

# Full code for Example 2 of the paper ‘climate4R: An Ecosystem of R packages for Climate Data Access, Post-processing and Bias Correction’

*M. Iturbide, J. Bedia, S. Herrera, J. Baño, J. Fernández, M. D. Frías, R. Manzananas, D. San Martín, E. Cimadevilla, A.S. Cofiño, J. M. Gutiérrez*

*2018-03-23*

## Contents

### 1 Example 2: CORDEX Ensembles via the User Data Gateway 1

First we create color palettes

## 1 Example 2: CORDEX Ensembles via the User Data Gateway

### 1.0.1 Climate data loading from UDG

The function `loginUDG` setup the UDG credentials for data access in the current R session:

```
loginUDG("userUDG", "pswrdUDG")
```

In this case we will load Spain02 and CORDEX daily data for the Iberian Peninsula:

```
lon <- c(-10, 5)
lat <- c(36, 44)
```

#### 1.0.1.1 E-OBS observation data

Loading with `loadGridData`:

```
TX <- loadGridData(eobs,
  var = "tasmax",
  season = 1:12,
  lonLim = lon,
  latLim = lat,
  years = 1971:2000,
  dictionary = "dicEOBS.dic")
```

#### 1.0.1.2 Daily CORDEX data loading

If the data is to be loaded from the UDG, we can use function `UDG.datasets` to print the inventory of available **UDG** datasets, where the name, type and url are specified. In this particular case, we can use the “name” of the desired dataset instead of passing the complete url to `loadGridData`. For example, if we are interested in loading observations from the **E-OBS** dataset [[@haylock\\_european\\_2008](#)] and porjection data from the **CMIP5** [[@taylor\\_overview\\_2011](#)] **MPI** model, thus, next we filter the names returned by `UDG.datasets` using an appropriate pattern:

```
ds <- UDG.datasets()
ds$name[grepl("CORDEX44.*historical", ds$name)]
```

```
## [1] EUROCORDER44_ICHEC-EC-EARTH_r12i1p1_RCA4_v1_historical
## [2] EUROCORDER44_CERFACS-CNRM-CM5_r1i1p1_RCA4_v1_historical
## [3] EUROCORDER44_ICHEC-EC-EARTH_r1i1p1_RACMO22E_v1_historical
## [4] EUROCORDER44_ICHEC-EC-EARTH_r3i1p1_HIRHAM5_v1_historical
## [5] EUROCORDER44_IPSL-CM5A-MR_r1i1p1_RCA4_v1_historical
## [6] EUROCORDER44_MOHC-HadGEM2-ES_r1i1p1_RCA4_v1_historical
## [7] EUROCORDER44_MOHC-HadGEM2-ES_r1i1p1_RACMO22E_v1_historical
## [8] EUROCORDER44_MPI-ESM-LR_r1i1p1_CCLM4-8-17_v1_historical
## [9] EUROCORDER44_MPI-ESM-LR_r1i1p1_RCA4_v1_historical
## [10] EUROCORDER44_IPSL-CM5A-MR_r1i1p1_WRF331F_v1_historical
## [11] EUROCORDER44_MPI-CSC-MPI-M-MPI-ESM-LR_r1i1p1_REMO2009_v1_historical
## [12] EUROCORDER44_MPI-CSC-MPI-M-MPI-ESM-LR_r2i1p1_REMO2009_v1_historical
## [13] AFRCORDEX44_KNMI-ICHEC-EC-EARTH_r1i1p1_RACMO22T_v1_historical
## [14] AFRCORDEX44_SMHI-CCCma-CanESM2_r1i1p1_RCA4_v1_historical
## [15] AFRCORDEX44_SMHI-CNRM-CERFACS-CNRM-CM5_r1i1p1_RCA4_v1_historical
## [16] AFRCORDEX44_SMHI-CSIRO-QCCCE-CSIRO-Mk3-6-0_r1i1p1_RCA4_v1_historical
## [17] AFRCORDEX44_SMHI-ICHEC-EC-EARTH_r12i1p1_RCA4_v1_historical
## [18] AFRCORDEX44_SMHI-IPSL-IPSL-CM5A-MR_r1i1p1_RCA4_v1_historical
## [19] AFRCORDEX44_SMHI-MIROC-MIROC5_r1i1p1_RCA4_v1_historical
## [20] AFRCORDEX44_SMHI-MOHC-HadGEM2-ES_r1i1p1_RCA4_v1_historical
## [21] AFRCORDEX44_SMHI-MPI-M-MPI-ESM-LR_r1i1p1_RCA4_v1_historical
## [22] AFRCORDEX44_SMHI-NCC-NorESM1-M_r1i1p1_RCA4_v1_historical
## [23] AFRCORDEX44_SMHI-NOAA-GFDL-GFDL-ESM2M_r1i1p1_RCA4_v1_historical
## 153 Levels: AFRCORDEX44_KNMI-ICHEC-EC-EARTH_r1i1p1_RACMO22T_v1_historical ...

ensemble.h <- as.character(ds$url[grepl("EUROCORDER44.*historical", ds$name)])[1:6]

TXh.list <- lapply(ensemble.h, function(x)
  loadGridData(dataset = x,
    var = "tasmax",
    season = 1:12,
    lonLim = lon,
    latLim = lat,
    years = 1971:2000,
    dictionary = "dicCDX.dic"))
```

```
ensemble.f <- as.character(ds$url[grepl("EUROCORDER44.*rcp85", ds$name)])[1:6]
TXf.list <- lapply(ensemble.f, function(x)
  loadGridData(dataset = x,
    var = "tasmax",
    season = 1:12,
    lonLim = lon,
    latLim = lat,
    years = 2071:2100,
    dictionary = "dicCDX.dic"))
```

## 1.0.2 Data transformation

### 1.0.2.1 Spatial and temporal consistency among members:

```
# Temporal intersection
lapply(TXh.list, function(x) getShape(x))
```

```
## [[1]]
##   time   lat   lon
## 10958    26    32
##
## [[2]]
##   time   lat   lon
## 10958    26    32
##
## [[3]]
##   time   lat   lon
## 10958    26    32
##
## [[4]]
##   time   lat   lon
## 10958    26    32
##
## [[5]]
##   time   lat   lon
## 10950    26    32
##
## [[6]]
##   time   lat   lon
## 10748    26    32

TXh.list <- do.call("intersectGrid.time", list(TXh.list, which.return = 1:6))
TXf.list <- do.call("intersectGrid.time", list(TXf.list, which.return = 1:6))
```

