Final: SIO 221B, Data Analysis

due Friday, December 13, 2002
Open book. Open note. Do not discuss these questions with each other. I will be happy to address any questions you might have.

1. (10 points) Consider a small data set that can be expressed as a matrix:

$$A = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 2 & 1 & 2 & 1 \\ 3 & 3 & 2 & 4 \\ -1 & 1 & -2 & 2 \end{bmatrix}$$

What is the rank of this matrix? How much of the variance in the data can be expressed by the first EOF? How much by the first two EOFs?

2. (20 points) You would like to estimate the meridional eddy temperature flux $\langle v'T' \rangle$ using a current meter and a thermistor. How long should you deploy your instruments in order to be able to distinguish the mean flux from zero (using 5% confidence limits)? At present you have made some preliminary measurements that have yielded the following information (some of which may prove irrelevant):

a. The lagged covariance of temperature can be fit to a Gaussian with an e-folding scale of 35 days.

b. The lagged covariance of velocity can be fit to a Gaussian with an e-folding scale of 10 days.

c. The lagged covariance of $\{v'T'\}$ can be fit to a Gaussian with an e-folding scale of 30 days.

d. The standard deviations of v, T, and $\{v'T'\}$ are all 2 in the appropriate units.

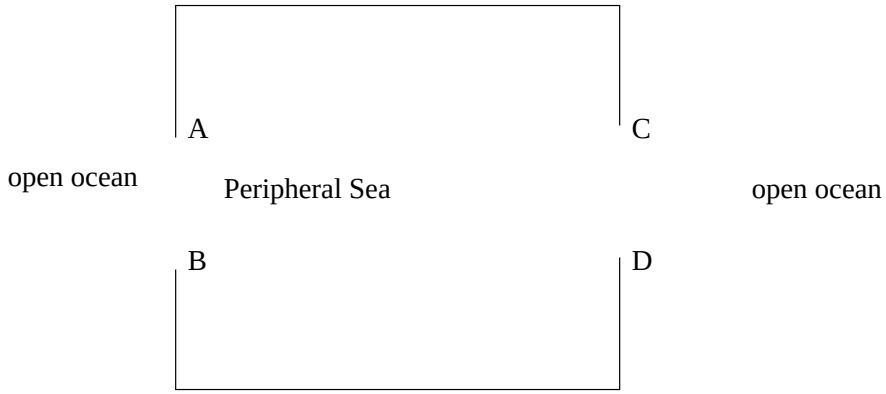
e. The means of v, T, and $\{v'T'\}$ appear to be about 1.

3. (20 points) You have collected the following data, which are irregularly spaced in time. In this part of the world, the covariance of temperature is assumed to be Gaussian with an

t (yrs)	T (°C)	σ_T (°C)
1920.	17.1	4.
1922	15.	4.
1923	14.8	2.
1923.5	14.4	2.
1927	19.	1.

e-folding timescale of 2 years. The variance of the signal is assumed to be 4 at zero time lag, and the variance of the measurement noise varies depending on the method used to collect each observation, as indicated in the table. Use the formalism of objective mapping to estimate the temperature at this location in 1925 and the uncertainty of your estimate. (Hint: be sure to remove the mean before doing the calculation.)

4. (20 points) On planet X, the Peripheral Sea has two entrances, both of which are deep, flat-bottomed channels that are substantially wider than the local Rossby radius, as shown in the figure. You have collected four hydrographic stations, one on either side of the the channel entrances, at point A, B, C, and D.



How can you estimate the total transport through each of the channels? Set up the matrix equations that you would use to solve for this transport. What additional information would you need to determine your solution?

5. (30 points) After carefully analyzing numerical model output for the North Pacific, oceanographer X has published a paper stating that fluctuations at 60°N are inversely correlated with fluctuations at 40°N. Moreover, in the model output, these fluctuations appear to be linked to the Pacific Decadal Oscillation (PDO).

You'd like to examine whether the behavior of the ocean resembles the model. Unfortunately, you are stranded on a desert island, and the only data that you have available are irregularly spaced XBT profiles and a time series of monthly PDO variations. The friend who sent you the PDO index rather cryptically remarked that she hoped you had a good singular value decomposition package, since she was sure you'd need either EOFs or a least-squares fitting technique to make sense of this data.

Write a clear discussion, explaining how you could use (a) EOFs and (b) least-squares fitting to analyze your data. Note that there may be more than one approach to this problem (and EOFs and least-squares fitting are not the only ways to treat this data).

For the EOF calculation, how would you actually go about sorting the data into a usable form? What would the elements of your matrix be, how would you handle this matrix, and how would you interpret the results of your analysis? What are the limitations of this analysis?

For the least-squares fitting, how would you organize your data and carry out your calculations? What quantities might you fit to what other quantities? What could you hope to learn from this analysis? What problems might you encounter?