EOG DNB Profile Datasets

By Feng Chi, Hsu (fengchihsu@mines.edu)
Created 10/15/2020

Objective

This document aims to help users navigate through the file collections produced by Earth Observation Group (EOG) for Day Night Band (DNB) profiles. EOG conducted DNB profile for selected regions, and the results are stored in file formats including comma separated values (CSV), keyhole markup language (KML), and PNGs. This document will explain the purpose of the directory structure, and the files stored in them. Users will be able to understand and make use of those files with the help of this document.

Background

EOG has a long history of working with nighttime remote sensing data. The birth of JPSS satellite series further expands the possibility of this field, especially with the advanced DNB sensor. EOG used to produce annual nighttime light (NTL) products, and recently begin making monthly ones. Nevertheless, it is always interest in investigating NTL in even finer temporal steps in large scale. Therefore, EOG developed infrastructure to extract pixel values and related metadata for given location from all S-NPP file archive. On top of that, a set of statistical analysis is performed with the temporal profile extracted for each point. This provides a novel gateway for researchers to investigate fine details of NTL changes overtime.

Technology

The two pillars for making this dataset possible are the database which hosts all VIIRS image and make it able to query and extract data by request. Another program is written to perform additional statistical analysis on top of the data extracted from the database.

Figure 1 shows how the database works with the analysis program.

