A broadband flux scale for low frequency radio telescopes

Anna M. M. Scaife and George H. Heald

ABSTRACT

We present parameterized broadband spectral models valid at frequencies between 30-300 MHz for six bright radio sources selected from the 3C survey, spread in Right Ascension from 0 – 24 hours. For each source, data from the literature are compiled and tied to a common flux density scale. These data are then used to parameterize an analytic polynomial spectral calibration model. The optimal polynomial order in each case is determined using the ratio of the Bayesian evidence for the candidate models. Maximum likelihood parameter values for each model are presented, with associated errors, and the percentage error in each model as a function of frequency is derived. These spectral models are intended as an initial reference for science from the new generation of low frequency telescopes now coming on line, with particular emphasis on the Low Frequency Array (LOFAR).

Key words: Radiation continuum:general – methods:observational – methods:statistical

1 INTRODUCTION

In order to quantitatively combine and contrast data from independent telescopes and surveys, often at multiple frequencies, it is necessary to have a standard calibration scale to form comparisons. This is especially important at frequencies below 300 MHz and above 15 GHz where the widely used Baars et al. (1977) radio flux density scale is incomplete. For the new generation of low frequency telescopes such as the Low Frequency Array (LOFAR; van Haarlem et al. in prep) it is becoming increasingly necessary to provide a broadband spectral reference for initial science, so that both archival and future measurements can be quantitatively compared to these new data. In addition to an absolute scaling, such telescopes require a well-defined set of calibrators spread in right ascension (RA) to allow for quasi-simultaneous broadband calibration of field observations. Here we present a set of parameterized models for six broad multi-frequency calibrators covering frequencies from 30 – 300 MHz and RAs from 0 – 24 hrs. We focus on the northern sky, and in particular on the applicability to LOFAR. This set of calibrators forms a flux scale that will be the basis of a major effort to develop an all-sky, broadband calibration catalog. The initial description given here will be continuously refined as new LOFAR data accumulate.

2 CALIBRATION OF LOW FREQUENCY TELESCOPES

Radio interferometers operating at low frequencies face a substantial calibration challenge. Strong ionospheric phase corruptions are common, especially below 100 MHz. For large-scale survey work, it is important that the processing of raw visibility data from the telescope can be automated. In order to jump-start such an automatic calibration and imaging process for any arbitrary field, a pre-existing model of the brightest sources in the field of view is required. Such a model must be intrinsically frequency dependent, since modern radio telescopes are inherently broadband in nature, with tremendous fractional bandwidths. For example, LOFAR routinely observes from 30 – 240 MHz, and is capable of observing as low as 10 MHz. Over such a broad range, the flux scales must be tied to a well-understood set of reference sources with spectral energy distributions that are well understood across the full bandpass. In the case of LOFAR, the production of such an all-sky broadband catalog is the key goal of the Multifrequency Snapshot Sky Survey (MSSS; Heald et al. in prep).

The reference sources which form the basis of the broadband flux scale must be selected for suitability as high-quality calibration targets. Several factors are relevant. First, the source should dominate the visibility function. In addition to high flux density, separation of contaminating flux from sources away from the pointing centre (‘off-beam’) can be improved in two further ways (i) averaging in time and frequency to smear out the contributions of off-beam sources on longer baselines; and (ii) the “demixing” technique (van der Tol et al. 2007), which has been adopted for use with LOFAR data. Secondly, the source should be compact compared to the angular resolution of the instrument, to allow simple morphological calibration models. Well-known sources such as Cyg A and Cas A have extremely complex morphologies, making calibration of an array with arcsecond angular resolution difficult. Thirdly, these calibrators must be spread in right ascension (RA) to allow for quasi-simultaneous broadband calibration with field observations.

With these considerations in mind we searched the 3C (Edge et al. 1959) and revised 3C (3CR; Bennet et al. 1962) catalogues
for an initial list of bright compact sources, with the criteria that (1) they must be at declinations greater than 30°, (2) they must have a flux density at 178 MHz greater than 20 Jy and (3) they must have an angular diameter less than 20 arcseconds (compact compared to the naturally weighted resolution of the Dutch LOFAR array).

These criteria result in an initial sample of six sources, of which we exclude one based on other data from the literature showing more substantial extension than indicated in 3CR (3C69; Pooley & Henbest 1974) and we include one additional source based on other data from the literature indicating that the extension listed in 3CR is an overestimate (3C286; e.g. Pearson et al. 1985). The final sample is listed in Table 1. Source extensions from 3CR are listed in Column [5] for each object. We note that high resolution observations (e.g. Akujar & Garrington 1995) confirm that the source structure in 3C48 is on sub-arcsecond scales, whilst 3C147, 3C286 and 3C295 have structure on scales < 6″. 3C196 has two dominant components separated by about 6″, as well as complex diffuse structure with a (precessing) jet morphology (Reid et al. 1995). The structure in 3C380 is known to be dominant on scales of > 30″, making it the most extended object in this sample (Reid et al. 1995).

