Downscaling using Analogs

Carmen Álvarez-Castro[1,2]
[1]Universidad Pablo de Olavide, UPO, Seville, Spain
[2]Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici, CMCC, Bologna, Italy

mcalvcas@upo.es
carmen.alvarez-castro@cmcc.it

CMCC
Centro Euro-Mediterraneo
sui Cambiamenti Climatici

UNIVERSIDAD
PABLO DE OLAVIDE
SEVILLA

www.cmcc.it

1

Downscaling using Analogs

This method of downscaling tries to make use of the information we have on the dependency between large and local scales.

Experience provides clear examples of a dependency between large-scale conditions and local weather elements such as rainfall and temperature. For instance, the rainfall in many places is connected with the El Niño Southern Oscillation phenomenon (ENSO) or the North Atlantic Oscillation (NAO +/–).

The global climate models are able to predict the characteristics of large-scale conditions reasonably well but have some issues at local scale conditions since we need higher resolution, so we will use some method of statistical downscaling to increase the resolution considering the dynamics at large scale such as Analogs
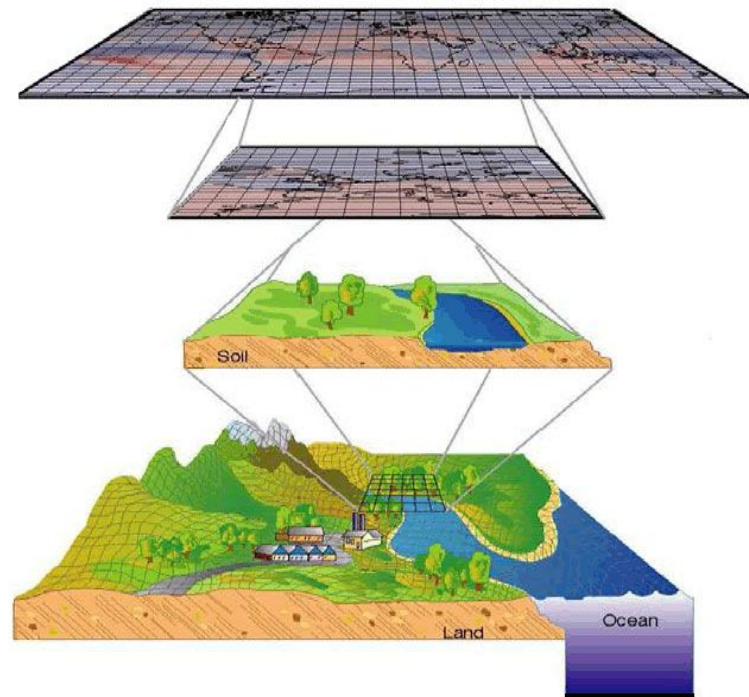
Analogs Method

Predictors: SLP and/or SST at Large Scale
Predictands: Precipitation, Temperature at local scale

Day of interest: **Analogs**
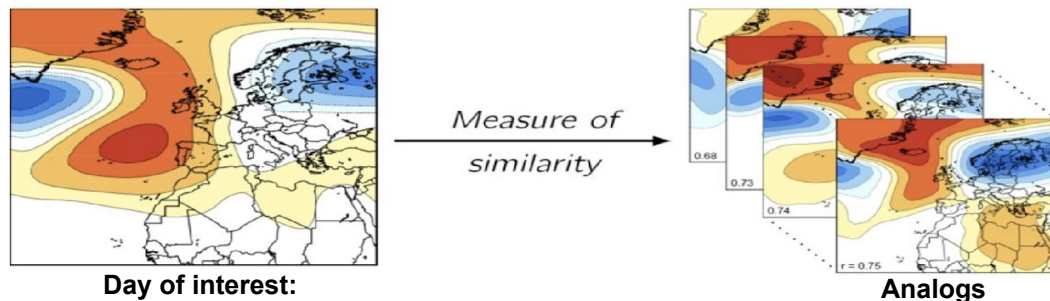
The **analogs** are days within the database which have a similar large scale pattern to the day of interest (Yiou et al, 2013, Jezequel et al, 2017).

The **downscaled field** is taking selecting the field of interest (i.e. precipitation, temperature) in the best analog day (1986/05/07) to the day of interest (2021/05/01) The best analog is defined as the day with minimum distance to the day of interest
at (1) large scale, (2) large scale and local scale, (3) large scale, local scale and maxima correlation (CST_Analogs and Analogs functions by Alvarez-Castro et al within Perez-Zanon et al, 2022 and CSTools, CRAN).

