Radio Telescope Data Center
Smithsonian Astrophysical Observatory

SMA Data Archiving and Reduction

Holly Thomas
SMA Advisory Comittee
18 July 2018

https://www.cfa.harvard.edu/rtdc/

The RTDC

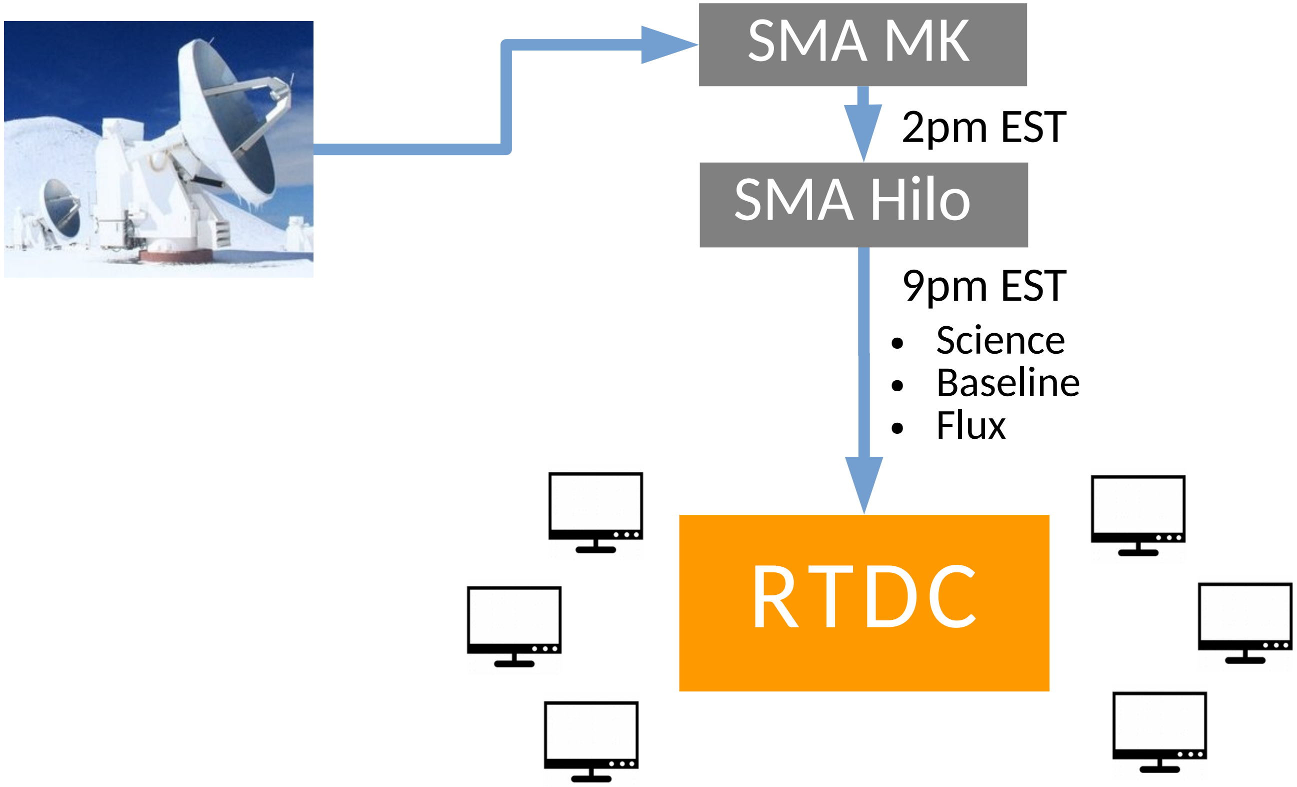
What we do

- Archive data from the SMA, CfA Millimeter-wave Telescope, and Antarctic Submillimeter Telescope & Remote Observatory (AST/RO).
- Supply hardware and software for the reduction of radio interferometric data (ALMA, VLA, VLBI, SMA).

For the SMA

- Store a copy of SMA data archive
- Maintain software required for SMA data processing
- Two web-based archives for external users
- Provide SMA data reduction information

