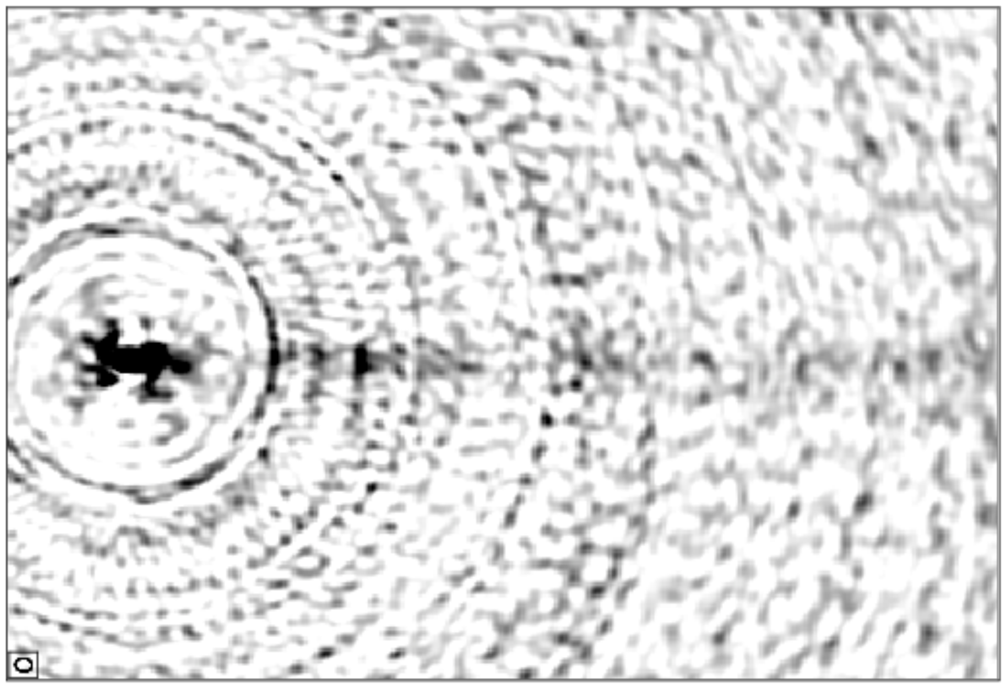
# Continuum error recognition and image analysis



acknowledgement: Anita Richards (JBCA), Robert Laing (ESO)

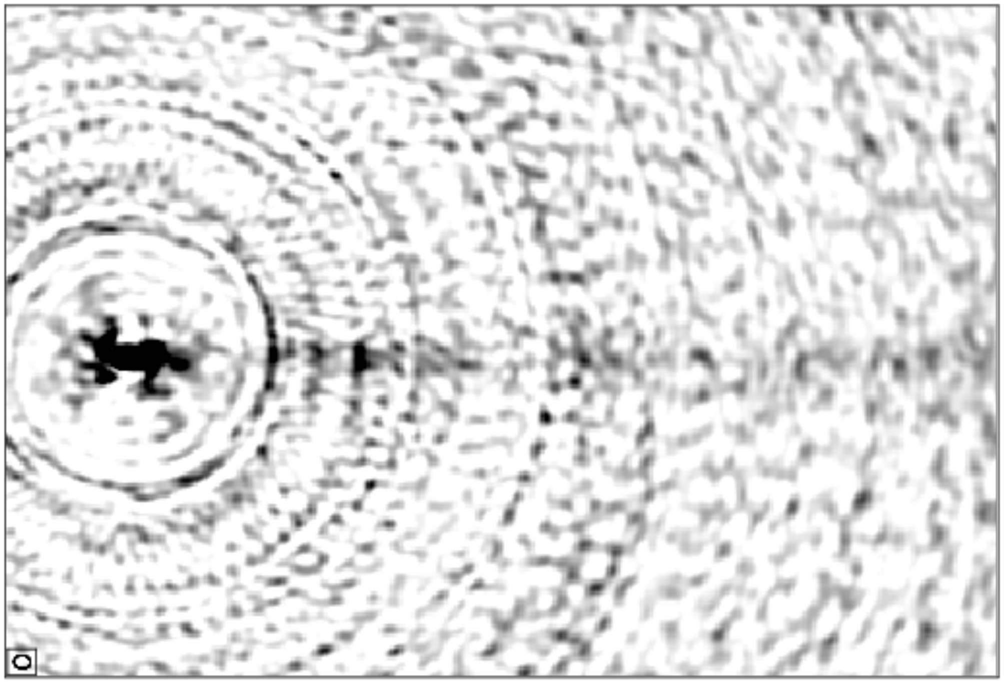
# Outline

- Error recognition: how do you recognise and diagnose residual errors by looking at images

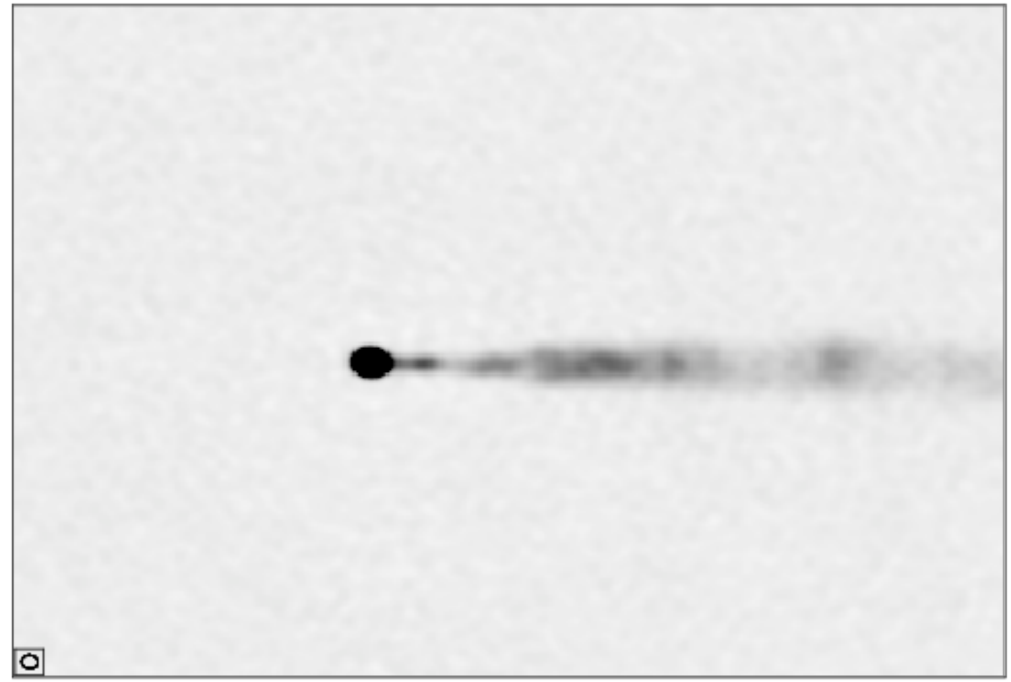
- Image analysis: how to extract scientifically useful numbers from images

- Unless otherwise specified, this talk is about continuum imaging in full polarisation .... but many ideas also apply to spectral-line work.

# Is my image good enough?



No



Yes

# How can I tell?

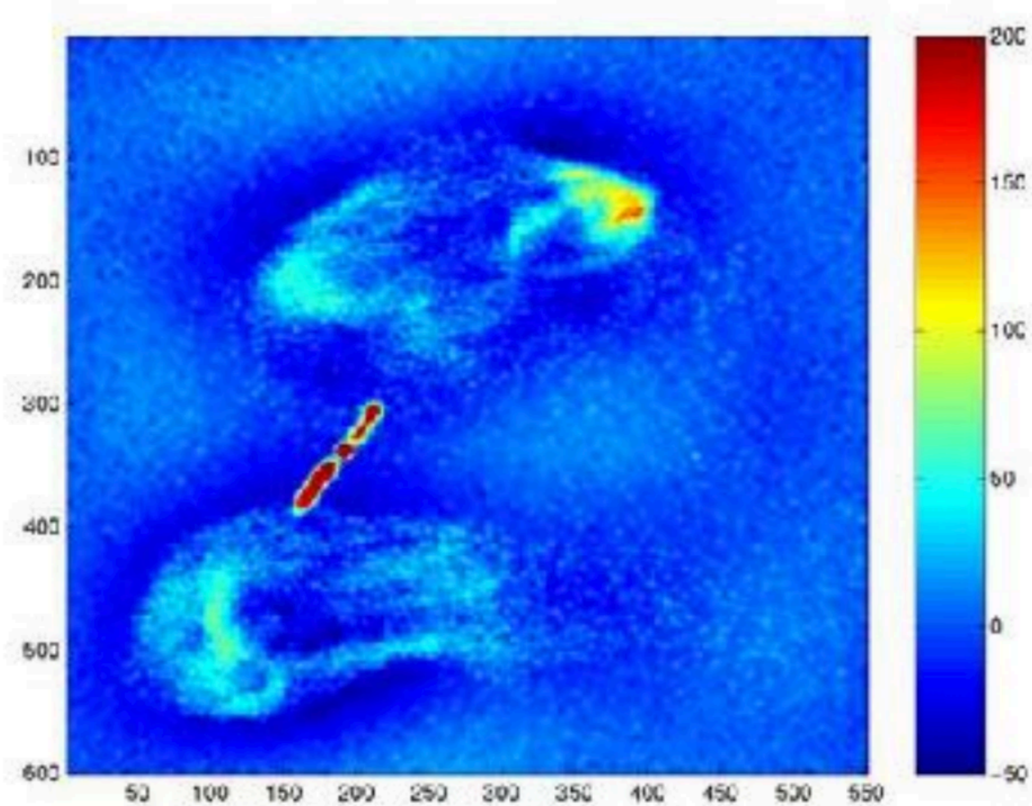
- **Look at the off-source rms noise**
  - Use on-line calculators (e.g. JVLA, ALMA) or formulae

$$S_{\rm rms} = \frac{2kT_{\rm sys}}{A_{\rm eff}\sqrt{N_A(N_A-1)t_{\rm int}\Delta\nu}}$$

  - Measure rms with (e.g.) casaviewer or imstat
  - Does the image rms increase near bright sources?
  - Is the noise random or are there ripples?
- **Are there obvious artefacts?**
  - Coherent I features < 4σ
  - Rings, streaks etc.
- **Properties of artefacts**
  - Additive (constant over the field) or multiplicative (scales with brightness)?
  - Symmetric or antisymmetric around bright sources?

# How can I tell?

- Large-scale negative structures
  - Negative “bowl' around source structure
  - Large-scale sinusoidal ripples
- **Missing short spacings**
- Need more data

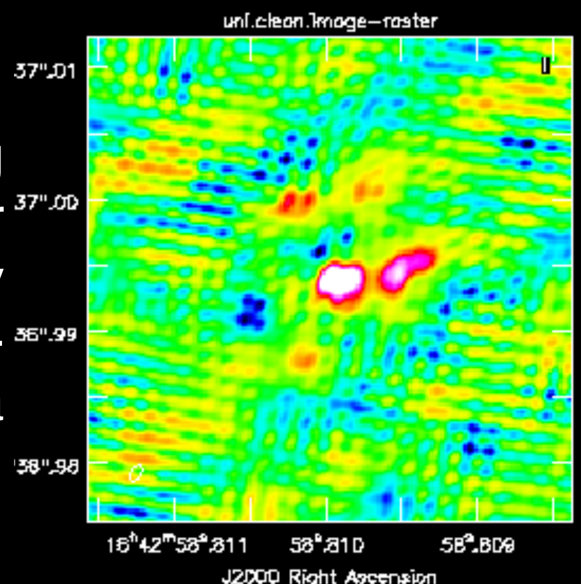



Paul Rayner 2001

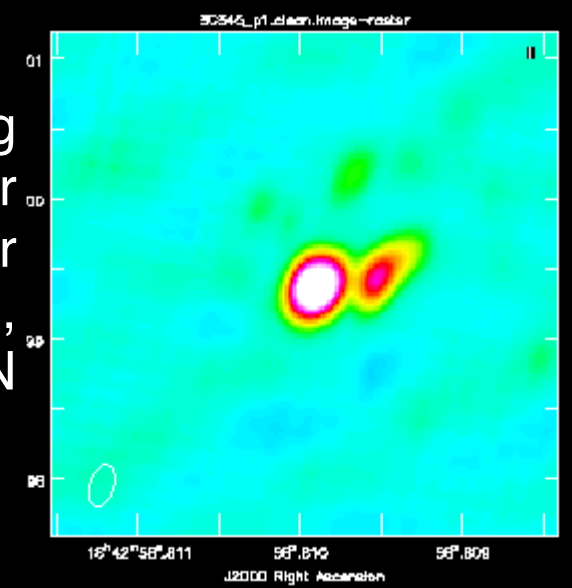
# How can I tell?

- Unnatural small-scale on-source structure
  - Diffuse structure looks spotty
  - Short-wavelength sinusoidal ripples
- Deconvolution errors
  - often associated with poor u,v coverage

Uniform weighting
Unsuitable for sparse uv sampling, partly-calibrated data




Natural weighting
Coarser resolution but far fewer artefacts, better S/N



# Possible causes: imaging problems

- Is the image big enough?
  - Confusing sources outside the image
    - Make a wider-field, tapered image and look
    - Look in standard catalogues (NVSS)
- Are the pixels small enough to sample the beam?
  - Are bright point sources accurately located on pixels?
- Wide-field issues (calculate expected effects)
  - Averaging time too long? (Azimuthal smearing $\propto$ radius)
  - Spectral channels too wide? (Radial smearing $\propto$ radius)
  - w-term (non-coplanar baselines)?
  - Ionosphere (single field > isoplanatic patch)?
  - Pointing/antenna position errors (see calibration talk)?

# Possible causes: imaging problems

- Missing short spacings
- Primary beam effects
- Deconvolution errors, especially with sparse u-v coverage
    - Resolution too high?
    - Poor choice of weighting?
    - Bad choice of CLEAN boxes (too small, too large, ...)
    - Insufficient CLEANing
    - single-scale CLEAN not good enough
- Source variability during the observations
    - Unexpectedly bad artefacts
        - Look at visibilities v. time ....

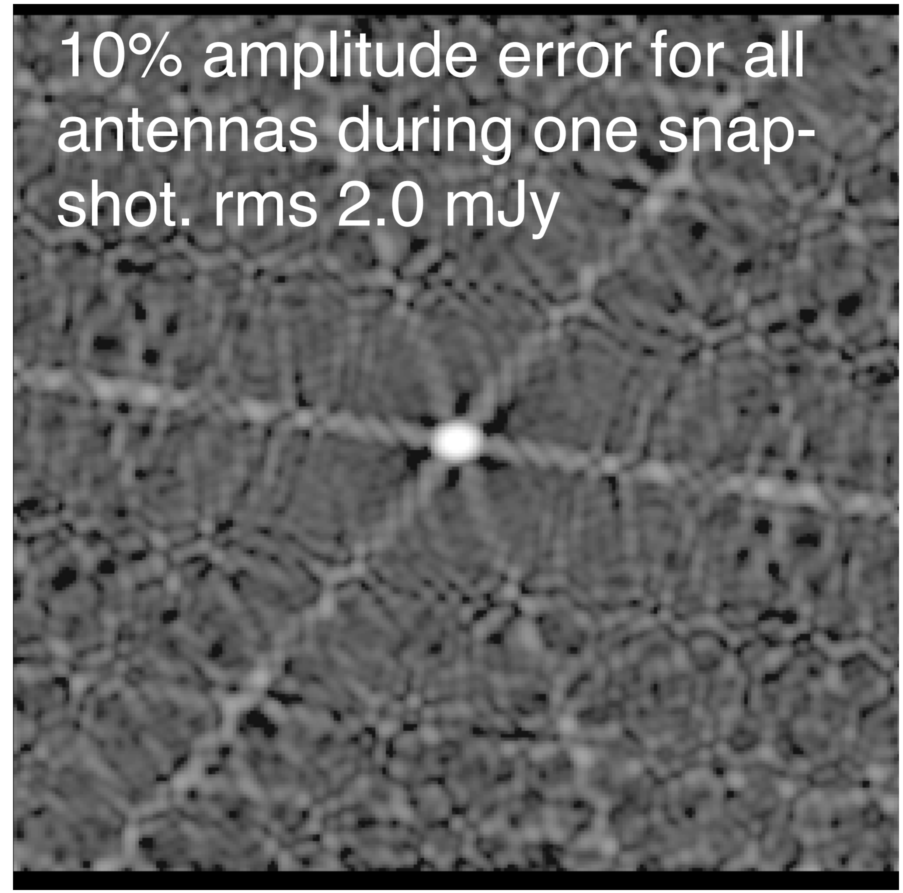
# Errors in the image and (u, v) planes

- Errors obey Fourier relations between (u, v) and image planes
  - Particularly helpful in recognising additive errors
  - e.g. single very high visibility: sinusoidal fringe
- u-v **amplitude** errors cause **symmetric** errors in the image plane
- u-v **phase** errors cause **antisymmetric** errors in the image plane

