Multi-frequency polarimetry of a complete sample of extragalactic radio sources

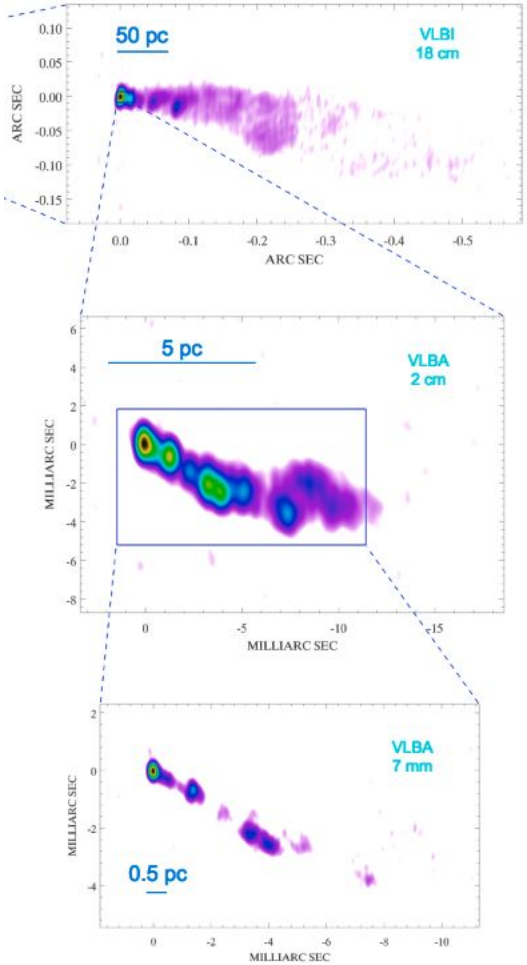
- Products and provenance information -

V. Galluzzi
INAF- Osservatorio Astronomico di Trieste

ESCAPE WP4 Provenance Workshop (7-8 September 2020)

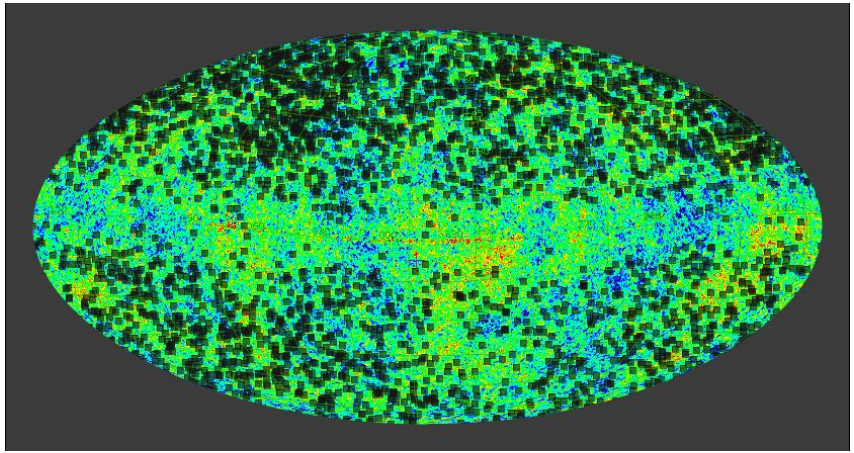
Scientific context

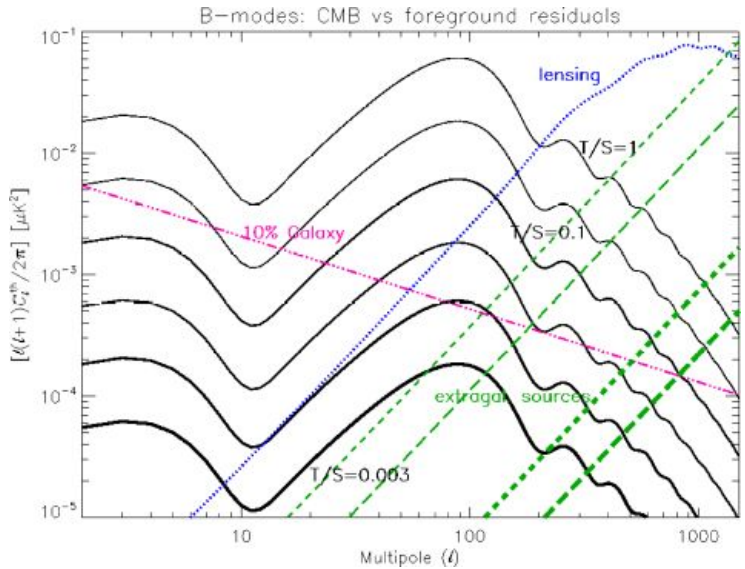
- The high-frequency (> 20 GHz), bright flux density (> 200 mJy) extragalactic radio sources (ERSs) population is dominated by blazars.
- Multi-frequency (and multi-epoch) polarimetry is invaluable to study magnetic fields and plasma in the inner and unresolved regions of their relativistic jets



Scientific context

- ERSs constitute an important foreground for CMB, such studies are also crucial for Cosmology, e.g. in the search for primordial B-modes associated to inflation in the Early Universe





