Applied Time Series Analysis
FS 2011 – Week 14

Marcel Dettling

Institute for Data Analysis and Process Design

Zurich University of Applied Sciences

marcel.dettling@zhaw.ch

http://stat.ethz.ch/~dettling

ETH Zürich, May 30, 2011

Applied Time Series Analysis
FS 2011 – Week 14

Non-Linear Models: ARIMA and SARIMA

Why?

We have seen that many time series we encounter in practice show trends and/or seasonality. While we could decompose them and model the stationary part, it might also be attractive to directly model a non-stationary series.

How does it work?

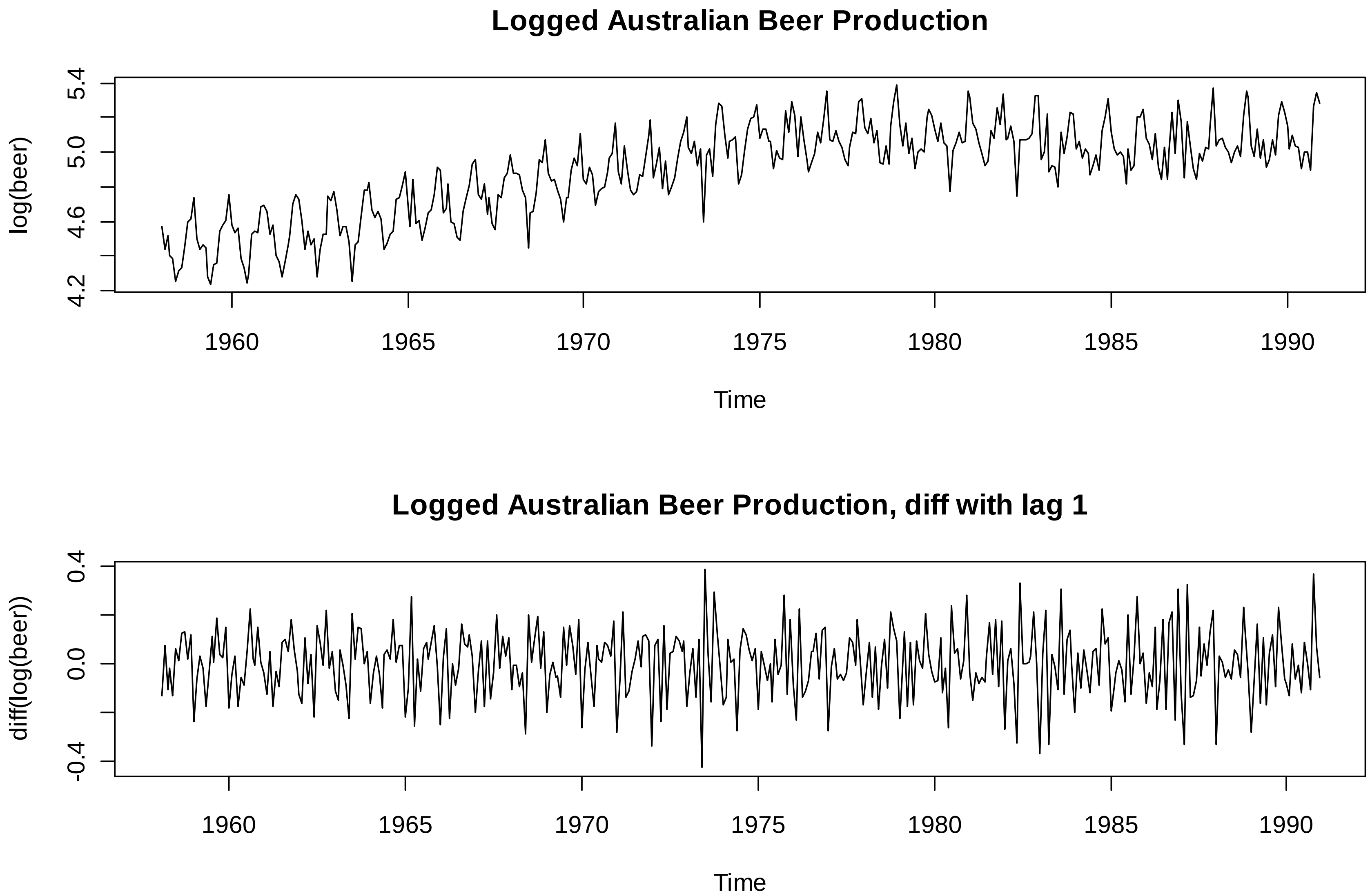
There is a mechanism, "the integration" or "the seasonal integration" which takes care of the deterministic features, while the remainder is modeled using an ARMA(p,q).

There are some peculiarities!

→ **see blackboard!**

Applied Time Series Analysis
FS 2011 – Week 14

Example: Australian Beer Production



ARIMA(p,d,q)-Models

Idea: Fit an ARMA(p,q) to a time series where the d^{th} order difference with lag 1 was taken before.

Example: If $Y_t = X_t - X_{t-1} = (1-B)X_t \sim ARMA(p,q)$, then $X_t \sim ARIMA(p,1,q)$

Notation: With backshift-operator B()

$$\Phi(B)(1-B)^d X_t = \Theta(B)E_t$$

Stationarity: ARIMA-models are usually non-stationary!

Advantage: it‘s easier to forecast in R!

Fitting and Forecasting

We start by fitting an ARIMA(0,1,1) to the beer series:

```
fit <- arima(log(beer), order=c(0,1,1))

> fit

Call: arima(x = log(beer), order = c(0, 1, 1))

Coefficients:              ma1
                       -0.2934
                s.e.    0.0529

sigma^2 estimated as 0.01734
log likelihood = 240.28,  aic = -476.57
```

