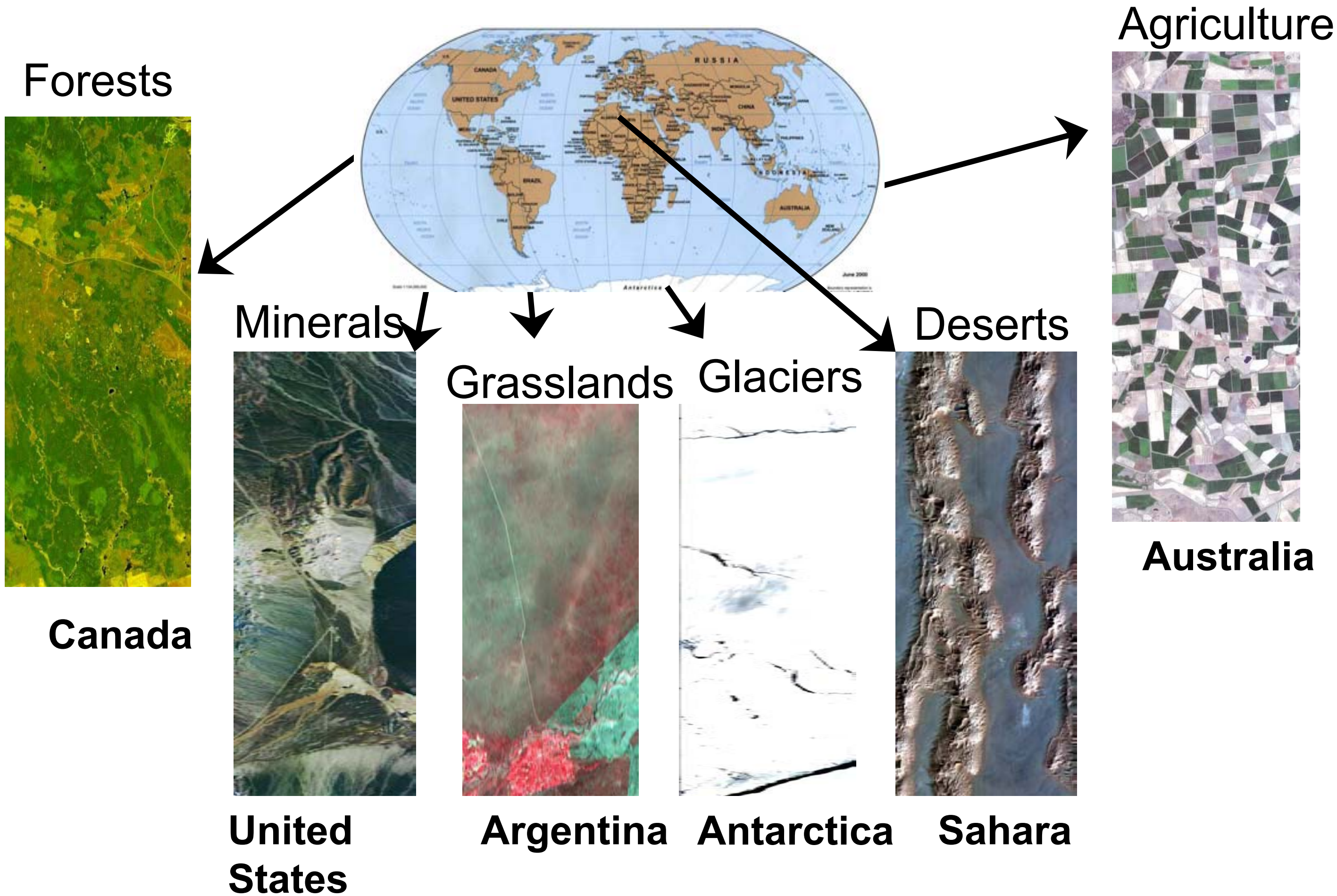


Hyperspectral Image Processing and Analysis for Land Cover Characterization

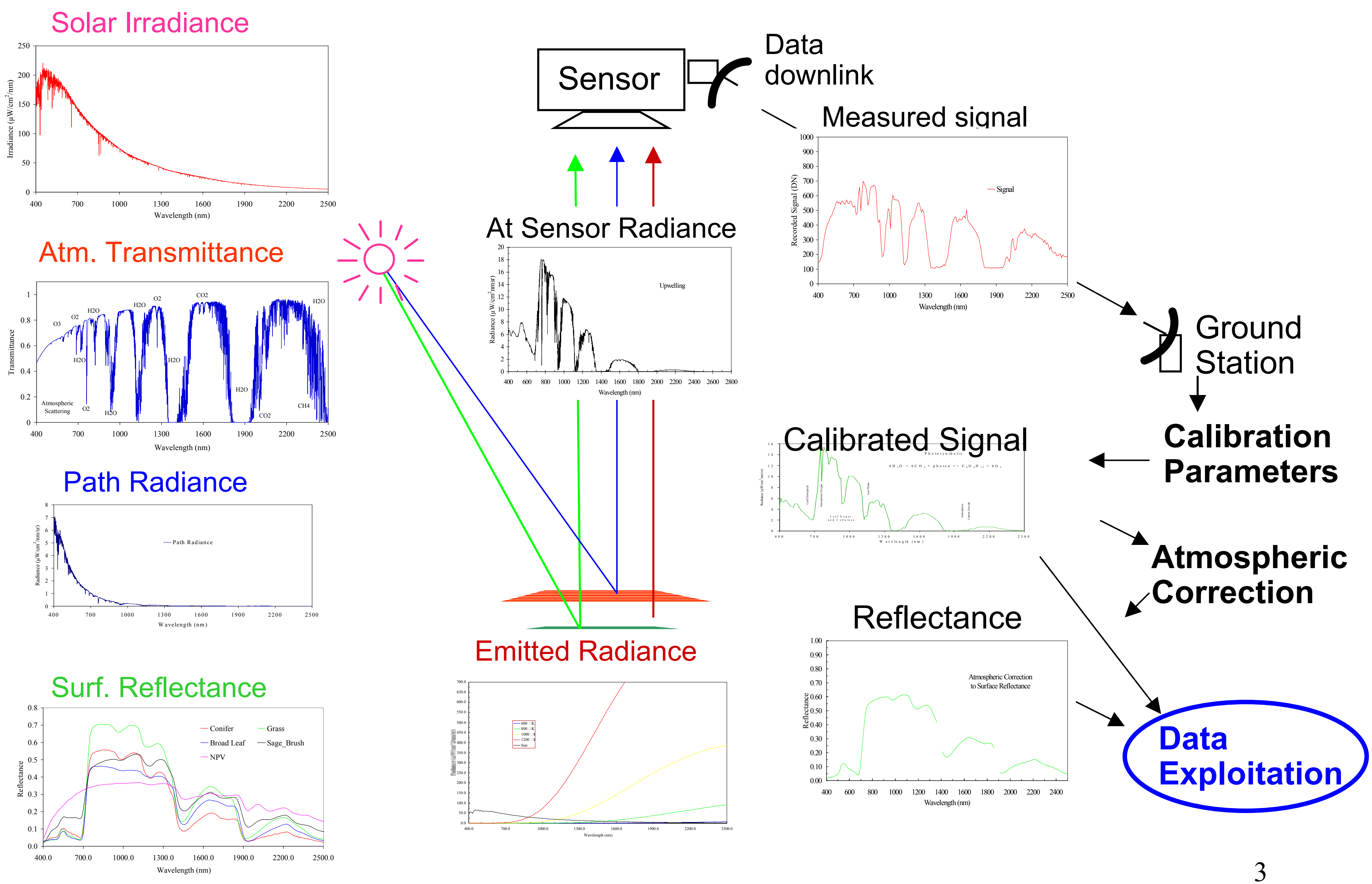
Dr. Jay Pearlman and Dr. Melba Crawford

September 19, 2004

Course Overview

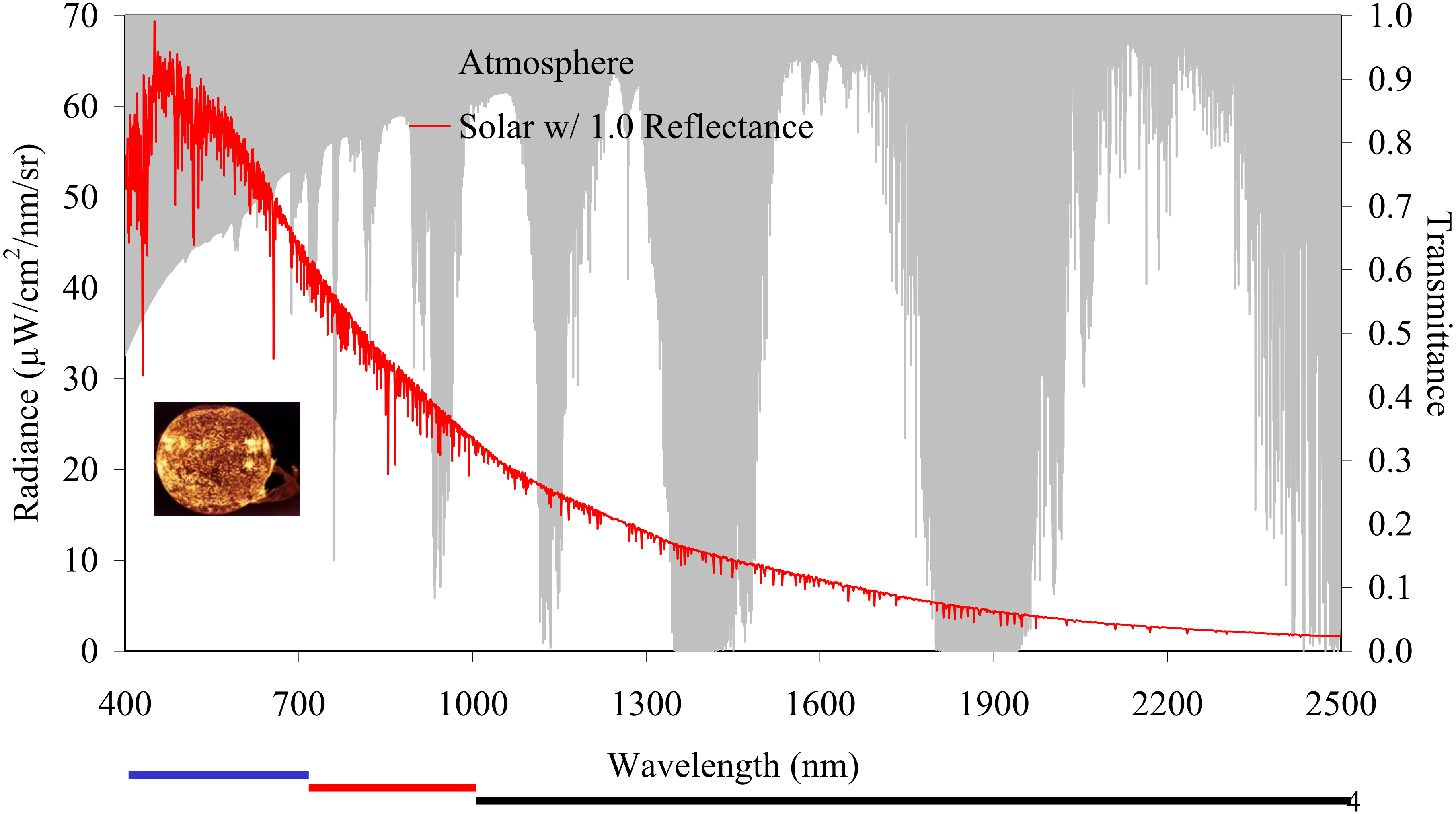
- Introduction to EO Remote Sensing
  - Photon end to end flow
  - Alternative sensor configurations
- Hyperion Instrument – design and trades
- Sensor Calibration
- Data Processing and Corrections
- Classification of Hyperspectral Data
  - Input space reduction
    - Feature selection and extraction
  - Output space reduction
    - Multiclassifier Systems
- A View Forward

EO Remote Sensing Process Overview

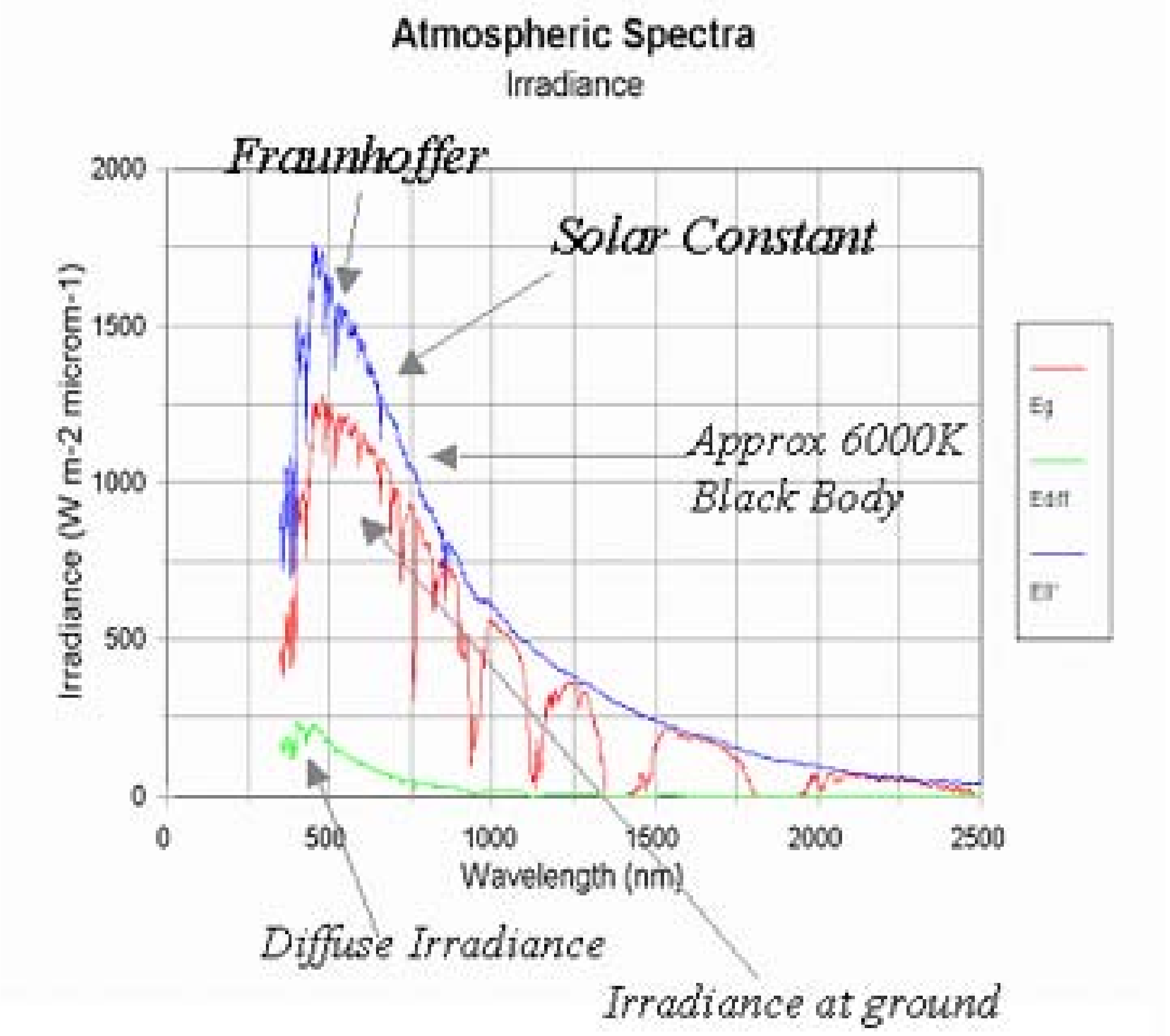


Courtesy of R. Green

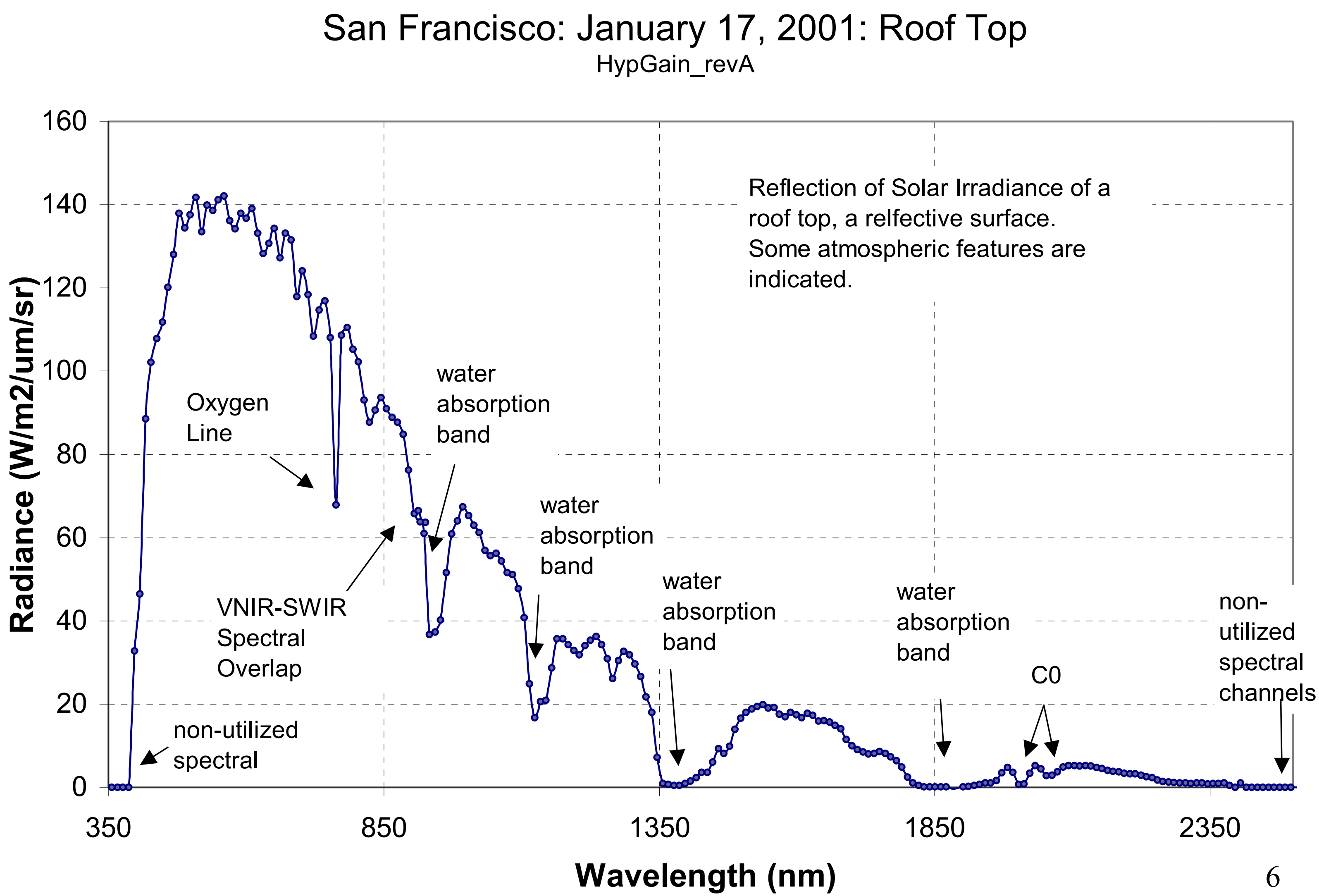
Solar Spectrum and Atmospheric Transmission



Solar Irradiance



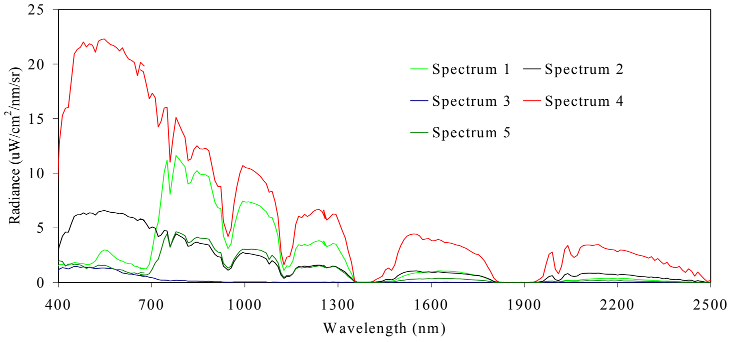
Reflection from a Metal Roof

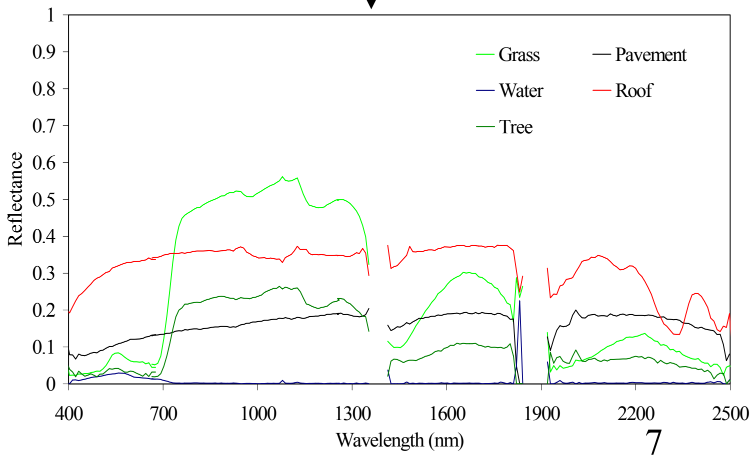


Radiance To Reflectance Inversion

$$Lt = \mu F_0 \rho_a / \pi + \mu F_0 T_d \rho_s T_u / \pi$$

Lt is the at sensor radiance
μ is the cosine of the solar zenith angle
F_0 is the exo atmospheric irradiance
ρ_a is the upward reflectance of the atmosphere
T_d is the downward transmittance
ρ_s is the reflectance of the surface
T_u is the upward transmittance of the atmosphere



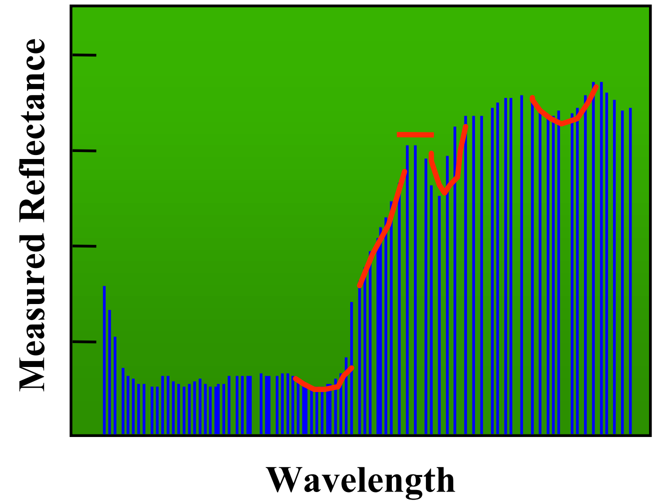


Alternative Sensor Configurations

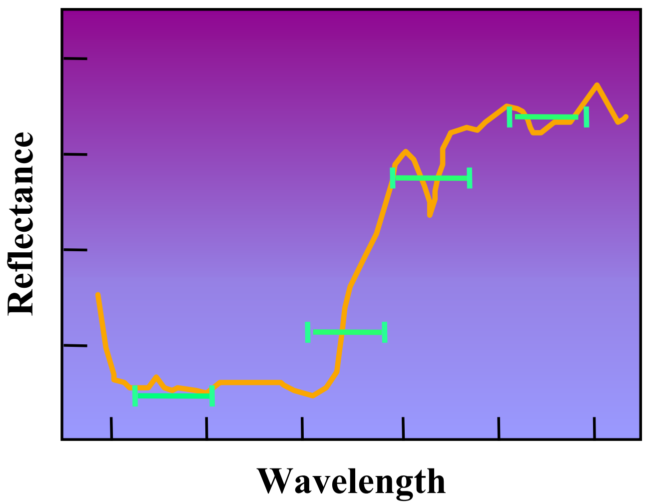
- Multispectral vs. Hyperspectral
- Whiskbroom vs. Pushbroom

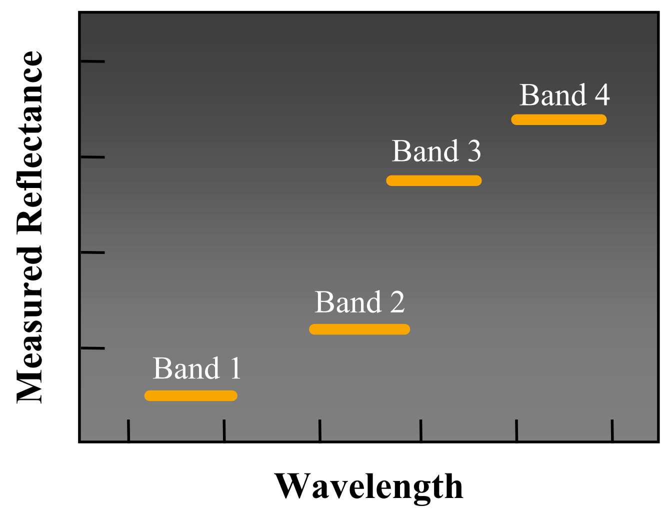
Hyperspectral and Multispectral Perspectives

Hyperspectral Imaging
Hundreds of bands



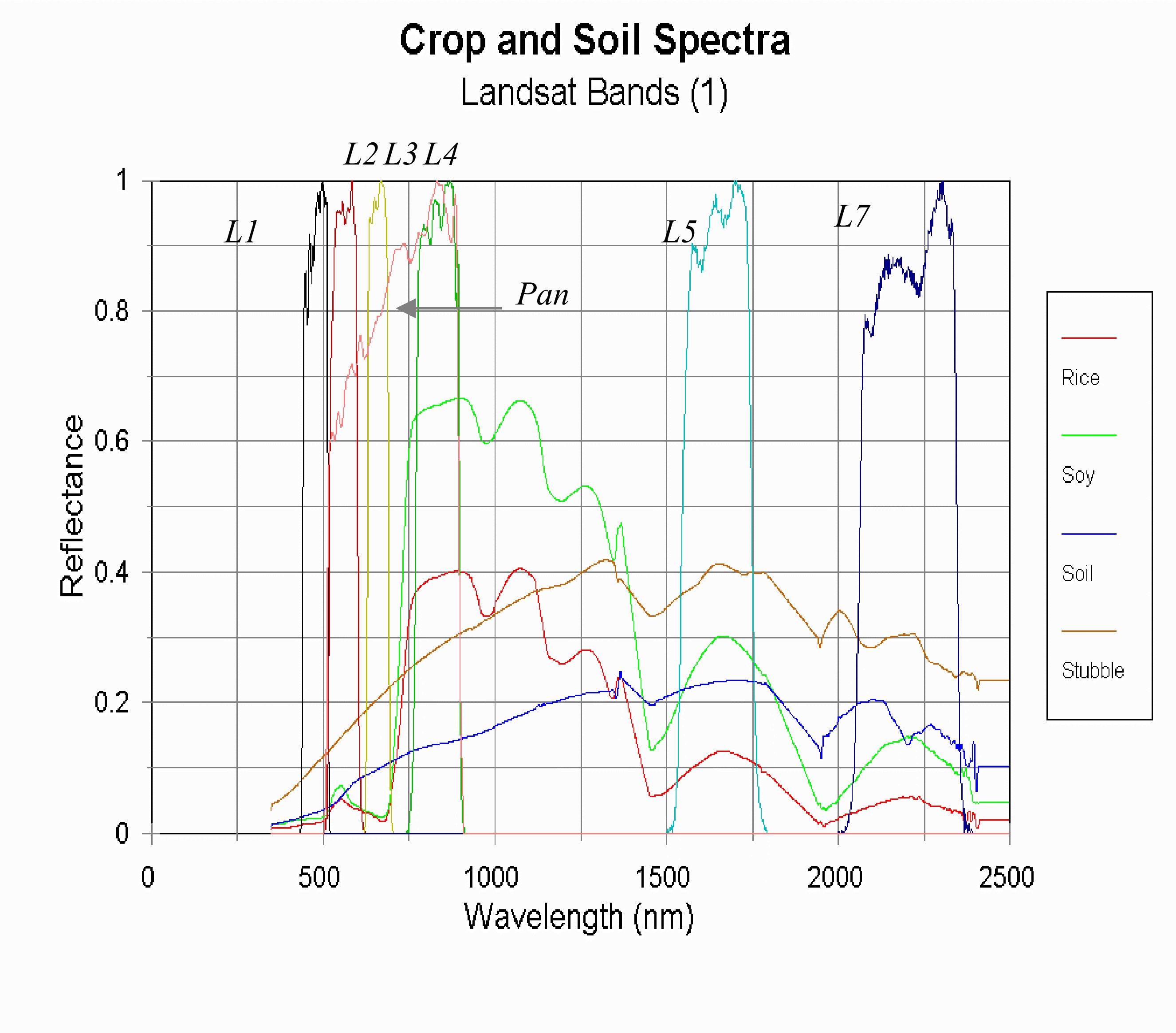
Spectral characteristic of scene





Multispectral Imaging
Few bands

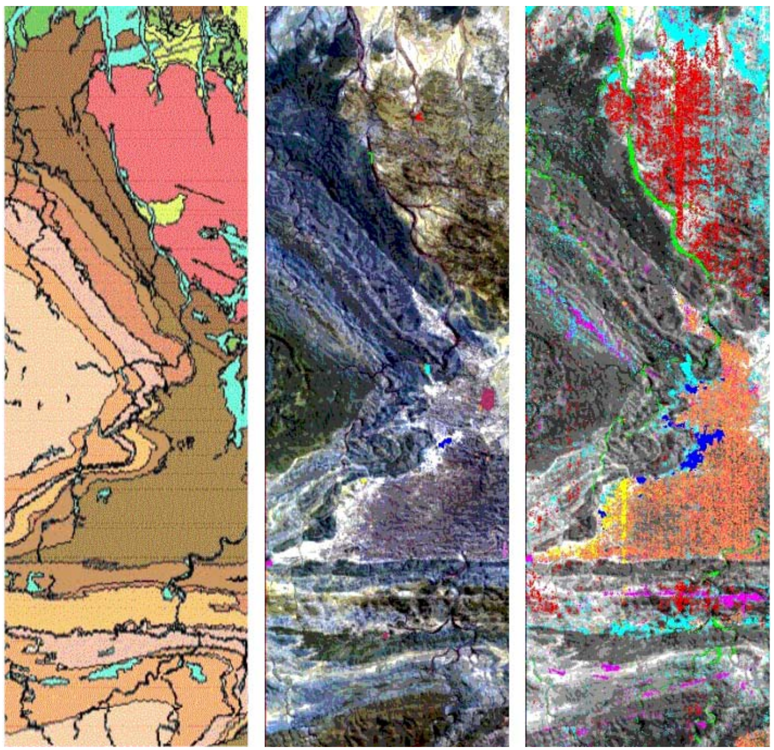
Landsat Picks the Windows – But Misses a Lot



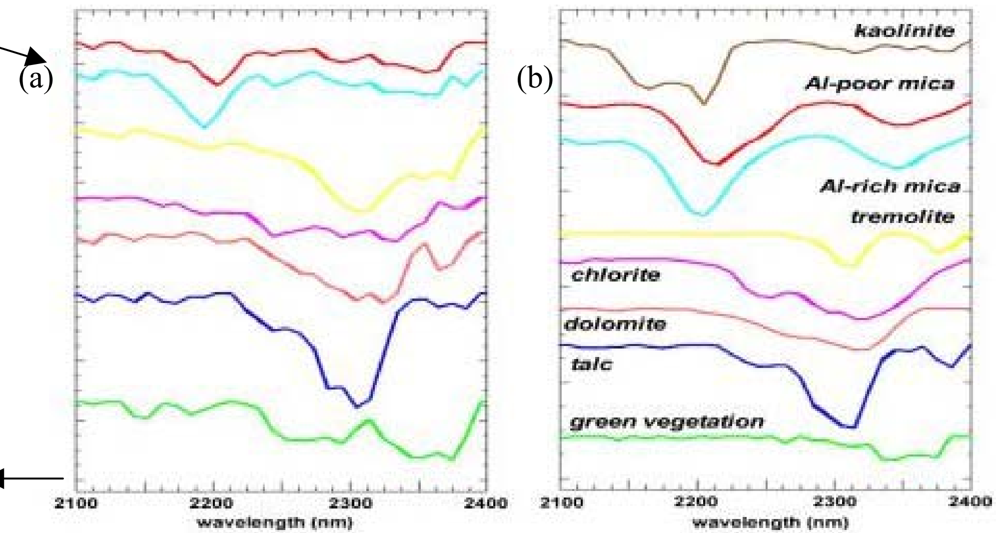
Hyperion Maps Mt. Fitton Geology

Hyperion-based apparent reflectance compares with library reference spectra

(1) (2) (3)



Hyperion Spectra **Reference Spectra**



Hyperion surface composition map agrees with known geology of Mt. Fitton in South Australia

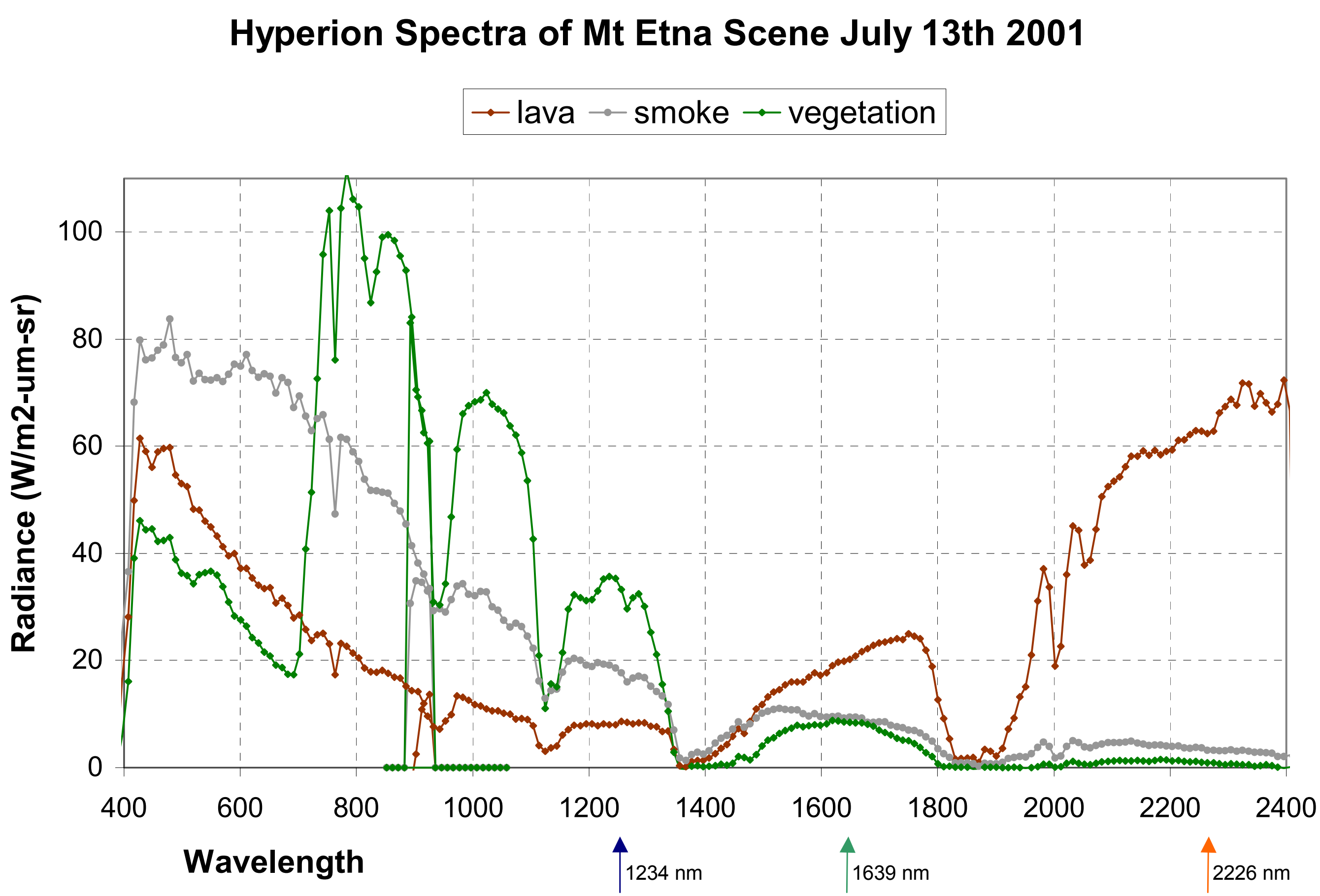
(1) Published Geologic Survey Map

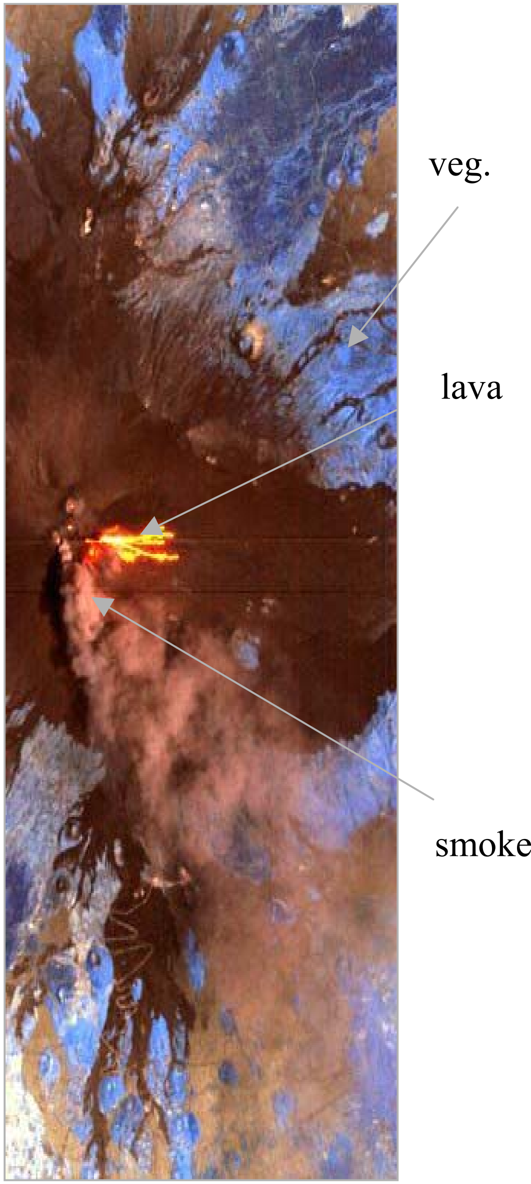
(2) Hyperion three color image (visible) showing regions of interest

