

Data Processing Status

R. Y. Shuping, *Data Processing Lead*

Space Science Inst.
USRA-SOFIA

SOFIA Data Products

Defined in the Data Processing Plan for SOFIA SIs :

Level 1: raw SI data in standardized format (FITS)

Level 2: corrected for instrument artifacts (e.g. dark current, bad pixels, etc...)

Level 3: flux calibrated (using FITS keywords; Jy/pix)

Level 4: high-order products possibly combining multiple observations
(e.g. mosaics, spectral cubes)

Since Last SUG (Oct 2014)

- Completed all remaining Level 3 processing for OC2 (Imaging/Grism, Nov 2014)
- Completed all processing for OC2H (Feb 2015) **on-schedule**
- Released v1 of EXES pipeline (Mar 2015)
- Supported both EXES and FIFI-LS commissioning flights.
- Hired new scientist, Ed Chambers (Univ. Cologne), to support GREAT
- Hired new engineer, Jennifer Holt (Keck), to work on FIFI-LS pipeline (and QA database).

Data processing team has produced ~270 GB L2/3 data (>62000 files) for OC1 and 2.

Operational Status

- All processing for FORCAST/FLITECAM in OC2 complete.
 - *Data available in archive; GIs notified.*
- *Supporting EXES OC3A processing.*
 - *Header fixes and manual processing in work.*
- Reprocessing:
 - Update FORCAST grism response curve uncertainty spectra for OC1 -- Processing DONE, notification imminent
 - Reprocess data from 01_0042 (PI: Sankrit) – processing DONE, notification imminent.
 - Reprocessing OC1B FORCAST imaging data to correct problem with large chop throws – IN WORK

SI Pipeline Readiness

SI Pipeline	Status
FORCAST Imaging	Automatic pipeline in operation.
FORCAST Grism	Automatic pipeline in operation.
FLITECAM Imaging	Automatic pipeline in operation.
FLITECAM Grism	Automatic pipeline in operation.
EXES	Manual pipeline tested/deployed; automatic mode under test with OC3A data.
FIFI-LS	Preliminary pipeline received; alpha version complete; <i>algorithm updates and DPS IT&V underway.</i>
HAWC+	Preliminary pipeline for HAWC received; <i>update for HAWC+ expected in May (TBC).</i>
GREAT	<i>Processing handled by SI team; results archived at SSC.</i>
HIPO	N/A

DPS Pipeline Development for CY 2015

- FORCAST:
 - Minor engineering updates for OC2H (DONE)
 - Minor engineering updates for OC3 (IN WORK)
 - Develop slit response function (IN WORK)
 - Review/test wavecals vs. TAAS data
 - Improve Droop correction in LWC
- FLITECAM:
 - Complete integration with DPS infrastructure (DONE)
 - Minor Updates for OC3 (IN WORK)
 - Update wavecals for LM+A grism (WAITING)
 - Update wavecals for 1" slit (imminent)
- FIFI-LS:
 - Complete IT&V with DPS (IN WORK)
 - Update telluric correction & flux calibration algorithms (WAITING)
 - Release v1 in time for OC3K (Oct 2015)
- EXES:
 - Finish/release v1.0.0 (DONE)

DPS Infrastructure Development for CY 2015

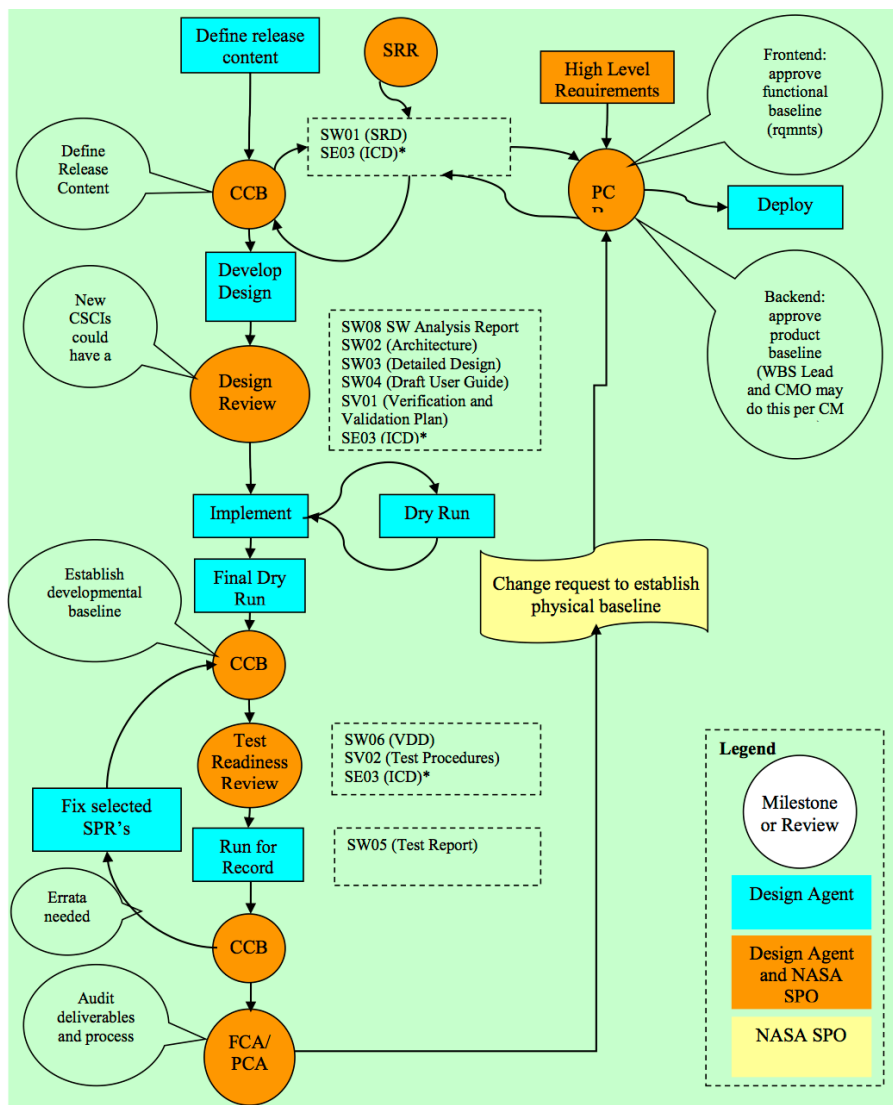
New capabilities/tools:

