Science & Technology Facilities Council
Rutherford Appleton Laboratory

A Study of the Cloud Flagging in Reprocessed (A)ATSR Data

Caroline Cox
October 2013

Contents

Reference Documents

Document Ref	Document
AD1	IDEAS-VEG-OQC-MEM-1158
AD2	Impact of the new gross cloud test thresholds on the aatsr cloud flagging
AD3	Improvements to the AATSR IPF relating to Land Surface Temperature Retrieval and Cloud Clearing over Land

1 Introduction

The third reprocessing of the (A)ATSR dataset was performed in mid-2013. The updates and improvements to the dataset include the following [AD1]:

- Updated IPF processor, v6.05
- New SST (Sea Surface Temperature) ADF incorporating the SST coefficients from the ARC project.
- New CH1 ADF incorporating improved colocation between nadir and forward views and improved absolute geolocation.
- New CL1 ADF incorporating improved GCT (gross cloud test) coefficients [AD2] and minor correction to the 1.6 Micron Histogram Cloud Test (correction of range_weight_limit parameter to 2.5 from 40).
- New PC1 ADFs (processing configuration data) to support IPF 6.05.
- New DEM (Digital Elevation Model) ADF which contains validity range spanning the whole Envisat mission.
- New dynamic VC1 files which contain a long-term drift correction will improve the visible calibration.
- New OSV files (orbit state vector) files will be used over the operationally used AUX_FRO files as they are of higher resolution and more complete.

Cloud testing is routinely performed on the Level 1 brightness temperature or reflectance data. Table 1 below provides an overview of the various cloud tests. The tests are given in the order in which they are applied, except for the snow and sun glint tests which are not strictly cloud tests, but the results of these are output in the same file as the cloud flags. The table shows which channels and views the cloud tests use, the conditions in which they are applied and the dependency on the results of other tests. Several of the changes to the (A)ATSR dataset that are listed above are bound to have an effect on these cloud flags, most obviously the new CL1 file for the GCT and the VC1 file affecting the visible calibration. However, the geolocation, colocation, a different DEM plus others may also alter the cloud flag of individual pixels.

The aim of this study is to address the following:

- Check whether the changes made to the cloud tests have been implemented correctly in the v2.1 reprocessed data (the GCT in particular)

- Check the reprocessing has not had any adverse effects on the cloud flags

- Quantify and assess the effect of the reprocessing on the cloud flagging.

Cloud Test	**Channel used**							**Views**		**Day/night**	**Dependence on other tests**	**Surface**
	12	**11**	**3.7**	**1.6**	**0.87**	**0.67**	**0.55**	**Separate**	**combined**			
Gross	X							X		Both	None	Both
Thin cirrus	X	X						X		Both	None	Both
Med/hi	X		X					X		Night	None	Both
Fog/low stratus		X	X					X		Night	None	Both
11 SCT	X	X						X		Both	GCT Thin cirrus, Med/hi Fog	Sea
1.6 histogram				X				X		Day	Sunglint GCT Thin cirrus 11 SCT	Sea
1.6 SCT[1]				X				X		Day	GCT Thin cirrus	Sea

											11 SCT 1.6 histogram	
11/12	X	X							X	Day	None	Sea
11/3.7		X	X						X	Day	None	Sea
IR histogram	X	X						X		Day	All but visible	Sea
Visible					X	X	X	X		Day	None	Land
Sunglint	NA	NA	NA	NA	NA	NA	NA	X		Day	None	Both
Snow		X		X	X		X	X		Day	None	Both

Table 1 An overview of the cloud tests performed on (A)ATSR data.

[1] If sun glint is detected in the 1.6um histogram test, a spatial coherence test is applied to the pixels in the sub-area. The 1.6 SCT can therefore be thought of as part of the 1.6 histogram test, however as it still has its own flag, it will still be treated independently.

2 Overview of cloud flag changes in reprocessed data

A one month sample of data has been analysed for each of the (A)ATSR instruments. The current version of each dataset, v2.0, has been compared with the reprocessed data, v2.1 using various software tools. Statistics on the differences between versions 2.1 and 2.0 have been calculated and particular frames of interest have been identified. For each instrument, a summary is provided of the overall changes to the cloud flag in reprocessed data in sections 2.1 to 2.3.

2.1 AATSR

The AATSR data from March 2012 was selected as the sample month as it was the first month of AATSR reprocessed data to be released. There were 35165 image frames used in this study.

Overall in v2.1 **nadir** images, there are 4.5% **fewer** pixels that have been classified as cloudy when compared with v2.0. In the **forward** images, there were 5.2% **fewer** pixels that have been classified as cloudy when compared with v2.0.

We can look at the percentage difference in the number of cloudy pixels for the v2.1 and v2.0 images as a function of cloud test where the percentage difference is defined below.

Percentage difference = (number of v2.1 cloudy pixels – number of v2.0 cloudy pixels)/number of v2.0 pixels

Figure 1 reveals that in the nadir images, the NDVI visible flag has been most affected by the reprocessing. It shows that the 1.6 SCT, GCT and the 11,12um histogram test act to identify more pixels as cloudy in v2.1, whereas all the other tests find fewer pixels to flag in v2.1. In the forward images, we see a similar trend to the nadir, except that there are significantly more sun-glinted pixels identified, the sign of the 1.6um histogram differences have changed and there are significantly more cloudy pixels identified by the 1.6 SCT in the forward images.

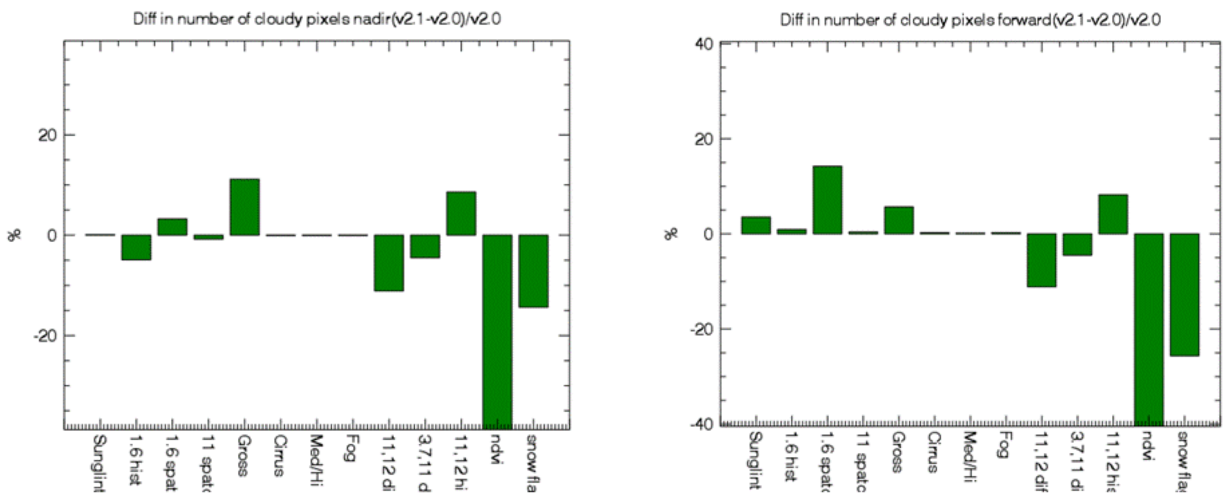

Figure 1. Bar plots of differences in the percentage of cloudy pixels in the v2.0 and v2.1 AATSR dataset for nadir and forward views.

The differences in the number of cloudy pixels in the images gives a good overall idea of how the reprocessing has affected the resulting flags, however, it does not always give the full extent of the changes. For example, v2.0 may have flagged an area of cloud of *n* pixels, and v2.1 may have flagged a completely different area of cloud covering *n* pixels in the same image frame. The straight difference would falsely lead one to conclude that there had been no effect from the reprocessing, as it does not reveal that two completely different areas had been flagged. In order to identify this, we need to look at the number of pixels flagged in v2.0 that have not been flagged in v2.1, and vice versa. We shall call these quantities the number of uniquely flagged pixels.

The total number of uniquely flagged pixels in the images as a percentage of the total number of image pixels is 0.30% (0.37%) in v2.0 and 0.12% (0.15%) in v2.1 for nadir (forward). Therefore 0.42% (0.52%) of pixels in each set of nadir (forward) images has changed classification with the reprocessing. This number does not seem significantly high. Even if any single cloud test is showing a big difference in the classification of cloud, the effect on the final cloud flag will be minimised as it only takes one test to have already flagged a pixel as cloudy for the pixel to be classified as cloudy. However, for those cases where the cloud flag has changed, we need to know why it has changed and whether this leads to an improvement in the cloud flag or not.

In the bar plot in Figure 2, plotted in red are the number of pixels that have been flagged in v2.0 but not in v2.1, as a function of each test. In blue, we plot the number of pixels that have been flagged in v2.1 but not in v2.0. Plotted in pink is the sum of these numbers which gives the total number of pixels that have changed classification. In orange, the difference between the numbers of uniquely flagged pixels in each image is shown. If there are near equal red and blue pixels then it is likely that that the cloud edges are being reclassified where for example changes in colocation have taken place, rather than large areas of cloud changing classification as a result of calibration improvements for example. Specific cases will be investigated.

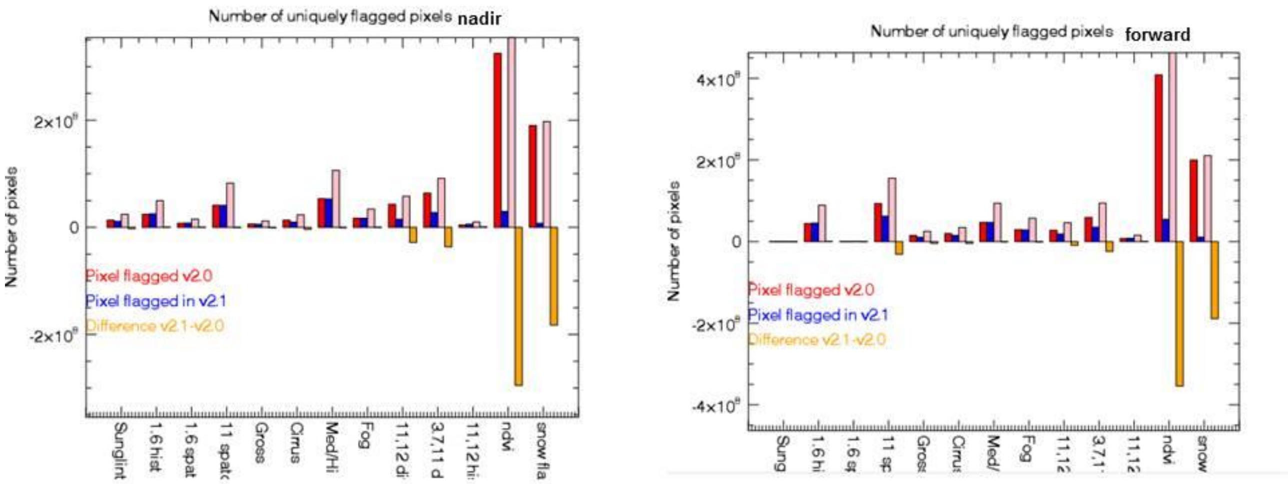

Figure 2. Bar plots of numbers of uniquely cloudy pixels in the v2.0 and v2.1 AATSR dataset for nadir and forward views. Red bars shows the pixel is flagged in v2.0 but not in v2.1, blue bars shows the pixel is flagged in v2.1 but not in v2.0, orange shows the difference in red and blue, and pink shows the sum.

2.2 ATSR-2

The ATSR-2 data from March 2003 was selected as the sample month so that it is from the same time of year as the AATSR dataset. There were 32936 image frames used in this study. The v2.1 image files are a different length to the v2.0 files as in the reprocessed version the data is now provided between ANX to ANX.

Overall in v2.1 **nadir** images, there are 0.6% **more** pixels that have been classified as cloudy when compared with v2.0. In v2.1 **forward** images, there are 0.4% **more** pixels that have been classified as cloudy when compared with v2.0. The sign of the differences are not expected to be consistent with those of AATSR as a different calibration has been applied to the instruments in the reprocessed datasets and the GCT improvements have not been applied to ATSR-1 or 2.

In Figure 3 the percentage difference in the number of cloudy pixels for the nadir v2.1 and v2.0 images as a function of cloud test is shown. The NDVI visible cloud test, like for AATSR, is the most affected by the reprocessing. The snow flag is also significantly altered, particularly in the forward images. In the nadir, the 11/12 diff test, 3.7/11 diff test and the 11/12 histogram test are also significantly changed, and less so in the forward images. Figure 4 also shows the effect of the reprocessing using the number of uniquely flagged pixels.

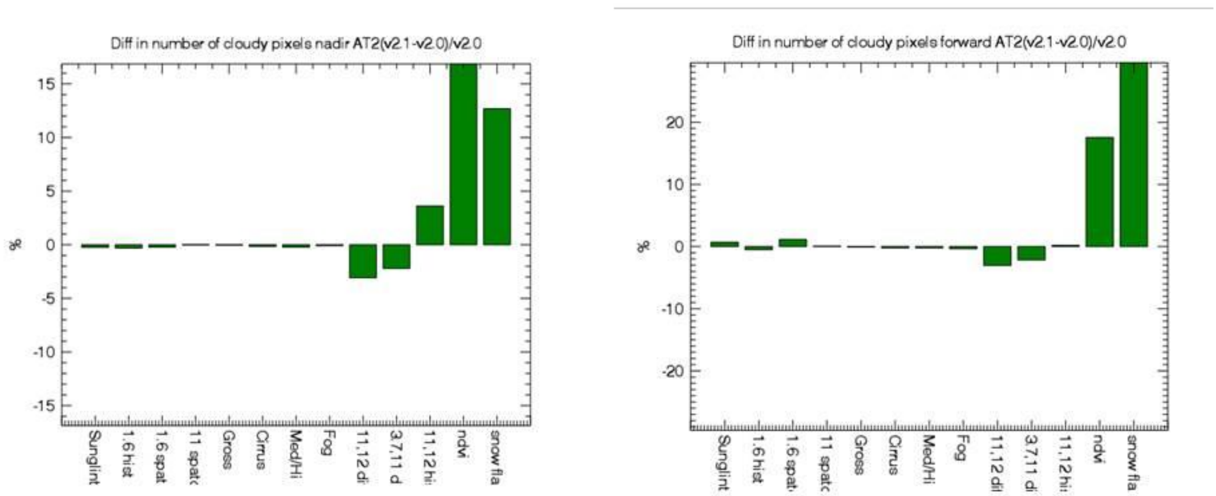

Figure 3 Bar plots of differences in the percentage of cloudy pixels in the v2.0 and v2.1 ATSR-2 dataset for nadir and forward views.

