GBT+WSRT Observations of NGC 4631 Working Group

Kelley Hess (ASTRON & Kapteyn Astronomical Institute, Netherlands)
Blazej Nikiel-Wroczyñski (Jagiellonian University, Kraków & Leiden Observatory
Nickolas Pingel (Australian National University, Canberra, Australia)
Pedro Salas (Green Bank Observatory, WV, USA)
Aleksandar Shulevski (API - Amsterdam & ASTRON - Dwingeloo, Netherlands)
Vaibhav Vaidya (Leiden, Netherlands)
September 17, 2019

This memo was prepared as part of the workshop "Improving Image Fidelity on Astronomical Data: Radio Interferometer and Single-Dish Data Combination," held on 12-16 Aug 2019 at the Lorentz Center in Leiden, The Netherlands.

1 Dataset overview

The provided high-resolution neutral hydrogen (HI) data are from the Westerbork Synthesis Radio Telescope (WSRT) Hydrogen Accretion in LOcal GAlaxieS (HALOGAS). See [1] for a full survey description. The single dish counterpart data are from the Green Bank Telescope [2]. The 'feather' and 'Faridani' image combination techniques utilize the publicly available cubes, while the joint devolution methods use a single spectral channel at 1417.31 MHz of the calibrated visibilities. Table 1 summarizes the data properties, and Figure 1 presents HI column density images of the original images

Table 1: Observational details

Parameter	Value
Phase center	J2000 12h42m08.01s +32d32m29.4s
Rest frequency [MHz]	1420.405
V_{Helio} [km s^{-1}]	605
ΔV (channel width) [km s^{-1}]	4.12
Velocity range imaged [km s^{-1}]	395 - 828
Map size	1.13×1.13 deg^2
WSRT Noise [mJy/beam]	0.3
GBT Noise [mJy/beam]	5
WSRT Beam Size [□″]	44.96×39.14 (PA=15°)
GBT Beam size [□″]	546×546
Range of WSRT baselines [m]	36 - 2700
GBT Diameter [m]	100

2 Combination Methods

2.1 Method 1: Feather

We utilized several software packages including `CASA`'s `FEATHER`, `AIPS`'s IMERGE, `miriad`'s `IMMERGE`, and `uvcombine` to implement the feather combination technique. A visual inspection of the resulting

