Hyperspectral data quality overview - 2011

This report describes issues that should be considered when further processing any of the 2011 Airborne Research Survey Facility (ARSF) datasets. The document will be updated over the course of the year, with the latest version available at:

http://arsf-dan.nerc.ac.uk/trac/wiki/Reports

Geo-referencing accuracy

ARSF currently deliver data at level 1 (calibrated sensor data) rather than level 3 (geo-referenced). This allows users to generate level 2 products (e.g. atmospherically corrected radiances) if they wish, and/or map to any projection or datum that suits.

The quality of the geocorrection for each project is described in the documentation supplied with the project and is normally of the order of a couple of meters (approximately 1 pixel). Where a vector overlay or other ground truth information is available, ARSF provide an indication of the average error. If you need higher accuracy, please contact us arsf-processing@pml.ac.uk. It may be possible to tune specific flight lines for higher accuracy or we can provide instructions on how to make your own alignments.

Timing errors

Due to an error in the handling of synchronisation between the navigation system and the Eagle and Hawk sensors, small timing errors (order of 0.05s) may occur. The consequence of timing errors is to cause scan lines to be positioned incorrectly and manifest visually as "wobbles" in the imagery. The wobbles are correlated to, but out of sync with, movements of the aircraft. An example is shown in Figure 1 below.

This issue has been extensively investigated and demonstrated to be a fault in the Specim systems. Specim are working with ARSF to provide upgrades and improvements to correct this issue, but have not yet succeeded.

Therefore we endeavour to correct all timing errors prior to delivery. As this is a manual process and relies on finding suitable visible features in the imagery, some errors may still remain. If any are found, please contact us at arsf-processing@pml.ac.uk.

Figure 1a: timing error in an Eagle line

Figure 1b: corrected version of above (0.13 seconds difference)

Sensor calibration (2011)

Prior to the start of the 2011 flying season, the Eagle and Hawk instruments were calibrated at Specim and also twice at NERC's facility in collaboration with the Field Spectroscopy Facility. The Specim calibration has been verified to be correct against the NERC measurements and is being used for 2011 data. In subsequent years, ARSF will use the NERC calibrations, which are more completely documented.

Wavelength calibration accuracy

The Specim wavelength calibration for Eagle was found to agree with the most recent NERC calibration to an average of approximately 0.15nm (~10% of a bandwidth) at the wavelengths of the calibration lamps used.

The Specim wavelength calibration for Hawk was found to agree with the NERC calibration to an average of approximagely 0.07nm (~1% of a bandwidth) at the wavelength of the calibration lamps used.

Radiometric calibration

The Specim radiometric calibration was found to agree with the NERC calibration to within ~3% for Eagle and ~1% for Hawk. This compares with uncertainty in the lamp used for the NERC calibration of between 2-10% for Eagle and 1-12% for Hawk (depending on wavelength – the edges of the spectral range have a higher uncertainty). We believe this lamp accuracy to be comparable to that achieved by other spectral calibration facilities such as NIST. The error characteristics of the lamp used for calibration by Specim are not known to NERC.

As with the 2009 and 2010 data, the signal to noise ratio appears to degrade at the low and high wavelength limits of both Eagle and Hawk. For example, a comparison of pixels near the high Eagle wavelengths and low Hawk wavelengths, over dark targets such as water reveals a mismatch (see Figure 2). Caution is advised when examining spectral responses at the edges of the usable range. Specifically, Eagle data below 450nm or above 900nm and Hawk data below 1100nm and above 2400nm should be treated with caution.

