Zürcher Hochschule
für Angewandte Wissenschaften

School of
Engineering

Projektarbeit (Wirtschaftsingenieurswesen)

Allokationstool für Asset Management von Mobiliar

Autoren Pascal Simon Bühler
Philipp Rieser

Hauptbetreuung Prof. Dr. Marc Wildi

Industriepartner Mobiliar

Datum 19.12.2020

Zürcher Hochschule
für Angewandte Wissenschaften

School of
Engineering

Erklärung betreffend das selbständige Verfassen einer Bachelorarbeit an der School of Engineering

Mit der Abgabe dieser Bachelorarbeit versichert der/die Studierende, dass er/sie die Arbeit selbständig und ohne fremde Hilfe verfasst hat. (Bei Gruppenarbeiten gelten die Leistungen der übrigen Gruppenmitglieder nicht als fremde Hilfe.)

Der/die unterzeichnende Studierende erklärt, dass alle zitierten Quellen (auch Internetseiten) im Text oder Anhang korrekt nachgewiesen sind, d.h. dass die Bachelorarbeit keine Plagiate enthält, also keine Teile, die teilweise oder vollständig aus einem fremden Text oder einer fremden Arbeit unter Vorgabe der eigenen Urheberschaft bzw. ohne Quellenangabe übernommen worden sind.

Bei Verfehlungen aller Art treten die Paragraphen 39 und 40 (Unredlichkeit und Verfahren bei Unredlichkeit) der ZHAW Prüfungsordnung sowie die Bestimmungen der Disziplinarmassnahmen der Hochschulordnung in Kraft.

Ort, Datum:	Unterschriften:
…………………………………	………………………………………………………....
	………………………………………………………….
	………………………………………………………….

Das Original dieses Formulars ist bei der ZHAW-Version aller abgegebenen Bachelorarbeiten zu Beginn der Dokumentation nach dem Titelblatt mit Original-Unterschriften und -Datum (keine Kopie) einzufügen.

Contents

Abstract

1. Introduction

The main purpose of trading is buying and selling stocks, bonds, or other financial instruments with increasing the returns of the investments in mind while maintaining relatively low risk. With the help of a trading strategy, an investor can try to improve his performance. One can simply divide the strategies into passive and active. The praised and well established passive strategy buy-and-hold takes no short price movements into account. Positioning and trading based on these short price movements are considered active trading.

This paper applies time-series analysis to these short price movements to create active trading strategies. The objective of these developed strategies is to outperform the buy-and-hold strategy.

1.1. The data used in this paper

The dataset which will be analyzed in this paper contains 4 tradeable indexes, a visualization of the data is shown below in figure 1.

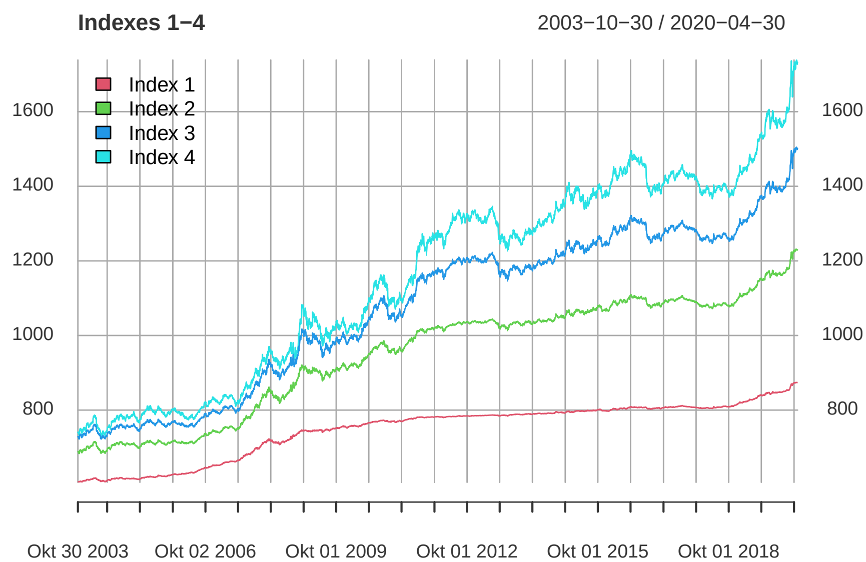


Figure 1: Visualization of the 4 indexes

Each time-series has 4306 observations and starts from October 2003 to April 2020. In all indexes is an upward drift observable, during the time period of the great recession (2008) is a slight bump visible. Also later in 2013 and 2016 are small break-ins evident. More interesting is the up and down behavior at the end of the series during the Covid19 pandemic.

In addition, to the indexes, the dataset contains 8 different interest rates of treasury bonds which will be used for further analysis. A few key-values of the interest rates are shown in the following table 1.

Table 1: Summary of the 8 interest rates.

	Maturity	Mean	Volatility	Min.	Max.
Interest 1	3M	4.09	4.95	-0.28	16.27
Interest 2	6M	4.47	5.05	0.01	16.73
Interest 3	1Y	4.64	4.22	0.20	10.51
Interest 4	2Y	5.42	4.58	0.49	16.58
Interest 5	3Y	6.16	4.33	0.75	16.51
Interest 6	5Y	7.41	3.82	1.06	16.44
Interest 7	7Y	9.50	3.31	1.71	16.63
Interest 8	10Y	11.61	2.93	3.13	17.47

A typical characteristic of interest rates is shown in the given data. A bond with longer maturities is often associated with higher returns compared with those with shorter maturities. An investor which invests in short-term treasury bonds will have his gain earlier but will be confronted with a lower return.

A more in depth analysis of the given dataset will follow in section 3.1.

1.2. Ojective of this paper

The objective of this paper is to trade these 4 indexes with an active trading strategy. The main objective is to outperform the passive buy-and-hold strategy. Methods such as the Moving-Average-Filter or the ARMA-GARCH-Model provide signals for either long or short the position to maximize the return of the investments in these indexes.

The performance of these strategies are build open various different parameters and conditions. The lengths of the filters applied to a Moving-Average may result in different solutions. Models could perform differently for any given length of the in-sample or out-of-sample scope. The necessity of including a historical crisis in the starting-sample can decide if a model performs better or worse than another. The correct validation of model parameters could have a significant impact on the forecasts.

In addition to all criteria and conditions, the strategies can be further adjusted by composing different weighted portfolios. Estimated predicted volatility can be used to modulate the position size to mitigate the risk.

