Multi-Frequency Synthesis with the ATCA

R.J. Sault

May 8, 1992

Introduction

Multi-frequency, or bandwidth, synthesis, is the practice of combining continuum data at several frequencies during the mapping stage, to produce a single output map. The u-v coordinates for the data are calculated from their actual frequencies (rather than some mean frequency), and convolved onto the gridded u-v plane as separate visibilities.

There are a number of advantages to this:

- The different frequencies will result in differing u-v spacings giving better u-v coverage.
- A wider instantaneous observing bandwidth will reduce the noise level in the final map.
- Dividing a single wide bandwidth into many small channels, and gridding them at their correct u-v coordinates reduces bandwidth smearing effects.

In the case of the ATCA, a single observation could result in several frequencies in three ways. These are either by time switching between different frequencies, or by observing two frequency bands simultaneously, or by using the different correlator channels in forming one map. Indeed, because the correlator produces many channels for continuum observations, the normal ATCA imaging mode should be to use multi-frequency synthesis.

When combining data at several frequencies, it is generally assumed that the source spectral index is small, and that it produces negligible distortion in the resultant processing. For high dynamic range ATCA maps, this may be a poor assumption at 20 cm (where the fractional bandwidth is 10%), when two widely differing frequency bands are being observed, or when the frequencies are being intentionally time-switched and are widely separated.

Conway, Cornwell and Wilkinson (1990) presented a technique to account for source spectral index in the deconvolution process. Additionally they analysed the affect that typical spectral indices have on the map quality. They found that for fractional bandwidths of 10%, spectral index effects need to be accounted for to produce maps with dynamic ranges greater than 1000.

We define the spectral index of a source, α, by the power law relation:

$$I(\nu) = I(\nu_0)(\frac{\nu}{\nu_0})^{\alpha}. \tag{1}$$

Generally the spectral index varies with position, but we drop this explicit dependence on position for notational simplicity. For a spectral variation other than a power law, we can define the spectral index as:

$$\alpha = \frac{\nu}{I}\frac{\partial I}{\partial \nu}. \tag{2}$$

This memo describes some experience of multi-frequency synthesis with the ATCA. It presents a new multi-frequency synthesis deconvolution algorithm, and outlines the multi-frequency synthesis software available in Miriad.

Multi-Frequency Deconvolution

To understand how the spectral index of a source can be accounted for in the deconvolution stage, consider a noiseless observation of a single point source at the phase centre. Assume that it has a linear brightness variation with frequency. A visibility for this source is given by two parameters: the source brightness, I, at some arbitrary reference frequency, ν_0, and the brightness derivative with frequency. That is

$$V(\nu) = I(\nu_0) + \frac{\partial I}{\partial \nu}\Delta\nu. \tag{3}$$

Consider the multi-frequency synthesis dirty map that we would form from an observation of this source. We can divide its response into two parts. If we assumed that the brightness at the reference frequency is unity, and the brightness derivative is zero (i.e. spectral index of zero) then we could predict the response of the source. This would be the normal dirty beam, $B_0(l,m)$. Similarly, we could predict the response of the source if we made the unphysical assumption that the brightness at the reference frequency is zero, and that the brightness derivative is

$$\frac{\partial I}{\partial \nu} = \frac{1}{\nu_0} \tag{4}$$

(the reason for this choice should become apparent). That is, we make a map from visibilities with the values

$$V(\nu) = \frac{\Delta\nu}{\nu_0}. \tag{5}$$

Conway *et al.* call this image the "spectral dirty beam", or $B_1(l,m)$. For fractional bandwidths of 10%, its peak value is typically less than 0.01.

Because of the linearity of the system, the response of the actual source will be a weighted sum of these two individual images. The weights are simply the values of the sources brightness and ν_0 times the brightness derivative. So the response of a point source with an arbitrary linear spectral variation will be

$$I(\nu_0)B_0(l,m) + \nu_0\frac{\partial I}{\partial \nu}B_1(l,m). \tag{6}$$

Generalising to an arbitrary source with linear spectral variation, the dirty map, $M(l,m)$, will be given by

$$M = I(\nu_0) * B_0 + \nu_0\frac{\partial I}{\partial \nu} * B_1, \tag{7}$$

where "$*$" represents convolution. With the exception of ν_0, all quantities in the above equation are functions of position (for notational brevity, we will not show this dependence explicitly).

Although the brightness is a desired quantity, we are generally interested in the spectral index rather than the scaled derivative. However, the scaled derivative and the spectral index are closely related, namely

$$I(\nu_0)\alpha = \nu_0\frac{\partial I}{\partial \nu}. \tag{8}$$

In the deconvolution process, we will be interested in solving for $I(\nu_0)$ and $I(\nu_0)\alpha$. Again for notational brevity, we will not explicitly include ν_0 in the equations that follow, but it should be remembered that we are solving for I and $I\alpha$ at the reference frequency ν_0.

Summarising the above discussion, the dirty map, M, will be a combination of two parts,

$$M = I * B_0 + (I\alpha) * B_1. \tag{9}$$

Although spectral variation is normally modeled by a power law (rather than a linear variation), clearly a linearised approximation to the power law would be adequate when the frequency range is not large. Additionally Conway *et al* have shown that an alternate expansion of the power law, about the geometric mean frequency, results in an approximation which can cope with a wider frequency range. Conway *et al.* in general, give a more rigorous treatment than the more intuitive approach given here.