(3) Hyperion surface composition map using SWIR spectra above

Courtesy of CSIRO, Australia

Mt. Etna Sample Spectra





Hyperspectral Sensor Configurations

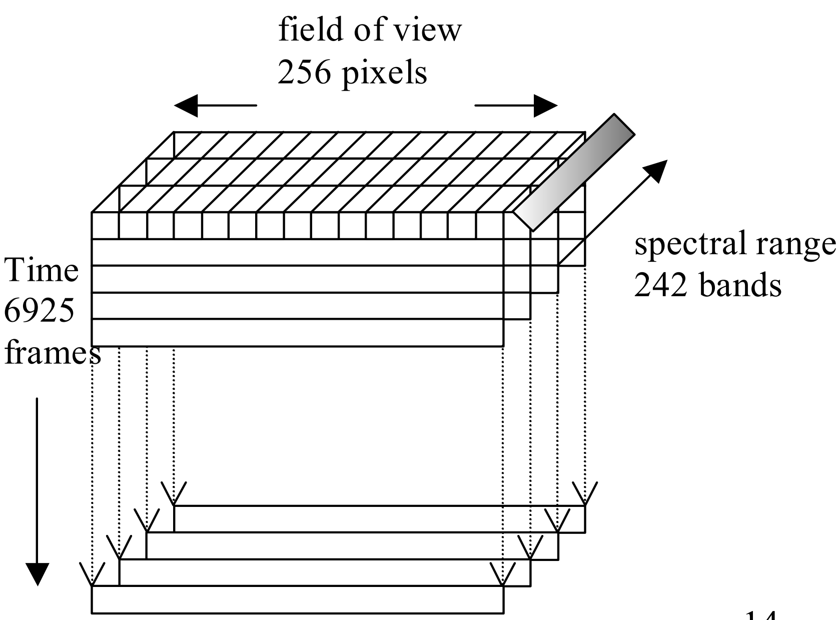
Characteristic	Whiskbroom	Pushbroom
Detectors	Linear array	2-dimensional array
Relative pixel dwell time	Shorter	**Longer**
Spatial uniformity	**Easier**	Harder
Calibration	**Easier**	Harder
Application domain	Airborne	**Space and Airborne**
Bandwidth	Similar	Similar

Hyperion Configuration

- Pushbroom configuration with common fore optics
  - 256 field-of-view locations, defines 7.7 km swath width, 30 meters each
  - 6925 frames of data, defines 185 km swath length, 30 meters each. (sample based on length of data), frame rate timed with spacecraft velocity

Hyperion Data Cube

X-swath width
Y-swath length
Z-spectra signature



The Hyperion Sensor

Hyperion Imaging Spectrometer

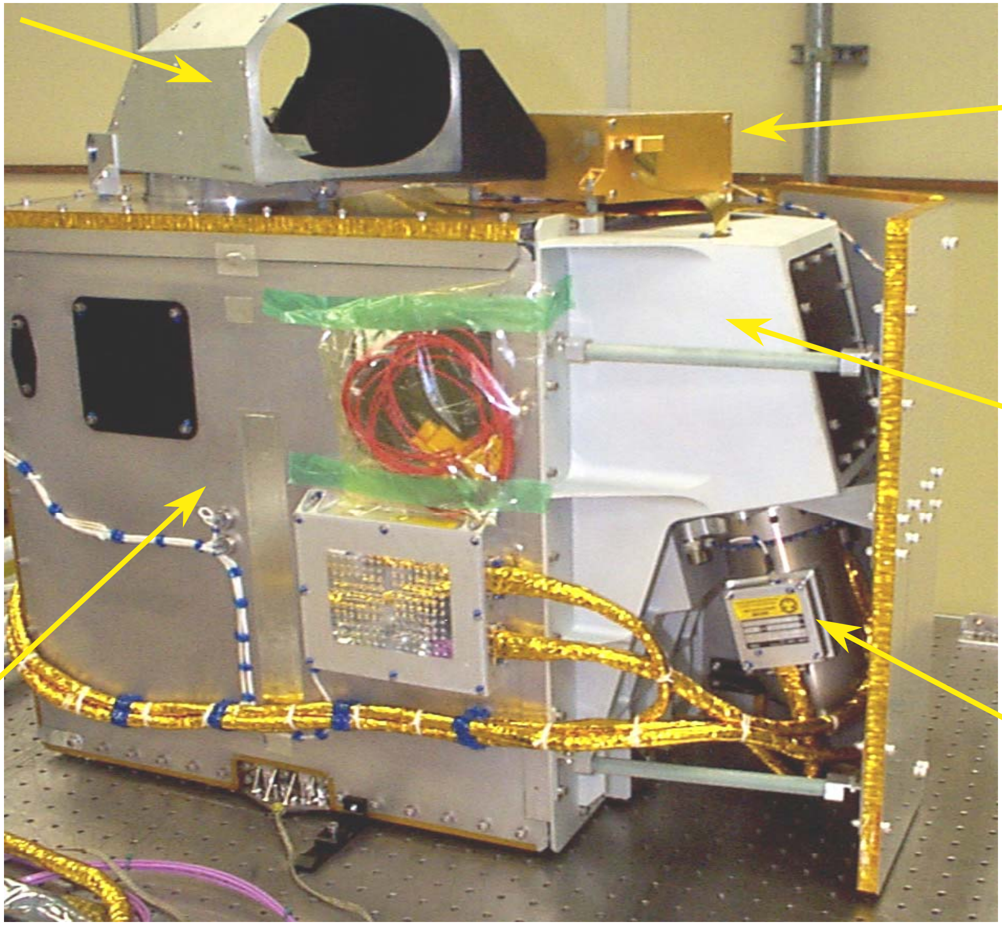
Hyperion is a push-broom imager

- 220 10nm bands covering 400nm - 2500nm
- 6% absolute rad. accuracy
- Swath width of 7.5 km
- IFOV of 42.4 μradian
- GSD of 30 m
- 12-bit image data
- Orbit is 705km alt (16 day repeat)

Hyperion Technologies

CALIBRATION
(spectral/
pushbroom)

DATA
RATES



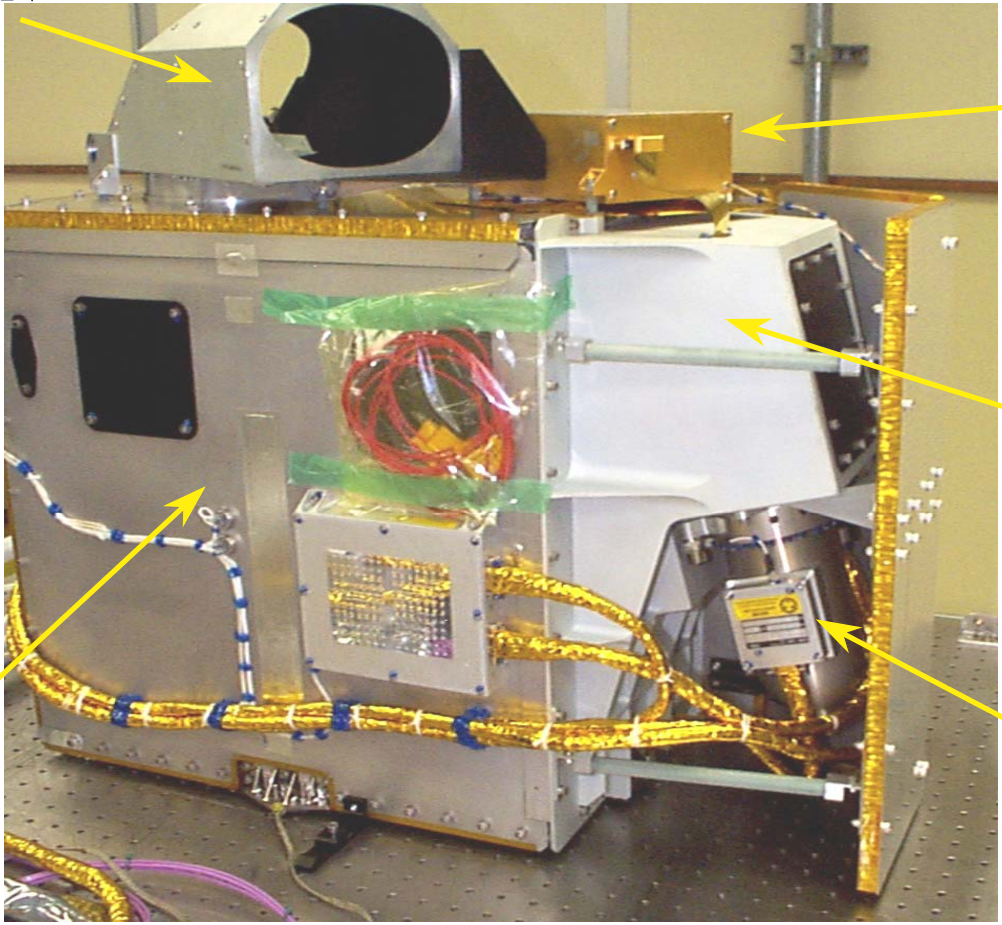
SPECTROMETER
(curved grating)

Reflecting
TELESCOPE

Pulse Tube
CRYOCOOLER

Hyperion Sensor Assembly

Baffle / CALIBRATION



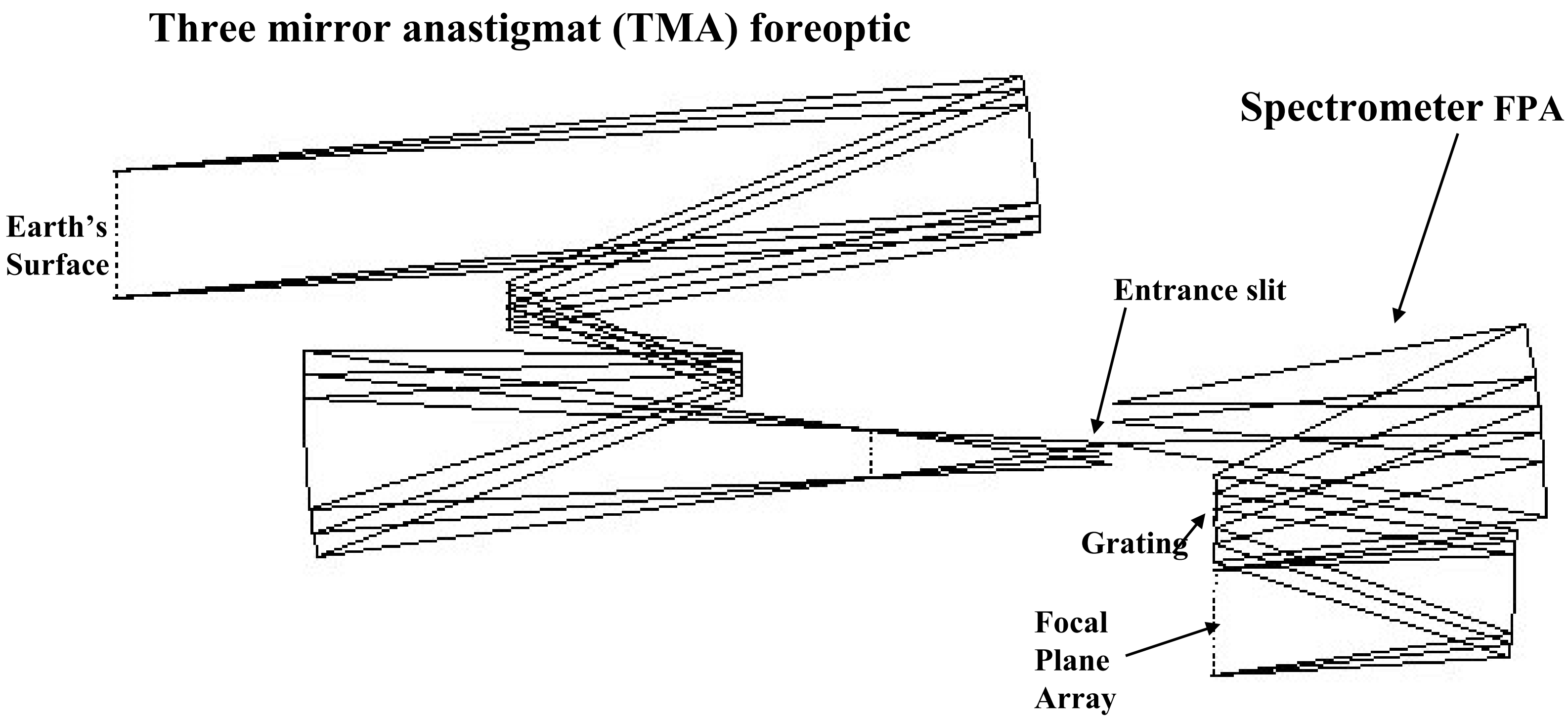
Signal processor

Spectro-meter

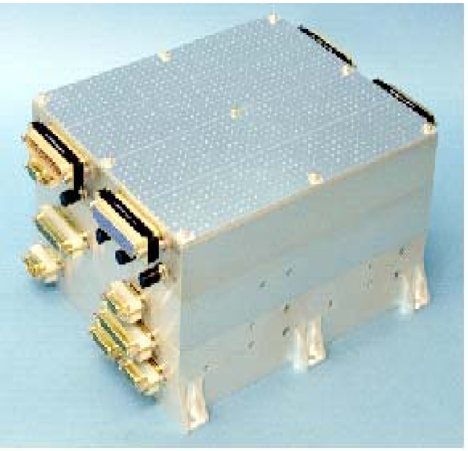
Reflecting TELESCOPE

Pulse Tube CRYOCOOLER

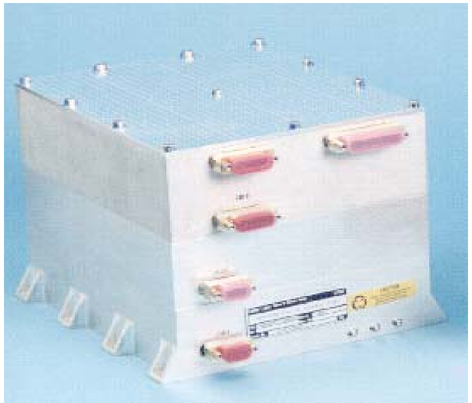
Hyperion Optical System



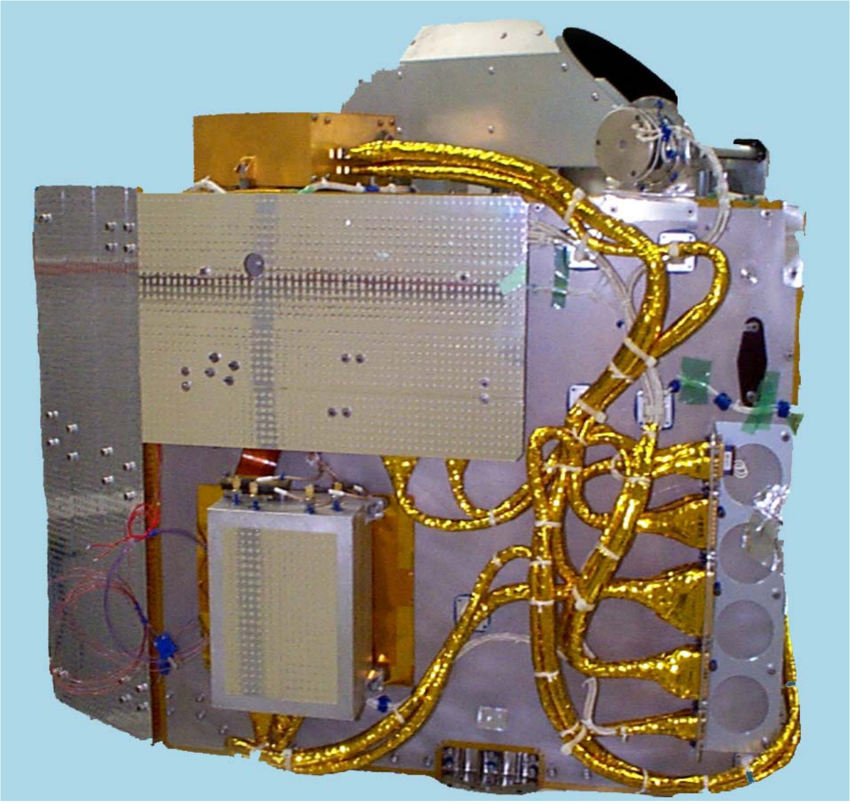
Hyperion Subassemblies



Hyperion Electronics Assembly (HEA)

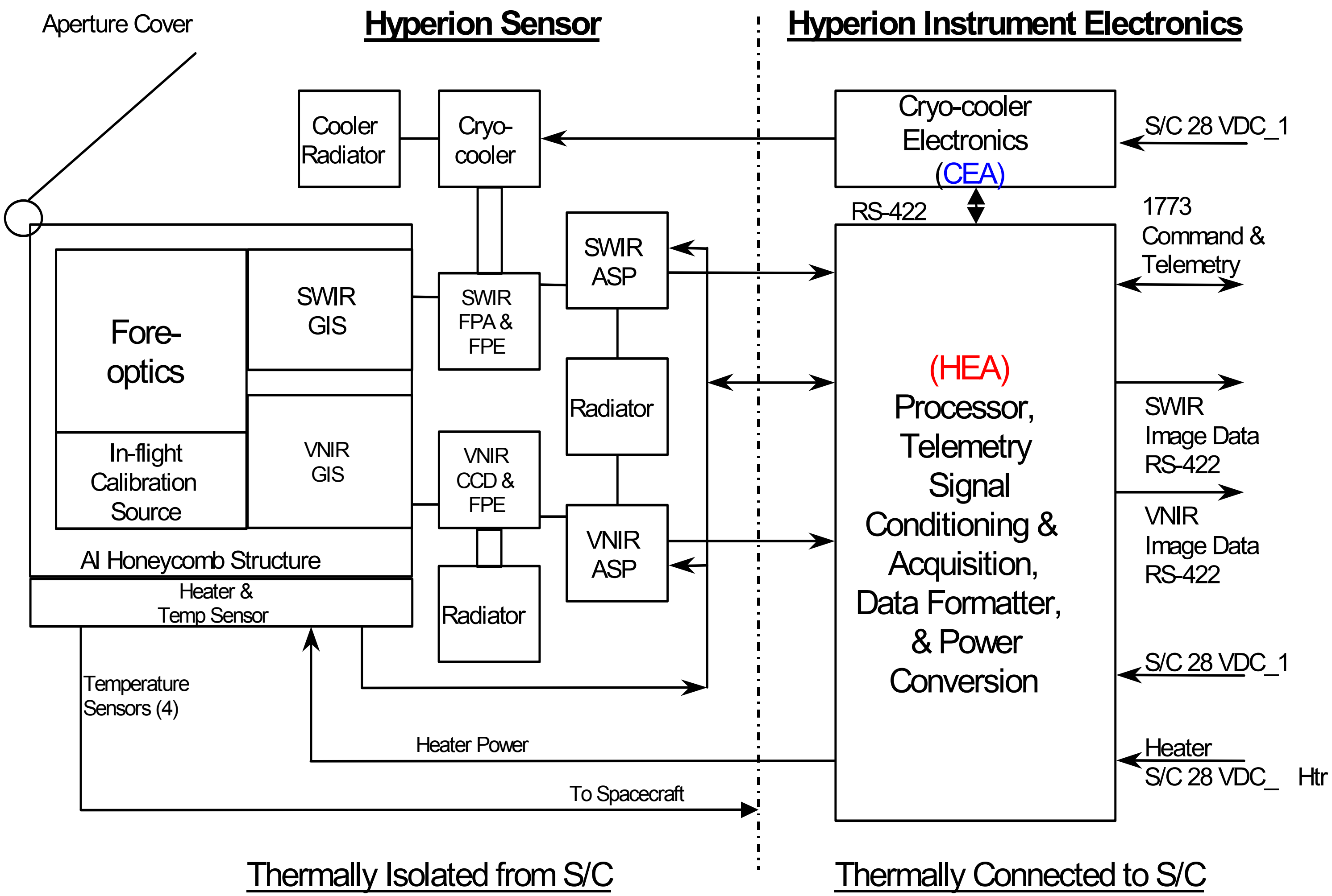


Cryocooler Electronics Assembly (CEA)



Hyperion Sensor Assembly (HSA)

Hyperion Block Diagram

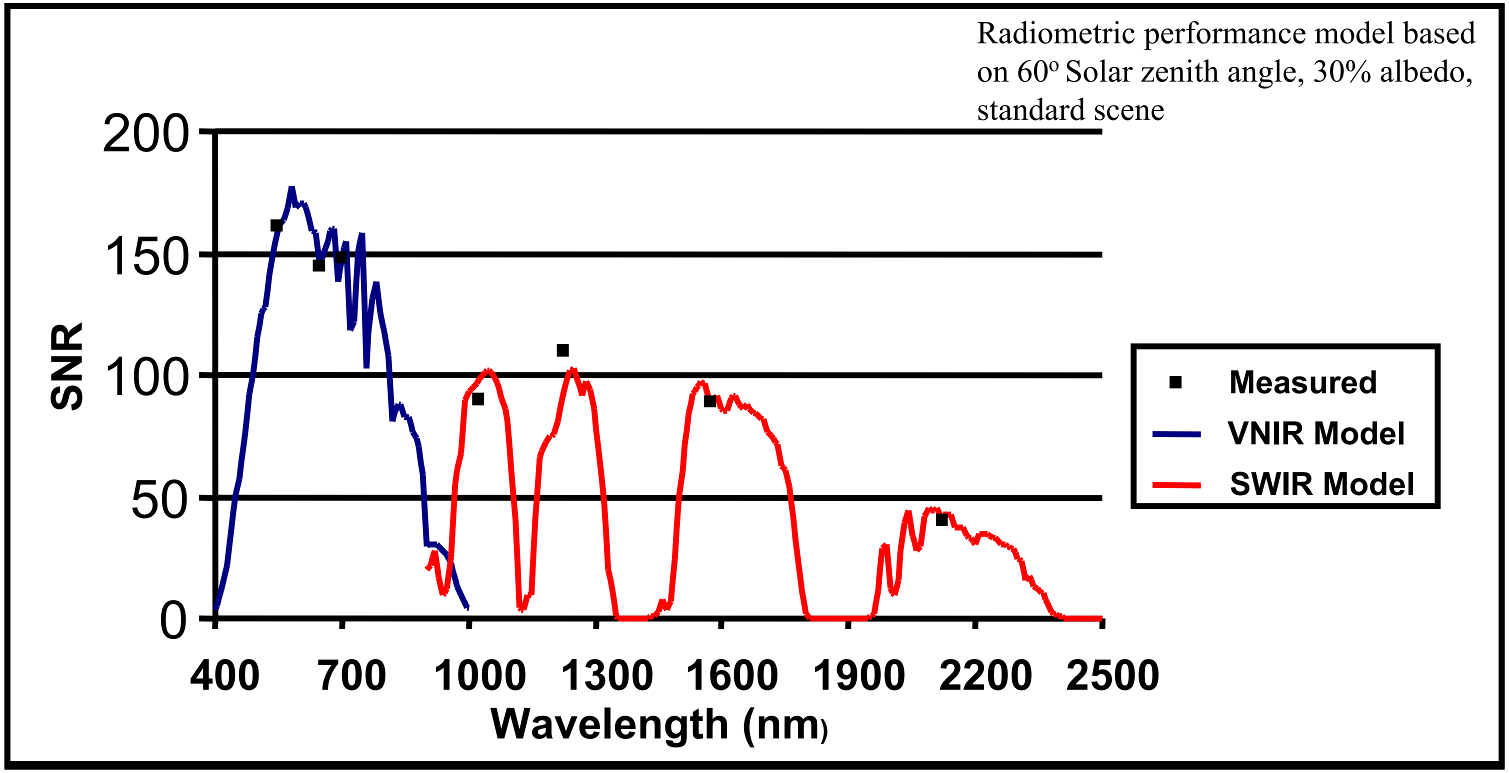


Hyperion Key Characteristics

	Characteristic	Pre-launch	On-orbit
OPTICAL	GSD (m)	29.88	30.38
	Swath (km)	7.5	7.75
	VNIR MTF @ 630nm	0.22-0.28	0.23-0.27
	SWIR MTF @ 1650nm	0.25-0.27	0.28
	VNIR X-trk Spec. Error	2.8nm@655nm	2.2nm
	SWIR X-trk Spec. Error	0.6nm@1700nm	0.58
RADIO-METRIC	Abs. Radiometry(1Sigma)	<6%	3.40%
	VNIR SNR (550-700nm)	144-161	140-190
	SWIR SNR (~1225nm)	110	96
	SWIR SNR (~2125nm)	40	38
SPECTRAL	No. of Spectral Channels	220	200 (L1)
	VNIR Bandwidth (nm)	10.19-10.21	**
	SWIR Bandwidth (nm)	10.08-10.09	**

** not measured

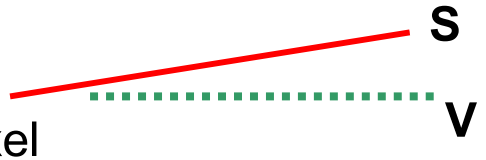
Hyperion SNR



Hyperion Measured SNR						
550 nm	650 nm	700 nm	1025 nm	1225 nm	1575 nm	2125 nm
161	144	147	90	110	89	40

Spectrometer Overlays

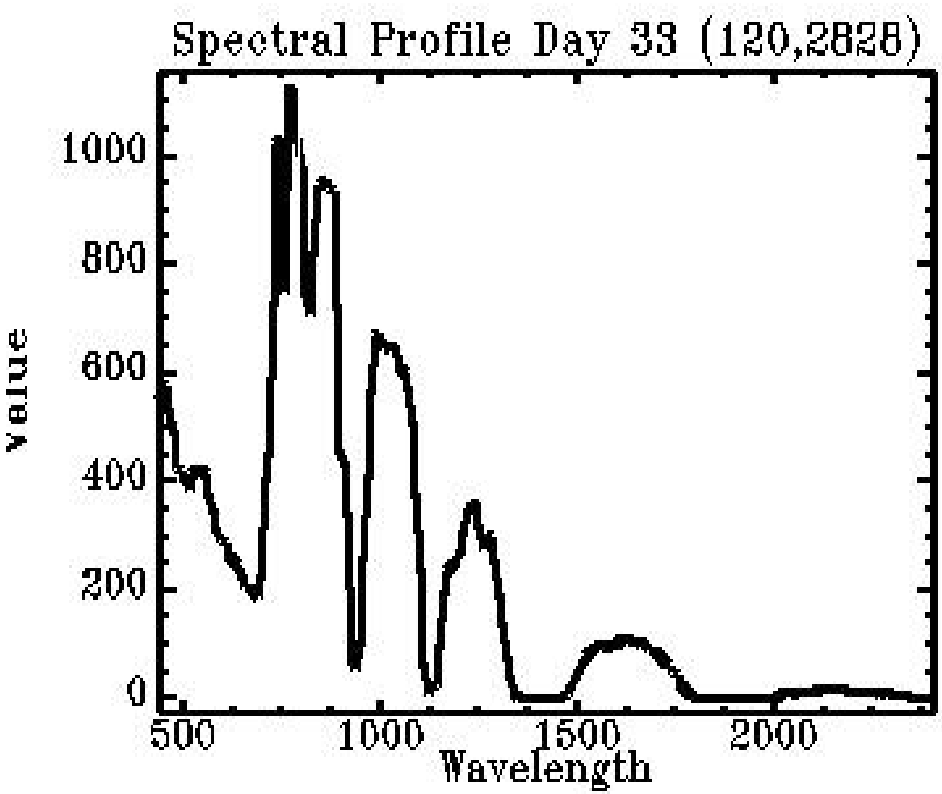
Level 1 data: 438-926nm and 892-2406nm
Bands 9-57 and 75 - 225;
SWIR is West of VNIR and rotated CCW by one pixel



Band	Center(nm)	FWHM(nm)
50	**854.66**	**11.27**
51	**864.83**	**11.27**
52	**875**	**11.28**
53	**885.17**	**11.29**
54	**895.34**	**11.3**
55	**905.51**	**11.31**
56	**915.68**	**11.31**
57	**925.85**	**11.31**
58	936.02	11.31
59	946.19	11.31
71	852	11.17
72	862.09	11.17
73	872.18	11.17
74	882.27	11.17
75	**892.35**	**11.17**
76	**902.44**	**11.17**
77	**912.53**	**11.17**
78	**922.62**	**11.17**
79	**932.72**	**11.17**
80	**942.81**	**11.17**

VNIR

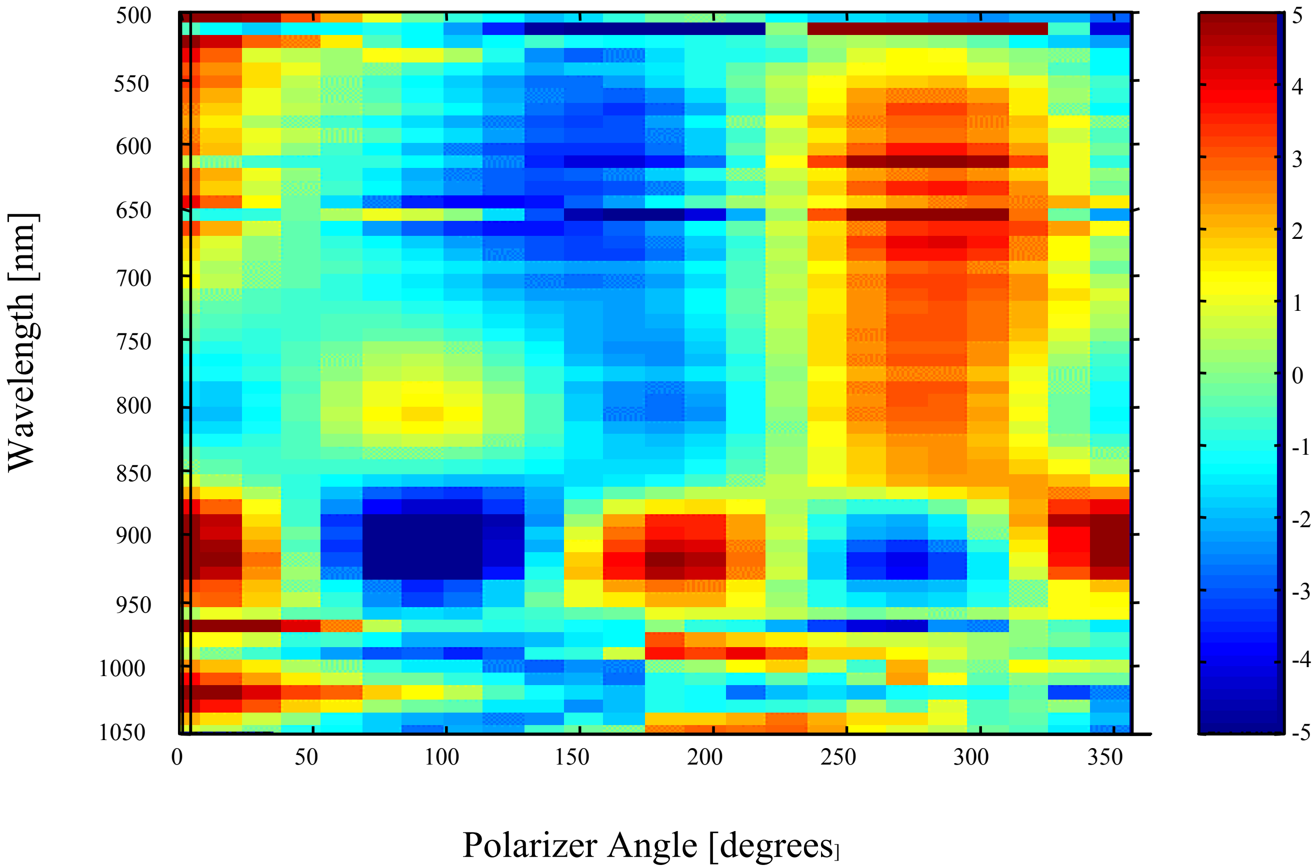
SWIR



Polarization - VNIR

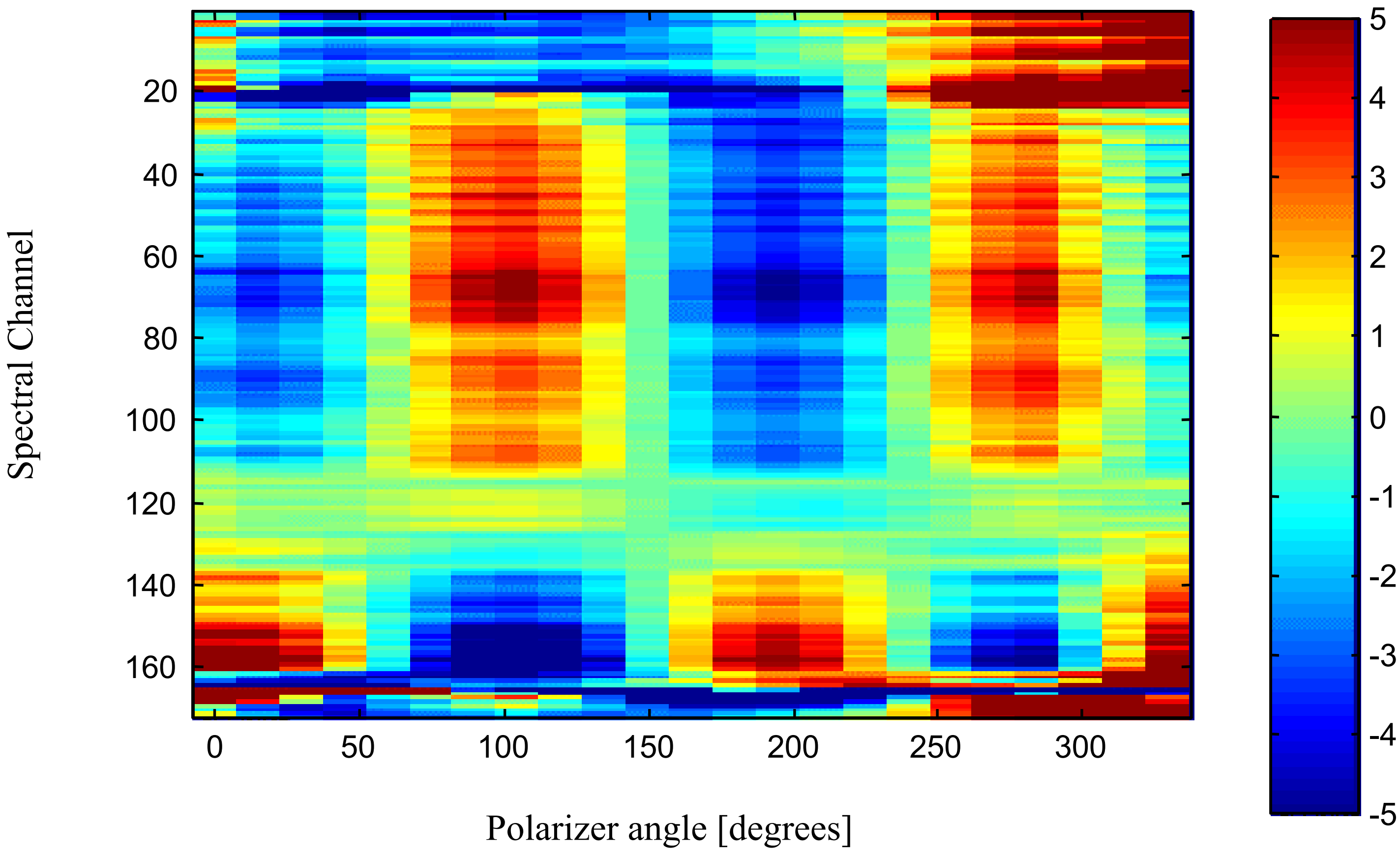
AVERAGED OVER FOV

In Units of Percent Polarization



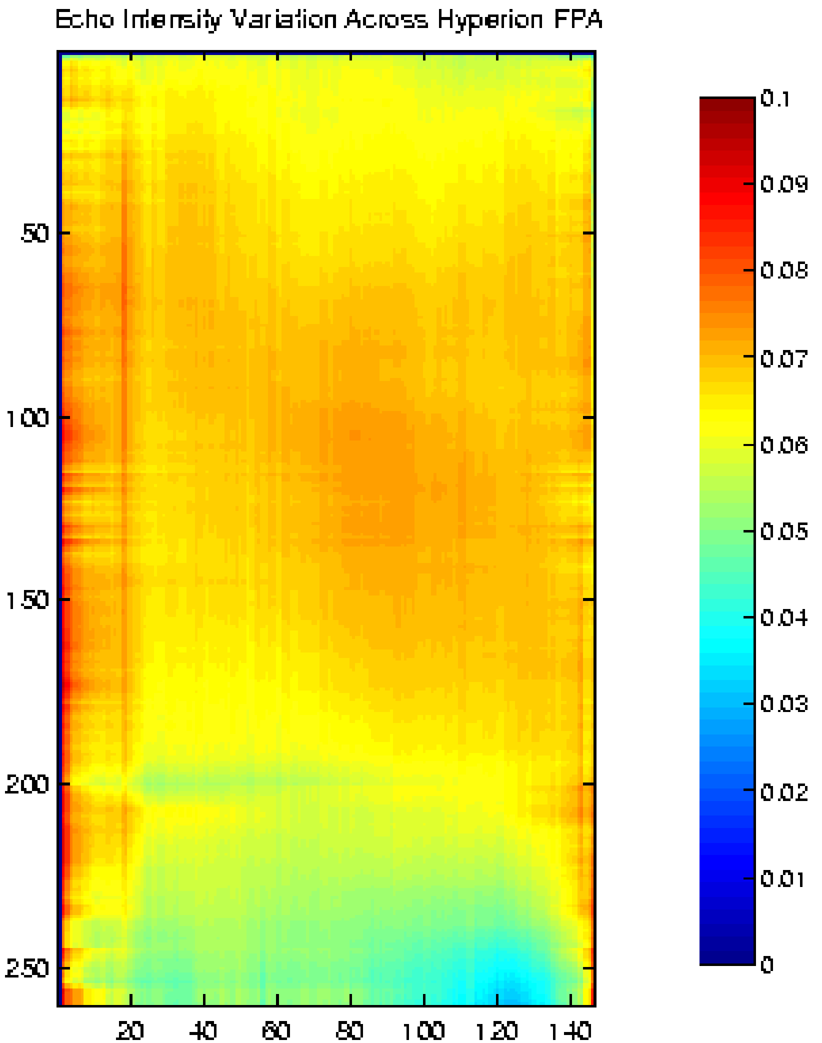
Polarization - SWIR

AVERAGED OVER FOV
In Units of Percent Polarization

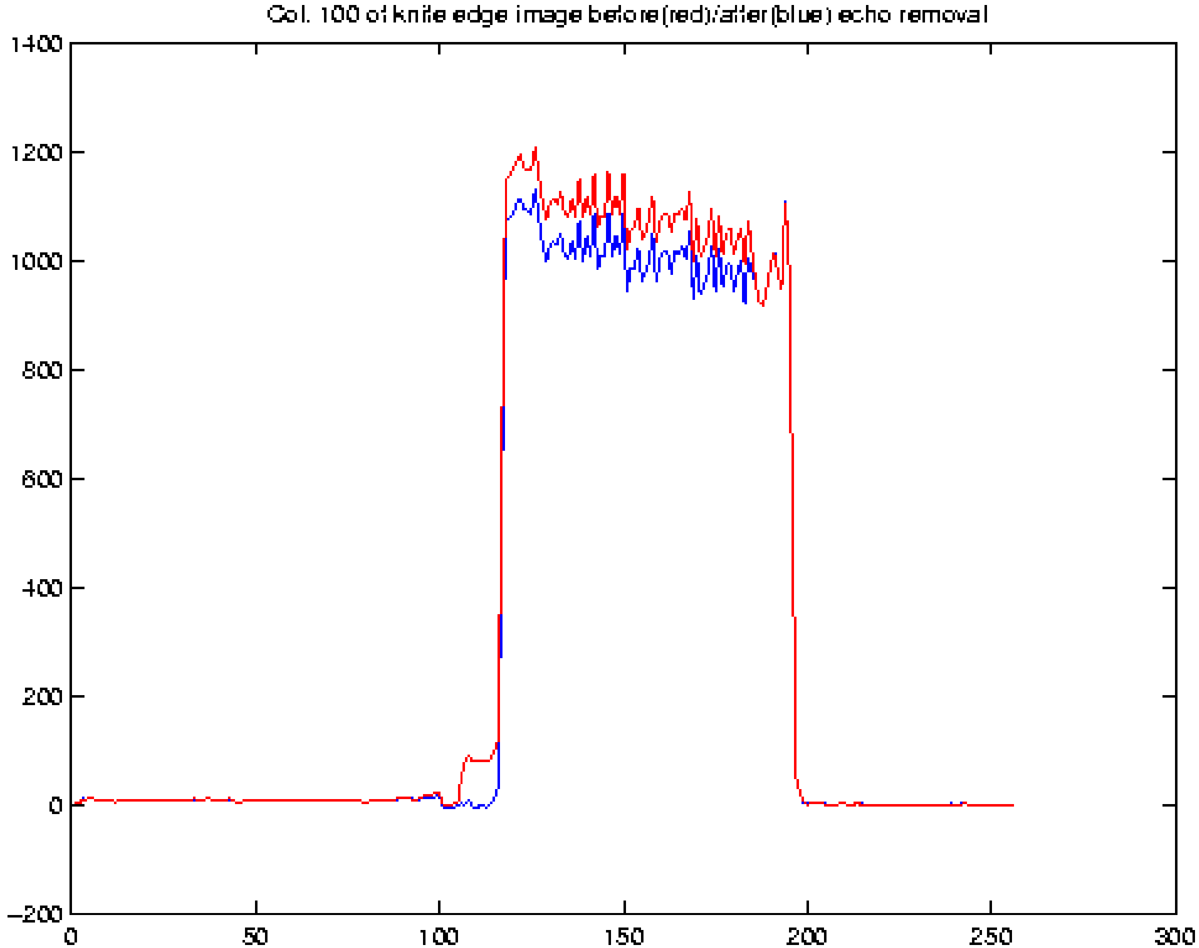


Echo Variation Map

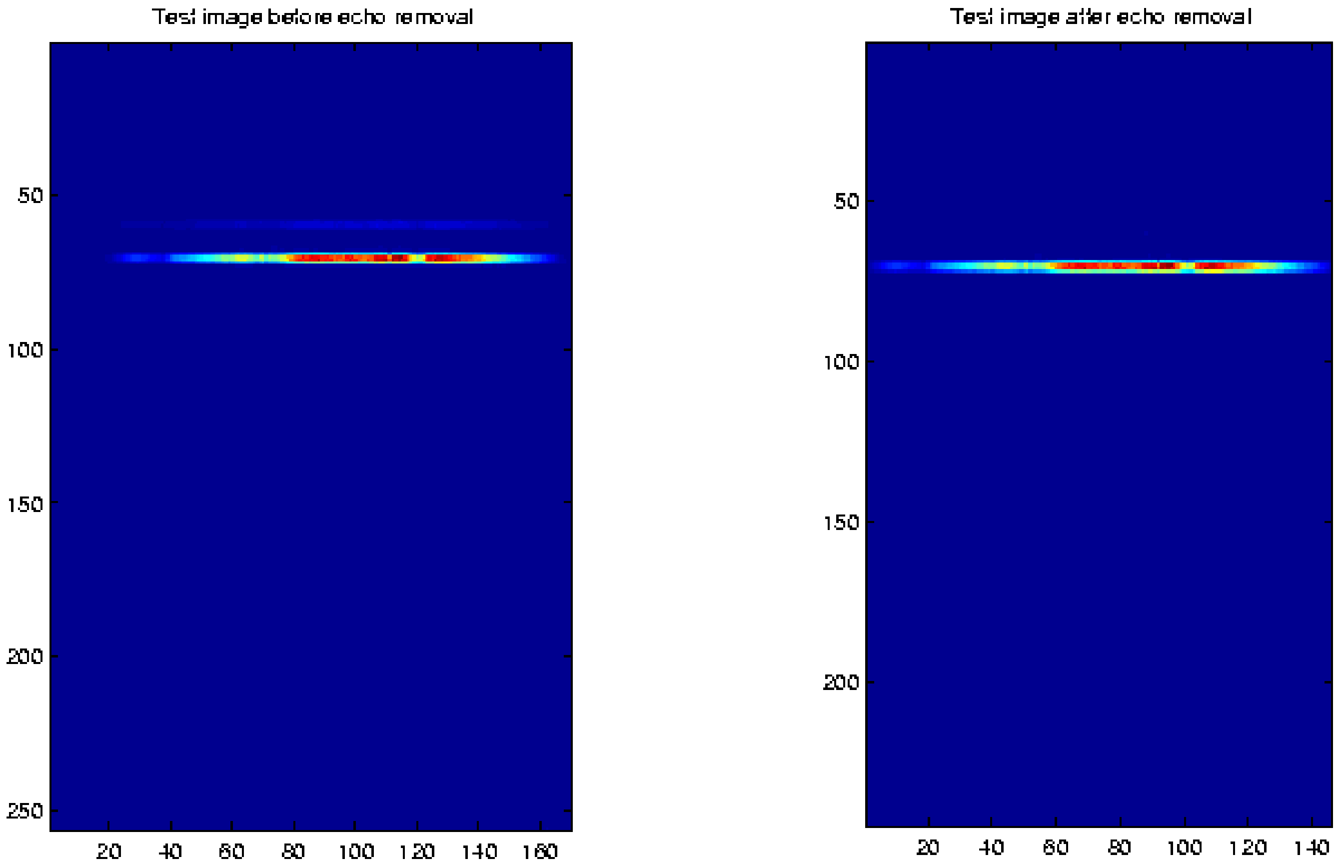
- Ratio images for each intensity level were interpolated to create maps of echo variation across the entire FPA.
- Based on this map a technique for echo removal was implemented.
- The removal algorithm consists of subtracting a shifted and scaled version of each frame from itself. Scaling is a multiplication of the image by the echo variation map.



