Data Processing Status

R. Y. Shuping, *Data Processing Development Lead*

SSI/USRA-SOFIA

W. Vacca, *Data Processing Science Lead*

USRA-SOFIA

SOFIA Users Group   Apr 28, 2014

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., flats, darks, bad pixels)

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

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

SI Pipeline Readiness

SI Pipeline	Status
FORCAST Imaging	Automatic pipeline in operation.
FORCAST Grism	Automatic pipeline in operation.
FLITECAM Imaging	Ready; auto operation pending IT&V w DPS.
FLITECAM Grism	Ready; auto operation pending IT&V w DPS.
GREAT	Manual scripts in operation; Level 3 only.
EXES	alpha version complete and under test.
FIFI-LS	Preliminary pipeline received; alpha version complete; development currently on hold.
HAWC+	Preliminary pipeline for HAWC received; will need additional development for HAWC+.
HIPO	N/A

SOFIA Users Group   April 2014

OC2 Level 2 Pipeline Processing

FORCAST (Imaging/Grism):

- Chop/Nod subtraction
- Non-linear response correction
- Bad-pixel removal
- Droop & "jailbar" (cross-talk) correction
- Optical distortion correction (I)
- Field rotation/alignment (I)
- Flatfield correction (G)
- Wavelength calibration (G)

FLITECAM (Imaging/Grism):

- Nod subtraction
- Flatfield correction
- Bad-pixel removal
- Optimal spectral extraction (G)
- Wavelength Calibration (G)

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 ~ 5-8); grism standards are A0V's, as used for ground-based NIR spectroscopy

Processing Status

Series	Config	L1 ODC	L1 DCS	L2 Proc	L2 QA	L2 DCS	L2 GI	L3 Proc	L3 QA	L3 DCS	L3 GI
OC1B -- FORCAST 108--110, 6/2013	imaging grism						9/27/13 9/27/13				9/27/13
OC1C -- GREAT 115--123, 7/2013		12/31/13								2/26/14	
OC1D -- FORCAST 128--132, 9/2013	imaging grism	10/19/13 10/19/13					1/7/14 5/2/14				1/7/14
OC1F -- FORCAST 135, 10/2013	imaging grism	11/1/14 11/1/14					1/7/14 5/2/14				1/7/14
OC1G -- GREAT 140--143, 2/2014		3/13/14									
OC2A -- FLITECAM 146--149, 2/2014	imaging grism	3/13/14 3/14/14			4/25/14 5/2/14		5/9/14		4/25/14		5/2/14
OC2B -- FORCAST 146--149, 2/2014	imaging grism	3/13/14 3/14/14			4/25/14			5/2/14			5/9/14

Legend:
- Complete (blue)
- In Work: no issues (green)
- In Work: minor issues (yellow)
- Waiting (gray)

SOFIA Users Group   April 2014

OC1 Results

Processed 99 GI AORs (24 grism, 75 imaging)

- All flights auto-pipelined successfully*: <20min per flight.
- Auto-pipelines generally produced good results; but there were some failures (particularly for grisms in OC1B).
- Many datasets re-processed automatically to improve quality:
    - Nearly all grism observations for OC1B; 20% in OC1D/F
    - ~40% of imaging observations
- Handful (10) reprocessed manually.

QA: 7 AORs had generally poor results, despite re-processing.

Flux Calibration: (Imaging):

- Standards: alpha Boo, beta And, beta Peg, beta UMi
- 7 AORs had missing or bad standards for one filter
- Reproducibility is good (but limited dataset…)
    - Dispersion of cal factors *within* flight: <10%
    - Dispersion of cal factors *across* flights: <5%.

* Once raw FITS headers were corrected.

OC1 Processed Data Issues

General Issues:

- FITS header errors that require fixing before processing and rely on handwritten logs for correction.
- FORCAST Grism: no telluric correction or flux calibration.
  - *Will correct/calibrate once process is established.*
- FOR_G329 observations failed due to leak in blocker.
  - *Fixed for OC2*
- Accuracy of WCS is poor:
  - Incorrect coordinates used occasionally when updating TA pointing.
  - Issue with WCS keywords in raw data (chopper offset)
  - Error in FORCAST pipeline (v1.0.2): *now corrected*.
- OC1B: Incorrect normalization for large chop throws
  - *GIs notified (3); reprocessing in-work*

Isolated Issues:

- Chopper smear (short chop settle time) in OC1B;
  - *Corrected in OC1D/F.*
- Vignetting
- Some raw images with high backgrounds (wrong frame time); flux cal not attempted.
- Loss of pointing.

OC2 Status

OC2A (FLITECAM): SN observation and some GI programs

- Imaging processing done: data release imminent.

- Grism processing underway: data release NLT May 15.

- Known Issues:
    - Optical distortion has not been measured/corrected; WCS accurate to 0.7" at ref pixel, ~3" near field edge.

OC2B (FORCAST): GI Programs

- Imaging: L2 processing complete; QA and flux calibration in-work.

