MSS: What's the ROI*?

Landsat Science Team Meeting
January 12, 2016
Virginia Polytechnic Institute

Dennis Helder
…and the Landsat cal team…

South Dakota State University
Image Processing Lab

*Return on Investment

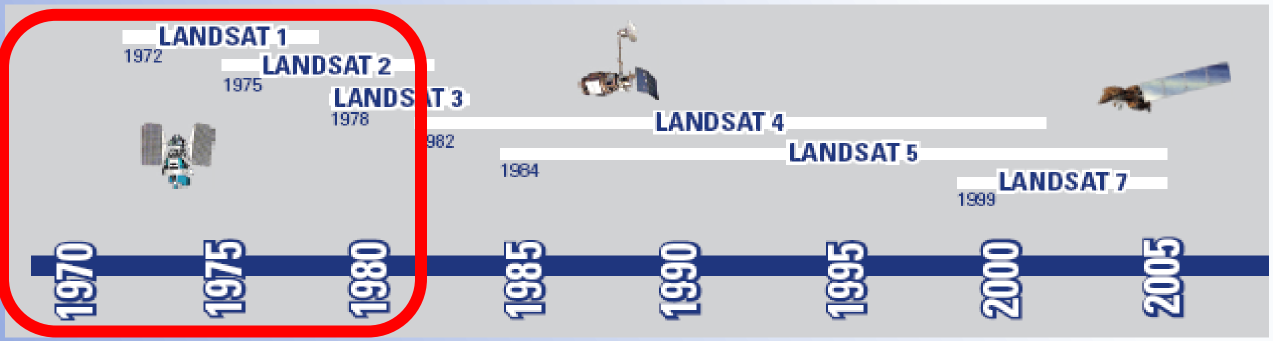
Outline

- The Return
- The Investment
  - Spectral
  - Calibration
  - Reflectance
  - Surface (maybe even a return???)

An old slide… but still an important story

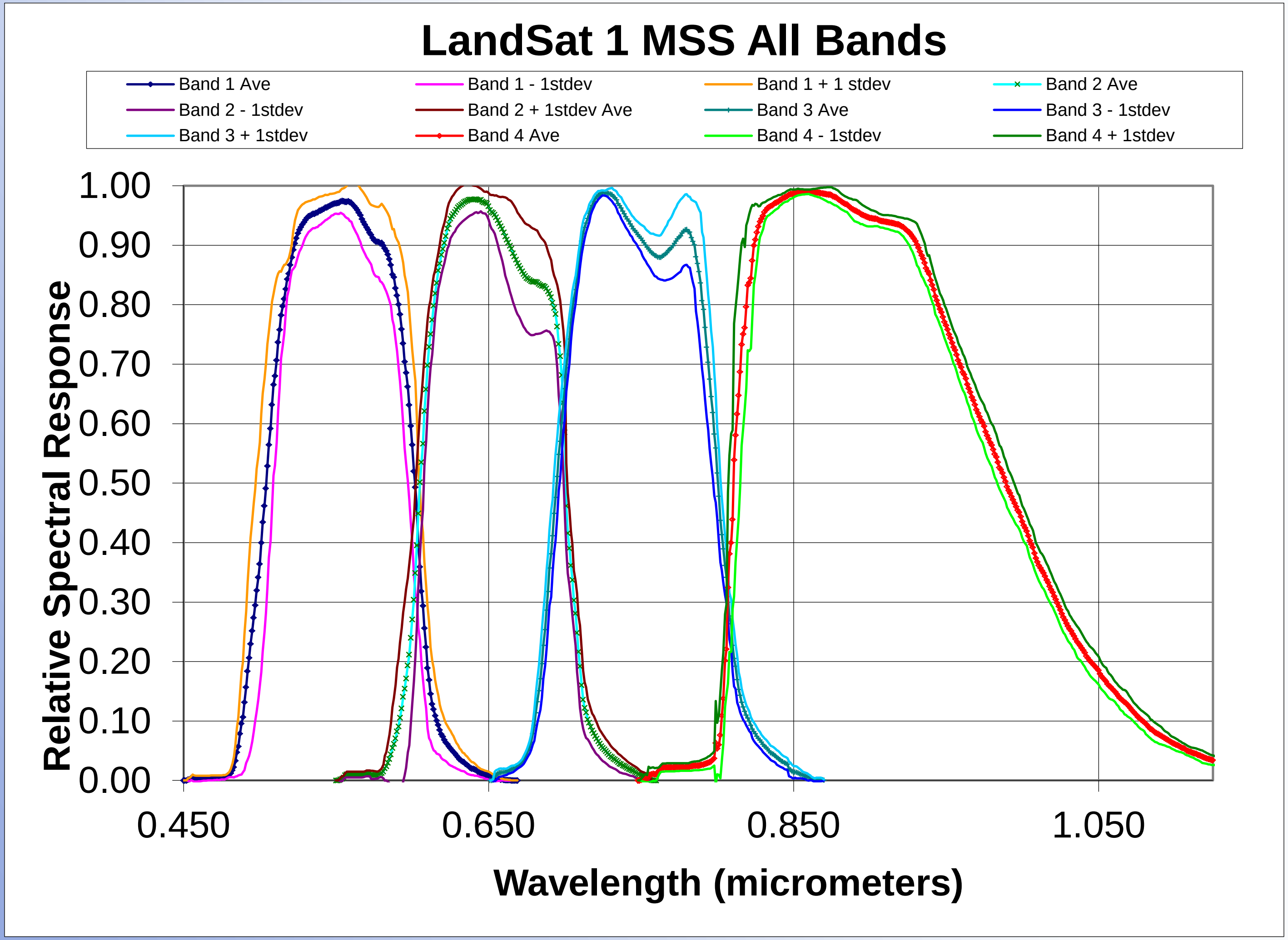
MSS: The Return

- **Landsat Program** - the longest running Earth observing program
- **Current and Past Landsat Imaging Systems**
  - 6 satellites (Landsat-1 through -5, and -7), Now 7!
  - no fewer than 8 instruments (5 MSS, 2 TM, and 1 ETM+ sensors) Now 9… or more
- MSS data in the USGS archive provides a historical record of the Earth's land surface from the early 1970's to the early 1990's, and was acquired from 5 MSS instruments onboard Landsat-1 through -5.
- For the multi-temporal studies, it is critically important that these 5 MSS instruments be cross-calibrated with one another.



