Iberia01 - A new gridded dataset of daily precipitation and temperatures over Iberia.

Paper notebook - submitted to Earth System Science Data (ESSD)

S. Herrera, R.M. Cardoso, P.M.M. Soares, F. Espírito-Santo & J. M Gutiérrez

2019-10-04

Abstract

Iberia01 is a gridded dataset of daily precipitation and mean, maximum and minimum daily temperatures over the Iberian Peninsula covering the period 1971-2015 at a 0.1º spatial resolution. This dataset has been developed based on previous works done for Spain (Spain02) and Portugal (PT02). The present annex includes the code, in R, to reproduce some of the figures shown in the paper Herrera et al. 2019, Earth System Science Data (ESSD). DOI - http://dx.doi.org/10.20350/digitalCSIC/8641

Contents

1 R PACKAGE REQUIREMENTS

Iberia01 is provided in NetCDF format and, therefore, can be accessed using a number of packages and libraries. In this example, we build on the climate4R (https://github.com/SantanderMetGroup/climate4R) R framework for Climate Data Access and Post-processing which allows to easily read the data, compute indices and validation metrics and plot the results. To ensure reproducibility, it is recommend to install the same versions of the packages that

produced this document, even though in most cases more recent versions exist (see the last Session Information Section).

1.1 Installing climate4R

The *climate4R* R framework can be installed using conda or docker. Individual packages can be also installed from the GitHub repositories using the function `install_github` from package `devtools`. For, instance, to install `transformeR` v0.0.14, it suffices with pointing to the specific version tag after the `@` symbol.

Thus, in order to install the proper versions of the packages of the *climate4R* bundle needed to run these examples:

```r
devtools::install_github(c("SantanderMetGroup/transformeR", "SantanderMetGroup/convertR",
                           "SantanderMetGroup/loadeR", "SantanderMetGroup/visualizeR"))

rm(list = ls())
options(java.parameters = "-Xmx8000m")
library(loadeR, lib.loc = "~/R/x86_64-pc-linux-gnu-library/3.5/")
```

1.2 R package for data loading

The R package `loadeR` will perform the data loading task, including authentication against the Santander User Data Gateway (UDG) server (see the loadeR wiki) and enabling the access to any remote dataset via OPeNDAP, as well as creating and accesing datasets locally.

Prior to data access through the Santander Climate Data Service (via the User Data Gateway), authentication is required (see instructions for registration, https://github.com/SantanderMetGroup/loadeR/wiki/Loading-data-from-the-User-Data-Gateway-(UDG)). Once a user/password have been obtained, the authentication is performed in one step:

```r
loginUDG(username = "---", password = "---") # Replace "---" by the personal user/password
```

1.3 R package for data post-processing and plotting

In addition, the R package `transformeR` enables climate data transformation (plotting, aggregation, subsetting, PCA/EOF analysis . . .). Further details on its capabilities and installation instructions are available in the wiki page. Moreover, the package `visualizeR` provides different ploting functions.

```
library(transformeR, lib.loc = "~/R/x86_64-pc-linux-gnu-library/3.5/")

library(convertR, lib.loc = "~/R/x86_64-pc-linux-gnu-library/3.5/")

library(visualizeR, lib.loc = "~/R/x86_64-pc-linux-gnu-library/3.5/")
```

1.3.1 Colour Palettes:

The package `RColorBrewer` is used to replicate the spectral color palette used in the paper in the correlation maps. Next, the palette `veri.colors` is defined, that will be used in the verification maps:

```
library(RColorBrewer, lib.loc = "~/R/x86_64-pc-linux-gnu-library/3.5/")
cols <- brewer.pal(n = 11, name = "Spectral")
veri.colors <- colorRampPalette(rev(cols))
```

2 Daily Precipitation Climatology:

First, we obtain the daily precipitation climatology and the differences between both datasets, Iberia01 and E-OBS v17.0. To this aim, the dataset and geographical domain should be defined:

```
eobsNcML <- "http://opendap.knmi.nl/knmi/thredds/dodsC/e-obs_0.25regular/rr_0.25deg_reg_v17.0.nc
ib01NcML <- "http://meteo.unican.es/tds5/dodsC/Iberia01/Iberia01_v1.0_010reg_aa_3d.ncml"
## Defining the temporal and geographical domain:
latLim <- c(34, 44)
lonLim <- c(-10, 6)
years <- c(1971:2010)
```

2.1   E-OBS v17.0:

First, we obtain the climatology for the E-OBS v17.0 dataset. To this aim we should know the variables included in the dataset

```
di <- dataInventory(eobsNcML)
## Variables:
names(di)
```

and load the corresponding data:

```
pr.eobs <- loadGridData(eobsNcML, "rr", dictionary = FALSE, lonLim = lonLim, latLim = latLim,
                        season = c(1:12), years = years, time = "DD", aggr.d = "sum")
setGridUnits(pr.eobs, "mm", var = "rr")
pr.eobs <- climatology(pr.eobs, clim.fun = list(FUN = "mean", na.rm = TRUE))
spatialPlot(pr.eobs, backdrop.theme = "countries", at = seq(-0.5,6,0.5),
            col.regions = veri.colors(51), colorkey = list(space = "bottom"),
            main = "Precipitation (1971-2010)")
```