- Update grouping algorithm/implementation (DONE)
- Auto pipeline timeout support (DONE)
- Automatic email notification tool (DONE)
- Deploy prototype internal QA database (IN WORK)
- Display QA comments in DCS archive (IN WORK)
- Update WCS QA reporting

Management Tools

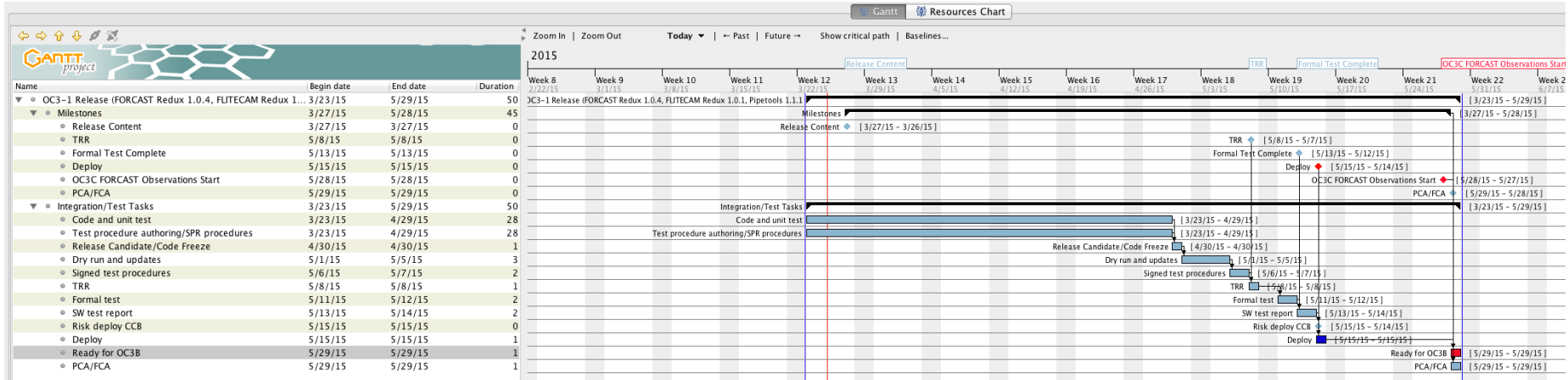
- All documents stored/shared internally using local Wiki.
- All Issues (developmental/operational) tracked using JIRA and assigned to specific software versions for release.
- Development schedules follow NASA SMP using Gantt charts to track progress.
 - Driven primarily by flight series start dates.
- Operational schedule/resources estimated using Gantt charts.
- Operational status tracked using Kanban Board (Trello)
 - Allows easy tracking/notification of processing on a flight-by-flight basis (web/mobile apps) .
 - Collects deadlines, responsibility, status, and notes all in one place.
 - *No deliverables or sensitive info stored there; purely tracking tool.*