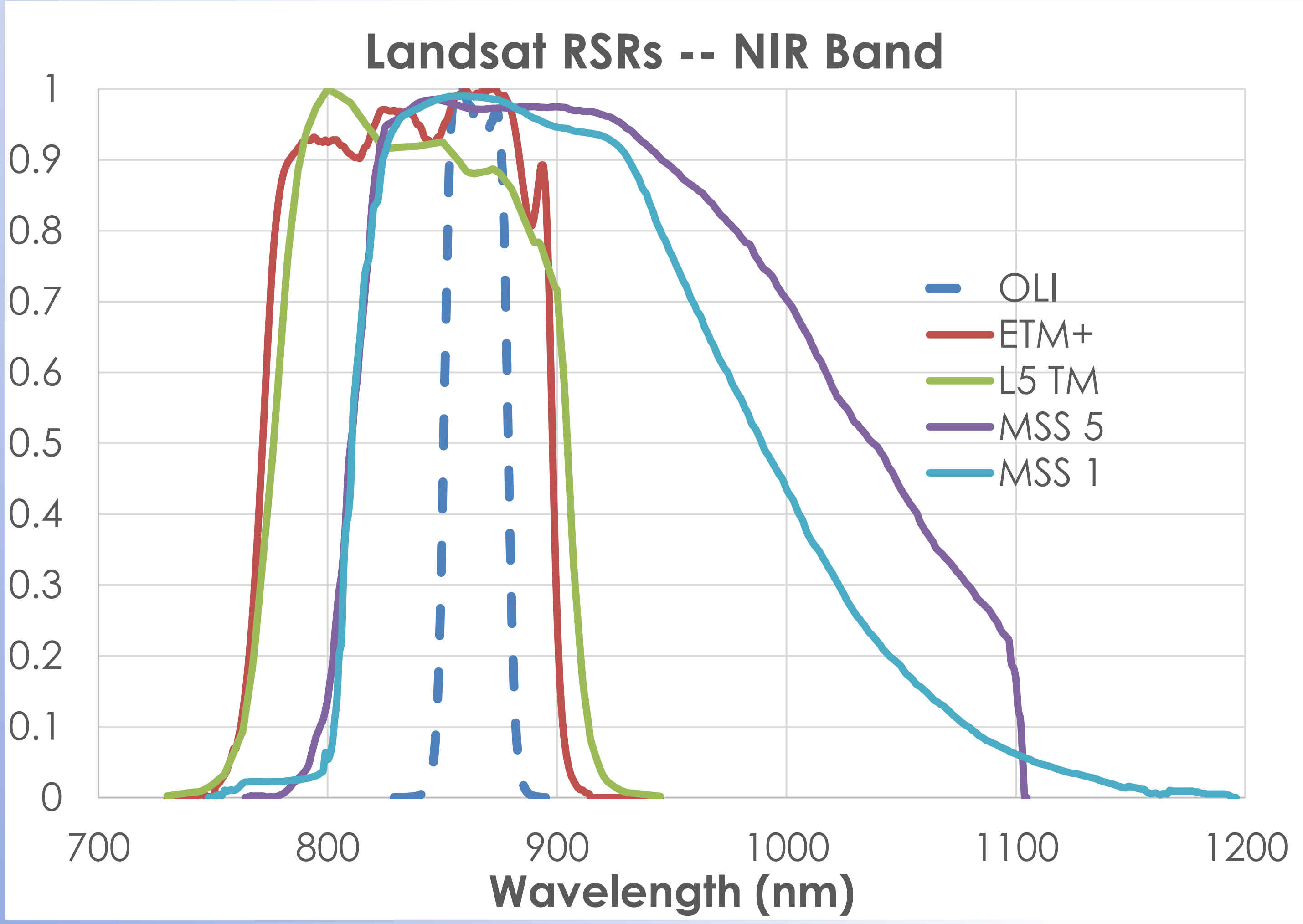
MSS is the only game in town to get back to the '70s… !!

The Investment: Spectral



Detectors in one band of one MSS sensor don’t measure the same thing!

The Investment: Spectral



Landsat sensors don't measure the same thing!

The Investment: Calibration

- Can MSS sensors be calibrated well enough to be useful?
  - Impact of Landsat 8 OLI
  - Impact of reflectance-based calibration
  - Impact of improvement in cross-calibration methods

MSS Cross-Calibration

Landsat 4 MSS to Landsat 5 MSS

Band	Atmospheric effects	Filter Spectral effects	BRDF effects	Total uncertainty estimates
1	4%	0.5%	1%	4%
2	3%	0.2%	1%	4%
3	4%	0.9%	2%	5%
4	11%	0.4%	2%	12%

Landsat MSS cross-calibration nominally propagates 4-12% errors from sensor to sensor with one notable exception—MSS 4 to MSS3.

Landsat 3 MSS to Landsat 4 MSS

Band	Atmospheric effects	Filter Spectral effects	BRDF effects	Total uncertainty estimates
1	0%	1.6%	0%	2%
2	0%	0.7%	0%	1%
3	0%	1.2%	0%	1%
4	0%	1.3%	0%	1%

Landsat 1 MSS to Landsat 2 MSS

Band	Atmospheri c effects	Spectral Filter effects	BRDF effects	Total uncertain estimates
1	4%	0.7%	1%	4%
2	3%	1.1%	1%	4%
3	4%	0.5%	2%	5%
4	11%	0.5%	2%	12%

Landsat 2 MSS to Landsat 3 MSS

Band	Atmospheric effects	Filter Spectral effects	BRDF effects	Total uncertainty estimates
1	4%	1.1%	1%	4%
2	3%	0.8%	1%	4%
3	4%	1.4%	2%	5%
4	11%	0.3%	2%	12%

Radiometric Calibration of the Landsat MSS Sensor Series, Helder et al., TGRS Vol. 50, No. 6.

Landsat Absolute Calibration

- Previous radiance-based calibration used Landsat 7 as the 'Gold Standard
- Current in-progess reflectance-based calibration uses Landsat 8 OLI as 'Gold Standard' and incorporates cross-cal improvements

Landsat Absolute Radiometric Calibration Uncertainties (%)								
	L7 ETM+	L5 TM	L4 TM	MSS 5	MSS 4	MSS 3	MSS 2	MSS 1
Blue	5	7	9	8	9	9	10	11
Green	5	7	9	8	9	9	10	11
Red	5	7	9	9	10	10	11	12
NIR	5	7	9	14	18	18	22	25
SWIR 1	5	7	9					
SWIR 2	5	7	9					

Forty-year calibrated record of earth-reflected radiance from Landsat: A review, Markham & Helder, Rem. Sens. Env., Vol 122, July 2012, pp. 30-40.

Landsat Absolute Radiometric Calibration Uncertainties (%)									
	L8 OLI	L7 ETM+	L5 TM	L4 TM	MSS 5	MSS 4	MSS 3	MSS 2	MSS 1
Blue	**3**	4	5	7	6	7	7	8	9
Green	**3**	4	5	7	6	7	7	8	9
Red	**3**	4	5	7	6	8	8	9	11
NIR	**3**	4	5	7	7	14	14	18	22
SWIR 1	**3**	4	5	7					
SWIR 2	**3**	4	5	7					

Reflectance based Cross-calibration of MSS and TM sensors—current status

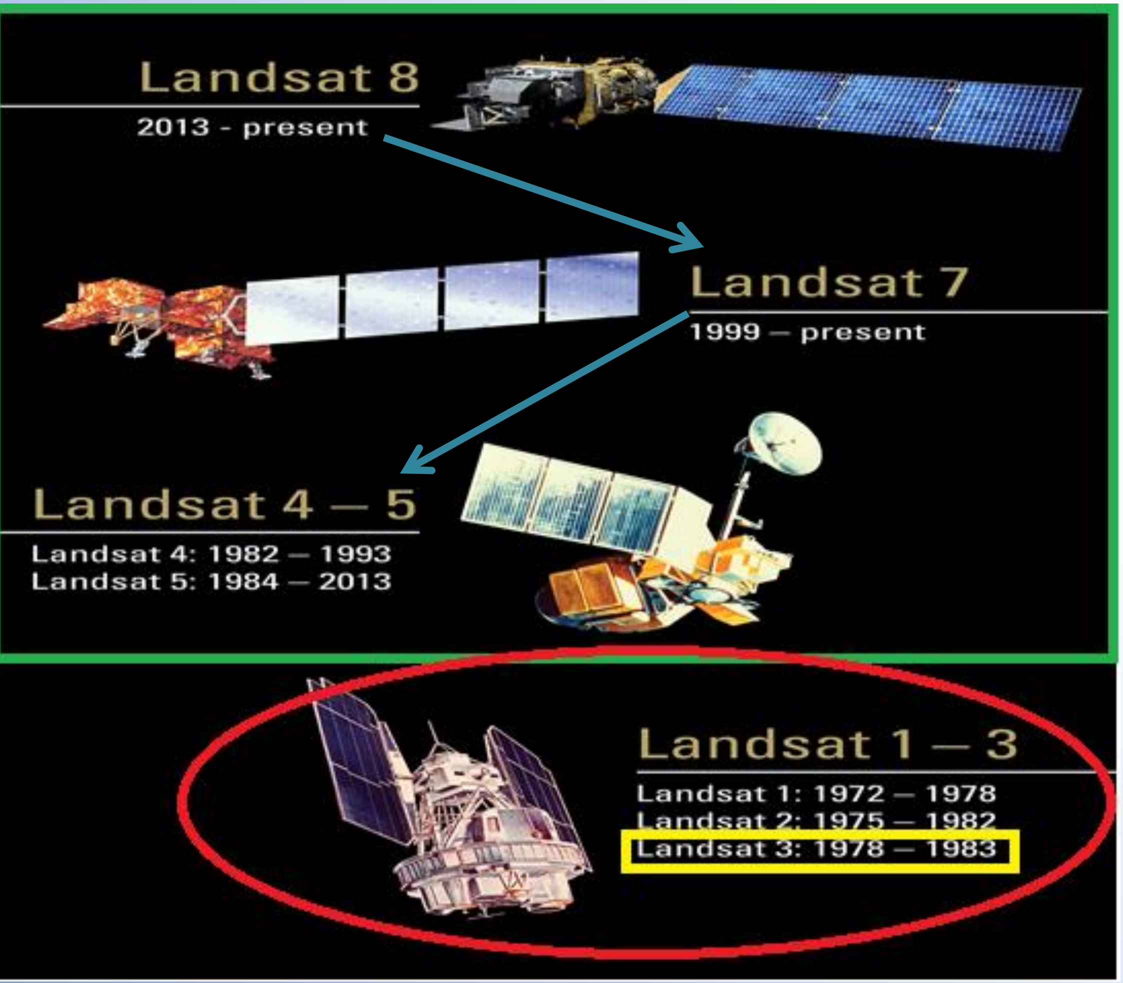
Sandeep Kumar Chittimalli
Larry Leigh, Dennis Helder
12/08/2015

