

The Effelsberg-Bonn HI Survey: Milky Way gas[★]

First data release

B. Winkel^{1,2}, J. Kerp¹, L. Flöer¹, P. M. W. Kalberla¹, N. Ben Bekhti¹, R. Keller², and D. Lenz¹

¹ Argelander-Institut für Astronomie (AIfA), Auf dem Hügel 71, 53121 Bonn, Germany

² Max-Planck-Institut für Radioastronomie (MPIfR), Auf dem Hügel 69, 53121 Bonn, Germany
e-mail: bwinkel@mpi.fr.de

Received 20 July 2015 / Accepted 26 October 2015

ABSTRACT

Context. The Effelsberg-Bonn HI Survey (EBHIS) is a new 21-cm survey performed with the 100-m telescope at Effelsberg. It covers the whole northern sky out to a redshift of $z \sim 0.07$ and comprises HI line emission from the Milky Way and the Local Volume.

Aims. We aim to substitute the northern-hemisphere part of the Leiden/Argentine/Bonn Milky Way HI survey (LAB) with this first EBHIS data release, which presents the HI gas in the Milky Way regime.

Methods. The use of a seven-beam *L*-band array made it feasible to perform this all-sky survey with a 100-m class telescope in a reasonable amount of observing time. State-of-the-art fast-Fourier-transform spectrometers provide the necessary data read-out speed, dynamic range, and spectral resolution to apply software radio-frequency interference mitigation. EBHIS is corrected for stray radiation and employs frequency-dependent flux-density calibration and sophisticated baseline-removal techniques to ensure the highest possible data quality.

Results. Detailed analyses of the resulting data products show that EBHIS is not only outperforming LAB in terms of sensitivity and angular resolution, but also matches the intensity-scale of LAB extremely well, allowing EBHIS to be used as a drop-in replacement for LAB. Data products are made available to the public in a variety of forms. Most important, we provide a properly gridded Milky Way HI column density map in HEALPix representation. To maximize the usefulness of EBHIS data, we estimate uncertainties in the HI column density and brightness temperature distributions, accounting for systematic effects.

Key words. surveys – ISM: atoms – techniques: spectroscopic

1. Introduction

In 2009 we initiated a new northern-hemisphere HI survey with the 100-m telescope at Effelsberg, Germany, to succeed the Leiden/Dwingeloo Survey (LDS; Hartmann & Burton 1997) done with the 25-m Dwingeloo telescope. Outstanding in terms of sensitivity and sky coverage compared to any prior endeavor, the LDS was later merged with the Instituto Argentino de Radioastronomía Survey (IAR; Arnal et al. 2000; Bajaja et al. 2005, for $\delta < -27.5$) to form the Leiden/Argentine/Bonn Survey (LAB; Bajaja et al. 1985; Kalberla et al. 2005) – the first full-sky Milky Way HI survey that was corrected for stray radiation (SR). The high-quality SR correction made the LAB survey one of the most important HI data bases to date. The references to the seminal article by Kalberla et al. (2005) reveal the tremendous legacy value of the LAB: one of its most frequent uses is to correct high-energy observations for galactic foreground extinction. However, the coarse angular resolution of LAB may cause significant uncertainties when it is applied to evaluating the intensity attenuation of unresolved sources.

With the development of 21-cm multibeam receivers in the late 1990s, it finally became feasible for the 100-m class single-dish observatories to survey significant portions of the sky with the resulting better angular resolution in reasonable amounts of observing time. The first such project was the very successful HI Parkes All-Sky Survey (HIPASS, Barnes et al. 2001),

which mapped the southern sky out to a radial velocity of $\sim 13\,000\text{ km s}^{-1}$ with the 64-m Parkes telescope. Limitations in the spectrometer meant that the velocity resolution is only 18 km s^{-1} , so too coarse to be useful for many studies of the Milky Way (MW) and its halo. Therefore, a second large-area Parkes HI survey was initiated in 2005: the Galactic All-Sky Survey (GASS; McClure-Griffiths et al. 2009; Kalberla et al. 2010; Kalberla & Haud 2015). It recorded data with narrow bandwidth, hence higher spectral resolution.

Another recent project is the GALFA-HI survey (Peek et al. 2011) currently going on with the Arecibo 300-m dish. The sheer size of the telescope makes GALFA-HI unique in terms of sensitivity. For the Galactic velocity regime, a dedicated backend is delivering spectral resolution up to 0.2 km s^{-1} . The downside is the limited area on the sky that is accessible to the Arecibo telescope.

In late 2008 a seven-beam 1.4-GHz (*L*-band) receiver was installed at the 100-m telescope at Effelsberg. Attached to the seven-feed array are state-of-the-art digital FFT-type spectrometers (FFTS, Stanko et al. 2005; Klein et al. 2012), allowing to observe Galactic and extra-galactic HI simultaneously for the first time with sufficient spectral resolution. The FFTS not only offers great spectral resolution ($\sim 1\text{ km s}^{-1}$) but also allows high-speed ($\sim 1\text{ s}$) storage of spectra. Both properties are beneficial for removing time-dependent radio frequency interference (RFI) from the data during post-processing.

With this new receiving system, we conducted a major HI survey of the sky north of declination $\delta \gtrsim -5^\circ$, called Effelsberg-Bonn HI Survey (EBHIS). To allow for an early first

[★] EBHIS Milky Way HI data is only available at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.128.5) or via <http://cdsarc.u-strasbg.fr/viz-bin/qcat?J/A+A/585/A41>

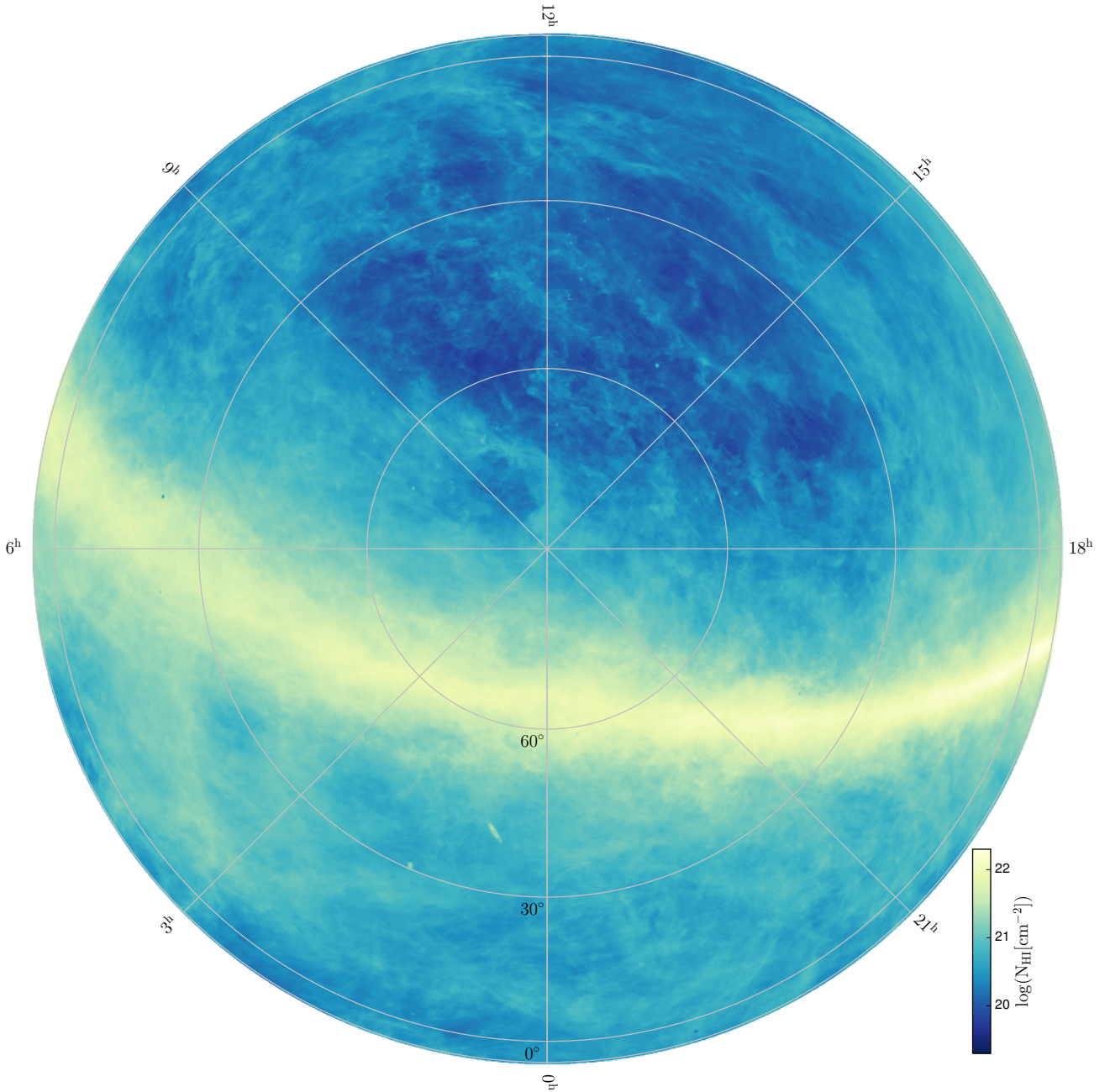


Fig. 1. EBHIS HI column density map, as integrated over the velocity interval $-600 \leq v_{\text{lsr}} \leq 600 \text{ km s}^{-1}$.

data release, survey observations were divided into two runs, each mapping the northern sky completely with a net integration time of about 35 s per beam. In April 2013 we finished first-coverage observations. This data has now been processed and is made available to the scientific community. Meanwhile, measurements for the second coverage are ongoing. Once finished, EBHIS sensitivity will improve by about 30%.

The main science goals of EBHIS have already been discussed in detail in Kerp et al. (2011). Here, we concentrate on assessing the quality of the final data products and on the way these will be made available to the astronomical community.

In Sect. 2 we briefly recapitulate the properties of EBHIS. The data processing pipeline was explained in detail in Winkel et al. (2010); however, several steps were further improved, such as RFI flagging, baseline-fitting, and correction for stray radiation, which we present in Sect. 3. Section 4 comprises the data quality evaluation of EBHIS. We compare EBHIS with the two

recent Galactic HI all-sky surveys, LAB and GASS, and discuss current limitations of the data. In Sect. 5 we study ensemble uncertainties of the observed column density and brightness temperature distributions. Section 6 explains how EBHIS data are made accessible to the astronomical community, most importantly the all-sky HI column density map (Fig. 1). We conclude with a summary and outlook.

2. Survey description

Observations of the first coverage were performed in on-the-fly mapping mode, scanning in right ascension, α , along lines of constant declination, δ . Since the parallactic angle changes during each scan line, the seven-feed array¹ needs to be rotated

¹ One dual-circular central feed and six dual-linear off-axis feeds in a hexagonal layout. The separation of the off-axis feeds from the optical axis is 15'.

accordingly to ensure a regular scan pattern of the offset feeds. Furthermore, the feed array is rotated by an additional $19^\circ.1$ with respect to the scanning direction such that the scan tracks of the individual feeds have equal separations.

The sky north of $\delta > -5^\circ$ was divided into 25 deg^2 sections for a total of 915 individual maps. The polar cap $\delta \gtrsim 85^\circ$ was observed in the Galactic coordinate frame, because near the equatorial pole, the Az–El mount of the 100-m telescope would significantly affect the possible mapping speed in the equatorial system. Hereafter we follow the Effelsberg observatory parlance, where one observation/map is called “scan”, consisting of several “subscans” (the individual scan lines).

Strong observing time constraints were applied for each field:

1. The southeastern quadrant is preferred in azimuth to maximize the angular distance to a mobile-communication broadcast tower in the vicinity of the telescope (at $Az \approx 290^\circ$).
2. We need to avoid large differential changes of parallactic angle during a scan line because the rotation speed of the receiver box at the 100-m dish is limited.
3. The angular distance to the Sun has to be kept as large as possible during daytime observations to avoid “solar ripples” in the spectral baselines (see also Barnes et al. 2005, and Sect. 4.4.1). These are most likely caused by reflections of solar emission off the telescope structure. The exact paths of reflection are difficult to predict, and it is therefore nearly impossible to avoid solar interference by means of observation scheduling. Because of that, several fields had to be observed again for our survey.

The fast-Fourier transform (FFT) spectrometers allow storing a full spectrum every 500 ms, which is an advantage to identifying time-variable RFI. Fast read-out is also beneficial, because the dump time restricts the maximum mapping speed because each independent beam area on the sky needs to be sampled with at least two data points along the scanning direction (Shannon 1949). To allow for a reasonably fast completion of the first sky coverage, we chose a relatively high scan speed of $240''$ per second, pushing the azimuthal engines of the 100-m almost to their limits for sources above an elevation of 60° . We spread the observed bandwidth of 100 MHz over 16384 channels, yielding a spectral resolution of 6.1 kHz. For the Milky Way data release, only the velocity range of $|v_{\text{lsr}}| \leq 600 \text{ km s}^{-1}$ is considered, which is about the frequency interval from 1417.5 MHz to 1423.5 MHz.

After gridding, the median noise level in the data cubes is approximately 90 mK but differs significantly from field to field (see Sect. 5.1). This is mainly a consequence of different elevation angles during the observations causing different levels of stray radiation from the ground and atmosphere. A second important effect is the seasonal changes in ambient temperature (compare Winkel et al. 2010).

In Table 1 we compile important survey parameters of EBHIS in comparison with the most relevant other recent single-dish Milky Way HI surveys LAB, GASS, and GALFA-HI. Apart from the T_{B} noise level, which is usually quoted for full spectral resolution data, we also provide a normalized noise value, $\sigma_{\text{rms}}^{\text{norm}}$, scaled to a velocity resolution of 1 km s^{-1} . The pure noise level, however, only quantifies the instrument’s ability to detect gas, which entirely fills the observing beam. For the many unresolved or point-like objects in the surveys, a better sensitivity proxy is the velocity-integrated flux density, $S_{\text{HI}}^{\text{int}}$. In the table we quote numbers for the 5σ flux-density detection limit, $S_{\text{HI}}^{\text{lim}}$, of an unresolved object with a Gaussian profile having a line width

Table 1. Comparison of basic parameters of recent HI surveys of the Milky Way.

	LAB	GASS	GALFA*	EBHIS	
δ	Full	$\leq 1^\circ$	$-1^\circ \dots 38^\circ$	$\geq -5^\circ$	
ϑ_{fwhm}	36'	16.1	4.0	10.8	
$ v_{\text{lsr}} $	$\leq 460^\dagger$	≤ 470	≤ 750	≤ 600	km s^{-1}
Δv	1.03	0.82	0.18	1.29	km s^{-1}
δv	1.25	1.00	0.18	1.44	km s^{-1}
σ_{rms}	80	57	325	90	mK
$\sigma_{\text{rms}}^{\text{norm}}$	89	57	140	108	mK
Γ_{mb}	0.132	0.649	10.52	1.434	K Jy^{-1}
$N_{\text{HI}}^{\text{lim}}$	3.9	2.5	6.1	4.7	10^{18} cm^{-2}
$S_{\text{HI}}^{\text{lim}}$	16.1	2.1	0.3	1.8	Jy km s^{-1}

Notes. The table quotes the declination range, δ , angular resolution ϑ_{fwhm} , velocity interval, v_{lsr} , channel separation, Δv , spectral resolution, δv , brightness temperature noise level, σ_{rms} , as well as the normalized noise level, $\sigma_{\text{rms}}^{\text{norm}}$, the data would have at a common spectral resolution of 1 km s^{-1} . The main beam sensitivity, Γ_{mb} , is the conversion factor between brightness temperature and flux density, i.e., $T_{\text{B}} = \Gamma_{\text{mb}} S$. The two bottom rows quote theoretical 5σ detection limits (velocity-integrated intensity) of an object with a Gaussian profile of 20 km s^{-1} line width (FWHM). (*) Numbers given are for the shallowest mode. Effective integration time varies across the survey area. Fields observed in comensal mode with AGES (Minchin et al. 2007, $\sigma_{\text{rms}}^{\text{norm}} = 33 \text{ mK}$) and ALFALFA (Giovannelli et al. 2005; Haynes et al. 2011, $\sigma_{\text{rms}}^{\text{norm}} = 60 \text{ mK}$) are much deeper (J. Peek, priv. comm.). (†) Northern (LDS) part only goes out to $v_{\text{lsr}} \leq 400 \text{ km s}^{-1}$.

(FWHM) of $w_{50} = 20 \text{ km s}^{-1}$. Likewise, a limiting column density, $N_{\text{HI}}^{\text{lim}}$, was calculated without assuming point likeness. Both quantities, $N_{\text{HI}}^{\text{lim}}$ and $S_{\text{HI}}^{\text{lim}}$, scale with $\sqrt{w_{50}}$ and $\sigma_{\text{rms}}^{\text{norm}}$.

3. Data reduction revisited

In the following we report on major modifications of the EBHIS data reduction software over the procedures reported in Winkel et al. (2010).

