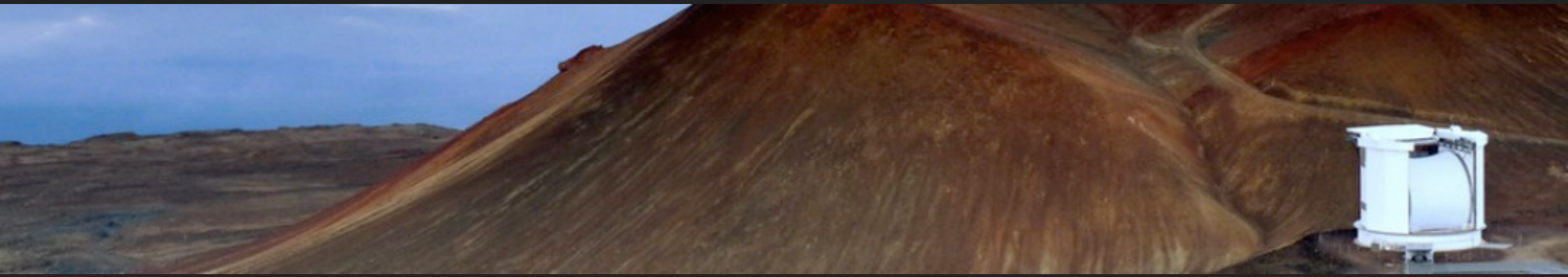


SARAH GRAVES, GRAHAM BELL,
MALCOLM CURRIE, DAVID BERRY

The JCMT SCUBA-2 Legacy Release: Unexpected Benefits and Lessons Learned

THE JAMES CLERK MAXWELL TELESCOPE

- 15m submillimetre telescope on Maunakea, Hawaii
- 30 years old in spring 2017
- Historically was operated by the Joint Astronomy Centre (JAC) for the UK, Canada and the Netherlands.
- Since March 2015 operated by East Asian Observatory.