SMA Data Flow



SMA MK

2pm EST

SMA Hilo

9pm EST

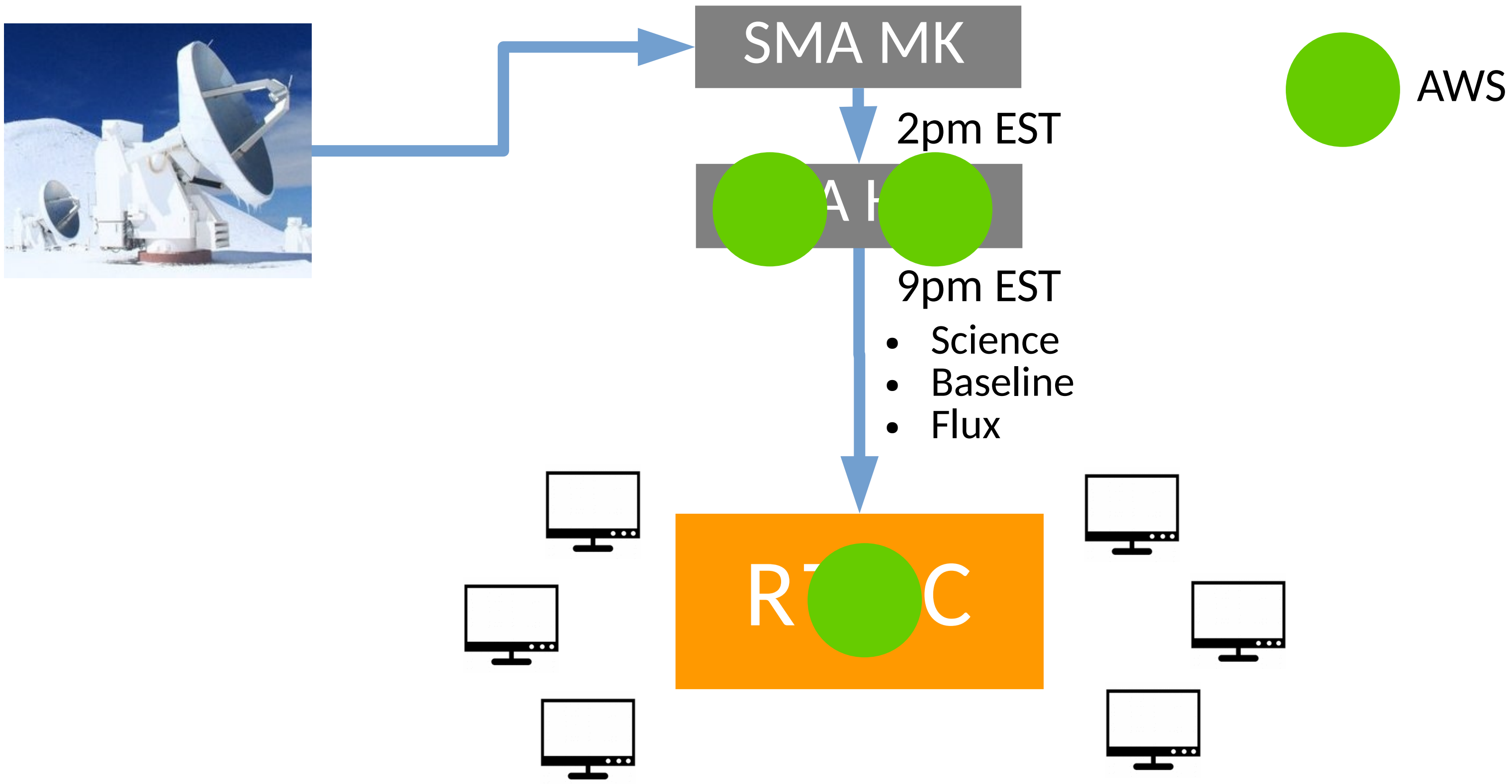
- Science
- Baseline
- Flux

RTDC

At the RTDC

- All science data
- Past 5 years of flux and baseline data

SMA Data Flow



SMA MK

2pm EST

A H

9pm EST

- Science
- Baseline
- Flux

RT C

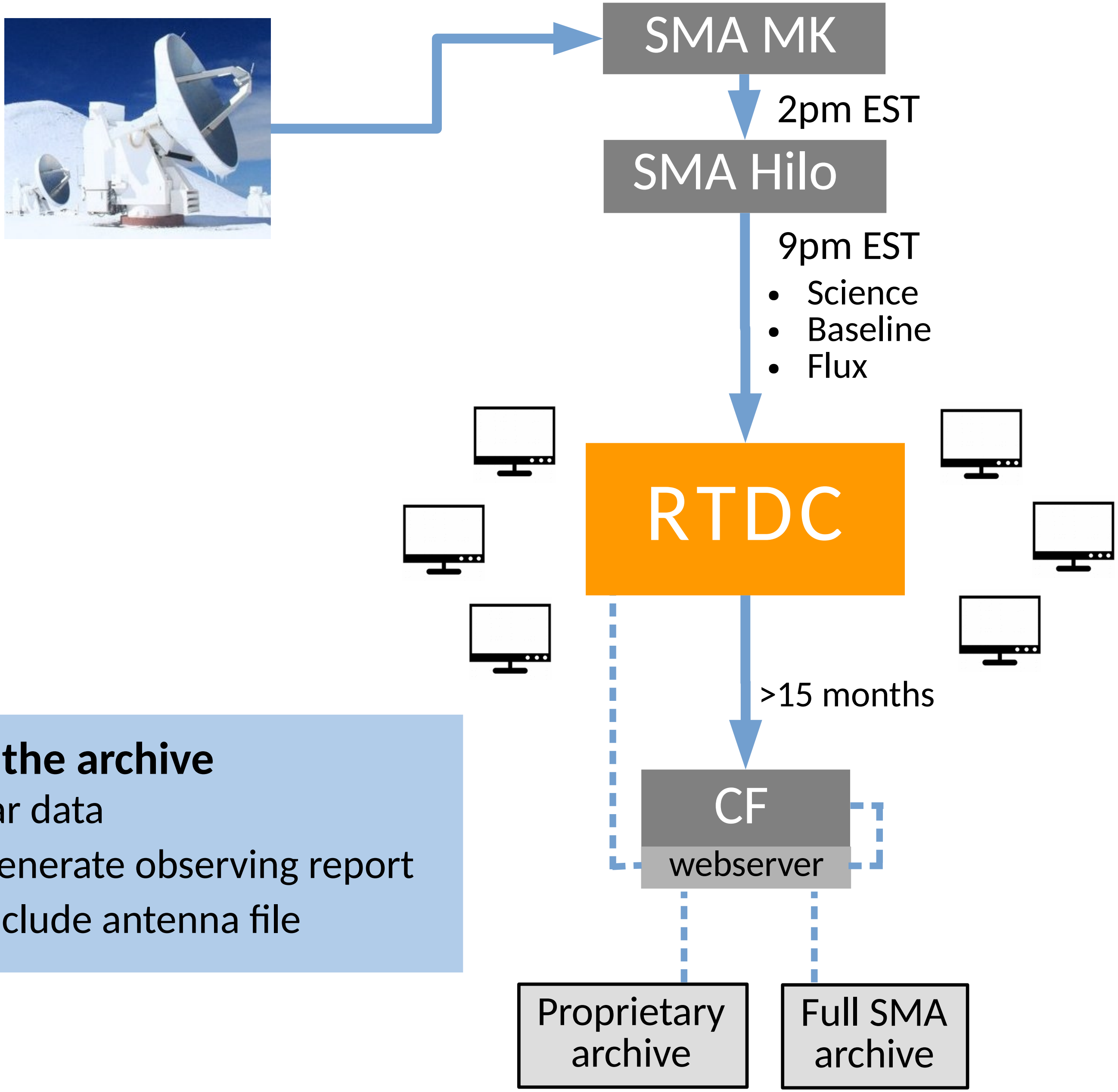
AWS

At the RTDC

- All science data
- Past 5 years of flux and baseline data

● SMA data

SMA Data Flow



SMA MK

2pm EST

SMA Hilo

9pm EST

- Science
- Baseline
- Flux

RTDC

>15 months

CF

webserver

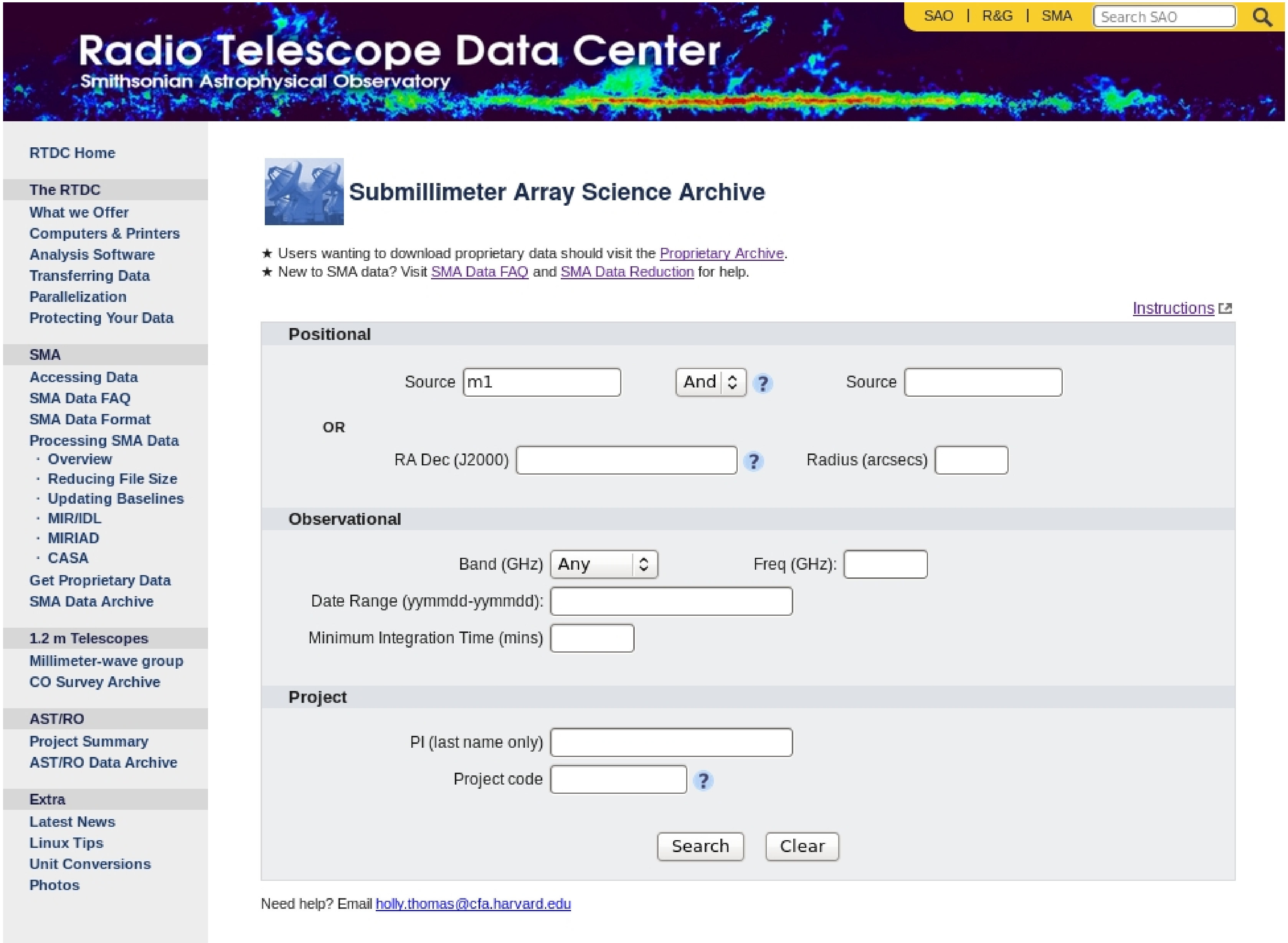
Proprietary archive

Full SMA archive

For the archive

- Tar data
- Generate observing report
- Include antenna file

SMA Archives


SAO | R&G | SMA
Search SAO
Radio Telescope Data Center
Smithsonian Astrophysical Observatory
RTDC Home
The RTDC
What we Offer
Computers & Printers
Analysis Software
Transferring Data
Parallelization
Protecting Your Data
SMA
Accessing Data
SMA Data FAQ
SMA Data Format
Processing SMA Data
· Overview
· Reducing File Size
· Updating Baselines
· MIR/IDL
· MIRIAD
· CASA
Get Proprietary Data
SMA Data Archive
1.2 m Telescopes
