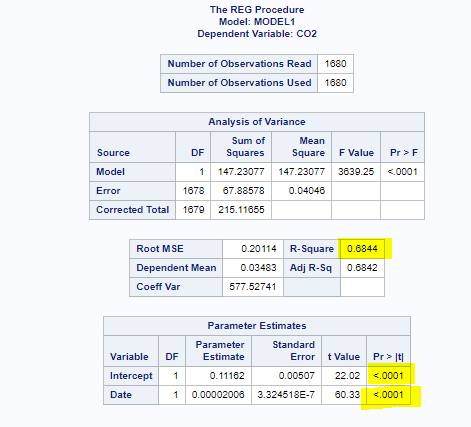
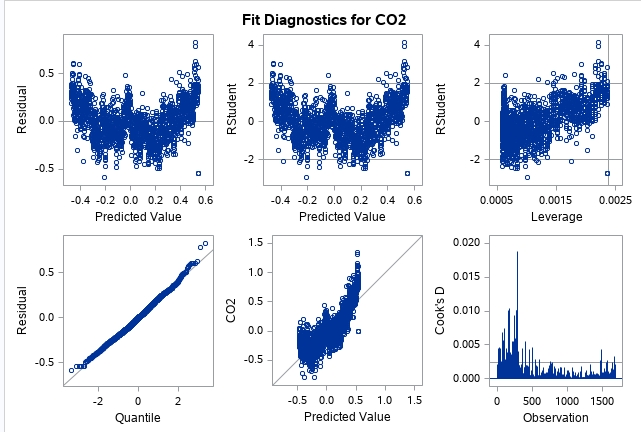
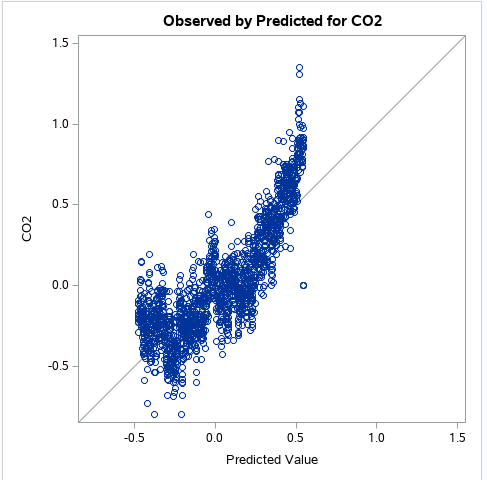


Durbib- Watson D between 0-2 means there is a positive correlation at 92% of 1st Order Correlation(1 time unit lag).

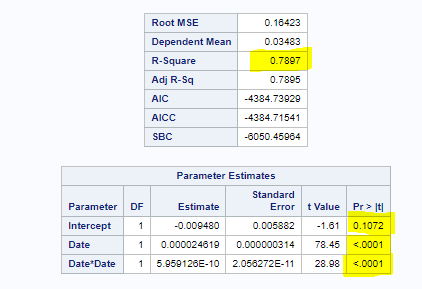


The summary shows a linear regression of CO2 emissions vs time. The p-value(<0.05) suggests that the model is significant. The model also defines 68% of the variability in data(R-square=0.6844). The equation for our regression model will be:

CO2 emissions = 0.00002008\*Date+0.11162



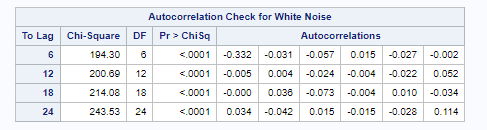
The residual values doesn’t seem random but normal. We can also see that for higher values of CO2 emissions (>0.5) the variance increases. So, we will check a squared model to see if that explains the data prediction better.



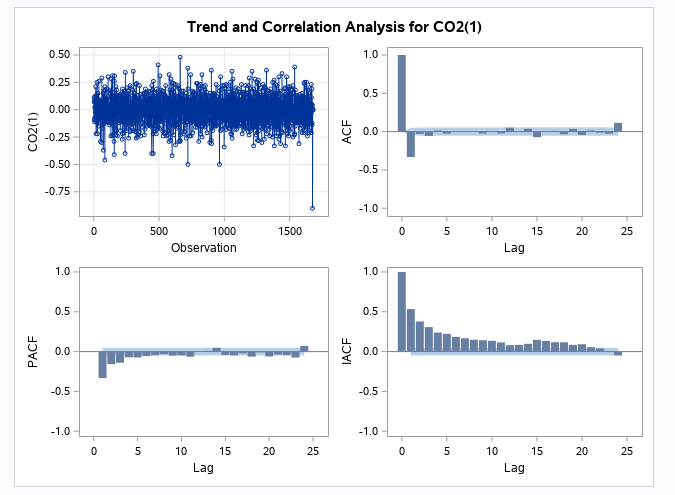
The squared regression shows better R square of 78% and p-values show that Date and Date^2 coefficients are significant but intercept is not. New equation will be:

CO2 emissions = 5.9E-10\*(Date^2) +0.0000246\*Date-0.009480

Now we will run the time series using ARIMA that includes Auto Regression and Moving Average. To run an ARIMA model we need to define p (lags in Auto Regression), d (non-seasonal difference) and q (lagged forecast errors in Moving Average). These attribute will be defined by checking for seasonality, ACF and PACF plots.



