# ERROR RECOGNITION



### Twelfth Synthesis Imaging Workshop 2010 June 8-15



### 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**

### Rats!!

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 Je

#### How can the problems be





### HIGH QUALITY IMAGE

### **Great!!**

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

#### What were defects?

Two antennas had 10% calibration errors, and one with a 5 deg error, plus a few outlier points.

#### This part of the lecture.

How to find the errors and remove them.





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





### **IMAGE DISPLAYS (1)**

												Piz	cel	va.	lue:	s													
				2	35				3	245		155			255					265					275				
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	0	0	0	0	1	1	0	0	0	0	0	0	0
283	0	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	5	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	t	2	4	8	12	17	22	23	24	22	27	77	136	87	19	2	0	0	0
271	0	0	0	0	0	0	0	0	1	1	2	- 4	8	15	21	28	35	36	37	33	43:	126	217	132	28	3	0	0	0
269	0	0	0	0	0	0	0	0	1	3	4	8	15	25	34	44	54	54	53	48	61:	173	288	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	78:	199	316	177	34	3	0	0	0
265	0	0	0	0	0	0	0	1	3	.7	14	24	40	60	77	97:	109	102	93	74	79:	191	28.9	15.5	29	3	0	0	0
263	0	0	0	0	0	0	1	2	- 5	11	22	37	58	86	108	130;	137	123:	105	79	73:	154	220	113	20	2	0	0	0
261	0	ö	0	0	0	1	1	3	8	17	33	54	81	116	139	156:	156	133:	107	75	61;	106	140	69	12	2	0	0	0
259	0	0	0	0	0	1	2	5	12	24	45	72	105	143:	162	170:	161	131	99	66	47	64	75	36	6	1	0	0	0
257	0	0	0	0	0	2	- 4	8	18	32	58	88:	124	160:	171	169:	152	118	86	55	36	36	36	16	з	1	0	0	0
255	0	0	0	0	1	2	7	16	27	42	70:	101:	135	162:	164:	1561	34:	001	71	44	27	20	16	7	1	0	0	0	0
253	0	0	0	0	1	4	15	34	43	51	77:	105:	133	150:	146:	1351	12	81	56	34	19	11	7	3	1	0	0	0	0
251	0	0	0	0	1	8	34	73	70	59	79:	100:	120	130	122	110	88	61	41	24	12	6	3	1	0	0	0	0	0
249	0	0	0	1	2	14	69	141	112	65	73	87:	100	106	96	83	64	43	27	14	7	3	1	1	0	0	0	0	0
247	0	0	0	1	3	23	121	238	167	69	62	69	77	81	70	58	42	26	16	8	3	1	0	0	0	0	0	0	0
245	0	0	0	1	3	34	180	338:	217	69	48	52	56	67	47	36	25	15	8	3	1	0	0	0	0	0	0	0	0
243	0	0	0	1	4	42	22.2	40.2:	24.2	65	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	44	22.9	398:	228	56	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	196	327	179	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	139	223	118	26	11	9	8	6	4	2	1	1	0	0	0	0	0	0	0	0	0	0	0
235	0	0	0	0	2	18	82	127	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	0	0	0	0	4	17	23	11	3	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
229	Ó	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	0	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 (2)



### 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 (3)**



### TV-based displays are most useful and interactive:

Grey-scale shows faint structure, but not good for high dynamic

ange and somewhat unbiased view of source



A pseudo contours





### DATA DISPLAYS(1)

#### List of u-v Data

Source=	J0121+11	Freq=	8.	43-	4858511	Sort-	= TB	1	RR	
Vis #	IAT	Ant	Su	$\mathbf{F}\mathbf{q}$	U(klam)	V(klam)	W(klam)	Amp	Phas	Wt
2191	0/22:35:08.23	2 5- 6	1	0	94220	23776	100371	0.614	-16	1.0000
3971	0/22:43:43.34	1 5-6	1	0	97659	24517	96844	0.508	-13	1.0000
6431	0/23:07:05.18	5 5- 6	1	0	106307	26661	86632	0.154	17	1.0000
6611	0/23:07:14.98	8 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	1 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	1 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	8 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	8 5-6	1	0	114156	29082	75098	0.266	134	1.0000
10861	0/23:39:02.08	8 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.



Often need precise times in order to flag the





### DATA DISPLAYS(2)









### DATA DISPLAYS(3)



#### 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.

lease check wei they are reason





### 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 RULE 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 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).