Millimeter-wave group
CO Survey Archive
AST/RO
Project Summary
AST/RO Data Archive
Extra
Latest News
Linux Tips
Unit Conversions
Photos
Instructions
Search
Clear

Submillimeter Array Science Archive

★ Users wanting to download proprietary data should visit the Proprietary Archive.
★ New to SMA data? Visit SMA Data FAQ and SMA Data Reduction for help.

Positional

Source: m1 — And — Source:

OR

RA Dec (J2000): — Radius (arcsecs):

Observational

Band (GHz): Any — Freq (GHz):

Date Range (yymmdd-yymmdd):

Minimum Integration Time (mins):

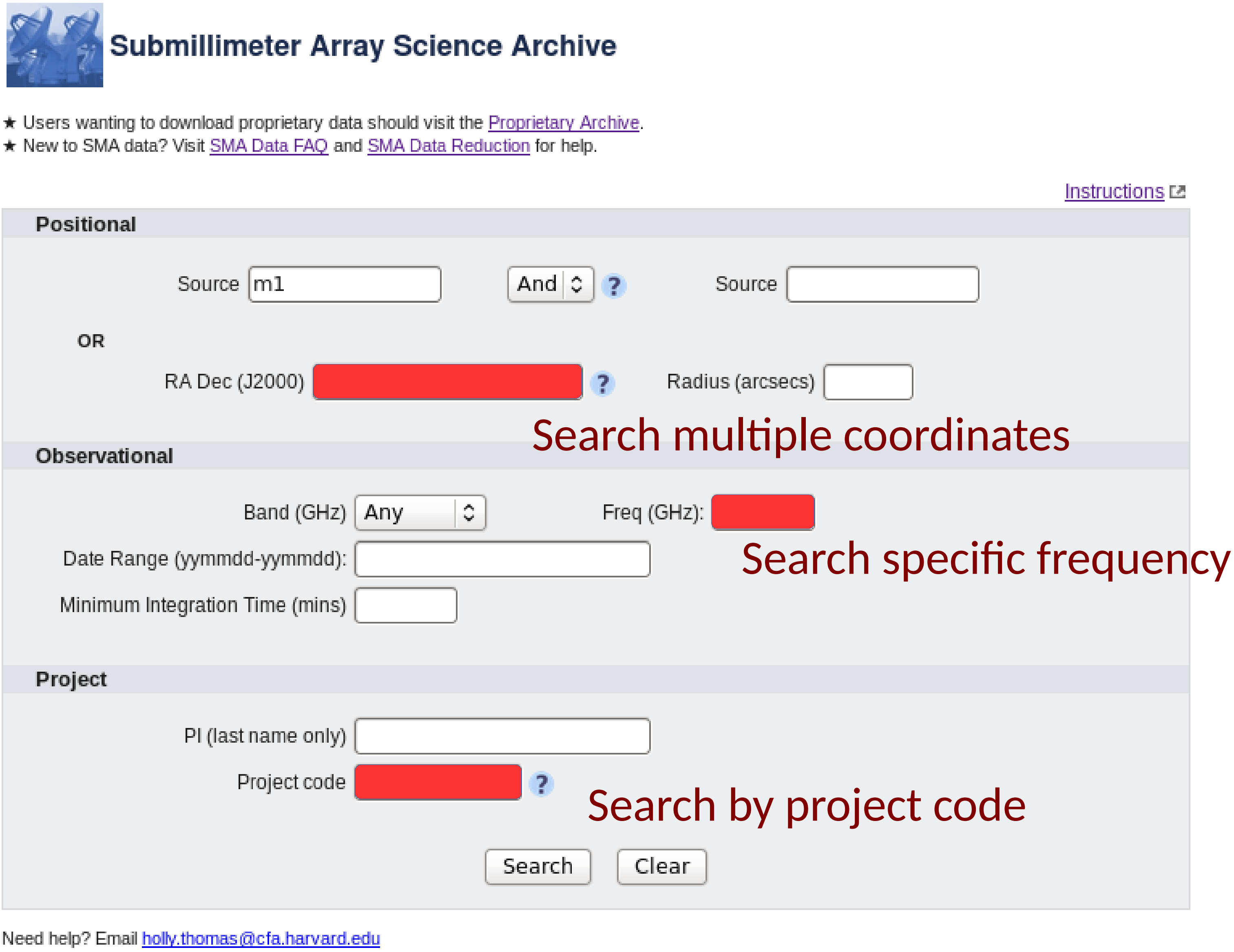
Project

PI (last name only):

Project code:

Need help? Email holly.thomas@cfa.harvard.edu

SMA Archives



Submillimeter Array Science Archive

★ Users wanting to download proprietary data should visit the Proprietary Archive.
★ New to SMA data? Visit SMA Data FAQ and SMA Data Reduction for help.

Instructions

Positional

Source: m1 — And — Source:

OR

RA Dec (J2000): — Radius (arcsecs):

Search multiple coordinates

Observational

Band (GHz): Any — Freq (GHz):

Date Range (yymmdd-yymmdd):

Minimum Integration Time (mins):

