Applied Time Series Analysis
SS 2013 – Week 11

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 6, 2013

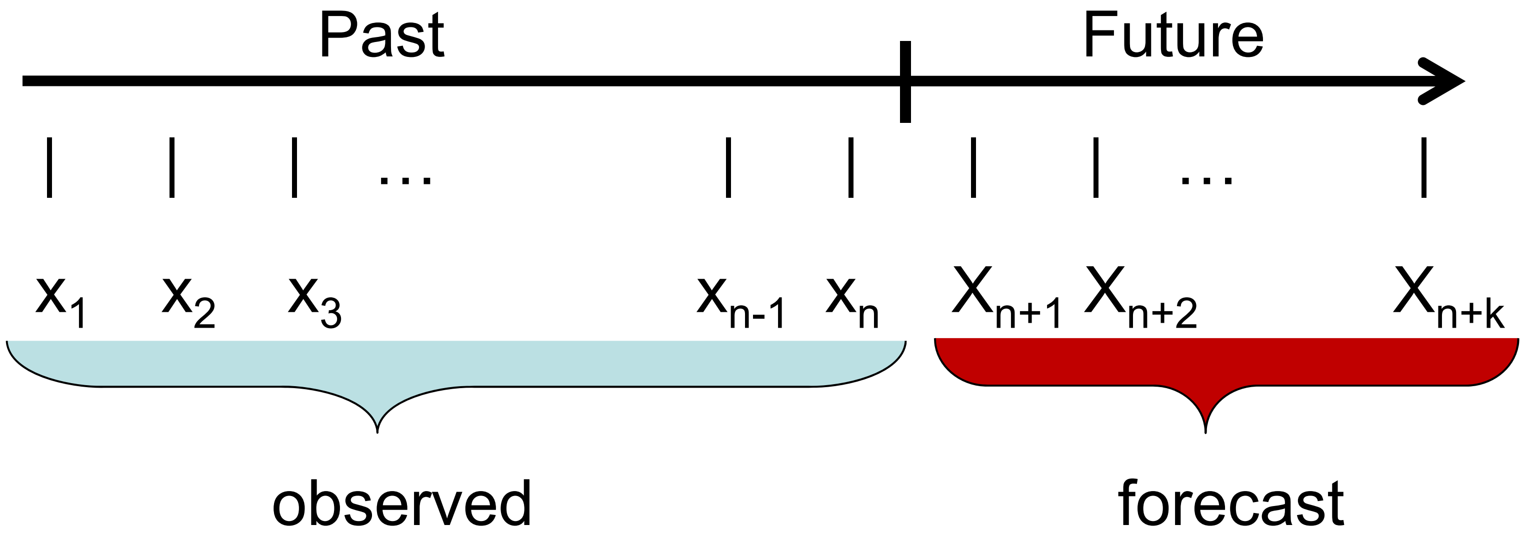
Applied Time Series Analysis
SS 2013 – Week 11

Forecasting with Time Series

Goal: Prediction of future observations with a measure of uncertainty (confidence interval)

Note:
- will be based on a stochastic model
- builds on the dependency structure and past data
- is an extrapolation, thus to take with a grain of salt
- similar to driving a car by using the rear window mirror

Forecasting, More Technical



observations

$$\{x_1, \ldots, x_n\} = X_1^n$$

estimates

$$\left\{\hat{X}_{n+1,n}, \ldots, \hat{X}_{n+k,n}\right\}$$

Applied Time Series Analysis
SS 2013 – Week 11

Sources of Uncertainty in Forecasting

There are 4 main sources of uncertainty:

1) Does the data generating model from the past also apply in the future? Or are there any breaks?

2) Is the AR(p)-model we fitted to the data $\{x_1, \dots, x_n\}$ correctly chosen? What is the "true" order?

3) Are the parameters $\alpha_1, \dots, \alpha_p, \sigma_E^2$ and m accurately estimated? How much do they differ from the "truth"?

