UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

The LOFAR Two-metre Sky Survey: Deep Fields Data Release 2

I. The ELAIS-N1 field

Shimwell, T.W.; Hale, C.L.; Best, P.N.; Botteon, A.; Drabent, A.; Hardcastle, M.J.; Jelić, V.; de Jong, J.M.G.H.J.; Kondapally, R.; Röttgering, H.J.A.; Tasse, C.; van Weeren, R.J.; Williams, W.L.; Bonafede, A.; Bondi, M.; Brüggen, M.; Brunetti, G.; Callingham, J.R.; De Gasperin, F.; Duncan, K.J.; Horellou, C.; Iyer, S.; de Ruiter, I.; Małek, K.; Nair, D.G.; Morabito, L.K.; Prandoni, I.; Rowlinson, A.; Sabater, J.; Shulevski, A.; Smith, D.J.B.; Sweijen, F.

DOI
10.1051/0004-6361/202452930

Publication date
2025

Document Version
Final published version

Published in
Astronomy and Astrophysics

License
CC BY

Link to publication

Citation for published version (APA):
Shimwell, T. W., Hale, C. L., Best, P. N., Botteon, A., Drabent, A., Hardcastle, M. J., Jelić, V., de Jong, J. M. G. H. J., Kondapally, R., Röttgering, H. J. A., Tasse, C., van Weeren, R. J., Williams, W. L., Bonafede, A., Bondi, M., Brüggen, M., Brunetti, G., Callingham, J. R., De Gasperin, F., ... Sweijen, F. (2025). The LOFAR Two-metre Sky Survey: Deep Fields Data Release 2: I. The ELAIS-N1 field. *Astronomy and Astrophysics*, *695*, Article A80. https://doi.org/10.1051/0004-6361/202452930

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:10 Oct 2026

A&A, 695, A80 (2025)
https://doi.org/10.1051/0004-6361/202452930
© The Authors 2025

Astronomy & Astrophysics

The LOFAR Two-metre Sky Survey: Deep Fields Data Release 2

I. The ELAIS-N1 field

T. W. Shimwell[1,2,⋆], C. L. Hale[3,4], P. N. Best[4], A. Botteon[5], A. Drabent[6], M. J. Hardcastle[7], V. Jelić[8], J. M. G. H. J. de Jong[2], R. Kondapally[4], H. J. A. Röttgering[2], C. Tasse[9,10], R. J. van Weeren[2], W. L. Williams[11], A. Bonafede[12,5], M. Bondi[5], M. Brüggen[13], G. Brunetti[5], J. R. Callingham[1,2], F. De Gasperin[5], K. J. Duncan[4], C. Horellou[14], S. Iyer[15], I. de Ruiter[15], K. Małek[16], D. G. Nair[17,18], L. K. Morabito[19,20], I. Prandoni[5], A. Rowlinson[15,1], J. Sabater, A. Shulevski[1,21,15], D. J. B. Smith[7], and F. Sweijen[19]

[1] ASTRON, The Netherlands Institute for Radio Astronomy, Postbus 2, NL-7990 AA Dwingeloo, The Netherlands
[2] Leiden Observatory, Leiden University, PO Box 9513, 2300 RA Leiden, The Netherlands
[3] Astrophysics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK
[4] Institute for Astronomy, School of Physics and Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK
[5] INAF-Istituto di Radioastronomia, Via P. Gobetti 101, 40129 Bologna, Italy
[6] Thüringer Landessternwarte, Sternwarte 5, D-07778 Tautenburg, Germany
[7] Centre for Astrophysics Research, University of Hertfordshire, Hatfield, Herts AL10 9AB, UK
[8] Ruđer Bošković Institute, Bijenička cesta 54, 10000 Zagreb, Croatia
[9] GEPI, Observatoire de Paris, Université PSL, CNRS, 5 Place Jules Janssen, 92190 Meudon, France
[10] Department of Physics & Electronics, Rhodes University, PO Box 94, Grahamstown 6140, South Africa
[11] SKA Observatory, Jodrell Bank, Lower Withington, Macclesfield SK11 9FT, United Kingdom
[12] DIFA – Università di Bologna, Via P. Gobetti 93, 40129 Bologna, Italy
[13] Hamburg Observatory, University of Hamburg, Gojenbergsweg 112, 21029 Hamburg, Germany
[14] Chalmers University of Technology, Onsala Space Observatory, SE-43992, Onsala, Sweden
[15] Anton Pannekoek Institute for Astronomy, University of Amsterdam, Postbus 94249, 1090 GE Amsterdam, The Netherlands
[16] National Centre for Nuclear Research, Pasteura 7, 02-093 Warsaw, Poland
[17] Astronomy Department, Universidad de Concepción, Casilla 160-C, Concepción, Chile
[18] Max Planck Institute for Radio Astronomy, Auf dem Hügel 69, 53121 Bonn, Germany
[19] Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK
[20] Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK
[21] Kapteyn Astronomical Institute, University of Groningen, Landleven 12, 9747 AD Groningen, The Netherlands

Received 8 November 2024 / Accepted 28 December 2024

ABSTRACT

We present the final 6″ resolution data release of the ELAIS-N1 field from the LOw-Frequency ARray (LOFAR) Two-metre Sky Survey Deep Fields project (LoTSS Deep). The 144 MHz images are the most sensitive achieved to date at this frequency and were created from 290 TB of data obtained from 505 h on-source observations taken over 7.5 years. The data were processed following the strategies developed for previous LoTSS and LoTSS Deep data releases. The resulting images span 24.53 square degrees, and using a refined source detection approach, we identified 154 952 radio sources formed from 182 184 Gaussian components within this area. The maps reach a noise level of 10.7 μJy beam^{-1} at 6″ resolution, where approximately half of the noise is due to source confusion. In about 7.4% of the image, our limited dynamic range around bright sources results in a further >5% increase in the noise. The images have a flux density scale accuracy of about 9%, and the standard deviation of offsets between our source positions and those from Pan-STARRS is 0.2″ in RA and Dec for high significance detections. We searched individual epoch images for variable sources, identifying 39 objects with considerable variation. We also searched for circularly polarised sources, achieving three detections of previously known emitters (two stars and one pulsar) whilst constraining the typical polarisation fraction plus leakage to be less than 0.045%.

Key words. techniques: image processing – catalogs – surveys – radio continuum: general

1. Introduction

The European Large Area ISO Survey-North 1 (ELAIS-N1; Oliver et al. 2000) field is one of the most widely studied extra-galactic fields that can be observed by the LOw-Frequency ARray (LOFAR; van Haarlem et al. 2013) at high elevations, corresponding to optimal sensitivity. It has a wealth of data at radio wavelengths (see e.g. Sect. 2.1 of Sabater et al. 2021) and across the rest of the electromagnetic spectrum (see e.g. Sect. 2.3.1 of Best et al. 2023). This extensive dataset allows

⋆ Corresponding author; shimwell@astron.nl

Open Access article, published by EDP Sciences, under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
This article is published in open access under the Subscribe to Open model. Subscribe to A&A to support open access publication.

for the classification of radio sources and the determination of their physical properties. Multi-wavelength efforts have revealed that the vast majority of detectable extragalactic radio sources are star-forming galaxies and active galactic nuclei (AGN; e.g. Smolčić et al. 2017a, Whittam et al. 2022, Best et al. 2023 and Das et al. 2024). Using deep multi-epoch observations from the most sensitive operating radio interferometers, the evolution of these source populations can be probed out to a high redshift ($z > 4$), and aspects such as the time variability or polarisation properties of any detected emission can be examined (e.g. Kondapally et al. 2021, 2022, Duncan et al. 2021, Callingham et al. 2021a, Smolčić et al. 2017b, Whittam et al. 2022, Cochrane et al. 2023 and Piras et al. 2024).

ELAIS-N1 together with three other well-studied extragalactic fields – Boötes (Jannuzi & Dey 1999), the Lockman Hole (Lockman et al. 1986), and the Euclid Deep Field North (Euclid Collaboration: Scaramella et al. 2022), which covers the North Ecliptic Pole (NEP) – form the LOFAR Two-metre Sky Survey Deep Fields (LoTSS Deep; see Best et al. 2023). Each of these fields has been targeted with between 376 and 505 h of LOFAR high band antenna (HBA) observations aiming to reach RMS noise levels on the order of 10–20 μJy beam^{-1} at 120–168 MHz when imaging at an angular resolution of 6″. Initial images of the fields reaching RMS noise levels as low as 17.1 μJy beam^{-1} and using up to 163.7 h of data have already been published in the first data release from LoTSS Deep (Sabater et al. 2021, Tasse et al. 2021, Bondi et al. 2024). These studies were accompanied by descriptions of the techniques used to produce the radio maps (Tasse et al. 2021), and for all fields except the NEP, which was processed later, multi-wavelength catalogues (Kondapally et al. 2021), redshifts (Duncan et al. 2021), and host galaxy properties have been derived (Best et al. 2023 and Das et al. 2024). LoTSS Deep has already facilitated many studies in areas such as radio source populations (e.g. Mandal et al. 2021), AGN (e.g. Mingo et al. 2022, Kondapally et al. 2022, 2023, Calistro Rivera et al. 2024 and de Jong et al. 2024), star formation (e.g. Bonato et al. 2021, Smith et al. 2021, Cochrane et al. 2023), cosmology (e.g. Hardcastle et al. 2021, Gloudemans et al. 2021), galaxy clusters (e.g. Osinga et al. 2021), magnetic fields (e.g. Piras et al. 2024), and even flare stars (e.g. Callingham et al. 2021a).

Excitingly, the data are also being used to probe the low frequency sky at sub-arcsecond resolution by making use of the international LOFAR stations (e.g. Morabito et al. 2022), which were not included in the first LoTSS Deep data release due to data processing limitations but for which data are present in the majority of the observations taken. Including these stations provides significant additional collecting area and sensitivity to compact sources whilst allowing confusion noise limitations to be mitigated. The first high resolution wide-field studies using the LoTSS Deep data have recently been released, with Sweijen et al. (2022) mapping 8 h of Lockman Hole data, de Jong et al. (2024) mapping 32 h of ELAIS-N1 data (RMS of 14 μJy beam^{-1}), and upcoming studies by Escott et al. (in prep.) and de Jong et al. (in prep.) using 8 h and at least 216 h of Bootes and ELAIS-N1 data, respectively. These very high resolution images have a lower surface brightness sensitivity but are complemented by intermediate (∼1.2″) resolution images, such as those from Ye et al. (2024) and de Jong et al. (2024), to allow for the detection of emission from more diffuse objects. Another limitation of LoTSS Deep is the lack of accurate radio spectral information for many of the detected sources, particularly at lower frequencies. Aiming towards rectifying this, a first study by Williams et al. (2021) has been carried out with the LOFAR low band antenna (LBA), and there are plans for deeper LBA studies in the future potentially reaching frequencies as low as 10 MHz (e.g. Groeneveld et al. 2024). Finally, the fields are also being mapped in linear polarisation with Herrera Ruiz et al. (2021) and Piras et al. (2024), searching deep ELAIS-N1 datasets (48 and 176 h respectively) for emission from discrete radio sources, whilst Jelić et al. (2014) and Šnidarić et al. (2023) have examined diffuse polarised emission from our own Galaxy using 8 h and 150 h of ELAIS-N1 data, respectively.

Here we release the final 6″ resolution LoTSS Deep ELAIS-N1 image. In this study, we use all 505 h of data taken for the LoTSS Deep ELAIS-N1 project to form a very deep synthesised image with an RMS noise level of 10.7 μJy beam^{-1} at 144 MHz. This is the most sensitive map to date at this frequency. A significant (≈50%) fraction of the image noise is from source confusion, implying that this image is approaching the sensitivity limit at 6″ resolution. We use the deep 6″ resolution map to characterise image properties, including the flux density scale, astrometric accuracy, dynamic range, and the variation between individual epoch images. We also perform analyses to examine the level of confusion noise, quantify the variability of sources in the field between epochs, and search for circularly polarised emission (linear polarisation will be included in a future study) and for very diffuse emission in the field.

LoTSS Deep is one of several efforts to approximately match the μJy beam^{-1} sensitivity levels at gigahertz frequencies of the deepest radio surveys (e.g. Smolčić et al. 2017a, van der Vlugt et al. 2021, and Heywood et al. 2021) but over a larger area of sky (tens of square degrees) in order to obtain more representative samples of sources and include rarer objects. The MeerKAT International GigaHertz Tiered Extragalactic Exploration (MIGHTEE; Jarvis et al. 2016) survey has similar aims but at gigahertz frequencies, and the first images from that survey have already been released (Heywood et al. 2022 and Hale et al. 2024). LoTSS Deep at 6″ and MIGHTEE at 5″ reach comparable depths, as both are severely limited in the deepest regions of the image by confusion noise. This limitation further highlights the need for higher resolution wide-field studies, such as those by Sweijen et al. (2022), de Jong et al. (2024), Ye et al. (2024), and Morabito et al. (2022), that can potentially probe far deeper by mitigating confusion noise.

In Sect. 2, we provide an overview of the observations, data processing, and cataloguing. In Sect. 3, we examine the quality of the ELAIS-N1 deep image. In Sect. 4, we analyse selected aspects of the dataset to demonstrate its scientific potential and limitations. In Sect. 5, we describe the data release products before summarising in Sect. 6.

2. Observations and data calibration

The ELAIS-N1 field has been observed with the LOFAR HBA 'dual inner' observing mode for a total of 505 h, consisting of 64 individual exposures each using between 54 and 75 independent stations taken over a period of 7.5 years (from 2014 May 19 to 2021 Oct 30) through four different LoTSS Deep observing campaigns. In this observing mode the core HBA sub-stations each act as an independent station and only the inner 24 tiles of the remote stations are used to match the area of the core HBA sub-stations to make the primary beam the same for all stations in the Netherlands. Each observation is bookended by 10 min scans of a calibrator source (always 3C295 and generally also 3C147). An overview of the dataset and resulting images is given in Table 1 and details of the individual observations are provided

in Table S1. We note that here we have already excluded 24 h of LOFAR cycle 0 observations (project LC0_019) which had poor noise levels or data issues (see Sabater et al. 2021 for further details) and 47.4 h of LOFAR cycle 15 observations (project LC15_003) which have poor sensitivity as they were deliberately observed at low elevation due to a different science aim.

The data were processed following the procedure described by Tasse et al. (2021). This approach was also used for the previous public data release of the LoTSS Deep fields (Tasse et al. 2021, Sabater et al. 2021, Kondapally et al. 2021, Duncan et al. 2021 and Best et al. 2023) and is used for the wider area LoTSS (e.g. LoTSS-DR2; Shimwell et al. 2022). The data are calibrated and imaged accounting for effects that are assumed constant throughout the observation as well as those that vary with time and/or position across the field of view (primarily the ionosphere and imperfections in the beam model). For completeness, a brief summary of the processing is given below.

After the observations were conducted the data were flagged, averaged and for more recent observations (those after Sept 2018) they were also compressed by the observatory from 32 bits per float to 12 bits per float (saving approximately a factor of 4 in storage volume with negligible impact on data quality for our purpose). These operations were done using the software packages DP3 (Dijkema et al. 2023), AOFlagger (Offringa et al. 2012) and Dysco (Offringa 2016). The data are then sent to the LOFAR Long Term Archive (LTA)[1] with these particular datasets being stored either at the Forschungszentrum Jülich[2] or the SURF[3] site. These data are presently retrievable through the LTA web interface, except for data from LC2_024 which are hosted elsewhere but are still available for download[4].

