The LOFAR Two-metre Sky Survey - II. First data release

How to cite:


For guidance on citations see FAQs.

© 2019 ESO

Version: Version of Record

Link(s) to article on publisher’s website:
http://dx.doi.org/doi:10.1051/0004-6361/201833559
oro.open.ac.uk
The LOFAR Two-metre Sky Survey


(Affiliations can be found after the references)

Received 4 June 2018 / Accepted 12 September 2018

ABSTRACT

The LOFAR Two-metre Sky Survey (LoTSS) is an ongoing sensitive, high-resolution 120–168 MHz survey of the entire northern sky for which observations are now 20% complete. We present our first full-quality public data release. For this data release 424 square degrees, or 2% of the eventual coverage, in the region of the HETDEX Spring Field (right ascension 10h45m00s to 15h30m00s and declination 45°00′00″ to 57°00′00″) were mapped using a fully automated direction-dependent calibration and imaging pipeline that we developed. A total of 325 694 sources are detected with a signal of at least five times the noise, and the source density is a factor of ~10 higher than the most sensitive existing very wide-area radio-continuum surveys. The median sensitivity is $S_{14{	ext{ MHz}}} = 71\, \mu$Jy beam$^{-1}$ and the point-source completeness is 90% at an integrated flux density of 0.45 mJy. The resolution of the images is 6″ and the positional accuracy is within 0.2″. This data release consists of a catalogue containing location, flux, and shape estimates together with 58 mosaic images that cover the catalogued area. In this paper we provide an overview of the data release with a focus on the processing of the LOFAR data and the characteristics of the resulting images. In two accompanying papers we provide the radio source associations and deblending and, where possible, the optical identifications of the radio sources together with the photometric redshifts and properties of the host galaxies. These data release papers are published together with a further ~20 articles that highlight the scientific potential of LoTSS.

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

1. Introduction

Surveys that probe deeply into new parameter space have enormous discovery potential. The LOFAR Two-metre Sky Survey (LoTSS; Shimwell et al. 2017) is one example: it is an ongoing survey that is exploiting the unique capabilities of the LOw Frequency ARray (LOFAR; van Haarlem et al. 2013) to produce a sensitive, high-resolution radio survey of the northern sky with a frequency coverage of 120–168 MHz (see Fig. 1). The survey was primarily motivated by the potential of low-frequency observations to facilitate breakthroughs in research areas such as the formation and evolution of massive black holes (e.g. Wilman et al. 2008; Best et al. 2014) and clusters of galaxies (e.g. Cassano et al. 2010; Brunetti & Jones 2014). However, there are many other important scientific drivers of the survey, and there is active research in areas such as high redshift radio sources (e.g. Saxena et al. 2017), galaxy clusters (e.g. Botteon et al. 2018; Hoang et al. 2017; de Gasperin et al. 2017; Savini et al. 2018; Wilber et al. 2018), active galactic nuclei (e.g. Brienta et al. 2017; Morabito et al. 2017; Williams et al. 2018), star forming galaxies (e.g. Calistro Rivera et al. 2017), gravitational lensing, galactic radio emission, cosmological studies (Raccanelli et al. 2012), magnetic fields (e.g. Van Eck et al. 2018), transients and recombination lines (e.g. Oonk et al. 2017).
The LoTSS survey is one of several ongoing or recently completed very wide-area low-frequency radio surveys that are providing important scientific and technical insights. Other such surveys include the Multifrequency Snapshot Sky Survey (MSSS; Heald et al. 2015), TIFR GMRT Sky Survey alternative data release (TGSS-ADR1; Intema et al. 2017), GaLactic and Extragalactic All-sky MWA (GLEAM; Wayth et al. 2015; Hurley-Walker et al. 2017), LOFAR Low-band Sky Survey (LoLSS; de Gasperin et al. 2019), and the Very Large Array Low-frequency Sky Survey Redux (VLSSr; Lane et al. 2014). However, LoTSS is designed to push further into new territory. This survey aims to provide a low-frequency survey that will remain competitive even once the Square Kilometre Array (Dewdney et al. 2009) is fully operational, and will not be surpassed as a low-frequency wide-area northern sky survey for the foreseeable future. The LoTSS can provide the astrometric precision that is required for robust identification of optical counterparts (see e.g. McAlpine et al. 2012) and a sensitivity that, for typical radio sources, exceeds that achieved in existing very wide area higher frequency surveys such as the NRAO VLA Sky Survey (NVSS; Condon et al. 1998), Faint Images of the Radio Sky at Twenty-Centimeters (FIRST; Becker et al. 1995), Sydney University Molonglo Sky Survey (SUMSS; Bock et al. 1999; Mauch et al. 2003), and WESterbork Northern Sky Survey (WENSS; Rengelink et al. 1997) and rivals forthcoming higher frequency surveys such as the Evolutionary Map of the Universe (EMU; Norris et al. 2011), the APERture Tile In Focus survey (e.g. Röttgering et al. 2011) and the VLA Sky Survey (VLASS\footnote{https://science.nrao.edu/science/surveys/vlass}). More specifically the primary observational objectives of LoTSS are to reach a sensitivity of less than $100 \mu Jy$ beam$^{-1}$ at an angular resolution, defined as the full width half maximum (FWHM) of the synthesised beam, of $\sim 6''$ across the whole northern hemisphere, using the High Band Antenna (HBA) system of LOFAR (see Fig. 1).

In the first paper of this series (Paper I: Shimwell et al. 2017) we described LoTSS and presented a preliminary data release. In that release the desired imaging specifications were not reached, as no attempt was made to correct either for errors in the beam models or for direction-dependent ionospheric distortions, which are severe in these low-frequency data sets. However, there has since been substantial improvements in the quality, speed, and robustness of the calibration of direction-dependent effects (DDEs) and imaging with the derived solutions (see e.g. Tasse 2014; Yatawata 2015; van Weeren et al. 2016; Tasse et al. 2018). Furthermore, LOFAR surveys of smaller areas of sky have demonstrated that the desired imaging specifications of LoTSS are feasible by making use of direction-dependent calibration (e.g. Williams et al. 2016; Hardcastle et al. 2016). These new insights have facilitated the first full quality public data release (LoTSS-DR1), which we present here in Paper II of this series.

As part of this series we also attempt to enrich our radio catalogues by locating optical counterparts using a combination of likelihood ratio cross matching and visual inspection (discussed in Paper III of this series: Williams et al. 2019). In addition, where counterparts are successfully located, we provide photometric redshift estimates and host galaxy properties (Paper IV: Duncan et al. 2019). In the near future, to improve on the redshifts for many sources, the William Herschel Telescope Enhanced Area Velocity Explorer (WEAVE; Dalton et al. 2012, 2014) multi-object and integral field spectrophotograph will measure redshifts of over a million LoTSS sources as part of the WEAVE-LOFAR survey (Smith et al. 2016).

In Sects. 2 and 3 we describe the observations, the data processing procedure for the present data release, and the quality of the resulting images. In Sect. 4 we give a brief overview of the optical cross matching and the photometric redshift estimation. Finally, we outline some upcoming developments in Sect. 5 before concluding in Sect. 6.

2. Observations and data reduction

We describe the status of LoTSS observations in the first subsection. The second subsection outlines the direction-independent calibration of the data; at present, the main challenge is retrieving and processing the large volume of archived data. The third subsection describes the direction-dependent calibration and imaging, where the focus is on the development and execution of a robust and automated pipeline. The final subsection summarises the mosaicing and cataloguing of the DR1 images.

2.1. Observation status

The ambitious observational objectives for LoTSS are outlined in Fig. 1. To achieve these objectives at optimal declinations, LoTSS observations are conducted in the HBA\textsubscript{Dual\_Inner} configuration with 8 h dwell times and a frequency coverage of 120–168 MHz. The entire northern sky is covered with 3168 pointings. By exploiting the multi-beam capability of LOFAR and observing in 8-bit mode two such pointings are observed simultaneously. As of May 2018, approximately 20% of the data have now been gathered and a further 30% are scheduled over the next two years (see Fig. 2); a total of approximately 13 000 h observing time are required to complete the entire survey with the present capabilities of LOFAR.

As in Shimwell et al. (2017), in this paper we focus on 63 LoTSS data sets (2% of the total survey) in the region of the HETDEX Spring Field that were observed between 2014 May 23 and 2015 October 15. Each 8 h observation was bookended by 10 min calibrator observations (primarily 3C 196 and 3C 295) and the data are archived with a time resolution of 1 s and a frequency resolution of 16 channels per 195.3 kHz sub-band (SB).
by the observatory\(^5\). This high time- and frequency-resolution data is kept to reduce time and bandwidth smearing to a level that is tolerable for future studies that will exploit the international baselines of LOFAR (only antennas within the Netherlands are used for the primary objectives of LoTSS). The high spectral resolution \((R \sim 5000−7000 \text{ or } 22–31 \text{ km s}^{-1} \text{ velocity resolution})\) of the data is also facilitating spectral line (Emig et al. 2019) and spectro-polarimetric studies.

