OCO D-55206

# Orbiting Carbon Observatory – 2 (OCO-2)



# Level 1B
# Algorithm Theoretical Basis

Version 1.2 Rev 1
June 10, 2015

National Aeronautics and
Space Administration

**JPL**

Jet Propulsion Laboratory

California Institute of Technology

Pasadena, California

# ORBITING CARBON OBSERVATORY (OCO) - 2
## LEVEL 1B
## Theoretical Basis Document

Annmarie Eldering    Jet Propulsion Laboratory, California Institute of Technology
Randy Pollock    Jet Propulsion Laboratory, California Institute of Technology
Richard Lee    Jet Propulsion Laboratory, California Institute of Technology
Robert Rosenberg    Jet Propulsion Laboratory, California Institute of Technology
Fabiano Oyafuso    Jet Propulsion Laboratory, California Institute of Technology
David Crisp    Jet Propulsion Laboratory, California Institute of Technology
Lars Chapsky    Jet Propulsion Laboratory, California Institute of Technology
Robert Granat    Jet Propulsion Laboratory, California Institute of Technology

Approved by:

_____

Version 1.3
June 10, 2015

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

## Document History

| Version | Revision | Date | Description/Comments |
|---------|----------|------|----------------------|
| 1.0 | 0 | April 14, 2014 | Initial version of OCO-2 L1B ATBD |
| 1.1 | 0 | December 30, 2014 | Updated in preparation for L1B public release |
| 1.2 | 0 | March 17, 2015 | Version submitted to L1B review panel |
| 1.2 | 1 | March 30, 2015 | Version delivered to DAAC alongside L2 release |
| 1.3 | 0 | June 10, 2015 | Version updated for release with version 7 and 7r data |

**TABLE OF CONTENTS**

# LIST OF FIGURES

## LIST OF TABLES

# 1    Scope of This Document and Background

The OCO-2 Data User's Guide discusses the OCO-2 mission objectives and measurements, including instrument characteristics. Please refer to that document for details of the data collection and naming convention. The focus of this algorithm theoretical basis document (ATBD) is on the Level 1B data and the process used to transform the inherent instrument measurements (L1A data) into radiometrically calibrated spectra (L1B data, Figure 1-1). The reader is encouraged to consult the published articles referenced in this document for details concerning the analyses of the instrument test data.
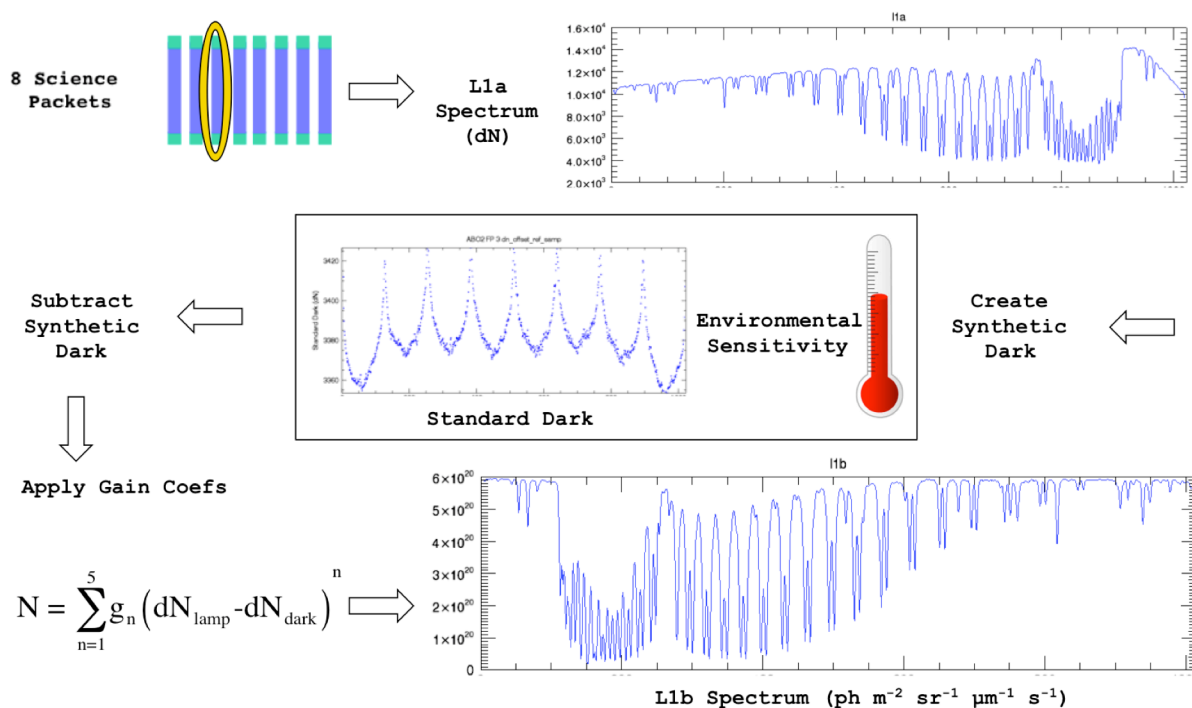


**Figure 1-1.** Overview of the ground processing data flow that results in L1B data.

## 1.0    Updates since Version 1.2, Rev 1 (30 March 2015)

After the release of the OCO-2 Build 6 data, one mistake and one flawed process were discovered in the algorithm used to adjust the gain coefficients used to convert L1A counts to L1B radiances.  Both were restricted to samples where the instrument bad pixel map had changed between ground testing and flight.  Since we have yet to update the A-band bad pixel map, these issues affected only the $CO_2$ bands.

The issues were corrected by rederiving the polynomial gain coefficients described in Section 2.2.  In Build 6, when a change in the gain coefficients was necessitated by a change in the bad pixel map, values would be calculated from the original instrument preflight test results.  While the primary testing was done in "sample mode", i.e. with the instrument performing onboard averaging of the data (see Section 1.1.2), a small amount of the testing was done in single pixel mode.  The frame rate in single pixel mode is a factor of 28 lower than in sample mode, so, given testing time constraints, these measurements by necessity have larger uncertainty.  Our original

approach, when updating the bad pixel map, was to generate new, synthetic sample DN values by applying the new bad pixel map to the single pixel data and then summing after applying the single pixel gains. This approach has the advantage of accounting for differences in quantum efficiency and linearity on a pixel-by-pixel basis, rather than assuming these characteristics to be uniform across the pixels in a sample. Contrary to our expectations, however, capturing these differences did not noticeably improve sample gain accuracy. In fact, any potential benefits from this approach were outweighed by the negative effects of mixing coefficients derived from data sets with different signal-to-noise characteristics. We use cubic polynomials to describe the nearly linear response of the FPAs. Performing the fits with the higher uncertainty single pixel data led to a different balance of the small cubic and quadratic terms, with the signs of these terms flipping back and forth unpredictably from sample to sample. At higher signal levels, radiances from the two populations of gains would then diverge. This separation of neighboring columns was interpreted by the retrieval algorithm as atmospheric information, thereby reducing the quality of the retrieval.

In Build 7, we now simply apply a scaling factor to the original calibration coefficients whenever a bad pixel map change results in a change in the number of effective pixels in a sample. More specifically, if a new bad pixel singlet or doublet is flagged in a sample, the onboard averaging of neighboring pixels (see Section 2.0) is assumed to compensate for the change, and no change to the gain coefficients is made. On the other hand, if a change in the bad pixel map results in $n \geq 3$ contiguous bad pixels in a sample, then the onboard software excludes those pixels from the sum, and the number of effective pixels in the sample drops from 20 to $20 - n$. If this happens, then the original gain coefficient is simply inflated by a factor of $^{20}/_{(20-n)}$ to correct for the missing pixels. Tests have recently shown that this leads to more accurate and consistent radiances in the vast majority of samples. The few where it does not work well are being marked as bad samples in the L1B data products and should be ignored by the user.

The mistake in the B6 data came about from an attempt to correct for slight errors in the footprint-to-footprint calibration. Since OCO-2 flies a short portion of each orbit with the slits aligned with the direction of spacecraft motion, it is possible to use the mean continuum radiances observed in a small, appropriately chosen latitude band as a relative calibration among the footprints. Applying this technique revealed a small ($< 2\%$) variation among the footprints that was consistent across all three bands. We ascribe this to uncharacterized spatial variations in the illumination inside the integrating sphere used for preflight testing. To correct for this effect, each gain coefficient is adjusted with a constant multiplier for each footprint, making the eight footprints more consistent while not changing the mean radiometric response of the spectrometer as a whole. Unfortunately, in the original software, this step was performed before the gain coefficents were adjusted for bad pixel map changes. This led to an offset of nearly one percent in a sizeable subset of the samples. Given that the instrument signal to noise ratio is typically several hundred to one, these offsets were much larger than the instrument noise and created significant problems in the XCO2 retrivals. Since the distribution of bad pixels is not uniform in either the spectral or spatial directions of the FPAs, this error introduced a significant, spurious footprint-dependent bias.

Following the discovery of these issues in the Build 6 calibration, extensive testing was performed to validate the changes in preparation of the gain coefficients, and to evaluate each ancillary file delivery. The team is confident that the resulting Level 2 products will be of much better quality in Build 7.
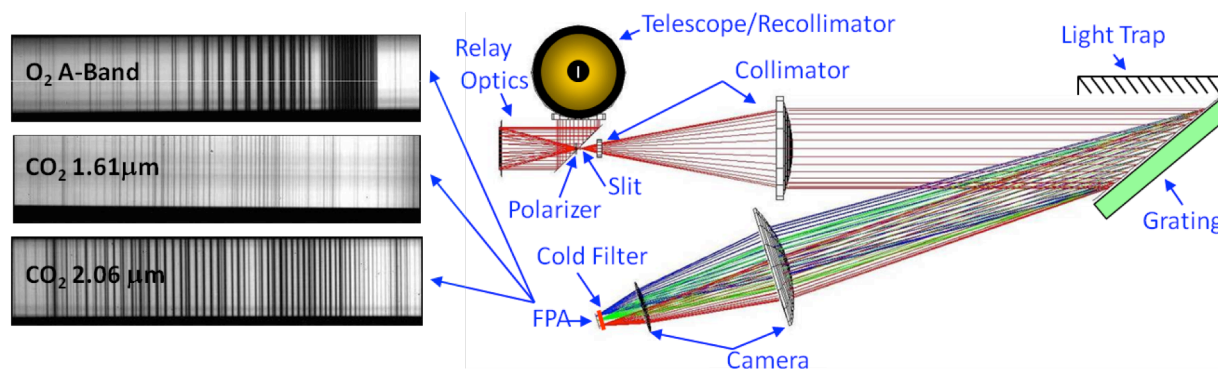
**Figure 1-2.** The OCO-2 instrument showing the major optical components and optical path (**right**) and images of spectra recorded by the FPA in the three spectral channels (**left**).

## 1.1   Instrument Characteristics

### 1.1.1   Instrument Optical Path

The OCO-2 instrument incorporates three, co-boresighted, long-slit, imaging grating spectrometers optimized for the molecular oxygen ($O_2$) A-band at 0.765 microns (μm) and the carbon dioxide ($CO_2$) bands at 1.61 and 2.06 μm. The three spectrometers use similar optical designs and are integrated into a common structure to improve system rigidity and thermal stability. They share a common housing and a common F/1.8 Cassegrain telescope (Figure 1-2). Light entering the telescope is focused at a field stop and then re-collimated before entering a relay optics assembly. There, it is directed to the three spectrometers by dichroic beam splitters, and then transmitted through a narrowband pre-disperser filter. The pre-disperser filter for each spectral range transmits light with wavelengths within ~1% of the central wavelength of the $CO_2$ or $O_2$ band of interest and rejects the rest. The light is then refocused onto the spectrometer slits by a reverse Newtonian telescope.

Each spectrometer slit is about 3 mm long and about 25 μm wide. These long, narrow slits are aligned to produce co-boresighted fields of view that are ~0.0005 radians wide by ~0.0146 radians long. Because the diffraction gratings efficiently disperse only the light that is polarized in the direction perpendicular to the long axis of the slit, a linear polarizer was included in front of the slit to reject the unwanted polarization before it enters the spectrometer, where it could contribute to the scattered light background. Once the light traverses a spectrometer slit, it is collimated by a two-element refractive collimator, dispersed by a gold-coated, reflective, planar, holographic diffraction grating, and then focused by a two-element camera lens onto a focal plane array (FPA), after traversing a second, narrowband filter. The narrowband filter just above the FPA is cooled to ~180 K to reject thermal emission from the instrument.

An onboard calibrator (OBC) has been integrated into the telescope baffle assembly [Crisp et al., 2008]. This system consists of a calibration "propeller" that carries an internal reflective diffuser on one end and an external transmission diffuser on the other. The reflective diffuser is illuminated by one of three lamps installed internal to the baffle assembly to provide flat field illumination to monitor changes in individual pixel response. It is also used when acquiring "dark frames" that are used to monitor the zero-level offset of the FPAs. The transmission diffuser is used when observing the Sun. Routine observations of the Sun are acquired just after the spacecraft crosses the northern terminator on approximately 12 orbits per day. These measurements are used to monitor changes in the radiometric calibration of the instrument.
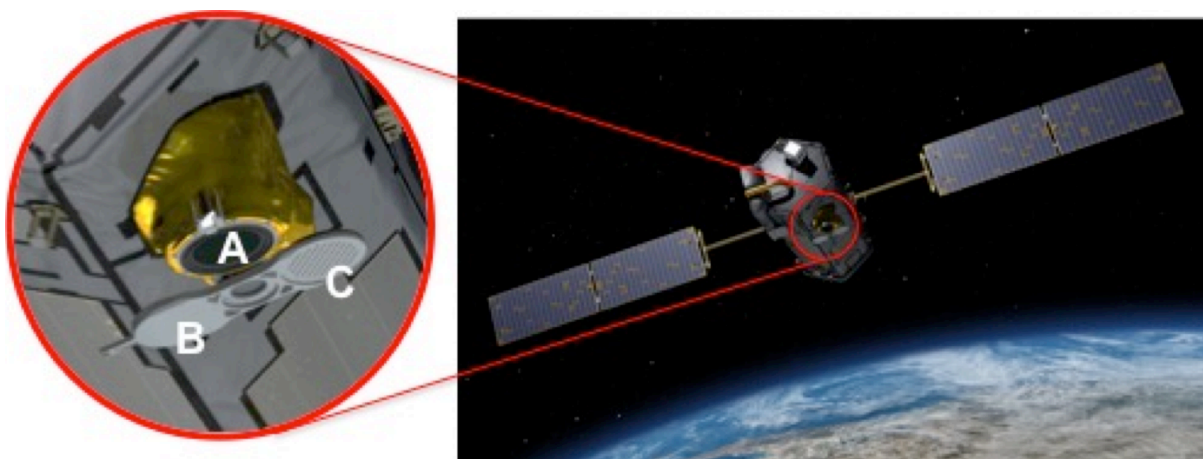
**Figure 1-3.** The Onboard Calibrator (OBC) is attached to the telescope baffle and can be rotated to place diffusers in front of the aperture. **(A)** is the entrance aperture of the telescope, **(B)** is the internal diffuser used for lamp and dark calibration, and **(C)** is the transmission diffuser used for solar observations.

Approximately once per month, science observations are replaced by a full dayside orbit of solar spectra. This samples ~ ±5 km/sec of Doppler shifts which provide high resolution sampling of solar lines to monitor the instrument line shape. For normal science observations, the propeller is in the position shown in Figure 1-3 so that light from the Earth enters the telescope directly.

### 1.1.2 Focal Plane Array Detectors

The spectrometer optics produce an image of a spectrum on a 1024 by 1024 pixel FPA with 18 μm pixels (Figure 1-4 and Figure 1-5). The grating disperses the spectrum onto 1016 of the 1024 FPA columns (4 columns are blanked out on each side of the FPA) in the direction perpendicular to the long axis of the slit. The full width at half-maximum (FWHM) of the slit image on the FPA (also called the 'Instrument Line Shape' or ILS) is sampled by 2 to 3 pixels in the direction of dispersion. The 3 mm long slit limits the spatial field of view to only ~190 pixels in the dimension orthogonal to the direction of dispersion. Science measurements are restricted to the middle ~160 of these 190 pixels.

For normal science operations, the FPAs are continuously read out at 3 Hz. A "rolling readout" scheme has been adopted for reading out and resetting the FPAs, obviating the need for a physical shutter and gaps between the exposures. To reduce the downlink data volume and increase the signal to noise ratio, ~20 adjacent pixels in the FPA dimension parallel to the long axis of the slit (i.e., the spatial direction in Figure 1-4a) are summed on board to produce up to eight spatially-averaged spectra along the slit. Each ~20 pixel sum is referred to as a "spectral sample." The angular field of view is defined as by the 1016 spectral samples is defined as a "summed footprint."

The along-slit angular field of view of each of these spatially-averaged spectral samples is ~1.8 milliradians (0.1° or ~1.3 km at nadir from a 705 km orbit). The angular width of the narrow dimension of the slit is only 0.14 milliradians, but the focus of the entrance telescope was purposely blurred to increase the effective full width at half maximum of each slit to ~0.5 milliradians to simplify the boresight alignment among the 3 spectrometer slits and to minimize the impact of bad pixels. Because it takes 0.333 seconds to scan across the 220 active rows of the array and the spacecraft is moving during the rolling readout, the surface footprints are shaped liked parallelograms rather than squares when the slit is oriented orthogonal to the ground track (Figure 1-4b).

In addition to the eight spatially-binned, 1016-element spectra, each spectrometer returns up to 20 columns (colors) from each FPA without any onboard spatial binning to sample the full along-slit spatial resolution. Each of these full-resolution "color slices" images a 220-pixel wide region of the FPA that includes the full length of the slit (~190 pixels) as well as a few pixels beyond the ends of the slit (Figure 1-4a). These full-spatial-resolution "color slices" are used to detect spatial variability within each of the spatially summed spectral samples and to quantify the thermal emission and scattered light within the instrument. The locations of the 20 color slices can be specified by commands from the ground. Note that the first and last four of the 220 pixels in each of the color slices do not contain valid radiometric information, but are included in the data products for engineering reasons.

