 59. Detection Map for SDC 3 (CH₄)

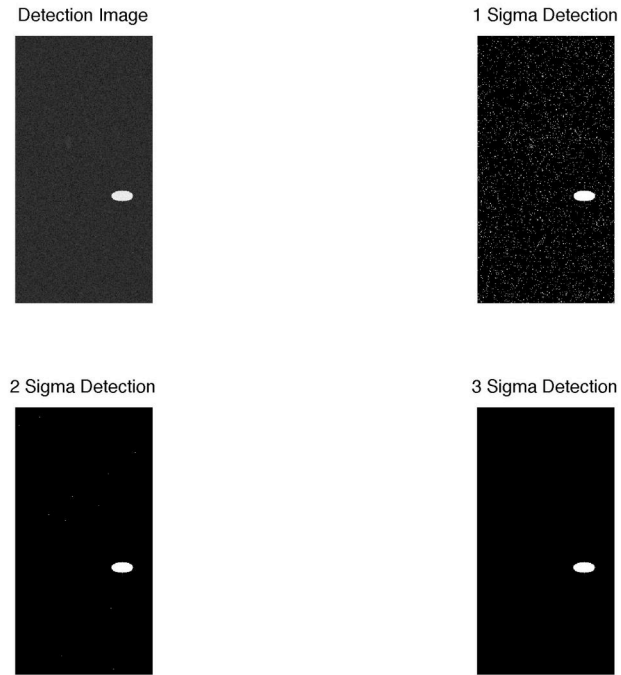


Figure 60. WPS Classification Map for SDC 3 (CH_4)

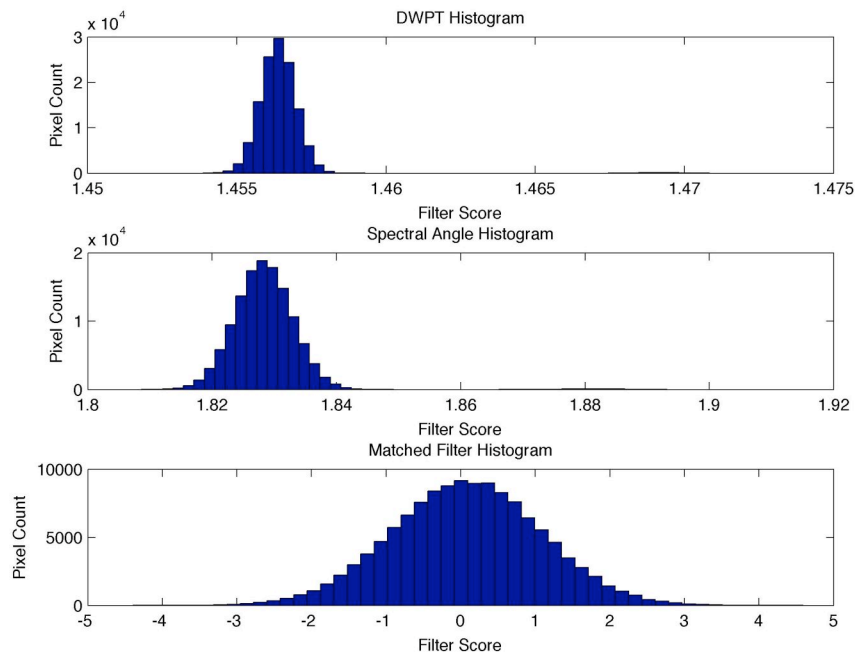


Figure 61. Histogram for SDC 3 (CH_4)

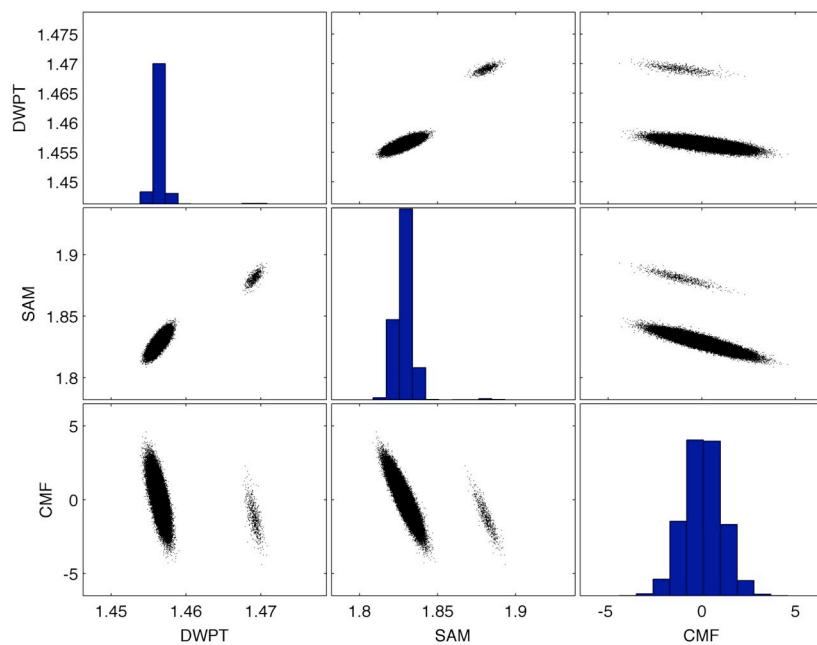


Figure 62. Scatterplot for SDC 3 (CH₄)

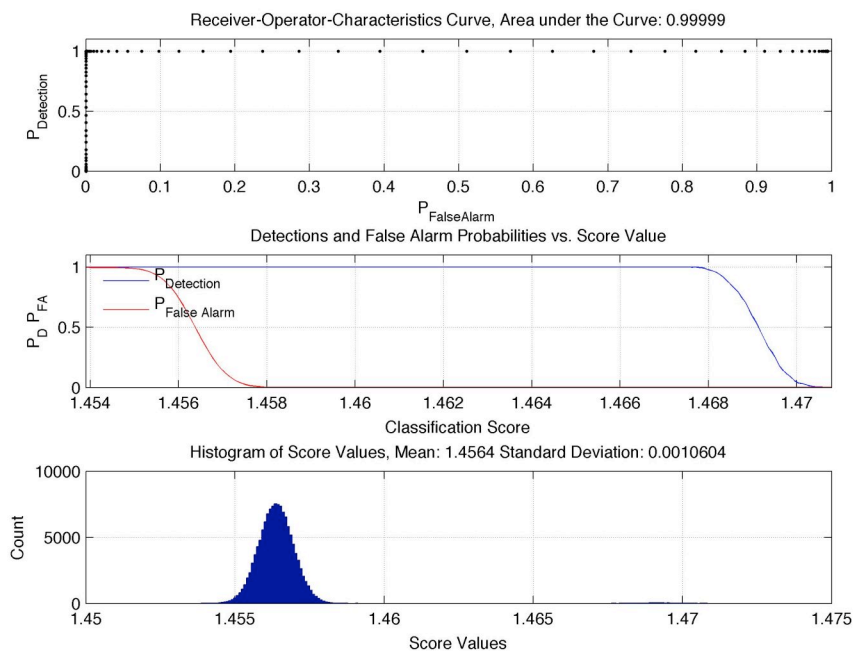


Figure 63. WPS ROC Curve for SDC 3 (CH₄)

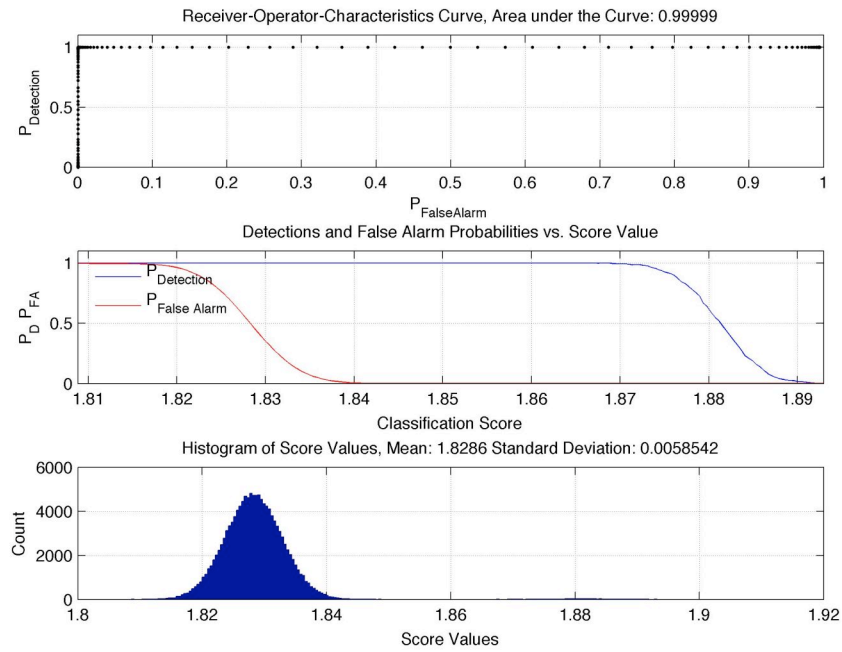


Figure 64. SAM ROC Curve for SDC 3 (CH_4)

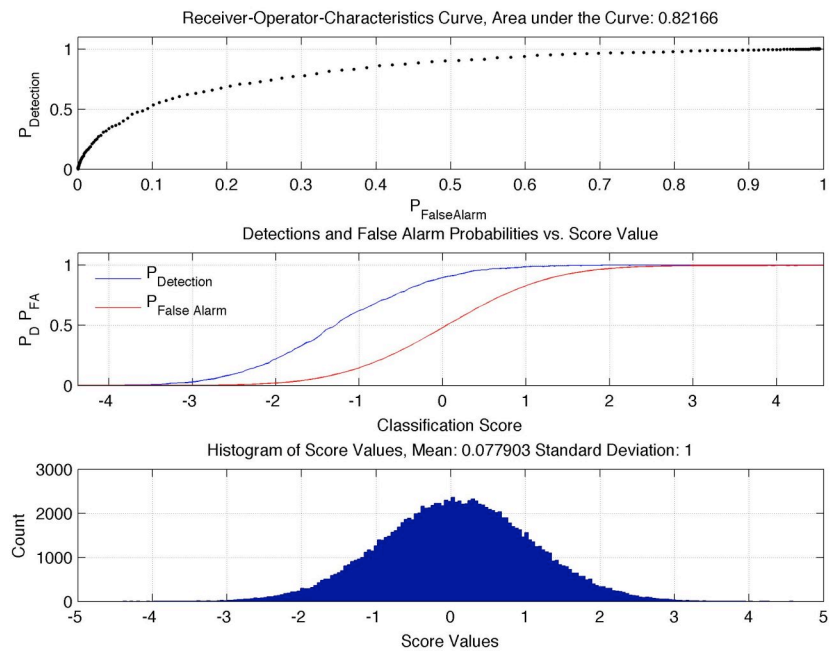


Figure 65. CMF ROC Curve for SDC 3 (CH_4)

SDC 4 co-locates the target plumes of SDC 3. The background noise level is the same as SDC 3. SDC 4 results are similar to SDC 3 results in that all three algorithms do well against SO_2 while CMF performance is degraded against CH_4 . SO_2 results are shown in Figures 66-69. There are no other noteworthy differences between SDC 3 and SDC 4 (Results illustrated in Figures 71-74). Again, only the WPS ROC curve is shown for SO_2 as SAM and CMF have similar performance, see Figure 70. For CH_4 all three ROC curves are shown in Figures 75-77.

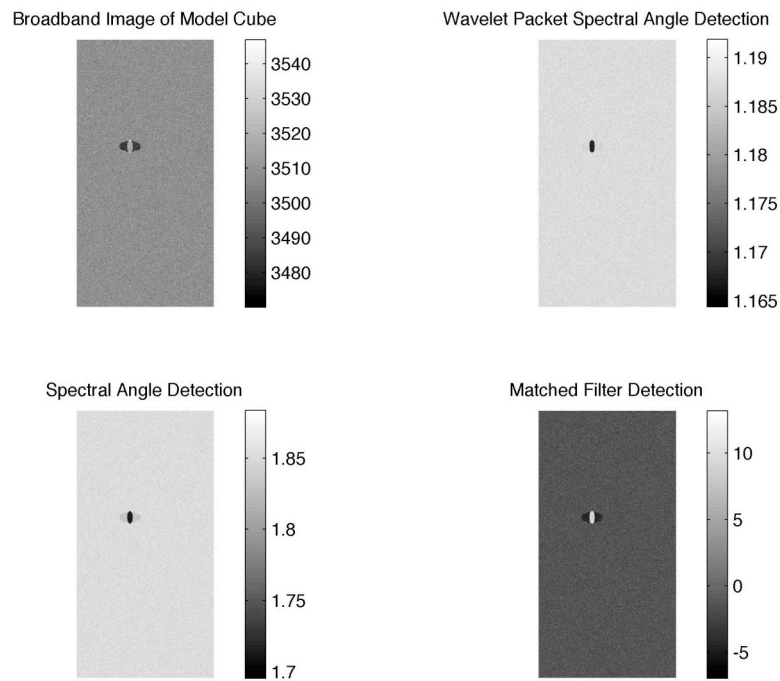


Figure 66. Detection Map for SDC 4 (SO_2)

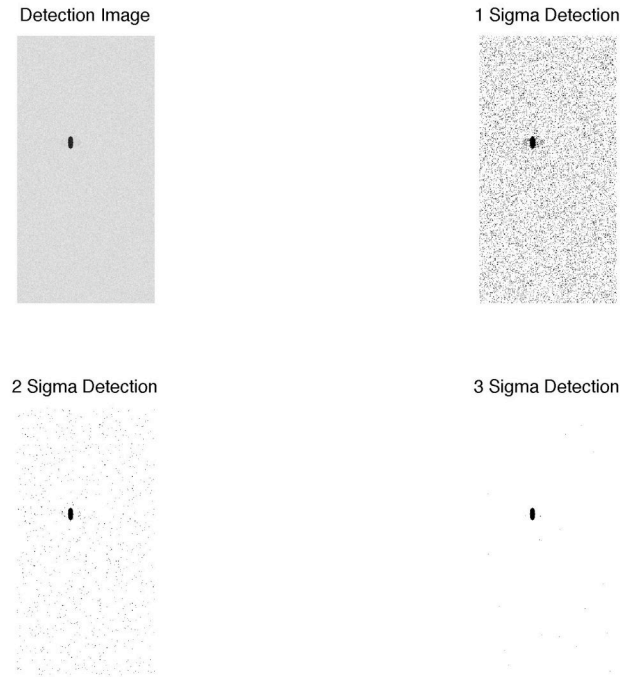


Figure 67. WPS Classification Map for SDC 4 (SO_2)

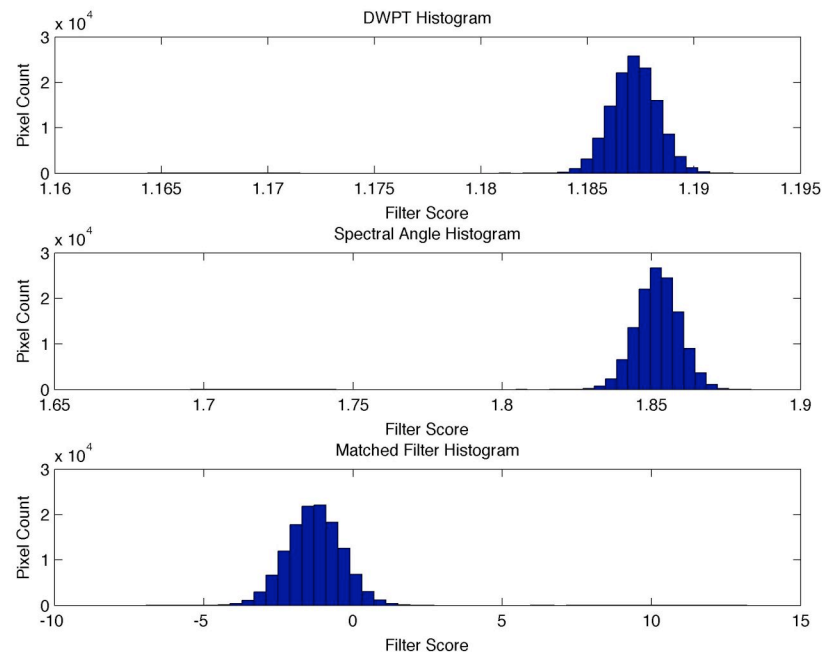


Figure 68. Histogram for SDC 4 (SO_2)

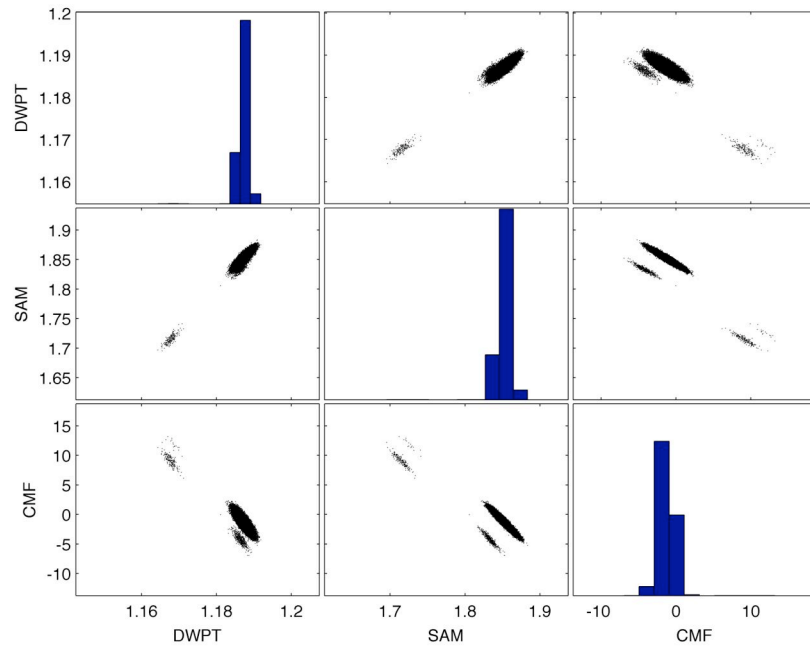


Figure 69. Scatterplot for SDC 4 (SO₂)

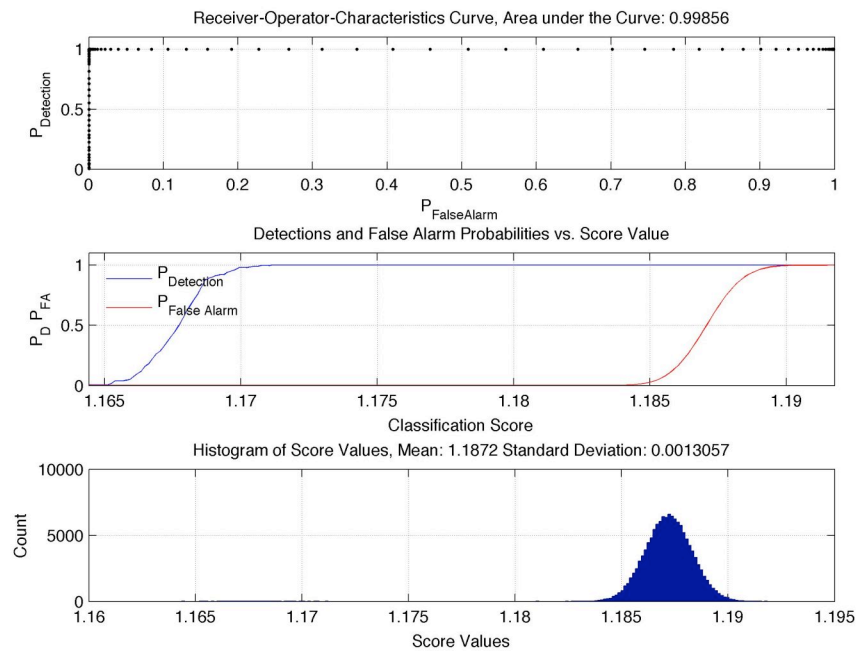


Figure 70. WPS ROC Curve for SDC 4 (SO₂)

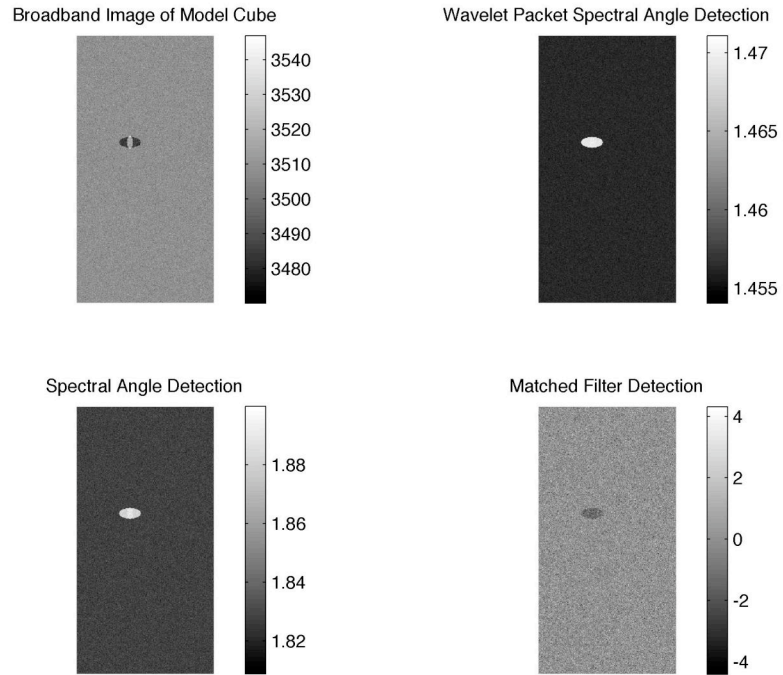


Figure 71. Detection Map for SDC 4 (CH_4)

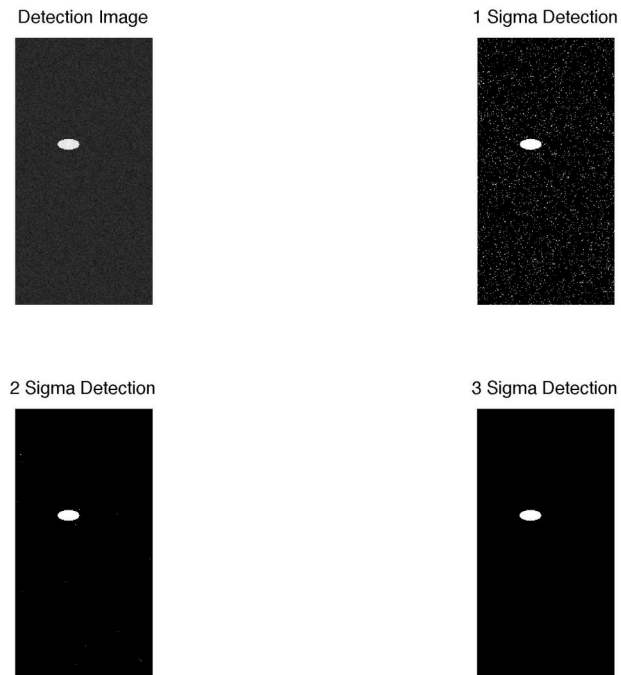


Figure 72. Classification Map for SDC 4 (CH_4)

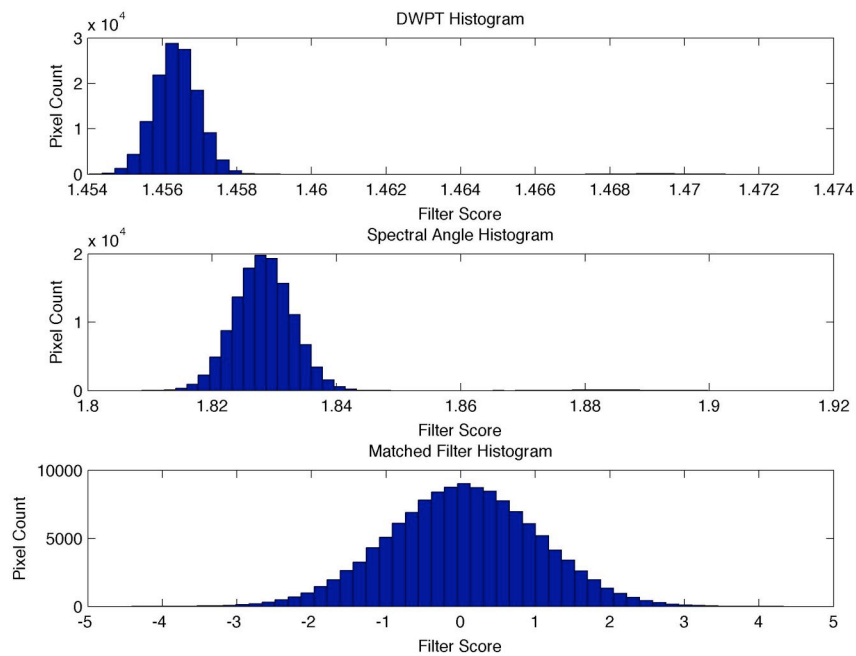


Figure 73. Histogram for SDC 4 (CH_4)

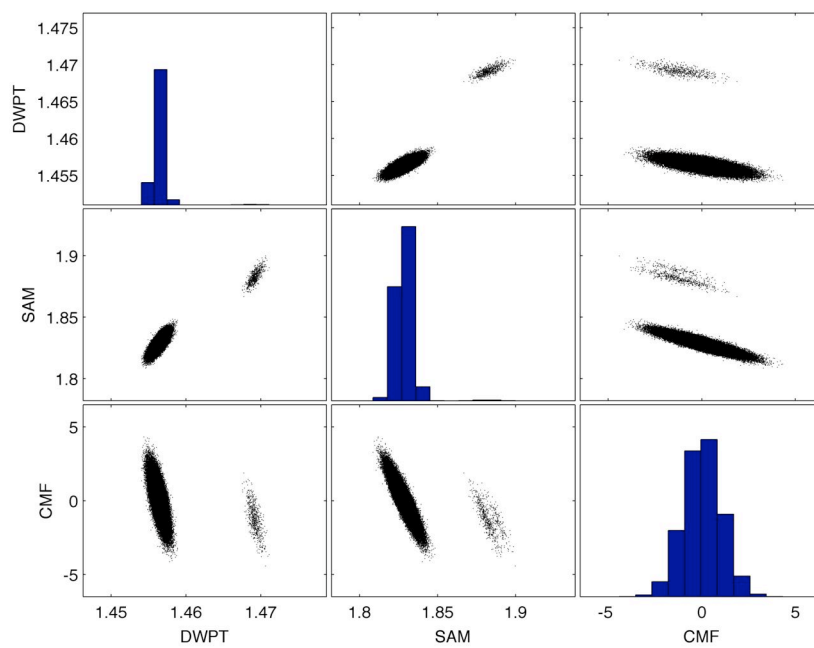


Figure 74. Scatterplot for SDC 4 (CH_4)

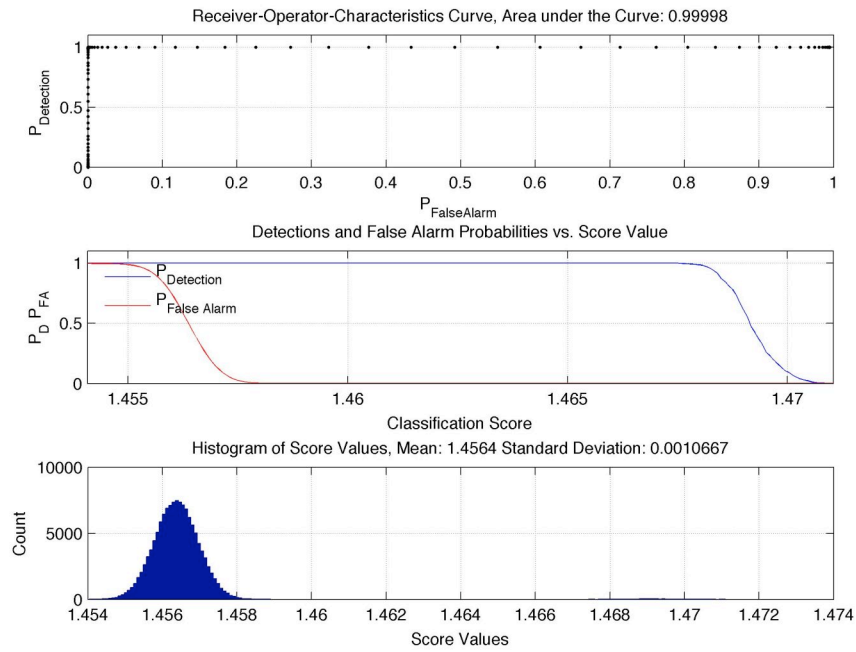


Figure 75. WPS ROC Curve for SDC 4 (CH₄)

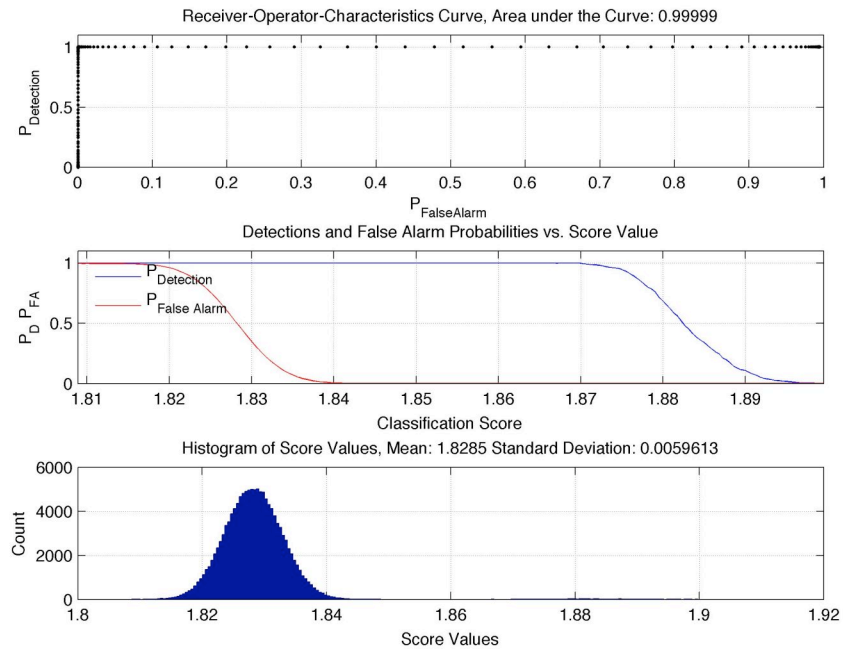


Figure 76. SAM ROC Curve for SDC 4 (CH₄)

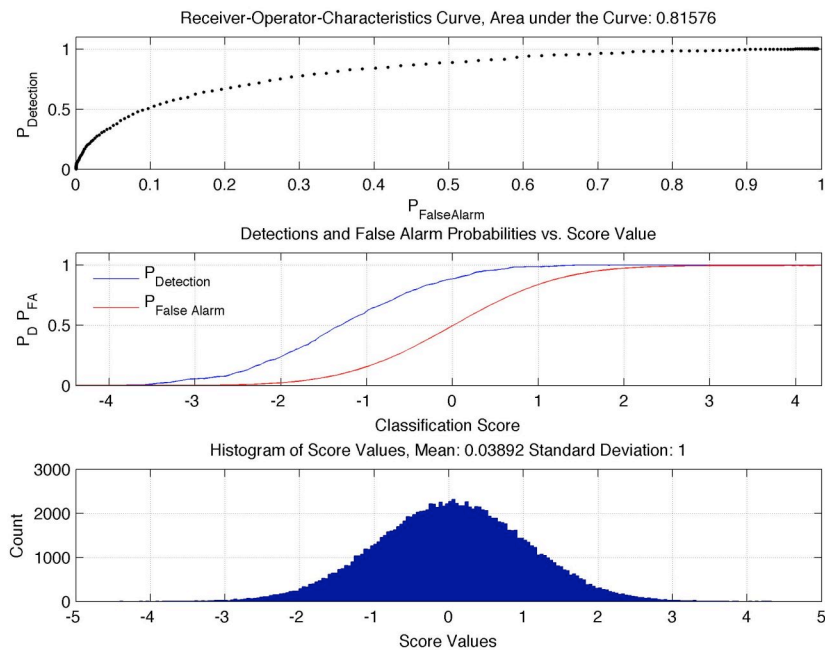


Figure 77. CMF ROC Curve for SDC 4 (CH₄)

SDC 4A increases the background noise level of SDC 4. SO₂ detection performance continues to do well for all algorithms. All algorithms continue to separate SO₂ from the background and the other target plume both visually and statistically. But, SDC 4A has pushed the limits of all three algorithms in CH₄ detection. None of the algorithms do well in detecting CH₄. Although visual detection is still good for all three, the data statistics indicate significant degradation. In the CH₄ ROC curves, WPS appears to have a slight advantage over SAM and CMF as evidenced when areas under the curve are compared. All other statistics indicate no separation. SO₂ results are presented in Figures 78-84. CH₄ results are presented in Figures 85-91.