The LTA datasets are large (290 TB) and we processed them on compute clusters local to the archives (JUWELS[5] and Spider[6]) to minimise data transport. Firstly the 3C295 calibrator observations corresponding to each target observation were processed through the PreFactor[7] calibrator pipeline (recently migrated to the LOFAR Initial Calibration pipeline; LINC[8]). This derives time-independent corrections for the bandpass, polarisation alignment (offset between XX and YY phase) and clock offsets between different stations. The target data were then processed with the PreFactor/LINC target pipeline which applies the derived calibrator solutions in addition to a single ionospheric Faraday Rotation Measure correction across the field (Mevius 2018) before calibrating them against a sky model derived from the TIFR GMRT Sky Survey alternative data release (TGSS-ADR1; Intema et al. 2017). The target pipeline also performs additional flagging of data contaminated by radio frequency interference and in our case removes the international LOFAR stations (which are not used for this study) and averages the data to 8 s time resolution and 0.0975 MHz spectral resolution (a factor of up to 64 averaging) before Dysco compressing to 16 bits per float (adding negligible additional noise in our case). These averaging parameters were chosen to minimise the data volume whilst keeping time and bandwidth smearing effects tolerable (see e.g. Sect. 3.1 of Shimwell et al. 2019). The PreFactor/LINC calibrator and target pipelines are described in detail by de Gasperin et al. (2019).

Table 1. ELAIS-N1 LoTSS Deep observations as well as the final image and calibrated data.

Coordinates (J2000)	16:11:00 55:00:00
Calibrated data frequency resolution (MHz)	0.0975
Calibrated data frequency range (MHz)	115–177
Calibrated data (calibration solutions) volume (TB)	4.7 (0.1)
Calibrated data time resolution (s)	8
On-source integration time (h)	505
Calibrated data flagged fraction	0.17
Image thermal noise (μJy beam^{-1})	7.5 at 6″ resolution
Image central RMS (μJy beam^{-1})	10.7 at 6″ resolution

Prior to performing direction and time-dependent calibration and imaging of the target field data we first subtracted sources outside the region of interest using the dataset with direction-independent calibration only. This decreases contamination from the sidelobes of these sources and allows us to focus on a smaller region of the sky in subsequent calibration and imaging steps, which lowers the computational cost. Furthermore, a tighter tessellation of facets (corresponding to a smaller patch of ionospheric and beam variations) for which we derive direction-dependent calibration solutions can be used without significantly increasing the number of calibration directions (and degrees of freedom), which helps maintain robust calibration and mitigates against the absorption of unmodelled emission. A similar approach is typically applied to LOFAR HBA low-declination fields where the N–S primary beam extension is very large (e.g. Hale et al. 2019). To choose an appropriate region for this source subtraction, we tested subtracting sources outside of square regions ranging from 4.2° × 4.2° to 7.0° × 7.0° for target observation 798146 (selected since it is one of the best quality observations; see Table S1). For each test region size, a low resolution (around 45″) 25° × 25° image of the observation was created, and sources outside the desired region were subtracted from the *uv*-data using the clean component models derived during the imaging. The subtracted data were then calibrated and imaged using DDF-pipeline[9], which corrects for ionospheric distortions and other direction-dependent or independent errors in the data. The pipeline uses kMS (Tasse 2014 and Smirnov & Tasse 2015) for simultaneous direction-dependent calibration and DDFacet (Tasse et al. 2018) for imaging with these solutions applied. All DDF-pipeline runs were conducted with 45 directions and between 10 000 × 10 000 and 17 000 × 17 000 1.5″ pixels to encompass the area where sources have not been subtracted. The quality of the resulting images was assessed by comparing the number of sources detected within 1.5° of the pointing centre. Here we found that when the region remaining after subtraction was too small (in this case less than 5.0° × 5.0°) the calibration was more susceptible to diverging in some directions (those that contained the least flux density) and the source density on the images was on average lower. For the tests with regions larger than this the image source density was comparable (between

[1] https://lta.lofar.eu/
[2] http://www.fz-juelich.de
[3] https://www.surf.nl/
[4] https://github.com/nudomarinero/AWS_elais-n1_public_data
[5] https://www.fz-juelich.de/en/ias/jsc/systems/supercomputers/juwels
[6] https://doc.spider.surfsara.nl/en/latest/Pages/about.html
[7] https://github.com/lofar-astron/prefactor
[8] https://linc.readthedocs.io/en/latest/
[9] https://github.com/mhardcastle/ddf-pipeline

7500 and 7750) with the 5.83° × 5.83° box providing the largest number of sources.

The direction-independent subtraction of sources outside the central 5.83° × 5.83° box was repeated for all epochs. The sky model derived from the target observation 798146 DDF-pipeline run was then used to derive direction and time-dependent calibration solutions in 45 directions for the other 71 target datasets. This calibration was done by providing DDF-pipeline with an input sky model and facet layout. This approach speeds up the processing significantly as it reduces the number of calibration cycles that is done from 8 to 4 and the number of imaging cycles from 7 to 2 (see Algorithm 1. of Tasse et al. 2021). In this mode the pipeline only performs the following steps: fast direction-dependent and direction-independent calibration from the provided sky model; imaging with calibration solutions to derive an updated sky model; fast and slow direction-dependent calibration using the updated sky model; and final imaging with both fast and slow solutions applied. The computational cost for the various processing steps for all the data on an AMD Rome node with 64 cores was approximately as follows: Prefactor/LINC 300 h; direction-independent source subtraction 1000 h; direction-dependent and independent calibration off a sky model 1500 h; imaging to create an updated sky model 380 h; final direction-dependent fast and slow calibration 1400 h; final imaging 120 h. We note that the imaging steps required processing all the data together but calibration steps were run simultaneously either for individual datasets (LINC/PreFactor and source subtraction) or portions of them (after LINC/PreFactor the data are stored in 1.95 MHz blocks which are independently calibrated by kMS). Assuming that 20 nodes are used for all steps that can run simultaneously and that 1 node is used for imaging the total runtime is approximately 1 month.

Prior to creating the final images we explored a range of different imaging parameters with the intention of creating images with optimal point-source sensitivity and surface brightness sensitivity. To find the appropriate settings we created 1000 different images spanning a wide range of robust parameters (−2.0 to 0.0 in intervals of 0.05) and outer uv-plane tapers (0 to 50 000λ with intervals of 2000λ) using 122–124 MHz data from observation 798146; the full range of robust parameters was not explored because the DDFacet fitting of the synthesised beam is somewhat unstable at robust parameters above 0 due to the highly non-Gaussian shape of the synthesised beam. In order to test the imaging parameters, we only produced Stokes V images, which allowed us to study just the thermal noise. This removed the need for deconvolution, eliminated any influence from confusion noise, and mitigated the dynamic range limitations associated with bright Stokes I sources. For each of the 1000 images we calculated the sensitivity (root mean square noise, $\sigma_{\rm RMS}$) as well as the brightness temperature sensitivity ($\frac{\lambda^2}{2k_B\Omega_{\rm bm}}\sigma_{\rm RMS}$ where λ is the wavelength, $\Omega_{\rm bm}$ is the synthesised beam solid angle and k_B is the Boltzmann constant). The measured sensitivities as a function of resolution (derived by DDFacet fitting the Point Spread Function, PSF, with a Gaussian) are shown in Fig. 1 where the values are scaled to expectations from the full 505 h ELAIS-N1 dataset. We also show the expected level of source confusion, which we define as the anticipated flux density level at which there is one source per 10 resolution elements (see Sect. 4.2 for details on the confusion noise). The parameters we have chosen for imaging (marked on the figure) at 6″ resolution are a compromise between resolution and sensitivity.

The final 6″ resolution Stokes I image is shown in Fig. 2. The sensitivity at the field centre of the image is 10.7 μJy beam^{-1} where approximately half of this is due to confusion noise – see Sect. 4.2.

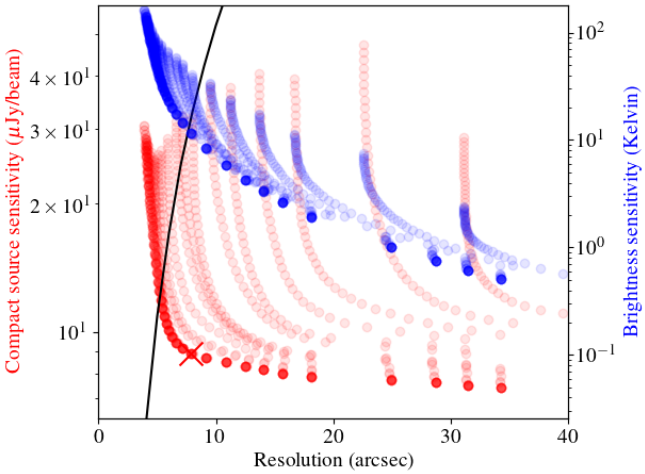

Fig. 1. Expected compact source sensitivity (red) and brightness sensitivity (blue) as a function of imaging resolution for the 122–124 MHz data from observation 798146. The resolution is varied by altering the visibility weightings with robust and tapering settings. The black curve shows the anticipated confusion noise (defined as 1 source per 10 resolution elements). The red cross shows the location of the 6″ resolution imaging parameters we have used (robust = −0.5, no taper). Only a small dataset was used for this analysis to limit the computational cost.

2.1. Source cataloguing

We use PyBDSF (Mohan & Rafferty 2015; version 1.10.3) for the detection and characterisation of sources within the image. PyBDSF was also used in the previous LoTSS Deep data release (Sabater et al. 2021; Tasse et al. 2021) and the wider area LOFAR surveys (Shimwell et al. 2019, 2022; de Gasperin et al. 2023). However, the ELAIS-N1 LoTSS Deep image from this release has an especially high density of sources and is severely confusion noise limited (e.g. Fig. 2). To account for this we refine the set up of PyBDSF compared to previous studies in order to accurately describe the noise in the image, optimise the number of sources detected and improve the characterisation of complex emission.

In our previous setup PyBDSF calculated an RMS map using a 2D sliding box across the image. This map was then used to identify (and fit) islands of contiguous emission above a threshold level. With this method, sources are not removed from the image prior to the calculation of the RMS map. Therefore, for images with a high source density the RMS map is elevated compared to if sources were removed from the image prior to the calculation of the RMS map. This will affect the detection threshold across the image and, as such, will reduce the completeness of the resulting source catalogue. To mitigate this, we developed a multi-step PyBDSF process where we first derive an RMS map that better reflects the image noise and then supply this RMS map to PyBDSF to generate the catalogue for this work. We detail this process in Appendix A.

Our final catalogue contains 154 952 sources across the image formed from 182 184 Gaussian components. For over 50% of these sources this is their first detection at radio wavelengths. The catalogue spans 24.53 square degrees, within the 30% level of the power primary beam, and extends down to a peak detection significance threshold of $4\sigma_{\rm thresh}$ and an island

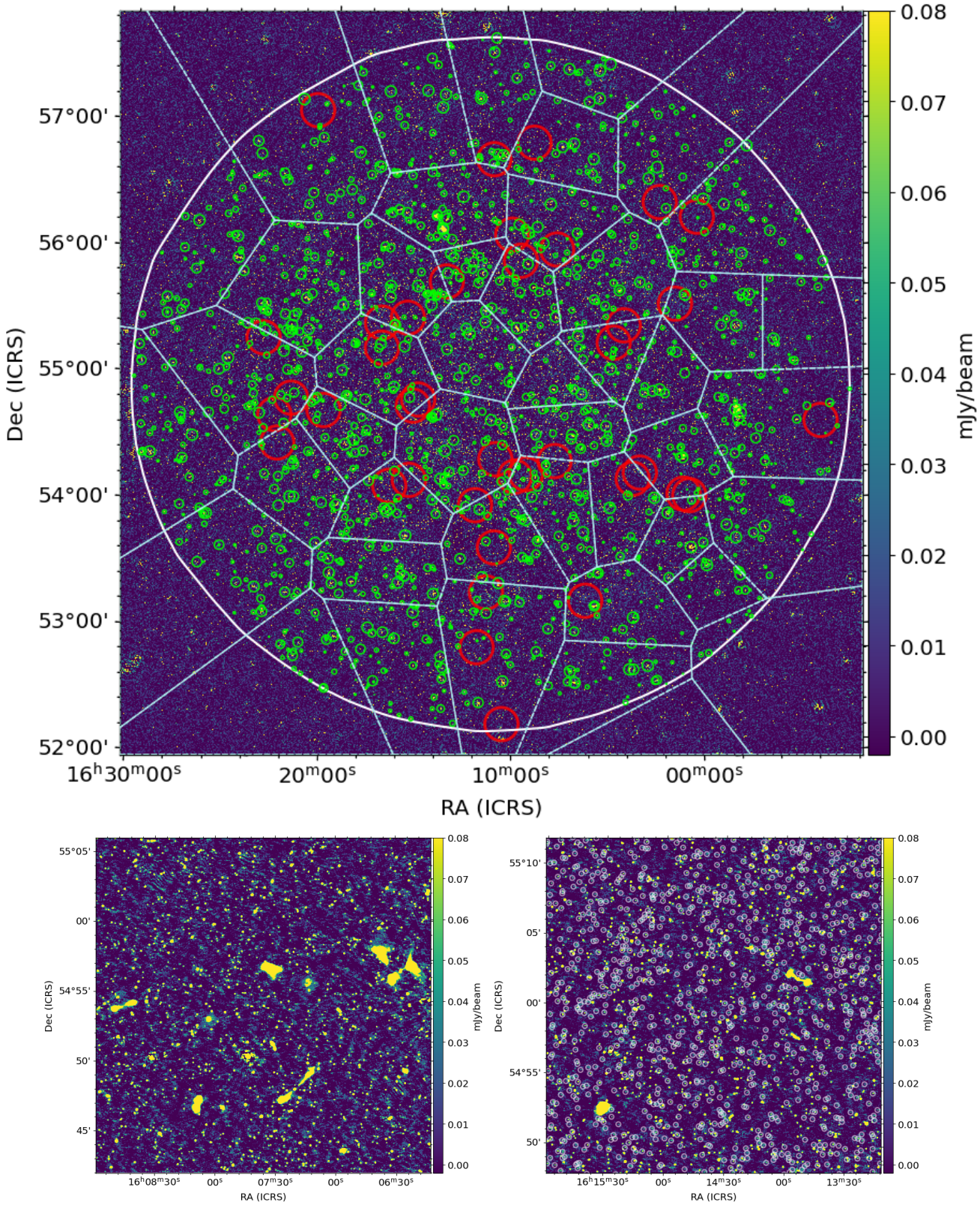

Fig. 2. The final 6″ resolution LoTSS Deep image of ELAIS-N1. The top panel shows the full depth (10.7 μJy beam^{-1}), full area (5.83° × 5.83°) image. The 30% level of the power primary beam is shown in white (approximately a circle of radius 2.8°) and encompasses 24.53 square degrees. The boundaries of the 45 facets used for direction-dependent calibration are shown with light blue lines. The green circles show regions where the noise is limited by the dynamic range (see Sect. 3.4) and the larger red circles show the locations of the most variable sources identified in the field (see Sect. 4.3). The bottom panels show two example 0.4° × 0.4° regions in more detail to highlight the source density (over 10 000 per square degree) and the variation in morphology of detected sources. The grey markers on the bottom right panel show sources in the new ELAIS-N1 catalogue that were not detected by Sabater et al. (2021).

boundary threshold of $3\sigma_{\rm isl}$ (using the PyBDSF parameters `thresh_pix` and `thresh_isl` respectively).

2.2. Source detection threshold

In addition to refining the source cataloguing procedure we also revisited the detection threshold. In previous LoTSS and LoTSS Deep data releases we used a $5\sigma_{\rm thresh}$ peak threshold for cataloguing. Given that in LoTSS Deep fields we are able to match a high fraction of radio sources with optical counterparts (achieving >97% identification rate in LoTSS Deep; see Kondapally et al. 2021), we can ascertain whether radio detections are likely to be genuine at a given detection threshold. We therefore investigated how varying our detection threshold influenced the number of spurious sources introduced.

