Plans For the STScI STIS Pipeline I: Overview

Stefi Baum and Phil Hodge
July 1995

ABSTRACT

In this ISR we describe the overall design of the STScI STIS calibration pipeline, from generic conversion through the STSDAS calibration software, `calstis`*. Included are a description of the pipelines's goals in priority order, a discussion of the need for 'contemporaneous calibration' (a new aspect of the STIS pipeline relative to the existing ones), and a description of the output products from the STIS pipeline, including a brief description of the FITS image extension files to which the science data will be written. In addition the calibration switches and files used to control the calibration of data in the pipeline are presented, the high level design of the* `calstis` *software is described, and the flow of different types of STIS data through the STIS pipeline are illustrated.*

1. Introduction

 In this ISR we describe the overall design of the STScI STIS calibration pipeline, from generic conversion through the STSDAS calibration software, `calstis`. Subsequent ISRs will describe the algorithms used to calibrate STIS data (see STIS ISR 95-007 for a description of two dimensional image reduction through flat fielding by `calstis`).

This document is organized as follows:

- Section 2 presents the "Goals of the STScI STIS Pipeline", page 2
- Section 3 describes "Contemporaneous Calibrations", page 3
- Section 4 discusses "Associated Exposures and STIS Datasets", page 4
- Section 5 describes "STIS Output Products", page 6
- Section 6 discusses "Science Data and FITS Image Extensions", page 9
- Section 7 presents the "Calibration Switches and Reference Files", page 12
- Section 8 describes the "High Level Design of `Calstis`", page 17
- Section 9 describes "Data Flow Through `Calstis`", page 26

2. Goals of the STScI STIS Pipeline

STIS is a versatile instrument capable of filtered imaging, first order and echelle spectroscopy (See the STIS Mini-Handbook 1995, for more information about the instrument). STIS's CCD operates from the near IR through the near-uv (2500-10000A) using a CCD. Two MAMA detectors, a Cs_2Te detector (NUV-MAMA) operating from 1650-3100 A and a CsI detector (FUV-MAMA) operating from 1150-1700A operate through the ultraviolet. STIS produces time integrated image data, when used in ACCUMULATE operating mode and an x,y,t stream when used in TIME TAG operating mode. STIS can produce images with the full pixel dimensions of the detector; it can also be used to bin pixels in x or y, and it can be used to read out from only a specified 'subarray' of the detector pixels.

The STScI STIS Pipeline (trans->generic conversion->calstis->archive) has as its goals in order of priority:

1. production of STIS datasets (Section 4 and Section 5) as the output of generic conversion, with all science data in FITS image extension files or FITS binary tables

2. two dimensional image reduction through flat fielding and photometric calibration of all data taken in ACCUM modes for GO configurations (prism data is the only exception - it will be processed through two dimensional image reduction but will not be flatfielded)

3. spectroscopic 2-D calibration (application of two-dimensional dispersion solutions, correction for pixel to pixel distortions, assignment of absolute wavelengths, and conversion from counts to flux units) of all ACCUM mode spectroscopic data taken in GO configurations (*excluding* prism data),

4. extraction of a one dimensional, background subtracted spectrum for all spectral observations taken using 'short' slits (<1" in the cross dispersion direction) and for all spectral point source observations taken using long slits, and production of a two dimensional rectified spectrum, with linear axes in dispersion and distance along the slit) for all long slit spectral observations taken in ACCUM mode,

5. cosmic ray rejection for crsplit observations and geometric correction of all imaging data taken in GO configurations,

6. production of a time integrated uncalibrated (but Doppler corrected as appropriate) *image* from TIMETAG operating mode data which is then passed through standard ACCUM mode reduction (as above)

7. production of a calibrated *time tag data stream* for data taken in TIMETAG operating mode which has been passed through the first phase of image reduction (through flat fielding for spectroscopic data, through photometric calibration for imaging data).

Calibration of data taken in Engineering Only modes will be performed only for those configurations which do not require additional calibration reference files or additional

software logic. Observations which do not satisfy this constraint will be processed through generic conversion and written as uncalibrated FITS data files to the archive. All exposures taken in Engineering Only mode will have the ENGMODE keyword in the science data header set to T (true) - the default value of this keyword is F (false) for data taken in GO supported configurations.

3. Contemporaneous Calibrations

The STIS pipeline will perform three types of contemporaneous calibration:

- determination of the zero point of the wavelength scale (center of the spectrum in wavelength space) and location of the aperture (center of the spectrum in sky space), using a line lamp exposure (wavecal)
- removal of cosmic rays from crsplit CCD exposures
- determination of the target location in a long slit, using the target acquisition images.

To support contemporaneous calibration, the pipeline must be able to link together separately commanded exposures. The mechanism for this linkage and the implications for the data produced for the GO and archive are described in more detail in Section 4, page 4.

wavecals

On STIS, the optical path from source to detector passes through the aperture (slit) wheel (where the filters are housed for imaging) and then through the Mode Select Mechanism which houses the first order gratings, the cross dispersers for use with the echelles and the mirrors for imaging work. Irrepeatability in the MSM causes the center of the spectrum (as defined by the aperture and wavelength centers) to fall on a slightly different location in the detector each time there is a movement of the MSM (the MSM induced offsets in x and y have been measure in component testing to be roughly plus or minus 2 pixels). In addition, for MAMA observations, the aperture location on the detector will be deliberately shifted each month, to ensure equalization of extracted charge across the detector.

To allow calibration of the zero point of the aperture location and the zero point of the wavelength scale for spectroscopic observations, a line lamp observation (so called *wavecal*) will be taken each time the MSM is moved or each new orbit. The wavecal observations for first order modes will be taken with the 'HITM' (Hole in The [CM-2] Mirror) system, which passes the light from the sky along with the light from the lamp - therefore the wavecal exposure itself must be processed to remove the source spectrum before it is used to calibrate the science data.

crsplits

Observers using the STIS CCD will want to crsplit (break into two or more separate exposures) observations over a certain duration, to allow identification of cosmic rays which

will otherwise compromise their images and spectra. The STScI STIS pipeline will process crsplit exposures through the pipeline together, producing a single cosmic ray rejected image from the exposures, after they have been individually processed through flat fielding. The single crsplit image is then passed through the subsequent calibration steps (e.g., photometric calibration, geometric correction in the case of images, spectral calibration and extraction of a one dimensional spectrum, see Section 9, page 26).

target acquisitions

The flight target acquisition software will center the acquisition target in short slits in both the dispersion and cross dispersion direction. However it will not center a target in the cross dispersion direction in a long slit. Due to the MSM uncertainty, this means that the location of a target in the long slit will be known only to +- 2 pixels (the MSM uncertainty). For observations of stars (with sufficient signal), there is no impact because the alignment is obvious between the two sets of exposures and a one dimensional spectrum will be extracted as part of the pipeline in any case. For observations of extended diffuse targets, however, or low signal stars, this means that the absolute location of the target in the slit is known only to 2 pixels, and that relative centering of spectra taken using different target acquisitions will not be known to better than ~ 4 pixels.

To allow inter-observation comparisons for long slit observations of extended objects, it will be necessary for the pipeline to locate the offset between the aperture center on the detector in the target acquisition exposure and the predicted aperture center (which is equivalent to determining the target location within the long slit). This information is needed to correctly populate header keywords in the subsequent science observations so that cross-observation comparisons can be made. The derivation of these two header keywords will be performed during generic conversion and the values placed in the header of the uncalibrated science dataset for all datasets to which the target acquisition applies. [1] Note that in addition to being utilized in generic conversion for the science exposure data, the target acquisition exposure themselves will be processed through calstis, and archived as their own datasets (see below).

4. Associated Exposures and STIS Datasets

As described above (see Section 3, page 3) the STIS pipeline will need to process separately commanded exposures together through calibration in order to produce the final calibrated science output product.   By definition, an 'associated set' is that set of expo-

---

1. The target acquisition exposure will automatically be linked with the concomitant science exposures by TRANS, so that it can be utilized during the generic conversion processing of the science data.

sures which are processed through the pipeline together (e.g., one associated set might be composed of a wavecal exposure and three science exposures). For STIS we make a further distinction between those linked exposures needed by generic conversion, which includes the target acquisition exposure for long slit spectroscopic observations, and those needed only by the calibration portion of the pipeline (i.e., `calstis`), which does not include the target acquisition exposure.

