Data Processing for Observing Cycle 1

R. Y. Shuping, Data Processing Development Lead

W. Vacca, Data Processing Science Lead
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 (e.g. BSCALE/BUNIT keywords, MJy/sr)

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

- Provide science-grade data products to GIs and science community;
  - process/analyze data;
  - quality assurance;
- Curate processed data in the archive and re-process as needed.
- Develop pipelines and other needed software tools.
- Support SI teams on pipeline development/delivery.
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.
System Context

SOFIA Science Center
Data Processing System

DRAFT Context Diagram
DO NOT DISTRIBUTE

Pipeline Operator/Scientist
Processing/Grouping Parameters
Processing Status

Data Processing System
CI: SI Processing Tools
SI Pipeline
Analysis Tools
COTS
CI: Pipeline Tools
DCS Core Services

Data Cycle System
DCS Core Services
DB Management Apps
DBs
Persistent Store

Processed/Validated Datasets Datamanifests
Processed Data Notifications (?)
Query Params

Requires access to persistent store

See Shuping et al. (ADASS 2012) for more on the DCS.
Data Processing System by instrument

- **FORCAST imaging**: Ready for operations
- **FORCAST grisms**: Operational; commissioning updates in–work
- **FLITECAM imaging**: Ready for operations
- **FLITECAM grisms**: Operational; waiting for additional wave cals.
- **GREAT**: Learning from Basic Science Level 3 deliveries (2013 Mar)
- **FIFI–LS**: Preliminary pipeline received; agreement with SI team to revisit
- **EXES**: Negotiating with SI team to obtain source code in summer
- **HAWC+**: HAWC pipeline installed and tested; tabled until HAWC+ ready
- **HIPO**: No need for reduced or calibrated products
OC1 Level 2 Pipeline Processing

- **FORCAST (Imaging/Grism):**
  - DRIP/FG
  - 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)
  - Optimal spectral extraction (G)
  - Wavelength calibration (G)

- **FLITECAM (Imaging/Grism):**
  - FDRP/FSpextool
  - Nod subtraction
  - Flatfield correction
  - Bad-pixel removal
  - Optimal spectral extraction (G)
  - Wavelength Calibration (G) (TBD)
Flux Calibration for FORCAST and FLITECAM

- On each flight, observations of standard stars will be obtained at each altitude (as in Basic Science).
  - See backup slides for standard star selection
- 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.
FORCAST Basic Science Flux Calibration Results

- Method worked well for calibrating FORCAST Basic Science data (Herter et al. 2013, in press)
  - RMS Response Uncertainty: ~6.1% (flight-by-flight)
  - Anecdotal results for a few Basic Science targets indicate that the calibration is good to <20% (3 sigma)
  - Currently analyzing standards from Basic Science to evaluate overall accuracy.

Herter et al. (2013)
Agreement with existing datasets is good.
Data Distribution Timeline

- **Level 2 data** available to GI within **2 weeks** of flight series completion
  - *OC1B Exception*: delivery within 4 weeks
- **Level 3 (Flux Calibrated) data** available to GI within **4 weeks** of flight series completion
  - *OC1B Exception*: delivery within 6 weeks
- Data distribution via SOFIA Archive
  - GI data released to public after 1–year (from archive insertion).
  - Standards and calibration parameters released immediately.
System Development Status

- Requirements and design reviews complete
- Development hardware/network ready and in-use
- Production hardware/network almost ready, awaiting formal test
  - Dev environment serves as backup
- V&V with commissioning data underway
- Development of standard operating procedures underway
- Formal system testing coming up in prep for Cycle 1