Source extraction down to $4\sigma_{\rm thresh}$ using the refined PyBDSF procedure results in ≈30% more sources than using the old procedure with a $5\sigma_{\rm thresh}$ threshold. The increase in numbers is due to both the lower detection threshold and a higher completeness achieved at $5 < {\rm S/N} \leq 6$ due to the deeper RMS map (where we define S/N as the catalogued integrated flux density, S_I, divided by the integrated flux density uncertainty, σ_{S_I}). If we were to use a $3\sigma_{\rm thresh}$ threshold this would increase the source number by a further ≈20%. To determine whether it is appropriate to use a threshold below $5\sigma_{\rm thresh}$, we examined the optical counterparts of the radio sources, as it allowed us to ascertain which radio detections are likely genuine. To do this, we first created a very deep radio catalogue by repeating the cataloguing procedure in Sect. 2.1 but using a lower detection limit of `thresh_isl=2.0` and `thresh_pix=3.0`. We then considered only radio sources with major axis sizes <10″ from the catalogue (25% of $3 < {\rm S/N} \leq 4$ sources, 19% of $4 < {\rm S/N} \leq 5$ sources and 16% of $5 < {\rm S/N} \leq 6$ sources), as these are more suitable for the likelihood ratio (LR) cross-matching analysis which determines the ratio of the probability that a optical source is associated with a particular radio source, to the probability that it is unrelated (e.g. Sutherland & Saunders 1992; Smith et al. 2011). We used the multi-wavelength catalogue and the LR method, both as described by Kondapally et al. (2021), to determine the multi-wavelength cross-matches for these radio sources.

In Fig. 3, we plot the fraction of sources that have a robust cross-match (defined based on the LR value that achieves a completeness and reliability of >99%; see Kondapally et al. 2021) as a function of S/N. The colour bar corresponds to the number of sources in each bin. The blue horizontal line corresponds to the overall cross-match rate for the full radio sample of ∼96%. This is less than 100% and is explained by the fact that the host galaxies for a small fraction of the sources will be too faint to be detected and hence will be missing from the multi-wavelength catalogue, whilst some radio sources may form part of a larger, more complex multi-component source that would not be expected to have a cross-match based on the use of the LR method alone. We find that at high S/Ns the cross-match rate is close to the blue line and then decreases below S/N ∼ 5; this is caused by some of the additional lower significance sources not being genuine radio sources. Based on this analysis, we find that at $4 < {\rm S/N} \leq 5$, around 2–3% of the radio sources are likely not genuine. As the S/N decreases further down to S/N ∼ 3, the fraction of matches declines more rapidly, and we find that a considerable fraction (∼10%) of the additional sources detected are likely not genuine, which would add significant contamination to the catalogue. We therefore adopted the use of a $4\sigma_{\rm thresh}$ detection threshold for source extraction. This allowed us to catalogue 30% more sources than our previous $5\sigma_{\rm thresh}$ approach and the vast majority of these additional sources we believe are genuine detections of faint radio sources. We do expect a small increase in the fraction of non-genuine sources by detecting down to $4\sigma_{\rm thresh}$ compared to $5\sigma_{\rm thresh}$ and therefore urge caution when conducting studies that include these faint sources. We also note that our upcoming efforts for visual classification and inspection will help identify the spurious sources (see Williams et al. 2021; Kondapally et al. 2021).

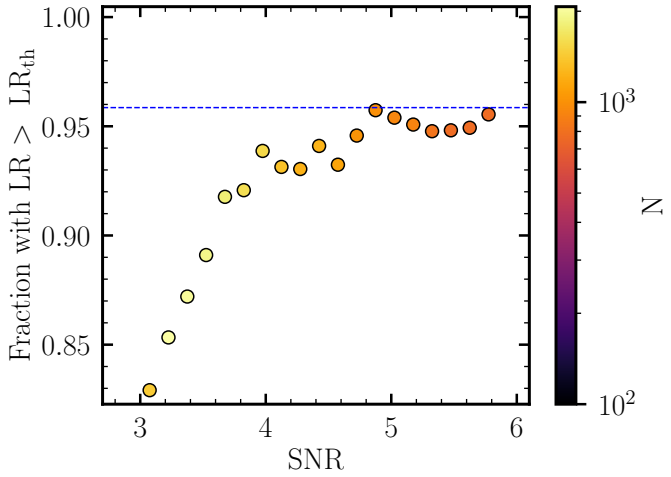

Fig. 3. Fraction of small (major axis < 10″) radio sources with a likelihood ratio match as a function of the S/N (S_I/σ_{S_I}). The colour of the points corresponds to the number of sources in each S/N bin. The blue horizontal line corresponds to the overall LR identification rate obtained. We find that lowering the detection threshold to a peak detection significance of $4\sigma_{\rm thresh}$ results in a large fraction (∼30%) of additional, genuine radio sources compared to adopting a $5\sigma_{\rm thresh}$ threshold. However, this also leads to a small increase in false detections (by ∼2–3%).

3. Image quality

The processing strategy used for this data release follows the same approach used for the previous ELAIS-N1 data release in Sabater et al. (2021), with the exception of the additional source subtraction we performed to remove sources outside the field of view. However, the amount of data used is approximately three times larger (505 h compared to 163 h) and due to ionospheric activity and telescope performance we expect substantial variations in quality between different observations. Below we assess several key image quality properties and, where possible, quantify appropriate uncertainties on these for this data release.

3.1. Source extensions

Slightly extended sources and point-like sources cannot be definitively distinguished in our images due to both measurement errors and imperfections in the image quality. For example, for a source to have the exact shape of the restoring beam it must be represented by a single pixel in our clean component model image. However, even for isolated seemingly compact sources we find that the flux density on the clean component model image is generally distributed between many neighbouring pixels. Typically, high S/N sources only have about 20–50% of their integrated flux density at the position of the peak pixel whereas low S/N sources generally have a higher fraction of their integrated flux density in a single pixel. This is likely

a consequence of only deconvolving to a certain depth. Furthermore, in our images the blurring of sources noticeably increases as a function of distance from the field centre. This is primarily due to time and bandwidth smearing but also the outer facets are larger and thus have greater ionospheric variations within the facet leading to poorer calibration. During the self-calibration this phase blurring is difficult to rectify because it is present in the sky models being used for the calibration.

Thus, to establish a criterion to distinguish unresolved from extended sources we must account for all sources being inherently blurred. To do this we first select and characterise a sample of 45 676 isolated seemingly compact sources from the full catalogue of 154 953 sources. We define these sources as those that are more than 15″ from their nearest neighbour, have PyBDSF code 'S' corresponding to being fit by a single Gaussian, a major axis below 15″, lie in the deconvolution mask and within 30% of the power primary beam and have $\frac{S_I}{\sigma_{S_I}} > 5$. Firstly we plotted the ratio of the integrated flux density, S_I, to peak brightness (S_P) versus distance from the pointing centre for this population and found the median $\frac{S_I}{S_P}$ in different radial bins. From these median values we derived a best fit of $\frac{S_I}{S_P} = 1.038 \times e^{0.094 \times D}$ where D is the distance from the field centre in degrees. We then modified all S_P values using this function (i.e. multiplying by $1.038 \times e^{0.094 \times D}$), which on average corrects for the offset in $\frac{S_I}{S_P}$ from unity at the field centre as well as the radial dependence, thus giving a more constant $\frac{S_I}{S_P}$ for all sources. Finally, we attempted to define a distinction between extended and unresolved sources using our modified catalogue and the behaviour of $\frac{S_I}{S_P}$ (a proxy for source extension) versus $\frac{S_I}{\sigma_{S_I}}$.

For our defined population of seemingly compact sources in our modified catalogue, the distribution of $\frac{S_I}{S_P}$ at a given $\frac{S_I}{\sigma_{S_I}}$ is skewed with an excess of sources at high $\frac{S_I}{S_P}$ values. Furthermore, as $\frac{S_I}{\sigma_{S_I}}$ increases, the number of sources significantly decreases, which makes it challenging to fit the distribution of $\frac{S_I}{S_P}$. Both of these aspects are expected even for a population of perfect compact sources with the skewness arising from the existence of a correlation between uncertainties. Specifically, if, due to measurement errors, the recorded S_I is higher than the real source S_I then the S_P will also likely be higher than its true value. To somewhat mitigate these issues and attempt to characterise the behaviour of seemingly compact sources in our images we conducted a suite of simulations.

We generated mock images by injecting a population of sources with flux densities taken from a given distribution of integrated flux densities and morphologies derived from our real CLEAN component models (from our DDFacet model images) of the seemingly compact sources identified using the criteria above. To do this the CLEAN component models were injected into a blank image and scaled using the distribution of integrated flux densities taken from the Mandal et al. (2021) counts above 0.43 mJy beam^{-1} and extended to lower flux density (1 μJy beam^{-1}) using the number counts from the T-RECS (Bonaldi et al. 2023) simulations (see Sect. 4.2). The CLEAN component model used for a particular injection was randomly selected from all the available models of sources that are close in integrated flux density to the source we are injecting. We then scaled the CLEAN component models to precisely match the intended integrated flux density. The simulated image containing all the deconvolution models was then convolved with the restoring beam (6″) and added to the PyBDSF residual image to produce a mock image that we then catalogued using PyBDSF following the procedure in Sect. 2.1. To reproduce the radial effects observed in the real images we inject sources at the same radial distance as they had in the real observations. We then again derive a function that describes the offset and radial dependence in $\frac{S_I}{S_P}$ for the simulated image and modify the catalogued S_P values to on average correct for this.

Repeating this procedure a number of times (we performed 20 runs) provides us with a large number of sources and allows us to better constrain the distribution of $\frac{S_I}{S_P}$ for seemingly compact sources at a given $\frac{S_I}{\sigma_{S_I}}$. An example histogram showing the distributions of simulated sources together with the seemingly compact sample and the full LoTSS sample for a moderate $\frac{S_I}{\sigma_{S_I}}$ is shown in Fig. 4. It is clear from this example that separating unresolved and extended sources cannot be done definitively. We chose to separate sources where the survival fraction (the complement of the cumulative distribution function) of the simulated seemingly compact sources was a factor of 5 times lower than the survival fraction of the full LoTSS catalogue (as shown in the figure, comparable values are found if we use the real seemingly compact source sample instead of the simulated one). This implies that sources with $\frac{S_I}{S_P}$ equal to or exceeding this value are five times more likely to be associated with the tail of the full LoTSS catalogue $\frac{S_I}{S_P}$ rather than the simulated seemingly compact one. This is shown as a function of $\frac{S_I}{\sigma_{S_I}}$ in the bottom panel of Fig. 4, and the best fit sigmoid function to this is

$$R_5 = 0.10 + \left(\frac{0.93}{1 + \left(\frac{S_I/\sigma_{S_I}}{14.95} \right)^{3.99}} \right). \quad (1)$$

Using these constraints we find that for the full LoTSS Deep catalogue where a radial dependence correction has been applied (see above) a total of 14 025 (9%) out of 154 952 sources are extended with the rest being unresolved at our resolution.

We further validate our unresolved criteria through comparison with the 0.3″ resolution de Jong et al. (2024) catalogue. At this high resolution the surface brightness sensitivity is low (see e.g. Fig. 1) but the sensitivity to compact sources, at 14 μJy beam^{-1}, is only 30% higher than in our catalogue. We therefore expect de Jong et al. (2024) to be primarily sensitive to compact sources and this is demonstrated in Fig. 4 (see also e.g. Sweijen et al. 2022). Here only 11% of sources crossmatched (within 1″) between the full de Jong et al. (2024) catalogue and our full catalogue are deemed extended at 6″ by our criteria, where almost all of these possibly extended sources have high S_I/σ_{S_I}. If we only examine the fainter sources in our catalogue ($S_I/\sigma_{S_I} < 15$) that are detected by de Jong et al. (2024) then just 0.3% of these are classified as extended at 6″ resolution by our criteria. This gives confidence in our criteria to separate unresolved from extended sources. The comparison also highlights that large numbers of the sources we define as unresolved are highly resolved by LOFAR when utilising the full ILT 0.3″ resolution. For example, only 22% of sources with $S_I/\sigma_{S_I} > 15$ in our catalogue are detected at 0.3″ by de Jong et al. (2024) with the rest being undetected likely due to surface brightness sensitivity limits. If we remove sources classified as extended at 6″ resolution this increases to 35% indicating that our unresolved criteria only slightly helps select sources that are still compact at 0.3″.

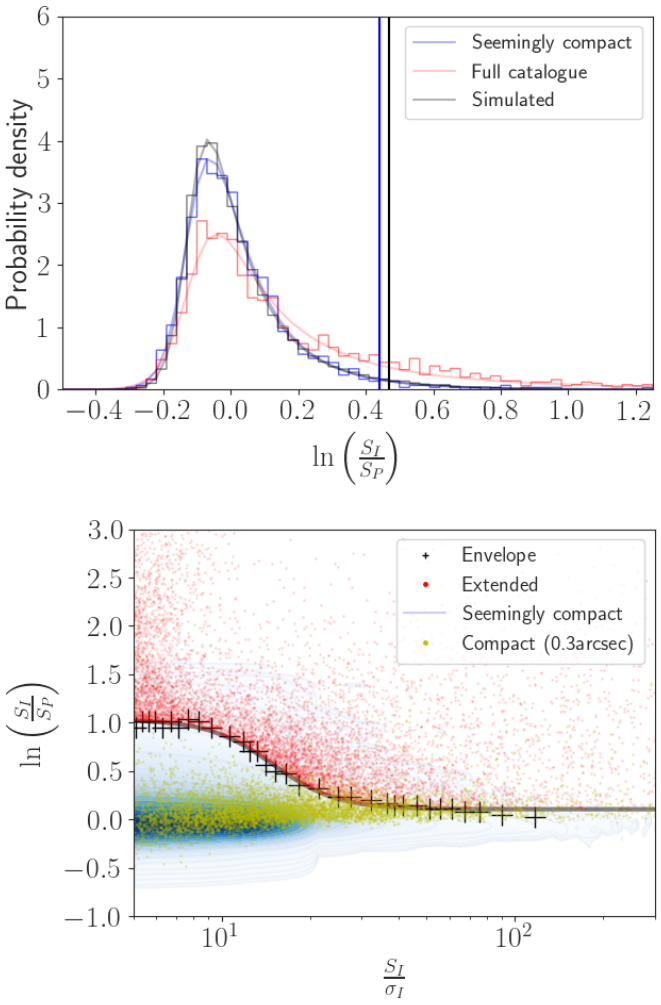

Fig. 4. Separation of extended and unresolved sources using the ratio of integrated flux density to peak brightness. Top: Histograms showing the probability density (bin width 0.03) of $\ln\left(\frac{S_I}{S_P}\right)$ for sources with $15.8 < \frac{S_I}{\sigma_{S_I}} < 17.3$. This corresponds to 999 seemingly compact LoTSS sources (blue), 1673 LoTSS sources in the full catalogue (red) and 35 286 sources in our simulations of seemingly compact sources (grey) described in Sect. 3.1. The fits to the histograms use a normal-inverse Gaussian distribution and are shown with faint solid lines. The vertical lines show the cutoffs we have chosen to separate unresolved from extended sources, these are defined as the location where the complement of the Cumulative Distribution Function for the full catalogue is equal to five times that of the simulated seemingly compact (black) or seemingly compact source (blue) distributions. Hence sources with $\ln\left(\frac{S_I}{S_P}\right)$ above this value are five times more likely to lie in the tail of the red curve compared to the black or blue curves. We use this definition as a boundary to separate unresolved from extended sources. Bottom: Separation values for unresolved and extended sources derived from the simulations as a function of $\frac{S_I}{\sigma_{S_I}}$ are shown with black crosses which have been fit with a sigmoid function (grey). The blue density contours and colour-scale show the seemingly compact LoTSS source population, whereas the red points show the locations of the sources in the full LoTSS catalogue that are extended according to the criteria we have used. The yellow points show the locations of sources in our catalogue that are also detected by de Jong et al. (2024) at 0.3″ resolution.

3.2. Flux density scaling and uncertainty

In Sabater et al. (2021) the flux density scale of the ELAIS-N1 LoTSS Deep DR1 image was carefully examined using auxiliary measurements from VLSSr at 74 MHz (Lane et al. 2014, TGSS-ADR at 150 MHz (Intema et al. 2017), 6C at 151 MHz (Hales et al. 1990), the ELAIS-N1 GMRT survey at 325 MHz (Sirothia et al. 2009), WENSS at 325 MHz (Rengelink et al. 1997), the ELAIS-N1 GMRT survey at 610 MHz (Garn et al. 2008), NVSS at 1.4 GHz (Condon et al. 1998) and FIRST at 1.4 GHz (Becker et al. 1995). The procedure used by Sabater et al. (2021) and the large amount of auxiliary data in the region allowed them to achieve an uncertainty of the flux density scale of their LOFAR maps of 6.5% in the central region. This is better than typically obtained with the LOFAR HBA where the accuracy is generally limited to 10–20%. This limitation is thought to be partly a consequence of inaccuracies in the LOFAR beam model as it is found that transferring calibration solutions derived from calibrators (whose models are on the Scaife & Heald 2012 flux density scale) does not necessarily result in an accurate flux density scale of the target field (e.g. Hardcastle et al. 2016 and Shimwell et al. 2022).