Profiles Before(red)/After(blue) Echo Removal



Hyperion Images Before / After Echo Removal



Anomalous Pixel Collage –
Possibilities including striping

Band	Swath Pix	Band	Swath Pix	Band	Swath Pix	Band	Swath Pix	Band	Swath Pix
5	114	12	6	32	114	99	92	168	256
5	141	12	114	33	114	**116**	**127**	**168**	**256**
6	6	13	114	34	114	**116**	**127**	**169**	**12**
6	68	14	114	39	177	**116**	**138**	**169**	**12**
6	172	14	247	40	13	**116**	**138**	**169**	**23**
7	7	15	114	48	20	**119**	**240**	**169**	**23**
7	68	16	114	49	20	**119**	**240**	**169**	**23**
7	179	17	114	50	20	**120**	**240**	**190**	**113**
7	185	18	114	51	20	**120**	**240**	**190**	**113**
8	7	19	114	52	13	121	196	**190**	**113**
8	12	20	114	52	33	125	54	**200**	**8**
8	68	21	114	53	33	125	115	**200**	**8**
8	114	22	114	57	13	125	160	**200**	**8**
8	121	23	114	**61**	**93**	125	164	**201**	**8**
9	6	24	114	**72**	**95**	127	40	**201**	**8**
9	68	25	114	**72**	**95**	127	66	**201**	**8**
9	114	26	114	75	2	127	213	**203**	**104**
10	6	27	47	**94**	**82**	128	30	**203**	**104**
10	114	27	114	**94**	**82**	128	96	**203**	**115**
10	131	28	47	**94**	**93**	128	126	**203**	**115**
10	199	28	114	**94**	**93**	165	148	203	116
11	6	29	114	**99**	**81**	**165**	**248**	222	98
11	114	30	114	**99**	**81**	**168**	**245**	225	186
11	199	31	114	**99**	**92**	**168**	**245**		

Bad Pix 3 - L1B1; CSIRO – Jupp; Dark Image – Pearlman; PFC - Han

Hyperion Nominal Data Modes

Mode	Cover Position	Data collect	Comment
Standby	Closed	None	Default mode for active state
Dark calibration	Closed	Minimum 100 frames	Performed as close as possible to imaging, before and after
Lamp calibration	Closed	Minimum 100 frames	Performed after second dark calibration; two radiance levels
Solar calibration	Open 37 degrees	Minimum 1 second Nominal 1 cube	Performed over North Pole only to keep cover out of ALI keep-out zone; yaw maneuvers required
Lunar calibration	Fully open (135^o)	Minimum: 1 second Nominal: 1 cube	Performed on dark side of earth; off-track spacecraft pointing required
Ground calibration	Fully open (135^o)	Minimum 1 second Nominal 1 cube	Ground target selected
Imaging	Fully open (135^o)	Minimum 1 second Nominal 9 cubes	Nominal data collect is equivalent to Landsat scene, and takes 27 seconds.

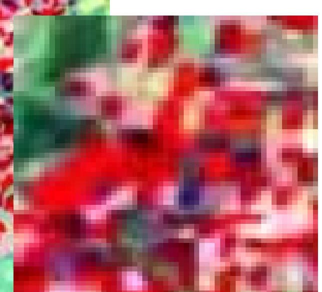
Performance Tradeoffs

- Spatial Resolution
- Spectral resolution

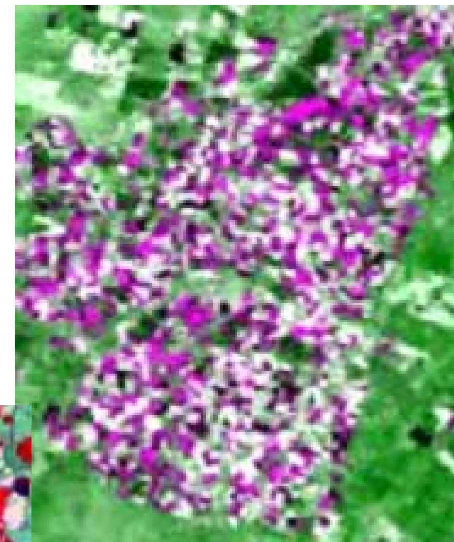
Coleambally Image Collection – 30m to 1km

Landsat

SAC-C



MODIS



250m

MODIS

1000m

Comparison of Hyperspectral Instruments

Parameter	Lewis HSI	Hyperion
Volume (L x W x H, cm)	43x69x94	39x75x66
Weight (Kg)	39.5	49
Avg Power (W)	66	51
Peak Power (W)		126
Aperture (cm)	12	12
IFOV (mrad)	**0.057**	**0.043**
Crosstrack FOV (deg)	0.84	0.63
Wavelength Range (nm)	380 - 2450	400 - 2450
Spectral Resolution (nm)	**5.1/6.45**	**10**
No. Spectral Bands	**384**	**220**
Digitization	12	12
Frame Rate (Hz)	**237**	**225**
Typical SNR	**100 - 200**	**65 - 130**
Spectral Calibration (nm)	1	1
Radiometric Calibration	<6%	<6%

Hyperion Calibration
Ground Test and On-Orbit Validation

- System performance assessment strategy
- Present pre-flight and on-orbit measurement techniques
  - Absolute Radiometric Calibration
  - Spectral Calibration
  - Image Quality Characterization

Strategy

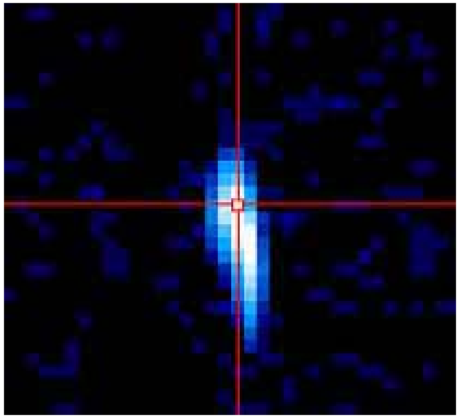
- Pre-Flight:
  - Establish fundamental characteristics of the instrument and assess requirement compliance
  - Establish instrument performance through the build, environmental test and spacecraft integration phases
  - Provide solid foundation for on-orbit comparison

- On-orbit:
  - Determine on-orbit performance and compare with pre-flight performance
  - Define data collects that can be used to assess identified performance parameters
  - Acquire and analyze data collections; and *assess accuracy of technique*
  - Compare on-orbit results with pre-flight

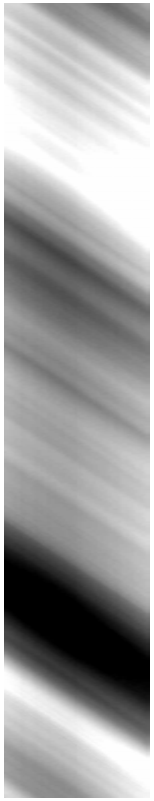
Special Targets for Characterization

Searchlights
-California

Planets
-Venus



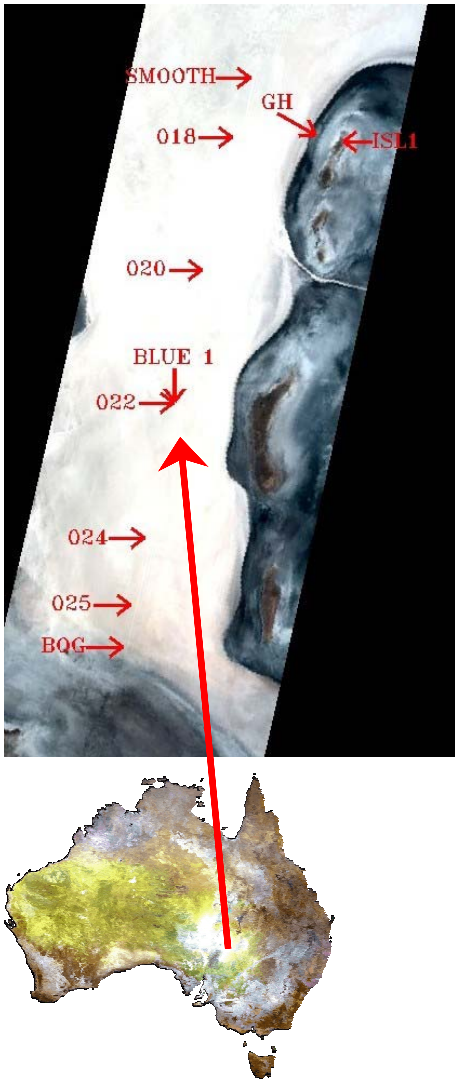
Gas Flares
-Moomba



90 deg
Yaw

Desert Sites used for Vicarious Calibration

Lake Frome



RR Valley

Arizaro/Barreal Blanco

the good

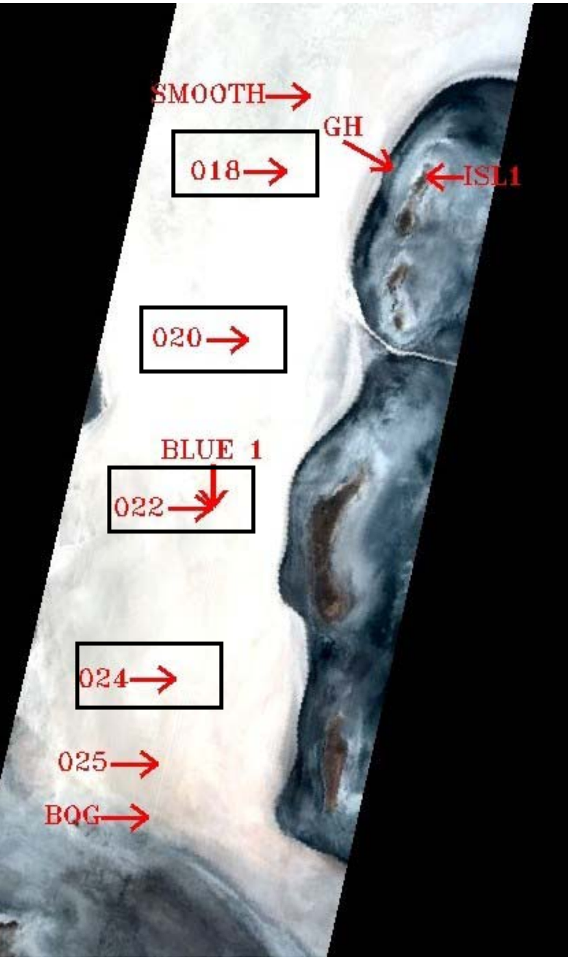
the bad

and

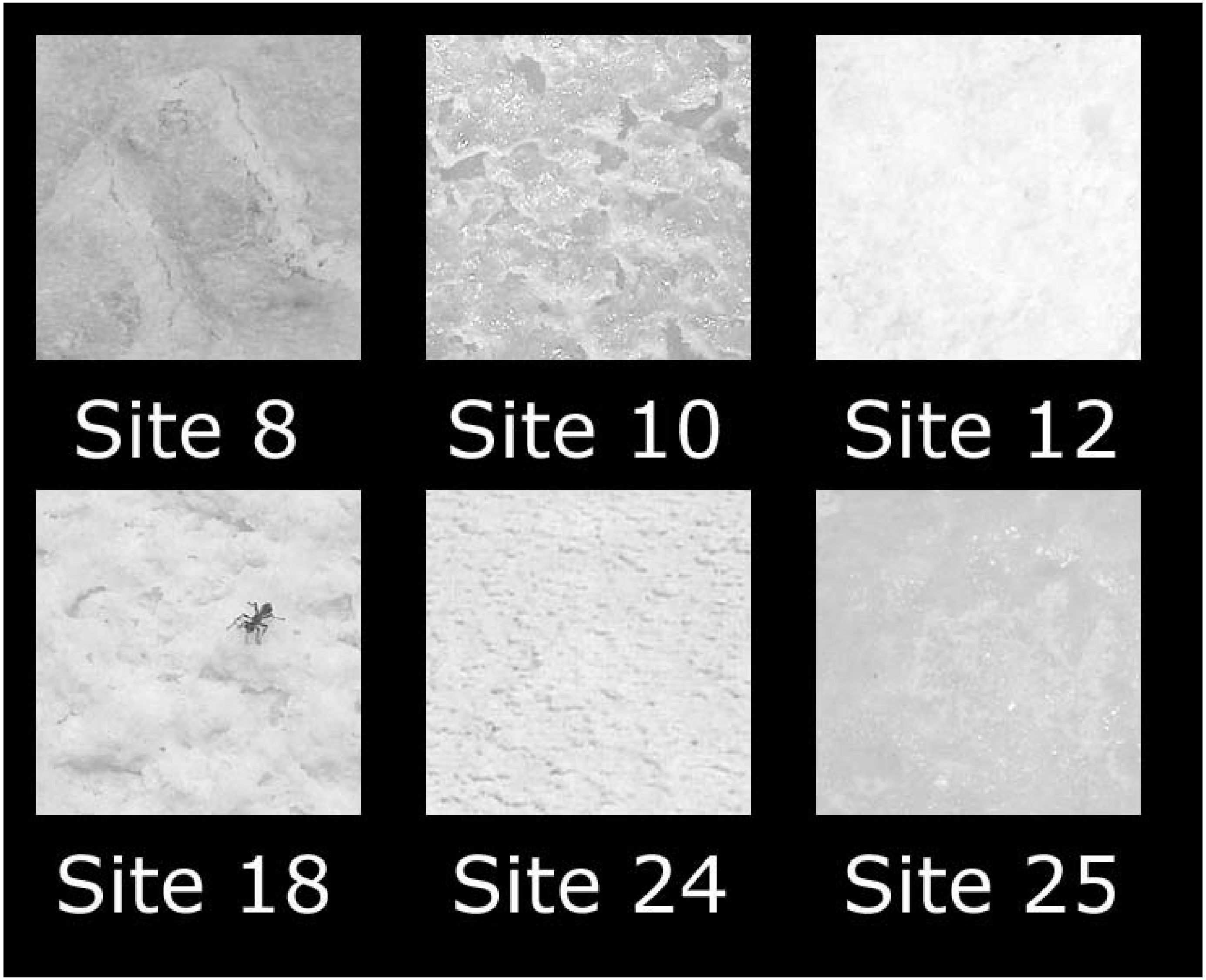
the gotchas

The world of real life operations

Images of Lake Frome

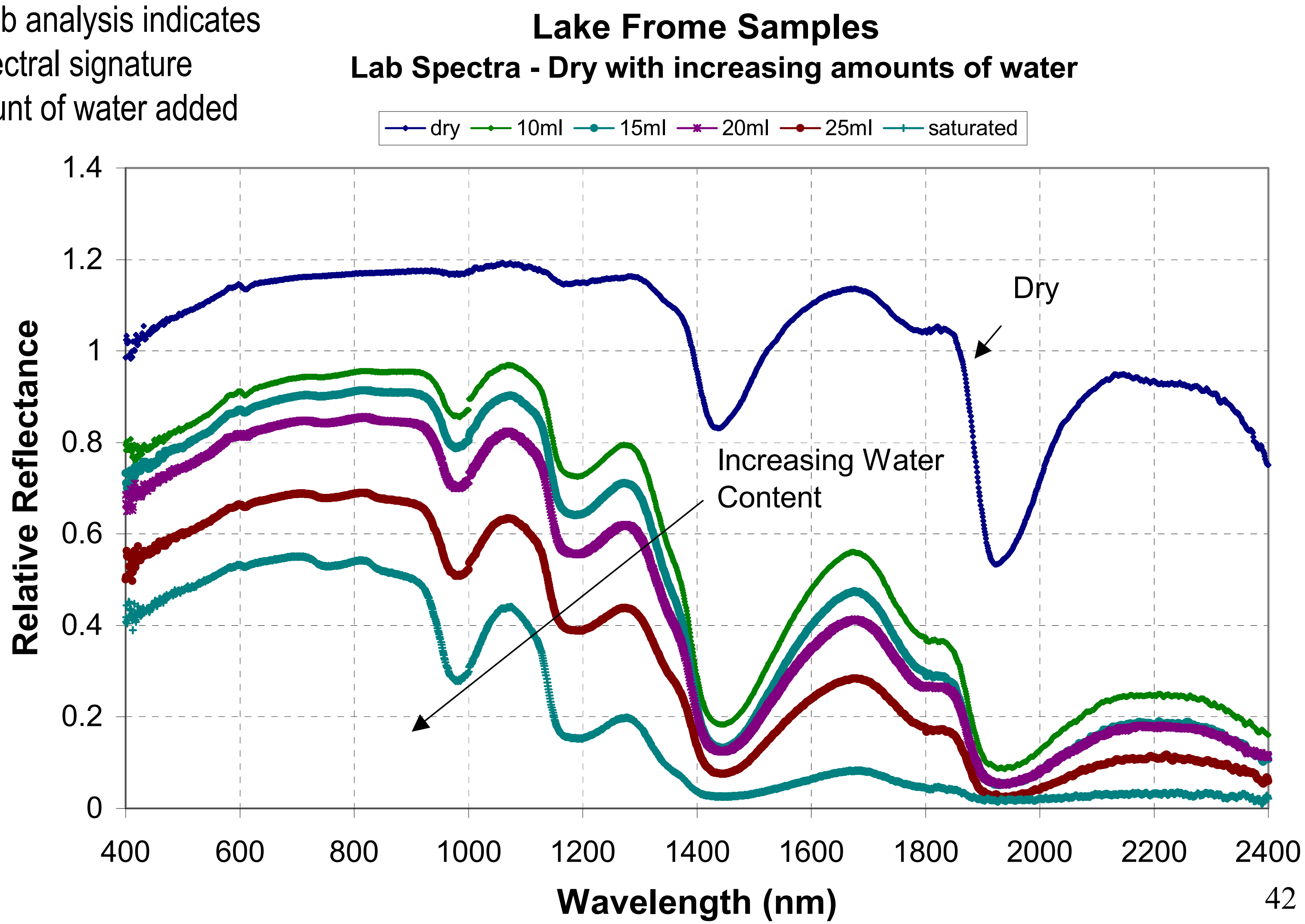


The Salt Surface



Impact of Moisture on Salt Signature

Approximate lab analysis indicates changes of spectral signature based on amount of water added



The Locusts That Did Not Make It Over

Calibration signature
Or
New culinary delicacy?

Arizaro Dry Lake bed

Ground Calibration

Coleambally Collection History

Date(GMT)	Julian Day(GMT)	Path/ Row	X-track Position	In-track Position	Comments
23-Dec-00	358	93/84	100	3148	Cum(South);VNIR
1-Jan-01	001	92/84	82	2131	Clear
9-Jan-01	009	93/84	-	-	UARDRY
17-Jan-01	017	92/84	-	-	Cloudy
25-Jan-01	025	93/84	95**	2481**	High Clouds(North)
2-Feb-01	034	92/84	36	2599	Clear
9-Feb-01	041	93/84	-	-	Cloudy
18-Feb-01	049	92/84	41	3782	Shadows(East)
25-Feb-01	056	93/84	-	-	
6-Mar-01	065	92/84	95	3977	Clear
13-Mar-01	072	93/84	104	2926	Clear
22-Mar-01	082	92/84	-	-	Cloudy

ARIZARO, Argentina

High altitude (~12,000ft) dry salt lakebed.

Surface: extremely rough, locally uniform, and bright.

Some years no rain at all!

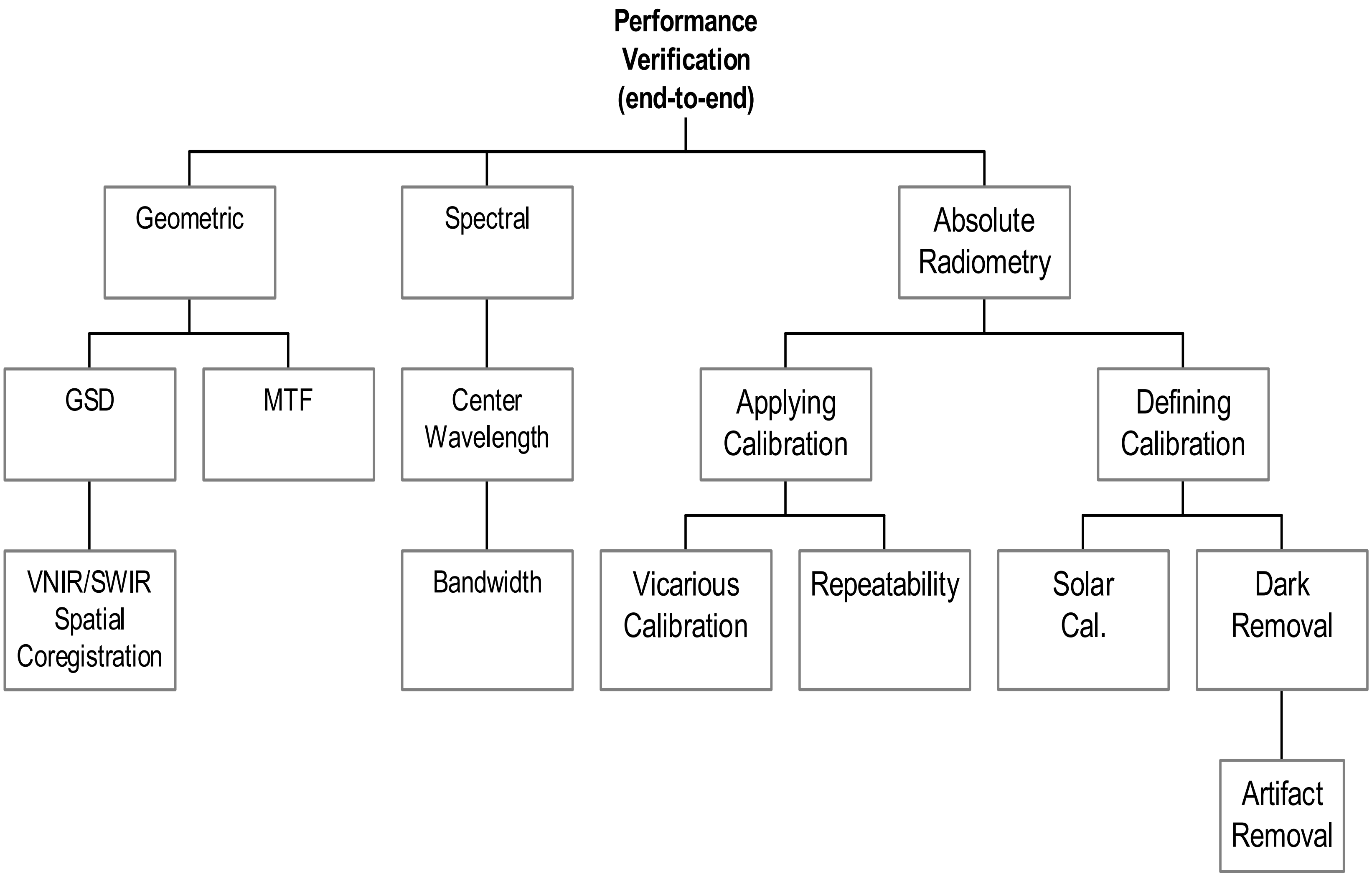
It rained the day we arrived

The good news is that conditions were good for the experiment on the 7th of February 2001.

Calibration & Validation Can Be Fun

Let’s Address Calibration

System Performance Verification Strategy



Absolute Radiometric Calibration

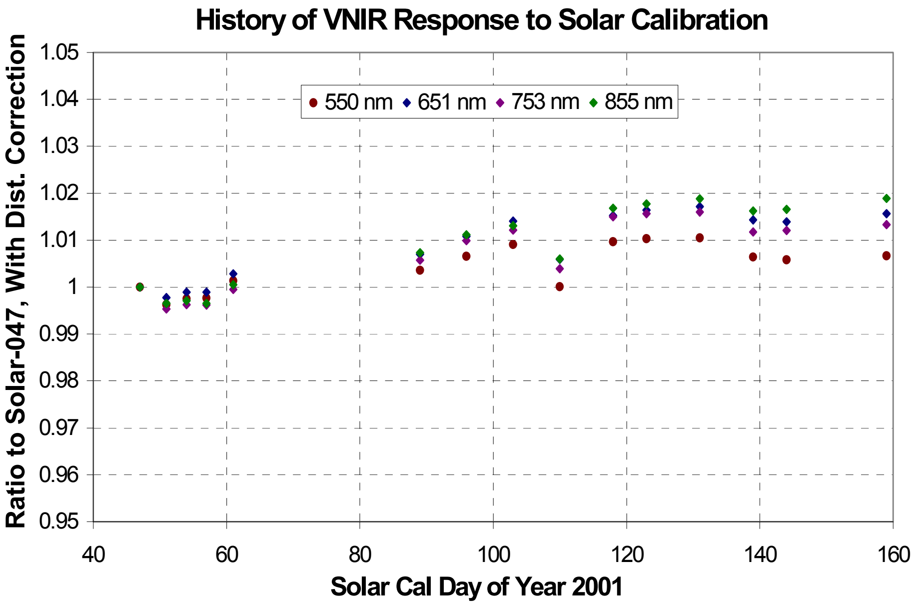
Key Factors Impacting Calibration

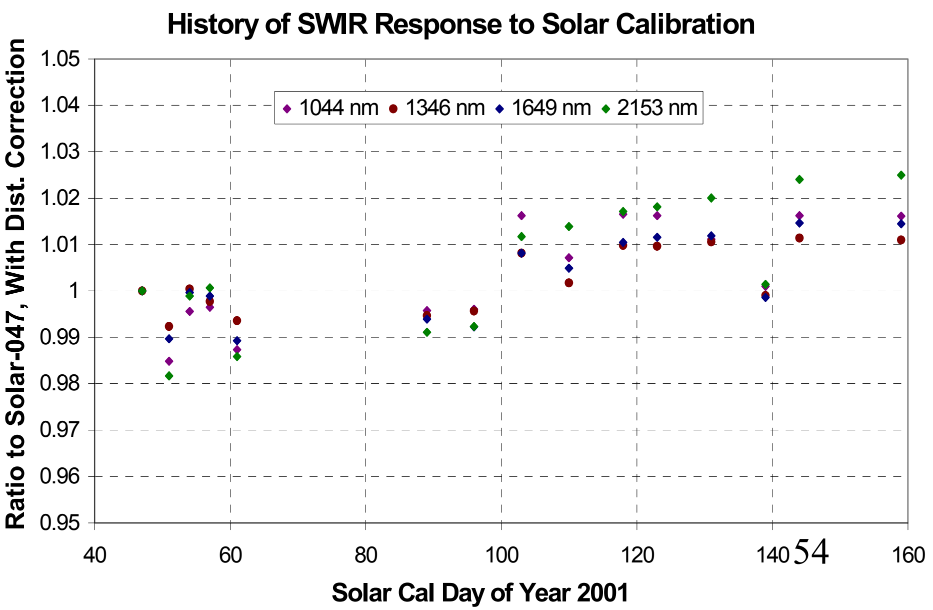
	Absolute Knowledge	**Intermediate Properties**	**Spacecraft Pointing**	**Strengths**
Solar Calibration	Models avail to community VNIR more accurate then SWIR	Diffuse reflectance of Hyperion cover	Critical to modeling intermediate properties	Uniform across field-of-view Constant
Lake Frome (vicarious)	Based on ground truth measurements	Atmospheric effects must be modeled	Depends on surface	User oriented effort
Lunar Calibration	Based on Lunar models	none	Spacecraft scans moon. Relative moon, sun, sat angle	No intermediate properties. Constant

On-Orbit Radiometric Calibration

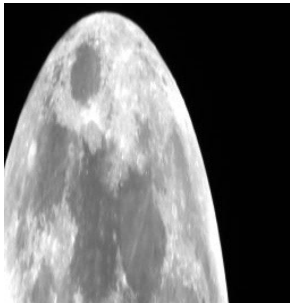
- Solar Calibration
  - Absolute Comparison: VNIR within 2%, SWIR 5-8% low; SWIR has larger uncertainty due to solar model and BRDF model of cover surface
  - Used to correct for pixel-to-pixel corrections
  - Included in repeatability assessment 0.6% for VNIR, 1.6% for SWIR
  - Used to define noise level as a function of signal level to determined SNR
- Lunar Calibration
  - Used to eliminate artifacts of diffuser
- Vicarious Calibration and Cross Calibration
  - Calibration accuracy in 4-10% range

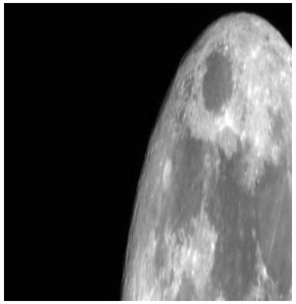
Solar Calibration Trending

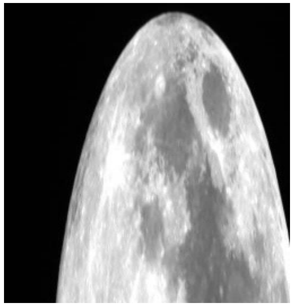




Lunar Calibration Trending


Day 038


Day 069

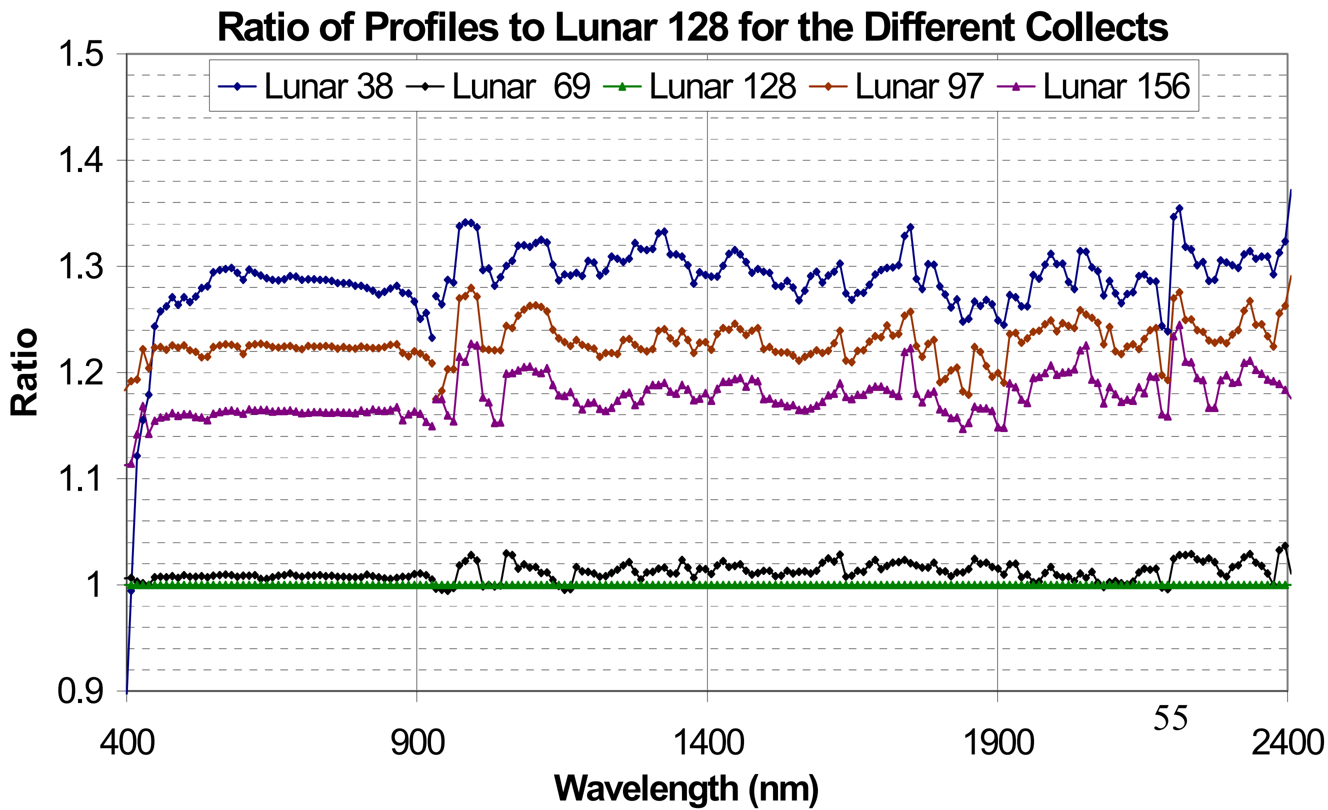

Day 097


Day 128

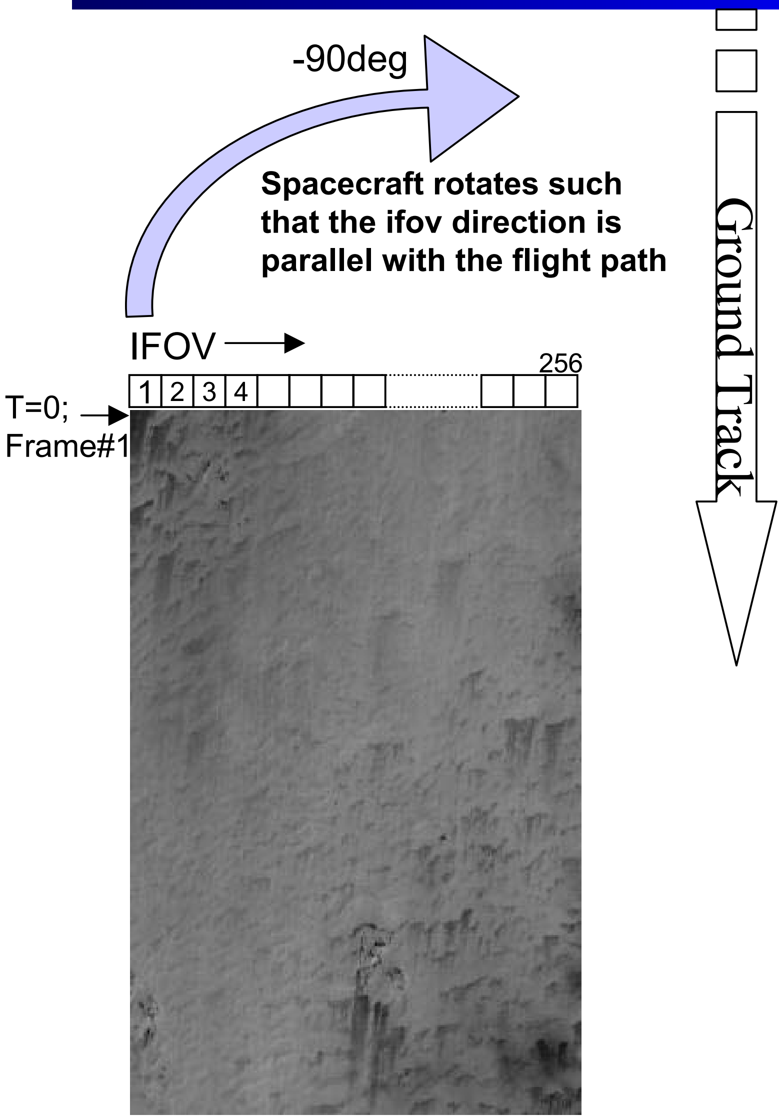

Day 156

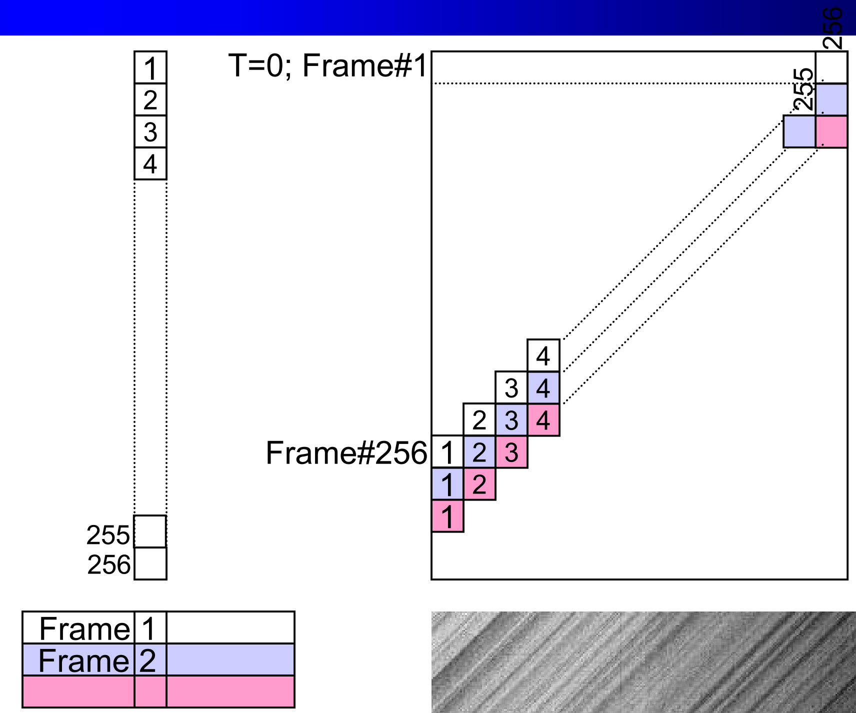
Intensity Dependent on Moon-Sun-Spacecraft Relative Angles

