Recibido: 07/03/2025
Aceptado: 23/04/2025
Disponible en línea: 23/05/2025
Publicado: 31/07/2025

REVISTA DE TELEDETECCIÓN
Asociación Española de Teledetección
(2025) 66, 23510
ISSN 1133-0953
EISSN 1988-8740
https://doi.org/10.4995/raet.2025.23510

Landsat Collection 2: key information and recommendations for data users and product developers

Xavier Pons *[1, 2], Cristina Cea [1], Óscar González-Guerrero [1], Jordi Cristóbal [1]

[1] *Grumets Research Group. Departament de Geografia, Edifici B. Universitat Autònoma de Barcelona. 08193 Bellaterra, Catalonia, Spain.*

[2] *CREAF, Edifici C, Universitat Autònoma de Barcelona, 08193 Bellaterra, Catalonia, Spain.*

Abstract: The United States Geological Survey (USGS) effort to provide coherent data for the Landsat series from various perspectives (e.g., geometric, radiometric, or metadata) is, without a doubt, admirable, especially considering the vast volume of data and the continuous scientific and technical challenges over many decades. Landsat Collection 2, initiated in 2020, represents the latest effort in this direction. This paper presents a detailed explanation of some important changes compared to previous distributions. The text highlights aspects of good practices (e.g., the choice of distribution format or the explicit coding of saturated pixels), radiometric inconsistencies (e.g., in areas of scene overlap or images taken on close dates), and decisions that pose difficulties for the user community (e.g., termination of the distribution of lower processing level products,or inconsistency of NoData values). It also addresses the significant differences in radiances between data processed by ESA (CEOS) and provides justification for the decision to change the meaning of the traditional DN, resulting in shifts in radiance rescaling factors (scale and offset) throughout the year. Furthermore, the paper offers alternatives for some problematic aspects of thermal infrared data processing. Ultimately, the aim is to assist other users and contribute to the debate on best practices in remote sensing image processing.

Key words: Landsat, Collection 2, radiance, TOA reflectance, BOA reflectance.

Landsat Colección 2: Información clave y recomendaciones para usuarios de los datos y desarrolladores del producto

Resumen: El esfuerzo del Servicio Geológico de los Estados Unidos (USGS) para proporcionar datos coherentes para la serie Landsat desde varios puntos de vista (p. ej., geométrico, radiométrico o en metadatos) es, sin duda, admirable, especialmente teniendo en cuenta el enorme volumen de datos y los continuos desafíos científicos y técnicos acaecidos a lo largo de tantas décadas. La Colección 2 de Landsat, iniciada en 2020, es el último esfuerzo en esta dirección. Este artículo presenta una explicación detallada de algunos cambios importantes con respecto a distribuciones anteriores. El texto revela aspectos de buenas prácticas (p. ej., la elección del formato de distribución o la codificación explícita de píxeles saturados), de incoherencias radiométricas (p. ej., en áreas de superposición de escenas, o en imágenes de fechas cercanas) y de decisiones que presentan dificultades para la comunidad de usuarios (p. ej., la finalización de la distribución de los productos de nivel de procesamiento inferior o la inconsistencia de los valores NoData). También se cuestiona la enorme diferencia de radiancias en relación

To cite this article: Pons, X., Cea, C., González-Guerrero, O., Cristóbal, J. 2025. Landsat Collection 2: key information and recommendations for data users and product developers. *Revista de Teledetección*, *66*, e23510. https://doi.org/10.4995/raet.2025.23510

* *Corresponding author: xavier.pons@uab.cat*

con los datos procesados por la ESA (CEOS), y se aporta justificación sobre la decisión de cambiar el significado de los tradicionales DN, lo que da lugar a cambios intra-anuales en los factores de conversión de radiancia. Además, se ofrecen algunas alternativas para algunos aspectos problemáticos del procesamiento de datos en el infrarrojo térmico. En definitiva, el trabajo pretende constituir un soporte a otros usuarios, así como aportar al debate sobre las mejores prácticas en el procesamiento de imágenes de teledetección.

Palabras clave: Landsat, Colección 2, radiancia, reflectancia TOA, reflectancia BOA.

1. Introduction

Since July 23, 1972, Landsat data have been accessible through a wide variety of media storage formats (mainly CCT tapes, CD-ROM, and Internet downloads), processing types (NLAPS, LPGS), and image formats (e.g., CEOS images distributed over the years through agreements between the USGS and ESA, GeoTIFF). In recent decades, the USGS has made several efforts to provide a consistent series from various perspectives (e.g., geometric, radiometric, metadata), as distributions with significant inconsistencies have been common (e.g., Cristóbal et al., 2004; Chen et al., 2021). During these years, there have also been frequent changes in various aspects, including basic elements such as metadata formats (currently in an MTL file, which has been evolving; for further details, see USGS Landsat Metadata MTL File Changes: https://www.usgs.gov/landsat-missions/july-9-2014-landsat-metadata-mtl-file-changes), or the coding of NoData as 0 or 255 in 8-bit radiometric resolution sensors. NLAPS *versus* LPGS processing is discussed in Cristóbal et al. (2004) and at https://www.usgs.gov/landsat-missions/national-land-archive-production-system. According to this last source, NLAPS was used in 99 Landsat 4 TM scenes and in 10138 Landsat 5 TM scenes (listed at the same page) and cannot be processed and added to the Collections-based level 1 inventory. From 1999 on, NLAPS is no longer used.