2.2. Direction-independent calibration

The publicly available LOFAR direction-independent calibration procedure was described in detail by van Weeren et al. (2016) and Williams et al. (2016) and makes use of the LOFAR Default Preprocessing Pipeline (DPPP; van Diepen & Dijkema 2016) and Williams et al. (2016) and makes use of the LOFAR direction-independent calibration pipeline on compute facilities amongst LoTSS team members for analysis and further processing. We are therefore working on implementing the direction-independent calibration pipeline each SB of the calibrator is sent to an available compute node where it is flagged for interference with AOFLAGGER (Offringa et al. 2012), averaged to two channels per 195 kHz SB and 8 s, and calibrated using a model of the appropriate calibration source, which has a flux density scale consistent with that described in Scaife & Heald (2012). We note that the Scaife & Heald (2012) flux density scale is consistent with the Perley & Butler (2017) scale to within \(\sim 5\%\) but that there are larger discrepancies (5–10\%) when comparing with the Baars et al. (1977) scale (see Scaife & Heald 2012 and Perley & Butler 2017 for details). Using a single compute node the resulting 244 calibration tables are combined and used to derive time-independent amplitude solutions, XX and YY phase offsets, and clock offsets for each station. Similarly, on separate compute nodes, the 244 single SB target files are each flagged, corrected for ionospheric Faraday rotation\(^8\), calibrated using the calibrator solutions, and averaged to a resolution of two channels per 195 kHz SB and 8 s. In the final step of the direction-independent calibration pipeline, the data for each contiguous 10-SB block are sent to different compute nodes where they are each combined to a single file that is phase calibrated against a sky model for the target field, which is generated from the TGSS-ADR1 catalogue (Intema et al. 2017). This produces 25 10-SB measurement sets for the target field, but the six highest frequency SBs are empty because there are only 244 SBs in the highest frequency measurement set.

For the bulk processing of data on the SURFsara facilities we made use of PiCaS\(^9\), a CouchDB based token pool server for heterogeneous compute environments. The PiCaS server allows millions of tasks to be scheduled on heterogeneous resources to monitor these tasks via a web interface and to provide easy access to logs and diagnostic plots, which helps ensure that our data quality is high. Examples of these diagnostic plots for the HETDEX Spring Field data are shown by Shimwell et al. (2017). We also make use of archiving and distribution facilities at SURFsara, allowing us to store the direction-independent calibrated data products (which are reduced from 16 TB to \(\sim 500\) GB per pointing) and freely distribute these amongst LoTSS team members for analysis and further processing.

The SURFsara Grid processing facilities enable high-throughput processing of large data sets stored on the local LTA site, however the LoTSS data sets are disseminated to all three LTA sites. Since the LTA sites are not linked to each other with a high bandwidth connection, the transfer speed to download data from the Forschungszentrum Jülich and Poznań LTA sites to SURFsara \((\sim 200 \text{ MB s}^{-1})\) is currently a bottleneck in our processing. We are therefore working on implementing the direction-independent calibration pipeline on compute facilities local to each of the LTA sites.

\(^{\text{5}}\) \(\sim 100\) of the early LoTSS observations were averaged to 2 s and 24.4 kHz.

\(^{\text{6}}\) https://github.com/lofar-astron/prefactor using commit dd68c57.

\(^{\text{7}}\) https://lta.lofar.eu/

\(^{\text{8}}\) https://github.com/lofar-astron/RMextract

\(^{\text{9}}\) http://doc.grid.surfsara.nl/en/latest/Pages/Practices/picas/picas_overview.html
2.3. Direction-dependent calibration and imaging

A robust, fast, and accurate calibration and imaging pipeline is essential to routinely create high-fidelity LoTSS images with a resolution of 6′′ and a sensitivity of 100 μJy beam−1. However, the necessity to correct DDEs, which are primarily ionospheric distortions and errors in the station beam model of the HBA phased array stations, adds significant complications to this procedure. These DDEs can be understood in terms of Jones matrices (Hamaker et al. 1996) and to correct for these matrices, which not only depend on direction but also on time, frequency, and antenna, they must be derived from the visibilities and applied during imaging. Various approaches have been developed to estimate the DDE (e.g. Cotton et al. 2004; Interna et al. 2009; Kazemi et al. 2011; Noordam & Smirnov 2010; van Weeren et al. 2016; Yatawatta 2015) but for this work we developed KillMS (kms; Tasse 2014; Smirnov & Tasse 2015) to calculate the Jones matrices and ddf (Tasse et al. 2018) to apply these during the imaging. Our software packages and the pipeline are publicly available and documented10. Below we briefly outline the calibration and deconvolution procedures before describing the pipeline in more detail.

2.3.1. Calibration of direction-dependent effects

One of the main difficulties in the calibration of DDE is the large number of free parameters that must be optimised for when solving for the complex-valued Jones matrices. The consequences of this are that finding the solutions can become prohibitively computationally expensive and that ill-conditioning can introduce systematics in the estimated quantities, which have a negative impact on the image fidelity.

To tackle the computational expense, Salvini & Wijnholds (2014), Tasse (2014), and Smirnov & Tasse (2015) have shown that when inverting the Radio Interferometric Measurement Equation (RIME; see e.g. Hamaker et al. 1996; Smirnov 2011) the Jacobian can be written using Wirtinger derivatives. The resulting Jacobian is remarkably sparse, which allows for short-cuts to be used when implementing optimisation algorithms such as Levenberg–Marquardt (see for example Smirnov & Tasse 2015). In particular, the problem can become antenna separable, and to solve for the Jones matrices associated with a given antenna in kms, only the visibilities involving that antenna are required at each iterative step. The computational gain can be as high as n2 (where n is the number of elementary antennas).

To reduce ill conditioning, kms uses the Wirtzinger Jacobian together with an Extended Kalman Filter (EKF) to solve for the Jones matrices (Tasse, in prep.). Instead of optimising the least-squares residuals as a Levenberg–Marquardt (LM) procedure would, the EKF is a minimum mean-square error estimator and is recursive (as opposed to being iterative). In practice, the prior knowledge is used to constrain the expected solution at a given time. While an LM would produce independent “noisy” estimates, the EKF produces smooth solutions that are more physical and robust to ill-conditioning.

To further improve the calibration, kms produces a set of weights according to a “lucky imaging” technique in which the weights of visibilities are based on the quality of their calibration solutions (Bonnasieux et al. 2018), so visibilities with the worst ionospheric conditions are weighted down in the final imaging.

---

10 See https://github.com/saopicc for kms and ddfacacet, and https://github.com/mhardcastle/ddf-pipeline for the associated LoTSS-DR1 pipeline.

2.3.2. Wide field spectral deconvolution

The ddfacet imager (Tasse et al. 2018) uses the kms-estimated direction-dependent Jones matrices and internally works on each of the directions for which there are solutions to synthesise a single image. To do this, several technical challenges had to be overcome. For example, the dependence of the Jones matrices on time, frequency, baseline, and direction, together with time- and frequency-dependent smearing, lead to a position dependent point spread function (psf). Therefore, although ddfacet synthesises a single image, each facet has its own psf that takes into account the DDE and time and bandwidth smearing whilst ensuring that the correct deconvolution problem is inverted in minor cycles.

Furthermore, to accurately deconvolve the LoTSS images, which have a large fractional bandwidth and a wide field of view, spectral deconvolution algorithms must be used to estimate the flux density and spectra of modelled sources whilst taking into account the variation of the LOFAR beam throughout the bandwidth of the data. The computational cost of this deconvolution can be high and therefore throughout our processing we make exclusive use of the subspace deconvolution (ssd) algorithm, an innovative feature of ddfacet (see Tasse et al. 2018 for a description). As opposed to clean and related algorithms, where a fraction of the flux density is iteratively removed at each major iteration, ssd aims at removing all the flux density at each major cycle. This is done in the abstracted notion of subspaces—in practice islands—each representing an independent deconvolution problem. Each one of these individual subspaces is jointly deconvolved (all pixels are simultaneously estimated) by using a genetic algorithm (the ssd-α, flavour of ssd), and parallelisation is done over hundreds to thousands of islands. A strength of ssd is that we can minimise the number of major cycles, by always recycling the sky model from the previous step. In practice the sky model generated in the preceding deconvolution step of the pipeline is then used to initialise the sky model in the next deconvolution. In other words, a proper dirty image is only formed at the very first imaging step and, thanks to ssd, the LoTSS-DR1 pipeline can work only on residual images and update the spectral sky model at each deconvolution step.

2.3.3. The LoTSS-DR1 pipeline

The LoTSS-DR1 pipeline has many configurable parameters including resumability, taking into account time and bandwidth smearing, bootstrapping the flux density scale off existing surveys, correction of facet-based astrometric errors, user specified deconvolution masks, and substantial flexibility in calibration and imaging parameters. The pipeline is suitable for the analysis of various LOFAR HBA continuum observations, including interleaved observations or those spanning multiple observing sessions. The entire pipeline takes less than five days to image one LoTSS pointing when executed on a compute node with 512 GB RAM (the minimum required for the pipeline is 192 GB) and four Intel Xeon E5-4620 v2 processors, which have eight cores each (16 threads) and run at 2.6 GHz.

