Data Processing Status

R. Y. Shuping, Data Processing Development Lead
SSI/USRA-SOFIA

W. Vacca, Data Processing Science Lead
USRA-SOFIA

SOFIA Users Group Meeting Oct 20, 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

<table>
<thead>
<tr>
<th>SI Pipeline</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>FORCAST Imaging</td>
<td>Automatic pipeline in operation.</td>
</tr>
<tr>
<td>FORCAST Grism</td>
<td>Automatic pipeline in operation.</td>
</tr>
<tr>
<td>FLITECAM Imaging</td>
<td>Ready; auto operation pending IT&amp;V w DPS.</td>
</tr>
<tr>
<td>FLITECAM Grism</td>
<td>Ready; auto operation pending IT&amp;V w DPS.</td>
</tr>
<tr>
<td>GREAT</td>
<td>Manual scripts in operation; Level 3 only.</td>
</tr>
<tr>
<td>EXES</td>
<td>alpha version complete and under test.</td>
</tr>
<tr>
<td>FIFI–LS</td>
<td>Preliminary pipeline received; alpha version complete; <em>bug–fixes and algorithm updates underway</em>.</td>
</tr>
<tr>
<td>HAWC+</td>
<td>Preliminary pipeline for HAWC received; will need additional development for HAWC+.</td>
</tr>
<tr>
<td>HIPO</td>
<td>N/A</td>
</tr>
</tbody>
</table>
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.
Since last SUG (April 2014)

- Completed all Level 2 processing for OC2.
- Completed all Level 3 processing for OC2 imaging.
  - Accuracy of <5% for both FORCAST and FLITECAM
- Development team won Ames Honor Award
- Lost Miguel Charcos-Llorens, pipeline scientist/engineer (currently seeking replacement).
- Presented 3 poster papers at ADASS 24 (Calgary)
- Completed FORCAST “browse quality” calibration procedure (see Vacca’s talk)

Data processing team has produced 220 GB L2/3 data (>56000 files) for OC1 and 2.
Remaining L3 products for FORCAST Grism Mode.
Remaining L3 products for FORCAST Grism Mode.
FORCAST Grism Flux Calibration

- Support for telluric correction and flux calibration added to FSpextool for FORCAST, with 'Flux Calibrate' step added to the Redux interface.
  - OC1 and G2xG1 response curves in work.
- Verification: Tested on all observations of alpha Boo in OC2 – verified against model of alpha Boo.
- Release of FORCAST-Redux pipeline complete
- Readiness Review passed on Oct 17.
- OC2 processing to start immediately with OC1 processing to follow ASAP.
- Expect FORCAST Grism Level 3 data to be available to GIs (via DCS archive) by end of October.
Top 4 Data Processing Issues

1. QA is time-consuming…
   - …but trend is positive – OC2 D/F QA was very quick.
   - See Backup Slides for QA process details

2. 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.

3. 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.

4. FIFI–LS pipeline and calibration procedure still in–work
   - Will likely need to process initial FIFI–LS flight series manually.
WCS Issues

- **FORCAST**
  - Investigation of OC2 observations is on-going (F167 and F178).
  - 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 ~0.7” at ref pixel; ~3” at field edge (due to coma and distortion).
Development for FY2015

Pipelines:
- FORCAST:
  - OC1 grism response curves (IN WORK)
  - New wavecals for G4 (need on-sky observations)
- FLITECAM:
  - Complete IT&V with automated system (IN WORK).
- FIFI-LS:
  - Testing/bug-fixes with commissioning data (IN WORK)
  - Algorithm updates
  - IT&V with Processing System
- EXES:
  - IT&V with Processing System
  - New observing modes (TBC)

New capabilities:
- Metadata/reporting subsystem (requirements/prototyping IN–WORK)
- Additional Re-processing Functionality (IN–WORK)

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.
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:
  - $\alpha$ Boo, $\alpha$ Cet, $\alpha$ CMa, $\alpha$ Tau, $\beta$ And, $\beta$ UMi, $\gamma$ Dra, $\sigma$ Lib
  - Asteroids (especially important for calibrating filter “blue leaks”)

- For FLITECAM, imaging standard stars chosen from Cohen et al. (2003) list of “Supertemplate” stars ($\sim$22 K–M giants with $K \sim 5$–8); grism standards are A0V’s, as used for ground-based NIR spectroscopy