Forecasting Surface Temperatures to help UNEP Implement Special Sustainability Programs for Target Countries

Group B7

61710175	Aneesh Chandran
61710726	Vaibhav Mathur
61710314	Pradeep Kumar Grandhi
61710251	Nishikant Mishra
61710309	Divya Dewan
61710269	Mahesh Panse

Table of Contents

Executive Summary

Increasing global land temperature has become a major cause of concern for countries around the world. Remarkably, this is the third consecutive year a new global annual temperature record has been set. The average global temperature across land and ocean surface areas for 2016 was **0.94°C** (**1.69°F**) above the 20th century average of **13.9°C** (**57.0°F**) (1). Increasing global temperature is an indicator of global climate change. This phenomenon affects energy consumption, precipitation cycles and crop production apart from a rise in sea levels which can displace people living near the coastal areas. Despite alarming evidences and consistent warnings from scientific communities to curb the impact of rising temperatures, the governmental organizations have not responded at the desired pace.

United Nations Environmental Programme (UNEP), a UN body which coordinates its environmental activities by assisting countries in implementing environmentally sound policies and practices, wants to introduce its **Special Sustainability Programs** for countries which are at most risk due to the rising surface temperature.

Problem Description:

Business Problem: Successful implementation of the special sustainability program is an important factor in help UNEP reduce the impact of increasing global temperatures. To help implement these measures, UNEP wants to study their impact and use the results to convince member countries to implement them in the future.

Forecasting Problem: UNEP has approached us to create a forecasting model to predict land temperatures for the selected countries. The goal is to:

- Forecast yearly average temperature for next 5 years (2013-2017) to sensitize the countries about the risk of climate change and convince them to deploy the existing sustainability measures in a more aggressive manner
- Forecast monthly average temperatures for next 24 months (Jan'13-Dec'14) as benchmarks to study the effect of new sustainability programs (using test and control groups)

Success criteria:

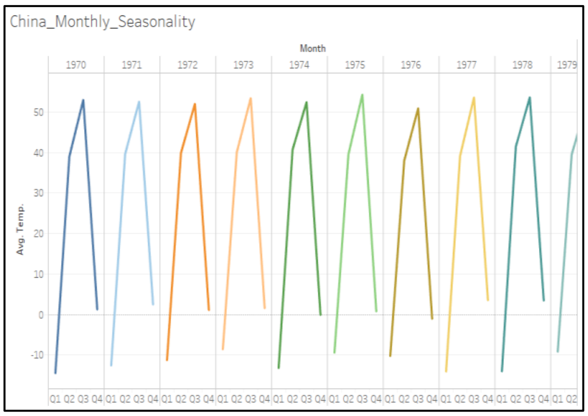
Actual temperatures recorded in the future should be as close to the forecasts as possible. Any positive deviation would signify the country is enforcing the sustainability measures properly.

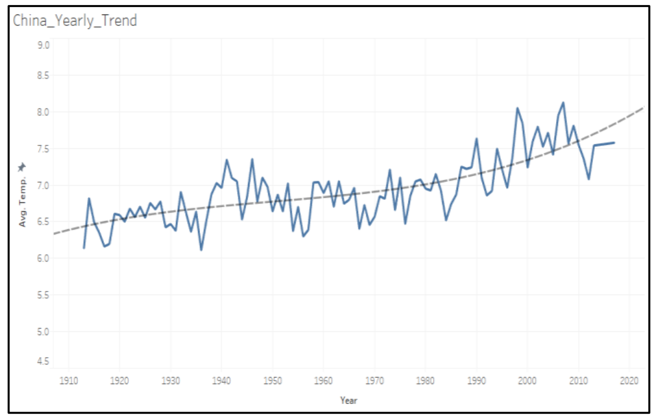
Data Description:

UNEP has provided us with data for monthly average temperatures from Jan' 1750 to Dec' 2012 for 10 countries. These countries were directly selected by UNEP based on several factors such as industrialization in the country, developing/developed status of the country, CO_2 emissions etc. Here we have used the data for a single country, China, to forecast its land temperatures. In addition, we analyzed the data for the remaining 9 countries including India, US, UK, South Korea, France, Germany, Saudi Arabia, Australia and Canada.