SW Development Lifecycle

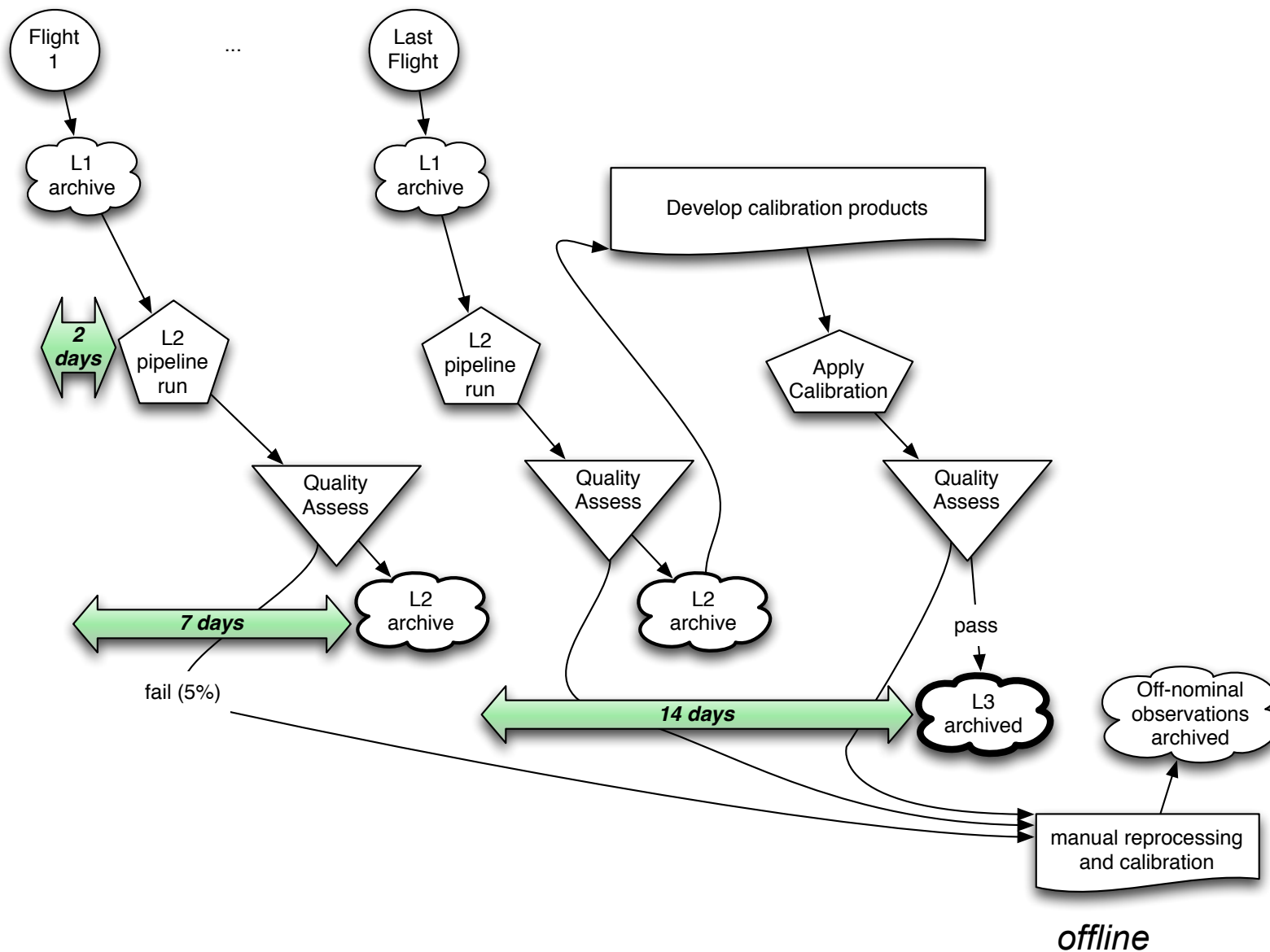


SW Engineering Lifecycle
 SOFIA Science Project SW Management Plan
 SCI-AR-PLA-PM20-2004 (Rev C.2)