The Conway Algorithm

Given that the response of a point source with spectral index, can be represented by the weighted sum of a dirty beam and a spectral dirty beam, a CLEAN-like multi-frequency synthesis algorithm can consist of locating components of these beams in the dirty map. The Conway *et al.* algorithm does this by iteratively performing two steps (hence its name as 'double deconvolution'). The first step consists of a series of normal CLEAN iterations (CLEANing with the normal dirty beam) until some depth is reached. That is, CLEAN M with B_0. The second step consists of CLEANing the remaining residuals with B_1. The first step determines components of B_0, whereas the second determines B_1 components.

There is a complication in the second step; the spectral dirty beam does not have a central peak. Indeed, Conway *et al.* have a preference for arranging the weights, etc, so that the central pixel of the spectral dirty beam is zero. To circumvent this, they correlate both the residuals and the spectral dirty beam, with the spectral dirty beam, and then CLEAN with these in the normal way. That is, they CLEAN $M \star B_1$ with $B_1 \star B_1$ (where $\star$ represent correlation).

This correlation with the spectral dirty beam is nothing mysterious – the correlation of B_1 with itself will always have a dominant central lobe. In general, if one wants to locate a signal of known form, buried in (gaussian) noise, then an optimum least squares approach is to correlate the noise with the signals form. The peak value in the result will give the location. Indeed, as Conway *et al.* note, normal CLEAN algorithms should CLEAN $M \star B_0$ with $B_0 \star B_0$. Rarely (if ever) is this done, because for a uniformly weighted beam (ignoring grid correction effects), B_0 will be proportional to $B_0 \star B_0$, and correlating with B_0 would simply introduce a scaling.

An Alternate Algorithm

An alternate approach is eliminate the "double deconvolution" nature of the Conway scheme, and to determine the contributions of B_0 and B_1 simultaneously. For simplicity, one dimensional notation will be used. A least squares approach would be as follows. If there is a CLEAN component at pixel j of the dirty map, M, and the beam centre is at 0, we would find a_0 and a_1 which minimise

$$\epsilon^2 = \sum_i (M(i+j) - a_0 B_0(i) - a_1 B_1(i))^2. \tag{10}$$

Solving this minimisation gives:

$$\begin{aligned} a_0 &= (A_{11}(0)A_{M0}(j) - A_{01}(0)A_{M1}(j))/\Delta \\ a_1 &= (A_{00}(0)A_{M1}(j) - A_{01}(0)A_{M0}(j))/\Delta \end{aligned} \tag{11}$$

where

$$\Delta = A_{00}(0)A_{11}(0) - A_{01}(0)A_{10}(0) \tag{12}$$

and

$$\begin{aligned} A_{00}(j) &= \sum_i B_0(i)B_0(i+j) \\ A_{11}(j) &= \sum_i B_1(i)B_1(i+j) \\ A_{01}(j) &= \sum_i B_0(i)B_1(i+j) \\ A_{10}(j) &= \sum_i B_1(i)B_0(i+j) \\ A_{M0}(j) &= \sum_i B_0(i)M(i+j) \\ A_{M1}(j) &= \sum_i B_1(i)M(i+j) \end{aligned}$$

In addition to determining the contributions of the B_0 and B_1 beams, any CLEAN-like algorithm must also determine the location of the next component. It is tempting to simply search for the peak residual.

as regular CLEAN does. A second approach is to find the location, j, that will minimise ϵ^2 – that is, simultaneously find j, a_0 and a_1 which minimises the error measure. It can be shown that these two approaches are identical for a normal CLEAN with a uniformly weighted beam. However, they are distinct for weighting schemes other than uniform weighting, as well as the multi-frequency synthesis deconvolution case. At least in the multi-frequency deconvolution case, it was found that the second method gives better deconvolutions.

The second approach notionally involves, at every pixel location, solving for a_0 and a_1 and computing the resultant ϵ^2. Some algebraic manipulation shows that this is equivalent to finding the location, j, which maximises

$$A_{M0}(j)^2 A_{11}(0) + A_{M1}(j)^2 A_{00}(0) - A_{M0}(j)A_{M1}(j)(A_{01}(0) + A_{10}(0)) \tag{13}$$

Implementational Issues

The location of a component, and determining the contribution of the B_0 and B_1 beams involves simple expressions of the A coefficients. Whereas A_{00}, A_{01}, A_{10} and A_{11} are fixed, A_{M0} and A_{M1} will vary as the CLEAN process progresses. All the A coefficients are quantities correlated with the B_0 and B_1 beams. A comparatively inexpensive algorithm presents itself:

1. Initially calculate:

$$\begin{aligned} A_{M0} &= M \star B_0 \\ A_{M1} &= M \star B_1 \\ A_{00} &= B_0 \star B_0 \\ A_{01} &= B_0 \star B_1 \\ A_{10} &= B_1 \star B_0 \\ A_{11} &= B_1 \star B_1. \end{aligned}$$

   These would be computed using FFTs.

