ERROR RECOGNITION & IMAGE ANALYSIS



Amy Mioduszewzki (NRAO)

Based on the lecture given by Gustaff van Moorsel at NRAO's 2010 Synthesis Imaging Workshop

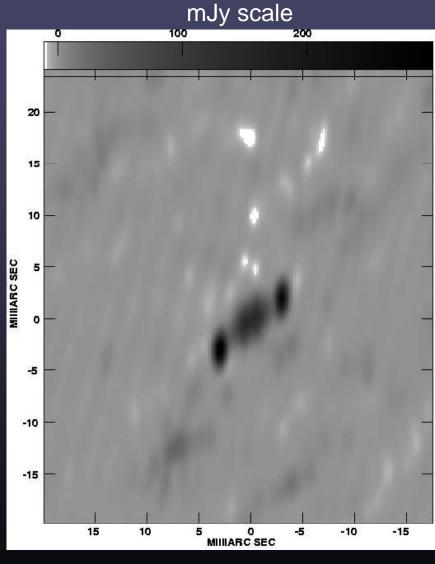
INTRODUCTION

- Why are these two topics 'Error Recognition' and 'Image Analysis' in the same lecture?
 - -- 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 almost infinite ways in which useful insight, information and parameters can be extracted from the image.
- Perhaps the two topics are related to the reaction one has when looking at an image after 'good' calibration, editing, self-calibration, etc.
- If the reaction is:

OBVIOUS IMAGE PROBLEMS

Yuck!!

- This can't be right. This
 is either the most
 remarkable radio source
 ever, or I have made an
 error in making the
 image.
- Clear signs of problems:
 - Image rms > expected rms
 - Unnatural features in the image
- How can the problems be found and corrected?



milliarcsec

HIGH QUALITY IMAGE

Great!!

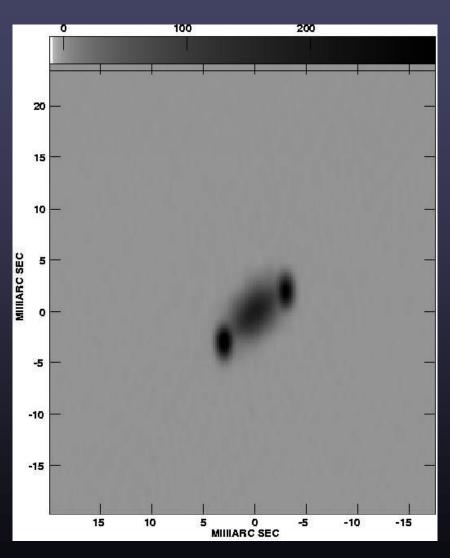
After lots of work, I can finally analyze this image and get some interesting scientific results.

What were defects?

- 1. Two antennas had 10% calibration errors
- 2. one with a 5 deg error
- 3. A few outlier data points.

This part of the 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). Self-calibration is sometimes 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, a digression on data and image display. First: Images

IMAGE DISPLAYS (1)

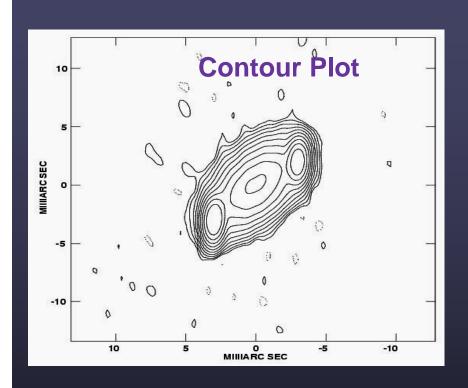
Pirel values																														
		235					245				255				265						2									
287	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
285	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Q.	0	0	0	1	1	0	0	0	0	0	0	0
283	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0
281	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	2	3	3	.3	4	3	1	0	0	0	0
279	0	-	0	0	0	0	0	0	0	0	0	0	0	0	1	1	2	3	4	4	5	8	8	12	8	3	1	0	0	0
277	0		0	0	0	0	0	0	0	0	0	0	0	1	2	3	5	7	7	8	9	9	19	32	22	6	1	0	0	0
275	0		0	0	0	0	0	0	0	0	0	0	1.	2	4	6	9	13	14	15	14	16	40	72	47	12	2	0	0	0
273	0		0	0	0	0	0	0	0	1	1	1	2	4	8	12	17	22:	23	24	22	27	77:	136	87	19	2	0	0	0
271	0	:	o-	0	Q.	0	0	0	0	1	1	2	4	8,	15	21	28	35	36	37	33	431	26:	217	132	28	3	0:	0	0
269	0		Ō.	0	0	0	Ō	0	0	- 1	3	4	8	15	25	34	44	54.	54	53	48	611	73:	298	168	34	3	0	0	0
267	0		0	0	0	0	0	0	1	2	4	8	14	25	40	52	67	79	77	74	63	781	993	316	177	34	3	0	0	0
265	0		0	0	0	0	0	0	1	3	. 7	14	24	40	60	77	97 :	1091	102	93	74	791	912	289	155	29	3	0	0	0:
263	0		Ú.	0	0	0	0	1	2	5	11	22	37	58	86:	108:	13 0:	137 1	123	105	79	731	542	220	113	20	2	0	0	0
261	0	1	Ø.	0	0	0	1	1	3	8	17	33	54	81	116:	139:	15 6:	LS 6:	133:	107	75	6 11	.06:	140	69	12	2	0	0	O:
259	0		0	0	0	0	1	2	5	12	24	45	721	08:	1431	62:	170:	l61 :	L31	99	66	47	64	78	36	6	1	0	0	0
257	0		0	0	0:	0	2	4	8	18	32	58	881	24:	1601	71:	l69:	L\$2 :	118	86	88	36	36	36	16	3	1	0	0	0
255	0	;	Ò	0	0	1	2	7	16	27	42	703	1011	35:	1621	641	.561	341	00.	71	44	27	20	16	7	1	0	0:	0	0
253	0		0	0	0	- 1	4	15	34	43	51	77:	1051	33:	1501	461	351	12	81	56	34	19	11	7	.3	1	0	0	0	0
251	0		0	O .	Q-	1	8	34	73	70	59	793	1001	20:	1301	221	10	88	61	41	24	12	6	3	1	0.	0	0	0	0
249	0		0	0	1	2	14	690	1411	12	68	73	871	00:	106	96	83	64.	43	27	14	7	3	1	1	0	0	0	0	0:
247	0		0	0	1	3	23:	121:	2381	ķ7	69	62	69	77	81	70	59	42	26	16	8	3	1	0	0	0	0	0	0	0
245	0	1	0	0	1	3	34:	1800	33.82	217	69	48	52	86	57	47	36	25	15	8	3	1	. 0	0	0	0	0	0	0	0
243	0	-	0	0	1	4	42:	22:24	40 2 2	42	68	36	37	38	37	29	21	14.	7	4	1	0	0	0	0	0	0	0	0	0
241	0		0	0	1	4	443	229:	3982	28	66	26	25	25	22	16	11	7	3	1	0	0	0	0	0	0	0	0	0	0
239	0		0	0	1	3	39:	1960	32.7 i	79	41	18	16	15	12	8	5	3	1	1	0	0	0	0	0	0	0	0	0	0
237	0		0	0	1	3	28:	1390	2231	18	26	11	9	8	6	4	2,	1	1	0	0	0	0	0	0	0	0	0	0	0
235	0	1	0	0	0	2	18	82:	1.27	64	14	6	5	4	3	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
233	0	:	0	0	0	1	9.	40	60	29	7	3	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0:	0	0
231	0		Ō.	0	0	0	4	17	23	11	3	1	1.	0	0	0	0	Q.	0	0	0	0	0	Ů.	ø	0	0	0	0	0
229	0		0	0	0.	0	2	6.	7	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.	0	0	0	0
227	0		0	0	0	0	1	2	2	.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
225	0		0	0	0	0	Ű	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
223	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

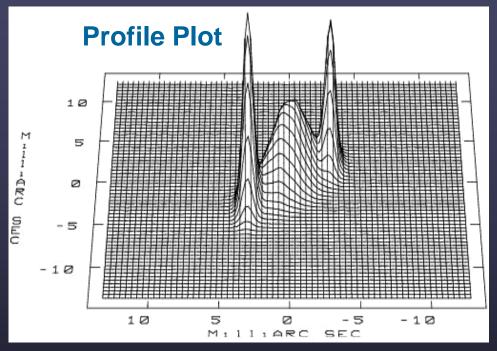
Digital image

Numbers are proportional to the intensity

Good for very slow links; rarely used anymore

IMAGE DISPLAYS (1)