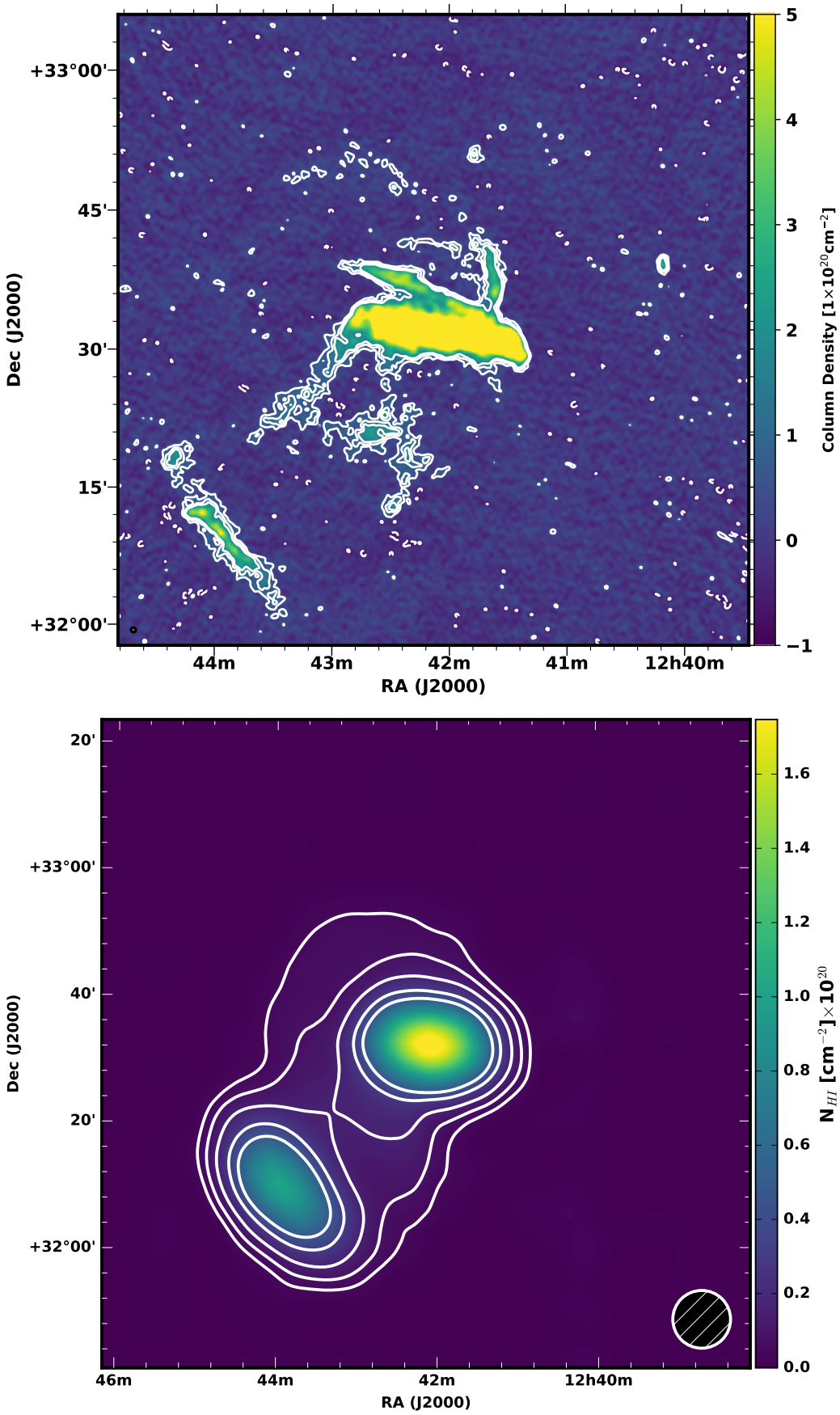
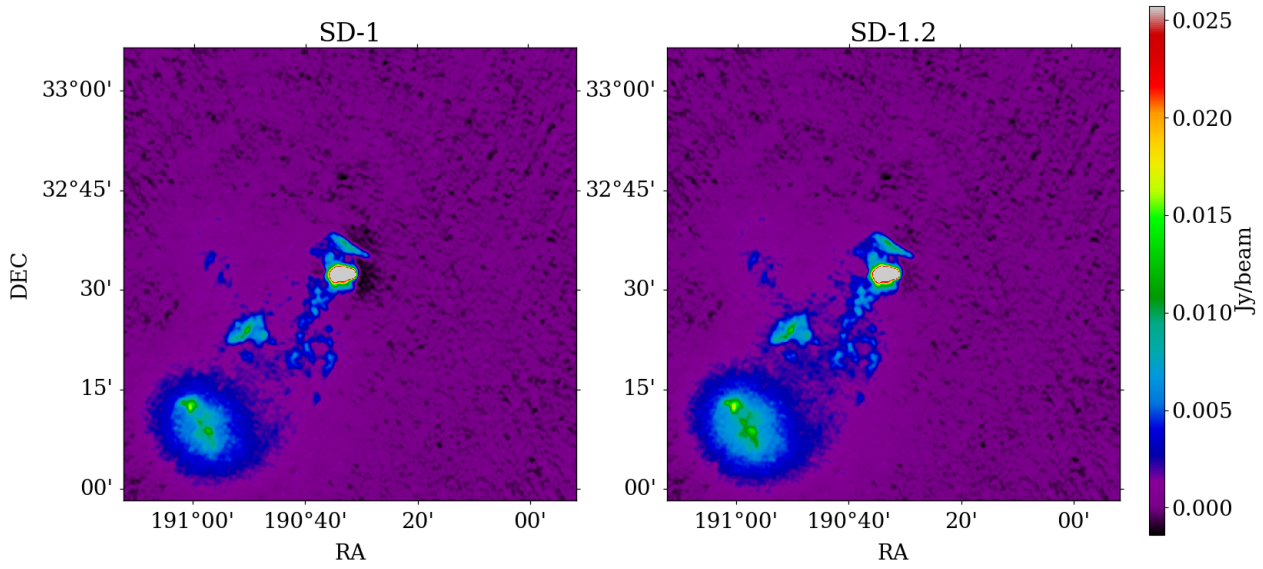