2. Compute equation 13 at all points, and so find its maximum.

3. Determine the contributions of the B_0 and B_1 beams, a_0 and a_1, using equation 11. The values of a_0 and a_1 are saved in component lists.

4. Update A_{M0} and A_{M1}.

$$\begin{aligned} A_{M0}(i+j) &\leftarrow A_{M0}(i+j) - a_0 A_{00}(i) - a_1 A_{10}(i) \\ A_{M1}(i+j) &\leftarrow A_{M1}(i+j) - a_0 A_{01}(i) - a_1 A_{11}(i). \end{aligned}$$

   This corresponds to the component subtraction step in normal CLEAN.

5. If it has not "converged", go back to step 2. Otherwise finish up.

If the data contain a single frequency, then B_1 is zero and the above algorithm degenerates into a standard CLEAN of $M \star B_0$ with $B_0 \star B_0$.

Though this algorithm has been described using a Högbom CLEAN scheme, it is equally well suited to a Clark scheme. It is not as well suited to the Steer CLEAN algorithm, however, unless we could assume that large regions of the dirty map have the same spectral index.

The preceding section has glossed over performing the correlations with the B_0 and B_1 beams. It was found that correlating with the entire B_0 and B_1 beams introduces edge effects which result from the finite extent of the images and the cyclic correlation which results from using FFTs. It was found preferable to correlate with only a subportion of B_0 and B_1 beams. A_{M0} and A_{M1} will be the full-sized dirty maps correlated with the subportion of the B_0 and B_1 beams. Similarly A_{00}, A_{01}, A_{10} and A_{11} will be a full

size beam correlated with a subportion of the beam. Ideally this subportion should be as large as the region being CLEANed, but in practice smaller regions can be tolerated. In a normal Clark CLEAN, the beam needs to be twice the size of the region being CLEANed. For this MFS algorithm, the beam would preferably be as large as three times the size of the region being CLEANed. Additionally we have not commented on the fact that A_{01} and A_{10} should be mirror images. When we correlate with only a subportion of the beam, A_{01} and A_{10} will not be mirror images – though their central pixels will have the same value.

Calibration Issues

To derive accurate spectral index information, care must be taken in the flux calibration steps. This is particularly so if there are multiple signal paths and the data from the different paths are being calibrated separately. It is easy for flux scale mismatches to dominate the spectral index effects.

On a different tack, Conway *et al.* note that if the visibility data are dominated by structure with a particular spectral index, α_0, then it is best to scale the visibility data by

$$(\frac{\nu}{\nu_0})^{-\alpha_0} \tag{14}$$

before mapping. This will tend to eliminate the spectral index from the data, and so reduce spectral effects. The residual effects will be due to the smaller remaining part of the spectral index.

Other Spectral Effects

The deconvolution algorithms described above approximate the spectral variation with frequency as a linear one. In the preceding sections we have interpreted this spectral variation as a spectral index. However there are other effects which can produce a spectral variation. Although the weighting schemes that Conway *et. al.* discuss have been optimised for a power law spectral variation, the deconvolution schemes will also work for any spectral variation that is approximately linear. These effects may well dominate any spectral index effect.

Primary Beam Effects

As the attenuation due to the primary beam is a function of frequency, it causes an effect in multi-frequency synthesis quite similar to source spectral index. However it is not of consequence near the pointing centre and it is a predictable effect. We can model the primary beam response, P, by a gaussian form:

$$P(\theta,\nu) = \exp(-a(\frac{\theta}{\theta_0})^2(\frac{\nu}{\nu_0})^2). \tag{15}$$

Here θ is the angular distance from the pointing centre, and θ_0 is the primary beam FWHM at the reference frequency ν_0, and $a = 4\log(2)$. This will result in an apparent spectral index of

$$\alpha = -8\log(2)(\frac{\theta}{\theta_0})^2(\frac{\nu}{\nu_0})^2. \tag{16}$$

At the half power point ($\theta = \theta_0/2$), and at the reference frequency, this corresponds to an apparent spectral index of -1.4.

Provided the fractional bandwidth is not too large, the apparent spectral index will be the sum of the true source spectral index and a part due to the primary beam variation. Because the primary beam effect is predictable, the apparent spectral index can be corrected to give the true spectral index.

Faraday Rotation

When imaging Stokes Q and U, Faraday rotation will cause a spectral variation. For a rotation measure R_m, speed of light c, intrinsic polarisation angle λ_0 and total linearly polarised intensity P (a real number) with spectral index α, then $Q+iU$ will vary as

$$Q(\nu)+iU(\nu) = P(\nu_0)(\frac{\nu}{\nu_0})^\alpha \exp(i(2R_m\frac{c^2}{\nu^2}+\lambda_0)) \tag{17}$$

It is useful to express this variation as an "apparent complex spectral index", α_C, being the ratio

$$\begin{aligned}\alpha_C &= \frac{\nu}{Q+iU}\frac{\partial(Q+iU)}{\partial\nu} \\ &= \alpha - 4i\frac{c^2}{\nu^2}R_m.\end{aligned} \tag{18}$$

This quantity can be readily computed from the results of the multi-frequency deconvolution of Q and U maps. For a rotation measure of 30 rad/m^2 and frequency of 4.7 GHz, this results in an imaginary part of the complex spectral index of $-0.49i$. As always, a linear approximation is being made – certainly π ambiguities are not dealt with. However the linear approximation should be reasonable provided

$$|\frac{\Delta\nu}{\nu_0}\frac{c^2}{\nu_0^2}R_m| < 0.2. \tag{19}$$

Implementation in Miriad

The following subsections briefly describe the tasks in the Miriad package that are useful in performing multi-frequency synthesis.

