Harmonized Landsat Sentinel-2 (HLS) Product User Guide

Product Version 2.0

J.G. Masek, J. Ju, M. Claverie, S. Skakun, J.-C. Roger,
E. Vermote, B. Franch, Z. Yin, J. L. Dungan

Principal Investigator: Dr. Jeffrey G. Masek, NASA/GSFC
Correspondence email address: Jeffrey.G.Masek@nasa.gov

Acronyms

AROP	Automated Registration and Orthorectification Package
BRDF	Bidirectional Reflectance Distribution Function
BT	Brightness temperature
CMG	Climate Modelling Grid
ETM+	Enhanced Thematic Mapper Plus
GDAL	Geospatial Data Abstraction Library
GLS	Global Land Survey
HDF	Hierarchical Data Format
HLS	Harmonized Landsat and Sentinel-2
KML	Keyhole Markup Language
MGRS	Military Grid Reference System
MSI	Multi-Spectral Instrument
NBAR	Nadir BRDF-normalized Reflectance
OLI	Operational Land Imager
QA	Quality assessment
RSR	Relative spectral response
SDS	Scientific Data Sets
SR	Surface reflectance
SZA	Sun zenith angle
TM	Thematic Mapper
TOA	Top of atmosphere
UTM	Universal Transverse Mercator
WRS	Worldwide Reference System

1 Introduction

The Harmonized Landsat and Sentinel-2 (HLS) project is a NASA initiative and collaboration with USGS to produce compatible surface reflectance (SR) data from a virtual constellation of satellite sensors, the Operational Land Imager (OLI) and Multi-Spectral Instrument (MSI) onboard the Landsat-8 and Sentinel-2 remote sensing satellites respectively. The combined measurement enables global land observation every 2-3 days at moderate (30 m) spatial resolution. The HLS project uses a set of algorithms to derive seamless products from OLI and MSI: atmospheric correction, cloud and cloud-shadow masking, spatial co-registration and common gridding, view angle normalization and spectral bandpass adjustment. The HLS data products can be regarded as the building blocks for a "data cube" so that a user may examine any pixel through time and treat the near-daily reflectance time series as though it came from a single sensor.

The HLS suite contains two products, S30 and L30, derived from Sentinel-2 L1C and Landsat L1TP (Collection 2) input, respectively. They are gridded into the same MGRS tiles with a 30m pixel size.

2 New in v2.0

HLS v2.0 builds on v1.4 by updating and improving processing algorithms, expanding spatial coverage, and providing validation. Particular updates are as follows:

- *Global coverage*. All global land, including major islands but excluding Antarctica, is covered.
- *Input data*. Landsat 8 Collection-2 (C2) data from USGS are used as input; better geolocation is expected as C2 data use the Sentinel-2 Global Reference Image (GRI) as an absolute reference.
- *Atmospheric correction*. A USGS C version of LaSRCv3.5.5 is applied for both Landsat 8 and Sentinel-2 data for computational speedup. LaSRCv3.5.5 has been validated for both Landsat 8 and Sentinel-2 within the CEOS ACIX-I (Atmospheric Correction Inter-Comparison eXercise, http://calvalportal.ceos.org/projects/acix).
- *QA band.* The QA band is generated exclusively by and named after Fmask, consistently for the two HLS products (S30 and L30). Like in v1.4 aerosol thickness level from atmospheric correction is also incorporated into the QA band.
- *BRDF adjustment*. BRDF adjustment mainly normalizes the view angle effect, with the sun zenith angle largely intact. This adjustment is applied to the Sentinel-2 red-edge bands as well.
- *Sun and view angle bands are provided.*
- *Product format*. The product is delivered in individual Cloud Optimized GeoTIFF (COG) files to allow for spectral and spatial subsetting in applications.
- *Temporal Coverage and Latency*. Version v2.0 moves toward "keep up" processing. The intent is to continually update products with <2-day latency. Users are cautioned however that HLS is still a research product.

3 Products overview

3.1 Input data

The Operational Land Imager (OLI) sensor is a moderate spatial resolution multi-spectral imager onboard the Landsat-8 satellite, in a sun-synchronous orbit with a 705 km altitude and a 16-day

repeat cycle. The sensor acquires data with a 15-degree field of view resulting in approximately a 185 km image swath. The OLI sensor has 9 solar reflective bands and the data are co-registered with the data from the 2-band TIRS (Thermal Infrared Sensor) instrument onboard the same Landsat-8 satellite (Irons et al., 2012). The native spatial resolution is 30 m for OLI and 100 m for TIRS, but TIRS data are resampled to 30 m for distribution. HLS v2.0 uses Landsat-8 Collection-2[1] Level-1 top-of-atmosphere (TOA) product as input: for "keep-up" processing, the Real-Time data with geolocation RMSE <= 12 m (i.e. Tier-1 equivalent) are used and, for back processing, Tier-1 data are used. The Real-Time TOA OLI data have the same quality as the tier-based data do, but the Real-Time TIRS data may have lesser geolocation and radiometric quality.

The Sentinel-2 Multi-Spectral Instrument (MSI) is onboard the Sentinel-2A and -2B satellites in a sun-synchronous orbit with a 786 km altitude and a combined 5-day repeat cycle. The sensor has a 20.6° field of view corresponding to an image swath of approximately 290 km. The spatial resolution varies with the spectral bands: 10 m for the visible bands and the broad NIR band, 20 m for the red edge, narrow NIR and SWIR bands, and 60 m for the atmospheric bands (Drusch et al., 2012). HLS v2.0 uses the Level-1C (L1C) Top of Atmosphere data as input. Table 1 provides an overview of Landsat 8 and Sentinel-2 characteristics.

Table 1: Input data characteristics

		Landsat 8/OLI-TIRS	**Sentinel-2A/MSI**	**Sentinel-2B/MSI**
Launch date		February 11, 2013	June 23, 2015	March 7, 2017
Equatorial crossing time		10:00 a.m.	10:30 a.m.	10:30 a.m.
Spatial resolution		30 m (OLI) / 100 m (TIRS)	10 m / 20 m / 60 m (see spectral bands)	
Swath / Field of view		180 km / 15°	290 km / 20.6°	
Spectral bands (central wavelength)	Ultra blue	443 nm	443 nm (60 m)	
	Visible	482 nm, 561 nm, 655 nm	490 nm (10 m), 560 nm (10 m), 665 nm (10m)	
	Red edge	-	705 nm (20 m), 740 nm (20 m), 783 nm (20 m)	
	NIR	865 nm	842 nm (10 m), 865 nm (20 m)	
	SWIR	1609 nm, 2201 nm	1610 nm (20 m), 2190 nm (20 m)	
	Cirrus	1373 nm	1375 nm (60 m)	
	Water Vapor	-	945 nm (60 m)	
	Thermal	10.9 µm, 12 µm	-	

[1] Landsat Collections — https://www.usgs.gov/landsat-missions/landsat-collections

3.2 Overall HLS processing flowchart

The same processing methods are applied to generate S30 and L30 (Fig. 1). LaSRC is used for atmospheric correction, Fmask for cloud masking (QA). Landsat data are gridded into the tiles that MSI use; all pixels are resampled to 30 m. Surface reflectance is corrected for view angle effect. MSI bandpasses are adjusted to the Landsat ones. A detailed description of processing methods can be found in Section 4.

