Tutorial for Using getDataFromCubeZone.py to Download Data from Given Coordinates

Prerequisites

Before you start, ensure you have the following:

- 1. Python installed on your computer.
- 2. The **getDataFromCubeZone**.py script, which is located in the agwise/AgWise_Data/getData_fromCube directory of the CGLABS repository.

Script Arguments

The script takes several arguments to define the data you want to download. Here's a breakdown of each argument:

- --startDate STARTDATE: The start date of the data cube in the format 'm-d' (e.g., '05-11' for May 11).
- --endDate ENDDATE: The end date of the data cube in the format 'm-d' (e.g., '08-30' for August 30).
- --xmin XMIN: Minimum X coordinate (longitude).
- --xmax XMAX: Maximum X coordinate (longitude).
- --ymin YMIN: Minimum Y coordinate (latitude).
- --ymax YMAX: Maximum Y coordinate (latitude).
- --variable VARIABLE: The variable to download (e.g., 'chirps-precipitation').
- --numberofyears NUMBEROFYEARS: Number of years to download the data for (in the format 'm-d').
- --lastYear YEAR: The year of the last data that we need to evaluate
- --downloadpath DOWNLOADPATH: The path where the data should be downloaded.

Example Command

Here's an example of how to run the script with all the necessary arguments:

bash

python **getDataFromCubeZone**.py --startDate='05-11' --endDate='08-30' --xmin=-90 --xmax=-83 --ymin=12 --ymax=16 --variable='chirps-precipitation' --downloadpath='/tmp/' --numberofyears=40

In this example:

- The data will be downloaded starting from May 11 to August 30.
- The geographical range is between longitudes -90 and -83, and latitudes 12 and 16.

- The variable 'chirps-precipitation' is specified for downloading.
- The data for the past 40 years will be downloaded.
- The data will be saved to the /tmp/ directory.

Runnig the code using LINUX Terminal

1. Open Your Terminal or Command Prompt

First, open your terminal (Linux or macOS) or command prompt (Windows).

2. Navigate to the Script Directory

Navigate to the directory where **getDataFromCubeZone**.py is located. Use the cd command to change directories. For example:

bash

cd path/to/agwise/AgWise_Data/getData_fromCube

Replace path/to/ with the actual path to the script.

3. Run the Script

Use the python command to run the script with the appropriate arguments. Here's the command again for reference:

bash

python **getDataFromCubeZone**.py --startDate='05-11' --endDate='08-30' --xmin=-90 --xmax=-83 --ymin=12 --ymax=16 --variable='chirps-precipitation' --downloadpath='/tmp/' --numberofyears=40

4. Check the Download Path

Once the script has finished running, navigate to the download path you specified (e.g., /tmp/) to check if the data has been successfully downloaded.

Tips

- Ensure that you have write permissions to the download path you specify.
- Verify the coordinates and dates to ensure they are within the valid range for the data cube you are accessing.
- If you encounter any errors, check the script documentation or help message by running:

bash
python **getDataFromCubeZone**.py --help

This will display detailed information about each argument and how to use them.

Tutorial: Download CSV Files Per Year Using getDataFromCubePoint.py

This tutorial will guide you through using the getDataFromCubePoint.py script to download wind speed data in CSV format for a specified range of years and per point. The script will fetch data for each year individually and save it to your specified download path.

Step 1: Understanding the Command

Here's the command we'll be working with:

cd /home/jovyan/agwise/AgWise_Data/getData_fromCube

hash

python getDataFromCubePoint.py --startDate "01-02" --endDate "05-30" --variable "agera5-windSpeed" --numberofyears 5 --lastYear 2023 --downloadpath "/tmp/" --x 2.269928 --y 2.269928

This command will download wind speed data in CSV format for each year from 2019 to 2023 (5 years), covering the period from January 2nd to May 30th each year. The data will be downloaded to the /tmp/ directory.

Step 2: Breakdown of Command Arguments

- --startDate "01-02": The start date of the period for which you want to download data (in MM-DD format).
- --endDate "05-30": The end date of the period for which you want to download data (in MM-DD format).
- --variable "agera5-windSpeed": The variable you are interested in. Here, it's wind speed from the agera5 dataset.
- --number of years of data you want to download.
- --lastYear 2023: The last year of the data range. The script will fetch data from 2023 back to 5 years.
- --downloadpath "/tmp/": The directory where the CSV files will be saved.
- --x 2.269928: The x-coordinate (longitude) of the location for which you want the data.
- --y 2.269928: The y-coordinate (latitude) of the location for which you want the data.

Step 3: Prerequisites

Before running the command, ensure you have the following prerequisites:

- 1. **Python Installed**: Make sure you have Python installed on your system.
- 2. **getDataFromCubePoint.py Script**: Ensure you have the getDataFromCubePoint.py script available in your working directory.

Step 4: Running the Command

1. **Open Terminal**: Open your terminal (Command Prompt, PowerShell, or any terminal application).