Challenging will be finding the most optimal model in this wide field of conditions and parameters. The buy-and-hold strategy will be used as a benchmark to be compared with the developed active trading strategies. Computing and comparing the Sharpe ratios of each model can serve as an indicator to rely on for better or worse models.

2. Theory

It is assumed that the reader of this paper already has basic knowledge of the mathematical principles of time series analysis. Therefore, this section will only briefly describe the mathematical models and processes.

2.1. Time-Series

Almost anything with a data point to a given timestamp can be named a time-series. The monthly gross domestic product, the weekly US gasoline production, or daily stock price movements. In this paper lies the focus of the analysis of financial time series. Due to trades often only take place during the week, there are gaps in the time series on the weekends, an exception would be the trading of cryptocurrencies like Bitcoin which are also tradeable at the weekends.

A series of data points with more or less equidistant time points t with the sample length of T, is called a time-series $x_t, t = 1, ..., T$ [1]. The analysis of a time-series x_t involves creating a reasonable model that can be utilized to perform forecast predictions.

2.1.1. Stationarity

In order to fit a suitable model with a given time series x_t, the assumptions of stationarity must be met. In this practical application, only the following weak-stationarity properties are required.

$$E[x_t] = \mu \tag{1}$$

$$Var(x_t) = \sigma_x^2 \tag{2}$$

$$Cov(x_t, x_{t-k}) = R(k) \tag{3}$$

Many financial time-series are subject to shift, trends or changing volatility. In figure 2 are the stock prices of Alphabet Inc Class A (Google) visualized. This time-series shows a clear upwards drift and towards the end the volatility increases.

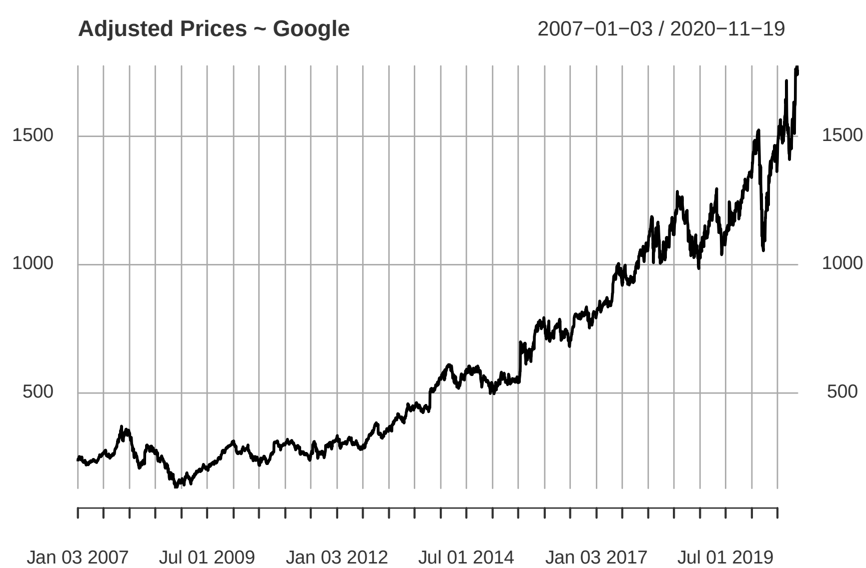


Figure 2: Visualization of the adjusted prices of the Alphabet Inc Class A Stock.

To improve the violated properties the first difference can be applied and additionally a logarithmic transformation can be performed [2]. The log-returns transformation can only performed to strict positive data.

$$\text{LogReturn} = \log(x_t) - \log(x_{t-1})$$

The result is the so-called log-returns.

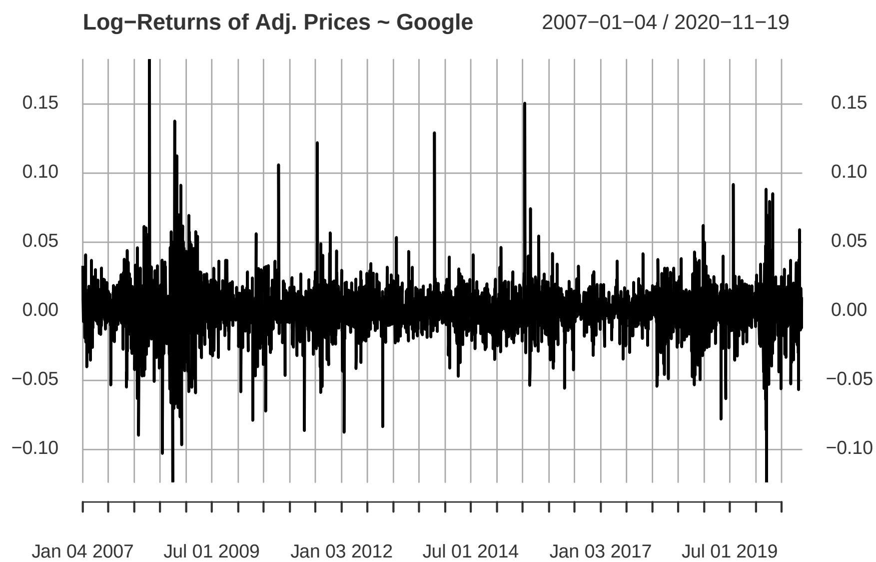


Figure 3: Visualization of the Log-Returns

Applying the transformation to the data causes the drift to disappear, but the series still contains stronger and weaker volatile phases. This effect often occurs in non-stationary financial data and is called volatility cluster. This special property is used for the modelling of forecast models, which will be discussed in chapter 2.2.

In the following examples, we will only work with a section of the time series, as it often makes no sense to look long into the past. The further a value lies in the past, the smaller its influence on a future value will be.

2.1.2. Autocorrelation

The autocorrelation function (ACF) reveals how the correlation between any two data points of the time series changes as their separation changes [3]. More precisely, acf measures the dependence between x_t and $x_{t\pm k}$ at lag k. The partial autocorrelation (PACF) measures the dependency between x_t and x_{t-k} at lag k [1]. For stationary time series, ACF can be used to identify the model order of a MA-process, PACF for AR-processes.

In the following figure 4 are ACF and PACF of the non-stationary adjusted Google stock visualized. Both graphics show the typical pattern of a non-stationary time series. The plot above shows the dependence structure of the time series. This means that it takes a long time until the series changes. Often a large value is followed by another large value, which indicates a strong trend. This property of the series can be seen in figure 2 as the long upward drift. The plot below indicates a significant partial autocorrelation at lag $k = 1$.