To simplify things for users, the STScI STIS pipeline will package all associated exposures *needed by `calstis`* into a single dataset, where by definition a dataset is a set of files that share a common rootname (have the same first 9 characters of their file name).[2] Thus each STIS dataset will contain all the inputs, exclusive of the static calibration reference files from CDBS, needed to process the science data through calibration (or recalibration by the user in the IRAF/STSDAS environment).

The STIS pipeline will join into a single dataset the following sets of separately commanded exposures

- crsplit science exposures with one another, to allow the pipeline to produce a cosmic ray rejected image for CCD imaging and spectroscopic observations. The separate exposures in the crsplit set are written as subsequent image extensions (Section 5, page 6) in the science file (rootname_sci.fit) by generic conversion.

- repeatobs science exposures with one another, to keep in one physical file, all the sequential time resolved image/spectra. In addition to calibrating each repeatobs science exposure individually, the STIS pipeline will produce a single integrated image/spectrum from the time series. The separate exposures in the repeatobs set are written as subsequence image extensions in the science file (rootname_sci.fit) by generic conversion.

- wavecal exposures (taken to determine the zero point of the wavelength scale and the aperture location on the detector) with each spectroscopic set of science exposures. The wavecal exposures are written to their own file in the dataset by generic conversion (rootname_wav.fit)

All of these associated sets are generated from single Proposal log sheet lines: the wavecals are automatically produced by TRANS[3], the crsplits are generated from a single line

---

2. TRANS populates a database relation (qbundle) in the PMDB which generic conversion uses to reconstruct the associated set when the data comes down from the telescope (a memo describing the associations and the relations which support them in under construction by Ed Hopkins and the SM97 Keyword Working Group).

3. The automatic wavecals can only be turned off using Engineering only parameters in which case TRANS will associate the user requested wavecal with the science exposure.

specified as CRPSLIT, and the repeatobs observations are generated from the NEXP=n command on the logsheet line.

5. STIS Output Products

All input and output products to and from `calstis` will be FITS files. The ACCUM mode science data will be stored as FITS image extensions, one dimensional extracted calibrated spectra will be stored as FITS binary tables. The time tag data stream will be written as a FITS binary table (readable back into IRAF as a qpoe file using the `fits2qp` task - it is anticipated that the existing x-ray software for processing of event data will prove useful in the analysis of time tag data). See Section 6, page 9 below for a more detailed description of the FITS Image Extensions and how they will be used.

The full list of uncalibrated output files which generic conversion can produce for STIS observations is given in Table 1. Any given observation will typically have a subset of these files (see Table 4). The final output product a user will typically want for different astronomical targets/science aims for spectroscopic observations is given inTable 5

Table 1. Uncalibrated file extensions (suffixes) in STIS datasets

File Name	Explanation	File Type	Type of Observation Where will be Present
_sci.fit	uncalibrated science data	FITS image extension file	all science and target acquisition observations (ACCUM, TIMETAG, ACQ, ACQ/PEAK)
_wav.fit	uncalibrated wavecal data	FITS image extension file	all spectroscopic observations (ACCUM or TIMETAG)
_spt.fit	science support file planned observational parameters from the PMDB and engineering keywords extracted from the UDL.	FITS image extension file	all observations (ACCUM or TIMETAG)
_wsp.fit	wavecal support file planned observation parameters from the PMDB and engineering keywords extracted from the UDL	FITS image extension file	
_asn.fit	association table (defines associated exposures)[a]	FITS binary table	all spectroscopic observations and any associated imaging observations
_tag.fit	uncalibrated time tag data stream	FITS binary table	all TIMETAG

a. The association table identifies all of the exposures which are members of the associated dataset by their original rootname and lists the role that the exposure plays in the association (e.g., wavecal, crsplit).

The full set of intermediate output files from `calstis` is given in Table 2

Table 2. Intermediate file extensions in STIS datasets

File Name	Explanation	File Type	Type of Observation Where will be Present
_flt.fit	flat fielded science data	FITS image extension file	all science and acquisition observations
_crj.fit	cosmic ray rejected science image	FITS image extension file	all crsplit spectroscopic observations
_map.fit	2-D wavelength map extension 1 contains the wavelength map[a] extension 2 contains the spatial map[b]	FITS image extension file	all spectroscopic observations
_pht.fit	photometric throughput table gives conversion to flux units as a function of wavelength	Fits binary table	all observations
_cwv.fit	flat fielded and source subtracted wavecal	FITS image extension file	all spectroscopic observations

a. assigns wavelength to each pixel in the _flt.fit (and _crj.fit) file.
b. assigns distance along the slit to each pixel in the _flt.fit (and _crj.fit) file.

The full list of possible final output products from calstis are given in Table 3

Table 3. Calibrated file extensions in STIS datasets

File Name	Explanation	File Type	Type of Observation Where will be Present
_cal.fit	fully calibrated two-dimensional data: geometrically corrected imaging data rectified spectroscopic image (with linear wavelength and space scales)	FITS image extension file	all science observations, except echelle
_x1d.fit	background subtracted 1-D spectra	FITS binary table	all spectroscopic observations of point sources or through short slits
_sum.fit	summed calibrated image for time resolved repeatobs	FITS image extension file	all repeatobs observations
_s1x.fit	summed extracted spectra for repeatobs	FITS binary table	all spectroscopic repeatobs observations of point source or through short slits
_ctg.fit	calibrated time tagged stream	FITS binary table	time tag data only
_trl.fit	trailer file (log on pipeline processing)	FITS binary table	all

Summarized by observation mode the files a user will see for a given observation are shown in Table 4.

Table 4. Typical STIS Output Products by Observation Type

Observation Type	Uncalibrated Files	Intermediate Files	Calibrated Files
ACCUM, IMAGING	_sci, _spt, _asn	_flt, _pht	_cal, _trl
ACCUM, IMAGING, CRSPLIT	_sci, _spt, _asn	_flt, _crj, _pht	_cal, _trl
ACCUM, IMAGING, REPEATOBS	_sci, _spt, _asn	_flt, _pht	_cal, _sum, _trl
ACCUM, SPECTROSCOPIC	_sci, _wav, _asn _spt, _wsp	_flt, _map, _pht, _cwv	_cal, _x1d, _trl
ACCUM, SPECTROSCOPIC, CRSPLIT	_sci, _wav, _asn _spt, _wsp	_flt, _crj, _map, _pht, _cwv	_cal, _x1d, _trl
ACCUM, SPECTROSCOPIC, REPEATOBS	_sci, _wav, _asn _spt, _wsp	_flt, _map, _pht, _cwv	_cal, _x1d, _trl, _sum, _s1d
TIMETAG IMAGING and SPECTROSCOPIC	ACCUM extensions plus _tag	ACCUM extensions	ACCUM extensions plus _ctg

Table 5. Science Output Products for Different Targets and Science Goals for Spectroscopic Observations

Target Type	Science Aim	File	Rationale
point source (star/quasar)	fully trust pipeline reduction	_x1d	1-d extracted spectrum (flux, error, data quality, wavelength, background, per pixel)
point source (star/quasar)	re-perform extraction	_flt or _crj, _map and _pht	uninterpolated flat fielded image and wavelength and space 'distortion' maps
extended object (nebulae/galaxy)	proceed directly to analysis	_cal	fully rectified 2-D spectrum simplest to work with directly
extended object (nebula/galaxy)	work from uninterpolated data	_flt or _crj, _map, _pht	more difficult to work with using existing software but retains full uninterpolated information

The full size of a typical CAL class dataset (inclusive of uncalibrated, intermediate, and calibrated outputs) is roughly 50 Megabytes for spectroscopic 1024 x 1024 format observations, and 25 Megabytes for imaging 1024 x 1024 format observations.

