Pilot Data Issues

Version 1.0 – 27/03/2020 – T. Westmeier

1. Introduction

The WALLABY pilot survey field in the direction of the Hydra cluster consists of two adjacent ASKAP tiles each of which is made up of two slightly shifted footprints for the purpose of minimising the variation in sensitivity across the field. The four footprints in total were calibrated, imaged, mosaicked and provided to the WALLABY team during the first week of March 2020 in the form of a FITS data cube with a size of approximately 1.1 TB, covering approximately 12 × 6 degrees on the sky and the frequency range of 1295.5 to 1439.5 MHz.

We subsequently ran the SoFiA 2 source finding pipeline on the entire data cube in the frequency range of about 1309 to 1418 MHz, corresponding to a redshift range of $z < 0.085$. While overall the results of the source finding run are very promising and indicative of science-quality data across most of the field, we uncovered several minor issues with the data cube that are described and illustrated in this memo.

2. Continuum subtraction

SoFiA still detects a significant number of radio continuum sources in the data cube, indicating that residual continuum emission is still present in the data as a result of insufficient continuum subtraction, in particular in the direction of bright sources.

A comparison with NVSS images at 1.4 GHz reveals a strong frequency-dependence of the level of residual continuum emission. At 1.4 GHz, NVSS sources are picked up by SoFiA above a flux density limit of about 300 mJy. At 1.3 GHz, however, this limit drops to below 100 mJy, resulting in far more more continuum sources being detected at the low-frequency end of the band than at the high-frequency end.

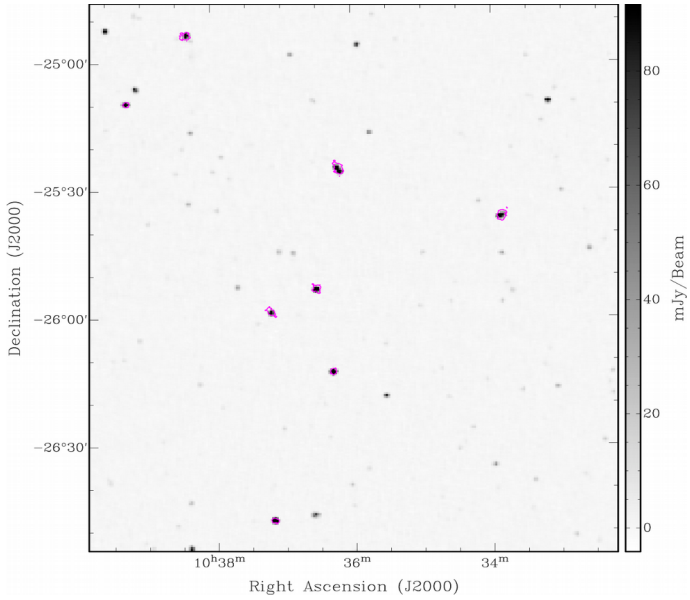

Figure 1: NVSS 1.4 GHz continuum image of a small region of the Hydra field with contours of SoFiA detections overlaid in purple. Apparently, all detections made by SoFiA in this region are due to residual continuum emission.

The issue is illustrated in Fig. 1 which shows an NVSS 1.4 GHz radio continuum image of a small region of the Hydra field overlaid with contours from the moment-0 map created by running SoFiA on just over 1000 channels at the low-frequency end of the band. While SoFiA detects quite a few sources in that region, all of the detections appear to be due to continuum emission in the data cube rather than H I line emission.

An example spectrum towards one of the residual continuum sources in the Hydra field, PKS J1028–2644, which has a flux density of about 0.9 Jy at 1.4 GHz, is shown in Fig. 2. There appears to be a strong periodic ripple in the direction of that source, the amplitude of which appears to be increasing towards the low-frequency end of the band. This would explain why the low-frequency end is more affected by residual continuum emission than the high-frequency end.