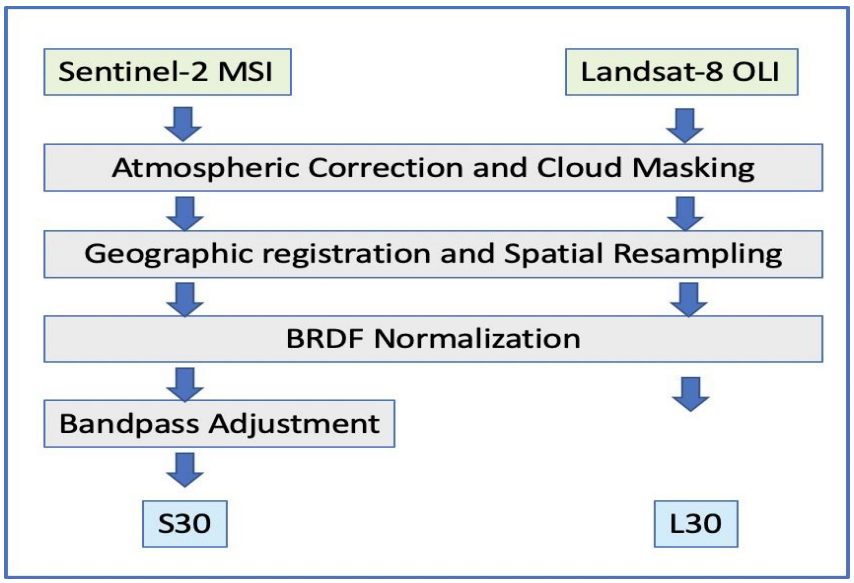

Figure 1: HLS science algorithm processing flow

3.3 Products specifications

The HLS product specifications are given in Table 2.

Table 2: HLS products specifications

Product Name	**L30**	**S30**
Input sensor	Landsat-8 OLI/TIRS	Sentinel-2A/B MSI
Spatial resolution	30 m	30 m
BRDF-adjusted	Yes (except band 09)	Yes (except bands 09, 10)
Bandpass-adjusted	No (HLS uses OLI bandpass)	Adjusted to OLI-like (except red edge, water vapor and cirrus bands)
Projection	UTM	UTM
Tiling system	MGRS (110*110)	MGRS (110*110)

3.4 Spectral bands

All Landsat-8 OLI and Sentinel-2 MSI reflective spectral bands nomenclatures are retained in the HLS products (Table 3).

Table 3: HLS spectral bands nomenclature

Band name	OLI band number	MSI band number	HLS band code name L8	HLS band code name S2	Wavelength (micrometers)
Coastal Aerosol	1	1	B01	B01	0.43 – 0.45*
Blue	2	2	B02	B02	0.45 – 0.51*
Green	3	3	B03	B03	0.53 – 0.59*
Red	4	4	B04	B04	0.64 – 0.67*
Red-Edge 1	–	5	–	B05	0.69 – 0.71**
Red-Edge 2	–	6	–	B06	0.73 – 0.75**
Red-Edge 3	–	7	–	B07	0.77 – 0.79**
NIR Broad	–	8	–	B08	0.78 –0.88**
NIR Narrow	5	8A	B05	B8A	0.85 – 0.88*
SWIR 1	6	11	B06	B11	1.57 – 1.65*
SWIR 2	7	12	B07	B12	2.11 – 2.29*
Water vapor	–	9	–	B09	0.93 – 0.95**
Cirrus	9	10	B09	B10	1.36 – 1.38*
Thermal Infrared 1	10	–	B10	–	10.60 – 11.19*
Thermal Infrared 2	11	–	B11	–	11.50 – 12.51*

* from OLI specifications
** from MSI specifications

3.5 Output projection and gridding

HLS has adopted the tiling system used by Sentinel-2. The tiles are in the Universal Transverse Mercator (UTM) projection and are 109,800 m (110 km nominally) on a side. The tiling system is aligned with the UTM-based Military Grid Reference System (MGRS). The UTM system divides the Earth's surface into 60 longitude zones, each 6° of longitude in width, numbered 1 to 60 from 180° West to 180° East. Each UTM zone is divided into latitude bands of 8°, labeled with letters C to X from South to North. A useful mnemonic is that latitude bands N and later are in the Northern Hemisphere. Each 6°×8° polygon (grid zone) is further divided into the 110 km ×110 km Sentinel-2 tiles labeled with letters. For example, tile 11SPC is in UTM zone 11, latitude band S (in Northern Hemisphere), and labeled P in the east-west direction and C in the south-north direction within grid zone 11S. Users should note that there is horizontal and vertical overlap of around 8-10 km between two adjacent tiles in the same UTM zone. For the two adjacent tiles both straddling a UTM zone boundary, the overlap of may be much greater. A KML file produced by ESA showing the location of all Sentinel-2 tiles is available at: https://sentinel.esa.int/documents/247904/1955685/S2A_OPER_GIP_TILPAR_MPC__20151209T095117_V20150622T000000_21000101T000000_B00.kml

One trivial difference from the ESA gridding is that HLS inherits the USGS UTM convention of keeping the Y coordinate for the Southern Hemisphere negative, therefore with no need for hemisphere specification. In contrast, many spatial data handling tools use a convention of adding 10,000,000 meters to make the southern coordinate positive (i.e. use of a false northing 10,000,000) and thus need to indicate which hemisphere to avoid confusion. These tools can recognize the USGS convention.

4 Description of Algorithms

4.1 Atmospheric correction

The same atmospheric correction algorithm, Land Surface Reflectance Code (LaSRC) developed by Eric Vermote (NASA/GSFC) (Vermote et al., 2016), is applied to data from both sensors. LaSRC is based on the 6S radiative transfer model and a heritage from the MODIS MCD09 products (Vermote and Kotchenova 2008) as well as the earlier LEDAPS algorithm implemented for Landsat-5 and Landsat-7 (Masek et al. 2006). A detailed description of the method is given in Vermote et al. (2016), and results of surface reflectance validation for Landsat 8 and Sentinel-2 within CEOS ACIX-I are provided in Doxani et al. (2018).

LaSRC uses atmospheric inputs (ozone, water vapor) from MODIS to correct for gaseous absorption and surface pressure based on topographic elevation to correct for molecular (Rayleigh) scattering. Aerosol optical thickness (fixed continental type) is derived via an image-based algorithm using the ratio of the red and blue spectral bands (Vermote et al., 2016). The output is directional surface reflectance. HLS also includes the two thermal infrared bands from the Landsat 8 TIRS sensor in the L30 product – these values are not atmospherically corrected, but are rescaled apparent brightness temperature (no atmosphere, unity emissivity).

HLS 2.0 uses a C version of LaSRC v3.5.5 implemented by USGS, mainly for computational speedup.

4.2 Spatial co-registration of input data

Our objective in HLS is to maintain the geodetic accuracy requirement of the Sentinel-2 images (<20 m error, 2σ) and improve the multi-temporal co-registration among Sentinel-2 images and between Sentinel-2 and Landsat 8 images (<15 m 2σ) for the 30 m products. This specification supports time series monitoring of small fields, man-made features, and other spatially heterogeneous cover types.

Co-registration is less of a concern in HLS v2.0, as newly processed input with better geolocation becomes available, but we describe the methodology as it still has relevance for earlier MSI L1C data. Before HLS v2.0, two issues impeded a direct registration of Landsat 8 and Sentinel-2 imagery without additional processing. First, while the relative co-registration of Landsat 8/OLI Collection-1 imagery was quite accurate (<6.6m, Storey et al. 2014), the absolute geodetic accuracy varied with the quality of the Global Land Survey 2000 (GLS2000) ground control around the world. In some locations, the GLS geodetic accuracy was in error by up to 38 m (2σ, Storey et al. 2016). As a result, Sentinel-2/MSI and Landsat 8/OLI Level-1 products did not align to sub-pixel precision for those locations (Storey et al. 2016). Second, an error in the yaw characterization for the MSI L1C images processed before v02.04 (May 2016) caused misregistration between the edges of MSI images acquired from adjacent orbits (ESA 2018). The misregistration of up to 2.8 pixels at 10 m resolution between Sentinel-2A images from adjacent orbits has been observed by Skakun et al. (2017) and Yan et al. (2018). Although the issue was fixed with L1C version 02.04 (yielding a measured absolute geolocation of less than 11m at 95.5% confidence, ESA 2018), archived Sentinel-2 data from 2015-2016 will continue to have this error until the entire archive is reprocessed by ESA.

