UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

AARTFAAC flux density calibration and Northern hemisphere catalogue at 60 MHz

Kuiack, M.; Huizinga, F.; Molenaar, G.; Prasad, P.; Rowlinson, A.; Wijers, R.A.M.J.

DOI
10.1093/mnras/sty2810

Publication date
2019

Document Version
Final published version

Published in
Monthly Notices of the Royal Astronomical Society

Link to publication

Citation for published version (APA):
Kuiack, M., Huizinga, F., Molenaar, G., Prasad, P., Rowlinson, A., & Wijers, R. A. M. J. (2019). AARTFAAC flux density calibration and Northern hemisphere catalogue at 60 MHz. *Monthly Notices of the Royal Astronomical Society*, *482*(2), 2502-2514. https://doi.org/10.1093/mnras/sty2810

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, Singel 425, 1012 WP 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:09 Mar 2023

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **482,** 2502–2514 (2019)
Advance Access publication 2018 October 18

doi:10.1093/mnras/sty2810

AARTFAAC flux density calibration and Northern hemisphere catalogue at 60 MHz

Mark Kuiack,[1★] Folkert Huizinga,[1] Gijs Molenaar,[2] Peeyush Prasad,[1] Antonia Rowlinson[1,3] and Ralph A. M. J. Wijers[1]

[1]*Anton Pannekoek Institute, University of Amsterdam, Science Park 904, NL-1098 XH Amsterdam, the Netherlands*
[2]*Department of Physics and Electronics, Rhodes University, PO Box 94, Grahamstown 6140, South Africa*
[3]*ASTRON, The Netherlands Institute for Radio Astronomy, Postbus 2, NL-7990 AA Dwingeloo, the Netherlands*

Accepted 2018 October 7. Received 2018 October 4; in original form 2018 June 14

ABSTRACT

We present a method for calibrating the flux density scale for images generated by the Amsterdam ASTRON Radio Transient Facility And Analysis Centre (AARTFAAC). AARTFAAC produces a stream of all-sky images at a rate of 1 s in order to survey the Northern hemisphere for short duration, low-frequency transients, such as the prompt doneelectromagnetic counterpart to gravitational wave events, magnetar flares, blazars, and other as of yet unobserved phenomena. Therefore, an independent flux density scaling solution per image is calculated via bootstrapping, comparing the measured apparent brightness of sources in the field to a reference catalogue. However, the lack of accurate flux density measurements of bright sources below 74 MHz necessitated the creation of the AARTFAAC source catalogue, at 60 MHz, which contains 167 sources across the Northern hemisphere. Using this as a reference results in a sufficiently high number of detected sources in each image to calculate a stable and accurate flux scale per 1 s snapshot, in real-time.

Key words: methods: data analysis – catalogues – surveys – radio continuum: general.

1 INTRODUCTION

The Amsterdam ASTRON Radio Transient Facility And Analysis Center (AARTFAAC) is an all-sky radio monitor, built as a parallel computational back-end to LOFAR (the Low-Frequency Array; van Haarlem et al. 2013). It operates primarily in LOFAR's low band (10–90 MHz) with an all-sky field of view, but can also operate in the high band (110–240 MHz) albeit only with an High Band Array (HBA) tile field of view (30° FWHM at 150 MHz). It can be used to monitor the radio sky on time scales upwards of 1 s, as often as is practicable within LOFAR observing constraints and data storage limitations. Its core science goal is to search for rare, bright transients at the lowest radio frequencies, which have proved to be quite elusive (see, e.g. Bell et al. 2014; Obenberger et al. 2015; Carbone et al. 2016; Rowlinson et al. 2016), but a few have been found in the otherwise poorly explored regime accessible to AARTFAAC (Hyman et al. 2005; Stewart et al. 2016; Murphy et al. 2017); time scales of seconds to hours, and fluxes above several jansky in the LOFAR low band (10–90 MHz). In this regime, coherently emitting objects will dominate, and thus any sources found will represent fairly extreme or exotic physics (see, e.g. Pietka et al. 2015).

However, many other applications are possible, such as detecting very high-redshift EoR signals (Patil et al. 2017), monitoring the state of the ionosphere and phenomena in it (Loi et al. 2015a,c), and monitoring meteor showers (e.g. Obenberger et al. 2014). And of course, many terrestrial phenomena such as RFI, air planes, and satellites are detected that need to be distinguished from more distant astrophysical signals and imaging artefacts before science analysis can start.

At the moment, AARTFAAC all-sky monitoring is limited to times when LOFAR is Low Band Array (LBA) mode,[1] but more importantly, due to the fact that the search for interesting objects has not yet been automated, and storage and offline search of the very large volumes of raw data it generates is not practical. In this paper, we describe the next step in achieving the goal of continuous all-sky monitoring with AARTFAAC. Previously, we described the basic properties of the AARTFAAC system and its real-time calibration and technical commissioning (Prasad et al. 2014), as well as the system design and correlator (Prasad et al. 2016), and TraP, the transients detection pipeline also used for LOFAR transient searches (Swinbank et al. 2015). In this paper, we describe the results of the first practical commissioning runs. The basic goal

★ E-mail: m.j.kuiack@uva.nl

[1]An upgrade is planned and funded that will allow simultaneous LBA and HBA observations

© 2018 The Author(s)
Published by Oxford University Press on behalf of the Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article-abstract/482/2/2502/5136428 by Universiteit van Amsterdam user on 02 April 2020

of these runs was to collect enough data to cover the full range of local sidereal times over a significant period of time (so as to cover a range of ionospheric and RFI conditions) to investigate practical strategies of bad data rejection, source extraction, and flux calibration to enable future real-time operation. We collected over 30 h of data, which we will show is a good compromise between getting a manageable dataset to experiment with extensively and sampling a sufficient range of conditions. In future work, we will first develop strategies for separating events of interest in large datasets from artefacts and known variability, before implementing the full intended data analysis (data taking, correlation, calibration, imaging, flux extraction, feature recognition, and alert generation) into a streaming pipeline that can function in real-time.

To increase the sensitivity to fainter objects, the brightest sources in the sky: Cygnus A, Cassiopeia A, Taurus A, Virgo A, The Sun, hereafter referred to as the 'A-team', and most of the diffuse Galactic plane emission, is removed during AARTFAAC calibration and imaging. The images have, thus far, not yet been properly flux-calibrated (Prasad et al. 2014). Instead, the pixel values in the resulting images are relative, with a scaling related to a normalization of the total power received before 'A-team' subtraction. An accurate characterization of transient phenomena, however, requires reference to a common physical scale. While studies of variability also require that each extracted source measurement is made within a comparable reference frame. This is only possible once each image has been corrected such that the pixel values refer to the physical units of flux density, janskys.

Radio flux density calibration is done by reference to catalogues of stable, well-studied calibrator sources. For example, a typical radio observation includes observing a calibrator source before and after observing the target. So the scaling of the calibrator data, the gain solution, is applied to the target data. In that case, the gain solution is assumed not to have changed substantially on the timescale of the observation.

Unfortunately for AARTFAAC, the only large surveys below 100 MHz are The Very Large Array Low-frequency Sky Survey Redux (VLSSr; Lane et al. 2014) at 74 MHz and The Eighth Cambridge (8C) Survey (Rees 1990) at 38 MHz. There is therefore a gap across nearly the entire frequency range of AARTFAAC. This clearly represents an opportunity for AARTFAAC to make an important contribution, with unique flux density measurements of the brightest sources in the 38–74 MHz range.

We therefore report on the method used to flux density calibrate AARTFAAC images in real-time, for our future transient search campaigns, and the resulting catalogue of bright sources.

First, in Section 2 particular technical details of AARTFAAC and the calibration observations are given. Secondly, in Section 3 we describe the catalogue bootstrapping method used to accurately flux density calibrate the images in real-time. Then, we report on the performance and stability of the method in Section 4, and analyses of the typical systematic uncertainties. Next in Section 5, the characteristics of the first AARTFAAC catalogue of persistent sources at 60 MHz are discussed. And lastly conclusions are given in Section 6.

2 DATA DESCRIPTION

By creating an all-sky image every second, AARTFAAC has the capability of generating a large amount of data. Therefore, the intended operational mode is to perform a transient search on the stream of images, saving only those data where an interesting event has been detected. However, in order to test the calibration method and fully characterize the data quality a set of observations was recorded and stored for analysis offline.

Additionally, full LST coverage was required to generate the catalogue of calibrator sources across the Northern hemisphere. Therefore, nearly 33 h of observations were recorded to test the flux density calibration method and generate the AARTFAAC catalogue. This allowed the analysis of sources for many hours, across separate observations, while maintaining a manageable data volume. These observations were recorded between 2016 August and December, as outlined in Table 1. During this period of time, the final stages of commissioning with the real-time imaging pipeline were completed, leaving only the image calibration.

In its present form, the AARTFAAC system shares the six core stations, known as the 'Superterp', with LOFAR, which is located near the village of Exloo in the Netherlands. It was designed to operate in parallel with regular LOFAR observations by splitting the antenna signals from the stations and rerouting them to the AARTFAAC correlators and imaging servers.

Each core LOFAR station consists of two subarrays: the HBA, which has a bandpass from 120 to 240 MHz, and the LBA, with a bandpass from 10 to 90 MHz. AARTFAAC currently only uses the LBA.

These LBA stations are made of 96 pairs of orthogonal droop dipoles distributed with a roughly Gaussian density distribution. Their simple antenna design, two wires attached at 45° to a central post over a metal mesh ground plane, offer a full sky field of view. Unfortunately, due to current computational constraints only data from 48 of the 96 antennas are processed. This subset of antennas may be distributed in one of four operating modes.

(i) INNER: antennas within 30 metres of the station centre.
(ii) OUTER: antennas 30–87 metres.
(iii) 2 SPARSE modes: either odd or even numbered antennas distributed throughout the station.

