

Readme.md for Zenodo Dataset “Year-round North Atlantic-European Weather Regimes in ERA5 reanalysis”

Information

[doi:10.5281/zenodo.17080146](https://doi.org/10.5281/zenodo.17080146)

author contact: Christian Grams weatherregimes@gmail.com

Date: 12 September 2025

Version: 1.0

Description

This dataset contains the original files for a life-cycle definition of 7+1 year-round North Atlantic-European weather regimes following the definition of Grams et al. (2017) [doi:10.1038/nclimate3338](https://doi.org/10.1038/nclimate3338), updated to ECMWF ERA5 reanalysis as described in Hauser et al. (2024) [doi:10.5194/wcd-5-633-2024](https://doi.org/10.5194/wcd-5-633-2024). It is supplementary data to the publication Grams (2025, *in preparation for Weather Clim. Dynam.*) which will contain an in-depth technical explanation and analysis of key characteristics and trend in these regimes. A key novelty in the regime definition are the objectively identified regime life cycles which allow process-oriented predictability studies (e.g. Hauser et al. 2024).

The dataset has been originally computed for the data period 1979-2019. Using the backward extension of ERA5 (Soci et al., 2024, [doi:10.1002/qj.4803](https://doi.org/10.1002/qj.4803)), it has been extended to 1950. Using the near-realtime update it is extended forward until 20250726_21 in release V1.0. For the lifetime of ERA5 it is planned to continuously update the data irregularly at least once per year and upon request. Future updates will ensure full backward compatibility.

This dataset also contains a Jupyter notebook with examples of simple data analysis in python. The notebook aims to facilitate an easy start of the work with the data. Christian thanks Seraphine Hauser and Dominik Büeler for help with coding this ipynb and reformatting the data for easier accessibility. For more technical information, please refer to the remainder of this Readme.md. Users are advised to get familiar with this Readme.md prior to using this data.

References

Bell, B., Hersbach, H., Simmons, A., Berrisford, P., Dahlgren, P., Horányi, A., et al.: The ERA5 global reanalysis: Preliminary extension to 1950. Q J R Meteorol Soc, 147(741), 4186–4227, [doi:10.1002/qj.4174](https://doi.org/10.1002/qj.4174), 2021.

Grams, C. M.: A life-cycle definition of year-round weather regimes: characteristics and trends in the North-Atlantic European region, *in preparation for Weather Clim. Dynam.*, 2025.

Grams, C. M., Beerli, R., Pfenninger, S., Staffell, I., and Wernli, H.: Balancing Europe's wind power output through spatial deployment informed by weather regimes, Nat. Clim. Change, 7, 557–562, [doi:10.1038/nclimate3338](https://doi.org/10.1038/nclimate3338), 2017.

Hauser, S., Teubler, F., Riemer, M., Knippertz, P., and Grams, C. M.: Life cycle dynamics of Greenland blocking from a potential vorticity perspective, Weather Clim. Dynam., 5, 633–658, [doi:10.5194/wcd-5-633-2024](https://doi.org/10.5194/wcd-5-633-2024), 2024.

Hersbach H, Bell B, Berrisford P, et al.: The ERA5 global reanalysis. Q J R Meteorol Soc, 146: 1999–2049. [doi:10.1002/qj.3803](https://doi.org/10.1002/qj.3803), 2020.

Soci, C., Hersbach, H., Simmons, A., Poli, P., Bell, B., Berrisford, P., et al.: The ERA5 global reanalysis from 1940 to 2022. Q J R Meteorol Soc, 150(764), 4014–4048, [doi:10.1002/qj.4803](https://doi.org/10.1002/qj.4803), 2024.

USE OF THIS DATA:

A full explanation of the method and key characteristics of the regimes and trends of regime life cycles will be provided in Grams (2025, *in preparation for Weather Clim. Dynam.*). So far the methodology has been described and data already used in the #LSDPatKIT teams's various papers. The best available description so far can be found in the initial paper Grams et al. 2017 [doi:10.1038/nclimate3338](https://doi.org/10.1038/nclimate3338) (methodology section and supplement) and in Hauser et al. 2024 [doi:10.5194/wcd-5-633-2024](https://doi.org/10.5194/wcd-5-633-2024). The latter also explains the slight modifications with the update on ERA5. Please cite these paper along with the repository version of Grams (2025) in the meantime and until the latter is fully published.

