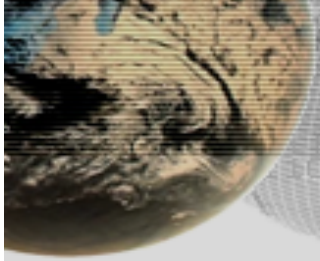


http://www.meteo.unican.es

Downscaling tools for adapting climate predictions to the user's needs: Challenges for their integration into climate services

A.S. Cofiño, J.M. Gutiérrez, J. Fernández, J. Bedia, M. Vega, S. Herrera, M.D. Frías, M. Iturbide, M.E. Magariño, and R. Manzanas

http://www.meteo.unican.es/udg-wiki

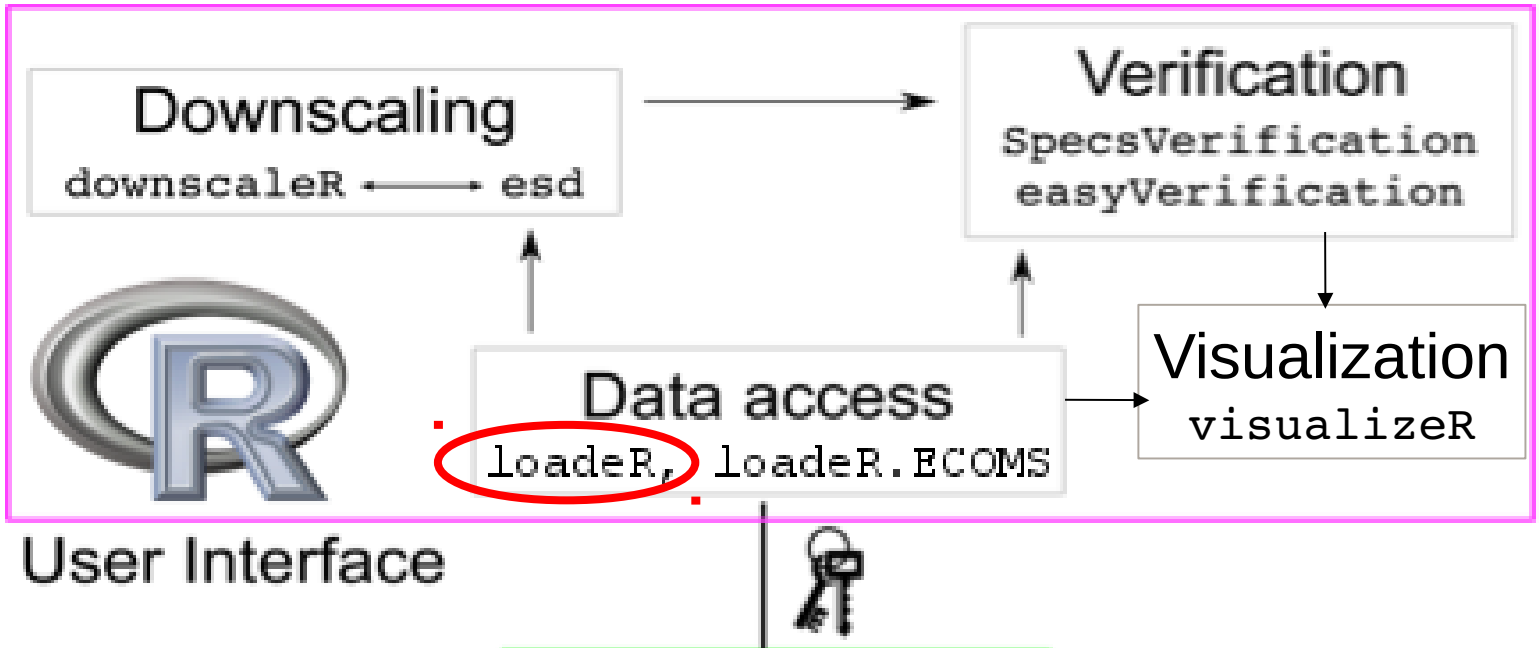
Joaquin Bedia
bediaj@unican.es
Santander Meteorology Group

Intl. Conference on Climate Science and Climate Serices
05-07 Oct 2016. Met Office, Exeter, UK

Contents

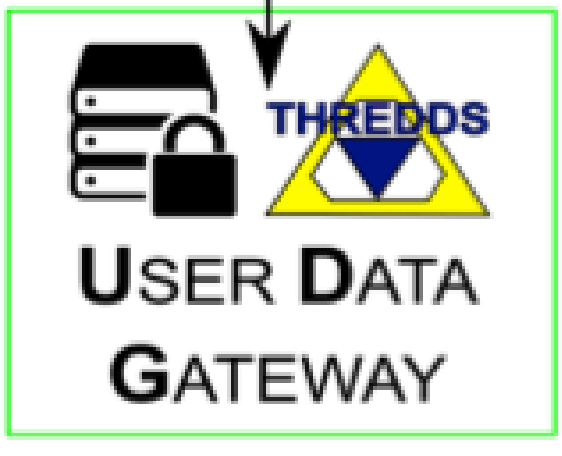
Contents

An R-based integrated framework for (remotely) accessing and processing climate data in the era of climate services



Existing **visualization**, **validation** and **downscaling** packages are transparently linked to UDG using **common data structures**.

Public and **restricted data** via virtual catalogs, allowing **harmonization** (a single vocabulary) and **data collocation**.



The User Data Gateway (UDG) is a **THREDDS** server with two in-house layers for:

1) Authentication

2) R-based data access.

Contents

ECOMS-UDG datasets

UDG provides **harmonized** access to locally stored daily data:

- **observations** (ECA, GSN, WFDEI),
- **reanalysis** (NCEP-R1, JRA55, ERA-Interim) and
- **CMIP5** data (several GCM projections).

and also to any other OPeNDAP **remote data**.

ECOMS-UDG is an extension for seasonal forecasting data, including hindcasts from state-of-the-art models:
ECMWF-System4, NCEP-CFSv2, UKMO-Glosea5.

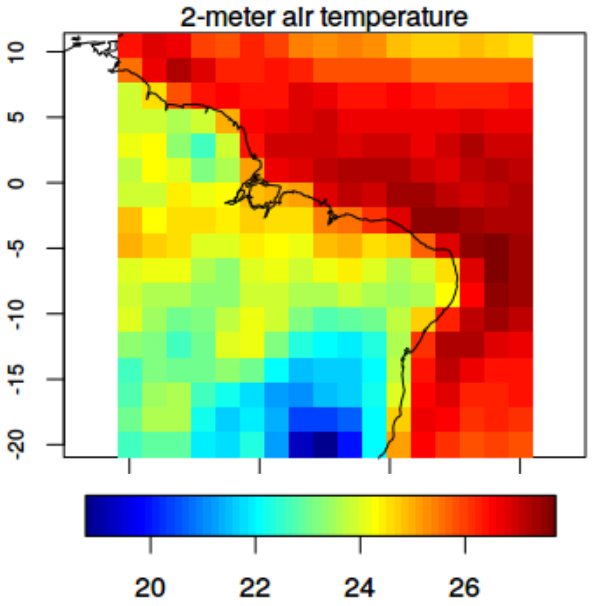
User-tailored design (SPECS and EUPORIAS) including variables needed for impact studies, mostly at surface level: **Precip, temp, wind speed, humidity, radiations, SLP**, but also **upper-air** information at 1000,850,700,500,300,200 mb (for statistical downscaling).

