JPL D-57134

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

Level 4 Water Use Efficiency (WUE) Algorithm Specification Document

Preliminary
June 8, 2018

Laura A. Jewell and Gregory H. Halverson
ECOSTRESS Algorithm Development Team
Jet Propulsion Laboratory
California Institute of Technology

Joshua B. Fisher, ECOSTRESS Science Lead
ECOSTRESS Algorithm Development Team
ECOSTRESS Science Team
Jet Propulsion Laboratory
California Institute of Technology

National Aeronautics and
Space Administration

JPL

Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, California 91109-8099
California Institute of Technology

Document Change Log

Revision	Date	Sections Changed	Author
Draft	06/01/2018	All	Laura Jewell
Preliminary	06/08/2018	All	Gregory Halverson Laura Jewell

Document TBD, TBR, TBS

Section/Page	Description	Date

Contacts

Readers seeking additional information about this document may contact the following ECOSTRESS Algorithm Development team members:

Laura A. Jewell
MS 168-200
Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena CA 91109
Email: Laura.A.Jewell@jpl.nasa.gov
Office: (818) 354-0645

Joshua B. Fisher
MS 233-305C
Jet Propulsion Laboratory
4800 Oak Grove Dr.
Pasadena, CA 91109
Email: jbfisher@jpl.nasa.gov
Office: (818) 354-0934

TABLE OF CONTENTS

1.0 INTRODUCTION

1.1 Identification

This is the Algorithm Specification Document (ASD) for Level 4 (L4) Water Use Efficiency (WUE) data product of the ECOsystem Spaceborne Thermal Radiometer Experiment on Space Station (ECOSTRESS) mission. The WUE product is generated from data acquired by the ECOSTRESS radiometer instrument according to the algorithm described in the ECOSTRESS Level 4 (WUE) Algorithm Theoretical Basis Document (ATBD) (JPL D-94649).

1.2 Purpose and Scope

This Algorithm Specification Document describes the Level 4 process used to generate the WUE product according to the PT-JPL algorithm. This includes an outline of the software used to compute evapotranspiration, and a description of the workflow to assimilate ECOSTRESS data products and the various ancillary data sets required for the L4 product.

The purpose of this ASD is to describe, in computer-science terms, the remote sensing algorithms that produce the ECOSTRESS end-user L4 data products. The science basis of an algorithm is not covered in an ASD, but is described in a corresponding ATBD (see section 1.4).

The ASD provides a software description of those algorithms as implemented in the operational ground system, the Science Data Operations System (SDOS). The intent of an ASD is to capture the "as-built" operational implementation of the algorithm. An individual ASD describes the process used in the creation of a single level of data product.

1.3 Mission Overview

The ECOSTRESS instrument measures the temperature of plants and uses that information to better understand how much water plants use and how they respond to stress.

ECOSTRESS addresses three overarching science questions:

- How is the terrestrial biosphere responding to changes in water availability?
- How do changes in diurnal vegetation water stress impact the global carbon cycle?
- Can agricultural vulnerability be reduced through advanced monitoring of agricultural water consumptive use and improved drought estimation?

The ECOSTRESS mission answers these questions by accurately measuring the temperature of plants. Plants regulate their temperature by releasing water through pores on their leaves called stomata. If they have sufficient water, they can maintain their temperature. However, if there is insufficient water, their temperatures rise. This temperature rise can be measured with a sensor in space. ECOSTRESS uses a multispectral thermal infrared (TIR) radiometer to measure the surface temperature, deployed on the International Space Station. The instrument will measure radiances at 5 spectral bands in the 8-12.5 μm range with approximately 70 by 70 meter of spatial resolution on the ground.

1.4 Applicable and Reference Documents

“Applicable” documents levy requirements on the areas addressed in this document. “Reference” documents are identified in the text of this document only to provide additional information to readers. Unless stated otherwise, the document revision level is Initial Release. Document dates are not listed, as they are redundant with the revision level.

1.4.1 Applicable Documents

ECOSTRESS Project Level 3 Science Data System Requirements (JPL D-94088).