6. Science Data and FITS Image Extensions

All STIS science data, excluding the time tag data stream and 1-D extracted spectra, will be stored as a series of science, data quality, and error array triplets in a FITS image extension file. As described in Section 4, a single STIS dataset can contain multiple science exposures; these will be stored in the extensions of a single FITS file (rootname_sci.fit, see Table 4). Each image extension header will have a keyword (EXPNAME) identifying the original commanded IPPSSOOOT of that exposure. The STIS wavecal (_wav.fit) will have a similar structure.

The format adopted for FITS image extension files is as follows:

- primary header unit - *extension 0*

- primary data unit - **empty** - *extension 0*

- first science image extension (EXTVER = 1, EXTNAME = SCI) - *extension 1*

- first error image extension (EXTVER=1, EXTNAME = ERR) [header plus data] - *extension 2*

- first data quality image extension (EXTVER=1, EXTNAME= DQ) (header plus data) - *extension 3*

- second science image extension (EXTVER=2, EXTNAME=SCI) - *extension 4*

- second error image extension (EXTVER=2, EXTNAME = ERR) - *extension 5*

- second data quality extension (EXTVER=2, EXTNAME=DQ) - *extension 6*

- etc...

A few important points to note are:

- The primary header data unit contains keywords which apply to all the image extensions (i.e., which are constant across the image extensions).

- The primary data unit is always left blank; this indicates that the science images in the image extensions are of equal status - one is not primary over the others.

The Error and Data Quality Arrays

Each science image will have an accompanying error array. Noise model keywords in the header of the error image extension (NOISEMOD and NOISCOF1 through NOISCOF5)

define the noise model. The values of NOISEMOD are tbd. The uncalibrated error array will have the noise model keywords set in the header and contain a zero size array (NAXIS=0). The pipeline I/O interface would then interpret the header keywords to construct an image in memory with the appropriate values. This considerably reduces the sizes of the science files. A similar scheme may also be used for the calibration reference files.

The error in the uncalibrated data is given by

$$\sigma = \sqrt{\sigma_c^2 + \langle R|g\rangle^2}$$

where

- R is the observed DN minus the electronic bias of the pixel

- g is the CCD gain factor (1 for MAMA)

- σ_c is the white noise component (e.g., readout noise).

Calstis will propagate statistical errors throughout the reduction process.

In addition to propagating statistical errors, calstis will also track conditions such as saturated pixels, photocathode blemishes and hot pixels, which are not well modelled by statistical errors. A data quality flag will be assigned to each pixel in the data quality extension. The flagged conditions are set as specific bits in the 16-bit integer words in the data quality file: thus more than one data quality condition can be flagged simultaneously, via the bitwise logical OR operation. Table 6 gives the specific conditions which are flagged. Generic conversion will fill the data quality file only when there is missing (data-lost) or dubious (softerr) data. If no such errors were taken, generic conversion will produce an empty data quality file whose header has NAXIS=0. When reading the data quality file the i/o interface will expand it to an array of appropriate size and value (See Internal Memo by A. Farris)

Table 6. STIS Data Quality Flags

FLAG	condition	QUALITY INDICATED
1	softerr	reed solomon decoding error
2	datalost	data replaced by fill values
4	detectorprob	bad detector pixel (e.g., bad column or row, mixed science and bias for overscan)
8	datamasked	data masked by occulting bar or beyond aperture
16	hotpix	pixel having dark rate > tbd times mean
32	largeblem	large blemish, depth > TBD percent of flat field
64	tbd	
128	tbd	
256	satpixel	saturated pixel, count rate at 90% of max possible - local non-linearity turns over and is multivalued; pixels within 10% of turnover and all pixels within 4 pixels of that pixel are flagged
512	calibdefect	bad pixel in reference file
1024	smallblem	small blemish, depth <> TBD percent of flat field
2048	tbd	
4096	tbd	
8192	datareject	data rejected in input pixel during image combination
16384	tbd	
32768	tbd	
65536	tbd	

Using Image Extensions in IRAF/STSDAS - An Example

A simple and common example is a CCD spectroscopic observation, in which the science exposure has been crsplit into 3 separate exposures. A single FITS science file (_sci.fit) will store the data from all three of the associated science exposures. This _sci.fit file will have:

- a primary header which applies to the set of science exposures and describes the full observation set (e.g., gives the total exposure time from the three crsplit exposures), accompanied by an empty primary data unit

- the first image extension header (containing the exposure specific keyword information like the exposure duration and start time for that particular exposure) plus the associated science data for the first crsplit exposure

- the error file (header plus data) for the first crsplit exposure.

- the data quality (header plus data) for the first crsplit exposure

These 3 image extensions will then be followed by similar triplets of science, error, and data quality extensions for each of the other two crsplit science exposures, resulting in a single FITS file containing 3 science exposures within 9 image extensions.

User Access Within IRAF/STSDAS

Within IRAF/STSDAS, using the FITS image extension kernel, users will be able to examine any given image extension using standard IRAF/STSDAS task and a well defined syntax (see STSDAS document describing FITS image extensions, Greenfield et al, 1995). For example, to display the science image extension for the second crsplit image in the example described above, the user would type

>disp o1234567t_sci.fit[sci,2] or

>disp o1234567t_sci.fit[4]

7. Calibration Switches and Reference Files

The calibration switches control the processing of the uncalibrated data through `cal-stis`. These switches appear in the primary header unit of the science file, the wavecal file, and the target acquisition file. The switches which appear in the science header are different for the 4 basic classes of STIS data (CCD Imaging, CCD Spectroscopic, MAMA Imaging and MAMA Spectroscopic) so that only the switches relevant to a specific observation will appear in its header. The portion of the science header containing the calibration switches and associated reference files for each of these classes is shown in Figure 1 through Figure 4. Note that the steps and switches for the spectroscopic reduction are subject to change; the final detailed design for the spectroscopic calibration portion of `calstis` is not yet defined.

Figure 1: CCD Imaging Header showing calibration switches and files

```
        / CALIBRATION SWITCHES: PERFORM, OMIT, COMPLETE

   DQICORR  = '     '      /data quality initialization
   ATODCORR= '     '       /correct for A to D conversion errors
   BLEVCORR = '     '       /subtract bias level computed from overscan image
   BIASCORR = '     '      /subtract bias image
   DARKCORR= '     '       /subtract dark image
   FLATCORR = '     '      /flat field image
   SHADCORR= '     '       /apply shutter shading correction
   GEOCORR  = '     '       /perform geometric correction for imaging modes
   PHOTCORR= '     '       /convert to absolute flux units
   CRCORR   = '     '      /combine exposures to reject cosmic rays
   RPTCORR  = '     '      /coadd individual repeat exposures


         /CALIBRATION REFERENCE FILES

   BPIXTAB   = '     '       /bad pixel table
   ATODFILE = '     '       /analog to digital correction file
   BIASFILE  = '     '       /bias image file name
   DARKFILE = '     '        /dark image file name
   PFLTFILE  = '     '       /pixel to pixel flat field file name
   DFLTFILE  = '     '        /delta flat field file name
   LFLTFILE  = '     '       /low order flat field file name
   SHADFILE  = '     '        /shutter shading correction file
   DISTFILE   = '     '        /spatial distortion image
   CCDGNTAB= '     '        /CCD relative gain parameters
   PHOTTAB  = '     '        /photometric throughput table
```