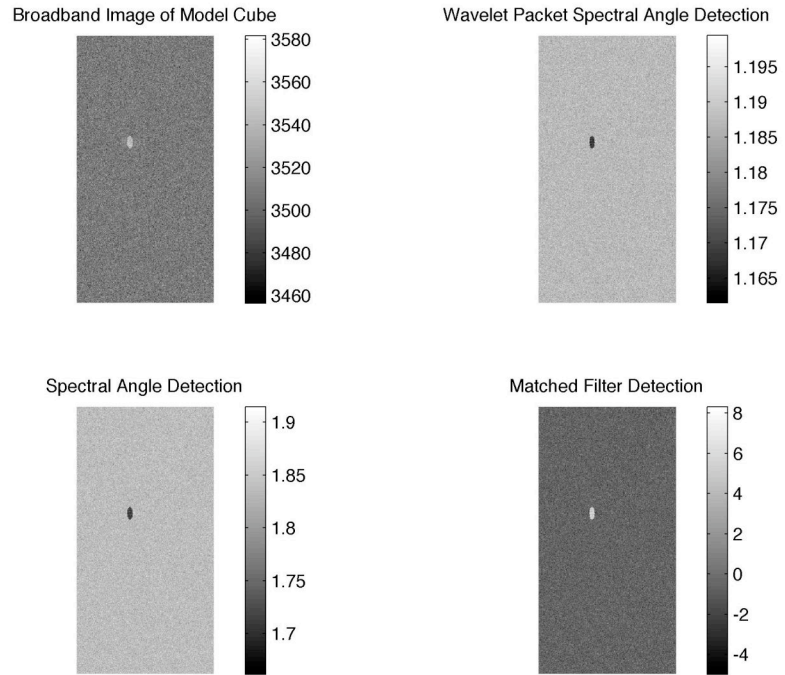


Figure 78. Detection Map for SDC 4A (SO_2)

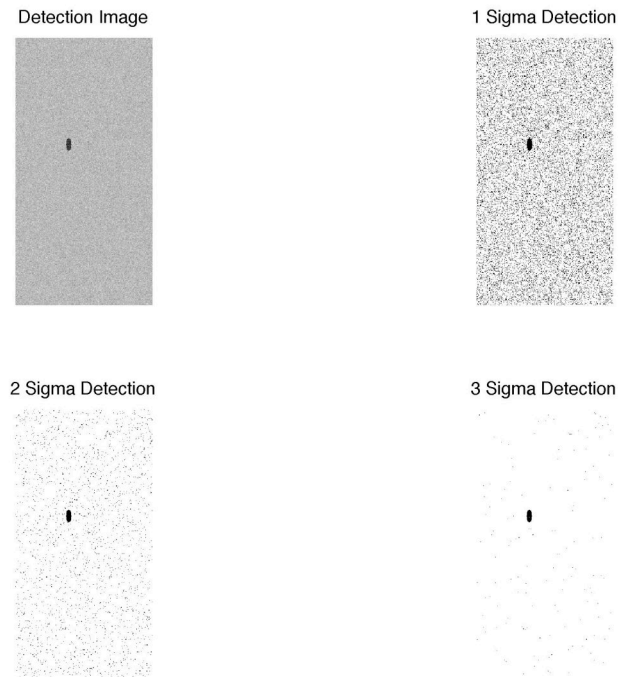


Figure 79. WPS Classification Map for SDC 4A (SO_2)

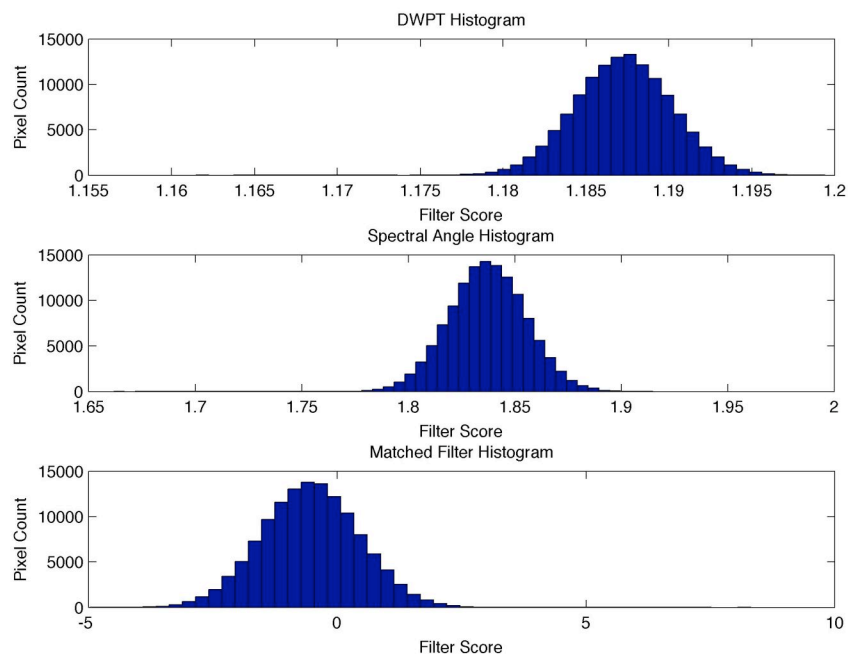


Figure 80. Histogram for SDC 4A (SO_2)

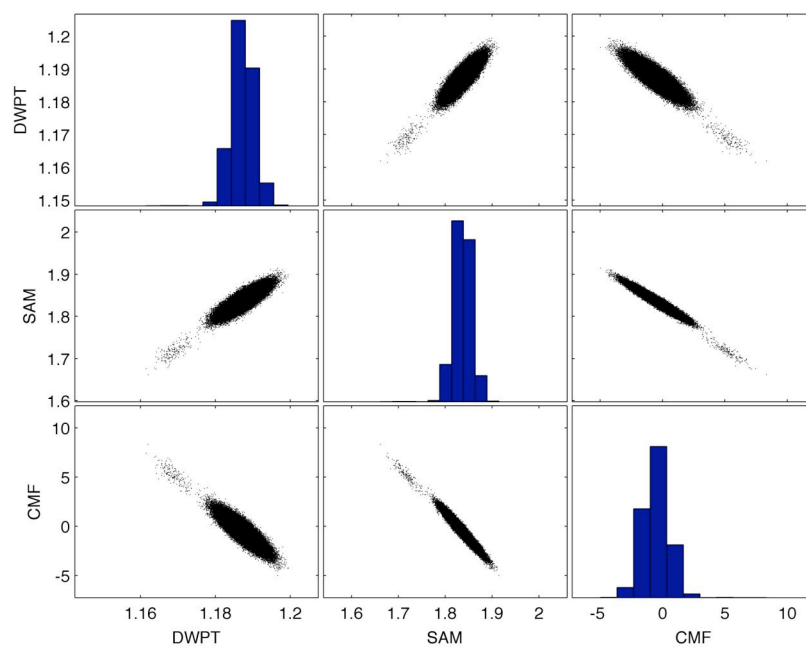


Figure 81. Scatterplot for SDC 4A (SO_2)

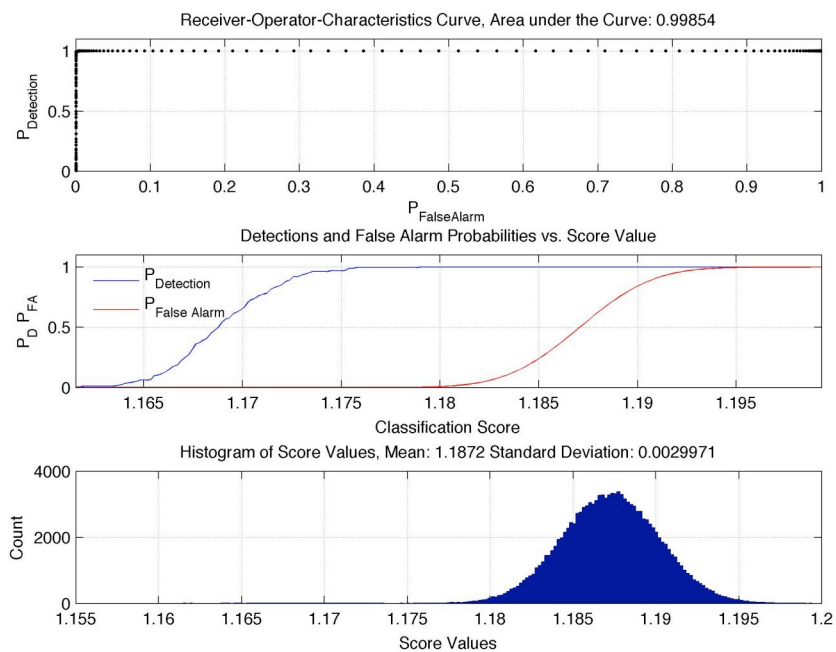


Figure 82. DWPT ROC Curve for SDC 4A (SO_2)

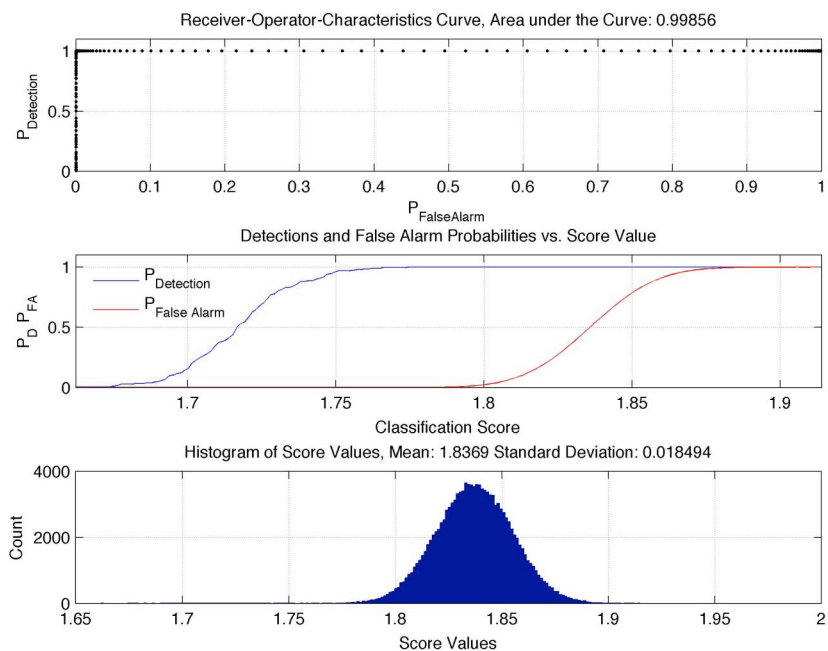


Figure 83. SAM ROC Curve for SDC 4A (SO_2)

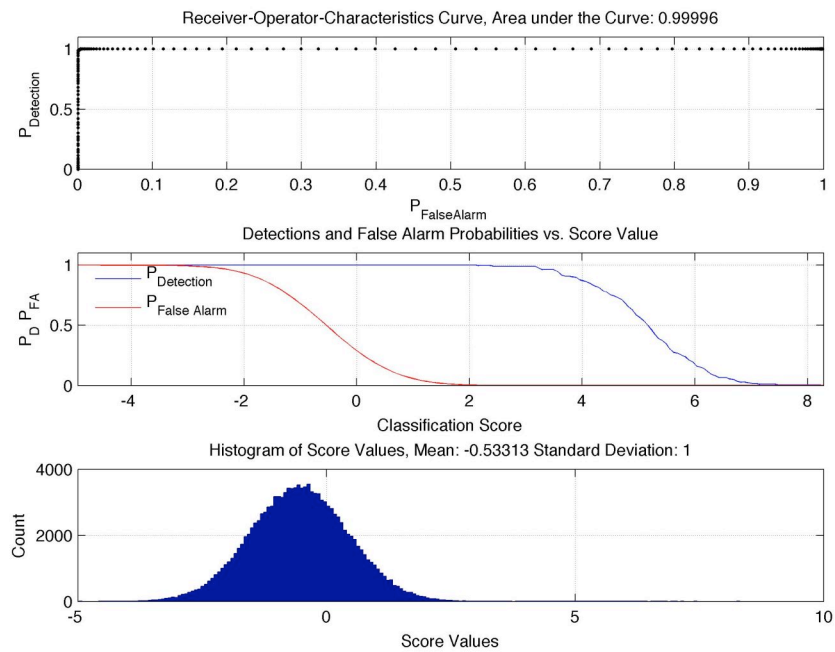


Figure 84. CMF ROC Curve for SDC 4A (SO_2)

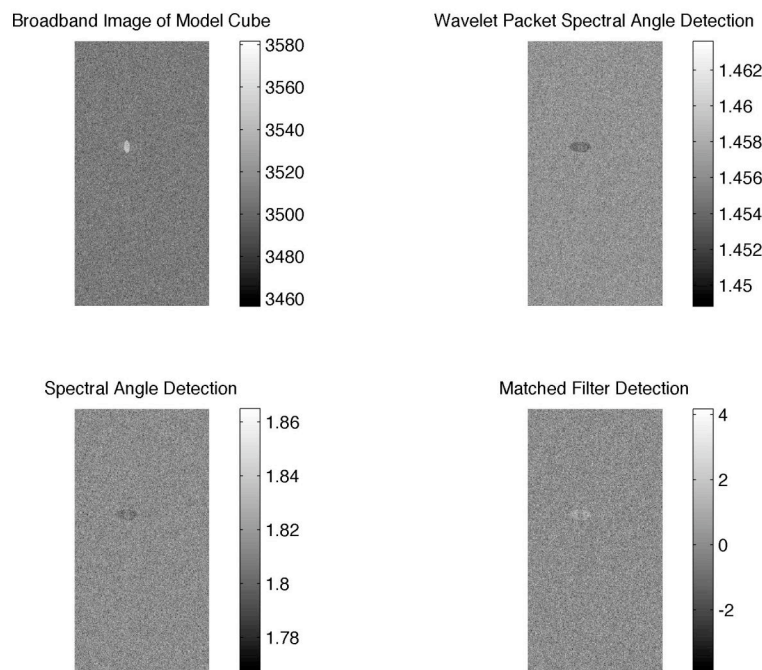


Figure 85. Detection Map for SDC 4A (CH_4)

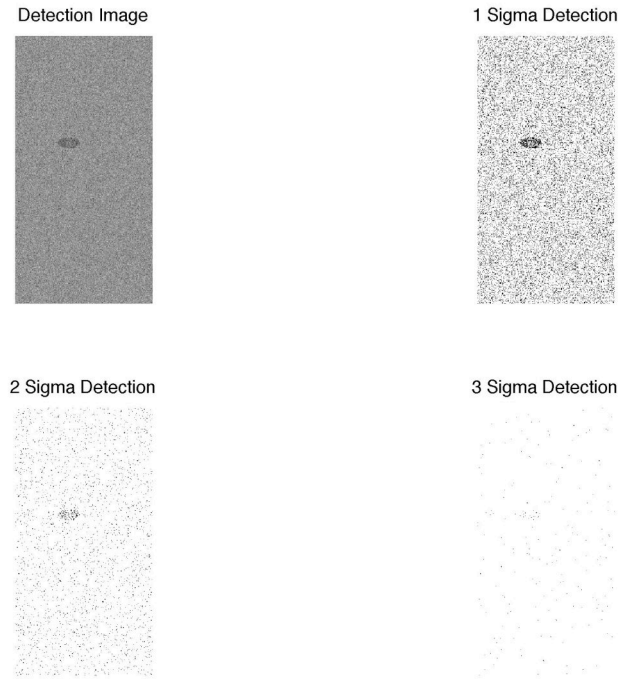


Figure 86. WPS Classification Map for SDC 4A (CH_4)

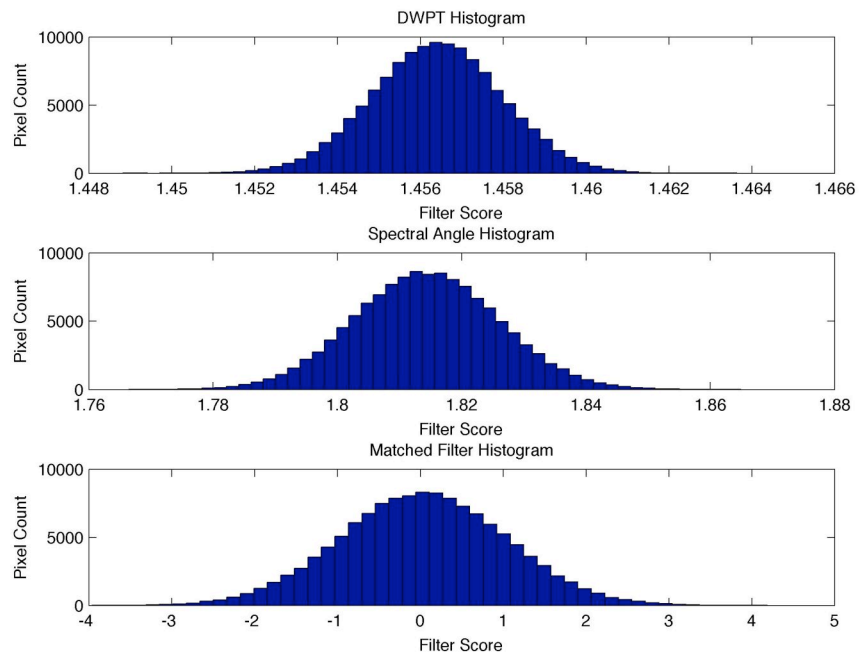


Figure 87. Histogram for SDC 4A (CH_4)

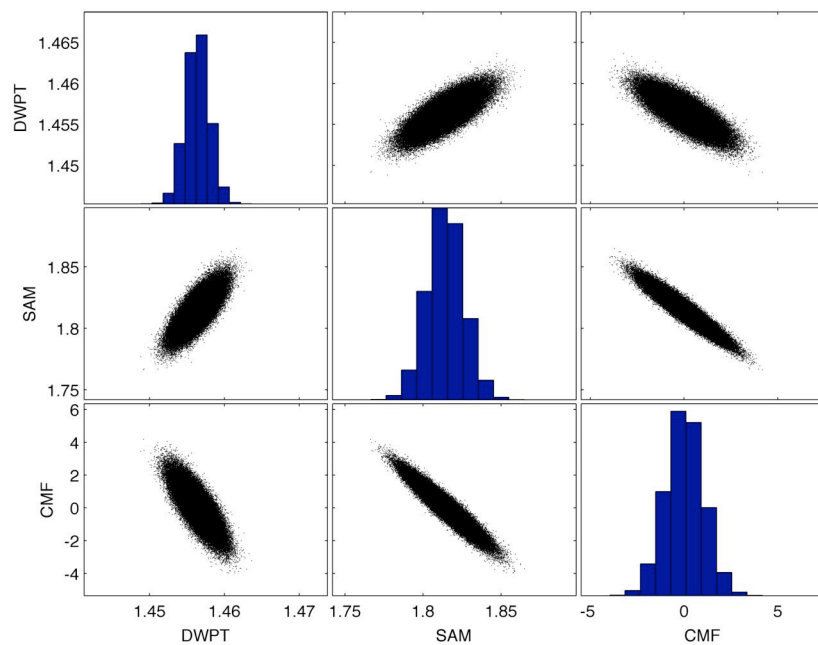


Figure 88. Scatterplot for SDC 4A (CH₄)

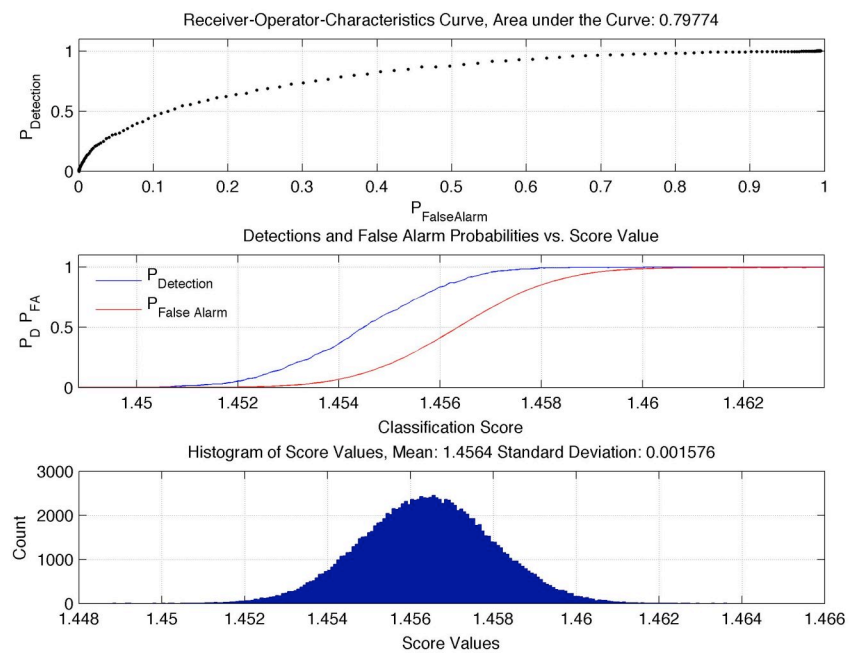


Figure 89. WPS ROC Curve for SDC 4A (CH₄)

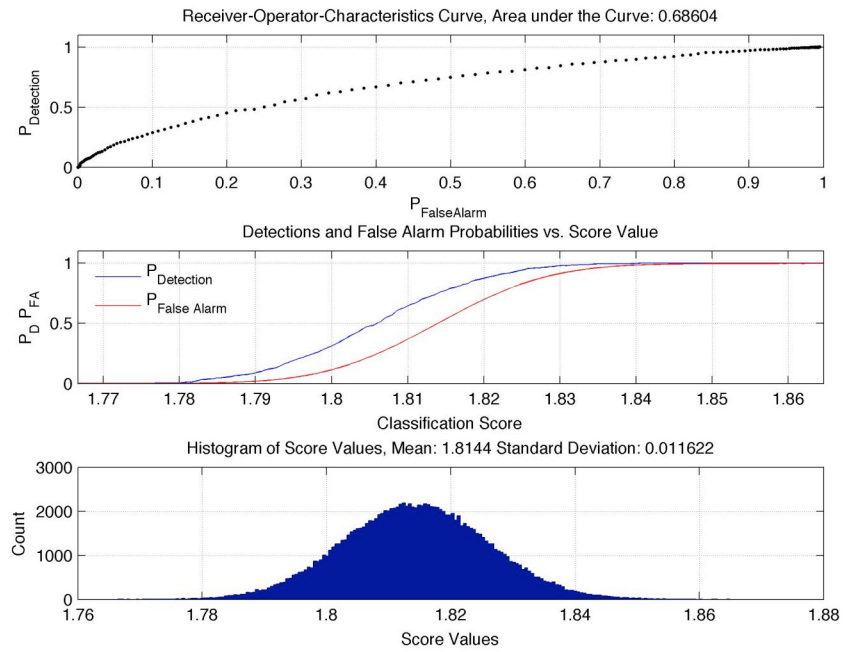


Figure 90. SAM ROC Curve for SDC 4A (CH_4)

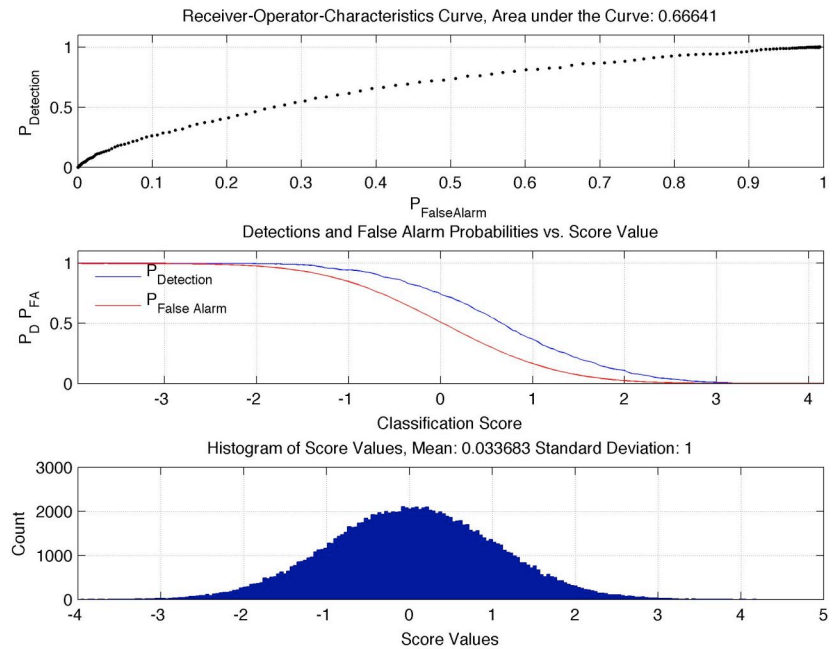


Figure 91. CMF ROC Curve for SDC 4A (CH_4)

3.2 Airborne Hyperspectral Imager Results

The focus of the AHI data analysis is plume detection in the wavelet packet subspace. In typical LWIR analysis, the detection step is followed by an identification step. The identification step uses linear regression to confidently identify the spectra. This step is left for discussion in Chapter 4, while the results herein focus on the ability of the WPS to detect a potential gas emission.

The three AHI flightlines analyzed represent both known and unknown gas emissions. Flightline 121247 contains a known SO₂ plume emitted from a power generation facility. Flightlines 105948 and 153924 are collected over petroleum refineries and were expected to have gas plumes in them. In the cases where unknown plumes may be present an additional identification step is applied to the suspected gas plume to assist in gas identification.

Results for all three include the following figures:

1. Detection Planes for SAM, CMF and WPS – These represent a detection for a specific target gases.
2. Selected Detection Planes (Color Maps) – These are the planes that best represent a suspected gas plume.
3. Scatter Plot – The scatter plots illustrate the general differences in the algorithm scores for the selected detection plane.

4. Classification Map – The classification map includes the 1σ , 2σ , and 3σ classified WPS detection planes and indicates the plume selected for identification. These figures are provided in Chapter 4.

AHI 121247 has a known SO₂ plume. All three algorithms were used with multiple target gases to visually compare detection planes. The SAM results were disappointing since no visible plumes are evident see Figure 92. Both CMF and WPS did have potential plumes in multiple image planes as shown in Figure 93 and Figure 94. The strongest visually for CMF was in the isobutane (C₄H₁₀) and propane (C₃H₈) images followed by the SO₂ image. The strongest visually for WPS image planes were the sulfur dioxide and isobutane planes. The SO₂ planes were selected for the identification step. The identification process with linear regression is discussed in Chapter 4. Visual comparison across the SO₂ detection planes, (SAM, WPS, and CMF) clearly show superior performance of the WPS plane versus the other two, Figure 95. A scatter plot for the selected detection plane is shown in Figure 96. The SO₂ detection plane was chosen for the identification step because of the visual impact of the WPS image as well as the *a priori* knowledge of SO₂.

