The GALFA-HI DR1-S data set

J. E. G. Peek, Kevin A. Douglas, Min-Young Lee,

Jana Grcevich, Carl Heiles, Eric Korpela & Jacob Cychosz

1. Basics of the data set

The Galactic Arecibo L-Band Feed Array Survey in Hi First Spring Data Release (GALFA-Hi DR1-S) is a partial map of the 21-cm Galactic sky available to the Arecibo 305m telescope, spanning 80° $< \alpha <$260°, -1° $< \delta <$33°, primarily at b > 0 (see Figure 1). The data were taken with the ALFA instrument from Summer 2005 to Summer 2008 using the GALSPECT spectrometer. These data were not taken as part of a single, monolithic survey, but rather stitched together from individual observing projects as part of the 'jigsaw-puzzle approach' of the GALFA-Hi survey (Goldsmith & GALFA-Consortium 2003). Observing projects included in this map are *Bipolar Flows* (a2011; Josh Peek, PI), *The Turn On GALFA Survey* (a2059c; Mary Putman and Snezana Stanimirovic, PIs), *Disk-Halo Interface*, (a2060; Mary Putman, PI), *A Cold Cloud Boundary* (a2220: Carl Heiles, PI), *HI Halo Clouds in the Galactic Outskirts* (a2221; Leonidas Dedes, PI). The data were taken in 1616 total hours of observing and represent 2.7 TB of raw information. The data were reduced by the authors from December 2008 to May 2009 using facilities at Space Science Laboratory and Arecibo Observatory.

The final maps are separated into standard GALFA-Hi survey cubes. Each cube covers $512' \times 512'$ (8° 32′ on a side), with 1′ pixels. The cubes are in the `RA---CAR DEC--CAR` projection, with the projection center at [180, 0]. This is to say that x-axis is small-circle α, the y-axis is δ, and the projection has orthogonal graticules. The data set comprises 107 × 2 data cubes in a 8° grid in small-circle α and δ. There is therefore a small overlap of 16′ on each edge of each cube with the adjacent cubes. Each survey cube region is covered by two cubes, the wide (`W`) cube and the narrow (`N`) cube. The wide cubes cover the entire observing band (~ -750 km s^{-1} $< v_{LSR} <\sim$ 750 km s^{-1}), but with a degraded resolution of 0.736 km s^{-1} . The narrow cubes cover a reduced band (~ -188km s^{-1} $< v_{LSR} <\sim$ 188km s^{-1}), but with the full resolution of 0.184 km s^{-1} . Note that we have not included the declination slice format in this data release, which would cover the entire band at full resolution. Slices can be produced upon request.

2. Methods of the data reduction

The data were reduced using the methods outlined in Peek & Heiles (2008) (PH08), with the GALFA-Hi Standard Reduction (GSR) pipeline, version 2.5. Since the publication of PH08, some advancements have been made in the GSR pipeline. A despiker code has been implemented that rejects narrow-band radio-frequency interference (RFI) before the data are doppler corrected. The