Figure 2: CCD Spectrographic Header showing calibration switches and files

```
          / CALIBRATION SWITCHES: PERFORM, OMIT, COMPLETE

DQICORR   = '     '        /data quality initialization
ATODCORR = '               /correct for A to D conversion errors
BLEVCORR = '     '          /subtract bias level computed from overscan image
BIASCORR  = '     '        /subtract bias image
DARKCORR = '     '         /subtract dark image
FLATCORR  = '     '        /flat field image
SHADCORR = '     '          /apply shutter shading correction
CRCORR    =                /combine exposures to reject cosmic rays
WAVECORR= '     '          /use wavecal to adjust wavelength zeropoint
DFNDCORR = '     '         /find spectrum zeropoint in dispersion
XFNDCORR = '     '         /find spectrum zeropoint in cross dispersion
FINDCORR  = '     '        /find spectral order positions
DISPCORR  = '     '        /apply 2-dimensional dispersion solutions
SGEOCORR= '     '           /correct for small scale geometric distortion
HELCORR   = '     '         /convert to heliocentric wavelengths
FLUXCORR=                  /convert to absolute flux units
X2DCORR   = '     '        /extract 2-d spectral image and wavelength map
RECTCORR= '     '          /rectify 2-D spectral image
X1DCORR   = '     '        /attempt 1-D spectral extraction
EMETHOD  = '     '         /1-D extraction method
BACKCORR= '     '          /subtract background, sky, or interorder
RPTCORR  = '     '         /add individual repeat exposures


          /CALIBRATION REFERENCE FILES
BPIXTAB   = '     '        /bad pixel table
ATODFILE  = '     '         /analog to digital correction file
BIASFILE   = '     '         /bias image file name (CCD only)
DARKFILE   = '     '         /dark image file name
PFLTFILE   = '     '         /pixel to pixel flat field file name
DFLTFILE   = '     '         /delta flat field file name
LFLTFILE   = '     '         /low order flat field file name
SHADFILE   = '     '          /shutter shading correction file
WAVECAL1= '     '          /wavecal image file name taken prior to obs.
WAVECAL2= '     '          /wavecal image file name taken after obs.
APERTAB  = '     '          /aperture description table
APLOCTAB = '     '          /aperture location table
SLITTAB    = '     '         /slit description table
PLATETAB = '     '          /plate scale table
SLOCTAB  = '     '          /spectral order location table
DISPTAB   = '     '          /dispersion coefficient table
IANGTAB   = '     '          /incidence angle correction table
OFFSTAB   = '     '          /internal/external wavelength offset table
SDSTFILE  = '     '          /small scale distortion image
APTHRTAB= '     '           /relative aperture throughput table
PHOTTAB  = '     '          /absolute sensitivity table
BACKTAB  = '     '          /background subtraction parameter table
INTORTAB = '     '          /echelle interorder scattered light corrections
```

Figure 3: MAMA Imaging Header showing calibration switches and files

```
              / CALIBRATION SWITCHES: PERFORM, OMIT, COMPLETE

LORESCNV= '       '        /convert MAMA data to Lo-Res before processing
TINTCORR = '      '        /integrate time tag data in time
TTAGCORR= '       '        /output calibrated time tag stream, TIMETAG only
DQICORR  = '      '        /data quality initialization
GLINCORR =                 /correct for global detector non-linearities
LINFLAG   =                /flag local and global detector non-linearities
DARKCORR= '       '        /Subtract dark image
FLATCORR = '      '        /flat field data
GEOCORR  = '      '        /perform geometric correction for imaging modes
PHOTCORR= '       '        /convert to absolute flux units
RPTCORR  = '      '        /add individual repeat exposures

                     /CALIBRATION REFERENCE FILES

BPIXTAB  = '      '         /bad pixel table
MLINTAB  = '      '         /MAMA linearity correction table
DARKFILE = '      '          /dark image file name
PFLTFILE  = '      '         /pixel to pixel flat field file name
DFLTFILE  = '      '         /delta flat field file name
LFLTFILE  = '      '         /low order flat
DISTFILE  = '      '         /spatial distortion image
PHOTTAB= '       '           /photometric throughput table
```

Figure 4: MAMA Spectroscopic Header showing calibration switches and files

```
            / CALIBRATION SWITCHES: PERFORM, OMIT, COMPLETE

TINTCORR = '        '        /integrate time tag stream to form accum image.
TTAGCORR= '        '        /output calibrated time tag stream
LORESCNV = '        '        /convert MAMA data to Lo-Res before processing
DQICORR   = '      '        /data quality initialization
GLINCORR  = '      '        /correct for global detector non-linearities
LINFLAG    = '     '        /flag local and global detector non-linearities
DARKCORR = '       '         /subtract dark image
FLATCORR = '       '        /flat field data
WAVECORR= '        '        /use wavecal to adjust wavelength zeropoint
DFNDCORR = '       '         /find spectrum zeropoint in dispersion
XFNDCORR = '       '         /find spectrum zeropoint in cross dispersion
FINDCORR  = '      '        /find spectral order positions
DISPCORR  = '      '        /apply 2-dimensional dispersion solutions
SGEOCORR = '       '        /correct for small scale geometric distortion
HELCORR    = '     '        /convert to heliocentric wavelengths
ERIPCORR  = '      '        /apply echelle ripple correction
FLUXCORR =                  /convert to absolute flux units
X2DCORR  = '        '        /extract 2-D spectral image and wavelength map
RECTCORR = '       '         /rectify 2-D spectral image
X1DCORR  = '       '        /attempt 1-D spectral extraction
EMETHOD  = '       '        /1-D extraction method
BACKCORR= '        '        /subtract background, sky, or interorder
RPTCORR  =                  /add individual repeat exposures


              /CALIBRATION REFERENCE FILE

BPIXTAB   = '      '        /bad pixel table
MLINTAB  = '      '        /MAMA linearity correction table
DARKFILE  = '      '         /dark image file name
PFLTFILE   = '     '         /pixel to pixel flat field file name
DFLTFILE   = '     '         /delta flat field file name
LFLTFILE   = '      '         /low order flat field
WAVECAL1= '       '        /wavecal image file name taken prior to obs.
WAVECAL2= '       '        /wavecal image file name taken after obs.
APERTAB   = '      '         /aperture description table
APLOCTAB = '       '         /aperture location table
SLITTAB    = '      '        /slit description table
PLATETAB = '       '        /plate scale table
SLOCTAB  = '       '        /spectral order location table
DISPTAB   = '      '        /dispersion coefficient table
IANGTAB   = '      '        /incidence angle correction table
SDSTFILE  = '      '         /small scale distortion image
OFFSTAB  = '       '        /internal/external wavelength offset table
ERIPTAB   = '      '        /echelle ripple coefficients
APTHRTAB= '        '        /relative aperture throughput table
PHOTTAB  = '       '        /absolute sensitivity table
BACKTAB  = '       '        /background subtraction parameter table
INTORTAB = '       '        /echelle interorder scattered light corrections
```

8. High Level Design of `Calstis`

`Calstis` will be broken into a series of "stages". Each stage will be a separate IRAF task (and could be a separate executable), and each will perform one or more calibration steps. `Calstis` is broken into a new stage whenever it writes an output file, either an intermediate file or the final calibrated file. This permits each stage to have just two or three parameters, which would be the names of one or two input files and the name of the output file(s). When running in the pipeline, the names of the input and output files for each stage will be constructed from the rootname with fixed suffixes. OPUS determines which stage to run using the OPUS process resource file, which specifies for each calstis return code, the next processing step to run.