Earlier HLS versions used the Automated Registration and Orthorectification package (AROP, Gao et al. 2009) to register Landsat imagery to a “master” Sentinel-2 image for each tile (see

Claverie et al., 2018 for details). However, HLS L30 of v2.0 is based on the USGS Collection 2 Landsat data, which now uses the Sentinel-2 Global Reference Image (GRI) as an absolute control. As a result, the improved Landsat ground control in Collection 2 eliminates the need for AROP for L30 production. However, cubic convolution resampling is still needed for L30 production because USGS aligns the UTM coordinate origin with a pixel center while ESA aligns it with a pixel corner. In addition, AROP is still required for 2015-2016 Sentinel-2 input data due to the yaw steering issue described above. When the entire Sentinel-2 archive is consistently processed to be on a collection basis, the use of AROP for early Sentinel-2 data will be retired.

4.3 Quality assessment mask

HLS provides per-pixel masking of cloud, cloud shadow, snow, water, and aerosol optical thickness levels. In earlier versions of HLS, the cloud mask was a union of cloud masks accompanying the Level-1 input, the internal cloud mask from atmospheric correction code LaSRC, and the cloud mask by Fmask (Zhu et al. 2015). In HLS v2.0, the cloud mask was generated exclusively by Fmask 4.2, an update of Fmask 4.0 reported in Qiu et al. (2019). Aerosol optical thickness level created during atmospheric correction is also incorporated into the per-pixel quality assessment mask, like in HLS v1.4.

4.4 View and illumination angles normalization

The L30 and S30 Nadir BRDF-Adjusted Reflectance (NBAR) products are surface reflectance normalized for the view angle and the illumination angle effect, using the *c*-factor technique by Roy et al. (2016). The view angle is set to nadir for all pixels in normalization. The illumination angle for a tile is set to the mean value of the solar zenith angles at the tile center at the respective times when Landsat-8 and Sentinel-2 overpass the tile center's latitude on the day; this angle is derived using the code described in Li et al (2018).

The BRDF normalization uses a set of constant BRDF coefficients, derived from 12-month MODIS 500 m global BRDF product (MCD43) (more than 15 billion pixels). The derived BRDF coefficients are applied to OLI and MSI bands equivalent to MODIS ones. The technique has been evaluated using off-nadir (i.e. in the overlap areas of adjacent swaths) ETM+ data (Roy et al. 2016) and MSI data (Roy et al. 2017). For the normalization of MSI red-edge spectral bands that have no MODIS equivalents, the linearly interpolated BRDF coefficients from the enclosing MODIS red and NIR wavelength bands are used (Roy et al 2017). BRDF coefficients for the three kernels (isotropic, geometric, and volumetric) are shown in the Table 4. The kernel definitions are described in the ATBD of MOD43 product (Strahler et al. 1999).

Table 4: BRDF coefficients used for the c-factor approach (Roy et al. 2016 and 2017)

Band name	**HLS band code name L8**	**HLS band code name S2**	**Equivalent MODIS band**	**f_{iso}**	**f_{geo}**	**f_{vol}**
Coastal/Aerosol	B01	B01	3	0.0774	0.0079	0.0372
Blue	B02	B02	3	0.0774	0.0079	0.0372
Green	B03	B03	4	0.1306	0.0178	0.0580
Red	B04	B04	1	0.1690	0.0227	0.0574

Red-Edge 1	–	B05	-	0.2085	0.0256	0.0845
Red-Edge 2	–	B06	-	0.2316	0.0273	0.1003
Red-Edge 3	–	B07	-	0.2599	0.0294	0.1197
NIR Broad	–	B08	2	0.3093	0.0330	0.1535
NIR Narrow	B05	B8A	2	0.3093	0.0330	0.1535
SWIR 1	B06	B11	6	0.3430	0.0453	0.1154
SWIR 2	B07	B12	7	0.2658	0.0387	0.0639

$$\rho(\lambda, \theta^{Norm}) = c(\lambda) \times \rho(\lambda, \theta^{sensor}) \quad (2)$$

$$c(\lambda) = \frac{f_{iso}(\lambda)+f_{geo}(\lambda)\times K_{geo}(\theta^{Norm})+f_{vol}(\lambda)\times K_{vol}(\theta^{Norm})}{f_{iso}(\lambda)+f_{geo}(\lambda)\times K_{geo}(\theta^{sensor})+f_{vol}(\lambda)\times K_{vol}(\theta^{sensor})} \quad (3)$$

where: $\theta^{Norm} \Leftrightarrow (\theta_v = 0, \theta_s = \theta_s^{out}, \Delta\varphi = 0)$

$\theta_{sensor} \Leftrightarrow (\theta^{sensor} = \theta_v^{sensor}, \theta_s = \theta_s^{sensor}, \Delta\varphi = \Delta\varphi^{sensor})$

The BRDF effect is caused predominantly by the view angle variation and secondarily by the solar angle variation. The normalization of the solar zenith angle is out of two considerations. First, Landsat-8 and Sentinel-2 overpass the same latitude 30 minutes apart; on these rare days when the Landsat-8 and a Sentinel-2 overpass the same ground location, the solar zenith will be different due to the 30-minute time difference. Second, since the solar zenith angle increases from east to west within a swath, the solar zenith angle for the overlapping area of two swaths can be different due to the tile's relative location change within the swaths. These points are illustrated for a tile near the Equator where the solar zenith angle changes most dramatically in these cases.

Tile 19NGA with its center at 0.41N and 66.71W is right on the equator. The mean solar zenith angle for the Landsat-8 image on the tile is greater than that for a temporally close Sentinel-2 image because Landsat-8 overpasses 30 minutes earlier (Fig. 2). There is also temporal oscillation in solar zenith angle in a sensor's image time series. The mean solar zenith angle of each Sentinel-2 granule in 2019 follows two curves, which in fact originated from two adjacent orbits, not from the coexistence of S2A and S2B; when the tile is located to the east of the nadir view in the original image swath, the solar zenith angle was smaller (the lower curve of Sentinel-2 in Fig. 2) than in a temporally close image when the tile is located to the west of the nadir view (the upper curve of Sentinel-2). The observed solar zenith angle oscillates day to day between the two orbits. Similar pattern was present in the Landsat images over this tile, which was also observed from two adjacent Landsat orbits (Fig. 2). The solar zenith angle temporal oscillation in Landsat time series is smaller because Landsat image swath is narrower than Sentinel-2's (185 km vs 290 km).

The solar zenith angle used in normalization is the mean of the solar zenith angles at the respective times that Landsat-8 and Sentinel-2 overpass a tile center's latitude. This prescribed solar zenith angle is calculated using the software provided by Li et al (2018). The idea is based on the fact that a sensor overpasses the same latitude at the same local solar time and therefore the solar zenith angle will be the same at nadir for the same latitude on the same day. The angle normalization takes into account all types of variations presented in Figure 2, but at the same time allows for the smooth change of daily solar zenith angle due to daily solar declination change.

For high-latitude tiles with their centers above the highest latitude that Landsat-8 and Sentinel-2's nadir view can reach (81.8 degrees and 81.38 degrees respectively), the NBAR solar zenith angle prescribed by Li et al (2018) cannot be applied, and the mean observed solar zenith angle in the tile is used instead. Only about 20 land tiles fall into this category.