Figure 1: HI column density iamges of the priginal WSRT (top) and GBT (bottom) cubes. The contours in the WSRT map begin at 5×10^{19} cm^{-2} and continues at 10 and 15 times that level, while the contours in the GBT images begin at 4×10^{18} cm^{-2} and continues at 2, 4, 8, and 12 times that level.

combined images from each respective software package provided an initial assessment of the quality. Setting `sdfactor=1.2`, which is a factor applied to the low-resolution data to correct for calibration errors, in `CASA`'s `FEATHER` produced the best results: no negative bowl around the source introduced, no excess "halo" emission, most likely coming from the SD data. Letting `miriad`'s `IMMERGE` solve for the required scaling factor by where the real and imaginary amplitudes of the low-resolution and high-resolution fitting a line between the real and imaginary amplitudes of the resolution-scaled low-resolution and high-resolution data in the Fourier plane within a given annulus of $u-v$ ranges ranges also produced a value of 1.2. However, in both cases, the total flux measured in a region encompassing the target galaxy was greater than the total flux in the GBT image. Changing the `effdishdiam` in `CASA`'s `FEATHER` task re-introduces the negative bowl. The channel maps at 1419.31 MHz shown in Figure 2 summarizes the effect of increasing the scaling factor from 1.0 to 1.2: the negative bowl around NGC 4631 is mitigated and more extended emission is visible. However, it is clear that the emission surrounding the companion, NGC 4656, is dominated by the contribution from the GBT and unphysical. This is due to its location well outside the primary beam of the WSRT. The WSRT data had not been primary-beam corrected, and we have used We have used `CASA v5.6`. **BNW inserts text from here...** Similar efforts undertaken with the `AIPS` task `IMERG` yielded comparable results as with the default setup of `FEATHER` in `CASA`. Choosing different boundaries of the spatial frequency overlap and/or a different scale factor also did not produce an appreciably better map.



*

Figure 2: Feathered images, `sdfactor`=1.0 (left) and `sdfactor` 1.2 (right).

2.2 Method 2: Joint Deconvolution (tp2vis)

We have encountered issues with how CASA extracts images from multi-channel ones. CASA always sets CRPIX3=0.0, so `qac_clean` crashes. Using a `Miriad` produced image, `qac_tp_vis` and

qack_clean work.

To use a different dish size for your total power data you have to set:

```
qac_vp(1)
dish = 100
qac_tpdish('ALMATP', dish)
qac_tpdish('VIRTUAL', dish)
tp2vis('NGC4631_GBT_Jy_REGRID_1417.31MHz.im', 'test.ms', 'WSRT.ptg', maxuv=100)
```

It is necessary to keep the order outlined above (initialise voltage pattern table BEFORE setting the sizes of telescope dish). Doing it the other way around results in an incompatible amplitude scale applied for the single dish data. Figure 3 presents the amplitude and weights as a function of u-v distance for the merged data set. The agreement between the GBT pseudo-visibilities generated by tp2vis and WSRT in the overlapping u-v range of ∼40 to 100 m indicates an acceptable merger of the two data sets.

We imaged the merged measurement using qac_clean and CASA's clean separately. The results of a series of different input parameter tests are summarized below:

- Qualitative result: the image made from the merged dataset is similar to the one made using WSRT data only; there are no large-scale structures (both real and artificial) visible; the only visible difference is in the smallest-scale structures (different noise fluctuations pattern?);
- Quantitative result: total flux of the observed structure is NOT higher in the merged map, than in the WSRT one; the exact difference is non-significant;
- Difference between the qac_clean and clean: the initial test suggested that the latter behaves better; however, we performed series of subsequent tests to isolate the effects of particular parameters. The difference arises because of the different gridding kernel used by default by qac_clean (mosaic) instead of standard). Unfortunately, if both tasks are being run with exactly the same set of parameters, the results do not differ at all, suggesting that the VP table is not taken into account; this problem has been already reported here: https://github.com/teuben/QAC/issues/8
- tclean has an optional parameter that should allow to inject a VP table; it is not used by the qac_clean task.

A comparison image of a qac_clean (with standard and mosaic modes) and casa imager is shown in Figure 4.