3.1. Improved RFI flagging

3.1.1. Automated flagging algorithm

Flöer et al. (2010) described how to make the best use of the fact that all seven feeds of the receiver are exposed to the same RFI environment. Following their approach improves RFI rejection efficiency significantly. We briefly review the basic RFI flagging workflow. Our automated flagging pipeline is optimized for three distinct types of RFI encountered at the Effelsberg 100-m telescope:

1. Near-constant narrow-band spikes, typically affecting one or two spectral channels;
2. Intermittent broad-band events affecting a hundred up to thousands of spectral channels;
3. Extremely strong RFI caused by the L3 mode of the GPS satellite system.

To detect the first two types of RFI, we use very similar detection strategies. Because we have 14 independent measurements (7 feeds with 2 polarization channels each) of the RFI environment at any given time, we use coincidence flagging to distinguish man-made interference from astronomical signals. Here

we assume that a signal being present in more than one feed at the same time does not have an astronomical origin, because all seven feeds are pointed at different locations on the sky. This assumption is violated if an astronomical source is very extended, as for diffuse HI emission from the Milky Way disk. We therefore also use matched filtering adapted to the typical appearance of RFI in a time-frequency plot. This enables a reliable RFI detection with the exception of the innermost part of the bright Galactic emission, where contributions at all spatial frequencies are present that cannot be reliably suppressed by matched filtering.

To process each individual observation, we first average the two orthogonal polarizations for each feed. Since RFI is often strongly polarized, either by reflection or intrinsically, there are cases where an RFI event is only present in a single polarization channel. Keeping the polarizations separate would then weaken our assumption that all feeds are exposed to the same RFI environment. While reducing the sensitivity gained from requiring signal coincidence, we still gain a factor of $\sqrt{2}$ in sensitivity from the averaging. After averaging, the subscans of each observation are processed independently, allowing multicored CPUs to be exploited for processing multiple subscans in parallel.

To detect narrow-band RFI, we average each subscan in time, yielding seven spectra per subscan. Because narrow-band RFI events are typically confined to fewer than three channels at full spectral resolution, we remove any large scale spectral component by subtracting a median-filtered version of each spectrum. To flag a certain channel as contaminated, we use combinatorial thresholding. Instead of marking a channel as containing RFI once its level exceeds a certain threshold t_1 , we require that a channel exceeds a lower threshold t_N in N feeds simultaneously. Assuming the noise in the data is Gaussian with standard deviation σ , the thresholds for t_N that have the same statistical significance as an individual threshold t_1 are calculated by solving the equation

$$1 - \Phi(t_1) = [1 - \Phi(t_N)]^N \quad (1)$$

where $\Phi(x)$ is the cumulative distribution function of the standard normal distribution. Here we also make use of the fact that all RFI events have positive intensities.

Using this type of combinatorial thresholding, we can impose the coincidence requirement and with that make better use of the available data to optimize our detection efficiency. We choose a lowest required coincidence level of $N = 4$. To flag a certain channel, we compare the data from all feeds to all successive thresholds for $N = 4 \dots 7$ and flag the channel if the criterion is met for any N . An example for the RFI detection quality is provided in Fig. 2. It shows a time–frequency plot (spectrogram) of one subscan of raw data (one feed) in the top panel. The bottom panel shows the same data but with all narrow-band events, detected by the RFI flagger, marked.

To detect broad-band interference, we apply a three-point median filter in the time domain of the data to suppress all persistent signals. We furthermore smooth the data in the spectral domain with a Gaussian filter adapted to the typical extent of the broad-band events. With the prepared data we again perform combinatorial thresholding across the seven feeds.

Since the GPS L3 mode is extremely bright and radiates always at a center frequency of 1381.05 MHz, we simply compare the root mean square (rms) in a 1 MHz window around this frequency to the rms in neighboring frequencies. If the rms differs by a factor of two, we flag a 10-MHz-wide window around this frequency. This removes the brightest part of the interference, but GPS L3 features extended spectral side lobes, which

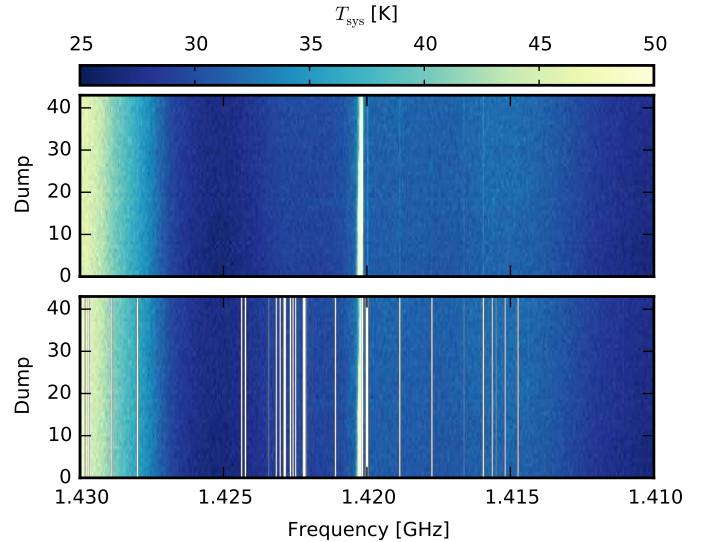


Fig. 2. Narrow-band flagging result. *Top panel:* spectrogram of raw data, *bottom panel:* same, but with flags visualized as vertical lines.

can reduce the fidelity of the data over a large percentage of the 100-MHz band (about 50%).

The data affected by broad-band and GPS interference is excluded from further processing, i.e., flagged. Typically, this does not limit the sensitivity of the Galactic survey, because broad-band interference mostly occurs at the low-frequency end of the 100-MHz-wide observing band. Also, GPS interference usually does not affect the radial velocities $-600 < v_{\text{LSR}} < +600 \text{ km s}^{-1}$. Because narrow-band interference in EBHIS data is numerous and often has near constant intensities, we attempt to subtract the RFI contribution instead of completely flagging an affected channel. This removes most of the narrow-band interference in the Galactic survey with the exception of a few narrow-band signals that mix with the bright HI emission from the Milky Way disk. Here, a reliable subtraction is not possible and the interference is left untouched. From our perspective, this is better than flagging the spectral channels in question completely because leftover narrow-band RFI signals have amplitudes (typically $\lesssim 0.5 \text{ K}$) that are small compared to the bright MW disk emission.

In the final data cubes, the Doppler correction causes frequency shifts of narrow-band RFI signals. In the topocentric frame they are mostly fixed in frequency (a few kHz change during hours), whereas after applying the LSR correction, shifts of up to two spectral channels occur. Therefore, residual narrow-band RFI can appear as “wave-like” patterns when visually inspecting consecutive channel maps in the data cube (see Sect. 4.2).

3.1.2. Manual RFI flagging

The RFI described is present in all observations. Apart from these regular signals, there are additional glitches that are not easily recognized by automated algorithms. To ensure consistently high data quality, each observation was also inspected manually. A simple graphical user interface was developed to compute spectrograms (images of the time-frequency plane). The user can quickly iterate over the different subscans, feeds, and polarization channels to search for RFI and other defects in the data. By mouse-clicking on a spectral dump, an RFI flag (bad

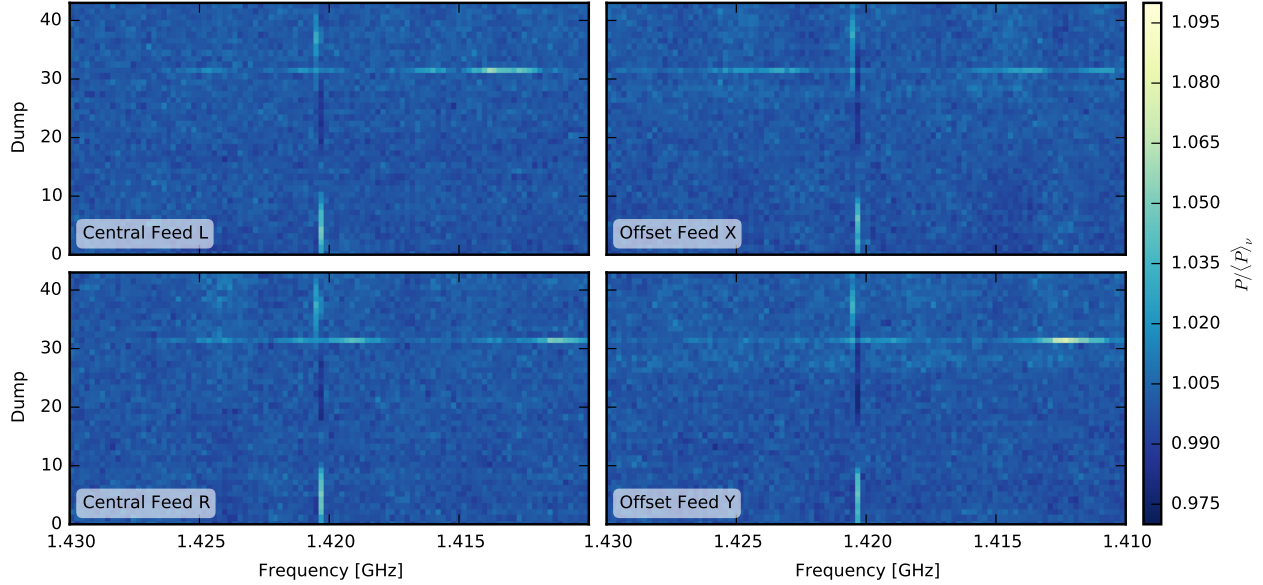


Fig. 3. Broadband RFI event, which is only visible in a single spectral dump (31). The *left panels* show the two polarizations (L and R) of the central feed, and *right panels* contain offset feed data (X and Y). For improved visualization, data is binned in frequency (32-fold, $\Delta\nu = 195$ kHz), and each spectral dump is divided by the median spectrum. The Milky Way HI emission line appears slightly displaced from its rest frequency owing to the LSR Doppler correction, which was not applied here.

spectrum) is generated in a data base. To improve the RFI-to-noise ratio, data is binned in frequency (usually 32-fold, giving a channel width of 195 kHz). To remove the bandpass shape, which would otherwise dominate the spectrogram, each spectral dump is divided by the median spectrum of the current subscan. A side effect is that the Milky Way emission line is only shown as a residual signal. For fields closer to the Galactic disk, where the Milky Way is highly structured, this can make it difficult to distinguish faint broad-band RFI from astronomical features.

In the following we briefly discuss the three most frequent types of broad-band interference that were encountered in EBHIS Milky Way data. Figure 3 shows an example of a wide-band burst that is polarized, varies with frequency, and couples differently into the individual feeds. It lasted less than 500 ms and affected the full 100 MHz band. Another glitch that is extending over the full 100 MHz frequency range is plotted in Fig. 4 (top panel). Here, most of the affected spectrum was more negative than the median spectrum, which is most likely caused by a high-intensity out-of-band RFI event, saturating the front end. Intermodulation products are detected all over the spectrum, for instance, at 1425 MHz. The bottom panel of Fig. 4 depicts two short bursts of events, which are a few 100-kHz wide, superposed on Milky Way emission. In this example, RFI and MW can be distinguished well, but closer to the Galactic plane this becomes more difficult. The time between both events is 12 s – which is a typical delay for Radar pulses. Sometimes up to three consecutive bursts were observed, but never more. We speculate that the 100-m dish may have received reflected radar pulses, such as from airplanes. However, the original emitter was never identified.

3.2. New 2D baseline fitting

Winkel et al. (2010) proposed adaptive Gaussian smoothing to calculate baseline solutions for each individual calibrated spectrum. Even though this approach performed well in most cases, we improved the robustness of the baseline solution by

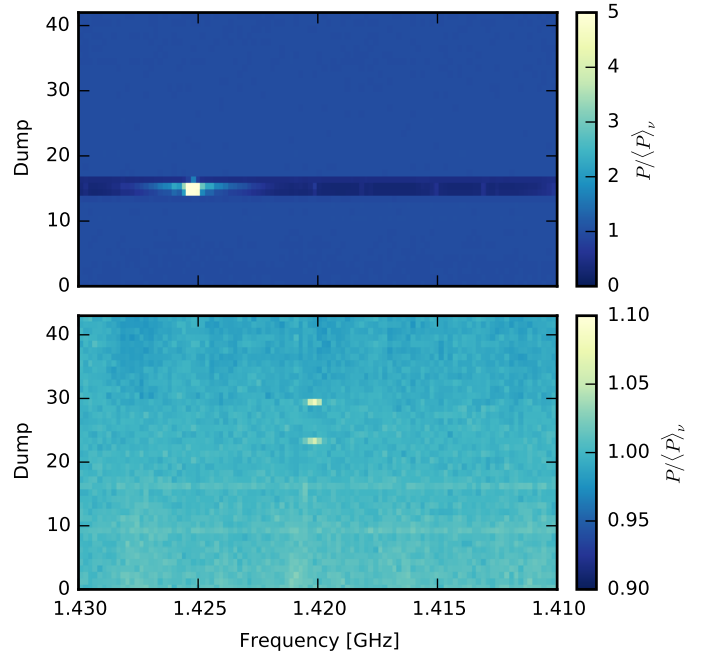


Fig. 4. Examples of RFI. *Top panel:* a very intense out-of-band event temporarily saturates the receiver and causes intermodulation products (e.g., at 1425 MHz). *Bottom panel:* two short bursts of RFI, likely radar pulses, with a width of a few 100 kHz.

extending the fitting procedure to the time-frequency domain. The revised algorithm made use of the fact that the system temperature in the baseline changes only slowly with time.

Further tests revealed, however, that faint sources were sometimes surrounded by negative baseline artifacts. The Gaussian-smoothing baseline algorithm relies on proper flagging or masking of sources. Otherwise, the flux contribution of sources will be smeared into the baseline solution itself. In this context, a source is thought to be anything that superposes the

baseline, such as HI clouds in the Milky Way and its halo or other galaxies (both in continuum and spectral line emission). The baseline algorithm itself tries to mask sources, for example, by iterative flagging of 3σ outliers. However, since baseline-fitting is done on the raw spectral data that have about 500 mK rms, faint sources are not always reliably detected.

It turned out that in these cases, polynomials are better suited for baseline estimation. They can describe underlying fluctuations well, but produce shallower troughs when no mask is provided. Because the Gaussian-smoothing baseline method benefited from extension to the time-frequency plane (t, f), we implemented 2D polynomial fitting, where the baseline, y_b , is described by

$$y_b = \sum_{i,j \geq 0} \alpha_{i,j} f^i t^j. \quad (2)$$

Baselines in EBHIS have a complicated structure, especially along the frequency axis. To avoid unreasonably high polynomial orders, we fit the data on tiles of 1024 spectral channels times the number of dumps per subscan (about 40). To suppress sharp gaps in the baseline solution at the tile edges, we interleave tiles (overlap: 512 channels) and interpolate between each of two adjacent solutions using sigmoid thresholds. Good results are achieved with polynomial orders of ten in spectral and two in temporal direction, while allowing only one cross term, $\alpha_{1,1}$, to describe mild tilts in the 2D baseline.

To improve the baseline solution even further, we initially used the LAB survey to provide a spectral mask around MW emission. This is supplemented with information from the HyperLEDA (Makarov et al. 2014) database² containing extragalactic HI objects and NVSS (Condon et al. 1998) to flag strong continuum sources (≥ 1.5 Jy). Weaker continuum sources are handled well by subtracting the average T_{sys} level from each spectral dump prior to computing the baseline. Finally, a full-resolution datacube of EBHIS itself is used to mask HI emission from the MW and its halo in an iterative scheme.

3.3. Frequency dependency of T_{cal}

Winkel et al. (2010) explain the intensity calibration of EBHIS data incorporating the frequency-dependence of the system temperature, T_{sys} . There, the temperature, T_{cal} , of the noise diode, which is fed into the receiver, was assumed to be constant over frequency. In the meantime, based on the methods proposed in Winkel et al. (2012b), we could derive the frequency dependence of T_{cal} using continuum calibration sources (e.g., 3C 48, NGC 7027). Unfortunately, such a measurement is time-consuming, since each beam must be positioned individually onto the continuum source.

As a faster alternative, we used an absorber – placed around the seven feeds – to obtain a simultaneous and almost RFI-free measurement of all noise diodes. Using the hot absorber alone cannot provide absolute calibration of the T_{cal} spectra. However, it is well suited to assessing the frequency-dependent behavior. The absolute value of T_{cal} at 1.42 GHz can then be measured using observations of the IAU standard position/source S 7 at $v_{\text{lsr}} \approx 0$ km s⁻¹, as explained in Winkel et al. (2010). This procedure was repeated several times during the whole observing campaign and yielded consistent results for the T_{cal} spectra of all 14 channels (7 feeds with 2 polarizations each).

3.4. T_{sys} -based weighting scheme

For EBHIS data, the system temperatures vary between the different feeds and naturally also between the phases where the noise diode is switched on and off. It is possible to achieve a slightly lower final noise level of the data by weighting each spectrum with its associated system temperature (see Winkel et al. 2012b). Since we do not use frequency switching to remove the bandpass from our data, we can reconstruct the full frequency-dependent system temperature during calibration and baseline calculations.