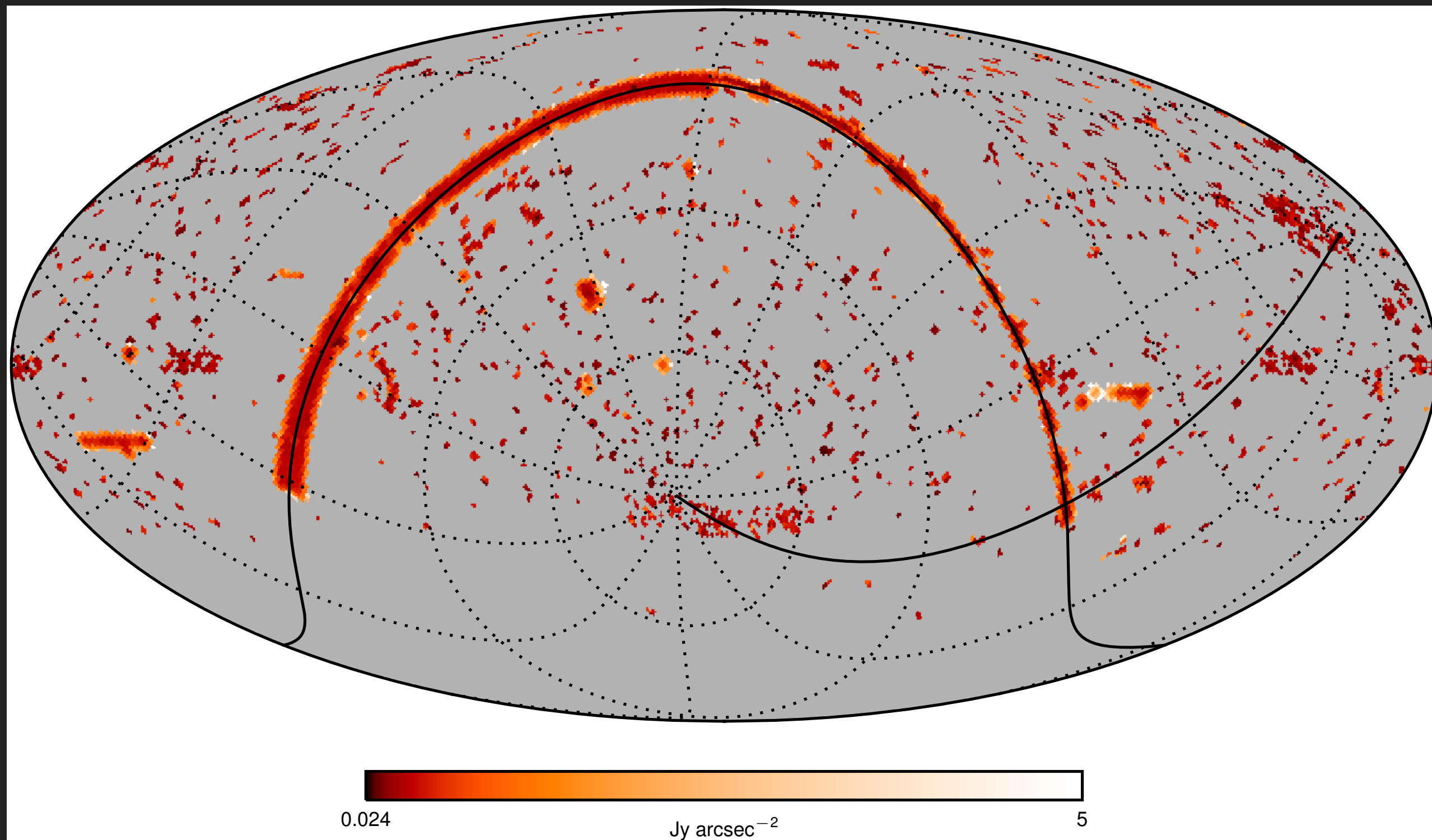
SCUBA-2

- ▶ 10,000 pixel bolometer camera operating simultaneously at 450 and 850 microns
- ▶ 13" main beam at 850 μ m
- ▶ 45 arcmin² field of view.
- ▶ Raw data are gridded and reduced using the Starlink SMURF makemap software (Chapin+2013)
- ▶ Different DR configurations used depending on the nature of emission in the field (e.g. blank field/point sources/complex filaments)



JCMT LEGACY RELEASE: SCUBA-2 850 OBSERVATIONS

- ▶ Aim: Providing consistent reductions, co-adds and catalogues of all public data (Bell+ 2014): diverse PI projects, calibration observations and surveys.
- ▶ Compare with non-observatory produced SCUBA legacy catalogue (DiFrancesco+ 2008)
- ▶ First batch of products released September 2015 (observations from 2011 February 2 to 2013 August 1) <http://www.eaobservatory.org/jcmt/science/archive/lr1/>
- ▶ 2nd batch planned for release in Autumn/Winter 2016. Includes all data taken before EAO handover, 2015 March 1st.
- ▶ 5828 hours/12404 observations/1356 square degrees in co-added tiles.
- ▶ 1.37 square degrees detected emission, available as catalogues & MOCs
- ▶ Trade offs in reduction method: insensitive to large scale structure.



BUILDING ON WHAT CAME BEFORE

Before work started on this project, many required pieces/systems for this work already existed at the observatory.

- ▶ JCMT Science archive (Economou+ 2015) at CADC
 - ▶ this already contained both raw and pipeline reduced data, and coadds from single projects on single night.
 - ▶ During same time as this work, the JSA was moved to CAOM-2
- ▶ Software: Starlink and the ORAC-DR pipeline (Jenness&Economou 2015)
- ▶ Systems: OMP: metadata and QA for all observations, with web and script interfaces. Stable and established system predating SCUBA-2.