Day of interest: 20210501

**Best Analogs:**

| date	dist	corr
19860507	146	0.90
19920520	356	0.78
19980511	425	0.74
20010515	478	0.65
20130510	553	0.61
20090523	740	0.59

Downscaling using Analogs in seasonal forecast

- Example using Analogs in Temperature for JJA 2022 in seasonal forecast. Downscaling **from 1° to 0.25° of horizontal resolution**.
- Using the method of Analogs we can **downscale and bias correct** the data of seasonal forecast models since is based on observations/reanalyses.
- Although there is a limitation of a gridded dataset for observation, using **Analogs** we can downscale different variables and regions.
- Since is a downscaling based on the dynamics, a **large scale region** should be selected in advance. In the case of this example, the Mediterranean region, we have selected the North Atlantic region (based on a previous study of correlation not shown here)
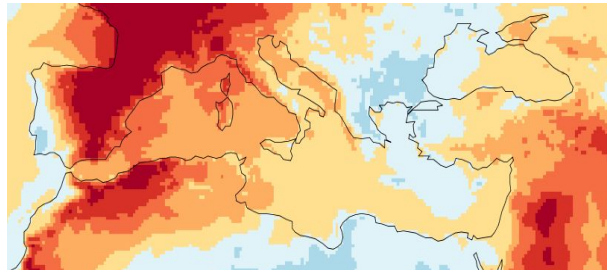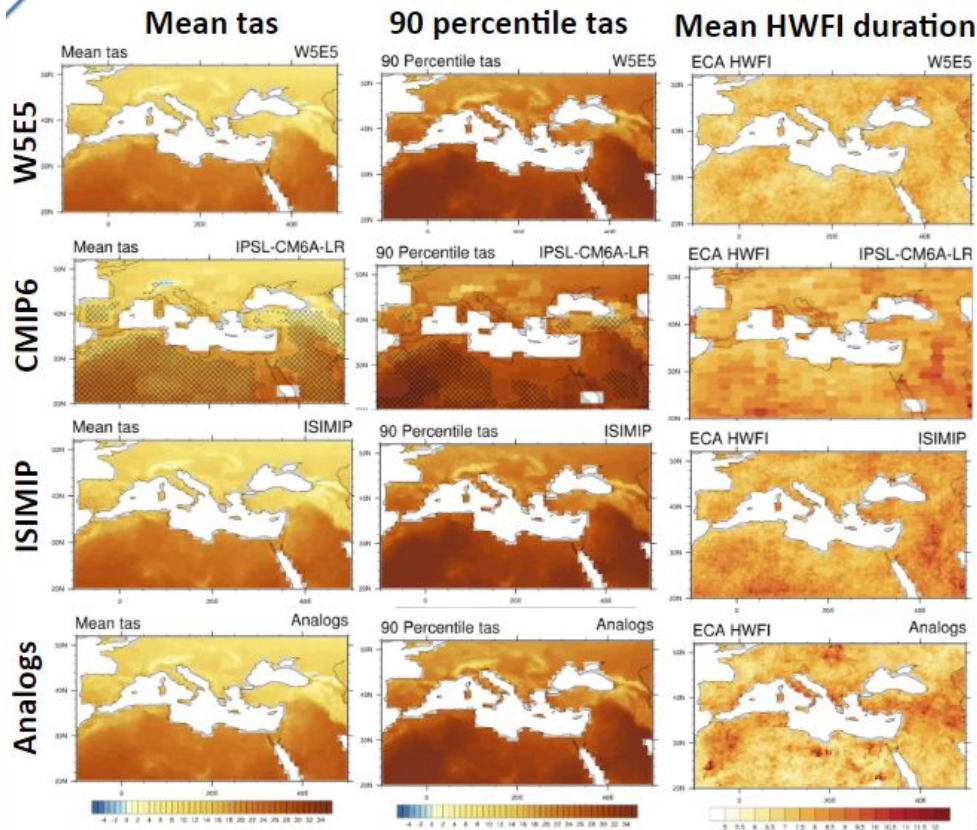
Temperature
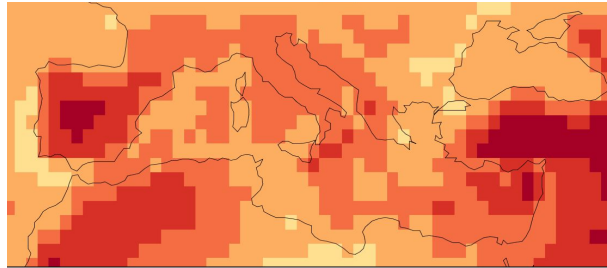

CMCC
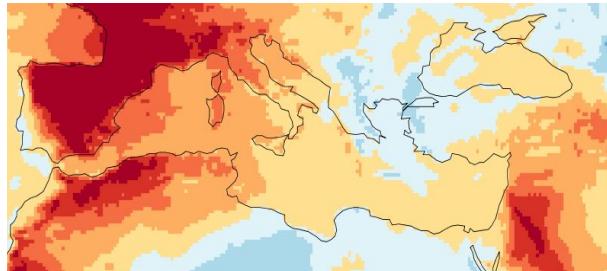
Downscaling using ANALOGS

OBS

(anom °C)

Downscaling using Analogs in Climate projections

Annual mean near-surface temperature and 90 percentile near-surface temperature are averaged over the period 1979-2014. Crossed regions are areas where difference between model and observation is larger than the sum of model and observation interannual variability (i.e. standard deviation). The mean duration of warm spell days index (ECA-HWFI) is calculated using the values of number of event and event duration computed with CDO. The timeseries show the normalized tas (absolute value divided by timeseries mean value) and standard HWFI. **Analogs are computed using the large scale field SLP of the region -80W-50E,20N-70N and criteria 1 of function Analogs within CSTools Rpackage**