# (u, v) or image plane?

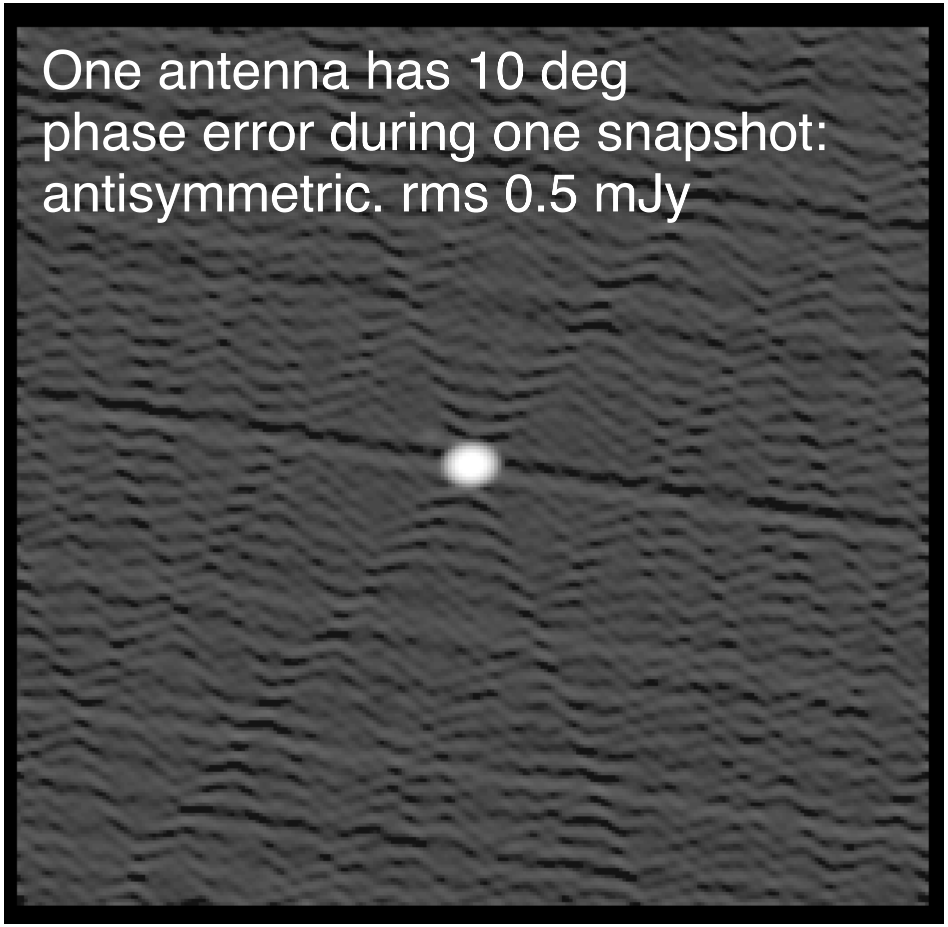
- Find the outliers in the u-v plane first
  - Big (MJy) points have big effects on the image (these should have been flagged, but mistakes happen)
  - A fraction f of bad data points with reasonable amplitudes give fractional error ~f in the image
- Low-level, but persistent errors are often easier to see in the image plane
- Rule of thumb: 10 deg phase error ≡ 20% amplitude error

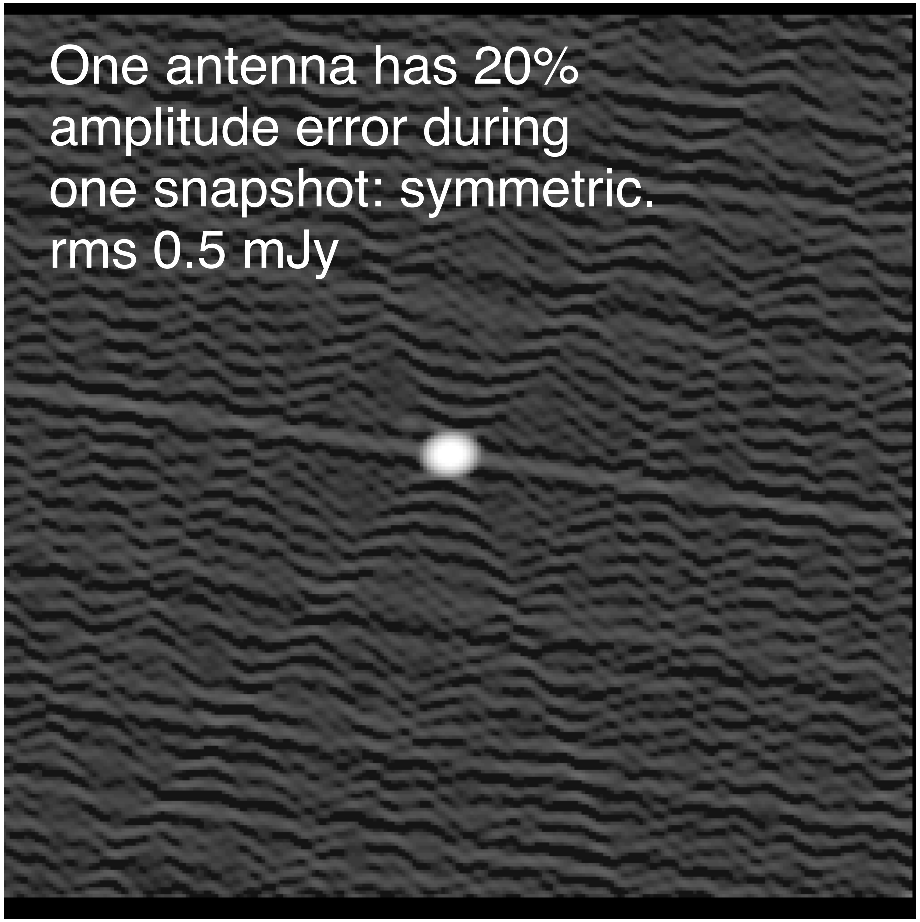
# Amplitude errors: all antennas

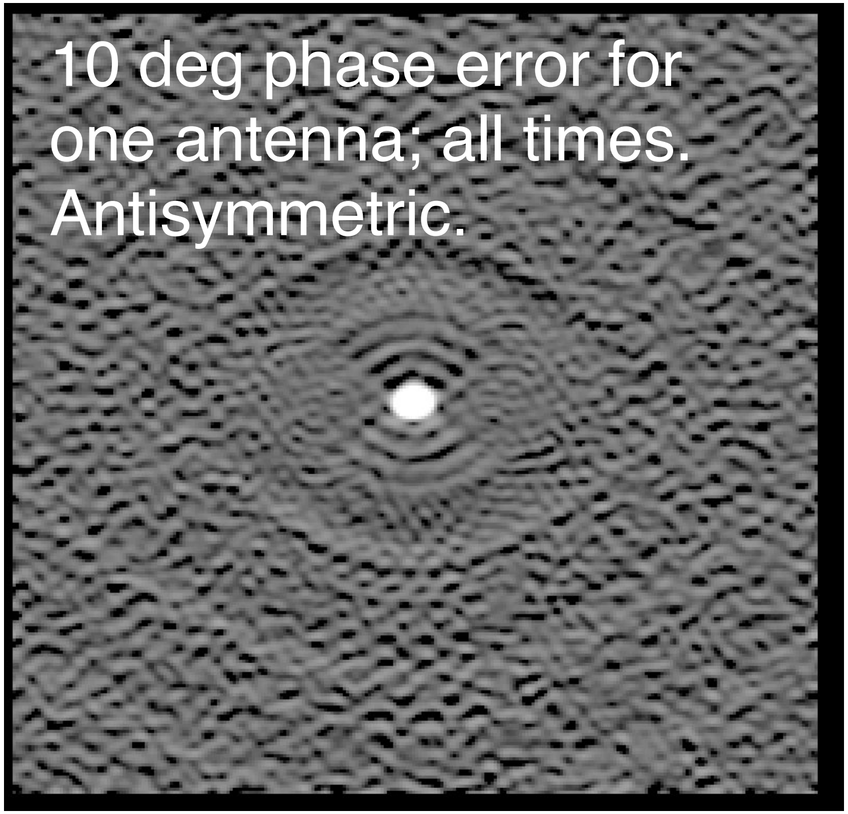

Error pattern looks like the dirty beam for a single VLA snapshot

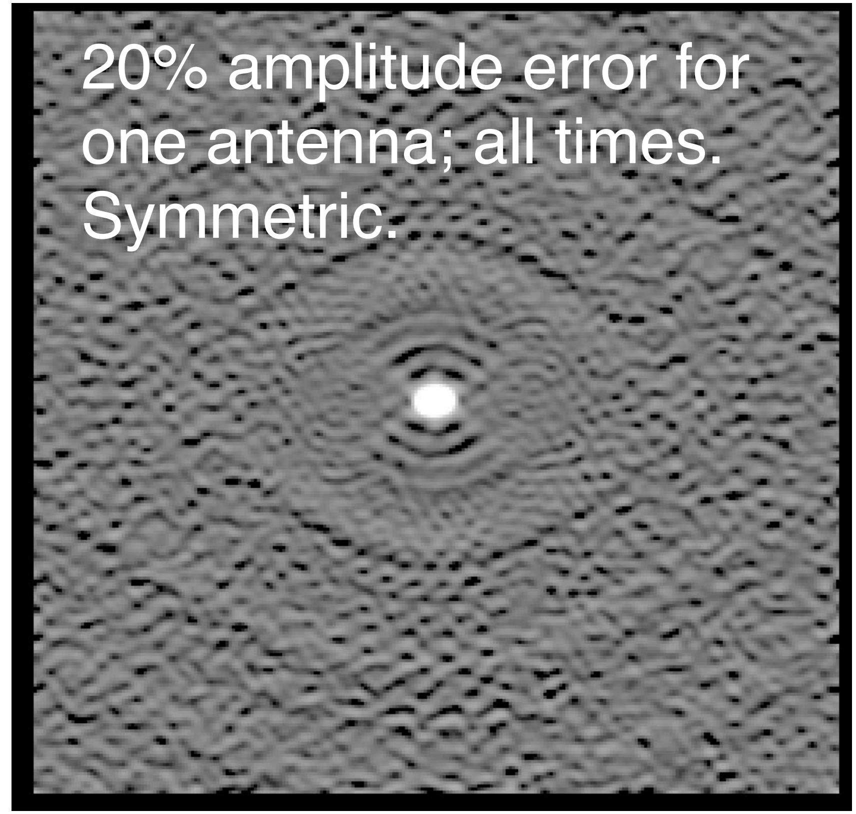
# One antenna in error at one time



# One antenna in error: all times

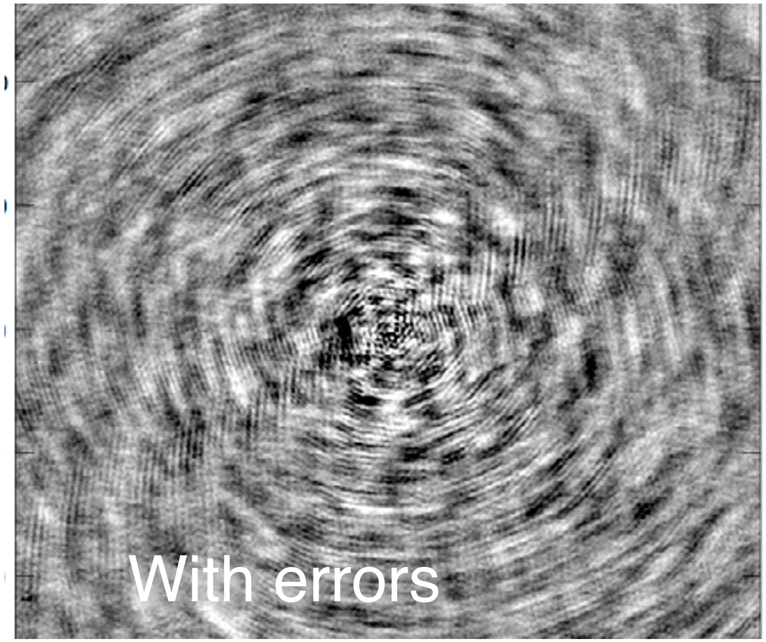


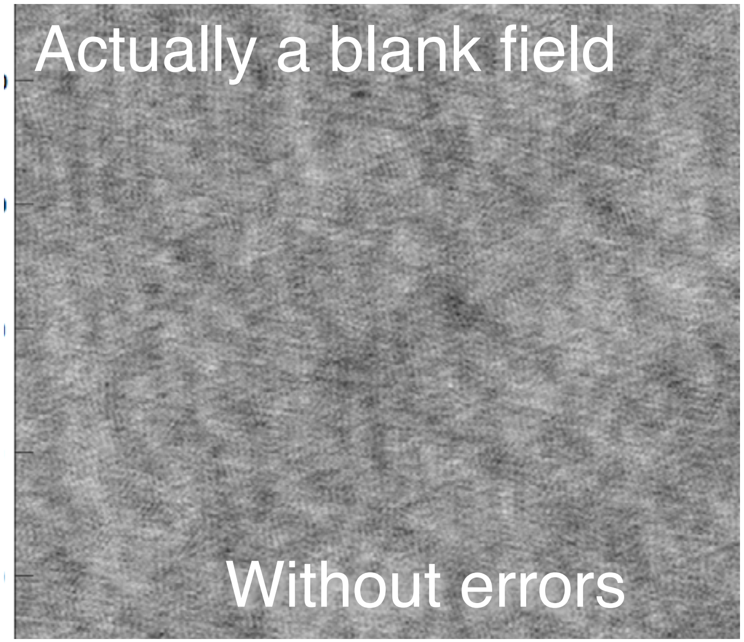




Multiplicative

Can diagnose by excluding one antenna in turn and re-imaging if hard to see in uv data

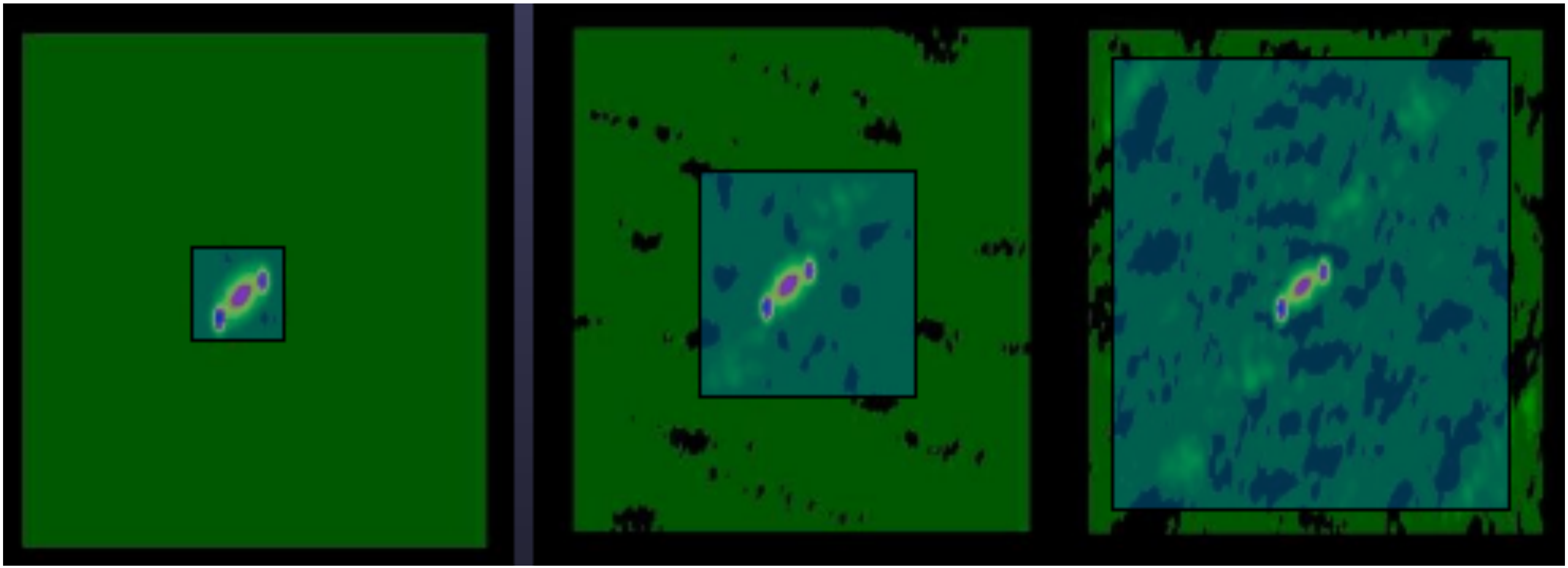
# Baseline-dependent errors




Most errors e.g. noisy receiver, bad weather, affect all the data for one (or more) antennas, i.e. all baselines to that antenna. Occasionally, bad baselines are caused by correlator errors, and for high dynamic range, by pol. leakage or bandpass responses. Errors are additive and per-baseline solutions need high S/N