Search specific frequency

Project

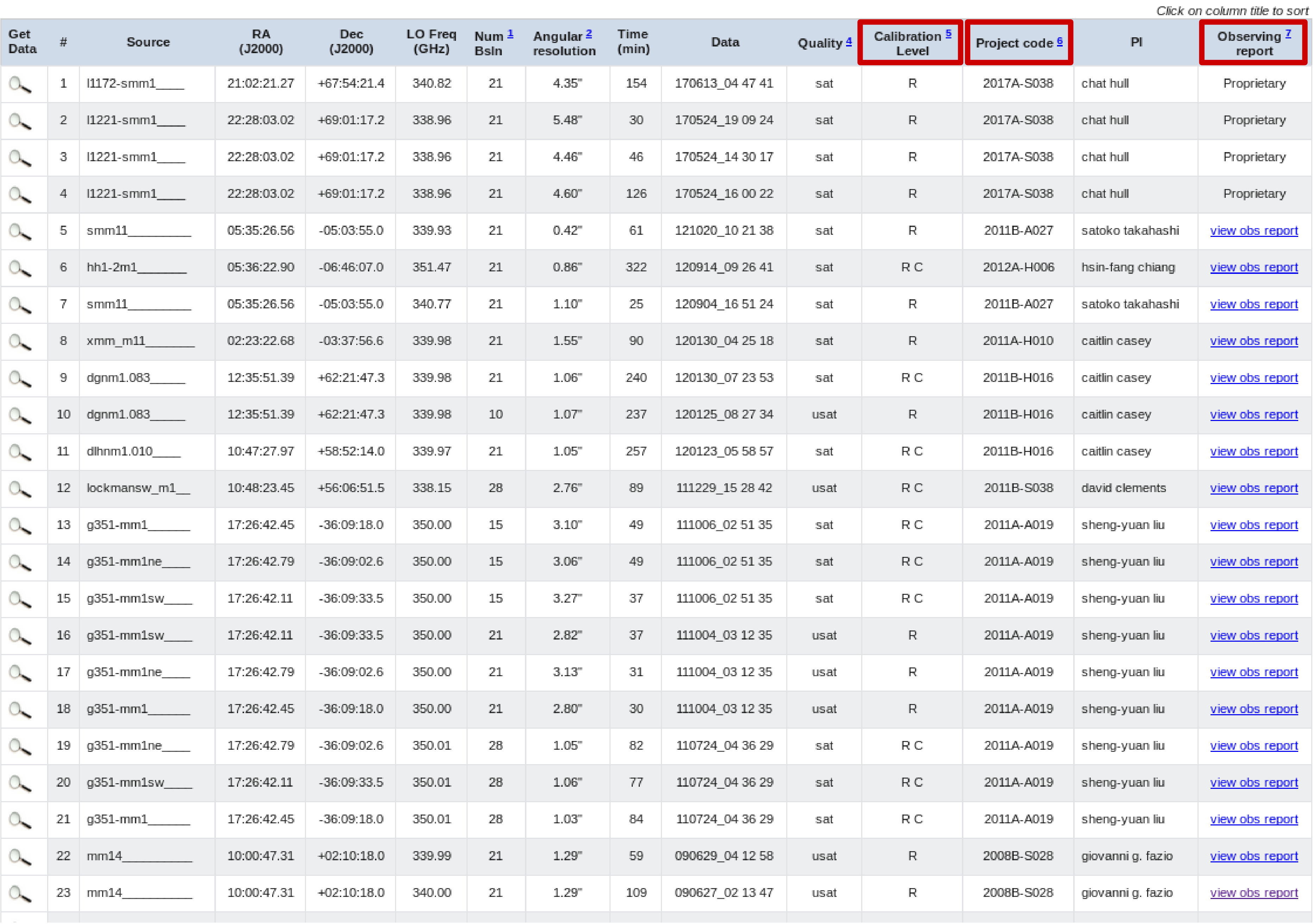
PI (last name only):

Project code:

Search by project code

Need help? Email holly.thomas@cfa.harvard.edu

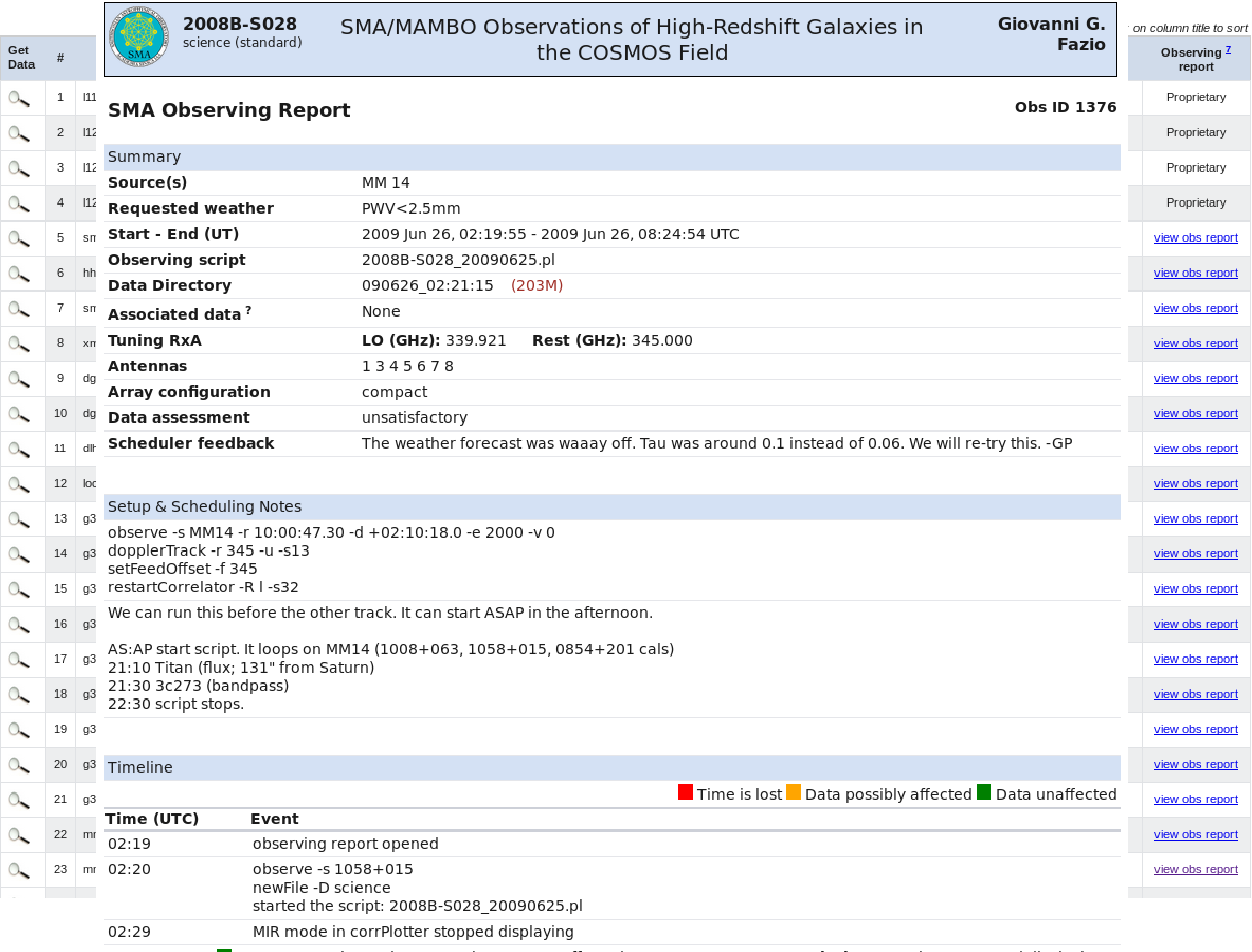
SMA Archives



Click on column title to sort

Get Data	#	Source	RA (J2000)	Dec (J2000)	LO Freq (GHz)	Num[1] Bsln	Angular[2] resolution	Time (min)	Data	Quality[4]	Calibration[5] Level	Project code[6]	PI	Observing[7] report
🔍	1	l1172-smm1____	21:02:21.27	+67:54:21.4	340.82	21	4.35"	154	170613_04 47 41	sat	R	2017A-S038	chat hull	Proprietary
🔍	2	l1221-smm1____	22:28:03.02	+69:01:17.2	338.96	21	5.48"	30	170524_19 09 24	sat	R	2017A-S038	chat hull	Proprietary
🔍	3	l1221-smm1____	22:28:03.02	+69:01:17.2	338.96	21	4.46"	46	170524_14 30 17	sat	R	2017A-S038	chat hull	Proprietary
🔍	4	l1221-smm1____	22:28:03.02	+69:01:17.2	338.96	21	4.60"	126	170524_16 00 22	sat	R	2017A-S038	chat hull	Proprietary
🔍	5	smm11_________	05:35:26.56	-05:03:55.0	339.93	21	0.42"	61	121020_10 21 38	sat	R	2011B-A027	satoko takahashi	view obs report
🔍	6	hh1-2m1_______	05:36:22.90	-06:46:07.0	351.47	21	0.86"	322	120914_09 26 41	sat	R C	2012A-H006	hsin-fang chiang	view obs report
🔍	7	smm11_________	05:35:26.56	-05:03:55.0	340.77	21	1.10"	25	120904_16 51 24	sat	R	2011B-A027	satoko takahashi	view obs report
🔍	8	xmm_m11_______	02:23:22.68	-03:37:56.6	339.98	21	1.55"	90	120130_04 25 18	sat	R	2011A-H010	caitlin casey	view obs report
🔍	9	dgnm1.083_____	12:35:51.39	+62:21:47.3	339.98	21	1.06"	240	120130_07 23 53	sat	R C	2011B-H016	caitlin casey	view obs report
🔍	10	dgnm1.083_____	12:35:51.39	+62:21:47.3	339.98	10	1.07"	237	120125_08 27 34	usat	R	2011B-H016	caitlin casey	view obs report
🔍	11	dlhnm1.010____	10:47:27.97	+58:52:14.0	339.97	21	1.05"	257	120123_05 58 57	sat	R C	2011B-H016	caitlin casey	view obs report
🔍	12	lockmansw_m1__	10:48:23.45	+56:06:51.5	338.15	28	2.76"	89	111229_15 28 42	usat	R C	2011B-S038	david clements	view obs report
🔍	13	g351-mm1______	17:26:42.45	-36:09:18.0	350.00	15	3.10"	49	111006_02 51 35	sat	R C	2011A-A019	sheng-yuan liu	view obs report
🔍	14	g351-mm1ne____	17:26:42.79	-36:09:02.6	350.00	15	3.06"	49	111006_02 51 35	sat	R C	2011A-A019	sheng-yuan liu	view obs report
🔍	15	g351-mm1sw____	17:26:42.11	-36:09:33.5	350.00	15	3.27"	37	111006_02 51 35	sat	R C	2011A-A019	sheng-yuan liu	view obs report
🔍	16	g351-mm1sw____	17:26:42.11	-36:09:33.5	350.00	21	2.82"	37	111004_03 12 35	usat	R	2011A-A019	sheng-yuan liu	view obs report
🔍	17	g351-mm1ne____	17:26:42.79	-36:09:02.6	350.00	21	3.13"	31	111004_03 12 35	usat	R	2011A-A019	sheng-yuan liu	view obs report
🔍	18	g351-mm1______	17:26:42.45	-36:09:18.0	350.00	21	2.80"	30	111004_03 12 35	usat	R	2011A-A019	sheng-yuan liu	view obs report
🔍	19	g351-mm1ne____	17:26:42.79	-36:09:02.6	350.01	28	1.05"	82	110724_04 36 29	sat	R C	2011A-A019	sheng-yuan liu	view obs report
🔍	20	g351-mm1sw____	17:26:42.11	-36:09:33.5	350.01	28	1.06"	77	110724_04 36 29	sat	R C	2011A-A019	sheng-yuan liu	view obs report
🔍	21	g351-mm1______	17:26:42.45	-36:09:18.0	350.01	28	1.03"	84	110724_04 36 29	sat	R C	2011A-A019	sheng-yuan liu	view obs report
🔍	22	mm14__________	10:00:47.31	+02:10:18.0	339.99	21	1.29"	59	090629_04 12 58	usat	R	2008B-S028	giovanni g. fazio	view obs report
🔍	23	mm14__________	10:00:47.31	+02:10:18.0	340.00	21	1.29"	109	090627_02 13 47	usat	R	2008B-S028	giovanni g. fazio	view obs report

SMA Archives



: on column title to sort

Get Data	#		Observing report
🔍	1	l11	Proprietary
🔍	2	l12	Proprietary
🔍	3	l12	Proprietary
🔍	4	l12	Proprietary
🔍	5	sm	view obs report
🔍	6	hh	view obs report
🔍	7	sm	view obs report
🔍	8	xm	view obs report
🔍	9	dg	view obs report
🔍	10	dg	view obs report
🔍	11	dlı	view obs report
🔍	12	loc	view obs report
🔍	13	g3	view obs report
🔍	14	g3	view obs report
🔍	15	g3	view obs report
🔍	16	g3	view obs report
🔍	17	g3	view obs report
🔍	18	g3	view obs report
🔍	19	g3	view obs report
🔍	20	g3	view obs report
🔍	21	g3	view obs report
🔍	22	mr	view obs report
🔍	23	mr	view obs report

2008B-S028
science (standard)

SMA/MAMBO Observations of High-Redshift Galaxies in the COSMOS Field

Giovanni G. Fazio

SMA Observing Report

Obs ID 1376

Summary

Source(s)	MM 14
Requested weather	PWV<2.5mm
Start - End (UT)	2009 Jun 26, 02:19:55 - 2009 Jun 26, 08:24:54 UTC
Observing script	2008B-S028_20090625.pl
Data Directory	090626_02:21:15 (203M)
Associated data ?	None
Tuning RxA	**LO (GHz):** 339.921 **Rest (GHz):** 345.000
Antennas	1 3 4 5 6 7 8
Array configuration	compact
Data assessment	unsatisfactory
Scheduler feedback	The weather forecast was waaay off. Tau was around 0.1 instead of 0.06. We will re-try this. -GP

Setup & Scheduling Notes

```
observe -s MM14 -r 10:00:47.30 -d +02:10:18.0 -e 2000 -v 0
dopplerTrack -r 345 -u -s13
setFeedOffset -f 345
restartCorrelator -R l -s32
```

We can run this before the other track. It can start ASAP in the afternoon.

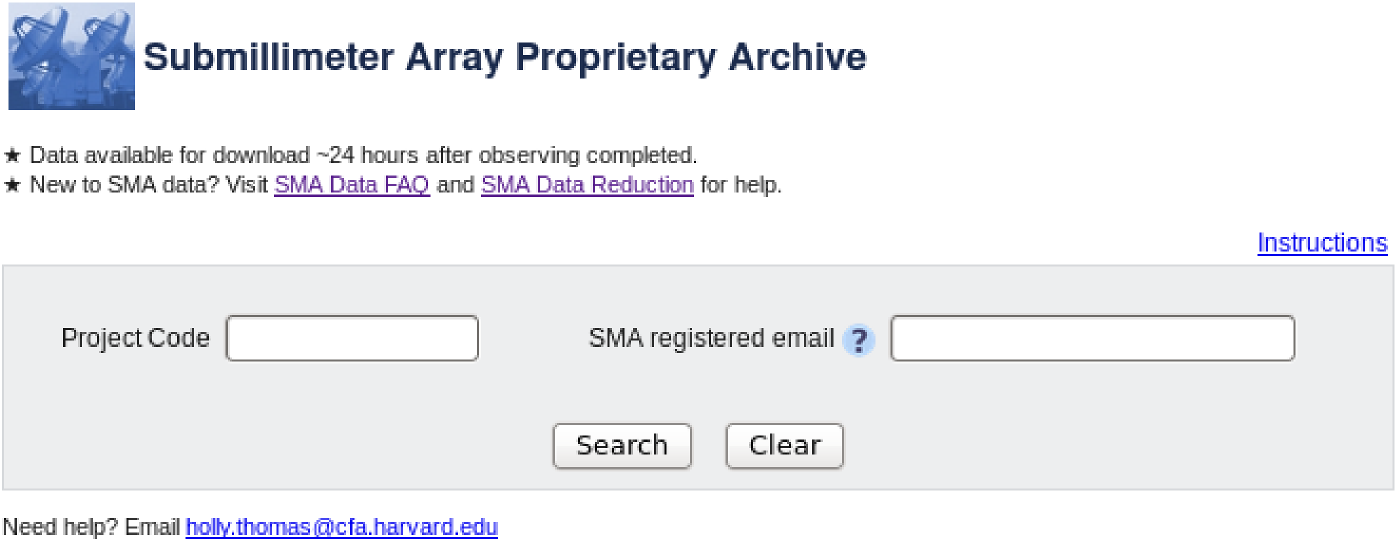
AS:AP start script. It loops on MM14 (1008+063, 1058+015, 0854+201 cals)
21:10 Titan (flux; 131" from Saturn)
21:30 3c273 (bandpass)
22:30 script stops.

