A FLUX SCALE FOR SOUTHERN HEMISPHERE 21 cm EPOCH OF REIONIZATION EXPERIMENTS

Daniel C. Jacobs1, Aaron R. Parsons2,3, James E. Aguirre2, Zaki Ali2, Judd Bowman4, Richard F. Bradley5,6,7, Chris L. Carilli8, David R. DeBoer3, Matthew R. Dexter3, Nicole E. Gugliucci9, Pat Klima6, Dave H. E. MacMahon3, Jason R. Manley9, David F. Moore4, Jonathan C. Pober2, Irina I. Stefan10, and William P. Walbrugh9

1 School of Earth and Space Exploration, Arizona State University, Tempe, AZ, USA
2 Astronomy Department, University of California, Berkeley, CA, USA
3 Radio Astronomy Lab., University of California, Berkeley, CA, USA
4 Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA, USA
5 Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, VA, USA
6 National Radio Astronomy Observatory, Charlottesville, VA, USA
7 Department of Astronomy, University of Virginia, Charlottesville, VA, USA
8 National Radio Astronomy Observatory, Socorro, NM, USA
9 Square Kilometer Array, South Africa Project. Cape Town, South Africa
10 Cavendish Laboratory, Cambridge, UK

Received 2013 June 20; accepted 2013 August 15; published 2013 October 4

ABSTRACT

We present a catalog of spectral measurements covering a 100–200 MHz band for 32 sources, derived from observations with a 64 antenna deployment of the Donald C. Backer Precision Array for Probing the Epoch of Reionization (PAPER) in South Africa. For transit telescopes such as PAPER, calibration of the primary beam is a difficult endeavor and errors in this calibration are a major source of error in the determination of source spectra. In order to decrease our reliance on an accurate beam calibration, we focus on calibrating sources in a narrow declination range from $-46^\circ$ to $-40^\circ$. Since sources at similar declinations follow nearly identical paths through the primary beam, this restriction greatly reduces errors associated with beam calibration, yielding a dramatic improvement in the accuracy of derived source spectra. Extrapolating from higher frequency catalogs, we derive the flux scale using a Monte Carlo fit across multiple sources that includes uncertainty from both catalog and measurement errors. Fitting spectral models to catalog data and these new PAPER measurements, we derive new flux models for Pictor A and 31 other sources at nearby declinations; 90% are found to confirm and refine a power-law model for flux density. Of particular importance is the new Pictor A flux model, which is accurate to 1.4% and shows that between 100 MHz and 2 GHz, in contrast with previous models, the spectrum of Pictor A is consistent with a single power law given by a flux at 150 MHz of $382 \pm 5.4$ Jy and a spectral index of $-0.76 \pm 0.01$. This accuracy represents an order of magnitude improvement over previous measurements in this band and is limited by the uncertainty in the catalog measurements used to estimate the absolute flux scale. The simplicity and improved accuracy of Pictor A’s spectrum make it an excellent calibrator in a band important for experiments seeking to measure 21 cm emission from the epoch of reionization.

Key words: dark ages, reionization, first stars – catalogs – instrumentation: interferometers

Online-only material: color figures

1. INTRODUCTION

Numerous radio telescopes are now exploring the prospects of using measurements of highly redshifted 21 cm emission to inform our understanding of cosmic reionization in the redshift range $z > 6$, corresponding to radio frequencies below 200 MHz (see reviews in Furlanetto et al. 2006; Morales & Wyithe 2010; Pritchard & Loeb 2012). These facilities include telescopes aiming to measure the global temperature change of 21 cm emission during the epoch of reionization (EoR), such as the Compact Reionization Experiment, the Zero-spacing Interferometer (Raghu et al. 2011), and the Experiment to Detect the Global EoR Signature (Bowman & Rogers 2010). Interferometers are also aiming to measure the power spectrum of 21 cm EoR emission, including the Giant Metre-wave Radio Telescope (GMRT; Paciga et al. 2011),11 the LOw Frequency ARray (LOFAR; Yatawatta et al. 2013; van Haarlem et al. 2013),12 the Murchison Widefield Array (MWA; Bowman et al. 2013; Tingay et al. 2013),13 and the Donald C. Backer Precision Array for Probing the Epoch of Reionization (PAPER; Parsons et al. 2010; Pober et al. 2013).14

Given the immense science potential in detecting 21 cm emission from the EoR, a great deal of research has focused on measuring the spectral and spatial variation of foreground emission in the 100–200 MHz band ($z = 6–13$ in the 21 cm line), which dwarfs the 21 cm signal by orders of magnitude (Furlanetto et al. 2006). In particular, the spectral properties of extragalactic point sources are important both because they are valuable calibration references and because they are strong foreground emitters that must be removed from 21 cm EoR measurements. With the scarcity of measured foreground properties in the 100–200-MHz frequency band over large areas of the sky

11 http://gmrt.ncra.tifr.res.in/
12 http://www.lofar.org/
13 http://www.mwatelescope.org/
14 http://eor.berkeley.edu/
(de Oliveira-Costa et al. 2008), continued foreground characterization is a vital step en route to any 21 cm EoR detection. At these low frequencies, the southern sky is much less well known than the north; catalog source fluxes at 150 MHz are inaccurate at the 20% level for $\delta < -20^\circ$ (Slee 1995; Vollmer et al. 2005). Both PAPER and the MWA are located in the southern hemisphere at radio-quiet reserves being prepared for the upcoming Square Kilometre Array. Therefore, the most extensive surveying work is now being conducted by the EoR experiments themselves (Jacobs et al. 2011; Williams et al. 2012; Bernardi et al. 2013).