→ Link to available variables and datasets

Contents

Defining obs/reanalysis data chunk

```
library(loadeR.ECOMS)
loginUDG(username = 'jDoe', password = '******')
NCEPr1 <- loadECOMS(dataset = 'NCEP',
                    var = 'tas',
                    lonLim = c(-60,-30),
                    latLim = c(-20,10),
                    season = 3:5,
                    years = 1991:2010)
```



Define verification times
Season: MAM
Period: 1991-2000

Any other OpeNDAP server can be accessed with the same functions.

EXAMPLE: To access EOBS from KNMI's server:

```
dataset = 'http://opendap.knmi.nl/[...]/rr_0.25deg_reg_v12.0.nc'
```

Defining a prediction data chunk

Initialization times ("runtimes")

Verification times ("forecast times")

Seasonal Forecast (prediction)

seasons

2000 2001

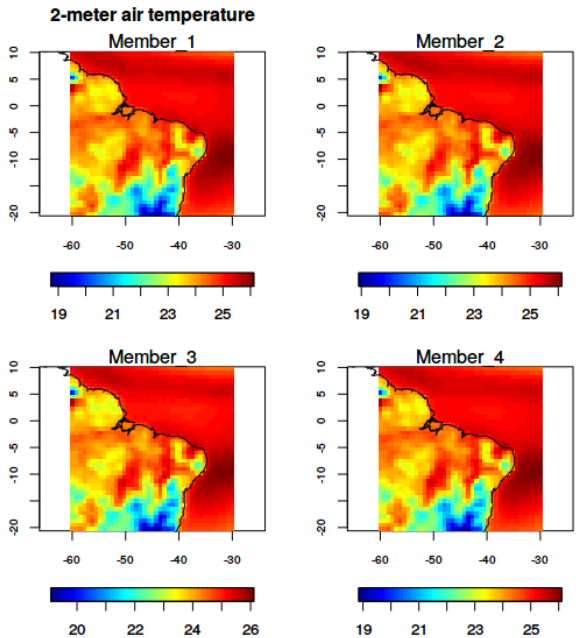
Target ➔ season

```
S4 <- loadECOMS(dataset = 'System4_seasonal_15',
                var = 'tas',
                lonLim = c(-60,-30),
                latLim = c(-20,10),
                members = 1:4,
                leadMonth = 2,
                season = 3:5,
                years = 1991:2010)
```

Define runtime
January
Initializations

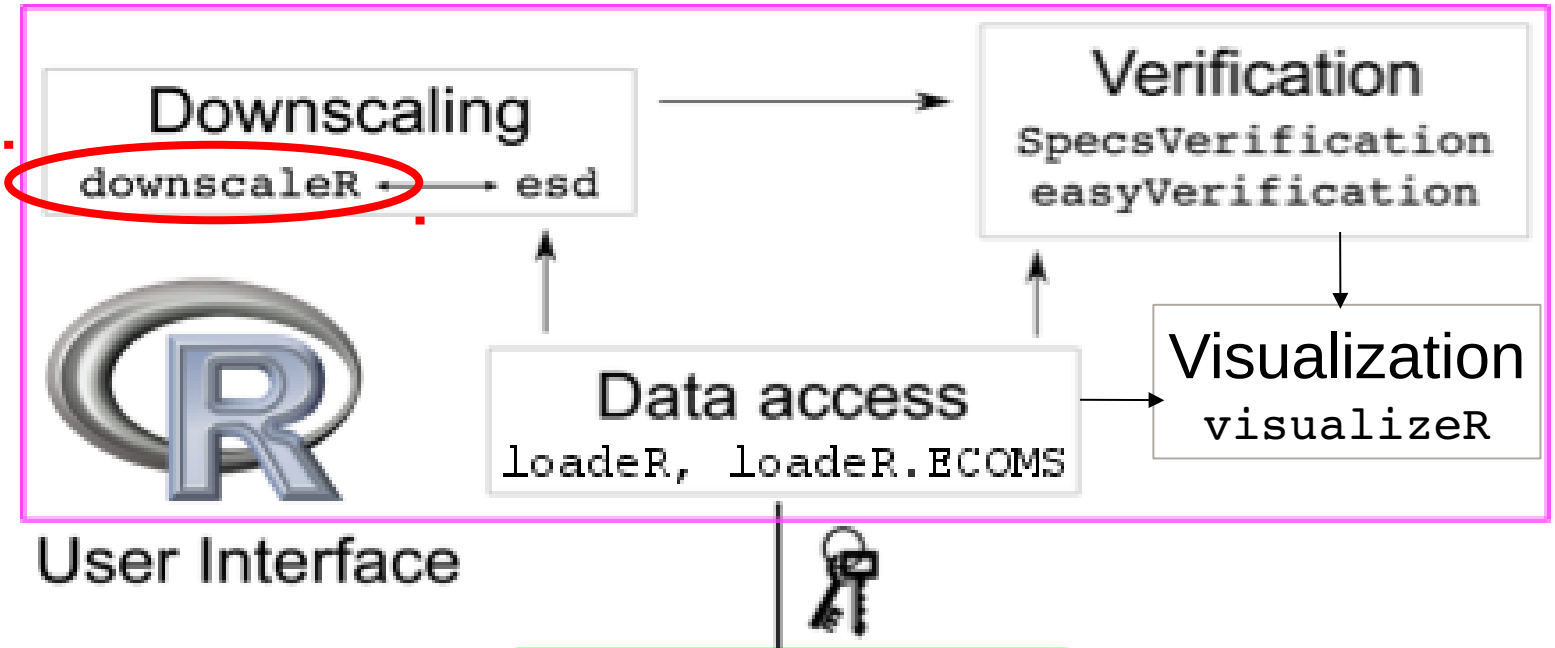
Define members
First 4

Define verification times
Season: MAM
Period: 1991-2000



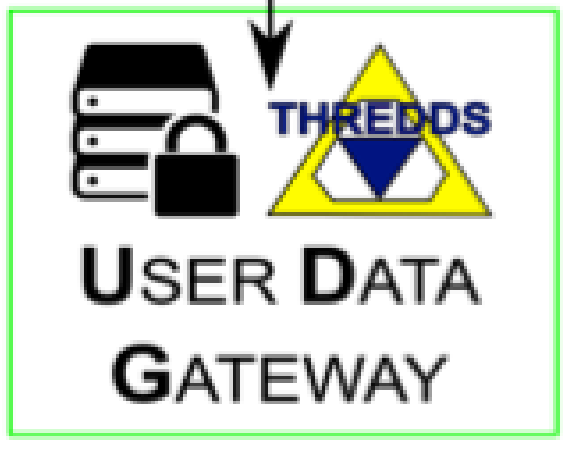
Contents

An R-based integrated framework for (remotely) accessing and processing climate data in the era of climate services



Existing **visualization**, **validation** and **downscaling** packages are transparently linked to UDG using **common data structures**.

