RFI at $u = 0$ in MeerKAT observations

F. M. Maccagni, J. Healy, W. J. G. de Blok, P. Serra

January 28, 2022

1 Stripes in the HI-datacubes

Coherent horizontal stripe artefacts are visible in most datacubes of MHONGOOSE, MeerKAT Fornax Survey (MFS) and open-time observations (50 hours, 30 and 10, hours of integration time, respectively. This affects the analysis of low S/N HI gas that can lie over and between these stripes. Besides being visible in the full-survey (55 hours) datacubes, these stripes are picked up by the source-finders when setting a detection threshold of 2σ (Fig. 1).

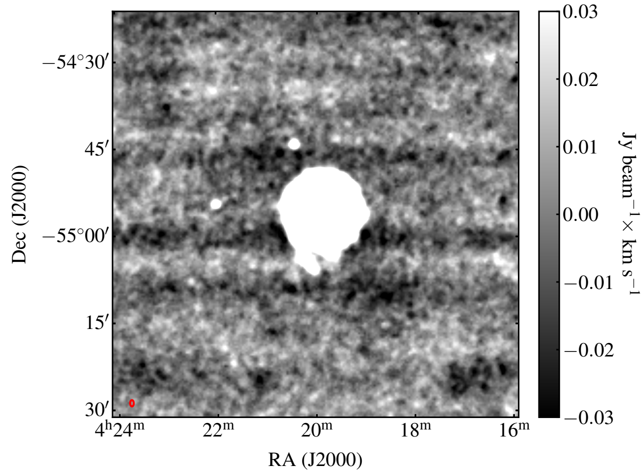

Figure 1: Moment-0 map built by `SoFiA` with a threshold cutoff on the S/N of 2σ.

These artefacts are present also in single MeerKAT 5 hours observations both in the wide-band and in the narrow-band spectral setup (26 kHz and 3 kHz channel widths, respectively). Fig. 2 show images from datacubes of 32k open-time and narrow-band MHONGOOSE observations, where 20 or 200 channels were binned together. The signal of the stripes increases when spectrally binning the visibilities. Stripes are visible in a single channel of the 32k observations, but are not visibile in single channels of narrow-band observations (see Sect. 4 for further details).

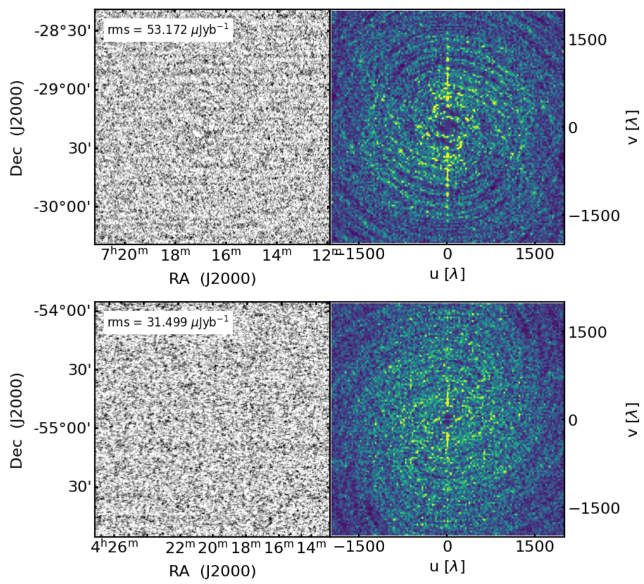

Figure 2: *Top Panels.* Image (and its FFT) of a wide-band observation built binning together 20 line-free channels. *Bottom Panels.* Images (and its FFT) of a narrow-band MHONGOOSE observation built binning together 200-channels.

2 RFI at $u = 0$

A periodic function in the image plane corresponds to a double point function in the (u, v) plane, because the Fourier transform of a sine function is the sum of two opposite imaginary δ functions. Consequently, to investigate the importance of the stripes in the (u,v) plane we calculate the FFT of the 200 channels ($\sim$ 650 kHz) images. The FFT images shown in the right panels of Fig.2 are the amplitudes of the FFT of the 650 kHz map. The images affected by the stripes by the horizontal stripes have in the FFT high amplitudes along the vertical direction ($u = 0$ and $v = \pm V$).

