# **MTech KE Unit 7 Time Series Forecasting**

Bharat Nagaraju (A0178258N)

Vigneshram Andiappan Selvaraj (A0178215A)

# **Index**

1. Introduction
2. Analysis
3. Conclusion

# **Introduction**

In this exercise we will be using different techniques of Time series forecasting to predict future oil price based on previously observed values. The data set used here is oil price which is spread over a time period from Jan 1986 to Jan 2006. It is discrete, and the interval between each point is constant. As per the exercise we will try different techniques and see which model suits the best for this dataset.

Time Series Forecasting methods

Before we choose the model we need to analyse a few characteristics of the data and then choose an appropriate model.

**Trend**: A long-term increase or decrease in the data. This can be seen as a slope (it doesn’t have to be linear) roughly going through the data.

**Seasonality**: When the data is affected by seasonal factors such as (hour of day, week, month, year, etc.) then seasonality exist. Seasonality can be observed with nice cyclical patterns of fixed frequency.

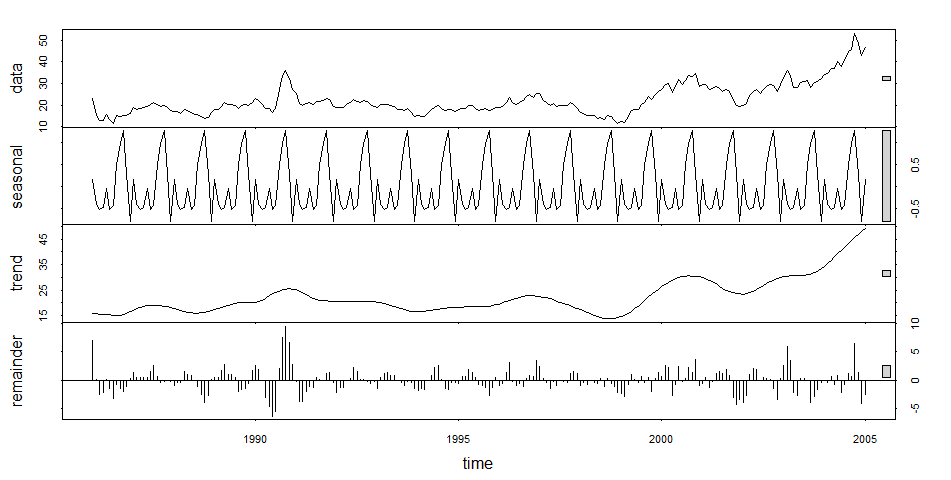
**Cyclicity** : When the data is characterized by rises and falls that are not of a fixed frequency then we can confirm the data has cycles. These fluctuations are usually due to economic conditions, and are often related to the “business cycle”. The duration of these fluctuations is usually tends to elongate for more than 2 years.

Decomposition of time series

Decomposition provides a useful abstract model for thinking about time series generally and for better understanding problems during time series analysis and forecasting. This is not used directly to do time series forecast but to analyse the data for time series forecast. Time series decomposition involves thinking of a series as a combination of level, trend, seasonality, and noise components. This is an important step needs to be done as perquisite to help in choosing the appropriate model. The below results show that data has trend but no seasonal.

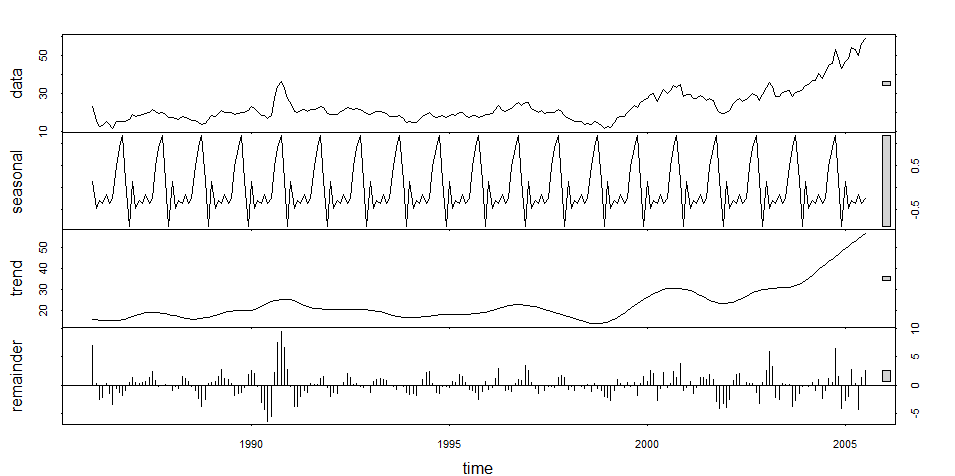
Below is analysis of data from Jan 1986 to Jan 2005.