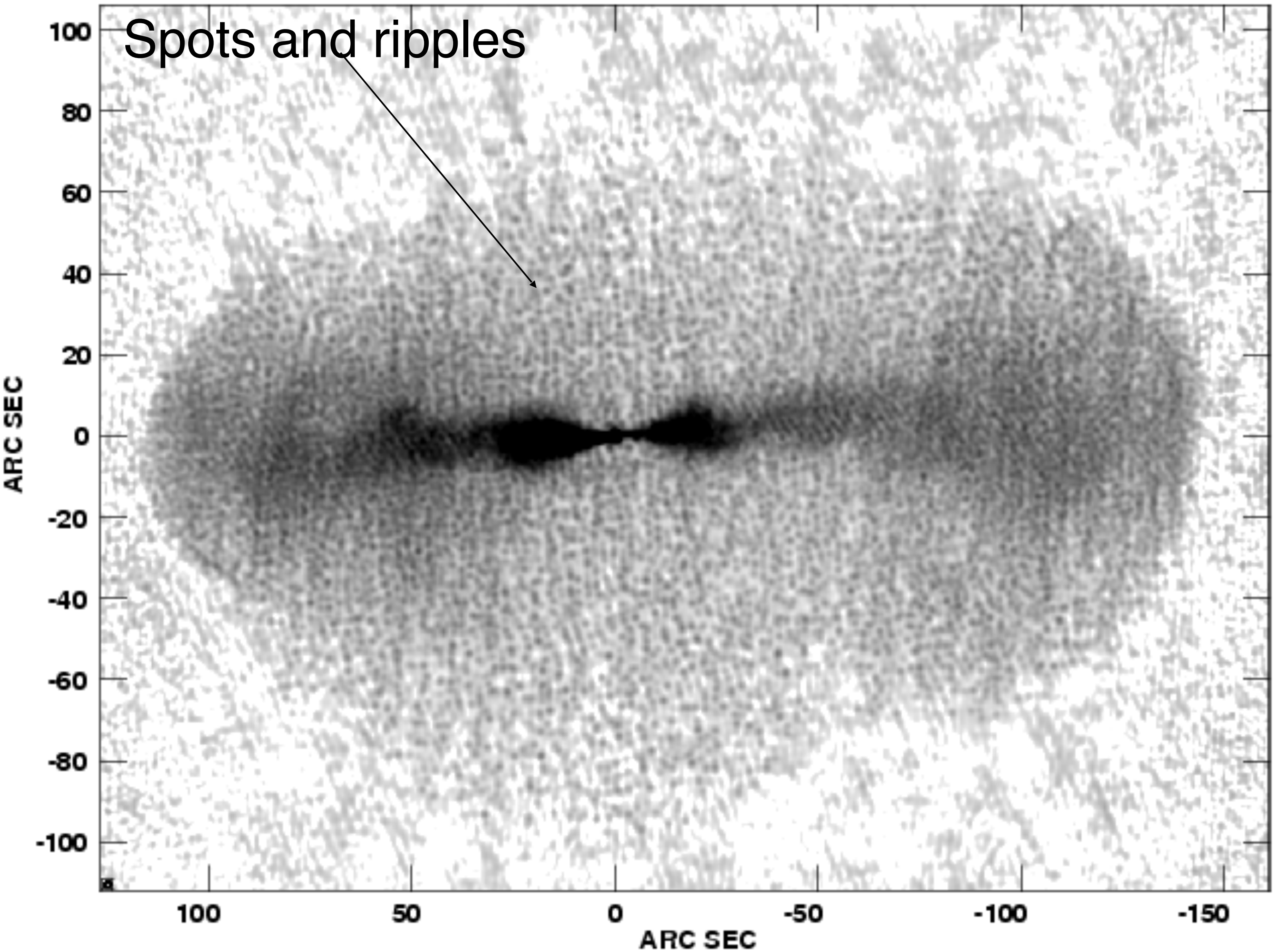
# CLEAN boxes too big!



Correct | Too big | Far too big

CLEAN functions best when the area in which it finds components is restricted. If unsure, do a small number of iterations per cycle and increase mask size if needed

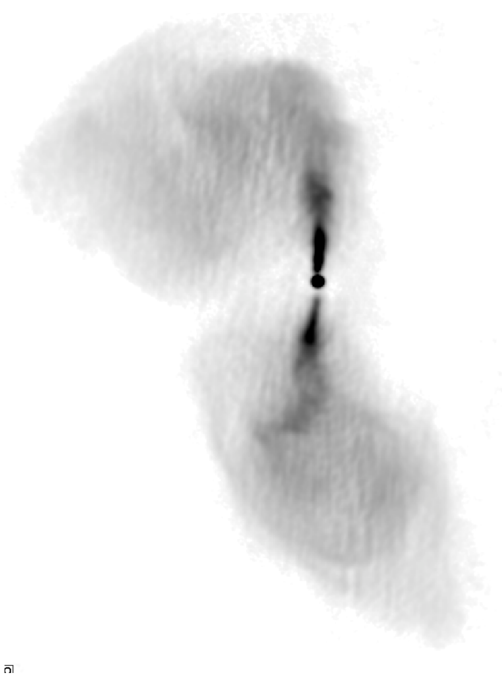
# Deconvolution Errors

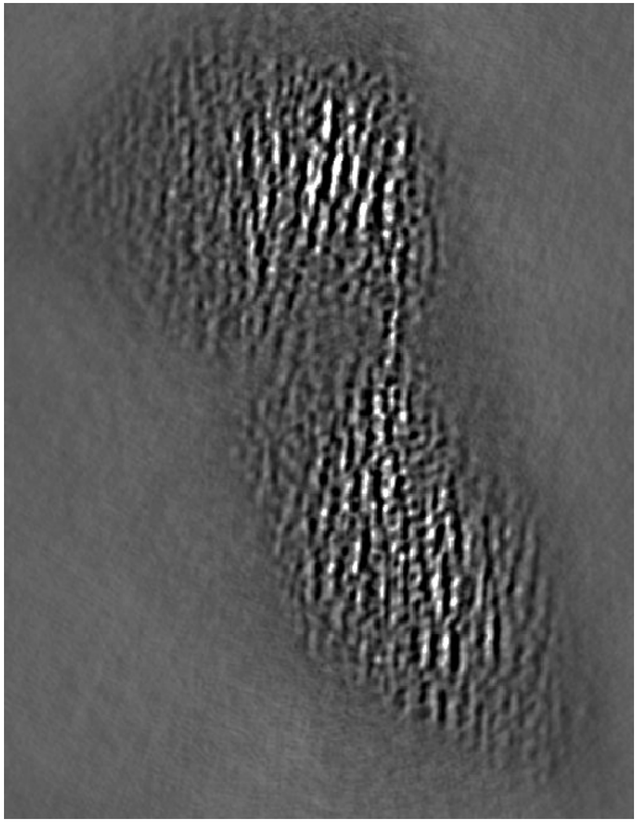



VLA A+B+C configurations. Short spacings OK, but with poor A-configuration coverage

Conventional CLEAN fails: try multi-resolution CLEAN or MEM or reduce the resolution

