Neutral Hydrogen in and around Galaxies in the SKA Era
Proceedings IAU Symposium No. 392, 2026
D.J. Pisano, K. M. Mogotsi, J. Healy & S. Blyth, eds.
doi:10.1017/S174392132500050X

WALLABY Pilot Survey and Data Release

Tobias Westmeier,[1,2] Barbara Catinella,[1,2] Nathan Deg,[3] Bi-Qing For,[1,2] Karen Lee-Waddell,[1,4] Chandrashekar Murugeshan,[4,2] Kristine Spekkens[3,5] and Lister Staveley-Smith[1,2]

[1]ICRAR, The University of Western Australia, 35 Stirling Highway, Crawley WA 6009, Australia
email: tobias.westmeier@uwa.edu.au
[2]ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D)
[3]Department of Physics, Engineering Physics & Astronomy, Queen's University, Kingston ON K7L 3N6, Canada
[4]CSIRO Space and Astronomy, PO Box 1130, Bentley WA 6102, Australia
[5]Department of Physics and Space Science, Royal Military College of Canada, P.O. Box 17000, Station Forces Kingston ON K7K 7B4, Canada

Abstract. The Wide-field ASKAP L-band Legacy All-sky Blind Survey (WALLABY) is a large HI survey being carried out with the Australian SKA Pathfinder (ASKAP). WALLABY will image about 14,000 square degrees of the sky in the 21-cm line of neutral hydrogen and is expected to detect approximately 200,000 galaxies out to a redshift of about 0.1. What makes WALLABY unique is the combination of a large galaxy sample with sufficiently high angular resolution (30″ FWHM) to spatially resolve a significant fraction of those galaxies. This enables entirely new scientific studies, such as the search for HI polar ring galaxy candidates or the extraction of resolved galaxy HI scaling relations for statistical samples. While the full survey is still underway, catalogues, images and spectra of nearly 2400 galaxies detected during the pilot phase have already been publicly released. More data releases are anticipated over the next few years as the survey progresses.

Keywords. radio lines: galaxies, surveys, catalogues

1. Introduction

The Wide-field ASKAP L-band Legacy All-sky Blind Survey (WALLABY; Koribalski et al. 2020) is a large interferometric 21-cm neutral hydrogen survey currently being carried out with the Australian Square Kilometre Array Pathfinder (ASKAP; Hotan et al. 2021). The survey was originally designed to image the entire sky south of a declination of $\delta = +30°$ with the primary aim to detect the 21-cm HI emission line from approximately half a million galaxies out to a redshift of $z = 0.26$. As a result of the lower-than-expected sensitivity of the phased array feeds on ASKAP (Hotan et al. 2021) the sky coverage of WALLABY had to be reduced to about 14,000 square degrees ($\approx 1.4\pi$ sr) while retaining the originally envisaged depth of the survey of about 1.6 mJy RMS per beam and spectral channel.

As shown in Fig. 1, the new survey footprint of WALLABY will cover a limited declination range of $-60° \lesssim \delta \lesssim -15°$. In addition, WALLABY will image a few separate regions of interest with excellent multi-wavelength coverage, namely the Virgo cluster region, the GAMA G02 field (Driver et al. 2009; Liske et al. 2015) and two narrow equatorial strips aligned with the Hyper Suprime-Cam Subaru Strategic Program (HSC–SSP; Aihara et al. 2018). In addition, the increased level of radio frequency interference from global

© The Author(s), 2026. Published by Cambridge University Press on behalf of International Astronomical Union.

https://doi.org/10.1017/S174392132500050X Published online by Cambridge University Press

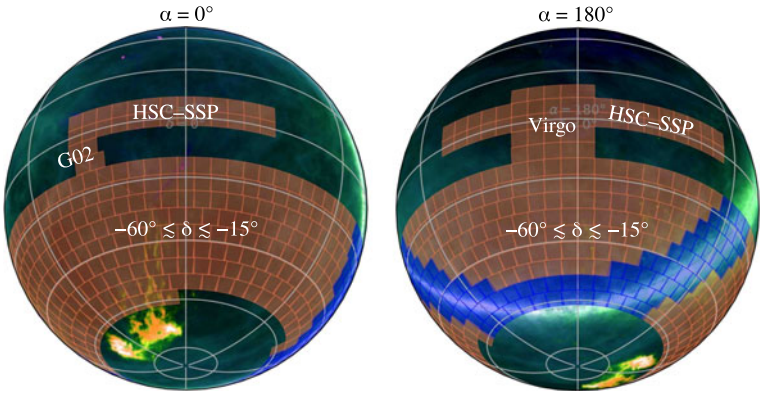


Figure 1. Location of the 552 WALLABY survey tiles on the sky in equatorial coordinates. The left-hand map is centred on $\alpha = 0°$, the right-hand map on $\alpha = 180°$. Tiles located inside the Zone of Avoidance are shown in blue colour, all other tiles in brown. The background image shows an integrated HI flux map from the HI4PI survey (HI4PI Collaboration 2016) for orientation, revealing the locations of the Galactic plane and the Magellanic system.