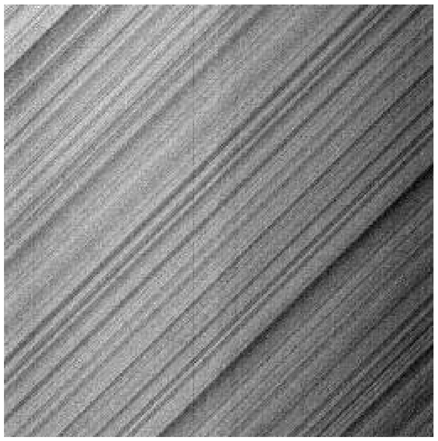
VNIR and SWIR track intensity changes



Description of the 90 Degree Yaw Collect





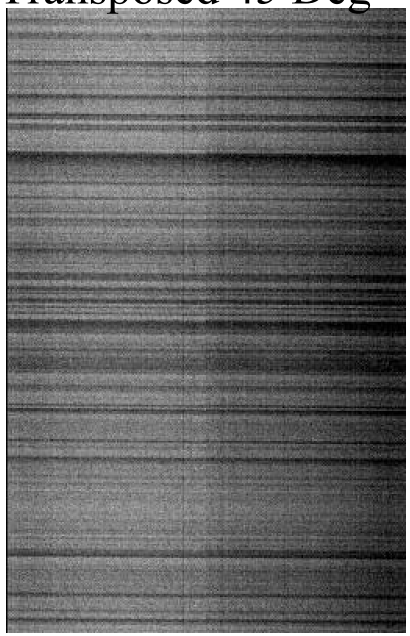


The target spot is viewed by pixel 256 during frame 1, and then by pixel 255 during frame 2, and so on…

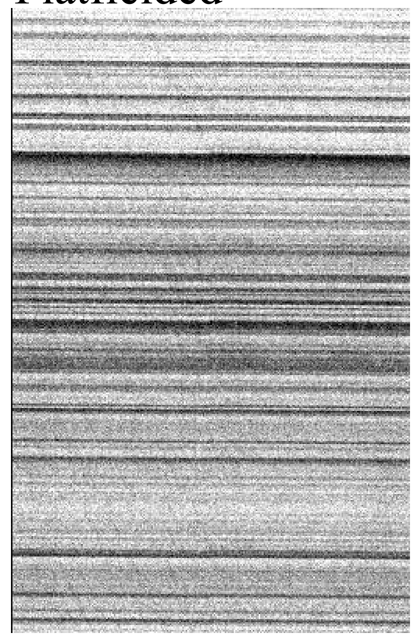
Analysis for Yaw Data

Level 1R

Transposed 45 Deg



Flatfielded



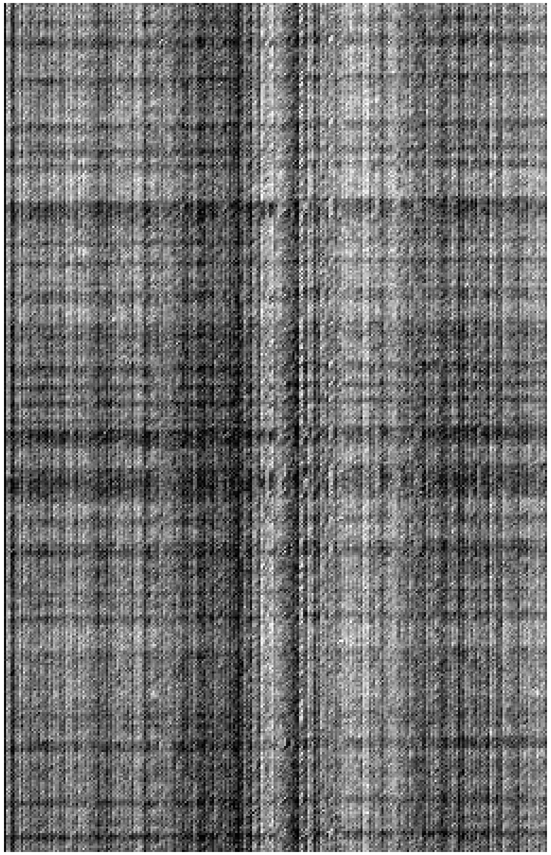
Flatfielded
(using 2001197 coefficients)

Differences between the sets of correction coefficients for the entire vnir focal plane are within ±2.5%

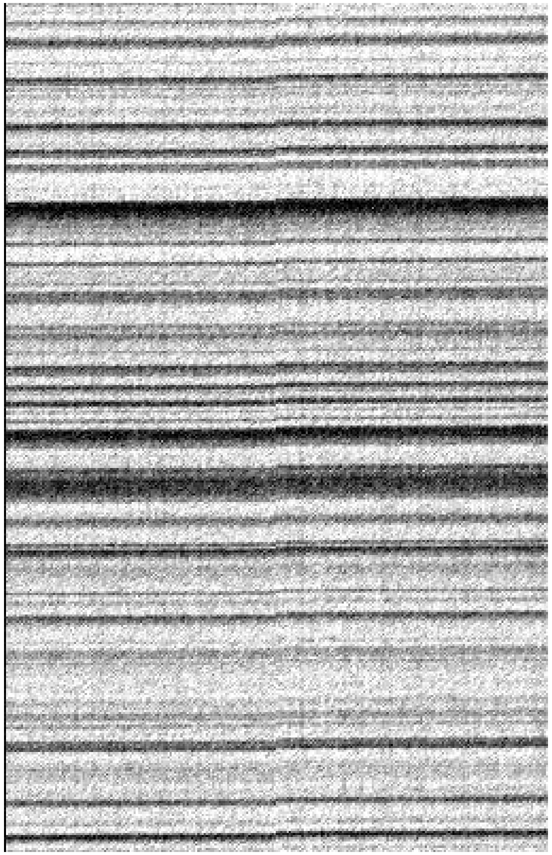
Hyperion Band 20 (VNIR)

Analysis for Yaw Data

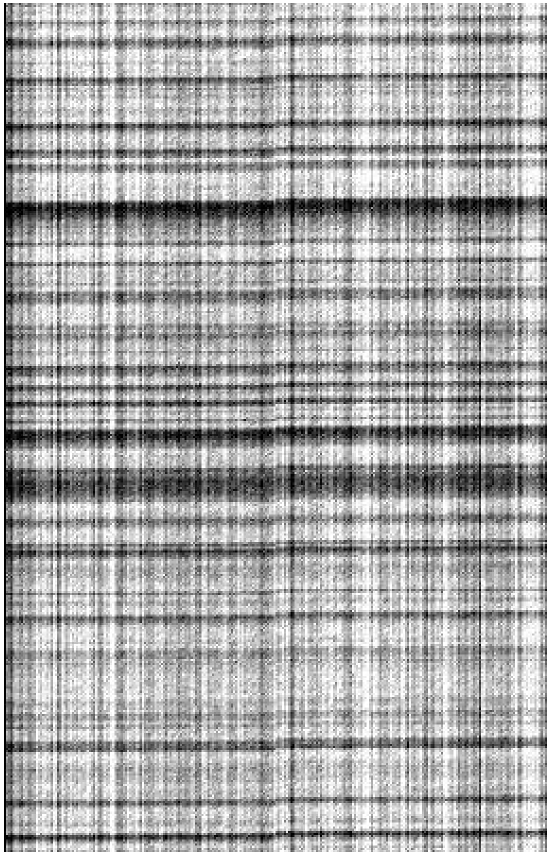
Transposed 45 Deg



Flatfielded



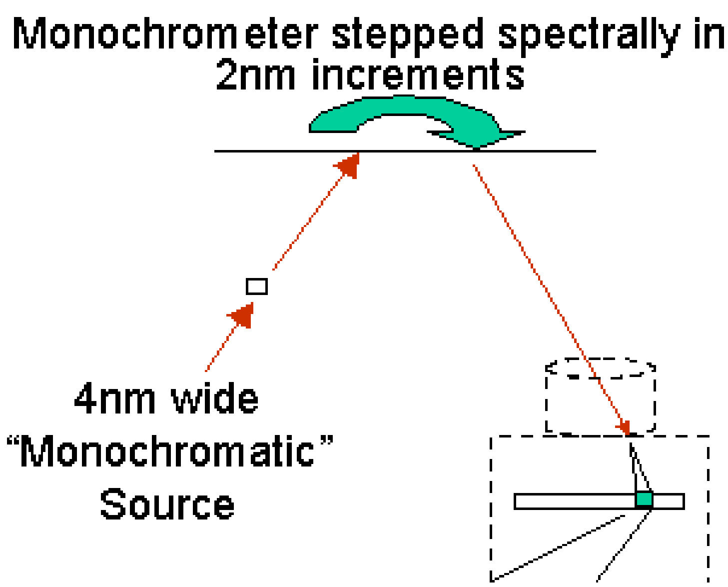
Flatfielded
(using 2001197 coefficients)

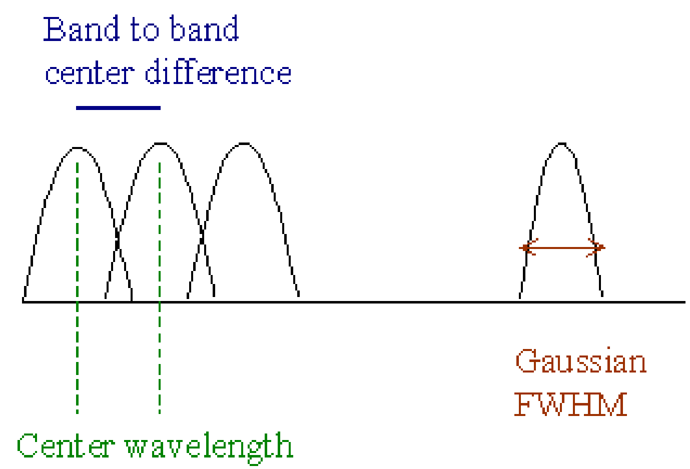


Hyperion Band 91 (SWIR)

Spectral Calibration

Pre-Flight Spectral Calibration





Monochrometer profiles used to define center wavelength and bandwidth at discrete locations

Spectral response modeled as a gaussian with a center wavelength and full width half max

Pre-Flight Spectral Calibration

- Center wavelength and bandwidth
  - Measured at discrete locations 20 VNIR locations, 25 SWIR locations.
  - Used to define the center wavelength and bandwidth for every VNIR and SWIR pixel, 256 field-of-view locations and 242 spectral bands.

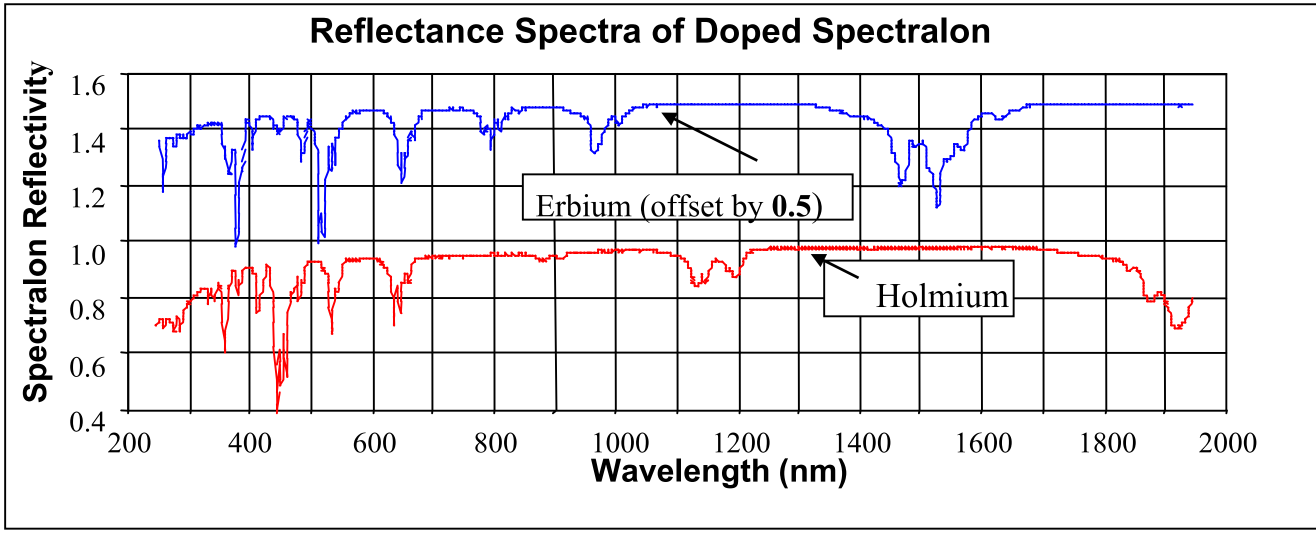
- Dispersion (nm/pixel)
  - Spacing of spectral channels, Hyperion dispersion (~10nm/pixel) closely matches the bandwidth (10 nm)

- Cross-track spectral difference
  - Maximum wavelength difference across field-of-view for a single spectral channel,
  - VNIR = 2.6-3.6 nm
  - SWIR = 0.40- 0.97 nm.

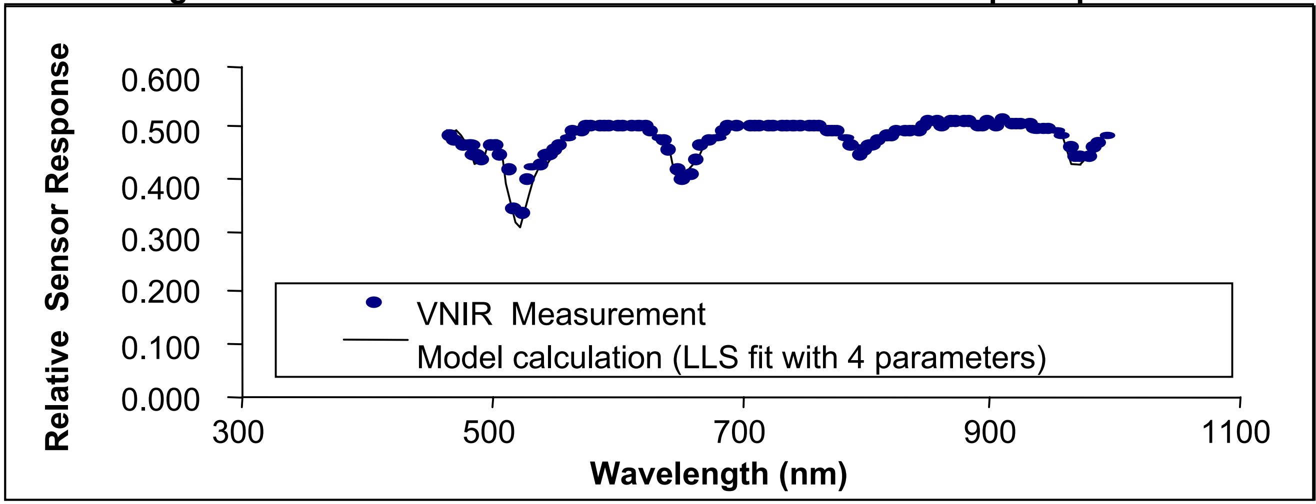
Pre-Flight Calibration Algorithm Development

Process:

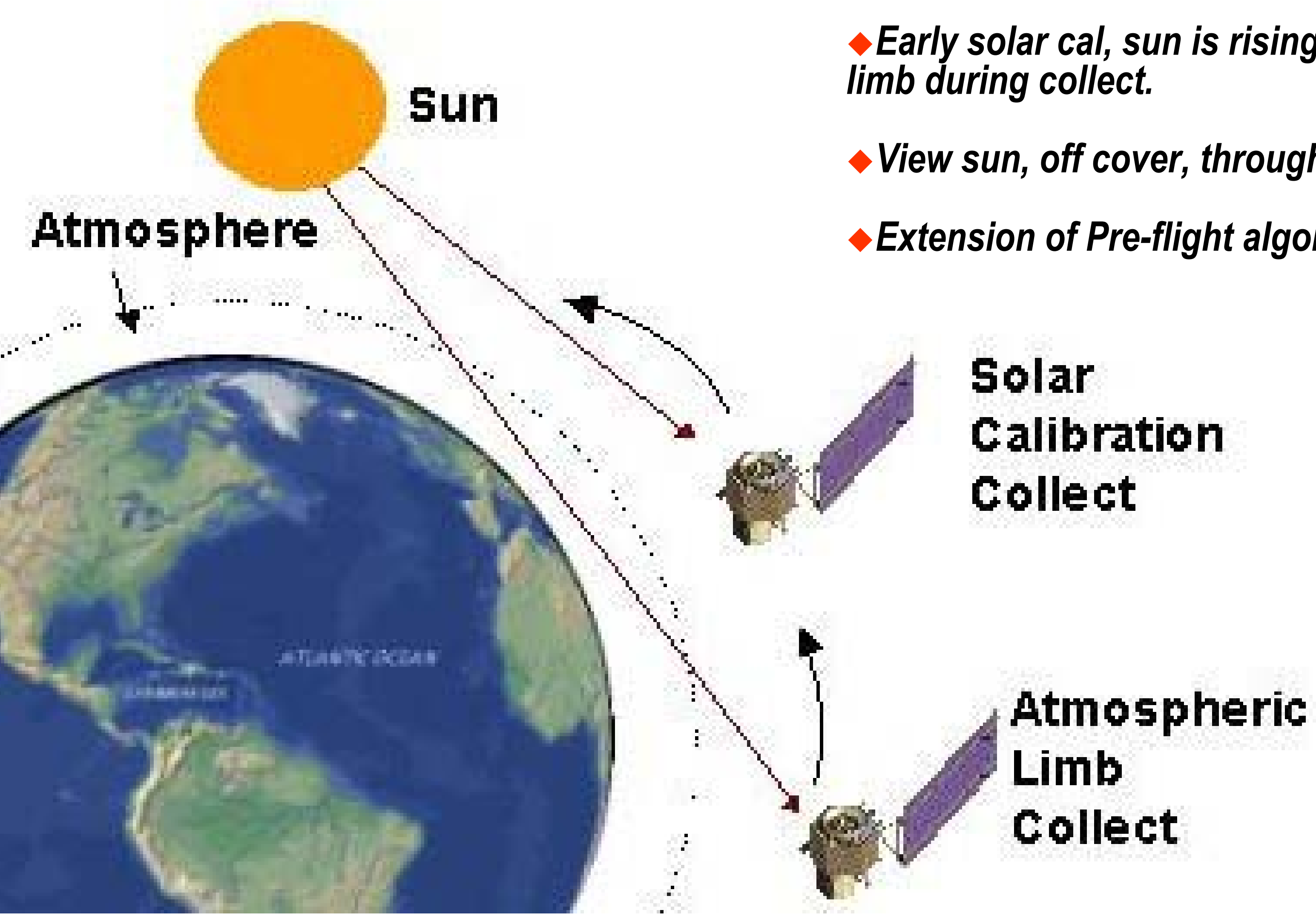
1. Image contains reference spectra uniform across the field of view. (pre-flight: doped spectralon)
2. High resolution reference spectra convolved with sensor spectral response function
3. Resulting reference spectra aligned with Hyperion measured spectra to determine spectral calibration.



High resolution scans of Holmium and Erbium Oxide doped Spectralon



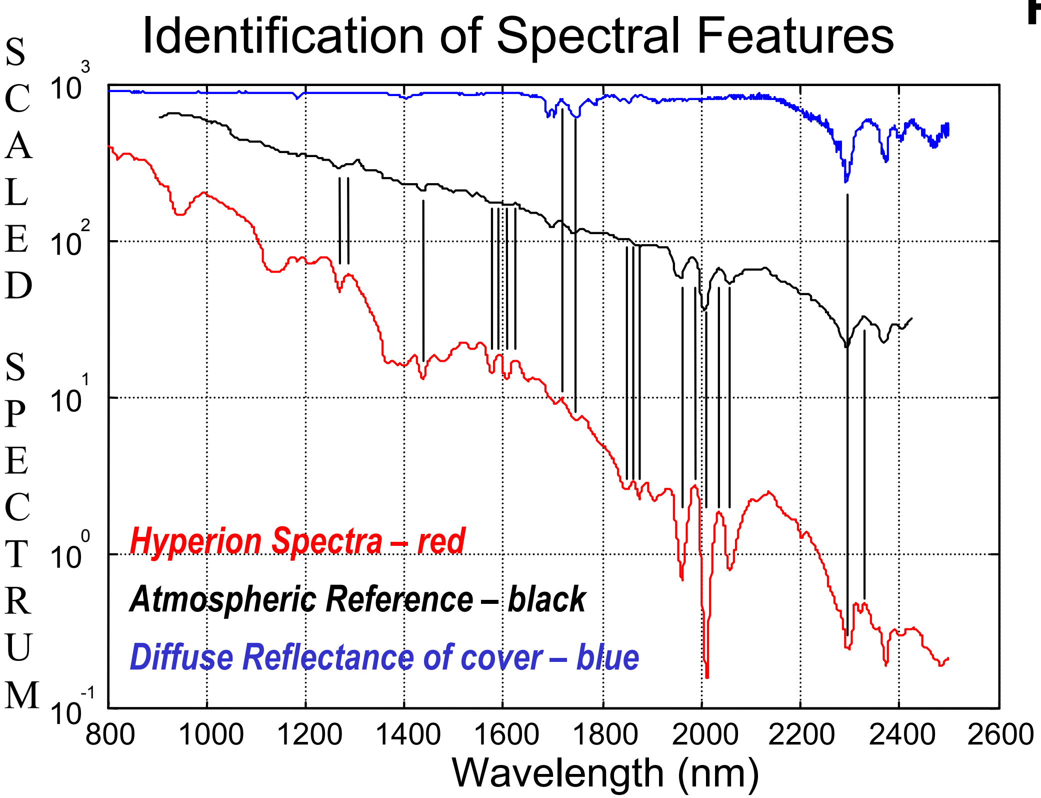
On-Orbit Spectral Calibration



- *Early solar cal, sun is rising through earth's limb during collect.*
- *View sun, off cover, through atmosphere*
- *Extension of Pre-flight algorithm development*

Spectral Calibration

SWIR



Process:

1.) Create Pseudo-Hyperion Spectra from reference: Modtran-3 for atmosphere, and Cary 5 & FTS measurements for diffuse reflectance of the cover

2.) Correlate Spectral Features: band number units of Hyperion max/min correlated with reference wavelength of max/min

3.) Calculate Band to Wavelength map: apply low order polynomial to fit the data over the entire SWIR regime

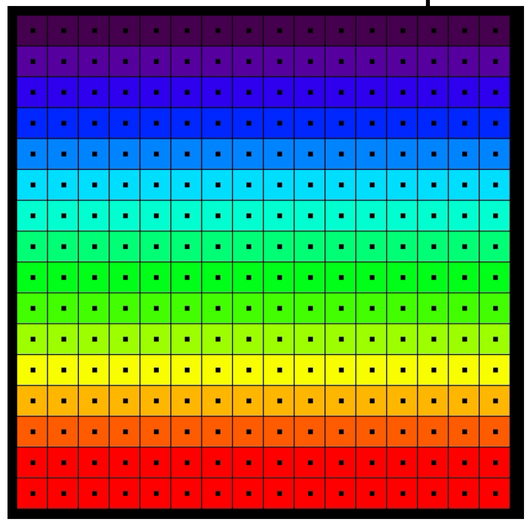
VNIR

- ***Spectral calibration based on two lines: one solar line (520 nm) and an oxygen line (762.5nm)***

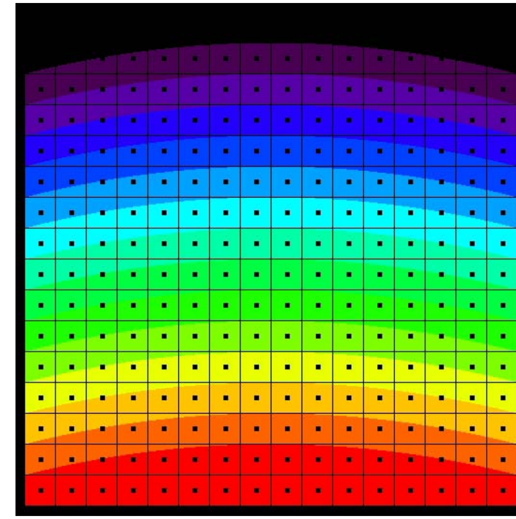
Spectral-Spatial-Uniformity

Depiction

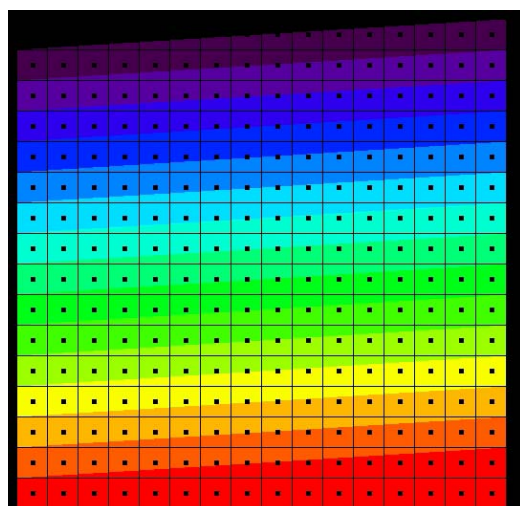
- Grids are the detectors
- Spots are the IFOV centers
- Colors are the wavelengths



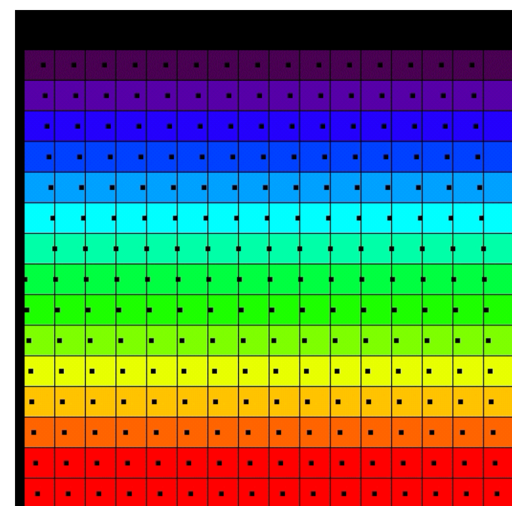
Requirement



Failure by Frown(aka smile)

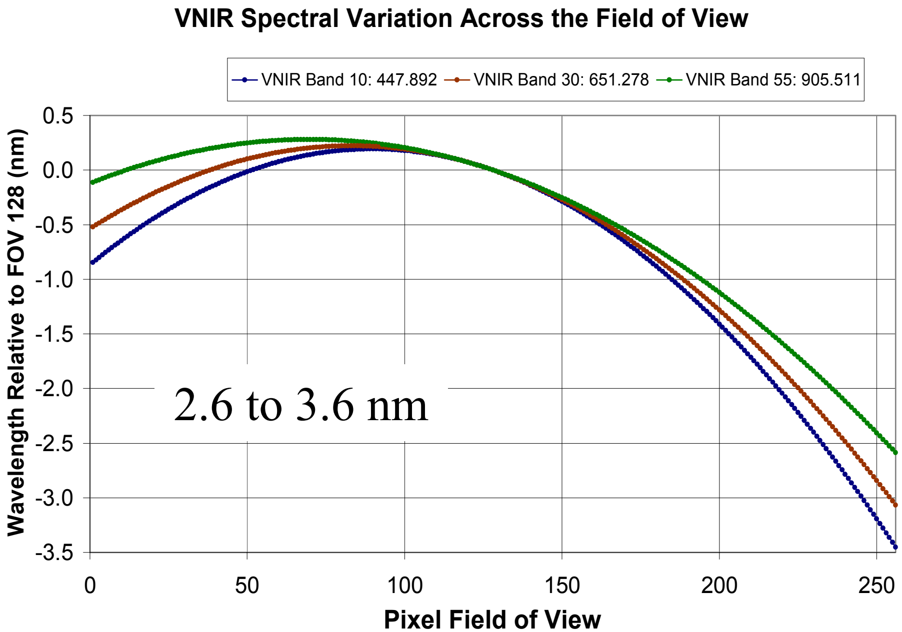


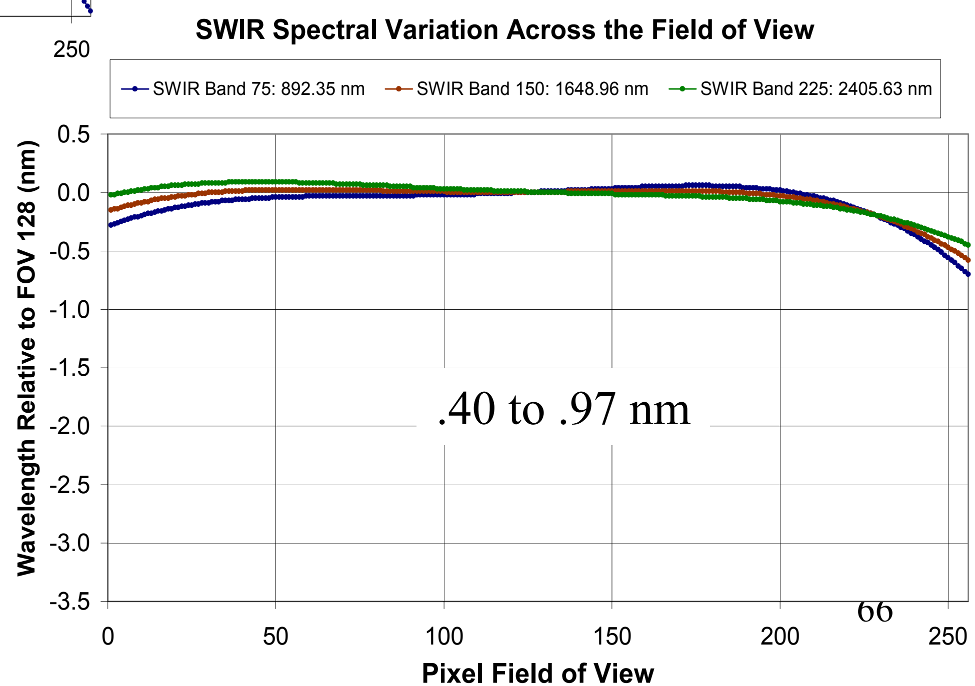
Failure by Twist



Failure by Spectral-IFOV-Shift

Characteristics of Spectral Calibration

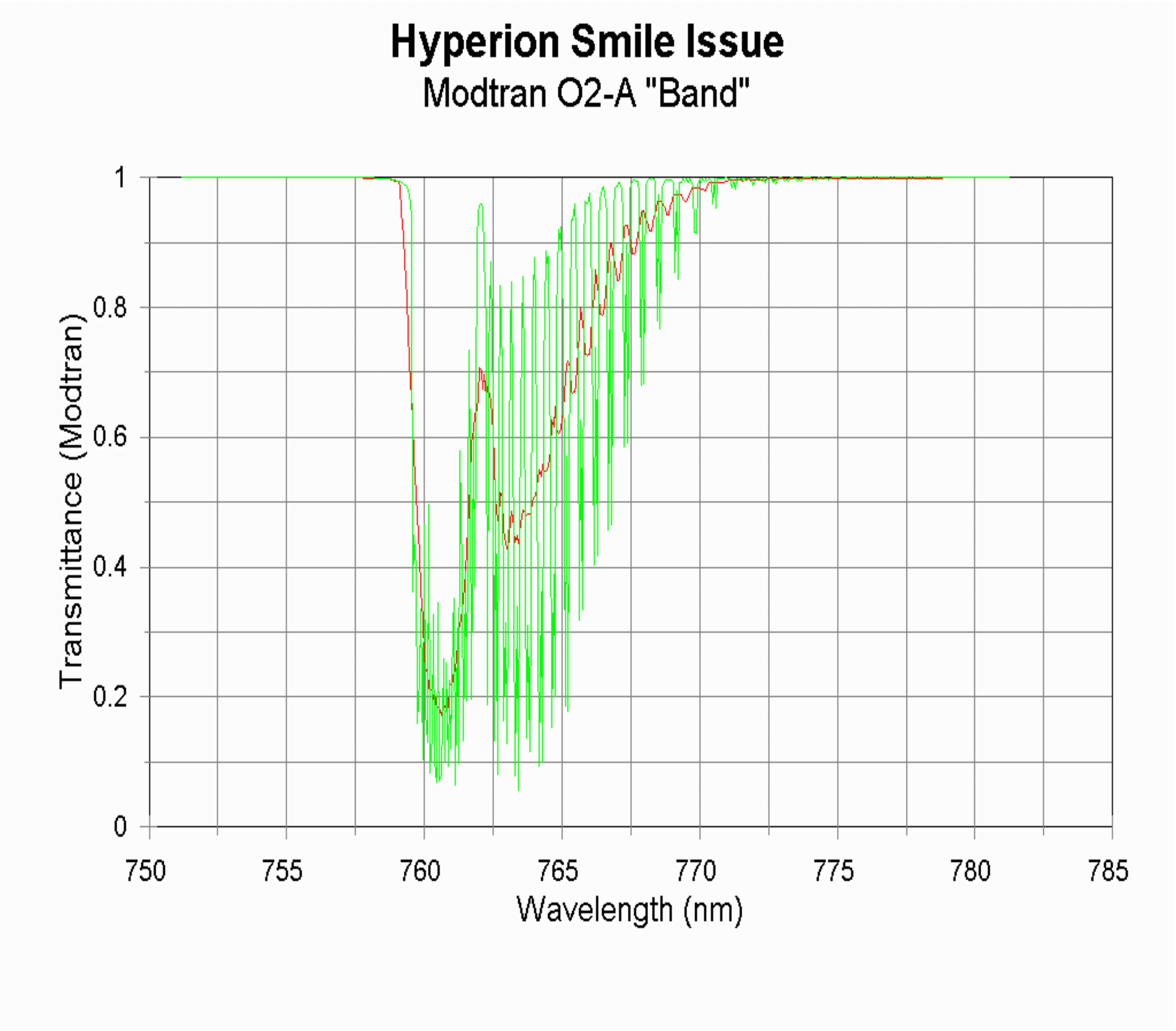




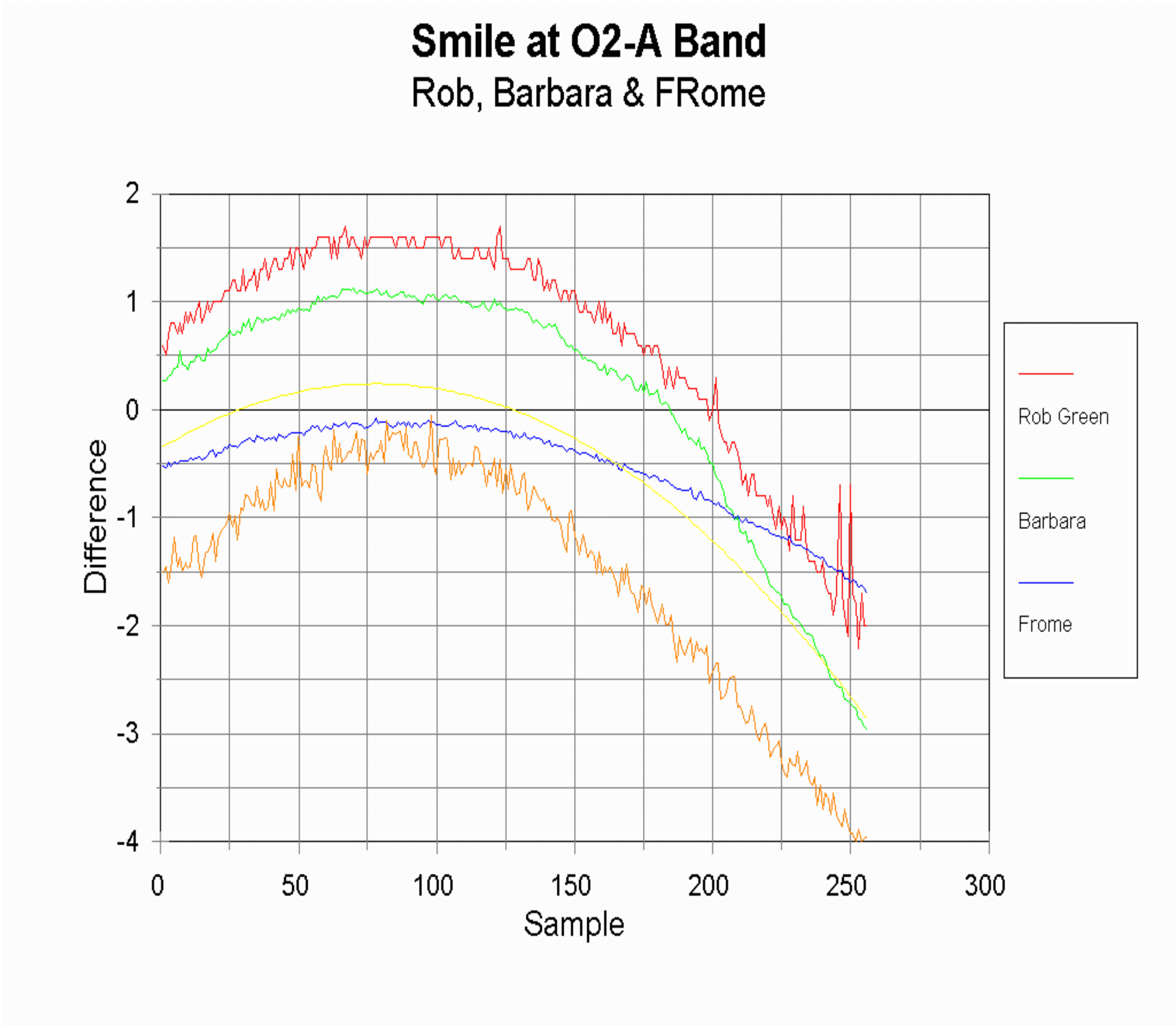
Understanding Frown

- Why is this important if it is only a few nm?
  - Some spectral effects such as movement of the chlorophyll edge due to stress is of the same order
  - Reproducibility of results due to variations in pointing

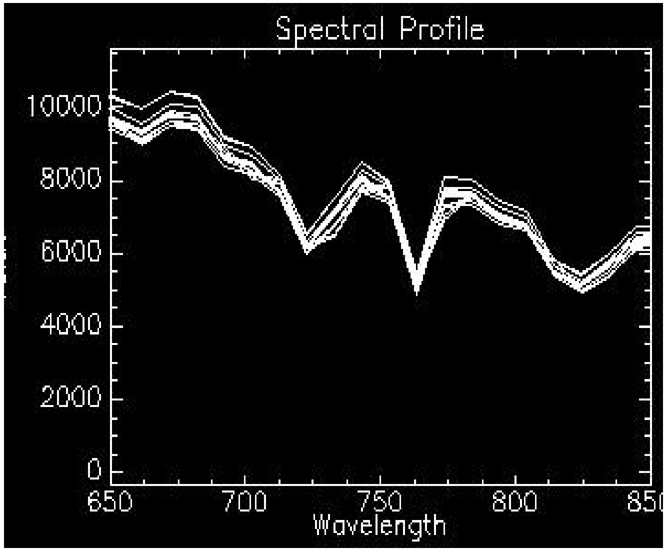
The Oxygen A Bands at 1 cm-1 & 1 nm Steps



Different Answers – Continuum?