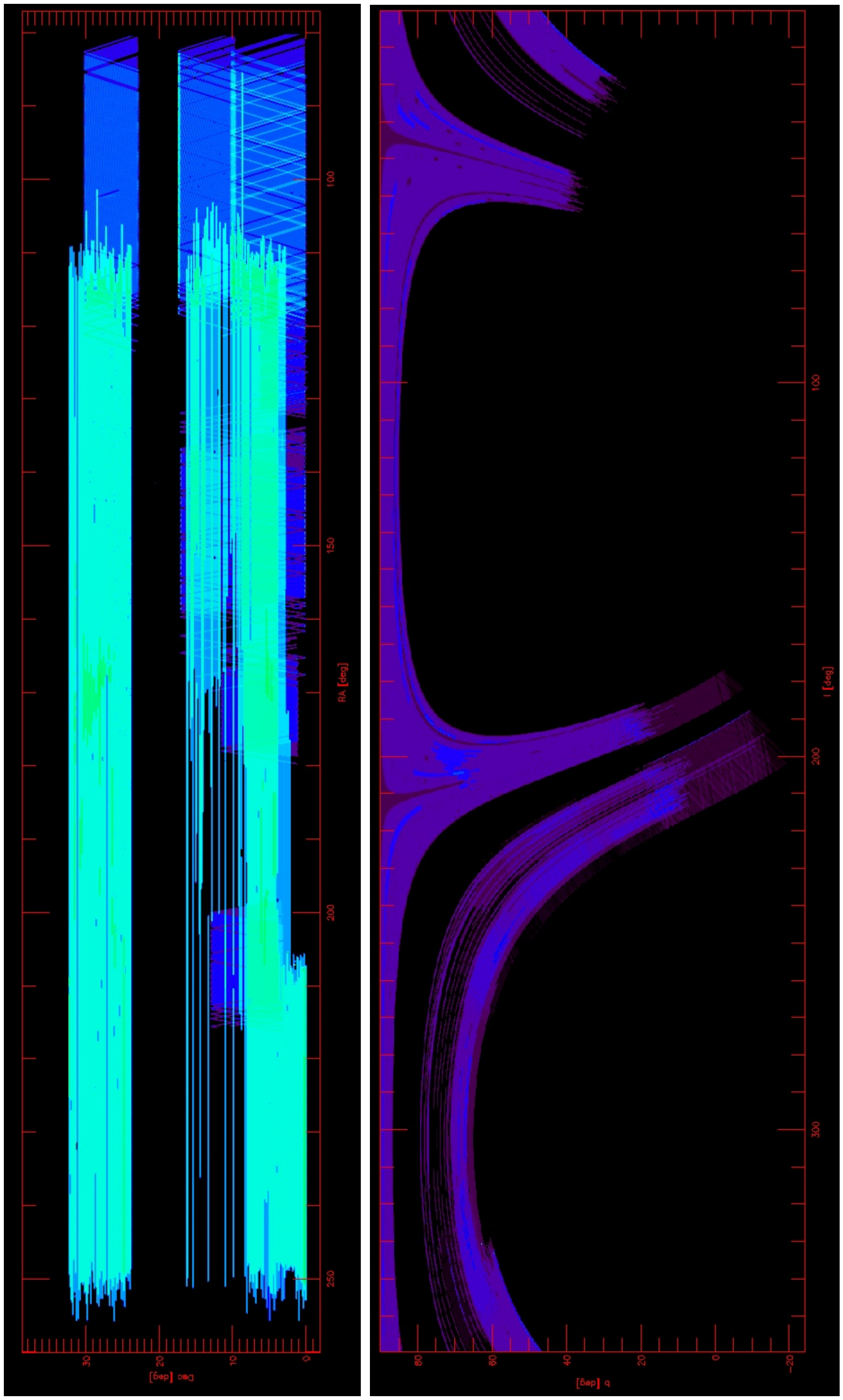

Fig. 1.— The DR1-S data set in equatorial coordinates (left) and Galactic coordinates (right).

baseline ripple reduction software (SPMOD) now has a more sophisticated algorithm for fitting baseline ripple, which reduces the effects of changing declination in basketweave observations and the contamination of HI in the baseline ripple fit. Note that due to the very large number of days, the gains were only corrected to zeroeth order, which is to say an overall offset for each beam and each day of observation. Other changes to the reduction software were mostly for optimization and organization and should not effect the data quality.

3. Known Contaminants

The GALFA-HI DR1-S data set is not without defects. We list the known problems here.

- STANDARD RANDOM NOISE The noise level varies depending upon the integration time per beam, but has a maximum noise level of ∼160 mK over the 0.736 km s^{-1} wide-channel in regions only covered by basketweave observations and minimum noise level < 50 mK over the same channel in regions covered by drift and basketweave observations.

- BASELINE RIPPLE The baseline ripple effects were significantly reduced by the SPMOD routines, but noticeable ripples remain, particularly in data binned in velocity and data that contain little HI signal. The typical amplitude of the striation generated by this ripple is about 30 mK RMS, although it does vary significantly, region-to-region. The effects of the baseline ripple fitting may change somewhat in the Galactic plane ($b \sim 0^\circ$), so take extra precaution when interpreting this data.