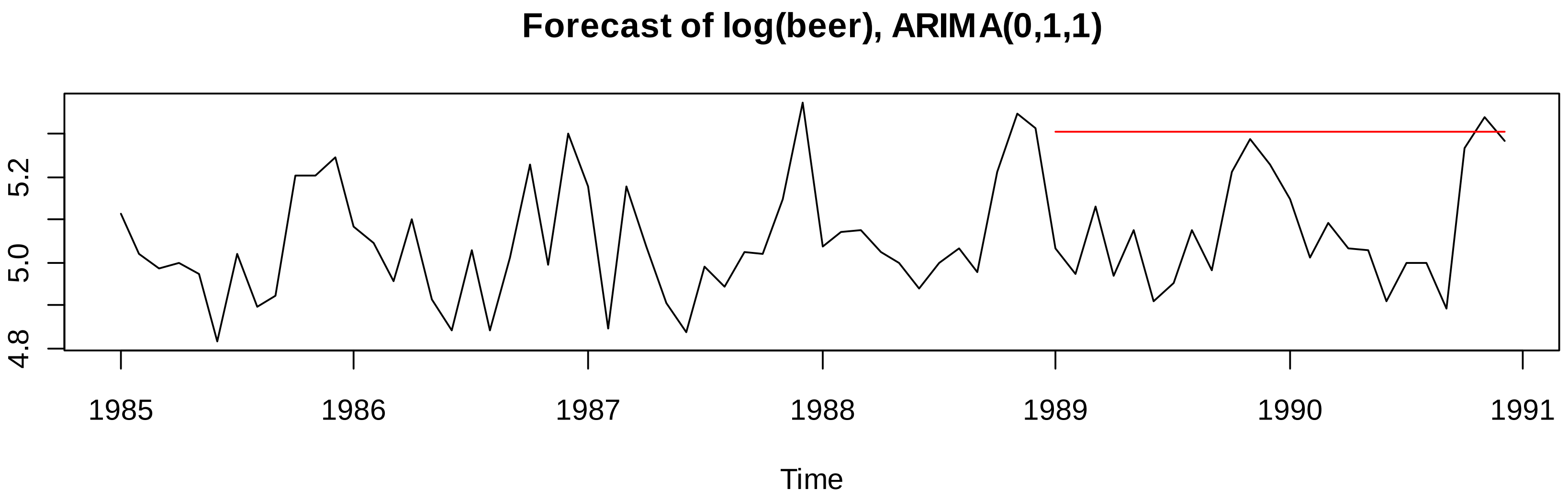
Fitting and Forecasting

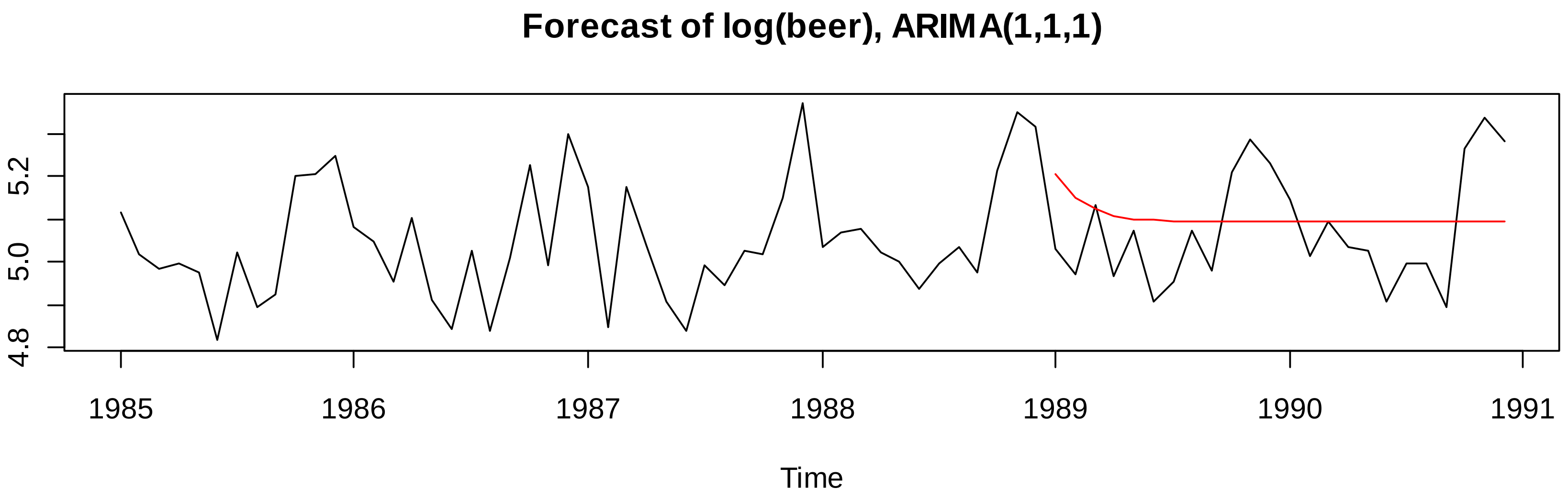
We start by fitting an ARIMA(0,1,1) to the beer series:

```
> fit.111 <- arima(log(beer), order=c(1,1,1))

> fit.111

Call: arima(x = log(beer), order = c(1, 1, 1))

Coefficients:            ar1      ma1
                      0.5094  -0.9422
                s.e.  0.0469   0.0125

sigma^2 estimated as 0.01491
log likelihood = 269.51,  aic = -533.01
```

