Data Analysis

"I have got some data, so what now?"

Naomi McClure-Griffiths
CSIRO Australia Telescope National Facility
2 Oct 2008

Outline

- Non-imaging analysis
- Parameter estimation
  - Point source fluxes, positions
  - Extended source flux
- Image combination
  - Spectral index
  - Polarization - see McConnell and Perley talks
- Visualisation
- 3-D datasets
  - Moments
- Understanding the limitations of your data

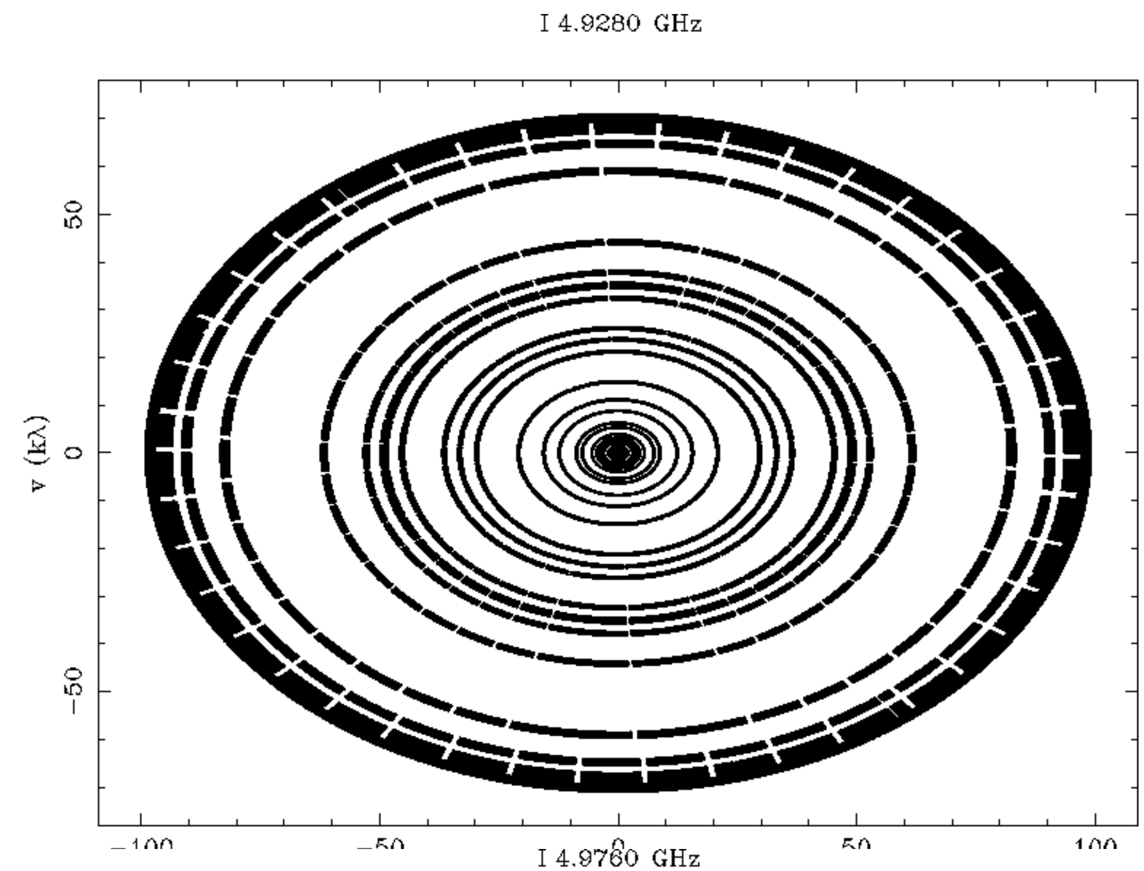
Step 1: What do you want to know from your observations?

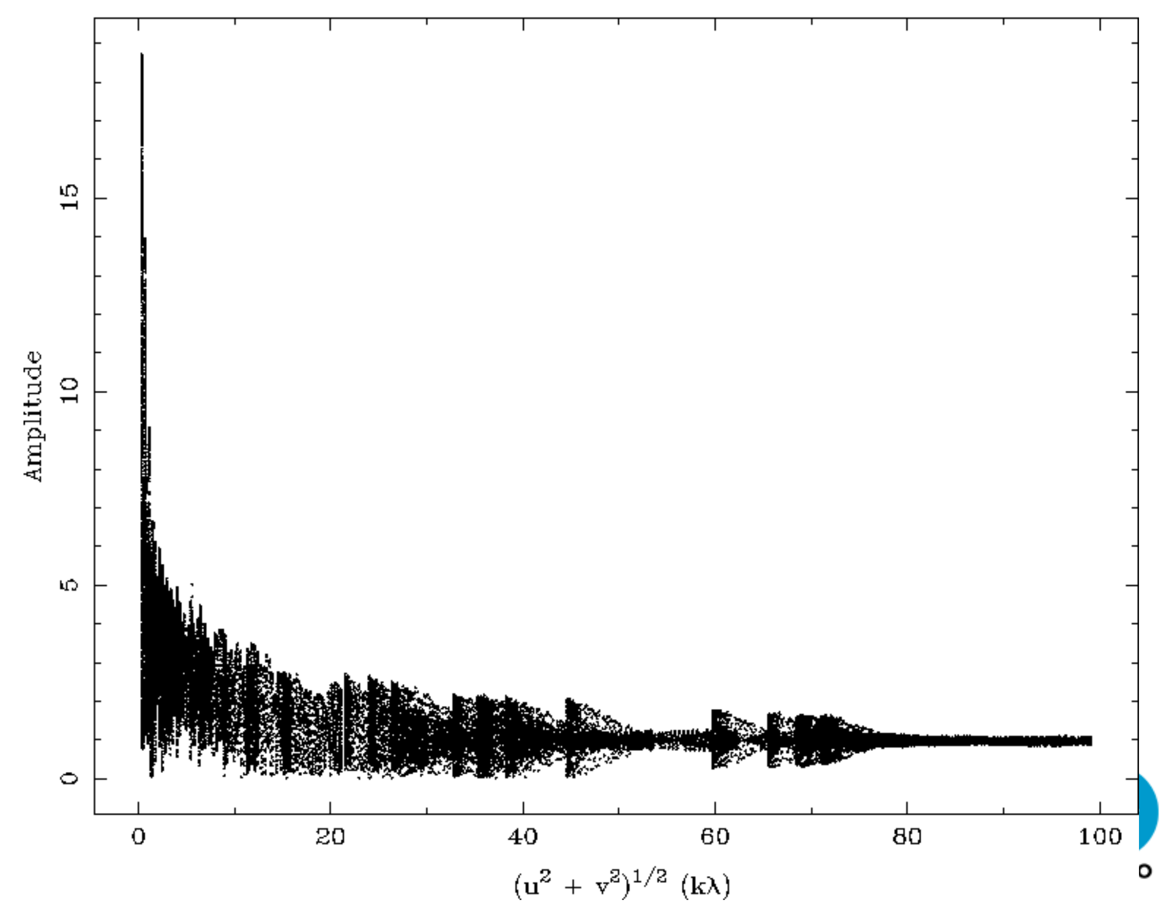
- Variability with time?
  - u-v amplitude vs time
- Source position/flux/ extent?
  - Characterising a source
  - Model fitting
  - Extend source flux
- Source morphology?
  - data visualisation
- Spectral line velocity, intensity, width?
- Gaussian fitting
- Moment analysis
- Emission mechanism?
- Spectral index
- Magnetic field properties?
  - Polarization angle & intensity, rotation measure

CSIRO

Step 2: Look at your u-v data

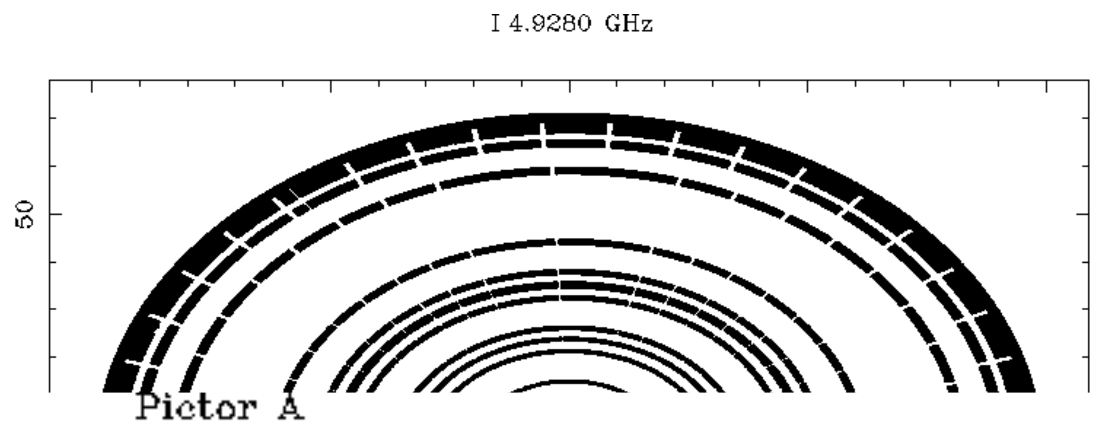
- You can tell a lot without even making an image, so look at your u-v data:
  - Amplitude vs. time
  - Amplitude vs. u-v distance
- At times this may be the only way to get something from your data, particularly if:
  - You have poor u-v coverage
  - You are only interested in variability

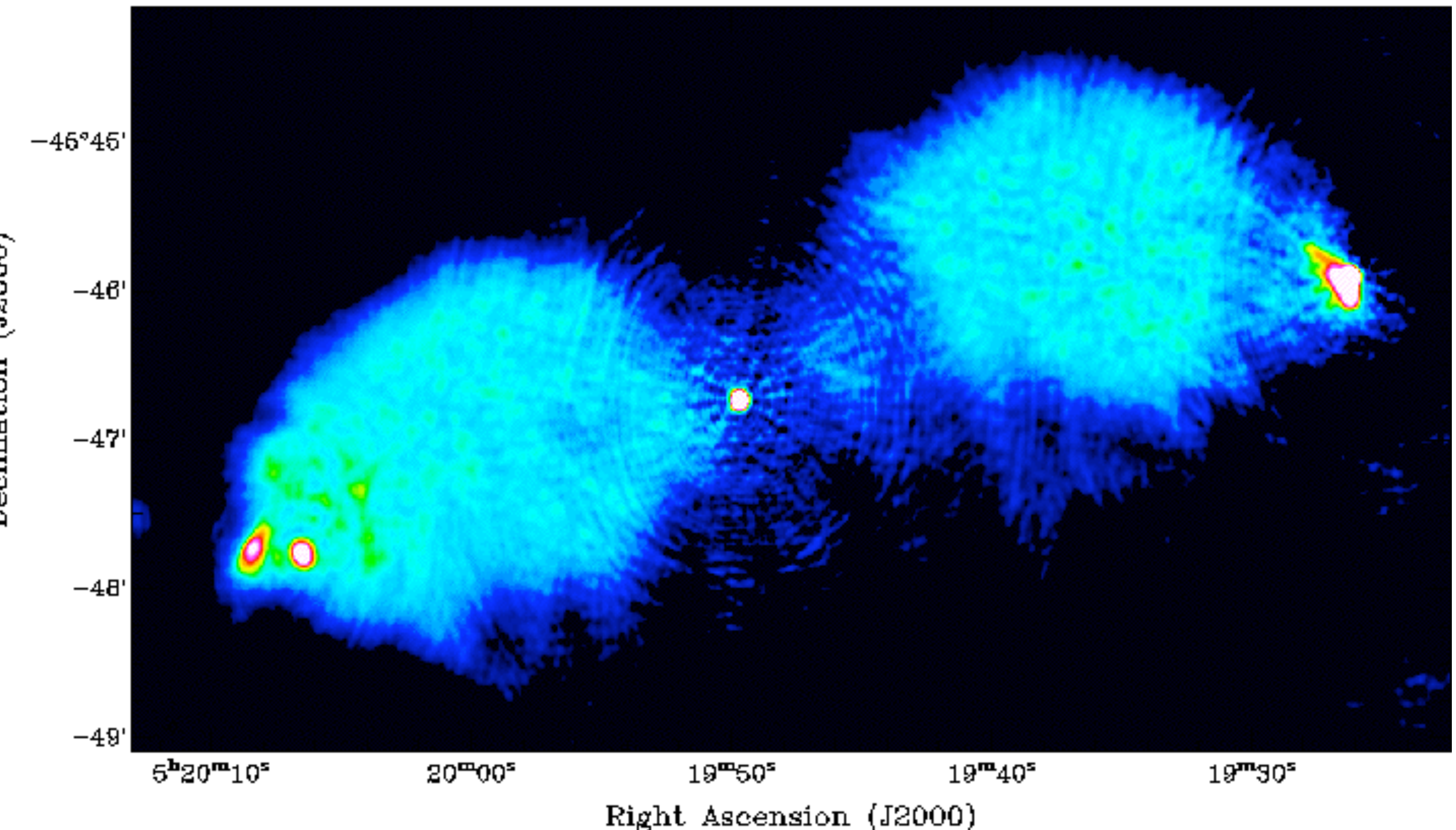


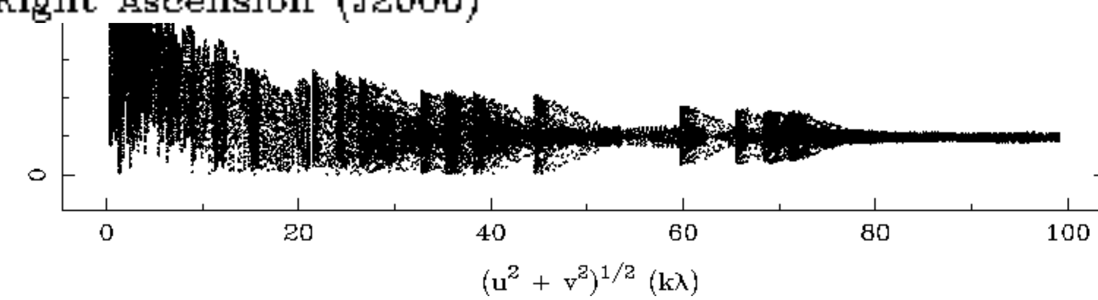


Step 2: Look at your u-v data

- You can tell a lot without even making an im your u-v data
  - Amplitude v
  - Amplitude v
- At times this way to get sc your data, pa
  - You have po
  - You are only variability







Non-imaging Analysis

- Synthesis imaging is an inverse-problem:
  - We have limited data, a known instrument and we try to infer the sky distribution
- The forward problem is much simpler:
  - Given a known sky distribution and a known instrument we can predict the visibilities.
- If you have a relatively simple source you can fit a model:
  - Define a parametric model for the sky distribution
  - Predict the visibilities for your telescope
  - Adjust the parameters of the model to fit your visibilities

Model Fitting

- The model, *F*, should fit the data, *V*

$$V(u,v) = F(u,v;a_1,...,a_m) + \text{noise}$$

- The likelihood is (for Gaussian errors)

$$L \propto \prod_{i=1}^{n} \left\{ \exp\left[-\frac{1}{2} \left(\frac{V_i - F(u_i,v_i;a_1,...,a_m)}{\sigma_i} \right)^2 \right] \right\}$$

- Maximise -Log(*L*) or minimise the chi-squared for the model fit

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{V_i - F(u_i,v_i;a_1,...,a_m)}{\sigma_i} \right)^2$$

Model Fitting

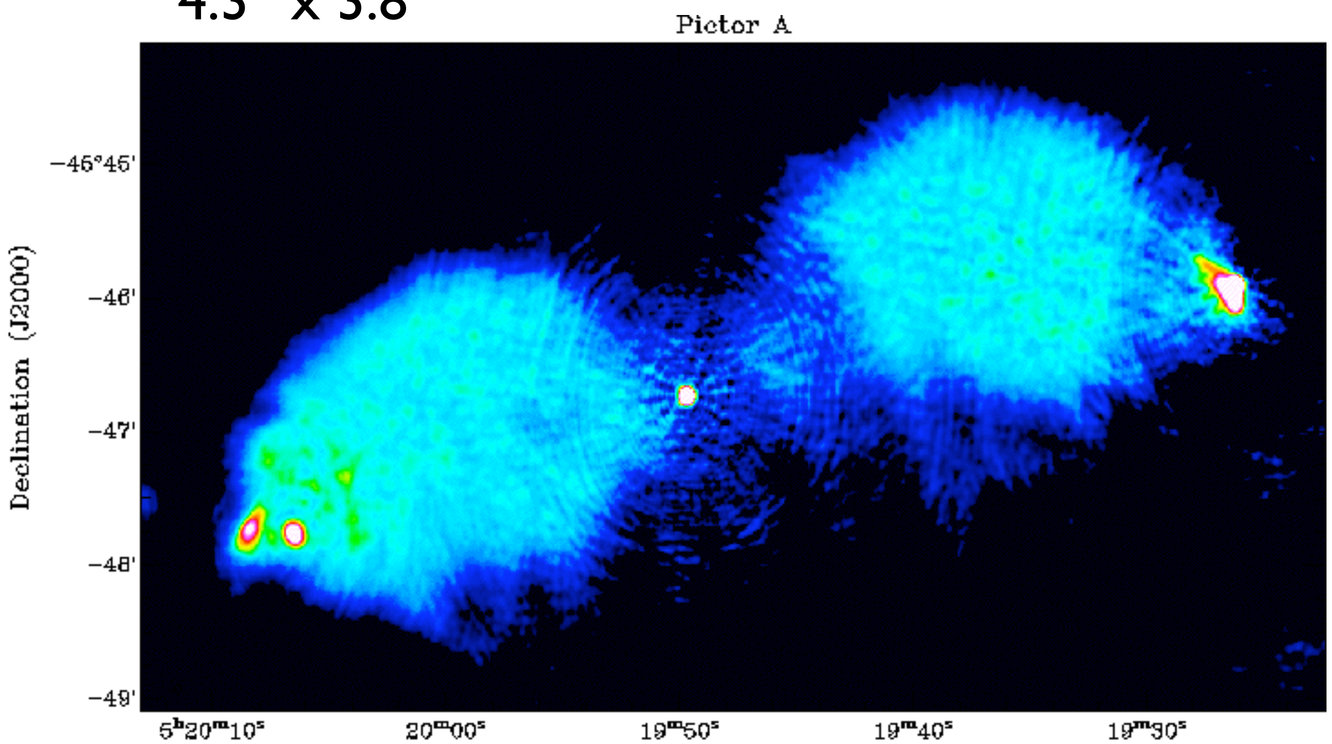
- Model fitting works best for sky brightness distributions that can be represented by a simple model with few parameters
  - Checking your calibrator - is it a point source?
  - Finding positions and fluxes for a few point sources
    - Estimates found this way often are more reliable than estimates from a deconvolved image because the errors are more predictable
  - Modelling and subtracting a bright source before imaging and deconvolution can give a better deconvolution
- Some miriad tasks that do model fitting are:
  - uvflux: fits a simple point source
  - uvfit: fits point sources, gaussians, disks, etc.