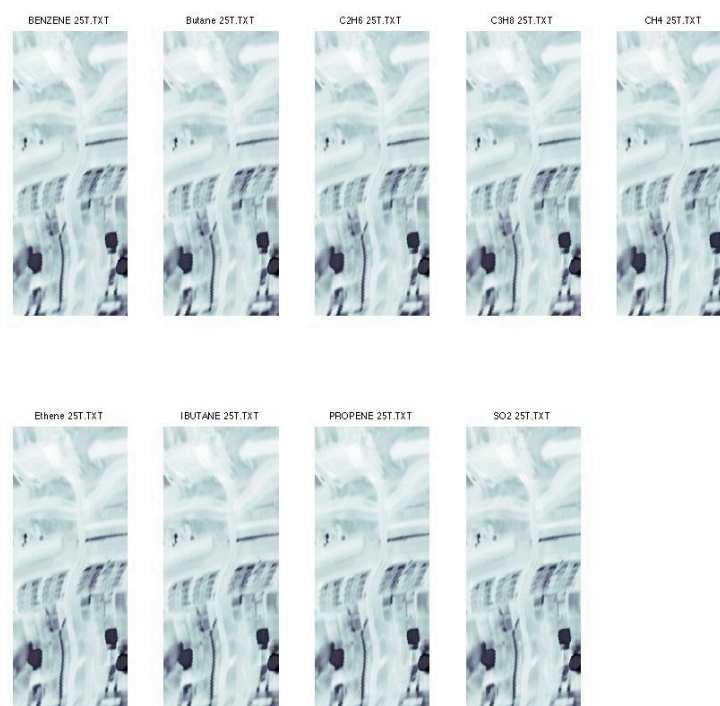


Figure 92. SAM Detection Planes for AHI 121247

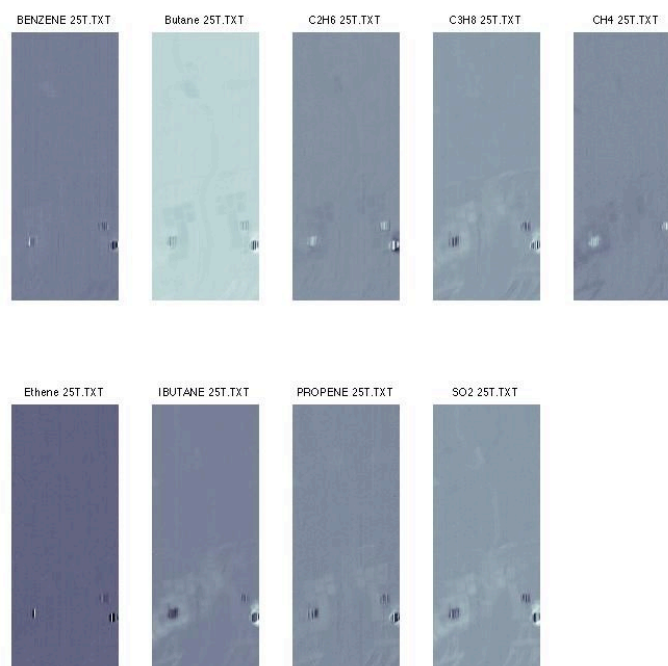


Figure 93. CMF Detection Planes for AHI 121247

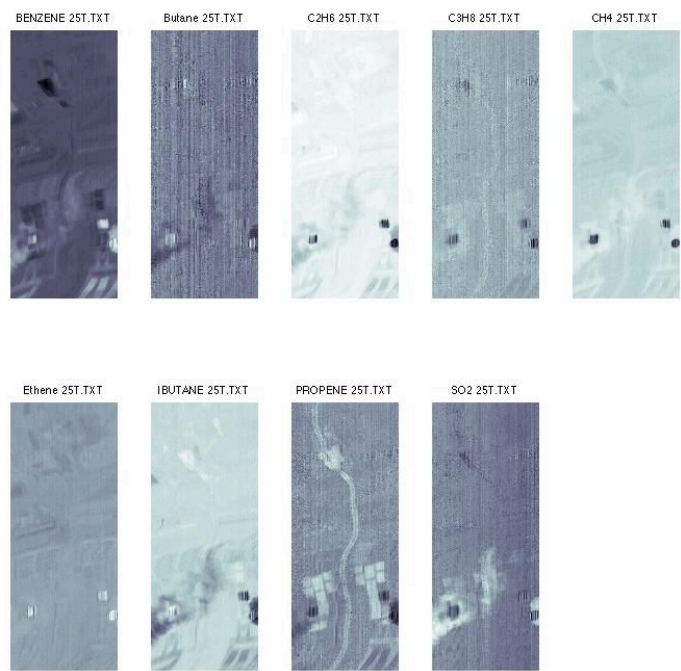


Figure 94. WPS Detection Planes for AHI 121247

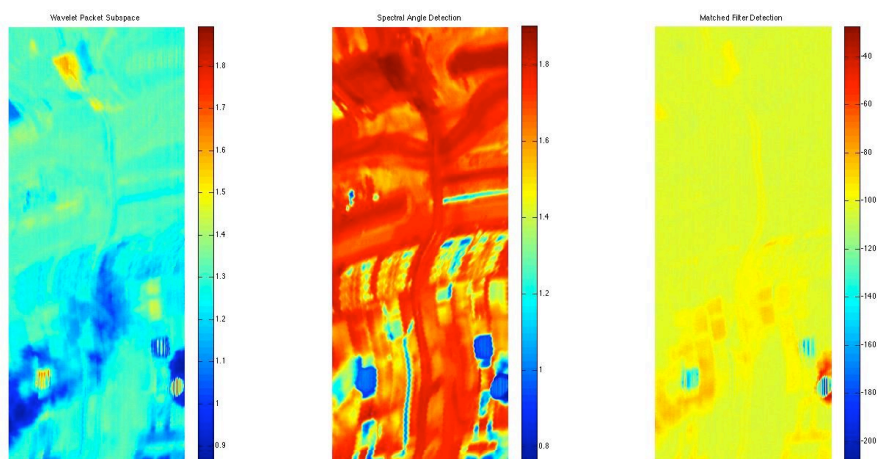


Figure 95. Selected Detection Planes for AHI 121247 (SO_2)

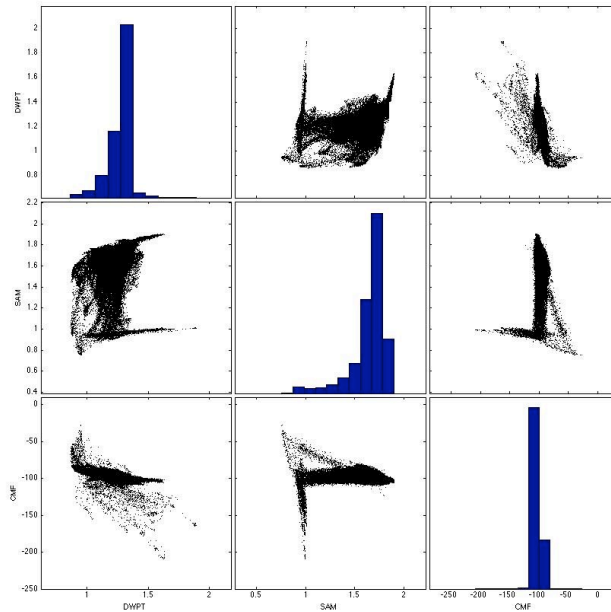


Figure 96. Scatterplot for AHI 121247

AHI 150652 is data collected over an animal feeding operation and AHI 105948 is data collected over a petroleum refinery. Although this data had plumes in them [60], neither of the published detections could be verified. Neither plume was detected with any of the algorithms. Another unknown plume in AHI 105948 was detected and identified.

For the two flight lines with unknown plumes (AHI 105948 and AHI153924), the typical two-step detection and identification process were applied. First, potential gas plumes were selected using visual review of the individual detection planes. The potential plumes were then identified using a liner regression technique [41]. A larger library of gases was used in the detection process to facilitate plume detection.

SAM results were again disappointing in that no planes could be used for plume detection, Figure 97. Both CMF and WPS had more potential plumes, but selection of a real plume versus a background feature was challenging (Figure 98 and Figure 99). Review of both the linescan image and the detection planes were required to visually separate potential plumes from ground features. The large slanted feature across the middle of the scene is an industrial stack. A small dark area, just below the stack, visible in the benzene CMF and WPS planes, was selected for identification. Visual comparison of all detection planes did not indicate any significant advantage of CMF or WPS (Figure 100). A scatterplot for the benzene detection plane is shown in Figure 101.

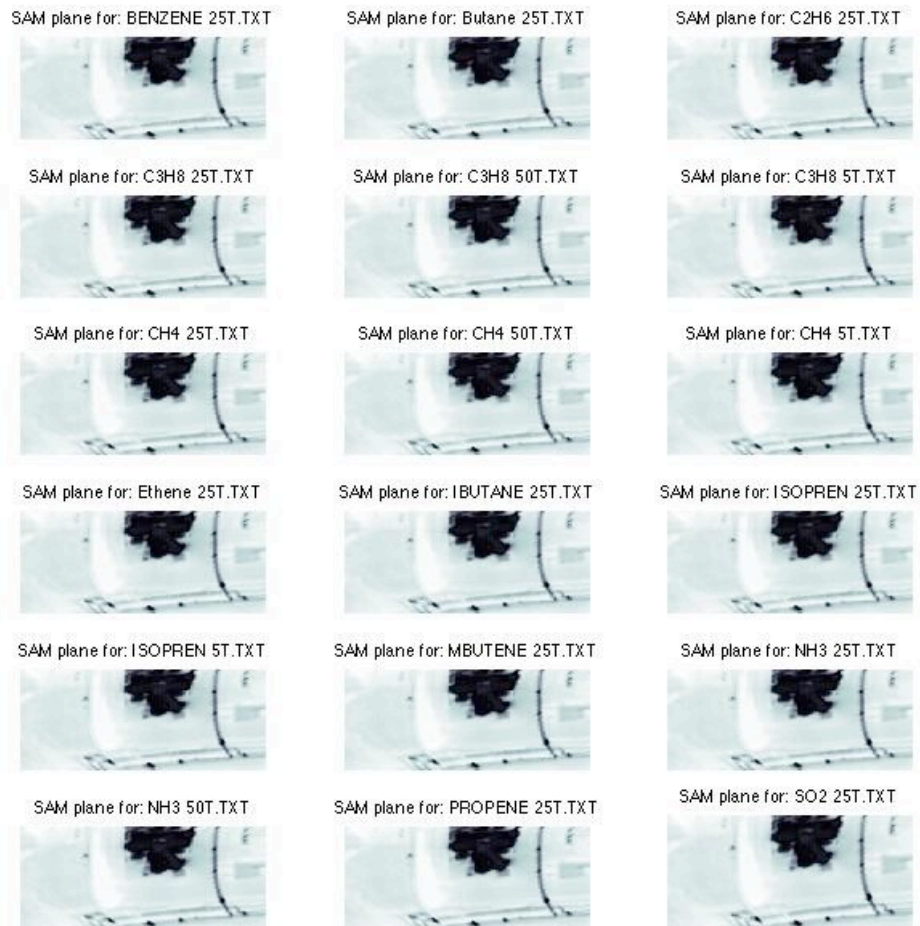


Figure 97. SAM Detection Planes for AHI 105948

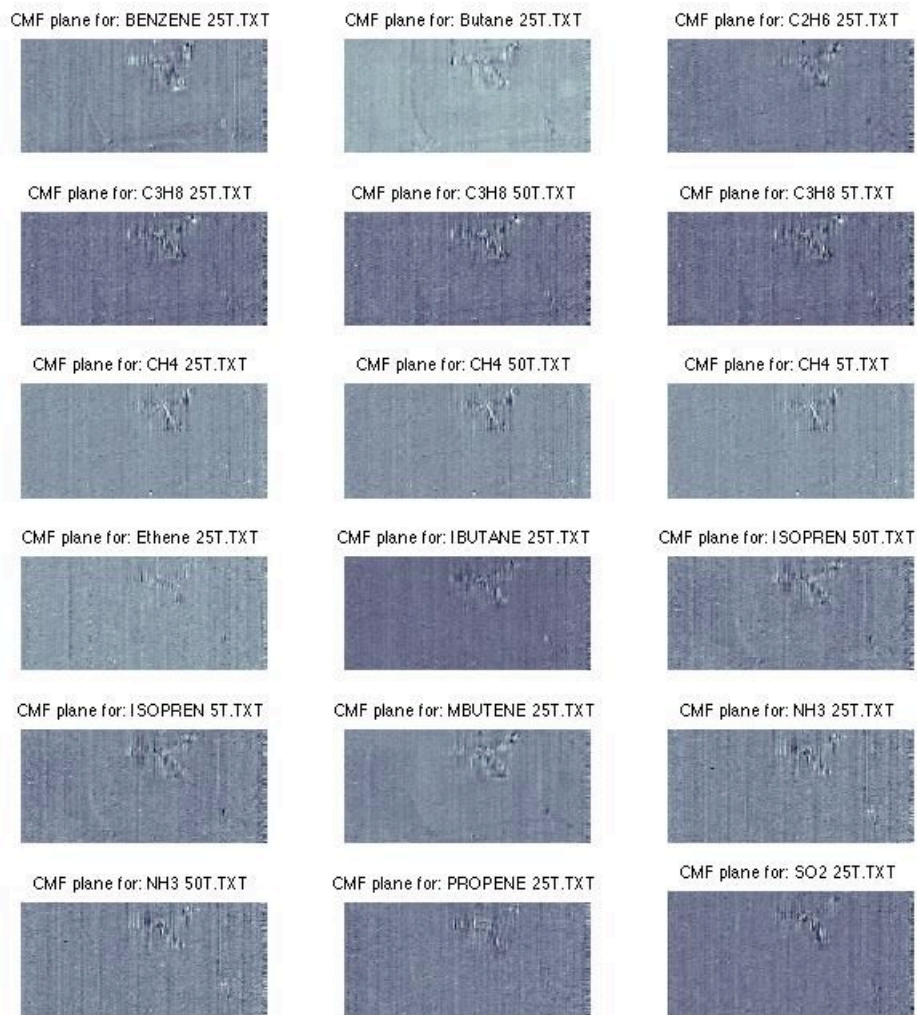


Figure 98. CMF Detection Planes for AHI 105948

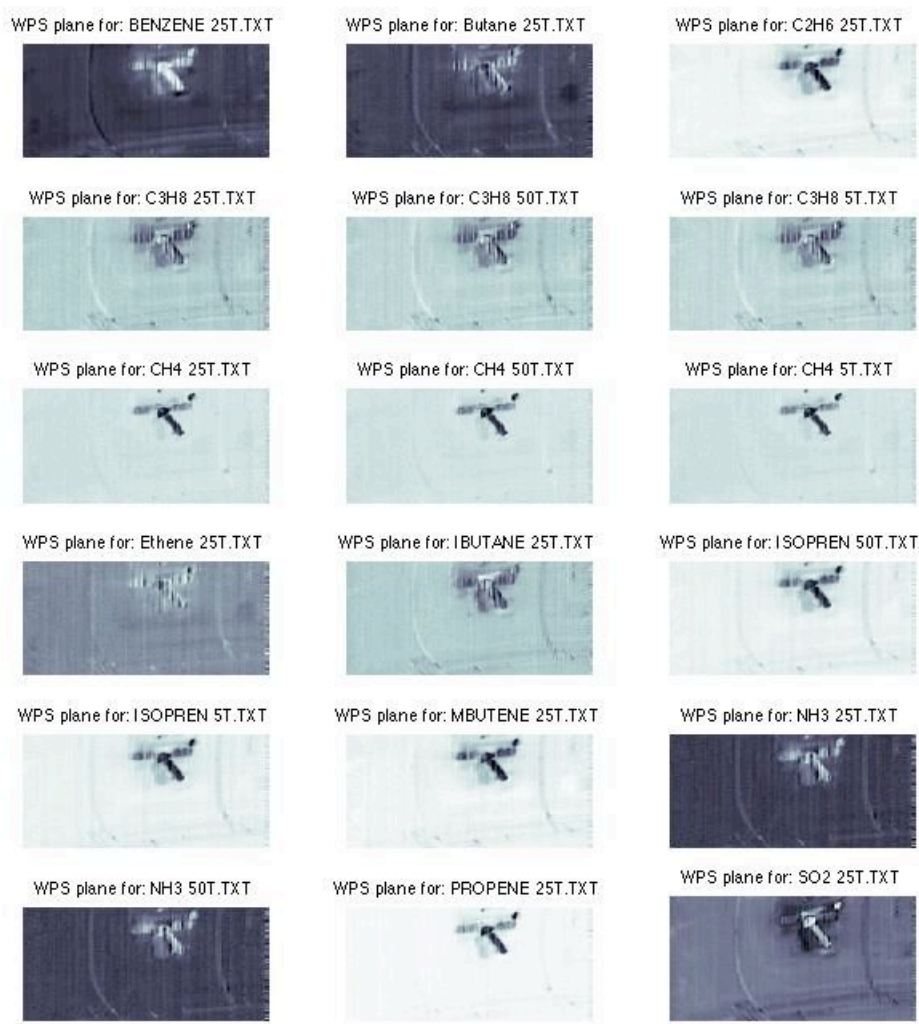


Figure 99. WPS Detection Planes for AH1 105948

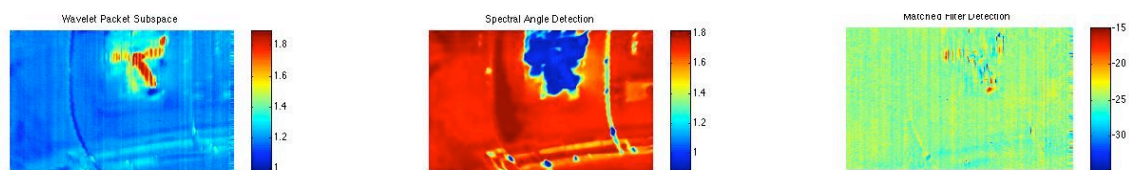


Figure 100. Selected Detection Planes for AH1 105948 (Benzene)

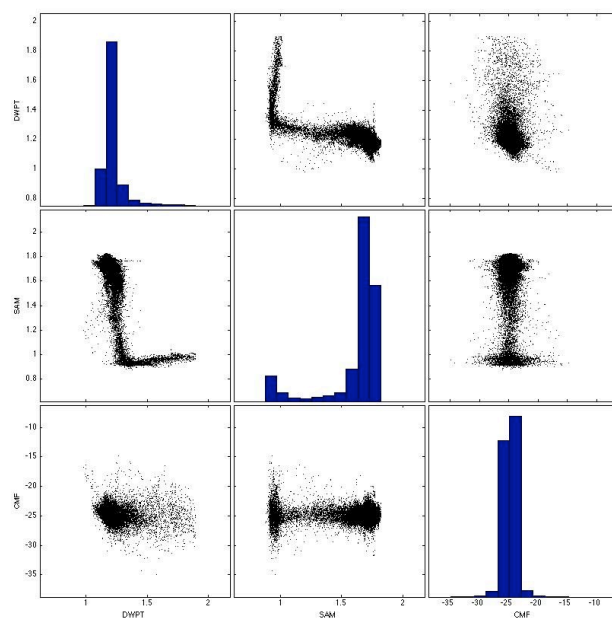


Figure 101. Scatterplot for AHI 105948

AHI 153924 is another petroleum refinery with suspected gas plumes. The same two-step approach as 105948 was taken for plume detection. In this case the AHI metadata included a potential set of gases released from this location. Gases that were available in the PNNL library were used for the detection step.

As with the previous AHI data, SAM performance was disappointing with no visible identification possible, (Figure 102). Both WPS and CMF exhibited a strong potential plume in the benzene image, Figure 103 and Figure 104. The visual strength of the color detection planes for CMF and WPS led to selection of the potential plume in the benzene detection images, Figure 105. Statistics of the image also indicate strong benzene detection scores (2σ to 3σ classification). The potential emission is located in

the upper half of the image near the center. This corresponds to a large tank in the linescan image. Visual comparison of the WPS and CMF planes lead to no significant advantage of one method to the other. But, used together and in conjunction with the statistics of the filter scores, this appears to be a strong result. Identification in Chapter 4 also leads to a potential benzene emission. A scatterplot for the benzene detection plane is shown in Figure 106.

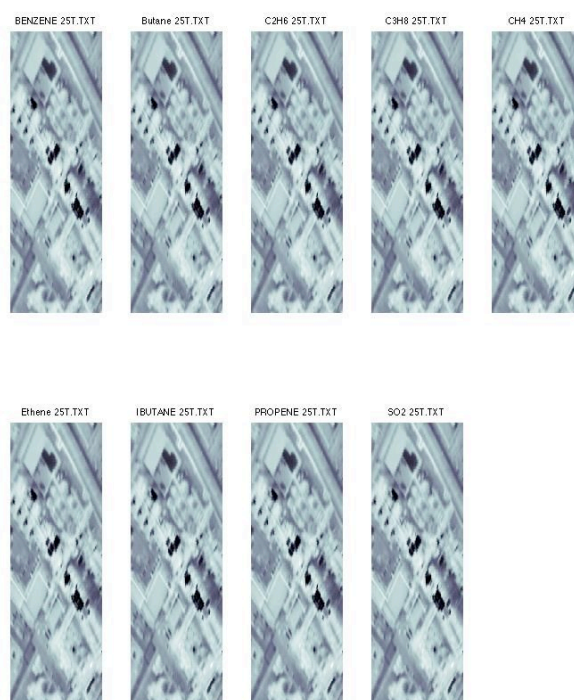


Figure 102. SAM Detection Planes for AHI 153924

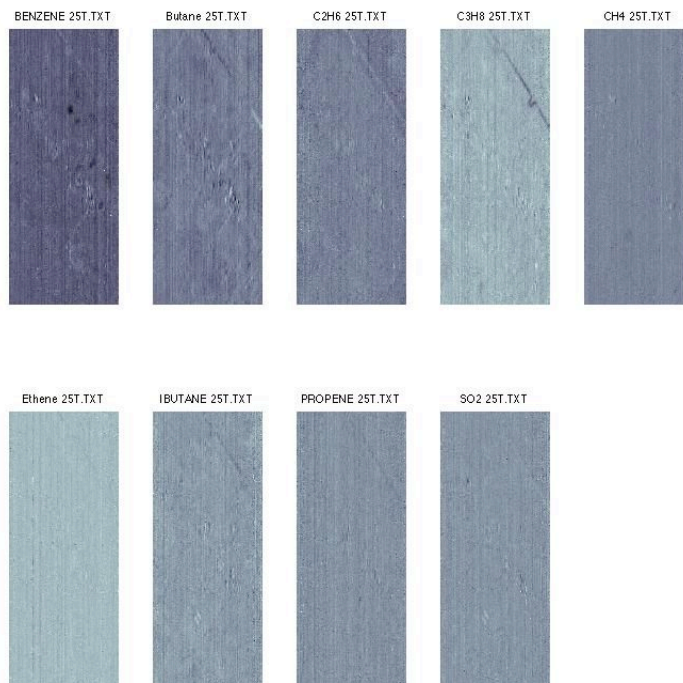


Figure 103. CMF Detection Planes for AHI 153924

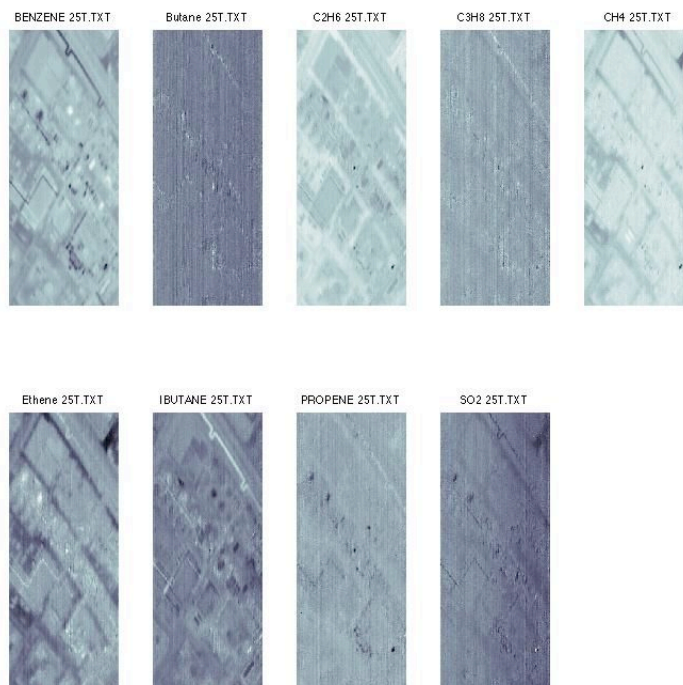


Figure 104. WPS Detection Planes for AHI 153924

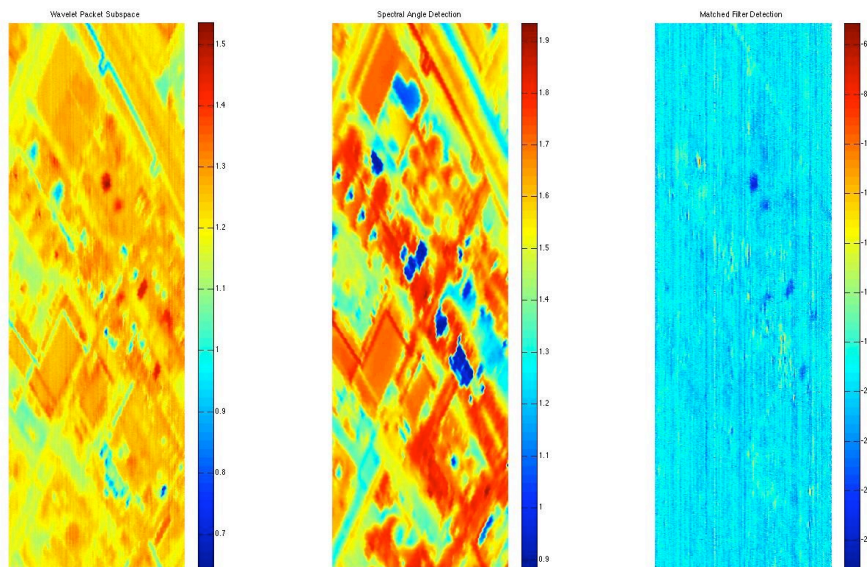


Figure 105. Selected Detection Planes for AHI 153924 (Benzene)

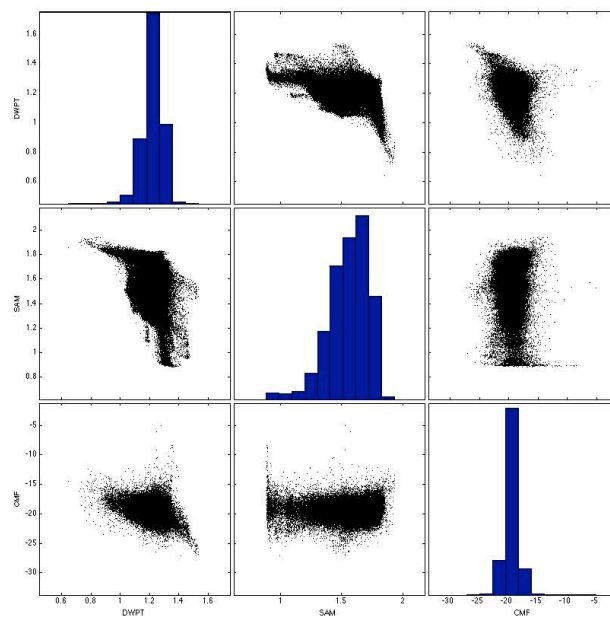


Figure 106. Scatterplot for AHI 153924

3.3 Advanced Infrared Sounder Data

The focus of analysis of AIRS data is the same as AHI data, the detection of gas emissions in the wavelet packet subspace. Given the spatial resolution of AIRS, gas emission events of a global nature, such as volcanic eruptions or large scale fires are suitable for this type of detection. The AIRS analysis focuses on sulfur dioxide detection from natural and man-made vents. Detection is attempted with the same algorithms as the previous data sets. Identification of the sulfur dioxide is through visual comparison and correlation with other published results. The wavelet packet subspace process is applied in the same process as the previous data sets. Detection images, classification maps, scatterplots, and histograms are used to illustrate the results.

The first data set, data granule 169, contains a large volcanic plume from Soufriere Hills volcanic eruption in Montserrat, July 2003. Comparison of the three algorithms shows varying levels of visual detection with WPS having the most striking image in Figure 107. The classification map for WPS demonstrates excellent detection beyond 3σ when compared to a TOMS SO₂ detection of the same event (Appendix 1). SAM and CMF had limited success beyond 1σ , (See Figure 108, Figure 109, Figure 110, and Figure 111). Appendix 1 includes some initial work regarding this data set. In the previous analysis a single level wavelet transform was used for a visual detection. In this case, the full wavelet packet subspace process with the SO₂ target spectrum best-basis has been applied. Statistics of the detection scores indicate that although SAM and CMF could separate some of the SO₂ pixels, the WPS separated many more of the target pixels

from the rest of the scene. A georegistered image, Figure 114, clearly indicates the source of the plume from Montserrat. A comparison of pixel and SO₂ spectra show good visual correlation between the pixel and reference spectrum, Figure 113. Identification of the SO₂ is discussed in Chapter 4.

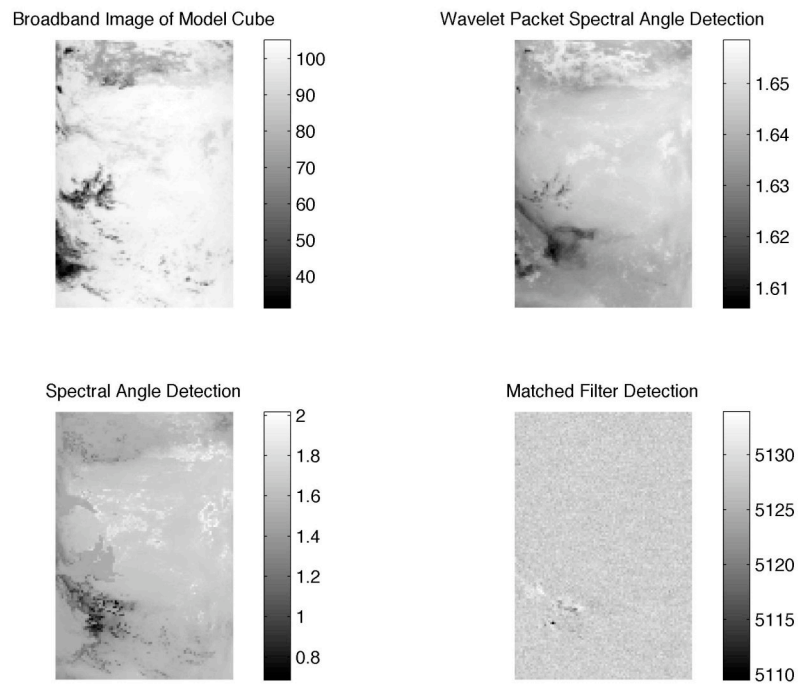


Figure 107. Detection map for AIRS Granule 169 (SO₂)

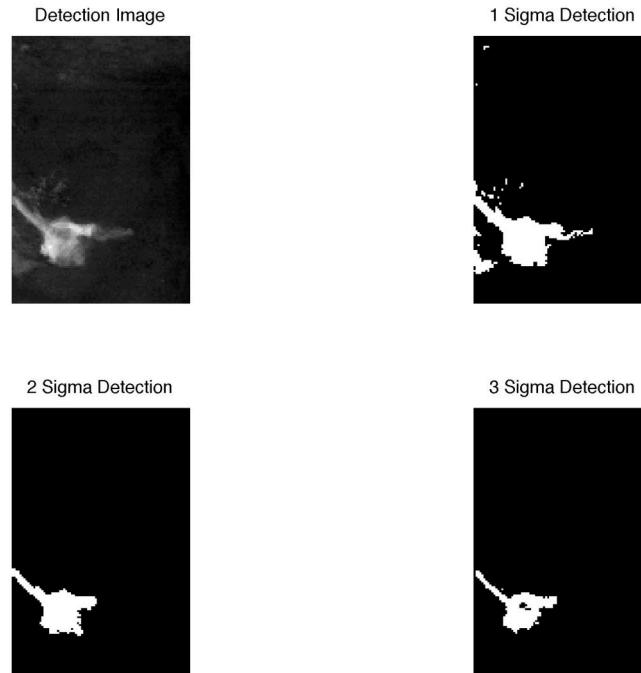


Figure 108. WPS Classification Map for AIRS Granule 169 (SO₂)

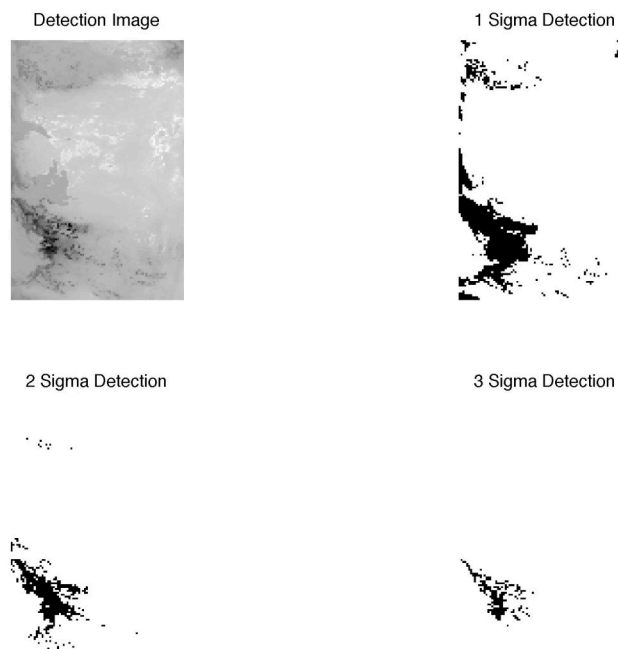


Figure 109. SAM Classification Map for AIRS Granule 169 (SO₂)

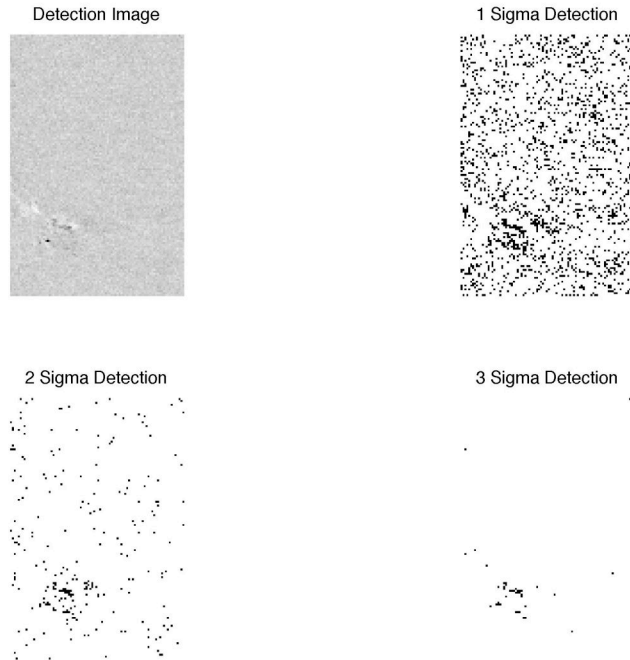


Figure 110. CMF Classification Map for AIRS Granule 169 (SO_2)