When running outside the pipeline, however, the input file names will be flexible in order to permit executing a given stage more than once, passing the previous output back in as input, with different switches set to perform. This design assures that it is easy for users to (re-)preform isolated stages of the pipeline on their STIS data. In the IRAF environment users will be able to run the entire pipeline sequence on their data, or they will be able to run isolated stages of the process (e.g., if they wish to redetermine the wavelength scale zeropoint or to re-extract the 1-D spectrum, changing the input parameters), entering their own input and output file names.
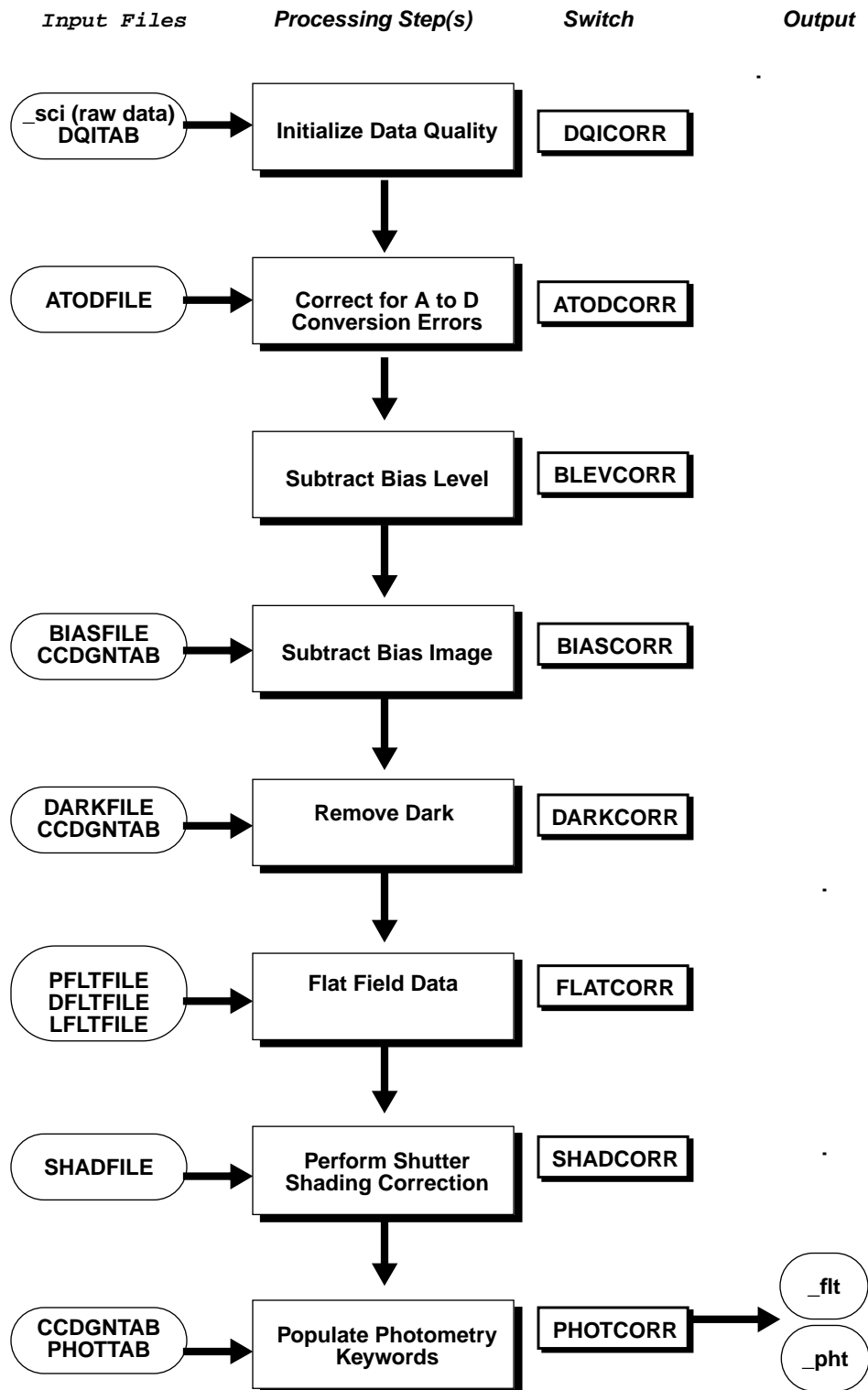
The 10 separate stages in the `calstis` pipeline are:

- `calstis-1` which performs two-dimensional image reduction, producing a flat fielded (photometrically calibrated in the case of imaging data) image (->_flt.fit)

- `calstis-2` which performs cosmic ray rejection for crsplit data (->_crj.fit)

- `calstis-3` which geometrically corrects imaging data (->_cal.fit for imaging data)

- `calstis-4` which processes the flat fielded wavecal observation and determines the center (aperture and central wavelength location) of the spectrum from the processed wavecal observation (->_cwv.fit)

- `calstis-5` which performs basic spectroscopic calibration, determining the mapping between detector pixels in the flat fielded image and wavelengths and distance along slit and determining the wavelength dependent conversion from counts to physical flux units (->_map.fit, _pht.fit)

- `calstis-6` which performs one dimensional spectral extraction, producing a background subtracted one dimensional spectrum (->_x1d.fit)

- `calstis-7` which produces a fully rectified two-dimensional calibrated spectral image, with wavelength running linearly in the y direction, and distance along slit running linearly in the x direction (->_cal.fit)

- `calstis-8` which sums repeatobs calibrated output files to produce a single integrated image and or 1-D spectrum (->_sum.fit, _sx1.fit)

- `calstis-9` which integrates the time-tag data stream in time, correcting for doppler motion as needed, to produce a _sci.fit file for time tag data. The _sci file is then passed through the ACCUM mode pipeline software (`calstis1-8`), as appropriate) (-> _sci.fit for TIMETAG only)

- `calstis-10` (or `calstis-1'`) which processes the time tag data stream for TIMETAG operating mode observations through 2-D image reduction, outputting a flat fielded or photometrically calibrated (in the case of image data) time tag data stream (->_ctg.fit).
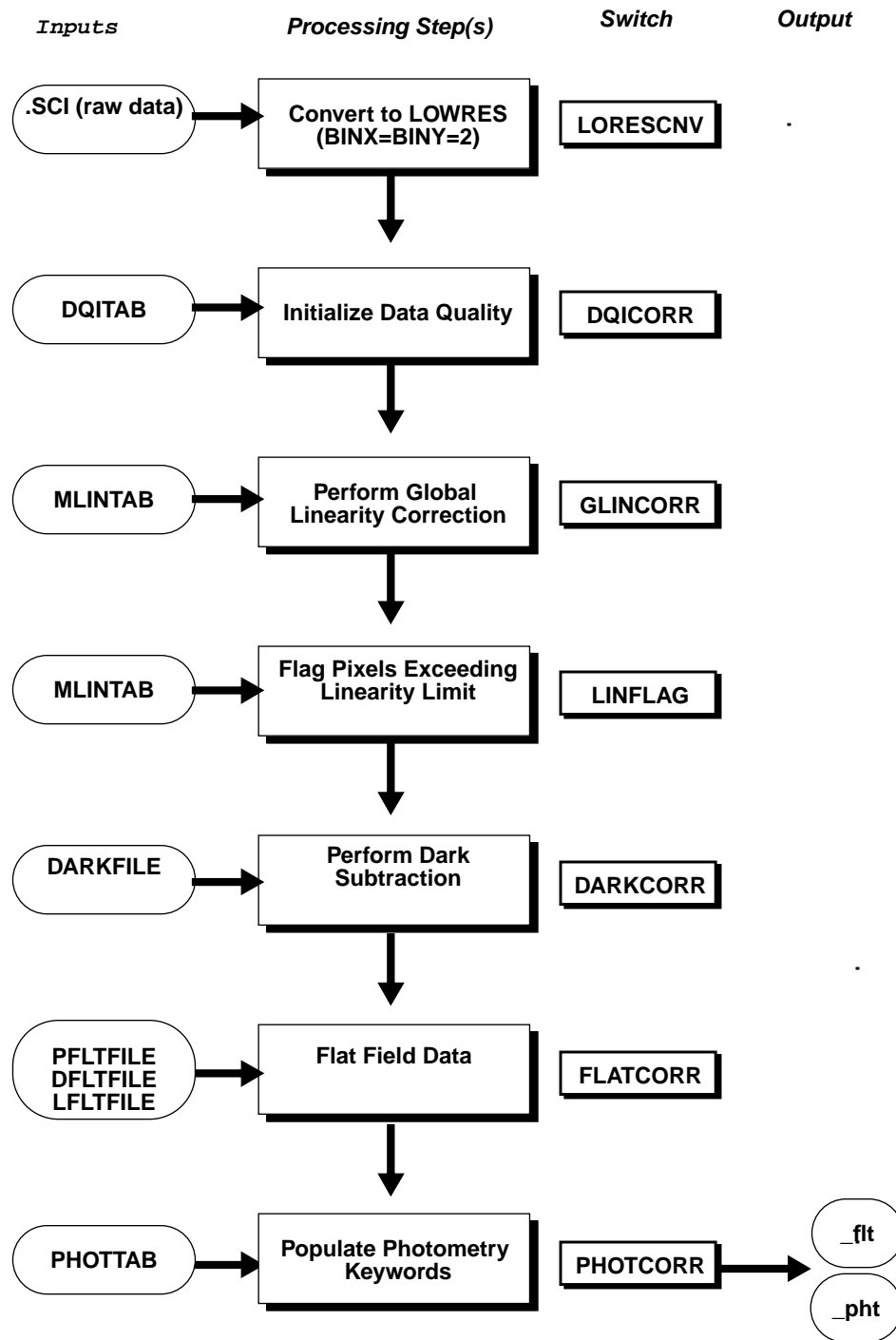
The sequence of operations performed in each of 'stages' 1-8 are illustrated in Figure 5 through Figure 14. As noted earlier, the design for the spectroscopic calibration portion of the `calstis` pipeline is still not finalized. so the diagrams for the spectroscopic portion describe only our present understanding, which is subject to change. `Calstis-10` has the same sequence of steps as does `calstis-1.`