LOFAR LBA observes predominately in the OUTER configuration due to the larger number of longer baselines providing better UV filling of the superterp. In comparison using the INNER configuration results in dipoles that are more tightly clumped in the centre of the station, leaving more space between the stations. The OUTER configuration utilizes the outer ring of station dipoles that maximizes point source sensitivity and reduces diffuse background emission. Additionally, regular LOFAR LBA observations will sum the antenna signals with a phase delay applied for the target pointing. However, because these phase delays are not applied at the stations during LBA observations, AARTFAAC has access to the raw signal from all 48 dipoles in operation, and is sensitive to the entire visible sky during all LBA observations.

The physical specifications of AARTFAAC are summarized in Table 2. Currently, in the standard operating mode of AARTFAAC a 1 s integrated Stokes I (1024 × 1024 resolution, 4.1 MB) fits image is created every second by integrating all 16 available subbands. This is a reduction from the initial total raw visibility rate 660 MB s^{-1}, including all subbands, which is reduced to 10 MB s^{-1} after calibration by averaging the 63 frequency channels that comprise each LOFAR subband. These calibrated visibilities are stored in the AARTFAAC archive for offline processing, and the upcoming transient survey. However, in the future only those data that are found to contain an interesting transient event will be stored.

In order to maximize sensitivity while reducing RFI, the subbands are configured in two continuous sets of eight subbands, 57.52–59.08 MHz and 61.04–62.6 MHz. This is near enough the peak sensitivity around 57 MHz (van Haarlem et al. 2013), while

Downloaded from https://academic.oup.com/mnras/article-abstract/482/2/2502/5136428 by Universiteit van Amsterdam user on 02 April 2020

Table 1. The set of observations used to test the flux density calibration method and generate the first AARTFAAC catalogue at 60 MHz. The start and end of each observation are given as the UTC of the first and last image, as well as the LST centre of the image, both to the nearest minute. During an observation data, blocks may be flagged and removed either by the correlator, visibility calibration, or imaging pipeline. Then, the images were filtered based on the average image pixel rms. Outliers are clearly the result of improper calibration, poor A-team subtraction, or bright RFI.

Start date	Start–end (UTC)	Start–end (LST)	Good images #	Unflagged data Per cent
2016 Aug 31	15:10–17:43	14h18m–16h52m	8839	96.3
2016 Sep 05	16:47–19:45	16h15m–19h14m	10 358	97.0
2016 Sep 07	03:40–09:38	03h14m–09h14m	21 291	99.1
2016 Sep 30	09:31–11:23	10h36m–12h29m	2703	40.2
2016 Nov 12	06:32–19:53	10h26m–23h50m	40 145	83.5
2016 Nov 13	20:00–22:57	00h01m–02h58m	5031	47.4
2016 Nov 14	08:27–15:33	09h03m–16h56m	23 084	90.3
2016 Dec 10	22:55–02:49	04h43m–08h37m	9794	70.0
Total	32:56:46		121 245	

Table 2. AARTFAAC system design specifications, from table 1 of Prasad et al. (2014). Here, the subband and processed bandwidth values are updated to reflect the current operational capabilities of AARTFAAC.

Parameter	AARTFAAC LBA	Comment
Array elements	288 inverted V antennas	Dual polarized elements
Frequency range	10–90 (MHz)	
Field of view	π (sr)	FWHM of beam
Total effective area	2617^a (m^2)	
T_{sys}	3600 ($\nu^{-2.55}$K)	
Angular resolution	60 (arcmin)	
Subband resolution	195 (kHz)	
Processed bandwidth	3.12 (MHz)	
Temporal resolution	1 (s)	

avoiding frequencies that have been observed to have a higher RFI occupancypercentage (see fig. 6; Offringa et al. 2013). With this configuration, a pixel rms < 10 Jy is achieved over 40 per cent of the Northern hemisphere, while 90 per cent achieves rms < 21 Jy.

3 IMAGE CALIBRATION

The AARTFAAC real-time calibration and imaging pipelines, as they are currently implemented, output all sky snapshot images at a rate, and integration time, of 1 s (Prasad et al. 2014). Yet, before the images can be used for transient detection or variability analysis, two corrections must be made. First, a direction-dependent rescaling, which corrects the images based on the antenna response pattern, also known as the primary beam. Then, a direction-independent rescaling, which transforms the pixel values from arbitrary units to a flux density in janskys.

Both of these corrections are important before the image stream can be analysed for transient or variable sources. Clearly, without accounting for any variation in the antenna response across the sky, where sensitivity peaks at zenith, the brightness of all sources would appear variable as they track across the field of view.

Similarly, applying a reliable flux density scaling to each image is vital for measuring variability. Given that the pixel values in raw images are arbitrary, with an unknown influence from the subtraction of the A-team sources, it would be difficult to determine whether any variation is intrinsic or an artefact of the calibration and imaging. Furthermore, searching for a transient source in an image with arbitrary scaling would make it impossible to determine the shape of the light curve, spectral index, or whether a candidate is indeed astrophysical at all.

These characteristics are also critical when devising further observations as they dictate what sensitivity and spectral coverage are necessary to maximize the likelihood of a follow-up detection. Additionally, in the cases where detections are made but follow-up observations yield no result, the shape and flux density distributions of transient light curves would be useful to model possible progenitor populations, and to compare our results with those of other low-frequency surveys. This is the case for many FRB studies to date (Petroff et al. 2015).

3.1 Beam model

The beam model is an approximation of the direction-dependent sensitivity, across the field of view, of the array. Thus, correcting for this differential gain response pattern ensures that the light curves of detected sources are flat, as they move across the field of view.

This pattern is the result of many physical factors including: observing frequency, the geometry of the stations and dipoles, their mutual interactions with each other, and the effect of the local terrain. Given these complicated interactions, it is modelled by simulating the full station layout of all of the dipoles that form the six stations on the LOFAR Superterp, with accurate placement and orientation, across the frequency spectrum.

Therefore, the beam response shape has been modelled at frequencies between 30 and 70 MHz in 5 MHz intervals. This covers the spectral range of AARTFAAC with sufficient accuracy, since the model does not change rapidly with frequency. These models were generated using WIPL-D, an electromagnetic simulation software package. Additionally, Fig. 2 illustrates the symmetry of the model at 60 MHz, about the zenith.

The left image of Fig. 1 shows an example raw AARTFAAC image. Although the background and noise appear flat across the image, sources decrease in apparent brightness as the sensitivity drops towards the horizons. The sensitivity peaks at zenith and decreases towards the horizon. Therefore, the images are corrected by dividing the raw image by the image of the beam model, normalized such that the gain at zenith is 1. The shape of the normalized beam model at 60 MHz is illustrated in the middle panel of Fig. 1, and the resulting beam corrected image is shown on the right. Given that the sensitivity near the horizon is lower, after the correction is applied, the noise near the horizon is also multiplied. However, the

Downloaded from https://academic.oup.com/mnras/article-abstract/482/2/2502/5136428 by Universiteit van Amsterdam user on 02 April 2020

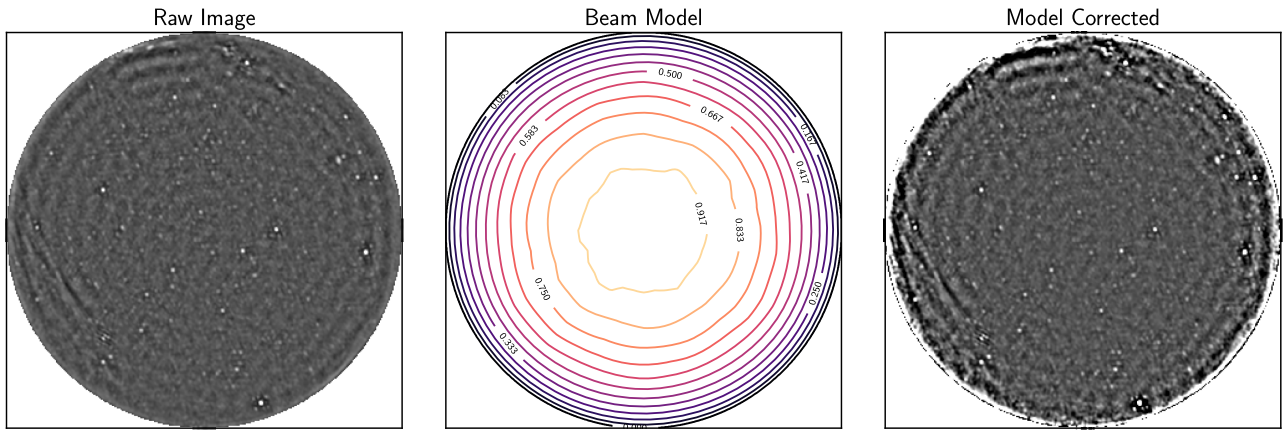

Figure 1. Left: an AARTFAAC all-sky image, output by the imaging pipeline, after A-team subtraction. Centre: a normalized beam model, for AARTFAAC at 60 MHz. The model shows the shape of the direction-dependent gain across the field of view with maximum gain at zenith, decreasing towards the horizon. Right: the same image with the correction applied. Note, these images have not been flux density calibrated so the pixel scaling is arbitrary.

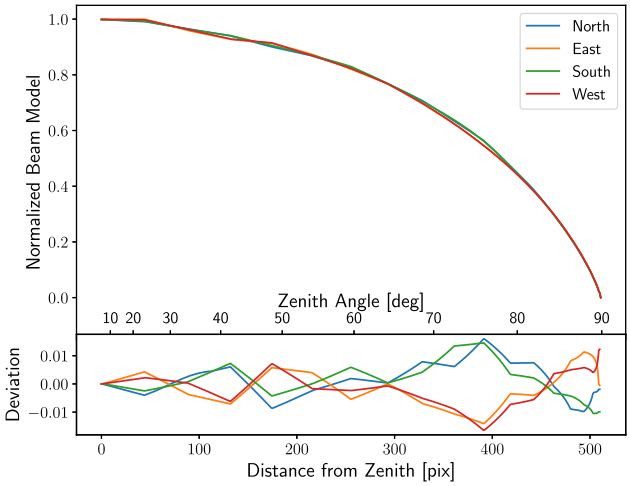

Figure 2. Comparing the beam model gain pattern at 60 MHz, from zenith to horizon, in the four compass directions. The values are normalized to their maximum value, which is at zenith. The beam is highly symmetric about the zenith, with relative deviations from perfect symmetry below 1 per cent out to 70°.

