

---

**ecmwf***modelsDocumentation*  
**Release 0.9.1.post1.dev5+gc80c277.d20230222**

**TU Wien**

**Feb 22, 2023**



# CONTENTS

<b>1</b>	<b>Citation</b>	<b>3</b>
<b>2</b>	<b>Installation</b>	<b>5</b>
<b>3</b>	<b>Supported Products</b>	<b>7</b>
<b>4</b>	<b>Contribute</b>	<b>9</b>
<b>5</b>	<b>Downloading ERA5 Data</b>	<b>11</b>
<b>6</b>	<b>Downloading ERA Interim Data</b>	<b>13</b>
<b>7</b>	<b>Reading data</b>	<b>15</b>
<b>8</b>	<b>Conversion to time series format</b>	<b>17</b>
8.1	Reading converted time series data . . . . .	18
<b>9</b>	<b>Contents</b>	<b>19</b>
9.1	Downloading ERA5 Data . . . . .	19
9.2	Downloading ERA Interim Data . . . . .	20
9.3	Reading data . . . . .	20
9.4	Conversion to time series format . . . . .	21
9.5	Reading converted time series data . . . . .	22
9.6	License . . . . .	22
9.7	Developers . . . . .	23
9.8	Changelog . . . . .	23
9.9	ecmwf_models . . . . .	25
<b>10</b>	<b>Indices and tables</b>	<b>41</b>
	<b>Python Module Index</b>	<b>43</b>
	<b>Index</b>	<b>45</b>



Readers and converters for data from the [ECMWF reanalysis models](#). Written in Python.

Works great in combination with [pytesmo](#).



## CITATION

If you use the software in a publication then please cite it using the Zenodo DOI. Be aware that this badge links to the latest package version.

Please select your specific version at <https://doi.org/10.5281/zenodo.593533> to get the DOI of that version. You should normally always use the DOI for the specific version of your record in citations. This is to ensure that other researchers can access the exact research artefact you used for reproducibility.

You can find additional information regarding DOI versioning at <http://help.zenodo.org/#versioning>





## INSTALLATION

Install required C-libraries via conda. For installation we recommend [Miniconda](#). So please install it according to the official installation instructions. As soon as you have the conda command in your shell you can continue:

```
conda install -c conda-forge pandas pygrib netcdf4 pyresample xarray
```

The following command will download and install all the needed pip packages as well as the ecmwf-model package itself.

```
pip install ecmwf_models
```

To create a full development environment with conda, the *yml* files inside the folder *environment/* in this repository can be used. Both environments should work. The file *latest* should install the newest version of most dependencies. The file *pinned* is a fallback option and should always work.

```
git clone --recursive git@github.com:TUW-GEO/ecmwf_models.git ecmwf_models
cd ecmwf_models
conda env create -f environment/latest.yml
source activate ecmwf_models
python setup.py develop
pytest
```



## SUPPORTED PRODUCTS

At the moment this package supports

- **ERA Interim** (deprecated)
- **ERA5**
- **ERA5-Land**

reanalysis data in **grib** and **netcdf** format (download, reading, time series creation) with a default spatial sampling of 0.75 degrees (ERA Interim), 0.25 degrees (ERA5), resp. 0.1 degrees (ERA5-Land). It should be easy to extend the package to support other ECMWF reanalysis products. This will be done as need arises.



## CONTRIBUTE

We are happy if you want to contribute. Please raise an issue explaining what is missing or if you find a bug. Please take a look at the [developers guide](#).



## DOWNLOADING ERA5 DATA

ERA5 (and ERA5-Land) data can be downloaded manually from the [Copernicus Data Store \(CDS\)](#) or automatically via the CDS api, as done in the download module (`era5_download`). Before you can use this, you have to set up an [account at the CDS](#) and setup the [CDS key](#).

Then you can use the program `era5_download` to download ERA5 images between a passed start and end date. `era5_download --help` will show additional information on using the command.

For example, the following command in your terminal would download ERA5 images for all available layers of soil moisture in netcdf format, between January 1st and February 1st 2000 in grib format into `/path/to/storage`. The data will be stored in subfolders of the format `YYYY/jjj`. The temporal resolution of the images is 6 hours by default.

```
era5_download /path/to/storage -s 2000-01-01 -e 2000-02-01 --variables swv11 swv12 swv13 ↵  
↵swv14
```

The names of the variables to download can be its long names, the short names (as in the example) or the parameter IDs. We use the `era5_lut.csv` file to look up the right name for the CDS API. Other flags, that can be activated in `era5_download` are:

- **-h** (**-help**) : shows the help text for the download function
- **-p** (**-product**): specify the ERA5 product to download. Choose either ERA5 or ERA5-Land. Default is ERA5.
- **-keep** (**-keep\_original**) : keeps the originally downloaded files as well. We split the downloaded, monthly stacks into single images and discard the original files by default.
- **-grib** (**-as\_grib**)  
[download the data in grib format instead of the default nc4] format (grib reading is not supported on Windows OS).
- **-h\_steps** : full hours for which images are downloaded (e.g. `-h_steps 0` would download only data at 00:00 UTC). By default we use 0, 6, 12 and 18.





## DOWNLOADING ERA INTERIM DATA

**ERA-Interim has been decommissioned. Use ERA5 instead.**

ERA-Interim data can be downloaded manually from the ECMWF servers. It can also be done automatically using the ECMWF API. To use the ECMWF API you have to be registered, install the `ecmwf-api` Python package and setup the ECMWF API Key. A guide for this is provided by [ECMWF](#).

After that you can use the command line program `eraint_download` to download images with a temporal resolution of 6 hours between a passed start and end date. `eraint_download --help` will show additional information on using the command.

For example, the following command in your terminal would download ERA Interim soil moisture images of all available layers (see the [Variable DB](#)) in netcdf format on the default gaussian grid for ERA-Interim ( $0.75^\circ \times 0.75^\circ$ ) into the folder `/path/to/storage` between January 1st and February 1st 2000. The data will be stored in subfolders of the format `YYYY/jjj`, where `YYYY` describes the year and `jjj` the day of the year for the downloaded files.

```
eraint_download /path/to/storage -s 2000-01-01 -e 2000-02-01 --variables swv11 swv12 ↵  
↵swv13 swv14
```

Additional optional parameters allow downloading images in netcdf format, and in a different spatial resolution (see the `-help` function and descriptions for downloading ERA5 data)



## READING DATA

After downloading the data for ERA Interim or ERA5 via `eraint_download` resp. `era5_download`, images can be read with the `ERA5GrbDs` and `ERA5NcDs` (for grib and netcdf image stacks), respectively the `ERA5GrbImg` and `ERA5NcImg` (for single grib and netcdf images) classes. The respective functions for reading images are defined in `ecmwf_models.erainterim.interface` `ecmwf_models.era5.interface`.

The following examples are shown for ERA5 data, but work the same way with the respective ERA Interim functions.

