

Error Recognition

Greg Taylor (UNM)

With help from:

Urvashi Rao, Sanjay Bhatnagar, Gustaaf van Moorsel,



Sixteenth Synthesis Imaging Workshop

16-23 May 2018



INTRODUCTION

- How are these two topics – ‘Error Recognition’ and ‘Image and Non-Imaging Analysis’ related?
 - **Error recognition** is used to determine defects in the (visibility) data and image during and after the ‘best’ calibration, editing, etc.
 - **Image analysis** describes the many ways in which useful insight, information and parameters can be extracted from the image.
 - **non-imaging analysis** describes how to extract information directly from the (u,v) data
- Perhaps these topics are related to the reaction one has when looking at an image after ‘good’ calibration, editing, self-calibration, etc.
- If the reaction is:

Fantastic Discovery or an Obvious Defect?

VLBA observations of SgrA* at 43 GHz

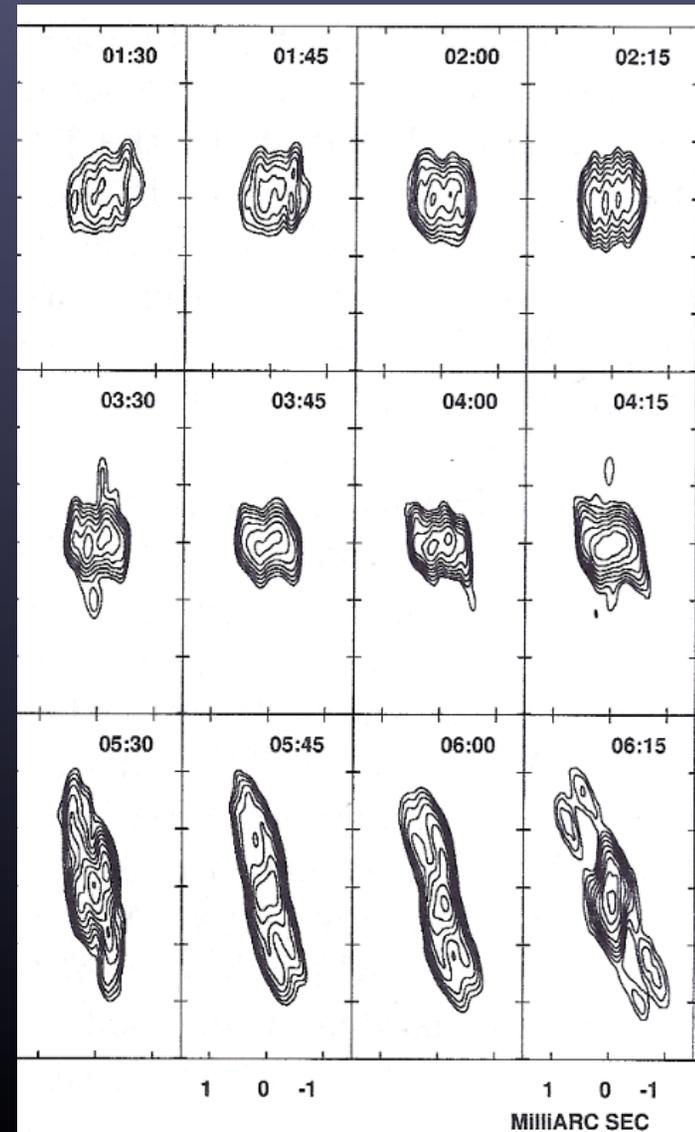
This can't be right. Either SgrA* has bidirectional jets that nobody else has ever seen or:

Clear signs of problems:

Image rms > expected rms

Unnatural features in the image

How can the problems be found and corrected?



Miyoshi et al. 2005

milliarcsec

HIGH QUALITY IMAGES

Reality

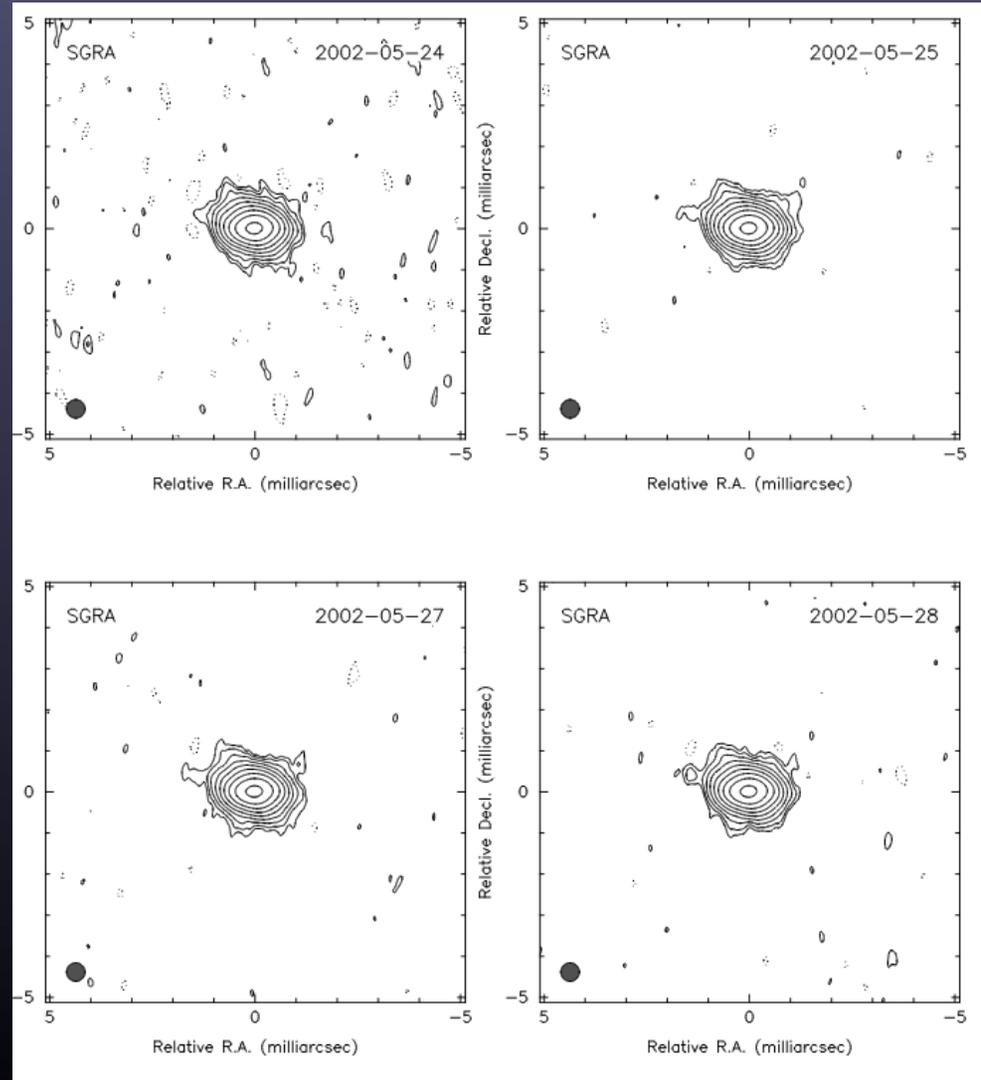
With care we can obtain good images.

What were defects?

Two antennas had ~30% calibration errors at low elevations.

This lecture.

How to find the errors and remove them.



milliarcsec

GENERAL PROCEDURE

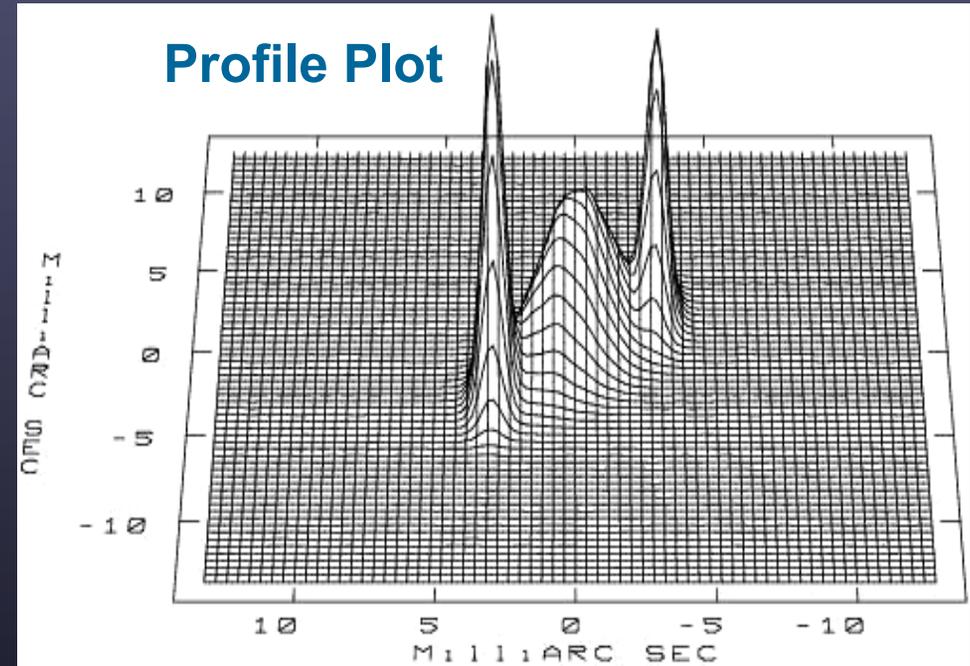
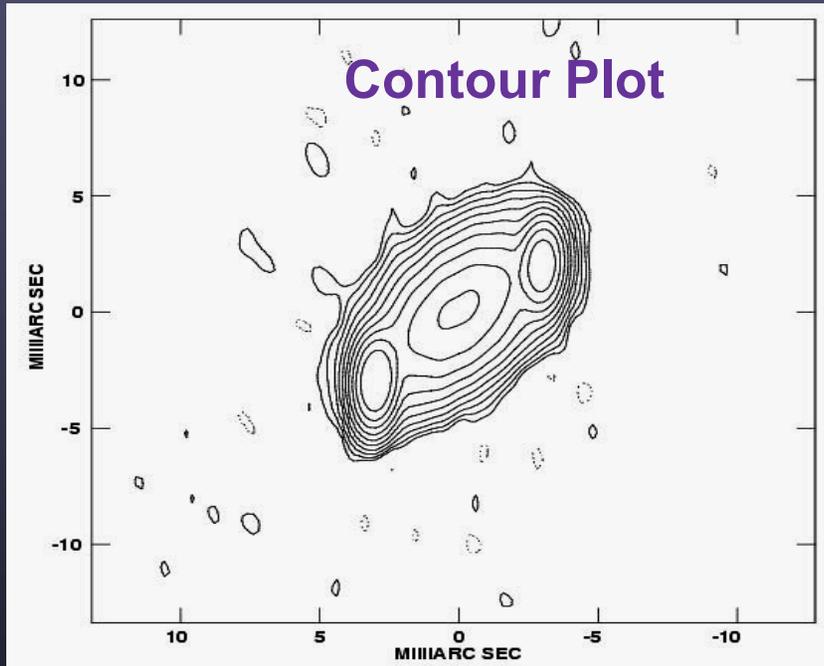
We assume that the data have been edited and calibrated reasonably successfully (earlier lectures) including self-calibration if necessary.

So, the first serious display of an image leads one—

- to inspect again and clean-up the data repeating some or all of the previous reduction steps.
 - removal of one type of problem can reveal next problem!
- once all is well, proceed to image-analysis and obtaining scientific results from the image.

But, first a digression on data and image display. First:
Images

IMAGE DISPLAYS (2)

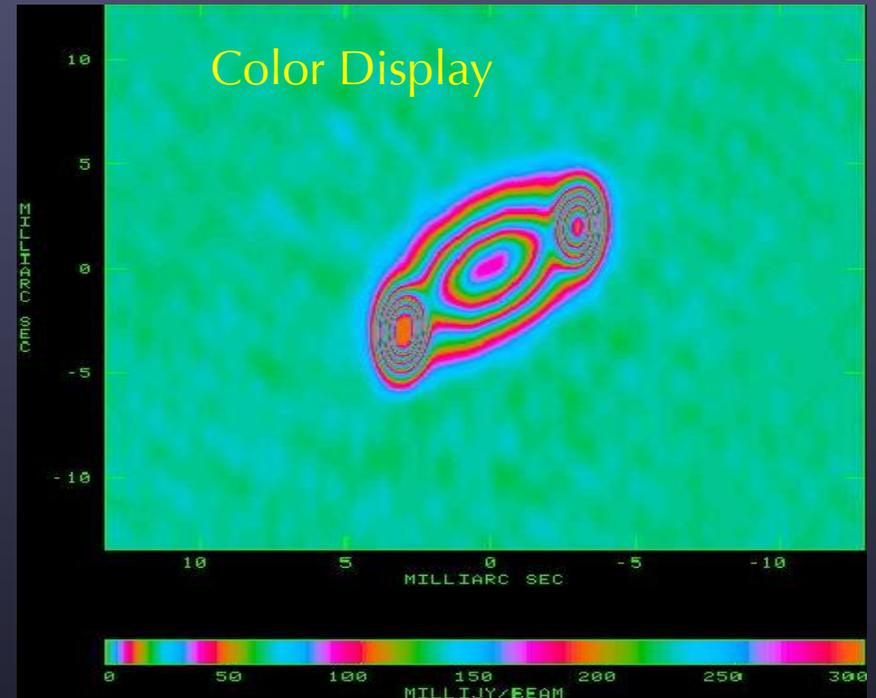
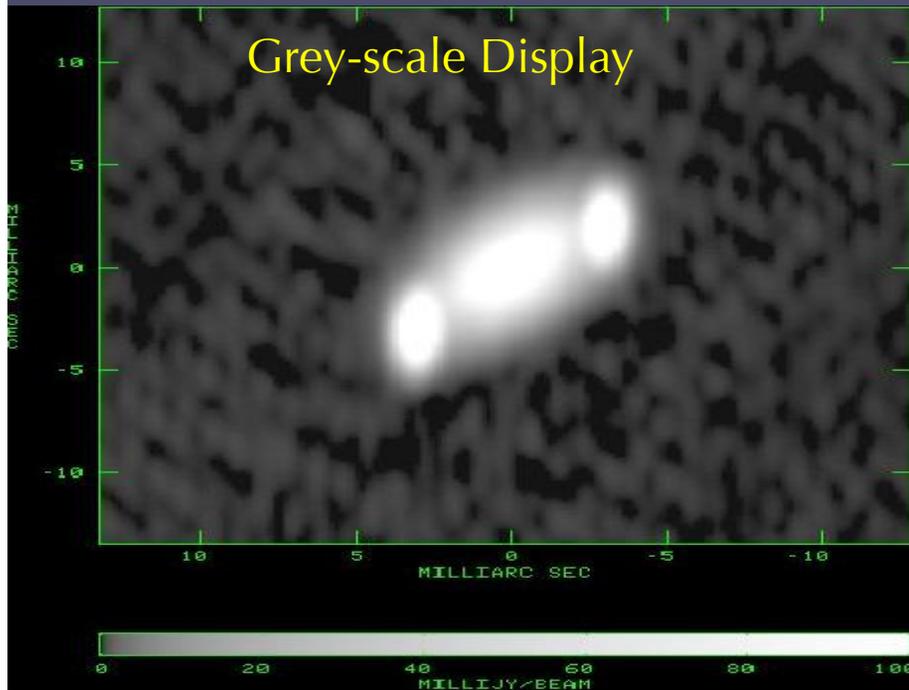


These plots are easy to reproduce and print

Contour plots give good representation of faint emission.

Profile plots give a good representation of the bright emission.

IMAGE DISPLAYS (3)



TV-based displays are most useful and interactive:

Grey-scale shows faint structure, but not good for high dynamic range and somewhat unbiased view of source
Color displays more flexible; e.g. pseudo contours

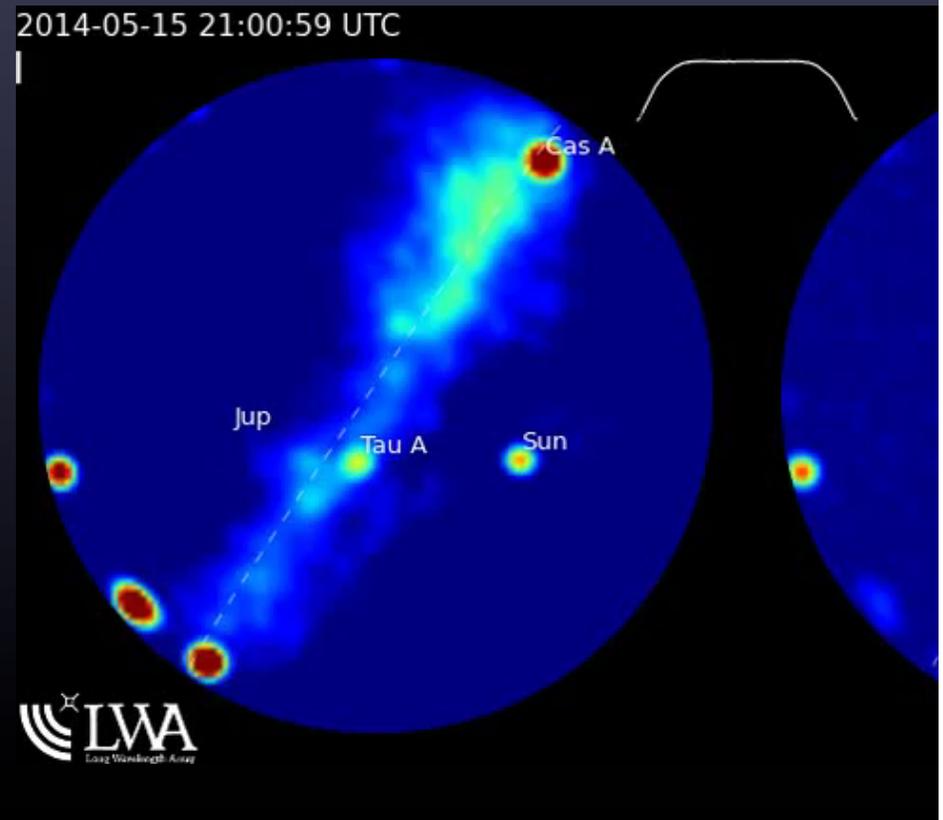
Movies and Radio Frequency Interference (RFI)

Great pressure from wireless devices (especially smartphones)
Auction in 2015 of 40 MHz raised 45 billion \$\$\$!
Auction in 2017 of 70 MHz within 580 – 700 MHz raised 19 billion
Next up? maybe 5G networks at ~60 GHz
Dynamic allocation/Shared use of spectrum
Passive use is still useful!