In this instrument design, the spectrometer slits, the grooves on the diffraction gratings, and columns of the FPAs must be well aligned to ensure that a fixed series of rows on the FPA will sample the same angular field of view (or spatial footprint) throughout the spectral range recorded by the FPA. For the OCO-2 instrument, perfect alignment of the FPAs with the other optical components was not possible due to a physical obstruction discovered late in the instrument assembly process. The focal plane arrays are therefore slightly rotated (or "clocked") with respect to the slit and grating. Consequently, a given geographic position does not map onto a single row of pixels on a sensor, but instead varies (roughly linearly) with spectral position (i.e., column). To compensate for this, and record the same spatial information across the entire spectrum, the starting row index for each spectral sample can be adjusted in increments of one
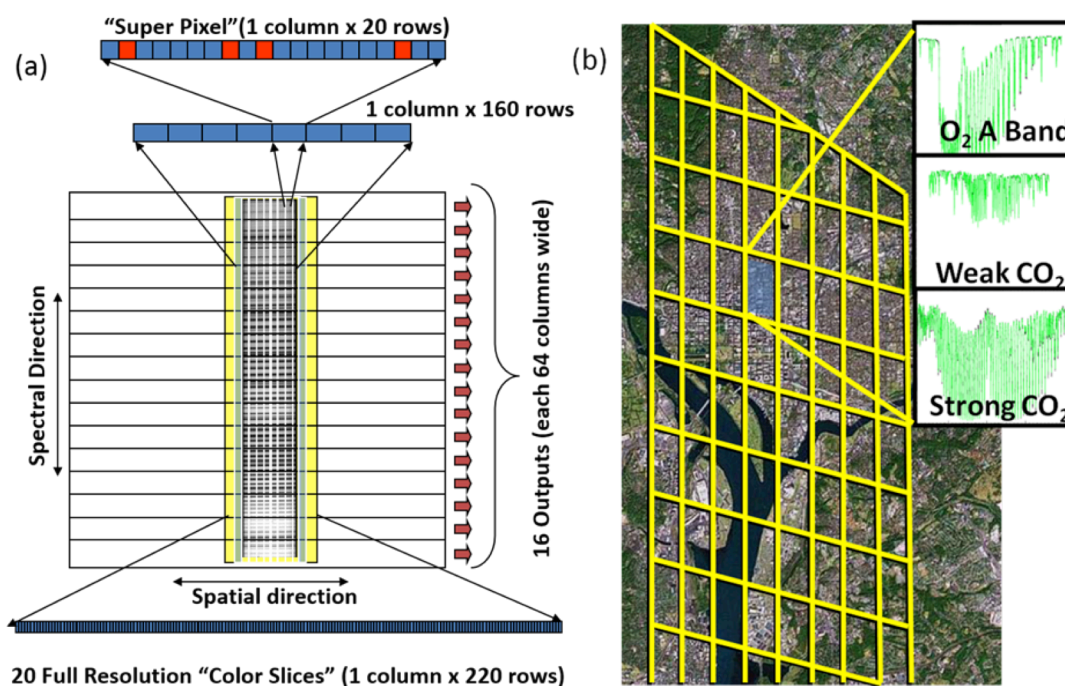


**Figure 1-4. (a)** The illumination and readout scheme used for the OCO-2 FPAs, showing the direction of spectral dispersion from bottom to top, and the spatial dimension from left to right. The ~160 illuminated pixels in the spatial dimension are summed into eight 20-pixel "footprints." If one or more of the pixels in a footprint is "bad" (red pixels), it is eliminated from the sum. If one or two contiguous pixels in a column are bad, then they are replaced with the average of their good neighbors. If three or more are contiguous, then they are replaced with zeros. This algorithm is only applied in the spatial direction—spectral information is never mixed. One of the 20 full-resolution "color slices" is also shown at the bottom. **(b)** Spatial layout of 8 cross-track footprints for nadir observations over Washington, D.C. Each footprint is shaped like a parallelogram, rather than a square, because of the rolling readout of the FPAs and the spacecraft motion.

pixel. This corresponds to about 1/20th of a summed footprint (Figure 1-5). This approach introduces little error in spatially homogeneous scenes, but can produce radiance discontinuities in spectra of scenes with strong intensity variations near the edge of a footprint. These discontinuities are corrected as part of the calibration process (see Section 4.4).
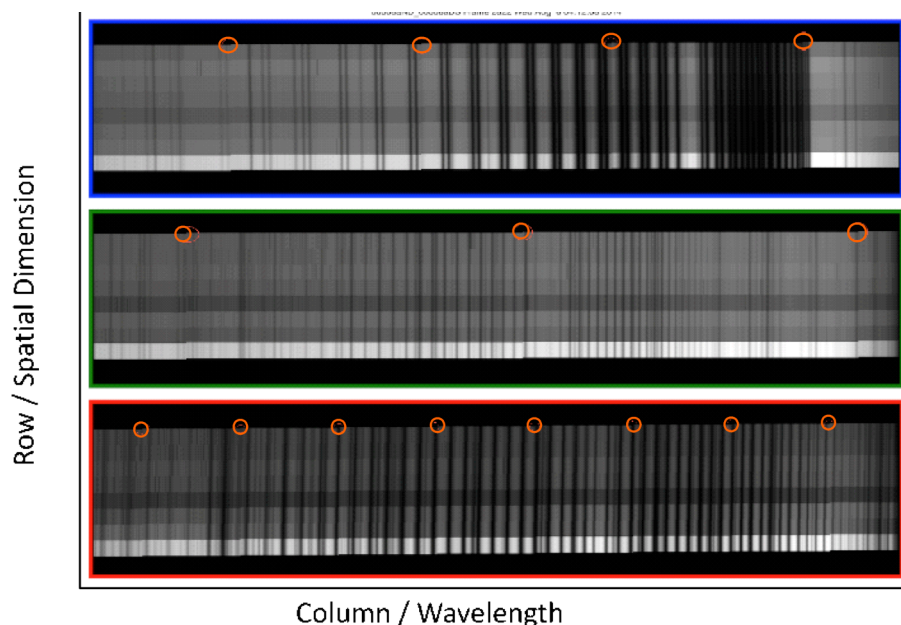


**Figure 1-5.** Images of the $O_2$ A-band **(top)**, 1.61 **μm** $CO_2$ band **(middle)** and the 2.06 **μm** $CO_2$ band **(bottom)** in "summed mode" taken from the OCO-2 first light frame over Papua New Guinea. The 8 spatially-summed footprints in each band are shown from bottom (footprint 1) to top (footprint 8). In this 0.333 second frame, footprint 1 was contaminated by a cloud and appears brighter than the rest. This frame clearly shows the rotation, or "clocking" of each FPA columns with respect to the dark $O_2$ and $CO_2$ absorption lines. The $O_2$ lines are tilted slightly counterclockwise, while the 1.61 and 2.06 **μm** lines are tilted clockwise. The columns where the clocking corrections are applied are highlighted with red circles. Soundings collected in regions with strong spatial gradients in illumination can produce radiance discontinuities at these wavelengths.

### 1.1.3 Bad Pixels and Bad Samples

OCO-2 is flying OCO flight-spare FPAs, which were delivered in 2006. A few percent of the 225,280 pixels in the active area of these arrays are either dead, or respond to light or thermal changes in a way that is not consistent with the majority of the other pixels. These pixels must be identified and excluded from the 20-pixel sums that are performed onboard, or they will contaminate the resulting 20-pixel "spectral samples" that constitute the spectra returned for the eight footprints by each spectrometer. Spectral samples with too many bad pixels, or with other known issues (e.g., optical artifacts on the short-wavelength ends of all three bands) are marked as bad in the L1B product, and should only be used with caution (almost all users should ignore them in higher-level processing).

The bad pixel identification process was started during prelaunch testing, but must be updated on orbit because additional pixels have degraded during the 2.5 year instrument storage period prior to launch, and others are damaged on-orbit by cosmic radiation, thermal cycling of the FPAs, and other effects. To identify bad pixels, the calibration team routinely collects dark and lamp data using "single pixel" mode, which returns all of the pixels in the active region of each array. (Single-pixel results cannot be returned routinely for science observations because it requires

~9 seconds to read out an FPA at full resolution and therefore would lead to a ~97% reduction in obtained spectra.) These single-pixel calibration data are then analyzed to identify bad pixels and add them to a "bad pixel map" that is uploaded to the spacecraft for use in the pixel-summing process. As of June 8, 2015, this has been done four times since launch (see Table 1-1). As the bad pixel maps improve, a larger faction of the spectral samples in each spectrum can be used to generate L2 products.

**Table 1-1.** – History of Bad Pixel Maps as of June 3, 2015. The version numbers show the three bads (A-band / Weak CO2 / Strong CO2). As can be seen, the A-band bad pixel map has not changed so far in the mission..

| Bad Pixel Map Version # | Date Loaded | Notes |
|---|---|---|
| 5 / 5 / 5 | 2014-07-02T09:56:00.000Z | This bad pixel map was derived during preflight tests in 2012 and used during the OCO-2 checkout period. No science data are being delivered from this period. |
| 5 / 6 / 6 | 2014-09-05T20:06:06.290Z | These updates to the $CO_2$ band bad pixel maps removed pixels that had lost sensitivity to light between initial characterization and launch. |
| 5 / 7 / 7 | 2014-11-10T14:56:20.120Z | These bad pixel maps went through ground validation but version 8 were ready before they could be loaded. No data were taken in space using these maps. |
| 5 / 8 / 8 | 2014-11-21T16:16:54.188Z | First updates to $CO_2$ bad pixel maps to mask out pixels with very high noise or very unstable dark current. |
| 5 / 9 / 9 | 2015-02-13T15:53:08.785Z | Improvement to 5 / 8 / 8 based on initial review of data |
| 5 / 10 / 10 | 2014-07-02T09:56:00.000Z | Further refinements to remove additional noisy/unstable pixels. |

## 1.1.4 Cosmic Ray Artifacts

Cosmic rays rarely produce permanent damage to the OCO-2 FPAs, but energetic particles produce ion trails as they traverse the FPAs, producing spurious intensity spikes. The largest effects are seen in the $O_2$ A-band, because the FPA in this channel is more sensitive to cosmic ray artifacts. Primary and secondary cosmic rays with a broad range of energies are occasionally seen everywhere along the orbit path but they are most common in the vicinity of the South Atlantic Anomaly (SAA), where up to 2% of an $O_2$ A-band spectral range can be contaminated in individual frame. The approximate geographical extent of the SAA as it impacts the OCO-2 mission can be seen in Figure 1-6. The characteristic spectral signature of cosmic rays is shown in Figure 1-7.

An algorithm has been developed to identify and screen these events in this L1B data delivered



meas-mod < 10sigma (A-band)

| 1000 | 1001 | 1002 | 1003 | 1004 | 1006 | 1007 | 1008 | 1009 | 1011 | 1012 | 1013 | 1014 | 1016 |
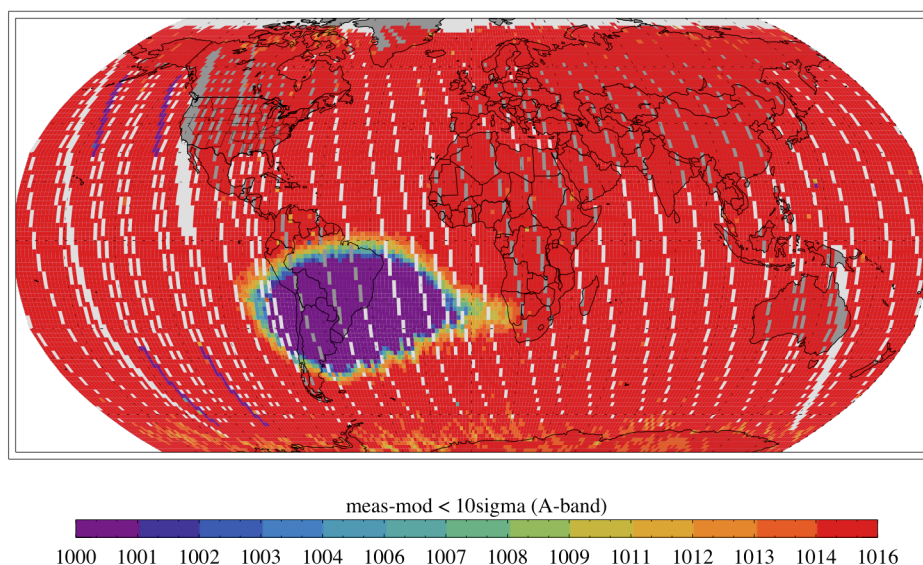
**Figure 1-6.** The number of spectral samples in the $O_2$ A band that are not contaminated by 10-sigma cosmic rays events clearly shows the geographical extent of the South Atlantic Anomaly (SAA).
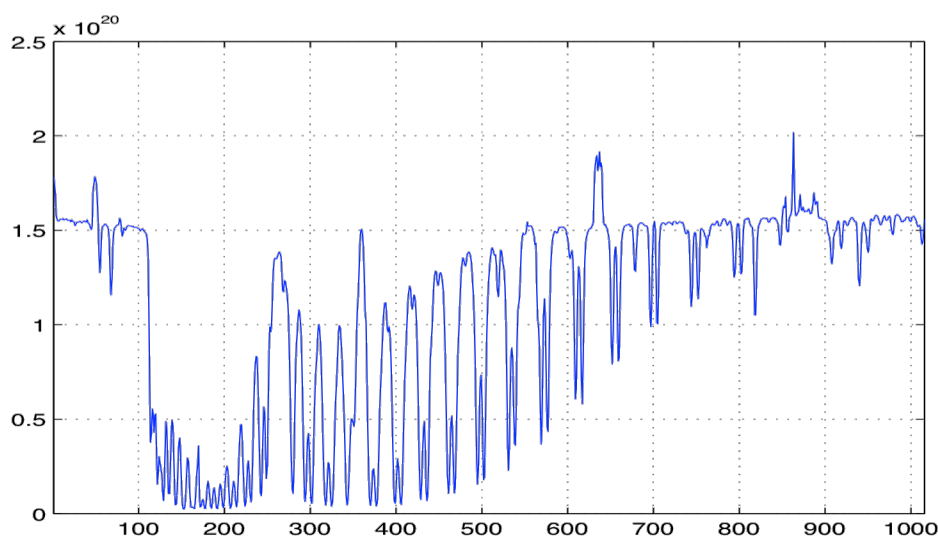


**Figure 1-7.** An example showing the impact of a radiation event on spectra. Note the three spikes with unusually high intensities.

to the science community. This algorithm includes the following steps:

1. For each L1 granule, a singular value decomposition is performed independently for each footprint and spectral band
2. Using the 40 leading eigenvectors, a least squares approach fits the optimal linear combination of all 40 eigenvectors to match individual spectra within the granule. This represents an empirical fit to all spectra at a computational speed that is not prohibitive for L1B processing. For strong outliers (>10 sigma), spectral samples are flagged and refitted after excluding outliers that strongly perturb the linear fit
3. The fit mismatch (residual), weighted by the expected detector noise is calculated "(meas-mod)/noise" and saved for each spectral sample.

In the L1B HDF5 files, these weighted residuals are saved in the group SpikeEOF as rounded integer values with all residuals less than 3 three times the noise level (3 sigma) set to 0 to minimize file size. The variables used to record the weighted residuals in the L1B files are:

SpikeEOF/spike_eof_weighted_residual_*band* (with *band* = o2, strong_co2, weak_co2)

The total number of spectral samples contaminated in by cosmic rays in any one spectrum are recorded in the variables:

SpikeEOF/spike_eof_bad_colors_*band* (with *band* = o2, strong_co2, weak_co2)

The OCO-2 L2 FP algorithm uses the weighted residuals in the L1B file to flag individual spectral samples as bad if their weighted residuals exceeds +6 sigma, and if the sounding falls within a geographic boundary box around the SAA (latitude −50 to 0, longitude −90 to 10 degrees). We use a one-sided threshold value as cosmic ray hits can only cause a positive anomaly in measured spectra. L1B users are advised adopt a similar approach for flagging spectral samples contaminated by cosmic rays before attempting any fit. We also advise users to apply this threshold only in the vicinity of the SAA.

## 1.2   Radiometric Overview

The focal planes are read out in unitless data numbers (dNs), and these are referred to as the L1A Spectrum (see Figure 1-1). The raw DN values include contributions from the bias and dark current and any photoelectrons collected by each pixel or spectral sample in active area of the FPA. The first data processing step is to apply dark correction, which removes the bias and dark current from each spectral sample. As discussed in Rosenberg et al. [2015], this is important because a DN offset can introduce a linearity error. During thermal-vacuum testing, a set of data was collected with the light source shuttered such that no light entered the instrument. These dark data were combined to produce high signal-to-noise single pixel and summed "synthetic darks." These data are used to create a baseline for trending the instrument performance, and were used for dark corrections early in the mission. While on orbit, dark data are routinely collected during the eclipse phase of the vast majority of orbits. The dark data collected over several orbits are combined to improve their signal to noise ratio and are used to correct the zero-illumination response of the data included in the L1B product based on the predicted optical and focal plane temperatures.. The methods used to create and apply the in-flight synthetic darks are described in greater detail below.

The next step is to apply gain curves, which describe the relationship between DN and calibrated radiance. The preflight gain curves are described in more detail in Rosenberg et al. [2015]. The sections below describe the approach for in-flight updates.

## 1.3    Spectral Response Overview

The instrument line shape (ILS) describes the spatial distribution of the light on the FPA that is produced when light from the slit is collimated, dispersed into a spectrum by the grating, and reimaged on the FPA. In the OCO-2 instrument, the ILS changes from footprint to footprint, and across each summed spectrum. The OCO-2 ILS was characterized during pre-launch testing, using techniques similar to those used for the original OCO instrument [Day et al. 2011; Frankenberg et al. 2014]. To properly use the L1B data, the instrument line shapes (ILS) must be known. This information is contained in the L1BSc files in two key fields: ils_delta_lambda and ils_relative_response. ILS are covered in more detail below, and their use is described in the Data User's Guide.