In the (u,v)-plane of the observed visibilities is very difficult to identify the ones that generate the stripes in the datacubes. We find that binning the channels together, and gridding the visibilities (through weighting and tapering) is key to identify the stripes. Considering only the amplitudes of the FFT and not the phases also increases the prominence of the stripes. Hence, in a gridded FFT of a continuum-subtracted image it is much easier to identify the vertical line (i.e. the stripes) rather than in the original visibilities.

RFI at $u = 0$ is visible in the observations of all MHONGOOSE galaxies and of the MFS but is not constant between observations, or even within a single observation. Fig. 3 shows that for J0419-54, observing tracks with low elevation ($\sim 40°$, top panels) have brighter stripes than tracks with high elevations ($> 70°$, bottom panels).

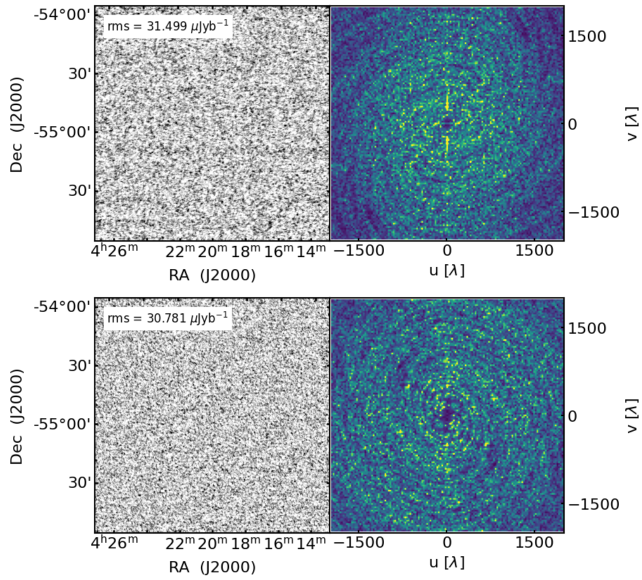

Figure 3: *Top Panels.* Image and its FFT of a low elevation track. *Bottom Panels.* Image and its FFT of a high elevation track.

The elevation dependence of the stripes' brightness is present also within the same observing track (Fig. 4, left panels). Each 5.5h MHONGOOSE observing track is subdivided in 5 on-target scans with a durations of just under 1 hour, scans at low elevations (3,5,7) show brighter stripes than scans at high elevations (9,12).

