Assignment: Comparing climate distribution data

Your task is to write a short report (no more than a few sides including plots) to show how a climate variable is projected to change, according to the given model output, in Durban. The focus is on comparing distribution data – the topic of the lecture (see Vula). Please submit your assignments on Vula by May 30th.

Use the data and R files available at http://www.csag.uct.ac.za/~jdaron/EGS4023.

Step 1

Open R and change the directory to your current work directory by selecting file/change dir...

(NB: In the R console, type "dir()" to see the files in your current working directory.)

Step 2

Decide on the variable you would like to analyse. Choose either monthly precipitation, mean monthly maximum temperature or mean monthly minimum temperature.

For the chosen variable download the observed data, control model data (1979 to 1999) and one file of projected data: for the projected data choose either emissions scenario A2 or B1 and either the period 2046 to 2065 or 2081 to 2100.

(NB: make sure you save the .txt files in your current working directory)

Step 3

Import the data into the workspace. Use the file "analysis.R" which is set up to import observed mean monthly maximum temperatures. Change the necessary labels so it is appropriate for your selected variable and data (e.g. if you chose precip, replace everything that says "maxT"). Once you have altered the R script, save it to your work directory and type in the command line:

> source("analysis.R") ... OR ... > source("file_I_have_just_renamed.R")

When importing the model data, notice that there are 10 different models in the output data files. Therefore each column needs to be assigned a unique label (e.g. "precip_model1") so when using the file "analysis.R", change parts of the first two lines so that they read something like:

```
model_maxT_durban<-read.table(file="name_of_data_file.txt", col.names=c("year", "month", "maxT_model1", "maxT_model2", ...."maxT_model10"))

attach(model_maxT_durban)
```

Then you need to ensure that the last 12 lines of code are repeated 10 times (copy and paste saves a lot of time, as does find and replace (CTRL+H)) but change the labels appropriately so they read something like:

```
jan_maxT_model1<-c(maxT_model1[jan_seq])

feb_maxT_model1<-c(maxT_model1[feb_seq])

mar_maxT_model1<-c(maxT_model1[mar_seq])

...

jan_maxT_model2<-c(maxT_model2[jan_seq])
```

```
feb_maxT_model2<-c(maxT_model2[feb_seq])

mar_maxT_model2<-c(maxT_model2[mar_seq])

...
```

Step 4 - Observations

Plot the **observed** data as a series of 12 boxplots (one for each month) in the same plot area. The line of code you need to do this can be found in the file "others.R". Save the figure (File/save_as...) into an appropriate filetype. This will be the first figure to include in the report. Give a brief description of what the plot shows.

Step 5 – Control period

Choose a month to focus on. If you have chosen March, for example, then compare the March control model data (1979 to 1999) to the observed data (1979 to 1999) using the Kolmogorov-Smirnov test and the Student-T test (commands given in "others.R"). Remember you have 10 models (only 9 models if you selected B1 2081 to 2100), so you should get 10 values of the KS statistic and 10 values of the t statistic. In the report, produce a table of the 10 KS values and 10 t values and discuss the results: are the models close to the observed data? Comment on which test is more informative for the data you have, and why?

Step 6 – Future climate

So far you will have chosen a variable (e.g. precip), a month (e.g March), a scenario (e.g. A2) and a future period (e.g. 2046 to 2065). Now select the three "best models"; this will be guided by Step 5. Produce boxplots which compare the observed *March precip* (for example) with the three control model runs of *March precip* and the three future projections of *March precip*. Do these three models indicate any significant change in the variable? Discuss any limitations in drawing conclusions based on this analysis and comment on the value of using more than one model.

Final step

Compile everything you have into a short report and submit on Vula.

Things to remember

Avoid spaces in R commands. It is best to label data in a format using the underscore: e.g. "my_climate_data". (this is a general rule for any programming!)

Be very careful with brackets, colons, commas etc in the R scripts...if a command doesn't work usually it is something simple.

Don't worry if you get the message: "The following object(s) are masked from 'obs_maxT_durban (position 3)': month, year...". I have no idea what this means but it doesn't influence your ability to make plots!

The task shouldn't take much more than two to three hours so if you find yourself spending a lot of time on it then come and find me...you may be missing some shortcuts.

Finally, if you have problems or queries then ask your peers or come and find me in room 6.03. You can always email me at jdaron@csag.uct.ac.uk. Have fun!