Timeline

🟥 Time is lost 🟧 Data possibly affected 🟩 Data unaffected

Time (UTC)	Event
02:19	observing report opened
02:20	observe -s 1058+015 newFile -D science started the script: 2008B-S028_20090625.pl
02:29	MIR mode in corrPlotter stopped displaying

SMA Archives


Project Code
SMA registered email
?
Search
Clear

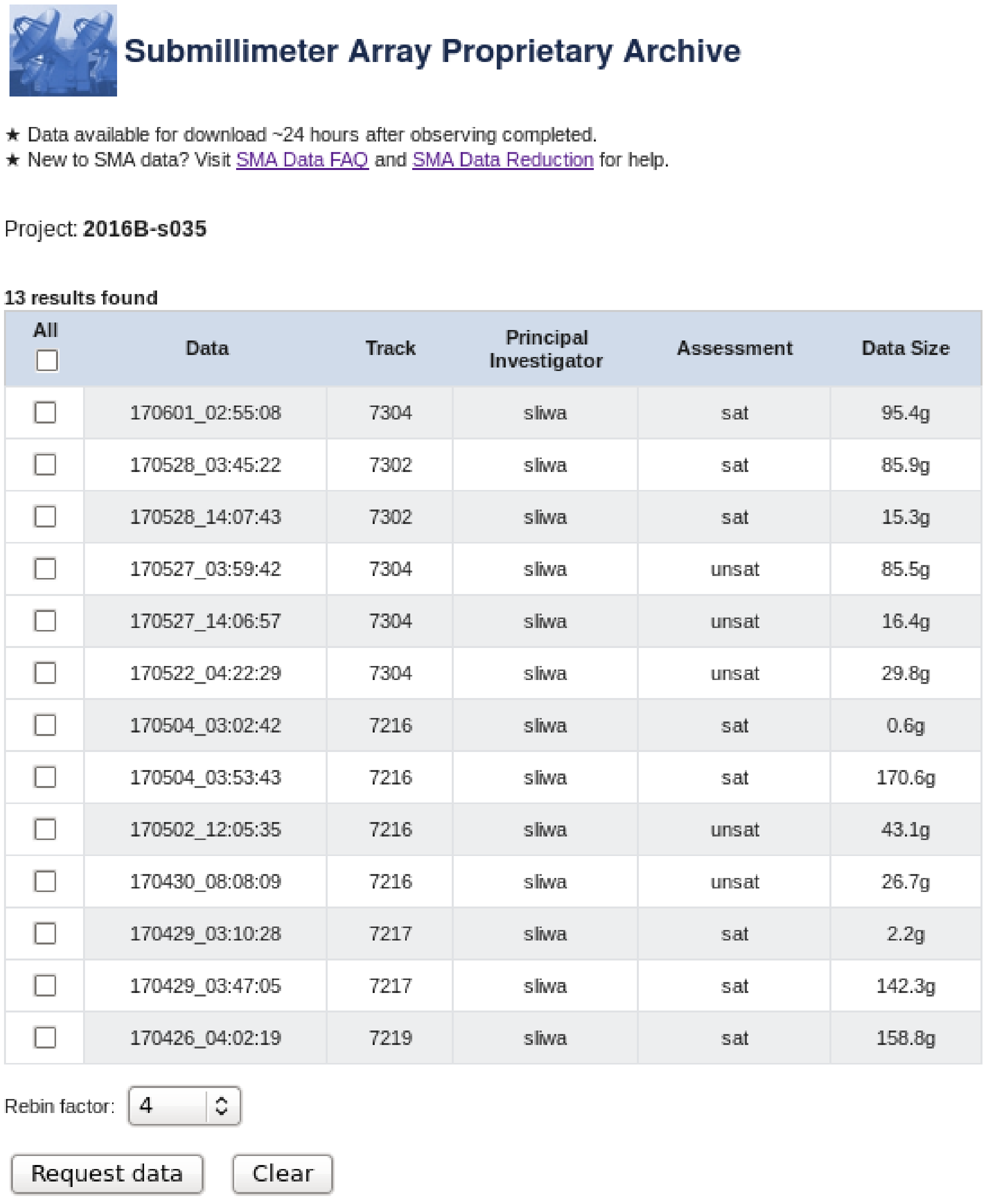
Submillimeter Array Proprietary Archive

- ★ Data available for download ~24 hours after observing completed.
- ★ New to SMA data? Visit SMA Data FAQ and SMA Data Reduction for help.

Instructions

Need help? Email holly.thomas@cfa.harvard.edu

SMA Archives



Submillimeter Array Proprietary Archive

★ Data available for download ~24 hours after observing completed.
★ New to SMA data? Visit SMA Data FAQ and SMA Data Reduction for help.

