iEMSs Conference 2020
International Environmental Modelling and Software Society

Rainfall forecasting over a Mediterranean catchment using Empirical Statistical Downscaling techniques

Khouloud Gader, AhlemGara, Marnik Vanclooster, Slaheddine Khlifi, Mohamed Slimani

Email: khouloud.gader@yahoo.fr

INTRODUCTION

Context:

Tunisia as a Mediterranean region ➞ ✓Climate variability
✓Tunisia a country particularly vulnerable to Climate Change

Climate change

Climate modeling

Scenario of green gas emission

Global Climate Model

Regional Climate Model

Why Downscaling?

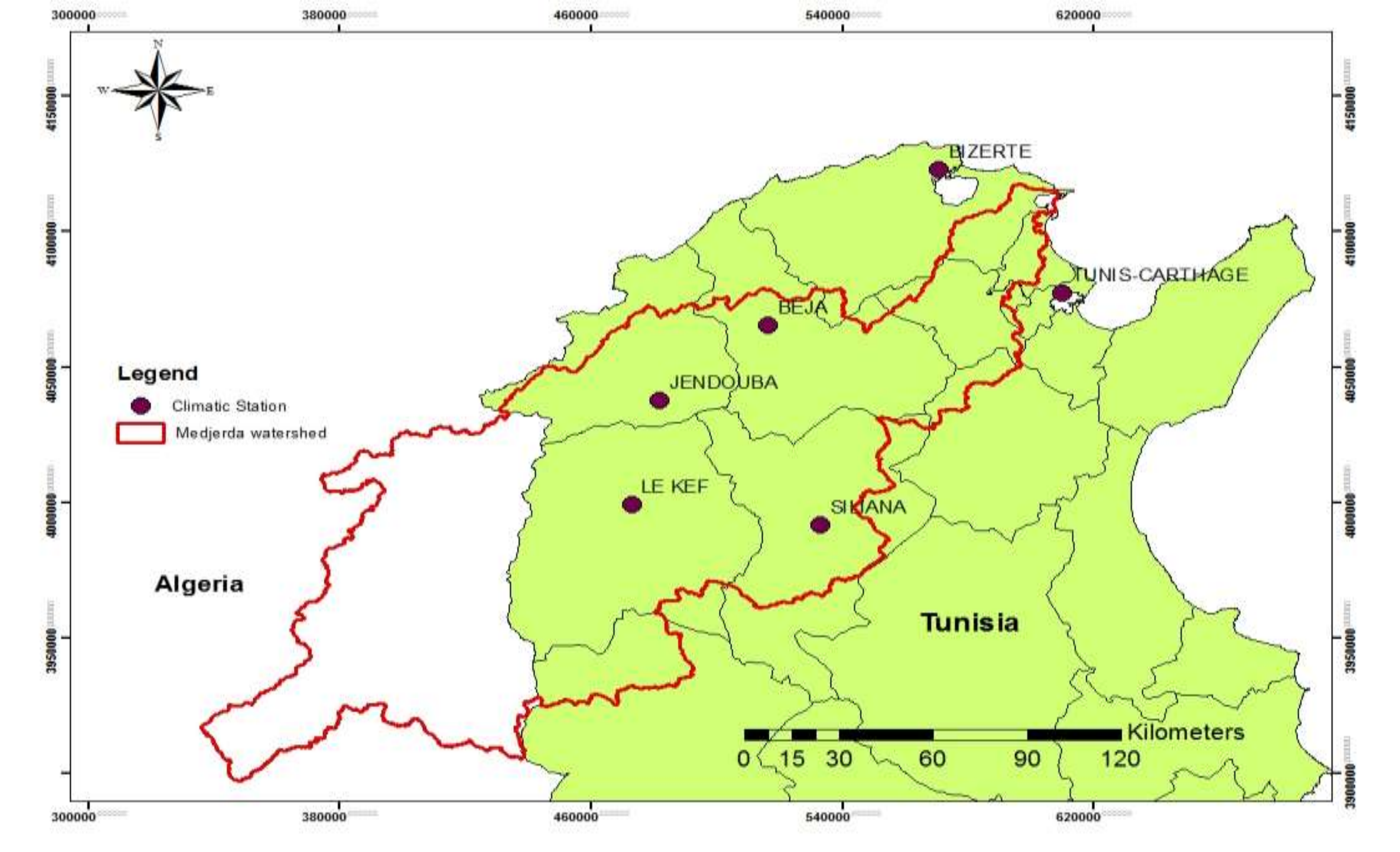
Objectives

This study consists:

To apply the Empirical Statistical Downscaling method to obtain precipitation projections for NE Tunisia by 2100 from the RCP 4.5 and RCP 8.5 scenario of GCM MIROC5