2.3 Method 3: SDINT

We generated a PSF image of the GBT using the provided make_gauss_beam.py script. We then had to re-ordered the axes to ensure the spectral axis was listed last. Since the raw visibilties only contained a single spectral channel, we set these methods in the runsdint.py script:

```
deconvolver='mtmfs'
specmode='mfs'
nterms=1
refreq=''
```

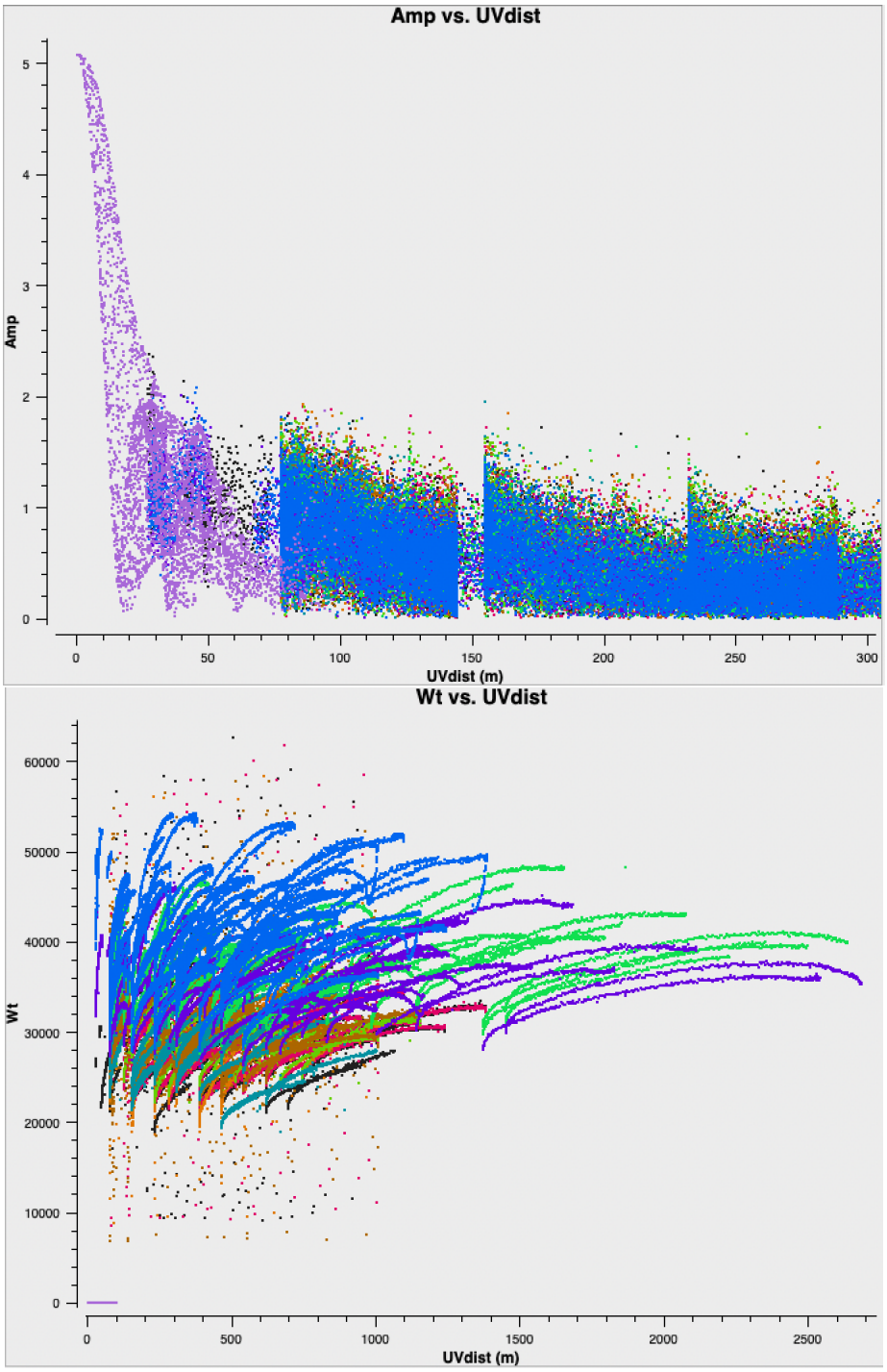


Figure 3: Amplitude (top) and weights (bottom) vs. `u-v` distance from the merged data set from `tp2vis`. Data is colored by observation: purple show GBT pseudo-visibilities, while all other colors represent calibrated WSRT visibilities.