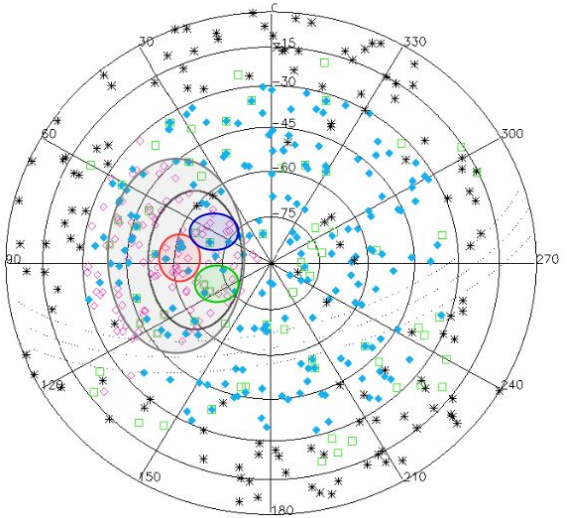
Observations

ATCA observations (project id: C2922, PI: M. Massardi)

Epoch	Allocated time	Frequencies	# objects	region
Sep. 2014	21 h	[5.5;38] GHz	53	b < - 75°
Mar.-Apr. 2016	26 h	[5.5;38] GHz	51	-75° ≤ b < -65°
	14 h	2.1 GHz	104	b < - 65°
July 2016*	5 h	33-35 GHz	35	b < -75°

- Faint PACO
- Spectrally-selected PACO
- ATCA calibrators
- Bright PACO

ALMA Cycle 3 observations
(project id: 2015.1.01522.S , PI: V. Galluzzi)

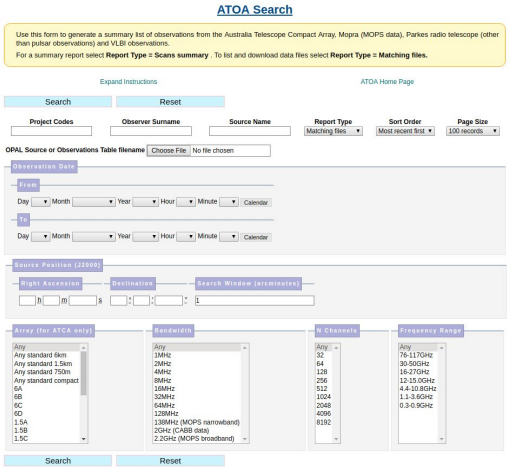


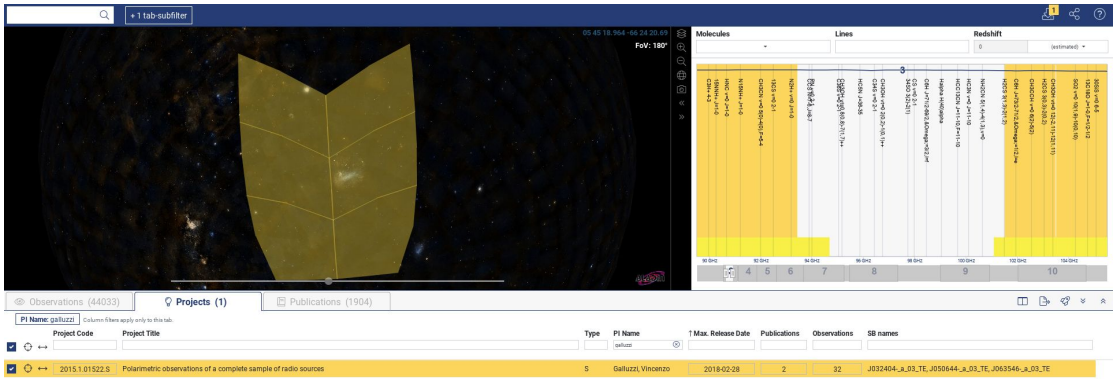
SG – Epoch	Time allocated	Frequencies	# objects	Region (J2000)
1 – 24/08/2016	3.0 h	97.5 GHz (4 spw)	14	(05:06:44 , -61:09:41)
2 – 22/09/2016	3.5 h	97.5 GHz (4 spw)	9	(06:35:46 , -75:16:16)
3 – 27/09/2016	3.5 h	97.5 GHz (4 spw)	9	(03:24:04 , -73:20:47)

Data products

- Level 0:
  a. raw visibilities (ATCA observations available through the Australia Telescope Online Archive, ATOA, in the MIRIAD RPFITS format; ALMA observations available through the ALMA Science Archive, ASA, in the ASDM format);

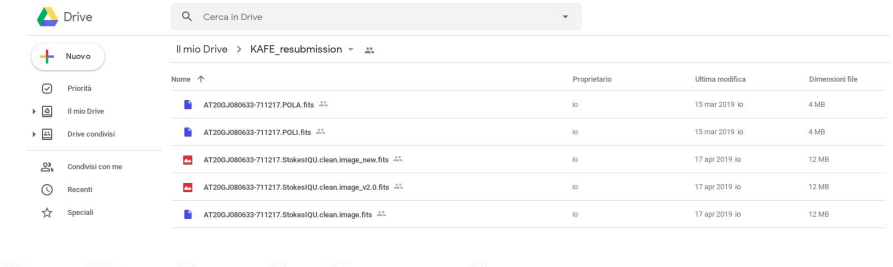
- Level 2:
  b. calibrated visibilities (available on request for ATCA observations in the MIRIAD UVFITS format, while for ALMA observations they can be generated from raw observations through the “script for PI” in the MS format);