# Downscaling using Analogs in Climate projections

Mean HWFI duration
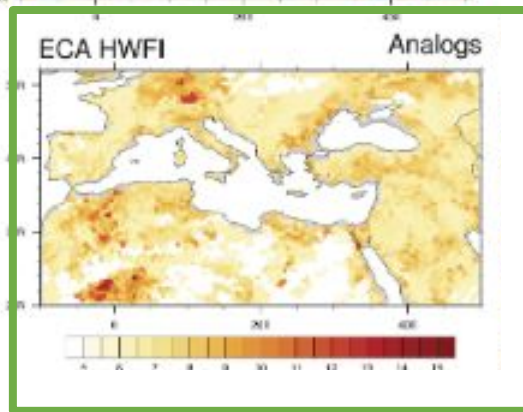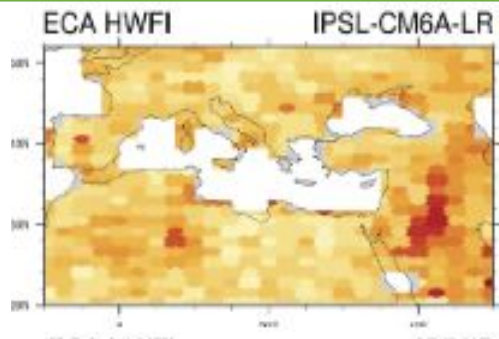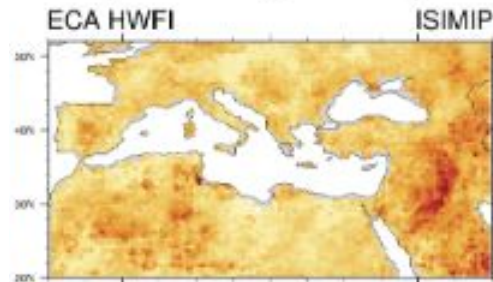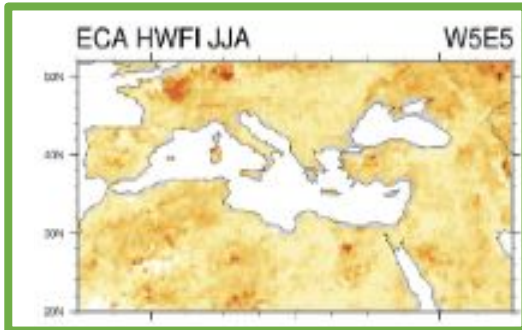
Applications of Analogs to Climate projections for a warm spell days index in JJA (HWFI)

Analogs are computed using the large scale field SLP of the region -80W-50E,20N-70N and criteria 1 of function Analogs within CSTools Rpackage

Using Analogs not only for downscaling but for…..

Attribution studies, Bias correction or as method of machine learning

Analogs are regions or period with similar climatic conditions to the future conditions of the selected location or period so we can perform Attribution studies or use the method as a method of machine learning and as a method to bias correct our data.
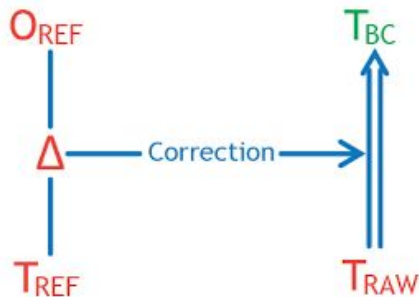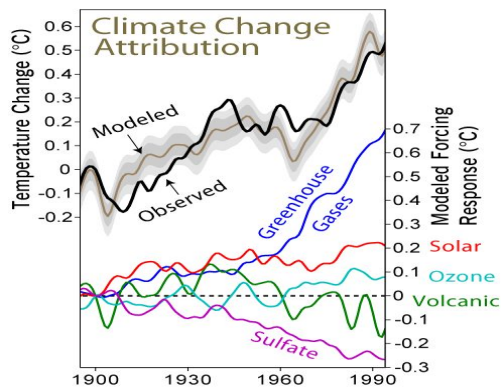
Yiou et al, 2017, Faranda et al 2020, Jezequel et al, 2017, Turco et al, 2017, Vrac et al, 2020

Searching for Analogs: Functions CST_Analogs, Analogs

Developing methodologies to extract usable information from predictions. We will produce tools for prediction verification**, calibration, downscaling, ensemble member combination and selection** that will be publicly released via a **toolbox** and shared among partners and users.