In the following section 2.2. the characteristics of the autocorrelation function can be used for the verification of ARIMA and ARCH-processes.

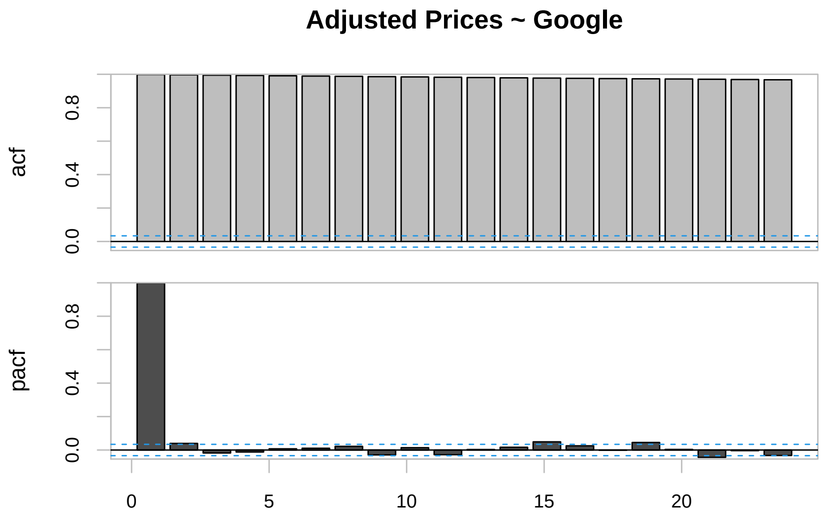


Figure 4: Acf and Pacf of the Adjusted Prizes of Google.

2.2. Models

The following processes are used to determine certain properties and characteristics of a time series so that they are transformed into a model. The goal is to fit the time series as well as possible in order to create reliable forecasts.

2.2.1. ARIMA

An ARIMA(p,d,q) process is defined as follows.

$$x_t = c + a_1 x_{t-1} + ... + a_p x_{t-p} + \epsilon_t + b_1 \epsilon_{t-1} + ... + b_q \epsilon_{t-q} \tag{4}$$

- p and q are the AR- and MA-model orders
- a and b are the AR- and MA-model parameters
- d is the differential parameter
- ϵ_t is a white noise sequence
- x_t is the given data $x_1, ..., x_T$

The mean of an ARIMA-process can be computed as:

$$\mu = \frac{c}{1 - a_1 - ... - a_p}$$

ARIMA processes can be divided into 4 different models. Choosing a model that best represents the time series is a difficult task. The goal is to find the best possible model with as few parameters as possible.

The previously introduced ACF and PACF can help to determine the orders of simple models. Provided that the time series is stationary, the model orders can be determined directly. For an AR(p)-process (ARIMA(p,0,0)), the ACF plot will gradually decrease and simultaneously the PACF should have a sharp drop after p significant lags. For an MA(q)-process (ARIMA(0,$0.q$)) the opposite is true, the ACF should show a sharp drop after a certain q number of lags while PACF should show a gradual decreasing trend. If both ACF and PACF show a gradual decreasing pattern, then the ARIMA(p,0,q)-process should be considered for modeling [4]. If the time series is not stationary, differentiation can be considered (ARIMA(p,d,q)).

The application of an analysis method to Google prices finds an ARIMA(1,1,0) as the optimal model. This makes sense if you look back at figure 4. Long dependency structures in the ACF plot indicating an AR(p) process and at the same time after lag=1=p the PACF has a strong drop. The differential operator d=1 transforms the non-stationary series into a stationary one.

To convince yourself of the quality of the model, you can use the Ljung-Box statistics shown in figure 5. For the lags where the p-values are above the 5% line, the forecasts are reliable. Here the values from the 6 lag are below the significant line, but since one only wants to make short-term forecasts (for example 1 day into the future), the lag is sufficient up to 5. If you want to forecast further than 5 days into the future, you might have to adjust the ARIMA model.

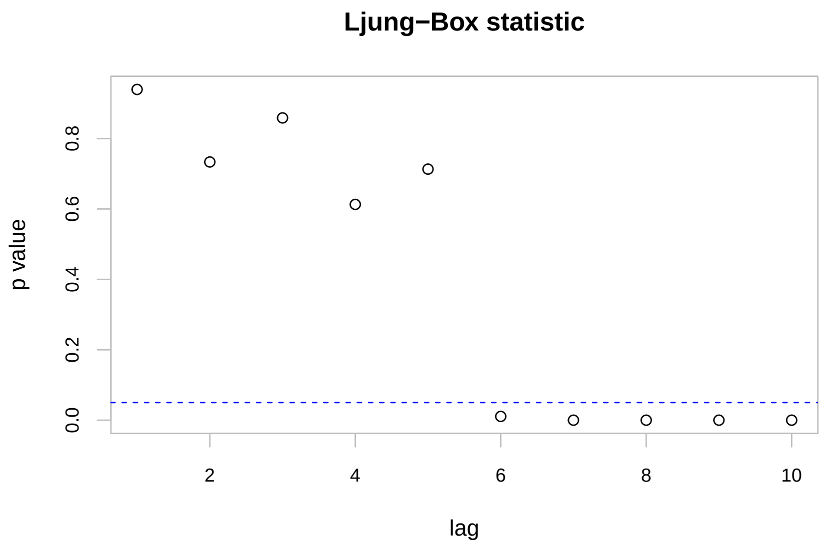


Figure 5: Ljung-Box statistic of Google log returns.

In figure 6 you can see the prediction of the model. The whole representation is shifted by one day so that one can compare the model with a true value. The green dot is the actual value of the time series. The red dots indicate the upper and lower 95% interval limits respectively. These indicate that a future value will be within this band. The blue dot is the point forecast predicted by the model.

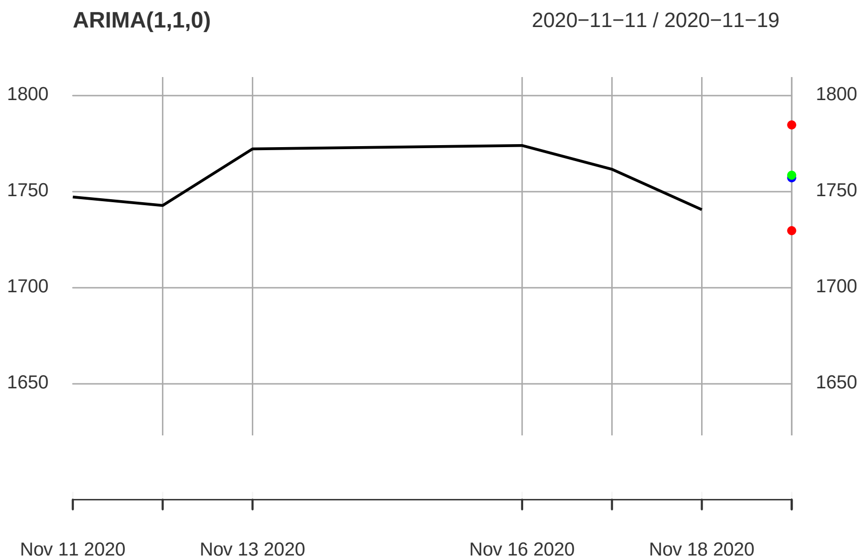


Figure 6: ARIMA-Forecast.