ECOSTRESS Science Data Management Plan (JPL D-94607)

423-ICD-005 ICD Between ECOSTRESS SDS and LPDAAC

ECOSTRESS Level 1 Algorithm Theoretical Basis Documents (JPL D-94641, D-94642)

ECOSTRESS Level 1 Algorithm Specification Document

ECOSTRESS Level 2 Algorithm Theoretical Basis Documents (JPL D-94643, D-94644)

ECOSTRESS Level 2 Algorithm Specification Document

ECOSTRESS Level 3 (ET_PT-JPL) Algorithm Theoretical Basis Document (JPL D-94645)

ECOSTRESS Level 3 (ET_PT-JPL) Product Specification Document (JPL D-94636)

ECOSTRESS Level 4 (WUE) Algorithm Theoretical Basis Document (JPL D-94649)

ECOSTRESS Level 4 (WUE) Product Specification Document (JPL D-94653)

1.4.2 Reference Documents

1.5 ECOSTRESS Data Products

The ECOSTRESS mission will generate 13 different distributable data products. The products represent four levels of data processing, with data granules defined as an image scene. Each image scene consists of 44 scans of the instrument mirror, each scan taking approximately 1.181 seconds, and each image scene taking approximately 52 seconds. Each image scene starts at the beginning of the first target area encountered during each orbit. Each orbit is defined as the equatorial crossing of an ascending International Space Stations (ISS) orbit.

ECOSTRESS Level 0 data include spacecraft packets that have been pre-processed by the Ground Data System (GDS). Level 1 products include spacecraft engineering data, the time-tagged raw sensor pixels appended with their radiometric calibration coefficients, the black body pixels used to generate the calibration coefficients, geolocated and radiometrically calibrated at-sensor radiances of each image pixel, the geolocation tags of each pixel, and the corrected spacecraft attitude data. Level 2 products include the land surface temperature and emissivities of each spectral band retrieved from the at-sensor radiance data, and a cloud mask. Level 2 products also appear in image scene granules. Level 3 products contain evapotranspiration data derived from Level 2 data. Level 4 products contain evaporative stress index and water use efficiency derived from Level 3 data.

The ECOSTRESS products are listed in Table 1-1. This document will discuss only the Level 4 WUE product.

Table 1-1: ECOSTRESS Distributable Standard Products

Product type	**Description**
L0	Level 0 "raw" spacecraft packets
L1A_ENG	Spacecraft and instrument engineering data, including blackbody gradient coefficients
L1A_RAW_ATT	Uncorrected spacecraft ephemeris and attitude data
L1A_PIX	Raw pixel data with appended calibration coefficients
L1B_GEO	Geolocation tags, sun angles, and look angles, and calibrated, resampled at-sensor radiances
L1B_ATT	Corrected spacecraft ephemeris and attitude data
L2_LSTE	Land Surface temperature and emissivity
L2_CLOUD	Cloud mask
L3_ET_PT-JPL	Evapotranspiration retrieved from L2_LSTE using the PT-JPL Algorithm
L3_ET_ALEXI	Evapotranspiration generated using the ALEXI/DisALEXI Algorithm
L4_ESI_PT-JPL	Evaporative Stress Index generated with PT-JPL
L4_ESI_ALEXI	Evaporative Stress Index generated with ALEXI/DisALEXI
L4_WUE	Water Use efficiency
L3_L4_QA	Quality Assessment fields for all ancillary data used in L3 and L4 products

2.0 ALGORITHM AND DATA DESCRIPTION

2.1 The Level 4 role in the ECOSTRESS data system

The ECOSTRESS Level 4 WUE process requires the Level 3 Evapotranspiration data products as well as the Level 1 geolocation information and ancillary data products from other sources (see section 2.2.3). The L4 WUE product is computed directly from the L3 ET product.

2.2 Input data sets

2.2.1 Requirements on inputs

All input data for the L4 product must be geolocated and time-ordered, and therefore have latitude, longitude, and time associated with each pixel. We do not pose strict requirements on the input products regarding format, projection, and spatio-temporal range. Given the wide range of sources of input data (presented in section x), the fetch and reprojection pipeline described in section x is designed to take care of processing this input data into formats, projections, and subsets as required by the application code described in section x.

2.2.2 Attributes of ECOSTRESS input products

ECOSTRESS input data products used for the L4 WUE data product are:

- Level 1B – Geolocation
- Level 3 – Evapotranspiration

2.2.3 Attributes of ancillary input products

Other ancillary input data products used for the L4 WUE data product are:

- MODIS Land

See the table below for more detailed information.