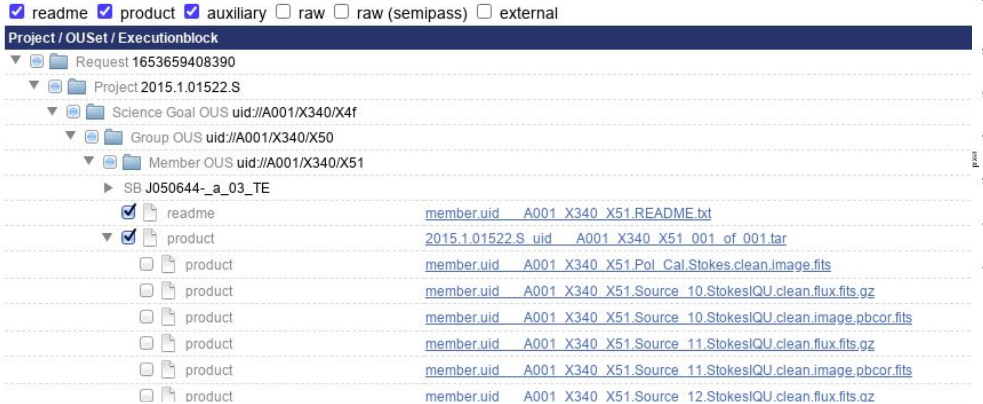


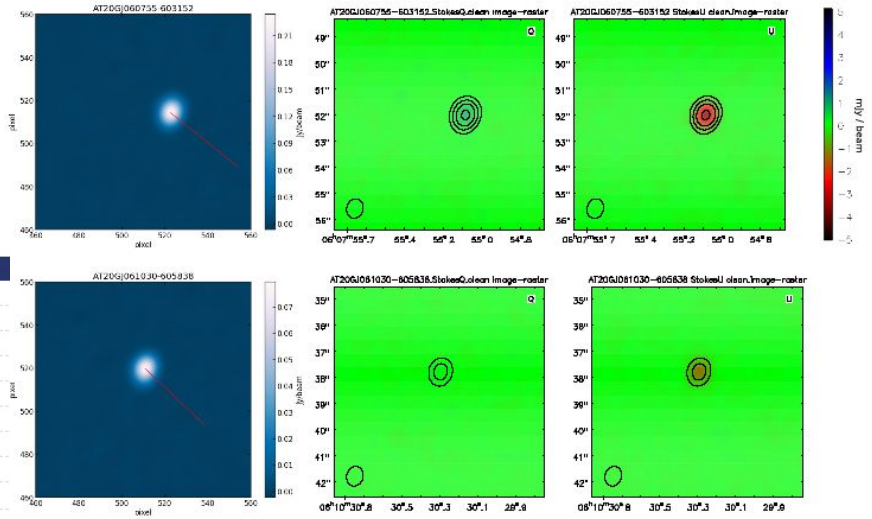


Data products

- Level 3:
  c. Stokes' I, Q and U maps for each source in the sample (FITS files available on request; for ALMA observations, thumbnails are published as supplementary material of a dedicated paper, while IQU CASA image cubes can be downloaded from ASA);







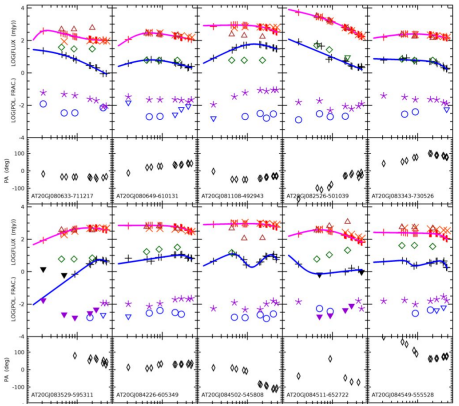
Data products

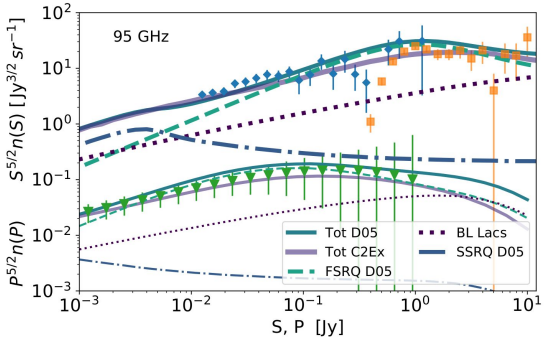
- Level 4:
  - d. various catalogues reporting full-Stokes' flux densities and parameters resulting from SED fitting (published as supplementary material of papers or available through VizieR);
  - e. various plots and tables reporting statistical estimators (e.g. mean and quartiles) of spectral indexes, polarization fractions and rotation measures (the code used for generating them has been mainly written in IDL and is available on request);
  - f. normalized source counts in total intensity and polarization at 20 and 95 GHz (tables are in the papers; the code is publicly available in a gitlab repository);
  - g. contamination forecasts due to extragalactic radio sources on CMB angular power spectra at 20 and 100 GHz (plots are reported in dedicated papers; the code is publicly available in the same gitlab repository).