South Dakota State University
Image Processing Lab

Background

- **The main objective is to transfer Landsat 8 TOA reflectance calibration back in time for consistent calibration of entire Landsat archive.**
- The basic idea is to be able to generate TOA reflectance from Landsat products without going through at-sensor radiance in a way which minimizes the uncertainty due to ESUN. This work also includes recalibrating some of the earlier Landsats based on Landsat 8.
- Some of the cross-cal and validation results were successfully shown in the previous TIM calibration meeting where calibration from Landsat 8 OLI to Landsat 4MSS looked consistent, but Landsat 3 calibration did not look consistent with time.

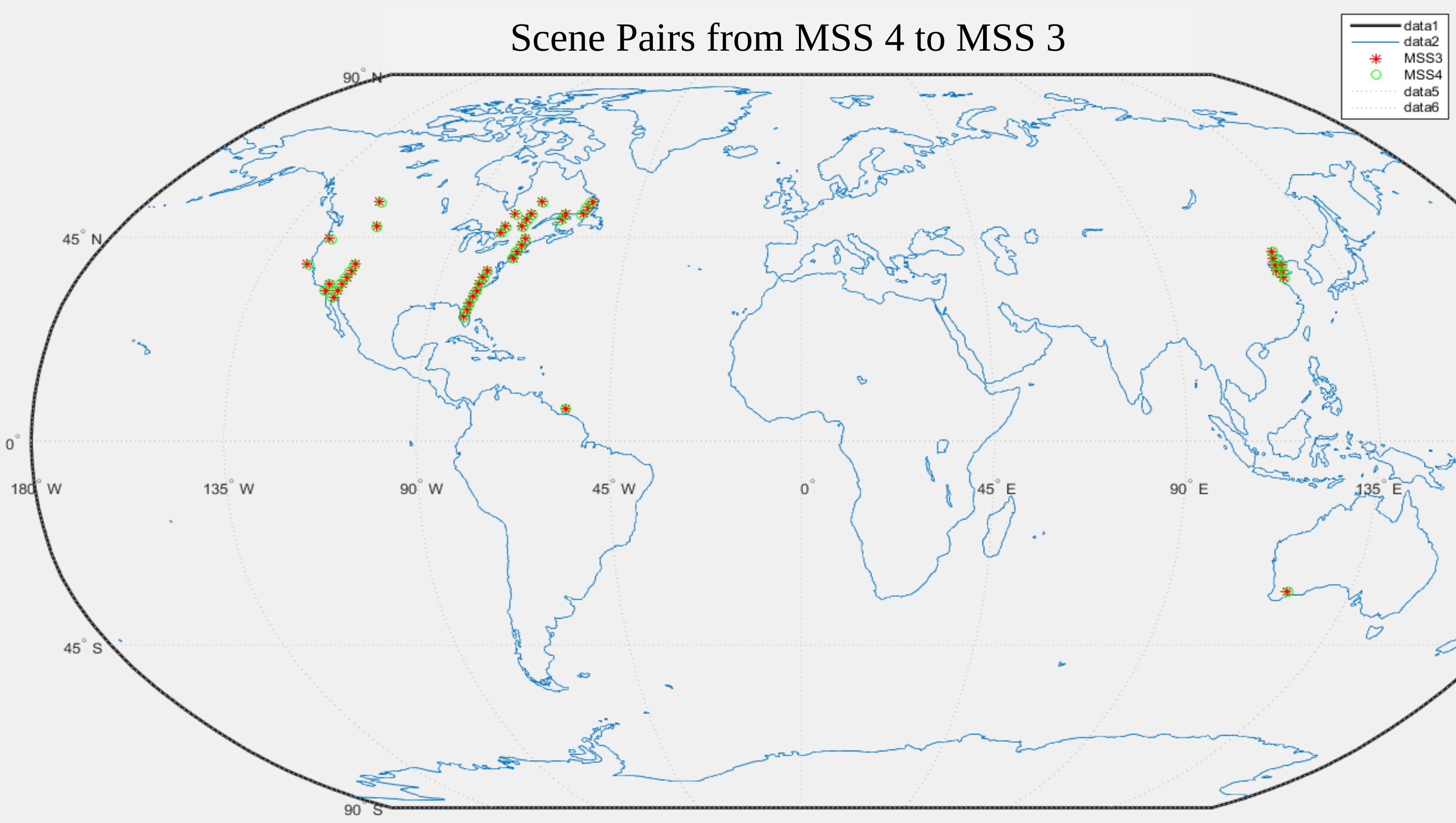
Where we were at the last LST meeting



'Reflectance' based Cross-Cal and its validation for Landsats 4-8, except Landsat 4 TM, was shown in the last TIM.

Landsat 3 data looked inconsistent, so we looked at it closer. Significant bias in MSS-R data lead us to look at MSS-A and MSS-P data to calibrate MSS 3.

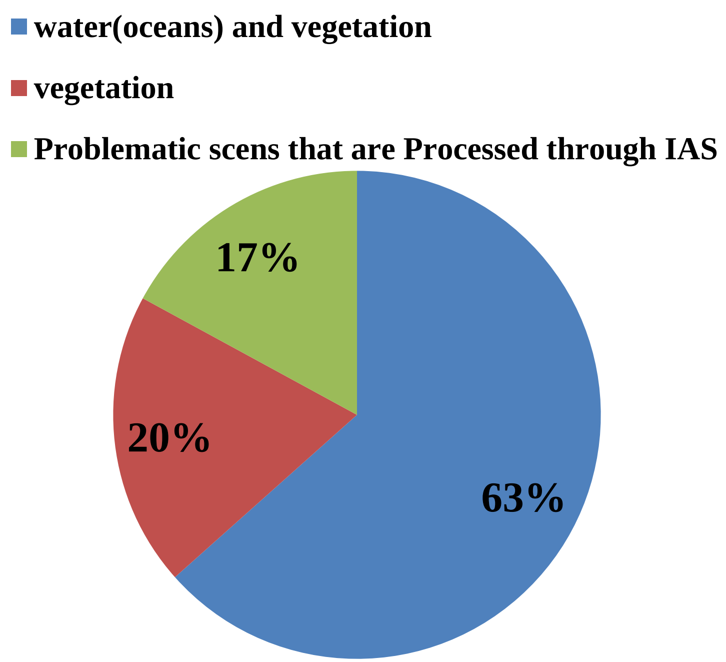
As part of the initial work we have done an exhaustive search to find scene pairs for Landsat 4 MSS and Landsat 3 MSS. We came up with 45 scene pairs of cloud free data.



