COS Data Processing Improvements Based on HST SMOV Results

Thomas B. Ake[1], P. Hodge, A. Aloisi, R. I. Diaz, P. Ghavamian, C. Keyes, D. Massa, S. Niemi, C. Oliveira, R. A. Osten, C. R. Proffitt[1], B. York

Space Telescope Science Institute

D. Sahnow

Johns Hopkins University

S. Béland, S. Penton, E. Burgh, K. France

University of Colorado

J. McPhate

University of California, Berkeley

Abstract. After the Cosmic Origins Spectrograph (COS) was installed onboard the Hubble Space Telescope (HST) in May 2009, it underwent an extensive calibration and characterization check-out during the Servicing Mission Observatory Verification (SMOV) period. The results from this program were used to update reference files and make changes to CALCOS, the COS data processing software. Improvements to the standard data products are discussed. For the FUV channel, we have begun development of a flat-field correction. As an intermediate step, grid wire shadows are now ignored when combining FP-POS exposures. Pulse-height filtering has been activated to reduce background features. For the NUV channel, vignetting corrections are incorporated in the flat field file. For both channels, improvements have been made in the wavelength scales, flux calibration, and data quality flagging. Additional data are included in the FITS products to allow users to perform customized processing.

1. Introduction

This paper describes several areas in which CALCOS and the COS calibrated data products have been improved since SMOV due to our better understanding of the instrument, changes to the reference files, and updates to the CALCOS software. Some improvements were planned before launch and were merely awaiting the arrival of SMOV data. Others were in response to investigations of the on-orbit behavior of the instrument. We highlight several advancements that involved either changes to the CALCOS software or changes to the way the software operates in standard pipeline processing. The paper is organized as follows. In Section 2, we explain changes to the handling of the grid wire shadows in merged X1DSUM spectra for the FUV detector, and, in Section 3, the filtering of TIME-TAG events by pulse height amplitude. In Section 4 we discuss improvements to the wavelength calibration. Section 5 presents new capabilities for customized processing and off-line tools for working with COS data, and Section 6, other improvements and plans for future work. Topics that

---

[1]Computer Sciences Corporation

are described in more detail elsewhere at this workshop are: FUV flat fields (Ake et al. 2010), spectral resolution (Ghavamian et al. 2010), wavelength calibration (Oliveira et al. 2010), changes in flux calibration (Osten et al. 2010), far ultraviolet performance (Osterman et al. 2010), target acquisition (Penton et al. 2010), detector performance (Sahnow et al. 2010), and dark count rate (Zheng et al. 2010).

A description of the CALCOS processing steps and reference files can be found in Chapter 3 of the COS Data Handbook (*http://www.stsci.edu/hst/cos/documents/datahandbook*). The latest information on CALCOS changes and dates of installations can be found at *http://www.stsci.edu/hst/cos/pipeline/CALCOSReleaseNotes*.

2. FUV QE Grid Wires

The FUV detector employs a wire grid above the microchannel plate (MCP) stacks to improve quantum efficiency. The wires cast shadows on the MCPs that appear as regularly spaced depressions in extracted spectra (X1D files), about 20% deep every 835 (segment A) or 841 (segment B) pixels. The affected locations can be identified in the CALCOS data products by a data quality flag, DQ=4.

Prior to launch, it was expected that the grid wire shadows could be removed with a flat field, but a methodology for flat fielding the FUV detector is still being developed. In the meantime, X1DSUM processing has been modified so that the grid wire regions are not included in the average (Figure 1). This is accomplished by including DQ=4 in the default value for the SDQFLAGS keyword, the parameter that controls which data will be ignored when FP-POS exposures are merged into the X1DSUM spectrum (Section 6 and Table 3). The thermal and geometric distortion corrections must be accurate to correct grid shadows with a flat field, but excluding the flagged regions is effective regardless of the shadow shape. The regions can also be made slightly wider than the actual shadows to accommodate misregistration. If data were included regardless of the DQ flags, the grid wire shadows would be reduced in depth, but would appear in more places. By excluding DQ=4 when forming the average, the signal-to-noise ratio is lower than in neighboring regions, but the grid wire shadows are effectively removed. For observations with two FP-POS positions, the contributions from only one exposure are present in the X1DSUM in the shadows; for four FP-POS steps, typically three exposures contribute to those locations. If no FP-POS stepping was performed, the X1DSUM spectrum would be a "sum" of only one exposure, and the spectrum will have gaps where the shadows lie. The data in the shadow regions are still availiable in the individual X1D spectra.