For example, you can read the image for a single variable at a specific date. In this case for a stack of downloaded image files:

```
# Script to load a stack of downloaded netcdf images
# and read a variable for a single date.
from ecmwf_models.era5.interface import ERA5NcDs
root_path = "/path/to/netcdf_storage"
ds = ERA5NcDs(root_path, parameter='swvl1')
data = ds.read(datetime(2010, 1, 1, 0))

# Script to load a stack of downloaded grib images
# and read a variable for a single date.
from ecmwf_models.era5.interface import ERA5GrbDs
root_path = "/path/to/grib_storage"
ds = ERA5GrbDs(root_path, parameter='swvl1')
data = ds.read(datetime(2010, 1, 1, 0))
```

You can also read multiple variables at a specific date by passing a list of parameters. In this case for a set of netcdf files:

```
# Script to load a stack of downloaded netcdf images
# and read two variables for a single date.
from ecmwf_models.era5.interface import ERA5NcDs
root_path = "/path/to/storage"
ds = ERA5NcDs(root_path, parameter=['swvl1', 'swvl2'])
data = ds.read(datetime(2000, 1, 1, 0))
```

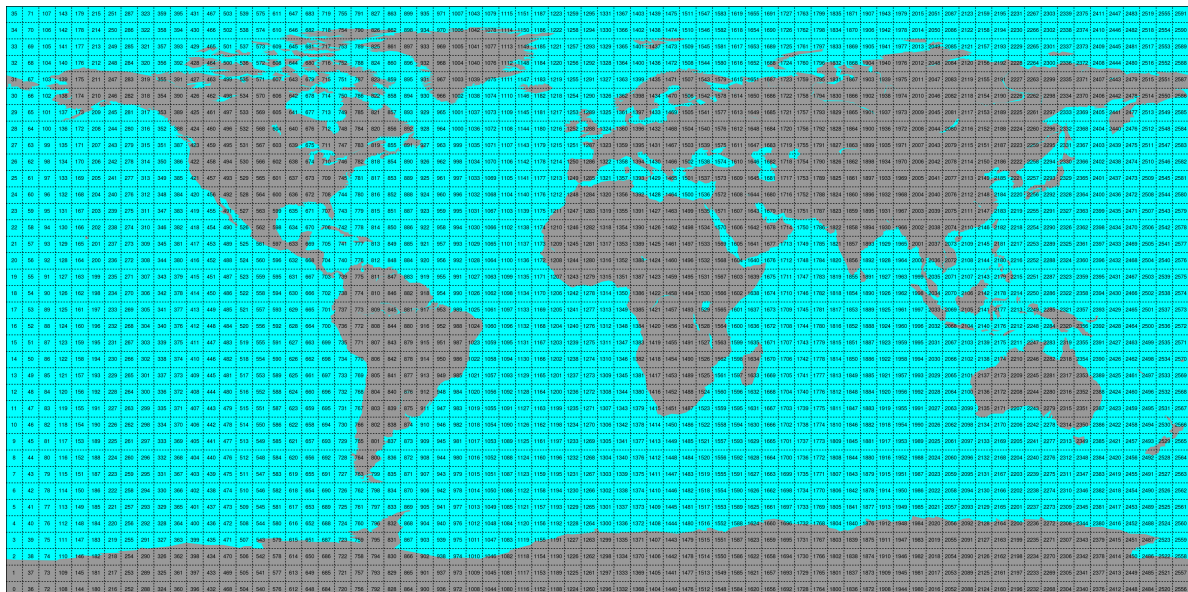
All images between two given dates can be read using the `iter_images` methods of all the image stack reader classes.



## CONVERSION TO TIME SERIES FORMAT

For a lot of applications it is favorable to convert the image based format into a format which is optimized for fast time series retrieval. This is what we often need for e.g. validation studies. This can be done by stacking the images into a netCDF file and choosing the correct chunk sizes or a lot of other methods. We have chosen to do it in the following way:

- Store only the reduced gaussian grid points (for grib data) since that saves space.
- Store the time series in netCDF4 in the Climate and Forecast convention **Orthogonal multidimensional array representation**
- Store the time series in 5x5 degree cells. This means there will be 2566 cell files and a file called `grid.nc` which contains the information about which grid point is stored in which file. This allows us to read a whole 5x5 degree area into memory and iterate over the time series quickly.



This conversion can be performed using the `era5_reshuffle` (respectively `eraint_reshuffle`) command line program. An example would be:

```
era5_reshuffle /era_data /timeseries/data 2000-01-01 2001-01-01 swl1 swl2
```

Which would take 6-hourly ERA5 images stored in `/era_data` from January 1st 2000 to January 1st 2001 and store the parameters “swl1” and “swl2” as time series in the folder `/timeseries/data`. If you time series should have a different resolution than 6H, use the `h_steps` flag here accordingly (images to use for time series generation have to be in the downloaded raw data). The passed names have to correspond with the names in the downloaded file, i.e. use the variable short names here. Other flags, that can be used in `era5_reshuffle` are:

- **-h (-help)** : Shows the help text for the reshuffle function
- **-land\_points** : Reshuffle and store only data over land points.
- **-h\_steps (-as\_grib)** : full hours for which images are reshuffled (e.g. `-h_steps 0` would reshuffle only data at 00:00 UTC). By default we use 0, 6, 12 and 18.
- **-imgbuffer** : The number of images that are read into memory before converting them into time series. Bigger numbers make the conversion faster but consume more memory.

Conversion to time series is performed by the [repurpose](#) package in the background. For custom settings or other options see the [repurpose documentation](#) and the code in `ecmwf_models.reshuffle`.

## 8.1 Reading converted time series data

For reading time series data, that the `era5_reshuffle` and `eraint_reshuffle` command produces, the class `ERATs` can be used. Optional arguments that are passed to the parent class (`OrthoMultiTs`, as defined in `pynetcf.time_series`) can be passed as well:

```
from ecmwf_models import ERATs
# read_bulk reads full files into memory
# read_ts takes either lon, lat coordinates to perform a nearest neighbour search
# or a grid point index (from the grid.nc file) and returns a pandas.DataFrame.
ds = ERATs(ts_path, ioclass_kws={'read_bulk': True})
ts = ds.read_ts(45, 15)
```

Bulk reading speeds up reading multiple points from a cell file by storing the file in memory for subsequent calls. Either Longitude and Latitude can be passed to perform a nearest neighbour search on the data grid (`grid.nc` in the time series path) or the grid point index (GPI) can be passed directly.

## CONTENTS

## 9.1 Downloading ERA5 Data

ERA5 (and ERA5-Land) data can be downloaded manually from the [Copernicus Data Store \(CDS\)](#) or automatically via the CDS api, as done in the download module (`era5_download`). Before you can use this, you have to set up an account at the CDS and setup the CDS key.

Then you can use the program `era5_download` to download ERA5 images between a passed start and end date. `era5_download --help` will show additional information on using the command.

For example, the following command in your terminal would download ERA5 images for all available layers of soil moisture in netcdf format, between January 1st and February 1st 2000 in grib format into `/path/to/storage`. The data will be stored in subfolders of the format `YYYY/jjj`. The temporal resolution of the images is 6 hours by default.

```
era5_download /path/to/storage -s 2000-01-01 -e 2000-02-01 --variables swv11 swv12 swv13 ↵  
↵swv14
```

The names of the variables to download can be its long names, the short names (as in the example) or the parameter IDs. We use the `era5_lut.csv` file to look up the right name for the CDS API. Other flags, that can be activated in `era5_download` are:

- **-h (-help)** : shows the help text for the download function
- **-p (-product)**: specify the ERA5 product to download. Choose either ERA5 or ERA5-Land. Default is ERA5.
- **-keep (-keep\_original)** : keeps the originally downloaded files as well. We split the downloaded, monthly stacks into single images and discard the original files by default.
- **-grib (-as\_grib)**  
[download the data in grib format instead of the default nc4] format (grib reading is not supported on Windows OS).
- **-h\_steps** : full hours for which images are downloaded (e.g. `-h_steps 0` would download only data at 00:00 UTC). By default we use 0, 6, 12 and 18.

## 9.2 Downloading ERA Interim Data

**ERA-Interim has been decommissioned. Use ERA5 instead.**

ERA-Interim data can be downloaded manually from the ECMWF servers. It can also be done automatically using the ECMWF API. To use the ECMWF API you have to be registered, install the `ecmwf-api` Python package and setup the ECMWF API Key. A guide for this is provided by [ECMWF](#).

After that you can use the command line program `eraint_download` to download images with a temporal resolution of 6 hours between a passed start and end date. `eraint_download --help` will show additional information on using the command.