Precipitation (1971-2010)
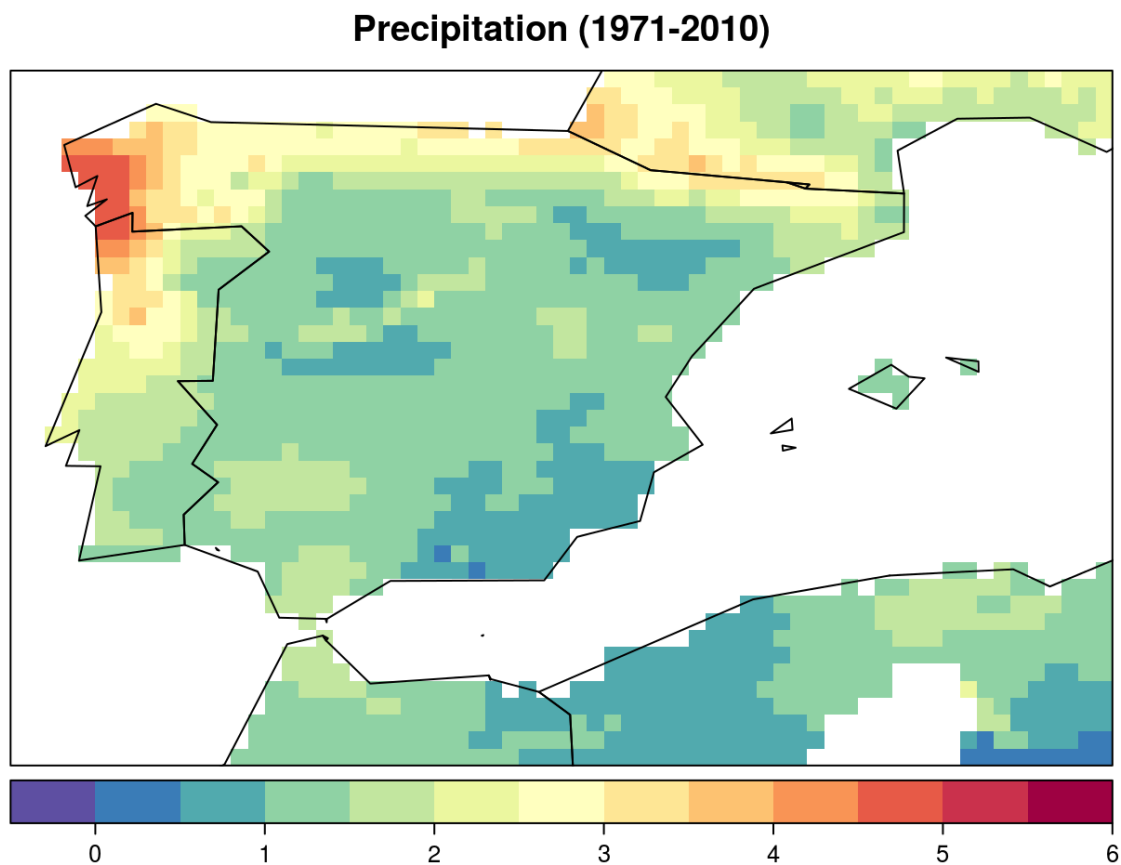
2.2 Iberia01:

Now, we consider the Iberia01 dataset:

```
di <- dataInventory(ib01NcML)
## Variables:
names(di)

pr.Ib01 <- loadGridData(ib01NcML, "pr", dictionary = FALSE, lonLim = lonLim, latLim = latLim,
                        season = c(1:12), years = years, time = "DD", aggr.d = "sum")
setGridUnits(pr.Ib01, "mm", var = "pr")
pr.Ib01 <- climatology(pr.Ib01, clim.fun = list(FUN = "mean", na.rm = TRUE))
spatialPlot(pr.Ib01, backdrop.theme = "countries", at = seq(-0.5,6,0.5),
            col.regions = veri.colors(51), colorkey = list(space = "bottom"),
            main = "Precipitation (1971-2010)")
```