2.2.2. ARCH & GARCH

The volatility clustering mentioned in section 2.1.1 can be handled with an auto-regressive conditional heteroscedastic process.

$$\begin{aligned}\epsilon_t &= \log(x_t) - \log(x_{t-1}) \\ \epsilon_t &= \sigma_t u_t \\ \sigma_t^2 &= c\sigma^2 + \sum_{k=1}^{m} \beta_k \epsilon_{t-k}^2\end{aligned} \tag{5}$$

with

- x_t is the original data (often non-stationary)
- ϵ_t is the stationary log-return
- u_t is independent and identically distributed (iid) and standardized random variable
- σ^2 is the unconditional variance of the process ϵ_t.
- σ_t^2 is the conditional variance of the process ϵ_t.

The ARCH(m)-process can be generalized by adding the lagged conditional variances to the equation 5. One gets the GARCH(n,m)-process.

$$\begin{aligned}\epsilon_t &= \log(x_t) - \log(x_{t-1}) \\ \epsilon_t &= \sigma_t u_t \\ \sigma_t^2 &= c\sigma^2 + \sum_{j=1}^{n} \alpha_j \sigma_{t-j}^2 + \sum_{k=1}^{m} \beta_k \epsilon_{t-k}^2\end{aligned} \tag{6}$$

With a GARCH(n,m)-process it is possible to model the volaclusters of a time series. The GARCH(1, 1) model has become widely used in financial time series modelling and is implemented in most statistics and

Table 2: Coefficients GARCH(1,1).

	Estimate	Std. Error	p-Value
ω	0.0000109392	1.852364e-06	3.515074e-09
α_1	0.0743302768	1.090412e-02	9.314549e-12
β_1	0.8944674459	1.385232e-02	0.000000e+00

econometric software packages. Those models are favored over other stochastic volatility models by many economists due to their relatively simple implementation [5].

For an optimal model some conditions must be fulfilled. Suppose you want to model the Google time series with a GARCH(1,1).

In table 2 the estimated coefficients of the process can be seen. The p-values are all lower than 0.05 and thus indicate that they are essential for the model. (Note: $\omega = c\sigma^2$)

The following parameter restrictions are also examined:

$$c + \sum_{j=1}^{n} \alpha_j + \sum_{k=1}^{m} \beta_k = 1 \tag{7}$$

with

$$c > 0, \alpha_k \geq 0, j = 1, ..., n, \beta_k \geq 0, 1, ..., m$$

To satisfy formula 7, c needs to be determined from ω. First calculate the unconditional variance.

$$\sigma^2 = \frac{\omega}{1 - \alpha_1 - \beta_1}$$

Calculate c with:

$$c = \frac{\omega}{\sigma^2}$$

and then check for the restriction in 7.

For the coefficients of GARCH(1,1) the restrictions are fulfilled. You can see that $c = 1 - \alpha_1 - \beta_1$. So this restriction can be determined easily with:

$$\sum_{j=1}^{n} \alpha_j + \sum_{k=1}^{m} \beta_k < 1 \tag{8}$$

If the parameter restrictions are not fulfilled, complications may arise, the forecast of the conditional variance $\hat{\sigma}_t^2$ may diverge to the unconditional variance σ^2 of the process.

Furthermore, the Ljung-Box statistics are important for the standardized residuals. Looking back at formula 6 standardized residuals u_t are proportional to the conditional volatilities σ_t, which should lead to the log returns ϵ_t. The conditional volatilities map the volacluster in the time series. To achieve the best possible model, one does not want to find these volacluster effects in the standardized residuals, but only in the conditional volatilities. The Ljung-Box statistics check this property. In figure 7 Ljung-Box statistics of the u_t and the u_t^2 are shown.

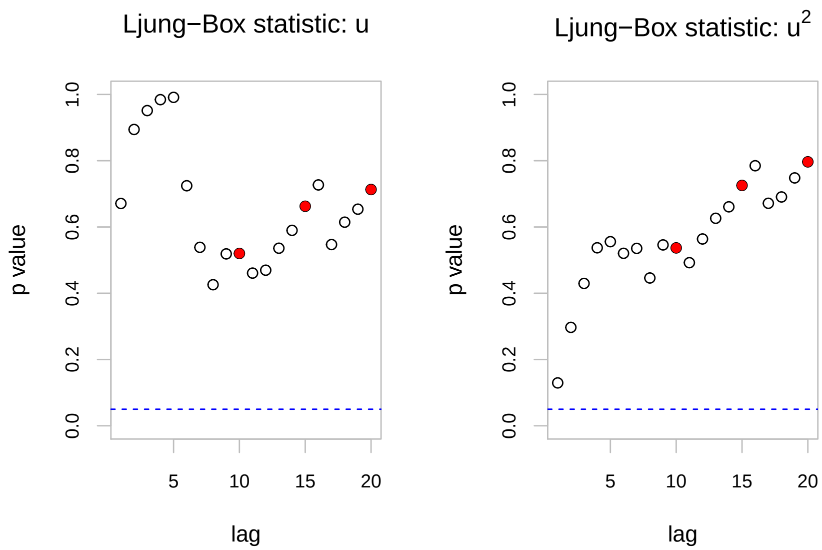


Figure 7: Ljung-Box statistic of the standardized residuals.

The plot shows reliable statistics for forecasts from the Google-GARCH(1,1) model. For all lags up to 20, the p-values are above the 5% line and thus hypothesis tests are discarded. However, if for a given lag=k the p-values would fall below the 5% line, then forecasts would only be reliable up to a forecast horizon k.

If one wants to improve the the standardized residuals u_t, an ARIMA part would have to be added to the existing model (see 2.2.3). This can again be optimized with different model orders. If you want to improve the squared standardized residuals u_t^2, then you should modify the GARCH model order.