For example, the following command in your terminal would download ERA Interim soil moisture images of all available layers (see the [Variable DB](#)) in netcdf format on the default gaussian grid for ERA-Interim (0.75°x0.75°) into the folder `/path/to/storage` between January 1st and February 1st 2000. The data will be stored in subfolders of the format `YYYY/jjj`, where `YYYY` describes the year and `jjj` the day of the year for the downloaded files.

```
eraint_download /path/to/storage -s 2000-01-01 -e 2000-02-01 --variables swvl1 swvl2,
↪swvl3 swvl4
```

Additional optional parameters allow downloading images in netcdf format, and in a different spatial resolution (see the `-help` function and descriptions for downloading ERA5 data)

## 9.3 Reading data

After downloading the data for ERA Interim or ERA5 via `eraint_download` resp. `era5_download`, images can be read with the `ERA5GrbDs` and `ERA5NcDs` (for grib and netcdf image stacks), respectively the `ERA5GrbImg` and `ERA5NcImg` (for single grib and netcdf images) classes. The respective functions for reading images are defined in `ecmwf_models.erainterim.interface` `ecmwf_models.era5.interface`.

The following examples are shown for ERA5 data, but work the same way with the respective ERA Interim functions.

For example, you can read the image for a single variable at a specific date. In this case for a stack of downloaded image files:

```
# Script to load a stack of downloaded netcdf images
# and read a variable for a single date.
from ecmwf_models.era5.interface import ERA5NcDs
root_path = "/path/to/netcdf_storage"
ds = ERA5NcDs(root_path, parameter='swvl1')
data = ds.read(datetime(2010, 1, 1, 0))

# Script to load a stack of downloaded grib images
# and read a variable for a single date.
from ecmwf_models.era5.interface import ERA5GrbDs
root_path = "/path/to/grib_storage"
ds = ERA5GrbDs(root_path, parameter='swvl1')
data = ds.read(datetime(2010, 1, 1, 0))
```

You can also read multiple variables at a specific date by passing a list of parameters. In this case for a set of netcdf files:

```
# Script to load a stack of downloaded netcdf images
# and read two variables for a single date.
```

(continues on next page)



(continued from previous page)

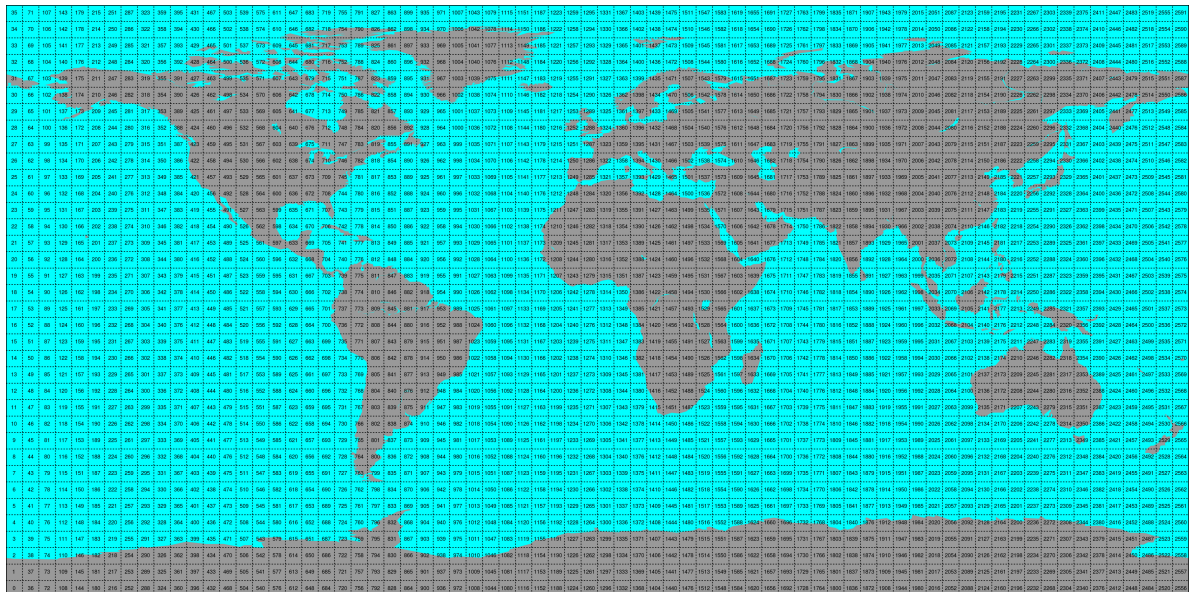
```
from ecmwf_models.era5.interface import ERA5NcDs
root_path = "/path/to/storage"
ds = ERA5NcDs(root_path, parameter=['swvl1', 'swvl2'])
data = ds.read(datetime(2000, 1, 1, 0))
```

All images between two given dates can be read using the `iter_images` methods of all the image stack reader classes.

## 9.4 Conversion to time series format

For a lot of applications it is favorable to convert the image based format into a format which is optimized for fast time series retrieval. This is what we often need for e.g. validation studies. This can be done by stacking the images into a netCDF file and choosing the correct chunk sizes or a lot of other methods. We have chosen to do it in the following way:

- Store only the reduced gaussian grid points (for grib data) since that saves space.
- Store the time series in netCDF4 in the Climate and Forecast convention **Orthogonal multidimensional array representation**
- Store the time series in 5x5 degree cells. This means there will be 2566 cell files and a file called `grid.nc` which contains the information about which grid point is stored in which file. This allows us to read a whole 5x5 degree area into memory and iterate over the time series quickly.



This conversion can be performed using the `era5_reshuffle` (respectively `eraint_reshuffle`) command line program. An example would be:

```
era5_reshuffle /era_data /timeseries/data 2000-01-01 2001-01-01 swvl1 swvl2
```

Which would take 6-hourly ERA5 images stored in `/era_data` from January 1st 2000 to January 1st 2001 and store the parameters “swvl1” and “swvl2” as time series in the folder `/timeseries/data`. If you time series should have a different resolution than 6H, use the `h_steps` flag here accordingly (images to use for time series generation have to be in the downloaded raw data). The passed names have to correspond with the names in the downloaded file, i.e. use the variable short names here. Other flags, that can be used in `era5_reshuffle` are:

- **-h (-help)** : Shows the help text for the reshuffle function
- **-land\_points** : Reshuffle and store only data over land points.
- **-h\_steps (-as\_grib)** : full hours for which images are reshuffled (e.g. `-h_steps 0` would reshuffle only data at 00:00 UTC). By default we use 0, 6, 12 and 18.
- **-imgbuffer** : The number of images that are read into memory before converting them into time series. Bigger numbers make the conversion faster but consume more memory.

Conversion to time series is performed by the [repurpose](#) package in the background. For custom settings or other options see the [repurpose documentation](#) and the code in `ecmwf_models.reshuffle`.

## 9.5 Reading converted time series data

For reading time series data, that the `era5_reshuffle` and `eraint_reshuffle` command produces, the class `ERATs` can be used. Optional arguments that are passed to the parent class (`OrthoMultiTs`, as defined in `pynetcf.time_series`) can be passed as well:

```
from ecmwf_models import ERATs
# read_bulk reads full files into memory
# read_ts takes either lon, lat coordinates to perform a nearest neighbour search
# or a grid point index (from the grid.nc file) and returns a pandas.DataFrame.
ds = ERATs(ts_path, ioclass_kws={'read_bulk': True})
ts = ds.read_ts(45, 15)
```

Bulk reading speeds up reading multiple points from a cell file by storing the file in memory for subsequent calls. Either Longitude and Latitude can be passed to perform a nearest neighbour search on the data grid (`grid.nc` in the time series path) or the grid point index (GPI) can be passed directly.

## 9.6 License

The MIT License (MIT)

Copyright (c) 2021 TU Wien, Department of Geodesy and Geoinformation

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the “Software”), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED “AS IS”, WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

## 9.7 Developers

- Sebastian Hahn <sebastian.hahn@geo.tuwien.ac.at>
- Christoph Paulik <cpaulik@vandersat.com>
- Wolfgang Preimesberger <wolfgang.preimesberger@geo.tuwien.ac.at>

## 9.8 Changelog

### 9.8.1 Unreleased changes in master branch

- Fix issue with too large data requests to download, added command line arg to specify max request size.

### 9.8.2 Version 0.9.1

- Fix returned status code in case of partial data availability.