Step 3: Image your data appropriately

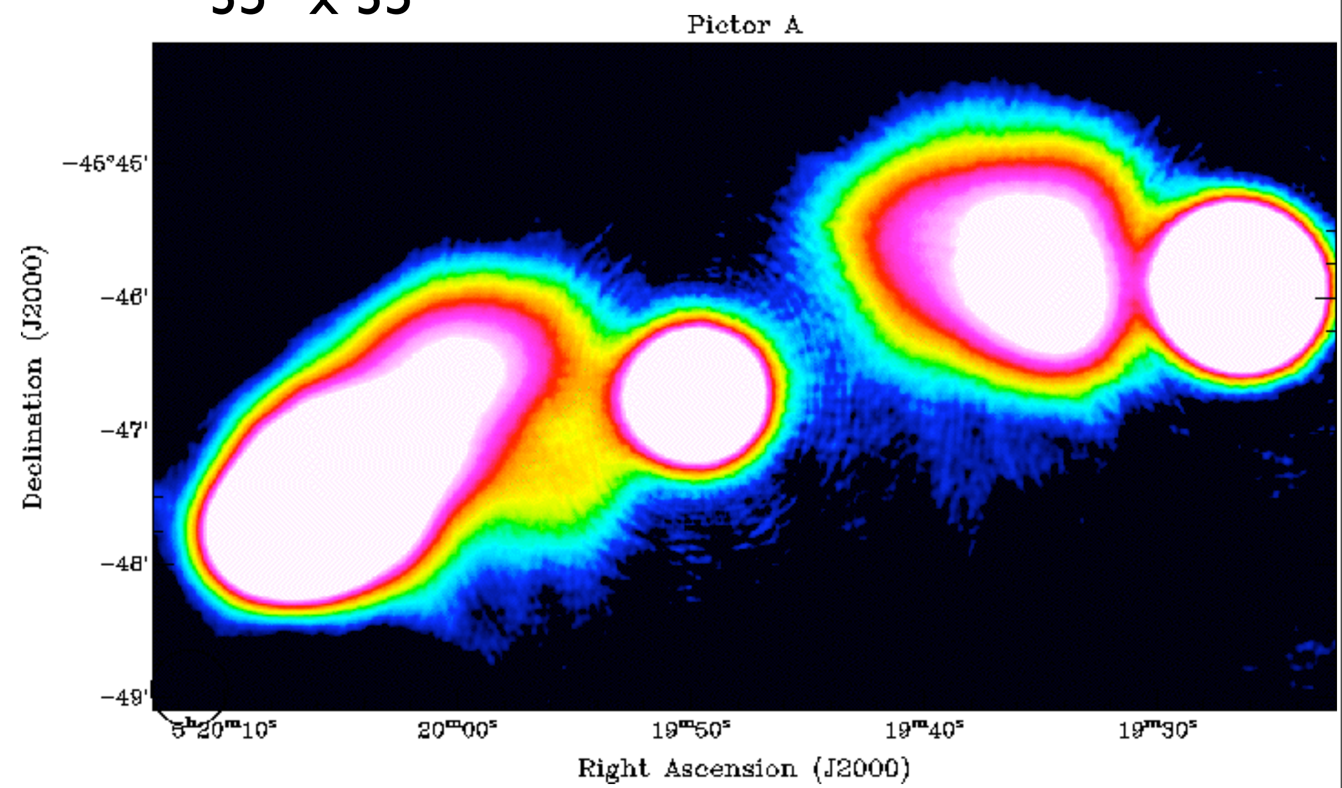
- Do you want to know the position of a source?
  - Image with as small a beam as possible
- Do you want to know the total flux?
  - Image with as large a beam as possible
- Do you want to know the peak flux?
  - Image with a small beam, run a high-pass filter over the image
- Do you want to look for low level emission?
  - Smooth the image

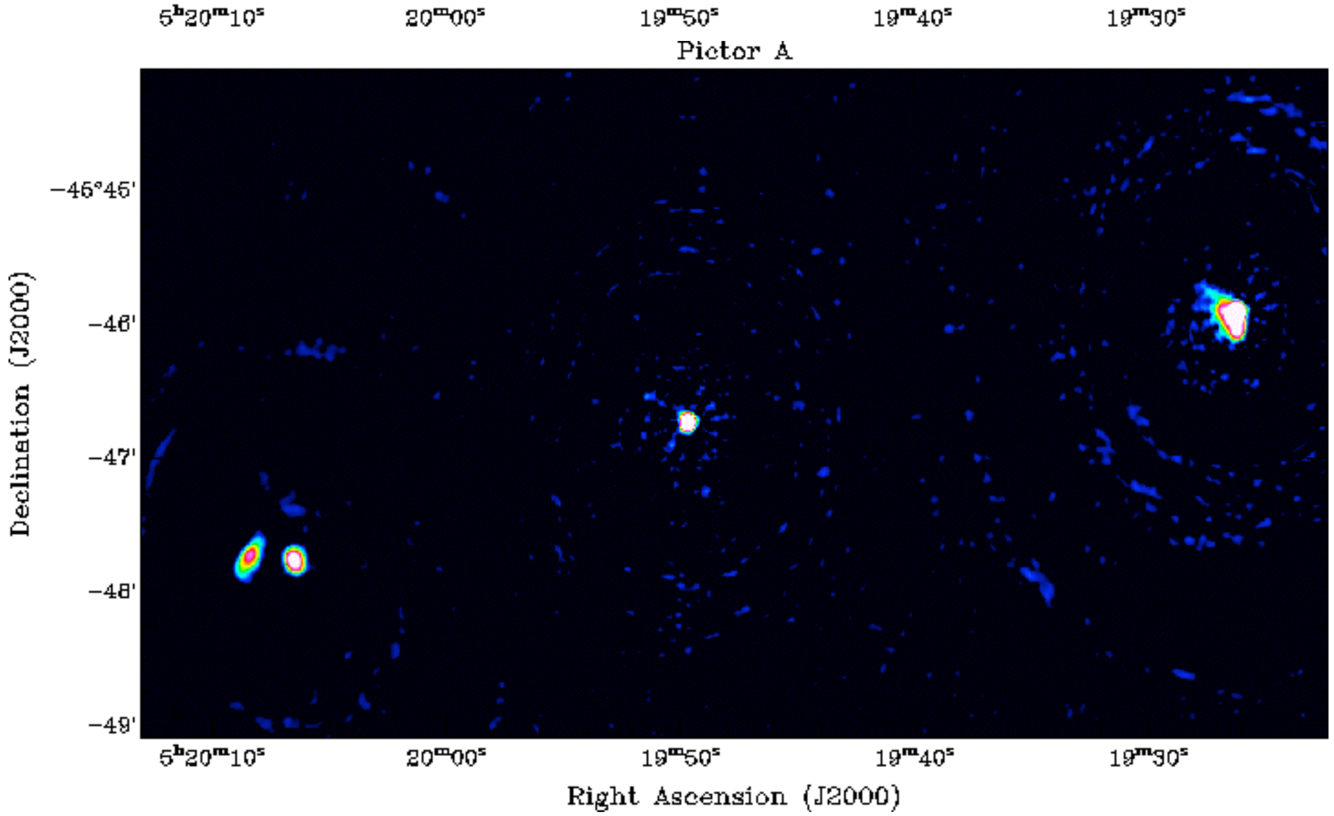
Image Manipulation

4.3" x 3.8"



35" x 35"



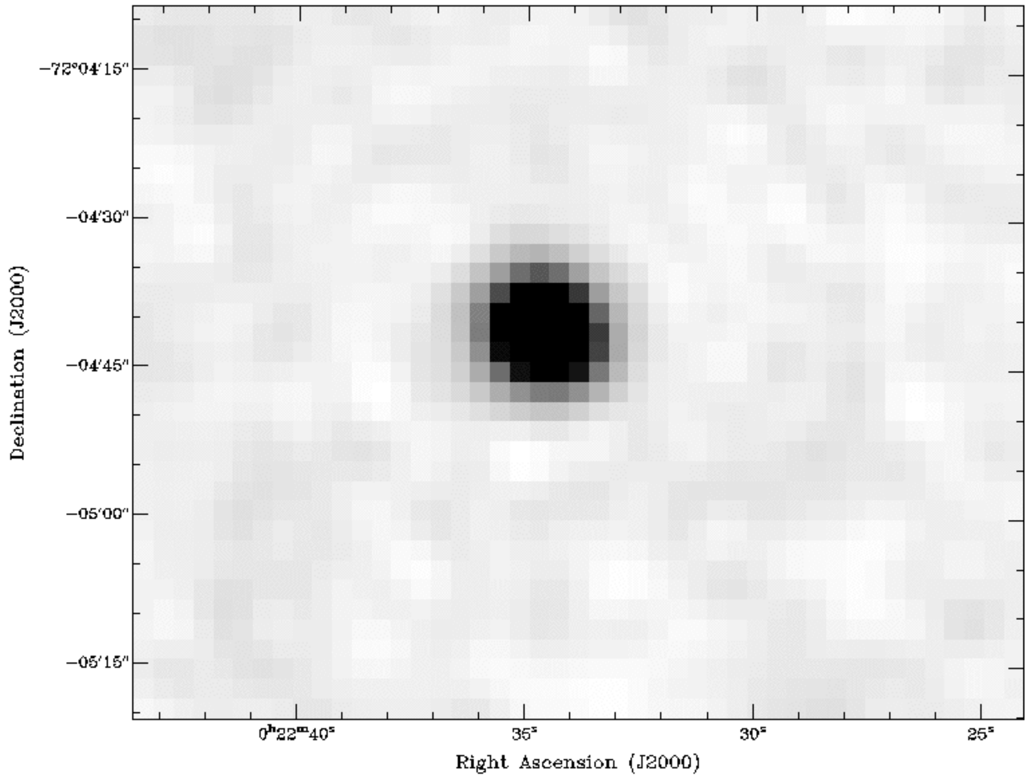


4.3" x 3.8" high-pass filter

Each image emphasizes a different aspect of the data:

- Hot spots plus large scale emission
- Total flux
- Peak intensities

Simple source parameters



- If you want to know the flux and position of a simple source you have options:
  - Estimate from *kvis* peak value
  - Use miriad's *maxfit* & *imfit*

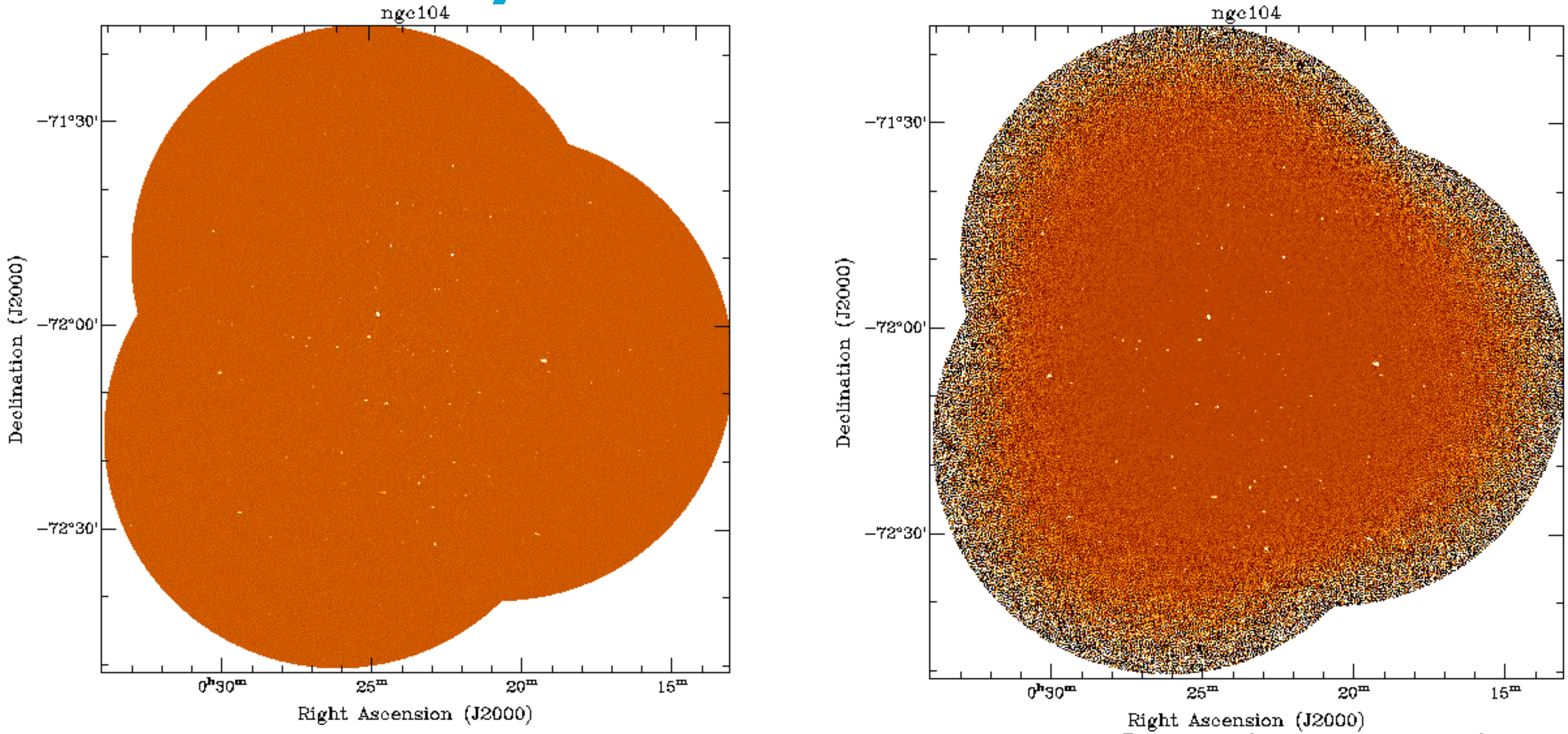
```
imfit% inp
  Task:   imfit
  in         = 47tuc_sub.mir
  region     = boxes(219,333,263,368)
  clip       =
  object     = gaussian
  spar       =
  fix        =
  out        =
  options    =
imfit% go
ImFit: version 1.0 30-Jun-99
-------------------------------------------------
RMS residual is 1.27E-05

Using the following beam parameters when
deconvolving and converting to integrated flux
  Beam Major, minor axes (arcsec):     8.92     8.42
  Beam Position angle (degrees):     44.4

Scaling error estimates by  4.6 to account for
noise correlation between pixels

Source  1, Object type: gaussian
  Peak value:                 9.5209E-04 +/-  2.4637E-05
  Total integrated flux:      9.9578E-04
  Offset Position (arcsec):    -418.425    11.647
  Positional errors (arcsec):     0.100     0.097
  Right Ascension:                0:22:34.634
  Declination:                   -72:04:41.060
  Major axis (arcsec):          9.266 +/-   0.243
  Minor axis (arcsec):          8.472 +/-   0.222
  Position angle (degrees):     55.72 +/-       11.85
  Deconvolved Major, minor axes (arcsec):    2.670     0.380
  Deconvolved Position angle (degrees):     69.8
-------------------------------------------------
```

Primary Beam Correction

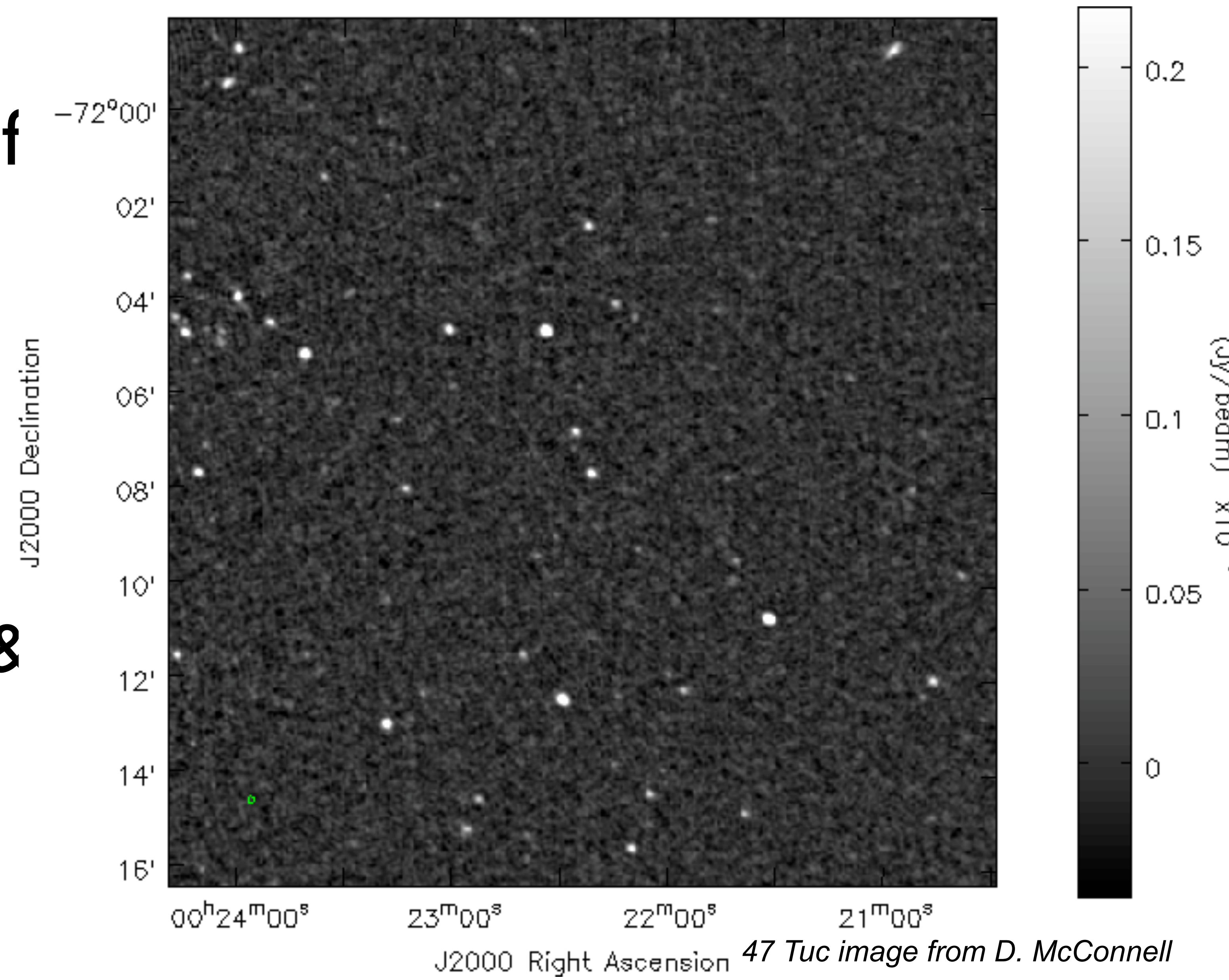


Primary beam corrected

- Beware the limitations of your image!
  - Your image is weighted by the primary beam sensitivity
  - Divide by the primary beam to restore sources to their expected brightness increases the noise at the edge of the image

Source Finding & Characterisation

- There are a number of tools for automated source finding and cataloging:
  - miriad's *imsad*
  - casapy's *findsources* & *fitsky*
- Limitations when images have artefacts or sources are extended



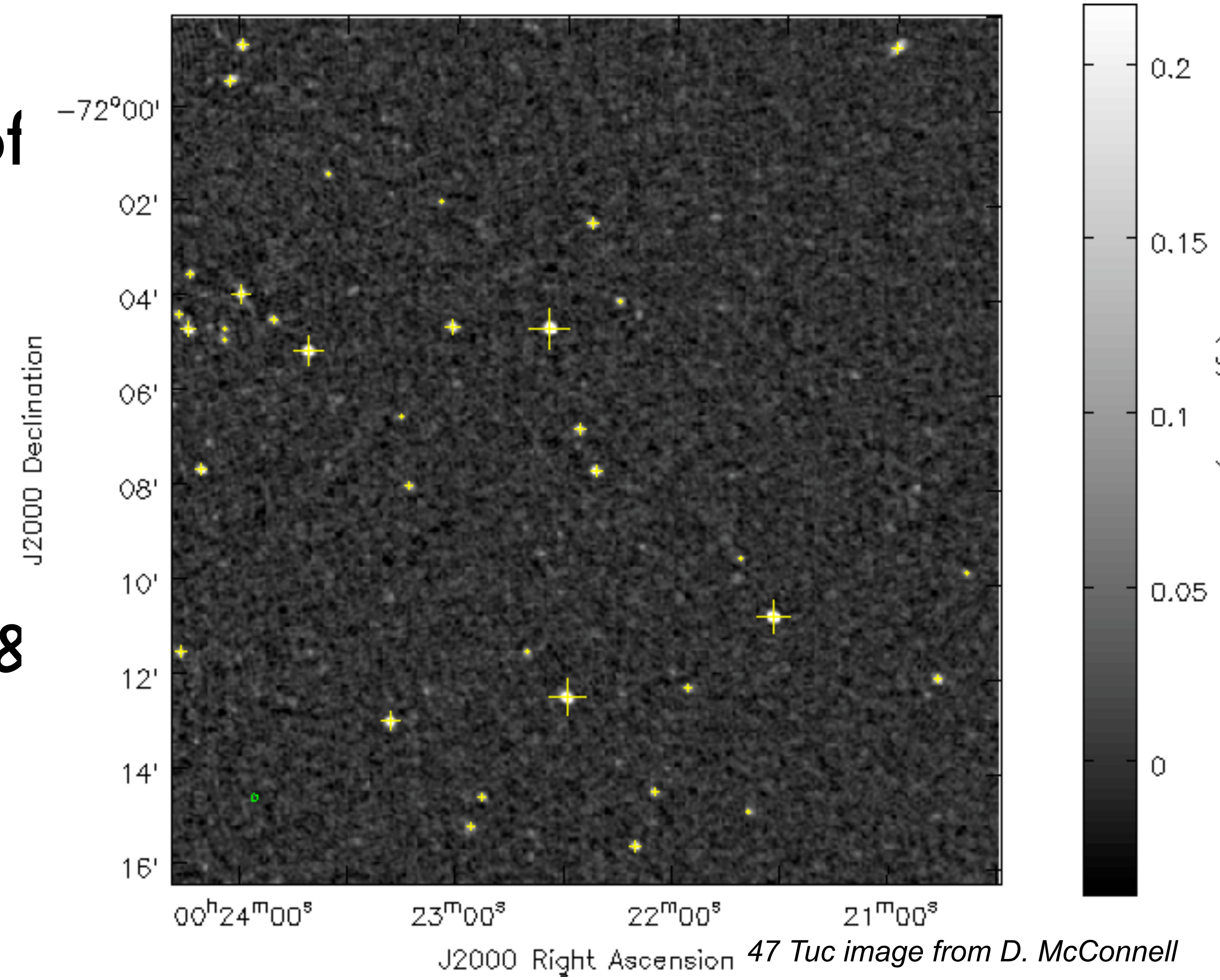
47 Tuc image from D. McConnell

Example from casapy *findsources*:

```
ia.open("47tuc.sub.img")
cl=ia.findsources(nmax=100,cutoff=0.1)
```