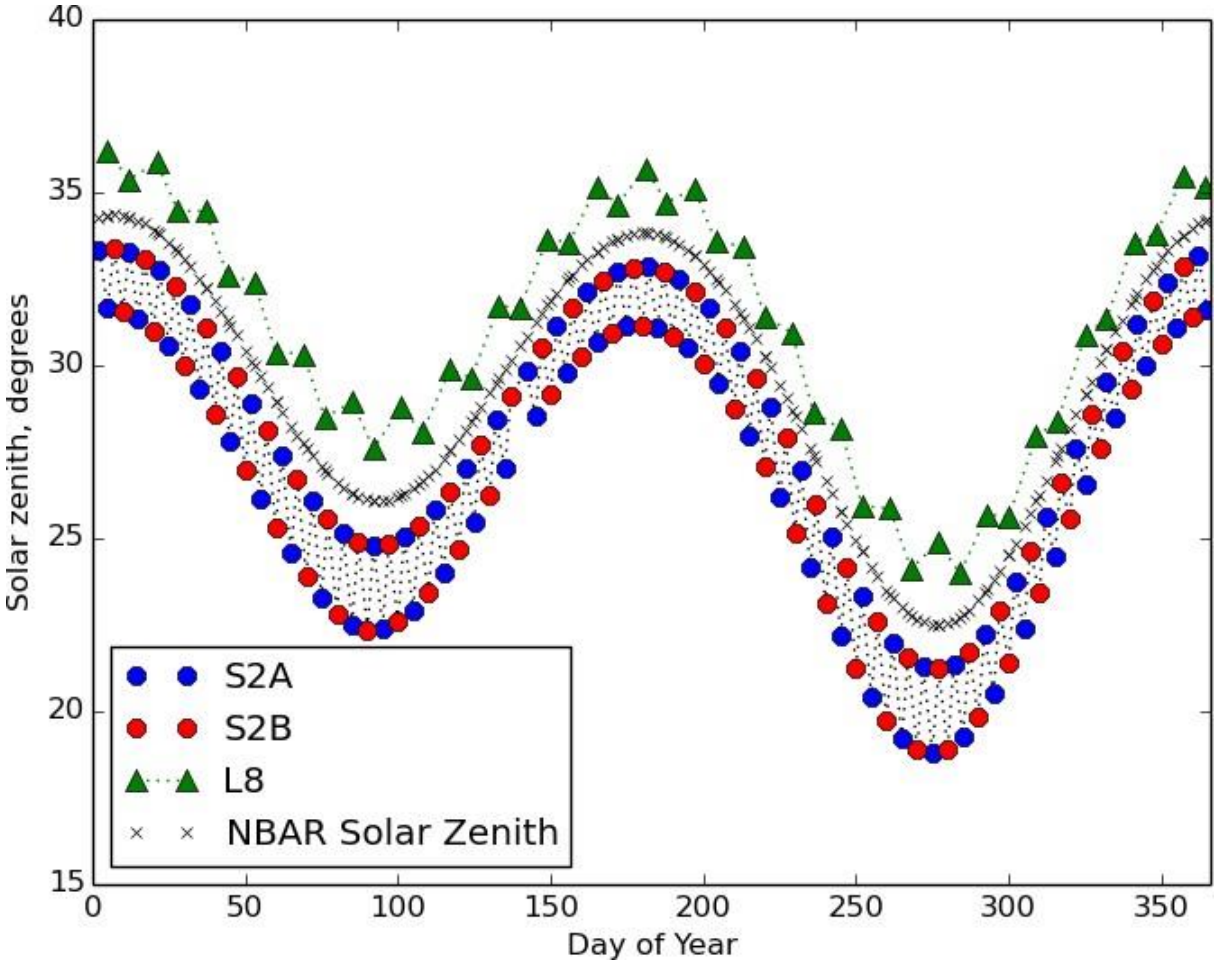

Figure 2. The observed mean solar zenith angle in each tiled Sentinel-2 and Landsat-8 image and the solar zenith angle used in each image's BRDF normalization, for an equatorial tile 19NGA in 2019. The observed mean solar zenith angle in a Landsat image is higher than that in a temporally close Sentile-2 image because Landsat overpasses 30 minutes earlier. There is also day-to-day oscillation in mean observed solar zenith angle for each sensor due to the alternating observation from two adjacent orbits.

4.5 Bandpass adjustment

MSI and OLI have slightly different bandpasses for equivalent spectral bands, and these differences need to be removed in HLS products. The OLI spectral bandpasses are used as reference, to which the MSI spectral bands are adjusted. The bandpass adjustment is a linear fit between equivalent spectral bands. The slope and offset coefficients were computed based on 500 hyperspectral spectra selected on 160 globally distributed Hyperion scenes processed to surface reflectance and used to synthesis MSI and OLI bands. MSI's RSRs correspond to the version v2.0 (Claverie et al., 2018). The spectral differences between MSI onboard Sentinel-2A

(S2A) and Sentinel-2B (S2B) are accounted. Note that the S2A CA and Blue bands RSRs correspond to S2B RSRs. The coefficients are given in Table 5, and scatterplots are given in Figure 3 and Figure 4.

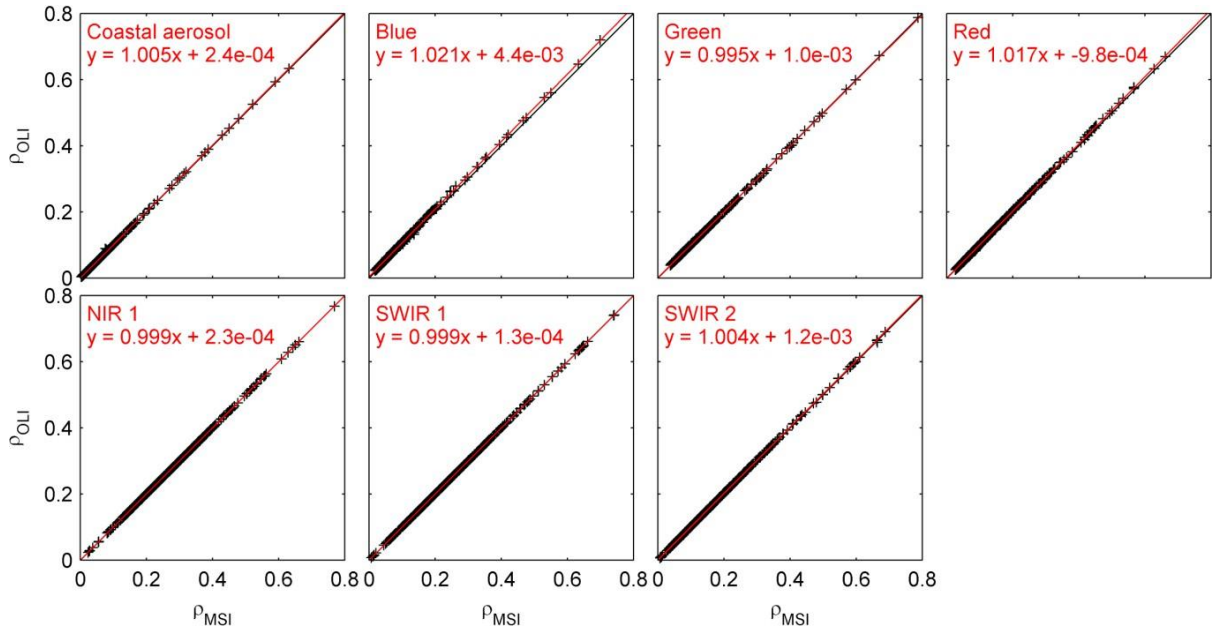

Figure 3: Sentinel-2A MSI vs OLI surface reflectance for the seven equivalent bands, using a synthetic dataset built with 500 surface reflectance spectra from Hyperion.

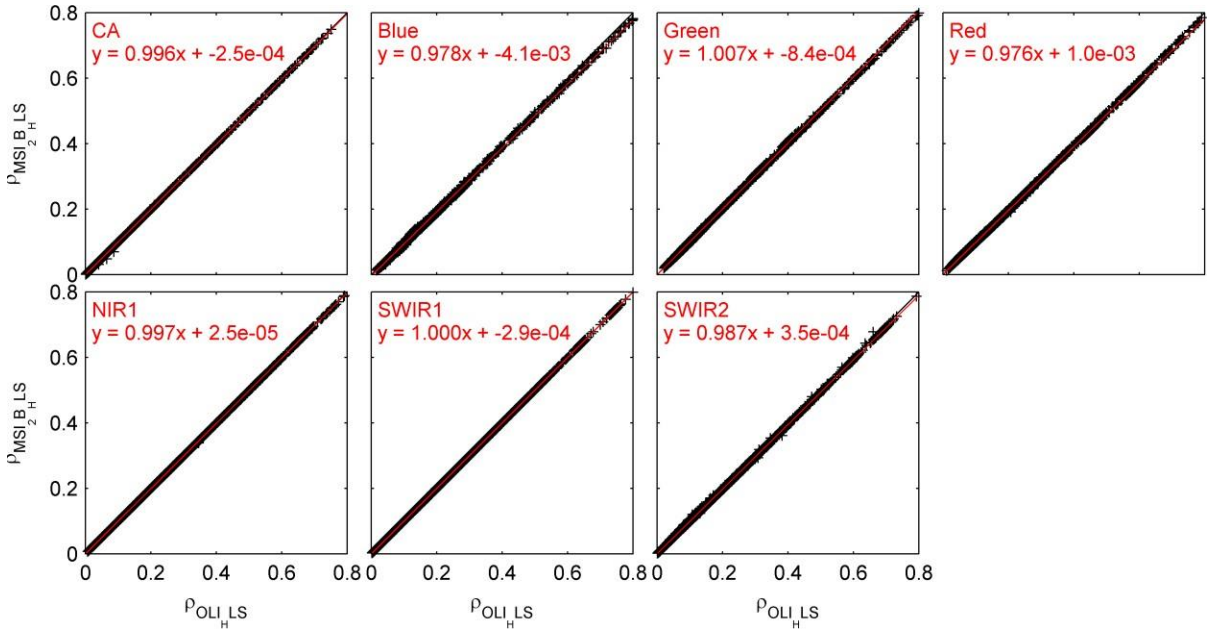

Figure 4: Same as Figure 3, but for Sentinel-2B.