The wavelength corresponding to each spectral sample must also be known. For OCO-2, this information is defined by a polynomial fit to the spectral dispersion. Dispersion coefficients are provided in the L1BSc files (dispersion_coef_samp). In the L2 retrieval of geophysical quantities, the Doppler shift is accounted for and the dispersion is corrected for each sounding.

## 1.4    In-Flight Measurements

Several types of data are collected to monitor the stability of the instrument data (Table 1-2). Where drifts are discovered, they lead to updated calibration parameters.

**Table 1-2.** Summary of inflight calibration measurements.

| Calibration Type | Approach, Purpose, and Results |
|---|---|
| Solar Calibration | • Approach: View the Sun through the transmission diffuser when in the vicinity of the northern terminator (shortly after final science measurements for a given orbit)<br>• Frequency: ~6x/day each in single pixel and summed modes<br>• Purpose: Provide absolute radiance standard<br>• Results: Very high precision (<<0.1% over a few hours)—stability of the calibrator over the mission life is within the lunar calibration repeatability (~1%) |
| Lunar Calibration | • Approach: Point the instrument boresight at a right ascension/declination location and allow the spacecraft orbital motion to scan the lunar disk<br>• Frequency: ~1x per lunar cycle (usually at ~75% waxing phase)<br>• Purpose:<br>  – Provide pointing validation<br>  – Provide relative radiometric calibration (to trend other calibration data, e.g., solar calibrator degradation)<br>• Results: Very high precision (<<0.1% over a few hours)—stability of the calibrator over the mission life is within the lunar calibration repeatability (~1%) |
| Lamp Calibration | • Approach: View one of three onboard lamps using the internal diffuser<br>• Frequency:<br>  – Lamp #1: ~6x/day each in single pixel & summed modes<br>  – Lamp #2: ~12x/year in both single pixel & summed mode<br>  – Lamp #3: ~3x/year in both single pixel & summed mode<br>• Purpose: Provide relative radiance standard with no spectral lines<br>• Results: Slow reduction in light output and color temperature thought to be due to lamp aging, but provides means to fill in solar lines observed in solar calibration |

| Calibration Type | Approach, Purpose, and Results |
|---|---|
| Dark Calibration | • Approach: Collect data with lamps off and internal diffuser blocking aperture<br>• Frequency: ~20x/day each in single pixel and summed modes<br>• Purpose: Monitor fixed pattern changes and fit out optics and focal plane temperature dependencies<br>• Results: Dark subtraction residuals no more than 20% background noise of instrument |
| Full Orbit Darks | • Approach: Collect data with lamps off and internal diffuser blocking aperture during the day side of the orbit<br>• Frequency: ~1x/month or after each decontamination cycle (note that when this occurs, an orbit of science data is lost)<br>• Purpose: Validate that dark model developed with data taken in eclipse is useful for correcting science data offsets<br>• Results: Data on day side is statistically similar to dark data taken in eclipse |
| Doppler Diffuse | • Approach: Maintain solar calibration during entire dayside of an orbit<br>• Frequency: ~1x/month or after each decontamination cycle (note that when this occurs, an orbit of science data is lost)<br>• Purpose: Take advantage of ~±5 km/s Doppler shift of the solar lines to validate the instrument line shape stability<br>• Results: Full width at half-maximum of ILS stability to <1% (repeatability of measurement) |
| Vicarious Calibration | • Approach: Track Railroad Valley, NV as spacecraft flies over collecting 1000s of spectra within a few km of the center of the playa<br>• Frequency: ~10x/year but frequency is driven by weather and field campaigns to collect simultaneous data at the site (note that when this occurs, ~1/4 orbit of science data is lost)<br>• Purpose: Validate absolute calibration of radiometric output<br>• Results: Differences between OCO-2 and ground measurements <10% but further work is needed to adapt GOSAT efforts [Kuze et al., 2011, 2013] to OCO-2's smaller footprint. Expect offsets to drop with refinements in analysis |

## 2   Radiometry

### 2.0   Bad Pixel Map

The application of the bad pixel map is invisible to users of the standard L1B product. The following is presented for information only.

The bad pixel map is a 1024 x 220 x 3 byte array where a value of 1 indicates that a pixel's signal should not be used. Bad pixels are corrected by the flight computer in space, while producing sample mode ("summed") data. The single pixel readings are never returned to the ground during nominal science operations (with the exception of the color slices). Within each ~20-pixel spectral sample, a single or double bad pixel is replaced by the average of the closest neighboring good pixels on each side. This is a good approximation because the spatial imaging of the telescope has been blurred to yield a point spread function with a full width of 3 – 6 pixels. If there are three or more contiguous bad pixels within a single spectral sample, those values are set to 0 and not included in the sum. As a result, a pixel can have a weight of 0, 1, 1.5, 2, 2.5, or 3 in the nominal 20 pixels that constitute a sample. Many sample mode calibration fields (gain, noise model, dark correction coefficients) must be updated with each bad pixel map change to reflect the new weighting.

### 2.0.1   Bad Pixel Map Usage

The bad pixel map is used only in SUMMED mode. The bad pixel map is not applied to data collected in single pixel mode. The bad pixel map also is not applied to the color slices. These data must be corrected separately by the user.

### 2.0.2   Summed Mode Pixel Map Processing

All bad pixels identified in a bad pixel map are replaced, as described above, before the onboard instrument software applies any pixel summation algorithms.

### 2.0.3   Bad Pixel Map/Footprint Relationship

The OCO-2 instrument digitizes an array 220 spatial pixels (rows) by 1024 spectral pixels (columns). Summed science data uses the "middle" 160 spatial pixels. This relationship is shown in Figure 2-1. The eight footprints are each approximately twenty pixels wide, with footprint limits that vary by band as shown in Table 2-1.

**Table 2-1.** Footprint limits.

| Footprint | ABO2 | WCO2 | SCO2 |
|:---:|:---:|:---:|:---:|
| 1 | 33 – 51 | 35 – 54 | 26 – 45 |
| 2 | 52 – 71 | 55 – 73 | 46 – 65 |
| 3 | 72 – 91 | 74 – 93 | 66 – 85 |
| 4 | 92 – 111 | 94 – 113 | 86 – 15 |
| 5 | 112 – 131 | 114 – 133 | 106 – 125 |
| 6 | 132 – 151 | 134 – 153 | 126 – 145 |
| 7 | 152 – 171 | 154 – 172 | 146 – 165 |
| 8 | 172 – 191 | 173 – 192 | 166 – 185 |
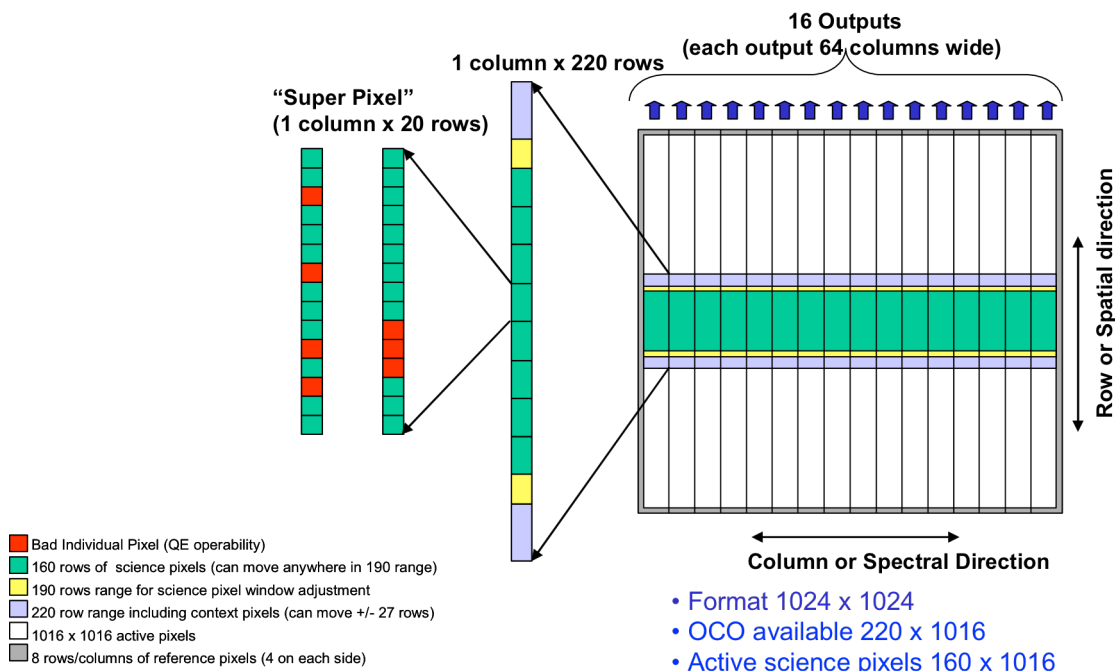
**Figure 2-1.** Focal plane readout definition.

The pixel map is always referenced to the 220 x 1024 digitized array. Thus, using the example, the smallest ABO2 band pixel map row value should be 33, and the largest pixel map row value should be 191. In order to be most useful, the pixel map limits should correspond to the limits defined in the footprint start and end limits.

### 2.0.4 Bad Pixel Mitigation Calculations

Bad pixels are corrected as shown in Table 2-2. Bad pixel correction is only performed in the spatial direction (along columns). These corrections are performed prior to the footprint summing process, without regard to footprint limits or boundaries. No interpolation or averaging is performed in the spectral direction.

**Table 2-2.** Bad pixel mitigation.

| Bad Pixel Configuration | Mitigation |
|---|---|
| Single bad pixel | Replaced with the average of the two spatially adjacent good pixel values |
| Double bad pixel | Replaced with the average of the two spatially adjacent good pixel values; the same value is used to replace both bad pixel values |
| Triple bad pixel | All bad pixels are replaced with zero values |

### 2.1 Dark Subtraction

Variations in the temperature of the optics and FPAs contribute to the dark background signal. The as-measured DN values can be corrected for these effects using the dark subtraction equation:

$$dn_{corrected} = \left( dn_{raw} - dn_{ref} \right) + c_{optics}(T_{optics}(t) - T_{ref\_optics}) + c_{FPA}(T_{FPA}(t) - T_{ref\_FPA})$$
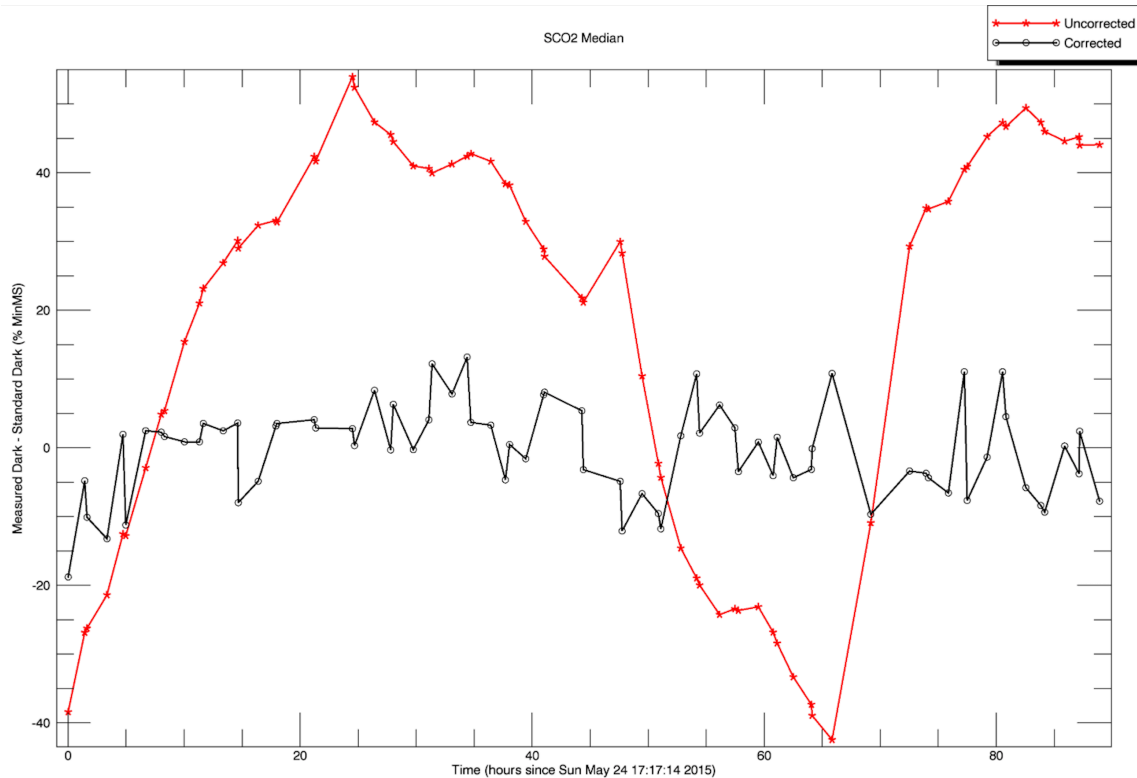
**Figure 2-2.** SCO2 dark signal level is sensitive to changes in the optical bench temperature, but these fluctuations diminish greatly after correction. Dark signal is expressed as a fraction of the noise floor as defined in Table 2-3.

where

- $dn_{corrected}$ —DN after dark correction
- $dn_{raw}$—as-measured DN
- $dn_{ref}$—DN detected with no illumination
- $c_{optics}$ —optics temperature dependence coefficients
- $T_{optics}(t)$ —optics temperature as a function of time (in L1B data product as FrameTemperatures/temp_smooth_optical_bench_grating_mz)
- $T_{ref\_optics}$ —the optics temperature that will yield dnref
- $c_{FPA}$ —FPA temperature dependence coefficients
- $T_{FPA}(t)$ —$O_2$, weak $CO_2$, and strong $CO_2$ FPA temperatures as a function of time (in L1B data product as FrameTemperatures/temp_smooth_fpa_[o2, weak_co2, strong_co2])
- $T_{ref\_FPA}$ —the $O_2$, weak $CO_2$, and strong $CO_2$ FPA temperatures that will yield dn$_{ref}$

$dn_{raw}$, $dn_{ref}$, $c_{optics}$, and $c_{FPA}$ have dimensions of 104 x 8 x 3—a value for each column, footprint, and band (Figure 2-3). These quantities are included in the Ancillary Radiometric Product (ARP) but are not provided in the L1B files.

The bias and dark current for an individual pixel is depend on voltage and temperature variations within the instrument. To minimize the science impact of these changes, dark correction performed in the conversion from L1A to L1B includes simultaneous adjustments for two temperatures. To avoid introducing errors due to noise in the temperature sensors, temperatures are smoothed before they are used in this correction algorithm. Note that all temperatures except

for the FPA temperatures are only updated every 15 seconds. Temperatures are smoothed over an entire orbit using a linear fit, performed before the orbit is broken into several files at mode change boundaries. One impact of this approach is a poor temperature fit (and thus dark correction) for orbits that include adjustments to the optics or focal plane temperatures. This will be addressed in future reprocessing.

**Table 2-3.** *MinMS* values.

| Band | *MinMS* value (photons/m$^2$/sr/μm) |
|---|---|
| $O_2$ A-band | $7.50*10^{16}$ |
| Weak $CO_2$ | $2.15*10^{16}$ |
| Strong $CO_2$ | $2.15*10^{16}$ |

## 2.2 Radiometric Gain

Because atmospheric absorption is inferred from the ratio between continuum and line core radiances, the radiometric calibration must account for any nonlinearities in the instrument gain. Although the instrument response appears almost linear, up to a 5th order polynomial of radiance vs. signal is provided for to describe any nonlinear effects:

$$Radiance = k \sum_{i=0}^{5} c_i \left( dn_{dark\_corrected} \right)^i$$

The scaling terms $k$ are stored in the variable gain_degrad_samp (k) and were set to 1.0 at launch. In-flight, the gain_degrad_samp terms are used to track changes in the instrument response throughout the mission through the vicarious calibration experiments described in
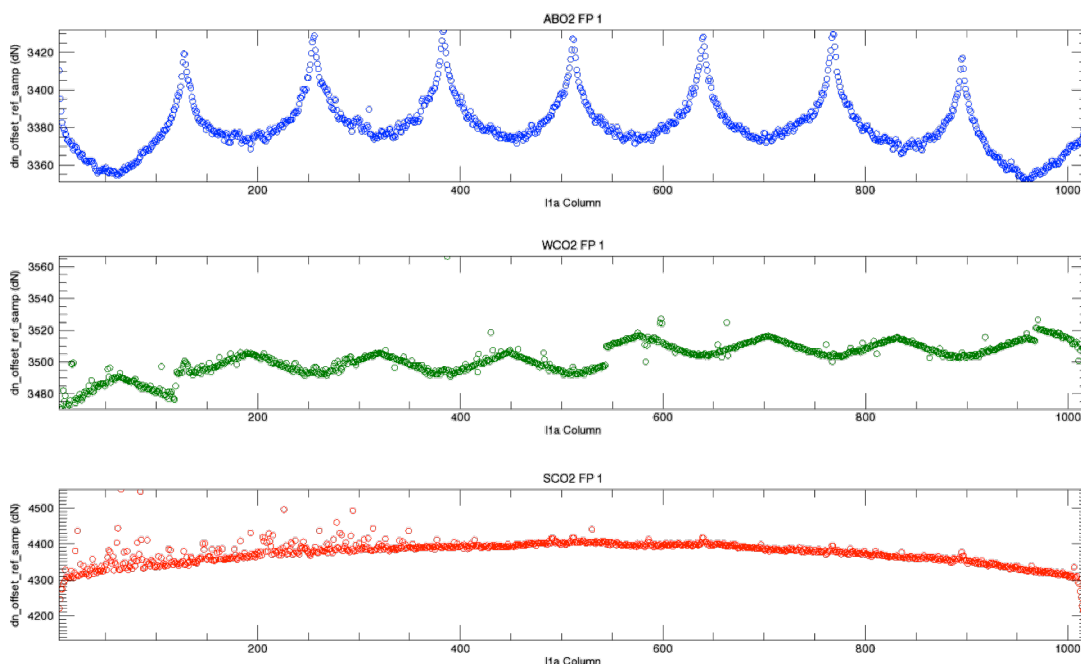


**Figure 2-3.** Example of a term in the dark subtraction ($dn_{raw}$). Spectral patterns in dark coefficients reflect features of the detectors (ABO2, WCO2) & optics (SCO2), and remove these artifacts from L1B spectra.