Previous work uses a distinct colour scheme for each of the regimes, which reflect the relation between regimes. We kindly ask to use this colour scheme and the correct order of regimes as described below. For convenience the data records provided in this dataset are already pre-sorted in this order. We provide RGB triplets as well. This ensures, that readers recognize the link of your study to other studies using this regime definition. Order and colours are briefly explained below, and in more detail further down in this Readme.

original colour scheme and weather regime data index in correct order

abbreviation of regimes in preferred order and data files "AT ZO ScTr AR EuBL ScBL GL no"

corresponding regime indices in *text* files "1 2 3 4 5 6 7 0"

wr_metadata_rgb_new_colors.txt

Long name	Name	Color	RGB
Atlantic trough	AT	indigo	(75, 0, 130)
Zonal	ZO	red	(255, 0, 0)
Scandinavian trough	ScTr	darkorange	(255, 140, 0)
Atlantic ridge	AR	gold	(255, 215, 0)
European blocking	EuBL	yellowgreen	(154, 205, 50)
Scandinavian blocking	ScBL	darkgreen	(0, 100, 0)
Greenland blocking	GL	blue	(0, 0, 255)
No regime	no	grey	(128, 128, 128)

Versions

V1.0

Initial release containing the original regime definition and the expansion to the period 19500111_00 - 20250726_21 .

Content of dataset

detailed explanations follow below

./	
Readme.md	this readme file
./scripts_first_steps	
WR_read_example.ipynb	Jupyter notebook with five examples for data usage
WR_read_example.html	Jupyter notebook as HTML with output after running successfully
fct_wrera_db.py	function for reading regime attribution data, used in Jupyter notebook
fct_wrlcera_db.py	function for reading regime life cycle data, used in Jupyter notebook
./wr_data/	the full regime data
WR_LCtribution.txt	weather regime attribution time series
WR_lifecycle_information_*.txt	weather regime life cycles
WRI_projections.txt	weather regime index time series
WRI_std_params.txt	mean & stddev. of projection for weather regime index IWR calculation
normweights_Z0500.txt	calendar time dependent normalisation weights for normalising Z0500 anomalies
EOFs_WRs.nc	EOF patterns as 2D fields in netCDF format and normalization weights
Clusters_WRs.nc	non-normalised regime patterns in netCDF format on global domain
Normed_Z0500-patterns_EOFdomain.nc	normalised regime patterns in netCDF format in EOF domain needed for projection
./example_data/	additional data for .ipynb examples
CLIM_Z0500_year_1979-2019.nc	2D field of mean Z500 year-round 1979-2019
Z_N161_20250601_00.nc	Lanczos-filtered Z500 data for 20250601_00

Data description for ERA5 year-round 7 regime life cycle definition.

Configuration

- 1979-2019 reference climatology for Z500 anomaly computation and IWR scaling
- 10-day low-pass filter (Lanczos)
- period 19790111_00 – 20191231_21 used for EOF clustering and original life cycle definition
- 7 leading EOFs explaining 74,4% of variance
- regime projection (regime index IWR), and life cycles extended to cover the data period 19500111_00 - present with irregular updates at least once per year during ERA5 life time

Changes to earlier ERA-Interim definition (Grams et al. 2017)

- normalization weights using the 'lat-weighted spatial mean (in EOF domain) of the grid-point based 30d temporal stddev of Z0500'. This

improves summer regime identification.

- EOF-clustering performed for 6h data (insensitive to 6h, 12, 48h time interval)
- IWR (projection) computed for 3h data
- Z500 normweight for 3,9,15,21 UTC uses normweight of 0,6,12,18 UTC

Weather regime attribution and regime index time series

Regime Attribution time series

The following TEXT file contains the time series of the unambiguous attribution of each three-hourly time step to one of the seven regimes or no regime following either the original contribution to the EOF clusters (only 1979-2019), the maximum standardised regime projection (regime index IWR), or the full life cycle definition.