Precipitation (1971-2010)
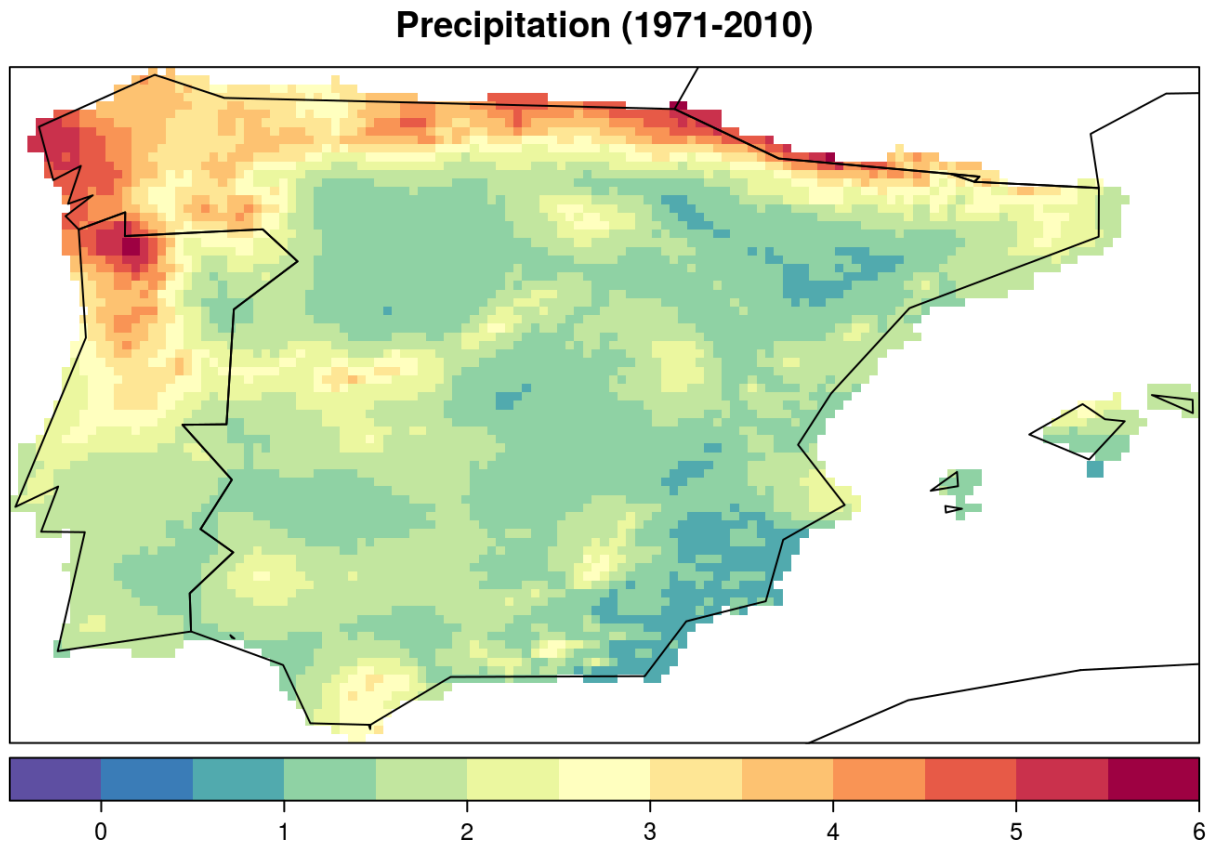
In order to obtain the difference between both datasets, we interpolate E-OBS v17.0 to the grid of the Iberian01 dataset:

```
## E-OBS climatology is regridded to the Iberian01 grid:
pr.eobs <- interpGrid(pr.eobs, new.coordinates = getGrid(pr.Ib01), method = "nearest")
```