METHODOLOGY

Study area



Location of the chosen climatic stations in the Medjerda catchment

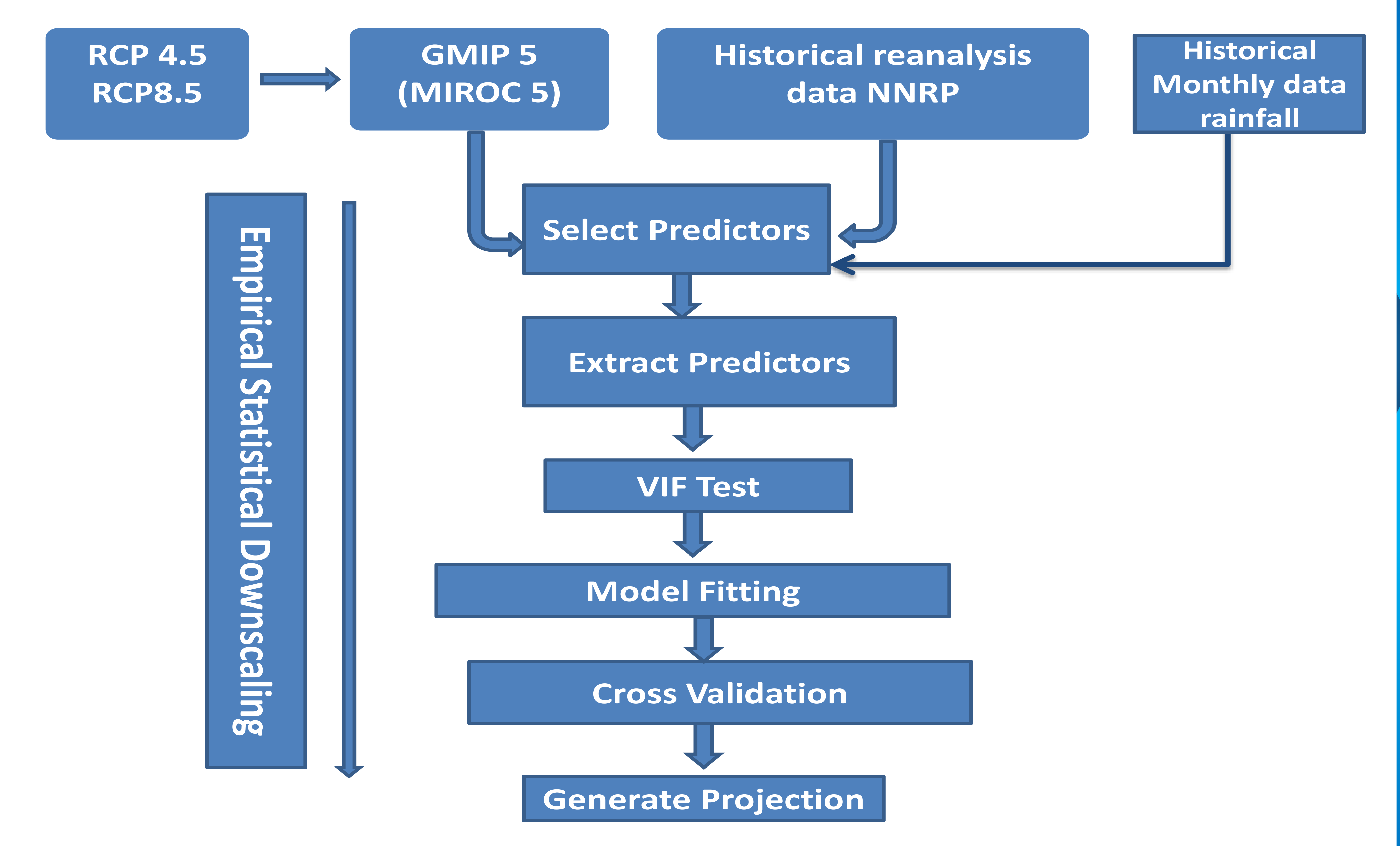
NCDC station IDs

Station	Id	Latitude	Longitude
Tunis carthage	607150	10.233	36.833
Beja	607230	9.183	36.733
Jendouba	607250	8.8	36.483
Siliana	607340	9.367	36.067
Bizete	607140	9.8	37.25
Le Kef	607320	8.700	36.133

Climate modeling

What is ESD?

Empirical Statistical Downscaling (ESD) is a method to predict local climate variables, based on the relationship between the local and large-scale climate variables



Flowchart of the ESD

Data

- Reanalysis data from the NCEP/NCAR Reanalysis Project

- Historical and RCP emissions data (RCP 4.5, RCP8.5)

- CMIP5 output from MIROC5 for historical and future conditions

- Historical monthly rainfall data during 30 years, from 1981 until 2010. these data are obtained from the national climate data center NCDC and then the National Institute of Meteorology "INM" database was used to fill the missing values.

Statistical characteristics of the selected precipitation series

	Siliana	Tunis Carthage	Beja	Bizerte	Jendouba	Le kef
Min	0	0	0	0	0	0
Max	172	218	246	308	207	235
1st Quartile	12	9	13	10	13	12
Median	26	29	36	38	29	27
3rd Quartile	48	53	71	78	56	52
Average	36	38	47	52	39	36
Variance	1044	1491	2015	2774	1288	1170
standard deviation	32	39	45	53	36	34
coefficient of variance	0,91	1,02	0,96	1	0,92	0,94

P-Value of the Mann-Kendall Trend Test

	Jan	Feb	Mar	April	May	June	July	Aug	Sept	Oct	Nov	Dec
Tunis Carthage	0.8	0,5	0,1	0,9	0,8	0,4	0,7	0,8	0,0	0,9	0.9	0,6
Beja	0,5	0,5	0,9	0,6	0,9	0,9	0,4	0,3	0,0	0,8	0,8	0,1
Jendouba	0,8	0,7	0,9	0,1	0,4	0,7	0,5	0,1	0,4	0,7	0,7	0,8
Siliana	0,1	0,5	0,8	0,0	0,1	0,2	0,4	0,8	0,7	0,6	0,5	0,4
Bizerte	0,3	0,5	0,8	0,1	0,6	0,6	0,9	0,9	0,0	0,5	0,2	0,7
Le Kef	0,9	0,3	0,5	0,1	0,0	0,2	0,2	0,0	0,0	0,5	0,9	0,4

Tools used during the climate modelling process

GCM Model	• **MIROC 5**
RCM Model	• **GLM, SGLM, GAM, SGAM,RF, ANN**
Predictor	• Sea level pressure (**slp**) • near-surface air temperature (**air**) • atmosphere temperature at level 500 hpa (**air 500**) • geopotential height at level 500 hpa (**hgt 500**) • near-surface relative humidity (**rhum**) • precipitation predictors (**prate**)
Criteria for model fitting	**COR, MAD, MAE, RMSE, MSE, Cross Validation**

To choose the appropriate model, we used a machine learning algorithm to test six different types of statistical models:

Linear regression	Random forest regression	neural network
• Generalized Linear Models (**GLM**), • Selected GLM (**SGLM**), • Generalized Additive Models (**GAM**), • Selected GAM (**SGAM**),	• Random forest (**RF**)	• artificial neural networks (**ANN**)

All models were compared with each other and with three reference models:

- ✓ average model (mean predictions from models),
- ✓ null/climatology model (mean rainfall),
- ✓ memory/persistence model (last year's rainfall).

Downscaling process

1. Analysis of correlation patterns based on the general circulation patterns most influencing the regional climate
2. Determination of areas of significant high correlation
3. Extraction of data from NCEP reanalyses and MIROC5 GCM outputs.
4. Extraction of the Predictors table from regression models
5. Objective selection using the variance inflation factor (VIF). This is an approach that identifies the collinearity between predictors
6. Model fitting
7. Cross validation: last step of validation of the choice of the best model
8. Projection

RESULTS AND DISCUSSION

In this research study, the ESD approach was used to generate regression-based rainfall projections for the horizon 2100 at the level of six main climate stations in the Medjerda catchment in Tunisia for the two RCP emissions scenarios (RCP4.5 and RCP 8.5)

so we have performed the downscaling process 12 times.

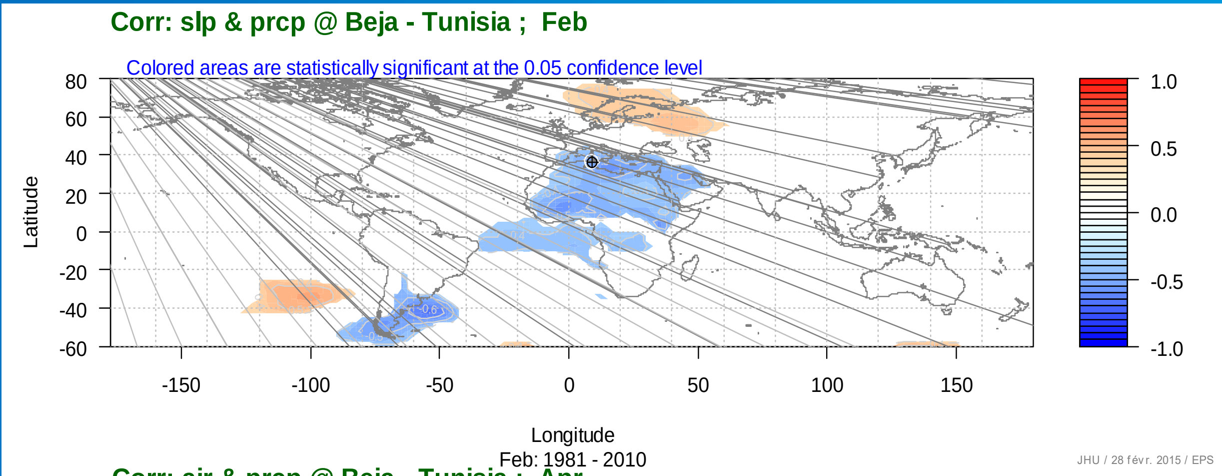
ESD approaches applied on a test station (case of precipitation at the Beja station)