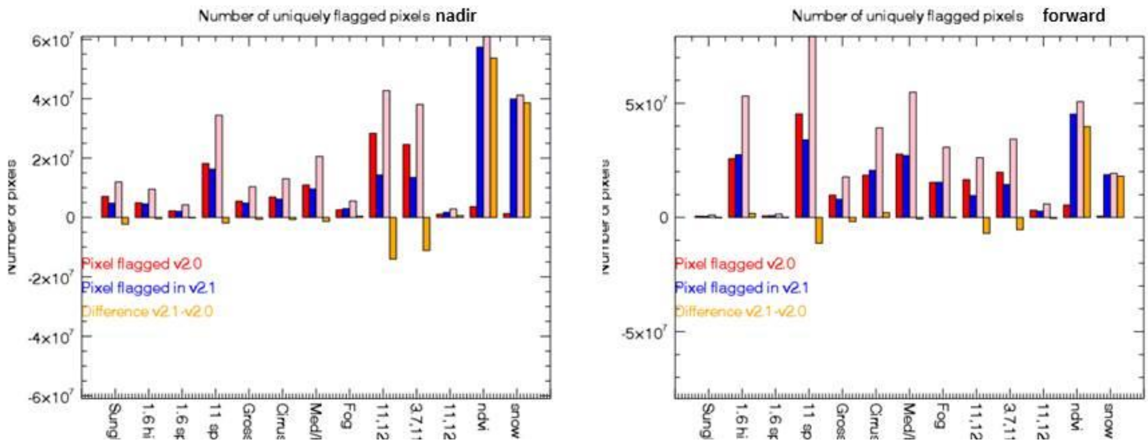

Figure 4. Bar plots of numbers of uniquely cloudy pixels in the v2.0 and v2.1 ATSR-2 dataset for nadir and forward views. Red bars shows the pixel is flagged in v2.0 but not in v2.1, blue bars shows the pixel is flagged in v2.1 but not in v2.0, orange shows the difference in red and blue, and pink shows the sum.

2.3 ATSR-1

The month of March from 1994 has been selected as the sample month to use for the ATSR-1 cloudy flagging study. ATSR-1 differs from ATSR-2 and AATSR in that it only has the 1.6um visible channel. Therefore, the visible cloud test and snow test is not applied to this dataset and we can expect fewer changes in the cloud flagging as a result. Also, the 3.7um channel failed in May 1992 and so there is no data from this channel after this point. This means the medium/high cloud test, fog and 11/3.7um tests were not applied to the dataset used here.

Overall in v2.1 **nadir** images, there are 0.28% **fewer** pixels that have been classified as cloudy when compared with v2.0. In the **forward** images, there are 0.42% **fewer** pixels that have been classified as cloudy when compared with v2.0. The most significant changes for ATSR-1 data are to the only tests which use the visible channel data, the 1.6um histogram test and the 1.6um SCT. The 11/12um difference test also shows significant changes as can be seen in Figure 5 and Figure 6.

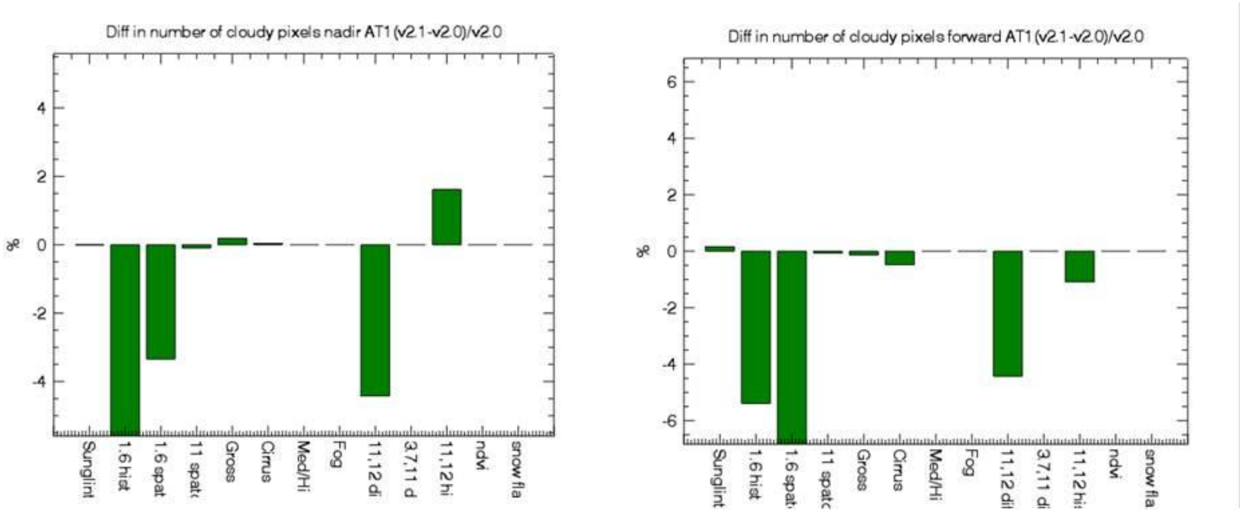

Figure 5 Bar plots of differences in the percentage of cloudy pixels in the v2.0 and v2.1 ATSR-1 dataset for nadir and forward views. Note that the med/hi, fog, 3.7/11 and NDVI tests (and also the sunglint and snow flag) are not applicable for this month of ATSR-1 data.

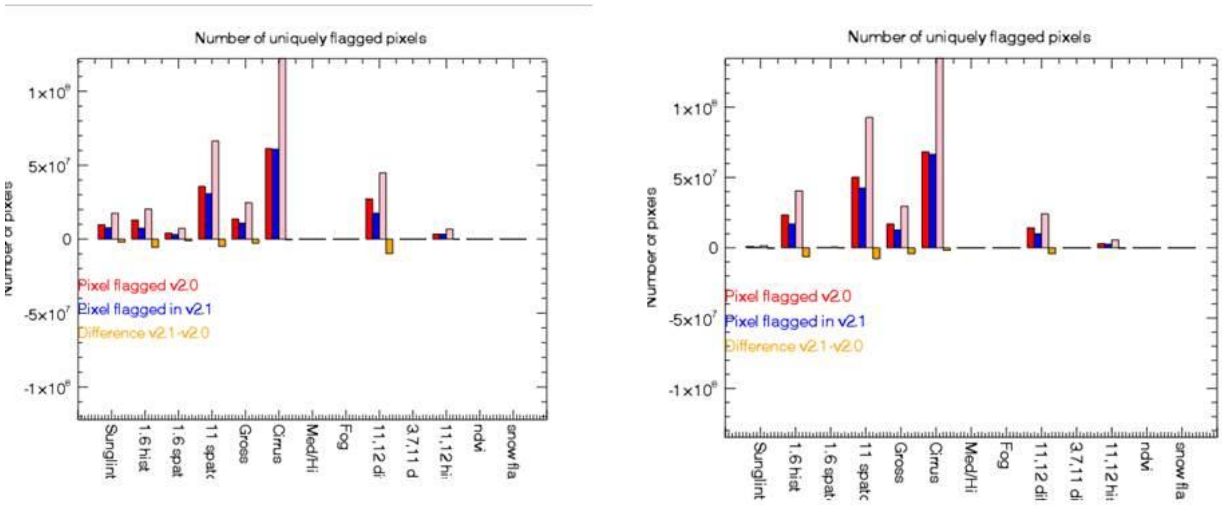

Figure 6. Bar plots of numbers of uniquely cloudy pixels in the v2.0 and v2.1 ATSR-1 dataset for nadir and forward views. Note that the med/hi, fog, 3.7/11 and NDVI tests (and also the sunglint and snow flag) are not applicable for this month of ATSR-1 data.

3 Investigating individual cloud tests

The changes to the cloud flagging of each individual cloud test are now investigated. They are described in order of their application.

3.1 Gross cloud test

In the AATSR test dataset, the GCT flag displayed differences in cloud flagging between v2.1 and v2.0 of 11% (6%) in the nadir (forward) images. For ATSR-2 and ATSR-1 data there was less than 0.1% differences in the nadir and forward images. The changes were more significant for AATSR compared with the other instruments as this was the only one for which the thresholds in the GCT were altered. Improvements to the ATSR-1 and ATSR-2 GCT are expected for the next reprocessing.

The changes that have been applied to the AATSR GCT since reprocessing involve re-calculating the 12um forward and nadir view thresholds [AD2]. An improved climatology was used when calculating these thresholds and at some latitudes this has caused the threshold to change by of the order of 10K. The thresholds are plotted for March in Figure 7. For this month, the thresholds have in general increased in v2.1 compared with v2.0, meaning overall more cloud should be detected than before as there should be more pixels with BTs below the new higher threshold temperature. This is indeed what is found and the statistics for AATSR data are summarised in Table 2.

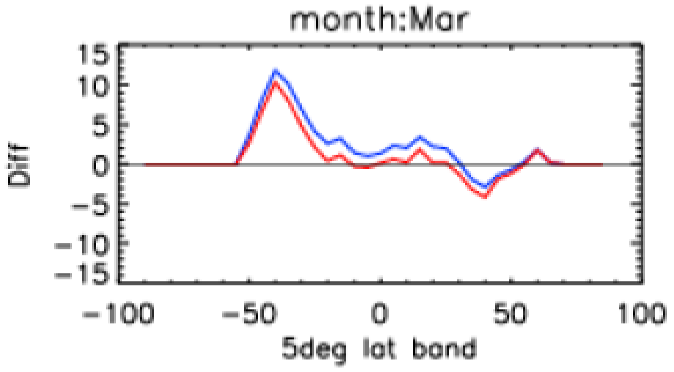

Figure 7. A plot showing the difference between the GCT thresholds in the v2.1 and v2.0 (v2.1-v2.0) processor. Blue is for nadir, red is for forward view. This plot is for March thresholds, the test month under study.

	V2.0		**V2.1**		**Difference**	
	Nadir	***Forward***	***Nadir***	***Forward***	***Nadir***	***Forward***
Pixels flagged by GCT	39.60%	40.33%	44.02%	42.63%	11.15%	5.69%
Uniquely flagged pixels by GCT ONLY	0.025%	0.058%	0.020%	0.043%	-22.13%	-25.40%

Table 2. A summary of the changes to the GCT flag in AATSR data.

Overall, more pixels have been flagged by the GCT in v2.1 compared with v2.0 in the nadir and forward images, in agreement in what is expected. Looking at cases where the pixels have been flagged in one image but not the other by only the GCT, we see that there are fewer uniquely flagged pixels in v2.1 for both the nadir and forward views. This implies therefore that although more pixels have been flagged using the new GCT, of the flags which will have an impact on the final cloud flag, the new GCT has actually been responsible for reducing the amount of cloud flagged. The new test is changing the result of those pixels that in v2.0 the GCT was the only test setting the cloud flag, perhaps incorrectly so. Now the test has been improved and those flags have been changed to clear.

We can look at the percentage of pixels that have changed classification as a function of latitude to see it follows the latitude trend of the changes to the BT threshold shown in Figure 7. If we calculate the difference in the number of cloudy pixels determined by the GCT in the v2.0 and v2.1 images within 5 degree latitude bins and find the value of these as a percentage of the total number of pixels in the image, the results are plotted in Figure 8 for the nadir and forward views. There is significant spike at around -40 to -50 degrees latitude, in line with the maximum difference between v2.0 and v2.1 GCT thresholds shown in Figure 7. The forward view data also shows a smaller peak at around 40 degrees latitude to coincide with the negative difference in the GCT thresholds at this latitude which is more pronounced for the forward view than the nadir view.

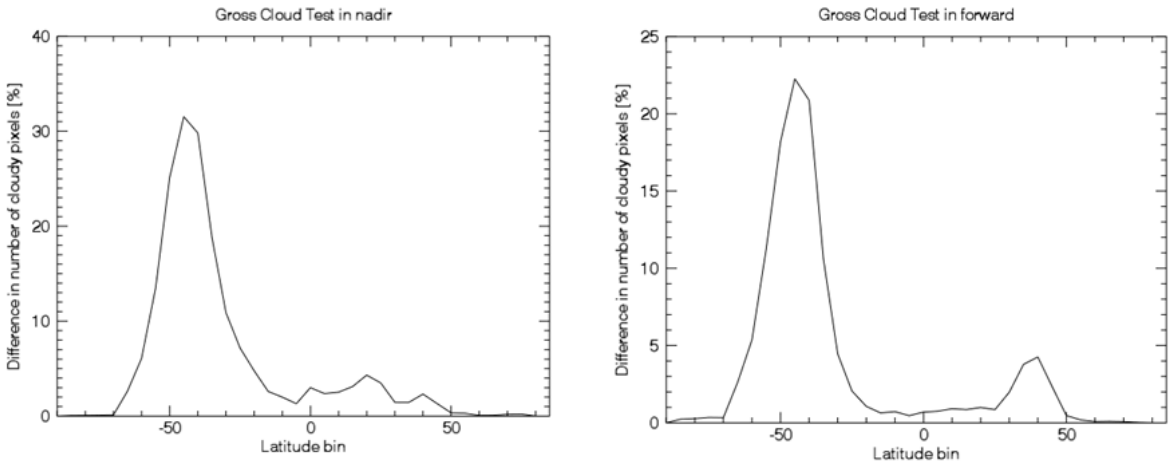

Figure 8. The latitude dependence of the impact of the new GCT on reprocessed cloud flagging is shown here. The difference in the number of cloudy pixels between v2.1 and v2.0 is plotted against latitude.