Public and **restricted data** via virtual catalogs, allowing **homogenization** (a single vocabulary) and **data collocation**.



downscaleR is fully integrated with the loadeR data structures to perform downscaling and bias correction

- **`downscaleR`** has been designed to work with daily data (seasonal predictions, multidecadal projections). Extends the **`loadeR`** capabilities for **data manipulation**
  - Regridding/interpolation, subsetting and aggregation
  - PCA/EOF analysis
- **Bias correction**/calibration (including cross-validation)
  - Scaling, ISI-MIP, qq-mapping (various forms), parametric.
- **Perfect-prog downscaling** (including cross-validation)
  - Analogs, regression (linear and generalized linear), weather typing...
- **Parallel execution** options

→ *https://github.com/SantanderMetGroup/downscaleR*

Contents

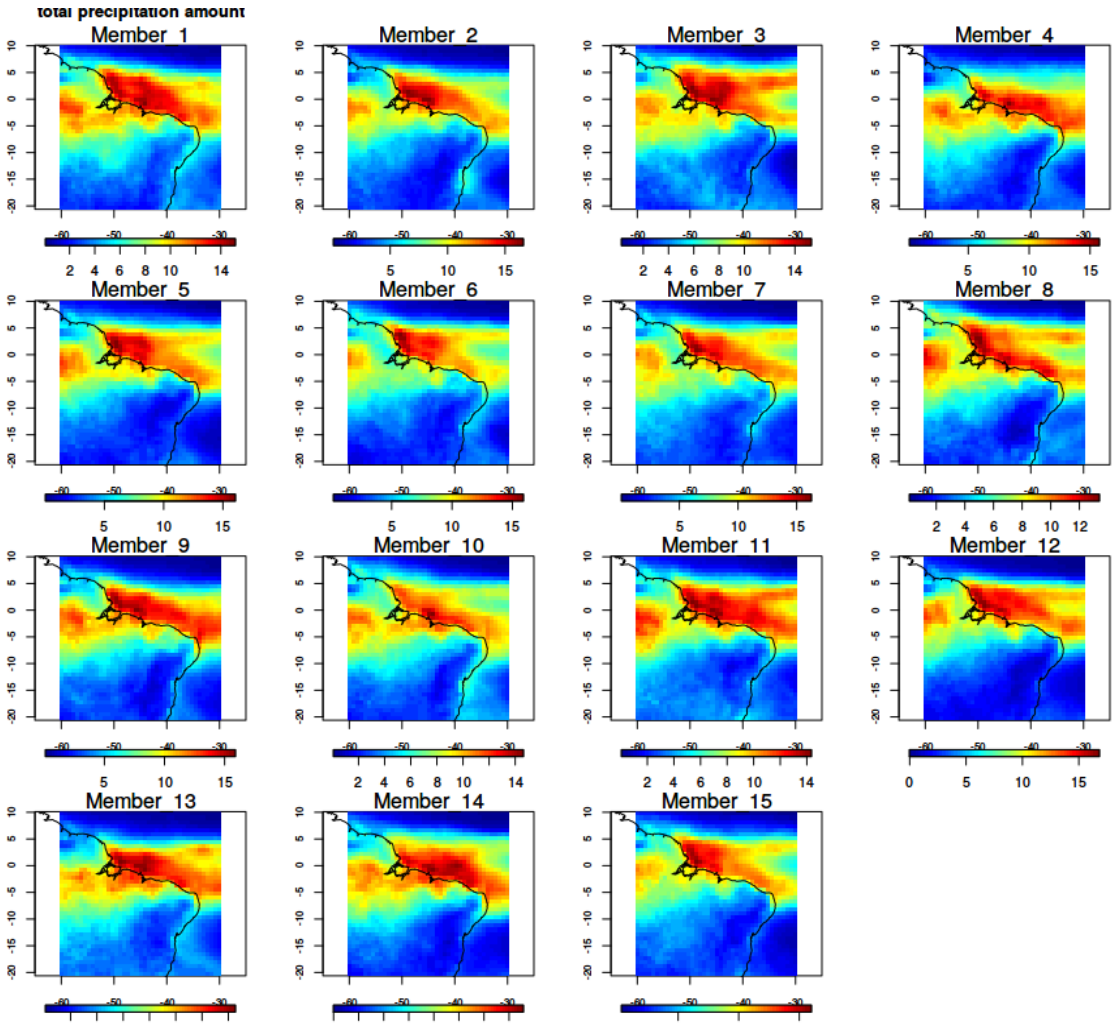
Downscaling of System4 MAM precipitation forecast over NE Brazil (January initialization, 15 members)

SPECS D52.2 http://www.specs-fp7.eu/wiki/images/d/d0/SPECS_D52.2.pdf

A relatively complex task usually involving many intermediate steps from data loading to analysis of the results...

… made EASY (10 commands)

… and **fully** REPRODUCIBLE



Data loading...

```
predictor <- c("psl", "ta@850", "hus@850")
# Loading NCEP (Predictors)
NCEP.psl <- loadECOMS(dataset = "NCEP" , var = predictor[1],
                      lonLim = c(-60,-30), latLim = c(-20,10),
                      season = 3:5, years = 1981:2010)
# Same for ta@850 and hus@850
# Loading System4 predictions (Predictand, precip)
S4 <- loadECOMS(dataset = "System4_seasonal_15" , var = 'tp',
                lonLim = c(-60,-30), latLim = c(-20,10),
                season = 3:5, years = 1981:2010,
                leadMonth = 2)
```

… data preprocessing ....

```
predictor <- c("psl", "ta@850", "hus@850")
# Loading NCEP (Predictors)
NCEP.psl <- loadECOMS(dataset = "NCEP" , var = predictor[1],
                      lonLim = c(-60,-30), latLim = c(-20,10),
                      season = 3:5, years = 1981:2010)
# Same for ta@850 and hus@850
# Loading System4 predictions (Predictand, precip)
S4 <- loadECOMS(dataset = "System4_seasonal_15" , var = 'tp',
                lonLim = c(-60,-30), latLim = c(-20,10),
                season = 3:5, years = 1981:2010,
                leadMonth = 2)

# Predictor dataset
NCEP <- makeMultiGrid(NCEP.psl, NCEP.ta850, NCEP.hus850)
# Computing EOFs and PCs
ncep.eof <- computeEOF(NCEP, n.eofs = 15)
# Interpolating S4 to the NCEP grid, and rescaling
S4 <- interpGrid(S4, getGrid(NCEP), method = 'nearest')
S4.sc <- rescaleMonthlyMeans(pred = NCEP, sim = S4)
```

… downscaling and cross-validation ...

```
predictor <- c("psl", "ta@850", "hus@850")
# Loading NCEP (Predictors)
NCEP.psl <- loadECOMS(dataset = "NCEP" , var = predictor[1],
                      lonLim = c(-60,-30), latLim = c(-20,10),
                      season = 3:5, years = 1981:2010)
# Same for ta@850 and hus@850
# Loading System4 predictions (Predictand, precip)
S4 <- loadECOMS(dataset = "System4_seasonal_15" , var = 'tp',
                lonLim = c(-60,-30), latLim = c(-20,10),
                season = 3:5, years = 1981:2010,
                LeadMonth = 2)

# Predictor dataset
NCEP <- makeMultiGrid(NCEP.psl, NCEP.ta850, NCEP.hus850)
# Computing EOFs and PCs
ncep.eof <- computeEOF(NCEP, n.eofs = 15)
# Interpolating S4 to the NCEP grid, and rescaling
S4 <- interpGrid(S4, getGrid(NCEP), method = 'nearest')
S4.sc <- rescaleMonthlyMeans(pred = NCEP, sim = S4)

# Downscaling with Generalized Linear Models
down <- downscale(obs = tp.wfdei, pred = ncep.eof,
                  sim = s4.sc, cross.val = "loocv",
                  method = "glm", n.pcs = 15,
                  parallel = TRUE)
```

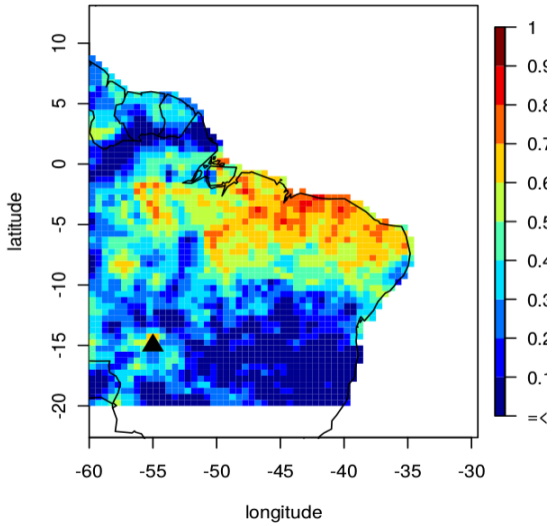
… and verification and visualization.