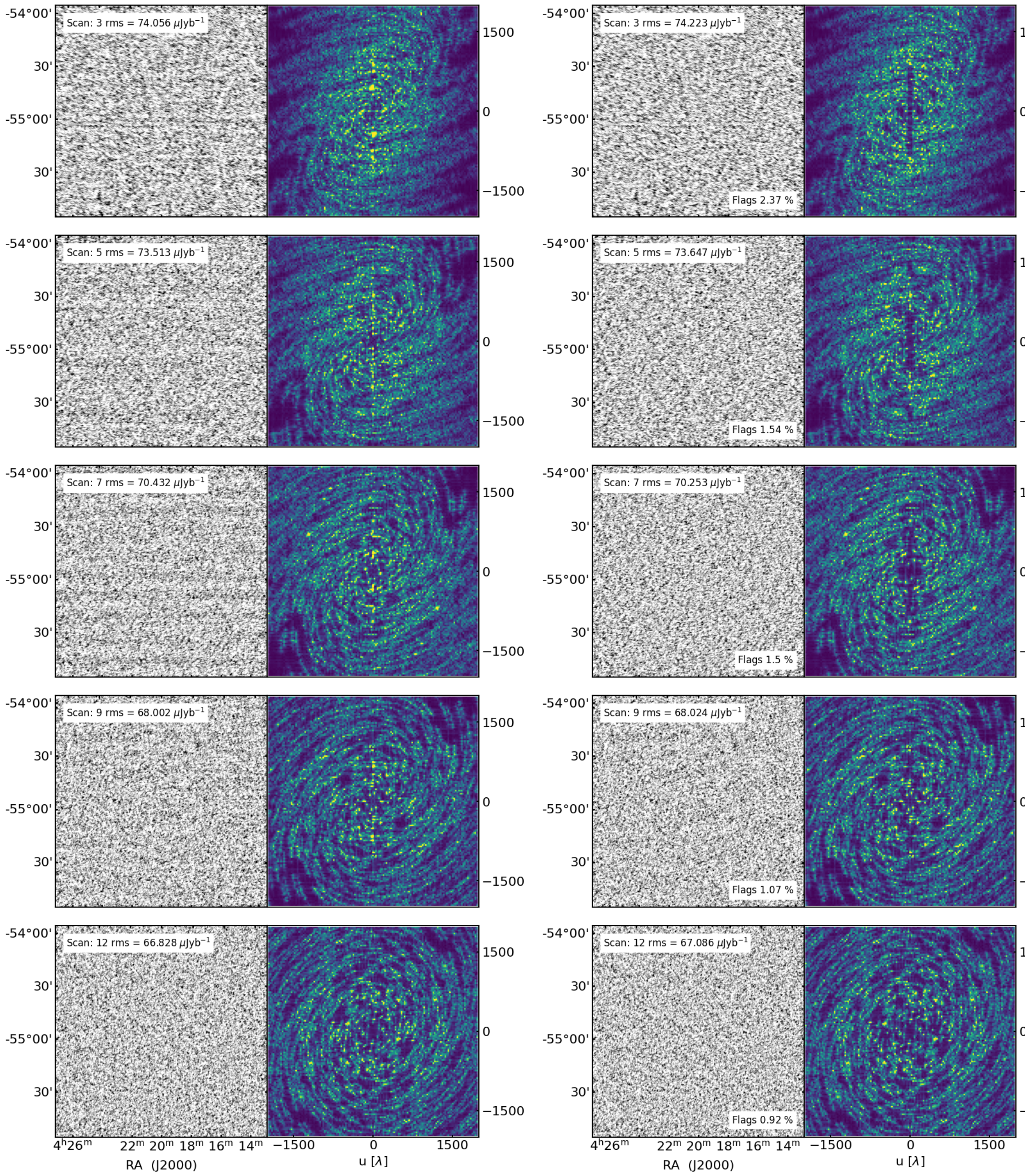

Figure 4: *Left panels.* Image (and its FFT) of each scan of a narrow-band observation. *Right panels.* Image (and its FFT) of the same visibilities after flagging the RFI at $u = 0$ as illustrated in Sec 5.

3 Elevation/azimuth dependence

To gain a better insight about when the stripes appear in the visibilities, we study the presence of the stripes in 550 scans of the MHONGOOSE galaxies and two tracks of the MFS (J0338-3523) and open-time observations (J1432-4411). Given their heterogeneous distribution in the sky, we can investigate if the presence of the stripes depends on the azimuth and elevation.

For each scan on the target galaxies, we create a 200 channel (~ 650 kHz) continuum subtracted line-free image and its FFT. We image using robust=1.5 and tapering=60″ to maximise the stripes' brightness. In the FFT we compute the 99.99 percentile of the amplitudes. When a bright vertical line is present in the (u,v) plane (i.e. a stripe in the image), more than 55% of the points above this percentile are located along the v-axis ($u = 0$). Fig. 5 shows the azimuth and elevation for the 550 scans of our sample. The colour coding of the points indicates the presence of the stripes ($\gtrsim$ 55%, red colours) or not ($\lesssim$ 55%, blue colours). The colour coding of the az/el slices of Fig. 5 gives the median percentage value of the observations within them. Observations at high elevations (> 70°) have less prominent stripes than scans at low elevations. Stripes also show a slight dependence on azimuth and they are less present in observations towards the North ($315 \lesssim az \lesssim 45$, counter-clockwise).

In our analysis, observations taken at the same elevation and azimuth, do not always show the same level of stripes, this may be partly caused by the fact that we investigate the presence of the stripes on data that has already been flagged by different automated routines throughout the standard processes needed to generate continuum-subtracted visibilities.