One significant complication to improving the state of affairs in foreground characterization is that many 21 cm EoR experiments, including LOFAR in the northern hemisphere and PAPER and MWA in the southern hemisphere, are designed for drift-scan observations or steered via phased array. This design decision has largely been driven by the simplicity and cost effectiveness of phased and/or correlated dipoles to achieve the aggressive sensitivity requirements for measuring the 21 cm power spectrum of reionization (Parsons et al. 2012; Beardsley et al. 2013; Jelić et al. 2008). Adding to the challenge, these telescopes cover much wider fields of view ($>10^\circ$) and bandwidths ($\sim 100\%$ fractional) than traditional dish telescopes. Because they do not physically point, flux calibration for such arrays relies heavily on an accurate model of the primary beam response to correct for the apparent flux scale that varies across the sky. This direction dependent gain is currently uncertain to 10% or more (Poher et al. 2012) and comprises a large fraction of the 20% flux uncertainty between current telescopes (Jacobs et al. 2013).

In this paper, we set out to significantly improve the accuracy of spectral measurements between 100 and 200 MHz for a set of bright sources in the declination range $-46^\circ$ to $-40^\circ$ that are of particular value for southern hemisphere 21 cm EoR experiments such as PAPER and the MWA. Using the fact that, for this restricted declination range, sources transit through a nearly identical primary beam response pattern, we are able to avoid one of the most debilitating sources of error in these measurements: the primary beam.

In Section 2, we provide some background on uncertainty in early EoR-band catalogs and explain our choice of calibrators. In Section 3, we describe our approach to measuring source spectra with drift-scan observations and deriving an absolute flux scale from catalog data. In Section 4, we detail the instrumental setup, observations, and analysis method followed. Sections 4.6 to 4.8 detail our approach to fitting a global flux scale and spectral models for each source. We use these fits in Section 5 to understand how well the PAPER data agree with previous measurements and we conclude in Section 6.

2. BACKGROUND

Historically, the best southern hemisphere EoR band data were obtained by Slee (1995) with the Culgoora Circular Array and various higher frequency measurements with Parkes. These data are typically uncertain to 20% or more and provide little coverage of the EoR band beyond a single narrow-band data point. More recent surveys include narrowband surveys by the GMRT and Mauritius (Pandey & Shankar 2005), a deep survey of the region near Hydra A by the 32 antenna MWA prototype (Williams et al. 2012), and a widefield survey by PAPER, also with 32 elements (Jacobs et al. 2011). Several sub-channels were provided in the Williams catalog, although with 60%–80% error bars—large compared with the 30% uncertainty on their wideband measurements. These sub-channels cover the band and spatial scales relevant to EoR measurements but are limited by the accuracy of the primary beam (Jacobs et al. 2013), as well as by the lack of precise in-band flux calibrators.

The response of the primary beam is of critical importance to EoR measurements. Differences between the polarization responses cause leakage of polarized signals into the total intensity measurement, possibly corrupting the EoR power spectrum (Moore et al. 2013). The primary beam shape is also critical to measuring and subtracting foregrounds. (Bernardi et al. 2013; Sullivan et al. 2012; Morales et al. 2012) a process which, to be effective, must be done to better than 1% precision for the brightest sources; Liu et al. 2009; Bowman et al. 2009. A method for decoupling uncertain fluxes from the uncertain beam has been described by Poher et al. (2012). In simulations, the method was able to achieve $3\%$–$10\%$ accuracy in measuring the primary beam, depending on the number of antennas and other variables, emphasizing the need for many repeated measurements of each alt-az pointing (which were found by assuming $180^\circ$ symmetry). A further investigation is under way to improve and implement this method; this investigation would be greatly aided by precise flux measurements unaffected by the primary beam uncertainty.

EoR measurements from PAPER and the MWA in the southern hemisphere have focused on the coldest regions where Galactic foregrounds are minimal, with the majority of possible observing time falling around R.A. $= 4$ hr, decl. $= -30$. The brightest and least-resolved calibrator in this region is Pictor A ($5^h 19^m 49^s 1, -45^\circ 46^\prime 45^\prime\prime$). Pictor A is a nearby Fanaroff–Riley type II radio galaxy similar to Cygnus A. At $\sim 400$ Jy, Pictor A is bright and sufficiently distant from other bright sources to make it eminently suitable as both a phase and flux calibrator. Its apparent size of $\sim 8^\prime$ is smaller than the scales being probed by current EoR instruments, making it suitable for precision calibration with only a modest level of resolution effects. However, like most other sources, precise flux measurements in the EoR band are not available. The previous best EoR band measurement is uncertain to 12% and appears to imply spectral flattening in the EoR band (Perley et al. 1997).

Establishing an accurate spectrum for Pictor A is of particular importance for PAPER—a dedicated EoR experiment that employs drift-scanning, dual-polarization dipole antennas tuned for efficient operation over a 120–170 MHz band. PAPER is located in the South African Karoo desert on the Square Kilometer Array South Africa reserve, 100 km north of the small town of Carnarvon. The PAPER array has grown from 16 elements deployed in early 2009 to a 64 element imaging array in 2011 (see Figure 2). Since 2011 November, it has been arranged in a maximally redundant grid configuration to make deep power spectral integrations (Parsons et al. 2012). Though highly sensitive as a power spectrum instrument, the maximally redundant array has a broad point-spread function in the image domain, severely limiting the number of sources that can be used for flux calibration. Drift scanning across the sky with a 45$^\circ$ FWHM primary beam, there are very few unresolved, bright sources that are far from the Galactic plane. Pictor A is bright and well enough separated from other emission to dominate the visibilities for a good fraction of the EoR observing season, making it a desirable source to use for flux calibration.