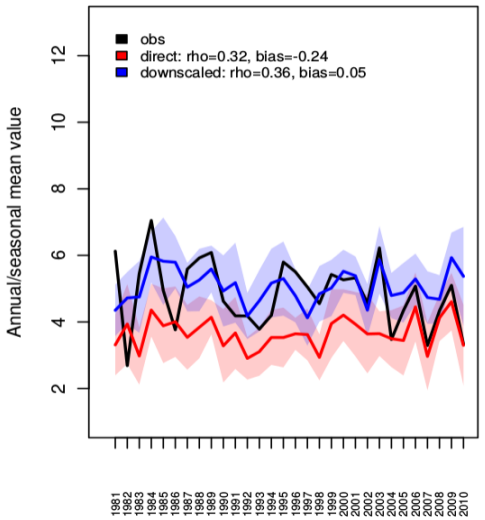
```
predictor <- c("psl", "ta@850", "hus@850")
# Loading NCEP (Predictors)
NCEP.psl <- loadECOMS(dataset = "NCEP" , var = predictor[1],
                      lonLim = c(-60,-30), latLim = c(-20,10),
                      season = 3:5, years = 1981:2010)
# Same for ta@850 and hus@850
# Loading System4 predictions (Predictand, precip)
S4 <- loadECOMS(dataset = "System4_seasonal_15" , var = 'tp',
                lonLim = c(-60,-30), latLim = c(-20,10),
                season = 3:5, years = 1981:2010,
                LeadMonth = 2)

# Predictor dataset
NCEP <- makeMultiGrid(NCEP.psl, NCEP.ta850, NCEP.hus850)
# Computing EOFs and PCs
ncep.eof <- computeEOF(NCEP, n.eofs = 15)
# Interpolating S4 to the NCEP grid, and rescaling
S4 <- interpGrid(S4, getGrid(NCEP), method = 'nearest')
S4.sc <- rescaleMonthlyMeans(pred = NCEP, sim = S4)

# Downscaling with Generalized Linear Models
down <- downscale(obs = tp.wfdei, pred = ncep.eof,
                  sim = s4.sc, cross.val = "loocv",
                  method = "glm", n.pcs = 15,
                  parallel = TRUE)
```

```
# Analysis of results
quickDiagnostics(down)
```

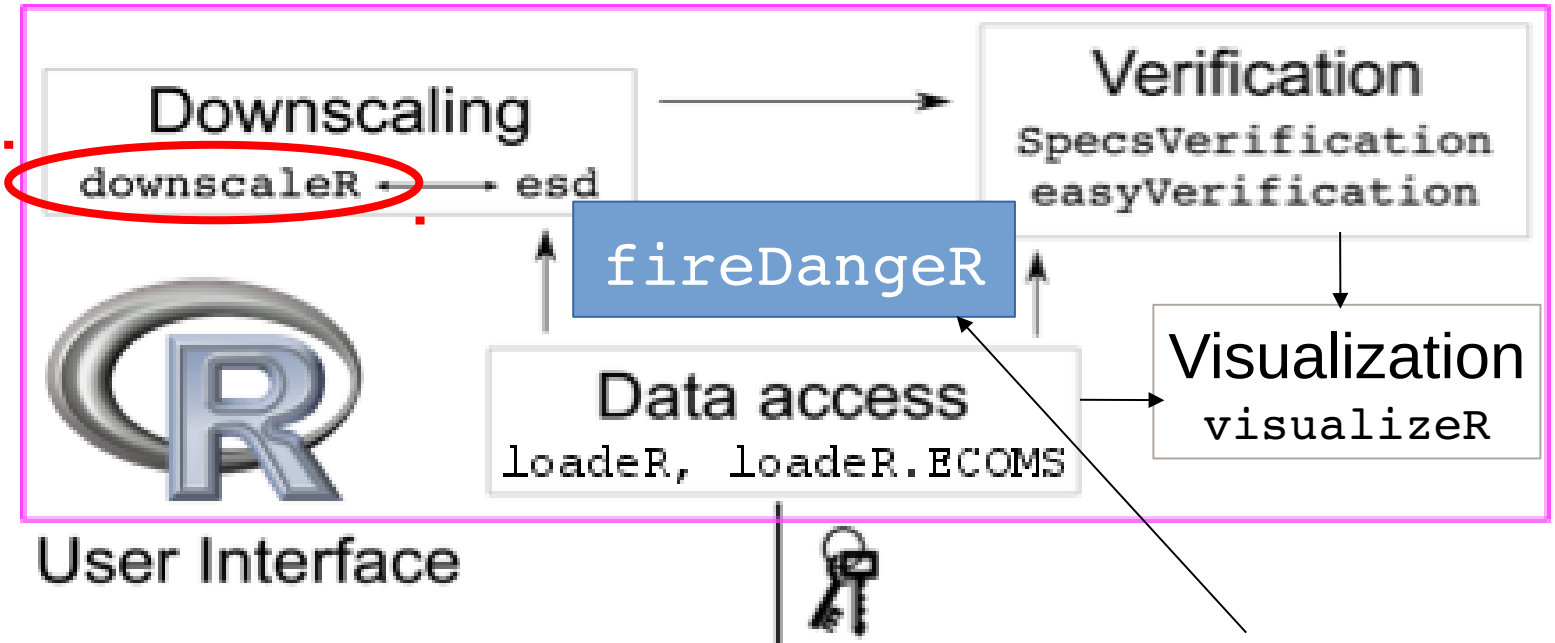




down
obs
S4

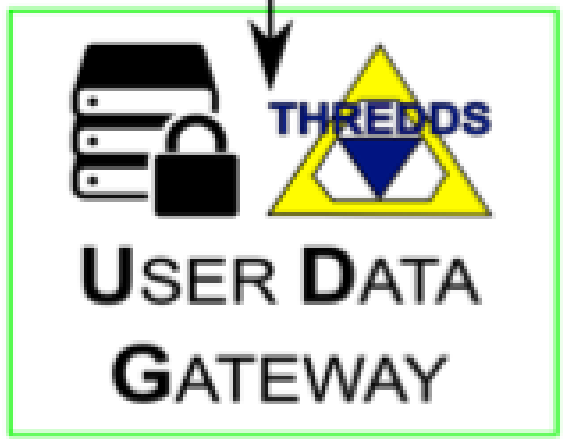
Contents

An R-based integrated framework for (remotely) accessing and processing climate data in the era of climate services



Existing **visualization**, **validation** and **downscaling** packages are transparently linked to UDG using **common data structures**.

Public and **restricted data** via virtual catalogs, allowing **homogenization** (a single vocabulary) and **data collocation**.



