Draft version September 17, 2022
Typeset using LaTeX default style in AASTeX631

Pipeline Improvements and Preliminary Results of the NURF Survey

Bailee Wolfe[1]

[1] *National Radio Astronomy Observatory*
1003 Lopezville Rd
Socorro, NM 87801, USA

ABSTRACT

With the ever increasing demand for observations and science at higher frequencies, there is a strong need for a high frequency VLBI calibrator survey. The NURF Survey offers an efficient way to follow up on known reference sources at higher frequencies and search for new reference sources by collaborating with USNO to make use of VLBA antennas left idle during certain observations and making use of automated procedures for source selection, scheduling, and a pipeline for data reduction. The sources targeted for this survey include pre-established International Celestial Reference Frame sources along with sources from VLASS and Gaia catalogues. These sources are intended to guide future ICRF work and for use as phase reference sources for the VLBA.

1. BACKGROUND

The International Celestial Reference Frame (ICRF) (Ma et al. 1997) is a catalog of bright, extra-galactic radio sources acting as a reference frame from which we can determine our position and the position of other objects in space. The sources in the ICRF are all quasars, as they need to be bright, relatively compact, and sufficiently far away so that their proper motions and parallax angles are negligible, allowing them to be used as 'stationary' references. Additionally, these sources make excellent phase reference sources for Very Long Baseline Interferometry (VLBI) radio observations, due to their brightness and compactness.

Historically, the ICRF has focused mostly on sources in S/X band, with few K band sources included, providing adequate sky coverage in the lower bands, with only sparse coverage in K band, as can be seen in Fig. 1. The increasing need for accurate astrometry, and thus a more accurate reference frame, means that there is a need for new reference frames which includes data at higher frequencies. With current science demands also pushing for VLBI observations at higher frequencies, it is necessary to expand the catalogue of available sources at K Band.

2. NURF SURVEY

The NRAO/USNO Reference and Flux Calibrator Survey (NURF) is a collaboration between the US Naval Observatory (USNO) and the National Radio Astronomy Observatory (NRAO) which aims to follow up on ICRF sources at higher frequencies, as well as adding sources from a cross match of VLASS and Gaia catalogues. This expands the catalogue of phase reference sources into higher frequencies for better resolution and accuracy and to search for potential phase calibrator sources for the Very Long Baseline Array (VLBA) in order to better meet the needs of current science. The survey targets around 6000 sources and carries out K band observations and simultaneous S and X band observations.

2.1. *Observing Time, Source Selection, and Efficiency*

Sources are selected from the current catalog of ICRF sources and also cross-correlated from the Gaia (Gaia Collaboration et al. 2022) and VLASS (Lacy 2019) surveys to primarily target quasars or suspected quasars. Gaia is a mission led by the European Space Agency with a goal of accurately mapping our galaxy, while also following up on quasars used in reference frames with hopes of detecting new extragalactic sources. The VLA Sky Survey (VLASS) is a radio all-sky survey led by NRAO. By cross-matching sources in Gaia and VLASS catalogues we hope to find quasars which will be bright and visible in radio, making them suitable candidates for detection with the VLBA. By targeting these sources the NURF Survey aims to follow up all ICRF sources and also fill in any gaps left in S, X and K band as shown in Fig. 2

