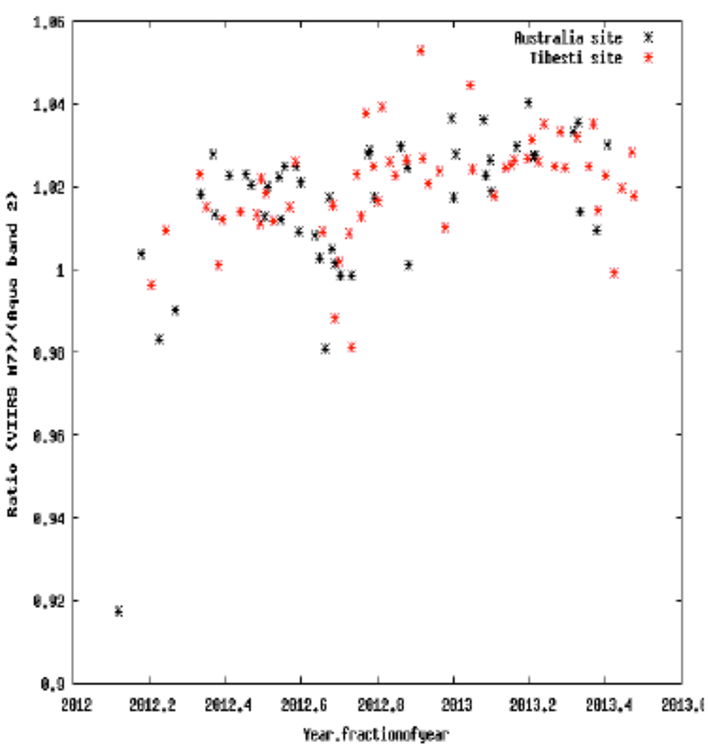
SENTINEL 2 PROGRESS REPORT -- Grant NNX11AE18G

Curtis Woodcock (Boston University)
Eric Vermote (now with NASA Goddard)

Summary
In preparation for the Sentinel 2 mission, the activities related to calibration, atmospheric correction and cloud/cloud shadow mask have continued using data from the Landsat sensor but also from the VIIRS, Formosat-2, SPOT4-Take5 and PROBA-V data. Additionally, we have worked on testing the cloud and cloud shadow detection capabilities of Sentinel-2 by using the subset of spectral bands that are available on Landsat 8. In particular, we have tested the relative performance for cloud and cloud shadow identification of the following 3 scenarios:

Heritage Landsat (optical and thermal)
Anticipate Sentinel-2 (optical and a cirrus band)
Landsat 8 (optical, a cirrus band, and thermal)

Calibration
The VIIRS and Aqua data have been used in an algorithm to inter-compare the calibration of VIIRS and Aqua. We verify the VIIRS calibration by comparing VIIRS SR products corrected for BRDF effect to the Aqua surface reflectance CMG product (0.05 degree resolution) corrected for BRDF effects. VIIRS is also projected on the CMG grid. The method has been updated to routinely monitor two sites (see figure 1a and 1b)



Site over Australia

Site over Tibesti

Figure 1: Example of VIIRS continuous calibration monitoring in band M7 (Near-Infrared) over two sites, one over Australia (50km x 50km) (black symbol) and one over Tibesti (~40km x 40km) (red symbol) (left side). The true color images of the sites are shown on the right side.

Atmospheric correction / Surface reflectance product

We have tested methods for comparison and adjustment of surface reflectance product acquired by different sensor for SENTINEL 2. Those methods rely on BRDF and spectral correction and have been previously tested on FORMOSAT 2 and MODIS for a dataset of about 100 cases during the 2006-2010 period (previous report). This inter-comparison has been extended to several other sensors, Landsat 5 and 7 (figure 2a-b), SPOT4-Take5 dataset which is a simulator for sentinel 2A,B (Figure 3) and PROBA-V (figure 4).