### 9.8.3 Version 0.9

- ERA5T images are downloaded and sliced correctly
- Code formatting with yapf

### 9.8.4 Version 0.8

- Program *era5\_download* returns exit code now (PR #27);
- Program *era5\_resuffle* can now take a bounding box to reshuffle spatial subsets;
- TravisCI was replaced by Github Actions;
- Pyscaffold 4 is used; contributing guide added; pre-commit added;
- Code formatting with black (line length 79);

### 9.8.5 Version 0.7

- Update pyscaffold structure
- Drop support for python2
- Travis deploy to pypi

### 9.8.6 Version 0.6.1

- Fix bug when creating 0.1 deg grid cells (floating point precision)
- Missing variables in grib files are now replaced by empty images.
- Read variable names from grib files from cfVarNameECMF instead of short\_name field

### 9.8.7 Version 0.6

- Add support for downloading, reading, reshuffling era5-land
- Add support for reading, reshuffling points over land only (era5 and era5-land)
- Add function to create land definition files
- Test with pinned environments

### 9.8.8 Version 0.5

- Change default time steps to 6 hours.
- Add more tests, also for download functions
- Update documentation, add installation script
- Fix bugs, update command line interfaces, update dependencies
- Separate download programs for ERA5 and ERA Interim
- Change the ERA5 download api to use cdsapi instead of ecmwf api
- Update package structure to better separate between the ERA products
- Add look-up-table file for more flexibility in variable names passed by user
- Update readme

### 9.8.9 Version 0.4

- Add ERA5 support (download, reading, TS conversion)
- Add netcdf support for ERA5 and ERA-Interim download (regular grid)
- Add new grid defintions: netcdf download in regular grid, grib in gaussian grid
- Add Download with spatial resampling (grib and nc)
- Update GRIB message storing (per day instead of per message)
- Add tests for splitting downloaded files, ERA5 reading, ERA5 reshuffling, generated grids
- Add new test data

### 9.8.10 Version 0.3

- Fix help text in ecmwf\_repurpose command line program.
- Fix reading of metadata for variables that do not have 'levels'
- Fix wrong import when trying to read the reformatted time series data.

### 9.8.11 Version 0.2

- Add reading of basic metadata fields name, depth and units.
- Fix reading of latitudes and longitudes - where flipped before.
- Fix longitude range to -180, 180.
- Add conversion to time series format.

### 9.8.12 Version 0.1

- First version
- Add ERA Interim support for downloading and reading.

## 9.9 ecmwf\_models

### 9.9.1 ecmwf\_models package

#### Subpackages

#### ecmwf\_models.era5 package

#### Submodules

#### ecmwf\_models.era5.download module

Module to download ERA5 from terminal in netcdf and grib format.

```
class ecmwf_models.era5.download.CDSStatusTracker
```

```
    Bases: object
```

```
    Track the status of the CDS download by using the CDS callback functions
```

```
    handle_error_function(*args, **kwargs)
```

```
    statuscode_error = -1
```

```
    statuscode_ok = 0
```

```
    statuscode_unavailable = 10
```

`ecmwf_models.era5.download.default_variables(product='era5')`

These variables are being downloaded, when None are passed by the user

#### Parameters

**product** (*str*, optional (default: 'era5')) – Name of the era5 product to read the default variables for. Either 'era5' or 'era5-land'.

`ecmwf_models.era5.download.download_and_move(target_path, startdate, enddate, product='era5', variables=None, keep_original=False, h_steps=(0, 6, 12, 18), grb=False, dry_run=False, grid=None, remap_method='bil', cds_kwds={}, stepsize='month', n_max_request=1000) → int`

Downloads the data from the ECMWF servers and moves them to the target path. This is done in 30 day increments between start and end date.

The files are then extracted into separate grib files per parameter and stored in yearly folders under the target\_path.

#### Parameters

- **target\_path** (*str*) – Path where the files are stored to
- **startdate** (*datetime*) – first date to download
- **enddate** (*datetime*) – last date to download
- **product** (*str*, optional (default: ERA5)) – Either ERA5 or ERA5Land
- **variables** (*list*, optional (default: None)) – Name of variables to download
- **keep\_original** (*bool* (default: False)) – keep the original downloaded data
- **h\_steps** (*list or tuple*) – List of full hours to download data at the selected dates e.g [0, 12]
- **grb** (*bool*, optional (default: False)) – Download data as grib files
- **dry\_run** (*bool*) – Do not download anything, this is just used for testing the functions
- **grid** (*dict*, optional (default: None)) – A grid on which to remap the data using CDO. This must be a dictionary using CDO's grid description format, e.g.:

```
grid = {
    "gridtype": "lonlat",
    "xsize": 720,
    "ysize": 360,
    "xfirst": -179.75,
    "yfirst": 89.75,
    "xinc": 0.5,
    "yinc": -0.5,
}
```

Default is to use no regridding.

- **remap\_method** (*str*, optional (default: 'bil')) – Method to be used for regridding. Available methods are: - "bil": bilinear (default) - "bic": bicubic - "nn": nearest neighbour - "dis": distance weighted - "con": 1st order conservative remapping - "con2": 2nd order conservative remapping - "laf": largest area fraction remapping
- **cds\_kwds** (*dict*, optional (default: {})) – Additional arguments to be passed to the CDS API retrieve request.

- **n\_max\_request** (*int, optional (default: 1000)*) – Maximum size that a request can have to be processed by CDS. At the moment of writing this is 1000 (N\_timestamps \* N\_variables in a request) but as this is a server side settings, it can change.

**Returns**

**status\_code** – 0 : Downloaded data ok -1 : Error -10 : No data available for requested time period

**Return type**

`int`

`ecmwf_models.era5.download.download_era5(c, years, months, days, h_steps, variables, target, grb=False, product='era5', dry_run=False, cds_kwds={})`

Download era5 reanalysis data for single levels of a defined time span

**Parameters**

- **c** (*cdsapi.Client*) – Client to pass the request to
- **years** (*list*) – Years for which data is downloaded ,e.g. [2017, 2018]
- **months** (*list*) – Months for which data is downloaded, e.g. [4, 8, 12]
- **days** (*list*) – Days for which data is downloaded (range(31)=All days) e.g. [10, 20, 31]
- **h\_steps** (*list*) – List of full hours to download data at the selected dates e.g [0, 12]
- **variables** (*list, optional (default: None)*) – List of variables to pass to the client, if None are passed, the default variables will be downloaded.
- **target** (*str*) – File name, where the data is stored.
- **geb** (*bool, optional (default: False)*) – Download data in grib format
- **product** (*str*) – ERA5 data product to download, either era5 or era5-land
- **dry\_run** (*bool, optional (default: False)*) – Do not download anything, this is just used for testing the functionality
- **cds\_kwds** (*dict, optional*) – Additional arguments to be passed to the CDS API retrieve request.

**Returns**

**success** – Return True after downloading finished

**Return type**

`bool`

`ecmwf_models.era5.download.main(args)`

`ecmwf_models.era5.download.parse_args(args)`

Parse command line parameters for recursive download

**Parameters**

**args** (*list*) – Command line parameters as list of strings

**Returns**

**clparams** – Parsed command line parameters

**Return type**

`argparse.Namespace`

`ecmwf_models.era5.download.run()`

`ecmwf_models.era5.download.split_array(array, chunk_size)`

Split an array into chunks of a given size.

**Parameters**

- **array** (*array-like*) – Array to split into chunks
- **chunk\_size** (*int*) – Size of each chunk

**Returns**

**chunks** – List of chunks

**Return type**

list

`ecmwf_models.era5.download.split_chunk(timestamps, n_vars, n_hsteps, max_req_size=1000, reduce=False, daily_request=False)`

Split the passed time stamps into chunks for a valid request. One chunk can at most hold data for one month or one day, but cannot be larger than the maximum request size.

**Parameters**

- **timestamps** (*pd.DatetimeIndex*) – List of daily timestamps to split into chunks
- **n\_vars** (*int*) – Number of variables in each request.
- **max\_req\_size** (*int, optional (default: 1000)*) – Maximum size of a request that the CDS API can handle
- **reduce** (*bool, optional (default: False)*) – Return only the start and end of each subperiod instead of all time stamps.
- **daily\_request** (*bool, optional (default: False)*) – Only submit daily requests, otherwise monthly requests are allowed (if the max\_req\_size is not reached).