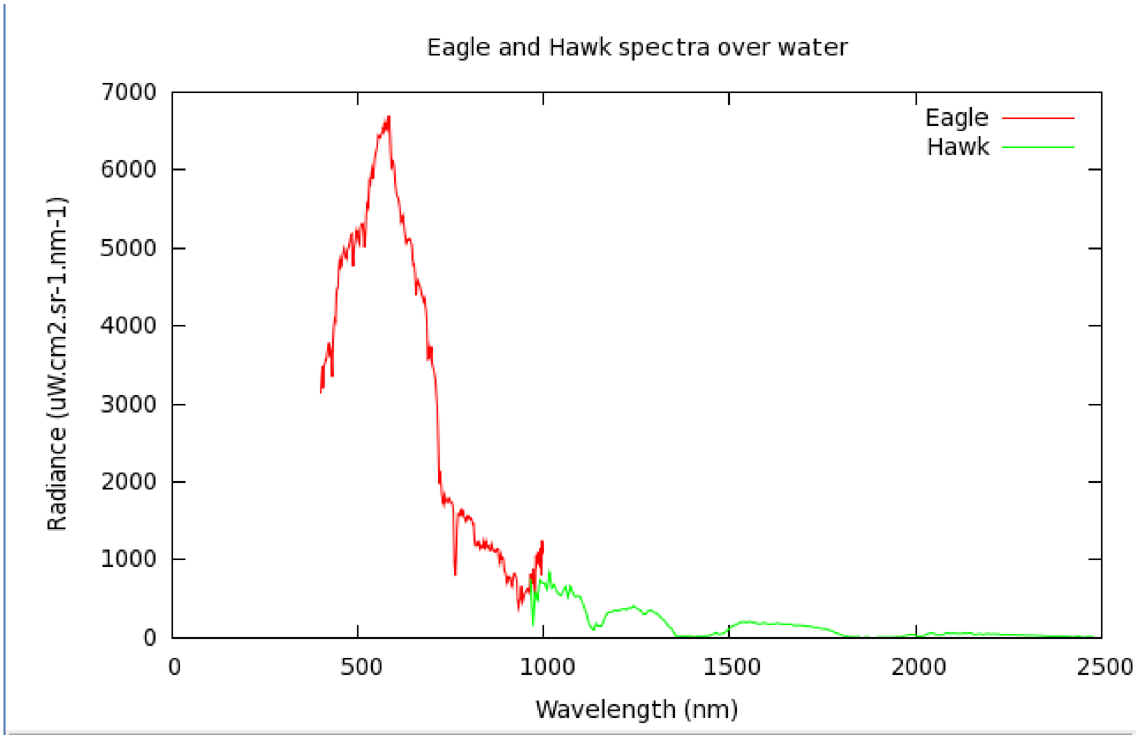

Figure 2a: Eagle vs. Hawk spectra over water (dark target), showing mismatch at the overlap point (2011 data)