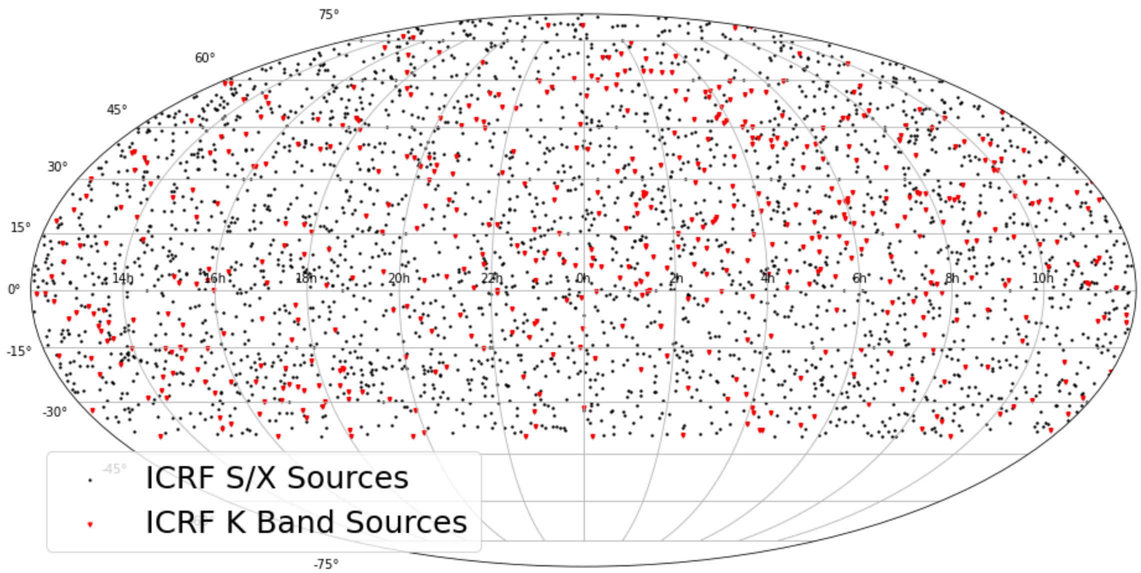

Figure 1. This plot shows the ICRF sky coverage in S and X bands, represented as black circles and coverage in K band, represented as red triangles. Where S band corresponds to 2-4 GHz, X band corresponds to 8-12 GHz, and K band corresponds to 20-25 GHz.

A survey of this size would typically take up large amounts of observing time on a telescope, however this issue is avoided by taking observations using antennas left idle during USNO projects that do not require the full ten VLBA dishes. NURF typically observes during a USNO project which makes use of the outer five VLBA antenna, leaving five shorter baseline antennas for our use, although the specific sites and number of antennas used for any given observation varies since NURF observations are taken during any USNO observation that does not use the full VLBA. These short baseline observations are advantageous as most sources will likely appear to be point-like when observed on these baselines, thus sources which are detected in K-band in these observations are also likely to be detected with the full array. In terms of efficiency, the short baselines allow for easier calibration and scheduling, as the antenna which are close together are more likely to experience similar weather conditions. A typical observation targets around 15 sources per hour with around 1 minute on source per band which includes time to slew to source. This provides relatively low time on source, but a good calibrator source should be easily detectable with a minute on source.

There is a need for automation and efficiency in this project, primarily due to the fact that it is not the primary focus of any of the scientists involved. This has lead to automated scripts used for source selection and scheduling. An observation starts by randomly selecting a source from the least observed sources in the database and then the next closest least observed sources are chosen. Data reduction is also automated using a pipeline, which is discussed in section 2.2.

2.2. *Pipeline*

The pipeline is written as a C Shell script which loads in the data from a given location on the user's drive, and then calls to AIPS to calibrate, image, and model fit the data. The user is prompted to provide the segment code, which is a two letter code unique to each observation or epoch. They then are asked to select a strong fringe finder source