Likely changes:

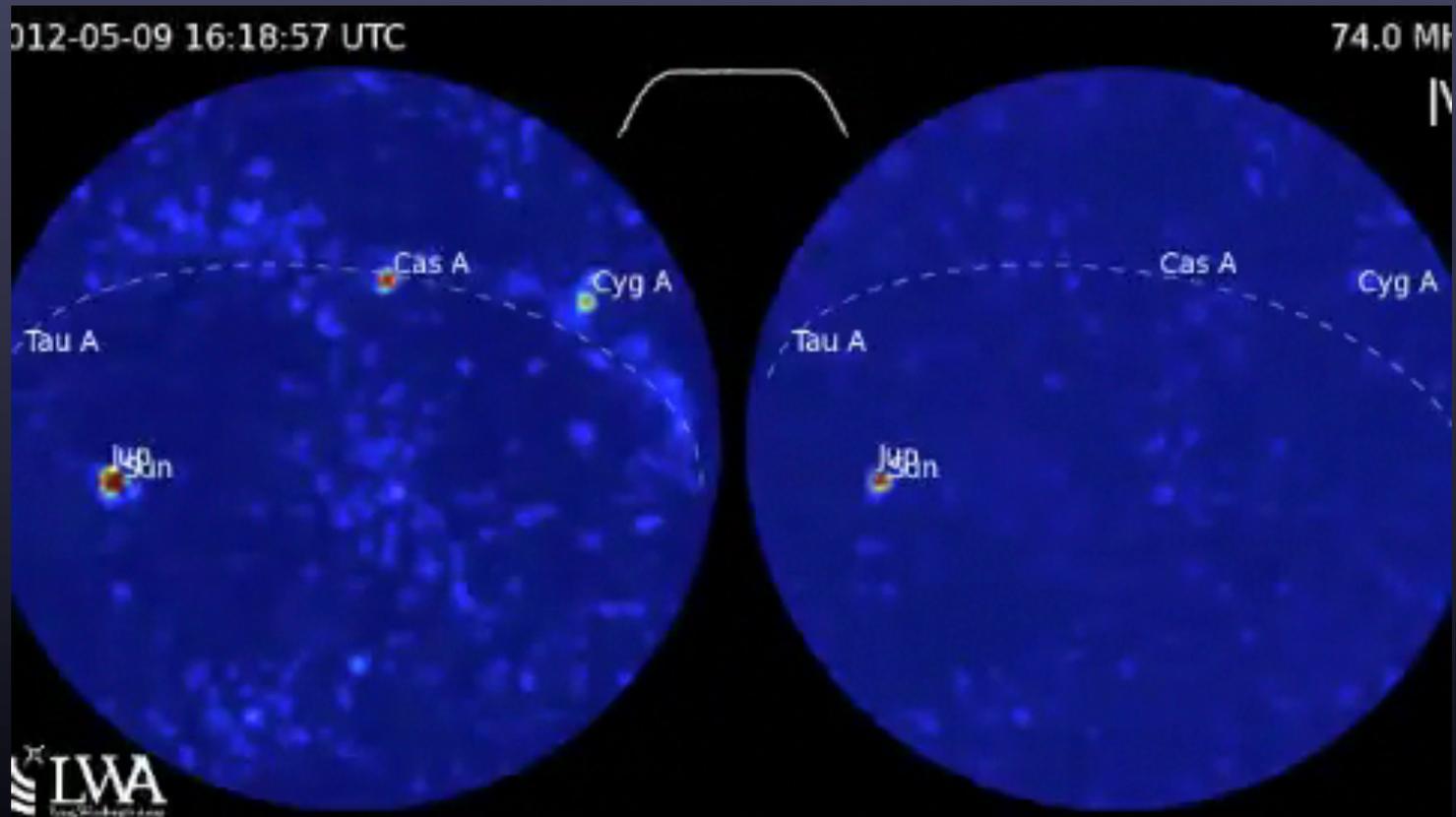
608 – 614 MHz mobile comm.
1616 – 1627 MHz mobile comm.
1755 – 1850 MHz mobile comm.
2155 – 2200 MHz mobile comm
4400 – 4940 MHz AMT (planes)
5150 – 5250 MHz unlicensed devices
5725 – 5850 MHz unlicensed devices
37000 – 38600 satellite/mobile com
57240 – 59400 WiGig
71000 – 76000 automobile radar
81000 – 86000 automobile radar

LWA1 with ~250 antenna stands



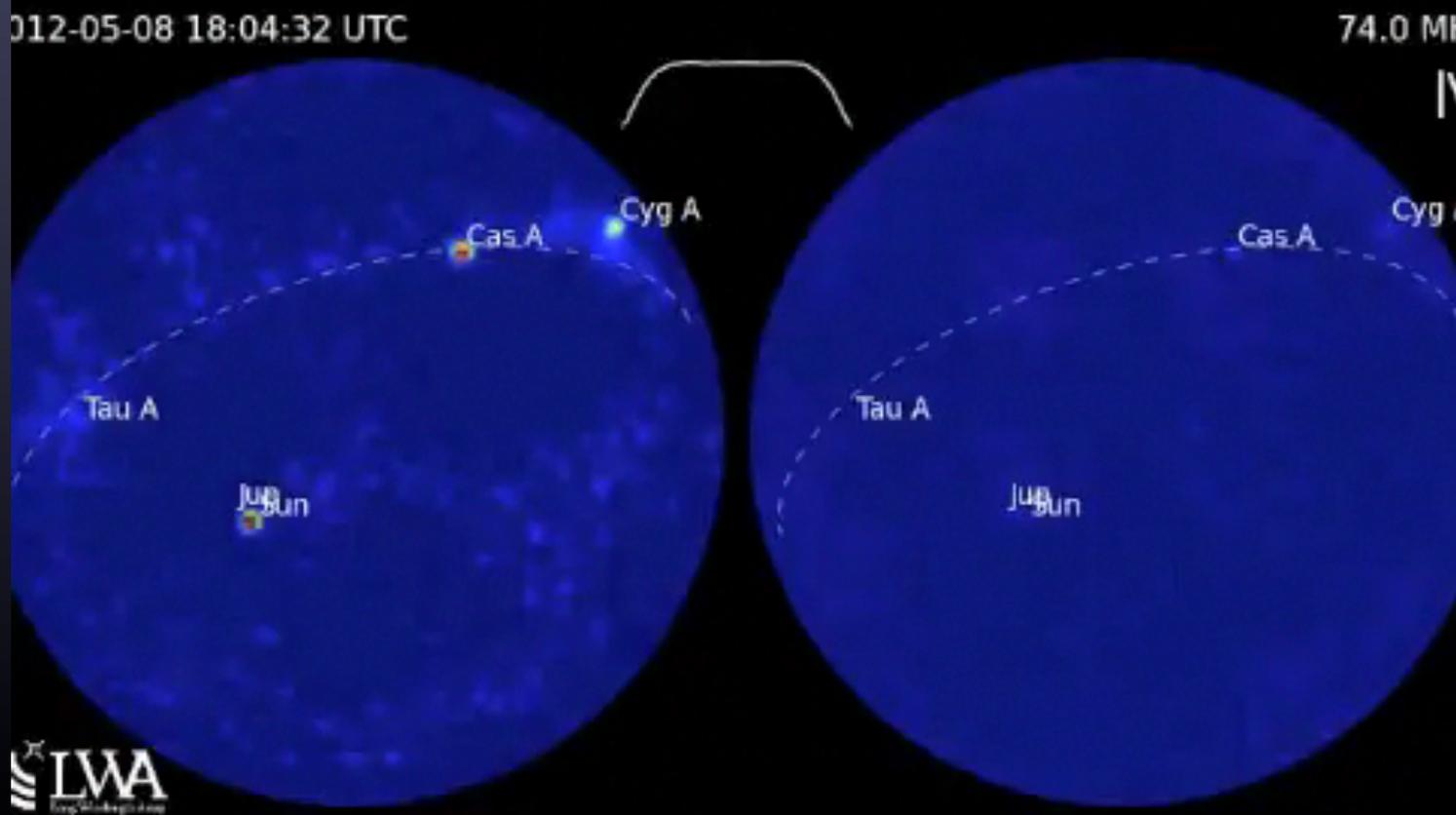
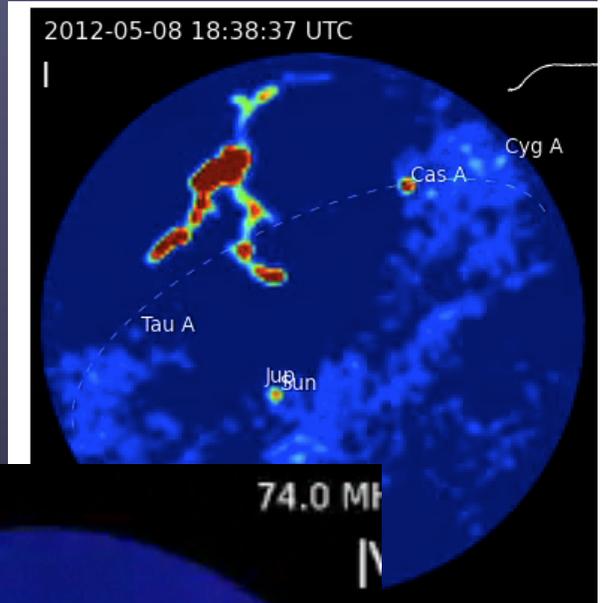
Movies and Solar Interference

Watch out for the Active Sun



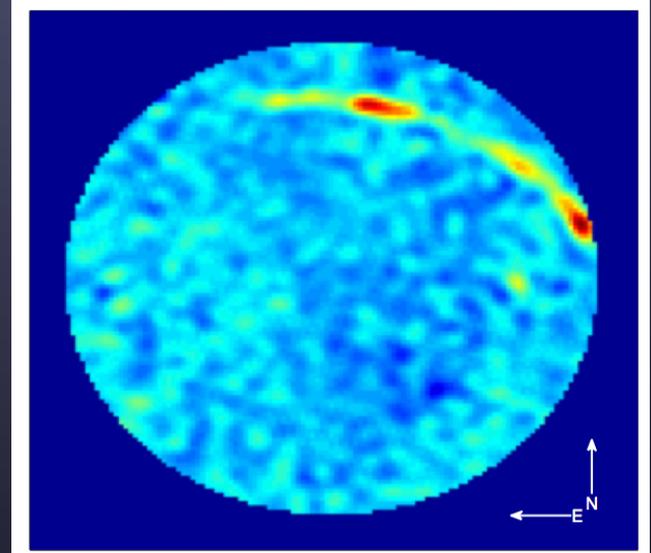
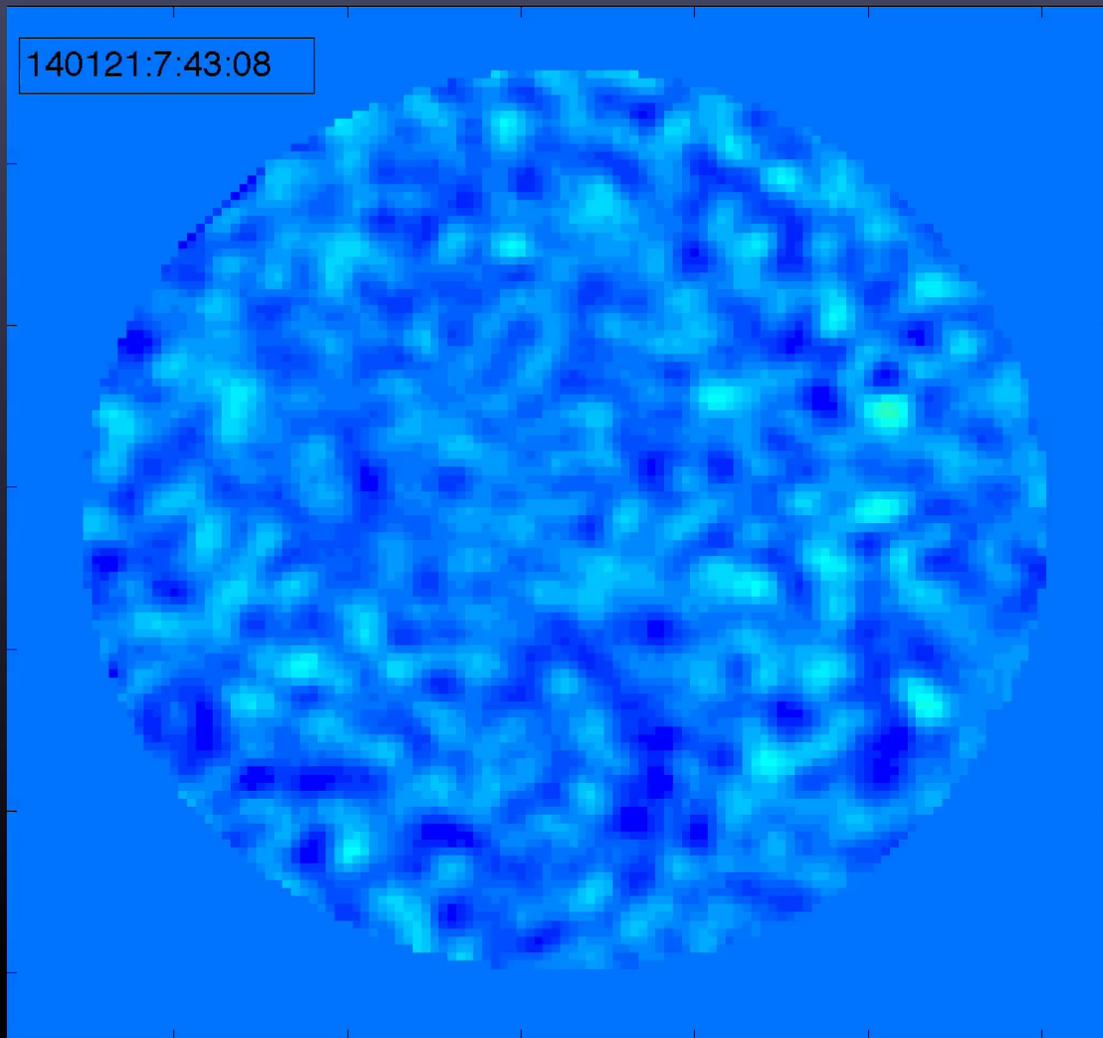
Lightning

Thunderstorm season on the Plains ...



Fireballs

Discovery of broad-band emission from meteors



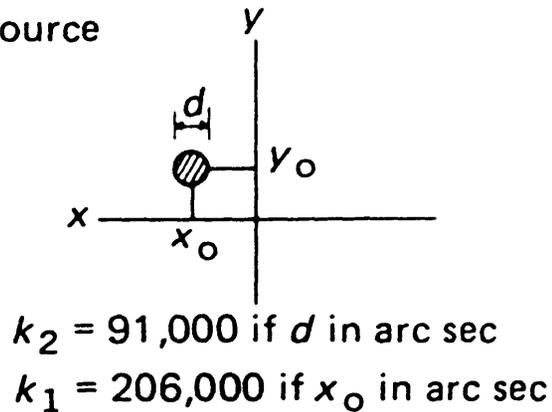
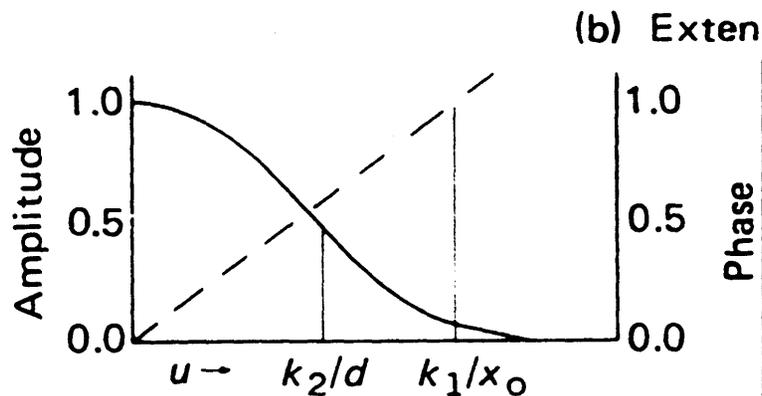
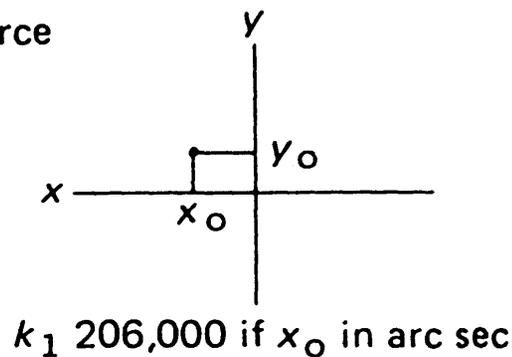
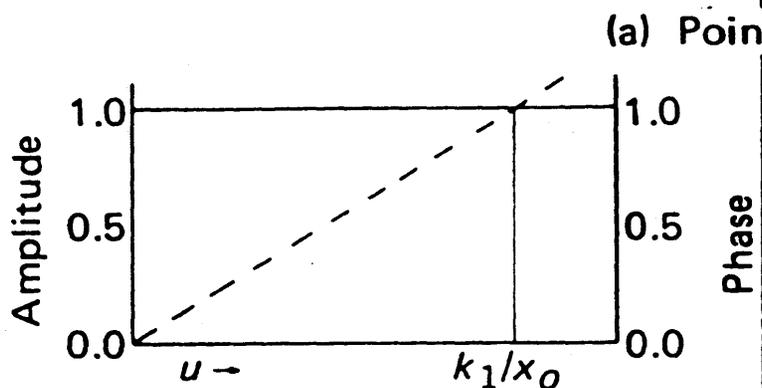
These happen about 1/week
Peak is ~ 3000 Jy at 40 MHz

Obenberger et al. 2014

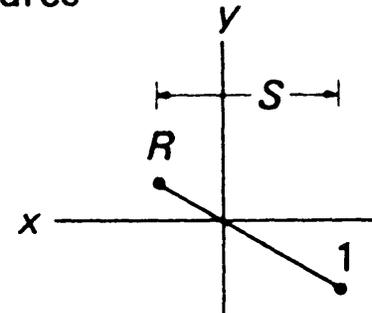
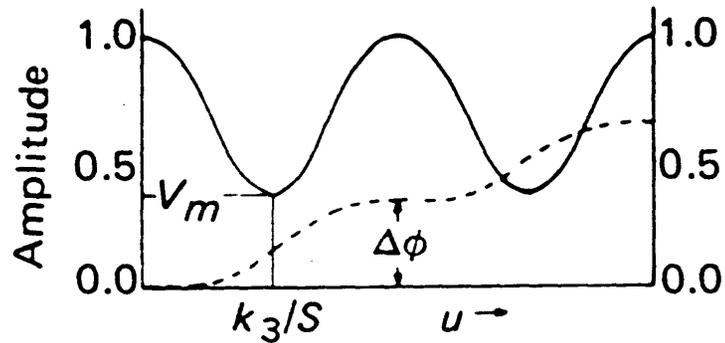
Visibility <-> FT <-> Brightness