Initiated in 2020, Collection 2 (Crawford et al., 2023), hereafter referred to as C2, is the second major reprocessing effort on the Landsat archive (Collection 1, C1, was developed in 2016; access to C1 ended December 30th, 2022. For further details on C1 *vs* C2, please refer to https://d9-wret.s3.us-west-2.amazonaws.com/assets/palladium/production/s3fs-public/atoms/files/Landsat-C1vsC2-2021-0430-LMWS.pdf). C2 has undergone several improvements through advancements in data processing, algorithm development, data access and distribution capabilities. Although C2 contains level 1 data from Landsat 1–9, level 2 and 3 science products are only offered for Landsat 4–9.

This article will address considerations of interest (mainly radiometric) or those that are unclear on the USGS websites (pages accessible from USGS Landsat Collection 2**:** https://www.usgs.gov/landsat-missions/landsat-collection-2) or in the recent article by Crawford et al. (2023). Although with a more focused approach on the validity of the C2 data in aquatic studies, it is necessary to indicate the complementary nature of Maciel et al. (2023) study. Given the importance of the historical Landsat series, these considerations may also be useful to the broader user and product developer communities to improve the best practices in remote sensing image processing. Although these considerations are mainly related to solar spectrum data, in some cases they also include thermal infrared data. The article is complemented with some final considerations regarding image metadata and geometry, among other topics**.**

2. Image radiometry considerations

2.1. Digital Numbers (DN) no longer have a classical meaning in all sensors

In C2, the strategy from C1 of modifying the DN originally captured by the sensor to new values has continued. These new values are referred to as Q_{cal}, or ‘Quantized and Calibrated Standard Product Pixel Values’. Previously, and still today for most sensors with similar characteristics, data captured by the sensor were scaled to DN at the sensor through a known linear function. In other words, radiances (L, typically in W/(m²·sr), or in W/(m²·sr·µm) if the spectral magnitude, L_λ, is provided) were written as integer values, using the appropriate radiometric resolution to avoid precision losses. Consequently, the inverse linear function, with slope *a* and Y-intercept *b*, could also be provided to retrieve L:

$$L = a \cdot DN + b \tag{1}$$

Usually, *a* and *b* were kept constant over time, until a further recalibration.

However, in the new Landsat collections this is no longer the case except for Landsat 1-7 level 1 (https://www.usgs.gov/landsat-missions/using-usgs-landsat-level-1-data-product): DN are rewritten to Q_{cal} to be able to provide rescaling factors from Q_{cal} to exoatmospheric reflectance, or "*Top of the Atmosphere*" (TOA) reflectance (R_{TOA} hereafter) in the level 1 product (containing Q_{cal} level 1), or at ground level, or "*Bottom of the Atmosphere*" (BOA) reflectance (R_{BOA}) in the level 2 product. Obviously, Q_{cal} level 2 values written in the level 2 product are different from the Q_{cal} values in the level 1 product.

The rescaling factors to reflectance (not to radiance), REFLECTANCE_MULT_BAND_# (the slope) and REFLECTANCE_ADD_BAND_# (the Y-intercept) (where # is the spectral band number), are adjusted to be always the same over time and even for all spectral bands in level 1 for Landsat 8-9 and for all platforms that are processed to level 2 (Landsat 1-3 are excluded: see section 3.3). It is important to note that, in the MTL files, different rescaling factors are documented with the same names, but in different "sections" (called groups): to TOA in the LEVEL1_RADIOMETRIC_RESCALING group (USGS, 2020, 2023), and to BOA in the LEVEL2_SURFACE_REFLECTANCE_PARAMETERS group (USGS, 2023).

For instance, under the section tagged as:

GROUP=LEVEL2_SURFACE_REFLECTANCE_PARAMETERS

rescaling factors are written as follows:

REFLECTANCE_MULT_BAND_1 = $2.75 \cdot 10^{-5}$

REFLECTANCE_MULT_BAND_2 = 2.75·10-5

REFLECTANCE_MULT_BAND_3 = $2.75 \cdot 10^{-5}$

REFLECTANCE_MULT_BAND_4 = $2.75 \cdot 10^{-5}$

REFLECTANCE_MULT_BAND_5 = $2.75 \cdot 10^{-5}$

REFLECTANCE_MULT_BAND_6 = $2.75 \cdot 10^{-5}$

REFLECTANCE_MULT_BAND_7 = $2.75 \cdot 10^{-5}$

REFLECTANCE_ADD_BAND_1 = -0.2

REFLECTANCE_ADD_BAND_2 = -0.2

REFLECTANCE_ADD_BAND_3 = -0.2

REFLECTANCE_ADD_BAND_4 = -0.2

REFLECTANCE_ADD_BAND_5 = -0.2

REFLECTANCE_ADD_BAND_6 = -0.2

REFLECTANCE_ADD_BAND_7 = -0.2

However, the drawback of this strategy in level 1 is that the rescaling factors to L (RADIANCE_MULT_BAND_# and RADIANCE_ADD_BAND_#) are no longer constant between dates, as they were in previous Landsat products.