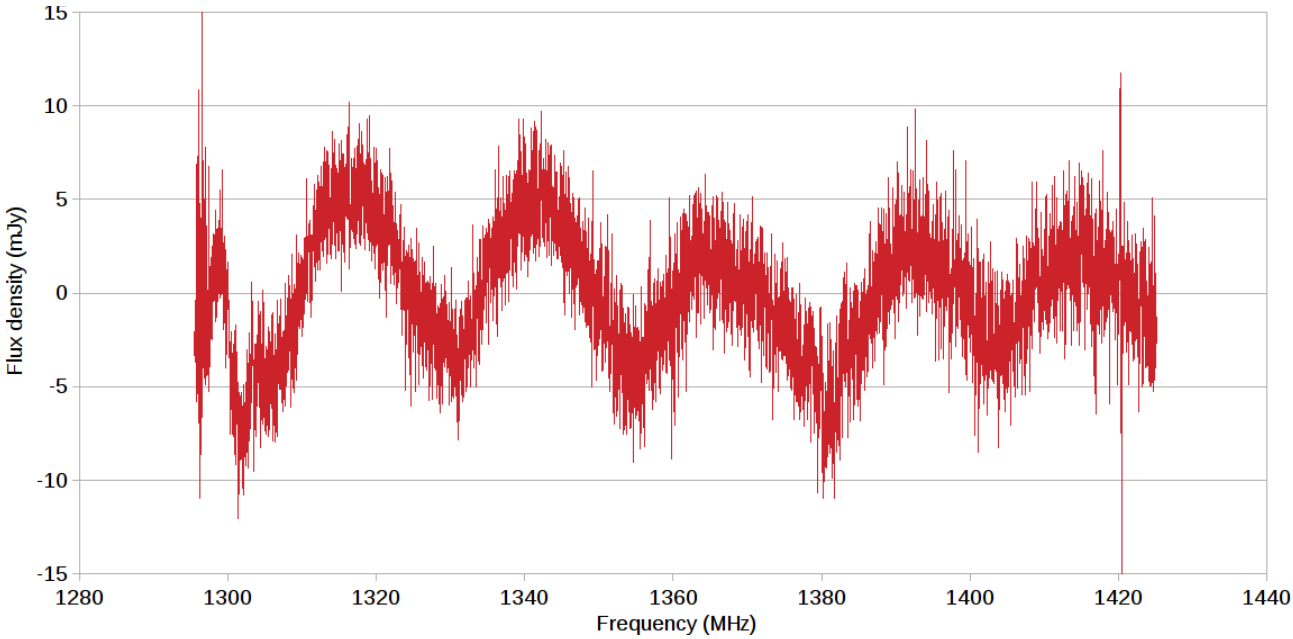

Figure 2: Spectrum in the direction of the source PKS J1028–2644 (~0.9 Jy at 1.4 GHz). A strong periodic ripple is seen that appears to increase in amplitude towards the low-frequency end of the spectrum.

3. Deconvolution

There is evidence of insufficient deconvolution of the H I emission from one of the brightest and most extended galaxies in the Hydra field, NGC 3137, which is of scientific interest due to its long H I tail of presumably tidal origin. A moment-0 image of the region around NGC 3137 is shown in Fig. 3.

While the galaxy itself and its long H I tail are dominating the central regions of that image, faint and extended filamentary emission is visible to either side of the galaxy and tail. Inspection of the actual data cube suggests that this faint emission is caused by residual sidelobes of NGC 3137 and its tail, in particular as the emission appears to be locked in phase space relative to the H I emission from the galaxy/tail and contains regions of negative signal as well.

As the sidelobe emission if rather faint, it is difficult to accurately estimate the residual sidelobe levels, although a rough comparison suggests that they are at or below 5% of the peak flux density of the disc of the galaxy. These findings suggest that the region around NGC 3137 has either not been deconvolved at all, or that insufficient deconvolution has resulted in the presence of residual sidelobes.