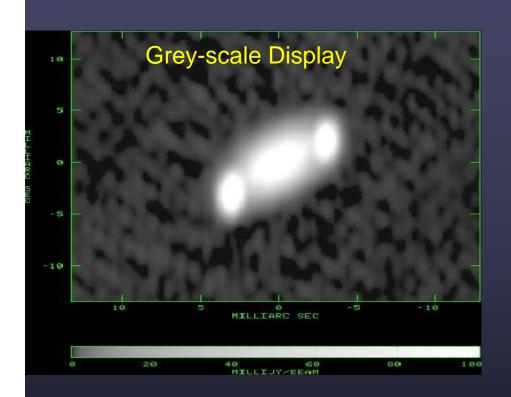


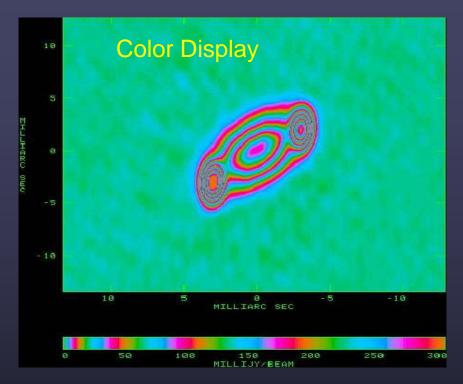


These plots are easy to reproduce and printed

- Contour plots give good representation of faint emission.
- Profile plots give a good representation of the bright emission and faint ripples.

IMAGE DISPLAYS (2)





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

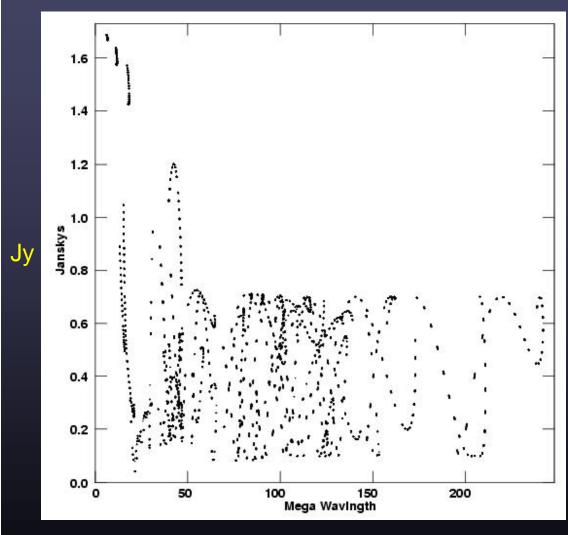
DATA DISPLAYS(1)

List of u-v Data

Source=	J0121+11		Freq=	8.	434	1858511	Sort:	= TB	1	RR	
Vis #	TAT		Ant	Su	Fq	U(klam)	V(klam)	W(klam)	Amp	Phas	Vt.
2191	0/22:35:	08.22		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

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

DATA DISPLAYS(2)

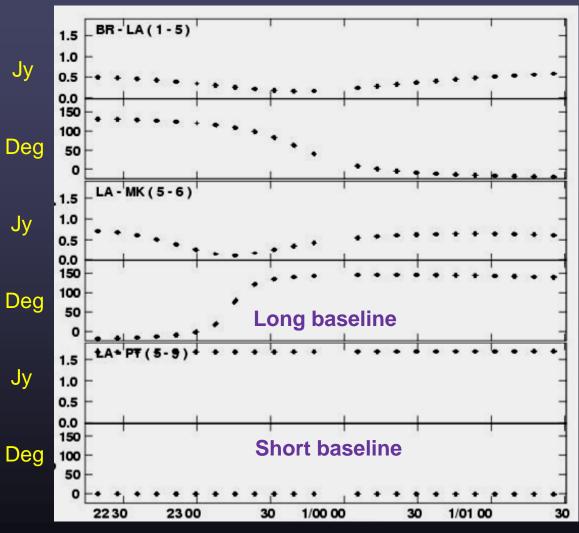


Visability amplitude vs. u-v distance plot

- General trend of data
- Useful for strong sources
- This is model from image showed earlier
 - Large fuzzy structure caused rise in flux towards short spacings
 - Oscillations at longer spacings suggest close double

Mega Wavelength

DATA DISPLAYS(3)



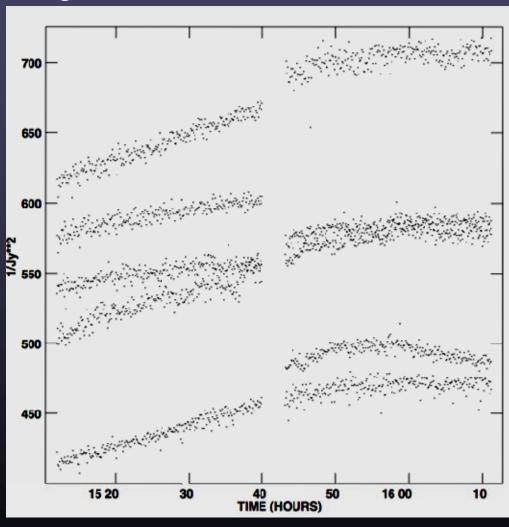
Amplitude and phase vs. time on various baselines

- Good for determining the continuity of the data
- Should be relatively smooth with time
- Outlier are obvious

Time in d/hh mm

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.
- Check weights to make sure they are reasonable.

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

Errors obey Fourier transform relationship

- •Narrow feature in uv plane ↔ wide feature in image plane
- •Wide feature in uv plane ↔ narrow feature in image plane
 - Note: easier to spot narrow features
- Orientations are orthogonal
- •Data uv amplitude errors ↔ symmetric image features
- •Data uv phase errors ↔ 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

GOLDEN RULES OF 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 in the image plane other than a slight increase in noise)

Persistent small data errors are a bigger problem:

- 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:

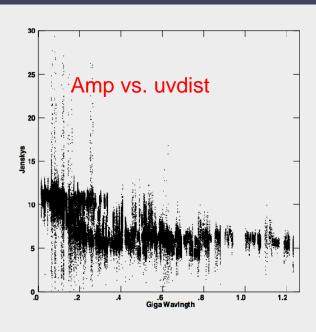
Perfect data but unstable algorithms. Common but difficult to discern

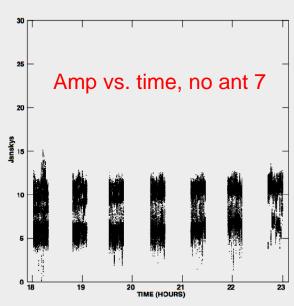
ERROR RECOGNITION IN THE U-V PLANE

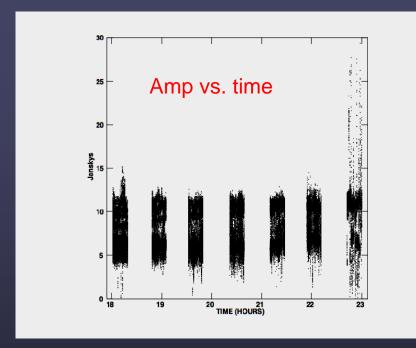
Editing obvious errors in the u-v plane

- •Mostly consistency checks assume that the visibility cannot change much over a small change in u-v spacing
- •Also, double-check gains and phases from calibration processes. These values should be relatively stable.