To understand the meaning of this change, level 1 product is considered. Although the reasoning is analogous for the level 2 product, conceptual differences inherent to level 2 must be considered.

In Equation 1 by Pons et al. (2014), R_{TOA} can be computed in the same way as USGS applies in its product, following:

$$R_{TOA} = \pi \cdot L \cdot \frac{d^2}{E_0} \tag{2}$$

where d is the Earth-Sun distance in astronomical units and E_0 is the exoatmospheric solar irradiance, typically in W/m². Moreover, for simplification, this computation ignores the cosine of the solar zenith angle, assumes transmittances (τ) equal to 1 and discards any atmospheric correction [atmospheric radiance, L_a=0] since this is a TOA calculation.

The data thus calculated from R_{TOA} (which will theoretically be between 0 and 1, both included) are scaled between the minimum and maximum value defined by the USGS and, consequently, Q_{cal} can be obtained through a new linear transformation:

$$Q_{cal} = R_{TOA} \cdot m + n \tag{3}$$

By isolating R_{TOA} we can deduce the rescaling factors that will allow converting Q_{cal} to R_{TOA}:

$$\mathrm{REFLECTANCE\_MULT\_BAND\_\#} = \frac{1}{\boldsymbol{m}} \quad (4)$$

$$\mathrm{REFLECTANCE\_ADD\_BAND\_\#} = \frac{-\boldsymbol{n}}{\boldsymbol{m}} \quad (5)$$

From Q_{cal}, it is also possible to obtain L by calculating a new slope and a new Y-intercept. However, this calculation will be dependent on d^2, which varies for each date. Note that this dependency is not the usual practice in most remote sensing data products.

Indeed, combining the Equation 2 and Equation 3 Q_{cal} can be rewritten as:

$$Q_{cal} = R_{TOA} \cdot m + n = \left(\pi \cdot L \cdot \frac{d^2}{E_0}\right) \cdot m + n \quad (6)$$

Then L can be rewritten as:

$$L = \frac{E_0(Qcal - n)}{m \cdot \pi \cdot d^2} = \frac{E_0}{m \cdot \pi \cdot d^2} \cdot Q_{cal} - \frac{n}{m \cdot \pi \cdot d^2} = \left(\frac{1}{d^2}\right) \cdot \frac{E_0}{m \cdot \pi} \cdot Q_{cal} - \left(\frac{1}{d^2}\right) \cdot \frac{n}{m \cdot \pi} \quad (7)$$

And finally, the rescaling factors can be computed as follows:

$$\mathrm{RADIANCE\_MULT\_BAND\_\#} = \left(\frac{1}{d^2}\right) \cdot \frac{E_0}{m \cdot \pi} \quad (8)$$

$$\mathrm{RADIANCE\_ADD\_BAND\_\#} = \left(\frac{1}{d^2}\right) \cdot \frac{n}{m \cdot \pi} \quad (9)$$

It can be observed that both the slope and the Y-intercept for obtaining radiances exhibit an annual cyclic shift depending on the date, as the Earth-Sun distance squared (d^2) changes for each date. Figure 1 illustrates this behavior, reflecting the well-known variation of the square of the Earth-Sun distance by approximately ±3.3% throughout the annual cycle.

The existence of a simple transformation using the REFLECTANCE_MULT_BAND and REFLECTANCE_ADD_BAND parameters to obtain R_{TOA} from Q_{cal} values may seem practical. However, users must be aware that the same perfectly invariant surface under identical atmospheric conditions on

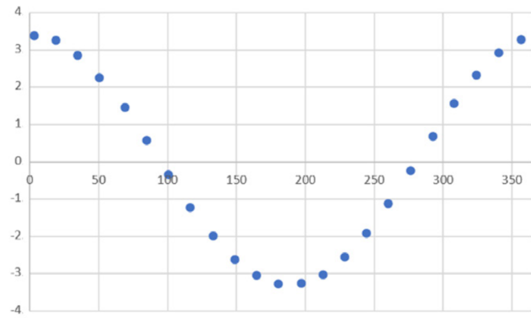

Figure 1. Slope rescaling factor annual dynamics (in %) for L conversion. Example extracted from MTL files of band 4 in 24 images between Jan 3rd and Dec 21st of 2022 in scene 197-031.

different dates will yield different R_{TOA} values. This is because the coefficients provided in the MTL files do not account for the Sun angle, as explained in Equation 2 and stated in the USGS Landsat Level-1 Data Product documentation (https://www.usgs.gov/landsat-missions/using-usgs-landsat-level-1-data-product). Remember that these constant values are not available in Landsat 1-7 level 1 and a classical processing must be applied. Regarding level 2, BOA coefficients are constant in all spectral bands in TM, ETM+ and OLI sensors.