Interpolate to eobs:

```
# Interpolation
TXh.list <- lapply(TXh.list, function(x) interpGrid(x, getGrid(TX)))
TXf.list <- lapply(TXf.list, function(x) interpGrid(x, getGrid(TX)))

m <- TX$Data[1,,]*0
mask.hist <- array(dim = c(getShape(TXh.list[[1]]["time"], dim(m))),
  for (i in 1:dim(mask.hist)[1]) mask.hist[i,,] <- m
mask.rcp <- array(dim = c(getShape(TXf.list[[1]]["time"], dim(m))),
  for (i in 1:dim(mask.rcp)[1]) mask.rcp[i,,] <- m

# Apply mask
TXh.list <- lapply(TXh.list, function(x)
  gridArithmetics(x, mask.hist, operator = "+"))
TXf.list <- lapply(TXf.list, function(x)
  gridArithmetics(x, mask.rcp, operator = "+"))
```

### 1.0.2.2 Create a multimember grid:

```
# Create a multimember grid
TXh.ens <- do.call("bindGrid.member", TXh.list)
TXf.ens <- do.call("bindGrid.member", TXf.list)
```

Note that spatialPlot recognizes a multi-member grid and displays a map for each member:

```
spatialPlot(climatology(TXf.ens), at = seq(5, 33, 1), backdrop.theme = "countries",
  col.regions = colorRampPalette(colstx), layout = c(3, 2), as.table = TRUE)
```

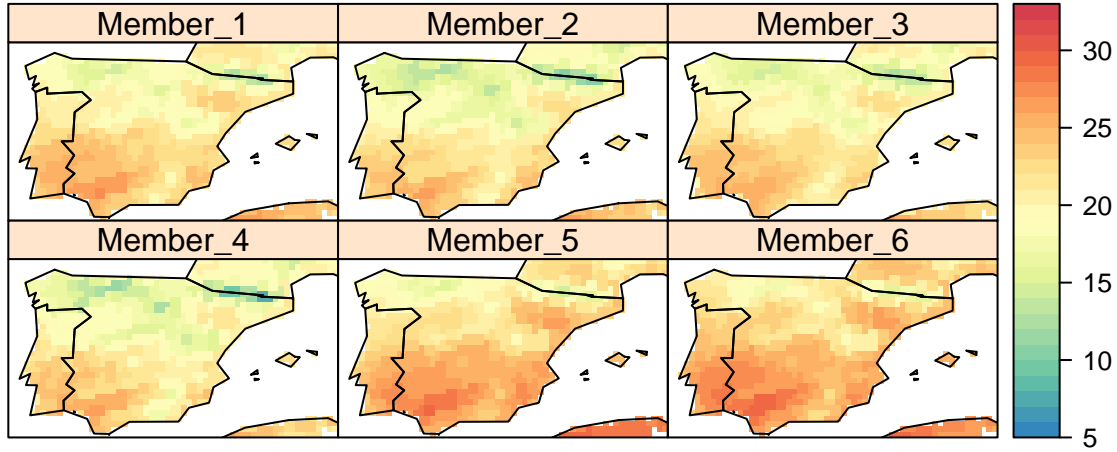


Figure 1: Maximum temperature (°C) in Iberia for an ensemble of 6 CORDEX RCMs under the RCP8.5 scenario and for future period 2071-2100. Fig. 8(above) in the manuscript.

Temporal plot

calculate bias for maximum temperature

*# Create a multimember grid*

```
TXh.list.ann <- lapply(TXh.list, function(x)
  aggregateGrid(x, aggr.y = list(FUN = "mean", na.rm = TRUE)))
TX.ann <- aggregateGrid(TX, aggr.y = list(FUN = "mean", na.rm = TRUE))
TXh.list.bias <- lapply(TXh.list.ann, function(x)
  gridArithmetics(x, TX.ann, operator = "-"))
```

```
TXh.bias.ens <- do.call("bindGrid.member", TXh.list.bias)
```

```
spatialPlot(climatology(TXh.bias.ens), at = seq(-10, 10, 1), backdrop.theme = "countries",
  col.regions = colorRampPalette(colsbias), layout = c(3, 2), as.table = TRUE)
```

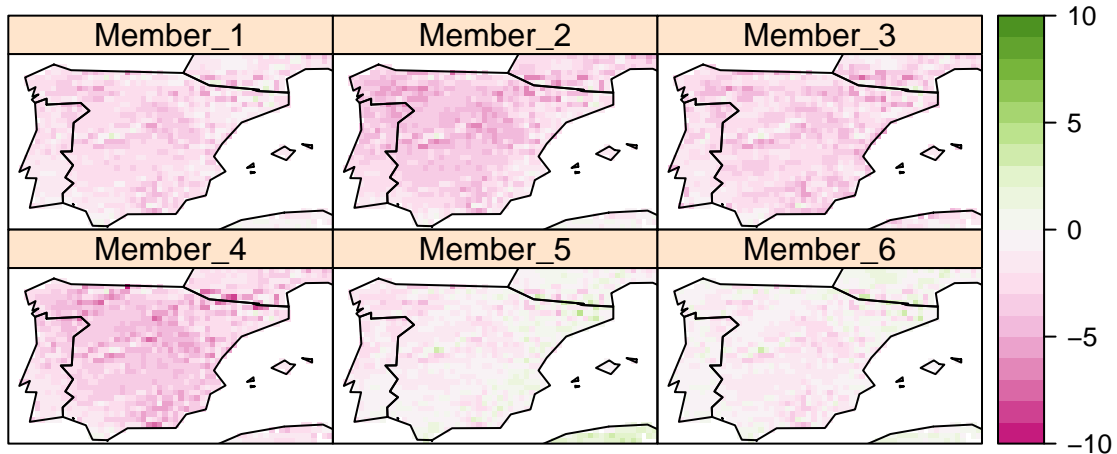


Figure 2: Bias of the maximum temperature (°C) in Iberia for an ensemble of 6 CORDEX RCMs w.r.t. E-OBS in the historical period 1971-2000. Fig. 8(below) in the manuscript.