Oxygen Line from Lake Frome Image



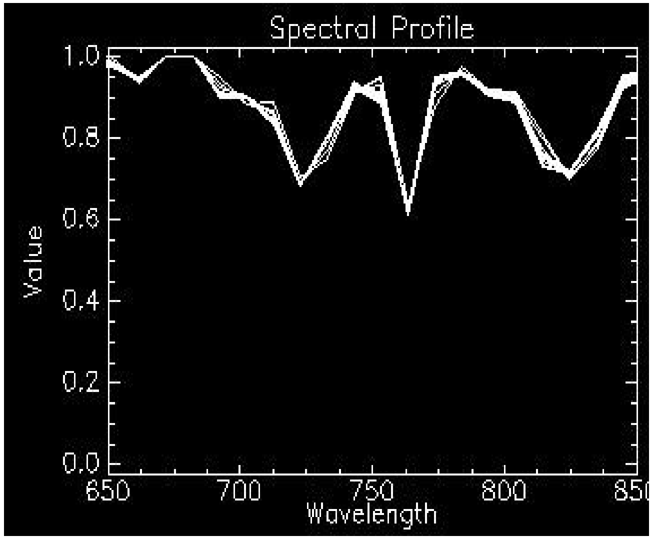


Image Quality

Image Quality

Cape Canaveral

Jan. 13 2001

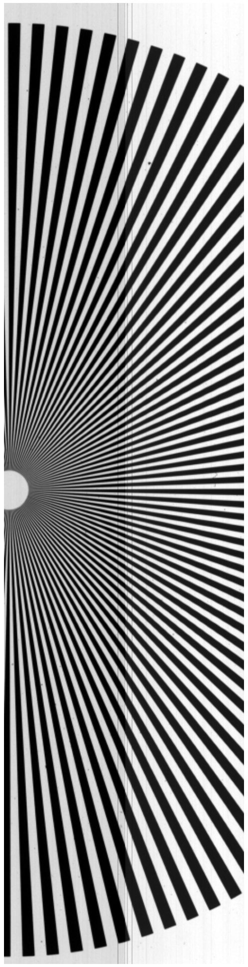
- Modulation Transfer Function
  - Pre-flight used knife edge and slit to measure Cross track direction, Along-track was Cross-Track*2/pi
  - On-orbit used Ice Shelf & Lunar Limb (knife edge) and bridge (slit) to measure Cross-Track and Along-Track directly.
- Co-registration of VNIR and SWIR
  - Pre-flight used test bed to project a slit with a broad spectrum at multiple locations
  - On-orbit used combination of edges (Lunar, Ross), point sources (clouds, flares), ground control points
- Ground Sample Distance
  - Pre-flight measured IFOV using test bed
  - On-orbit triangulated marked features in well mapped scene

MTF Approach

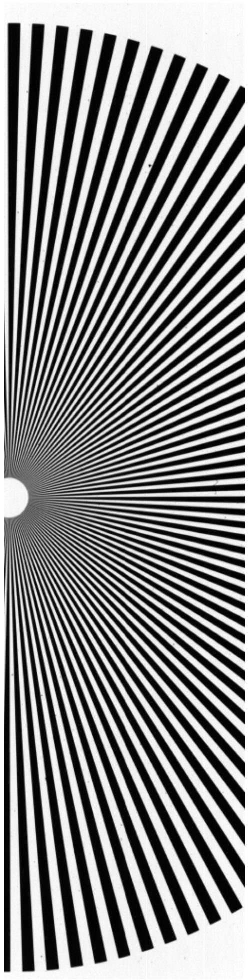
- Calculate cross-track and in-track MTF using a step response and impulse response example

- Results of on-orbit analysis give good agreement with the pre-launch laboratory measurements

VNIR Imaging of "Starburst Pattern"



Level 0 data



Level 1 data

Pattern used to validate optics performance and level 1 processing

MTF Example: Cross-track Bridge

Dec 24, 2000. Bridge is the Mid-bay bridge near Destin, Florida.

Bridge width (13.02 m) acquired and utilized in the MTF processing.

Bridge angle small, every 5th line used to develop high resolution bridge image.

MTF result at Nyquist is between 0.39 to 0.42; pre-flight measurement was 0.42.

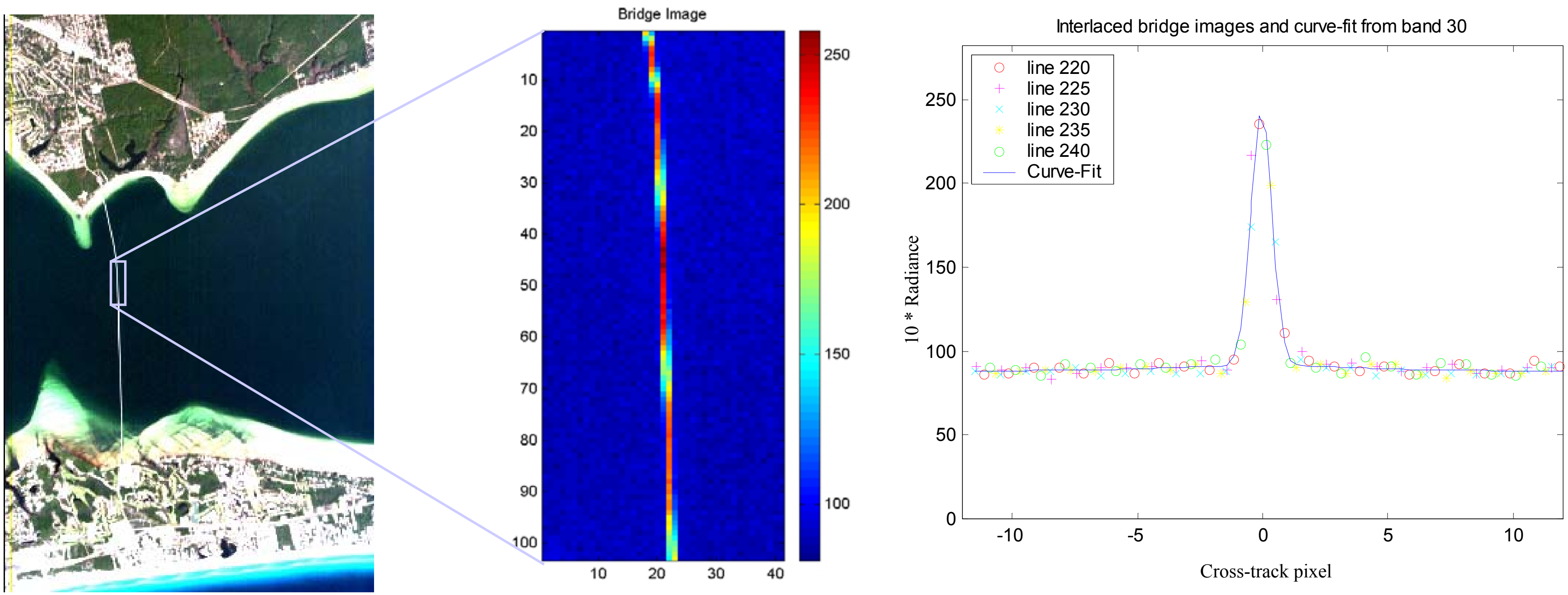
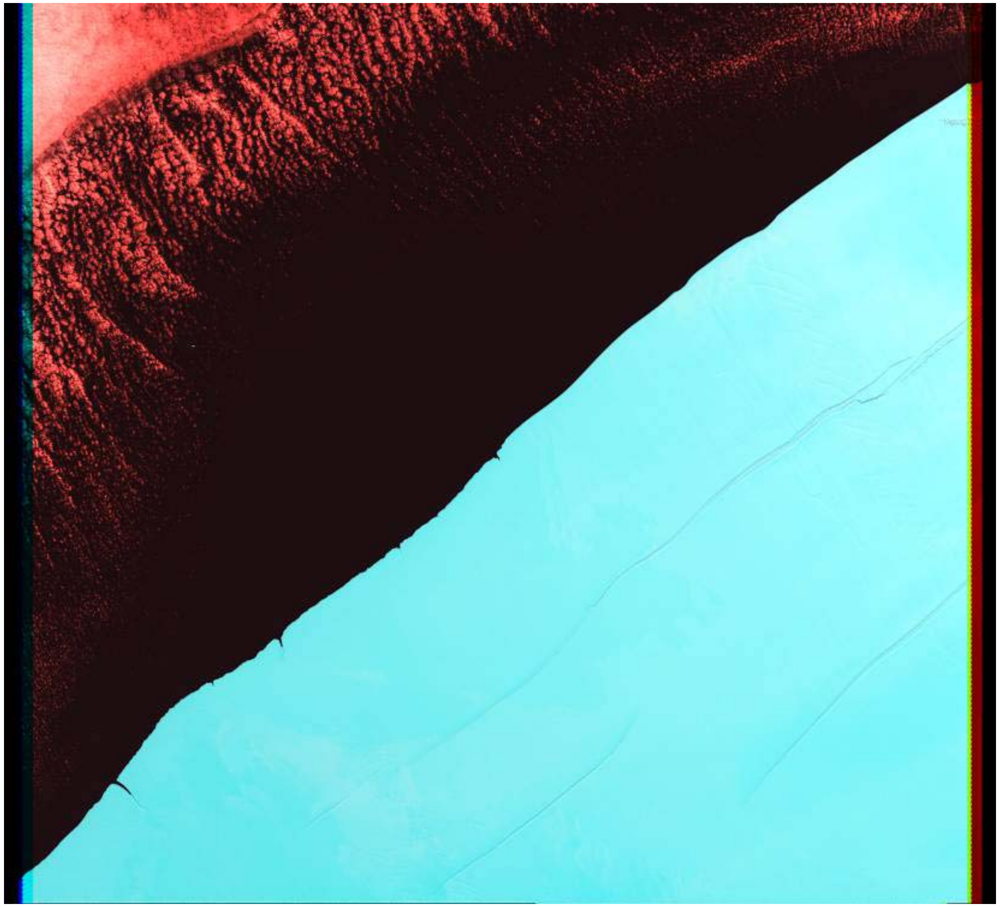
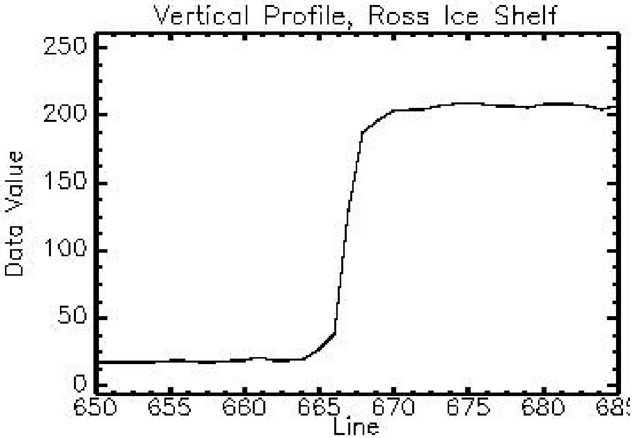
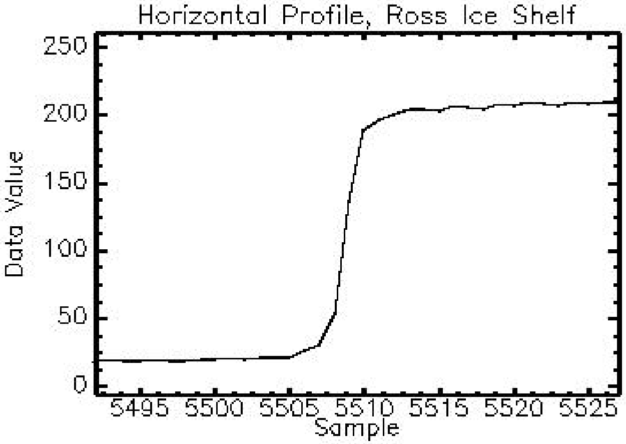


Image Quality Example

Vertical and Horizontal MTF can be calculated from diagonal edge







MTF Calculation Process Summary

- MTF measurement on-orbit requires collection of scenes
  - step response: Ross Ice Shelf, Lunar Collect
  - impulse response: Bridge (Eglin, Cape Canaveral)

- Data Processing steps (simplified):

1.) Define Edge Spread Function (ESF): Interlace adjacent lines from an object that is at a slight angle to the spacecraft motion

- allows over sampling (sub-pixel) sampling

2.) Calculate Line Spread Function (LSF):

- edge technique: *calculate an error function curve-fit to the ESF to derive the LSF, OR take the band-limited derivative of the ESF with a Tukey window.*

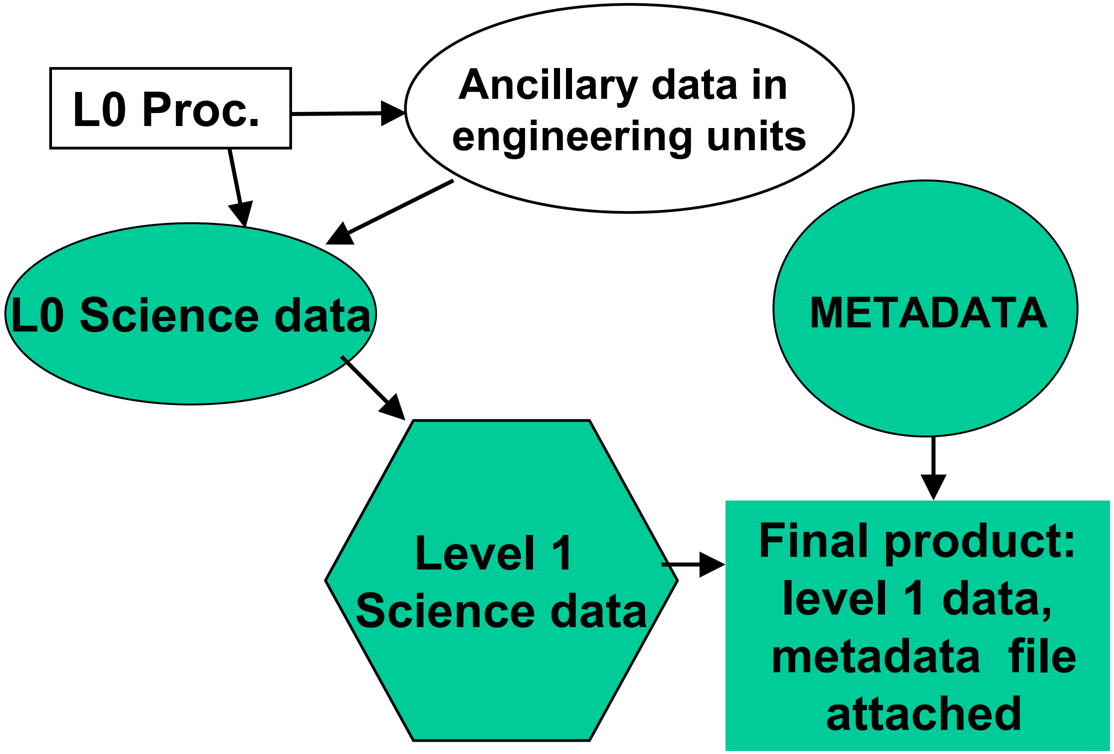
- slit technique: *LSF is obtained from interlacing adjacent lines and de-convolving the profile with the slit width (in the frequency domain)*

3.) MTF is the Fourier Transform of the LSF

What’s Next

- We have looked at instrument characteristics and calibration
- Now what do we do to “correct” the data?
  - Radiometric
  - Geometric
  - Atmospheric
  - Uniformity (striping)
  - Other nuances
- How well can all this work?

Hyperion Data Flow

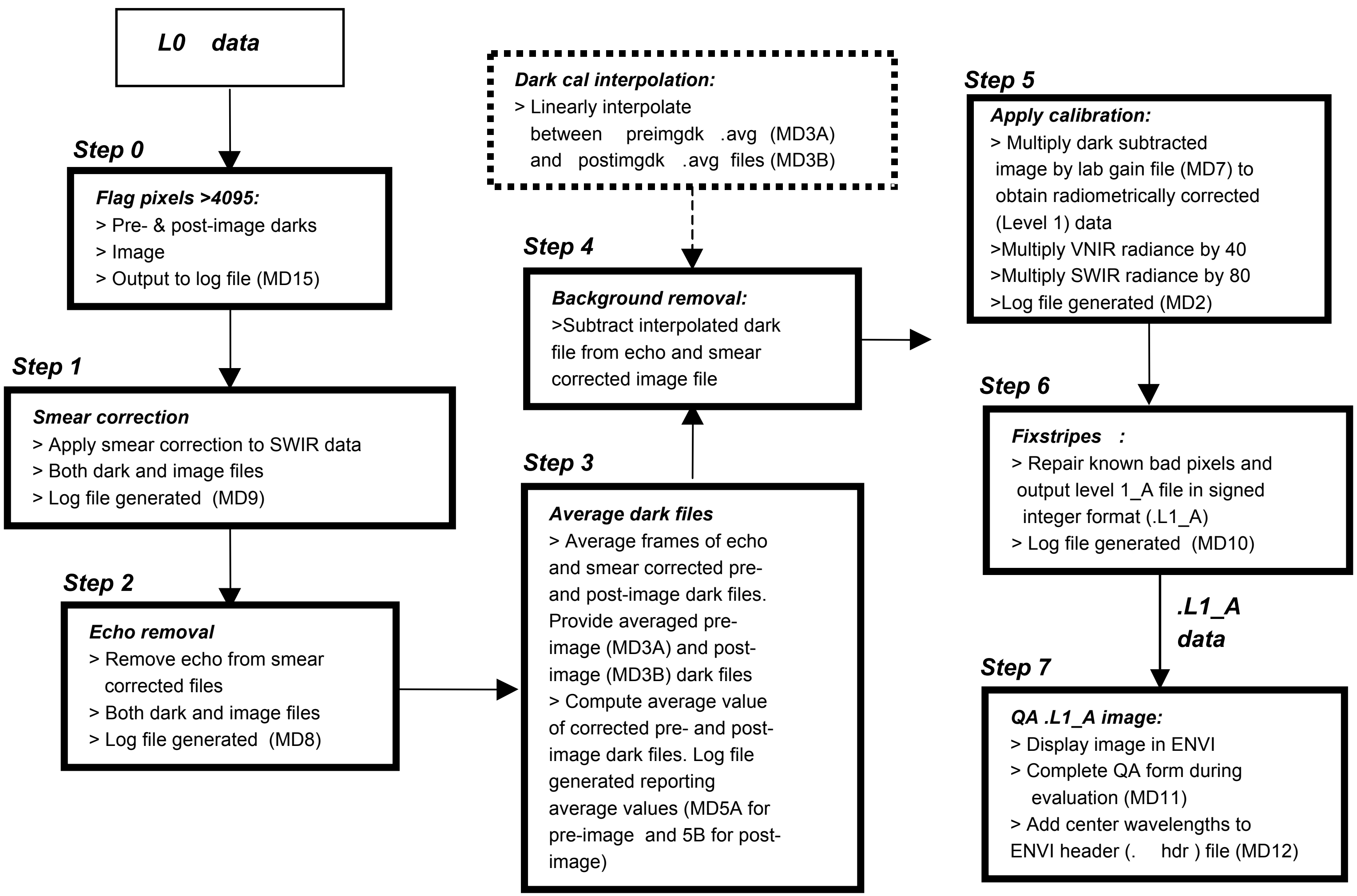


Science Data: Level 0 or Level 1 (radiometrically corrected) data products with VNIR and SWIR data frames combined. Includes solar, lunar calibrations, earth images, dark and light calibrations

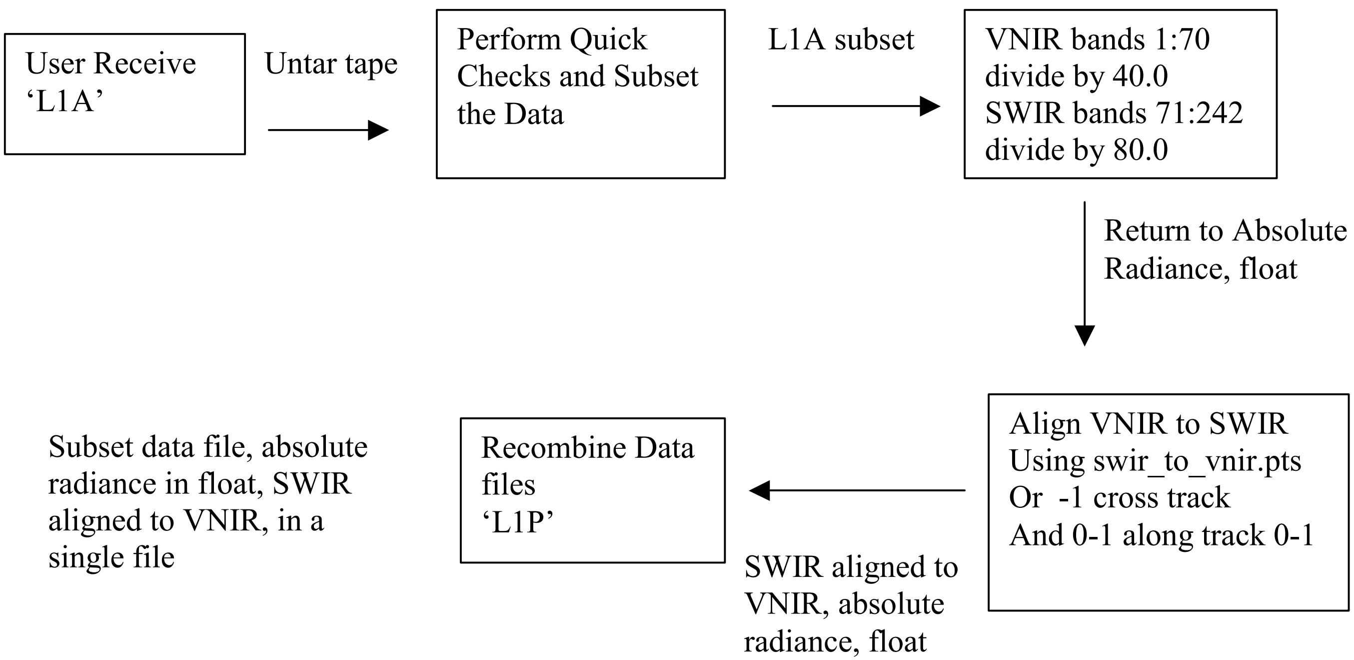
Metadata: Data about the science data. Information to support higher level processing, e.g., pre-flight characterization data

Ancillary Data: Supporting data derived from spacecraft telemetry during image collection

Level 1 Data Processing Flow



Hyperion Processing: One Step Further

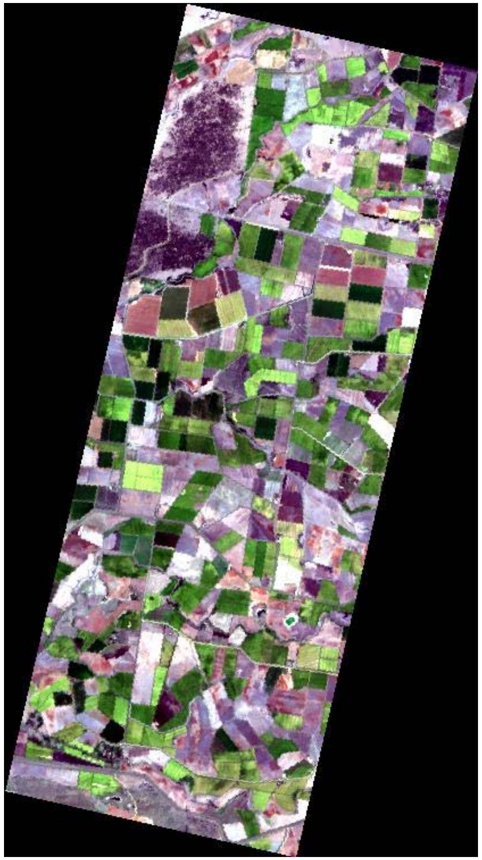


Data Processing Levels

Level 1A:

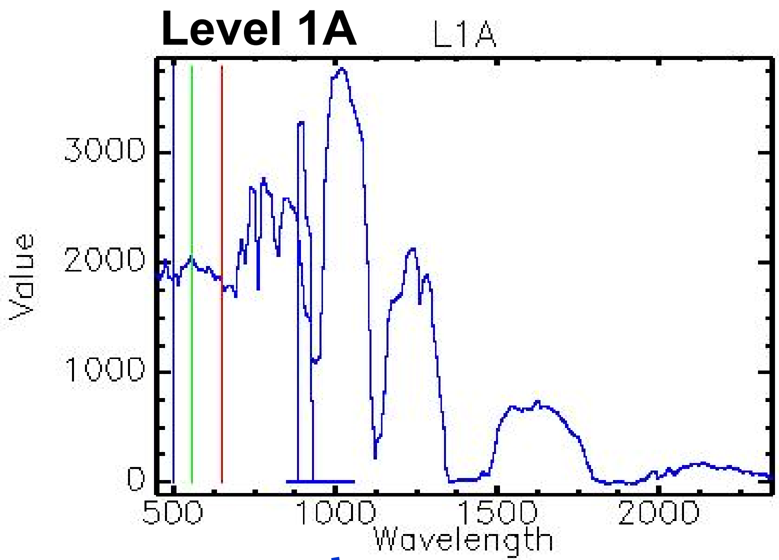
delivered data

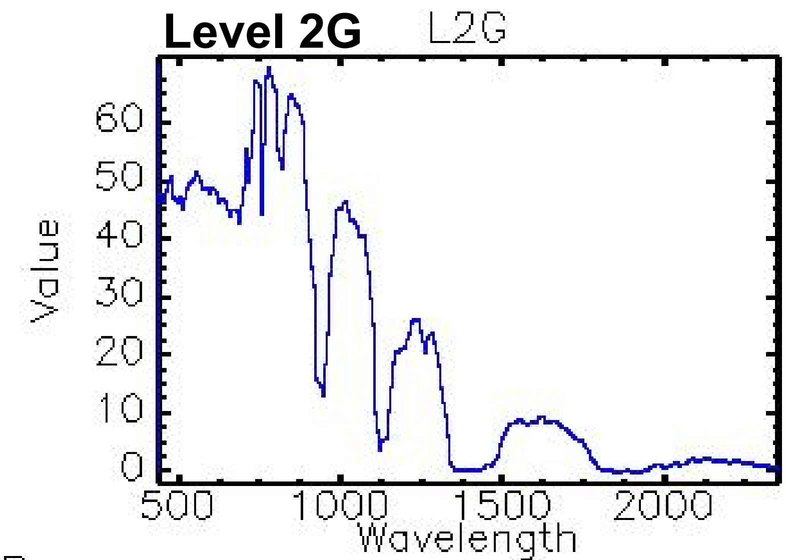
Level 1P:
Radiometric and VNIR/SWIR overlay



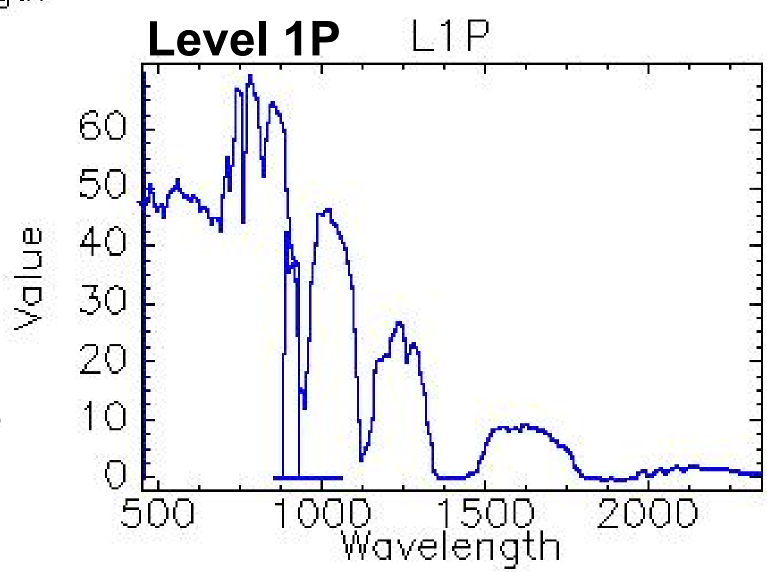
Level 2G:
Geometric and VNIR/SWIR merge

Data Evolution - Processing L1A –> L2G



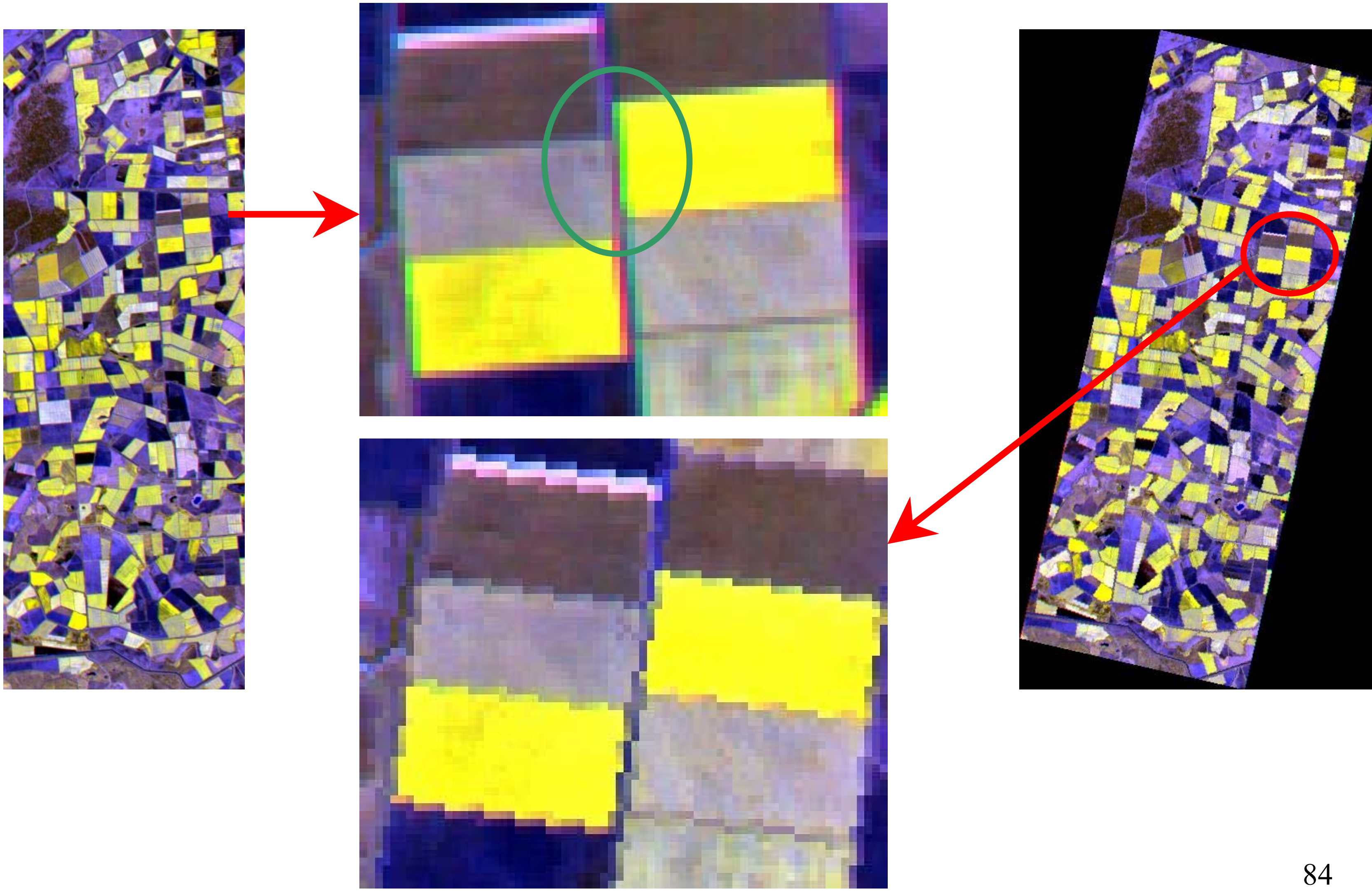


Level 1A: delivered data
Level 1P: Radiometric and VNIR/SWIR overlay
Level 2G: Geometric and VNIR/SWIR merge



VNIR/SWIR merge:
Bands 9-55 & 77-220

Geo-Correction: L1A → L2G



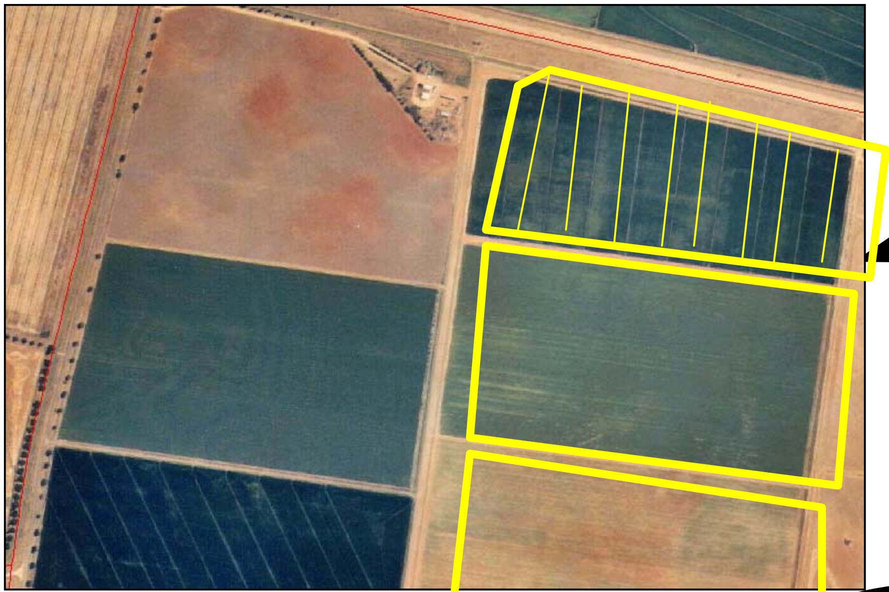
Methods / Geo-correction

A time series of polynomials allowing for geo-correction were developed by simultaneously fitting Ground Control Points (GCPs) collected in multiple images (date and sensor) using the MOSMOD module of microBRIAN. These polynomials allow the images to be resampled to geographic coordinates.

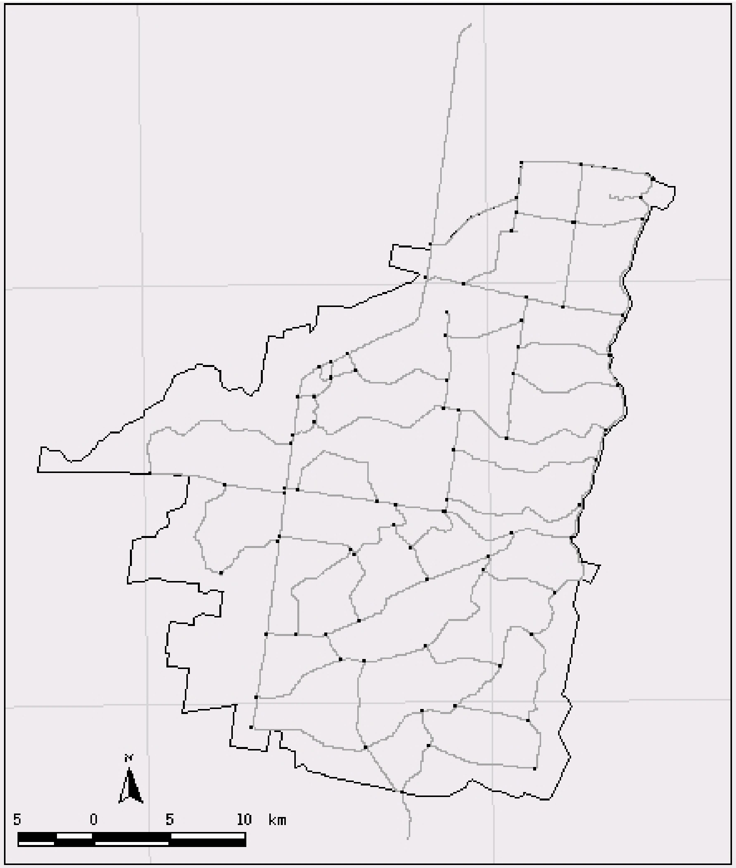
Fifty GCPs were collected between the base Aerial photography, an ETM image acquired 02-January-2001 (ETMjan), Hyperion VNIR, acquired 02-January-2001, 03-February-2001, 07-March-2001 (HypVNIRjan, HypVNIRfeb, HypVNIRmar), and Hyperion SWIR for those same dates (HypSWIRjan, HypSWIRfeb, HypSWIRmar).

MOSMOD both optimizes the images to geographic coordinates and the images internal registration.

GIS Data



High resolution (2m) digital aerial photographs acquired January 2001 used for creating positionally accurate field and rice bay GIS datasets.