# Multi-scale CLEAN helps

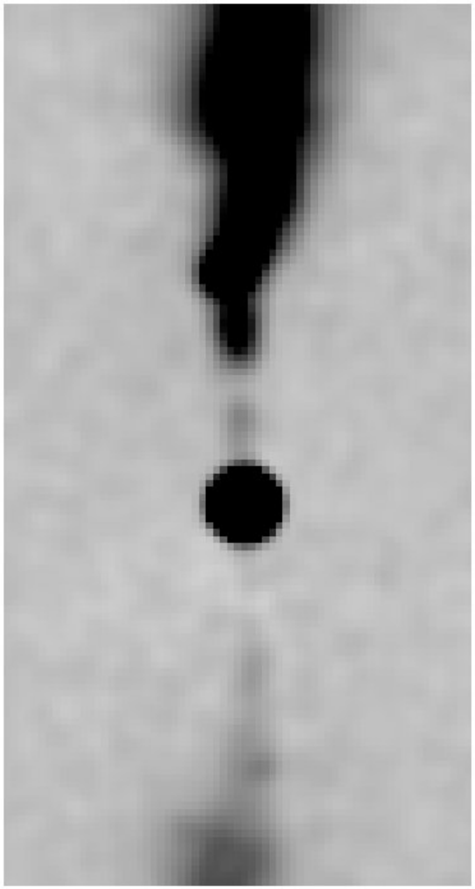




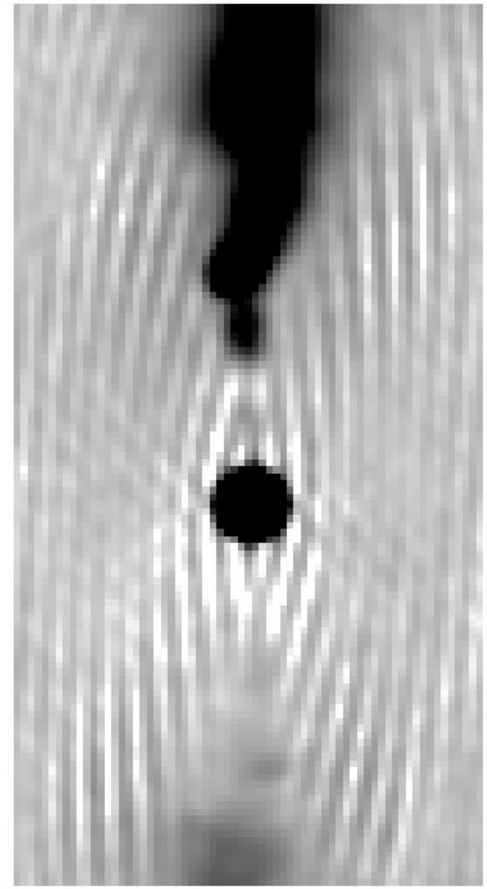
1-scale 3-scale 1-scale – 3-scale

Multi-scale CLEAN has removed a high-frequency ripple

# Point source not on a pixel

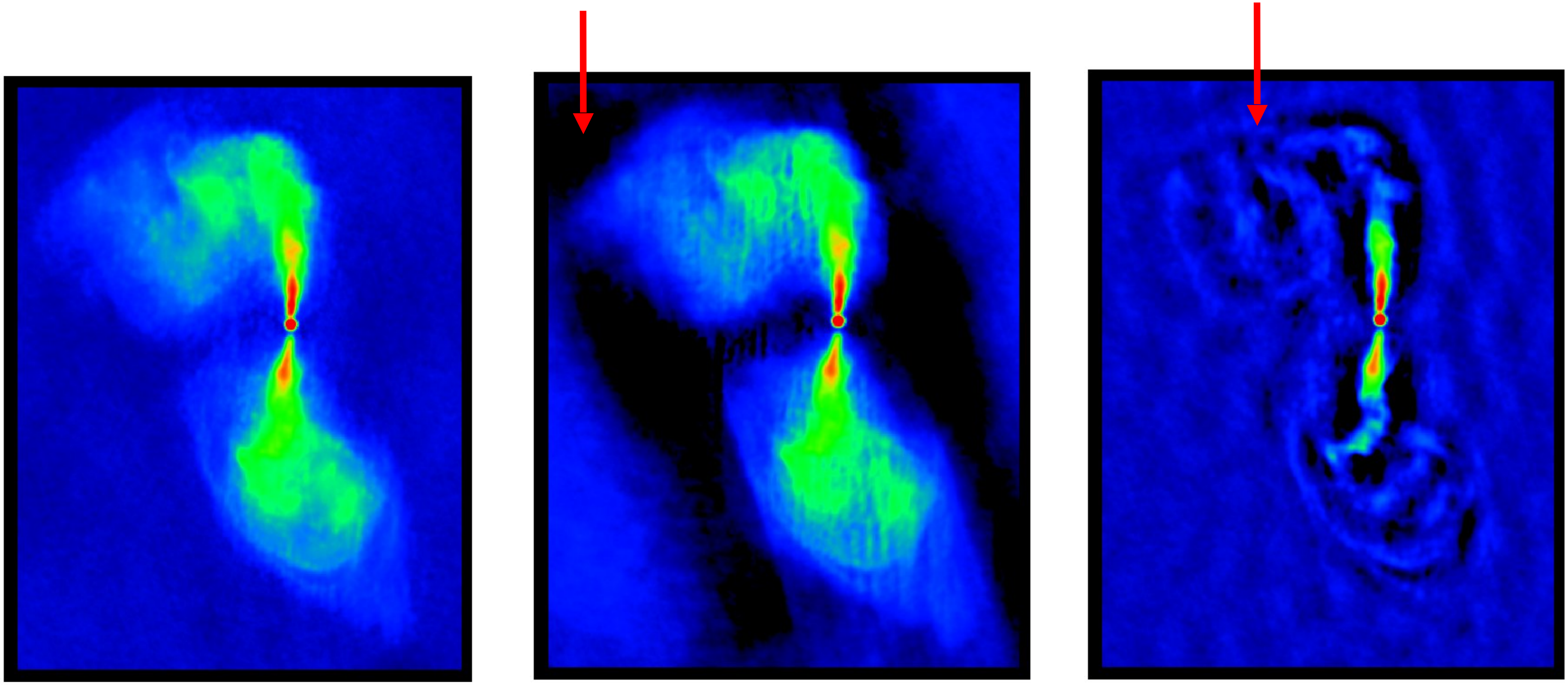


Compact source
on a pixel



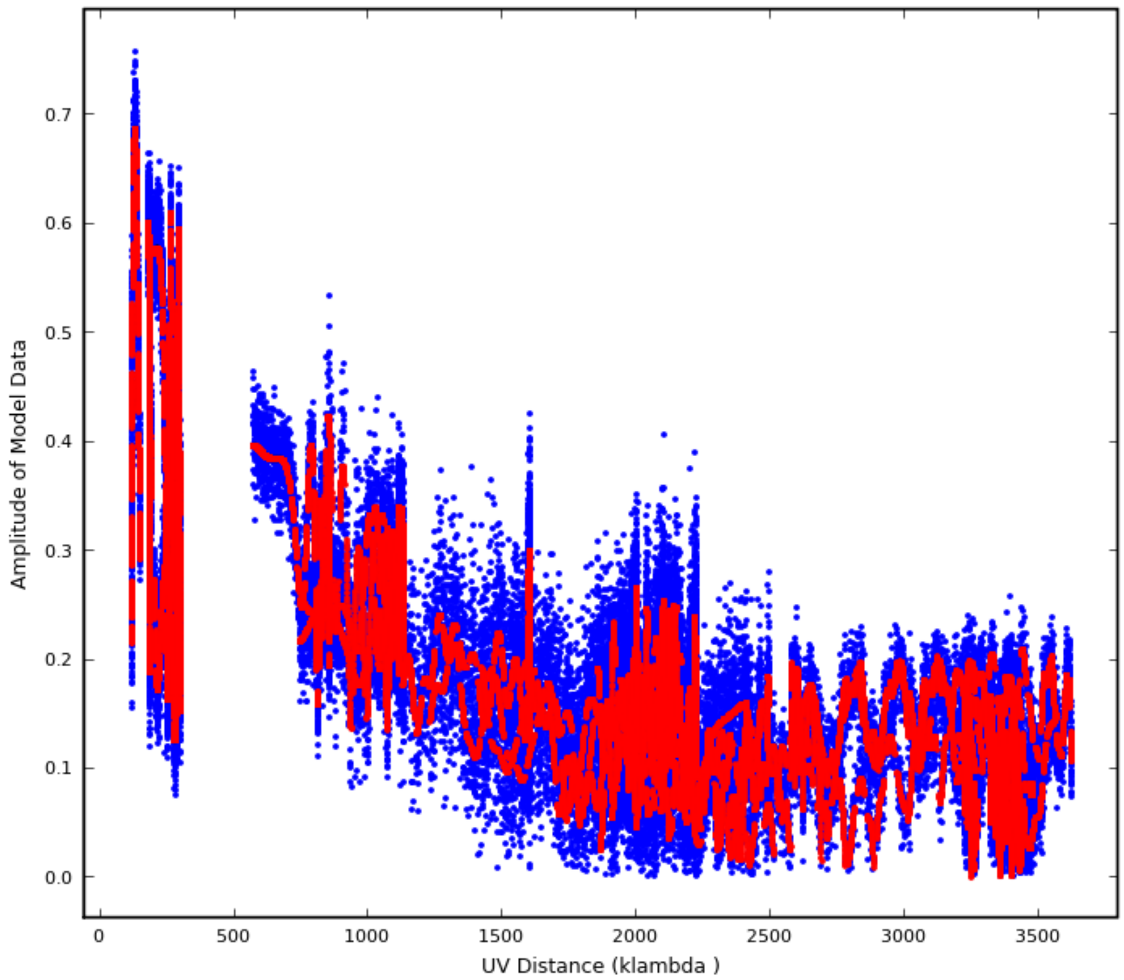
Mis-centred
by 0.5 pixel

# Missing short spacings




uv range < 225 kλ    uv range 2–225 kλ    uv range 10–225 kλ

# Does the model fit the data?



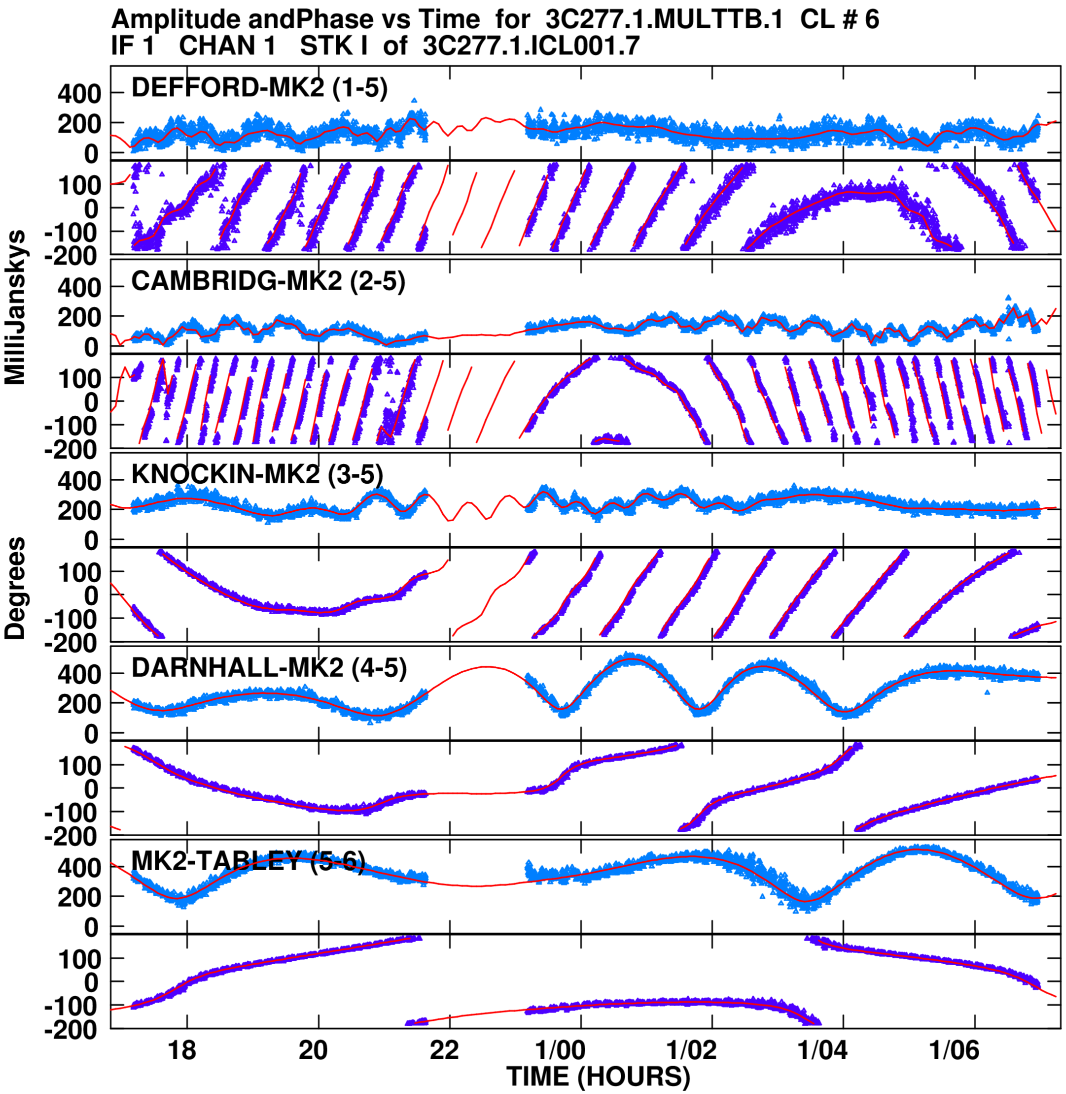

Plot amplitude against uv distance

Data

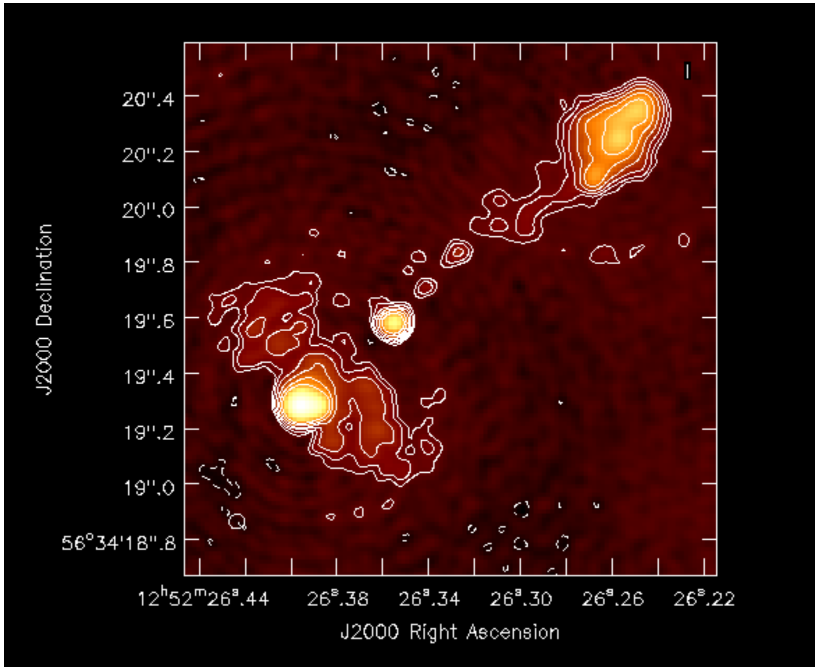
Model

Amplitudes are consistent

# Does the model fit the data?

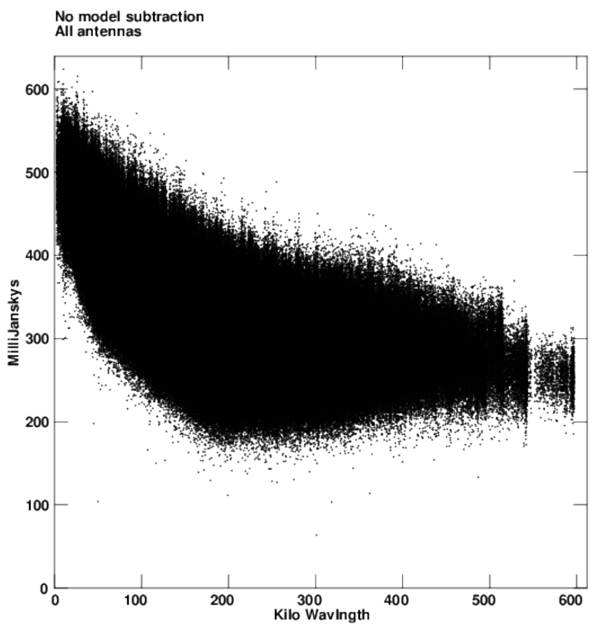


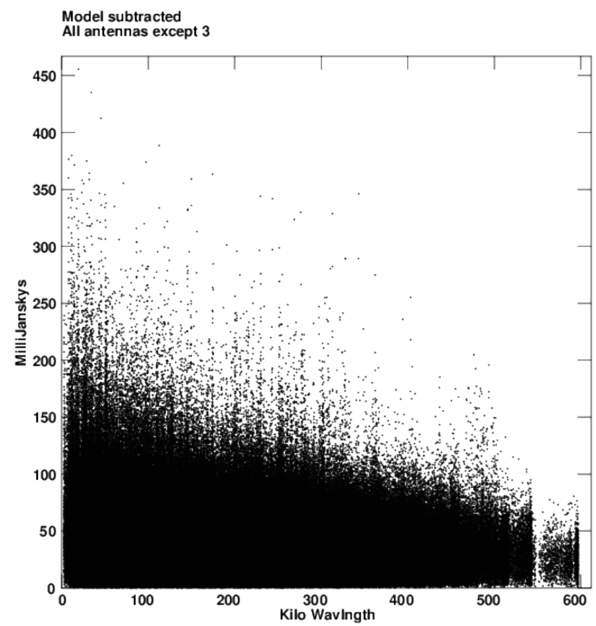

Phase fits

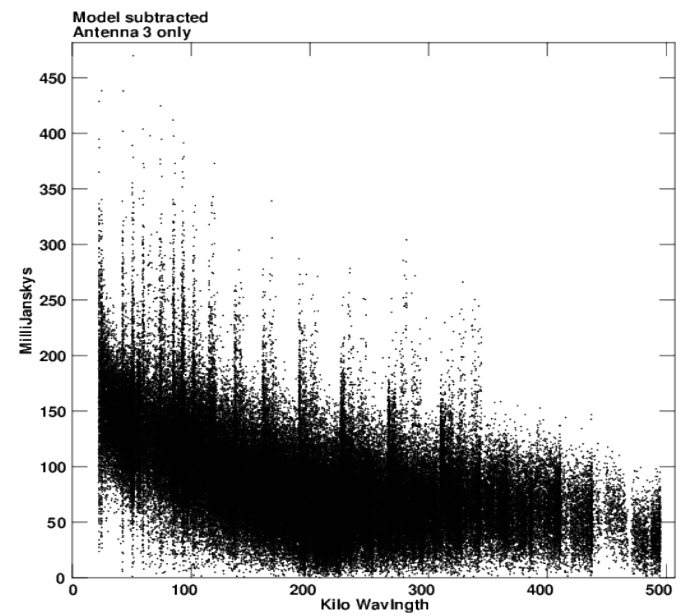


# Does the model fit the data

The case of a bright source with a low-level error




Error present (all antennas plotted) Nothing obvious

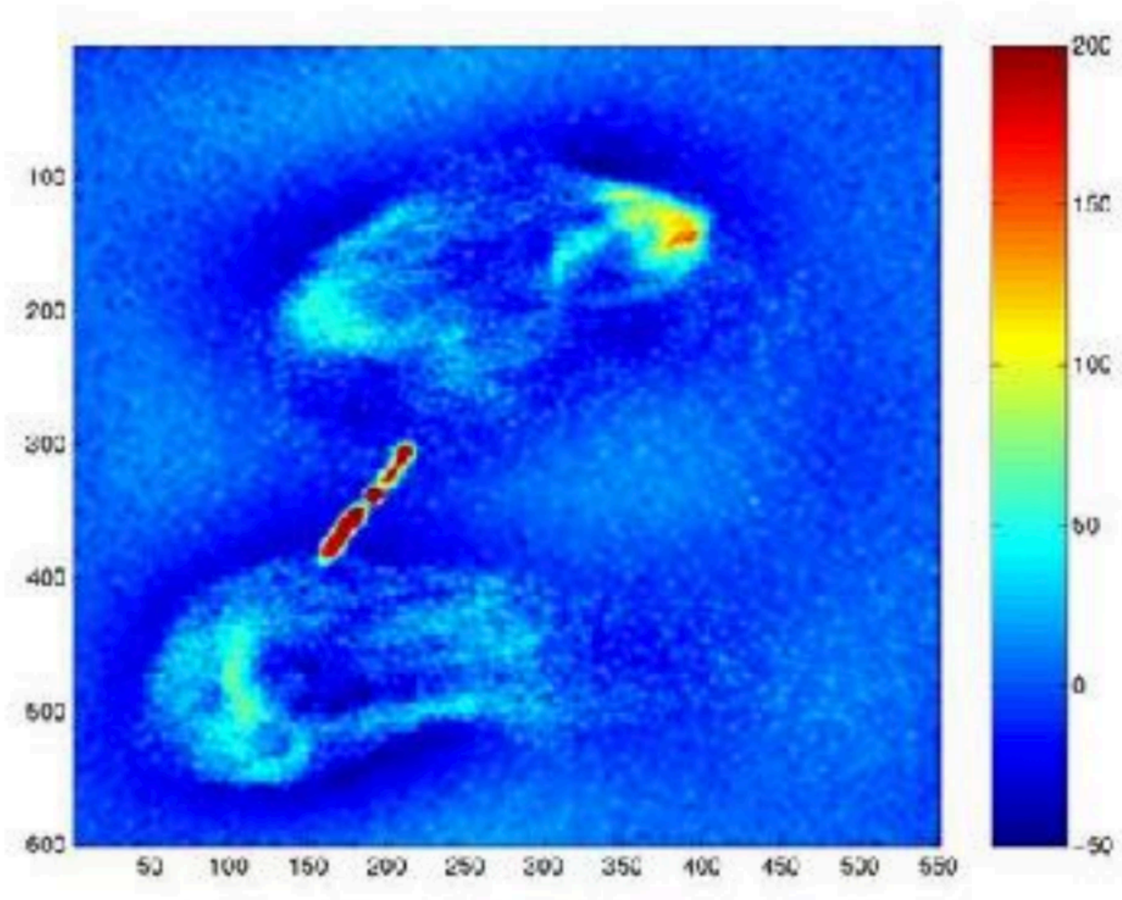
Model subtracted (all except antenna 3 plotted). Some discrepant data

Model subtracted (antenna 3 only). Mis-scales data clearly visible.