Section 1.4.

The $c_i$ terms are stored in the variable gain_preflight_samp, ($c_i$). At launch, these are set based on a combination of the best five radiometric calibration tests performed during final instrument testing at JPL. While the algorithm allows an offset term ($i = 0$), the $c_0$s are all set to zero as dark subtraction is made in the earlier computational step. No significant improvement was observed after cubic correction, so the $c_4$s and $c_5$s are also set to zero. The determination of the gain_preflight_samp coefficients is described in detail in Rosenberg et al. [2015]. These values are being updated to minimize footprint-to-footprint variations. This is done using data from "streak flats," obtained by orienting the long axis of the slit parallel to the orbit track, such that all footprints record the same scene.

An example of the application of this equation (for the strong $CO_2$ band, column 500, footprint 3) is:

$$Radiance = 0 + 2.898{\times}10^{15} * dn + 1.902{\times}10^{9} * dn^2 + 9.559{\times}10^{3} * dn^3 + 0 * dn^4 + 0 * dn^5$$

# 3    Signal-to-Noise Ratio (SNR)

## 3.0    Calculation of SNR

The noise values are not stored directly in the L1B HDF5s, but they can be calculated using a few fields in the L1BSc data file and the following formulas for the noise equivalent radiance and consequently the SNR:

$$NEN = \frac{MaxMS}{100} \cdot \sqrt{\left| \frac{100 \cdot N}{MaxMS} \right| \cdot C_{photon}^2 + C_{background}^2}$$

$$SNR = \sqrt{\frac{100\,N^2}{MaxMS * (C_{background}^2 \frac{MaxMS}{100} + C_{photon}^2 N)}}$$

where

- $N$ —radiance value
- $MaxMS$ —maximum measurable signal per band (see Table 3-1 below)
- $C_{photon}$ —first coefficient of L1BSc/InstrumentHeader/snr_coef ([0, *, *, *])
- $C_{background}$ —second coefficient of L1BSc/InstrumentHeader/snr_coef ([1, *, *, *])

The user should be aware that the $C_{photon}$ term captures the noise for an uniform scene – any additional varience that comes from a non-uniform scene causing changes in the instrument response based on pixel-to-pixel variabililty in either quantum efficiency or linearity is not included.  Our experience has been that this approximation does not lead to significant errors in the Level 2 retrival of XCO2, but users investigating other questions should confirm that this approximation is valid for their application. Figure 3-1 shows an example set of coefficients.

**Table 3-1.** *MaxMS* values.

| Band | *MaxMS* value (photons/m$^2$/sr/µm) |
|---|---|
| O$_2$ A-band | $7.00*10^{20}$ |
| Weak CO$_2$ | $2.45*10^{20}$ |
| Strong CO$_2$ | $1.25*10^{20}$ |
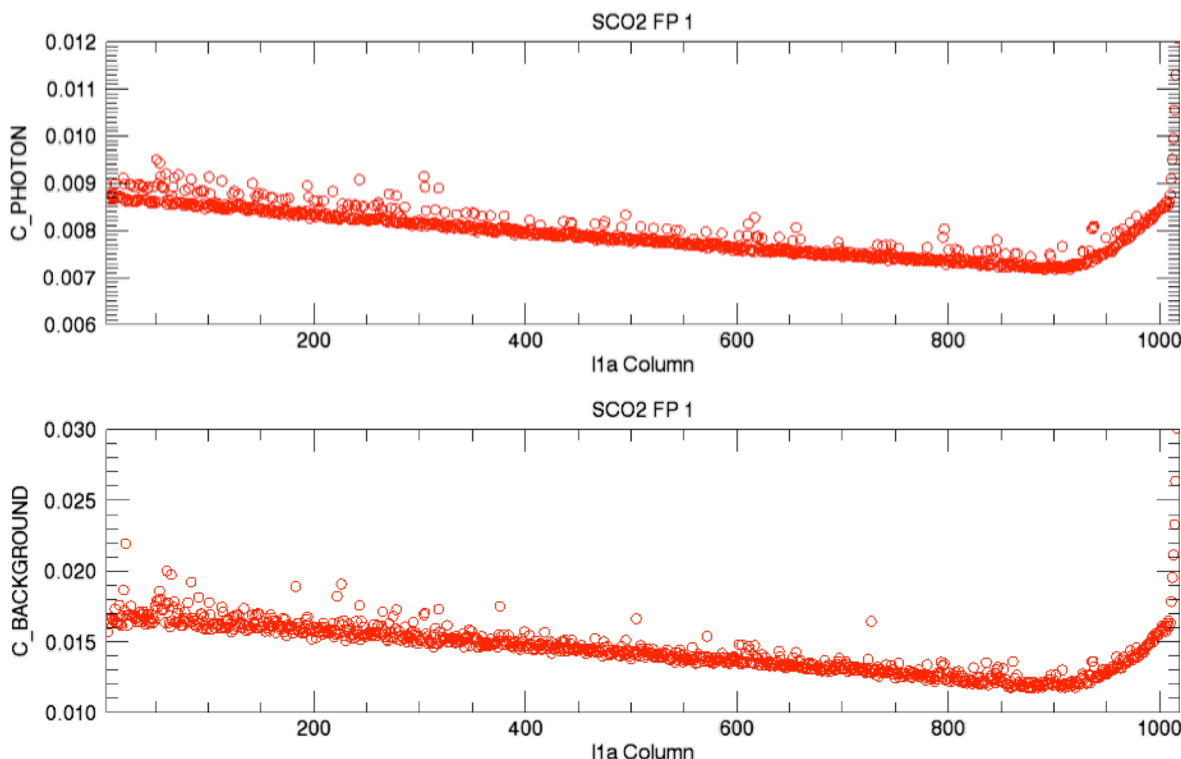
**Figure 3-1.** Example of signal-to-noise coefficients.

## 3.1 Development of SNR Coefficients

The initial snr_coef are based on a combination of the data from the best five radiometric calibration tests during final instrument testing at JPL. While this model assumes a perfectly linear detector, which is not the case, it still fits the data reasonably well. The only significant errors observed are at extremely low signal levels, where this SNR expression overestimates the radiance uncertainty because we do not currently include a term for uncertainty due to dark subtraction but its presence is unlikely to significantly change the radiance uncertainty estimates.

## 3.2 In-Flight Updates of SNR Coefficients

The dark noise (background term) will be updated based on the dark calibration data, measured directly during dark calibration periods (~12 two-minute dark calibration data collections each day).

Photon noise is assessed using solar calibration and lamp calibration data (our only consistent radiance scenes). We assume that the photon noise changes if the gain_degrad_samp term changes and adjust accordingly.

## 3.3 Bad Sample Mask

The third entry of L1BSc/InstrumentHeader/snr_coef (zero-based indices [2, *, *, *]) is used to identify bad samples that should be excluded by the retrieval algorithms. The third entry can range between 0 and 15 and is the sum of the following byte codes:
- 0 = Good

- 1 = Radiometric Problem (e.g., jumping pixel that needs to be identified for future bad pixel map, linearity problems)
- 2 = Spatial Problem (e.g., low wavelength 100 columns)
- 4 = Spectral Problem (e.g., ILS not trustworthy - last couple of samples on the $CO_2$ bands where the mask shadows the FPA)
- 8 = Polarization Problem (no examples yet)

Note that the fixes in the process of updating the gain coffeicents for bad pixel map changes (see Section 1.0) has led to a significant reduction in the number of bad samples in the Build 7 data products.

# 4    Spectral Response

## 4.0    Spectral Dispersion Coefficients

The dispersion coefficients express the relationship between the spectral element index (an individual FPA column or spectral sample) and its associated wavelength (see Figure 4-1). These data are contained in L1BSc/InstrumentHeader/dispersion_coef_samp. Note that this quantity does not account for Doppler shifts or dispersion adjustments that are applied in the L2 retrieval stage. The coefficients are used as follows:

$$\lambda = \sum_{i=0}^{5} c_i \cdot \text{column}^i.$$

Here, column refers to the column number in the L1BSc files (1 to 1016), which is raised to the power, $i$.

An example calculation of the wavelength grid is:

$$\lambda = 0.757633 + 1.75265 \times 10^{-5} \cdot \text{column}^1$$
$$- 2.91788 \times 10^{-9} \cdot \text{column}^2 + 3.29430 \times 10^{-13} \cdot \text{column}^3$$
$$- 2.72386 \times 10^{-16} \cdot \text{column}^4 + 7.66707 \times 10^{-20} \cdot \text{column}^5$$

Our experience has been that the wavelengths calculated by this process are very good for nadir
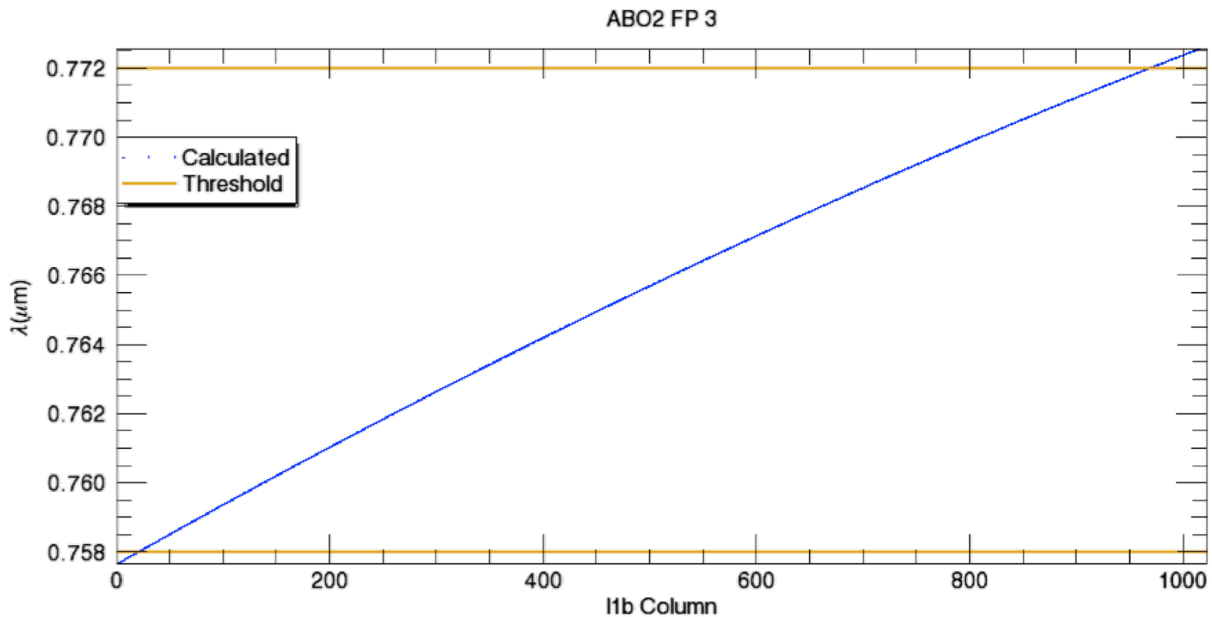


**Figure 4-1.** An example of the conversion of the focal plane column to wavelength scale. The blue line shows the column number-to-wavelength transformation. The yellow lines indicate the required wavelength range for this band.

data, but there appears to be a small change in the offset term for glint data – probably due to a tiny change in the thermal gradiant on the spectrometer. The impact is most pronounced in the Weak CO2 channel but is still only about 5% of a pixel registration error. Our current file format has no means to capture this change. We recommend that users who care about getting better than six significant figures of accuracy in the wavelengths allow their analysis to fit out small shift and stretch terms in each spectra (in addition to the corrections required to account for the Doppler shift between the spacecraft and atmosphere). Attempts to better capture this will be explored in future data releases.

## 4.1   Instrument Line Shape

The central challenge of the OCO-2 spectral calibration is determining not just a single ILS, but rather the ILS for every spectral pixel index, footprint, and band. The three bands, eight footprints per band, and 1016 spectral samples per footprint yield in theory 24,384 individual ILS functions. This is in contrast to, e.g., the Thermal and Near Infrared Sensor for Carbon Observations-Fourier transform spectrometer (FTS) aboard the Greenhouse Gases Observing Satellite (GOSAT), which in theory has just two ILS functions (one per polarization), as described by Yoshida et al. [2012]. However, the physics of the OCO-2 instrument design dictates that the ILS and centroid wavelength response (dispersion) of OCO-2 should vary smoothly in the spectral dimension across each band. The details of the measurements and analysis for determining the ILS are reported in Lee et al. [2015].

In the data product, for each band, footprint, and spectral element (3 x 8 x 1016 pixels), there are two 200-element lookup tables: L1BSc/InstrumentHeader/ils_delta_lambda and L1BSc/InstrumentHeader/ils_relative_response. These curves describe the response of each spectral element of the instrument versus wavelength, and can be used to convolve high spectral resolution spectra for comparison with OCO-2 spectra.

Initial determination of the ILS was performed using tunable diode lasers that were stepped through a range of wavelengths covering the OCO-2 spectral range. The final ILS was then optimized and validated by comparing solar spectra recorded simultaneously by OCO-2 and an FTS, as described by O'Dell et al. [2011]. As reported in Lee et al. [2015], the resulting ILS profiles (Figure 4-2) showed agreement between the two spectra to approximately 0.2% rms, satisfying the preflight calibration requirement of < 0.25% rms.

Post launch spectral residuals in the retrievals have been carefully analyzed for evidence of ILS errors, using on-orbit data. The solar calibration observations have provided an opportunity to identify and trend relative changes in the ILS. However, since the flight diffuser may make subtle changes to the ILS, an absolute measurement may not be possible. As of June 2015, on-orbit examination of the spectral residuals indicates a stable ILS.

## 4.2    Treatment of Polarization

As noted above, the OCO-2 instrument measures a single linear component of the polarization of the incoming light. The axis of polarization accepted by the instrument depends on the yaw angle of the satellite and other instrument parameters. For convenience, this is given in the L1B file via two sets of parameters. The first is the polarization angle of the axis of the accepted polarization, and the second is a set of "Stokes Coefficients" $\{m_1, m_2, m_3, m_4\}$ which define the response of the instrument to Stokes vector of the incoming light:

$$I_{meas} = m_1 I + m_2 Q + m_3 U + m_4 V$$

The polarization angle $\phi_{pol}$ is defined such that $m_2 = \frac{1}{2} \cos(2 \phi_{pol})$ and $m_3 = \frac{1}{2} \sin(2 \phi_{pol})$, and physically represents the angle between the axis of accepted polarization of the instrument and the reference plane for polarization. Stokes parameters $Q$ and $U$ also depend upon the chosen reference plane for polarization. For OCO-2, we choose the local meridian plane to be this reference plane for polarization, which is the plane containing the local normal unit vector and the vector pointing from the target field-of-view (FOV) to the satellite (Figure 4-3). Stokes parameter $Q$ is then defined as:

$$Q = I_{||} - I_{\perp}$$

i.e., $Q$ is the difference in intensities of the light polarized parallel and perpendicular to this
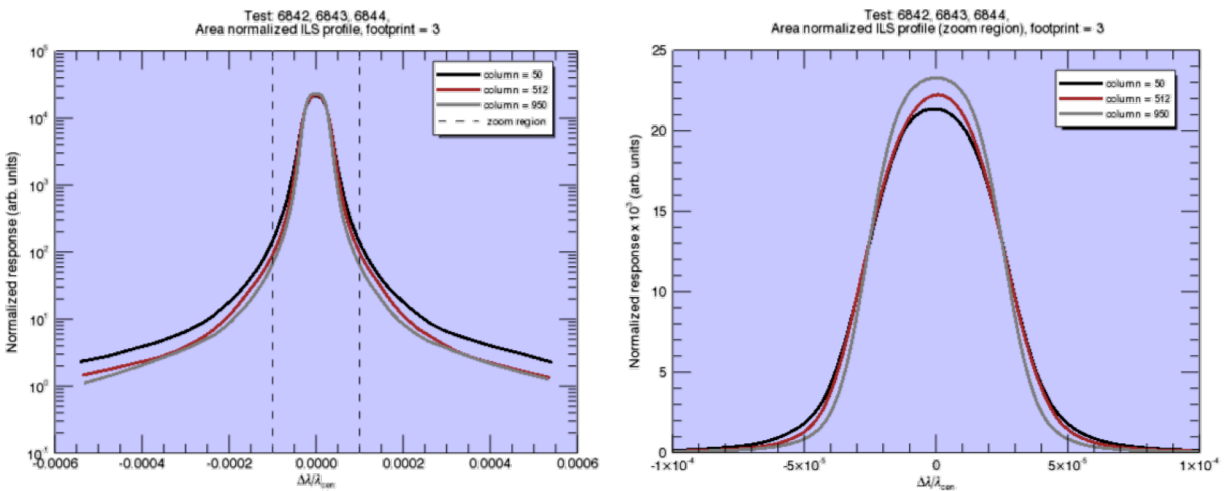


**Figure 4-2.** Example ILS profiles for the $O_2$ A band at three different spectral pixel indexes. **(Left)** Semi-log plot of the ILS. **(Right)** Linear plot of the ILS core, demarcated by the dashed vertical lines in the left plot.
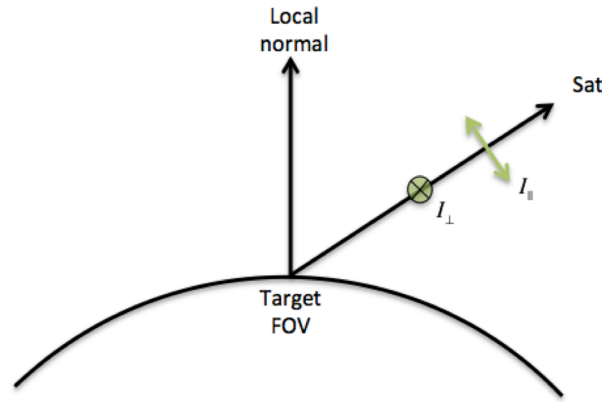
**Figure 4-3.** Geometry of the reference plane for polarization, i.e., the local meridian plane. This plane is formed by the local normal unit vector and the ray from the target FOV to the satellite. The direction of propagation of measured light is the latter. Two example polarizations are shown depicting light polarized both parallel (‖) and perpendicular (⊥) to this plane. The polarization angle of the instrument is defined as the angle between the axis of accepted polarization and the reference plane for polarization.

plane. $U$ is similar but requires a rotation by 45°:

$$U = I_{45} - I_{135}$$

The definition of Stokes parameter $U$ requires a definition of the "sense of rotation." Excerpting from Schutgens et al. [2004]:

> "The sense of rotation is usually defined by demanding that the three unit vectors ⊥, ‖, and prop form a right-hand coordinate frame, see Figure 2. Chandrasekhar [1960], van de Hulst [1981], and Hovenier and de Haan [1985] assumed ⊥ × ‖ = prop, while Azzam and Bashara [1987] and Slijkhuis [2000] assumed ‖ × ⊥ = prop. The difference in Stokes vector is the sign of $U$ (and hence χ 180° −χ). Note that POLDER data [Bréon and CNES Project Team, 1997] uses the latter convention, while SCIAMACHY data [Slijkhuis, 1998] currently uses the first."