Table 2-1: L4 WUE Ancillary Input Products

Product	**Description**	**Provider**
MOD17A2H	Gross Primary Productivity 8-Day L4 Global 1km	LP-DAAC

2.3 Output data sets

2.3.1 Attributes of output products

The ECOSTRESS Level 4 WUE output product fields are described in the table below:

Table 2-2: L4 WUE Output fields

Field Name	**Type**	**Unit**	**Field Data (per pixel, 5400 * 5632)**
GROUP	**WATER USE EFFICIENCY**		
WUEavg	Float	GPP/ET	Average Water Use Efficiency

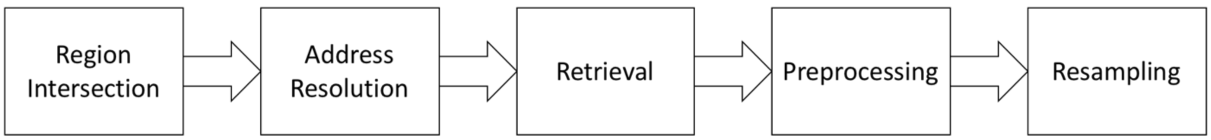
3.0 SOFTWARE DESIGN

3.1 Overview

The ECOSTRESS pipeline prepares source data from a multitude of remote sensing and reanalysis data products. These source data are retrieved, preprocessed, and resampled to an ECOSTRESS scene. These preprocessed data are run through a series of models to produce the ECOSTRESS product.

3.2 Description of major code sections

3.2.1 Pre-processing pipeline



The downloader PGE begins by calculating which source data granules spatio-temporally intersect the given ECOSTRESS scene. A polygon surrounding the convex hull of latitude/longitude pairs contained in the L1B geolocation arrays is taken as the destination region. Source regions from Landsat 8 in terms of Landsat path/row are determined by intersecting the target ECOSTRESS polygon with the static set of polygons in the WRS2-Descending region extent system. A static set of polygons in the MODIS Sinusoidal Grid are also intersected against the target footprint of the ECOSTRESS scene to select the horizontal and vertical index pairs for MODIS land products.

Selection of swath granules for the MODIS atmospheric products requires knowledge of the Terra satellite's orbit. To calculate the orbit of Terra, the Two-Line Element (TLE) observed by the Air Force at the time of ECOSTRESS overpass is queried from *spacetrack.com*. This TLE is used to calculate the swath ground track of Terra in a 12-hour radius surrounding the ECOSTRESS overpass time using the *pyephem* library. The boundary of potentially matching MODIS swath granules is calculated by dividing Terra's swath ground track into 5-minute increments beginning at UTC midnight and applying a constant cross-track swath width. The polygons surrounding these temporally intersecting MODIS swaths are then intersected against the target ECOSTRESS polygon to determine the set of MODIS swath granules that should be retrieved.

With the required granules of source data known, remote servers are crawled to resolve these granules to URLs of available granules for retrieval. Source Landsat 8 files are addressed on Amazon Web Services as individual GeoTIFF files for each band. Addresses for MODIS land data are queried from the Land Processes Distributed Active Archive Center (LP-DAAC) Inventory API. Addressed for MODIS atmosphere granules are obtained by crawling the CSV listings of the Level-1 and Atmosphere Archive & Distribution System Distributed Active Archive Center (LAADS DAAC) HTTP Interface.

With the addresses of required granules recorded, the downloader PGE makes calls to WGET to retrieve these files from the remote servers, organized into a consistent directory structure on the local filesystem. In order to conserve on storage requirements, all runs of the ECOSTRESS downloader retrieve to a collective local mirror. WGET is able to efficiently retrieve only files that it has not already retrieved.

The preprocessor PGE loads and preprocesses the retrieved source data granules. Both Landsat and MODIS record their datasets in 16-bit integer analogs to floating point values. These Digital Numbers (DNs) in the Landsat bands are converted to reflectance using linear parameters read from the Landsat metadata, and these reflectance values are used to calculate NDVI and albedo. Atmospheric and land inputs are ready-made as products in MODIS but are also linearly converted into real units and filtered by accompanying quality flags.

With source data preprocessed at their original locations, these data are ready to be projected onto the target ECOSTRESS scene. Since it is only possible to project between gridded surfaces, intermediate gridded surfaces must be created to represent the source MODIS swaths and destination ECOSTRESS swath. The *pyresample* library is used to resample by nearest neighbor between swaths and their representative grids. The *rasterio* library is used to warp data between gridded surfaces, smoothed by cubic interpolation. These techniques are combined to mosaic preprocessed source data over the extent of the ECOSTRESS scene and resample this mosaic onto the ECOSTRESS geolocation provided by the L1B product. Finally, these data are stored as the inputs for the PT-JPL and disALEXI PGEs.

This data-store of model forcings passed to the model PGEs is organized by variable and date into a directory structure and written in Hierarchical Dataset Format (HDF5) using the *h5py* library. HDF5 was chosen as an intermediate file format to accommodate geolocated/non-gridded data.