# QUIZ!

Credit: Emil Lenc

# What's wrong with this image?



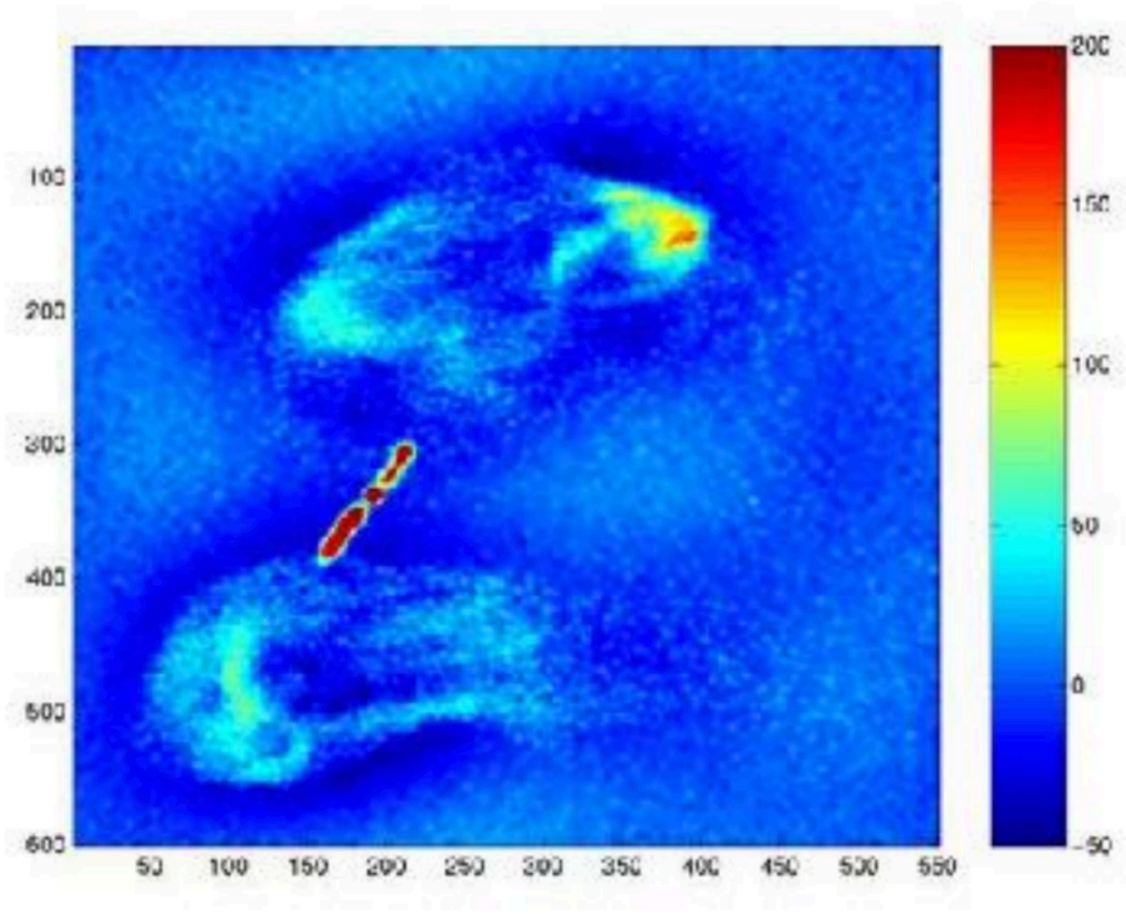

A: Missing short baselines

B: Missing long baselines

C: Alien entity

D: RFI

# What's wrong with this image?



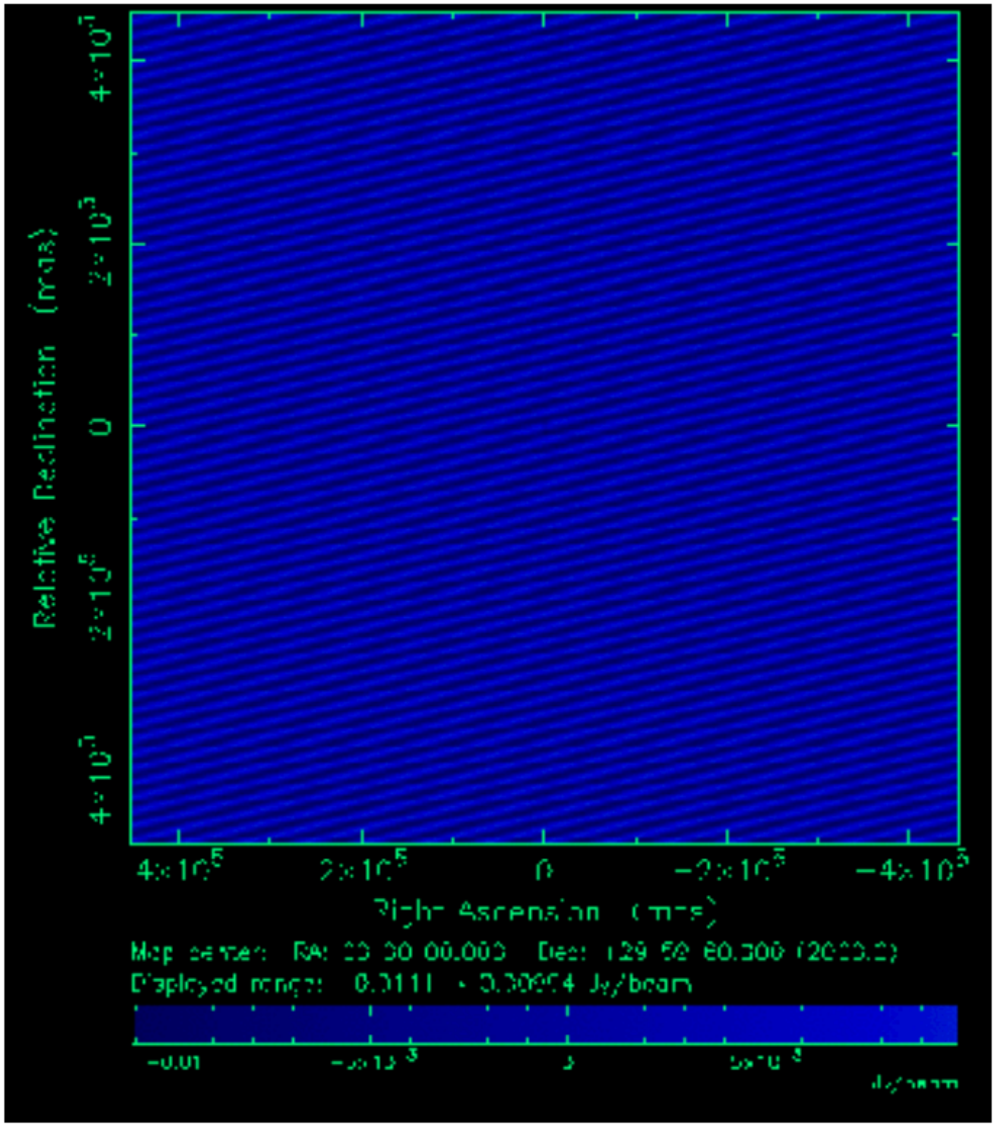

**A: Missing short baselines**

B: Missing long baselines

C: Alien entity

D: Phase errors

# What's wrong with this image?



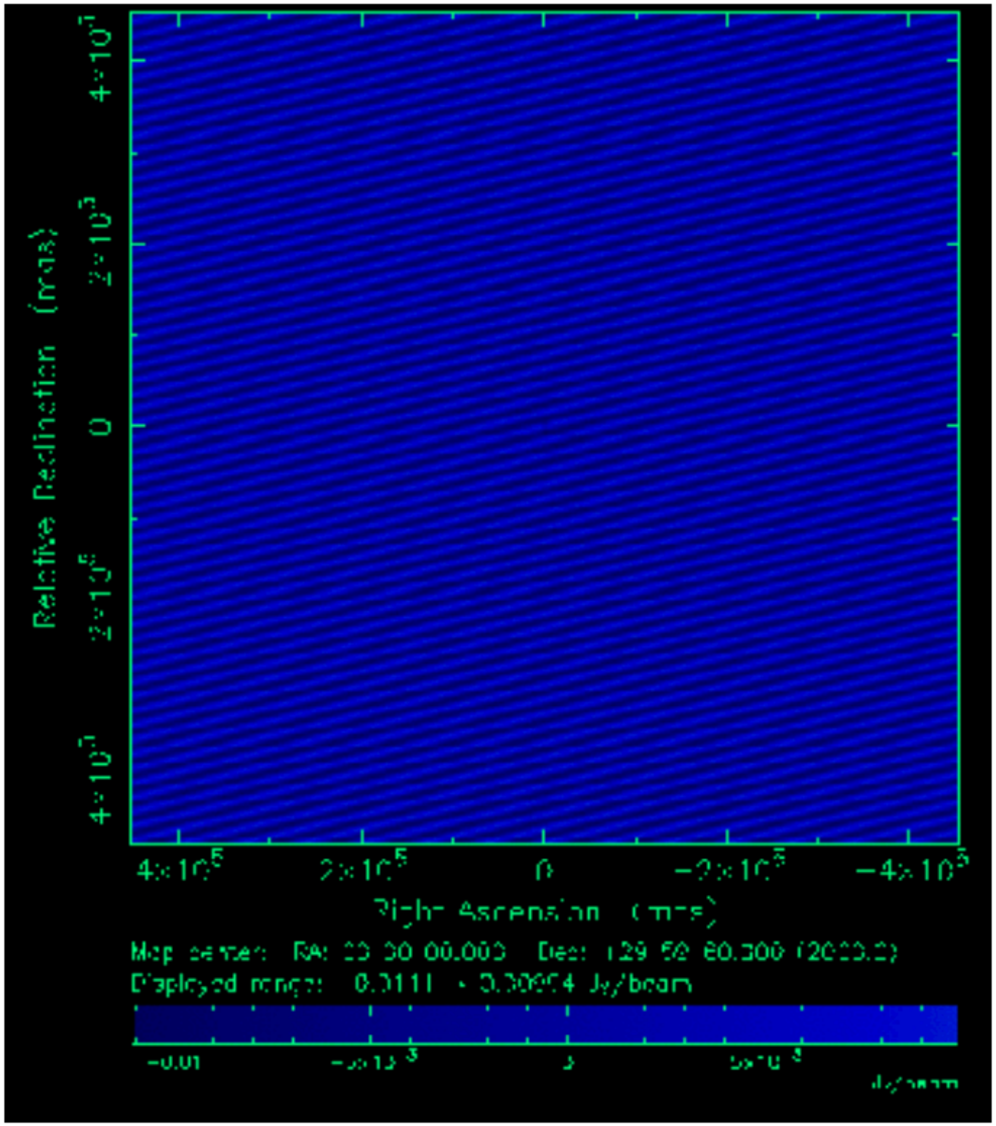

A: Bandwidth smearing

B: Source on edge of field

C: Left the blinds closed

D: RFI

# What’s wrong with this image?



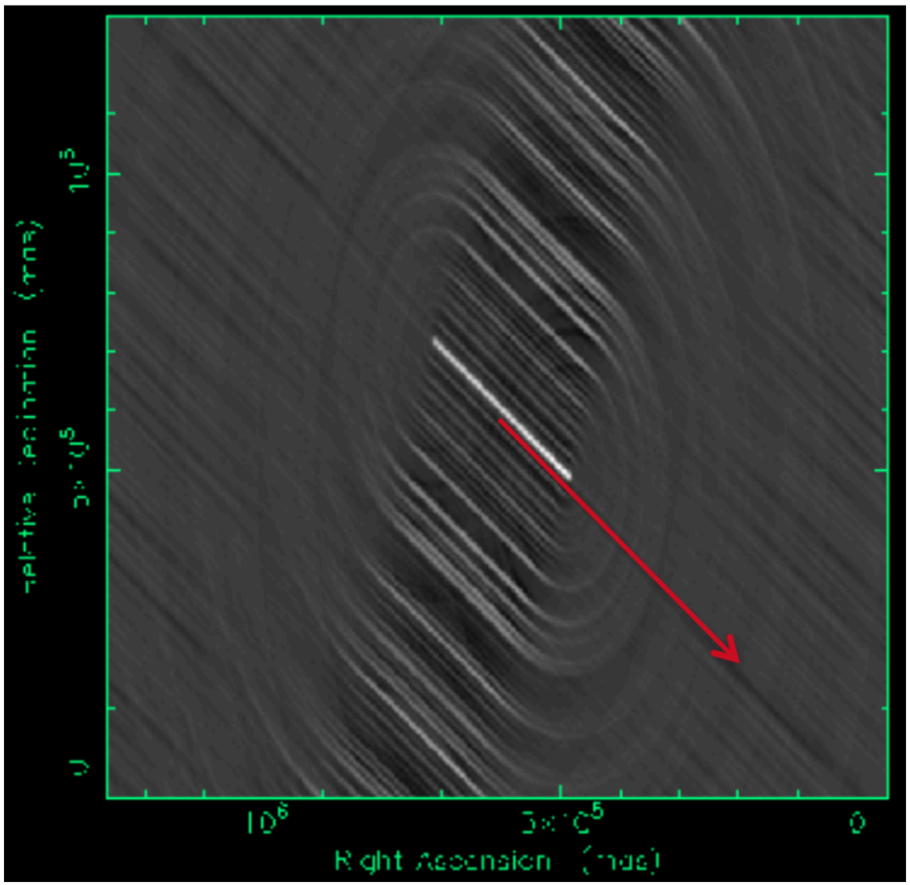

A: Bandwidth smearing

B: Source on edge of field

C: Left the blinds closed

**D: RFI**

# What’s wrong with this image?



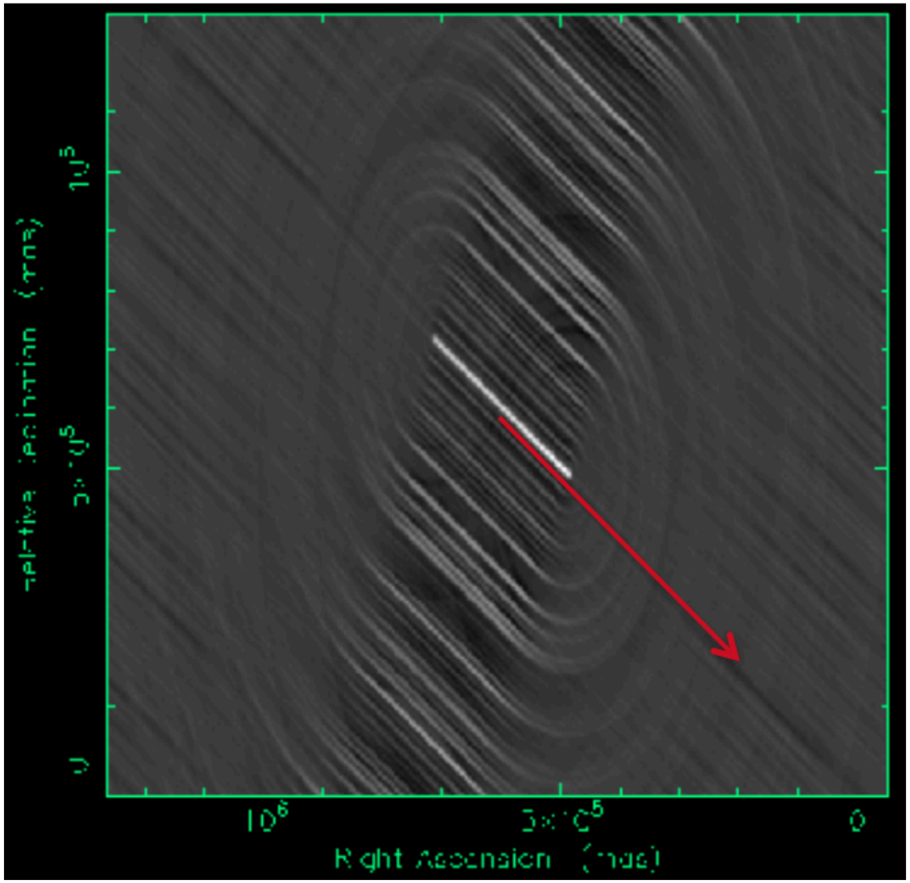

A: Bandwidth smearing

B: Amplitude error

C: Cosmic ray

D: Deconvolution error

# What’s wrong with this image?



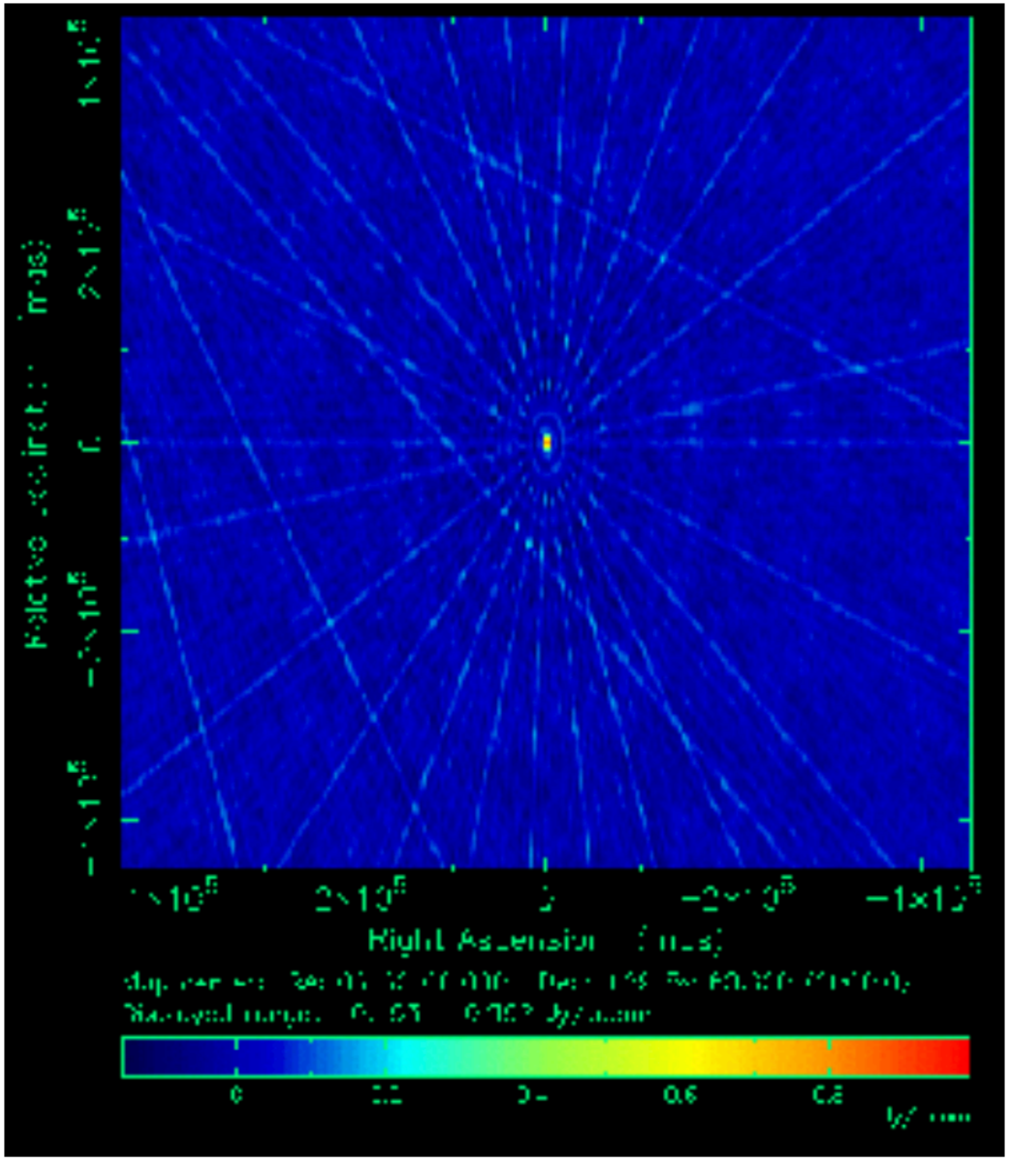
**A: Bandwidth smearing**

B: Amplitude error

C: Cosmic ray

D: Deconvolution error