3.1.1 Examples of images showing significant changes in flag

The scenes for which only the GCT has flagged any uniquely cloudy pixels in the v2.0 and v2.1 image have been identified and the most significant areas of change are presented here to establish whether the new GCT has improved the cloud clearing. Most cases do seem to show that the new GCT moves in the correct direction.

In the 5 cases presented below, screenshots of the Cloudview and QUASAR tools are shown. Within the Cloudview tool, the image to the left is the v2.0 frame, with brown indicating land, white indicating the overall cloud flag and pink showing the GCT flag. In some cases, the image to the right shows the cloud flags of the v2.1 image where the same colour scheme applies, but in most cases in the right hand side image of the Cloudview tool, the difference between the v2.0 and v2.1 images is shown, with red showing areas where the v2.0 image has cloud but v2.1 does not, and blue showing where the v2.1 image has cloud but v2.0 does not. The QUASAR tool simply shows an image of a selected channel.

1. **North Pacific**

Image: `ATS_TOA_1PRUPA20120315_232539_000065273112_00360_52529_4062.N1`

Lat: 48
Lon: 160
Start pixel: 15360
Surface: Sea
Net difference in number of uniquely cloudy pixels flagged by this test alone: `-34357`
Nadir

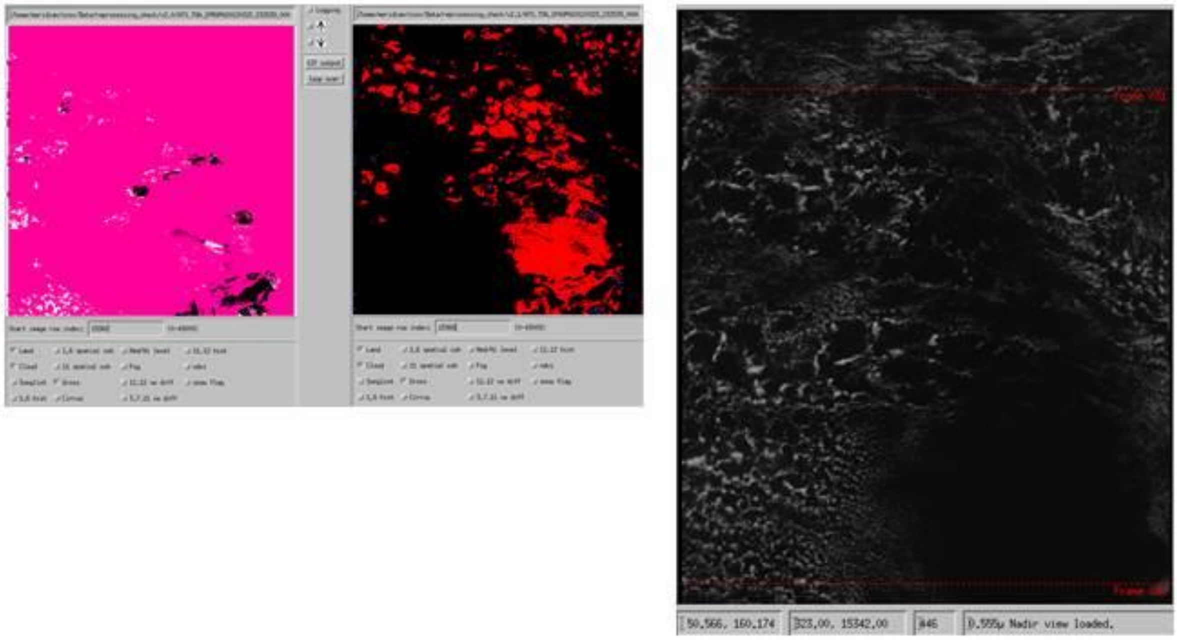

Figure 9. The frame of interest is shown in the cloudview tool on the left and the visible 0.55um channel in the QUASAR tool on the right. At 48 degrees latitude, the nadir GCT threshold has been reduced and therefore we would expect less cloud to be identified at this latitude. This frame is in agreement with what we expect and demonstrates that the new GCT has probably correctly determined that there is a clear area in the frame. Based upon the appearance of the visible channels, there does not look to be any cloud where new GCT has removed the cloud flag.

2. West Coast of Japan

Image: `ATS_TOA_1PRUPA20120322_004552_000065273113_00016_52616_4151.N1`
Lat: 43.5
Lon: 138
Start pixel: 15872
Surface: Coast
Net difference in number of uniquely cloudy pixels flagged by this test alone: `-33749`
Nadir

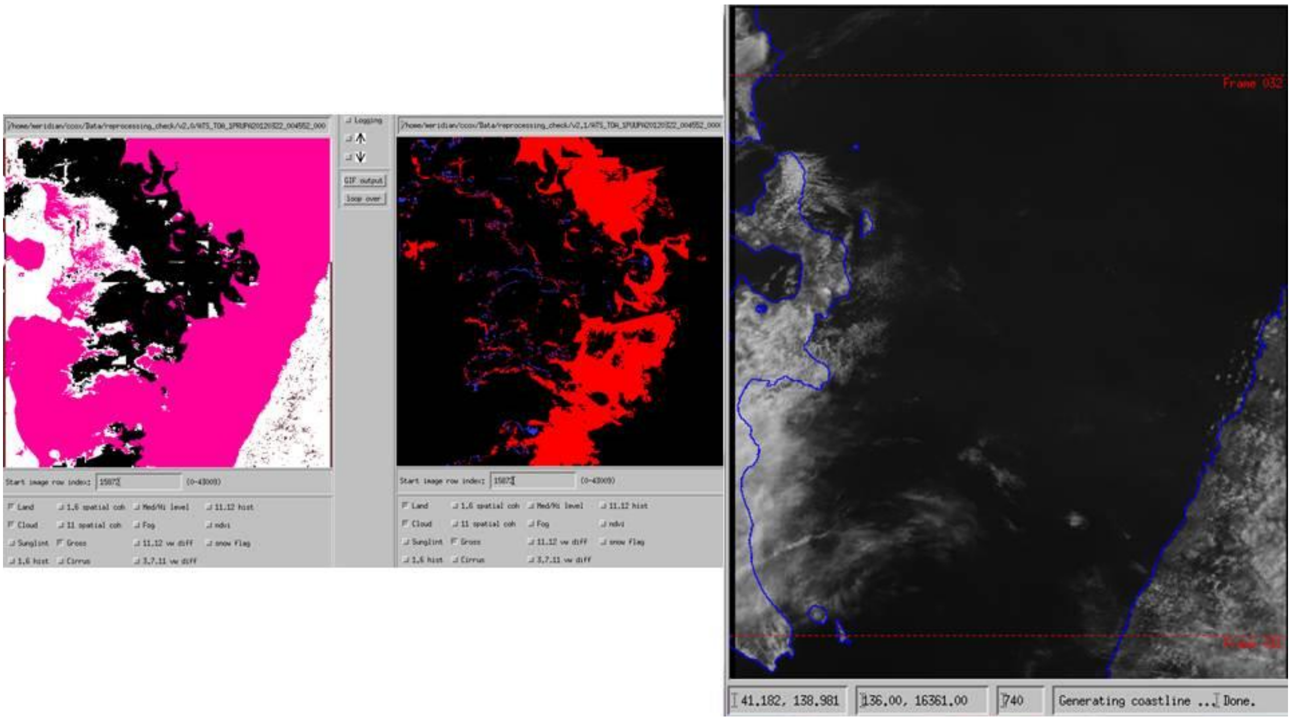

Figure 10. The frame of interest is shown in the cloudview tool on the left and a visible channel in the QUASAR tool on the right. Again, at 43.5 latitude, the nadir GCT threshold has been reduced and therefore we would expect less cloud to be identified at this latitude. This frame is in agreement with what we expect. The visible images look clear where the cloud has been removed by the new GCT, demonstrating the new GCT has likely correctly removed the flag in the reprocessed image.

3. Indian Ocean

Image: ATS_TOA_1PRUPA20120303_040254_000065273112_00176_52345_3741
Lat:-41
Lon: 70
Start pixel: 25088
Surface: Sea
Net difference in number of uniquely cloudy pixels flagged by this test alone: 31137
Nadir

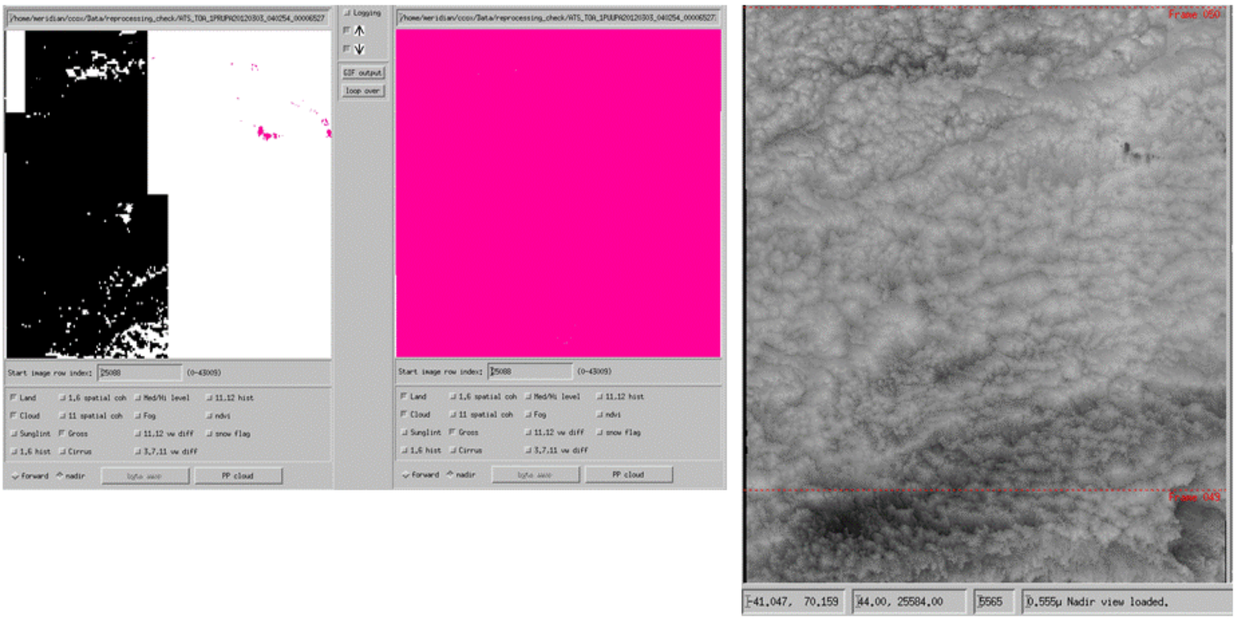

Figure 11. The frame of interest is shown in the cloudview tool on the left and the 0.555um visible channel in the QUASAR tool on the right. This case shows a large area to the left of the image that had not been identified as cloudy by any test in v2.0. In the reprocessing, the whole frame is flagged as cloudy, and the GCT is the only test to have flagged the previously clear half of the frame as shown by the pink colour. On visual inspection of a visible channel, it does obviously appear that there is cloud all over the scene. For this month and latitude, the difference in the GCT threshold is at its greatest and so finding a significant difference here can be expected.

4. West coast Chile

Image: ATS_TOA_1PRUPA20120322_140743_000065273113_00024_52624_4160.N1
Lat:-20
Lon: -78
Start pixel: 23040
Surface: Sea
Net difference in number of uniquely cloudy pixels flagged by this test alone: 30904
Nadir

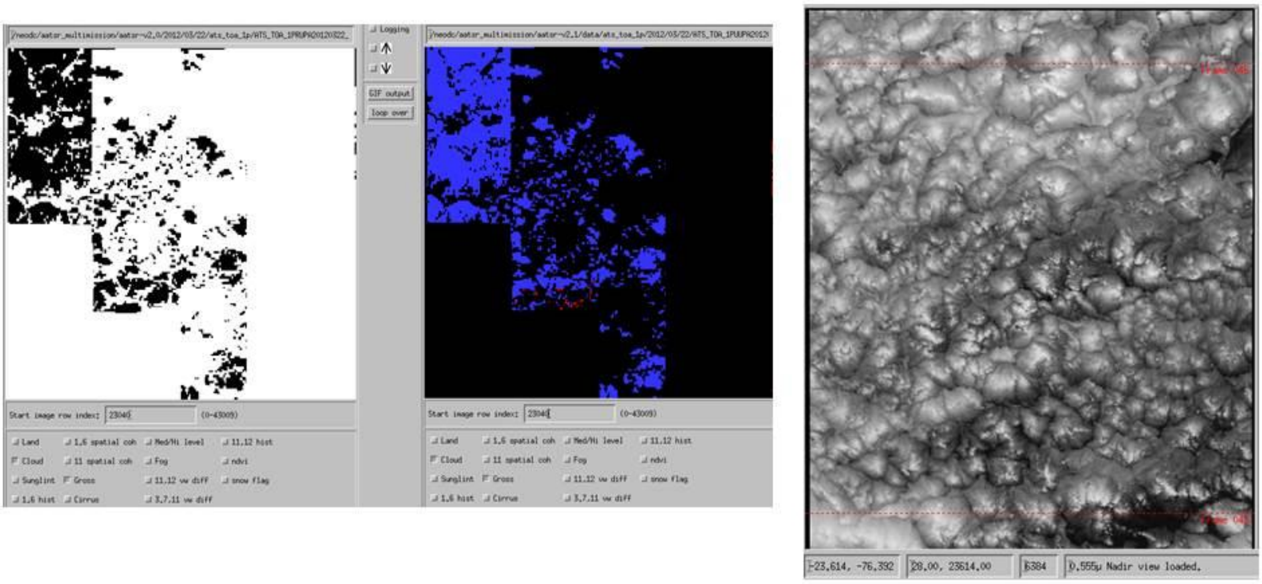

Figure 12 In this case there is an area of cloud flagged in v2.1 that was not there in v2.0. The visible channels clearly show that there is cloud over the scene.