Now an optimal model has been found and a forecast can be made. Since a GARCH(n,m) process is white noise sequence the expected value $E[\epsilon_{T+h}|\epsilon_T, ..., \epsilon_1] = \mu = 0$ can be assumed (if the mean value in the fit object was also estimated and is significant, then the expected value is the estimated μ).

Calculating the forecast variance is a recursive process. With increasing model order the calculation becomes more and more difficult. For this work the rather simple calculation for a GARCH(1,1) model is sufficient. One receives:

$$\hat{\sigma}^2_{T+h} = \omega + (\alpha_1 + \beta_1)^2 \hat{\sigma}^2_{T+h-1}$$

If the parameter restriction from formula 8 is true, the forecast variance converges with the increasing forecast horizon to the unconditional variance of the process.

$$\hat{\sigma}^2_{T+h} = \frac{\omega}{1 - \alpha_1 - \beta_1} = \sigma^2$$

The 95% forecast interval is calculated as follows:

$$E[\epsilon_{T+h}|\epsilon_T, ..., \epsilon_1] \pm 1.96\sqrt{\hat{\sigma}^2_{T+h}}$$

Applying

bla bla blublibli

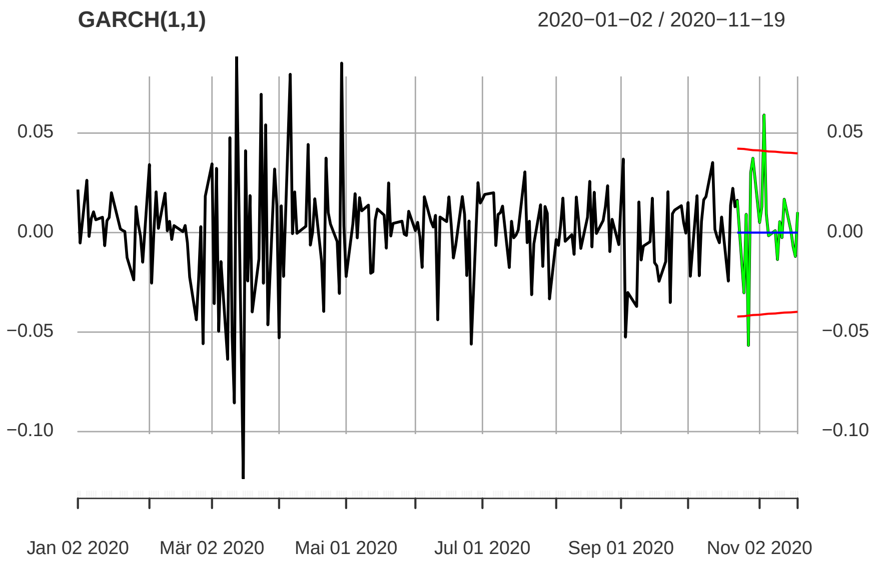


Figure 8: GARCH-Forecast.

One can see the divergence to the unconditional standard deviation $\sigma =$ asd 0.018724

2.2.3. ARIMA-GARCH

Another process is the combination of ARIMA and GARCH processes.

$$\begin{aligned} y_t &= \mu + a_1 y_{t-1} + ... + a_p y_{t-p} + \epsilon_t + b_1 \epsilon_{t-1} + ... + b_q \epsilon_{t-q} \qquad (9) \\ \epsilon_t &= \sigma_t u_t \\ \sigma_t^2 &= c\sigma^2 + \sum_{j=1}^{n} \alpha_j \sigma_{t-j}^2 + \sum_{k=1}^{m} \beta_k \epsilon_{t-k}^2 \qquad (10) \end{aligned}$$

Is called the mean-equation 9

Is called the variance-equation 10

[6]

2.3. Moving Average Filters

moving average filters are basically used to identify trends and smooth out price fluctuations. As a commonly used tool moving average filters are very simple in its usage, historical data was summarized and divided bi the length of the filter. Many differnt indicators are built on the Moving average principle, mostly theyre used in combinations of different lengths to create new indicators. in the following section we introduce some of the popular indicators based on the MA principle.

The actual challenge in using Moving average filters is to figure out which length of the filter brings the most useful information.

2.3.1. Equally-weighted Moving Average or SMA

SMA stands for Simple Moving Average which. depending on the length of the filter(l) l obervations since the last noted observation will be considered. theyre getting summarized and divided by the filterlength equals the EqMA. For every timestep a new observation is considered and the last one eliminated.

EqMA

$$y_t = \frac{1}{L} \sum_{k=0}^{L-1} x_{t-k} \tag{11}$$

- L = filterlength
- x = original series price e.g.

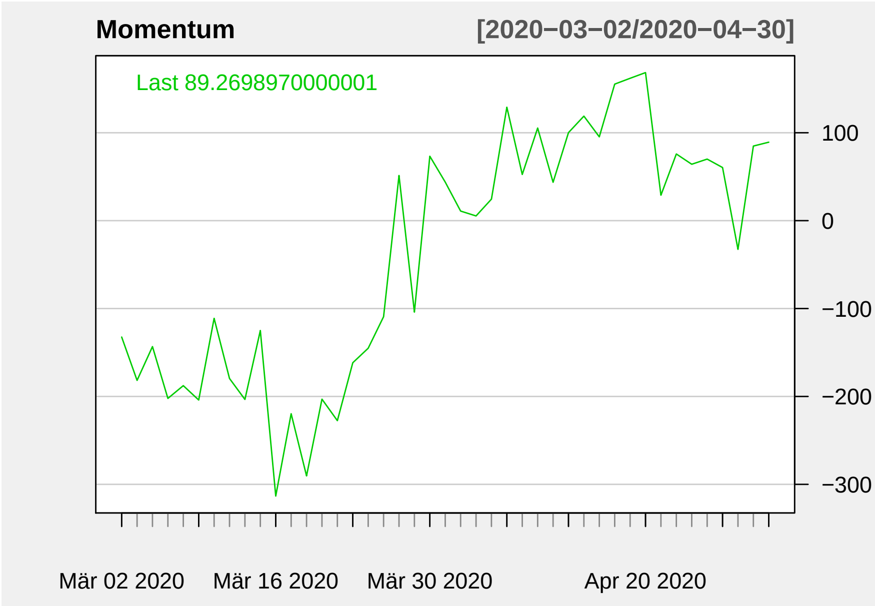
2.3.1.1 Momentum

Momentum is an indicator wether a market is bullish or bearish, it measures the “speed” of the trend direction in the market. for a timespan k the last price p_k, k timesteps ago is subtractet from the last price. This is equivalent to applying and EqMA to the price differences.

