**DSC 275/475: Time Series Analysis and Forecasting (Fall 2021)**

**HW #3 (Total points: 50)**

***Instructions/Suggestions:***

* For this homework, you can use external libraries/functions to implement the time-series models.
* ***For R users,*** 
  + Recommend using the “Arima” function in the “forecast” package (<https://www.rdocumentation.org/packages/forecast/versions/8.13/topics/Arima>). In this function, set “include.constant” to be TRUE to match the standard equation for ARIMA we are using in the lectures.
* ***For Python users,*** 
  + For Q1 and Q2:
    - Use the “ARIMA” function from “statsmodels.tsa.arima\_model”. Use the “summary” method to obtain the model parameters. You may get a warning message about deprecation when using “ARIMA” from this package but that is OK and you can ignore that.
  + For Q2, you can use the *predict* method after fitting the ARIMA model to obtain the model output values. Make sure to specify the argument *typ='level'* when using this function.

1. The data provided in the file Measurement\_Q1.xls exhibits a linear trend. Apply the following models to the data. **(20 pts)**
2. Develop an IMA(1,1) model for the data. Display the model parameters obtained as your output **(8 pts)**
3. Compute and plot the first difference of the data. **(2 pts)**
4. Now, develop an MA(1) model on the first difference. Display the model parameters obtained as your output **(8 pts)**
5. Based on the model parameters you obtained in (a) and (c), comment on how the two models are related. **(2 pts)**
6. Consider the global mean surface air temperature anomaly data provided in GlobalAirTemperature.xls. **(15 pts)**
7. Apply an IMA(1,1) model to this data. Calculate the SSE by comparing the model output with the data. **(6 pts)**
8. Now, apply an IMA(1,2) model. Calculate the SSE. **(6 pts)**
9. Comment whether model (a) or (b) is better suited for this data based on the SSE? **(3 pts)**
10. Review the dataset in the file Measurement\_Q3.xls which contains measurements recorded annually over close to 50 years. **(15 pts)**
11. Plot the time series, ACF, and Partial AutoCorrelation Function (PACF). **(6 pts)**
12. Compute the first difference and plot its ACF and PACF. **(6 pts)**
13. Based on the results in (b) above, what model order, i.e. *p,d,q* in ARMA(*p,d,q*), would you recommend for the above time series? Provide a justification for your answer. **(3 pts)**