Source Finding & Characterisation

- There are a number of tools for automated source finding and cataloging:
  - miriad's *imsad*
  - casapy's *findsources & fitsky*
- Limitations when images have artefacts or sources are extended



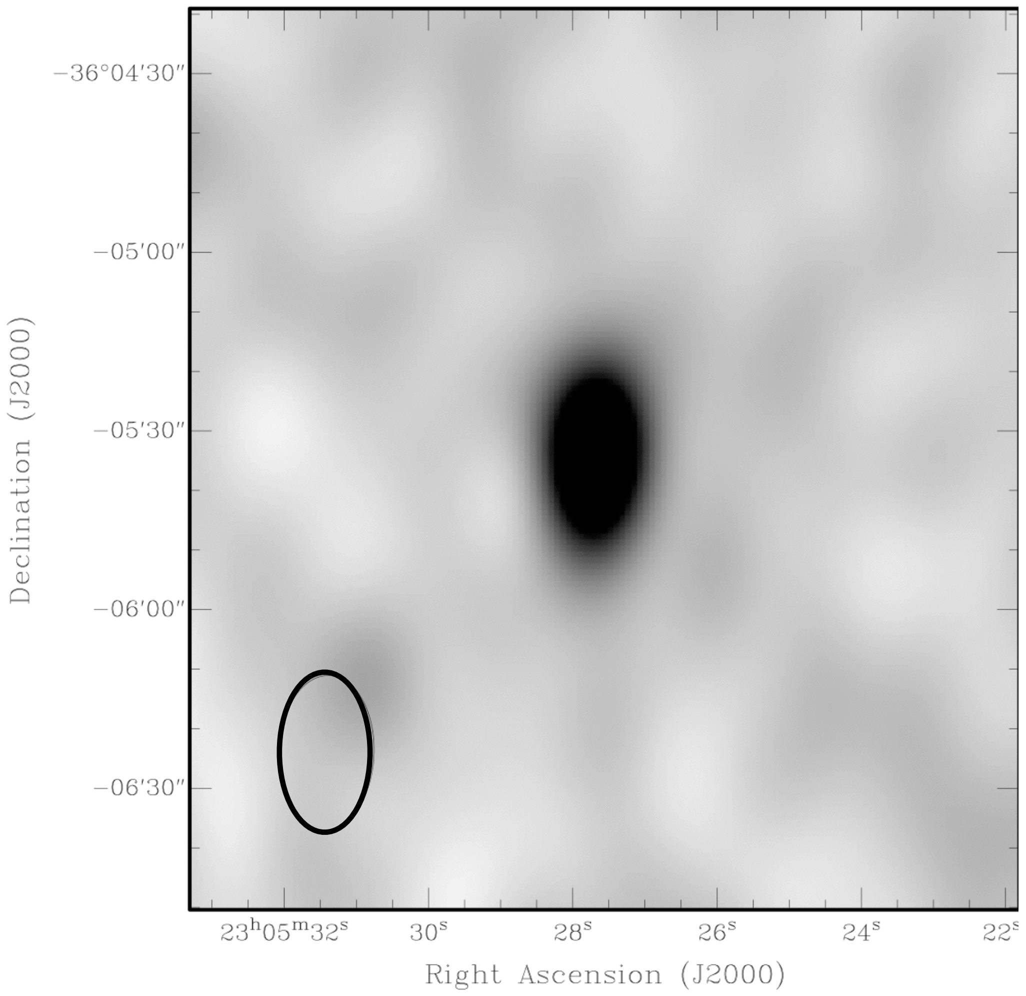
47 Tuc image from D. McConnell

Example from casapy *findsources*:

```
ia.open("47tuc.sub.img")
cl=ia.findsources(nmax=100,cutoff=0.1)
```

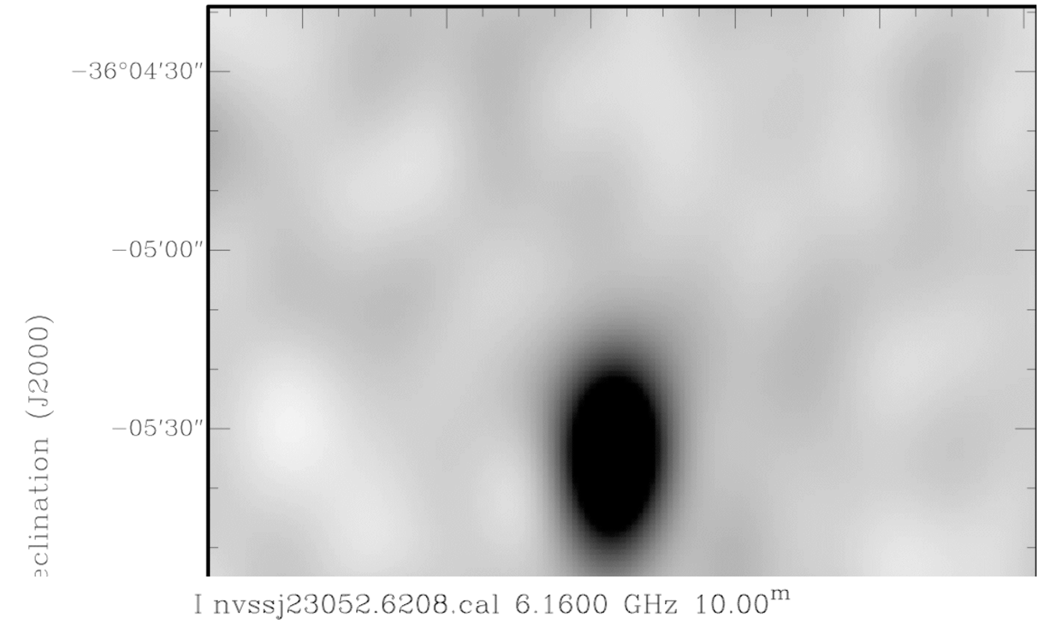
Source size: the subtleties

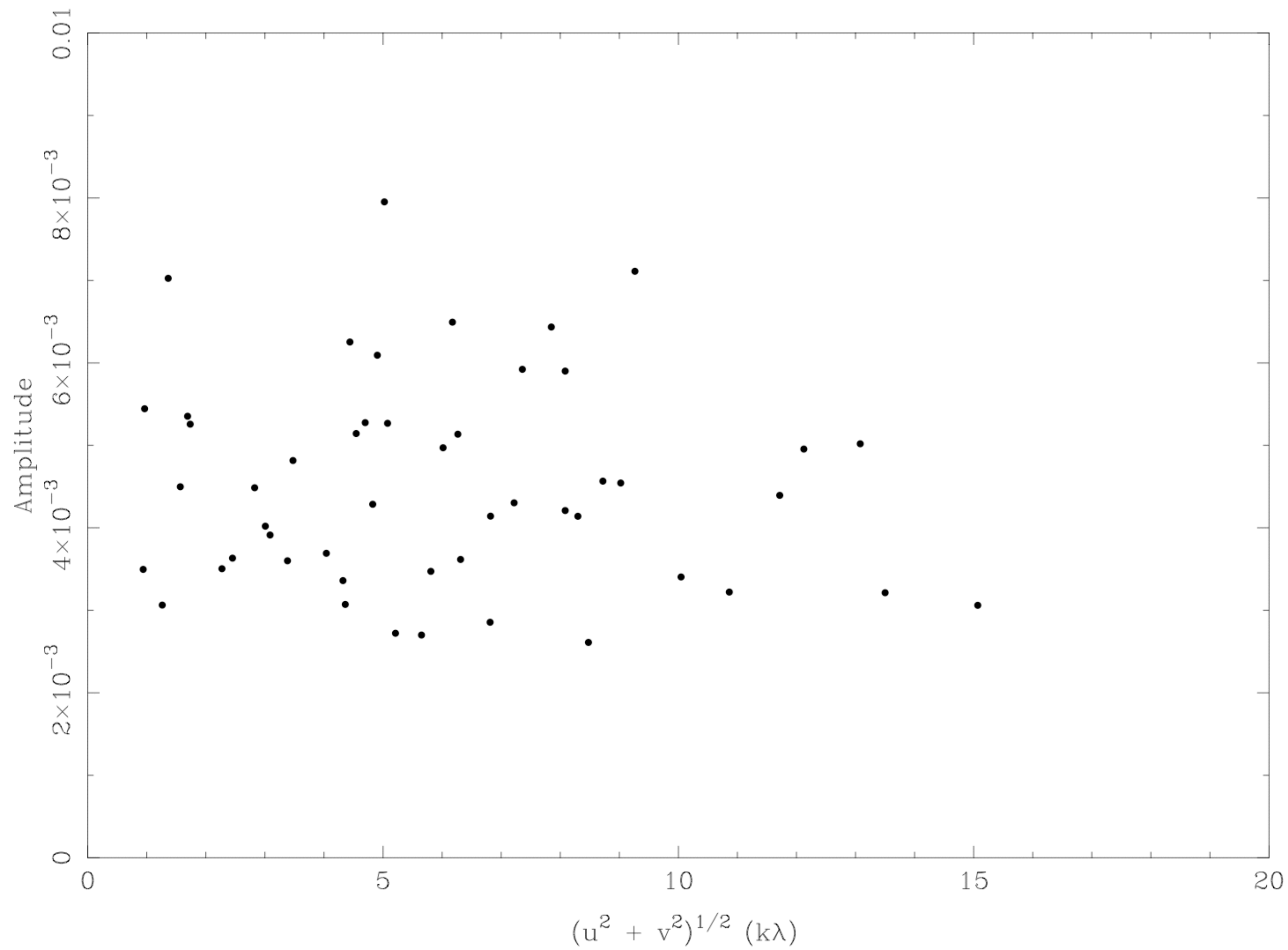
- When imaged the source appears unresolved
  - From the image the source size is < 17" x 21"



Source size: the subtleties

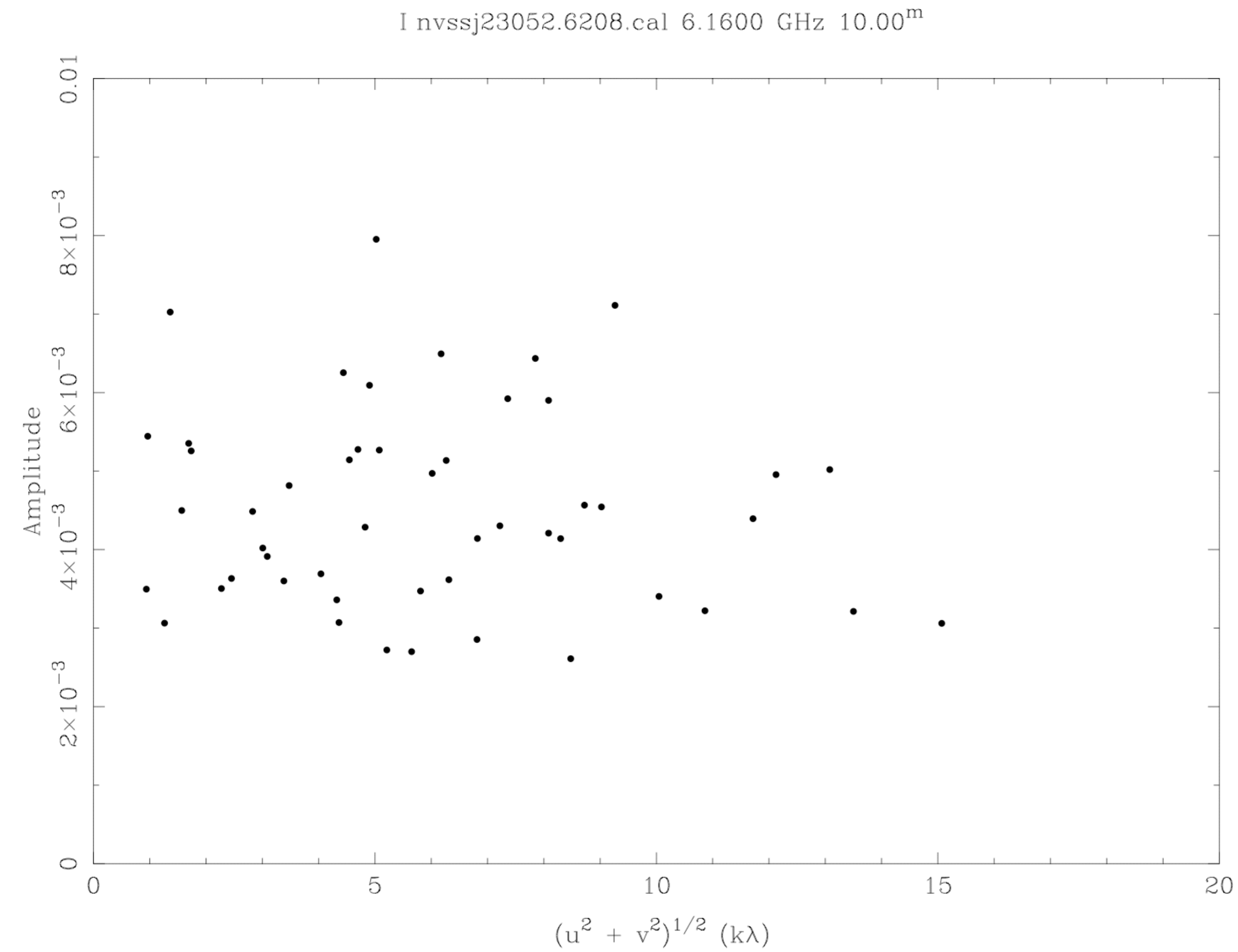
- When imaged the source appears unresolved
  - From the image the source size is < 17” x 21”
- The visibilities tell you that the amplitude is flat to ~ 15 kλ





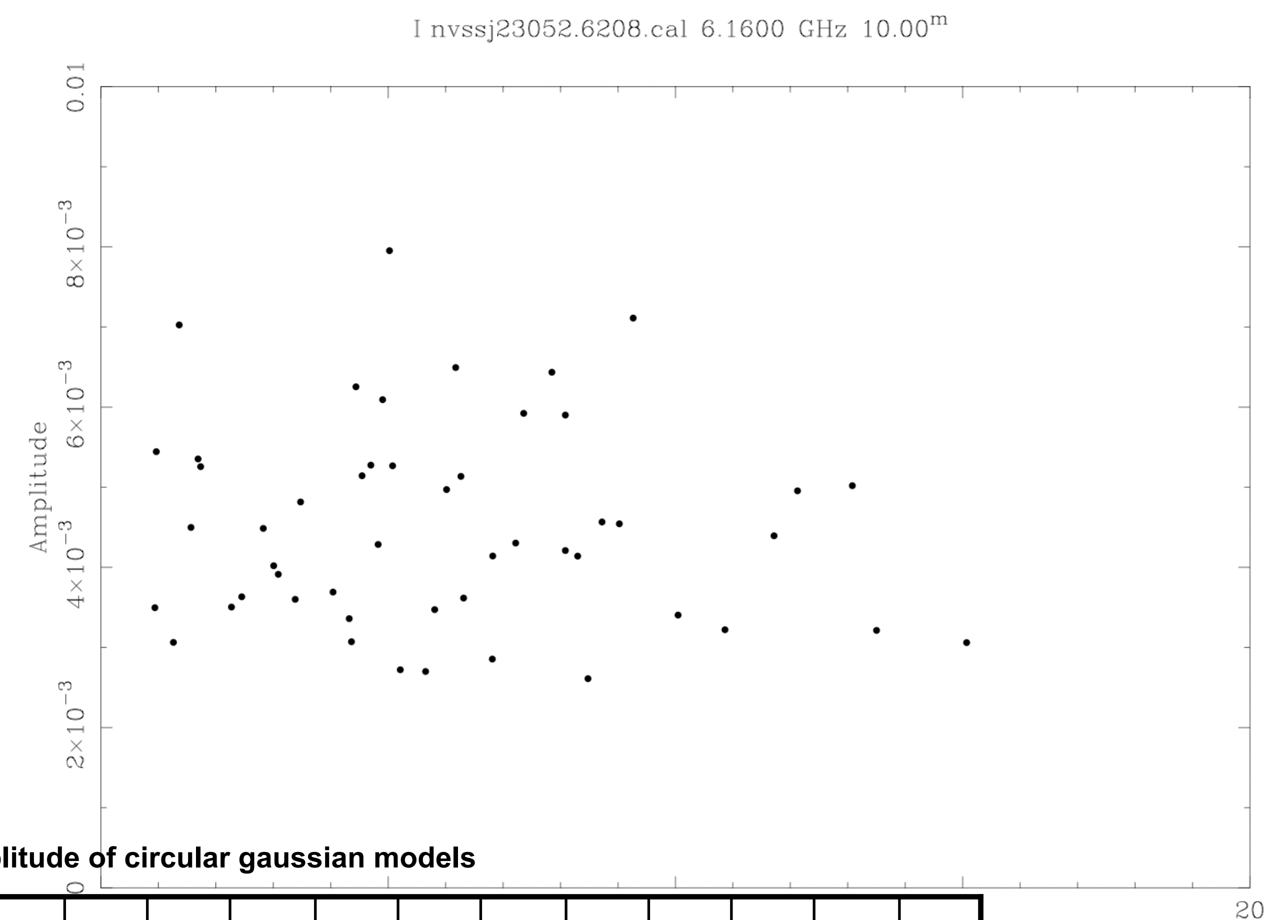
Source Size: the subtleties

- But by comparing the amplitudes vs uvdist with models you can actually constrain the source size