2.3 Differences between E-OBS v17.0 and Iberia01:

- Bias:

```
bias <- gridArithmetics(pr.eobs, pr.Ib01, operator = "-")
spatialPlot(bias, backdrop.theme = "countries", at = seq(-3,3,0.5),
            col.regions = veri.colors(51), colorkey = list(space = "bottom"),
            main = "Bias (1971-2010)")
```
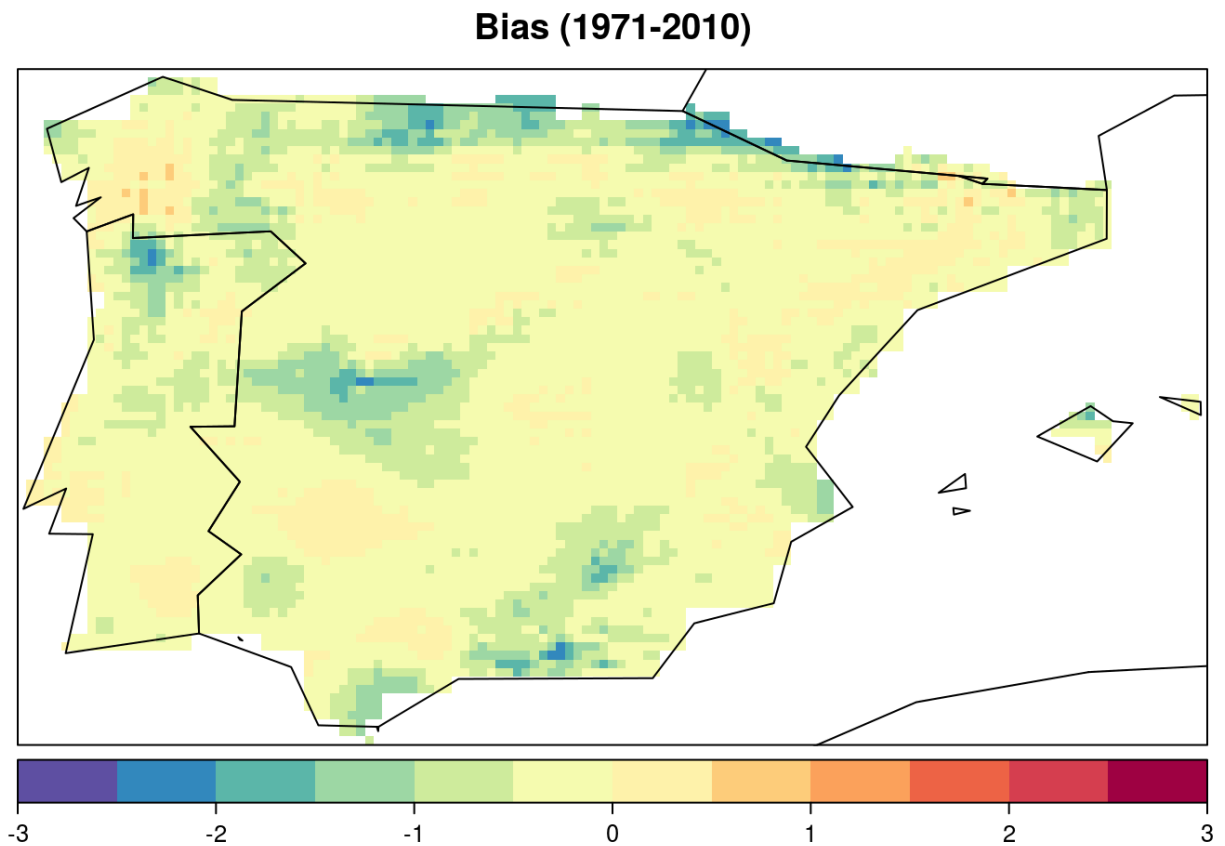
Bias (1971-2010)
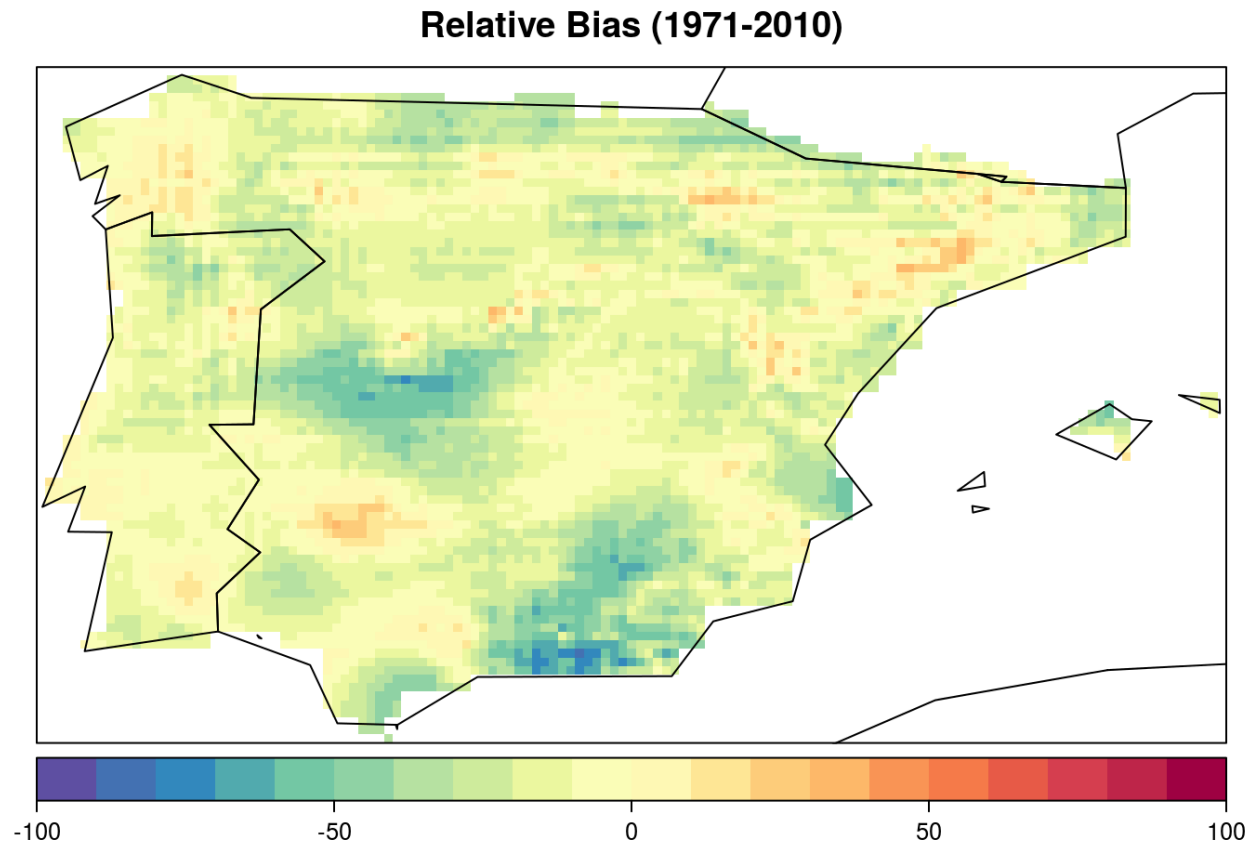