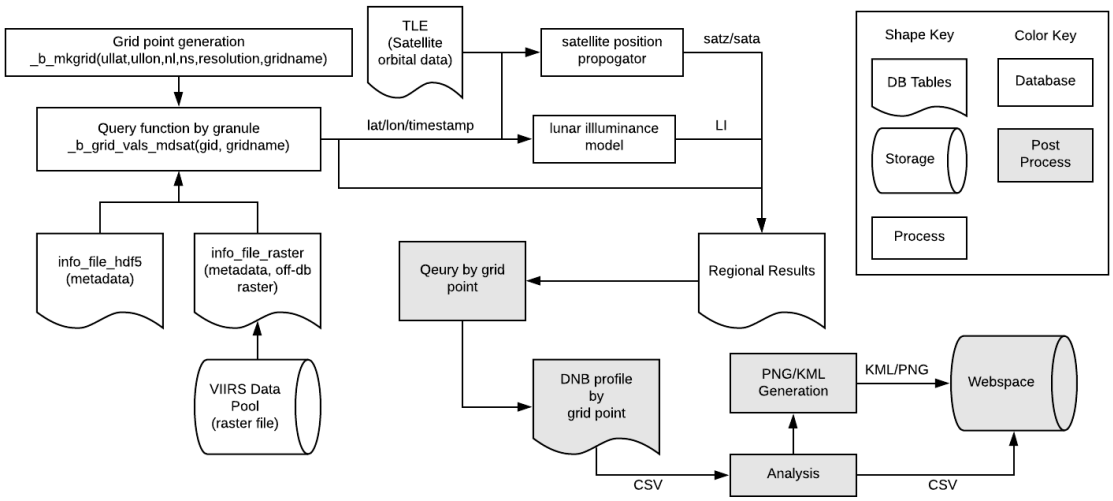


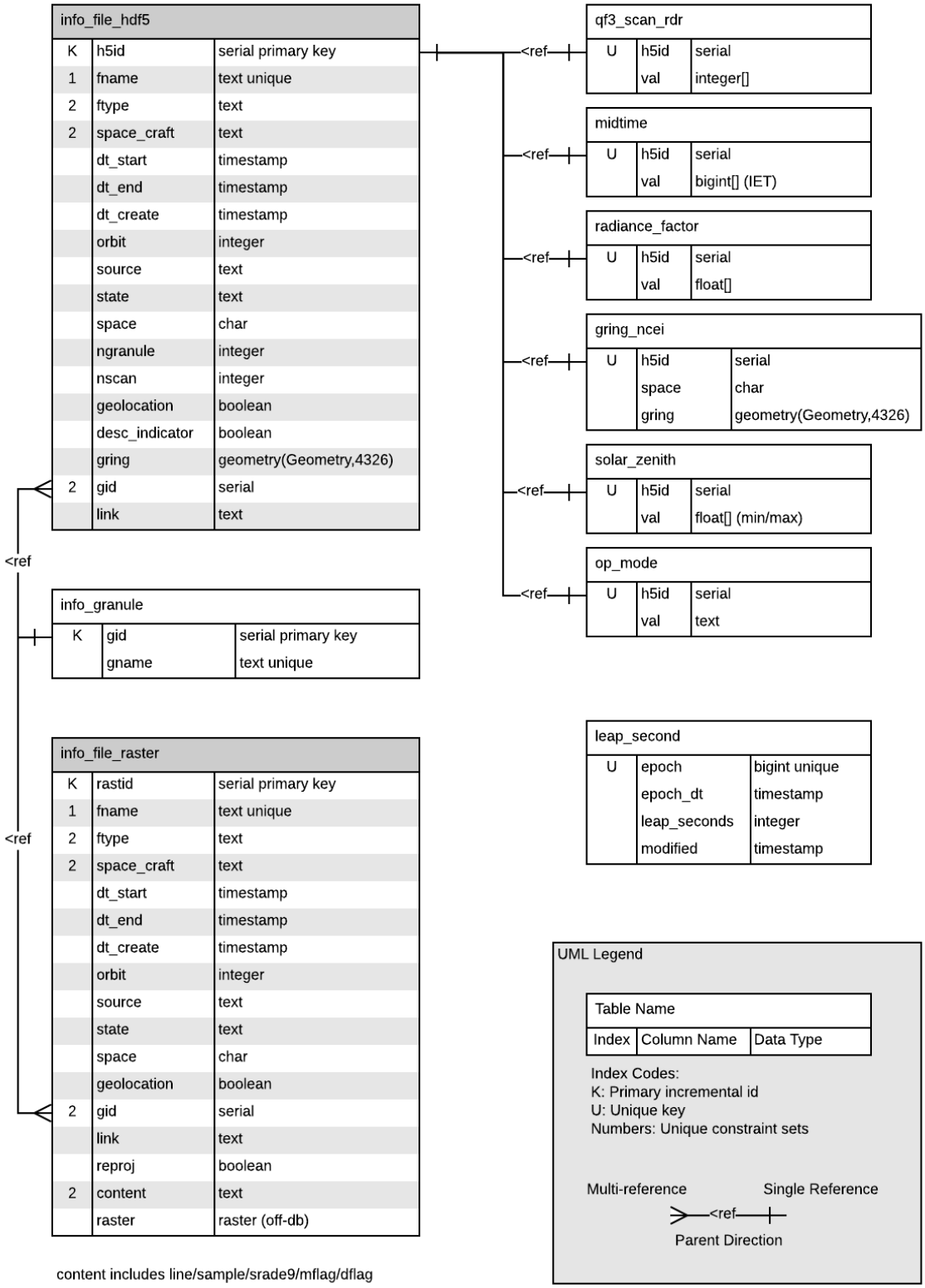
Figure 1 Flow chart of major steps in VIIRS image database and temporal profile analysis program

VIIRS Image Database

The database is built with PostgreSQL. The database itself stores metadata for all metadata for related HDF5 files and raster files. The actual pixel arrays are stored as out-DB raster types, which means the database only stores the path to the file, but not the pixel array itself. This enables EOG to archive the raster file along with other files for the same granule, while avoid double the storage space required. Also, the raster can be updated if necessary, without having to change the database records.

The database schema is designed as shown in Figure 2. The two most important tables are info_file_hdf5 and info_file_raster, which store all meta data for VIIRS HDF5 files and derived raster files respectively. Both tables have foreign key of gid (granule ID) pointing to the table info_granule. There are subsequent tables with foreign key h5id (HDF5 ID) pointing to info_table_h5. That means, if one of the granule record is removed from info_granule, all subsequent records related to that granule will be removed from the database.