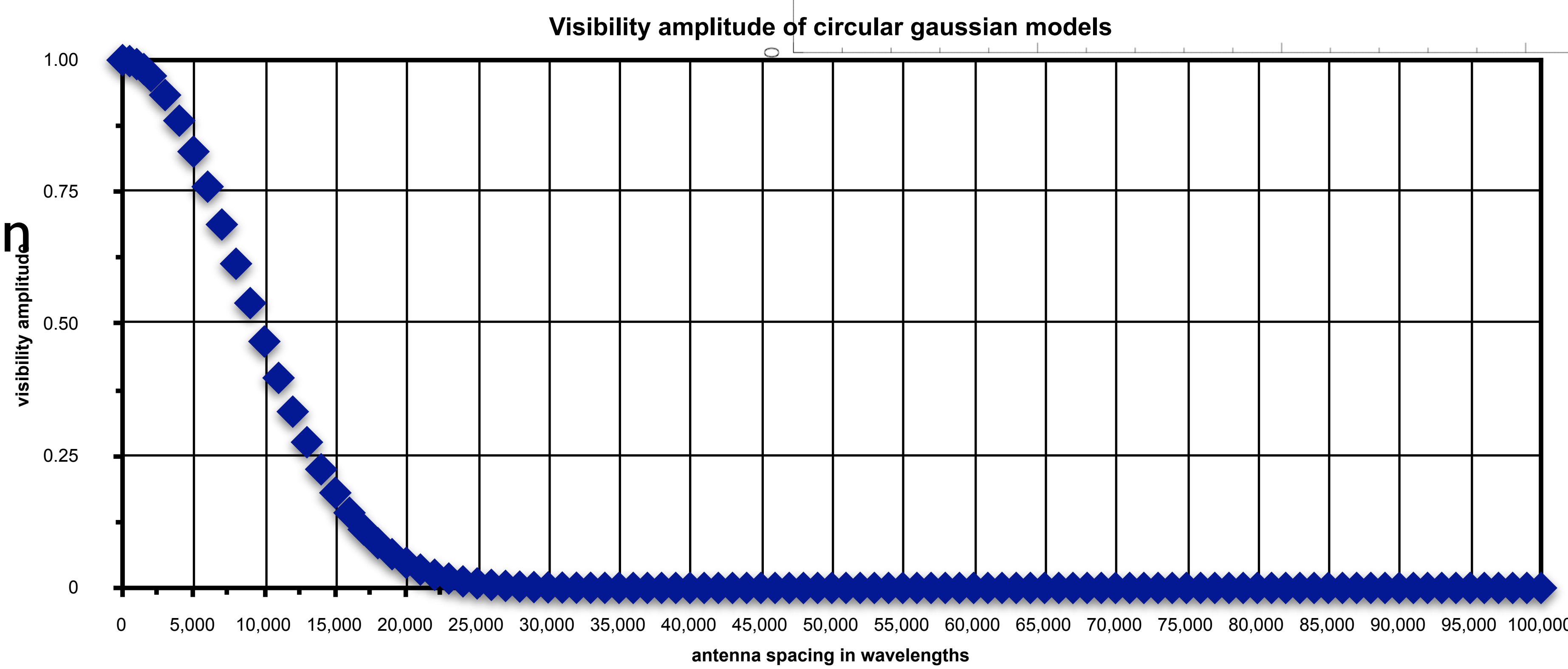


Source Size: the subtleties

- But by comparing the amplitudes vs uvdist with models you can actually constrain the source size

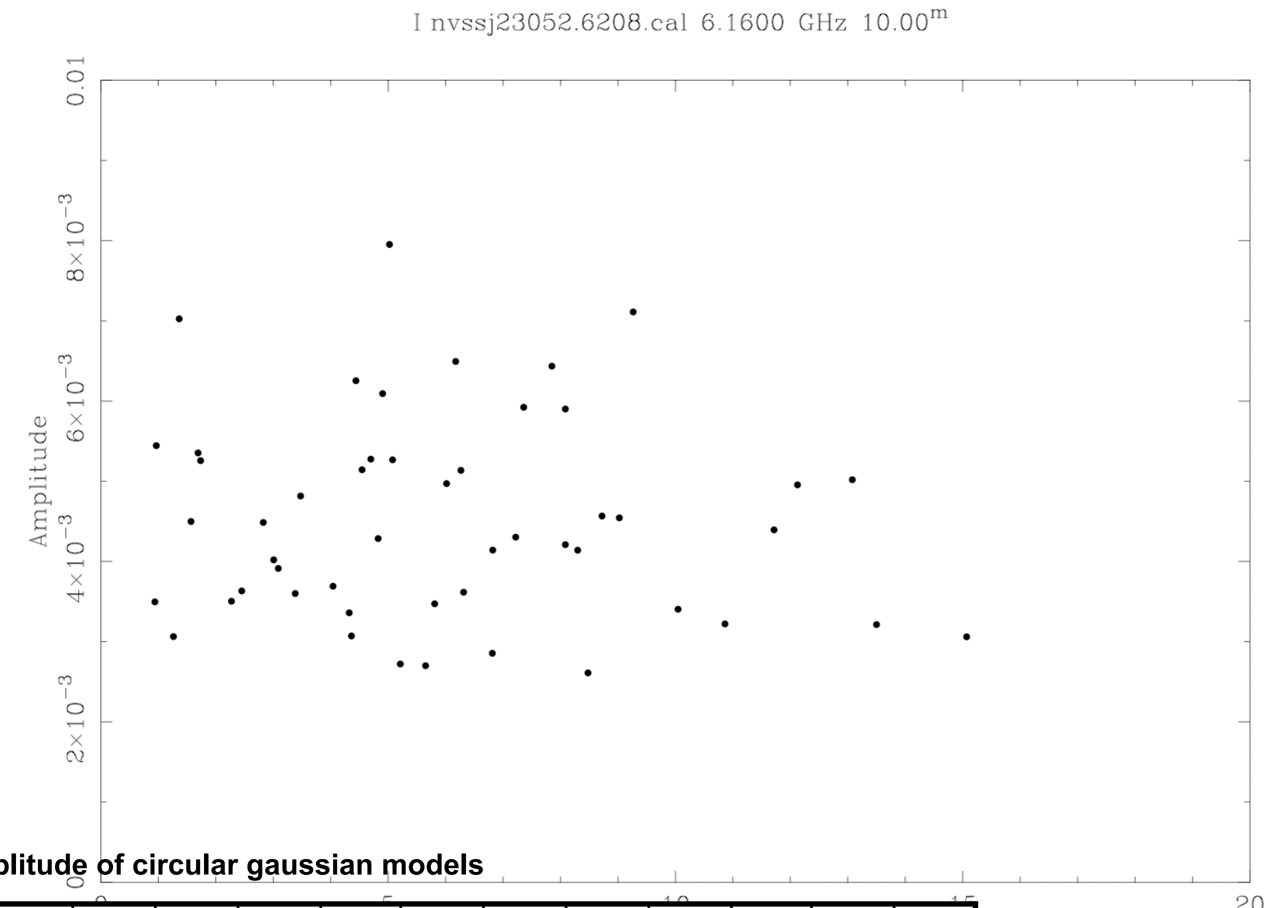


18” gaussian

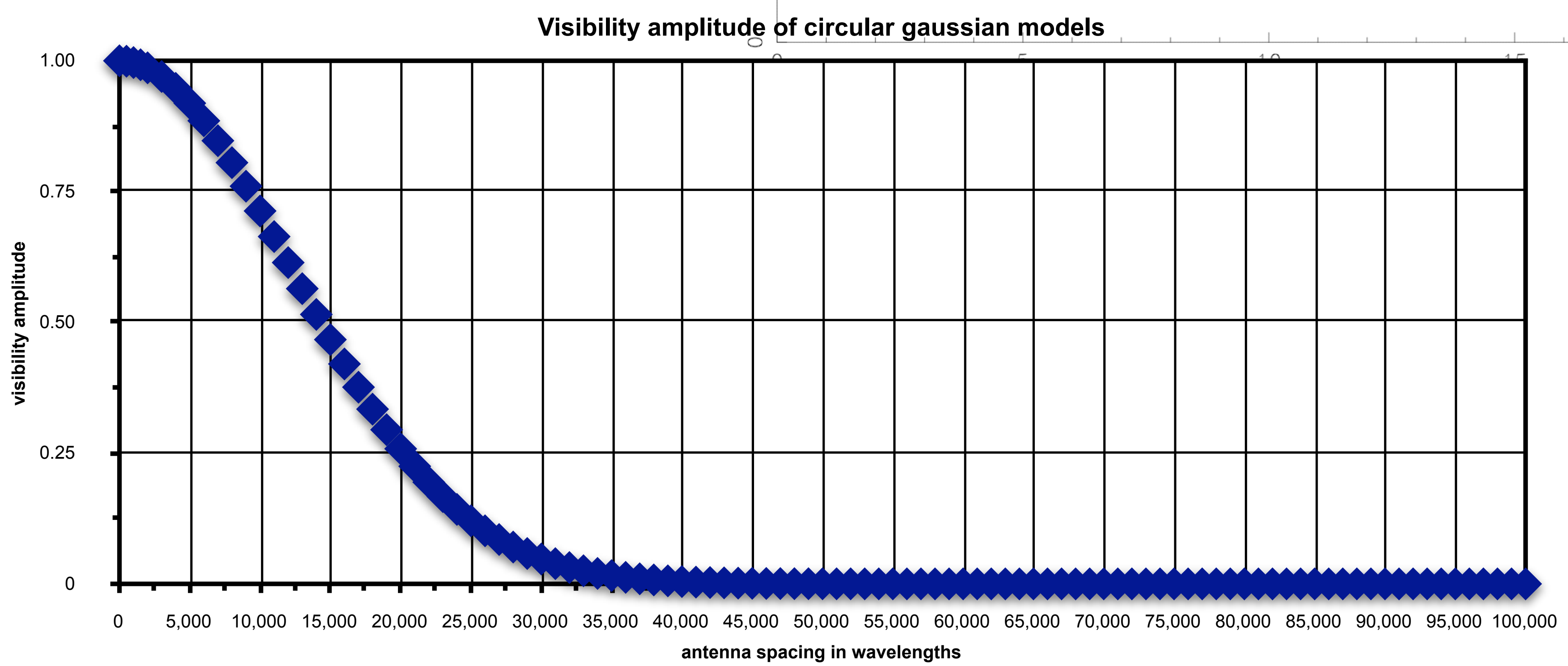


Source Size: the subtleties

- But by comparing the amplitudes vs uvdist with models you can actually constrain the source size

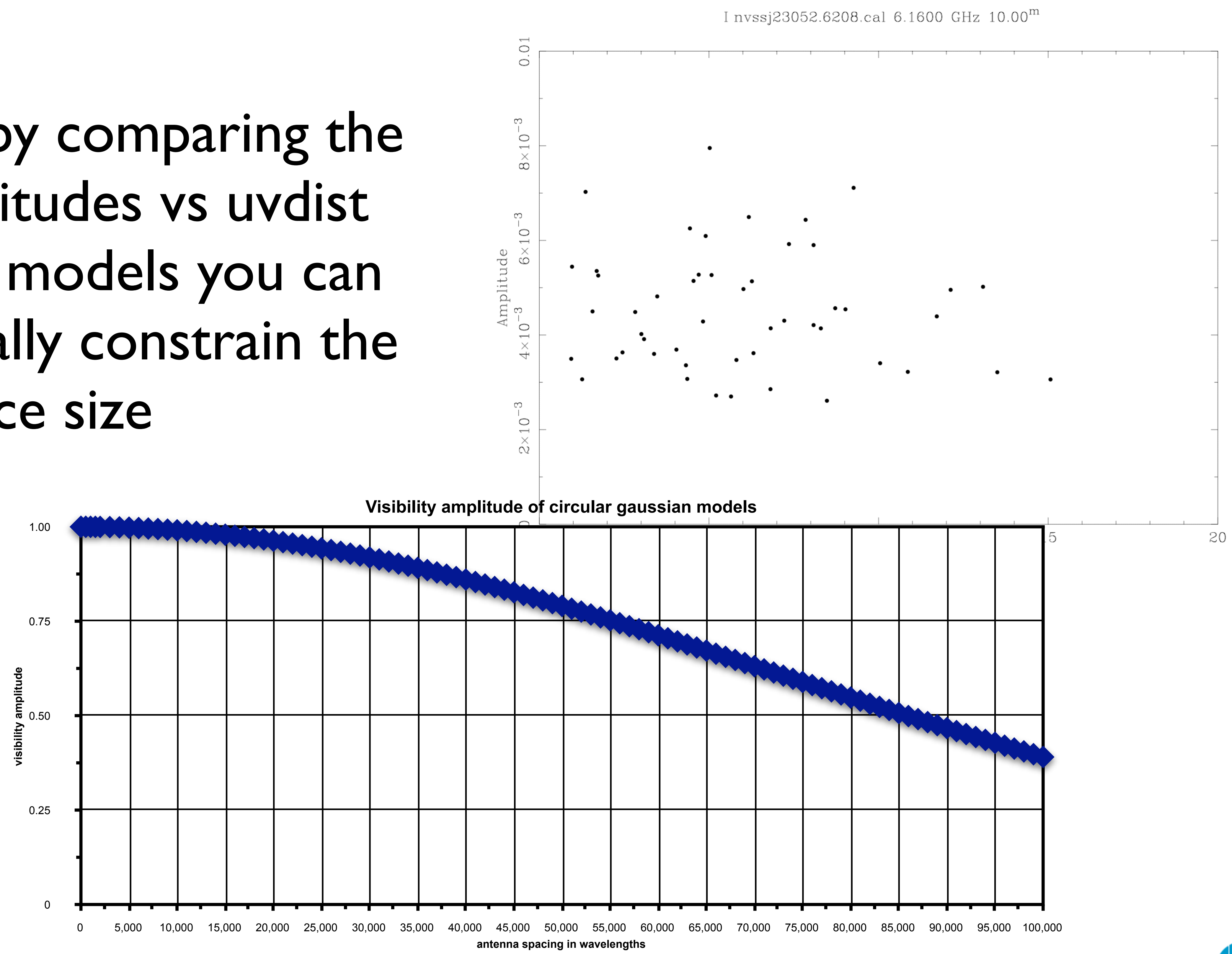


12" gaussian



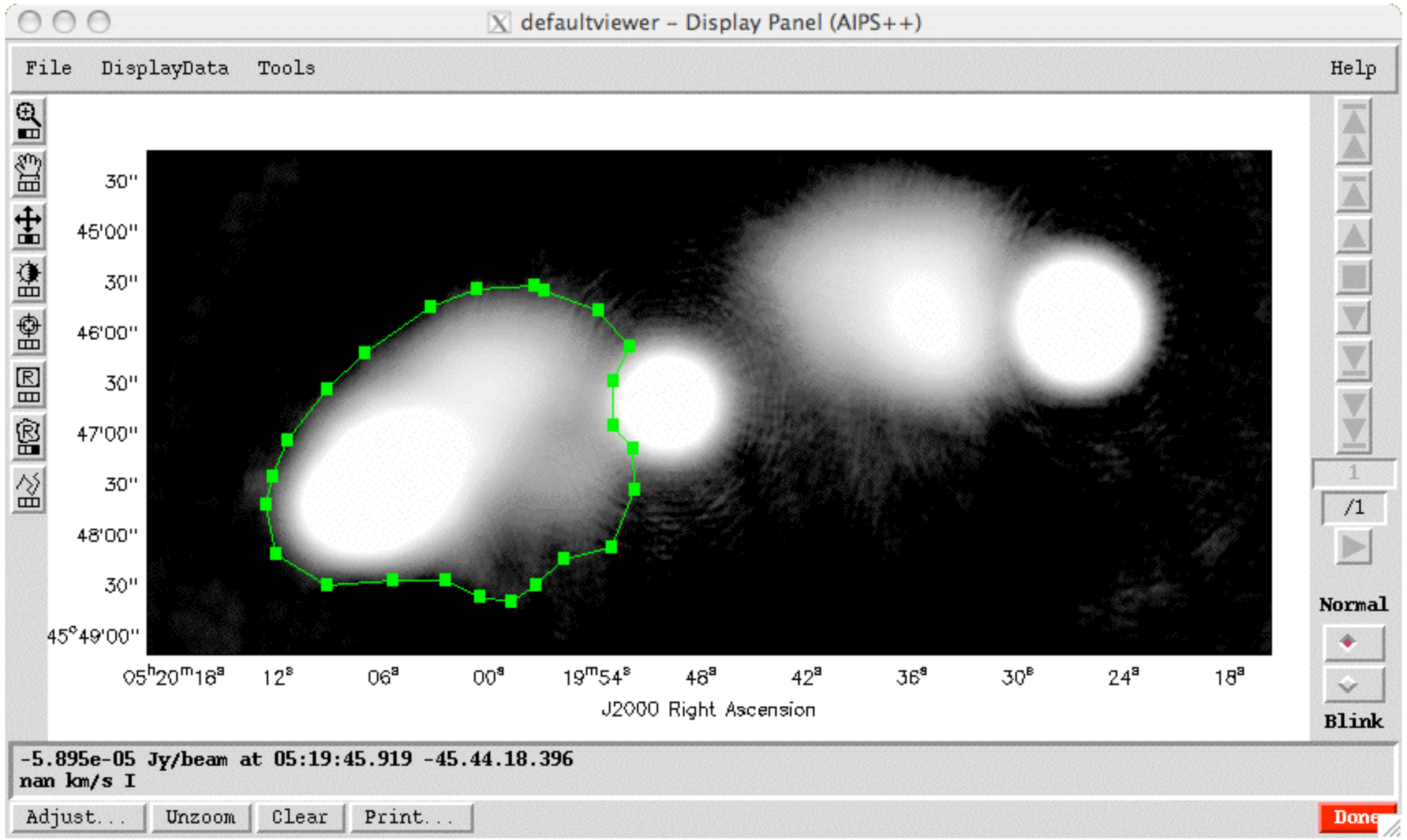
Source Size: the subtleties

- But by comparing the amplitudes vs uvdist with models you can actually constrain the source size



2” gaussian

Extended Source Flux

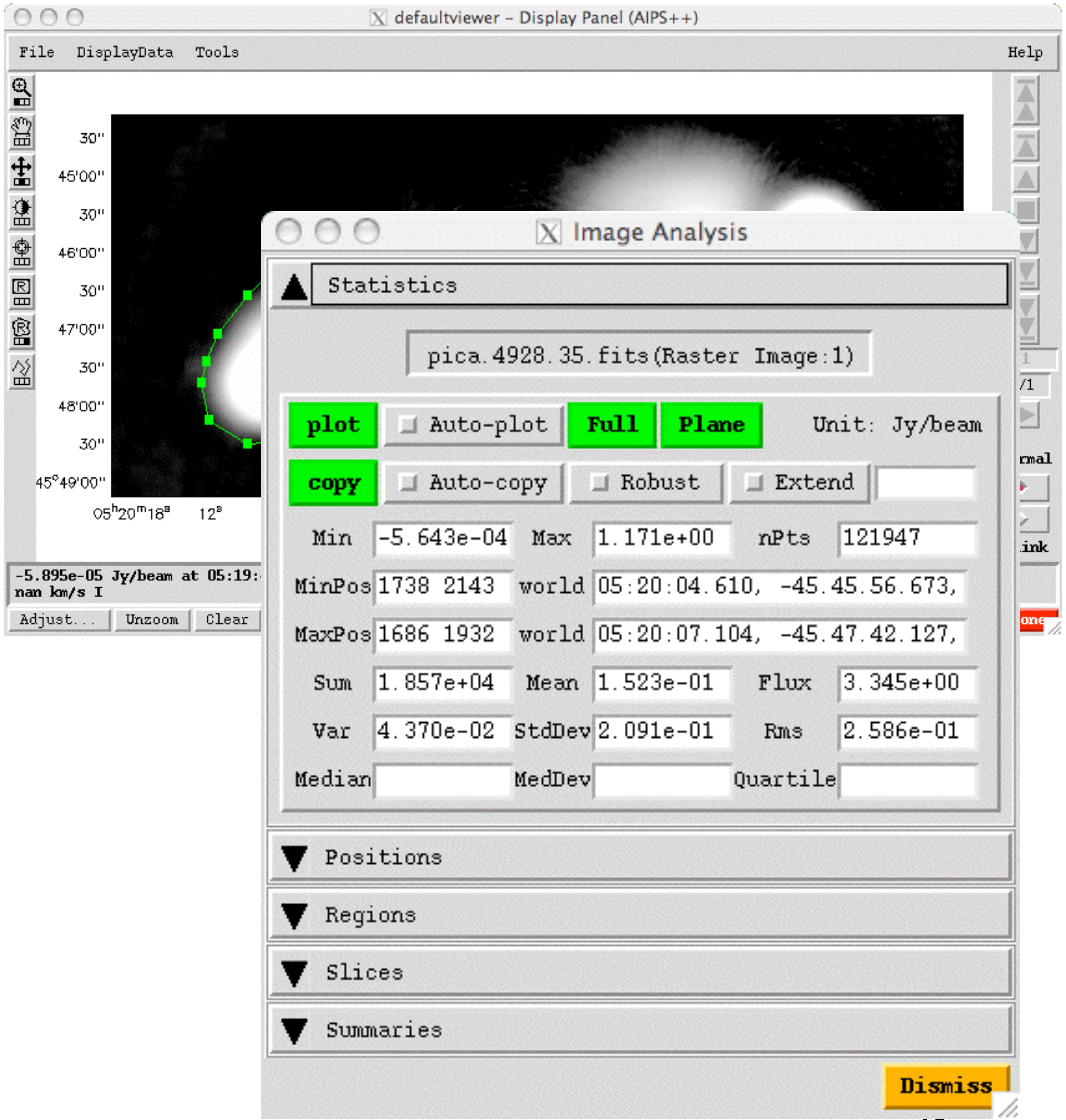


- To estimate the flux for extended sources, integrate over the whole source:

$$S_\nu[\mathrm{Jy}] = \frac{\int I_\nu[\mathrm{Jy\ bm}^{-1}]d\Omega}{\int d\Omega}$$