- Relative Bias:

```
## Relative Bias

C <- 100

relBias <- gridArithmetics(pr.eobs, pr.Ib01, pr.Ib01, C, operator = c("-", "/", "*"))

spatialPlot(relBias, backdrop.theme = "countries", at = seq(-100,100,10),
            col.regions = veri.colors(51), colorkey = list(space = "bottom"),
            main = "Relative Bias (1971-2010)")
```

Relative Bias (1971-2010)

3 Mean Daily Temperature:

Once the climatologies for some precipitation indices have been obtained we show the climatology of mean temperature. Note that, in this case, the datasets should be redefined, at least for E-OBS v17.0:

```
eobsNcML <- "http://opendap.knmi.nl/knmi/thredds/dodsC/e-obs_0.25regular/tg_0.25deg_reg_v17.0.nc
ib01NcML <- "http://meteo.unican.es/tds5/dodsC/Iberia01/Iberia01_v1.0_010reg_aa_3d.ncml"
## Defining the temporal and geographical domain:
latLim <- c(34, 44)
lonLim <- c(-10, 6)
years <- c(1971:2010)
```

3.1 E-OBS v17.0:

First, we obtain the climatology for the E-OBS v17.0 dataset. To this aim we should know the variables included in the dataset

```
di <- dataInventory(eobsNcML)
## Variables:
names(di)
```

and load the corresponding data:

```
tg.eobs <- loadGridData(eobsNcML, "tg", dictionary = FALSE, lonLim = lonLim, latLim = latLim,
                        season = c(1:12), years = years, time = "DD", aggr.d = "mean")
tg.eobs <- climatology(tg.eobs, clim.fun = list(FUN = "mean", na.rm = TRUE))
spatialPlot(tg.eobs, backdrop.theme = "countries", at = seq(0,25,1),
            col.regions = veri.colors(51), colorkey = list(space = "bottom"),
            main = "Mean Temperature (1971-2010)")
```

Mean Temperature (1971-2010)

3.2 Iberia01:

Now, we consider the Iberia01 dataset:

```
di <- dataInventory(ib01NcML)
```

```
## Variables:
```

```
names(di)
```