# What’s wrong with this image?



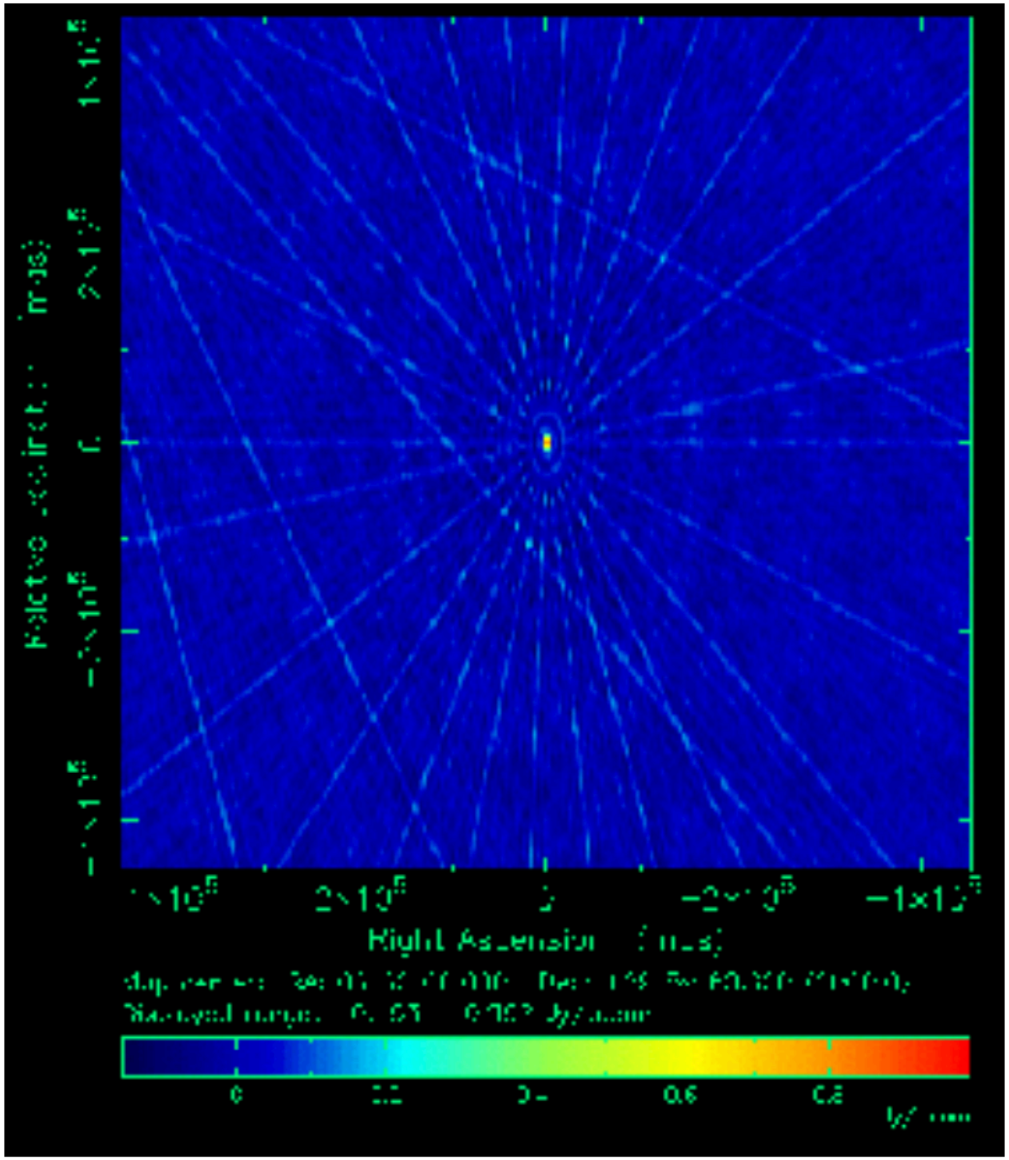
A: Phase error

B: Amplitude error

C: Deconvolution error

D: Source on edge of field

# What’s wrong with this image?



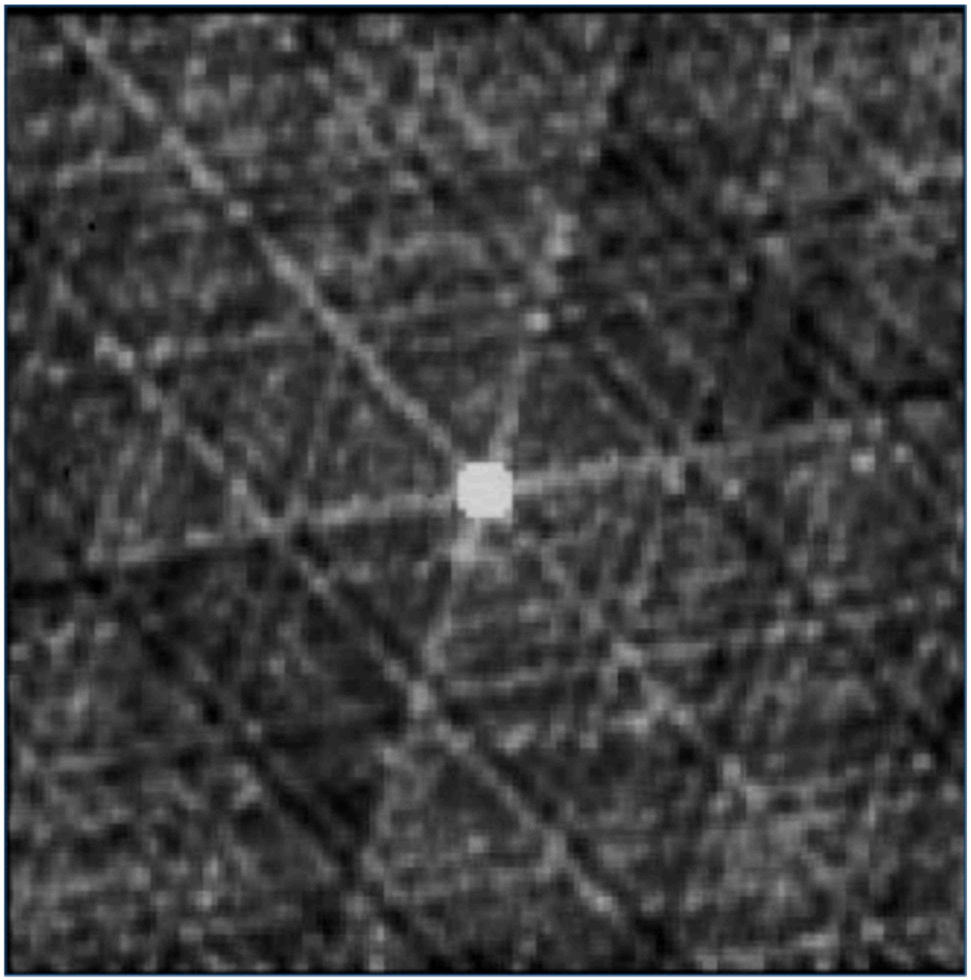

A: Phase error

B: Amplitude error

C: Deconvolution error

**D: Source on edge of field**

# What's wrong with this image?



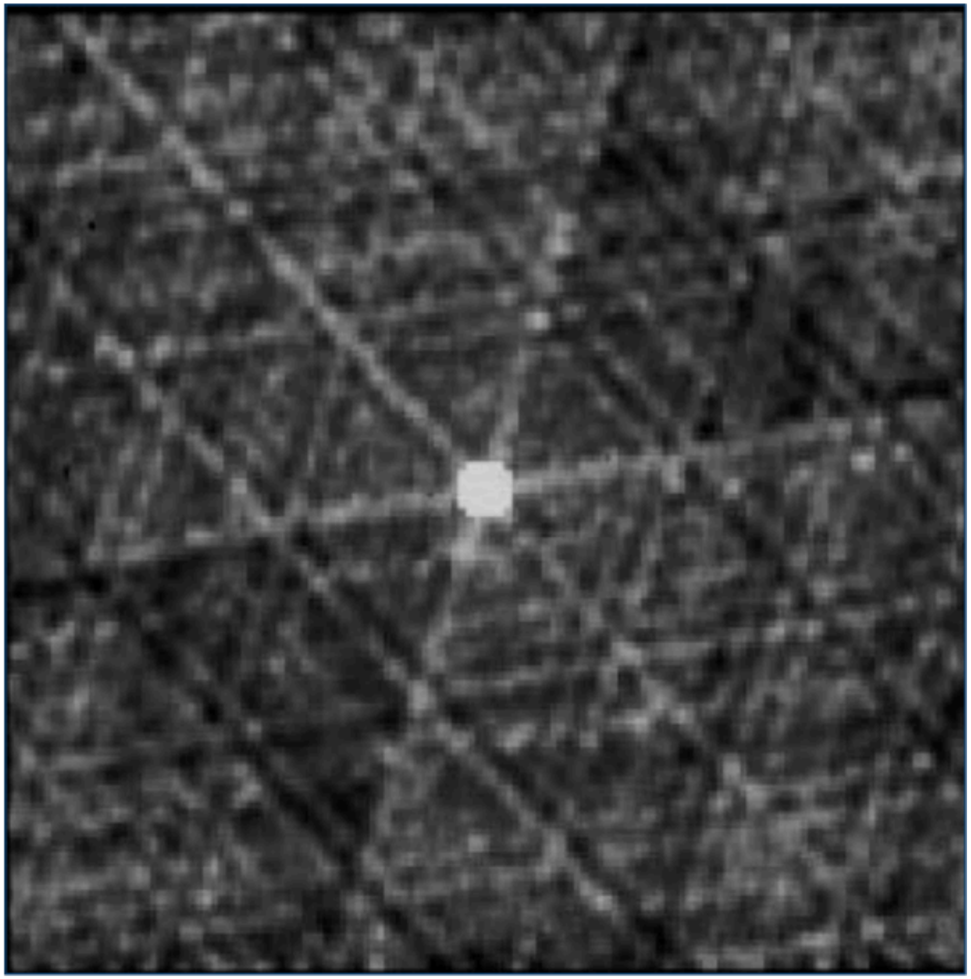
A: RFI

B: Amplitude error

C: Deconvolution error

D: Wrong type of tartan

# What's wrong with this image?



A: RFI

**B: Amplitude error**

C: Deconvolution error

D: Wrong type of tartan

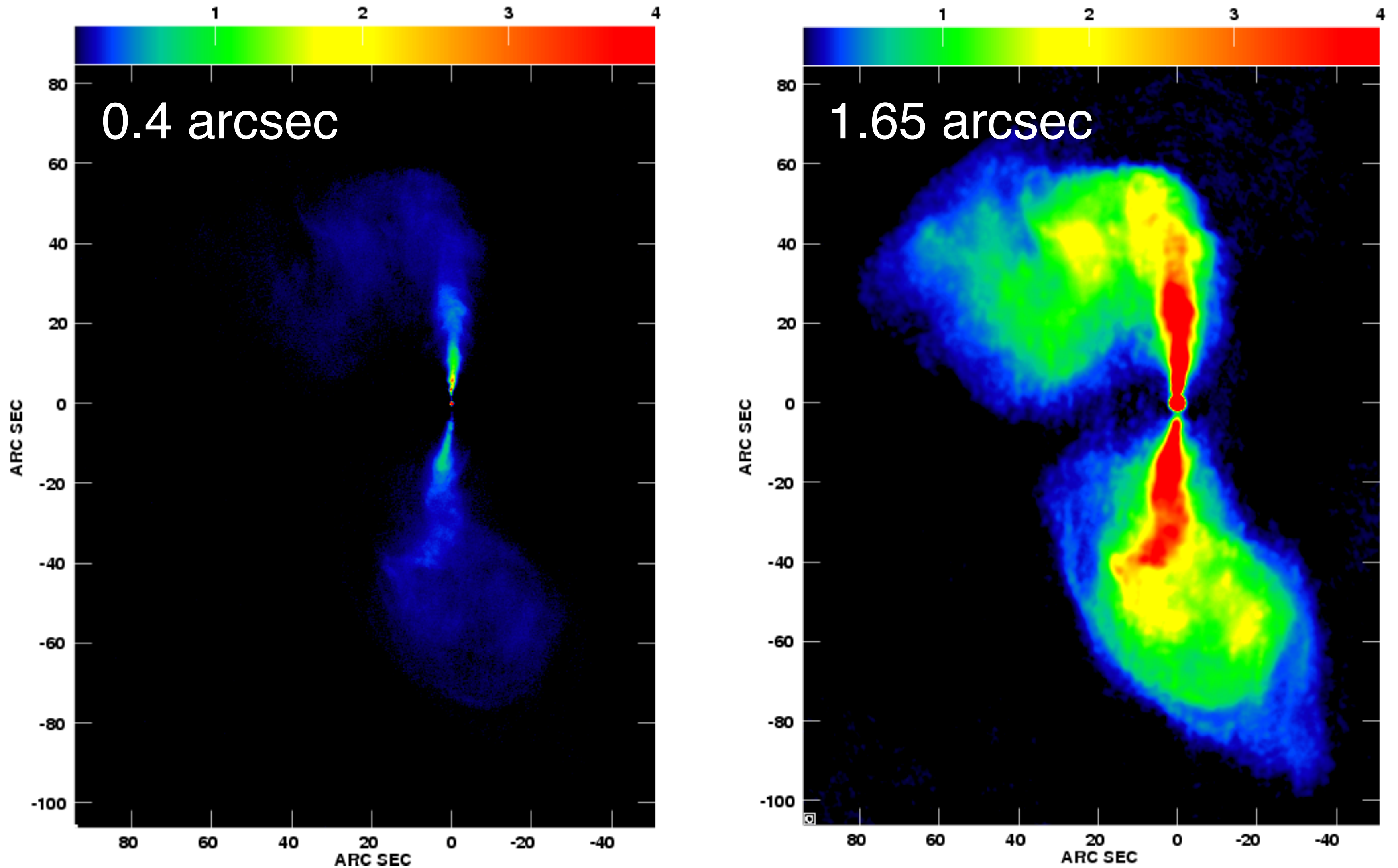
# Summary of error recognition

- (u,v) plane
  - Look for outliers (high or low) – flagging tutorial
  - Subtract the best model – check residuals in amplitude and phase
- Image plane
  - Do the defects look like the dirty beam?
  - Additive or multiplicative?
  - Symmetric or antisymmetric?
  - Missing spacings?
  - Deconvolution errors?