5. East China Sea Coast

Image: ATS_TOA_1PRUPA20120315_014237_000065273112_00347_52516_4049.N1
Lat: 38
Lon: 122
Start pixel: 16384
Surface: Coast
Net difference in number of uniquely cloudy pixels flagged by this test alone: -29255
Nadir

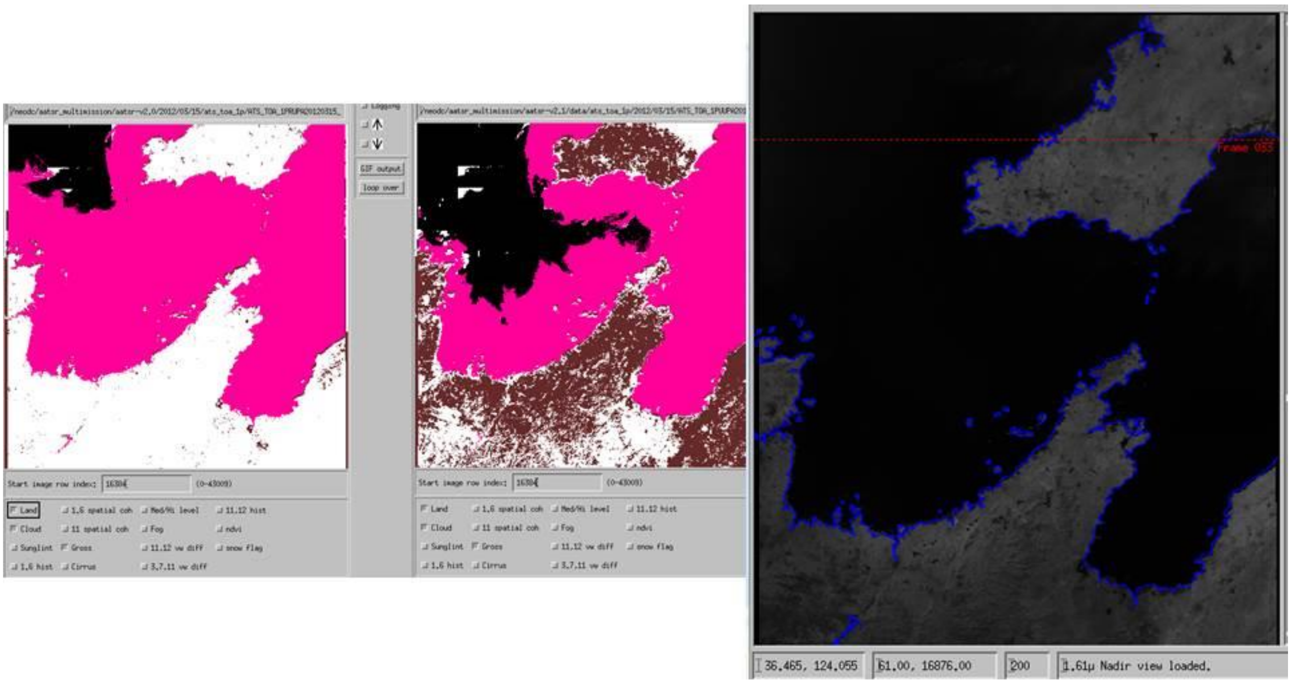

Figure 13. In these images we can see that there is less cloud in v2.1. All visible images look clear, and it seems like there is even more cloud that could have been taken away by the cloud tests.

3.1.2 Cases from November 2011

The original study which proposed the new GCT thresholds (AD2) tested them using prototype processor (PP) on several scenes from 11 Nov 2011 where SST anomalies were present. Several of the same frames used in that study have now been identified in the reprocessed AATSR dataset and are compared with the frames from the PP to ensure that they are in agreement. As there are other changes made to the v2.1 data that will not have been made to the PPv2.1 data such as co/geolocation, they will not be in exact agreement and will show differences around the cloud edges, but we would expect the gross area of cloud to be the same. Note that the image start pixels are not the same for the PP images and the OP images, but the same geographical areas are shown.

1. Case 1 in AD2, Sea of Okhotsk.

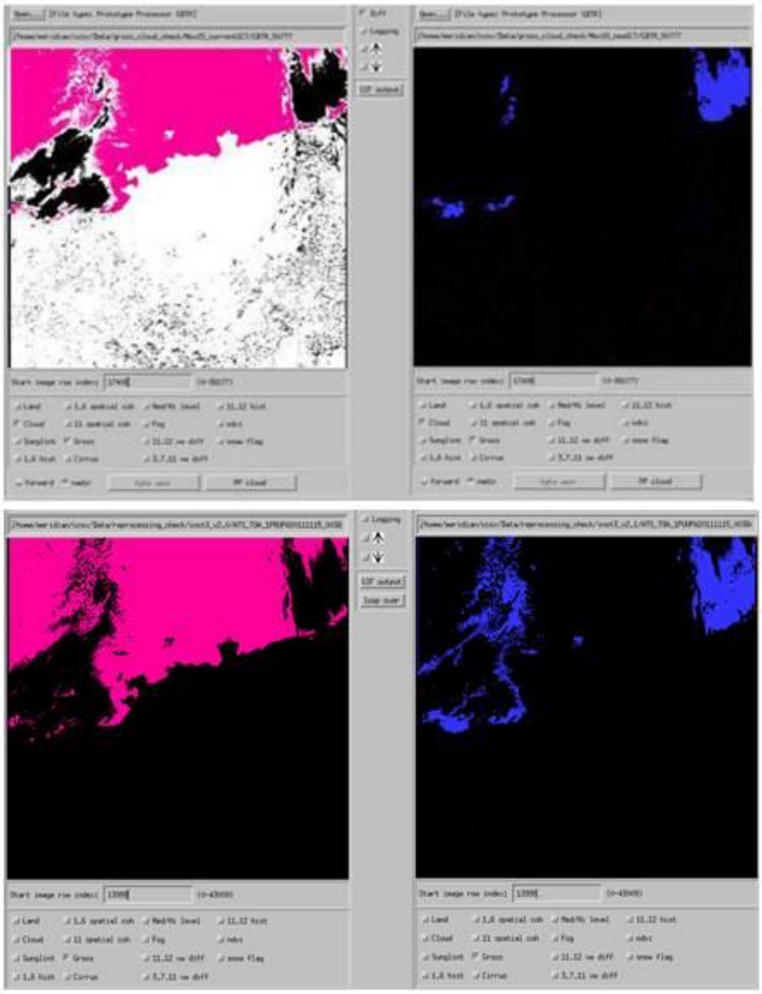

Figure 14 Case 1. Top plot shows the image taken from AD2. LHS is the v2.0PP GCT flag in pink, and RHS is the difference between that and the PP with the new GCT thresholds. The bottom image shows the v2.0 flags in the LHS and the difference between that and the v2.1 reprocessed data on the RHS. The v2.0PP/v2.1PP difference and v2.0/v2.1 difference plot is not exactly the same. There are more blue areas in v2.0/v2.1 compared to v2.0PP/v2.1PP around the cloud edges from geolocation changes in v2.1. However, it is clear that the same area of cloud to the top left of each image has been identified by both.

2. Case2 of AD2, Vladivostok.

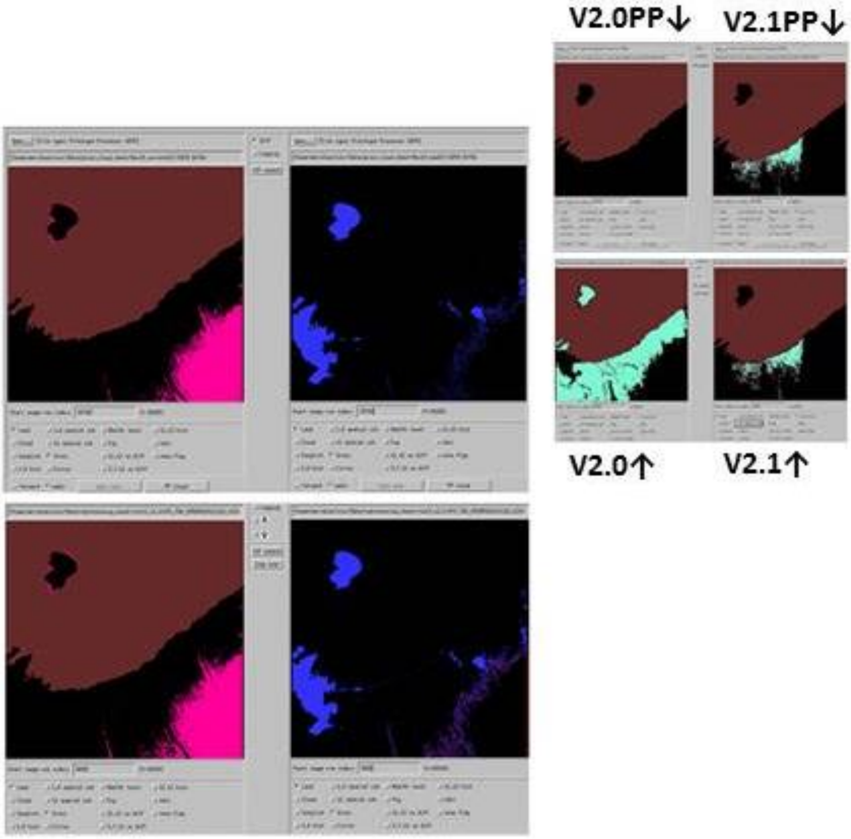

Figure 15. Case 2. The images to the LHS are the v2.0PP/v2.1PP and v2.1/v2.1 as shown in the previous figure. The GCT in v2.1 matches that of v2.1PP. The 11,12 hist test (green) flag is also included and shown on the RHS. The upper plots show the v2.0PP and v2.1PP, and the lower plots show the v2.0 and v2.1. Strangely, here the v2.0 does not match that of the v2.0pp with the 11,12 hist test result indicating that the PP is doing something different wrt the operational prosessor. From pers. comm. with Jack Abolins, it appears that this flipping of cloud flag from the 11,12 histogram test with the PP versus OP has been observed before and this test (along with some other tests TBD) does not seem to be very robust to the initial conditions used in the processor. This issue has never been properly investigated. It is discussed again later in this document.

3. Case 5 in AD2, West Coast of Alaska

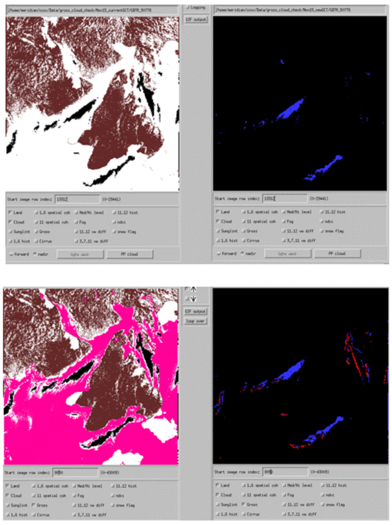

Figure 16. Case 5. Top plot shows the image taken from AD2. LHS is the v2.0 GCT flag in pink, and RHS is the difference between that and the PP with the new GCT thresholds. The bottom image shows the v2.0 flags again in the LHs and the difference between that and the reprocessed data on the RHS. The v2.0PP/v2.1PP difference and v2.0/v2.1 difference plot is almost identical except for some cloud edging effects.

4. Case 6 in AD2, Sea of Okhotsk

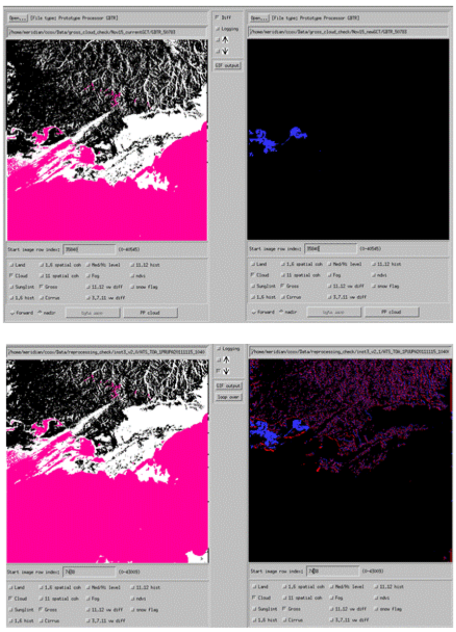

Figure 17. Case 6. Top plot shows the image taken from AD2. LHS is the v2.0 GCT flag in pink, and RHS is the difference between that and the PP with the new GCT thresholds. The bottom image shows the v2.0 flags again in the LHS and the difference between that and the reprocessed data on the RHS. The v2.0PP/v2.1PP difference and v2.0/v2.1 difference plot is identical apart from cloud edging effects.

5. Case 7 of AD2, South Atlantic Ocean

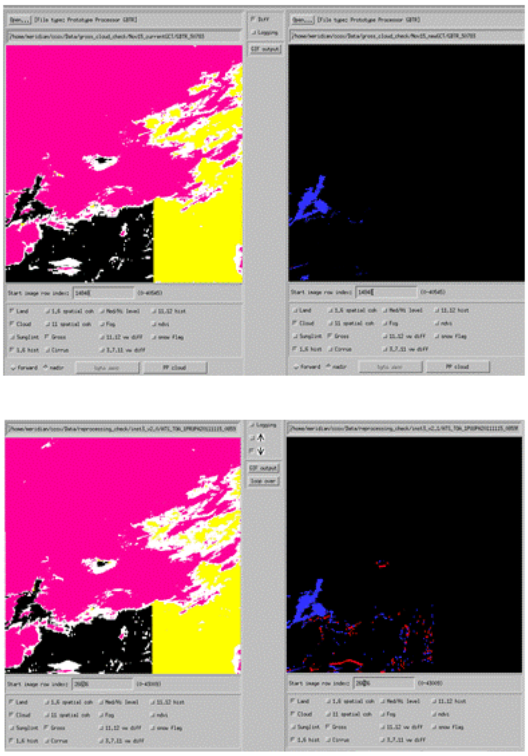

Figure 18. Case 7. Top plot shows the image taken from AD2. LHS is the v2.0 GCT flag in pink, and RHS is the difference between that and the PP with the new GCT thresholds. The bottom image shows the v2.0 flags in the LHs and the difference between that and the reprocessed data on the RHS. The v2.0PP/v2.1PP difference and v2.0/v2.1 difference plot is identical apart from cloud edging effects.