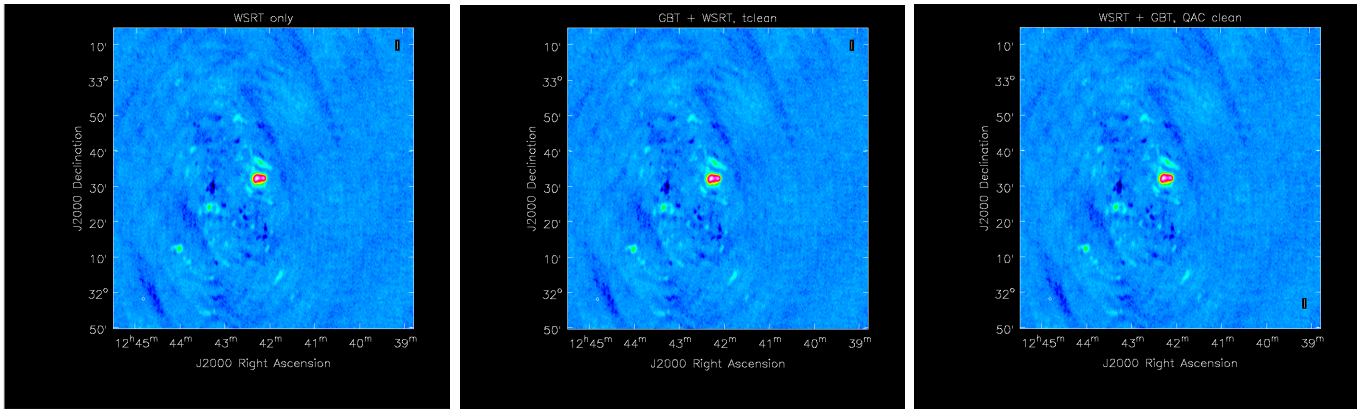


Figure 4: Comparison plots for `tp2vis`. Left to right: WSRT data only, GBT+WSRT imaged with tclean, and GBT + WSRT imaged with qac_clean. Same cleaning parameters (niter, gridder, size, etc.) were chosen for all three images.

Unfortunately, the resulting combined images are blank. The most likely cause is that our input WSRT measurement set is in a format — perhaps related to the single spectral channel — that is not expected within the `SDINT` source code. It would be instructive to try again with calibrated visibilities spanning at least a few spectral channels.

2.4 Method 4: Faridani

(Image Plane) The recent method outlined by [3] combines single-dish and cleaned interferometric data sets in the image domain. The basic procedure is outlined by the equations

$$I_{\rm comb} = I_{\rm WSRT}^{\rm high-res} + \alpha\,(I_{\rm GBT} - I_{\rm WSRT}^{\rm conv})\,, \tag{1}$$

where α is the ratio of the WSRT and GBT beam areas. Advantages of this method include avoiding a Fourier Transform, which introduces sharp artifacts due to Gibbs ringing in the `u-v` domain when dealing with extended emission in the image plane, and a robust flux scaling of the cleaned interferometric data to the single-dish data. We utilized a model of the GBT beam at L-Band (shown in Figure 5 to convolve the WSRT data to GBT resolution, in order to ensure an accurate flux comparison.

The resulting images, which are presented in the next Section, share similarities to those produced by feather — albeit with a more significant bowl around NGC 4631. While convolving the WSRT to GBT resolution instead with a Gaussian with a full-width at half-max of $9.1'$ — the typical Gaussian approximation given the GBT's 100 m diameter and aperture illumination taper — mitigated the prominence of the negative bowl, the artifact still persists. Inspection of the residual image (i.e., the second term in Equation 1 reveals that the negative bowl stems from the additional flux at the center of the convolved WSRT image. The artificial additional flux in the convolved WSRT map demonstrates that these data and the regridded GBT data are not truly at equal angular resolutions, indicating we are not making comprehensive assumptions about the exact shape of the GBT beam.

3 Comparison and Evaluation

Several 'goodness' tests were discussed during the workshop including: comparing the total flux, comparing power spectra, and generating a Q-parameter map. Examples of these tests applied to several of the combined maps are summarized below.