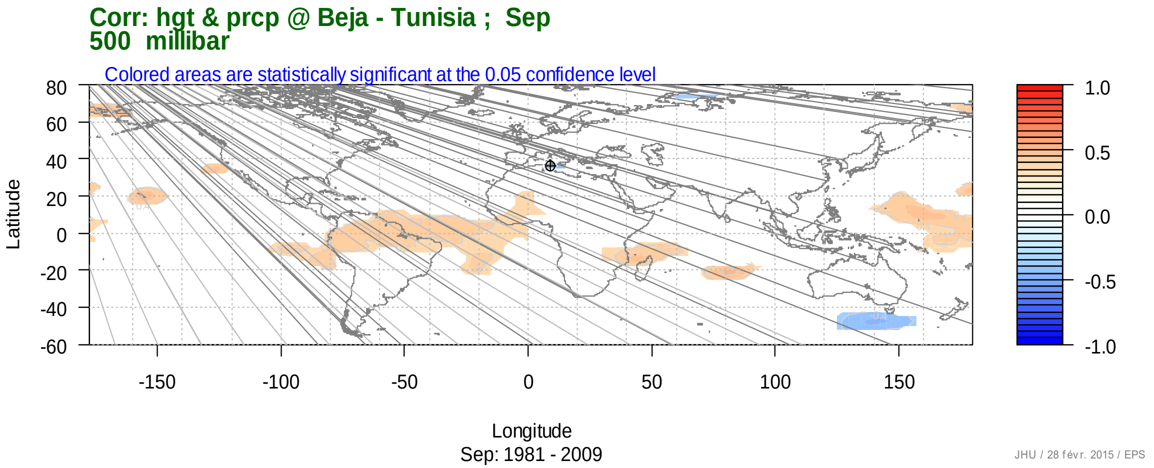
Precipitation mean data for a 30-year period from 1981 to 2010 are converted to R format by entering the station location data.

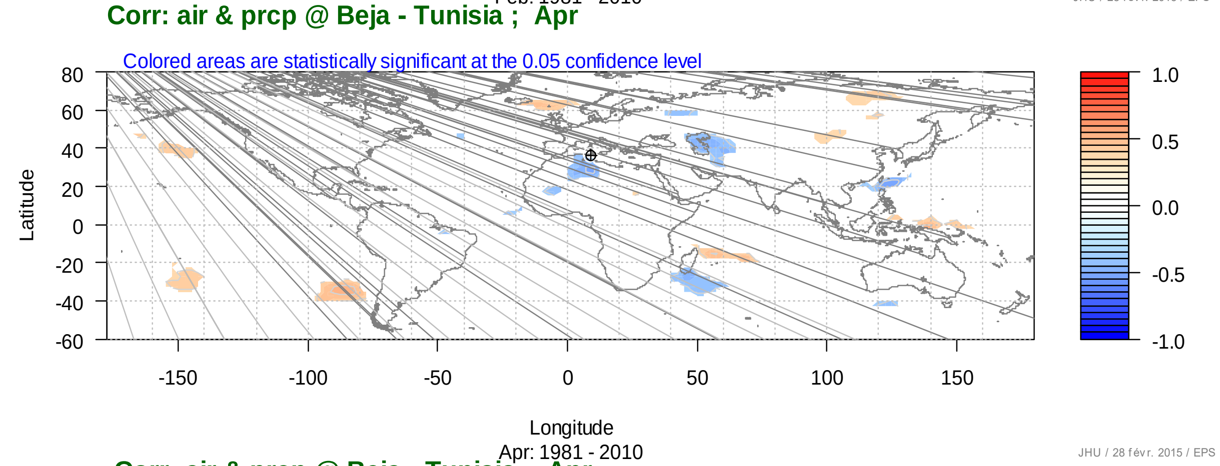
Station	NCDC Id	Latitude	Longitude	Altitude
Beja	607230	9.183	36.733	159 m

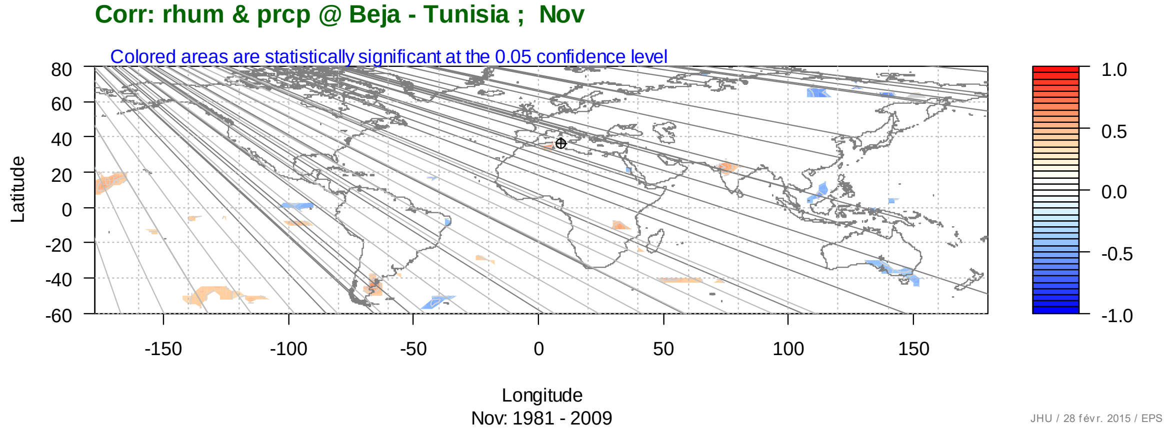
Extract predictors

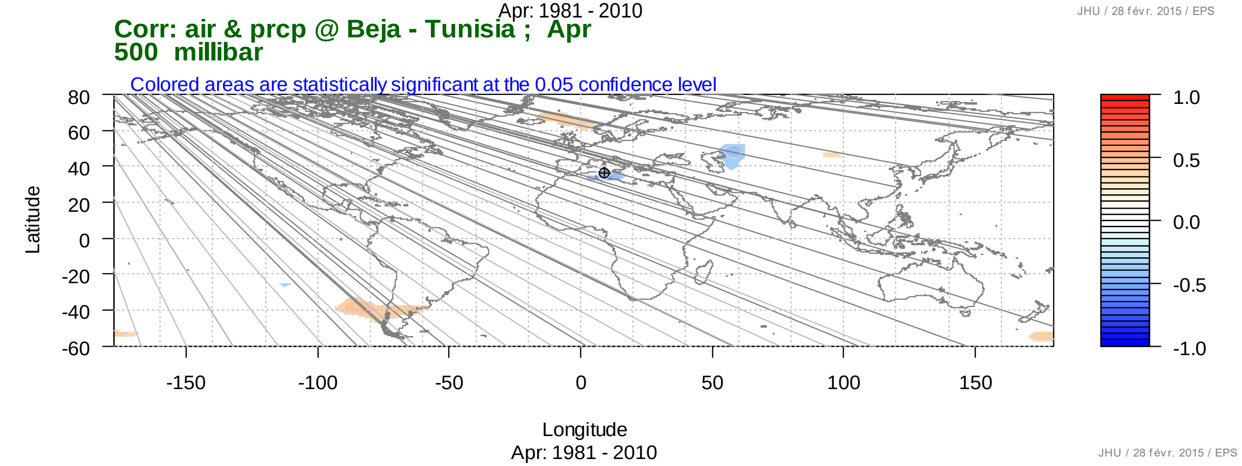
- The principle of this approach is that GCMs are better at capturing the large-scale variability of pressure, temperature, humidity, ...
- We used historical reanalysis data to identify regions of correlation and extracted time series of mean data from the reanalysis data set and from the GCM outputs.
- Predictor extraction is based on **correlation maps** that allow visualizing correlations between large-scale variables and local precipitation (we have 12 maps for each predictor that represents each month).
- Based on these correlation maps, we have defined the areas of high correlation of the **large-scale variables(Predictors)** that have physical significance for use in projections.