Here, "prop" denotes a unit vector pointing in the direction of propagation. For OCO-2, we use the former convention of Chandrasekhar [1960], van de Hulst [1981], and Hovenier and de Haan [1985], also adopted by the SCIAMACHY team [Slijkhuis, 1998], i.e., we assume that ⊥ × ‖ = prop. This convention is used both in the definition of the polarization angle and Stokes coefficients provided in the OCO-2 L1B files, as well as the L2 radiative transfer solvers.

Even in nadir mode, the target field of view never sits directly at the subsatellite point. It is always offset slightly, and there is a different geometry for each of the eight OCO-2 footprints in a given frame. Therefore the polarization reference plane, defined above, can be significantly different for each of the eight footprints in a given frame, and therefore $m_Q$ and $m_U$ may vary significantly from one footprint to the next in a given frame.

In previous documents, we indicated that for OCO-2 the polarization reference plane would be defined as the principal plane, which is formed by the ray from the target to the Sun and the target to the satellite. This frame, while simpler in the sense that the definitions of $Q$ and $U$ would not vary much in nadir mode from one footprint to the next, is typically not used in the RT community. Therefore, we have avoided it in favor of the local meridian plane. In glint mode, the

principal and local meridian planes are nearly identical, but in nadir and target modes they can differ significantly.

Further, the orientation of the axis of accepted polarization is not always horizontal to the ground (i.e., perpendicular to the local meridian plane). While this was indeed the intent for OCO-2, a problem with the actual polarization plane of the instrument was discovered on orbit. A mitigation strategy was employed to yaw the spacecraft by 30° in glint mode, but not in nadir mode (as of the time of this writing). In target mode, because of the complicated slewing of the spacecraft, the polarization angle can vary widely.

In the special case of glint measurements in the principal plane, when $\phi_{pol} = -30°$, the measured intensity reduces to:

$$I'_{\parallel} = 0.5\ I + 0.25\ Q - 0.433\ U$$

However, in nadir and target modes, the polarization angle may vary with the footprint index from near 0° to near 180°.

Finally, we note that in the L1B software, identical target locations are assumed for all three spectral bands, and for each channel in a given band. Therefore, the polarization angle and Stokes coefficients take a single set of values for a given sounding.

## 4.3   Clocking Correction

As noted in Section 1.0, in this instrument design, the spectrometer slits, the grooves on the diffraction gratings, and columns of the FPAs must be well aligned to ensure that a fixed series of rows on the FPA will sample the same spatial footprint throughout the spectral range recorded by each FPA. For the OCO-2 instrument, perfect alignment of the FPAs was not possible due to a physical obstruction discovered late in the instrument assembly process. The focal plane arrays are therefore slightly rotated (or "clocked") with respect to the slit and grating (Figure 4-4). Consequently, a given geographic position does not map onto a single row of pixels across the entire spectral range sampled by the FPA. Instead, it changes from row to (roughly linearly) with spectral position (i.e., column). This is not an issue in spatially homogeneous regions, but can introduce spectral radiance discontinuities for scenes with substantial variations in illumination across a given spatial sample. Identifying and correcting these discontinuities is complicated, somewhat, but the fact that instrument sums ~20 spatial pixels into each spectral sample that is transmitted to the ground. For example, in lunar calibration data, where radiances can decrease from a maximum value to the noise level over a span of four or five pixels and where there are no obscuring absorption features, discontinuities are clearly discernable above the noise level in nearly every measurement. Methods for identifying and correcting these discontinuities are described in the following section.

Because fields of view must ultimately be mapped onto discrete pixels, the start and end positions of each footprint are adjusted in single pixel increments (i.e., clocking jumps) in row space at selected columns (Figure 4-5). These jumps can introduce discontinuities in the mean measured radiances for a spatial footprint if there are sharp gradients in the illumination near its boundaries. To compensate for these discontinuities, we have developed a simple clocking correction algorithm that enforces continuity in radiance across single pixel footprint shifts. The principal idea behind the algorithm is that although we do not have access to the full matrix of radiances over all rows and columns, we can use the entire row (i.e., spatial) information of a subset of the twenty color slices as a proxy for the spatial variation of intermediate columns, particularly those columns where clocking jumps occur. More specifically, at a small number of color slice positions we compute what the magnitude of the discontinuity in radiance would be if the slice position coincided with a particular clocking jump. We then make the assumption that this discontinuity is a valid approximation over some spectral neighborhood of the color slice and that the results from all the color slices can then be smoothly joined in some way.

The selection of which columns comprise the set of color slices can be changed by ground commands. One could place color slices adjacent to the columns with the clocking jumps. However, this is problematic for this application because some of these columns include sharp atmospheric or solar absorption features, and these features can be shifted by 2–3 columns by Doppler shifts associated with the spacecraft motion. If a color slice is placed in a region where the radiance changes rapidly (e.g., near an absorption line), the variation of radiance along a given row may be more indicative of the local spectral structure than of the spatial features of the measurement. Therefore, it is very important to place color slices where spectral variation is limited, i.e., in continuum regions. The availability of pixels in clear continuum regions varies
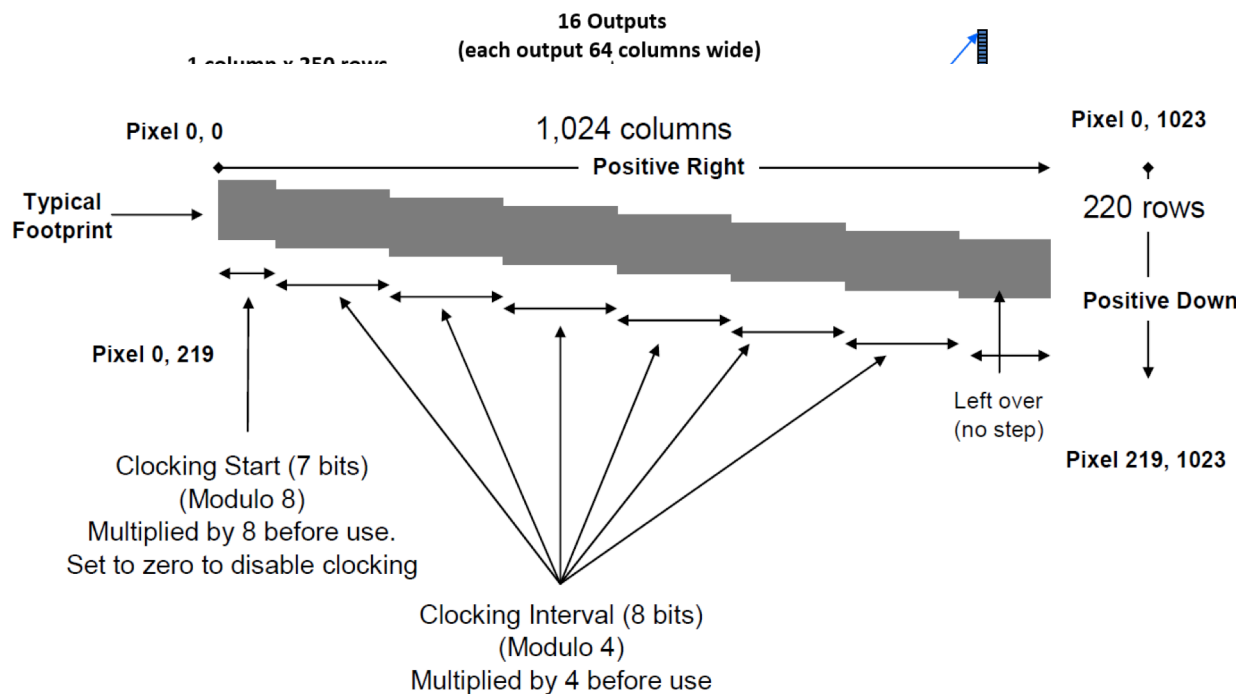


**Figure 4-5.** The "clocking" scheme used to sample the same spatial information (vertical dimension) across the entire spectral range (horizontal) for a rotated FPA.

substantially from band to band. For the $O_2$ A-band ($O_2$ A), continuum regions were selected just beyond the measurable high J value lines of the P and R branches as well as between the two branches. A similar approach was adopted for the weak $CO_2$ band near 1.61 μm (WCO2). In contrast, the deeply saturated strong $CO_2$ band near 2.06 μm (SCO2) offers no spectral regions as transparent as in the other two bands, and none of the color slices used by the clocking algorithm lies in between the P and R branches.

Figure 4-6 through Figure 4-11 illustrate the nature of the clocking error and the choices of color slices for each of the three OCO-2 bands. For each band, there is a large figure which contains three different graphic elements. First are the nearly horizontal solid lines which delineate the boundaries of the eight footprints. The discrete jumps in the footprint boundaries illustrate the clocking effect. Note that the magnitude of the clocking effect varies among the bands, ranging from 8 discontinuities in the most severe case (SCO2) to 2 in the least (WCO2). Second, the dashed vertical lines indicate the current selection of all twenty color slices (as of October 2014). *Only those slices that are labeled by an index are used by the clocking algorithm.* Finally, a representative spectrum is superimposed to provide some context for the color slice selection.

Generally, the color slices used by the clocking algorithm are located in spectral regions with little absorption. Not all twenty color slices are used for clocking; other color slices have been located within strong absorbtion features to assess the spatial inhomogeneity due to clouds and aerosols or to monitor solar lines. In the $O_2$ A-band, slices 16, 17, and 18 lie in the neighborhood of a solar line and were mistakenly included in the clocking correction. However, it was found that corrections using these slices correlate very strongly with those using slices 15 and 19, which do lie in the continuum, so that the inclusion of these color slices in the clocking



**Figure 4-6.** Schematic illustrating the clocking for the $O_2$ A band along with the current color slice selection. Horizontal, piecewise-constant solid curves delineate footprint boundaries. Dashed vertical lines indicate color slice selection. Only those that are numbered are used for clocking. Finally, a typical spectrum of the band is overlaid for reference.

correction algorithm is unlikely to have significant impact. Nevertheless, future versions of the clocking algorithm may remove these slices from consideration. Slices 3, 4, 5, and 15 from the SCO2 band may in the future be eliminated or replaced, since they lie close to weak absorption lines. Further testing is required to determine what difference their omission would make.

**Figure 4-7.** The color slices used in the clocking correction for the O$_2$ A band (shown in Figure 4-6) are aggregated into five "groups." Each panel depicts a single group.



**Figure 4-8.** Schematic illustrating the clocking for the WCO2 band along with the current color slice selection. Piecewise-constant solid curves delineate footprint boundaries. Dashed vertical lines indicate color slice selection. Only those that are numbered are used for clocking. Finally, a typical spectrum of the band is overlaid for reference.

**Figure 4-9.** The color slices used in the clocking correction for the WCO2 band (shown in Figure 4-8) are aggregated into four "groups." Each panel depicts a single group.
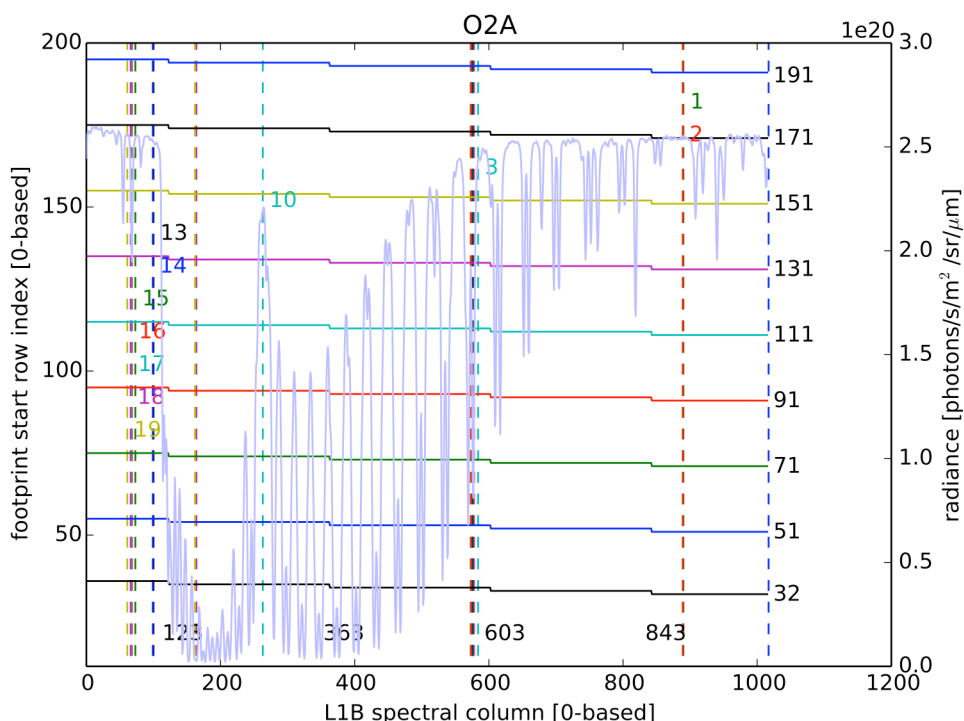