3. FUV Pulse Height Filtering

The FUV detector in TIME-TAG mode transmits a pulse height value between 0 and 31 for each event. The pulse height amplitudes (PHA) from photons have a different distribution from background events. Noise events typically have very large or very small PHAs. The PHACORR step in CALCOS flags TIME-TAG events that have PHAs outside specified limits. Such events are excluded when constructing spectra, thereby reducing the background. Prior to launch, CALCOS was set to accept all PHA values, so no filtering would occurr.

During SMOV, FUV segment B was found to exhibit artifacts that resemble emission lines. These features have PHA=0, which is outside the normal range for true photon events. To eliminate these features and reduce detector noise, CALCOS now uses PHA values 4–30 for TIME-TAG exposures. See Figure 2.

When an observation is taken in ACCUM mode, the individual events and information about them (such as PHA) are not preserved, so ACCUM observations cannot be filtered by CALCOS. Thus the features in segment B may be present in the extracted spectra. Since
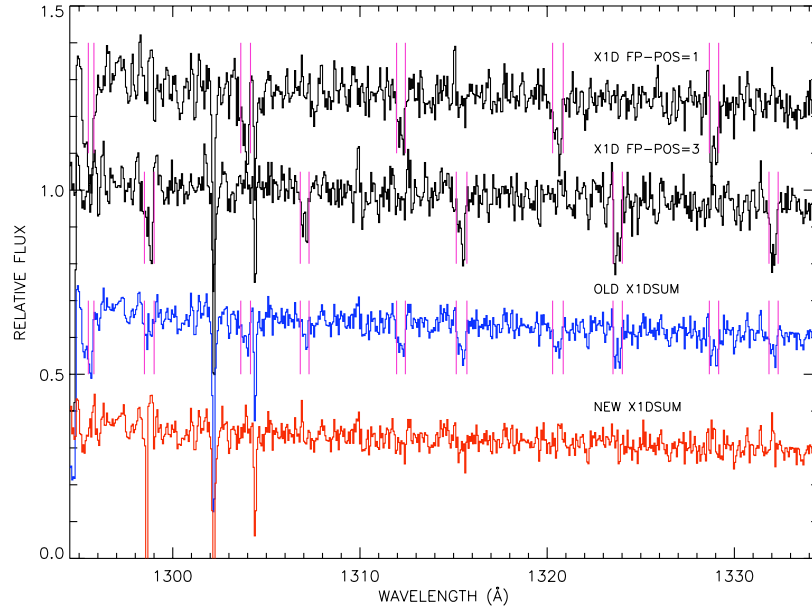
Figure 1: FUV G130M spectra of WD0320–539, offset by an arbitrary amount, taken at two FP-POS settings (X1D spectra) and coadded as an X1DSUM spectrum product. Only two lines, O I λ1302 and Si II λ1304, are real. The top two spectra are the individual exposures. *Violet* bars indicate grid wire regions. The *blue* spectrum is the average of the two (old X1DSUM). The shadows are reduced by a factor of 2, but occur in twice as many places. The *red* spectrum is the current CALCOS processing, where the grid wire regions in the individual spectra are ignored in the coaddition.

ACCUM mode exposures are taken only for high count rate objects, the artifacts will likely be overwhelmed by the target spectrum, but may affect a low flux area such as the bottom of an absorption line. The regions where these features are found are listed in the FUV bad pixel table and are flagged by CALCOS with DQ=4096.

Since PHA filtering can eliminate valid photons with pulse heights at the extremes of their distribution, the filtering process changes the flux calibration by a small amount, as can be seen in Fig. 2. For the time being, CALCOS is using the same sensitivity curves for TIME-TAG and ACCUM data, i.e. with no PHA filtering. CALCOS will be modified in the future to select flux calibrations based on the range of pulse heights that were used in the PHACORR step.