3 FLUX SCALES

The data used for spectrum fitting are listed in Table 2. In order to provide a common flux scaling, these data have been revised onto the flux scale of Roger, Bridle & Costain (1973; hereafter RBC) below 325 MHz. This scale has been chosen to avoid the suggested issues (e.g. Rees 1990a) with the secular decrease in the flux density of Cas A at low frequencies (< 100 MHz) inherent in the widely used Baars et al. (1977; B77) scale.

At low radio frequencies most data are tied to the RBC or Kellerman, Pauliny-Toth & Williams (1969; KPW) scale. The correction factors for moving between these scales at ν < 325 MHz are listed in Table 2. At ν > 325 MHz the RBC and KPW scales are in agreement and consequently such data, where calibrated on the B77 scale, are corrected using a polynomial fit to the correction factors listed in B77 onto the KPW scale. Data from WENSS (Rengelink et al. 1997) have been corrected using an average correction factor to bring them onto the B77 scale and a further scaling to bring them onto the RBC scale. The 6C, 8C and MIYUN surveys are calibrated on the RBC scale in their original form, and the Bologna survey (Colla et al. 1970) is calibrated on the KPW scale which is consistent with the RBC at 408 MHz. Data from Aslanyan et al. (1968) are scaled using the ratio of the stated flux densities for the calibrator sources (3C348 & 3C353) in the original paper to the predicted values at 60 MHz from the spectral models for these sources in RBC. Data from Scott & Shakeshaft (1971) are corrected onto the scale of Artyukh et al. (1969) and then onto the RBC scale using the factors listed in Tables III & IV of RBC, this is subject to the caveat that the difference in flux densities from 81.5 to 86 MHz is assumed to be negligible compared to the uncertainty in these factors (∼ 3 per cent). Where applied, the scaling factors in each case are listed in Table 2. The original flux densities for the sources from the 3C catalogue (Edge et al. 1959) have not been included in the model fitting. The large size of the errors associated to these data is such that they have no influence on the parameter estimation.

Additional data are available at 12.6 – 25 MHz from the UTR-1 telescope (Braude et al. 1970a,b), calibrated on the Gravoko scale. These data have not been used in the fitting, primarily because the discrepancy between the Gravoko and RBC flux scales is not only frequency but also flux density dependent and there is no complete revision scale available. For a discussion see RBC.

4 SPECTRAL MODEL

A spectral model of the form

\[ \log S = \log A_0 + A_1 \log \nu + A_2 \log^2 \nu + \cdots \]

was used. The model was applied in linear frequency space, i.e.

\[ S[\text{Jy}] = A_0 \prod_{i=1}^{N} 10^{A_i \log^i [\nu/150\text{MHz}]} \]

in order to retain Gaussian noise characteristics. Both determination of the optimal order (N) of polynomial model and maximum likelihood parameter estimation were performed using a Markov Chain Monte Carlo (MCMC) implementation. We used a simulated annealing method, through the METRO documentation (Hobson & Baldwin 2004), to employ a Bayesian inference approach, where Bayes'
The Bayesian evidence, $\Pr(D|H) \equiv Z$, is a factor required for normalizing the posterior over the prior volume, such that

$$Z = \int \mathcal{L}(\Theta) \Pi(\Theta) d^M \Theta,$$

where $M$ is the dimensionality of the prior volume, here $M = N + 1$. For parameter estimation the evidence factor can be neglected as it is independent of the model parameters. Maximum likelihood (ML) or maximum a posteriori (MAP) parameter values can be obtained by sampling the normalized distribution in each case to determine the peak in parameter space. However, in model selection the evidence becomes important for ranking different models based on a common dataset. It can be seen from the previous equation that the evidence represents the average of the likelihood over the prior, and therefore favors models with high likelihood values throughout the parameter space and penalizes models with regions of very low likelihood. This is equivalent to numerically implementing Occam’s razor, whereby larger evidence values are returned for simple models (i.e. fewer parameters) with compact parameter spaces, compared to more complex models - unless the more complex model provides a significantly better fit to the data.

Selecting between models, say $H_0$ and $H_1$, based on their evidence can be done using the ratio,

$$\frac{\Pr(H_0|D)}{\Pr(H_1|D)} = \frac{\Pr(D|H_0)\Pr(H_0)}{\Pr(D|H_1)\Pr(H_1)} = \frac{Z_0 \Pr(H_0)}{Z_1 \Pr(H_1)},$$

where $\Pr(H_0)/\Pr(H_1)$ is the ratio of prior probabilities. This ratio can be set before any conclusions have been drawn from the data; in many cases there is no reason to favor one particular model a priori and consequently this factor can be set to unity. In this circumstance the model selection can be based solely on the ratio of evidences.