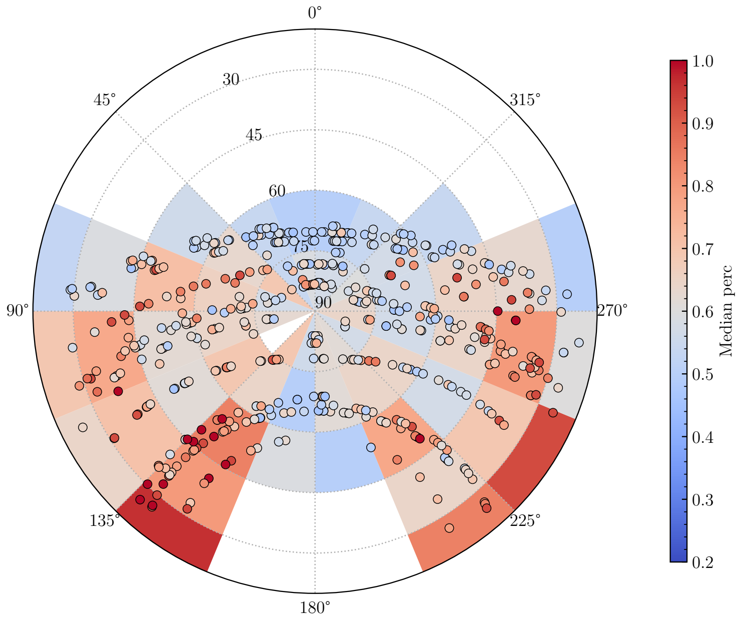

Figure 5: Azimut (longitude) vs. Elevation (latitude) of 550 scans of the MHONGOOSE observations. The colour codes what fraction of all FFT pixels above the 99.99 percentile are at u=0, from none (dark blue) to all (dark red).

4 RFI at u=0 when rms $\lesssim 900\,\mu$Jy beam^{-1}

In our analysis, we create images ~ 650kHz-wide channels to maximise the brightness of the stripes. In the setup of the MHONGOOSE and MFS data reduction this is a large enough line-free frequency range to estimate the extent of the stripe and build a mask of it. This mask is then applied also to the other channels, including those where HI is present (and where, because of the line signal, is not possible to identify the stripes in the FFT).

Besides binning the channels together, also weighting the image and tapering play a role in highlighting the stripes in the FFTs. After several trials we find that the best grid in the (u,v) plane to identify the stripes is given by images generated with robust= 1.5 and gaussian tapering = 60″.

The stripes become first visible when the noise in the binned images fall below a noise of $\lesssim 900\mu$ Jy beam^{-1}, independently of the original channel width of the input images. When imaging a wide-band 5 hour observation, stripes are visible in the scans of single channels (26.5kHz wide, see scan 5 and 8 of Fig. 6 left). Similarly in narrow-band observations, stripes are present in some scans when we bin together four 6.5kHz-wide channels (26.5kHz wide image, see scan 3 and 5 of Fig. 6, right).

In a 5-hour MFS observation, the stripes have a peak amplitude of 60μJy beam^{-1}, while the noise is ~ 100μJy beam^{-1}. The stripes add coherently: in 50 hours the noise goes down to 30μJy beam^{-1}while the stripes stay the same. Their amplitude is then ~ 2 times the noise level. This explains why `SoFiA` picks up the stripes at this threshold (Fig. 1).