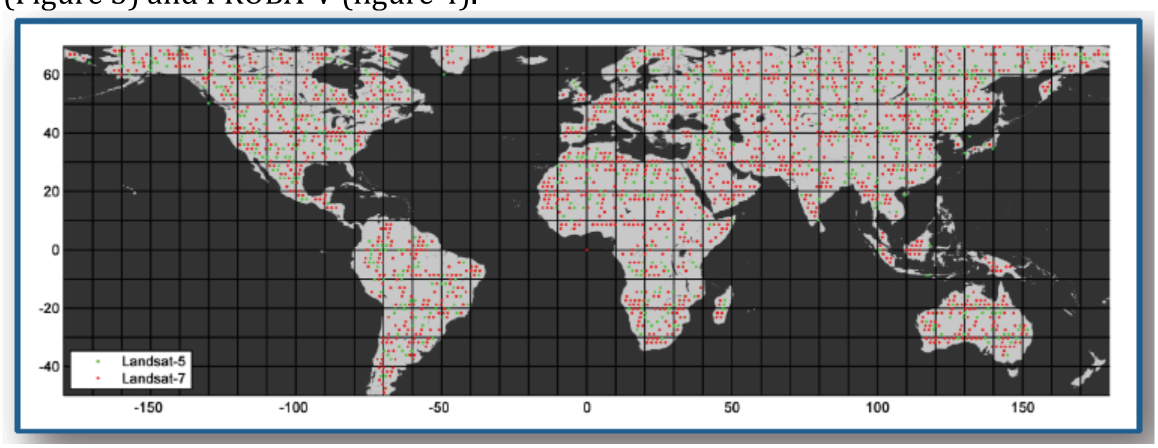


Figure 2a: Distribution of the 2125 sites used for the inter-comparison of the Landsat 5 and 7 and MODIS surface reflectance. A total of 6552 were analyzed during 2002-2013