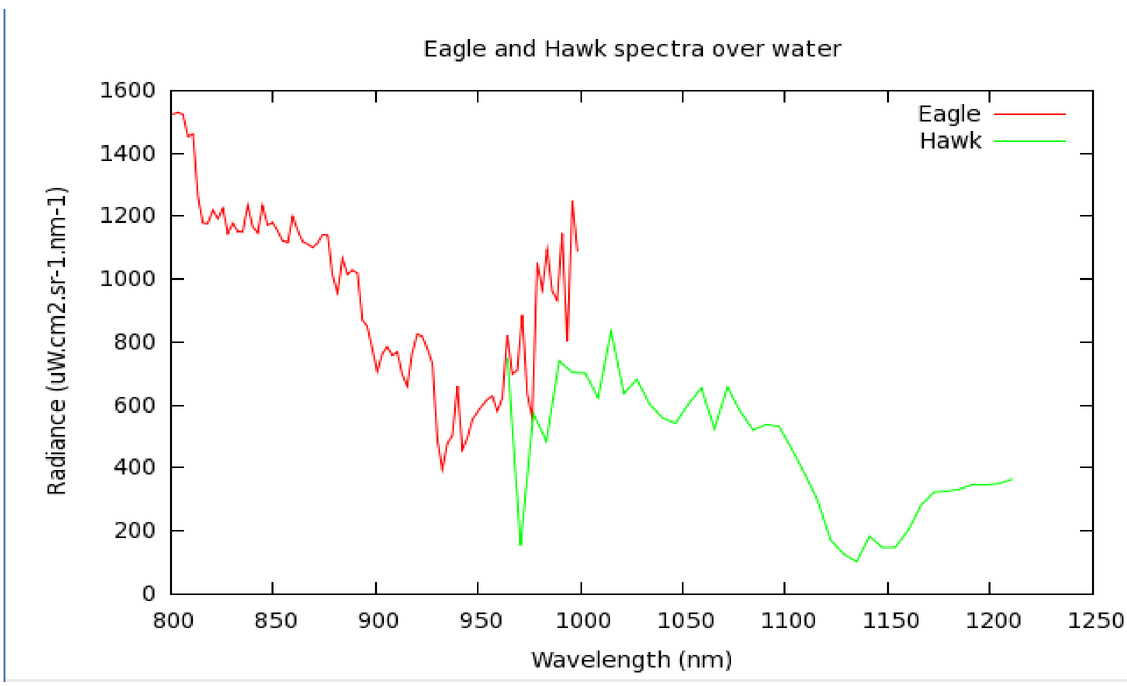

Figure 2b: zoomed up version of Figure 2a

Overflowed pixels

The instruments have a limited dynamic range and must be set to capture data over the appropriate range of signal strength. For example, if the area of interest is dark, then the instrument will be configured to capture as much low light detail as possible. This configuration is set based on operator experience, the principal investigator's indication of the areas of importance and the prevailing conditions. Inevitably, some pixels are unexpectedly bright – e.g. sunglint over water or part of a cloud. These pixels may exceed the maximum capture level and overflow. Typically they are not in areas of interest, but should be accounted for. The accompanying mask file will contain an overflow flag value in the level 1 equivalent pixel.

In Hawk, overflows are marked for just the pixel/band in question. However, Eagle uses a frame transfer CCD, where data are read out in rows. Incoming light continues to accumulate in unread rows during the transfer and is removed by “smear correction” software, which relies on data from one row to correct the next. If a pixel overflows, information is lost and all subsequent pixels in that column cannot be fully corrected. In Eagle, the net effect is that an overflow at 600nm will cause all bluer bands (600nm -> ~400nm) to be under-corrected for that spatial pixel. In this case, the mask file will contain a “smear affected” flag value for the equivalent pixel position. When Eagle data with overflows are delivered, we mask all bands (in the mask file) following an overflow as they will incorporate some unknown additional light. If you would prefer your actual level 1 files to be masked out rather than use the separate mask file please contact arsf-processing@pml.ac.uk.

Smear correction

The Eagle uses a CCD that shifts data out line by line at the end of a frame. While this readout process is quick, additional light still falls onto the detector during the readout period. Currently this is corrected for by subtracting a small amount of light measured in the previous line(s) as they are read-out. This procedure assumes the light input is unchanged during the integration and read-out, but this is a good approximation. However, the sensor is often run with a bandset that doesn't record all of the lines.

The Eagle CCD is 1024x1024, with nominal sensitivity from ~200nm to 1200nm, with readout progressing from red (1200nm) to blue (200nm). In operation, only the middle ~500 bands are recorded (~450-950nm), partly due to low sensitivity in the other regions, but also because there are significant internal reflections/second order effects (which is normal). Figure 3a shows a view of the amount of light falling on different parts of the detector. The red box shows the approximate area that is recorded in normal conditions. Internal reflections can be clearly seen, although they have been highly enhanced to make them visible. The amount of light in the central region greatly overwhelms that of the reflection, although their contribution to error can still be significant in weakly illuminated bands.