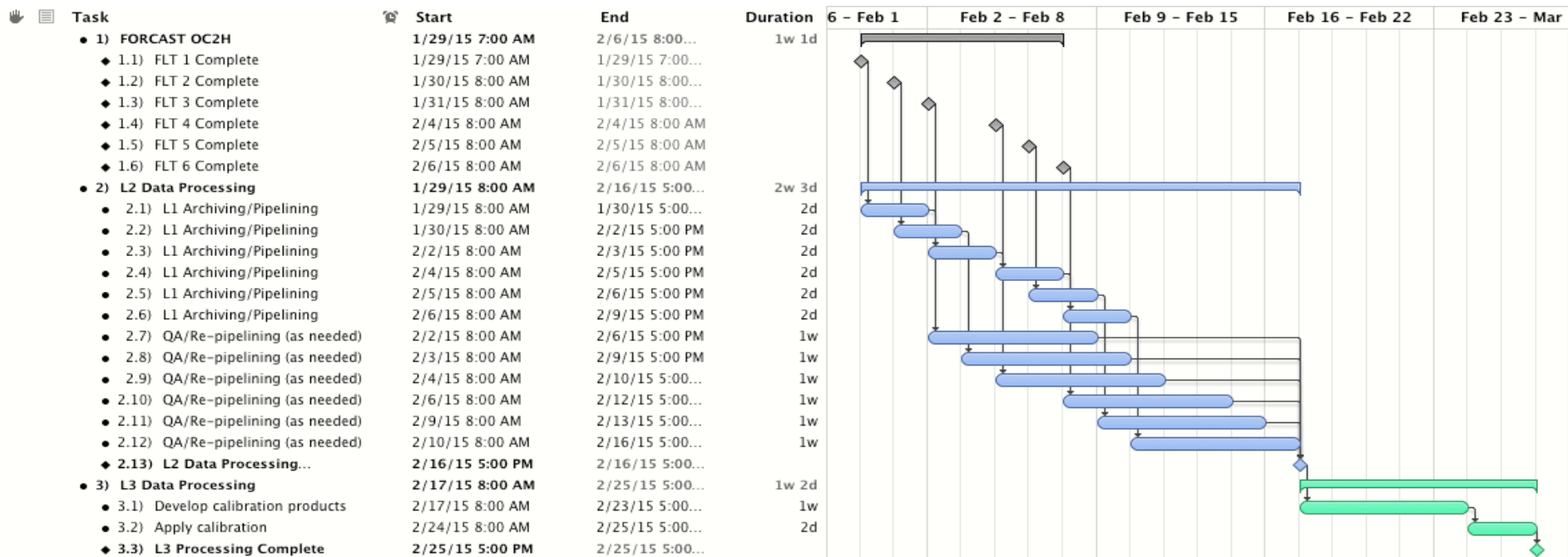
Development Lifecycle Schedule



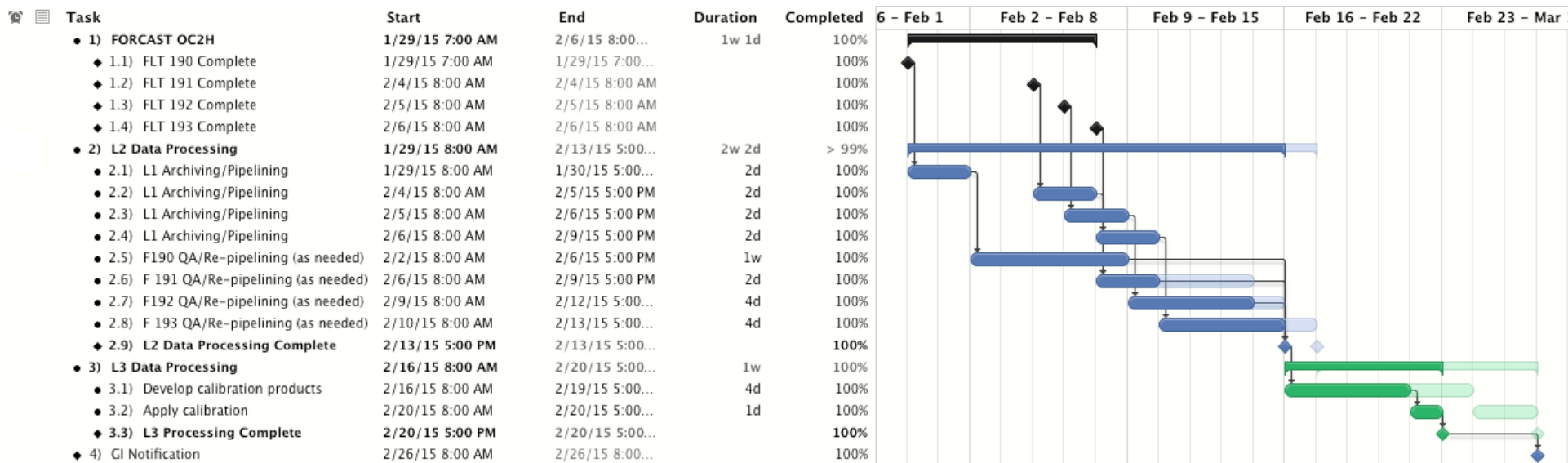
Data Processing Operational Flow



OC2H Processing Schedule (Planned)



OC2H Processing Schedule (Actual)



OC2H QA: Data Disposition

Flight	N_Obs*	Nominal (%)	Usable (%)	Problem (%)**	Fail (%)	Notes
190 Imaging	74	90	0	0	10	Focus issues.
190 Grism	13	75	0	25	0	Problems w/ Juno flux calibration.
191 Imaging	60	85	15	0	0	
191 Grism	24	79	0	21	0	Problems w/ Juno flux calibration.
192 Imaging	82	100	0	0	0	
192 Grism	28	60	0	30	10	Problems with Juno flux calibration; Juno not in slit.
193 Imaging	83	97	3	0	0	
193 Grism	18	76	0	24	0	Problems w/ Juno flux calibration.

* *All observations, including standards, GTO, DD, etc...*

** *No "PROBLEM" data distributed for GI programs; used for standards/GTO data only.*

No off-line reprocessing needed for OC2H.