Table 1. Overview of updated WALLABY survey parameters.

Parameter	Value	Unit	Notes
Number of ASKAP tiles	552		
Sky coverage	14,000	deg^2	1.4π sr
Maximum baseline	2	km	
Angular resolution	30	arcsec	
Frequency range	1295.5–1439.5	MHz	
Redshift range	$\lesssim 0.1$		
Frequency resolution	18.5	kHz	
Velocity resolution	3.9	$\mathrm{km\,s^{-1}}$	at $z = 0$
Integration time per tile	16	h	2 epochs of 8 h each
RMS noise	1.6	mJy	per beam and channel
Column density sensitivity	8.9×10^{19}	$\mathrm{cm^{-2}}$	$5\sigma \times 0.1$ MHz at $z = 0$

navigational satellites below a frequency of about 1310 MHz has forced us to restrict the redshift range of the survey to $z \lesssim 0.1$ by processing and imaging only the upper half of the 288 MHz instantaneous bandwidth of ASKAP. Despite the limited redshift range and sky coverage, WALLABY is still expected to detect approximately 200,000 galaxies. The updated WALLABY survey specifications are summarised in Table 1.

WALLABY imaging only uses ASKAP baselines shorter than 2 km in combination with adequate weighting and tapering to achieve a constant angular resolution (FWHM of the synthesised beam) of close to $30''$ across the processed bandwidth of 1295.5 to 1439.5 MHz. At a distance of 10 Mpc this corresponds to a linear scale of about 1.5 kpc, meaning that WALLABY will at least marginally resolve a significant fraction of the detected galaxies (see Koribalski et al. 2020 for details). The spectral channel width of $1/54$ MHz ≈ 18.5 kHz corresponds to a velocity resolution of roughly 4 $\mathrm{km\,s^{-1}}$ at $z = 0$ which is sufficient for extragalactic objects, but may not be ideal for studies of the cold neutral medium in the Milky Way and Magellanic system.

What makes WALLABY unique compared to previous HI surveys is the fact that it combines large galaxy sample sizes with a sufficiently small beam to spatially resolve a significant fraction of those galaxies. According to the WALLABY reference simulation presented in Koribalski et al. (2020), about 2,300 galaxies are expected to be resolved by more than 5 beams across the major axis (after correcting for the reduced sky coverage of 1.4π sr), with tens of thousands of galaxies expected to be marginally resolved. This enables the WALLABY team to address a wide range of

https://doi.org/10.1017/S174392132500050X Published online by Cambridge University Press

cutting-edge science cases, including global and spatially resolved HI scaling relations (Reynolds et al. 2023; Deg et al., in prep.; Lee et al., in prep.), gas stripping in groups and clusters (Wang et al. 2021; Reynolds et al. 2021, 2022; For et al. 2021; Murugeshan et al. 2021; Wang et al. 2022; Kim et al. 2023), the Tully–Fisher relation (Courtois et al. 2023; Mould et al. 2024), the search for so-called "dark" galaxies (Wong et al. 2021; O'Beirne et al. 2024), the study of ultra-diffuse galaxies (For et al. 2023) and polar ring galaxies (Deg et al. 2023), galaxy morphometrics (Holwerda et al. 2023) and the study of HI associated with the host galaxies of Fast Radio Bursts (Glowacki et al. 2023).

2. WALLABY Pilot Survey

Prior to the start of full survey observations in early 2023 the WALLABY team carried out a pilot survey from 2019 to 2022. The main purpose of the pilot survey was to test and refine our survey strategy and post-processing pipelines, but the target fields were chosen such that they would be of scientific interest at the same time. Phase 1 of the pilot survey (2019–2020) targeted three 60 $\deg^2$ fields in the direction of the Hydra and Norma galaxy clusters and the NGC 4636 galaxy group. Phase 2 of the pilot survey targeted two 30 $\deg^2$ fields in the direction of the NGC 4808 galaxy group and the Vela supercluster in addition to a large 120 $\deg^2$ mosaic of the region around the NGC 5044 galaxy group in the supergalactic plane.