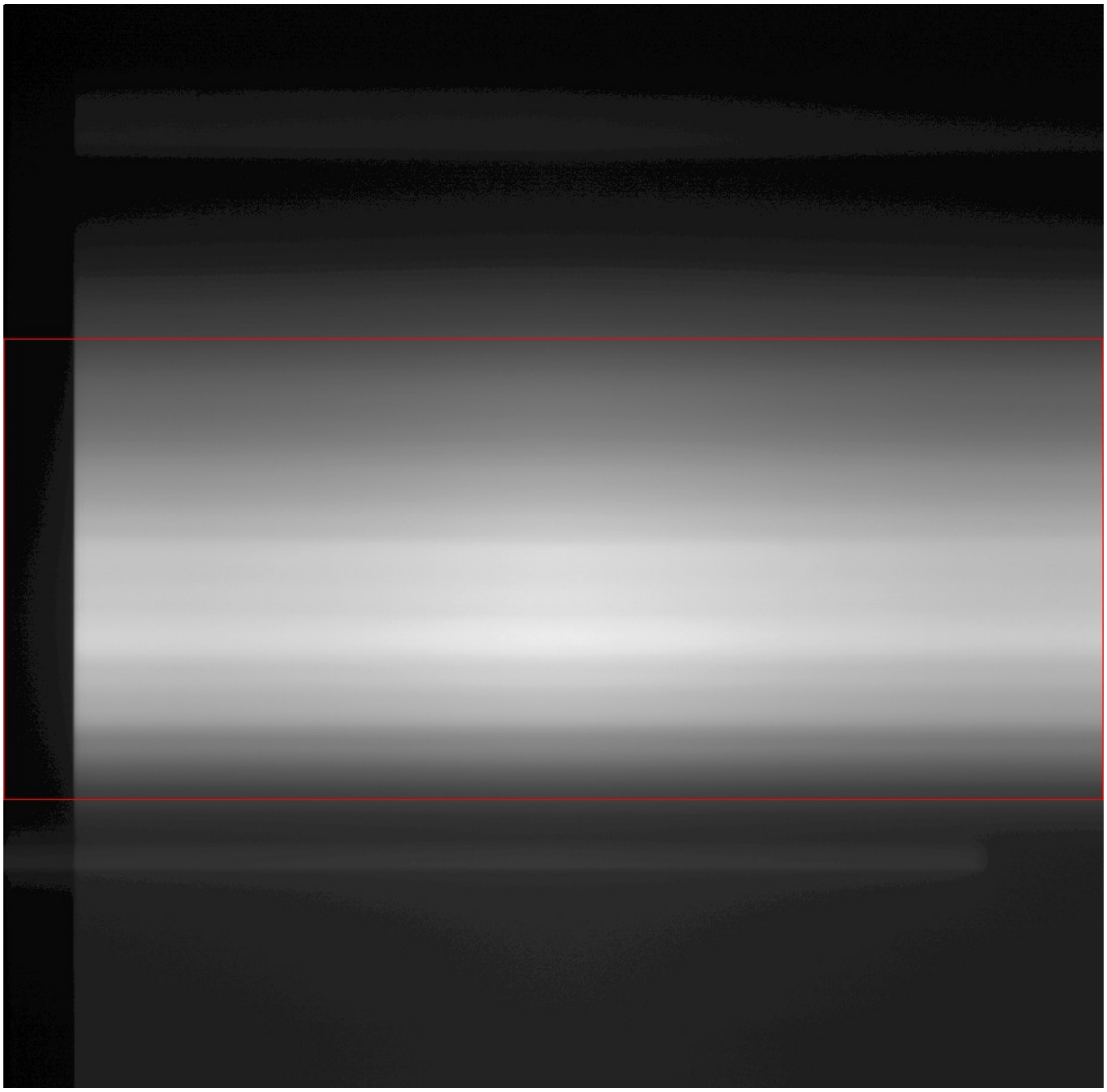

Figure 3a: highly enhanced view of the Eagle CCD, showing recorded bands (inside the red box) and light outside this region.