**Figure 4-10.** Schematic illustrating the clocking for the SCO2 band along with the current color slice selection. Piecewise-constant solid curves delineate footprint boundaries. Dashed vertical lines indicate color slice selection. Only those that are numbered are used for clocking. Finally, a typical spectrum of the band is overlaid for reference.
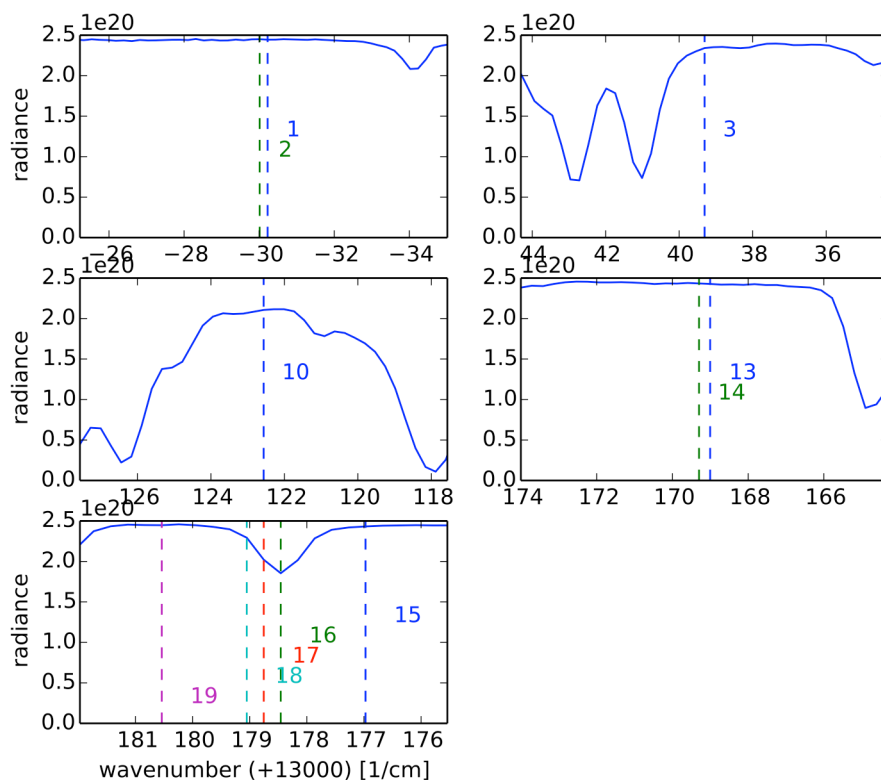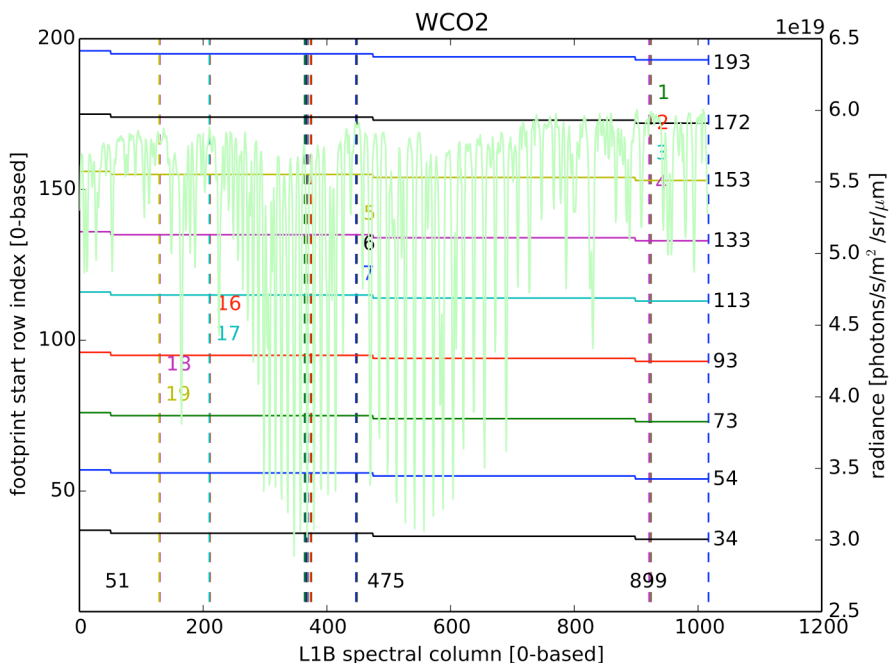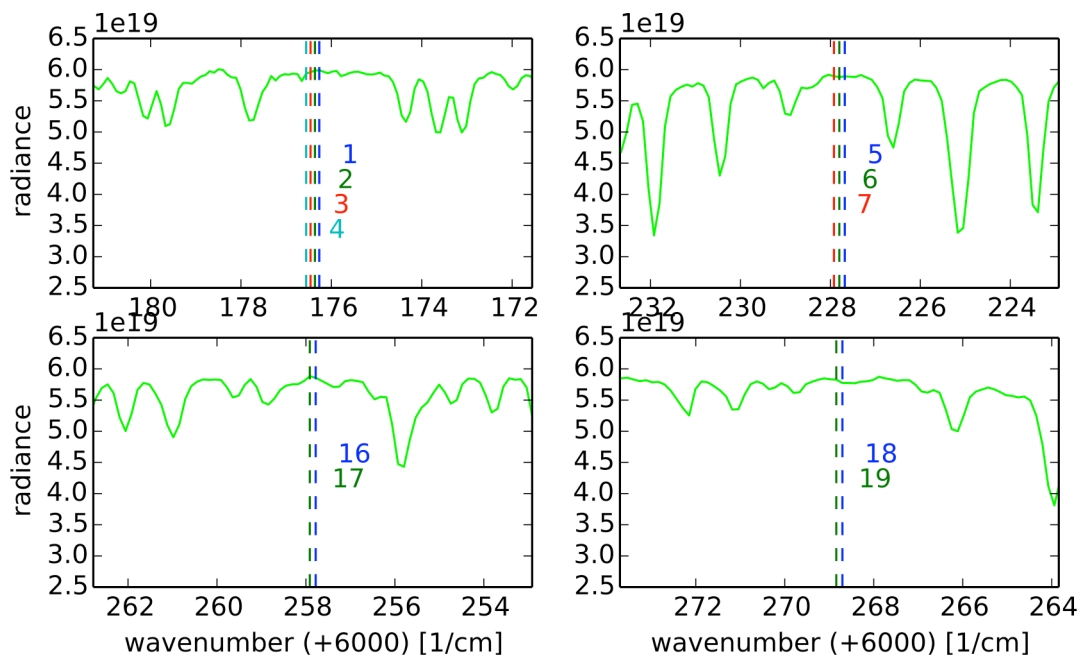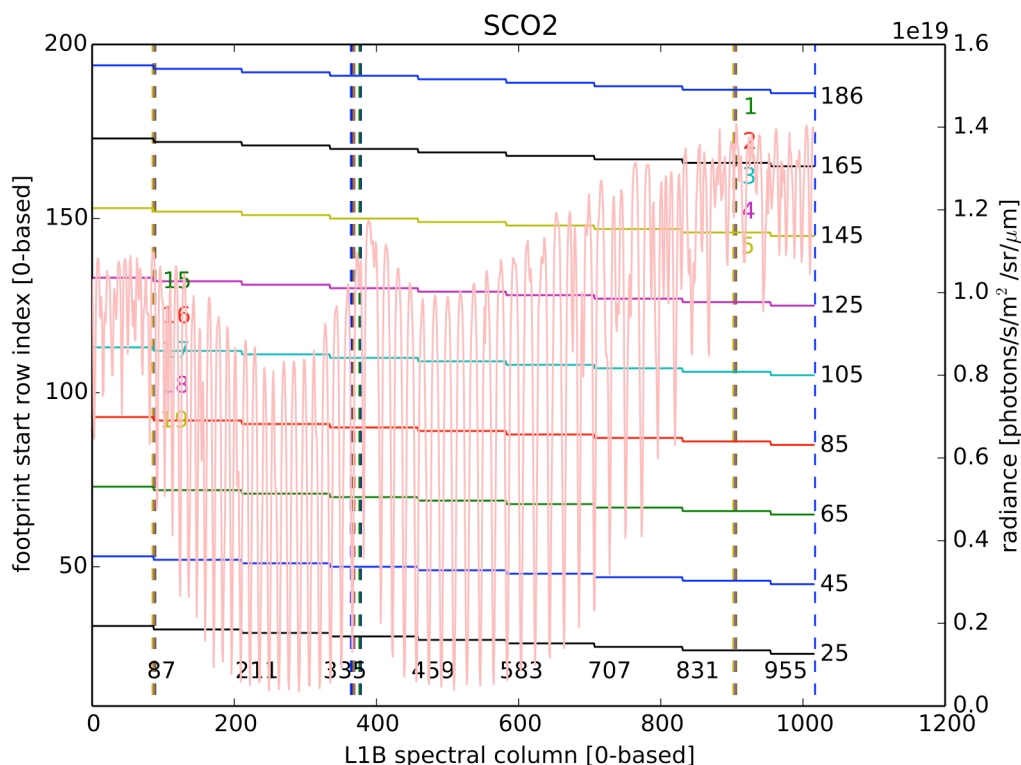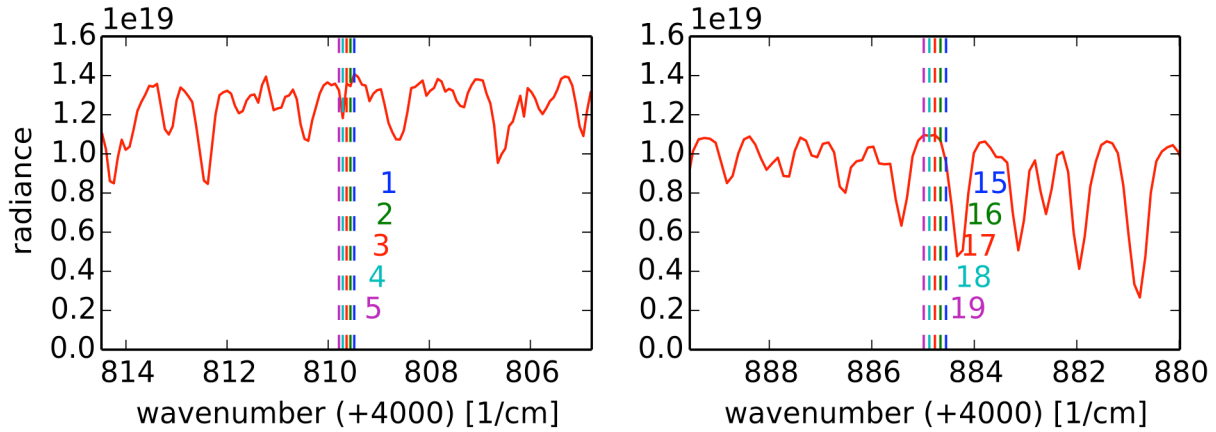
**Figure 4-11.** The color slices used in the clocking algorithm for the SCO2 band (shown in Figure 4-10) are aggregated into two "groups." Each panel depicts a single group.

### 4.3.1 Clocking Algorithm Description

#### 4.3.1.1 First Stage: Calculating Average Radiances

The algorithm used to compute modified L1B radiances proceeds in two stages. In the first stage, for each footprint-color slice pair, $(f, c)$, average radiances over all possible footprint start and end positions are computed and saved to a file:

$$\bar{R}(f,c,j) = \frac{1}{M_{f,c}} \sum_{r \in FP(f)+j} R(r,c)$$

In this expression, $R(r, c)$ is the radiance as a function of row $r$ and color slice $c$, and $M_{f,c}$ corresponds to the number of rows contained in the footprint-color slice pair. The sum is over all rows r contained in footprint f (as defined in the last column) but shifted by j pixels. In other words, if in the last column, a footprint is comprised of pixels {p, …, p+19}, sums are performed over pixels {p+j, …, p+19+j}, for all j from 0 to the number of clocking jumps contained in the band. Thus, these mean radiances represent averages over all the possible sets of pixels that may comprise a footprint.

Anomalous (e.g., negative) values must be eliminated so that they do not skew the averages, since they would lead to potentially large errors in the corrections. Therefore, a heuristic filtering is performed in which "bad" pixels are replaced with an average of nearest neighbors (in row space). Pixels are marked "bad" if they satisfy at least one of three conditions: (1) they are marked as "bad" in the ARP, (2) they have a negative value, or (3) the relative spatial gradient exceeds some threshold and the absolute spatial variation is more than twice the noise level σ, defined by:

$$\sigma = \frac{R_{max}}{100} \sqrt{\frac{100 R(\lambda)}{R_{max}} \eta_0^2 + \eta_1^2}$$

where $R(\lambda)$ is the radiance, $R_{max}$ is an estimate of the maximum expected signal, and $\eta_j$ are the signal-to-noise ratio (SNR) coefficients defined in the L1B file. The threshold is an adjustable

parameter and taken to be a relative increase of 1.6 (or decrease of 1/1.6) from one row to the next. *The current scheme fails if there are two or more consecutive bad pixels.* This almost never happens for the O$_2$ A-band, but does occur occasionally in the SCO2 band. Future implementations of the algorithm will take such pathological cases into account.

Table 4-1 provides a glossary of symbols used in the clocking algorithm description.

**Table 4-1.** Glossary of symbols.

| Symbol | Definition |
|---|---|
| $\bar{R}(f,c,j)$ | Mean radiance, over rows $\{r + j \mid r \epsilon f\}$ in footprint f and color slice c |
| $M_{f,c}$ | Number of pixels in color slice c, footprint f |
| $R(f,c)$ | Radiance at row r in color slice c |
| $r(f,G,j)$ | Ratio needed to correct clocking at jump j for footprint f and group G |
| $N_{f,G}$ | Normalization factor for footprint f and group G to preserve mean radiance over entire spectrum |
| $h_G(\lambda)$ | Weighting function for group G at column $\lambda$ |
| $C_G(f,\lambda)$ | Correction function for group G, footprint f at column $\lambda$ |
| $C(f,\lambda)$ | Final correction function applied to L1B radiance for footprint f, column $\lambda$ |

Figure 4-12 shows a large but not extreme example of the sort of spatial variation that might appear within a single frame. The grey, dotted vertical lines indicate footprint boundaries. Individual solid colored curves correspond to the *corrected* radiance within particular color slices, while all uncorrected radiance measurements are denoted by grey circles. Footprints 3 and 4 (counting from the left and starting from zero) show significant differences in both radiance and slope between their left and right boundaries. Such differences will produce large clocking corrections. Note also that there are three uncorrected points whose nearly vanishing values are replaced by averages of their neighbors. Without this correction the mean radiances within a footprint would be underestimated by ~5%. If the "bad" pixels occur near footprint boundaries and were to remain uncorrected, the clocking algorithm would produce erroneous corrections.
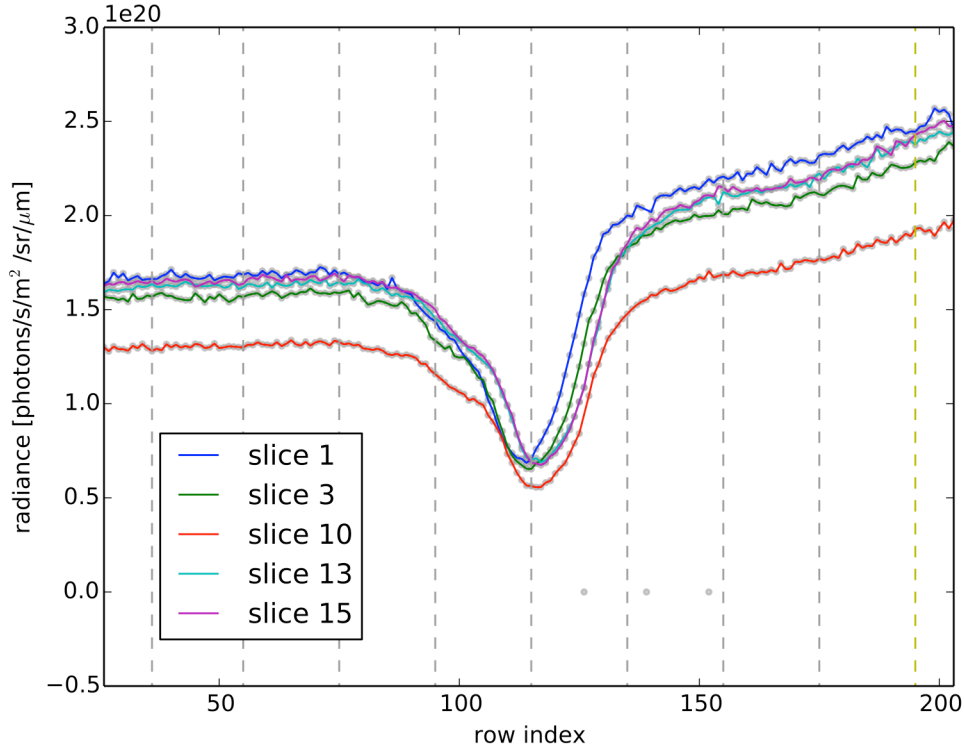
**Figure 4-12.** $O_2$ A band radiances measured for a few select color slices as a function of row index, which serves as a proxy for actual spatial variation within a scene. Solid curves are corrected values in which "bad" pixel data have been removed. Grey circles represent original measured data. For this particular case, three bad pixels have been replaced by nearest neighbor averages. Grey vertical lines represent footprint boundaries. The large differences between the values of the radiance on the boundaries of footprints 3 and 4 (starting from zero on the left) will produce large "clocking errors," or spectral discontinuities in the measurement.

### 4.3.1.2   Second Stage: Computing Ratios and Final Correction Function

After computing and saving the radiance means, the algorithm proceeds to the second stage in which column radiances are modified. First, the set of color slices is partitioned into "groups", which consist of clusters of spectrally close columns. The motivation for this partition is statistical. Using color slices that are spectrally very close should produce similar discontinuities; averaging over a small ensemble of strongly correlated slices reduces statistical error. The aggregation of color slices into groups follows a simple rule: two slices are in the same group if they are separated by fewer than $n$ pixels, where $n$ is taken to be 6 by default but is adjustable. (If $n$ were chosen to be 0, for example, the groups would all be singletons.) In the future this rule may be further refined, possibly by allowing for aggregations that are dependent on footprint. Illustrations of the current selection of color slices and their aggregation into groups for each band are shown in Figure 4-7, Figure 4-9, and Figure 4-11.

We now combine the mean radiances computed in the first stage into a group dependent quantity by summing the mean radiances over all color slices contained in a group. Then, for each group, the ratio of the mean radiance at each jump is computed:

$$r(f,G,j) = \frac{1}{N_{f,G}} \frac{\sum_{c \in G} \bar{R}(f,c,j)}{\sum_{c \in G} \bar{\bar{R}}(f,c,j+1)}$$

In this expression the summation is over all color slices, $c$, within a group, $G$, and $N_{f,g}$ is a normalization factor defined below. (All radiance variables depend also on frame index, which has been omitted for notational clarity.) Neglecting the normalization factor for the moment, these ratios define multiplicative factors that if applied to only one side of a jump, produce a continuous radiance across the jump.

We now need to interpolate (and possibly extrapolate) to all spectral points. The ratio described above specifies a multiplicative factor to be a applied at a jump position, p under the constraint that the multiplicative factor to be applied at position p+1 is identically equal to unity. To extend the multiplicative factor to all columns, we linearly interpolate from a value of one from the last clocking jump to the value given by the ratio as defined above. The resulting correction function, $C_G(f, \lambda)$, is a saw tooth defined for each group $G$ and at every spectral point $\lambda$. The left panel of Figure 4-13 shows example correction functions for a particular footprint for the $O_2$ A-band. Plotted are several curves of slightly different shade of blue, corresponding to correction functions using different groups. In this example, all ratios at all jumps were found to lie between 0.95 and 0.97, and for each group a correction function is defined by linearly interpolating between a jump location, where the function's value is the ratio and one pixel to the right of the previous jump location, where the function's value equals one.  To maintain roughly the same average radiance across the entire band, this correction function is normalized by a factor $N_{f,G}$ such that its mean over the entire spectral domain is equal to unity. The right panel of Figure 4-13 features red curves corresponding to the blue ones in the left panel, but after renormalization.

At this point, one has multiplicative correction functions for each color slice grouping. These
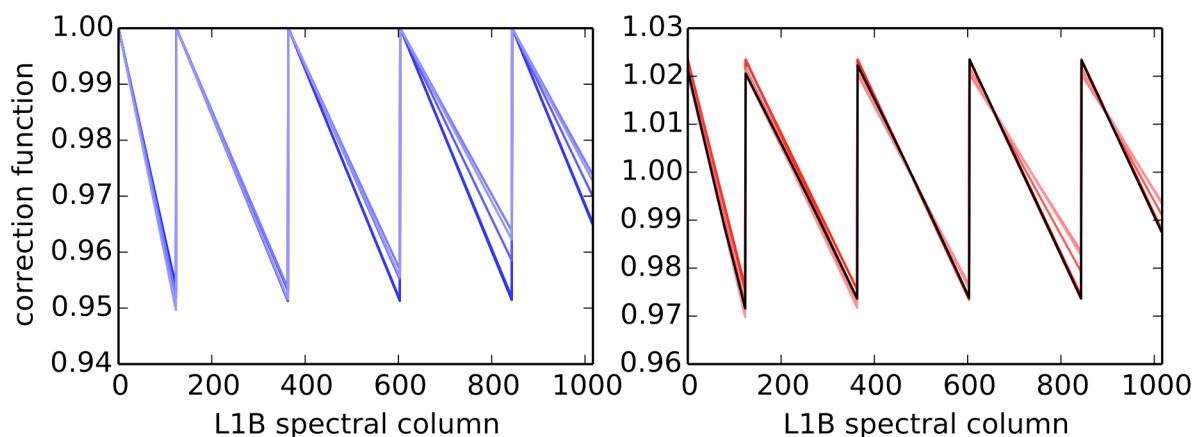


**Figure 4-13.** Example of a clocking correction function for the $O_2$ A band. The blue curves on the **left** correspond to different pre-normalized correction functions determined by different groups of color slices for a single frame and footprint. To (roughly) preserve the value of the average radiance, the correction functions are normalized such that the mean over columns equals one, as shown in the red curves on the right. Multiplication by weighting functions for each group yields the final correction function, depicted as the black curve on the **right**.

functions need to be combined to produce a single correction function, $C(f, \lambda)$, for each frame-footprint pair. Because the corrections are expected to be most accurate for columns that are spectrally close to a group, we employ simple hat functions to create a weighted average:

$$C(f, \lambda) = \sum_G h_G(\lambda) C_G(f, \lambda)$$

where $h_G(\lambda)$ are the triangular weighting functions, which fall linearly from a maximum value of one at the average column position of group $G$ to a value of zero at the average positions of the neighbors of $G$. If $G$ is the first or last group—that is, it does not have a neighbor to the left or to the right—then $h_G(\lambda)$ has a constant value of one from the average column position to the end of the detector. Note that the hat functions depend only on the choice of color slices and their grouping and have the property that:

$$\sum_G h_G(\lambda) = 1$$

The weighting functions for the current selection of color slices and groupings for the O$_2$ A-band is shown in Figure 4-14, where each color corresponds to a different group, $G$.