Top 4 Data Processing Issues

1. **FIFI-LS pipeline and calibration procedure still in-work**

- *Supporting Christof Iserlohe (Stuttgart) on algorithm updates (flats, telluric correction, flux calibration, etc...).*

2. **QA is time-consuming**

- Trend is positive for FORCAST – OC2 D/F/H QA (FORCAST) took only a couple days per flight.
- Still an issue for EXES and FIFI-LS (Level 1)
- *See Backup Slides for QA process details*

3. **SI Configuration Changes**

- Changes in SI often cause changes to pipelines which require formal test and release according to NASA software management plan; can impact processing schedule.
- *Mitigated by new “beta release” procedure.*

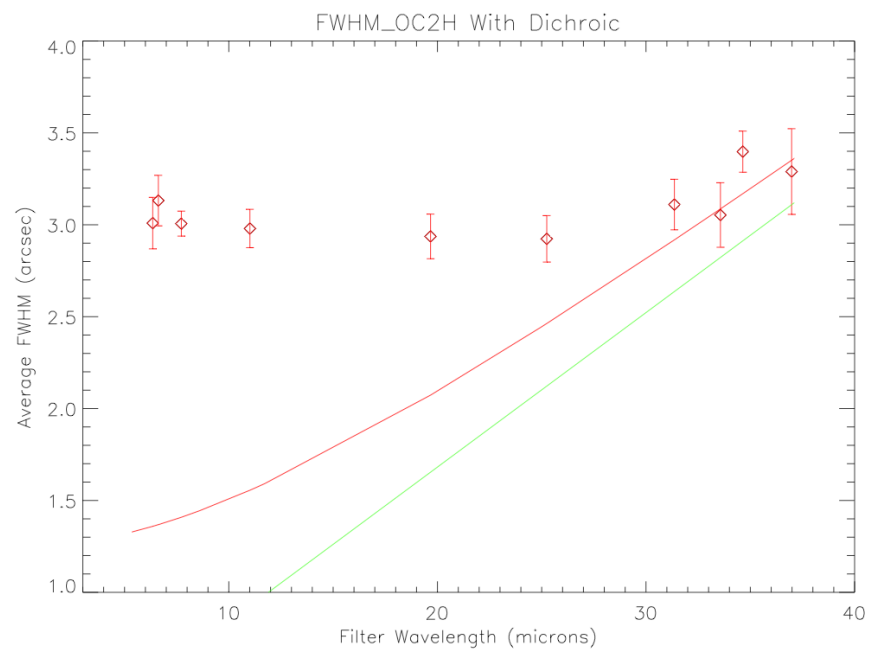
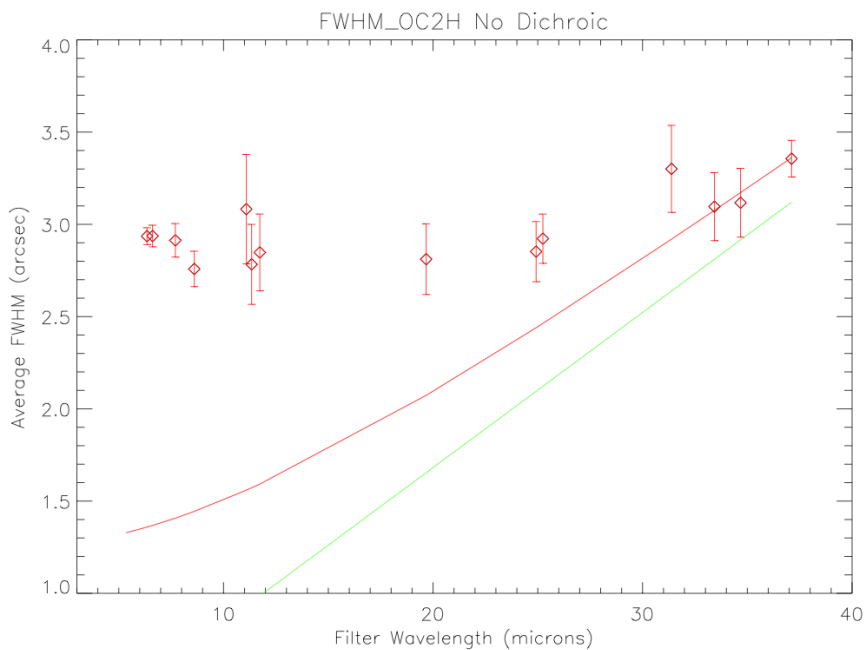
4. **Processing staff (~6 FTE) matrixed into other observatory activities (e.g. flight ops).**

- Short Term: Sometimes conflicts with processing schedule/deadlines.
- Long Term: Benefits pipeline operations due to increased familiarity with SI data and observing modes/strategies.