- DISTORTED BASELINES Due to a (since remedied) glitch in the software pipeline, some regions at very low Galactic latitude ($|b| < 2^\circ$) have what the authors refer to as a ‘shadow’. This comes from an over-correction of the baseline by a vestigial piece of code in the presence of strong continuum radiation (Galactic synchrotron). The distortion is a parabolic shape with frequency, with amplitude typically a few ×0.1 K. This quirk has not been fixed in the data, but is a simple $T = a \times V_{LSR}^2$ parabolic correction, if others wish to remove this effect in post-processing.

- BEAM GAINS VARIATION Over the course of a given observation, and from observation to observation, beam gains can vary. In this reduction we attempt to fit out the effects of the latter but not the former. It is difficult to measure the residual effects on the final maps, but the scale appears smaller than that of the baseline ripple.

- RECEIVER GLITCHES The receivers will also occasionally ‘glitch’, and spectra will be distorted for many minutes at a time (see Fig 2). Particularly bright glitches are flagged out, but dimmer glitches are not flagged and can contaminate integrated spectra.

- OVERALL MISCALIBRATION We attempt to calibrate the entire data set to the LAB survey (Kalberla et al. 2005) by averaging spectra over large regions and scaling to the LAB survey, which we take to be correct. Overall amplitude calibration to the LAB survey is accurate to about 2%, based on the variation of our calibration across different regions of our maps.

- POSSIBLE VELOCITY OFFSET We velocity calibrate to local standard of rest within the GSR package using local Arecibo meta-data, rather than calibrating to other data sets as we do for amplitude calibration. Therefore, we can determine how consistent our data are with the LAB in the velocity domain by finding the RMS difference between the aggregate GALFA-HI DR1-S spectrum over some region and the aggregate LAB spectrum over the same region as a function of some arbitrary shift δv in the GALFA-HI DR1-S velocity axis. If the velocities are exactly aligned, a minimum in the RMS should occur at δv = 0 km s^{-1}. In fact, it occurs relatively consistently near δv= 0.357 km s^{-1}, as shown in Figure 4. Other comparisons are shown in Table 3. As this time it is not known which data sets represent correct HI velocities, but we note that the LAB data are inconsistent with 4 independently reduced data sets by a similar offset.

Survey 1	Survey 2	Targets	Offsets (1-2), km s^{-1}
LAB	GALFA-HI	9 regions in DR1-S	0.276 – 0.448
LAB	GALFA-HI	b $\sim$ 90° region in DR1-S	0.034
HT2003	GALFA-HI	3C225b, 3C310, 3C315	-0.378, 0.025 and 0.126
LDS	HT2003 (HCRO)	Cas A	0.270
LDS	Heiles (GBT)	Cas A	0.203
LDS	LAB	many regions	0
L-Band PH09	GALFA-HI	Pig 1	0.07

Table 1: Differences in the velocities of the targets from one survey to another. The L-Band PH09 observations have a possible offset of one bin, ± 0.161 km s^{-1}. We do not know what frequency the LAB/LDS data has for the 21 cm line.

- STRAY RADIATION There is also significant contamination to the spectra from stray radiation. In this case, the stray radiation is Galactic HI signal that leaks into the observed spectrum from far beyond the main beam through scattering off of aperture blockage, the ground screen, etc. The effects of this can be seen by subtracting the LAB data for a region from the GALFA-HI data for the same region. We have plotted these residuals in Figure **??** for 9 different regions across DR1-S, which shows that there is typically between 0.1K and 0.2 K of stray radiation across about 50 to 100 km s^{-1} of baseline.

- SIDELOBES Corrections for the first (nearest) sidelobe are implemented in the reduction, which seems to successfully decrerase the effective resolution of the observations. One unintended consequence of this correction is that any errors in the spectra will be spread around somewhat, creating a small 'halo' of radius $\sim 5'$ around said errors. Also note that first sidelobe correction is not in effect at the edge of the observed region or at edge of each data cube.