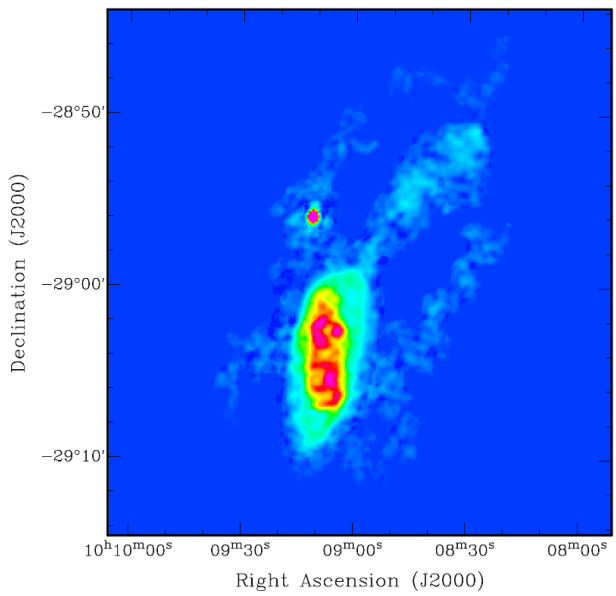

Figure 3: Moment-0 image of the region around NGC 3137 from SoFiA, showing the galaxy itself, its extended H I tail to the upper-right and possible sidelobes to either side of the galaxy and its tail. The bright point source seen against the upper-left sidelobe is a background continuum source, PKS 1006–286.