We can use aggregateGrid to for example calculate the multimember mean and deviation

```
TXf.ens.mean <- aggregateGrid(TXf.ens, aggr.mem = list(FUN = mean, na.rm = TRUE))
TXf.ens.sd <- aggregateGrid(TXf.ens, aggr.mem = list(FUN = sd, na.rm = TRUE))

spatialPlot(climatology(TXf.ens.mean), at = seq(5, 33, 1),
            col.regions = colorRampPalette(colstx), backdrop.theme = "countries")
```

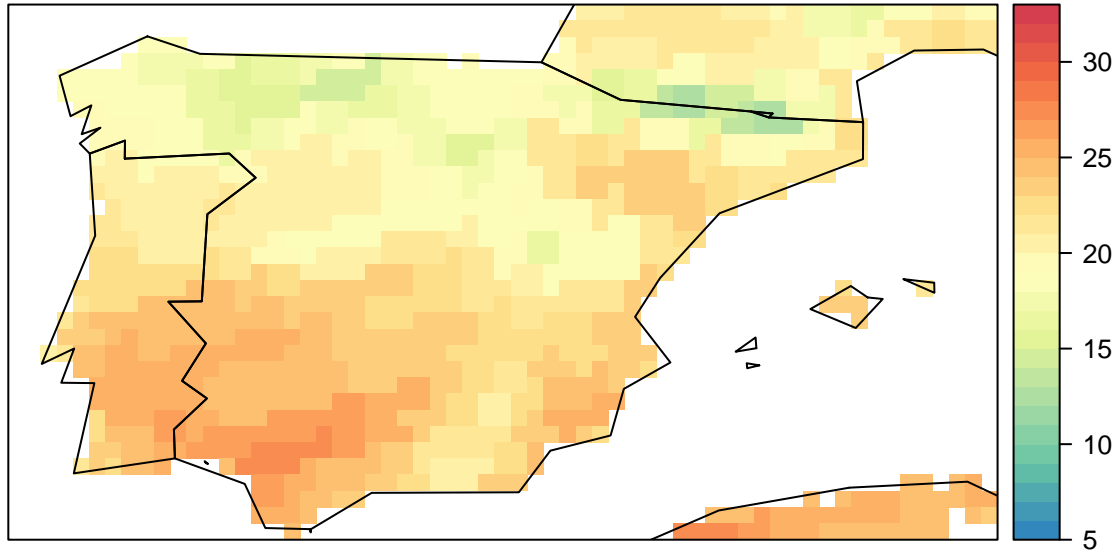


Figure 3: Ensemble mean of maximum temperature (°C) in Iberia for 6 CORDEX RCMs under the RCP8.5 scenario and for future period 2071-2100. Not shown in the manuscript.

```
spatialPlot(climatology(TXf.ens.sd), at = seq(0, 6, .5),
            col.regions = colorRampPalette(colssd), backdrop.theme = "countries")
```

### 1.0.3 ETCCDI index calculation from raw data

```
SUf.ens <- climdexGrid(tx = TXf.ens, index.code = "SU")

SUf.ens.mean <- aggregateGrid(SUf.ens, aggr.mem = list(FUN = mean, na.rm = TRUE))
SUf.ens.sd <- aggregateGrid(SUf.ens, aggr.mem = list(FUN = sd, na.rm = TRUE))

spatialPlot(climatology(SUf.ens.mean), backdrop.theme = "countries",
            at = seq(0, 260, 10), col.regions = colorRampPalette(colsindex))

spatialPlot(climatology(SUf.ens.sd), set.max = 50, backdrop.theme = "countries",
            at = seq(0, 50, 5), col.regions = colorRampPalette(colssd))
```

### 1.0.4 Bias correction

15 min operation

```
TXf.ens.bc <- biasCorrection(TX,
                            TXh.ens,
                            TXf.ens,
                            window = c(30, 7),
```

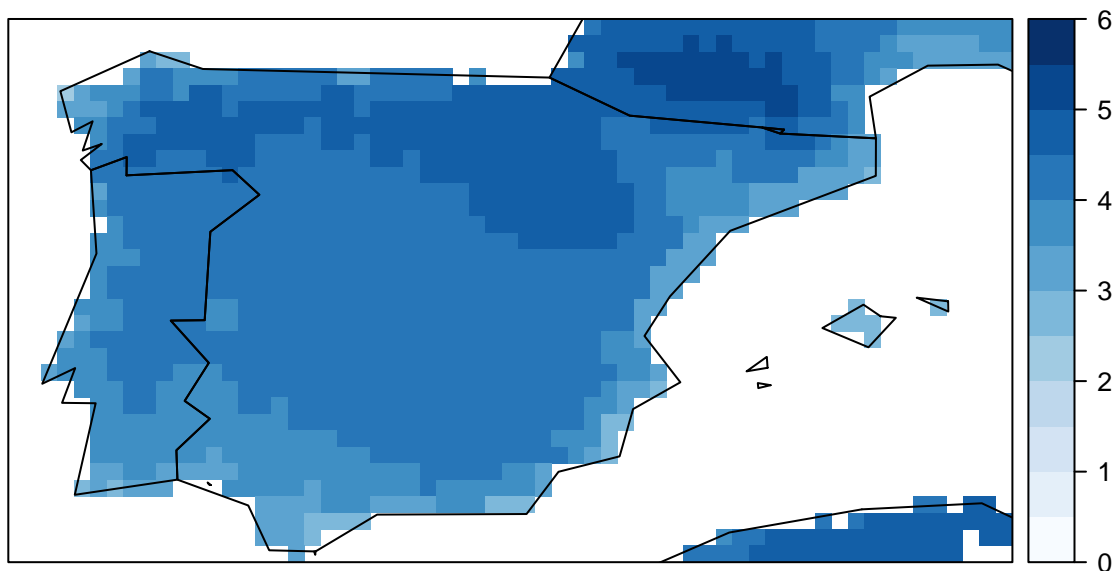


Figure 4: Ensemble standard deviation (sd) of maximum temperature ( $^{\circ}\text{C}$ ) in Iberia for 6 CORDEX RCMs under the RCP8.5 scenario and for future period 2071-2100. Not shown in the manuscript.