The pipeline operates on the direction-independent calibration products which, for each pointing, are 25 10-SB (1.95 MHz) measurement sets with a time and frequency resolution of 8 s and two channels per 195 kHz SB. The pipeline first removes severely flagged measurement sets (those with ≥80% of data flagged) and selects six 10-SB blocks of data that are evenly spaced across the total bandwidth for imaging. This quarter of the data is self-calibrated to gradually build up a model of the
radio emission in the field, which is then used to calibrate the full data set. A brief outline of the steps of LoTSS-DR1, which are shown in Fig. 3, is as follows:

Step 1 Direction-independent spectral deconvolution and imaging (6 $\times$ 10 SB);
Step 2 Sky model tesselation in 45 directions;
Step 3 Direction-dependent calibration (6 $\times$ 10 SB, kms with EKF);
Step 4 Bootstrapping the flux density scale;
Step 5 Direction-dependent spectral deconvolution and imaging (6 $\times$ 10 SB, phase-only solutions, three major cycles);
Step 6 Direction-dependent calibration (6 $\times$ 10 SB, kms with EKF)
Step 7 Direction-dependent spectral deconvolution and imaging (6 $\times$ 10 SB), one major cycle, amplitude, and phase solutions);
Step 8 Direction-dependent calibration (24$\times$10 SB, kms with EKF);
Step 9 Direction-dependent spectral deconvolution and imaging (24$\times$10 SB, two major cycles, amplitude, and phase solutions);
Step 10 Facet-based astrometric correction.

The $\text{facet}$ is used in Step 1 to image the direction-independent calibrated data using the ssd algorithm, which allows us to rapidly deconvolve very large images. The present implementation of ssd requires a deconvolution mask and we use $\text{facet}$ to automatically generate one based on a threshold of 15 times the local noise, which is re-evaluated at every major cycle. The mask created during the deconvolution is supplemented with a mask generated from the TGSS-ADR1 catalogue to ensure that all bright sources in the field are deconvolved even when observing conditions are poor and automatically masking the sources is challenging. The image produced from the 60 SB data set consists of 20 000 $\times$ 20 000 $1.5''$ pixels, has a restoring beam of 12$''$, and the noise varies between 0.25 mJy beam$^{-1}$ and 2 mJy beam$^{-1}$ depending on the observing conditions and source environment. From this image a refined deconvolution mask is created and used to reduce the number of spurious components in the ssd component model of the field by filtering out those that lie outside the region within the refined mask.

At Step 2 the resulting sky model is used to define 45 facets that cover the full 8.3$''$ $\times$ 8.3$''$ region that has been imaged. The ssd component model is used for the first direction dependent
calibration of the 60 SB data set (Step 3). This calibration is done using kms, which creates an amplitude and phase solution for each of the 45 facets every 60 s and 1.95 MHz of bandwidth, and the data are reimaged. Throughout the pipeline, in order not to absorb unmodelled sky emission into the kms calibration solutions (in particular faint extended emission seen by a small number of baselines), we always calibrate the visibilities using only baselines longer than 1.5 km (corresponding to scales of ∼4.5′).

After this initial direction-dependent calibration we bootstrap the LoTSS-DR1 flux densities in Step 4 following the procedure described by Hardcastle et al. (2016). This not only improves the accuracy of our flux density estimates but also decreases amplitude errors that can occur owing to imperfections in the calibration across the bandwidth. In this step each of the six 10-SB blocks imaged in the previous step is imaged separately at lower resolution (20′′) using aoracet which applies the direction dependent phase calibration solutions. A catalogue is made from the resulting image cube using the Python Blob Detector and Source Finder (PyBDSF; Mohan & Rafferty 2015) where sources are identified using a combined image created from all the planes in the cube and the source flux density measurements are extracted from each plane using the same aperture. Sources within 2.5′ of the pointing centre that are at least 100′′ from any other detected source and have a integrated flux density exceeding 0.15 Jy are positionally cross matched with the VLSSr and WENSS catalogues using matching radii of 40′′ and 10′′, respectively. The WENSS catalogue used has all the flux densities scaled by a factor of 0.9 which, as described by Scait & Heald (2012), brings it into overall agreement with the flux density scale we use. Correction factors are then derived for each of the six 10-SB blocks to best align the LoTSS-DR1 integrated flux density measurements with VLSSr and WENSS assuming the sources have power-law profiles across this frequency range (74 MHz–325 MHz). During the fitting, sources that are poorly described by a power law are excluded to remove, for example incorrect matches or sources with spectral curvature. From the 70±14 matched sources per field the correction factors derived for each of the six 10-SB blocks are typically 0.85±0.1 and these are extrapolated linearly to the entire 25 10-SB data set. The six 10-SB 20′′ resolution images are also stacked to provide a lower resolution (20′′) image that has a higher surface brightness sensitivity than the higher resolution images. This image is used to identify diffuse structures that are prevalent in LOFAR images, but may not be detected at sufficient significance in the higher resolution imaging. These extended sources are then added to the mask to ensure that they are deconvolved in later imaging steps. Sources are classified as extended sources if they encompass a contiguous region larger than 2000 pixels with all pixels having a signal above three times the local noise of the image.

After the bootstrap derived corrections factors are applied the 60 SBs of data are imaged with the direction-dependent phase solutions applied in aoracet in Step 5. As explained above, for efficiency reasons ssd is initiated with the ssd components from the direction-independent imaging, which allows us to deconvolve deeply with three major ssd iterations. The image size and resolution are the same as in Step 3 but the input mask is improved because it is a combination of that obtained from the direction-independent imaging, the mask generated from the TGSS-ADR1 catalogue, and the low-resolution mask created from the bootstrapping: at this point the auto-masking threshold is also lowered to ten times the local noise. Again, once the imaging is complete the image is masked and the mask is used to reduce spurious entries in the ssd component model. The noise levels in this second imaging step range from 130 µJy beam−1 to 600 µJy beam−1. In Step 6 this new model is input into kms which calculates improved direction-dependent calibration solutions for each of the 45 facets every 60 s and 1.95 MHz of bandwidth.

A third imaging step is performed on the 60 SBs of data (Step 7), this time applying both the phase and amplitude direction-dependent calibration solutions but otherwise following the same procedure as before. This produces images with noise levels ranging from 100 µJy beam−1 to 500 µJy beam−1 and a final ssd component model that is used to calibrate the entire 240 SBs of the data set with kms (Step 8).

The full bandwidth is imaged at both low and high resolution in aoracet with the newly derived phase and amplitude solutions applied (Step 9). The low-resolution image has a resolution of 20′′ and a significantly higher surface brightness sensitivity than when imaging at higher resolution. In this low-resolution image ssd is not initiated with a previously derived model because the uv-data used in the imaging are different as an outer a cut of 25.75 km is applied. To deconvolve deeply we perform three separate iterations of the low-resolution imaging, each time improving the input mask and lowering the automasking threshold. The noise level of the final 20′′ resolution images ranges from 100 µJy beam−1 to 400 µJy beam−1, which corresponds to a brightness temperature of 9 K–35 K.

The full bandwidth high-resolution imaging is performed with a resolution of 6′′. The deconvolution mask that has been gradually built up through the self-calibration of the 60 SB data set, as well as that from the lower resolution imaging from the full bandwidth, and an auto masking threshold in aoracet of five times the local noise allow for a very deep deconvolution. This is performed with two separate runs of aoracet with a masking step in between to ensure that the local noise is well estimated and faint sources (signal-to-noise (S/N) ≥5) are masked. The resulting high-resolution images have noise levels that vary from 60 µJy beam−1 to 160 µJy beam−1. Once the deconvolution is complete the images are corrected for astrometric errors in aoracet which can apply astrometric corrections to each of the facets independently (Step 10). The astrometric corrections applied vary from 0.9′′ to 4.4′′ with a median of 0.8′′ and are derived from cross-matching the LOFAR detected sources in each facet with the Pan-STARRS catalogue (Flewelling et al. 2016). The errors on the derived offsets vary from 0.1′′ to 4.8′′ with a median 0.2′′.

During the cross-matching a histogram of the separations between all Pan-STARRS sources within 60 arcsec of compact LOFAR sources is made for each facet. This typically consists of ~140 Pan-STARRS sources per LoTSS-DR1 source and an average of 190 radio sources per facet. If all sources in the facets are systematically offset, then this histogram should have a peak at the value of the offset between the LoTSS-DR1 and Pan-STARRS sources. To search for the location of this peak and estimate the RA and Dec offsets and their corresponding errors in each facet, we use a Markov Chain Monte Carlo (MCMC) method and uninformative priors. In this procedure the emcee package (Foreman-Mackey et al. 2013) is used to draw MCMC samples from a Gaussian function plus a background where the initial parameter estimates are derived from the observed LOFAR and Pan-STARRS position offsets. The likelihood function is calculated using a gamma distribution with a shape parameter defined by the observed LOFAR and Pan-STARRS position offsets. The posterior probability distribution is calculated taking into account the uninformative priors (background, offset, and Gaussian peak greater than zero and a Gaussian
standard deviation less than 5′′) that are put on the offset Gaussian function and background level.

The pipeline is very robust and with no human interaction the processing failed for only 5 of the 63 fields in the HETDEX Spring Field region, thus providing 58 images in this region. One (P2) of these failures was due to exceptionally bad ionospheric conditions and the other four (P31, P210+52, P214+52, and P215+50) were due to the proximity of very bright sources (3C 280 and 3C 295).