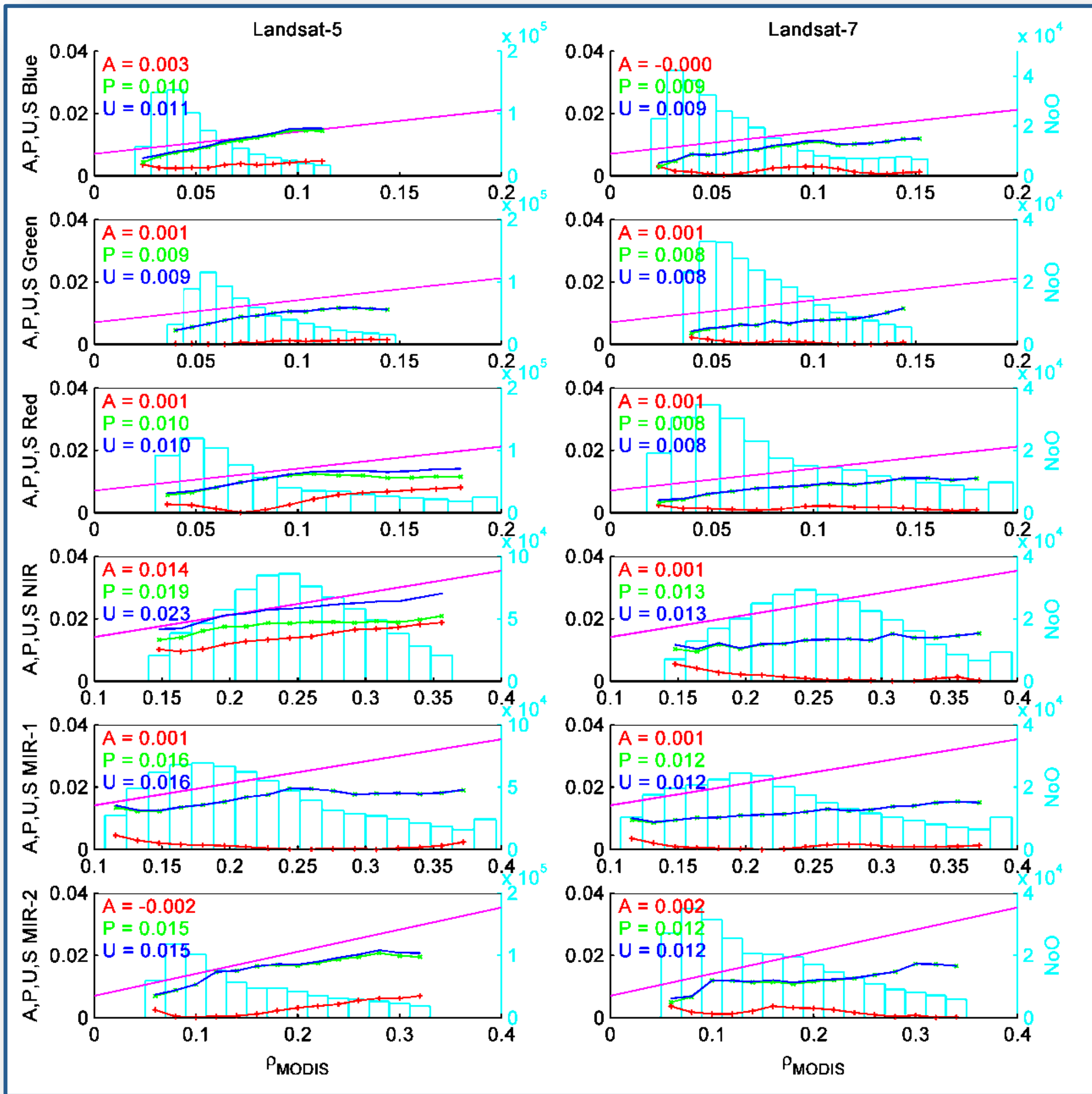


Figure 2b: Results of the inter-comparison Landsat 5and 7 with MODIS Terra summarized by the Accuracy (bias), Precision and Uncertainty (combined error) metrics.

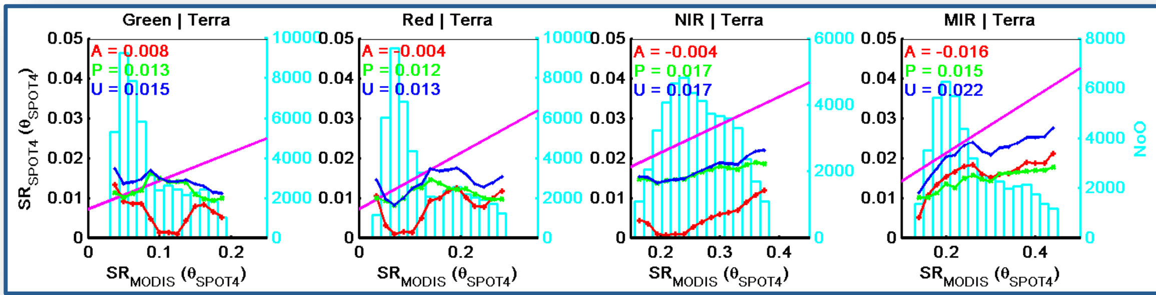


Figure 3: Results of the inter-comparison of SPOT4-Take 5 reflectance dataset with MODIS Terra summarized by the Accuracy (bias), Precision and Uncertainty (combined error) metrics. 750 images were acquired over 42 sites during February to June 2013 period to simulate the repeat cycle of Sentinel 2AB.