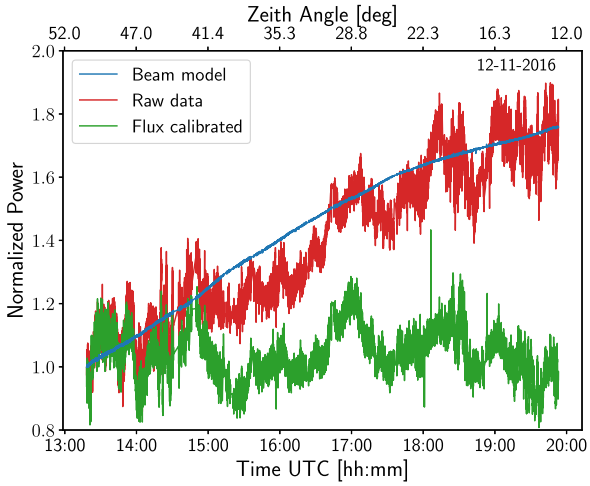

Figure 3. Example light curve for a source observed for ∼7 h, normalized to the first data point. The raw data (red) before calibration clearly shows the shape of beam pattern along the track of the source across the sky (blue), while after calibration (green) the curve is much flatter.

mean flux density of sources will be constant as they rotate through the beam, as illustrated in Fig. 3.

Although the beam model has been observationally verified during LOFAR commissioning (van Haarlem et al. 2013), AARTFAAC is able to perform an additional test, using sources detected across the field of view, observed over hours as they move across the sky. After calibrating the data, the extracting flux measurements of each source at different locations on the sky were compared to the mean of their light curve. No position dependent deviations, which would indicate an improperly modelled beam, were observed within our detection region of 50 degrees from zenith.

3.2 Reference catalogues

Given the real-time streaming nature of AARTFAAC observations, an algorithm that computes the flux scaling per image, using only the image itself, is preferred. This excludes traditional flux density scale calibration techniques, such as intermittently observing another field with a calibrator source.

Fortunately, with our field of view encompassing the entire sky, enough bright sources are present in each image to compute the scaling in real-time. However, this is only possible if accurate apparent fluxes at the observing frequency are known. Therefore, with a population of enough bright sources covering the Northern hemisphere, the apparent brightness of those sources extracted in each image can be compared with the expected flux density, and the conversion scale factor computed. This is described fully in Section 3.3.

Furthermore, utilizing the greatest number of calibrators ensures that the variability of any one source does not dominate the calibration solution. Additionally, ionospheric fluctuations are phenomena localized on the sky, as density fluctuations travel through the field of view, and therefore decorrelate on larger angular sizes. To this end, gathering together a catalogue of all sources with a signal-to-noise ratio $> 5\sigma$, and broadband spectra across our entire observing range, and field of view, would allow accurate and stable flux density calibration at any observing frequency.

Downloaded from https://academic.oup.com/mnras/article-abstract/482/2/2502/5136428 by Universiteit van Amsterdam user on 02 April 2020

Recently, several catalogues of calibrator sources have been published with accurate broadband spectra in the LOFAR LBA range, 30–80 MHz.

One example, Scaife & Heald (2012) contains six bright sources from the Third Cambridge Catalogue, 3C, with spectral models between 30 and 300 MHz. Unfortunately, with only six sources spanning the Northern hemisphere, AARTFAAC images would not contain enough calibrator sources to ensure a robust scaling solution at all observing times. However, the analytic spectral models across the full LBA band and the fact that these sources appear in the other catalogues adds a useful intercatalogue check.

Secondly, Perley & Butler (2017) published modelled spectra for 20 sources with flux density measurements down to 74 MHz. Four of these are the A-team sources that are subtracted from the images before this flux density calibration step, leaving 16 sources. Of these, 6 have a declination $<-10^{\circ}$, and are therefore outside, or too near the edge, of our field of view. This leaves eight sources that are viable flux density calibrators. Even still, rarely are more than three of these sources visible in the sky simultaneously, which is preferred to have stable flux density fit solution at all times.

Lastly, the catalogue published by Helmboldt et al. (2008) contains spectra for 388 sources selected from VLSSr that are brighter than 15 Jy at 74 MHz. The spectral shapes of all sources are described with either a single power law, or if enough data are present, by the function $Y = A + BX + C\exp(DX)$, where $Y = \log(F_\nu/1\ \mathrm{Jy})$ and $X = \log(\nu/74\ \mathrm{MHz})$, which describes a spectral turnover of the flux density, F_ν, at lower frequencies, ν, in some sources.

In addition to these, the full VLSSr catalogue (Lane et al. 2014) was used to follow up sources that are detected in AARTFAAC images with $>5\sigma$ signal to noise but are not associated with any object in the Helmboldt et al. (2008) catalogue. These might result from two or more sources in VLSSr with <15 Jy that are sufficiently close together to be confused at AARTFAAC's resolution.

The lower resolution of AARTFAAC and the densely packed array allows us to see much more diffuse emission than the VLSSr. For this reason, the supernova remnant catalogue by Green (2014) was also used for source association. These objects are of interest to us because they are bright at low frequencies, and their larger angular size reduces the effect of ionospheric scintillation. However, due to the frequency at which these flux densities are given (1.4 GHz) and the much narrower beamwidth, it is impossible to simply extrapolate and compare with AARTFAAC measurements.

The spectral models published by Perley & Butler (2017) are much more accurate in the frequency range observed with AARTFAAC. However, this catalogue does not contain enough sources across the Northern hemisphere to ensure that three or more sources are observable simultaneously, which is a requirement, used to ensure a more stable scaling solution in the presence of scintillation. And while Helmboldt et al. (2008) present a catalogue with many more sources, the simpler spectral models result in a much greater uncertainty in the flux density predictions below 74 MHz. Neither catalogue was therefore sufficient to accurately compute a flux density scale at all times. This necessitated the creation of the AARTFAAC low-frequency catalogue, which is outlined in detail in Section 5.

3.3 Flux density scale

The AARTFAAC catalogue was created by first modelling the Perley & Butler (2017) source fluxes at lower frequencies. This was done by using a Monte Carlo method where 10 000 sets of flux density measurements for a given source at different frequencies

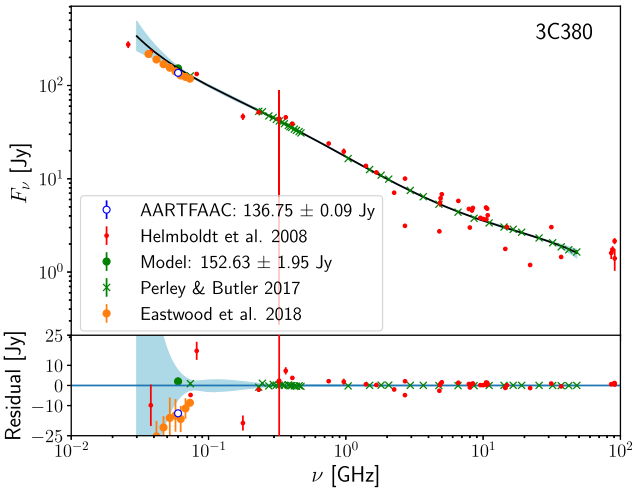

Figure 4. The spectral model of 3C 380 is generated from the flux density measurements by Perley & Butler (2017). The uncertainty in the model (light blue region) is calculated via Monte Carlo random sampling of sets of measurements based on their reported uncertainty. This gives the model value at our measurement frequency, 60 MHz (green filled circle). The AARTFAAC catalogue flux density value (open circle) and Helmboldt et al. (2008) measurements (red dots) and Eastwood et al. (2018) measurements (orange filled circles) are compared to the model (black line). The residuals show the difference between the measured values and the model.

were generated using the flux density measurements and uncertainties available in the supplementary online data. Next, a least squares fit was done to the spectral models with the same polynomial degrees as those in Perley & Butler (2017):

$$\log(F_\nu) = a_0 + a_1 \log(\nu_G) + a_2 \log(\nu_G)^2 + a_3 \log(\nu_G)^3 + \ldots, \quad (1)$$

where F_ν is the flux density in Jy and ν_G is the frequency in GHz. Finally, the resulting functions are then evaluated at the AARTFAAC observing frequency. This provides predicted fluxes with accurate uncertainties. The spectral model for one such source, 3C 380, is shown in Fig. 4, along with the flux density measurements from Helmboldt et al. (2008) and Perley & Butler (2017). Here, the model predicted flux density at 60 MHz, and the final AARTFAAC catalogue flux density, after averaging the data from all observations, is shown not to be in agreement, within their mutual uncertainties. However, when comparing with the flux density as measured by Eastwood et al. (2018), the differences are much smaller for all sources, except 3C 286. Therefore, because of the similarity of the LOFAR-LBA and OVRO-LWA antenna design, but different imaging and calibration method, it is clear that the measured flux density at 60 MHz is accurate.

The AARTFAAC catalogue is then made by bootstrapping the flux density scale from these sources in the following way.

First, an observation of a few hours was considered at a time when five calibrator sources were visible.

Then the predicted flux density values at 60 MHz, obtained from the reference catalogue, were compared to their sigma-clipped light curves extracted from the observation. Iteratively, clipping the flux density measurements greater than 3σ from the mean reduces the possible effect of RFI or imaging artefacts.