CSTools -> Rpackage   version 5.0.1. available in CRAN   Last update June 2023

Package 'CSTools'

June 6, 2023

Title Assessing Skill of Climate Forecasts on Seasonal-to-Decadal Timescales

Version 5.0.1

Description Exploits dynamical seasonal forecasts in order to provide information relevant to stakeholders at the seasonal timescale. The package contains process-based methods for forecast calibration, bias correction, statistical and stochastic downscaling, optimal forecast combination and multivariate verification, as well as basic and advanced tools to obtain tailored products. This package was developed in the context of the 'ERA4CS' project 'MEDSCOPE' and the 'H2020 S2S4E' project and includes contributions from 'ArticXchange' project founded by 'EU-PolarNet 2'. 'Pérez-Zanón et al. (2022) <doi:10.5194/gmd-15-6115-2022>'.

https://cran.r-project.org/package=CSTools

Analogs: Analogs based on large scale fields
CST_Analogs: Downscaling using Analogs based on large scale fields

https://cran.r-project.org/web/packages/CSTools/vignettes/Analogs_vignette.html

Pérez-Zanón et al 2022

Hands on Session

In this example, the seasonal temperature forecasts, initialized in october, will be used to **perform a downscaling in the Balearic Islands temperature using the ecmwf system 4 seasonal forecasting system from ECMWF**, by computing Analogs in Sea level pressure data (SLP) in a larger region (North Atlantic).

The first step will be to load the data we want to downscale (i.e.seasonal forecast) in the large region (i.e North Atlantic) for temperature (predictand) and SLP (predictor) and same variables and region for a higher resolution data (ERA5).

In a second step we will interpolate the model to the resolution of ERA5.

In a third step we will find the analogs using one of the three criterias.
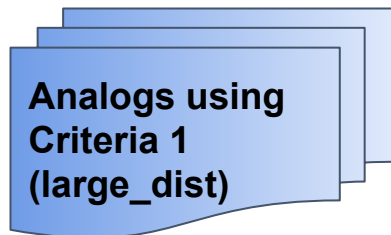
In a four step we will get the downscaled dataset in the region selected (local scale, in this case Balearic Islands).
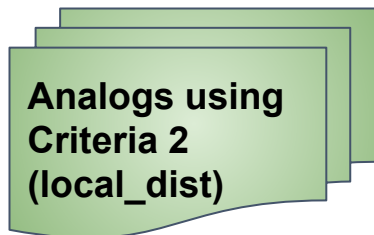
Introduction of the function

For instance if we want to perform a temperature donwscaling in Balearic Island for October we will get a daily series of temperature with 1 analog per day, the best analog.

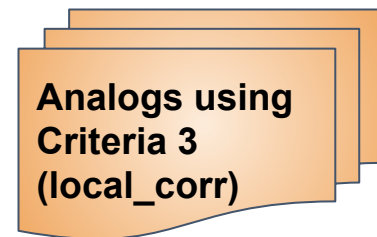**How we define the best analog for a certain day?**

This function offers three options for that:

Analogs using Criteria 1 (large_dist)

Analogs using Criteria 2 (local_dist)

Analogs using Criteria 3 (local_corr)

Hands on Session

Analogs using Criteria 1 (large_dist)

(1) The day with the minimum Euclidean distance in a large scale field: using i.e. pressure or geopotencial height as variables and North Atlantic region as large scale region. The Atmospheric circulation pattern in the North Atlantic (LargeScale) has an important role in the climate in Spain (LocalScale). The function will find the day with the most similar pattern in atmospheric circulation in the database (obs, slp in ERA5) to the day of interest (exp, slp in model). Once the date of the best analog is found, the function takes the associated temperature to that day (obsVar, tas in ERA5), with a subset of the region of interest (Balearic Island).

Analogs using Criteria 2 (local_dist)

(2) Same that (1) but in this case we will search for analogs in the local scale (Balearic Island) instead of in the large scale (North Atlantic). Once the date of the best analog is found, the function takes the associated temperature to that day (obsVar, t2m in ERA5), with a subset of the region of interest (Balearic Island).

Analogs using Criteria 3 (local_corr)

(3) Same that (2) but here we will search for analogs with higher correlation at local scale (Balearic Island) and instead of using SLP we will use t2m.
I

Hands on Session

- In particular the Analogs Method uses a nonlinear approach that follows the methodology of Yiou et al. 2013.

- An efficient implementation of Analogs is provided for CSTools by the **CST_Analogs() function.**

- **Two datasets** are used to illustrate how to use the function.

    - The first one could be entirely run by the users since it is using <u>data samples provided along with the package.</u>

    - The second one uses <u>data that needs to be downloaded or requested.</u>

Hands on: Setting the environment

load the last versions of R and the packages -> module load R/4.2.0

```
# To run this example several libraries are required
# install.packages('CSTools', 'zeallot', 's2dv','ClimProjDiags', etc)
# attention to the version of the packages

# load libraries:
library(CSTools) # version CSTools_5.0.1
library(s2dv) # version s2dv_1.4.1
library(ClimProjDiags) #version ClimProjDiags_0.1.3
library(zeallot) #version zeallot_0.1.0
library(fields) #version fields_13.3
library(multiApply) #version multiApply_2.1.4
library(abind) #version abind_1.4-5
library(s2dverification) # version s2dverification_2.10.3

# in order to check the versions of the packages
sessionInfo()
```

Hands on: Examples

```
/home/tr100/CST_Analogs
```

```
data_for_Analogs.rdat
MEDCOF_CST_Analogs_Alvarez-Castro.pdf
MEDCOF_examples_CST_Analogs.R
```

https://cran.r-project.org/web/packages/CSTools/vignettes/Analogs_vignette.html

For further information, discussion, comments, etc
carmen.alvarez-castro@cmcc.it or mcalvcas@upo.es