Project: **2016B-s035**

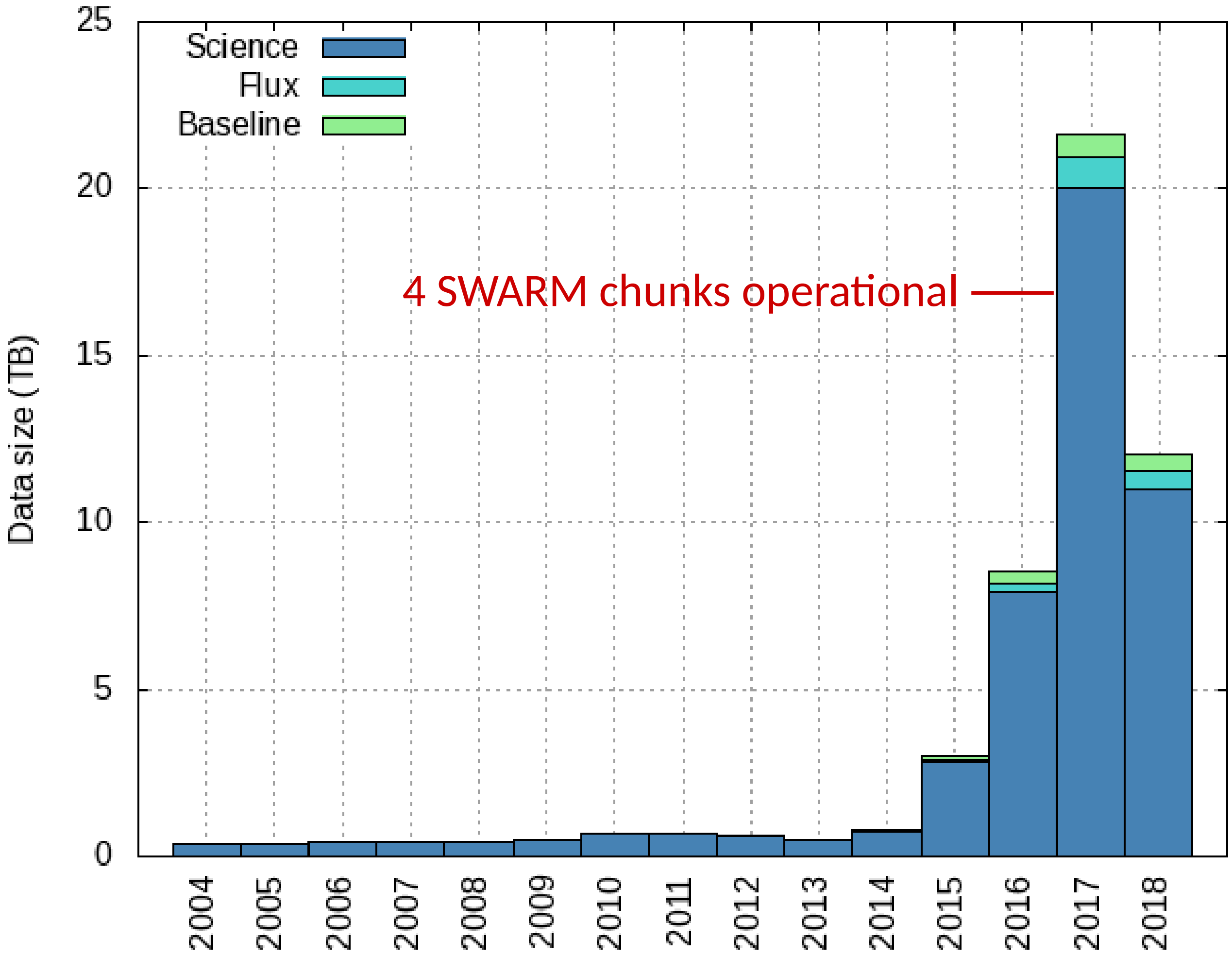
13 results found

All ☐	Data	Track	Principal Investigator	Assessment	Data Size
☐	170601_02:55:08	7304	sliwa	sat	95.4g
☐	170528_03:45:22	7302	sliwa	sat	85.9g
☐	170528_14:07:43	7302	sliwa	sat	15.3g
☐	170527_03:59:42	7304	sliwa	unsat	85.5g
☐	170527_14:06:57	7304	sliwa	unsat	16.4g
☐	170522_04:22:29	7304	sliwa	unsat	29.8g
☐	170504_03:02:42	7216	sliwa	sat	0.6g
☐	170504_03:53:43	7216	sliwa	sat	170.6g
☐	170502_12:05:35	7216	sliwa	unsat	43.1g
☐	170430_08:08:09	7216	sliwa	unsat	26.7g
☐	170429_03:10:28	7217	sliwa	sat	2.2g
☐	170429_03:47:05	7217	sliwa	sat	142.3g
☐	170426_04:02:19	7219	sliwa	sat	158.8g

Rebin factor: 4

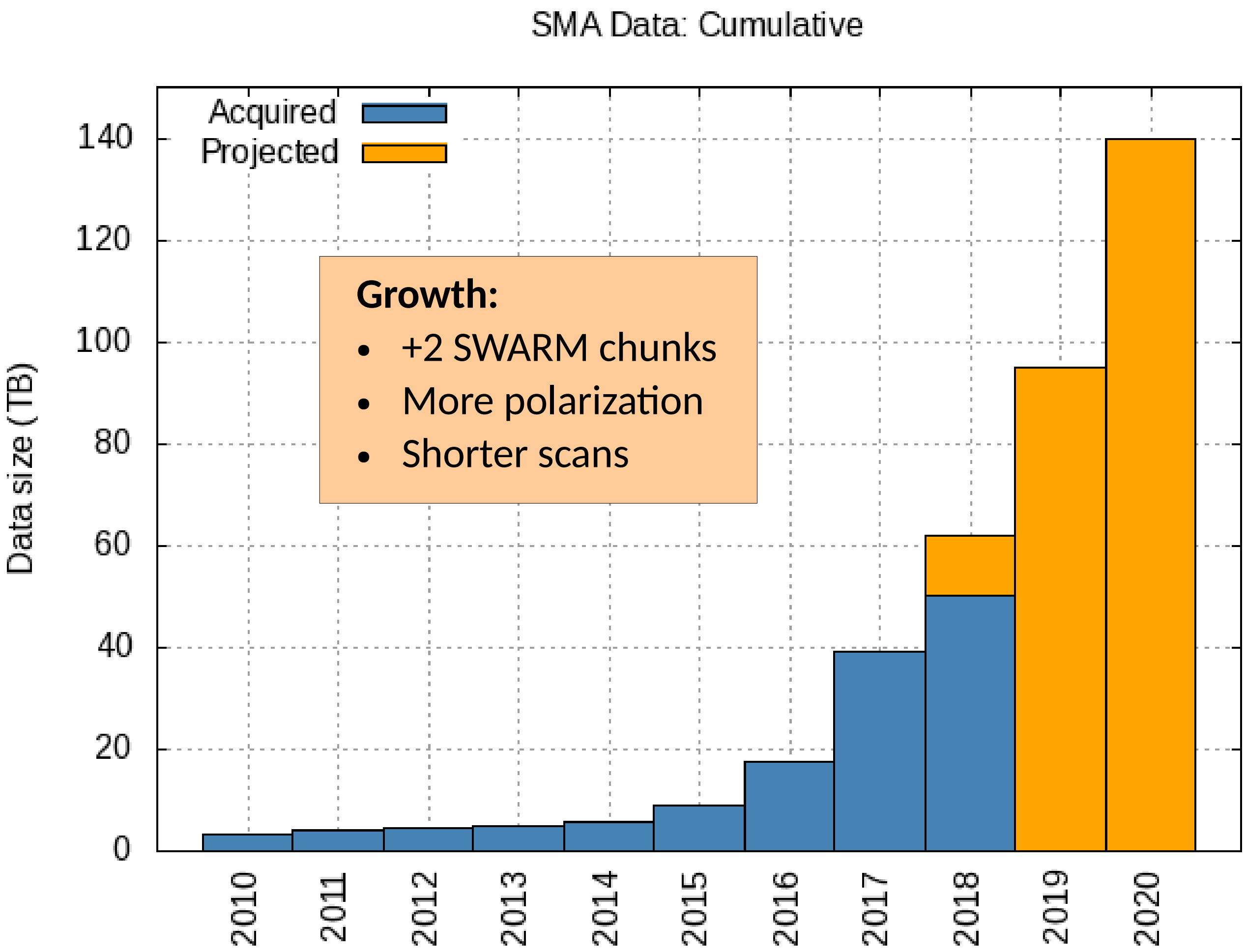
Linked from SMA Observer Center

SMA Data



4 SWARM chunks operational

SMA Data



SMA Data: Cumulative

Growth:

- +2 SWARM chunks
- More polarization
- Shorter scans

- Current size of archive ~ 50 TB
- Feb 2018 – New storage server (160TB)

SMA Data

amazon web services

Google Cloud

Microsoft Azure

Cloud benefits

- Wont degrade
- Fully expandable
- Cheaper long-term

Testing AWS for SMA – use TBD

- All data or non-proprietary data only?
- Private (long-term backup) + public (packaged for archive)
- Archive provides link to cloud