VISIBILITY AMPLITUDE PLOTS







Amp vs. uvdist shows outliers

Amp vs. time shows outliers in last scan

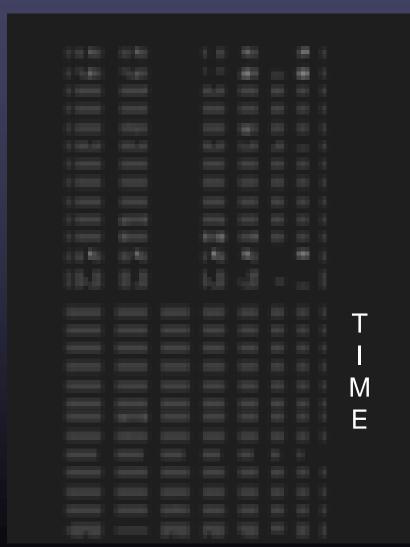
Amp vs. time without ant 7 should good data

(3C279 VLBA data at 43 GHz)

VISIBILITY AMPLITUDE RASTERS

BASELINE

Ant 1 2 3 4 5 6 7 8



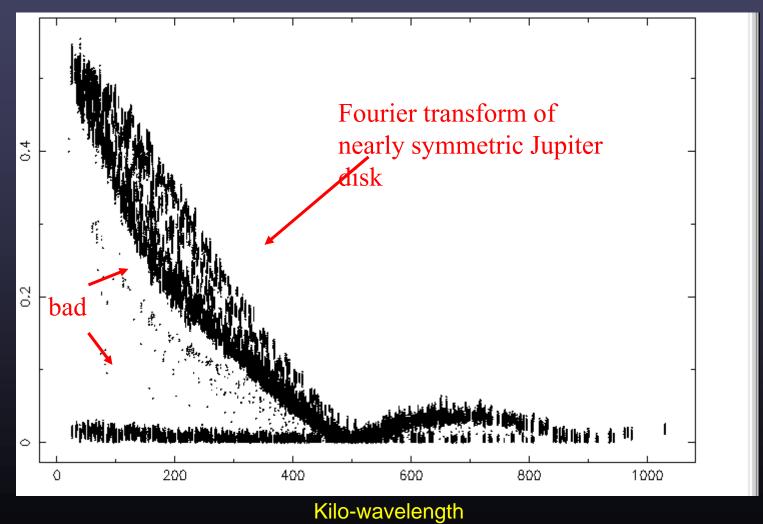
(Last two scans from last slide)

Use AIPS task TVFLG or CASA viewer

- •Raster scan of baseline versus time immediately shows where the bad data are
- Pixel range is 5 to 20 Jy
- Bad data can be flagged with an interactive clipping control

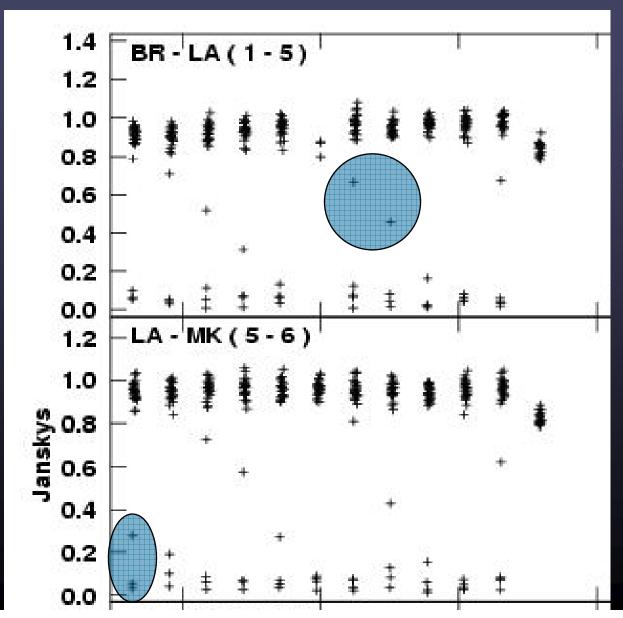
Example Edit (2)

AIPS WIPER or CASA plotms



Jansky

Drop-outs at Scan Beginnings



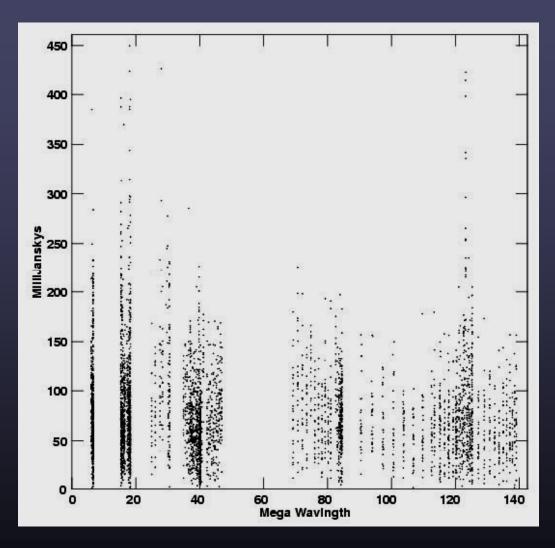
Often the first few points of a scan are low. E.g. antenna not on source.

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

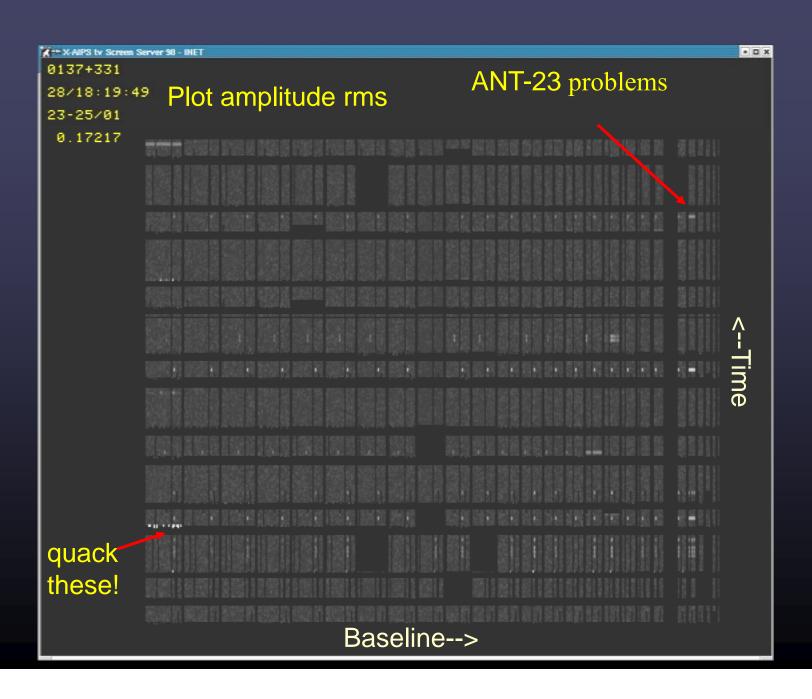
Editing Noise-dominated Sources

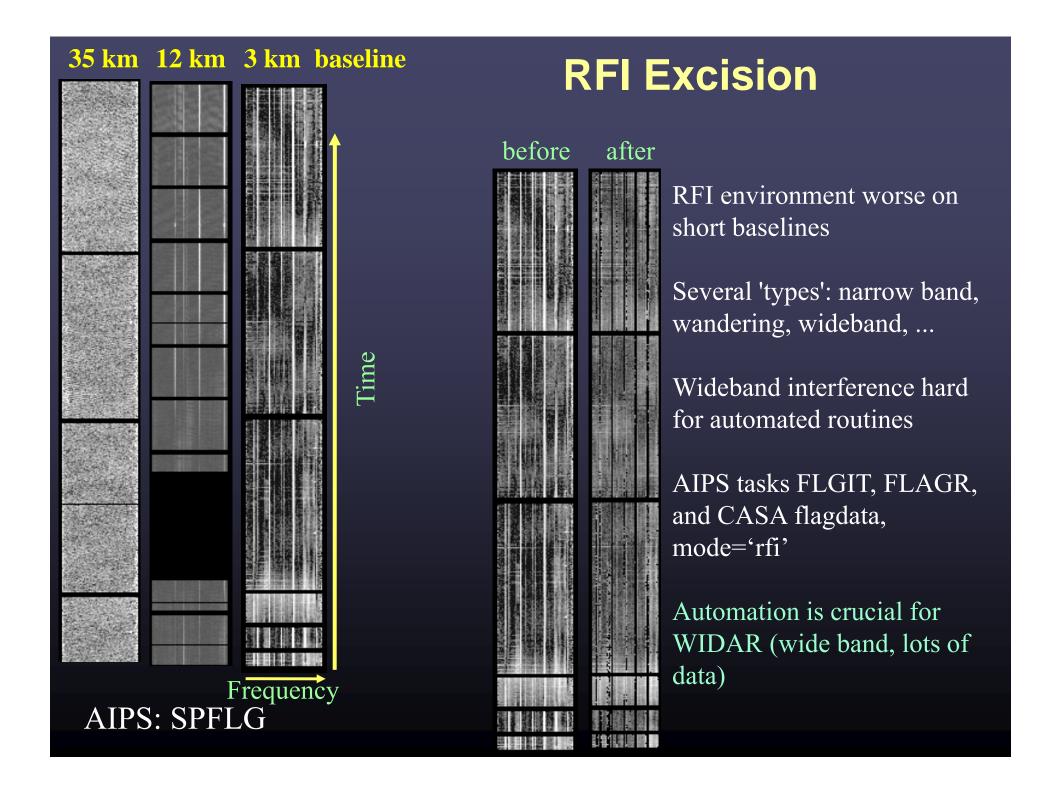


No source structure information is detected. Noise dominated.