Light curves are generated by extracting the source fluxes from each image with the Python Source Extractor (Carbone et al. 2018), then the sources extracted from each image are associated with the extractions from previous images in a running catalogue database.

Downloaded from https://academic.oup.com/mnras/article-abstract/482/2/2502/5136428 by Universiteit van Amsterdam user on 02 April 2020

This is performed using the LOFAR Transient Pipeline (TraP; Swinbank et al. 2015 and references therein).

The durations of the observations used for generating the catalogue were each longer than 2 h. This ensures that scintillation effects shown in Fig. 3, which occur on a 15–20 min time scale (Loi et al. 2015b), are averaged out.

We then calculate the scaling solution via linear regression, weighting the sources according to their measurement uncertainty. The single scale factor calculated for that observation was then applied to the mode of the flux density measurements of the other persistent sources, detected above 5σ.

Those inferred values and our measurement of the original reference sources are added to the AARTFAAC catalogue. By adding the additional calibrators, and using the new AARTFAAC catalogue as a reference for the other observations where fewer Perley & Butler (2017) calibrators are visible, a more accurate scale factor for each additional observation can be calculated.

To summarize, in each new observation, all light curves longer than 2 h are extracted, the scale factor for the observation is calculated, and it is applied to the new sources, then the new sources are added to the catalogue. The number of data points for each source is recorded so that when a source, which already exists in the catalogue, is re-observed, a mean weighted by the number of measurements is used to calculate the updated flux density value.

Now that the entire Northern hemisphere has been observed, and all persistent sources detectable above 5σ, flux densities of 40–50 Jy for much of the sky, have been added, the AARTFAAC catalogue can be used to flux density calibrate any individual AARTFAAC image. This is possible because there are 30–50 observable sources above this threshold at any time. This greater number of sources, across the full field of view, provides a stable flux density solution.

Finally, the AARTFAAC catalogue was used to correct the flux density scale on the full 33 h of data. This was done in a streaming mode, calculating the linear scaling solution for each image independently, so that the intended use case could be analysed by verifying light curves extracted from the data. In doing so, we observed that the scaling solution did have a dependence on LST, which was expected due to the scaling of the raw visibilities according to the total power of objects in the field of view, but that the scaling solutions, at a given LST, were stable across the months spanned by the set of observations.

4 ANALYSIS OF FLUX-CALIBRATED DATA

AARTFAAC produces a snapshot image of the entire sky at a rate of 1 per second. To this stream of images a correction for the antenna response pattern, as well as a scale factor, per image, is applied. This enables the creation of a reliable light curve for each source. In the previous section, the method for correcting each image by applying the beam model and calculating the scale factor, required to scale the pixel values to flux densities in janskys, was discussed.

Each of these corrections influence the light curves on different time scales; the varying sensitivity of the antenna will modulate the apparent brightness over a long period of time as the source moves across the sky, whereas the flux density scaling is corrected on each image independently and would therefore have its effect on the timescale of ionospheric fluctuations.

Turbulence in the ionosphere causes the apparent brightness, as well as the position and shape, of sources to fluctuate. This reduces the instantaneous accuracy of measurements from individual AARTFAAC images. Fortunately, this is overcome by observing each source for a sufficiently long time that the mean value of the light curve converges.

The length of time for which each source must be observed depends in general on the typical timescale of ionospheric scintillation. For example, if a source is observed many times, but for a shorter duration than the timescale of these fluctuations, the variance of the light curves, and thus the uncertainty in the flux density will be dominated by the ionospheric fluctuations.

As an example, Fig. 5 illustrates this for the Tycho supernova remnant. Large sources of synchrotron emission, such as supernova remnants generally do not intrinsically vary in brightness, making them useful tools to study the systematic effects on our flux density measurements.

For a pure Gaussian noise process, the standard error in the mean, SEM, defined as the standard deviation of the means calculated from subsets of the data, decreases proportional to the number of samples in the subset, $\sigma_{\mathrm{m},N} \propto N^{-1/2}$. This is the green 'Theoretical' line in Fig. 5(a), scaled by the average uncertainty in each individual integrated flux density measurement, Δf_{int}.

However, Fig. 5(d) illustrates that AARTFAAC flux density measurements are not a pure Gaussian process. In fact, despite the larger angular size of Tycho, its light curve, shown in Fig. 5(c), reveals that the measured flux density is still significantly modulated by the ionosphere.

There are therefore two regimes, timescales much less than, or much greater than the timescale of ionospheric fluctuations, 10^2–10^3 s, which represent the dominant sources of uncertainty in AARTFAAC flux density measurements.

First, the statistical uncertainty in each individual source fit due to the image noise. For timescales much less than the fluctuations in the ionosphere ($<10^2$ s), these measurements are highly correlated and thus not independent. This is due to the fact that AARTFAAC images are confusion noise-limited. The noise profile is therefore below what is estimated by the individual source fits. This is evident in the left-hand side of Fig. 5(a), where the 'Sequential sample' curve, generated by calculating the mode of subsamples from the light curve sequentially over increasing time, is below the pure Gaussian estimate.

Secondly, the variation of the brightness due to the electron density fluctuations in the ionosphere. Again, the noise profile differs from the Gaussian estimate for light curves much longer than the typical ionospheric timescale ($>\sim 10^2$ s). The variation caused by the ionosphere causes fluctuations that are greater than the estimated uncertainties from each image. This effect causes the estimated error to cross above the Gaussian estimate.

Furthermore, observing a source for many hours will result in significant motion across the sky, and through the beam of the antenna. Fig. 3 illustrates an example of correcting for the antenna response on a source, whose light curves have been normalized to the first data point. Clearly, in the raw light curve (red) the increasing sensitivity of the antenna is visible as the source rises in the sky towards zenith. Alongside this, the beam response pattern along the path of the source (blue) illustrates how as the sensitivity increase towards zenith and the beam centre, explaining the dramatic increase in the apparent brightness. When the beam model is divided out, a much flatter calibrated light curve (green) remains. In fact, the residual variability in the light curve shown is predominantly due to turbulence in the ionosphere, causing the apparent brightness fluctuations.

Consequently, these noise characteristics indicate that observations shorter than a few minutes duration may not yield an accurate average flux density value, regardless of the noise properties

Downloaded from https://academic.oup.com/mnras/article-abstract/482/2/2502/5136428 by Universiteit van Amsterdam user on 02 April 2020

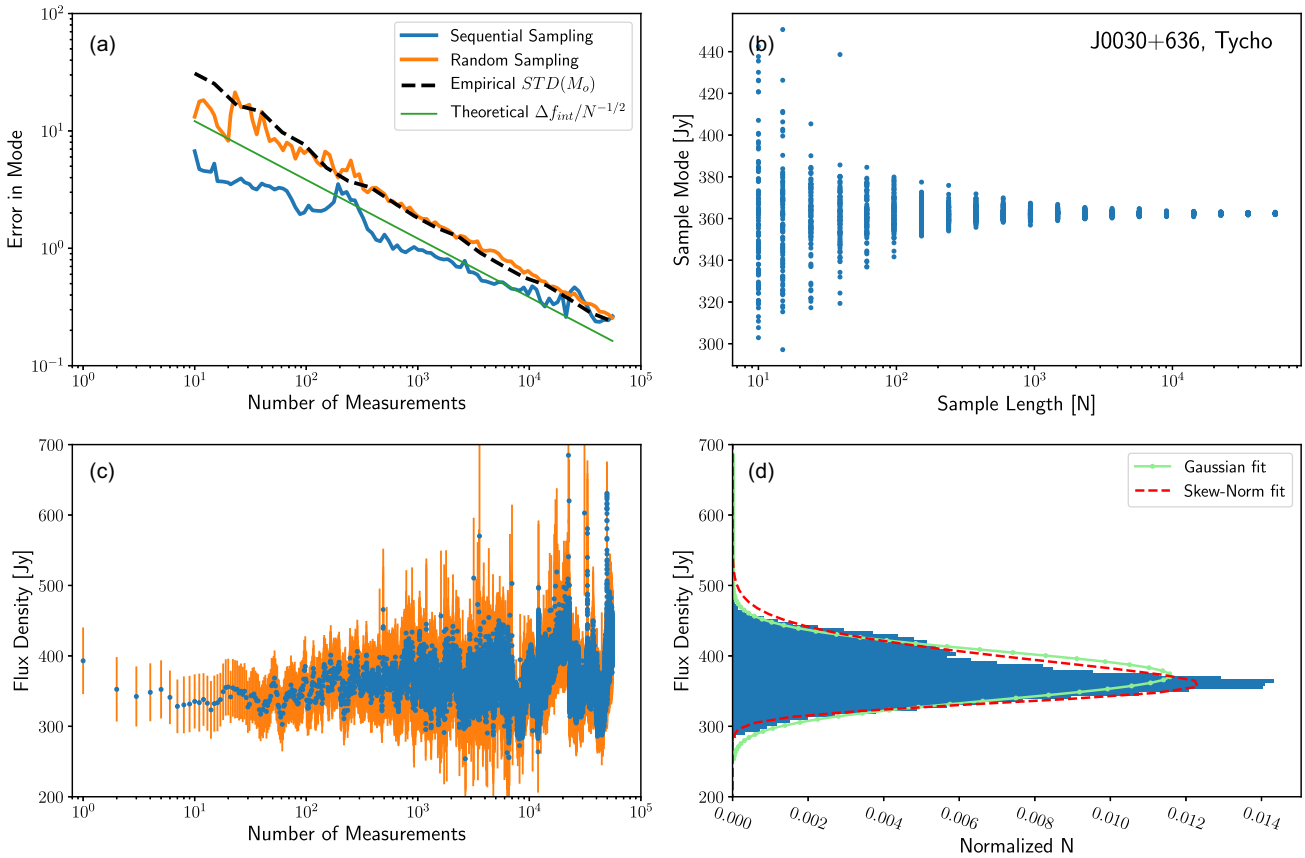

Figure 5. The noise profile of the Tycho supernova remnant. The source is intrinsically stable due to its large size, and emission mechanism, so fluctuations in its measured flux density are due strictly to either source fit statistics or ionospheric fluctuations. Figure a shows the error in the mode as a function of sample size, for different methods of sampling the data, and different methods of estimating the uncertainty. The sequential sampling curve (blue) was generated by taking subsets of the measurements from the light curve in order. This shows how the uncertainty is affected by the systematic sources of error over time. The random sampling curve (orange) was generated by randomly sampling, with replacement, from the entire light curve and estimating the uncertainty in the mode of that subsample. The empirical (black dashed) was calculated by randomly sampling multiple subsets of the light curve, then calculating the standard deviation in the modes of those subsets, then repeating this for a number of sample lengths. These populations of mode, from different subsets, are plotted in figure b. Lastly, the theoretical line (green) shows how the standard error in the mean value of the data would be expected if the noise were purely Gaussian, independent and equal to the average uncertainty in the fit from each image. Figure c shows the entire light curve of Tycho from our set of observations, with the uncertainty of each measurement in orange. Lastly, figure d presents a histogram of the data, and a comparison to a normal and skew normal distribution.

