JPL D-96967

# ECOsystem Spaceborne Thermal Radiometer Experiment on Space Station (ECOSTRESS)

# Level-1B Rad PGE Algorithm Specification Document

Mike Smyth

ECOSTRESS Software Team

June 11, 2018

**National Aeronautics and Space Administration**

**Jet Propulsion Laboratory**
**California Institute of Technology**
**Pasadena, California**

This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

# Contacts

Readers seeking additional information about this document may contact the following ECOSTRESS Science Team members:


- Mike Smyth
  MS 169-321
  Jet Propulsion Laboratory
  4800 Oak Grove Drive
  Pasadena, CA 91109
  Email: Mike.M.Smyth@jpl.nasa.gov
  Office: (818) 354-9812

- Jordan Padams
  MS 321-351
  Jet Propulsion Laboratory
  4800 Oak Grove Drive
  Pasadena, CA 91109
  Email: Jordan.H.Padams@jpl.nasa.gov
  Office: (818) 354-3130

# List of Acronyms

| | |
|---|---|
| ALEXI | Atmosphere–Land Exchange Inverse |
| ARS | Agricultural Research Service |
| ATBD | Algorithm Theoretical Basis Document |
| Cal/Val | Calibration and Validation |
| CDL | Cropland Data Layer |
| CFSR | Climate Forecast System Reanalysis |
| CONUS | Contiguous United States |
| DisALEXI | Disaggregated ALEXI algorithm |
| DPU-IO | Digital Processing Unit Input/Output |
| ECOSTRESS | ECOsystem Spaceborne Thermal Radiometer Experiment on Space Station |
| ET | Evapotranspiration |
| EVI-2 | Earth Ventures Instruments, Second call |
| FPIE | Focal Plane Interface Electronics |
| FSWT | Flight Software Time (in GPS time) |
| GET-D | GOES Evapotranspiration and Drought System |
| GPS | Global Positioning System |
| HRSL | Hydrology and Remote Sensing Laboratory |
| ISS | International Space Station |
| L-2 | Level 2 |
| L-3 | Level 3 |
| LTAR | Long-Term Agroecosystem Research |
| MODIS | MODerate-resolution Imaging Spectroradiometer |
| NASS | National Agricultural Statistics Service |
| NLCD | National Land Cover Dataset |
| NOAA | National Oceanographic and Atmospheric Administration |
| PGE | Product Generation Executive |
| PM | Penman-Monteith |
| PSD | Product Specification Document |
| RMSD | Root Mean Squared Difference |
| SDS | Science Data System |
| SEB | Surface Energy Balance |
| TIR | Thermal Infrared |
| TSEB | Two-Source Energy Balance |
| USDA | United States Department of Agriculture |

# Contents

# 1 Introduction

The ECOSTRESS mission will provide high-resolution multi-spectral thermal infrared imagery to support field-scale mapping of evapotranspiration (ET) or consumptive water use. The thermal data will be converted to Level 2 (L-2) radiometric land surface temperature (LST) and emissivity products by JPL as described in the Surface Temperature Algorithm Theoretical Basis Document (ATBD).

## 1.1 Objective

The purpose of this Algorithm Specification Document (ASD) is to describe the computer processing system that will be used to generate Level 1B (L1B) Radiance files from the ECOSTRESS L1A CAL files.

## 1.2 Scope

This document describes the L1B Rad Product Generation Executive (PGE) implemented at the ECOSTRESS Science Data System (SDS) to generate L1B Radiance image files.