One difficulty remains: the HI emission line itself contributes to the system temperature at the appropriate frequencies. In theory, one had to add each individual HI profile to the underlying continuum T_{sys} spectrum to obtain the proper weights. This would be disadvantageous in practice because the individual 500 ms-line profiles are very noisy; $\sigma_{\text{rms}} \sim 0.5$ K. Such weighting would introduce an additional noise component. Therefore, we decided to compute the average HI line profile per subscan and per beam and use it as the system temperature input for the weighting. Strictly speaking, this attempt is an oversimplification for maps where the T_{sys} level varies significantly during the subscan, for instance, owing to large changes in elevation. However, since all feeds experience the same gradient in system temperature, only a higher order effect is introduced into the weighting, and that can be neglected for our purpose. Introducing the T_{sys} -based weighting scheme improves the final rms level by about 1 to 2%.

3.5. Stray-radiation correction

Galactic emission at 21-cm is seen in all directions, but predominantly from the Galactic plane extending across the whole sky. Antenna side lobes, pointing to such regions of high flux density, will eventually produce artificial signals unrelated to the intensity received by the primary beam. This so-called stray radiation varies with time and season (van Woerden 1962) and is most critical for observations of high-galactic latitude objects with faint HI emission. In extreme cases, stray radiation may provide a larger contribution to a measured spectrum than the true brightness temperature of the observed sky position.

The side-lobe structure of a telescope depends on antenna type and design. For a paraboloidal reflector, the extended side lobes, called stray cones, are mainly caused by the support legs that carry the prime focus cabin and/or secondary mirror (Kalberla et al. 1980a, their Fig. 6). Radiation is also received from regions outside the rim of the reflector, the spillover region. Reflecting surfaces within the aperture can also cause additional side lobes (see Sect. 3.5.4).

The impact of stray radiation can be minimized by reducing the number of scattering surfaces within the telescope aperture. Prime examples are the Bell Labs horn reflector antenna (Stark et al. 1992) and the Green Bank Telescope (GBT; Prestage et al. 2009). For these telescopes the stray radiation is reduced significantly in comparison to a standard paraboloid; i.e., the main beam efficiency is increased from about 70% to 90%. However, the remaining 10% of side-lobe efficiency can still be serious enough that a numerical correction of the observations is unavoidable (Boothroyd et al. 2011).

First attempts by van Woerden (1962) to correct observations of the Dwingeloo telescope for stray radiation were limited by a lack of sufficient computing power. Kalberla et al. (1980a,b) provided the first corrections for the Effelsberg 100-m telescope. Lockman et al. (1986) obtained similar results by bootstrapping

² <http://leda.univ-lyon1.fr/>

from the Bell Labs Survey (Stark et al. 1992), which is affected only a little by stray radiation. Subsequently, Hartmann et al. (1996) and Kalberla et al. (2005) developed a correction for the resurfaced Dwingeloo dish, Higgs & Tapping (2000) for the 26-m Telescope at the Dominion Radio Astrophysical Observatory, Bajaja et al. (2005) for the 30-m telescope at Villa Elisa, Kalberla et al. (2010) for the Parkes 64-m dish, and Boothroyd et al. (2011) for the 100-m GBT.

3.5.1. Basics

The antenna temperature T_A observed by a radio telescope is given for each of the individual receivers by a convolution of the true brightness temperature distribution T on the sky with the beam pattern P of the antenna:

$$T_A(x, y) = \int P(x - x', y - y') T(x', y') dx' dy'. \quad (3)$$

Here we use an approximation in Cartesian coordinates to simplify the expression of the convolution integrals. In general, T_A is time- and frequency-dependent, spherical coordinates have to be used, and the integration needs to take the observable part of the sky (i.e., the horizon) into account, as well as ground reflectivity. For the main beam and all sidelobes, atmospheric attenuation and refraction need to be considered, too.

For the pattern P of the antenna we use the normalization

$$\int P(x, y) dx dy = 1 \quad (4)$$

and split the antenna diagram into the main beam area (MB) and the stray pattern (SP):

$$T_A(x, y) = \int_{\text{MB}} P(x - x', y - y') T(x', y') dx' dy' + \int_{\text{SP}} P(x - x', y - y') T(x', y') dx' dy'. \quad (5)$$

By defining the main beam efficiency,

$$\eta_{\text{MB}} \equiv \int_{\text{MB}} P(x, y) dx dy, \quad (6)$$

of the telescope, Eq. (5) can be rewritten as

$$T_B(x, y) = \frac{T_A(x, y)}{\eta_{\text{MB}}} - \frac{1}{\eta_{\text{MB}}} \int_{\text{SP}} P(x - x', y - y') T(x', y') dx' dy'. \quad (7)$$

To determine the main beam-averaged brightness temperature, T_B , we need to know the antenna pattern, P , with sufficient accuracy, but also the true brightness temperature T on the sky, which a priori is unknown. Apparently we need an approximation for T .

3.5.2. How to solve Eq. (7)

There are two unknowns in Eq. (7), T_B and T . Furthermore, the solution depends on the choice of the main beam area MB. We first consider conditions for a solution of this equation. Replacing T_B on the lefthand side with T leads to a Fredholm equation of the second kind, which can be solved unambiguously for $\eta_{\text{MB}} > 0.5$ (Kalberla et al. 1980a). State-of-the-art telescopes have main beam efficiencies $\eta_{\text{MB}} \gtrsim 0.7$, defined for the region

within the first minimum of the antenna diagram. The existence of a solution of Eq. (7) is therefore warranted. A standard procedure for finding the solution is to use successive approximations, as proposed by Bracewell & Roberts (1954).

Knowing these basic limitations, we can develop a strategy for deriving correct brightness temperatures. Most important is to find ways to detect stray radiation effects in the observations and, likewise, to separate regions in the antenna diagram that cause contributions. As far as possible, we measure parameters of the antenna diagram P , next we model near-side-lobe structures (Sect. 3.5.3). For side lobes far from the main beam, we study details of the telescope construction and employ ray-tracing to find the critical regions in the pattern (Sect. 3.5.4).

In most cases it is impossible to determine accurate absolute side-lobe levels from ray tracing alone, but by varying individual free parameters (intensities for 69 side-lobe structures in the final version of the far side-lobe model), we get satisfactory solutions. A simple but powerful test for identifying a critical region in the antenna diagram is to correct the observations with an unreasonably high side-lobe level, which then causes overcorrections. The resulting spurious negative features in corrected profiles and data cubes can be easily spotted by eye.

Characteristic of stray radiation errors is the variability of spurious line components that also led to the detection of stray radiation in the first place (van Woerden 1962; Kalberla et al. 1980b). We make use of this effect. An iterative determination of side-lobe levels requires a sufficiently large sample. We use EBHIS for declinations $0^\circ < \delta < 60^\circ$ and velocities $-100 < v_{\text{l sr}} < +100 \text{ km s}^{-1}$.

Coherent regions in $(l, b, v_{\text{l sr}})$ are selected that contain little or no HI emission but are sensitive to stray radiation effects. First we use LAB data to verify the absence of significant emission features. Next we determine peak deviations and independently the rms scatter over larger emission-free $(l, b, v_{\text{l sr}})$ regions with the aim of minimizing these errors. Our analysis is hampered by the fact that we only have a single sky coverage, so it is not possible to compare different EBHIS observations at the same position. Stray radiation problems are, however, recognizable as discontinuities between individual fields that cause a patchy or blocky structure (compare Fig. 11, top panel). Our strategy is to improve the SR corrections by successive approximations, modifying antenna parameters but also the correction algorithm itself (Sect. 3.6), as well as the estimate for the brightness temperature distribution T in Eq. (7).

3.5.3. The multibeam antenna pattern: near side lobes

To solve Eq. (7), it is necessary to work out the complete antenna response, P , for each of the seven beams of the Effelsberg multifeed system. For practical reasons we distinguish between near side lobes around the main beam within a distance of 4° and far side lobes farther out. Both regions cover approximately half of the antenna's solid angle outside the main beam.

Using strong radio continuum sources such as Cas A, it is hardly possible to measure accurate side-lobe structures below -40 dB . Such observations are typically restricted to distances $\lesssim 1 \text{ deg}$ from the main beam but cover only $\lesssim 50\%$ of the near side-lobe beam's solid angle required in Eq. (7). We therefore model the side lobes out to radial distance of 4 deg from the main beam in a similar way to what was previously done for GASS.

The far-field pattern of an antenna is approximated by the auto-correlation of the aperture plane distribution. We used the measured feed-horn response pattern (Keller et al. 2006) and

the telescope geometry, including shadowing caused by the focus cabin and the feed support legs to derive the complex aperture distribution function for each feed. The pattern was then calculated by Fourier transformation in a similar way to Baars (2007, their Sect. 2.2).

For the seven beams, we distinguish between the central feed and the six surrounding feeds in a hexagonal layout. The offset beams have a radial offset of 17.3 cm from the optical axis. As mentioned in Sect. 2, the seven-beam receiver is rotated during observations to account for the changing parallactic angle that would otherwise distort the resulting scan pattern in the equatorial coordinate frame. Unfortunately, this means that the feed horns rotate relative to the focus cabin's support legs. This causes the already complex aperture distribution function to change with time. For the sake of processing speed, we accounted for this rotation only in steps of 5° . We also tested a less accurate stepping of 10° , which was previously used to correct data from GASS (Kalberla et al. 2010), but differences are barely noticeable. We conclude that the approximation of the beam rotation within $\pm 2.5^\circ$ (EBHIS) or $\pm 5^\circ$ (GASS) is accurate enough to solve Eq. (7) for the near side lobes without noticeable uncertainties. Utilizing the four-fold symmetry of the antenna aperture, in total 18 antenna diagrams for the offset feeds need to be provided during the solution of Eq. (7). As an example, we plot in Fig. 5 the diagram for the central feed and an offset feed showing the coma lobe at a position angle of 60° .

For the correction algorithm, we averaged the previously modeled side-lobe intensities on a fixed grid containing 2160 cells in cylindrical coordinates for the inner 4° of the antenna diagrams. The cells have an azimuthal extent of 5° and a radial extent of 0.125° , covering the radial range of 0.25° (first minimum) up to 4° .

3.5.4. The far side lobes

The far side lobes are determined by details of the telescope structure. It is difficult to measure these side lobes (e.g., Higgs 1967; Hartsuijker et al. 1972), so we used ray-tracing to model them. While in the near side-lobe range, details of each of the individual antenna diagrams need to be considered, and it is sufficient to use a single common far side-lobe diagram for all of the receivers. For the subsequent discussion of far-side-lobe effects, we consider the telescope as a transmitting system.

The most important side lobes that are far away from the main beam of the 100-m telescope, were determined by Kalberla et al. (1980a, see their Fig. 3). These are the stray cones, caused by reflection of plane waves from the primary mirror at the feed support legs. The second most important structure is the spillover at the edge of the main reflector. Its irradiation level depends on the edge taper of the primary feeds.

In addition, we considered spherical waves originating in the feeds and determined their reflections at the support legs. Figure 6 shows these structures. The eightfold symmetry is caused by the four legs. Each leg consists of four main tubes; two of them, separated by 1.98 m, are visible from the feed. These side lobes have an interesting structure, but it turned out that they are rather unimportant.

The Effelsberg 100-m telescope has a Gregorian secondary focus with several feed systems in an apex cabin. The seven-feed system is situated in primary focus, and the roof of the apex cabin is closed during EBHIS observations. This roof causes reflections that are offset from the main axis by 40° . The side lobes are easily determined from the geometry of the apex roof. However, that these reflections are oriented in the east-west

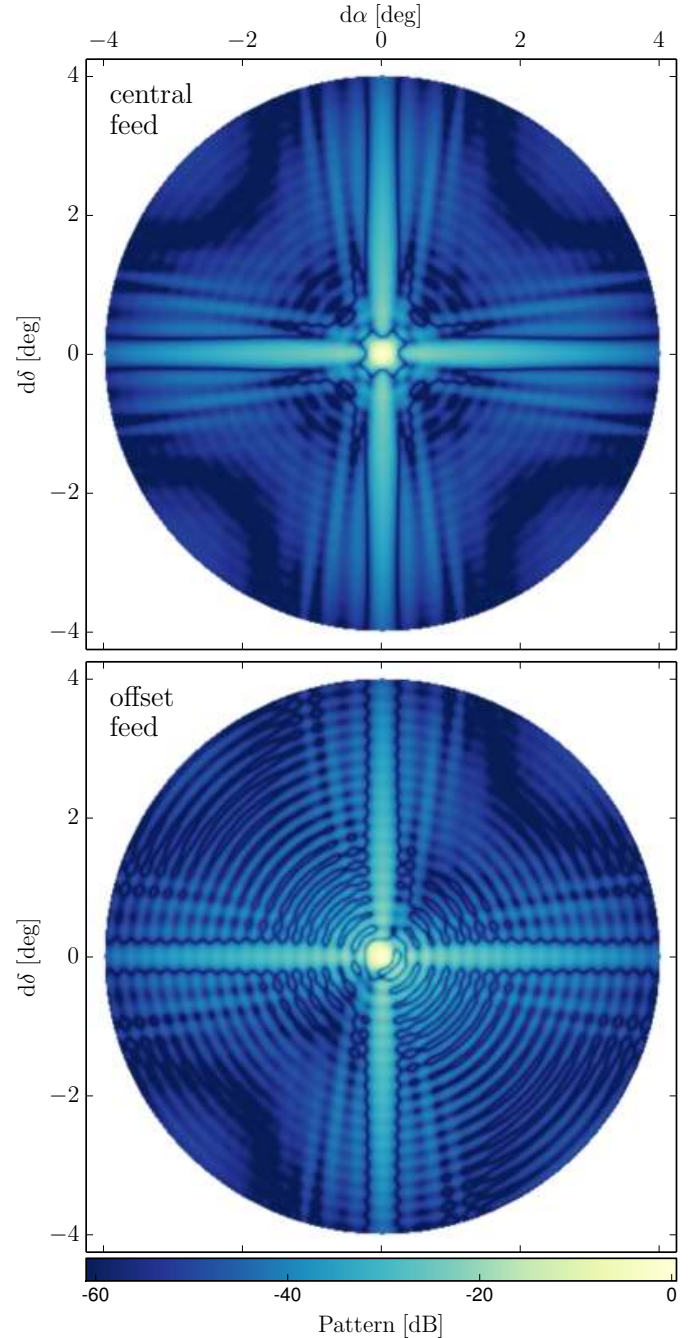


Fig. 5. Synthetic antenna patterns within a radius of 4° for the central beam (*top*) and an offset beam (*bottom*) with a coma lobe at a position angle of 60° .

direction makes the situation very uncomfortable since it can produce multiple reflections. A part of the rays being reflected off the apex roof can hit feed support legs located east or west of the roof. This causes secondary stray cones. Only a fraction of these secondary stray cones are reflected onto the sky. The other part is reflected to the main mirror, causing caustics. A minor fraction of the secondary stray cones can hit the fence at the rim of the reflector before escaping. These features were determined by ray-tracing. Figure 7 shows the resulting pattern. We disregarded further reflections caused by support legs to the north and south of the focus cabin.

The apex lobes discussed here differ significantly from side-lobe properties of the original telescope design. Initially the roof

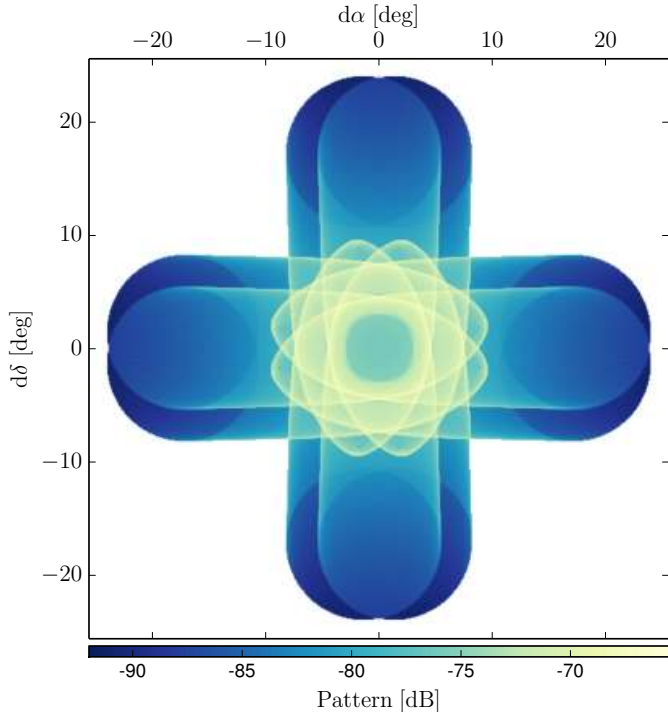


Fig. 6. Side lobes caused by reflections of spherical waves from the feed at the support legs of the 100-m (intensities relative to main beam).

of the apex cabin was designed so that reflections occurred in four triangular lobes in between the support legs (see Fig. 3 of Kalberla et al. 1980a). The new construction allows fast switching between primary and secondary focus but causes extremely complicated side-lobe structures far off the main beam that can only partly be modeled. We also do not account for minor construction details like stairs, cable ducts, and cross ties of the prime-focus support legs.