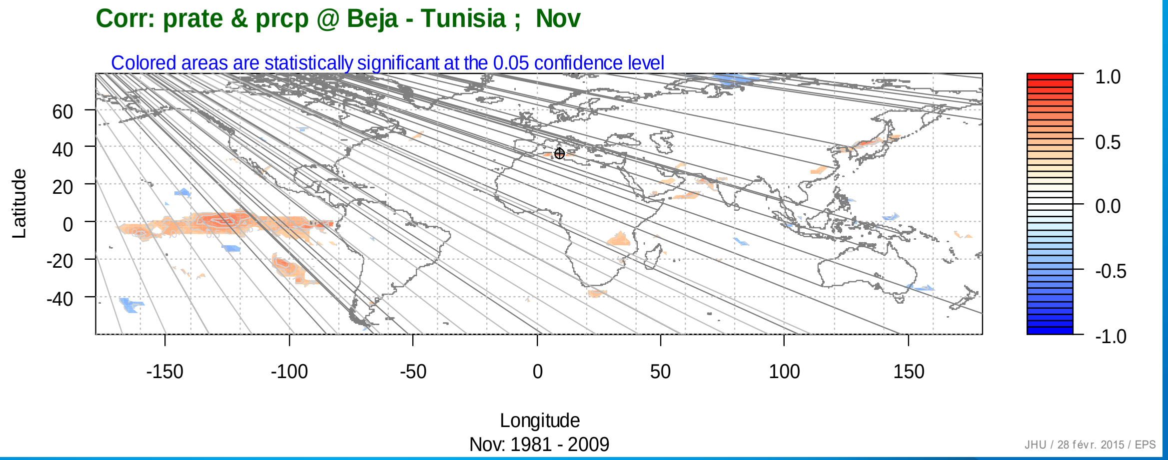

Corr: slp & prcp @ Beja - Tunisia ; Feb
Colored areas are statistically significant at the 0.05 confidence level
Latitude
Longitude
Feb: 1981 - 2010
JHU / 28 févr. 2015 / EPS


Corr: hgt & prcp @ Beja - Tunisia ; Sep
500 millibar
Colored areas are statistically significant at the 0.05 confidence level
Latitude
Longitude
Sep: 1981 - 2009
JHU / 28 févr. 2015 / EPS


Corr: air & prcp @ Beja - Tunisia ; Apr
Colored areas are statistically significant at the 0.05 confidence level
Latitude
Longitude
Apr: 1981 - 2010
JHU / 28 févr. 2015 / EPS

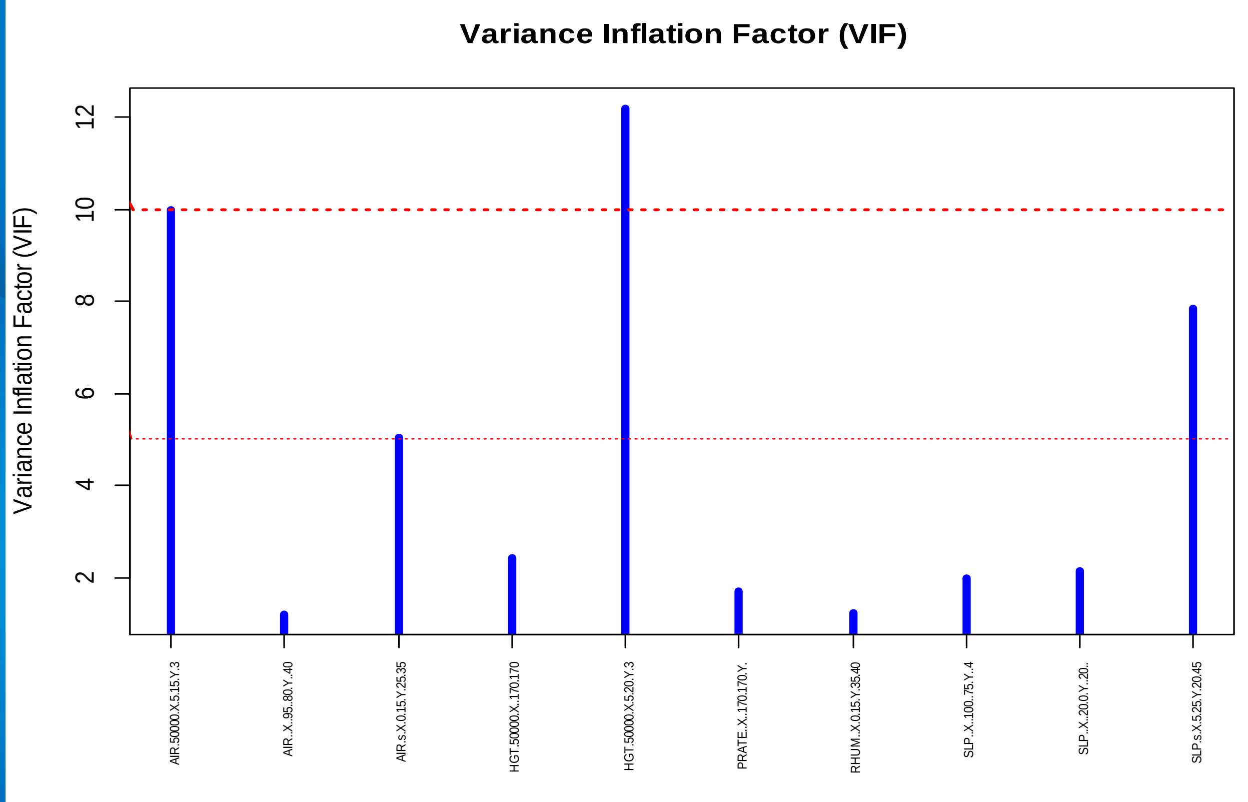

Corr: rhum & prcp @ Beja - Tunisia ; Nov
Colored areas are statistically significant at the 0.05 confidence level
Latitude
Longitude
Nov: 1981 - 2009
JHU / 28 févr. 2015 / EPS