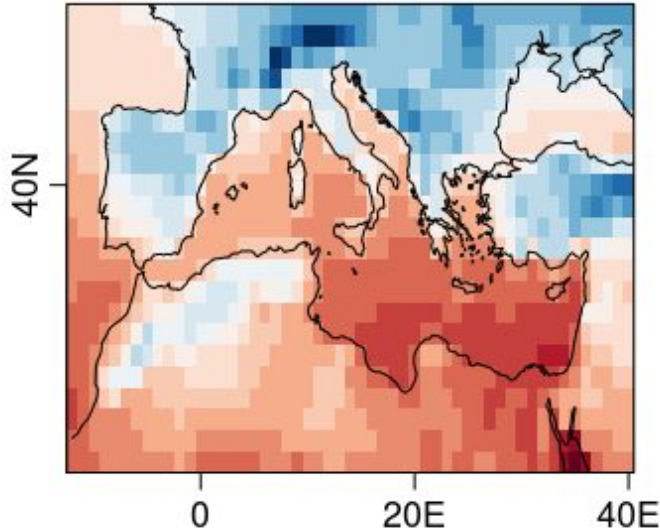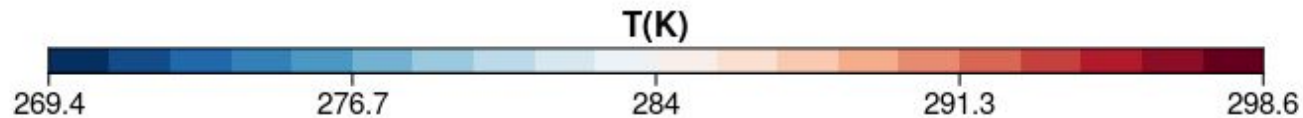# Example 1: using data from CSTools

Analogs sdate November 2000
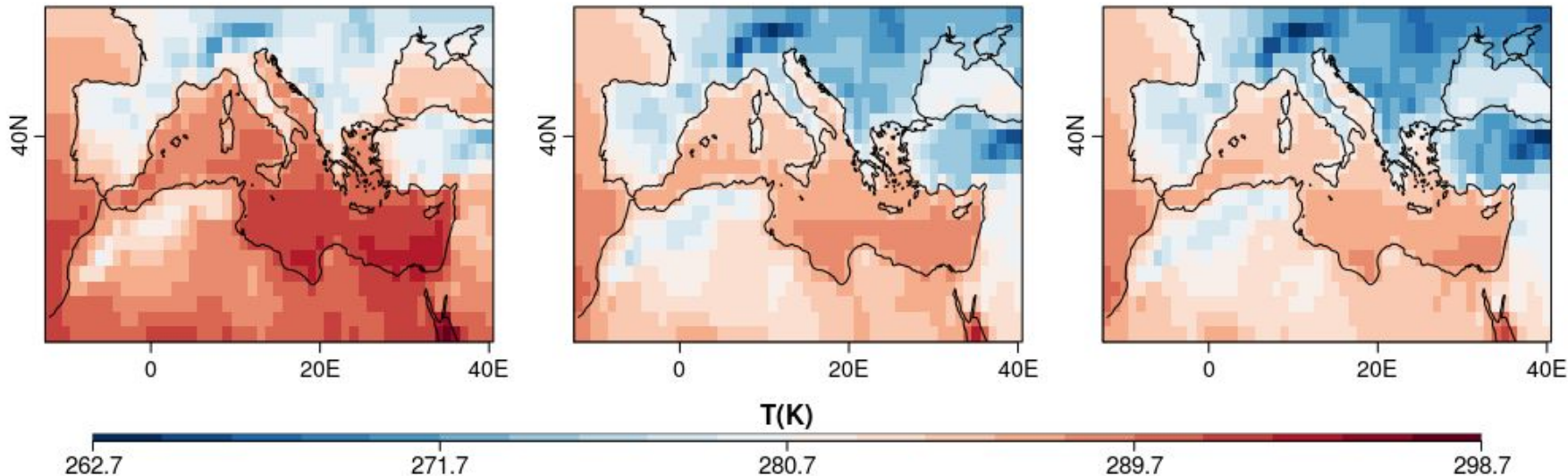
Example 1: using data from CSTools
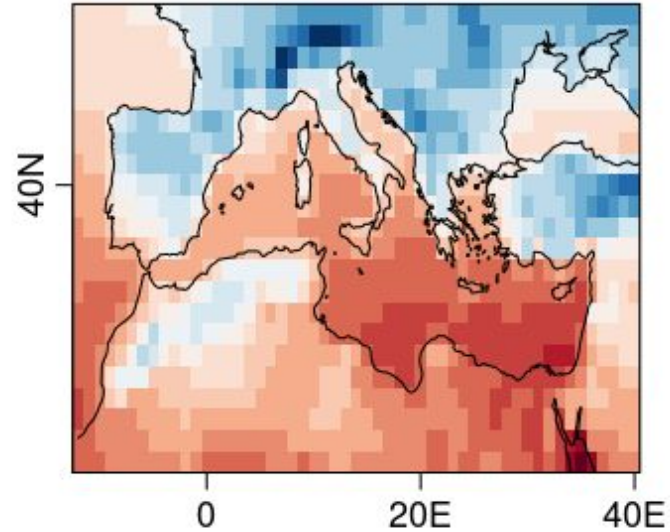
Example 2: Load data using CST_Load or loading Rdat previously downloaded: Criteria 1