- If the source is on a background you may need to estimate its contribution and subtract
- For simple sources you can model the source to find the flux

Extended Source Flux



- To estimate the flux for extended sources, integrate over the whole source:

$$S_\nu[\mathrm{Jy}] = \frac{\int I_\nu[\mathrm{Jy\ bm^{-1}}]d\Omega}{\int d\Omega}$$

- If the source is on a background you may need to estimate its contribution and subtract
- For simple sources you can model the source to find the flux

Brightness Temperature

- It is often useful to characterise emission in terms of the brightness temperature, T_b, in Kelvin
  - Temperature of an equivalent blackbody having the intensity, I_ν, in Jy/Bm at a given frequency, ν

- The Rayleigh-Jeans approximation applies so for flux density, S, and synthesised beam, Ω,

$$I_\nu = \frac{S}{\Omega} = \frac{2\nu^2}{c^2} k T_b$$

Brightness Temperature II

- The synthesised beam is modelled as a circular Gaussian with FWHM of θ:

$$\Omega = \frac{\pi\theta^2_{\rm HPBW}}{4\ln 2}$$

- So the brightness temperature of a feature that fills the beam is:

$$T_b = \frac{2\ln 2}{\pi}\frac{\lambda^2}{k}\frac{I_\nu}{\theta^2_{\rm HPBW}}$$

- Or in sensible units: is:

$$T_b = 1.36\frac{\lambda[\rm cm]^2 I_\nu[\rm mJy\ Bm^{-1}]}{\theta_{\rm HPBW}[\rm arcsec]^2}$$

Estimating errors

- As with all measurements, they are meaningless without an estimate of the error
- Error estimates need to made for:
  - Source fluxes
  - Source positions
  - Source sizes
  - Combined images, such as polarization images
- Beware of errors returned from model fits as these are usually formal fitting errors
- Image flux statistics can be estimated from the image itself:
  - miriad's imstat
  - kvis "s" key
  - casapy statistics

Position and size errors

- Absolute positional accuracy depends on the quality of your data/calibration
  - Precision of the position of the phase cal (secondary)
  - Separation of phase cal from source - the closer the better!
  - Weather, phase stability
  - Signal-to-noise
  - Beware of positions after self-cal
- Relative positions and motions
  - Limited by signal to noise

Rough Error Estimates:

P = Component Peak Flux Density

σ = Image rms noise

P/σ = signal to noise = S

B = Synthesized beam size

θ_i = Component image size

ΔP = Peak error = σ

ΔX = Position error = $B / 2S$

$\Delta\theta_i$ = Component image size error
= $B / 2S$

θ_t = True component size
= $(\theta_i^2 - B^2)^{1/2}$

$\Delta\theta_t$ = Minimum component size
= $B / S^{1/2}$

From E. Fomalont

Errors on extended source flux estimates

- Beware of missing short spacings:
  - total flux measurements will be lower limits
- Your estimation of the background level and the region you choose to integrate over determines your error in flux measurements

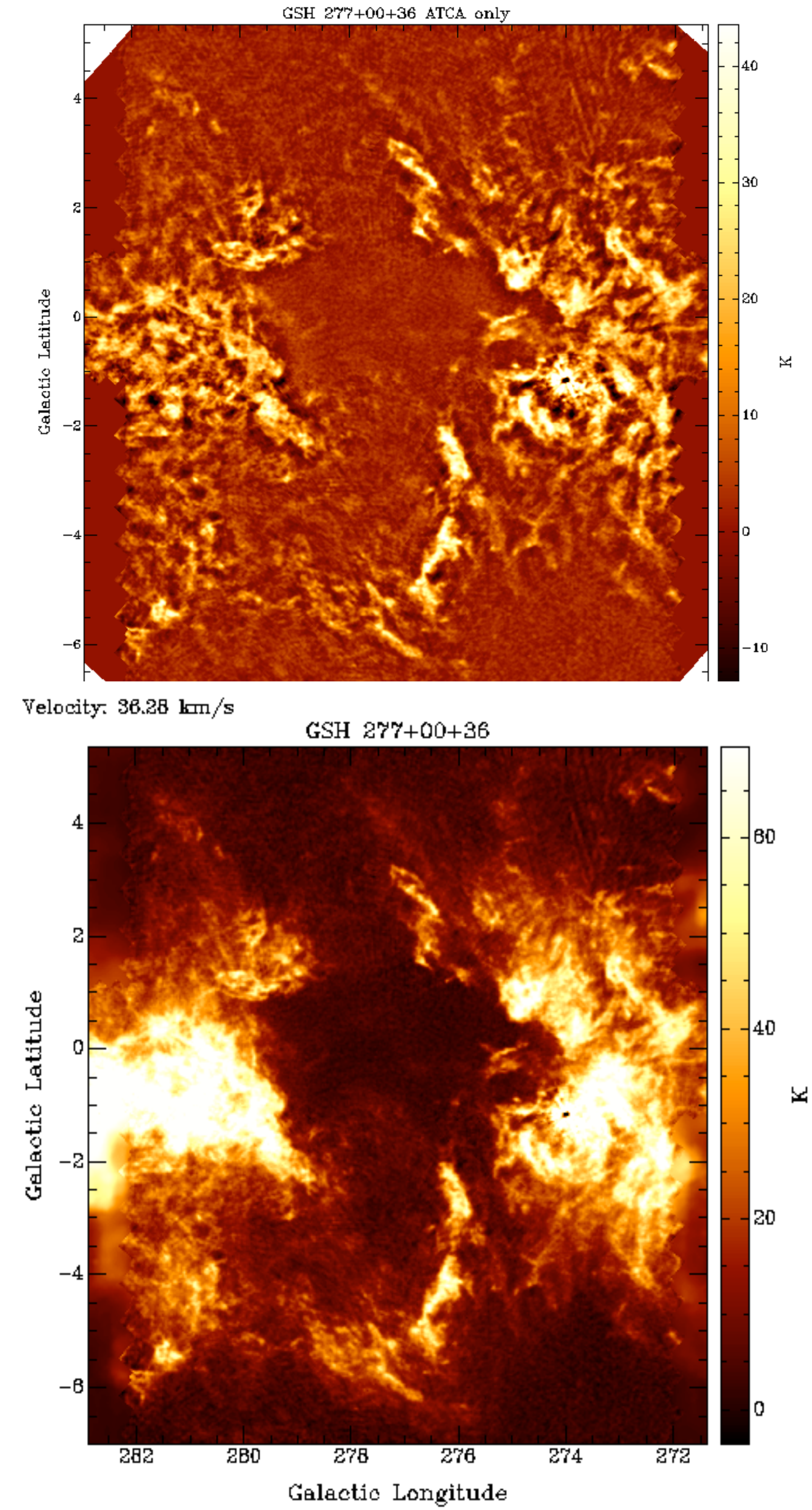
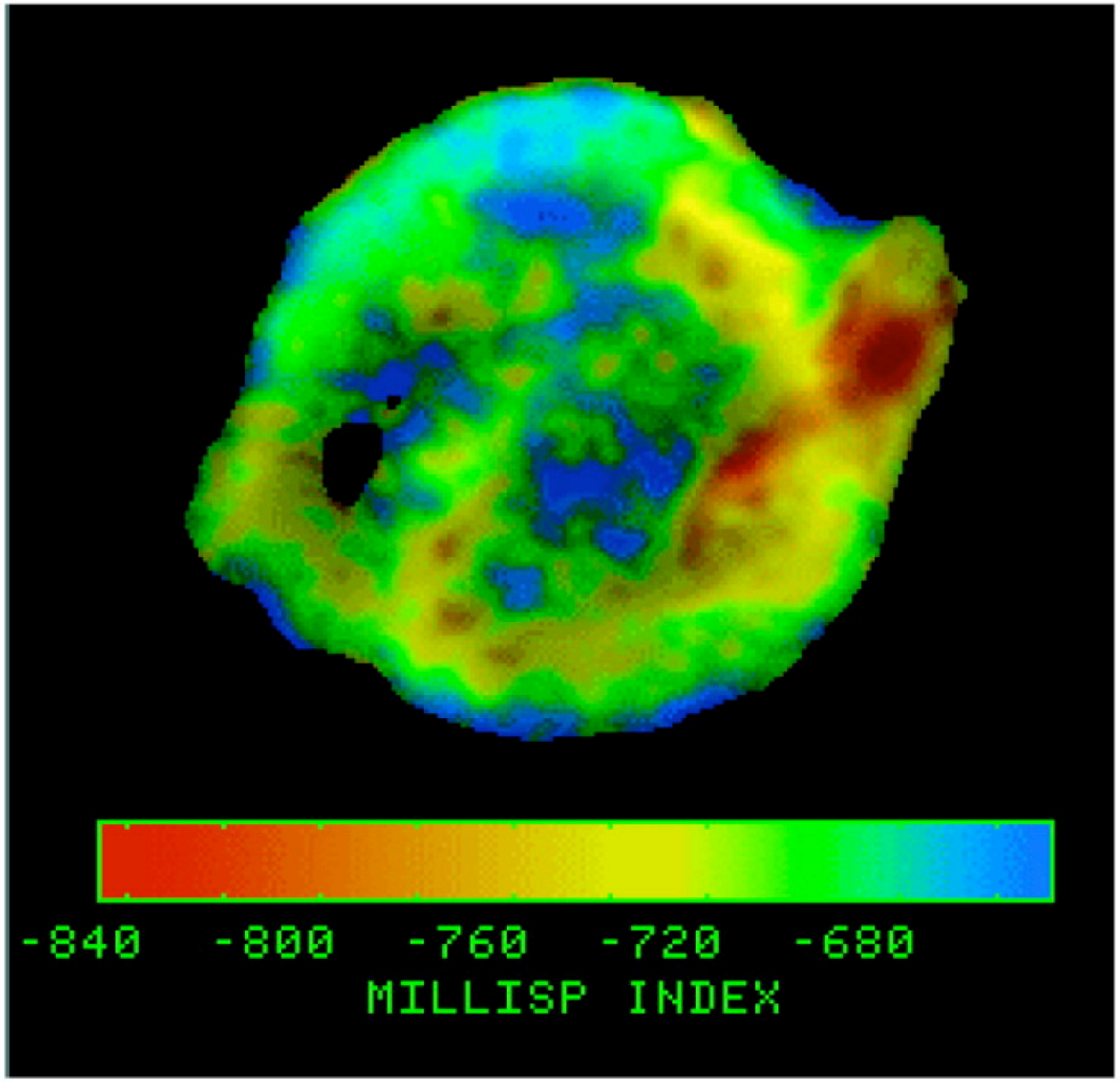


Image combination: spectral index

- Data at multiple frequencies are useful for studying the spectral index:

$$\alpha = \frac{\log(S_{\nu_1}/S_{\nu_2})}{\log(\nu_1/\nu_2)}$$

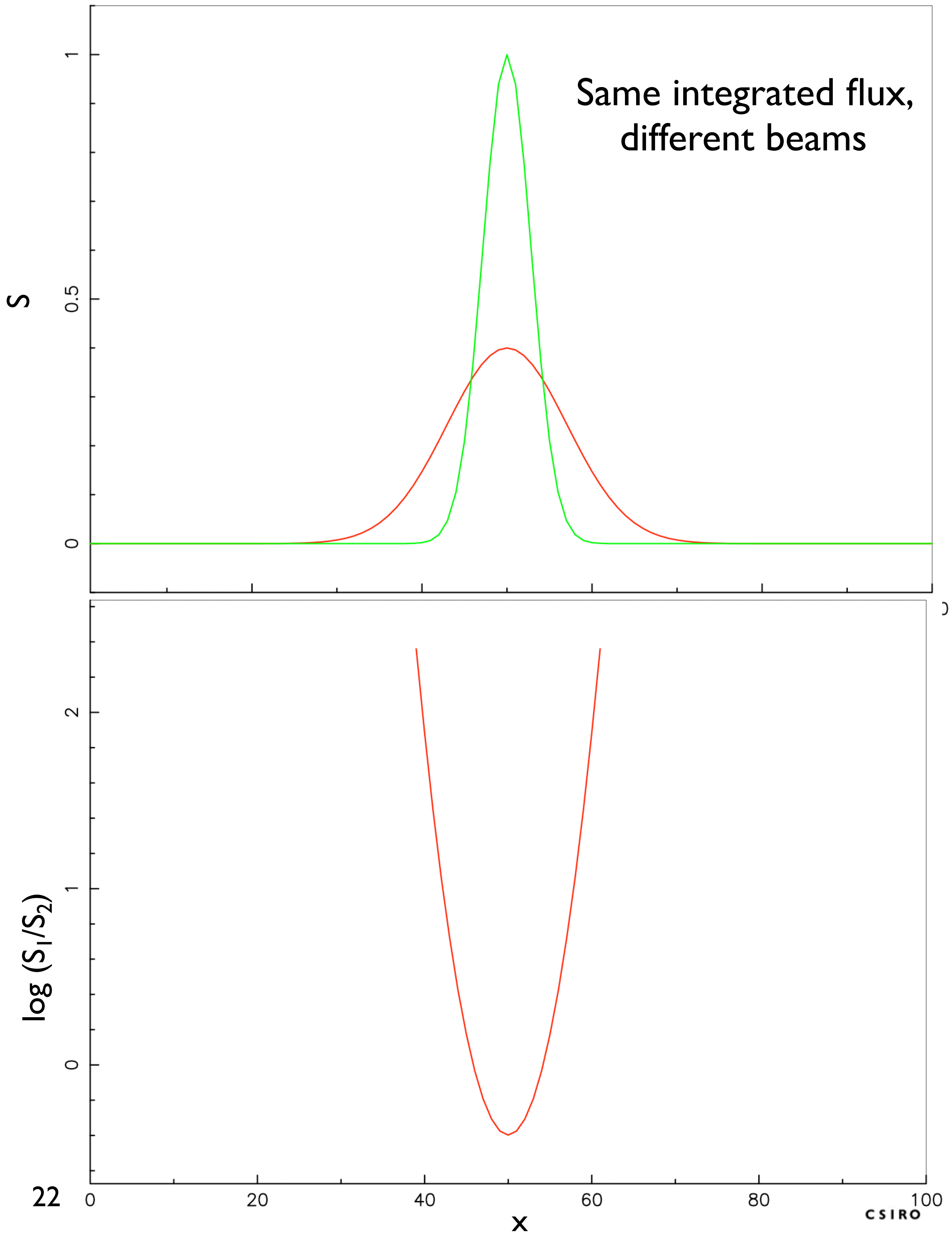
- Spectral index of emission tells about the emission mechanism, i.e. thermal, synchrotron
- Can tell us about the energy spectrum of the radiating particles
- Be warned: spectral index maps of radio emission are dependent on the baselines used



Tycho's SNR
DeLaney et al. (2002)

Spectral Index Pitfalls

- It is important to ensure that you have the same size restoring beam for both images



Spectral Index Pitfalls

- It is important to ensure that you have the same size restoring beam for both images
- Images must be aligned or you get a false gradient

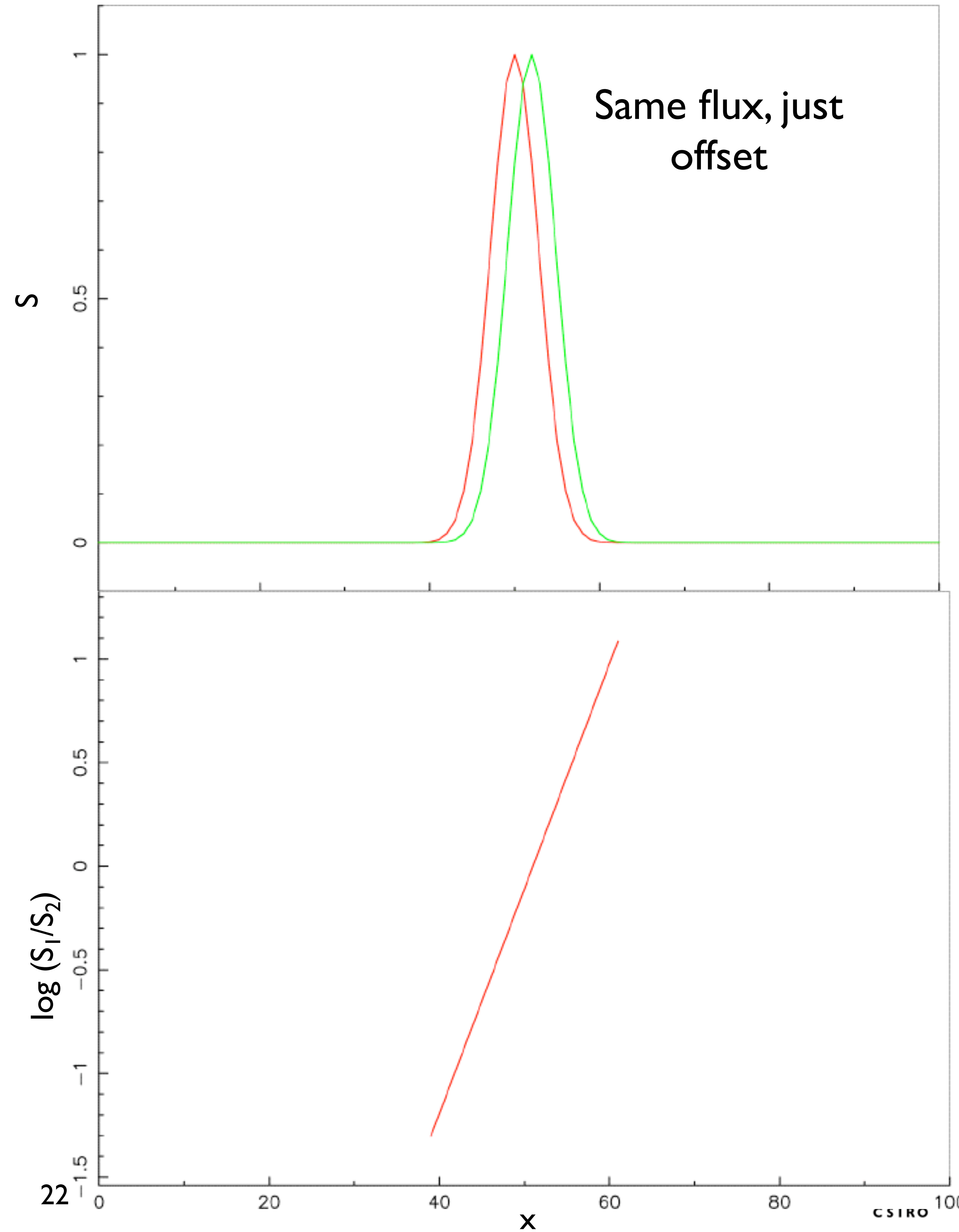


Image combination: Polarization

- Your telescope provides you with images of Stokes parameters I, Q, U,V
- You might want to know the polarized intensity:

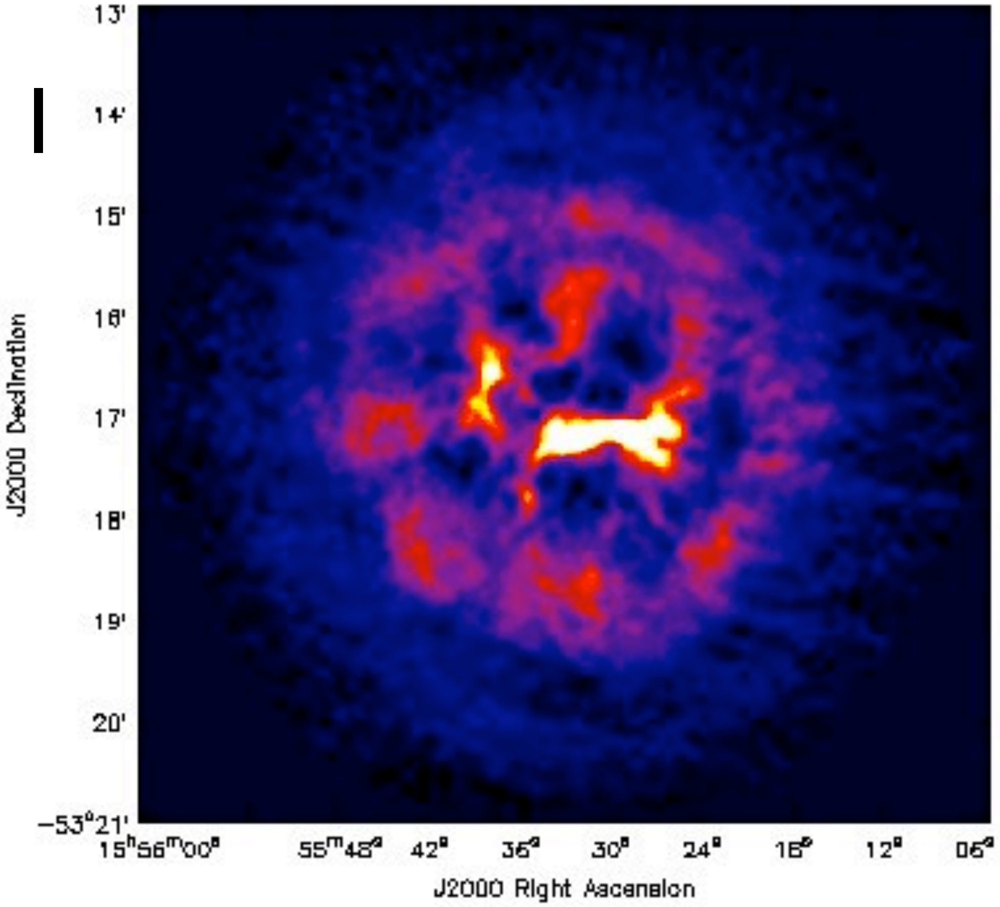
$$PI = \sqrt{Q^2 + U^2}$$

- Polarization angles of your image:

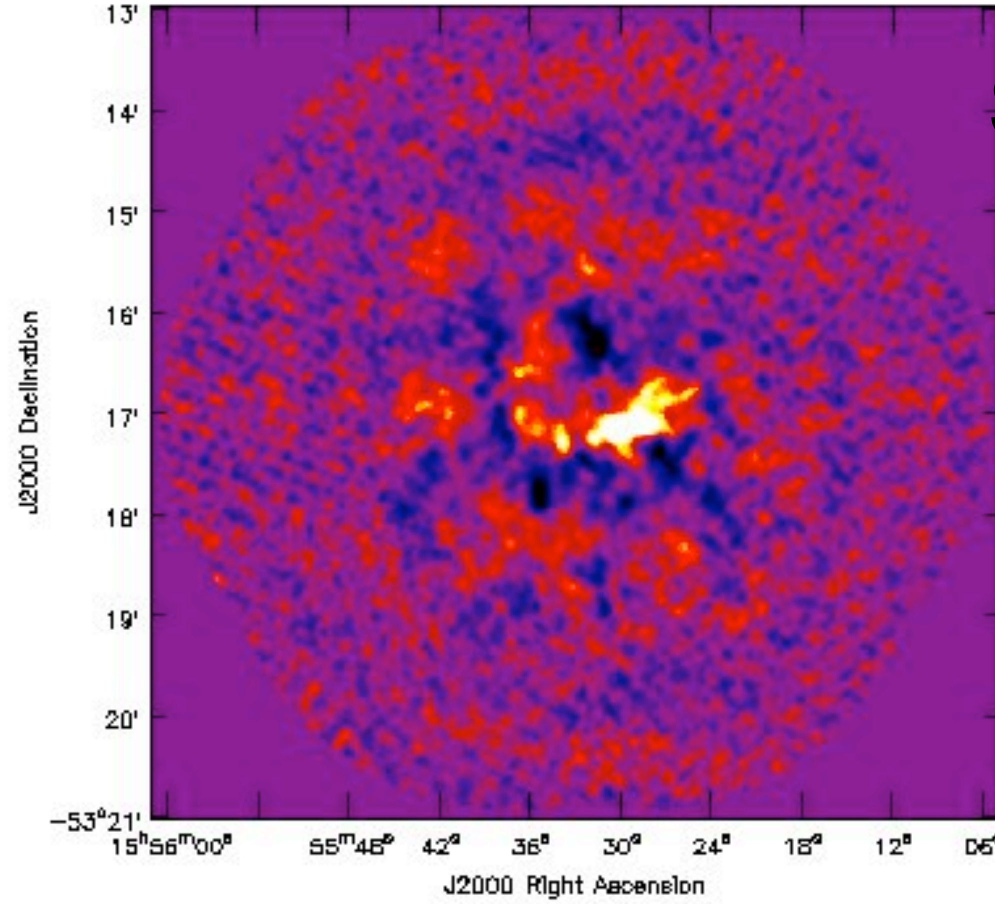
$$\chi = 1/2 \tan^{-1}(U/Q)$$

- Note that PI is a positive definite quantity so the noise is no longer Gaussian
- For low-signal to noise images you need to debias
- Various polarimetric tasks can do this combination for you:
  - e.g. Miriad's impol

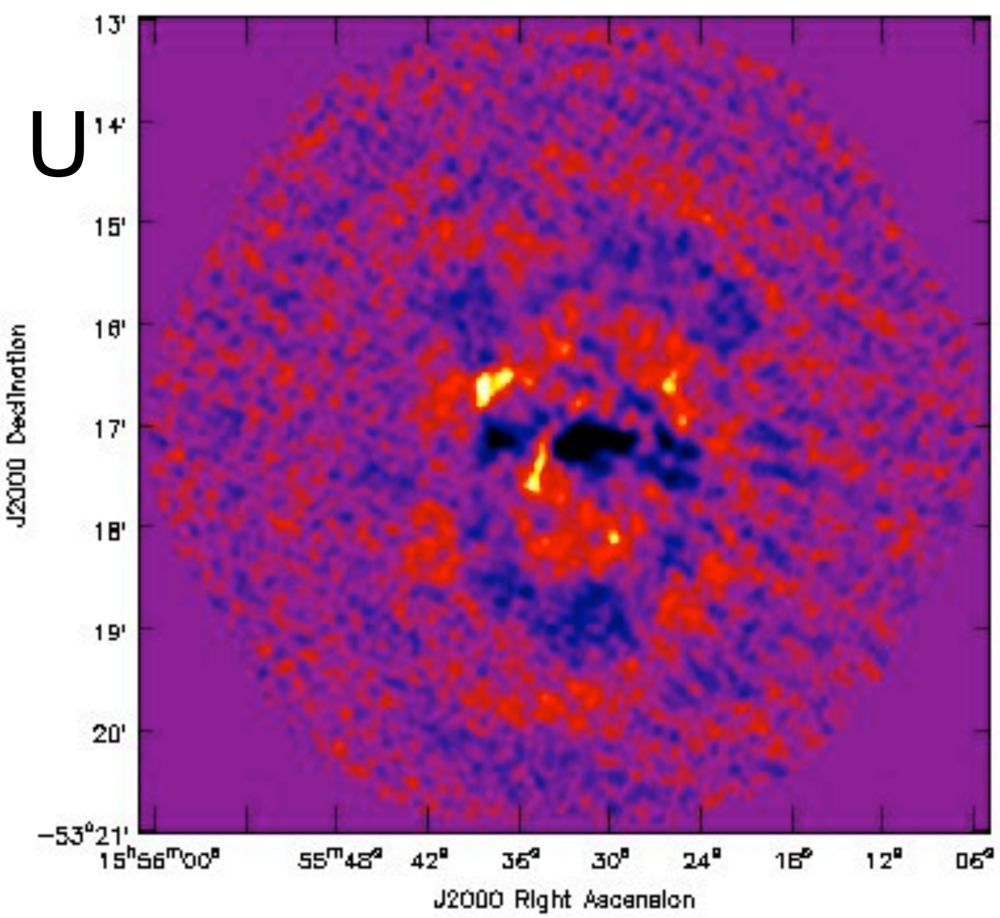
Stokes I



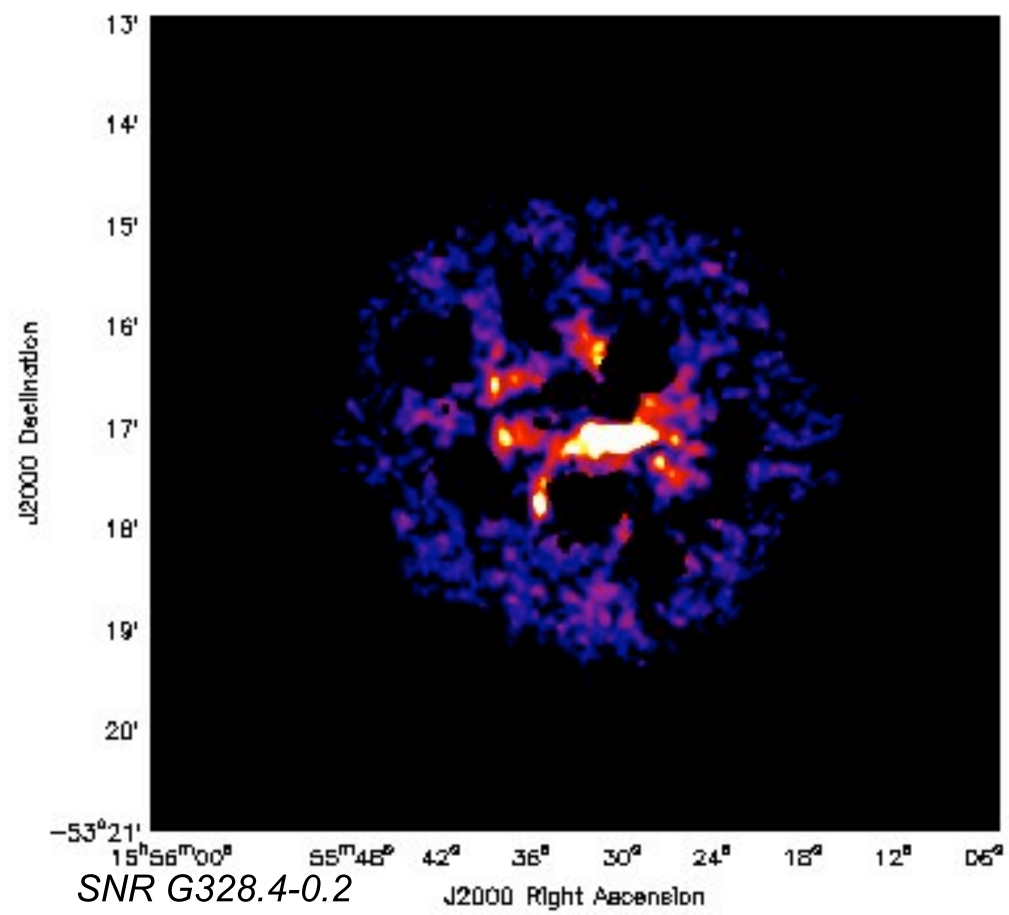
Stokes Q



Stokes U



PI



SNR G328.4-0.2
Johnston, McClure-Griffiths & Koribalski (2002)

CSIRO

Polarization Analysis II

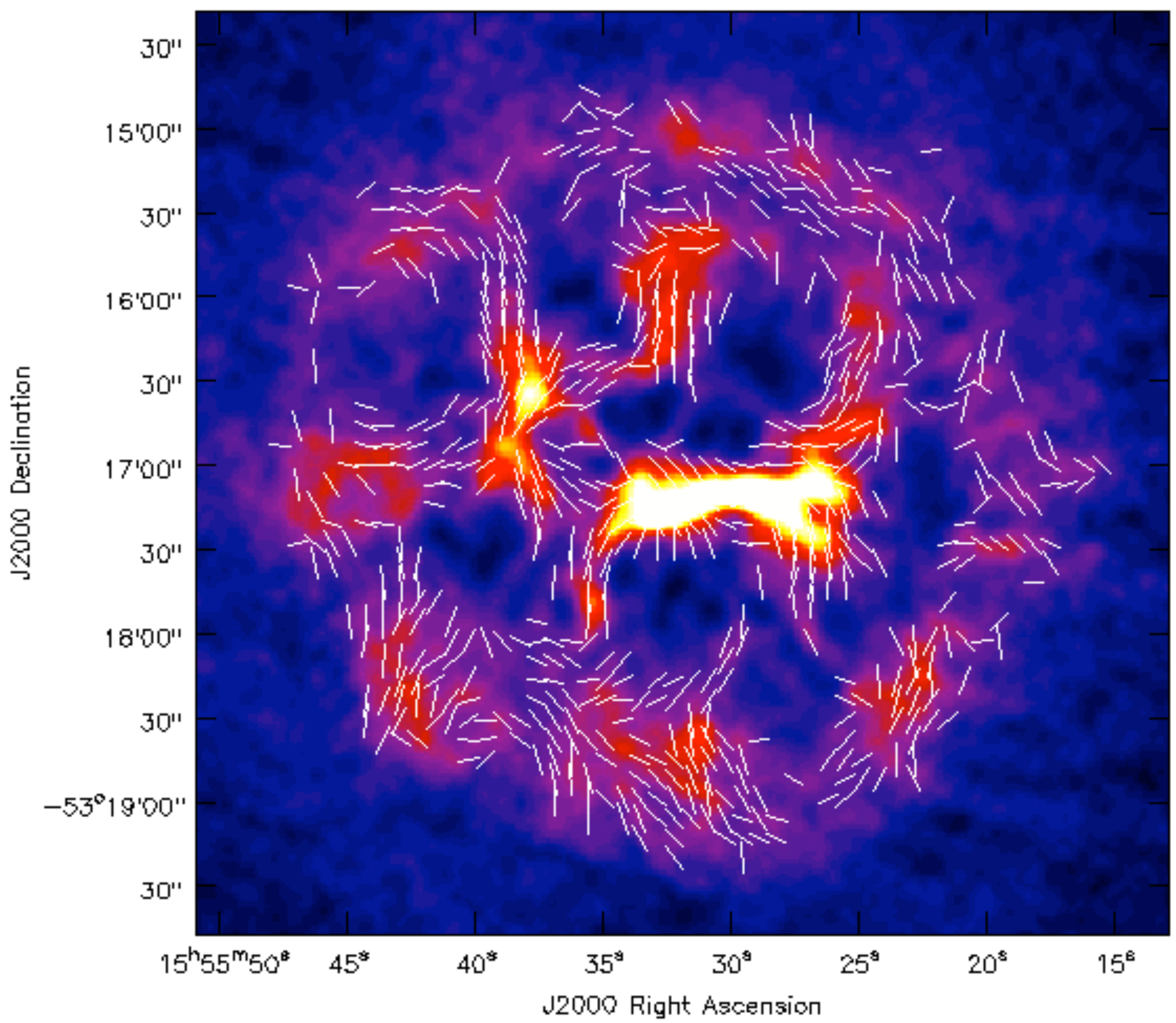
- From the PA at different frequencies we can determine the rotation measure:

$$RM = \frac{\chi(\lambda_1) - \chi(\lambda_2)}{\lambda_1^2 - \lambda_2^2}$$

- And the intrinsic position angle:

$$\chi = \chi_0 + RM\lambda^2$$

- The intrinsic position angle can be displayed as vectors
  - Note the B-field direction is rotated 90 deg from the PA