2.4. Mosaicing and radio source cataloguing

The LoTSS pointings tile the sky following a spherical spiral distribution (Saff & Kuijlaars 1997); they are typically separated by 2.58′ and have six nearest neighbours within 2.8′. With the FWHM of the HBA Dual Inner station beam being 3.40′ and 4.75′ at the top (168 MHz) and bottom (120 MHz) of the LoTSS frequency coverage, respectively, there is significant overlap between the pointings. To produce the final data release images, a mosaic has been generated for each of the 58 pointings that was successfully processed. For each pointing the images of the (up to six) neighbouring pointings are reprojected to the frame of the central pointing using the astrometry-based reproject code and then all seven (or fewer) pointings are averaged using the appropriate station beam and the central image noise as weights in the averaging. During the mosaicing of the high-resolution images, facets with uncertainties in the applied astrometric corrections (derived as described in Sect. 2.3.3) larger than 0.5′′ are excluded to ensure that the final maps have a high astrometric accuracy. This criterion is also a good proxy for image quality and allows us to identify and remove any facets that diverged during the processing due to poor calibration solutions. Once the images of the neighbouring pointings are combined the mosaiced map is blanked to leave just the pixels that lay within the 0.3 power point of the station beam of the central pointing. An example region from a mosaic is shown in Fig. 4 and the noise map of the entire mosaiced region is shown in Fig. 5.

To produce a catalogue of the radio sources we performed source detection on each mosaic using PyBDSF. The sources were detected with a 5σ peak detection threshold and a 4σ threshold to define the boundaries of the detected source islands that were used for fitting. The background noise variations were estimated across the images using a sliding box algorithm, where a box size of 30 × 30 synthesised beams was used except in the regions of high S/N sources (≥150) where the box size was decreased to just 12 × 12 synthesised beams; this box size was tuned to more accurately capture the increased noise level in these regions. The PyBDSF wavelet decomposition functionality was also utilised to better characterise the complex extended emission in the images. The resulting catalogues of the individual mosaics were combined and duplications were removed by only keeping sources that are detected in the mosaic to which they are closest to the centre.

In the concatenated catalogue the columns kept from the PyBDSF output are the source position, peak brightness, integrated flux density, source size and orientation, and the statistical errors from the source fitting for each of these. In addition we keep the source code which describes the type of structure fitted by PyBDSF (see Table 1 caption for the definition of these) and the local root mean square noise estimate. We append columns that provide the mosaic identity, number of pointings that contribute to the mosaic at the position of the source, fraction of those in which the source was in the deconvolution mask, and whether or not the source is believed to be an artefact (see Williams et al. 2019 for a description of artefact identification). The fraction of the source in the deconvolution mask is calculated by finding the mask value (1 or 0) at the centre of each Gaussian component for every source in all of the contributing pointings and using the effective integration times to calculate the weighted average. To find the masked fraction for a source that consists of multiple Gaussian components, we use the integrated flux densities of each component as weights and assign the weighted average of the masked fraction of these components to the source. These final parameters, together with the mosaiced residual images, which are also provided, allow users to assess the quality of the deconvolution for sources. This is particularly important for faint sources that may not be in the masks and also for extended sources where, because of the integral of the dirty beam exceeding that of the restoring beam, the apparent flux density in dirty images is substantially larger than in deconvolved images. Example entries from the catalogue are shown in Table 1 and a selection of some of the more spectacular sources in our images are represented in Fig. 6.

3. Image quality

The observations used in this data release were conducted between 2014 May 23 and 2015 October 15 and the varying observing conditions significantly impact the image quality even after direction-dependent calibration, which reduces the impact of ionospheric disturbances. In this section, we assess the derived source sizes, astrometric precision, flux-density uncertainty, dynamic-range limitations, sensitivity, and completeness, and briefly discuss some remaining calibration and imaging artefacts.

3.1. Source extensions

Identifying unresolved sources using the PyBDSF-derived measurements is complicated by several factors. For example, astrometric errors in the mosaiced images cause an artificial broadening of sources, the varying quality of calibration blurs the sources by differing amounts, time averaging and bandwidth smearing can artificially extend sources, and the extent to which a source is deconvolved impacts its measured size. To accurately quantify all this would require realistic simulations in which compact sources are injected into the uv-data taking into account DDEs. Furthermore, as the precise criteria for distinguishing resolved sources varies between facets and observations, a prohibitively large number of these simulations would need to be performed. Our calibration and imaging pipelines are continuing to evolve and hence such a large undertaking is beyond the scope of this present study. An alternative approach would have been to inject point sources into our maps and use these to characterise the source finding algorithm; however, such a simulation would not account for distortions in source morphologies caused by calibration inaccuracies. Instead we attempted to assess whether or not sources are resolved by looking at the extensions of real sources that we assert are unresolved and we used these to define an average criterion with which additional unresolved sources can be identified across the entire mosaic.

To create a sample of unresolved sources the LoTSS-DR1 catalogue was first filtered to contain only isolated sources, which we define as being sources with no other LoTSS-DR1 source within 45′. Any sources that were not in the deconvolution mask in every pointing in which they are detected were also excluded. From the remaining entries we then selected only sources that are classified as “S” by PyBDSF; this source
Fig. 4. Top figure: example of a LoTSS-DR1 image; bottom figures: same region in NVSS (left) and FIRST (right). The black and red circles overlaid on the FIRST image show FIRST and TGSS-ADR1 sources, respectively. In this region there are 689 LoTSS-DR1 sources, 71 FIRST sources, 46 NVSS sources, and 16 TGSS-ADR1 sources. The resolution of the LoTSS-DR1 image is 6′′ and the sensitivity in this region is approximately 70 µJy beam⁻¹. This field is dominated by the spectacular galaxy NGC 4258, which in the LoTSS-DR1 image has an extent of over 3000 synthesised beams, together with the smaller edge-on spiral galaxy NGC 4217.
code corresponds to sources that are the only objects within a PyBDSF island and are well fit with a single Gaussian. Finally, as described below, we imposed a cut on the major axis of the LoTSS-DR1 sources to limit the maximum extent of the low-frequency emission.

We emphasise that, owing to imperfect calibration, most truly unresolved sources in the LoTSS-DR1 catalogue do not have an integrated flux density to peak brightness ratio of 1.0 or a fitted major axis size of 6″ (i.e. a size equal to the restoring beam). For example, the approximately 50 seemingly compact, bright (S/N in excess of 500) sources that meet the above criteria all have measured sizes in the FIRST catalogue of less than 5″ and we can therefore assert these are either unresolved or barely resolved. However, in the LoTSS-DR1 catalogue these sources have a median ratio of the integrated flux density to peak brightness equal to 1.12 with a median absolute deviation of 0.04. Furthermore, for seemingly compact LoTSS-DR1 sources that are detected with a lower S/N there is significantly more variation in the measured integrated flux density to peak brightness ratio. To characterise this, and separate extended from compact sources, we derived an envelope with the functional form \( \frac{S_{\text{int}}}{S_{\text{peak}}} = 1.25 + A(\frac{S_{\text{peak}}}{\text{rms}})^B \), which encompasses 95% of the LoTSS-DR1 sources that meet the above criteria (see Fig. 7). The factor of 1.25 was derived from the median plus three times the median absolute deviation of the integrated flux density to peak brightness ratio of the seemingly compact high S/N (≥500) sources. We used this envelope to define a boundary between compact and extended sources.

The fitted envelope is dependent upon the cut used on the major axis of the LoTSS-DR1 sources and we explored the impact of this by varying that selection criterion from 10″ to 20″ (see Fig. 7). We find that this has little impact on the classification of sources with S/N of more than 100 as either extended or compact; however, it has a much larger impact on sources with lower S/Ns. Whilst there is no definite value to use for this cut, we chose a 15″ limit on the LoTSS-DR1 major axis, which gives a best fit envelope of \( \frac{S_{\text{int}}}{S_{\text{peak}}} = 1.25 + 3.1(\frac{S_{\text{peak}}}{\text{rms}})^{0.53} \). There are a total of 280 000 LoTSS-DR1 sources within this envelope and we define these as compact. As a cross check we note that 19 500 of these sources correspond to entries in the FIRST catalogue and in that catalogue 88% of them are less than 5″ in size, indicating that they are also compact at higher frequencies.

### 3.2. Astrometric precision

The astrometry of our images is originally set by our phase calibration based on the TGSS-ADR1 catalogue. However, during direction-dependent calibration the astrometry can shift between regions because of the varying precision of the calibration models that are built up in different facets. For example, after direction-dependent calibration of a LOFAR data set Williams et al. (2016) found −1″ offsets that varied systematically across their field, but they were able to correct these using the positions in the FIRST catalogue to provide a LOFAR HBA image with a standard deviation in the RA and Dec offsets from FIRST of just 0.4″. In our processing we also refined the astrometric accuracy after the self-calibration cycle is complete by correcting each facet independently using positions in the Pan-STARRS optical catalogue. Furthermore, during the mosaicing we do not include facets that have an uncertainty in the estimated astrometric correction of greater than 0.5″ to ensure high astrometric accuracy (see Sect. 2).