Key Characteristics: Through initial EDA, we find that the data has the following characteristics:

1. Trend: The yearly data (average over 12 months) follows a 3rd order polynomial trend
2. Seasonality: The monthly average land temperature data follows a monthly seasonality with a slightly increasing upwards linear trend.



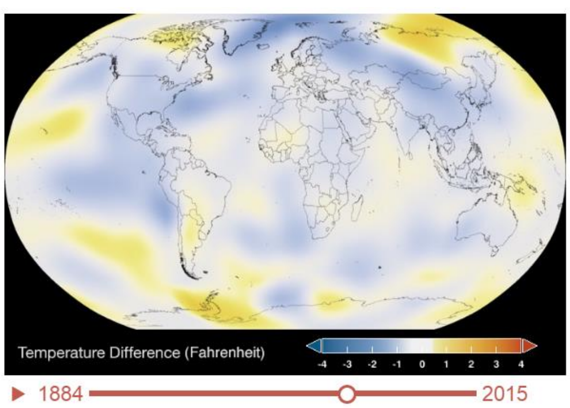


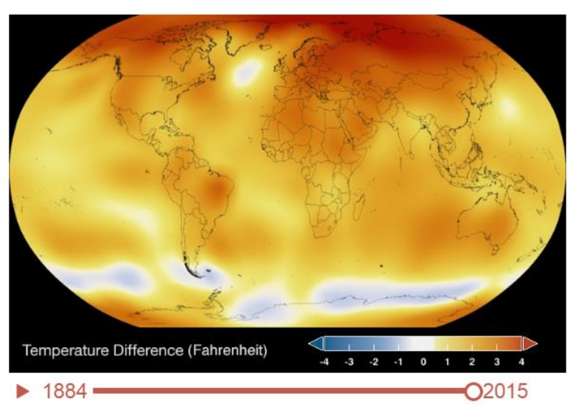
Technical Summary

Data Preparation:

As per the business goal defined earlier in the report we had to produce two kinds of forecasts viz. monthly (24 forecast periods from Jan’ 2013- Dec’ 2014) and yearly (5 forecast periods from 2013-2017).

- For monthly forecasts, we considered data from 1970-2012 since the change in temperature became visibly important after 1970. Through EDA, we concluded that there was slight linear trend visible along with monthly seasonality. Hence we decided to use Naive (Lag-12), Holt-Winters (Additive), and Multiple Linear Regression to test the data and forecast future values.





- For yearly forecasts, we considered data for the past 100 years from 1913-2012 to help produce forecasts for next 5 years since this was the period when affect was significantly visible due to rapid industrialization. Through EDA, we concluded that yearly data had a 3rd order polynomial trend in most cases. Hence we selected Naïve (Lag-1), Double Exponential, and Multiple Linear Regression to test the data and forecast future values.

Performance Metrics:

The performance metric we used in this case was the Mean Average Percentage Error (MAPE). This was because we are not interested in comparing countries against each other, but the change in the temperature of a particular country w.r.t its past temperature. This was important since UNEP is trying to identify individual countries to implement its sustainability measures. We also calculated the RMSE for both training and validation period separately for each method and these values were compared to check the overfitting (if any) of the data. The model with least MAPE and no overfitting (through RMSE) was further used to forecast monthly and yearly values.

After generating forecasts with the best model, we also checked for autocorrelations using ACF plot. If there was any autocorrelation (lag 1 or lag 2) we used the ARIMA model to calculate the "errors of errors". Next we ran ACF plot of residuals of ARIMA to check if any autocorrelation still existed. The final forecasted values included the values of best fit model + ARIMA along with respective confidence and prediction intervals.