Pilot observations were carried out using the full ASKAP array of 36 antennas and hence reach the same sensitivity and $30''$ resolution as the full survey. The ASKAPsoft data reduction pipeline (Guzman et al. 2019) produces calibrated, continuum-subtracted and deconvolved HI data cubes for each individual 8-hour footprint which are made available to the WALLABY team for post-processing and the creation of value-added data products such as mosaicked data cubes of each field or source catalogues.

Source finding was carried out using version 2 of the Source Finding Application (SoFiA; Serra et al. 2015; Westmeier et al. 2021) which is available on GitLab.† To improve the completeness and reliability of the resulting source catalogue, each detection reported by SoFiA was visually confirmed by looking at the HI intensity and velocity maps, the integrated spectrum and an optical image from the Digitized Sky Survey. Detections deemed to be false positives were removed from the catalogue. In addition to the source catalogue, SoFiA also provides additional data products with each detection, including a small data cube and source mask, moment maps and an integrated HI spectrum. Please see Westmeier et al. (2022) and Murugeshan et al. (2024) for more details.

In addition to source finding, we also attempted to model the kinematics of spatially resolved galaxies by fitting tilted ring models to all galaxies with a major axis size greater than two times the beam size or a signal-to-noise ratio of $\log_{10}(\mathrm{SNR}) \geq 1.25$. Initially, Fully Automated TiRiFiC (FAT; Kamphuis et al. 2015) and 3D-Barolo (Di Teodoro & Fraternali 2015) were used for tilted-ring fitting. The scripts developed for this purpose, referred to as the WALLABY Kinematic Analysis Proto-Pipeline (WKAPP), are available on GitHub.‡ The WALLABY team has now also developed its own tilted-ring fitting software called 3KIDNAS (pronounced /ɪˈkɪdnəz/) which is available on GitHub.§ For more information on the kinematic modelling please see Deg et al. (2022).

Overall, about 2350 galaxies were detected in both phases of the pilot survey combined, including 592 detections from phase 1 and 1760 detections from phase 2. Hence, WALLABY has already collected the largest sample of (at least marginally) resolved and uniformly observed HI detections available to date. The higher detection rate in phase 2

† SoFiA: https://gitlab.com/SoFiA-Admin/SoFiA-2
‡ WKAPP: https://github.com/CIRADA-Tools/WKAPP
§ 3KIDNAS: https://github.com/NateDeg/3KIDNAS

https://doi.org/10.1017/S174392132500050X Published online by Cambridge University Press

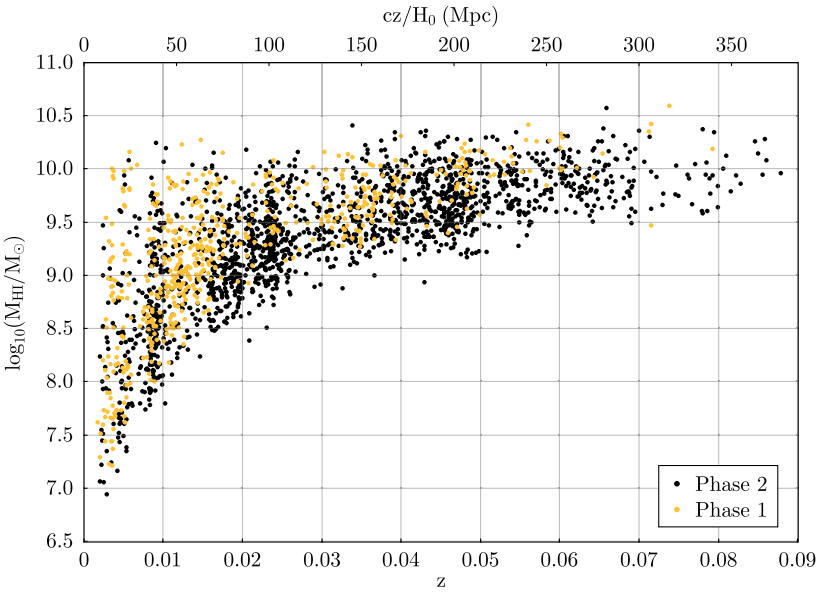


Figure 2. HI mass, $M_{\rm HI}$, plotted against barycentric redshift, z, and Hubble distance, cz/H_0, of the galaxies detected in phase 1 (orange) and 2 (black) of the WALLABY pilot survey. A Hubble constant of $H_0 = 70$ km s^{-1} Mpc^{-1} is assumed in all calculations.

is largely due to improvements in data quality and source finding strategy, although cosmic variance may have contributed to some degree. The number of resolved galaxies with successful kinematic modelling is 109 and 126 for phase 1 and 2 of the pilot survey, respectively. Fig. 2 shows the HI mass of all pilot survey detections as a function of redshift. Galaxies down to $M_{\rm HI} \approx 10^7\ M_\odot$ were detected by WALLABY in the local volume, but WALLABY's detection rate is expected to peak in the redshift range of about 0.03 to 0.04 with a median HI mass of approximately $10^{9.5}\ M_\odot$ (Murugeshan et al. 2024) which is just below the turnover point of the HI mass function (Jones et al. 2018).

