Wave & Wind Hindcast Data Downloader

This repository contains a Python script to download and process **ERA5 reanalysis data** using the **Climate Data Store (CDS) API** provided by ECMWF. The script:

- `download_era5_data.py` - Downloads **wave height, wave direction, peak wave period, wind speed, and wind´ direction**.

Program Options

The program offers two primary modes of operation:

1. **Download Mode:**
    - **Purpose:** Retrieves the raw hourly ERA5 reanalysis data directly from the CDS API.
    - **Functionality:**
        - Connects to the CDS API and downloads data in **GRIB format**.
        - Stores the downloaded raw GRIB files in the `grib/` folder.
        - Implements a robust retry mechanism with **exponential back-off** to handle API failures.
    - **When to Use:** Choose this option if you require the original, unprocessed dataset or need to re-run processing steps later.
2. **Processing Mode:**
    - **Purpose:** Processes the raw GRIB files into a structured format for analysis.
    - **Functionality:**
        - Extracts key meteorological and oceanographic variables such as significant wave height (`swh`), mean wave direction (`mwd`), peak wave period (`pp1d`), 10m wind speed (`wind`), and 10m wind direction (`dwi`).
        - Compiles the extracted data into a **CSV file** saved as `results/download_era5_data.csv`.
        - Utilizes logging (recorded in `download_era5_data.log`) to track processing steps and capture any errors.

- **When to Use:** Opt for this mode if you need the data in a readily analyzable CSV format or require processed data for further research.

Features

- Downloads hourly **ERA5 reanalysis** data.
- Retrieves selected meteorological and oceanographic variables.
- Uses a retry mechanism with **exponential back-off** to handle API failures.
- Saves processed data in **CSV format** for further analysis.
- Uses **logging** to record download and processing steps.

Files

File	Description
`download_era5_data.py`	Retrieves wind and ocean wave data.
`download_era5_data.log`	Log file storing execution details.
`grib/`	Folder for storing raw GRIB files.
`results/download_era5_data.csv`	Processed data in CSV format.

About the ERA5 Wave Model

This dataset is derived from the **ECMWF Reanalysis v5 (ERA5) wave model**, which provides hourly estimates of essential climate variables spanning from 1940 to the present. The ERA5 wave model is a component of the ERA5 dataset, developed by the **European Centre for Medium-Range Weather Forecasts (ECMWF)**.

ERA5 Wave Model Highlights:

- Uses **state-of-the-art** numerical weather prediction models and data assimilation techniques.
- Provides **hourly data** at a **31 km horizontal resolution** globally.
- Includes **wind-wave interactions, swell propagation, and wave generation** mechanisms.
- Incorporates **satellite observations, buoy measurements, and reanalysis techniques** to improve accuracy.
- Supplies a comprehensive **historical dataset** for research, operational forecasting, and climate applications.

More details can be found at:

- ERA5 Single Levels Dataset
- ECMWF ERA5 Overview

Installation

Install Dependencies

Ensure you have **Python 3.x** installed. Then install the required libraries using **Conda**:

```
conda install -c conda-forge eccodes cdsapi pygrib pandas tqdm
```

Alternatively, use **pip**:

```
pip install cdsapi pygrib pandas tqdm
```

Set Up CDS API Key

1. Register for an **ECMWF account** at: CDS Registration
2. Obtain your **API key** from: CDS API
3. Create a `.cdsapirc` file in your home directory (`~/.cdsapirc` on Linux/Mac, `C:\Users\YourName\.cdsapirc` on Windows):

```
url: https://cds.climate.copernicus.eu/api/v2
key: YOUR-USER-ID:YOUR-API-KEY
verify: 1
```

Usage

Run the Script

Use the following command to run the script:

```
python "download_era5_data.py"
```

Configurable Parameters

The script retrieves data for **Leixões Costeira, Porto (Portugal)** with coordinates **(41.31666°N, -8.983333°W)**. You can modify these values in the script:

```
LONGITUDE = -8.983333
LATITUDE = +41.31666
```

It downloads data from **1940 to 2025**. To change the time range, update:

```
START_YEAR = 1940
END_YEAR = 2025
```

Variables Retrieved

Variable	Short Name	Description
`swh`	140229	Significant height of combined wind waves and swell
`mwd`	140230	Mean wave direction
`pp1d`	140231	Peak wave period
`wind`	140245	10m wind speed
`dwi`	140249	10m wind direction

Data Storage

The downloaded data is stored in:

- **GRIB files** in the `grib/` folder.
- **Processed CSV data** in `results/download_era5_data.csv`.

A sample CSV row looks like:

```
datetime,swh,mwd,pp1d,wind,dwi
1940-01-01 00:00:00,2.5,280,8.0,5.2,220
```

References