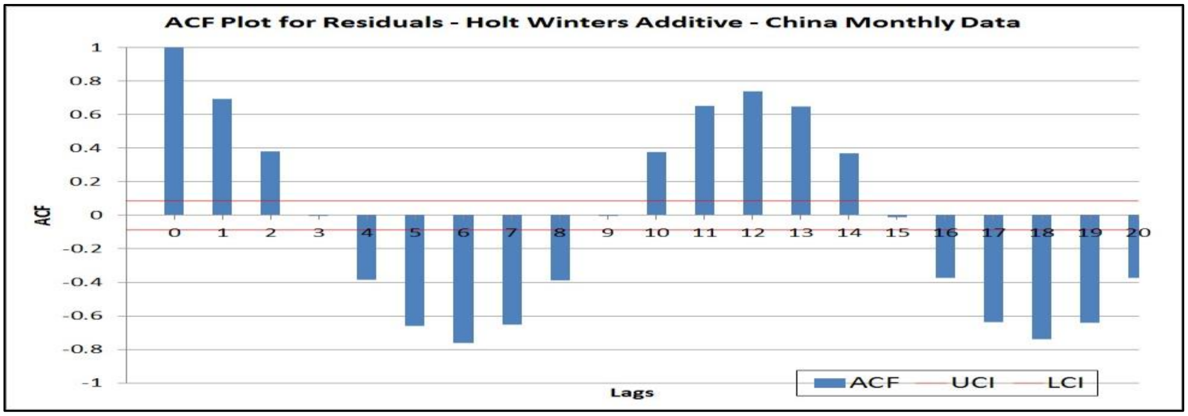
Test Case - China:

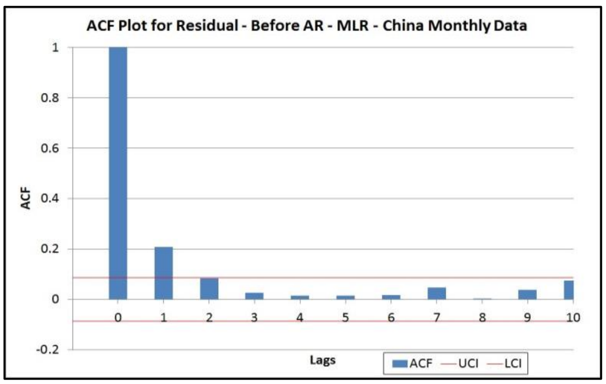
Like most other countries, the monthly data for China had a linear trend with monthly seasonality. Applying the 3 models to this data, we identified that Holt Winter's had the lowest MAPE. However, on checking autocorrelations for the residuals using ACF plot, we observed seasonality and concluded that Holt-Winter's was in fact not capturing the seasonality well. Therefore, we decided to use MLR (with second best MAPE) for forecasting monthly values.

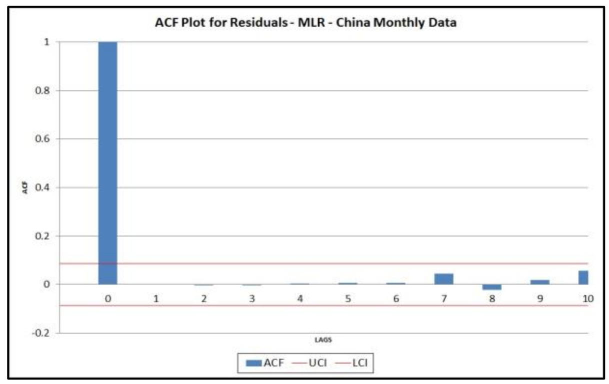
CHINA MONTHLY DATA MODEL SUMMARY

	Naïve	MLR	Holt-Winters Additive Default	Holt-Winters Additive Optimized
Training				
MSE	1.32887519	0.549697658	0.67699064	2.839440686
RMSE	1.15276849	0.741415982	0.822794409	1.685064001
MAE	0.905880952	0.550918175	0.623796791	1.397333538
MAPE	0.548453672	0.518731594	0.429176332	0.626202646
Validation				
MSE	1.739458417	1.369293337	1.318671129	0.731695631
RMSE	1.318885293	1.17016808	1.148334067	0.855392092
MAE	0.93125	0.83496656	0.840084218	12.31034205
MAPE	0.439235792	0.440407215	0.433434446	0.230347863

Notice that even though the difference in the RMSE for training and validation data was significant, the MLR was predicting the data very well and hence was a good choice to select. Refer Appendix-1 for plot of all models fitted to the original dataset. Refer Appendix-2 for the time plot of residuals for MLR.