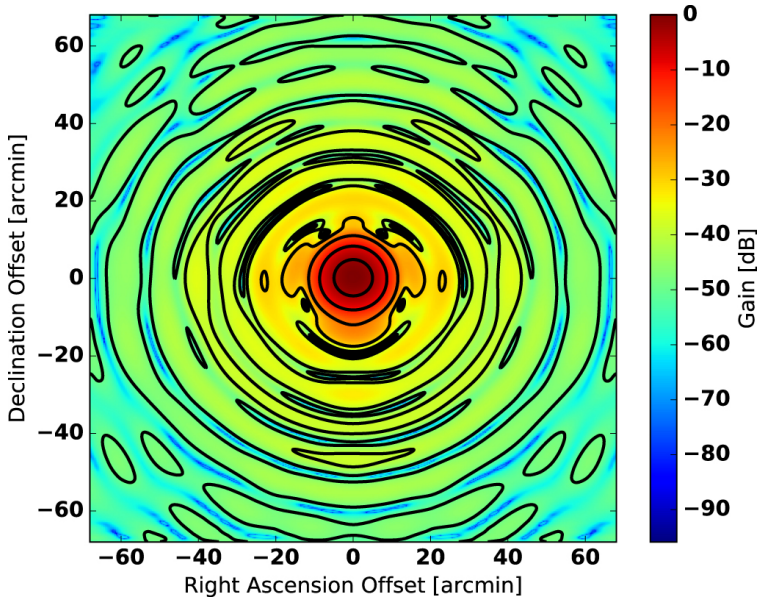


Figure 5: The GBT Beam Model L-Band used to convolve the WSRT image to GBT resolution as part of the Faridani method. Taken from Figure 1 of [2].

3.1 Total Flux

The total flux of the combined image should compare well with the total flux measured in single dish, indicating the combination is at least reasonable. Figure 6 compares the total flux as a function channel number for the cubes generated from several feather methods and cubes produced through the Faridani method, which utilized a the GBT L-Band beam model and Gaussian approximation as the convolution kernels to put the WSRT and GBT data at similar angular resolution. It is clear that the original WSRT data was missing a significant fraction of flux, especially at the receding velocities near the systemic velocities of the companion and most of the extended emission revealed by the GBT. All feather methods, save for `IMMERGE`, recover the total amount of flux well. The mismatch between the GBT and IMMERGE is due to tapering the GBT input cube to match the residual primary beam response in the WSRT image. Turning this option off brings the curves into agreement.

The relative offset between the WSRT and GBT profiles is small at approaching velocities where most of the emission originates from NGC 4631 before increasing substantially at the velocities corresponding to the extended emission. While such an offset would be clear if one just measured a total flux value, a plot such as Figure 6 will reveal where a given combination method is struggling.

3.2 Power Spectrum

The power spectrum is a quantitative measure of how structure is distributed in an image as a function of length scale. Ideally, the power at large angular scales between the combined image and single dish image will be in good agreement. Figure 7 compares power spectra derived from the the various feather methods with the `sdfactor` set to 1.2 at a single channel centered at 1417.31 MHz. The resulting feathered images from `miriad`'s `IMMEGE` and `uvcombine` recover almost exactly the same amount of large-scale power, while a non-negligible amount of relative power is missing from the map produced by `CASA`'s `FEATHER`; the relative power at small-scales is exactly the same for all three methods. This plot demonstrates the power spectrum is a useful tool in gauging the relative