Over 466km of linear road network and 129 well-defined points digitised with a Differential Global Positioning System (DGPS). These datasets can be used for geo-referencing Hyperion time series

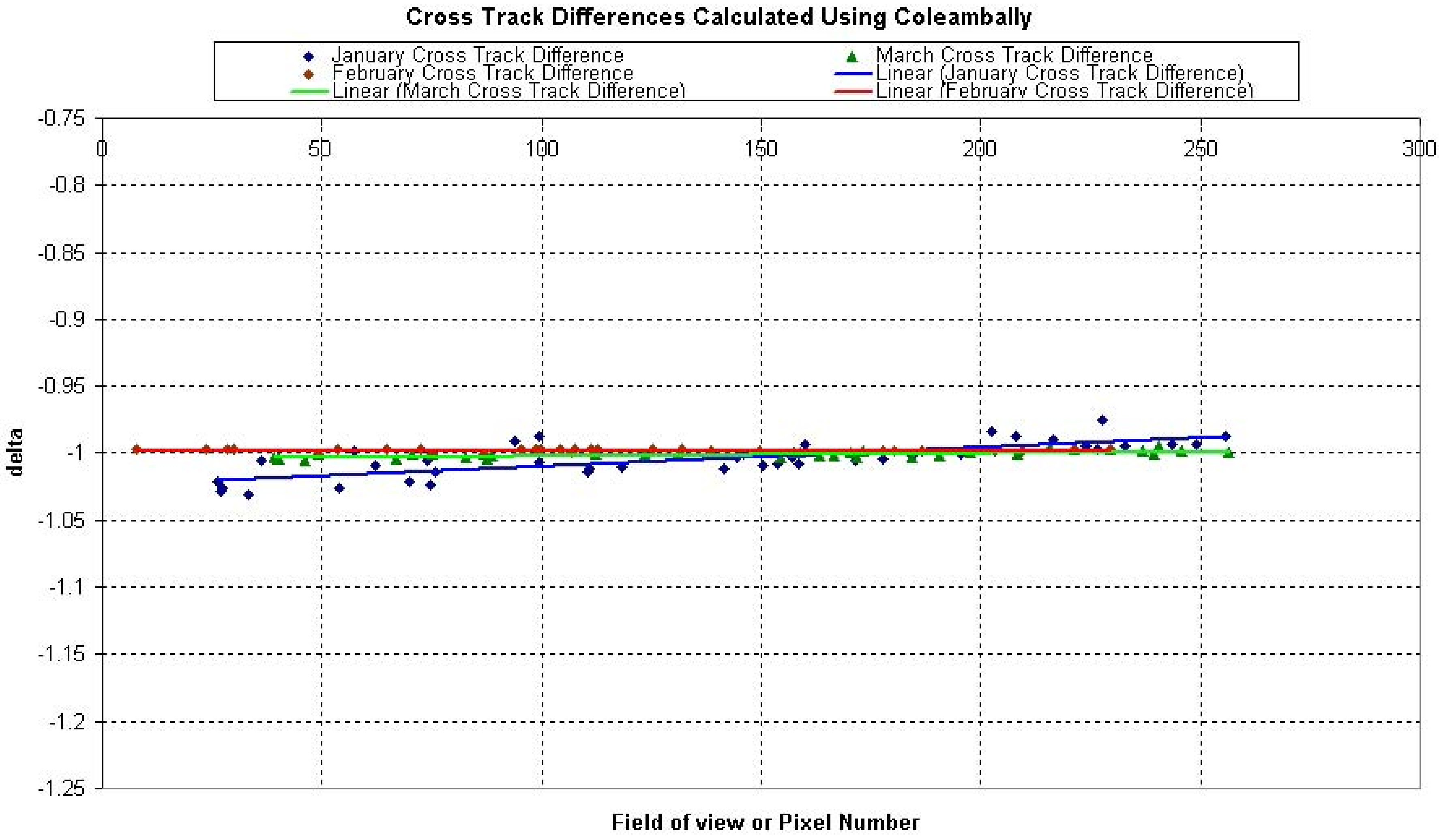
Coleambally Geo - Correction

- Six Hyperion ‘images’ (VNIR and SWIR for 02 Jan; 03 Feb; and 07 Mar 2001); GCPs collected in each Hyperion ‘image’ and 2m air photos

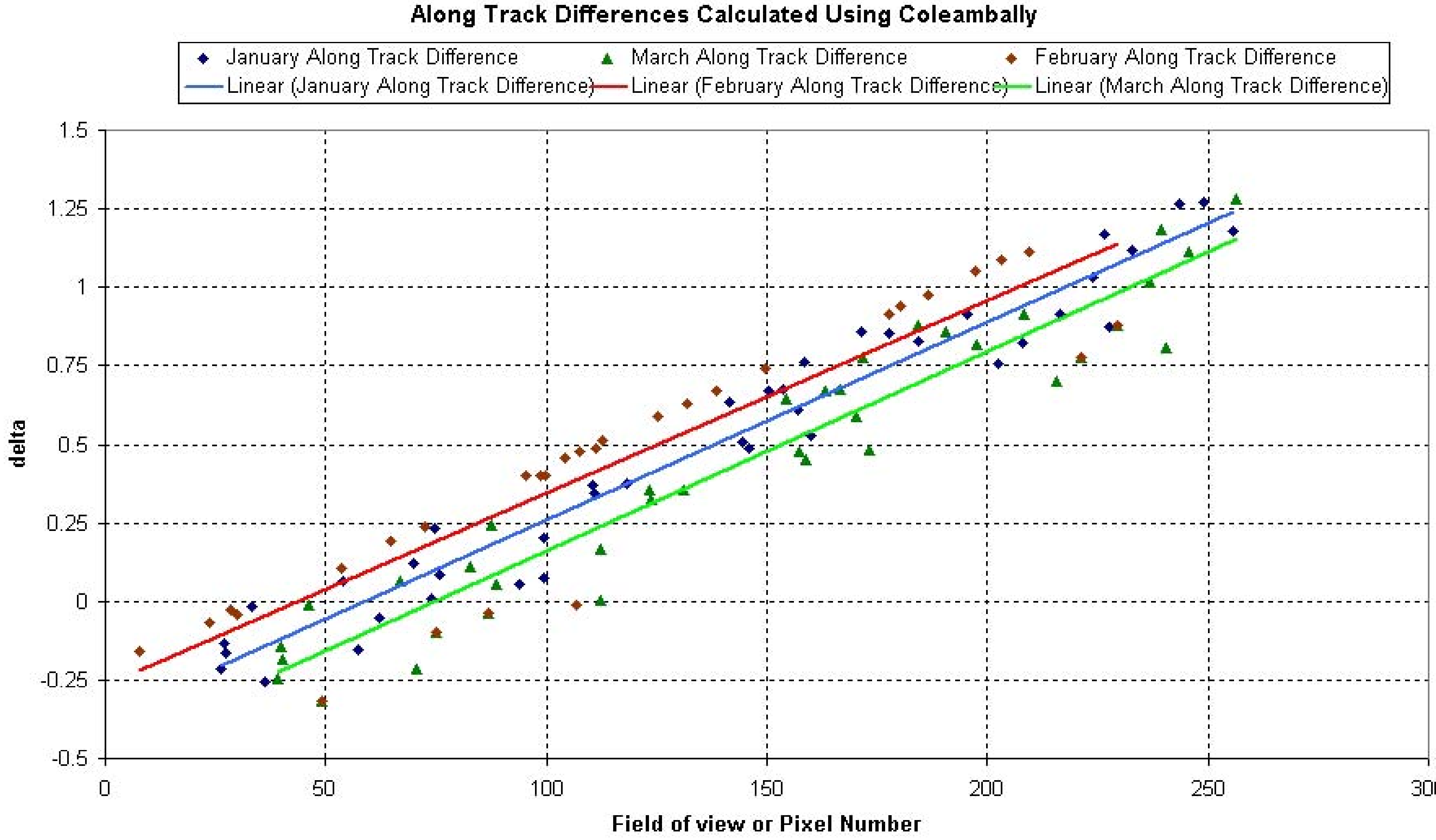
- Prior to fitting, outlier GCPs were identified and redefined; a linear polynomial was selected and a transitive chain method (MOSMOD) was used

Pixel Size
Avg X = 30.77m; Avg RMS error = 2.52 m
Avg Y = 30.49m; Avg RMS error = 6.87 m
Registration
Cross-track 12.9m (SD 0.6m)
Along-track 11.6m (SD 2.1m)

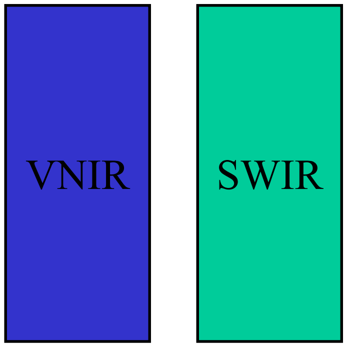
Preliminary Results / Geo-correction



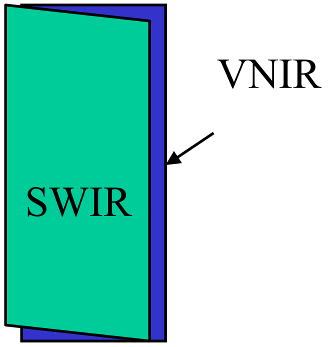
Preliminary Results / Geo-correction



Aligning the VNIR and SWIR



- VNIR and SWIR are independent spectrometers and focal plane arrays

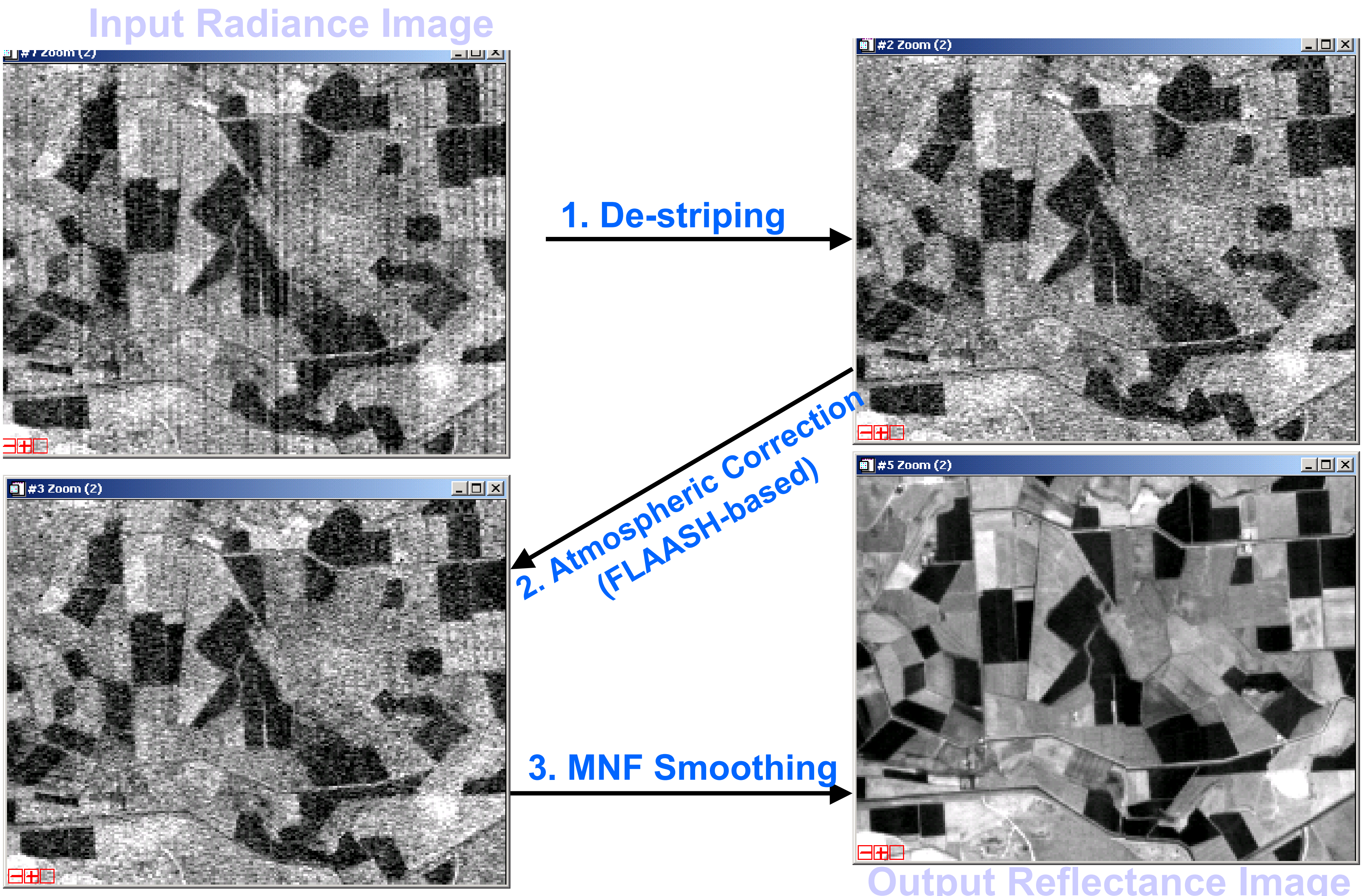


- For the user to align the SWIR to the VNIR apply

  - constant cross track shift of –1 pixel

  - linear along track shift of 0 pixels at field-of-view 0 and 1 pixel at field of view 256.

Reflectance Processing
Hyperion Band 223 (2385 nm)



Should Atmospheric Correction be Performed?

- Hyperspectral remote sensing has effects of atmosphere as well as earth materials
- The result is easily recognized and its worst effects can be avoided
- Should you atmospherically correct?
  - Some applications do not need it (classification, MNF, CVA, exploration)
  - Some can benefit from it (standardization)
  - Some need it (modeling)
- The data are there to use – do not be put off nor wait until all the “problems” have been solved!

A Simple Atmospheric Model

Ground Reflectance

Irradiance

Path Radiance

At-Sensor Radiance

$$L_t = \frac{1}{\pi} E_g T \frac{\rho_t}{1 - s\rho_t} + L_p$$

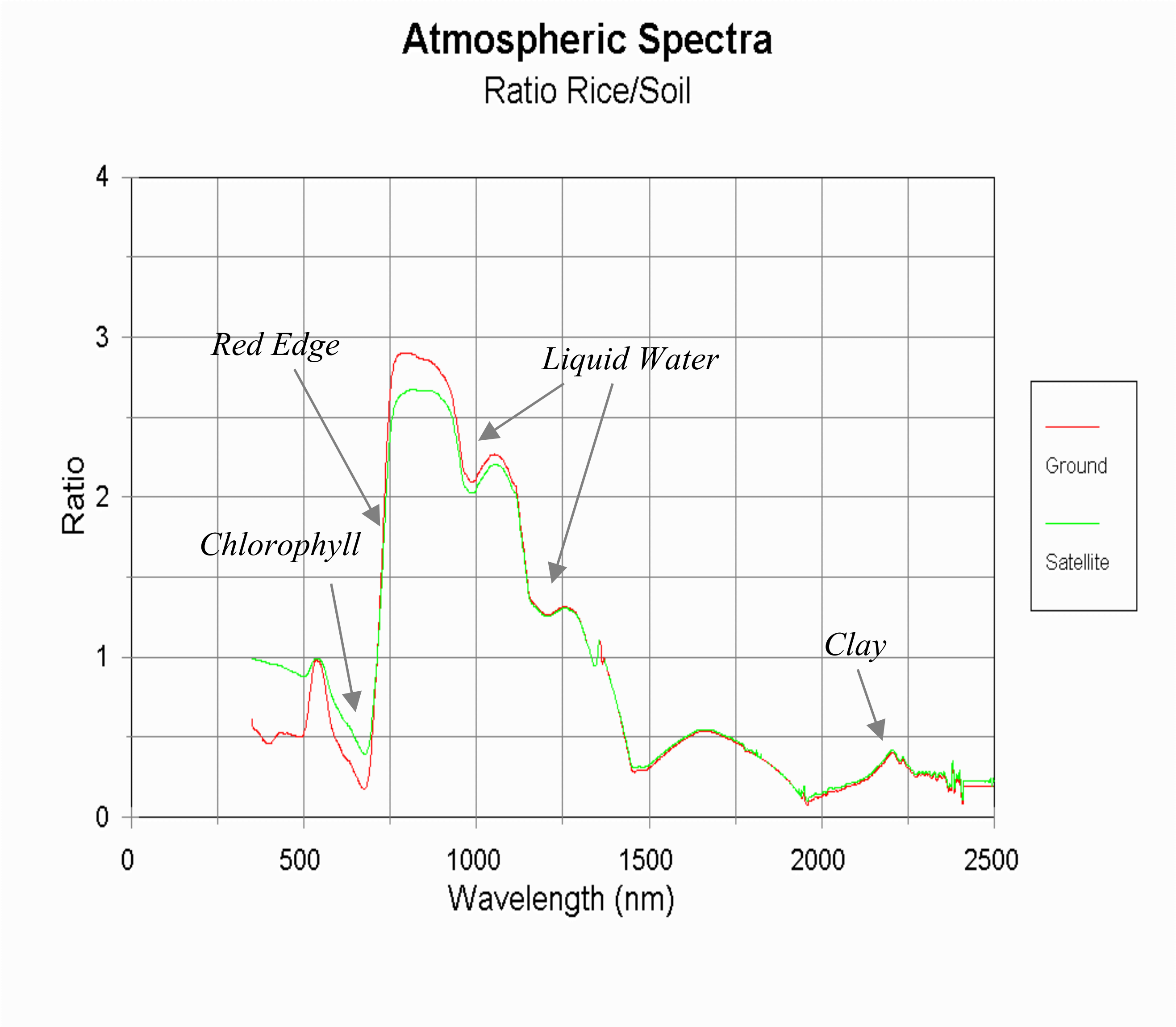
Atmosphere Reflectance

Transmittance

$$Modelling: \qquad \rho_t \rightarrow L_t$$

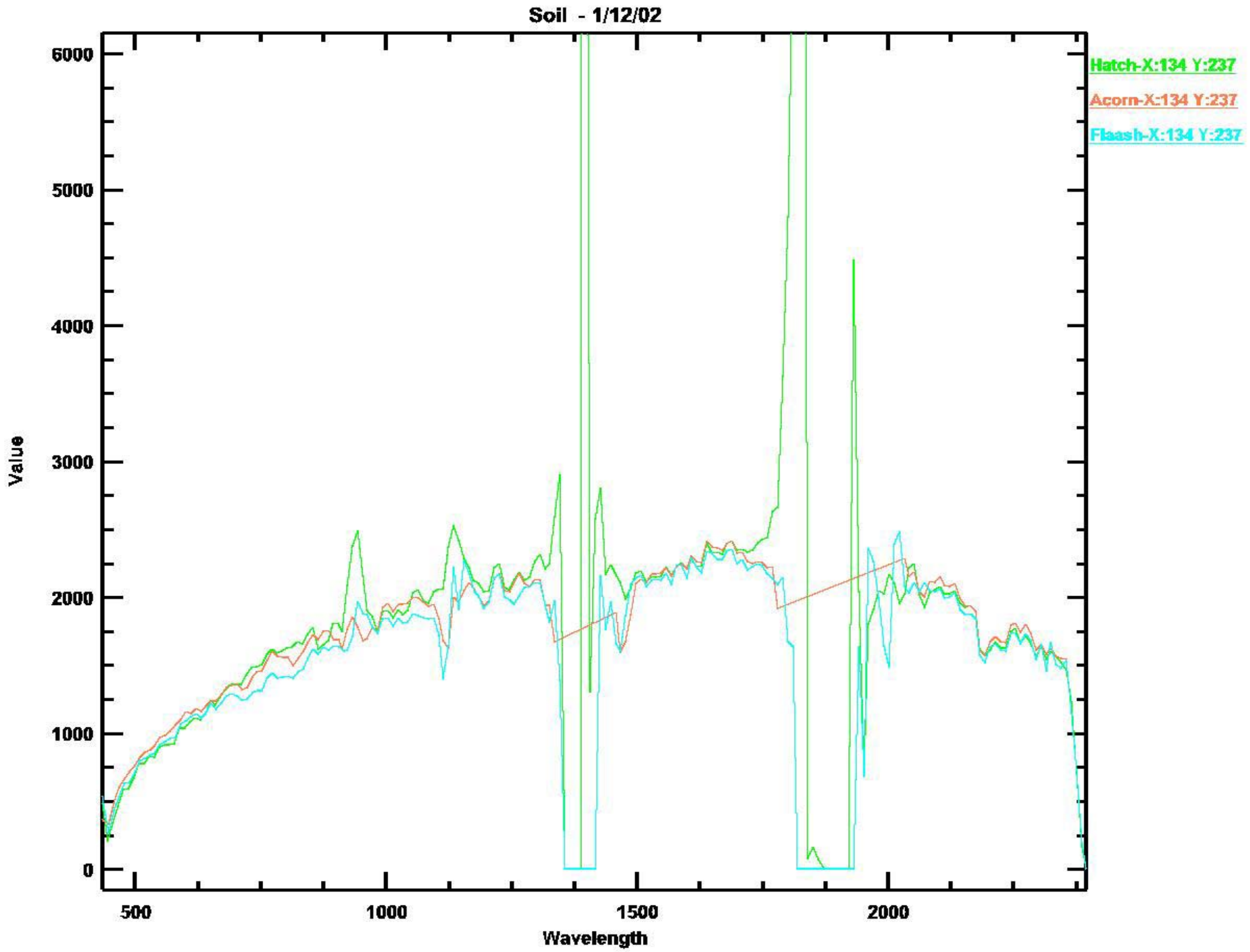
$$Atmospheric\ Correction: \qquad L_t \rightarrow \rho_t$$

Not so frightening after all?! (DLBJ)



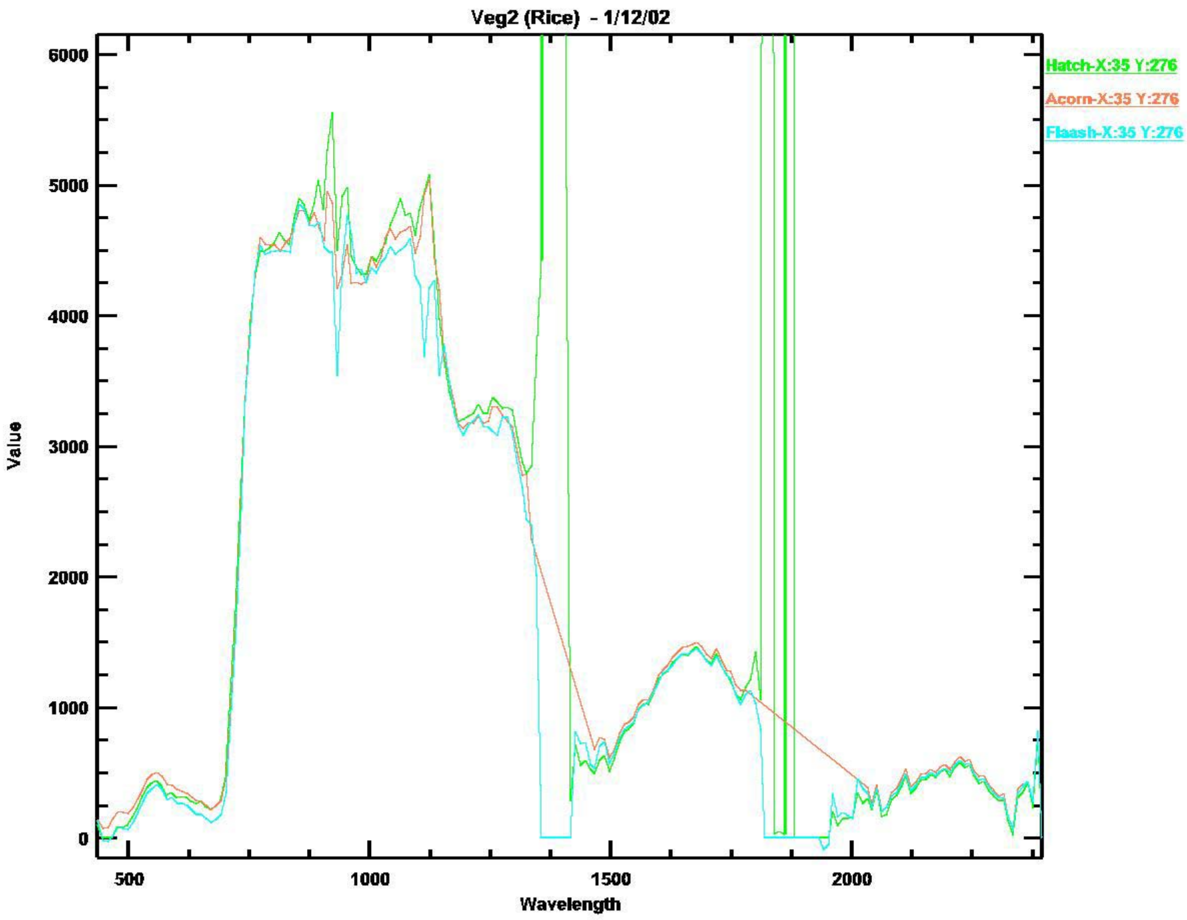
Atmospheric Correction Codes

Hatch, FLAASH, ACORN



Atmospheric Correction Codes

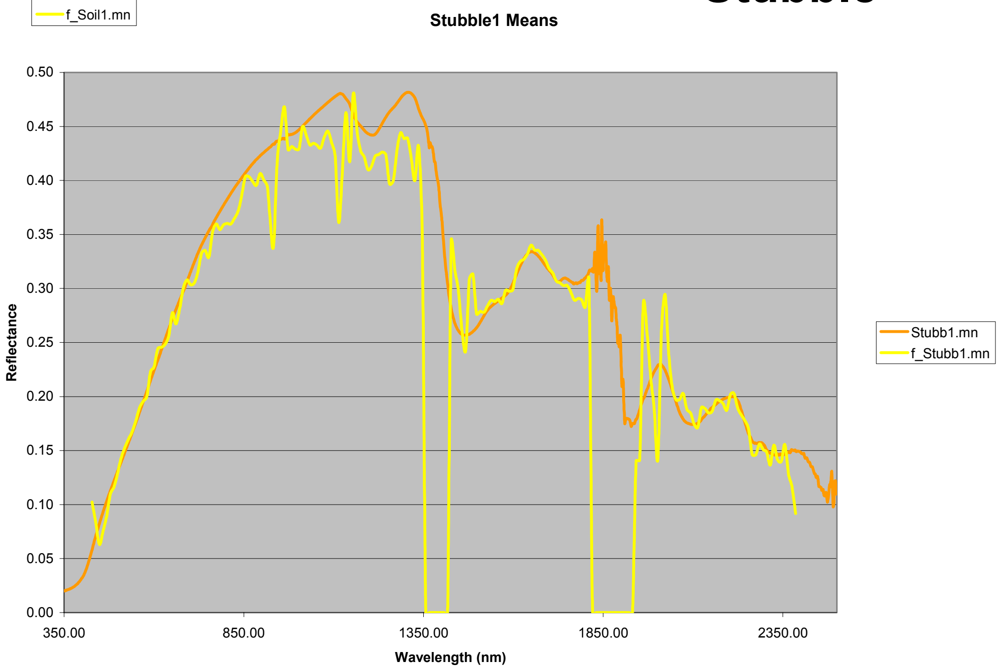
Hatch, FLAASH, ACORN



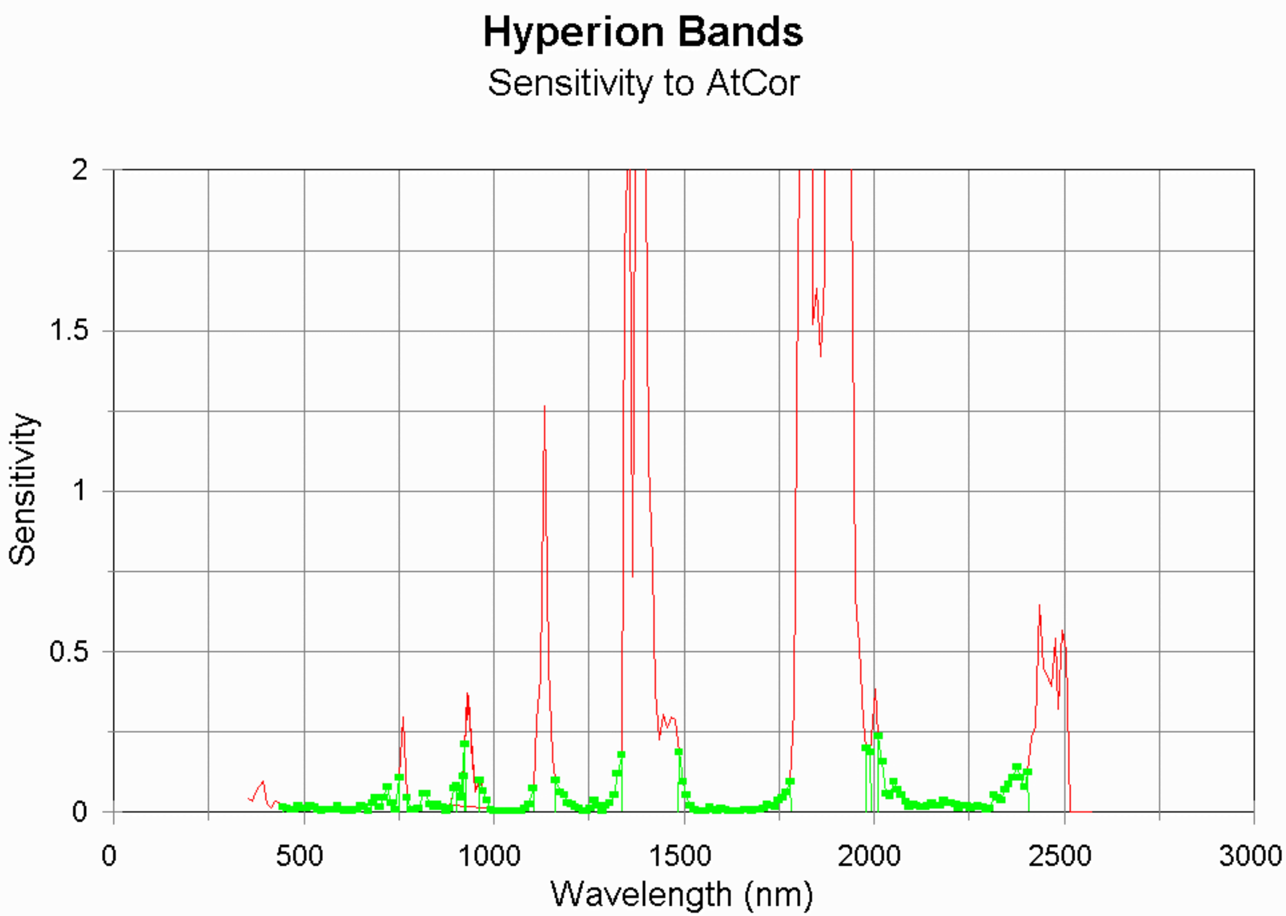
FLAASH – ASD Comparison

Soil

Stubble



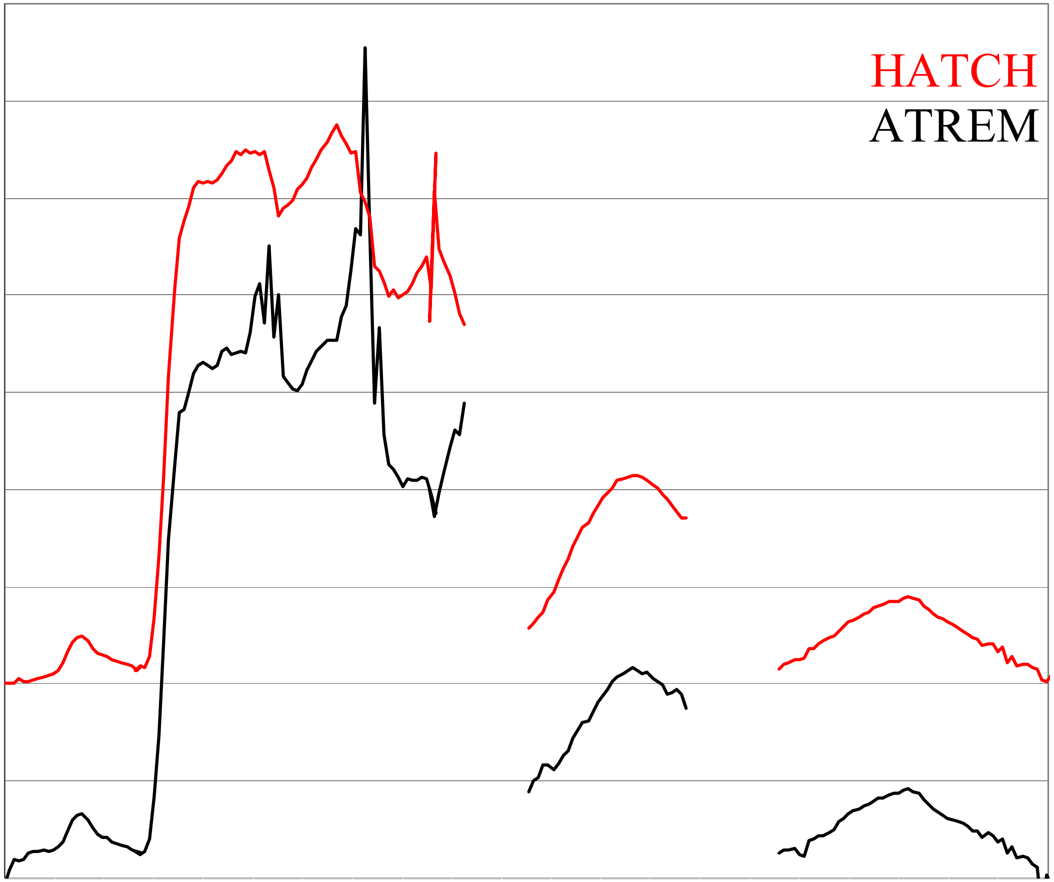
Atmospheric Avoidance



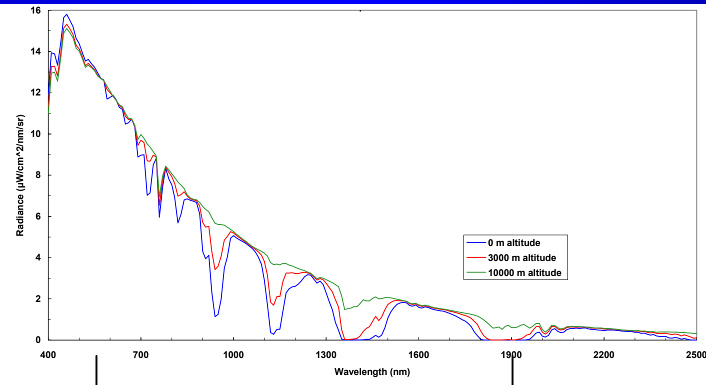
Bands	Wavelength (nm)
10-57	448-926
81-97	953-1114
101-119	1155-1336
134-164	1488-1790
182-221	1972-2365

"Stable" set of bands can be considered if application permits

Yerington, NV: pixel 62, 35



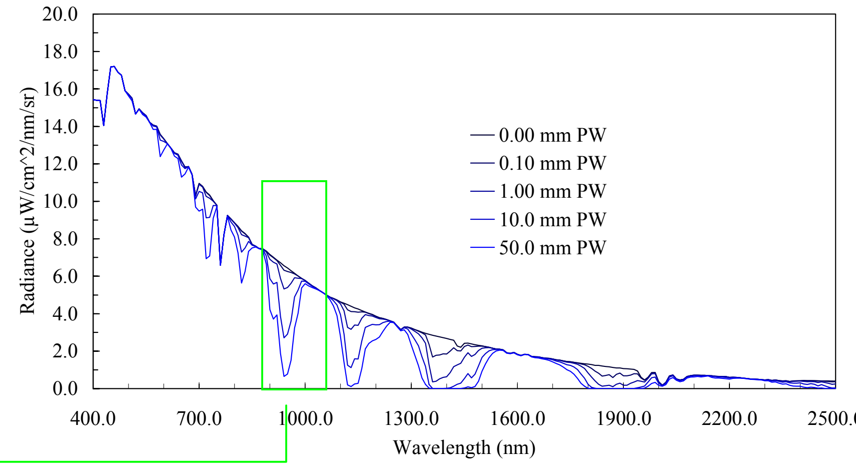
Cirrus Cloud Detection Over Mojave Desert



Visible Image

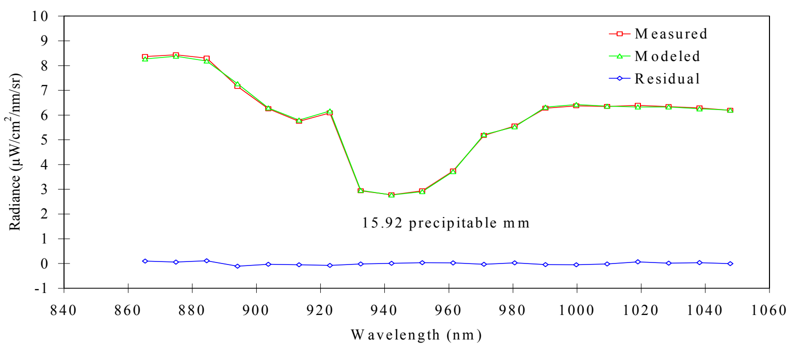
Image from 1380 nm

Accounting For Water Vapor

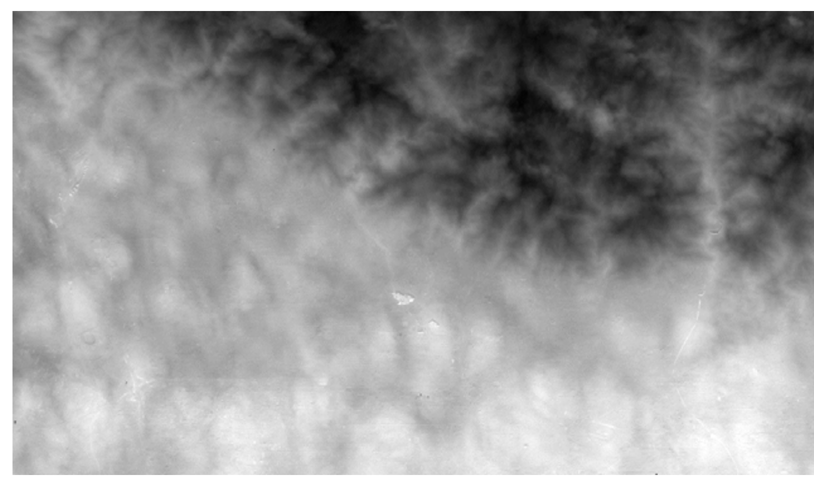


Modeled

Measured

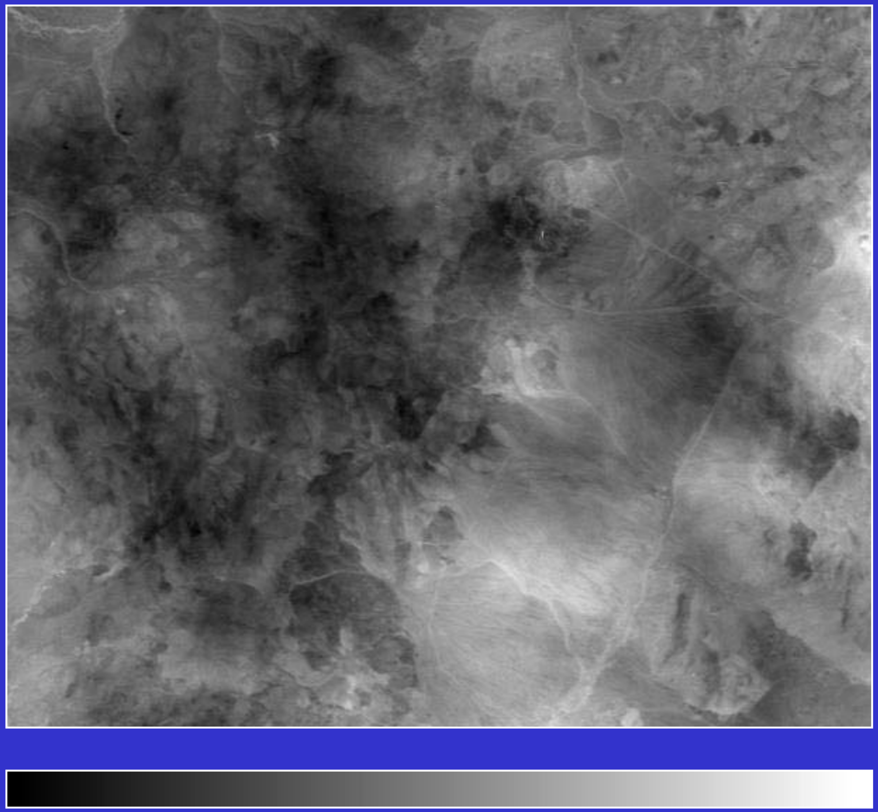


Water Vapor
Parameter map

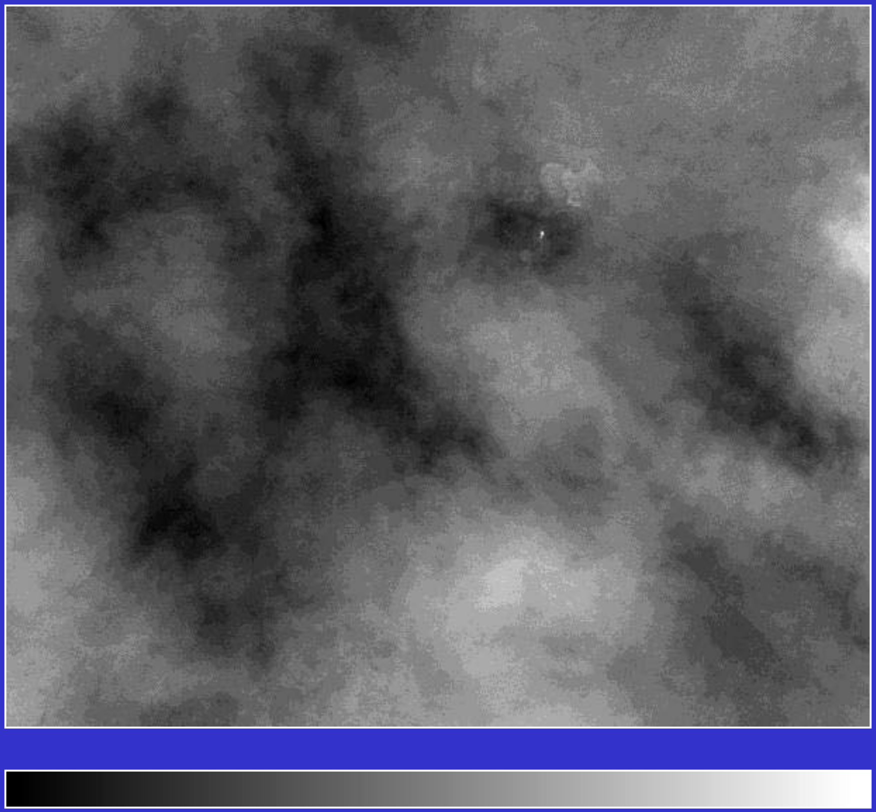


Yerington, NV: Water Vapor Images

ATREM

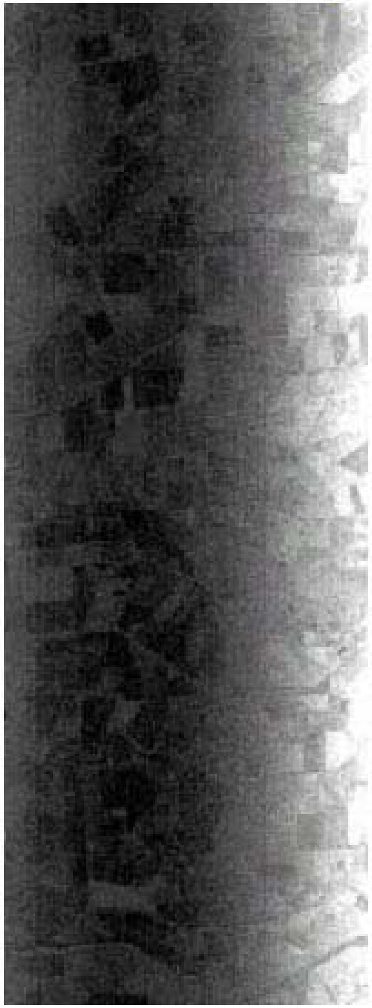


HATCH

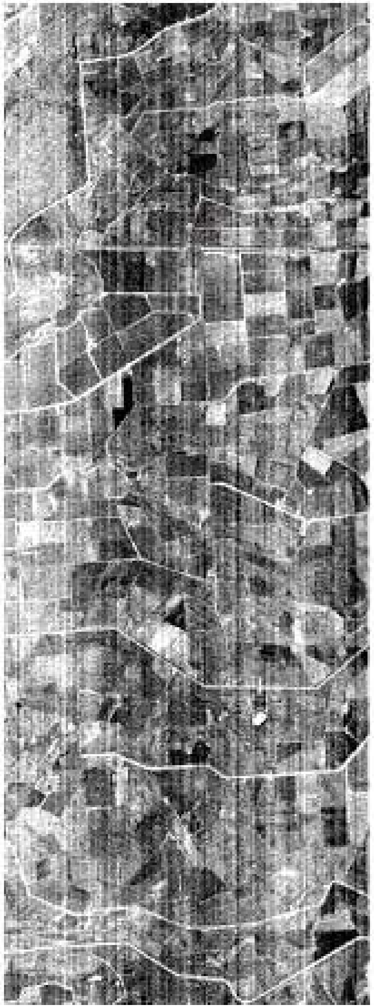


Stripes in the Image

RGB (42, 21,15)



MNF 1 Radiance



MNF 15 Radiance

- Stripes may not impact analyses
- Three spatial frequency scales occur:
    - High
    - Medium
    - Low
- Is there a temporal effect?