# Image Analysis

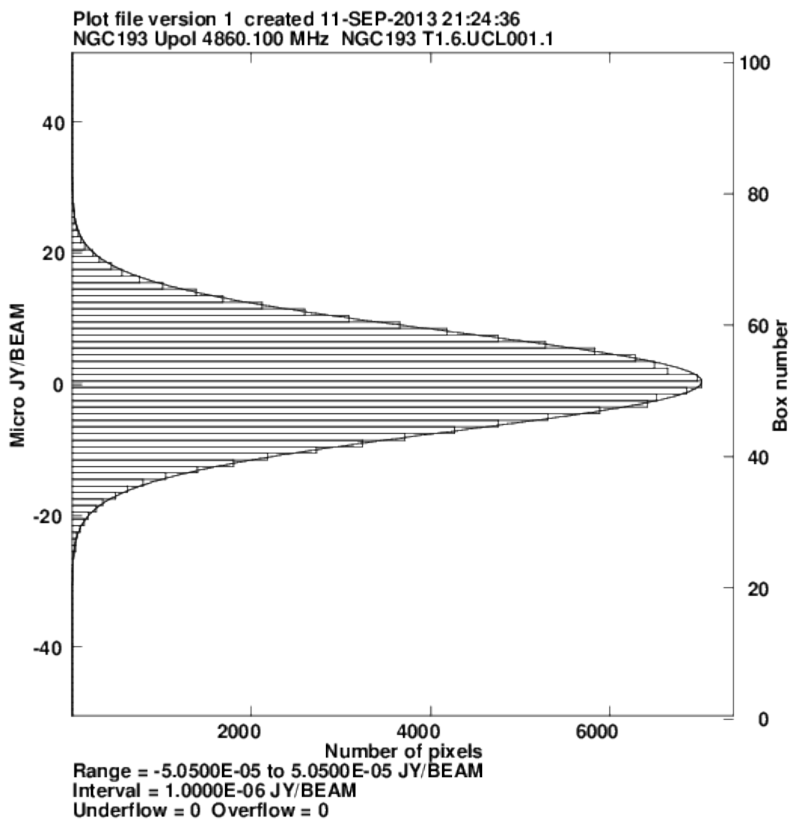
- Given: a well-calibrated dataset producing a high-quality image (or, in general, image cube)
- How can we extract scientifically useful numbers?
- This is a very open-ended problem, depending on:
  - image complexity
  - scientific goals
- Selected topics (excluding spectral line):
  - Picking the correct resolution
  - Parameter estimation
  - Comparing images: spectra, polarization etc.; registration
  - Getting images into your own code

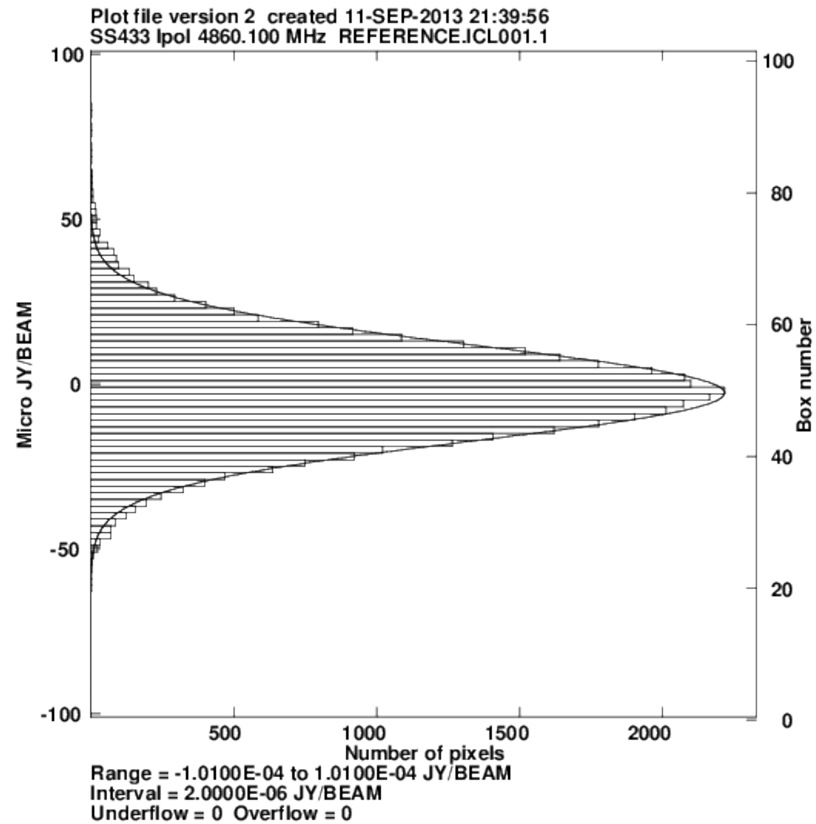
# Match the resolution to the problem




Exactly the same dataset, imaged with different Gaussian tapers

# Measure the off-source noise distribution




Good case: rms = 7.5μJy; Gaussian noise with zero mean

Excess noise above Gaussian tail

# Estimating the flux density of an extended source

- Use a **low-resolution** image, cleaned deeply
  - The beam areas of the restored CLEAN components and residuals are not the same in general.
- Sum the flux density over some area (rectangular, polygonal, ...) – casa imstat, viewer.
- Remember that the total flux density is $\sum I/B$, where *B* is the integral over the beam. For a Gaussian,

$$B = \pi(\mathrm{FWHM}/\mathrm{pixel})^2/4\ln(2)$$

The reduction packages will calculate this for you.

- The reason is that the images are normalised so that a point source of flux density 1 Jy gives a **peak** response of 1 Jy/beam on the image.

# Component fitting

- Image plane: fit 2D elliptical Gaussian components
  - Assume source components are ~Gaussian
  - Size estimation quite straightforward
  - casa imfit (AIPS SAD for very many components)
- u-v plane
  - More accurate for small numbers of ~point-like sources
  - Can fit models slightly more complex than point-like
  - Accounts for imperfect sampling, but not good for very complex brightness distributions
- Error estimates
  - Analytic (position errors derived from phase noise)
  - From fitting routines
  - By simulation

# Error estimates for Gaussian fits

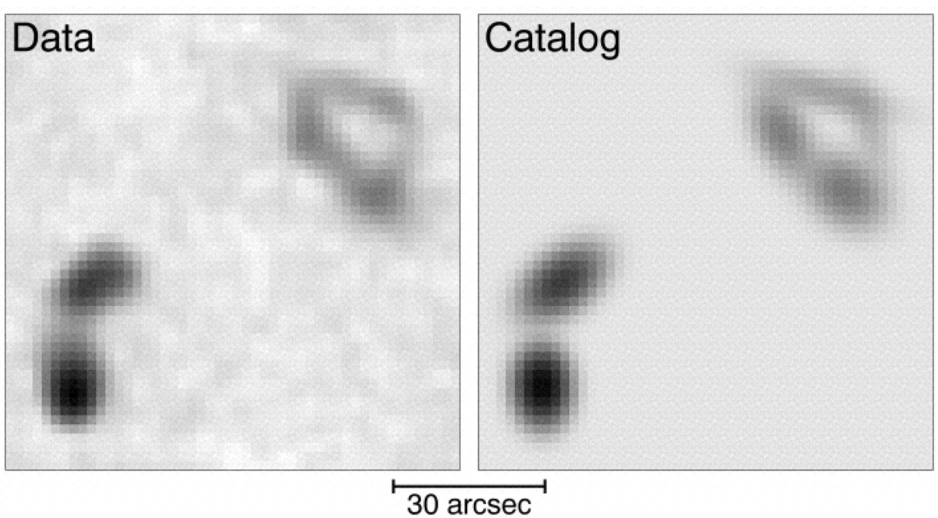
- Definitions
  - P = peak component flux density
  - $\sigma$ = image rms noise
  - $\theta_B$ = CLEAN beam size
  - $\theta_{obs}$ = component size
  - S/N = P/$\sigma$ = signal/noise

Assume Gaussian distribution of random noise in the image plane – well-calibrated, well-imaged maps

- rms errors
  - Error on peak flux density = $\sigma$
  - Position error = $\theta_B/2$ (S/N) (or worse for sparse arrays)
  - True component size $\theta = (\theta_{obs}^2 - \theta_B^2)^{1/2}$
  - Minimum measurable component size = $\theta_B/S^{1/2}$
    - S/N >100 is needed to determine a size $<\theta_B/10$.

# Automated image fitting

- Automated routines can be used to locate and fit sources (essential for surveys). SAD in AIPS is a good example. casa is weak in this area.
- Also adapt routines used in optical astronomy (e.g. SExtractor)
    - beware incorrect noise model
- Often worthwhile to make Monte Carlo simulations to assess realistic errors in position and (especially) flux density (e.g. add model point sources).




Catalog plot shows sizes, position angles and flux densities of components fitted automatically to images from the FIRST survey.

# Basic image arithmetic

- Standard packages allow mathematical operations on one or more images (casa immath, various AIPS tasks):
  - Sum, product, quotient, ...
  - Spectral index α ($S \propto \nu^{\alpha}$)
  - Faraday rotation measure
  - Optical depth
- Can create masks (e.g. S/N)
- Other image manipulations (spatial filtering, etc.) are also possible
- Fitting to images at more than 2 frequencies under development
  - casa rmfit, spxfit

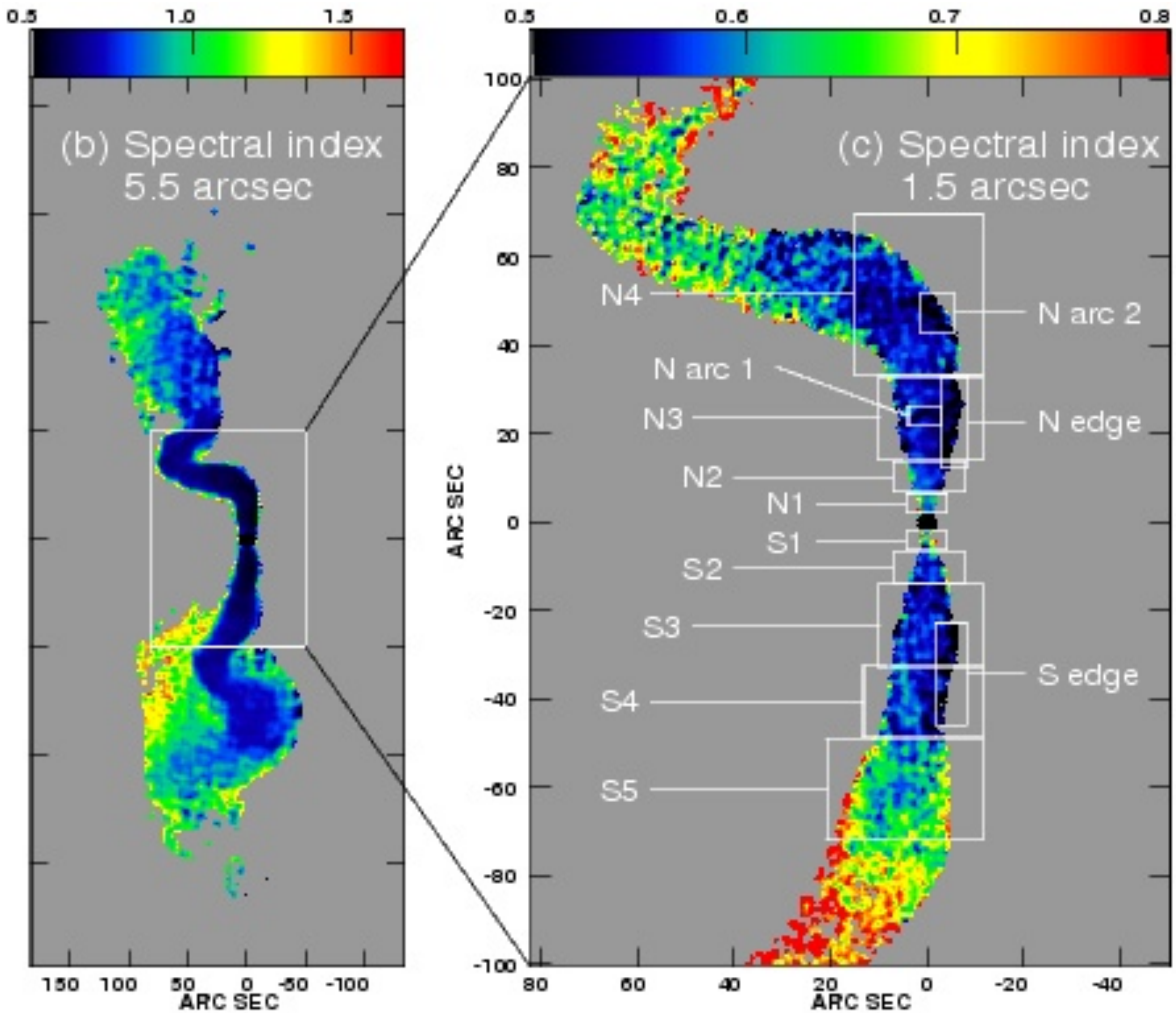
# Basic image manipulation

- Often useful to make subimages
  - casa imsubimage
- Smooth images (e.g. if the restoring beam is not quite what you want)
  - Gaussian or user-supplied kernel
  - casa imsmooth
- Regridding images
  - Often needed (e.g.) to align two images with different coordinates, pixel sizes for comparison
  - casa imregrid

# Comparing images at different frequencies

- Match the resolutions
  - Pick appropriate weighting and Gaussian taper to get approximately the same dirty beam FWHM
  - Restore with the same beam
  - Possible to combine uv data and use MFS imaging with spectral index or even curvature
- Error propagation
  - Gaussian random noise in the image plane is the best case: you can only do worse
  - Be careful near edges of the source and sharp brightness gradients