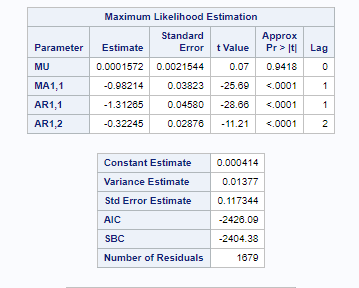
The Autocorrelation check is used to test for white noise. If the p-value is significant, we can say that the data is correlated else the data is independent. Here the tables shows the autocorrelations at different lags and p-values suggest that the data is correlated.



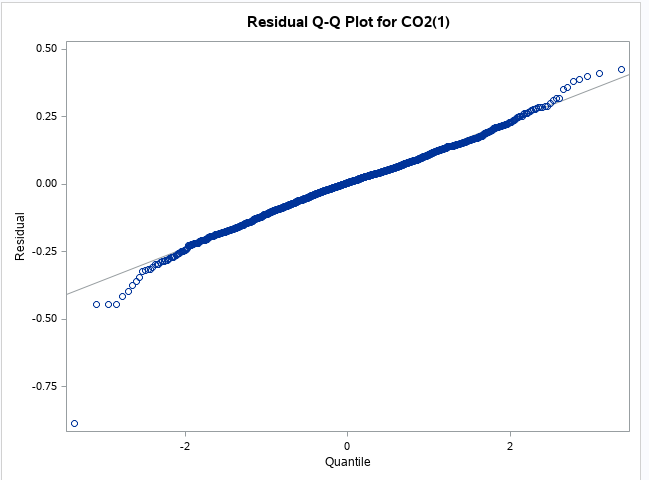
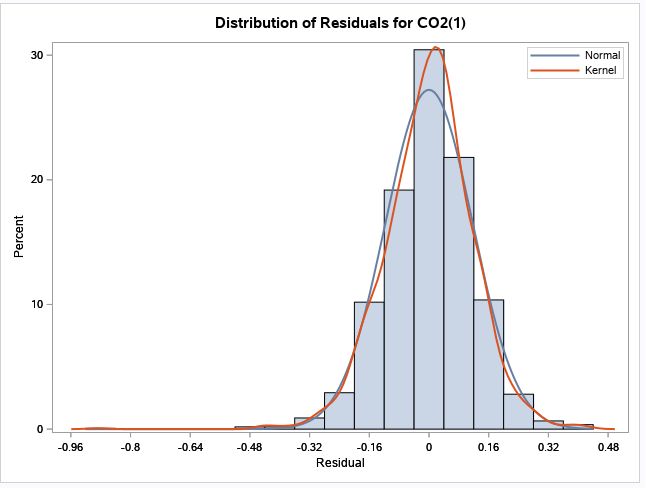
The ACF and PACF plots are used to identify p (lag for auto regression) and seasonality. We can see that ACF plot starts with a positive value and then continues with negative values till 12. But there is no pattern following.



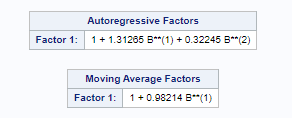
So, we can say that AR is explained very well using lag-1. Also, the PACF plot cuts off at 2. We will iterate through different pdq values and get the best estimates with lowest AIC score.



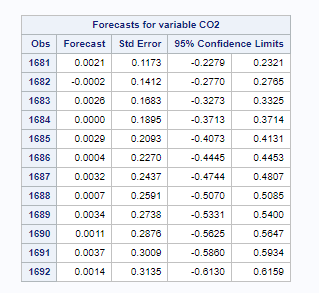
The pdq (2,1,1) shows better AIC -2426 compared to other pdq values as well as squared regression. The p-value <0.05 also signifies that the parameters we have selected are good. We will predict using these parameters



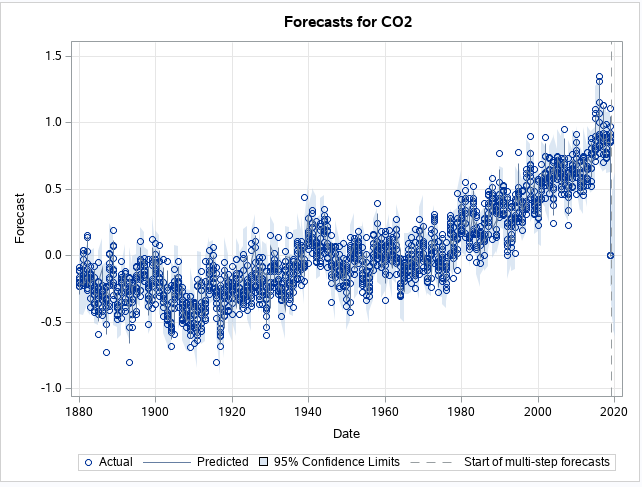
The distribution of residuals are normal unlike regression and squared regression



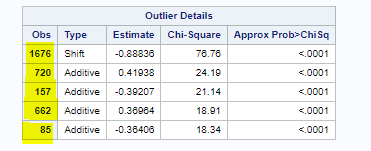
The tables show the equation for Autoregression and Moving Average prediction of ARIMA model



This tables shows the forecast of next 12 months of data



Graphical Forecast highlighted by line at the end and connected with the existing data. So this plot shows the complete trend of historical data+predicted data



The last table shows the outliers with row number and values of the observations.

Link to participate as a peer reviewer: <https://centre.humdata.org/call-for-reviewers/>

Do the panelists have frameworks, playbooks or checklists that humanitarians involved in forecasting should go to when planning and executing on humanitarian and development work?