Apercal-The Apertif calibration pipeline

Adebahr, B.; Schulz, R.; Dijkema, T.J.; Moss, V.A.; Offringa, A.R.; Kutkin, A.; van der Hulst, J.M.; Frank, B.S.; Vilchez, N.P.E.; Verstappen, J.; Adams, E.K.; de Blok, W.J.G.; Denes, H.; Hess, K.M.; Lucero, D.; Morganti, R.; Oosterloo, T.; Pisano, D.-J.; Ivashina, M.V.; van Cappellen, W.A.; Connor, L.D.; Coolen, A.H.W.M.; Damstra, S.; Loose, G.M.; Maan, Y.; Maccagni, F.M.; Mika, A.; Mulder, H.; Oostrum, L.C.; Orrú, E.; Smits, R.; van der Schuur, D.; van Leeuwen, J.; Vohl, D.; Wijnholds, S.J.; Ziemke, J.

Link to publication

Apercal - The Apertif Calibration Pipeline

B. Adebahr[a,b,*], R. Schulz[b], T. J. Dijkema[b], V. A. Moss[c,d,b], A. R. Offringa[b,f], A. Kutkin[b,h], J. M. van der Hulst[f], B. S. Frank[i,j,k],
N. P. E. Vilchez, J. Verstappen[f,b], E. K. Adams[b,f], W. J. G. de Blok[b,k,f], H. Denes[b], K. M. Hess[b,f], D. Lucero[g], R. Morganti[b,f],
T. Oosterloo[b,f], D.-J. Pisano[l], M. V. Ivashina[m], W. A. van Cappellen[b], L. D. Connor[n,b], A. H. W. M. Coolen[b], S. Damstra[b],
G. M. Loose[b], Y. Maan[b], F. M. Maccagni[q], A. Mika[b], H. Mulder[b], L. C. Oostrum[b,n,p], E. Orrú[b], R. Smits[b], D. van der Schuur[b],
J. van Leeuwen[b,n], D. Vohl[b], S. J. Wijnholds[b], J. Ziemke[b,o]

[a]*Astronomisches Institut der Ruhr Universität Bochum (AIRUB), Bochum, Germany*
[b]*ASTRON, Dwingeloo, The Netherlands*
[c]*CSIRO Astronomy and Space Science, Sydney, Australia*
[d]*University of Sydney, Sydney, Australia*
[e]*Cavendish Laboratory, Cambridge, UK*
[f]*Kapteyn Astronomical Institute, University of Groningen, Groningen, The Netherlands*
[g]*Department of Physics, Virginia Tech, Blacksburg, VA, USA*
[h]*Astro Space Center of Lebedev Physical Institute, Profsoyuznaya Str. 84/32, 117997 Moscow, Russia*
[i]*South African Radio Astronomy Observatory (SARAO), 2 Fir Street, Observatory, 7925, South Africa*
[j]*Inter-University Institute for Data Intensive Astronomy and Department of Astronomy, University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa*
[k]*University of Cape Town, Cape Town, University Private Bag Rondebosch, 7701, South Africa*
[l]*Department of Physics and Astronomy, West Virginia University, White Hall, Box 6315, Morgantown, WV 26506, USA*
[m]*Dept. of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden*
[n]*Anton Pannekoek Institute, University of Amsterdam, Postbus 94249, 1090 GE Amsterdam, The Netherlands*
[o]*Rijksuniversiteit Groningen Center for Information Technology, P.O. Box 11044, 9700 CA Groningen, the Netherlands*
[p]*Netherlands eScience Center, Science Park 140, 1098 XG, Amsterdam, The Netherlands*
[q]*INAF – Osservatorio Astronomico di Cagliari, Via della Scienza 5, 09047 Selargius, CA, Italy*

Abstract

Apertif (APERture Tile In Focus) is one of the Square Kilometre Array (SKA) pathfinder facilities. The Apertif project is an upgrade to the 50-year-old Westerbork Synthesis Radio Telescope (WSRT) using phased-array feed technology. The new receivers create 40 individual beams on the sky, achieving an instantaneous sky coverage of 6.5 square degrees. The primary goal of the Apertif Imaging Survey is to perform a wide survey of 3500 square degrees (AWES) and a medium deep survey of 350 square degrees (AMES) of neutral atomic hydrogen (up to a redshift of 0.26), radio continuum emission and polarisation. Each survey pointing yields 4.6 TB of correlated data. The goal of *Apercal* is to process this data and fully automatically generate science ready data products for the astronomical community while keeping up with the survey observations. We make use of common astronomical software packages in combination with Python based routines and parallelisation. We use an object oriented module-based approach to ensure easy adaptation of the pipeline. A Jupyter notebook based framework allows user interaction and execution of individual modules as well as a full automatic processing of a complete survey observation. If nothing interrupts processing, we are able to reduce a single pointing survey observation on our five node cluster with 24 physical cores and 256 GB of memory each within 24h keeping up with the speed of the surveys. The quality of the generated images is sufficient for scientific usage for 44 % of the recorded data products with single images reaching dynamic ranges of several thousands. Future improvements will increase this percentage to over 80 %. Our design allowed development of the pipeline in parallel to the commissioning of the Apertif system.

Keywords:
techniques: image processing, surveys, telescopes, pipelines, data reduction

1. Introduction

Radio astronomy is currently moving into a survey era with most of the new or upgraded facilities being designed as survey telescopes. The technological advances over the past few decades in receiver design allow the usage of broadband receivers with fine spectral resolution. In addition, Phased Array Feeds (PAFs) are enlarging the field of view by an order of magnitude by forming tens of adjacent beams simultaneously. The increase in computing power and easy access to powerful computing facilities allows the use of these receiver technologies for radio interferometric imaging. The amount of data produced by modern radio interferometric observations already reaches several terabytes for observation times of a few hours posing a new challenge to the processing of the data.

The Apertif system is an upgrade of the receiver system of

_____
[*]Corresponding author
Email address: adebahr@astro.rub.de (B. Adebahr)

the old Westerbork Synthesis Radio Telescope (WSRT) located in the Netherlands with PAF technology. Until January 2021 the receivers were operating at a central L-band frequency of 1280 MHz, which was moved to 1370 MHz afterwards. The instantaneous bandwidth is 300 MHz. 384 individual subbands consisting of 64 spectral channels each provide a spectral resolution of 12.2 kHz. 40 individual adjacent beams are formed on the sky achieving an instantaneous sky coverage of 6 square degrees. Voltages are correlated to generate data in all four Stokes parameters. These specifications make Apertif an ideal survey instrument for continuum, polarisation and HI-surveys of the Northern sky. For a detailed overview of the Apertif system we refer the reader to van Cappellen et al. (2021).

The deluge of data produced by this new generation of radio telescopes requires an automatic and/or semi-automatic approach. Several calibration pipelines have been designed over the last few years with different purposes in mind. Many aim for general data reduction of a broad range of different imaging observations, mostly data from a specific facility, like e.g. the CASA pipeline for the Jansky Very Large Array (JVLA) and the Atacama Large Millimeter Array (Kent et al., 2018; Muders et al., 2014), DP3[1] and prefactor[2] for the Low Frequency Array (LOFAR), FLAGCAL for the Giant Metrewave Radio Telescope (GMRT) (Prasad and Chengalur, 2012), ASKAPsoft[3] for the Australian Square Kilometre Array Pathfinder (ASKAP), CARACal[4] (formerly MeerKATHI) (Józsa et al., 2020a,b) for the MeerKAT telescope[5] or rPICARD for Very Long Baseline Interferometry (Janssen et al., 2019). Others are tailored towards individual surveys like the VLA Sky Survey (VLASS) pipeline (Marvil et al., 2019) or on a specific aspect like e.g. the improvement of the image quality of LOFAR data using a direction dependent calibration technique (van Weeren et al., 2016; Tasse et al., 2018) or even completely new approaches like stochastical imaging (Yatawatta, 2020). An example for the inclusion of older software packages into a modern pipeline is the Obit-based[6] MeerKAT SDP pipeline[7]. First steps have been undertaken to use supercomputing architecture for these pipelines (Mechev et al., 2019; Sabater et al., 2017; Drabent et al., 2019). The goal of Apercal, the dedicated imaging and calibration pipeline for the Apertif system is to provide an easy to use framework generating science ready images to the user in an automatic way while keeping modularity and transparency.

Apertif conducts an imaging and a time-domain survey (Adams and van Leeuwen, 2019). Here we focus on the imaging survey which comprises two tiers: The Apertif Wide Extragalctic Survey (AWES) and the Apertif Medium deep Extragalactic Survey (AMES) which are designed to cover 3500 and 350 square degrees, respectively. These surveys are able to probe the neu-

tral hydrogen content of the Universe up to a redshift of 0.26 and the faint general and polarised radio source population. The footprints of the surveys are vastly overlapping with the ongoing LOFAR LoTSS (Shimwell et al., 2017) and future WEAVE (Smith et al., 2016) and LOFAR LBA Sky Survey (LoLSS) (de Gasperin et al., 2021). In addition, the survey areas have been chosen such that the observing efficiency is optimised and that there is good overlap with regions that have important ancillary photometric and spectroscopic data in the optical and the IR. SDSS[8] (Blanton et al., 2017), Pan-STARRS[9] (Chambers et al., 2016), and WISE[10] (Wright et al., 2010) provide optical and near IR photometric coverage over large portions of the surveys while Herschel-Atlas [11] (Valiante et al., 2016) provides far IR photometric coverage over a portion of AMES. ManGA[12] (Yan et al., 2016) and WEAVE[13] (Dalton et al., 2014) provide targeted spectroscopic coverage of optically selected galaxies and/or radio selected targets. HetDex[14] (Hill et al., 2008) will provide untargeted spectroscopic coverage over a large area of the northern sky. These synergies allow multi-wavelength studies for a statistically significant number of objects.

The Apertif system is installed on 12 (RT2-RTD) of the 14 dishes of the existing WSRT array (van Cappellen et al., 2021) providing maximum baselines of 2.4 km. The array is aligned in an East-West-configuration, which requires nearly full synthesis 12 hr tracks for sufficient (u,v)-coverage to allow good imaging performance. A single survey pointing produces approximately 4.6 TB of raw visibility data. This large data volume provides a challenge to the calibration and analysis of the data, making the traditional hands-on approach unfeasible. Over the last decades several data reduction pipelines have been developed using software packages like The Netherlands East West Synthesis Telescope Array Reduction (NEW-STAR)[15], the Astronomical Image Processing System (AIPS) (Wells, 1985) and the Multichannel Image Reconstruction Image Analysis and Display (MIRIAD) (Sault et al., 1995).

The similarity of the Apertif data to the old WSRT data, together with the experience we collected from previous WSRT data reduction schemes, allowed us to design *Apercal*, a pipeline which is using modern astronomical software and suits the specific needs of the Apertif system.

In this paper we describe the workflow of *Apercal* including the routines used for reducing the data. The challenges of building a stable software package with exchangeable modules, automatic execution and user-interaction are highlighted. In Sect. 2 we give an overview of the Apertif surveys. We describe the design and workflow of the Apercal pipeline in Sect. 3 followed by detailed explanations of the individual modules and subroutines in Sect. 4. An additional *Apercal* independent mosaicking routine is described in Sect. 5 The final data products and their

---

[1] https://github.com/lofar-astron/DP3

[2] https://github.com/lofar-astron/prefactor

[3] https://www.atnf.csiro.au/computing/software/askapsoft/sdp/docs/current/pipelines/introduction.html

[4] https://github.com/caracal-pipeline/caracal

[5] For some recent publications CARACal was also used for reducing VLA and GMRT data

[6] https://www.cv.nrao.edu/~bcotton/Obit.html

[7] https://github.com/ska-sa/katsdpcontim

[8] www.sdss.org

[9] panstarrs.ifa.hawaii.edu/pswww/

[10] https://wise2.ipac.caltech.edu/docs/release/allwise/

[11] www.h-atlas.org

[12] www.sdss.org/surveys/manga/

[13] www.ing.iac.es/weave/about.html

[14] hetdex.org

[15] https://github.com/lofar-astron/Newstar

ingest into the Apertif archive is described in Sect. 6. The quality assessment of the data products is illustrated in Sect. 7. We summarise and give an outlook to future improvements of *Apercal* in Sect. 8.

2. Apertif survey operations

In this section we give a brief overview of how Apertif imaging observations are executed and of completed and ongoing survey operations. For a more detailed description we refer the reader to Adams et al. in prep. and Hess et al. in prep.

Each Apertif imaging observation consists of two different types of observations: Calibrator and target observations. For each of the 40 beams in a target observation usually two calibrator observation are performed. Calibrator observations consist of short scans of 3-5 minutes on a standard reference flux calibration source (3C147 or 3C196) and a standard reference polarised calibration source (3C138 or 3C286). These observations are necessary for an absolute flux, polarisation leakage and polarisation angle calibration and are not used for any imaging. Due to the equatorial mount of the WSRT dishes, which keeps the beam pattern fixed on the sky while tracking, and the good amplitude and phase stability of the Apertif system, these scans only have to be scheduled every 12 to 24 hrs.

Target survey fields are observed uninterrupted for 11.5 hrs and are bracketed by two calibrator observations, a flux calibrator and a polarised calibrator source since full polarimetric calibration needs unpolarised as well as polarised calibrator solutions. The 11.5 hr exposure was found to be the best compromise between sufficient (u,v)-coverage guaranteeing the image quality and sensitivity for spatial scales up to several arcminutes, which is needed for the science goals of the surveys, and the optimal usage of telescope time for executing the survey.

3. Apercal design

Apercal is a radio astronomical data reduction pipeline for Apertif imaging observations which is based entirely on python2.7 (Van Rossum and Drake Jr, 1995). The pipeline in its final stage will be frozen to this python version. Additional software also incorporating data analysis routines and exchanges of pipeline modules with improvements to the pipeline routines will be developed in python3.

Each individual calibration step in the pipeline is encapsulated into a single module. This approach, by defining the input and output formats of the data and using independent data sub-directories, guarantees the independence of each individual module. It allows the exchange of single routines or even entire pieces of the software with minimal effort and impact onto the rest of the pipeline. Each module consists of individual subroutines, which are handling specific sub-calibration and -imaging tasks within a module. Each subroutine can be imported and executed independently. The whole pipeline is compatible with the Python Jupyter notebook framework[16], so

that user-interaction is simple and pipeline parameters can be changed on-the-fly. During the commissioning of the Apertif system details and parameters of the data reduction still had to be evaluated and were subject to several changes. At this stage the modular *Apercal* structure allowed a minimisation of the workload for the pipeline developers and commissioning team.