In addition, we will develop a stand-alone utility which runs also a separate task in iraf/stsdas which extracts two values (the x and y offsets of the source in the slit) from the target acquisition. Generic conversion will use this task to derive the values of keywords it populates in the uncalibrated science file (_sci.fit) for long slit spectroscopic observations. This step will not run as part of `calstis`, but will be available within STSDAS as a stand-alone tool.
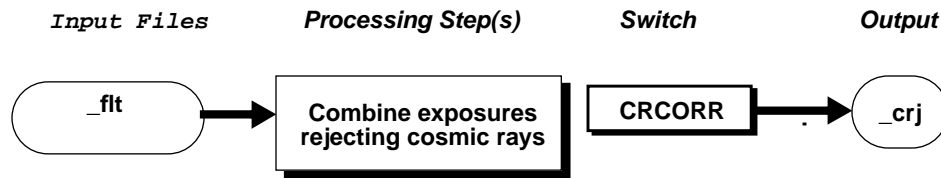
Figure 5: . Flow through `calstis-1`, two-D image reduction, for *CCD* data
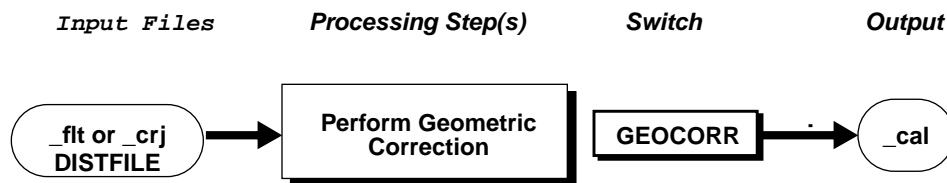
Figure 6: Flow Through `calstis-1`, two-D image reduction for *MAMA* data

Inputs	*Processing Step(s)*	*Switch*	*Output*

.SCI (raw data) → **Convert to LOWRES (BINX=BINY=2)** — LORESCNV

DQITAB → **Initialize Data Quality** — DQICORR

MLINTAB → **Perform Global Linearity Correction** — GLINCORR

MLINTAB → **Flag Pixels Exceeding Linearity Limit** — LINFLAG

DARKFILE → **Perform Dark Subtraction** — DARKCORR

PFLTFILE DFLTFILE LFLTFILE → **Flat Field Data** — FLATCORR

PHOTTAB → **Populate Photometry Keywords** — PHOTCORR → _flt _pht

Figure 7: Flow through `calstis-2`, cr-rejection for crsplit exposures

Input Files	*Processing Step(s)*	*Switch*	*Output*

_flt → Combine exposures rejecting cosmic rays → CRCORR → _crj

Figure 8: Flow through `calstis-3`, geometric calibration for imaging data

Input Files	*Processing Step(s)*	*Switch*	*Output*

_flt or _crj DISTFILE → Perform Geometric Correction → GEOCORR → _cal

Figure 9: Flow through `calstis-4`, wavecal processing

Input Files	Processing Step(s)	Switch	Output

_fwv and _flt or _crj → **Subtract Source from wavecal** — sclamp=HITM

PLATETAB SLOCTAB, SLITTAB → **Locate Spectrum, Update zeropoint offset in science header** — DFNDCORR

APLOCTAB → **Locate Aperture, Update zeropoint offset in science header** — XFNDCORR → _flt *or* _crj

(headers updated)
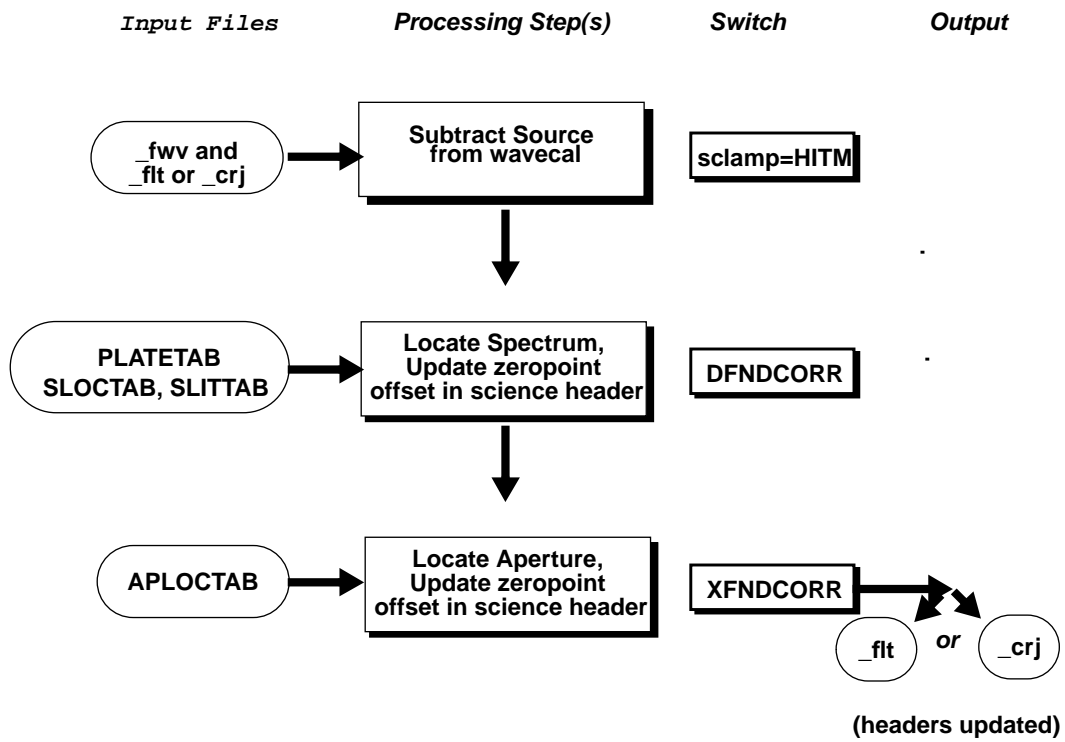
Figure 10: Flow through `calstis-5`, Basic Spectroscopic Reduction

Input Files	Processing Step(s)	Switch	Output

_crj or _flt
LFLTFILE

ERIPTAB → **Apply Echelle Ripple Correction** — **RIPCORR** → **_flt**

DISPTAB IANGTAB → **Apply 2-D Dispersion Solutions** — **DISPCORR**

SDSTFILE → **Correct for small scale geometric distortion** — **SDSTCORR**

Extract 2-D Flux Image and Wave Map — **X2DCORR**

OFFSTAB → **Convert to Heliocentric wavelengths** — **HELCORR**

PHOTTAB APTHRTAB → **Convert to Absolute Units** — **FLUXCORR** → **_pht** / **_map**

Figure 11: Flow Through `calstis-6`, 1-D extraction.