4) The stochastic variability coming from the innovation E_t

→ **we will here restrict to short-term forecasting!**

Applied Time Series Analysis
SS 2013 – Week 11

How to Forecast?

Probabilistic principle for point forecasts:

$$\hat{X}_{n+k,n} = E\left[X_{n+k} \mid X_1^n \right]$$

→ we forecast the expected value, given our observations

Probabilistic principle for prediction intervals:

$$Var\left(X_{n+k} \mid X_1^n \right)$$

→ we use the conditional variance

MA(1) Forecasting: Summary

- We have seen that for an MA(1)-process, the k-step forecast for k>1 is equal to m.

- In case of k=1, we obtain for the MA(1)-forecast:

$$\hat{X}_{n+1,n} = m + \beta_1 \cdot E[E_n \mid X_1^n]$$

  The conditional expectation is (too) difficult to compute

- As a trick, we not only condition on observations 1,…,n, but on the infinite past:

$$e_n := E[E_n \mid X_{-\infty}^n]$$

MA(1) Forecasting: Summary

- We then write the MA(1) as an AR(∞) and solve the model equation for E_n :

$$E_n = \sum_{j=0}^{\infty} (-\beta_1)^j \cdot (X_{n-j} - m)$$

- In practice, we plug-in the time series observations x_{n-j} where available. For the „early" times, where we don't have observations, we plug-in $\hat{m}$.

- This is of course only an approximation to the true MA(1)-forecast, but it works well in practice, because of:

$$|\beta_1| < 1$$

Applied Time Series Analysis
SS 2013 – Week 11

ARMA(p,q) Forecasting

As with MA(1)/MA(q) forecasting, we face problems with

$$E[E_{n+1-j} \mid X_{-\infty}^{n}]$$

which is difficult to compute. We use the same tricks as for MA(1) and obtain

$$\hat{X}_{n+k,n} = m + \sum_{i=1}^{p} \alpha_i (E[X_{n+k-i} \mid X_{-\infty}^{n}] - m)$$
$$+ E[E_{n+k} \mid X_{-\infty}^{n}] - \sum_{j=1}^{q} \beta_j E[E_{n+k-j} \mid X_{-\infty}^{n}]$$

where …

ARMA(p,q) Forecasting

…where

$$E[X_t \mid X_{-\infty}^{n}] = \begin{cases} x_t & \text{if } t \le n \\ \hat{X}_{t,n} & \text{if } t > n \end{cases}$$