Apercal configuration parameters are read in by the pipeline via a simple text file. An example is given in Appendix A. Information in this file consist of the location of the data and the calibration parameters for the individual modules. If a parameter is not given in a file, a default value is read from a master configuration file.

Communication and transfer of parameters between modules is performed via a parameter file inside which the outcome of each individual calibration step is saved and can be accessed. An example can be found in Appendix B. Actual logging of information, warning or error messages from the pipeline is carried out via the python logging module. After an execution of a subroutine the parameter file is updated on disk saving the exact parameters used and the success or failure of the task. Entries in the parameter file are organised in a bibliography structure with a string-keyword and an attached data structure of any needed format. Each module and subroutine has certain prerequisites, which are mostly based on the quality and success or failure of a previous calibration step. Prerequisites for the execution of a subroutine or module are checked by loading the parameter file and checking the data values attached to the called bibliography entry. If prerequisites are met, the module is either executed or skipped. This logic is very important for a stable and smooth execution of the pipeline.

The similarity of the reduction strategy and the format of Apertif data to the old WSRT data allowed us to use common astronomical software packages in our pipeline instead of developing our own radio astronomical calibration software. This development philosophy offered the advantage of using reliable and fully tested calibration routines while minimising the workload for software development. Therefore, *Apercal* uses third-party astronomical software, most notably AOFlagger (Offringa et al., 2012b), the Common Astronomical Software Application (CASA) (McMullin et al., 2007), Multichannel Image Reconstruction Image Analysis and Display (Miriad) (Sault et al., 1995) and the Python Blob Detector and Source Finder (PyBDSF) (Mohan and Rafferty, 2015).

Installation of python based modules was performed using the pip-repositories[17]. For CASA the available binary installation from the NRAO webpage was used[18]. PyBDSF was installed via the KERN-repository[19]. MIRIAD was compiled with some minor changes (see Appendix C) using the native ATNF version available on their website[20].

Data processing of Apertif data currently takes place on a computing cluster with five individual nodes with identical specifications. Each node is equipped with 24 physical cores

---

[16]https://jupyter-notebook.readthedocs.io/en/stable/notebook.html

[17]https://pypi.org/project/pip/

[18]https://casa.nrao.edu/

[19]https://kernsuite.info/

[20]https://www.atnf.csiro.au/computing/software/miriad/

with hyperthreading for a total of 48 CPUs, 256 GB memory and 33 TB storage space. Each new observation is first sent to the Apertif Long Term Archive (ALTA) from where it is automatically retrieved and copied to the cluster. The visibilities of each beam of an observation are stored as an individual dataset. This strategy was found to be optimal in terms of storage and processing since individual beams can be and are treated as independent observations. The 40 beams of an observation, with a data size of 115 GB each, are therefore distributed evenly over four of the nodes leaving the fifth node available for testing, development and further data analysis. Hyperthreading was activated for routines which supported the feature and proved to accelerate the processing.

Apercal currently produces data products such as calibration solution tables, fully calibrated target datasets and images in all Stokes parameters as well as line cubes for AWES. Additional processing, e.g. the co-addition of different observing epochs, outside of *Apercal*, is necessary for AMES. Science-ready data products will become public[21] after a limited propriety time via ALTA[22]. *Apercal* is publicly available on GitHub[23].

4. Pipeline modules

In the following section we describe the functionality, the input and output, and the different algorithms used for processing data for each individual pipeline module. An overview of the whole reduction pipeline is given in Fig. 1. The whole processing can be divided into four different steps: The triggering of the pipeline involving downloading and preparation of the data, the flagging and cross-calibration, the self-calibration and imaging, and as a last step, the validation and ingest into the Apertif archive. The progress of the pipeline in between and within modules is saved in a stats-file. An overview of the runtime for each individual module is shown in Fig. 2.

4.1. Autocal

Autocal is an automatic routine which executes the start of the reduction pipeline as soon as new observations have been found on ALTA. It automatically identifies the flux and polarisation calibrator data products for a given target field dataset. The necessary information for an *Apercal* pipeline run are then collected and written to a configuration file. After that *Apercal* is triggered and starts downloading the relevant data from ALTA.

Notifications for observers on duty are provided at the start of a pipeline run, in case of failed reduction steps and for successful runs. In case of failed processing steps, the observers are inspecting the data and processing logs for possible reasons. Once the reason has been found and corrected, the failed reduction steps are restarted.

After *Apercal* finishes a reduction *Autocal* automatically triggers the quality assessment (QA) pipeline and ingests the processed data products back into ALTA, with notifications for each stage. Coordinating activities in this way via *Autocal* enables us to moderate the usage of the cluster.

4.2. Prepare

Apercal defines a directory structure for processing where each module uses its own subdirectory to access data and save outputs. All of the following modules (except *convert*) use a single subdirectory, so that individual steps can easily be deleted and restarted. The naming of the directories can be adjusted to the needs of the users with keywords in the configuration file.

The main tasks of the *prepare* module are the setup of the directory structure and the download of the data. Once the module is executed given an input target and calibrator datasets, it checks the availability of the data on the local disc. In case data is not locally available, the module checks the availability on ALTA via an iRODS framework[24]. If successful, a Python routine is used[25] to download the data to the local disc and to place it in the appropriate position of the directory structure. After a dataset has been successfully copied from ALTA or located on disc, the correctness of the file is checked via a checksum.

A minimum of a target dataset and a flux calibrator needs to be present per beam for this step to be successful. This condition ensures that, if no flux calibrator is available, the execution of the pipeline is stopped. On the other hand, the pipeline will continue when a polarised calibrator is not available. In this case, the polarisation calibration within the *crosscal* module is omitted and no polarisation imaging is performed.

4.3. Preflag

The *preflag* module handles all pre-calibration flagging of the data. It is separated into three different operations: The flagging of data with issues known *a priori*, additional flagging options, and fully automatic flagging to identify and mitigate spurious radio frequency interference (RFI). The first two operations use the DRIVE-CASA Python wrapper (Staley and Anderson, 2015) to parse commands to CASA while the last one uses the AOFlagger routines (Offringa et al., 2012b,a, 2010).

The subroutines for *a priori* known issues cover three distinct operations: First the data are checked for shadowing. This occurs when the illumination of a dish is hampered by other close-by dishes if a source is observed at a low declination. Timestamps of shadowing can be calculated using a given dish diameter, the position of the dishes and the position of the observed source. Data points positively identified to suffer from shadowing are then automatically flagged. The next step is a mitigation of the effect of the steep bandpass edges of the individual subbands of the Apertif system. The first two and last channels of each 64 channel subband are flagged. Thirdly, due to a finite impulse response (FIR) filter in the channel filterbank,
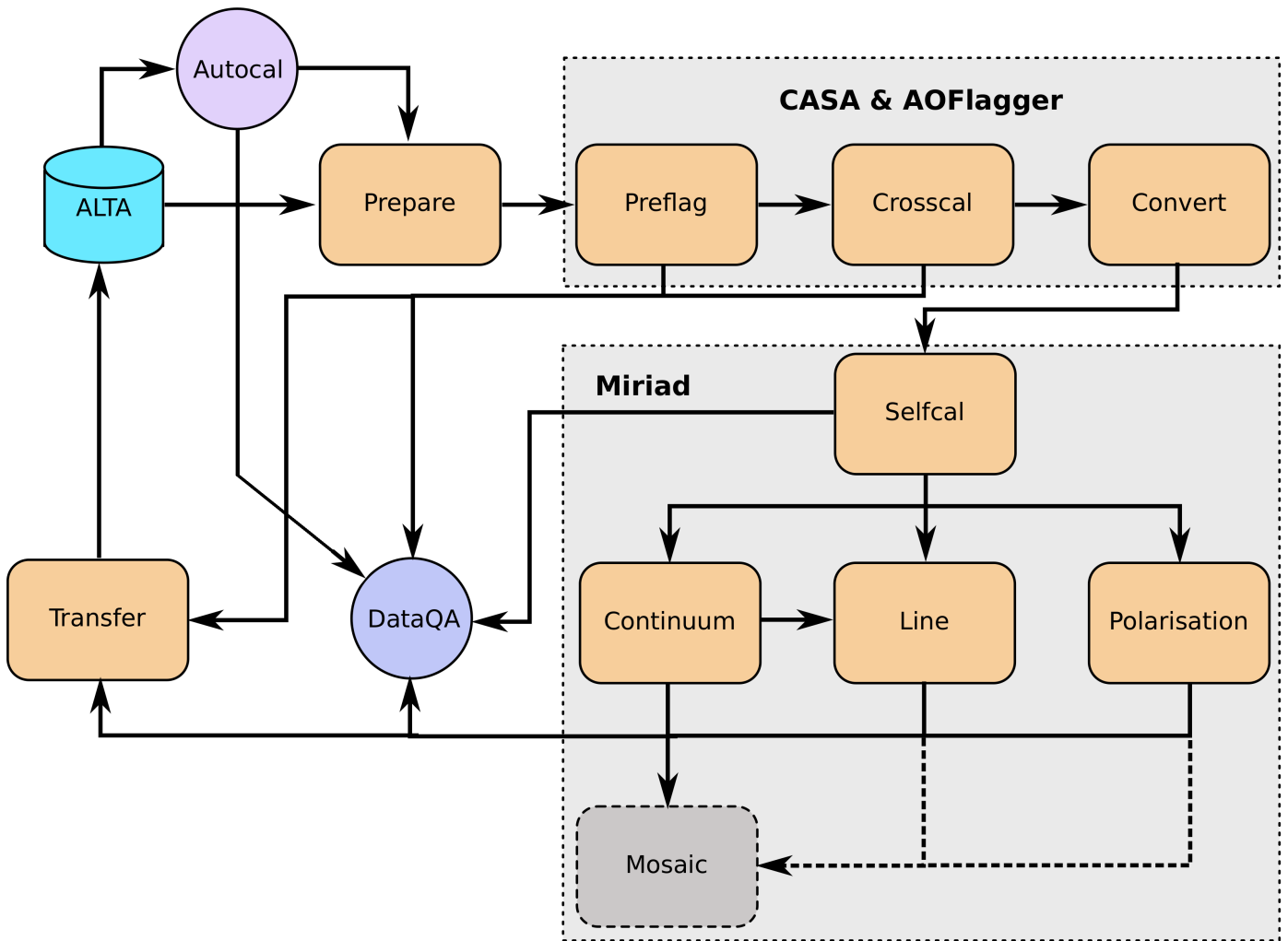

Figure 1: Overview of the structure of Apercal. Each rectangular box represents a single module. The grey boxes encapsulate the astronomical software packages used within the individual modules. Arrows illustrate the data and workflow within the pipeline. The dashed arrows are routines which are currently developed. The individual modules are described in detail in the following sections: *Autocal*: Sect. 4.1, *Prepare*: Sect. 4.2, *Preflag*: Sect. 4.3, *Crosscal*: Sect. 4.4, *Convert*: Sect. 4.5, *Selfcal*: Sect. 4.6, *Continuum*: Sect. 4.7, *Line*: Sect. 4.8, *Polarisation*: Sect. 4.9, *Transfer*: Sect. 4.11, *Mosaic*: Sect. 4.10 and Sect. 5, *DataQA*: Sect. 7
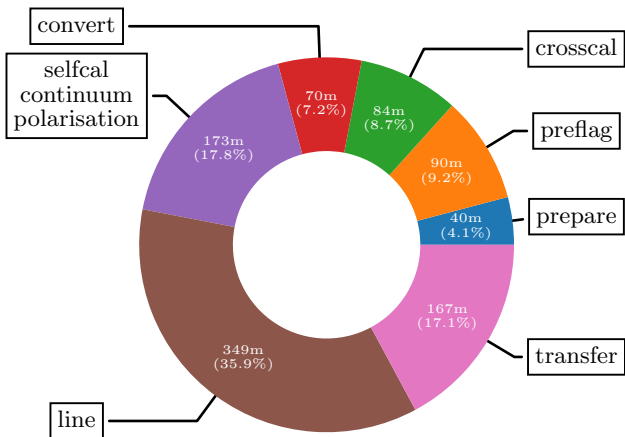
Figure 2: Pie chart showing the runtimes of the individual *Apercal* modules for an example reduction of one beam. The complete runtime was 972 minutes.

artificial signal is generated in channels 16 and 48 of each subband (see van Cappellen et al. (2021)). This causes an artificial source on the order of several µJy in the pointing centre of each image. To mitigate this effect, an option exists to flag these channels automatically.

Subroutines for the additional flagging options encompass the removal of auto correlations, entire antennas, specific cross-correlations, individual baselines, channel and/or time ranges. Any flagging commands not covered by the standard commands can be parsed to a file using the standard CASA-syntax. These options are supposed to be used either when known elements within the Apertif system are not working or a user identifies additional issues during calibration of the data. For the latter case a complete recalibration of the data with the specified additional flags is usually performed. The data ranges to flag for the above mentioned subroutines are specified in the configuration file. More flagging criteria, based on the calibration solutions, are evaluated in the *crosscal* step (see Sect. 4.4).

The last step of the module uses the AOFlagger routines to automatically identify and flag any unknown and previously not flagged RFI in the calibrator and target datasets. Automated detection of interference is challenging because of the large variety in observation properties: calibrators for the surveys are only observed for 3-5 minutes and show high signal-to-noise ratios, whereas the target fields are observed for 11.5 hours that normally have low signal-to-noise ratios. In the first case, low-level RFI that potentially lasts longer than the calibrator run is hard to identify. In the second case, more sensitive detection is possible, but care has to be taken not to overflag the data. This is especially the case for observations with strong HI-line emission, which could easily be false identified as RFI. We designed a flagging strategy which suits both types of observations and avoids flagging strong line emission.

The designed AOFlagger strategy for Apertif makes use of a low-pass filter in combination with the Sumthreshold and Scale Invariant Rank operations to detect interference (Offringa et al., 2010, 2012b). Previously set flags written to the MS and are

used by AOFlagger for flagging the data further. An example of our strategy applied to Apertif data is shown in Fig. 3. Details and further results of the AOFlagger strategy will be discussed in Offringa et al. (in prep.).