To determine the resulting astrometric accuracy of our mosaic catalogue we performed a simple nearest neighbour cross match in which we took the closest Pan-STARRS, WISE, and FIRST counterpart that lies within 5″ of each of the compact LoTSS-DR1 sources that were identified using the procedure described in Sect. 3.1. We then created histograms of the RA and Dec offsets and fit these with a Gaussian, where the location of the peak and the standard deviation correspond to our systematic position offset and the total uncertainty; these total errors are a combination of errors in the LoTSS-DR1 positions from the source finding software, the real astrometric errors in the LoTSS-DR1 positions, and the errors in the positions of objects in the cross-matched surveys (which were selected owing to their high astrometric accuracy). The astrometry of the Pan-STARRS

---

**Fig. 5.** Noise image of the LoTSS-DR1 where the median noise level is 71 µJy beam⁻¹. As described in Sect. 3.4 many of the regions with high noise levels are caused by dynamic-range limitations. Sources from the revised 3C catalogue of radio sources (Bennett 1962) are overplotted as black circles to show the location of potentially problematic objects.
Example of entries in the LoTSS-DR1 source catalogue.

<table>
<thead>
<tr>
<th>Source ID</th>
<th>RA (h:m:s)</th>
<th>Dec (°:′′)</th>
<th>S$_{\text{peak}}$ (mJy beam$^{-1}$)</th>
<th>S$_{\text{int}}$ (mJy beam$^{-1}$)</th>
<th>S$_{\text{max}}$ (mJy beam$^{-1}$)</th>
<th>S$_{\text{min}}$ (mJy beam$^{-1}$)</th>
<th>DC$_{\text{Major}}$ (°)</th>
<th>DC$_{\text{Minor}}$ (°)</th>
<th>PA (°)</th>
<th>Type</th>
<th>Mosaic</th>
<th>Number Pointings</th>
<th>Masked Fraction</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILTJ1018.07</td>
<td>04:00:07.9</td>
<td>-00:00:00</td>
<td>0.1 ± 0.1</td>
<td>0.1 ± 0.1</td>
<td>0.1 ± 0.1</td>
<td>0.1 ± 0.1</td>
<td>0.1 ± 0.1</td>
<td>0.1 ± 0.1</td>
<td>0.1 ± 0.1</td>
<td>S</td>
<td>P291+55</td>
<td>2</td>
<td>0.00</td>
<td>×</td>
</tr>
<tr>
<td>+491636.3</td>
<td>191.621</td>
<td>0.6 ± 0.4</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>S</td>
<td>P191+55</td>
<td>4</td>
<td>0.00</td>
<td>×</td>
</tr>
<tr>
<td>+492701.3</td>
<td>191.621</td>
<td>0.6 ± 0.4</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>S</td>
<td>P161+55</td>
<td>3</td>
<td>0.00</td>
<td>×</td>
</tr>
<tr>
<td>+492701.3</td>
<td>191.621</td>
<td>0.6 ± 0.4</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>S</td>
<td>P161+55</td>
<td>3</td>
<td>0.00</td>
<td>×</td>
</tr>
<tr>
<td>+492701.3</td>
<td>191.621</td>
<td>0.6 ± 0.4</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>S</td>
<td>P161+55</td>
<td>3</td>
<td>0.00</td>
<td>×</td>
</tr>
<tr>
<td>+492701.3</td>
<td>191.621</td>
<td>0.6 ± 0.4</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>S</td>
<td>P161+55</td>
<td>3</td>
<td>0.00</td>
<td>×</td>
</tr>
<tr>
<td>+492701.3</td>
<td>191.621</td>
<td>0.6 ± 0.4</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>S</td>
<td>P161+55</td>
<td>3</td>
<td>0.00</td>
<td>×</td>
</tr>
<tr>
<td>+492701.3</td>
<td>191.621</td>
<td>0.6 ± 0.4</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>S</td>
<td>P161+55</td>
<td>3</td>
<td>0.00</td>
<td>×</td>
</tr>
<tr>
<td>+492701.3</td>
<td>191.621</td>
<td>0.6 ± 0.4</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>S</td>
<td>P161+55</td>
<td>3</td>
<td>0.00</td>
<td>×</td>
</tr>
<tr>
<td>+492701.3</td>
<td>191.621</td>
<td>0.6 ± 0.4</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>S</td>
<td>P161+55</td>
<td>3</td>
<td>0.00</td>
<td>×</td>
</tr>
<tr>
<td>+492701.3</td>
<td>191.621</td>
<td>0.6 ± 0.4</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>S</td>
<td>P161+55</td>
<td>3</td>
<td>0.00</td>
<td>×</td>
</tr>
</tbody>
</table>
catalogue was determined using a combination of 2MASS and Gaia positions and the typical standard deviation of the offsets from Gaia positions is less than 0.05″ (Magnier et al. 2016). The WISE catalogue has a positional uncertainty of 0.2″ (Cutri et al. 2012) in RA and Dec with respect to the 2MASS Point Source Catalog for sources detected at high significance, and the FIRST survey has astrometric uncertainties of 0.1″ with respect to the absolute radio reference frame (White et al. 1997).

We cross-matched a total of 7100 sources from the LoTSS-DR1 catalogue to all three comparison sources and we found...
that, for these sources, there is a systematic positional offset from Pan-STARRS of less than 0.03″ and the standard deviation of the offsets is less than 0.2″ in both RA and Dec (see Fig. 8). Similarly, in comparison to WISE, we found the same sources have a systematic offset of less than 0.01″ and a standard deviation of less than 0.27″ in both RA and Dec. When comparing to FIRST, the systematic offsets are less than 0.02″ and the standard deviation is approximately 0.3″ in both RA and Dec. The direction of the derived systematic offsets varies when comparing the LoTSS positions with the three different surveys. We also examined the astrometric accuracy of our mosaic catalogue as a function of the LoTSS-DR1 peak brightness. We checked the accuracy of the catalogue to better estimate the real astrometric errors in the LoTSS-DR1 positions as bright (≥20 mJy), compact sources typically have errors in their derived positions of less than 0.05″. For the compact LoTSS-DR1 sources above 20 mJy the fitted standard deviation to a Gaussian of the RA and Dec (FWHM (σ ≈ 0.2″)) of the Gaussian functions that are fitted to the histograms of the offsets. Similar plots showing the same LoTSS-DR1 sources cross-matched with WISE or FIRST sources show comparable systematic offsets and standard deviations of less than 0.27″ and 0.3″, respectively.

3.3. Accuracy of the flux density scale

Owing to inaccuracies in the existing LOFAR beam models, transferring amplitude solutions derived from calibrators to the target field data does not generally result in an accurate flux density scale for the target field. For example, Hardcastle et al. (2016) found the errors in the flux density scale to be up to 50% and that the peak of the fitted Gaussian functions were displaced by between 0.05″ and 0.12″. These numbers give an indication of the varying astrometric accuracy across the HETDEX Spring Field region. We note that, as was found in the mosaiced images, these astrometric errors vary with the S/N of the detections and this explains why the individual pointings have apparently larger astrometric errors than the mosaiced images.

We found that the standard deviation of the Gaussian fitted to a histogram of the RA and Dec astrometric offsets from Pan-STARRS varied from 0.31″ to 0.54″ with an average of 0.39″ and that the peak of the fitted Gaussian functions were displaced by between 0.05″ and 0.12″. These numbers give an indication of the varying astrometric accuracy across the HETDEX Spring Field region. We note that, as was found in the mosaiced images, these astrometric errors vary with the S/N of the detections and this explains why the individual pointings have apparently larger astrometric errors than the mosaiced images.
that in some circumstances the flux density scale can drift during the processing; however, 60% percent of the fields remain within 5% of the original bootstrapped derived values.

For further verification of the flux density scale we compared the catalogued integrated flux density in the compact source LoTSS-DR1 catalogue to those in the TGSS-ADR1 catalogue. The TGSS-ADR1 measurements were not used during the bootstrapping to allow for this comparison. Furthermore, the TGSS-ADR1 flux density scale is not tied to the flux density scales of VLSSr or WENSS as the survey was instead calibrated directly against bright, well-modelled sources, on the Scaife & Heald (2012) flux density scale. For the 835 compact sources with LoTSS-DR1 integrated flux densities in excess of 100 mJy the median ratio of the integrated LoTSS-DR1 flux densities to the TGSS-ADR1 flux densities is 0.94 and the standard deviation of 0.14 (see the left panel of Fig. 9). However, at integrated flux densities below 100 mJy, where the point-source completeness of the TGSS-ADR1 catalogue decreases to less than 90% and detections are not always at very high significance, there is substantially more scatter in the ratio of TGSS-ADR1 to LoTSS-DR1 integrated flux densities with a standard deviation of 0.27.

Part of the scatter in the TGSS-ADR1 and LoTSS-DR1 integrated flux density ratios is from variations in the quality of the images of various LoTSS-DR1 pointings. To examine the consistency of our measurements we compared the integrated flux density of compact sources in catalogues derived from each of the pointings used in LoTSS-DR1 with the TGSS-ADR1 catalogue. The median ratio of the LoTSS-DR1 integrated flux densities to the TGSS-ADR1 integrated flux densities varies from 0.75 to 1.15 with a median of 1.0 and a standard deviation of 0.08. The discrepancy between this median integrated flux density ratio, which is derived from individual LoTSS-DR1 pointings, and corresponding value for the entire mosaic (0.94), appears to be a consequence of the mosaicing. Sources with apparently low LoTSS-DR1 integrated flux densities more often reside in pointings with apparently low noise levels that are more highly weighted during the mosaicing procedure. Furthermore, we made use of the large overlap between pointings to examine flux density scale variations and found that the standard deviation of the median ratio of the integrated flux density between pointings is 0.2 and, whilst the maximum discrepancy in the integrated flux density measurements is 55%, 80% of the ratios are within 20% of unity.