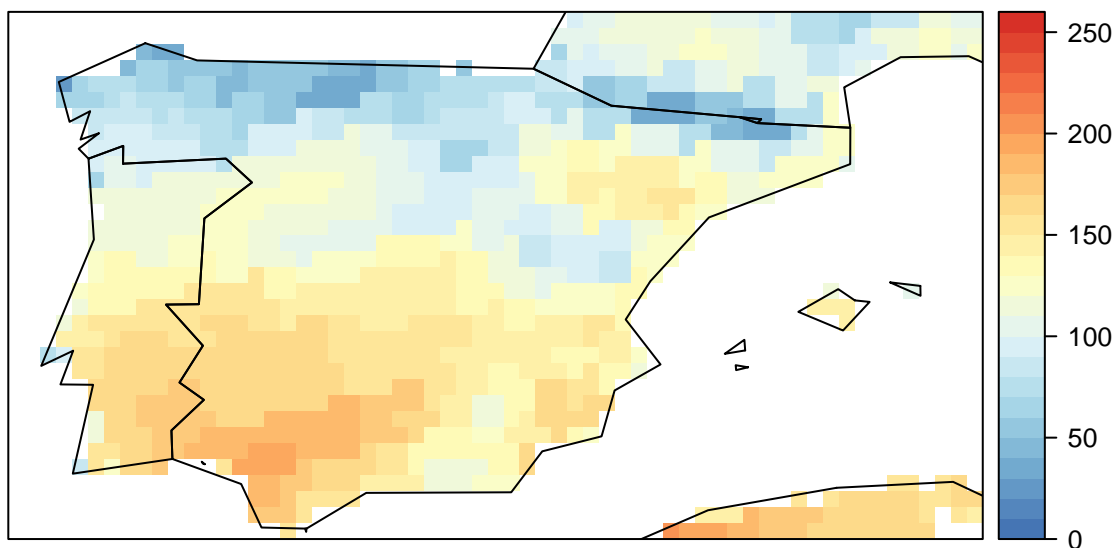


Figure 5: Summer days in Iberia for the future period 2071-2100 computed from the original RCM daily maximum temperature data. The figure shows the ensemble mean. Fig. 6(above, left) in the manuscript.

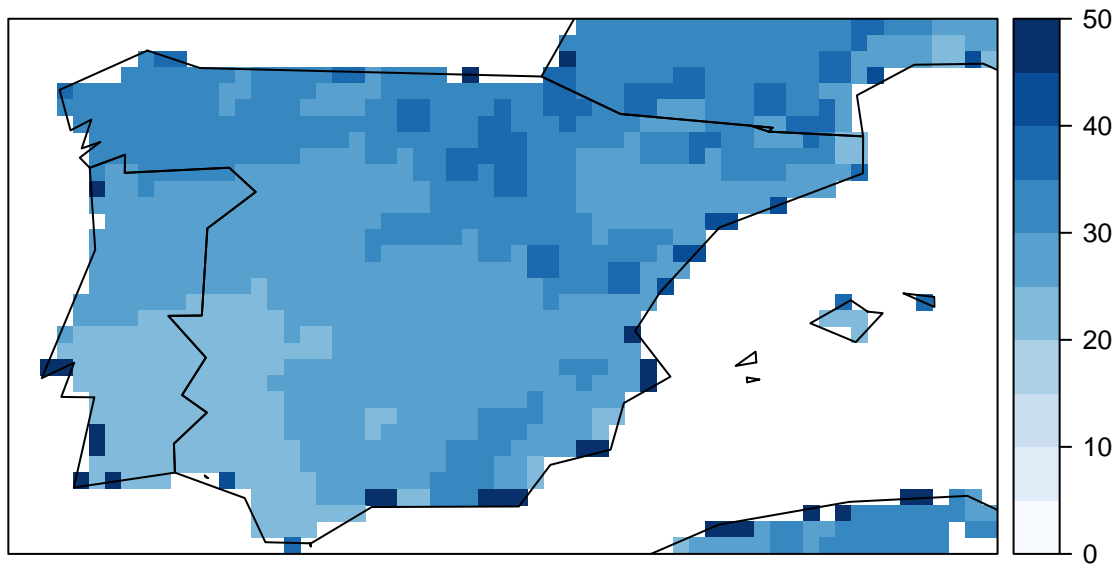


Figure 6: Standard deviation of summer days in Iberia for the future period 2071-2100 computed from the original RCM daily maximum temperature data. The figure shows the spread of the ensemble. Fig. 6(above, right) in the manuscript.

```

                                extrapolation = "constant",
                                method = "eqm")

TXf.ens.bc.mean <- aggregateGrid(TXf.ens.bc, aggr.mem = list(FUN = mean, na.rm = TRUE))
TXf.ens.bc.sd <- aggregateGrid(TXf.ens.bc, aggr.mem = list(FUN = sd, na.rm = TRUE))

spatialPlot(climatology(TXf.ens.bc.mean), backdrop.theme = "countries",
            at = seq(5, 33, 1), col.regions = colorRampPalette(colstx))

spatialPlot(climatology(TXf.ens.bc.sd), backdrop.theme = "countries",
            at = seq(0, 6, .5), col.regions = colorRampPalette(colssd))

```

### 1.0.5 ETCCDI index calculation from bias corrected data

```

SUf.ens.bc <- climdexGrid(tx = TXf.ens.bc, index.code = "SU")

SUf.ens.bc.mean <- aggregateGrid(SUf.ens.bc, aggr.mem = list(FUN = mean, na.rm = TRUE))
SUf.ens.bc.sd <- aggregateGrid(SUf.ens.bc, aggr.mem = list(FUN = sd, na.rm = TRUE))

spatialPlot(climatology(SUf.ens.bc.mean), backdrop.theme = "countries",
            at = seq(0, 230, 10), col.regions = colorRampPalette(colsindex))

spatialPlot(climatology(SUf.ens.bc.sd), backdrop.theme = "countries",
            at = seq(0, 50, 5), col.regions = colorRampPalette(colssd))