To again obtain a high accuracy in the flux density scale we simply align our flux density scale with the catalogue of Sabater et al. (2021). To do this, from our catalogue we first select isolated (those without a nearest neighbour within 15″), unresolved (defined in Sect. 3.1) sources that have PyBDSF code 'S'. This leaves 97 678 sources (which we refer to throughout as the unresolved, isolated sources in our catalogue) to cross-match with the Sabater et al. (2021) catalogue. Of these, two sources match (within 5″) multiple sources in the Sabater et al. (2021) catalogue and are discarded. If we then do a further cut to remove sources not detected in the Sabater et al. (2021) catalogue and also low significance sources in either catalogue ($S_I/\sigma_{S_I} < 10$), then this reduces the number of cross-matched sources to 18 390. We find a median ratio of 0.89 between the integrated flux densities from Sabater et al. (2021) and those from the new catalogue (differences on this scale are typical and reflect LOFAR HBA flux density scale calibration inaccuracies). Our map and catalogue are scaled by this value in order to put them on the same flux density scale as Sabater et al. (2021). This flux density scale is also consistent with de Jong et al. (2024), and after removing sources detected in either image at $S_I/\sigma_{S_I} < 10$, the median of our integrated flux density measurements over those from de Jong et al. (2024) is 0.97. We note that due to time and bandwidth smearing as well as calibration errors there are larger variations in the peak brightness (see Sect. 3.1 and de Jong et al. (2024) Sect. 6.2).

For completeness, and because new auxiliary data are available, we again attempt to verify our flux density scale and further assess its accuracy. Instead of cross-matching our catalogue with all possible radio wavelength catalogues covering the region, we only used the deeper, higher resolution catalogues for cross-matching. These are the 54 MHz LoLSS catalogue (see de Gasperin et al. 2023 and de Jong et al. 2024), as well as the catalogues from the 325 MHz (Sirothia et al. 2009) and 610 MHz (Ocran et al. 2020 and Ishwara-Chandra et al. 2020) GMRT ELAIS-N1 surveys and the 1.4 GHz FIRST survey (Becker et al. 1995). The LoLSS catalogue used was not available at the time of the Sabater et al. (2021) study and was created by de Jong et al. (2024) who processed a LoLSS dataset covering ELAIS-N1 using the standard data processing strategy for that survey (see de Gasperin et al. 2023). The flux density scales applied during the creation of the catalogued images that we compare with is consistent with the Scaife & Heald (2012) flux density scale, with the exception of the 325 MHz catalogue. To correct the 325 MHz catalogue we scaled it according to the 3C286 flux densities listed by Scaife & Heald (2012) compared

to the values for this calibrator listed by Sirothia et al. (2009). We then created a cross-matched catalogue using the appropriate cross-matching radius given in Table 3 of Sabater et al. (2021) for comparison surveys in common. For LoLLS we used a 6″ cross-matching radius and for the 610 MHz GMRT survey we used 10″. Only 63 sources exist in the LoTSS Deep catalogue that are detected at all frequencies covered by the auxiliary catalogues (if we were to remove sources based on the unresolved, isolated criteria given above we would be left with just 7 sources so we do not apply those cuts here).

For each of these sources we then fitted only the measurements in the auxiliary catalogues using a second-order polynomial to describe the log-log spectrum over frequency (see Fig. 5) whilst taking into account the errors on the various flux density measurements by repeating the fitting a number of times but with flux densities drawn from a Gaussian distribution centred on the measured source flux density with a width reflecting the uncertainty. For these we assumed a 5% flux density scale error for FIRST (White et al. 1997), 6% for LoLSS (de Gasperin et al. 2023), and 10% for both GMRT ELAIS-N1 surveys (Chandra et al. 2004). During this process we found the spectrum of six sources was poorly fit by a 2nd order polynomial and excluded these from the remainder of the analysis. We also note that the Sirothia et al. (2009) values were found to be systematically high by 25%, so these were scaled down – it is mentioned by Sirothia et al. (2009) that their integrated flux densities were about 30% higher than WENSS, which is consistent with the offset we find. The fitted polynomials were subtracted from the source measurements to give residuals for each fit at each frequency. From these residuals we found that the LoTSS Deep flux densities on average agree with the fitted polynomial giving confidence in the flux density scale but there is a standard deviation of 9%, which reflects an upper bound on the uncertainty. This derived uncertainty is comparable to what was found by Sabater et al. (2021) for the previous LoTSS Deep data release. We note that our derived uncertainty is influenced by the uncertainty used for the comparison surveys, for example our uncertainty would increase to 11% if we were to increase the errors on the comparison surveys to 10%, 10% and 15% for FIRST, LoLSS and the GMRT surveys respectively.

3.3. Astrometric accuracy

During the data processing we refine our astrometry by aligning each individual facet used for calibration with Pan-STARRS DR2 (Flewelling et al. 2020) following the procedure described in Shimwell et al. (2019). This results in a typical shift per facet of magnitude $0.44 \pm 0.20''$ in RA and $0.44 \pm 0.16''$ in Dec where these corrections are applied to the *uv*-data by DDFacet when constructing the final image. To validate the astrometric precision of our final catalogues we again use Pan-STARRS DR2, which has a typical astrometric accuracy within 0.03″ (see e.g. Magnier et al. 2020 and Makarov et al. 2017). To do this we construct a cross-matched catalogue by performing a simple nearest neighbour cross-match between the unresolved, isolated sources in our catalogue and the mean Pan-STARRS DR2 epoch positions. In the cross-matching, we allowed up to 1″ separation and only used Pan-STARRS DR2 sources that are detected in more than five single-epoch images to eliminate spurious detections.

Using all sources in this cross-matched catalogue we first examined the offset between the optical and radio positions as a function of S_I/σ_{S_I}. Flux density bins were constructed such that they each contained 1000 sources and histograms of the RA and Dec offsets within each bin were fit with a Gaussian function.

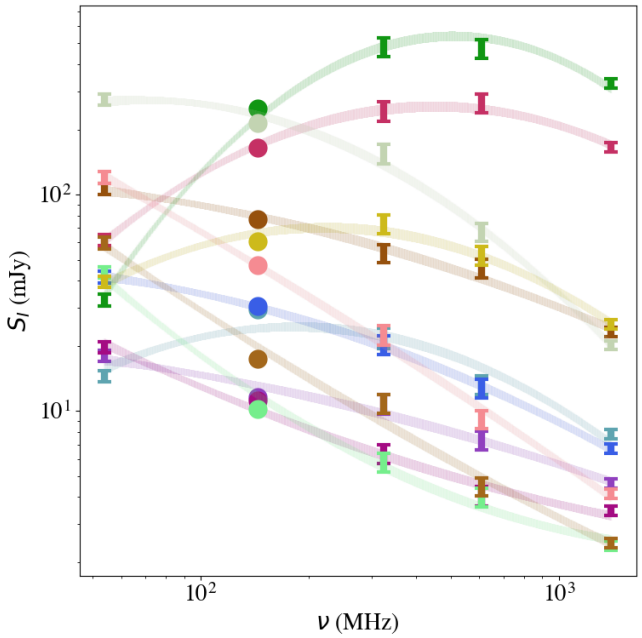

Fig. 5. Selection of 12 of the 55 sources used to verify the deep ELAIS-N1 flux density. Each source is shown in a different colour. The circles show the LoTSS Deep measurements, and the error bars show the LoLSS, GMRT 325 MHz, GMRT 610 MHz, and FIRST measurements with the associated error in the flux density scale. The lines show the derived second-order polynomial fits: $\log S_I(\nu) = \log S_0 + A \log(\nu/\nu_0) + B \log(\nu/\nu_0)^2$, where S_0, A and B are the best-fit parameters. The line thickness shows the uncertainty in the fit.

The centre of the Gaussian function indicates the accuracy of our alignment with Pan-STARRS DR2 whilst the standard deviation is scatter due to, for example, signal-to-noise effects and calibration imperfections. We found that for high significance ($S_I/\sigma_{S_I} > 15$) detections the measured standard deviation is 0.2″ or less. For lower significance detections ($S_I/\sigma_{S_I} < 5$), where signal-to-noise effects play a larger role, this increases to 0.5″. The width of this fitted Gaussian as a function of S_I/σ_{S_I} can be described by $\sigma_{\mathrm{RA}} = \sqrt{\left(\frac{1.83}{S_I/\sigma_{S_I}}\right)^2 + 0.17^2}$ and $\sigma_{\mathrm{Dec}} = \sqrt{\left(\frac{1.40}{S_I/\sigma_{S_I}}\right)^2 + 0.17^2}$. The centre of the fitted Gaussian (i.e. the systematic positional offset) has a less clear dependence on S_I/σ_{S_I} but for the previously defined high significance source sample it was found to be 0.04″ in RA and 0.03″ for Dec. For comparison with our measurements we performed a simple simulation where a population of perfect point-like sources were injected into the PyBDSF residual image and this simulated image was catalogued using PyBDSF. Here we again used the source counts derived from observations by Mandal et al. (2021) but extended to lower flux densities using the T-RECS simulations (Bonaldi et al. 2023). In the simulations we found comparable curvature in the offsets as a function of S_I/σ_{S_I} but far lower offsets at high S_I/σ_{S_I}. The results of the simulation together with the measurements from our images are shown in Fig. 6.

We also examined the offsets as a function of RA and Dec using the high significance source sample from the matched Pan-STARRS catalogue. The width of each RA or Dec bin was chosen so that each bin contained 200 sources. As shown in Fig. 7 the variations in the centres and standard deviations of fitted Gaussians are small although the outermost regions tend

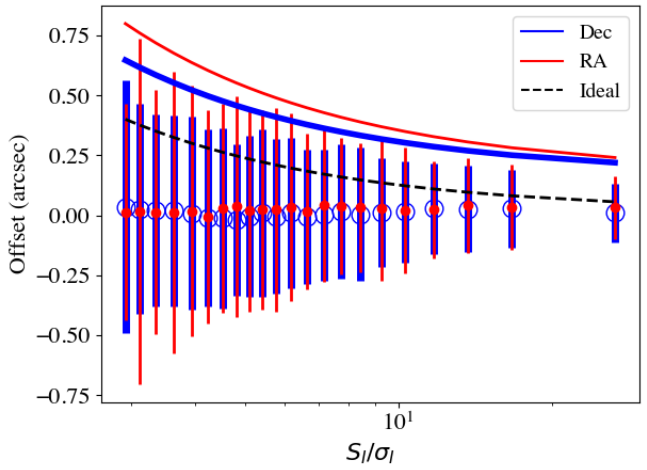

Fig. 6. Right ascension (red) and Dec (thick blue) astrometric offsets as a function of S_I/σ_{S_I}. Each flux density bin was chosen to contain 1000 sources and the histograms of the distribution of RA or Dec offsets between our catalogue and Pan-STARRS DR2 were fit with a Gaussian function. The points and error bars show the centre of the fit and ± the standard deviation respectively. Curves fitting the derived standard deviations of RA and Dec as a function of S_I/σ_{S_I} are also shown with the same colour scheme. The black dashed curve shows the behaviour observed through a simple simulation of perfectly compact sources injected into our residual image and characterised with PyBDSF.

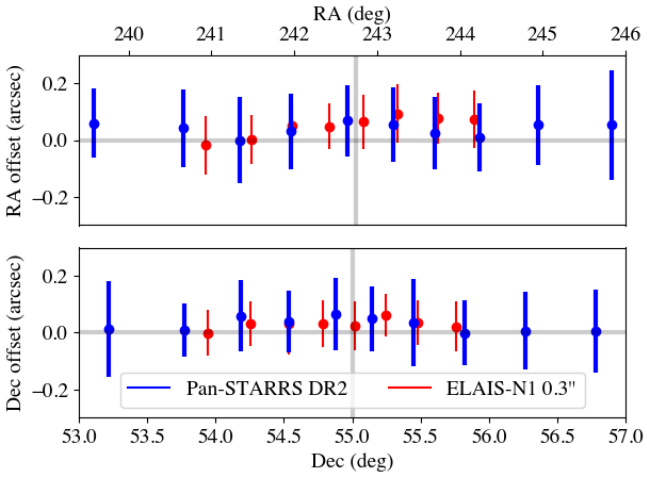

Fig. 7. Right ascension and Dec offsets between our catalogued positions and Pan-STARRS DR2 (thick blue) and the ELAIS-N1 0.3″ image (red) from de Jong et al. (2024) as a function of RA and Dec respectively. The error bars show ± the standard deviation of a Gaussian fitted to histograms of the offsets and the points show the centre of the fitted Gaussian. The grey vertical line shows the field centre and the grey horizontal line shows 0 offset.

to have larger standard deviations. Finally, for completeness we also repeated this analysis but comparing with the 0.3″ resolution LOFAR image from de Jong et al. (2024). In this analysis we used the same cross-matching procedure as used for Pan-STARRS DR2 and for consistency applied a $S_I/\sigma_{S_I} > 15$ cut to our catalogue and kept 200 sources per RA or Dec bin. There is again good astrometric alignment and from all sources in the cross-matched catalogue we find the fitted Gaussian centres of RA and Dec offsets to be 0.05″ and 0.03″ respectively. The derived standard deviations of 0.11″ in RA and 0.10″ in Dec were even smaller than found through the comparison with Pan-STARRS DR2. The results from this cross-matching are shown in Fig. 7 and the variations are again quite small over the smaller field of view imaged by de Jong et al. (2024).

We note that the LR catalogue from Kondapally et al. (2021) was not used for astrometric verification because it does not cover the full LOFAR field of view. Furthermore, when it was constructed it made use of a preliminary Pan-STARRS data before the astrometry was refined by comparison with GAIA data (see Magnier et al. 2020). As a result the multi-wavelength positions in Kondapally et al. (2021) are on average 0.12″ in RA and 0.1″ in Dec offset from both Pan-STARRS DR2 positions and our new radio catalogue. Work is ongoing to construct an updated multi-wavelength catalogue with deeper optical datasets for identifications of the LoTSS Deep DR2 sources and this will have improved astrometry.

3.4. Dynamic range

To assess the dynamic range limitations of our final deep image we examine the noise enhancement within a 5′ radius around the 635 very isolated (more than 15″ from the nearest source of any flux density and more than 300″ from the nearest source with $S_I > 10$ mJy) sources with apparent integrated flux densities exceeding $S_I > 5$ mJy, where the region examined is fully within the 30% of the power primary beam. For each source we measure the median absolute deviation (MAD) of the pixel values in annular regions of increasing distance from the source. We fit these profiles with a Gaussian function, remove sources where this is a poor fit and divide all the fitted Gaussian functions by the measured local RMS. We group the sources into 10 apparent integrated flux density bins each containing the same number of sources. At each radial bin we calculate the median of the MAD values derived from all the sources in the appropriate integrated flux density bin as well as the bootstrap errors. For each apparent integrated flux density bin we then fit these median MAD profiles with a Gaussian and find the point at which the median MAD is only 5% higher than the background noise and calculate the area within this region. Finally, we find that the area impacted by dynamic range versus integrated flux density is well described by an exponential function, allowing us to derive that for sources with S_I greater than 5 mJy,

$$A_{\mathrm{Dyn},5\%} = -5.37 \times 10^{-4} + (6.62 \times 10^{-3} \times (1 - e^{-22.7 S_I})), \quad (2)$$

where $A_{\mathrm{Dyn},5\%}$ is the dynamic range limited area in square degrees and S_I is the integrated flux density in Jansky. We can use this to mask our image in order to contain only regions not substantially impacted (<5%) by dynamic range. We find that 7.4% of the map within 30% of the power primary beam has noise levels that are at least 5% higher than the local RMS due to dynamic range limitations. These regions are shown in Fig. 2.