Below is analysis of data from Jan 1986 to Jul 2005.





Then we did the Dicker Fuller test to check for data stationary. The tests produced p value of 0.7897 which should be below 0.05 for stationary. This obviously means we cannot use the data as it is in the ARIMA model. Hence we need to do some transformation of data which is differencing it (making it stationary) before using for ARIMA model.

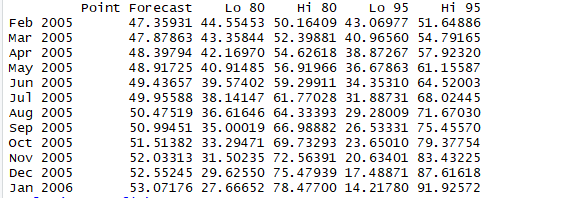
Exponential Smoothening (Holt-Winters)

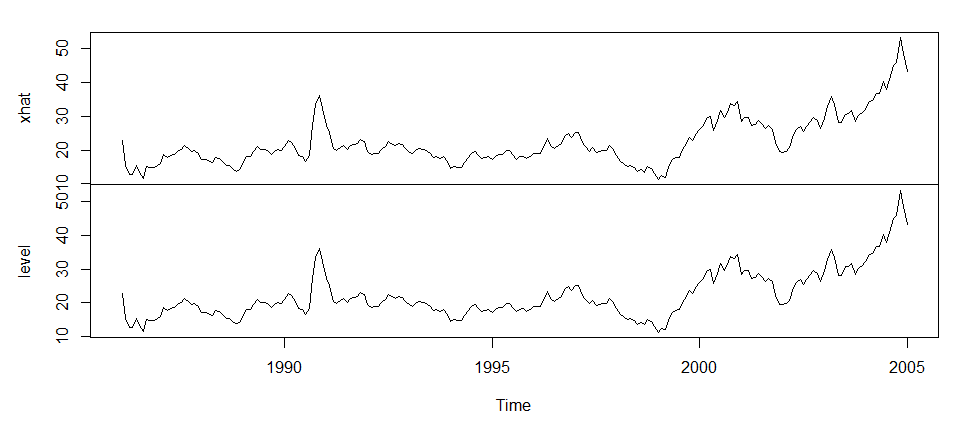
Exponential smoothing is a time series forecasting method for univariate data that can be extended to support data with a systematic trend or seasonal component. Forecasts produced using exponential smoothing methods are weighted averages of past observations, with the weights decaying exponentially as the observations get older. In other words, the more recent the observation the higher the associated weight. For oil price as we see there is some drastic increase in the recent years and hence we wanted to give more weightage to the recent years than the past. This is the reason why we had chosen a big value for alpha.