15 Known during daylight hours as Culgoora Radio Heliograph.
16 http://tgss.ncra.tifr.res.in/
3. APPROACH

In this section, we describe our general approach to controlling the impact of beam model errors on measured spectra derived from drift-scan observations of Pictor A and a selection of known, bright sources. Our approach uses a set of “source tracks” as a function of frequency and time. Each source track is a beam, formed by phasing measured visibilities toward known source locations as they drift through the primary beam and summing over antenna pairs, as described in Section 4.3. Since sources fall at different positions within the primary beam, it is generally not possible to relate the fluxes of sources at different positions to one another without an accurate beam model. To date, the accuracy of source flux measurements in the southern hemisphere in the 100–200 MHz band have largely been limited by the accuracy of these beam models.

We mitigate this problem by selecting sources within a narrow declination range, so that source tracks represent nearly identical cuts through the beam response pattern. Using this fact, the relative amplitudes of sources can be deduced with minimal reliance on an accurate beam model. In our analysis, a prior beam model is only necessary for extrapolating across a narrow declination range (we estimate the errors associated with this extrapolation in Section 4.7) and for approximating optimal inverse-variance weighting when averaging over source tracks to determine a single spectrum.

This approach addresses several sources of error that were identified in Jacobs et al. (2013) as originating from uncertainty in the primary beam model. First, under most imaging schemes, each source flux is measured at just a few points in the primary beam. As errors tend to vary across the beam, it is often difficult to decouple flux uncertainty from beam uncertainty. Using the drift-scan beam-forming technique, each source is measured thousands of times at a variety of different primary beam values to provide a complete sample of the flux variance due to primary beam variation. Second, the flux calibration was found to vary over the sky due to the dual uncertainties of primary beam response and prior catalog. To reduce our exposure to flux calibration variation, we limit our observations to sources passing within 5° of pointings that we can directly calibrate to a single reference source. A third limitation, identified in Williams et al. (2012), was the increase of uncertainty toward the edge of the beam. To minimize our sensitivity to both beam and noise uncertainty, we weight the source track by an additional factor of the primary beam model.

The remainder of the analysis in this paper relates to calibrating the relative source amplitudes that we establish as an absolute flux scale as a function of frequency. Given the shortcomings of existing in-band measurements of sources in the southern hemisphere, we must bootstrap this flux scale from an ensemble of calibrator sources that exhibit simple power-law southerm declination, with a maximum of 10% over a 5° declination range.

3.1. Source Selection

As PAPER is a drift-scan instrument, each declination describes a distinct path through the primary beam. The flux time series of a beam-form provides a detailed sample of the primary beam relative to the peak of the trace; however, a model is still required to calibrate between declinations. Below, we argue that by averaging over long tracks, we minimize our susceptibility to localized primary beam error and find that limiting our source selection to within a declination range of 5° around a calibrator is a good compromise. In Section 4.7, we estimate the error resulting from this limited scope application of the primary beam model.

We choose sources from the Molonglo Reference Catalog (MRC; Large et al. 1981) that are within 5° in declination of a common calibrator (J2331−416) and more than 10° from the Galactic plane and Centaurus A. We limit the flux extrapolated from 408 MHz to be greater than 10 Jy, assuming a power-law spectral index of −1.37 This selection isolates 32 sources in a narrow stripe that passes through the majority of the southern EoR fields. See Figure 1 for a map of the sources relative to the Galactic plane and other structures mapped by PAPER using the same observations presented here. See Table 1 for a complete listing of names and positions.

3.2. Analysis Overview

The spectra reported here are measured by beam-forming—phasing the visibilities to the target location and summing over baselines to produce a spectral time series of “perceived” flux $S_p$. We define a perceived flux as the true source flux density, $S$, illuminated by the true beam pattern $B$:

$$S_p = B(t)S,$$

where $B$ is a function of time because the observations are taken as a drift scan. The desired source spectrum is then isolated from neighboring sources by delay filtering (Parsons & Backer 2009). We then average the spectrum in the time domain, weighting by a model of PAPER’s primary beam response ($B_M$); as the perceived flux is already weighted once by the true beam $B$, this weighted sum is an effective approximation of inverse-variance weighting (Pober et al. 2012). Mathematically, our estimated source flux density is given by

$$S_{est} = \frac{\sum B(t_M)S_p}{\sum B(t_M^2)} = \frac{\sum B(t_M)B(t)}{\sum B(t_M^2)}S \equiv g(\delta)S,$$

where the subscript $M$ indicates a model beam and $S$ without a subscript is the true flux density of the source. The net time integrated, weighted beam response $g$ is purely a function of declination $\delta$. The rate at which this factor changes with declination defines the declination range over which our calibrator source can be used to set the gain scale of the observations. To simulate this effect, we multiplied a measurement of the primary beam (Pober et al. 2012) by electromagnetic simulations (Parsons et al. 2010) and find that $g$ diverges slowly with declination, with a maximum of 10% over a 5° declination range. Thus, to first order, we can use our calibrator source to calibrate sources within 5° of declination and remove most of the variation with declination. We estimate this spectral dependent $g$ at a reference declination of 41°36′ using J2331−416, fitting a spectral index model to the catalog data points below 2 GHz. This

17 Most radio sources in this band have power-law spectra $S(\nu) = (\nu/\nu_0)^\alpha$ and a typical spectral index of $\alpha = −1$. 