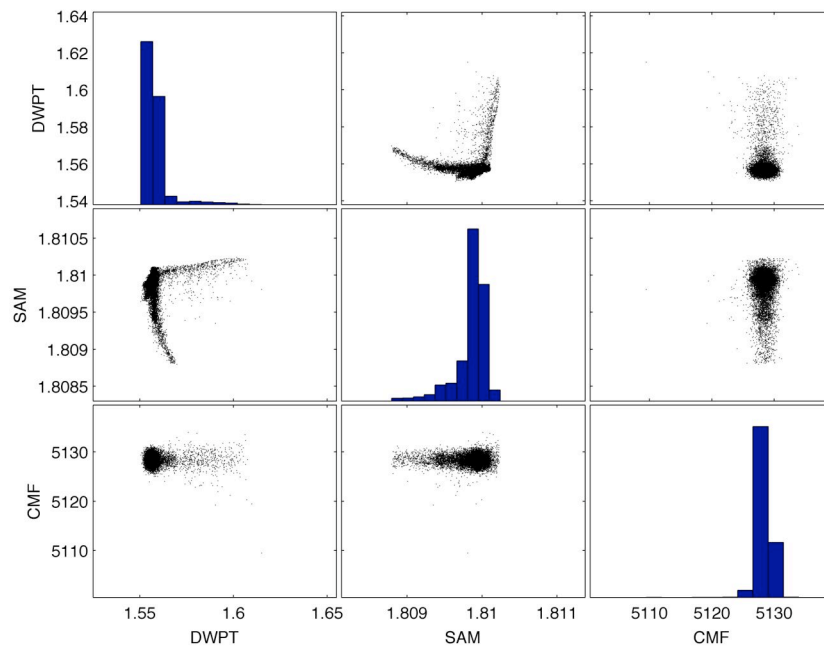


Figure 111. Scatterplot for AIRS Granule 169 (SO_2)

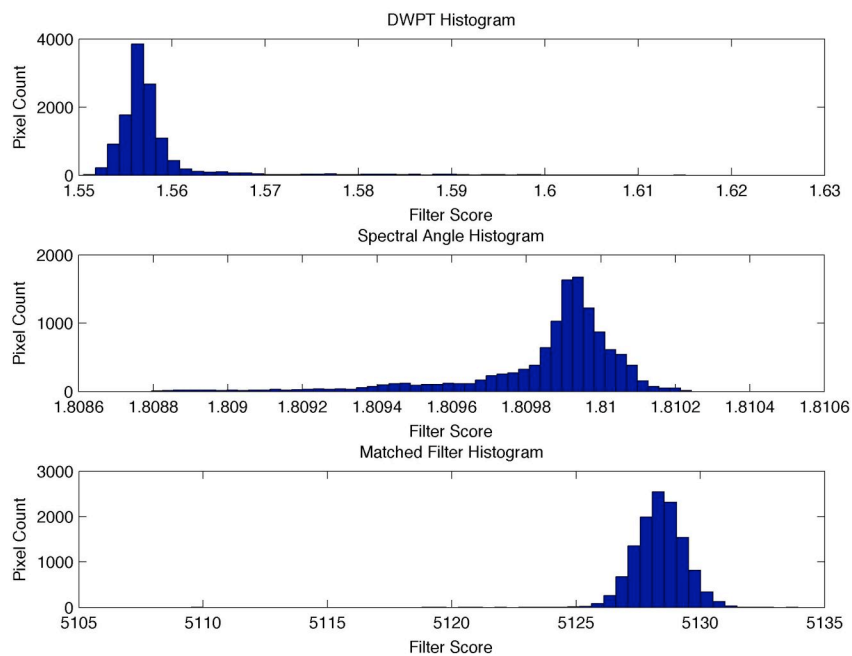


Figure 112. Histogram for AIRS Granule 169 (SO₂)

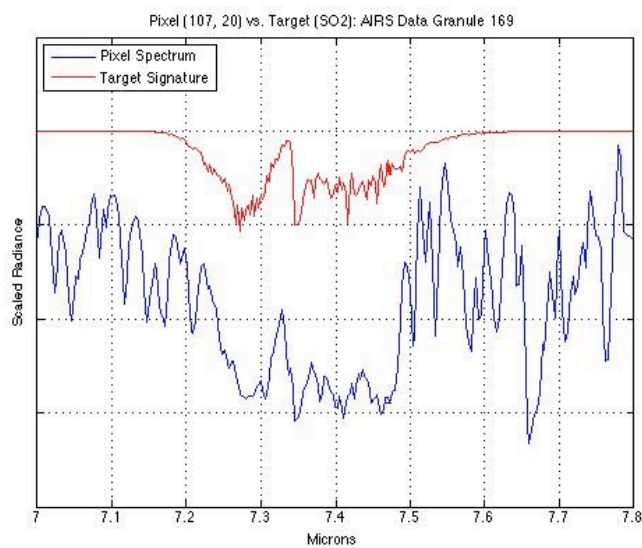


Figure 113. SO₂ Spectral Comparison for AIRS 169

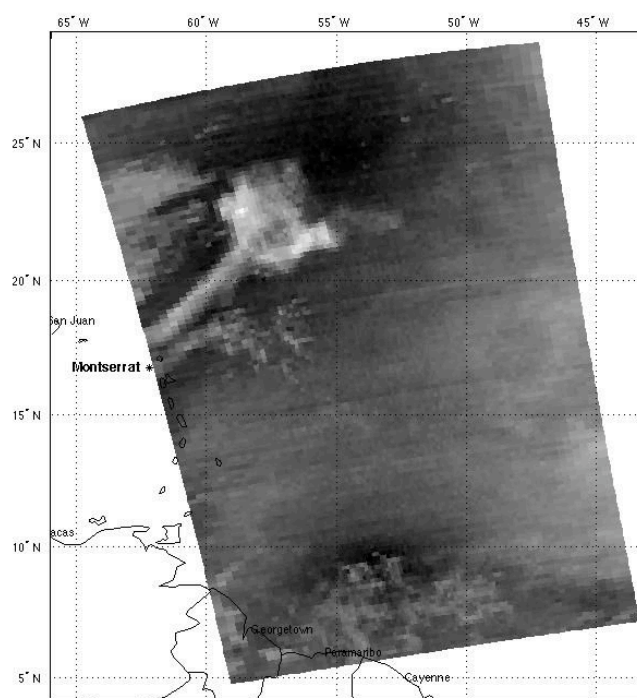


Figure 114. Georegistered Image of AIRS 169
(Montserrat Volcanic Eruption, 2003)

AIRS data granules 102 and 107 are from June 29, 2003 and July 3, 2003 respectively. These data granules contain the Al-Mishraq sulfur fire of 2004 [81]. The same process for plume detection was applied to these data sets. In both cases the WPS detection process clearly indicates a potential SO_2 emission, Figure 115 and Figure 116. Classification at 3σ also indicates a strong score for both 102 and 107, see Figure 117 and Figure 118. SAM and CMF detection images are marginal at best. Statistics for the detection scores show good separation by WPS for SO_2 and the background while SAM and CMF scores are virtually inseparable, see Figure 119 and Figure 120. (Data granule 107 figures are provided in Appendix 3.) Target spectrum and pixel spectrum visual

comparison for 102 and 107 are not as definitive as 169, (not shown). The SO₂ identification applied is discussed further in Chapter 4. Georegistered images provide evidence that AL-Mishraq is the source of the emission and potential SO₂, Figure 121 and Figure 122

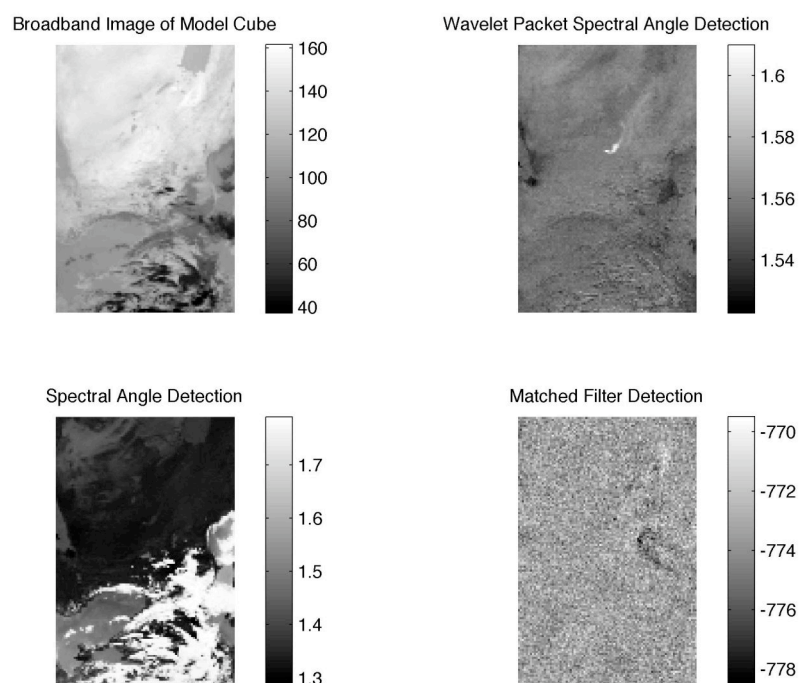


Figure 115. Detection map for AIRS Granule 102 (SO₂)

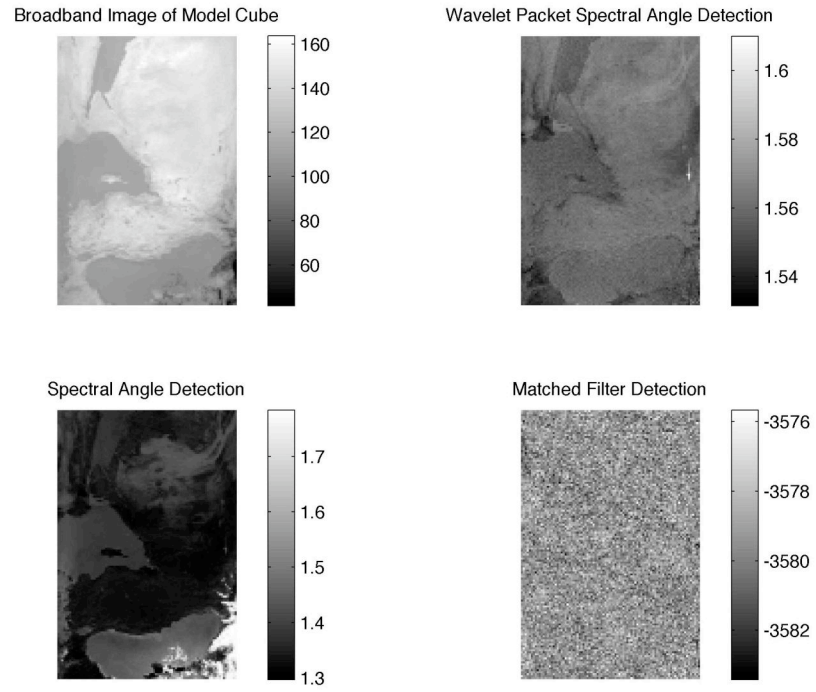


Figure 116. Detection map for AIRS Granule 107 (SO₂)

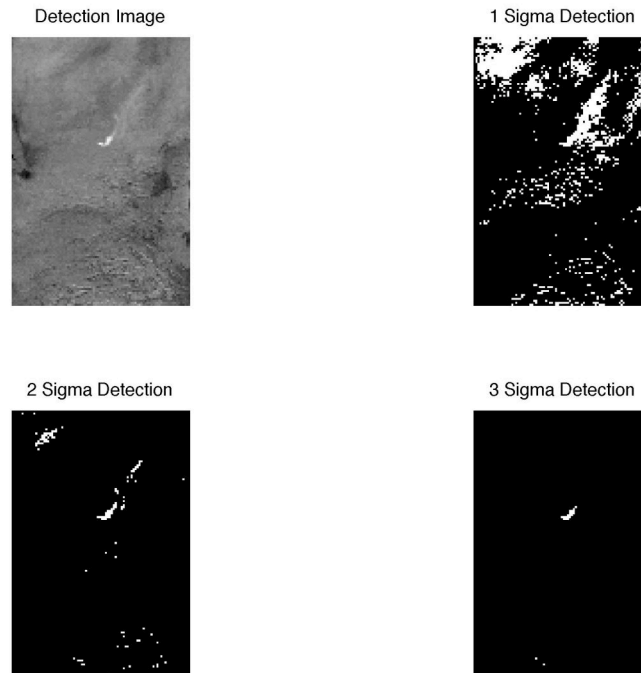


Figure 117. WPS Classification Map for AIRS Granule 102 (SO₂)

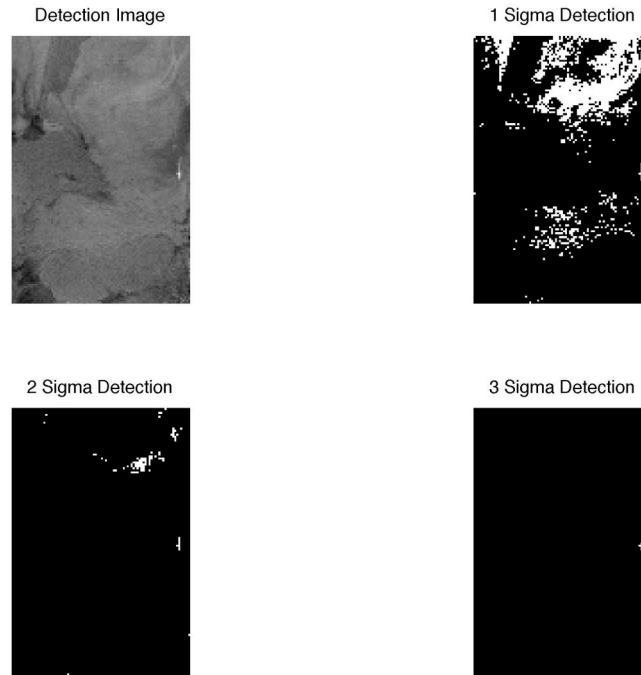


Figure 118. WPS Classification Map for AIRS Granule 107 (SO_2)

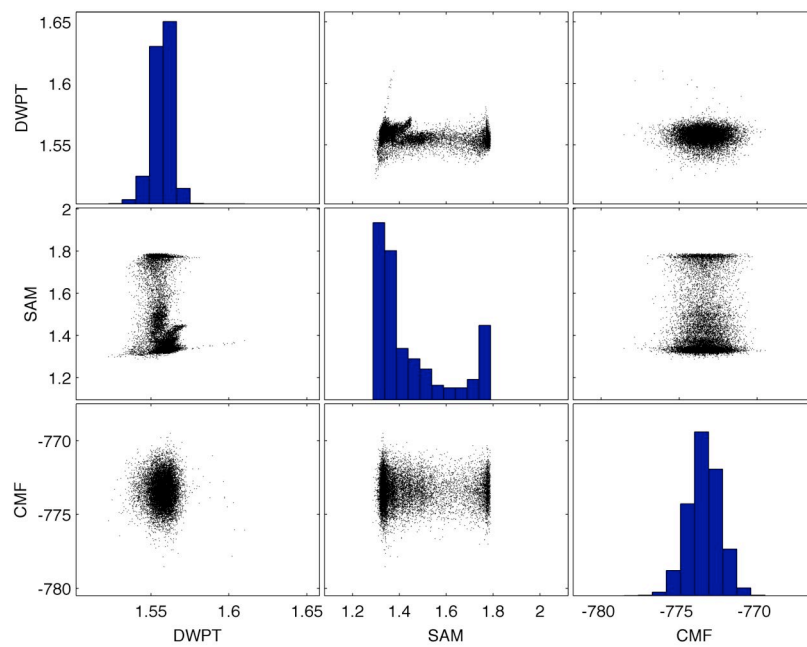


Figure 119. Scatterplot for AIRS Granule 102 (SO_2)

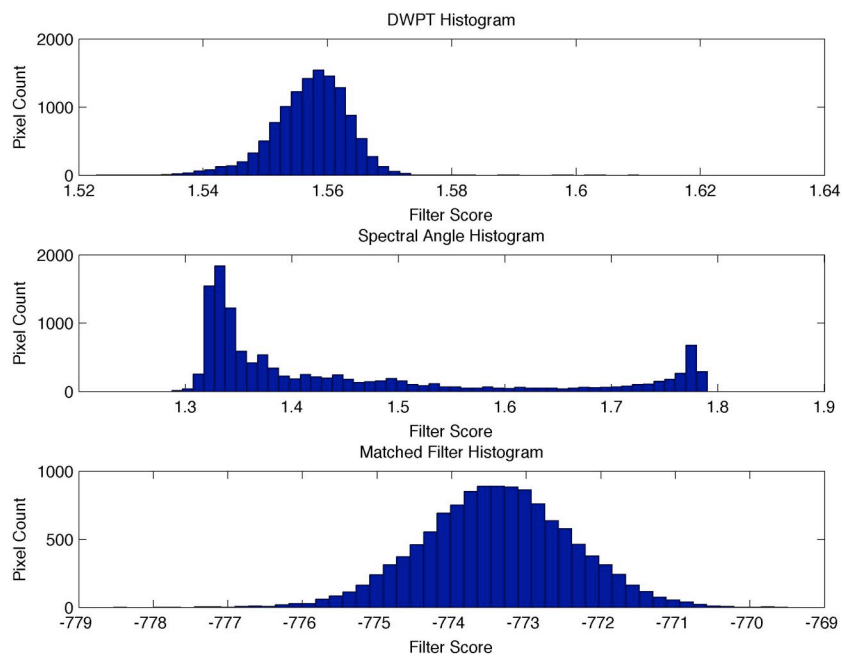


Figure 120. Histogram for AIRS Granule 102 (SO₂)

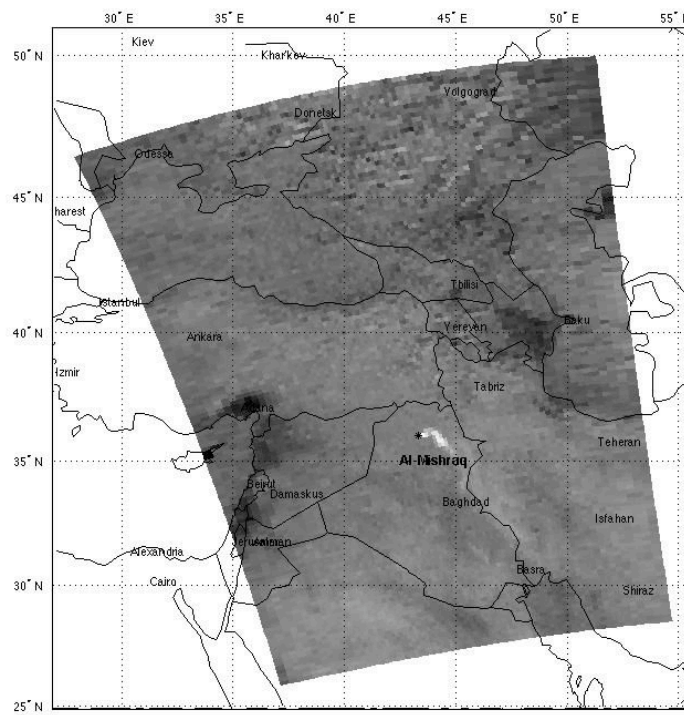


Figure 121. Georegistered Image for AIRS 102

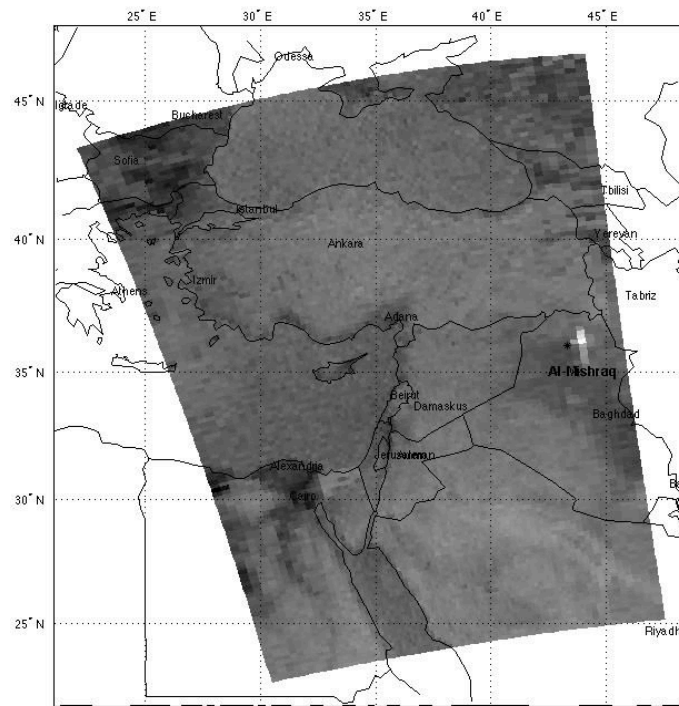


Figure 122. Georegistered Image for AIRS 107

4 Discussion

Chapter 4 will discuss the results. It will document factual relationships, their significance and other observations. Exceptions in the process will be noted as well as any theoretical implications and initial conclusions. The chapter will progress through the overall research process. Starting with the simulated data cubes, stops are made along the way to discuss wavelet selection, wavelet node and coefficient selection, and potential improvements to the wavelet packet subspace approach.

4.1 Construction of Simulated Data Cubes

In construction of the simulated cubes, the primary driver was development of the radiance model. The simulated data cubes focused on AHI wavebands because of AHI's spectral and spatial resolution. This was done for two reasons. First, the primary focus of the simulated data was to evaluate the WPS algorithm in a relatively simple model. The second is, AIRS's high spectral and low spatial resolution. AIRS requires a complex atmospheric and surface model to define each pixel.

For the AHI spatial resolution of several square meters, it could be assumed that any one pixel would contain homogenous surface emission, a homogenous target gas signature in some known relative quantity, and homogenous atmospheric upwelling. For

the AIRS spectral and spatial resolution these assumptions are not necessarily valid. Here each assumption is evaluated in slightly more detail.

For AIRS, the source of surface emissions are the contributions from many types of land surface materials. This is due to its large spatial resolution of > 14 km by 14 km per pixel. These surface emissions also include rivers, lakes, and the ocean. Because of its varied material type, the surface emission is more like a graybody emission than the single material assumed to be in an AHI pixel. One could argue that at the pixel level, the homogenous surface emission is valid in either case. When comparing the total area collected, each AIRS pixel must be modeled as individual graybodies with different surface emissivities and temperatures. For an AHI flightline over a similar background (i.e. forest, desert, or urban backgrounds), it may be assumed that most of the background is similar, requiring only a single surface emissivity and temperature for the entire scene.

This same argument easily moves into atmospheric upwelling. In a single 14 km by 14 km pixel, the atmosphere may contain, clouds, rain, and open sky, while in a pixel of several square meters (~ 25 m²), a homogenous atmosphere is easily assumed. Across the entire AHI flightline there generally exists only open sky and cloud covered areas. These areas can be spatially subset during analysis and the atmosphere assumed to be homogenous in each case.

For the homogenous target spectrum assumption, a single AIRS pixel covers a minimum of 196 km². Gas emissions at that scale are global events and may occur in only a single pixel. A relatively small gas plume in an AHI flightline will cover many pixels with similar distributions in several pixels. A single expected relative quantity and

temperature for each scene is much simpler to estimate than modeling multiple 196 km² areas.

For the simulated data cubes the simplest assumptions are made. A single surface upwelling with the same temperature and emissivity is used across the scene. A single atmospheric upwelling is also applied across the scene. The gas plume inserted is at the same homogenous temperature and relative quantity for each pixel in the plume. For data cubes with noise (lower S/C) random noise is added to the background. The background combines surface emission and atmospheric upwelling prior to adding the plume signature. The plume does not add another noise component. This model best represents the AHI assumptions vice the AIRS assumptions.

This does lead us to the validity of the radiance model across high and low spectral and spatial resolutions. Fundamentally the radiance model is valid in either case. The basic physics remains the same for any variation of spectral or spatial resolution. The significant difference occurs in the pixel-by-pixel modeling of the surface emission and atmospheric upwelling. For the AHI case, the assumption has already been made for uniform surface emission and upwelling across the scene. For the AIRS case, each pixel should have an individually calculated surface emission and atmospheric upwelling given its large spatial resolution.

The analysis conducted on the AIRS data does not develop individual models for each pixel. It takes the same assumption as the AHI analysis that the surface emission and atmospheric upwelling are uniform across the scene. This is done to alleviate the computational burden of 12150 LBLRTM atmospheric model runs per AIRS data granule

and to simplify the characterization of each pixel's surface and atmospheric properties. In practice or production, individual runs and pixels are independent and this number of atmospheric model runs is an easily parallelizable problem. The single LBLRTM run per data cube is performed for all three algorithms evaluated. Thus, no algorithm has any advantage or disadvantage given this assumption.

The most significant contribution of the radiance model is its concept of thermal contrast. Thermal contrast is a function of: surface temperature, ambient air temperature, plume temperature, plume column density, and ambient air column density. Once calculated, it applies a scalar value to the target spectra, which drives the intensity of the plume and its basic character as an absorption or emission spectrum.

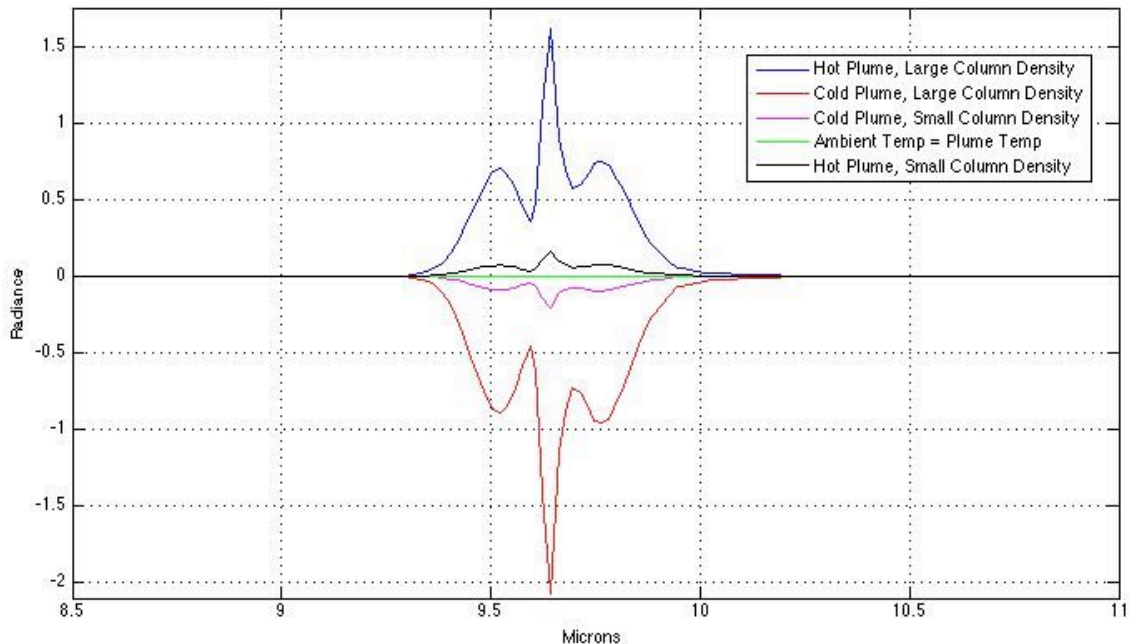


Figure 123. Variation of Target Signature (Benzene) Caused by Thermal Contrast

Figure 123 shows a benzene spectrum that has been sampled to AHI wavelengths. Multiple thermal contrasts have been applied. The emission signatures, (plume temperature greater than ambient), are positive on the radiance scale. The more intense blue line represents a greater column density of gas than the green line. Both lines are at similar plume and ambient temperatures of 295 K and 292 K respectively.

The red and magenta lines are absorption signatures (plume temperature less than ambient) with similarly varied column densities. The more intense absorption (red) is at the same column density as the blue line, and the magenta line is at the same column density as the black line. The plume temperature in this case (both red and magenta) is 288 K with ambient air temperature remaining at 292 K.

As we can see from the figure, the absorption or emission characteristic of the target signature is driven by the relationship between plume temperature and ambient air temperature. Intensity of the signature is driven by the column density of the combined plume and ambient air. It is significant to note that when the plume temperature is at the ambient air temperature, there is zero thermal contrast and thus no target signature (green line).

4.2 Wavelet Selection Process

With the simulated data constructed and with the radiance model well understood, the simulated cubes were then used to evaluate algorithm performance and to assist in the wavelet selection process. In Chapter 2, the wavelet selection process describes a brute-

force method to identify the best wavelet per target signature. Few definitive conclusions are made with the brute-force method for the four target gases under study. The quantitative and qualitative observations that are made include both shorter mother wavelets appear to generate more near zero coefficients, and differing target spectra do not appear to drive wavelet choice.

To evaluate the pragmatic case, several variations of the simulated data cubes with several mother wavelets were run. As the results in Chapter 3 demonstrate, algorithm performance was not stressed until SDC 4A. SDC 1 through SDC 3 did not reveal any specific metric that could discriminate performance between individual mother wavelets. An intermediate simulated data cube was constructed that emulated SDC 4A but with a higher S/C ratio of $\sim -33\text{dB}$ for CH_4 . This configuration was developed after a trial and error period of manipulating thermal contrast and evaluating ROC curves, histograms, P_d/P_{fa} , and area under the ROC curves. The area under the ROC curve proved to be a discriminator for wavelet performance and CH_4 was the stressing case. The results follow in Table 8.

Table 8. Pragmatic Wavelet Selection Results

<i>Algorithm/Wavelet</i>	<i>Area under the ROC Curve</i>
Spectral Angle Mapper	0.94220
Clutter Matched Filter	0.94608
Daubechies 1	0.91032
Daubechies 2	0.92708
Daubechies 3	0.92526
Coiflet 2	0.92396
Coiflet 3	0.90056
Symlet 2	0.92708
Symlet 3	0.92526
Biorthogonal 1.1	0.91032

Longer wavelets were attempted, but their computational time in the MatLab environment proved to be prohibitive (see section 4.7). The general trend identified was that a shorter, not necessarily the shortest, mother wavelet had slightly better performance than longer wavelets. Symlet and Daubechies had identical performance, while Coiflet and Biorthogonal wavelets were slightly behind. Selected mother wavelets were then applied to SDC 4A with the following results.

Table 9. SDC 4A Wavelet Selection Results

<i>Algorithm/Wavelet</i>	<i>Area under the ROC Curve</i>
Spectral Angle Mapper	0.68604
Clutter Matched Filter	0.66641
Daubechies 2	0.79774
Daubechies 3	0.79099
Symlet 2	0.79774
Coiflet 2	0.78693

Again, the Daubechies 2 and Symlet 2 had slightly better performance. In this case, the WPS algorithm outperformed SAM and CMF. This result holds significant promise for the analysis of real hyperspectral data in the wavelet packet subspace.

Depth of the wavelet packet decomposition was chosen adaptively. Depth was based on the best level tree of the target signature. When the WPS process was applied to the simulated data cubes, best level trees were determined to be $N=2$ for CH_4 and $N=5$ for SO_2 . When larger depths were manually chosen, performance decreased. No results were tabulated for manually chosen depths.

The pragmatic results agree with the Chapter 2 quantitative and qualitative observations. With the pragmatic results for SDC 4A, the Daubechies 2 mother wavelet with adaptive depth selection was used to calculate the results for SDC 1-3, the AHI flightlines, and the AIRS data granules.