All you can do is remove outlier points above 0.3 Jy. Precise level not important as long as large outliers removed.

USING TVFLG (VIEWER) DISPLAY on a source





ERROR RECOGNITION IN THE IMAGE PLANE

Some Questions to ask:

Noise properties of image:

Is the rms noise about that expected from integration time? 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 4xrms noise

Does the image have characteristics in the dirty beam?

Image-making parameters:

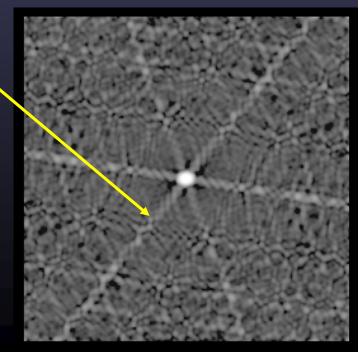
Is the image big enough to cover all significant emission? Is cell size too large or too small? ~4 points per beam okay Is the resolution too high to detect most of the emission?

EXAMPLE 1 Data bad over a short period of time

Results for a point source using VLA. 13-5min observations 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

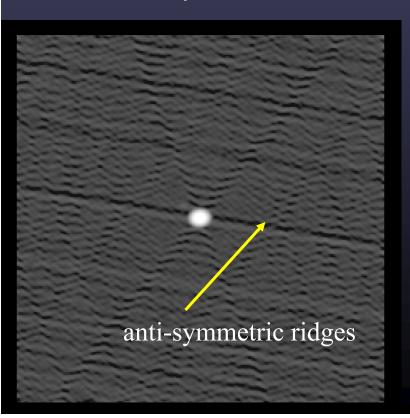
6-fold symmetric pattern due to VLA "Y". Image has properties of dirty beam.



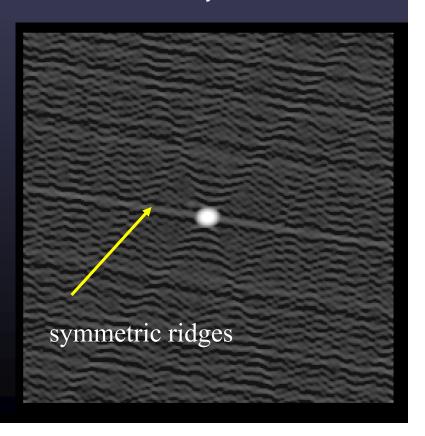
EXAMPLE 2 Short burst of bad data

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

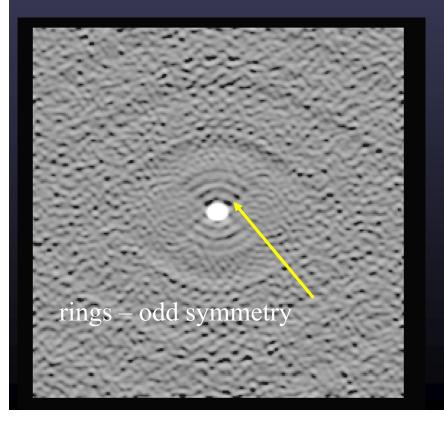


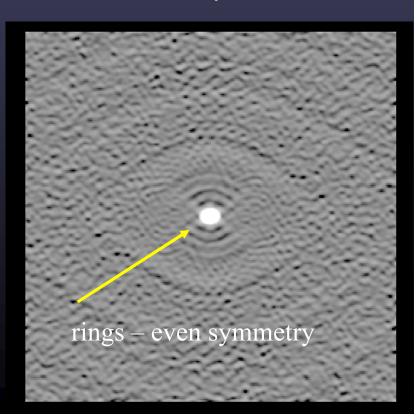
EXAMPLE 3Persistent errors over most of observations

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

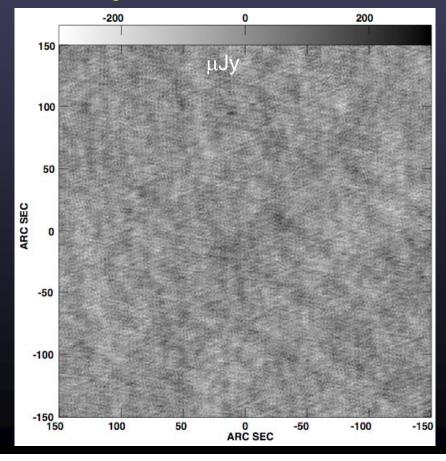
Occasionally correlators produce ghost signals or cross talk signals Occurred during change-over from VLA to EVLA system

Symptom: Garbage near phase center, dribbling out into image

Image with correlator offsets

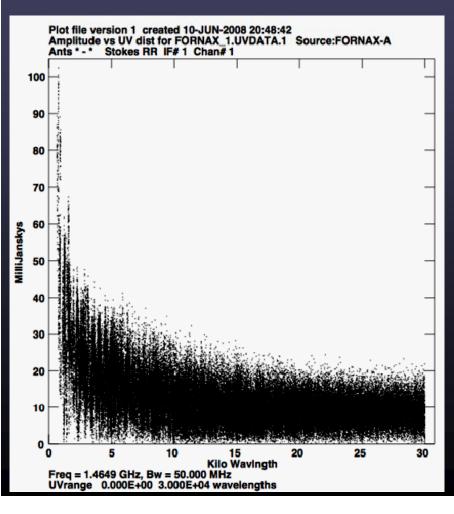
-200 200 ARC SEC ARC SEC

Image after correlation of offsets



DECONVOLUTION ERRORS

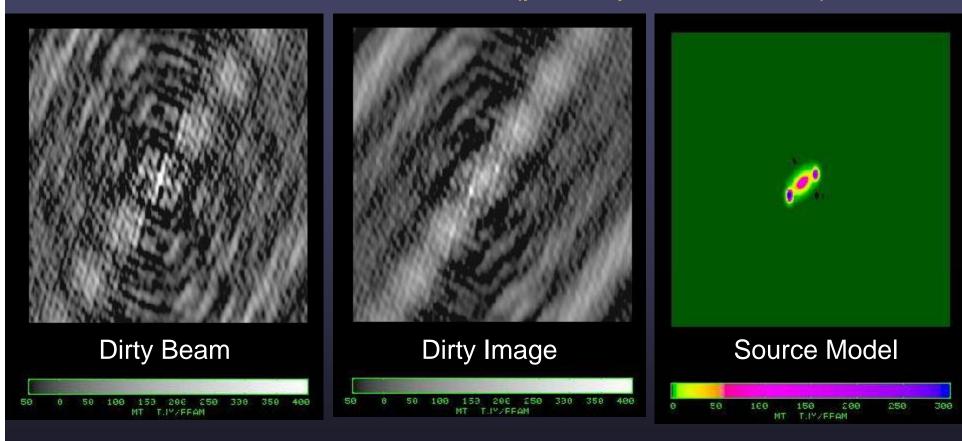
Even if the data are perfect, image errors and uncertainties will occur because the (u-v) coverage is not adequate to map the source structure.



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

Get shorter spacing or single-dish data

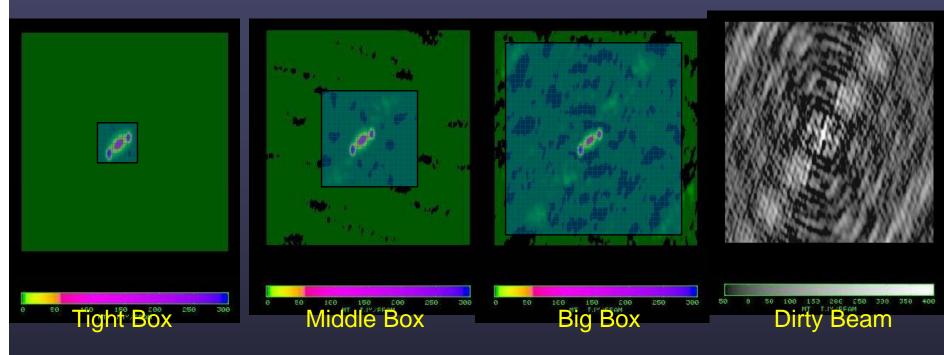
DIRTY IMAGE and BEAM (point spread function)



The dirty beam has large, complicated side-lobe structure. It is often difficult to recognize any details on the dirty image. An extended source exaggerates the side-lobes.