Applied Time Series Analysis
FS 2011 – Week 14

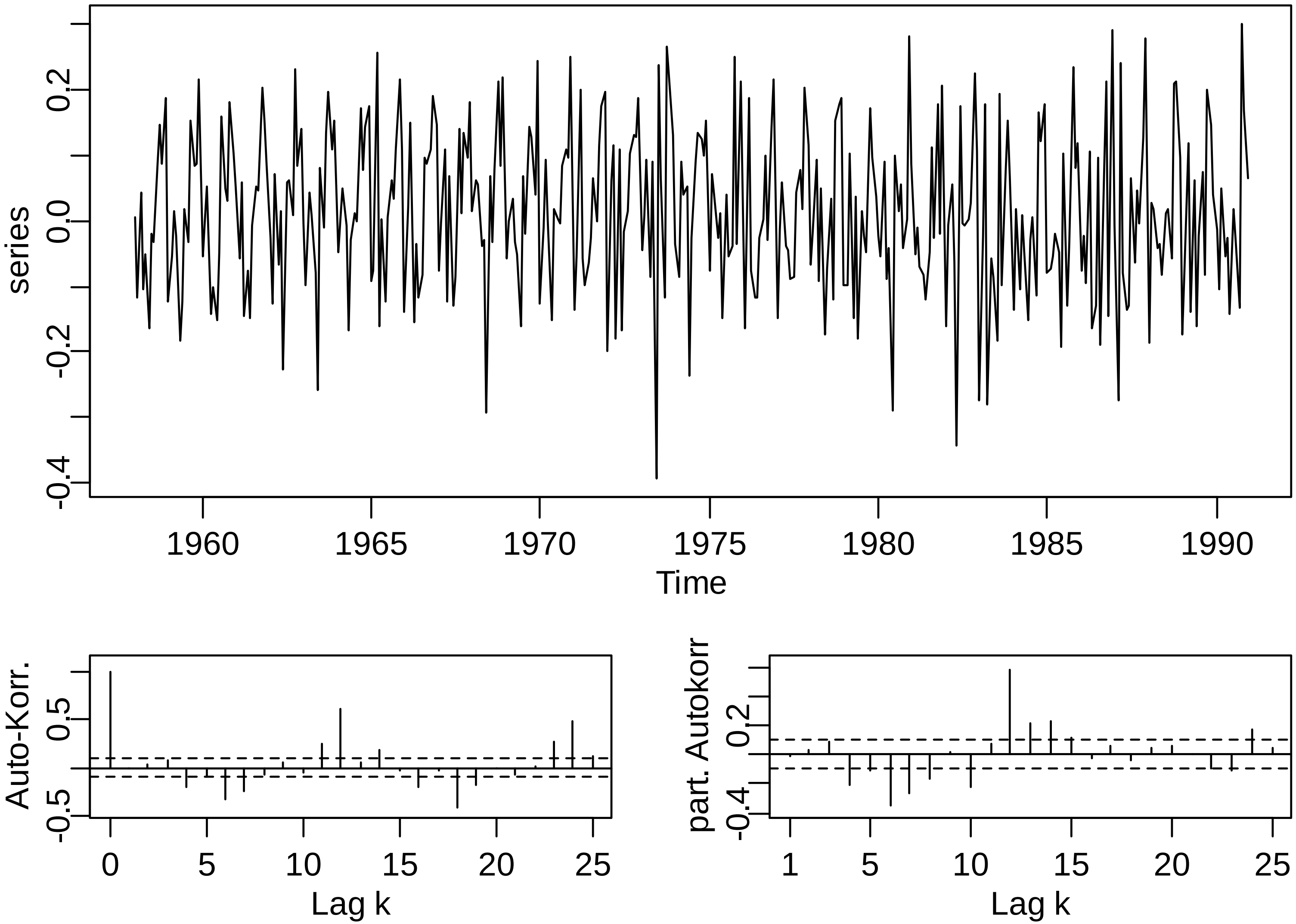
Forecasting Australian Beer Production





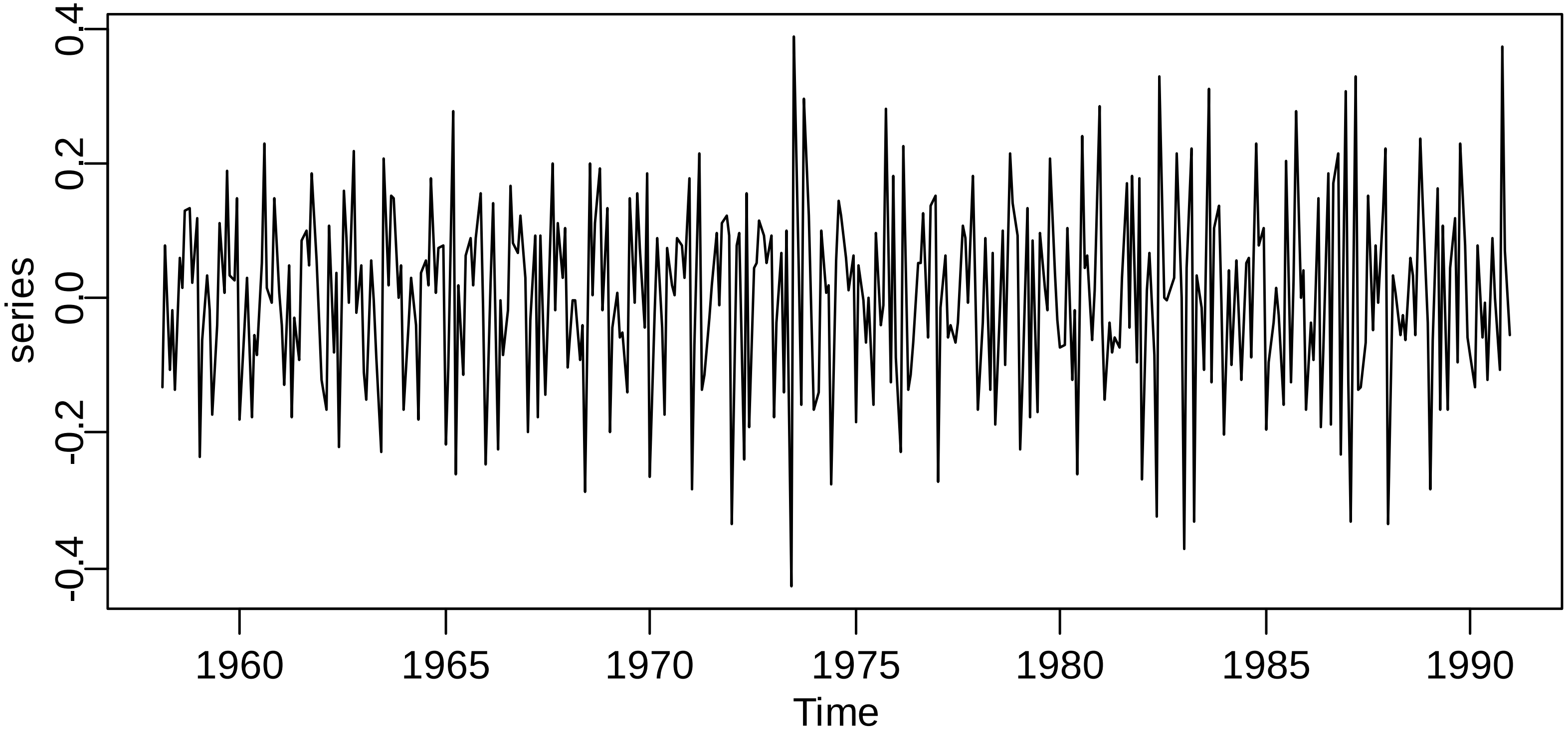
Applied Time Series Analysis
FS 2011 – Week 14

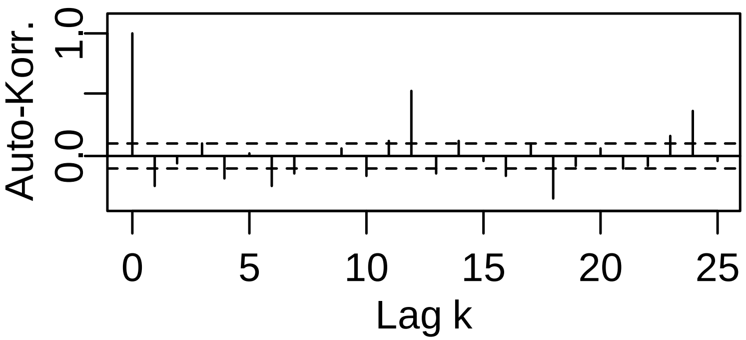
Residual Analysis of the ARIMA(1,1,1)