4.   Wavelength Assignment

CALCOS uses PtNe lamp data, which are taken during TIME-TAG exposures or, for AC-CUMs, just before or after the exposure, to convert X pixel values to wavelengths through WAVECORR. The lamp data are extracted as spectra and are cross-correlated with reference file spectral templates to compute offsets to the grating dispersion relations. The offsets are then used as a correction in assigning wavelengths to the pixels.

Prior to obtaining on-orbit data during SMOV, the reference files for the lamp templates and dispersion relations were based on the default FP-POS=3 setting. An exposure for any FP-POS setting was cross-correlated with the one template for that central wavelength. During SMOV, sufficient data were obtained at all four FP-POS settings to create template lamp spectra separately for all FP-POS settings for every central wavelength. This then allows better registration of an exposure's wavecals with the wavelength reference frame.
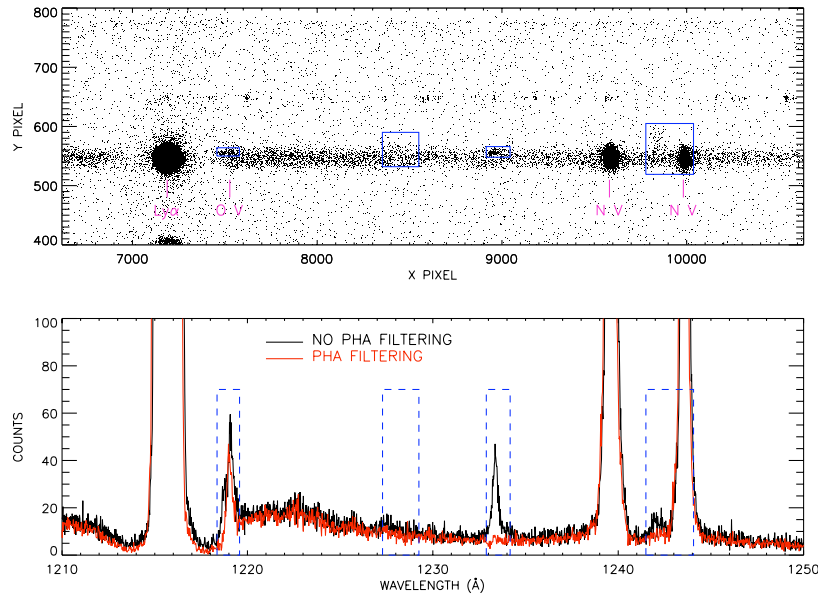
Figure 2: FUV G130M spectrum of Lin 358 showing pseudo-emission features in segment B. Upper panel: COUNTS image showing areas (*blue*) in the FUV bad pixel table where the artifacts occur. Lower panel: Extracted spectra with (*red*) and without (*black*) event pulse height filtering, illustrating how the filtering removes the contaminating features and lowers the extracted counts somewhat. The effects of the artifacts range from distorting some line profiles to inducing false emission lines. Note that CALCOS cannot eliminate these features in ACCUM exposures.

In addition, for the NUV detector prior to SMOV, the PtNe data from the three stripes were added together (pixel by pixel, i.e. without regard to wavelength) to form a single collapsed spectrum prior to cross-correlation. This was done because it was originally thought that the shifts of all the stripes would be nearly the same, and combining the spectra would help to fill in regions where one stripe had few lines with good signal. But in fact the shifts can differ by several pixels from one stripe to another, and these differences lead to incorrect wavelengths as well as misalignments between FP-POS spectra in X1DSUM files. The current version of CALCOS uses a separate lamp template for each FP-POS position (with known offsets from the FP-POS=3 dispersion relation), and for NUV, each stripe is processed separately.

The wavelength calibration has been modified to make it more robust as well as more accurate. The shift between the wavecal spectrum and template lamp spectrum is now found by minimizing χ^2, rather than by maximizing the cross correlation. When a bright emission line is partially truncated by an edge of the NUV detector, the distorted line shape can bias the shift that is determined. CALCOS checks for such truncated lines and sets the wavecal data to zero in their vicinity. The improvement in wavelength accuracy for COS spectra is illustrated in Figure 3. See COS ISRs 2010–05 and –06 for details (Oliveira, 2010a and 2010b).