4.4. Crosscal

Cross-calibration is performed independently for each calibrator-target-pair of observations. The cross-calibration step solves for the bandpass, gain, delay and polarisation leakage solutions of the flux calibrator. Because the flux calibrator is unpolarised, the cross-hand delay and polarisation angle solutions are derived from the polarised calibrator. All *crosscal* calibration steps use the standard CASA routines in release version 4.7.0-1-el7. In case a polarisation calibrator has not been successfully observed or its dataset has not passed the *preflag* module, polarisation leakage, polarisation angle and cross-hand delay solutions are not determined. This criterion is evaluated on a per-beam basis. In such a case, later stages of the pipeline omit polarisation imaging, but still perform self-calibration, line and continuum imaging to maximize the amount of the generated data products. Bandpass, polarisation leakage and polarisation angle solutions are derived on a per-channel basis to mitigate any effects within the observed bandwidth. For the unpolarised calibrators the flux density scale from Perley and Butler (2017) is used while for the additionally needed information for the polarised calibrators, such as the polarisation angle, degree of polarisation and rotation measure, Perley and Butler (2013) is used. We want to note, that for linear feeds CASA is not properly solving for polarisation angle corrections. This results in an exchange of the Stokes Q- and U-values for the two polarised calibrator sources used. This behaviour was found to be constant in time and over frequency, so that a correction can be performed after imaging the polarisation data products. Therefore, it is carried out during the polarisation mosaic step (see 5.2).

Calibrator data are automatically checked for problematic dish-beam combinations. Problems here arise from individual receiver elements in the PAF, which are malfunctioning due to broken connectors, cables, or electronics. Beams are formed in a certain direction by weighting the signal of the different elements accordingly (van Cappellen et al., 2021). Beams where the weights for malfunctioning elements are high, result, if used for calibration and imaging, in higher noise levels in the images or even a complete divergence of the calibration solutions. These problematic beams are spotted most easily in the auto-correlation data. The amplitude and spectral behaviour of the Apertif auto-correlations is known well, so that differences from the expected values were used to develop a metric for an automatic detection algorithm.

The currently implemented metric checks on a per beam and dish basis for the flux calibrator after a first cross-calibration, that not more than 50% of the auto-correlation data show amplitudes of more than 1500 K (see Fig. 4). This value was derived from our experiences, where significant artefacts in the images become apparent. In addition, the bandpass phase solutions of the flux calibrator are investigated after calibration.
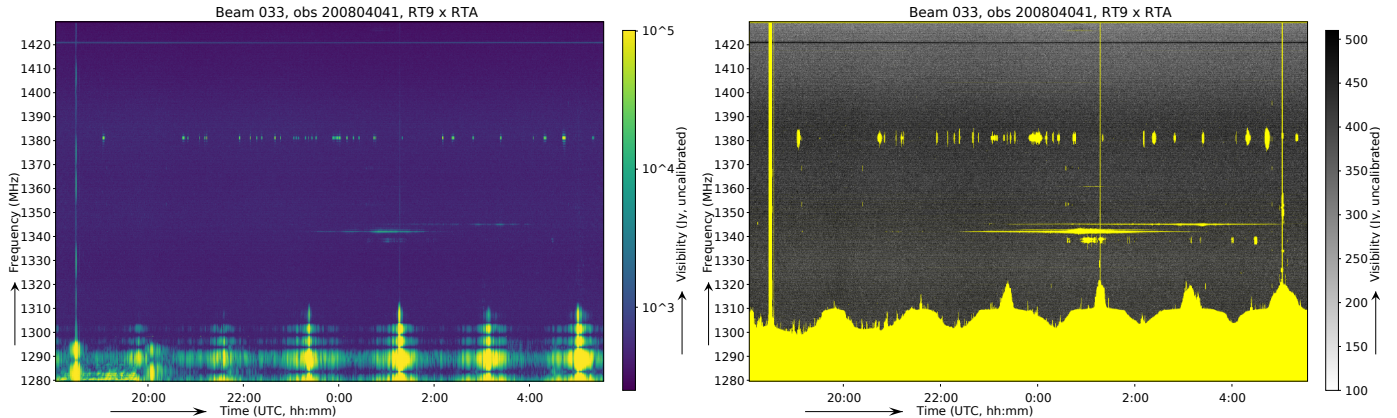
Figure 3: Time vs frequency plots of a typical Apertif 11.5 hr integration. The scale is representing the amplitude of the visibilities. The left hand side shows the original data without any flagging. Automatic flags determined by AOFlagger are shown on the right hand side for the same data coloured in yellow. The HI-line at 1421 MHz is successfully omitted during the flagging.

A standard deviation of more than 15 deg is marking the calibration as failed. If one of the above mentioned criteria applies to a dish-beam combination, the specified data is marked and flagged automatically. The flags are then extended to the rest of the datasets of the observation, namely the target and polarisation calibrator data. This whole process is performed in an iterative way, where the whole cross-calibration is redone and results are checked by the metric described above. A maximum of four crosscal iterations are allowed after which the *crosscal* module gives a final result. The pipeline is stopped for beams not passing this stage so that further processing on that particular beam ceases, but other beams passing the stage continue to be processed through the pipeline. Beams passing these checks are used to calibrate the target fields by applying their calibrator solutions to the target field data. Since the Apertif system does not possess a diode injected gain correcting system as the old WSRT MFFE backend did, the stability of the crosscal solutions was checked by applying different calibrator solutions several days apart from each other to target datasets. None of the final datasets showed systematic offsets or differences of target field source fluxes by more than 10 %. Any further processing of the calibrator datasets stops here and the following modules only focus on the target data.

4.5. Convert

Based on the familiarity of the developers with suitable software packages for the further reduction steps combined with the experience gained from the reduction of archival WSRT data using mostly MIRIAD, we chose to stay with this reduction package. We used the native ATCA MIRIAD version with slight modifications to be able to read and calibrate the Apertif data (see Appendix C).

Since MIRIAD is not able to access the Measurement Set[26] (MS)-format native to CASA, we need to convert the file format before doing any further reduction. Unfortunately a task for a direct conversion from MS to MIRIAD format is not available,

so that we have to first convert to the UV-FITS standard (Cotton, 1990) and from there to MIRIAD. For this purpose we use the CASA task EXPORTUVFITS followed by the MIRIAD routine FITS.

4.6. Selfcal

Self-calibration is a standard procedure in radio interferometric data reduction to enhance the dynamic range of images. Small changes in the processing of the signals in the receiver electronics (e.g. temperature changes) and ionospheric and tropospheric variations of the Earth's atmosphere cause changes of the received amplitudes and phases over the time of the target observations. These usually slowly changing variations cannot be compensated by the bracketing calibrator observations and therefore need self-calibration. This technique, developed more than 40 years ago and nicely summarised in Pearson and Readhead (1984), involves the calibration of the data by using a model generated by imaging the data itself.

The task of the *selfcal* module is to solve for the antenna and feed time-based variations of the target data within a self-regulating algorithm using the self-calibration technique. To guarantee the stability of the self-calibration process and the processing within a reasonable time frame, several preliminary steps are executed within the *selfcal* module before the actual self-calibration starts.

First, the target data is averaged down in frequency by a factor of 64 over the 64 channels of each subband resulting in a frequency resolution of 0.78 MHz and therefore a bandwidth smearing of a < 1%-level. We also do not expect any strong amplitude or phase variations within this frequency span due to ionospheric variations or from the receiver hardware. It is important to note that this frequency averaged dataset is only used for continuum and polarisation imaging in later stages of the pipeline and any HI-line imaging is performed on a dataset with the original resolution where the derived self-calibration gains are interpolated in frequency and applied.

In order to mitigate any influence of strong HI-line emission or residual RFI on the self-calibration solutions we generate an

---

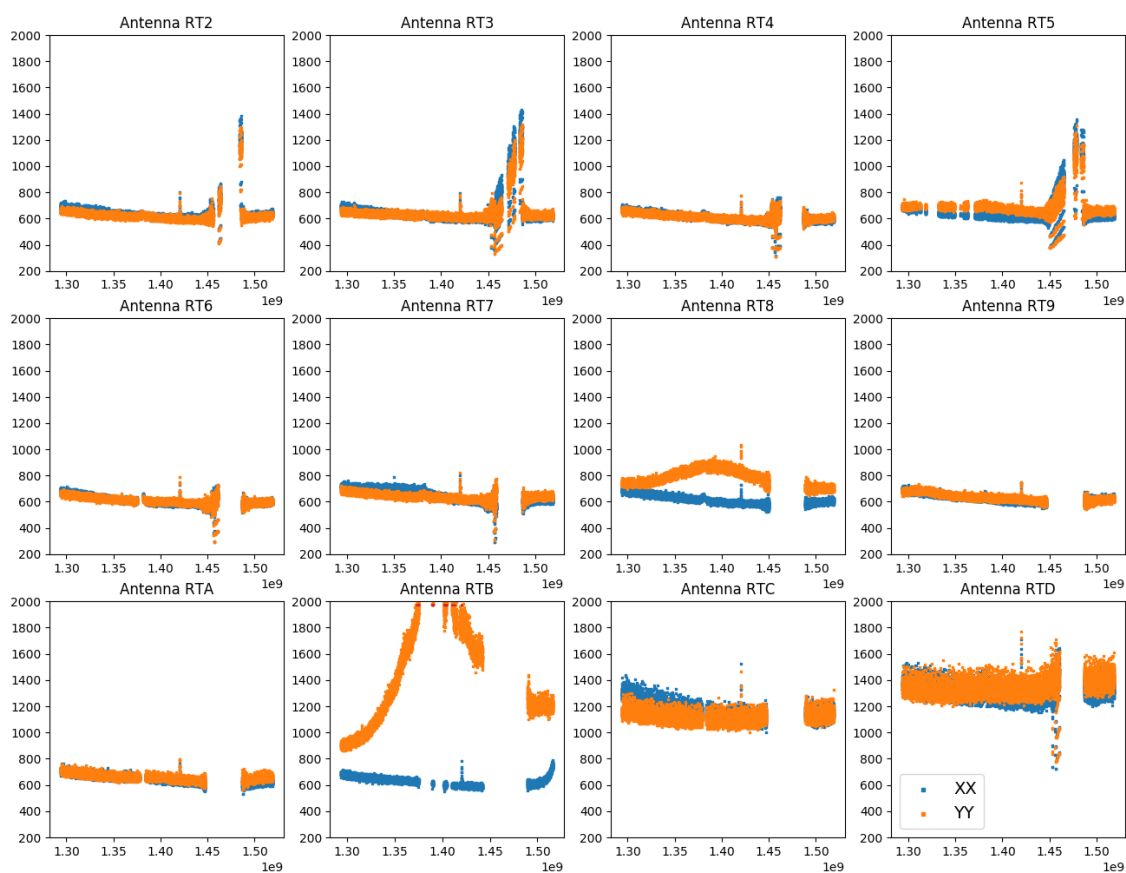[26]https://casa.nrao.edu/Memos/229.html

Figure 4: Example plot for the quality assurance pipeline. Amplitudes of the autocorrelation visibilities are shown vs frequency for Beam 06 of one observation. Each subplot shows the XX (orange) and YY (blue) autocorrelations for one antenna. In this example the YY-autocorrelation values for antenna RTB show an atypical slope and curvature and more than 50 % of the values are larger than 1500 K, so that all of its visibilities were flagged for this beam. RTC and RTD show constantly higher values, but both antennas are below 1500 K, so that we were still able to use the data from these antennas.

image cube out of the averaged data. For each of the 64 images in the cube its standard deviation is measured. We then use an outlier detection algorithm to locate the channels affected by either above mentioned reasons and flag them in the averaged dataset. As above, these flags are only used for the continuum and polarisation imaging later in the pipeline and not for HI-line imaging.

The performance of the self-calibration is often strongly dependent on the first image passed to the solver. In order to start the self-calibration with an image of an optimal initial quality, we use the information provided by radio continuum surveys at the same wavelength. For this purpose we first query the catalogue of the Faint Images of the Radio Sky at Twenty-Centimeters (FIRST) Survey (Becker et al., 1995). Since this survey does not cover the whole Apertif survey footprints information for fields outside of the FIRST footprint are collected from the NRAO Very Large Array Sky Survey (NVSS) (Condon et al., 1998). We observe with a fractional bandwidth of ≈ 20%, so that we need to account for the spectral index and primary beam variations over frequency. For acquiring a spectral index for the sources in our skymodel we query the Westerbork Northern Sky Survey (WENSS) catalogue (Rengelink et al., 1997) and cross-match. Since WENSS has inferior resolution compared to the other two surveys, we account for multiple source matches by summing the fluxes of the individual components to derive the spectral index and assign the same value to all of them. We then account for the primary beam response of the Apertif system by using the primary beam model of the WSRT before the Apertif upgrade as an approximation (Popping and Braun, 2008). The final skymodel is then generated by directly fourier transforming the catalogue source fluxes and positions into the (u,v)-domain with the MIRIAD task UVMODEL. We then derive phase-only calibration solutions, which ensures that all our images are aligned to the same common reference frame given by the above mentioned surveys. In addition the resolution of this parametric skymodel is not limited by the pixel raster of the images, but rather the fitted position of the sources. The solution interval for this parametric calibration is usually on the order of several minutes, which is set in the configuration file. We performed several tests on this method by comparing sources detected in finally self-calibrated images of Apertif observations with archival values of the NVSS- and FIRST-catalogues and archival WSRT-observations. Differences were always on the order of a only few percent with no obvious systematic effects. A quantitative analysis of source position and flux accuracy for the Apertif surveys will be published in Denes et al. in prep. and Kutkin et al. in prep.

The next step involves the actual self-calibration iterations. To improve the stability of the pipeline a multi-frequency image of circular polarisation (Stokes V) is generated at the beginning of each iteration. Since the circular polarised sky is essentially empty any sources in such an image would hint to severe calibration problems. Therefore, the image statistics can be analysed for following a normal distribution resembling Gaussian noise. This was implemented using the method developed by D'Agostino (1971) and D'Agostino and Pearson (1973), which combines information of the skewness and kurtosis. If these

values exceed a certain number given in the configuration file, the self-calibration is aborted.