and

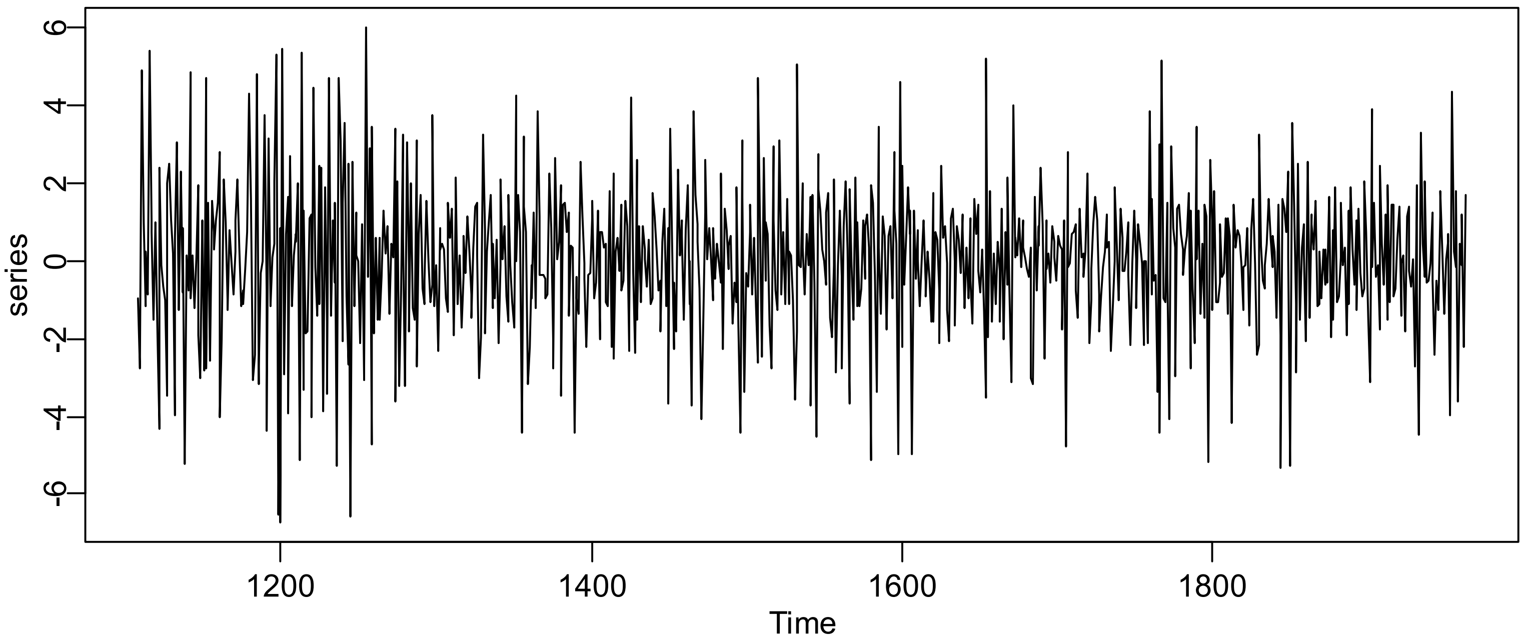
$$E[E_t \mid X_{-\infty}^{n}] = \begin{cases} e_t & \text{if } t \le n \\ 0 & \text{if } t > n \end{cases}$$

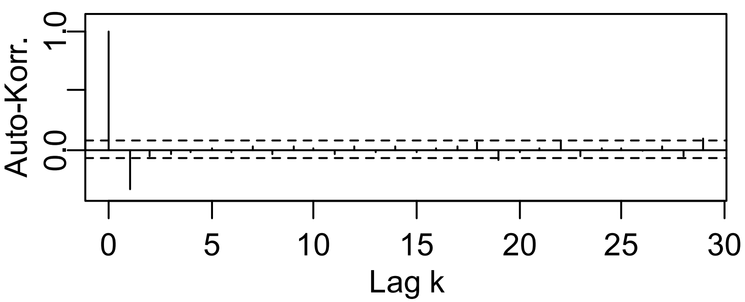
with

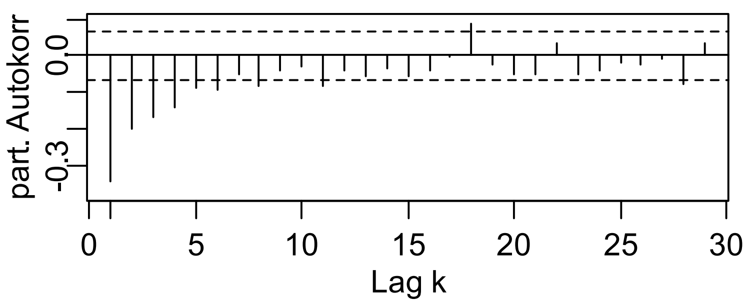
$$e_t = x_t - m - \sum_{i=1}^{p} \alpha_i (x_{t-i} - m) + \sum_{j=1}^{q} \beta_j e_{t-j}$$