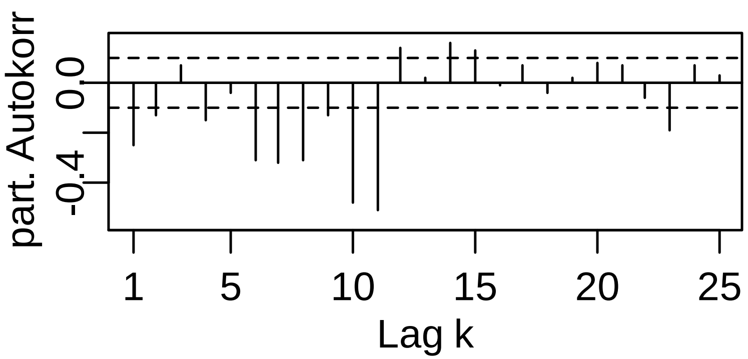


Applied Time Series Analysis
FS 2011 – Week 14

ACF/PACF: Differenced Original Series







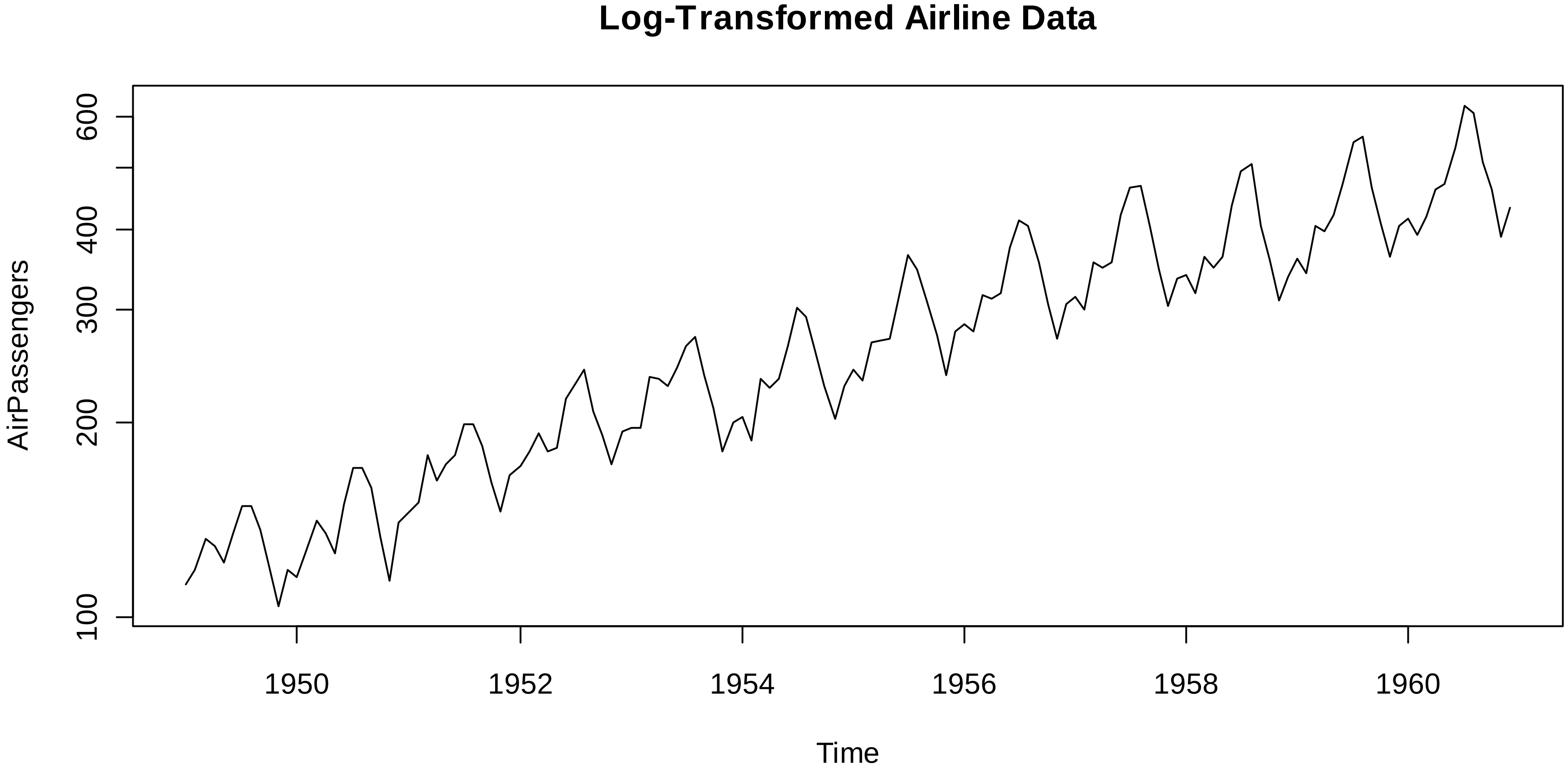
Beer Series: Results

- From the residuals, we cleary observe that the fitted models are not adequate, and thus the forecasts are not to be trusted
- It seems as if we failed to include the seasonality into the model. This is visible from the residuals.
- However, this is also visible from ACF/PACF of the original (differenced) series. This is where we made the "mistake" in the first place.
- We need more complex models which can also deal with seasonality. They exist, see the following slides...

Applied Time Series Analysis
FS 2011 – Week 14

SARIMA(p,d,q)(P,D,Q)^s

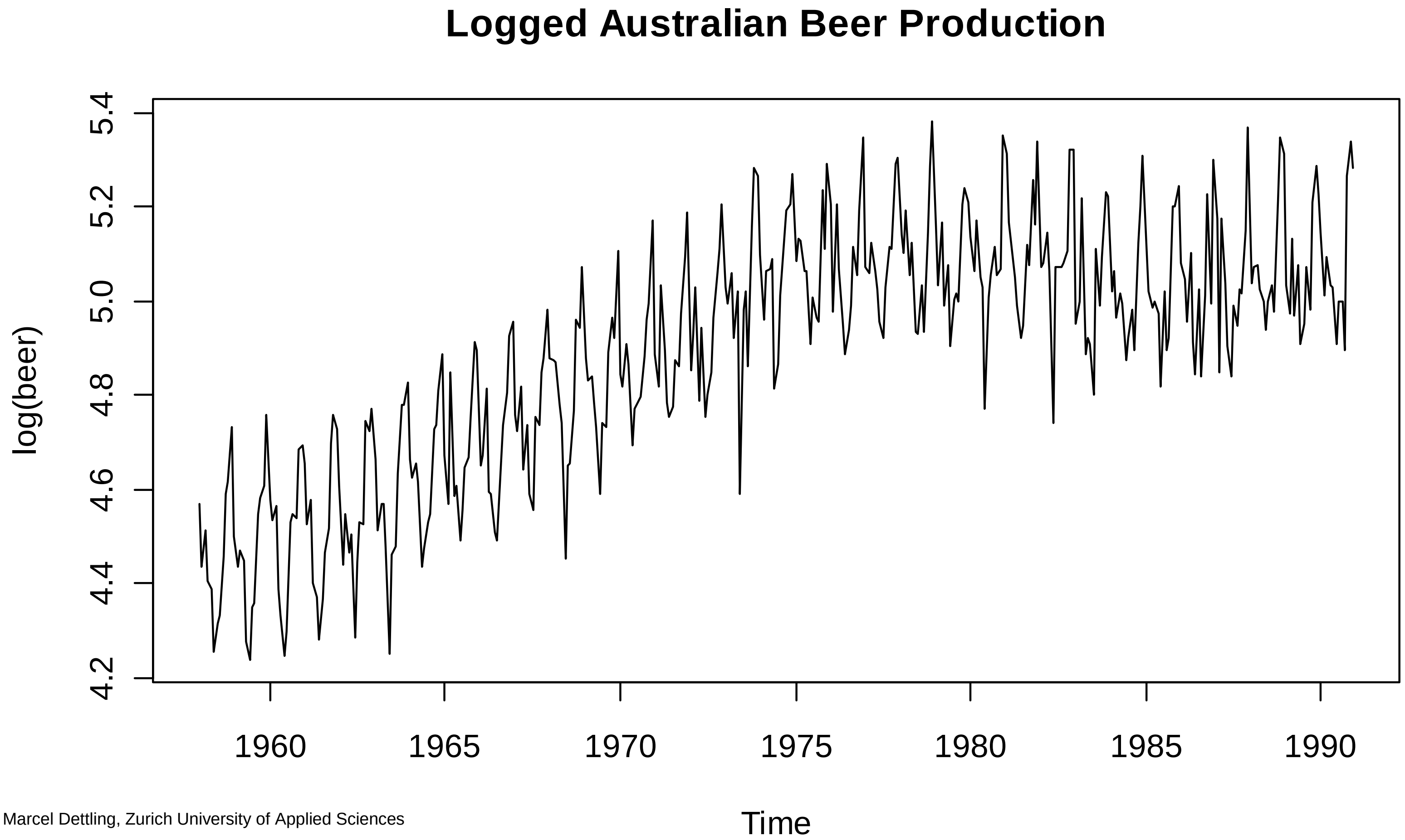
= a.k.a. Airline Model. We are looking at the log-trsf. airline data



Applied Time Series Analysis
FS 2011 – Week 14

SARIMA(p,d,q)(P,D,Q)^s

or at the log-transformed Australian Beer Production



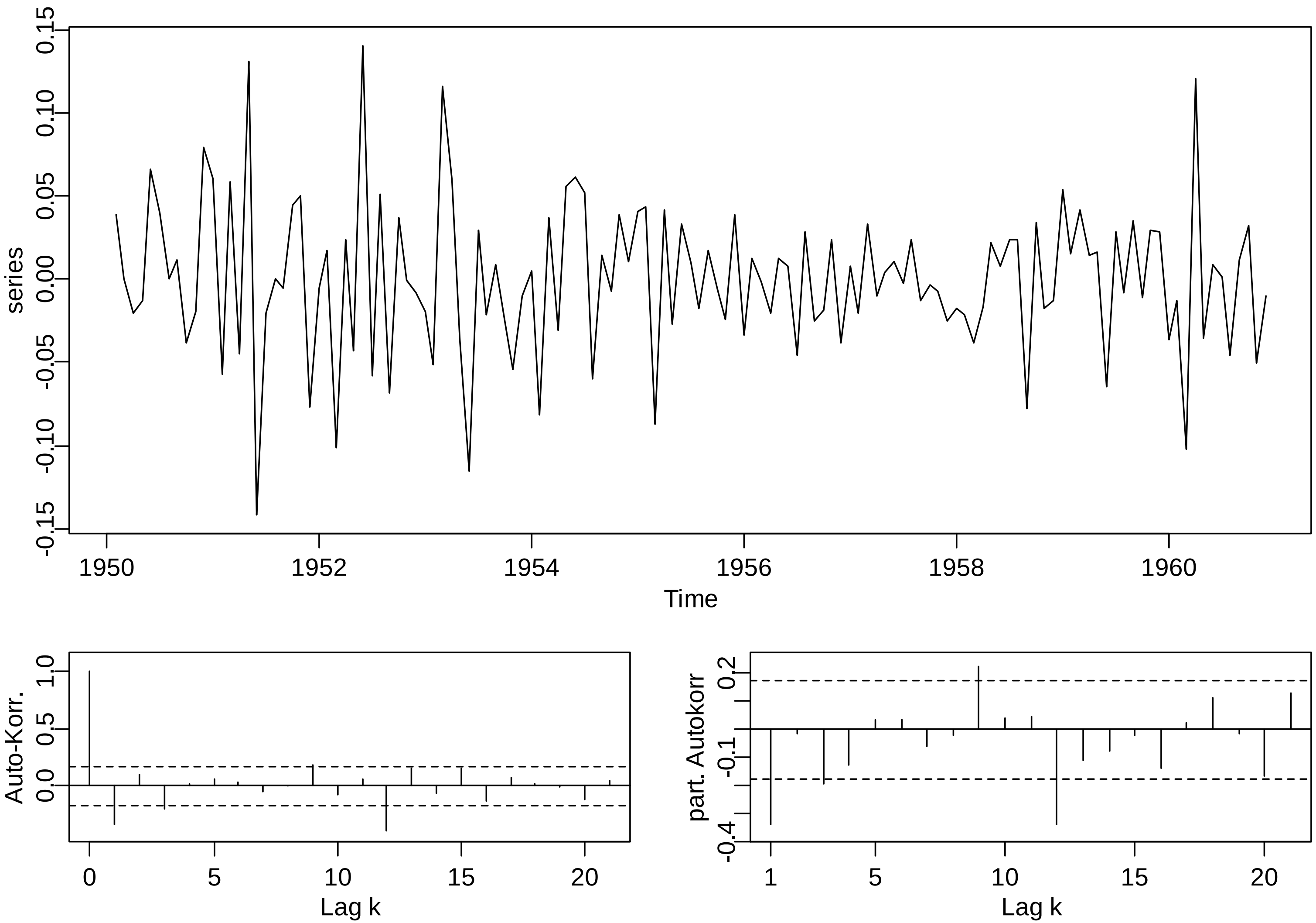
Applied Time Series Analysis
FS 2011 – Week 14