2.2. Inconsistent treatment in overlapping areas between distribution units

When analyzing an overlapped area in consecutive image pairs, it is common to find pixels that, with the same geographic position, exhibit different radiometric values between the two scenes. This inconsistency is already present in the level 1 product and will propagate to all subsequent products (level 2 and beyond). For instance, on June 10th, 2004, differences in pixels of interest from a Landsat 5 TM image of path 197 and rows 030 and 031 (WRS-2) showed discrepancies greater than 1 Q_{cal} in multiple spectral regions simultaneously. As examples, this pattern was found in the following pixels (longitude, latitude): snow (2.468796, 42.510419), cloud (2.519669, 42.495735), water (2.102762, 42.571561), coniferous vegetation (2.550657, 42.499102) and urban area (2.873622, 42.416712). Another example can be found in Landsat 9 OLI-TIRS at

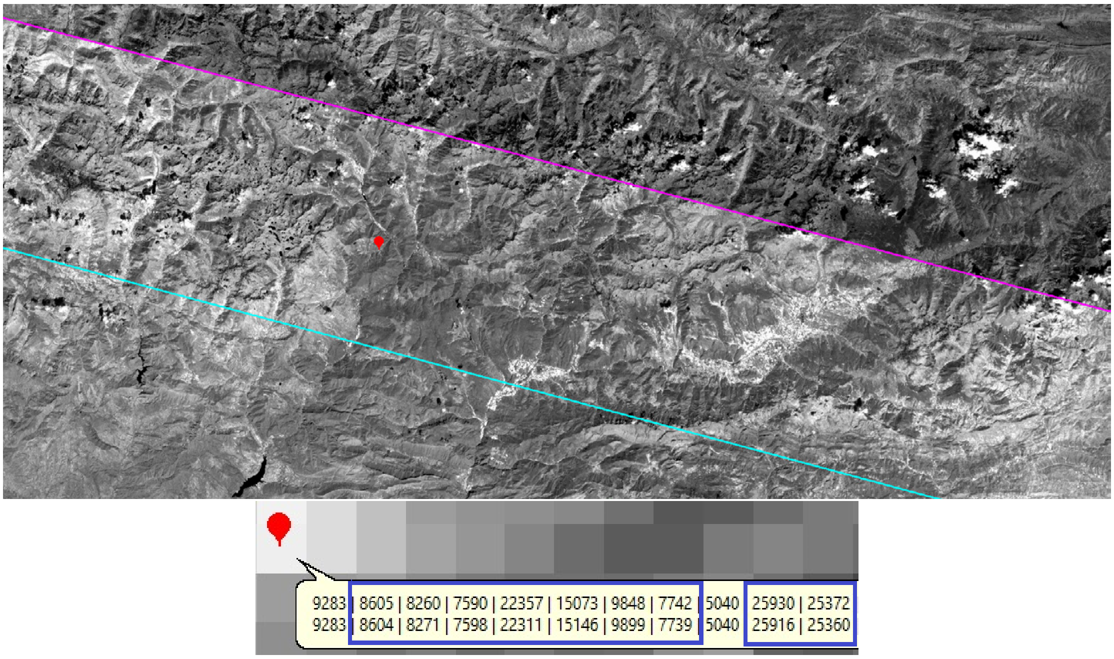

Figure 2. Upper panel: Overlapping area between two consecutive scenes (path 198 and rows 031 and 032, WRS-2) of a Landsat 9 OLI-TIRS Collection 2 image from August 30th, 2022 and a point of interest. Lower panel: Point of interest values, in the level 1 product, highlighting (in deep blue) the differences in most bands of the two scenes; values are in Qcal. Note that the point of interest is not at any image border (the lower border of the upper scene is the blue line; the upper border of the lower scene is the fuchsia line). Band values are presented in spectral order, left to right: 1-Coastal/Aerosols, 2-Blue, 3-Green, 4-Red, 5-Near Infrared, 6- Shortwave Infrared 1, 7- Shortwave Infrared 2, 8- Panchromatic, 9-Cirrus, 10- Longwave Infrared 1, 11- Longwave Infrared 2. The upper record corresponds to the 198 031 scene, while the lower record to the 198 030 scene.

(1.165870, 42.510689) in Figure 2, in this case in path 198 and the same rows than the previous examples.

As no source of evidence to support the usage of the upper or lower scene is known from the authors of this paper, our suggestion is to produce the mosaic by averaging the pairs of values. This can be done, for example, using a layer calculator producing NoData when operating two cells if one of them, or both, are NoData, and having a special cover function, CVR(arg1, arg2) that properly manages NoData and data values in that way: the cell from arg1 is selected, unless that cell is NoData, then the cell from arg2 is selected. Through the expression CVR(([A]+[B])/2, CVR([A],[B])), the resulting cell will be the average if it can be computed, or, on the contrary the cell of the first or second argument (NoData is only obtained if the cell is NoData in both layers). This can be done, for example with the MiraMon CalcImg module (https://www.miramon.cat/help/eng/msa/CalcImg.htm) that also has a more high-level function CVRM(arg1, arg2), producing the same result with the more compact syntax CVRM([A], [B]).