The overlap of the cubes allows for continuously corrected maps in the observed region. The second (and further) sidelobes still have an effect on the data at a lower level, as evidenced in Mary Putman's work on M33 (Putman et al. 2008). This contamination is usually undetectable without extremely strong HI amplitude gradients, as are seen at the edge of resolved galaxies.

REFERENCES

Goldsmith, P., & GALFA-Consortium. 2003, http://alfa.naic.edu/galactic/docs/

Heiles, C., & Troland, T. H. 2003, ApJS, 145, 329

Kalberla, P. M. W., Burton, W. B., Hartmann, D., Arnal, E. M., Bajaja, E., Morras, R., & Pöppel, W. G. L. 2005, A&A, 440, 775

Peek, J. E. G., & Heiles, C. 2008, ArXiv e-prints

Putman, M. E., Peek, J. E. G., Muratov, A., Gnedin, O., Hsu, W., Douglas, K. A., Heiles, C., Stanimirovic, S., Korpela, E. J., & Gibson, S. J. 2008, ArXiv e-prints

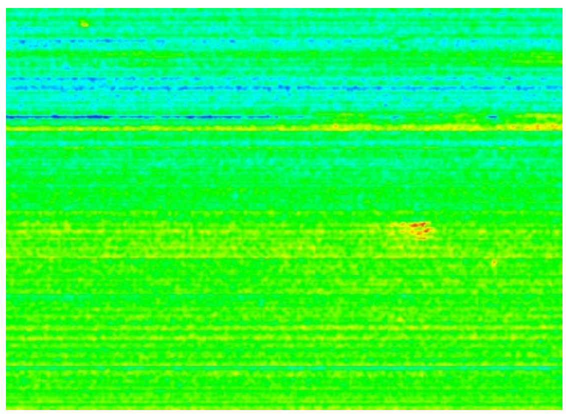

Fig. 2.— An example of a glitch in the DR1-S data set (red part of image), as shown in a 40 km s^{-1} integrated image at 400 km s^{-1} LSR. The striation is caused by baseline ripple.