WR_LCattribution.txt

WR index after Michel and Riviere (2011) of filtered data N161 low pass > 10days, Z0@500, normed: intersection times, 0V

Cluster Class Index 0-7: no AT ZO ScTr AR EuBL ScBL GL

time in h since 19790101_00 | YYYYMMDD_HH | EOF attribution | max WR index | lifecycle WR index

-253968	19500111_00	0	6	0
-253965	19500111_03	0	6	0
-253962	19500111_06	0	6	0
.
.
.
240	19790111_00	3	3	3
243	19790111_03	0	3	3
246	19790111_06	3	3	3
249	19790111_09	0	3	3
.
.
.

The file contains 5 columns, the 5th is most relevant:

1. hour since 19790101_00
2. date in yyyyymmdd_hh
3. eof attribution -> timestep contributing to cluster XY based on EOFs (only EOF clustering period 11.1.1979-31.12.2019 only for 0 6 12 18 UTC times, other times are "0")
4. max WR index (Michel and Rivière, 2011). -> based on weather regime projection in physical space
5. LC attribution based on WR index (Michel and Rivière, 2011). -> THIS IS THE LIFECYCLE attribution INCLUDING NO REGIME. USE THIS FOR MOST PURPOSES.

Regime indices and regime order:

The data indices 0-7 in TEXT files refer to the following regimes: NAME (ABBREVIATION IN FILES)

0-7: no AT ZO ScTr AR EuBL ScBL GL

0=no regime [only 5th column] 1=Atlantic Trough (AT) 2=Zonal regime(ZO) 3=Scandinavian Trough(ScTr) 4=Atlantic Ridge (AR) 5=European Blocking (EuBL) 6=Scandinavian Blocking (ScBL) 7=Greenland Blocking (GL)

As in the Jupyter notebook, please write your code so that the regimes appear sorted in the following way (this will group related regimes). **Please also use this order when plotting panels / grouping regimes.** Reason: related regimes are grouped next to each other, and cyclonic / blocked are grouped together. In all previous works we use this order. More information on the reasoning in Grams (2025).

abbreviation of regimes in preferred order and data files "AT ZO ScTr AR EuBL ScBL GL no"

corresponding regime indices "1 2 3 4 5 6 7 0"

We also kindly ask you to use the following colour codes for colouring regimes in plots. Reason again is that the mixed colours reflect relations between regimes, and this is the colour scheme has already been widely used (better intercomparison of studies). The RGB Table and python matplotlib names used in Bueeler et al. 2021 [doi:10.1002/qj.4178](https://doi.org/10.1002/qj.4178) and Osman et al. 2023 [doi:10.1002/qj.4512](https://doi.org/10.1002/qj.4512) are listed below and an example

provided in the Jupyter notebook.

original colour scheme

wr_metadata_rgb_new_colors.txt

Long name	Name	Color	RGB
Atlantic trough	AT	indigo	(75, 0,130)
Zonal	Z0	red	(255, 0, 0)
Scandinavian trough	ScTr	darkorange	(255,140, 0)
Atlantic ridge	AR	gold	(255,215, 0)
European blocking	EuBL	yellowgreen	(154,205, 50)
Scandinavian blocking	ScBL	darkgreen	(0,100, 0)
Greenland blocking	GL	blue	(0, 0,255)
No regime	no	grey	(128,128,128)

colour blind friendly colour scheme close to original

We recommend this setting for better distinction of ScTr-EuBL and ZO-ScBL used in Gerighausen et al. (2025, in press) [doi:10.48550/arXiv.2408.04302](https://doi.org/10.48550/arXiv.2408.04302). Tested with <https://colororacle.org/>.

Long name	Name	Color	RGB
Atlantic trough	AT	indigo	(75, 0,130)
Zonal	Z0	rose	(204,102,119)
Scandinavian trough	ScTr	sand	(221,204,119)
Atlantic ridge	AR	gold	(255,215, 0)
European blocking	EuBL	teal	(68,170,153)
Scandinavian blocking	ScBL	green	(17,119, 51)
Greenland blocking	GL	blue	(0, 0,255)
No regime	no	grey	(128,128,128)

full colour blind friendly scheme

see <https://www.nki.nl/about-us/responsible-research/guidelines-color-blind-friendly-figures/>

Long name	Name	Color	RGB
Atlantic trough	AT	wine	(136, 34, 85)
Zonal	Z0	rose	(204,102,119)
Scandinavian trough	ScTr	sand	(221,204,119)
Atlantic ridge	AR	olive	(153,153, 51)
European blocking	EuBL	teal	(68,170,153)
Scandinavian blocking	ScBL	green	(17,119, 51)
Greenland blocking	GL	indigo	(51, 34,136)
No regime	no	grey	(128,128,128)

The following data are stored in sub-folder ./wr_data.

