



## CDO and the CMIP standard

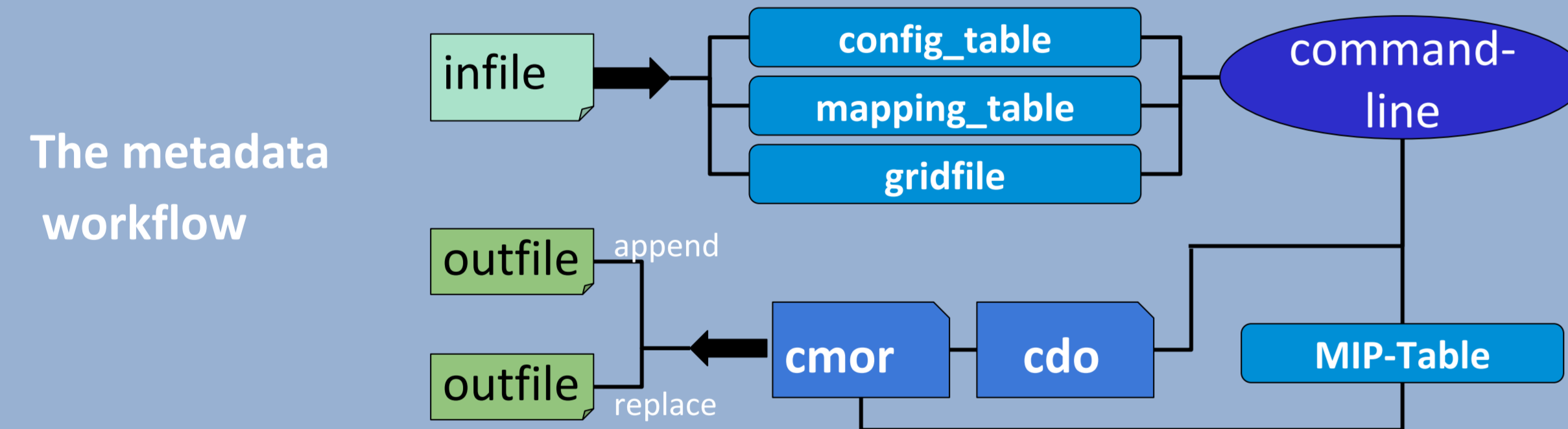
The Climate Data Operators tool (CDO) is a collection of command line operators to manipulate and analyze gridded model data. The data standard of the Coupled Model Intercomparison Project (CMIP) is widely used and simplifies the analysis and comparison of climate models.

The CDO's metadata model has been adapted to the CMIP standard such that CDO is able to process all CMIP related information.

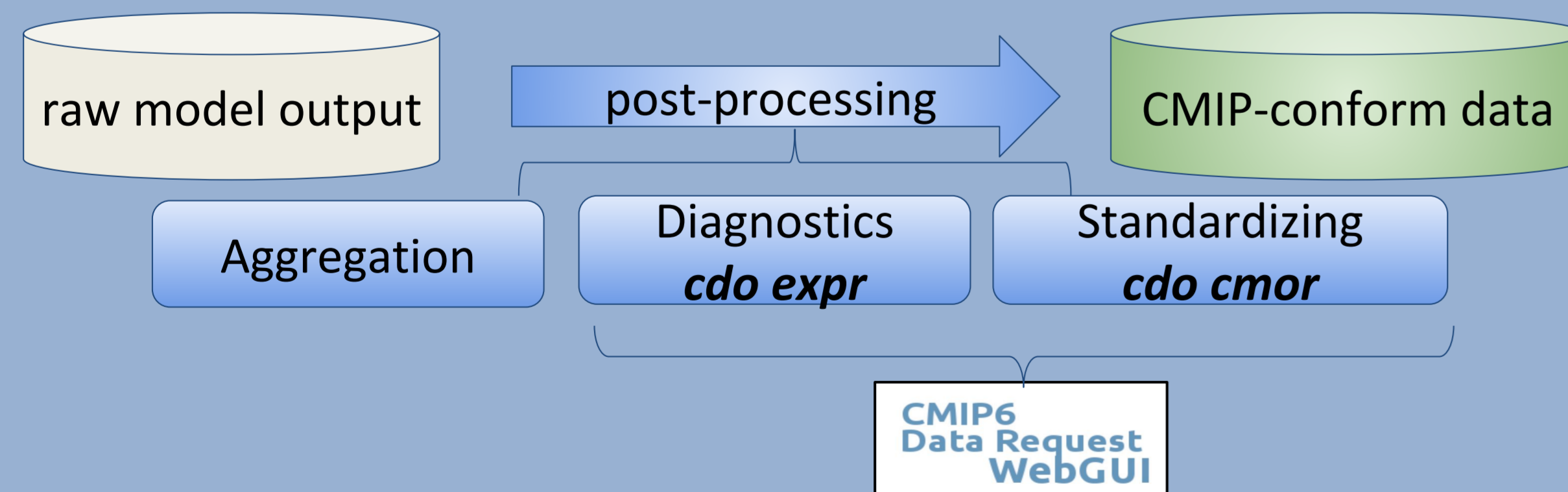
With the **cdo cmor operator**, CDO can now be used for any step of the project's data workflow consisting of *data production*, *data standardisation* as well as *data analysis*. The Climate Model Output Rewriter (CMOR) is a software package which rewrites climate data to be CMIP compliant.

```

The user interface
cdo cmor, MIP-table, \
    info= config_table \
    mapping_table= mapping_table \
    grid_info= gridfile
infile
"File stored in: $DRS/outfile"
    
```



