

Help on module read_datasets:

NAME

Read_datasets

AUTHOR

Katie Whitman

FILE

operational-sep/library/read_datasets.py

DATA

__email__ = 'kathryn.whitman@nasa.gov'

__maintainer__ = 'Katie Whitman'

VERSION

V0.1

FUNCTIONS

check_paths()

: Check that the paths that hold the data and output exist. If not, create.

check_data(startdate, enddate, experiment, flux_type)

: Check that the files containing the data are in the data directory. If the files for the requested dates aren't present, they will be downloaded from the NOAA website. For SEPTEM (RSDv2) data, if missing, the program prints the URL from which the data can be downloaded and unzipped manually.

The RSDv2 data set is very large and takes a long time to read as a single file. This program will generate files containing fluxes for each year for faster reading.

check_ephin_data(startdate, enddate, experiment, flux_type)

: Check for SOHO/COSTEP/EPHIN data on your computer. If not there, download from <http://ulysses.physik.uni-kiel.de/costep/level3/l3i/> 30 minute data will be downloaded. Intensities are in units of $(\text{cm}^2 \text{ s sr mev/nuc})^{-1}$

First available date is 1995 12 8 (DOY = 342).

The files are available in daily or yearly format.

check_for_bad_data(dates, fluxes, energy_bins)

- : Search the data for bad values (flux < 0) and fill the missing data with an estimate flux found by performing a linear interpolation with time, using the good flux values immediately surrounding the data gap.

check_goes_data(startdate, enddate, experiment, flux_type)

- : Check that GOES data is on your computer or download it from the NOAA website. Return the filenames associated with the correct GOES data.

check_sepem_data(startdate, enddate, experiment, flux_type)

- : Check if SEPTEM data is present on the computer. Break into yearly files if needed. Return SEPTEM filenames for analysis.

define_energy_bins(experiment, flux_type)

- : Define the energy bins for the selected spacecraft or data set.
If the user inputs their own file, they must set the user_energy_bins variable in library/global_vars.py.
User may select options to apply Sandberg et al. (2014) effective energies for GOES EPS by specifying "S14" and/or apply Bruno (2017) effective energies for GOES-13 or -15 P6, P7 and HEPAD by specifying "Bruno2017"

do_interpolation(i, dates, flux)

- : If bad fluxes (flux < 0) are found in the data, find the first prior data point and the first following data point that have good flux values. Perform linear interpolation in time:
$$F(t) = F_1 + (t - t_1) * (F_2 - F_1) / (t_2 - t_1)$$

This subroutine does the calculation for a single instance of bad data that corresponds to array index i.

extract_date_range(startdate, enddate, all_dates, all_fluxes)

- : Extract fluxes only for the dates in the range specified by the user.

find_goes_data_dimensions(filename)

- : Input open csv file of GOES data. Identifies the start of the data by searching for the string 'data:', then returns the number of header rows and data rows present in the file.

get_west_detector(filename, dates)

- : For GOES-13+, identify which detector is facing west from the orientation flag files. Get an orientation for each data point.
EPEAD orientation flag. 0: A/W faces East and B/E faces West.
1: A/W faces West and B/E faces East. 2: yaw-flip in progress.

read_in_ephin(experiment, flux_type, filenames1)

- : Read in EPHIN files from your computer.

read_in_files(experiment, flux_type, filenames1, filenames2, filenames_orien)

- : Read in the appropriate data files with the correct format. Return an array with dates and fluxes. Bad flux values (any negative flux) are set to -1. Format is defined to work with the files downloaded directly from NOAA or the RSDv2 (SEPTEM) website as is.

The fluxes output for the GOES-13+ satellites are always from the westward-facing detector (A or B) by referring to the orientation flags provided in the associated orientation file. Data taken during a yaw flip (orientation flag = 2) are excluded and fluxes are set to -1.

Note that the EPS detectors on GOES-08 and -12 face westward. The EPS detector on GOES-10 faces eastward. GOES-11 is a spinning satellite.

read_in_goes(experiment, flux_type, filenames1, filenames2, filenames_orien)

- : Read in GOES data from your computer. User may specify option to choose corrected or uncorrected GOES fluxes.

read_in_sepem(experiment, flux_type, filenames1)

- : Read in SEPTEM data files from the computer.

read_in_user_files(filenames1)

- : Read in file containing flux time profile information that was specified by the user.

The first column MUST contain the date in YYYY-MM-DD HH:MM:SS format. The remaining flux columns to be read in are specified by the user in the variable user_col at the very beginning of this program.

The date column should always be considered column 0, even if you used whitespace as your delimiter. The code will consider the date format YYYY-MM-DD HH:MM:SS as one column even though it contains whitespace.

Any number of header lines are allowed, but they must be indicated by # at the very beginning, including empty lines.

Be sure to add the energy bins associated with your flux columns in the

subroutine `define_energy_bins` under the "user" is statement.

`make_yearly_files(filename)`

: Convert a large data set into yearly files.