Input Files	*Processing Step(s)*	*Switch*	*Output*
_flt or _crj _map, _pht	Find Spectrum to Extract	FINDCORR	. .
BACKTAB	Determine Background Spectrum	BACKCORR	
	Extract 1-Dimensional Spectrum	X1DCORR	_x1d
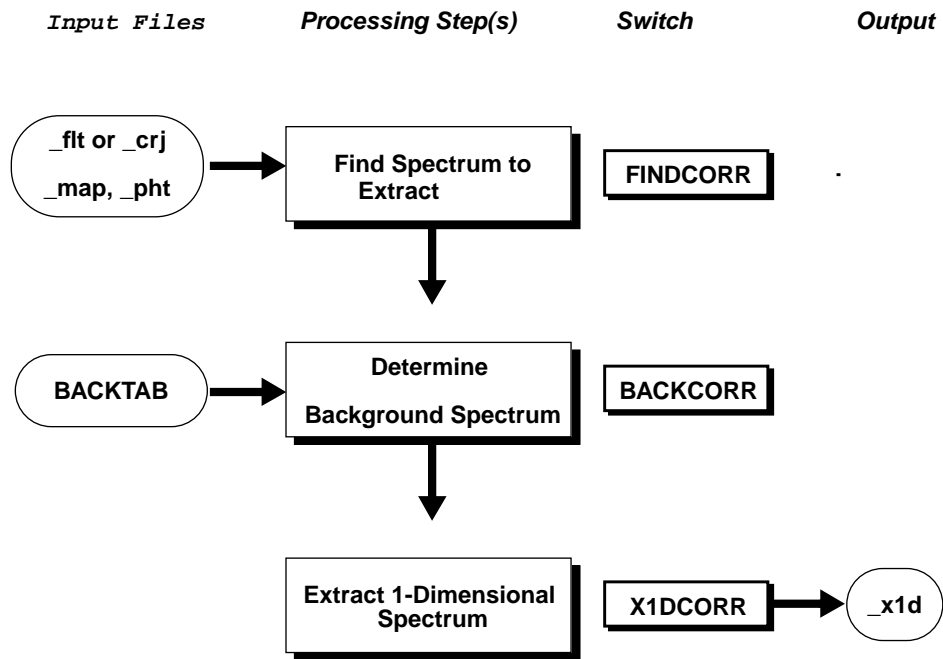
Figure 12: Flow Through `calstis-7,` 2-D rectification

Input Files　　　*Processing Step(s)*　　　*Switch*　　　*Output*

```
_flt or _crj
_map, _pht   →   Rectify 2-D Spectrum
                 Linearize in Wavelength   X2DCORR  →   _cal    .
```
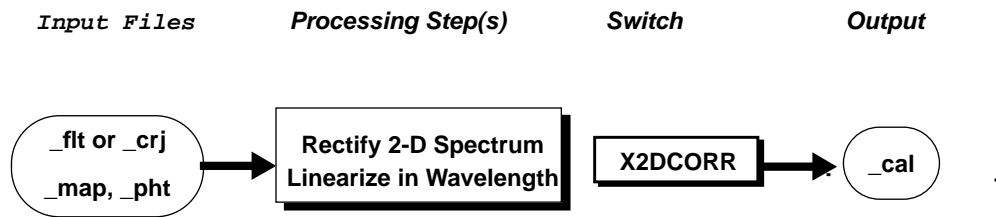
Figure 13: Flow through `calstis-8,` summing of repeatobs exposures.

Input Files　　　*Processing Step(s)*　　　*Switch*　　　*Output*

```
                                                        _sum
_cal, _x1d   →   Sum Repeat         RPTCORR  →
                 Exposures                              _sx1
```
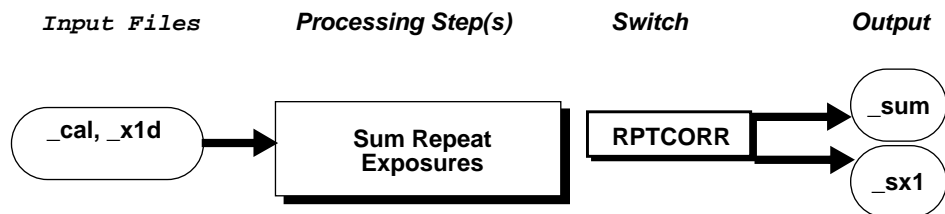
Figure 14: Flow Through `calstis-9,` integrate time tag data stream

Input Files　　　*Processing Step(s)*　　　*Switch*　　　*Output*

```
_tag   →   Integrate Time Tag
           Data in Time        TINTCORR  →   _sci    .
```
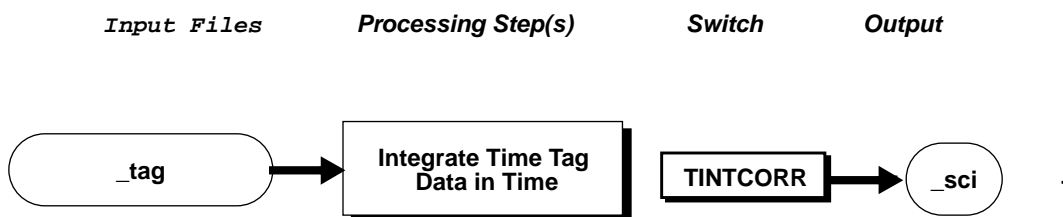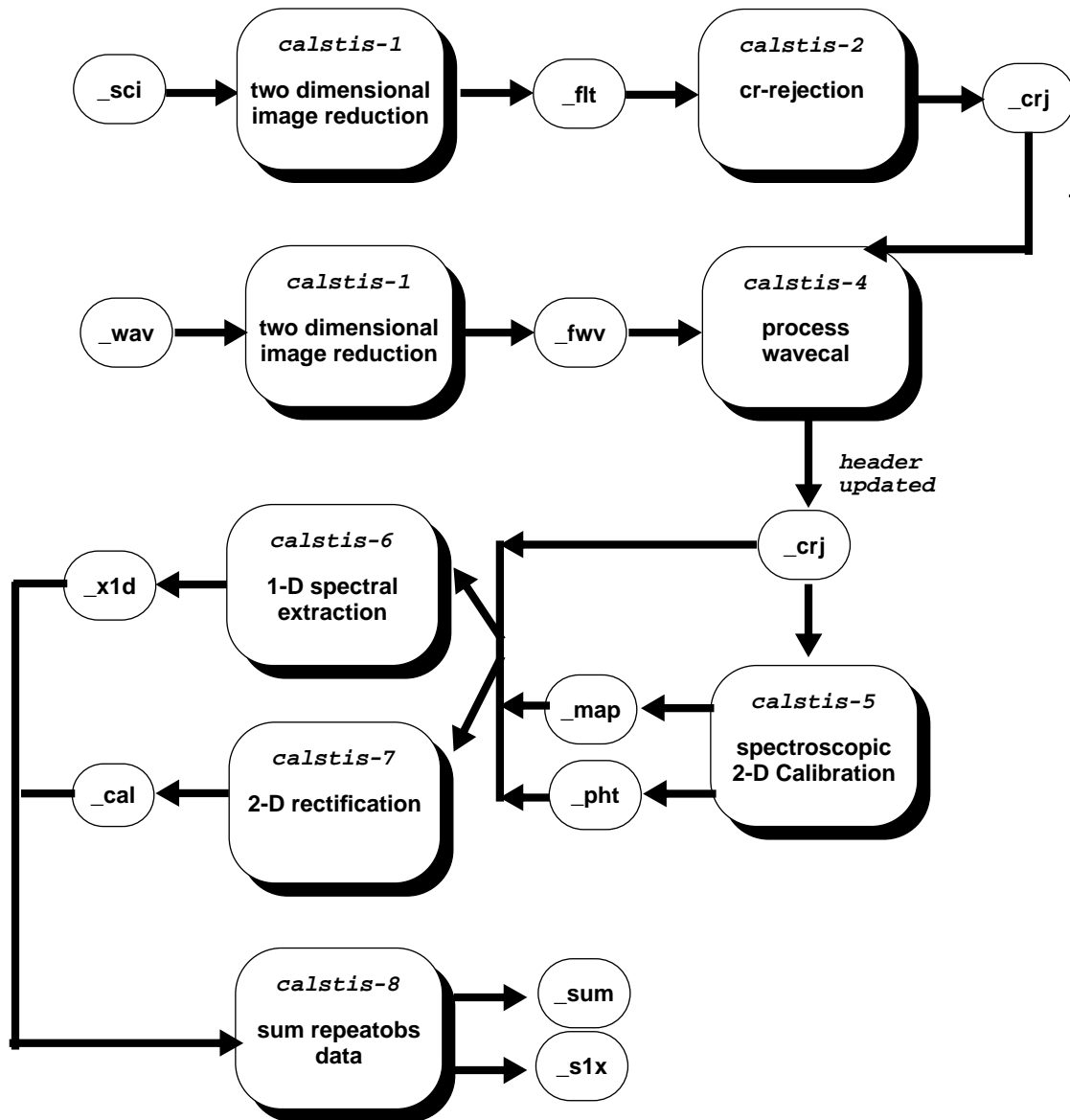
9. Data Flow Through `Calstis`