Mean Temperature (1971-2010)
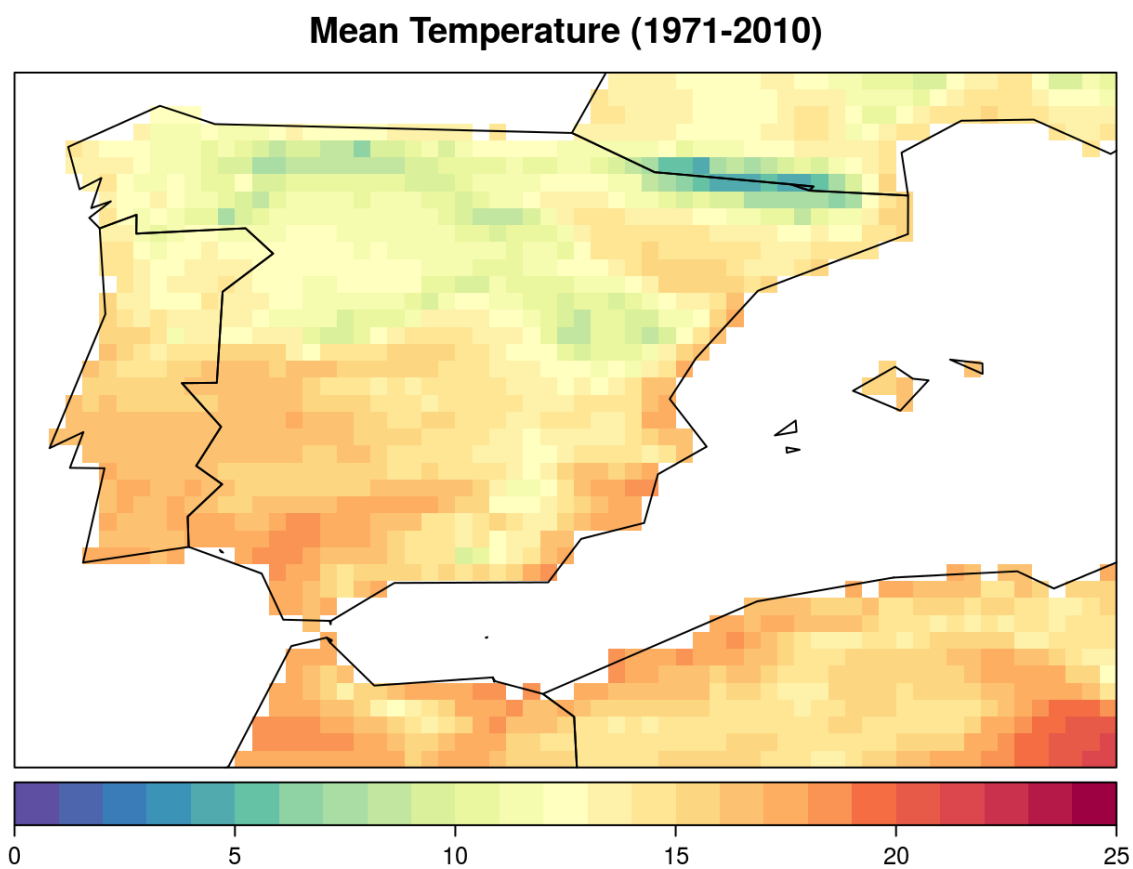

```
tg.Ib01 <- loadGridData(ib01NcML, "tas", dictionary = FALSE, lonLim = lonLim, latLim = latLim,
                        season = c(1:12), years = years, time = "DD", aggr.d = "mean")
tg.Ib01 <- climatology(tg.Ib01, clim.fun = list(FUN = "mean", na.rm = TRUE))
spatialPlot(tg.Ib01, backdrop.theme = "countries", at = seq(0,25,1),
            col.regions = veri.colors(51), colorkey = list(space = "bottom"),
            main = "Mean Temperature (1971-2010)")
```
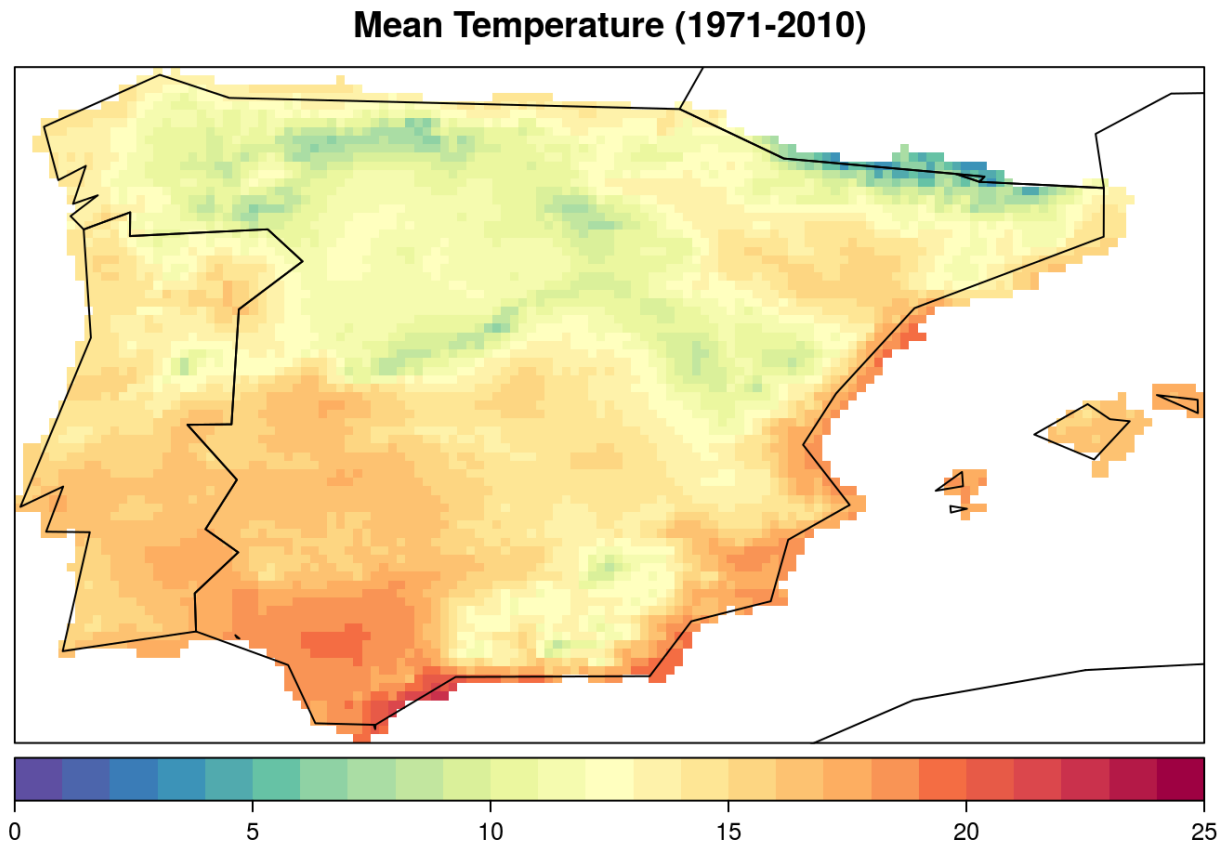
In order to obtain the difference between both datasets, we interpolate E-OBS v17.0 to the grid of the Iberian01 dataset:

```
## E-OBS climatology is regridded to the Iberian01 grid:
tg.eobs <- interpGrid(tg.eobs, new.coordinates = getGrid(tg.Ib01), method = "nearest")
```