J/MNRAS/465/4085/catalog — PACO radio sources complete sample polarimetry (Galluzzi+ 2017)
Post annotation — Full catalogue of I and P flux density, polarisation fractions, angles, and fit parameters *(53 rows)*

start AladinLite | plot the output | query using TAP/SQL

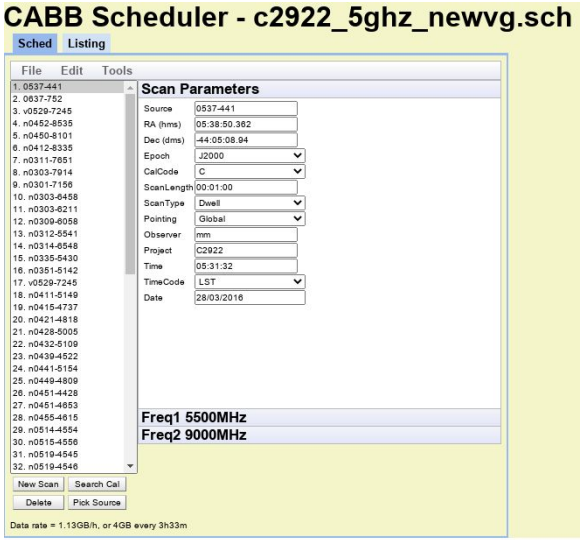
Full	_RAJ2000 "h:m:s"	_DEJ2000 "d:m:s"	AT20G	RAJ2000 deg	DEJ2000 deg	Flag	l_	m5500 %	l_	m9000 %	l_	m18000 %	l_	m24000 %	l_	m33000 %	l_	m38000 %
1	03 24 04.02913	-73 20 47.0033	J032404-732047	051.016788	-73.3463898			0.29		0.59		0.96		1.84		2.31		2.59
2	03 32 43.11027	-72 49 04.8853	J033243-724904	053.179626	-72.8180237			1.41		2.29		2.04		2.18		2.72		3.22
3	03 40 28.17987	-67 03 16.8199	J034028-670316	055.117416	-67.0546722			0.43		0.91		0.52	<	0.31	<	0.07	<	0.53
4	03 55 47.85993	-66 45 33.7007	J035547-664533	058.949416	-66.7593613			1.40				1.64		0.92		0.33		0.28
5	04 08 20.54907	-65 45 08.2123	J040820-654508	062.085621	-65.7522812			0.12		0.62		3.05		4.48		5.58		6.01
6	04 08 48.75000	-75 07 20.1122	J040848-750720	062.203125	-75.1222534	a		1.35		4.40		4.32		3.06		2.28		2.80
7	04 25 06.90133	-66 46 50.1103	J042506-664650	066.278756	-66.7805862					0.33		1.14		1.53		2.60		2.18
8	04 28 10.86927	-64 38 23.6134	J042810-643823	067.045289	-64.6398926			3.23				3.16		5.51		5.11		3.42
9	04 33 34.05960	-60 30 14.1998	J043334-603014	068.391915	-60.5039444			2.13		1.75		2.39		1.77		0.47		0.49
10	04 40 47.91867	-69 52 17.5030	J044047-695217	070.199661	-69.8715286			0.40		0.51		1.95		2.13		3.21		3.88
11	05 04 01.68993	-60 49 52.9918	J050401-604953	076.007041	-60.8313866					1.39		2.52		3.19		4.44		4.72
12	05 05 43.52967	-79 14 39.5930	J050543-791439	076.431374	-79.2443314		<	0.07		0.62		0.89		0.80		0.81		1.48
13	05 06 43.98913	-61 09 40.9846	J050644-610941	076.683288	-61.1613846			4.05		3.94		5.46		5.54		5.40		5.31
14	05 07 54.69927	-61 04 42.9942	J050754-610442	076.977914	-61.0786095			1.64		1.52		4.22		5.14		6.64		6.90
15	05 16 37.95960	-72 37 07.8074	J051637-723707	079.158165	-72.6188354			3.75		2.59		1.35		1.30		2.01		2.28
16	05 16 44.88093	-62 07 06.0910	J051644-620706	079.187004	-62.1183586			6.73		6.61		6.54		6.83		7.55		7.83
17	05 22 34.42547	-61 07 57.1361	J052234-610757	080.643439	-61.1325378			3.45		2.29		1.45		1.64		1.69		2.40