In this work, for each model, priors were assumed to be uniform and separable and ML (MAP) parameters were determined initially using the METRO sampling algorithm. Once parameter values had been determined, the evidence in each case, $Z$, was calu-

### Table 2. References for data used in spectral fitting. Column [1] frequency; column [2] reference; column [3] correction factor applied to original data for conversion to RBC flux scale.

| Freq. [MHz] | Ref. | factor |
|------------|------|--------|
| 10 MHz     | Bridle & Purton 1968 | 1.20$^*$ |
| 22.25 MHz  | Roger, Bridle & Costain 1973 | - |
| 38 MHz     | Kellerman, Pauliny-Toth & Williams 1969 | 1.18$^*$ |
| 60 MHz     | Aslanyan et al. 1968 | 1.04$^*$ |
| 86 MHz     | Art'yukh et al. 1969 | 0.94$^*$ |
| 151 MHz    | Baldwin et al. 1985 (6C) | - |
| 178 MHz    | Kellerman, Pauliny-Toth & Williams 1969 | 1.09$^*$ |
| 232 MHz    | Zhang et al. 1997 (MIYUN) | - |
| 325 MHz    | Rengelink et al. 1997 (WENSS) | 0.90$^*$ |
| 408 MHz    | Colla et al. 1970 | - |
| 750 MHz    | Kellerman, Pauliny-Toth & Williams 1969 | - |
| 960 MHz    | Kovalev et al. 1997 | 0.96$^*$ |
| 1400 MHz   | Kellerman, Pauliny-Toth & Williams 1969 | - |

$^*$ from RBC; $^*$ from B77; $^*$ see text for details.
lated over a $\pm 3\sigma$ prior volume centered on the ML parameter values, with $\sigma$ determined for each parameter directly from the posterior distribution. The evidence calculation was repeated multiple times in each case in order to assess the variance of the evidence. Evidence ratios (also known as Bayes factors, or the odds) were then used to determine the optimal polynomial fit based on the Jeffreys scale (Jeffreys 1961), see § 4.1. In practice we take $\Delta \ln Z > 1$ as our threshold for selecting the best model; this choice is justified in Section 5.

### 4.1 Requirements for model selection

The requirement to use a model of increased complexity (i.e. polynomial of higher order) depends upon the degree to which the evidence increases relative to the next lowest order, see Column [9] of Table 3. On the original Jeffreys scale (Jeffreys 1961) an increase of a factor of 3 (i.e. $\Delta \ln Z \geq 3$) is considered substantial evidence to prefer the higher order model and can be considered equivalent to a 99.7% confidence result. Revised versions of the Jeffreys scale (e.g. Gordon & Trotta 2007) divide the level of support into categories where it is considered as either ‘inconclusive’ ($\Delta \ln Z < 1$), ‘weak’ ($1 \leq \Delta \ln Z \leq 2.5$), ‘moderate’ ($2.5 \leq \Delta \ln Z \leq 5$), or ‘strong’ ($\Delta \ln Z \geq 5$).

### 5 RESULTS

The ML parameters and evidence values for each polynomial fit to the source spectra are listed in Table 3. An example of the different orders is shown for 3C48 in Fig. 1. The poor fit of the linear and 2$^\circ$ polynomial model is evident by eye. This is also reflected in the values of the evidence for these models: an evidence ratio, and hence difference in the logarithm of the evidence, of $\ln Z_3 - \ln Z_2 > 100$ indicates a definitive preference; a difference of $\ln Z_2 - \ln Z_1 = 3.16$ is substantial evidence for preferring the 3$^\circ$ model above the 2$^\circ$ model. The fractional evidence ratio, $\ln Z_4 - \ln Z_3 = -3.55$, between the 4$^\circ$ and 3$^\circ$ polynomial models indicates that the 3$^\circ$ model is still preferred. In this case the goodness-of-fit is not diminished by the 4$^\circ$ model, but there is no evidence in the data to support the use of the extra parameter and hence the model is penalized.

In general the results for this sample are easily interpreted, with Bayes factors of $\Delta \ln Z > 3$ clearly indicating a preferred order of polynomial in most cases. When comparing different polynomial order fits to the 3C295 and 3C380 datasets the Bayes factors are less conclusive than in other cases. A difference of $\Delta \ln Z = 1.6$ between the third and fourth order models in the case of 3C295 is intermediate to the ‘weak support’ category. Although the support for moving to the higher order model is weak, it is not inconclusive and so in the context of the work here we choose to prefer the fourth order model. In the case of 3C380, a value of $\Delta \ln Z = 0.64$ is securely in the inconclusive category and so we prefer the lower order model in this instance. Best fitting spectral models for the six calibrator sources are shown in Fig. 2.