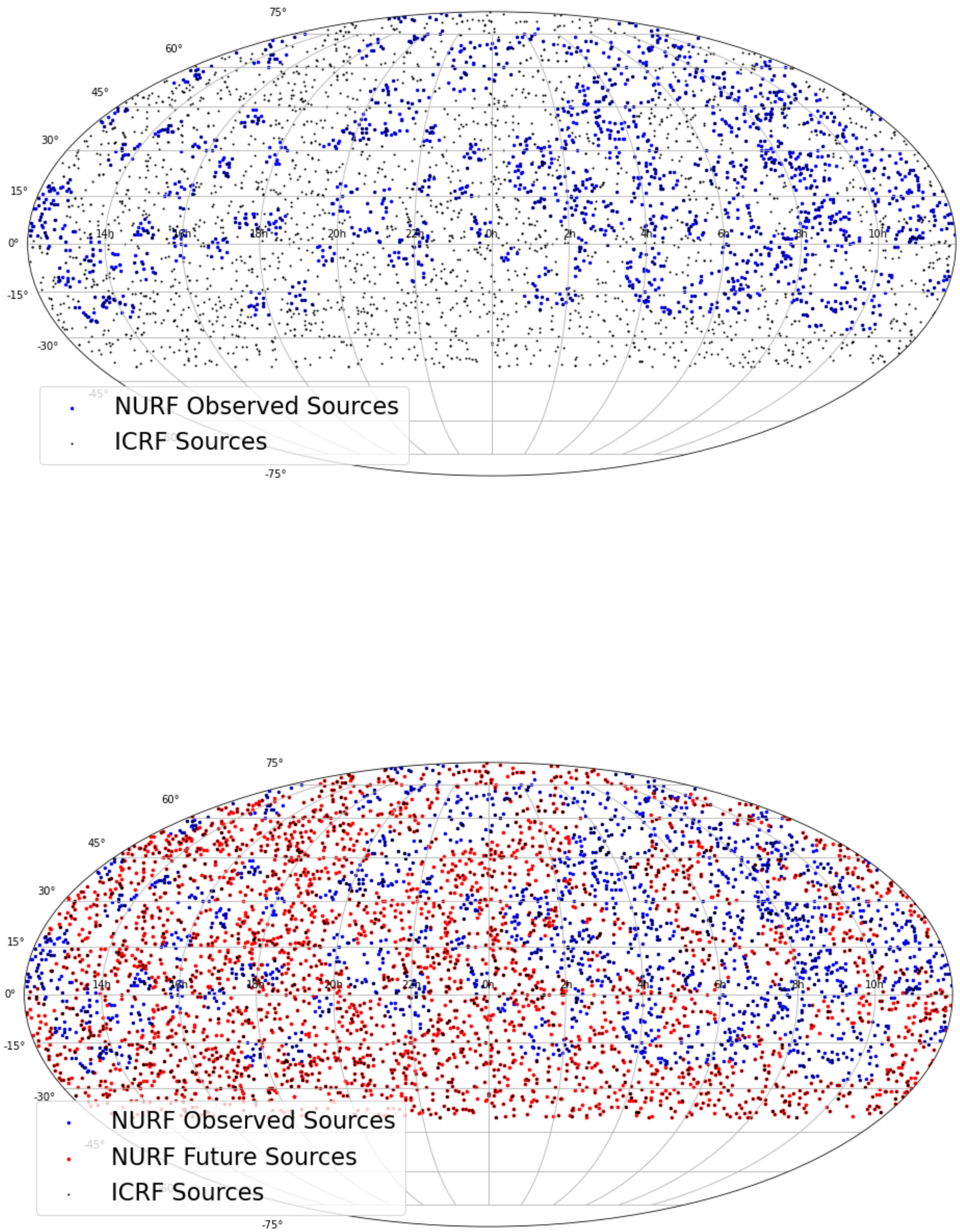

Figure 2. The top panel shows ICRF coverage in black and the observed NURF sources in blue, while the bottom panel shows the former along with the future NURF observations in red.

from a given plot which shows amplitude over the entire observation time with each source labeled, with the goal of allowing the user to select a bright source to use to calibrate instrumental delay and determine bandpasses .

Once the fringe finder source is selected, the script calls AIPS using unique AIPS numbers for each band, to prevent any potential confusion in the system. Once AIPS is called, the first 10 seconds of each scan are flagged to allow for slew time. The certified AIPS VLBA pipeline, VLBARUN (Sjouwerman et al. 2005) is then called to carry out the necessary data calibration and to separate the data out by source. After this, fluxes are obtained for each source by creating images of the sources using IMAGR, auto-cleaning the images and then fitting the flux using the IMFIT task. This outputs the peak flux in the image field for each source in each band and a signal to noise ratio for each calculation. Another procedure which fits fluxes using UV data is also included at this point in the pipeline, which will be covered in more detail in a later section. The results of the fitting procedures are written out into text files which can then be read back into the database to access later.

2.2.1. *Pipeline Considerations*

Various difficulties were considered in the creation of the survey, many of which were particularly relevant in constructing the pipeline. In these observations there are no planned phase reference sources, which does not allow for positional information to be obtained from any of the sources observed. However, the NURF survey plans to follow up any detections with phase reference observations to determine accurate positions. Other considerations included how to handle flagging and multi-band delay for these data, creating automated processes for imaging and model fitting, and handling duplicate sources in the data base.

3. IMPROVEMENTS

In response to the difficulties considered in the creation of the pipeline, some improvements have had to be made in various aspects from the methods for obtaining source flux and storing outputs in an accessible way.

3.1. *UV Model Fitting*

A new procedure was created to make use of the UVFIT task in AIPS, which fits a gaussian to the UV data for a given source and outputs a flux, error, and RMS. This procedure was integrated in to the pipeline to be used along with the image fitting task to allow for more certainty in detections.

UV model-fitting was added due to the relatively short time on source (less than a minute) and few baselines, resulting in quite sparse UV coverage. This leads to a very poor point spread function in the dirty beam which, in turn, makes imaging and fitting difficult. By fitting using the UV data we stand a better chance of being able to extract fluxes from sources.