4.3 Observations from the Simulated Data Cubes

Significant observations from analysis of the simulated data include:

1. The concept of thermal contrast drives the target spectrum characteristics regardless of algorithm
2. Shorter wavelets perform better than longer wavelets both by measured metrics and by computational time

3. Wavelet packet decomposition depth should be driven adaptively by target spectrum best-basis
4. A best basis cost function of the L^1 Norm, although chosen arbitrarily, appears to be adequate for characterizing a gas spectrum in a wavelet packet library
5. In one specific controlled case, WPS has outperformed both SAM and CMF. This case was the most stressing (lowest S/C) case among the simulated data.

4.4 Comparison of WPS Approach to SAM and CMF

Spectral Angle Mapper is not generally a method used for trace gas detection. In most analysis some type of matched filter or orthogonal projection is used to separate the background from the weak target spectrum. SAM is used as part of this comparison to demonstrate how the wavelet packet subspace does indeed increase the background and target separation leading to improved target detection. The comparison follows.

SAM is applied in spectral space to each pixel of the simulated data cube. In the wavelet packet subspace approach, a wavelet packet best basis is chosen adaptively based on the target spectrum of interest. This best basis defines the wavelet decomposition applied to the spectral data and selects the nodes and coefficients used to match the target to the data. The two patterns, wavelet packet decomposed target coefficients and wavelet packet decomposed pixel coefficients, are compared to each other by measuring the spectral angle between them. Without the wavelet packet decomposition, this pattern match reduces to the original SAM algorithm. Thus, comparing the two matching scores

quantifies the difference or similarity that the transformation to wavelet packet space applies. The similarities and difference are most clearly seen in the scatter plots of Chapter 3.

The SDC scatter plot results illustrate the correlation between SAM and the WPS filter scores. All of the scatter plots illustrate a nearly linear relationship between the two sets of scores. This implies a strong positive correlation. This is somewhat expected as the wavelet packet decomposition is a linear transformation of the spectral data into another space.

WPS and CMF scores demonstrate a similar trend but with a strong negative correlation. The sign of the correlation is due to several factors. These include target spectra in emission or absorption and the statistics of the data cube as represented in the data covariance. What is important is that the behavior of the scores does not vary for unknown reasons. The correlation characteristics are similar for all cases and the separation between background and target clusters is consistent.

Interestingly, multiple clusters appear in the CMF vs. SAM scatter plots for the multiple plume cases. Most likely this is the other plume in the scene. For WPS comparisons, the second plume is part of the background cluster. This behavior is consistent with the idea that the target has been separated from the background through the best basis. The second plume has become part of the background.

For all cases where the noise increases and the target signature decreases, e.g. SDC 2A and 4A, the separation between the target cluster and background cluster decreases (as expected). For CH₄ in SDC 4A, none of the algorithms have any separation between the

target and background and thus the target is not detectable. The WPS algorithm behaves in a similar fashion as SAM and CMF.

For the weaker CH₄ plume in SDC 3, 4, and 4A, CMF performance falls off. This may be due to the presence of both plumes in calculation of the covariance matrix. It is expected in application of the CMF that the covariance matrix represents the background only. This result suggests that the data used for calculation of the covariance matrix should never include any potential target signatures. This concept is applied in analysis of the AHI data.

The WPS algorithm has several intrinsic qualities worth noting at this point. The WPS algorithm is target signature based. A wavelet packet library is decomposed for the target spectrum and a best-basis chosen based on the L^1 Norm cost function. The best-basis is represented by nodes of the wavelet packet tree. Each node contains the coefficients for that scale and location of the wavelet packet decomposition. Significant differences that define the WPS algorithm are:

1. The best basis nodes are an orthogonal subspace that best represents the target spectrum for the given cost function (L^1 Norm).
2. The target vector is represented by the concatenation of the coefficients from the best basis nodes.
3. The target vector can implicitly reduce the dimensionality of the spectral space. That is, the total number of coefficients used for pattern matching may be significantly less than the original number of spectral bands.

4. The background is implicitly suppressed by eliminating any best basis nodes of the background vector wavelet packet decomposition and by using the target best basis for matching.
5. Significant features of the target signature are captured in scale and location through the best basis nodes.

SAM does not possess any of these qualities. CMF does represent an orthogonal subspace that is rotated to match the variance of the data. CMF also suppresses the background component through calculation of the inverse covariance and applies it to the pixel vector (data vector). Separating the WPS from CMF explicitly is the orthogonal subspace chosen for matching best represents the target spectrum based on an entropy criterion (i.e., cost function). The CMF uses only variance. Also, any entropy criterion may be used to select the subspace. At least $2^{N/2}$ subspaces are possible (N = number of bands used for the target spectrum).

4.5 Analysis of the AHI Flightlines

The AHI flightlines were analyzed using two different analysis methods. The first method was focused on duplicating the plume detections provided with the AHI data set. The second method was focused on detecting and identifying other potential plumes in the data. A manual review of the data led to both spatial and spectral sub-setting. Bands

below 7.5 μm and above 11.5 μm were assessed as bad or inconsistent among pixels.

Samples 1 to 8 in the spatial domain were also considered bad.

Method 1 was applied to FL 121247 and FL 150652. FL 150652 contains a potential NH_3 (ammonia) plume and FL 121247 contains a potential SO_2 plume. The analysis method is straightforward. Each algorithm was applied to the data using several library spectra. Visual and statistical results were compared among the three algorithms.

Method 2 was applied to FL 105948 and FL 153924. Flightline 153924 had no known plumes in it. FL 105948 had a potential benzene plume [60] but this plume was not detected nor analyzed. This analysis method had two steps 1) apply each algorithm using several library spectra and visually compare results 2) select a potential plume from step 1 and use linear regression to identify the most significant chemical. These two steps are the typical detection and identification steps used in LWIR hyperspectral gas analysis.

Visual comparison of the WPS and CMF detections to the EPA provided SO_2 detections are striking (Figure 124). SAM detection was weak in all cases. The CMF algorithm partially detects the SO_2 while WPS performed significantly better. CMF had stronger detections in the isobutane matched filter plane. WPS also had a strong isobutane detection, noted in Figure 94. Covariance calculation for CMF used data from another section of the same flightline where no SO_2 was detected. WPS SO_2 detection best matched the provided EPA image. The EPA analysis applied a principal component method with matched filter. Specifics of the procedure are unknown.

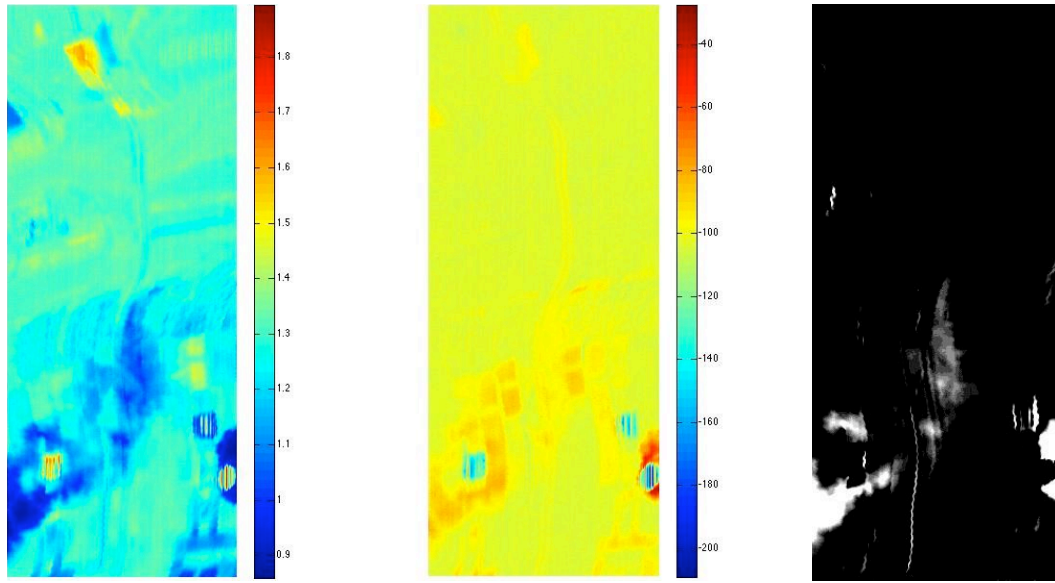


Figure 124. Comparison of SO₂ Detections FL121247
WPS – Left, CMF – Center, EPA Provided – Right

To quantify the strength of the potential SO₂ plume, a classification map was created showing a strong plume beyond 2σ in Figure 125. A preliminary linear regression was applied in this case (although not part of the discussed analysis method). The regression result was not as strong as expected with R^2 of ~ 0.4 and an *F-Statistic* of 45. Pearson correlation between SO₂ and the plume spectra was calculated at 0.697. These results indicate there are most likely other gases mixed in the plume with SO₂. The EPA provided SO₂ identification was a visual comparison of spectra only.

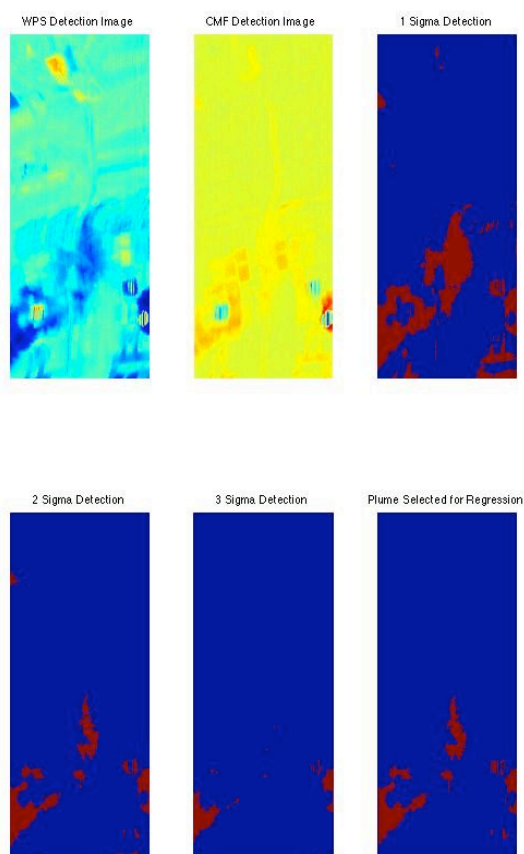


Figure 125. FL 121247 SO₂ Plume Classification Map

FL150652 is a collect over a CAFO facility and has a potential NH₃ plume. After applying all three methods, no plume was detected. Modifications to the CMF and WPS algorithms and the atmospheric model were attempted with no success. (Modifications included selection of different backgrounds for the covariance matrix calculation and adjustments to the MODTRAN model.) included The potential NH₃ plume was previously identified as very weak [60]. Individual pixel examination of the data and comparison to the published results led to no conclusions regarding the presence of ammonia. No results are provided herein.

FL 105948 and FL 153924 were selected for analysis for their potential for gas plumes. Both flightlines were collected over major petroleum refineries and were expected to contain multiple plumes. Metadata existed for FL 153924, which indicated known releases of various gases. In both cases all three algorithms were applied to detect plumes. No significant changes were noted in any of the algorithms. The initial depth of the wavelet packet decomposition was reduced from 8 to a range of 2-4 to speed detection time.

It should be noted that the initial depth is used to perform a full wavelet decomposition of the target spectra. The best basis is selected from the full decomposition tree. The level of the best basis tree is generally not the initial depth. For example the SO₂ target depth selected was 5 while for CH₄ the target depth was 2. The depth chosen should match the bandwidth of the spectral features of interest (adaptive depth selection).

FL 105948 was the most challenging in regards to plume detection. Ultimately a small area just south of the industrial stack was chosen as a potential plume. This area was identified in both the CMF and WPS detections. SAM had little success in any of the flightlines for plume detection. FL 105948 had few potential plume detections. A trial and error process was applied in selecting a potential plume and then applying a linear regression for identification. Plumes were selected by examining the means and standard deviations of the filter scores. Pixels greater than 2σ or 3σ away from the data mean were selected as plume pixels. Figure 126 is a classification map built from the WPS benzene detection plane.

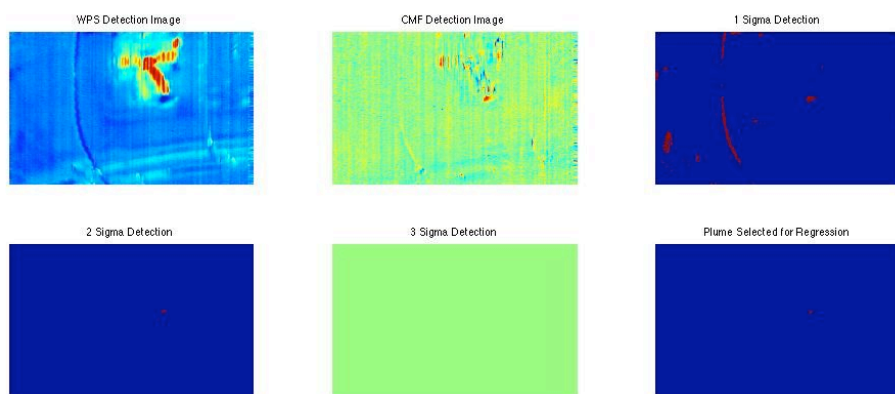


Figure 126. FL 105948 WPS Plume Detection

A Google Earth image was used to provide context to the scene. The plume is located just south of the industrial stack in the upper left quadrant. The target library for detection consisted of chemicals listed in the AHI metadata for the FL 153924 site and that were available in the PNNL Library. A table of chemicals used for detection and identification is given in Appendix 4.



Figure 127. Google Earth Image for FL 105948

A constrained linear regression was applied to the mean pixel spectrum of the plume. Figure 128 plots the plume and background spectra. A sum to 1 constraint was applied. The regression library applied did not contain all of the potential chemicals that may be present in the plume or scene. These results represent linear regression against the target spectrum with only the library spectra that were available. Also, an average background spectrum was inserted into the library to represent a background endmember. The results indicate a potential methane plume, which is not unreasonable for this location. R^2 for the regression was strong at 0.74 with a *F-Statistic* of 45. There were no other chemicals with significantly high regression coefficients. Visual review of the plume and methane spectrum indicate a single feature at about 8.25 μm . This is potentially a plume that contains methane.

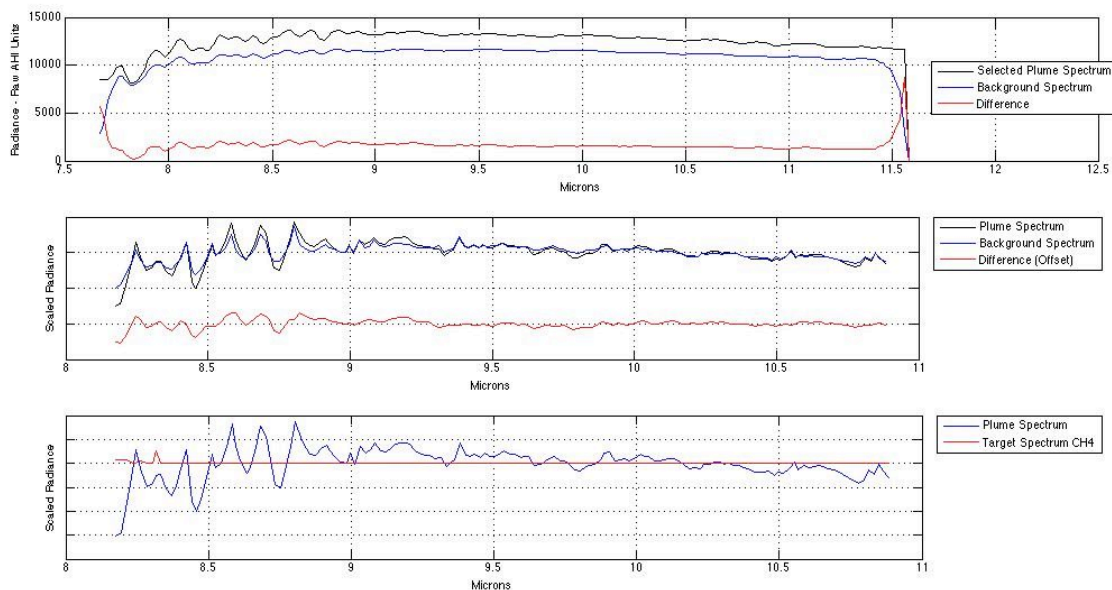


Figure 128. FL 105948 Spectral Identification of CH₄

WPS and CMF analysis of FL 153924 led to several potential plumes. The strongest appeared in the benzene detection plane for both WPS and CMF. The linescan image provided sufficient context to determine that the potential plume was directly over a large chemical storage tank. The results of the classification map based on the benzene detection plane proved a strong plume beyond 3σ , see Figure 129. The same constrained linear regression was applied to the mean plume pixel as FL 105948. The library used was also identical. The regression result indicates a benzene plume. R^2 is calculated at 0.858 with an *F-Statistic* of 72. Visual comparison of the mean plume pixel and target spectrum illustrates the difference between background and plume pixel representing the major features of benzene between 8.5 μm and 11 μm , Figure 130. Benzene releases are indicated by the metadata for this site. The regression scores coupled with the strong WPS and CMF benzene detection, metadata, and contextual data, indicate the strong possibility for a benzene plume.

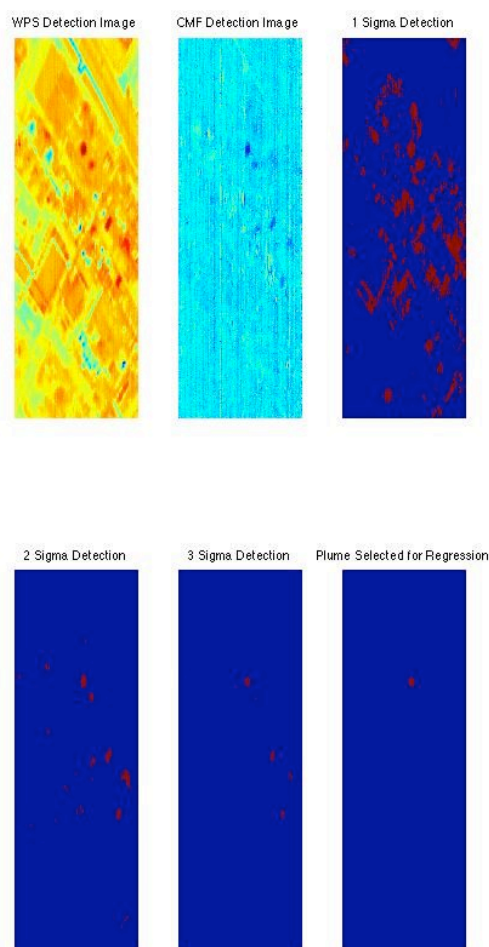


Figure 129. FL 153924 WPS Plume Detection

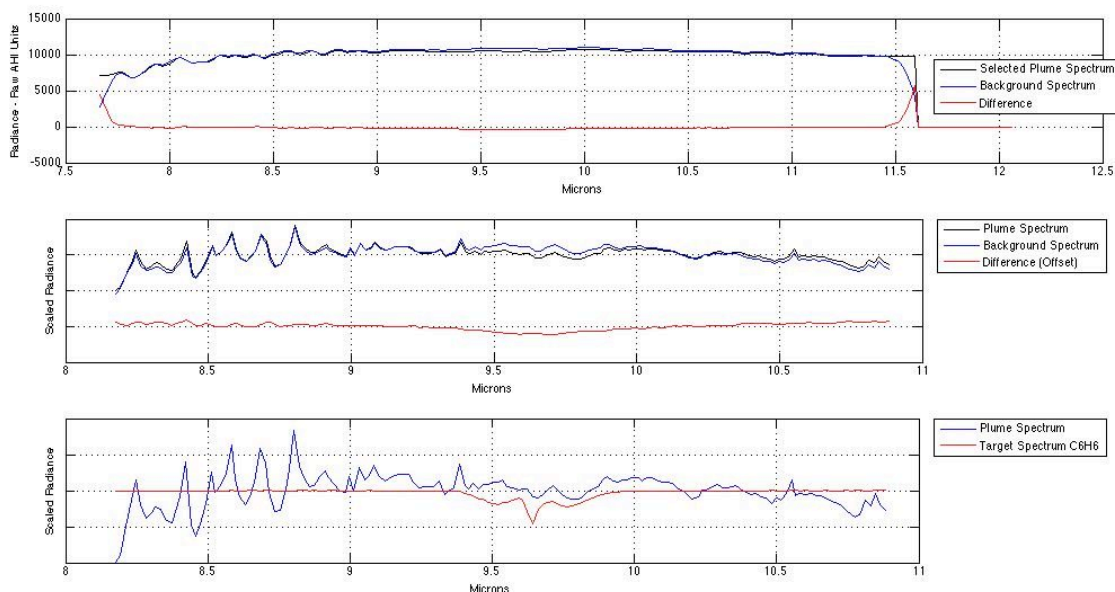


Figure 130. FL 153924 Spectral Identification of Benzene

The AHI results demonstrate that WPS plume detection has equivalent results to a similar CMF based process. In some cases visualization and detection of the plume may be improved with the WPS. The combination of both the WPS and CMF detection planes can improve plume identification and its subsequent chemical identification. The SAM algorithm, as the results in Chapter 3 illustrate, was not effective in plume detection. The SAM result clearly demonstrates that the wavelet packet subspace does indeed transform the pixel spectrum into a subspace that more easily detects the target spectrum.

Full identification through linear regression remains a separate topic and is not addressed by the WPS algorithm. The negative result of FL 150652 requires further study and a better understanding of the principal components approach used to detect the NH_3 plume in that case.

4.6 Validation of Atmospheric SO₂ Detection and Identification in AIRS Data

Linear regression techniques are typically applied to high spatial, low spectral resolution airborne hyperspectral data for gas identification. In a lower spatial resolution case such as AIRS, linear regression may not be the most suitable method for identification. Given the AIRS pixel resolution of 14 km, any pixel spectrum represents the mixing of gasses in over 196 km². A linear regression would need to take into account all of the possible atmospheric constituents that may be present across that area. Even if such information is known, the error associated with such small relative quantities of gas would quickly negate any reasonable result. Modeling and calculating the atmosphere at this larger scale is best left to complex atmospheric models such as MODTRAN or LBLRTM and detailed atmospheric analysis.

Many times gases are detected and identified in the atmosphere by examining the difference in brightness temperature of specific bands in the spectrum of interest. These bias difference methods take advantage of spectral absorption and emission features of the gas and are used for aerosol, particulate, and gas detection and identification [83; 84]. For SO₂, a simple approach would be to ratio the deep absorption at 7.35 μm with the significantly less absorption at 7.32 μm. This type of band math is commonly applied to quickly reveal features of interest.

To validate our WPS detection of atmospheric SO₂ we apply the bias difference methods applied by Carn in his measurement of volcanic SO₂ from Mt Etna [81]. The results are visually compared. The two bias differences are 1285-1345 cm⁻¹ (7.78-7.43

μm) and $1228\text{-}995\text{cm}^{-1}$ ($8.14\text{-}10.05\ \mu\text{m}$), (BD1 and BD2 respectively). These bias differences are calculated with the brightness temperature for each wavelength. The difference between the two brightness temperatures results in the difference image. Each of the bias difference methods uses the difference between a region with a significant absorption feature and a region with no absorption. For completeness, pixel spectra for each suspected plume are compared to the target spectrum in Figure 134.

AIRS 169 is the July 2003 Montserrat eruption. The WPS method and BD1 ($1285\text{-}1345\ \text{cm}^{-1}$) agree very well. The BD2 ($1228\text{-}995\text{cm}^{-1}$) image does not appear to perform well. The BD2 feature in the SO_2 spectrum is significantly weaker than the BD1 feature. This particular SO_2 plume has been identified with other sensors such as TOMS in the ultraviolet spectrum. This result demonstrates the WPS algorithm's ability to detect SO_2 . In applying the WPS in this case, depth was selected adaptively as previously discussed. Additional target spectra were assessed for this scene, but none detected the plume as clearly as SO_2 .

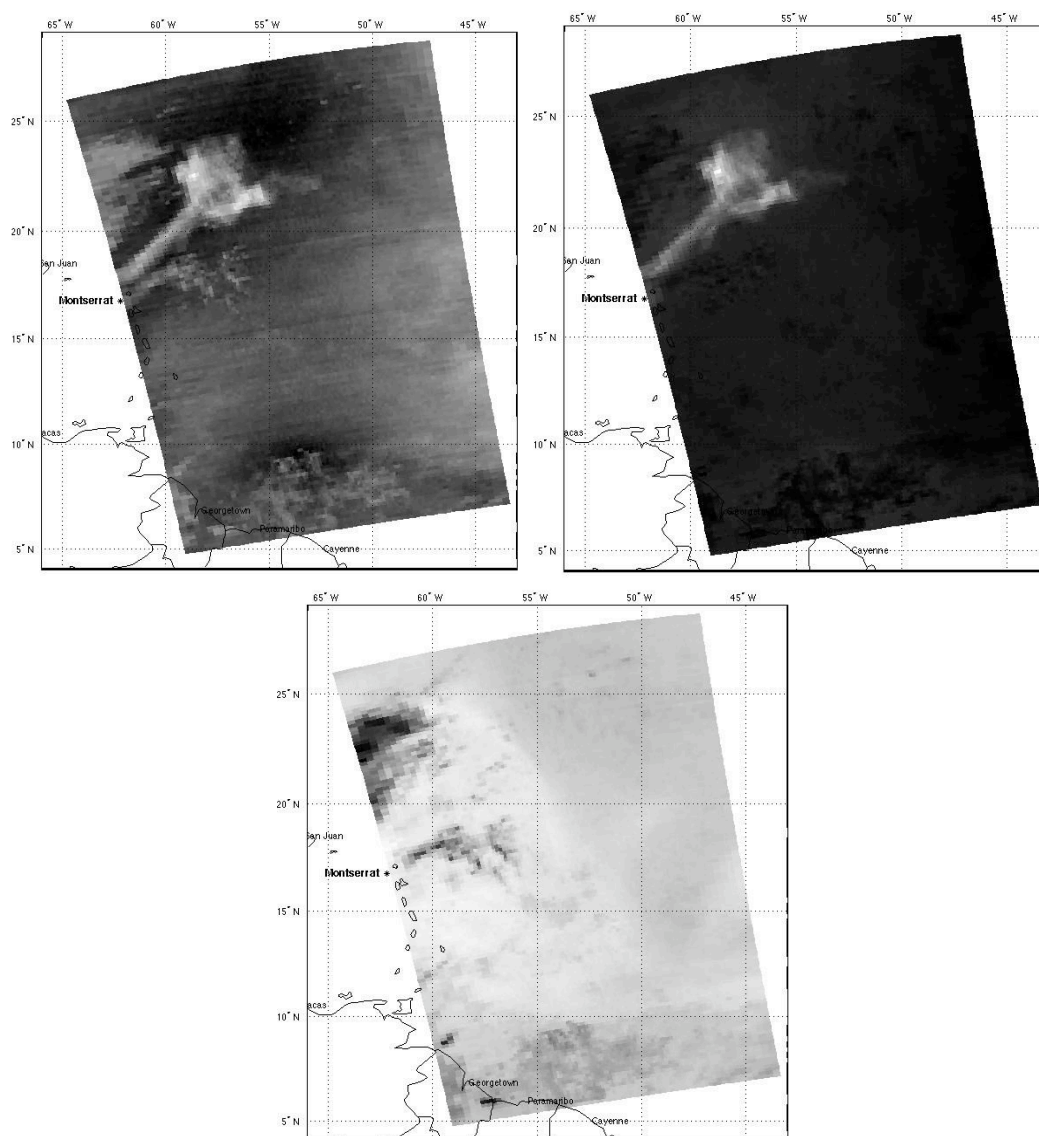


Figure 131. AIRS Data Granule 169 SO₂ Detection Comparison
 Top Left: WPS, Top Right: 1285-1345 cm⁻¹, Bottom: 1228-995 cm⁻¹

AIRS 102 contains the June 29, 2003 Iraqi sulfur fire. Again the WPS and BD1 methods detect the SO₂ plume very well. In the BD2 method there is a small darker area that coincides with the SO₂ plume of the other methods. Two of the three methods appear to agree on the location of SO₂ while the BD2 case is marginal.

AIRS 107 also contains the July 2, 2003 Iraqi sulfur fire event. In this case the WPS detection visually performs the best. The BD2 method has somewhat more contrast than the BD1 method near the source of the fire at Al-Mishraq.

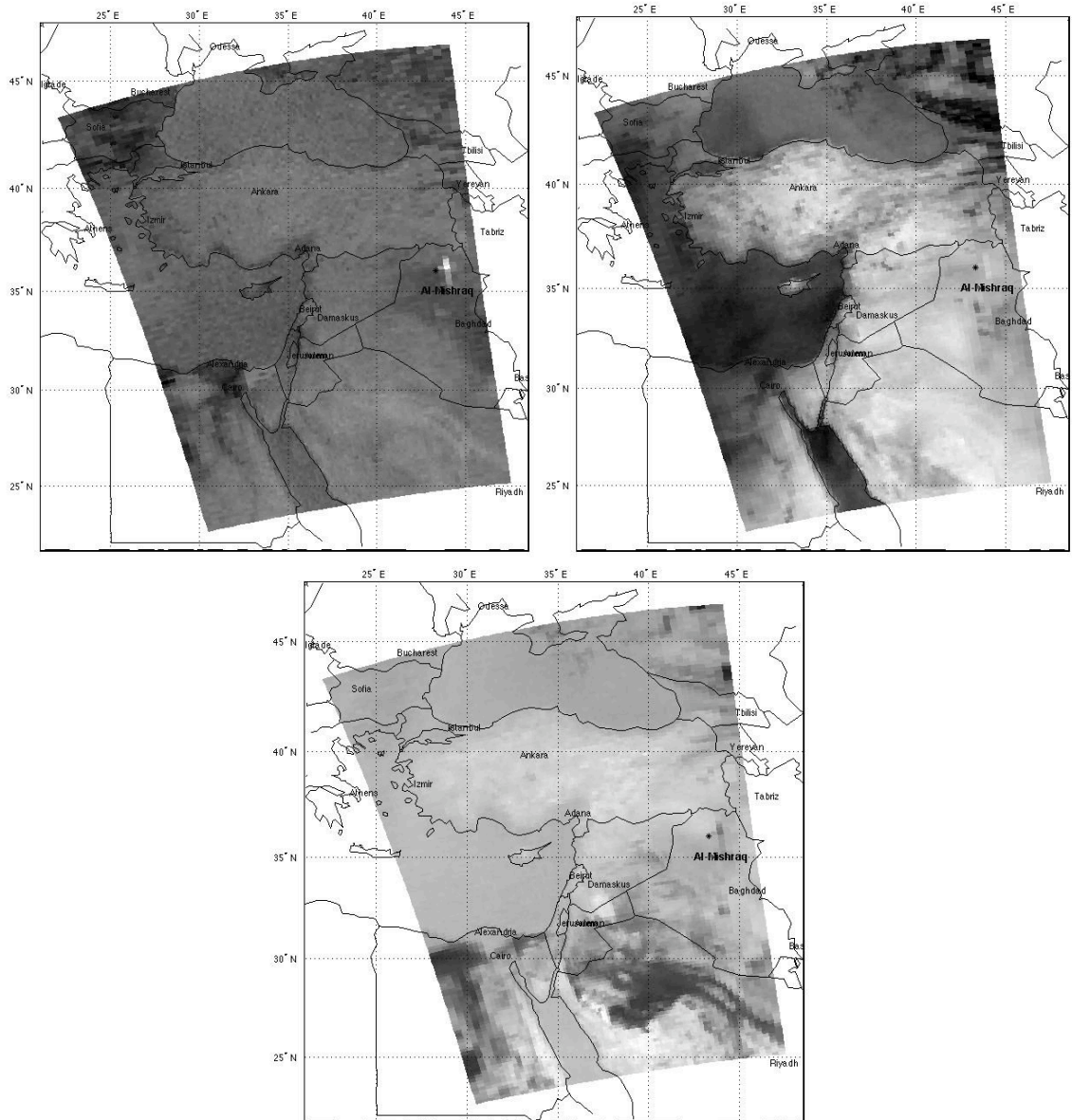


Figure 133. AIRS Data Granule 107 SO₂ Detection Comparison
 Top Left: WPS, Top Right: 1285-1345 cm⁻¹, Bottom: 1228-995 cm⁻¹

In both the AIRS 102 and 107 cases, initial results with the WPS method were inconclusive. A more detailed examination led to the wavelet coefficients selected for target signature pattern matching. The nodes selected by the process were not identical to the nodes selected for AIRS 169. Ten nodes were selected for 102 and 107, while six nodes were selected for 169. Two variables were identified that differed between the cases. 1) bad bands identified in AIRS metadata, and 2) atmospheric model runs (different run for each case). Both of these variables can change the target spectrum used for the best basis either by changing the wavelengths used or by changing the intensity of the spectral features. Since AIRS 169 was a proven case of SO₂ detection, those nodes were applied to the other data sets to generate the results of Chapter 3.