```

### 1.0.6 SU difference corrected non-corrected:

```

#A single location
SU <- climdexGrid(tx = TX, index.code = "SU")

```

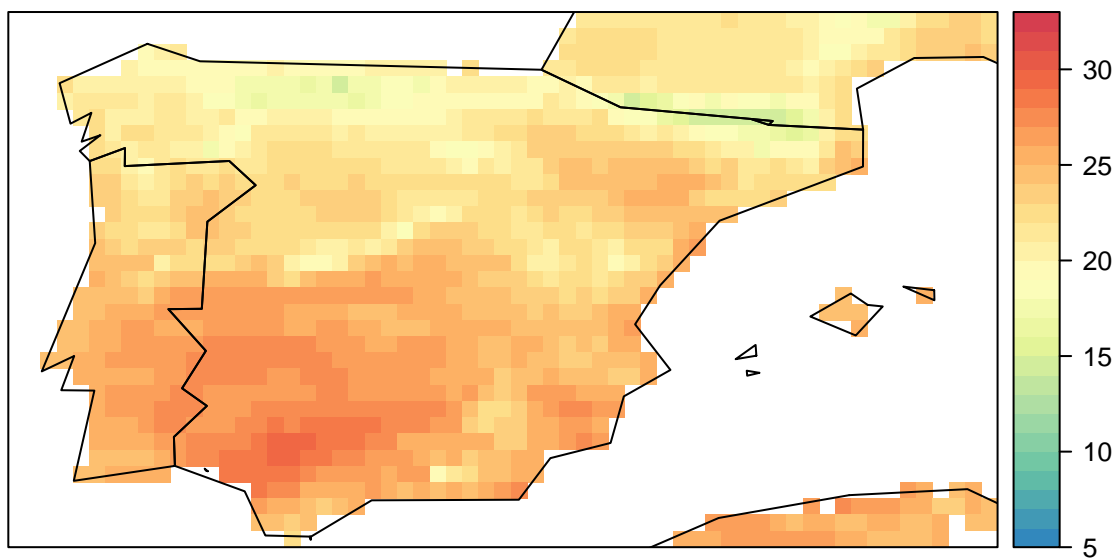


Figure 7: Ensemble mean of bias corrected maximum temperature ( $^{\circ}\text{C}$ ) in Iberia for 6 CORDEX RCMs under the RCP8.5 scenario and for future period 2071-2100. Not shown in the manuscript.

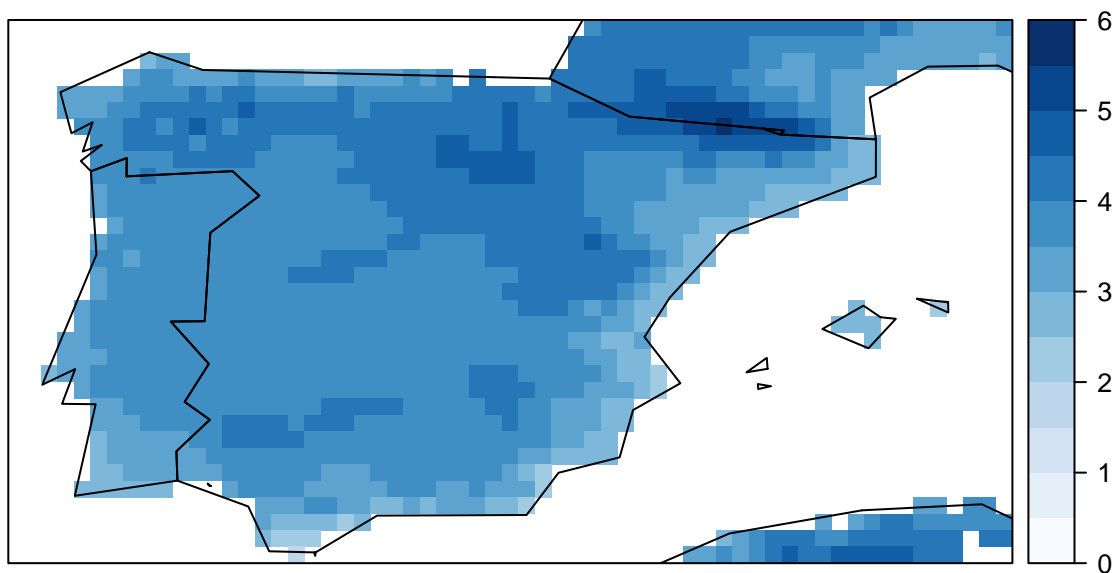


Figure 8: Ensemble standard deviation (sd) of bias corrected maximum temperature ( $^{\circ}\text{C}$ ) in Iberia for 6 CORDEX RCMs under the RCP8.5 scenario and for future period 2071-2100. Not shown in the manuscript.



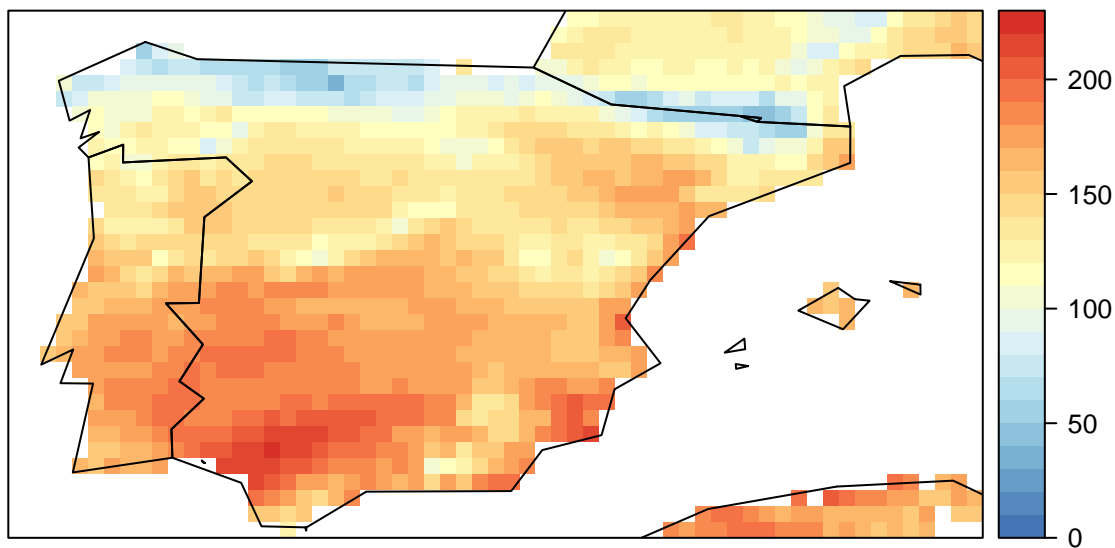


Figure 9: Summer days in Iberia for the future period 2071-2100 computed from the bias corrected RCM daily maximum temperature data. The figure shows the ensemble mean. Fig. 6(below, left) in the manuscript.