5. Customized Processing

CALCOS is not only used in the STScI pipeline, but runs in the STSDAS environment. This allows users to customize their own data processing. Also it is expected that GOs will develop their own tools to work either in concert with or in addition to CALCOS processing. Some updates have been made to CALCOS since SMOV to facilitate user processing.
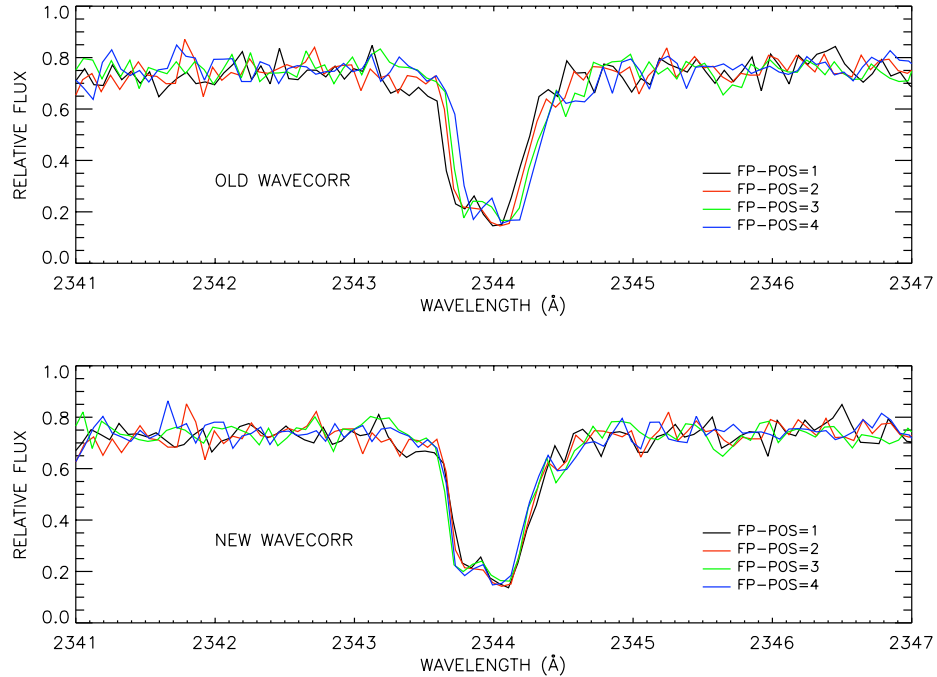
Figure 3: NUV G225M spectra of Feige 48 with the pre-SMOV and current CALCOS versions of WAVECORR. The FP-POS spectra with the old processing show wavelength misalignment errors that have been removed with the current technique.

The corrected time-tag event list (CORRTAG file) is the primary intermediate product most useful for customized processing. Two modifications to CALCOS have been made for this purpose so far. CALCOS can now accept CORRTAG files as input, not just raw files. This lets a user do some of the processing by setting some but not all of the calibration switch keywords to PERFORM and running CALCOS on the raw file. Then calibration switches can be reset in the CORRTAG header, the CORRTAG data can be modified (e.g. by applying one's own flat field), and then CALCOS can be rerun with the modified CORRTAG file as input. The CORRTAG file itself has also been modified; a new wavelength column has been added to the table. After wavelength calibration is computed as part of preparing for spectral extraction, CALCOS assigns a wavelength to each event in the file. This will allow users to combine and extract spectra directly without converting the data into images. Table 1 lists the columns in the updated version of the CORRTAG file.

There are two new tasks (in addition to CALCOS) in the hstcos package in STSDAS in PyRAF. The **x1dcorr** task extracts a 1-D spectrum (or spectra, if NUV), given a COR-RTAG file as input. This task creates FLT and COUNTS files from the CORRTAG file and then extracts the spectrum from the FLT and COUNTS files, using the same code as in CALCOS. The **splittag** task takes a CORRTAG file and splits it by time into multiple files in the same format, i.e. CORRTAG. Then running CALCOS on the output files from **splittag** will generate a spectrum for each of those time intervals.