We also searched for trends between the source integrated flux density measurements and the distance from the LoTSS pointing centres (see Fig. 9). Using the 835 bright compact sources in the mosaic catalogue that were cross-matched with TGSS-ADR1 we found no strong dependence of the ratio of the LoTSS-DR1 integrated flux density to the TGSS-ADR1 integrated flux density on the distance from the closest LoTSS pointing: the inner bin has a ratio of 0.95 and the outer has a ratio of 0.92. For the peak brightness the radial dependence is slightly stronger with the inner bin at 0.86 and the outer bin at 0.81. To assess the impact at further distances we look at the peak brightness to integrated flux density ratio of compact sources in the LoTSS-DR1 catalogues derived from individual pointings. Given that our data are averaged to two channels per SB and 8 s, it may be expected that time-averaging and bandwidth-smearing effects are non-negligible in the LoTSS-DR1 mosaics; for example, we estimate using the formulas given by Bridle & Schwab (1989) that at 6′ resolution the time-averaging and bandwidth smearing are as shown in Fig. 10. However, oraract has a facet-dependent PSF which, for deconvolved sources, accounts for the impact of smearing. As a result the ratio of the peak brightness to integrated flux density in our LoTSS-DR1 images does not have as strong a dependence on distance from the nearest pointing centre as found in other studies that used imagers that do not...
was determined to create an average noise profile as a function of distance. These average noise profiles for various integrated flux density ranges are shown in Fig. 11.

The area in square degrees of sky that surrounds bright sources and has a noise level more than 15% higher than the noise in the wider region depends on the source integrated flux density according to approximately $0.1(e^{-0.007S} − 1) = 0.002$, where $S$ is the integrated flux density in mJy. From this equation, and removing overlapping regions, we calculated that the noise is limited by the dynamic range of our maps (i.e. the noise is more than 15% higher than the noise level in regions uncontaminated by bright sources) for 32 square degrees of the 424 square degrees that were imaged, i.e. 8% of the total area of the survey. Similarly, we calculated the area with even more enhanced noise levels of 50% and 100% higher than the noise level in contaminated regions as 3% and 2%, respectively.

3.5. Sensitivity

The latitude of LOFAR is 52°54′32″, putting the HETDEX Spring Field region, which has a declination ranging from 47° to 55°, close to the optimal location where the projected area of the HBA dipoles and hence the sensitivity of the array is at its highest. The entire LoTSS-DR1 6″ resolution mosaic of the HETDEX Spring field region covers an area of 424 square degrees and the median noise level across the mosaic is $71 \mu$Jy beam$^{-1}$; 65%, 90%, and 95% of the area has noise levels below $78 \mu$Jy beam$^{-1}$, $115 \mu$Jy beam$^{-1}$, and $147 \mu$Jy beam$^{-1}$, respectively (see Fig. 12). These variations are due to varying observing conditions, telescope performance (e.g. missing stations or a higher level of interference), pointing strategy, and imperfections in the calibration and imaging procedure. The impact of the calibration and imaging procedure is particularly evident around bright sources in which the noise is limited by the dynamic range, as discussed in Sect. 3.4. The variations due to the observing conditions are also significant and the noise level on images of the individual pointings varies from $60 \mu$Jy beam$^{-1}$ to $160 \mu$Jy beam$^{-1}$. The sensitivity variations due to the mosaicking strategy in this region are much smaller. We find that the average mosaic noise as a function of distance from the closest pointing centre (just including regions covered by more than one pointing) only varies from $72 \mu$Jy beam$^{-1}$ to $78 \mu$Jy beam$^{-1}$ with a minimum at ~1° from a pointing centre and a maximum at ~55° from the nearest pointing centre. By comparison, the LoTSS-DR1 20″ resolution mosaic has higher noise levels due to the uv-cut applied in the imaging step. In this case the median noise level is $132 \mu$Jy beam$^{-1}$, 65%, 90%, and 95% of the area has noise levels below $147 \mu$Jy beam$^{-1}$, $223 \mu$Jy beam$^{-1}$, and $302 \mu$Jy beam$^{-1}$, respectively.

The contribution of confusion noise to the total noise level that is measured on our 6″ resolution images is also small. To quantify this we followed the approach of Franzen et al. (2016) and injected a broken power-law distribution of point sources convolved with a 6″ Gaussian into a blank image. As in Franzen et al. (2016) the power law used for sources with an integrated flux density in excess of 6 mJy was \( S = 69988 S^{1.54} \text{Jy sr}^{-1} \) in agreement with Euclidean normalised differential counts at 154 MHz derived by Intema et al. (2011), Ghosh et al. (2012), and Williams et al. (2013). For fainter sources we fitted a power law of \( S = 82 S^{-2.41} \text{Jy sr}^{-1} \) to the deep 150 MHz counts presented in Williams et al. (2016) and, whilst these counts reach a depth of 700 \( \mu \text{Jy} \), for simplicity we assumed they hold to an integrated flux density limit of 10 \( \mu \text{Jy} \). Given that the counts are thought to decrease towards

**Fig. 10.** Integrated flux density to peak brightness ratios of compact LoTSS-DR1 sources as a function of separation from the pointing centre for catalogues derived from individual LoTSS-DR1 pointings. The thick red curve shows the approximate ratio expected from time and bandwidth smearing assuming that unresolved sources have a ratio of 1.0. The effects of time and bandwidth smearing are taken into account during deconvolution in nOracet. The red points show the median ratios within bins of distance; the 95% confidence intervals are ~0.02 and were estimated by the bootstrap method. The horizontal errors bars give the bin width and the vertical dashed line shows the median distance between LoTSS pointings.

The dynamic range in our images is limited and bright sources have an impact on the image noise properties in a non-negligible fraction of the area that has been mapped. Whilst there are many factors that impact the dynamic range, our testing of the data processing procedure has indicated that the amplitude normalisation scheme that we used certainly plays a significant role. Other contributors include the layout and size of the facets and the quality of the models that are built up during the self-calibration procedure.

To assess the dynamic-range limitations we examined pixels on mosaics of the final nOracet residual images in 5″ wide annuli around compact LoTSS-DR1 sources that were identified in Sect. 3.1. A profile of the pixel standard deviation within every annulus was determined for each of these sources out to a radius of 500″. Each profile was fit with a Gaussian function plus a constant, which we assume is the level of the noise in the surrounding region and we used this to normalise the measurements. Within each distance bin, we averaged together all normalised noise measurements of sources within a given integrated flux density range and the mean and standard deviation
such low flux densities (e.g. Wilman et al. 2008) this should result in a conservative estimate for the confusion noise. From the pixel values in the simulated image we derived the probability of deflection [P(D)], which is highly skewed with an interquartile range of 18\% ability of deflection [P(D)], which is highly skewed with an interquartile range of 18\% ability of deflection [P(D)], which is highly skewed with an interquartile range by a factor of 1.349, which gives a confusion noise estimate at 6\" of 14\mu Jy beam\(^{-1}\), which is significantly lower than the rms levels obtained. Our lower resolution images, however, are much more severely impacted by confusion noise and when repeating the analysis at 20\" our confusion noise estimate is 85\mu Jy beam\(^{-1}\). We note that the very faint sources do not have a large impact on the sigma for the P(D) distributions; for example assuming the counts instead extend to 1\mu Jy assumes 5.1 million sources rather than 200 000 sources per square degree but increases the 20\" resolution confusion noise estimate by only 5\% to 89\mu Jy beam\(^{-1}\). The power-law indices assumed in the calculations, however, play a more significant role; for example, again following Franzen et al. (2016), if for the sources between 10\mu Jy and 6\mu Jy we assume \(\frac{dN}{dS} = 6998 S^{-1.54}, 1841 S^{-1.8}, 661.8 S^{-2.0}\) or 237.9 S\(^{-2.2}\) Jy\(^{-1}\) sr\(^{-1}\) we estimate 20\" resolution P(D) sigma values of 1\mu Jy beam\(^{-1}\), 10\mu Jy beam\(^{-1}\), 24\mu Jy beam\(^{-1}\), and 47\mu Jy beam\(^{-1}\).

Several of the early LoTSS observations were conducted in a manner in which two neighbouring pointings were observed simultaneously, including 10 observations (thus 20 pointings) in this data release. In these circumstances a minor impact on the sensitivity in the overlapping regions of the simultaneously observed pointings is correlated noise. In an attempt to quantify the impact we examined pixel values in the overlapping regions of pointings by reprojecting the images to a common frame and ignoring regions containing sources (defined as those with values more than 1\sigma). The Pearson correlation coefficient calculated from these noise pixels was generally found to be 0.03–0.13 for pointings observed simultaneously but typically less than 0.03 for pointings observed at separate times. We also compared noise levels in mosaiced regions that contained data from two simultaneously observed pointings with regions where all contributing pointings were observed at different times. We found that regions where simultaneously observed pointings contribute to the mosaics have a median noise level that is ~2\% higher than other regions. The LoTSS observations of neighbouring pointings have not been conducted simultaneously since these very early observations.