3.2. *Database*

Initially, the database was created solely as a way of generating schedules automatically by tracking observations. However it was determined that the outputs from the pipeline could also be stored in this format, making output data easier to access. The existing database containing all the target sources along with positions and source catalogue information required several edits to allow for storing of output data. This makes it possible to interact with the results of the observations, and display information publicly once a front end is developed. To this end, a method for removing duplicate sources, and a script for parsing the information from both fitting procedures and displaying it within the database was created.

It was necessary to remove duplicate sources as we had many listed twice in our database, leading to the same source being observed multiple times, and resulting in issues within the pipeline related to calibrating and separating out these duplicate sources. In response to this we removed all sources within two arcseconds of each other with a preference for removing sources with less observations. These duplicate sources are due to the positional accuracy limits of Gaia and VLASS, and the fact our targets were selected from multiple catalogues.

Additionally, the database needed to be updated to support storing the outputs from the IMFIT and UVFIT procedures. The flux densities, errors, snrs, and names from both the image-plane and UV-plane fitting procedures are all written out into the database for the respective source in each band.

4. PRELIMINARY RESULTS

In order to form a basis for comparison of the results of the UVFIT procedure and IMAGR flux results, data for 44 sources from three different epochs (DL, DJ, and EJ) were calibrated and imaged. These data were calibrated

