METR 5433 - Advanced Statistical Meteorology
Spring 2020 – Homework #1
<u>**Due:**</u> 11 February 2020

(#1) This problem examines the statistical relationship between 500 hPa heights and the El Niño-Southern Oscillation (ENSO) during boreal winter (December, January and February). For this question, you will use two files: (1) *ENSO.txt*, which contains monthly-mean standardized values of an index measuring ENSO strength and (2) *Z500.nc*, which contains gridded monthly-mean 500 hPa geopotential heights globally from 1950 - 2019. The steps for the analysis and final product will be as follows:

- Calculate 500 hPa height anomalies, use the 1981 – 2010 base period as your climatology / mean.
- Calculate the December - February (DJF) mean 500 hPa geopotential height anomaly field and ENSO index from each complete winter season in the dataset.
- Generate composite DJF-mean 500 hPa height anomaly plots for times when the ENSO index exceeds 1σ (i.e., Positive ENSO) and exceeds -1σ (i.e., negative ENSO). Use shaded contours for plotting.
- Compute the composite difference (Positive ENSO - Negative ENSO). Identify regions where the composite difference is significant at the $p < 0.1$ level using either a t-test or a Monte Carlo test (try out both to see potential differences!). Consider each event as a degree of freedom for the calculation. Outline or stipple these significant regions on the composite difference plot.
- Assemble the plots for presentation as a 3-panel plot, with the positive ENSO composite on top, the negative ENSO composite in the middle, and the composite difference on the bottom. Make the plot only for the Northern Hemisphere (i.e., 20-90 N).

Provide a short (1-2 paragraphs) physical interpretation of your results. Discuss the regions of statistical significance in particular and what implications these may have on understanding the extratropical influence of ENSO on global weather patterns. Also explain how you did your significance testing and if there is a difference in the two statistical techniques used.

(#2) The file *BostonDailyMaxPrecip.txt* contains the daily maximum precipitation per year from official Boston records (1893-2018). The data provided are in three columns: (1) The maximum daily precipitation in inches (in); (2) The date of the event; (3) The maximum daily precipitation in millimeters (mm).

(a) Using the data provided, fit a Gumbel distribution to the Boston precipitation data in mm. Use the methods of moments to estimate the parameters for the distribution.

(b) Make a plot with a histogram of the annual daily-maximum precipitation (in mm) and the fitted Gumbel distribution overlaid on the histogram. On your plot, indicate the values of ζ and β used for the fit.

(c) Using your fitted distribution, determine what the probability is that Boston will experience a year with a daily maximum precipitation amount (i) greater than 130 mm (about 5 in) and (ii) less than 50 mm (about 2 in).

(d) Calculate the daily maximum precipitation amounts corresponding to a 1-in-50 year, 1-in-100 year, and 1-in-500 year event in Boston.

Note: In python, you may want to look up the function stats.gumbel_r from the scipy library to check answers.

(#3) A popular topic in wintertime subseasonal and seasonal forecasting is using the state of the stratospheric polar vortex to make skillful predictions of tropospheric weather regimes. One mode of climate variability that forecasters seek to predict is the *Northern Annular Mode* (NAM) in both the stratosphere and the troposphere. For stratospheric levels, strongly positive (negative) values of the NAM index (e.g., NAM at 50 hPa; NAM_{50}) typically mean that the stratospheric polar vortex is stronger (weaker) than normal - i.e., climatological westerlies in the stratosphere are stronger (weaker) than average. In the troposphere, strongly positive (negative) values of the near-surface NAM index (e.g., NAM at 1000 hPa; NAM_{1000}) indicate that the mean position of the polar jet stream is more poleward (equatorward) than normal, suggesting warmer (colder) than normal temperatures across the Northern Hemisphere mid-latitudes. In this exercise, you will investigate how the probability density function (PDF) of NAM_{1000} may change given the background state of the stratospheric circulation.

Download the files NAM50.txt and NAM1000.txt, which contain daily-mean standardized values of the NAM index at 10 hPa and 1000 hPa, respectively. Restrict your analyses to the extended boreal winter season (i.e., November - April).

Calculate the climatological PDF of NAM_{1000} (i.e., all NDJFMA days). Then, on the same plot, overlay (in different colors or linestyles) two other *conditional* PDFs:

- The PDF of NAM_{1000} for days when $\text{NAM}_{50} > +1\sigma$.
- The PDF of NAM_{1000} for days when $\text{NAM}_{50} < -2\sigma$.

Be sure to take note of the number of samples in each PDF and what distribution you used for the fitting, based on the histograms. Justify your choice. Note - the plots of the PDFs should be your *fitted* PDFs, not the histograms. Then, calculate the probability that the NAM_{1000} index will be $< -1\sigma$ in the general climatology and in each of the two conditions above. Finally, use the non-parametric Kolmogorov–Smirnov test[1] to test if the two conditional PDFs are significantly different from each other at the $p < 0.05$ level. (i.e., they are two distinct distributions).

Provide a 1-2 paragraph discussion of your findings above, including whether or not this analysis supports the idea of using the strength of the stratospheric polar vortex as a potential predictor for Northern Hemisphere surface weather regimes.

[1] You can read more about the K-S test online: e.g., http://www.itl.nist.gov/div898/handbook/eda/section3/eda35g.htm. The test is designed to check if two PDFs are significantly different from each other, which is useful for conditional composites. Look up the K-S test for your specific programming language for details.