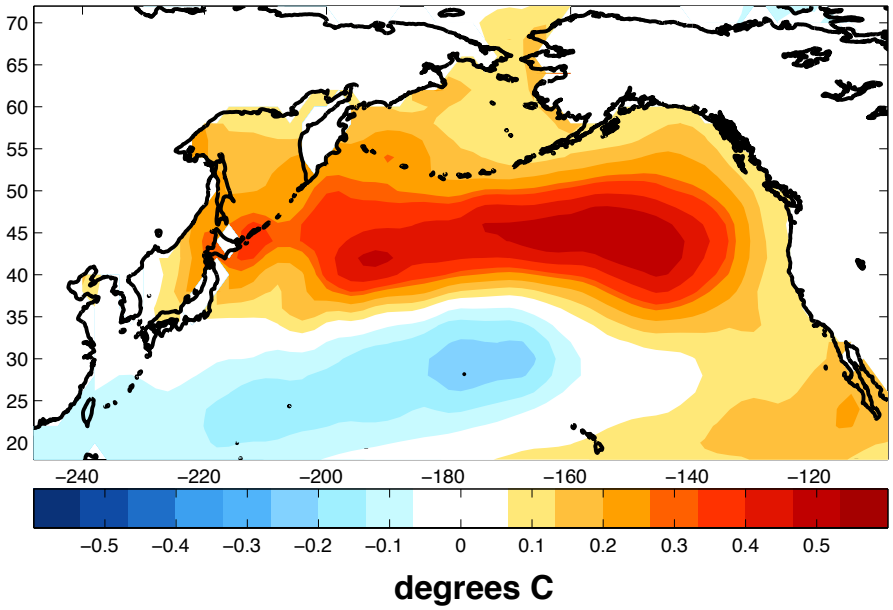
# ADVANCED ENVIRONMENTAL DATA ANALYSIS

## FINAL EXAM 2022

You have a set of sea surface temperature SST(*lon,lat,time*) and sea level pressure SLP(*lon,lat,time*) observations. You also given a pattern of SST anomalies below, where the anomalies are defined with respect to the seasonal cycle in the SST.




(a) Compute a map of the fraction of variance in the sea surface temperature anomalies (SSTa) that is is explained by this pattern. Show the linear algebra used for the computation as pseudo code.

(b) Compute how this SSTa pattern is correlated with the sea level pressure anomalies (SLPa) data over the North Pacific. Make a spatial map of the correlation coefficients with SLPa associated with this SSTa pattern. Show the linear algebra used for the computation as pseudo code.

You will find that the correlations in the SLPa are not very strong and you are left wondering if the SLPa that you have found is dynamically linked to the SSTa pattern. In the higher latitude the SSTa are typically forced by the SLPa variations.

$$\frac{dSSTa}{dt} = \alpha \cdot SLPa - \gamma SSTa$$

(c) Use this this simple AR1 model to reconstruct the temporal variability of the SSTa pattern using the SLPa pattern as a forcing function. Plot the results of your model reconstruction of the SSTa temporal variability along with the correlation skill. Describe how you picked the coefficient gamma in the model. *(hint: you need to develop time indices for the SSTa and SLPa patterns to test the model)*

(d) Develop a significance test of the correlation skill of the model using a Monte Carlo approach by assuming that the time variability of SLPa is like white noise. Make a histogram of the probability distribution function of the cross-correlation coefficients and label the value of correlation for the 95% significance threshold. Explain briefly in words how you implemented the Monte Carlo test.

(e) Now compute the correlation skill of the model reconstruction on the following timescale (i) periodicities longer than 8 year, (ii) periodicities shorter than 2 years, and (iii) periodicities between 2-8 years. For each of these frequency bands make a plot of the temporal variability and indicate if the correlation coefficient is above the 95% significance threshold using the Monte Carlo significance test you developed in (d). (*note: fit these three results in one page*)

**MATLAB data:**

Download file: FINAL_EXAM.mat

This file contains the monthly mean values of SLP and SST.

```
SLP =
        SLP: [57x23x708 double]
        lon: [57x23 double]
        lat: [57x23 double]
       mask: [57x23 double]
    datenum: [1x708 double]
       year: [708x1 double]
      month: [708x1 double]

SST =
        SST: [71x28x708 double]
        lon: [71x28 double]
        lat: [71x28 double]
       mask: [71x28 double]
    datenum: [1x708 double]
       year: [708x1 double]
      month: [708x1 double]

SSTA_PATTERN: [71x28] % same grid as the SST data
```

**QUESTION 2**

In the previous question you have evaluated how the SSTa pattern is dynamically linked to the atmospheric forcing of the SLPa. We are now going to re-approach this question using Empirical Orthogonal Functions (EOFs).

Compute the EOFs of the spatial covariance between SLPa and SSTa. Make sure to renormalize the EOFs so that the spatial patterns retain the units and the principal components (PCs) are normalized by their standard deviation.

(a) Write down how you computed the the EOFs for this problem and how you derived the principal components.

(b) Make a plot of the explained variance for the first 20 eigenvectors along with their error bars. Indicate what value you have used for the effective degrees of freedom (DOF*) in your error bas calculation. How many modes are distinct? Are there any sets of significant modes that are not independent? (*note: assess the DOF* from the autocorrelation function of the SSTa PC1*)

(c) Find out which of the SSTa modes PCs is more strongly correlated with the time variability of the SSTa pattern of QUESTION 1 (just pick the one mode that is most strongly correlated). Compare the two time series and indicate the correlation. Define this mode as mode X.

(d) Make a map of the spatial expressions of the SSTa and SLPa for mode X.

(e) Compare the PCs of the SSTa and SLPa for mode X. What is their correlation? How does the correlation change if you compare the SSTa PCX with the output of the simple dynamical model (QUESTION 1c) forced by SLPa PCX? Make plots and indicate the correlations and their significance.