Assignment 1: SIO Pier temperature and salinity time series
Due October 20, 2020

Figure 2. (a) The original 1,000-ft) (305 m) Scripps Pier (right) built in 1916, alongside the new 1,084-ft (330 m) pier during construction, 1988. (b) Pier sampling well, Claude W. Palmer, 1949. (Photographs from SIO Photographic Laboratory, UC San Diego Digital Collections).

Figure from Rasmussen et al. (2019)

Daily measurements of temperature and salinity have been collected from the SIO Pier since 1916, resulting in one of the longest continuous hydrographic time series in the Pacific Ocean. Samples have been collected at the end of the pier near the surface and bottom, providing some indication of vertical structure in the nearshore water column.

The measurements represent a spot sample taken at varying times during the day, so aliasing of the daily heating cycle is a concern, as are internal waves that particularly impact the near-bottom temperature readings. Rasmussen et al. (2019) describe adjustments to the data to account for time of day sampling. We'll start by looking at the raw data, and later compare our results with the adjusted time series of the near-surface temperatures.

Assignment:

1) Load the mat file SIO_PIER_1916_201905.mat. The data are from the site https://shorestations.ucsd.edu/data-sio/. The file variables are:

t: time in Matlab serial date number, or days since January 0, 0000
Ts: surface temperature (°C) at ~ 0.5-m depth
Tb: bottom temperature (°C) at ~ 5-m depth
Ss: surface salinity ~ 0.5-m depth
Sb: bottom salinity ~ 5-m depth

Let's get a first impression of the data. Construct the following plots:

a) Each variable over the entire record
b) Each variable over a one-year period
c) same as a) and b) but for the differences Ts-Tb and Ss-Sb

Describe any notable features of the data that stand out to you.

2) On a time series plot of Ts, indicate the max and min daily values ever recorded.

3) The daily values are interesting, but given all the sampling issues, let's average the data down a bit to emphasize the low frequency variability. Compute and plot the monthly-averaged time series for each variable. These should look very similar to the daily time series, just smoother. A running average is a basic type of filtering technique.

4) Let's get a sense of what a typical year looks like. Compute and plot the seasonal cycle for each variable (i.e., the mean, over all years, of each month). For practice, try using subplot and put all of this on one plot. Also see if you can label the x axis to indicate the month ('Jan', 'Feb', etc).

5) The seasonal range is so large, that it tends to obscure nonseasonal variability, which often gives insight into various oceanographic phenomena (El Ninos, heat waves, trends, etc.). Compute and plot the anomaly for each variable by removing the seasonal cycle from the monthly time series.

6) How are these variables distributed? Plot histograms (normalized to a PDF) of the Ts and Ss anomaly time series. Compare to a Gaussian distribution by overplotting a Gaussian PDF using the sample mean and standard deviation of each time series.

7) Can you see evidence of long-term change using histograms? Compare histograms (normalized to a PDF) of the Ts and Ss anomaly time series for the first 20 years of the record compared to the last 20 years.

8) To emphasize how things have changed, let's estimate an exceedance probability. Assuming a Gaussian distribution, what is the probability that the max monthly Ts in the first 20 years of the record is exceeded in the last 20 years of the record?

9) Let's look at some trends. Compute year-averaged time series from the monthly Ts and Ss anomalies, as well as Ts-Tb (a rough measure of stratification). Compute the linear trends for each time series and give the 95% confidence interval. Assume that each year represents an independent data point. Which trends are significantly different than zero?

10) Compute a 2nd-order polynomial to each time series. Do any of the series exhibit a significant acceleration?

11) There appears to be a noticeable uptick in temperature since the mid 1970s. Compute the linear trend and 95% confidence of each series since 1975. How does the post-1970s trend compare to the pre-1970s trend?

12) It's always dangerous to extrapolate trends, but let's see how different these trend estimates are going forward in time. Extrapolate Ts to 2100 using the linear fit over the entire record, the 2nd order fit, and the linear fit since 1975. Plot each of these estimates on a single time series along with the monthly-averaged anomaly. Comment on these results. When would these projections start to diverge?