The side-lobe levels from the spillover lobes were estimated from the edge taper of the receiver feed. For the stray cones we used previous results from Kalberla et al. (1980a), but in the new software the radial side-lobe levels of the cones are modeled by a Gaussian (2° FWHM). All far side-lobe components were adjusted individually in a way similar to what is described by Kalberla et al. (2005), when we were searching for a consistent solution to Eq. (7).

3.6. The correction algorithm

Our correction algorithm is based on Kalberla et al. (1980a) and was extended later for multibeam systems (Kalberla et al. 2010). Some further improvements were made, the most important ones concerning the all-sky brightness temperature distribution, T_B , that is needed for the deconvolution according to Eq. (7). As previously mentioned, the basic strategy was to utilize the LAB data set to calculate an initial guess of the SR correction, i.e. to solve the side-lobe-pattern integral in Eq. (7). After optimizing free parameters in the antenna pattern for the LAB-based SR correction, we switched over to use EBHIS itself and GASS data for the input sky. Again, several iterations were processed to find the optimal solution.

For a deconvolution on the sphere, an equal pixel-area representation of T_B is desirable. Here we used the Hierarchical Equal Area isoLatitude Pixelisation (HEALPix Górski et al. 2005) scheme. HEALPix is a versatile structure for the pixelization

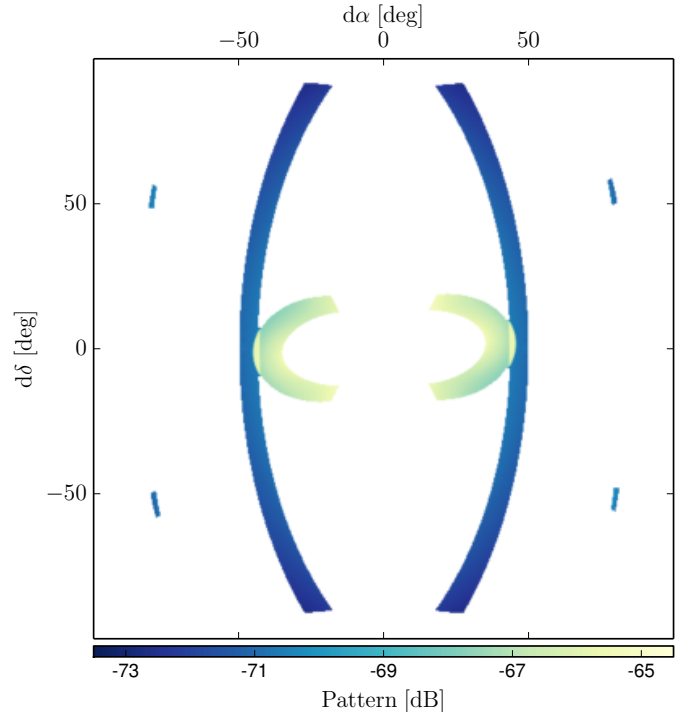


Fig. 7. Side lobes caused by reflections from the roof of the secondary focus cabin onto the feed support legs (intensities relative to main beam). The large vertical feature is a secondary stray cone. The inner lobes are caustics caused after reflections from two tubes of the feed support legs back to the primary mirror. The four spot-like side lobes are caused by the fence around the rim of the paraboloid.

of data on the sphere. Most important for our needs is that this scheme allows the spatial resolution of the T_B spectra to be easily matched to the needs in accuracy during the correction. A single parameter, N_{side} , is sufficient to define the resolution of T_B (Górski et al. 2005, Table 1). Here $N_{\text{side}} = 256$ corresponds to a size of $13.7'$ per pixel, which in most cases is adequate for near side-lobe correction. For $N_{\text{side}} = 64$, a pixel size of $55.0'$ is given, which would be more suitable for far side-lobe corrections.

Side-lobe correction computing speed depends critically on the number of the processed T_B spectra. High-resolution single-dish data for the correction of the far side lobes are only needed in regions with significant fluctuation in brightness temperature. The HEALPix scheme allows a flexible scaling, and we have chosen $N_{\text{side}} = 128$ for regions with column densities $N_{\text{HI}} > 5 \times 10^{21} \text{ cm}^{-2}$, $N_{\text{side}} = 64$ for $10^{21} < N_{\text{HI}} < 5 \times 10^{21} \text{ cm}^{-2}$, and $N_{\text{side}} = 32$ for $N_{\text{HI}} < 10^{21} \text{ cm}^{-2}$ (see Fig. 8).

A near side-lobe correction demands a high-resolution approximation of the true-sky brightness temperature T_B according to Eq. (7). We initially used the LAB survey, resampled to $N_{\text{side}} = 256$ and later to $N_{\text{side}} = 512$. After this preliminary SR correction based on low-resolution data, we employed the SR-corrected EBHIS dataset itself ($N_{\text{side}} = 1024$, i.e., full-angular resolution, $\theta_{\text{pix}} = 3.4'$). GASS data were supplemented on the southern sky on the same HEALPix grid. The last step was repeated several times in conjunction with a recomputation of the 2D baseline models (Sect. 3.2). Thus we increased the accuracy step-by-step for the near side-lobe correction. In addition we improved the accuracy of the deconvolving kernel function: for each beam from initially 396 to finally 2160 samples of the side-lobe structure. The beam rotation was traced initially within ± 5 deg, finally accurate to ± 2.5 deg.

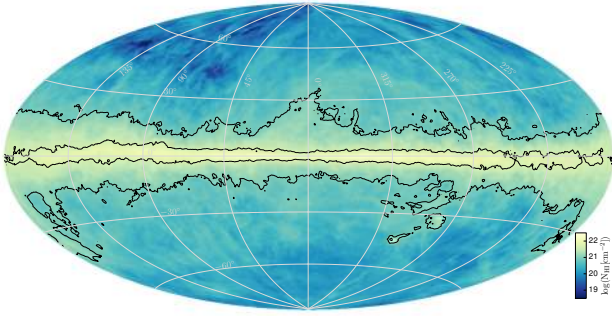


Fig. 8. Column density distribution from the LAB survey. The isophotes indicate $N_{\text{HI}} = 10^{21} \text{ cm}^{-2}$ and $N_{\text{HI}} = 5 \times 10^{21} \text{ cm}^{-2}$, limiting ranges with different HEALPix resolutions that were used for far-side-lobe stray radiation correction.

All these improvements in the correction procedure led to a more precise and efficient algorithm but eventually did not improve the quality of the solution, as desired. Because we cannot model all parameters of the antenna diagram with sufficient detail, residual SR features appear in the data. Unaccounted for are substructures of the feed support legs such as stairways and cable channels. In calculating caustics (Fig. 7), we only took three scattering surfaces into account and disregarded further scattering at the feed support legs. Also some substructures of the roof of the apex cabin were ignored. We did not attempt to model reflections from the subreflector mount.

In Appendix A.1 we show maps of the near-side-lobe, far-side-lobe, and total SR correction (in HI column densities, Fig. A.2), which was applied to the EBHIS data. To allow comparison, the three figures share the same colorbar scale as Fig. 1. It is remarkable that the far-side-lobe correction in several low-column density regions at higher Galactic latitudes is greater than the reconstructed column densities in these fields, which underlines the importance of SR correction.

4. Data quality

In this section we assess the data quality of EBHIS in detail. As with many other large-scale single dish surveys, there are still residual artifacts visible in the data. One important reason is the sheer amount of data that makes automatic processing necessary. In Winkel et al. (2010) and Sect. 3, we presented how many problems with the raw data had to be dealt with. In general the developed data reduction software proved to be very robust. In the following we comprehensively discuss the cases where our data processing pipeline reaches its limits. This is not because we consider the final data products lacking in quality compared to, for instance, GASS and LAB, but instead we want the potential users of EBHIS to be able to work efficiently and be well-informed with the data. Furthermore, we thoroughly cross-checked the EBHIS intensity calibration against LAB and GASS.

4.1. Calibration: consistency check vs. LAB and GASS

To have an independent consistency check of the overall brightness temperature calibration, baseline solution, and residual SR, we compared the EBHIS data with the LAB survey. To avoid distortion of the results by gridding effects and angular undersampling of the sky in LAB, we took the original pointing positions of LAB and computed a weighted average of EBHIS dumps surrounding the associated positions.

In Fig. 9 (left panel) we show the comparison of total column densities between both surveys. The relation between the two is almost perfectly a one-to-one correlation, based on a linear model fit to the data:

$$N_{\text{HI}}^{\text{EBHIS}} = 1.0023(4) \cdot N_{\text{HI}}^{\text{LAB}} - 0.3(3) \times 10^{18} \text{ cm}^{-2}. \quad (8)$$

To account for both errors in x and y , orthogonal distance regression (Boggs & Rogers 1990) is used (provided by the SciPy module odr; Jones et al. 2001). Proper statistical column density errors, used as weights for the regression, are found using error propagation of the frequency-dependent rms accumulated in the column density integral.

Likewise, the center panel of Fig. 9 displays the relationship of brightness temperatures between EBHIS and LAB. Again, a one-to-one correlation is found:

$$T_{\text{B}}^{\text{EBHIS}} = 1.0000(6) \cdot T_{\text{B}}^{\text{LAB}} + 12.2(5) \text{ mK}. \quad (9)$$

By computing the intensity-weighted velocity (Moment-1) for each spectrum, it is possible to test the correctness of the frequency/velocity scale (see Fig. 9, right panel). The median velocity mismatch between EBHIS and LAB is 0.07 km s^{-1} . This indicates a well-behaved relationship. For an insignificant number of positions, we observe outliers. We attribute this to residual RFI and other artifacts in either of the data sets, which can have a strong impact on the Moment-1 calculation. We minimize the potential impact of artifacts in the data on the Moment-1 calculation by applying a mask with a threshold of 2 K.

While the two surveys agree extremely well on average, it has to be noted that the brightness-temperature scatter about the linear relation is relatively high. We attribute this to the LAB survey because a comparison of EBHIS with GASS reveals much lower scatter; see Appendix A.2, where we provide comparison plots for EBHIS, GASS, and LAB computed for the common data slice ($-4.5^\circ \leq \delta \leq -0.2^\circ$). Unfortunately, we could not find the cause of this enhanced scatter. One well-known problem of the LAB survey is that it is not fully sampled in the spatial domain. But since we compare spectra on the original pointing positions, the angular sampling is irrelevant here. For completeness, we also show similar comparison plots for GASS vs. LAB on the southern hemisphere in Appendix A.2.

Another important finding, which we made during our tests, is that EBHIS and GASS (second data release), as well as LDS (the northern part of LAB) and GASS, showed a discrepancy in calibration by about 4%. On the other hand, the IAR contribution to LAB ($\delta \leq -30^\circ$) matches the GASS flux-density scale, which means that there is an inherent inconsistency in the LAB data between the northern and southern parts. This finding led to a revised intensity calibration scale for the third GASS data release (Kalberla & Haud 2015). The rescaled GASS is very consistent with EBHIS and LDS in terms of column densities and brightness temperatures (see Figs. A.3 and A.4). As a consequence, however, the IAR part of LAB became inconsistent with GASS (Fig. A.5).

4.2. RFI

According to Sect. 3.1.1, the automatic RFI flagger is able to find even faint narrow-band interference. As described earlier, we attempt to subtract near-constant-amplitude narrow-band events from the data to minimize loss of samples caused by the flagging. This subtraction is done subscan-wise per receiver feed and polarization channel. For each flagged channel in the average spectrum, we calculated the difference to the mean value of adjacent

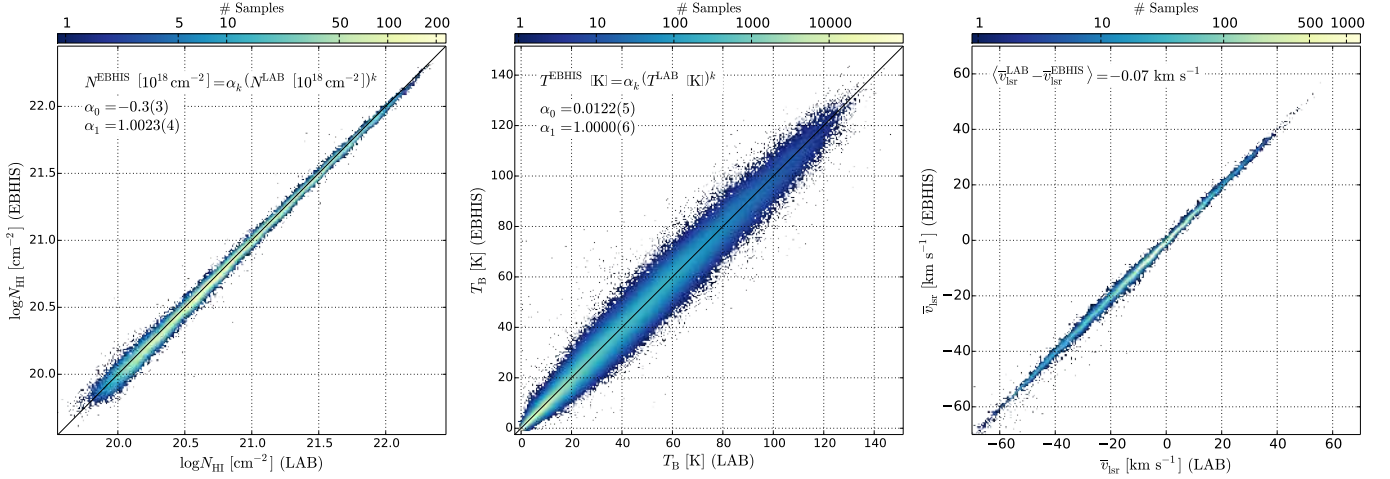


Fig. 9. Comparison of the LAB survey and EBHIS. *Left panel:* column densities; *center panel:* brightness temperatures; and *right panel:* intensity-weighted velocities (Moment-1). Black solid lines represent the best-fit relation (see text).

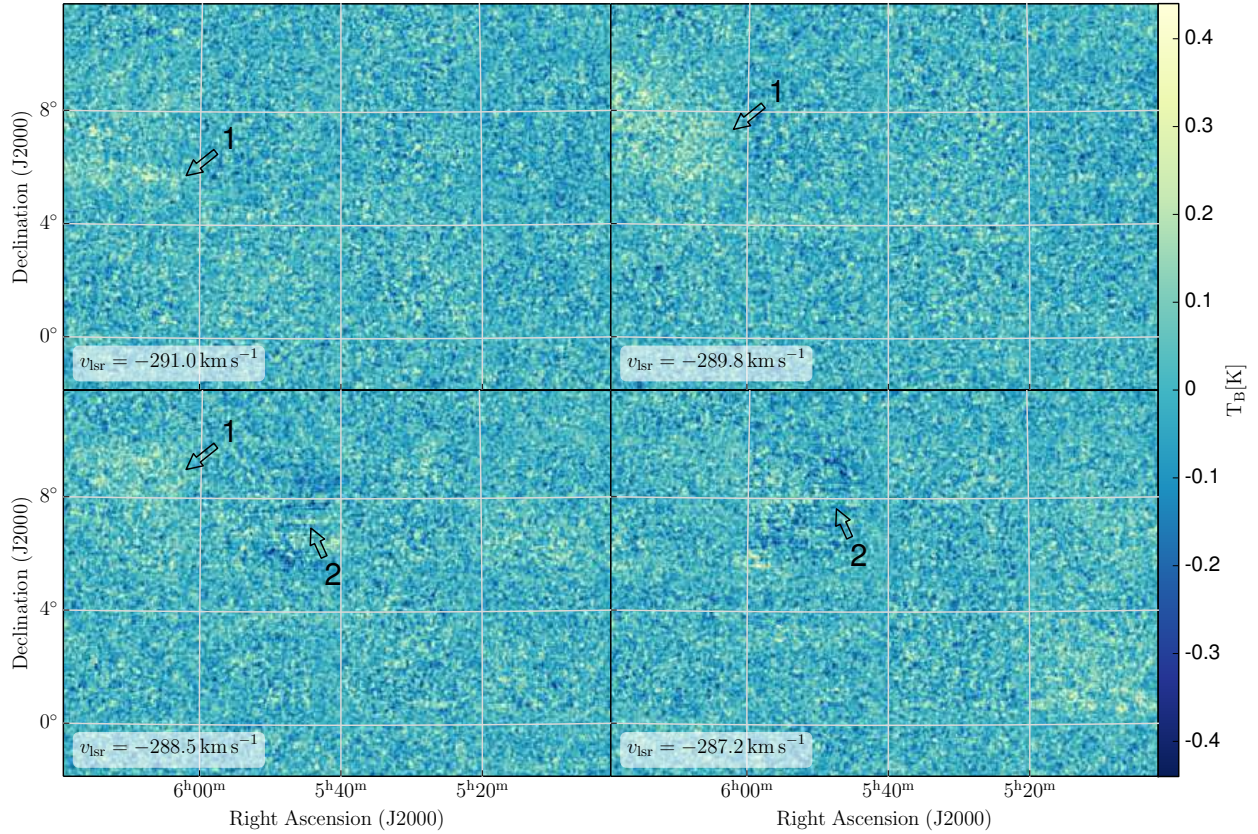


Fig. 10. Example of residual artifacts in the data caused by imperfect subtraction of near-constant-amplitude narrowband RFI. Arrows (1) and (2) denote two different features that appear as “wave-like” patterns. The average residual intensities of both events are about 75 mK and -45 mK , respectively, which is below the 1σ noise level of EBHIS.

channels and corrected for the offset. To distinguish whether an RFI event has constant intensity, we measured the rms of the time series in the spectral channel relative to interference-free regions. Both procedures are, of course, affected by noise. In case of inaccurate RFI intensity estimates or wrong classification, residual artifacts will be visible in the data cubes. The LSR Doppler shift correction, which is subsequently applied, can shift these artifacts slightly along the spectral axis, with the size of the shift being a function of spatial position and observing time.