Examination of the nodes and coefficients also revealed that some nodes significantly drove the magnitude of the angle between the vectors. This is a common situation when examining the data in feature space. Eliminating those coefficients or normalizing the magnitude of the coefficients across the entire target vector, will mitigate this situation. Normalization of the coefficients across the vector will also ensure the information in the nodes selected is not lost.

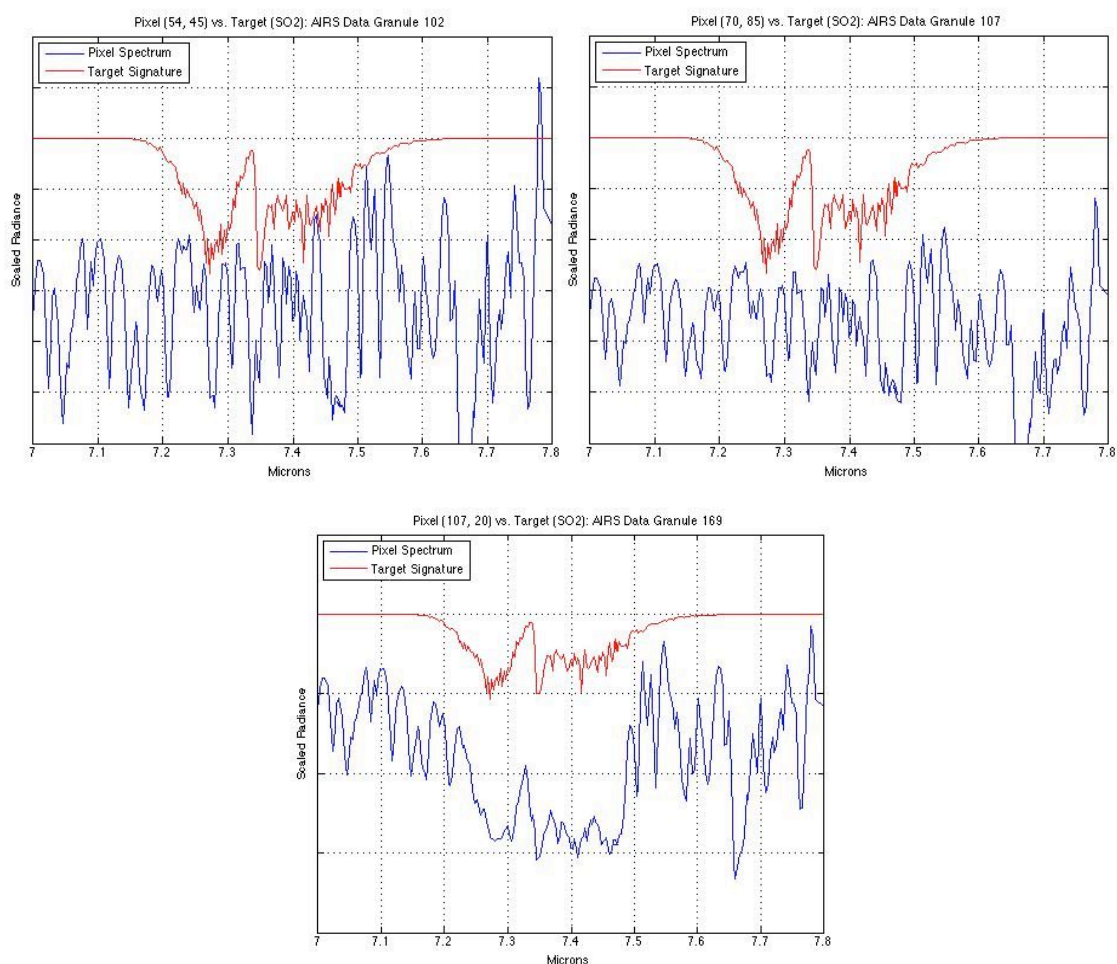


Figure 134. Pixel Spectra for AIRS SO₂ Detections

Examination of the pixel spectra leads to quick confirmation of the SO₂ plume in AIRS 169. AIRS 102 and AIRS 107 are less convincing. Linear regression was attempted on the AIRS spectra. Results were inconclusive and are not reported. The combination of the BD methods for 102 and 107 and SO₂ detections by other sensors on these dates and at these locations [81] provides convincing evidence that this is an SO₂ plume.

4.7 Computational Improvements and Issues

Initial order of operations estimates for the CMF and WPS approach suggest that the number of operations for each approach are similar in magnitude. The total magnitudes are the same, but the difference occurs in the number of operations per pixel. For the CMF approach an inverse covariance matrix, $O(N^3)$, must be calculated once for the entire data cube. Then, a multiplication of that matrix against each pixel is performed, $O(N^3) + O(m \times n)$. In the WPS approach, a wavelet packet decomposition is applied to each pixel, $O(N \log(N))$ for the decomposition, multiplied by the total number of pixels or $O(N \log(N)) \times (m \times n)$.

In the MatLab environment, this similar order of magnitude in operations did not equate to similar computational times. In general, wavelet decomposition times were at least an order of magnitude greater than the covariance matrix calculation. An AIRS CMF run took on the order of 2-4 minutes, while a similar WPS run was of the order 20-30 minutes. The following were observed during processing:

1. On a dual CPU processor, only a single processor was utilized for either CMF or WPS (multi-threading was enabled).
2. The CPU that was utilized was not at full capacity during computation of the WPS. It was at full capacity for CMF.
3. CMF required significantly more memory for computation (1.5 GB for CMF vs. 400 MB for WPS).

4. There was a significant amount of disk activity during the WPS processing.

The CMF processing times are not unexpected. All of the processing is performed in memory with significant operations reduced to large matrix multiplies and calculation of the pseudoinverse, both of which are memory and CPU intensive. CMF processing requires access to all of the data cube to perform its operations. Analyzing a data cube of 500 MB will fill at least 500 MB of memory with the data alone.

The WPS processing observations must be attributed to the MatLab environment. MatLab internal wavelet packet decomposition functions were used and the amount of background processing and bookkeeping performed is unknown. Since wavelet decomposition is a recursive matrix multiply, background processing, other bookkeeping, and multiple disk reads/writes are driving the computational time difference.

Significant differences were also noted in computational time between a dual 64-bit PowerPC processor and a dual-core 32-bit Intel processor. In many cases the 32-bit Intel Processor was significantly faster in wavelet processing than the PowerPC processor. It's assumed that the MatLab environment is optimized for one platform vice the other. In general, computational times on both a dual processor (dual-core) Intel platform and dual processor PowerPC platform were disappointing.

Regardless of the computational time, the original objective of the research has been met. The WPS approach only requires access to one pixel of data at a time. The pixel-by-pixel method for trace gas detection has been demonstrated. This approach is an easily parallelizable problem.

4.8 Image Context and Visualization Aspects

Analysis of both AHI and AIRS data required significant context to understand the results. In the AHI analysis, the first step was to identify the location of each flightline. Google Earth was used with aircraft GPS data to identify which flightline may be of interest. The other AHI metadata provided general locations and types of gases that may be present at specific facilities. Detailed examination of the location of a potential plume was required to determine if there was a source for the plume and if the identified gas was reasonable for that location. Again, Google Earth was used to get higher resolution imagery for areas of interest.

AIRS image context was slightly different. Given its global coverage, the data and its results cover wide areas. Without some geospatial context, detection and identification leads to results with low utility. These results must be mapped to a location before the context can be understood. The detection images indicate SO₂. But georegistering that result confirms that the SO₂ is the Al-Mishraq sulfur fire.

To facilitate context and understanding of the results, Google Earth is used to visualize the results of automated AIRS processing. CMF and WPS automated results are converted to KML files that can be opened in Google Earth. With the demanding computational time requirements in the prototype environment, weekly analysis of AIRS global data is performed. Results are pushed to a website where both georegistered images and KML files are available (www.earthintelligence.net).

5 Conclusions

The WPS approach is a new method for the hyperspectral data analysis toolbox. The results demonstrate that the WPS approach can be used to detect and identify trace gases in the atmosphere. Since the WPS approach is a single pixel analysis method, it facilitates the parallelization of hyperspectral data analysis. This is a tremendous benefit over matched filter methods, which require calculation of the covariance matrix from the data population. A significant benefit of this single pixel approach is its potential application real-time collection and analysis systems. The original research goals were to:

1. develop a wavelet packet hyperspectral detection method, and
2. develop an easily parallelizable method for trace gas detection in spectral data.

By demonstrating that the WPS approach performs as well as state-of-the-art gas detection methods, both goals #1 and #2 have been met.

5.1 Wavelet Packet Subspace for Spectral Trace Gas Detection

The results demonstrate that in both simulated and real hyperspectral and ultra-spectral data the WPS approach can detect trace gases. When compared to state-of-the-art

algorithms, WPS performance is comparable to and in some case exceeds those algorithms. This performance is consistently demonstrated in all three types of data through both statistical and visual methods. The AHI and AIRS results demonstrate that in both the high-spatial/low-spectral resolution and low-spatial/high-spectral resolution case, the WPS approach performs well. This is due to the target-based detection and adaptive best-basis approach that are implicit in the WPS method.

The initial selection of the Daubechies 2 mother wavelet followed a quantitative and qualitative selection process. Pure numerical results of the process led to a general direction for selection of the mother wavelet. The pragmatic results demonstrated that Daubechies 2 would perform best as the mother wavelet.

Selection of the *L1 Norm* as the cost function for the best-basis was much more subjective. The pragmatic results also indicated this cost function would perform well. Best-basis selection will require more study.

The WPS approach finds an orthogonal subspace that best represents the target of interest. In doing so it implicitly suppresses background contributions of the scene. This approach parallels the fundamental problem in hyperspectral gas detection, separation of the target from the background.

5.2 Computational Improvements

The order of operations for the WPS is comparable to state-of-the-art detection methods. Of significant difference is the pixel based approach of WPS. This approach

transforms the memory intensive processing of matched filter techniques into an easily parallelizable problem. This EPP is a much more appropriate solution to real-time spectral collection and analysis systems. Real-time systems can more easily make use of their multi-processor designs and architectures by simply transmitting and processing individual pixel vectors as opposed to very large matrices.

5.3 Context & Visualization

Hyperspectral analysis made significant use of Google Earth to provide context and visualization. Geospatial tools such as Google Earth are critical in developing high-confidence results. Without geospatial context and/or high resolution imagery, atmospheric gas detections may make little sense. Validation of many of the results herein would not have been possible without geospatial context. Google Earth is now used as part of an automated detection process for AIRS data developed to implement the WPS approach.

5.4 Future Research

This work has applied a well known signal processing technique to hyperspectral data analysis. There are numerous variations in both orthogonal subspaces and cost functions that the wavelet packet subspace could apply. This work has focused on a basic

approach driven by the target characteristics of LWIR hyperspectral data. Future research in this area will include:

1. Application of the WPS to more LWIR data with different spectral and spatial resolutions,
2. Application of the WPS to reflective hyperspectral data analysis,
3. Further study of wavelet selection for reflective data,
4. Study of different cost functions and their impact on algorithm, performance, and
5. Implementation of the WPS on a high performance computing platform for real-time hyperspectral analysis.

It is hoped that this research encourages others in the hyperspectral community to examine the WPS approach and to use it for analysis of spectral data.

Appendix 1. Previous Wavelet Based Hyperspectral Analysis

Experiment 1 performs classification of hyperspectral data using in-scene pixels as the target of interest. This case uses AVIRIS data of Moffett Field, CA in the reflective region of the electro-magnetic spectrum. A comparison is made of existing hyperspectral classification methods using reflectance values versus the same classification methods on Daubechies 4 discrete wavelet transform coefficients. This case also compares the impact of dimension reduction on classification methods.

Experiment 2 uses an AIRS data set in the emissive region of the electro-magnetic spectrum. This data set is of the July 13, 2003 volcanic eruption in the Soufriere Hills of Montserrat. This experiment attempts to visually identify features in the data similar to how a Principal Components Analysis is used to identify dominant features in multi-spectral or hyperspectral data. Experiment 2 attempts to detect the volcanic ash cloud in the data prior to identifying SO₂ in the ash cloud. The detection step visually compares the ability of the DWT to a PCA to detect the ash cloud.

Experiment 3 analyzes an application of the Discrete Wavelet Packet Transform (DWPT). The same AVIRIS data set from Experiment 1 is utilized. A pixel of grass and asphalt is linearly mixed to demonstrate the improvement in classification between spectral space and the DWPT.

Experiment 1: Reflective In-Scene Classification with DWT

Using an AVIRIS data set of Moffett Field, CA, four types of pixels using literal information, water, grass (fairway grass), rooftop, and road surface (probably asphalt) are chosen. Figure 1 is a close-up image of the regions chosen as known spectra. Figure 2 is a plot of the spectra for the known materials, blue-water, green-grass, magenta-rooftop, red-asphalt. These regions have been chosen for two reasons: 1) Spectral diversity, vegetation vs. non-vegetation, and water, 2) Image features. All four regions were easy to identify in image space. Once these regions were chosen, three classification methods are used to find additional pixels, which can be classified as one of the four regions of interest.



Figure1. Moffet Field Regions of Interest
(Blue - Water, Yellow - Roof, Magenta - Runway, Green - Grass)

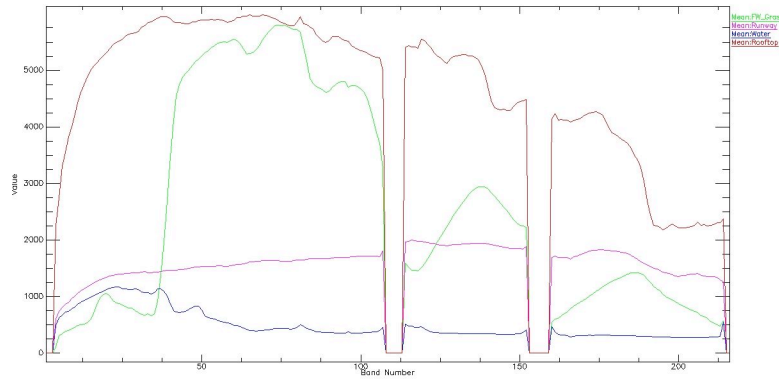


Figure 2. Spectra for Regions of Interest
(Blue - Water, Yellow - Roof, Magenta - Runway, Green - Grass)

Classification methods come in two basic types, distance measures, and statistical measures. For this experiment we will use both distance measures, (Spectral Angle and Euclidean Distance) and a statistical measure (Spectral Matched Filter).

The baseline classification maps for spectral angle, Euclidean distance, and matched filter are shown in leftmost column of Figure 3. The baseline case applies classification methods in spectral space using all the wavebands available. The fairway grass region in green is present in at least 3 other areas, which may also be golf courses. The rooftops in yellow clearly stand out in the upper right as well as the asphalt region from the airport. The water region is spread out through the middle of the image.

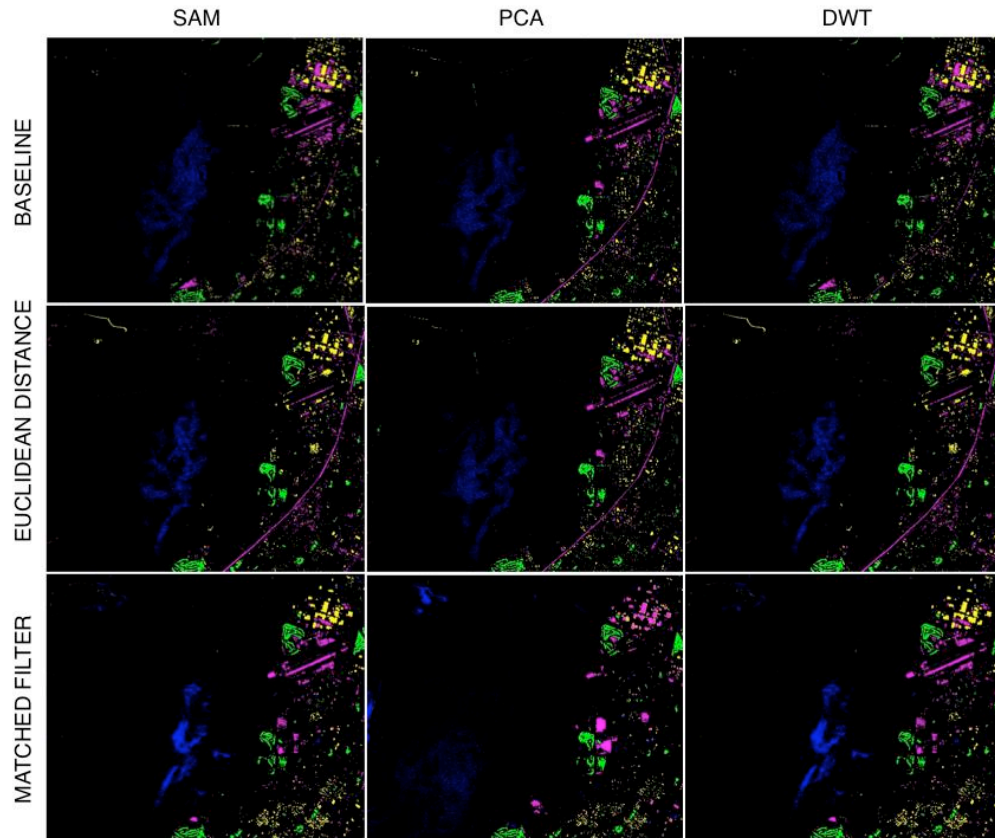


Figure 3. Classification Map Comparison, Spectral Space, PCA, DWT

In comparing the baseline cases we see that Euclidean distance finds the major roadway across the lower right of the image, while the other methods do not. The matched filter results highlights the airport features, and seems to provide a cleaner view of the golf course features.

Classification maps for the principal components case are shown in the center column of Figure 3. These figures show the all bands case with no dimension reduction. In comparing this case to the baseline case, asphalt features show up more prominently in the spectral angle and Euclidean distance maps than in the baseline case. Although the

baseline case shows more asphalt around the rooftops in the upper right. Water features are much less prominent in this case.

The matched filter results for this case (all bands) show some questionable results. Rooftop features appear to be miss-classified as asphalt, and two large features show up in the middle right portion of the image. These results may be due to Hughes Phenomena or an error in the processing. Water features are very dissimilar.

The wavelet transform case classification maps (all bands) are shown in the rightmost column of Figure 3. The wavelet transform case shows excellent correlation with the baseline case. The asphalt features in all three classification methods are very similar, as are the water, rooftop, and grass. Reviewing raw tabulated data, wavelet results for spectral angle and Euclidean distance was excellent.

In addition to these results, classification was also performed on a reduced dimensionality set of data. This step is typical of hyperspectral dimension reduction using PCA. Of the original 224 bands, reduced sets of 2, 4, 8, 16, and 32 bands are used and the same classification methods are performed. (Spectral matched filter compares only 16 and all bands.) Results are tabulated in Table 1. Overall classification accuracy for all four regions (asphalt, grass, rooftop, water) is captured in the *Overall* result column.

Table 1. Euclidean Distance and Spectral Angle Comparison

PCA - Spectral Angle				Wavelet Transform - Spectral Angle		
Bands	Overall	Commision	Omission	Overall	Commision	Omission
2	93.5	70.98	71.47	94.39	24.11	23.98
4	96.35	35.18	35.43	95.79	19.79	19.51
8	97.01	29	29.93	97.68	9.58	9.79
16	97.01	29	29.93	98.22	6.28	5.5
32	99.86	0.06	0	98.76	3.1	4.13
all	97	29.42	29.42	100	0	0
PCA - Euclidean Distance				Wavelet Transform - Euclidean Distance		
Bands	Overall	Commision	Omission	Overall	Commision	Omission
2	98.06	1.25	1.02	97.65	15.64	15.91
4	99.03	0.64	0.45	98.36	5.8	5.89
8	99.51	0.29	0.06	98.89	0.96	0.58
16	99.88	0	0.06	99.13	0.77	0.42
32	96.99	29.43	29.52	99.31	0.38	0.13
all	100	0	0	100	0	0
PCA - Spectral Matched Filter				Wavelet Transform - Spectral Matched Filter		
Bands	Overall	Commision	Omission	Overall	Commision	Omission
Compared to Baseline Euclidean Distance						
16	95.22	8.61	9.57	95.54	10.92	11.49
all	95.23	8.61	9.57	95.42	18.05	17.73
Compared to Baseline Spectral Angle						
16	94.64	33.06	33.83	94.94	33.41	33.9
all	94.64	33.06	33.83	95.02	36.22	36.04
Compared to Baseline Spectral Matched Filter						
16	95.71	17.37	18.56	96.36	17.53	18.37
all	95.71	17.37	18.56	99.99	0	0

Experiment 2: Emissive In-Scene Detection with DWT

Experiment 2 attempts to detect a volcanic ash cloud. A typical use of PCA in multi-spectral or hyperspectral data is to visually detect principal features in the data. In a reflective data set this may be water features, or roads, etc. The data set used in this experiment is AIRS data from a 2003 eruption in Montserrat.

Various satellite sensors were able to capture the volcanic ash cloud as it traveled through the atmosphere. This included Landsat ETM, TOMS and AIRS. Volcanic Ash Clouds are a significant hazard to aviation and detection of volcanic sulfur dioxide is a method used to track the dispersion of volcanic ash. TOMS (Total Ozone Mapping Spectrometer) measured the sulfur dioxide content of the volcanic ash cloud, see Figure 4. This data will be used to visually verify the ash cloud.

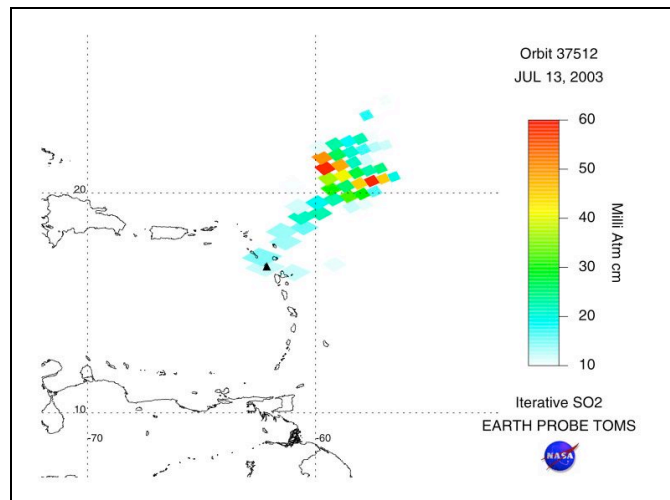


Figure 4. TOMS SO₂ Measurement of Soufriere Hills Eruption⁴

The AIRS data cube were both transformed by the PC rotation and wavelet transform. AIRS data includes topography and latitude/longitude for each pixel. This provided a reference point to aid in identifying where the source of the volcanic cloud. (This AIRS data cube is spatially inverted from typical geographic coordinates.) This data cube has no underlying land masses and has a complete ocean background. In browsing the PC cube we see no significant features in the first 100 bands. In browsing

⁴ <http://toms.umbc.edu/>

the DWT cube we see a stunning image in band 25. Comparing this to the TOMS image, we clearly see a shape similar to the SO₂ cloud, see Figure 5.

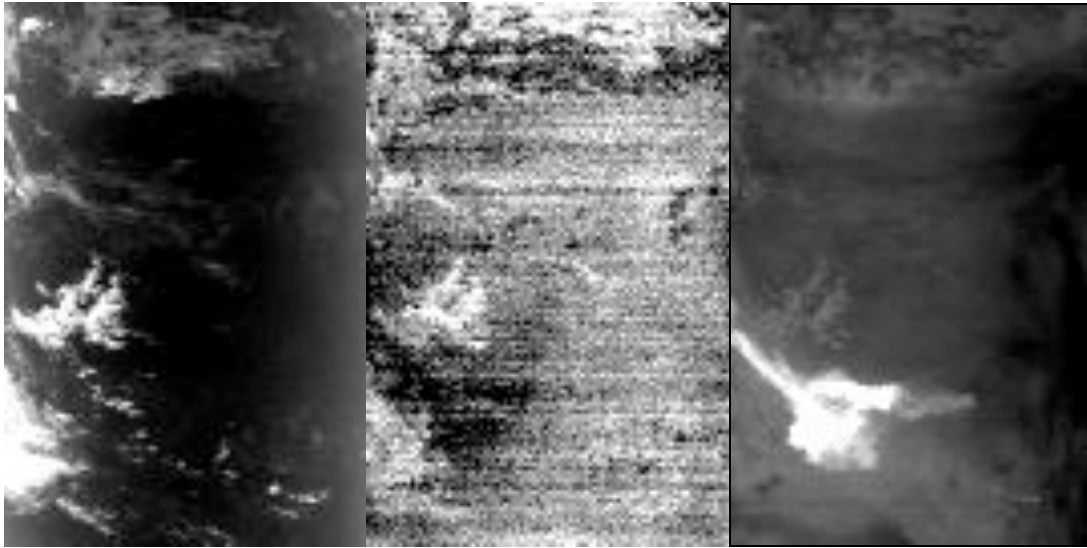


Figure 5. Original Image, PCA, and DWT Results for Detection Compared Original Data (Left), PCA Rotation (Center), DWT Coefficients (Right)

Experiment 3: Hyperspectral Classification with DWPT

Experiment 3 uses the same AVIRIS data as Experiment 1. From this data 2 pixel spectra are selected: 1) grass and 2) asphalt. These spectra are linearly mixed from 0% (all asphalt) to 100% (all grass) in 5% increments (95% asphalt, 5% grass). Three classification metrics are used to compare results. Spectral Angle (SA: angle between n-dimensional vectors), Euclidean Distance (ED: distance between n-dimensional vectors Measure), and Pearson Product Moment Correlation (PMC: a linear regression coefficient). Spectral Angle and Euclidean Distance were chosen because of their widespread application in hyperspectral analysis. Pearson was chosen because it is

insensitive to the gain and offset variations that Spectral Angle (SA) and Euclidean Distance (ED) are subject to.

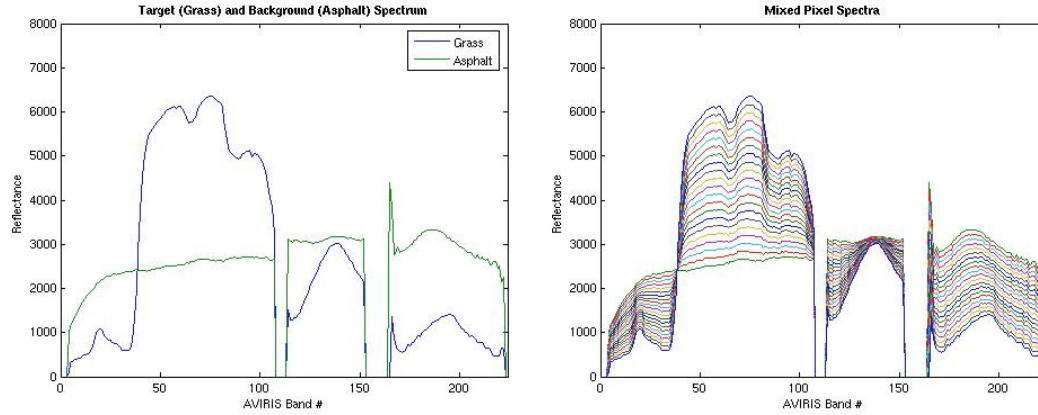


Figure 6. Target and Background Spectrum

As the pixels are linearly mixed, in Figure 6, they are classified against the target spectrum of grass. Thus at 100% grass all classification methods should show their maximum correlation, i.e. 0.0 for Spectral Angle, 0.0 for Euclidean Distance, and 1.0 for Pearson Product Moment Correlation (PMC). This classification is done in two spaces, 1) Spectral Space, 2) Wavelet Coefficient Space. In spectral space the classification is straightforward, the mixed pixel spectrum is compared to the target spectrum (grass), and a scalar value is calculated.

In Wavelet Coefficient Space, a target based approach is used. A DWPT (Daubechies 5) is applied to the target spectrum, (grass spectrum). From this basis set, a best basis is chosen using Log Energy as a cost function^{1,1}. The same DWPT is applied to the mixed pixel spectrum. The best basis developed from the target spectrum is then

used to select wavelet coefficients from both the target spectrum and the mixed pixel spectrum. These two sets of wavelet coefficients are then compared using the three classification methods.

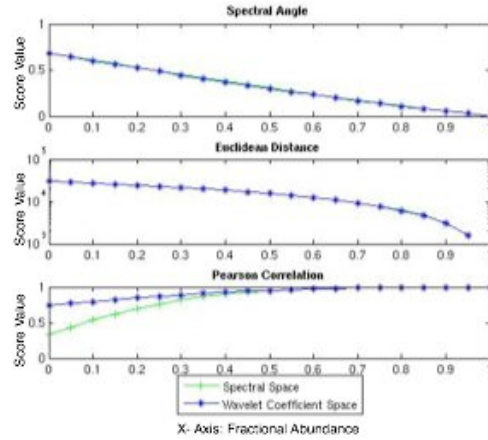


Figure 7. Results Comparison No Noise Case

The results in Figure 7 show SA and ED to be identical. PMC for low percentage of target material (0.0-0.3) shows better correlation for Wavelet Coefficients than in spectral space. This is not unexpected as the basis is chosen from the significant features of the target spectrum. This higher correlation indicates the asphalt spectrum has features that are very similar to the target (grass). In reviewing the two spectra, this is clear. The next set of results adds 5% of random noise to the mixed pixel, and then 30%.

For the 5% noise case (Figure 8, left), Wavelet Coefficient Space demonstrates slight improvement for all classification methods. For the 30% case (Figure 8, right), the improvement in Wavelet Coefficient Space is much more pronounced.