In the previous section, we described the 10 steps which comprise `calstis`. A given science observation will not typically pass through all 10 steps. The steps which a given science observation passes through will be controlled by the calibration switches in the header of the science or wavecal file (see Section 7, page 12). Figure 15 through Figure 17 show the path through `calstis` taken by data for each of 4 basic observation classes; CCD imaging, MAMA imaging, CCD spectroscopy, and MAMA spectroscopy. As noted earlier, the design for the spectroscopic calibration portion of the `calstis` pipeline is still not finalized and is subject to change.

For MAMA data which are taken in time tag mode, the time tag stream will be passed through `calstis-9` to integrate in time to produce a _sci file, and then the _sci file will proceed through reduction as if it was an ACCUM mode image. However, in addition, the _tag.fit file which contains the time integrated science image output from generic conversion will be processed through two-dimensional image reduction as a time tagged data stream through `calstis-10` and be output as a calibrated time tagged stream (ctg.fit). Development of `calstis-10` is of lowest priority (it is anticipated that `calstis-10` may not be ready for the start of SMOV).
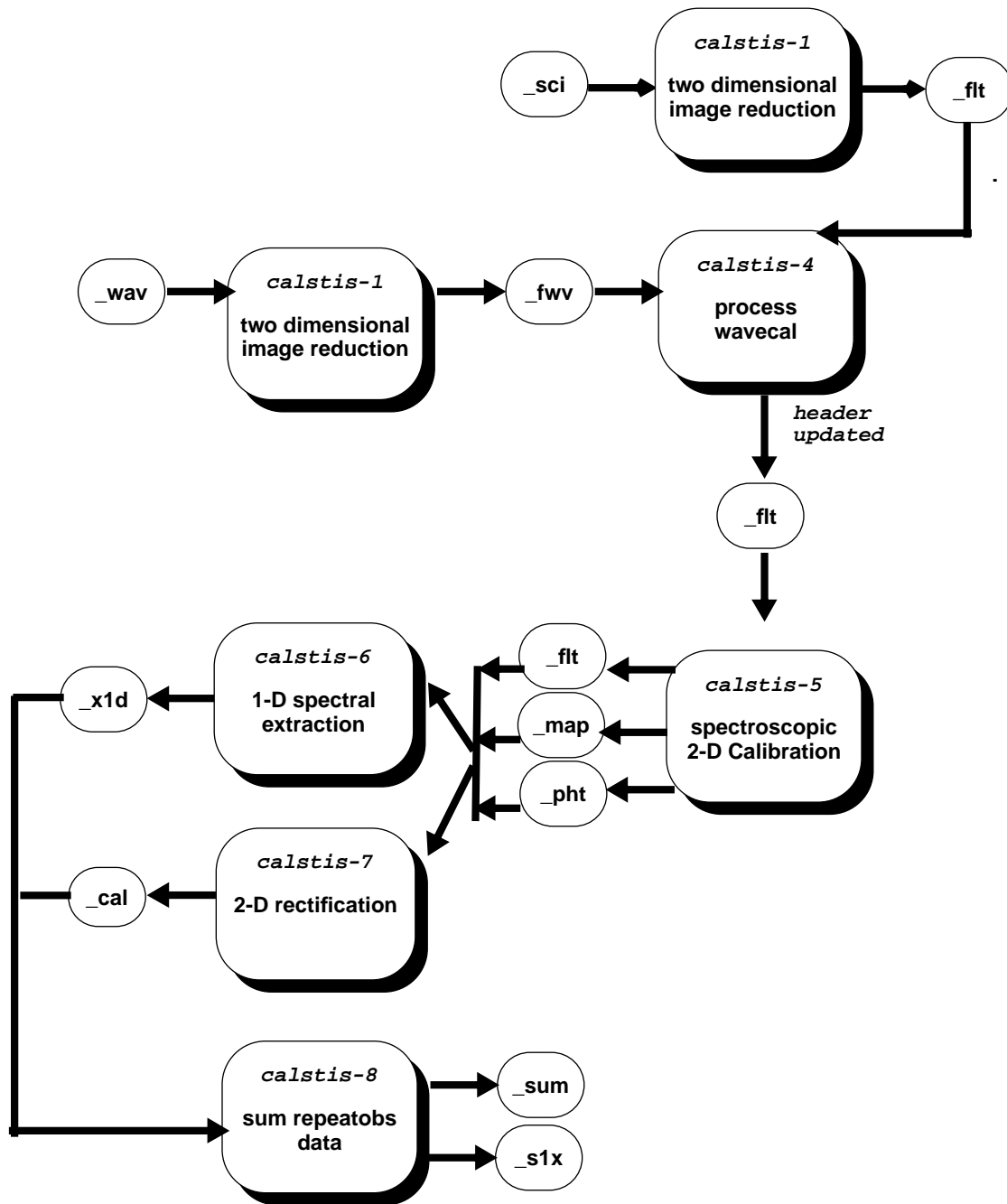
Figure 15: Flow through `calstis` for CCD and MAMA Imaging Data
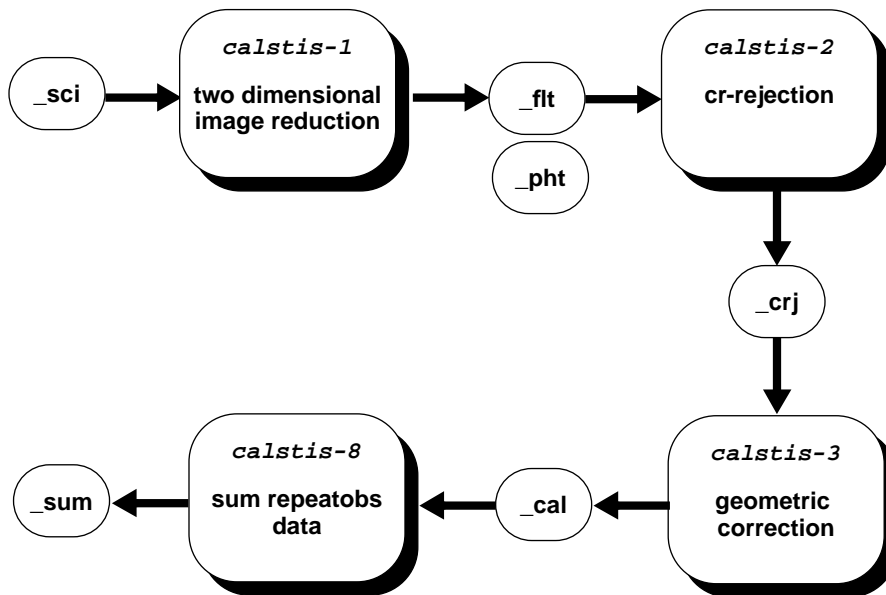
Figure 16: Flow through `calstis` for CCD Spectroscopic Data.

Figure 17: Flow through `calstis` for MAMA Spectroscopic Data.

_sci → *calstis-1* **two dimensional image reduction** → _flt

_wav → *calstis-1* **two dimensional image reduction** → _fwv → *calstis-4* **process wavecal**

header updated

_flt

calstis-5 **spectroscopic 2-D Calibration** → _flt, _map, _pht

calstis-6 **1-D spectral extraction** → _x1d

calstis-7 **2-D rectification** → _cal

calstis-8 **sum repeatobs data** → _sum, _s1x