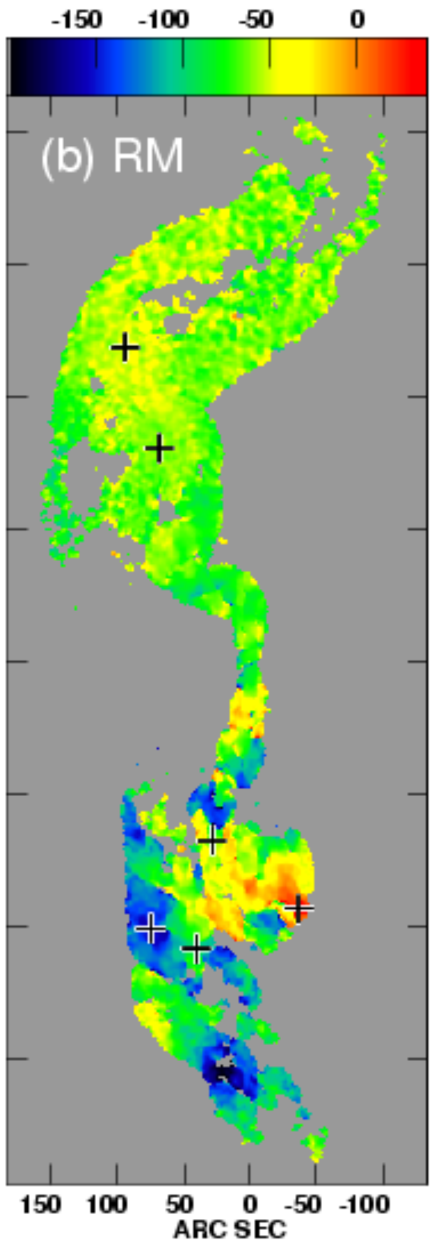
# Spectral index




Flux density changes with frequency; synchrotron follows power-law

$$S \propto \nu^{-\alpha}$$

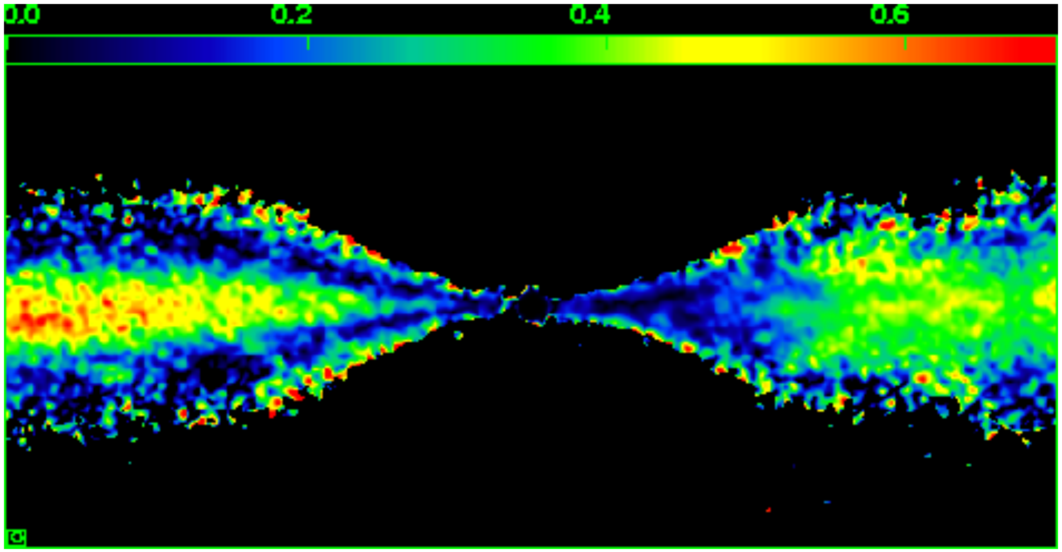
# Rotation Measure



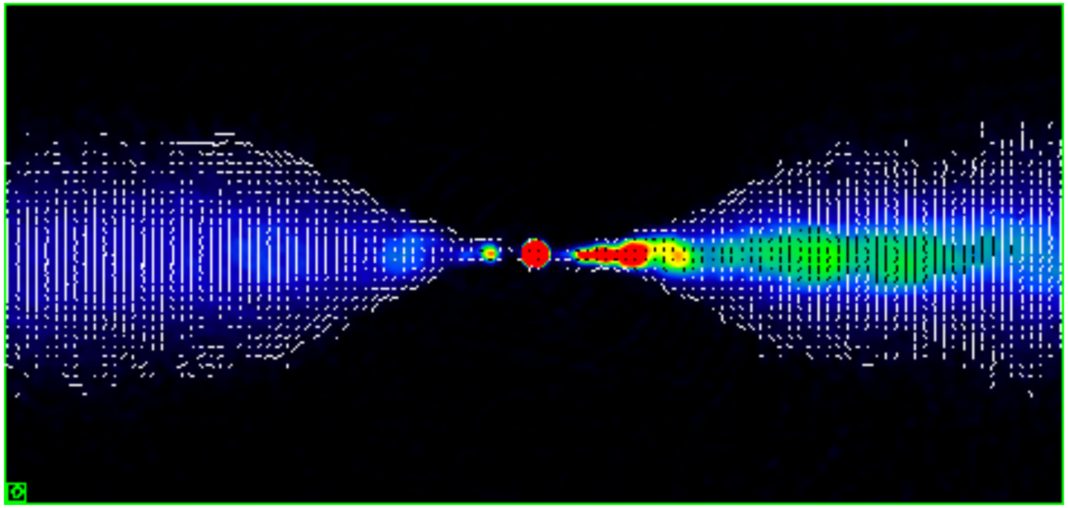

Polarisation position angle of radio wave changes as it passes through B field

$$PA = PA(0) + RM\ \lambda^2$$

# Displaying polarisation data



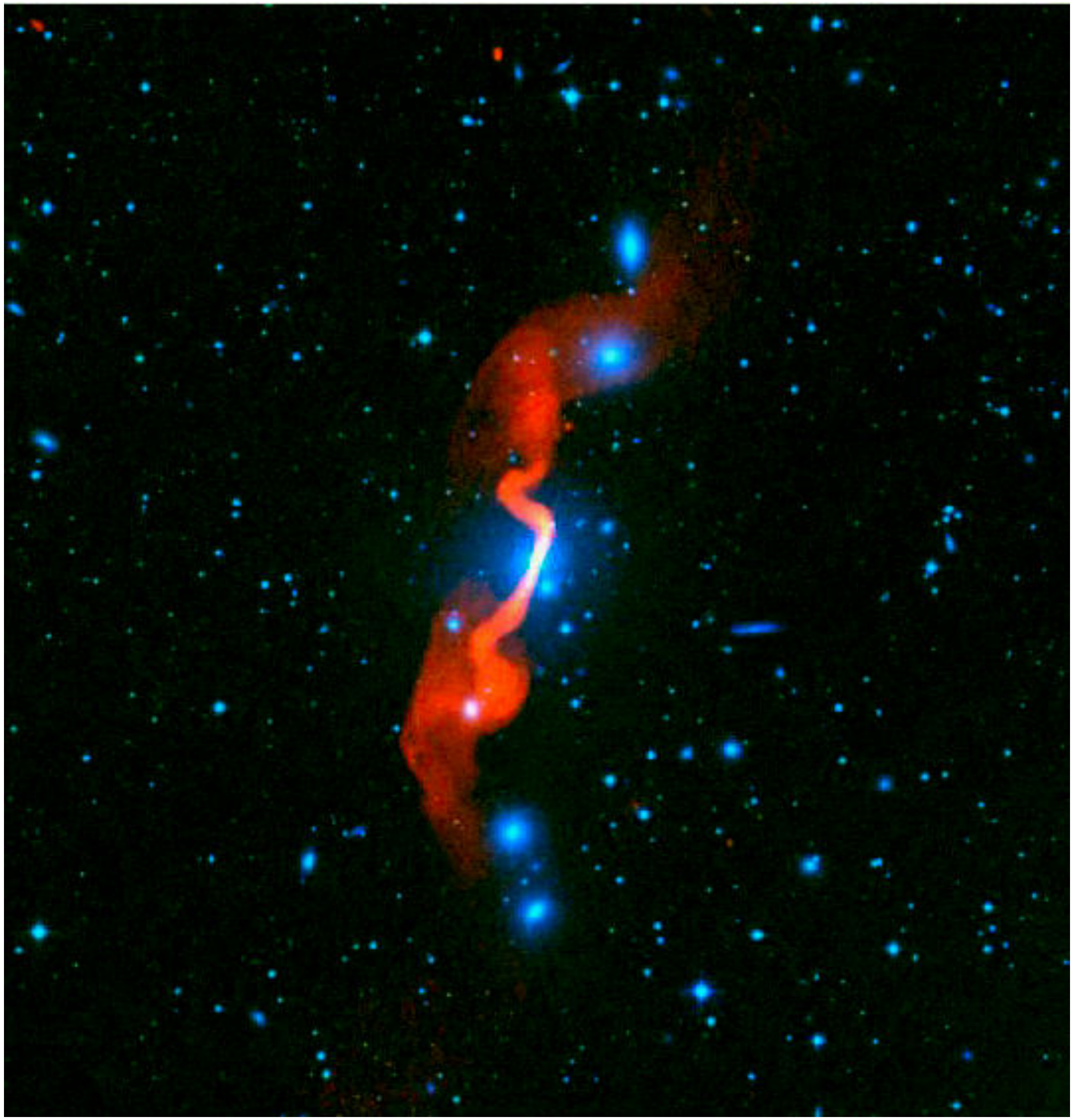

Fractional polarisation
p = P/I



Vectors; lengths ∝ p,

directions B-field, after correction for Faraday rotation.

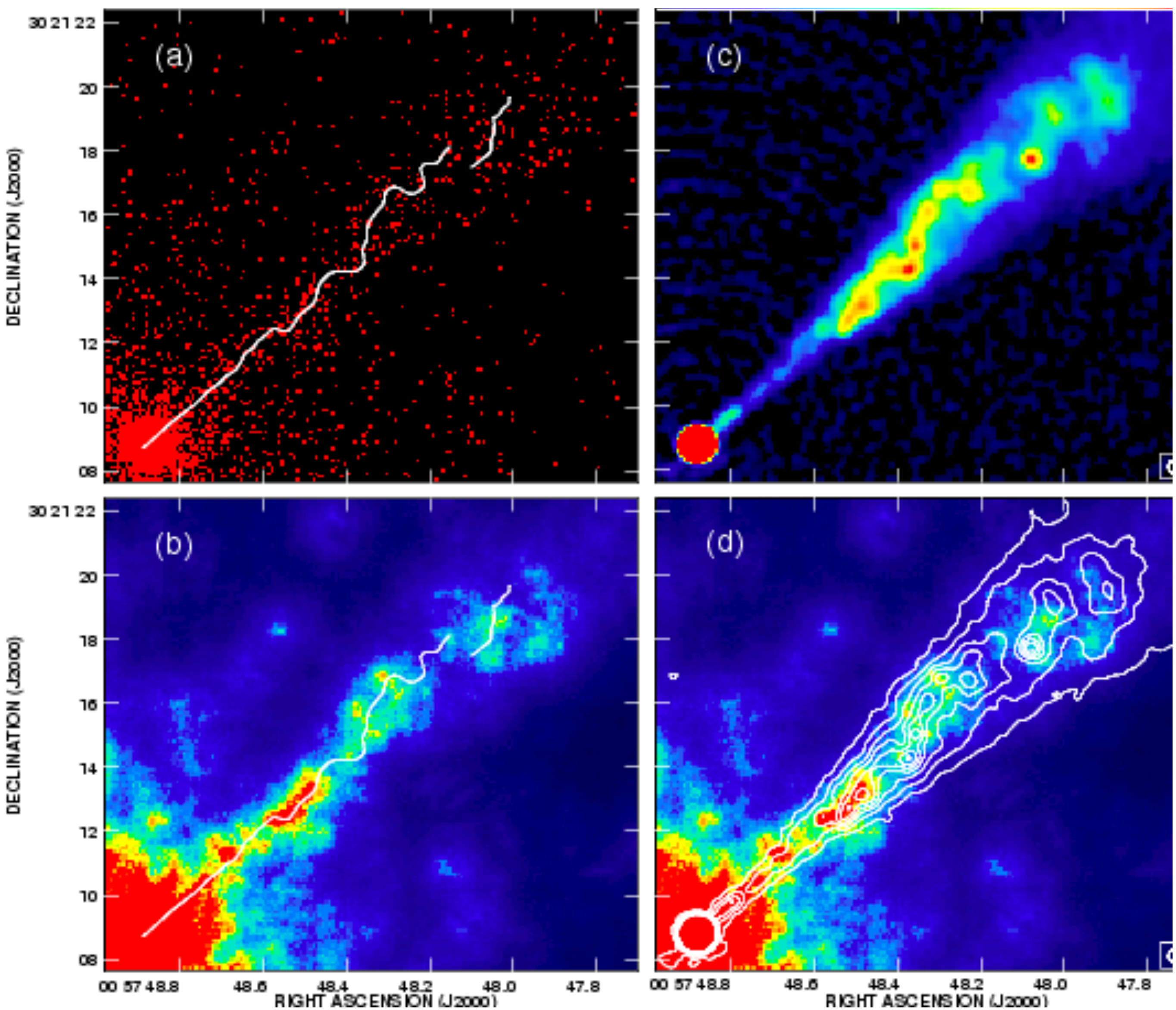
# Regridding: radio – optical overlay



1.4GHz radio (VLA) in red

Optical (DSS) in blue

# Regridding: radio – X-ray overlay



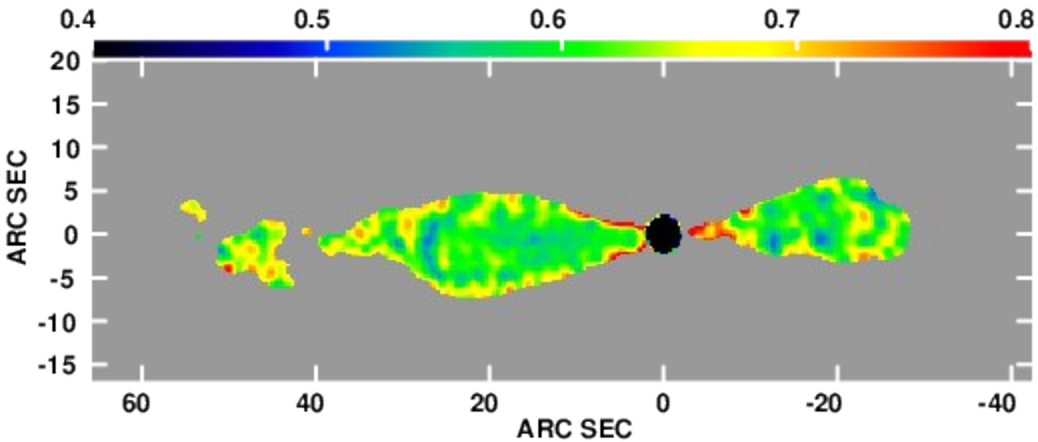

Radio (0.35 arcsec, 4.9 GHz, VLA)

X-ray (0.6 arcsec, 0.5 – 8 keV)

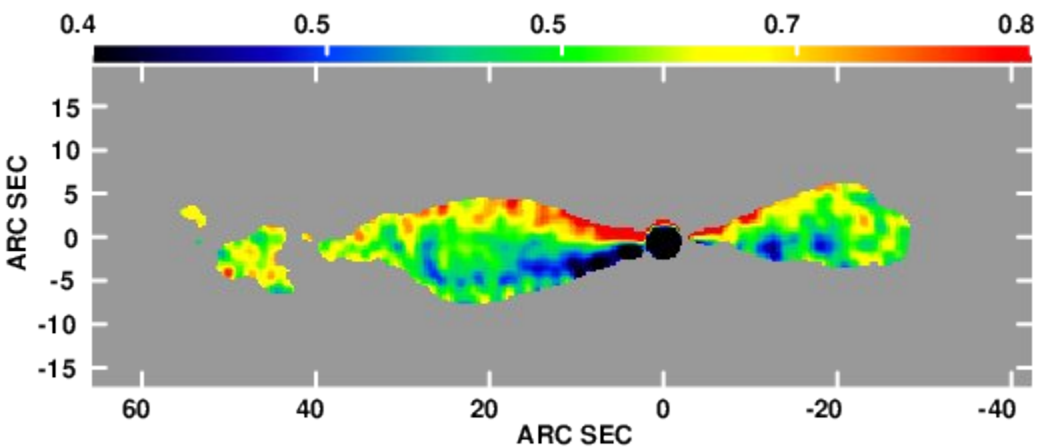
# Issues in image registration

- Rationale for image combination
  - Multiwavelength comparisons
  - Proper motions
- Regridding
  - Tools available (casa imregrid)
- Accuracy of registration
  - Need accurate pre-self-cal phase calibration: ideally,
    - calibrator is close to the target
    - use the same phase calibrator for all observations
      - Use in-beam phase referencing if possible
  - Beware changes in structure with frequency
  - N.B.: images at other wavebands may have less accurate absolute astrometry

# Registration errors




Spectral index between 1.365 and 4.9 GHz

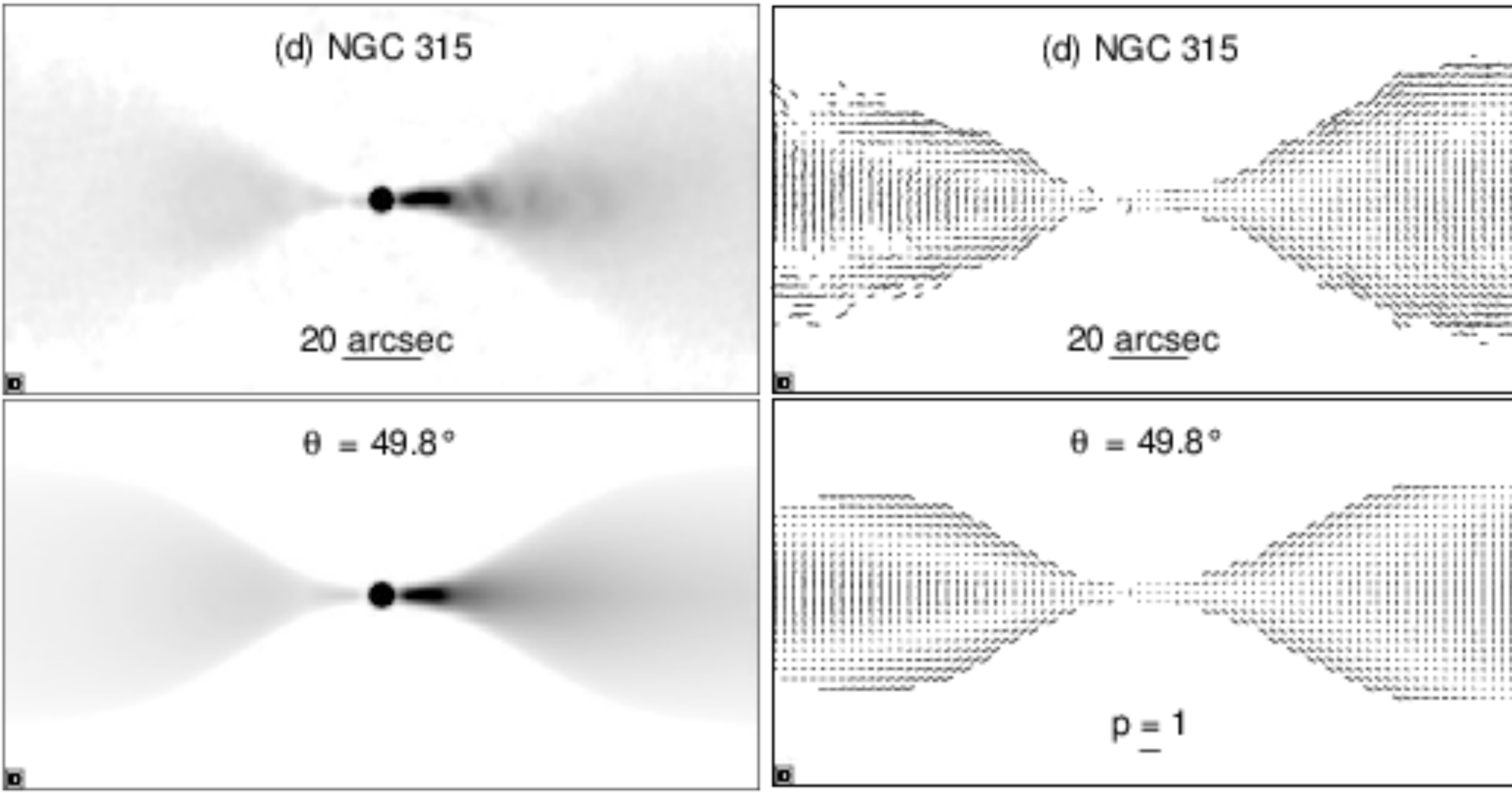



Relative shift of 0.2 FWHM

# Getting data into your own code

- If the standard packages do not do what you want, do not be frightened of importing images into your own programs.
- FITS interchange standard
  - can be read and written by all radio astronomy packages
  - mostly standard for images (uv less so)
  - well documented interfaces to common programming languages (python, C, fortran, IDL, ...)
  - and even to graphics manipulation packages (gimp)
- Easy read/write from casa to python arrays

# Example: jet modelling




Total intensity

Vectors p/apparent B field

# Summary of continuum image analysis

- Match the resolution to the problem
- For simple images, fit component parameters and derive errors
- Image comparison
  - Simple mathematical operations are easy
  - Regridding and interpolation often required
  - Registration is an issue
  - Noise propagation
- Straightforward to read images into your own code for more sophisticated modelling