Approaches for De-Striping

1. Set a column mean (intensity) equal to the global mean

2. Set the column mean and standard deviation equal to the global mean and standard deviation

3. Use a locally-based mean and standard deviation in spatial dimensions

4. Use a locally-based mean and standard deviation in spatial *and* spectral dimensions

Effects of ‘Global’ and ‘Local’ De-Striping

- Typical spatial scales are pixel, surface features and swath

- ‘Global’ removes streaks and low-frequency effects (but ‘Global’ can bias the results) – it depends on the surface cover characteristics along the column

- ‘Local’ removes spikes – especially in the VNIR but leaves in the low frequency effects such as the smile radiance effect.

- VNIR and SWIR noise are different. Should be treated separately.

Striping

VNIR

SWIR

De-streaking Approach

(m_{ij}, s_{ij}) *are column averages for sample i band j*

$(\overline{m}_{ij}, \overline{s}_{ij})$ *are reference values for them*

$$x_{ijk} \rightarrow \alpha_{ij} x_{ijk} + \beta_{ij} \qquad \textit{sample i band j line k}$$

$$\alpha_{ij} = \frac{\overline{s}_{ij}}{s_{ij}} \qquad \beta_{ij} = \overline{m}_{ij} - \alpha_{ij} m_{ij}$$

Selection of reference values determines whether the de-striping is local or global

Global De-striping MNF1 and 15

Original Data

MNF 1 Radiance

MNF 15 Radiance

De-Streaked Data

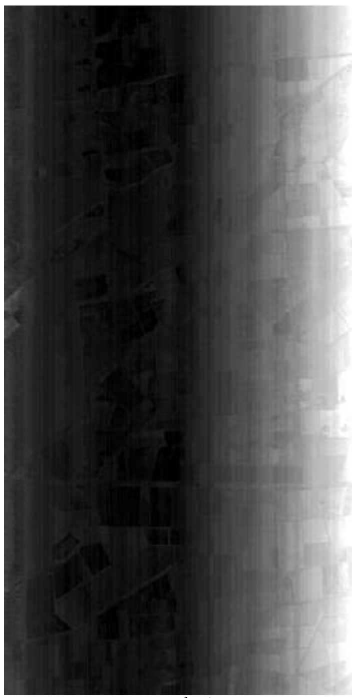
MNF 1 Radiance

MNF 15 Radiance

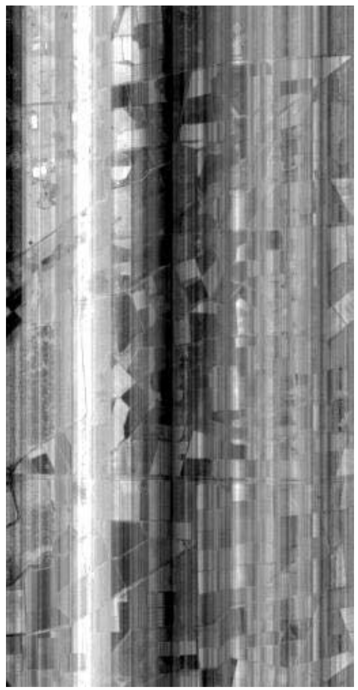
Effects of ‘Global’ and ‘Local’ De-Striping

Change Analysis

•Global De-streaked minus Original Radiance

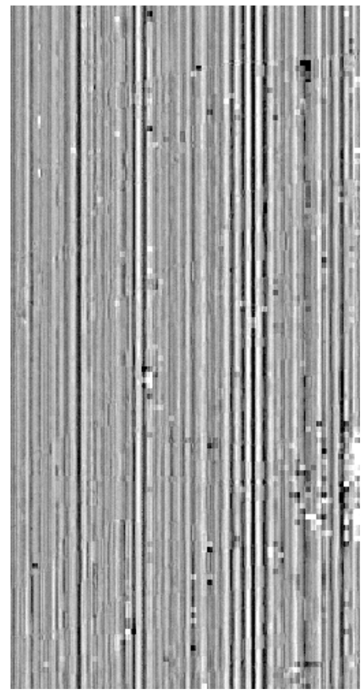


•MNF Band 1

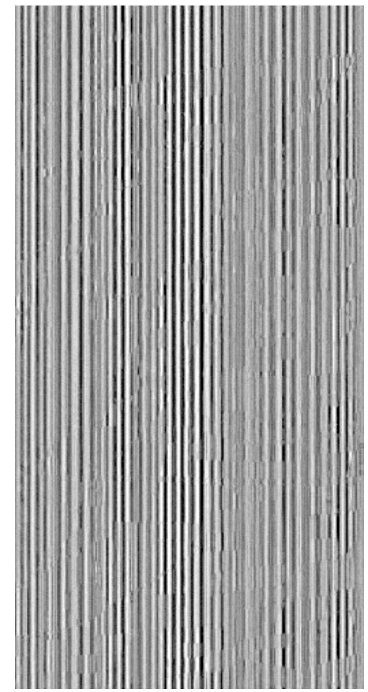


MNF Band 5

Local De-streaked minus Original Radiance



MNF Band 1



MNF Band 5

•Global De-streaking removes ‘smile’ effect and streaks

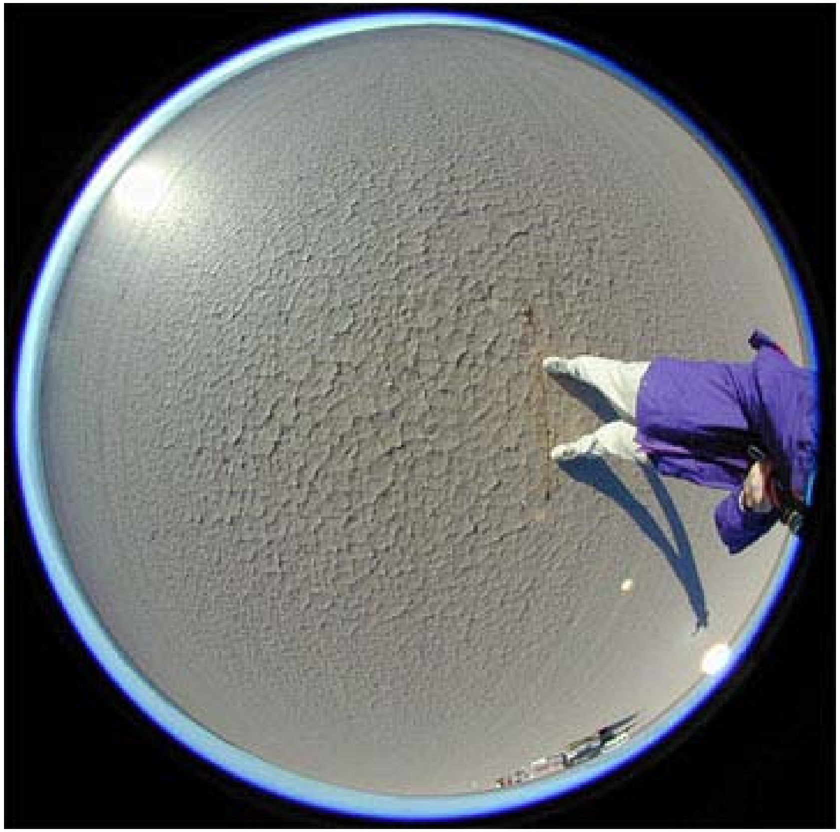
•But also alters mid frequency spatial effects in the data

Local de-streaking removes stripes but leaves the ‘smile’ effect in

Other Thoughts

Other Effects - Directional Illumination

Midday



Early Morning

Through the Glass Darkly – or “no free lunch”

- Hyperspectral data is fighting for photons
- Small pixels & many narrow bands lead to lower SNR
- The technology is good but there is a limit

- Landsat FWHM ranges from 70 in VNIR to 300-400 in the SWIR

- Hyperion Bands have an FWHM and step of 10 nm

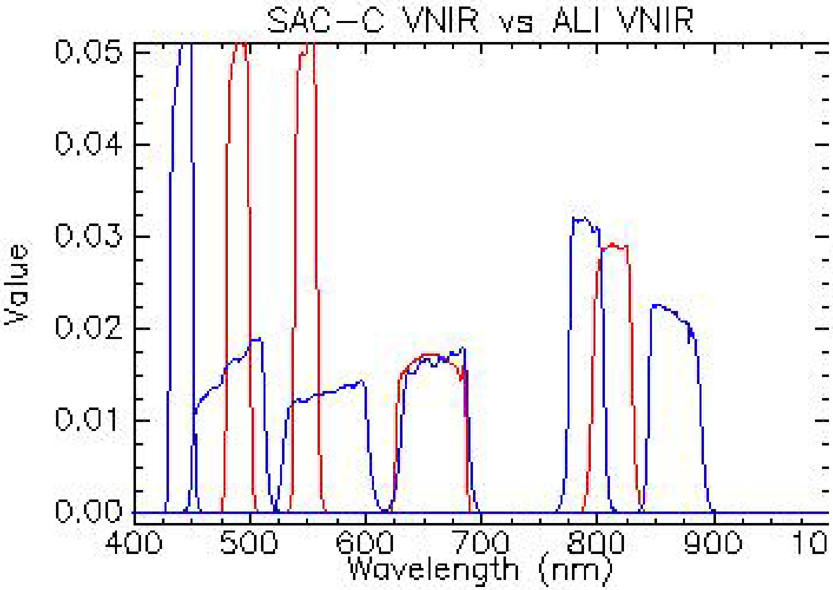
Sensor Characteristics

SAC-C ETM+ ALI MODIS

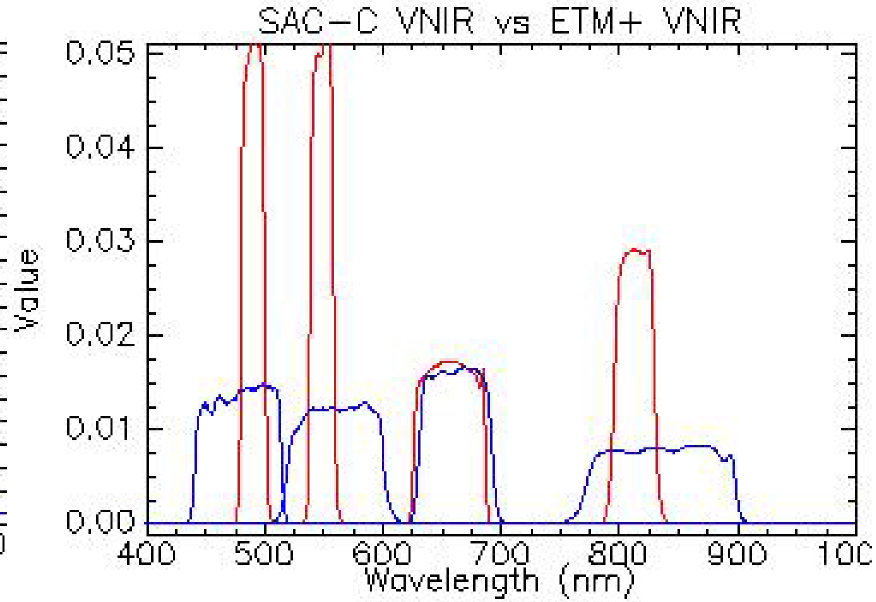
SAC-C	CWL (nm)	FWHM (nm)	ETM	CWL (nm)	FWHM (nm)	ALI	CWL (nm)	FWHM (nm)	MODIS Land	CWL (nm)	FWHM (nm)
						Band 1p	441.6	18.9			
Band 1	490.5	19.3	Band 1	478.7	67.2	Band 1p	484.8	52.6	Band 3	465.7	17.6
Band 2	548.8	19.2	Band 2	561.0	77.6	Band 2	567.2	69.7	Band 4	553.7	19.7
Band 3	656.6	58.0	Band 3	661.4	60.0	Band 3	660.0	55.8	Band 1	646.3	41.9
Band 4	812.9	34.2	Band 4	834.6	120.8	Band 4	790.0	31.0	Band 2	856.5	39.2
						Band 4p	865.6	44.0			
						Band 5p	1244.4	88.1	Band 5	1242.3	24.7
Band 5	1598.5	106.6	Band 5	1650.2	190.5	Band 5	1640.1	171.2	Band 6	1629.4	29.7
			Band 7	2208.1	251.3	Band 7	2225.7	272.5	Band 7	2114.2	52.2
			Pan	719.9	319.5	Pan	591.6	144.3			

SAC-C VNIR Bands vs ALI, ETM+

ALI

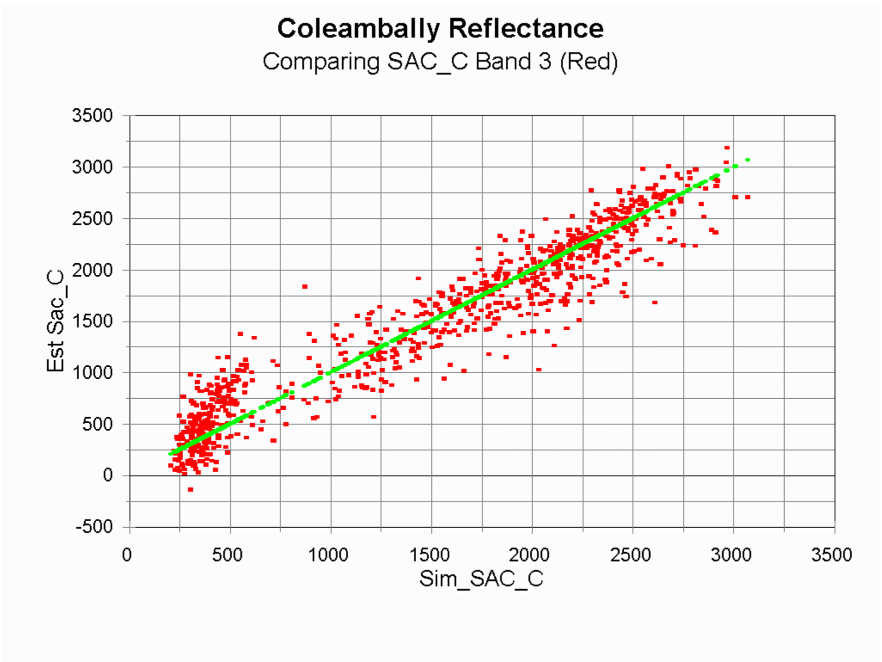


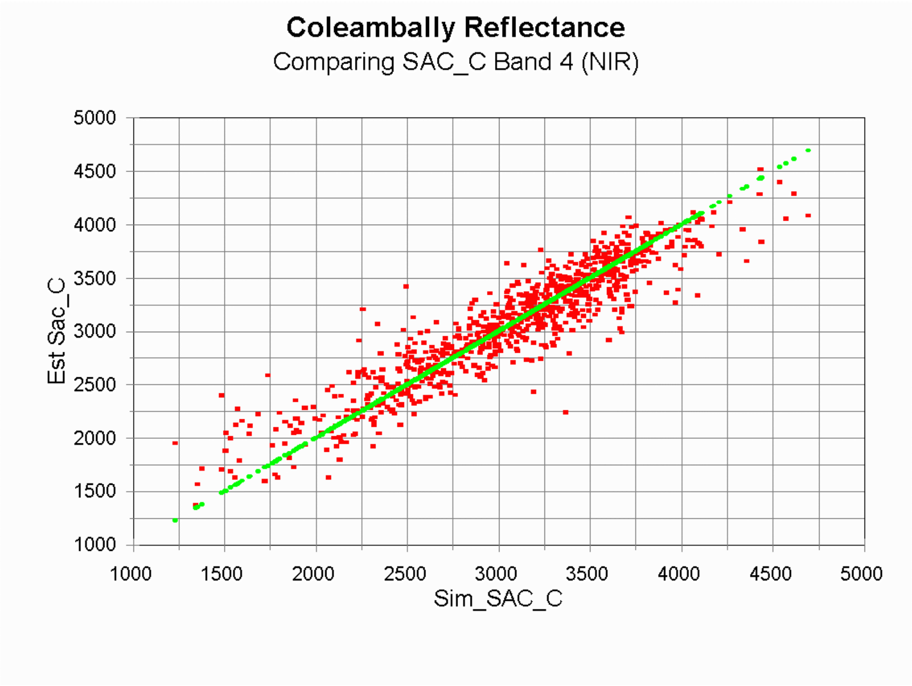
ETM+



Red: SAC-C
Blue: Other Instrument

Binned Hyperion vs SAC-C



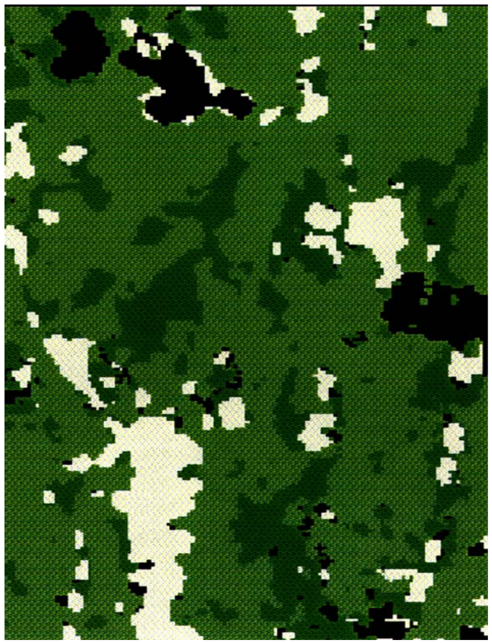


Hyperspectral Imaging Applications & Benefits

Application	Existing Satellite Capabilities (SPOT, LandSat)	Hyperion Capability	Perceived Benefits
Mining/Geology	Land cover classification	Detailed mineral mapping	Accurate remote mineral exploration
Forestry	Land cover classification	Species ID Detail stand mapping Foliar chemistry Tree stress	Forest health/infestations Forest productivity/yield analysis Forest inventory/harvest planning
Agriculture	Land cover classification Limited crop mapping Soil mapping	Crop differentiation Crop stress	Yield prediction/commodities crop health/vigor
Environmental Management	Resource meeting Land use monitoring	Chemical/mineral mapping & analysis	Contaminant Mapping Vegetation Stress

Hyperspectral Image Provides Forestry Detail

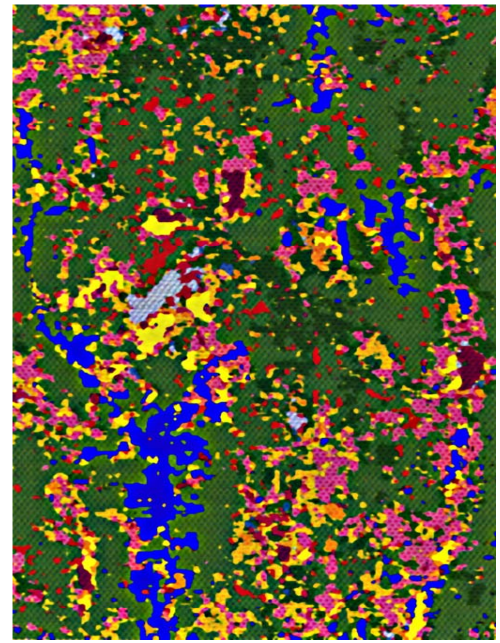
LandSat Analysis



Legend

- No Data
- Hardwood
- Softwood
- Grass / Fields

Hyperspectral Analysis



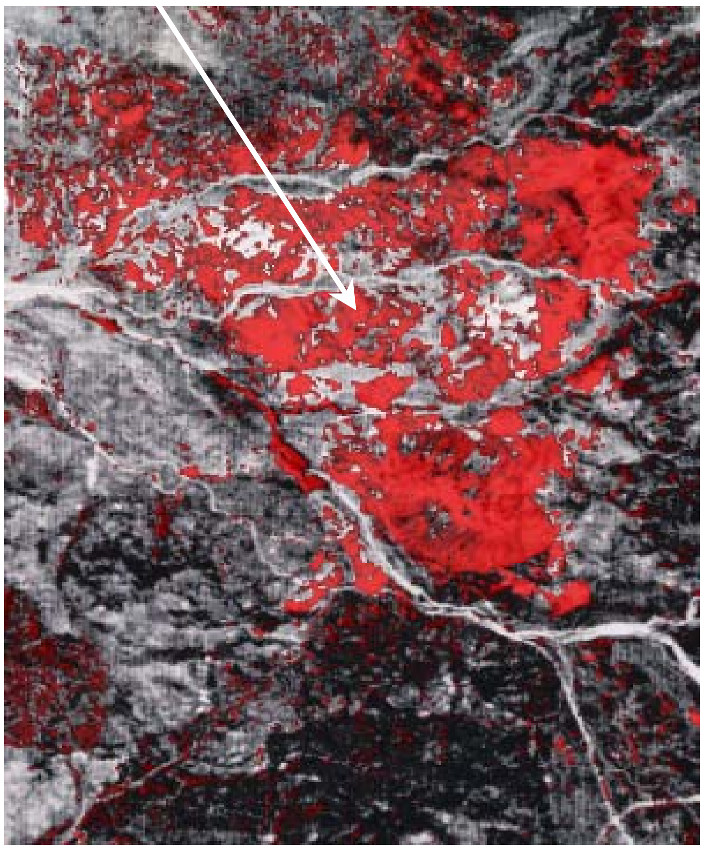
Legend

- No Data Open field
- Red Maple
- Red Oak
- Mixed Hardwood
- Hardwood/ Conifer Mix
- White Pine
- Hemlock/ Hardwood Mix
- Mixed Conifer
- Norway Spruce
- Red Pine
- Spruce Swamp
- Hardwood Bog

Analysis by Mary Martin University of New Hampshire

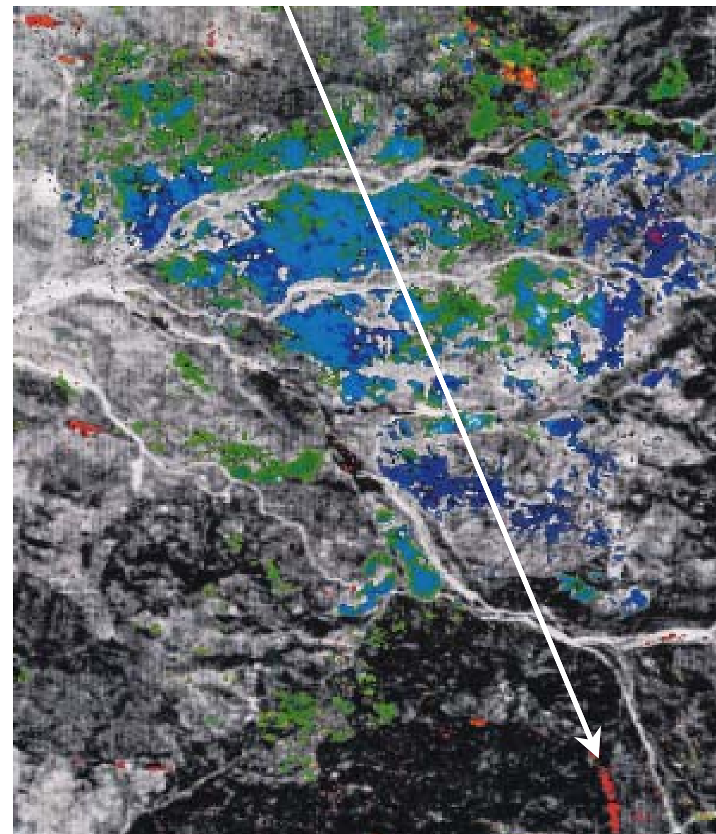
Hyperspectral Image Provides Geological Data

GEOTHERMAL AREA
(no specific mineral information)



MULTISPECTRAL ANALYSIS

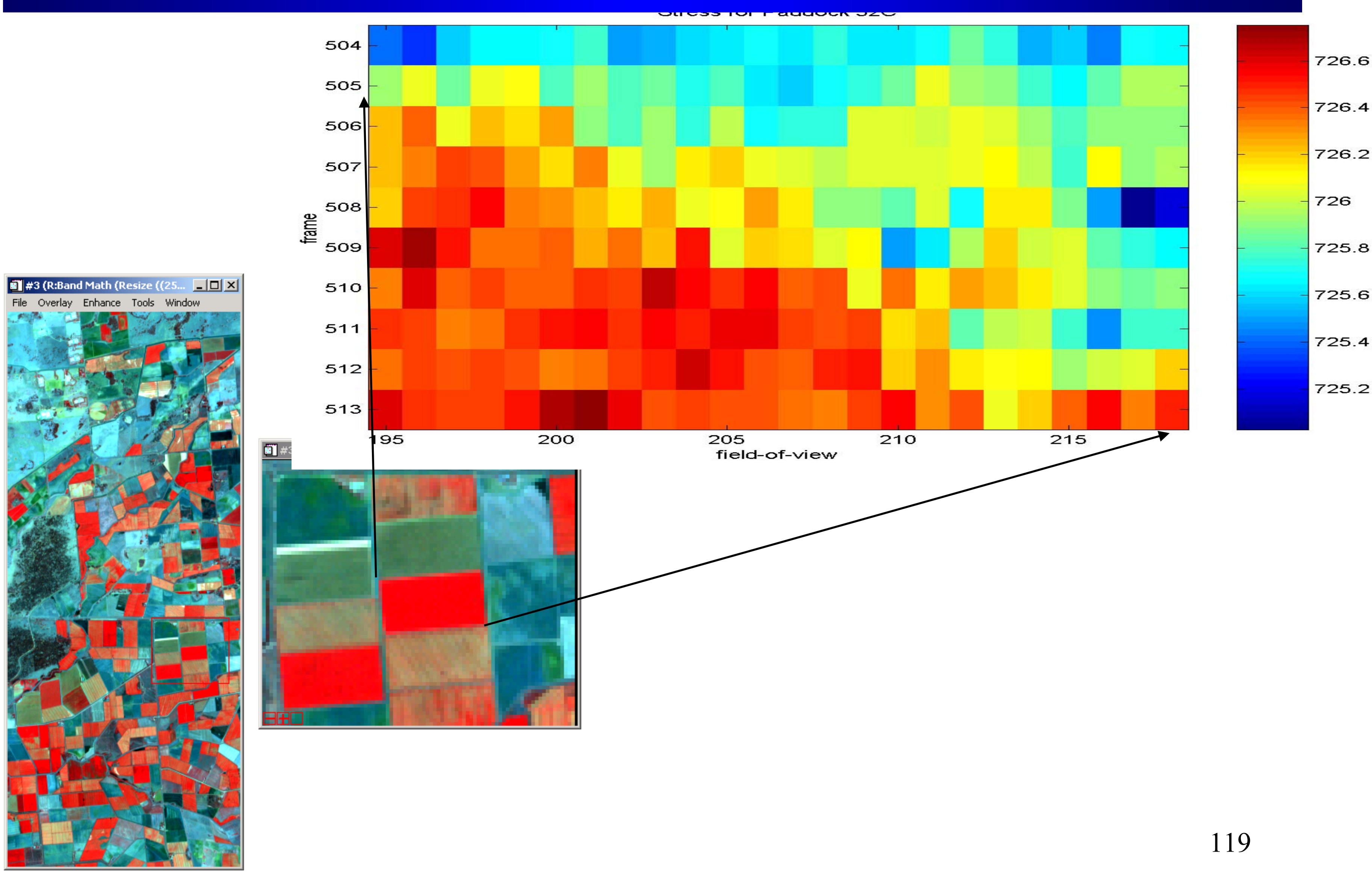
CALCITE
(gold bearing quartz)



HYPERSPECTRAL ANALYSIS

Analysis courtesy AIG Limited Liability Company

Analysis of Red Edge Shift- sample process



Now, Let's Classify the Data!

Pixel-Based Supervised Classification Methods for Hyperspectral Data

Important Issues for Classification of Hyperspectral Data

- Impact of noise and calibration
  - Lower SNR than corresponding multispectral data
  - Striping in data acquired by pushbroom sensors
  - Data susceptible to atmospheric artifacts and sensor anomalies
- Large input space
  - Enormous number of parameters to estimate
  - Sparse data in hyperspectral domain
  - Adjacent bands are often highly correlated
  - Quantity of training data typically limited
- Potentially large output space
  - Possible overlapping spectral signatures

Approaches to Resolve Issues

- High dimensional input spaces
  - Simplify or reduce dimension of input space via extraction
    - Subset feature selection
    - Projection of original features on a lower dimensional space
  - Regularize covariance matrix of input space
  - Utilize unlabelled samples via semi-supervised learning
  - Develop ensembles of classifiers
- High dimensional output spaces
  - Output space decomposition
- Utilization of nonparametric classifiers

Dealing with High Dimensional Input Spaces

Feature Selection/Extraction

- Goals
  - Reduce no. observations required to train the classifier
  - Reduce computational complexity
  - Improve classification accuracy
- Desired Properties of Extractors
  - Class dependent
  - Exploit band ordering
  - Transformations should maximize discrimination among classes

- Selection vs extraction
  - Selection of subset of original features
    - Less redundancy, but potentially some information loss
    - Better interpretability (domain knowledge)
    - Possibly improved generalization
  - Extraction
    - Typically computationally superior to selection
    - No loss of information as all features retained
    - Resultant features unrelated to physical phenomena

Approaches to Feature Extraction

- Principal Component Transform [Anderson (1984)]
  - Orthogonal linear combinations of the p original bands with maximum variance. At pixel x,

$$\boldsymbol{Z}(x) = \{Z_i(x), i = 1, \ldots p\}^t$$

  - Formulated as an optimization problem

$$Max \;\; a_i^t \boldsymbol{\Sigma} a_i$$

subject to

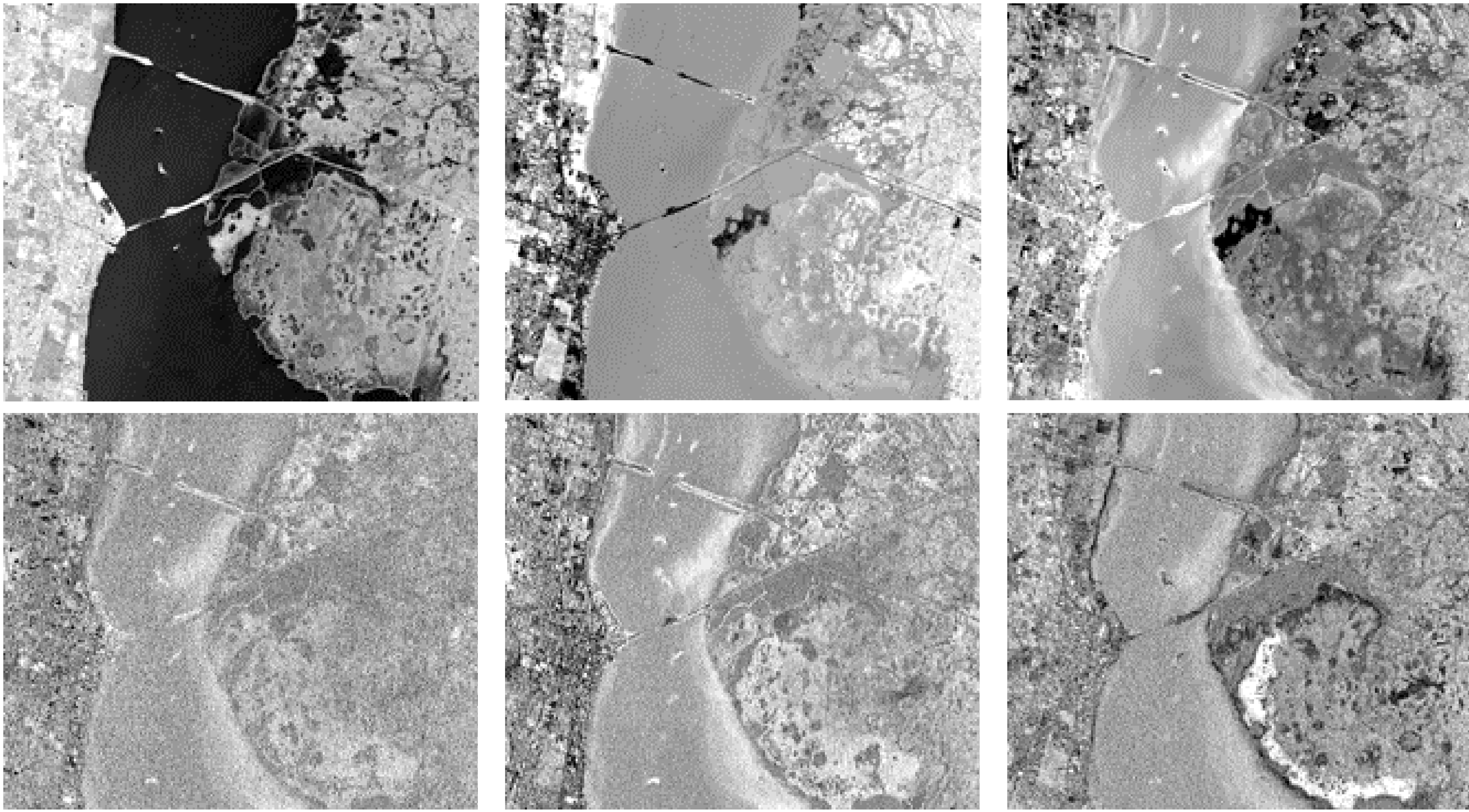
$$a_i^t a_i = 1, \quad i = 1, \ldots . p$$

$$E\left\{a_i^t \mathbf{Z}(x) \mathbf{Z}(x)^t a_j\right\} = 0 \quad \forall i \neq j, \quad j = 1, \ldots i$$

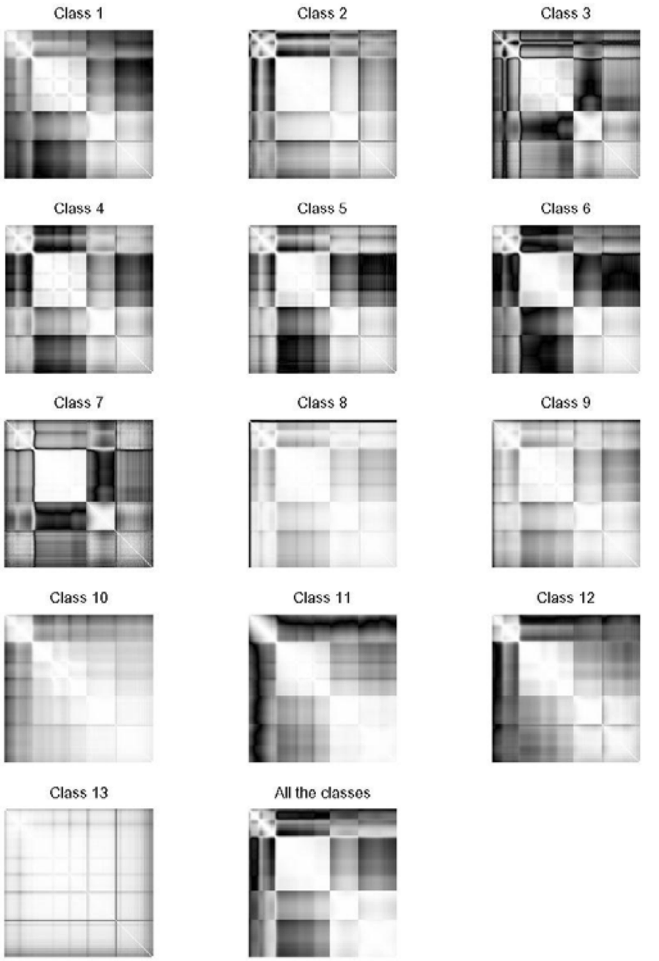
  - Solution is a subset of successively ordered eigenvalues of the covariance matrix Σ; a_i^* are corresponding eigenvectors