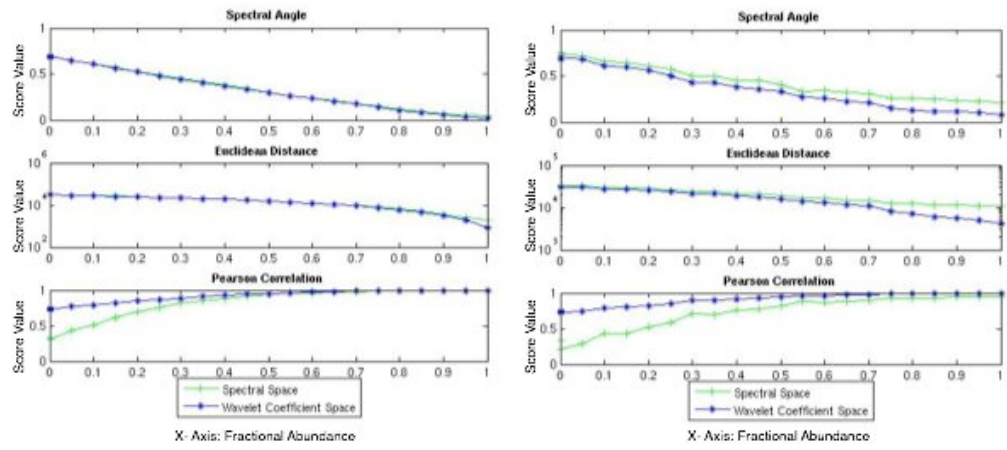


Figure 8. Results Comparison 5% (Left) and 30% (Right) Random Noise Case

Appendix 2. Simulated Data Cube Images and Spectra

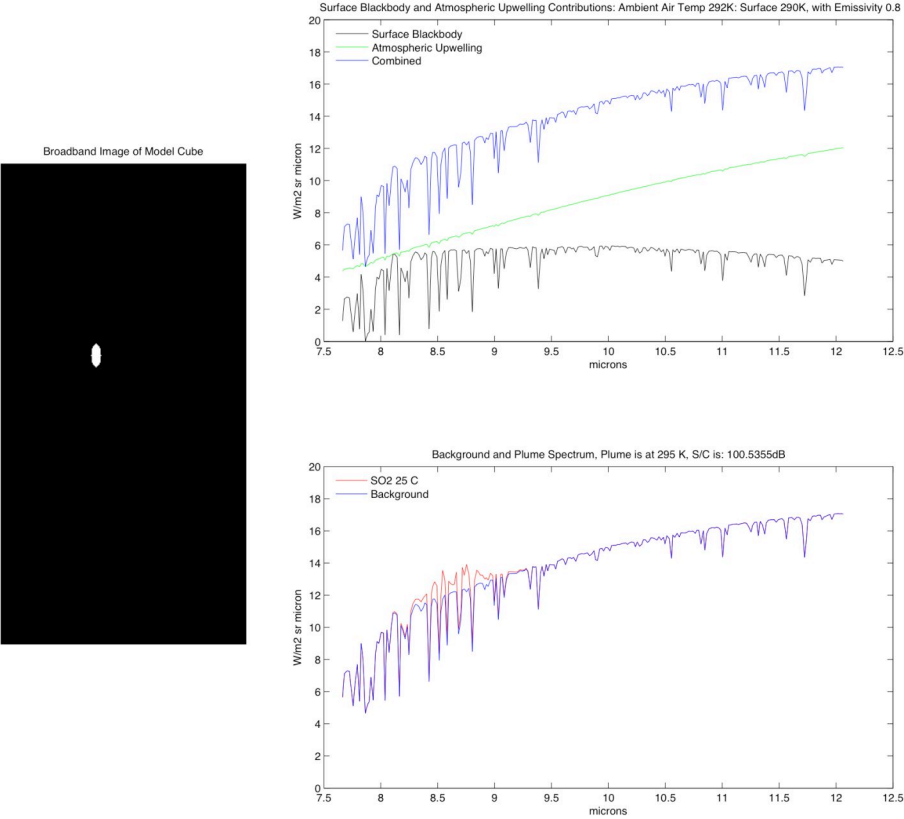


Figure 1. Simulated Data Cube 1

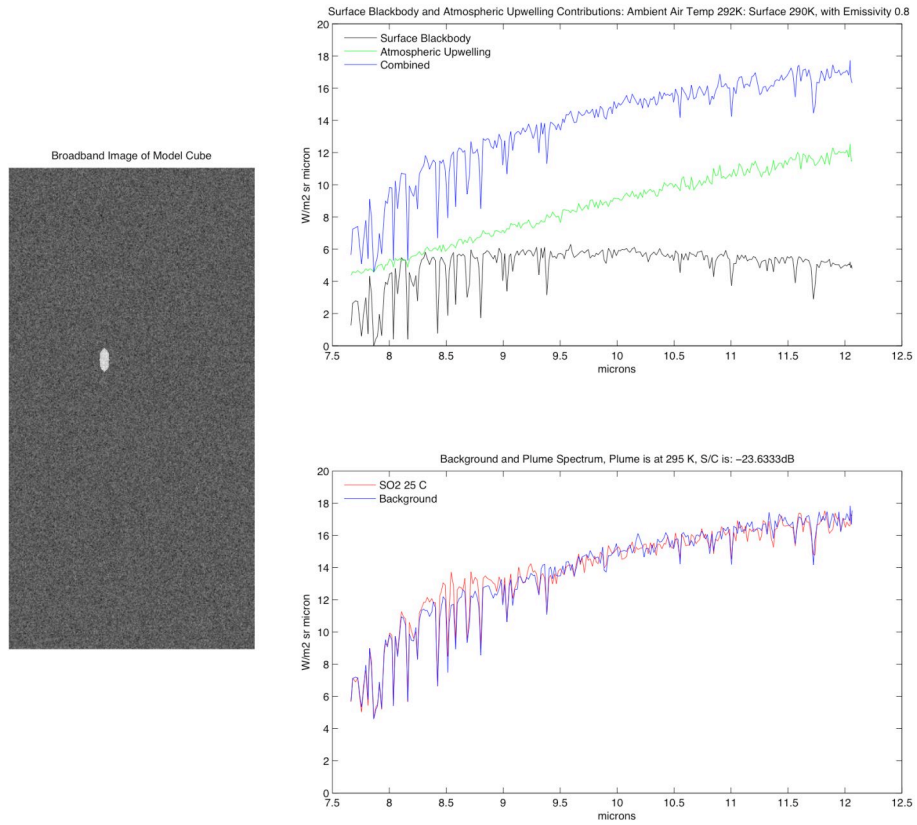


Figure 2. Simulated Data Cube 2

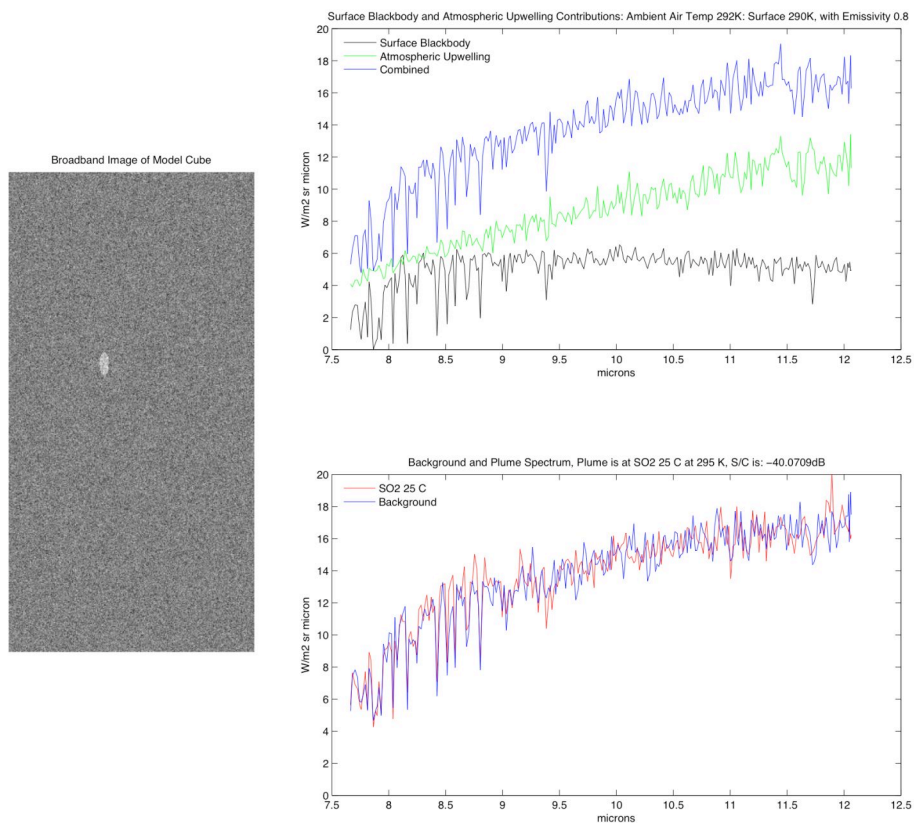


Figure 3. Simulated Data Cube 2A

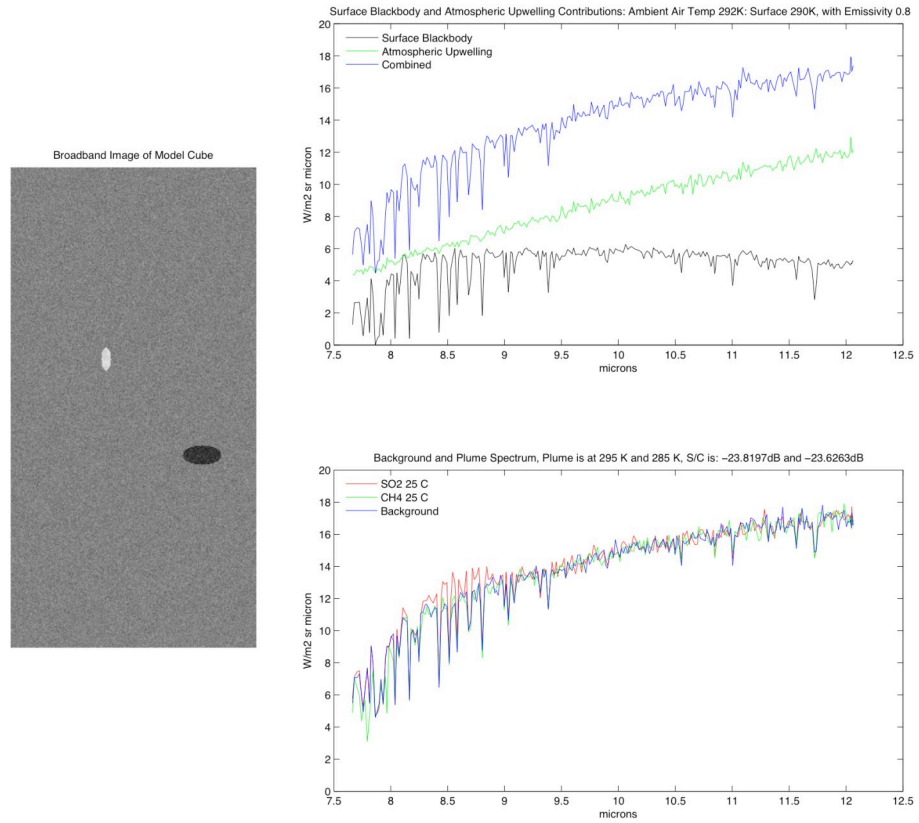


Figure 4. Simulated Data Cube 3

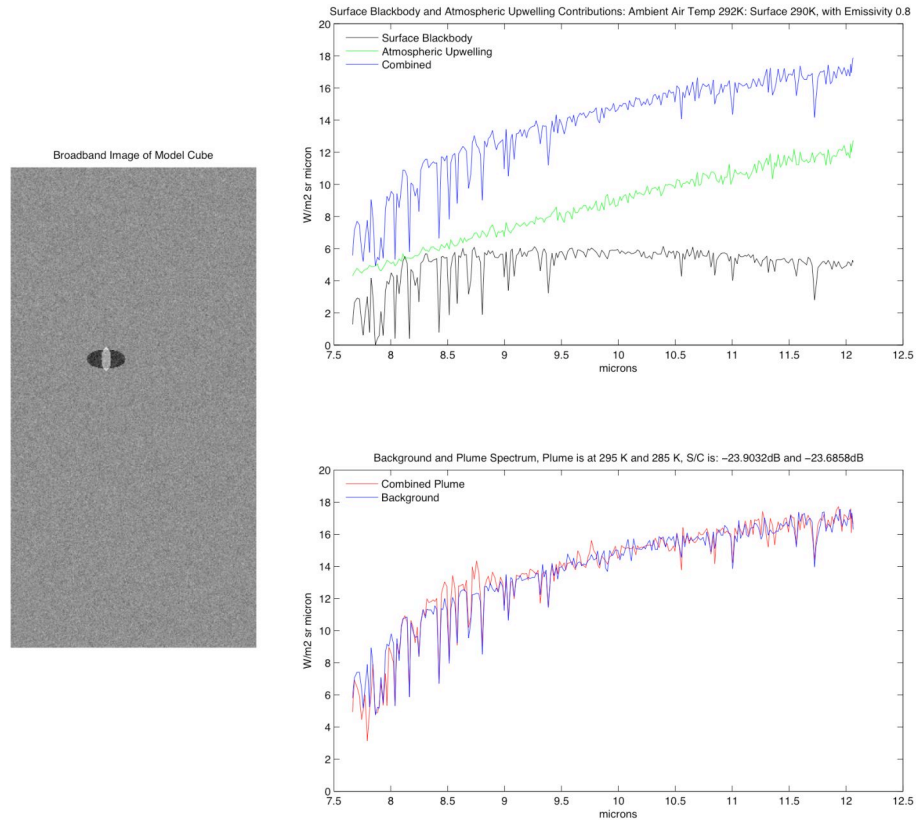


Figure 5. Simulated Data Cube 4

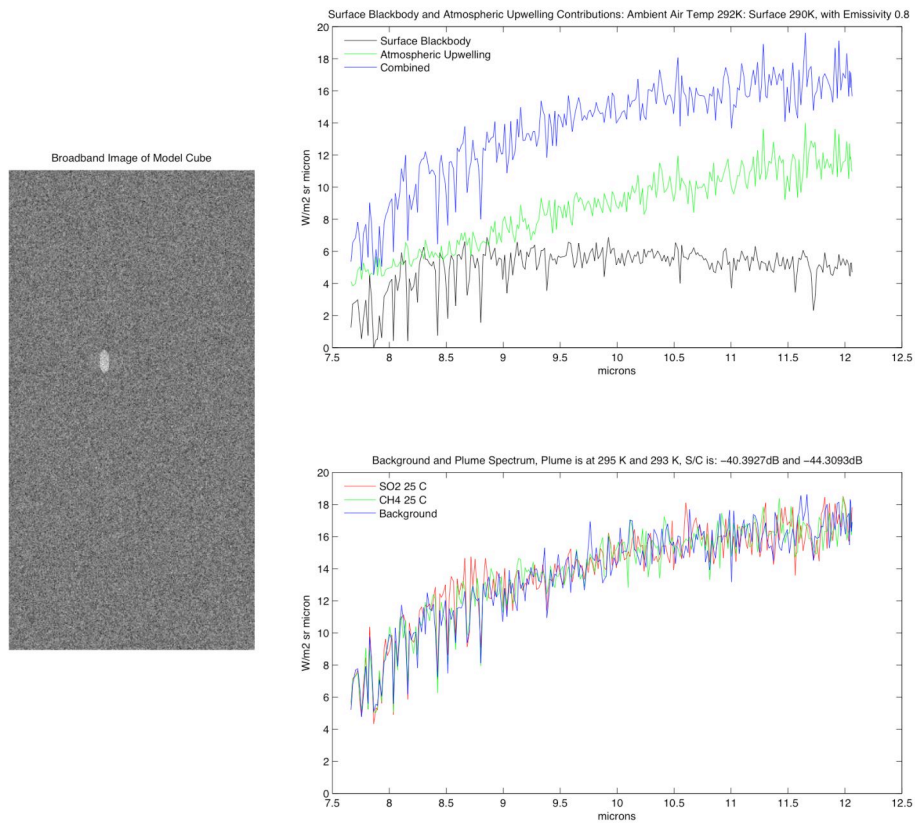


Figure 6. Simulated Data Cube 4A

Appendix 3. Additional Wavelet Selection Results

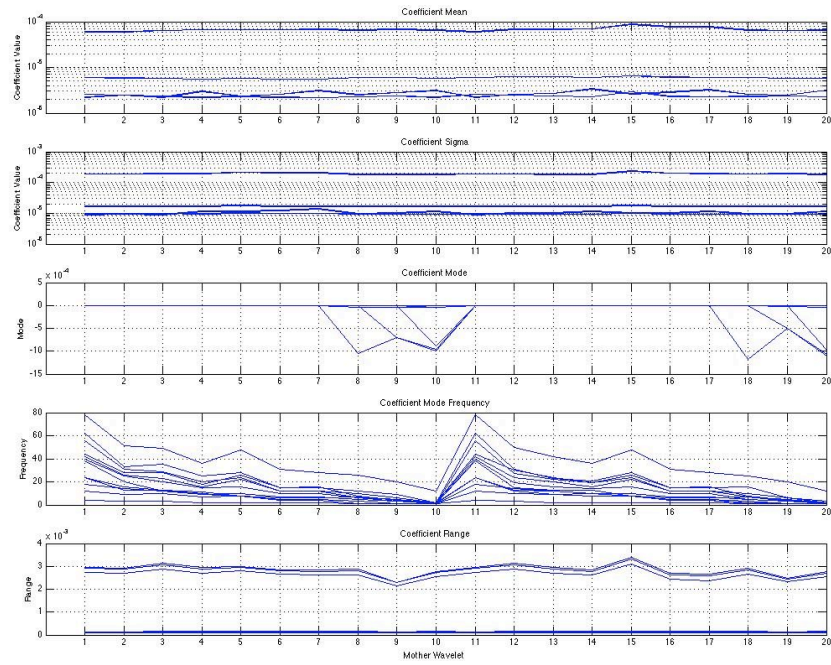


Figure 1. Wavelet Coefficient Statistics for AHI Wavelengths (Biorthogonal Wavelets)

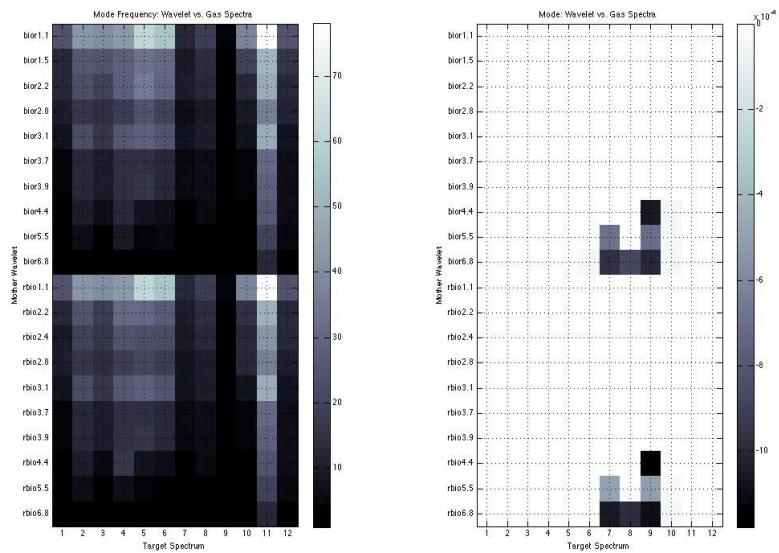


Figure 2. Wavelet Mode Image for AHI Wavelengths (Biorthogonal Wavelets)

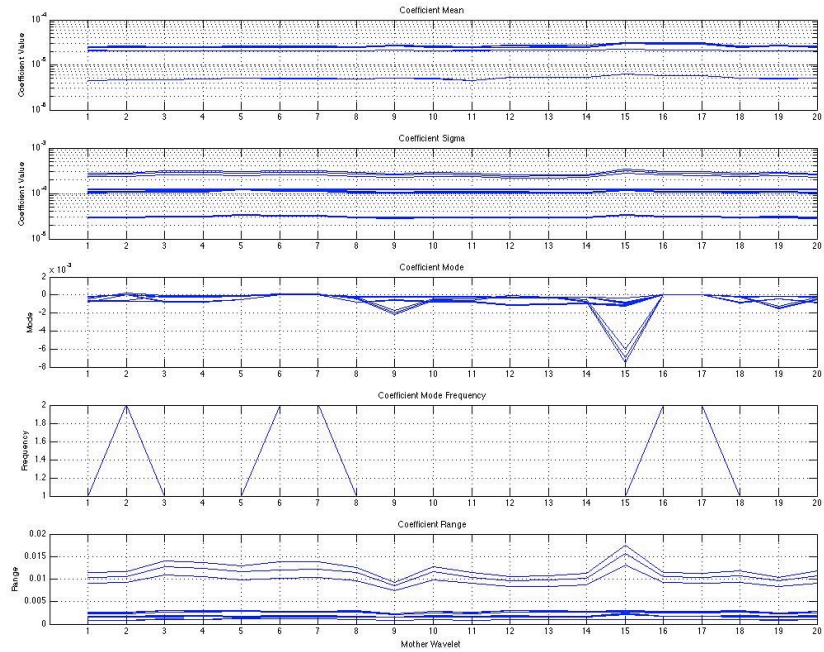


Figure 3. Wavelet Coefficient Statistics for AIRS Wavelengths (Biorthogonal Wavelets)

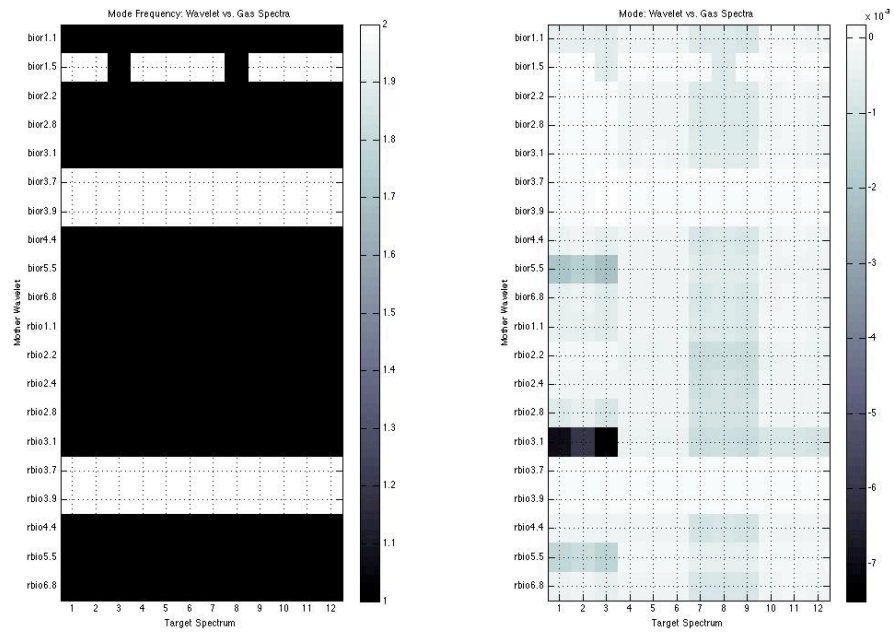


Figure 4. Wavelet Mode Image for AIRS Wavelengths (Biorthogonal Wavelets)

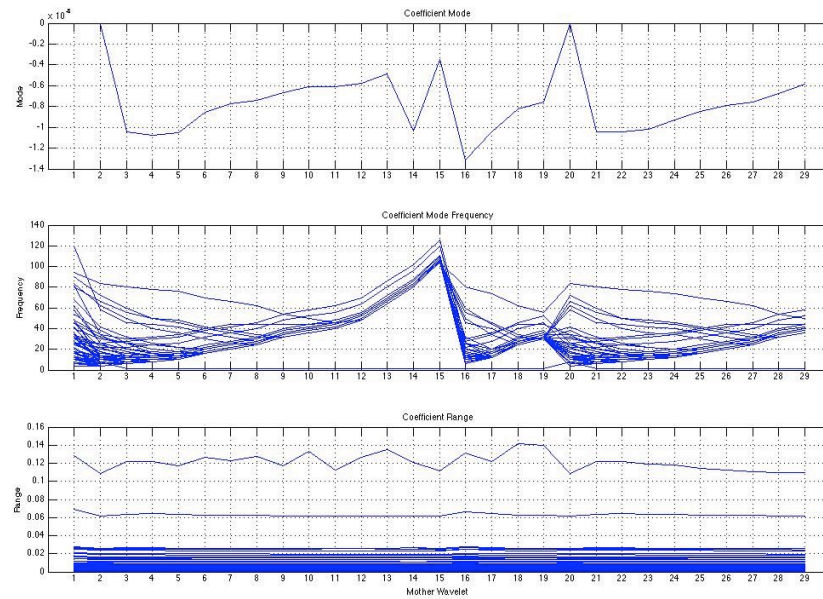


Figure 5. Wavelet Coefficient Statistics for AHI Wavelengths (Orthogonal Wavelets Wavelets – All Spectra)

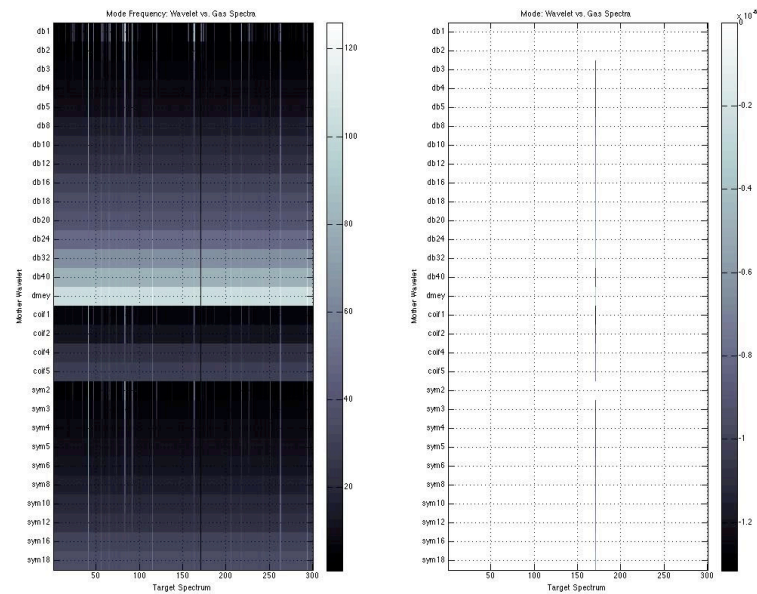


Figure 6. Wavelet Mode Image for AHI Wavelengths (Orthogonal Wavelets – All Spectra)

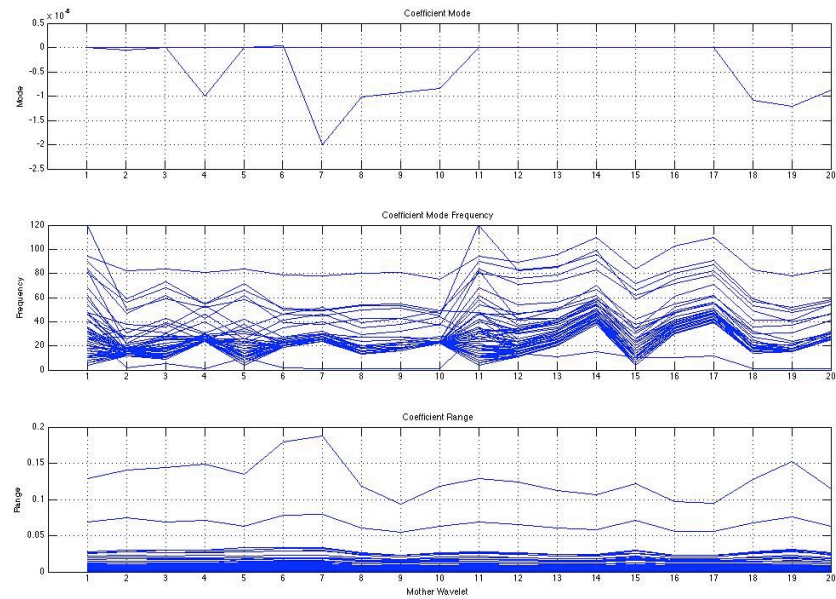


Figure 7. Wavelet Coefficient Statistics for AHI Wavelengths (Biorthogonal Wavelets Wavelets – All Spectra)

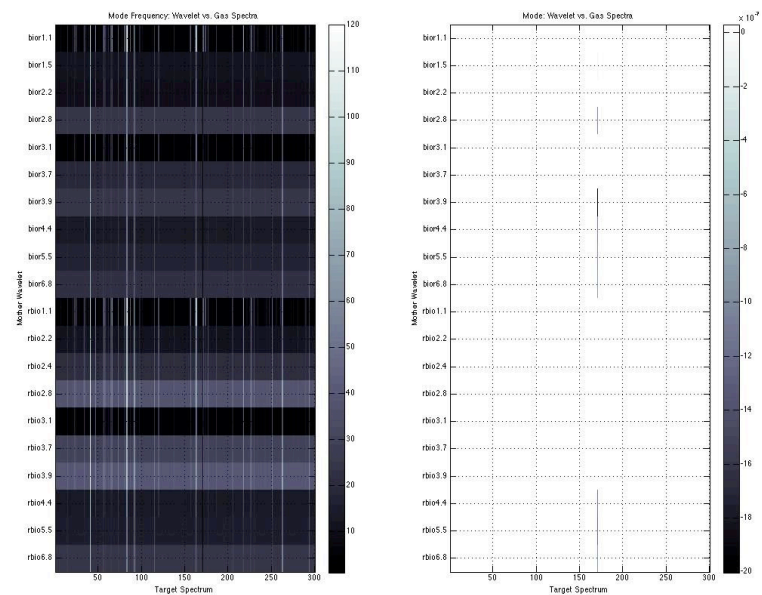


Figure 8. Wavelet Mode Image for AHI Wavelengths (Biorthogonal Wavelets Wavelets – All Spectra)

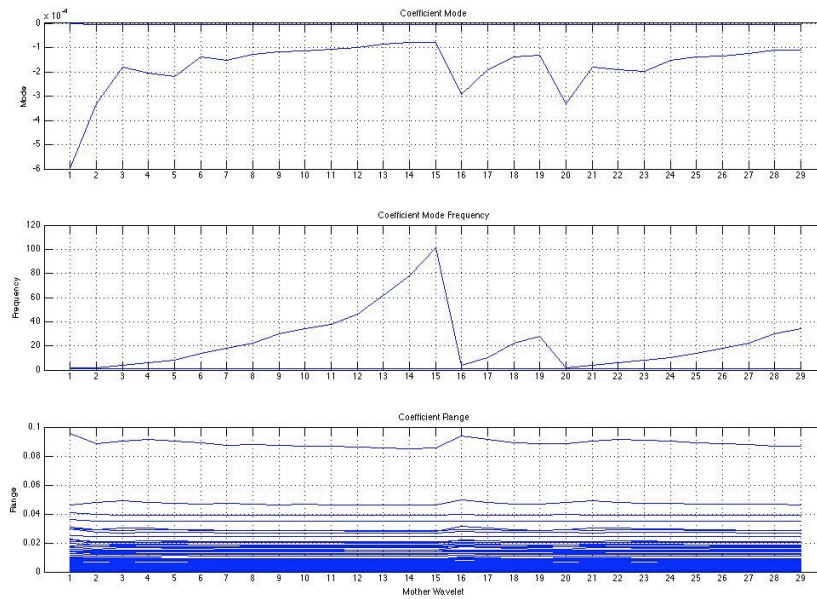


Figure 9. Wavelet Coefficient Statistics for AIRS Wavelengths (Orthogonal Wavelets Wavelets – All Spectra)