SARIMA(p,d,q)(P,D,Q)^s

We perform some differencing… (→ **see blackboard**)

Applied Time Series Analysis
FS 2011 – Week 14

ACF/PACF of SARIMA(p,d,q)(P,D,Q)^s



Modeling the Airline Data

Since there are "big gaps" in ACF/PACF:

$$Z_t = (1+\beta_1 B)(1+\gamma_1 B^{12})E_t$$
$$= E_t + \beta_1 E_{t-1} + \gamma_1 E_{t-12} + \beta_1\gamma_1 E_{t-13}$$

This is an MA(13)-model with many coefficients equal to 0, or equivalently, a SARIMA(0,1,1)(0,1,1)^{12}.

Note: Every SARIMA(p,d,q)(P,D,Q)^s can be written as an ARMA(p+sP,q+sQ), where many coefficients will be equal to 0.

Applied Time Series Analysis
FS 2011 – Week 14

SARIMA(p,d,q)(P,D,Q)^s

The general notation is:

$$Y_t = (1-B)^d (1-B^s)^D X_t$$

$$\Phi(B)\Phi(B^s)Y_t = \Theta(B)\Theta^s(B^s)E_t$$

Interpretation:

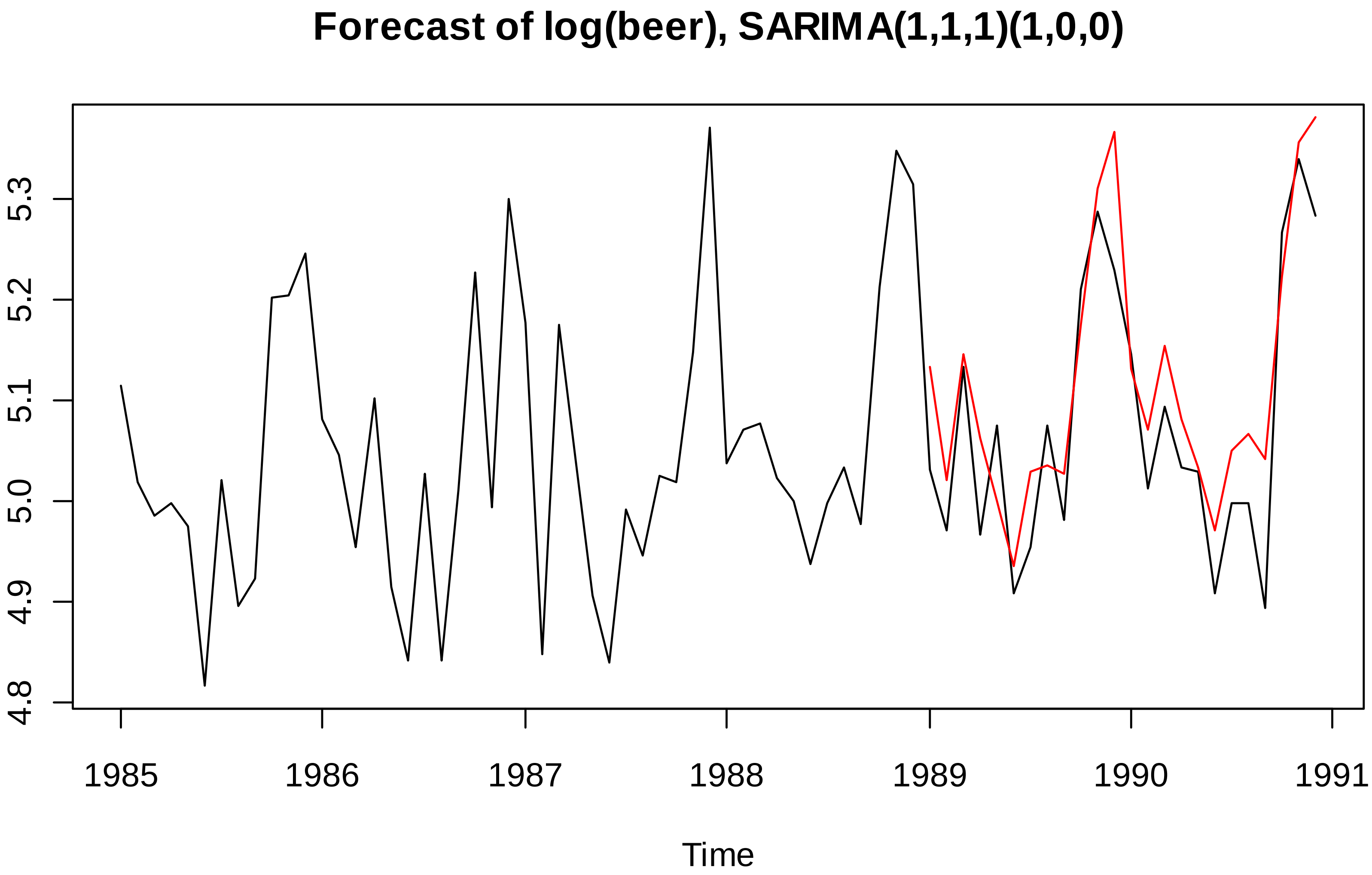
- one typically chooses d=D=1
- s = periodicity in the data (season)
- P,Q describe the dependency on multiples of the period

