This document is supplementary information to travel with the file Landsat\_C2\_GSL\_collated\_DSWE1\_center\_meta\_v2024-08-14.csv. This file was created in a copy of the template repository [ROSSyndiacate/Landsat\_C2\_SRST\_template](https://github.com/rossyndicate/Landsat_C2_SRST_template) currently located [here](https://github.com/steeleb/Landsat_C2_SRST_GSL). Please reach out to B Steele ([b.steele@colostate.edu](mailto:b.steele@colostate.edu)) with questions.

The Landsat pull was set up to obtain all available Landsat data from the Landsat archive from missions 4, 5, 7, 8, 9 using the polygons provided (“Shapefile HM.shp”) to define the ‘center’ point from which to extract surface reflectance and surface temperature data. I gave the 4 areas a row id (“r\_id”), see the .shp file in the repo data\_in/GSL\_dropone.shp – there was an empty polygon that I had to remove, hence the name. The center locations can be found in data\_acquisition/out/user\_polygon\_centers.csv. Buffer distance was set to include all pixels within a 120 m radius when pixels were ‘confident water’ (DSWE1, [Jones, 2019](https://doi.org/10.3390/rs11040374)). This workflow is based on the workflow established for Collection 1, but with more rigorous masking for clouds of all types, fill values, unrealistic values, and data quality.

For the purpose of this project, I have applied correction coefficients to the band data from John Gardener’s lab (using the method described by [Roy, et al 2014](https://www.sciencedirect.com/science/article/pii/S003442571400042X)), which are subject to change and at this point are not citable. The surface temperature data have not been bias-adjusted, so if you want to use those, know that you may want to do some investigation for how they line up with any *in situ* data before using. [Herrick and Steele, et al 2023](https://doi.org/10.1002/ecs2.4357) can give you some additional insight on how to use those data. You could probably use them for trend analysis, just know that the absolute value shown is not precise.

These data have not been thoroughly checked for data consistency. While the masking process removes most errant data, you may need to apply some additional QA measures to assure the data is good to go for analysis. ([This script has some suggestions for QA](https://github.com/steeleb/Landsat_C2_SR_GLORIA/blob/main/c_data_download_collation/src/baseline_QAQC_RS_data.R).) We have scene viewers in the GEE code GUI if you notice strange values so that you can see the image they are taken from. These are still under development, but should give you a good sense of whether there are wonky data points due to contamination within the image. GSL happens to be one of the areas that we list as examples, so checking out images should be extra easy.

[Landsat 5/7 image viewer](https://code.earthengine.google.com/77c02954d9e54327abf797e5398b3f52?hideCode=true)

Landsat 8/9 image viewer

Please follow up if you intend to publish any resultant analyses so we can provide you with proper DOI’s and/or citations for this workflow including John’s DOI for the correction coefficients.

Column definitions/units

Column prefixes

|  |  |  |
| --- | --- | --- |
| **Column prefix** | **Definition** | **Data Source** |
| med\_\* | Median value within defined buffer | GEE workflow |
| min\_\* | Minimum value within defined buffer | GEE workflow |
| max\_\* | Maximum value within defined buffer | GEE workflow |
| sd\_\* | Standard deviation of the mean within defined buffer | GEE workflow |
| mean\_\* | Mean value within defined buffer | GEE workflow |
| kurt\_\* | Kurtosis of the distribution within defined buffer | GEE workflow |

Optical Bands

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Column name, prefix, or suffix** | **definition** | **Inter-mission handoff applied** | **Applicable mission** | **Unit** | **Data Source** | **Use Case** |
| \*\_Blue | Aggregated statistic for blue band reflectance | no | all | unitless | GEE workflow | Algorithm development, QA |
| \*\_Green | Aggregated statistic for green band reflectance | no | all | unitless | GEE workflow | Algorithm development, QA |
| \*\_Red | Aggregated statistic for  red band reflectance | no | all | unitless | GEE workflow | Algorithm development, QA |
| \*\_Nir | Aggregated statistic for near infrared band reflectance | no | all | unitless | GEE workflow | Algorithm development, QA |
| \*\_Swir1 | Aggregated statistic for shortwave infrared 1 band reflectance | no | all | unitless | GEE workflow | Algorithm development, QA |
| \*\_Swir2 | Aggregated statistic for shortwave infrared 2 band reflectance | no | all | unitless | GEE workflow | Algorithm development, QA |
| \*\_Aerosol | Aggregated statistic for aerosol band reflectance | no | LS 8/9 only | unitless | GEE workflow | Algorithm development, QA |
| \*\_corr7 | Optical band value with LS7 handoff coefficient applied, see associated flag columns for data use | Yes, correction relative to LS7 | LS 5-9, LS8 correction applied to LS9 | unitless | “NW\_CLP\_RS” workflow | Algorithm\_Development, QA |