## Integration into the CMIP6 post-processing workflow



A highly modular infrastructure allows the user to customize the post-processing. The GUI supplies the user with tools to easily configure and automate the generation of post-processing script fragments. Raw model output information is specified in a variable mapping database.

## CDO and Climate Extremes Indices

The Expert Team on Climate Change Detection and Indices (ETCCDI) defines 27 core Climate Extremes Indices (CEI) and recommends methods to calculate them. These methods have been implemented in the CDOs and are made available in a set of etccdi-operators. Output of these operators has been validated with the results of the reference tool RCLimindex. The CEIs can be subdivided into four categories for which changes were made in the CDO code. These categories are discussed in the following.

### Absolute values and threshold exceedances

- Operators allow to set the output frequency (yearly or monthly).

### Duration CEIs

- CDO will not stop counting consecutive days when years change.

Figure 1 shows temporal averaged values of the CWD index. CWD is the maximum of consecutive wet days per year for historical simulation r1i1p1f1 of the earth system model MPI-ESM1-2-HR. A wet day is a day with a precipitation sum over the day above 1mm.

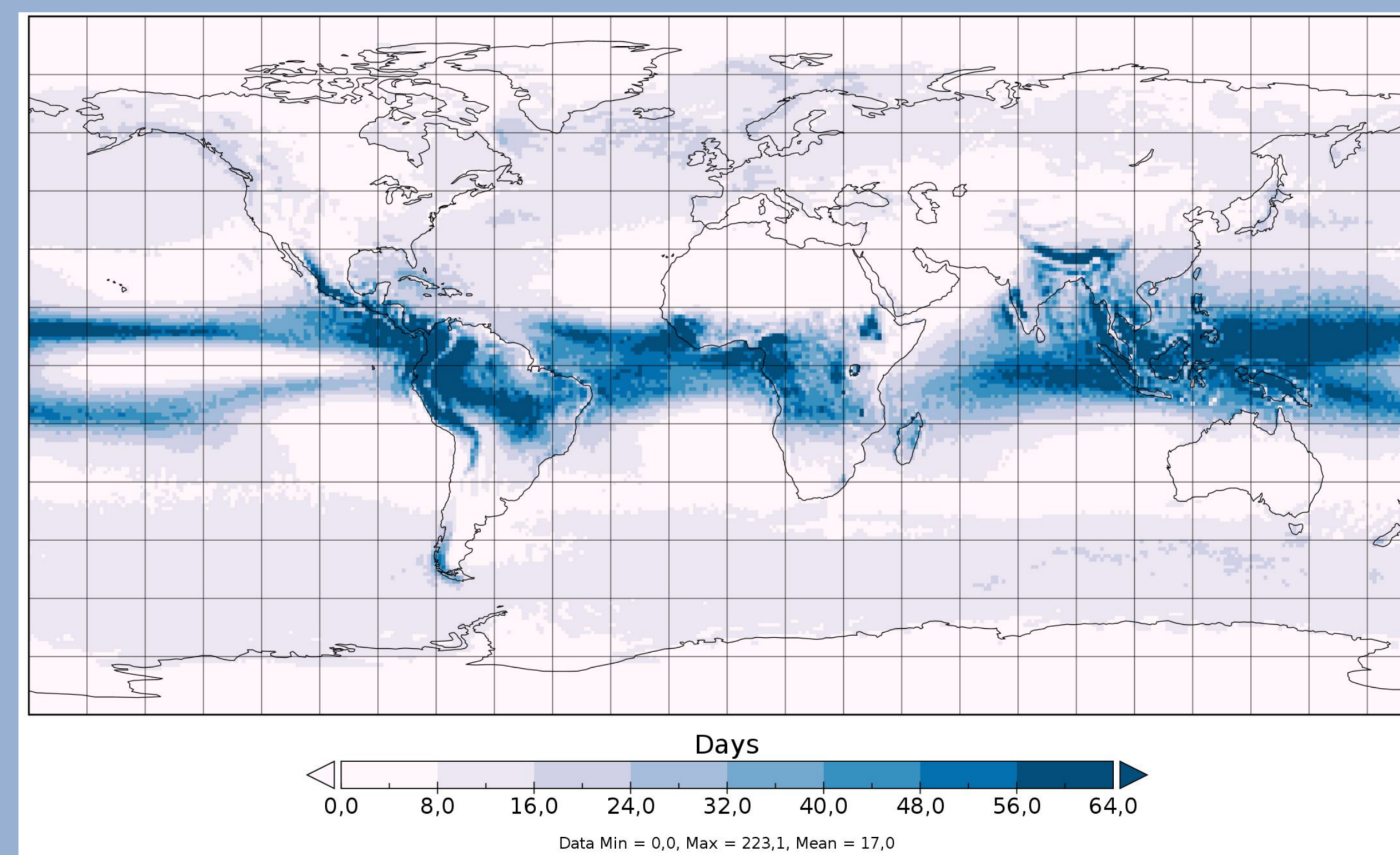


Fig.1: CWD index mean over 1961-1990 of one realization

### Percentile based CEIs

- Percentile calculations take the edges of the time series into account.
- The R percentile calculation method "type 8" is applied if sufficient memory is available.
- Bootstrapping is applied for the reference period.



The etccdi operators will be published in the next CDO release 1.9.9. A dev version can be installed via conda.

```
git clone https://gitlab.dkrz.de/k204210/cdo_cei.git
```