The sum of the product of these weighting functions and the curves in red in Figure 4-15 yields
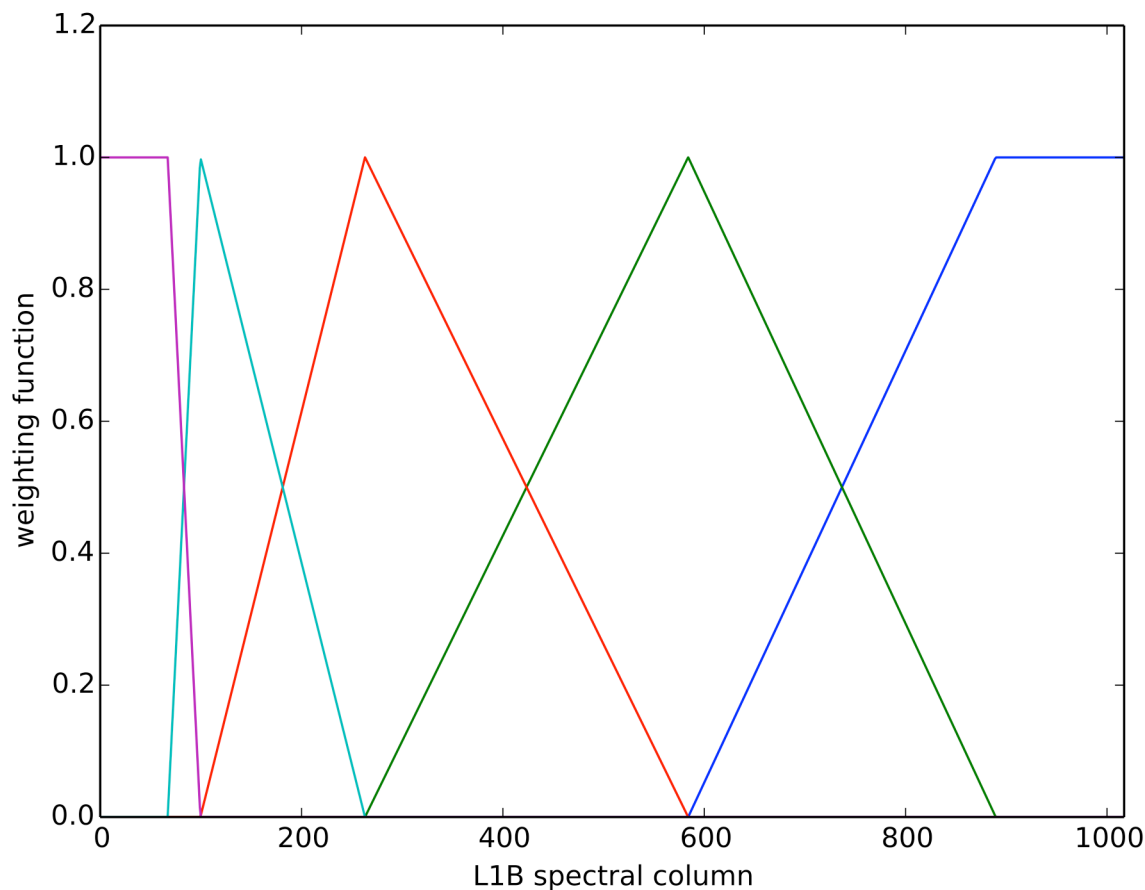


**Figure 4-14.** Weighting functions for each group of color slices for the O$_2$ A band.

the weighted average correction function, shown in black in Figure 4-15. This weighted correction is applied to the original L1B radiances to produce the corrected radiances:

$$R_{corrected}(f, \lambda) = R(f, \lambda)C(f, \lambda)$$

In practice, most corrections for non-lunar scenes are fairly small; the difference between the original and corrected radiances is barely perceptible, although discontinuities may be more evident in the residuals of retrievals. A typical "large" correction is on the order of several percent. To facilitate analysis of the effect of the clocking correction a simple scalar metric, $\mu(f)$, depending only on sounding is also saved to the L1B output:

$$\mu(f) = max_\lambda |C(f, \lambda) - 1|$$

Figure 4-15 shows the effect of applying the correction illustrated in the preceding figures to the measured radiance of the $O_2$ A-band. The black/red curves denote original/corrected radiances.
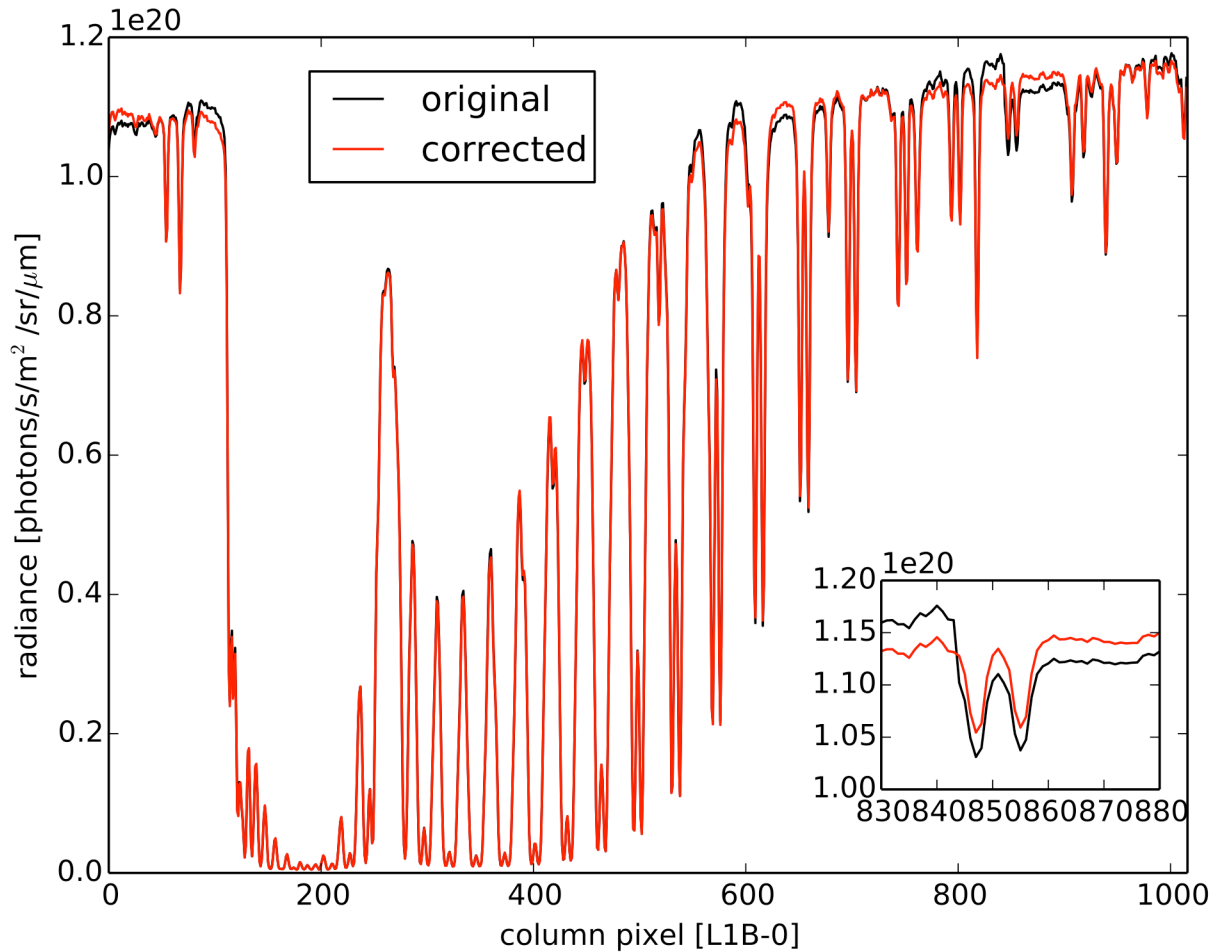


**Figure 4-15.** An example $O_2$ A spectrum, both before and after clocking correction. The inset shows a close up in the neighborhood of a jump, where the correction is largest.

Figure 4-15 shows that a clocking jump occurs at column 843. The inset illustrates the elimination of the discontinuity at that pixel to produce an even continuum region about the $O_2$ doublet.

## 4.3.2  Discussion

The behavior of the clocking correction varies significantly among the three bands. In this section, we restrict discussion to a comparison of the effect of the correction on a representative L1B product (orbit 1600); its impact on actual $X_{CO2}$ retrievals will not be considered here. Orbit 1600, shown in Figure 4-16 contains both land (33%) and ocean frames (66%).

Figure 4-17 shows the cumulative distribution of the magnitude of the correction (as defined above) for each band and terrain type (i.e., land and water). For example, for the $O_2$ A-band roughly 90% of all frame-footprint pairs over land require a correction of less than 0.5%. While the statistics for the $O_2$ A and WCO2 bands are similar over land, the SCO2 bands requires larger corrections: the $90^{th}$ percentile for the SCO2 occurs at a correction magnitude of 1%. The reason for this difference is unclear but may be related to the much larger number of "bad" pixels in the SCO2 band. We note that for each case shown in Figure 4-17 at least 40% of scenes require no correction at all. This result is an artifact of the noise cutoff threshold discussed in the previous section; corrections that are of the same order of magnitude as the noise are artificially set to zero.

One useful metric to assess the soundness of the correction of the algorithm is to examine the consistency of the predicted corrections among the color slices. Figure 4-18, Figure 4-19, and Figure 4-20 show for the $O_2$ A, WCO2, and SCO2 bands, respectively, the correlation coefficient, $\rho$, of the ratios $r$, defined in the previous section:

$$\rho(f, G, G') = \frac{\sum_{s,j} r(s, f, G, j) r(s, f, G', j)}{\left(\sum_{s,j} r^2(s, f, G, j)\right)^{1/2} \left(\sum_{s,j} r^2(s, f, G', j)\right)^{1/2}}$$

where the summation over a frame index variable $s$ has been made explicit and instead of the employing the grouping described previously, we have taken each group $G$ to contain a single color slice. It was shown previously that a significant fraction of corrections are either small or identically zero. Since large relative differences of very small quantities may not be meaningful, we impose a restriction to the summation in the equation above: for a given group (i.e., color slice) $G$ we sum over only those corrections that are larger than a cutoff of 0.5%. Note that because the set of summands for a group $G$ will not be identical to the set of summands for a group $G'$, the correlation coefficient defined above will not generally be a symmetric function of $G$ and $G'$.
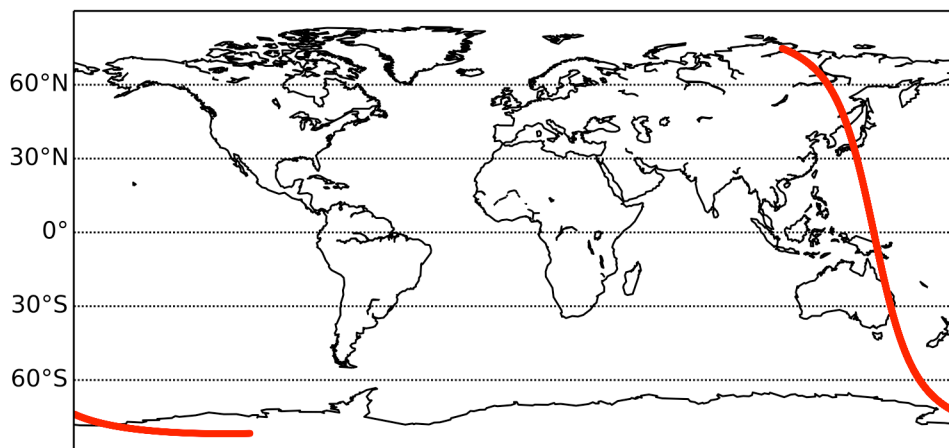
**Figure 4-16** Orbit used in discussion of the clocking correction.

Figure 4-18 shows the correlation matrix by footprint for the $O_2$ A-band. We first note that the correlations are rather strong in all cases, everywhere exceeding 0.87. Also, as expected, color slices that are spectrally close have stronger correction correlations. This finding justifies our aggregation of color slices into groups. However, note that there is a dependence on footprint: spectrally distant color slices show greater consistency for lower footprint indices than for higher indices. That is, for higher footprint indices a more diverse sampling of color slices is required for the correction algorithm. This result is most apparent for the $O_2$ A-band but is evident for the other bands as well. Finally, it was noted earlier that slices 16, 17, and 18 (corresponding to colors 68, 67, and 66) actually lie near solar lines and therefore ought not to have been included in the correction algorithm. It was asserted that their inclusion, nonetheless, would not significantly modify the final outcome. The strong correlation illustrated in Figure 4-18 between those color slices and all the other slices on the shortwave end justifies that assertion.

Figure 4-19 shows correlations between color slices for the WCO2 band. Generally, the correlation matrix elements are even stronger than for the O$_2$ A-band even for spectrally distant slices. This is likely due to the low degree of clocking in the WCO2 band, which suffers only three jumps over the whole band as compared to four in the O$_2$ A-band and eight for the SCO2 band. However, there are some oddly weaker correlations involving color 211 in footprints 5 and 6 and color 922 in footprints 0 and 1. We suspect these arise from a greater number of instances where consecutive "bad" pixels foil the correction algorithm.
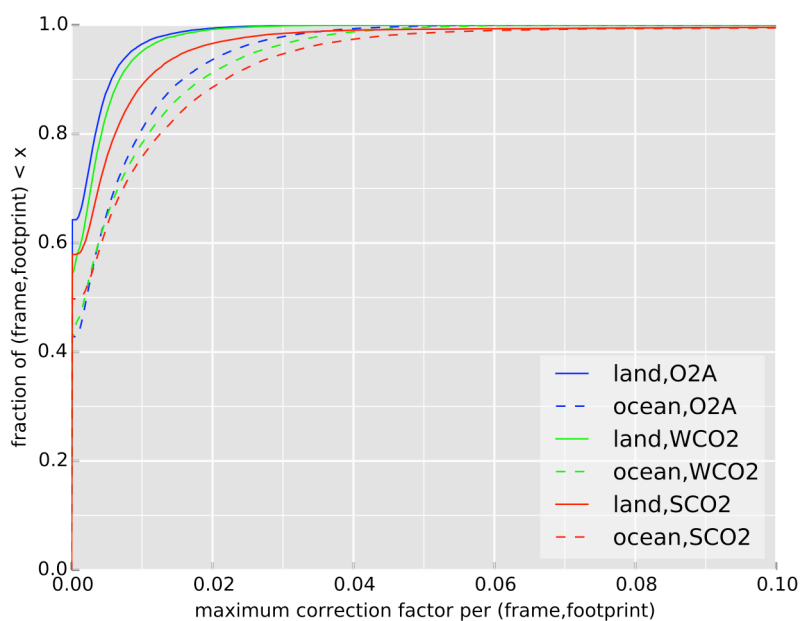