Consequently, any light falling in the 950-1200nm area cannot be corrected for and smear resulting from that light will remain in the final image. This erroneous light will have the greatest effect where the "true" signal is lowest (e.g. absorption bands). We have run some simulations and the error is naturally worse in bands where there is little light (the red and blue ends of the spectral range, where signal versus noise is lowest, which possibly contributes to why these are poorer quality) and worse the shorter the integration time (below 10ms integration times, errors rapidly increase).

Figure 3b shows estimates of the error introduced by smear into a real-world dataset. The values for unrecorded smear were taken from a calibration experiment and applied to real data from Ethiopia (day 299a/2008) at a variety of integration times. This will likely cause overestimates of the error, as the calibration lamp is brighter than real-world collection conditions, but it is indicative of the relative magnitudes of the error to be expected. As can be seen, the error is dependent on the signal at a particular band, with higher error (2-3%) at the edges of the spectral range, but also with spikes of 1% error in the absorption bands.

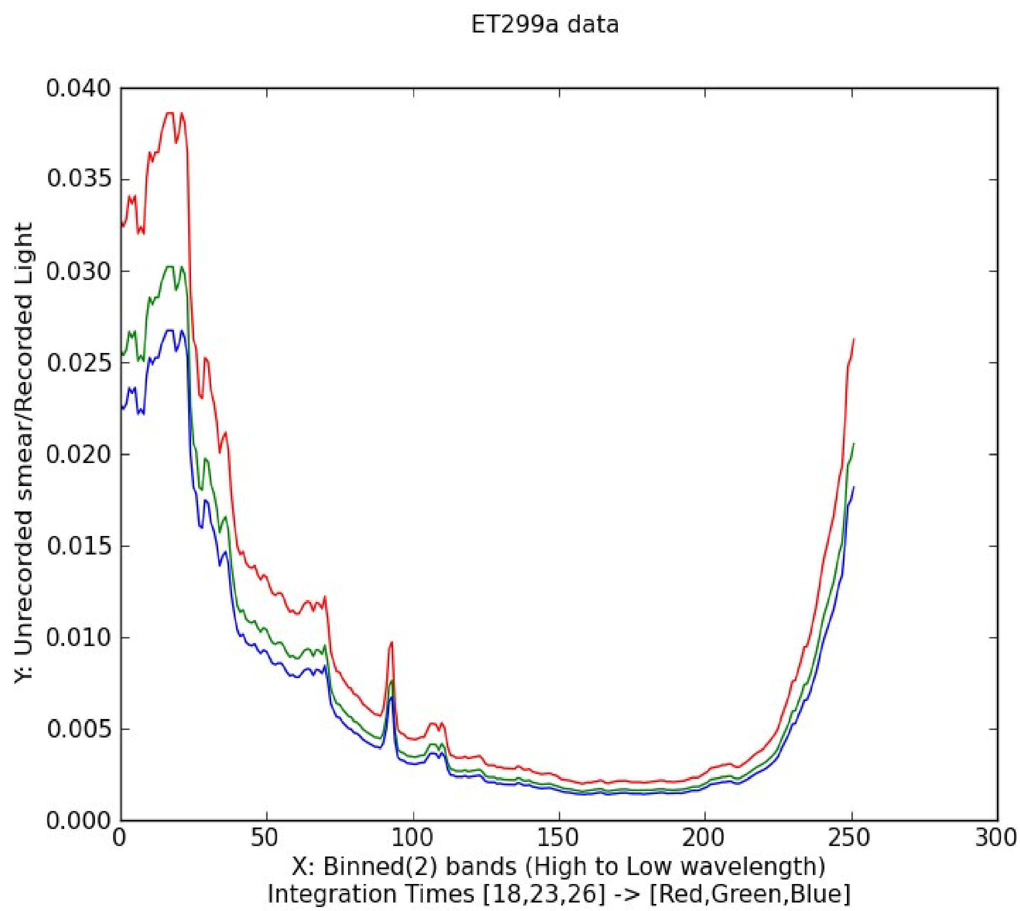

Figure 3b: estimated real-world error due to uncorrected smear, using calibration lamp data to provide estimates of the smear and real Ethiopia data (day 299a/2008). Note that the x-axis is band-number, not wavelength (band 0 = ~1000nm, band 250 = ~400nm).