3.6. Completeness

To thoroughly estimate the completeness of the survey, sources of varying flux densities and positions should be injected into simulated data sets that include realistic DDEs. However, in the absence of such simulations, we instead inject sources into the direction-independent calibrated data sets taking into account the direction-dependent corrections that are applied in that specific direction to correct for the ionospheric and beam errors. After these data sets are processed with the pipeline and the injected sources are catalogued and their properties are compared to the parameters of the sources that were injected. We note that this procedure assumes that the direction-dependent corrections, with which the fake sources are injected, accurately describe the real DDEs. Given the computational cost of our calibration and imaging and that our pipelines will be improved for future data releases, we only performed this analysis for 10 SBs of data from one pointing following the procedure outlined below:

Step 1 Obtain the final direction dependent calibration solutions from a 240 SB run of the LoTSS-DR1 pipeline;

Step 2 Create a simulated image of 300 delta functions drawn from a power-law distribution \(\frac{dN}{dS} \propto S^{-1.6}\) and use the model to predict the visibilities for this model, corrupted by the same direction dependent distortions and add these to real direction independent calibrated data in the 10-SB data set;

Step 3 Execute the LoTSS-DR1 pipeline on the 10-SB simulated data set;

For comparison, following the approach described in Heald et al. (2015), we also estimated the completeness by injecting 300 point-like sources with integrated flux densities drawn from a power-law distribution \(\frac{dN}{dS} \propto S^{-1.6}\) into the final restored image of 1 of the 10SB runs produced in Step 3. To improve the statistics the realistic simulations were
Fig. 12. Estimated noise variations on the direction-dependent calibrated LoTSS-DR1 images. The red line shows the cumulative area of the mosaiced region that has an estimated noise less than a given value. The histogram shows the distribution of noise estimates within the mosaiced region.

Fig. 13. Estimated point-source completeness for a 10 SB (1/24th of the data) for a single LoTSS-DR1 pointing. The red line shows the completeness above a given integrated flux density and the blue line shows the fraction of sources detected at a specific integrated flux density value. Because of the large number of sources injected during the simulation the Poisson errors are negligible but the errors bars reflect the standard deviation of the measurements as a function of position across the mosaic.

However, performing such simulations with 10 SBs of data from a single pointing suggests that the shape of completeness curves derived from realistic simulations is similar to that obtained from injecting sources into calibrated images (Fig. 13). Therefore, to approximate the completeness of the entire LoTSS-DR1 we only used the less computationally expensive approach of injecting point sources into residual images.

From each of the 58 mosaic images a residual image is generated using PyBDSF as a byproduct of the LoTSS-DR1 catalogue creation. Into each of these residual maps we inject 6000 sources with integrated flux densities drawn from a power-law distribution \( \frac{dN}{dS} \propto S^{-1.6} \) and ranging from 0.1 mJy to 10 Jy. This procedure is repeated 50 times for each of the mosaiced images to ensure a statistically robust measurement. The fraction of sources recovered above an integrated flux density limit, or the point-source completeness, varies with integrated flux density as shown in Fig. 14 and is 65% at 0.18 mJy, 90% at 0.35 mJy, and 95% complete at 0.45 mJy. However, we emphasise that, as shown in Fig. 13, the real integrated flux density level for the completeness levels is likely a factor of \( \sim 1.3 \) higher (thus 90% at 0.45 mJy).

3.7. Image artefacts

In the LoTSS-DR1 mosaics there are several different types of artefacts. The low-level positive and negative haloes are particularly prominent around some sources; these haloes can be difficult to distinguish from real emission and make it challenging to precisely characterise faint diffuse emission. These artificial haloes are believed to be a product of having a minimum \( \phi \)-distance on the baselines used in the calibration; we suspect this expedient, implemented to avoid modelling out extended emission, can cause the amplitude solutions of the antennas with more short baselines to become slightly discrepant from the more remote antennas. For comparison with our images, we note that several diffuse objects within the region covered by this data release have been processed using a different direction dependent calibration algorithm (Facet Calibration; van Weeren et al. 2016). This procedure does not use a large minimum \( \phi \)-distance.
in the calibration and the images do not suffer from artificial haloes; see the maps presented in, for example Brüggen et al. (2018), Savini et al. (2018), and Wilber et al. (2018).

In some fields there are also clear amplitude calibration artefacts that are primarily a consequence of the amplitude normalisation scheme that we used during the direction-dependent calibration. Some fields that were observed in bad conditions also have clear phase errors that are dependent on both our calibration solution interval (1 min) and the size of the facets. Finally, whilst we attempted to ensure that our masks encompass extended sources, there are instances in which faint diffuse emission has still not been fully deconvolved.

As described in Sect. 5.1, in future data releases we plan on improving upon each of these issues. However, for this data release, to aid with the identification of artefacts, we provided mosaics of the final residual maps to accompany our deconvolved continuum images along with the artefact flag resulting from the source (dis-)association and host galaxy identification work of the companion paper Williams et al. (2019).

4. Public data release

In this section, we summarise the products that form the first LoTSS public data release\(^{11}\). These products consist of the mosaiced images that have been described in this paper in addition to the catalogue that we derived from the direct application of P\(\nu\)BDSF to the mosaiced 6″ resolution images. In some cases, P\(\nu\)BDSF does not perfectly represent the radio source population: large extended radio sources may split across several different catalogue entries in the P\(\nu\)BDSF catalogue, or alternatively two closely separated but physically distinct radio sources may be merged into a single catalogue entry by P\(\nu\)BDSF. Therefore, to enhance the scientific value of the released LoTSS-DR1 catalogues we attempted to associate or deblend the catalogued components of radio emission into actual radio sources where necessary, and also to identify the optical counterparts of all sources. If an optical counterpart has been located we also estimated its photometric redshift. For completeness, these projects that add value to publicly released LoTSS-DR1 catalogue are briefly summarised below, but for a full description see Papers III and IV in this series (Williams et al. 2019; Duncan et al. 2019).

4.1. Mosaics and raw P\(\nu\)BDSF catalogue

We released both the 6″ and 20″ resolution 120–168 MHz mosaiced images that were created following the direction dependent calibration procedure described in Sect. 2. These mosaics cover 424 square degrees in the region of the HETDEX Spring Field (see Fig. 5) and have the quality shown in Figs. 4 and 6 and described in detail in Sect. 3. We released 6″ and 20″ mosaiced residual images to help assess the reliability of the morphology of extended structures; these images show the quality of deconvolution and properties of the background noise.

The raw P\(\nu\)BDSF catalogue that was released was created from the 6″ resolution mosaiced images; this catalogue is described in Sect. 2.4. This catalogue contains 325,694 radio sources, has a source density of 770 sources per square degree and a point-source completeness of 90% at an integrated flux density of 0.45 mJy (see Sect. 3.6). To aid the interpretation of the catalogue completeness we released the P\(\nu\)BDSF derived noise maps of the 6″ mosaics.

4.2. Source (dis-)association and optical counterparts

For most radio sources the expected host galaxy position is well defined by the properties of the radio source and it is therefore appropriate to use a statistical method to identify the counterparts in Pan-STARRS and WISE. For this we employ a likelihood ratio method (e.g. Richter 1975; de Ruiter et al. 1977; Sutherland & Saunders 1992). However, for a number of complex sources, such methods are either not possible or unreliable, so we employ a human visual classification scheme based on the Zooniverse\(^{12}\) framework. Sources in the raw P\(\nu\)BDSF catalogue are first sorted based on their catalogue characteristics and selected either for visual (dis-)association and identification or for likelihood ratio cross-matching by means of a decision tree. The details of how these decisions are made and full details of the likelihood ratio and visual classification methods are given by Williams et al. (2019). Using this procedure, counterparts were identified for 71% of the radio sources. These source characteristics and visual inspection procedure are also very useful in flagging probable artefacts in the P\(\nu\)BDSF catalogue. Again, details are given by Williams et al. (2019), but the final column of Table 1 provides a flag highlighting the P\(\nu\)BDSF sources identified as probable artefacts based on that work.

4.3. Photometric redshift estimation

Knowing the redshift of a source is a fundamental requirement for extracting key physical properties from continuum radio observations, such as luminosity or physical size, and for understanding the host galaxy (e.g. its stellar mass). Although future optical spectroscopy campaigns such as WEAVE-LOFAR\(^{13}\) (Smith et al. 2016) will target more than 10\(^6\) 150 MHz-selected sources and provide high-precision spectroscopic redshifts and accurate source classifications for a large portion of the LoTSS population, existing spectroscopic redshifts, largely from SDSS, are available only for a very small subset of sources. Therefore, photometric redshifts (photo-z) are a vital method for identifying the physical properties of radio sources and we produced photo-z estimates for all plausible counterparts in the combined Pan-STARRS/All-WISE catalogue that was used for host-galaxy identification in the previous section. Full details of the photo-z estimation, which combines template-based and machine-learning estimates, are presented in a companion release paper (Duncan et al. 2019).