## 1.3 References

Reference 1: ECOSTRESS Level-1 Focal Plane Array and Radiometric Calibration Algorithm Theoretical Basis Document (ATBD), JPL D-94803

Reference 2: ECOSTRESS Level-1 Product Specification Document (PSD), JPL D-94634

Reference 3: ECOSTRESS L1A Calibration Algorithm Specification Document, JPL D-96966

Reference 4: ECOSTRESS Level-1B Resampling and Geolocation Algorithm Theoretical Basis Document (ATBD), JPL D-94641

# 2 Algorithm Description & Software Design

## 2.1 Data System Context

The ECOSTRESS processing levels are conceptually described as:

- Level 0 Processing prepares incoming datasets for higher-level processing
- Level 1 Processing generates engineering data products and calibrated, geolocated science measurements
- Level 2 Processing generates ECOSTRESS science results
- Level 3 and 4 Processing generate physical retrievals of target variables (ET and reference ET ratio)

## 2.2 The L1B Rad PGE Role in the ECOSTRESS Data System

The L1B Rad PGE is one of four PGEs within the L1 context (Figure 1). It uses the DN, Gain and Offset form L1A CAL PGE and the ephemeris and attitude data from the L1A Raw PGE to generate radiance images. In addition, it co-registers the 6 ECOSTRESS bands. The SWIR channel is not calibrated, so the data is left as DNs. However, the SWIR data is co-registered with the other 5 ECOSTRESS bands.

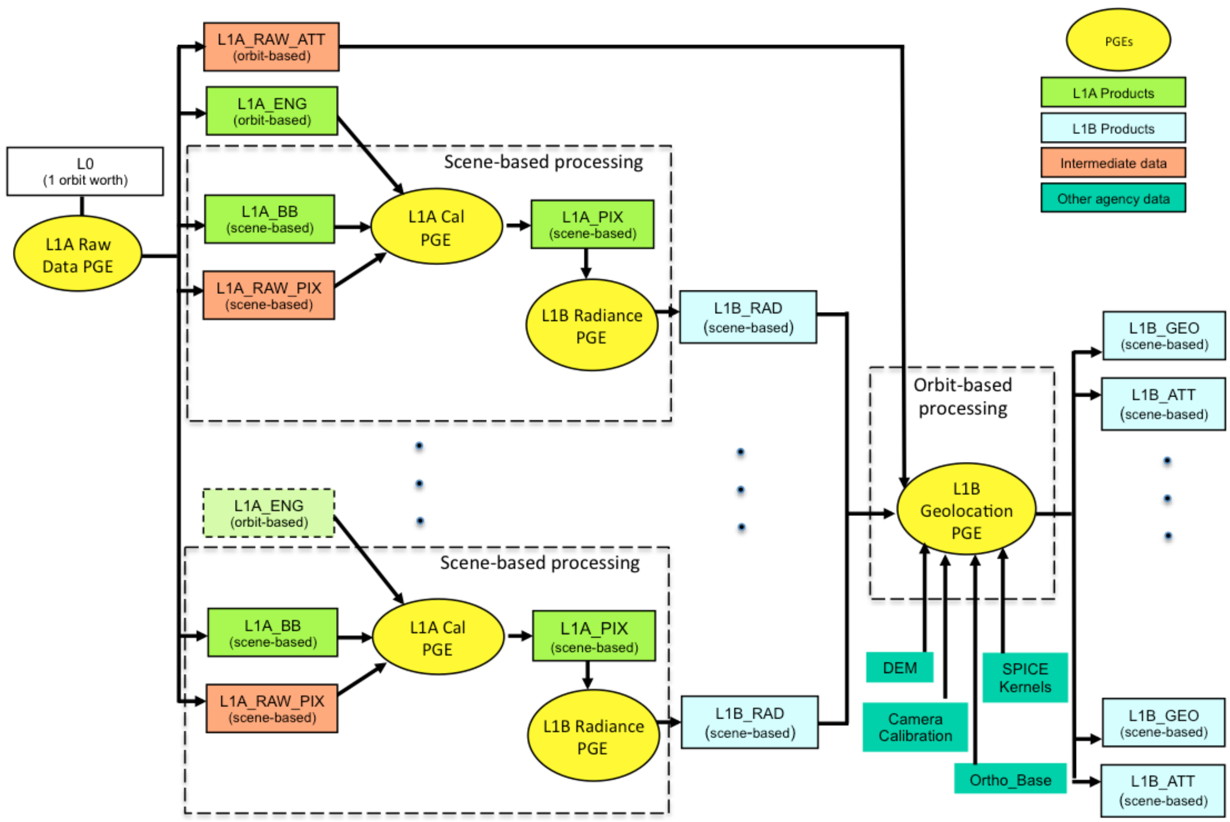


Figure 1: ECOSTRESS Level 1 (L1) Contextual Flow Diagram.