3.5. Individual observation imaging quality

To compare the individual image quality we first aligned all individual images to the same flux density scale by following the full procedure outlined in Sect. 3.2 but using our deep catalogue for comparison rather than the Sabater et al. (2021) catalogue. We then constructed a catalogue containing only isolated unresolved sources in the deep field catalogue that are detected in all 64 of the individual epoch images. To create this catalogue we further filtered our unresolved, isolated deep field catalogue so that it just contains $\frac{S_I}{\sigma_{S_I}} > 5$ sources that are not in the 7.4% of the image impacted by dynamic range limitations (Sect. 3.4).

We then cross-matched the individual epoch catalogues (unfiltered) with the filtered deep catalogue using a cross-matching radius of 4″ and remove sources not present in all epochs. This yielded a catalogue of 560 sources that have integrated flux densities between 0.55 and 7.55 mJy.

The quality of the images derived from each observation varies substantially, as is apparent from Table S1, which shows details of the individual epoch images and their corresponding observations. To further assess the relative quality of the images we examined the 560 sources that are detected in all 64 epochs. Histograms of the integrated flux density ratios of the sources in each epoch compared to the deep image are shown in Fig. 8. Whilst the median values of the flux density ratios (0.97 ± 0.02) show the flux density scales of the images are well aligned the widths of the histograms vary significantly. For each epoch we calculated the median absolute deviation (MAD) of the flux density ratios and identified six pointings (namely L233804, L346136, L346154, L346454, L347494 and L782679) with particularly poor integrated flux density values compared to the deep image (MAD values more than 1.5 times the median MAD). The actual image RMS values at shown in Fig. 8. The figure also shows these RMS values adjusted to reflect variations in bandwidth, integration time and flagging fraction (blue crosses), as well as the inverse of the source sky density within 30% of the power primary beam (red circles). Here we see that the six epochs with poor integrated flux density values have consistently high RMS values and a low number of sources.

If we exclude the six epochs identified as poor then 1447 sources are detected in every one of the remaining 58 images. These sources are located across the image with a mean of 33 sources per facet and a standard deviation of 22 amongst the facets. Using these sources we attempted to quantify flux density scale variations in the individual epoch images as a function of position. To do this we examined the distribution of the ratios of the integrated flux densities in the deep epoch to those in the individual epoch images for every facet. We compared these distributions to those derived from simple simulations using the same sources and errors but adding an additional fractional error to the source flux that is fixed for a given observation and facet (i.e. mimicking a flux density scale misalignment between the observations that is allowed to vary between facets). We then varied the amount of additional fractional error in the integrated flux density until the standard deviation of a Gaussian function fitted to the simulated distribution matches the real distribution. The results of this are shown in Fig. 9, which indicates that for regions further from the pointing centre there are generally larger discrepancies on the integrated flux density measurements between observations: a crude linear fit indicates an additional error on the integrated flux density of $\sigma_{S_I,A} = \frac{D-0.76}{10} S_I$ for $D > 0.76$ where D is distance from the pointing centre in degrees. This effect may be partly due to calibration issues (e.g. we note that facets further from the pointing centre also tend to be larger, implying that the ionospheric calibration may be less precise) or possibly uncertainty in the LOFAR beam model could play a role. The mean additional flux density scale uncertainty of all the facets is 5%. This level represents the residual flux density scale uncertainty between individual epoch images after they have been aligned with the flux density scale of our deep image.

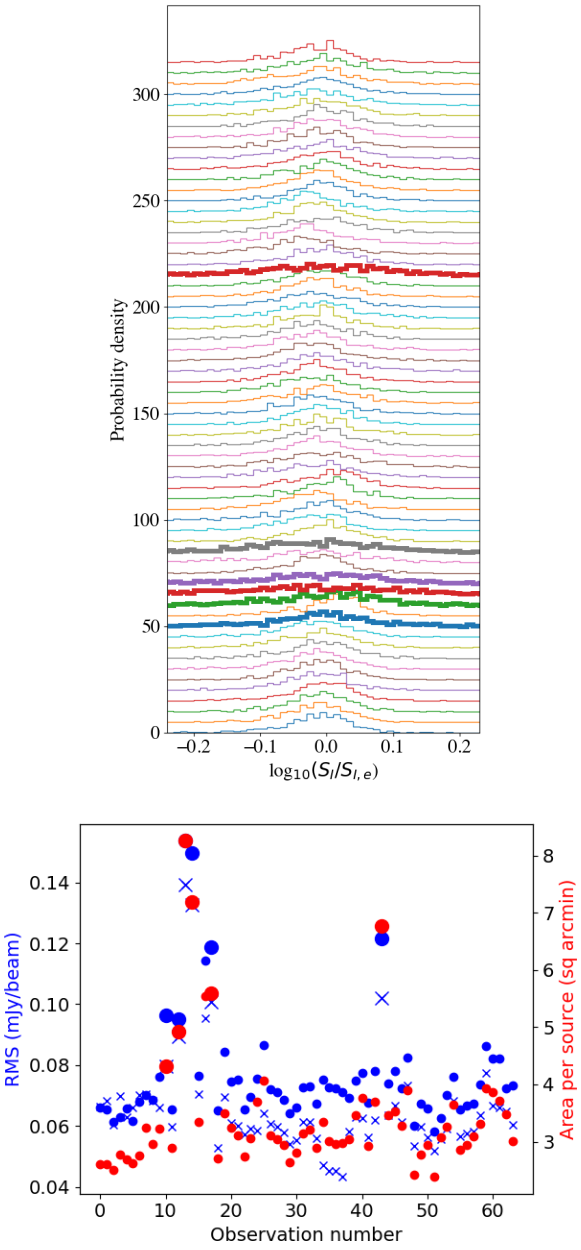

Fig. 8. Variations in the characteristics of the individual epoch ELAIS-N1 images. The top panel shows the probability density histograms of the integrated flux density ratio for the 64 individual epochs catalogues compared to measurements from our final continuum image. The y-axis of each epoch is offset by five for display purposes. The six thick lines correspond to observations L233804, L346136, L346154, L346454, L347494, and L782679, which have poor flux density values and were identified as having either a median flux density ratio or standard deviation further than 1.5 sigma from the other observations. In the bottom panel we show the RMS measured for each of the 64 epochs with blue circles and the blue crosses show the RMS scaled to a consistent 90° elevation, 231 subbands and 8 h duration without any flagging (see Table S1). The red circles show the area per detected source in each of the epoch images (i.e. inverse of source density). The larger circles and crosses correspond to the six poor quality observations.

4. Image analysis

Here we discuss several properties of the dataset to help demonstrate its scientific potential, its limitations and the improvements it provides upon previous work. We evaluate the sensitivity and its limitations from confusion noise, and the ability of the data to discover variable and polarised sources and to probe faint diffuse emission.

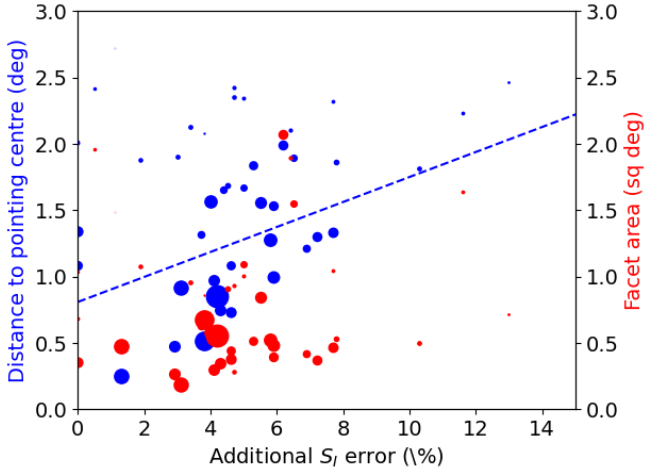

Fig. 9. Additional fractional uncertainty in S_I for each facet plotted against distance to the pointing centre (blue) and the facet area (red). The additional uncertainty is derived by comparing the distribution of ratios of integrated flux density measurements in each epoch to the deep integration for the 1447 sources detected in the 58 good epochs with a simple simulation using the same sources and errors but allowing for flux density scale offsets between pointings and facets. The consistency of our epoch image integrated flux density measurements decreases with increasing distance from the facet centre and the dashed blue line shows a linear fit to the data using the number of sources in the facet (proportional to marker size) for weighting.

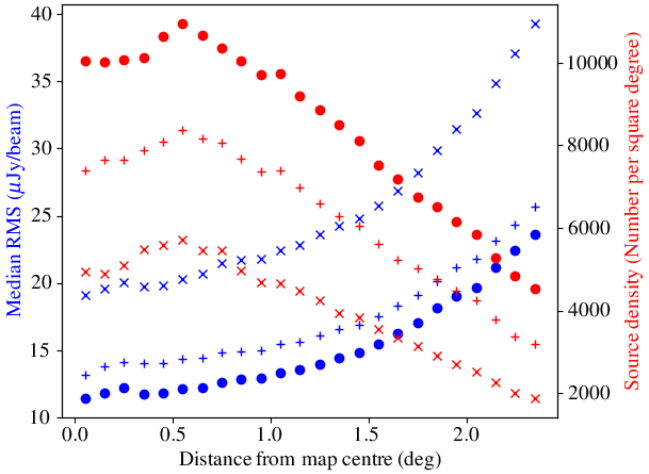

Fig. 10. Source density (red) and the image RMS (blue) as a function of distance from the pointing centre. The x symbols show the values derived from the previous release of the LoTSS Deep ELAIS-N1 field (Sabater et al. 2021), the + symbols show the values derived from our image using the exact same PyBDSF parameters as Sabater et al. (2021) and the circular symbols show the values derived using our revised PyBDSF strategy described in Sect. 2.1.

4.1. Comparison with Sabater et al. (2021)

For a simple quantitative comparison with Sabater et al. (2021) we create a catalogue from our image using exactly the same PyBDSF parameters as used in this previous study. We examine the source density and image noise as a function of distance from the pointing centre and the results are shown in Fig. 10. The RMS in our new image is a mean factor of 0.68 times the RMS in the previous data release image with little radial dependence (minimum factor is 0.65 and maximum factor is 0.71) and generally the source density is 1.57 times higher (minimum of 1.43 and maximum of 1.71). Using the refined PyBDSF parameters that detect sources at a lower level of significance and better characterise extended structures we again significantly increase (by a mean factor of 1.34) the number of sources in the catalogue derived from our new image compared with using the old PyBDSF parameters on the same image. Furthermore, adopting this new approach means additional faint sources are removed from the residual image from which the RMS is calculated. This results in a moderate decrease (mean factor of 0.87) in the measured RMS of our new image compared to when using the old PyBDSF parameters on the same image.

4.2. Confusion noise

Confusion noise is a consequence of the population of faint sources with flux densities below the detection/deconvolution threshold. Assuming that this population consists solely of point sources then each undetected source will contribute to the image as the data PSF multiplied by the flux density of the source with both the source and associated sidelobes adding noise to the image. When a large number of sources contribute to the confusion noise then these contributions form a substantial background of noise that has a highly skewed and non-Gaussian distribution with a non zero mean (e.g. Condon et al. 2012, Zwart et al. 2015).

As was already apparent in the work of Sabater et al. (2021) the sensitivity of the 6″ resolution LoTSS Deep images is impacted by confusion noise even at a sensitivity of 20 μJy beam^{-1}. To quantify the impact of confusion noise as a function of image depth we examine how the image noise decreases as a function of observing time and assess how it deviates from expectations, or comparable measurements, without confusion noise. To explore this it is prohibitively computationally expensive to reimage the large dataset with various amounts of integration time removed. Instead we stacked the individual epoch DDFacet residual Stokes I images to mimic different total integration times. To do this we randomly chose a number (2–64) of individual epoch images and stacked them together in the image plane using the weighted mean ($\sigma^{-2}_{V_{\rm RMS}}$, where $\sigma_{V_{\rm RMS}}$ is root mean square Stokes V image noise) that accounts for variation in the image quality from factors such as observing elevation and ionospheric conditions or calibration quality. The total noise on each stacked image was then estimated by finding the 68.27 percentile of a function fitted to the distribution of the stacked DDFacet residual map noise pixels after removing regions from the image impacted by dynamic range (see Sect. 3.4). The thermal noise contribution to the total noise on each stacked image was found from Stokes V images stacked using the same epochs and weighting. In Fig. 11 we demonstrate that the noise properties measured from the stacked individual epoch Stokes V dirty maps (the data were processed before DDFacet had Stokes V deconvolution functionality) is approximately equal to the noise estimated from the weights of the individual images (the standard deviation of the ratio of the measured noise to the estimated noise is 0.02). This behaviour is expected as the noise in these maps does not suffer from confusion, since sources are generally not circularly polarised and our levels of leakage are low (see Sect. 4.4).

The analysis is complicated by the fact that each individual epoch image is only deconvolved to a certain threshold (five times the local noise of typically ∼70 μJy beam^{-1}), and therefore

sources below this threshold are present in our stacked DDFacet residual images and contribute to the confusion noise in those maps. This is different from how the deep image was created as there we used all the data during imaging and deconvolved to a far lower threshold (five times the local noise of approximately ∼10.7 μJy beam^{-1}). To account for this during the stacking of our DDFacet residual images we searched for sources that were present in the deep image unresolved source catalogue and if these sources were detected in the stacked DDFacet residual image with a significance exceeding 3 times the local noise (obtained from a weighted stacking of the individual epoch PyBDSF noise maps) then we subtracted them in the image plane, assuming that the sources are point-like and have the response of the PSF of the central facet (DDFacet calculates a PSF for each facet). We did this by subtracting the peak brightness multiplied by the normalised PSF centred on the source position. This subtraction means that these sources and associated sidelobes no longer contribute to the measured confusion noise in the stacked DDFacet residual image measurements, thus approximately mimicking the joint deconvolution of the stacked data.

Our measurements showing the noise estimates on our stacked DDFacet residual images compared to the thermal noise estimates from the Stokes *V* images are shown in Fig. 11. We also show the measurements from the stacked DDFacet residual images prior to the image-plane deconvolution to demonstrate the impact of that procedure. In this plot we clearly see that as more DDFacet residual images are stacked the sensitivity continues to improve but that the measured stacked DDFacet residual image noise deviates further from the corresponding thermal noise estimate. This is because, even though the confusion noise is decreasing as more data are stacked due to the image-plane deconvolution removing detected sources from the stacked DDFacet residual and the number of undetected sources below the detection threshold decreasing, there is still a population of faint sources that remain in the stacked DDFacet residual images that form the confusion noise.

The shape of the curve in the Stokes *I* DDFacet residual image measurements in Fig. 11 can be compared to simulations that attempt to reconstruct the confusion noise based on an assumed source population and the true image PSF. At low flux densities there is uncertainty about the source population as it has not yet been measured but it can be estimated from simulations and existing counts. In Fig. 12 we show a fit of the differential source counts normalised to a non-evolving Euclidean model derived from the previous LoTSS Deep release by Mandal et al. (2021). At low integrated flux densities (≲0.43 mJy) the normalised counts begin to rapidly decline but, given statistical limitations due to the depth of the previous data release, the uncertainty becomes increasingly large. Instead of following this rapid decline we extrapolate the source counts from Mandal et al. (2021) at 0.43 mJy assuming the number counts down to 1 μJy from the T-RECS simulated source population at 144 MHz (Bonaldi et al. 2023).

In our simulations we take the PSF of the central facet together with our approximation of the real source counts and for each source inject the appropriately scaled PSF into the stacked Stokes *V* map (i.e. the thermal noise map) at a random location. To mimic DDFacet residual images, we then remove sources (PSFs) that exceed a detection threshold (5σ) as these would have been deconvolved during the imaging and would no longer be in the DDFacet residual image. We continue subtracting sources until the image noise converges.