Regime projection time series

The following TEXT file contains the time series of the standardised regime projection, which is the regime index IWR, for each of the 7 regimes at each three-hourly time step.

WRI_projections.txt

WR index after Michel and Riviere (2011) of filtered data N161 low pass > 10days, Z0@500 normed: intersection times

time in h since 19790101_00, YYYYMMDD_HH, WR index for regimes: AT Z0 ScTr AR EuBL ScBL GL

-253968	19500111_00	-0.05433900	0.09829105	-1.38683009	-0.32397965	0.58707279	0.80997479	0.00612640
-253965	19500111_03	-0.08838309	0.09768835	-1.34896004	-0.29978162	0.60675180	0.79657406	-0.00862692
-253962	19500111_06	-0.12188707	0.09700651	-1.30812359	-0.27494192	0.62419289	0.78114247	-0.02308116
.
.
.
240	19790111_00	0.86587346	0.75074303	1.35784054	-0.45948577	-1.18579698	-1.07059395	-0.34373909
243	19790111_03	0.85441613	0.72267860	1.31580353	-0.44945070	-1.16703987	-1.03862977	-0.32952571
246	19790111_06	0.83927369	0.69352823	1.27102554	-0.43822104	-1.14390016	-1.00381625	-0.31613618
249	19790111_09	0.82096398	0.66378951	1.22433603	-0.42610571	-1.11709785	-0.96681082	-0.30378893

.
.
.

This file contains the standardised projection of instantaneous 10-day low-pass filtered normalized Z500 anomalies in each of the regime following Michel and Rivière (2011), see Grams (2025) for details. It describes a current flow situation in terms of resemblance to each of the 7 cluster mean EOF patterns in physical space. It is used to objectively identify the regime life cycles.

The files contains 9 columns

- 1. hour since 19790101_00
- 2. date in `yyyymmdd_hh`
- 3. 3 to 9 weather regime index (IWR) following Michel and Rivière, 2011 **correctly ordered**: AT ZO ScTr AR EuBL ScBL GL

Part 2 of the Jupyter notebook generates the example time series plot below showing each IWR time series and the identified active life cycles (in bold). On the bottom the “dominant active life cycle” is marked. The latter corresponds to the unambiguous “LC attribution” (5th column in `WR_LCtribution.txt`) and is the active LC with maximum projection at that time (if two or more LC coexist).

`era5_tseries_20241101_00_20250331_21.png`

Weather regime life cycle files

The TEXT files `WR_lifecycle_information*.txt` contain the objective life cycle definition.

The header for regimes 1-7 look as follows, for no regime (index 0) it is reduced and defined differently (see below):

LIFECYCLE INFORMATION														
AT														
clsfd EOF : 7958/59860 (13.2944%)														
total mxI : 31802/220728 (14.4078%)														
number	onset	sat start	mx	sat end	decay	dctr	dcto	dctoID	dctoDATE	onfr	onto	onfrID	onfr	
0	19500131_09	19500202_18	19500204_12	19500214_09	19500216_03	AT	none	-999	19500216_06	ScBL	ScBL	0	19500	
1	19500313_06	19500316_03	19500316_18	19500317_06	19500320_18	AT	none	-999	19500320_21	GL	GL	0	19500	
2	19500709_12	19500709_21	19500723_06	19500801_06	19500801_06	AT	ScBL	2	19500804_03	none	AT	-999	19500	
3	19500906_06	19500913_12	19500914_09	19500915_15	19500919_09	AT	none	-999	19500919_12	ZO	ZO	1	19500	
4	19510108_06	19510109_06	19510111_12	19510112_21	19510114_00	AT	none	-999	19510114_03	none	AT	-999	19510	
5	19510310_00	19510313_15	19510315_00	19510316_12	19510320_09	AT	ScTr	5	19510322_18	GL	GL	5	19510	
...														

key parameters defining the life cycle

- **number**: exclusive ID of the lifecycle
- **onset**: onset date
- **sat start**: begin of saturation stage date (not used)
- **mx**: maximum stage date
- **sat end**: end of saturation stage date (not used)
- **decay**: decay date