Map-Making: INVERT

The option "**mfs**" in the inputs to the map-making program, INVERT, instructs it to grid all correlations (regardless of frequency) into a single map using the correct u and v coordinates for each frequency. The correlations to use can be specified using the '**line**' and '**select**' parameters. The '**sbeams**' parameter is used to form spectral dirty beams. '**Sbeams**' is set to indicate the highest order spectral beam to produce, as described by Conway *et al.*. The zeroth and first order beams are the normal dirty beam and spectral dirty beam respectively. Higher order dirty beams can be produced (second, third, etc), though they are not used by any current Miriad software. The default is to produce only the normal dirty beam (i.e. the zeroth order beam, or **sbeams=0**). The reference frequency is the geometric mean of the frequencies of the input correlations, and the "logarithmic expansion" is used in forming the spectral beams (as suggested by Conway *et al.*). Because of the effects of the weighting scheme, the central pixel of the spectral dirty beam will be zero only for naturally weighted maps. All the beams are stored as different planes in the output **beam** data-set.

Cleaning: MFCLEAN and RESTOR

MFCLEAN is the Miriad task which implements the multi-frequency deconvolution scheme described above. It can perform either Clark or Högbom iterations. The inputs to MFCLEAN include a dirty map and beam, formed by INVERT, using the "**mfs**" option and **sbeams** set to 1 or greater (that is, the the beam data-set must contain at least the normal dirty beam and the spectral dirty beam). MFCLEAN produces an output model consisting of two planes – the I and $I\alpha$ planes. In other respects, it takes much the same arguments as the CLEAN task. Note however, that the beam should be three times the size of the region being cleaned (normal CLEAN requires it to be twice the size).

RESTOR takes the output model from MFCLEAN and produces either a residual map, or a "restored" map. The restored map will be a single plane corresponding to the source at the reference frequency.

Spectral Index and Rotation Measure: MFSPIN

Given the model produced by MFCLEAN, MFSPIN computes a spectral index-image. Optionally it will correct the spectral index for the effect of the primary beam using a gaussian model. It can also produce a rotation measure image, if fed Stokes Q and U models.

Self-Calibration: SELFCAL and GPSCAL

Task SELFCAL assumes that antenna gain is solely a function of antenna number, i and j, and is independent of frequency. Consequently all frequencies can be used in the solution process. The antenna gains, g_i, are found which minimise the difference between the true visibilities, V_{ij}, and the model visibilities, $\hat{V}_{ij}$, viz

$$\epsilon^2 = \sum |\hat{V}_{ij} - g_i g_j^* V_{ij}|^2.$$

For data within a given solution interval, the summation is taken over all baselines and frequencies. Normally SELFCAL uses the rule that the model image must contain a plane at the same frequency as the true visibility, before it will generate a model visibility. Giving the "`mfs`" option allows SELFCAL to relax this rule in one of two ways. If the model image consists of a single plane, then option "`mfs`" causes SELFCAL to assume the source has zero spectral index. If the model image consists of two planes (as produced by MFCLEAN), then SELFCAL assumes that the second plane is an $I\alpha$ plane. Either way, SELFCAL can now determine the model visibility at any frequency.

The assumption that the gains are independent of frequency should be treated with some caution. If all the visibilities for a given baseline go through the one signal path, then the instrumental phase for the different visibilities should be equal (assuming the signal path is well equalised). However atmospheric phase should vary linearly with frequency. Provided the fractional bandwidth is small, it is still a reasonable approximation that the gains are independent of frequency. Indeed this is the normal approximation used in deriving calibration solutions from a "channel 0" data-set.

If the visibilities for a given baseline go through multiple signal paths (e.g. the two IF system for the ATCA), then the instrumental phases for the different paths could be significantly different. Tentative results for the two IF system of the ATCA show that, apart from an offset, the phase of the two signal paths track each other as the ratio of the frequencies (i.e. the path length difference is constant). See Figure 1 (by Kesteven) for an example. The offset between them can be eliminated either by adjusting the phases of the two paths to zero at the start of the observation, or by the primary calibration process. Provided the offset has been eliminated, and provided the frequencies are not too different, it can be useful to approximate the gains as independent of frequency.

If the phases vary significantly between the frequencies, then the different frequencies need to be self-calibrated individually. Ideally there would be a self-calibration task which knew that the phase errors at different frequencies tracked as the ratio of the frequencies.