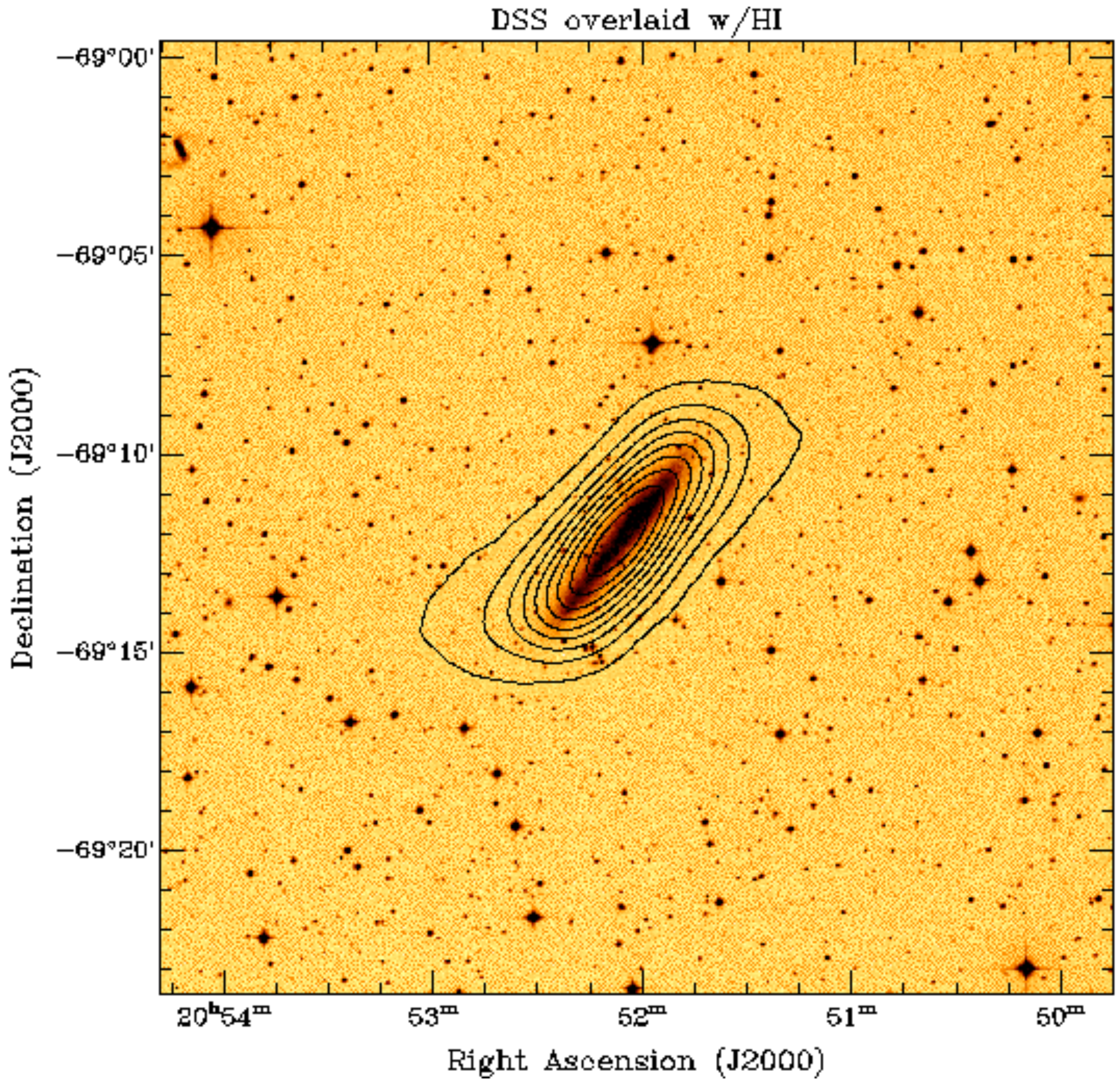


B-field vectors

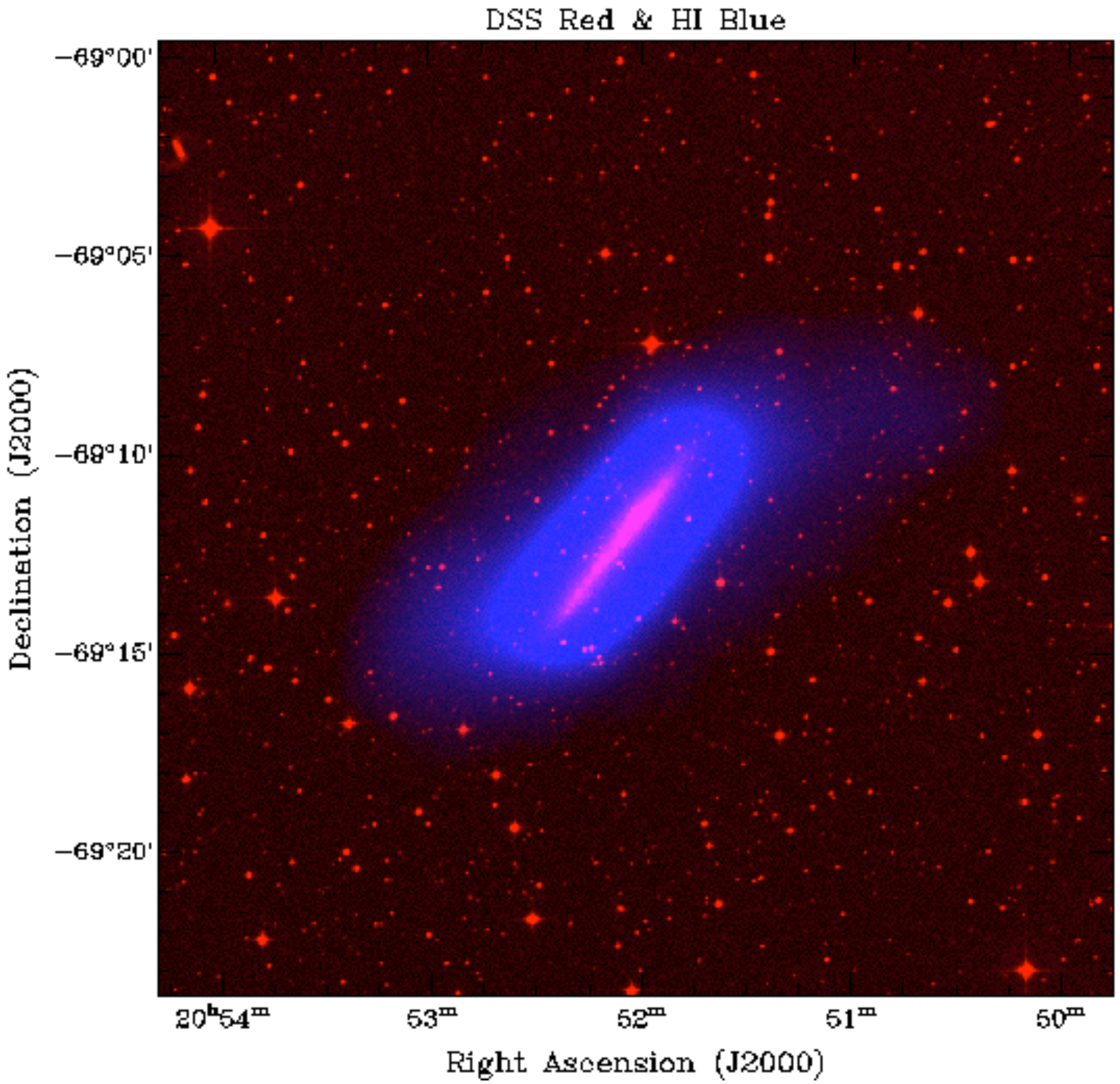
Data Visualisation

- The tools available are extensive:
  - Miriad cgdisp: produces good publication quality plots and is scriptable but it's a bit clunky
  - Kvis: excellent for quick looks, easy to overlay images and get ballpark estimates of noise, flux, etc. Not as good for publication quality plots and can't handle large areas of sky
  - casapy viewer: very good for coordinate systems and some nice analysis tools, not very intuitive
  - DS9: aimed mainly at optical/xray images, but it's quite flexible and handles coordinates well

Data Visualisation



Overlays: DSS Red plate w/ HI total intensity contours



RGB: DSS Red plate with HI total intensity

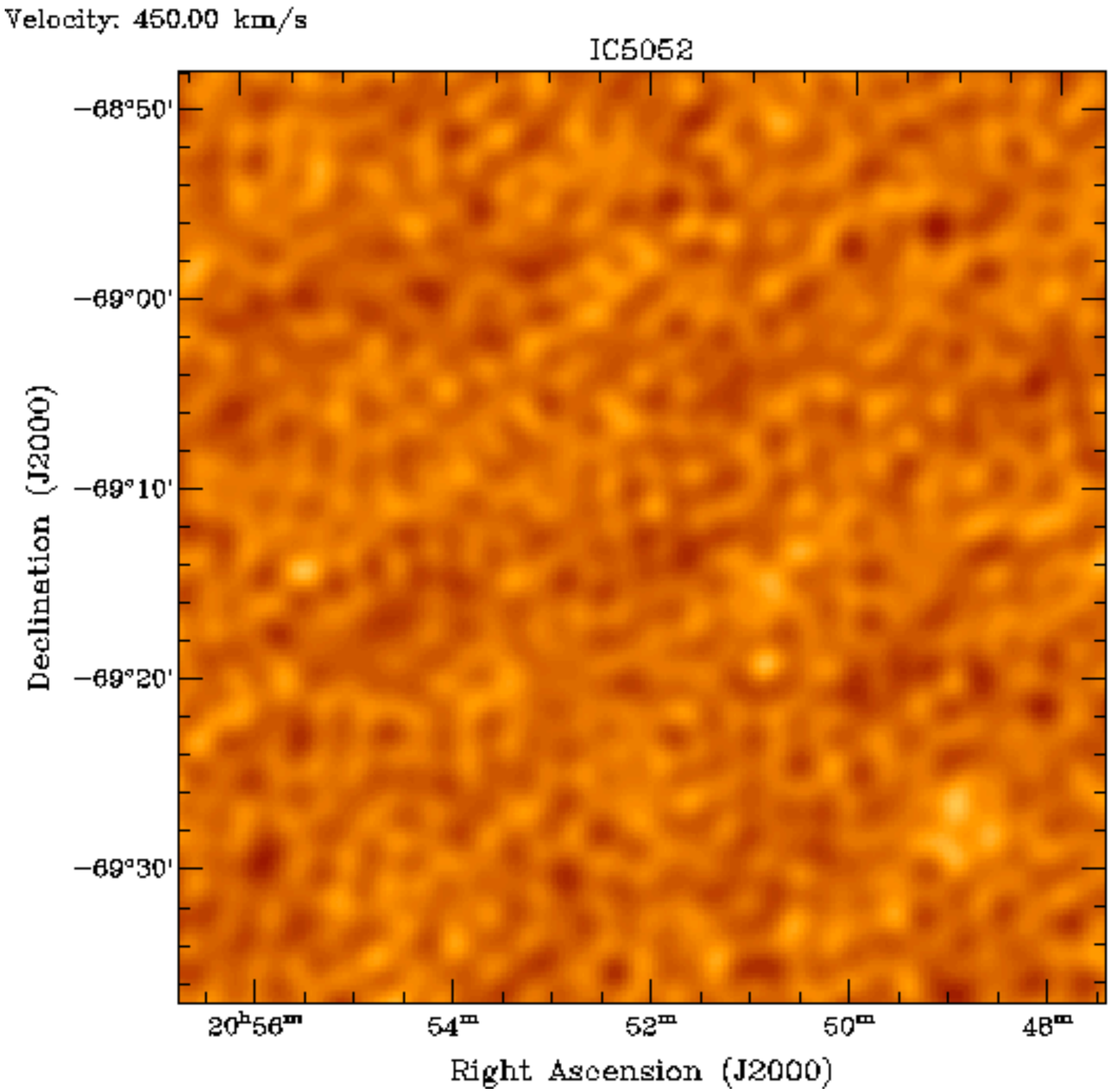
IC 5052 from LVHIS (Koribalski et al. in prep.)

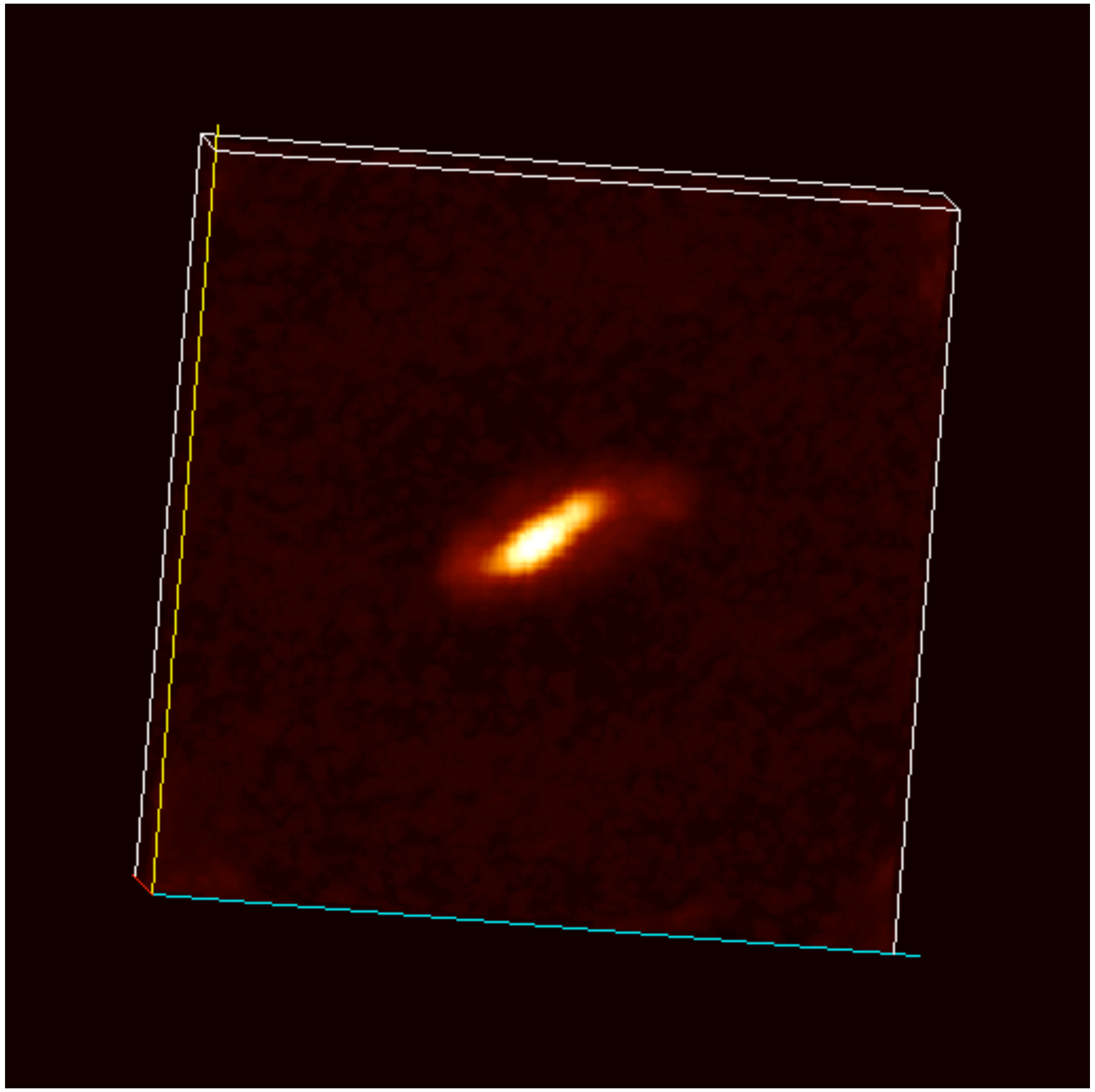
3-d data visualisation

IC 5052 from LVHIS (Koribalski et al. in prep.)

3-d visualisation is tricky, usually use time to probe the 3rd axis - for quantitative analysis need to reduce the dimensionality

3-d data visualisation

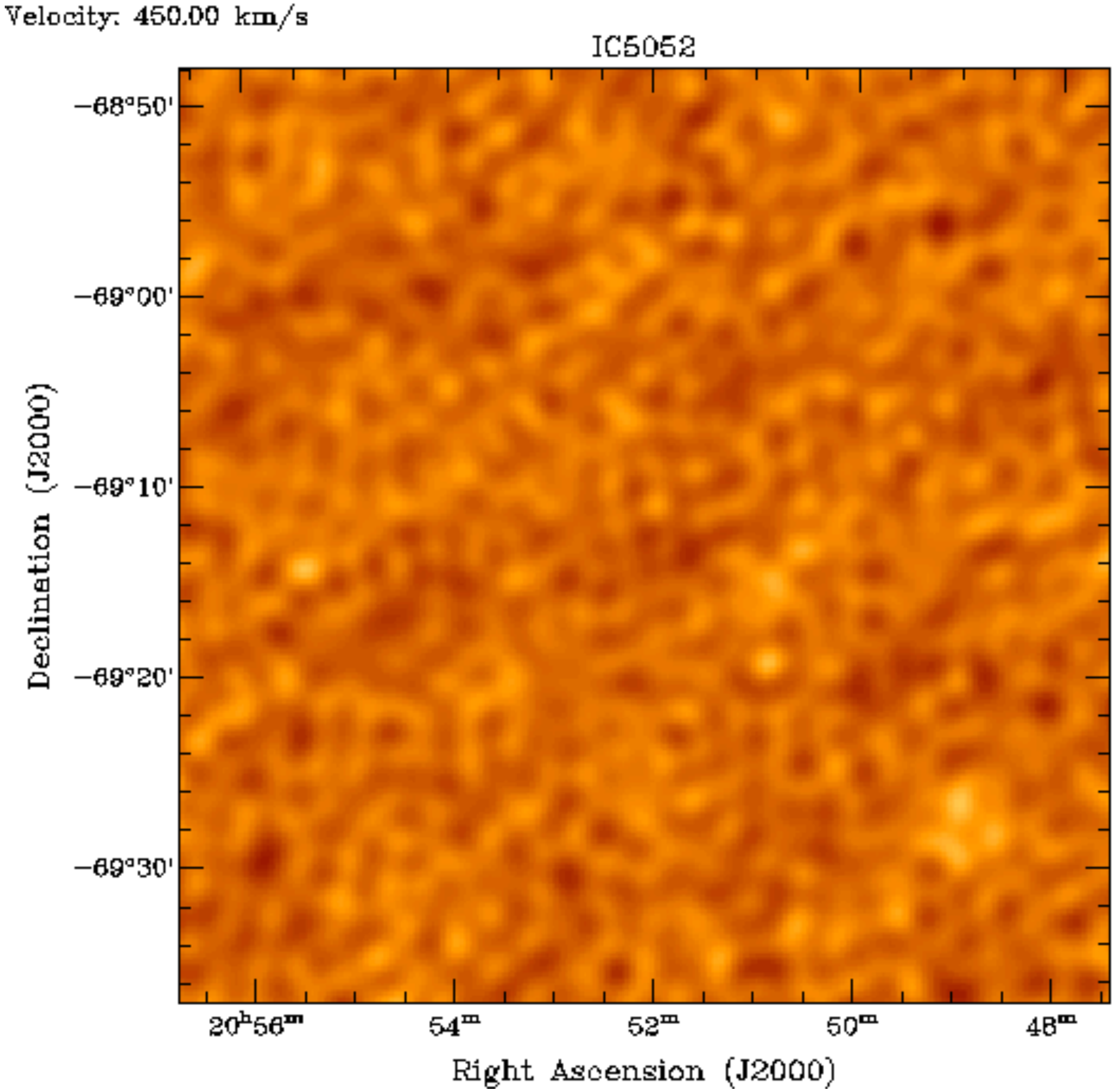


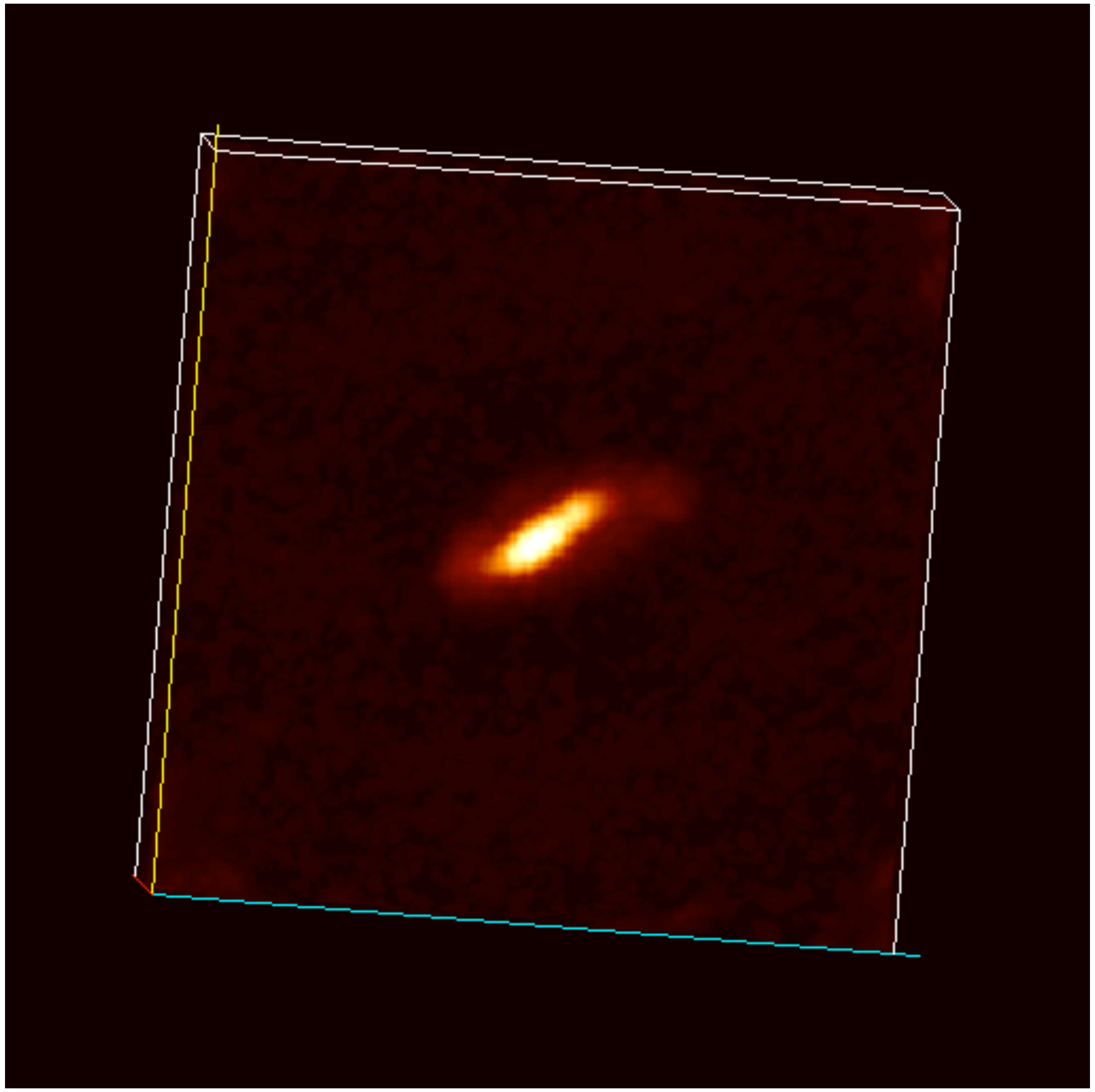


IC 5052 from LVHIS (Koribalski et al. in prep.)

3-d visualisation is tricky, usually use time to probe the 3rd axis - for quantitative analysis need to reduce the dimensionality

3-d data visualisation





IC 5052 from LVHIS (Koribalski et al. in prep.)