Transitions within a time window and based on life cycle (use these)

- **dctr**: regime type of active dominant life cycle at *decay* (not used)
- **dcto**: within 4 days after the decay (dc, dc+96h), type of first active dominant lifecycle or none
- **dctoID**: within 4 days after the decay (dc, dc+96h), ID of first active dominant lifecycle or -999 for none
- **dctoDATE**: within 4 days after the decay (dc, dc+96h), date when the other active dominant lifecycle is identified for the first time. For none this is *dt* after the *decay*
- **onfr**: up to 4 days prior to onset (on-96h,on), type of first (backward looking) active dominant lifecycle or none
- **onto**: regime type of active dominant lifecycle at *onset* (not used)
- **onfrID**: up to 4 days prior to onset (on-96h,on), ID of first active dominant lifecycle or -999 for none
- **onfrDATE**: up to 4 days prior to onset (on-96h,on), date when the other active dominant lifecycle is identified for the first time. For none this is *dt* before the *onset*

Immediate transitions based on dominant active life cycle (not used)

- **trfr**: regime type of active dominant life cycle when the current life cycle becomes dominant for the first time. It can be the life cycle itself (in the case that max projection is reached for the first time but another projection was larger without contributing to a LC (not persistent

enough)).

- *trfrID*: ID of the *trfr* LC
- *trfrDATE*: date when the *trfr* life cycle was dominant for the last time (this is *-dth* (one time step) before the considered LC becomes dominant for the first time)
- *trto*: regime type of active dominant life cycle when the current life cycle does no longer have the strongest projection for the first time. It can be the life cycle itself (in the case that another regime index is higher but this regime does not become an active life cycles (not persistent enough)).
- *trtoID*: ID of the *trto* LC
- *trtoDATE*: date when the *trto* life cycle is dominant for the first time

For the no regime a simplified file is contained:

TEXT file `WR_lifecycle_information_no.txt` contains a simplified life cycle indicating the begin and end of a no regime period.

Mean and standard deviation of regime projection for weather regime index calculation

The weather regime index IWR is defined as the standardised projection. This requires the static mean and standard deviation of the projection which are computed for the period 1979-2019. TEXT file `WRI_std_params.txt` contains this data, with each column containing the mean (3rd row) and standard deviation (4th row) for each of the regimes as labelled in 2nd row.

```
mean and std proj. (Michel and Riviere, 2011) 19790111_00 to 20191231_18
AT ZO ScTr AR EuBL ScBL GL
mean -0.0023841809 -0.0078344550 -0.0015250760 0.0021150061 0.0042730025 0.0089835953 0.0049633454
stdv 0.1711293906 0.2818541229 0.2004271448 0.2492269725 0.2292326987 0.2332117558 0.3415428102
```

Normalisation weights for 500hPa geopotential height

Text file `normweights_Z0500.txt` contains the calendar time dependent normalisation weights, see description of variable `normwgt(t time)` below.

EOF patterns as 2D field

NETCDF file `EOFs_WRs.nc` contains the spatial EOF patterns of the leading 20 EOFs and some important meta information:

- Variable `EOF(eof, lat, lon)`: spatial patterns of EOF1 to EOF20 for normalised 10-day low-pass filtered 500 hPa geopotential height anomalies six-hourly from 19790111_00-20191231_18.
- Variable `normwgt(t time)`: the calendar time dependent normalisation weight (in 6-hourly time steps, starting 00 UTC 1 January, ending 18 UTC 31 December for a leap (!) year), which is the spatial mean of the “running” standard deviation computed at a given calendar time for all 1979-2019 time steps in a +-15day window. The data array is valid for a leap year, meaning the array index 236 refers to 00 UTC 29 February, index 240 to 00 UTC 1 March. 29 February has to be overstepped for a non-leap year (see example for the if-statement in Part 4 of the .ipynb).
- Variable `time`: corresponding time for `normwgt` in hours since 180-01-01 00 UTC (1980 is an exemplary leap year)
- Variable `variance_exp1(eof)`: explained variance by each of the 20 EOFs

