

Progress in Constructing the SOFIA Data Center

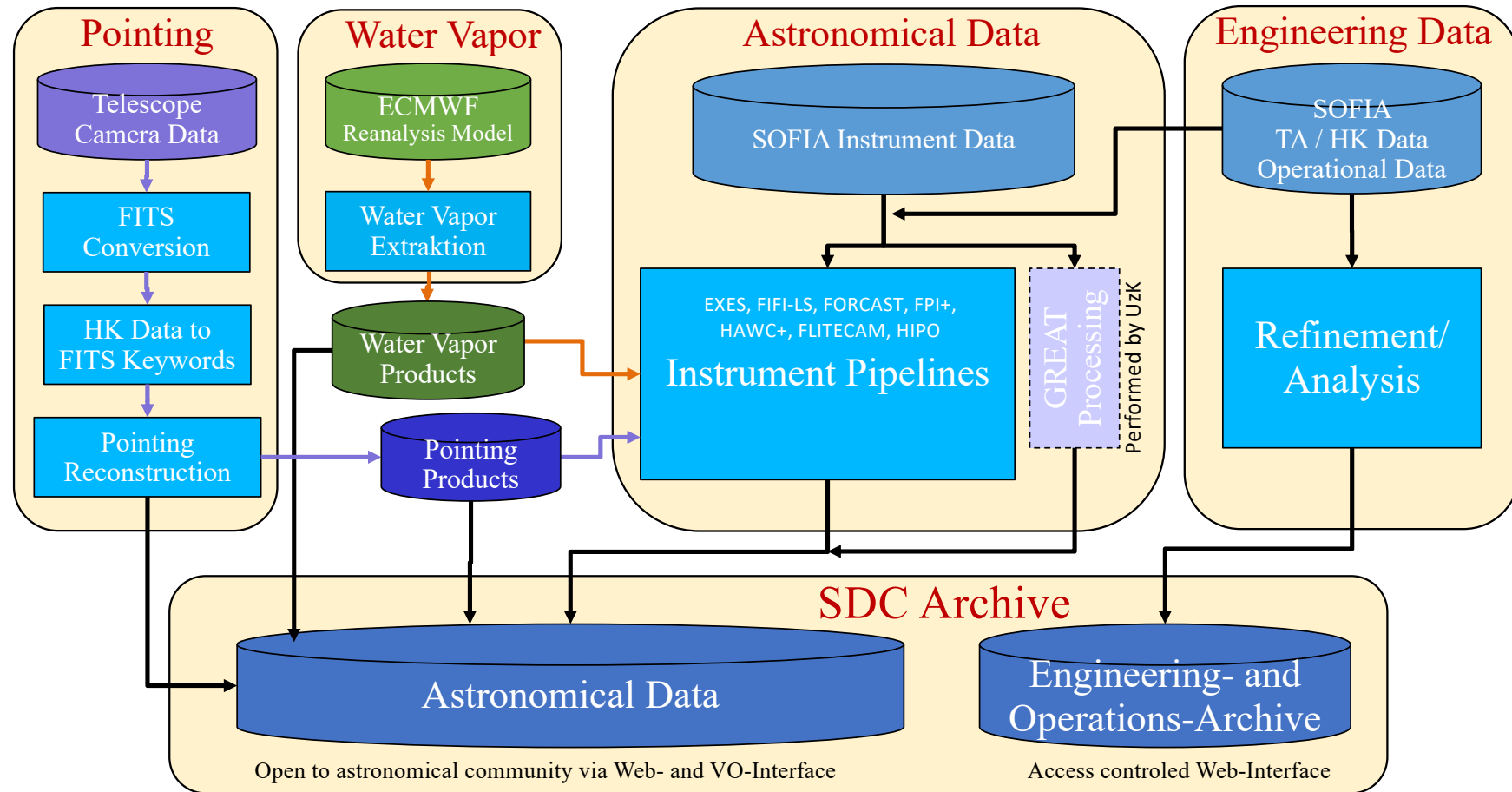
B. Schulz, A. Bryant, J. Früh, B. Greiner, M. Hütwohl, C. Iserlohe,
T. Keilig, B. Knieling, K. Schindler, A. Siggelkow, M. Wiedemann, O. Zeile