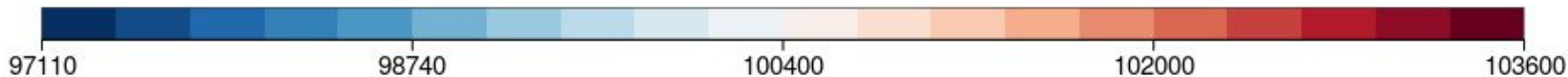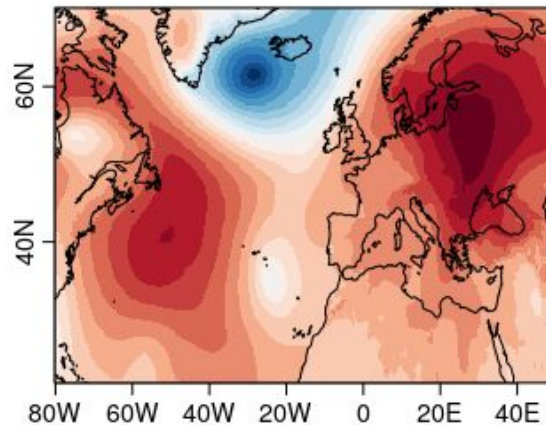Two variables and criteria Large [scale] Distance

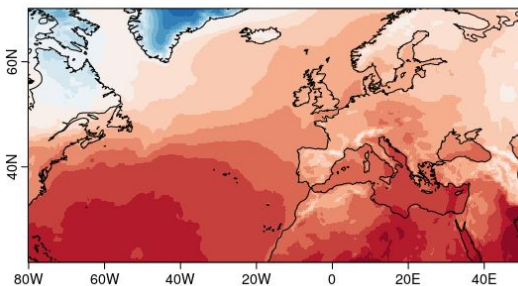First member
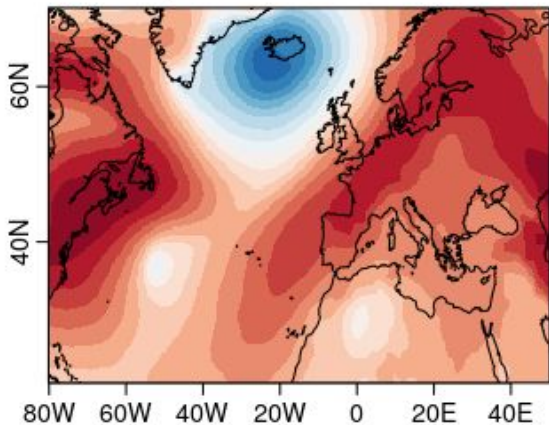
Exp PSL 15-10-2000   Obs PSL 15-10-2000   Obs PSL 7-10-2005
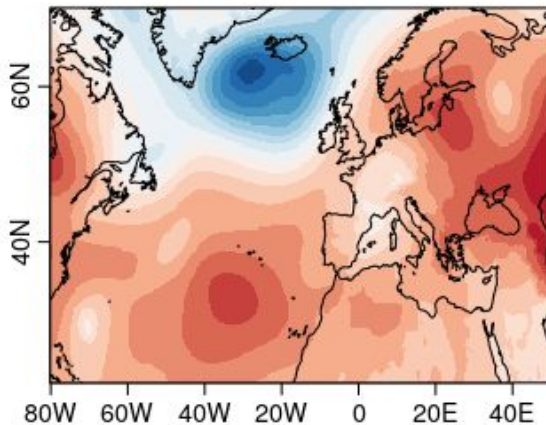
97110   98740   100400   102000   103600
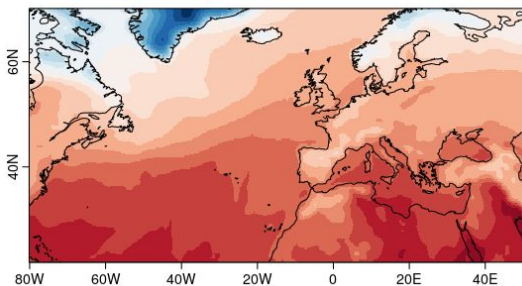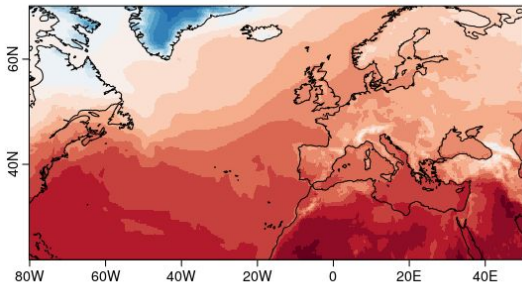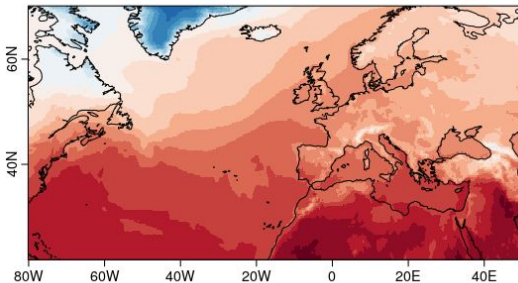
Example 2: Load data using CST_Load or loading Rdat previously downloaded: Criteria 1
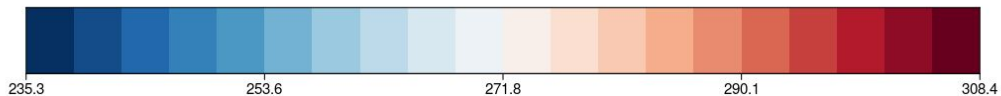
Two variables and criteria