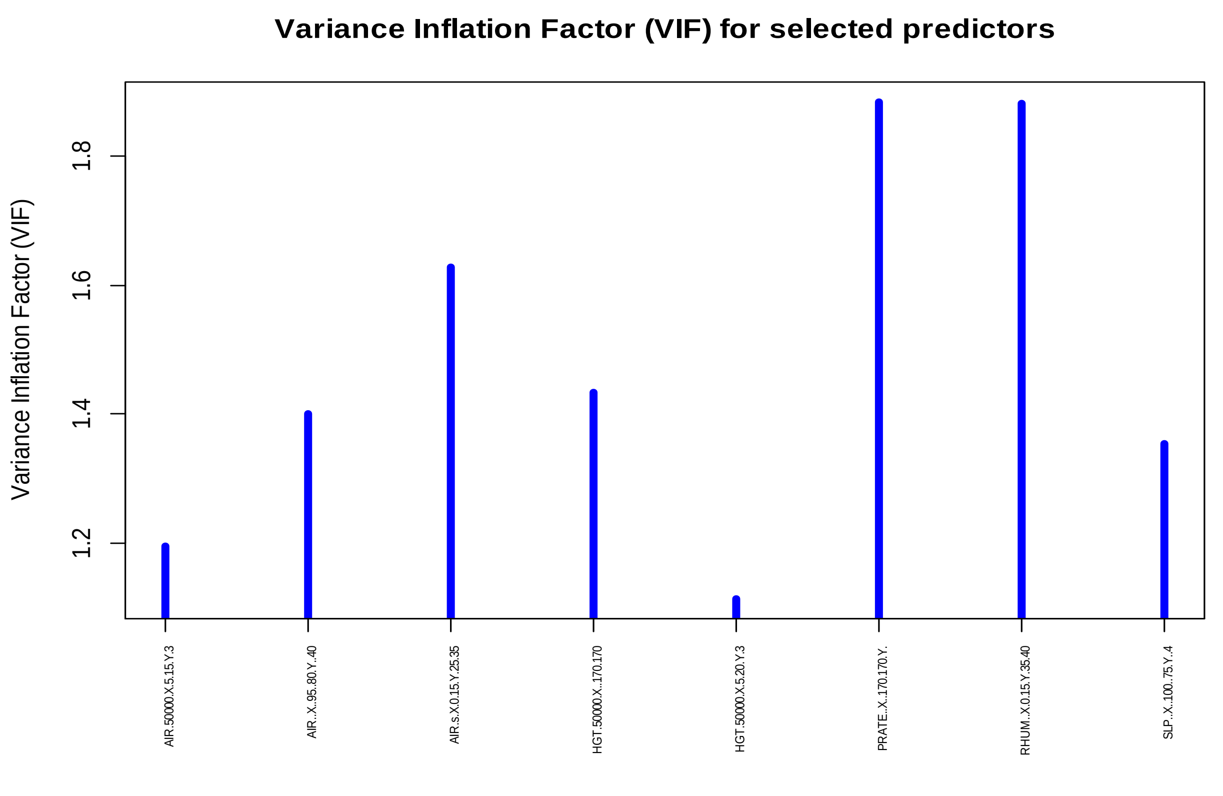

Corr: air & prcp @ Beja - Tunisia ; Apr
500 millibar
Colored areas are statistically significant at the 0.05 confidence level
Latitude
Longitude
Apr: 1981 - 2010
JHU / 28 févr. 2015 / EPS


Corr: prate & prcp @ Beja - Tunisia ; Nov
Colored areas are statistically significant at the 0.05 confidence level
Latitude
Longitude
Nov: 1981 - 2009
JHU / 28 févr. 2015 / EPS

The next step is to adjust the models that allows studying the significance of the predictors and the consistence of the developed models.

In order to avoid the problem of collinearity between predictors, we used an objective prediction selection algorithm “VIF test” that eliminates the variables responsible for high VIF values.





Model adjustment

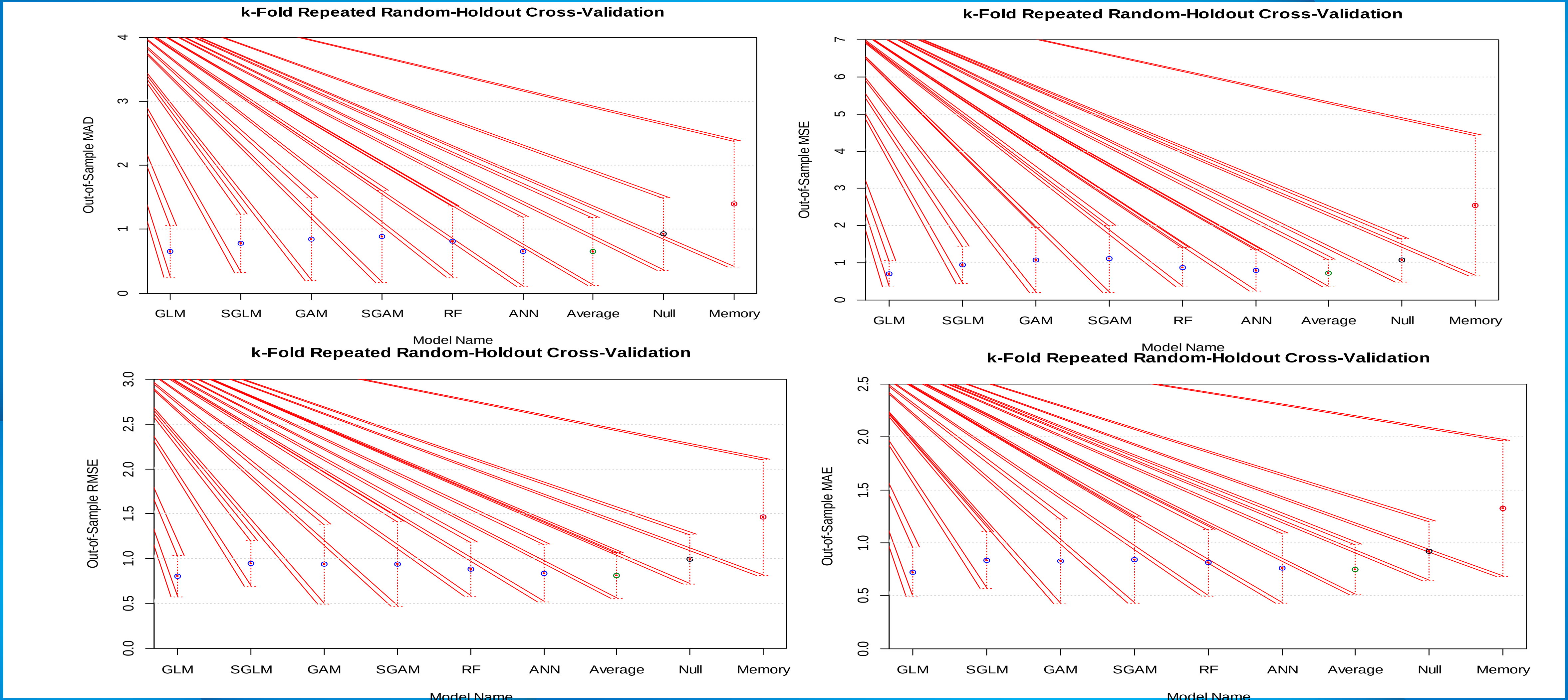
Model fitting

	COR	MAE	RMSE	MSE	MAD
GLM	**0.87**	**0.34**	**0.5**	**0.25**	**0.45**
SGLM	0.7	0.58	0.7	0.49	0.77
GAM	0.86	0.41	0.5	0.25	0.62
SGAM	0.78	0.50	0.6	0.35	0.61
RF	0.69	0.65	0.77	0.59	0.93
ANN	0.51	0.89	0.98	0.96	1.07
Average	0.84	0.51	0.6	0.36	0.73
Null	NA	0.87	0.98	0.96	1.07

Better models have lower error measures and higher correlation coefficients

Predictive skill

➢The statistical significance of differences between models was calculated using the multiple comparison test “***The t-test of Bonferroni***” and it confirm that the GLM outperform the others models in forcasting the precipitation at the level of Beja Sation



The final simplest and best-performing model is the GLM

Data simulation and projection

At this stage, we have used the results obtained in the previous steps:

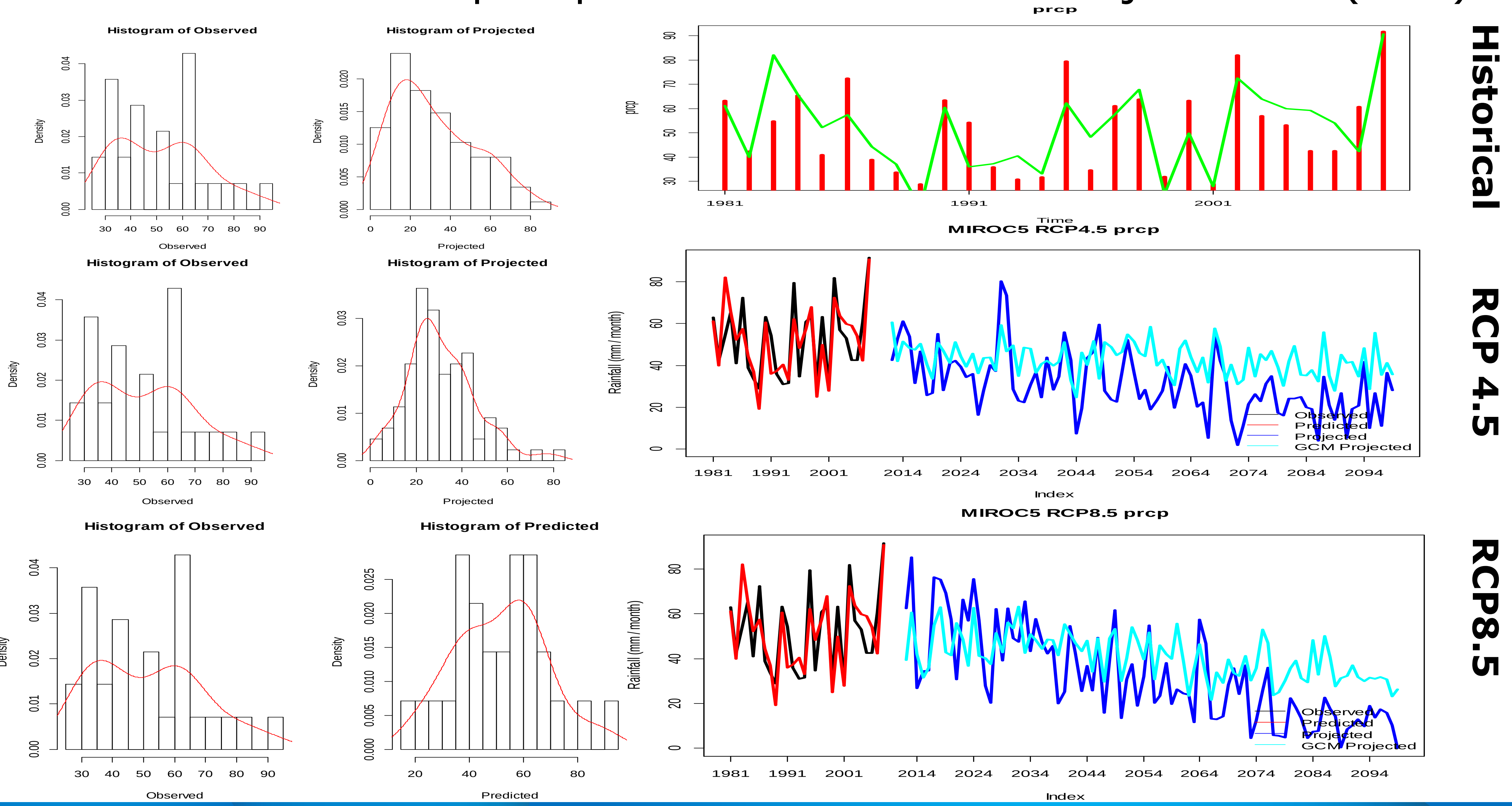
- Reanalysis data file
- MIROC5 GCM projection data file according to RCP4.5 scenario
- The selected model (GLM in the case of precipitation at the Beja station)

we repeat the same steps to extract the files corresponding to the RCP 8.5 scenario.

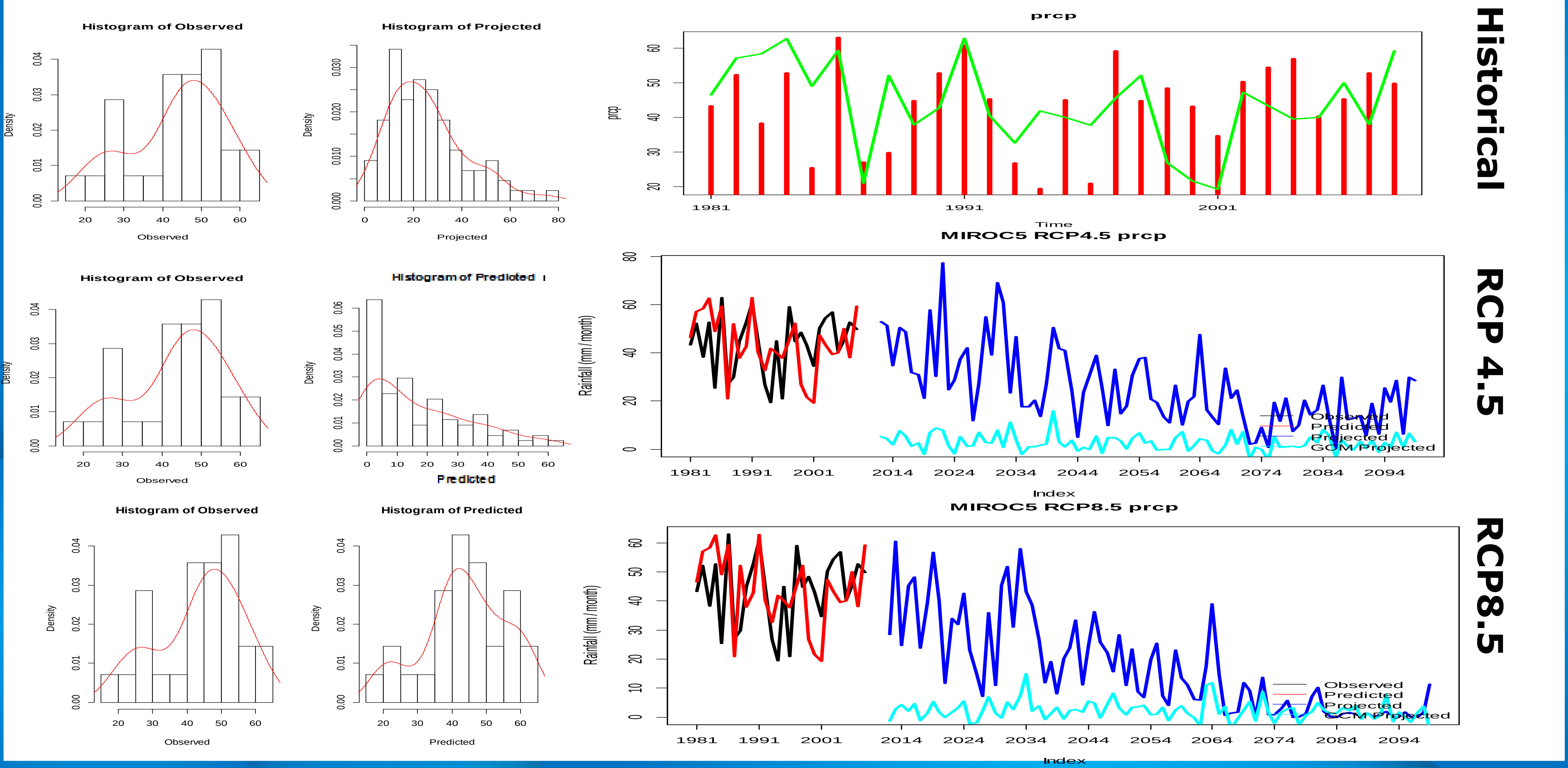
→ This led to the projection of precipitation at the Beja station by 2100 according to the two scenarios RCP4.5 and RCP 8.5

The same process (case of precipitation at the Beja station) is used to carry out the projections of precipitation by 2100 for the two scenarios RCP 8.5 and RCP 4.5 at the level of other stations.