Figure 3 explains how and where the database got the values required in each table from VIIRS HDF5 files. Figure 4 shows a higher level view of how data are ingested and processed in EOG Near Realtime (NRT) system.



content includes line/sample/srade9/mflag/dflag

Figure 2 The schema of VIIRS Image Database

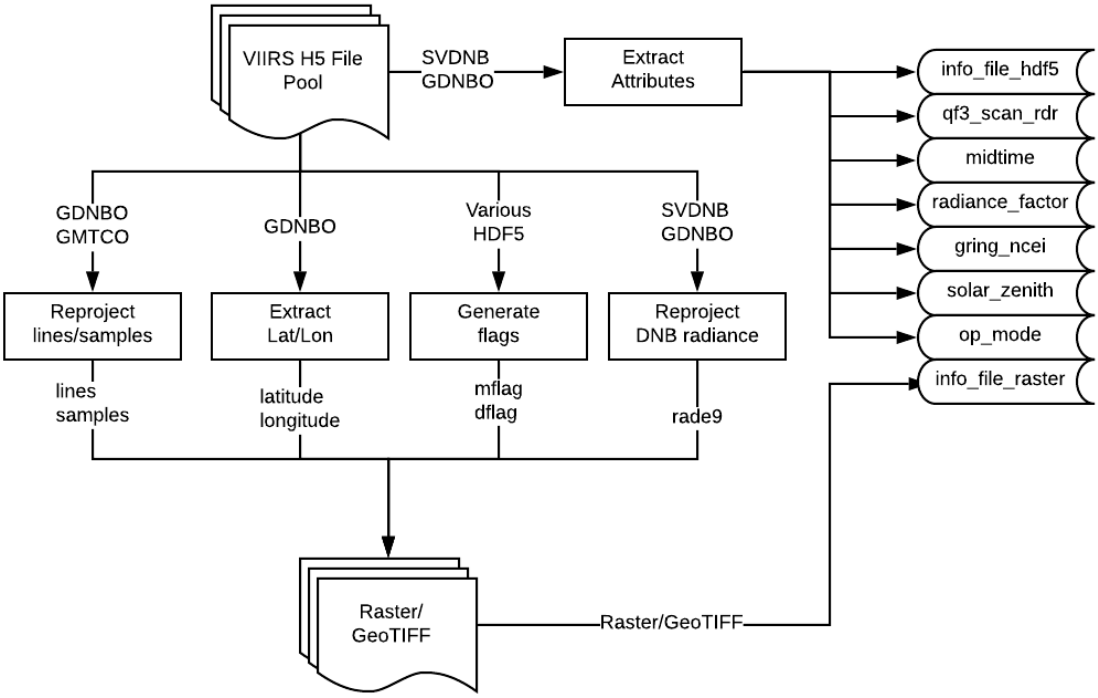


Figure 3 The data flow from VIIRS HDF5 files to VIIRS Image Database

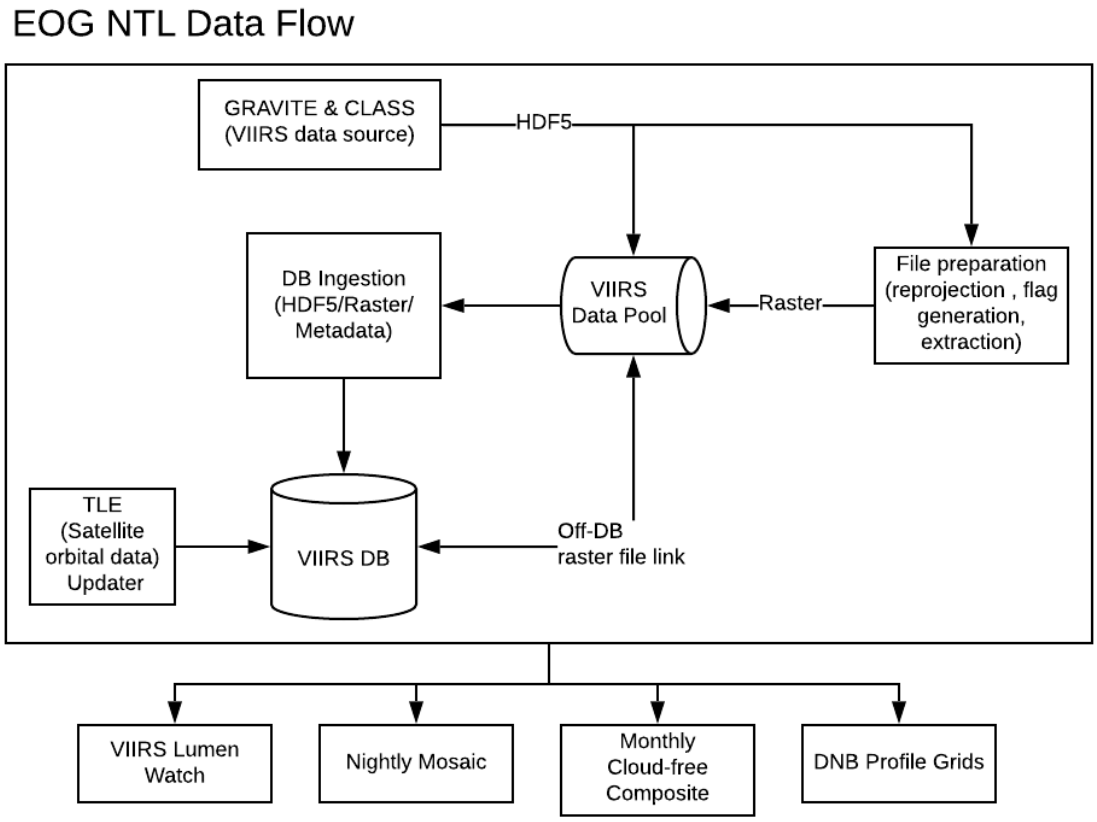


Figure 4 Higher level data flow in EOG NRT system

Temporal Profile Extraction from VIIRS Image Database

To extract the temporal profile, the first step is to define the grid points for the region of interest (ROI). See another document for how grid points are defined. Pixels from DNB images will be associated to the nearest grid point. See another document (in planning) to learn how the query function is designed in PostgreSQL DB to perform the extraction.

The extracted DNB profile will be supplied with satellite geometry and lunar illuminance (LI) values. These two sets of data are calculated on-the-fly using the timestamp, two-line-element (TLE) table, and coordinate of the given pixel location. The summary of algorithm to approximate satellite geometry using SGP4 model is shown in Figure 5. The LI is calculated with the model developed by US Naval Observatory in document Circular NO. 171 titled "Computer Programs for Sun and Moon: Illuminance with Contingent Tables and Diagrams" by P. M. Janiczek and J. A. DeYoung published in 1987.