in each individual image. However, the histogram of flux density values measured from the entire observation, shown in Fig. 5(d), illustrates that the noise profile is, by appearance, nearly Gaussian, after observing the source for a period significantly longer than the timescale of the ionosphere. This is to be expected given the central limit theorem. Therefore, making prolonged observations results in both accurate and precise flux density measurements.

4.1 Flux density measurement statistics

As the noise was expected to be Gaussian, with potential systematic influences from either the ionosphere or an incorrectly modelled beam pattern, a skewed normal Gaussian distribution was fit to each light curve,

$$f(x) = \frac{2}{\omega}\phi\left(\frac{x-\xi}{\omega}\right)\Phi\left(\alpha\left(\frac{x-\xi}{\omega}\right)\right), \tag{2}$$

where ϕ is the standard normal probability distribution and Φ is its cumulative distribution, and the transformation $x \rightarrow \frac{x-\xi}{\omega}$, to the fitted parameters ξ, the location, ω the scale, and α, the shape, which defines skewness. The skewness increases with the absolute value of α, with more weight in the left tail when $\alpha < 0$ and in the right tail when $\alpha > 0$, when $\alpha = 0$ the skewed normal distribution becomes the normal distribution, and ξ is simply the mean, and ω the standard deviation.

After fitting equation (2) to each of the light curves, it was found that the skewness was most frequently positive, with a larger tail in the distribution towards higher flux. This could indicate that the variation, introduced by the ionosphere, preferentially focuses the light into shorter bright caustics. These move along the ground, similar to the light on the bottom of a swimming pool.

Consequently, simply integrating over long observations, either by simple average, or calculating the median would yield results biased by the preference for outliers with increased brightness. Hence, the mode of the distribution is the most robust measurement of the central tendency of each source, and therefore the most accurate description of its flux density.

However, the mode of the skew normal distribution is not analytic, but can be approximated numerically,

$$M_o \approx \xi + \omega m_o(\alpha), \tag{3}$$

Downloaded from https://academic.oup.com/mnras/article-abstract/482/2/2502/5136428 by Universiteit van Amsterdam user on 02 April 2020

where the ξ, ω, and α, are the fit parameters location, scale, and shape, of the distribution. The function $m_o(\alpha)$ is the degree to which the skew modifies the mode from the mean, which for a normal distribution is equal to 0. This is given by

$$m_o(\alpha) \approx \mu_z - \frac{\gamma_1 \sigma_z}{2} - \frac{\mathrm{sgn}(\alpha)}{2} \exp\left(-\frac{2\pi}{|\alpha|}\right), \tag{4}$$

where $\sigma_z = \sqrt{1-\mu_z^2}$, such that

$$\mu_z = \sqrt{\frac{2}{\pi}}\delta,$$

for

$$\delta = \frac{\alpha}{\sqrt{1+\alpha^2}},$$

and where γ_1 is the skewness,

$$\gamma_1 = \frac{4-\pi}{2} \frac{\left(\delta\sqrt{2/\pi}\right)^3}{\left(1-2\delta^2/\pi\right)^{3/2}}. \tag{5}$$

The parameters were fit using the Bayesian inference MCMC package PyMC3. This method randomly samples the parameters from defined prior distributions, then computes the likelihood. The uncertainty estimate was output by the PyMC3 package (Salvatier, Wiecki & Fonnesbeck 2015). By defining the mode as a deterministic function of the fit parameters, PyMC3 gives the resulting uncertainty of the mode directly, as well as producing a covariance matrix for the parameters, and plots of the marginal posterior probability distributions.

The robustness of the mode and correctness of the uncertainty calculation was tested in two ways.

First, subsets of the data of varying length were randomly sampled with replacement from the light curve. Then the mode of each subset was calculated. Fig. 5(b) shows the variance in the mode of each sample of Tycho flux density measurements, taken from different length intervals of the total observation. As the length of these intervals was increased, the standard deviation of the mode of each subset was calculated, and plotted as the dashed black line in Fig. 5(a).

Then, the 'random sampling' curve in Fig. 5(a) shows the estimation of the uncertainty in the mode of a randomly drawn sample from the light curve. Comparing these uncertainty estimates, the PyMC3 estimate from the random sample, with the empirically measured standard deviation of the modes from a number of different subsets, and the estimate calculated from the fit parameter uncertainties added in quadrature, provides an additional independent verification of the reported uncertainty.

As previously argued, this catalogue presents the mode of the skew normal distribution as the most robust measurement of the flux density of each source, under the influence of a turbulent ionosphere. The observed tendency towards positive skewness indicates that longer integrations, or simply averaging successive shorter integrations, would yield a positive bias in the inferred flux density.

Given that AARTFAAC images are generated and calibrated at a rate of one per second, well below the typical time scale of ionospheric scintillation, it is possible to observe a large number of flux density measurements, fit the distribution, and calculate the mode. However, other low-frequency surveys typically make much longer integrations to reduce the image noise level. Therefore, the quantity that should be compared is the mean, rather than the mode. The mean, μ, of the distributions fit can be calculated from the parameters given in the catalogue as follows:

$$\mu = \xi + \omega\delta\sqrt{\frac{2}{\pi}}. \tag{6}$$

Lastly, the uncertainty presented in the flux densities are the statistical uncertainty in the modes of each distribution. As shown in Fig. 5(a), follow up measurements would agree within the stated uncertainty if the duration of the observations is sufficient. However, observations shorter than the ionospheric time scale could only be expected to agree within the standard deviation, described by

$$\sigma = \sqrt{\omega^2\left(1-\frac{2\delta^2}{\pi}\right)}. \tag{7}$$

5 CATALOGUE

5.1 General properties

The aim of this catalogue is to fill the gap between the VLSSr at 74 MHz (Lane et al. 2014) and the 8C catalogue at 38 MHz. Indeed, for many sources below the 60° declination limit of the 8C survey, this catalogue contains the lowest frequency flux density measurement. The source detection region extends to 50° from zenith. As a result of correcting the effect of the primary beam, the background noise increases from zenith towards the horizon. However, within 50° from zenith the noise is roughly constant or increases slowly. Beyond 50° however, the background noise is greater than twice the noise at zenith and increases dramatically. Given that zenith is towards 52.°9 declination at the LOFAR superterp, the survey covers the full Northern hemisphere. This is illustrated in Fig. 6, where the coverage area of the AARTFAAC catalogue is compared to the Perley & Butler (2017), Helmboldt et al. (2008), and the 8C (Rees 1990) catalogues.

This catalogue will also be beneficial as an independent check for the calibration of low frequency, wide field images generated by the LWA, and other LOFAR-LBA surveys such as the forthcoming MSSS. Additionally, the technique presented here can be implemented for real-time streaming calibration of the Southern hemisphere arrays MWA and SKA-LOW.

5.2 Position

The uncertainties in the position measurements by AARTFAAC are large relative to other surveys due to the lower resolution of the images. Fortunately, given the threshold to which we detect sources, the resulting number density in the sky is such that this does not cause a problem when associating measurements of any source across the set of images in an observation. Correspondingly, within an association radius of 1° any AARTFAAC source can be uniquely matched. Moreover, since the primary goal is the creation of a catalogue for flux density calibration to compare with future transients, highly accurate source positions are not essential.

Nevertheless, the best estimate of the position of each source was measured. This was done by taking the mean, weighted by the uncertainty in the fitted position from each extraction.