Problem that we faced with MSS 3 data

Among 45 scene pairs, 10 scene pairs can be used to calibrate MSS 3. Only MSS-P and MSS-A data is processed through IAS and after a brief study we found that 17% (7 scene pairs) were problematic. They show high striping.

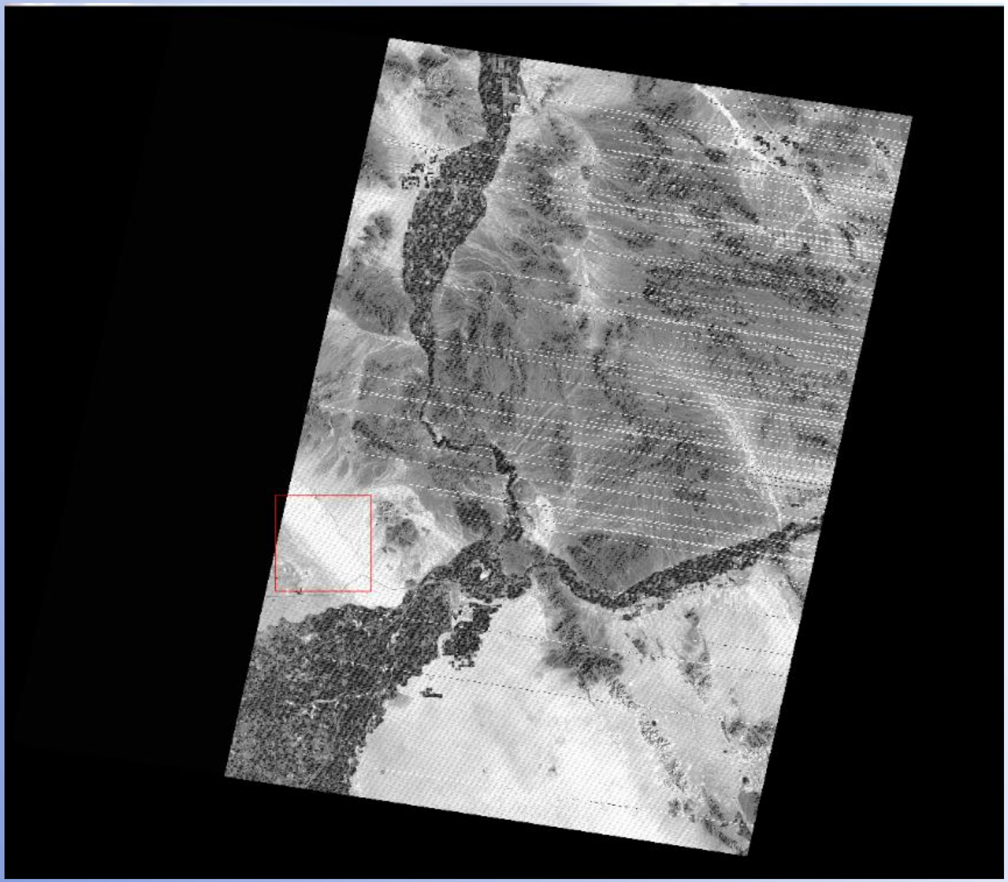
Number of MSS4 to MSS3 Scene Pairs found: 45

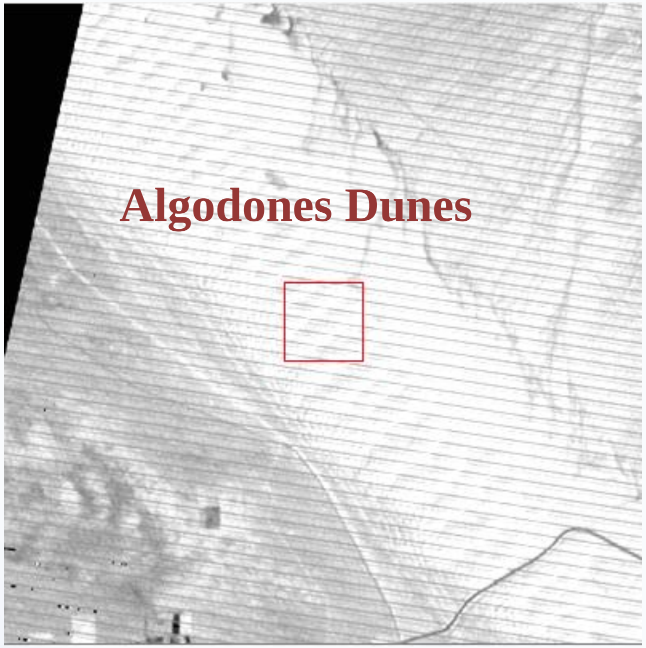


The pie chart shows percentage of MSS 3 data with stripes and that needs to be corrected by IAS so a better calibration of MSS 3 can be done.
As of now, we don’t have enough data to calibrate MSS 3.

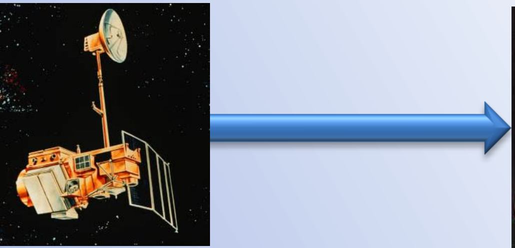
Sample scene from Algodones dunes showing strips

This is in MSS-A data format





Part of the Algodones Dunes which is zoomed showing high percentage of stripes



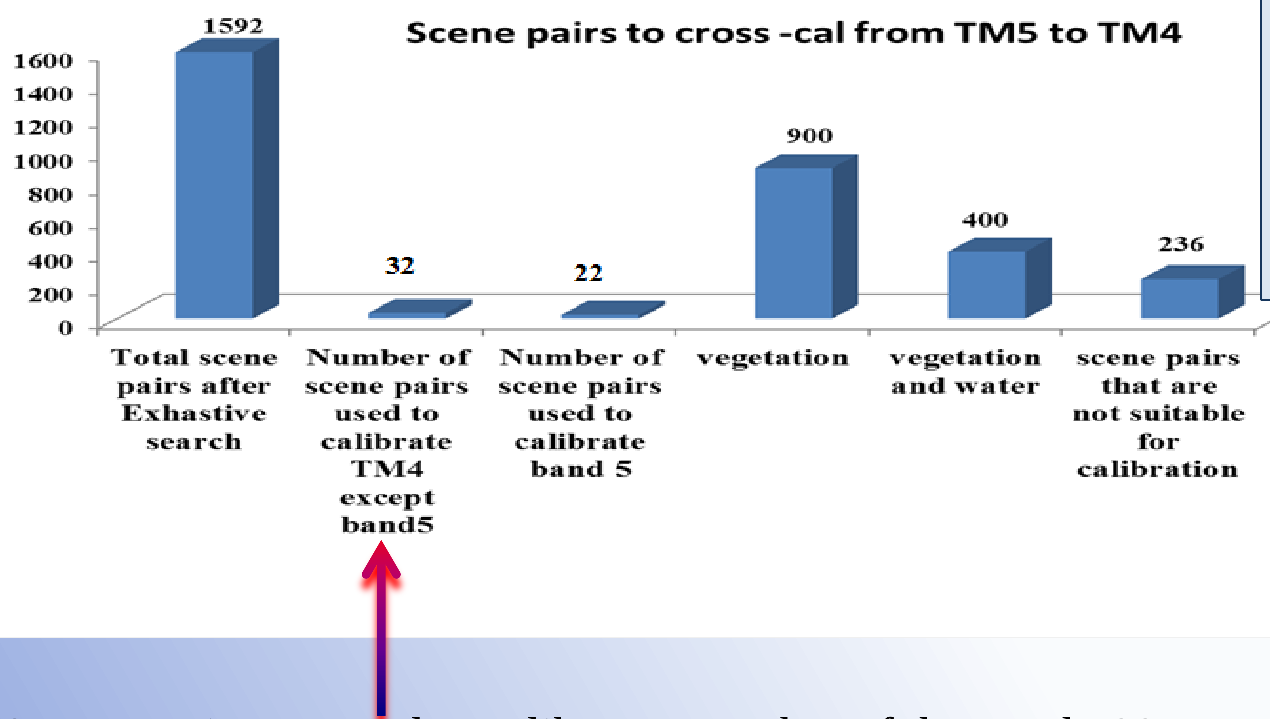
Landsat-5 to Landsat-4

Radiometric and Reflectance based Cross-Cal from Landsat 5 TM to Landsat 4 TM

Procedure

- Exhaustive search to find near–coincident scene pairs for Landsat 5 TM and Landsat 4 TM.
- Started with L1T product, converted to Radiance, scaled it by day 1 launch gain in the CPF which brought it back to DN. Then corrected for gain, bias subtraction, gain drift, etc.
- Multiple ROIs with an optimal size of 3x3km to 5x5km were selected from near-coincident scene pairs. Some darker ROIs were also selected for a better distribution.
- SBAF was derived using Hyperion.
- Linear regression approach was used to find gain and bias.
- Statistical tests were done to check the significance of gain and bias.