SELFCAL cannot convert between raw polarisations and Stokes parameters. The visibility data and the model must be of the same polarisation type. Additionally, SELFCAL assumes that the gains of the X and Y polarisation channels are the same. This generally appears to be a good assumption. Task GPSCAL is similar to SELFCAL, except that it can convert Stokes parameter models into raw polarisation parameters (GPSCAL can take I, Q, U and V models). Neither does it assume that the gains of the X and Y polarisation channels are equal.

An Example Observation

An ATCA observation of the source 1733-56 (by Morganti, Killeen and Tadhunter) is useful to demonstrate multi-frequency techniques. Simultaneous observations were made at 4.74 and 5.7 GHz, using two configurations. The two observations were "snapshots" resulting in 16 cuts of 10 to 15 minutes each. At the two frequencies, the middle 17 channels of the correlator output were averaged into "channel-0" data-sets, giving a 68 MHz bandwidth. The two frequencies were calibrated independently. The total uv

coverage for the observation (including the tracks at both 4.74 and 5.7 GHz) is given in Figure 2. The final map, after several INVERT/MFCLEAN/SELFCAL iterations is shown in Figure 3. It was found that better maps resulted by assuming the SELFCAL gains were the same at the two frequencies (i.e. self-calibrating both frequencies together, rather than individually).

This observation is a good candidate for using multi-frequency techniques. The source is strong, containing a 560 mJy peak central core. So, despite modest uv coverage, a map of reasonable dynamic range results. The rms noise away from the source is 0.65 mJy/beam, whereas residual errors near the central core have peak values of 3.9 mJy/beam. The theoretical noise is 0.10 mJy/beam – phase errors appear to limit the dynamic range.

Stokes Q, U and V maps were made and MFCLEANed, using the same SELFCAL gains. The Q and U maps showed very little leakage of I (less than 0.1%). V showed a leakage of I of 0.3% at the central core. The rms noise away from the source was 0.18 mJy/beam. Despite the source being moderately polarised (Q and U are the order of 30 mJy/beam in the hotspots), GPSCAL failed to improve over the SELFCAL gain solutions.

The spectral indices derived from I and $I\alpha$ were typically -0.5 for the northern hotspot, 0.0 for the central core and -0.7 for the southern hotspot. These spectral indices appeared to be believable to the 0.1 level. and are consistent with what would be expected for a source of this kind. These spectral indices have been corrected for the effect of the primary beam. As the source is moderately large primary beam effects will contribute an apparent spectral index of -0.8 in the southern-most hotspot. Additionally the spectral indices were corrected for an apparent error in the primary calibration (discovered after the data reduction), which introduced a spectral index of -0.6. The spectral index for the southern hotspot appears steeper than might be expected, which may be due to the simple gaussian model used to correct for primary beam effects.

Spectral index and rotation measure images were derived from the Q, $Q\alpha$, U and $U\alpha$ images. The spectral index was consistent with that of I (but of poorer quality). The average rotation measure was -35 rad/m^2, but the error in this would appear to be high (the rms scatter was 30 rad/m^2).

To compare this against standard processing, four additional clean maps were made using the normal CLEAN algorithm. The best SELFCAL gains previously determined were used in these experiments. The first two maps were made using just the 4.74 and 5.7 GHz data, respectively, whereas the third was the average of the first two maps. The fourth was formed using data from both frequencies (gridding using the correct uv coordinates), but cleaned using standard CLEAN. For these four maps, the rms error level away from the source was 0.67, 0.78, 0.76 and 1.0 mJy/beam respectively, and the deepest negative near the central core was -10, -10, -10 and -8.2 mJy/beam respectively. Figure 4 shows these four maps. This should be compared with the version using multi-frequency techniques. Figure 5, which has an rms error level away from the source of 0.65 mJy/beam and a minimum of -3.9 mJy/beam. The grey-scales are saturated at the 1% level of the core flux. Although multi-frequency synthesis processing has not affected the error level far from the sources, it has significantly reduced errors near the sources. Additionally we have given an "unfair" advantage to the "normal" techniques by using the best SELFCAL gains derived from multi-frequency synthesis processing. Presumably the inferior models resulting from the "normal" techniques would have resulted in inferior SELFCAL gains.

References

J.E Conway, T.J. Cornwell, P.N. Wilkinson, 1990,"Multi-frequency synthesis a new technique in radio interferometric imaging". *Mon. Not. R. astr. Soc.* 246, 490.