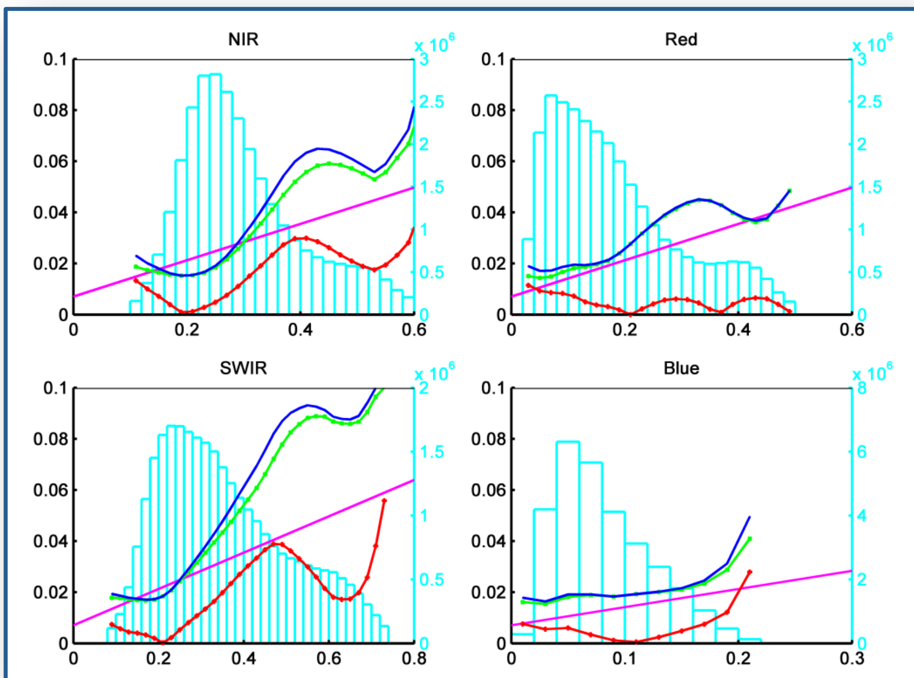


Figure 4: Results of the inter-comparison of Proba-V reflectance dataset with MODIS Terra summarized by the Accuracy (bias), Precision and Uncertainty (combined error) metrics. A global comparison was performed during 35 days at the end of 2013 .

Cloud/Cloud Shadow Mask

Following the validation technique used by the VIIRS Cloud Mask Team, we have started analyzing the performance of the cloud mask by using machup dataset between MODIS Aqua and the CALIOP instrument. The CALIPSO mission and in particular the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) provides a unique unbiased opportunity to evaluate cloud mask products. Despite its relatively narrow footprint (330m to 5km depending on the altitude of the layer sensed), CALIOP is acquiring data about 2 minutes after MODIS Aqua which makes it ideal for cloud mask evaluation. As a prototype, we have analyzed one day of global matchups (58409 matchups) acquired on March, 2, 2007 shown on figure 5. The CALIOP track is shown in red, note that we have limited our analysis to land only. Table 1 summarize the results of the analysis, for different scenarios we report the performance of MOD35 and ICM as the % of missed clouds (leakage), and the % of false detection (False Det.). The threshold used for cloud optical thickness was set to 0.1, as those sub-visible cloud are probably interpreted as aerosol and corrected as part as the atmospheric correction. From table 1, it is clear that ICM and MOD35 have comparable performances globally but when excluding high latitudes the leakage of ICM is better than MOD35. By design the ICM is conservative and tuned to the no-snow condition. Furthermore if excluding snow condition in ICM (using the quality flag provided in the product), the performance is further improved (see ICM Global Case1 collumn). Finally, by also using the cloud adjacent quality flag available in ICM (see ICM Global Case2 collumn) the leakage drops to close to 2%.
This prototype although promising is preliminary and a detailled protocol for use of the CALIOP data will need to be agreed on between different cloud mask teams. Also it will need to be extended to more cases at least covering a year of matchup data to establish regional /seasonnal statistically robust estimates of performances.