Figure 10 shows the visual signature of the residual RFI that appears as “wave-like” structure along the frequency axis. We marked two distinct features in four consecutive spectral planes with arrows to emphasize the change in position. For visualization purposes we show two strong artifacts. For most of the residual RFI, an experienced astronomer might not even recognize the features in a single channel map, but when sliding through the data cube, the trained human brain can identify such faint changes.

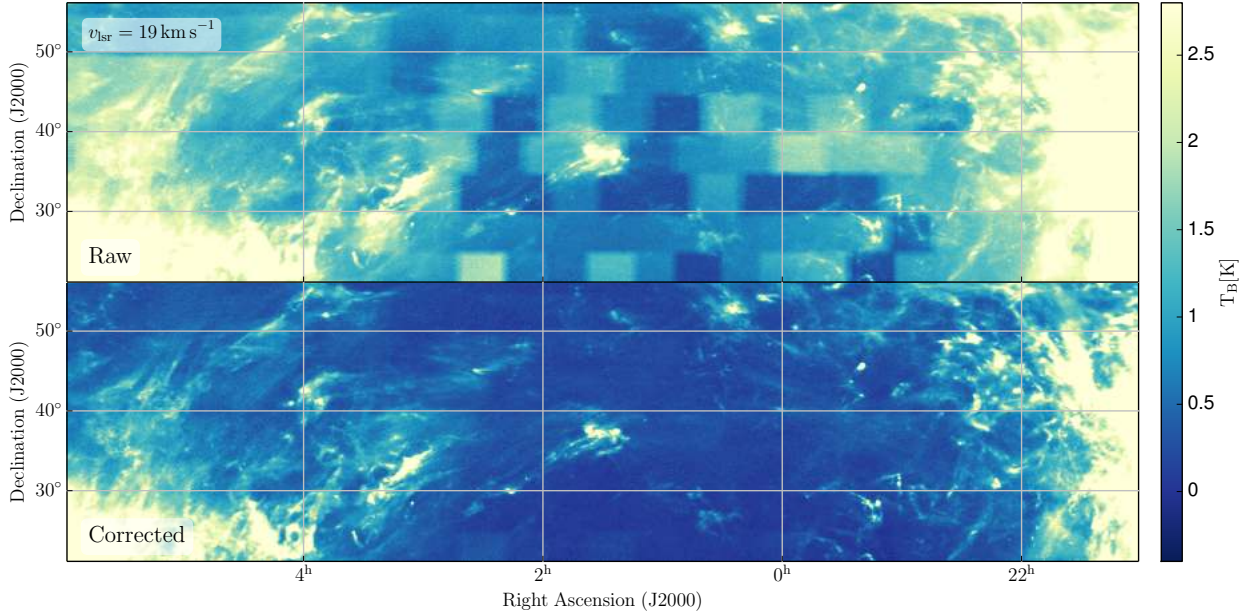


Fig. 11. Example for a region that is significantly affected by stray radiation. *Top panel:* brightness temperature distribution without stray-radiation correction applied. *Bottom panel:* same as top panel after correction for stray radiation. The blocky appearance in the top panel is caused by the $5^\circ \times 5^\circ$ observing grid.

4.3. Stray radiation

The quality of the SR correction can in principle be assessed by repeated observations of each field and calculation of the scatter between the measurements (Kalberla et al. 1980a). To apply this method, we simply lack data, because each field was observed only once. It is possible, however, to use the small overlaps between the individual fields, which were usually observed on different dates or even seasons. Still, only two realizations of the SR are recorded per overlap, meaning that this method can only be used to measure the average scatter of the ensemble of all fields. Likewise, comparison to other datasets, such as the LAB survey or GASS, can be used to estimate deviations. However, the difference between SR correction and baseline fitting errors cannot be worked out sufficiently with such an all-sky statistical attempt, making this approach not well-suited to our SR-correction optimization. For an analysis of the ensemble distributions, we refer to Sect. 5 where the overall uncertainties of EBHIS data are studied in more detail.

Nevertheless, for the SR correction, a quality measure was needed during the iterative manual optimization. A more local approach was chosen, with visual inspection of larger regions with relatively weak line emission, as in the field shown in Fig. 11. Here, residual far side-lobe SR errors manifest as patches associated with the individual observed fields (see Fig. 11 top panel). The amplitudes of these patches is then minimized during the optimization. It is helpful to look at the largest remaining amplitudes of the SR patches. Here, we find errors of up to 200 mK, which is only twice as large as the noise level, σ_{rms} , of EBHIS data at full resolution. Furthermore, the nominal noise level of large emission-line free regions is used as a measure of correction quality. As SR artifacts are predominant close to MW disk velocities, the number of such regions is limited, though.

By looking at spectral-channel maps, it is easier to differentiate between SR errors and baseline effects, because the typical velocity width of SR errors is smaller than that of baseline errors, although there is a regime of a few km s^{-1} that could be

explained by both. For the visual inspection, we mostly used data that is averaged over four velocity channels, reducing the system noise by a factor of about two.

A thorough analysis of EBHIS SR correction quality is also done in Martin et al. (2015) in the context of their GHIGLS project (HI mapping at intermediate Galactic latitude using the Green Bank Telescope). In total 15 selected relatively low-column-density fields were observed. For three of the targets (NEP, NCPLEB, and DRACO; see Martin et al. 2015, for definition), we provided EBHIS prerelease data cubes for comparison. The best-fit relation between EBHIS and GHIGLS flux-density scale is reported as

$$T_{\text{B}}^{\text{EBHIS}} = 1.009(11) \cdot T_{\text{B}}^{\text{GHIGLS}}. \quad (10)$$

Using this scale factor, individual and averaged spectra (per field) are compared. In contrast to the GBT data, EBHIS can have substantial SR contribution even at high-velocity-cloud velocities owing to the different antenna type. For the SR-corrected spectra, however, both data sets agree very well. We refer the interested reader to Martin et al. (2015) for more details.

4.4. Baselines

As discussed in Sect. 3.2, we initially had problems with baseline depression in regions around faint emission that was not correctly masked during the fitting process. Such artifacts could be largely suppressed using 2D polynomial fits in combination with improved masking procedures. However, the nature of any baseline fitting process makes the baseline estimate within a masked region uncertain, because one can only interpolate between the surrounding data points. This becomes a serious problem if the mask covers several tens to hundreds of spectral channels and high polynomial orders are used. Owing to the complexity of the baseline, we have to apply relatively high polynomial orders (see Sect. 3.2). From our experience, at low Galactic latitudes where the mask can be up to a few hundred spectral channels wide, uncertainties become significant. Not knowing the true baseline

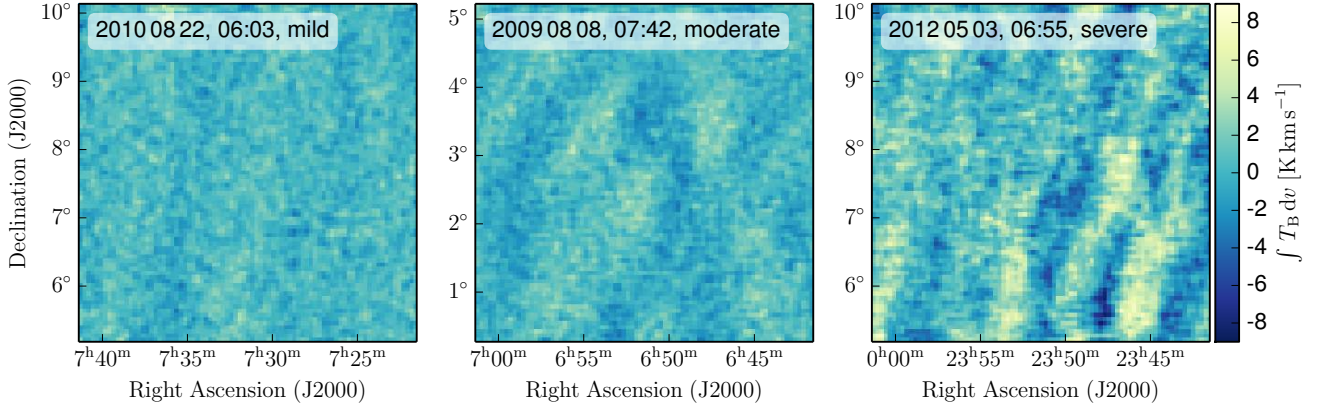


Fig. 12. Examples for the classification of solar interference. Fields with moderate or severe ripples were observed again. To emphasize the effect, each panel shows the velocity-integrated T_B over 100 spectral channels.

(and there is no way to measure it accurately³), we can only try to assess resulting baseline errors in a statistical sense (see Sect. 5).

There are two additional effects that can cause baseline problems: solar interference and strong continuum sources. We describe them in the following.

4.4.1. Solar interference

In the L band, the Effelsberg 100-m is vulnerable to solar radio emission, which can produce strong time-variable sinusoidal intensity variations, “ripples”, in the spectra. To avoid this so-called solar interference, EBHIS observing sessions are mainly scheduled during night-time. Still, about a tenth of the maps have been recorded during the daytime (mainly sunsets/-rises at beginning/end of a session). For these, we maximized the angular distance to the Sun; however, in some cases the data is still affected to some extent. This is caused by scattering of the solar radiation at the telescope structure, i.e., entering the receiving system via far-side lobes.

In Fig. 12 we give an example of how the resulting ripples appear in the final data cubes. To make them visible, 100 spectral channel maps without line emission were summed up. Three different fields, all observed during sunrise, show various severity levels from mild (rms increased by less than 50%), moderate (50% to 100%), to severe distortion ($\geq 100\%$). It is interesting to note, that – by chance – the distance of the three field centers to the Sun was about 40° , and also the bearing was comparable ($\sim 55^\circ$ for left and right panels, $\sim 70^\circ$ for center panel). Despite the similar conditions, the ripple differs significantly in strength, demonstrating that the degradation depends sensitively on the exact path of incident of solar rays.

In the majority of the fields affected by solar interference, the ripples are difficult to spot in the full-resolution data cube. Despite that, we decided to re-observe all maps that were classified as “moderate” or “severe” cases.

³ Using a hot or cold load would in principle allow removing some of the complexity in the underlying system temperature curve, allowing smaller polynomial orders. However, for the 21-cm seven-feed receiver, the load would need to be placed in front of the feeds – because the first stages of the front end are thought to produce most of the variations – which is simply not possible for practical reasons given the physical size of the feed assembly.

4.4.2. Standing-wave pattern

Like most modern receiving systems, the L -band seven-beam receiver at Effelsberg utilizes low-noise amplifiers (LNAs) that have a much higher bandwidth than what is processed by the intermediate-frequency system and the back-ends. Such broadband LNAs can make a system vulnerable to “out-of-band” RFI at usually higher⁴ frequencies than recorded.

Furthermore, the L -band receiver is also used for ongoing pulsar surveys. For these projects, a larger bandwidth of 250 MHz is desired (e.g., Kramer & Champion 2013; Barr et al. 2013). Therefore, the seven-feed system was designed to process frequencies above 1260 MHz, and the 100 MHz bandpass filter required for EBHIS is only inserted at the very end of the IF chain.

As a result, the receiver feeds and front-end LNAs are sensitive to radiation in the frequency range of about 1260 MHz to 2+ GHz. For the receiver designers, GSM at 1.8 GHz was a major concern, because it could easily drive the receiver into saturation. As a countermeasure, the seven-beam system makes use of waveguide resonance cavities, which provide a bandpass-like behavior around the frequency range of interest (for details see Keller et al. 2013).

A side effect of this is a multimodal wave-like pattern across the spectrum. It can be well-fitted using our baseline algorithms, though the necessary polynomial order in the spectral domain needs to be relatively high (compare Sect. 3.2). A more serious side effect, however, is that strong continuum sources cause a change in the pattern. The baseline solution, which is calculated for each subscan as a whole (2D fit, see Sect. 3.2), is then not working well. As a consequence, bright continuum sources (above a few Jy) produce a residual standing-wave-like pattern in the data cubes, since the average continuum level itself is subtracted from the spectral line cubes.

The residual ripples can fortunately be described by low-order polynomials. As a post-processing step, we usually apply a third-order polynomial fit across the full velocity range, i.e., not on the subtiles used for 2D baselining. For NVSS sources in excess of 1.5 Jy, we use an adaptive scheme, where the polynomial order is increased (maximum: 8) until the residual is sufficiently flat; i.e., the rms in the residual has converged. This approach produces very good baselines even for ~ 100 -Jy-class sources (see Fig. 13). For M 1 (top panel) and M 87 (bottom panel)

⁴ This is because a waveguide acts as high-pass filter, i.e., all radiation above a certain cutoff frequency can enter the system.

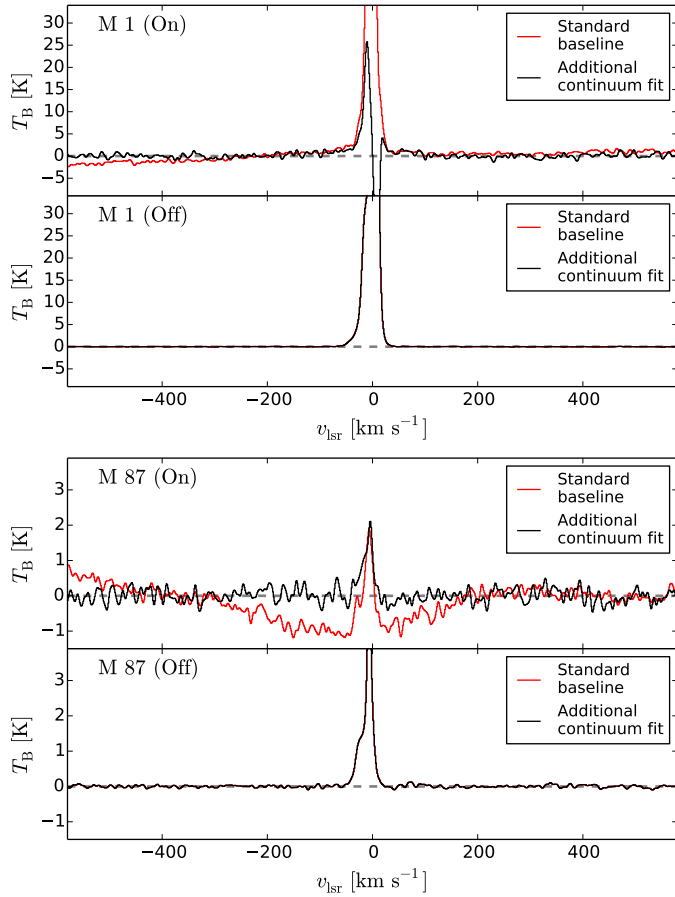


Fig. 13. Two of the brightest continuum sources, M 1 (Tau A, *top panel*) and M 87 (Vir A, *bottom panel*). For comparison we also show OFF positions below each ON source spectrum. Such strong continuum sources can cause residual ripples in the final data (red curves). Therefore, we employ an additional adaptive low-order polynomial baseline fitting, producing sufficiently flat baselines (black curves).

and respective OFF positions in the direct vicinity, the spectra with and without continuum-polynomial-fit post-processing are shown. In all cases, the additional treatment produces very flat baselines.

In Fig. 13 two additional effects, which are associated with continuum radiation, can be studied. First, the brightness temperature noise is significantly higher toward bright continuum sources, caused by the much higher system temperature. Second, for M 1 (ON), the emission line itself appears completely different in the post-processed version. In fact, the EBHIS pipeline measures the (line-free) rms in each spectral dump and flags outliers as bad. Therefore, in the final data product, some original spectral dumps were not accounted for in the gridding process – those being located closest to the bright continuum source. As these contain the strongest absorption imprint, the associated red curve in Fig. 13 is missing parts of the HI absorption feature. The effect is also visible for the M 87 case but to a lesser degree.