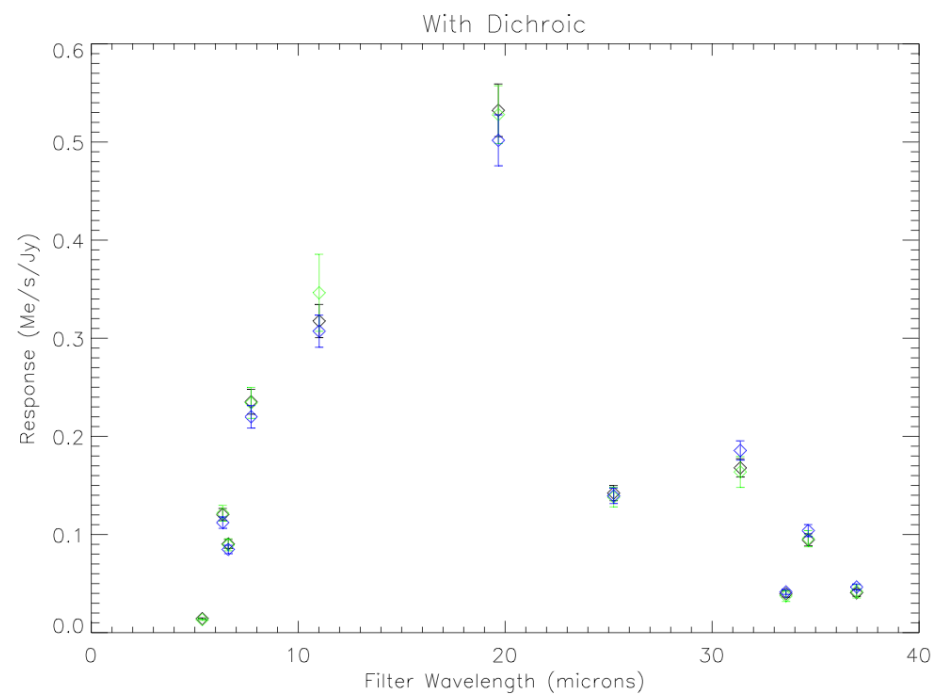
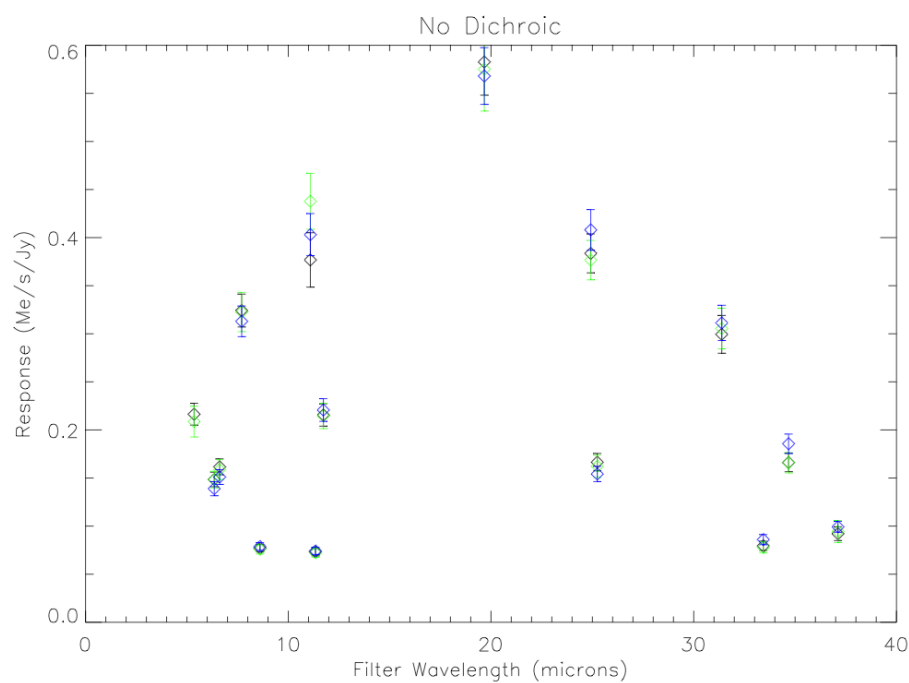
Backup Slides

FORCAST Measured Image Quality (OC2)



Output from flux calibration analysis.

FORCAST Imaging Response



Output from flux calibration analysis.