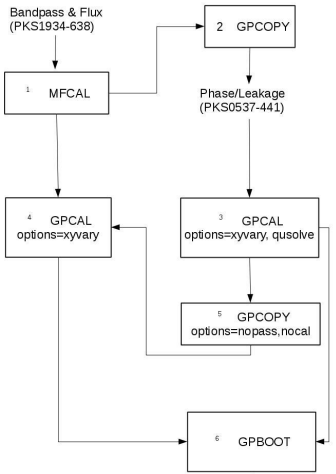




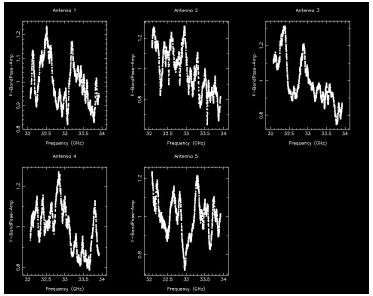
Ancillary products

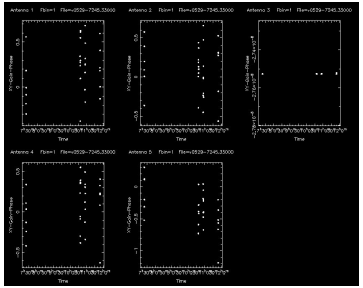
h. ATCA observation schedules (available on request or directly accessible to all registered ATCA observers).
i. calibration scripts (exemplar MIRIAD scripts for ATCA data are also published in the PhD thesis) and imaging scripts;
j. diagnostics plots produced during the execution of scripts (image files in png or jpeg);
k. observations and processing log files (human-readable .txt files).






```
frequency="18000"
BandLea_cal="n0538-4405"
PhaFlu_cal="v0529-7245"

mfcal vis=$BandLea_cal.$frequency refant=3 interval=0.1
gpcal vis=$BandLea_cal.$frequency refant=3 interval=0.1 nfbin=2 options="xyvary,qusolve"
gpcopy vis=$BandLea_cal.$frequency out=$PhaFlu_cal.$frequency options="nocal"
gpcal vis=$PhaFlu_cal.$frequency refant=3 interval=0.1 nfbin=2 flux="1.229,4.871E-03,-1.090E-02,2.099E-03" options="xyvary,qusolve,nopol"
#mfboot vis="$PhaFlu_cal.$frequency,$PhaFlu_cal.$frequency" select="source($PhaFlu_cal)" flux="1.2622,18.0,0.0009" device=/xs

while read src; do
  if [ $src  != $PhaFlu_cal ] && [ $src != $BandLea_cal ]; then
   gpcopy vis=$PhaFlu_cal.$frequency out=$src.$frequency
   fi
   uvspec vis=$src.$frequency options=nobase stokes="i" axis="ch,amp" device="Plots@$frequency/fluxesI_$src@$frequency"/png
   uvspec vis=$src.$frequency options=nobase stokes="q" axis="ch,amp" device="Plots@$frequency/fluxesQ_$src@$frequency"/png
   uvspec vis=$src.$frequency options=nobase stokes="u" axis="ch,amp" device="Plots@$frequency/fluxesU_$src@$frequency"/png
   uvspec vis=$src.$frequency options=nobase stokes="v" axis="ch,amp" device="Plots@$frequency/fluxesV_$src@$frequency"/png
   uvplt  vis=$src.$frequency stokes=i options=nobase axis="uvdist,amp" device="Plots@$frequency/AmpVsUVdist_$src@$frequency"/png
   uvflux vis=$src.$frequency stokes="i,q,u,v"
done < IIIslotIIIGT_18e24GHz_indexlist_g0529.log

####################Diagnostica_1#####################################################
uvplt vis=$PhaFlu_cal.$frequency stokes="i,q,u,v" axis="real,imag" options="nobase,nofqav" device="Plots@$frequency/Diagnostics/scatter_plot_flux_cal@$frequency"/png
######################################################################################

####################Diagnostica_2#####################################################
uvplt vis=$PhaFlu_cal.$frequency axis="time,parang"  options=nobase device="Plots@$frequency/Diagnostics/PhLcal_parang@$frequency"/png
varplt vis=$PhaFlu_cal.$frequency yaxis=chi device="Plots@$frequency/Diagnostics/Phase_chi@$frequency"/png
uvflux vis=$PhaFlu_cal.$frequency stokes="i,q,u,v"
uvplt vis=$PhaFlu_cal.$frequency stokes="i" axis="real,imag" options="nobase" device="Plots@$frequency/Diagnostics/scatter_plot_I_Phase_cal@$frequency"/png
uvplt vis=$PhaFlu_cal.$frequency stokes="q,u,v" axis="real,imag" options="nobase" device="Plots@$frequency/Diagnostics/scatter_plot_QUV_Phase_cal@$frequency"/png
gpplt vis=$PhaFlu_cal.$frequency yaxis=phase options=xygains device="Plots@$frequency/Diagnostics/phase_mod_xygains@$frequency"/png
gpplt vis=$PhaFlu_cal.$frequency yaxis=amp options=bandpass device="Plots@$frequency/Diagnostics/bandpass_phase_cal@$frequency"/png
gpplt vis=$PhaFlu_cal.$frequency yaxis=phase options=gains device="Plots@$frequency/Diagnostics/phase_cal_gains@$frequency"/png
```