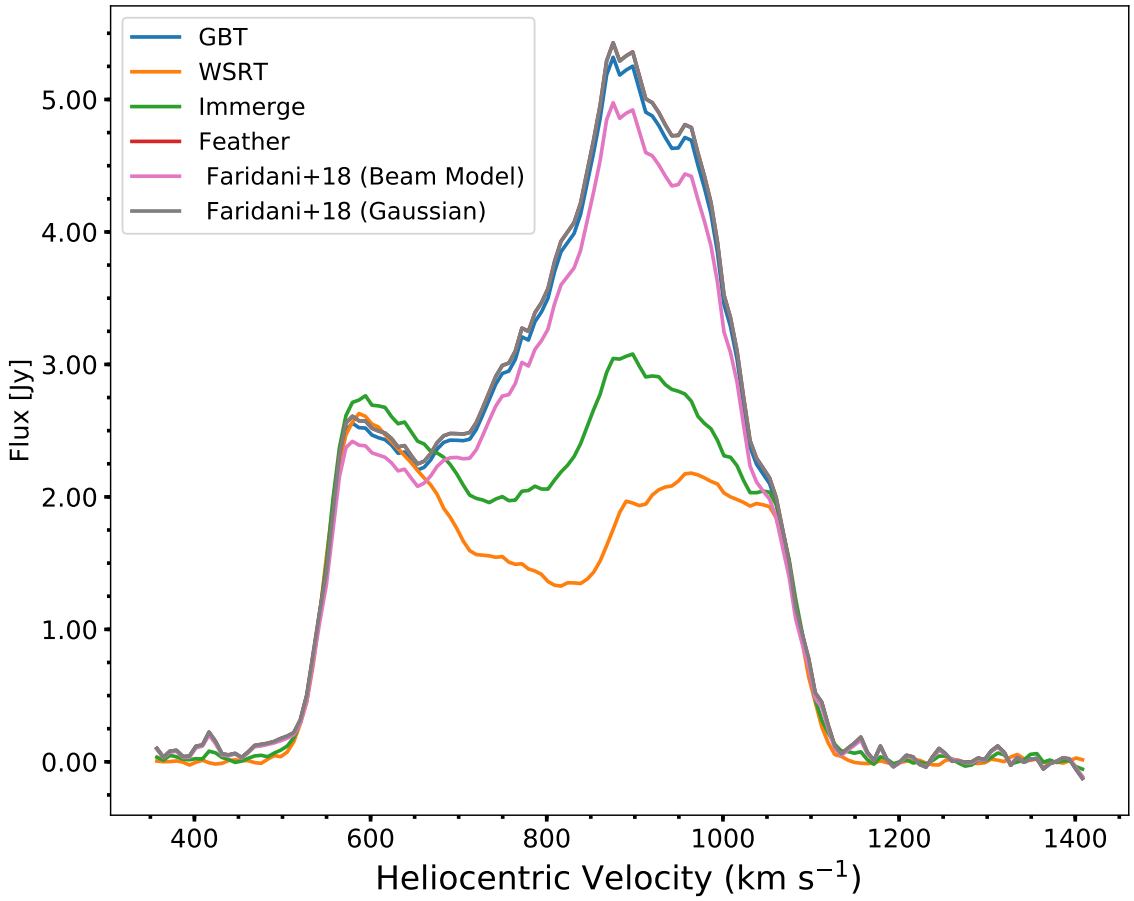


Figure 6: Flux as a function of channel from cubes generated with different combination methods. The 'Feather' profile refers to CASA's FEATHER task with `sdfactor=1.2`.