2.3. Acceptable reflectance discrepancies at level 1 and excessive at level 2

When analyzing pixels from the same area captured a few days apart within the same year, and with no evidence of sudden changes in cover, it is expected that discrepancies should only be due to atmospheric causes, given that the illumination and observation angles are highly similar. For instance, when examining a pseudo-invariant fully concrete pixel in a flat area from an image of path 197 and row 031 (longitude: 2.226218, latitude: 41.410353), pixel differences reach 1.9 units of % R_{TOA} in the Near Infrared (NIR) between December 5th and January 3rd. This might be acceptable due to their TOA nature. However, this difference reaches 6.2 units of % R_{BOA} in the level 2 product, which is unexpected. Furthermore, when calculating the extreme differences in R_{TOA} and R_{BOA} for this pseudo-invariant pixel in the Blue, Green, Red, NIR, and Shortwave Infrared (SWIR) bands of 1.6 and 2.2 µm throughout the year, differences up to 16.1% (SWIR-2) in R_{TOA}, with an overall average of 13.0% across all bands, and up to 6.7% in NIR for R_{BOA}, with an overall average of 4.7%, can be observed (Table 1). Therefore, these discrepancies, along with those noted in subsection 2.6, suggest the potential benefits of employing atmospheric adjustment methods based on pseudo-invariant zones (Pons et al., 2014) or representative field radiometry data obtained through the use of UAV (Padró et al., 2018), as alternatives to the official product.

Table 1. Variation Range of R_{TOA} and R_{BOA} for a pseudo-invariant pixel (concrete) at three different times in the same year (longitude: 2.226218, latitude: 41.410353, scene 197 031). Data in % reflectance, corresponding to the point shown in Figure 3.

	TOA	BOA
B	10.2	2.8
G	10.7	3.9
R	11.4	5.0
NIR	14.5	6.7
SWIR1	15.3	5.6
SWIR2	16.1	4.3
Mean:	13.0	4.7
Median:	12.9	4.7

2.4. Large differences with ESA-CEOS radiometric values in some sensors

The task of radiometric homogenization between Landsat sensors that has been carried out up to the C2 is very interesting. However, when comparing the L values of identical images (same dates) from ESA-CEOS products, one can observe significant discrepancies in images from some instruments (Table 2). The pertinent question is whether these differences are justified by the temporal homogenization performed by the USGS. The largest differences were found in TM, while in ETM+ they were small. In our exploration, the coefficients in ETM+ were very similar or identical, while in Landsat 5 TM they were different. In ETM+, CEOS DN values were practically identical to C2 Q_{cal}, but this was not the case in TM scenes. Since the L_λ differences were almost always positive, it implies a systematic bias in which the C2 values are almost always higher than those of ESA-CEOS. While addressing this question is beyond the scope of this article, it is essential to highlight this potential issue to the Landsat user and producer communities.

2.5. Problems in estimating surface temperature

Estimating surface temperature using the new image processing techniques in Landsat Collection 2 faces several challenges and

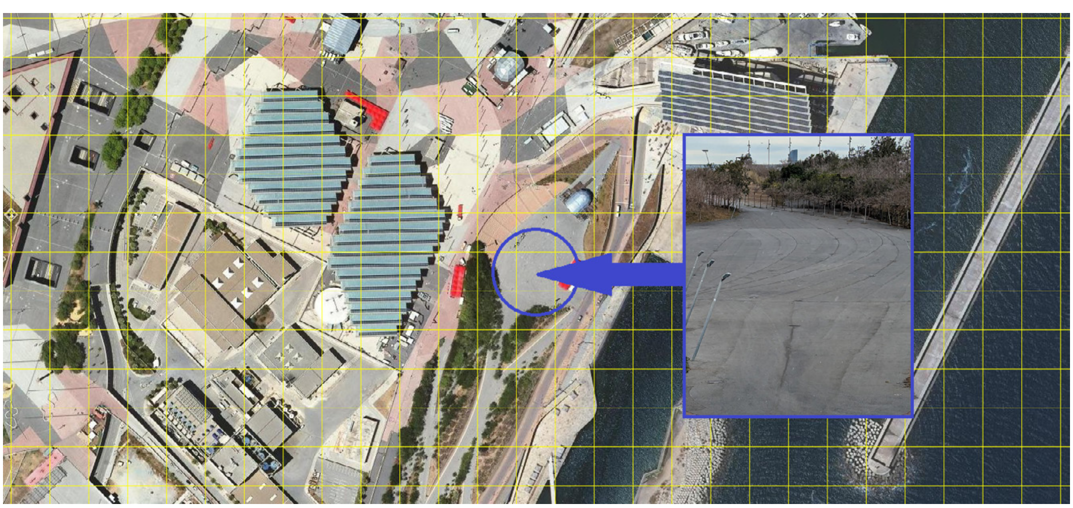

Figure 3. Full concrete pseudo-invariant area used to compute the results shown in Table 1. Orthophoto: Institut Cartogràfic i Geològic de Catalunya. Field photo: X. Pons.

Table 2. Medians of L_λ (units: W/(m²·sr·µm)) in the Blue, Green, Red, Near Infrared, and Shortwave Infrared bands (1.6 µm and 2.2 µm) for a large area of 7526 km² in scene 197 031 on the four specified dates (yyyymmdd) for level 1 products. Values are provided for ESA (CEOS) and Earth Explorer (EE) images. The differences between them and the percentage differences relative to ESA (CEOS) are also included.