5.3 Reference source fit

The accuracy with which the fluxes of the modelled reference sources are measured from the images, after calibration, is a validation of the models. This is due to the fact that, as more sources are added to the catalogue and incorporated into the calibration

Downloaded from https://academic.oup.com/mnras/article-abstract/482/2/2502/5136428 by Universiteit van Amsterdam user on 02 April 2020

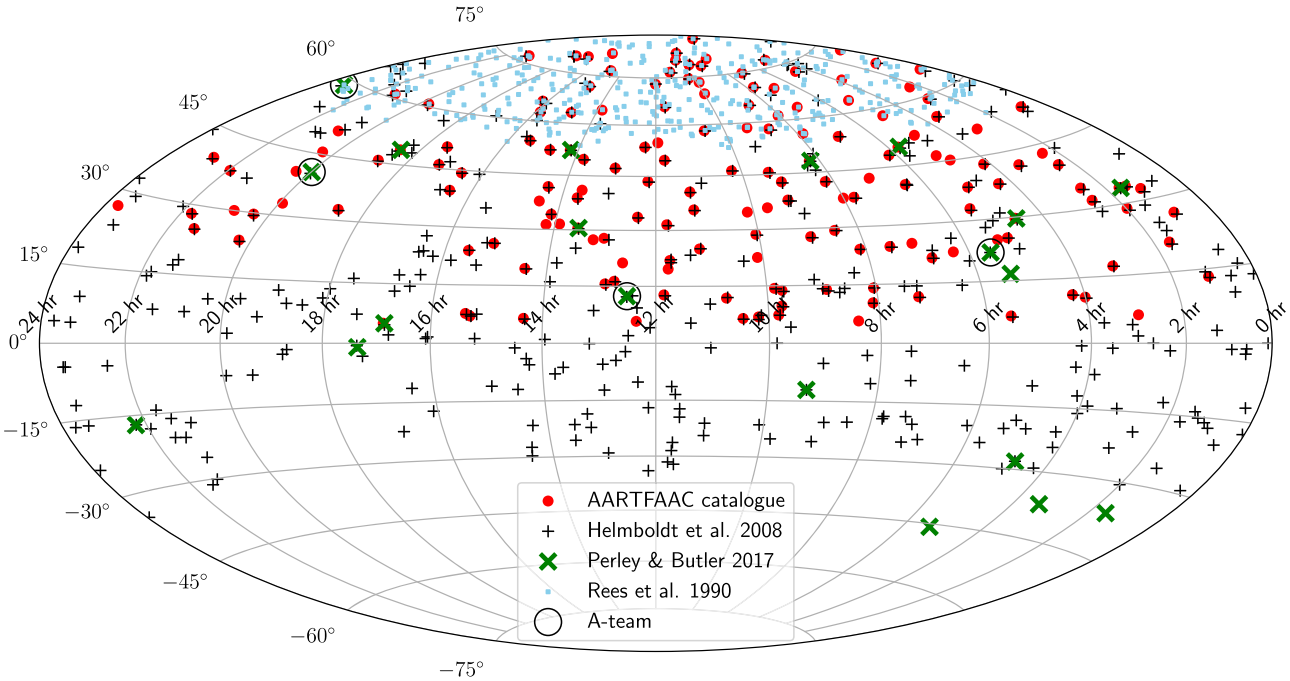

Figure 6. A comparison of the 167 AARTFAAC catalogue sources, and the primary reference sources: 23 in Perley & Butler (2017) of which eight are observable above a declination of 0°, and 388 from Helmboldt et al. (2008) of which 120 are associated with AARTFAAC sources.

scheme, the inferred flux density of a single incorrectly modelled source would diverge from the initial estimate, due to the influence of the other correctly modelled sources on the flux density scaling fit for the entire image.

In order to illustrate the resulting goodness of fit between the modelled reference fluxes used in calibration and the resulting measurements, the spectrum from Perley & Butler (2017), with our AARTFAAC data point (open circle), is shown in Fig. 4. Here, the AARTFAAC measured flux density does not agree, within the uncertainty of the reference model (blue region). In fact, Fig. 7 and Table 3 show that the only source for which the reference model and AARTFAAC measurement agree within their combined uncertainty is Hercules A, the brightest source.

However, a strong linear relationship between the reference flux density and the measured flux density is illustrated in Fig. 7. This illustrates the linear response of the array to sources from 50 to over 800 Jy and suggests that the relative brightnesses of the models are not completely accurate at 60 MHz.

Lastly, our flux density values were compared to those measured by the Owens Valley Long Wavelength Array (OVRO-LWA; Eastwood et al. 2018). The OVRO-LWA flux density values compared in Table 3 are derived by interpolating between the measured values, provided in their supplementary online material, and the uncertainties calculated using a Monte Carlo method to randomly draw a population of flux density measurements within the reported uncertainty range. Notably, the sources 3C 48 and 3C 380 both agree within mutual uncertainties. This is interesting due to the similarity of the instruments, but very different method for calibration and imaging.

Lastly, it is apparent that the flux density measured here for 3C 286 is significantly higher than what is reported by all three of the reference catalogues. It is unclear what could cause this for a single source. 3C 286 is a very well-known calibrator. In an

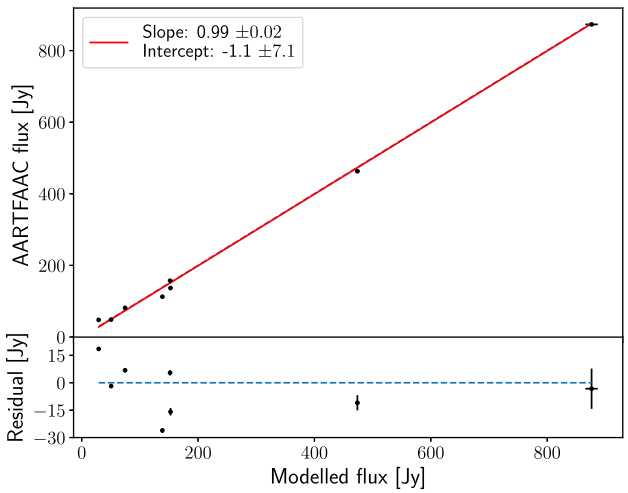

Figure 7. Comparing the modelled flux density of the initial calibrators, from Perley & Butler (2017) to their AARTFAAC measured flux density. This shows excellent relative brightness agreement across the sources. The exception is the faintest source 3C 286, with a modelled flux density of 29.15 ± 0.25 Jy and a measured flux density of 47.69 ± 0.12 Jy. Measurements reported by Helmboldt et al. (2008) also have a large scatter around the model across the frequency range. This may be due to systematic differences among the different surveys.

attempt to explain the additional flux, measurements by a single dish instrument, in which 3C 286 is unresolved (Ott et al. 1994) were compared. However, this study yielded results that agree with the measurements of Perley & Butler (2017) using the Very Large Array, indicating that we are not observing additional diffuse flux as in the case of the Tycho supernova remnant. Additionally, given that 3C 286 is at a high Galactic latitude, 10° north of the North Galactic

Downloaded from https://academic.oup.com/mnras/article-abstract/482/2/2502/5136428 by Universiteit van Amsterdam user on 02 April 2020

Table 3. A comparison of the difference between the AARTFAAC inferred flux density, and values modelled from the spectral fits presented in the reference catalogues. The modelled flux density, and associated uncertainties, for the Perley & Butler (2017) catalogue were generated via a Monte Carlo method, by fitting a spectrum to random samples of the flux density measurements. These values were used as the initial flux density scale for bootstrapping to the entire AARTFAC catalogue. Additionally for comparison, the modelled flux density from Scaife & Heald (2012) are shown. These flux density estimates, and uncertainties, were generated using the coefficients and their uncertainty in the spectral model published. As such, the uncertainty in these flux density estimates at 60 MHz is much higher. Lastly, our values are compared to the results of Eastwood et al. (2018), interpolated to 60 MHz.

Source name	Scaife and Heald model (Jy) @ 60 MHz	Perley and Butler model (Jy) @ 60 MHz	AARTFAAC measured (Jy) @ 60 MHz	Eastwood et al. interpolated (Jy) @ 60 MHz
3C 48	77 ± 5	74.55 ± 0.64	81.36 ± 0.33	83.7 ± 2.29
3C 123	–	473.75 ± 3.87	462.85 ± 0.33	–
3C 147	43 ± 4	50.59 ± 0.66	48.71 ± 0.14	44.89 ± 1.02
3C 196	151 ± 5	151.84 ± 1.25	157.30 ± 0.20	–
3C 286	33 ± 2	29.15 ± 0.25	47.69 ± 0.12	32.63 ± 0.43
3C 295	134 ± 11	138.65 ± 1.13	112.55 ± 0.18	125.22 ± 3.44
3C 380	156 ± 4	152.63 ± 1.95	136.75 ± 0.09	134.49 ± 3.70
Hercules A	–	876.31 ± 10.55	873.07 ± 0.37	–

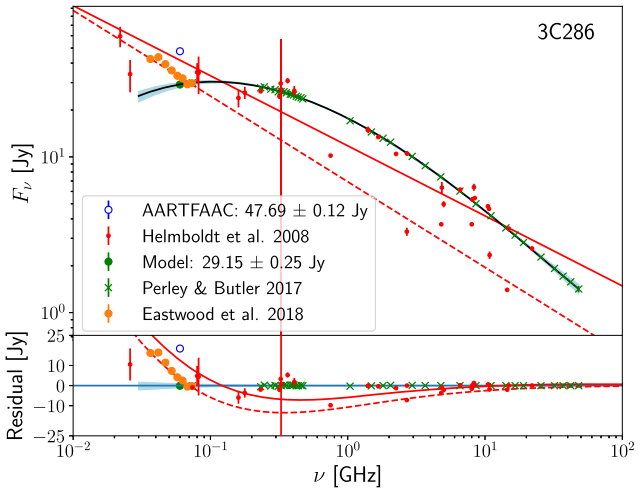

Figure 8. A comparison of the collected reference measurements of the flux density of 3C 286, compared to the AARTFAAC measurements (open circle). The Perley & Butler (2017) measurements (green x's) and modelled spectra (black line) was used for the initial catalogue bootstrapping. While values reported by Eastwood et al. (2018, orange circle) are direct measurements at frequencies comparable to AARTFAAC with the OVRO-LWA, an instrument with a similar design to AARTFAAC. However, the flux densities reported by Helmboldt et al. (2008, red dots) show the high variance in published flux densities across the spectrum. Two simple power laws (solid and dashed red lines) fitted by Helmboldt et al. (2008) are also shown. Interestingly, both AARTFAAC and OVRO-LWA agree better with the simple power laws shown here.

Spur, it is unlike likely the additional flux is the result of diffuse Galactic emission removed by simply imposing the minimum baseline of 10λ. Further, there is no correlation between sky location and deviations between the modelled and measured flux densities. When viewing the AARTFAAC measured flux density alongside the flux density values and simple spectra reported by Helmboldt et al. (2008), illustrated in Fig. 8, the difference does not appear as anomalous. In fact, the power-law spectra fit by Helmboldt et al. (2008) predict a flux density at 60 MHz of 42 Jy. While these data and spectra are less precise than those measured by Perley & Butler (2017), it is notable that both the results from AARTFAAC and the OVRO-LWA are better predicted by these spectra.