Momentum

$$y_t = p_t - p_{t-K} \tag{12}$$

- p_t = prices of the series
- K = Lag



2.3.2. Exponentially-weighted Moving Average

Since not all observations are having the same influence of future value we can apply a weight to past observations. One method will be exponentially weighted Moving average. So we chose an optimal parameter to give past observations weights decreasing by alpha

A skillfull trader chose an optimal α to increase the performance of the measurement. Weights could also be given individually by adding a weight vector to the filter.

EMA

$$y_t = \frac{1}{\sum_{k=0}^{m} \alpha^k} \sum_{k=0}^{m} \alpha^k x_{t-k} \tag{13}$$

- m = filterlength
- α = Parameter to weigh the observations

2.3.3. Moving Average Crossings

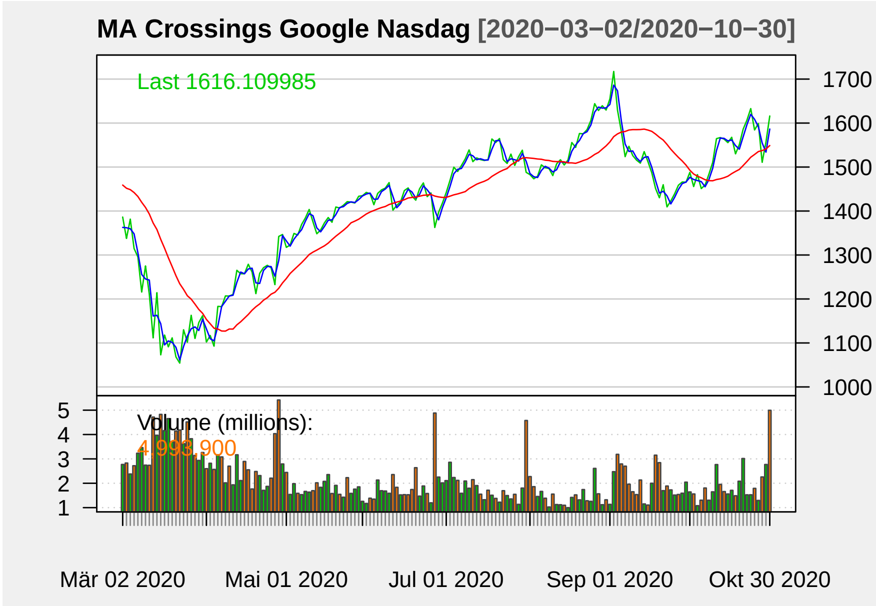
$$y_t = \frac{1}{L_1} \sum_{k=0}^{L_1-1} x_{t-k} - \frac{1}{L_2} \sum_{k=0}^{L_2-1} x_{t-k} \tag{14}$$

- L_1 = filterlength 1

- L_2 = filterlength 2
- $0 < \alpha < 1$ = Parameter to weigh the observations

Moving average crossings are basically just different MA's with different lengths applied to a time-series. The points the filters then cross will be used as a trading signal to go long, short or hold.

An easy example of Ma average crossings with 2 mas of different length is visualized in # 2.



2.4. Real Strength Index

2.5. Sharpe Ratio

Sharpe ratio is a very powerful and widely used ratio to measure performance. It describes return per risk.

$$\text{SharpeRatio} = \frac{R_p - R_f}{\sigma} \tag{15}$$

- R_p = Return of Portfolio
- R_p = Risk free Rate, mostly treasury bonds
- σ_p = standard deviation of portfolios excess return (risk)

2.6. Carry

carry trades are trading strategies were usually money is borrowed at a lower interest rate, than the investment is giving in return. the risk of this strategy is based in the currency risk.

2.7. Value

2.8 Bollinger bands

Bollinger bands are a analysis tool founded by John Bollinger. It contains a moving average and an upper and lower band . The bands are defined by adding a constant$K$ times a standard deviation σ_t to the *Moving Average* for the upper , and subtracting it for the lower band.

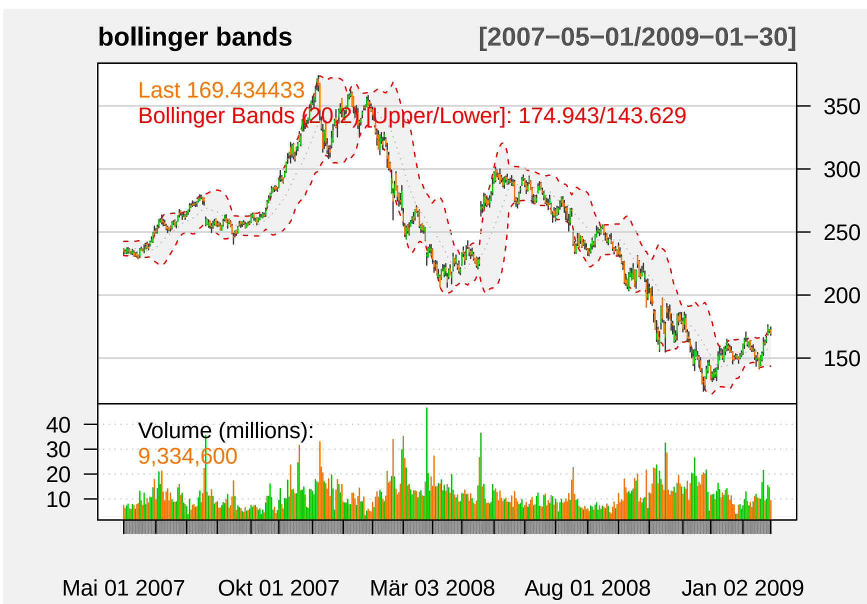
$$\mathrm{U_t} = MA_t + K\sigma, L_t = MA_t - K\sigma \tag{16}$$

the variance from bollingers theory is calculated by:

$$\sigma_{\mathrm{t}}^2 = \frac{1}{N}\sum_{k=0}^{N-1}(x_{t-k} - MA_t)^2 \tag{17}$$

The calculated σ_t could be problematic because its derived from the original series and increases with the level, its non stationary. Therefore an other method to calculate the standart deviation could be used. As done in section 2.2.2. σ could be provided by a GARCH, which would handle the volatility.