3
Figure 1. PAPER map of the radio sky centered on the south pole showing relative positions of targeted sources. The map was made using the July half of the data used in this catalog and the procedure described in Jacobs et al. (2011) to produce a lightly CLEANed map. The circular artifact around the south celestial pole is due to residual instrumental cross-talk and appears to be highly localized below declinations of 75° where PAPER's sensitivity is low. "×"s mark measured source locations and have sizes scaled according to disagreement with previous data (inverse "improvement index" from Section 5); a smaller size indicates a better agreement with previous data. Orange dots indicate sources with no model overlap at 76% in Figure 14. (A color version of this figure is available in the online journal.)

Table 1

| Name          | R.A.  | Decl. | S125  | S135  | S145  | S155  | S165  |
|---------------|-------|-------|-------|-------|-------|-------|-------|
|               | (°)   | (°)   | (Jy)  | (Jy)  | (Jy)  | (Jy)  | (Jy)  |
| Pictor A      | 80.09 | −45.78| 455.3 | 13.3  | 409.6 | 15.8  | 389.2 |
| 0003—428      | 1.68  | −42.5 | 14.5  | 1.7   | 13.1  | 1.4   | 12.8  |
| 0007—446      | 2.8   | −44.3 | 23.2  | 2.2   | 21.1  | 2.2   | 20.2  |
| 0008—421      | 2.89  | −41.81| −0.2  | 0.8   | −0.3  | 1.0   | 0.9   |
| 0039—445      | 10.7  | −44.16| 29.2  | 3.0   | 27.6  | 2.2   | 26.4  |
| 0043—424      | 11.73 | −42.05| 48.6  | 4.2   | 45.2  | 3.5   | 43.9  |
| 0048—447      | 12.87 | −44.4 | 16.2  | 1.7   | 14.6  | 1.3   | 13.7  |
| 0049—433      | 13.22 | −43.03| 20.9  | 2.5   | 19.0  | 2.2   | 20.8  |
| 0103—453      | 16.48 | −45.02| 31.3  | 3.0   | 29.0  | 2.7   | 27.1  |
| 0201—440      | 31.05 | −43.76| 13.6  | 2.2   | 12.6  | 1.5   | 11.7  |
| 0438—436      | 70.17 | −43.53| 9.5   | 1.4   | 6.0   | 0.7   | 6.6   |

technique levels the bandpass due to spectral dependence of the primary beam and allows us to place our measured source flux estimates on approximately the same flux scale. Finally, we average the spectra from a resolution of 400 kHz to 10 MHz bins, adding the individual errors in quadrature for our uncertainty.

At this point, we have a set of spectra that have been calibrated relative to each other to the best of our ability. Now, they must be put on a global flux scale. Meanwhile, our error estimate does not include an estimate of the flux calibration error. To address both points, we fit for a global flux scale that incorporates the best available catalog data. These catalogs have all been set, as best as possible, to the Baars et al. (1977) scale, to which all other points are tied. Using a Bayesian analysis, we compute the variation in the flux scale implied by comparing many sources to prior
catalog models. By including sources from across the R.A./decl.
range, the variation in this flux scale fit estimates the overall
uncertainty in flux resulting from primary beam or prior catalog
uncertainty and folds in the error due to extrapolation beyond
the frequency range of the Baars scale, which is technically
limited to 300 MHz and above.

4. DATA REDUCTION

Measurements are derived from observations using two nights
on the east–west dipole arms of 64 PAPER antennas deployed
in a minimum-redundancy imaging configuration (see Figure 2) at the Karoo Radio Observatory site in South Africa.
A 100–200 MHz band was correlated with 2048 frequency
channels and integrated for 10.7 s before visibilities were
stored. Observations included here were conducted on 2011
July 4, running from JD2455748.17 to JD2455748.72 and 2011
October 17, running from JD2455852.2 to JD2455852.6. In both
seasons, two antennas were omitted owing to a malfunctioning
signal path. This 20 hr long combined data set provides near complete hour angle coverage for the entire
24 hr of right ascension. The July portion of the data set comes
from the same observing campaign described by Pober et al.
(2013) and Stefan et al. (2013).

4.1. Data Compression

In pre-processing, we use delay/delay-rate (DDR) filters
(Parsons & Backer 2009) to identify radio-frequency interfer-
ence (RFI) events as part of a data compression technique that
reduces the data volume by over a factor of 40. A more
detailed description may be found in Appendix A of Parsons et al.
(2013), where these authors use the same compression method
on PAPER power spectrum observations. Here, we summarize
the process.

First, we remove known RFI transmission bands and analog
filter edges and then we flag outliers in the frequency and time
derivative at $4\sigma$ to remove RFI events. Next, we suppress sky-
like emission by applying a DDR filter to remove delays and
delay rates within the horizon limit of a 300 m baseline
(the maximum length of any PAPER baseline). We derive a second
set of RFI flags by masking $4\sigma$ outliers in these residuals and apply these flags back to the unfiltered data. Finally, we compress the data by applying a DDR filter to preserve emission
within the horizon limit of a 300 m baseline, deconvolve to
suppress flagging artifacts, and down sample the result in the
time/frequency domain to the critical Nyquist rate of the DDR
filter. The result is that the original 2048 frequency channels
become 203 and the 60 original time samples per 10 minute file
become 14.

4.2. Per-antenna Gain Calibration

Small time- and antenna-dependent gain variations with time
introduce significant systematic effects that degrade instrument
performance. Temporal variations are dominated by changes
in ambient temperature. We use measurements of ambient
temperature versus time on the balun amplifier of a fiducial
antenna and near receiver amplifiers to divide out the predicted
gain variation versus temperature that these amplifiers are
known to exhibit. Inspecting the auto correlations, we
find that the two are highly correlated with a best-fit temperature
autocorrelations to their temperature difference (Figure 3), we
use the October observations, when the temperatures remained
nearly constant, as a reference.