Initially we have done an exhaustive search to find scene pairs from TM 5 to TM 4 and we found 1592 scene pairs of cloud free data. Then we filtered out the scenes which are not suitable for calibration. Here is the summary chart.



Scene Pair Criteria

≤ 8 days apart

Temporally stable

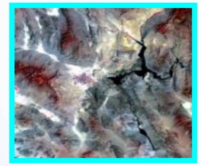
Spatially stable

Bright

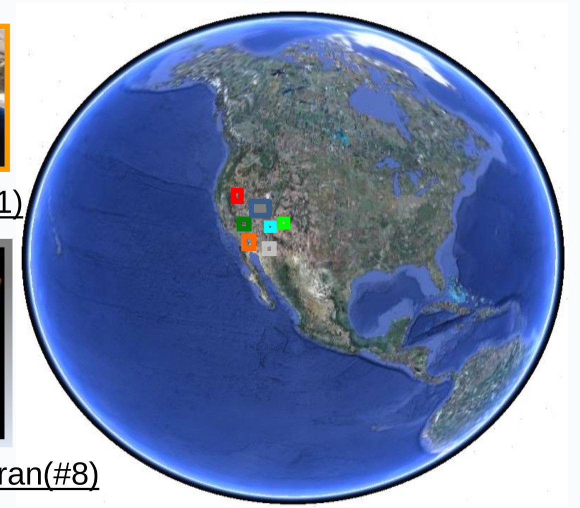
Homogeneous

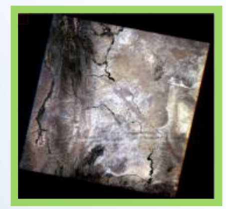
Overall 32 scene pairs are used to calibrate TM4, but of them only 22 scene pairs were used to calibrate band 5 as this band is saturated over the Arabian desert.

22 scene pairs within USA


Ivanpah Playa(#4)

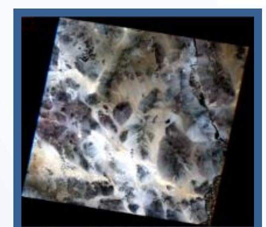

Sonora Desert(#1)

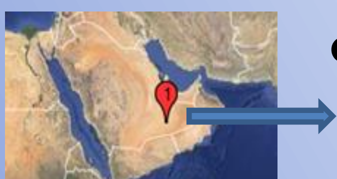



Rock Springs(#2)


Near to Sonoran(#8)


Algodones Dunes(#3)


Kelso Dunes (#4)



Overall 10 scene pairs from Arabian deserts 1,2,3

Overall 32 near-coincident scene pairs from TM 5 to TM 4 with time differences of 8 days were used to calibrate TM 4.

Procedure to calibrate 4 TM using 5TM

Reflectance Cal equation

$$\rho_{5TM,\lambda} = L_{5TM,\lambda} * (L_{5,\text{CPF average\_band gains}}(t) / L_{5,\text{corrected linear drif}}\ gains(t)) * g_{7_{june1,1999,L7-L5Cross\_cal}}) / g_{5TM,\rho,\lambda}$$

$\rho_{5TM,\lambda}$: **Band Specific Reflectance as seen by L5 TM**

$L_{5TM,\lambda}$: **Band Specific Radiance as seen by L5 TM**

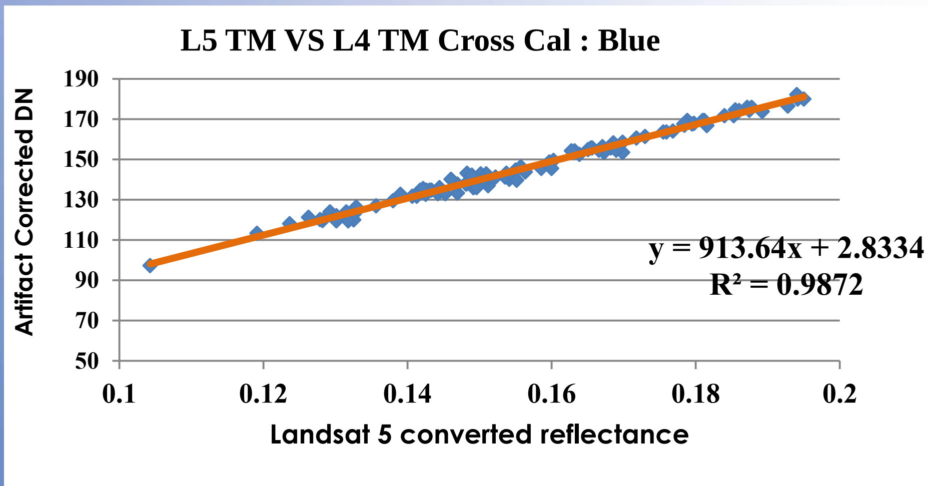
$g_{5TM,\rho,\lambda}$: **Band** Specific L5 TM Reflectance gain to convert from Reflectance space to DN space

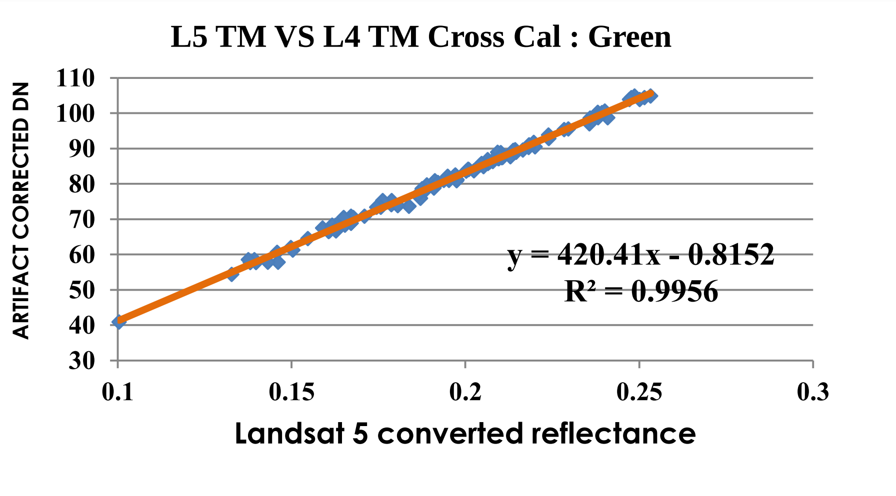
Reflectance Cal Equation to take Reflectance from 5TM to 4TM DN