3.2.2 Computation of Water Use Efficiency

The PT-JPL PGE loads the preprocessed forcing data prepared by the preprocessor PGE and runs them through a sequence of surface energy balance models. All variables, including geolocation, are loaded from the forcing data store created by the preprocessor using the HDF5 library. The Forest Light Environmental Simulator (FLiES) calculates incoming shortwave radiation from atmospheric observations. This is accomplished by querying a multidimensional array of precomputed FLiES model output as a lookup table. The Breathing Earth Systems Simulator (BESS) ingests the ECOSTRESS LST as outgoing longwave radiation. This balance is closed to produce net radiation, the total energy available at the surface. This instantaneous net radiation value is sinusoidally integrated from sunrise to sunset to produce a daytime daily average net radiation.

The Priestley-Taylor Jet Propulsion Laboratory (PT-JPL) model calculates transpiration, interception, and soil evaporation from net radiation using the Priestley-Taylor formula constrained by observed land surface properties. These partitions are summed as instantaneous latent heat flux. The evaporative fraction derived from this instantaneous value is multiplied by the daily net radiation to calculate a daily latent heat flux in watts per square meter. A constant latent heat of vaporization and the length of time from sunrise to sunset are taken into account in converting this value to a total mass of evaporated water in kilograms. The 8-day Gross Primary Production (GPP) given by MOD17 is divided into a daily average GPP in grams of carbon. The WUE product is calculated as the ratio of daily GPP in grams of carbon to daily ET in kilograms of water. This ratio is written to the WUE product using the HDF5 library.

4.0 OTHER CONSIDERATIONS

4.1 Error handling

Error handling is managed using Python's exception handling framework. Exception classes were designed for all foreseeable disruptions to data processing, each of which is assigned a unique exit code that is returned to the Process Control System (PCS).

4.2 Dependencies on existing software

The Anaconda distribution of Python 3.5 was chosen as the platform for the downloader/preprocessor subsystem because of its good support for geospatial data processing. The conda environment manager included in Anaconda provides seamless dependency management for the scientific, geospatial, and data management packages required for ECOSTRESS. These package dependencies include: `pandas`, `affine`, `ephem`, `fiona`, `geopandas`, `h5py`, `lxml`, `netcdf4`, `nose`, `numpy`, `pyproj`, `python-dateutil`, `rasterio`, `requests`, `shapely`, `six`, `spacetrack`, `bs4`, `usgs`, `homura`, `termcolor`, `geocoder`, `scipy`, `pyresample`, `future`, `pygments`, `xmltodict`, `scikit-image`, `jsonmerge`, `rasterstats`, and `untangle`.

The C++ language was chosen to implement the PT-JPL model code to maximize the efficiency of processing large arrays. The main dependency linked to by the PT-JPL PGE is the Hierarchical Data Format (HDF) library, which is required for writing the ECOSTRESS product to file.

4.3 Assumptions and limitations

This software assumes daytime acquisition. It is designed not to run for granules acquired at night.

4.4 Quality assessment and recording

No original quality flags are produced. Instead the quality flags of the source data products are resampled by nearest neighbor onto the geolocation of the ECOSTRESS scene. A quality flag array for each input dataset, when available, is collected into a combined QA product accompanying the PT-JPL product file.

5.0 APPENDIX A: ABBREVIATIONS AND ACRONYMS

ALEXI	Atmospheric-Land Exchange Inversion
ASD	Algorithm Specifications Document
ATBD	Algorithm Theoretical Basis Document
BESS	Breathing Earth Systems Simulator
DAAC	Distributed Active Archive Center
DisALEXI	ALEXI Disaggregation algorithm
ECOSTRESS	ECOsystem Spaceborne Thermal Radiometer on Space Station
ESI	Evaporative Stress Index
ET	Evapotranspiration
FLiES	Forest Light Environmental Simulator
GDS	Ground Data System
GPP	Gross Primary Production
HDF	Hierarchical Data Format
Hz	Hertz
ICD	Interface Control Document
ISS	International Space Station
JPL	Jet Propulsion Laboratory
Km	kilometer
L0 – L4	Level 0 through Level 4
LP	Land Processes
LST	Land Surface Temperature
LSTE	Land Surface Temperature and Emissivity
m	meter
MODIS	Moderate Resolution Imaging Spectroradiometer
NASA	National Aeronautics and Space Administration
NCEP	National Centers for Environmental Protection
netCDF	Network Common Data Format
PCS	Process Control System
PGE	Product Generation Executable
PSD	Product Specifications Document
PT-JPL	Priestly-Taylor-JPL
QA	Quality Assurance
SDOS	Science Data Operations System
SDS	Science Data System
TBD	To Be Determined
TBR	To Be Reconciled
TBS	To Be Specified
USDA	United State Department of Agriculture
USGS	United States Geological Society
UTC	Coordinated Universal Time
WUE	Water Use Efficiency
XML	Extensible Markup Language