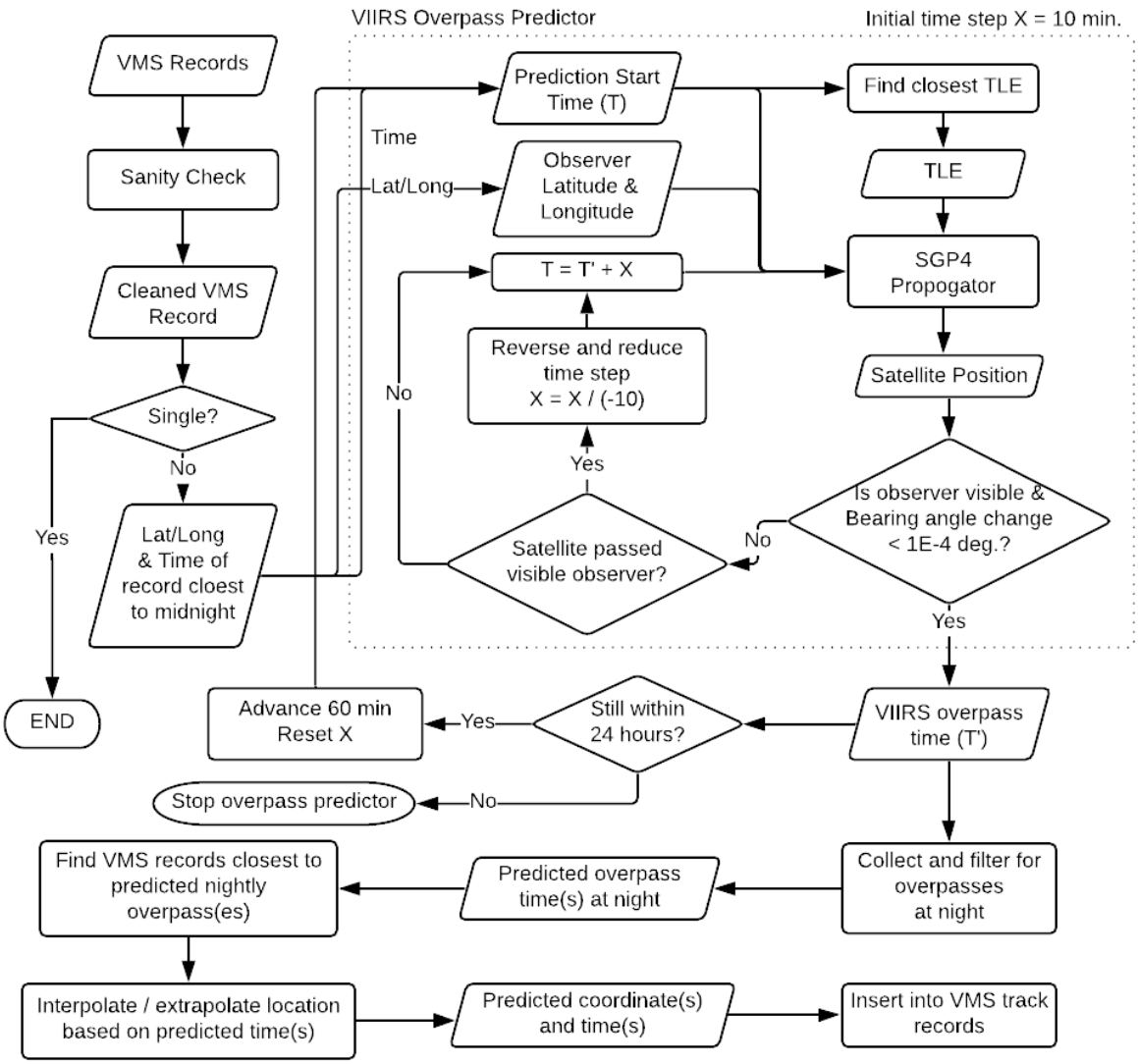


Figure 5 VIIRS Overpass Predictor flow

Temporal Profile Analysis Program

The extracted DNB temporal profile receives additional processing to reveal its statistical meaning and attributes. The details can be found in the published paper "Indicators of Electric Power Instability from Satellite Observed Nighttime Lights".

Datasets

The resulted dataset are stored in the format of CSV, PNG, and KML.

Regional Naming Convention

The regional naming is defined as below.

[country code]_[region ID]

- Country Code: ISO-3166-3 country code.
- Region Name: Region identifier name, see Table 1 for details.

Table 1 Region names example

Country Code	Region ID	Comment
bgd	dhaka	Bangladesh Dhaka
bgd	rohingya_refugee_camp	Bangladesh Rohingya Refugee Camp
chn	beijing	China Beijing
chn	shanghai	China Shanghai
chn	shdisney	China Shanghai Disney Land
chn	wuhan	China Wuhan
ind	charpakri_bihar	India Charpakri Bihar
ind	delhi_airport	India Delhi Airport
ind	gaya_bihar	India Gaya Bihar
ken	nairobi	Kenya Nairobi
kor	seoul	South Korea Seoul
prk	pyongyang	North Korea Pyongyang
usa	bakken	USA Bakken Oil Field
usa	blackrock	USA Blackrock
usa	houston	USA Houston
usa	ncal	USA North California (Santa Rosa)
usa	whitesands	USA Whitesands
yem	marib	Yemen Marib
yem	sanaa	Yemen Sana'a
zwe	harare	Zimbabwe Harare

Regional Grid

Each region contains a collection of points. Each point falls on the 15 arc second grid. The global grid is defined with upper left point at 180.0W, 90.0N. See grid definition document for details. Each regional point is given a [ptid] as x+[grid_width]*y.