→ **see blackboard**

Zurich University of Applied Sciences

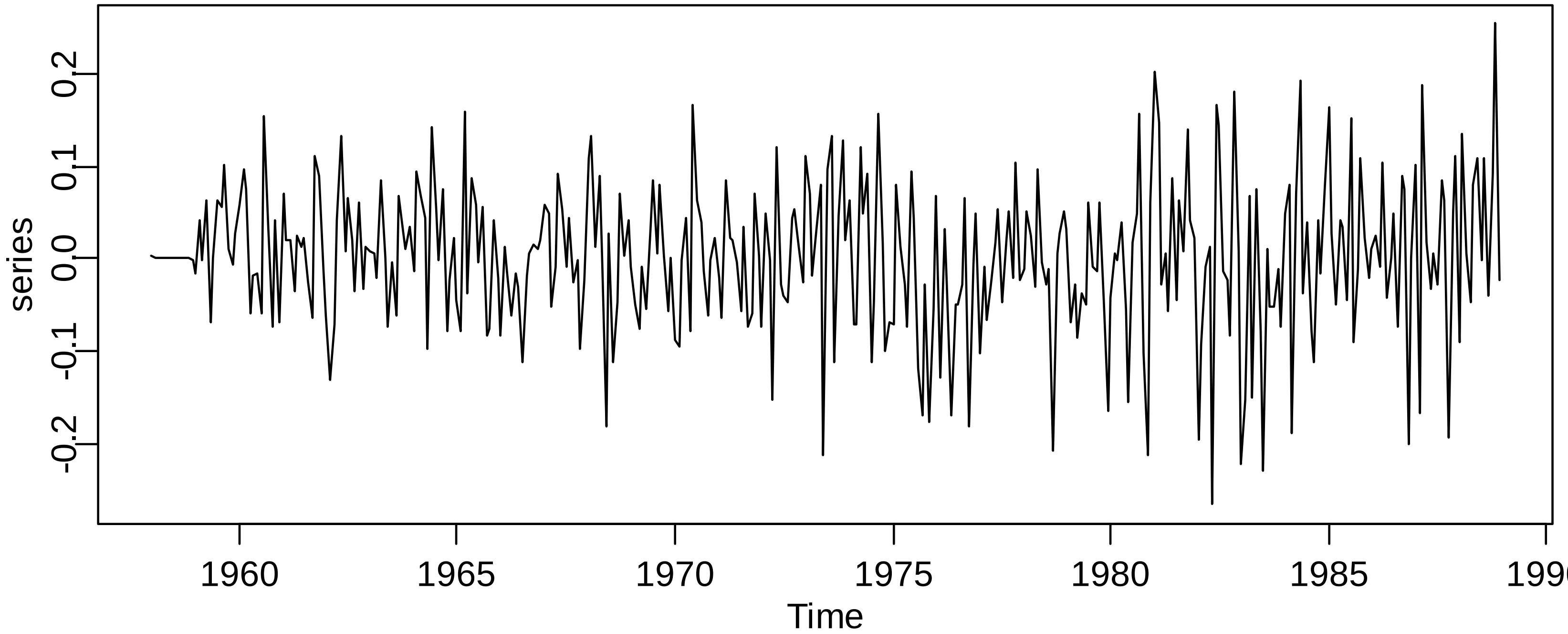
Applied Time Series Analysis
FS 2011 – Week 14

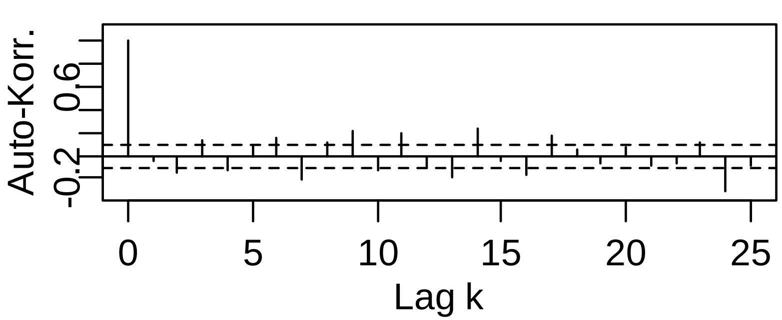
Forecasting Australian Beer Production

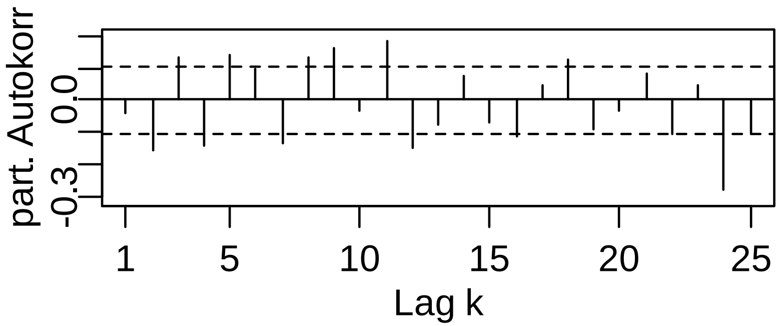


Applied Time Series Analysis
FS 2011 – Week 14

Residual Analysis of SARIMA(1,1,1)(1,1,0)







Applied Time Series Analysis
FS 2011 – Week 14

Outlook to Non-Linear Models

What are linear models?

Models which can be written as a linear combination of X_t
i.e. all AR-, MA- and ARMA-models

What are non-linear models?

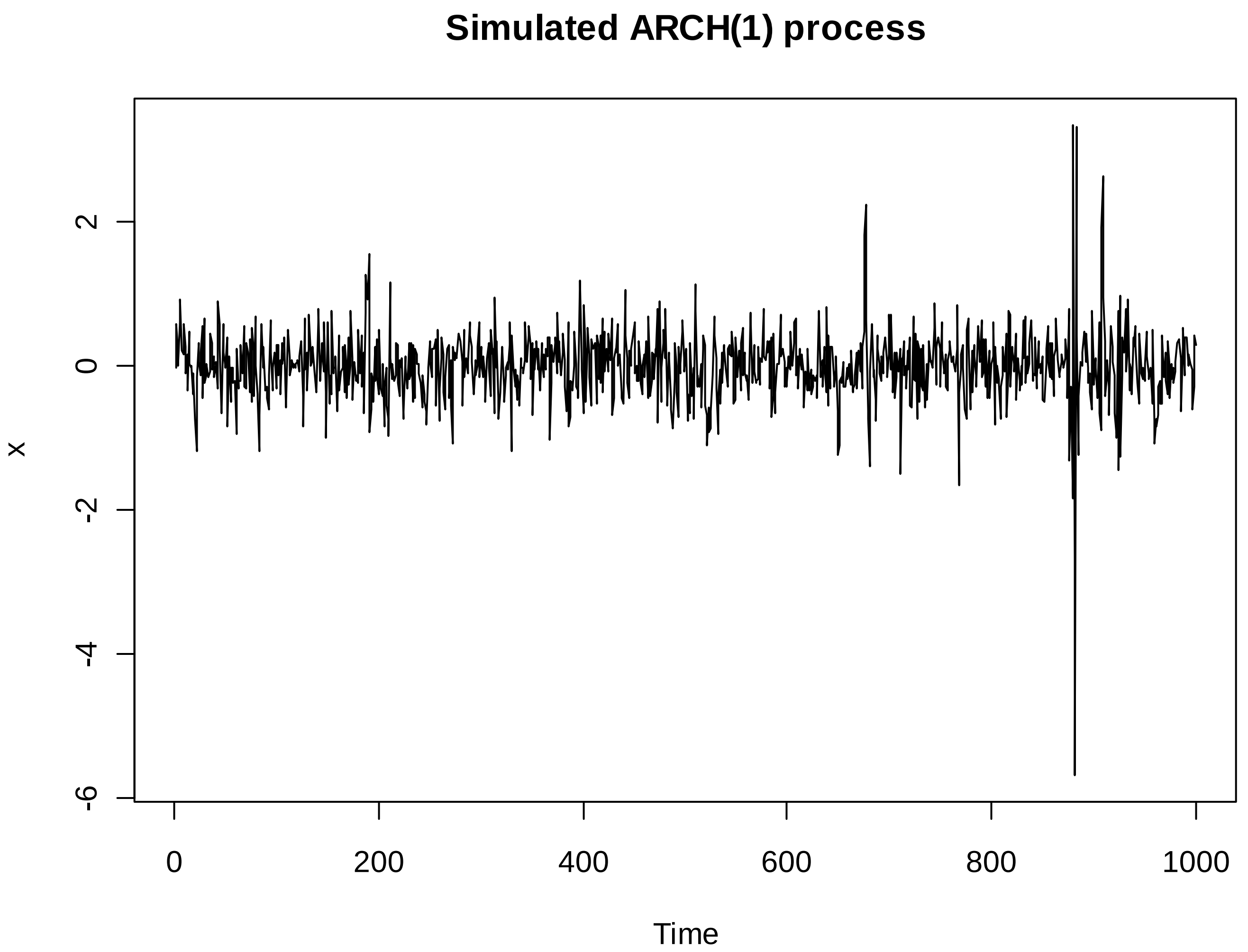
Everything else, e.g. non-linear combinations of X_t,
terms like X_t^2 in the linear combination, and much more!

