Project
SF2943 Time Series Analysis

Sargon Danho sargond@kth.se
Xing Guan xguan@kth.se
Edward Sikorski sikorski@kth.se
Shadman Ahmad shadmana@kth.se
Keziban Colak, keziban@kth.se

May 16, 2018

1 Analysis of a non-financial time series: Milk Production Time Series Analysis

1.1 Introduction

This report will analyze a data set of milk production, through an extensive time series analysis. Stationarity, trend and seasonality are three factors that will be analyzed to see if a model can be built to forecast future milk production. The dataset contains nearly two decades of monthly dairy production from the state of California, between the years 1995 to 2013 [1]. For further analyze we divided 68% of the data to a training set and 32% to a test set.

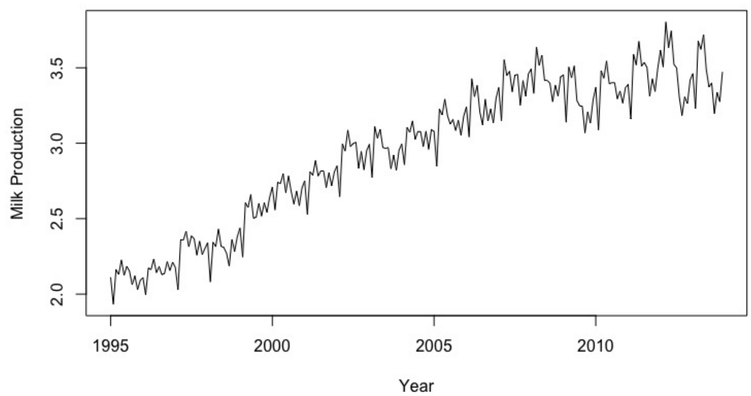


Figure 1: Monthly milk production in billion lbs.

1.2 Analysis of stationarity

The procedure of the time series analysis in this paper is based on a general approach. First we plot the series to check whether there are trend, seasonality, sharp changes in behavior or any outliers. By looking at the plot of the dataset, see fig 1, we can state that our dataset is non-stationary and has *trend and likely seasonality* as well.

A time series is classified as stationary series if the mean is zero and if the covariance does not change over time. In order to build a model, the time series must be stationary. When the stationary criterion is fulfilled, a stochastic model can be build up for prediction of the time series. If the criterion is violated, stationarity can be achieved by removing trend and seasonality of the data. As well as differencing.

There exist various tests and tools that can be used to check stationarity, but in this report the Augmented Dickey-Fuller test was used. The ADF-test is a unit root test. It can be tested by null hypothesis H_0 which state that it exist a unit root hence the data is not stationary w.r.t trend and seasonality. The data got a $p - value = 0.602$, which says that we can not reject the null hypotheses thus our data is not stationary. Applying ADF on the remainder of the training dataset, the ADF value drops down to -5.70 and the p value was equal to 0.01 which indicate the data is stationary. A similar ADF-test for differencing was done by using the R function "ndiffs" on the original data, which showed that one differencing was needed to obtain stationarity ($d = 1$).

To determine the degree dependence in the observed data and to choose a model, the sample autocorrelation function (sample ACF) can be used. The sample ACF is one of the important tools and is used to determine the lags of the model. Another tool that is useful when choosing a model is the partial autocorrelation function (PACF) that is the autocorrelation between two points after removing the linear dependence on data. By varying the moving average and autoregressive degrees q and p, respectively, the number of AR and MA lags are determined [2]. The following figure shows the decomposition of the training set into trend, seasonal and remainder components.

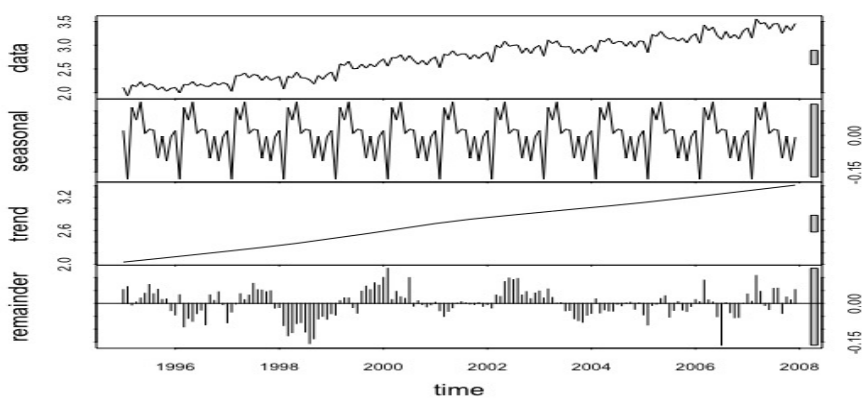


Figure 2: Seasonality, trend and remainder from the dataset.