	20040219 (Landsat 5 TM)				20040728 (Landsat 5 TM)			
	CEOS	EE	EE-CEOS	%vsCEOS	CEOS	EE	EE-CEOS	%vsCEOS
B	28.62	38.30	9.69	33.84	42.47	57.45	14.98	35.26
G	20.70	27.57	6.87	33.21	33.62	43.50	9.88	29.39
R	14.92	19.71	4.79	32.13	23.78	31.19	7.41	31.16
NIR	29.45	32.65	3.20	10.88	62.03	68.57	6.54	10.54
SWIR1	3.68	4.32	0.65	17.63	7.76	9.14	1.38	17.78
SWIR2	0.75	0.90	0.15	20.31	1.54	1.82	0.28	18.14

	20020426 (Landsat 7 ETM+)				20040610 (Landsat 5 TM)			
	CEOS	EE	EE-CEOS	%vsCEOS	CEOS	EE	EE-CEOS	%vsCEOS
B	61.55	62.33	0.78	1.27	38.86	52.09	13.23	34.05
G	47.92	48.72	0.80	1.67	30.10	40.61	10.51	34.91
R	34.16	34.79	0.62	1.82	19.75	27.02	7.26	36.78
NIR	63.72	64.69	0.97	1.52	67.73	74.70	6.97	10.29
SWIR1	8.57	8.72	0.14	1.68	7.54	8.66	1.11	14.75
SWIR2	1.71	1.71	0.00	0.17	1.31	1.55	0.24	18.49

limitations. Currently, data from the ASTER Global Emissivity Dataset (GED) are used to estimate the emissivity required for calculating surface temperature. However, various issues in its accurate estimation, such as blockiness artifacts due to nearest neighbour resampling in the GED product, missing time series data, or vegetation emissivity anomalies, result in the failure to estimate ST for certain pixels or images in the C2 ST product. Additionally, the surface temperature products may exhibit a “blocky” appearance over small surface targets with contrasting thermal infrared signatures. This artifact arises from the nearest neighbour interpolation of the 100-meter ASTER GED products to the finer 30-meter Landsat grid. Another challenge is the geometric mis-registration between the ASTER NDVI and Landsat data, which can cause visual artifacts in the surface temperature products, particularly in regions with significant thermal contrast. Furthermore, the ASTER GEDv3 dataset used for emissivity is static and represents average emissivity from 2000-2008, necessitating adjustments for current land surface conditions at the time of the Landsat overpass. These adjustments can introduce inaccuracies, making it difficult to achieve precise surface temperature estimates. To address these issues, alternative algorithms can be employed to calculate vegetation or bare soil emissivity from Landsat solar spectrum data, such as the method proposed by Sobrino et al. (2008). These and other problems are documented at USGS Landsat Collection 2 Known Issues. These limitations underscore the complexities involved in accurately estimating surface temperature with the new image processing methods in Landsat Collection 2 (https://www.usgs.gov/landsat-missions/landsat-collection-2-known-issues#ST).

2.6. Additional details

As a new feature in C2, it is noteworthy that two layers are provided to indicate saturated pixels, which is beneficial for excluding them in precise radiometric approximations. Additionally, several issues related to the radiometric correction algorithm are documented at USGS Landsat Collection 2 Known Issues (https://www.usgs.gov/landsat-missions/landsat-collection-2-known-issues#SR), although a detailed discussion of these issues is beyond the scope of this brief article. It is, however, disappointing that, as acknowledged in the referenced source, the coding of NoData values is still conflated with zero radiometric values, which is unrelated to the inherent challenges of radiometric correction in certain situations. Nevertheless, the wide unsigned integer data type [0, 65535] used for distributing the imagery allows for a well-defined flag value (e.g., 65535), which is distinct from reasonable radiometric values.

3. Other considerations

3.1. Limitations in distribution options: consequences in geometry and radiometry

USGS distribution platforms are excellent for discovering and downloading data, but they have ceased offering the classic products that are closest to raw data. While it is true that Landsat users rarely had access to completely raw data, less processed products allowed for improvements in georeferencing (using finer Digital Elevation Models, DEM) and, importantly, avoided the use of bicubic interpolation as the resampling method. This method is notably the least respectful of the original radiometry, which is particularly problematic in categorical thematic mapping applications. Mixing the radiometry of 16 pixels, even when done in a weighted manner, creates a suboptimal situation for classifiers, especially in areas with high radiometric contrast (e.g., sandy beaches and water).

On other occasions, very poorly georeferenced images are encountered (e.g., LT05_L1GS_198031_20060927_20200831_02_T2 in the Iberian Peninsula, as well as numerous images of Greenland, due to the poor quality of the DEM used). These images are labelled as '*tier 2*' as a warning. However, in some cases, such as the aforementioned image, proper correction is entirely feasible, as demonstrated by previous processing carried out by the authors of this paper. Unfortunately, this is no longer possible as the original image is no longer available. While it is understood that the USGS cannot geometrically correct each image with the utmost precision, the current policy prevents users who are capable of doing so from achieving a higher level of accuracy in specific areas.