3.3 Differences between E-OBS v17.0 and Iberia01:

- Bias:

```
bias <- gridArithmetics(tg.eobs, tg.Ib01, operator = "-")
spatialPlot(bias, backdrop.theme = "countries", at = seq(-5,5,0.5),
           col.regions = veri.colors(51), colorkey = list(space = "bottom"),
           main = "Bias (1971-2010)")
```


Bias (1971-2010)
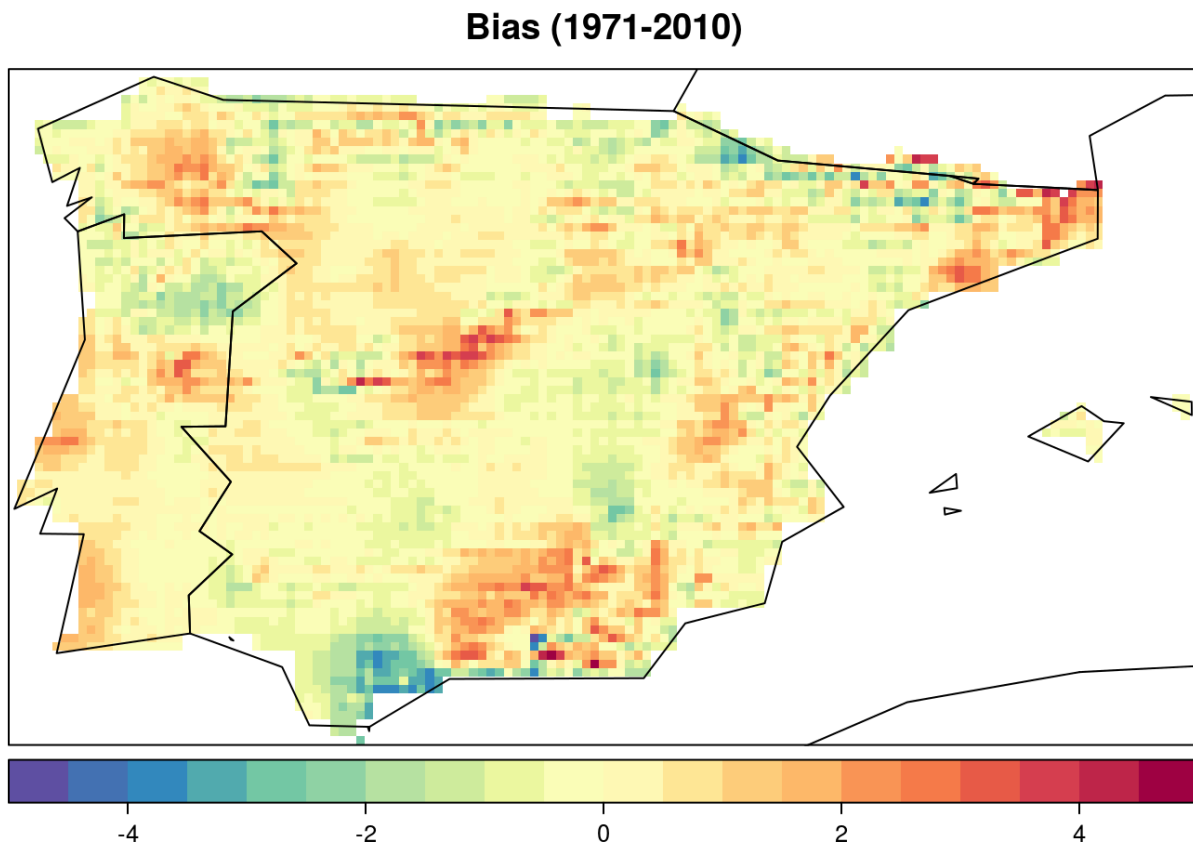
4 Additional packages

4.1 Functions for Extreme Value Distributions (evd):

In order to illustrate the analysis included in Herrera et al. 2019 the evd package should be installed
to obtain the Generalized Extreme Value (GEV) distribution.

```r
install.packages("evd")

library(evd, lib.loc = "~/R/x86_64-pc-linux-gnu-library/3.5/")
```

In order to extend the analysis shown in this document to other indices, the reader can consider the following R packages to easily estimate other climate and drought indices.