The last part of the figure is the remainder that is isolated in order to analyse the data. Afterward, the ACF and PACF plots of the remainder data is plotted shown in the figures below.

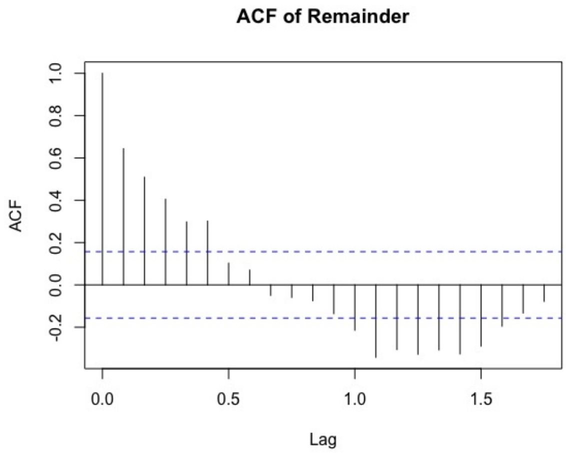


Figure 3: Autocorrelation function of training.

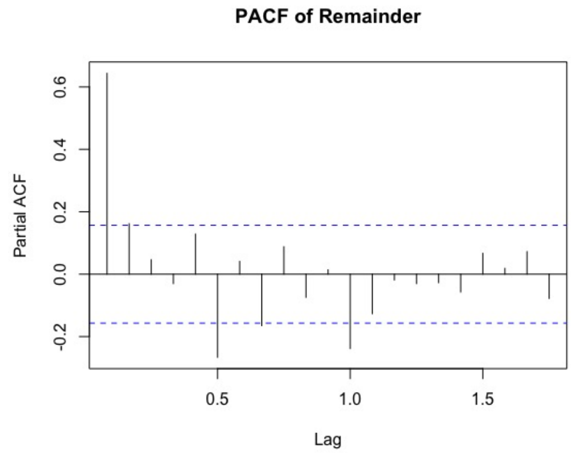


Figure 4: Partial autocorrelation function of training.

1.3 Model Structure and Building

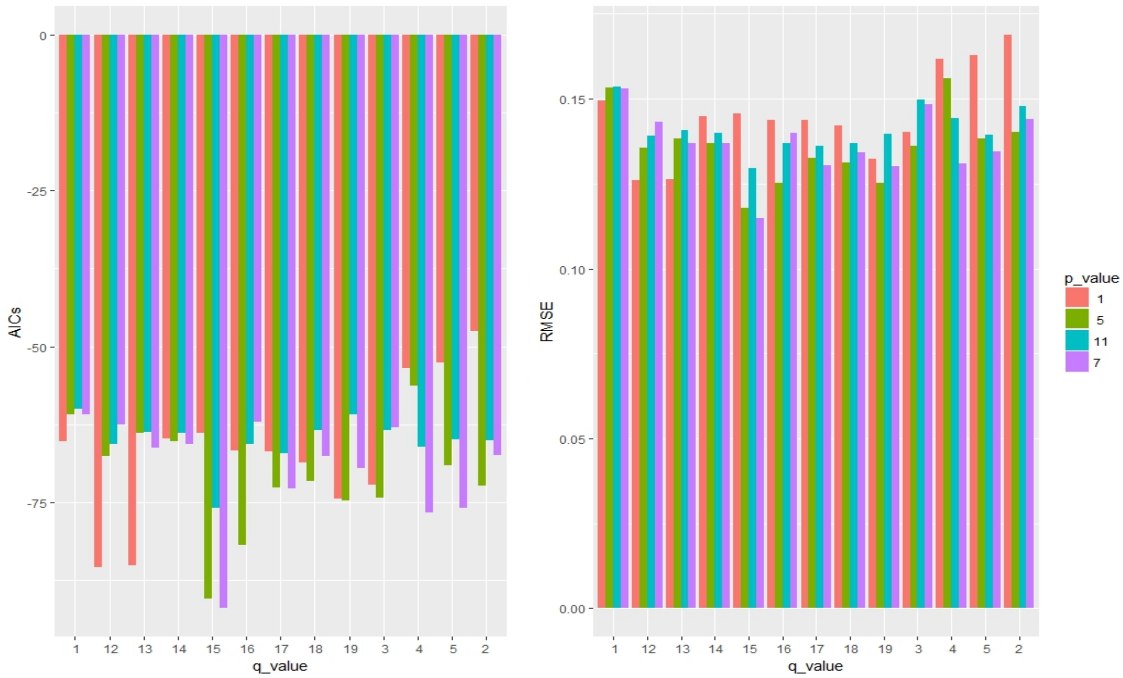
ARIMA, short for Autoregressive Integrated Moving Average, is a general approach of ARMA. In this report, a non-seasonal ARIMA model is used, which is denoted as *ARIMA(p,d,q)*. The *p-value* indicates the number of lags in autoregressive model, the *d-value* is the degree of differencing and the *q-value* is the order of the moving average model [2].