Optical Band Correction Flags

|  |  |  |
| --- | --- | --- |
| **Column name, prefix, or suffix** | **definition** | **Data Source** |
| flag\_\*\_7 | Indication of whether or not the band value was beyond the Rrs range used in the corrections to relative Landsat 7 values | “NW\_CLP\_RS” workflow |
| flag\_\*\_8 | Indication of whether or not the band value was beyond the Rrs range used in the corrections to relative Landsat 8 values | “NW\_CLP\_RS” workflow |

Thermal Band

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Column name, prefix, or suffix** | **definition** | **Inter-mission handoff applied** | **Applicable mission** | **Unit** | **Data Source** | **Use Case** |
| \*\_SurfaceTemp | Aggregated surface temperature, note prefix for summary statistic | N/A | all | Degrees Kelvin | GEE workflow | Algorithm development, QA |
| \*\_temp\_qa | Uncertainty of the surface temperature estimate | N/A | all | 1 = 0.01 Degrees Kelvin | GEE workflow | QA |

Pixel-level Metadata

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Column name, prefix, or suffix** | **definition** | **Applicable mission** | **Unit** | **Use Case** |
| prop\_clouds | proportion pixels with clouds within defined buffer, these have been masked, but this is a useful measure for knowing how contaminated the neighboring pixels are with clouds | all | Proportion of 30x30m pixels | QA |
| prop\_hillShadow | proportion of pixels NOT shadowed by surrounding topography included in aggregation – only unshadowed pixels are included in the data summaries | all | Proportion of 30x30m pixels | QA |
| mean\_hillShade | average hillshade value (0-364) within specified area included in aggregation | all | degree | QA |
| pCount\_dswe1 | number of pixels in a 120m radius that are considered 'confident' open water | all | count of 30x30m pixels | QA |
| pCount\_dswe1a | number of pixels in a 120m radius that are considered ‘confident’ open water or that meet the threshold for algal response | all | count of 30x30m pixels | QA |
| pCount\_dswe3 | number of pixels in a 120m radius that are considered confident ‘vegetated water’ | all | count of 30x30m pixels | QA |
| pCount\_gt0 | number of pixels in a 120m radius that are any DSWE value greater than 0 (water of any type) | all | Count of 30x30m pixels | QA |
| \*\_cloud\_dist | distance to cloud (for any pixel in the 90m radius) | all | 1 = 0.01 km | QA |
| prop\_medHighAero | proportion of pixels within defined buffer that have medium of high aerosol applied in the LaSRC atmospheric correction | LS 8/9 only | proportion of 30x30m pixels | QA |

Scene-level Metadata

| **Column name, prefix, or suffix** | **definition** | **Applicable mission** | **Unit** | **Use Case** |
| --- | --- | --- | --- | --- |
| system.index | Landsat-specific unique system identifier for the scene | all | N/A | N/A |
| WRS\_PATH | WRS2 path of the image | all | N/A | N/A |
| WRS\_ROW | WRS2 row of the image | all | N/A | N/A |
| mission | Spacecraft mission identifier | all | N/A | N/A |
| date | Date of image acquisition | all | N/A | Matching |
| UTC\_time | Time of image acquisition | all | Time in UTC | Matching |
| CLOUD\_COVER | percent of scene with cloud cover | all | percent | QA |
| IMAGE\_QUALITY | scale of 0-9 quality of image for reflectance | all (this was originally IMAGE\_QUALITY\_OLI for 8/9) | unitless | QA |
| IMAGE\_QUALITY\_TIRS | scale of 0-9 quality of image for thermal data |  | unitless | QA |
| SUN\_AZIMUTH | Sun azimuth angle in degrees for the image center location at the image center acquisition time. A positive value indicates angles to the east or clockwise from the north. A negative value indicates angles to the west or counterclockwise from the north. |  | degrees | QA |
| SUN\_ELEVATION | Sun elevation angle in degrees for the image center location at the image center acquisition time. A positive value indicates a daytime scene. A negative value indicates a nighttime scene. Note: For reflectance calculation, the sun zenith angle is needed, which is 90 - sun elevation angle. |  | degrees | QA |

Location/Data Handling Metadata

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Column name, prefix, or suffix** | **definition** | **Applicable mission** | **Unit** | **Use Case** |
| r\_id | rowid for GEE process, artifact of processing | all | N/A | Matching with location information |
| source | Original pull .csv exported from GEE | All | N/A | QA |
| DSWE | DSWE mask used for GEE analysis | all | 1 (confident open water) or 3 (confident vegetated water) | Upstream data subsetting |
| Latitude | Latitude of point for GEE analysis | all | Decimal degree, WGS84 | N/A |
| Longitude | Longitude of point for GEE analysis | all | Decimal degree, WGS84 | N/A |
| dist | Calculated distance of the center point from shore | All |  | QA |