5. Uncertainties of EBHIS data

Uncertainties of the EBHIS data can be subdivided into three main categories: pure thermal noise, calibration uncertainties, and errors in the baseline (which includes SR). Brightness temperature rms can be determined easily from emission line-free regions in the HI data under the assumption that the noise is

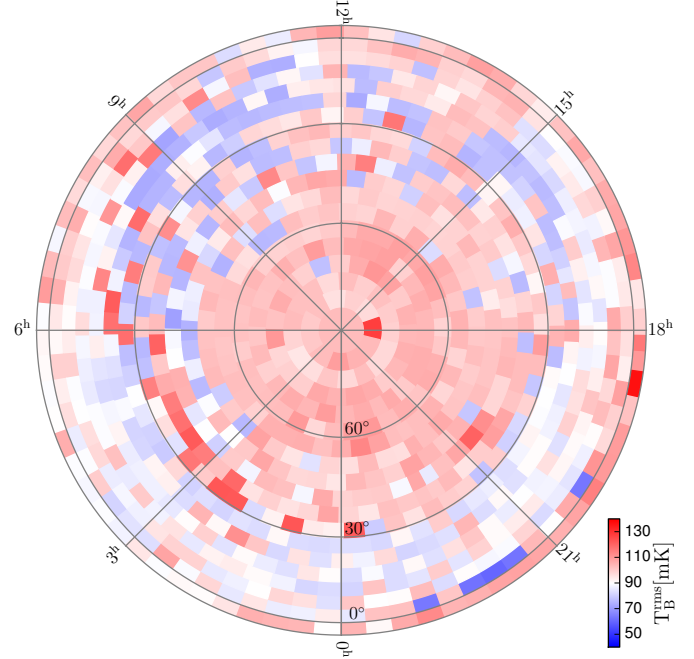


Fig. 14. EBHIS noise, σ_{rms} , at full resolution per observed field.

normal-distributed. In Sect. 4.1 we have already shown that intensity calibration is unbiased with respect to the established LAB survey data. Here, we have a closer look at the typical scatter in the calibration and on systematic effects caused by baseline subtraction and SR correction, which is a more complicated task.

5.1. Noise map

EBHIS data are taken at all seasons, thus at different ambient temperatures, elevation angles, and system temperatures. This means that final rms noise levels in the spectral data can differ substantially. Even though the MW covers a substantial fraction of the Galactic EBHIS data, each individual spectrum contains enough emission line-free spectral channels to robustly compute the rms from the data itself. We applied an iterative procedure of repeatedly computing the standard deviation σ_{rms} in a spectrum and subsequently clipping outliers in excess of $3\sigma_{\text{rms}}$. Usually, this scheme converges after three or four iterations. In Fig. 14 we display the average rms, measured this way, per observed field. The overall median rms value is $\sigma_{\text{rms}} = 90$ mK.

The rms in spectral channels containing HI line emission will be higher, because line emission contributes to the system temperature and as such increases noise. Likewise, the baseline system temperature is not constant such that the rms is slightly frequency-dependent (typically $\lesssim 1\%$ over the MW velocity interval).

5.2. Uncertainties from EBHIS–LAB comparison

In the previous sections, we have already discussed how difficult it is to disentangle the various uncertainties in the final data products (HI column densities or brightness temperature spectra). In particular, the stray-radiation contribution can hardly be separated from baseline-fit errors. Therefore, we attempt to quantify the uncertainty of the column density distribution as a whole using empirical methods.

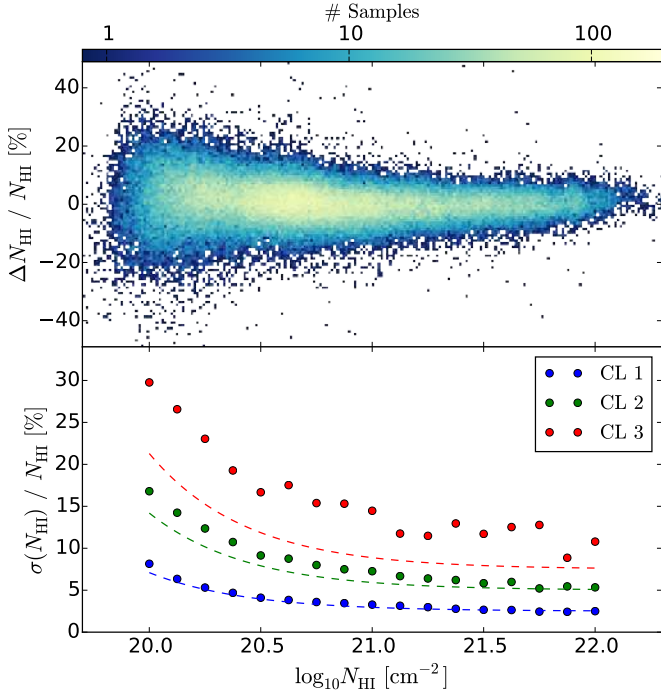


Fig. 15. Uncertainty of the EBHIS–LAB column density distribution. *Top panel:* relative column density difference of EBHIS with respect to LAB. *Bottom panel:* the 68.3% (CL-1), 95.4% (CL-2), and 99.7% (CL-3) confidence intervals from the distribution in the top panel. Because the difference between two distributions has a wider spread than the two constituents, we rescale the difference ΔN_{HI} to the equivalent standard deviation $\sigma(N_{\text{HI}}) = \Delta N_{\text{HI}}/\sqrt{2}$. The dashed lines represent the theoretical expectation, for the case where the error is caused by pure thermal noise, as measured in emission-line free regions, plus an additional contribution of 2.5% that represents the typical intensity calibration uncertainties.

One method to do this was already utilized in Sect. 4.1. By comparing EBHIS column densities to the LAB survey we not only can check on the correct flux-density scale, but the relative ensemble error can also be inferred from the scatter about the best-fit linear relation in the $N_{\text{HI}}^{\text{EBHIS}} - N_{\text{HI}}^{\text{LAB}}$ or $T_{\text{b}}^{\text{EBHIS}} - T_{\text{b}}^{\text{LAB}}$ plots (Fig. 9), for example, as a function of column density or brightness temperature. As an example, in Fig. 15 (top panel), a histogram of the percentaged column density differences, $\Delta N_{\text{HI}} \equiv N_{\text{HI}}^{\text{EBHIS}} - N_{\text{HI}}^{\text{LAB}}$, is shown as a function of $\log N_{\text{HI}}$.

Because we work with the difference of two noisy distributions, the uncertainties of the two contributing data sets have to be smaller. If both distributions had the same standard deviation, scaling the measured width of the resulting distribution by $\sqrt{2}$ would provide a good estimator of the original width. The LAB and EBHIS surveys both have different noise levels. Taking the ratio of rms values into account, one could easily correct for this. However, systematic uncertainties need not necessarily be distributed in the same way, for instance, their ratio could be different. It is not even clear whether the uncertainty distributions of both surveys have a similar shape. Therefore, we assume that both data sets share the same magnitude of systematic effects. This is justified further by the rms values of LAB and EBHIS, which are comparable, and in our experience the quality of the data reduction scales with the noise level.

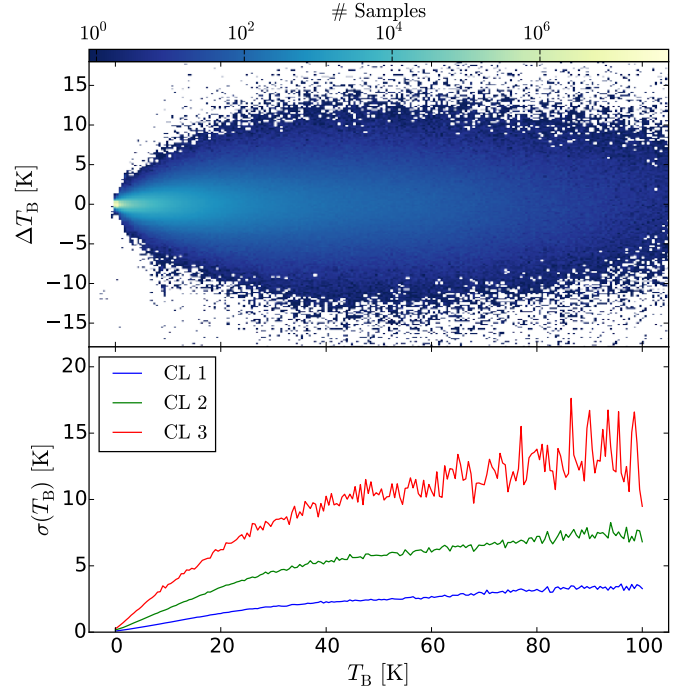


Fig. 16. Uncertainty of the EBHIS–LAB brightness temperature distribution. *Top panel:* brightness temperature difference of EBHIS with respect to LAB. *Bottom panel:* the 68.3% (CL-1), 95.4% (CL-2), and 99.7% (CL-3) confidence intervals from the distribution in the top panel. Similar to Fig. 15, we use $\sigma(T_{\text{B}}) = \Delta T_{\text{B}}/\sqrt{2}$.

5.2.1. Column density uncertainties

Applying the aforementioned scaling by $\sqrt{2}$, we can estimate the standard deviation of the EBHIS column density distribution, again as a function of $\log N_{\text{HI}}$. This is plotted in the lower panel of Fig. 15 (filled circles), as calculated in bins of width 0.125 dex. We show three confidence intervals, 68.3% (CL-1), 95.4% (CL-2), and 99.7% (CL-3), as calculated from the distribution in the top panel. The dashed lines represent the theoretical expectation, in the case where the uncertainties are caused by thermal noise in addition to a 2.5% increase. This additional contribution reflects the typical S7-based intensity calibration uncertainties (Winkel et al. 2010).

The CL-1 theoretical curve matches the measured distribution width very well. Only the lowest two data points ($\log N_{\text{HI}} \lesssim 20.25$) show a slight excess, most likely caused by the imperfect SR- and baseline correction. For the two other confidence levels, the estimated distributions are wider than predicted, meaning that the systematic errors are not perfectly normal-distributed.

5.2.2. Brightness temperature uncertainties

Following the same approach, we evaluated the brightness temperature uncertainties (see Fig. 16). For very low values of T_{B} , the errors are consistent with the rms noise level in EBHIS. At the highest T_{B} values, the uncertainty approaches about 2.5% of the signal brightness temperature, which is compatible with the column density case.

5.3. Uncertainties from EBHIS–EBHIS comparison

A particular drawback of the EBHIS–LAB comparison is the impossibility of getting distinct systematic errors for each of the

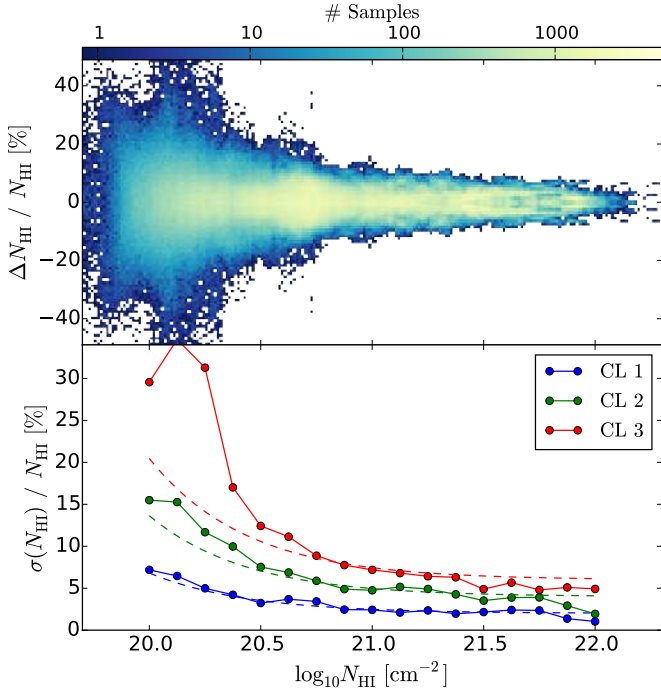


Fig. 17. As in Fig. 15 but computed from column density differences of EBHIS in the overlap regions of the observed fields.

two surveys. It remains unclear, whether EBHIS or LAB or both surveys together determine the scatter in the uncertainty distributions. To circumvent this problem, we tried to use EBHIS data alone for a similar study. Utilizing the overlaps of the 915 observed fields, a scatter histogram can again be inferred. One particular difficulty is that the beam weights in the overlap regime are lower than in the inside of the fields, and the rms is likewise higher⁵. Furthermore, by using overlap areas alone, the sampling of data points on the sky becomes highly irregular.

This effect is visible in Figs. 17 and 18 where the distribution of N_{HI} and T_{B} deviations is much less Gaussian-like than in Figs. 15 and 16. Nevertheless, the confidence intervals (lower panels) are even slightly narrower than inferred from the EBHIS–LAB comparison. In particular, the $\sigma(T_{\text{B}})$ inferred from EBHIS alone is lower and discloses a roughly linear behavior, $\sigma(T_{\text{B}}) \approx 0.02 \cdot T_{\text{B}}$ for $T_{\text{B}} \gg \sigma_{\text{rms}}$.

5.4. Discussion of derived uncertainty curves

Among the published data products, the Milky Way column density map is probably of highest scientific interest. In this section, we presented the first-ever attempt to quantify the systematic and statistical uncertainties of the N_{HI} distribution.

Although the measured N_{HI} distributions are not perfectly normal distributed, it is probably fair to say that uncertainties are mostly determined by the 2.5% intensity calibration error already reported in Winkel et al. (2010). This is a typical value for single-dish HI observations. Only the lowest column densities, $N_{\text{HI}} \lesssim 2 \times 10^{20} \text{ cm}^{-2}$, show a slight excess, which we attribute to baseline and SR-correction problems. However, compared to the pure thermal noise and calibration uncertainties, their impact is minor.

⁵ If two neighboring observations are combined, both of them contribute to the overlap so that the rms in the overlap is even slightly smaller than usual.

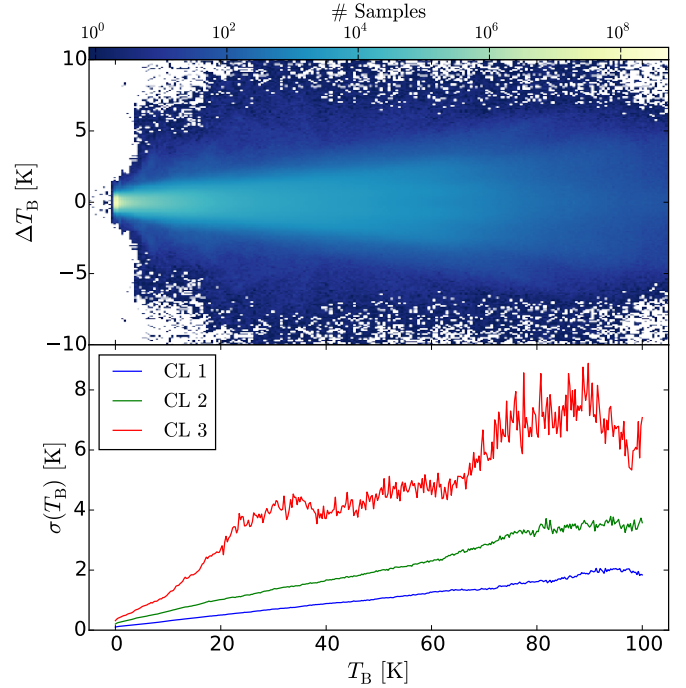


Fig. 18. As in Fig. 16 but computed from brightness temperature differences of EBHIS in the overlap regions of the observed fields.

Also the T_{B} error distributions reveal interesting results. While $\sigma(T_{\text{B}})$ from EBHIS alone is approximately 2% over the full range, it shows a completely different behavior in the EBHIS–LAB plot, where it is higher at medium values of T_{B} . We attribute this to the LAB survey, but the reason for this remains unclear.

It is important to note that for small T_{B} , the predictions from the uncertainty distributions should not be used. Here, one should work with the local brightness temperature rms that can easily be measured for each sight line and with high accuracy.

The results indicate that EBHIS performs somewhat better than LAB, but to be conservative, it can be safely assumed that EBHIS is at least on a par with LAB in terms of systematic uncertainties. With the completion of the second sky coverage of EBHIS data, it will be possible to assess the ensemble uncertainties to much better precision from EBHIS alone.

5.5. Polynomial parametrization of uncertainties

For convenience, we calculated parametrizations of the uncertainty values displayed in Figs. 15 to 18 based on piece-wise polynomials. For the N_{HI} errors, Table 2 contains column density intervals with respective polynomial coefficients. It is

$$\frac{\sigma(N_{\text{HI}})}{N_{\text{HI}}} [\%] = \sum_{k=0}^2 a_k (\log N_{\text{HI}})^k \quad (11)$$

for each bin $\log N^l \leq \log N_{\text{HI}} < \log N^u$.

Likewise, Table 3 provides coefficients for the brightness temperature uncertainties, with

$$\sigma(T_{\text{B}}) = \sum_{k=0}^2 a_k \left(\frac{T_{\text{B}}}{100 \text{ K}} \right)^k \quad (12)$$

and bin intervals $T_{\text{B}}^l \leq T_{\text{B}} < T_{\text{B}}^u$.

Table 2. Coefficients for a piece-wise polynomial parametrization of N_{HI} uncertainty curves in Figs. 15 and 17.

CL	$\log N^l$	$\log N^u$	a_0	a_1	a_2
LAB–EBHIS (Fig. 15)					
1	19.0	20.5	7851.92	-768.187	18.800
1	20.5	23.0	279.89	-24.953	0.561
2	19.0	20.5	6490.52	-625.672	15.099
2	20.5	23.0	723.14	-64.890	1.466
3	19.0	20.5	-3167.04	344.222	-9.219
3	20.5	23.0	1090.50	-96.962	2.176
EBHIS–EBHIS (Fig. 17)					
1	19.0	20.5	-181.71	26.872	-0.871
1	20.5	23.0	109.16	-8.740	0.175
2	19.0	20.5	-9280.02	936.857	-23.603
2	20.5	23.0	324.47	-27.141	0.569
3	19.0	20.5	-125472.84	12466.934	-309.591
3	20.5	23.0	2112.36	-193.710	4.451

Table 3. Coefficients for a piece-wise polynomial parametrization of T_{B} uncertainty curves in Figs. 16 and 18.