## CDO features: Application in python

A wrapper around the CDO binary enables the usage of CDO in python. This allows users to combine state-of-the-art python analysis tools with CDOs. On the rhs, a quick plot for the CWD index (see left column) is created in a jupyter notebook with pyplot.

```

cwo_values = cdo.etccdi_cwd(input=prHamburg,
output="cwoHamburg.nc",
returncdf=True).variables[
"consecutive_wet_days_index"
].per_time_period[0:]
cwo_values = cwo_values.flatten()

import matplotlib.pyplot as plt
plt.hist(cwo_values, bins=[5,5,6,5,7,5,8,5,
9,5,10,5,11,5,12,5,13,5])
plt.grid()
plt.xlabel("Largest number of consecutive
wet days per year")
plt.ylabel("Frequency")
plt.show()
    
```

### Brackets, pipes and expressions

The flexibility of CDOs arise from options to customize and combine their application. The following use case exemplifies how a suitable setting can lead to an optimized execution. Each color in the call on the rhs corresponds to one of the listed points.

```

"Merge a time series and multiply with 3!"

cdo mergetime \
infile1 infile2 tempfile
cdo mulc,3 tempfile outfile

cdo -P 8 cmor, Amon, info=info.txt \
-expr, "tas=3*invar" \
[mergetime infile1 infile2 ]
outfile
    
```

- A number of OpenMP threads can be set
- cdo cmor standardizes the input.
- cdo expr evaluates an expression which can substitute operator tasks
- Operators that process an arbitrary number of inputs (e.g. mergetime) can be piped by using brackets
- Pipes minimize I/O

## CMIP6 services at DKRZ

- Data Request post-processing infrastructure
- Data Quality Assurance
- Data Citation and PID support
- Hosting an ESGF-node and the CMIP Data Pool which provides fast and flexible access to about 2-5 PB of CMIP6 data
- HPC resources and software environments for interactive data analysis
- Hosting evaluation platforms like the freva evaluation system