## 2.3 Input Data Sets

The following input data files are required:

- L1A_PIX: Containing the raw Digital Number (DN) imagery in HDF5 format of the ground target as expressed in:
  - Five TIR bands and
  - SWIR DN image with Dark Current subtracted
- L1A_RAD_GAIN: An intermediate file passed from L1A Cal PGE to the L1B Rad PGE
  - Five TIR image radiance Gain files
  - Five TIR image radiance Offset files
- L1A_RAW_ATT: Ephemeris and Attittude information
  - Ephemeris
  - Velocity
  - Attitude Quaternion

## 2.4 Output Data Sets

The following output data files are created:

- L1B_RAD: Contains the TIR bands as radiance, and the SWIR band as DN
  - 5 co-registered TIR bands as radiance data

- 5 Data Quality Indicators images
- SWIR DN image with Dark Current subtracted

# 3 Overview of Design

The L1B Rad PGE converts raw ground target imagery in Digital Numbers (DN) to calibrated units of radiance (Watt/m2/sr/um). The actual calculation of the gain and offset to apply is done in the previous L1A Cal PGE, and the values are passed in an intermediate L1A_RAD Gain file (see Reference 1 and Reference 3). This is done for the 5 TIR bands. The SWIR band is not calibrated, so the SWIR DNs are left as DN values.

By the design of the instrument, the ground footprint of each of the ECOSTRESS pixels is not square, instead it is nominally 35-meter x 70-meter. In the L1B Rad PGE we average two samples in the Focal Plane Array (FPA) to produce one nominally 70-meter x 70-meter pixel. This takes the 11264 x 5400 L1A_PIX image into a 5632 x 5400 image with nominally square ground footprint pixels.

Also, by design of the instrument, the 6 ECOSTRESS bands are not co-registered, instead they are offset from each other on the ground (see Reference 4). The L1B Rad PGE does a band to band registration to a common band (Band 4 – TIR 9.060-micron). The data is resampled to coregister the 6 bands to the common band.

Finally, a defect was introduced by an early vibration test performed on the ECOSTRESS instrument. There are 16 lines of data that are left as a DN value of 0 for 3 bands (Band 1 – TIR 8.285-micron, Band 5 – TIR 12.001-micron, and the SWIR band 1.66-micron). This is the so called “striping” defect in the instrument data. To reduce the effect of this defect, the 2 TIR bands are filled in with “best guess” values (the SWIR band is left unchanged). This best guess is the predicted value from a neural net trained on the full 5 bands.

Finally, data quality indicator fields are added to indicate pixels where we have uncorrected striping, interpolated values to fill in striping, bad or missing data, or pixels that are not seen (due to the bands not being co-registered when the data was collected).

# 4 Detailed descriptions

The following describes how the L1B Rad PGE processing is done.

## 4.1 Convert to Radiance Units

We convert to radiance units by applying the gain and offset calculated in the L1A Cal PGE to the DN values. For each band $b$, we calculate:

$$Rad_b = G_b * DN_b + Offset_b$$

## 4.2 Band to Band Co-registration

This calculation is done from each band to resample it to the reference band (Band 4 – TIR 9.060-micron):

1. For each band b (including the SWIR band):
    a. For each scan index
        i. Calculate band to band tie-points. This is done by taking a 10x30 evenly spaced grid:
            1. Determine the height of the center of the scan by projection the reference band to the surface.
            2. For evenly spaced grid points in the scan image:

a. Calculate the location of the surface at the fixed height for the reference band.

b. Project back to the instrument for band b.

c. Add a tie-point between the reference band (the original evenly spaced grid point) and the image coordinate calculated for band b.

ii. Fit a quadratic geometric model to the tie-points, giving the mapping from band b to the reference band.

iii. Use the geometric model to warp band b to the reference band. We use nearest neighbor resampling to preserve fill values in the original L1A Cal DN image. For the SWIR band we resample the DN values, for the TIR bands we resample the radiance calculated in section 4.1.