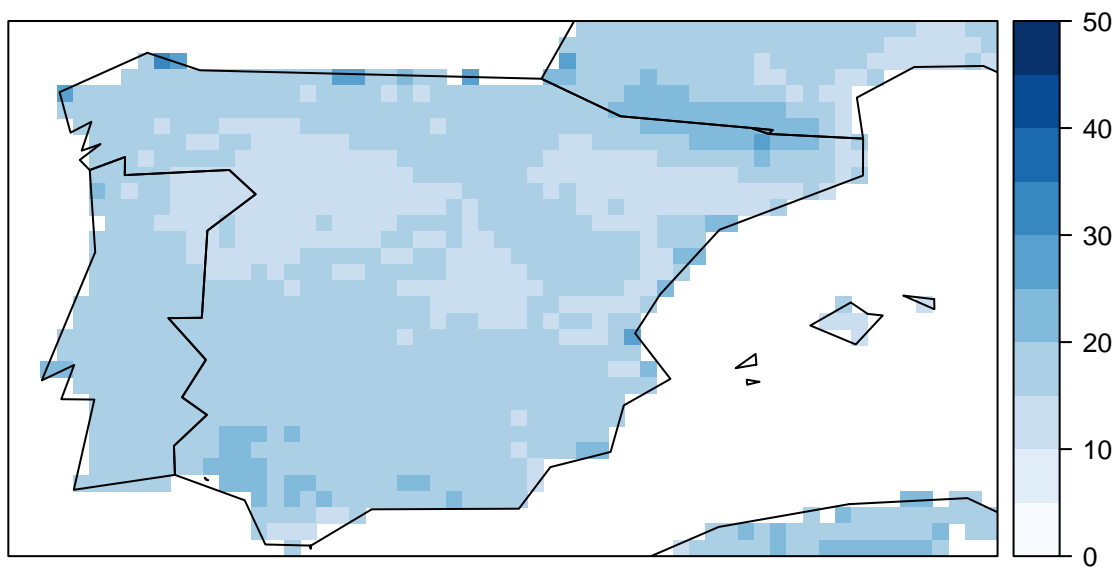


Figure 10: Standard deviation of summer days in Iberia for the future period 2071-2100 computed from bias corrected RCM daily maximum temperature data. The figure shows the spread of the ensemble. Fig. 6(below, right) in the manuscript.

```

SUh.ens <- climdexGrid(tx = TXh.ens, index.code = "SU")

SU.Z.eobs <- subsetGrid(SU, latLim = 41.64, lonLim = -0.89)
SU.Z.cdx <- lapply(list(SUh.ens, SUf.ens, SUf.ens.bc), function(x)
  subsetGrid(x, latLim = 41.64, lonLim = -0.89))
SU.Z.1m <- lapply(SU.Z.cdx, function(x)
  subsetGrid(x, members = 1))
SU.Z <- c(list(SU.Z.eobs), SU.Z.cdx, SU.Z.1m)

cols <- c("black", rep(c("red", "red", "blue"), 2))
temporalPlot(SU.Z,
  cols = cols,
  lty = rep(c(1,3), each = 4),
  lwd = 0.8,
  xyplot.custom = list(ylim = c(70, 220), ylab = "",
    key = list(space = "top", lines = list(lty = c(rep(1,4), 3),
      col = c(cols[1:4], cols[1]),
      lwd = 0.8),
    text = list(c("E-OBS", "CDX_hist", "CDX_rcp85", "CDX_rcp85.
      "1st_member"),
      cex = .7),
    columns = 3, rows = 3)))

```

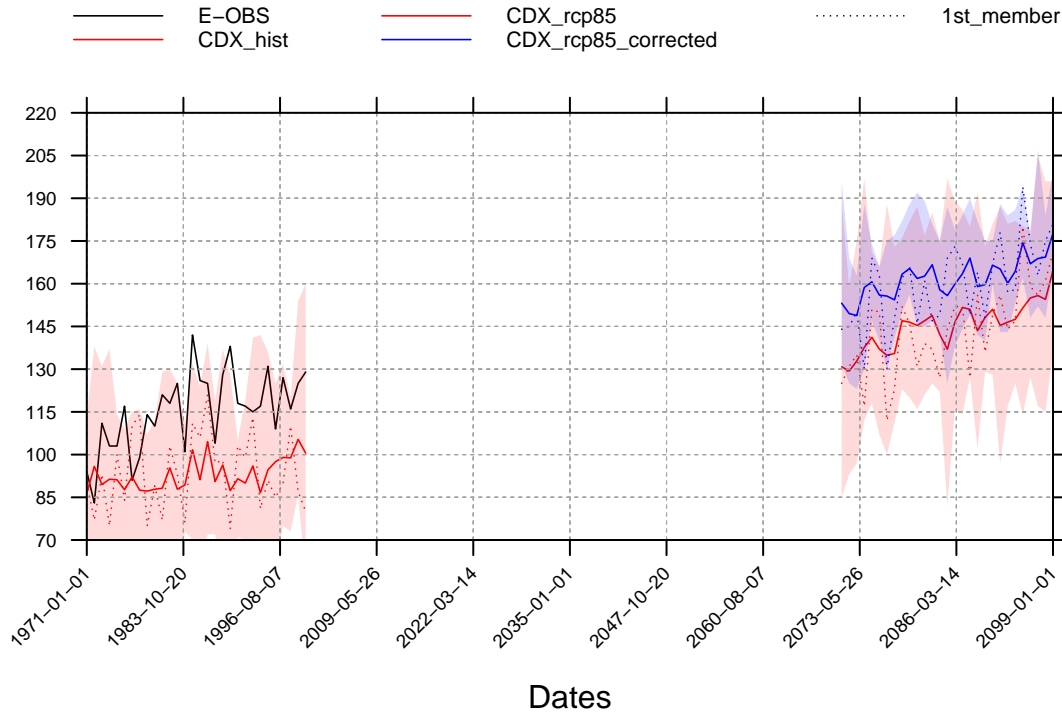


Figure 11: Annual summer days time series for a single gridbox (the one closest to Zaragoza, Spain) computed from (red) the original RCM daily maximum temperature data, and (blue) daily maximum temperature bias corrected data using E-OBS (black). When it comes to CORDEX data, continuous lines correspond to the ensemble mean and the shadowed area to the range (uncertainty). Dashed lines correspond to the 1st member of the ensemble. Fig. 7 in the manuscript.