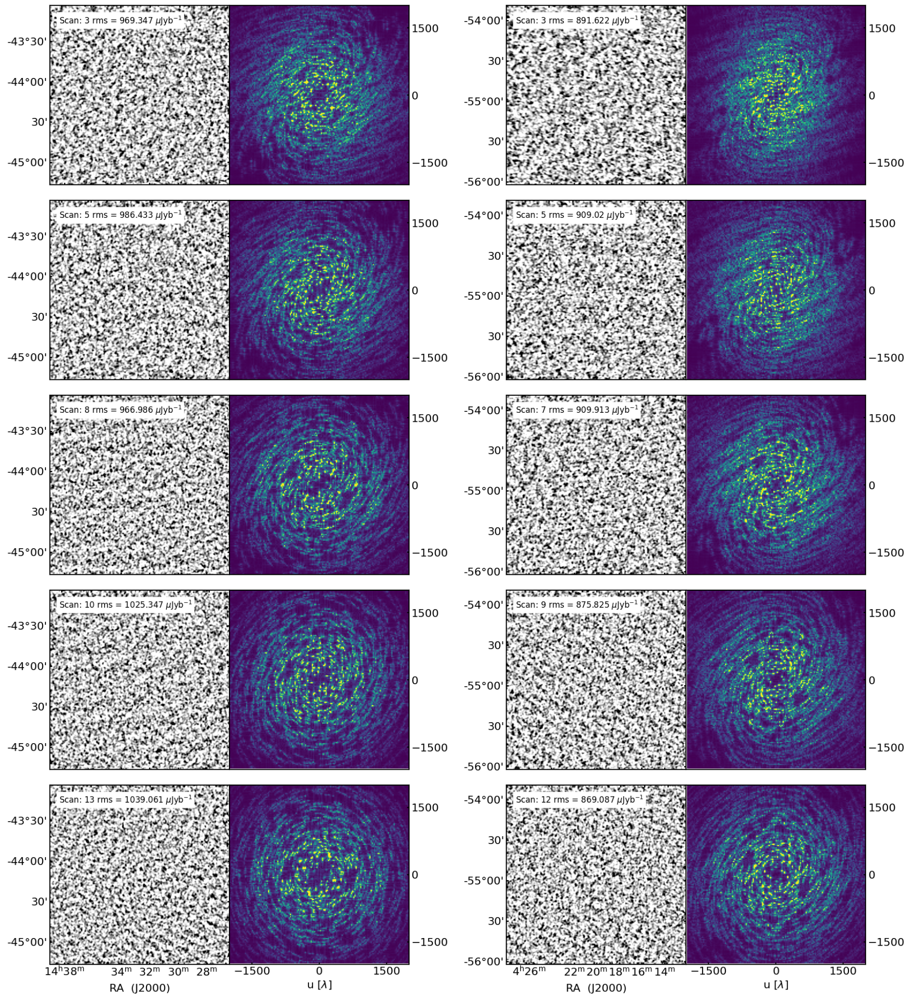

Figure 6: *Left panels.* Single channel image (and its FFT) of each scan of a 32k open-time observation. Stripes are visibile in scan 5 and 8 *Right panels.* Image (and its FFT) of a narrow-band observation where we binned four 6.5kHz-wide channels together ($\Delta\nu$ ~ 26 kHz). Stripes are visibile in scan 3 and 5.

5 Flagging strategy

The strategy used to identify the presence of the stripes in each scan can be used to flag the continuum-subtracted visibilities and remove these artefacts, without losing too many visibilities in the central regions of the (u,v)-plane. By selecting all visibilities corresponding to the 99.99 percentile of the amplitudes in the FFT of a 200 line-free channel ($\sim$ 650 kHz) image (as illustrated in Sect. 3), we flag on average per scan $\sim$ 1.4% of the visibilities. By eye, the stripes are not visibile anymore in the 200 channel images and in the FFTs of the single scans (Fig. 4, right), and in the 200 channel image of the full tracks (Fig. 7, bottom left). Nevertheless, when we create a cube at low spatial resolution ($\Theta \sim 60''$) binning three channels together, low-level stripes are still present, which compromise the reliability of our low S/N HI detections (see the moment-0 map in Fig. 7, right).

