

Seasonal forecast of extreme events

Hands on session

Objectives

- ★ Basics of R language
- ★ Learn how to use s2dverification

1 Load files, compute basic statistics, plot map and time series

Exercise 1 – First steps with R and s2dverification

- Open R (typing R in the terminal).
- Load the needed library, with the following command:

```
library(s2dverification)
```

- The documentation of s2dverification is available online here:
http://ic3.cat/wikicfu/img_auth.php/S2dverification.pdf
- You can see the list of available functions in the package by typing:

```
help(package=s2dverification)
```

- To see the help of a specific function, you can type:

```
help(Load)
```

- Here I give you an example of how to use of Load function:

```
library(s2dverification)
Data=Load("tas",
  c("EnsCmccSeas","EnsIfmSeas","EnsEcmwfSeas","EnsMetfrSeas","EnsUkmoSeas"),
  obs = "ERAint",sdates=c("19930501","19940501"), nleadtime = 7, leadtimemin = 1,
  leadtimemax = 7, storefreq = "monthly", sampleperiod = 1, nmember=9,
  output = "areave", lonmin = 190, lonmax = 240, latmin = -5,latmax = 5)
```

- You can copy paste it in R, run and check everything is working fine. Once it is finished have a look at the dimension of *Data* by typing this command in R:

```
dim(Data$mod)
dim(Data$obs)
```

- With the help of the Load function try to understand what the different dimensions are ?
- Now run another time Load with the same parameters, but change *output = "areave"* to *output = "lonlat"*.
- What are the dimension of *Data* now ?
- What happens if your remove the *lonmin*, *lonmax*, *latmin*, *latmax* options ?
- What happens if your change the *leadtimemin*? and *leadtimemax* option?

Exercise 2 – First program plot the 2m-temperature skill over Europe

- In the directory named *handson* create a directory named *R* and go to this directory.
- Open a new file named *corrskill-europe.R*.
- At the beginning of the file, load the *s2dverification* package:

```
library(s2dverification)
```

- If you are using emacs and you don't have color highlighting just execute this line in the terminal and reopen the file.

```
/usr/local/bin/ictp-install ess
```

- Generate a list of start dates to be used in the *sdates* argument of *Load* which includes all May startdate between 1979 and 2005 (you can have a look at the help of the function *seq* and *paste* to do it).
- Based on the example of the previous exercise, try to Load the leadtimes 2, 3 and 4 of 2m-temperature (*tas*) of all ENSEMBLES models, over Europe (20W70E-25N75N) for all May startdates between 1979 and 2005.
- Calculate the June-July-August seasonal mean using the *Mean1Dim* function for both models and observations.
- With the same function, calculate the ensemble mean of each model.
- With the *Corr* function, calculate the time correlation in June-July-August between each model and the ERA-interim reanalysis.
- Using the *PlotEquiMap* function, plot a map of the correlation coefficient for each model and save it in a postscript (You can use the *postscript* and *dev.off* functions). You can choose the following interval and the following color bar:

```
min=-1
max=1
int=(max-min)/20
interval=seq(min,max,int)

color=c("blue4", "blue3", "blue", "dodgerblue3", "dodgerblue2",
        "dodgerblue1", "steelblue1", "cadetblue2", "cadetblue1",
        "white", "white", "gold", "goldenrod", "chocolate", "orangered", "firebrick1",
        "firebrick3", "firebrick", "firebrick4", "red4")
```

or generate another one using the *brewer.pal* function.

- If you want to generate a multipanel plot you can use the *layout* function (you can have a look at this document for more help: <http://seananderson.ca/courses/11-multipanel/multipanel.pdf>). Be careful, some feature of *PlotEquiMap* are not available with multipanel plots.

Exercise 3 – Compare the skill over over land the mediterranean region

- In the directory named *handson/R*, open a new file named *Timeserie-skill-mediter.R*.
- The mediteranean region 3E25E-36N44N cover both sea and land, but we would like to calculate the skill only over land. For this, you need to use a land-sea mask. The land-sea masks of ENSEMBLES and ERA interim are here:

```
ls /afs/ictp.it/public/c/cprodhom/land_sea_mask_ERAIN.T.nc
ls /afs/ictp.it/public/c/cprodhom/land_sea_mask_ENSEMBLES.nc
```

- To create masks usable in the *maskobs* and *maskmod* arguments of the *Load* function, you can use the following lines:

```
fnc=open.ncdf ("/afs/ictp.it/public/c/cprodhom/land_sea_mask ERAINT.nc")
maskERA1=get.var.ncdf (fnc, "LSM")
close.ncdf (fnc)
maskERA1[which(is.na(maskERA1))]=1

fnc=open.ncdf ("/afs/ictp.it/public/c/cprodhom/land_sea_mask ENSEMBLES.nc")
maskmod=get.var.ncdf (fnc, "LSM")
close.ncdf (fnc)
maskmod[maskmod > 0.5]=1
maskmod[maskmod <= 0.5]=0

listmaskmod=list (maskmod,maskmod,maskmod,maskmod,maskmod)
```

- Use the `Load` function with `maskobs` and `maskmod` arguments to Load the data averaged over the land in the mediterranean region for all ENSEMBLES models and ERAint reanalysis, for all may start dates between 1979 and 2005.
- With the `Clim` and `Ano` functions of `s2dverification`, calculate the anomalies of both models and observation.
- Calculate the ensemble mean of the models.
- Calculate the Root Mean Square Skill Score (RMSSS) for all leadtimes, with the `RMSSS` function of `s2dverification`.
- Use the `PlotVsLTime` function to plot this score and save it in a postscript. `PlotVsLTime` expects a third dimension of the score of size 4 (lower confidence interval, score, upper confidence interval and significance level), while with `RMSSS` you will have only 2 values (score and pvalue). To do it you can use the following commands to resize your score matrix:

```
skillreshape=array(dim=c(5,1,4,7))
skillreshape[,2,]=skill[,1,]
skillreshape[,4,]=skill[,2,]
```

- Do the same plot, but instead of plotting a line for the p-values, mark with a dot the significant values (p-values under 0.05). To do this, you can have a look at the tutorial: <http://www.harding.edu/fmccown/r/>.

Exercise 4 – Load the extreme variables with R

- When the data of exercises 3 and 4 are calculated, try to use the 2 procedures you have written in exercise 6 and 7 to calculate the skill over Europe and mediterranean region for extreme variables.