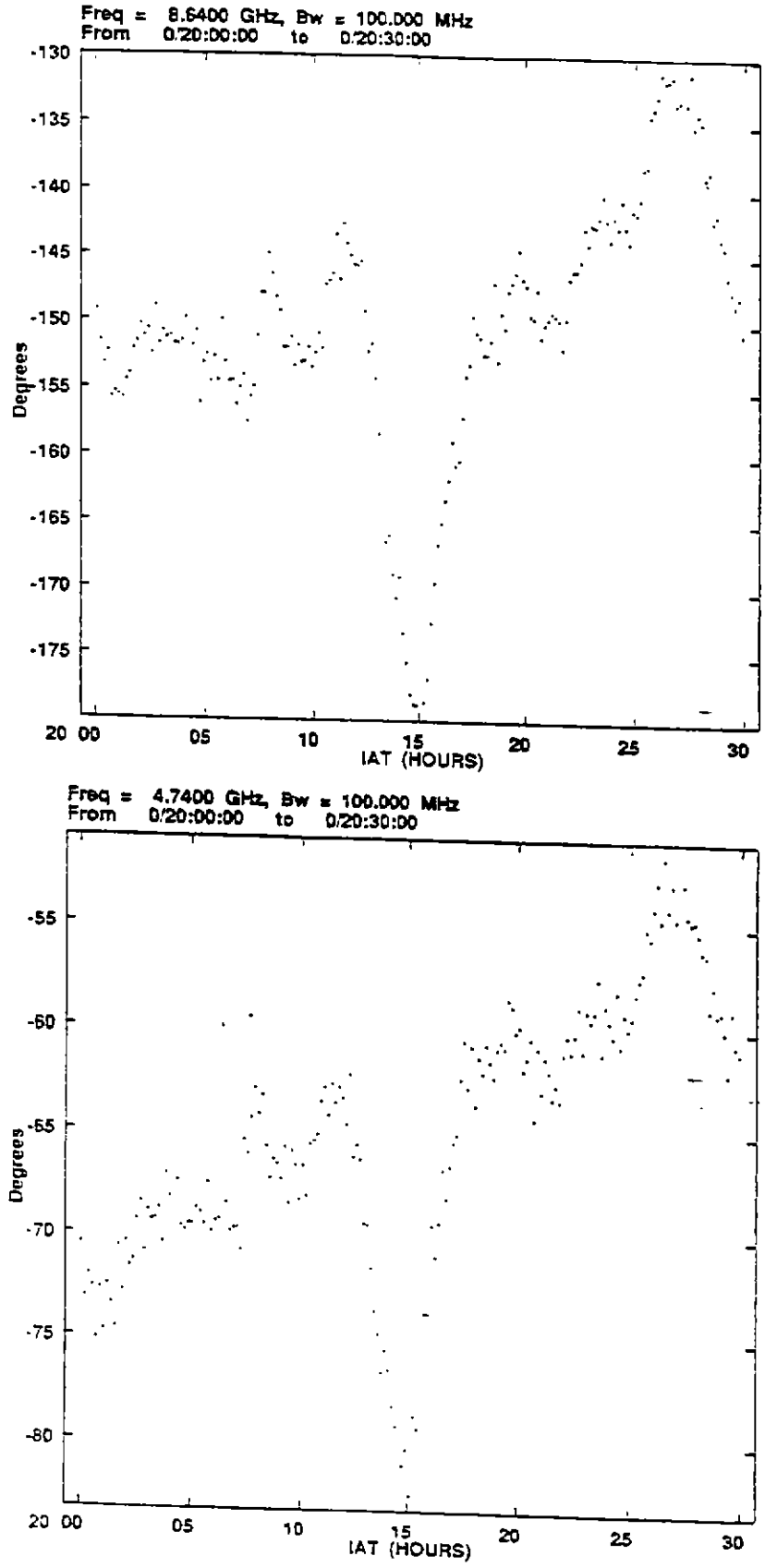

Figure 1 Phase Errors, at 4.74 and 8.64 GHz Note that, apart from an offset and scale factor, the phases track each other well.