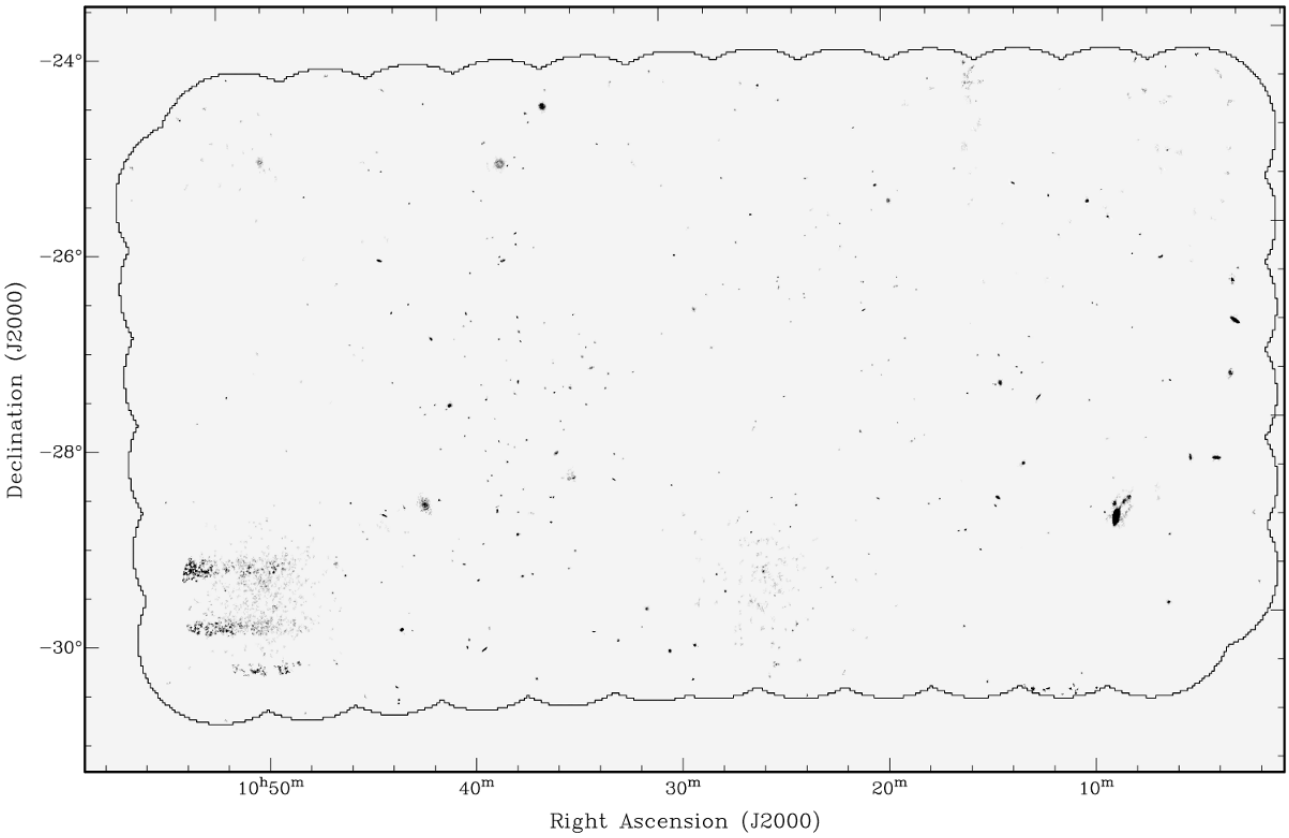

Figure 4: Combined moment-0 map from SoFiA 2 over the velocity range of cz ~500 to 14,000 km/s. Significant artefacts are visible in the lower-left corner of the map, with minor artefacts along the top and bottom edges as well.

4. Flagging and artefacts

Overall, SoFiA 2 still picks up quite a few artefacts across the entire data volume spanned by the observation. While several bad channels are taken out by SoFiA's auto-flagging algorithm, regions of large-scale artefacts remain, as seen in a moment-0 map of the entire field shown in Fig. 4. The most notable example is an apparently bad beam creating significant artefacts near the lower-left corner of the mosaic, thereby hampering our source finding efforts in that region.

However, several other strange artefacts appear across other parts of the field and in particular near the upper and lower edge of the mosaic. At first glance these artefacts would appear like regions of stochastic noise. However, a closer inspection reveals that those noise regions got picked up by SoFiA 2 because of a general offset of the entire bandpass that is too small to be noticed in an individual line of sight, but appears prominently in a spectrum averaged over a larger region covering the full extent of the artefact (Fig. 5).

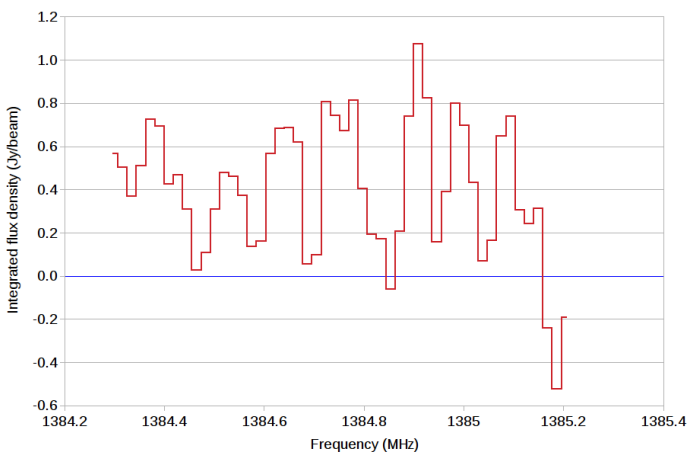

Figure 5: Integrated spectrum over a larger region around one of the small-scale, noise-like artefacts found in some regions of the field. The significant positive offset is clearly visible after application of a Hann filter of width 3.