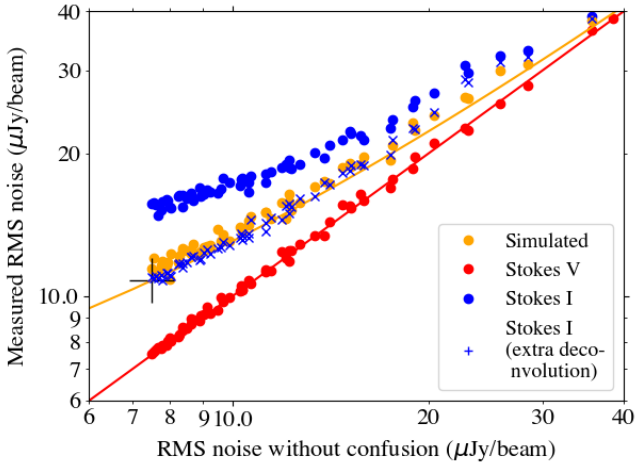

Fig. 11. Impact of confusion noise at 144 MHz and 6″ resolution. The blue crosses and circles show the noise measured from our stacked Stokes *I* residual images with and without image-plane deconvolution of sources that would be detected and deconvolved if the *uv*-data were imaged together. The orange circles show the measurements from our simulations that are fit with $\sigma_{I_{\rm RMS}} = 0.9\sigma_{V_{\rm RMS}} + 5.7\times10^{-4}\sigma^2_{V_{\rm RMS}} + 3.96$. Here $\sigma_{I_{\rm RMS}}$ is the measured Stokes *I* sensitivity (root mean square noise) and $\sigma_{V_{\rm RMS}}$ is the sensitivity without confusion, both are in micro Janskys per beam. In red are the measurements from our stacked Stokes *V* maps. The *y*-axis shows the total noise measured from stacked DDFacet residual images whereas the *x*-axis shows the thermal noise estimated from the weights of the Stokes *V* images for datasets included in that stacked image. The red line shows where these are equal and the large black cross is the sensitivity of our final ELAIS-N1 6″ resolution image.

In Fig. 11 we plot the results of our simulations together with the measurements from the stacked DDFacet residual images as well as the derived noise estimate on our final jointly deconvolved map where we have used the weights from the Stokes *V* image stacked map to estimate the thermal noise. Through both our image-plane deconvolved stacked images and our simulations we are able to reproduce the noise achieved in the final jointly deconvolved map (implying the source counts provide comparable confusion noise to what is observed) as well as the trend for increasing deviations from thermal noise with decreasing sensitivity. Our final image reaches 10.7 μJy beam^{-1} sensitivity where approximately half (defined as $\sigma_{I_{\rm CONF}} = \sqrt{\sigma^2_{I_{\rm RMS}} - \sigma^2_{V_{\rm RMS}}}$) of this is from confusion noise, $\sigma_{I_{\rm CONF}}$ (i.e. we would have reached an RMS noise level of 7.5 μJy beam^{-1} if there were no confusion). Extrapolating our measurements we find that if we were to double the integration time, confusion would become even more dominant – an estimated total noise of 8.8 μJy beam^{-1} or $\sigma_{V_{\rm RMS}} = 5.3$ μJy beam^{-1} (i.e. noise without confusion) and confusion noise of $\sigma_{I_{\rm CONF}} = 7.0$ μJy beam^{-1}.

4.3. Compact source variability

We first constructed a catalogue containing all isolated unresolved sources in the field that we wish to search for variability. To do this we begin by searching for sources that are only detected in any of the 58 individual epoch images (excluding the six poor quality epochs identified in Sect. 3.5) and not in the deep image. For this we filter each individual epoch catalogue so that it just contains $\frac{S_I}{\sigma_{S_I}} > 5$ unresolved isolated sources, where again 'unresolved' is defined as derived in Sect. 3.1 together with having PyBDSF source code 'S' and 'isolated' is defined as

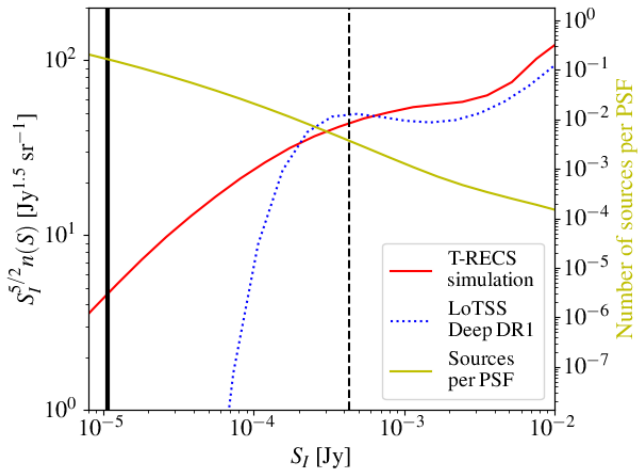

Fig. 12. Euclidean normalised source counts at 144 MHz (y-axis on the left) from previous surveys and simulations. The dashed blue line shows the counts from Mandal et al. (2021) who used the previous LoTSS Deep fields data release. The solid red line shows the counts from a T-RECS (Bonaldi et al. 2023) simulation at 144 MHz. The Mandal et al. (2021) counts drop rapidly at low flux densities due to the limited depth of the previous LoTSS Deep data release images. Hence in our confusion noise simulations (Sect. 4.2), we have extrapolated the Mandal et al. (2021) counts below 0.43 mJy (shown by vertical dashed black line) with the T-RECS counts. The yellow line shows the number of sources above a given flux density per 6″ restoring beam (y-axis on the right) and the thick vertical black line shows the sensitivity of our final deep ELAIS-N1 6″ image. Source counts derived from LoTSS Deep DR2 will be presented in a future study.

having no nearest neighbours within 15″. We cross-match each of these filtered individual catalogues with the deep catalogue using a cross-matching radius of 4″ and find that there are only two unresolved isolated sources detected in the individual fields (both in L811736) that are not also in the deep catalogue, but one of these is clearly an artefact and associated with a nearby (0.1°) bright (2.8 Jy) source so it is removed.

We then filtered the deep catalogue (154 952 entries plus one entry for the source found just in L811736) using the same criteria as above, which removed 83 722 due to the S/N cut, a further 11 924 due to the nearest neighbour cut, 12 110 from the S_Code cut, 4096 by the size cut, and 3365 due to dynamic range cuts. This left 39 736 sources amongst which we could search for variability. For each source and each epoch, we recorded the catalogued S_I and S_P and the corresponding errors, which are all scaled by the individual field scaling factors (derived in Sect. 3.5) to align with the deep field catalogue. If no source was detected in a particular individual epoch catalogue, we instead derived S_I and S_P and associated errors from the primary beam corrected individual epoch images. This was done by taking the pixel value corresponding to the source position in the deep catalogue as well as the local RMS around this position. Both values were scaled by the appropriate individual field scaling factor to align with the deep field catalogue. For all sources we also recorded the sum of the pixels within the PyBDSF deep image mask island for that particular source scaled by the individual field scaling factors ($S_{I,\mathrm{ap}}$) and associated error ($\sigma_{\mathrm{ap,tot}}$). Out of the 39 736 sources, 37 676 of these are detected in at least one of the 58 individual epoch observations whilst only 1447 are detected in all epochs. The 2060 sources not detected in any of the individual epochs were not examined further.

For each of the 37 676 sources in this cross-matched catalogue, we calculated the commonly used variability parameters V and η (see e.g. Rowlinson et al. 2019), which are proxies for the amount of variability and the variability significance, respectively. These parameters are defined as

$$V = \frac{1}{\overline{S}} \sqrt{\frac{N}{N-1}\left(\overline{S^2} - \overline{S}^2\right)}, \tag{3}$$

$$\eta = \frac{N}{N-1}\left(\overline{\omega S^2} - \frac{\overline{\omega S}^2}{\bar{\omega}}\right), \tag{4}$$

where V is the ratio of the standard deviation of the flux densities to the mean flux density, η is the reduced χ^2 relative to a model with constant flux density, N is the number of measurements (58), $\omega = 1/\sigma^2_{I,\mathrm{tot}}$, overbars denote mean and an additional uncertainty that we define as 5% of the flux density (see Sect. 3.5) is added in quadrature to the σ_{S_I} derived by PyBDSF to give the total error on integrated flux density ($\sigma_{S_I,\mathrm{tot}}$).

If a source is one of the 1447 detected in all epochs we calculated the variability parameters using the catalogued epoch S_I, S_P and $S_{I,\mathrm{ap}}$ values as S. Additionally, we repeated the variability calculation 1000 times drawing S from a Gaussian distribution centred on S_I with width $\sigma_{S_I,\mathrm{tot}}$ and a further 1000 calculations drawing from around each of $S_{I,\mathrm{ap}}$ and S_P instead of S_I. For each flux density type we then conservatively selected the lowest derived V and η for that source from all the calculations (i.e. least amount of variability given the measurements).

Most sources, however, are only detected in a few epochs. For these sources we performed the same analysis as above but for the epochs where the source is not detected we drew S from a uniform probability distribution between 0 and the upper limit (e.g $5\times\sigma_{I,\mathrm{tot}}$, $5\times\sigma_{P,\mathrm{tot}}$ or $5\times\sigma_{\mathrm{ap,tot}}$). We set the error on the flux density values of these undetected sources as being the standard deviation of a population of flux densities drawn from a uniform distribution between 0 and the upper limit. Again for these sources we conservatively selected the lowest derived V and η from all the calculations. In total we thus have three minimum V and η values for each source corresponding the S_I, S_P, and $S_{I,\mathrm{ap}}$ measurements.

Figure 13 shows the lowest of the three minimum variability parameter sets for each of the 37 676 sources we have examined – that is, these are the lowest values of both of the variability parameters for each source that we are able to derive given the measurements (see Fig. 14 for an example). They are colour-coded by the fraction of epochs a particular source is detected in. For demonstration purposes we have also included the well known flare star binary CR Draconis (see e.g. Callingham et al. 2021a) even though it was excluded from the sample as it is in a dynamic-range limited area of the map. We also included the millisecond pulsar PSR J1552+5437 (Pleunis et al. 2017), which was excluded due to PyBDSF fitting the emission with multiple Gaussians. As can be seen from the shape of the distribution the sources that are detected in almost all epochs show a slightly increasing variability significance (η) with variability (V). This is a consequence of the 5% fractional error that we have placed on the flux density – fainter sources tend to have larger variability (V) because the source fluxes are dominated by measurement errors and the flux density changes significantly from epoch to epoch. For brighter sources that are detected in almost all epochs the fractional uncertainty dominates over the measurement errors. For these sources we find rapidly decreasing V with increasing flux density whereas η only slightly decreases with increasing flux density. If the calculations are repeated with

a 10% fractional error on the flux density this trend of increasing V with η for sources that are detected in most epochs becomes much stronger as η more rapidly decreases with increasing flux density. The weak trend we see with a 5% fractional uncertainty gives confidence that this fractional uncertainty approximately reflects the correct value for the data under the assumption that sources that are detected in more epochs are genuinely of similar variability to sources detected in fewer epochs.

The sources that are detected in a few of the epochs have a large variability, as for each undetected epoch we randomly assigned the flux density a value between zero and the 5σ upper limit. However, the variability significance (η) is low because the assigned flux densities in undetected epochs have a high degree of uncertainty (the standard deviation of a population of flux densities is in the range between zero and the upper limit).

Outliers in Fig. 13 with an unusually high η are our candidate variable sources. These are distributed across the entire image (see Fig. 2) and have a range of different levels of variability (V). The 38 sources with η above the 99.9 percentile are circled in red and those circled in blue are the six example sources that we show in Fig. 15. The sources with high η include the flare star CR Draconis (see e.g. Callingham et al. 2021a and the millisecond pulsar PSR J1552+5437 (see e.g. Pleunis et al. 2017). The small red dwarf star with an exoplanetary companion GJ 625 is a prominent outlier in V albeit with a lower η (circled blue but not red). The other 36 sources with high η were all classified as part of the Best et al. (2023) LoTSS Deep fields DR1 classification effort and were classed as 2 high-excitation radio galaxies (HERGs), 17 low-excitation radio galaxies (LERGs), 3 radio-quiet AGN (RQAGN), 11 star-forming galaxies (SFGs) and 3 unclassified. Furthermore, we note that all 36 of these sources are highly compact as they are detected in the 0.3″ resolution LOFAR image (de Jong et al. 2024) and have measured major axis below 1.2″ in that catalogue. If the emission is genuinely compact then a possibility is that refractive scintillation plays a significant role (e.g. Bell et al. 2019) but the modulation index (here defined as $M = \frac{\sigma_{S,\mathrm{epochs}}}{\overline{S}}$, where $\sigma_{S,\mathrm{epochs}}$ is the standard deviation of the epoch flux density measurements and $\overline{S}$ is the mean flux density measurement from the epochs) we measure for these 36 sources using all our data is comparably high (between 12% and 55%). Further investigation is required to more precisely characterise the variability and confirm its nature.

The cautious approach we have taken to search for variable sources in the field uses three different ways of measuring the flux density and searches for the minimum levels of variability given the measurements and errors. This approach was used to demonstrate that substantial populations of variable sources exist in these data and are identifiable despite the variations in the image characteristics between epochs (e.g. Fig. 8). There are many refinements that could be made to the procedure to identify larger numbers of variable sources in future studies, such as better characterisation of local variations in image quality, expanding the search to cover more sources, making fuller use of the image values, improving the identification of outliers in the variability plot and exploring different time periods (combining epochs or within epochs). For example, we can make more use of the data if we simply repeat the calculations but for epochs without detections we instead use Gaussian probability distributions centred on the image measurements of S_I, S_P and $S_{I,\mathrm{ap}}$ and with standard deviations set to the corresponding errors. By doing this we find that for sources with few epoch detections typically the η increases slightly (~15%) whilst the distribution in V broadens substantially and stretches to far higher values (~1.0). Overall though the sources selected as most significantly variable are

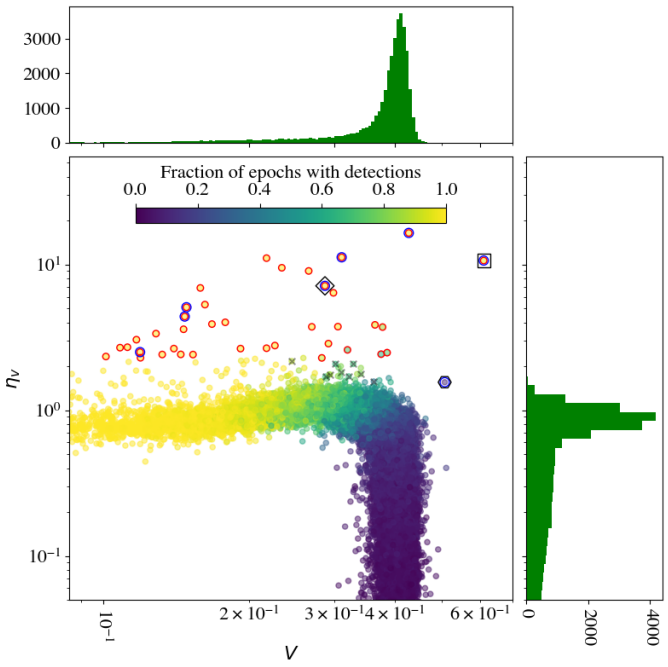

Fig. 13. Derived variability parameters V and η for the 37 676 sources detected in at least one epoch. The marker colour shows the fraction of epochs in which the source is detected. More variable sources have high values of V and η. The 38 candidate variable sources are circled in red and the sources shown in Fig. 15 are circled in blue. The box, diamond and hexagon show Cr Draconis, PSR J1552+5437 and GJ 625 respectively. The grey x's show additional candidate variable sources identified when using Gaussian rather than uniform probability distributions for non detections (see text).

largely the same as those previously identified (29 out of 38) with only the candidate variable sources close to the 99.9 percentile boundary shifting one way or the other. The main difference was that this approach resulted in more low signal-to-noise candidate variables that are detected in just a few epochs (marked with grey x's in Fig. 13), whereas the uniform probability distribution approach to non detections led to more candidates from sources that are detected in the majority of the epochs.

To allow for examination of variability within these data we include in this data release a catalogue containing the flux density and error measurements from each epoch in addition to the images from individual epochs and the derived minimum η and V values.