**Returns**

**chunks** – List of start and end dates that contain a chunk that the API can handle.

**Return type**

list

## ecmwf\_models.era5.interface module

This module contains ERA5/ERA5-Land specific child classes of the netcdf and grib base classes, that are used for reading all ecmwf products.

```
class ecmwf_models.era5.interface.ERA5GrbDs(root_path: str, parameter: Collection[str] = ('swvl1', 'swvl2'), h_steps: Collection[int] = (0, 6, 12, 18), product: Literal['era5', 'era5-land'] = 'era5', subgrid: Optional[CellGrid] = None, mask_seapoints: Optional[bool] = False, array_1D: Optional[bool] = False)
```

Bases: [ERAGrbDs](#)

```
class ecmwf_models.era5.interface.ERA5GrbImg(filename: str, parameter: Optional[Collection[str]] = ('swvl1', 'swvl2'), subgrid: Optional[CellGrid] = None, mask_seapoints: Optional[bool] = False, array_1D=False)
```

Bases: [ERAGrbImg](#)



```
class ecmwf_models.era5.interface.ERA5NCDS(root_path: str, parameter: Collection[str] = ('swvl1',
                                         'swvl2'), product: Literal['era5', 'era5-land'] = 'era5',
                                         h_steps: Collection[int] = (0, 6, 12, 18), subgrid:
                                         Optional[CellGrid] = None, mask_seapoints:
                                         Optional[bool] = False, array_1D: Optional[bool] = False))
```

Bases: [ERANCDS](#)

Reader for a stack of ERA5 netcdf image files.

#### Parameters

- **root\_path** (*str*) – Path to the image files to read.
- **parameter** (*list or str, optional (default: ('swvl1', 'swvl2'))*) – Name of parameters to read from the image file.
- **product** (*str, optional (default: 'era5')*) – What era5 product, either era5 or era5-land.
- **h\_steps** (*list, optional (default: (0, 6, 12, 18))*) – List of full hours to read images for.
- **subgrid** (*pygeogrids.CellGrid, optional (default: None)*) – Read only data for points of this grid and not global values.
- **mask\_seapoints** (*bool, optional (default: False)*) – Read the land-sea mask to mask points over water and set them to nan. This option needs the ‘lsm’ parameter to be in the file!
- **array\_1D** (*bool, optional (default: False)*) – Read data as list, instead of 2D array, used for reshuffling.

```
class ecmwf_models.era5.interface.ERA5NCImg(filename: str, parameter: Optional[Collection[str]] =
                                         ('swvl1', 'swvl2'), product: Literal['era5', 'era5-land'] =
                                         'era5', subgrid: Optional[CellGrid] = None,
                                         mask_seapoints: Optional[bool] = False, array_1D:
                                         Optional[bool] = False))
```

Bases: [ERANCImg](#)

## ecmwf\_models.era5.reshuffle module

Module for a command line interface to convert the ERA Interim data into a time series format using the repurpose package

```
ecmwf_models.era5.reshuffle.main(args)
```

```
ecmwf_models.era5.reshuffle.parse_args(args)
```

Parse command line parameters for conversion from image to time series.

#### Parameters

**args** (*list*) – command line parameters as list of strings

#### Returns

**args** – Parsed command line parameters

#### Return type

[argparse.Namespace](#)

```
ecmwf_models.era5.reshuffle.reshuffle(input_root, outputpath, startdate, enddate, variables,
                                     product=None, bbox=None, h_steps=(0, 6, 12, 18),
                                     land_points=False, imgbuffer=50)
```

Reshuffle method applied to ERA images for conversion into netcdf time series format.

### Parameters

- **input\_root** (*str*) – Input path where ERA image data was downloaded to.
- **outputpath** (*str*) – Output path, where the reshuffled netcdf time series are stored.
- **startdate** (*datetime*) – Start date, from which images are read and time series are generated.
- **enddate** (*datetime*) – End date, from which images are read and time series are generated.
- **variables** (*tuple or list or str*) – Variables to read from the passed images and convert into time series format.
- **product** (*str, optional (default: None)*) – Either era5 or era5-land, if None is passed we guess the product from the downloaded image files.
- **bbox** (*tuple optional (default: None)*) – (min\_lon, min\_lat, max\_lon, max\_lat) - wgs84. To load only a subset of the global grid / file.
- **h\_steps** (*list or tuple, optional (default: (0, 6, 12, 18))*) – Hours at which images are read for each day and used for reshuffling, therefore this defines the sub-daily temporal resolution of the time series that are generated.
- **land\_points** (*bool, optional (default: False)*) – Reshuffle only land points. Uses the ERA5 land mask to create a land grid. The land grid is fixed to 0.25\*0.25 or 0.1\*0.1 deg for now.
- **imgbuffer** (*int, optional (default: 200)*) – How many images to read at once before writing time series. This number affects how many images are stored in memory and should be chosen according to the available amount of memory and the size of a single image.

```
ecmwf_models.era5.reshuffle.run()
```

## Module contents

### ecmwf\_models.erainterim package

#### Submodules

#### ecmwf\_models.erainterim.download module

Module to download ERA Interim from terminal.

```
ecmwf_models.erainterim.download.default_variables() → list
```

These variables are being downloaded, when None are passed by the user

```
ecmwf_models.erainterim.download.download_and_move(target_path, startdate, enddate, variables=None,
                                                    keep_original=False, grid_size=None, type='an',
                                                    h_steps=(0, 6, 12, 18), steps=(0,), grb=False,
                                                    dry_run=False)
```

Downloads the data from the ECMWF servers and moves them to the target path. This is done in 30 days increments between start and end date to be efficient with the MARS system. See the recommendation for doing it this way in <https://software.ecmwf.int/wiki/display/WEBAPI/ERA-Interim+daily+retrieval+efficiency>

The files are then extracted into separate grib/nc files and stored in yearly folders under the target\_path.

#### Parameters

- **target\_path** (*str*) – Path to which to copy the extracted parameter files
- **startdate** (*datetime*) – First date to download
- **enddate** (*datetime*) – Last date to download
- **variables** (*list, optional (default: None)*) – List of variable ids to pass to the client, if None are passed, the default variable ids will be downloaded.
- **keep\_original** (*bool, optional (default: False)*) – Keep the original downloaded data
- **grid\_size** (*list, optional (default: None)*) – [lon, lat] extent of the grid (regular for netcdf, at lat=0 for grib) If None is passed, the default grid size for the data product is used.
- **type** (*str, optional (default: 'an')*) – Data stream, model to download data for (fc=forecast)
- **h\_steps** (*list, optional (default: [0, 6, 12, 18])*) – List of full hours to download data at the selected dates
- **grb** (*bool, optional (default: False)*) – Download data as grib files instead of netcdf files
- **dry\_run** (*bool*) – Do not download anything, this is just used for testing the functions

```
ecmwf_models.erainterim.download.download_eraint(target_path, start, end, variables, grid_size=None,
                                                type='fc', h_steps=(0, 6, 12, 18), grb=False,
                                                dry_run=False, steps=(0,))
```

Download era interim data

#### Parameters

- **target\_path** (*str*) – path at which to save the downloaded grib file
- **start** (*date*) – start date
- **end** (*date*) – end date
- **variables** (*list*) – parameter ids, see wiki
- **product** (*str, optional*) – Name of the model, “ERA-interim” (default) or “ERA5”
- **grid\_size** (*[float, float], optional*) – size of the grid in form (lon, lat), which the data is resampled to If None is passed the minimum grid for the accoring product is chosen
- **h\_steps** (*tuple, optional (default: (0, 6, 12, 18))*) – List of full hours to download data at the selected dates
- **grb** (*bool, optional (default: False)*) – Download data as grb files instead of nc files
- **dry\_run** (*bool*) – Do not download anything, this is just used for testing the functions

```
ecmwf_models.erainterim.download.main(args)
```

`ecmwf_models.erainterim.download.parse_args(args)`

Parse command line parameters for recursive download

**Parameters**

**args** (*list*) – Command line parameters as list of strings

**Returns**

**clparams** – Parsed command line parameters

**Return type**

`argparse.Namespace`

`ecmwf_models.erainterim.download.run()`

## ecmwf\_models.erainterim.interface module

This module contains ERA Interim specific child classes of the netcdf and grib base classes, that are used for reading all ecmwf products.