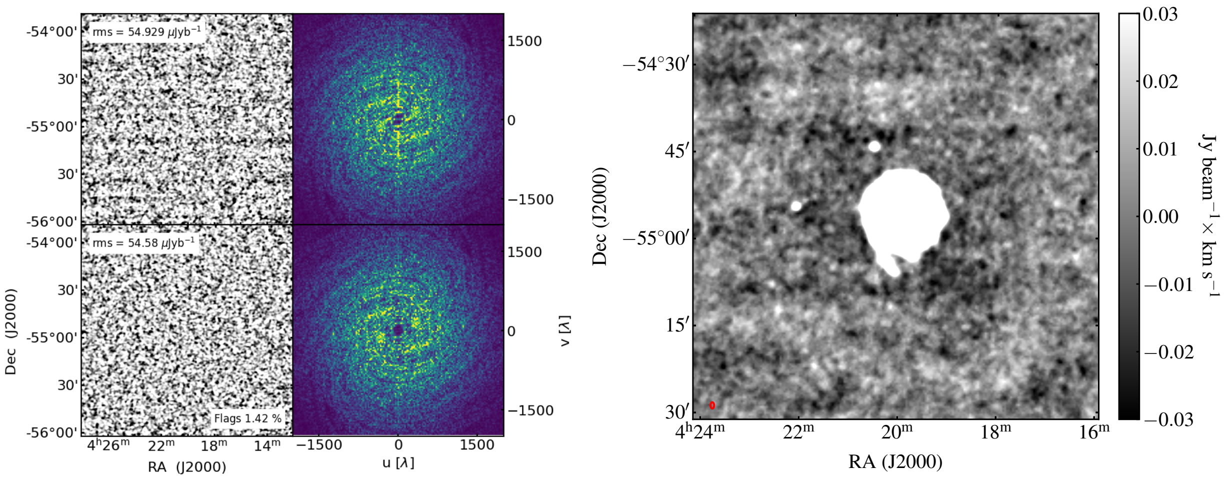

Figure 7: *Left Panels, Top.* 200 channel image (and its FFT) of a 5.5 hours MHONGOOSE observation before flagging the stripes. *Bottom.* Same as the top image after flagging 1.4% of the visibilities, with amplitudes in the FFT above the 99.99% of the distribution. *Right Panel.* Moment-0 map built by `SoFiA` with a threshold cutoff of 2σ after flagging the amplitudes above the 99.99 percentile.

In our H I studies, we use automated source finders (e.g. `SoFiA`) to identify the cold gas emission in the datacubes. Our source-finder strategies generally use low detection thresholds (e.g. $\sim 2\sigma$) and different smoothing kernels. Ideally, the stripes should not be included in these automated detection runs, while coherent low signal-to-noise line emission should. Hence, to decide if stripes have been sufficiently removed, we check if they are not visible by eye in the 55-hour datacubes of the MHONGOOSE observations and in their moment-0 map created by `SoFiA`, with a threshold of 2σ.

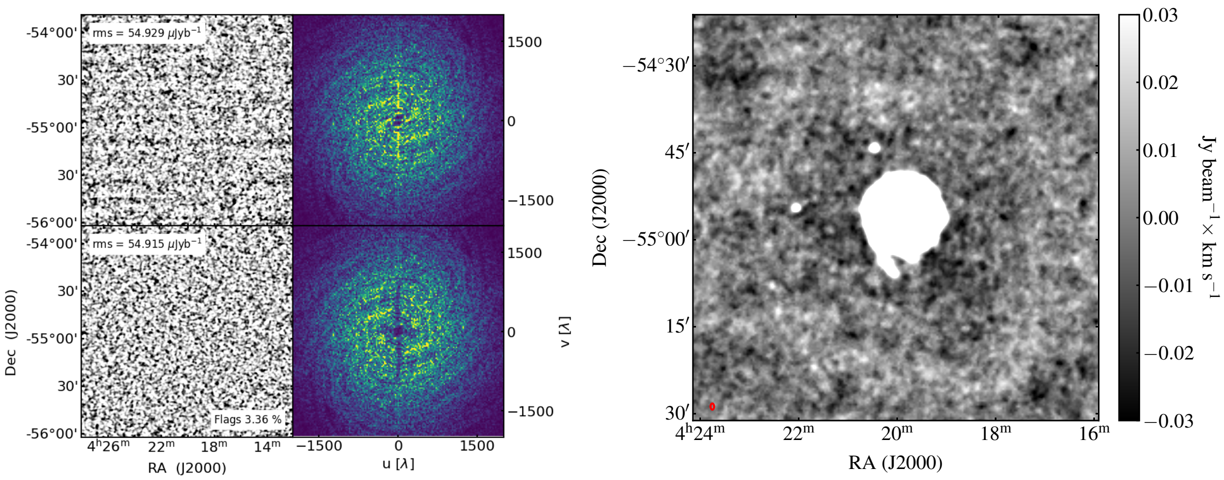

Figure 8: *Left Panels, Top.* 200 channel image (and its FFT) of a single track MHONGOOSE observation. *Bottom.* Same as the top image but flagging the visibilities with amplitudes in the FFT above $300 \times mad$. *Right Panel.* Resulting moment-0 map built by `SoFiA` with a threshold cutoff on the S/N of 2σ.