Large [scale] Distance

Example 2: Load data using CST_Load or loading Rdat previously downloaded: Criteria 2
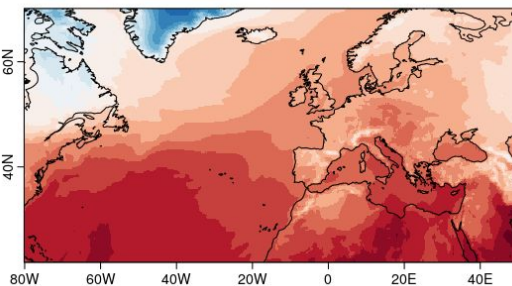
Two variables and criteria

Local [scale] Distance

Example 2: Load data using CST_Load or loading Rdat previously downloaded: Criteria 3

Two variables and criteria Local [scale] Correlation

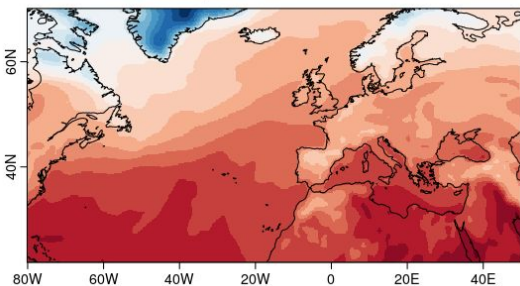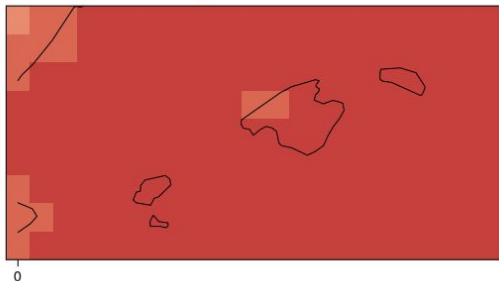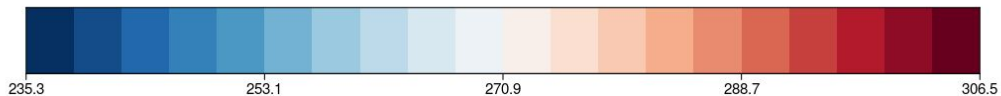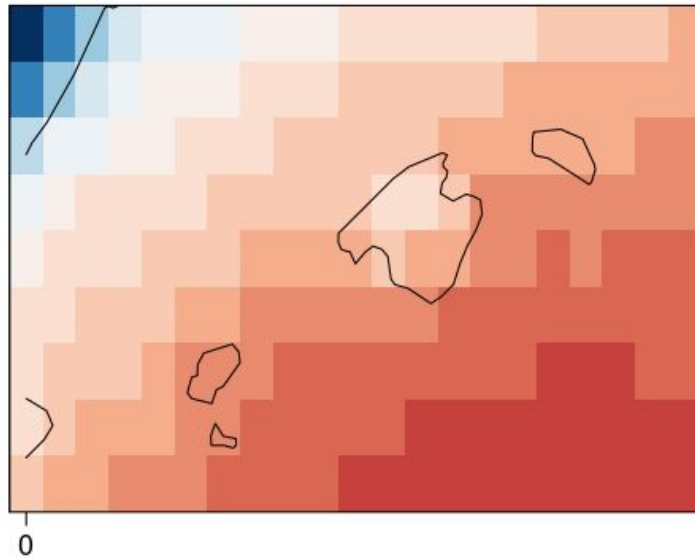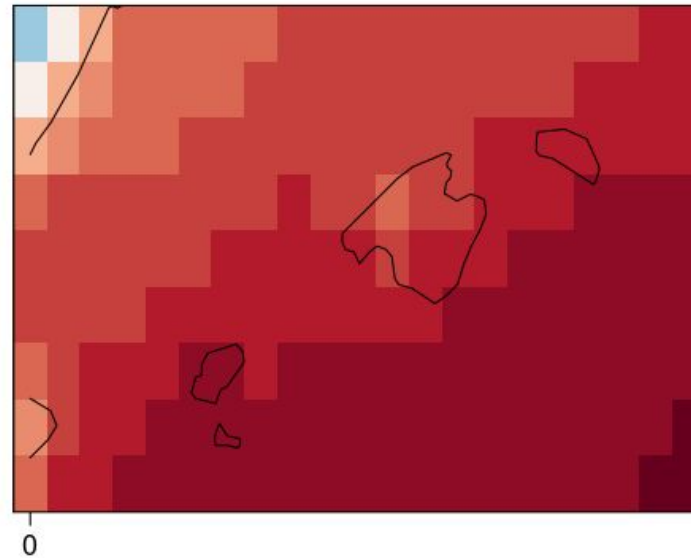# Example 3: Short examples with simple data

```r
# Example 1: Downscaling using criteria 'Large_dist' and a single variable:
expSLP <- rnorm(1:20)
dim(expSLP) <- c(lat = 4, lon = 5)
obsSLP <- c(rnorm(1:180), expSLP * 1.2)
dim(obsSLP) <- c(time = 10, lat = 4, lon = 5)
time_obsSLP <- paste(rep("01", 10), rep("01", 10), 1994 : 2003, sep = "-")
downscale_field <- Analogs(expL = expSLP, obsL = obsSLP,
                           time_obsL = time_obsSLP,time_expL = "01-01-1994")


# Example 2: Downscaling using criteria 'Large_dist' and 2 variables:
obs.pr <- c(rnorm(1:200) * 0.001)
dim(obs.pr) <- dim(obsSLP)
downscale_field <- Analogs(expL = expSLP, obsL = obsSLP, obsVar = obs.pr,
                           time_obsL = time_obsSLP, time_expL = "01-01-1994")

# Example 3: Downscaling using criteria 'Local_dist' and 2 variables:
# analogs of local scale using criteria 2
region = c(lonmin = -1 ,lonmax = 2, latmin = 30, latmax = 33)
Local_scale <- Analogs(expL = expSLP, obsL = obsSLP, time_obsL = time_obsSLP,
                       obsVar = obs.pr, criteria = "Local_dist",
                       lonL = seq(-1, 5, 1.5),latL = seq(30, 35, 1.5),
                       region = region,time_expL = "01-10-2000",
                       nAnalogs = 10, AnalogsInfo = TRUE)
```

```r
# Example 4: Downscaling using criteria 'Local_cor' and 2 variables:
exp.pr <- c(rnorm(1:20) * 0.001)
dim(exp.pr) <- dim(expSLP)
Local_scalecor <- Analogs(expL = expSLP, obsL = obsSLP, time_obsL = time_obsSLP,
                          obsVar = obs.pr, expVar = exp.pr,
                          criteria = "Local_cor", lonL = seq(-1, 5, 1.5),