## 1.0.7 Calculate other indexes

```
climindexShow()[,1:6]
```

### 1.0.7.1 Calculate CDD

```
eobs.pr <- "http://opendap.knmi.nl/knmi/thredds/dodsC/e-obs_0.25regular/rr_0.25deg_reg_v16.0.nc"
```

```
write(c("identifier,short_name,time_step,lower_time_bound,upper_time_bound,cell_method,offset,scale,deas  
      "tasmax,tx,24h,0,24,max,0,1,0,0,"  
      "pr,rr,24h,0,24,max,0,1,0,0,"), "dicEOBS.dic")
```

```
pr <- loadGridData(eobs.pr,  
                  var = "pr",  
                  season = 1:12,  
                  lonLim = lon,  
                  latLim = lat,  
                  years = 1971:2000,  
                  dictionary = "dicEOBS.dic")
```

```
write(c("identifier,short_name,time_step,lower_time_bound,upper_time_bound,cell_method,offset,scale,deas  
      "tasmax,tasmax,24h,0,24,max,-273.15,1,0,0,"  
      "pr,pr,24h,0,24,max,0,86400,0,0,"), "dicCDX.dic")
```

```
prh <- lapply(ensemble.h, function(x)  
             loadGridData(dataset = x,  
                           var = "pr",  
                           season = 1:12,  
                           lonLim = lon,  
                           latLim = lat,  
                           years = 1971:2000,  
                           dictionary = "dicCDX.dic"))  
prf <- lapply(ensemble.f, function(x)  
             loadGridData(dataset = x,  
                           var = "pr",  
                           season = 1:12,  
                           lonLim = lon,  
                           latLim = lat,  
                           years = 2071:2100,  
                           dictionary = "dicCDX.dic"))  
prh.t <- do.call("intersectGrid.time", list(prh, which.return = 1:6))  
prf.t <- do.call("intersectGrid.time", list(prf, which.return = 1:6))  
prh.ens <- do.call("bindGrid.member", prh.t)  
prf.ens <- do.call("bindGrid.member", prf.t)
```

```
CDDf.ens <- climindexGrid(pr = prf.ens, index.code = "CDD")  
CDDf.ens.interp <- interpGrid(CDDf.ens, getGrid(pr))  
CDDf.ens.mean <- aggregateGrid(CDDf.ens.interp, aggr.mem = list(FUN = "mean", na.rm = TRUE))  
CDDf.ens.sd <- aggregateGrid(CDDf.ens.interp, aggr.mem = list(FUN = "sd", na.rm = TRUE))  
  
m <- pr$Data[1,,]*0  
mask.cdd <- array(dim = c(getShape(CDDf.ens.mean)["time"], dim(m)))  
for (i in 1:dim(mask.cdd)[1]) mask.cdd[i,,] <- m  
CDDf.ens.mean <- gridArithmetics(CDDf.ens.mean, mask.cdd, operator = "+")
```

```
CDDf.ens.sd <- gridArithmetics(CDDf.ens.sd, mask.cdd, operator = "+")

spatialPlot(climatology(CDDf.ens.mean), backdrop.theme = "countries", at = seq(0, 60, 5),
  col.regions = colorRampPalette(colsindex))
```

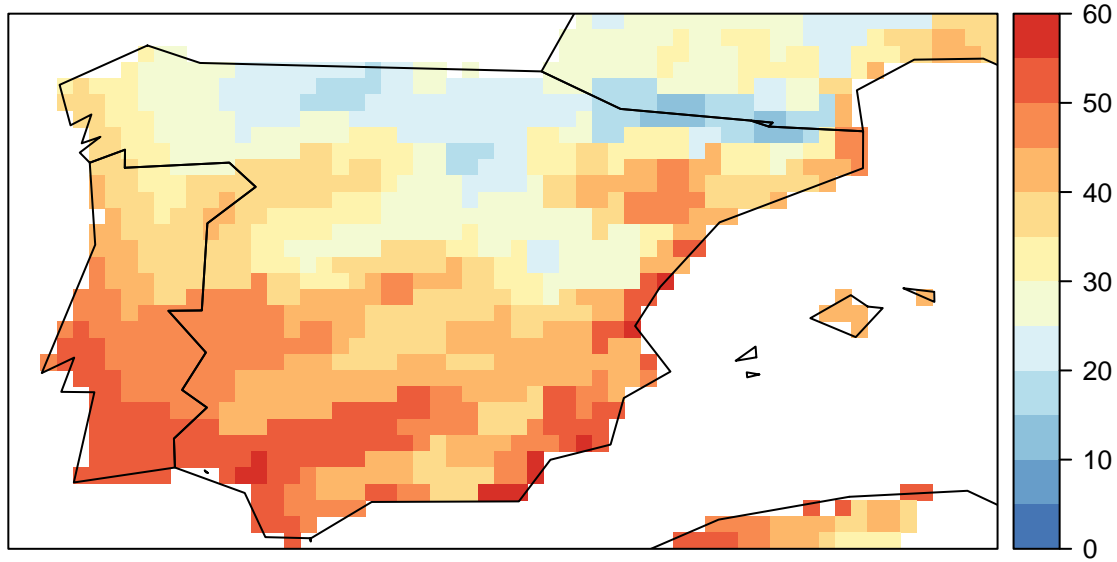


Figure 12: Consecutive dry days (CDD) in Iberia for the future period 2071-2100 computed from raw RCM daily precipitation data. The figure shows the the ensemble mean. Not shown in the manuscript.

```
spatialPlot(climatology(CDDf.ens.sd), backdrop.theme = "countries", at = seq(0, 60, 5),
  col.regions = colorRampPalette(colssd))

prf.ens.bc <- biasCorrection(y = pr,
  x = prh.ens,
  newdata = prf.ens,
  precipitation = TRUE,
  window = c(30, 7),
  extrapolation = "constant",
  method = "eqm",
  wet.threshold = 0.1)

CDDf.ens.bc <- climdexGrid(pr = prf.ens.bc, index.code = "CDD")
CDDf.ens.bc.mean <- aggregateGrid(CDDf.ens.bc, aggr.mem = list(FUN = "mean", na.rm = TRUE))
CDDf.ens.bc.sd <- aggregateGrid(CDDf.ens.bc, aggr.mem = list(FUN = "sd", na.rm = TRUE))

spatialPlot(climatology(CDDf.ens.bc.mean), backdrop.theme = "countries", at = seq(0, 60, 5),
  col.regions = colorRampPalette(colsindex))

spatialPlot(climatology(CDDf.ens.bc.sd), backdrop.theme = "countries", at = seq(0, 60, 5),
  col.regions = colorRampPalette(colssd))
```