Motivation for non-linear models?

- modeling cyclic behavior with quicker increase then decrease
- non-constant variance, even after transforming the series

Applied Time Series Analysis
FS 2011 – Week 14

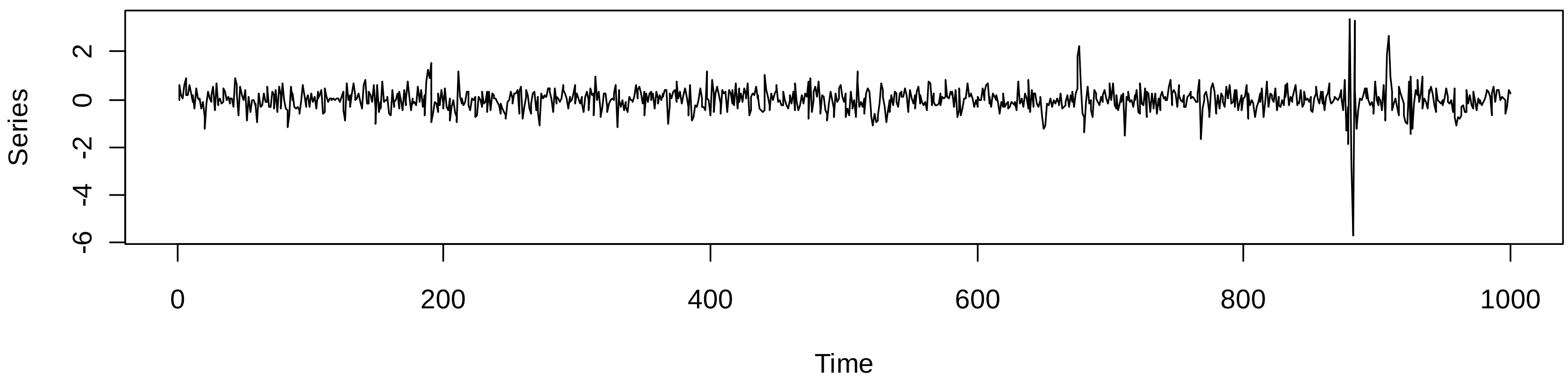
Simulated ARCH(1)-Process



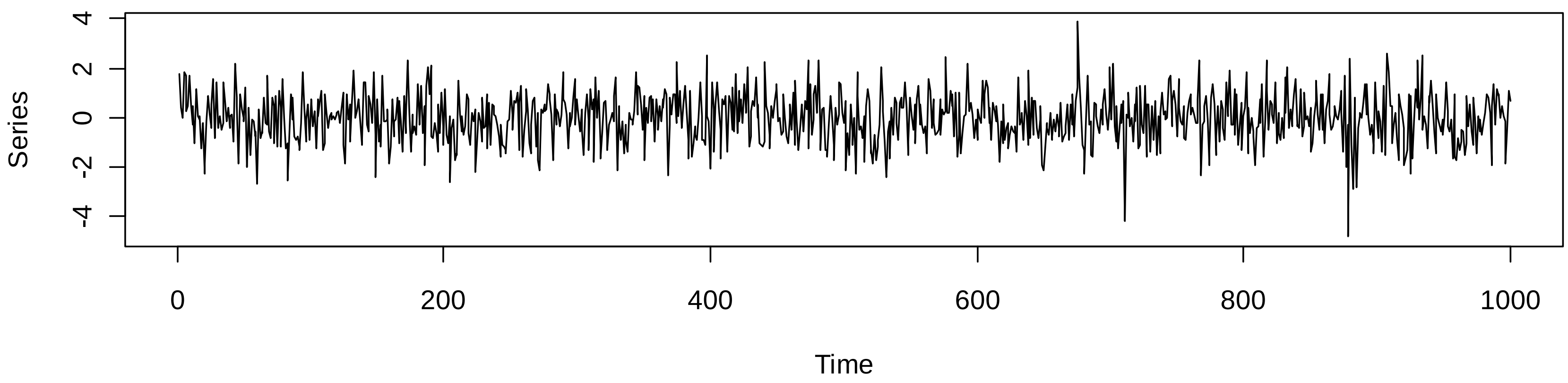
Applied Time Series Analysis
FS 2011 – Week 14

Residuals from a Fitted ARCH(1)