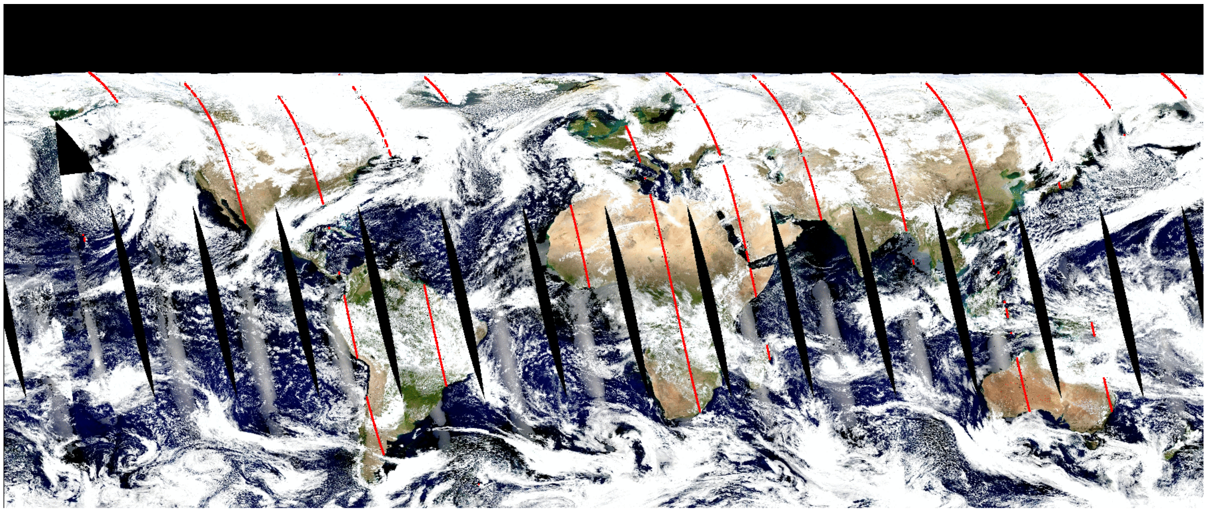


Figure 5: Aqua true color surface reflectance image for March, 2, 2007. The CALIOP track is shown in red, only matchups over Land are selected.

	MOD35 Global	MOD35 60S-60N	ICM Global	ICM 60S-60N	ICM Global Case1	ICM Global Case2
Leakage	6.1%	5.6%	5.8%	**4.0%**	**2.6%**	**2.1%**
False Det.	6.1%	6.4%	6.5%	6.7%	6.5%	6.5%

Table 1: Analysis of the performance of MOD35 and ICM under various scenarios. Global (Global), excluding latitude higher than 60N or lower than 60S (60S-60N), excluding cloud incorrectly detected as snow (ICM Global Case1) using the ICM snow quality flag, and finally further excluding ICM cloud adjacent quality flag (ICM Global Case2).

Cloud and Cloud Shadow Detection

We have tested different scenarios for cloud and cloud shadow detection by using subsets of the Landsat 8 bands. For what we call the “legacy Landsat” scenario, only the optical and one thermal band are used. For what we call the “anticipated Sentinel-2 capability” we have used the optical bands and the cirrus detection band. To compare the results we used the traditional Fmask algorithm for Landsat, and calibrated a new Fmask version for the anticipated Sentinel-2 capability. We ran them on the appropriate subset of bands from Landsat 8 for the 7 scenes indicated in Figure 6. The full results are shown for one scene from Oregon. Figure 7 is a simple color composite of the entire scene. Figure 8 shows the themal band, while Figure 9 shows the cirrus band on the left, and on the right it has been density sliced to illustrate the threshold used in the MODIS cloud detection algorithm. Notice that there are many cirrus clouds with values below the MODIS threshold. Figures 10 and 11 show the Fmask results for the heritage Landsat scenario and the anticipated Sentinel-2 scenario. Table 2 shows how the two results differ, and Table 3 shows an accuracy assessment of each scenario in the places where their results differ. These results for Oregon are shown as an example, as including all the tables and images would overwhelm the report.

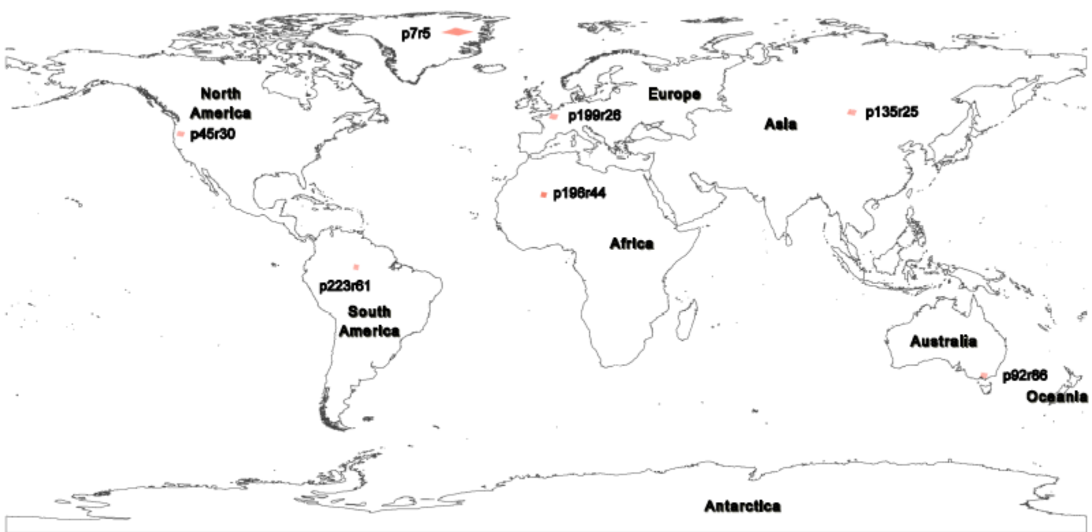


Figure 6. Locations of Landsat 8 images used to test cloud and cloud shadow capabilities anticipated for Sentinel-2.