After passing this check each iteration consists of inverting the (u,v)-data using the MIRIAD task INVERT followed by an automatic masking routine involving the source finder PyBDSF (Mohan and Rafferty, 2015) to limit the CLEAN algorithm (Högbom, 1974) to islands of emission from astronomical sources omitting any image artefacts. Within PyBDSF we use an adaptive rms box without a pre-generated rms map. The rms value up to which PyBDSF is limiting the source detection is adaptive and is described below. All other parameters are set to their default values.

We are aware that for linear feeds the Stokes Q emission can have a non-negligible effect on the calibration. Denes et al. in prep. showed that the contribution (leakage) for nearly all beams is only on the order of 1 % up to a primary beam response level of ∼ 30 %. Therefore, we perform all self-calibration and imaging is performed on the total intensity Stokes I data only. Image deconvolution is executed using the multi-frequency CLEAN-algorithm (Sault and Wieringa, 1994) implemented into the MIRIAD task MFCLEAN, which uses a first order polynomial to derive the spectral index of the sources within the imaged bandwidth. After cleaning, restored and residual images are created using the task RESTOR. The CLEAN model generated during the cleaning process is then used to derive new calibration solutions.

The CLEAN algorithm performs best for images which consist of point-sources and do not show calibration artefacts. In order to omit these artefacts and correctly generate masks for extended sources we use an iterative approach, where we set adaptive thresholds and constantly perform quality assurance incorporating different metrics as described in the following.

Once one of the adaptive threshold criteria is reached the cleaning process is stopped. For each full cleaning iteration within a self-calibration cycle three different thresholds for generating masks are calculated: the theoretical noise threshold T_{tn}, the noise threshold T_n and the dynamic range threshold T_{dr}.

The theoretical noise T in Jy/beam is determined by calculating the standard deviation from images generated in circular polarisation (Stokes V). Astronomical circular polarised sources are very rare on the sky and if present only at very low flux levels. Residual RFI on the other hand is often circular polarised and raises the noise levels of these images. The noise statistics of these images therefore represent the actual quality of the data and the theoretically reachable noise of the final images well. T_{tn} is given in units of Jy/beam and defined as

$$T_{tn} = T \cdot n_\sigma \tag{1}$$

where n_σ is the confidence interval for regarding islands of emission as real. This is usually set to $n_\sigma = 5$.

In order to guarantee a smooth convergence of the self-calibration skymodel the two additional thresholds T_{dr} and T_n set limits for the maximum dynamic range achievable in an image without reconstruction and the adaptation to image artefacts, respectively. The dynamic range threshold within a cycle is defined by the number of the current major cycle m, the initial

dynamic range DR_i and a factor defining how fast the threshold should increase DR_0 such as

$$T_{dr} = \frac{I_{max}}{DR_i \cdot DR_0^m} \qquad (2)$$

where I_{max} is the maximum pixel value in the residual image of the previous cycle. The parameter DR_i is dependent on the level of the first major sidelobe in the dirty beam. The ratio between the maximum and this value gives the maximum dynamic range by which an image can be cleaned before another cycle of image reconstruction needs to be performed.

The adaption of the noise threshold T_n for stopping each individual run of the CLEAN algorithm is given by

$$T_n = \frac{I_{max}}{((c_0 + n \cdot c_0)(m + 1))} \qquad (3)$$

where m and n are the number of major and minor cycles, respectively, and c_0 handles how aggressive the cycles are performed. For each individual run of MFCLEAN all three thresholds are calculated and the maximum is set as a limit for the generation of masks in PyBDSF for the current major cycle. The maximum number of major cycles is set in the configuration file. This is useful for cases where the theoretical noise threshold is approached very slowly or never reached due to remaining image artefacts. Then cleaning is performed within this mask down to a level of the derived masking level divided by the parameter c_1, which is usually set to $c_1 = 5$.

The length of the solution interval s for each self-calibration cycle is reduced for an increasing number of major cycles m. This is possible due to the improved and more complex CLEAN component model used and the better starting values for fitting solutions derived from earlier selfcal iterations. In addition a higher number of data points due to more baselines, feeds and solution intervals over frequency guarantee a better convergence of the self-calibration. Therefore, the solution interval is determined using the following equation:

$$s = \frac{t/n_s}{m} \qquad (4)$$

where t is the total observation time and

$$n_s = \sum_l \frac{I(l)}{SNR \cdot T \sqrt{n_B n_{nf} n_f}} \qquad (5)$$

where l is the number of clean components, I the flux of each individual clean component, SNR the needed signal-to-noise ratio, T the theoretical noise, n_B the number of baselines, n_{nf} the number of solution intervals over frequency and n_f the number of polarisations to solve for. Due to the larger number of degrees of freedom for amplitude and phase calibration in contrast to phase-only calibration the SNR is set to 3 for phase-only calibration and to 10 for combined amplitude and phase calibration. The above formalism ensures that solution intervals decrease during the self-calibration process consecutively while still containing enough signal-to-noise for a proper calibration.

The *selfcal* module first performs up to a given maximum number of iterations (typically five) of phase-only self-calibration. By summing the flux of the clean components the module calculates the total flux in the dataset and can estimate the available SNR. This allows to perform an additional self-calibration step for a combined amplitude and phase solution for datasets with high SNR improving their final image quality. The image quality of observations with low SNR would suffer from this step and are therefore excluded. If at any point during the process the theoretical noise limit is reached, *selfcal* performs only one last iteration of self-calibration.

During and after each imaging and cleaning cycle the dirty image, the cleaning mask, the clean component model image and the restored images are checked for any obvious problems resulting from a divergence of the calibration routines using the following criteria and metrics. The maximum value in dirty images of total intensity should always exceed the minimum. In addition images are inspected for Not a Number (NaN)-values. Both conditions are checked when starting a cleaning where new solutions were derived and applied to the data and a new dirty image is generated. Clean component images are checked for containing zero CLEAN components. In addition, the clean component images are checked for clean components with unrealistically strong negative or positive fluxes. The restored image is again checked for containing no NaN-values and for unrealistically strong positive or negative values on the order of 10^4Jy. These NaN-values are generated in the rare cases where PyBDSF, during the masking procedure, does not find islands of emission and CLEAN of MFCLEAN are executed with an empty mask. This can often happen in cases where the data is suffering from system failures or diverging calibration solutions. The final residual image should mostly consist of noise and is therefore checked for gaussianity using the aforementioned algorithms .

A combined amplitude and phase calibration (A&P) is not as stable as a phase-only calibration due to the increased degrees of freedom, so that this step can easily worsen the image quality due to diverging calibration solutions. Therefore, in addition to the quality assurance process described above, we added an additional metric to check the quality of the A&P calibration compared to the phase-only one. After the A&P calibration we generate another dirty image and compare the image statistics independently, namely its maximum, minimum and standard deviation with the dirty image of the last phase-only self-calibration cycle. If the ratio of one of those values exceeds a limit given in the config-file, which is usually based on a worsening of the image quality after a combined A&P-calibration step, the A&P calibration is assessed as failed and any following module will use the last successful phase-only self-calibration solutions. Calibration solutions are applied in the subsequent modules before any further imaging is performed. An example of a final total power self-calibrated image in comparison to one without self-calibration is shown in Fig. 5.

4.7. Continuum

The *continuum* module performs two different tasks to generate final deep continuum Stokes I images from the self-calibrated (u,v)-data. First it generates a deep multi-frequency image using the MFCLEAN task in MIRIAD and secondly several
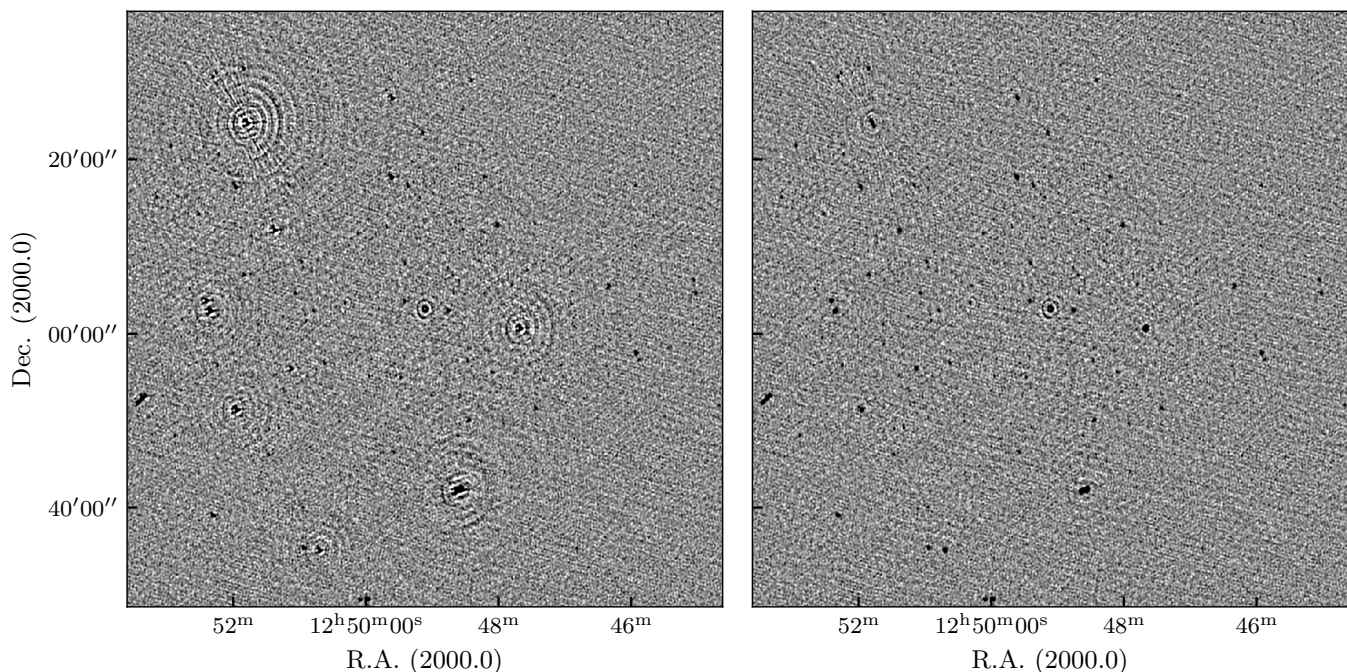
Figure 5: Example of a continuum image before phase self-calibration (left) and after (right). Artefacts around stronger sources are significantly reduced and all sources appear more focused.

individual images spanning narrower frequency ranges of usually ≈20 MHz over the full bandwidth using the task CLEAN are produced. The purpose of this is to generate an as-deep-as-possible total intensity image with the maximum possible resolution (given by the highest frequency) using the multi-frequency imaging approach and in addition derive reliable spectral indices and curvatures for as many sources as possible using the individual frequency images. In fact multi-frequency cleaning generates these images already, but their values are only reliable in cases of high signal-to-noise ratios. For both imaging steps we use a natural robust weighting (Cornwell and Wilkinson, 1981) to acquire the maximum possible resolution. Images generated with the default parameters have a size of 3073x3073 pixels with a sampling of 4" per pixel, which allows the imaging and cleaning of any sources up to the first sidelobe level of the primary beam response in order to minimise artefacts.

Cleaning and masking iterations are in both cases continued until the theoretical noise limit has been reached. Masking and validation of all continuum images is performed in the same way as described in Sect. 4.6.

4.8. Line

The *line* module first applies the derived self-calibration solutions to the non-averaged data. This is performed using the MIRIAD task GPCOPY. It automatically takes care of the different frequency resolution of the two datasets by interpolation.

The HI-line imaging is the most computing intensive task in the Apertif data reduction. The following strategy for handling the line data and the image cubes has been found to be most time efficient: imaging is performed by generating eight

individual cubes over the 300 MHz of bandwidth with a small amount of overlap in frequency. The overlap is necessary to avoid splitting the detected line emission of individual objects between two adjacent image cubes. In order to improve sensitivity and save processing time and disc space, data between 1130 MHz and 1416 MHz are averaged in frequency by binning three channels together resulting in a frequency resolution of 36.6 kHz corresponding to a velocity resolution of 7 km/s, which is usually sufficient for any extragalactic objects at these frequencies. The data at the highest frequencies which features the Galactic neutral hydrogen and small galaxies in the nearby Universe retain the full spectral resolution of 12.2 kHz or 2.3 km/s. All line data is in topocentric velocities, which is the native reference frame of the data. Conversions to other reference frames are performed during further analysis steps.

We want to note that the current flagging and imaging strategy is not optimised for calibrating and imaging the Galactic range of the data. This would need a joint deconvolution of several beams at the same time plus a combination with single dish data, for example from the HI4PI-survey (HI4PI Collaboration et al., 2016), to mitigate the effect of missing short spacings.

In order to generate image cubes containing only HI-line emission the continuum has to be subtracted. Several different approaches are possible here: the fitting of baselines to the amplitude of the data followed by subtraction, the subtraction of constant fluxes over frequency in the image domain and the direct subtraction of a continuum clean component model from the (u,v)-data. The best performance in terms of time consumption and fidelity was achieved with the last method, so that we decided to use this in *Apercal*. For this subtraction the final

11

clean component model of the *continuum* module is used.

The final line images are produced using a parallelisation approach where each individual frequency channel is imaged as an individual imaging task. For this most computing intensive part of the pipeline, the Python Pymp[27] utilisation of the OpenMP[28] framework was used. All resulting images are then combined into single line image cubes.

4.9. Polarisation

Polarisation imaging is performed in Stokes Q, U and V. Q and U fluxes from astronomical sources exhibit a sinusoidal dependence of the square of the observed wavelength, which is caused by Faraday rotation. In addition, Stokes Q, U and V fluxes can have negative values in contrast to Stokes I, which needs to be positive in all cases. These effects would lead to bandwidth depolarisation in case of multi-frequency imaging of Stokes Q and U over our full 300 MHz bandwidth (Burn, 1966). Therefore, we image Stokes Q and U as cubes, using the method described in Sect. 4.8, where one image is generated for a bandwidth of 6.25 MHz, which results into 48 images (subbands) for each Stokes parameter. This mitigates the effect of bandwidth depolarisation for most astronomical sources. In addition, this method allows the usage of the rotation measure synthesis technique (Brentjens and de Bruyn, 2005) in post-processing, so that the linear polarisation properties of the detected sources can be analysed in more detail than with the standard methods, which suffers from bandwidth depolarisation effects.