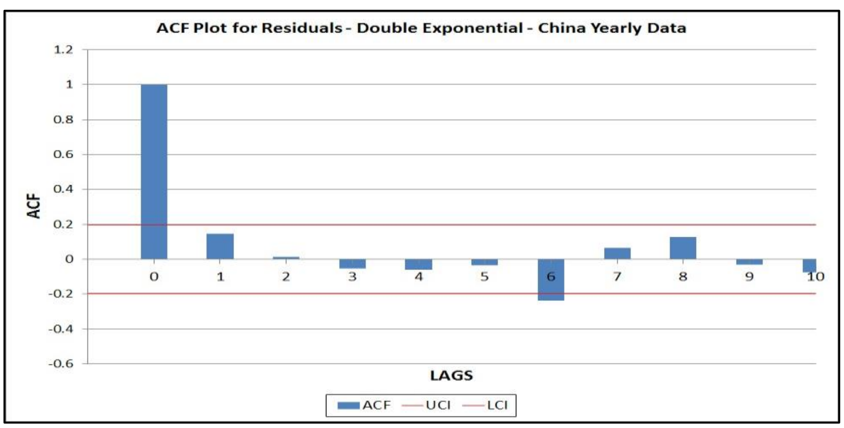




For yearly data with third degree polynomial trend and no seasonality we found that double exponential method produced the lowest Mape. Hence it was used to prepare the five year forecast.

CHINA YEARLY DATA MODEL SUMMARY

	Naïve	MLR	Double Exponential
Training			
MSE	0.110532522	0.066041852	0.079011753
RMSE	0.332464316	0.256986092	0.281090294
MAE	0.280136525	0.207182937	0.22977371
MAPE	0.040578212	0.030163702	0.033262614
Validation			
MSE	0.488730201	0.566910314	0.172235386
RMSE	0.699092413	0.752934468	0.415012513
MAE	0.655116667	0.673434571	0.335237588
MAPE	0.088855515	0.091701632	0.045998752



Conclusion and Recommendation

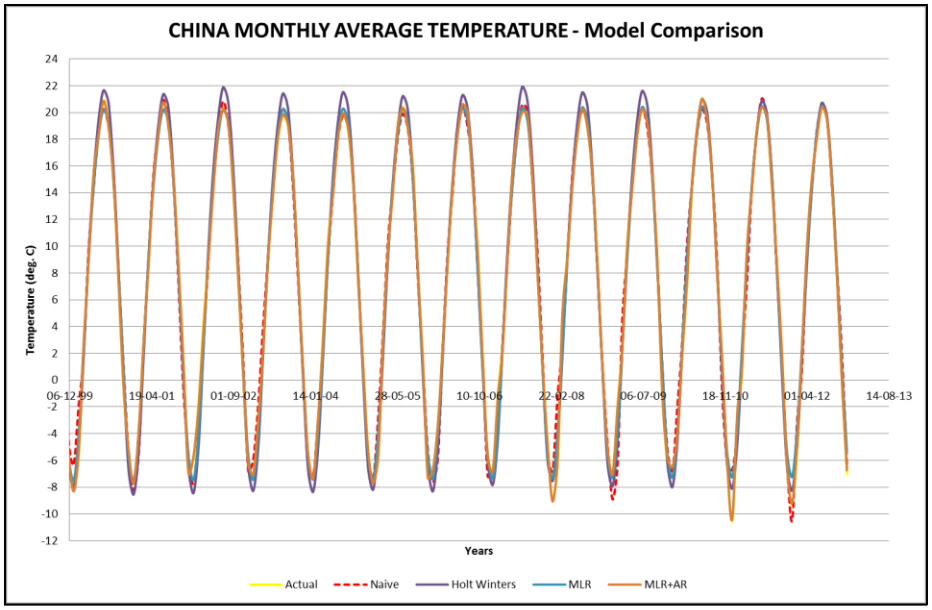
With the 24 month and 5-yearly forecasts provided to UNEP, we would recommend the following:

- To test the validity of the Special Sustainability Programs, UNEP should create test and control groups with test group following the Sustainability program and control group not following any of the measures. UNEP can then compare the average temperatures for test and control groups against forecasted temperatures and analyze how effective the Special Sustainability programs are.
- Assuming that the Special sustainability programs are successful, UNEP should keep adding the latest temperature data to the model and revise the forecasting model if required, so that the model becomes more robust and can forecast temperatures with more accuracy, thereby helping UNEP to identify other countries which should be part of the Special programs.
- Because the model doesn't take any external factors into consideration, UNEP could enhance the current model by identifying and adding factors that have strong correlation with average temperatures.
- Given the CO2 emissions have a strong correlation with the increasing temperatures, UNEP should introduce Carbon tax worldwide.
- A carbon tax is a fee for making users of fossil fuels pay for climate damage their fuel use imposes by releasing carbon dioxide into the atmosphere, and for motivating switches to clean energy.
- UNEP could also collaborate with agencies such as Energy Coordinating Agency (ECA) which specialize in providing infrastructural support to curb temperature increases.

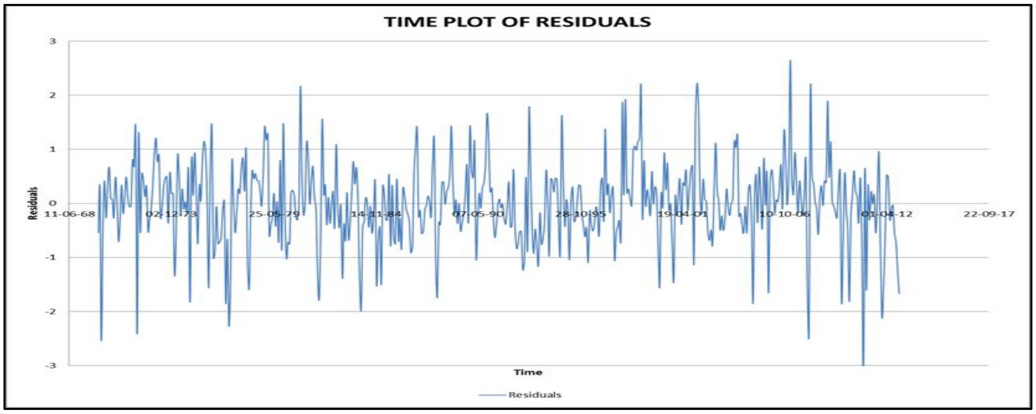
UNEP

Appendices

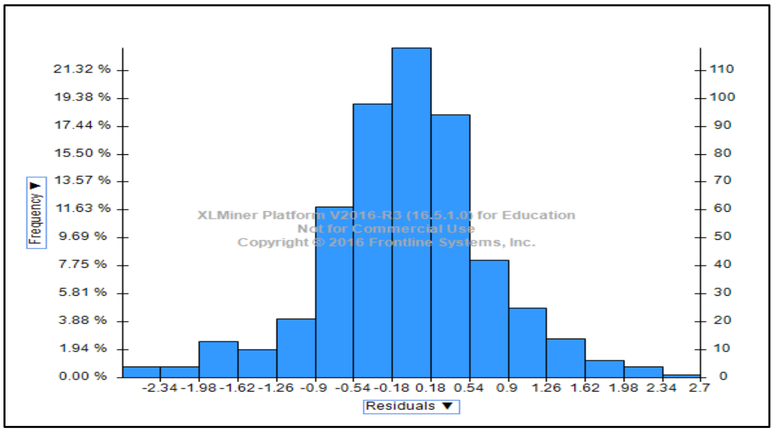
Appendix-1: Model Comparison for China



Appendix-2: Time-Plot for Monthly Residuals (MLR)