SOFIA Data Center, IRS, Universität Stuttgart

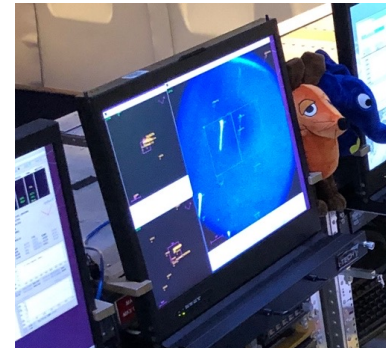
- The SOFIA Data Center started nominally 1. July 2024
- Goals
 - Create scientific resource for the time after the mission
 - Bridging expected 10-20 year gap in new FIR Data availability
 - Ensure quality of SOFIA data archive to secure scientific productivity
 - Enable data exploitation without need for expert knowledge
 - End June 2029 with subsequent archive transfer to DZA
 - Stay in sync with current SOFIA data repository at IRSA
- 2024 DSI wrap-up activities, data capture, project workshop, relocation from the US (LA County, Bay Area) to Stuttgart
- Beginning 2025 almost whole team in Stuttgart

- Engineering- and Operations-Archive
 - Ingest loose documents in unconfigured part
 - Transfer Data from Windchill to DaRUS
 - Configure user interface
 - Prepare User's Handbook
- Telescope Camera Imager Pipeline
 - Convert images (WFI, FFI, FPI+) to FITS Format
 - Add relevant header information from MCCS and TA housekeeping data
 - Provide WCS from plate solving
- Water Vapor Calibration
 - Pulled Water vapor levels from ECMWF ERA5 atm. reanalysis model
 - Atm. transmission correction in pipelines
- FIFI-LS Recalibration
 - Update response curves
 - Photo/Spectrometric resampling
- Pipeline Improvements and Data Reprocessing
 - (GREAT done in Cologne)
 - FIFI-LS (current main focus)
 - HAWC+
 - FORCAST
 - EXCES

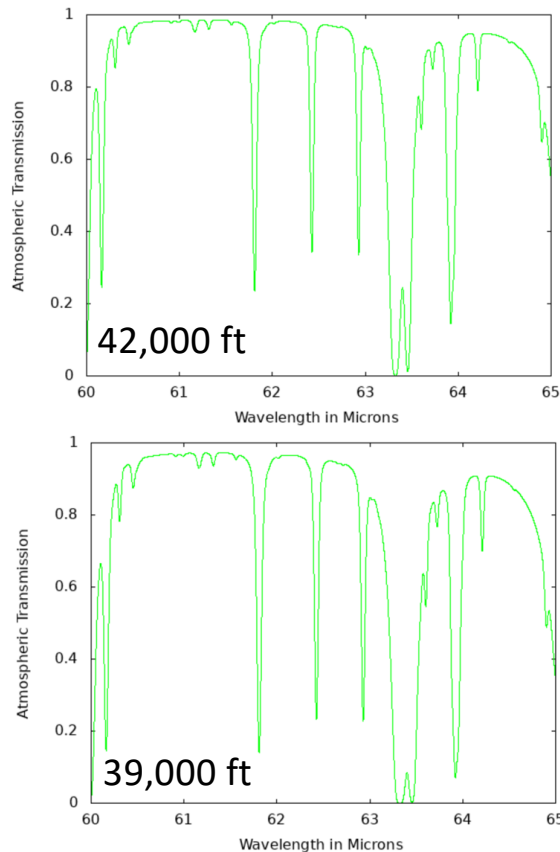
Reprocessing Workflow



- Pipeline in development
- Guide Cameras
 - Wide Field Imager – WFI
 - Fine Field Imager - FFI
 - Focal Plane Imager – FPI+
- ~90,000 images per flight, >700 flights
- Processing levels
 - Level 0: Extract from ark to FITS format, only few keywords included
 - Level 1: 280 Keywords from SOFIA housekeeping data added including WCS
 - Level 2: Updated WCS using plate solving algorithms
- Difficulty levels for plate solving
 - Un-chopped images
 - Chopped images
 - HAWC+ Lissajous-scans
- Final products:
 - Level 2 processed files
 - Lists with timestamped pointing deviations