We proceeded to our model building by looking at the ACF fig.3 and PACF fig.4 plots. We can see that we have correlation and picked the lags with values above the significance level $\pm 1,96/\sqrt{n}$, as p and q values.

1.4 Model Selection and Validation

For the next part, we wanted to see which model was most suitable to fit our data. We created different ARIMA models with all of the combinations of the extracted p and q- values and tested these on our test data set. As well as $d = 1$, we saw that we got an impaired performance if we excluded the differencing. Afterwards, we compared 2 types of measurements from each model, AIC and RMSE.

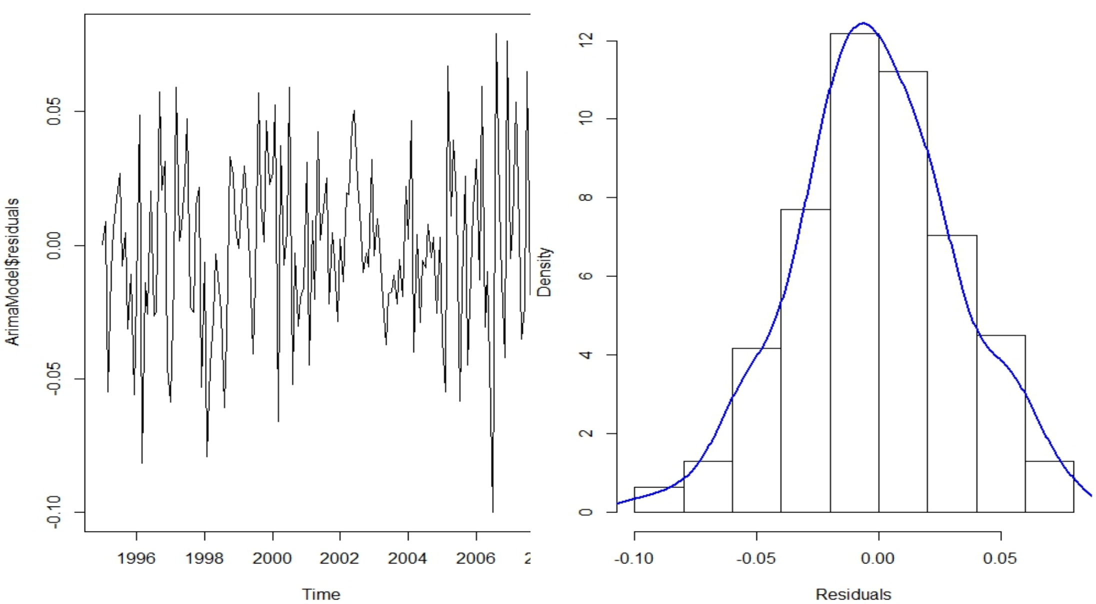
The basic notion of AIC (Akaike Information Criteria), is that by continually increasing the parameters to a model will give a better fit. But this gives rise to a problem called over-fitting which consequently make the model unnecessarily complex about the real underlying pattern. By looking at fig 5a we can see that the model ARIMA(7,1,15) has the lowest AIC value, thus is more appropriate for fitting in terms of containing information and not over-fit the data. Since, AIC is mostly used for model comparison and not estimating the absolute quality of the model itself. We then looked at the standard RMSE (root mean squared error) values from each model on the test set, to increase our confidence on our model selection. By looking at figure 5b we can conclude that same model with the lowest AIC also has the lowest RMSE value.