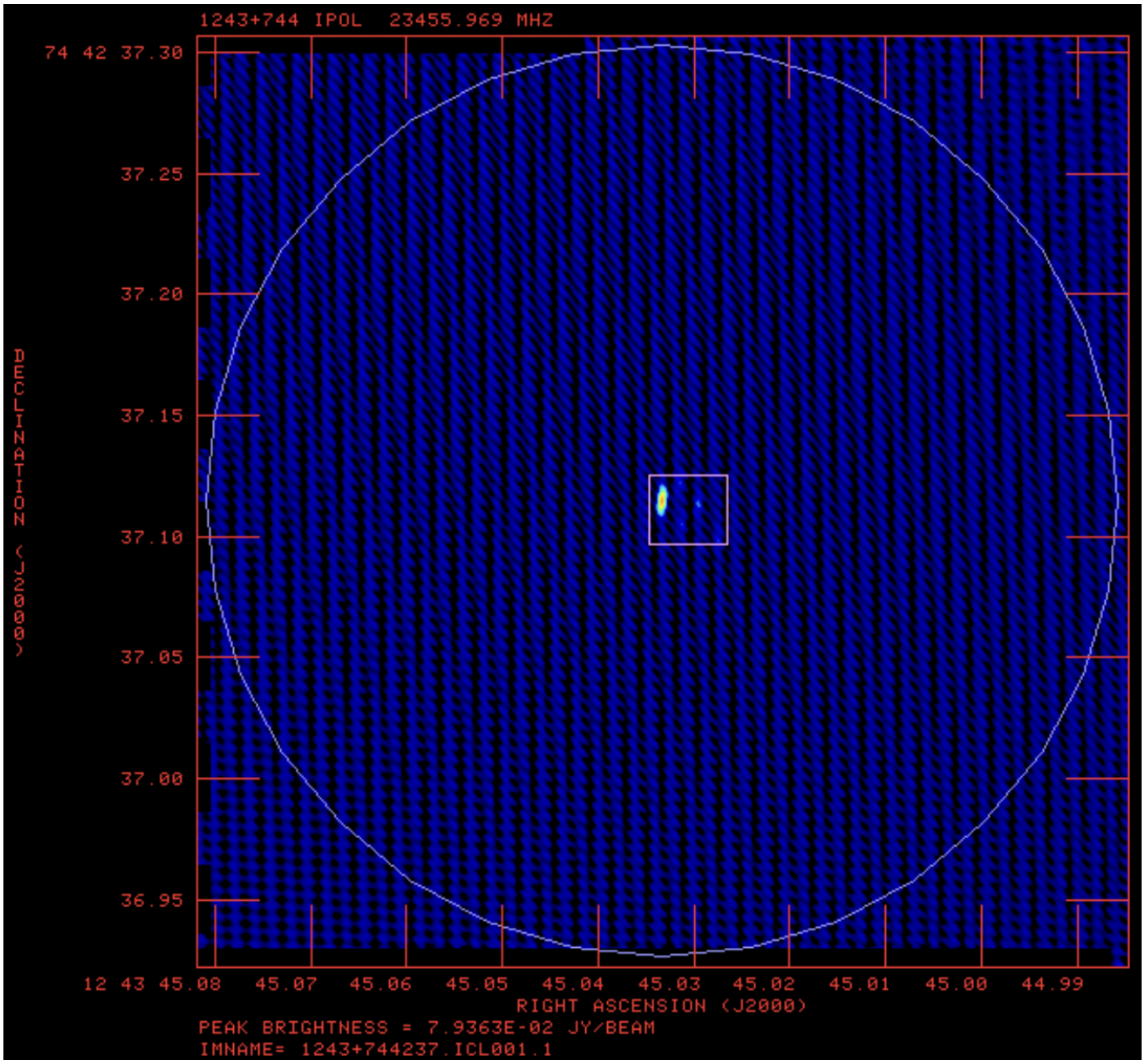

Figure 3. This figure shows a cleaned image in the user interface of AIPS of source 1243+744237. The image was created manually using IMAGR in AIPS. The bottom of the image shows the peak calculated flux density for the source to be 79.36 mJy/Beam. This result can be compared to the output of the UVFIT function and IMAGR. IMAGR was not able to detect this source, and UVFIT returned a flux density of 36 +/- 1 mJy.

using procedures from VLBARUN to correct for delay, fringe fit, and minimal flagging was also done on a case by case basis. From there, the sources were imaged and cleaned using IMAGR with boxes drawn around the approximate image center. An example of a cleaned image and peak flux calculation can be seen in Fig. 3. In reporting the fluxes calculated from the manual imaging, the total clean flux and peak flux were reported for each source.

Of the 44 sources manually imaged, 21 resulted in a detection. Of these same sources ran through the pipeline, IMAGR and UVFIT reported 6 and 21 detections, respectively. A detection for the UVFIT procedure was considered to be a flux greater than 5 sigma, using the error output by UVFIT which is calculated from the Gaussian fit. The results of the manual imaging allowed us to determine if the flux calculations and detection rate for the pipeline was reasonable and in agreement with previously determined flux densities.

In total, 130 sources were ran through the pipeline across 10 epochs. Of these 130 sources, IMAGR detected 20, and UVFIT detected 65. This yields an overall 5 sigma detection rate of 50% for the UVFIT procedure and 15% for IMAGR, see Table **??**. Due to the UVFIT error calculation being model dependent and not based on physical variables, this is likely an over estimate of the actual detection rate. However, this still is a promising result for the possibility of using UV fitting to detect observed sources. As all detections will be followed up with phase referenced observations to obtain good positions, leaning towards overestimating detections is the safer option.

Table 1. This table reports the number of detections and 5 sigma detection rates for UVFIT, IMAGR, and manual imaging. Only 44 sources were imaged manually, while 130 total sources were ran through the pipeline.

Model type	**Detections (3 epochs)**	**Detections (10 epochs)**	**Detection Rate**
UVFIT	21	65	50%
IMAGR	6	20	15%
Manual imaging	21	no data	no data

5. CONCLUSIONS

The NURF Survey offers an efficient way to increase the available reference and calibrator sources at higher frequency for VLBI observations. With the pipeline improvements showing promising detection rates for the UV fitting procedure, and future improvements under way for the imaging procedure, source selection, and scheduling, the NURF Survey will result in the detection of many more sources in K band, which will be stored in the database for ease of follow up in future observations.

ACKNOWLEDGEMENTS

I would like to thank the staff at NRAO and USNO for their support and for allowing me to be a part of this collaboration. I would like to extend an extra special thank you to Jay Blanchard and Lorant Sjouwerman for their advice and guidance during my time at NRAO.

REFERENCES

Gaia Collaboration, Vallenari, A., Brown, A. G. A., et al. 2022, arXiv e-prints, arXiv:2208.00211. https://arxiv.org/abs/2208.00211

Lacy, M. 2019, in American Astronomical Society Meeting Abstracts, Vol. 233, American Astronomical Society Meeting Abstracts #233, 424.01

Ma, C., Arias, E. F., Eubanks, T. M., et al. 1997, IERS Technical Note, 23, II

Sjouwerman, L. O., Mioduszewski, A. J., & Greisen, E. W. 2005, in Astronomical Society of the Pacific Conference Series, Vol. 340, Future Directions in High Resolution Astronomy, ed. J. Romney & M. Reid, 613