**WIND PREDICTION BASED ON TIME SERIES ANALYSIS**

By

Adithya R Ganesh Abhishek Sundaresan Abhishek Pal Gautham Krishnan

**Time Series:**

A**time series** is a sequence of measurements of the same variable(s) made over time. Usually the measurements are made at evenly spaced times - for example, hourly, monthly or yearly.

**AR Model:**

An **autoregressive model** is when a value from a time series is regressed on previous values from that same time series.

For example,  yt on yt−1:     yt=β0+β1yt−1+ϵt. ------ (1)

In this regression model, the response variable in the previous time period has become the predictor and the errors have the usual assumptions about errors in a simple linear regression model.

**Order of an AR Model:**

The **order** of an autoregression is the number of immediately preceding values in the series that are used to predict the value at the present time. So, equation (1) is a first-order autoregression, written as AR(1).

Similarly:                           yt=β0+β1yt−1+β2yt−2+ϵt.   --------(2)

This model is a second-order autoregression, written as AR(2), since the value at time t is predicted from the values at times t−1 and t−2.

More generally, a kth-order autoregression, written as AR(*k*), is a multiple linear regression in which the value of the series at any time *t* is a (linear) function of the values at times t−1,t−2,…,t−k.

**Autocorrelation and Partial Autocorrelation**

The coefficient of correlation between two values in a time series is called the **autocorrelation function** (**ACF**).

For example the ACF for a time series yt is given by:      Correlation(yt,yt−k).

This value of *k* is the time gap being considered and is called the **lag**. A **lag 1** autocorrelation (i.e., *k*= 1 in the above) is the correlation between values that are one time period apart. More generally, a **lag *k*** autocorrelation is the correlation between values that are *k* time periods apart.

The ACF is a way to measure the linear relationship between an observation at time *t* and the observations at previous times. If we assume an AR(*k*) model, then we may wish to only measure the association between yt and yt−k and filter out the linear influence of the random variables that lie in between (i.e., yt−1,yt−2,…,yt−(k−1)), which requires a transformation on the time series. Then by calculating the correlation of the transformed time series we obtain the **partial autocorrelation function** (**PACF**).

The PACF is most useful for identifying the order of an autoregressive model. Specifically, sample partial autocorrelations that are significantly different from 0 indicate lagged terms of y that are useful predictors of yt.

Once the auto regression model was implemented, the given dataset was cleaned to separate the 3 provided attributes. Later, the data was split into train and test. Training was done on 90% of the dataset and tested on the remaining 10%.

**Graphical Approach:**

Graphical approaches to assessing the lag of an autoregressive model include looking at the ACF and PACF values versus the lag.

In a plot of ACF versus the lag, if we see large ACF values and a non-random pattern, then likely the values are serially correlated.

In a plot of PACF versus the lag, the pattern will usually appear random, but large PACF values at a given lag indicate this value as a possible choice for the order of an autoregressive model. It is important that the choice of the order makes sense.

The following were the plots obtained when testing on the 3 given parameters.

The blue plot below shows the differential of the actual values present in the dataset and the red plot depicts the predicted values.

It can be seen that the first plot shows the change in wind speed. A mean square error of 0.070 was obtained for this case.

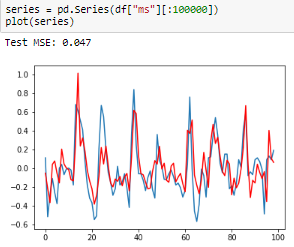


Figure 1 Wind speed prediction plot

In the second plot due to higher variation in wind direction, the prediction of wind direction at a particular time becomes harder. When the plots were made on the tested values, the prediction obtained had lower accuracy however it is noticed that the change in wind direction approximately follows the same behaviour as the actual values.

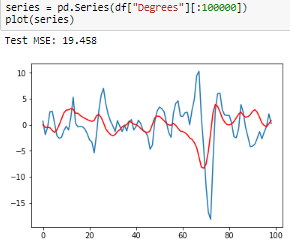


Figure 2 Wind direction prediction plot

The final plot depicts the change in temperature across the period of time. The model learnt across the years and was successful in predicting what the temperature may be at a point of time.

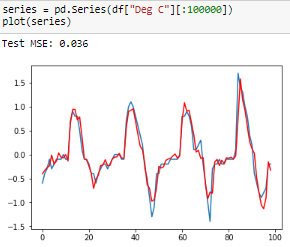


Figure 3 Temperature prediction plot

**Conclusion**

A model was successfully developed for predicting the given parameters. Auto Regression gave feasible results but for even better accuracy, deep learning techniques such as RNN are being implemented along with ARIMA model.