SMA Data Reduction

Reduction path

Calibration: MIR

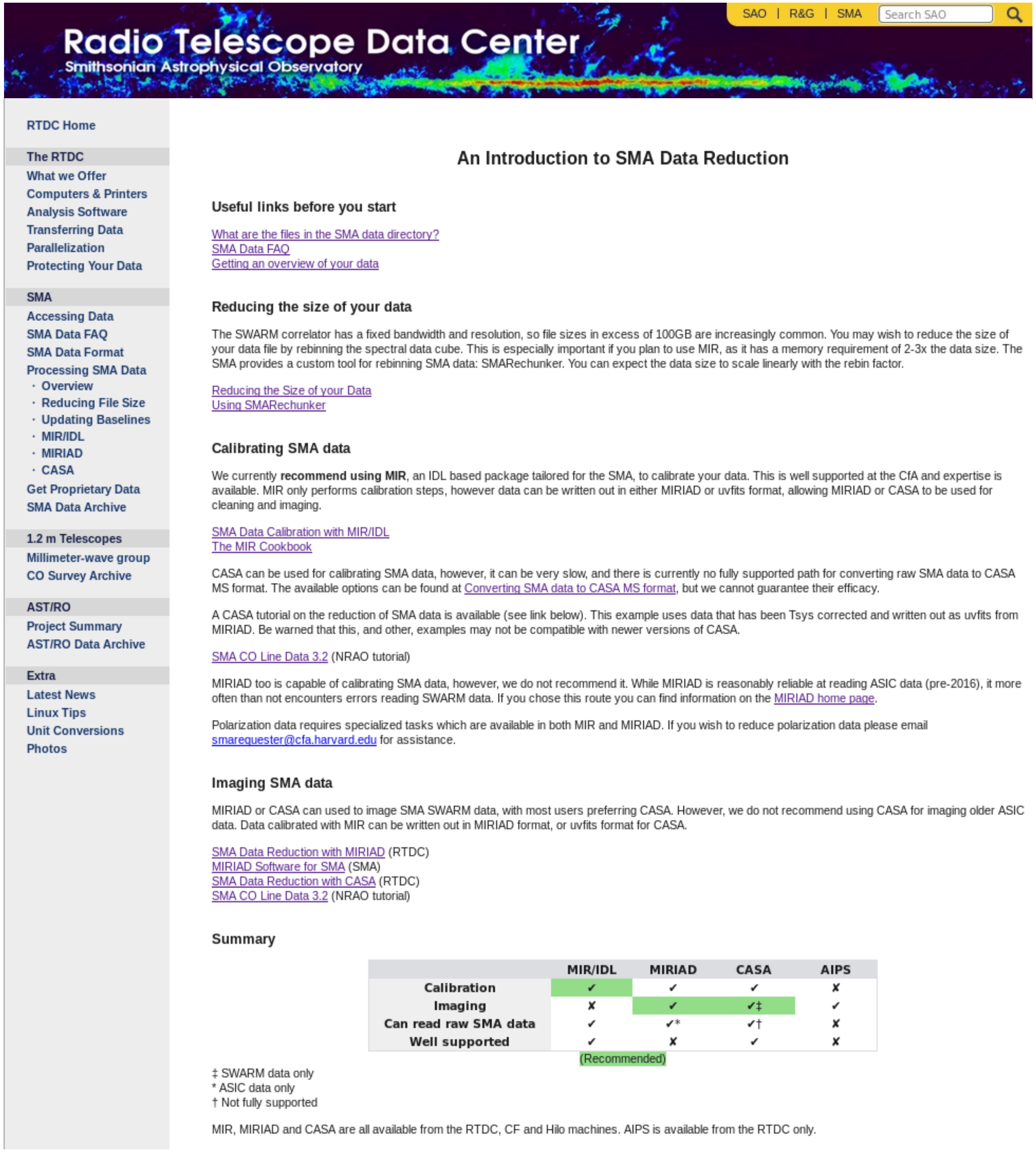
Imaging: CASA (MIRIAD)

Issues with MIR: memory/license

Future

- All reduction in CASA
- More work needed first
- CASA parallelization

Good documentation is key!


SAO | R&G | SMA
Search SAO
Radio Telescope Data Center
Smithsonian Astrophysical Observatory
RTDC Home
The RTDC
What we Offer
Computers & Printers
Analysis Software
Transferring Data
Parallelization
Protecting Your Data
SMA
Accessing Data
SMA Data FAQ
SMA Data Format
Processing SMA Data
· Overview
· Reducing File Size
· Updating Baselines
· MIR/IDL
· MIRIAD
· CASA
Get Proprietary Data
SMA Data Archive
1.2 m Telescopes
Millimeter-wave group
CO Survey Archive
AST/RO
Project Summary
AST/RO Data Archive
Extra
Latest News
Linux Tips
Unit Conversions
Photos

An Introduction to SMA Data Reduction

Useful links before you start

What are the files in the SMA data directory?
SMA Data FAQ
Getting an overview of your data

Reducing the size of your data

The SWARM correlator has a fixed bandwidth and resolution, so file sizes in excess of 100GB are increasingly common. You may wish to reduce the size of your data file by rebinning the spectral data cube. This is especially important if you plan to use MIR, as it has a memory requirement of 2-3x the data size. The SMA provides a custom tool for rebinning SMA data: SMARechunker. You can expect the data size to scale linearly with the rebin factor.

Reducing the Size of your Data
Using SMARechunker

Calibrating SMA data

We currently **recommend using MIR**, an IDL based package tailored for the SMA, to calibrate your data. This is well supported at the CfA and expertise is available. MIR only performs calibration steps, however data can be written out in either MIRIAD or uvfits format, allowing MIRIAD or CASA to be used for cleaning and imaging.

SMA Data Calibration with MIR/IDL
The MIR Cookbook

CASA can be used for calibrating SMA data, however, it can be very slow, and there is currently no fully supported path for converting raw SMA data to CASA MS format. The available options can be found at Converting SMA data to CASA MS format, but we cannot guarantee their efficacy.

A CASA tutorial on the reduction of SMA data is available (see link below). This example uses data that has been Tsys corrected and written out as uvfits from MIRIAD. Be warned that this, and other, examples may not be compatible with newer versions of CASA.

SMA CO Line Data 3.2 (NRAO tutorial)

MIRIAD too is capable of calibrating SMA data, however, we do not recommend it. While MIRIAD is reasonably reliable at reading ASIC data (pre-2016), it more often than not encounters errors reading SWARM data. If you chose this route you can find information on the MIRIAD home page.

Polarization data requires specialized tasks which are available in both MIR and MIRIAD. If you wish to reduce polarization data please email smarequester@cfa.harvard.edu for assistance.

Imaging SMA data

MIRIAD or CASA can used to image SMA SWARM data, with most users preferring CASA. However, we do not recommend using CASA for imaging older ASIC data. Data calibrated with MIR can be written out in MIRIAD format, or uvfits format for CASA.

SMA Data Reduction with MIRIAD (RTDC)
MIRIAD Software for SMA (SMA)
SMA Data Reduction with CASA (RTDC)
SMA CO Line Data 3.2 (NRAO tutorial)

Summary

	MIR/IDL	MIRIAD	CASA	AIPS
Calibration	✓	✓	✓	✗
Imaging	✗	✓	✓‡	✓
Can read raw SMA data	✓	✓*	✓†	✗
Well supported	✓	✗	✓	✗

(Recommended)

‡ SWARM data only
* ASIC data only
† Not fully supported

MIR, MIRIAD and CASA are all available from the RTDC, CF and Hilo machines. AIPS is available from the RTDC only.

Science Pipeline

Goals

- Provide PIs with calibrated datasets
- Build processed data archive
- Non-properietary data → calibrated + quicklook image

Testing the Options

- MIR calibration scripts
- Robust SMA → CASA conversion – CASA calibration script
- “In-house” calibration pipeline

Summary

To Date

- More useful webpages
- More functional archive
- Responding to storage requirements
- Exploring pipeline options

The Future

- Supply PIs with calibrated (imaged) data
- More integration between RTDC, SMA, and cloud services
- Clear documentation for users