$$\rho_{OLI} = a \times \rho_{MSI} + b \quad (5)$$

Table 5: Bandpass adjustment coefficients

			Sentinel-2A		Sentinel-2B	
HLS Band name	OLI band number	MSI band number	Slope (a)	Intercept (b)	Slope (a)	Intercept (b)
CA	1	1	0.9959	-0.0002	0.9959	-0.0002
BLUE	2	2	0.9778	-0.004	0.9778	-0.004
GREEN	3	3	1.0053	-0.0009	1.0075	-0.0008
RED	4	4	0.9765	0.0009	0.9761	0.001
NIR1	5	8A	0.9983	-0.0001	0.9966	0.000
SWIR1	6	11	0.9987	-0.0011	1.000	-0.0003
SWIR2	7	12	1.003	-0.0012	0.9867	0.0004

4.6 Spatial resampling

Spectral data resampling in L30 creation uses cubic convolution. With the improved geolocation in the Collection-2 Landsat-8 data, AROP-based temporal co-registration has phased out. But spatial resampling of Landsat-8 data is still needed in gridding the data into the MGRS tiles used by Sentinel-2, because even if the Landsat image and the intended MGRS tile are in the same UTM zone, ESA registers a corner of a Sentinel-2 pixel to the UTM coordinate origin, but USGS registers the center of a Landsat pixel to the UTM coordinate origin. A simple coordinate shift will not work. Moreover, when the Landsat data and the MGRS tile are in different UTM zones, reprojection and resampling must always be performed.

Spectral data resampling in S30 creation uses the simple area-weighted average because of the nesting relationship between the original 10 m, 20 m, 60 m input pixels and the desired output 30 m pixels. To produce an S30 pixel, a set of 3x3 10 m pixels are averaged with equal weights, 2x2 20 m pixels are averaged with weights 4/9, 2/9, 2/9, and 1/9 depending on the relative position, and a 60 m pixel is duplicated to produce 2x2 S30 pixels.

To create S30 from Sentinel-2 L1C input of processing baseline prior to 2.04 (approximately mid-2016), additional resampling is needed, before the area-weighted average, as part of the AROP-based temporal co-registration. Once LaSRC derives the surface reflectance, cubic convolution is applied to resample the 10 m/20 m/60 m spectral bands based on AROP-derived coordinate transform. Then the aera-weighted average described above is applied. Only a small proportion of L1C images are of processing baseline prior to 2.04.

Resampling of the QA layer is implemented differently. The QA layer is mostly based on cloud masking by Fmask, which outputs at 30 m over the WRS-2 path-row grid for Landsat and at 20 m over the MGRS grid for Sentinel-2. It also incorporates the qualitative aerosol optical thickness level assessment from the atmospheric correction. The resample of the QA is performed for each bit of the QA byte in turn:

- L30: while a 4x4 window is used in cubic convolution resampling of the spectral data, only the innermost 2x2 pixels are examined for QA resampling because 1) almost all the cubic convolution weights are in the 2x2 window, 2) a 4x4 window would artificially create a mixture situation. The "presence" rule is used: a QA bit value 1 in any of the 4 input pixels causes the output bit to be set to 1 for an L30 pixel and an output L30 QA bit is set to 0 only if all the 2x2 input pixels have 0 at that bit.

– S30: 10 m QA are resampled to 30 m using the same “presence” rule. A QA bit value 1 in any of the nine nesting input 10 m pixels causes the output bit to be set to 1 for S30. That is, a QA bit for S30 is set to 0 only if the input QA bit values for all the nine nesting 10 m pixels are 0.

If a resampling window contains a mixture of QA bits, the “presence” rule will make the output QA bits not mutually exclusive. For example, the output QA bits may indicate a pixel is cloud, cloud shadow and water at the same time. This is not a mistake; the mixture nature of HLS QA bits allows the users to select/discard data in a way they want.

The aerosol optical thickness level resample is special because it has two bits (bits 6-7 in QA). It is resampled in such a way that a higher aerosol thickness level dominates a lower aerosol level in output. For example, if any of the input pixels in the sampling window has “high aerosol,” the output aerosol level will be “high aerosol,” and if the highest aerosol level in the sampling window is “moderate aerosol,” the output will be “moderate aerosol,” and so on.

5 Spatial coverage

HLS v2.0 covers all the global land except Antarctica, as depicted in a land mask (Fig. 5) derived from the NOAA shoreline dataset (https://www.ngdc.noaa.gov/mgg/shorelines/data/gshhg/latest/). Antarctica is excluded because of low solar elevations which compromise the plane-parallel atmospheric correction.
Note that the data acquisition over some small oceanic islands by the Landsat and Sentinel-2 sensor may not be made regularly.

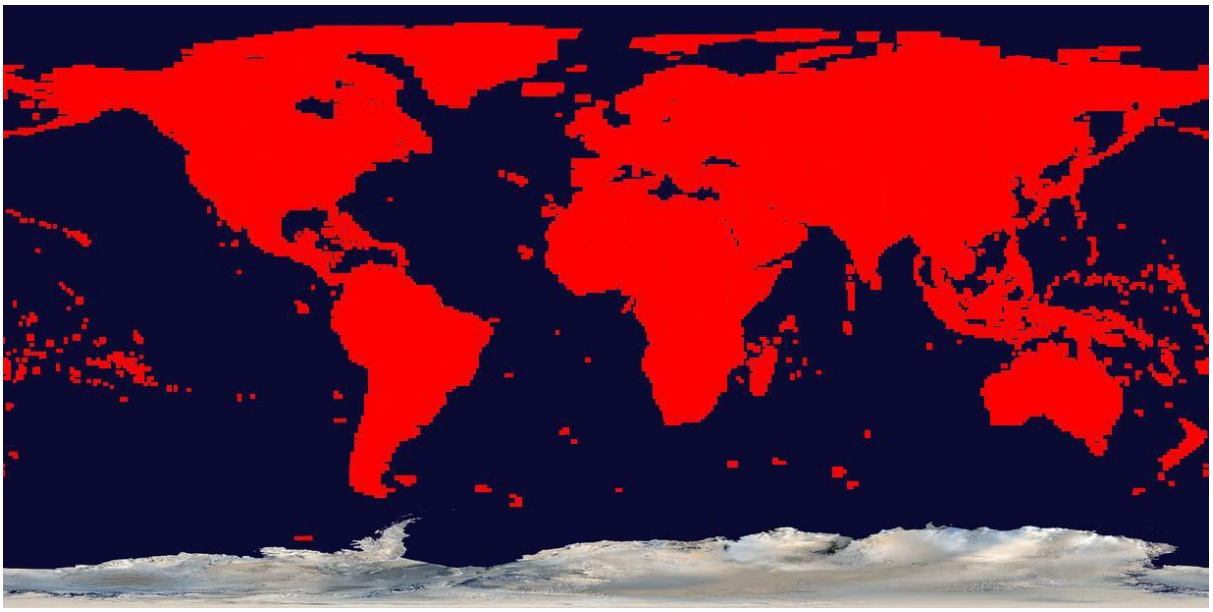

Figure 5: HLS v2.0 covers the global land, including major islands but excluding Antarctica.

