Toronto Weather Forecasting Using Time Series Techniques

Hrithik Kumar Advani
1005570035 || kumaradv
hrithik.kumaradvani1@mail.utoronto.ca
University of Toronto Mississauga

ABSTRACT

Time series analysis and weather forecasting are very important and strong tools in today's world. Many of the activities around the world rely heavily on the weather conditions of that particular place. Here ARMA models with trend added have been used to model Toronto Temperature dataset, which includes data for past three years. It was also used to forecast the future tempearaure and it was compare with the test data values. After choosing the best model, one step ahead forecast was done and it was close to the test data value giving a good fit for the model.

KEYWORDS

model, forecast, ARMA, residuals, time series

1 INTRODUCTION

Weather Prediction is the application of science and technology to predict atmospheric conditions ahead of time for a particular region [2].The mounting losses of property and income which result from extreme weather events, the ever increasing use of the atmosphere as a transportation route, man's ability to modify his atmospheric environment (deliberately or inadvertently), and the ever present control which the atmosphere (the rain, the snow, the wind, the humidity, and the sunshine) exerts over our "economy",are all significant in today's society [1]. The predictions of the atmospheric scientist, and in particular those of the weather forecaster, are, therefore, of considerable economic and social value [1].Weather forecast plays an important role in economic sector as well. For example, It is important for the farmers in agriculture sector to have a good idea of the weather forecast so that the farmer can plan the growth of the plants accordingly.

It even becomes more important to understand the weather patterns of the big and cold cities, like Toronto, Canada.There are number of reasons for the importance of weather forecasting of Toronto. In winter, It can get really cold in Toronto. Not only that, but at times, Toronto experiences extreme weather condition(snow storms, freezing rain).Emergency alert and warning systems are brought into place. Economic activities are disturbed, universities and schools are closed and much more. Knowing the weather forecast especially in the winter becomes really important for the citizens of the Toronto to take precautionary measures while going out of their homes. This certainly becomes important for the strategy makers too. They have to plan ahead for the safety of the people and their well being. In recent years, there have been improvements of the techniques for the weather forecasting an its awareness to the people. For example Canadaian Metrological Service has provided weather forecasts to the people. It seems that it is possible to forecast Toronto's temperature looking at the accuracy of the prediction of the metrological department, though not many studies have been done on this topic.

In this study, we aim to study the daily tempratures of Toronto for last three years. In particular we will be looking at the data from mid of March 2018 till mid of March 2021(1095 observations). We would like to model this data using time series techniques and forecast the temperature for next few days. Forecast results will be further analysed.

2 MODEL SPECIFICATION

Box-Jenkins frame work was used to arrive at the best model. The dataset was divided in to training and testing dataset with 90 to 10 percent ratio. The test data was taken from the end of the original dataset so that there are no missing values in the data. Figure 1 below shows the time series plot for the Toronto temperature dataset for the training data. There seems to be some periodic behaviour in the data. It would be interesting to see whether this periodic behaviour leads to some kind of seasonal behaviour or any trend behaviour.

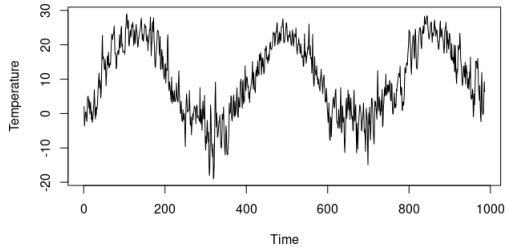

Figure 1: Time series plot for Toronto Temparture dataset

Figure 2 below shows the Sample Autocorreltaion function plot(ACF) of the data at different lags and Figure 3 shows the Sample Partial Autocorreltaion function plot(PACF) of the data at different lags. ACF values values die out rather slowly giving an indication of the non-stationary of the time series of the data. PACF plot also shows few significant(at 95% confidence level) values. Augmented Dickey-Fuller test was performed as a quantitaive measure to check for stationarity of the time series. p-value of 0.543 suggest that there is enough evidence to conclude the presence of stationarity in the data. To further investigate the level and trend stationarity in the data, KPSS test was run to account for such behaviour. Small p-values of 0.02 for level stationarity and 0.01 for trend stationarity suggests that there is not enough evidence to reach the conclusion that the data has either level or trend stationarity present.

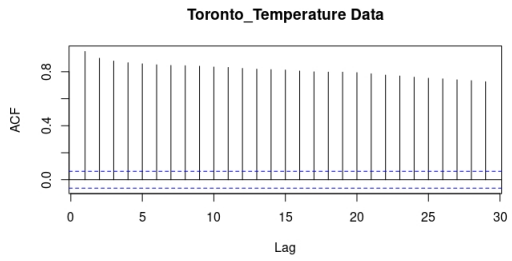

Figure 2: ACF plot for the data

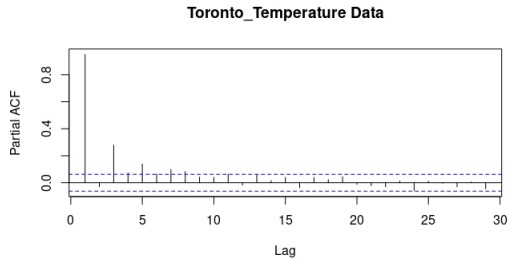

Figure 3: PACF plot for the data

After further investigation, It was found that cosine trend was present in the data. Cosine trend was removed from the data to make it stationary. Stationary model was used for further analysis. ARMA modelling technique was used in search of the best model. Several candidates were further checked.

3 FITTING AND DIAGNOSTICS

EACF was computed on 5x5 grid. Later on, the goodness of best models were evaluated using AIC and BIC criterion. Residuals plots were used to check the correlation between the residuals and QQ-plot and Shapiro test were conducted to see the normality of residuals. Those all were satisfied. Model with minimum harmonic mean error on the training data was considered to be the best model.

4 FORECASTING

As mentioned earlier, test data was used for forecasting purposes.Residual mean square error of 3.1 was achieved for the best model. Forecasted temperature of around -2.7 was predicted by the best model for one step ahead forecast. this forecast temperature is the one after the trend was added to the model.

5 DISCUSSION

As the original time series was non stationary, it was detrended to get the stationary time series. Analysis were carried out on the stationary time series to get the most appropriate model. After getting the most appropriate model, it was used for the forecasting propose of one step ahead forecast. The model was able to predict values closer to the test data values hence even graphically we can consider that model to be useful.

6 CONCLUSION

This study was conducted on the past temperature data and the forecasted values would not be beneficial currently. But these values can be used for future purpose. These can be checked against the observed values of the coming year on same date and can be compared to check for the similarities and differences in the weather data. These techniques and modelling from this study can be used to produce similar models which can be used to predict missing temperature values in future. This can also help policy and strategy makers for better planning.

REFERENCES

[1] W. J. Maunder. 1969. The consumer and the weather forecast. *Atmosphere* 7, 1 (1969), 15–22. https://doi.org/10.1080/00046973.1969.9676566 arXiv:https://doi.org/10.1080/00046973.1969.9676566

[2] ZhanJie Wang and A B M Mujib. 2017. The Weather Forecast Using Data Mining Research Based on Cloud Computing. *Journal of Physics: Conference Series* 910 (10 2017), 012020. https://doi.org/10.1088/1742-6596/910/1/012020