- Characteristics of Principal Component Transform
  - Developed to maintain variance of data in smaller set of orthogonal inputs
  - Order of components not necessarily related to data quality
  - Components have no physical relationship to the targets
  - PCT data are image content dependent (poor generalization)
  - Does not exploit the adjacency of correlated bands
  - Maximum variability has no inherent relationship to criterion for classification
  - Sensitive to outliers

Principal Components of AVIRIS Data



- Segmented Principal Component Transformations [Jia and Richards (1999)]
  - Detect boundaries in correlation matrix, compute PCT's and select subset. Further prune using Bhattacharya distance
  - Characteristics of SPCT
    - Restricted to two-class problems
    - Utilizes adjacent band correlation structure
    - Not related to discrimination measures
    - Sensitive to outliers



- Maximum Noise Fraction Transform [Green, Berman, Switzer, and Craig (1988)]
  - Uncorrelated linear combinations of bands that maximize "noise fraction" for each band.

$$\boldsymbol{Z}(x) = \{Z_i(x), i = 1, \ldots p\}^t = \boldsymbol{S}(x) + N(x)$$

where

$$\Sigma_Z = \Sigma_S + \Sigma_N \quad \Sigma_Z, \Sigma_N \text{ known}$$

  - Define "noise fraction" at each pixel x

$$NF_i(x) = Var\{N_i(x)\} / Var\{Z_i(x)\},\ i = 1, \ldots p$$

– Formulated as an optimization problem similar to PCT, except that orthogonal linear combinations are sought to maximize the Noise Fraction

$$Max\left[NF_i(x) = a_i^t \Sigma_N a_i \big/ a_i^t \Sigma_Z a_i\right]$$

subject to

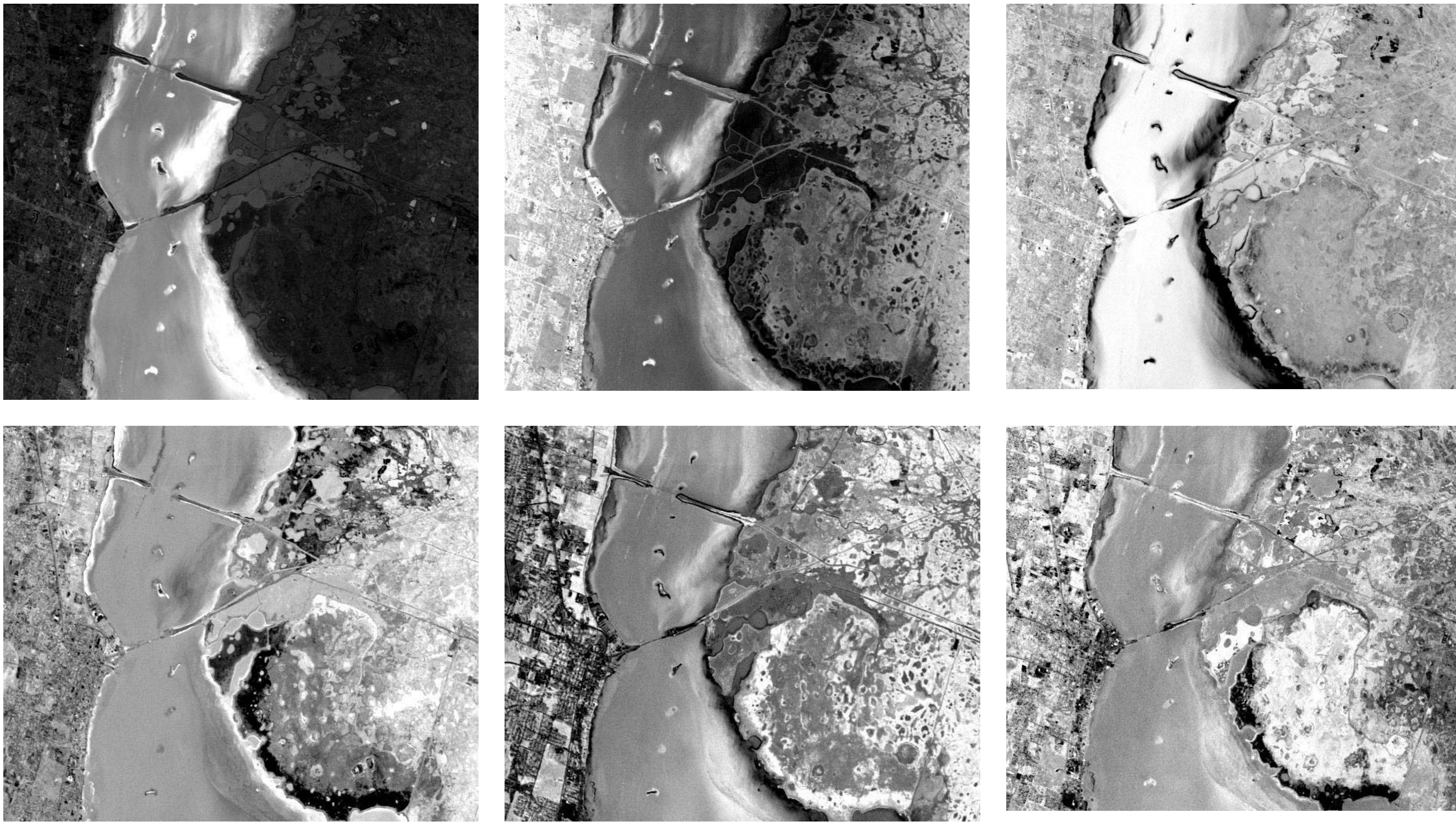
$$a_i^t \Sigma_Z a_i = 1 \quad i = 1,\dots p$$

$$E\left\{a_i^t Z(x) Z(x)^t a_j\right\} = 0 \quad \forall i \neq j,\ j = 1,\dots i$$

– Optimal weights: $a_i^* =$ eignvectors of $\Sigma_N \Sigma^{-1}$

- Characteristics of MNF Transform
  - Developed to filter noisy bands, then transform data back to original domain
  - Ordering related to quality of tranformed data
  - Invariant under data rescaling
  - Does not exploit the adjacency of correlated bands or class specific information
  - Resultant bands of transformed data *may* be related subjectively to phenomena, but are not related to discrimination between classes

MNF Bands for AVIRIS Data



- Fisher's Linear Discriminant [Anderson (1984)]
  - Seeks projection matrix $\boldsymbol{A^*}$ to maximize separation between classes $\Omega = \{\omega_i, i = 1,..C\}$

$$Max \; \frac{\left|\boldsymbol{A}^t \boldsymbol{B} \boldsymbol{A}\right|}{\left|\boldsymbol{A}^t \boldsymbol{W} \boldsymbol{A}\right|}$$

where

$$\boldsymbol{B} = \sum_{\omega\in\Omega} \left|X_\omega\right| \left(\mu_\omega - \bar{\mu}\right)\left(\mu_\omega - \bar{\mu}\right)^t \quad \boldsymbol{W} = \sum_{\omega\in\Omega} \sum_{x\in X_\omega} \left(x - \mu_\omega\right)\left(x - \mu_\omega\right)^t$$

$\boldsymbol{A^*}$ = generalized eigenvectors of $\boldsymbol{W}^{-1}\boldsymbol{B}$ $(\dim \leq C - 1)$

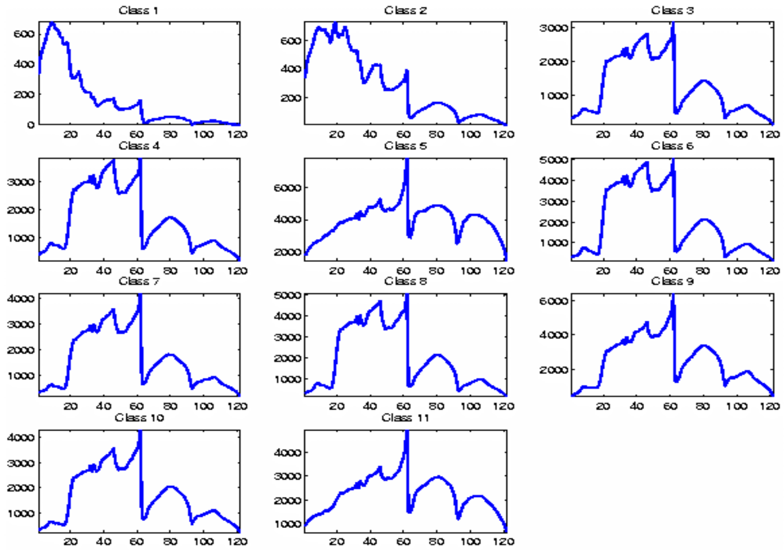
- Characteristics of Fisher’s Linear Discriminant
  - Requires a trained classifier
  - Poor discrimination well between classes with similar means
  - Classes with “different” variance dominate the combination
  - Ignores band ordering
  - Problematic for hyperspectral data with small training sample due to required matrix inversion

- Decision Boundary Feature Extraction [Lee and Landgrebe (1997)]
  - Computes a decision boundary that determines the projection of the hyperspectral space for a two-class problem
    - *C*-class problem solved using wtd sum of 2-class decision boundary feature matrices
  - Requires trained classifier
  - Ignores band ordering

- Projection Pursuit [Jimenez and Landgrebe (1999)]
  - Compute projection for two classes based on Bhattacharya distance
  - Subsets constrained to be smaller groups of adjacent bands
  - One set of features selected for all classes
  - C-class problem uses sum of discriminatory power for pairs (problematic for large no. of classes)

- Best Bases Band Aggregation [Kumar et al. (2002); Morgan et al. (2004)]
  - Adjacent bands merged based on correlation measure (or product of correlation and Fisher discriminant)
  - Implemented in Top-Down and Bottom-Up approaches
  - Characteristics
    - Exploits band adjacent correlations, maintains unique individual bands
    - Features are class specific
    - Affects signal to noise ratio in a non-uniform way
    - Mitigates impact of small training sample problem

- Polyline Feature Extraction [Henneguelle et al. (2003)]
  - Approximate mean spectrum via piece-wise polynomials (usually linear).



  - Select break points via top-down or bottom-up methods and determine parameter values (Linear: slope, midpoint of segment)
  - Parameters of piece-wise functions become input features to classifier

– Characteristics
  - Exploits band adjacency
  - Features are class dependent
  - Computationally efficient
  - Provides approximation to derivatives
  - Features robust for generalization
  - Sensitive to stopping criterion (no. of bands vs error in approximation)

- Discrete Fourier Transform
  - Describes band sequence via trigonometric functions
  - Subset of modes selected to represent features
  - Characteristics
    - Multiscale representation of spectral signature
    - Difficult to determine optimal modes to represent original problem
    - Computation unrelated to class discrimination
    - Modes may contain domain knowledge

- Discrete Wavelet Transform [Bruce et al. (2001, 2002); Kaewpijit et al. (2003)]
  - Multiscale method that accounts for both frequency and position
  - Describes band sequence via discrete wavelet transform
  - Implemented in conjunction with a selection method
  - Characteristics
    - Provides robust features
    - No. of inputs must be a power of 2
      - Accomplished via padding or interpolation
    - Criterion unrelated to class discrimination

Approaches to Feature Selection

- Selection techniques can be implemented individually, or in conjunction with extraction methods

- Methods that guarantee optimal solutions
  - Exhaustive search
  - Branch & bound (only if objective function is monotonic)

- Heuristic methods
  - Characteristics
    - Typically produce sub-optimal solutions
    - Developed to reduce computational complexity
    - Traditionally implemented via sequential forward selection or backward elimination

Heuristic Feature Selection Methods

- Sequential Floating Search Methods [Pudil et al. (1994); Serpico and Bruzzone (2001)]
  - Evaluate a number of features for elimination/forward selection for each feature (group) added/eliminated

- Tabu Search [(Zhang and Sun (2002); Korycinski et al. (2003)]
  - Developed for solution of combinatorial optimization problems
  - Non-greedy heuristic that adaptively and reactively adjusts search space, flexible in number of features evaluated at each iteration

- Various simulated annealing and genetic algorithms [Siedlecki and Sklansky (1988, 1989)]

Regularization of Covariance Matrix

- Methods to regularize estimated covariance matrix
  - Pseudo-inverse [Fukanaga (1990); Skurichina and Duin (1996); Raudys and Duin (1998)]
    - Poor performance when ratio of amount of training data to input dimension is small
    - Exhibits "peaking" effect at $|X| = d/2$
  - Shrinking and pooling covariance matrices [Tadjudin and Landgrebe (1999)]
    - Reduce variance of estimates, but introduce bias

Incorporation of Unlabeled Samples

- Semi-supervised learning
  - Augment existing training data with unlabeled samples [Jeon and Landgrebe (1999); Tadjudin and Landgrebe (2000); Jackson and Landgrebe (2001)]
    - Classify unlabeled observations and update estimates, usually via EM algorithm
    - Alternatively perturb sample data
  - Effective for generalization and knowledge re-use
  - Sensitive to original training samples, impacted by outliers

Dealing with Large Output Spaces

Multiclassifier Systems

- Multiclassifier Systems
  - Ensembles of base classifiers that are learned individually to produce an aggregated predictor
  - Characteristics
    - Component classifiers often weak
    - Decision boundaries for individual classifiers typically simple
    - *Generalization* typically superior, if classifiers are *diverse*
- Goals
  - Improve classification accuracy and generalization
  - Reduce complexity and improve interpretation
  - Enhance transferability of classifiers to new problems or beyond spatial extent of training data
  - Mitigate the impact of sensor artifacts
  - Improved utilization of small quantity of training data
- Challenges
  - Achieving highly diverse, but relevant classifiers
  - Maintaining interpretability

Multiclassifiers from Input Space Decomposition

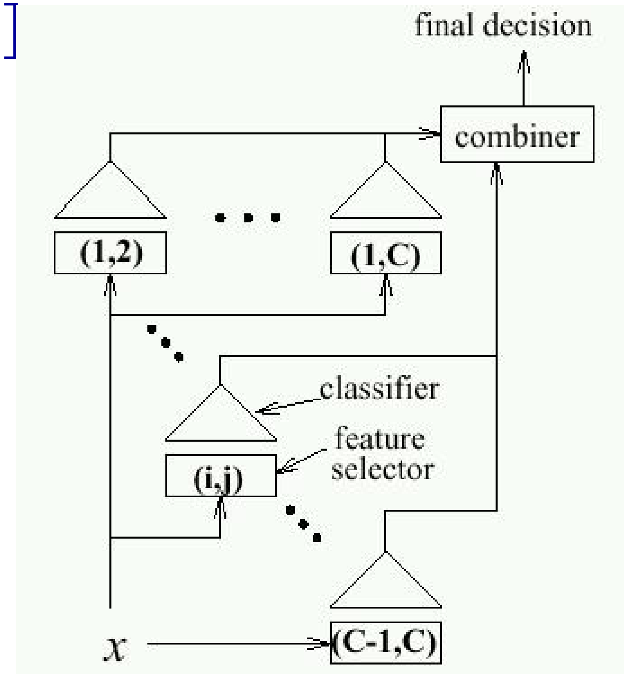
- Selecting sub-samples of original data, creating classifiers, and developing a classifier for each sample. Combine results from individual classifiers
  - Perform poorly for extremely small sample sizes as ensemble methods cannot overcome lack of diversity
- Bagging [Brieman (1996)]
  - Create multiple samples with replacement; develop individual classifiers
- Arcing
  - Adaptively reweighting and combining (e.g. boosting) [Freund and Schapire (1996); Dietterich (2000)]
  - Simple random sub-sampling [Ho (1998); Skurichina and Duin (2002)] (See previous slides for details)

- Random subspace selection methods [Ho (1998); Skurichina and Duin (2002)]
  - Utilized in conjunction with multiclassifier systems
    - Randomly sample input features for each classifier
    - Combine outputs of multiple classifiers via voting or maximizing aposteriori probabilities
  - Characteristics
    - Provides diversity and robust classifiers with good generalization
    - Reduces band redundancy, but gnores band adjacency, unless implemented with band aggregation (which reduces diversity)
    - Mitigates impact of small sample size problems and stabilizes parameter estimates
    - Computationally expensive
    - Does not provide domain knowledge on feature characteristics
    - Mitigates the impact of anomalous sensor behavior (striping)

Multiclassifier Systems: Output Space Decomposition

- C *one-vs-rest* problems [Anand et al. (1995)]
  - Classify individual classes vs the mixture of the "rest"
  - Can result unbalanced priors for Bayesian methods if sample sizes for some classes extremely small
  - Does not exploit class affinity, so convergence can be very slow

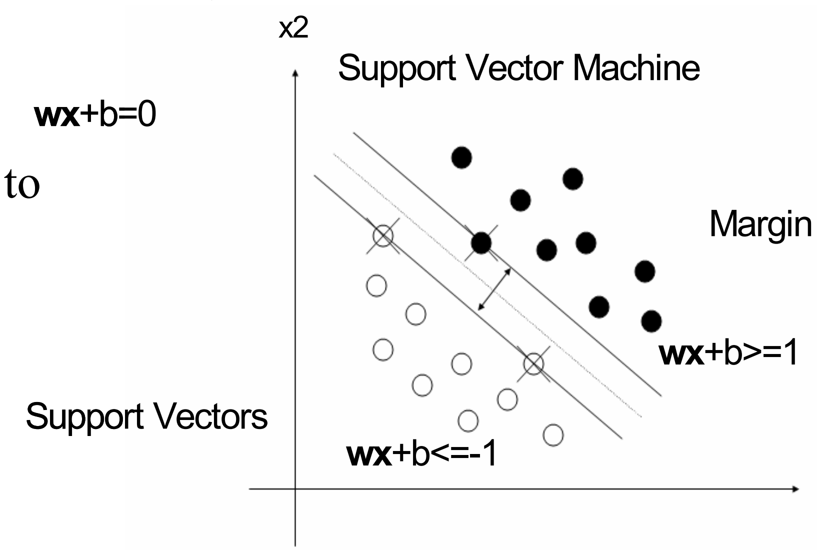
- Pairwise classifiers [Crawford et al.(1999); Kumar et al. (2001); Furnkranz (2002)]
  - Selects best set of features for class pairs, performs classification, then combines results for C-class problem via voting or maximum *aposteriori* rule
  - Can incorporate pairwise feature extraction
  - Yields $\binom{C}{2}$ classifiers
  - Potential coupling of outputs



- Error correcting output codes [Dietterich and Bakiri (1995); Rajan and Ghosh (2004)]
  - C-class problem encoded as $\bar{C} >> C$ binary classifiers which are combined by voting
  - Each class assigned according to $\bar{C} \times C$ code book
  - Observations labeled as class with code closest to code formed by outputs of $\bar{C}$ classifiers
  - Characteristics
    - Implemented with user-selected classifiers
    - Feature selection/extraction tunable to pairwise classifiers
    - Output dependent on classifier separation
    - Reduces impact of small sample sizes
    - Does not exploit natural class affinities

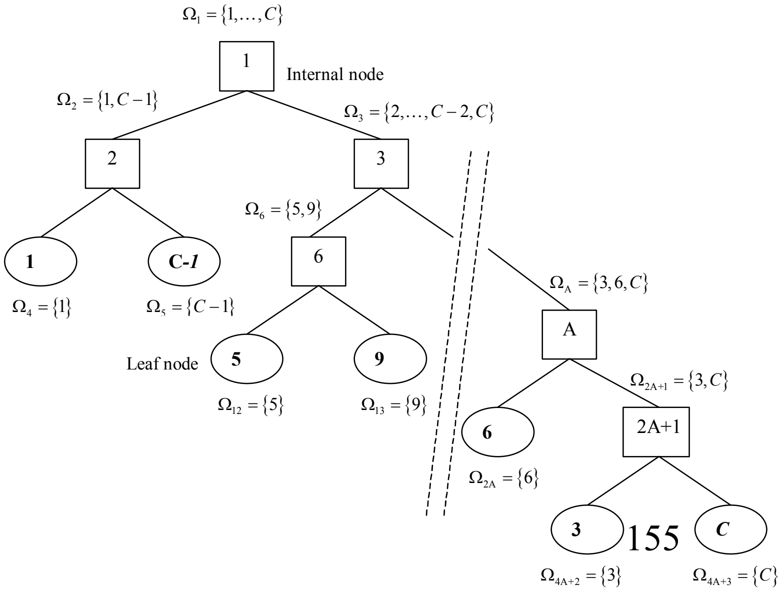
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	1	0	1	1	0	0	1	0	1	0	0	0	1
3	0	1	1	0	1	0	1	1	1	1	0	0	1	0
4	1	0	1	1	0	0	1	0	1	0	0	0	1	1
5	1	1	0	1	0	1	1	1	1	0	0	1	0	0
6	0	0	0	0	1	1	1	0	1	1	0	1	0	1
7	1	0	1	1	1	1	0	0	0	1	0	1	1	0
8	0	1	1	0	0	1	0	1	0	0	0	1	1	1
9	0	1	1	1	0	1	1	0	0	1	1	0	0	0
10	1	0	1	0	1	1	1	1	0	0	1	0	0	1
11	0	0	0	1	1	1	0	1	1	0	1	0	1	0
12	1	1	0	0	0	1	0	0	1	1	1	0	1	1
13	1	0	1	0	0	0	0	1	1	1	1	1	0	0
14	0	1	1	1	1	0	0	0	1	0	1	1	0	1
15	1	1	0	0	1	0	1	0	0	0	1	1	1	0
16	0	0	0	1	0	0	1	1	0	1	1	1	1	1

- Support vector machines [Vapnik (1995); Hsu and Lin (2002)]
  - Maximize the margin between two class samples instead of accuracies
  - Characteristics
    - Formulated as an optimization problem for a binary classifier
    - Nonparametric
    - Potentially high dimensional decision boundary
    - Classification accuracies dependent on typically slow tuning
    - Tends not to overtrain, which leads to high classification accuracies and good generalization



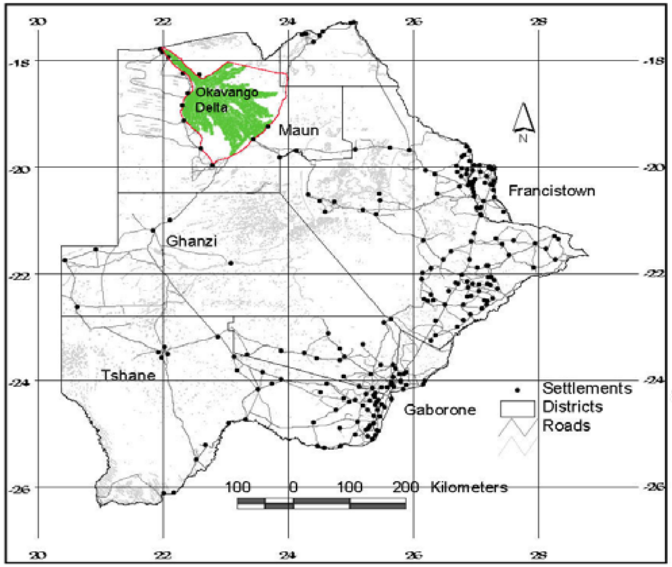
- Hierarchical Decomposition Methods
  - Characteristics
    - Adopt a sequential “divide and conquer” approach involving decomposition based on input or output spaces
    - Maintain natural groupings with useful domain knowlege
  - Hierarchical Decision Trees
    - Tree constructed sequentially based on series of binary questions and splitting rule which maximizes decrease in impurity of parent and child nodes
      - CART: Best split based on linear combinations of input features; Gini impurity index, post pruning [Brieman et al. (1984)]
      - C4.5: Impurity index based on entropy [Quinlan (1986)]

- Binary Hierarchical Classifier [Kumar et al. (2002); Morgan et al. (2004)]
  - Output space decomposition approach
    - *C* leaf nodes, *C*-1 internal nodes
  - Top down tree constructed via deterministic simulated annealing
  - Feature extraction/selection tunable to each internal node
  - Classifier specific to each internal node (Fisher discriminant, SVM)
  - Mitigates small sample problems
  - Exploits natural class affinities

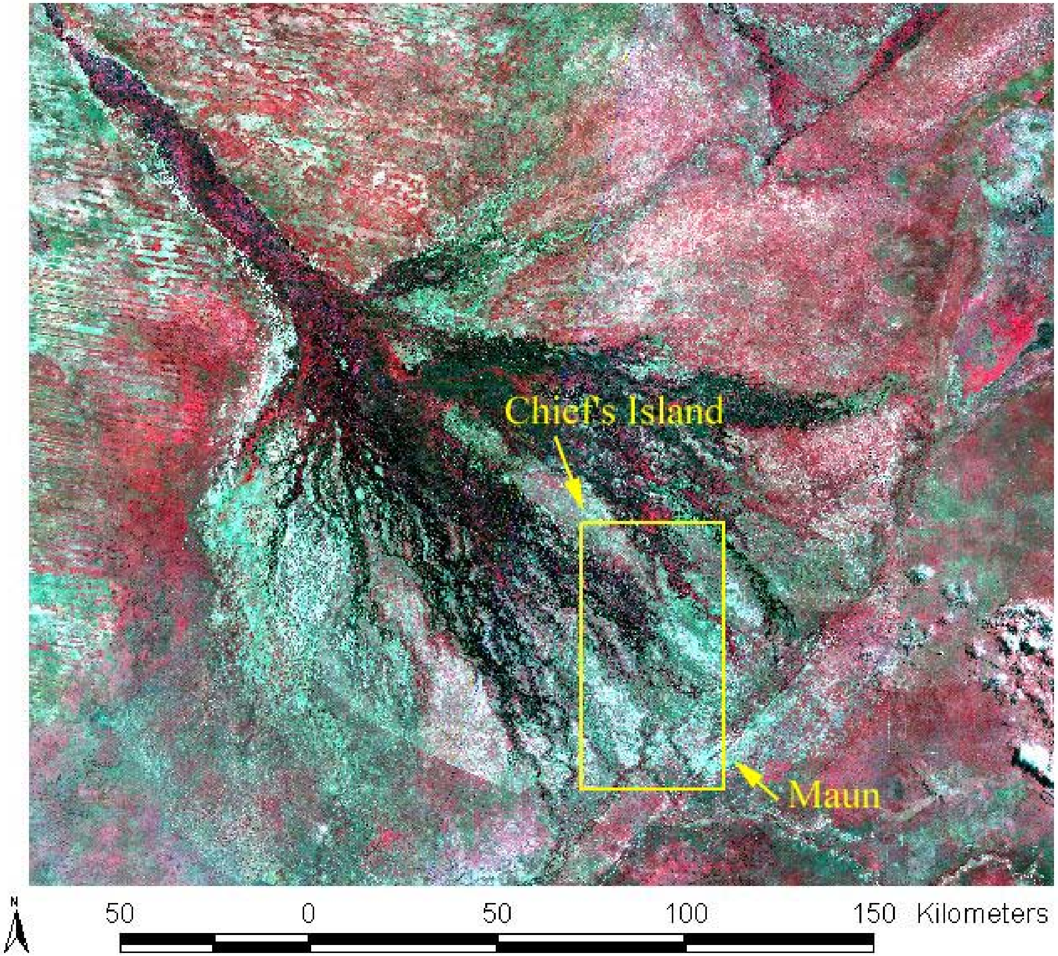


Let’s look at an example

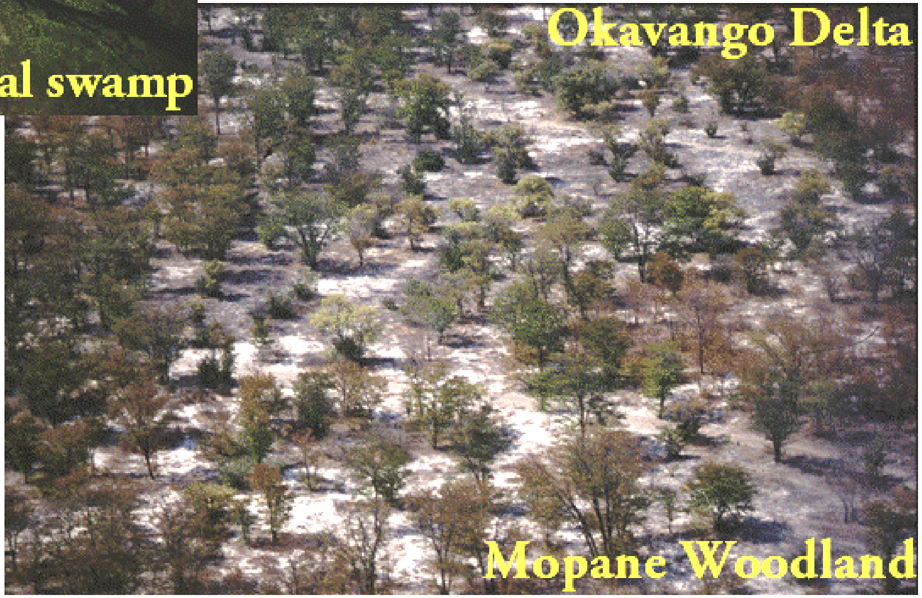
Study Site: Okavango Delta, Botswana



Republic of Botswana



Classification Results: Okavango Delta



Classification scheme consists of 14 landcover classes selected using IKONOS imagery, aerial photography, and field campaigns

Hyperion Hyperspectral Data Inputs

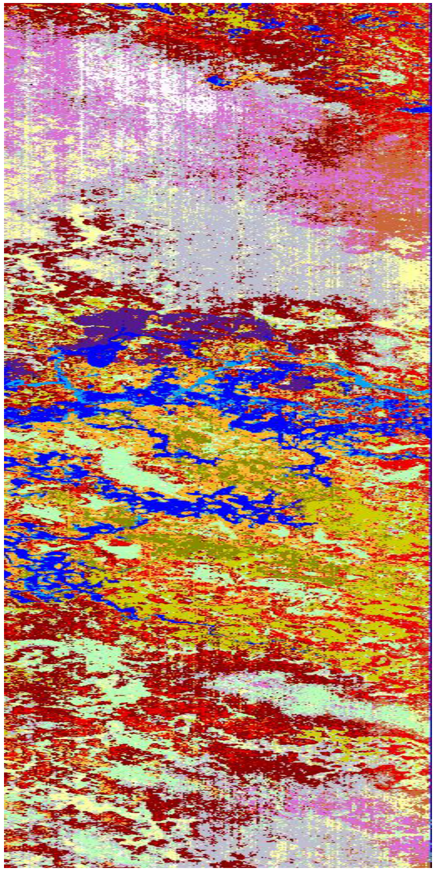
- Hyperion hyperspectral data
  - Acquired by Earth Observer-1 Satellite May 31, 2001
  - Converted to reflectances, destriped, georeferenced

Class #	Class Name	Sample Size
1	water	270
2	hippo grass	101
3	floodplain grasses1	251
4	floodplain grasses2	215
5	reeds	269
6	riparian	269
7	firescar	259
8	island interior	203
9	acacia woodlands	314
10	acacia shrublands	248
11	acacia grasslands	305
12	short mopane	181
13	mixed mopane	268
14	exposed soils	95

Design of Experiment for Random Forests

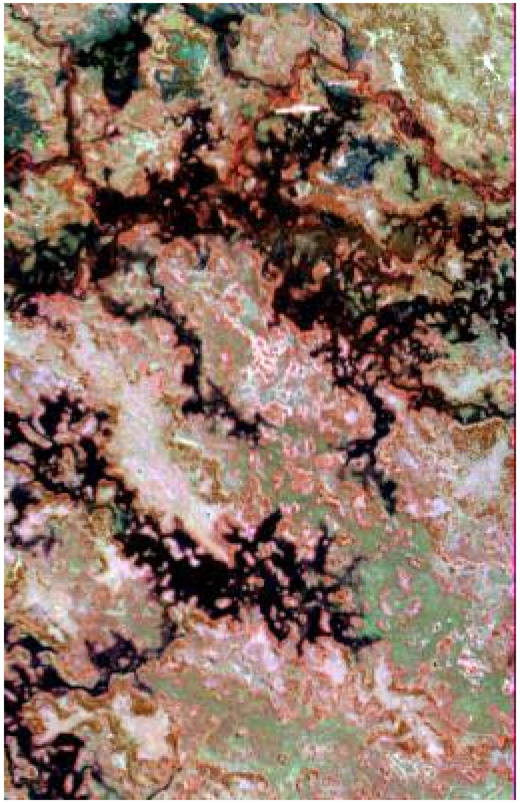
- Create 10 independent stratified samples of training/test data
  - Select 75%:25% of training data for training and test
  - Subsample from 75% to achieve 50, 15% sample size; test remains constant
  - Select spatially removed test sample
- Create BHC tree
  - Select bagged sample
  - Randomly sample input features
    - (D/5) or min(D/5, 20)
  - Develop tree
- Grow BHC Random Forest
- Compare to BB-BHC, RS-BHC, BB RS BHC, RF CART

Adequate Destriping is Critical for Classification

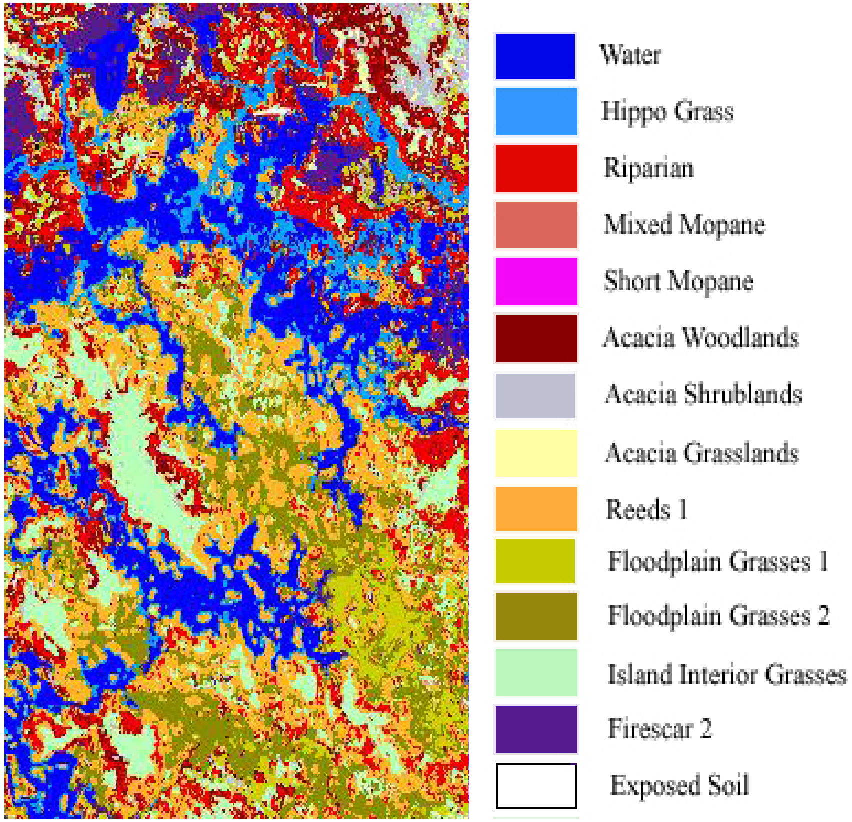


Classification Results with Poor Destriping

Classification Results: BHC Algorithm Okavango Hyperion

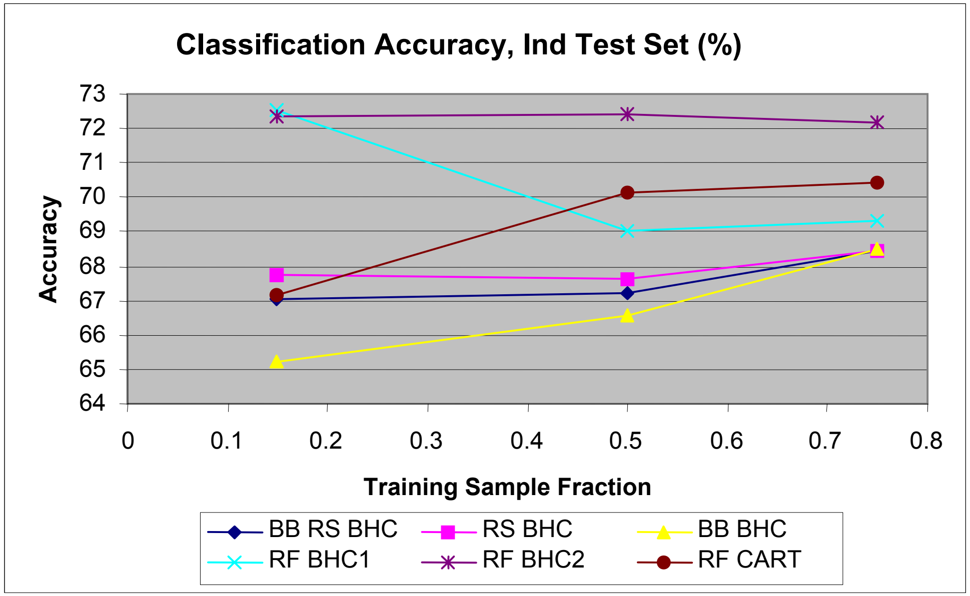


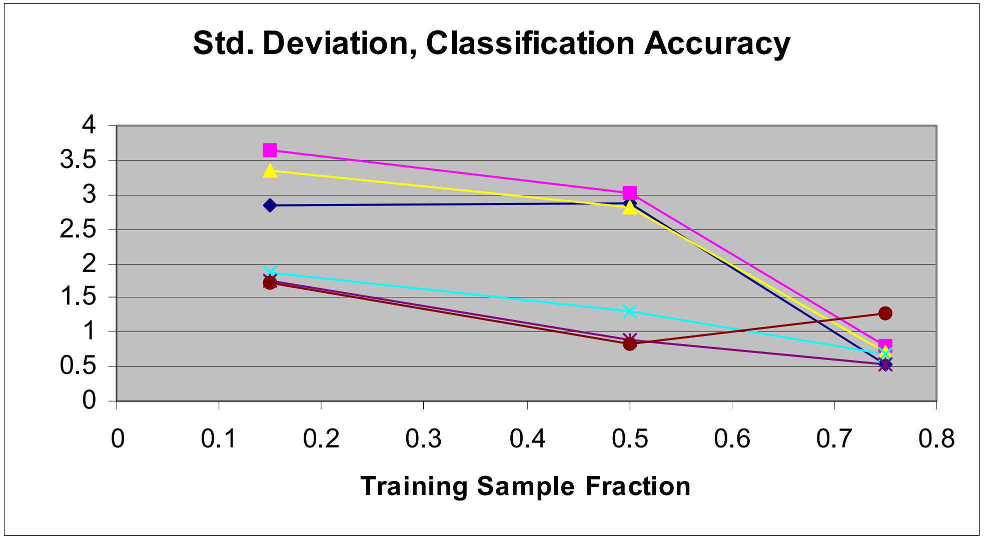
(a) Subset of Hyperion data (Bands 51, 149, 31),



(b) Classified image of Hyperion data

Classification Accuracies, Independent Test Set





Next Steps - an interactive discussion