`fireDangeR` is a package to compute **specific CIIs for the forestry and civil protection sectors** (EUPORIAS WP22)

EUPORIAS

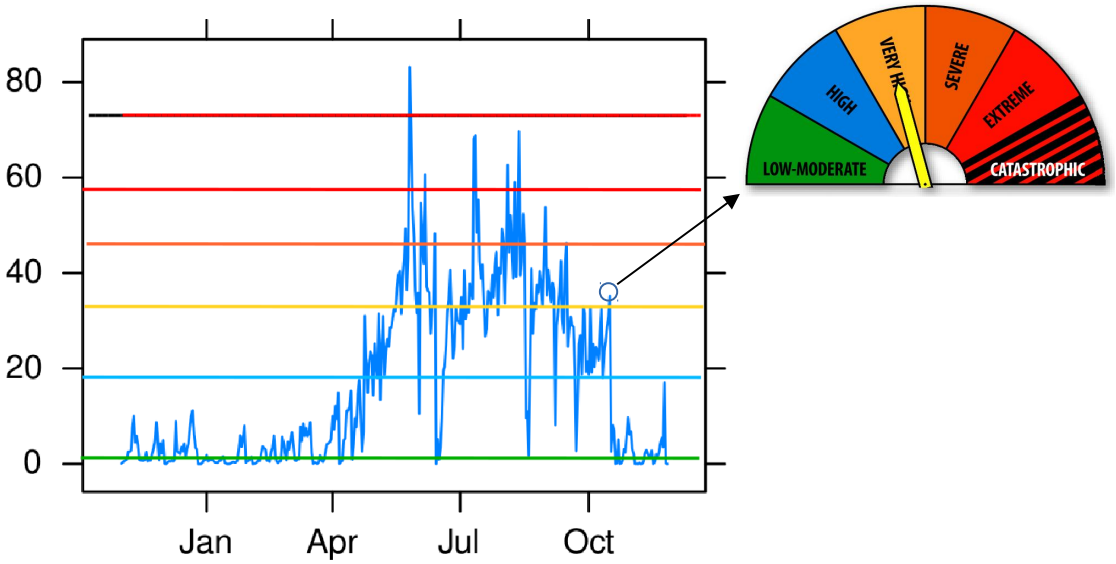
Santander Meteorology Group
A multidisciplinary approach for weather & climate

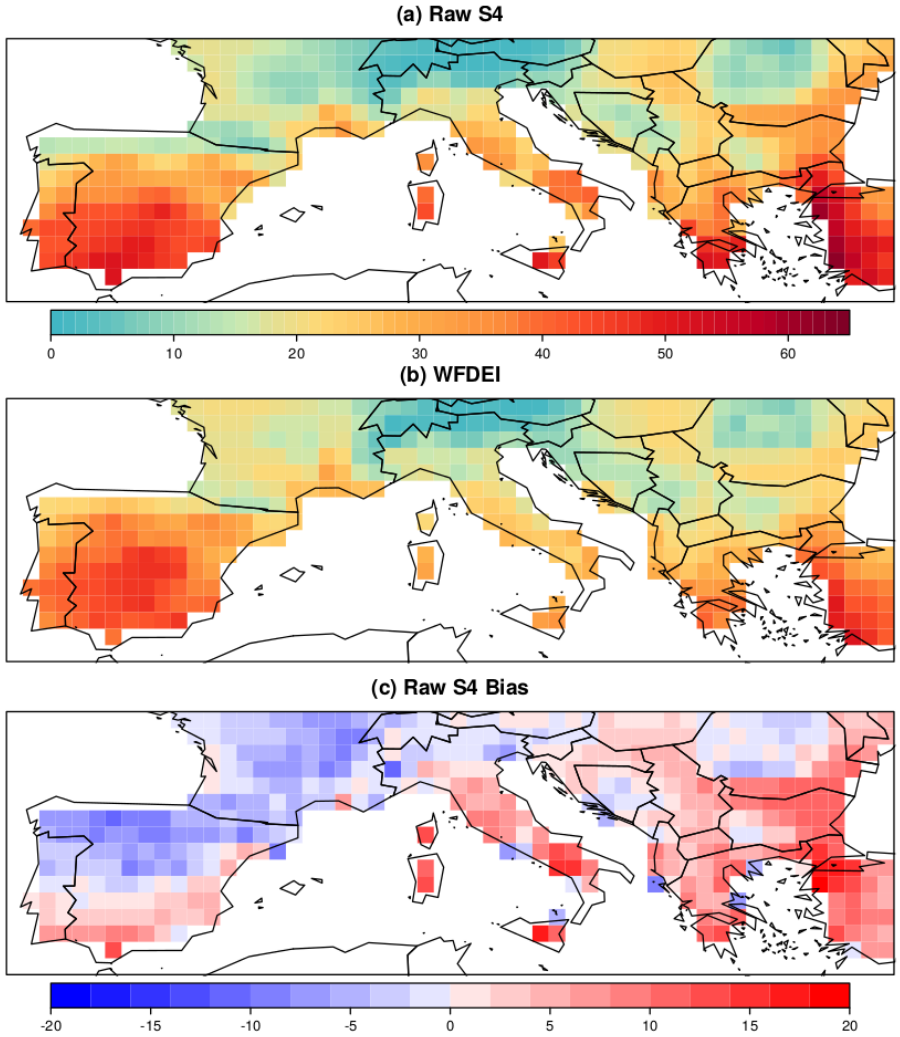
Worked example FWI Bias Correction

- FWI is a **daily**-based, **multivariable** CII rating the potential for fire ignition and spread given the atmospheric conditions

FWI: f(hurs, tas, wss, precip)

- Fire managers require **bias correction** for operational applicability
- Standard protocols and **user-friendly tools** are still lacking in a seasonal forecasting context





Bias correction (Empirical Quantile Mapping) of System4 JJAS FWI forecast over EU-MED region (May init., 15 members)

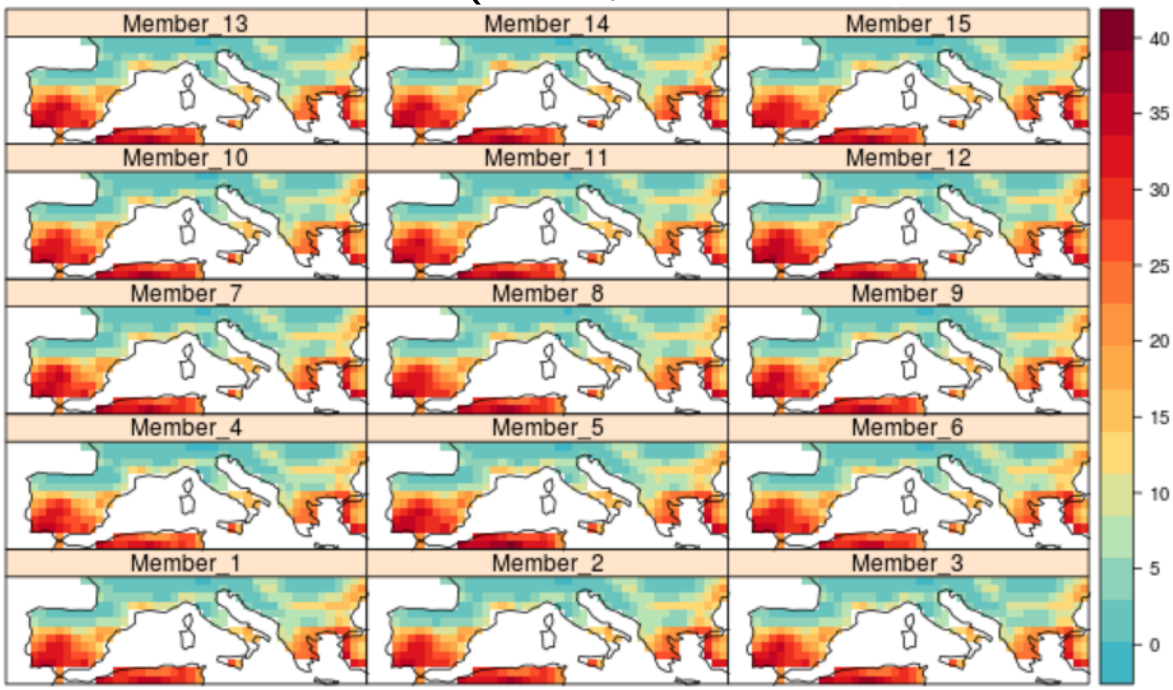
Worked example: https://github.com/SantanderMetGroup/fireDanger/wiki/CS_Bedia-et-al-2016-(submitted)

A relatively complex task usually involving many intermediate steps from data loading to FWI calculation and analysis of the results...