- Grism: waiting for wavelength calibration maps.
    - *Currently working with Luke to develop/check maps.*

Data Processing Delays

1. ***QA continues to take longer than expected.***
   - Still working out issues with raw data and processing algorithms.
   - *See Backup Slides for QA process details*
2. ***Instrument configurations are still settling.***
   - Changes in SI often cause changes to pipelines which require formal test and release according to NASA software management plan.
3. ***FORCAST Grism calibration still in-work.***
   - DPS Team working with grism PI (Keller)*.*
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.

SOFIA Users Group   April 2014

Development for OC2

Pipelines:
- FORCAST: Updates to improve integration with automated system (DONE); need G3xG4 wave cals (IN WORK)
- FLITECAM: Finalize using commissioning results (DONE).
  - IT&V with automated system needed (IN WORK).
- FIFI-LS: Develop/test alpha for commissioning (DONE)
  - testing with commissioning data still TBD (ON HOLD)
- EXES: Develop/test alpha for commissioning (DONE)

Infrastructure upgrades:
- Adding dedicated test environment (DONE)

New capabilities:
- Metadata/reporting subsystem (ON-HOLD)
- Re-processing tools (ON-HOLD)
- Complete FORCAST Grism telluric/flux calibration procedure (SUMMER)
- Develop FLITECAM Grism telluric/flux calibration procedure (SUMMER)

SPR maintenance/upgrades: on-going.

Backup Slides

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.

Data Processing Flow
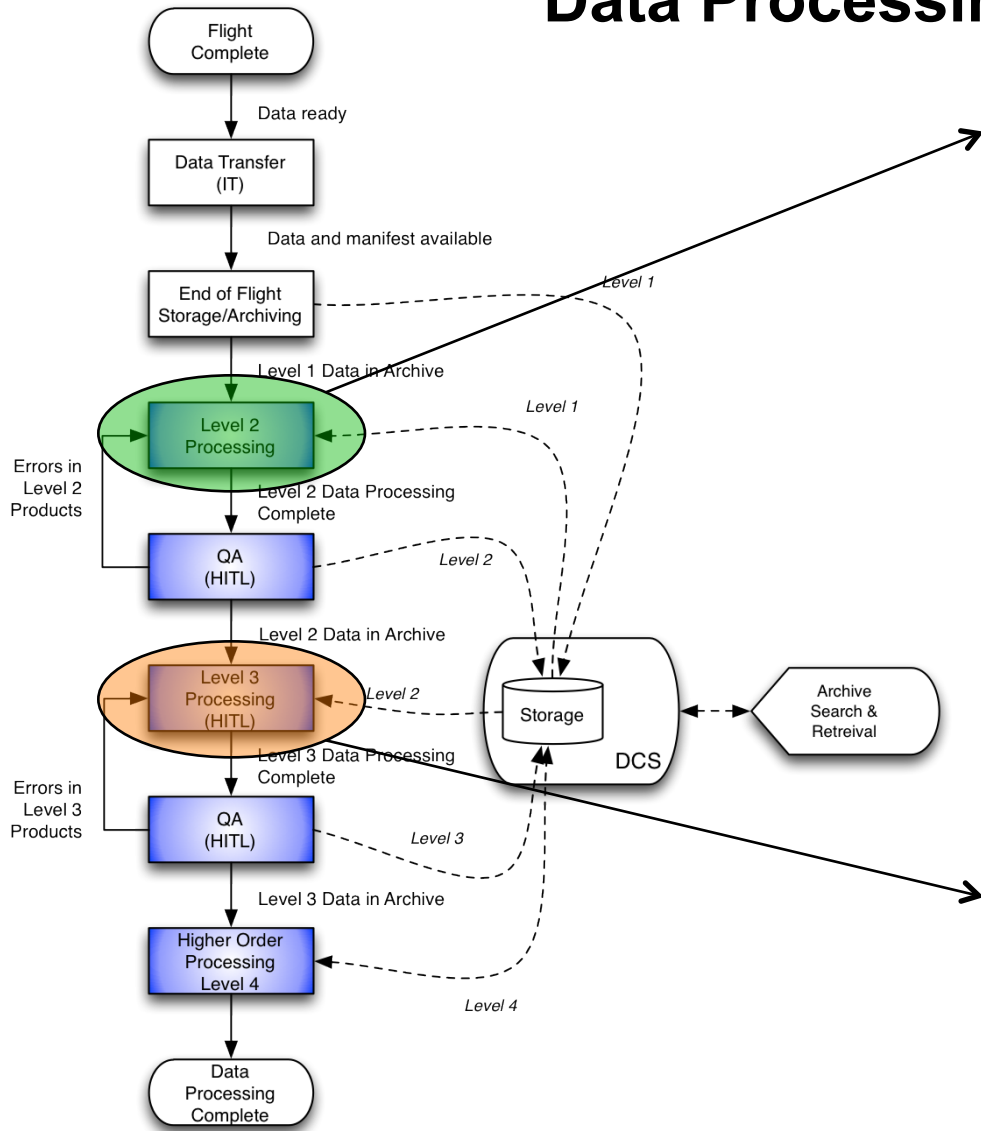
Two modes of Level 2 Processing:

- **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.