Visibility function

Brightness distribution



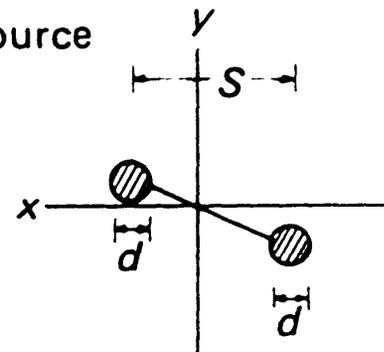
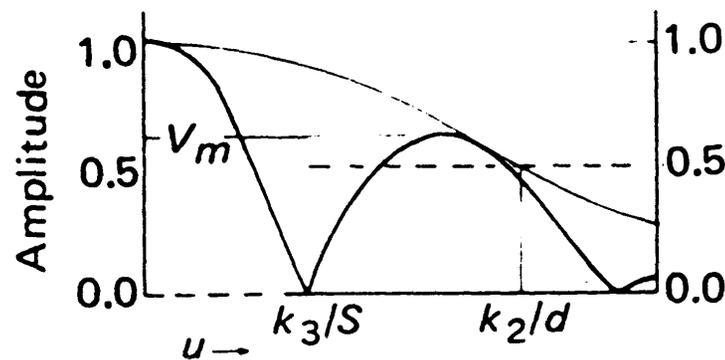
(c) Point double source



$$k_3 = 103,000 \text{ if } S \text{ in arc sec}$$

$$V_m = \frac{R - 1}{R + 1} ; \Delta\phi = \frac{1}{1 + R}$$

(d) Extended double source



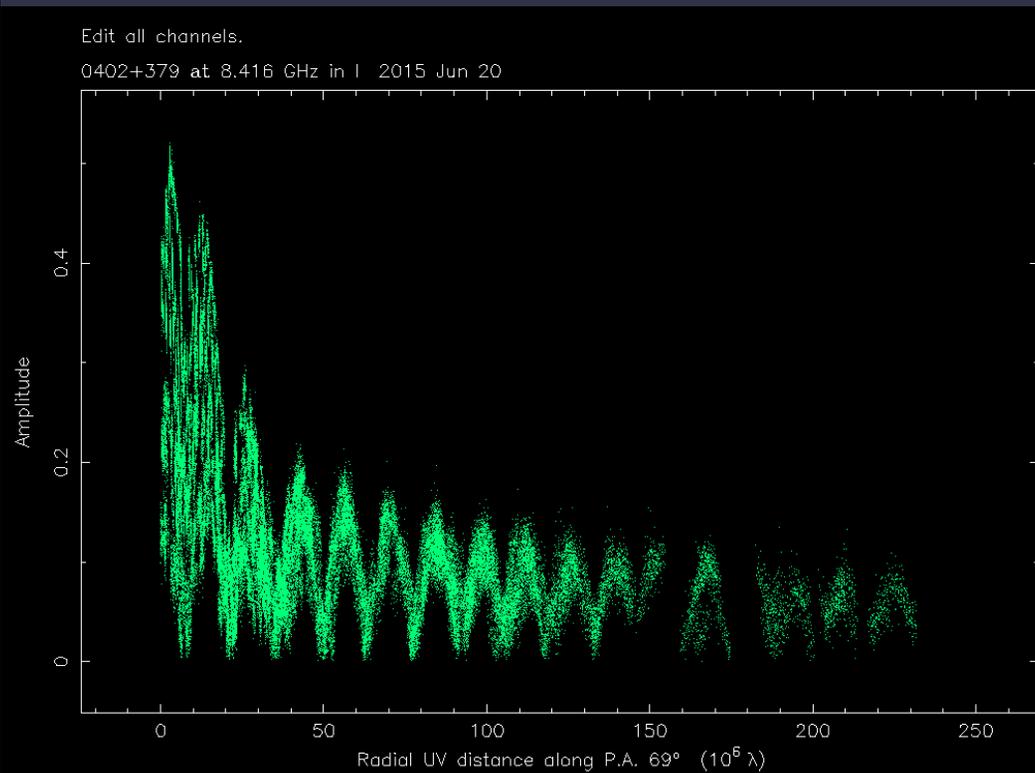
$$k_3 = 103,000 \text{ if } S \text{ in arc sec}$$

$$k_2 = 91,000 \text{ if } d \text{ in arc sec}$$

$$V_m \approx \exp \left\{ -3.57 \left(\frac{d}{S} \right)^2 \right\}$$

Trial model

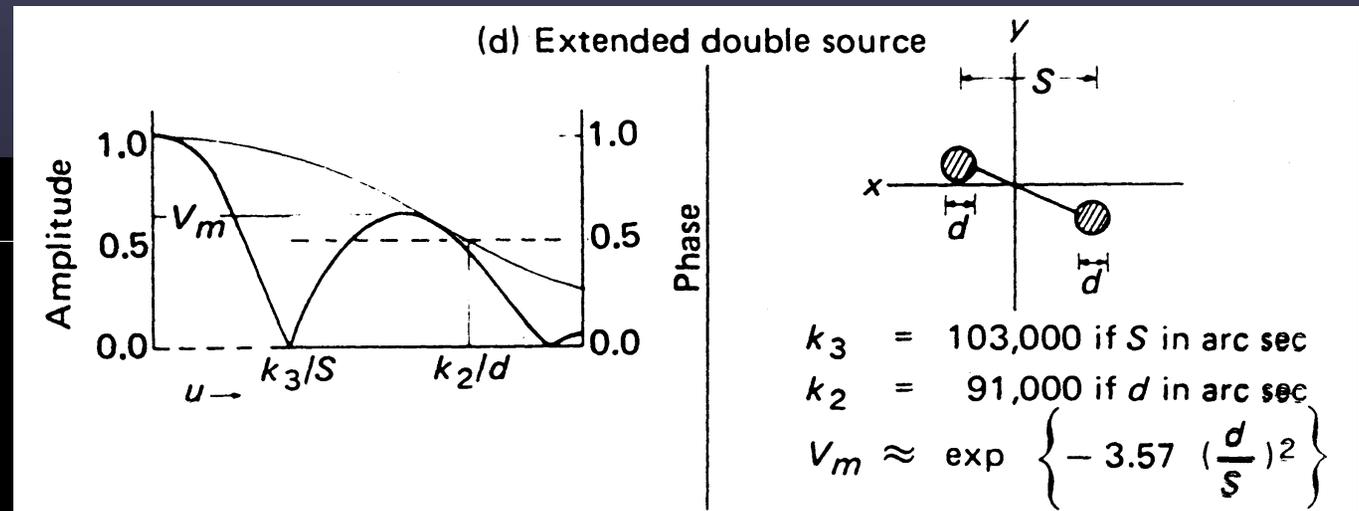
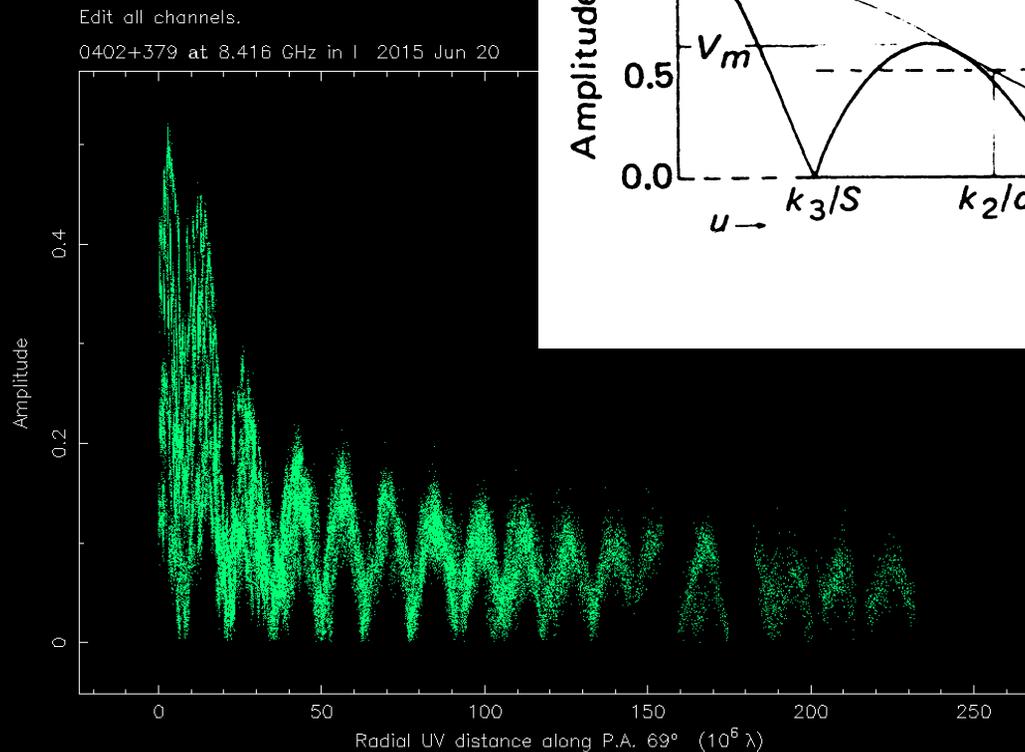
By inspection, we can derive a simple model:
Your GUESS?



Trial model

By inspection, we can derive a simple model:

Two roughly equal components, each 0.1 Jy, separated by about 15 milliarcsec in p.a. 69° , each about 0.5 milliarcsec in diameter (gaussian FWHM)



To be refined later...

DATA DISPLAYS(1)

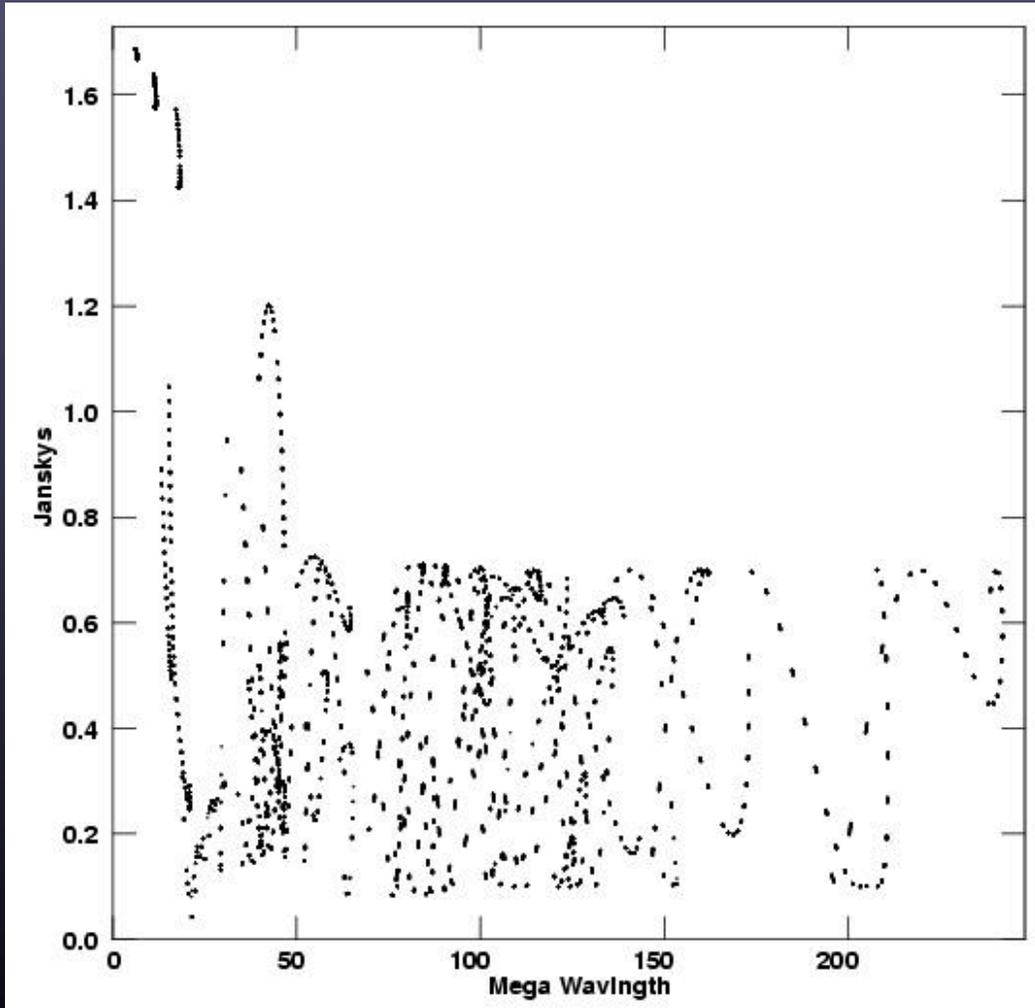
List of (u,v) Data

Source=	J0121+11	Freq=	8.434858511	Sort=	TB	1	RR			
Vis #	IAT	Ant	Su Fq	U(klam)	V(klam)	W(klam)	Amp	Phas	Wt	
2191	0/22:35:08.22	5- 6	1 0	94220	23776	100371	0.614	-16	1.0000	
3971	0/22:43:43.34	5- 6	1 0	97659	24517	96844	0.508	-13	1.0000	
6431	0/23:07:05.15	5- 6	1 0	106307	26661	86632	0.154	17	1.0000	
6611	0/23:07:14.98	5- 6	1 0	106364	26677	86557	0.152	17	1.0000	
6791	0/23:07:24.81	5- 6	1 0	106421	26692	86483	0.150	18	1.0000	
6971	0/23:07:34.64	5- 6	1 0	106477	26708	86408	0.148	19	1.0000	
7151	0/23:07:44.47	5- 6	1 0	106534	26724	86333	0.146	19	1.0000	
7331	0/23:07:54.30	5- 6	1 0	106591	26739	86259	0.144	20	1.0000	
7511	0/23:15:06.84	5- 6	1 0	109027	27438	82930	0.101	74	1.0000	
7691	0/23:15:16.67	5- 6	1 0	109081	27454	82854	0.101	75	1.0000	
7871	0/23:15:26.50	5- 6	1 0	109135	27470	82777	0.102	77	1.0000	
8051	0/23:15:36.33	5- 6	1 0	109189	27486	82701	0.102	78	1.0000	
8231	0/23:15:46.16	5- 6	1 0	109243	27502	82624	0.103	79	1.0000	
8411	0/23:15:55.99	5- 6	1 0	109297	27518	82547	0.104	81	1.0000	
9701	0/23:31:02.36	5- 6	1 0	114020	29035	75322	0.260	134	1.0000	
9791	0/23:31:06.29	5- 6	1 0	114040	29042	75290	0.261	134	1.0000	
10301	0/23:31:29.88	5- 6	1 0	114156	29082	75098	0.266	134	1.0000	
10861	0/23:39:02.08	5- 6	1 0	116320	29863	71379	0.348	139	1.0000	
10951	0/23:39:06.01	5- 6	1 0	116339	29870	71346	0.348	139	1.0000	
11171	0/23:39:15.84	5- 6	1 0	116384	29887	71264	0.350	139	1.0000	

Old School, but sometimes worth-while: e.g. , can search on e.g. Amp > 1.0, or large weight. Often need precise times in order to flag the data appropriately.

DATA DISPLAYS(2)

Jy



Mega Wavelength

Visibility Amplitude versus
Projected (u,v) spacing

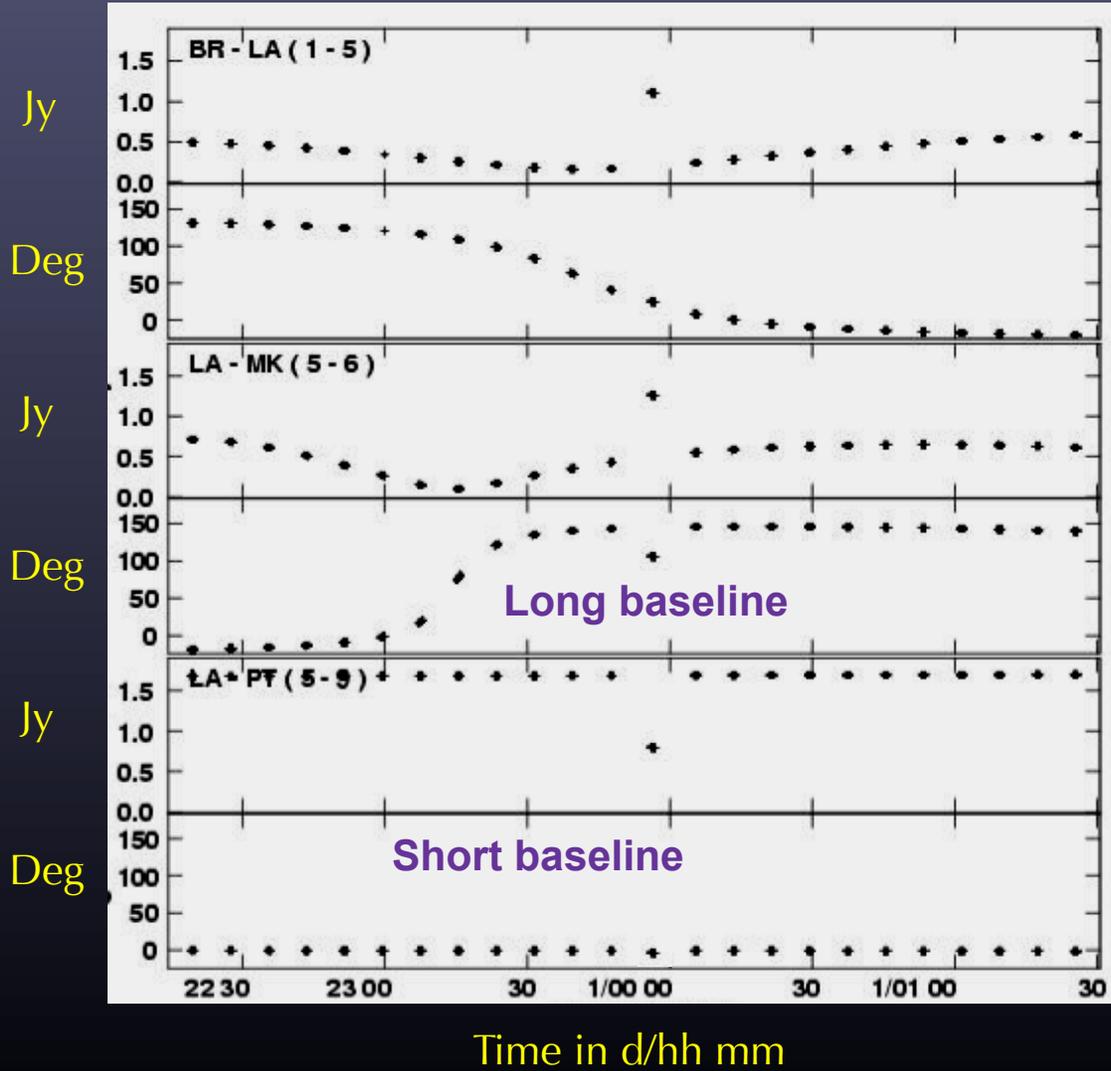
General trend of data.

Useful for relatively strong
sources.

Triple source model. Large
component causes rise at
short spacings.

Oscillations at longer spacings
suggest close double.