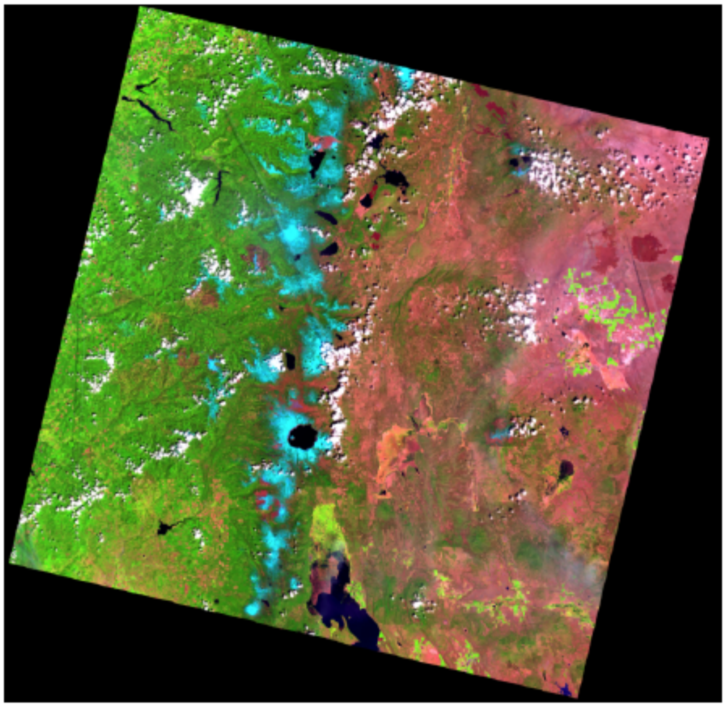


Figure 7. Landsat 8 image of Oregon with extensive coverage of cirrus clouds, which are largely invisible in this conventional color composite.

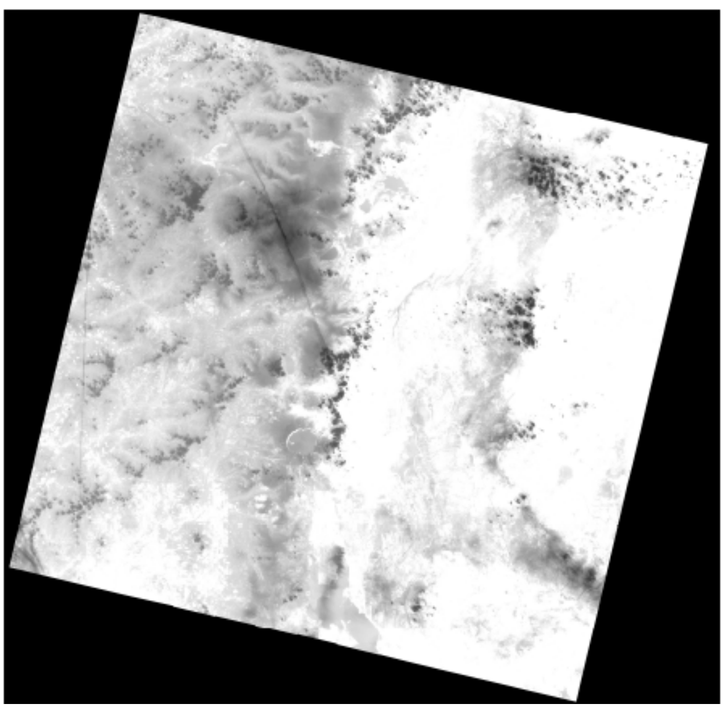


Figure 8. The thermal band from Landsat 8 for the Oregon scene.