5% in dirty beam becomes 20% for extended source

CLEANING WINDOW SENSITIVITY



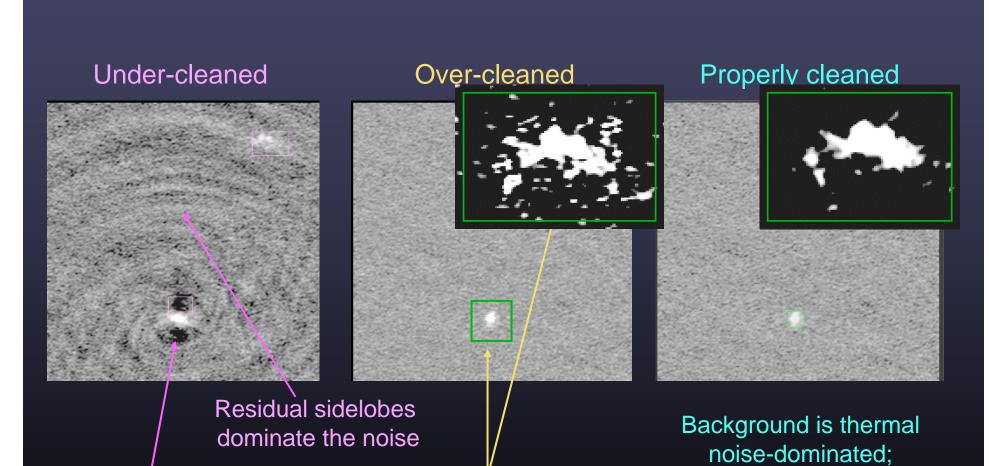
One small clean box

One clean box around all emission

Clean entire inner map quarter

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

How Deep to Clean?



Region's within

clean boxes

appear "mottled"

no "bowls" around

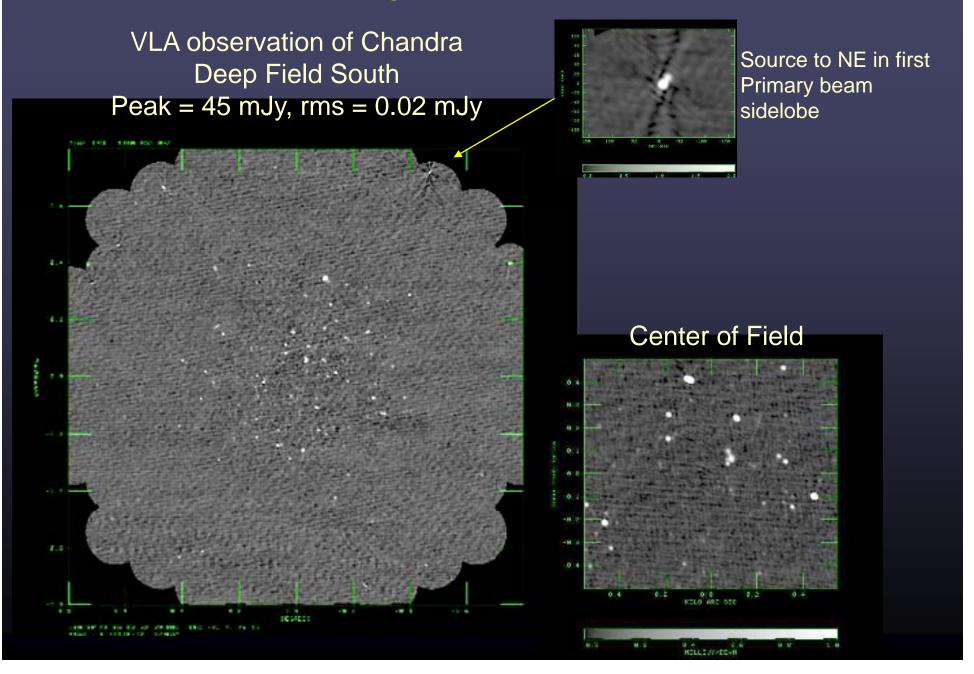
sources.

Emission from

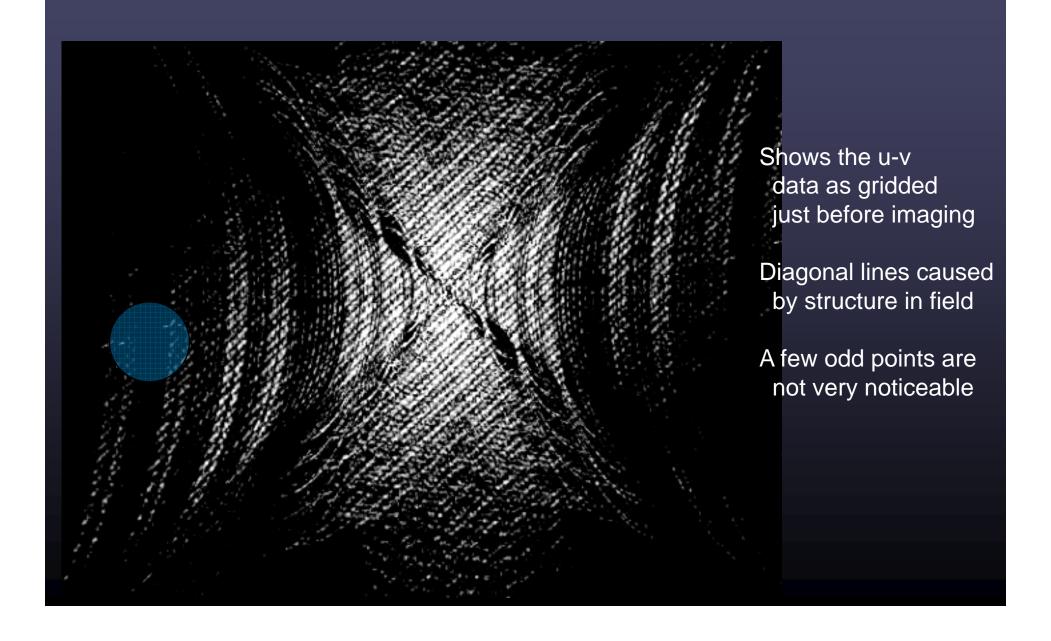
second source sits

atop a negative "bowl"

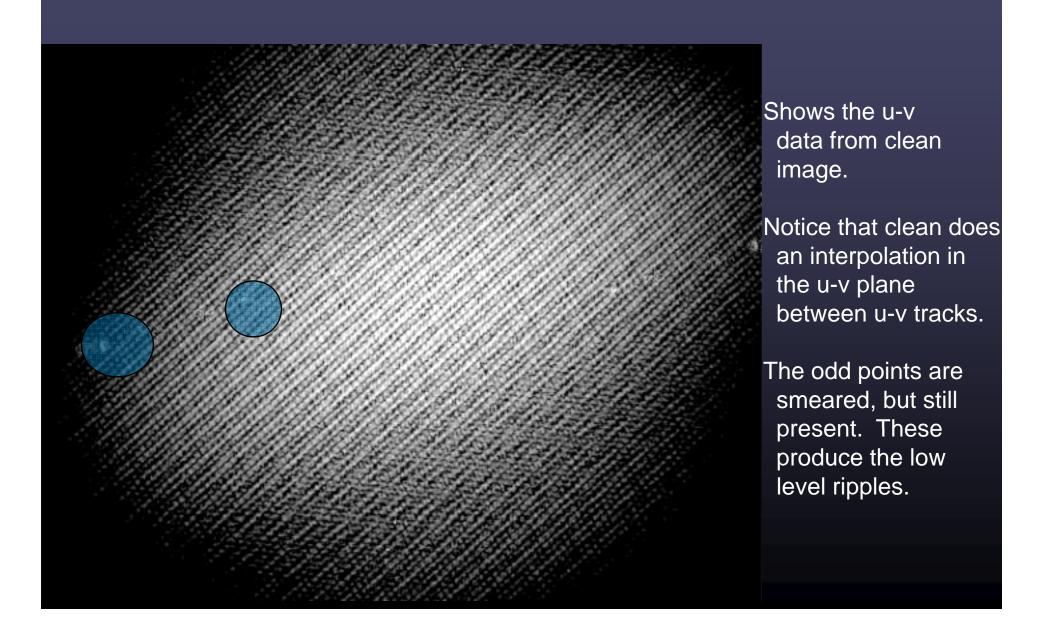
FINDING HIDDEN BAD DATA



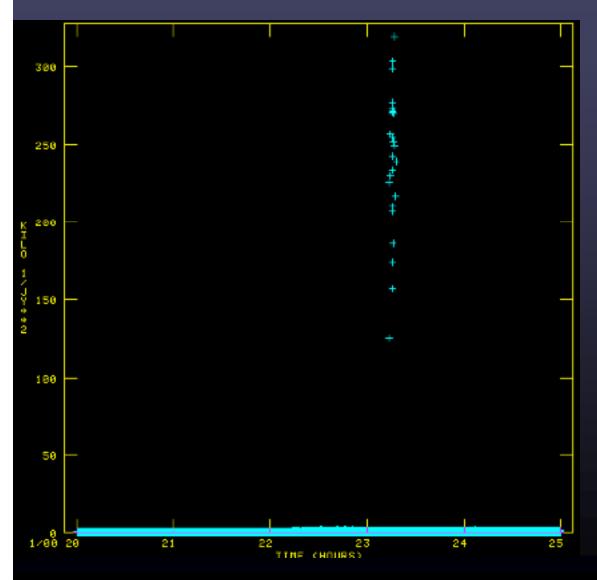
Fourier Transform Dirty Image



Fourier Transform Clean Image



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??