DATA DISPLAYS(3)



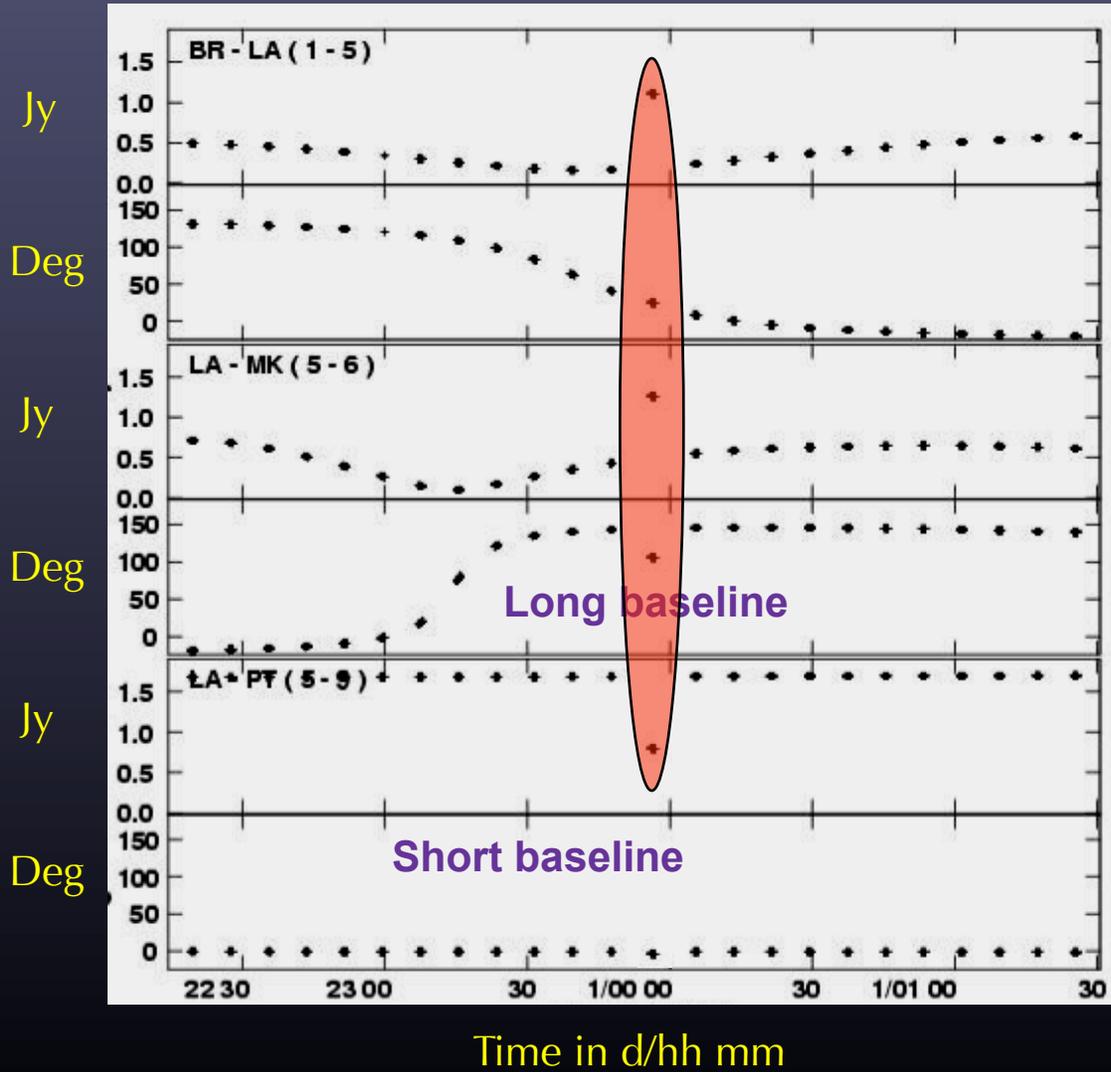
Visibility amplitude and phase versus time for various baselines

Good for determining the continuity of the data

Should be relatively smooth with time

Outliers are obvious.

DATA DISPLAYS(3)



Visibility amplitude and phase versus time for various baselines

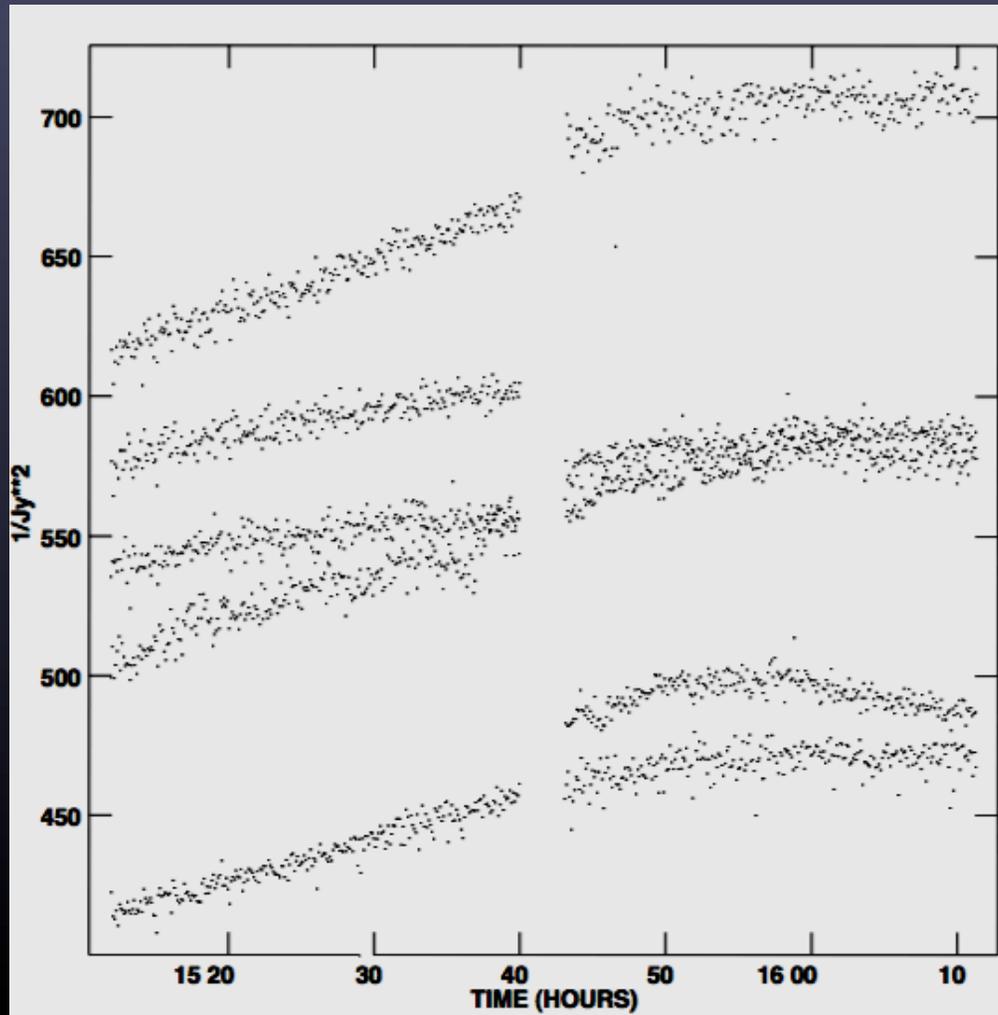
Good for determining the continuity of the data

Should be relatively smooth with time

Outliers are obvious.

DATA DISPLAYS(4)

Weights of antennas 4 with 5,6,7,8,9



All (u,v) data points have a weight. The weight depends on the antenna sensitivity, measured during the observations

The amplitude calibration values also modify the weights.

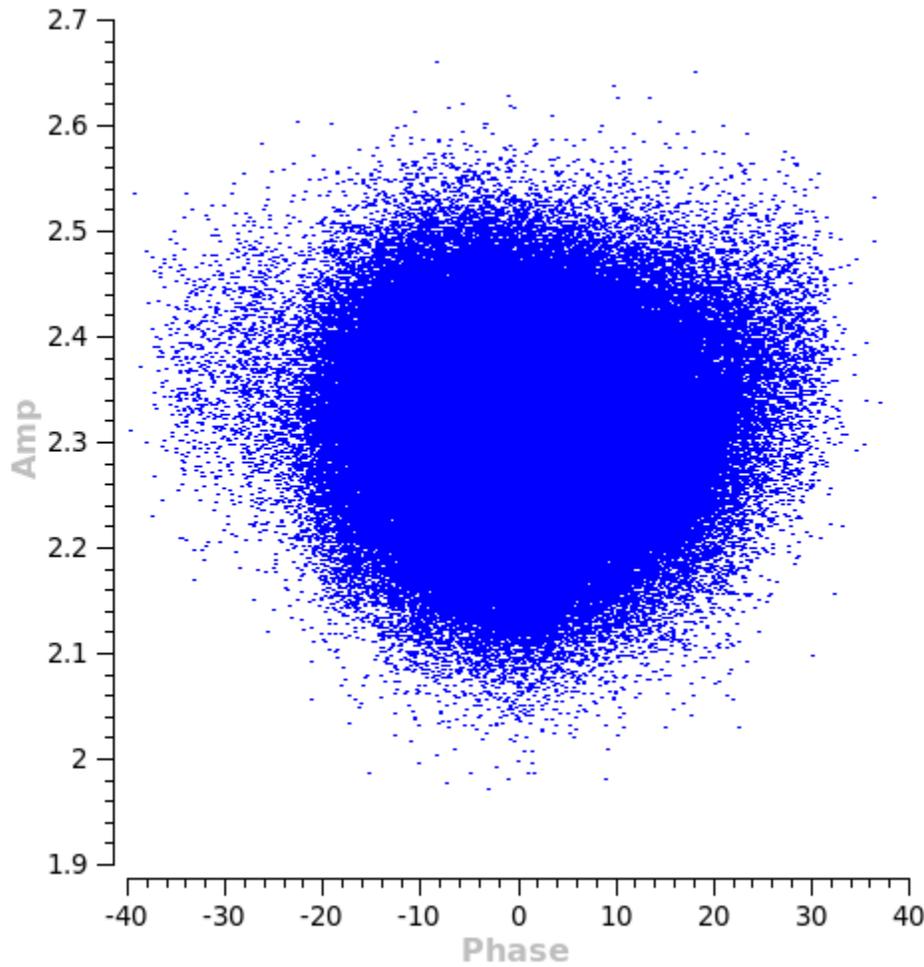
Occasionally the weight of the points become very large, often caused by subtle software bugs.

A large discrepant weight causes the same image artifacts as a large discrepant visibility value.

Please check weights to make sure they are reasonable.

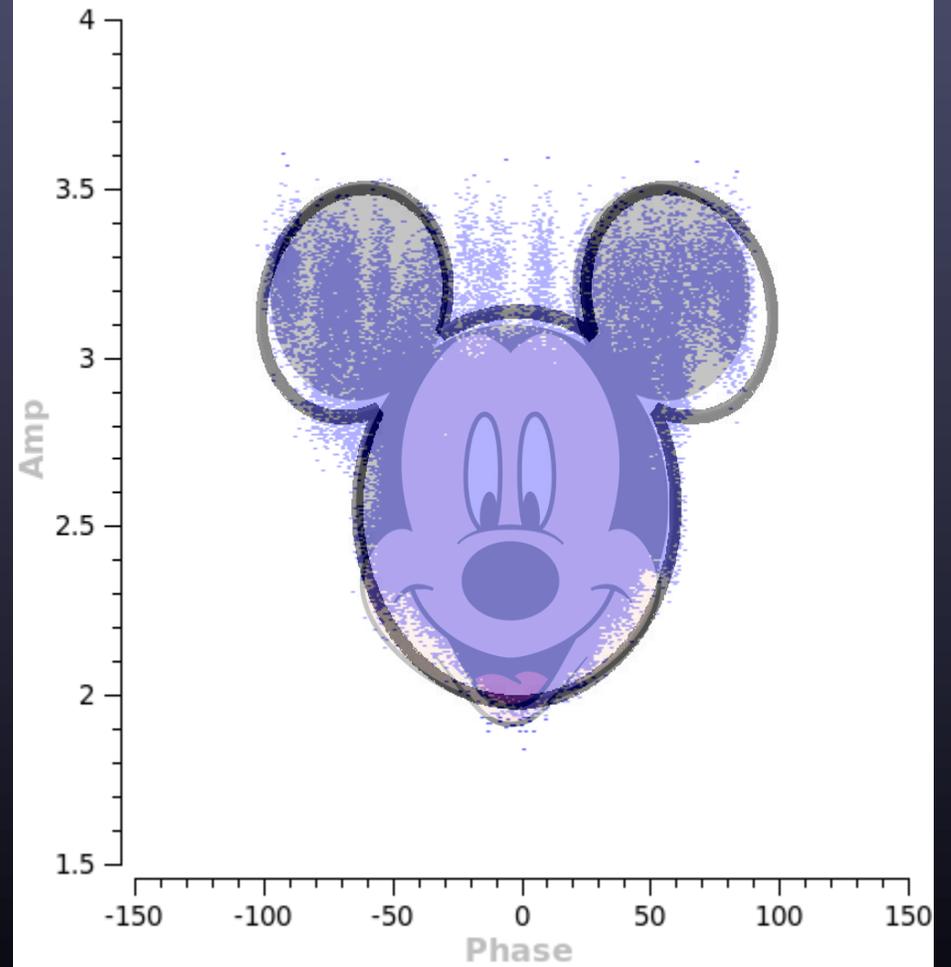
DATA DISPLAYS(5) – Amplitude vs Phase

Amp vs. Phase



Good

Amp vs. Phase Field: J0136+4751



Bad

IMAGE PLANE OR DATA (U,V) PLANE INSPECTION?

Errors obey Fourier transform relationship

Narrow feature in (u,v) plane \leftrightarrow wide feature in image plane

Wide feature in (u,v) plane \leftrightarrow narrow feature in image plane

Note: often easier to spot narrow features

Data (u,v) amplitude errors \leftrightarrow symmetric image features

Data (u,v) phase errors \leftrightarrow asymmetric image features

An obvious defect may be hardly visible in the transformed plane

A small, almost invisible defect may become very obvious in the transformed plane

Noise bumps can have sidelobes!

FINDING ERRORS

---Obvious outlier data (u,v) points:

100 bad points in 100,000 data points gives an 0.1% image error
(unless the bad data points are 1 million Jy)

LOOK at DATA to find gross problem (you'd be hard pressed to find it in the image plane other than a slight increase in noise)

---Persistent small data errors:

e.g. a 5% antenna gain calibration error is difficult to see in (u,v) data (not an obvious outlier), but will produce a 1% effect in image with specific characteristics (more later).

USE IMAGE to discover problem

---Non-Data Problems:

Data ok, but algorithms chosen aren't up to the task.

ERROR RECOGNITION IN THE (u,v) PLANE

Editing obvious errors in the (u,v) plane

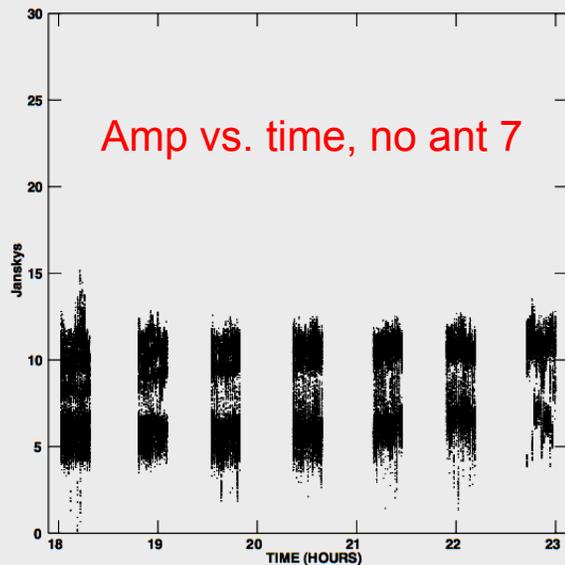
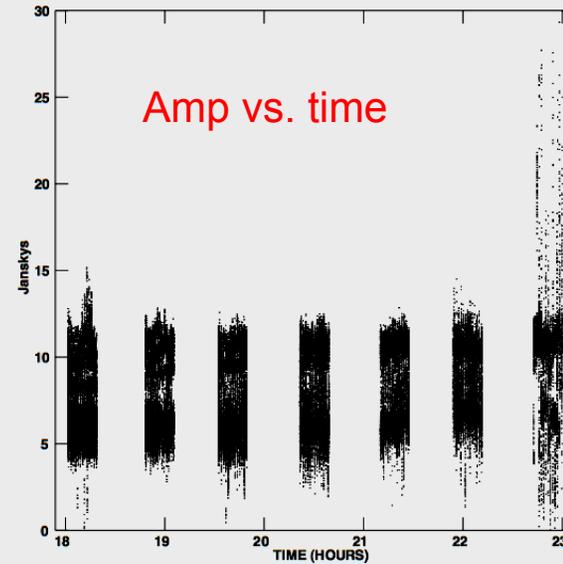
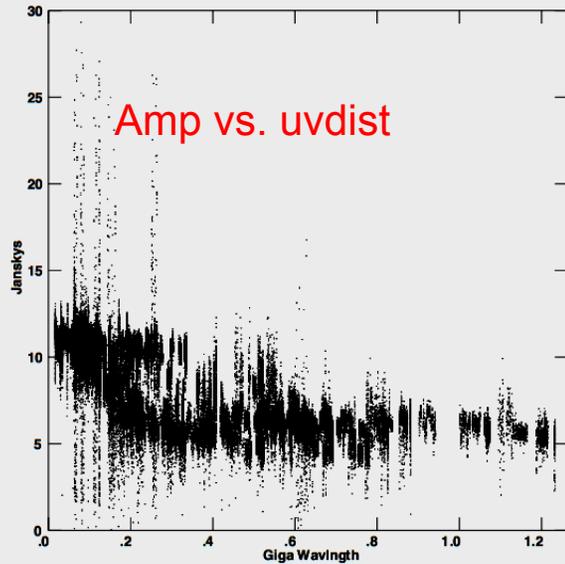
---Mostly consistency checks assume that the visibility cannot change much over a small change in (u,v) spacing

---Also, look at gains and phases from calibration processes. These values should be relatively stable.