Table 4. A list of sources for which a spectral turnover was predicted, which we do not observe. Here, the signifier from the AARTFAAC and VLSSr are given. In the supplemental online material, the AARTFAAC flux density measurements are compared alongside the measurements and spectral fits in fig. 1 of Helmboldt et al. (2008).

AARTFAAC label		VLSSr label
J0110 + 322	-	J0110 + 315
J1011 + 068	-	J1011 + 062
J1147 + 496	-	J1146 + 495
J1445 + 768	-	J1447 + 766
J1630 + 442	-	J1629 + 442

5.4 Spectral turnovers

Some of the spectral models fitted by Helmboldt et al. (2008) predict a turnover below the lowest frequency at which measurements were made. Nevertheless, the new flux density measurements made at 60 MHz clearly indicate that, instead, the spectral shape of at least six of these sources continue to rise. The spectra for the six sources, whose labels from both the AARTFAAC catalogue and VLSSr catalogue are listed in Table 4, can be seen in the supplemental online material, where the AARTFAAC measurement is plotted alongside the flux density measurements and spectral fits from Helmboldt et al. (2008). An example of these figures is shown in Fig. 9, where the flux density measurements from Klein et al. (1979) are plotted as well.

Further ongoing flux density studies, across the full observational spectrum of the LOFAR LBA, 10–90 MHz, could potentially reveal turnovers at a frequency lower than 60 MHz.

5.5 Blended sources

One issue with interferometers is that the minimum baseline length determines the sensitivity of the instrument to regions with large diffuse emission. Therefore, objects with a larger angular size will

Downloaded from https://academic.oup.com/mnras/article-abstract/482/2/2502/5136428 by Universiteit van Amsterdam user on 02 April 2020

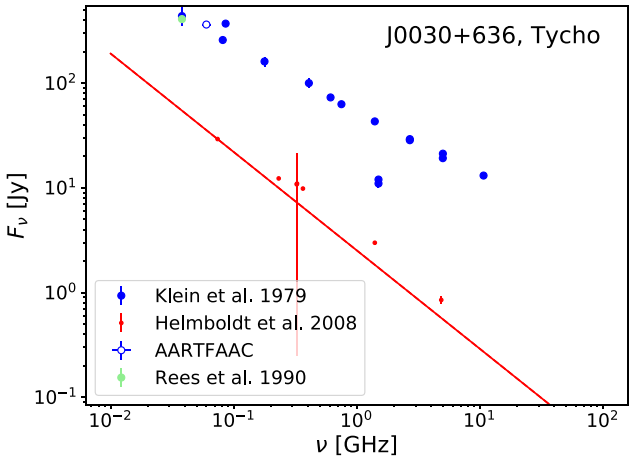

Figure 9. The spectra of the supernova remnant 1572, also known as Tycho. The measurements reported by Helmboldt et al. (2008, red dots) show a significantly reduced integrated flux density compared to the AARTFAAC measurement at 60 MHz. This is due to the higher resolution of these surveys resolving out power in the diffuse emission. Comparing this to the multifrequency single dish measurements by Klein et al. (1979, blue circles), the 8C catalogue (Rees 1990, green circle), and AARTFAAC (open circle) that fully recover flux density on this scale like a single dish instrument, as a result of their lower resolution.

have their diffuse emission, at least partially, resolved out by interferometers that achieve a high angular resolution. This reduces the total apparent flux density of diffuse sources when compared to measurements by a single dish instrument. Consequently, in order to remove the large, bright, diffuse structure of the Milky Way, we eliminate all baselines below 10λ. This effectively eliminates much of the Galactic emission that would otherwise be a dominant foreground.

In observing large diffuse sources, such as nearby supernova remnants, we see that AARTFAAC recovers the total integrated flux density as effectively as a single dish instrument. This effect is shown in Fig. 9 for the supernova remnant Tycho, where the flux density measurements (red dots) and spectral fit (red line) reported by Helmboldt et al. (2008) are compared with a multiwavelength analysis observed with a single dish instrument by Klein et al. (1979, blue circles). Additionally, the integrated flux density value from the 8C catalogue is plotted in green. Consequently, the sensitivity of AARTFAAC on larger angular scales presents an additional use case beyond transients; for example, to study Galactic emission, and diffuse regions around other sources such as Cassiopeia A and Cygnus A.

Due to the relatively low resolution of AARTFAAC, compared to the 8C and VLSSr surveys, some objects reported here as individual sources are known to be composed of two or more sources blended together. By comparing the AARTFAAC catalogue to VLSSr with integrated flux density greater than 10 Jy, we find that the sources listed in Table 5 are the result of multiple blended objects.

However, so long as the sum of these blended sources maintains a stable flux density, the component contribution of each is not important for us, since AARTFAAC does not resolve them independently. Therefore, no effort was made to de-blend the sources, and attempt to retrieve an accurate flux density for each individually. The blended sources, as they appear to AARTFAAC, are still useful for our calibration purposes. Although instruments with a higher resolution, including AARTFAAC after the currently planned extension that will incorporate six additional stations, may need to exclude these when calibrating, and measure the separate component fluxes.

Table 5. A list of sources in the AARTFAAC catalogue that are known to be composed of two or more VLSSr sources, blended together. A threshold of 10 Jy was used to filter the VLSSr catalogue in order to limit the number of objects compared. This threshold is motivated by the fact that a source below 10 Jy at 74 MHZ would, given the sensitivity of AARTFAAC, likely not have a strong contribution to the observed flux density.

AARTFAAC label	VLSSr fluxes (Jy)	AARTFAAC flux density (Jy)
J0027 + 643	28.24 ± 0.19	362.42 ± 0.27
–	17.62 ± 0.19	–
–	13.51 ± 0.20	–
J0110 + 134	56.78 ± 0.10	122.93 ± 0.23
–	23.69 ± 0.10	–
J0224 + 430	12.77 ± 0.15	96.47 ± 0.10
–	10.16 ± 0.16	–
J0400 + 105	23.88 ± 0.26	132.39 ± 0.15
–	10.85 ± 0.26	–
J0420 + 381	38.76 ± 0.21	167.93 ± 0.09
–	29.24 ± 0.21	–
–	13.44 ± 0.21	–
J0506 + 381	86.76 ± 0.24	237.15 ± 0.13
–	25.40 ± 0.24	–
–	14.56 ± 0.24	–
–	11.31 ± 0.24	–
J0627 + 401	16.76 ± 0.10	44.95 ± 0.06
–	16.31 ± 0.09	–
J0657 + 542	39.71 ± 0.09	56.75 ± 0.06
–	12.61 ± 0.10	–
J0829 + 292	19.55 ± 0.10	39.14 ± 0.08
–	15.38 ± 0.10	–
J0855 + 139	40.21 ± 0.17	64.32 ± 0.13
–	15.17 ± 0.18	–
J1144 + 218	30.70 ± 0.14	50.02 ± 0.16
–	16.40 ± 0.13	–
J1445 + 768	18.67 ± 0.13	45.58 ± 0.10
–	14.72 ± 0.14	–
J1506 + 259	28.97 ± 0.12	127.36 ± 0.13
–	11.99 ± 0.12	–
J1552 + 050	373.91 ± 0.49	873.07 ± 0.37
–	309.02 ± 0.51	–
J1840 + 797	22.01 ± 0.32	117.25 ± 0.14
–	18.60 ± 0.31	–
–	10.92 ± 0.31	–
J2247 + 397	34.83 ± 0.18	150.89 ± 0.15
–	25.58 ± 0.19	–
–	25.46 ± 0.18	–

5.6 Summary of flux calibration and measurement accuracy

From all the above, it is clear that a number of significant factors play a role in the accuracy with which AARTFAAC can report calibrated source fluxes: the stability and sensitivity of our instrument, the stability of the ionosphere, and the ability to relate our instrumental brightness measurements to well-studied flux calibrators. Here, we collect and summarize our findings on these aspects.

Instrumental flux stability. We collected data from all LSTs, over a 3-month period, and find that longer time average fluxes show no trends with time or Zenith angle at levels above 1 per cent, so indeed our measurements and instrumental calibration are quite stable (Figs 3 and 5). Furthermore, we see that for bright sources the error estimates on 1-s measurements are a bit larger than the

Downloaded from https://academic.oup.com/mnras/article-abstract/482/2/2502/5136428 by Universiteit van Amsterdam user on 02 April 2020

measured flux variations on short time scales (Fig. 5), so our error estimates are somewhat conservative.

Ionospheric effects. We have used data from representative ionospheric weather, however excluding periods of either strong RFI or ionospheric turbulence around the A-team sources that resulted the data being flagged by the correlator or calibration pipelines, and no images being created. We see that in these data, the dominant timescale for ionospheric variations is of the order of 15 min at our observing frequency (60 MHz) and the typical amplitude is 10–15 per cent, somewhat larger than the instantaneous flux measurement accuracy of bright sources (Fig. 5). We also find that these variations are spatially uncorrelated on scales more than 5°. This is why we employ the strategy of fitting instantaneous flux scales using all available AARTFAAC catalogue sources at any time: it decreases the uncertainty in the calibration scale factor and prevents the scintillation variations of a single calibrator to dominate the flux scale. Because we do this every second and monitor the variations, we are provided with an automatic assessment of ionospheric conditions, which is also used by LOFAR. These effects are much less in MWA at somewhat higher frequencies, see Loi et al. (2015c), indicating that even within the LOFAR low band the strength will vary significantly with frequency.