The estimated (probably overestimated) error ranges from <0.1% (long integration times, high signal bands) to ~70% error (worst possible case of short integration time, low signal bands). The following table shows the best and worst-case error estimates for real-world data at a variety of integration times. Figure 3c shows the distribution of integration times over several years of data collection, to give an indication of the likely impact.

Integration time (ms)	Best case error (% @ peak signal strength)	Worst-case error (% @ weakest signal)
1	5%	70%
4	~2%	20%
6	2%	10%
8	0.6%	9%

12	0.5%	6%
16	0.3%	4%
18	0.2%	4%
23	0.2%	3%
26	0.1%	2.5%
28	0.1%	2.5%
31	0.1%	2%

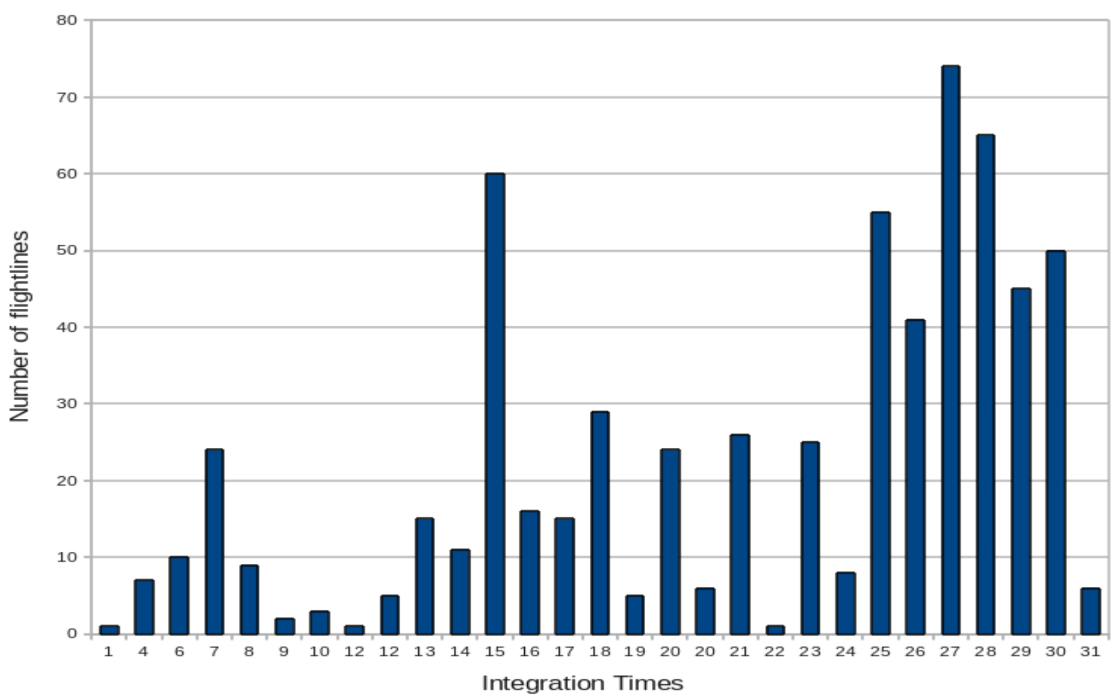

Figure 3c: distribution of integration times for flights in 2008-2010

Bad pixels

The Hawk instrument has a number of bad pixels that give inaccurate values. There are many different types of error (e.g. constant pixel values, uncorrected offset, duplicating neighbouring pixels, etc), and ~1% (about 600) of pixels are to be expected to be bad on the type of CCD used in the Hawk instrument. A list of known bad pixels is included in the delivery as an ASCII text file. The bad pixels will appear in level 1 datasets as straight lines along the direction of flight and as undulating lines in level 3 following the motion of the aircraft (e.g. Figure 4). Typically, they will only affect a single band and are difficult to detect. A complete solution for detecting and removing these started in 2010 and is on-going.