Stokes V is representing the circular polarisation, which does not show any sinusoidal behaviour. Due to this fact and since the circular polarised sky is very faint, we perform a multi-frequency synthesis for imaging Stokes V. This also allows to maximise the sensitivity of the produced images to detect possible circular polarised sources.

For all polarisation images cleaning is performed using the final mask generated by the multi-frequency imaging part of the *continuum* module. Since polarisation images usually only show very faint emission, the clean threshold is set by the standard deviation of the pixels in the image. This accounts, especially for the Stokes Q and U images, for the variations of the noise over the imaged bandwidth. Since we are only imaging a limited bandwidth at a time for Stokes Q and U we use the faster MIRIAD task CLEAN while for the imaging of the full bandwidth for Stokes V the task MFCLEAN is used. We want to note that due to the separate cleaning of Stokes Q- and U-images biases regarding the polarisation angles could be introduced into the final maps as has been shown by Pratley and Johnston-Hollitt (2016).

4.10. Mosaic

Once all data of an observation have been processed through the *continuum*, *line*, and *polarisation* modules, a *mosaic* module can be executed to generate a combined image of all beams of one observation. This step is currently not performed as part of the automatic pipeline.

Images are regridded to a common grid centred on the central beam of the observation and then corrected for the Apertif compound beam response. The compound beam response has been characterised using drift scans of a strong astronomical source over the whole field of view of the Apertif Phased Array Feed (for details see Denes et al. in prep.). The different beams of one observation have slightly varying synthesised beam sizes due to different flagging of the data and slightly different declination positions, so that all input images are convolved to the a given common beam.

The combination of the input images then follows an inverse square weighting based on the compound beam response and the background noise of the individual images. The background noise is estimated using the MIRIAD task SIGEST, which minimises the contribution of sources for the determination of the noise level.

Since all data of one observation has been taken using the same electronics, a correlation between the noise of different beams exists, which raises the noise level of the final mosaic. An option for including this correlation matrix during mosacking is implemented and will be used once the coefficients have been measured. First tests showed a minor change of the noise levels between correlated and uncorrelated data of adjacent beams of ∼ 2 %.

The *mosaic* module is currently only producing continuum mosaics using the central frequency of the observational setup for correcting the primary beam response. Additional features in the future include the implementation of the frequency and long-term time dependence of the beam pattern. The current implementation takes approximately an hour for generating a continuum mosaic. Future improvements will include enhancements to the speed of the module, which will then allow us to generate polarisation and line mosaics within an acceptable amount of time.

4.11. Transfer

This module converts the self-calibrated MIRIAD data to UVFITS format. Similar to *line* it applies the phase and, if available, amplitude self-calibration solutions to the non-averaged cross-calibrated data automatically before conversion. The data is then ready for archive ingest.

5. Additional mosaicking software

To allow users a convenient workflow with the provided Apertif data products, especially continuum images and polarisation and line cubes, we developed the *Apercal* independent mosaicking package *amosaic*[29]. This package is purely based on up-to-date Python3 routines and allows command line or Jupyter notebook execution. These specifications ensure that the astronomical community is able to work with the provided data products of the current and future Apertif data releases

---

[27] https://github.com/classner/pymp
[28] https://www.openmp.org/

[29] https://github.com/apertif/amosaic

without the need for installing or maintaining any radio astronomical calibration packages.

The current version of the software is able to mosaic Stokes I continuum images and Stokes Q and U cubes of single observations. In the following we want to describe the capabilities of the software and give an outlook on future improvements.

5.1. Continuum mosaicking

Continuum images for the mosaicking operation are accepted or rejected based on their noise, synthesised beam parameters and different quality metrics (see Adams et al. in prep for a detailed explanation). Accepted images are then primary beam corrected using externally supplied primary beam models. A description of the two different methods to measure the Apertif primary beam are given in Denes et al. in prep. and Kutkin et al. in prep.. Primary beam response images are adjusted to the frequency and size of the target image and then multiplied.

We determine the smallest common beam of all accepted images using the Python package *Radio Beam*[30] with small adjustments to the released code for better handling of small convolution kernels. The Python package *reproject*[31] is used to convolve all images to the calculated common beam.

For the combination of the images we use the *reproject_interp* algorithm inside the *reproject* package. Images are inversely square weighted according to their noise characteristics. The final mosaic is then clipped to a given level of the primary beam response and written to disc.

5.2. Polarisation mosaicking

Polarisation mosaicking is performed as described in Sect. 5.1 with adjustments needed for handling image cubes and polarisation specific analysis steps. Our calibration strategy introduces an offset in the polarisation angle calibration (see Sect. 4.4). To correct for this effect the final Stokes Q and U images are exchanged with each other. Then the image cubes are prepared for further processing in view of the RM-Synthesis technique (Brentjens and de Bruyn, 2005). Since this technique is mathematically a Fourier Transform which best works on regularly gridded data, all images in the cubes need to be on the same frequency grid. In addition, a variation of the input synthesised beam along the frequency axis would cause spatial variations in the resulting Faraday cubes leading to additional unpredictable biases for the scientific analysis. We therefore generated a routine to illustrate the final accepted beams and channels for given constrains. Limits for the mosaicing rotuines can be set on noise in individual channel images, their synthesised major and minor beam sizes and the maximum images for a certain beam these constrains are not met for. This routine can easily be used to optimise the input parameters for polarisation mosaics to maximise the output with regard to sensitivity and resolution. The individual image mosaics are generated independently in Stokes Q and U using the same procedure as described in Sect. 5.1. As a final step all image mosaics are combined into Stokes Q- and U-image cubes.

5.3. Future improvements

Additional features for the mosaicking software include the combination of different Apertif pointings. The mosaicking for line cubes is currently in development. Due to the large data volumes, which need to be processed, this also includes the parallelisation of the mosaicking software of different channels in the line image cubes at the same time. In addition, as described in Sect. 4.10, noise between beams of the same observation is correlated. The calculation of these correlation coefficients and their application to the images is another future improvement.

6. Final data products and archive ingest

Upon the successful completion of an *Apercal* run, the pipeline reports back to *Autocal* about which beams (if any) have failed on particular steps. All final data products need to be uploaded to ALTA on a beam-by-beam basis, and connected back to the original raw data via the target IDs. To guide this process, we built *Apergest*, which manages the ingest of processed data products. *Autocal* sends a trigger call to *Apergest* for a given target ID, which results in *Apergest* first referencing a list of expected data products which it then searches for on the cluster. If any data products are missing due to failures in particular steps of the pipeline, it simply continues the upload with what is available.

Once *Apergest* has identified all relevant files, it produces a file in JSON format with metadata information about each file which can be interpreted and translated by ALTA. We also include additional metadata, such as the *Apercal* pipeline version or *Apergest* version. The resultant JSON file is fed to an ALTA instance (which is run inside a Singularity container (Kurtzer et al., 2017)) in order to prepare the ingest. Once successfully prepared, *Apergest* starts the ingest which results in the actual upload of each data product. Finally, *Apergest* then deletes all processed data products which have been successfully uploaded.

7. Data quality assessment

We built the Python package *DataQA*[32] for quality assessment (QA) after the data were processed with *Apercal*. The aim of this approach is not to identify single problematic data products such as single images in a line cube or small outliers during calibration, but rather give an overall assessment of a complete observation and its reduced *Apercal* products. Often a failed run of *Apercal* or low quality imaging products can be traced back to a small amount of bad data and observations do not have to be repeated. The quality assessment of individual beams and the release of those for public usage will be described in Adams et al. in prep.

DataQA was developed specifically for this purpose. It produces diagnostics for each step of *Apercal* to support us in assessing the quality of the processing and the observation. For

---

[30]https://radio-beam.readthedocs.io/en/latest/
[31]https://reproject.readthedocs.io/en/stable/

[32]https://github.com/apertif/dataqa

easier access to all QA products, *DataQA* produces an html-based overview which can be viewed in the web browser. In the following, the QA products for each *DataQA* step are briefly described.

Prepare: DataQA retrieves diagnostic plots from ALTA that are based on the uncalibrated raw data. They are created for the target field and both calibrators for each beam of an Apertif imaging survey observation (see Fig. 6 for an example plot of a calibrator observation). These plots provide information on the impact of RFI, the quality of the tuning of the amplifiers in the system, the pointing accuracy as well as other potential issues with the system itself such as failing electronics during the observation.

Preflag: Apercal produces plots of the RFI flagging performed by *AOFlagger* for three baselines that represent a short, an intermediate and a long baseline of the WSRT (see Fig. 3 for an example). The plots are produced for each beam of the calibrators and the target. *DataQA* collects these plots and combines them for better inspection of the RFI flagging done across the different beams.

Crosscal: DataQA generates plots for all derived cross calibration solutions in the *crosscal* module (see Fig. 4 for an example). This allows an easy identification of malfunctioning frequency bands or dishes and residual RFI influencing the calibrator solutions.

Selfcal: For each self-calibration iteration and each beam the model, residual and restored images are shown in this part of the module. Due to the large amount of images, these are mostly used for closer inspection of certain beams. The most important usage of the images are the final model images. Due to image artefacts and the non-linear nature of the CLEAN-algorithm, sidelobes of sources are sometimes included in the cleaning masks. The CLEAN-components within these masks will then be considered as real sources for the following self-calibration iterations. These ring-like structures are easy to identify by eye. In addition to the individual images, the final phase and amplitude self-calibration solutions are plotted for each beam, time and frequency solution interval.

Continuum: The final clean component model images and restored and residual images are shown for each beam. The final quality of the images can easily be inspected by eye to spot any obvious artefacts in the images. For a quantitative quality assurance we use an *Apercal*-specific version of the ASKAP continuum validation tool[33]. This tool automatically identifies the stronger continuum sources in the final images and compares their position and fluxes to the NVSS database. This therefore serves as a check of the overall flux scale of the final images as well as their astrometry.

Line: A statistical approach is used for the quality assurance of the HI-line cubes due to the large amount of images, which would otherwise need to be inspected. The standard deviation for each individual image in the cube is determined and plotted vs. the frequency. This allows for an easy identification of individual low quality images as well as broad bands of missing data or images affected by broad band RFI.

QA overview and quality report: Currently for each observing week an observational support astronomer (OSA) is assigned the task of inspecting the data quality assessment plots. Each individual *DataQA* step is then ranked on a scale between 0 and 6 based on the number of successfully reduced beams. Additional constraints encompass the loss of the long baselines, if dishes RTC and RTD failed, which enlarges our synthesised beam size by a factor of two. Individual comments can be inserted by the OSA. The complete report is saved to disk, so that all numbers can be accessed at a later point in time.

8. Summary and outlook

Apercal is the dedicated data reduction pipeline for the imaging surveys of the SKA pathfinder telescope Apertif. We managed to create a data reduction framework using an object oriented Python-based approach. Well-established astronomical software packages are used to minimise the workload for the developers and optimise the timing and quality of the results. In addition, Apercal is partly parallelised using Python software packages, which make it easy to adjust to any multi-node hardware system. Currently Apercal is taking 24 hrs for a 11.5 hr target observation to reduce. Since only two thirds of the telescope time is dedicated to imaging observations and calibrator observations, beam weight measurements as well as maintenance produce an additional overhead of ~25 %, it is possible for our setup to keep up with the incoming data amount.

The modularity and independence of individual modules and routines allowed the usage and testing of the pipeline during commissioning of the Apertif system already. The whole pipeline and its prerequisites are publicly available, so that users within the astronomical community are able to improve and modify it to their needs.

In order to ensure the quality of the generated images and the stability of the pipeline two stages of quality assurance are used. The first one being fully automatic and executed during the reduction of the data. Each step in the reduction process including statistics and prerequisites for the following steps is saved in a stats-file. This allows the user to directly see at which stage the data reduction was successful or failed. The second step of quality assurance is executed after each individual pipeline run. Here plots are generated and solutions and images are visualised in a compressed format, so that failing observations or bad quality images are noticed by an observer within an optimal timeframe. Currently 44 % of the observed data processed by Apercal are on a science-ready quality level.

Future improvements to Apercal include an improved masking algorithm, an automatic direction dependent reduction step in full polarisation and direction dependent leakage corrections.

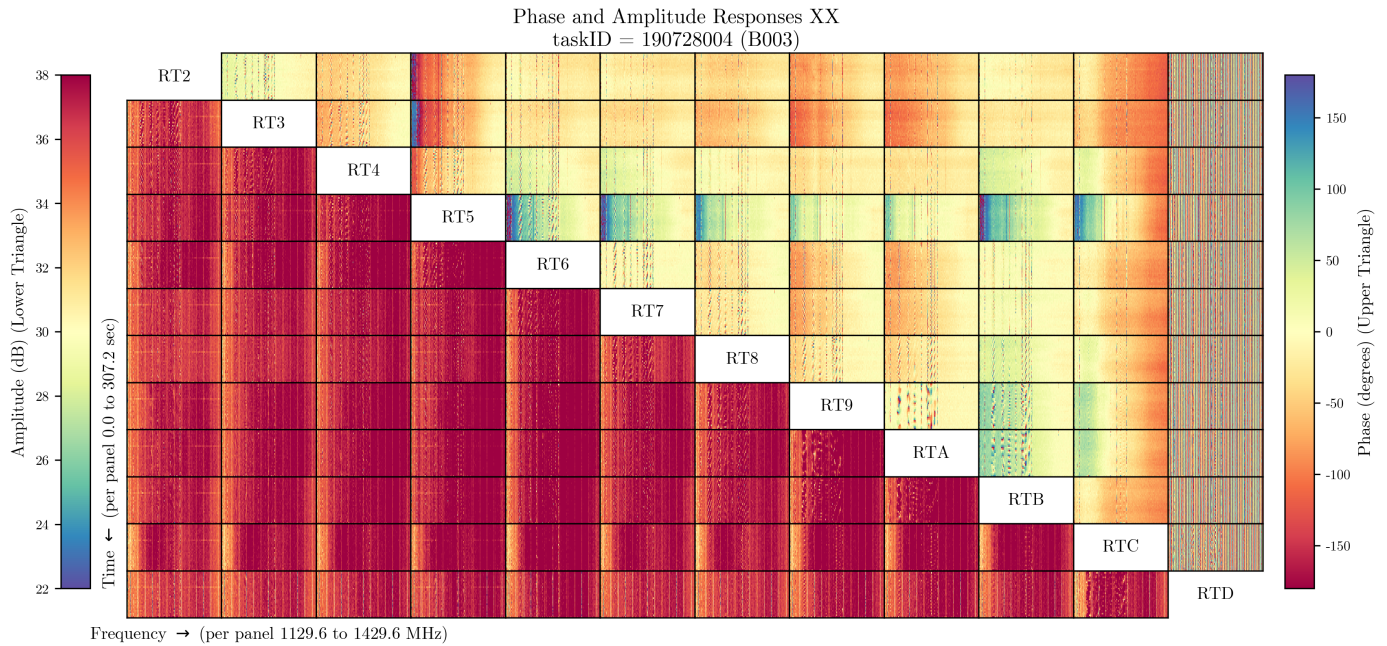[33]`https://github.com/Jordatious/ASKAP-continuum-validation`

Figure 6: Example plot of the raw uncalibrated data for the data quality assurance. Phase (top right triangle) and amplitude (bottom left triangle) are visualised as a colourscale. Each plot represents the visibilities of one baseline on a frequency vs. time basis. The difference in terms of RFI environment is clearly visible for the lower part of the observed freqeuncy band compared to the higher one. In this example antenna RTD was malfunctioning, which can easily be identified by the discontinuous phase behaviour.