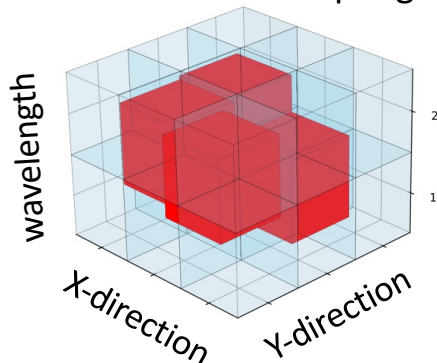
- Internal development at **GitHub Enterprise Instance of Uni Stuttgart**
 - Guest accounts for outside collaborators available on request.
- Public at: https://github.com/SOFIA-Data-Center/sofia_redux
- Forked from: https://github.com/SOFIA-USRA/sofia_redux and pushes from the internal repository
- **GitHub.com contributions** through issues or pull requests **welcome** and will be synced to internal
- Version releases will update existing PyPI package sofia-redux
 - currently outdated 1.3.3 from 2023
- **Unit tests** continuously checked with GitHub Actions CI/CD workflow (internal runner)



- Calculates spectral atmospheric transmission for Near and Far-IR Radiation (S. Lord, 1992)
- Essential for instruments FORCAST, FIFI-LS, HAWC+
- Uses own database of HITRAN line parameters from ~30 years ago
- Error found from HITRAN parameters in $^{17}\text{O}^{16}\text{O}$ isotopologue
 - Predicts too much absorption at certain wavelengths
 - Seems to be the only parameter with substantial issue
- Correction in collaboration with S. Lord
- Further plans
 - Update all other HITRAN parameters to newest values
 - Goal to set up public ATRAN server again
 - Original ATRAN server was switched off by NASA during the wind-down of the SMO
 - Allow Ozone as free parameter