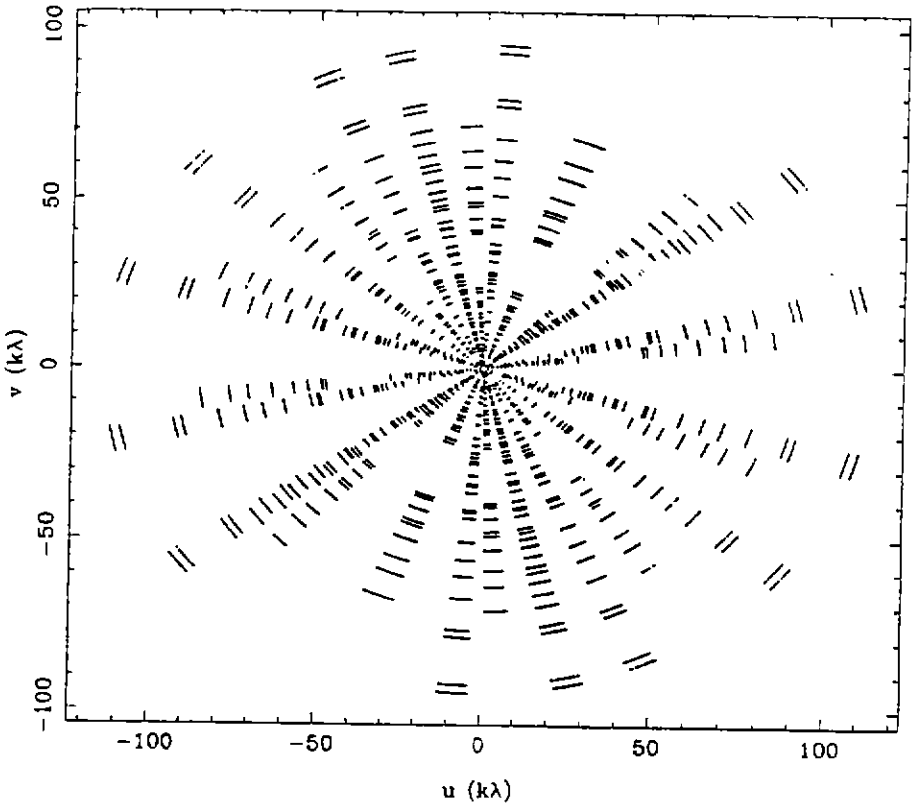

Figure 2: UV Coverage for 1733-56

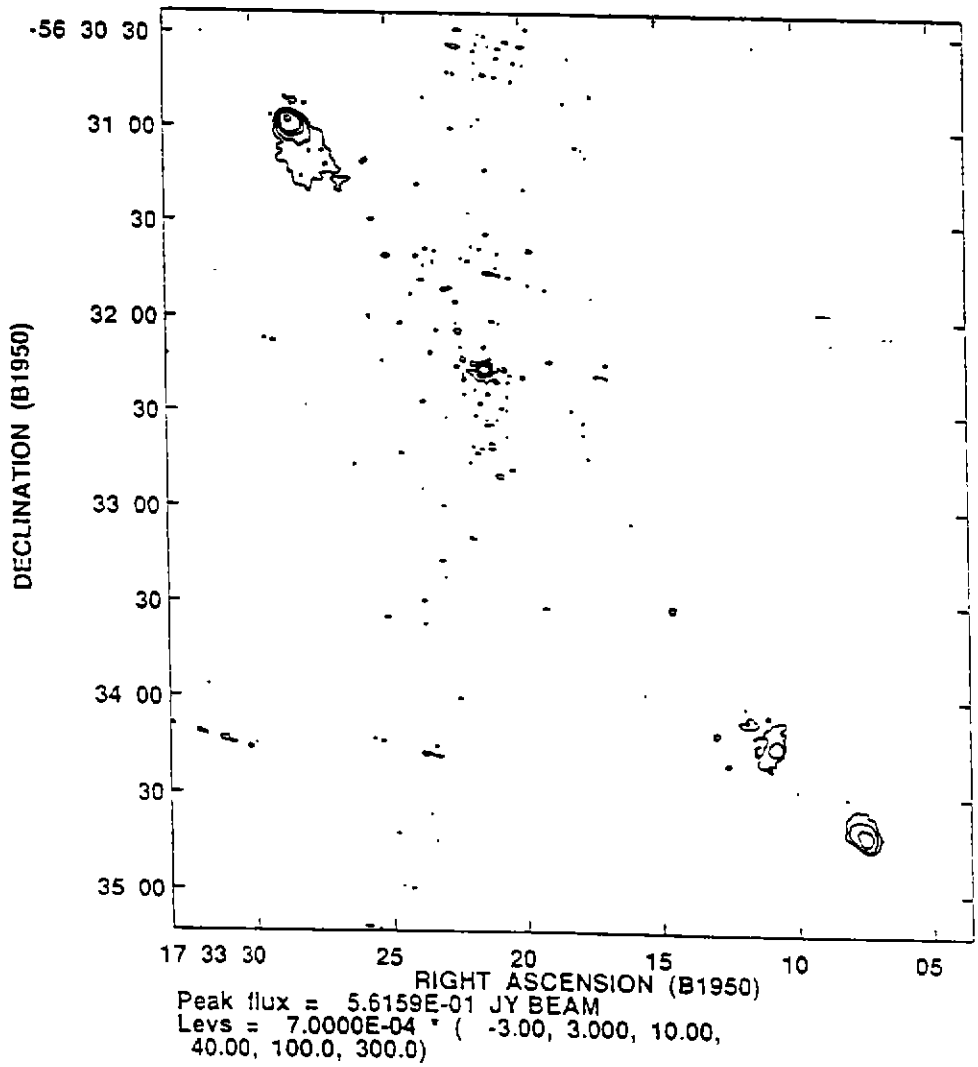

Figure 3: Source 1733-56, using MFCLEAN

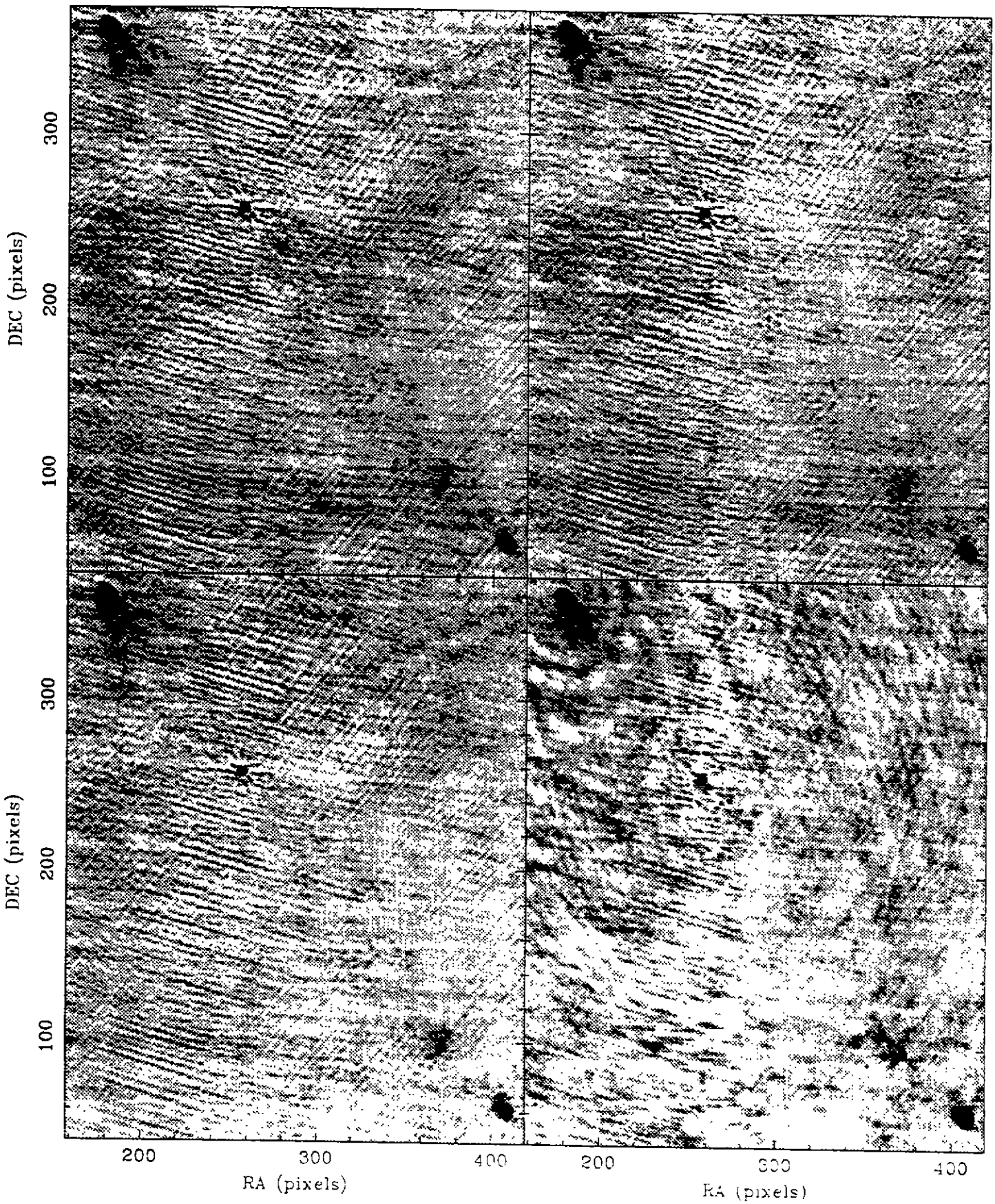

Figure 4. Source 1733-56 - Using Normal CLEAN. The upper