3.1.3 Summary

The GCT flags of the AATSR data have been checked and it is concluded that the new BT thresholds in the test have been correctly applied in v2.1 data and the results are as expected. Although no images are presented here, the ATSR-2 and ATSR-1 data also show slight differences in the GCT flag even with no change to the test threshold because of co/geolocation changes in the v2.1 data.

3.2 Thin Cirrus Test

The changes to the cloud flagging from this test are small, with differences of less than 0.5% for all instruments as it only uses the 11 and 12um BT for which the calibration has not changed. The differences are primarily due to the changes in the geolocation, as can be seen in the example in Figure 19. The red/green pixels are occurring along the cloud edges only. As many other tests tend to flag the same pixels, there are fewer instances where this test works alone and it accounts for less than 0.5% of all pixels. However, for those instances the changes to the number of pixels only flagged by this test has increased significantly in the reprocessed data.

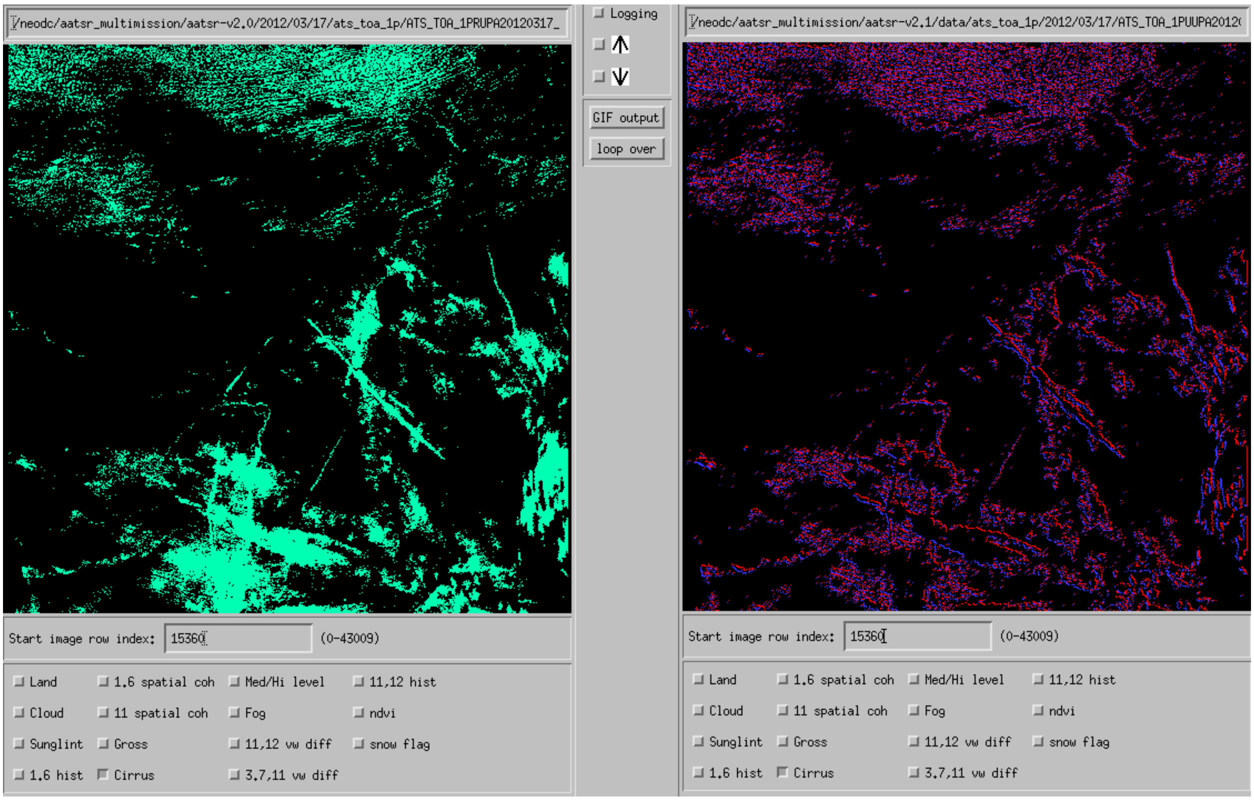

Figure 19. The thin cirrus flag is shown on the LHS for AATSR v2.0, and the difference plot on the RHS. This is the frame showing the most changes for the nadir view for AATSR. We can see the changes are happening at the cloud edges only.

3.3 Medium/High Cloud Test

Again, as for the cirrus test, the major differences found in the cloud flag are at cloud edges due to changes in geolocation as this test only uses 2 IR channels. An example is shown in Figure 20.

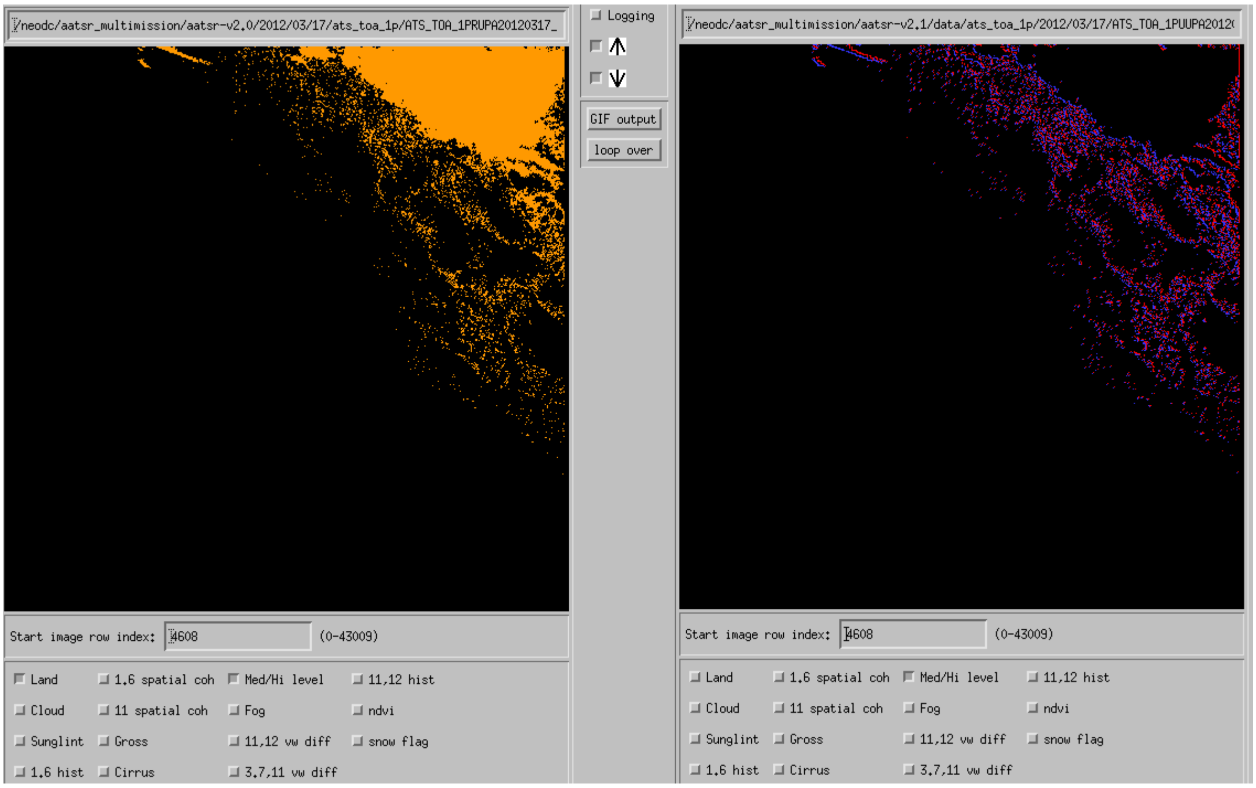

Figure 20. The medium/high cloud flag is shown on the LHS for AATSR v2.0, and the difference plot on the RHS. We can see the changes are happening at the cloud edges only.

3.4 Fog/Low Stratus Cloud Test

Again, as this test only uses the 11 and 3.7 channels and does not depend on any other tests, the major differences are at cloud edges due to changes in geolocation. An example is given in Figure 21.

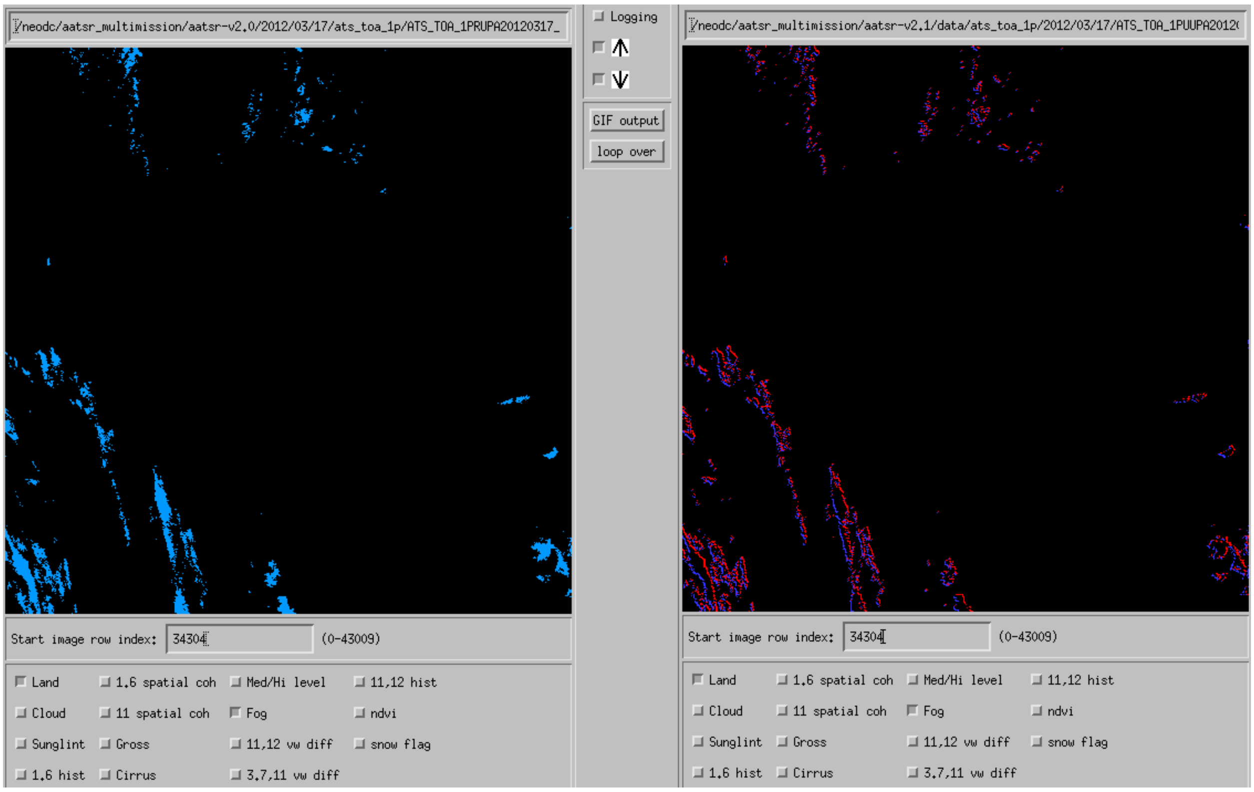

Figure 21. The flog/low stratus flag is shown on the LHS for AATSR v2.0, and the difference plot on the RHS. This is the frame showing the most changes for the nadir view for AATSR. We can see the changes are happening at the cloud edges only.

3.5 11micron Spatial Coherence Test

Changes to the cloud flagging in this test arise due to a couple of issues. Firstly, the test depends upon the result of preceding cloud tests (GCT, Thin cirrus, Med/hi, Fog). The GCT in particular may cause a group of pixels to undergo a change in the flag. If the pixels were previously not flagged by the GCT in v2.0, then are flagged in v2.1, the 11um SCT will not be applied to them. Conversely, if a group of pixels were flagged in v2.1 and are not flagged in v2.1, this group will go on to have the 11um SCT applied to them, and then may result in a cloud flag by this test that was not previously there. Secondly, a slight difference in the 11um and 12um BT due for example from colocation changes may lead to a more significant area of change since the test operates on blocks of pixels. Finally, if there are differences in the land map in the reprocessed data then this will also lead to differences in the SCT cloud flag as it is only applied over sea. In some cases entire frames have changed classification in ATSR-2 data. Still, the overall effect of these does not change the classification of significant numbers of pixels, as less than 0.5% change for all instruments.

Several frames have been identified where there is a significant change in the cloud classification that does not appear to be due to any of the reasons identified above. It is thought that these changes between v2.0 and v2.1 may have arisen due to sensitivity to the initial conditions used in the processor and this has been observed before for this test (pers. comm. Jack Abolins). A few examples where there is a large area of change within a frame have been identified here and are discussed below. The robustness of this test to initial starting conditions may want to be investigated further in the future.

1. Example 1

Image: `ATS_TOA_1PRUPA20120306_121404_000065273112_00224_52393_3789.N1`
Start pixel: `26112`
Net difference in number of uniquely cloudy pixels flagged by this test alone: `40915`

In this example from AATSR shown in Figure 22, there is a significant area of cloud that has been added by the 11um SCT in the reprocessed data. The cause of this is not due to changes in cloud flagging by the preceding tests as it is clear in the original image where the flag has now been set. There do not seem to be any data quality issues in v2.0 either.