6 Product formats

6.1 File format

HLS 2.0 products are in Cloud Optimized GeoTIFF (COG), one file per data layer to offer the flexibility of only downloading the needed data layers and, for cloud-based applications, the needed spatial subsets within a tile. The COG files are internally compressed.

L30 data are stored in directories such as HLS.L30.T17SLU.2020209T155956.v2.0/, which suggests L30 over tile 17SLU from data acquired on day 209 of 2020 specifically at UTC 155956. This example product consists of the following files:

HLS.L30.T17SLU.2020209T155956.v2.0.B01.tif
HLS.L30.T17SLU.2020209T155956.v2.0.B02.tif
HLS.L30.T17SLU.2020209T155956.v2.0.B03.tif
HLS.L30.T17SLU.2020209T155956.v2.0.B04.tif
HLS.L30.T17SLU.2020209T155956.v2.0.B05.tif
HLS.L30.T17SLU.2020209T155956.v2.0.B06.tif
HLS.L30.T17SLU.2020209T155956.v2.0.B07.tif
HLS.L30.T17SLU.2020209T155956.v2.0.B09.tif
HLS.L30.T17SLU.2020209T155956.v2.0.B10.tif
HLS.L30.T17SLU.2020209T155956.v2.0.B11.tif
HLS.L30.T17SLU.2020209T155956.v2.0.Fmask.tif
HLS.L30.T17SLU.2020209T155956.v2.0.SZA.tif
HLS.L30.T17SLU.2020209T155956.v2.0.SAA.tif
HLS.L30.T17SLU.2020209T155956.v2.0.VZA.tif
HLS.L30.T17SLU.2020209T155956.v2.0.VAA.tif
HLS.L30.T17SLU.2020209T155956.v2.0.cmr.xml
HLS.L30.T17SLU.2020209T155956.v2.0.json
HLS.L30.T17SLU.2020209T155956.v2.0.jpg

The filenames for individual spectral bands and Fmask cloud mask are self-explaining. Sun zenith angle (SZA), sun azimuth angle (SAA), view zenith angle (VZA) and view azimuth angle (VAA) file are also provided; see Section 6.4 for details. File HLS.L30.T17SLU.2020209T155956.v2.0.cmr.xml is the metadata file, HLS.L30.T17SLU.2020209T155956.v2.0.json contains the size and checksum value of each file, and HLS.L30.T17SLU.2020209T155956.v2.0.jpgis a natural-color browse image.

The UTC time in the filenames is the sensing time at the input Landsat-8 scene center. After gridding into the MGRS tiles it does not accurately indicate the sensing time over the tile. If two scenes overlap a MGRS tile, the sensing time of one the scenes are chosen by chance. So, this timing information is not accurate for the MGRS tile center; it is intended mainly as an identifier, not for quantitative analysis.

S30 data are stored in the same format. An example directory HLS.S30.T17SLU.2020117T160901.v2.0 may contain the following files:

HLS.S30.T17SLU.2020117T160901.v2.0.B01.tif
HLS.S30.T17SLU.2020117T160901.v2.0.B02.tif
HLS.S30.T17SLU.2020117T160901.v2.0.B03.tif
HLS.S30.T17SLU.2020117T160901.v2.0.B04.tif
HLS.S30.T17SLU.2020117T160901.v2.0.B05.tif
HLS.S30.T17SLU.2020117T160901.v2.0.B06.tif
HLS.S30.T17SLU.2020117T160901.v2.0.B07.tif
HLS.S30.T17SLU.2020117T160901.v2.0.B08.tif
HLS.S30.T17SLU.2020117T160901.v2.0.B8A.tif
HLS.S30.T17SLU.2020117T160901.v2.0.B09.tif
HLS.S30.T17SLU.2020117T160901.v2.0.B10.tif

HLS.S30.T17SLU.2020117T160901.v2.0.B11.tif
HLS.S30.T17SLU.2020117T160901.v2.0.B12.tif
HLS.S30.T17SLU.2020117T160901.v2.0.Fmask.tif
HLS.S30.T17SLU.2020117T160901.v2.0.SZA.tif
HLS.S30.T17SLU.2020117T160901.v2.0.SAA.tif
HLS.S30.T17SLU.2020117T160901.v2.0.VZA.tif
HLS.S30.T17SLU.2020117T160901.v2.0.VAA.tif
HLS.S30.T17SLU.2020117T160901.v2.0.cmr.xml
HLS.S30.T17SLU.2020117T160901.v2.0.json
HLS.S30.T17SLU.2020117T160901.v2.0.jpg

The UTC time in the S30 product filenames is the time the sensor begins to sense the sun-lit side of the earth for each orbit, not the exact sensing time over the tile center. When a sequence of observations is available on the same day at the high latitude, they can still be differentiated by this timing information.

6.2 L30

The product L30 contains Landsat-8 OLI surface reflectance and TOA TIRS brightness temperature gridded at 30 m spatial resolution in MGRS tiles. Table 6 lists all the data layers of the L30 product.

Table 6: All the data layers of the L30 product (SR = Surface Reflectance, NBAR = Nadir BRDF-normalized Reflectance, TOA Refl. = Top of Atmosphere Reflectance, TOA BT = Top of Atmosphere Brightness temperature).

Data layer	OLI band number	Units	Data type	Scale	Fill value	Spatial Resolution	Description
B01	1	reflectance	int16	0.0001	-9999	30	NBAR
B02	2	reflectance	int16	0.0001	-9999	30	
B03	3	reflectance	int16	0.0001	-9999	30	
B04	4	reflectance	int16	0.0001	-9999	30	
B05	5	reflectance	int16	0.0001	-9999	30	
B06	6	reflectance	int16	0.0001	-9999	30	
B07	7	reflectance	int16	0.0001	-9999	30	
B09	9	reflectance	int16	0.0001	-9999	30	TOA Refl.
B10	10	degree °C	int16	0.01	-9999	30	TOA BT
B11	11	degree °C	int16	0.01	-9999	30	
FMASK (*Table 9*)	-	none	uint8	-	255	30	Quality bits

S30

The product S30 contains MSI surface reflectance at 30 m spatial resolution.
Table 7 lists all the data layers of the S30 product.

Table 7: list of the SDS of the S30 product (SR = Surface Reflectance, NBAR = Nadir BRDF-Adjusted Reflectance, TOA Refl. = Top of Atmosphere Reflectance).

Data layer	MSI band number	Units	Data type	Scale	Fill value	Spatial Resolution	Description
B01	1	reflectance	int16	0.0001	-9999	30	NBAR
B02	2	reflectance	int16	0.0001	-9999	30	
B03	3	reflectance	int16	0.0001	-9999	30	
B04	4	reflectance	int16	0.0001	-9999	30	
B05	5	reflectance	int16	0.0001	-9999	30	
B06	6	reflectance	int16	0.0001	-9999	30	
B07	7	reflectance	int16	0.0001	-9999	30	
B08	8	reflectance	int16	0.0001	-9999	30	
B8A	8A	reflectance	int16	0.0001	-9999	30	
B09	9	reflectance	int16	0.0001	-9999	30	TOA Refl.
B10	10	reflectance	int16	0.0001	-9999	30	
B11	11	reflectance	int16	0.0001	-9999	30	NBAR
B12	12	reflectance	int16	0.0001	-9999	30	
FMASK (*Table 9*)	-	none	uint8	-	255	30	Quality bits

6.3 The sun and view angles

HLS v2.0 also provides the sun zenith/azimuth and view zenith/azimuth angles used in BRDF correction; in case a user may want to do BRDF correction differently. The S30 angle data is interpolated from the ESA-provided 5 km angles in a text form; HLS selects the view angle of the 2nd red-edge band and uses it on all bands. The L30 angle data is provided in the Collection-2 data; it is originally derived for the red band and is representative of all bands.

Table 8: Description of the sun and view angles.