Absolute calibration. To tie the AARTFAAC flux scale to more widely applicable radio flux calibrations, we compared our fluxes to a number of previously established radio catalogues. We had to fit models and interpolate, since very few previous measurements are available at 60 MHz (Section 5). In Table 3, we can see the precision and stability of our measurements is indeed very good compared to previous work, but that the calibrations of different very bright northern sources differ by a few to 10 per cent between papers, and in a few cases more (specifically our flux for 3C 286 seems anomalously high compared to other work). Absolute flux calibration at these radio frequencies thus appears to be mostly reliable to 10 per cent.

6 CONCLUSION

This work presents the method used for calibrating the flux density scale of AARTFAAC images in real-time for the upcoming transient and variability surveys. Additionally, the AARTFAAC catalogue of calibration sources is presented.

Due to the lack of a sufficient number of well-measured calibrator sources at low frequencies, and the requirement that AARTFAAC images are calibrated in real-time, a bootstrap algorithm was used. Hence, the AARTFAAC catalogue is fundamentally based on a flux density scaling derived from the spectral fits published by Perley & Butler (2017). Consequently, any systematic bias in the fluxes reported there could influence the AARTFAAC catalogue. Therefore, as a verification the AARTFAAC catalogue was compared to the larger, but less precise, catalogue of Helmboldt et al. (2008). Here, good agreement was found between the spectral fit extrapolated to 60 MHz and the measurements presented in this work. However, a tendency for AARTFAAC to measure more flux was observed. This is potentially explained by the much higher density of the LOFAR superterp resulting in higher sensitivity to diffuse emission. Therefore, it is shown that AARTFAAC is capable of filling the gap between 38 and 74 MHz and providing an accurate flux density scale for the calibration of densely packed LOFAR.

Additionally, statistical analysis of the times series of flux density measurements for each source resulted in insights into the effect of ionospheric variability. Significantly, it was observed that such variability preferentially skews the distribution of measurements in the positive direction, giving the average a positive bias. Consequently, it was found that the most robust method to mitigate this was to fit a skew normal distribution, and calculate the mode. In light of this, observations that attempt to achieve high sensitivity with long integrations, without correcting for ionospheric variability on short timescales, will also have a positive bias.

For our ability to detect new transients, the flux calibration is of minor importance, because this depends only on the ability to detect a new source against the local noise. The main limitations for this are (i) that we examine large volumes of data, so we typically have to set the threshold at eight times the local rms to avoid many false positives, and (ii) that we have to distinguish transients from RFI, terrestrial sources, and artefacts. But if we do detect a source, the instantaneous flux uncertainty on second time scales will be dominated by the ionosphere for bright sources, and roughly equally by the ionospheric and measurement noise near the threshold. The ionospheric uncertainty decreases significantly only when the sources stay on for longer than the typical variation time of 15 min.

In the future, this work may be extended by incorporating more data. Given that the primary science goal of AARTFAAC is to survey the sky for transient and variable sources, many hundreds of hours of observations will be made. This enormous amount of data will allow for extremely precise flux density values to be measured.

Lastly, observations across the entire frequency range of the LOFAR LBA 10–90 MHz would allow highly accurate spectra to be fit, however at the low end of the bandpass, Perley & Butler (2017) would no longer be a suitable reference catalogue. Lastly, the planned upgrade of AARTFAAC to include 12 stations of the LOFAR core, rather than the six currently in operation, will produce images of higher sensitivity and resolution. This will allow for the measurement of currently blended sources, and many that are now below our detection threshold.

ACKNOWLEDGEMENTS

This work is funded by the ERC Advanced Investigator grant no. 247295 awarded to Professor Ralph Wijers, University of Amsterdam. We thank The Netherlands Institute for Radio Astronomy (ASTRON) for support provided in carrying out the commissioning observations.

The authors would also like to thank the LOFAR science support for their assistance in obtaining and processing the data used in this work. We use data obtained from LOFAR, the LOFAR designed and constructed by ASTRON, which has facilities in several countries, which are owned by various parties (each with their own funding sources), and that are collectively operated by the International LOFAR Telescope (ILT) foundation under a joint scientific policy.

REFERENCES

Bell M. E. et al., 2014, MNRAS, 438, 352
Carbone D. et al., 2016, MNRAS, 459, 3161
Carbone D. et al., 2018, Astron. Comput., 23, 92
Eastwood M. W. et al., 2018, AJ, 156, 32
Green D. A., 2014, Bull. Astron. Soc. India, 42, 47
Helmboldt J. F., Kassim N. E., Cohen A. S., Lane W. M., Lazio T. J., 2008, ApJS, 174, 313
Hyman S. D., Lazio T. J. W., Kassim N. E., Ray P. S., Markwardt C. B., Yusef-Zadeh F., 2005, Nature, 434, 50
Klein U., Emerson D. T., Haslam C. G. T., Salter C. J., 1979, A&A, 76, 120
Lane W. M., Cotton W. D., van Velzen S., Clarke T. E., Kassim N. E., Helmboldt J. F., Lazio T. J. W., Cohen A. S., 2014, MNRAS, 440, 327

Downloaded from https://academic.oup.com/mnras/article-abstract/482/2/2502/5136428 by Universiteit van Amsterdam user on 02 April 2020

Loi S. T. et al., 2015a, Geophys. Res. Lett., 42, 3707
Loi S. T. et al., 2015b, Radio Sci., 50, 574
Loi S. T. et al., 2015c, MNRAS, 453, 2731
Murphy T. et al., 2017, MNRAS, 466, 1944
Obenberger K. S. et al., 2014, ApJ, 788, L26
Obenberger K. S. et al., 2015, J. Astron. Instrum., 4, 1550004
Offringa A. R. et al., 2013, A&A, 549, A11
Ott M., Witzel A., Quirrenbach A., Krichbaum T. P., Standke K. J., Schalinski C. J., Hummel C. A., 1994, A&A, 284, 331
Patil A. H. et al., 2017, ApJ, 838, 65
Perley R. A., Butler B. J., 2017, ApJS, 230, 7
Petroff E. et al., 2015, MNRAS, 447, 246
Pietka et al., 2015, MNRAS, 446, 3687
Prasad P., Wijnholds S. J., Huizinga F., Wijers R. A. M. J., 2014, A&A, 568, A48
Prasad P. et al., 2016, J. Astron. Instrum., 5, 1641008
Rees N., 1990, MNRAS, 244, 233
Rowlinson A. et al., 2016, MNRAS, 458, 3506
Salvatier J., Wiecki T., Fonnesbeck C., 2015, preprint (arXiv:1507.08050)
Scaife A. M. M., Heald G. H., 2012, MNRAS, 423, L30
Stewart A. J. et al., 2016, MNRAS, 456, 2321
Swinbank J. D. et al., 2015, Astron. Comput., 11, 25
van Haarlem M. P. et al., 2013, A&A, 556, A2

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

AARTFAAC-Helmboldt-comparison.pdf
AARTFAAC_catalogue_2.csv

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

APPENDIX A: SUPPLEMENTARY ONLINE MATERIAL: AARTFAAC CATALOGUE

Table A1. The first 10 source, as an example, of the AARTFAAC Northern hemisphere catalogue at 60MHz. The full catalogue as a machine readable csv is available online. The labels here are derived from the positions measured, as in the VLSSr, the positions are given in degrees with the uncertainty calculated from the standard deviation of the individual position measurements. The flux density values, and uncertainties, are calculated from the mode of a skew-normal distribution fit to the population of measurements. The parameters of the skew normal are given. Common names are given to matching sources where possible. Lastly, the number of images from which the light curve is generated is given.

Label	Ra (°)	Dec. (°)	Flux density (Jy)	Location ξ	Scale ω	Shape α	Name	Measurements #
J0007 + 725	1.96 ± 0.36	72.53 ± 0.12	86.17 ± 0.13	75.94 ± 0.09	22.17 ± 0.09	3.27 ± 0.05	CTA 1	72 070
J0027 + 642	6.78 ± 0.10	64.22 ± 0.10	362.42 ± 0.27	336.64 ± 0.31	48.75 ± 0.27	2.10 ± 0.04	Tycho SNR	55 465
J0044 + 521	11.19 ± 0.05	52.20 ± 0.11	70.75 ± 0.09	63.80 ± 0.08	14.30 ± 0.08	2.86 ± 0.05	3C 20	35 940
J0057 + 266	14.45 ± 0.06	26.66 ± 0.09	72.75 ± 0.14	65.78 ± 0.30	13.20 ± 0.20	1.17 ± 0.06	3C 28	18 137
J0058 + 683	14.64 ± 0.20	68.39 ± 0.11	91.20 ± 0.08	85.74 ± 0.26	11.45 ± 0.14	0.86 ± 0.05	3C 27	41 301
J0110 + 322	17.54 ± 0.09	32.26 ± 0.11	61.68 ± 0.12	55.01 ± 0.12	12.99 ± 0.11	2.39 ± 0.07	3C 34	19 807
J0110 + 134	17.59 ± 0.03	13.45 ± 0.09	122.93 ± 0.23	111.74 ± 0.34	20.57 ± 0.27	1.58 ± 0.06	3C 33	11 207
J0127 + 332	21.97 ± 0.12	33.30 ± 0.13	51.72 ± 0.14	47.55 ± 0.22	7.68 ± 0.17	1.54 ± 0.11	3C 41	5048
J0137 + 210	24.47 ± 0.04	21.08 ± 0.07	89.87 ± 0.18	79.84 ± 0.23	18.60 ± 0.19	1.83 ± 0.06	3C 47	16 127
J0139 + 332	24.84 ± 0.04	33.29 ± 0.10	81.36 ± 0.33	62.62 ± 0.26	39.52 ± 0.26	3.05 ± 0.07	3C 48	25 907

This paper has been typeset from a TEX/LATEX file prepared by the author.

Downloaded from https://academic.oup.com/mnras/article-abstract/482/2/2502/5136428 by Universiteit van Amsterdam user on 02 April 2020