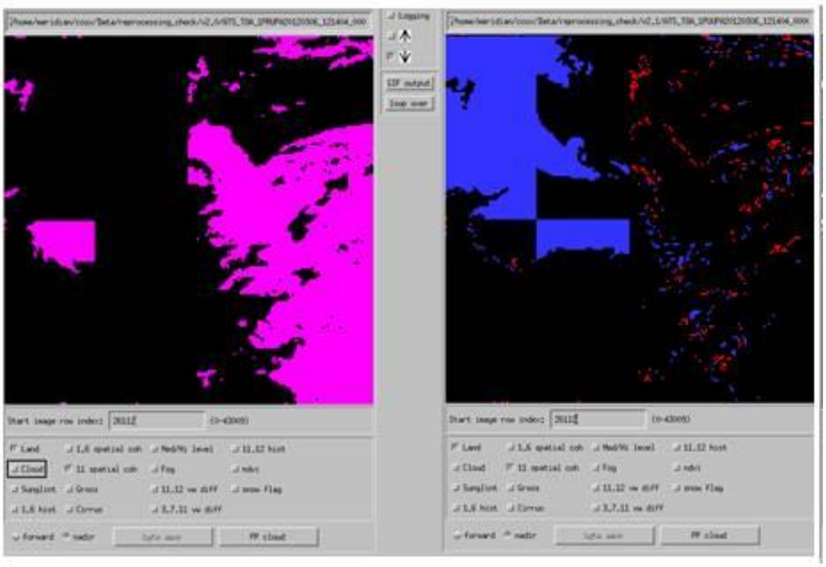

Figure 22. LHS image shows the 11SCT flag in v2.0 data in purple. RHS shows the differences between v2.0 and v2.1, with blue indicating v2.1 has cloud where v2.0 does not. The differences do not arise from flagging by the preceding cloud tests or changes in the land flag. It is the only test to flag the large area of cloud to the top left of the v2.1 image.

2. Example 2

Image: `ATS_TOA_1PRUPA20120317_235223_000065273112_00389_52558_4091.N1`
Start pixel: `17408`
Net difference in number of uniquely cloudy pixels flagged by this test alone: `-23079`

In this example from AATSR shown in Figure 23, there is a significant area of cloud for which the flag has been removed by the 11um SCT in the reprocessed data. The cause of this is not due to changes in cloud flagging by the preceding tests as it is clear in the original image where the flag has now been set. There do not seem to be any data quality issues in v2.0 either.

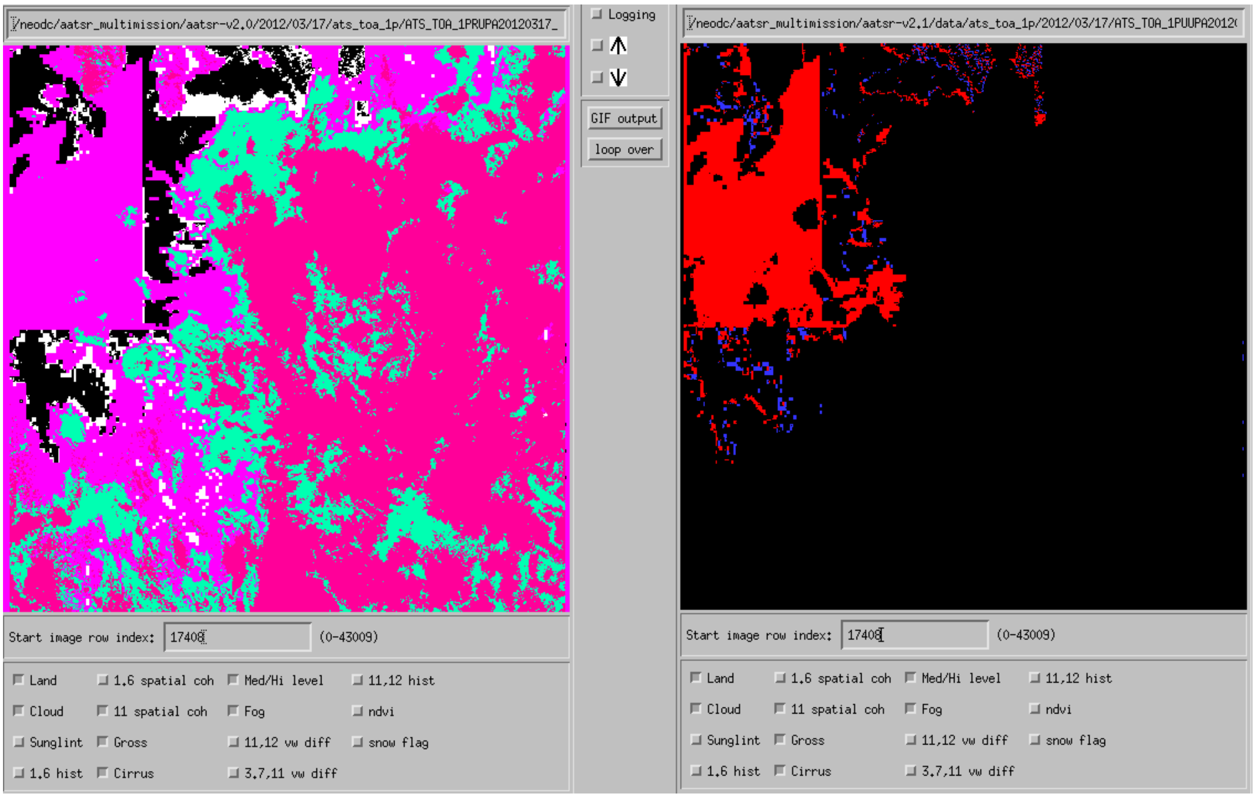

Figure 23. LHS image shows the 11SCT flag in v2.0 data in purple, cirrus in green and GCT in pink. RHS shows the differences between v2.0 and v2.1, with red indicating v2.1 has no cloud where v2.0 does. The differences do not arise from flagging by the preceding cloud tests or changes in the land flag.

3. Example 3

Image: `ATS_TOA_1PRUPA20120314_154114_000065273112_00341_52510_4043.`
Start pixel: `19456`
Net difference in number of uniquely cloudy pixels flagged by this test alone: `-17460`

This is another large area where the AATSR 11SCT flag has changed and is not affected by preceding tests, as shown in Figure 24.

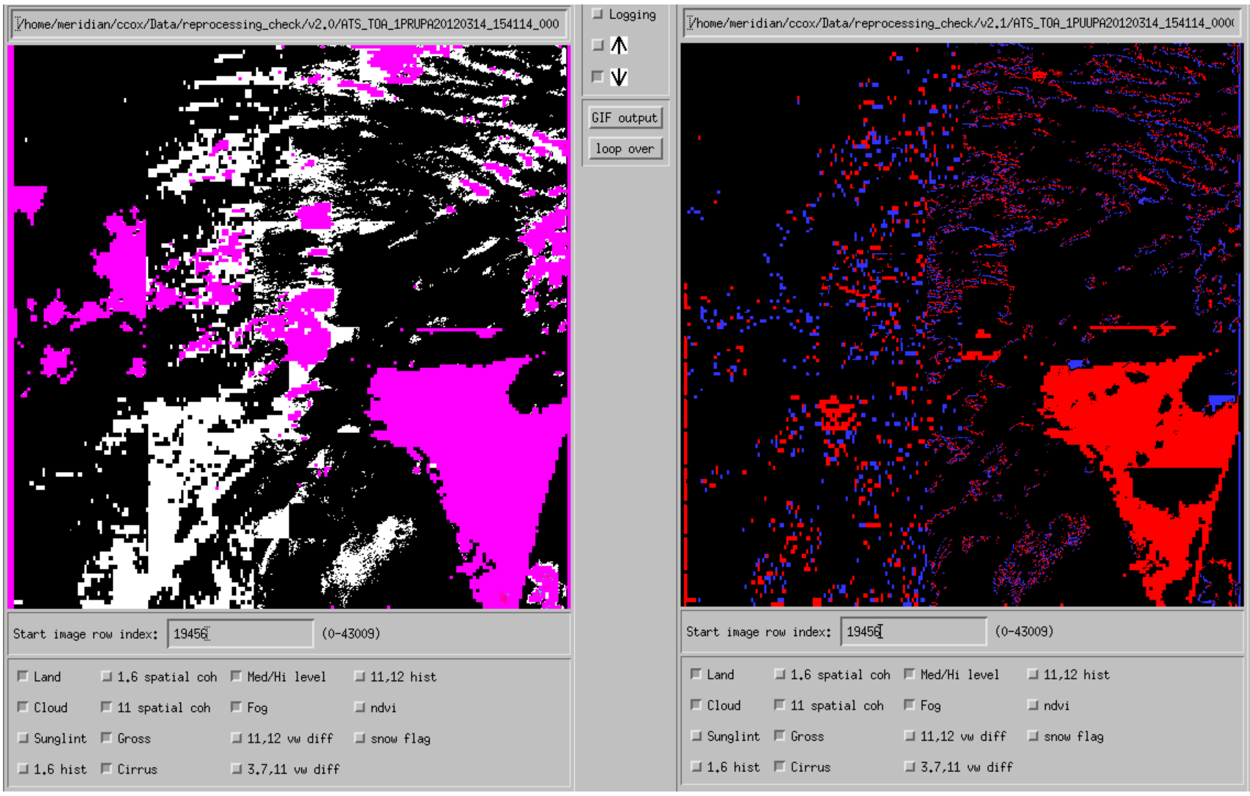

Figure 24. LHS image shows the 11SCT flag in v2.0 data in purple. RHS shows the differences between v2.0 and v2.1, with red indicating v2.1 has no cloud where v2.0 does. The differences do not arise from flagging by the preceding cloud tests or changes in the land flag.

3.6 1.6um Histogram Test

This test is responsible for most of the changes to the cloud flag in ATSR-1 data with -5% (-5%) nadir (forward). For ATSR-2 data it is -0.3% (-0.5%) nadir (forward) and for AATSR it is -5% (<1%) nadir (forward).

The 1.6um histogram test is applied to sea pixels only and so it will also be affected by changes in the land map which explains most of the ATSR-1 cases where large areas of cloud flag have been altered. The test relies on the result of some of the preceding tests which also has an impact on the change to this flag and the calibration of this channel has also changed which can explain the differences in cloud flag. The effect of changes to colocation can also be observed in the images as shown in Figure 25.

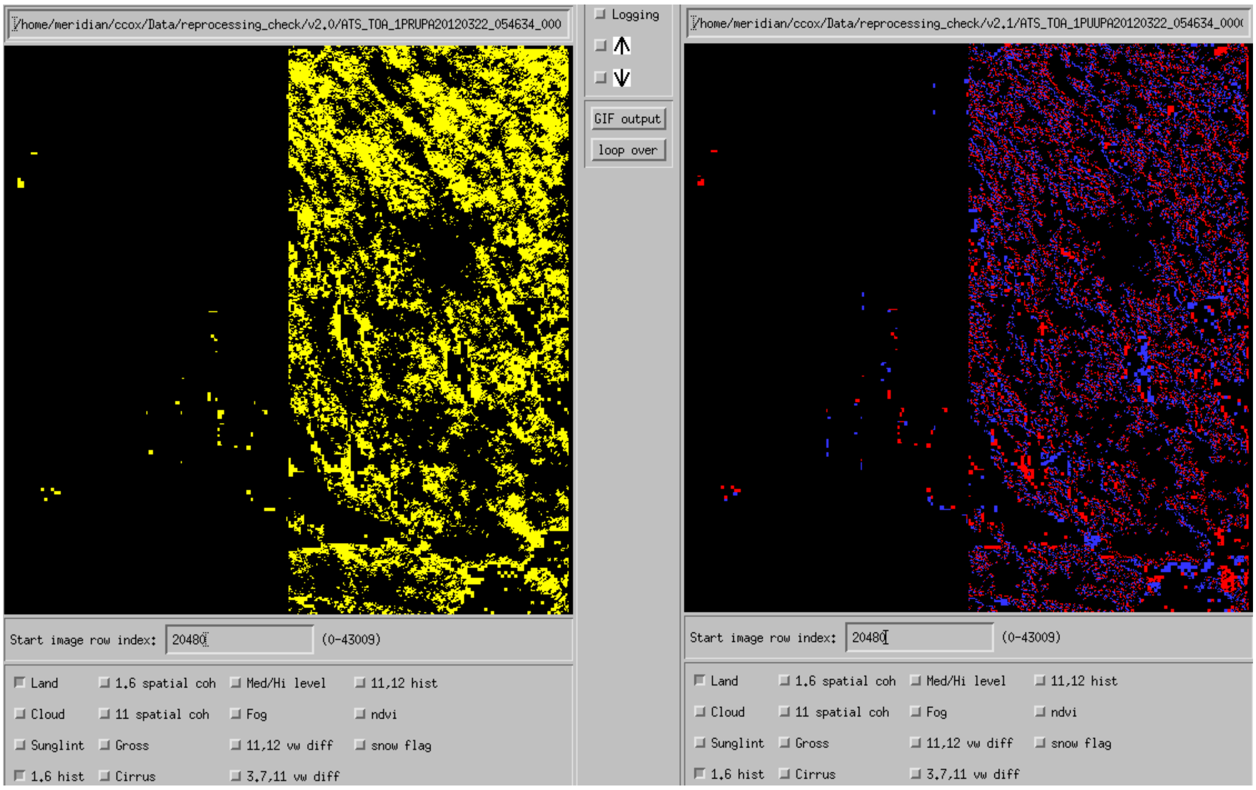

Figure 25. An example of AATSR data where the 1.6 histogram cloud flag has been changed around the cloud edges. LHS is the flag, and the RHS is the difference in v2.0 and v2.1 data.

3.7 1.6 Spatial Coherence Test

This test, together with the previous test, is responsible for most of the changes to the cloud flag in ATSR-1 data with -3% (-7%) nadir (forward). For ATSR-2 data it is -0.2% (1.1%) nadir (forward) and for AATSR it is 3% (14%) nadir (forward).

The 1.6um SCT is applied to sea only and so it will also be affected by changes in the land map which explains most of the ATSR-1 cases where large areas of cloud flag have been altered. The test relies on the result of some of the preceding tests (GCT, Thin cirrus, 11 SCT, 1.6 histogram) which also has an impact on the change to this flag and the calibration of this channel has also changed which can explain the differences in cloud flag. The effect of changes to colocation can also be observed in the images.