Recent (Independent) Results

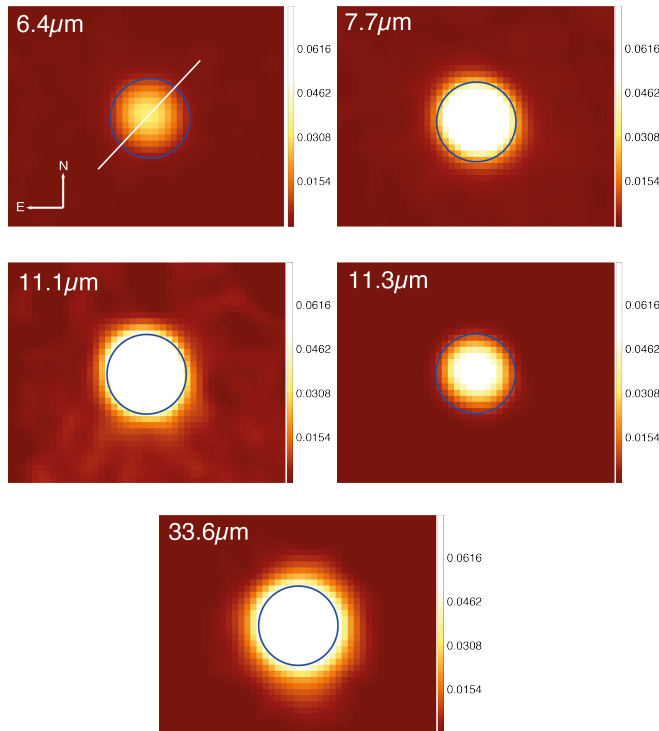
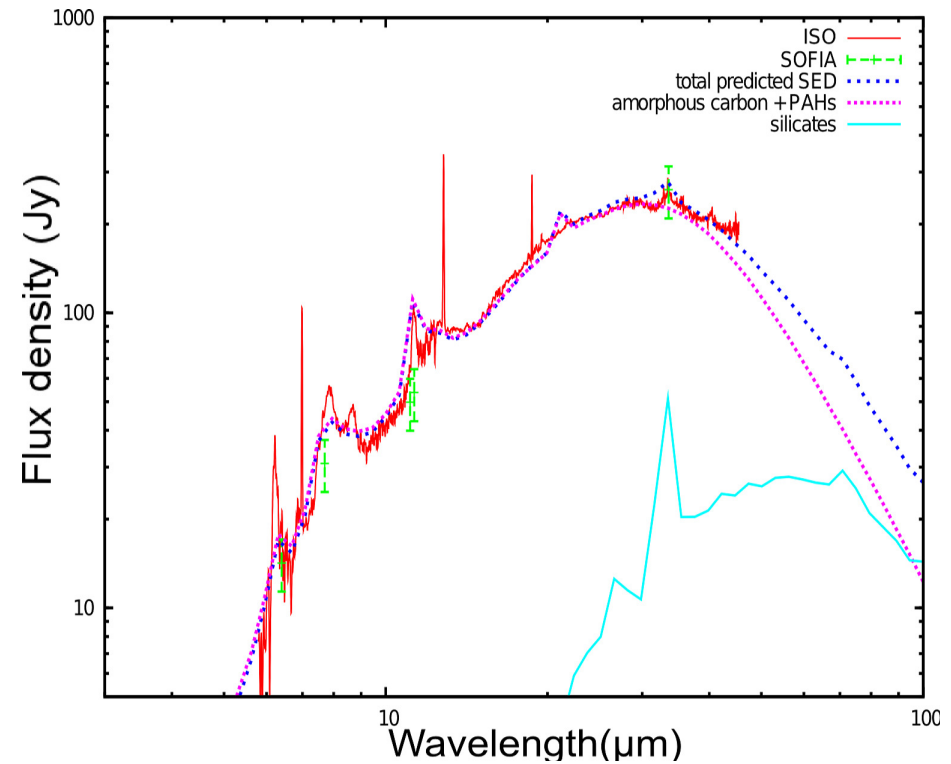
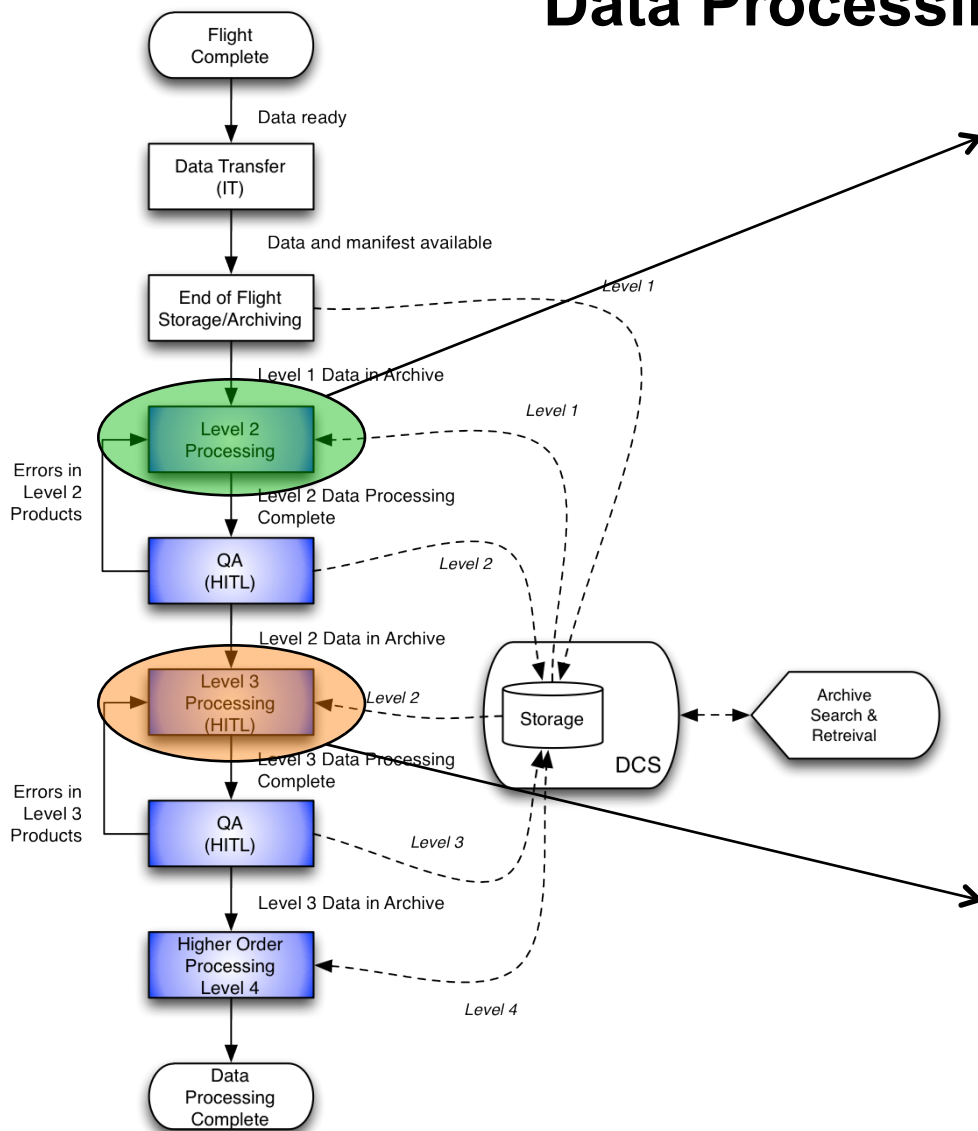