See Summer school lecture notes in 2002 by Myers

See ASP Vol 180, Ekers, Lecture 15, p321

VISIBILITY AMPLITUDE PLOTS



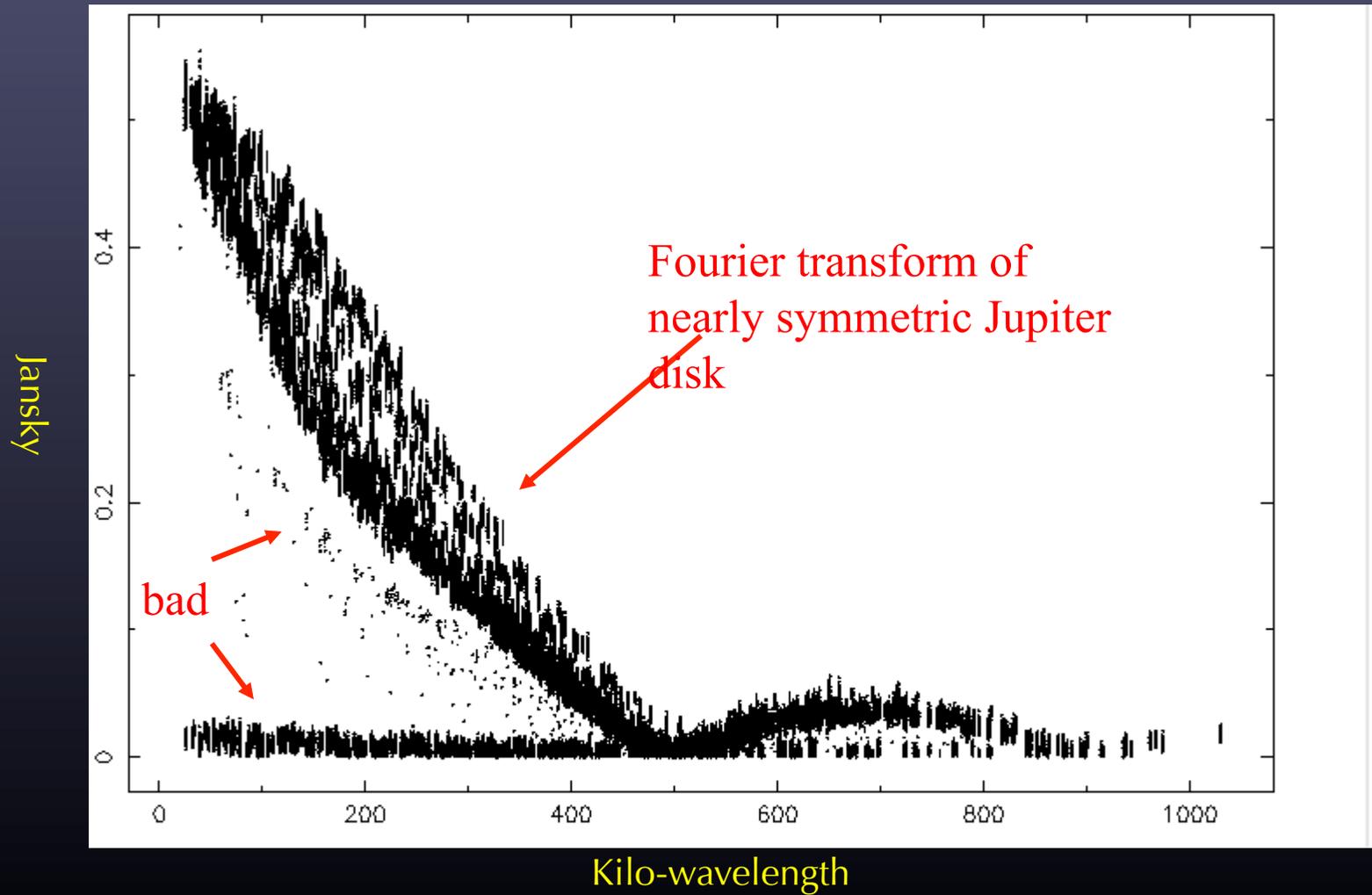
Amp vs. uvdist shows outliers

Amp vs. time shows outliers in last scan

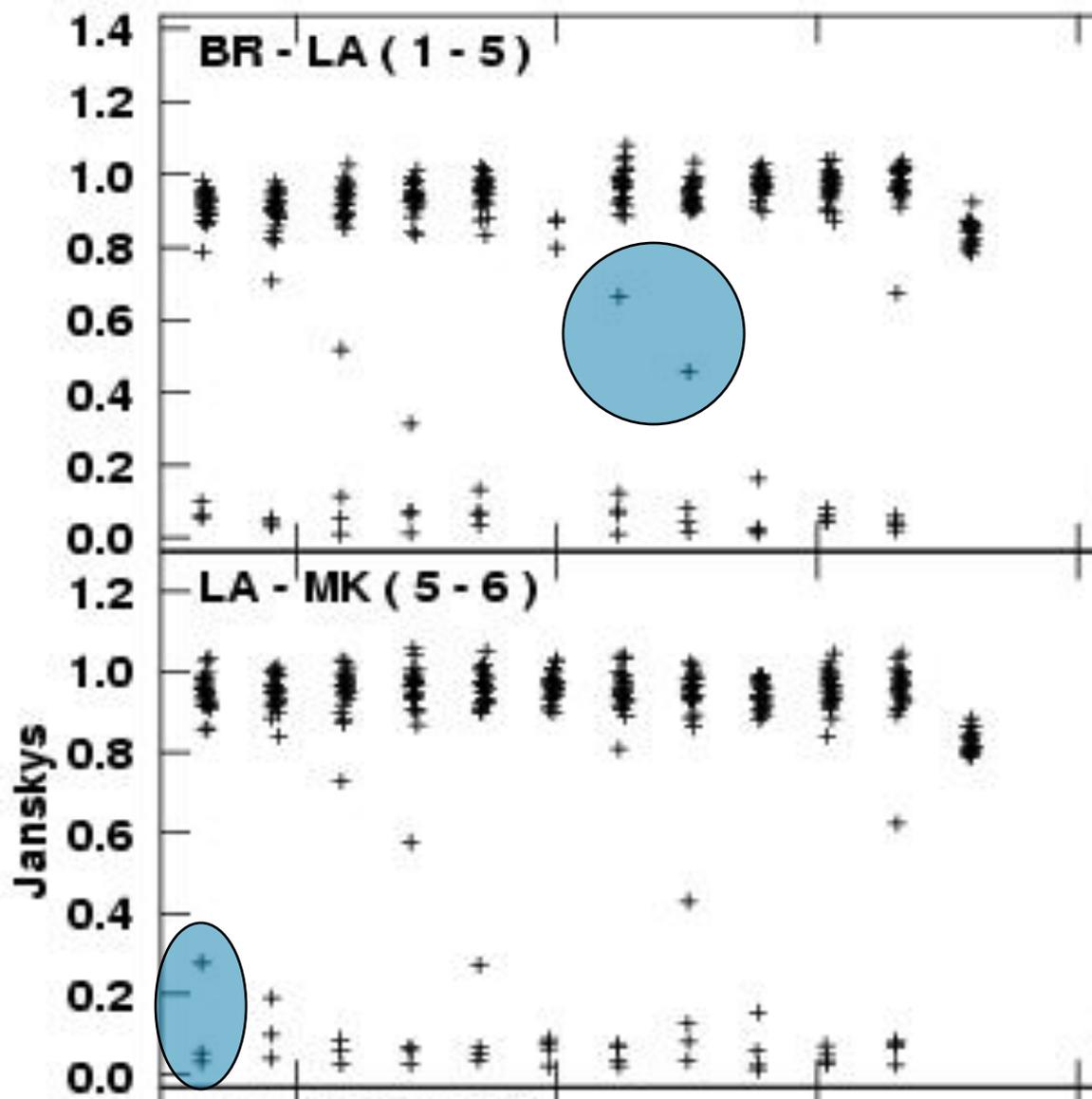
Amp vs. time without ant 7 shows good data

(3C279 VLBA data at 43 GHz)

Example Edit – plotms



Drop-outs at Scan Beginnings



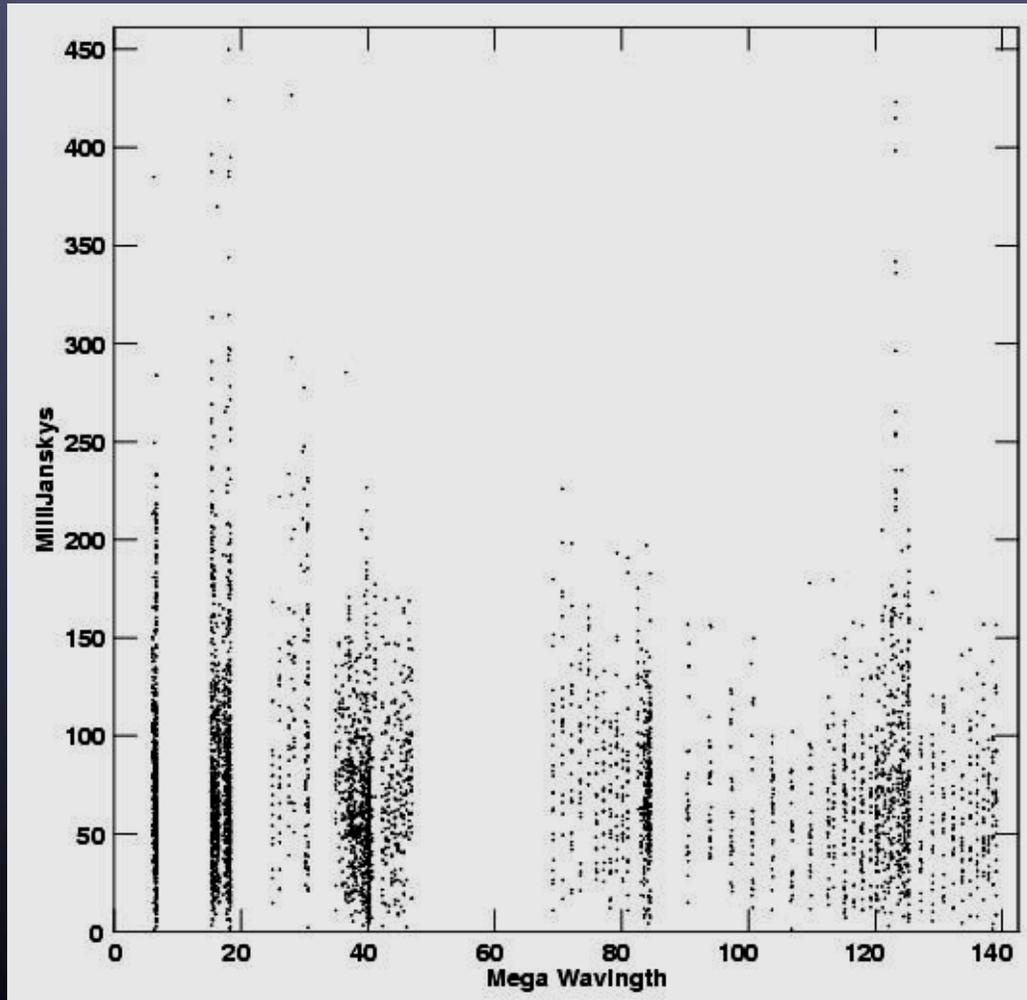
Often the first few points of a scan are low. E.g. antenna not on source.

Software can remove these points (aips, casa 'quack')

Flag extension:

Should flag all sources in the same manner even though you cannot see dropout for weak sources

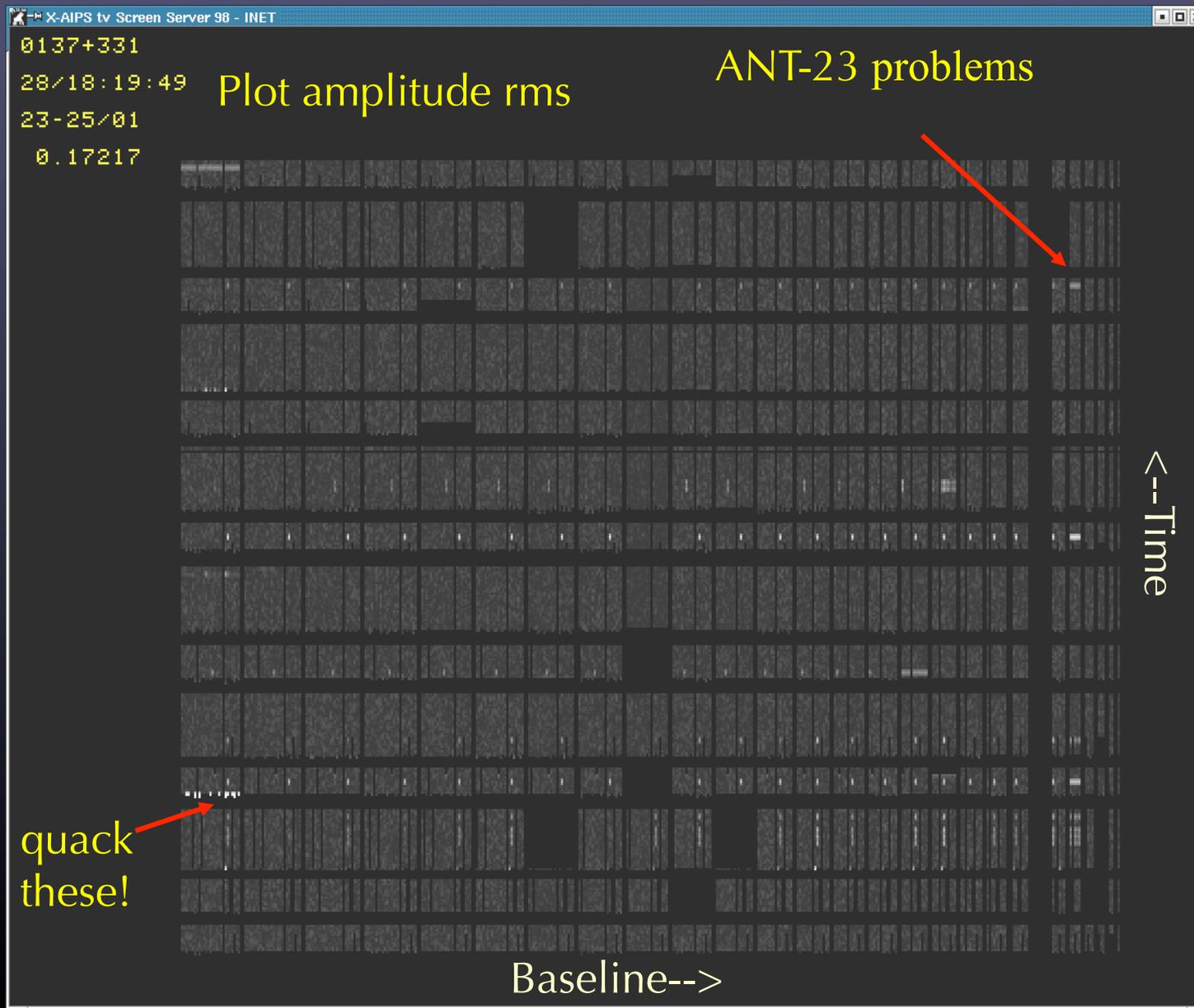
Editing Noise-dominated Sources



No source structure
information is detected.
Noise dominated.

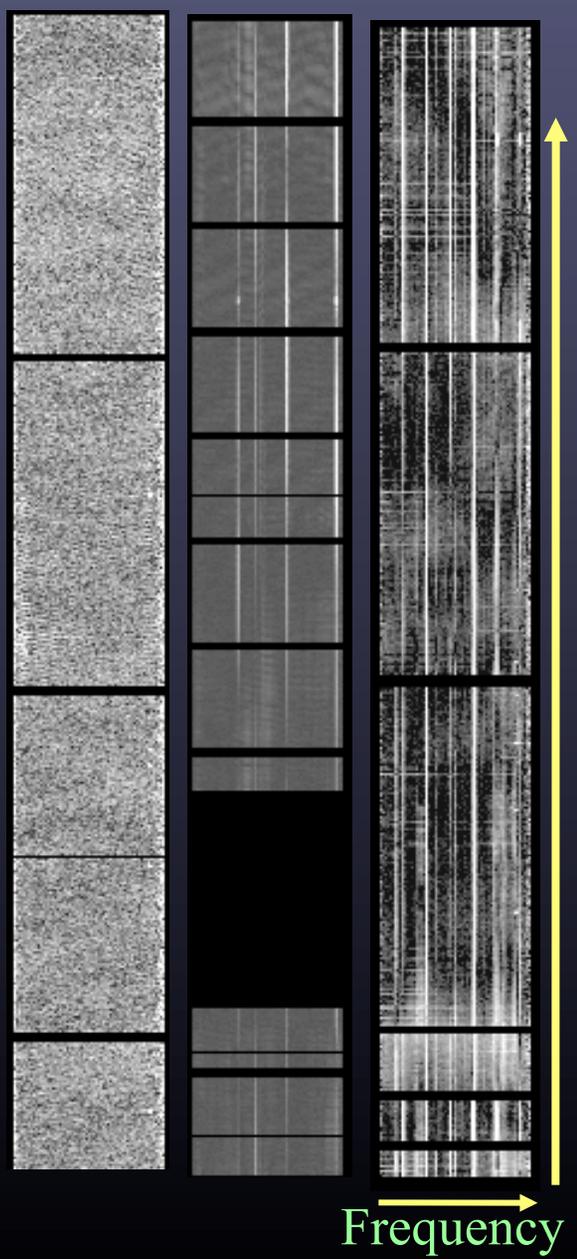
All you can do is quack and
remove outlier points
above $\sim 3\sigma$ (0.3 Jy).
Precise level not important
as long as large outliers are
removed.

USING TVFLG (VIEWER) DISPLAY on a source



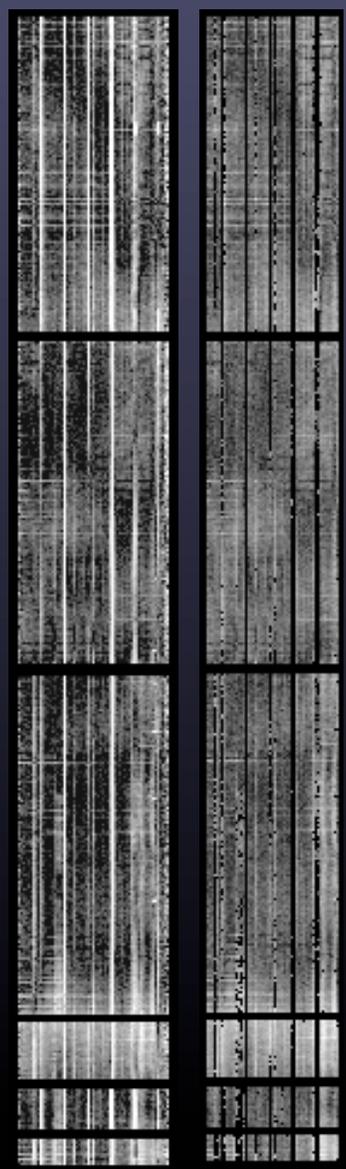
RFI Excision

35 km 12 km 3 km baseline



AIPS: SPFLG

before after



RFI environment worse on short baselines

Several 'types': narrow band, wandering, wideband, ...

Wideband interference hard for automated routines

AIPS tasks FLGIT, RFLAG, SPFLG and CASA flagdata, mode='rfi'

AOFlagger by Offringa

Automation is crucial for WIDAR (wide band, lots of data)

ERROR RECOGNITION IN THE IMAGE PLANE

Some Questions to ask:

Noise properties of image:

- Is the rms noise about expected from integration time + bandwidth?
- Is the rms noise much larger near bright sources?
- Are there non-random noise components (faint waves and ripples)?

Funny looking Structure:

- Non-physical features; stripes, rings, symmetric or anti-symmetric
- Negative features well-below 4σ noise
- Does the image have characteristics that look like the dirty beam?

Image-making parameters:

- Is the image big enough to cover all significant emission?
- Is cell size too large or too small? ~ 4 points per beam okay
- Is the resolution too high to detect most of the emission?

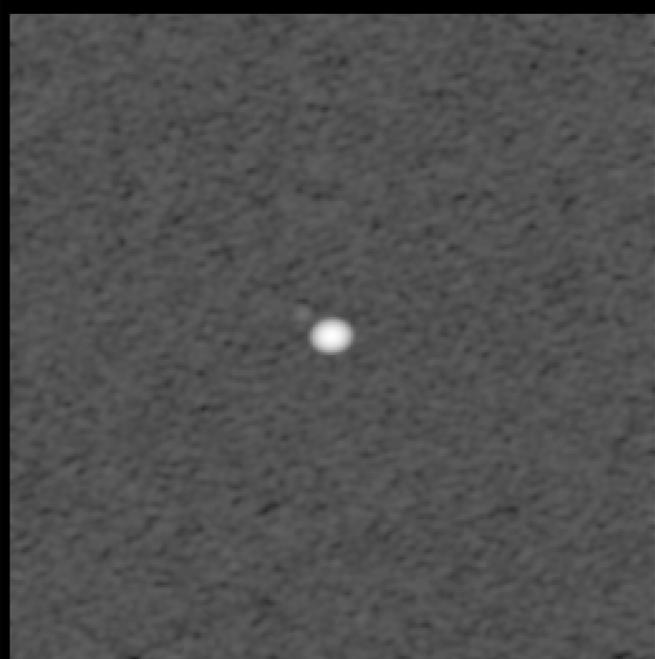
EXAMPLE 1

Data bad over a short period of time

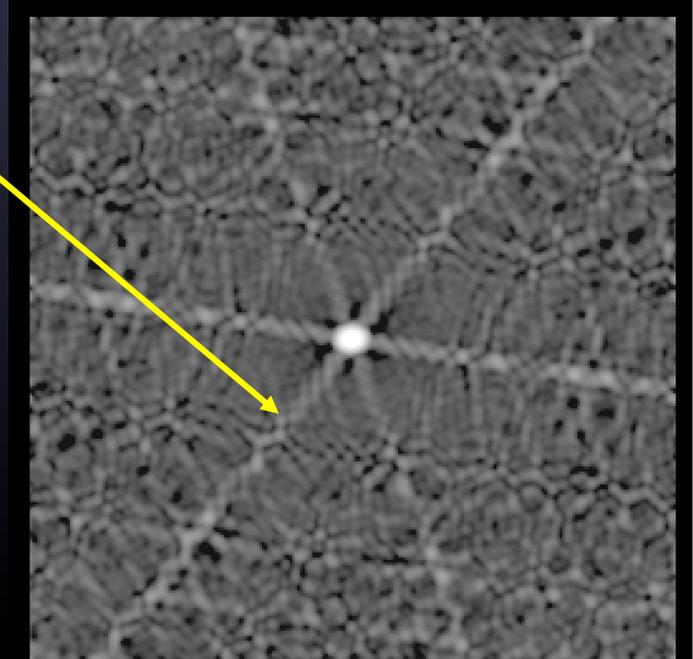
Results for a point source using VLA. 13 x 5min observation over 10 hr.
Images shown after editing, calibration and deconvolution.

no errors:
max 3.24 Jy
rms 0.11 mJy

10% amp error for all
antennas for 1 time period
rms 2.0 mJy



6-fold symmetric
pattern due to
VLA "Y".
Image has
properties of dirty
beam.