**class** `ecmwf_models.erainterim.interface.ERAIntGrbDs`(*root\_path: str, parameter: Collection[str] = ('swvl1', 'swvl2'), subgrid: Optional[CellGrid] = None, mask\_seapoints: Optional[bool] = False, h\_steps: Collection[int] = (0, 6, 12, 18), array\_1D: Optional[bool] = False*)

Bases: `ERAGrbDs`

**class** `ecmwf_models.erainterim.interface.ERAIntGrbImg`(*filename: str, parameter: Optional[Collection[str]] = ('swvl1', 'swvl2'), mode: Optional[str] = 'r', subgrid: Optional[CellGrid] = None, mask\_seapoints: Optional[bool] = False, array\_1D: Optional[bool] = False*)

Bases: `ERAGrbImg`

**class** `ecmwf_models.erainterim.interface.ERAIntNcDs`(*root\_path: str, parameter: Optional[Collection[str]] = ('swvl1', 'swvl2'), subgrid: Optional[CellGrid] = None, mask\_seapoints: Optional[bool] = False, h\_steps: Collection[int] = (0, 6, 12, 18), array\_1D: Optional[bool] = False*)

Bases: `ERANcDs`

**class** `ecmwf_models.erainterim.interface.ERAIntNcImg`(*filename: str, parameter: Optional[Collection[str]] = ('swvl1', 'swvl2'), mode: Optional[str] = 'r', subgrid: Optional[CellGrid] = None, mask\_seapoints: Optional[bool] = False, array\_1D: Optional[bool] = False*)

Bases: `ERANcImg`

## ecmwf\_models.erainterim.reshuffle module

Module for a command line interface to convert the ERA Interim data into a time series format using the repurpose package

`ecmwf_models.erainterim.reshuffle.main(args)`

`ecmwf_models.erainterim.reshuffle.parse_args(args)`

Parse command line parameters for conversion from image to time series.

### Parameters

**args** (*list*) – command line parameters as list of strings

### Returns

**args** – Parsed command line parameters

### Return type

`argparse.Namespace`

`ecmwf_models.erainterim.reshuffle.reshuffle(input_root, outputpath, startdate, enddate, variables, mask_seapoints=False, h_steps=(0, 6, 12, 18), imgbuffer=50)`

Reshuffle method applied to ERA images for conversion into netcdf time series format.

### Parameters

- **input\_root** (*str*) – Input path where ERA image data was downloaded to.
- **outputpath** (*str*) – Output path, where the reshuffled netcdf time series are stored.
- **startdate** (*datetime*) – Start date, from which images are read and time series are generated.
- **enddate** (*datetime*) – End date, from which images are read and time series are generated.
- **variables** (*list or str or tuple*) – Variables to read from the passed images and convert into time series format.
- **mask\_seapoints** (*bool, optional (default: False)*) – Mask points over sea, replace them with nan.
- **h\_steps** (*tuple, optional (default: (0, 6, 12, 18))*) – Full hours for which images are available.
- **imgbuffer** (*int, optional (default: 50)*) – How many images to read at once before writing time series. This number affects how many images are stored in memory and should be chosen according to the available amount of memory and the size of a single image.

`ecmwf_models.erainterim.reshuffle.run()`

## Module contents

### Submodules

#### ecmwf\_models.grid module

Common grid definitions for ECMWF model reanalysis products (regular gridded)

`ecmwf_models.grid.ERA5_RegularImgLandGrid(res_lat: float = 0.25, res_lon: float = 0.25, bbox: Optional[Tuple[float, float, float, float]] = None) → CellGrid`

Uses the 0.25 DEG ERA5 land mask to create a land grid of the same size, which also excluded Antarctica.

#### Parameters

- **res\_lat** (*float*, *optional* (default: 0.25)) – Grid resolution (in degrees) in latitude direction.
- **res\_lon** (*float*, *optional* (default: 0.25)) – Grid resolution (in degrees) in longitude direction.
- **bbox** (*tuple*, *optional* (default: None)) – (min\_lon, min\_lat, max\_lon, max\_lat) - wgs84 bbox to cut the global grid to.

`ecmwf_models.grid.ERA_IrregularImgGrid(lons: ndarray, lats: ndarray, bbox: Optional[Tuple[float, float, float, float]] = None) → CellGrid`

Create a irregular grid from the passed coordinates.

`ecmwf_models.grid.ERA_RegularImgGrid(res_lat: float = 0.25, res_lon: float = 0.25, bbox: Optional[Tuple[float, float, float, float]] = None) → CellGrid`

Create regular cell grid for bounding box with the selected resolution.

#### Parameters

- **res\_lat** (*float*, *optional* (default: 0.25)) – Grid resolution (in degrees) in latitude direction.
- **res\_lon** (*float*, *optional* (default: 0.25)) – Grid resolution (in degrees) in longitude direction.
- **bbox** (*tuple*, *optional* (default: None)) – (min\_lon, min\_lat, max\_lon, max\_lat) - wgs84 bbox to cut the global grid to.

#### Returns

**CellGrid** – Regular, CellGrid with 5DEG\*5DEG cells for the passed bounding box.

#### Return type

CellGrid

`ecmwf_models.grid.get_grid_resolution(lats: ~numpy.ndarray, lons: ~numpy.ndarray) -> (<class 'float'>, <class 'float'>)`

try to derive the grid resolution from given coords.

## ecmwf\_models.interface module

`class ecmwf_models.interface.ERAGrbDs(root_path, product, parameter=('swvl1', 'swvl2'), subgrid=None, mask_seapoints=False, h_steps=(0, 6, 12, 18), array_1D=True)`

Bases: MultiTemporalImageBase

Reader for a stack of ERA grib files.

#### Parameters

- **root\_path** (*string*) – Root path where the data is stored
- **product** (*str*) – ERA5 or ERAINT
- **parameter** (*list* or *str*, *optional* (default: ['swvl1', 'swvl2'])) – Parameter or list of parameters to read

- **expand\_grid** (*bool*, optional (default: *True*)) – If the reduced gaussian grid should be expanded to a full gaussian grid.

**tstamps\_for\_daterange**(*start\_date*, *end\_date*)

Get datetimes in the correct sub-daily resolution between 2 dates

**Parameters**

- **start\_date** (*datetime*) – Start datetime
- **end\_date** (*datetime*) – End datetime

**Returns**

**timesteps** – List of datetimes

**Return type**

list

```
class ecmwf_models.interface.ERAGrbImg(filename, product, parameter=('swvl1', 'swvl2'), subgrid=None,
                                     mask_seapoints=False, array_1D=True, mode='r')
```

Bases: ImageBase

Base class for reader for a single ERA Grib file.

**Parameters**

- **filename** (*str*) – Path to the image file to read.
- **product** (*str*) – ERA5 or ERAINT
- **parameter** (*list or str*, optional (default: `['swvl1', 'swvl2']`)) – Name of parameters to read from the image file.
- **subgrid** (*pygeogrids.CellGrid*, optional (default: *None*)) – Read only data for points of this grid and not global values.
- **mask\_seapoints** (*bool*, optional (default: *False*)) – Read the land-sea mask to mask points over water and set them to nan. This option needs the 'ism' parameter to be in the file!
- **array\_1D** (*bool*, optional (default: *False*)) – Read data as list, instead of 2D array, used for reshuffling.
- **mode** (*str*, optional (default: *'r'*)) – Mode in which to open the file, changing this can cause data loss. This argument should not be changed!

**close**()

Close file.

**flush**()

Flush data.