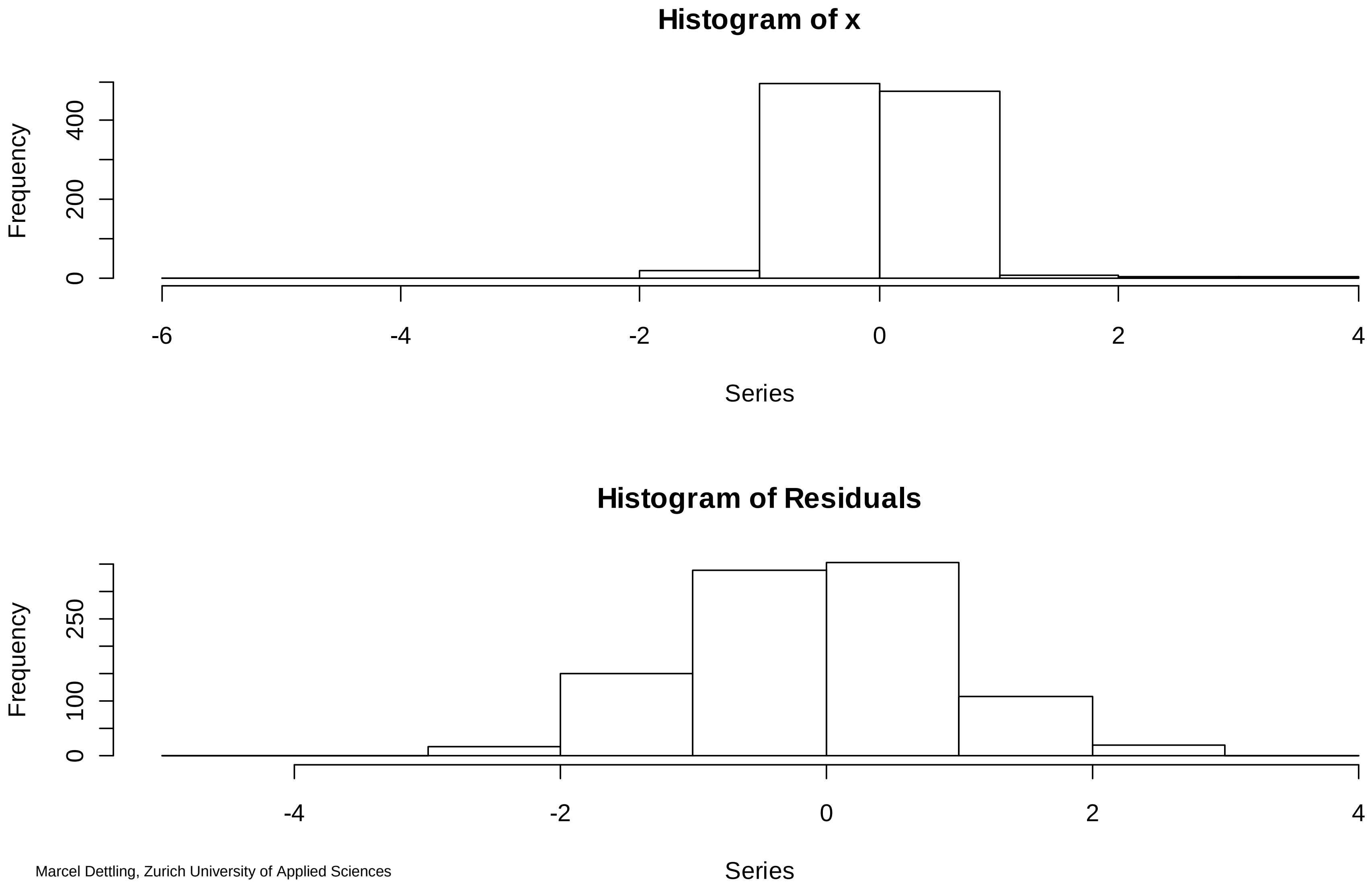
x



Residuals

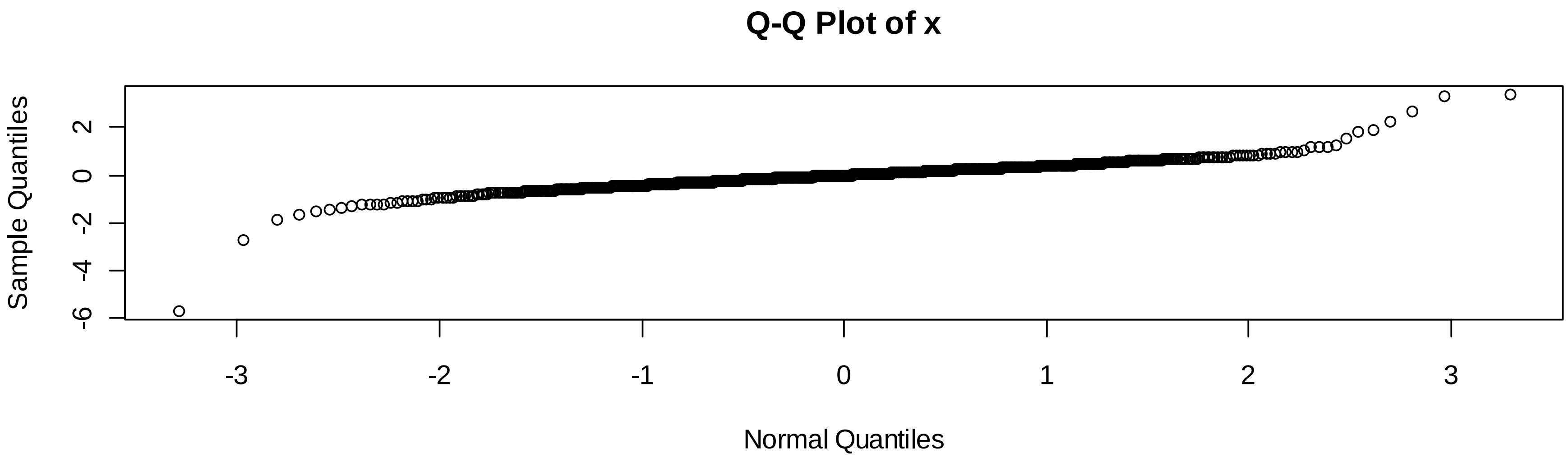


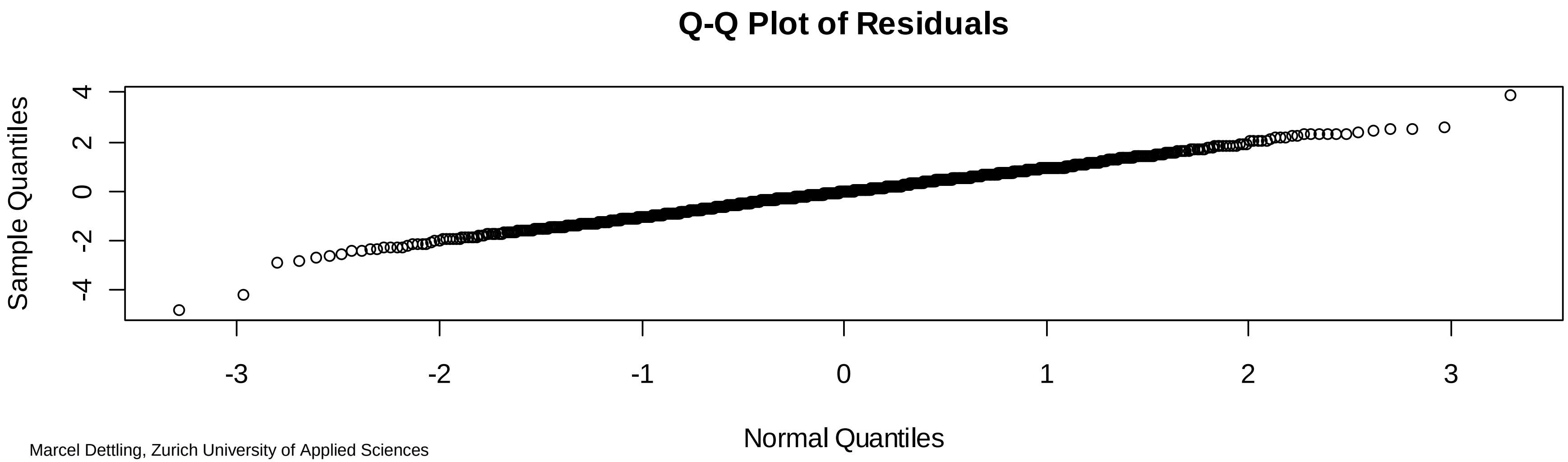
ARCH(1) is Long-Tailed



Applied Time Series Analysis
FS 2011 – Week 14

ARCH(1) is Long-Tailed

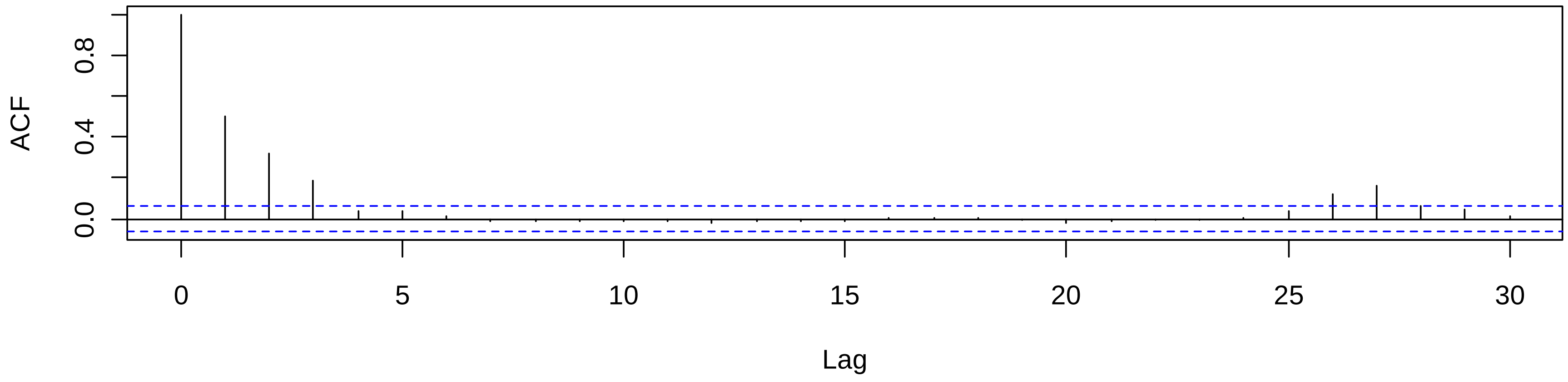




Applied Time Series Analysis
FS 2011 – Week 14

Dependency in the Squared Series

ACF of Squared x



ACF of Squared Residuals