First tests showed that these improvements increase the percentage of science ready data products to over 80 %.

---

[34] http://www.astropy.org

Appendix A. Configuration file

An example for a pipeline configuration file is shown below. Individual parameters are sorted in the usual execution order of the *Apercal* modules. Parameter names begin with the name of the module where the parameter is introduced and used for the first time. Parameters can have any format, which is supported by the Python ConfigParser package.

$\vdots$

```
[CROSSCAL]
crosscal_refant = 'RT2'
crosscal_initial_phase = True
crosscal_global_delay = True
crosscal_bandpass = True
crosscal_gains = True
crosscal_crosshand_delay = True
crosscal_leakage = True
crosscal_polarisation_angle = True
crosscal_transfer_to_cal = True
crosscal_transfer_to_target = True
crosscal_autocorrelation_data_fraction_limit = 0.5

[CONVERT]
convert_fluxcal = True
convert_polcal = True
convert_target = True
convert_removeuvfits = True
convert_removems = True
convert_averagems = True
```

$\vdots$

Appendix B. Parameter file

Parameters saved within the parameter file for the *preflag* module for a single beam (Beam 07) are shown below as an example. Parameters are saved as bibliographies and saved to disk as pickled numpy objects. The nomenclature for each parameter starts with the module name where it was generated, followed by the beam number and the actual parameter name. Parameter values can have all python supported formats and be of any desirable dimension. Helper functions to access or modify the parameter file can be found within the PARAM.PY subroutine in *Apercal*.

$\vdots$

```
preflag_B07_targetbeams_manualflag_auto False
preflag_B07_fluxcal_manualflag_time ['']
preflag_B07_fluxcal_manualflag_baseline ['']
preflag_B07_fluxcal_manualflag_antenna ['']
preflag_B07_polcal_manualflag_from_file False
preflag_B07_targetbeams_edges True
preflag_B07_fluxcal_ghosts False
preflag_B07_fluxcal_edges False
preflag_B07_fluxcal_manualflag_auto False
preflag_B07_targetbeams_manualflag_clipzeros False
preflag_B07_aoflagger_polcal_flag_status False
preflag_B07_fluxcal_shadow False
preflag_B07_polcal_manualflag_baseline ['']
preflag_B07_aoflagger_bandpass_status True
preflag_B07_polcal_edges False
```

```
preflag_B07_fluxcal_manualflag_clipzeros False
preflag_B07_polcal_manualflag_antenna ['']
preflag_B07_fluxcal_manualflag_corr ['']
preflag_B07_fluxcal_manualflag_channel ['']
preflag_B07_targetbeams_manualflag_time ['']
preflag_B07_targetbeams_manualflag_from_file False
preflag_B07_polcal_manualflag_auto False
preflag_B07_aoflagger_fluxcal_flag_status False
preflag_B07_polcal_ghosts False
preflag_B07_targetbeams_manualflag_corr ['']
preflag_B07_targetbeams_manualflag_antenna ['']
preflag_B07_polcal_manualflag_clipzeros False
preflag_B07_targetbeams_manualflag_channel ['0~1024']
preflag_B07_polcal_manualflag_channel ['']
preflag_B07_polcal_manualflag_time ['']
preflag_B07_targetbeams_shadow True
preflag_B07_fluxcal_manualflag_from_file False
preflag_B07_polcal_shadow False
preflag_B07_targetbeams_manualflag_baseline ['']
preflag_B07_aoflagger_targetbeams_flag_status True
preflag_B07_polcal_manualflag_corr ['']
preflag_B07_targetbeams_ghosts False
```

⋮

Appendix C. Changes to the MIRIAD source code before compilation

```
gpcopy.for: l.469  if (relax) int1 = max(int1, 1.0d0)
fits.for: l.454    call uvset(tno,'corr','r',0,0.0,0.0,0.0)
          l.461    call uvset(tno,'preamble','uv/time/baseline',0,0.0,0.0,0.0)
          l.463    call uvset(tno,'preamble','uvw/time/baseline',0,0.0,0.0,0.0)
mfcal.h   l.4      parameter(MAXDATA=1000000,MAXPOL=2)
uvio.c    l.3549   if(u*u + v*v <= limit && w <= 0) return(1);
maxdim.h           PARAMETER(MAXANT = 16)
                   PARAMETER(MAXCHAN = 32768)
                   PARAMETER(MAXWIN = 32)
                   PARAMETER(MAXWIDE = 32)
                   PARAMETER(MAXFBIN = 32)
maxdimc.h          #define MAXANT  16
                   #define MAXCHAN 32768
                   #define MAXWIN 32
                   #define MAXWIDE 32
```

References

Adams, E.A.K., van Leeuwen, J., 2019. Radio surveys now both deep and wide. Nature Astronomy 3, 188–188. doi:10.1038/s41550-019-0692-4.

Astropy Collaboration, Price-Whelan, A.M., Sipőcz, B.M., Günther, H.M., Lim, P.L., Crawford, S.M., Conseil, S., Shupe, D.L., Craig, M.W., Dencheva, N., Ginsburg, A., Vand erPlas, J.T., Bradley, L.D., Pérez-Suárez, D., de Val-Borro, M., Aldcroft, T.L., Cruz, K.L., Robitaille, T.P., Tollerud, E.J., Ardelean, C., Babej, T., Bach, Y.P., Bachetti, M., Bakanov, A.V., Bamford, S.P., Barentsen, G., Barmby, P., Baumbach, A., Berry, K.L., Biscani, F., Boquien, M., Bostroem, K.A., Bouma, L.G., Brammer, G.B., Bray, E.M., Breytenbach, H., Buddelmeijer, H., Burke, D.J., Calderone, G., Cano Rodríguez, J.L., Cara, M., Cardoso, J.V.M., Cheedella, S., Copin, Y., Corrales, L., Crichton, D., D'Avella, D., Deil, C., Depagne, É., Dietrich, J.P., Donath, A., Droettboom, M., Earl, N., Erben, T., Fabbro, S., Ferreira, L.A., Finethy, T., Fox, R.T., Garrison, L.H., Gibbons, S.L.J., Goldstein, D.A., Gommers, R., Greco, J.P., Greenfield, P., Groener, A.M., Grollier, F., Hagen, A., Hirst, P., Homeier, D., Horton, A.J., Hosseinzadeh, G., Hu, L., Hunkeler, J.S., Ivezić, Ž., Jain, A., Jenness, T., Kanarek, G., Kendrew, S., Kern, N.S., Kerzendorf, W.E., Khvalko, A., King, J., Kirkby, D., Kulkarni, A.M., Kumar, A., Lee, A., Lenz, D., Littlefair, S.P., Ma, Z., Macleod, D.M., Mastropietro, M., McCully, C., Montagnac, S., Morris, B.M., Mueller, M., Mumford, S.J., Muna, D., Murphy, N.A., Nelson, S., Nguyen, G.H., Ninan, J.P., Nöthe, M., Ogaz, S., Oh, S., Parejko, J.K., Parley, N., Pascual, S., Patil, R., Patil, A.A., Plunkett, A.L., Prochaska, J.X., Rastogi, T., Reddy Janga, V., Sabater, J., Sakurikar, P., Seifert, M., Sherbert, L.E., Sherwood-Taylor, H., Shih, A.Y., Sick, J., Silbiger, M.T., Singanamalla, S., Singer, L.P., Sladen, P.H., Sooley, K.A., Sornarajah, S., Streicher, O., Teuben, P., Thomas, S.W., Tremblay, G.R., Turner, J.E.H., Terrón, V., van Kerkwijk, M.H., de la Vega, A., Watkins, L.L., Weaver, B.A., Whitmore, J.B., Woillez, J., Zabalza, V., Astropy Contributors, 2018. The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package. Astronomical Journal 156, 123. doi:10.3847/1538-3881/aabc4f, arXiv:1801.02634.

Astropy Collaboration, Robitaille, T.P., Tollerud, E.J., Greenfield, P., Droettboom, M., Bray, E., Aldcroft, T., Davis, M., Ginsburg, A., Price-Whelan, A.M., Kerzendorf, W.E., Conley, A., Crighton, N., Barbary, K., Muna, D., Ferguson, H., Grollier, F., Parikh, M.M., Nair, P.H., Unther, H.M., Deil, C., Woillez, J., Conseil, S., Kramer, R., Turner, J.E.H., Singer, L., Fox, R., Weaver, B.A., Zabalza, V., Edwards, Z.I., Azalee Bostroem, K., Burke, D.J., Casey, A.R., Crawford, S.M., Dencheva, N., Ely, J., Jenness, T., Labrie, K., Lim, P.L., Pierfederici, F., Pontzen, A., Ptak, A., Refsdal, B., Servillat, M., Streicher, O., 2013. Astropy: A community Python package for astronomy. Astronomy and Astrophysics 558, A33. doi:10.1051/0004-6361/201322068, arXiv:1307.6212.

Becker, R.H., White, R.L., Helfand, D.J., 1995. The FIRST Survey: Faint Images of the Radio Sky at Twenty Centimeters. Astrophysical Journal 450, 559. doi:10.1086/176166.

Blanton, M.R., Bershady, M.A., Abolfathi, B., Albareti, F.D., Allende Prieto, C., Almeida, A., Alonso-García, J., Anders, F., Anderson, S.F., Andrews, B., Aquino-Ortíz, E., Aragón-Salamanca, A., Argudo-Fernández, M., Armengaud, E., Aubourg, E., Avila-Reese, V., Badenes, C., Bailey, S., Barger, K.A., Barrera-Ballesteros, J., Bartosz, C., Bates, D., Baumgarten, F., Bautista, J., Beaton, R., Beers, T.C., Belfiore, F., Bender, C.F., Berlind, A.A., Bernardi, M., Beutler, F., Bird, J.C., Bizyaev, D., Blanc, G.A., Blomqvist, M., Bolton, A.S., Boquien, M., Borissova, J., van den Bosch, R., Bovy, J., Brandt, W.N., Brinkmann, J., Brownstein, J.R., Bundy, K., Burgasser, A.J., Burtin, E., Busca, N.G., Cappellari, M., Delgado Carigi, M.L., Carlberg, J.K., Carnero Rosell, A., Carrera, R., Chanover, N.J., Cherinka, B., Cheung, E., Gómez Maqueo Chew, Y., Chiappini, C., Choi, P.D., Chojnowski, D., Chuang, C.H., Chung, H., Cirolini, R.F., Clerc, N., Cohen, R.E., Comparat, J., da Costa, L., Cousinou, M.C., Covey, K., Crane, J.D., Croft, R.A.C., Cruz-Gonzalez, I., Garrido Cuadra, D., Cunha, K., Damke, G.J., Darling, J., Davies, R., Dawson, K., de la Macorra, A., Dell'Agli, F., De Lee, N., Delubac, T., Di Mille, F., Diamond-Stanic, A., Cano-Díaz, M., Donor, J., Downes, J.J., Drory, N., du Mas des Bourboux, H., Duckworth, C.J., Dwelly, T., Dyer, J., Ebelke, G., Eigenbrot, A.D., Eisenstein, D.J., Emsellem, E., Eracleous, M., Escoffier, S., Evans, M.L., Fan, X., Fernández-Alvar, E., Fernandez-Trincado, J.G., Feuillet, D.K., Finoguenov, A., Fleming, S.W., Font-Ribera, A., Fredrickson, A., Freischlad, G., Frinchaboy, P.M., Fuentes, C.E., Galbany, L., Garcia-Dias, R., García-Hernández, D.A., Gaulme, P., Geisler, D., Gelfand, J.D., Gil-Marín, H., Gillespie, B.A., Goddard, D., Gonzalez-Perez, V., Grabowski, K., Green, P.J., Grier, C.J.,