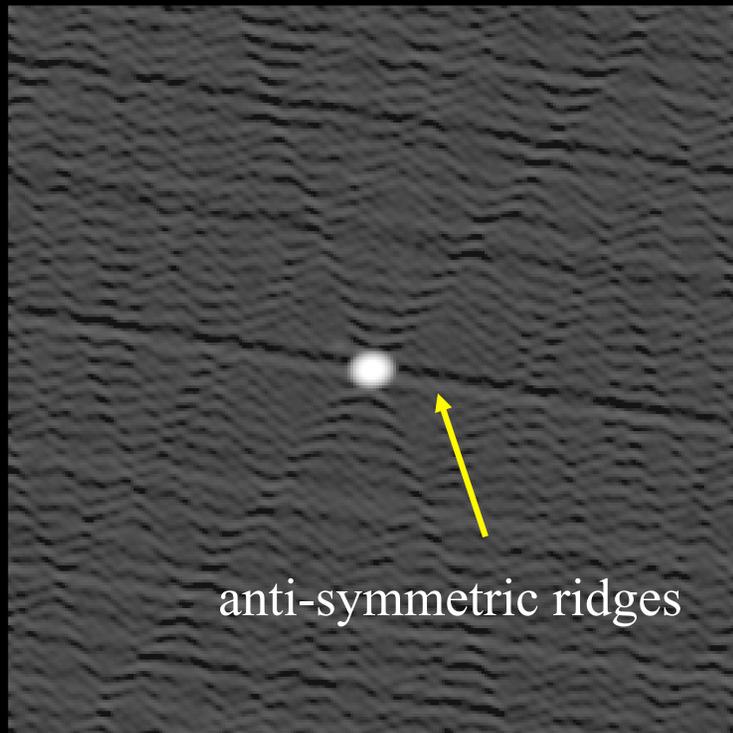


EXAMPLE 2

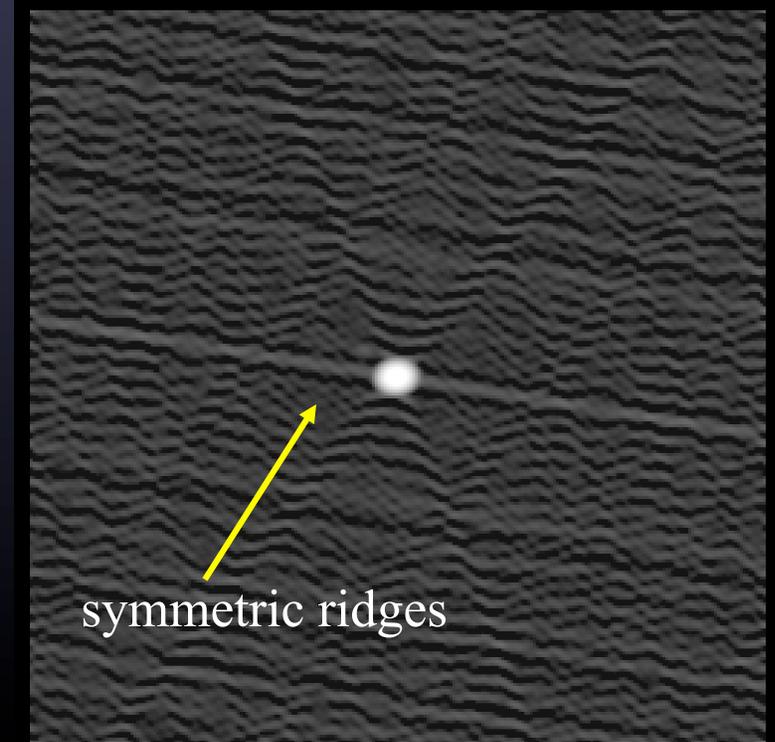
Short burst of bad data

Typical effect from one bad antenna

10 deg phase error for
one antenna at one time
rms 0.49 mJy



20% amplitude error for
one antenna at one time
rms 0.56 mJy



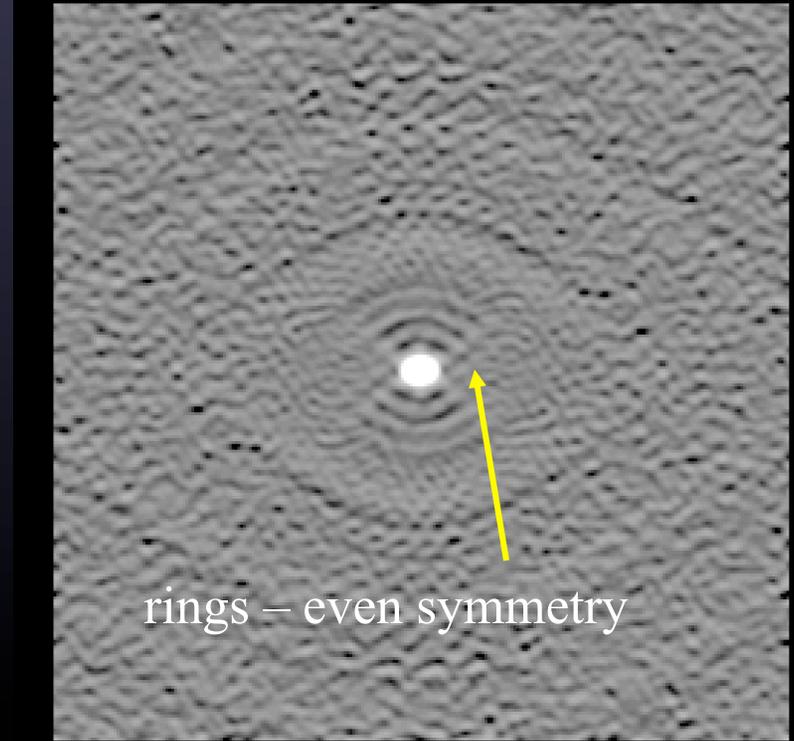
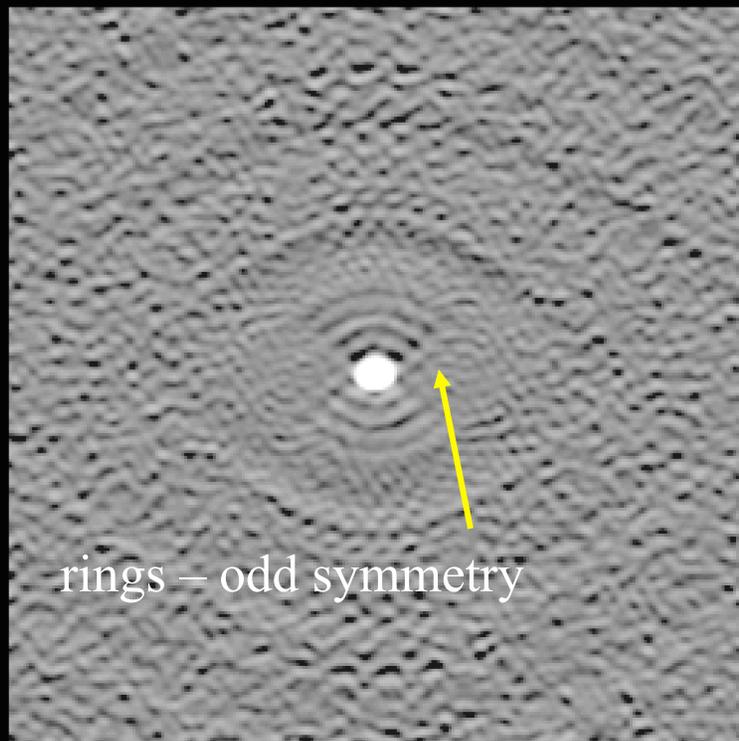
EXAMPLE 3

Persistent errors over most of observations

NOTE: 10 deg phase error to 20% amplitude error
cause similar sized artifacts

10 deg phase error for
one antenna all times
rms 2.0 mJy

20% amp error for one
antenna all times
rms 2.3 mJy



EXAMPLE 4

Spurious Correlator Offset Signals

Occasionally correlators produce ghost signals or cross talk signals
Occurred during change-over from VLA to EVLA system

Symptom: Garbage near phase center, dribbling out into image

Image with correlator offsets

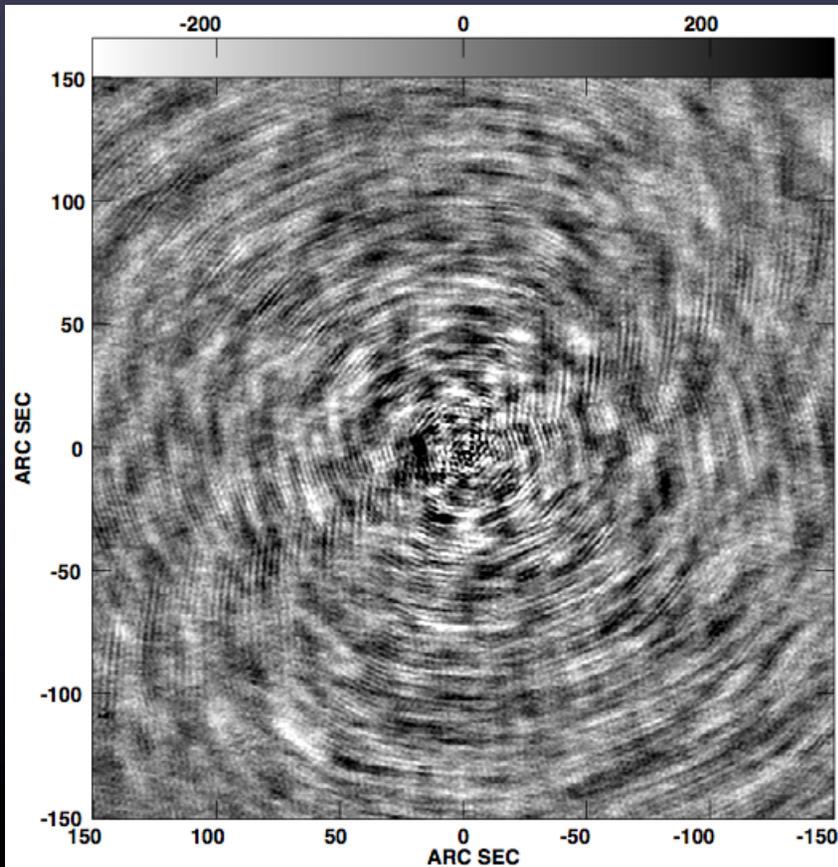
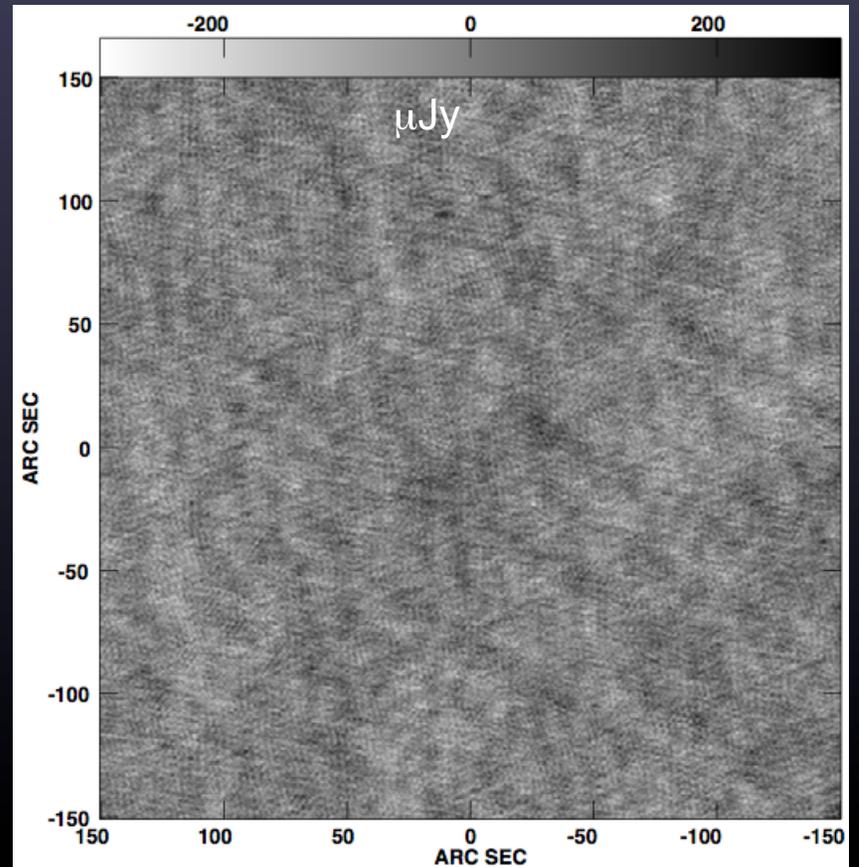


Image after correction of offsets



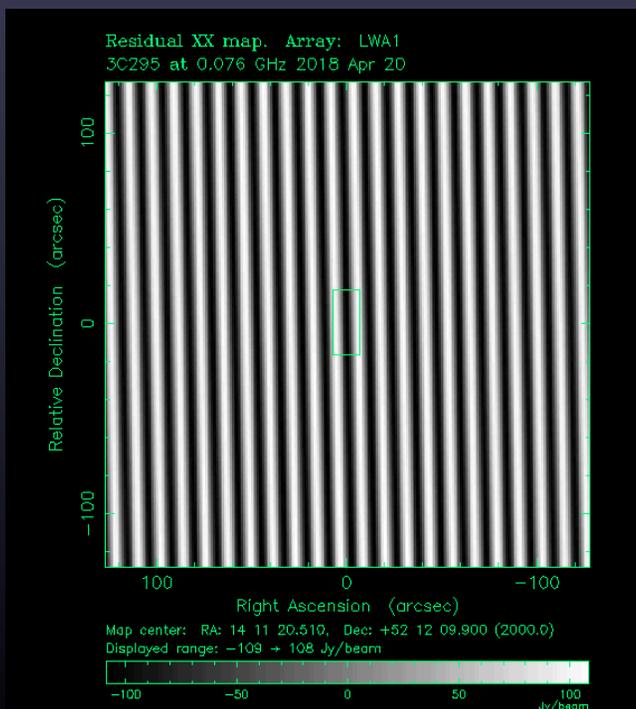
EXAMPLE 5

Bad visibility weight

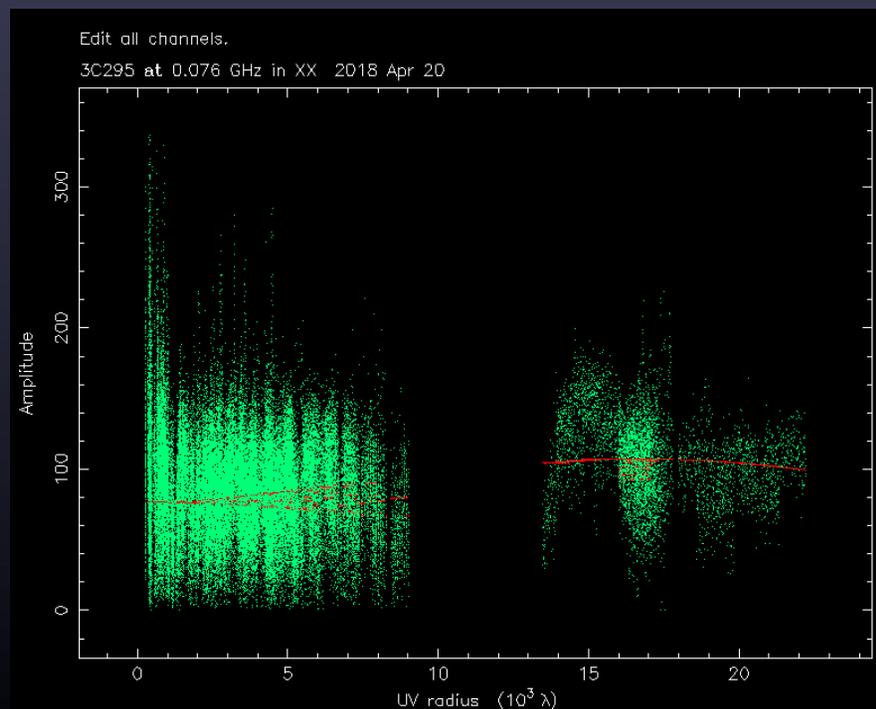
Occasionally visibility weights can be incorrect, producing severe Artifacts.

Symptom: Stripes from affected visibilities

Striping drowns out image

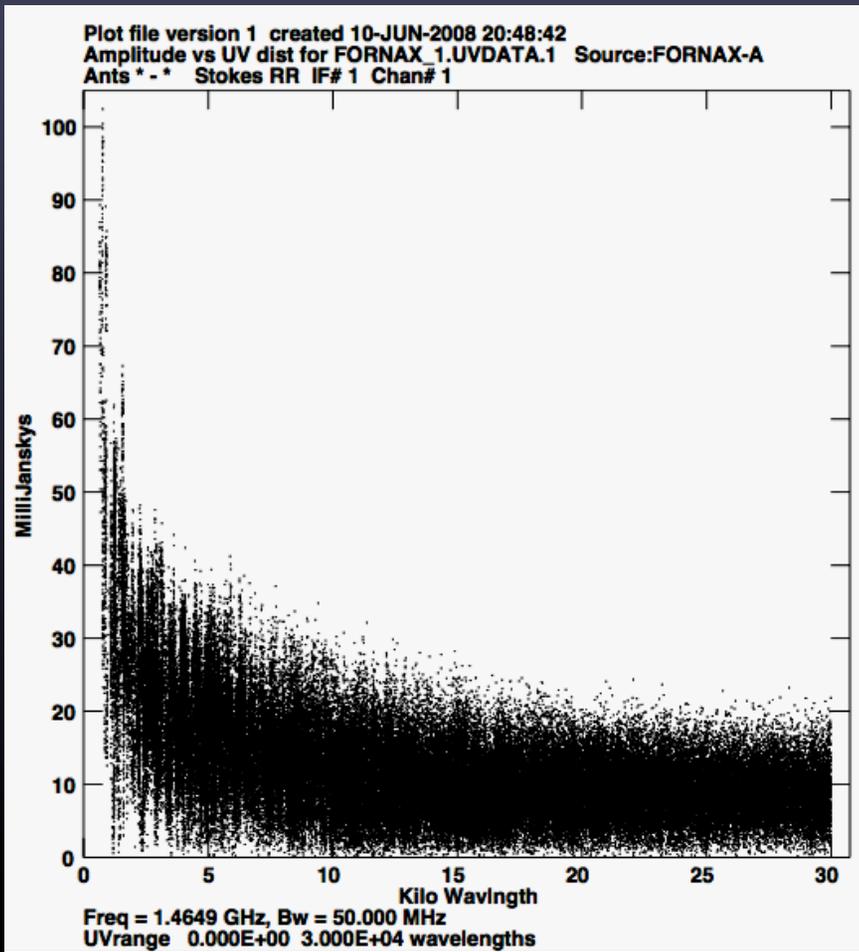


Visibility amplitudes vs u,v distance



DECONVOLUTION ERRORS

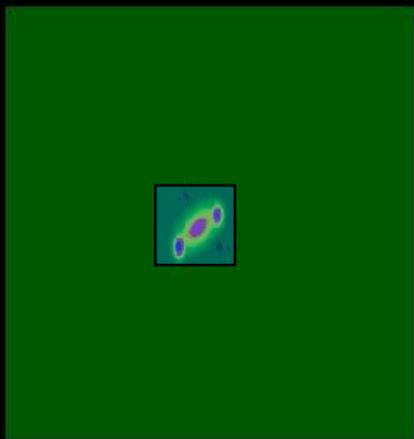
Even if the data are perfect, image errors and uncertainties can occur if the (u,v) coverage is not adequate to map the source structure.