3. Public Data Release

The WALLABY team aims to make all data products, including source catalogues, images, spectra and kinematic models, publicly available within 1–2 years of the calibrated data cubes being delivered to the team. Calibrated and mosaicked data cubes of all 8-hour WALLABY footprints are released through the CSIRO ASKAP Science Data Archive† (CASDA; Huynh et al. 2020) immediately after acceptance by the WALLABY team. Data cubes of individual footprints from the WALLABY pilot survey and the full WALLABY survey can be retrieved through their respective project codes of AS102 and AS202. These single-epoch footprints combine all 36 compound beams of ASKAP into a single, large mosaic covering about 30 deg^2. Hence, the resulting FITS data cubes are quite large (about 500 GB) and may be difficult to retrieve and handle without access to adequate storage and computing facilities. Full 16-hour WALLABY sky tiles require two separate 8-hour epochs to be mosaicked together and are not currently supplied by CASDA (except for the six pilot survey fields).

In addition to the calibrated data cubes produced by the ASKAP processing pipeline, the WALLABY team already publicly released value-added data products from phase 1 of the pilot survey in November 2022, including a source catalogue as well as images

† CASDA: https://data.csiro.au/domain/casdaObservation

https://doi.org/10.1017/S174392132500050X Published online by Cambridge University Press

and spectra of 592 HI detections found in the three phase 1 pilot survey fields (Westmeier et al. 2022) as well as kinematic models of 109 spatially resolved galaxies (Deg et al. 2022). A detailed description of the source catalogue and individual data products from this Public Data Release 1 (PDR1) is provided by Westmeier et al. (2022). The data are available either through CASDA or through the Canadian Astronomy Data Centre (CADC).† In addition to direct data access through a web browser, the source catalogue can also be accessed using VO-compliant tools such as TopCat (Taylor 2005) or the PyVO package in Python. More information on data availability and access methods is provided on the official WALLABY website.‡

Value-added data products from phase 2 of the pilot survey were released to the public in September 2024 (Murugeshan et al. 2024), including catalogues, images and spectra of 1760 HI detections and kinematic models of 126 spatially resolved galaxies. Public Data Release 2 (PDR2) for the first time also includes high-resolution data cubes of 80 galaxies from the HI Parkes All-Sky Survey (HIPASS; Meyer et al. 2004; Wong et al. 2006) found within the three phase 2 pilot fields. These are imaged at 12″ resolution to allow for more detailed morphological and kinematical studies, albeit at lower surface brightness sensitivity compared to the standard 30″ images.

Given the nature of the pilot survey, users of the data are advised that several teething issues are still affecting its quality. These include residuals from bright radio continuum sources, interference from ASKAP's on-dish calibrators, residual sidelobes from imperfect deconvolution and flux errors from uncertainties in the primary beam shape and partial deconvolution of faint sources. Please see Westmeier et al. (2022) and Murugeshan et al. (2024) for further details and possible mitigation strategies.

4. Outlook

Full WALLABY survey observations started in May 2023 (following a short observing trial in November 2022) and are currently ongoing. WALLABY was originally expected to be completed within five years, although significant delays during the first two years of observing have occurred which could lead to a slip in the survey completion date. Nevertheless WALLABY is expected to produce the largest interferometric sample of spatially resolved HI observations of galaxies in the local Universe over the next few years in the lead-up to SKA science operations. We plan to have incremental public data releases of value-added data products on an annual basis as survey observations progress to ensure that astronomers around the world have free access to WALLABY catalogues and images after a short period of 1–2 years during which team members have proprietary access to all value-added data.

Acknowledgements

This scientific work uses data obtained from Inyarrimanha Ilgari Bundara / the Murchison Radio-astronomy Observatory. We acknowledge the Wajarri Yamaji People as the Traditional Owners and native title holders of the Observatory site. CSIRO's ASKAP radio telescope is part of the Australia Telescope National Facility (https://ror.org/05qajvd42). Operation of ASKAP is funded by the Australian Government with support from the National Collaborative Research Infrastructure Strategy. ASKAP uses the resources of the Pawsey Supercomputing Research Centre. Establishment of ASKAP, Inyarrimanha Ilgari Bundara, the CSIRO Murchison Radio-astronomy Observatory and the Pawsey Supercomputing Research Centre are initiatives of the Australian

† CADC: https://www.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/en/search/

‡ WALLABY website: https://wallaby-survey.org/data/

https://doi.org/10.1017/S174392132500050X Published online by Cambridge University Press

Government, with support from the Government of Western Australia and the Science and Industry Endowment Fund.