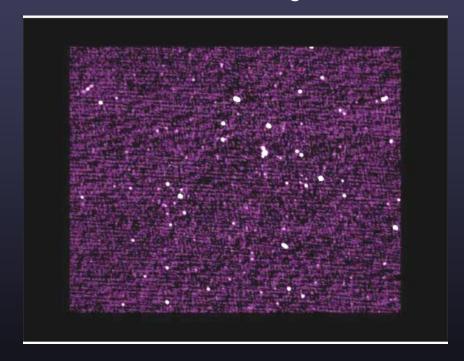
Sometimes in applying calibration produced an incorrect weight in the data. Not present in the original data.

These problems can sneak up on you. Beware.

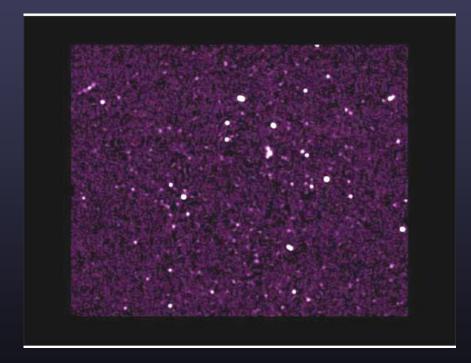
Improvement of Image

Removal of low level ripple improves detectability of faint sources

Before editing



After editing



SUMMARY OF ERROR RECOGNITION

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

UV data

Look for outliers in u-v data using several plotting methods.

Check calibration gains and phases for instabilities.

Look at residual data (uv-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?

IMAGE ANALYSIS



IMAGE ANALYSIS

- Input: Well-calibrated data-base producing a high quality image
- Output: Parameterization and interpretation of image or a set of images

This is very open-ended

Depends on source emission complexity Depends on the scientific goals

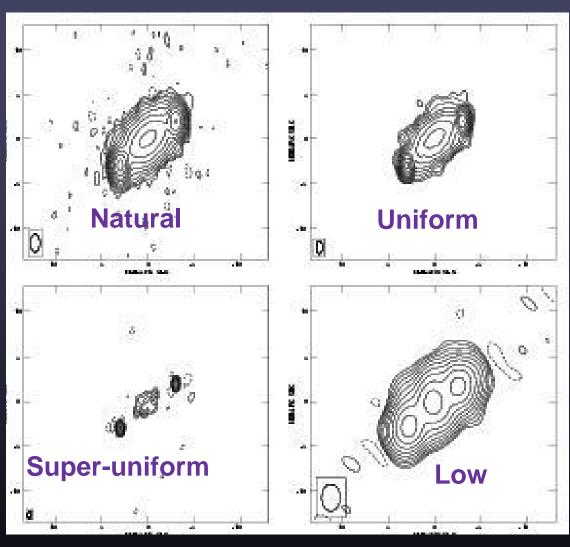
Examples and ideas are given.

Many software packages, besides AIPS and Casa (e.g., IDL, DS-9) are available.

IMAGE ANALYSIS OUTLINE

- Multi-Resolution of radio source.
- Parameter Estimation of Discrete Components
- Polarization Data
- Image Comparisons
- Positional Registration

IMAGE AT SEVERAL RESOLUTIONS



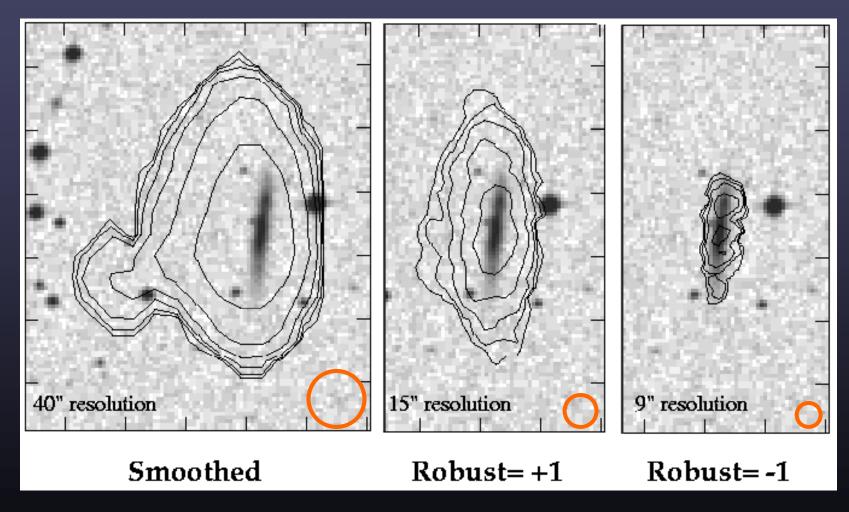
Different aspect of source structure can be see at various resolutions, shown by the ellipse in the lower left corner of each box.

SAME DATA USED FOR ALL IMAGES

For example,
Outer components are small
from SU resolution
There is no extended
emission from low resolution

Imaging and Deconvolution of Spectral Line Data:

Type of weighting in imaging



HI contours overlaid on optical images of an edge-on galaxy

PARAMETER ESTIMATION

Parameters associated with discrete components

Fitting in the image

- Assume source components are Gaussian-shaped
- Deep cleaning restores image intensity with Gaussian-beam
- True size * Beam size = Image size, if Gaussian-shaped.
 Hence, estimate of true size is relatively simple.

• Fitting in (u-v) plane

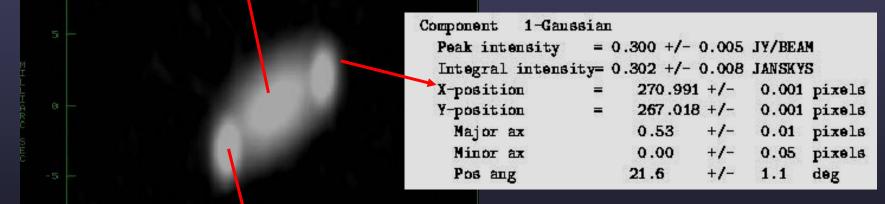
- Better estimates for small-diameter sources
- Can fit to any source model (e.g. ring, disk)

Error estimates of parameters

- Simple ad-hoc error estimates
- Estimates from fitting programs

IMAGE FITTING

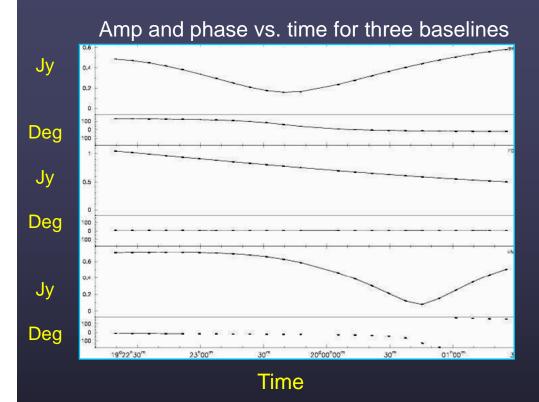
```
2-Gaussian
Component
  Peak intensity
                    = 0.104 +/- 0.005 JY/BEAM
  Integral intensity= 0.998 +/- 9,47
                                      JANSKYS
 X-position
                         255.986 +/-
                                       0.0029 pixels
  Y-position
                         257.033 +/-
                                       0.0032 pixels
                                       0.02
                                              pixels
   Major ax
                        19.99
                                 +/-
                                       0.03
                                              pixels
   Minor ax
                         9.98
                       135.3
                                 +/-
                                       0.1
   Pos ang
                                              deg
```