4.4. Stokes V emission

There are three known circularly polarised sources within the field, CR Draconis, GJ 625 and PSR J1552+5437. The circularly polarised emission from these sources was identified either in this ELAIS-N1 LOFAR dataset by Callingham et al. (2021b,a), and Sabater et al. (2021) or in LoTSS by Callingham et al. (2023). Here we search for additional Stokes V emission at the locations of each of the 39 676 unresolved sources identified in Sect. 4.3 and we also measure the emission from the three known Stokes V emitters to demonstrate our methodology. We used several search techniques to try to take advantage of the multi-epoch dataset whilst recognising that the handedness of Stokes V sources can vary with time and changes would cause the signal to flip from positive to negative or vice versa in the Stokes V images.

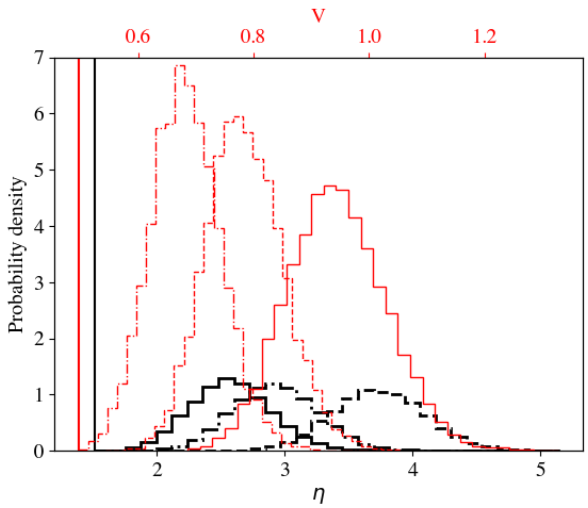

Fig. 14. Probability density distributions of the derived variability parameters V and η for GJ 625. In black we show η with the solid, dashed, and dot-dashed lines corresponding to the S_I, S_P, and $S_{I,\mathrm{ap}}$ measurements, respectively. The red shows the V distributions using the same line style. The red and black vertical lines show the lowest derived V and η parameters (i.e. least variation) that we adopt.

Firstly, we used PyBDSF to create noise images for each of our Stokes V individual epoch images and used these to measure the Stokes V noise at the positions of each source in each epoch. We then searched for Stokes V emission by finding the maximum absolute value of the Stokes V image in the pixels that lay both within the appropriate deep image Stokes I PyBDSF mask island and within 4″ of the catalogued deep image position. This search was performed for every epoch image in which the source was detected in Stokes I giving a total of 570 944 Stokes V measurements. For consistency we scaled our Stokes V measurements and errors by the individual field scaling factors that were derived to align the individual epochs with the deep catalogue. In Fig. 16 we show our measured Stokes V signal divided by noise and the corresponding Stokes I epoch measurements (PyBDSF catalogued S_P and σ_{S_P}). In total, six different sources are detected in Stokes V with an S/N exceeding five. The two most prominent detections are of the flare star CR Draconis (33 detections with Stokes V $5.3 < \mathrm{S/N} < 46.2$) and the small red dwarf star GJ 625 (3 detections with Stokes V $5.9 < \mathrm{S/N} < 7.0$). The pulsar PSR J1552+5437 is detected at a maximum significance of 4.8 whereas another four sources have Stokes V S/N levels very slightly exceeding five (all less than 5.2) in a single epoch, these are sources with identities ILTJ161318.70+535057.6 (L798164), ILTJ161227.28+544553.2 (L686962), ILTJ160926.33+564927.2 (L230461) and ILTJ160859.34+531359.5 (L816344). These 4 sources have no obvious highly circularly polarised counterparts such as stars or known pulsars and may be false positives. To assess this we approximated the expected number of false positives by finding the number of noise pixels exceeding 5 S/N in the Stokes V cutout images outside of the deep field PyBDSF mask regions (i.e. away from all sources). Between all the epoch images we found 158 > 5 S/N Stokes V pixels out of 0.4 billion pixels searched away from sources (i.e. 6.6×10^{-5}%) and all had $\mathrm{S/N} < 6$. By comparison, if we exclude CR Dra, GJ 625 and PSR J1552+5437, then in our search for Stokes V emission in pixels associated with detected sources we identified 12 S/N > 5 Stokes V pixels out of 21 million pixels searched (i.e. 5.6×10^{-5}%). Hence our detection rate at source locations is approximately equal to our false positive rate and we conclude that our 4 additional detections of Stokes V sources are likely spurious.

In order to see if other detections could be made we searched for signals in stacked Stokes V cutout images of small areas around each individual source. We did this via two approaches. Firstly, we simply stacked the images using a weighted average (using the Stokes V PyBDSF noise image and inverse-variance weighting) and characterising the significance of detections in the same way as was done for the individual epoch sources. This approach works well for sources that do not change handedness but signals from sources that change handedness would be diluted. In this approach we detected CR DRa and PSR J1552+5437 at high significance with S/N of 50.7 and 12.0 respectively. No other sources were detected above a S/N of 5.0 with the next most significant being GJ 625 with an S/N of 4.98. In the second approach the stacking was done using an inverse-variance weighting of the absolute values of the Stokes V cutout images from all the epochs for each source, where the absolute value was used to account for possible sign changes between epochs. To quantify the significance of the measurements from the Stokes V stacked cutout images we evaluated the distribution of pixels of the full field stacked absolute value Stokes V epoch images. The pixels are well described by a beta distribution and we can quantify significance using confidence intervals, e.g. the confidence intervals of 68.27%, 95.45% and 99.73% corresponding to an S/N of 1, 2 and 3 respectively. The results from this analysis are also shown in Fig. 16 with the Stokes V stacked S/N compared to the deep field $\frac{S_I}{\sigma_{S_P}}$. The stars CR Dra and GJ 625 were detected with moderate to high significance as was the pulsar PSR J1552+5437 (72.0, 5.5, 8.4 respectively) but no other sources were detected above a significance of 5.0. Repeating the analysis but instead only considering epochs in which sources are detected in Stokes I, we were able to detect the known Stokes V emitters at improved S/N (77.5, 13.2, 10.5 for CR Dra, GJ 625 and PSR J1552+5437 respectively) but again we found no other sources with a significance exceeding 5.0.

Finally we attempted to search for a statistical detection from the 39 674 sources without obvious Stokes V emission (i.e. excluding CR Dra, GJ 625 and PSR J1552+5437, which are genuine Stokes V emitters) through another stacking exercise. We use the Stokes V cutouts for all epochs for all of the 39 674 sources. We then calculated the inverse variance weighted average of the absolute value of all the Stokes V cutout images. To assess the significance of any signal in the stacked on-source Stokes V image we use offset cutouts taken around each source and add these together in the same way. We fitted a beta distribution to the stacked offset cutout image and again quantified significance with confidence intervals. In addition to this we also created an artificial leakage image where we took the Stokes I flux density value in a particular epoch convolved to the 6″ restoring beam and multiplied that by leakage levels in the range of 0.01% to 1% (previously for bright sources in LoTSS-DR2 the leakage level was found to be 0.056%; Shimwell et al. 2022), added it to the stacked offset Stokes V image and performed the same weighted stacking. Through this analysis we found no Stokes V signal in our stacked on-source image (S/N of 0.5) but we are able to detect the artificial leakage signals at an S/N exceeding 5.0 when the leakage is greater than 0.045%.

To summarise, three sources (CR DRa, GJ 625 and PSR J1552+5437) are detected clearly in Stokes V using several different approaches but no other convincing detections are made

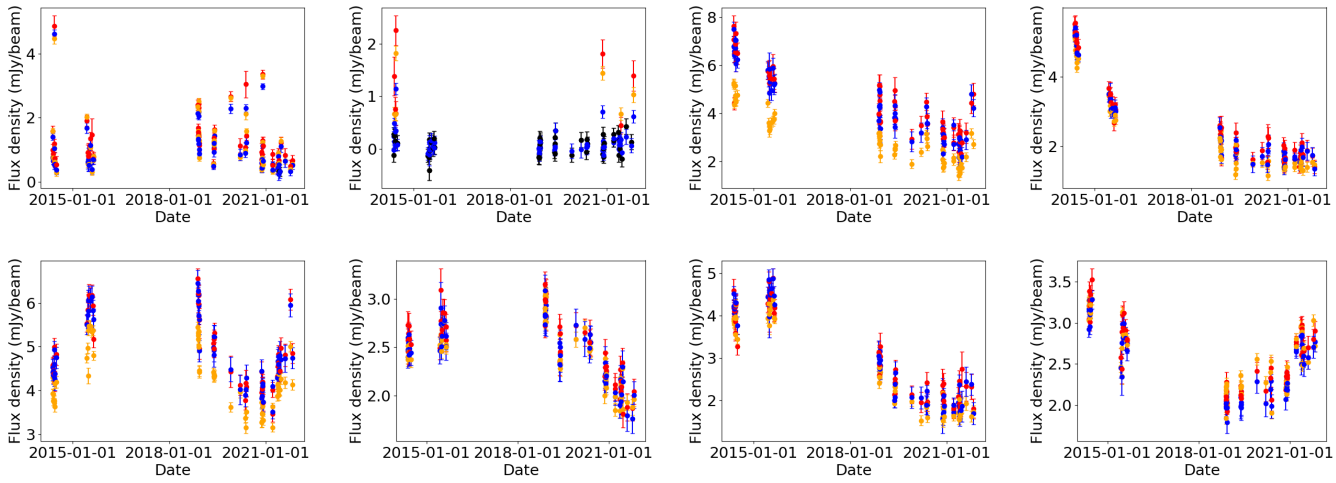

Fig. 15. For a selection of variable sources highlighted in blue in Fig. 13 we show the S_I (red), S_P (orange), $S_{I,\mathrm{ap}}$ (blue), and pixel values for non detections (black) as a function of time. Clockwise from the top left the panels, we show CR Draconis, GJ 625, PSR J1552+5437, and the others are classified as a RQAGN, LERG, SFG, LERG, and unclassified, respectively, by Best et al. (2023).

through either searching individual observations or stacking. A search for a statistical detection of Stokes V emission from all unresolved sources in the field also yielded a null result, implying that the typical level of emission from these sources when combined with leakage is less than 0.045%.

5. Public data release

The images and associated catalogues presented and analysed in this publication are publicly available and can be accessed via the LOFAR surveys webpage[10]. This consists of the following:

- A full depth (10.7 μJy beam^{-1}) high (6″) resolution Stokes I image out to the 30% power point of the LOFAR primary beam.
- A catalogue derived from the full depth 6″ resolution image that contains 154 952 radio sources and the corresponding 182 184 Gaussian components from PyBDSF (version 1.10.3). The columns are those provide by default by PyBDSF and include position, integrated flux density, peak brightness, size and estimated statistical errors on all parameters (i.e. not including radial effects or additional uncertainties on astrometry or flux density scale, which should be added as appropriate).
- High (6″) resolution Stokes I and Stokes V individual epoch images.
- A catalogue derived from the 6″ individual epoch images that contains 36 676 unresolved isolated sources with flux density measurements from each epoch and derived variability parameters.

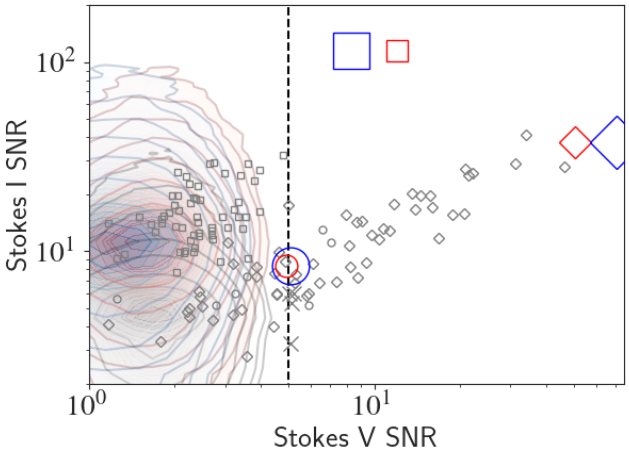

Fig. 16. Derived fractional circular polarisation for the 37 676 unresolved isolated sources detected in at least one epoch in a region of the map not impacted by dynamic range limitations. CR Draconis and PSR J1552+5437 are also included for demonstration purposes even though they fall outside the selection criteria. The grey (570 944 measurements), red and blue contours show the probability density function for the individual epoch measurements, stacked measurements and stacked absolute value measurements respectively. The lowest contours contain 99% of the measurements. The diamonds, circles, box and x's show CR Draconis, GJ 625, PSR J1552+5437 and other detections above 5 S/N respectively. The smaller grey symbols represent measurements from the individual epochs, the medium sized red symbols are from the stacked images and the large blue symbols are from the stacked absolute value images. Details of the S/N derivations are given in Sect. 4.4.

6. Summary

We have described the second and final 6″ resolution data release for the LoTSS Deep ELAIS-N1 field at 144 MHz. Our image contains 154 952 sources (consisting of 182 184 Gaussian components) in a 24.53 square degree region and reaches depths of 10.7 μJy beam^{-1}. The image is amongst the most sensitive radio wavelength images ever released over such a large area when accounting for the typical spectral index of radio sources. For comparison, assuming a typical spectral index of −0.8 (Böhme et al. 2023), our sensitivity is equivalent to 1.85 μJy beam^{-1} at 1284 MHz and is comparable to the depth of the MIGHTEE survey, which will span over 20 square degrees at that frequency (Heywood et al. 2022 and Hale et al. 2024). At higher frequencies, deeper images have been produced for small areas. For example, the COSMOS-XS survey covered 180 square arcminutes and reached a depth of 0.53 μJy beam^{-1} at 3 GHz (van der Vlugt et al. 2021), which is a factor of 1.8 deeper than the extrapolated LoTSS Deep sensitivity at this frequency

[10] https://lofar-surveys.org/

(0.97 μJy beam^{-1}), but again it covers over two orders of magnitude less sky area.

In this paper, we have characterised the general properties of the LoTSS Deep ELAIS-N1 image, finding that 9% of the sources appear to be resolved and 7.4% of the area is affected by dynamic range limitations. The flux density scale is thought to be good to better than 9%, and the standard deviation of the RA and Dec offsets between our catalogue and Pan-STARRS DR2 are less than 0.2″ for high-significance detections. We examined the contribution of confusion noise, finding that approximately half of the image noise is due to confusion and demonstrating that resolution is the most critical avenue to decreasing the noise level further. We searched for variable sources in the field and identified the 39 sources that are variable with high significance. We also searched for Stokes *V* emitting sources and did not detect any previously unknown sources but found three detections of known circularly polarised sources. We limited the circular polarisation signal of the population of our defined unresolved sources to be less than 0.045% of the total intensity signal including leakage.

The data from this release are publicly available and can be accessed via the collaboration's webpage. This includes deep images, individual epoch images, and source catalogues. Observations are now complete for three other LoTSS Deep fields, and publications examining the 6″, 0.3″, and polarisation properties are forthcoming. These will make use of up to ∼400 h of data for each field, which is approximately a factor of four larger than what was utilised for LoTSS Deep DR1. Future work will also present multi-wavelength identifications and characterisation for the radio sources detected in the LoTSS Deep fields, and through the William Herschel Telescope Enhanced Area Velocity Explorer survey of LOFAR selected sources (WEAVE-LOFAR; Smith et al. 2016), we will obtain optical spectroscopy for almost all sources in the central regions of these fields.

Data availability

All data products described in Sect. 5 are available on the LOFAR surveys webpage: https://lofar-surveys.org/. A table (Table S1) providing details of the individual observations and associated images is available via Zenodo, at https://doi.org/10.5281/zenodo.14603969.