While the origin of these artefacts are currently unknown, it would appear that they are associated with specific beams, e.g. beam 26 of footprint 2A near the lower-left corner of the map. The extended, filamen-

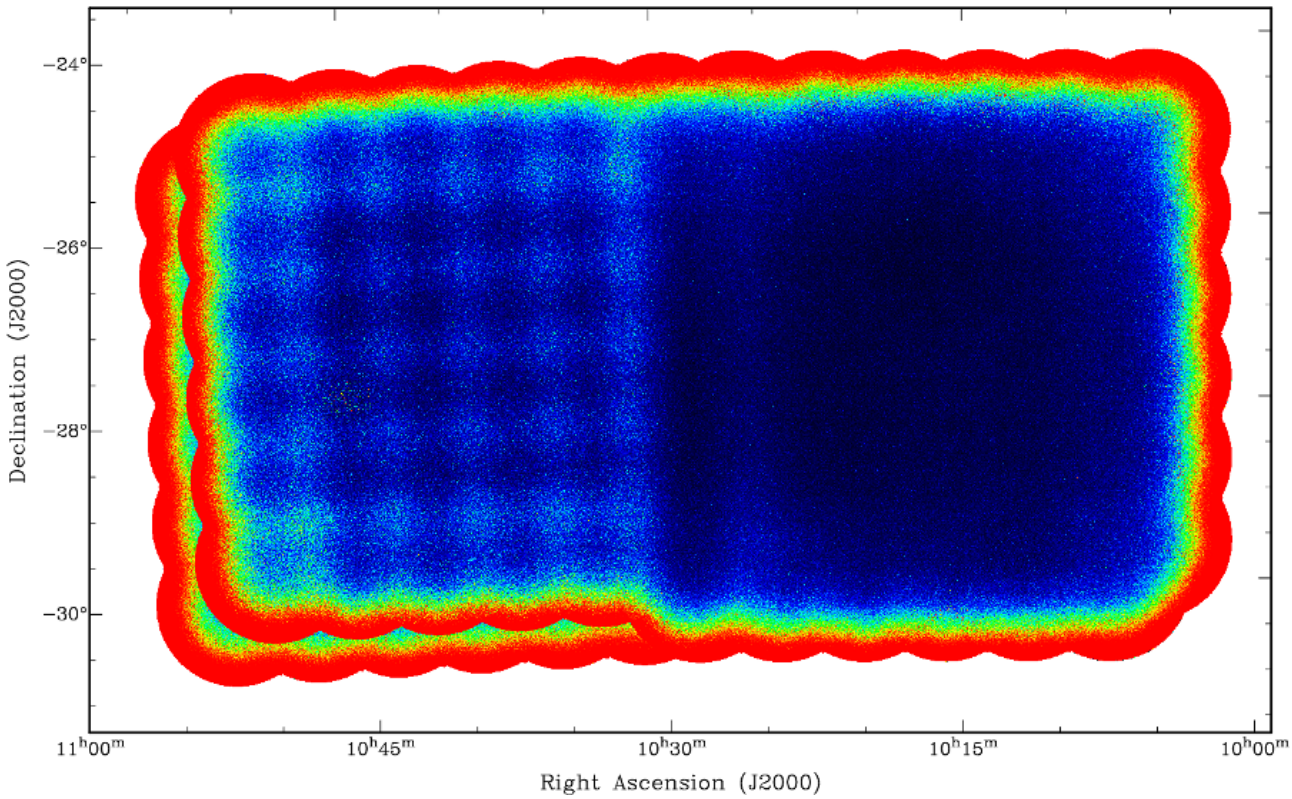

Figure 6: Map of peak flux density in the channel range of 750–1749. The interleaving footprints in the left half of the appear as if they had been pasted on top of one another due to missing data in certain frequency ranges.

tary artefacts seen in beam 26 all appear to be located in frequency ranges affected by correlator drop-outs (see Section 5), while the more compact artefacts scattered across various regions of the map (including in beam 26) seem to be spread across the entire frequency range.

5. Correlator issues

Fig. 6 shows a map of the peak flux density within the channel range of 750–1749. It reveals that the two interleaving footprints A and B in the left half of the field have not been mosaicked as expected. With the help of Karen Lee-Waddell and Matthew Whiting, this was identified as being the result of missing data in certain frequency ranges of one of the two footprints, possibly due to correlator drop-outs. This leaves just one footprint to contribute to the mosaic, thus resulting in a factor of $\sqrt{2}$ increase in the noise level.