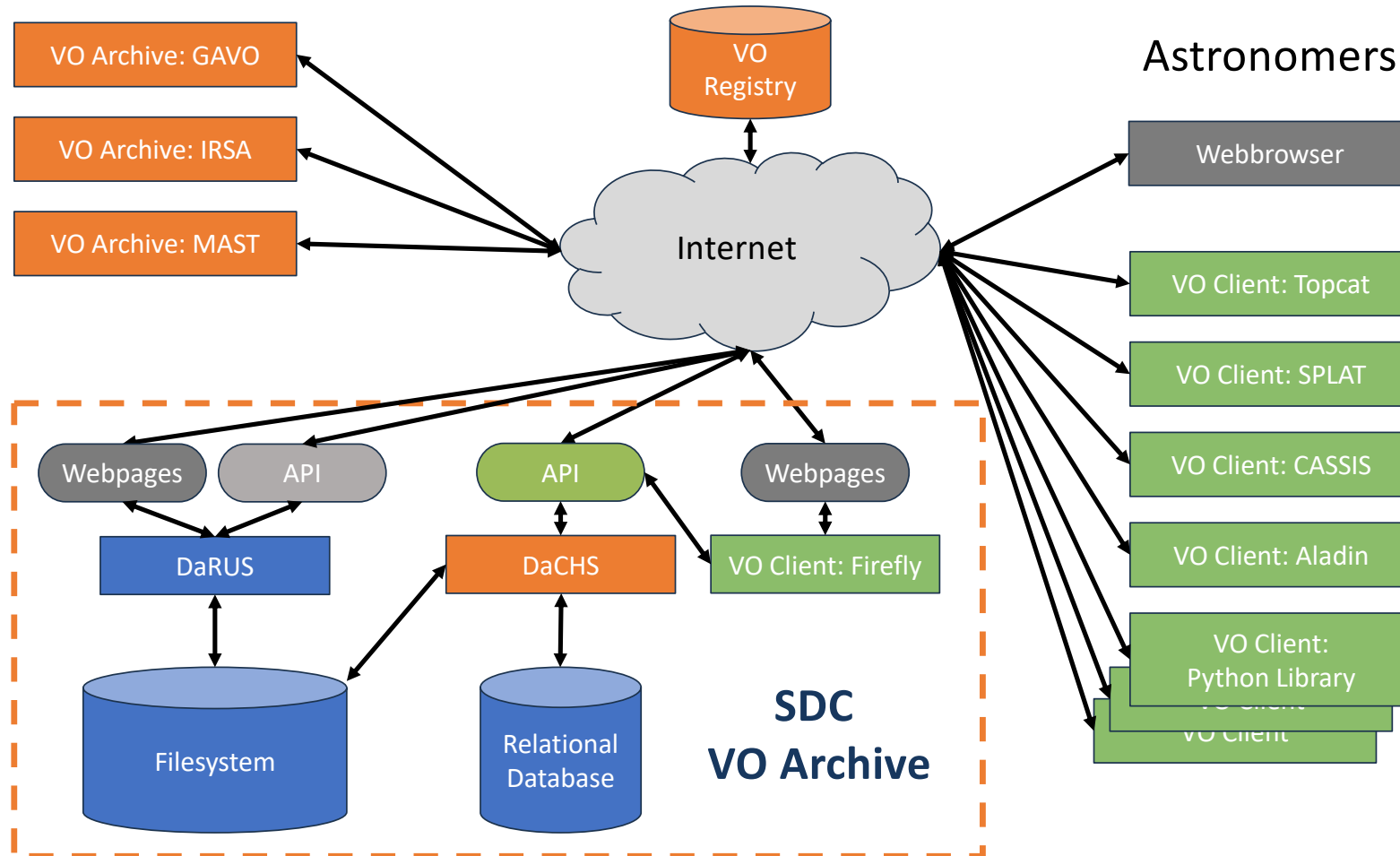
- Main pipeline goals:
 - Raw data consolidation (bad files)
 - Pointing stability and error checking
 - Full implementation of ECMWF telluric correction
 - Correction/recalculation of ATRAN atmospheric transmission model
 - Improved Flux calibration
 - Update groupings of Level 4 products
 - Combinations of observations
- Issues with current calibration
 - Several calibrators: Mars, Uranus, Callisto (+/- 15%)
 - Fadda et al. 2023
 - Limited calibration dataset (28 flights) under inhomogeneous conditions
 - Single Mars model scaled with distance
 - Ad-hoc scaling factors between flights
 - 3D interpolation in spectral data cube

Data Cube Resampling



Fixes

- Use Mars Model from E. Lellouch calculated for each observation
- New Drizzle-based interpolation in spatial and spectral direction ensures flux preservation
- Take advantage of corrections to ATRAN model
- Overall response curve corrections expected between 1% - 12%



- Good progress in all areas
- Additional effort for FIFI-LS recalibration necessary
- Delays in building data reprocessing, ATRAN server and archive setups
- Operations and Technology Archive on track to finalization this year
- Water vapor extraction from ECMWF model complete
- Next work packages:
 - DaCHS installation, Setup automatic reprocessing system, US-instruments, Identification of scientifically important data sets