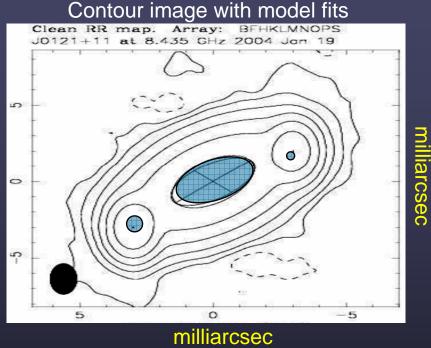


```
3-Gaussian
Component
 Peak intensity
                   = 0.393 +/- 0.004 JY/BEAM
 Integral intensity= 0.403 +/- 0.008 JANSKYS
 X-position
                        241.007 +/-
                                      0.001 pixels
 Y-position
                        241.988 +/-
                                      0.001 pixels
   Major ax
                        1.54
                                +/-
                                     0.01 pixels
   Minor ax
                        0.21
                                +/-
                                      0.01 pixels
                        3.6
                                +/-
                                      0.2
                                            deg
   Pos ang
```

AIPS task: JMFIT Casa tool imfit

(U-V) DATA FITTING





DIFMAP has good u-v fitting algorithm

Fit model directly to (u-v) data Compare model to data Contour display of image
Ellipses show true component
size. (super-resolution?)

See Greg Taylor's talk at the 12th Synthsis Imaging Workshop, "Non-image Data Analysis"

COMPONENT ERROR ESTIMATES

```
P = Component Peak Flux Density
```

$$\sigma$$
 = Image rms noise

$$P/\sigma$$
 = signal/noise = **S**

$$\theta_{i}$$
 = Component image size

$$\Delta P$$
 = Peak error = σ

$$\Delta X$$
 = Position error = $B/2S$

$$\Delta\theta_{i}$$
 = Component image size error = $B/2S$

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

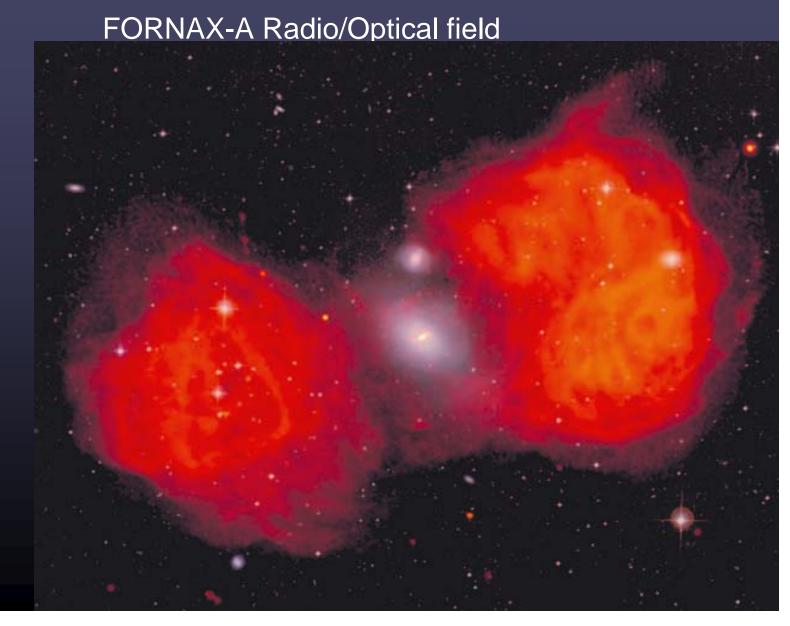
$$\Delta\theta_t$$
 = Minimum component size = $B/S^{1/2}$

Comparison and Combination of Images of Many Types

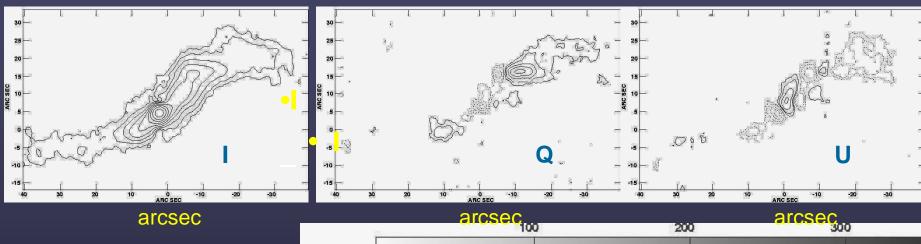
Radio is red Faint radio core in center of NGC1316

Optical in blue-white

Frame size is 60' x 40'

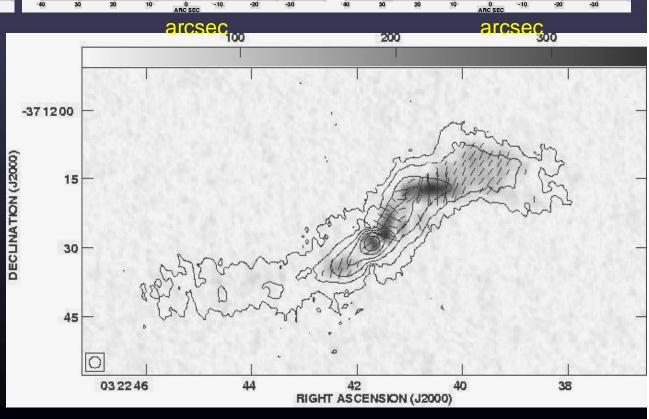


LINEAR POLARIZATION

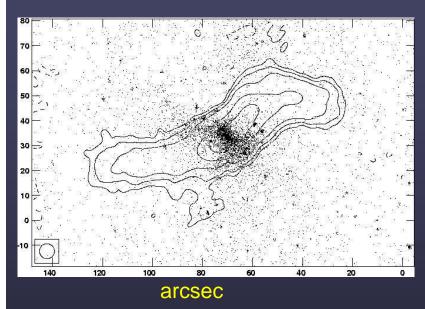


Multi-purpose plot

Contours: I,Q,U Pol Grey scale: P Pol sqrt (Q²+U²) - noise Line segments – P angle atan2(0.5*Q/U)