6.   Other Improvements

The COS reference files are updated periodically as calibration studies are completed. Table 2 lists the reference files used by CALCOS, sorted by their installation dates, since SMOV and through July 2010. Many files created pre-SMOV and presently in use were not expected

Table 1: Corrected Event List (CORRTAG) Columns

Parameter	Type	Description
TIME	FLOAT	Time of event from beginning of exposure (seconds)
RAWX	INT	Uncorrected position along dispersion (pixel)
RAWY	INT	Uncorrected position cross dispersion (pixel)
XCORR[†]	FLOAT	Thermally and geometrically corrected RAWX (FUV)
YCORR[†]	FLOAT	Thermally and geometrically corrected RAWY (FUV)
XDOPP	FLOAT	XCORR corrected for HST Doppler shift at this time
XFULL	FLOAT	XDOPP shifted to wavecal frame
YFULL	FLOAT	YCORR shifted to wavecal frame
WAVELENGTH[‡]	**FLOAT**	**XFULL wavelength assignment (Å)**
EPSILON	FLOAT	Deadtime, flat field correction factor
DQ	INT	Data quality flag
PHA	BYTE	Pulse height amplitude (0–31)

[†]For NUV, XCORR, YCORR are the same as RAWX, RAWY since no correction is needed
[‡]Newly added column

to change. Some processes (bad time and burst corrections) have not been needed. Time dependent sensitivity corrections began in July 2010.

Other enhancements have been incorporated into CALCOS since SMOV.

- Corrections for NUV vignetting have been integrated into the flat field file (Ake, Burgh & Penton 2010); in the future vignetting may be handled with its own reference file.

- Data quality flags have been redefined to mark additional detector characteristics (Table 3). One particular change is that the DQ=4 flag, which had only been applied to FUV grid wire locations, is now used for the NUV to flag regions corrected for vignetting. While CALCOS should exclude wire shadows when creating an FUV X1DSUM spectrum since no flat field is available, for the NUV channel the vignetted regions can be used in the sum since they are corrected by its flat field. The difference in processing is controlled by having the detector shadow bit in SDQFLAGS set to 1 for the FUV and 0 for NUV.

- The data quality extension of the FLT and COUNTS files are more accurately shifted by the wavecal offset and Doppler shift, which is important when combining FP-POS stepped spectra.

- The usefulness of the G140L grating below 1150 Å has been demonstrated during SMOV, requiring wavelength calibration changes since the PtNe lamps do not extend into that region.

- Several keywords have been added to preserve such information as the number of events flagged as bad or exposure time lost due to FUV bursts, and keywords have been added for spectral-location information.

Further improvements are planned for the future.

- The acceptable range of pulse heights will need to be changed to track the changes in FUV gain due to exposure to bright sources. Currently a single pulse height range is applied for an entire FUV detector segment; however, the gain varies depending on position on the detector, so the PHA filtering should also vary with position. Filtering

Table 2: Chronological Listing of Reference File Installations

File Type	FUV File Name	NUV File Name	Description
Pre-SMOV			
FLATFILE	n9n20182l_flat.fits		Flat field image (unused)
GEOFILE	s7g1700cl_geo.fits		FUV geometric correction
BADTTAB	s7o1739kl_badt.fits		FUV bad times (unused)
BPIXTAB		s7g1700pl_bpix.fits	Bad pixel locations
BRFTAB	s7g1700el_brf.fits		FUV reference frame
BRSTTAB	s7g1700fl_burst.fits		FUV burst rejection (unused)
DEADTAB	s7g1700gl_dead.fits	s7g1700ql_dead.fits	Deadtime correction
TDSTAB	t2314312l_tds.fits	t2314314l_tds.fits	Time dependent sensitivity (unused)
WCPTAB		t2314313l_wcp.fits	Wavelength calibration parameters
Sep 11, 2009			
FLATFILE		t9b18111l_flat.fits	Flat field image
XTRACTAB	t9b1608jl_1dx.fits	t9b1608kl_1dx.fits	Spectrum extraction regions
WCPTAB	t9b18112l_wcp.fits		Wavelength calibration parameters
Sep 14, 2009			
DISPTAB	t9e1307kl_disp.fits	t9e1307ll_disp.fits	Wavelength dispersion relations
Sep 17, 2009			
FLUXTAB		t9h1220sl_phot.fits	Sensitivity curves
Oct 19, 2009			
FLUXTAB	taj18183l_phot.fits		Sensitivity curves
Jan 29, 2010			
DISPTAB	u1t1616ml_disp.fits	u1t1616pl_disp.fits	Wavelength dispersion relations
LAMPTAB	u1t1616nl_lamp.fits	u1t1616ol_lamp.fits	PtNe lamp templates
WCPTAB	u1t1616ql_wcp.fits		Wavelength calibration parameters
PHATAB	u1t1616ll_pha.fits		Pulse height screening
Mar 10, 2010			
BPIXTAB	u381724gl_bpix.fits		Bad pixel locations
Apr 14, 2010			
XTRACTAB	u4d1930sl_1dx.fits	u4d1930tl_1dx.fits	Spectrum extraction regions
Apr 29, 2010			
FLUXTAB	u4t18348l_phot.fits		Sensitivity curves
Jun 28, 2010			
LAMPTAB	u6s1320rl_lamp.fits		PtNe lamp templates
DISPTAB	u6s1320ql_disp.fits		Wavelength dispersion relations
DISPTAB	u6s1320pl_disp.fits		Wavelength dispersion relations
Jul 13, 2010			
TDSTAB	u7d20377l_tds.fits	u7d20378l_tds.fits	Time dependent sensitivity