Comparing the average ratio between the July and October
autocorrelations to their temperature difference (Figure 3), we
find that the two are highly correlated with a best-fit temperature
coefficient of $(-0.058 \text{ dB C}^{-1})$. As we see in Figure 4,
this technique removes a significant amount of disagreement
between the two nights, with the peak difference decreasing
from 18% to 3%.

Matching of relative gains and phases between antennas was
done by fitting a per-antenna complex gain to portions of data
that have an easily modeled sky background. This calibration is
only relative between antennas; bandpass and absolute flux cali-
bration comes after time and frequency averaging in Sections 4.5

\begin{figure}
\centering
\includegraphics[width=0.8\textwidth]{figure2.png}
\caption{Antenna position in the 64 antenna, minimum-redundancy PAPER
array configuration at the Karoo Radio Observatory site in South Africa.
Data were obtained during 2011 July and October on a single "x" polarization
over a band between 120 and 180 MHz.}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=0.8\textwidth]{figure3.png}
\caption{Calibrating the temperature coefficient between the July and October
nights. Here, we plot the ratio of auto-correlations (averaged over frequency
and antenna) against the temperature difference between the two nights. The
very linear middle region has a slope of $-0.055 \text{ dB C}^{-1}$ while the full, more
complex range is fit by $-0.06 \text{ dB C}^{-1}$ (black line). We compromise by choosing
$-0.058 \text{ dB C}^{-1}$, which was found to significantly improve the agreement (see
Figure 4).}
\end{figure}
Figure 4. Gain differences between the July and October nights are dominated by temperature variation. Here, we see the effect on the average auto correlation before (orange) and after (black) application of the temperature-dependent gain. Note that the span of the data in the bottom fractional difference plot is limited to points occurring in both nights and does not cover the full range seen in the top plot.

(A color version of this figure is available in the online journal.)

and 4.7. The antenna delays and amplitudes were found by fitting a point source visibility model to Centaurus A, Pictor A, and Fornax A. Each source is imperfectly modeled by a single point source, but the solution differences are minimized by averaging over the three independent solutions. These same calibration solutions have been successfully applied in Pober et al. (2013) and Stefan et al. (2013), for power-spectrum analyses of foregrounds and imaging of Centaurus A, respectively.

The October per-antenna gains were found by computing the average ratio of the auto-correlations against the July data. Delays were held constant, as in Jacobs et al. (2011).

4.3. Beam-forming

Spectral time series are computed by beam-forming to the selected sky locations. A beam is formed by phasing baselines to the desired location and summing over baselines longer than 20 wavelengths.

These complex spectra are then fringe rate filtered to remove spectrally smooth sources that deviate by more than ±0.1 mHz from the fringe rate of the source in question (Parsons & Backer 2009; cf., the LOFAR de-mixing approach; Offringa et al. 2012), producing a time-dependent spectrum with minimal side lobes. This spectrum is then averaged over time, weighting by a model of the primary beam, as discussed above. The result is equivalent to a very long Earth rotation synthesis image with a single image pixel. The weighted contributions from each baseline (the $uv$ coverage) are shown in Figure 5. When combined with the filtering, this technique is a robust and simple method for measuring spectra of unresolved sources.

4.4. Compensation for Resolution Effects in Pictor A

The beam-forming method assumes the target is a point source. Our primary target, Pictor A, is slightly resolved by PAPER and merits closer attention. Pictor A is a double-lobed radio galaxy with a main lobe separation of 7'. As seen in Figure 5, it is nearly unresolved by PAPER’s 15' synthesized beam. However, given the high signal-to noise ratio of the observations, we see a 20% drop in flux on the longest ($\sim$300 m) baselines. We account for this fact by weighting baselines in the beam-form step according to a model of structure observed by Perley et al. (1997) at 330 MHz, which they found to be consistent with their more limited 74 MHz images as well as
detailed high-frequency maps. The normalized image is Fourier-transformed and sampled at the desired $uv$ spacing by spline interpolation. These samples are used to weight each baseline contribution in the baseline sum. The result is an estimate of the total integrated flux for Pictor A. Where the resolution is highest, at the top of the band, the net correction is 3%. At the bottom of the band, where the resolution size has grown to 19$''$, the correction is only 0.6%. The resulting spectrum is shown in Figure 8.

4.5. Bandpass Calibration

The resulting set of spectra were then calibrated with a model of J2331$-416$\textsuperscript{18} to remove the residual bandpass $g$ from Equation (2) caused by the net effect of the primary beam, as discussed in Section 3. This source was selected for its relatively high brightness, spectral smoothness, and large quantity of available catalog data. Each source track has a slightly different sample profile resulting from different amounts of flagged data at each time and channel. To account for these differences in the bandpass calibration, we build a set of calibration tracks that match the tracks for each source. Calibration of each source then proceeds with an optimally matched calibration spectrum.

4.6. Fitting a Power-law Model to the Spectra

There are a variety of previous data at multiple wavelengths that we want to use as calibration; we also want to compare our measurements with these data. Our method for accomplishing both of these steps is to assume a basic spectral model relating the different catalog data points and fit for spectral model and gain parameters.