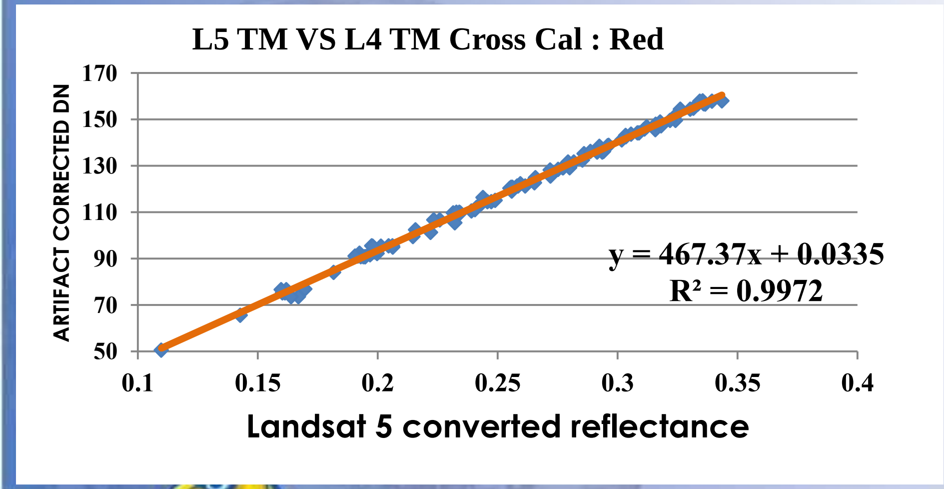
$$DN_{4TM,\lambda} = g_{4TM,\rho,\lambda} * (SBAF_{\frac{4TM}{5TM}} * \rho_{5TM,\lambda} * \frac{d^2{}_{4TM}}{d^2{}_{5TM}} * \frac{\cos\alpha_{4TM}}{\cos\alpha_{5TM}}) + b_{4TM,\rho,\lambda}$$

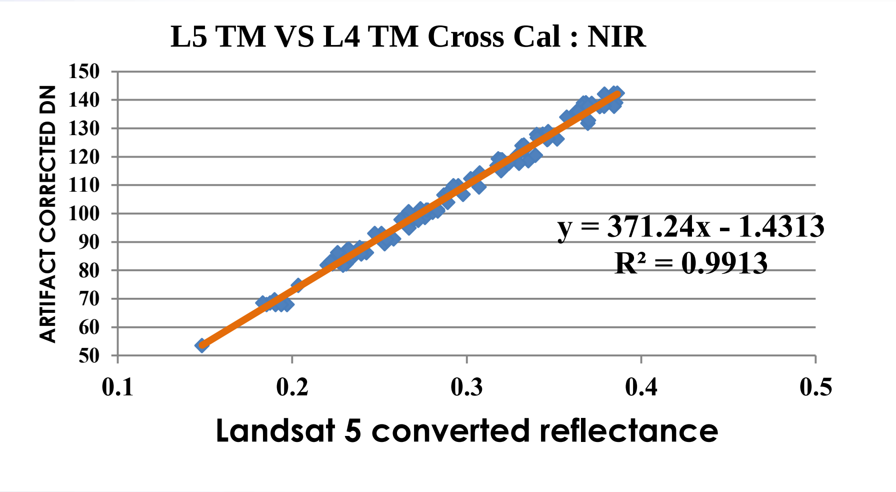
*Note DN(ACDN),4TM = L4 $Radiance * (L4TM, CPF\ day1\ launch\ Average\ band\ gains)$

Landsat 5 TM to 4 TM Cross-Calibration

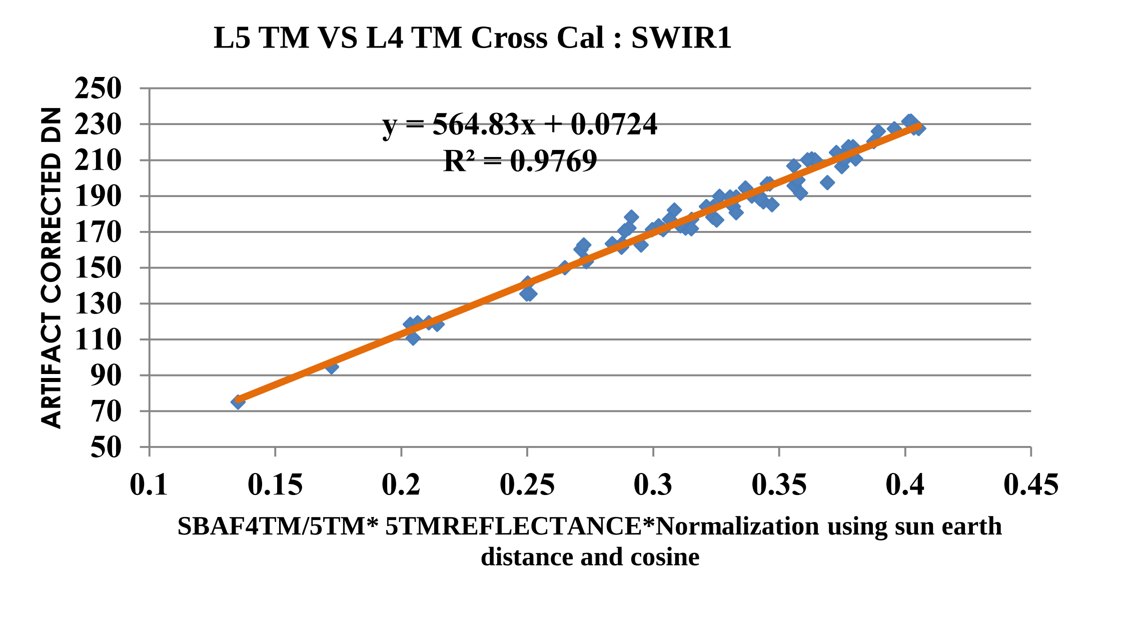


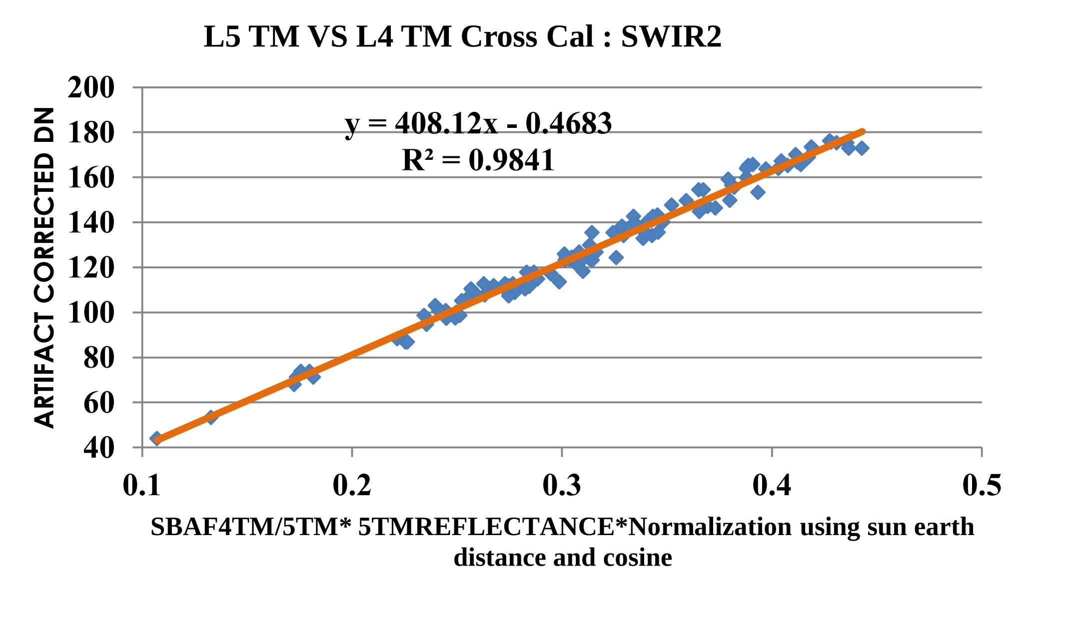






South Dakota State University
Image Processing Lab





The R-squared value for SWIR 1 and SWIR2 is low when compared to the remaining bands due to low signal and low atmospheric effects. But residual plots show that these are a good fit for both bands.