(a) AIC vs model

(b) RMSE vs model

Figure 5: AIC and RMSE values on testset for different models, where a bar corresponds to a model with q- value from x-axis and a p- value with a color



(a) Residuals on each time period

(b) Histogram of the residuals

Figure 6: Plots of the residuals of the model ARIMA(7,1,15)

Further to evaluate the traits of the model on the given data set. We did a residual analysis, were some assumptions must be confirmed by the residuals of the model. Fig 6a shows some sort of a white noise and the residuals are scattered around the mean zero which is expected. However, it appear to be some type of heteroscedasticity especially with the large spike between the years 2006 and 2008. But looking at the normal probability plot of the residuals fig 6b, we can clearly see that the residuals are normally distributed, which is another important assumption from the residual analysis.

1.5 Conclusions

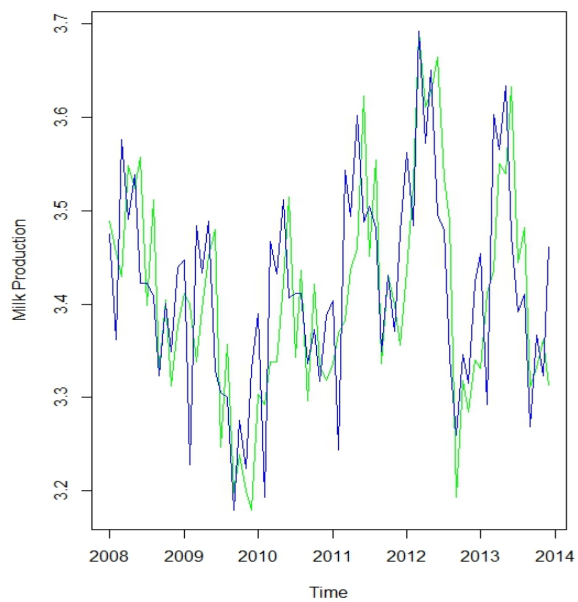


Figure 7: Plot of the milk production from test data. The blue is the original test data and the green is the estimation

The model gave an RMSE of 0.1147432 and an AIC value of -91.95 on the test set, and by the look of the similarities in fig 7, we can see that the model is a good fit for the data. Thus, an arima model with the parameters p = 7, d = 1, q = 15 is suitable to use when you want forecast this specific data containing quantitative production of milk.

1.6 References

[1] Brian Gould, Agricultural and Applied Economics, UW Madison,
URL : http://future.aae.wisc.edu/tab/production.html

[2] Brockwell, Peter J. and Davis, Richard A.: Introduction to Time Series and Forecasting, 3 ed, Springer, New York, 2016.

2 Discussion of the research article: Time-Series Analysis of Kentucky Coal Production

2.1 Summary

The research article focuses on forecasting and modeling of Kentucky coal production for 1983 by using time series data for years between 1976-1982. In order to simulate state level of monthly coal production, an autoregressive moving average (ARMA) model that satisfies all diagnostic requirements and estimations is applied to the monthly time series data. The autoregressive components in the ARMA model are of lag 1 and lag 12 and the moving average component is of lag 1. Comparison of model predictions for 1983 with actual monthly coal production in 1983 shows that the cyclical patterns were accurately simulated. In order to get short-range forecasting of different coal industry activities, similar models could be integrated into planning programs of state level. Compared with large scale econometric models, the ARMA model is very practical for short-term forecasting of coal production.

The state of Kentucky, as a leader of coal production in the United States, had a 16.7% share of total coal production in 1983 and therefore to federal energy policies, the future of Kentuckys coal industry was of vital importance. The energy crises of the 1970s were the reason to that the relationship between production and consumption from various sources was wanted to be modelled. The most of the developed models have been based on econometric analyses and in short range forecasting, they have not been useful.
A model specially well suited to short-term forecasting that is called as ARMA, autoregressive moving average, replicates past behaviour of a time series and for this reason an ARMA model was used for short-term forecasting of Kentucky coal production. The ARMA model used for the forecasting and the method for determining of the parameters were developed by Box and Jenkins in order to predict and control a time series. In order to determine moving average (MA) components, autocorrelations (ACFs) were analysed and partial autocorrelations (PACFs) to single out autoregressive (AR) components. By estimating and testing the parameters of a tentative model, the model could be diagnosed where the residuals from the model were sufficiently small. To estimate and evaluate different ARMA model, the Time-Series Program, BMDP, was used. Accordingly, a ARMA model that satisfied the stationarity requirement could be estimated and was used to get forecasting of Kentucky monthly coal production. By re-estimating the parameters and combining 1976-82 and 1983 data, the model could be enhanced. The updated model reflected improvements and satisfied all the requirements. Comparison between model forecast values and historical values for coal production showed that forecasts were accurate to within 5% of actual production for most of the months. The 1976-82 ARMA model overestimated 1983 coal production by 4% while the revised model provided better prediction.