Figure 1. SOFIA images of BD +30° 3639 using the 6.4 μ m, 7.7 μ m, 11.1 μ m, 11.3 μ m, and 33.6 μ m filters. To compare the size of each image, we have overplotted a blue circle of 5'' radius in all of them. The white line in the top left panel represents the PA of the cut made to all the images.



Data Processing Flow



Two modes of Level 2 Processing:

- **Manual:** operator runs established version of pipeline interactively in stand-alone environment (workstation) on a single observation.
- **“Automatic”:** pipeline is run automatically on data for a whole mission. Some provision for user-interaction will be made.

Level 3 Processing will be highly user interactive, utilizing both COTS and custom tools/ pipelines.

Quality Assurance Activities

Level 1:

- Validate header keyword values against observing log
- Spot check data (based on observing logs) for anomalies
- All discrepancies/changes tracked in JIRA
- Update headers and submit to DCS archive (scripts, DCS tools)
- Updated files are stored as new “revision” in DCS; only *latest* revision is displayed in DCS archive.

Level 2:

- Inspect L2 products for issues outlined in Pipeline Users Manual (for all datasets, including commissioning/GTO) and assess overall data quality.
- Document discrepancies/issues and recommend fix for re-processing in QA log sheets.
- Update DATAQUAL keyword for all products.
- Submit to DCS for final storage/archiving.

Flux Calibration for FORCAST and FLITECAM Imaging

- Commissioned SI calibration plans are under change control.
- For OC2, a new plan for FORCAST was approved: one calibrator will be observed per flight (all filters)
 - *Effect on calibration accuracy still TBD.*
- For imaging, fluxes and wavelengths have been derived for each standard star for each filter using a comprehensive model of the instrument throughput and atmospheric transmission
- Corrections for differences in airmass, altitude, and pwv, between targets and standards have been derived from ATRAN models for each passband and incorporated into the calibration software
- Calibration parameters applied to Level 2 data to produce Level 3 products; calibration params also stored in archive for reference.
- All standards obtained in-flight are used for calibration.

Standard Star Selection

- For FORCAST, standard stars chosen from list of Herschel standards for which good models covering the FORCAST bandpass are available:
 - α Boo, α Cet, α CMa, α Tau, β And, β UMi, γ Dra, σ Lib
 - Asteroids (especially important for calibrating filter “blue leaks”)
- For FLITECAM, imaging standard stars chosen from Cohen et al. (2003) list of “Supertemplate” stars (~ 22 K-M giants with $K \sim 5-8$); grism standards are A0V’ s, as used for ground-based NIR spectroscopy

WCS Issues

- FORCAST
 - C2NC2 and NMC observations have large WCS discrepancies, apparently related to nodding. After correction, WCS much more accurate.
 - FORCAST control software will need update to correctly account for nodding (and other minor discrepancies).
- FLITECAM
 - WCS solution in headers appears accurate to $\sim 0.7''$ at ref pixel; $\sim 3''$ at field edge (due to coma and distortion).