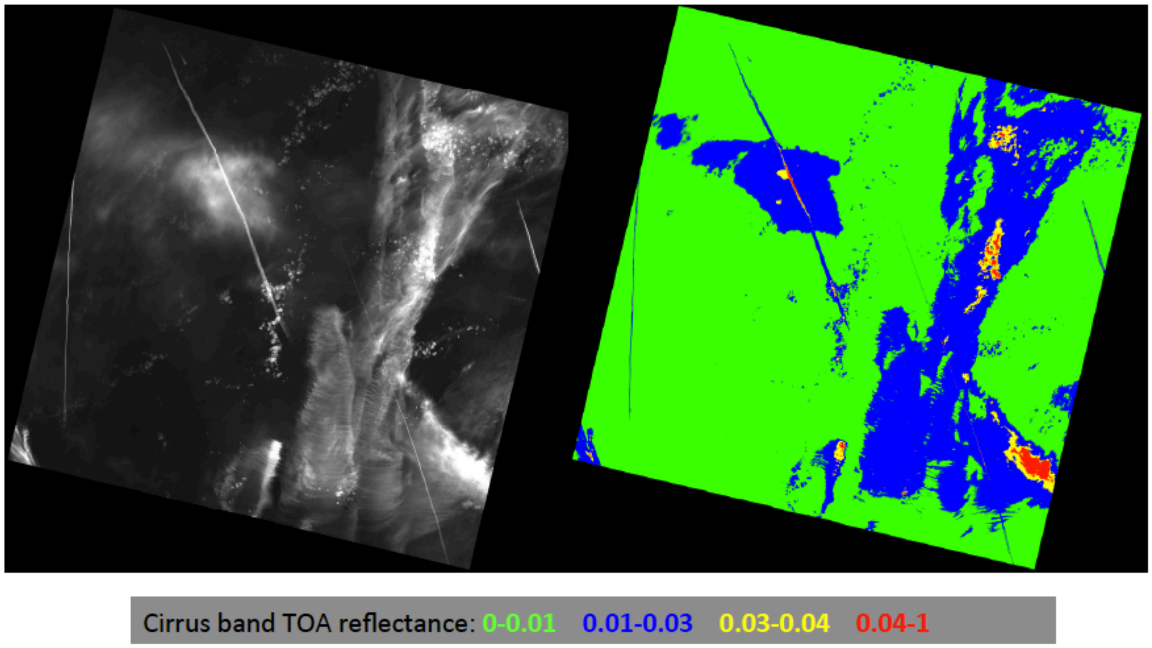


Figure 9. On the left is the cirrus band from Landsat 8, and on the right is a density slice that shows the areas that would be found using the threshold used in MODIS (red and yellow). Notice that there are many cirrus clouds that are below the MODIS threshold.

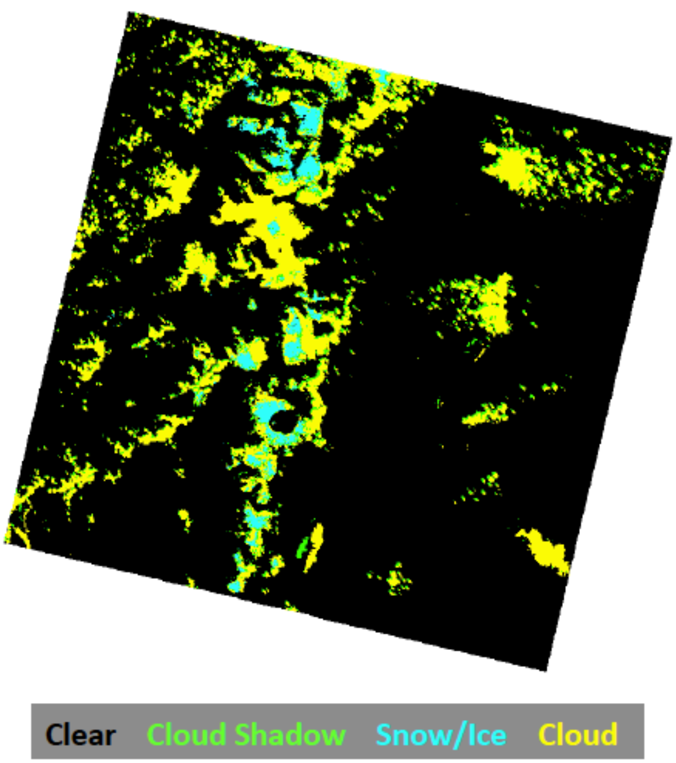
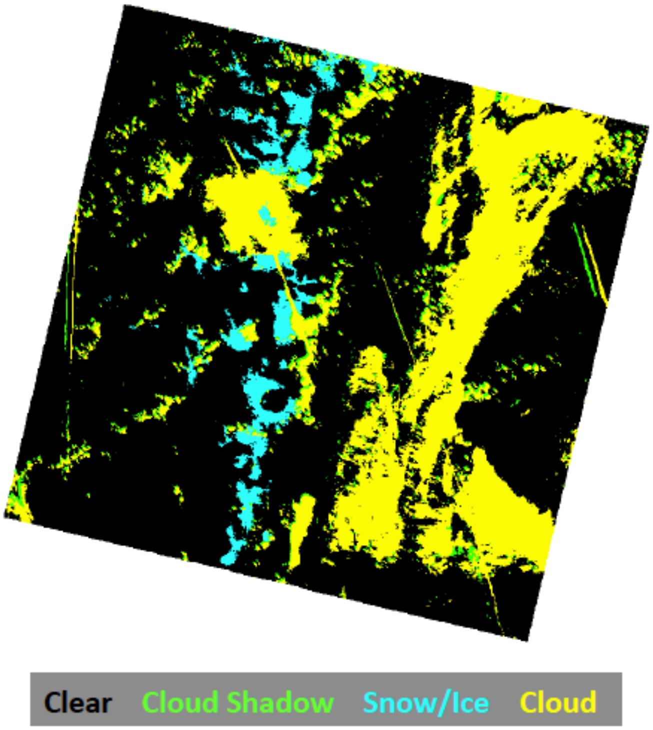