Gunn, J.E., Guo, H., Guy, J., Hagen, A., Hahn, C., Hall, M., Harding, P., Hasselquist, S., Hawley, S.L., Hearty, F., Gonzalez Hernández, J.I., Ho, S., Hogg, D.W., Holley-Bockelmann, K., Holtzman, J.A., Holzer, P.H., Huehnerhoff, J., Hutchinson, T.A., Hwang, H.S., Ibarra-Medel, H.J., da Silva Ilha, G., Ivans, I.I., Ivory, K., Jackson, K., Jensen, T.W., Johnson, J.A., Jones, A., Jönsson, H., Jullo, E., Kamble, V., Kinemuchi, K., Kirkby, D., Kitaura, F.S., Klaene, M., Knapp, G.R., Kneib, J.P., Kollmeier, J.A., Lacerna, I., Lane, R.R., Lang, D., Law, D.R., Lazarz, D., Lee, Y., Le Goff, J.M., Liang, F.H., Li, C., Li, H., Lian, J., Lima, M., Lin, L., Lin, Y.T., Bertran de Lis, S., Liu, C., de Icaza Lizaola, M.A.C., Long, D., Lucatello, S., Lundgren, B., MacDonald, N.K., Deconto Machado, A., MacLeod, C.L., Mahadevan, S., Geimba Maia, M.A., Maiolino, R., Majewski, S.R., Malanushenko, E., Malanushenko, V., Manchado, A., Mao, S., Maraston, C., Marques-Chaves, R., Masseron, T., Masters, K.L., McBride, C.K., McDermid, R.M., McGrath, B., McGreer, I.D., Medina Peña, N., Melendez, M., Merloni, A., Merrifield, M.R., Meszaros, S., Meza, A., Minchev, I., Minniti, D., Miyaji, T., More, S., Mulchaey, J., Müller-Sánchez, F., Muna, D., Munoz, R.R., Myers, A.D., Nair, P., Nandra, K., Correa do Nascimento, J., Negrete, A., Ness, M., Newman, J.A., Nichol, R.C., Nidever, D.L., Nitschelm, C., Ntelis, P., O'Connell, J.E., Oelkers, R.J., Oravetz, A., Oravetz, D., Pace, Z., Padilla, N., Palanque-Delabrouille, N., Alonso Palicio, P., Pan, K., Parejko, J.K., Parikh, T., Pâris, I., Park, C., Patten, A.Y., Peirani, S., Pellejero-Ibanez, M., Penny, S., Percival, W.J., Perez-Fournon, I., Petitjean, P., Pieri, M.M., Pinsonneault, M., Pisani, A., Poleski, R., Prada, F., Prakash, A., Queiroz, A.B.d.A., Raddick, M.J., Raichoor, A., Barboza Rembold, S., Richstein, H., Riffel, R.A., Riffel, R., Rix, H.W., Robin, A.C., Rockosi, C.M., Rodríguez-Torres, S., Roman-Lopes, A., Román-Zúñiga, C., Rosado, M., Ross, A.J., Rossi, G., Ruan, J., Ruggeri, R., Rykoff, E.S., Salazar-Albornoz, S., Salvato, M., Sánchez, A.G., Aguado, D.S., Sánchez-Gallego, J.R., Santana, F.A., Santiago, B.X., Sayres, C., Schiavon, R.P., da Silva Schimoia, J., Schlafly, E.F., Schlegel, D.J., Schneider, D.P., Schultheis, M., Schuster, W.J., Schwope, A., Seo, H.J., Shao, Z., Shen, S., Shetrone, M., Shull, M., Simon, J.D., Skinner, D., Skrutskie, M.F., Slosar, A., Smith, V.V., Sobeck, J.S., Sobreira, F., Somers, G., Souto, D., Stark, D.V., Stassun, K., Stauffer, F., Steinmetz, M., Storchi-Bergmann, T., Streblyanska, A., Stringfellow, G.S., Suárez, G., Sun, J., Suzuki, N., Szigeti, L., Taghizadeh-Popp, M., Tang, B., Tao, C., Tayar, J., Tembe, M., Teske, J., Thakar, A.R., Thomas, D., Thompson, B.A., Tinker, J.L., Tissera, P., Tojeiro, R., Hernandez Toledo, H., de la Torre, S., Tremonti, C., Troup, N.W., Valenzuela, O., Martinez Valpuesta, I., Vargas-González, J., Vargas-Magaña, M., Vazquez, J.A., Villanova, S., Vivek, M., Vogt, N., Wake, D., Walterbos, R., Wang, Y., Weaver, B.A., Weijmans, A.M., Weinberg, D.H., Westfall, K.B., Whelan, D.G., Wild, V., Wilson, J., Wood-Vasey, W.M., Wylezalek, D., Xiao, T., Yan, R., Yang, M., Ybarra, J.E., Yèche, C., Zakamska, N., Zamora, O., Zarrouk, P., Zasowski, G., Zhang, K., Zhao, G.B., Zheng, Z., Zheng, Z., Zhou, X., Zhou, Z.M., Zhu, G.B., Zoccali, M., Zou, H., 2017. Sloan Digital Sky Survey IV: Mapping the Milky Way, Nearby Galaxies, and the Distant Universe. Astronomical Journal 154, 28. doi:10.3847/1538-3881/aa7567, arXiv:1703.00052.

Brentjens, M.A., de Bruyn, A.G., 2005. Faraday rotation measure synthesis. Astronomy and Astrophysics 441, 1217–1228. doi:10.1051/0004-6361:20052990, arXiv:arXiv:astro-ph/0507349.

Burn, B.J., 1966. On the depolarization of discrete radio sources by Faraday dispersion. Monthly Notices of the RAS 133, 67.

Chambers, K.C., Magnier, E.A., Metcalfe, N., Flewelling, H.A., Huber, M.E., Waters, C.Z., Denneau, L., Draper, P.W., Farrow, D., Finkbeiner, D.P., Holmberg, C., Koppenhoefer, J., Price, P.A., Rest, A., Saglia, R.P., Schlafly, E.F., Smartt, S.J., Sweeney, W., Wainscoat, R.J., Burgett, W.S., Chastel, S., Grav, T., Heasley, J.N., Hodapp, K.W., Jedicke, R., Kaiser, N., Kudritzki, R.P., Luppino, G.A., Lupton, R.H., Monet, D.G., Morgan, J.S., Onaka, P.M., Shiao, B., Stubbs, C.W., Tonry, J.L., White, R., Bañados, E., Bell, E.F., Bender, R., Bernard, E.J., Boegner, M., Boffi, F., Botticella, M.T., Calamida, A., Casertano, S., Chen, W.P., Chen, X., Cole, S., Deacon, N., Frenk, C., Fitzsimmons, A., Gezari, S., Gibbs, V., Goessl, C., Goggia, T., Gourgue, R., Goldman, B., Grant, P., Grebel, E.K., Hambly, N.C., Hasinger, G., Heavens, A.F., Heckman, T.M., Henderson, R., Henning, T., Holman, M., Hopp, U., Ip, W.H., Isani, S., Jackson, M., Keyes, C.D., Koekemoer, A.M., Kotak, R., Le, D., Liska, D., Long, K.S., Lucey, J.R., Liu, M., Martin, N.F., Masci, G., McLean, B., Mindel, E., Misra, P., Morganson, E., Murphy, D.N.A., Obaika, A., Narayan, G., Nieto-Santisteban, M.A., Norberg, P., Peacock, J.A., Pier, E.A., Postman, M., Primak, N., Rae, C., Rai, A., Riess, A., Riffeser, A., Rix, H.W., Röser, S., Russel, R., Rutz,

L., Schilbach, E., Schultz, A.S.B., Scolnic, D., Strolger, L., Szalay, A., Seitz, S., Small, E., Smith, K.W., Soderblom, D.R., Taylor, P., Thomson, R., Taylor, A.N., Thakar, A.R., Thiel, J., Thilker, D., Unger, D., Urata, Y., Valenti, J., Wagner, J., Walder, T., Walter, F., Watters, S.P., Werner, S., Wood-Vasey, W.M., Wyse, R., 2016. The Pan-STARRS1 Surveys. arXiv e-prints , arXiv:1612.05560arXiv:1612.05560.

Condon, J.J., Cotton, W.D., Greisen, E.W., Yin, Q.F., Perley, R.A., Taylor, G.B., Broderick, J.J., 1998. The NRAO VLA Sky Survey. Astronomical Journal 115, 1693–1716. doi:10.1086/300337.

Cornwell, T.J., Wilkinson, P.N., 1981. A new method for making maps with unstable radio interferometers. Monthly Notices of the RAS 196, 1067–1086.

Cotton, W.D., 1990. A Programmer's Guide to the NRAO Astronomical Image Processing System (Charlottesville:NRAO), chap. 14.7, http://www.aoc.nrao.edu/aips/goaips.html.

D'Agostino, R., Pearson, E.S., 1973. Tests for departure from normality. Empirical results for the distributions of b2 and b1. Biometrika 60, 613–622. URL: https://doi.org/10.1093/biomet/60.3.613, doi:10.1093/biomet/60.3.613, arXiv:https://academic.oup.com/biomet/article-pdf/60/3/613/576955/60.3.613.pdf.

D'Agostino, R.B., 1971. An omnibus test of normality for moderate and large size samples. Biometrika 58, 341–348. URL: https://doi.org/10.1093/biomet/58.2.341, doi:10.1093/biomet/58.2.341, arXiv:https://academic.oup.com/biomet/article-pdf/58/2/341/732726/58-2-341.pdf.

Dalton, G., Trager, S., Abrams, D.C., Bonifacio, P., López Aguerri, J.A., Middleton, K., Benn, C., Dee, K., Sayède, F., Lewis, I., Pragt, J., Pico, S., Walton, N., Rey, J., Allende Prieto, C., Peñate, J., Lhome, E., Agócs, T., Alonso, J., Terrett, D., Brock, M., Gilbert, J., Ridings, A., Guinouard, I., Verheijen, M., Tosh, I., Rogers, K., Steele, I., Stuik, R., Tromp, N., Jasko, A., Kragt, J., Lesman, D., Mottram, C., Bates, S., Gribbin, F., Fernando Rodriguez, L., Delgado, J.M., Martin, C., Cano, D., Navarro, R., Irwin, M., Lewis, J., Gonzalez Solares, E., O'Mahony, N., Bianco, A., Zurita, C., ter Horst, R., Molinari, E., Lodi, M., Guerra, J., Vallenari, A., Baruffolo, A., 2014. Project overview and update on WEAVE: the next generation wide-field spectroscopy facility for the William Herschel Telescope, in: Ramsay, S.K., McLean, I.S., Takami, H. (Eds.), Ground-based and Airborne Instrumentation for Astronomy V, p. 91470L. doi:10.1117/12.2055132, arXiv:1412.0843.

de Gasperin, F., Williams, W.L., Best, P., Bruggen, M., Brunetti, G., Cuciti, V., Dijkema, T.J., Hardcastle, M.J., Norden, M.J., Offringa, A., Shimwell, T., van Weeren, R., Bomans, D., Bonafede, A., Botteon, A., Callingham, J.R., Cassano, R., Chyzy, K.T., Emig, K.L., Edler, H., Haverkorn, M., Heald, G., Heesen, V., Iacobelli, M., Intema, H.T., Kadler, M., Malek, K., Mevius, M., Miley, G., Mingo, B., Morabito, L.K., Sabater, J., Morganti, R., Orru, E., Pizzo, R., Prandoni, I., Shulevski, A., Tasse, C., Vaccari, M., Zarka, P., Rottgering, H., 2021. The LOFAR LBA Sky Survey I. survey description and preliminary data release. arXiv e-prints , arXiv:2102.09238arXiv:2102.09238.

Drabent, A., Hoeft, M., Mechev, A.P., Oonk, J.B.R., Shimwell, T.W., Sweijen, F., Danezi, A., Schrijvers, C., Manzano, C., Tsigenov, O., Dettmar, R.J., Brüggen, M., Schwarz, D.J., 2019. Realising the LOFAR Two-Metre Sky Survey – using the supercomputer JUWELS at the Forschungszentrum Jülich. arXiv e-prints , arXiv:1910.13835arXiv:1910.13835.

HI4PI Collaboration, Ben Bekhti, N., Flöer, L., Keller, R., Kerp, J., Lenz, D., Winkel, B., Bailin, J., Calabretta, M.R., Dedes, L., Ford, H.A., Gibson, B.K., Haud, U., Janowiecki, S., Kalberla, P.M.W., Lockman, F.J., McClure-Griffiths, N.M., Murphy, T., Nakanishi, H., Pisano, D.J., Staveley-Smith, L., 2016. HI4PI: A full-sky H I survey based on EBHIS and GASS. Astronomy and Astrophysics 594, A116. doi:10.1051/0004-6361/201629178, arXiv:1610.06175.

Hill, G.J., Gebhardt, K., Komatsu, E., Drory, N., MacQueen, P.J., Adams, J., Blanc, G.A., Koehler, R., Rafal, M., Roth, M.M., Kelz, A., Gronwall, C., Ciardullo, R., Schneider, D.P., 2008. The Hobby-Eberly Telescope Dark Energy Experiment (HETDEX): Description and Early Pilot Survey Results, in: Kodama, T., Yamada, T., Aoki, K. (Eds.), Panoramic Views of Galaxy Formation and Evolution, p. 115. arXiv:0806.0183.

Högbom, J.A., 1974. Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines. Astronomy and Astrophysicss 15, 417.

Janssen, M., Goddi, C., Falcke, H., van Rossum, D., van Bemmel, I., Kettenis, M., Small, D., Martí-Vidal, I., 2019. rPICARD: A CASA-based Calibration Pipeline for VLBI Data. arXiv e-prints ,

arXiv:1905.01905arXiv:1905.01905.

Józsa, G.I.G., White, S.V., Thorat, K., Smirnov, O.M., Serra, P., Ramatsoku, M., Ramaila, A.J.T., Perkins, S.J., Molnár, D.C., Makhathini, S., Maccagni, F.M., Kleiner, D., Kamphuis, P., Hugo, B.V., de Blok, W.J.G., Andati, L.A.L., 2020b. CARACal: Containerized Automated Radio Astronomy Calibration pipeline. arXiv:2006.014.

Józsa, G.I.G., White, S.V., Thorat, K., Smirnov, O.M., Serra, P., Ramatsoku, M., Ramaila, A.J.T., Perkins, S.J., Molnár, D.C., Makhathini, S., Maccagni, F.M., Kleiner, D., Kamphuis, P., Hugo, B.V., de Blok, W.J.G., Andati, L.A.L., 2020a. MeerKATHI – an end-to-end data reduction pipeline for MeerKAT and other radio telescopes. arXiv e-prints , arXiv:2006.02955arXiv:2006.02955.

Kent, B.R., Masters, J.S., Chandler, C.J., Davis, L.E., Kern, J.S., Ott, J., Schinzel, F.K., Medlin, D., Muders, D., Williams, S., Geers, V.C., Momjian, E., Butler, B.J., Nakazato, T., Sugimoto, K., 2018. The Very Large Array Data Processing Pipeline, in: American Astronomical Society Meeting Abstracts #231, p. 342.14.

Kurtzer, G.M., Sochat, V., Bauer, M.W., 2017. Singularity: Scientific containers for mobility of compute. PLoS ONE doi:10.1371/journal.pone.0177459.

Marvil, J., Chandler, C., Kent, B.R., Myers, S.T., Schinzel, F., 2019. Calibration of the VLA Sky Survey, in: American Astronomical Society Meeting Abstracts #233, p. 446.02.