EXTRA BENEFITS OF A LEGACY RELEASE

Primary reason is (of course) the scientific goal. However, there are a lot of additional benefits for an observatory.

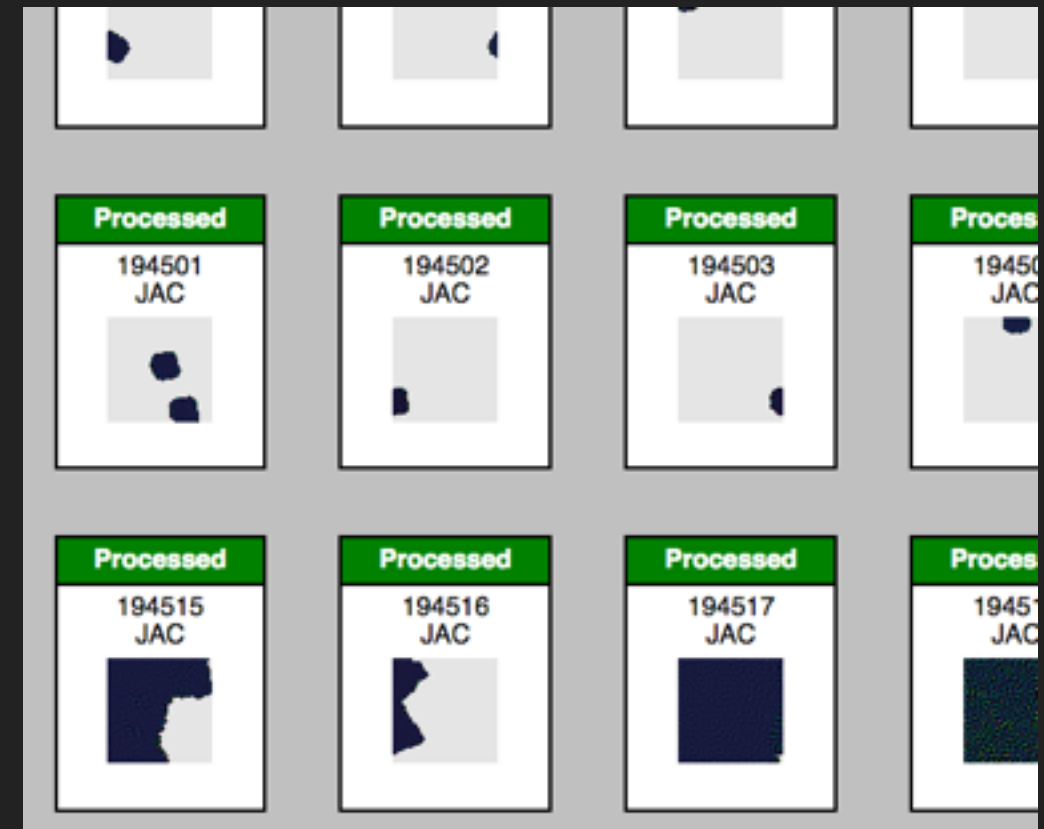
- ▶ Improvements to DR/Analysis Software
- ▶ Improvements to systems and processes that run your software, including QA processes.
- ▶ Discovery of problems/bugs.
- ▶ Enhanced understanding of what your users need

IMPROVEMENTS TO DR AND ANALYSIS SOFTWARE

- ▶ We required a DR configuration that would work well on different types of source emission:
 - ▶ Developed a SCUBA-2 reduction configuration for this end, which we use as our default, producing better quality reductions for users.
- ▶ Choice to use HEALPix tiles => HPX projection now well supported by Starlink/AST (Berry+2016 A&C), used by Starlink for its coordinate system support.
- ▶ Improvements to analysis software to support HPX (e.g. removing assumptions about pixel sizes/shape.)