- N = usually the filterlength and the length considered for σ are the same
- K = Constant usually equals 2
- σ_p = standard deviation of the series
- U_t = upper band
- L_t = lower band



2.9 Moving Average Convergence Divergence

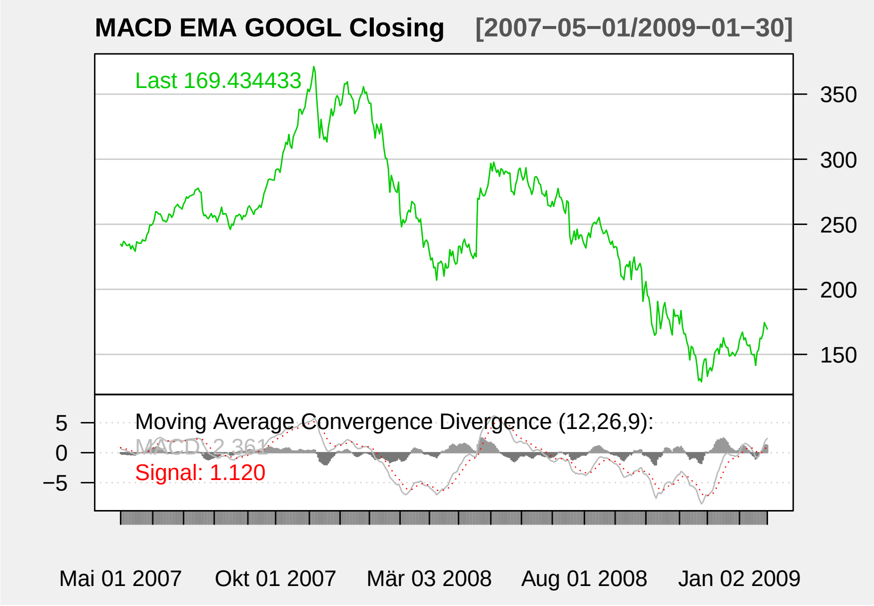
The MACD is also a commonly used filter. The basic principle is to Subtract a longer EMA with length L as in section 2.3.3 from a shorter EMA from length S then smooth the result with another EMA with length R. As a result with can use the crossing of the 2 generated curves for trading. As an alternative we could use SMA's instead of EMA's

$$macd_t = \frac{1}{\sum_{k=0}^{t-1} \alpha^k} \sum_{k=0}^{t-1} \alpha^k x - t - k - \frac{1}{\sum_{k=0}^{t-1} \beta^k} \sum_{k=0}^{t-1} \beta^k x - t - k \tag{18}$$

$$MACDsignal_t = \frac{1}{\sum_{k=0}^{t-1} \gamma^k} \sum_{k=0}^{t-1} \gamma^k x - t - k \tag{19}$$

- x_t = prices or log prices
- S = length of the $short_1$ EMA usually 12
- L = length of the $long_1$ EMA usually 26
- R = length of the "double smoothing ema" usually 9
- $\alpha = 1 - \frac{1}{S}$
- $\beta = 1 - \frac{1}{L}$
- $\gamma = 1 - \frac{1}{R}$

$_1$ short and long in the meaning s<l, not buy sell



3.1 Trading Signals Using signals to trade

in Section 2 we've learned different indicators and models for timeseries. these models and indicators are now used to trade the underlying asset. As an example we're taking the ma crossings, as mentioned 2.3.3. the points where the two MAs cross, are now used to create a trading signal. when the longer MA comes from below to the crossing we are going long the asset and if it approaches the point from above we're shorting the position. Technically We apply a 1 to a vector at each crossing, where we intend to buy and apply a -1 at the points we want to sell.

3. Methodology

In this section models were created trying to outperform the buy and hold strategy. starting with the usage of the simplest models , diffrent aproaches were chosen to fullfill the goal.

3.1. Time-Series Analysis

Plots of the timeseries, decomposition. Stationarity (refer to the theory section)