2. **Navigate to Script Location**: Use the cd command to navigate to the directory where your getDataFromCubePoint.py script is located.

bash cd path/to/your/script

Step 5: Verify the Downloaded Files

3. **Execute the Command**: Copy and paste the command into your terminal and press Enter.

bash Copiar código python getDataFromCubePoint.py --startDate "01-02" --endDate "05-30" --variable "agera5windSpeed" --numberofyears 5 --lastYear 2023 --downloadpath "/tmp/" --x 2.269928 --y 2.269928

1. **Navigate to the Download Path**: After the script completes, navigate to the /tmp/ directory (or the path you specified).

bash cd /tmp/

2. **Check the Files**: You should see CSV files for each year from 2019 to 2023 containing the wind speed data for the specified period.

Runnig the code using R

1. Using system commands in R:

This method involves calling the Python script from within your R environment using the system function.:

Steps:

 Construct the Python command: Create a string variable in R containing the Python script call with your desired arguments. Here's an example based on the provided code:

Fragmento de código

python_command <- "python **getDataFromCubeZone**.py --startDate='05-11' --endDate='08-30' --xmin=-90 --xmax=-83 --ymin=12 --ymax=16 --variable='chirps-precipitation' --downloadpath='/tmp/' --numberofyears=40"

2. **Execute the command:** Use the system function in R to run the constructed Python command.

system(python_command)

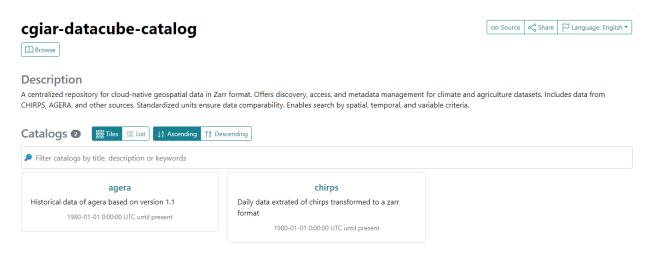
This will execute the Python script and download the data according to the specified arguments.

Guide: Finding Available Variables in the DataCube

This guide will walk you through the steps to identify available variables in the DataCube and how to use them in your Python scripts.

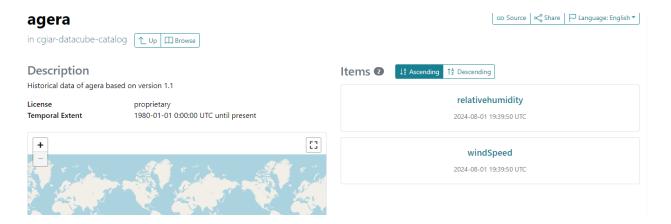
Step 1: Access the DataCube Catalog

 Open your web browser and navigate to the following URL: https://radiantearth.github.io/stac-browser/#/external/raw.githubusercontent.com/alliance-datascience/datacube-catalog/main/catalog/catalog.json



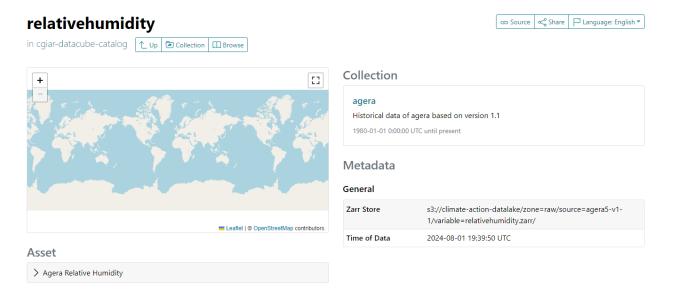
Step 2: Select Your Data Source

2. In the STAC Browser, you will see a list of data sources. Click on the source that interests you. This could be a dataset or a collection that contains the type of data you are looking for.



Step 3: Choose the Variable

After selecting the source, you will see a list of variables available within that source. Click on the variable you need to query. This will take you to a detailed view of that variable.



Step 4: View Collection, Metadata, and Asset

4. In the detailed view of the variable, you will see sections for Collection, Metadata, and Asset. Click on the **Asset** section. This section contains specific information about the variable.

Zarr



A measure of the amount of water vapor present in the air compared to the maximum amount of water vapor the air can hold at a specific temperature and pressure. Derived from Agera's dataset, this variable provides insights into atmospheric moisture conditions

Metadata

General

Authentication Enabled	yes
Source	https://cds.climate.copernicus.eu/cdsapp#!/dataset/sis- agrometeorological-indicators?tab=form
Update Frecuency	Monthy
Name R Integration	agera5-relative humidity
Resolution	0.1 x 0.1
Dimensions	- Lat: latitud Description: latitud of the data

Step 5: Identify the Variable Name

5. In the Asset section, look for the metadata field labeled **Name R Integration**. This is the name of the variable that you will use as a parameter in your Python scripts.

Name R agera5-relativehumidity
Integration