left panel uses 4.7 GHz data only. Upper right uses 5.7 GHz data only. Lower left is the average of the of the upper two. Lower right map all data together, but uses standard CLEAN

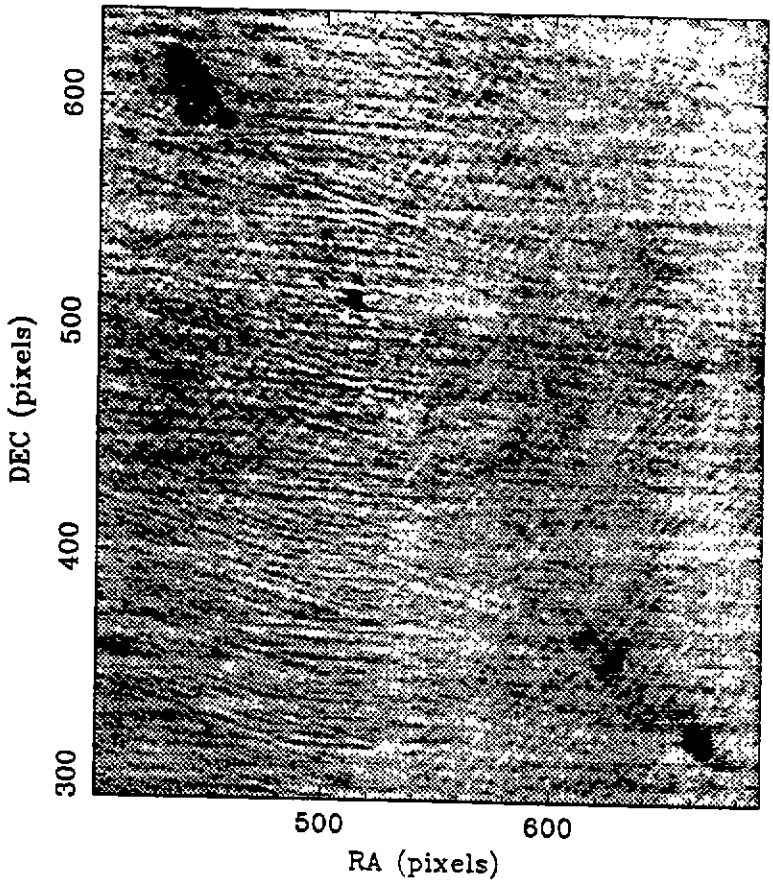

Figure 5: Source 1733-56 – Using MFCLEAN