Appendix-3: Histogram for Residuals (MLR)

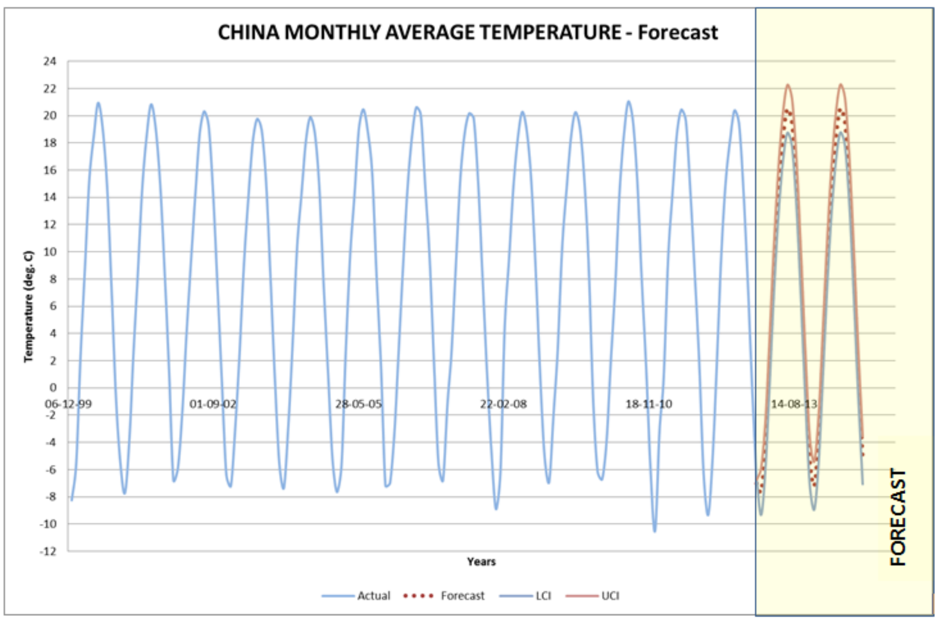


UNEP

Table-1: Monthly Forecast – China

Month	Predicted Value	95% Confidence Intervals		Error (AR1)	f % Confidence Level		Final Forecast	LCI	UCI
		Lower	Upper		Lower	Upper			
01-01-13	-7.206912791	-7.467260263	-6.946565319	-0.390615378	-1.852770957	1.071540201	-7.597528169	-9.320031219	-5.875025118
01-02-13	-3.681215116	-3.941562588	-3.420867644	-0.152440413	-1.643951794	1.339070967	-3.833655529	-5.585514382	-2.081796677
01-03-13	2.324366279	2.064018807	2.584713751	-0.04795706	-1.544605693	1.448691572	2.276409219	0.519413114	4.033405323
01-04-13	9.235645349	8.975297877	9.495992821	-0.016392792	-1.51352277	1.480737186	9.219252557	7.461775107	10.97673001
01-05-13	14.62976163	14.36941416	14.8901091	-0.005420054	-1.502607508	1.4917674	14.62434157	12.86680665	16.3818765
01-06-13	18.55701744	18.29666997	18.81736491	-0.001815767	-1.499009451	1.495377916	18.55520167	16.79766052	20.31274283
01-07-13	20.50871512	20.24836764	20.76906259	-0.000605132	-1.497799517	1.496589254	20.50810998	18.75056813	22.26565184
01-08-13	19.51759884	19.25725137	19.77794631	-0.000202087	-1.49739655	1.496992376	19.51739675	17.75985482	21.27493869
01-09-13	15.06259884	14.80225137	15.32294631	-6.74326E-05	-1.497261904	1.497127039	15.0625314	13.30498946	16.82007335
01-10-13	8.428296512	8.16794904	8.688643983	-2.25083E-05	-1.497216981	1.497171964	8.428274003	6.670732059	10.18581595
01-11-13	0.99969186	0.739344389	1.260039332	-7.51207E-06	-1.497201985	1.497186961	0.999684348	-0.757857596	2.757226293
01-12-13	-5.327982558	-5.58833003	-5.067635086	-2.50726E-06	-1.49719698	1.497191966	-5.327985065	-7.08552701	-3.570443121
01-01-14	-7.18277352	-7.445620918	-6.919926122	-8.36815E-07	-1.49719531	1.497193636	-7.182774357	-8.942816228	-5.422732486
01-02-14	-3.657075846	-3.919923244	-3.394228447	-2.79295E-07	-1.497194752	1.497194194	-3.657076125	-5.417117996	-1.897034254
01-03-14	2.34850555	2.085658151	2.611352948	-9.32172E-08	-1.497194566	1.49719438	2.348505456	0.588463585	4.108547328
01-04-14	9.259784619	8.996937221	9.522632018	-3.11121E-08	-1.497194504	1.497194442	9.259784588	7.499742717	11.01982646
01-05-14	14.6539009	14.3910535	14.9167483	-1.03839E-08	-1.497194483	1.497194463	14.65390089	12.89385902	16.41394276