3-d visualisation is tricky, usually use time to probe the 3rd axis - for quantitative analysis need to reduce the dimensionality

Reducing the dimensionality: moments

- Zeroth moment - integrated intensity:

$$I_{tot} = \Delta v \sum_{i=1}^{n} I(\alpha, \delta, v_i)$$

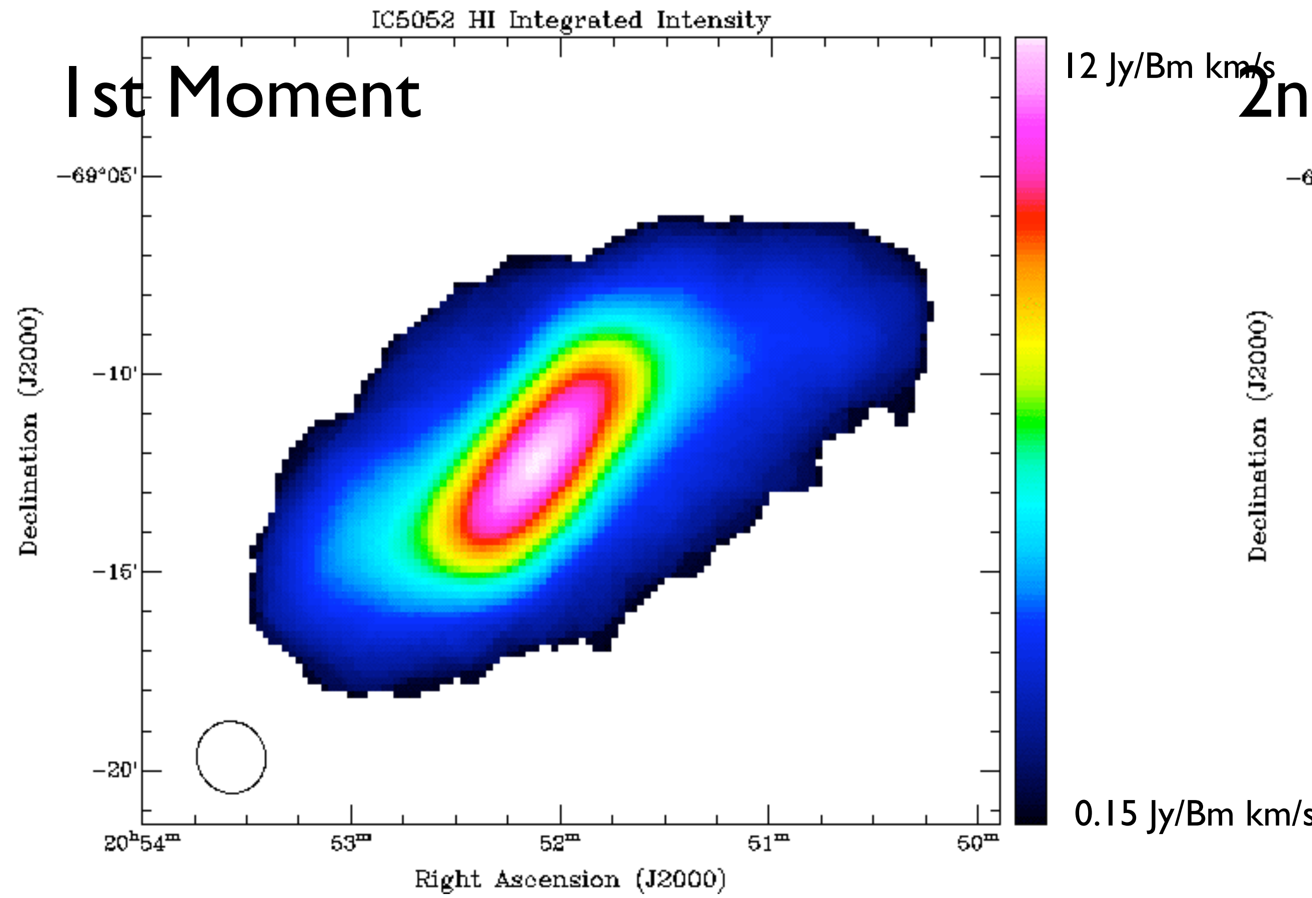
- First moment - intensity weighted coordinate:

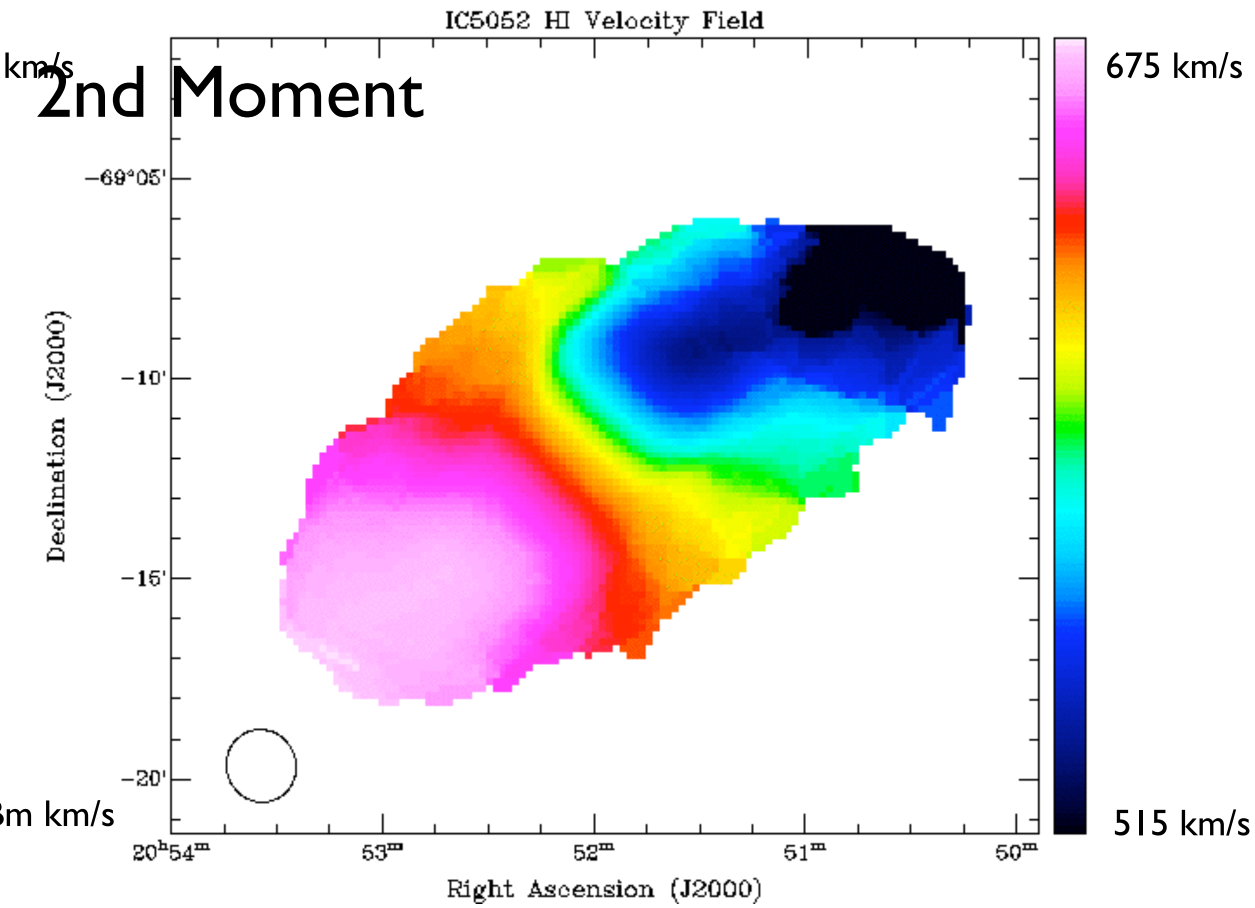
$$\bar{v} = \frac{\sum_{i=1}^{n} v_i I(\alpha, \delta, v_i)}{\sum_{i=1}^{n} I(\alpha, \delta, v_i)}$$

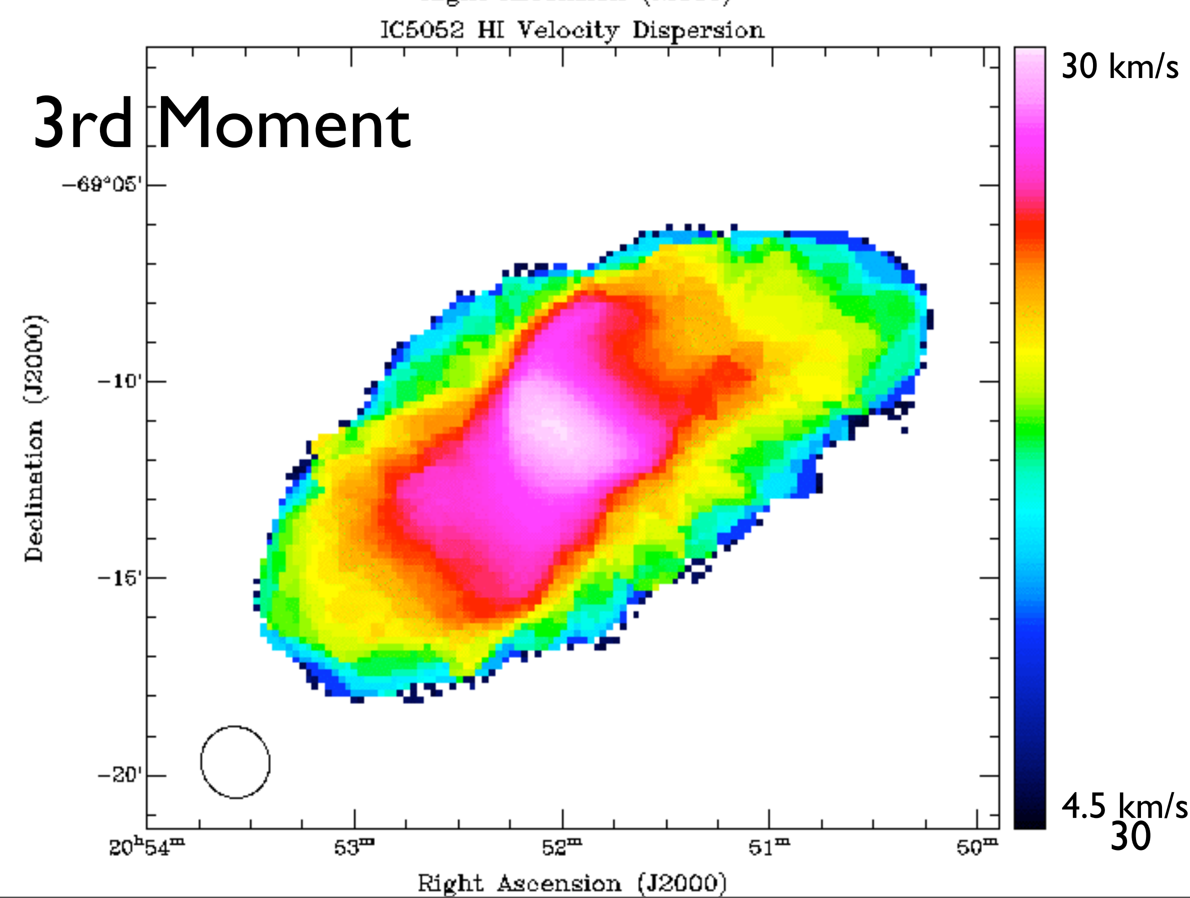
- Second moment - intensity weighted dispersion of coordinate:

$$\sigma_v(\alpha, \delta) = \sqrt{\frac{\sum_{i=1}^{n} I(\alpha, \delta, v_i)(v_i - \bar{v}(\alpha, \delta))^2}{\sum_{i=1}^{n} I(\alpha, \delta, v_i)}}$$

Moments







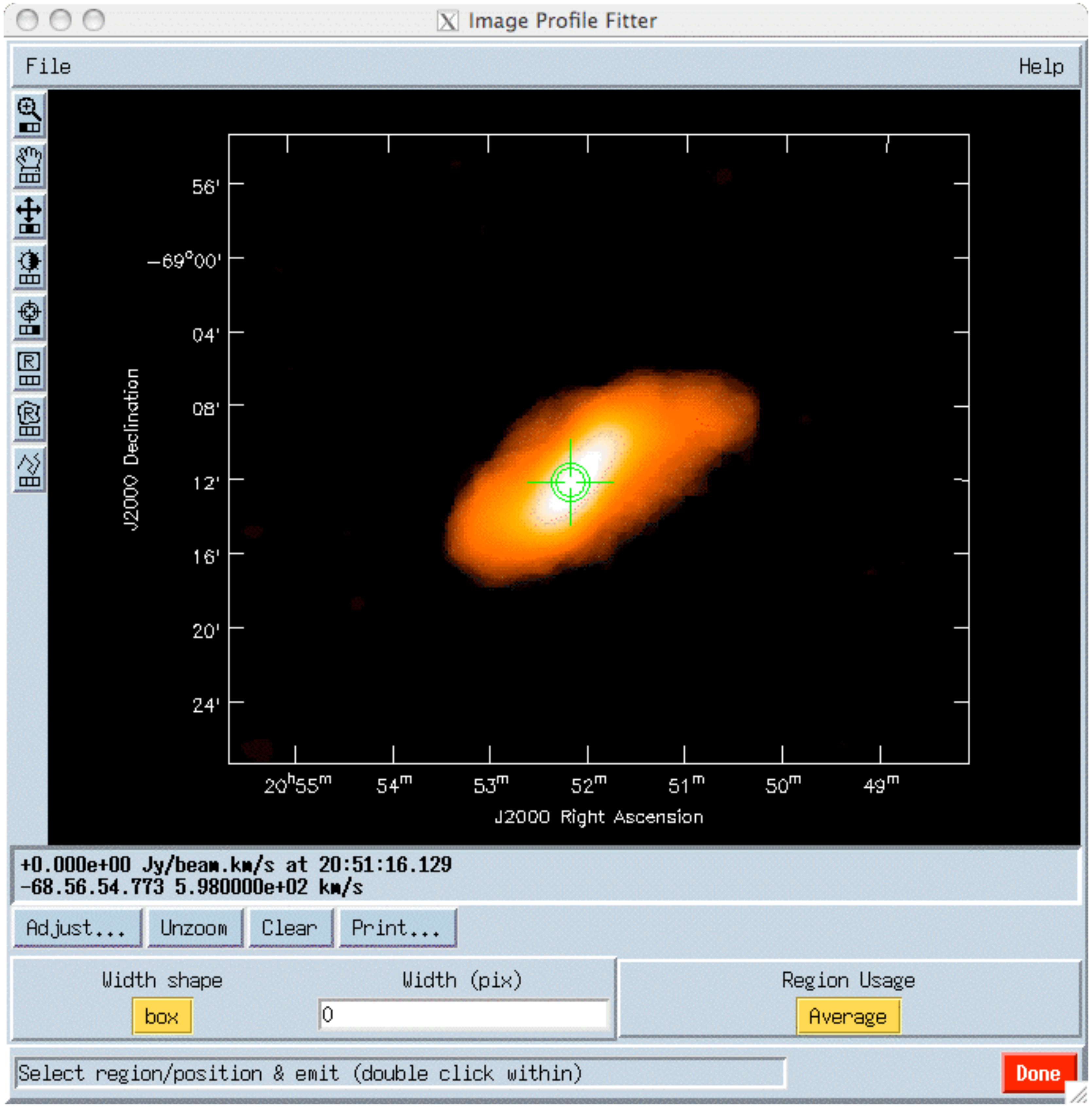
Moments are used to reduce the dimensionality for quantitative analysis

IC 5052 from LVHIS (Koribalski et al. in prep.)

Profile Fitting:

For spectral line analysis it is useful to quantify the strength of a line, its width and its position by fitting a model

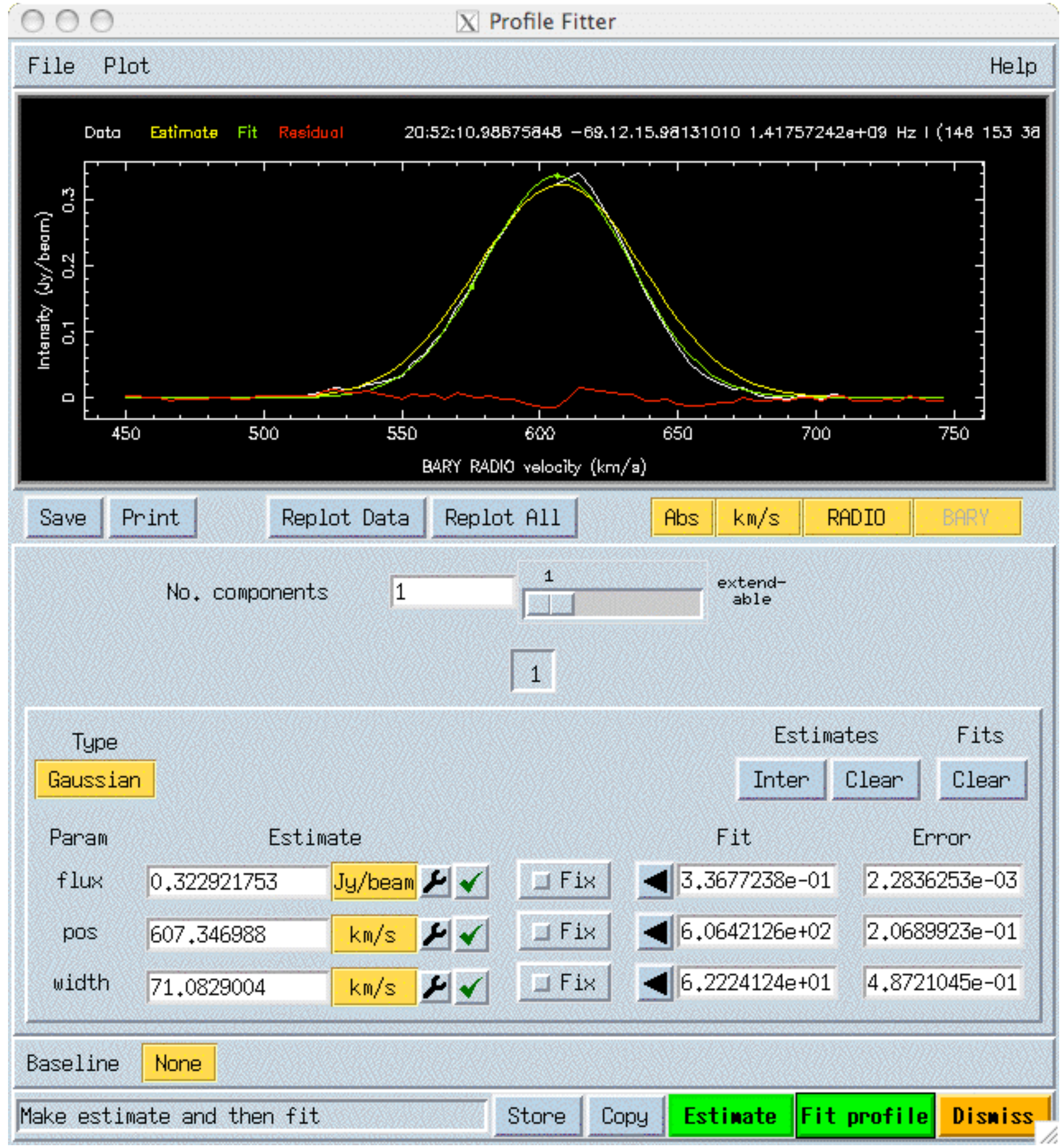
miriad: *gaufit*
casapy: *fitprofiles*



Profile Fitting:

For spectral line analysis it is useful to quantify the strength of a line, its width and its position by fitting a model

miriad: *gaufit*
casapy: *fitprofiles*



Summary

- Decide what you want to know before you do anything
  - The question you ask can determine:
    - Whether you image your data
    - How you restore your data
    - How you manipulate your data
- Look at your u-v data
  - Errors and limitations are sometimes easier to spot in the u-v domain
- Use all of the visualisation tools available to you, look at 3-D data from different perspectives
- For data analysis try to reduce the dimensionality, i.e. take slices, use moments

Summary II

- Make certain you understand your errors:
  - Estimate the rms in your image using kvis or miriad's *imstat*
  - Remember that errors returned by fitting processes are usually formal fitting errors, not the actual error
  - Keep in mind that the signal-to-noise of your image effects the accuracy with which you can determine positions, source sizes
  - Polarization images may need to be de-biased
- There are many techniques so don't be afraid to try a technique from a different field, e.g., use moments in image plane to find source sizes and centroids

Recommended Reading

- Fomalont, E. 1999 in "Synthesis Imaging in Radio Astronomy II, Eds. G. B. Taylor, C. Carilli, R. A. Perley, chapter 14
- Peason, T. J. 1999 in "Synthesis Imaging in Radio Astronomy II, Eds. G. B. Taylor, C. Carilli, R. A. Perley, Chapter 16
- MIRIAD manual!