Angle band	***Units***	***Data type***	***Scaling factor***	***Fill value***	***Spatial resolution***
Sun zenith	*degrees*	*uint16*	*0.01*	*40,000*	*30 m*
Sun azimuth	*degrees*	*uint16*	*0.01*	*40,000*	*30 m*
View zenith	*degrees*	*uint16*	*0.01*	*40,000*	*30 m*
View azimuth	*degrees*	*uint16*	*0.01*	*40,000*	*30 m*

6.4 Quality Assessment layer

HLS v2.0 products have one Quality Assessment (QA) layer, generated from Fmask 4.2, and named after Fmask. The Fmask integer output is converted to the bit representation (Table 9) as in HLS v1.4. The HLS processing dilates the Fmask cloud and cloud shadow by 5 pixels for L30 and S30 and labels the dilation as "Adjacent to cloud/shadow." The qualitative aerosol optical thickness level from atmospheric correction is also incorporated.

Table 9: Description of the bits in the one-byte Quality Assessment layer. Bits are listed from the MSB (bit 7) to the LSB (bit 0)

Bit number	Mask name	Bit value	Mask description
6-7	aerosol level	11	High aerosol
		10	Moderate aerosol
		01	Low aerosol
		00	Climatology aerosol
5	Water	1	Yes
		0	No
4	Snow/ice	1	Yes
		0	No
3	Cloud shadow	1	Yes
		0	No
2	Adjacent to cloud/shadow	1	Yes
		0	No
1	Cloud	1	Yes
		0	No
0	Cirrus	Reserved, but not used	NA

See Appendix A on how to decode the QA bits with simple integer arithmetic.

6.5 Metadata

Metadata about the L30 and S30 products is presented in the xmr.xml file.

6.5.1 Key metadata elements for L30 from the complete XML list include:

- LANDSAT_PRODUCT_ID
  The Landsat-8 input L1 scene product ID for processing backtracing. If two adjacent scenes from the same WRS path overlap the same MGRS tile, both product IDs are reported.
- SENSING_TIME
  The WRS scene center sensing time, carried over from the Level-1 metadata; not precisely represented the data gridded into the tile. When two scenes overlap the tile, the sensing time for each is retained.
- SPATIAL_COVERAGE
  The percentage of the tile with data
- CLOUD_COVERAGE
  The percentage of cloud and cloud shadow in observation based on Fmask
- SPATIAL_RESAMPLING_ALG
  Resampling algorithm in gridding Landsat data into the tile. Cubic convolution.
- HORIZONTAL_CS_NAME
  The map projection of the input Landsat scene or scenes. The UTM zone of the input Landsat scene may be different from that of L30
- ULX and ULY
  The UTM X/Y coordinate at the upper left corner of the tile
- ADD_OFFSET
  Value added to the spectral data before they are scaled to int16 reflectance data
- REF_SCALE_FACTOR
  Multiplier to be applied to the int16 reflectance data to get the unscaled reflectance
- THERM_SCALE_FACTOR
  Multiplier to be applied to the int16 thermal bands to get the temperature in Celsius
- ANG_SCALE_FACTOR
  Multiplier to be applied to the uint16 angle bands to get the angle in degrees
- FILLVALUE
  Pixel value in the spectral bands where no observation was made
- QA_FILLVALUE
  The QA pixel value where no observation was made
- ANG_FILLVALUE
  The angle pixel value where no observation was made
- MEAN_SUN_AZIMUTH_ANGLE
  The mean solar azimuth in the tile
- MEAN_SUN_ZENITH_ANGLE
  The mean solar zenith in the tile
- NBAR_SOLAR_ZENITH
  The solar zenith angle used in NBAR derivation.
- ACCODE
  The version of LaSRC used by HLS for L30
- TIRS_SSM_MODEL

Metadata carried over from Landsat L1 data, indicating the quality of the thermal data

- TIRS_SSM_POSITION_STATUS
  Metadata carried over from Landsat L1 data, indicating the quality of the thermal data
- IDENTIFIER_PRODUCT_DOI
  This L30 product's DOI.

6.5.2 Key metadata elements for S30 from the complete XML list include:

- PRODUCT_URI
  The input L1C granule URI, for processing backtracing
- SENSING_TIME
  Sensing time at the center of the granule, or for earlier L1C data, the time at the start of the datatake
- SPATIAL_COVERAGE
  The area percentage of the tile with data
- CLOUD_COVERAGE
  The percentage of cloud and cloud shadow in observation based on Fmask
- HORIZONTAL_CS_NAME
  The map projection of the input L1C data, same as that of S30
- ULX and ULY
  The UTM X/Y coordinate at the upper left corner of the tile
- SPATIAL_RESAMPLING_ALG
  Algorithm used in resampling 10 m/20 m/60 m data to 30 m: area weighted average. For L1C data prior to baseline 2.04, cubic convolution is used in co-registration before area weighted average
- ADD_OFFSET
  See above for L30
- REF_SCALE_FACTOR
  See above L30
- ANG_SCALE_FACTOR
  See above L30
- FILLVALUE
  See above for L30
- QA_FILLVALUE
  See above for L30
- ANG_FILLVALUE
  See above for L30
- MEAN_SUN_AZIMUTH_ANGLE
  See above for L30
- MEAN_SUN_ZENITH_ANGLE
  See above for L30
- MEAN_VIEW_AZIMUTH_ANGLE
  See above for L30
- MEAN_VIEW_ZENITH_ANGLE
  See above for L30
- NBAR_SOLAR_ZENITH
  The solar zenith angle used in NBAR derivation.
- MSI_BAND_01_BANDPASS_ADJUSTMENT_SLOPE_AND_OFFSET
  The slope and offset applied to the Sentinel-2 B01 reflectance in the linear bandpass adjustment

- MSI_BAND_02_BANDPASS_ADJUSTMENT_SLOPE_AND_OFFSET
- MSI_BAND_03_BANDPASS_ADJUSTMENT_SLOPE_AND_OFFSET
- MSI_BAND_04_BANDPASS_ADJUSTMENT_SLOPE_AND_OFFSET
- MSI_BAND_11_BANDPASS_ADJUSTMENT_SLOPE_AND_OFFSET
- MSI_BAND_12_BANDPASS_ADJUSTMENT_SLOPE_AND_OFFSET
- MSI_BAND_8A_BANDPASS_ADJUSTMENT_SLOPE_AND_OFFSET
- AROP_AVE_XSHIFT

  AROP-derived average coordinate shift in X direction relative to the reference image. Populated only for Sentinel-2 L1C data prior to processing baseline 2.04.
- AROP_AVE_YSHIFT

  AROP-derived average coordinate shift in Y direction relative to the reference image. Populated only for Sentinel-2 L1C data prior to processing baseline 2.04.
- AROP_NCP

  Number of control points identified by AROP. Populated only for Sentinel-2 L1C data prior to processing baseline 2.04.
- AROP_RMSE(METERS)

  Root mean squared error in AROP model fitting. Populated only for Sentinel-2 L1C data prior to processing baseline 2.04.
- AROP_S2_REFIMG

  Geolocation reference image name. Populated only for Sentinel-2 L1C data prior to processing baseline 2.04.
- ANGLEBAND

  For some earlier L1C data, the view angle information can be missing for some spectral bands. Therefore, a substitute band was used in NBAR derivation. This 0-based 13-element array indicates the ID of the substitute band for each band. No substitution is needed for later versions of L1C.
- ACCODE

  The version of LaSRC used by HLS for S30
- IDENTIFIER_PRODUCT_DOI

  This S30 product's DOI.

7 Known issues

The atmospheric correction over bright targets occasionally retrieves unrealistically high aerosol and thus makes the surface reflectance too low. High aerosol retrievals, both false high aerosol and realistically high aerosol, are masked when quality bits 6-7 are both set to 1 (see Table 9 in the User Guide); the corresponding spectral data should be discarded from analysis.

References

Claverie, M., Vermote, E., Franch, B., & Masek, J. (2015). Evaluation of the Landsat-5 TM and Landsat-7 ETM + surface reflectance products. *Remote Sensing of Environment*, *169*, 390-403.

Claverie, M., Ju, J., Masek, J.G., Dungan, J.L., Vermote, E.F., Roger, J.-C., Skakun, S.V., & Justice, C.O. (2018). The Harmonized Landsat and Sentinel-2 surface reflectance data set, in press, *Remote Sensing of Environment*.