The best results are given by refining the statistics by which we select the stripes in the FFT of the 100 channel image (*robust* = 1.5, taper= $60''$). The stripe is identified by the 'hot pixels' in the FFT. By measuring the median absolute deviation (mad = median($|a_{i,(U,V)}$ − median($a_{i,(U,V)})|$)) of the distribution of the amplitudes (a_i) in the FFT plane (U, V), we find that the cutoff (median($a_{i,(U,V)}$) + 300 × mad) allows us to efficiently get rid of the stripes in each scan, resulting in flagging on average 3% of the data (as shown in Figs. 8 left and Fig.9). The noise in the image is slightly lower, likely because removing coherent artefacts the distribution of the signal is closer to gaussian. The right panel of Fig. 8 shows that by this method the stripes are almost negligible also in the very sensitive moment-0 map made by SoFiA (cleaning errors are now dominating).

We will apply the flagging strategy illustrated above on all MHONGOOSE and MFS observations. In the continuum-subtracted MS dataset we will flag in each scan and over the entire spectral band the visibilities with (u,v)-coordinates that in the FFT of a 650kHz line-free image[1] have amplitudes median($a_{i,(U,V)}$) + 300 × mad.

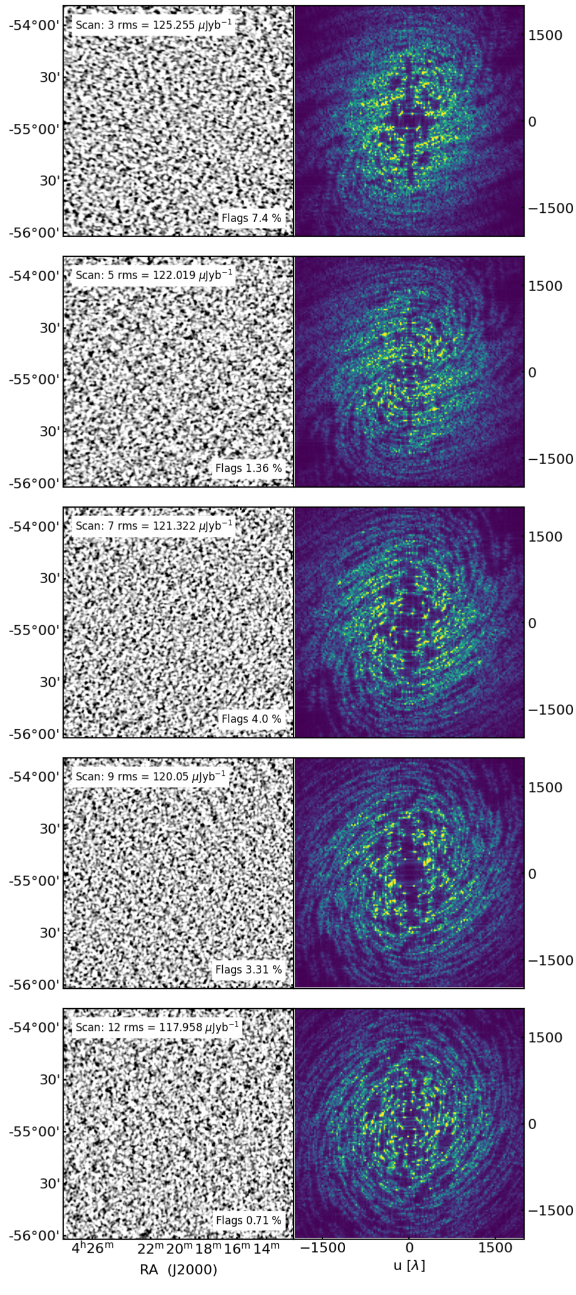

Figure 9: Image (and its FFT) of each scan of a 5.5 hours MHONGOOSE observation after flagging the visibilities identified as the RFI at u=0 by having amplitudes in the FFT above 300 × mad.

[1] generated with robust=1.5, taper 60 arcseconds, image size 400x400 pixels of 20″