… made EASY (a few commands)

… and **fully** REPRODUCIBLE

Mean FWI (JJAS, 1981-2010



Data loading...

```
inputvars <- c("hurs", "wss", "tas", "tp")
# Loading WFDEI (Observations)
obs.hurs <- loadECOMS(dataset = "WFDEI" , var = inputvars[1],
                      lonLim = c(-10,35), latLim = c(35,47),
                      season = 6:9, years = 1981:2010)
# Same for wss, tas and tp
# Loading System4 variables
s4.hurs <- loadECOMS(dataset = "System4_seasonal_15" , var = inputvars[1],
                     lonLim = c(-10,35), latLim = c(35,47),
                     season = 6:9, years = 1981:2010,
                     leadMonth = 1)
```

… FWI calculation ...

```
inputvars <- c("hurs", "wss", "tas", "tp")
# Loading WFDEI (Observations)
obs.hurs <- loadECOMS(dataset = "WFDEI" , var = inputvars[1],
                      lonLim = c(-10,35), latLim = c(35,47),
                      season = 6:9, years = 1981:2010)
# Same for wss, tas and tp
# Loading System4 variables
s4.hurs <- loadECOMS(dataset = "System4_seasonal_15" , var = inputvars[1],
                     lonLim = c(-10,35), latLim = c(35,47),
                     season = 6:9, years = 1981:2010,
                     leadMonth = 1)

obs.multigrid <- makeMultiGrid(obs.hurs, obs.tas, obs.tp, obs.wss)
s4.multigrid <- makeMultiGrid(s4.hurs, s4.tas, s4.tp, s4.wss)
obs.fwi <- fwiGrid(multigrid = obs.multigrid, landmask = NULL)
s4.fwi <- fwiGrid(multigrid = s4.multigrid, landmask = s4.landmask, parallel = TRUE)
```

… Empirical Quantile Mapping...

```
inputvars <- c("hurs", "wss", "tas", "tp")
# Loading WFDEI (Observations)
obs.hurs <- loadECOMS(dataset = "WFDEI" , var = inputvars[1],
                      lonLim = c(-10,35), latLim = c(35,47),
                      season = 6:9, years = 1981:2010)
# Same for wss, tas and tp
# Loading System4 variables
s4.hurs <- loadECOMS(dataset = "System4_seasonal_15" , var = inputvars[1],
                     lonLim = c(-10,35), latLim = c(35,47),
                     season = 6:9, years = 1981:2010,
                     leadMonth = 1)

obs.multigrid <- makeMultiGrid(obs.hurs, obs.tas, obs.tp, obs.wss)
s4.multigrid <- makeMultiGrid(s4.hurs, s4.tas, s4.tp, s4.wss)
obs.fwi <- fwiGrid(multigrid = obs.multigrid, landmask = NULL)
s4.fwi <- fwiGrid(multigrid = s4.multigrid, landmask = s4.landmask, parallel = TRUE)

hindcast_bc <- biasCorrection(y = obs.fwi,
                              x = s4.fwi,
                              newdata = NULL,
                              method = "eqm",
                              cross.val = "loocv",
                              window = c(30,5))
```

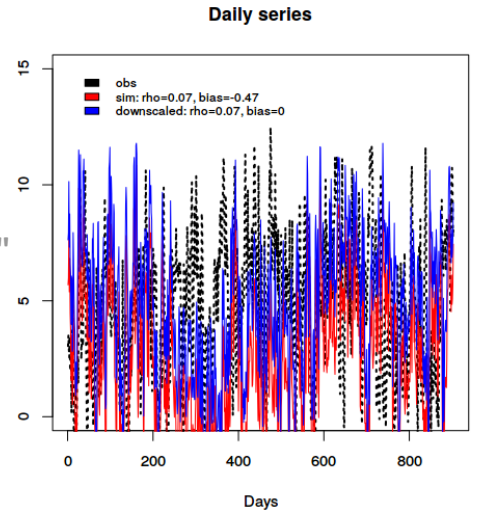
… and a quick overview of the correction effect

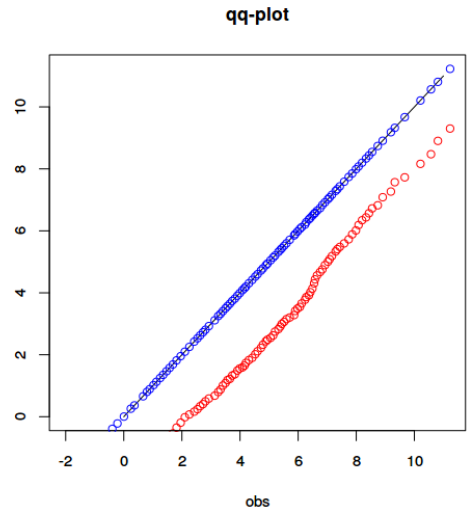
```
inputvars <- c("hurs", "wss", "tas", "tp")
# Loading WFDEI (Observations)
obs.hurs <- loadECOMS(dataset = "WFDEI" , var = inputvars[1],
                      lonLim = c(-10,35), latLim = c(35,47),
                      season = 6:9, years = 1981:2010)
# Same for wss, tas and tp
# Loading System4 variables
s4.hurs <- loadECOMS(dataset = "System4_seasonal_15" , var = inputvars[1],
                     lonLim = c(-10,35), latLim = c(35,47),
                     season = 6:9, years = 1981:2010,
                     leadMonth = 1)

obs.multigrid <- makeMultiGrid(obs.hurs, obs.tas, obs.tp, obs.wss)
s4.multigrid <- makeMultiGrid(s4.hurs, s4.tas, s4.tp, s4.wss)
obs.fwi <- fwiGrid(multigrid = obs.multigrid, landmask = NULL)
s4.fwi <- fwiGrid(multigrid = s4.multigrid, landmask = s4.landmask, parallel = TRUE)

hindcast_bc <- biasCorrection(y = obs.fwi,
                              x = s4.fwi,
                              newdata = NULL,
                              method = "eqm",
                              cross.val = "loocv"
                              window = c(30,5))

quickDiagnostics(y, x, hindcast_bc,
                 location = c(-5,42),
                 members = 1)
```





Contents

Integration with existing tools

`easyVerification` package. Common validation scores for seasonal forecasting. Tutorials and worked examples for integration with `loadeR`, `loadeR.ECOMS` and `downscaleR`

Available documentation with worked examples:
http://www.meteo.unican.es/work/downscaler/wiki/docs/ecoms_bias_correction.pdf

EUPORIAS

Verification of seasonal forecasts from the ECOMS User Data Gateway: a worked example

Joaquín Bedia[1] and Jonas Bhend[2]

[1]Santander Met Group. University of Cantabria - CSIC (Spain)
[2]Federal Office of Meteorology and Climatology MeteoSwiss, Zurich, Switzerland

version 2 - '2015-05-18'