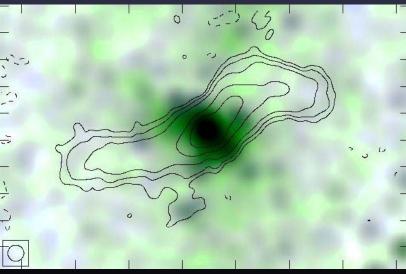


COMPARISON OF RADIO/X-RAY IMAGES



Contours of radio intensity at 5 GHz

Dots represent X-ray Intensity (photons) between 0.7 and 11.0 KeV

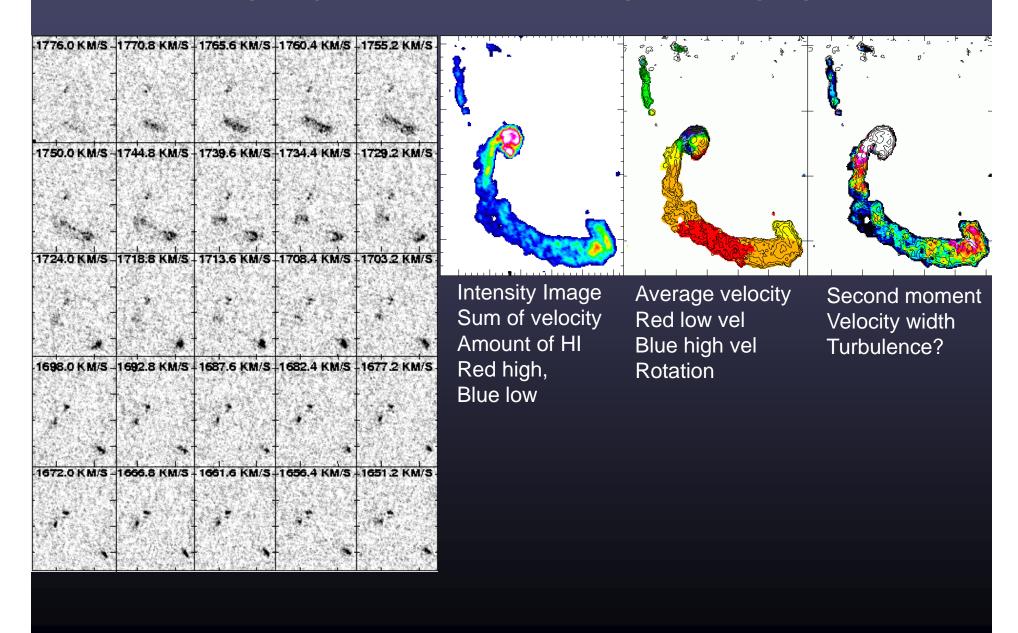


Contours of radio intensity at 5 GHz

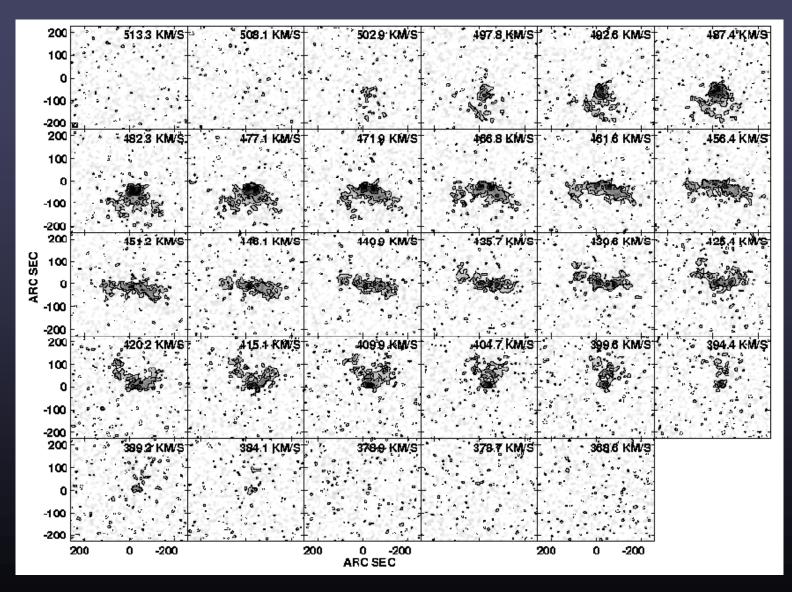
Color intensity represents X-ray intensity smoothed to radio resolution

Color represents hardness of X-ray (average weighted frequency)
Blue - soft (thermal)
Green - hard (non-thermal)

SPECTRAL LINE REPRESENTATIONS



Visualizing Spectral Line Data: Channel Images



Greyscale+contour representations of individual channel images

Visualizing Spectral Line Data: Channel Images

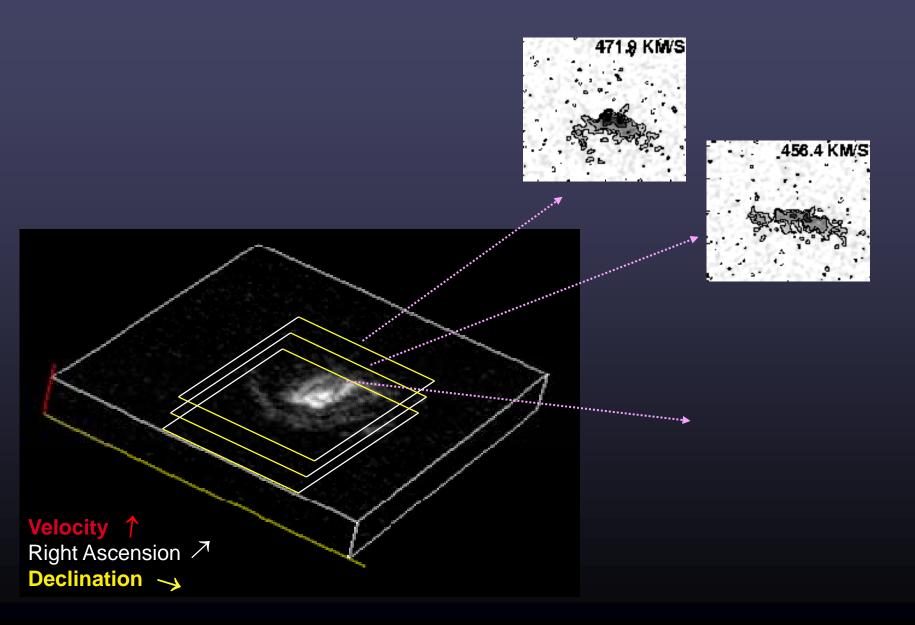


IMAGE REGISTRATION AND ACCURACY

 Separation Accuracy of Components on One Image due to residual phase errors, regardless of signal/noise:

Limited to 1% of resolution

Position errors of 1:10000 for wide fields, i.e. 0.1" over 1.4 GHz PB

Images at Different Frequencies:

Multi-frequency. Use same calibrator for all frequencies. Watch out at frequencies < 2 GHz when ionosphere can produce displacement. Minimize calibrator-target separation

• Images at Different Times (different configuration):

Use same calibrator for all observations. Daily troposphere changes can produce position changes up to 25% of the resolution.

Radio versus non-Radio Images:

Header-information of non-radio images often much less accurate than that for radio. For accuracy <1", often have to align using coincident objects.

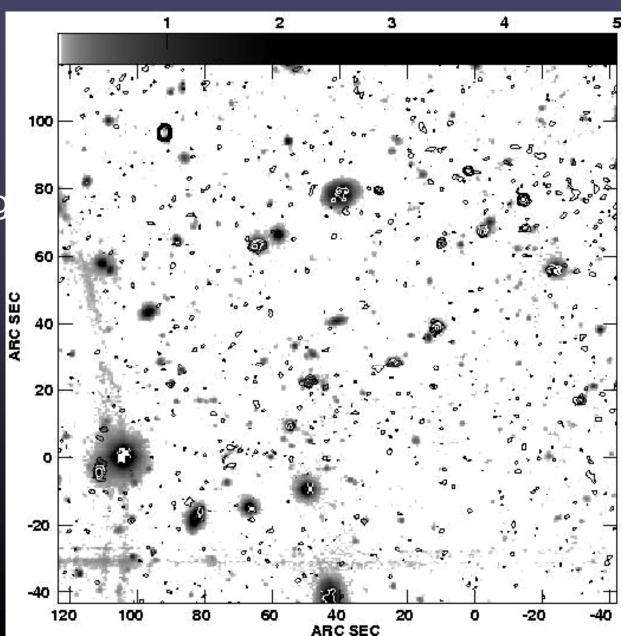
DEEP RADIO / OPTICAL COMPARISON

Grey-Scale:

Optical emission faintest is 26-mag

Contours:

Radio Emission faintest is 10 µJy



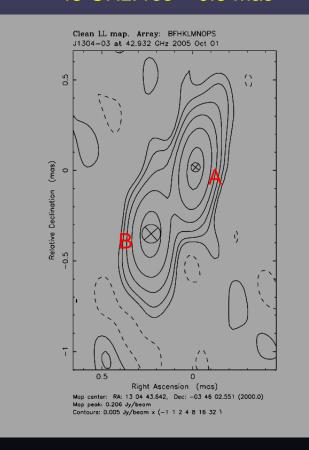
Radio Source Alignment at Different Frequencies

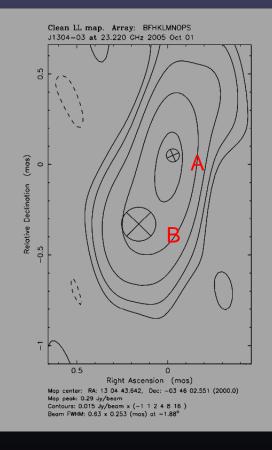
Self-calibration at each frequency aligns maximum at (0,0) point Frequency-dependent structure causes relative position of maximum to change Fitting of image with components can often lead to proper registration

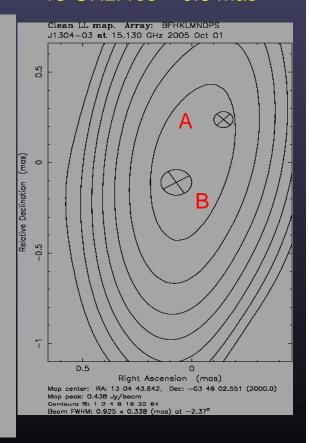
43 GHz: res = 0.3 mas

23 GHz: res = 0.6 mas

15 GHz: res = 0.8 mas







SUMMARY

- Analyze and display data in several ways
 Adjust resolution to illuminate desired interpretation, analysis
- Parameter fitting useful, but try to obtain error estimate
 Fitting in u-v plane, image plane
- Comparison of multi-plane images tricky (Polarization and Spectral Line)
 - Use different graphics packages, methods, analysis tools
- Registration of a field at different frequencies or wave-bands can be subtle.
 Often use ad hoc methods by aligning 'known' counterparts