A measurement of the noise level towards the centre of the left tile with SoFiA 2, shown in Fig. 7, reveals that two small frequency ranges at about 1325 and 1385 MHz are affected, as indicated by a systematically higher noise level. Each window is 4 MHz wide which is exactly the width of one correlator card, pointing at a correlator issue as the cause.

The right one of the two affected ranges is at a redshift of $z \sim 0.025$. Fortunately, this is beyond the far end of the Hydra cluster and in a redshift range without any overdensity in 6dF, so the missing data should not affect our science in any significant way.

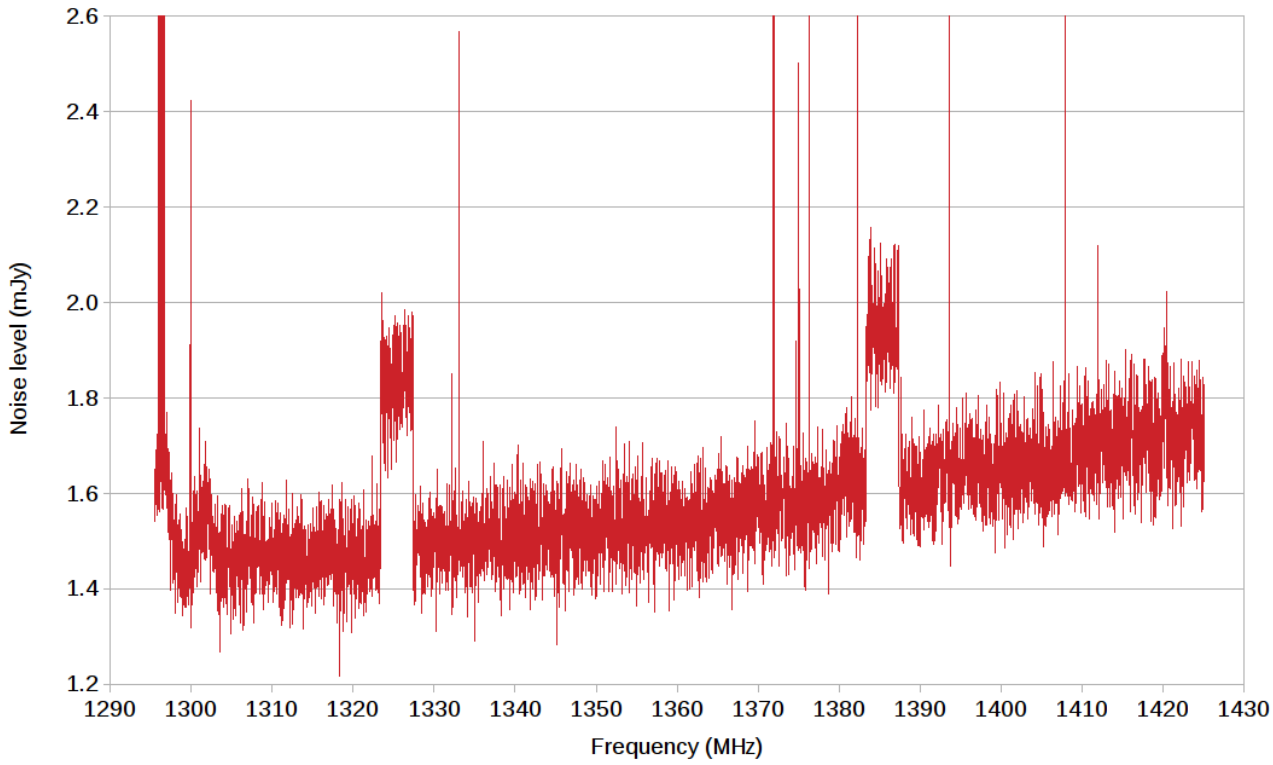

Figure 7: Noise level towards the centre of the left tile as a function of frequency. Two windows of missing data in one footprint at about 1325 and 1385 MHz are clearly visible in addition to several narrow RFI spikes.