IMPROVEMENTS TO SYSTEMS: DATA PROCESSING SOFTWARE

- ▶ Created a new data processing system (jsa_proc). Originally to coordinate and handle ORAC-DR reductions of our largest jobs.
- ▶ ... while we were developing this, CADC had to stop running our data processing on their machines.
- ▶ Because of the work to support legacy release, much shorter delay in restoring nightly pipeline processing.
- ▶ New data processing system a great success: handles processing, data transfer and meta data ingestion into CADC archive.

[illegible]

IMPROVEMENTS TO SYSTEMS: QA

- ▶ Requirement of release was to have a human eye okay individual science observation – this is also done for nightly reductions.

hpx-s2-850-r2-coadd : Job 194576

Job Info	
Job QA State	Unknown
State	Processed
Location	JAC
Parameters	COADD_JSA_TILES
Task	hpx-s2-850-r2-coadd
Mode	public
Tiles	
JSA Wrap DR Logs	jswrapdr_20...
PICARD Logs	picard_20160...
Job Info Page	job information
Children	198792
Parents	171650 8963 172933 8967 8973 172431 8977 8980 172058 555 183463 183465 1082 1219 1093 8908 8930 1230 1105 10453 173518 171483 1250 171621 1254 172009 481 495 496 8950 8953 172415

Good & go to next job

Set QA state for this job

☐ Invalid

☒ Unknown

☐ Bad

☐ Questionable


☐ Good

List any problems seen

Set QA & stay on page

Set QA & go to next

scuba-2-000862_healpix-850um



Parent Observation Info												
id	utdate	scan	instru-	sub-	project	sur-	scanmode	source	ob-	date_obs	Synched OMP Status	tau
481	2012-08-14	47	SCUBA-2	850	MULSC38	DBS	CURVY_PONG	NGC1333-S	science	2012-08-14 12:39:22	Good	0.07111
482	2012-08-14	50	SCUBA-2	850	MULSC38	DBS	CURVY_PONG	NGC1333-S	science	2012-08-14 14:18:36	Good	0.06886

- ▶ Built this QA into new processing system, and it was extensively used and improved while QA'ing > 12000 observations.
- ▶ Same system now used for nightly reductions – became easier to identify projects with poor config, and to identify patterns in QA failures.

DISCOVERY OF PROBLEMS

Observations from different PIs and surveys. aren't usually considered as a single data set – but they become one for the legacy release.

The analysis of these as a single data set reveals issues and areas to investigate:

- ▶ analysis of the pointing offsets: discovered updates to pointing source positions not copied into flux calibration MSBs.
- ▶ Calibration analysis: replicating analysis of Dempsey+2013 over a larger time period.
- ▶ Investigation of errors and poor QA found and corrected various mis flagged observations.

A BETTER UNDERSTANDING OF WHAT ARCHIVAL USERS NEED TO KNOW

Staff spent time tracking down problem observations : becomes clear which information is needed to do this.

- ▶ Observation and shift logs often required to understand data flags, but not available to archival users.
- ▶ Data quality flags weren't in initial CAOM-2, hence not visible to archive users (since added)
- ▶ Forces observatory to think about long-term archival use of observations, as opposed to immediate PI's needs.
 - ▶ test observations taken by instrument scientists with the dome closed –not always flagged as JUNK! Completely opaque to external users.

LESSONS LEARNED

- ▶ Errors in metadata can take up an amazing amount of time and thought to work around. It should be a very high priority to fix them!
- ▶ information contained in free form text is much better than nothing – but it is a time sink if you have to investigate.
- ▶ JCMT's systems (i.e. OMP) were designed by people considering long term access and archival use – without this archival structure this data release would have been far harder, and the results far worse.

- ▶ If you're thinking with a view to long-term access you care about a different set of problems and criteria – but most of the time these won't be top priority for observatory staff working on fixing immediate issues.
- ▶ Regular analysis over large timescales is fruitful – reveal changes, bugs etc – even though it can be lost amid day-to-day observatory work.