**read**(*timestamp=None*)

Read data from the loaded image file.

**Parameters**

**timestamp** (*datetime*, optional (default: *None*)) – Specific date (time) to read the data for.

**write**(*data*)

Write data to an image file.

**Parameters**

**image** (*object*) – pygeobase.object\_base.Image object

```
class ecmwf_models.interface.ERANcDs(root_path, product, parameter=('swvl1', 'swvl2'), subgrid=None,
                                     mask_seapoints=False, h_steps=(0, 6, 12, 18), array_1D=False)
```

Bases: MultiTemporalImageBase

Class for reading ERA 5 images in nc format.

#### Parameters

- **root\_path** (*str*) – Root path where image data is stored.
- **parameter** (*list or str, optional (default: ['swvl1', 'swvl2'])*) – Parameter or list of parameters to read from image files.
- **subgrid** (*pygeogrids.CellGrid, optional (default: None)*) – Read only data for points of this grid and not global values.
- **mask\_seapoints** (*bool, optional (default: False)*) – Use the land-sea-mask parameter in the file to mask points over water.
- **h\_steps** (*list, optional (default: [0, 6, 12, 18])*) – List of full hours for which images exist.
- **array\_1D** (*bool, optional (default: False)*) – Read data as list, instead of 2D array, used for reshuffling.

```
tstamps_for_daterange(start_date, end_date)
```

Get datetimes in the correct sub-daily resolution between 2 dates

#### Parameters

- **start\_date** (*datetime*) – Start datetime
- **end\_date** (*datetime*) – End datetime

#### Returns

**timestamps** – List of datetimes

#### Return type

*list*

```
class ecmwf_models.interface.ERANcImg(filename, product, parameter=['swvl1', 'swvl2'], subgrid=None,
                                       mask_seapoints=False, array_1D=False, mode='r')
```

Bases: ImageBase

Reader for a single ERA netcdf file.

#### Parameters

- **filename** (*str*) – Path to the image file to read.
- **product** (*str*) – ‘era5’ or ‘era5-land’ or ‘eraint’
- **parameter** (*list or str, optional (default: ['swvl1', 'swvl2'])*) – Name of parameters to read from the image file.
- **subgrid** (*pygeogrids.CellGrid, optional (default: None)*) – Read only data for points of this grid and not global values.
- **mask\_seapoints** (*bool, optional (default: False)*) – Read the land-sea mask to mask points over water and set them to nan. This option needs the ‘lsm’ parameter to be in the file!
- **array\_1D** (*bool, optional (default: False)*) – Read data as list, instead of 2D array, used for reshuffling.



- **mode** (*str*, optional (default: 'r')) – Mode in which to open the file, changing this can cause data loss. This argument should not be changed!

**close()**

Close file.

**flush()**

Flush data.

**read**(*timestamp=None*)

Read data from the loaded image file.

**Parameters**

**timestamp** (*datetime*, optional (default: None)) – Specific date (time) to read the data for.

**write**(*data*)

Write data to an image file.

**Parameters**

**image** (*object*) – pygeobase.object\_base.Image object

**class** ecmwf\_models.interface.ERATs(*ts\_path*, *grid\_path=None*, *\*\*kwargs*)

Bases: GriddedNcOrthoMultiTs

Time series reader for all reshuffled ERA reanalysis products in time series format. Use the read\_ts(lon, lat) resp. read\_ts(gpi) function of this class to read data for locations!

**Parameters**

- **ts\_path** (*str*) – Directory where the netcdf time series files are stored
- **grid\_path** (*str*, optional (default: None)) – Path to grid file, that is used to organize the location of time series to read. If None is passed, grid.nc is searched for in the ts\_path.
- **used** (Optional keyword arguments that are passed to the Gridded Base when) –

-----  
–

**parameters**

[list, optional (default: None)] Specific variable names to read, if None are selected, all are read.

**offsets**

[dict, optional (default: None)] Offsets (values) that are added to the parameters (keys)

**scale\_factors**

[dict, optional (default: None)] Offset (value) that the parameters (key) is multiplied with

ioclass\_kws: dict, (optional)

**read\_bulk**

[boolean, optional (default: False)] If set to True, the data of all locations is read into memory, and subsequent calls to read\_ts then read from cache and not from disk. This makes reading complete files faster.

**read\_dates**

[boolean, optional (default: False)] If false, dates will not be read automatically

but only on specific request useable for bulk reading because currently the netCDF num2date routine is very slow for big datasets.

## ecmwf\_models.utils module

Utility functions for all data products in this package.

**exception** `ecmwf_models.utils.CdoNotFoundError(msg=None)`

Bases: `ModuleNotFoundError`

`ecmwf_models.utils.get_default_params(name='era5')`

Read only lines that are marked as default variable in the csv file

### Parameters

**name** (*str*) – Name of the product to get the default parameters for

`ecmwf_models.utils.load_var_table(name='era5', lut=False)`

Load the variables table for supported variables to download.

### Parameters

**lut** (*bool, optional (default: False)*) – If set to true only names are loaded, so that they can be used for a LUT otherwise the full table is loaded

`ecmwf_models.utils.lookup(name, variables)`

Search the passed elements in the lookup table, if one does not exists, raise a Warning

`ecmwf_models.utils.make_era5_land_definition_file(data_file, out_file, data_file_y_res=0.25, ref_var='lsm', threshold=0.5, exclude_antarctica=True)`

Create a land grid definition file from a variable within a downloaded, regular (netcdf) era5 file.

### Parameters

- **data\_file** (*str*) – Path to the downloaded file that contains the image that is used as the reference for creating the land definition file.
- **out\_file** (*str*) – Full output path to the land definition file to create.
- **data\_file\_y\_res** (*float, optional (default: 0.25)*) – The resolution of the data file in latitude direction.
- **ref\_var** (*str, optional (default: 'lsm')*) – A variable in the data\_file that is the reference for the land definition. By default, we use the land-sea-mask variable.
- **threshold** (*float, optional (default: 0.5)*) – Threshold value below which a point is declared water, and above (or equal) which it is declared a land-point. If None is passed, then a point is declared a land point if it is not masked (numpy masked array) in the reference variable.
- **exclude\_antarctica** (*bool, optional (default: True)*) – Cut off the definition file at -60° Lat to exclude Land Points in Antarctica.

`ecmwf_models.utils.mkdate(datestring)`

Turns a datetime string into a datetime object

### Parameters

**datestring** (*str*) – Input datetime string

### Returns

**datetime** – Converted string

**Return type**

datetime

`ecmwf_models.utils.parse_filetype(inpath)`

Tries to find out the file type by searching for grib or nc files two subdirectories into the passed input path. If function fails, grib is assumed.

**Parameters**

**inpath** (*str*) – Input path where ERA data was downloaded to

**Returns**

**filetype** – File type string.

**Return type**

str

`ecmwf_models.utils.parse_product(inpath: str) → str`

Tries to find out what product is stored in the path. This is done based on the name of the first file in the path that is found.

**Parameters**

**inpath** (*str*) – Input path where ERA data was downloaded to

**Returns**

**product** – Product name

**Return type**

str

```
ecmwf_models.utils.save_gribs_from_grib(input_grib, output_path, product_name,
                                       filename_tmpl='{product}_AN_%Y%m%d_%H%M.grb',
                                       keep_original=True)
```

Split the downloaded grib file into daily files and add to folder structure necessary for reshuffling.

**Parameters**

- **input\_grib** (*str*) – Filepath of the downloaded .grib file
- **output\_path** (*str*) – Where to save the resulting grib files
- **product\_name** (*str*) – Name of the ECMWF model (only for filename generation)
- **filename\_tmpl** (*str*, *optional* (default: `product_OPER_0001_AN_date_time`)) – Template for naming each separated grib file

```
ecmwf_models.utils.save_ncs_from_nc(input_nc, output_path, product_name,
                                    filename_tmpl='{product}_AN_%Y%m%d_%H%M.nc', grid=None,
                                    keep_original=True, remap_method='bil')
```

Split the downloaded netcdf file into daily files and add to folder structure necessary for reshuffling.