4.2 R package for Climate Indices calculation:

climate4R.indices is the package to compute several indices within the climate4R framework, therefore is seamlessly integrated with the climate4R data structures, and provides support for parallel computing.

```r
devtools::install_github(c("SantanderMetGroup/climate4R.indices"))

library(climate4R.indices, lib.loc = "~/R/x86_64-pc-linux-gnu-library/3.5/")

indexShow()
```

```
##   code             longname   indexfun tn tx tm pr units
## 1   FD Number_of_frost_days     tn.th  1  0  0  0  days
## 2 TNth  Number_of_cold_days     tn.th  1  0  0  0  days
## 3 TXth   Number_of_hot_days     tx.th  0  1  0  0  days
## 4  GDD  Growing_Degree_Days     gdd.th  0  0  1  0  degC
## 5  CDD  Cooling_Degree_Days     cdd.th  1  1  1  0  days
## 6  HDD  Heating_Degree_Days     hdd.th  1  1  1  0  days
## 7    P           Percentile percentile  0  0  0  0  <NA>
##                                            description
## 1                      Count_of_days_when_TN<0degC
## 2                      Count_of_days_when_TN<THdegC
## 3                      Count_of_days_when_TX>THdegC
## 4 Accumulated_sum_of_degrees_of_and_when_TM>THdegC
## 5                      Count_of_days_being_TH=22degC
## 6                    Count_of_days_being_TH=15.5degC
## 7                          Climate_value_or_percentile
```

```
?indexGrid    # see the examples
```

―

5 SESSION INFORMATION AND PACKAGE VERSIONS

```
print(sessionInfo(package = c("loadeR", "convertR", "transformeR",
                             "visualizeR", "RColorBrewer")))
```

```
## R version 3.5.2 (2018-12-20)

## Platform: x86_64-pc-linux-gnu (64-bit)

## Running under: Ubuntu 14.04.3 LTS

##

## Matrix products: default

## BLAS: /usr/lib/atlas-base/atlas/libblas.so.3.0

## LAPACK: /usr/lib/lapack/liblapack.so.3.0

##

## locale:
##  [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
##  [3] LC_TIME=es_ES.UTF-8        LC_COLLATE=en_US.UTF-8
##  [5] LC_MONETARY=es_ES.UTF-8    LC_MESSAGES=en_US.UTF-8
##  [7] LC_PAPER=es_ES.UTF-8       LC_NAME=es_ES.UTF-8
##  [9] LC_ADDRESS=es_ES.UTF-8     LC_TELEPHONE=es_ES.UTF-8
## [11] LC_MEASUREMENT=es_ES.UTF-8 LC_IDENTIFICATION=es_ES.UTF-8

##

## attached base packages:
## character(0)

##

## other attached packages:
## [1] loadeR_1.4.15     convertR_0.1.2     transformeR_1.5.1
## [4] visualizeR_1.4.0  RColorBrewer_1.1-2

##
```

```
## loaded via a namespace (and not attached):
##  [1] pkgload_1.0.2          maps_3.3.0
##  [3] SpecsVerification_0.5-2 dotCall64_1.0-0
##  [5] sm_2.2-5.6             assertthat_0.2.1
##  [7] sp_1.3-1              latticeExtra_0.6-28
##  [9] grDevices_3.5.2        yaml_2.2.0
## [11] remotes_2.0.2          sessioninfo_1.1.1
## [13] backports_1.1.4        lattice_0.20-38
## [15] glue_1.3.1            base_3.5.2
## [17] RcppEigen_0.3.3.5.0    digest_0.6.20
## [19] colorspace_1.4-1       htmltools_0.3.6
## [21] Matrix_1.2-15          devtools_2.0.1
## [23] raster_3.0-2          padr_0.5.0
## [25] scales_1.0.0          processx_3.3.0
## [27] CircStats_0.2-6        dtw_1.21-3
## [29] proxy_0.4-23          datasets_3.5.2
## [31] usethis_1.4.0          withr_2.1.2
## [33] verification_1.42      pbapply_1.4-2
## [35] cli_1.1.0             magrittr_1.5
## [37] crayon_1.3.4          easyVerification_0.4.4
## [39] memoise_1.1.0          evaluate_0.13
## [41] ps_1.3.0              methods_3.5.2
## [43] fs_1.2.6              MASS_7.3-51.4
## [45] utils_3.5.2          pkgbuild_1.0.2
## [47] tools_3.5.2          data.table_1.12.2
## [49] prettyunits_1.0.2      stringr_1.4.0
## [51] munsell_0.5.0          stats_3.5.2
## [53] callr_3.1.1           climate4R.indices_0.0.1
## [55] akima_0.6-2           compiler_3.5.2
## [57] mapplots_1.5.1         evd_2.3-3
```

```
## [59] rlang_0.4.0          grid_3.5.2
## [61] RCurl_1.95-4.12      graphics_3.5.2
## [63] spam_2.3-0           bitops_1.0-6
## [65] tcltk_3.5.2          rmarkdown_1.11
## [67] boot_1.3-20          vioplot_0.3.2
## [69] codetools_0.2-16     curl_3.3
## [71] abind_1.4-5          R6_2.4.0
## [73] zoo_1.8-6            knitr_1.22
## [75] loadeR.java_1.1.1    udunits2_0.13
## [77] rprojroot_1.3-2      desc_1.2.0
## [79] rJava_0.9-11         stringi_1.4.3
## [81] parallel_3.5.2       Rcpp_1.0.2
## [83] fields_9.8-6         xfun_0.5
```