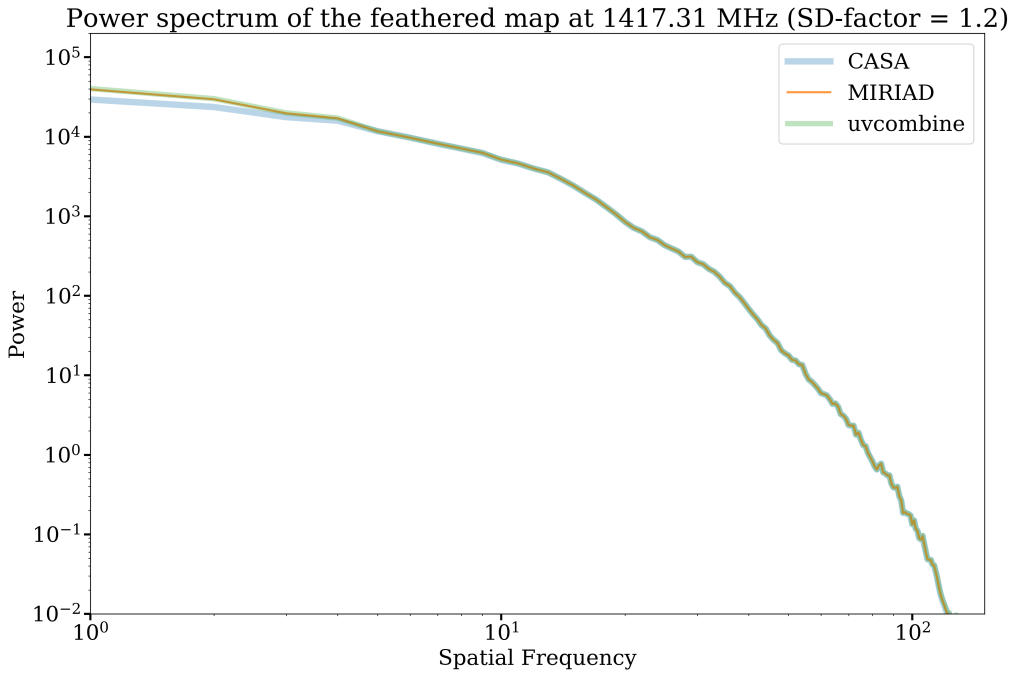


Figure 7: Power spectra from a single channel in the combined cubes generated from several feather methods.

performance of a combination method.

3.3 Q-Parameter Maps

The Q-parameter is defined as

$$Q = \frac{(I_{\rm GBT} - I^{\rm conv}_{\rm WSRT})}{I_{\rm GBT}}. \tag{2}$$

If Q is near zero, the resulting combination has recovered much of the large-scale structure present in the single-dish image. Q is computed on a pixel-by-pixel basis, allowing for a spatial comparison of the recovery of the large-scale structure. Example Q-parameter maps from the comparing a single channel from the original WSRT map, feathered map (with `CASA`), and Faradini map (with a beam model convolution kernel) are presented in FIGURE. In the original WSRT image, Q is relatively close to zero around NGC 4631 but deviates significantly away from zero towards the companion. The Q maps generated from the two combined cubes show Q is much closer to zero in these same regions. Interestingly, the deviation from zero in the Faridani Q-map around NGC 4631 traces exactly the negative bowl artifact.

4 Conclusions

We investigated several methods to combine single dish and interferometric data in both the image and `u-v` domain utilizing WSRT and GBT observations of the interacting galaxy NGC 4631. We successfully created combined images with several feather techniques and the joint deconvolution method, `tp2int`. Most likely to due the formatting of the provided visibility data, we were not able to create images using the `SDINT` package. When feathering in several software packages, the

`sdfactor` — a factor applied to the single dish data to put the two data sets on the same flux scale — needed to be set to 1.2 in order to mitigate a negative bowl artifact in the combined images; however, the total flux in the combined images would result in total flux measurements above what the GBT measured, indicating the calibration of the GBT data may be suspect.

We carried out several 'goodness' benchmarks — including comparing the flux along the spectral axis, measuring the respective power spectra, and generating Q-parameter maps — on the combined cubes generated from the various feather methods in order to characterize relative differences between the available software packages. The small difference revealed by these benchmarks indicate that such tests are useful to quantify the prominence of such as a negative bowl caused by a flux scale offset, for example.

A definitive and ubiquitous benchmark that describes how close a combination resembles the true sky distribution remains to be defined (if such a test is even possible). Applying the benchmarks outlined in this work and others developed as part of this workshop to the combination products of combination several methods should help one distinguish which technique is right for their specific data set. Additionally, when possible, the same combination methods should be applied to models of the object of interest with a known flux, morphology, and perhaps even kinematics to help converge on the most useful technique for one's science.

Bibliography

[1] G. Heald, G. Józsa, P. Serra, L. Zschaechner, R. Rand, F. Fraternali, T. Oosterloo, R. Walterbos, E. Jütte, and G. Gentile. The Westerbork Hydrogen Accretion in LOcal GAlaxieS (HALOGAS) survey. I. Survey description and pilot observations. , 526:A118, Feb 2011.

[2] N. M. Pingel, D. J. Pisano, G. Heald, T. H. Jarrett, W. J. G. de Blok, G. I. G. Józsa, E. Jütte, R. J. Rand, T. Oosterloo, and B. Winkel. A GBT Survey of the HALOGAS Galaxies and Their Environments. I. Revealing the Full Extent of H I around NGC 891, NGC 925, NGC 4414, and NGC 4565. , 865(1):36, Sep 2018.

[3] S. Faridani, F. Bigiel, L. Flöer, J. Kerp, and S. Stanimirović. A new approach for short-spacing correction of radio interferometric datasets. *Astronomische Nachrichten*, 339(1):87–100, Jan 2018.