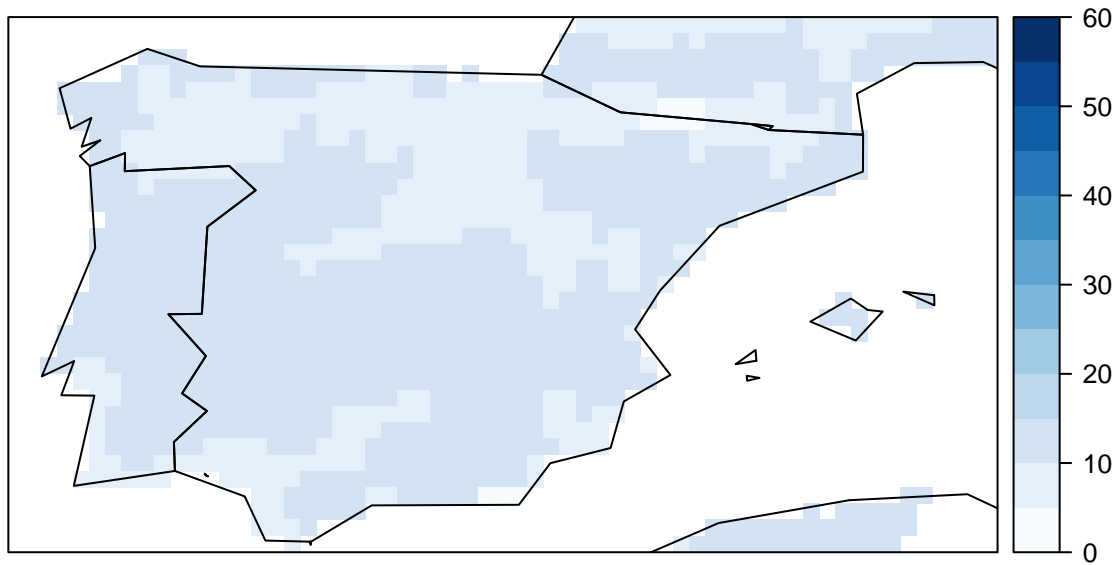


Figure 13: Standard deviation of consecutive dry days (CDD) in Iberia for the future period 2071-2100 computed from raw RCM daily precipitation data. The figure shows the spread of the ensemble. Not shown in the manuscript.

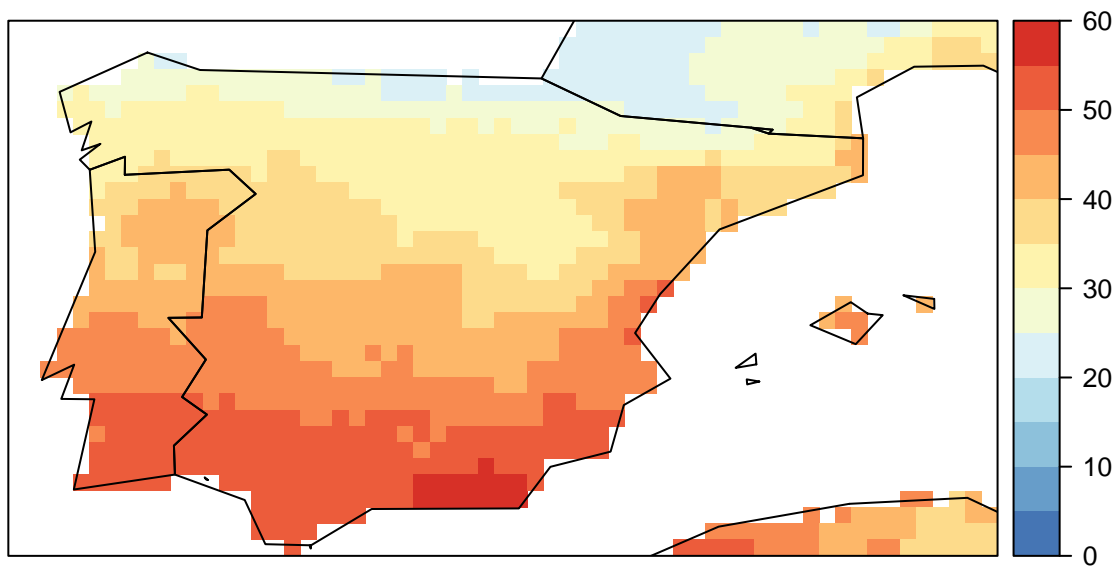


Figure 14: Consecutive dry days (CDD) in Iberia for the future period 2071-2100 computed from bias corrected RCM daily precipitation data. The figure shows the the ensemble mean. Not shown in the manuscript.

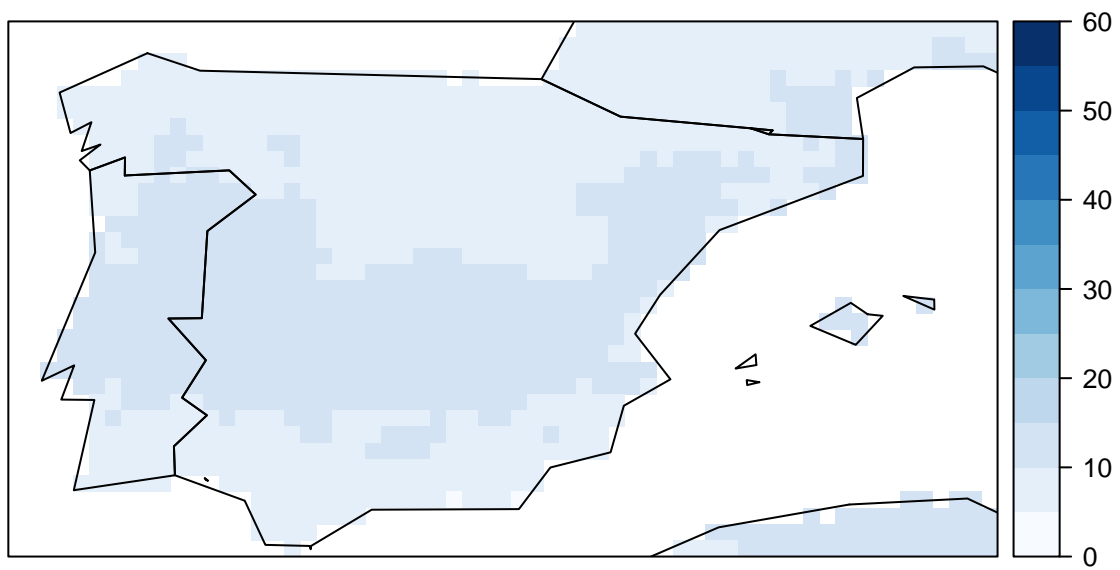


Figure 15: Standard deviation of consecutive dry days (CDD) in Iberia for the future period 2071-2100 computed from bias corrected RCM daily precipitation data. The figure shows the spread of the ensemble. Not shown in the manuscript.