Annual aggregation of daily data ...

```
fwimean_bc <- aggregateGrid(grid = hindcast_bc,
                            aggr.y = list(FUN = "mean", na.rm = TRUE))
fwimean_obs <- aggregateGrid(grid = obs.fwi,
                             aggr.y = list(FUN = "mean", na.rm = TRUE))
```

... detrending ...

```
fwimean_bc <- aggregateGrid(grid = hindcast_bc,
                            aggr.y = list(FUN = "mean", na.rm = TRUE))
fwimean_obs <- aggregateGrid(grid = obs.fwi,
                             aggr.y = list(FUN = "mean", na.rm = TRUE))
pred <- detrendGrid(fwimean_bc)
obs <- detrendGrid(fwimean_obs)
```

… ROC Skill score with `easyVerification` …

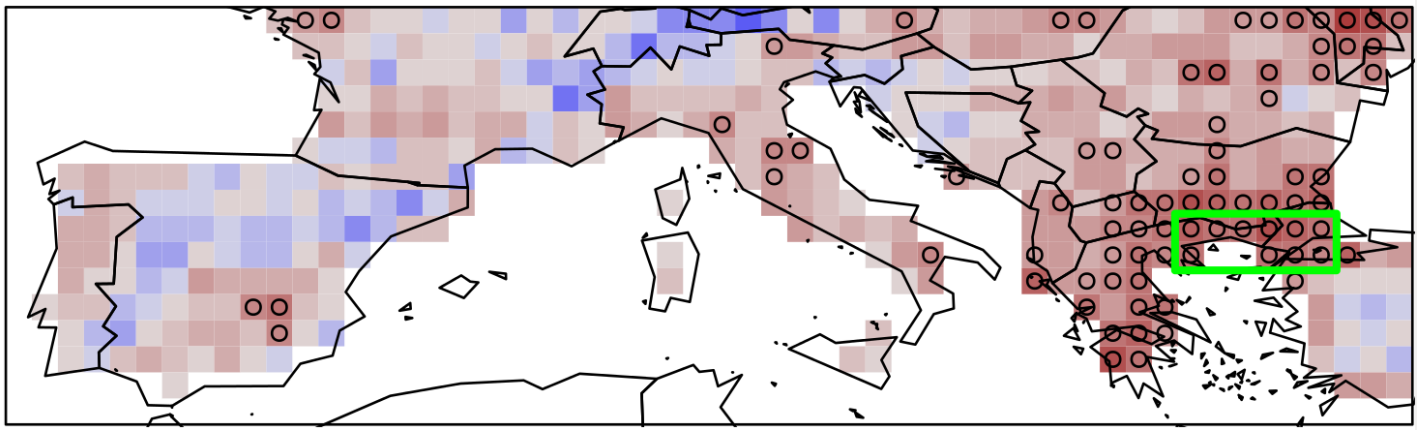
```
fwimean_bc <- aggregateGrid(grid = hindcast_bc,
                            aggr.y = list(FUN = "mean", na.rm = TRUE))
fwimean_obs <- aggregateGrid(grid = obs.fwi,
                             aggr.y = list(FUN = "mean", na.rm = TRUE))
pred <- detrendGrid(fwimean_bc)
obs <- detrendGrid(fwimean_obs)
rocss <- easyVerification::veriApply("EnsRocss",
                                     fcst = pred[["Data"]],
                                     obs = obs[["Data"]],
                                     prob = 2/3, ensdim = 1, tdim = 2)
```

… return to the loadeR – downscaleR grid format

```
fwimean_bc <- aggregateGrid(grid = hindcast_bc,
                            aggr.y = list(FUN = "mean", na.rm = TRUE))
fwimean_obs <- aggregateGrid(grid = obs.fwi,
                             aggr.y = list(FUN = "mean", na.rm = TRUE))
pred <- detrendGrid(fwimean_bc)
obs <- detrendGrid(fwimean_obs)
rocss <- easyVerification::veriApply("EnsRocss",
                                     fcst = pred[["Data"]],
                                     obs = obs[["Data"]],
                                     prob = 2/3, ensdim = 1, tdim = 2)
upper.tercile <- easyVeri2grid(easyVeri.mat = rocss$cat2,
                               obs.grid = obs)
```

… and plot the resulting ROCSS map

```
fwimean_bc <- aggregateGrid(grid = hindcast_bc,
                            aggr.y = list(FUN = "mean", na.rm = TRUE))
fwimean_obs <- aggregateGrid(grid = obs.fwi,
                             aggr.y = list(FUN = "mean", na.rm = TRUE))
pred <- detrendGrid(fwimean_bc)
obs <- detrendGrid(fwimean_obs)
rocss <- easyVerification::veriApply("EnsRocss",
                                     fcst = pred[["Data"]],
                                     obs = obs[["Data"]],
                                     prob = 2/3, ensdim = 1, tdim = 2)
upper.tercile <- easyVeri2grid(easyVeri.mat = rocss$cat2,
                               obs.grid = obs)
plotClimatology(upper.tercile,
                scales = list(draw = TRUE),
                backdrop.theme = "countries"
                main = "ROCSS", sp.layout = list(sig))
```



Contents

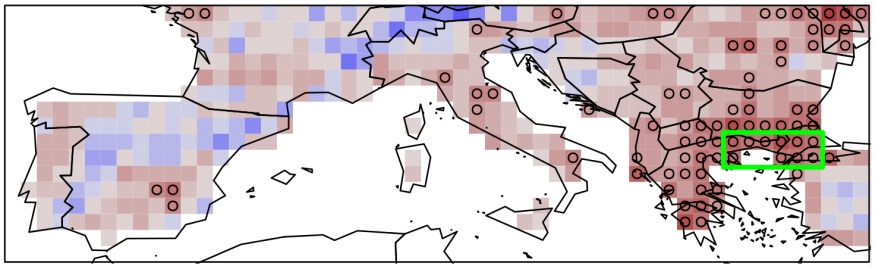
Integration with existing tools

`visualizeR` package. Special verification plots for forecast skill visualization. Tutorials and worked examples for integration with `loadeR`, `loadeR.ECOMS` and `downscaleR`

Documentation and worked examples in the wiki:
https://github.com/SantanderMetGroup/visualizeR/wiki

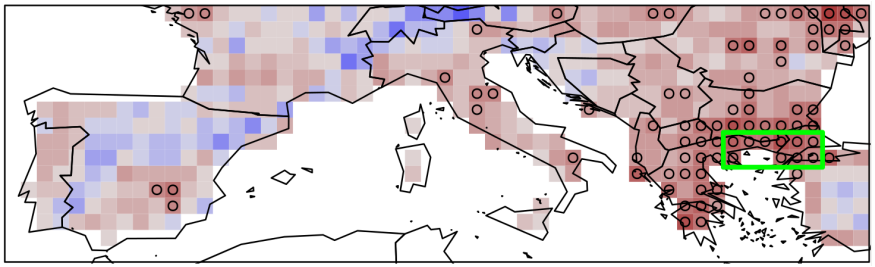
EUPORIAS

Spatial subset of a region of interest...