In addition normalisation weights are also provided in text file `normweights_Z0500.txt`.

```
>ncdump -h EOFs_WRs.nc
netcdf EOFs_WRs {
dimensions:
    eof = 20 ;
    latitude = 121 ;
    longitude = 241 ;
    time = 1464 ;
variables:
    int eof(eof) ;
    float latitude(latitude) ;
        latitude:standard_name = "latitude" ;
        latitude:units = "degrees_north" ;
    float longitude(longitude) ;
        longitude:standard_name = "longitude" ;
        longitude:units = "degrees_east" ;
    int64 time(time) ;
        time:standard_name = "time" ;
        time:units = "hours since 1980-01-01 00:00" ;
    float normwgt(time) ;
        normwgt:units = "geopotential metres" ;
    float EOF(eof, latitude, longitude) ;
        EOF:_FillValue = -999.f ;
        EOF:ntimes = 59860LL ;
```

```

        EOF:method = "no transpose" ;
        EOF:matrix = "covariance" ;
float variance_expl(eof) ;
        variance_expl:_FillValue = -999.f ;
}

```

Regime pattern as 2D field on global domain

NETCDF File `Clusters_WRs.nc` contains the non-normalised 10-day low-pass filtered 500hPa geopotential height anomaly (Z0500, 0 is an indicator for the Lanczos filter setup) for each of the regimes on the global domain at 0.5 degree horizontal grid spacing of. Units are geopotential metres (gpm). For an approximate calculation of the full field you can add the year-round climatology of 500 hPa geopotential height (e.g. for plotting, see example in Part 5 of the Jupyter notebook).

- Variable `Z0500_mean(wr, lat, lon)`: 10-day low-pass filtered 500hPa geopotential height anomaly mean of all 6-hourly time steps contributing to regime with data index `wr` according to EOF attribution
- Variable `Z0500_std(wr, lat, lon)`: 10-day low-pass filtered 500hPa geopotential height anomaly standard deviation of all 6-hourly time steps contributing to regime with data index `wr` according to EOF attribution
- global attribute `ClassNames`: String indicating the abbreviation of the regimes used for index 'wr' and thus the order of regime patterns in the file (here standard order "AT ZO ScTr AR EuBL ScBL GL")

```

> ncdump -h Clusters_WRs.nc
netcdf Clusters_WRs {
dimensions:
    wr = 7 ;
    latitude = 361 ;
    longitude = 720 ;
variables:
    int wr(wr) ;
    float latitude(latitude) ;
        latitude:standard_name = "latitude" ;
        latitude:units = "degrees_north" ;
    float longitude(longitude) ;
        longitude:standard_name = "longitude" ;
        longitude:units = "degrees_east" ;
    float Z0500_mean(wr, latitude, longitude) ;
        Z0500_mean:_FillValue = -999.f ;
        Z0500_mean:units = "geopotential metres" ;
        Z0500_mean:description = "mean of non-normalised 10-day low-pass filtered 500hPa geopotential height and
    float Z0500_std(wr, latitude, longitude) ;
        Z0500_std:_FillValue = -999.f ;
        Z0500_std:units = "geopotential metres" ;
        Z0500_std:description = "standard deviation of non-normalised 10-day low-pass filtered 500 hPa geopotent

// global attributes:
        :ClassNames = "AT ZO ScTr AR EuBL ScBL GL" ;
}

```

Regime pattern as normalised 2D field in EOF domain

NETCDF File `Normed_Z0500-patterns_EOFdomain.nc` contains the **normalised** 10-day low-pass filtered 500hPa geopotential height anomaly for each of the regimes in the **EOF domain**. This variable is needed and used for computing the projection and weather regime indices. Before computing the projection, you need the instantaneous time considered, by the normalisation weight, which is stored in the variable `normwgt(time)` in `EOFs_WRs.nc` and in the text file `normweights_Z0500.txt` (see before and example in Part 4 of the Jupyter notebook).

```

> ncdump -h Normed_Z0500-patterns_EOFdomain.nc
netcdf Normed_Z0500-patterns_EOFdomain {
dimensions:
    wr = 7 ;
    latitude = 121 ;
    longitude = 241 ;
variables:
    int wr(wr) ;
    float latitude(latitude) ;
        latitude:standard_name = "latitude" ;
        latitude:units = "degrees_north" ;
    float longitude(longitude) ;
        longitude:standard_name = "longitude" ;
        longitude:units = "degrees_east" ;
    float Z0500_mean(wr, latitude, longitude) ;