The extreme rise of visibility at the short spacings makes it impossible to image the extended structure. You are better off imaging the source with a cutoff below about 2 kilo-wavelengths

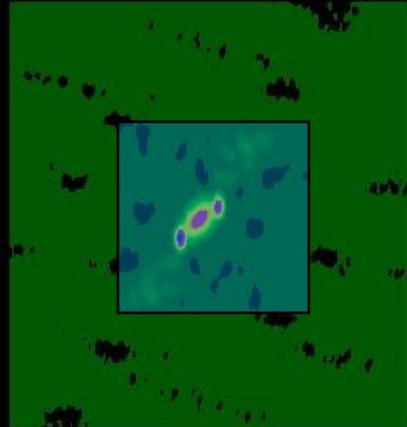
Get shorter spacing or single-dish data

CLEANING WINDOW SENSITIVITY



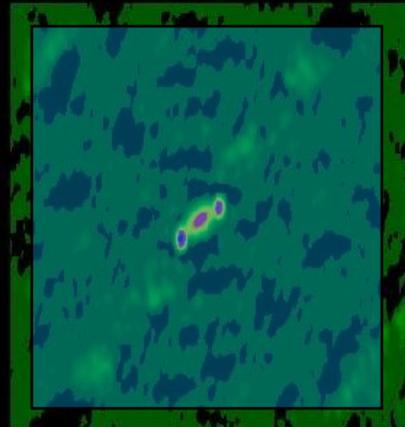
Tight Box

One small clean box



Middle Box

One clean box around all emission



Big Box

Clean entire inner map quarter

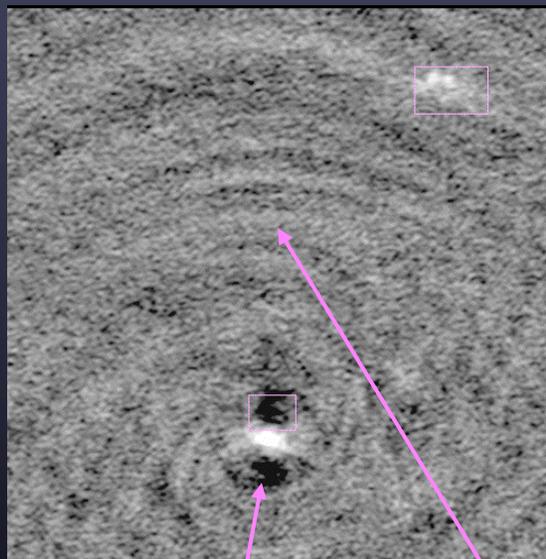


Dirty Beam

Make box as small as possible to avoid cleaning noise interacting with sidelobes

How Deep to Clean?

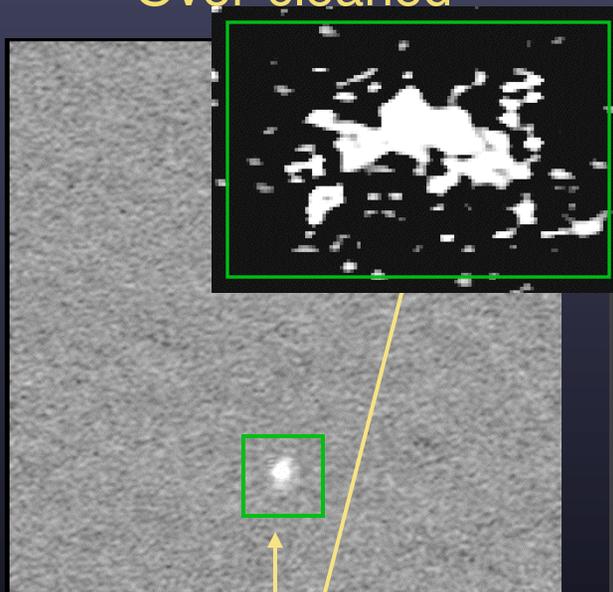
Under-cleaned



Emission from second source sits atop a negative "bowl"

Residual sidelobes dominate the noise

Over-cleaned



Regions within clean boxes appear "mottled"

Properly cleaned

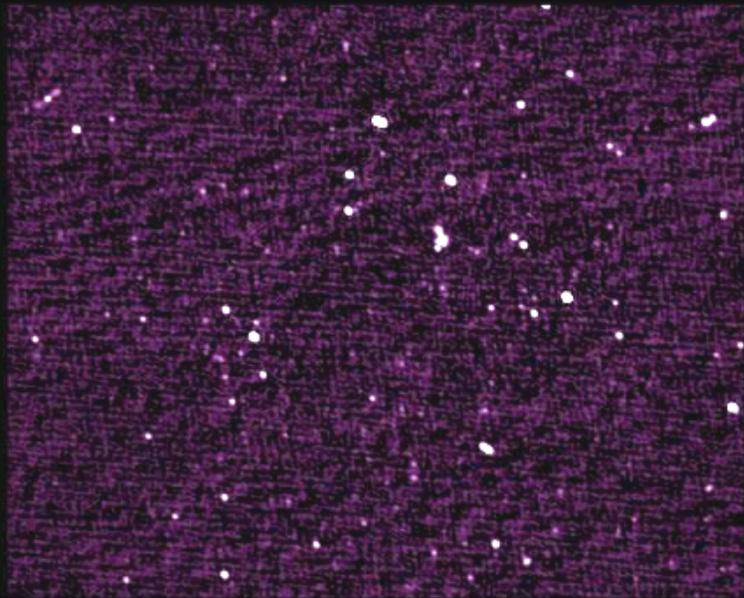


Background is thermal noise-dominated; no "bowls" around sources.

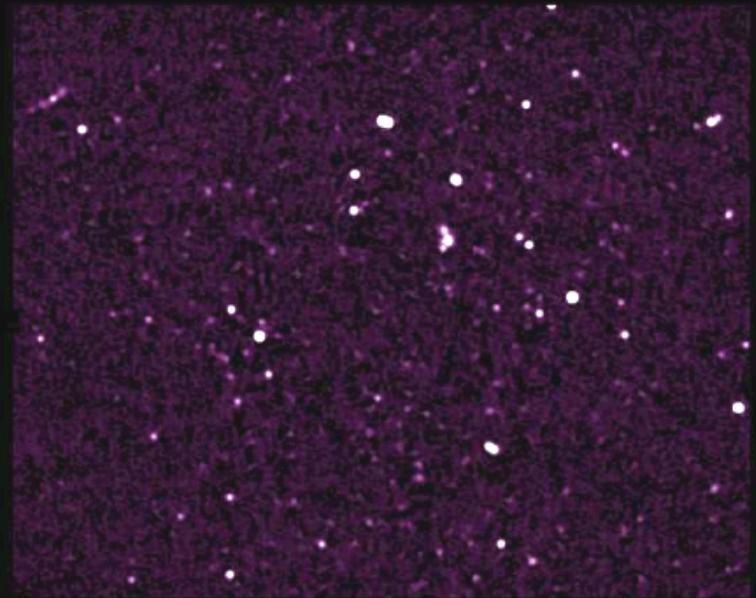
Improvement of Image

Removal of low level ripple improves detectability of faint sources

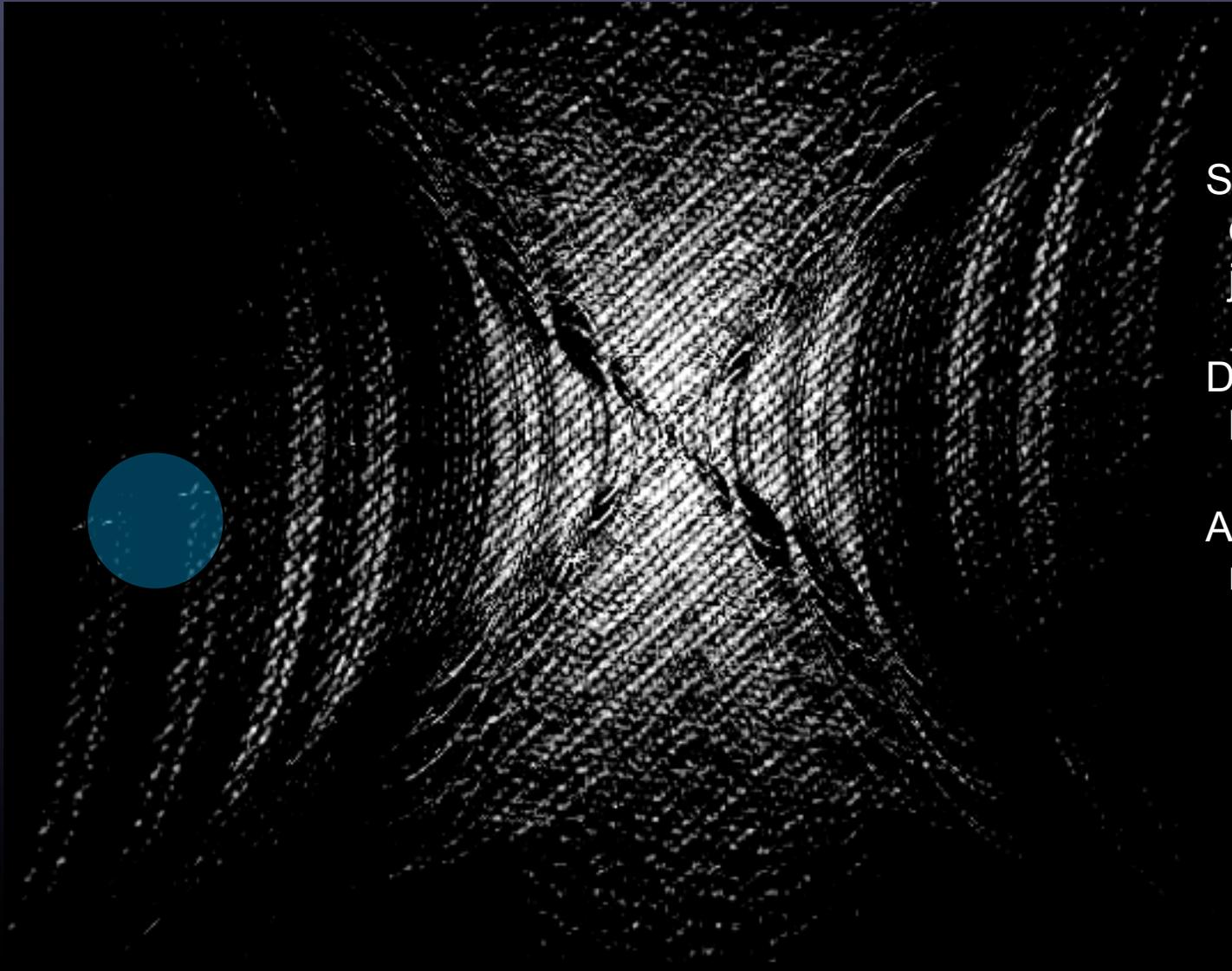
Before editing



After editing



Fourier Transform Dirty Image

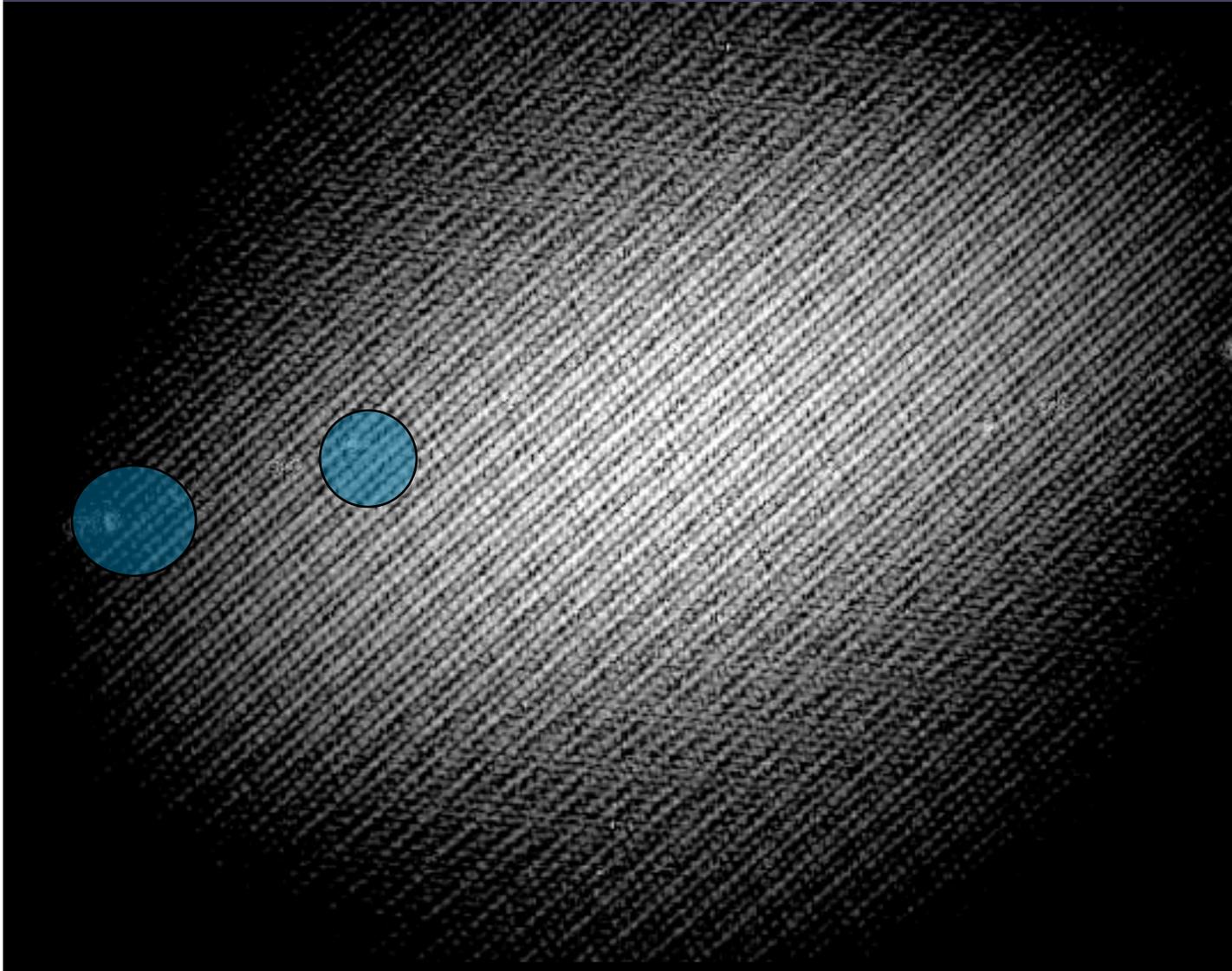


Shows the (u,v)
data as gridded
just before imaging

Diagonal lines caused
by structure in field

A few odd points are
not very noticeable

Fourier Transform Clean Image

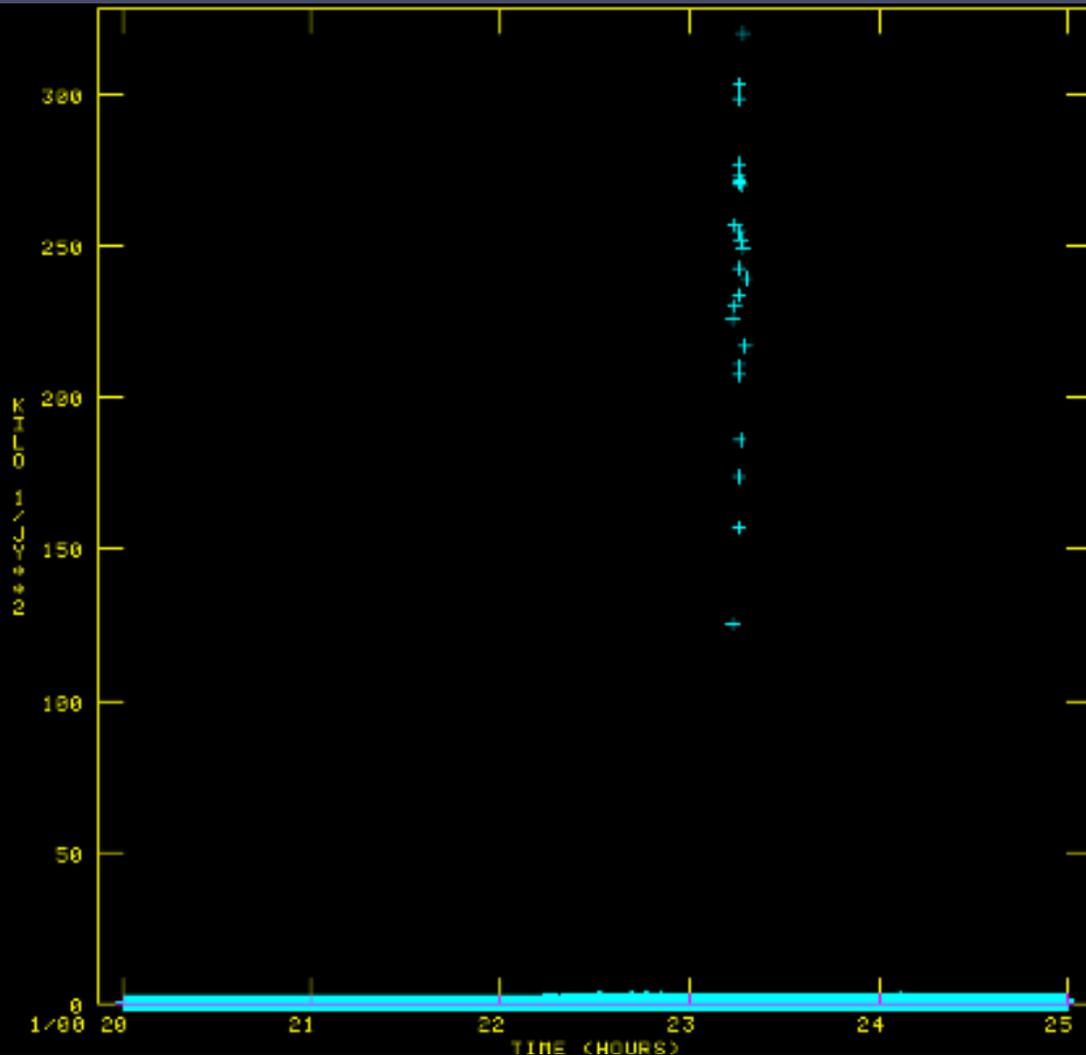


Shows the (u,v) data from clean image.

Diagonal lines still present. Notice that clean does an interpolation in the u,v plane between u,v tracks.

The odd points are smeared, but still present. These produce the low level ripples.