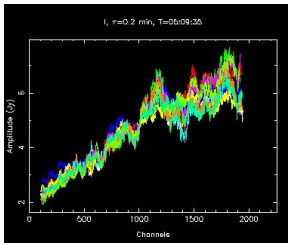


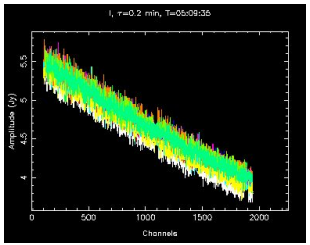


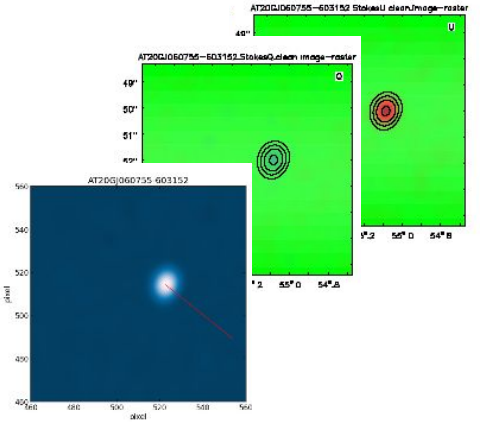
Provenance information: motivations

- The main reason is represented by the fact that radio polarimetry involves more calibration steps and presents more potential sources of systematics with respect to standard observations in total intensity. Thus, additional diagnostics is usually addressed to provide indicators about data quality (e.g. statistics of leakage terms solutions and Stokes' parameters flux densities)
- The variety of data products ranging from observations to more advanced ones (e.g., catalogues and source counts) particularly require traceability information.

```
gpcal: Revision 1.22, 2016/05/01 22:31:05 UTC

Applying bandpass corrections to 0537-441.5500
Number of antennae: 6
Reading the data ...
Number of solution intervals: 30
Total visibilities read: 450
Number visibilities accepted: 300
Iter= 1, Amplit/Phase Solution Error:   0.092
Iter= 1, Polarisation Solution Error:   0.022
Iter= 1, Overall Solution Error:        0.575
Iter= 2, Amplit/Phase Solution Error:   0.012
Iter= 2, Polarisation Solution Error:   0.000
Iter= 2, Overall Solution Error:        0.008
Iter= 3, Amplit/Phase Solution Error:   0.000
Iter= 3, Polarisation Solution Error:   0.000
Iter= 3, Overall Solution Error:        0.000
I flux density:   4.3039
Percent Q:  -1.084
Percent U:  -0.588
Leakage terms:
 Ant 1:Dx,Dy = ( 0.00042,-0.00371),( 0.00095,-0.00404)
 Ant 2:Dx,Dy = ( 0.00249, 0.00323),(-0.00131, 0.00252)
 Ant 3:Dx,Dy = ( 0.00457,-0.00583),(-0.00700,-0.00714)
 Ant 4:Dx,Dy = (-0.00842,-0.00802),( 0.00948,-0.00940)
 Ant 5:Dx,Dy = ( 0.00188, 0.01660),(-0.00118, 0.01579)
```