Statistical Test For Landsat 5 TM to 4TM Cross-Calibration

Null Hypothesis: Bias=0

		Band 1		
	Coefficients	Standard Error	t Stat	P-value
(Intercept)	2.833385	1.635377	1.732557	0.086227
Slope	913.637	10.35122	88.26373	<2e-16

		Band 2		
	Coefficients	Standard Error	t Stat	P-value
(Intercept)	-0.81517	0.542821	-1.50173	0.136286
Slope	420.4146	2.767523	151.9101	<2e-16

		Band 3		
	Coefficients	Standard Error	t Stat	P-value
(Intercept)	0.033524	0.63213	0.053033	0.957811
Slope	467.3745	2.446247	191.0577	<2e-16

		Band 4		
	Coefficients	Standard Error	t Stat	P-value
(Intercept)	-1.43131	1.02734	-1.39322	0.166612
Slope	371.2442	3.458519	107.342	<2e-16

		Band 5		
	Coefficients	Standard Error	t Stat	P-value
(Intercept)	0.072371	3.295017	0.021964	0.982537
Slope	564.8344	10.17117	55.53291	<2e-16

		Band 7		
	Coefficients	Standard Error	t Stat	P-value
(Intercept)	-0.46832	1.659553	-0.28219	0.778372
Slope	408.1243	5.15586	79.15737	<2e-16

- From the statistical test at 95% confidence, bias (Intercept) for all the bands were found to be insignificant, hence the fitted line is passed through zero to find the gain for the model.

MSS Surface Reflectance Product???
a.k.a.
SMACAA-ING the MSS Archive

Larry Leigh
South Dakota State University
Larry.Leigh@sdstate.edu

South Dakota State University
Image Processing Lab

MSS Surface Reflectance

- Brief overview of the SMACAA
- SMACAA scaled for cluster performance
- What's been processed to date
- How the output is generated and applied

Overview SMACAA

- SDSU Modtran Atmospheric Compensation for Any-time and Any-location (SMACAA)
  - SMACAA was borne out of the need to do atmospheric corrections over desert sites (PICS).
    - These locations tended to cause issues for other atmospheric correction methodologies, mainly because no “dark targets” existed in the field of view of the sensor.
  - The approach taken was to remove the need for any “scene content” information.
  - This really opens up locations that can be corrected but also the sensors we correct don’t require any specific band passes.

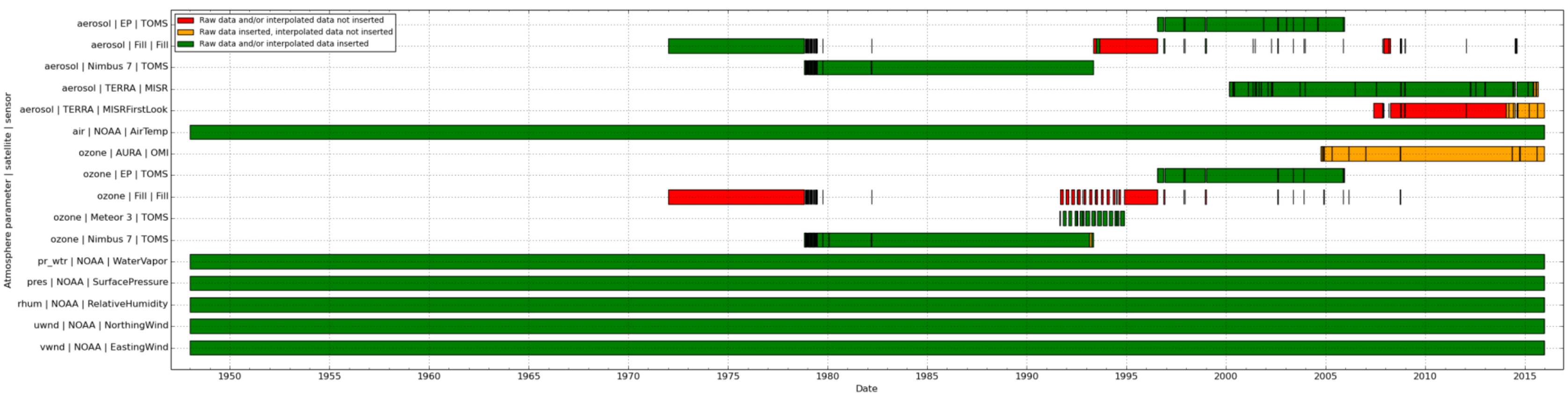
SMACAA DB

- To get around the need for specific scene content (for the purposes of determining atmospheric make up) requires “external” knowledge of what’s happening in the atmosphere.
  - First question that had to be answered is, does this information exist, does it cover the globe, and does it take us back to the launch on Landsat 1?

SMACAA DB

Current Status:

- All NOAA NCEP data (water vapor key)
  - **OPERATIONAL**: Automatic processing and ingest
- All OMI / TOMS Ozone ingested
  - **OPERATIONAL**: Automatic processing and ingest
- All MISR/TOMS Aerosol data ingested
  - **OPERATIONAL**: Automatic processing and ingest
- DEM – SRTM
  - **OPERATIONAL**: Automatic processing
- NOAA CO2
  - **OPERATIONAL**: Automatic processing



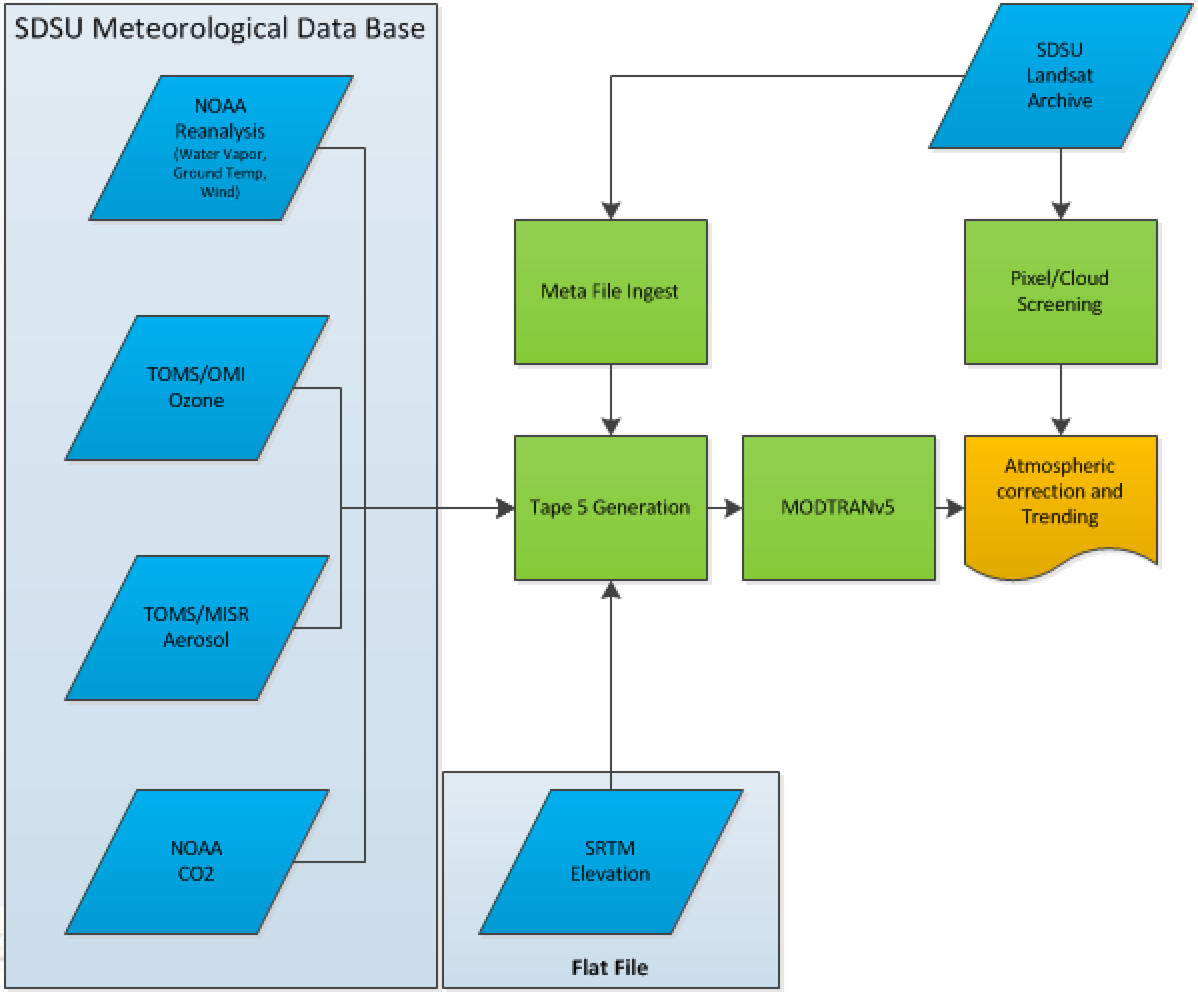
Current holding of the Database covers the minimal requirements to do the processing for any location on the planet and as far back as 1972, the data is automatically ingested / screened / and converted.