The same methodology can be applied to any coal industry activity that fulfils the stationarity requirement. To enhance the effectiveness of state planning programs short-term forecasting can be incorporated with longer-term econometric models.

2.2 Strengths and Weaknesses

We have decided to critically examine the report out of the following aspects: relevance, originality, accuracy of analysis and presentation.

Relevance

The relevance of the article is high, as it shows that the effectiveness of the ARMA model. It is easy to use and does not consider important key economic indicators, even if its results are accurate. The article also provides other applications of ARMA modeling, indicating other areas of interest to conduct further research in.

Originality

The originality of the article at this moment of time is low, as there is an abundance of similar articles on short-term forecasting using an ARMA model. However, as the article was published in 1986 we searched in Google Scholar for the keywords forecast, ARMA model and short- term, dated up until 1985 and we still found a large number of articles using the ARMA model to forecast time series. We conclude that the the article was not ground-breaking for its time, mostly as the model had been created over 30 years earlier.

Accuracy of analysis

The model forecasted for most of the months, a number that deviated no more than 5.The forecasts are quite reasonable, especially considering that the model does not consider any economic indicators. It would have been interesting to see the results of another model on the time series data, to see if a more accurate result could have been attained.

Presentation

The presentation of the report was well done, with suitable tables and figures. The order of content was well placed and easy to follow, for example how the model was updated after it was shown that the initial results were not statistical significant. One weakness of the report however, we found, was that we would have liked more tables on the complete model. Only a result on RMSE was given, and we would might have liked more on stationarity.

3 Problem 3

Part A1

The Naive Approach - To model the daily return, we use a normal r.v. from a historical data sample. The data set consisted of 50 closing prices of trading. Setting that Z is standard normal, we model the next day's closing price as:

$$S_1 \overset{d}{\approx} S_0 * e^{\tilde{\sigma} * Z} \tag{1}$$

The 5% downside movement can be found by using the inverse of the cumulative distribution function of Z. The results are shown in the table below.

Date (t=0)	S_0	$S_1^{0.05}$	$\hat{\sigma}$	$log(\frac{S_1^{0.05}}{S_0})$
2018-01-11	1630	1613	0.65%	-1.07%
2018-02-09	1500	1477	0.94%	-1.55%

Table 1: k=1, Naive model

For a *GARCH* model, the return X_t is a model of the standard deviation multiplied with Z_t, which is a standard normal i.i.d. process. Also, the standard deviation can be seen as:

$$\sigma_t^2 = a_0 + \sum_{i=1}^{p} a_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{2}$$

where the values α and β are fitted to the two year historical data. To determine the optimal p and q values, we used BIC and found that (1,1) was the optimal fit. The results for this model is found below:

Date (t=0)	S_0	$S_1^{0.05}$	σ_0	$log(\frac{S_1^{0.05}}{S_0})$
2018-01-11	1630	1615	1.04%	-1.97%
2018-02-09	1500	1468	0.99%	-2.19%

Table 2: k=1, Garch(1,1) model

Part A2 As we model only five days ahead, the model for the fifth day's closing price is $S_1 \overset{d}{\approx} S_0 * e^{\tilde{\sigma} * Z}$ which is calculated yet again through the inverse of the cumulative distribution function of Z. The *GARCH* model is done similarly as in the previous part, and the results are as follows:

Date (t=0)	S_0	$S_5^{0.05}$	$\hat{\sigma}$	$log(\frac{S_5^{0.05}}{S_0})$
2018-01-11	1630	1592	0.65%	-2.40%
2018-02-09	1500	1449	0.94%	-3.46%

Table 3: k=5, Naive model

Date (t=0)	S_0	$S_5^{0.05}$	σ_0	$log(\frac{S_5^{0.05}}{S_0})$
2018-01-11	1630	1580	1.04%	-3.11%
2018-02-09	1500	1429	0.99%	-4.83%

Table 4: k=5, Garch(1,1) model

Part B As mentioned before, we used the $GARCH$ model. One main advantage of the method is that it is able to model in rapid increase in volatility much better than the naive method. In the figure below, it is visualized how the GARCH process takes the sharp increase in volatility that occurred in the financial markets in the beginning of February. Previous to that, in January, volatility was low as markets steadily increased, and one might argue that the use of the $GARCH$-model is limited in those times.