CL	T_{B}^l	T_{B}^u	a_0	a_1	a_2
LAB–EBHIS (Fig. 16)					
1	0	3	0.082	4.217	51.244
1	3	10	0.086	5.255	12.418
1	10	40	-0.123	9.470	-8.800
1	40	150	1.391	2.165	-
2	0	3	0.168	10.207	157.108
2	3	10	0.116	16.675	-0.965
2	10	40	-0.203	21.741	-19.675
2	40	150	3.765	3.951	-
3	0	3	0.257	28.598	176.797
3	3	10	0.218	35.736	-17.435
3	10	40	0.256	36.996	-33.836
3	40	150	6.743	7.246	-
EBHIS–EBHIS (Fig. 18)					
1	0	3	0.104	0.031	-0.003
1	3	150	0.108	0.019	-
2	0	3	0.211	0.070	-0.009
2	3	150	0.235	0.036	-
3	0	3	0.331	0.119	-0.013
3	3	150	0.337	0.079	-

6. Data products

We make EBHIS freely available to the astronomical community via CDS. Several different data products are provided for convenience.

6.1. Column density maps

For many purposes one only desires the total HI column density for a (set of) given position(s), e.g., to correct X-ray flux densities for foreground absorption. Full-(northern-)sky column density maps are published in FITS format (Hanisch et al. 2001, and references therein) in two sky projections, zenith-equal-area (ZEA; see Greisen & Calabretta 2002; Calabretta & Greisen 2002, compare Fig. 1), and HEALPix (Górski et al. 2005), which is widely used for all-sky datasets. An unofficial version also exists of the LAB column density map in the HEALPix representation⁶. It was reprojected from the original plate-carrée projection to HEALPix using nearest-neighbor resampling (Land & Slosar 2007). For this, we find significant

⁶ See <http://lambda.gsfc.nasa.gov/product/foreground/>

deviations compared to our own calculations, which were derived from the original LAB observations rather than from an intermediary data product. We recommend not using the unofficial LAB HEALPix map any longer.

6.2. Spectral data

The full spectral data will be provided for download in FITS binary table format on a HEALPix grid (nside-1024, Galactic coordinates are used for indexing). Because the EBHIS Milky Way spectral data set has a size of about 50 GByte, we split it into smaller chunks, namely into the 192 HEALPix pixels of the nside-4 tessellation. Because EBHIS is only available for declinations above -5° , not all of the nside-4 pixels contain data, and these are therefore omitted. Additionally, smaller spectral data cubes of size $20^\circ \times 20^\circ$ are available for download in FITS image format.

7. Summary

We presented the first data release of the Effelsberg-Bonn HI Survey (EBHIS), elaborated on the recent changes to the data processing software and techniques compared to Winkel et al. (2010), and studied the quality of the resulting data products in great detail. EBHIS is a 21-cm HI line survey of the sky north of $\delta \geq -5^\circ$, comprising full spatial sampling and high sensitivity. At full spectral resolution of $\delta v = 1.44 \text{ km s}^{-1}$, EBHIS has an average noise level of $\sim 90 \text{ mK}$. Its angular resolution is $10''.8$. About 1000 h of observing time with the Effelsberg 100-m telescope were spent until the completion of the first coverage that forms the basis for the first data release. Owing to the large bandwidth of 100 MHz, not only is the Milky Way emission covered in EBHIS, but also HI in the Local Volume out to a redshift of $z \sim 0.07$. The presented data products comprise the Milky Way emission. The analysis of the extragalactic part of the survey will be published elsewhere.

EBHIS was designed to replace the northern-hemisphere portion of the important Leiden/Argentine/Bonn Survey (LAB; Kalberla et al. 2005), which has about a factor of 4 lower angular resolution and severely suffers from the insufficient spatial sampling (beam-by-beam only; compare Kerp et al. 2011, e.g., their Fig. 7).

Great effort went into developing new data reduction techniques, including but not limited to software radio-frequency-interference mitigation, a frequency-dependent flux-density calibration scheme, and robust baseline estimation in the 2D time-frequency domain. Furthermore, the very successful stray-radiation prediction algorithms, which were also employed for the LAB and Galactic All-Sky Survey (GASS; McClure-Griffiths et al. 2009; Kalberla et al. 2010; Kalberla & Haud 2015), have been adapted to the Effelsberg telescope.

We carefully assessed the data quality of the EBHIS spectra and inferred column densities. Our investigation revealed a mismatch in intensity calibration of GASS with respect to EBHIS and LAB, which led to a revised calibration scale in the third data release of GASS (Kalberla & Haud 2015). EBHIS and LAB show a near-perfect one-to-one relation in column densities and brightness temperatures. From a comparison of EBHIS and LAB, and from the overlaps of the EBHIS fields themselves, we estimated the uncertainties of the EBHIS column density distribution and brightness temperatures. These uncertainties do not necessarily represent those of individual lines of sight but rather the ensemble properties. Still, the N_{HI} distribution uncertainties – calculated for the first time for a large-area single-dish

survey of the neutral hydrogen – increase the value of the published N_{HI} map, which will probably be the primarily used data product.

Column density maps and full spectral data have been made freely available to the public.

8. Outlook

At the moment, we are undertaking observations to map the sky north of $\delta \geq 30^\circ$ for a second time, and we will also ask to extend this campaign to cover the full northern hemisphere again. This second coverage will boost the EBHIS sensitivity by 30%, will allow even more sophisticated data reduction methods like basket-weaving to be applied, and will help to suppress residual RFI and SR.

In [Winkel et al. \(2012a\)](#) we presented a fast linear least-squares approach to apply to the basket-weaving observing technique. The idea is to map a portion of the sky twice, with two orthogonal scanning directions. Since the underlying astronomical sky is basically constant, one can disentangle scanning effects from the desired original brightness distribution. With respect to HI, apparent scanning effects are mostly caused by inaccurate baseline solutions, e.g., if (weak) HI sources are not correctly flagged or windowed. This is a common problem in single-dish astronomy, as shown by [Calabretta et al. \(2014\)](#), among others.

However, for the continuum flux densities inherently present in the EBHIS raw data, scanning effects are mostly given by changes in the underlying system temperature, especially since the different feeds have different RX temperatures. With basket-weaving it will be possible to remove these scanning effects very efficiently and produce a 1.4-GHz Stokes- I map as byproduct of our survey. The feasibility of this goal has been presented in [Winkel et al. \(2012a\)](#). However, we remind the interested reader that our SR solution only applies for the HI sky and not the 1.4 GHz continuum sky. This is because the HI SR removal does not need to account for ground radiation, while this contribution is significant for continuum. Therefore, like the HIPASS continuum data set ([Calabretta et al. 2014](#)), the EBHIS continuum data will suffer substantially from SR effects.

Acknowledgements. EBHIS is based on observations with the 100-m telescope of the MPIfR (Max-Planck-Institut für Radioastronomie) at Effelsberg. We would like to thank the observatory staff for the substantial and continuous assistance carrying out the measurements. We are grateful to Axel Jessner for carefully proofreading the manuscript and for his valuable comments. The authors thank the Deutsche Forschungsgemeinschaft (DFG) for support under grant numbers KE757/7-1, KE757/7-2, KE757/7-3, and KE757/9-1. B.W. was partially funded by the International Max Planck Research School for Astronomy and Astrophysics at the Universities of Bonn and Cologne (IMPRS Bonn/Cologne). L.F. is also a member of IMPRS Bonn/Cologne. D.L. is a member of the Bonn–Cologne Graduate School of Physics and Astronomy (BCGS). This research has made use of NASA’s Astrophysics Data System and the HyperLeda database ([Makarov et al. 2014](#)). We would like to express our gratitude to the developers of the many C/C++ and Python libraries, made available as open-source software, which we have used: most importantly, NumPy ([van der Walt et al. 2011](#)) and SciPy ([Jones et al. 2001](#)), Cython ([Behnel et al. 2011](#)), Astropy ([Astropy Collaboration et al. 2013](#)), and the GNU Scientific library ([Gough 2009](#)). Figures have been prepared using matplotlib ([Hunter 2007](#)) and in part using the Kapteyn Package ([Terlouw & Vogelaar 2015](#)).

References

Arnal, E. M., Bajaja, E., Larrarte, J. J., Morras, R., & Pöppel, W. G. L. 2000, *A&AS*, 142, 35
 Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, 558, A33

Baars, J. W. M. 2007, *The Paraboloidal Reflector Antenna in Radio Astronomy and Communication*, *Astrophys. Space Sci. Libr.*, 348
 Bajaja, E., Cappa de Nicolau, C. E., Cersosimo, J. C., et al. 1985, *ApJS*, 58, 143
 Bajaja, E., Arnal, E. M., Larrarte, J. J., et al. 2005, *A&A*, 440, 767
 Barnes, D. G., Staveley-Smith, L., de Blok, W. J. G., et al. 2001, *MNRAS*, 322, 486
 Barnes, D. G., Briggs, F. H., & Calabretta, M. R. 2005, *Radio Science*, 40, 5
 Barr, E. D., Champion, D. J., Kramer, M., et al. 2013, *MNRAS*, 435, 2234
 Behnel, S., Bradshaw, R., Citro, C., et al. 2011, *Comp. Sci. Eng.*, 13, 31
 Boggs, P. T., & Rogers, J. E. 1990, in *Statistical analysis of measurement error models and applications*, *Contemporary Mathematics*, 112, 186
 Boothroyd, A. I., Blagrove, K., Lockman, F. J., et al. 2011, *A&A*, 536, A81
 Bracewell, R. N., & Roberts, J. A. 1954, *Austr. J. Phys.*, 7, 615
 Calabretta, M. R., & Greisen, E. W. 2002, *A&A*, 395, 1077
 Calabretta, M. R., Staveley-Smith, L., & Barnes, D. G. 2014, *PASA*, 31, 7
 Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, *AJ*, 115, 1693
 Flöer, L., Winkel, B., & Kerp, J. 2010, in *RFI Mitigation Workshop*, 42
 Giovanelli, R., Haynes, M. P., Kent, B. R., et al. 2005, *AJ*, 130, 2598
 Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, *ApJ*, 622, 759
 Gough, B. 2009, *GNU Scientific Library Reference Manual – 3rd edn.* (Network Theory Ltd.)
 Greisen, E. W., & Calabretta, M. R. 2002, *A&A*, 395, 1061
 Hanisch, R. J., Farris, A., Greisen, E. W., et al. 2001, *A&A*, 376, 359
 Hartmann, D., & Burton, W. B. 1997, *Atlas of Galactic Neutral Hydrogen* (Cambridge, UK: Cambridge University Press), 243
 Hartmann, D., Kalberla, P. M. W., Burton, W. B., & Mebold, U. 1996, *A&AS*, 119, 115
 Hartsuijker, A. P., Baars, J. W. M., Drenth, S., & Gelato-Volders, L. 1972, *IEEE Transactions on Antennas and Propagation*, 20, 166
 Haynes, M. P., Giovanelli, R., Martin, A. M., et al. 2011, *AJ*, 142, 170
 Higgs, L. A. 1967, *Bull. Astron. Institutes of the Netherlands Suppl. Ser.*, 2, 59
 Higgs, L. A., & Tapping, K. F. 2000, *AJ*, 120, 2471
 Hunter, J. 2007, *Comp. Sci. Eng.*, 9, 90
 Jones, E., Oliphant, T., Peterson, P., et al. 2001, *SciPy: Open source scientific tools for Python*, <http://www.scipy.org>
 Kalberla, P. M. W., & Haud, U. 2015, *A&A*, 578, A78
 Kalberla, P. M. W., Mebold, U., & Reich, W. 1980a, *A&A*, 82, 275
 Kalberla, P. M. W., Mebold, U., & Velden, L. 1980b, *A&AS*, 39, 337
 Kalberla, P. M. W., Burton, W. B., Hartmann, D., et al. 2005, *A&A*, 440, 775
 Kalberla, P. M. W., McClure-Griffiths, N. M., Pisano, D. J., et al. 2010, *A&A*, 521, A17
 Keller, R., Nalbach, M., Müller, K., et al. 2006, *Multi-Beam Receiver for Beam-Park Experiments and Data Collection Unit for Beam Park Experiments with Multi-Beam Receivers*, *Tech. Rep.*, Max-Planck-Institut für Radioastronomie
 Keller, R., Dohlus, M., & Hauth, W. 2013, in *Electromagnetics in Advanced Applications (ICEAA)*, *International Conference on*, 1323
 Kerp, J., Winkel, B., Ben Bekhti, N., Flöer, L., & Kalberla, P. M. W. 2011, *Astron. Nachr.*, 332, 637
 Klein, B., Hochgürtel, S., Krämer, I., et al. 2012, *A&A*, 542, L3
 Kramer, M., & Champion, D. J. 2013, *Class. Quant. Grav.*, 30, 224009
 Land, K., & Slosar, A. 2007, *Phys. Rev. D*, 76, 087301
 Lockman, F. J., Jahoda, K., & McCammon, D. 1986, *ApJ*, 302, 432
 Makarov, D., Prugniel, P., Terekhova, N., Courtois, H., & Vauglin, I. 2014, *A&A*, 570, A13
 Martin, P. G., Blagrove, K. P. M., Lockman, F. J., et al. 2015, *ApJ*, 809, 153
 McClure-Griffiths, N. M., Pisano, D. J., Calabretta, M. R., et al. 2009, *ApJS*, 181, 398
 Minchin, R. F., Auld, R., Davies, J. I., et al. 2007, in *IAU Symp.* 235, eds. F. Combes, & J. Palouš, 227
 Peek, J. E. G., Heiles, C., Douglas, K. A., et al. 2011, *ApJS*, 194, 20
 Prestage, R. M., Constantikes, K. T., Hunter, T. R., et al. 2009, *IEEE Proc.*, 97, 1382
 Shannon, C. E. 1949, *Proc. Institute of Radio Engineers*, 37, 10
 Stanko, S., Klein, B., & Kerp, J. 2005, *A&A*, 436, 391
 Stark, A. A., Gammie, C. F., Wilson, R. W., et al. 1992, *ApJS*, 79, 77
 Terlouw, J. P., & Vogelaar, M. G. R. 2015, *Kapteyn Package*, version 2.3, Kapteyn Astronomical Institute, Groningen, available from <http://www.astro.rug.nl/software/kapteyn/>
 van der Walt, S., Colbert, S., & Varoquaux, G. 2011, *Comp. Sci. Eng.*, 13, 22
 van Woerden, H. 1962, *De neutrale waterstof in Orion*, Ph.D. Thesis, Univ. of Groningen
 Winkel, B., Kalberla, P. M. W., Kerp, J., & Flöer, L. 2010, *ApJS*, 188, 488
 Winkel, B., Flöer, L., & Kraus, A. 2012a, *A&A*, 547, A119
 Winkel, B., Kraus, A., & Bach, U. 2012b, *A&A*, 540, A140

Appendix A: Supplementary figures

A.1. Additional sky maps

Figure A.1 contains a composite image combining N_{HI} column densities and intensity-weighted velocities (Moment-1). For the column density part, the same FITS map as in Fig. 1 was used. For the Moment-1 image, first a mask was produced by Gaussian-filtering the original data cube and applying an intensity threshold. Otherwise, low-column density spectra would lead to numerical artifacts in the Moment-1 calculation. Intensity-weighted velocities are encoded as colors in Fig. A.1. The resulting image is converted to the HSV color space and the saturation- and value subchannels (the “brightness”) are modified according to the column density in each pixel. To account for the high density contrast (two orders of magnitude) in the N_{HI} map, the brightness is modified differently as a function of velocity. Otherwise, the relatively fainter intermediate- and high-velocity gas would hardly be visible. The two-dimensional color wedge depicts the HSV transfer function.

In Fig. A.2 the SR correction, marginalized to column densities, is visualized. The bottom panel is the total SR contribution as predicted by our SR model. The upper panels contain

near- (left) and far-side-lobe (right) contributions. The maps are on the same grid as the EBHIS N_{HI} map in Fig. 1 and share the same intensity scale to allow comparison. Toward higher Galactic latitudes, i.e., lower column densities, the SR contribution of the far side lobes gets increasingly important and can even be larger than the true column density in extreme cases.

A.2. Flux-density calibration and velocity consistency

In Figs. A.3 to A.5 we present linear correlation diagrams for consistency checks of the intensity calibration (for the spectral integral, i.e., N_{HI} value, and brightness temperatures per voxel) and velocity axis accuracies. To avoid the side effects caused by the insufficient spatial sampling in LAB data, we took the original pointing positions of LAB and regridded EBHIS and GASS data to match. There is a common slice close to zero declination, $-4.5^\circ \leq \delta \leq -0.2^\circ$, where all three surveys provide data. For the comparison of LAB and GASS we split LAB into its constituents LDS (northern part, $\delta > -27.5$) and IAR survey ($\delta < -27.5$). Both subsamples show different behavior in terms of intensity calibration and velocity zero point.