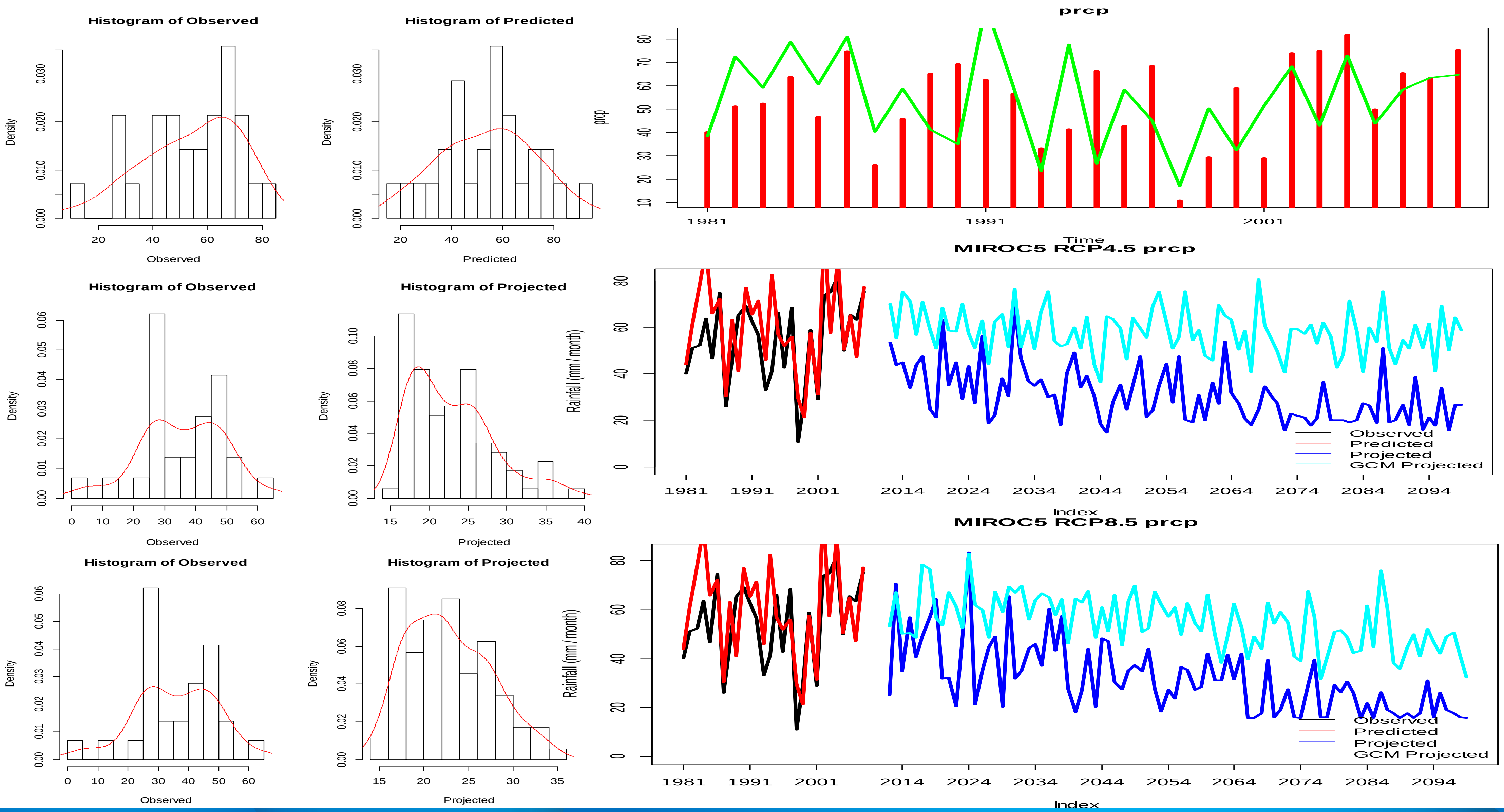
ESD simulation of precipitation at the level of Beja Station (GLM)



ESD simulation of precipitation at the level of Tunis Carthage Station(SGLM)



ESD simulation of precipitation at the level of Bizerte Station (RF)