Other sources of data that could help fill some of the data gaps exist, and can / will be added in the future. SMACAA is only as good as the data behind it.

*Kalnay et al.,The NCEP/NCAR 40-year reanalysis project, Bull. Amer. Meteor. Soc., 77, 437-470, 1996

Automated Process - Flow Chart

- Local MYSQL Databases are actively being updated via external links
  - NASA & NOAA
- All data is automatically pulled via Matlab scripts from the database
- Images located in SDSU Archive or META file dumps are ingested in bulk or individually, screened and processed.
- Finally all inputs are passed through Modtran, with the results being applied to the imagery or correction parameters generated for external application.



SMACAA Output

- The output is a set of 3 numbers / per band.

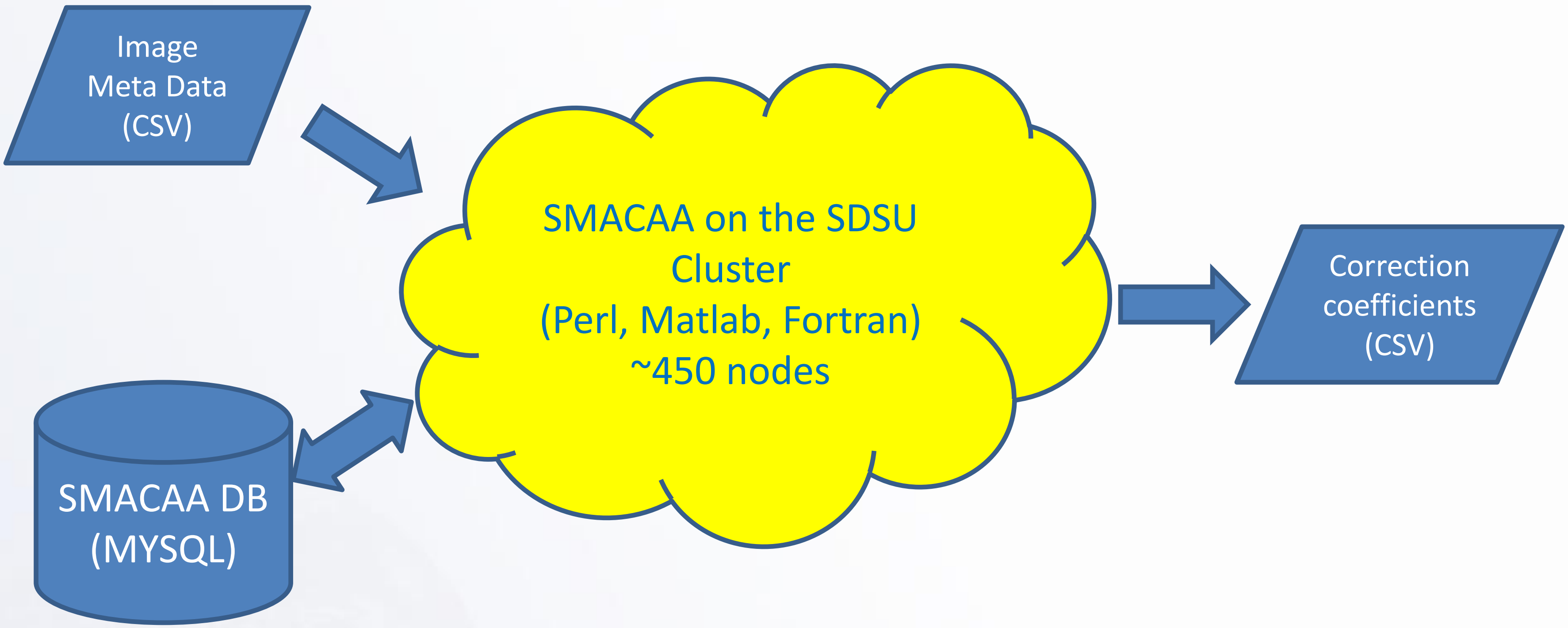
$$\text{GroundRef} = \frac{\text{TOARef} - \text{ATMOBias(band)}}{\text{ATMOGain(band)} + (\text{TOAREf} - \text{ATMOBias(band)}).* \text{ATMORho(band)}}$$

- At this time we are not trying to do pixel by pixel correction. While this is possible, it would require a bit of a rethink to the way data is generated.
    - This could impact scenes with high terrain variability
- These coefficients are currently being passed to Oregon State University for ingest into Google Earth Engine.

SMACAA on the SDSU Cluster

- As the workload was increasing, the need to scale up from one server to a cluster environment quickly became critical. This was a nontrivial change that required a bit of a rethink on how the system processed data. Namely, we needed to start processing images without images. This “blind processing” only requires the metadata for each scene, but that metadata must be accurate.
  - Metadata inputs include:
    - Lat/Lon, satellite view angles, date and time, and Spectral RSR
  - Metadata accuracy has posed problems, and is a concern.

SMACAA in Operation



- SMACAA currently lives on a 450 node cluster, meta file data is parsed out to each node, and each node communicates with the DB to pull out meteorological data. Run files are generated, these are parsed offline to create corresponding correction lookup files.
- Current runtime for a single SMACAA simulation ~ 30 minutes/node, so at full capacity an image is processed every 4 seconds

Current Status of the Processing

The good and the bad:

- Good
  - Currently ~ 1.2 Million "scenes" have been processed.
- Bad
  - due to initial "bad metadata files" there have been multiple runs of the same scenes
- Current goal is to do CONUS for all 5 sensors.
  - ~500,000 scenes ?
- As this process goes more operational need to communicate with EROS, and see if we can get a better set of metadata, and a better sense of what the total archive looks like. At this point we are simply pulling metadata from GLOVIS; is this the "best" way?

South Dakota State University
Image Processing Lab

Current stats

Sensor	SMACAA	CONUS Glovis	Global Glovis
MSS1	1702	37948	125737
MSS2	53445	58982	229223
MSS3	14215	25816	113937
MSS4	25388	26045	134917
MSS5	112890	286875	508317

Of the 1.2 million scenes due to reprocessing/updates/bad meta data we are left with ~125,000 images.

Think the bugs are worked out at this time, we understand the input data, all code has been "burned in" with the 1.2 million job run, and is automated. The remaining CONUS images should be processed in ~ 1 month.

Scaling out to the world: 3-6 months? (really think the above numbers only represent a portion of the data available at EROS, really big question mark for me)

Conclusion

- SMACAA is performing well in its new home on the SDSU cluster
- MSS Archive processing is still in progress, wish I could find the time to really polish this off (my "Friday night 'fun'")
- Scaling up for the future?
  - Having the system local solves problems
    - DB ingest is all PYTHON 😎
    - SMACAA process is a combination of PERL 😩 Matlab 😒and Fortran 😱 (modtran)
  - This means first either we need Matlab in it's new home, or the code needs to be converted (know of other "public" clusters with Matlab so may not be an issue)
  - Also Modtran would needs to be licensed for the new location (think this is probably simple)
  - So at the moment under same conditions would be easy to transfer to a larger cluster, just requires time and money **$$$**.

GO MSS!!

Thank You