Appendix-4: Monthly Forecast - China



Appendix-5: Time-Plot for Yearly Residuals (Double Exponential)

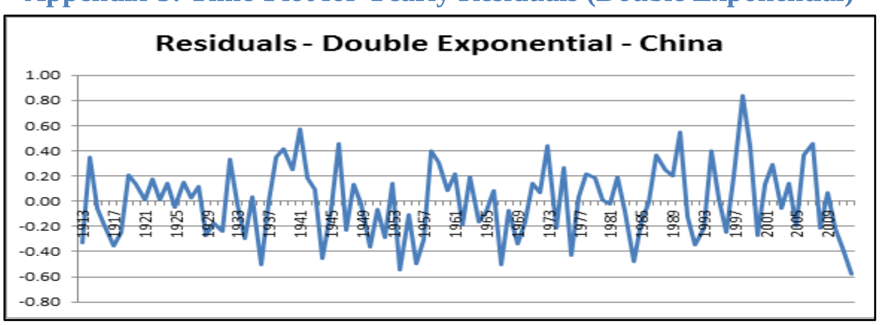
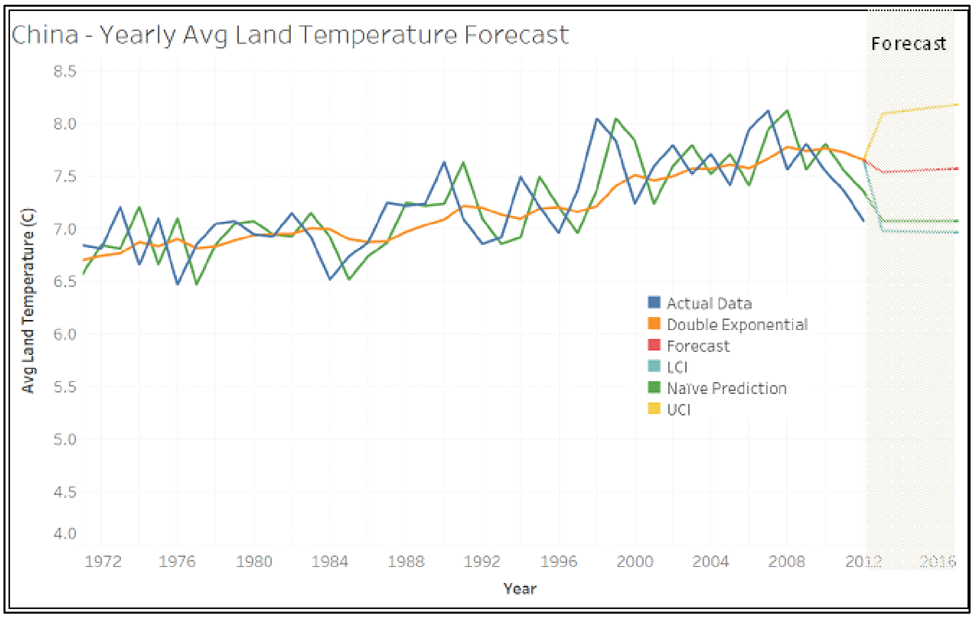