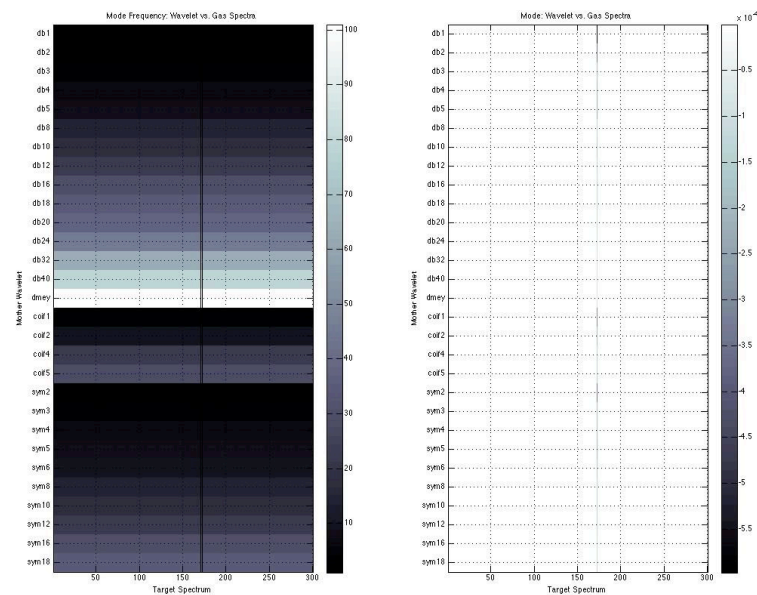


Figure 10. Wavelet Mode Image for AIRS Wavelengths (Orthogonal Wavelets Wavelets – All Spectra)

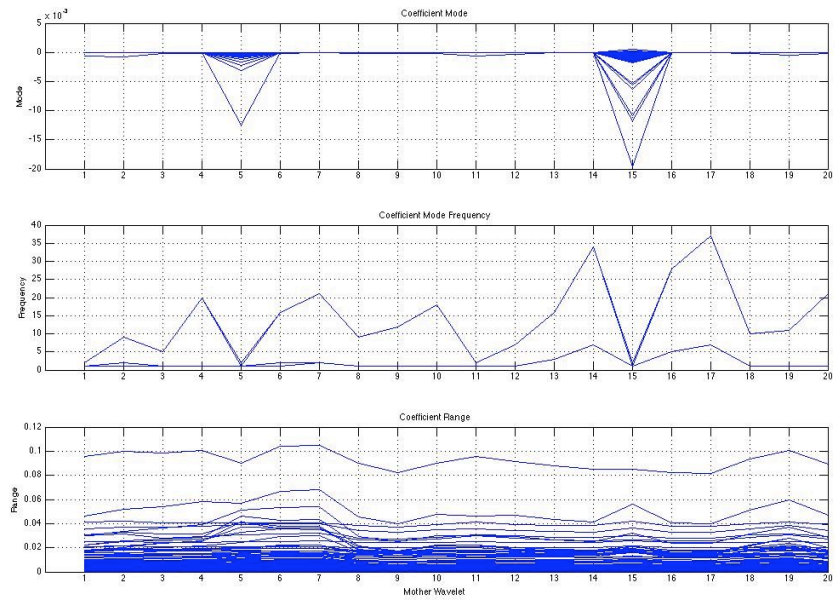


Figure 11. Wavelet Coefficient Statistics for AIRS Wavelengths (Biorthogonal Wavelets Wavelets – All Spectra)

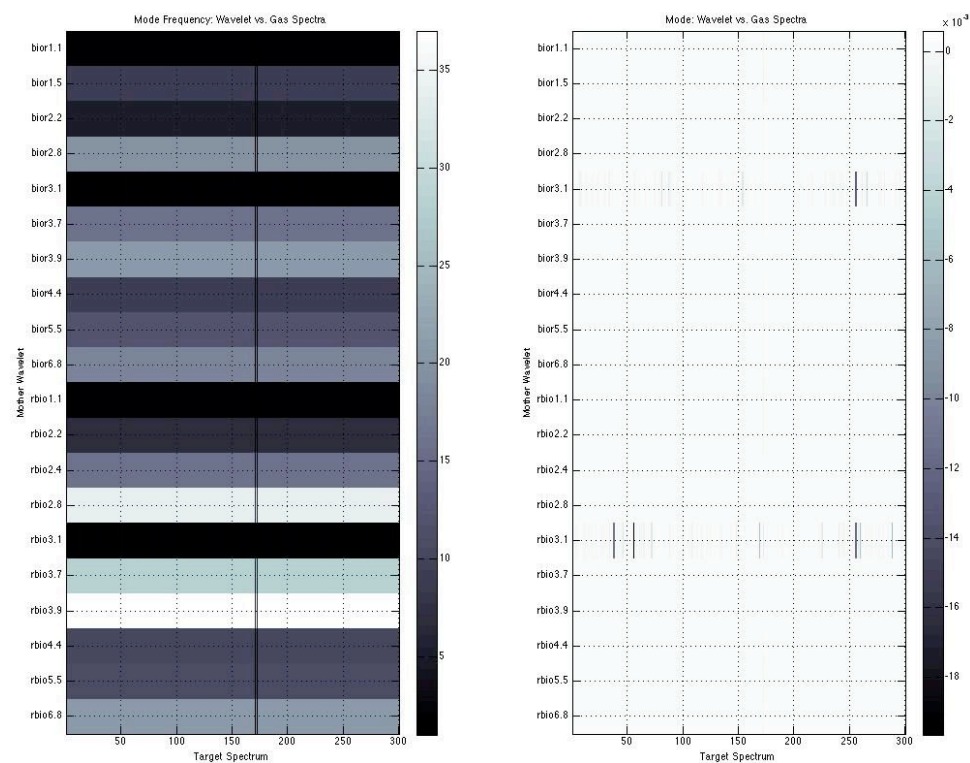


Figure 12. Wavelet Mode Image for AIRS Wavelengths (Biorthogonal Wavelets Wavelets – All Spectra)

Appendix 4. Additional AIRS Results

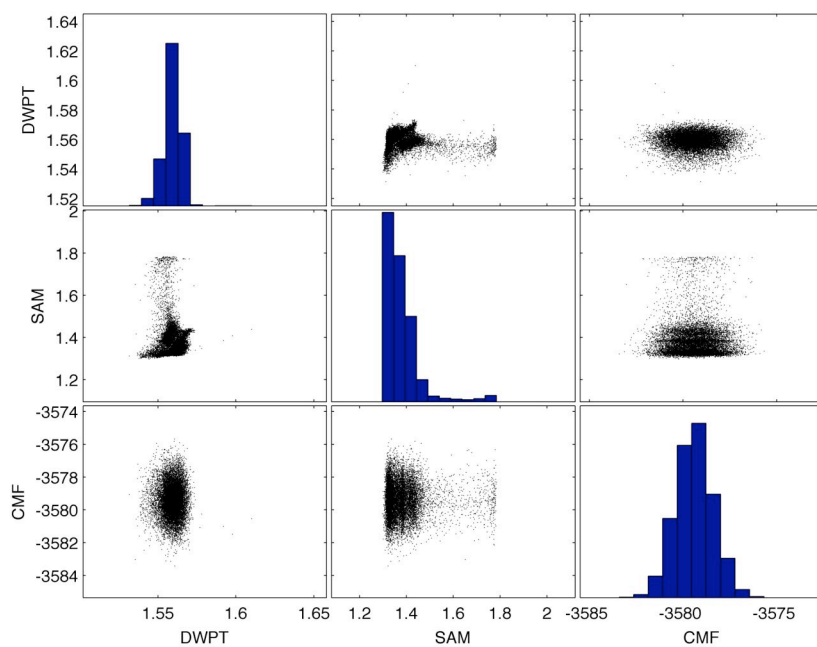


Figure 1. Scatter Plot for AIRS Data Granule 107

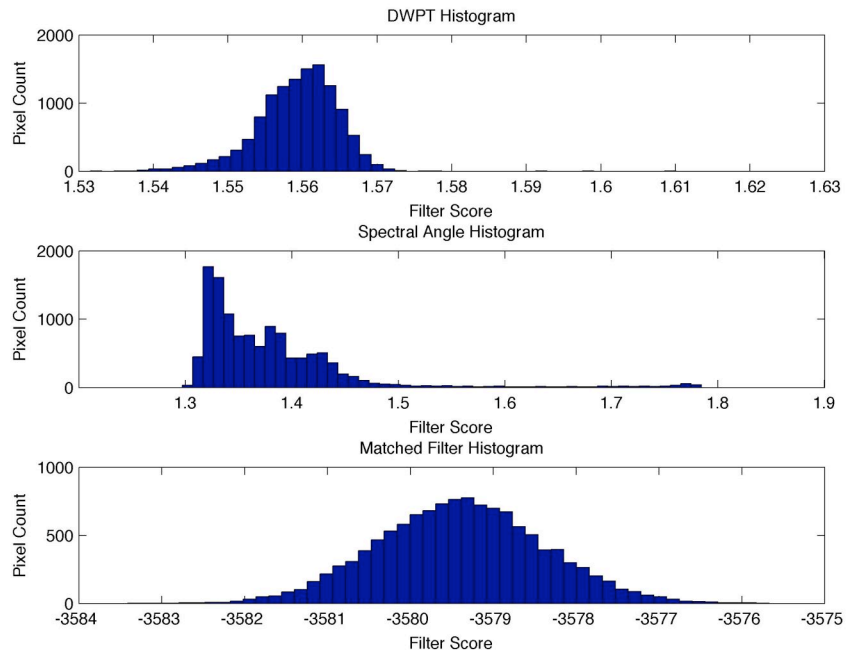


Figure 2. Histogram for AIRS Data Granule 107

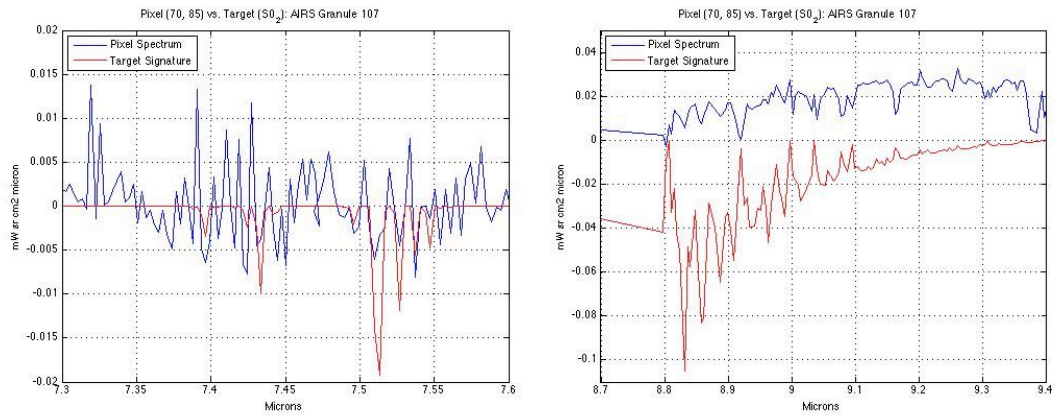


Figure 3. Plume SO_2 Spectra Comparison for AIRS Data Granule 107

Appendix 5. Detection and Regression Libraries

Table 1. Detection Library

Chemical Name	Chemical Formula	Temperature
Benzene	C ₆ H ₆	25C
Butane	C ₄ H ₁₀	25C
Ethane	C ₂ H ₆	25C
Methane	CH ₄	5C
Methane	CH ₄	25C
Methane	CH ₄	50C
Propane	C ₃ H ₈	5C
Propane	C ₃ H ₈	25C
Propane	C ₃ H ₈	50C
Ethene	C ₂ H ₄	25C
Isobutane	C ₄ H ₁₀	25C
Isoprene	C ₅ H ₈	5C
Isoprene	C ₅ H ₈	50C
Methylethylbutene	C ₂ H ₅ C(CH ₃)	25C
Ammonia	NH ₃	25C
Ammonia	NH ₃	50C
Propene	C ₃ H ₆	25C
Sulfur Dioxide	SO ₂	25C

Table 2. Regression Library

Chemical Name	Chemical Formula	Temperature
Benzene	C ₆ H ₆	25C
Butane	C ₄ H ₁₀	25C
Ethane	C ₂ H ₆	25C
Propane	C ₃ H ₈	25C
Methane	CH ₄	25C
Ethene	C ₂ H ₄	25C
Isobutane	C ₄ H ₁₀	25C
Sulfur Dioxide	SO ₂	25C

LIST OF REFERENCES

LIST OF REFERENCES

- [1] Landgrebe, D. A., "Some fundamentals and methods for hyperspectral image data analysis," Systems and Technologies for Clinical Diagnostics and Drug Discovery II, vol. 3603, pp. 104-113, 21 April 1999 1999.
- [2] Gomez, R. B., "Key Issues of Hyperspectral Sensing Technology Applications," Proceedings of the AIAA Space Technology Conference & Exposition, 28-30 September 1999 1999.
- [3] Shaw, G. A. and Burke, H.-h. K., "Spectral Imaging for Remote Sensing," Lincoln Laboratory Journal, vol. 14, pp. 3-28, 2003.
- [4] Meer, F. v. d. and Jong, S. M. d., Imaging spectrometry : basic principles and prospective applications. Dordrecht ; Boston: Kluwer Academic Publishers, 2001.
- [5] Shepherd, G. G., Spectral imaging of the atmosphere. San Diego, Calif.: Academic Press, 2002.
- [6] King, M. D. and Greenstone, R., 1999 EOS reference handbook : a guide to NASA's Earth Science Enterprise and the Earth Observing System. Greenbelt, Md.: NASA/Goddard Space Flight Center, 1999.
- [7] Granger, S. L., Leroy, S. S., Manning, E. M., Fetzer, E. J., Oliphant, R. B., Braverman, A., Sung-Yung, L., and Lambigtsen, B. H., "Development of Level 3 (gridded) products for the Atmospheric Infrared Sounder (AIRS)," in Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004 IEEE International, 2004, pp. 2506-2509 vol.4.
- [8] Li, J., Savtchenko, A., and Qin, J., "Digital imaging and gridding of AIRS Visible/NIR data," in Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004 IEEE International, 2004, pp. 4850-4853 vol.7.
- [9] Hapke, B., Theory of reflectance and emittance spectroscopy. Cambridge [England] ; New York, NY, USA: Cambridge University Press, 1993.

- [10] Rothman, L. S., Jacquemart, D., Barbe, A., Chris Benner, D., Birk, M., Brown, L. R., Carleer, M. R., Chackerian, J. C., Chance, K., Coudert, L. H., Dana, V., Devi, V. M., Flaud, J. M., Gamache, R. R., Goldman, A., Hartmann, J. M., Jucks, K. W., Maki, A. G., Mandin, J. Y., Massie, S. T., Orphal, J., Perrin, A., Rinsland, C. P., Smith, M. A. H., Tennyson, J., Tolchenov, R. N., Toth, R. A., Vander Auwera, J., Varanasi, P., and Wagner, G., "The HITRAN 2004 molecular spectroscopic database," vol. 96, pp. 139-204, 2005.
- [11] Andrews, D. G., An introduction to atmospheric physics. Cambridge, UK ; New York, NY: Cambridge University Press, 2000.
- [12] Thomas, G. E. and Stamnes, K., Radiative transfer in the atmosphere and ocean. Cambridge ; New York: Cambridge University Press, 1999.
- [13] Ward, P. D. and Brownlee, D., Rare earth : why complex life is uncommon in the universe. New York: Copernicus, 2000.
- [14] MacPherson, D., "Detecting and visualizing weak signatures in hyperspectral data," 2003, pp. x, 174 leaves.
- [15] Funk, C. C., Theiler, J., Roberts, D. A., and Borel, C. C., "Clustering to improve matched filter detection of weak gas plumes in hyperspectral thermal imagery," Geoscience and Remote Sensing, IEEE Transactions on, vol. 39, pp. 1410-1420, 2001.
- [16] Theiler, J. and Foy, B. R., "Effect of signal contamination in matched-filter detection of the signal on a cluttered background," Geoscience and Remote Sensing Letters, IEEE, vol. 3, pp. 98-102, 2006.
- [17] Villeneuve, P. V., Fry, H. A., Theiler, J. P., Clodius, W. B., Smith, B. W., and Stocker, A. D., "Improved matched-filter detection techniques," in Imaging Spectrometry V, Denver, CO, USA, 1999, pp. 278-285.
- [18] Pai-Hui, H. and Yi-Hsing, T., "Feature extraction of hyperspectral data using the Best Wavelet Packet Basis," in Geoscience and Remote Sensing Symposium, 2002. IGARSS '02. 2002 IEEE International, 2002, pp. 1667-1669 vol.3.
- [19] Jazaeri, A., "Enhancing hyperspectral spatial resolution using multispectral image fusion: A wavelet approach," United States -- Virginia: George Mason University, 2007.
- [20] Bruce, L. M. and Li, J., "Wavelets for computationally efficient hyperspectral derivative analysis," Geoscience and Remote Sensing, IEEE Transactions on, vol. 39, pp. 1540-1546, 2001.

- [21] Kaewpijit, S., Le Moigne, J., and El-Ghazawi, T., "A wavelet-based PCA reduction for hyperspectral imagery," in Geoscience and Remote Sensing Symposium, 2002. IGARSS '02. 2002 IEEE International, 2002, pp. 2581-2583 vol.5.
- [22] Schmidt, F., Doute, S., and Schmitt, B., "WAVANGLET: An Efficient Supervised Classifier for Hyperspectral Images," Geoscience and Remote Sensing, IEEE Transactions on, vol. 45, pp. 1374-1385, 2007.
- [23] Younan, N. H., King, R. L., and Bennett, H. H., "Hyperspectral data analysis using wavelet-based classifiers," in Geoscience and Remote Sensing Symposium, 2000. Proceedings. IGARSS 2000. IEEE 2000 International, 2000, pp. 390-392 vol.1.
- [24] Gupta, M. R. and Jacobson, N. P., "Wavelet Principal Component Analysis and its Application to Hyperspectral Images," in Image Processing, 2006 IEEE International Conference on, 2006, pp. 1585-1588.
- [25] Webb, A. R., Statistical pattern recognition, 2nd ed. West Sussex, England ; New Jersey: Wiley, 2002.
- [26] Walnut, D. F., An introduction to wavelet analysis. Boston: Birkhäuser, 2002.
- [27] Bruce, L. M., Morgan, C., and Larsen, S., "Automated detection of subpixel hyperspectral targets with continuous and discrete wavelet transforms," Geoscience and Remote Sensing, IEEE Transactions on, vol. 39, pp. 2217-2226, 2001.
- [28] Jun, Z. and Regentova, E., "Wavelet based feature reduction method for effective classification of hyperspectral data," in Information Technology: Coding and Computing [Computers and Communications], 2003. Proceedings. ITCC 2003. International Conference on, 2003, pp. 483-487.
- [29] Mallat, S. G., A wavelet tour of signal processing, 2nd ed. San Diego: Academic Press, 1999.
- [30] Environmental Research Institute of Michigan. Infrared Information and Analysis (IRIA) Center., Wolfe, W. L., Zissis, G. J., and United States. Office of Naval Research., The infrared handbook. Washington: The Office, 1978.
- [31] Richards, J. A. and Jia, X., Remote sensing digital image analysis : an introduction, 3rd ed. Berlin ; New York: Springer, 1999.

- [32] Yung, Y. L. and DeMore, W. B., Photochemistry of planetary atmospheres. New York: Oxford University Press, 1999.
- [33] Schott, J. R., Remote sensing : the image chain approach. New York: Oxford University Press, 1997.
- [34] Elachi, C., Introduction to the physics and techniques of remote sensing. New York: Wiley, 1987.
- [35] Twomey, S., Introduction to the mathematics of inversion in remote sensing and indirect measurements. Mineola, N.Y.: Dover Publications, 1996.
- [36] Heath, M. T., Scientific computing : an introductory survey, 2nd ed. Boston: McGraw-Hill, 2002.
- [37] Anderson, G. P., Pukall, B., Allred, C. L., Jeong, L. S., Hoke, M., Chetwynd, J. H., Adler-Golden, S. M., Berk, A., Bernstein, L. S., Richtsmeier, S. C., Acharya, P. K., and Matthew, M. W., "FLAASH and MODTRAN4: state-of-the-art atmospheric correction for hyperspectral data," in Aerospace Conference, 1999. Proceedings. 1999 IEEE, 1999, pp. 177-181 vol.4.
- [38] Clough, S. A., Shephard, M. W., Mlawer, E. J., Delamere, J. S., Iacono, M. J., Cady-Pereira, K., Boukabara, S., and Brown, P. D., "Atmospheric radiative transfer modeling: a summary of the AER codes," vol. 91, pp. 233-244, 2005.
- [39] Resmini, R. G., "Categorization of hyperspectral information (HSI) based on the distribution of spectra in hyperspace," in Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery IX, Orlando, FL, USA, 2003, pp. 581-590.
- [40] Borel, C. C., Clodius, W. B., and Villeneuve, P. V., "Synthetic hyperspectral data cubes for complex thermal scenes," in Imaging Spectrometry IV, San Diego, CA, USA, 1998, pp. 222-231.
- [41] Young, S. J., "Detection and Quantification of Gases in Industrial-Stack Plumes Using Thermal-Infrared Hyperspectral Imaging," Aerospace Report, vol. ATR-2002(8407)-1, 10 February 2002 2002.
- [42] Chang, C.-I., Hyperspectral imaging : techniques for spectral detection and classification. New York: Kluwer Academic/Plenum Publishers, 2003.
- [43] Landgrebe, D. A., Signal theory methods in multispectral remote sensing. Hoboken, N.J.: Wiley, 2003.

- [44] Sweet, J. N., "The spectral similarity scale and its application to the classification of hyperspectral remote sensing data," United States -- New York: State University of New York College of Environmental Science and Forestry, 2003.
- [45] Stuart, B. and Ando, D. J., Modern infrared spectroscopy. New York: Published on behalf of ACOL (University of Greenwich) by Wiley, 1996.
- [46] Smith, B. C., Quantitative spectroscopy : theory and practice. Amsterdam ; Boston: Academic Press, 2002.
- [47] Marion, R., Michel, R., and Faye, C., "Measuring trace gases in plumes from hyperspectral remotely sensed data," Geoscience and Remote Sensing, IEEE Transactions on, vol. 42, pp. 854-864, 2004.
- [48] Schaum, A. P., "Spectral subspace matched filtering," in Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VII, Orlando, FL, USA, 2001, pp. 1-17.
- [49] Arfken, G. B. and Weber, H.-J., Mathematical methods for physicists, 4th ed. San Diego: Academic Press, 1995.
- [50] Weisstein, E. W., "Moore-Penrose Matrix Inverse," From Mathworld --A Wolfram Web Resource, pp. <http://mathworld.wolfram.com/Moore-PenroseMatrixInverse.html>, 2002.
- [51] Sharpe, S. W., Sams, R. L., and Johnson, T. J., "The PNNL quantitative IR database for infrared remote sensing and hyperspectral imaging," in Applied Imagery Pattern Recognition Workshop, 2002. Proceedings. 31st, 2002, pp. 45-48.
- [52] Symonds, R. B. and Geological Survey (U.S.), "Compositions, origins, emission rates and atmospheric impacts of volcanic gases," Vancouver, WA: U.S. Dept. of the Interior U.S. Geological Survey, 1998.
- [53] EPA, "Six Common Air Pollutants," <http://www.epa.gov/oar/urbanair/so2/what1.html>.
- [54] Smith, B. C., Infrared spectral interpretation : a systematic approach. Boca Raton: CRC Press, 1999.
- [55] EPA, "EPA Climate Change - Methane," <http://www.epa.gov/methane/sources.html>.

- [56] EPA, "Technology Transfer Network Air Toxics Website - Benzene, <http://www.epa.gov/ttn/atw/hlthef/benzene.html>."
- [57] Lucey, P. G., Horton, K. A., Williams, T., Mignard, M., Julian, J., Kokubun, D., Allen, G., Winter, E. M., Schlangen, M., Kendall, W., Schaff, W., Stocker, A., Bowman, A. P., Fields, D. J., and Sayre, C. J., "An Airborne Hyperspectral Imager For Hyperspectral Mine Detection," <http://www.higp.hawaii.edu/ahi/>, 2001.
- [58] Winter, E. M., "The application of hyperspectral sensors to the detection of land mines," in Aerospace Conference, 2005 IEEE, 2005, pp. 2030-2034.
- [59] Mares, A. G., Olsen, R. C., and Lucey, P. G., "LWIR spectral measurements of volcanic sulfur dioxide plumes," in Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery X, Orlando, FL, USA, 2004, pp. 266-272.
- [60] Williams, D. J., Feldman, B. L., Williams, T. J., Pilant, D., Lucey, P. G., and Worthy, L. D., "Detection and identification of toxic air pollutants using airborne LWIR hyperspectral imaging," in Multispectral and Hyperspectral Remote Sensing Instruments and Applications II, Honolulu, HI, USA, 2005, pp. 134-141.
- [61] Strow, L. L., Hannon, S. E., De Souza-Machado, S., Motteler, H. E., and Tobin, D., "An overview of the AIRS radiative transfer model," Geoscience and Remote Sensing, IEEE Transactions on, vol. 41, pp. 303-313, 2003.
- [62] Lambrigtsen, B. H., "Calibration of the AIRS microwave instruments," Geoscience and Remote Sensing, IEEE Transactions on, vol. 41, pp. 369-378, 2003.
- [63] Strow, L. L., Hannon, S. E., Weiler, M., Overoye, K., Gaiser, S. L., and Aumann, H. H., "Prelaunch spectral calibration of the atmospheric infrared sounder (AIRS)," Geoscience and Remote Sensing, IEEE Transactions on, vol. 41, pp. 274-286, 2003.
- [64] Aumann, H. H., Chahine, M. T., Gautier, C., Goldberg, M. D., Kalnay, E., McMillin, L. M., Revercomb, H., Rosenkranz, P. W., Smith, W. L., Staelin, D. H., Strow, L. L., and Susskind, J., "AIRS/AMSU/HSB on the Aqua mission: design, science objectives, data products, and processing systems AIRS, the first hyper-spectral infrared sounder for operational weather forecasting," Geoscience and Remote Sensing, IEEE Transactions on, vol. 41, pp. 253-264, 2003.
- [65] Pagano, T. S., Aumann, H. H., Hagan, D. E., and Overoye, K., "Prelaunch and in-flight radiometric calibration of the Atmospheric Infrared Sounder (AIRS),"

- Geoscience and Remote Sensing, IEEE Transactions on, vol. 41, pp. 265-273, 2003.
- [66] Heli, W., Ping, Y., Jun, L., Baum, B. A., Hung-Lung, H., Platnick, S., Yongxiang, H., and Strow, L., "Retrieval of semitransparent ice cloud optical thickness from atmospheric infrared sounder (AIRS) measurements," Geoscience and Remote Sensing, IEEE Transactions on, vol. 42, pp. 2254-2267, 2004.
 - [67] Hagan, D. E. and Minnett, P. J., "AIRS radiance validation over ocean from sea surface temperature measurements," Geoscience and Remote Sensing, IEEE Transactions on, vol. 41, pp. 432-441, 2003.
 - [68] Goldberg, M. D., Qu, Y., McMillin, L. M., Wolf, W., Lihang, Z., and Divakarla, M., "AIRS near-real-time products and algorithms in support of operational numerical weather prediction," Geoscience and Remote Sensing, IEEE Transactions on, vol. 41, pp. 379-389, 2003.
 - [69] Toptygin, A. Y., Griбанov, K. G., Imasu, R., Bleuten, W., and Zakharov, V. I., "Seasonal methane content in atmosphere of the permafrost boundary zone in Western Siberia determined from IMG/ADEOS and AIRS/AQUA data," in Multispectral and Hyperspectral Remote Sensing Instruments and Applications II, Honolulu, HI, USA, 2005, pp. 508-514.
 - [70] Susskind, J., Barnet, C. D., and Blaisdell, J. M., "Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds," Geoscience and Remote Sensing, IEEE Transactions on, vol. 41, pp. 390-409, 2003.
 - [71] Strang, G. and Nguyen, T., Wavelets and filter banks. Wellesley, MA: Wellesley-Cambridge Press, 1996.
 - [72] Gorman, J. D., Subotic, N. S., and Thelen, B. J., "Robust material identification in hyperspectral data via multiresolution wavelet techniques," in Acoustics, Speech, and Signal Processing, 1995. ICASSP-95., 1995 International Conference on, 1995, pp. 2805-2808 vol.5.
 - [73] Bruce, L. M., Jiang, L., and Yan, H., "Automated detection of subpixel hyperspectral targets with adaptive multichannel discrete wavelet transform," Geoscience and Remote Sensing, IEEE Transactions on, vol. 40, pp. 977-980, 2002.
 - [74] Jiang, L., Mann Bruce, L., and Mathur, A., "Wavelet transform for dimensionality reduction in hyperspectral linear unmixing," in Geoscience and Remote Sensing

- Symposium, 2002. IGARSS '02. 2002 IEEE International, 2002, pp. 3513-3515 vol.6.
- [75] Bruce, L. M., Koger, C. H., and Jiang, L., "Dimensionality reduction of hyperspectral data using discrete wavelet transform feature extraction," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 40, pp. 2331-2338, 2002.
 - [76] Kaewpijit, S., Le Moigne, J., and El-Ghazawi, T., "Automatic reduction of hyperspectral imagery using wavelet spectral analysis," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 41, pp. 863-871, 2003.
 - [77] Lennon, M., Mercier, G., Mouchot, M. C., and Hubert-Moy, L., "Spectral unmixing of hyperspectral images with the independent component analysis and wavelet packets," in *Geoscience and Remote Sensing Symposium, 2001. IGARSS '01. IEEE 2001 International*, 2001, pp. 2896-2898 vol.6.
 - [78] Agarwal, A., LeMoigne, J., Joiner, J., El-Ghazawi, T., and Cantonnet, F., "Wavelet dimension reduction of AIRS infrared (IR) hyperspectral data," in *Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004 IEEE International*, 2004, pp. 3249-3252 vol.5.
 - [79] Arnold, P. S., Brown, S. D., and Schott, J. R., "Hyperspectral simulation of chemical weapon dispersal patterns using DIRSIG," in *Targets and Backgrounds VI: Characterization, Visualization, and the Detection Process*, Orlando, FL, USA, 2000, pp. 288-299.
 - [80] Jakubowski, M. K., Pogorzala, D., Hattenberger, T. J., Brown, S. D., and Schott, J. R., "Synthetic data generation of high-resolution hyperspectral data using DIRSIG," in *Imaging Spectrometry XII*, San Diego, CA, USA, 2007, pp. 66610G-11.
 - [81] Carn, S. A., A. J. Krueger, N. A. Krotkov, M. A. Gray "Fire at Iraqi sulfur plant emits SO₂ clouds " *Geophys. Res. Lett.*, vol. 31, 1029/2004 2004.
 - [82] Teolis, A., *Computational signal processing with wavelets*. Boston: Birkhäuser, 1998.
 - [83] Krueger, A. J., L. S. Walter, P. K. Bhartia, C. C. Schnetzler, N. A. Krotkov, I. Sprod, G. J. S. Bluth "Volcanic sulfur dioxide measurements from the total ozone mapping spectrometer instruments," *J. Geophys. Res.*, vol. 100(D7), pp. 14,057–14,076, 1995.

- [84] Carn, S. A., L. L. Strow, S. de Souza-Machado, Y. Edmonds, S. Hannon,
"Quantifying tropospheric volcanic emissions with AIRS: The 2002 eruption of
Mt. Etna (Italy)," *Geophys. Res. Lett.*, vol. 32, 1029/2004 2005.

CURRICULUM VITAE

Mark Anthony Zarate Salvador was born on June 10, 1968 in New York City, New York and is an American citizen. He graduated from Iona Preparatory School in New Rochelle, New York, in 1986. He received the Bachelor of Science in Aerospace Engineering from Syracuse University in 1990 and the Master of Science in Applied and Engineering Physics from George Mason University in 2001. He is employed by Logos Technologies as a Senior Scientist.