Issue tracked at: http://arsf-dan.nerc.ac.uk/trac/ticket/111

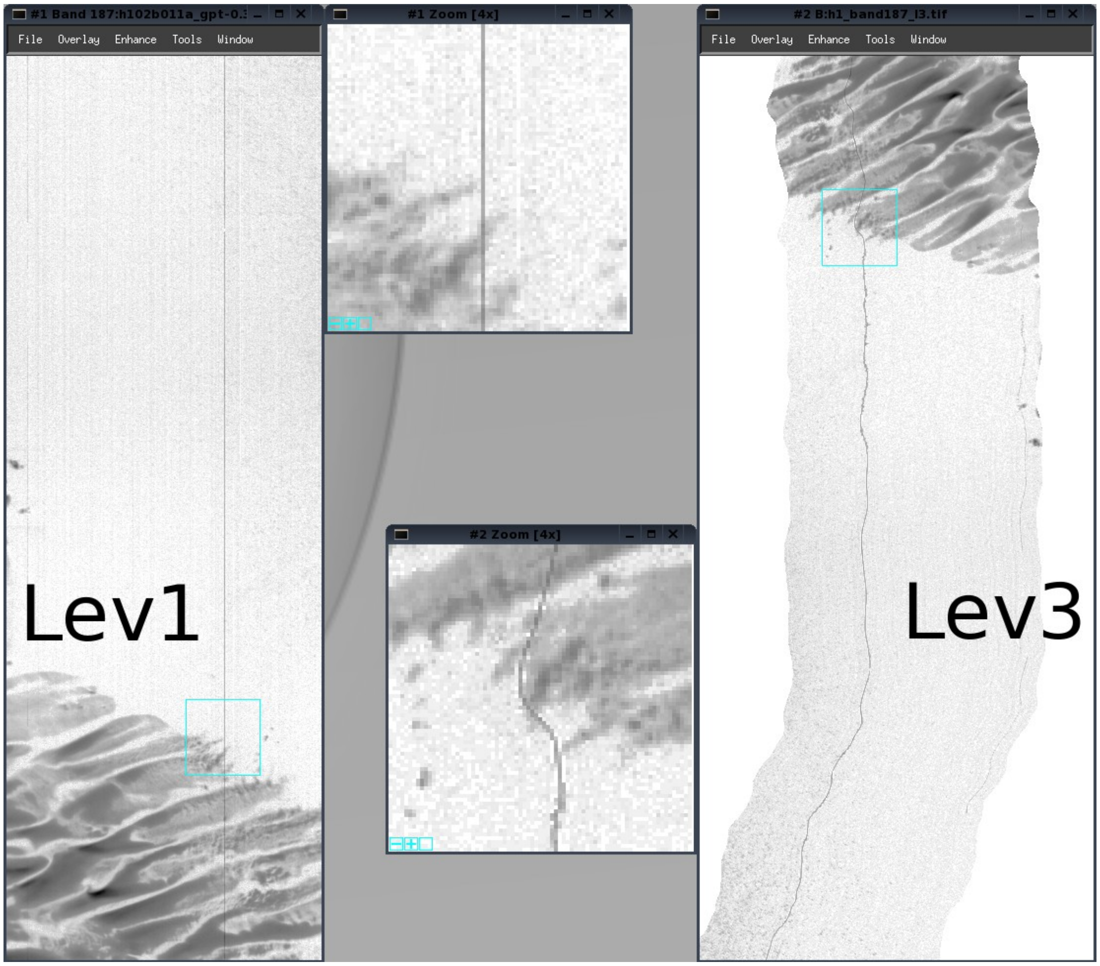

Figure 4: A bad pixel on Hawk band 187, in a scene over water (images inverted to improve contrast on paper)