Additionally, in some regions, users have access to DEM and/or Digital Surface Models (DSM) that are superior to those used in C2. This is particularly critical for radiometric corrections: computing the cosine of the incident angle benefits from more detailed surface modelling, and using the same DEM/DSM for both geometric and radiometric corrections is crucial to avoid overcorrections and undercorrections in ridges and valleys.

Finally, in countries where the UTM projection is not the official one, or where a datum significantly different from WGS84 is used (and is relevant for a spatial resolution such as that of Landsat), image resampling can cause undesirable issues. This criticism is, in fact, acknowledged by the USGS: regarding the level 2 products of most of the United States, which are in the Albers conic projection, the USGS states: '*Raw Landsat data are calibrated and reprojected directly into the U.S. ARD Albers projection, which means the U.S. ARD are not resampled twice'* (https://earth.esa.int/eogateway/documents/20142/0/Landsat-Collection-1-vs-Collection-2-Summary.pdf).

3.2. Quality Assessment bands

Numerous Quality Assessment (QA) metadata are still provided in C2 in image (band) format. However, the format remains non-trivial for many users due to the mixture of bits with different meanings within the same byte and the fact that some properties require the consideration of more than one bit at a time. Consequently, some software has developed *ad hoc* modules (e.g., ArcGIS: https://www.usgs.gov/landsat-missions/landsat-quality-assessment-arcgis-toolbox) or have separated specific bands for each quality assessment (e.g., MiraMon: Pons, 2025). These bands support significant compression, so distributing them in more user-friendly raster values would benefit the user community. It would also be preferable to codify pixels where the QA flags have no meaning with an explicit NoData value (e.g., all bits set to one).

Another issue relates to bands containing angles (sensor or Sun) (USGS 2022, p. 53). Problems were reported in Landsat Sensor Viewing Angle Bands Known Issues (https://www.usgs.gov/landsat-missions/landsat-sensor-viewing-angle-bands-known-issues), which have fortunately been resolved. However, it would be better to avoid coding NoData angles with 0, because this is a valid angle value, and to avoid coding angle values outside the sensed image area. Indeed, profile figures on the USGS Solar Illumination and Sensor Viewing Angle Coefficient Files webpage (https://www.usgs.gov/landsat-missions/solar-illumination-and-sensor-viewing-angle-coefficient-files) show these 0 values. Some software (e.g., MiraMon:

Pons, 2025) recodes cells affected by NoData (deduced from other bands) to an appropriate value (e.g., –32768, as the angle files have units of hundredths of degrees, positive or negative) and explicitly documents this in the user metadata file. While angle files are included in C2 products, they are only for band 4 (near the center of the OLI focal plane). The angle tool has options to create angle bands for every band in the product, including the TIRS bands. The Solar Illumination and Sensor Viewing Angle Generation for Landsat 4-7 can be obtained at https://landsat.usgs.gov/sites/default/files/documents/LANDSAT_ANGLES_15_3_0.tgz. Detailed instructions, including how to obtain the source code are at https://d9-wret.s3.us-west-2.amazonaws.com/assets/palladium/production/s3fs-public/atoms/files/L4-7_AngleBands_README-20200520.txt.

For Landsat 8/9 the tool is at https://landsat.usgs.gov/sites/default/files/documents/L8_ANGLES_2_7_0.tgz and the instructions and code at https://www.usgs.gov/media/files/landsat-8-angles-creation-tools-readme. Landsat TM and ETM+ Solar and View Angle Generation Algorithm Description Document (ADD) is available at https://d9-wret.s3.us-west-2.amazonaws.com/assets/palladium/production/s3fs-public/atoms/files/LSDS-1691_L4-5TM-L7-ETM_Solar-View-Angle-Generation_ADD-v1.pdf. Landsat 8 Solar and View Angle Generation Algorithm Description Document (ADD) is available at https://d9-wret.s3.us-west-2.amazonaws.com/assets/palladium/production/s3fs-public/atoms/files/LSDS-1928_L8-OLI-TIRS_Solar-View-Angle-Generation_ADD-v2.pdf.

3.3. Incomplete time series at level 2

Level 2 products are not available for Landsat 1-3 (USGS Landsat Collection, https://www.usgs.gov/landsat-missions/landsat-collection-2, and Crawford et al., 2023), resulting in a significant loss of information in the time series (lack of temporal completeness) during the years between Landsat 1 (1972) and Landsat 4 (1982).

3.4. Direct COG for geoservices

C2 is distributed in Cloud-Optimized GeoTIFF (COG) format, which is a significant advancement. This recent Open Geospatial Consortium (OGC) specification (Masó, 2023) enables images to be accessed directly from servers without the need for prior preparation, while still allowing the same files to be read efficiently from local environments.

3.5. Metadata leading to confusion

In the level 2 product, rescaling factors for L and R_{TOA} are included (see section 2.1). Upon considering their meaning, it can be deduced that they are not applicable to the image they document. Given the non-reversibility of the process of writing the new Q_{cal} proportional to the R_{BOA} using simple coefficients, their inclusion seems highly confusing and prone to errors, making them unsuitable for instructional purposes. For instance, it is not easy to understand why, in the same file, there are two values for the same key (e.g., REFLECTANCE_MULT_BAND_4=), or a value (e.g., RADIANCE_MULT_BAND_4=) that should not be applied to that product.