## 4.3 Create Square Pixels

Take the results from the band to band co-registration, and average in the line direction to give square pixels. The average should only include pixels that are not marked as fill in L1A Cal DN image:

```
for(i = 0; i < data.rows(); ++i)
   for(j = 0; j < data.cols(); ++j)
    v1 = raw(2*i,j);
    v2 = raw(2*i+1,j);
    if(v1 <= fill_value_threshold)
          if(v2 <= fill_value_threshold)
            data(i,j) = std::max(v1, v2);
          else
            data(i,j) = v2;
    else
          if(v2 <= fill_value_threshold)
            data(i,j) = v1;
          else
            data(i,j) = (v1 + v2) / 2;
```

## 4.4 Stripe Repair

We use the tensorflow/tflearn machine learning python library to predict the missing stripe data in bands 1 through 5 based on the values of band 2, 3, and 4. The algorithm for this is:

1. Create a training dataset by randomly selecting points:

```
 while counter < self.training_size:
   random_x_ind = random.randint(1, dataset.shape[0] - 2)
   random_y_ind = random.randint(1, dataset.shape[1] - 2)
   if missing_mask[random_x_ind, random_y_ind,
        band_number] > 0:
          continue
   grid_3x3x3 = dataset[(random_x_ind - 1):(random_x_ind + 2 ),
                   (random_y_ind - 1):(random_y_ind + 2 ),
                         1:4]
```

```
total_nan_in_grid = np.sum(np.isnan(grid_3x3x3))
if total_nan_in_grid > 0:
  continue

training_x[counter, :, :, :] = grid_3x3x3
training_y[counter] = dataset[random_x_ind, random_y_ind,
                     band_number]

counter= counter + 1
```

2. Create a 30 neuron neural net, and train on the training set. We want to predict y given x.
3. For each missing pixel in the stripes in band 1 and band 5:
    a. Create the predictor:

    ```
    grid_3x3x3 = dataset[ (row_index - 1):(row_index + 2 ),
                 (col_index - 1):(col_index + 2 ), 1:4 ]
    ```

    b. If grid_3x3x3 has any missing data, leave point as missing (we can’t predict it).
    c. If grid_3x3x3 does not have missing data, then use the neural net to predict the values for band 1 and band 5. Replace the radiance value in the output data with the predicted value.
    d. If we have replaced a pixel in band 1 or band 5, set the corresponding pixel in the data_quality_1 or data_quality_5 to “DQI_INTERPOLATED”.
    e. If we could not replace a pixel because grid_3x3x3 had missing data, then set data_quality_1 or data_quality_4 to “DQI_STRIPE_NOT_INTERPOLATED”.

## 4.5 Generating metadata

The PGE will generate both standard and product-specific metadata for use by the PCS to catalog and track each scene file.

The scene is marked as “Day” or “Night” based on the solar elevation angle of the center pixel of the image. If this is > 0 than the scene is marked “Day”, otherwise it is marked “Night”.

## 4.6 Saving ancillary data

The PGE also copies the fields “Time/line_start_time_j2000” and “FPIEncoder/EncoderValue” from the L1A_PIX file the L1B_RAD output file for use by later processing. The “EncoderValue” indicates a measure of scan mirror angle when each sample was acquired for each scan. The values is just directly read and copied from L1A_PIX to L1B_RAD.

The line start time in L1A_PIX is for the original, non-square image pixels, so there are 256 values for each of the 44 scans, or a total of 11264. Since we average the pixels, we read every other time from the L1A_PIX and write it to the L1B_RAD file, producing a total of 5632 times.

# 5 Other Considerations

## 5.1 Error handling

The L1B Rad PGE was designed to handle all L1B-expected problems, and to terminate with exit codes for other unexpected conditions. The PGE will return a value of "1" and exit if it finds conditions that prevent it from processing. It will return a value of "0" if the processing is successful. In addition to the standard SYSOUT log file, a formatted log file is created that summarizes the internal processing and provides additional details when a problem occurs.