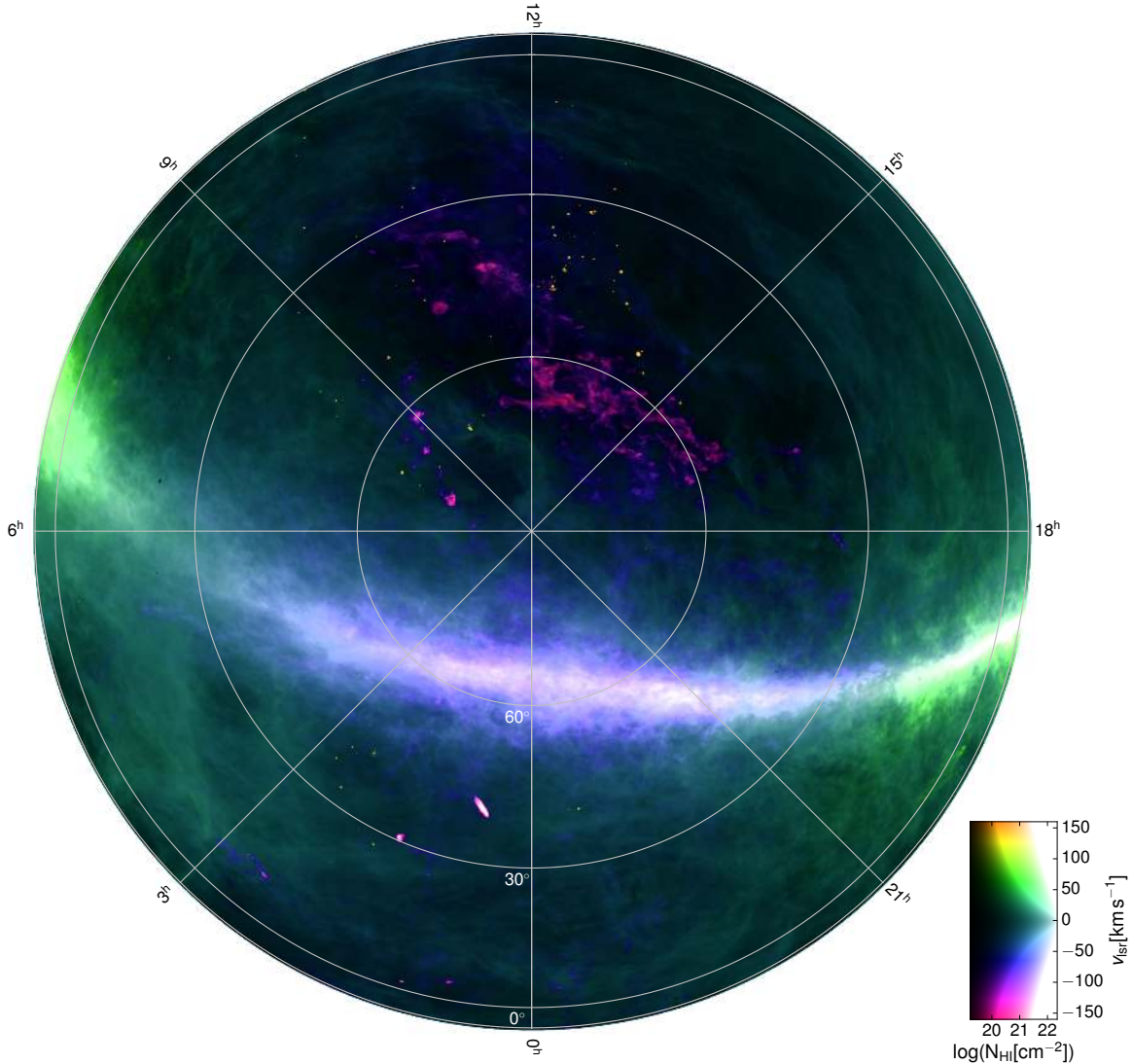


Fig. A.1. Composite image displaying N_{HI} column densities (encoded as brightness) and intensity-weighted velocities (Moment-1, encoded as different colors) for EBHIS. To make the weaker intermediate- and high-velocity gas more visible – compared to the bright MW disk emission – we modified the brightness scale as a function of velocity, as indicated in the color wedge.

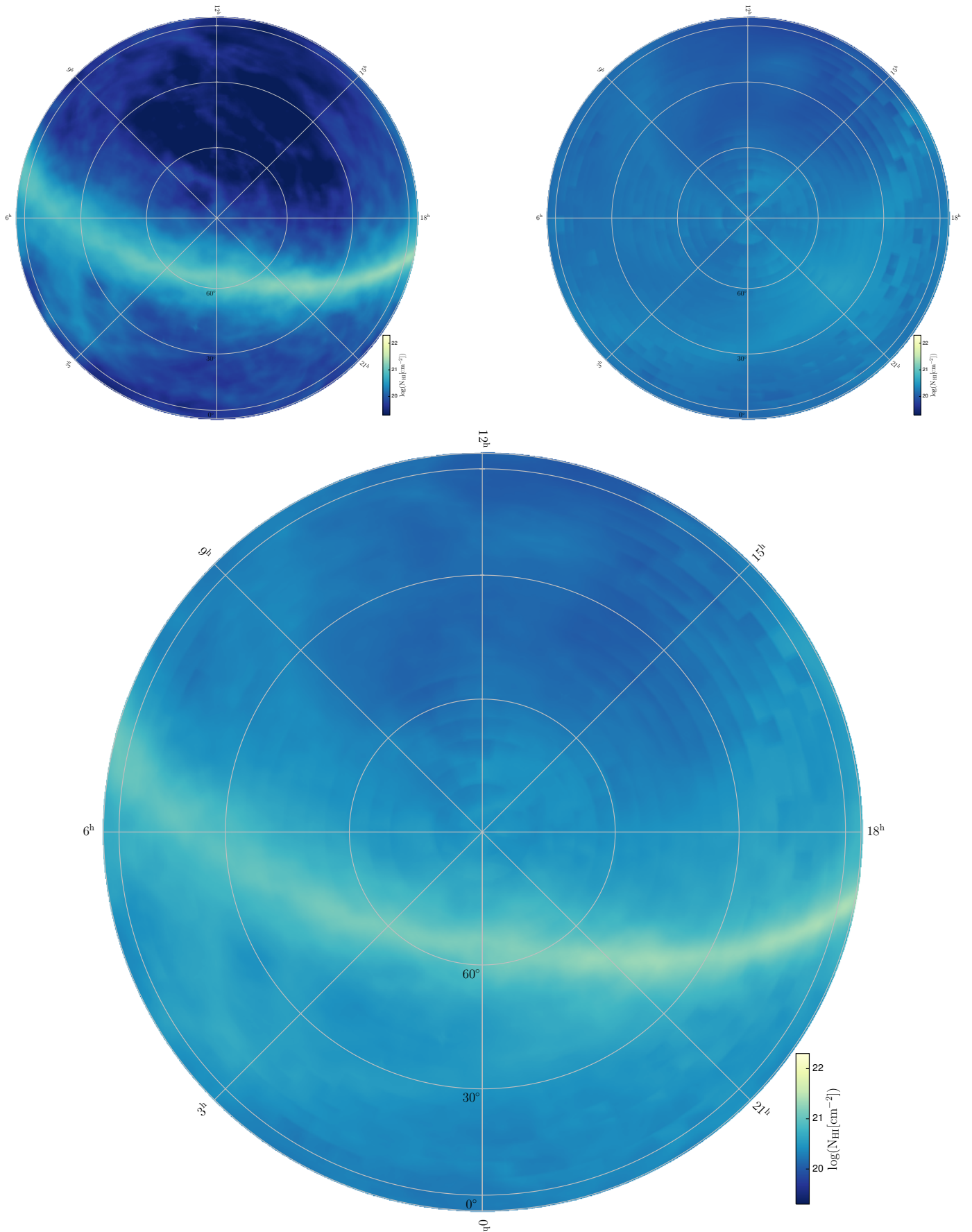


Fig. A.2. Stray-radiation correction maps showing the near- and far-side-lobe contribution (*top panels*), as well as the total amount of SR (*bottom panel*). The color scale matches the one used in Fig. 1 to allow direct comparison.

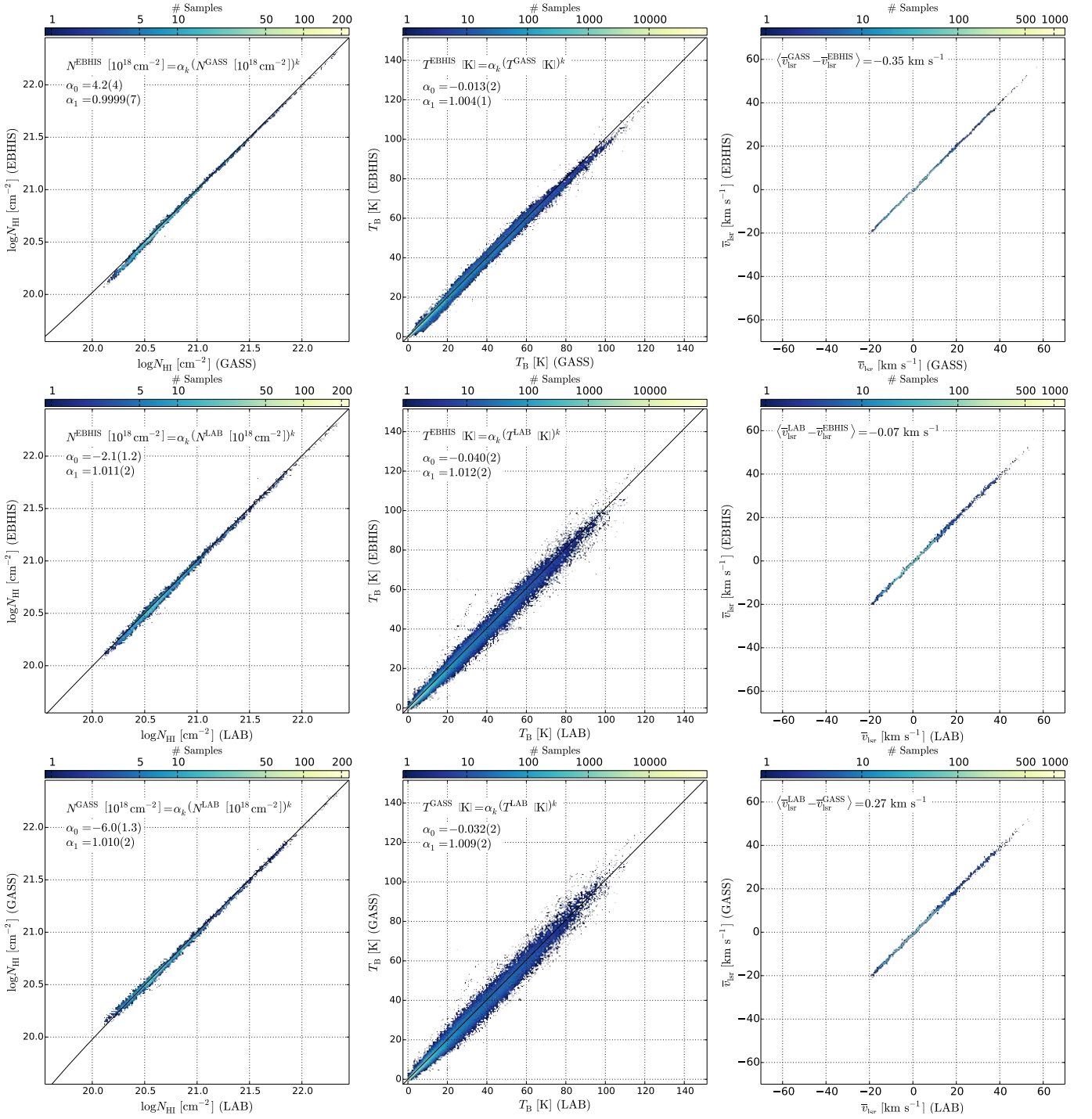


Fig. A.3. Comparison between EBHIS, GASS, and LAB in the overlap region $-4.5^\circ \leq \delta \leq -0.2^\circ$. *Left panels:* column densities; *center panels:* brightness temperatures; and *right panels:* intensity-weighted velocities (Moment-1).

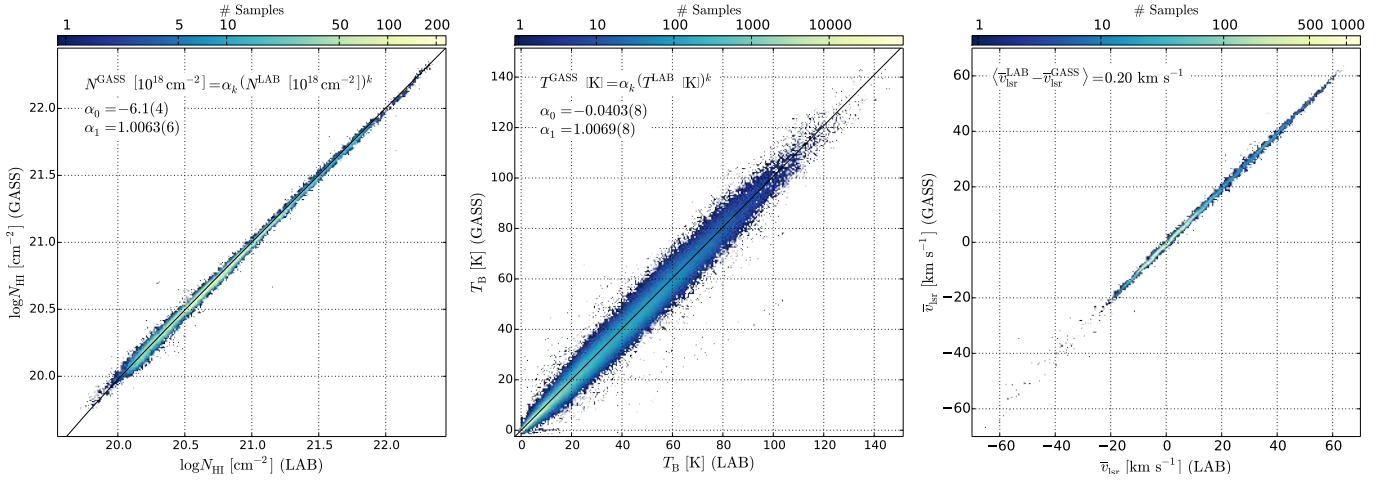


Fig. A.4. Comparison between GASS and LDS (LAB with $\delta > -30^\circ$). *Left panel:* column densities; *center panel:* brightness temperatures; and *right panel:* intensity-weighted velocities (Moment-1).

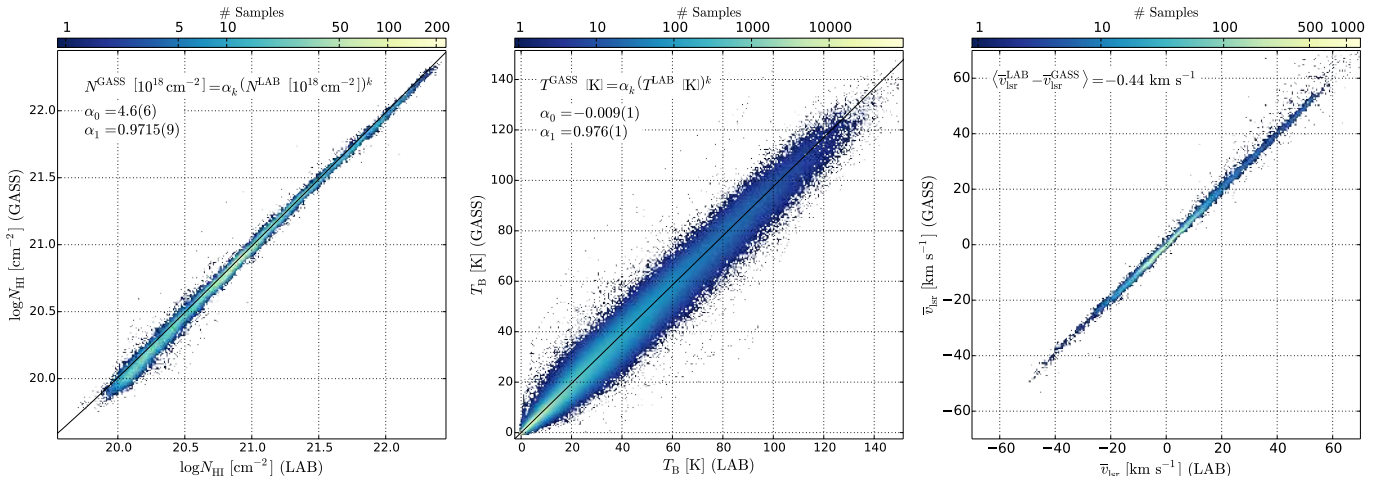


Fig. A.5. Comparison between GASS and IAR survey (LAB with $\delta < -30^\circ$). *Left panel:* column densities; *center panel:* brightness temperatures; and *right panel:* intensity-weighted velocities (Moment-1).