3.8 11/12um Difference Test

This test relies on looking at the differences between the forward and nadir views and so will be sensitive to the changes in the colocation. The differences between v2.1 and v2.0 cloud flags from this tests are -4% for ATSR-1, -3% for ATSR-2 and -11% for AATSR. An example of the change in cloud flag from an AATSR image are shown in Figure 26. Where there are many cloud edges, the number of pixels changing classification will be highest.

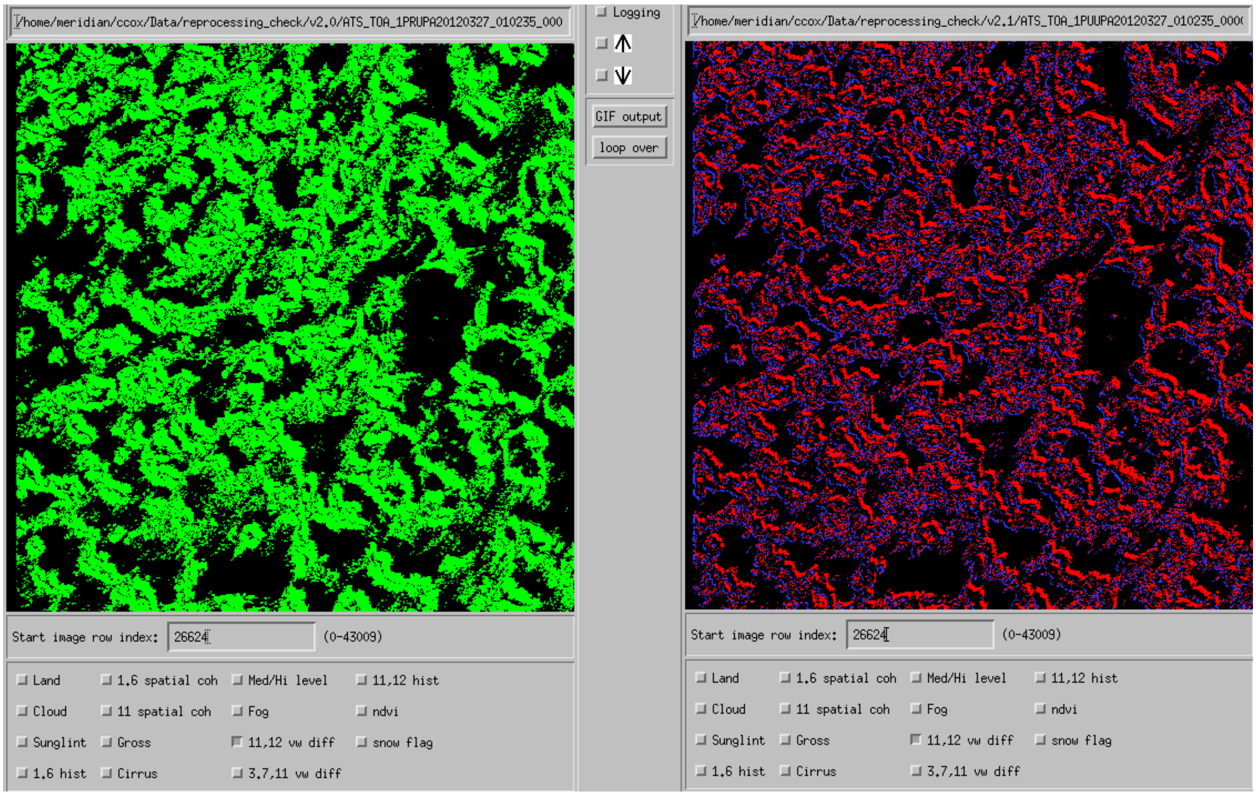

Figure 26 LHS shows the 11/12um diff cloud test flag in green for the v2.0 image and the RHS shows the difference with v2.1.

3.9 3.7/11um Difference Test

As for the 11/12um difference test, this test will be affected mainly by colocation changes. The difference for the AATSR data is -4% and for ATSR-2 it is -2%. An example image is shown in Figure 27.

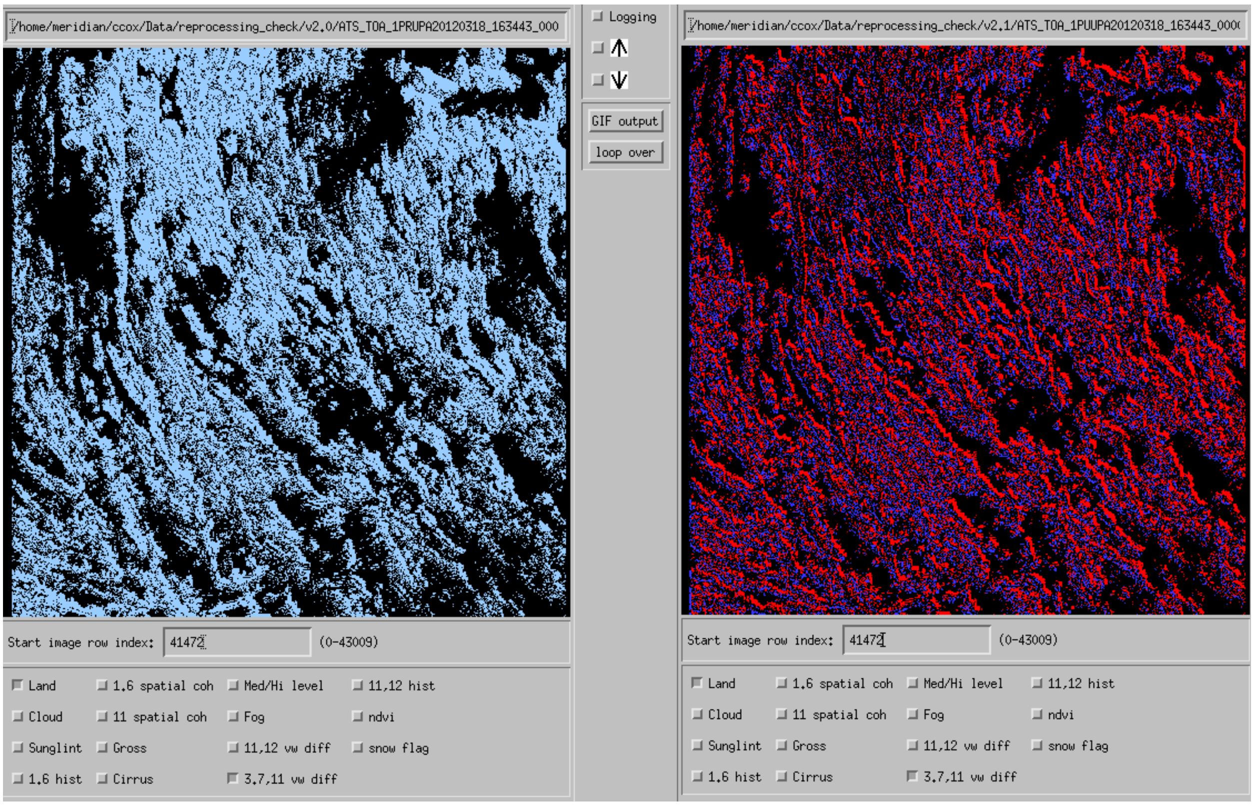

Figure 27. LHS shows the 3.7/11um diff cloud test flag in blue for the v2.0 image and the RHS shows the difference with v2.1.

3.10 Visible/NDVI Cloud Test

The visible cloud test is the one that has shown the most changes in the cloud flag as a result of reprocessing as it uses 3 of the 4 visible channels, all of which have had their calibration improved.

The visible/NDVI cloud test uses NDVI and NDVI-type ratios together with a look-up table of scene types to determine whether there is cloud over a particular land surface. Although in principal it could be used over the sea, currently it is only used over land. The scheme was based on unpublished work done at RAL by AD Stevens [AD3] A look-up table provides scene type 'zones' as functions of indices that are calculated using the 0.87, 0.67 and 0.55 micron channels. This look-up table was empirically derived using AATSR data. The zones are plotted in Figure 28 and definitions of them are given below.

- Zone 0: Ice
- Zone 1: Thin cloud over sea
- Zone 2: Very thin cloud over sea
- Zone 3: Cloud
- Zone 4: Sea
- Zone 5: Dirty water
- Zone 6: Desert
- Zone 7: Desert with sparse vegetation
- Zone 8: Vegetation, possibly dried or incomplete cover
- Zone 9: Thin cloud over vegetation
- Zone 10: Possibly slight cloud cover over vegetation
- Zone 11: Vegetation

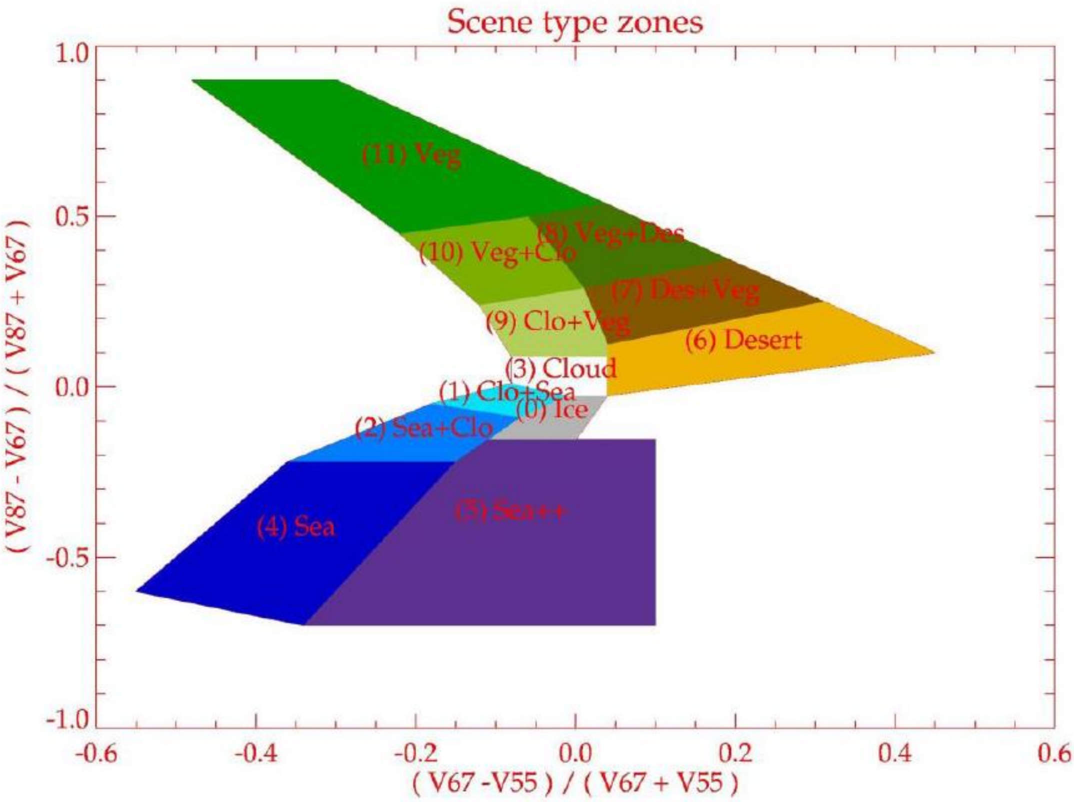

Figure 28 The visible cloud test zones as a function of NDVI and NDVI-type indices.

Any data that is found to lie outside of these zones is classed as 'no zone' and if invalid reflectance data is provided, the test returns an error code to indicate this. There is one pure cloud zone and 4 mixed cloud zones (zones 10, 9, 2, 1) for cloud over vegetation or sea, however, there do not seem to be any mixed zones for desert or ice, implying the scheme does not account very well for these mixed types. There is also no snow surface type, although these scenes may have been included in the ice surface type.

As these zones were derived empirically from pre-2007 AATSR, the change in the visible channel calibration undergone with the recent reprocessing will mean that the look-up table is no longer relevant to the current data. Naturally, many pixels which fall into the centre of the zone will not be affected and will remain in their original zone but as we will see, many pixels now fall into different zones leading to a change in the flag.

In AATSR data the v2.1 visible/NDVI cloud flag has led to significantly less cloud flagging, whereas in ATSR-2 it has led to significantly more cloud flagging. In many cases, entire frames have changed classification and many where we see the highest number of pixels are changing from cloud to desert. To investigate these changes in more detail, the numbers of pixels that have been classified in each zone within the v2.0 and v2.1 data have been calculated for the March images of AATSR data and the results are shown in Figure 29. Note that this test would normally only return a flag to indicate a cloudy of clear pixel rather than specific information on the surface type. This surface type has been found specifically for this study.