Applied Time Series Analysis
SS 2013 – Week 11

ARMA(p,q) Forecasting: Douglas Fir

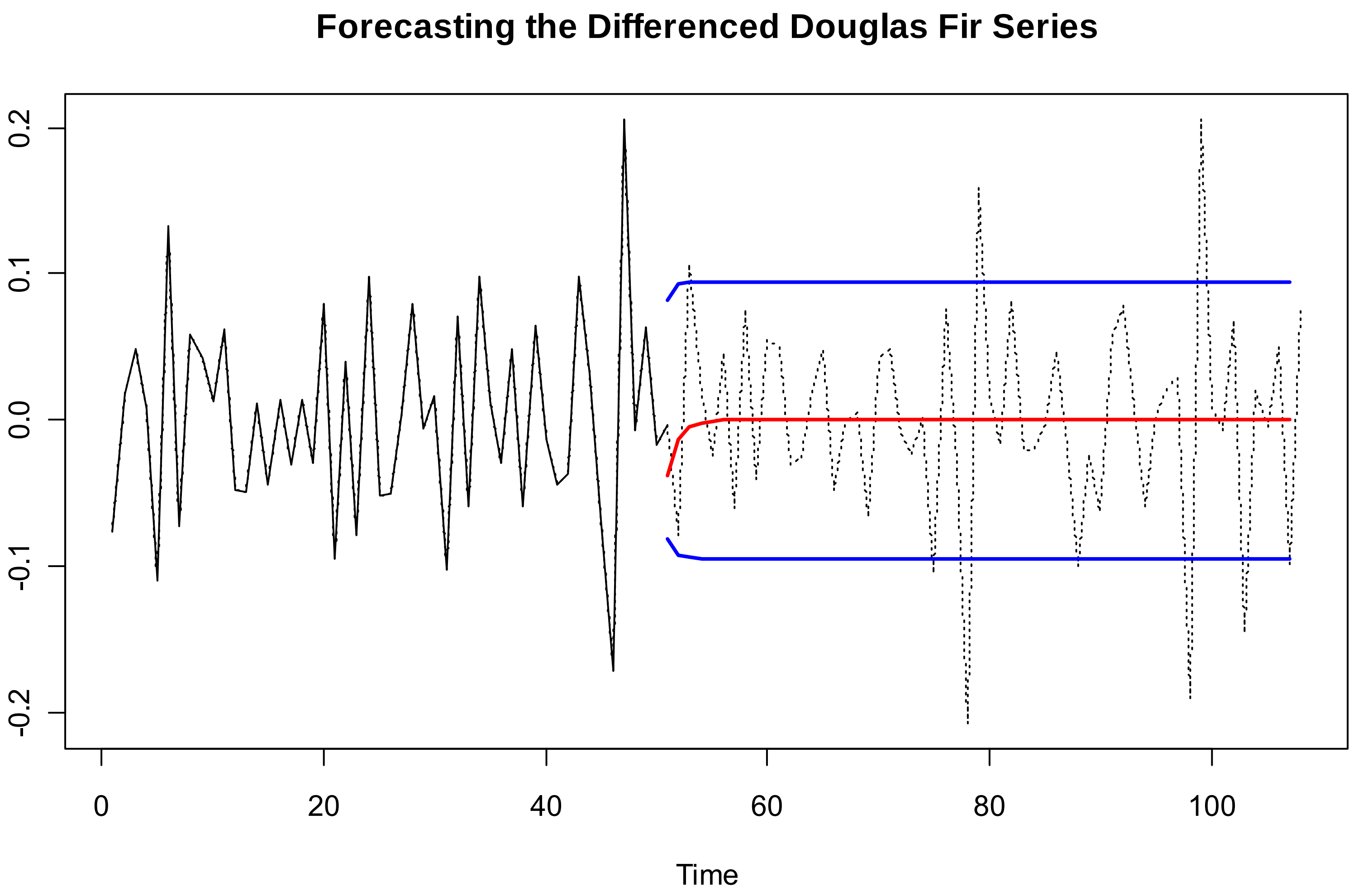






Applied Time Series Analysis
SS 2013 – Week 11

ARMA(p,q) Forecasting: Example



Forecasting Decomposed Series

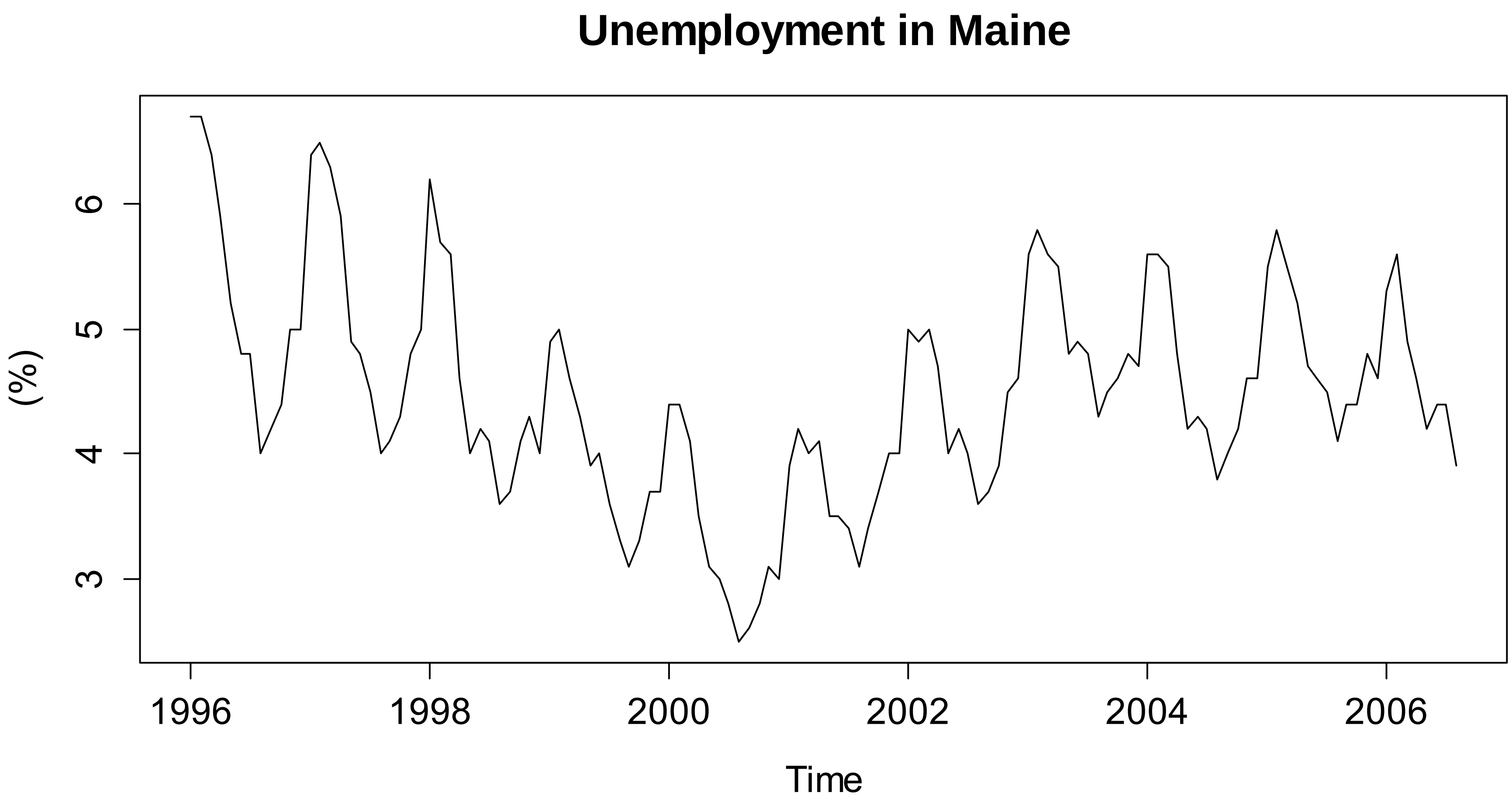
The principle for forecasting time series that are decomposed into trend, seasonal effect and remainder is:

1) **Stationary Remainder**
   Is usually modelled with an ARMA(p,q), so we can generate a time series forecast with the methodology from before.

2) **Seasonal Effect**
   Is assumed as remaining “as is”, or “as it was last” (in the case of evolving seasonal effect) and extrapolated.