In other words, the appropriateness of the "GROUP=LEVEL1…" in level 2 metadata files and the avoidance of the inclusion of radiance and reflectance rescaling factors, as they cannot be applied, could be reconsidered. Although the USGS (2023) states that '*these fields are for provenance*', the risk of confusion is significant

4. Conclusions

After examining the characteristics and example data of Landsat Collection 2, this paper highlights several important changes compared to previous distributions. While some of these changes represent improvements and good practices, others introduce radiometric inconsistencies, alter the usual interpretation of the image data, or require especial attention from the user community. Some recommendations are also given, such as providing imagery without geometric correction to allow the use of local, more precise, DEM and avoiding cubic convolution, which would result in more accurate radiometric corrections, or the possibility of directly adopting other datum and/or cartographic projections.

5. Acknowledgements

The authors would like to thank the USGS for its open data policy, as well as the effort to continue improving the Landsat time series of remote sensing images. This work has been carried out in part thanks to the SGR grant from Consolidated and Quality Research Groups of the Generalitat de Catalunya (SGR 2021 no. 00554) and by the Spanish MCIU Ministry DynaFun project (PID2023-152719OB-C21 MCIU/AEI/ERDF,EU). Xavier Pons is the recipient of an ICREA Academia Excellence in Research Grant (2023–2027).

References

Chen, F., Wang, C., Zhang, Y., Yi, Z., Fan, Q., Liu, L., Song, Y. 2021. Inconsistency among Landsat Sensors in Land Surface Mapping: A Comprehensive Investigation Based on Simulation. *Remote Sensing*. *13*(7), 1383. https://doi.org/10.3390/rs13071383

Crawford, C.J., Roy, .D.P., [32 additional authors] and Zahn, S. 2023. The 50-Year Landsat Collection 2 Archive. *Science of Remote Sensing*, *8*, 100103. https://doi.org/10.1016/j.srs.2023.100103

Cristóbal, J., Pons, X., Serra, P. 2004. Sobre el uso operativo de Landsat-7 ETM+ en Europa. *Revista de Teledetección*, *21*, 55-59. https://www.aet.org.es/revistas/revista21/AET21-11.pdf

Maciel, D.A., Pahlevan, N., Barbosa, C.C.F., de Novo, E.M.L.d.M., Paulino, R.S., Martins, V.S., Vermote, E., Crawford, C.J. 2023. Validity of the Landsat surface reflectance archive for aquatic science: Implications for cloud-based analysis. *Limnology and Oceanography Letters*, *8*, 850-858. https://doi.org/10.1002/lol2.10344

Masó, J. 2023. OGC Cloud Optimized GeoTIFF Standard. Open Geospatial Consortium. https://www.opengis.net/doc/is/COG/1.0

Padró, J.C., Muñoz, F.J., Ávila, L.A., Pesquer, L., Pons, X. 2018. Radiometric Correction of Landsat-8 and Sentinel-2A Scenes Using Drone Imagery in Synergy with Field Spectroradiometry. *Remote Sensing*, *10*, 1687. https://doi.org/10.3390/rs10111687

Pons, X., Pesquer, L., Cristóbal, J., González-Guerrero, O. 2014. Automatic and improved radiometric correction of Landsat imagery using reference values from MODIS surface reflectance images. *International Journal of Applied Earth Observation and Geoinformation*, *33*, 243-254. https://doi.org/10.1016/j.jag.2014.06.002

Pons, X. 2025. MiraMon. Geographic Information System and Remote Sensing software. Bellaterra: CREAF. https://www.miramon.cat/Index_eng.htm

Sobrino, J.A., Jiménez-Muñoz, J.C., Soria, G., Romaguera, M., Guanter, L., Moreno, J. 2008. Land Surface Emissivity Retrieval From Different VNIR and TIR Sensors. *IEEE Transactions on Geoscience and Remote Sensing*, *46*(2), 316-327. https://doi.org/10.1109 /TGRS.2007.904834

USGS. 2020. Landsat 8-9 OLI/TIRS Collection 2 Level 1 Data Format Control Book. Retrieved March 5, 2025, from https://www.usgs.gov/media/files/landsat-8-9-olitirs-collection-2-level-1-data-format-control-book

USGS. 2021. Landsat 8-9 Calibration and Validation (Cal/Val) Algorithm Description Document (ADD). Retrieved March 5, 2025, from https://d9-wret.s3.us-west-2.amazonaws.com/assets/palladium/production/s3fs-public/atoms/files/LSDS-1747_Landsat8-9_CalVal_ADD-v4.pdf

USGS. 2022. Landsat 9 Data Users Handbook. Retrieved March 5, 2025, from https://www.usgs.gov/media/files/landsat-9-data-users-handbook

USGS. 2023. Landsat 8-9 OLI/TIRS Collection 2 Level 2 Data Format Control Book. Retrieved March 5, 2025, from https://www.usgs.gov/media/files/landsat-8-9-olitirs-collection-2-level-2-data-format-control-book