### 5.1 Error budget

Errors on individual parameters for each fit were determined directly from the posterior distribution and are listed in Table 3. The uncertainty in the model due to these errors was derived analytically using differential error propagation and the 1$\sigma$ bound on the model in each case is illustrated in Fig. 2 as a blue shaded area. We illustrate the percentage error of each model as a function of frequency from 30 – 240 MHz.

#### 5.2 Notes on Individual Sources

3C380 This source has data at 10 MHz in Bridle & Purton (1968) but the very low flux density (168 Jy) indicates that the spectrum turns over sharply below 20 MHz. The effect of this turn over is marginal above 30, but would require significantly increased complexity in the model. Consequently these data have been excluded from the fit.

### 6 CONCLUSIONS

We have presented parameterized broadband spectral models for six bright radio sources selected from the 3C survey between 30-300 MHz, spread in Right Ascension from 0 – 24 hours. For each source, data from the literature have been compiled and tied to a common flux density scale. These data have then been used to parameterize an analytic polynomial spectral calibration model. The best fitting polynomial model order in each case has been determined using the ratio of the Bayesian evidence for the candidate models. Maximum likelihood parameter values with associated errors have been presented. The percentage error in each model as a function of frequency has been derived and is illustrated in § 5.1.
These spectral models are intended as an initial reference for science quality data from the new generation of low frequency telescopes, such as LOFAR, now coming on line. In this context we have shown that two of these sources lead to unacceptably high flux scale uncertainty at frequencies below 70 MHz (3C147 & 3C295), and we also note that 3C380 may be unsuitable for precision calibration at higher frequencies where its angular extent becomes an issue.

7 ACKNOWLEDGEMENTS

We thank Julia Riley and Ger de Bruyn for useful discussions and Rick Perley for his careful review of the manuscript. This research has made use of the Astrophysical CATalogs Support System (CATS) and the NASA/IPAC Extragalactic Database (NED).

REFERENCES

Akujor C. E., Garrington S. T., 1995, A&A, 112, 235
Artyukh V. S., Dagkesamanskii V. V., Vitkevich R. D., Kozhukhov V. N., 1969, SvA, 12, 567
Aslanyan A. M., Malumyan V. G., Sanamyan V. A., 1968, Ap, 4, 252
Baars J. W. M., Genzel R., Pauliny-Toth I. I. K., Witzel A., 1977, A&A, 61, 99
Baldwin J. E., Boysen R. C., Hales S. E. G., Jennings J. E., Waggett P. C., Warner P. J., Wilson D. M. A., 1985, MNRAS, 217, 717
Bennett A. S., 1962, MmRAS, 68, 163
Braude S. Y., Megn A. V., Ryabov B. P., Zhouchk I. N., 1970, ApkSS, 8, 275
Braude S. Y., Lebedeva O. M., Megn A. V., Ryabov B. P., Zhouchk I. N., 1970, ApL, 5, 129
Bradl A. H., Burton C. R., 1968, AJ, 73, 717
Colla G., et al., 1970, A&AS, 1, 281
Edge D. O., Shakeshaft J. R., McAdam W. B., Baldwin J. E., Archer S., 1959, MmRAS, 68, 37
Gordon C., Trotta R., 2007, MNRAS, 382, 1859
Hobson M. P., Baldwin J. E., 2004, Applied Optics, 43, 2651
Jeffreys, H., 1961, Theory of probability, Clarendon Press, Oxford, 3rd ed., 50, 69
Kellermann K. I., Pauliny-Toth I. I. K., Williams P. J. S., 1969, ApJ, 157, 1
Kovalyov Y. A., 1997, BSAO, 44, 50
Pearson T. J., Readhead A. C. S., Perley R. A., 1985, AJ, 90, 738
Pooley G. G., Henbest S. N., 1974, MNRAS, 169, 477
Reid A., Shone D. L., Akujor C. E., Browne I. W. A., Murphy D. W., Pedelty J., Rudnick L., Walsh D., 1995, A&A, 110, 213
Rees N., 1990a, MNRAS, 244, 233
Rees N., 1990b, MNRAS, 243, 637
Rengelink R. B., Tang Y., de Bruyn A. G., Miley G. K., Bremer M. N., Roettgering H. J. A., Bremer M. A. R., 1997, A&A, 124, 259
Rogers R. S., Costain C. H., Brindle A. H., 1973, AJ, 78, 1030
Rogers R. S., Costain C. H., Lacey J. D., 1969, AJ, 74, 366
Scott P. F., Shakeshaft J. R., 1971, MNRAS, 154, 199
van der Tol S., Jeffs B. D., van der Veen A. J., 2007, ITSP, 55, 4497
Zhang X., Zheng Y., Chen H., Wang S., Cao A., Peng B., Nan R., 1997, A&A, 121, 59