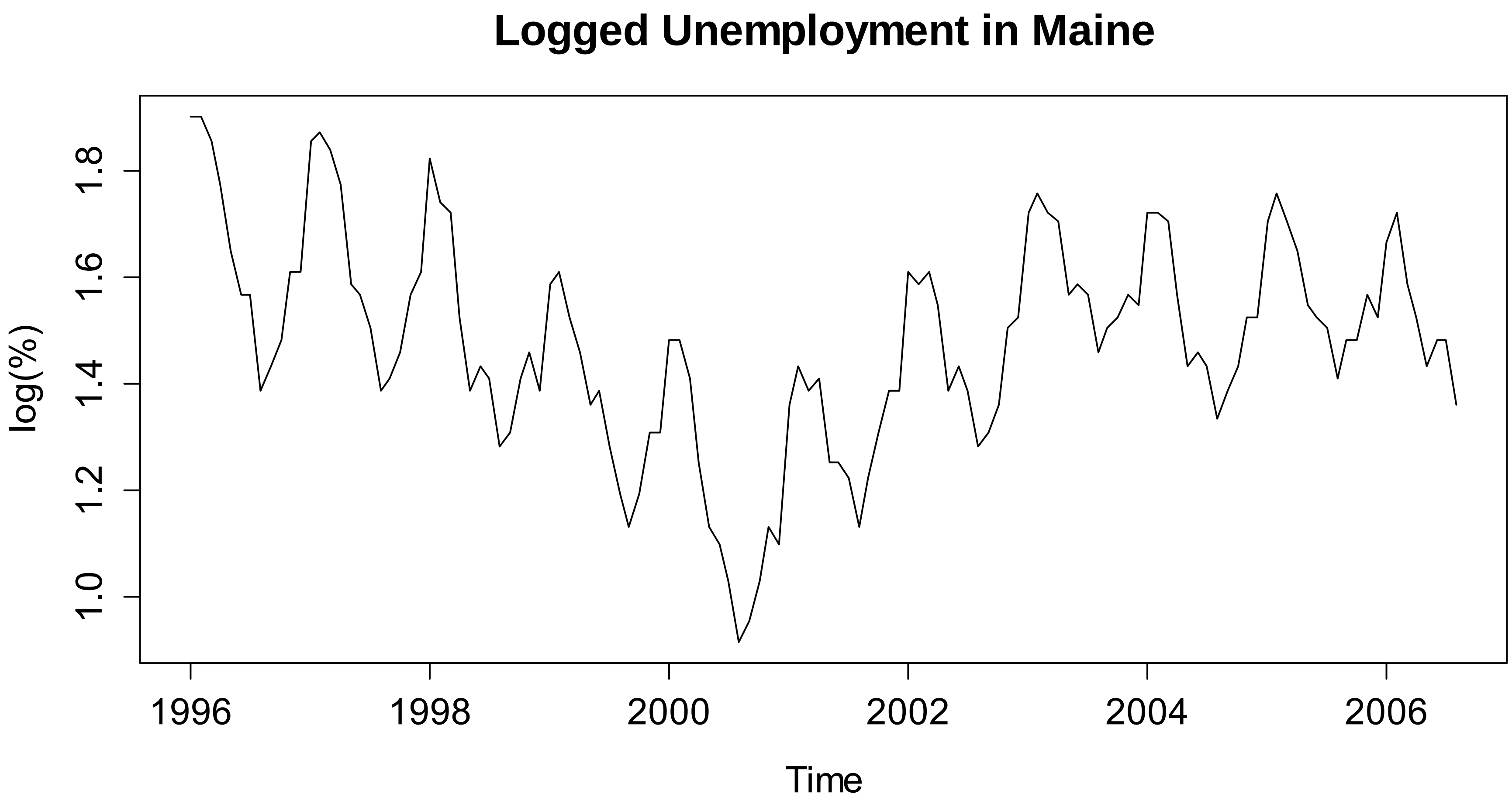
3) **Trend**
   Is either extrapolated linearly, or sometimes even manually.