4.7. Approximating the Absolute Flux Scale

At the output of the beam-forming and bandpass calibration step, the flux scale is tied to a model fit of the catalog values of J2331$-416$. The accuracy of this fit, and the implied uncertainty in the flux scale, limits the accuracy of the PAPER measurements. To refine the flux scale and estimate our flux scale uncertainty, we bootstrap a single global flux scale correction factor using six sources selected for their brightness, spectral linearity, and data availability.\textsuperscript{19} To build a more complete spectral model, we go beyond the data found in Vollmer et al. (2010) and include all spectral measurements below 2 GHz found in the NASA/IPAC Extragalactic Database (NED).\textsuperscript{20} These additional catalog measurements are primarily by Parkes and Molonglo (Kuehr et al. 1981), with the best precision coming from the Wills fluxes at 538 and 634 MHz (Wills 1975). Where error bars are not given, we assume an uncertainty of 25%.

Using the MCMC method, we fit spectral index models to the catalog fluxes simultaneously with a global PAPER flux scale factor using an MCMC chain to calculate the logarithm of the

\textsuperscript{18} Here, we set the spectrum to $S_{50} = 33$ Jy, $\alpha = -0.76$, derived by a linear least-squares fit to catalog measurements below 1 GHz (Slee 1995; Kuehr et al. 1981; Large et al. 1981; Burgess & Hunstead 2006). Further calibration on top of this somewhat arbitrary calibration model is explored in detail below.

\textsuperscript{19} Calibration sources: 2250$-412$, 2331$-416$, 1932$-464$, 0103$-453$, 0547$-408$, and 0043$-424$.

\textsuperscript{20} ned.ipac.caltech.edu: accessed 2013 April 1.
Marginalized posterior of the PAPER flux scale factor relative to J2331−416, given a joint fit to all the NED database listings for 2250−412, 2331−416, 1932−464, 0103−453, 0547−408, and 0043−424 below 2 GHz and the corresponding PAPER values. The peak of the distribution sets the overall flux scale, while the 76% width is an additional fractional error to be added quadratically to the per-source uncertainty. By averaging over many sources and the Baars calibrated data points, this uncertainty encapsulates errors due to both primary beam model inaccuracy as well as the average flux scale uncertainty due to extrapolating many absolutely calibrated measurements.

\[
\log L_s = \sum \left[ \frac{S_{\nu} - S_{150} \left( \frac{\nu}{150} \right)^{\alpha}}{2(\Delta S_{\nu})^2} \right] + \frac{[gS_{\nu,PAPER} - S_{150} \left( \frac{\nu}{150} \right)^{\alpha}]}{2(\Delta S_{\nu,PAPER})^2} \right] \]

which samples the posterior probability of the PAPER data ($S_{\nu,PAPER}$) and catalog values ($S_{\nu,cat}$), given a spectral index model for each source and a global flux scale factor ($g$). Marginalizing over the fitted flux scales, we find the resulting flux scale distribution function, which is shown in Figure 7. The 76% confidence limit on this flux scale, relative to the J2331−413 calibration, is $+0.11^{+0.05}_{-0.08}$ dB, or a 1.54% multiplicative error on every calibrated PAPER measurement. This flux scale is applied to the PAPER spectra with the errors added in quadrature.

The resulting Pictor A spectrum, plotted in Figure 8, ranges in precision from 3.8% at 135 MHz to 2.6% at 165 MHz. The rest of the calibrated spectra are plotted in Figure 9 and listed in Table 1.

4.8. Fitting Spectral Models

Finally, we compare all of our calibrated spectra with catalog measurements. Our method will be to first establish a best-fit model for the existing data, then add the PAPER data to the fit, and then assess the degree to which the PAPER data are supported by prior measurements (accuracy). We also assess the degree to which PAPER improves our knowledge of the spectrum (precision).

To establish a baseline model, we fit a spectral model to prior catalog data from the spectrally and spatially cross-matched meta-catalog by Vollmer et al. (2010) using the MCMC chain to sample the (log) likelihood, which assumes Gaussian measurement errors:

\[
\log L = \sum_{\nu} \frac{[S_{\nu} - S_{150} \left( \frac{\nu}{150} \right)^{\alpha}]}{\Delta S_{\nu}}. \tag{4}
\]

As when fitting the flux scale, we estimate the confidence interval of the resulting parameters as the boundary enclosing 76% of the samples. A detailed view of the resulting posterior is shown in gray on Figure 10. Most of the posteriors are characterized by steep sided, symmetric distributions. While many appear to have linearly correlated parameters, several posteriors display the classic banana shape associated with a non-linear correlation between parameters. For the rest of the sources, we focus on the 76% confidence levels, plotting the two-dimensional distributions in Figures 11–15 and listing the marginalized values in Table 2.

The fit is then performed again with both the PAPER and catalog data, shown in black in the same figures.
Figure 9. PAPER spectra of 31 sources compared with existing data from Vollmer et al. (2010) between 40 MHz and 2 GHz and a sampling of spectral fits drawn from MCMC samples at greater than 76% confidence. Sources used to bootstrap the flux calibration are noted with a "*" and are shown with the additional calibration data found in NED. Pictor A is shown separately in Figure 8.

Figure 10. Detailed view of the posterior probability distribution, as described in Section 4.8. Pictor A is shown on the left and 2323$-407$ is shown on the right. Confidence contours range from 0.2 to 0.8; the curves to the top and to the right are the marginalized distributions. Gray lines indicate the fit to catalog data and black lines indicate the fit to the joint catalog and PAPER data. The MCMC chain reveals the distinctive "banana" shape of the posterior due to correlations between the model parameters. We encourage the reader to compare these contours with the contours for this source shown in Figure 12. The fit parameters listed in Table 2 are the locations where the marginalized curve crosses 0.76.
5. RESULTS AND DISCUSSION

The majority of the 32 models fit using the new PAPER data were found to agree with models fit to previous data and all but two are suitable for comparison. 0008−421 was not detected by previous PAPER observations and exhibits a flattening at higher frequencies, most likely due to synchrotron self-absorption (Jacobs et al. 2011). Meanwhile, 1459−417 does not have enough measurements in the SPECFIND catalog on which to build a model for comparison. We exclude these two objects from further analysis.