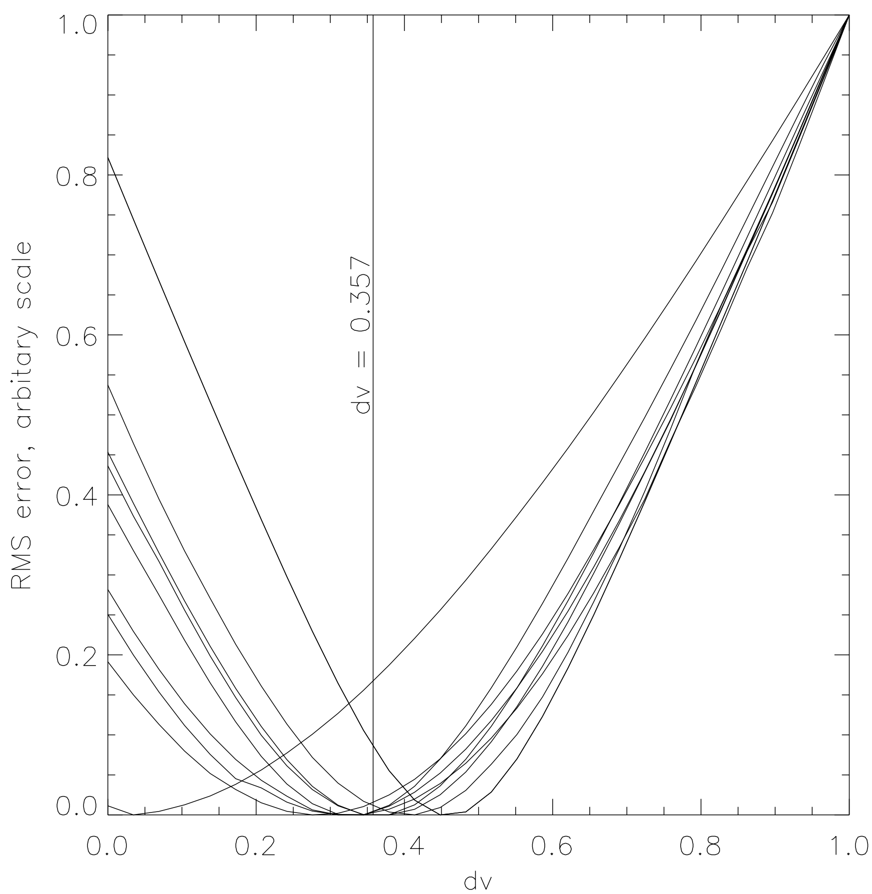

Fig. 3.— The RMS error velocity error, scaled to 0.0–1.0, for 9 cubes distributed over the DR1-S region. 8 of 9 are consistently around the value of 0.357. One cube, at b $\sim 90^{\circ}$, shows little offset, which may represent some spatial component to δv, though we cannot confirm this as the this cube has very little HI flux and thus acts as a poor velocity calibrator.

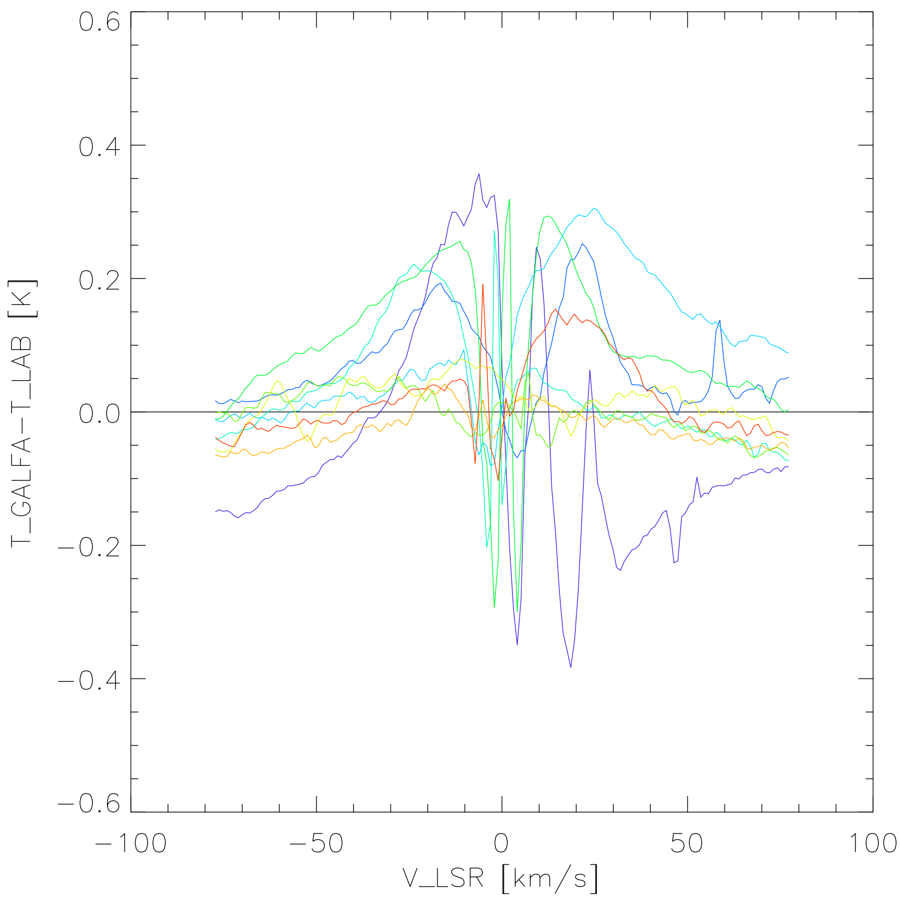

Fig. 4.— The residual spectra when the LAB data is subtracted from the GALFA-HI DR1-S data for 9 different $8^\circ \times 8^\circ$ regions. Note the general feature that these residual spectra rise toward $v_{LSR} = 0$, where the contamination from the Galactic HI will be greatest.