The MEDSCOPE Toolbox for Climate Forecasts postprocessing

Núria Pérez-Zanón (nuria.perez@bsc.es), Louis-Philippe Caron, Silvia Terzago, Bert Van Schaeybroeck, Lauriane Batté, M. Carmen Álvarez-Castro, Susana Corti, Marta Dominguez, Federico Fabiano, Silvio Gualdi, Jost von Hardenberg, Llorenç Lledó, Nicolau Manubens, Paola Marson, Stefano Materia, Eroteida Sánchez, Verónica Torralba, Deborah Verfaillie, and Danila Volpi

★ The relevance of CSTools is given by its state-of-the-art methods.

Documentation:
https://CRAN.R-project.org/package=CSTools

Video tutorials:
https://www.medscope-project.eu/products/tool-box/cstools-video-tutorials/

Gitlab project:
https://earth.bsc.es/gitlab/external/cstools

Pérez-Zanón et. al (under preparation)

Other key features are:

➢ the software structure in individual functions → allows to the users to design their own postprocessing chain of functions tailored to their needs
➢ each function is built in a nested design:
  ○ the external level → allows to retain data and metadata information during the analysis
  ○ the middle level → allows to parallel computing by setting one parameter ‘ncores’ thanks to multiApply package
  ○ the fundamental level → enables the compatibility with startR package
The MEDSCOPE Toolbox for Climate Forecasts postprocessing

Seasonal forecast ~10 MB
- 8 latitudes
- 8 longitudes
- 30 startdates
- 215 lead times
- 25 members

RainFARM

Downscaled seasonal forecast ~1 TB
- 800 latitudes
- 800 longitudes
- 30 startdates
- 215 lead times
- 250 members

Seasonal forecast Reference
- 256 latitudes
- 512 longitudes
- 30 startdates
- 1 lead times
- 51 members

Calibration ~30 MB

Test code: https://earth.bsc.es/gitlab/es/startR/-/blob/develop-RainFARMCase/inst/doc/usecase/ex2_10_rainFARM.R

Get results in the terminal.

Declare large data
```r
exp <- Start()
obs <- Start()
```

Define workflow
```r
wrap_cal <- function(obs, exp) {
calibrated <- CSTools::cal()
wf <- AddStep(Step())
}
```

Job execution
```r
res <- Compute(wf,
    chunks = list(lat = 2, lon = 2),
    cluster = list(host = "hord3",
        queue_type = "lsf",
        cores_per_job = 2))
```