McMullin, J.P., Waters, B., Schiebel, D., Young, W., Golap, K., 2007. CASA Architecture and Applications, in: Shaw, R.A., Hill, F., Bell, D.J. (Eds.), Astronomical Data Analysis Software and Systems XVI, p. 127.

Mechev, A.P., Shimwell, T.W., Plaat, A., Intema, H., Varbanescu, A.L., Rottgering, H.J.A., 2019. Scalability model for the LOFAR direction independent pipeline. Astronomy and Computing 28, 100293. doi:10.1016/j.ascom.2019.100293, arXiv:1906.11516.

Mohan, N., Rafferty, D., 2015. PyBDSF: Python Blob Detection and Source Finder. arXiv:1502.007.

Muders, D., Wyrowski, F., Lightfoot, J., Williams, S., Nakazato, T., Kosugi, G., Davis, L., Kern, J., 2014. The ALMA Pipeline. volume 485 of *Astronomical Society of the Pacific Conference Series*. p. 383.

Offringa, A.R., de Bruyn, A.G., Biehl, M., Zaroubi, S., Bernardi, G., Pandey, V.N., 2010. Post-correlation radio frequency interference classification methods. Monthly Notices of the RAS 405, 155–167. doi:10.1111/j.1365-2966.2010.16471.x, arXiv:1002.1957.

Offringa, A.R., de Bruyn, A.G., Zaroubi, S., 2012a. Post-correlation filtering techniques for off-axis source and RFI removal. Monthly Notices of the RAS , 2693doi:10.1111/j.1365-2966.2012.20633.x, arXiv:1201.5564.

Offringa, A.R., van de Gronde, J.J., Roerdink, J.B.T.M., 2012b. A morphological algorithm for improving radio-frequency interference detection. Astronomy and Astrophysics 539, A95. doi:10.1051/0004-6361/201118497, arXiv:1201.3364.

Pearson, T.J., Readhead, A.C.S., 1984. Image Formation by Self-Calibration in Radio Astronomy. Annual Review of Astron and Astrophys 22, 97–130. doi:10.1146/annurev.aa.22.090184.000525.

Perley, R.A., Butler, B.J., 2013. Integrated Polarization Properties of 3C48, 3C138, 3C147, and 3C286. Astrophysical Journals 206, 16. doi:10.1088/0067-0049/206/2/16, arXiv:1302.6662.

Perley, R.A., Butler, B.J., 2017. An Accurate Flux Density Scale from 50 MHz to 50 GHz. Astrophysical Journals 230, 7. doi:10.3847/1538-4365/aa6df9, arXiv:1609.05940.

Popping, A., Braun, R., 2008. The standing wave phenomenon in radio telescopes. Frequency modulation of the WSRT primary beam. Astronomy and Astrophysics 479, 903–913. doi:10.1051/0004-6361:20079122, arXiv:0712.2303.

Prasad, J., Chengalur, J., 2012. FLAGCAL: a flagging and calibration package for radio interferometric data. Experimental Astronomy 33, 157–171. doi:10.1007/s10686-011-9279-5, arXiv:1111.6415.

Pratley, L., Johnston-Hollitt, M., 2016. An improved method for polarimetric image restoration in interferometry. Monthly Notices of the RAS 462, 3483–3501. doi:10.1093/mnras/stw1377, arXiv:1606.01482.

Rengelink, R.B., Tang, Y., de Bruyn, A.G., Miley, G.K., Bremer, M.N., Roettgering, H.J.A., Bremer, M.A.R., 1997. The Westerbork Northern Sky Survey (WENSS), I. A 570 square degree Mini-Survey around the North Ecliptic Pole. Astronomy and Astrophysicss 124, 259–280. doi:10.1051/aas:1997358.

Sabater, J., Sánchez-Expósito, S., Best, P., Garrido, J., Verdes-Montenegro,

L., Lezzi, D., 2017. Calibration of LOFAR data on the cloud. Astronomy and Computing 19, 75–89. doi:10.1016/j.ascom.2017.04.001, arXiv:1704.05064.

Sault, R.J., Teuben, P.J., Wright, M.C.H., 1995. A Retrospective View of MIRIAD, in: Shaw, R.A., Payne, H.E., Hayes, J.J.E. (Eds.), Astronomical Data Analysis Software and Systems IV, p. 433. arXiv:astro-ph/0612759.

Sault, R.J., Wieringa, M.H., 1994. Multi-frequency synthesis techniques in radio interferometric imaging. Astronomy and Astrophysicss 108, 585–594.

Shimwell, T.W., Röttgering, H.J.A., Best, P.N., Williams, W.L., Dijkema, T.J., de Gasperin, F., Hardcastle, M.J., Heald, G.H., Hoang, D.N., Horneffer, A., Intema, H., Mahony, E.K., Mandal, S., Mechev, A.P., Morabito, L., Oonk, J.B.R., Rafferty, D., Retana-Montenegro, E., Sabater, J., Tasse, C., van Weeren, R.J., Brüggen, M., Brunetti, G., Chyży, K.T., Conway, J.E., Haverkorn, M., Jackson, N., Jarvis, M.J., McKean, J.P., Miley, G.K., Morganti, R., White, G.J., Wise, M.W., van Bemmel, I.M., Beck, R., Brienza, M., Bonafede, A., Calistro Rivera, G., Cassano, R., Clarke, A.O., Cseh, D., Deller, A., Drabent, A., van Driel, W., Engels, D., Falcke, H., Ferrari, C., Fröhlich, S., Garrett, M.A., Harwood, J.J., Heesen, V., Hoeft, M., Horellou, C., Israel, F.P., Kapińska, A.D., Kunert-Bajraszewska, M., McKay, D.J., Mohan, N.R., Orrú, E., Pizzo, R.F., Prandoni, I., Schwarz, D.J., Shulevski, A., Sipior, M., Smith, D.J.B., Sridhar, S.S., Steinmetz, M., Stroe, A., Varenius, E., van der Werf, P.P., Zensus, J.A., Zwart, J.T.L., 2017. The LOFAR Two-metre Sky Survey. I. Survey description and preliminary data release. Astronomy and Astrophysics 598, A104. doi:10.1051/0004-6361/201629313, arXiv:1611.02700.

Smith, D.J.B., Best, P.N., Duncan, K.J., Hatch, N.A., Jarvis, M.J., Röttgering, H.J.A., Simpson, C.J., Stott, J.P., Cochrane, R.K., Coppin, K.E., Dannerbauer, H., Davis, T.A., Geach, J.E., Hale, C.L., Hardcastle, M.J., Hatfield, P.W., Houghton, R.C.W., Maddox, N., McGee, S.L., Morabito, L., Nisbet, D., Pand ey-Pommier, M., Prandoni, I., Saxena, A., Shimwell, T.W., Tarr, M., van Bemmel, I., Verma, A., White, G.J., Williams, W.L., 2016. The WEAVE-LOFAR Survey, in: Reylé, C., Richard, J., Cambrésy, L., Deleuil, M., Pécontal, E., Tresse, L., Vauglin, I. (Eds.), SF2A-2016: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics, pp. 271–280. arXiv:1611.02706.

Staley, T.D., Anderson, G.E., 2015. Chimenea and other tools: Automated imaging of multi-epoch radio-synthesis data with CASA. Astronomy and Computing 13, 38–49. doi:10.1016/j.ascom.2015.08.003, arXiv:1505.08123.

Tasse, C., Hugo, B., Mirmont, M., Smirnov, O., Atemkeng, M., Bester, L., Hardcastle, M.J., Lakhoo, R., Perkins, S., Shimwell, T., 2018. Faceting for direction-dependent spectral deconvolution. Astronomy and Astrophysics 611, A87. doi:10.1051/0004-6361/201731474, arXiv:1712.02078.

Terlouw, J.P., Vogelaar, M.G.R., 2015. Kapteyn Package, version 2.3. Kapteyn Astronomical Institute. Groningen. Available from http://www.astro.rug.nl/software/kapteyn/.

Valiante, E., Smith, M.W.L., Eales, S., Maddox, S.J., Ibar, E., Hopwood, R., Dunne, L., Cigan, P.J., Dye, S., Pascale, E., Rigby, E.E., Bourne, N., Furlanetto, C., Ivison, R.J., 2016. The Herschel-ATLAS data release 1 - I. Maps, catalogues and number counts. Monthly Notices of the RAS 462, 3146–3179. doi:10.1093/mnras/stw1806, arXiv:1606.09615.

van Cappellen, W.A., Oosterloo, T.A., Verheijen, M.A.W., Adams, E.A.K., Adebahr, B., Braun, R., Hess, K.M., Holties, H., van der Hulst, J.M., Hut, B., Kooistra, E., van Leeuwen, J., Loose, G.M., Morganti, R., Moss, V.A., Orrú, E., Ruiter, M., Schoenmakers, A.P., Vermaas, N.J., Wijnholds, S.J., van Amesfoort, A.S., Arts, M.J., Attema, J.J., Bakker, L., Bassa, C.G., Bast, J.E., Benthem, P., Beukema, R., Blaauw, R., de Blok, W.J.G., Bouwhuis, M., van den Brink, R.H., Connor, L., Coolen, A.H.W.M., Damstra, S., van Diepen, G.N.J., de Goei, R., Dénes, H., Drost, M., Ebbendorf, N., Frank, B.S., Gardenier, D.W., Gerbers, M., Grange, Y.G., Grit, T., Gunst, A.W., Gupta, N., Ivashina, M.V., Józsa, G.I.G., Janssen, G.H., Koster, A., Kruithof, G.H., Kuindersma, S.J., Kutkin, A., Lucero, D.M., Maan, Y., Maccagni, F.M., van der Marel, J., Mika, A., Morawietz, J., Mulder, H., Mulder, E., Norden, M.J., Offringa, A.R., Oostrum, L.C., Overeem, R.E., Paragi, Z., Pepping, H.J., Petroff, E., Pisano, D.J., Polatidis, A.G., Prasad, P., de Reijer, J.P.R., Romein, J.W., Schaap, J., Schoonderbeek, G.W., Schulz, R., van der Schuur, D., Sclocco, A., Sluman, J.J., Smits, R., Stappers, B.W., Straal, S.M., Stuurwold, K.J.C., Verstappen, J., Vohl, D., Wierenga, K.J., Woestenburg, E.E.M., Zanting, A.W., Ziemke, J., 2021. Apertif, Phased Array Feeds for the Westerbork Synthesis Radio Telescope. arXiv e-prints , arXiv:2109.14234arXiv:2109.14234.

Van Rossum, G., Drake Jr, F.L., 1995. Python reference manual. Centrum voor Wiskunde en Informatica Amsterdam.

van Weeren, R.J., Williams, W.L., Hardcastle, M.J., Shimwell, T.W., Rafferty, D.A., Sabater, J., Heald, G., Sridhar, S.S., Dijkema, T.J., Brunetti, G., Brüggen, M., Andrade-Santos, F., Ogrean, G.A., Röttgering, H.J.A., Dawson, W.A., Forman, W.R., de Gasperin, F., Jones, C., Miley, G.K., Rudnick, L., Sarazin, C.L., Bonafede, A., Best, P.N., Bîrzan, L., Cassano, R., Chyży, K.T., Croston, J.H., Ensslin, T., Ferrari, C., Hoeft, M., Horellou, C., Jarvis, M.J., Kraft, R.P., Mevius, M., Intema, H.T., Murray, S.S., Orrú, E., Pizzo, R., Simionescu, A., Stroe, A., van der Tol, S., White, G.J., 2016. LOFAR Facet Calibration. Astrophysical Journals 223, 2. doi:10.3847/0067-0049/223/1/2, arXiv:1601.05422.

Wells, D.C., 1985. NRAO's Astronomical Image Processing System (AIPS). Springer US, Boston, MA. pp. 195–209. URL: https://doi.org/10.1007/978-1-4615-9433-8_18, doi:10.1007/978-1-4615-9433-8_18.

Wright, E.L., Eisenhardt, P.R.M., Mainzer, A.K., Ressler, M.E., Cutri, R.M., Jarrett, T., Kirkpatrick, J.D., Padgett, D., McMillan, R.S., Skrutskie, M., Stanford, S.A., Cohen, M., Walker, R.G., Mather, J.C., Leisawitz, D., Gautier, Thomas N., I., McLean, I., Benford, D., Lonsdale, C.J., Blain, A., Mendez, B., Irace, W.R., Duval, V., Liu, F., Royer, D., Heinrichsen, I., Howard, J., Shannon, M., Kendall, M., Walsh, A.L., Larsen, M., Cardon, J.G., Schick, S., Schwalm, M., Abid, M., Fabinsky, B., Naes, L., Tsai, C.W., 2010. The Wide-field Infrared Survey Explorer (WISE): Mission Description and Initial On-orbit Performance. Astronomical Journal 140, 1868–1881. doi:10.1088/0004-6256/140/6/1868, arXiv:1008.0031.

Yan, R., Bundy, K., Law, D.R., Bershady, M.A., Andrews, B., Cherinka, B., Diamond-Stanic, A.M., Drory, N., MacDonald, N., Sánchez-Gallego, J.R., Thomas, D., Wake, D.A., Weijmans, A.M., Westfall, K.B., Zhang, K., Aragón-Salamanca, A., Belfiore, F., Bizyaev, D., Blanc, G.A., Blanton, M.R., Brownstein, J., Cappellari, M., D'Souza, R., Emsellem, E., Fu, H., Gaulme, P., Graham, M.T., Goddard, D., Gunn, J.E., Harding, P., Jones, A., Kinemuchi, K., Li, C., Li, H., Maiolino, R., Mao, S., Maraston, C., Masters, K., Merrifield, M.R., Oravetz, D., Pan, K., Parejko, J.K., Sanchez, S.F., Schlegel, D., Simmons, A., Thanjavur, K., Tinker, J., Tremonti, C., van den Bosch, R., Zheng, Z., 2016. SDSS-IV MaNGA IFS Galaxy Survey—Survey Design, Execution, and Initial Data Quality. Astronomical Journal 152, 197. doi:10.3847/0004-6256/152/6/197, arXiv:1607.08613.

Yatawatta, S., 2020. Stochastic calibration of radio interferometers. Monthly Notices of the RAS 493, 6071–6078. doi:10.1093/mnras/staa648, arXiv:2003.00986.