Directory Structure

The files are stored in separate directories for each region. The directory structure is shown in Table 2.

Here we list some scenarios and which file to use.

- I want to see the complete DNB temporal profile without any filtration.

  Please contact EOG for access to full record with cloudy nights.

- I want to see the DNB temporal profile without low moon nights.

  You can find the cloud-free, low-moon night only DNB profile with satellite zenith angle normalization at:
  [region_name]/csv/[region_name]_[ptid].csv

- I want to see the complete DNB temporal profile with low moon nights.

  You can find the cloud-free, satellite zenith angle normalized, LI corrected hence include all lunar state (for [flavor] is "licorr") DNB profile temporal series at:
  [region_name]/csv/[region_name]_[ptid].[flavor].csv

- Which column should I use?

  For raw DNB radiance: use "dnbrad".

  For satellite zenith angle normalized DNB radiance: "rade9_mult_nadir_norm"

  For satellite zenith angle normalized and LI corrected: "rade9_dnb_licorr"
  T*here is no LI corrected but no zenith angle normalized DNB radiance, as LI correction requires the radiance to be satellite zenith angle normalized.*

  For timestamps of the record: "midscan"

  For pixel coordinates, it depends on which pixel you are looking for.

  - DNB pixel: use "dlat" and "dlon".
  - M-band pixel: use "mflag" and "mlon".
  - Location of 15 asec pixel: use "pt_lat" and "pt_lon".

- How can I use the flags column, I.e. "mflag", "dflag", and "vflag"?

  See another document for EOG flag definition. (In planning)

- I want to exam closely the autocorrelation function (ACF) series used to plot the ACF chart.

  You can find the ACF series for each [ptid] at:
  [region_name]/csv/[region_name]_[ptid].[flavor].acf.csv

- I want to see the statistical summary of the DNB profile.

  The statistical summary for each [ptid] can be found at:
  [region_name]/csv/[region_name]_[ptid].[flavor].stats.csv

These are one-row CSV for each [ptid], there is also statistic summary CSV for all points at: [region_name]/[region_name].[flavor].stats.csv

- I want to browse the results on Google Earth.

  You can use the KML file to show all regional points at:
  [region_name]/[region_name].[flavor].grid_segm.kml
  This is the parent KML with links pointing to children KMLs each having 2k points. Therefore the parent KML can be kept very lightweight.

- I want to find the images used in KML showing the analysis results.

  Those images are in PNG format, and can be found at:
  [region_name]/img/[region_name].[ptid].png

Table2 Directory structure for regional DNB profile dataset

Directory and file name	**Comment**
[region_name]	Regional directory
csv	CSV directory
[region_name]_[ptid].csv	Default DNB profile with satellite zenith angle correction (cloud-free, low-moon only)
[region_name]_[ptid].[flavor].acf.csv	ACF series CSV
[region_name]_[ptid].[flavor].csv	Cloud-free, full lunar state, LI and satellite zenith angle corrected DNB profile
[region_name]_[ptid].[flavor].stats.csv	Statistical summary one-row CSV
img	Image for KML directory
[region_name]_[ptid].png	Profile analysis result charts for KML
[region_name]_[ptid].[flavor].grid_segm.kml	Parent KML for the region
[region_name]_[ptid].[flavor].grid_segm_[part id].kml	Child KML for the region, 2k points each
[region_name].[flavor].stats.csv	Aggregated statistical summary CSV

PNG and KML Files

Please see this document on how to read the information displayed in the chart panel.

CSV Files

The headers for each type of CSV file is explained in the following tables.