3.2 using simple methods

3.2.1 sma signals to trade

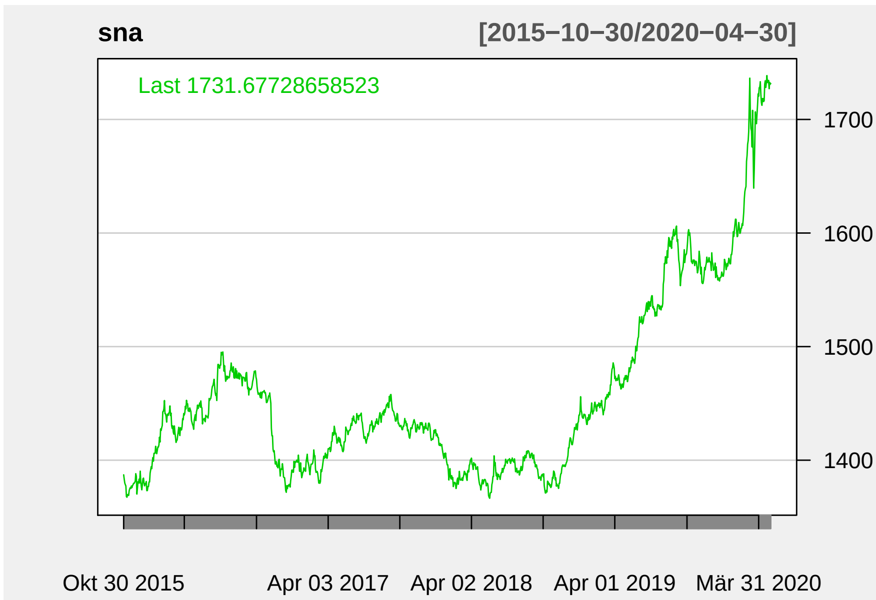


Figure 9: conversion data

```
## [1] 0.0001624021
```

Figure 10: conversion data

Figure 11: conversion data

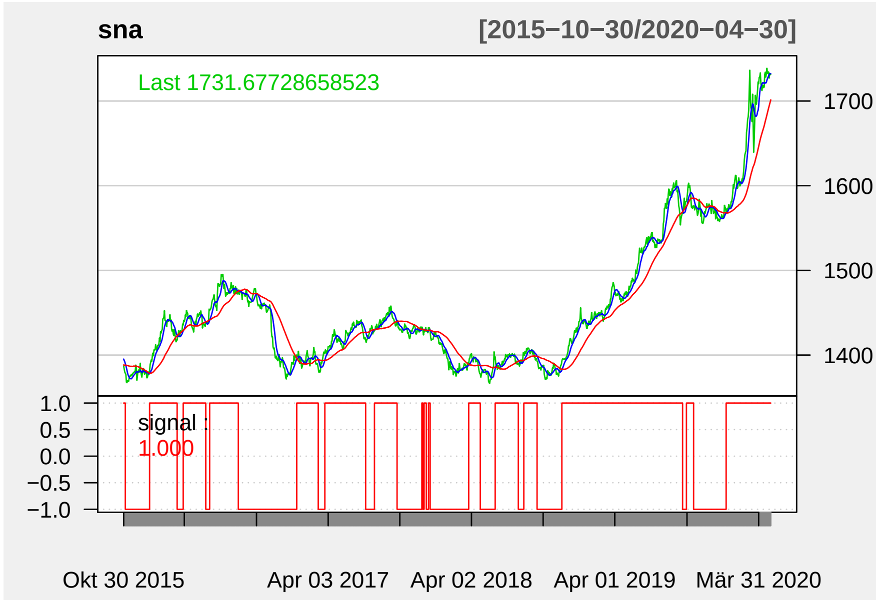


Figure 12: conversion data

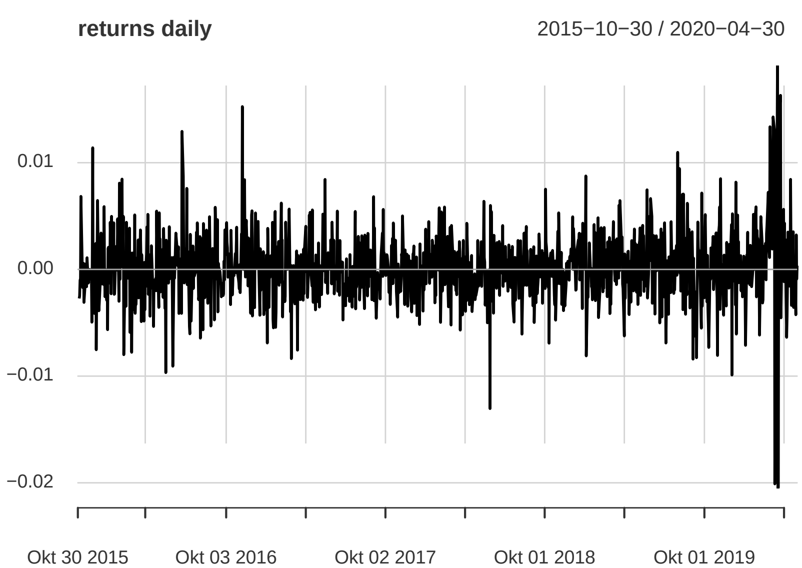


Figure 13: conversion data

Figure 14: conversion data

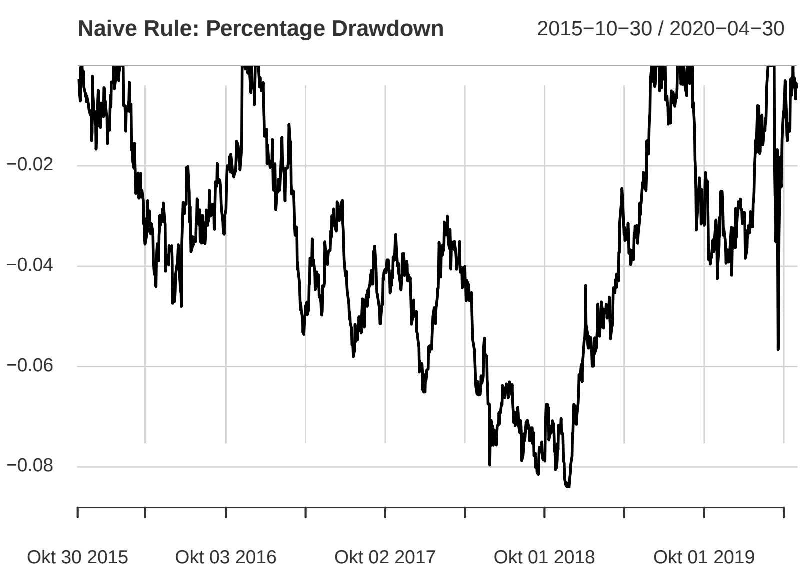


Figure 15: conversion data

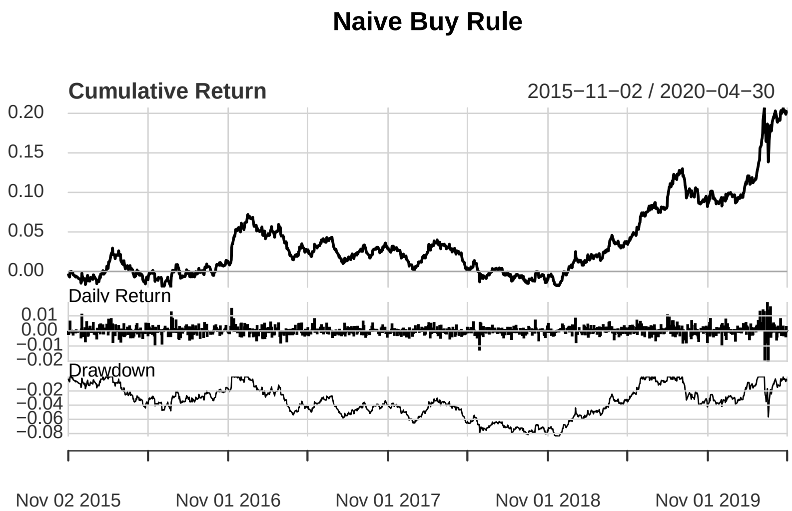


Figure 16: conversion data

4. Conclusion

5. References

[1] M. Wildi, *Econometrics 1: Time series analysis*. Winterthur: ZHAW, 2017, p. 221.

[2] M. Wildi, *Econometrics 3: Conditional heteroscedasticity models*. ZHAW, 2020, p. 56.

[3] A. H. A. L. Kholkin D. A. Kiselev, *Encyclopedia of materials: Science and technology (second edition)*. Elsevier, 2001, p. 10388.

[4] M. Masum, "Time series analysis: Identifying ar and ma using acf and pacf plots." https://towardsdatascience.com/identifying-ar-and-ma-terms-using-acf-and-pacf-plots-in-time-series-forecasting-ccb9fd073db8 (accessed Nov. 16, 2020).

[5] B. Williams, *GARCH(1,1) models*. Ruprecht-Karls-Universität Heidelberg, 2011, p. 42.

[6] M. Wildi, *An introduction to conditional volatility models*. Winterthur: ZHAW, 2020, p. 32.

Attachment