### ---Non-Data Problems:

Perfect data but unstable algorithms. Common but difficult





### 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.

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 should good data

(3C279 VLBA data at 43 GHz)





### **VISIBILITY AMPLITUDE RASTERS**



(Last two scans from previous slide)

Use AIPS task TVFLG, 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 – plotms (2)





Jansky





The University of New Mexico

### **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



Institute for Advanced Studies et los Alexas Nettosal Laboratory

#### 35 km 12 km 3 km baseline

## **RFI Excision**





RFI environment worse on short baselines

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

Wideband interference hard for automated routines

AIPS tasks FLGIT, FLAGR, and CASA flagdata, mode='rfi'





The University of New Mexico

# ERROR RECOGNITION IN THE IMAGE

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 observation over 10 hr. Images shown after editing, calibration and deconvolution.

no errors: max 3.24 Jy rms 0.11 mJy



6-fold symmetric pattern due to VLA "Y". Image has properties of dirty beam. 10% amp error for all antennas for 1 time period rms 2.0 mJy



### 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 (self-cal)



### 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 correlation of offsets



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.



New Mexico Tech

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?**



Residual sidelobes dominate the noise

#### Emission from





Regions within clean boxes appear "mottled"

# Background is thermal noise-dominated; no "bowls" around





The University of New Mexico

### FINDING HIDDEN BAD DATA

## Chandra Deechandra Deep Field South

#### Peak = 45 mJy, rms = 0.02 mJy



#### Source to NF in first Primary beam

Primary beam sidelobe

See Lectures Bhatnagar and Ott on Wide-field Imaging, and Perley on High Dynamic Range Imaging

Center of Field







### Fourier Transform Dirty Image



nows the u-v lata as gridded ust before imaging

agonal lines caused y structure in field

few odd points are ot 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.





### 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**



### Twelfth Synthesis Imaging Workshop 2010 June 8-15



### **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.





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



#### Smoothed

#### Robust=+1

Robust=-1



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)

(see non-imaging analysis)

Error estimates of parameters

Simple ad-hoc error estimates

stimates from fitting programs



New Mexico Tech CIENCE · ENGINEERING · RESEARCH · UNIVERSITY



### **IMAGE FITTING**

Component 2-Gaus	ssian	8				
Peak intensity	=	0.104 +/-	0.005	JY/BEAM		
Integral intens:	ity=	0.998 +/-	9,47	JANSKYS		
X-position	=	255.98	5 +/-	0.0029	pixels	
Y-position	=	257.03	3 +/-	0.0032	pixels	
Major ax		19.99	+/-	0.02	pixels	
Minor ax		9.98	+/-	0.03	pixels	
Pos ang		135.3	+/-	0.1	deg	

e	Component 1-Gan	issian				
	Peak intensity	= (	.300 +/-	0.005	JY/BEA	M
1	Integral intens	sity= (	.302 +/-	800.0	JANSKY	S
	X-position	=	270.991	+/-	0.001	pixels
	Y-position		267.018	+/-	0.001	pixels
	Major ax		0.53	+/-	0.01	pixels
	Minor ax		0.00	+/-	0.05	pixels
-5 -	Pos ang		21.6	+/-	1.1	deg
						-

Component 3-Gauss	ian	8			
Peak intensity	=	0.393 +/-	0.004	JY/BEAD	4
Integral intensit;	y= -	0.403 +/-	800.0	JANSKY	5
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
Pos ang		3.6	+/-	0.2	deg

- AIPS task: JMFIT
- Casa tool
- imfit





### (U-V) DATA FITTING



#### Time

#### DIFMAP has good u-v fitting algorithm

Fit model directly to (u-v) data

#### Contour display of image

Ellipses show true component

size. (super-resolution?)



I, "Non-image Data Analysis





### **COMPONENT ERROR ESTIMATES**

P/s = signal/noise = S

- **P** = Component Peak Flux Density
- **s** = Image rms noise
- **B** = Synthesized beam size
- *qi* = Component image size

DP = Peak error = s
DX = Position error = B / 2S
Dqi = Component image size error = B / 2S
qt = True component size = (qi2 - B2)1/2
Dqt = Minimum component size = B / S1/2

#### eg. S–100 means can determine size of B/10





### **Comparison and Combination of Images of Many Types**



Optical in blue-white

Frame size is 60' x 40'





2005 NSF Senior Review



### 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



arcsec



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





Intensity Image Sum of velocity Amount of HI Red high, Blue low Average velocity Red low vel Blue high vel Rotation Second moment Velocity width Turbulence?







### **Visualizing Spectral Line Data: Channel Images**







New Mexico Tech SCIENCE ENTITIES TO TECHTORING CONTOUR REPRESENTATIONS OF INDIVIDUAL CHARGE IN THE

### **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.

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 🖅 Jy





### Radio Source Alignment at Different Frequencies

23 GHz: res = 0.6 mas

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



ENCE . ENGINEERING . RESEARCH . UNIVERSITY

## Clean LL map. Array: BFHKLMNOPS J1304-03 at 23.220 GHz 2005 Oct 01 B Map center: RA: 13 04 43.642, Dec: -03 46 02.551 (2000.0) Map peak: 0.29 Jy/beam Contours: 0.015 Jy/beam x (-1 1 2 4 8 16 ) am FWHM: 0.63 x 0.253 (mas) at -1.88

#### 15 GHz: res = 0.8 mas



The University of New Mexico

nstitute for Advanced Studie

### IMAGE ANALYSIS: 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