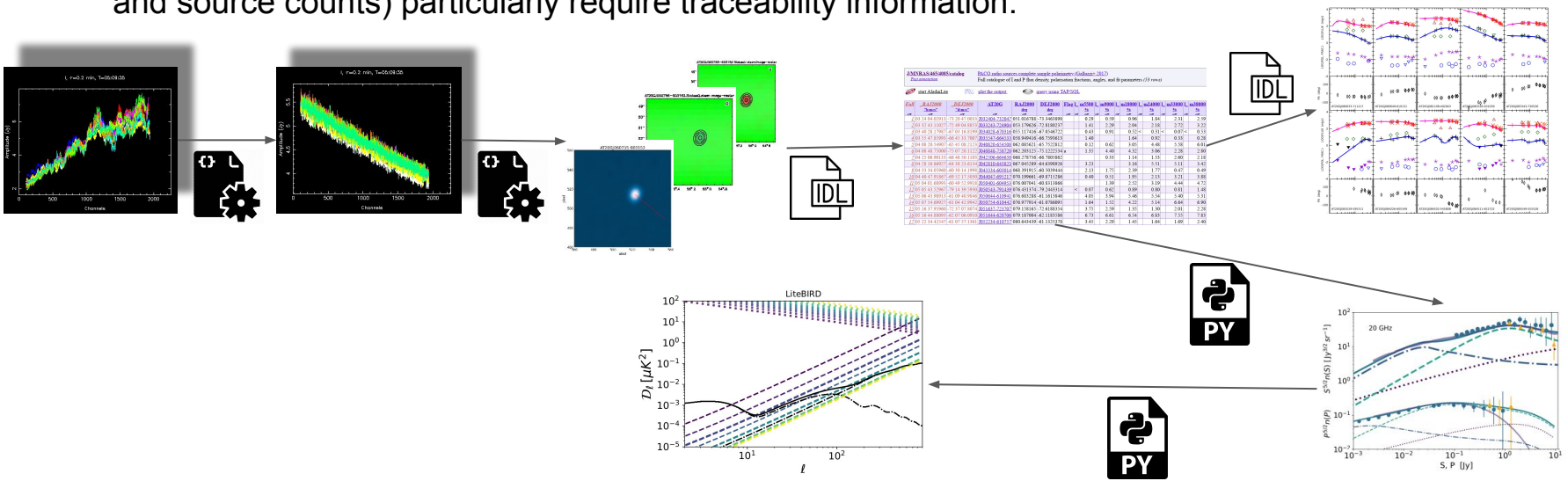


J/MNRAS/465/4085/catalog — PACO radio sources complete sample polarimetry (Galluzzi+ 2017)
Zoer annotation — Full catalogue of I and P flux density, polarisation fractions, angles, and fit parameters (53 rows)
start AladinLite — plot the output — query using TAP/SQL

Full	RAJ2000 "h:m:s"	DEJ2000 "d:m:s"	AT20G	RAJ2000 deg	DEJ2000 deg	Flag	l_m5500	l_m9000	l_m18000	l_m24000	l_m33000	l_m38000
							%	%	%	%	%	%
1	03 24 04.02913	-73 20 47.0033	J032404.-732047	051.016788	-73.3463898		0.29	0.59	0.96	1.84	2.31	2.59
2	03 32 43.11027	-72 49 04.8833	J033243.-724904	053.179626	-72.8180237		1.41	2.29	2.04	2.18	2.72	3.23
3	03 40 28.17987	-67 03 16.9199	J034028.-670316	055.117416	-67.0546722		0.43	0.91	0.52	0.31	0.07	0.53
4	03 55 47.85993	-66 45 33.7007	J035547.-664533	058.949416	-66.7593613		1.40		1.64	0.92	0.33	0.28
5	04 08 20.54907	-65 45 08.2123	J040820.-654508	062.085621	-65.7522812		0.12	0.62	3.05	4.48	5.58	6.01
6	04 08 48.75000	-75 07 20.1123	J040848.-750720	062.203125	-75.1222534	a	1.35	4.40	4.32	3.06	2.28	2.80
7	04 25 06.90133	-66 46 50.1103	J042506.-664650	066.278756	-66.7805862			0.33	1.14	1.53	2.60	2.18
8	04 28 10.86927	-64 38 23.6134	J042810.-643823	067.045289	-64.6398926		3.23		3.16	5.51	5.11	3.42
9	04 33 34.05960	-60 30 14.1998	J043334.-603014	068.391915	-60.5039444		2.13	1.75	2.39	1.77	0.47	0.49
10	04 40 47.91867	-69 52 17.5030	J044047.-695217	070.199661	-69.8715286		0.40	0.51	1.95	2.13	3.23	3.88
11	05 04 01.68993	-60 49 52.9918	J050401.-604953	076.007041	-60.8313866			1.39	2.52	3.19	4.44	4.72
12	05 05 43.52967	-79 14 39.5930	J050543.-791439	076.431374	-79.2443314	<	0.07	0.62	0.89	0.80	0.81	1.48
13	05 06 43.90913	-61 09 40.9846	J050644.-610941	076.683288	-61.1613846		4.05	3.94	5.46	5.54	5.40	5.31
14	05 07 54.69927	-61 04 42.9942	J050758.-610442	076.977914	-61.0786095		1.64	1.52	4.22	5.14	6.64	6.90
15	05 16 37.95960	-72 37 07.8074	J051637.-723707	079.158165	-72.6188354		3.75	2.59	1.35	1.30	2.01	2.28
16	05 16 44.88093	-62 07 06.0910	J051644.-620706	079.187004	-62.1183586		6.73	6.61	6.54	6.83	7.55	7.83
17	05 22 34.42547	-61 07 57.1361	J052234.-610757	080.643439	-61.1325378		3.45	2.29	1.45	1.64	1.69	2.40

Provenance information: motivations

- The main reason is represented by the fact that radio polarimetry involves more calibration steps and presents more potential sources of systematics with respect to standard observations in total intensity. Thus, additional diagnostics is usually addressed to provide indicators about data quality (e.g. statistics of leakage terms solutions and Stokes' parameters flux densities)
- The variety of data products ranging from observations to more advanced ones (e.g., catalogues and source counts) particularly require traceability information.



Provenance information: goals

1. ensure quality and reliability
2. keep the traceability of the data products
3. provide more information to the final user
4. reproduce or reprocess the data
5. debug a pipeline
6. find contact information
7. acknowledge people

Provenance information: sources

1. ensure quality and reliability –
2. keep the traceability of the data products ✓
3. provide more information to the final user ✓
4. reproduce or reprocess the data –
5. debug a pipeline ✗
6. find contact information ✓
7. acknowledge people ✓


Multifrequency polarimetry of a complete sample of PACO radio sources

Characterization of polarimetric and total intensity behaviour of a complete sample of PACO radio sources in the radio bands

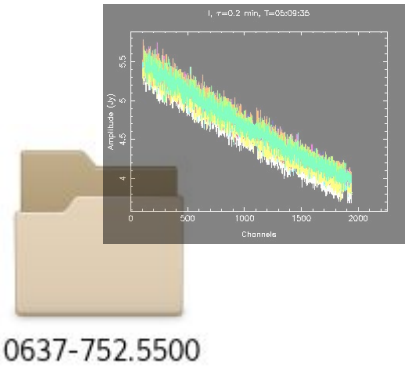
ALMA Band 3 polarimetric follow-up of a complete sample of faint PACO sources

<u>Published papers</u>

- human readable only;
- high level provenance information;
- the most complete source of provenance information.