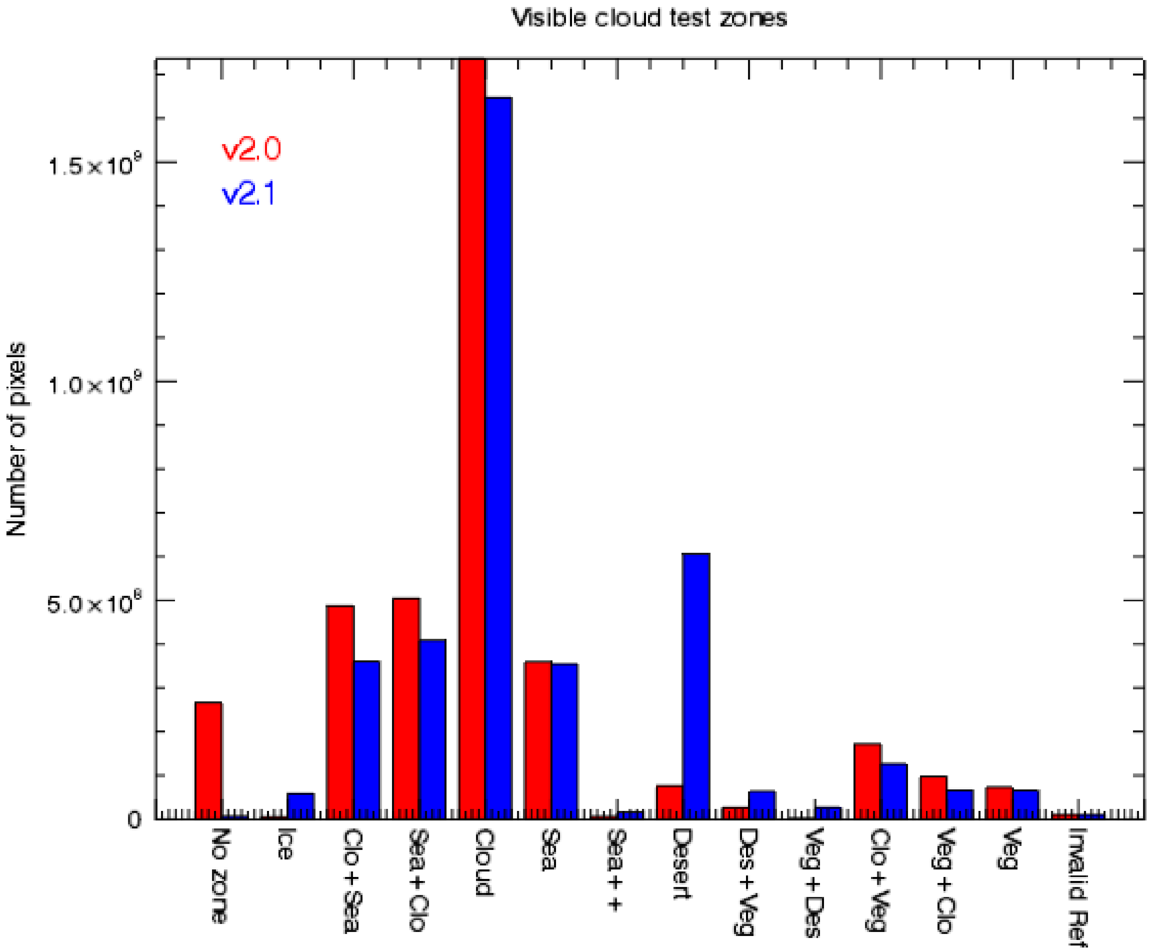

Figure 29. A bar plot showing the number of pixels classified into each visible test zone for v2.0 data (red) and v2.1 data (blue) for the March 2012 AATSR dataset.

Before analysing any differences between the v2.0 and v2.1 zone classifications, it is evident that there are significant numbers of pixels classified in both versions as being over sea, either with or without cloud. This immediately introduces concerns with the robustness of the test as it is known that all the pixels processed here are land pixels only.

The bar plot reveals that there are significantly less pixels in 'no zone' in v2.1 and there are significantly more pixels now in 'Desert', 'ice', 'Desert+veg' and 'Veg+desert'. If we look at which zone the 'no zone' v2.0 pixels have been reclassified as in v2.1 we see that they have spread into the zones on the left hand side of the scene type zone plot, with most of them moving to 'Cloud' as shown in Figure 30. The pixels that are now classed as Desert in v2.1 have moved mainly from zone 3, 'Cloud' in v2.0, as shown in Figure 31. The pixels that are now classed as Ice in v2.1 have moved mainly from zones 1, 2 and 3 in v2.0, as shown in Figure 32.

The bar plot in Figure 33 shows the numbers of pixels that have changed zones in order of the most numerous, with the top 20 displayed. In total 45% of all pixels have moved into a different zone, with the biggest change being from 'cloud' to 'desert'.

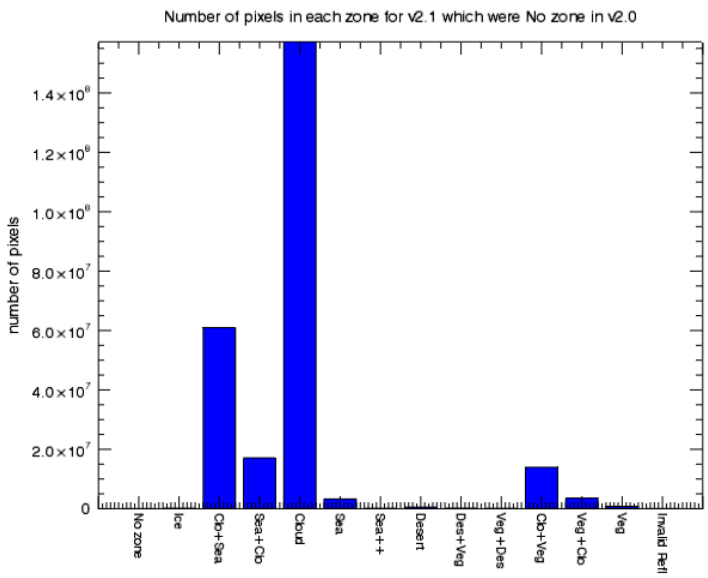

Figure 30. See title. Data is March 2012 AATSR.

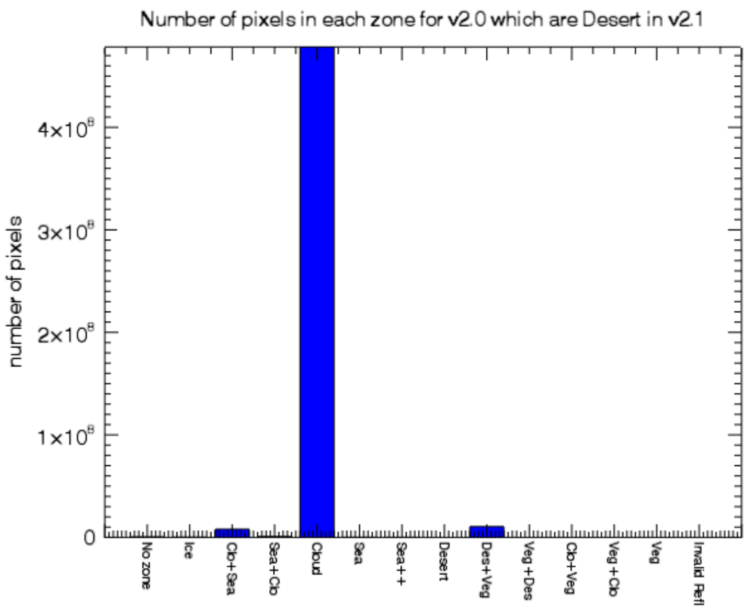

Figure 31. See title. Data is March 2012 AATSR.

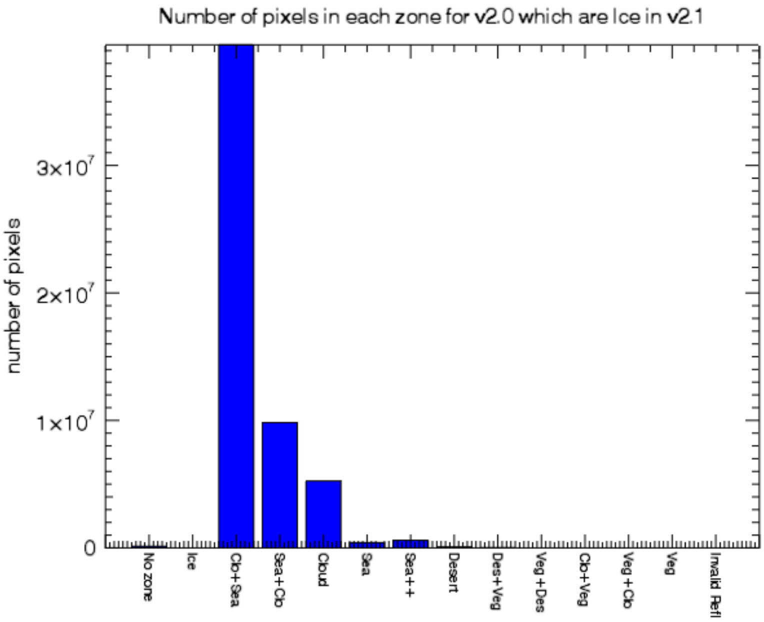

Figure 32. See title. Data is March 2012 AATSR.

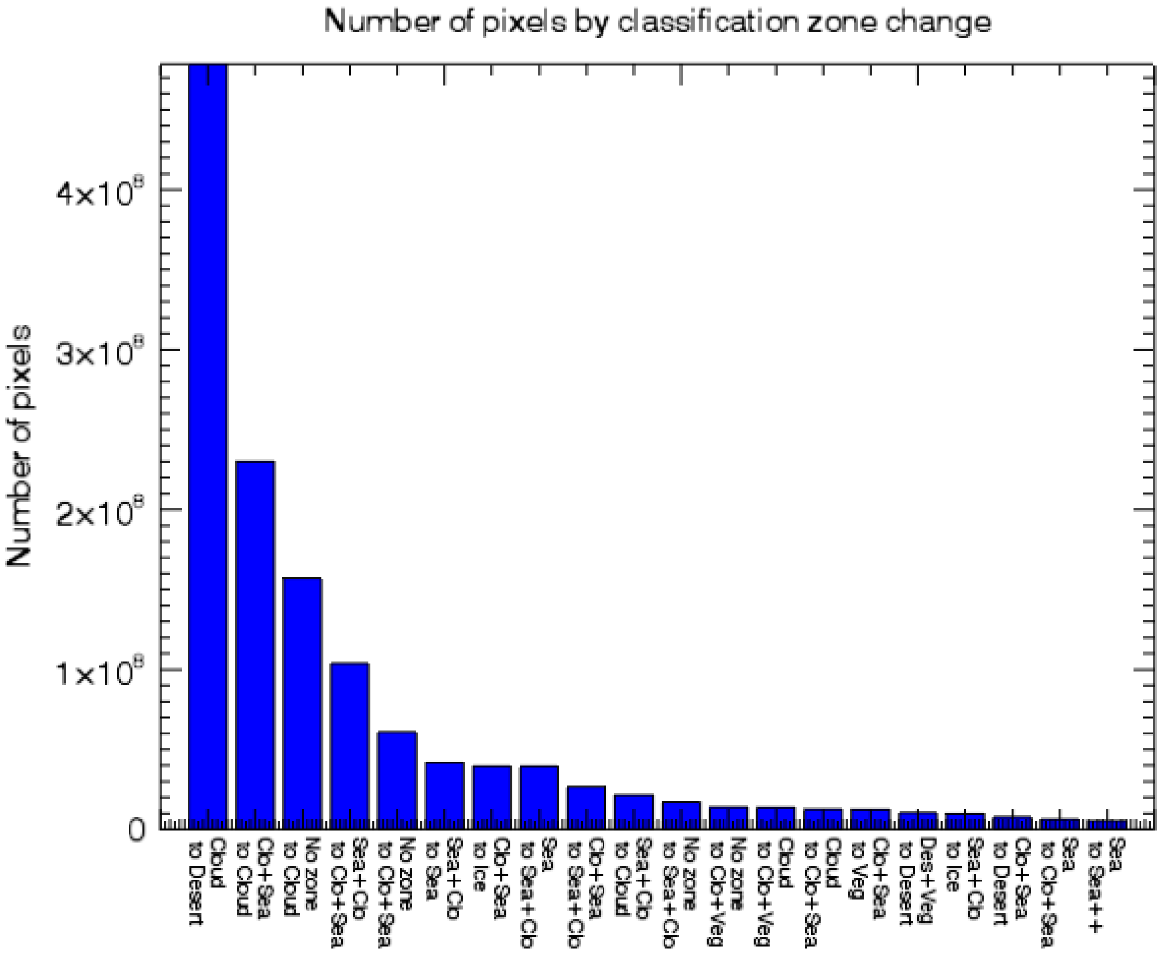

Figure 33. A bar plot showing the number of pixels that have moved between different zones. The 20 zone changes with the highest number of pixels are shown here.

Clearly, the changes in AATSR visible channel calibration are generally shifting the pixels into zones to the right with reference to the scene type zone plot in Figure 28. If we make the assumption that the visible cloud test correctly determined the surface scene type and cloud cover for v2.0 then with the reprocessing the test is no longer appropriate as the change to calibration has been sufficient to shift the zone types of 45% of all pixels, and changes the cloud flag of over 35% of pixels.

The same analysis has not been done for ATSR-2, but it is expected that ATSR-2 will show the same sort of shifts in zone classification, although they will be in a different direction as overall ATSR-2 showed more visible cloud flags in v2.1 data. In addition to changes to the visible calibration, ATSR-2 is affected by the new DEM.

This analysis has highlighted the limitations of the visible/NDVI cloud test and it is now not strictly applicable to the reprocessed (A)ATSR data. The look-up table could be recalculated for the next reprocessing using improved methods (e.g. simulating different scene types) and perhaps introducing new zones but this will take a significant amount of effort.

3.11 Snow Flag

Although not a cloud test, there have been significant changed to the snow flag in AATSR and ATSR-2 data following the reprocessing (ATSR-1 data does not use it). The snow flag was introduced as an experimental field that could be used to map snow. The algorithm it uses is based on published work from Hall et al and it uses the reflectance of the 0.87um channel and a snow index, NDSI which uses the 0.55 and 1.6um channels. The 11um BT is also used for a threshold test. Changes in this flag are expected due to the use of the visible channels. In theory, the improved visible channel calibration should improve the accuracy of this test, but as it is only an experimental field, no further analysis has been performed.

4 Summary

This study has investigated the changes to cloud flagging of the reprocessed (A)ATSR v2.1 dataset. The following issues have been highlighted for future work/discussion at the QWG.

- The new GCT thresholds look to be performing as expected and the improvements and effect on the v2.1 AATSR cloud flags highlight the need to re-calculate the ATSR-1 and 2 thresholds for the next reprocessing
- This study has shown up that on occasion, the 11um SCT show unexpected large shifts in the cloud flag which are likely to be due to sensitivity to initial starting conditions in the processor. The non-robustness of this test, while not a major issue, could still be investigated.
- The visible test flag has undergone large changes with the reprocessing and this study has highlighted that it may be worthy of further investigations.