Acknowledgements. This paper is based (in part) on data obtained with the LOFAR telescope (LOFAR-ERIC) under project code LC2_024, LC4_008, LT10_012 and LT14_003. LOFAR (van Haarlem et al. 2013) is the Low Frequency Array designed and constructed by ASTRON. It has observing, data processing, and data storage facilities in several countries, that are owned by various parties (each with their own funding sources), and that are collectively operated by the LOFAR European Research Infrastructure Consortium (LOFAR-ERIC) under a joint scientific policy. The LOFAR-ERIC resources have benefited from the following recent major funding sources: CNRS-INSU, Observatoire de Paris and Université d'Orléans, France; BMBF, MIWF-NRW, MPG, Germany; Science Foundation Ireland (SFI), Department of Business, Enterprise and Innovation (DBEI), Ireland; NWO, The Netherlands; The Science and Technology Facilities Council, UK; Ministry of Science and Higher Education, Poland. This research made use of the Dutch national e-infrastructure with support of the SURF Cooperative (e-infra 180169) and NWO (grants 2019.056 & 2023.036). The Jülich LOFAR Long Term Archive and the German LOFAR network are both coordinated and operated by the Jülich Supercomputing Centre (JSC), and computing resources on the supercomputer JUWELS at JSC were provided by the Gauss Centre for Supercomputing e.V. (grant CHTB00) through the John von Neumann Institute for Computing (NIC). This research made use of the University of Hertfordshire high-performance computing facility and the LOFAR-UK computing facility located at the University of Hertfordshire and supported by STFC [ST/P000096/1], and of the Italian LOFAR IT computing infrastructure supported and operated by INAF, and by the Physics Department of Turin university (under an agreement with Consorzio Interuniversitario per la Fisica Spaziale) at the C3S Supercomputing Centre, Italy. CLH acknowledges support from the Leverhulme Trust through an Early Career Research Fellowship. CLH also acknowledges support from the Oxford Hintze Centre for Astrophysical Surveys which is funded through generous support from the Hintze Family Charitable Foundation. PNB and RK are grateful for support from the UK STFC via grants ST/V000594/1 and ST/Y000951/1. AB acknowledges financial support from the European Union – Next Generation EU. MJH and DJBS acknowledge support from the UK Science and Technology Facilities Council (STFC) via grants ST/V000624/1 and ST/Y001249/1. RJvW acknowledges support from the ERC Starting Grant ClusterWeb 804208. MB acknowledges funding by the DFG under Germany's Excellence Strategy – EXC 2121 "Quantum Universe" – 390833306. FdG acknowledges support from the ERC Consolidator Grant ULU 101086378. KJD acknowledges support from the STFC through an Ernest Rutherford Fellowship (grant number ST/W003120/1). SI acknowledges the ASPIRE programme at the Anton Pannekoek Institute for Astronomy. IdR and AR acknowledge support from the NWO Aspasia grant (number 015.016.033). IdR, AR, and JMGHJdJ acknowledge support from the NWA CORTEX grant (NWA.1160.18.316) of the research programme NWA-ORC which is (partly) financed by the Dutch Research Council (NWO). D.G.N. acknowledges funding from Conicyt through Fondecyt Postdoctorado (project code 3220195). LKM is grateful for support from a UKRI FLF [MR/T042842/1]. IP acknowledges support from INAF under the Large Grant 2022 funding scheme (project "MeerKAT and LOFAR Team up: a Unique Radio Window on Galaxy/AGN co-Evolution"). FS appreciates the support of STFC [ST/Y004159/1].

References

Becker, R. H., White, R. L., & Helfand, D. J. 1995, ApJ, 450, 559
Bell, M. E., Murphy, T., Hancock, P. J., et al. 2019, MNRAS, 482, 2484
Best, P. N., Kondapally, R., Williams, W. L., et al. 2023, MNRAS, 523, 1729
Böhme, L., Schwarz, D. J., de Gasperin, F., Röttgering, H. J. A., & Williams, W. L. 2023, A&A, 674, A189
Bonaldi, A., Hartley, P., Ronconi, T., De Zotti, G., & Bonato, M. 2023, MNRAS, 524, 993
Bonato, M., Prandoni, I., De Zotti, G., et al. 2021, A&A, 656, A48
Bondi, M., Scaramella, R., Zamorani, G., et al. 2024, A&A, 683, A179
Calistro Rivera, G., Alexander, D. M., Harrison, C. M., et al. 2024, A&A, 691, A191
Callingham, J. R., Pope, B. J. S., Feinstein, A. D., et al. 2021a, A&A, 648, A13
Callingham, J. R., Vedantham, H. K., Shimwell, T. W., et al. 2021b, NatAs, 5, 1233
Callingham, J. R., Shimwell, T. W., Vedantham, H. K., et al. 2023, A&A, 670, A124
Chandra, P., Ray, A., & Bhatnagar, S. 2004, ApJ, 612, 974
Cochrane, R. K., Kondapally, R., Best, P. N., et al. 2023, MNRAS, 523, 6082
Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, AJ, 115, 1693
Condon, J. J., Cotton, W. D., Fomalont, E. B., et al. 2012, ApJ, 758, 23
Das, S., Smith, D. J. B., Haskell, P., et al. 2024, MNRAS, 531, 977
de Gasperin, F., Dijkema, T. J., Drabent, A., et al. 2019, A&A, 622, A5
de Gasperin, F., Edler, H. W., Williams, W. L., et al. 2023, A&A, 673, A165
de Jong, J. M. G. H. J., van Weeren, R. J., Sweijen, F., et al. 2024, A&A, 689, A80
Dijkema, T. J., Nijhuis, M., van Diepen, G., et al. 2023, Astrophysics Source Code Library [record ascl:2305.014]
Duncan, K. J., Kondapally, R., Brown, M. J. I., et al. 2021, A&A, 648, A4
Emig, K. L., White, G. J., Salas, P., et al. 2022, A&A, 664, A88
Euclid Collaboration (Scaramella, R., et al.) 2022, A&A, 662, A112
Flewelling, H. A., Magnier, E. A., Chambers, K. C., et al. 2020, ApJS, 251, 7
Garn, T., Green, D. A., Riley, J. M., & Alexander, P. 2008, MNRAS, 383, 75
Gloudemans, A. J., Duncan, K. J., Kondapally, R., et al. 2021, A&A, 648, A7
Groeneveld, C., van Weeren, R. J., Osinga, E., et al. 2024, NatAs, 8, 786
Hale, C. L., Williams, W., Jarvis, M. J., et al. 2019, A&A, 622, A4
Hale, C. L., Heywood, I., Jarvis, M. J., et al. 2024, MNRAS, 536, 2187
Hales, S. E. G., Masson, C. R., Warner, P. J., & Baldwin, J. E. 1990, MNRAS, 246, 256
Hardcastle, M. J., Gürkan, G., van Weeren, R. J., et al. 2016, MNRAS, 462, 1910
Hardcastle, M. J., Shimwell, T. W., Tasse, C., et al. 2021, A&A, 648, A10
Herrera Ruiz, N., O'Sullivan, S. P., Vacca, V., et al. 2021, A&A, 648, A12
Heywood, I., Murphy, E. J., Jiménez-Andrade, E. F., et al. 2021, ApJ, 910, 105
Heywood, I., Jarvis, M. J., Hale, C. L., et al. 2022, MNRAS, 509, 2150
Intema, H. T., Jagannathan, P., Mooley, K. P., & Frail, D. A. 2017, A&A, 598, A78
Ishwara-Chandra, C. H., Taylor, A. R., Green, D. A., et al. 2020, MNRAS, 497, 5383
Jannuzi, B. T., & Dey, A. 1999, ASPC, 191, 111
Jarvis, M., Taylor, R., Agudo, I., et al. 2016, MeerKAT Science: On the Pathway to the SKA, 6

Jelić, V., de Bruyn, A. G., Mevius, M., et al. 2014, A&A, 568, A101
Kondapally, R., Best, P. N., Hardcastle, M. J., et al. 2021, A&A, 648, A3
Kondapally, R., Best, P. N., Cochrane, R. K., et al. 2022, MNRAS, 513, 3742
Kondapally, R., Best, P. N., Raouf, M., et al. 2023, MNRAS, 523, 5292
Lane, W. M., Cotton, W. D., van Velzen, S., et al. 2014, MNRAS, 440, 327
Lockman, F. J., Jahoda, K., & McCammon, D. 1986, ApJ, 302, 432
Magnier, E. A., Schlafly, E. F., Finkbeiner, D. P., et al. 2020, ApJS, 251, 6
Makarov, V., Berghea, C. T., & Frouard, J. H. 2017, Technical Memorandum, AADD USNO
Mandal, S., Prandoni, I., Hardcastle, M. J., et al. 2021, A&A, 648, A5
Mevius, M. 2018, Astrophysics Source Code Library [record ascl:1806.024]
Mingo, B., Croston, J. H., Best, P. N., et al. 2022, MNRAS, 511, 3250
Mohan, N., & Rafferty, D. 2015, Astrophysics Source Code Library [record ascl:1502.007]
Morabito, L. K., Jackson, N. J., Mooney, S., et al. 2022, A&A, 658, A1
Ocran, E. F., Taylor, A. R., Vaccari, M., Ishwara-Chandra, C. H., & Prandoni, I. 2020, MNRAS, 491, 1127
Offringa, A. R. 2016, A&A, 595, A99
Offringa, A. R., van de Gronde, J. J., & Roerdink, J. B. T. M. 2012, A&A, 539, A95
Oliver, S., Rowan-Robinson, M., Alexander, D. M., et al. 2000, MNRAS, 316, 749
Osinga, E., van Weeren, R. J., Boxelaar, J. M., et al. 2021, A&A, 648, A11
Piras, S., Horellou, C., Conway, J. E., et al. 2024, A&A, 687, A267
Pleunis, Z., Bassa, C. G., Hessels, J. W. T., et al. 2017, ApJ, 846, L19
Rengelink, R. B., Tang, Y., de Bruyn, A. G., et al. 1997, A&AS, 124, 259
Rowlinson, A., Stewart, A. J., Broderick, J. W., et al. 2019, A&C, 27, 111
Sabater, J., Best, P. N., Tasse, C., et al. 2021, A&A, 648, A2
Scaife, A. M. M., & Heald, G. H. 2012, MNRAS, 423, L30
Shimwell, T. W., Röttgering, H. J. A., Best, P. N., et al. 2017, A&A, 598, A104
Shimwell, T. W., Tasse, C., Hardcastle, M. J., et al. 2019, A&A, 622, A1
Shimwell, T. W., Hardcastle, M. J., Tasse, C., et al. 2022, A&A, 659, A1
Sirothia, S. K., Dennefeld, M., Saikia, D. J., et al. 2009, MNRAS, 395, 269
Smirnov, O. M., & Tasse, C. 2015, MNRAS, 449, 2668
Smith, D. J. B., Dunne, L., Maddox, S. J., et al. 2011, MNRAS, 416, 857
Smith, D. J. B., Best, P. N., Duncan, K. J., et al. 2016, Proc. Annu. Meet. French Soc. Astron. Astrophys., 271
Smith, D. J. B., Haskell, P., Gürkan, G., et al. 2021, A&A, 648, A6
Smolčić, V., Novak, M., Bondi, M., et al. 2017a, A&A, 602, A1
Smolčić, V., Novak, M., Delvecchio, I., et al. 2017b, A&A, 602, A6
Šnidarić, I., Jelić, V., Mevius, M., et al. 2023, A&A, 674, A119
Sutherland, W., & Saunders, W. 1992, MNRAS, 259, 413
Sweijen, F., van Weeren, R. J., Röttgering, H. J. A., et al. 2022, NatAs, 6, 350
Tasse, C. 2014, ArXiv e-prints [arXiv:1410.8706]
Tasse, C., Hugo, B., Mirmont, M., et al. 2018, A&A, 611, A87
Tasse, C., Shimwell, T., Hardcastle, M. J., et al. 2021, A&A, 648, A1
van der Vlugt, D., Algera, H. S. B., Hodge, J. A., et al. 2021, ApJ, 907, 5
van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, A&A, 556, A2
White, R. L., Becker, R. H., Helfand, D. J., & Gregg, M. D. 1997, ApJ, 475, 479
Whittam, I. H., Jarvis, M. J., Hale, C. L., et al. 2022, MNRAS, 516, 245
Williams, W. L., de Gasperin, F., Hardcastle, M. J. H., et al. 2021, A&A, 655, A40
Ye, H., Sweijen, F., van Weeren, R. J., et al. 2024, A&A, 691, A347
Zwart, J. T. L., Santos, M., & Jarvis, M. J. 2015, MNRAS, 453, 1740

Appendix A: PyBDSF Source detection

Below we detail the multi-step PyBDSF process that we used to generate the catalogue for the ELAIS-N1 deep image.

– Firstly, we run PyBDSF over the image using a similar set up to previous LoTSS Deep data releases that makes use of wavelet decomposition to better fit emission of different scales. However, as we want to remove sources from the image to be able to more accurately measure RMS, we use a $4\sigma_{thresh}$ peak signal-to-noise ratio source threshold (using the PyBDSF parameter `thresh_pix`). From this we save the residual image generated (we denote this as `resid_map_1`). This uses the PyBDSF command:

```
bdsf.process_image(image, detection_image=
    app_image, thresh_isl=3.0, thresh_pix=4.0,
    rms_box=(150,15), rms_map=True, mean_map=
    'zero', ini_method='intensity',
    adaptive_rms_box=True, adaptive_thresh=150,
    rms_box_bright=(60,15), group_by_isl=False,
    group_tol=10.0, output_opts=True,
    output_all=False, atrous_do=True,
    atrous_jmax=4, flagging_opts=True,
    flag_maxsize_fwhm=0.5, advanced_opts=True,
    blank_limit=None, frequency=restfreq)
```

where `image` is the primary beam corrected image, `app_image` is
the image without a primary beam correction applied and `restfreq`
is the frequency of the survey in Hz.

– Next, we inject the brightest ($\geq 150\sigma_{thresh}$) sources back into the `resid_map_1` image to create `resid_map_with_bright`. This is done because PyBDSF uses a smaller sliding box size around bright sources in order to better characterise the elevated noise in these regions. By injecting the sources back into `resid_map_1` we ensure that subsequent PyBDSF runs use the same sliding box behaviour as previous runs.

– Then, we run PyBDSF as above on `resid_map_with_bright` and use this to calculate an RMS map (`rms_map_deep`) which is deeper than that obtained from the initial PyBDSF run because the source density of `resid_map_with_bright` is far lower than our original image.

– To produce a deeper and more complete catalogue we run PyBDSF over the original image using the same parameters as above, except passing it the RMS map (`rms_map_deep`) directly. This creates a residual image which we denote as `resid_map_deep`. The complete configuration run is:

```
bdsf.process image(image, detection_image=
    app_image, thresh_isl=3.0, thresh_pix=4.0,
    rms_box=(150,15), rms_map=True, mean_map=
    'zero', ini_method='intensity',
    adaptive_rms_box=True, adaptive_thresh=150,
    rms_box_bright=(60,15), group_by_isl=False,
    group_tol=10.0, output_opts=True,
    output_all=False, atrous_do=True,
    atrous_jmax=4, flagging_opts=True,
    flag_maxsize_fwhm=0.5, advanced_opts=True,
    blank_limit=None, frequency=restfreq,
    rmsmean_map_filename=[meanmap, rms_map_deep],
    rmsmean_map_filename_det=[meanmap,
    rms_map_deep_app])
```

– Finally, as a visual inspection of the PyBDSF residual maps indicates that complex structures are not always adequately modelled, we run PyBDSF on `resid_map_deep` to pick up emission which may have been missed by PyBDSF in the previous run using the configuration:

```
bdsf.process image(resid_map_deep,
detection_image=
    resid_map_deep_app, thresh_isl=3.0,
    thresh_pix=
    10.0, rms_box=(150,15), rms_map=True,
    mean_map='zero', ini_method='intensity',
    adaptive_rms_box=True, adaptive_thresh=150,
    rms_box_bright=(60,15), group_by_isl=False,
    group_tol=10.0, output_opts=True,
    output_all=False, flagging_opts=True,
    flag_maxsize_fwhm=0.5, advanced_opts=True,
    blank_limit=None, frequency=restfrq,
    rmsmean_map_filename=[meanmap, rms_map_deep],
    rmsmean_map_filename_det=[meanmap,
    rms_map_deep_app], flag_maxsize_bm=100)
```

– We add these sources to the final source and Gaussian component catalogues. We also update the `resid_map_deep` map accordingly so that it reflects the final residual image.