Table-2: Yearly Forecast - China

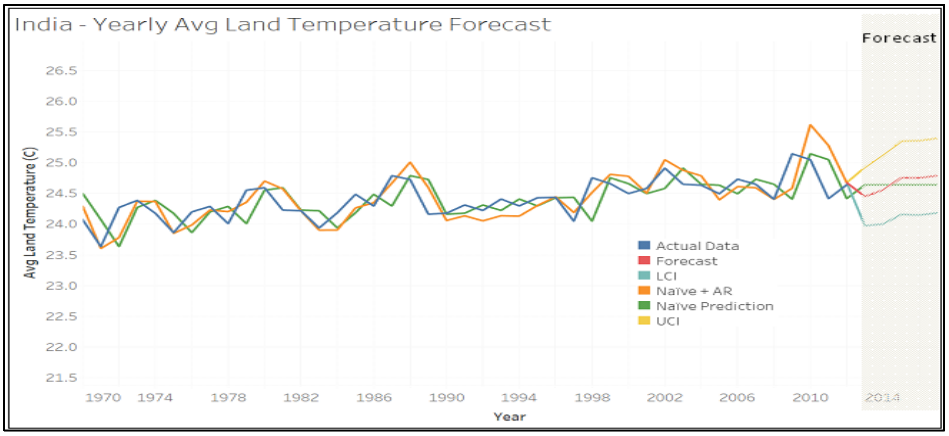
Year	Forecast	LCI	UCI
2013	7.53885641	6.981666999	8.096045821
2014	7.548172351	6.977824342	8.11852036
2015	7.557488292	6.974263752	8.140712832
2016	7.566804233	6.970966933	8.162641533
2017	7.576120174	6.967917466	8.184322881

Appendix-6: Yearly Forecast - China

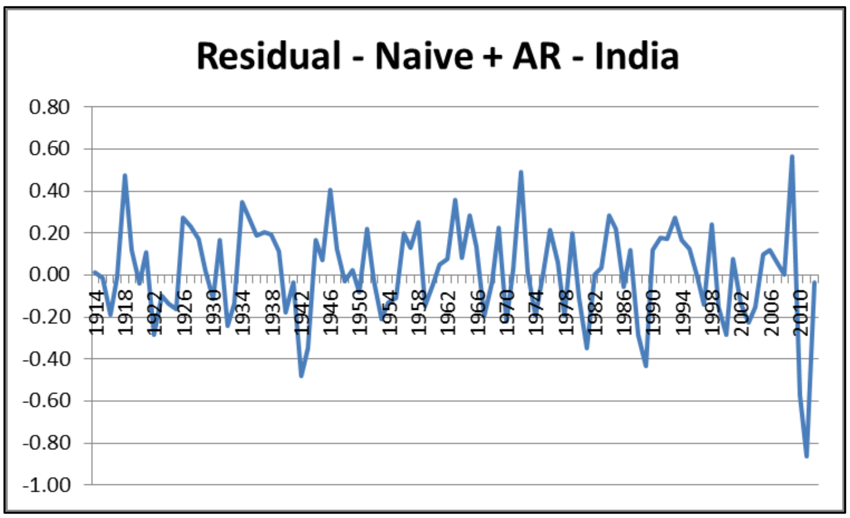


Test Case – India:

Appendix-7: Yearly Forecast - India



Appendix-8: Time-Plot for Yearly Residuals (Naïve+AR)



References:

1. S. (n.d.). Global Analysis - Annual 2016. Retrieved February 09, 2017, from https://www.ncdc.noaa.gov/sotc/global/201613
2. Global Warming Timeline. (n.d.). Retrieved February 09, 2017, from http://history.aip.org/climate/timeline.htm
3. UN Environment. (n.d.). Retrieved February 09, 2017, from http://www.unep.org/