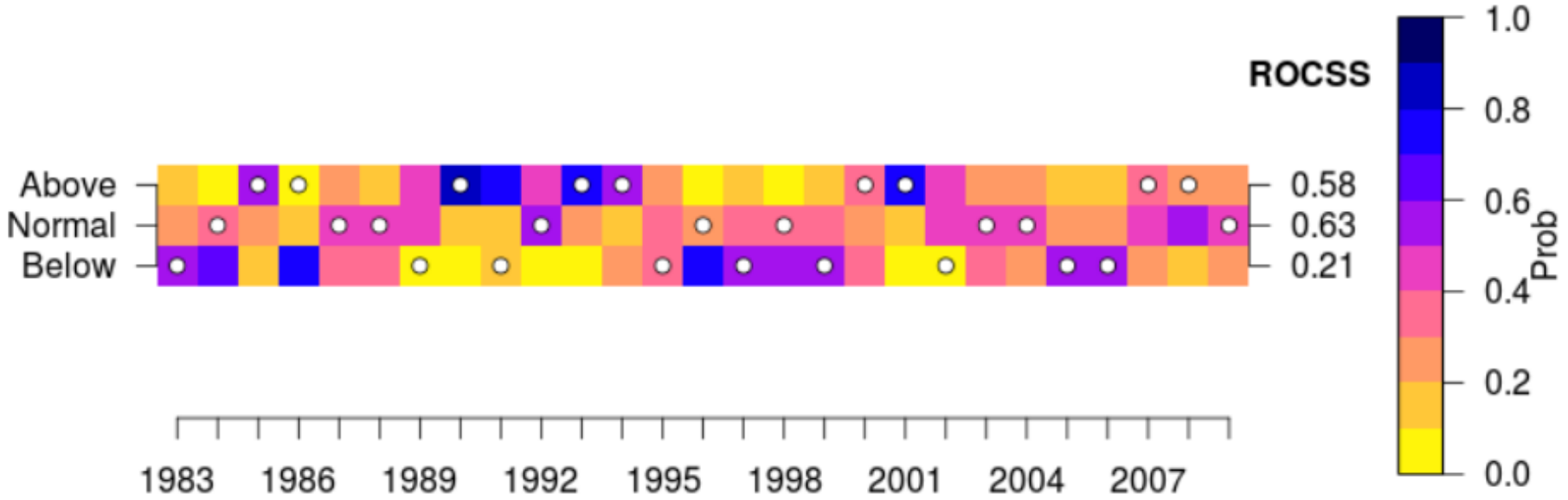


```
pred.sub <- subsetGrid(grid = pred,
                       latLim = c(40,42.5), lonLim = c(24,26.5))
obs.sub <- subsetGrid(grid = obs,
                      latLim = c(40,42.5), lonLim = c(24,26.5))
```

Tercile Validation Plot

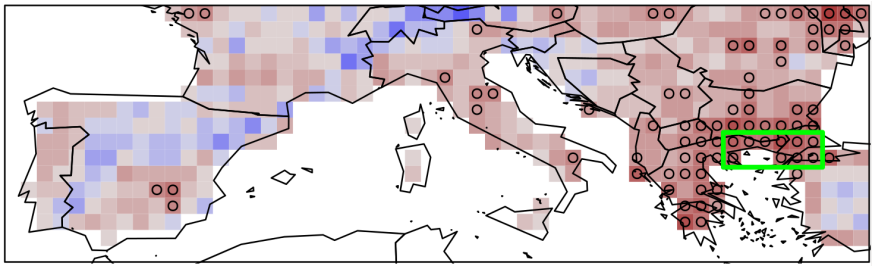


```
pred.sub <- subsetGrid(grid = pred,
                       latLim = c(40,42.5), lonLim = c(24,26.5))
obs.sub <- subsetGrid(grid = obs,
                      latLim = c(40,42.5), lonLim = c(24,26.5))
tercileValidation(pred = pred.sub, obs = obs.sub)
```

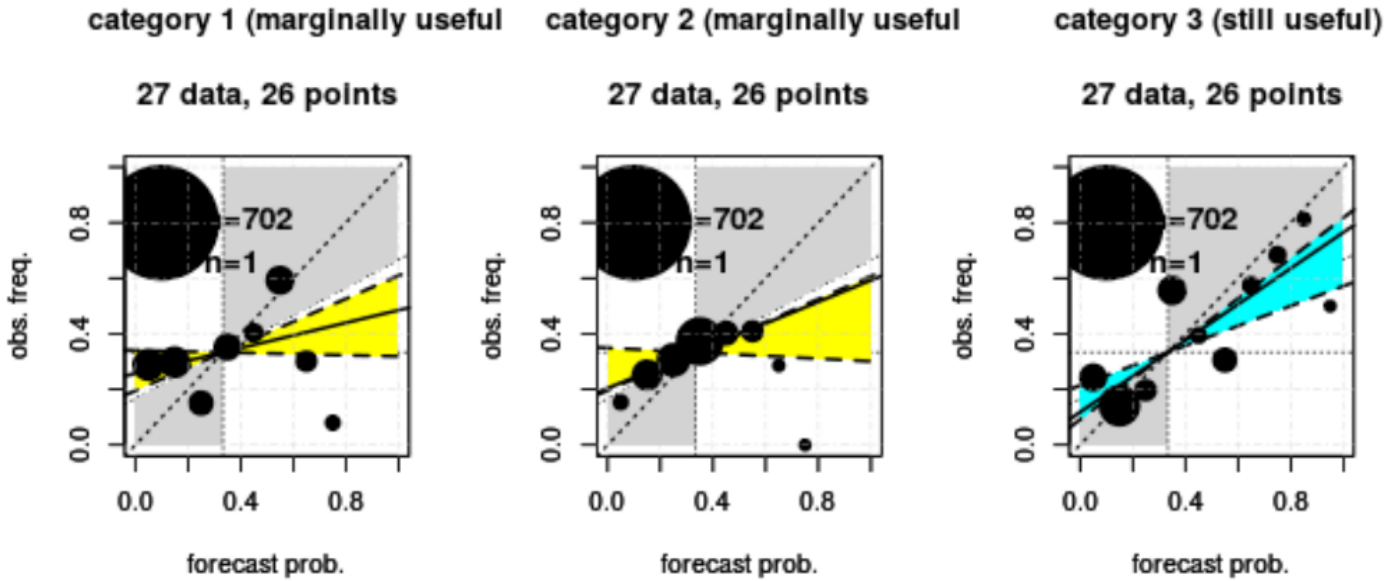


Implementation of tercile validation as pesented by:
Diez et al. 2011. doi:10.1111/j.1600-0870.2011.00523.x

Reliability Plots with reliability categories



```
pred.sub <- subsetGrid(grid = pred,
                       latLim = c(40,42.5), lonLim = c(24,26.5))
obs.sub <- subsetGrid(grid = obs,
                      latLim = c(40,42.5), lonLim = c(24,26.5))
tercileValidation(pred = pred.sub, obs = obs.sub)
reliabilityCategories(obs = obs.sub, pred = obs.sub, nbins = 3, sigboot = 0.75)
```



Implementation of reliability as presented by:
Weisheimer & Palmer 2014. doi:10.1098/rsif.2013.1162

Conclusions

- All datasets and variables required by ECOMS users centralized in a single **OpeNDAP Server** (The ECOMS-UDG) with two extra layers for
  - 1) authentication and
  - 2) harmonization

- A suite of **R packages** fully integrated allow performing many different tasks, such as:
  - User friendly access to the ECOMS-UDG
  - Data transformation
  - Downscaling and bias correction
  - Verification
  - Data visualization
  - Specific CII calculation
  - ….

OPEN-SOURCE BENEFITS
* Reproducibility
* Customizability
* Flexibility
* Interoperability
* Auditability
...

Contents

Links and packages

UDG wiki with instructions for registration

→ **http://www.meteo.unican.es/udg-wiki**

The ECOMS-UDG wiki

→ **http://www.meteo.unican.es/udg-wiki/ecoms**

Link to data access and downscaling packages:

→ **https://github.com/SantanderMetGroup/loadeR**

→ **https://github.com/SantanderMetGroup/downscaleR**

Thank you