12 months Prediction

Accuracy

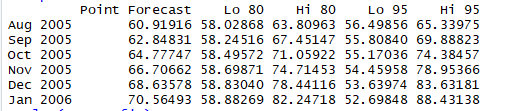


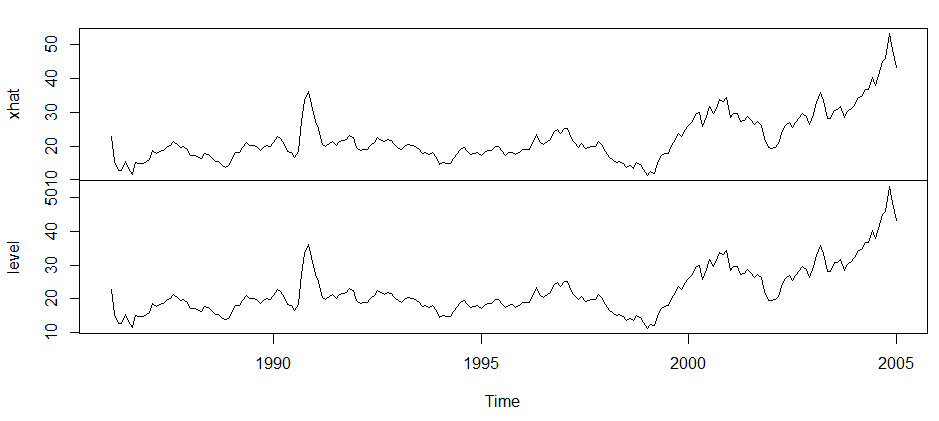




6 months Prediction

Accuracy





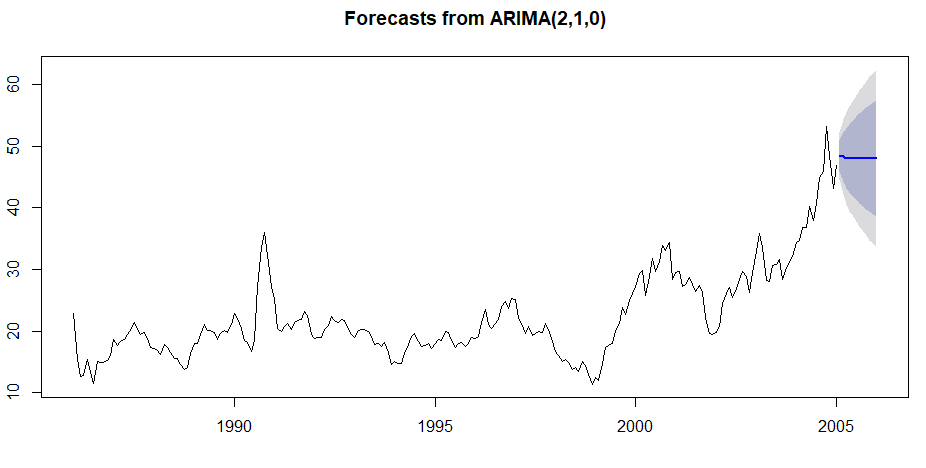
ARIMA

ARIMA, short for ‘Auto Regressive Integrated Moving Average’ is a class of models that explains a given time series based on its own past values, that is, its own lags and the lagged forecast errors, so that equation can be used to forecast future values.

12 months

Accuracy

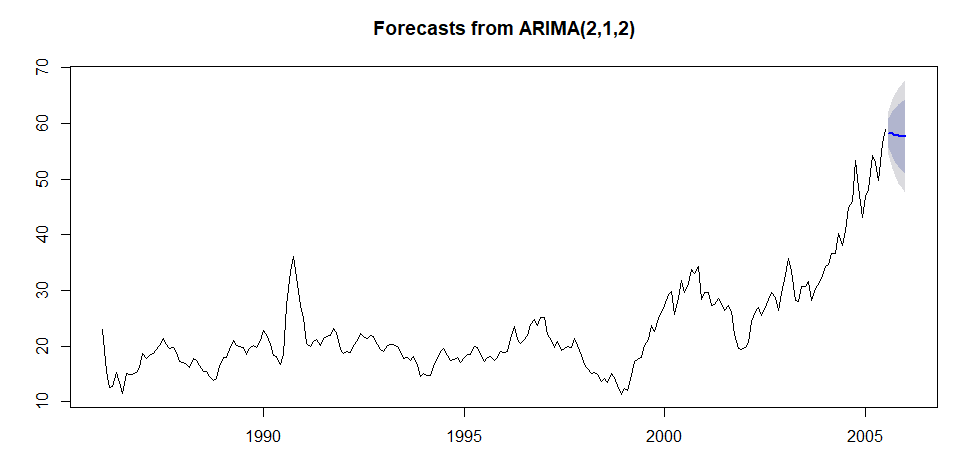




6 months Prediction

Accuracy





Logistic Regression

This method cannot be used for this dataset as we would need at least one more variable added to the time to make the logistic regression work.

**Comparison of Models for 12 months prediction**

As it can be seen ARIMA model after doing the first order differencing predicted the values slightly more accurate than the Exponential Smoothing. Also it can be seen that the accuracy of the exponential smoothing increased after more weightage was given to the recent values. But this can cause wrong prediction of future prices in cases where the recent values are either just abnormal spikes in the data.

**Comparison of Models for 6 months prediction**

It can be seen that Exponential Smoothing model gives much better result than the first order differenced ARIMA model. Predicted the values slightly more accurate than the Exponential Smoothing. But it is obvious that the accuracy of the exponential smoothing increased after more weightage was given to the recent values and when it is a short-term prediction. In cases of shorter period, exponential smoothing can be used to give better results.

**Models difference for 6 and 12 month prediction.**

For the comparison of the models on the 6 months prediction, though both Exponential Smoothing as well as ARIMA are able to achieve better results, Exponential Smoothing is better. Though there should be no seasonality in data for Exponential Smoothing still it performs better when the right weightage is given to the past data based on the analysis. The disadvantage with Holt’s Linear trend method is that the trend is constant in the future, increasing or decreasing indefinitely. For long forecast horizons, this can be problematic but for short term predictions it is better.