```

```

        Z0500_mean:_FillValue = -999.f ;
        Z0500_mean:units = "normalised with spatial average in EOF domain of 30-day running standard deviation a
        Z0500_mean:description = "WRs: AT Z0 ScTr AR EuBL ScBL GL" ;
float Z0500_stdd(wr, latitude, longitude) ;
        Z0500_stdd:_FillValue = -999.f ;
        Z0500_stdd:units = "normalised with spatial average in EOF domain of 30-day running standard deviation a
        Z0500_stdd:description = "WRs: AT Z0 ScTr AR EuBL ScBL GL" ;

// global attributes:
        :ClassNames = "AT Z0 ScTr AR EuBL ScBL GL" ;
}

```

Example data

This data is stored in sub-folder ./example_data.

Climatological mean of 500 hPa geopotential height

NETCDF File CLIM_Z@500_year_1979-2019.nc contains the year-round climatological mean (of all 3-hourly time steps 1979-2019) for 500 hPa geopotential height in variable Z@500(latitude, longitude). No filter or normalisation is applied.

```

> ncdump -h CLIM_Z@500_year_1979-2019.nc
netcdf CLIM_Z@500_year_1979-2019 {
dimensions:
    latitude = 361 ;
    longitude = 720 ;
variables:
    float latitude(latitude) ;
        latitude:standard_name = "latitude" ;
        latitude:units = "degrees_north" ;
    float longitude(longitude) ;
        longitude:standard_name = "longitude" ;
        longitude:units = "degrees_east" ;
    float Z@500(latitude, longitude) ;
        Z@500:long_name = "Z@500" ;
}

```

Example file for low-pass filtered geopotential height anomaly

NETCDF File Z0500_20250601_00.nc contains 10-day low-pass filtered 500 hPa geopotential height anomalies (abbreviated Z0) computed with respect to a 90-day running mean climatology. Different Lanczos filters are applied, but only filter 0 is relevant here. The time step is 00 UTC 1 June 2025. ERA5t is used.

- variable Z0: 10-day low-pass filtered data

```

> ncdump -h Z0500_20250601_00.nc
netcdf Z0500_20250601_00 {
dimensions:
    time = 1 ;
    lev = 1 ;
    lat = 361 ;
    lon = 720 ;
variables:
    float Z0(time, lev, lat, lon) ;
        Z0:_FillValue = -999.f ;
    double time(time) ;
        time:units = "hours since 1979-01-01 00:00" ;
        time:calendar = "standard" ;
    int lev(lev) ;

// global attributes:
        :domxmin = -180.f ;
        :domxmax = 179.5f ;
        :domymin = -90.f ;
        :domymax = 90.f ;
        :date = "20250601_00" ;
        :lev = 500 ;
        :lpass = 240, 120, 24 ;
        :history = "Lanczos filtered anomaly with filterwidth 161 timesteps (dt=3h); Z0 is low-pass filtered dat

```

Jupyter Notebook and python scripts with examples for working with the data

These files are stored in `./scripts_first_steps`.

`WR_read_examples.ipynb` contains a Jupyter Notebook with examples for facilitating an easy start to work with the data in Python. It shows how to open the regime data files and reading them into a dictionary. There are five example applications as described below and explained in more detail in the notebook annotations.

`WR_read_examples.html` is an HTML version, which shows how the output should look like, if the notebook runs successfully.

Christian thanks Seraphine Hauser and Dominik Büeler for help with coding this ipynb and providing auxiliary functions.

Part 1 uses `fct_wrera_db` (V Nov 2020) and `fct_wrlcera` (V Sep 2019) provided by Dominik Büeler to read the WR data. The former reads the file containing IWR time series and the categorical `maxIWR` and `LCattr` attributions. The latter generates LC objects.

Part 2 generates an time series plots of IWR, with the active life cycles marked, and a marker for the life cycle attribution.

Part 3 computes a frequency climatology in a given period and plots the frequency absolute as well as an anomaly.

Part 4 computes the projection in each of the regimes (regime indices), using the `netcdf` files with the normalised regime patterns and an instantaneous current Lanczos filtered geopotential height anomaly. It provides an example how to normalise the data prior to computing the regime indices.

Part 5 reads the regime patterns and the climatology for plotting the original regime patterns.

Christian M. Grams, 12 September 2025.