Assignment: Module 2

Learning Goals

Subject knowledge:

- The impacts of the ENSO on global precipitation (which provide the basis for the ENSO teleconnections and ENSO-based statistical prediction models)
- Understand the difference between composite analysis and correlation analysis

Analysis skills:

- Read, analyze and visualize a netCDF dataset
- Normalize a time series
- Composite analysis: a useful tool to extract strong signals associated with a climate mode
- Student’s t-test: one of the commonly used significance tests
- Construct correlation and significance maps
- Construct a multi-linear regression prediction model using the forward selection method to choose predictors
- Test a model performance using the leave-one-out cross validation.

Problem 1

In this problem we will examine the impacts of the ENSO on precipitation by constructing the composite mean anomalies of precipitation for the positive and negative phases of the ENSO, respectively.

- Given the monthly mean Nino3.4 index, calculate the normalized seasonal mean Nino3.4 index for 1979-2020 Dec-Feb (DJF), and identify the years with the Nino3.4 seasonal mean anomalies exceeding one standard deviation or more above the long-term mean, which are selected as the El Nino years.
- Calculate the DJF seasonal mean precipitation from the monthly mean precipitation data, and then calculate the seasonal mean anomalies by removing the long-term seasonal mean during 1979-2020.
- Construct the composite anomalies of DJF seasonal mean precipitation by averaging the precipitation anomalies over the selected El Nino years
- Perform the one-sample, two-sided t-test
- Plot the anomalies and highlight the anomalies exceeding the 95% confidence level at each grid point.
- Repeat the above analysis for the La Nina years, in which the Nino3.4 seasonal mean anomalies are one standard deviation or more below the long-term mean.

You may expect that the composite anomalies during the El Nino and La Nina years are like mirror images to each other (i.e., with the same spatial pattern but opposite polarity). Do the composite analyses here support this expectation?

Input data:

/data/zhuowang/a/zhuowang/ATMS521/Data/detrend.nino34.ascii.txt

/data/zhuowang/c/zhuowang/Data/CMAP/precip.mon.mean.nc

Sample Script:

/data/zhuowang/a/zhuowang/ATMS521/Sample_Scripts/comp_precip_DJF_incomplete.py

Domain of analysis: 75S-75N, 0-360E

Time period of analysis: 1979-2019

Output: Plots of composite anomalies of DJF seasonal mean precipitation for El Nino and La Nina years, respectively, with significance highlighted.

Problem 2

Another way to examine the impacts of the ENSO is to construct a correlation map: please calculate the Pearson correlation between the DJF seasonal mean Nino3.4 index and the DJF seasonal mean precipitation at each grid point. Plot the correlation map and highlight the correlation values exceeding the 95% confidence level, and answer the following questions.

1) Do you need to remove the long-term mean before calculating the Pearson correlation?
2) Both the composite analysis in Problem #1 and the correlation analysis in this problem illustrate the impacts of the ENSO on precipitation. What differences do the two analysis methods have?

Input data:

/data/zhuowang/a/zhuowang/ATMS521/Data/detrend.nino34.ascii.txt

/data/zhuowang/c/zhuowang/Data/CMAP/precip.mon.mean.nc

Domain of analysis: 75S-75N, 0-360E

Time period of analysis: 1979-2019

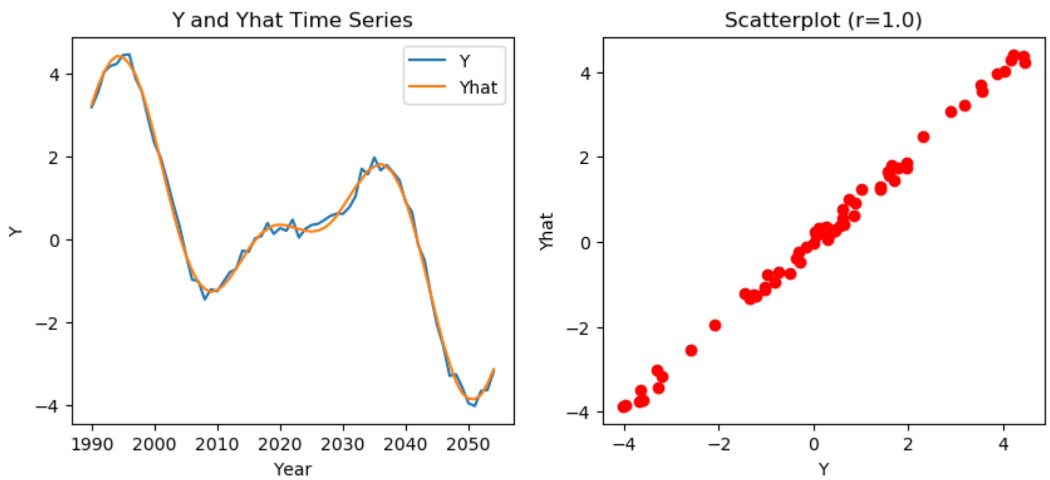
Output: A correlation map between the Nino3.4 index and the DJF seasonal mean precipitation, with significance highlighted.

Problem 3

"**MLR_sample.py**" is a simple example showing how to construct a multi-linear regression prediction model and assess the prediction skill using the leave-one-out cross validation method. The script calls a user-defined function in "**leave_one_out_MLR.py**". If X(M,N) denotes the predictors, and Y(M) denotes the predictand, where M is the length of the time series and K is the number of predictors. The script seeks a linear approximation

Y_hat=a1*x1+a2*x2+…+ak*xk+b

where ai is the linear regression coefficient (slope) for the predictor xi, and b is the intercept. The figure below shows the time series of the Y and Y_hat (left) and the scatter plot of Y and Y_hat (right).



Please develop a multi-linear prediction model for the Atlantic accumulated cyclone energy (ACE) during July-October using climate indices in the previous season (April-June or AMJ). We will choose predictors using the forward selection method and assess the prediction skill by calculating the anomaly correlation coefficient (ACC) between the predicted and observed ACE using the leave-one-out method.

- Read the ACE time series and the AMJ seasonal mean MDR SST time series from input data files.
- Construct the AMJ seasonal mean time series for Nino3.4, AMO and NAO using the monthly mean data.
- Calculate the Pearson correlation between the ACE and the different potential predictors, select the one with the strongest correlation as our first predictor, which is denoted as x1.
- Construct a simple linear regression model for the ACE using x1 and calculate the ACC between the predicted and the observed ACE using the leave-one-out method.
- Select additional predictors using the forward selection method and stop adding more predictors when the increase of ACC is less than 0.03.
- Plot the observed and predicted time series of ACE with the final set of predictors.

Input data: under "/data/zhuowang/a/zhuowang/ATMS521/Data"

- Nino34_monthly-1979-2018.txt: the monthly mean ENSO Nino3.4 time series
- AMO_monthly-1979-2018.txt: the monthly mean AMO time series
- NAO_monthly-1979-2018.txt: the monthly mean NAO time series
- RSST-1979-2018AMJ.txt: the AMJ seasonal mean time series of three SST indices: the Atlantic Main Development Region (MDR, 80-20W, 10-25N) SST, the tropical mean SST (Trop, 30S-30N), and the relative SST (RSST, MDR minus Trop).
- ace_index.1979-2018.txt: ACE time series

Sample Scripts:

- /data/zhuowang/a/zhuowang/ATMS521/Sample_Scripts/MLR_sample.py
- /data/zhuowang/a/zhuowang/ATMS521/Sample_Scripts/leave_one_out_MLR.py

Output: plot of the observed and predicted time series of ACE