Forecasting Decomposed Series: Example

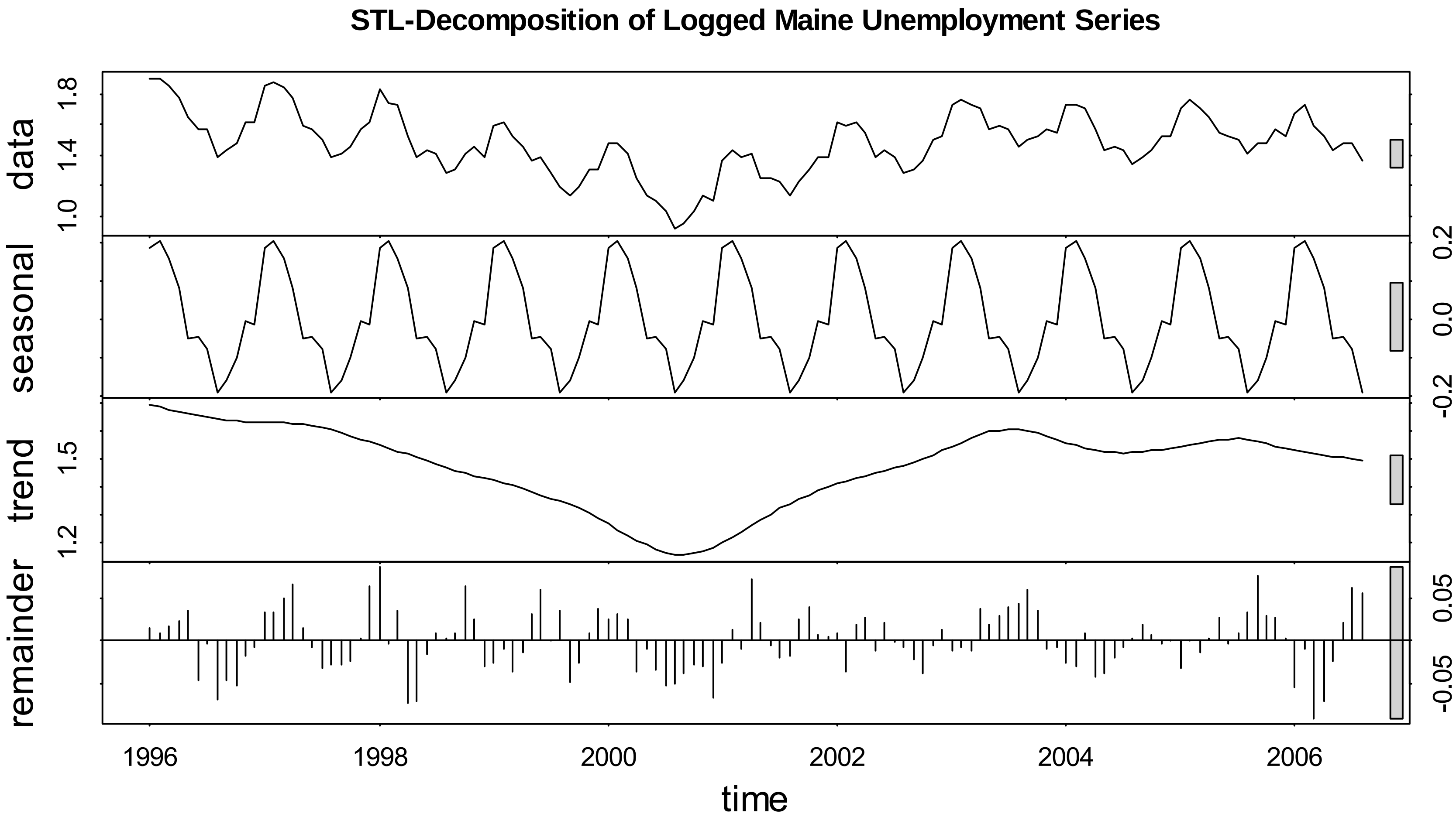


Applied Time Series Analysis