Parts of this research were supported by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013.

References

Aihara, H., Armstrong, R., Bickerton, S., et al. 2018, *PASJ*, 70, S8
Courtois, H. M., Said, K., Mould, J., et al. 2023, *MNRAS*, 519, 4589
Deg, N., Spekkens, K., Westmeier, T., et al. 2022, *PASA*, 39, e059
Deg, N., Palleske, R., Spekkens, K., et al. 2023, *MNRAS*, 525, 4663
Di Teodoro, E. M., Fraternali, F. 2015, *MNRAS*, 451, 3021
Driver, S. P., Norberg, P., Baldry, I. K., et al. 2009, *A&G*, 50, 5.12–5.19
For, B.-Q., Wang, J., Westmeier, T., et al. 2021, *MNRAS*, 507, 2300
For, B.-Q., Spekkens, K., Staveley-Smith, L., et al. 2023, *MNRAS*, 526, 3130
Glowacki, M., Lee-Waddell, K., Deller, A. T., et al. 2023, *ApJ*, 949, 25
Guzman, J. C., Whiting, M., Voronkov, M, et al. 2019, *Astrophysics Source Code Library*, ascl:1912.003
HI4PI Collaboration 2016, *A&A*, 594, A116
Holwerda, B. W., Bigiel, F., Bosma, A., et al. 2023, *MNRAS*, MNRAS, 521, 1502
Hotan, A. W., Bunton, J. D., Chippendale, A. P., et al. 2021, *PASA*, 38, e009
Huynh, M., Dempsey, J., Whiting, M. T., Ophel, M. 2020, *ASP Conf. Ser.*, 522, 263
Jones, M. G., Haynes, M. P., Giovanelli, R., Moorman, C. 2015, *MNRAS*, 477, 2
Kamphuis, P., Józsa, G. I. G., Oh, S.-H., et al. 2015, *MNRAS*, 452, 3139
Kim, S.-J., Oh, S.-H., Wang, J., et al. 2023, *MNRAS*, 519, 318
Koribalski, B. S., Staveley-Smith, L., Westmeier, T., et al. 2020, *Ap&SS*, 365, 118
Liske, J., Baldry, I. K., Driver, S. P., et al. 2015, *MNRAS*, 452, 2087
Meyer, M. J., Zwaan, M. A., Webster, R. L., et al. 2004, *MNRAS*, 350, 1195
Mould, J., Jarrett, T. H., Courtois, H, et al. 2024, *MNRAS*, 533, 925
Murugeshan, C., Kilborn, V. A., For, B.-Q., et al. 2021, *MNRAS*, 507, 2949
Murugeshan, C., Deg, N., Westmeier, T., et al. 2024, *PASA*, in press
O'Beirne, T., Staveley-Smith, L., Wong, O. I., et al. 2024, *MNRAS*, 528, 4010
Reynolds, T. N., Westmeier, T., Elagali, A., et al. 2021, *MNRAS*, 505, 1891
Reynolds, T. N., Catinella, B., Cortese, L., et al. 2022, *MNRAS*, 510, 1716
Reynolds, T. N., Catinella, B., Cortese, L., et al. 2023, *PASA*, 40, e032
Serra, P., Westmeier, T., Giese, N., et al. 2015, *MNRAS*, 448, 1922
Taylor, M. B. 2005, *ASP Conf. Ser.*, 347, 29
Wang, J., Staveley-Smith, L., Westmeier, T., et al. 2021, *ApJ*, 915, 70
Wang, S., Wang, J., For, B.-Q., et al. 2022, *ApJ*, 927, 66
Westmeier, T., Westmeier, T., Kitaeff, S., Pallot, D., et al. 2021, *MNRAS*, 506, 3962
Westmeier, T., Deg, N., Spekkens, K., et al. 2022, *PASA*, 39, e058
Wong, O. I., Ryan-Weber, E. V., Garcia-Appadoo, D. A., et al. 2006, *MNRAS*, 371, 1855
Wong, O. I., Stevens, A. R. H., For, B.-Q., et al. 2021, *MNRAS*, 507, 2905

https://doi.org/10.1017/S174392132500050X Published online by Cambridge University Press