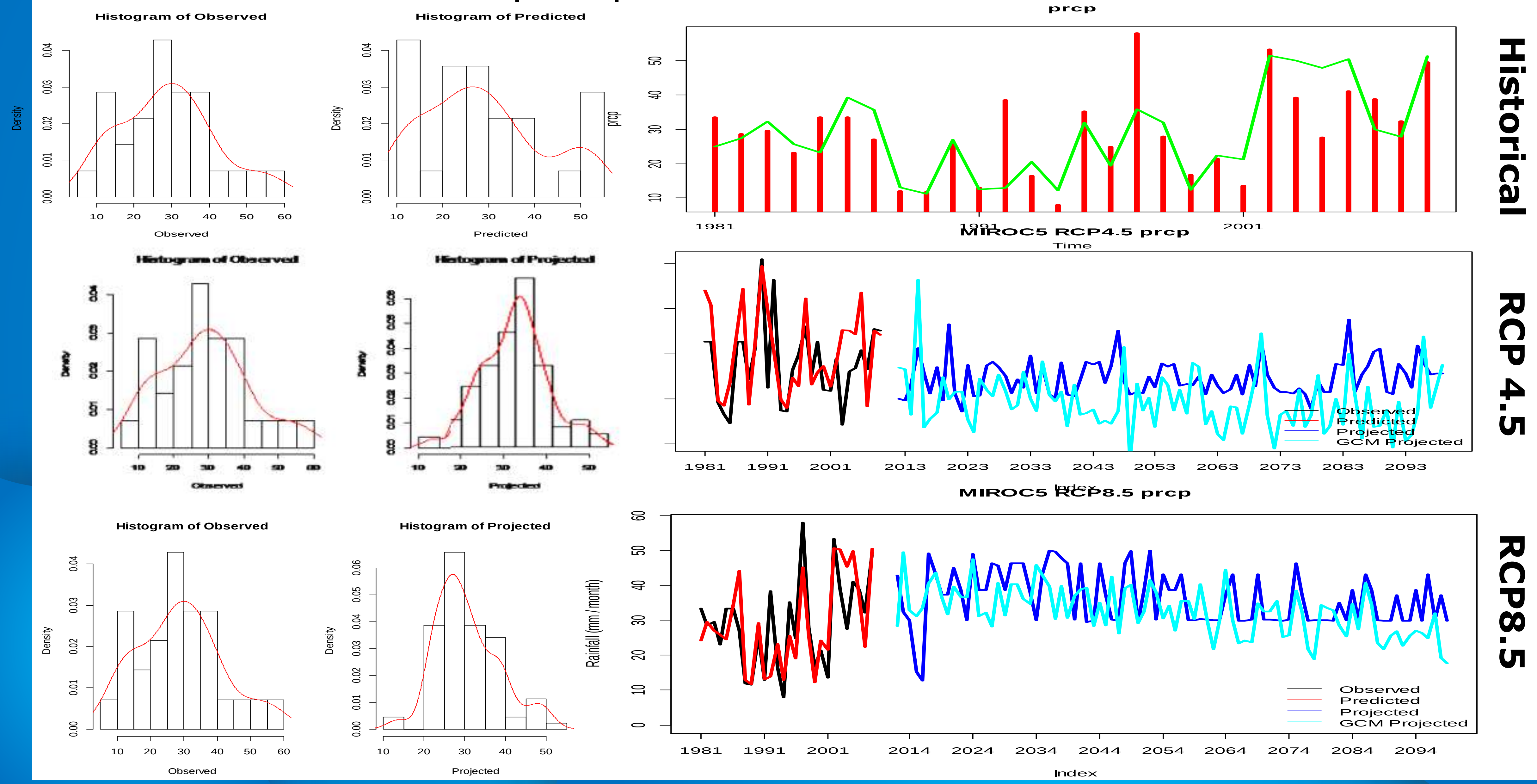


Historical

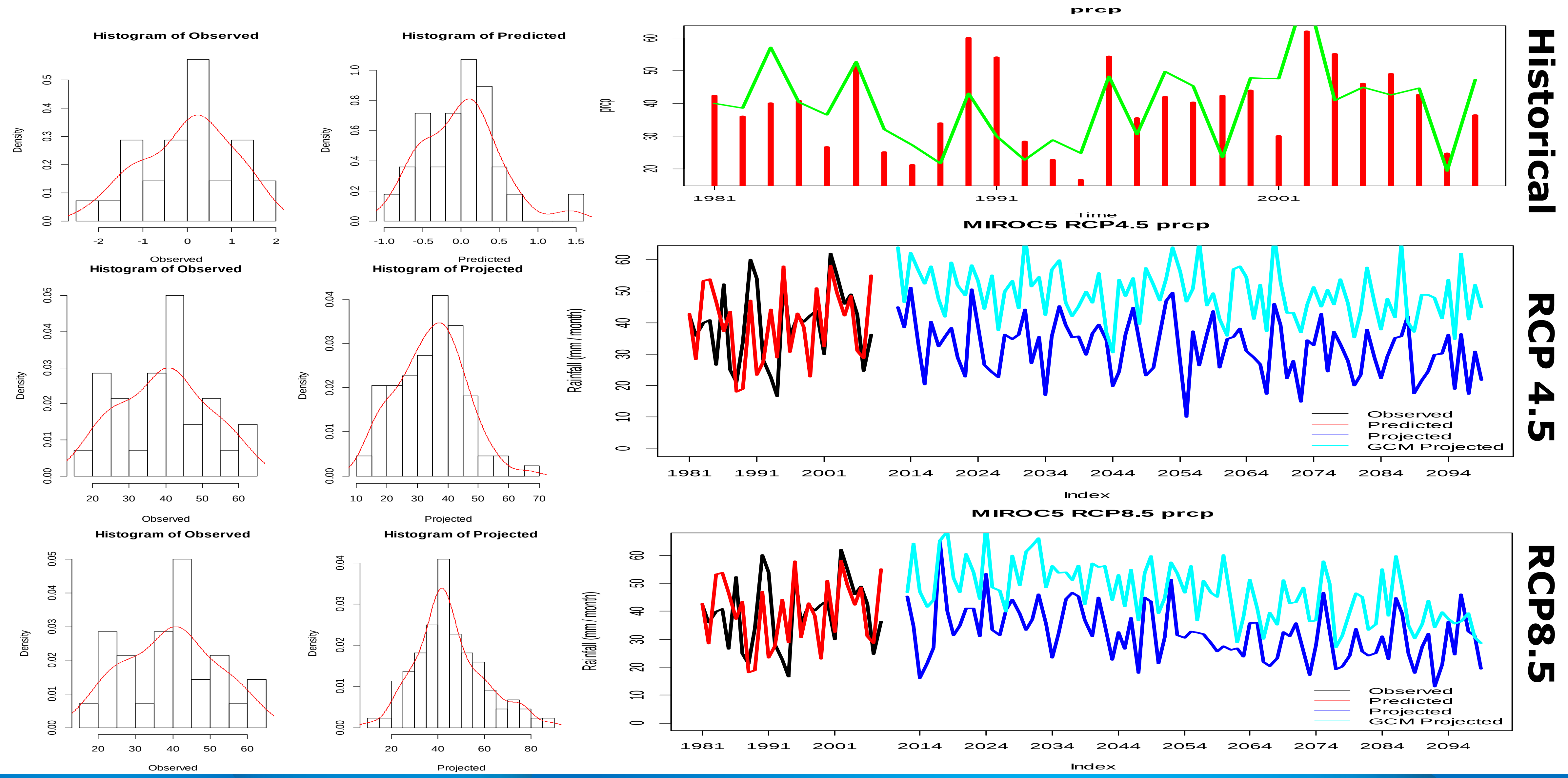
RCP 4.5

RCP8.5

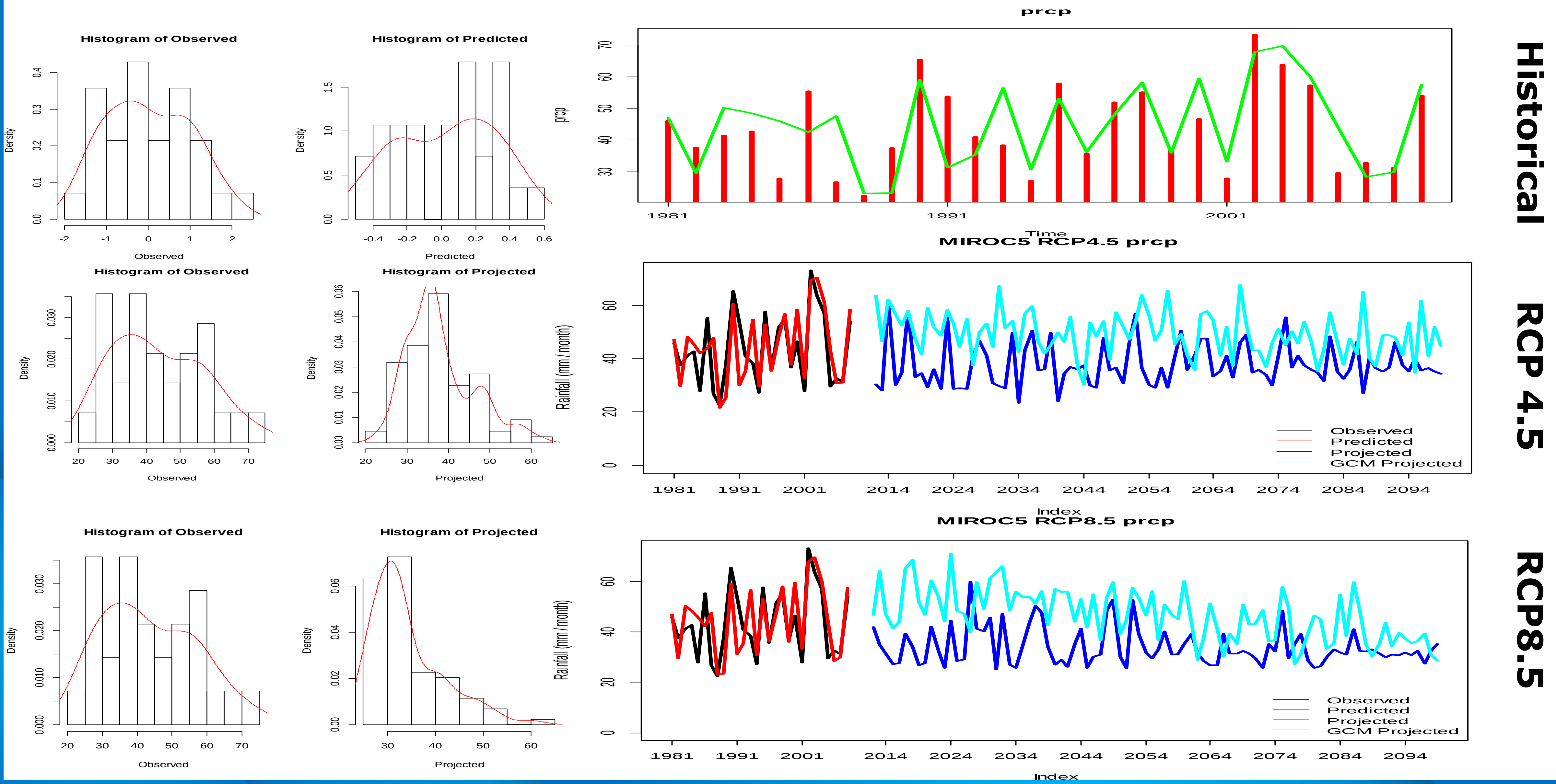
ESD simulation of precipitation at the level of Siliana Station RF



ESD simulation of precipitation at the level of Le Kef Station (SGLM)



ESD simulation of precipitation at the level of Jendouba Station (RF)



CONCLUSION

- Precipitation will undergo an robust and significant decrease.
- The decrease varies between 9% and 35% according to the RCP 4.5 and can reach more than 40% according the RCP 8.5.
- The results confirm the high spatial and temporal variability of precipitation in the Medjerda catchment and predict that this variability will increase in the future. Indeed, the decrease of the expected precipitation varies according to the different stations of the study area.

Thank you for your attention