Modelling and Predicting the Number of Airline Passengers

Assignment 3 – 02417 Time Series Analysis – Anders Hørsted (s082382)

In this report a model of the monthly number of airline passengers in the U.S. is build. The data set used to build the model is the actual number of passengers for every month between January 1995 and March 2002. To be able to give an estimate of the precision of the model, the data set is separated into two parts: A training set containing data for the period January 1995 to June 2001, and a test set containing data for the period July 2001 to March 2002. Throughout the modelling phase we work as if only the training set is available. The test set is used when the final model have been build, to compare the predictions of the final model, and the actual numbers in the test set. From now and until the section about measuring the model performance the training set is just referred to as "the data set" or "the data"

Data exploration

In this section the data set is introduced. First a plot of the data is created and shown in figure 1.

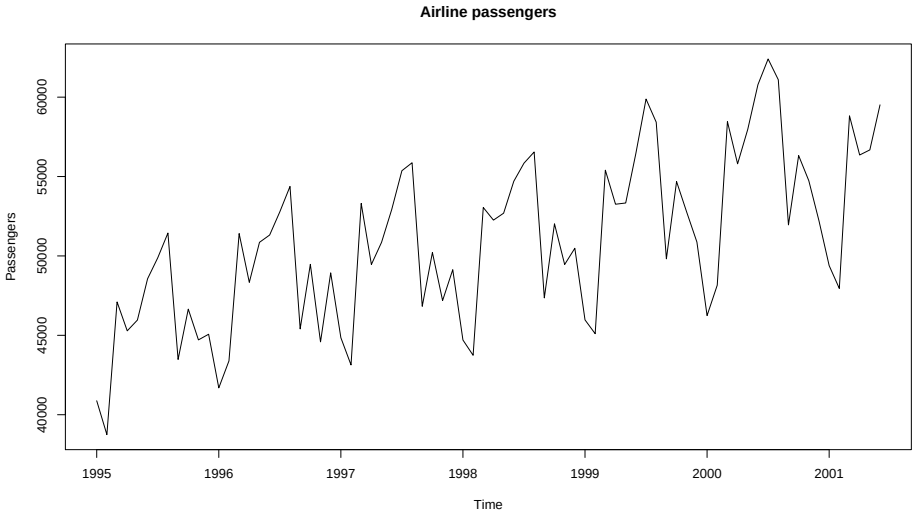

Figure 1: Plot of data set used for modelling

From the plot a general upward trend is recognized. This isn't surprising since the U.S. economy got stronger (REFERENCE!!!) in the period and therefore more airline

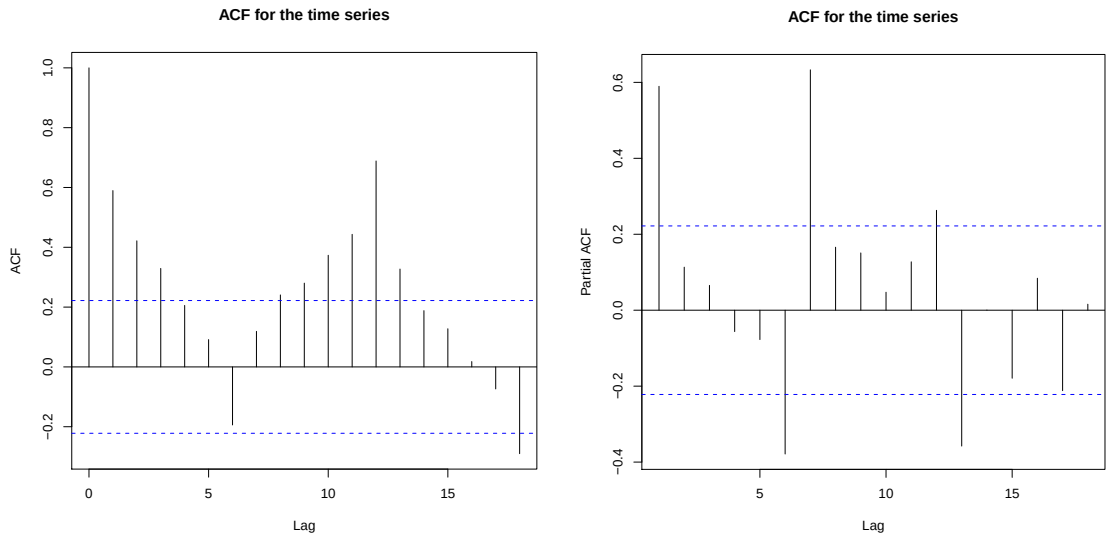

Figure 2: CAPTION!!!

passengers should be expected. Also a regular seasonal pattern can be seen which isn't surprising either. In the summer months e.g. we would expect more passengers than for the other months etc. The conclusion of these two observations is that the time series is non-stationary and this is something that should be coped with during the modelling phase. To support that the time series is non-stationary the estimated autocorrelation function and partial autocorrelation function are now plotted (see figure **??**)

Appendices

All R code used for this assignment is included here. All source code incl. latex code for this report can be found at `https://github.com/alphabits/dtu-fall-2011/tree/master/02417/assignment-3`

References

[1] Henrik Madsen, *Time Series Analysis*. Chapman & Hall/CRC, 1st Edition, 2008.