                          time_expL = "01-10-2000", latL = seq(30, 35, 1.5),
                          lonVar = seq(-1, 5, 1.5), latVar = seq(30, 35, 1.5),
                          nAnalogs = 8, region = region, AnalogsInfo = FALSE)

# Example 5: List of best analogs in the three criterias Large_dist,
Large_scale <- Analogs(expL = expSLP, obsL = obsSLP, time_obsL = time_obsSLP,
                       criteria = "Large_dist", time_expL = "01-10-2000",
                       nAnalogs = 7, AnalogsInfo = TRUE)
Local_scale <- Analogs(expL = expSLP, obsL = obsSLP, time_obsL = time_obsSLP,
                       time_expL = "01-10-2000", criteria = "Local_dist",
                       lonL = seq(-1, 5, 1.5), latL = seq(30, 35, 1.5),
                       nAnalogs = 7, region = region, AnalogsInfo = TRUE)
Local_scalecor <- Analogs(expL = expSLP, obsL = obsSLP, time_obsL = time_obsSLP,
                          obsVar = obsSLP, expVar = expSLP,
                          time_expL = "01-10-2000", criteria = "Local_cor",
                          lonL = seq(-1, 5, 1.5), latL = seq(30, 35, 1.5),
                          lonVar = seq(-1, 5, 1.5), latVar = seq(30, 35, 1.5),
                          nAnalogs = 7, region = region,
                          AnalogsInfo = TRUE)
```

Mediterranean Seasonal Climate Update

This is the Seasonal Climate Update for for a large region encompassing the Mediterranean basin (MSCU), which presents forecasts of Precipitation and Temperature, for the upcoming season, using the North American Multi-Model Ensemble (NMME) seasonal prediction system (Kirtman et al, 2014, Becker et al, 2014), the Copernicus Climate Change (C3S) seasonal prediction system (https://climate.copernicus.eu/seasonalforecasts) and the AEMET empirical model developed within the MEDSCOPE framework (Rodriguez-Guisado et al, 2019).

The MSCU is principally addressed to the MedCOF community and more generally, to users and stakeholders who may benefit from having at hands a summary, in graphical homogeneous form, of the seasonal predictions produced by some of the main international and well-documented multi-model systems.

https://www.medscope-project.eu/products/mediterranean-seasonal-climate-update/