Doxani, G., Vermote, E., Roger, J. C., Gascon, F., Adriaensen, S., Frantz, D., ... & Louis, J. (2018). Atmospheric correction inter-comparison exercise. *Remote Sensing*, *10*(2), 352.

Drusch, M. et al. (2012) Sentinel-2: ESA's optical high-resolution mission for GMES operational services, Remote Sensing of Environment, 120, 25-36.

ESA (2018). Sentinel-2 Data Quality Report S2-PDGS-MPC-DQR.

Franch, B., Vermote, E.F., Claverie, M., (2014a). Intercomparison of Landsat albedo retrieval techniques and evaluation against in situ measurements across the US SURFRAD network. *Remote Sensing of Environment*, 152, 627-637.

Franch, B., Vermote, E. F., Sobrino, J. A., & Julien, Y. (2014b). Retrieval of surface albedo on a daily basis: Application to MODIS data. *IEEE Transactions on Geoscience and Remote Sensing*, 52(12), 7549-7558.

Gao, F., Masek, J.G., & Wolfe, R.E. (2009). Automated registration and orthorectification package for Landsat and Landsat-like data processing. *Journal of Applied Remote Sensing, 3(1)*, 033515.

Irons, J.R., Dwyer, J.L, and J. Barsi (2012) The next Landsat satellite: The Landsat Data Continuity Mission, Remote Sensing of Environment, 122,11-21, 10.1016/j.rse.2011.08.026

Li, Z., Zhang, H.K., Roy, D.P., 2018, Investigation of Sentinel-2 bidirectional reflectance hot-spot sensing conditions, IEEE Transactions on Geoscience and Remote Sensing, 10.1109/TGRS.2018.2885967. (https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8594675)

Masek, J. G., Vermote, E. F., Saleous, N. E., Wolfe, R., Hall, F. G., Huemmrich, K. F., ... & Lim, T. K. (2006). A Landsat surface reflectance dataset for North America, 1990-2000. *IEEE Geoscience and Remote Sensing Letters*, *3*(1), 68-72.

Qiu S., Zhu Z., and He B., Fmask 4.0: Improved cloud and cloud shadow detection in Landsats 4-8 and Sentinel-2 imagery, Remote Sensing of Environment, (2019), doi.org/10.1016/j.rse.2019.05.024

Roy, D. P., Li, J., Zhang, H. K., Yan, L., Huang, H., & Li, Z. (2017). Examination of Sentinel-2A multi-spectral instrument (MSI) reflectance anisotropy and the suitability of a general method to normalize MSI reflectance to nadir BRDF adjusted reflectance. *Remote Sensing of Environment*, *199*, 25-38.

Roy, D.P., Zhang, H.K., Ju, J., Gomez-Dans, J.L., Lewis, P.E., Schaaf, C.B., Sun, Q., Li, J., Huang, H., & Kovalskyy, V. (2016). A general method to normalize Landsat reflectance data to nadir BRDF adjusted reflectance. *Remote Sensing of Environment, 176*, 255-271.

Roy, D.P., Li, Z., Zhang, H.K., 2017, Adjustment of Sentinel-2 multi-spectral instrument (MSI) red-edge band reflectance to nadir BRDF adjusted reflectance (NBAR) and quantification of red-edge band BRDF effects, Remote Sensing, 9(12), 1325. (http://www.mdpi.com/2072-4292/9/12/1325)

Schaaf, C. B., Gao, F., Strahler, A. H., Lucht, W., Li, X., Tsang, T., ... & Lewis, P. (2002). First operational BRDF, albedo nadir reflectance products from MODIS. *Remote Sensing of Environment*, *83*(1-2), 135-148.

Shuai, Y., Masek, J. G., Gao, F., & Schaaf, C. B. (2011). An algorithm for the retrieval of 30-m snow-free albedo from Landsat surface reflectance and MODIS BRDF. *Remote Sensing of Environment*, *115*(9), 2204-2216.

Skakun, S., Roger, J. C., Vermote, E. F., Masek, J. G., & Justice, C. O. (2017). Automatic sub-pixel co-registration of Landsat-8 Operational Land Imager and Sentinel-2A Multi-Spectral Instrument images using phase correlation and machine learning based mapping. *International Journal of Digital Earth*, *10*(12), 1253-1269.

Storey, J., Choate, M., & Lee, K. (2014). Landsat 8 Operational Land Imager On-Orbit Geometric Calibration and Performance. *Remote Sensing*, *6*, 11127-11152

Storey, J., Roy, D. P., Masek, J., Gascon, F., Dwyer, J., & Choate, M. (2016). A note on the temporary misregistration of Landsat-8 Operational Land Imager (OLI) and Sentinel-2 Multi Spectral Instrument (MSI) imagery. *Remote Sensing of Environment*, *186*, 121-122.

Strahler, A.H., Lucht, W., Schaaf, C.B., Tsang, T., Gao, F., Li, X., Lewis, P., & Barnsley, M. (1999). MODIS BRDF/Albedo Product: Algorithm Theoretical Basis Document Version 5.0. In M. documentation (Ed.). Boston.

Vermote, E., Justice, C. O., & Bréon, F. M. (2009). Towards a generalized approach for correction of the BRDF effect in MODIS directional reflectances. *IEEE Transactions on Geoscience and Remote Sensing*, *47*(3), 898-908.

Vermote, E., Justice, C., Claverie, M., & Franch, B. (2016). Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. *Remote Sensing of Environment*, *185*, 46-56.

Vermote, E. F., & Kotchenova, S. (2008). Atmospheric correction for the monitoring of land surfaces. *Journal of Geophysical Research: Atmospheres*, *113*(D23).

Yan, L., Roy, D. P., Li, Z., Zhang, H. K., & Huang, H. (2018). Sentinel-2A multi-temporal misregistration characterization and an orbit-based sub-pixel registration methodology. *Remote Sensing of Environment*, *215*, 495-506.

Zhang, H.K., Roy, D.P., & Kovalskyy, V. (2016). Optimal Solar Geometry Definition for Global Long-Term Landsat Time-Series Bidirectional Reflectance Normalization. *IEEE Transactions on Geoscience and Remote Sensing*, *54*, 1410-1418.

Zhu, Z., Wang, S., & Woodcock, C.E. (2015). Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4-7, 8, and Sentinel 2 images. *Remote Sensing of Environment*, *159*, 269-277.

Acknowledgment

We thank Feng Gao for providing and spending many hours adapting the AROP code for HLS with quick turnaround. We also thank Jan Dempewolf for offering his Python script which works around the GDAL-incompatible issue in HLS v1.3 and it has proved very useful for many people. We also thank Shuang Li, Min Feng, and Mark Broich for GDAL-test HLS v1.4.

Appendix A. How to decode the bit-packed QA

Quality Assessment (QA) encoded at the bit level provides concise presentation but is less convenient for users new to this format. This appendix shows how to decode the QA bits with simple integer arithmetic and no explicit bit operation at all. An analogy in the decimal system illustrates the idea. For example, given integer 3215, we want to get the digit of the hundreds place (i.e. 2). First divide the integer by 10^2 (i.e. 100) to get an integer quotient 32, then the digit of the ones place (the least significant digit) of the quotient is what we want. To get the ones digit, we compute 32 – ((32 / 10) * 10) and get 2. (Note that integer division 32/10 evaluates to 3, not 3.2.) The same idea applies to binary integers. Suppose we get a QA as a decimal number 100, which translates into binary 01100100, indicating that the aerosol level is low (bits 6-7), it is water (bit 5), and adjacent to cloud (bit 2). Suppose we want to find whether it is water, by examining the value of bit 5. It can be achieved in two steps:

- Divide 100 by 2^5 to get the quotient, 3, as a result of integer division
- Find the value of the least significant bit of the quotient by computing 3 – ((3/2) * 2), which is 1. (3 / 2 = 1 for integer division.)

The pixel is indeed water based given the decimal QA value. Note that Step 2 above is essentially an odd/even number test. When the quotient from Step one is odd, the bit in question is 1.