Inspecting the remaining 30 confidence contours, we find that, for the vast majority (90%), the PAPER measurements confirm the power-law extrapolation from higher frequencies. The remaining 10% either (1) have large PAPER error bars and therefore provide no new information or (2) do not agree with the spectral index model. To understand where most measurements fall on this spectrum, we numerically quantify the overall model improvement derived from the addition of the PAPER data as

$$\text{Improvement} = (\text{Precision increase})(\text{Catalog agreement})$$

$$= \left(\frac{1}{\text{Area}(\text{PAPER})} - \frac{1}{\text{Area}(\text{Cat})}\right)(\text{Area}(\text{PAPER} \cap \text{Cat})),$$

where the area is defined as the number of samples having a posterior probability above 76%.

Figure 11. Spectral model contours, as described in Figure 8. The catalog fit is shown in gray and the black lines fold in the PAPER data. Sources marked with a * were used to assess calibration error.
Figure 12. Spectral model contours, as described in Figure 8. The catalog fit is shown in gray and the black lines fold in the PAPER data. Sources marked with a * were used to assess calibration error.

Figure 13. Fits of the last three sources, as described in Figure 11. The catalog fit is shown in gray and the black lines fold in the PAPER data. Sources marked with a * were used to assess calibration error.
We quantify the fit precision increase as the change in the contour figure of merit, defined as the inverse area of the confidence contour. Meanwhile, catalog agreement is the fraction of the PAPER confidence interval that overlaps the catalog confidence interval. Thus, for example, in a PAPER fit that overlaps the catalog confidence contour by 41\% but increases in precision (confidence area shrinks) by a factor of three, the resulting improvement will be 0.123. This improvement index is included in the table of fit parameters (Table 2) and is used as the size scale on the map in Figure 1.

In these sources, the improvement index ranges from a maximum of 7.84 to −0.001. One source (1017−421) shows a slightly negative improvement, suggesting that PAPER data have added to the uncertainty (Figure 15), while two sources have exactly zero improvement, which indicates that the PAPER data have pulled the fit far from the model preferred fit.
by the catalog data (see Figure 14 for details on these sources).

However, the vast majority of sources (90%) have a positive improvement index, indicating strong confirmation of the extrapolated spectrum (see Figures 11–13). Pictor A is close to the top of this group with an improvement of 0.942—only two sources show stronger confirmation. The flux model is $S_{150} = 382 \pm 5.4$ Jy and $\alpha = -0.76 \pm 0.01$, a fractional error of 1.4%.

This fractional error represents an order of magnitude improvement in the accuracy of this primary flux calibrator. In fact, the uncertainty is so small that it is worth considering its validity. First, consider that we have fit the model to many data points, including seven measurements accurate to 3% (five PAPER, two Wills). If the errors were completely Gaussian, the net error would be $3\%/\sqrt{7} = 1.14\%$. Consider also that, where most surveys might include on the order of 10 independent measures of a bright source from different facets to arrive at a 5%–20% uncertainty, we have included thousands of independent measurements over the full horizon to horizon transit, while carefully controlling for systematic variation and resolution effects.

6. CONCLUSIONS

This paper presents a set of total flux measurements at multiple frequencies within a 100–200 MHz band, providing an absolute flux calibration for southern hemisphere EoR studies, as well as a modest set of verification fluxes for other sources suitable for constraining future primary beam studies. We have provided a measurement of Pictor A with enough precision to confirm a linear spectral index between 150 and 600 MHz. We apply the same filtered beam-forming method to measure the spectra of bright sources with similar primary beam responses.

The measurements provided here are the first calibrated, broadband spectra to cover the EoR band. Existing EoR band measurements are accurate to 20%, implying a 40% uncertainty in the absolute power spectrum level. The PAPER Pictor A spectrum is found to be accurate to $\sim3\%$, a factor of seven improvement over previous EoR band measurements.

This uncertainty includes the variation in each PAPER measurement ($\sim1\%$), variation between sources, and the errors resulting from extrapolating the Baars et al. (1977) scale beyond its original range. The last two sources of error are found simultaneously by fitting the spectral extrapolations of several flux calibrators at once; these sources account for about half of the error. Although previous measurements suggested the possibility of spectral curvature below 200 MHz, we have found no evidence supporting this hypothesis. With these measurements, we are able to confirm a single spectral index model for Pictor A between 120 MHz and 600 MHz.

A set of 31 additional verification sources were also targeted to provide additional characterization of the flux scale as well as an overall assessment of the catalog accuracy. Using a Bayesian analysis, we conclude that most of these sources are consistent with previous measurements, provide useful new constraints, and support the conclusion that the Pictor A flux is on the correct scale.

Direct measurements of the Pictor A spectrum are key to correctly setting the flux scale of PAPER, MWA, and future EoR experiments such as the Hydrogen Epoch of Reionization Array. These spectra provide tighter constraints on many of the EoR band fluxes, while limiting the pernicious effect of primary beam uncertainty. Future work will use these fluxes to further refine the primary beam models of these experiments, which is crucial for properly reconstructing both the image and the power spectrum flux. Although we have focused on a narrow declination range in this paper, the techniques we describe here may be applied over any declination range. Future work will also aim to extend this analysis to other declination ranges and to tie the flux calibration of these other declination ranges back to the absolute flux scale we derive here.