Table 3: Original and New COS Data Quality (DQ) Flag Assignments

Bit	Decimal	Old Description	New Description	SDQFLAGS[†] FUV	NUV
1	1	Reed-Solomon error	Reed-Solomon error	0	0
2	2	Brush mark	Detector flaw	0	0
3	4	Grid wire	Detector shadow	1	0
4	8	Spectrum near detector edge	Spectrum near detector edge	1	1
5	16	Dead spot	Dead spot	1	1
6	32	Hot spot	Hot spot	1	1
7	64	Burst	Burst	0	0
8	128	Out of bounds	Pixel out of bounds	1	1
9	256	Fill data	Fill data	0	0
10	512	PHA too low	PHA out of bounds	0	0
11	1024	PHA too high	Unused	0	0
12	2048	Bad time interval	Bad time interval	0	0
13	4096	Wavelength below 900 Å	Background feature	0	0
14	8192	Wrinkle (segment A)	Low gain area	0	0
15	16384	S distortion (segment A)	Unused	0	0

[†]An SDQFLAGS bit=1 excludes data with the corresponding DQ flag from the X1DSUM spectrum

by pulse height affects the flat field and flux calibration; CALCOS will need to check (via a header keyword) that the reference files are consistent with the PHA filtering currently being applied to the data.

- The FUV flat field calibration will likely be modified. Currently CALCOS divides by a 2-D image, as is done for most instruments. There is evidence that a 1-D flat for the relevant spectral extraction region works better, however, and if further investigation shows this to be the case, CALCOS will be modified to use this format for the flat field.

- CALCOS may be modified to check the high voltage, and calibration would be limited if the high voltage is out of bounds.

- The algorithm for detecting FUV bursts is based on what was done for FUSE (Far Ultraviolet Spectroscopic Explorer). If bursts are actually present in COS data, the burst-detection algorithm should be improved and customized for COS.

7. Summary

CALCOS and the COS data products continue to improve with time. Prior to SMOV, a prioritized list was created for CALCOS enhancements and bug fixes, and plans were in place to update reference files derived from on-orbit observations. A great deal was learned about the instrument during SMOV, and valuable calibration data were obtained during this period. Several new reference files were prepared based on SMOV data, including lamp template spectra, dispersion relations, spectrum extraction regions, and time-dependent flux calibration. Besides the planned changes to the software and data products, some unexpected instrument characteristics, such as the low-gain artifacts in the FUV detector and vignetting of the NUV channel, were corrected. Other updates, like FUV flat fielding and calibrations below 1150 Å, are still on the horizon.

References

Ake, T. B., Burgh, E. B. & Penton, S. V. 2010, COS ISR 2010-03, "COS Near-UV Flat Fields and High S/N Determination from SMOV Data"

Ake, T. B. et al. 2010, this volume

Ghavamian, P. et al. 2010, this volume

Oliveira, C. 2010a, COS ISR 2010-05, "SMOV: COS NUV Wavelength Calibration"

Oliveira, C. 2010b, COS ISR 2010-06, "SMOV: COS FUV Wavelength Calibration"

Oliveira, C. et al. 2010, this volume

Osten, R. et al. 2010, this volume

Osterman, S. et al. 2010, this volume

Penton, S. et al. 2010, this volume

Sahnow, D. et al. 2010, this volume

Zheng, W. et al. 2010, this volume