Figure 10. The heritage Landsat results from Fmask.



Sentinel / TM/ETM+	Clear Land	Clear Water	Cloud Shadow	Snow/Ice	Cloud
Clear Land	**66.00%**	0.00%	1.23%	0.06%	15.61%
Clear Water	0.00%	**1.25%**	0.02%	0.00%	0.10%
Cloud Shadow	1.20%	0.01%	**0.86%**	0.38%	0.54%
Snow/Ice	0.00%	0.00%	0.01%	**1.63%**	0.01%
Cloud	2.16%	0.01%	0.56%	1.64%	**6.74%**

Table 2. This table shows the difference between the heritage landsat cloud mask results and the anticipated result from Sentinel-2.

TM/ETM+ Scenario					
Ground Truth (Pixels)					
Class	Clear Land	Clear Water	Cloud Shadow	Cloud	User
Clear Land	**1**	0	0	190	0.52%
Clear Water	0	**0**	0	0	0%
Cloud Shadow	1	0	**0**	4	0%
Cloud	0	1	2	**1**	25%
Producer	50%	0%	0%	0.51%	Overall=**1%**

Sentinel Scenario					
Ground Truth (Pixels)					
Class	Clear Land	Clear Water	Cloud Shadow	Cloud	User
Clear Land	**0**	0	0	0	0%
Clear Water	0	**1**	0	0	100%
Cloud Shadow	1	0	**2**	14	11.76%
Cloud	1	0	0	**181**	99.45%
Producer	0%	100%	100%	92.82%	Overall=**92%**

Table 3. Accuracies of the two scenarios for places where the cloud and shadow results differ between the heritage Landsat scenario and the anticipated abilities of Sentinel-2 for the Oregon scene.

Rather than include all the tables for all the scenes, here is a summary of their accuracies in the places where the results for the heritage Landsat and anticipated Sentinel-2 capabilities.

	Heritage Landsat	Anticipated Sentinel-2
Oregon (path 45, row 30)	1%	92%
Australia (path 92, row 86)	25.5%	59.5%
Africa (path 196,row 44)	7.5%	78%
Mongolia (path 135, row 25)	17%	73%
Paris, France (path 199, row 26)	25.5%	70%
Amazon (path 233, row 61)	3%	91%
Polar (North path 7, row 5)	16.5%	76%

Results and Conclusions:

From these results it is easy to see that cloud and cloud shadow detection will be easier using Sentinel-2 data, particularly when cirrus clouds are present. The Landsat 8 scenario (optical, cirrus and thermal) is preferable, but for cloud and cloud shadow detection if you had to choose between a thermal band or a cirrus detection band, the latter would be better.