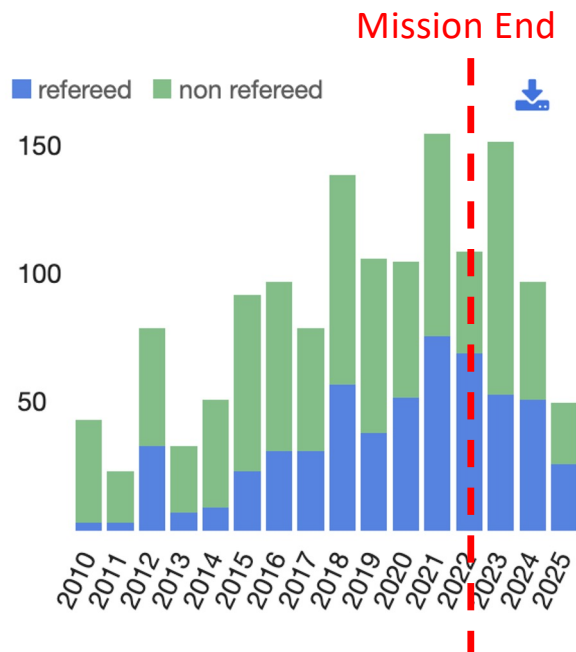
29.7.2023, PIMA Air and Space Museum, Tucson

EXTRA SLIDES

- Numbers
 - 45 pull requests merged
 - 150 commits with 8,930 additions and 3,649 deletions.
- Software compatibility
 - at least:
 - Python 3.11 to 3.13
 - NumPy 1.24.3 to 2.2
 - SciPy 1.9.3 to 1.16
 - Pandas 1.5 to 2.3
 - Matplotlib 3.3 to 3.10
 - Astropy 6.0 to 7.1
- Pending updates:
 - DS9 interaction
 - GUI framework: PyQt5 --> PyQt6
 - Pip based installs only, no Conda support from SDC
- FIFI-LS updates
 - Updated data package for bad pixels, flatfields, response curves
 - Interpolation between ATRAN spectra
 - New feature: Telluric Scaling
 - New feature: Pointing discard

REFEREED

☐ non-refereed 850
☐ refereed 562

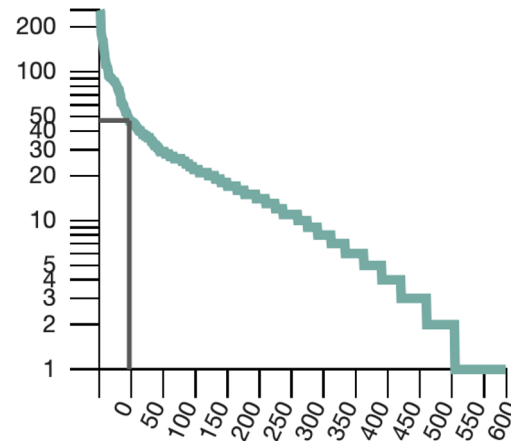


Cited

1,410 top cited records with 11,585 total citations

H-Index for results: 47

Y-axis: linear ☐ log ☒

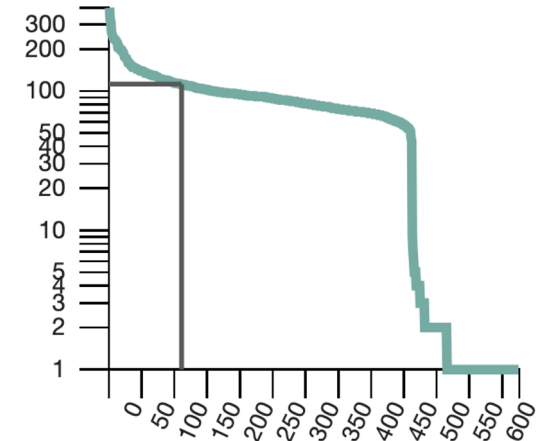


Reads

1,410 top read records with 46,399 total recent reads

H-Index for results: 111

Y-axis: linear ☐ log ☒



Source: Private ADS-Library B. Schulz

Data Reprocessing

https://github.com/SOFIA-Data-Center/sofia_redux/

https://pages.github.tik.uni-stuttgart.de/IRS-SOFIA-Data-Center/sofia_redux/