Table 3 Header Definition for DNB profile CSV

Column name	**Definition**	**Unit**
gid	Granule ID	Unitless
satellite	Satellite Name	Unitless
dlines	Line number of the pixel in DNB image	Unitless
dsamples	Sample number of the pixel in DNB image	Unitless
dnbrad	Raw DNB radiance	$nW/cm^2/sr$

dlat	Latitude of the pixel in DNB image	Degrees
dlon	Longitude of the pixel in DNB image	Degrees
dflag	Bitflag for pixel in DNB image	Unitless
mlines	Line number of the pixel in M-band image	Unitless
msamples	Sample number of the pixel in M-band image	Unitless
mflag	Bitflag for pixel in M-band image	Unitless
vflag	mflag + vflag	Unitless
ptid	Point ID	Unitless
pt_name	Point name	Unitless
pt_lat	Point latitude in 15-arcsecond grid	Degrees
pt_lon	Point longitude in 15-arcsecond grid	Degrees
pt_x	Point x location in regional grid	Unitless
pt_y	Point y location in regional grid	Unitless
pt_base	Base name of regional grid	Unitless
midscan	Mid-scan time of the scan	Timestamp
sata	Satellite azumuth angle	Degrees
satz	Satellite zenith angle	Degrees
scan_angle	Scan angle	Degrees
li	Lunar illuminance	Lux
satz_bin_n	Satellite zenith angle bin number	Unitless
satz_lut	Satellite zenith angle from LUT	Degrees
sample_aggzone_sd	Standard deviation DNB radiance in the aggregation zone	$nW/cm^2/sr$
satz_bin_mdn	Median satellite zenith angle in the bin	$nW/cm^2/sr$
rade9_bin_1q	25% quantile of DNB radiance in the bin	$nW/cm^2/sr$
rade9_bin_mdn	50% quantile of DNB radiance in the bin	$nW/cm^2/sr$
rade9_bin_3q	75% quantile of DNB radiance in the bin	$nW/cm^2/sr$
adjust_mult_nadir_norm	Adjust coefficient (multiplicative)	Unitless
rade9_aggzone_norm	Zenith angle normalized DNB radiance (additive)	$nW/cm^2/sr$
rade9_mult_nadir_norm	Zenith angle normalized DNB radiance (multiplicative)	$nW/cm^2/sr$
rade9_dnb_licorr	LI corrected DNB radiance	$nW/cm^2/sr$

cloud_state	Cloud state parsed from vflag	Unitless

Table 2 Header Definition for Statistic Summary CSV (one-row and aggregated)

Header Name	**Definition**	**Unit**
region	Region name ([country code]_[region_id])	Unitless
ptid	Serial ID of grid point (Start from 1)	Unitless
pt_lat	Latitude of grid point	Degrees
pt_lon	Longitude of grid point	Degrees
pt_x	X location of grid point in grid array (Start from 1)	Unitless
pt_y	Y location of grid point in grid array (Start from 1)	Unitless
lidnb_phi_skew	Delete	
lidnb_norm_r2	Delete	
all_li	Delete	
licorr_slope	Slope used for LI correction	Unitless
longterm_slope	Slope of longterm trend line	$nW/cm^2/sr/day$
longterm_intercept	Intercept of longterm trend line	$nW/cm^2/sr$
longterm_r2	R square value of longterm trend line	Unitless
all_skew	Skew value for full profile	Unitless
all_kurtosis	Kurtosis value for full profile	Unitless
all_mean	Mean radiance vlaue for full profile	$nW/cm^2/sr$
all_variance	Variance of radiance value for full profile	$nW/cm^2/sr$
all_median	Median of radiance value for full profile	$nW/cm^2/sr$
all_cdfslope	CDF slope value of radiance vlaue for full profile	Unitless
total_count	Count of records	Unitless
primary_lag	Primary lag	days
primary_acf	ACF at primary lag	Unitless
primary_prominence	Prominance at primary lag days	Unitless
secondary_lag	Secondary lag	days
secondary_acf	ACF at secondary lag	Unitless
secondary_prominence	Prominance at secondary lag days	Unitless
annual_cycle	Annual cycle detected	Boolean
pop2018	Population in the grid at 2018	Person

annual_lift	Lift index by year	nW/cm^2/sr
annual_mean	Mean radiance value by year	nW/cm^2/sr
annual_variance	Variance of radiance by year	nW/cm^2/sr
annual_skew	Skew vlue by year	Unitless
annual_kurtosis	Kurtosis value by year	Unitless
annual_cdfslope	CDF slope value by year	Unitless
annual_outage	Outage ratio by year	Unitless
annual_ald	ALD value by year	Unitless