5. Future prospects

In future data releases we will not only present maps from a significantly larger fraction of the sky, but there is also active development to improve many aspects of the LoTSS data processing; in the survey we observed almost 20% of the northern sky and in this work we only presented 10% of these data or 2% of the northern sky. For example, to tackle the large LoTSS data rates we are working with the LOFAR e-infra group to implement our direction-independent calibration pipeline on the Forschungszentrum Jülich and Poznań LTA sites. Furthermore, the observatory is beginning to utilise Dysco compression (Offringa 2016) to reduce the size of the archived data sets by a factor of approximately four. To improve the accuracy of the direction-independent calibration pipeline, amongst other things, the accuracy of the derived amplitude and clock

\(^{11}\) https://lofar-surveys.org

\(^{12}\) www.zooniverse.org

\(^{13}\) http://www.ing.iac.es/weave/weavelofar/
solutions are being increased. In the direction-dependent calibration and imaging pipeline there is significant work to improve the fidelity of the images and to implement the pipeline on the SURFsara Grid. To further enhance the scientific potential of our data products there is also active work to exploit the polarisation (e.g. Van Eck et al. 2018), wide fractional bandwidth, the longest baselines provided by the international stations, and to use the data for spectral line studies (Onck et al. 2017; Salas et al. 2018; Emig et al. 2019) and for searches for transient sources.

Discussing all of these future prospects in detail is beyond the scope of this article; however, in the following subsections we provide some details on several prospects, namely improving the direction-dependent calibration and exploiting the fractional bandwidth of LoTSS.

5.1. Reducing image artefacts

The LoTSS-DR1 processing strategy has produced sensitive and good quality LOFAR images, however it failed for 8% of the fields and, as described in Sect. 3.7, the final mosaics contain several different types of artefacts. Therefore, in an attempt to improve the images the development of the pipeline has been ongoing. The latest tests that use a refined recipe that still makes use of km and noracet for calibration and imaging, respectively, have shown that by removing the minimum \( \mu \)-distance in the calibration and instead smoothing the amplitude solutions with a low-order polynomial function and fitting the phase solutions with a function proportional to \( \nu^{-1} \) (which is, to first order, the phase behaviour introduced by free electrons in the ionosphere) the artificial haloes and holes can be effectively removed. Furthermore, these changes, together with other enhancements such as turning off the amplitude normalisation, improving the sky models used for calibration by increasing the depth of the deconvolution, and refining the direction-independent calibration by making use of accurate models derived from the direction-dependent imaging, have allowed us to decrease the failure rate of the pipeline, improve the dynamic range, and increase the number of sources detected. A demonstration of the improvements that are a result of these recent developments is shown in Fig. 15 and a refined version of the LoTSS processing pipeline will be fully described in a future publication.

5.2. Exploiting the large fractional bandwidth of LoTSS

With a fractional bandwidth of approximately 33%, LoTSS has the third largest fractional bandwidth of any very wide area radio continuum survey produced to date. Only MSSS (Heald et al. 2015) and the GLEAM (Wayth et al. 2015; Hurley-Walker et al. 2017) survey have observed the sky with larger fractional bandwidths, but both have significantly poorer angular resolutions and sensitivities (see Fig. 1). To demonstrate the scientific potential of the spectral information that can be derived from LoTSS, for a test field we divided a direction-dependent calibrated LoTSS data set into three parts, each with a width of 16 MHz, and generated a three-channel image with noracet. The integrated flux density measurements in each part of the bandwidth and the source association between the three images was done using PyBDSF. An example of some observed spectra, with comparison to other surveys, is shown in Fig. 16. In this demonstration field we were able to accurately derive (with 10% uncertainty or less) in-band spectra for compact, isolated (no other source within 100” of the LoTSS position) sources with integrated flux densities \( \geq 10 \text{ mJy} \), where the uncertainty estimate of the derived spectral indexes was obtained by comparing with spectral indexes measured from fitting to VLSSr and NVSS (\( \geq 50 \text{ mJy} \)) or to TGSS-ADR1 and NVSS for the fainter sources (\( \geq 10 \text{ mJy} \)).

Low-frequency spectral information is valuable for many science cases, such as for identifying low-luminosity peaked-spectrum sources (Callingham et al. 2017) and investigating the energy distribution of electrons in the emitting region of radio sources (e.g. Bonavera et al. 2011). For example, the source shown in the left panel of Fig. 16 is a peaked-spectrum source with a radio luminosity \( < 10^{24} \text{W Hz}^{-1} \), which is two orders of magnitude fainter than the median radio luminosity of previous peaked-spectrum samples (e.g. O’Dea 1998). Probing this population of low-luminosity peaked-spectrum sources could potentially identify sources powered by a short-lived outburst of the central activity that might not able to escape from the host galaxy (Czerny et al. 2009). Such sources could be the short-lived precursors needed to account for the overabundance of peaked-spectrum sources relative to the large-scale radio galaxies (Kunert-Bajraszewska & Labiano 2010).

The right panel of Fig. 16 is the spectrum of a source that shows a significant deviation from a standard power law. If such a deviation is not taken into account, it leads to orders of magnitude uncertainty in the estimate of the energy stored by the lobes of the radio galaxy (Duffy & Blundell 2012; Harwood et al. 2017).

Therefore, the spectral information that can be supplied by LoTSS will have diverse scientific impact, providing internal spectral index information to flux densities below the levels possible by cross-comparison with existing sky survey data. As a consequence of processing constraints this spectral information is not included with this current release, but we plan to include it in future releases.

6. Summary

In this publication we have described the first full quality LoTSS data release, which is available online\(^\text{1}\). We outlined how we managed the large LoTSS data rate and we introduced the completely automated direction-dependent calibration and imaging pipeline that we used to produce 120–168 MHz continuum images. The high-resolution (6") images we present cover 424 square degrees in the region of the HETDEX Spring Field and contain 325 694 sources that are detected with a significance in excess of five times the noise. This source density is a factor of at least ten higher than any existing very wide area radio continuum survey. As described in companion papers (Williams et al. 2019; Duncan et al. 2019) the LoTSS-DR1 catalogue has been enhanced by identifying the optical counterparts of radio sources and estimating their photometric redshifts. Finally, this data release is published together with ~20 articles to highlight the scientific potential of LoTSS.

The LoTSS-DR1 images have a median sensitivity of 71 \( \mu \text{Jy beam}^{-1} \) with approximately 10% of the mapped area being dynamic-range limited. For point sources, the survey is 90% complete at a peak brightness of 0.45 mJy beam\(^{-1}\). We examined the fidelity of our images and found that the astrometric accuracy is approximately 0.2" in both RA and Dec. The flux density scale is in overall agreement with other radio surveys and the uncertainty on the integrated flux density measurements is \( \sim 20\% \).

There are many opportunities to enrich the LoTSS data products through, for example polarimetric measurements or full

\(^{1}\) https://lofar-surveys.org
exploitation of the longest baselines in the international LOFAR array. We briefly demonstrated a few such possibilities including improvements to the calibration and imaging and the measurement of the in-band spectral index.

Acknowledgements. We thank the anonymous referee for his/her comments. This paper is based on data obtained with the International LOFAR Telescope (ILT) under project codes LC2.038 and LC3.008. 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, which are owned by various parties (each with their own funding sources) and are collectively operated by the ILT foundation under a joint scientific policy. The ILT 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; Department of Business, Enterprise and Innovation (DBEI), Ireland; NWO, The Netherlands; The Science and Technology Facilities Council (STFC), UK. Part of this work was carried out on the Dutch national e-infrastructure with the support of the SURF Cooperative through grant e-infra 160022 & 160152. The LOFAR software and dedicated reduction packages on https://github.com/apmechev/GRID_LRT were deployed on the e-infrastructure by the LOFAR e-infragroup, consisting of J. B. R. Oonk (ASTRON and Leiden Observatory), A. P. Mechev (Leiden Observatory) and T. Shimwell (ASTRON) with support from N. Danezi (SURFsara) and C. Schrijvers (SURFsara). This work made use of the University of Hertfordshire high-performance computing facility (http://uhpc.herts.ac.uk) and the LOFAR-UK computing facility located at the University of Hertfordshire and supported by STFC [ST/P000096/1]. The Data Center of the Nançay Radioastronomy Station acknowledges the support of the Conseil Régional of the Région Centre Val de Loire in France. We thank Forschungszentrum Jülich for storage and computing. This research made use of Astropy, a community-developed core Python package for astronomy (Astropy Collaboration 2013) hosted at http://www.astropy.org/, and of the astropy-based reproject package (http://reproject.readthedocs.io/en/stable/). We are grateful to Thomas Robitaille for support in adapting the reproject package to support the very large images used in this paper. HR, DNH, KJD, and RJvW acknowledge support from the ERC Advanced Investigator programme NewClusters 321271. AB acknowledges support from the ERC-Stg DRANOEL, no 714245. AOC gratefully acknowledges support from the European Research Council under grant ERC-2012-StG-307215 LODESTONE. RM and MB gratefully acknowledge support from the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013) /ERC Advanced Grant RADOLIFE-320745. RJvW acknowledges support from the ERC VIDI research programme with project number 639.042.729, which is financed by the Netherlands Organisation for Scientific Research (NWO).