Bad weighting of a few (u,v) points

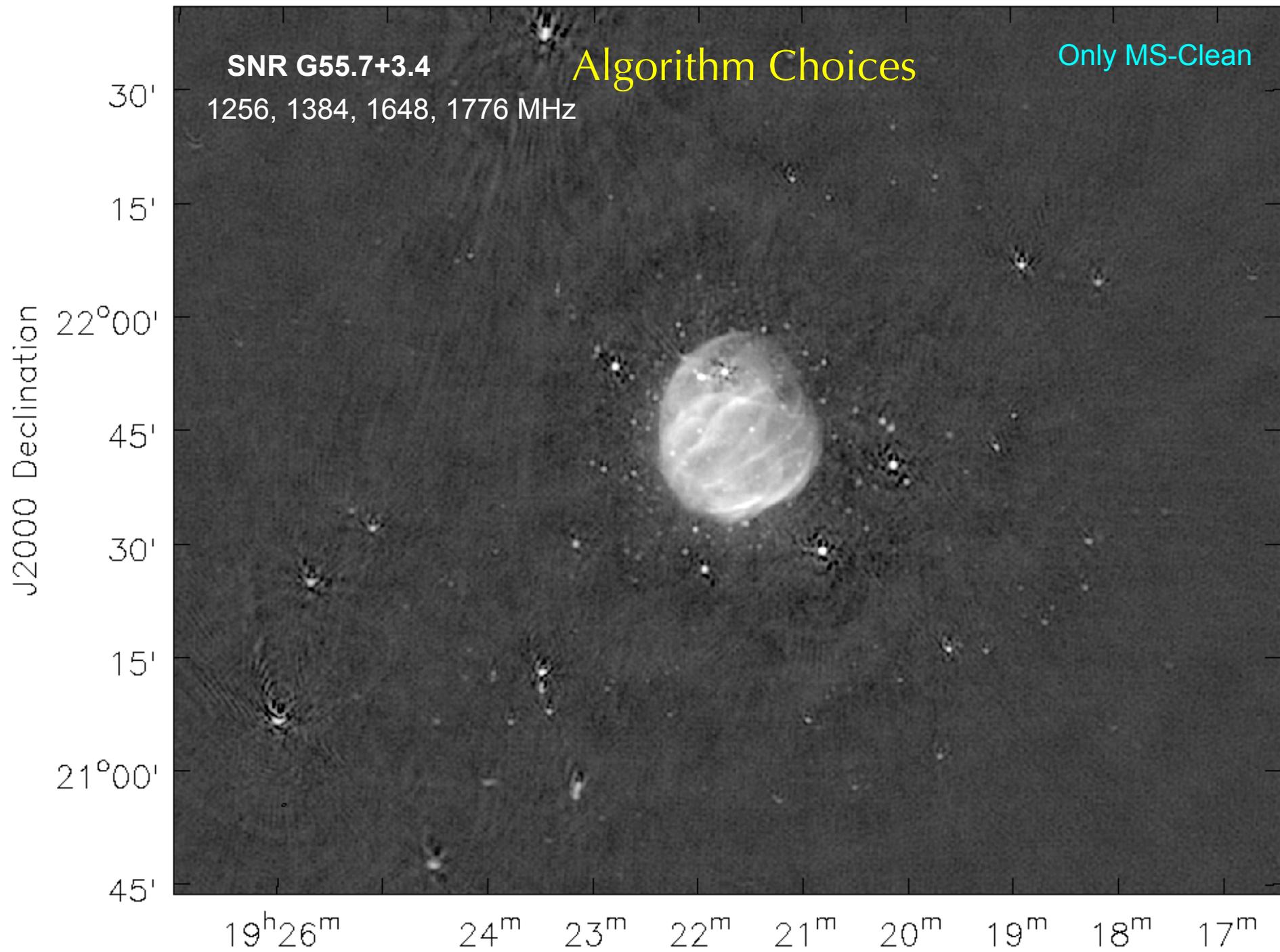


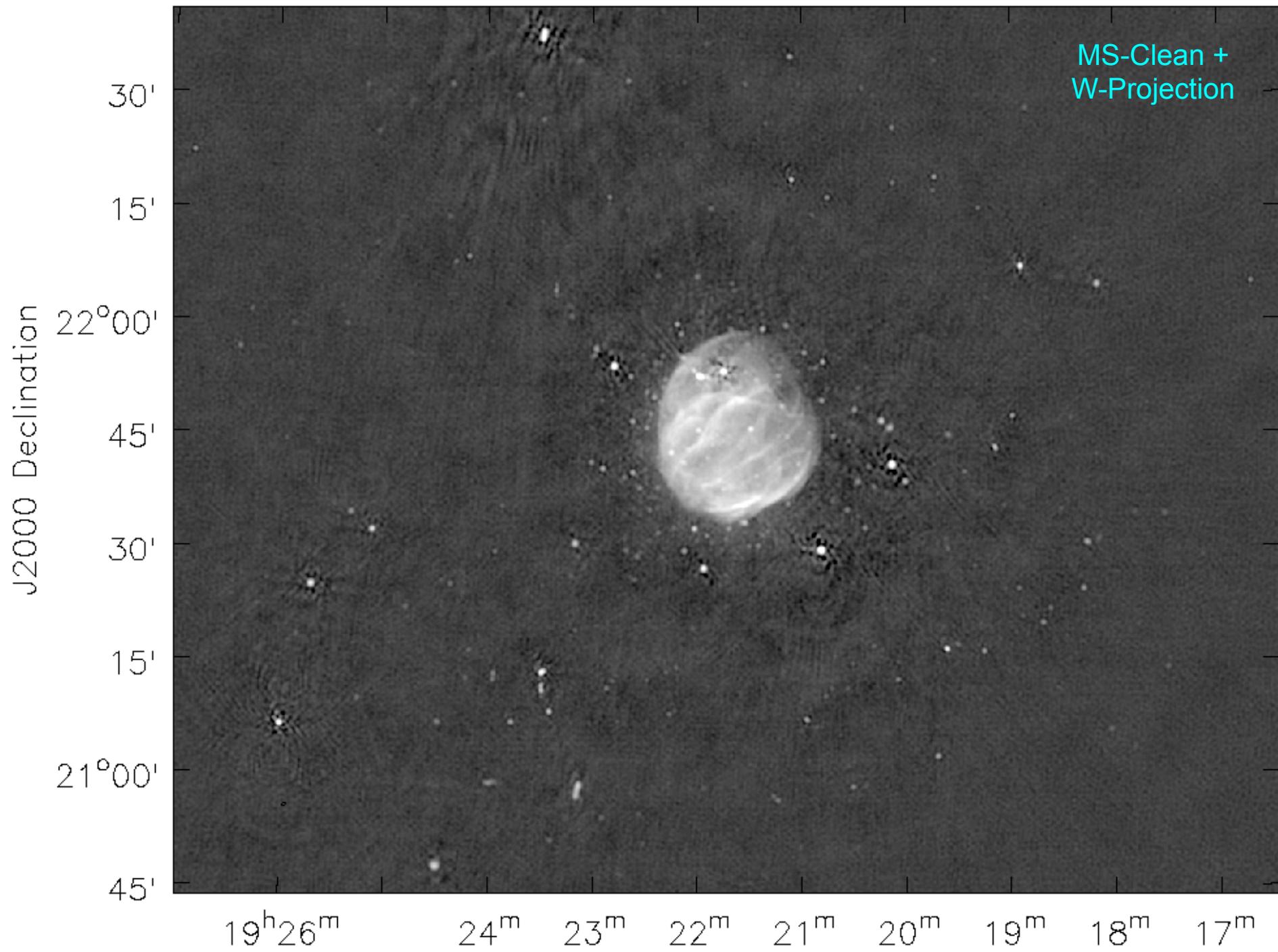
After a long search through the data, about 30 points out of 300,000 points were found to have too high of a weight by a factor of 100. Effect is <1% in image.

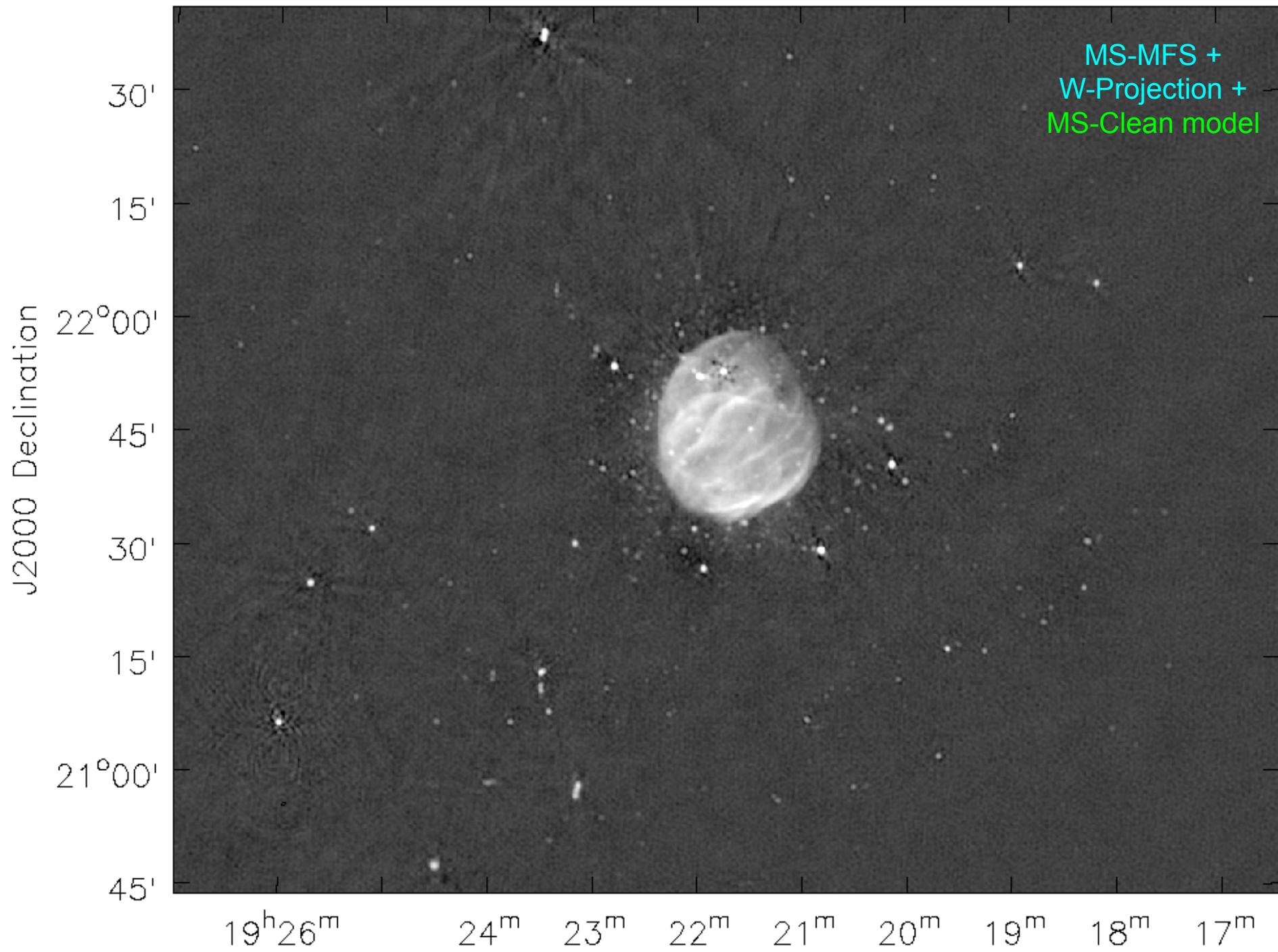
Cause??

Sometimes in applying calibration produced an incorrect weight in the data. Not present in the original data.

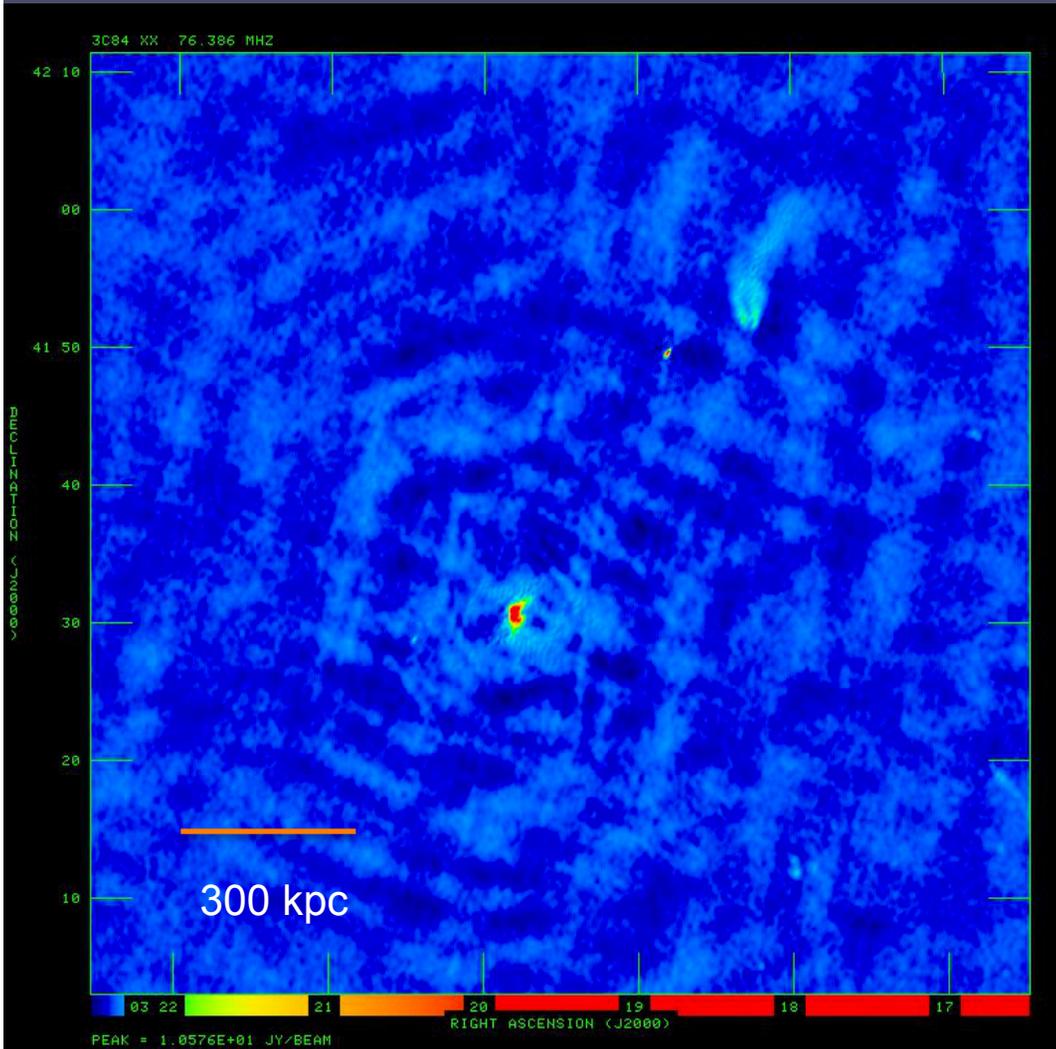
These problems can sneak up on you. Beware.







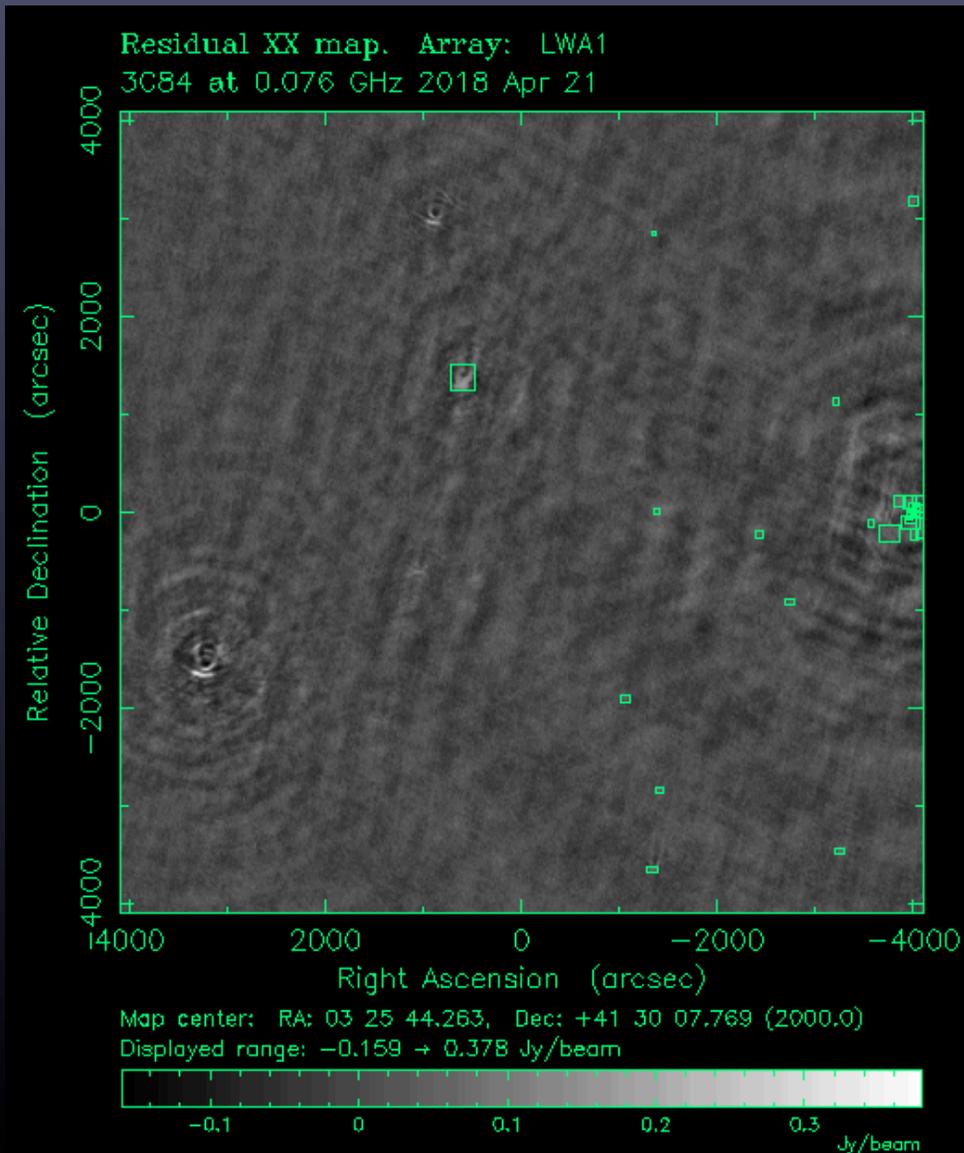
Are all sources included in the map/model?



At low frequencies the sky is filled with sources, make sure that you haven't left sources out of your model.

Strong sources outside the image can raise the noise significantly

Are all sources included in the map/model?



At low frequencies the sky is filled with sources, make sure that you haven't left sources out of your model.

Strong sources outside the image can raise the noise significantly

SUMMARY OF ERROR RECOGNITION

Source structure should be 'reasonable', the rms image noise as expected, and the background featureless. If not,

(u,v) data

Look for outliers in (u,v) data using several plotting methods.
Check calibration gains and phases for instabilities.
Look at residual data (u,v data - clean components)

IMAGE plane

Do defects resemble the dirty beam?
Are defect properties related to possible data errors?
Are defects related to possible deconvolution problems?
Are other corrections/calibrations needed?
Does the field-of-view encompass all emission?

Further Reading

- <http://www.nrao.edu/whatisra/>
- www.nrao.edu
- 2010 Lecture on Non-Imaging Analysis
- Synthesis Imaging in Radio Astronomy
- ASP Vol 180, eds Taylor, Carilli & Perley
- Numerical Recipes, Press et al. 1992