# Assignment: Module 4

**Learning Goals**

Subject knowledge:

- ENSO teleconnection and its asymmetry between the two hemispheres
- ENSO as an important source of predictability on the seasonal time scale
- NAO as the leading EOF mode of the mean SLP in boreal winter

Analysis skills:

- The empirical orthogonal function (EOF) analysis
- Reconstruct a field using leading EOF modes
- Two-sample, two-sided t-test
- Lead/lag composites
- Predict a field variable using the simple linear regression model
- Cosine weighting function

**Problem 1**

We will examine the ENSO teleconnection by constructing the composite anomalies of the 500-hPa geopotential height (H500) between the positive and negative phases of the ENSO. Please follow the steps below.

- Given the monthly mean Nino3.4 index, calculate the normalized seasonal mean Nino3.4 index for 1979-2020 Dec-Feb (DJF), and identify the El Nino and La Nina years using the criterion that the normalized Nino3.4 index has a magnitude >= 1.0.
- Calculate the DJF seasonal mean H500 from the monthly mean data, and identify the subgroups for El Nino years and La Nina years.
- Construct the composite mean of DJF seasonal mean H500 for El Nino years and La Nina years, respectively, and take the difference between the two composite means.
- Perform the two-sample, two-sided t-test.
- Plot the composite differences and highlight the values exceeding the 95% confidence level.

You should see wavetrain patterns extending from tropical Pacific to the extratropics, which represent teleconnection patterns associated with the ENSO.

**Input data**:

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

/data/zhuowang/b/shared/NNR2/monthly/hgt.mon.mean.nc

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

**Problem 2**

Using the normalized Nino3.4 index during DJF, construct the composite differences of H500 between El Nino and La Nina years in the following spring (i.e., March-May or MAM).

**Problem 3**

"**SLR_sample.py**" is a simple example showing how to construct a linear regression model and assess the model skill. Assuming X(M) and Y(M) denote the predictor and predictand time series, respectively, where M is the length of the time series, the sample script derives the predicted time series yhat and the anomaly correlation coefficient (ACC) between Y and yhat using the leave-one-out method.

Please develop a simple linear regression model to predict precipitation during Dec-Feb (DJF) using the Nino3.4 index in the previous season (i.e., Sep-Nov or SON). Please predict precipitation at every grid point over (75S-75N, 0-360E) and assess the prediction skill using the leaving-one-out cross validation method. Please follow the steps below.

- Calculate the seasonal mean Nino3.4 during 1979-2019 SON
- Calculate the seasonal mean precipitation anomalies during 1979-2019 DJF
- Create the predicted precipitation time series using a simple linear regression model with the leave-one-out method, and calculate the ACC between the predicted and observed precipitation during 1979-2019 DJF at each grid point. (Note that you will repeat the model development and cross validation practice for each grid point)
- Plot the observed and predicted precipitation side by side for 2010 DJF (i.e., 2010.12-2011.02) and 2015 DJF (2015.12-2016.02).
- Plot the map of ACC with the significant correlation highlighted. This map can be used to assess the model prediction skill, and it shows that the prediction skill varies strongly from region to region.

**Input data**:

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

CMAP monthly mean precipitation: /data/zhuowang/b/shared/CMAP/precip.mon.mean.nc

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

**Problem 4**

As discussed in a lecture, The NAO time series can be derived using the EOF analysis, in which the NAO is the leading ($1^{st}$) EOF mode of the SLP over the North Atlantic in boreal winter.

- Calculate the seasonal mean SLP during December through Feb (DJF) from 1979-2019 from the monthly mean data over the Atlantic sector (20-80N, 90W-40E) (Please note that the domain crosses the prime meridian)
- Extract the first two EOF modes and the corresponding Principal component (PC) time series. Since the grid cell area decreases poleward, please apply the square root of the cosine latitude as a weighting function. This way the covariance matrix is weighted by the cosine function and the decreasing grid cell area is properly considered.
- Plot the first two EOF modes and the normalized PC time series.
- Plot the fractional variance of the first 10 EOF modes as a bar chart.

**Input data**:

/data/zhuowang/b/shared/NNR2/monthly/mslp.mon.mean.nc

**Domain of analysis**: 20-80N, 90W-40E

**Sample Script**: /data/zhuowang/a/zhuowang/ATMS521/Sample_Scripts/
EOF_SLP_incomplete.py

**Hints**: Information on how to perform the EOF analysis can be found at:
https://ajdawson.github.io/eofs/latest/api/eofs.standard.html

**Problem 5**

As discussed in Module 3, the EOF analysis is often used for dimension reduction in the weather regime analysis. Repeat the weather regime in Problem 1 of Assignment 3, but prior to applying the K-means, please reduce the dimension of the H500 daily anomaly data by retaining only the first 10 EOF modes.

**Input data** (H500 from NNR2): /data/zhuowang/b/shared/NNR2/daily/hgt.YYYY.nc

**Time period of analysis**: Dec 1979-Feb 2021

**Output**: the daily weather regime labels; a figure showing the mean H500 for each weather regime; weather regime frequency

**Hints**: No sample script is provided for this one because your script can be based on Problem 1 in Assignment #3 and Problem 4 in this assignment. You can use the "**reconstructedField**" option: https://ajdawson.github.io/eofs/latest/api/eofs.standard.html