Provenance information: sources

1. ensure quality and reliability –
2. keep the traceability of the data products –
3. provide more information to the final user ✔
4. reproduce or reprocess the data –
5. debug a pipeline ✘
6. find contact information ✘
7. acknowledge people ✘



Data products

- human and machine readable info;
- traceability information may be really fine-grained;
- info typically limited to the considered data product.

```
HISTORY MFCAL: MfCal: $Revision: 1.20 $, $Date: 2017/05/08 02:31:10 $
HISTORY MFCAL: Executed on: 2019-12-13T16:50:41.0
HISTORY MFCAL: Command line inputs follow:
HISTORY MFCAL:   vis=0537-441.5500
HISTORY MFCAL:   flux=3.5155,4.0,0.0796
HISTORY MFCAL:   refant=2
HISTORY MFCAL:   interval=0.1
HISTORY GPCOPY: Miriad gpcopy: Revision 1.13, 2015/10/29 01:31:39 UTC
HISTORY GPCOPY: Executed on: 2019-12-13T16:50:47.0
HISTORY GPCOPY: Command line inputs follow:
HISTORY GPCOPY:   vis=j0529-7245.5500
HISTORY GPCOPY:   out=0537-441.5500
HISTORY GPCAL: gpcal: Revision 1.22, 2016/05/01 22:31:05 UTC
HISTORY GPCAL: Executed on: 2019-12-13T16:50:47.0
HISTORY GPCAL: Command line inputs follow:
HISTORY GPCAL:   vis=0537-441.5500
HISTORY GPCAL:   refant=2
HISTORY GPCAL:   interval=0.1
HISTORY GPCAL:   nfbin=2
HISTORY GPCAL:   options=xyvary,nopol
```

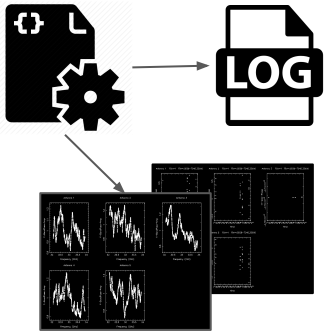
some output from GPCAL:

```
HISTORY GPCAL: I flux density:   3.6266
HISTORY GPCAL: I flux density:   3.6031
HISTORY GPCAL: I flux density:   3.6521
```

E.g., in the case of visibilities processed with MIRIAD, HISTORY keywords capture all the runned tasks

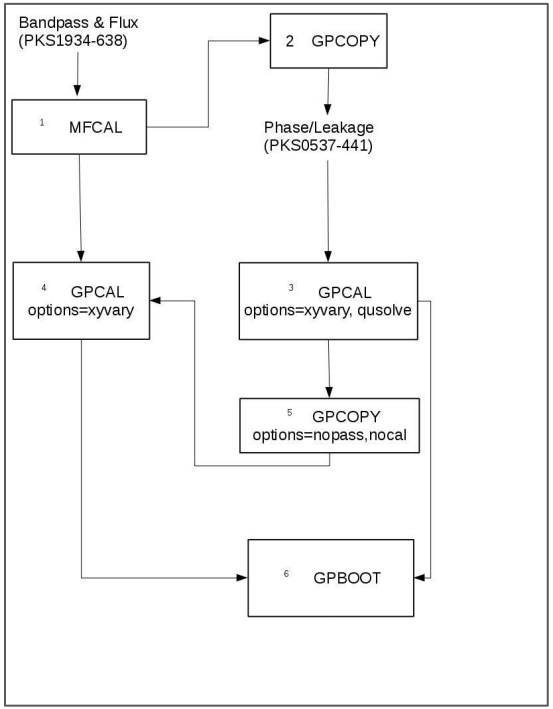
Provenance information: sources

1. ensure quality and reliability ✔
2. keep the traceability of the data products ✔
3. provide more information to the final user ✔
4. reproduce or reprocess the data ✔
5. debug a pipeline ✔
6. find contact information ✘
7. acknowledge people ✘



Ancillary products

- human readable info;
- traceability and quality metrics information may be really accurate;
- info typically limited to the data they are associated to.



Example of a calibration schema in MIRIAD adopted by the project

Conclusions

- Given the scientific scope of this (and of similar) project(s), the criticalities in data calibration and the variety of data products typically involved, the data discovery of such datasets could be significantly bolstered by showing to archive miners a clear tree diagram showing the products and a bunch of quality metrics metadata.
- As calibration, imaging and flux density extraction recipes are far from being definitive, a significant fraction of archive miners would be able to re-process data, hence the exploitation of the Provenance DM at an higher granularity level may be considered.