**Figure 4-17.** Cumulative distribution function of magnitude of the correction for orbit 1600. Results are separated by band and terrain type (i.e., land and water).
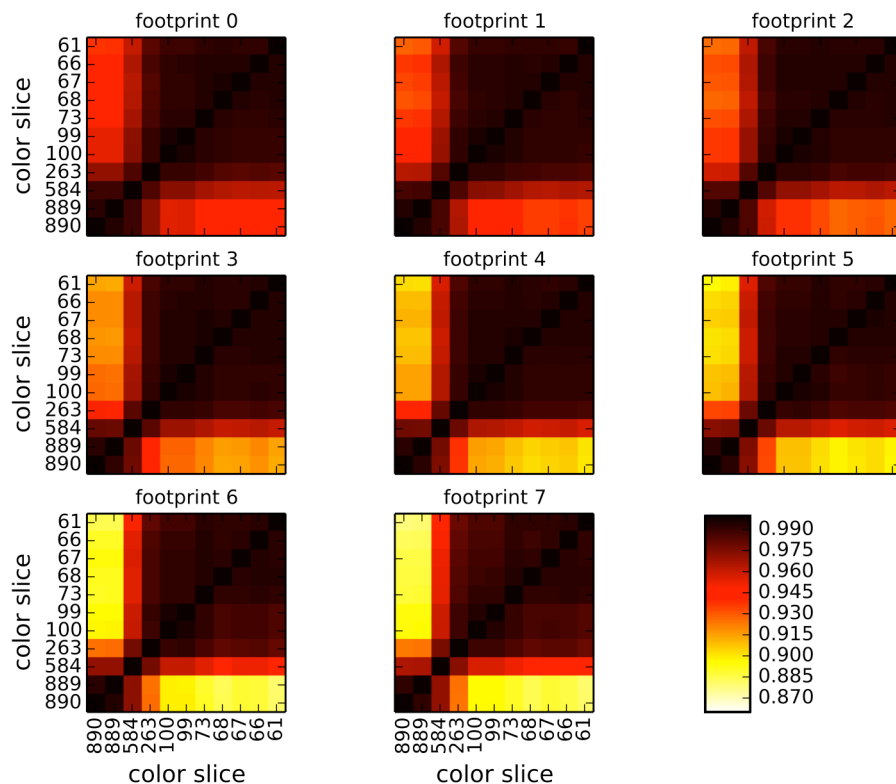
**Figure 4-18.** Correction correlation matrix for the $O_2$ A band.
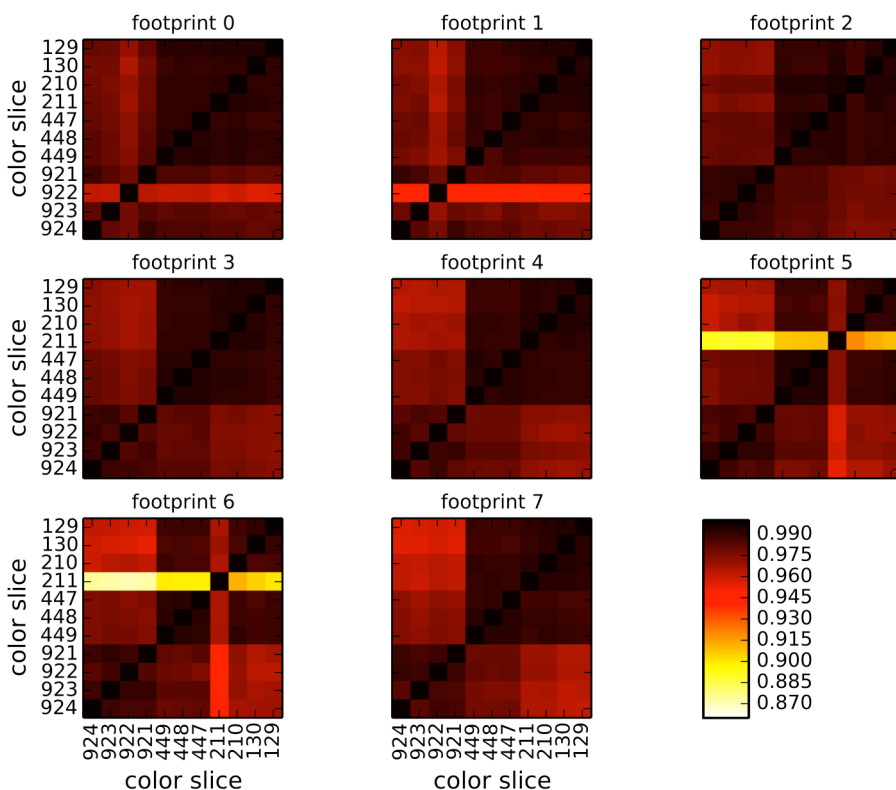


**Figure 4-19.** Correction correlation matrix for the WCO2 band.

Unlike the other two bands, the corrections prescribed by color slices in the SCO2 band are not generally very consistent, even for spectrally nearby slices as Figure 4-20 shows. We strongly suspect that this result stems from the much larger number consecutive "bad" pixels, which the algorithm fails to properly handle. Note that the group of color slices 1 – 5 (colors 89 through 85) that do generally produce consistent corrections lie on the shortwave end of the detector, where there are far fewer pixels that are marked as "bad" in the ARP. Furthermore, the large clocking effect in the SCO2 should weaken the correlation between the two groups on the long and short wavelength ends of the detector.

Finally, we consider the effect of the algorithm on the spectra themselves. One method to visualize the reduction of clocking-induced discontinuities is to perform a singular value decomposition on a set of spectra before and after correction and examine the resulting change in the right singular vectors:

$$S = U\Sigma V^*$$

In this expression the columns of S are comprised of the uncorrected spectra, $\Sigma$ is the diagonal matrix of singular values, and the columns of U and V are the left and right singular vectors. Figure 4-21 demonstrates the effect of the clocking algorithm on the set of spectra in orbit 1600. The first eight singular vectors are plotted both before (top row) and after (bottom row) correction. Discontinuities that are clearly visible in the sixth (yellow-green), seventh (orange), and eighth (red) singular vectors prior to correction are completely eliminated by the clocking algorithm. Because only a minority of spectra require significant correction, the size of the effect as measured by the ratio of singular values is fairly small; the ratio of sixth and first singular value is 0.0045 in this data set. The impact of the correction on the other bands is similar and shown in Figure 4-22 and Figure 4-23.

These results indicate that this correction approach will usually be adequate. In cases where it is not, it will produce a poor fit in the Level 2 retrieval, and the Level 2 post processing screening algorithm will usually reject the sounding.
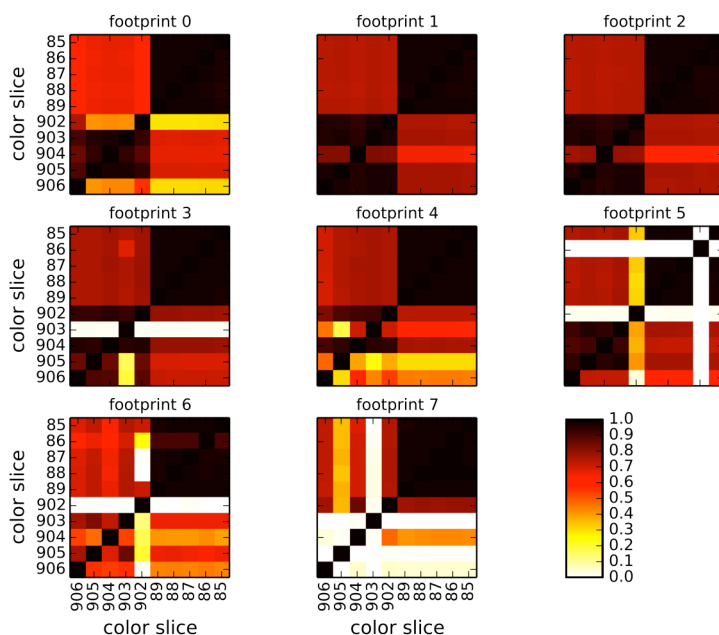


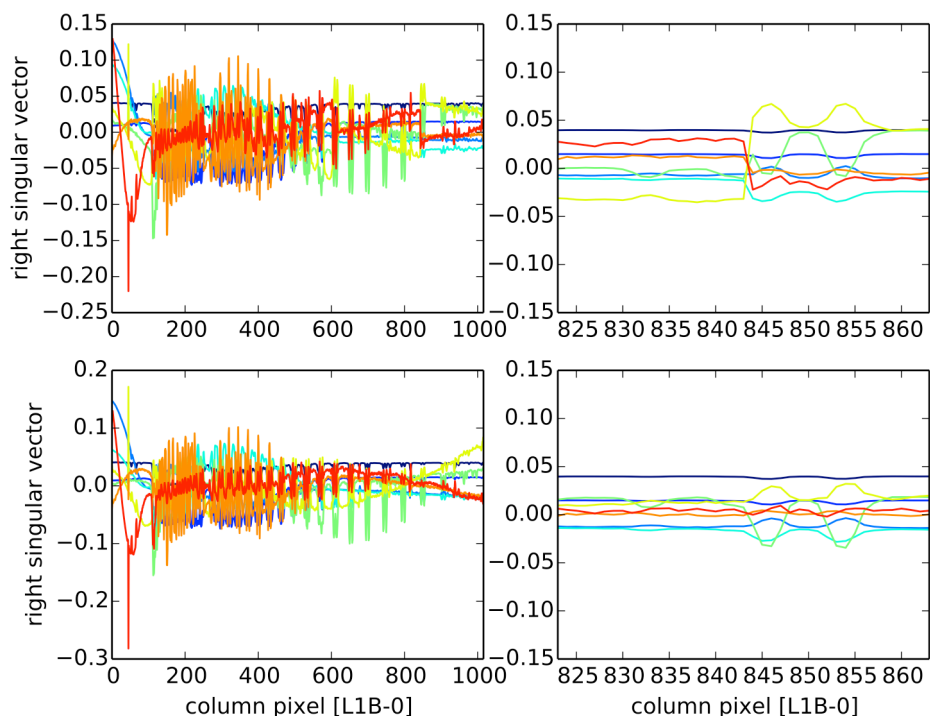**Figure 4-20.** Correction correlation matrix for the SCO2 band.

**Figure 4-21.** Right singular vectors corresponding to eight largest singular values from violet (largest) to red (smallest). The top/bottom rows are pre/post corrected vectors. The panels on the right are close up of clocking jumps at 843. The right singular vector corresponding to the sixth (yellow-green), seventh (orange), and eighth (red) largest singular values exhibit discontinuities at pixel 843 prior to correction.
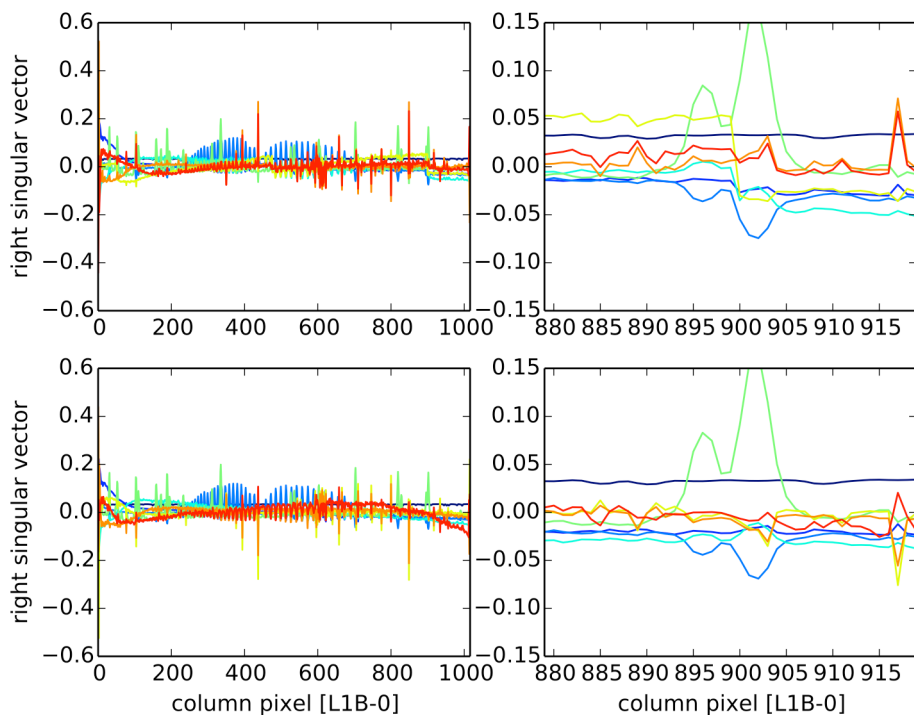


**Figure 4-22.** First eight right singular vectors for the WCO2 band. The clocking jump evident in the sixth one (in yellow-green) with singular value 0.28% of the largest value vanishes after the clocking correction.
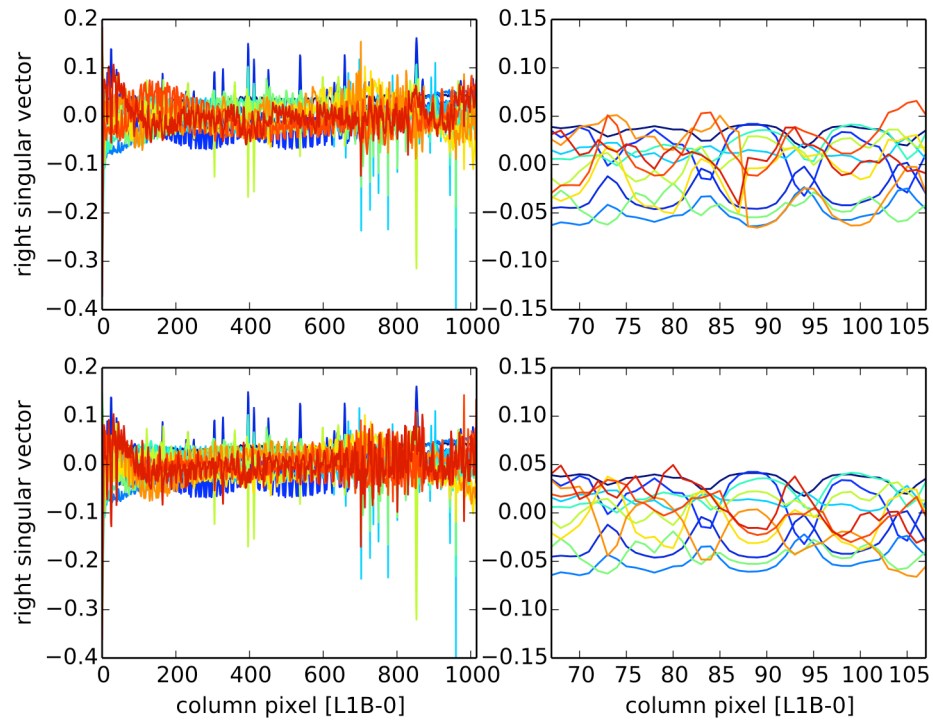
**Figure 4-23.** First twelve right singular vectors for the WCO2 band. The clocking jump evident in the eleventh one (in orange) with singular value 0.34% of the largest value vanishes after the clocking correction.

## 5    Ancillary Radiometric Product (ARP) Files

The key calibration parameters are stored in the ARP, which is not distributed publicly. These files are changed as needed and deemed necessary. There will be an ARP file for initial processing and an updated ARP for reprocessing to reflect any additional knowledge gained from the measured data set about dark subtraction or gain degradation. Each data product identifies the ARP used in Metadata/ARPAncillaryDatasetDescriptor.

# 6 Other Resources

There are a number of other project documents that the user should refer to as they work with the data.

4. Data User's Guide—this document discussed the publicly distributed data products, L1BSc, L2IDP, L2Std, and L2Dia. The key data fields are discussed, and tables are included that specify all of the fields in each data product

5. L2 ATBD—this ATBD steps through the physics and implementation of the Level 2 algorithm

6. ATBD for IMAP-DOAS and ABO2—these ATBDs describe the two methods of identifying potentially cloudy footprints, in what we refer to as the prescreening step. These data are then used for setting data quality and data selection levels

7. Published papers—there are a number of published papers describing the algorithm, application to GOSAT, prescreening steps, etc. Please see the most up-to-date list of publications on oco2.jpl.nasa.gov

# 7 References

Chandrasekhar, S., (1960), Radiative Transfer. Dover Publications, Dover, N.Y.

Crisp, D., C. E. Miller, P. L. DeCola, (2008), NASA Orbiting Carbon Observatory: measuring the column averaged carbon dioxide mole fraction from space, *J. Appl. Remote Sens., 2,* 023508; doi:10.1117/1.2898457.

Crisp, D. (2008), The Orbiting Carbon Observatory: NASA's First Dedicated Carbon Dioxide Mission, *Proc. SPIE,* 7106. doi: 10.1117/12.802194.

Day, J. O., C.W. O'Dell, R. Pollock, C. J. Bruegge, D. Rider, D. Crisp, and C. E. Miller, (2011), Preflight Spectral Calibration of the Orbiting Carbon Observatory, IEEE Trans. Geosci. Remote Sens., 49, 2793-2801.

Hovenier, J. W., and J. F. de Haan (1985), Polarized light in planetary atmospheres for perpendicular directions, *Astron. Astrophys.,*146, 185–191.

Kuze, Akihiko, et al. (2013), Long-Term Vicarious Calibration of GOSAT Short-Wave Sensors: Techniques for Error Reduction and New Estimates of Radiometric Degradation Factors. *IEEE Transactions on Geoscience and Remote Sensing,* 52.7: 1-14.

Kuze, Akihiko, et al. (2011), Vicarious calibration of the GOSAT sensors using the Railroad Valley desert playa. *IEEE Transactions on Geoscience and Remote Sensing*, 49.5: 1781-1795.

Lee et al. (2015), Preflight Spectral Calibration of the Orbiting Carbon Observatory-2, in preparation.

O'Dell, C. W., J. O. Day, R. Pollock, C. J. Bruegge, D. M. O'Brien, R. Castano, I. Tkatcheva, C. E. Miller, and D. Crisp (2011), Preflight Radiometric Calibration of the Orbiting Carbon Observatory *IEEE Transactions on Geoscience and Remote Sensing*, 49.6: 2438-2447. doi: http://dx.doi.org/10.1109/TGRS.2010.2090887.

Frankenberg, C., R. Pollock, R. A. M. Lee, R. Rosenberg, J.-F. Blavier, D. Crisp, C.W. O'Dell, G. B. Osterman, C. Roehl, P. O. Wennberg, and D. Wunch, (2014), The Orbiting Carbon Observatory (OCO-2): spectrometer performance evaluation using pre-launch direct sun measurements, Atmos. Meas. Tech., 7, 1–10. www.atmos-meas-tech.net/7/1/2014/doi:10.5194/amt-7-1-2014.

Rosenberg et al. (2015), Preflight Radiometric Calibration of Orbiting Carbon Observatory 2, in preparation.

Schutgens, A. J., Tilstra, L. G., Stammes, P. and Breon, F.-M. (2004), On the relationship between Stokes parameters $Q$ and $U$ of atmospheric ultraviolet/visible/near-infrared radiation, *J. Geophys. Res.*, 109 (D09205), doi:10.1029/2003JD004081.

Slijkhuis, S. (1998), SCIAMACHY level 0 to 1c processing algorithm theoretical basis document,Tech. Note ENV-ATB-DLR-SCIA-0041,Dtsch. Zent. für Luft- und Raumfahrt, Oberpfaffenhofen, Germany.

van de Hulst, H. C. (1981), Light Scattering by Small Particles , Dover, Mineola, N. Y.

Yoshida, Y., N. Kikuchi, and T. Yokota (2012), On-orbit radiometric calibration of SWIR bands of TANSO-FTS onboard GOSAT. *Atmospheric Measurement Techniques Discussions*, 5.4: 4711-4734.