The PAPER project is supported by the National Science Foundation (awards 0804508, 1129258, and 1125558) and a generous grant from the Mt. Cuba Astronomical Association. This work makes use of the “MCMC Hammer” emcee Python library (Foreman-Mackey et al. 2012; http://danfm.ca/emcee/) and the NASA/IPAC Extragalactic Database (NED), which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

REFERENCES

Baars, J. W. M., Genzel, R., Pauliny-Toth, I. I. K., & Witzel, A. 1977, A&A, 69, 99
Beardsley, A., Hazelton, B. J., Morales, M. F., et al. 2013, MNRAS, 429, L5
Bernard, G., Williams, C., Hewitt, J., et al. 2013, ApJ, 771, 105
Bowman, J., Cairns, I., Kaplan, D. L., et al. 2013, PASA, 30, 31
Browne, I., Morales, M., & Hewitt, J. 2009, ApJ, 695, 183
Browne, J. D., & Rogers, A. E. E. 2010, Natur, 468, 796
Burgess, A. M., & Hunstead, R. W. 2006, AJ, 131, 100
de Oliveira-Costa, A., Tagmark, M., Gaensler, B. M., et al. 2008, MNRAS, 388, 247
Foreman-Mackey, D., Hogg, D., Lang, D., & Goodman, J. 2012, PASP, 125, 306
Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, PhR, 433, 181
Hogg, D. W., Bovy, J., & Lang, D. 2010, arXiv:1008.4866H
Jacobs, D., Aguirre, J. E., Parsons, A. R., et al. 2011, ApJL, 734, L34
Jacobs, D. C., Bowman, J., & Aguirre, J. E. 2013, ApJ, 760, 5
Jelić, V., Zaroubi, S., Labropoulos, P., et al. 2008, MNRAS, 389, 1319
Kuehr, H., Witzel, A., Pauliny-Toth, I., & Nauber, U. 1981, A&A, 145, 367
Large, M. I., Mills, B. Y., Little, A. G., Crawford, D. F., & Sutton, J. M. 1981, MNRAS, 194, 693
Liu, A., Tagmark, M., & Zaldarriaga, M. 2009, MNRAS, 394, 1575
Mackay, D. C. 2003, Information Theory, Inference and Learning Algorithms (Cambridge: Cambridge Univ. Press)
Moore, D., Aguirre, J., Parsons, A., Jacobs, D., & Pofer, J. 2013, ApJ, 769, 154
Morales, M. F., Hazelton, B., Sullivan, I., & Beardsley, A. 2012, ApJ, 752, 137
Morales, M. F., & Wyithe, J. S. B. 2010, ARA&A, 48, 127
Offringa, A. R., de Bruyn, A. G., & Zaroubi, S. 2012, MNRAS, 422, 563
Otrupcek, R., & Wright, A. 1991, PASAu, 9, 170
Paciga, G., Chang, T.-C., Gupta, Y., et al. 2011, \textit{MNRAS}, 413, 1174

Pandey, V., & Shankar, N. U. 2005, URSI General Assembly Proceedings, http://www.ursi.org/Proceedings/ProcGA05/pdf/J03b.5(01214).pdf

Parashare, C. R. 2011, PhD thesis, Univ. Virginia

Parsons, A., Backer, D. C., Foster, G. S., et al. 2010, \textit{AJ}, 139, 1468

Parsons, A., Pober, J., McQuinn, M., Jacobs, D., & Aguirre, J. 2012, \textit{ApJ}, 753, 81

Parsons, A. R., & Backer, D. C. 2009, \textit{AJ}, 138, 219

Parsons, A. R., Liu, A., Aguirre, J. E., et al. 2013, \textit{ApJ}, submitted (arXiv:1304.4991v2)

Perley, R. A., Roser, H.-J., & Meisenheimer, K. 1997, \textit{A&A}, 328, 12

Pober, J., Parsons, A. R., Aguirre, J. E., et al. 2013, \textit{ApJL}, 768, L36

Pober, J. C., Parsons, A. R., Jacobs, D. C., et al. 2012, \textit{AJ}, 143, 53

Press, W. 1997, in Unsolved Problems in Astrophysics, ed. J. N. Bahcall & J. P. Ostriker (Princeton, NJ: Princeton Univ. Press), 49

Pritchard, J. R., & Loeb, A. 2012, \textit{RPPh}, 75, 6901

Raghunathan, A., Patra, N., Shankar, N. U., Ekers, R. D., & Subrahmanyan, R. 2011, General Assembly and Scientific Symposium, 2011 XXXth URSI, http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6051278

Sivia, D., & Skilling, J. 2006, Data Analysis: A Bayesian Tutorial (Oxford Univ. Press)

Slee, O. 1995, \textit{AuJPh}, 48, 143

Stefan, I. I., Carilli, C. L., Green, D. A., et al. 2013, \textit{MNRAS}, 432, 1285

Sullivan, I. S., Morales, M. F., Hazelton, B. J., et al. 2012, \textit{ApJ}, 759, 17

Tingay, S., Goeke, R., Bowman, J. D., et al. 2013, \textit{PASA}, 30, 7

van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, \textit{A&A}, 556, A2

Vollmer, B., Gassmann, B., Derrière, S., et al. 2010, \textit{A&A}, 511, 53

Williams, C. L., Hewitt, J. N., Levine, A. M., et al. 2012, \textit{ApJ}, 755, 47

Wills, B. J. 1975, \textit{AuJPh}, 38, 1

Yatawatta, S., de Bruyn, A. G., Brentjens, M. A., et al. 2013, \textit{A&A}, 550, 136