**Parameters**

- **input\_nc** (*str*) – Filepath of the downloaded .nc file
- **output\_path** (*str*) – Where to save the resulting netcdf files
- **product\_name** (*str*) – Name of the ECMWF model (only for filename generation)
- **filename\_tmpl** (*str*, *optional* (default: `"{product}_AN_%Y%m%d_%H%M.nc"`)) – Template for naming each separated nc file
- **keep\_original** (*bool*) – keep the original downloaded data

`ecmwf_models.utils.str2bool(v)`

Parse a string to True/False

**Parameters**

**v** (*str*) – String to parse, must be part of the lists below.

**Returns**

**str2bool** – The parsed bool from the passed string

**Return type**

bool

**Module contents**

## INDICES AND TABLES

- `genindex`
- `modindex`
- `search`



## PYTHON MODULE INDEX

### e

- `ecmwf_models`, 40
- `ecmwf_models.era5`, 30
  - `ecmwf_models.era5.download`, 25
  - `ecmwf_models.era5.interface`, 28
  - `ecmwf_models.era5.reshuffle`, 29
- `ecmwf_models.erainterim`, 33
  - `ecmwf_models.erainterim.download`, 30
  - `ecmwf_models.erainterim.interface`, 32
  - `ecmwf_models.erainterim.reshuffle`, 33
- `ecmwf_models.grid`, 33
- `ecmwf_models.interface`, 34
- `ecmwf_models.utils`, 38





## C

CdoNotFoundError, 38  
 CDSStatusTracker (class in *ecmwf\_models.era5.download*), 25  
 close() (*ecmwf\_models.interface.ERAGrbImg* method), 35  
 close() (*ecmwf\_models.interface.ERANcImg* method), 37

## D

default\_variables() (in module *ecmwf\_models.era5.download*), 25  
 default\_variables() (in module *ecmwf\_models.erainterim.download*), 30  
 download\_and\_move() (in module *ecmwf\_models.era5.download*), 26  
 download\_and\_move() (in module *ecmwf\_models.erainterim.download*), 30  
 download\_era5() (in module *ecmwf\_models.era5.download*), 27  
 download\_eraint() (in module *ecmwf\_models.erainterim.download*), 31

## E

*ecmwf\_models*  
 module, 40  
*ecmwf\_models.era5*  
 module, 30  
*ecmwf\_models.era5.download*  
 module, 25  
*ecmwf\_models.era5.interface*  
 module, 28  
*ecmwf\_models.era5.reshuffle*  
 module, 29  
*ecmwf\_models.erainterim*  
 module, 33  
*ecmwf\_models.erainterim.download*  
 module, 30  
*ecmwf\_models.erainterim.interface*  
 module, 32  
*ecmwf\_models.erainterim.reshuffle*  
 module, 33

*ecmwf\_models.grid*  
 module, 33  
*ecmwf\_models.interface*  
 module, 34  
*ecmwf\_models.utils*  
 module, 38  
 ERA5\_RegularImgLandGrid() (in module *ecmwf\_models.grid*), 33  
 ERA5GrbDs (class in *ecmwf\_models.era5.interface*), 28  
 ERA5GrbImg (class in *ecmwf\_models.era5.interface*), 28  
 ERA5NcDs (class in *ecmwf\_models.era5.interface*), 28  
 ERA5NcImg (class in *ecmwf\_models.era5.interface*), 29  
 ERA\_IrregularImgGrid() (in module *ecmwf\_models.grid*), 34  
 ERA\_RegularImgGrid() (in module *ecmwf\_models.grid*), 34  
 ERAGrbDs (class in *ecmwf\_models.interface*), 34  
 ERAGrbImg (class in *ecmwf\_models.interface*), 35  
 ERAIntGrbDs (class in *ecmwf\_models.erainterim.interface*), 32  
 ERAIntGrbImg (class in *ecmwf\_models.erainterim.interface*), 32  
 ERAIntNcDs (class in *ecmwf\_models.erainterim.interface*), 32  
 ERAIntNcImg (class in *ecmwf\_models.erainterim.interface*), 32  
 ERANcDs (class in *ecmwf\_models.interface*), 36  
 ERANcImg (class in *ecmwf\_models.interface*), 36  
 ERATs (class in *ecmwf\_models.interface*), 37

## F

flush() (*ecmwf\_models.interface.ERAGrbImg* method), 35  
 flush() (*ecmwf\_models.interface.ERANcImg* method), 37

## G

get\_default\_params() (in module *ecmwf\_models.utils*), 38  
 get\_grid\_resolution() (in module *ecmwf\_models.grid*), 34

## H

handle\_error\_function() (in module *ecmwf\_models.era5.download.CDSStatusTracker* method), 25

## L

load\_var\_table() (in module *ecmwf\_models.utils*), 38  
lookup() (in module *ecmwf\_models.utils*), 38

## M

main() (in module *ecmwf\_models.era5.download*), 27  
main() (in module *ecmwf\_models.era5.reshuffle*), 29  
main() (in module *ecmwf\_models.erainterim.download*), 31  
main() (in module *ecmwf\_models.erainterim.reshuffle*), 33  
make\_era5\_land\_definition\_file() (in module *ecmwf\_models.utils*), 38  
mkdate() (in module *ecmwf\_models.utils*), 38  
module  
  *ecmwf\_models*, 40  
  *ecmwf\_models.era5*, 30  
  *ecmwf\_models.era5.download*, 25  
  *ecmwf\_models.era5.interface*, 28  
  *ecmwf\_models.era5.reshuffle*, 29  
  *ecmwf\_models.erainterim*, 33  
  *ecmwf\_models.erainterim.download*, 30  
  *ecmwf\_models.erainterim.interface*, 32  
  *ecmwf\_models.erainterim.reshuffle*, 33  
  *ecmwf\_models.grid*, 33  
  *ecmwf\_models.interface*, 34  
  *ecmwf\_models.utils*, 38

## P

parse\_args() (in module *ecmwf\_models.era5.download*), 27  
parse\_args() (in module *ecmwf\_models.era5.reshuffle*), 29  
parse\_args() (in module *ecmwf\_models.erainterim.download*), 31  
parse\_args() (in module *ecmwf\_models.erainterim.reshuffle*), 33  
parse\_filetype() (in module *ecmwf\_models.utils*), 39  
parse\_product() (in module *ecmwf\_models.utils*), 39

## R

read() (*ecmwf\_models.interface.ERAGrbImg* method), 35  
read() (*ecmwf\_models.interface.ERANcImg* method), 37  
reshuffle() (in module *ecmwf\_models.era5.reshuffle*), 29  
reshuffle() (in module *ecmwf\_models.erainterim.reshuffle*), 33

run() (in module *ecmwf\_models.era5.download*), 27  
run() (in module *ecmwf\_models.era5.reshuffle*), 30  
run() (in module *ecmwf\_models.erainterim.download*), 32  
run() (in module *ecmwf\_models.erainterim.reshuffle*), 33

## S

save\_gribs\_from\_grib() (in module *ecmwf\_models.utils*), 39  
save\_ncs\_from\_nc() (in module *ecmwf\_models.utils*), 39  
split\_array() (in module *ecmwf\_models.era5.download*), 27  
split\_chunk() (in module *ecmwf\_models.era5.download*), 28  
statusCode\_error(*ecmwf\_models.era5.download.CDSStatusTracker* attribute), 25  
statusCode\_ok(*ecmwf\_models.era5.download.CDSStatusTracker* attribute), 25  
statusCode\_unavailable (*ecmwf\_models.era5.download.CDSStatusTracker* attribute), 25  
str2bool() (in module *ecmwf\_models.utils*), 39

## T

tstamps\_for\_daterange() (*ecmwf\_models.interface.ERAGrbDs* method), 35  
tstamps\_for\_daterange() (*ecmwf\_models.interface.ERANcDs* method), 36

## W

write() (*ecmwf\_models.interface.ERAGrbImg* method), 35  
write() (*ecmwf\_models.interface.ERANcImg* method), 37