SS 2013 – Week 11

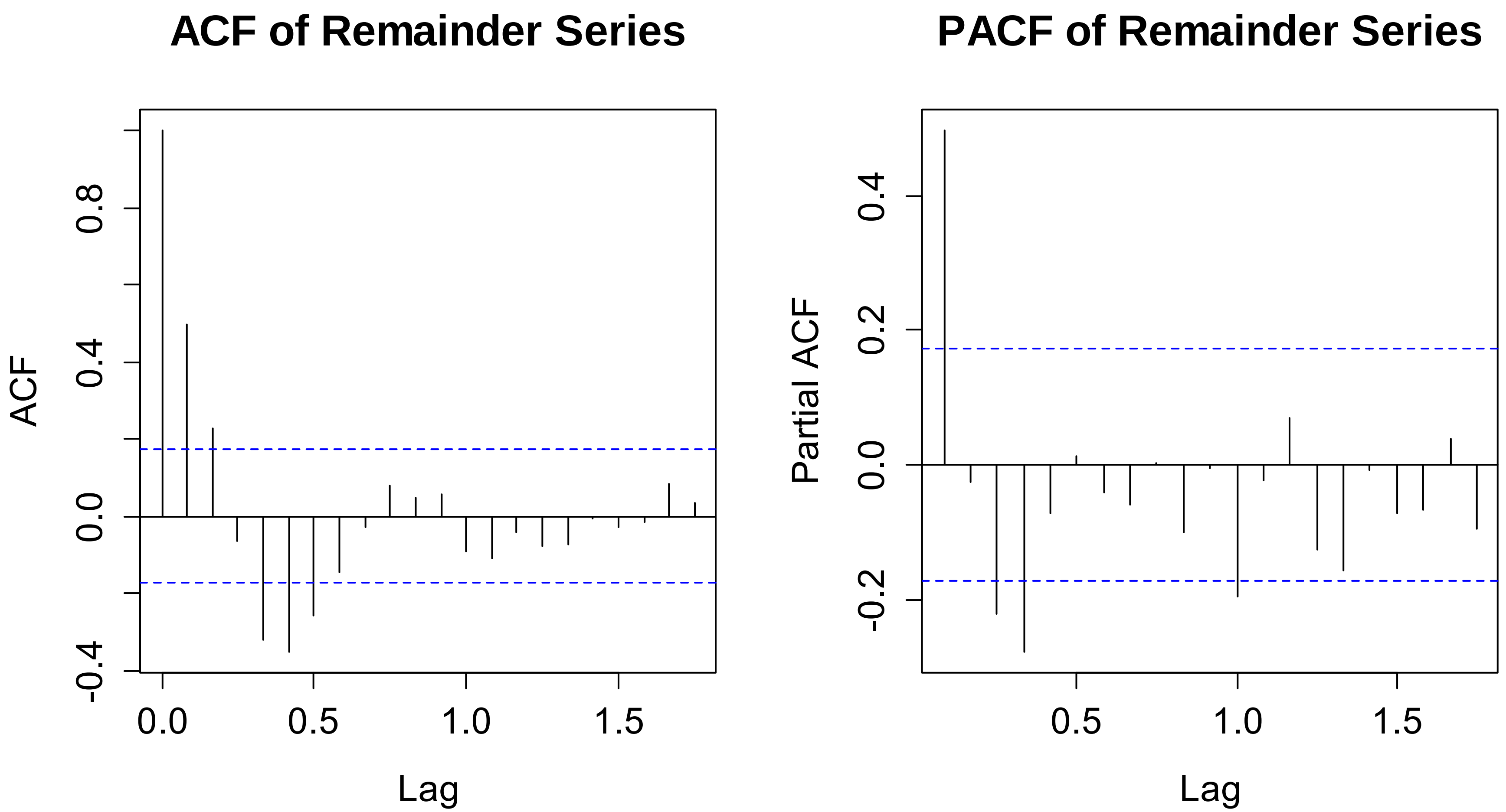
Forecasting Decomposed Series: Example



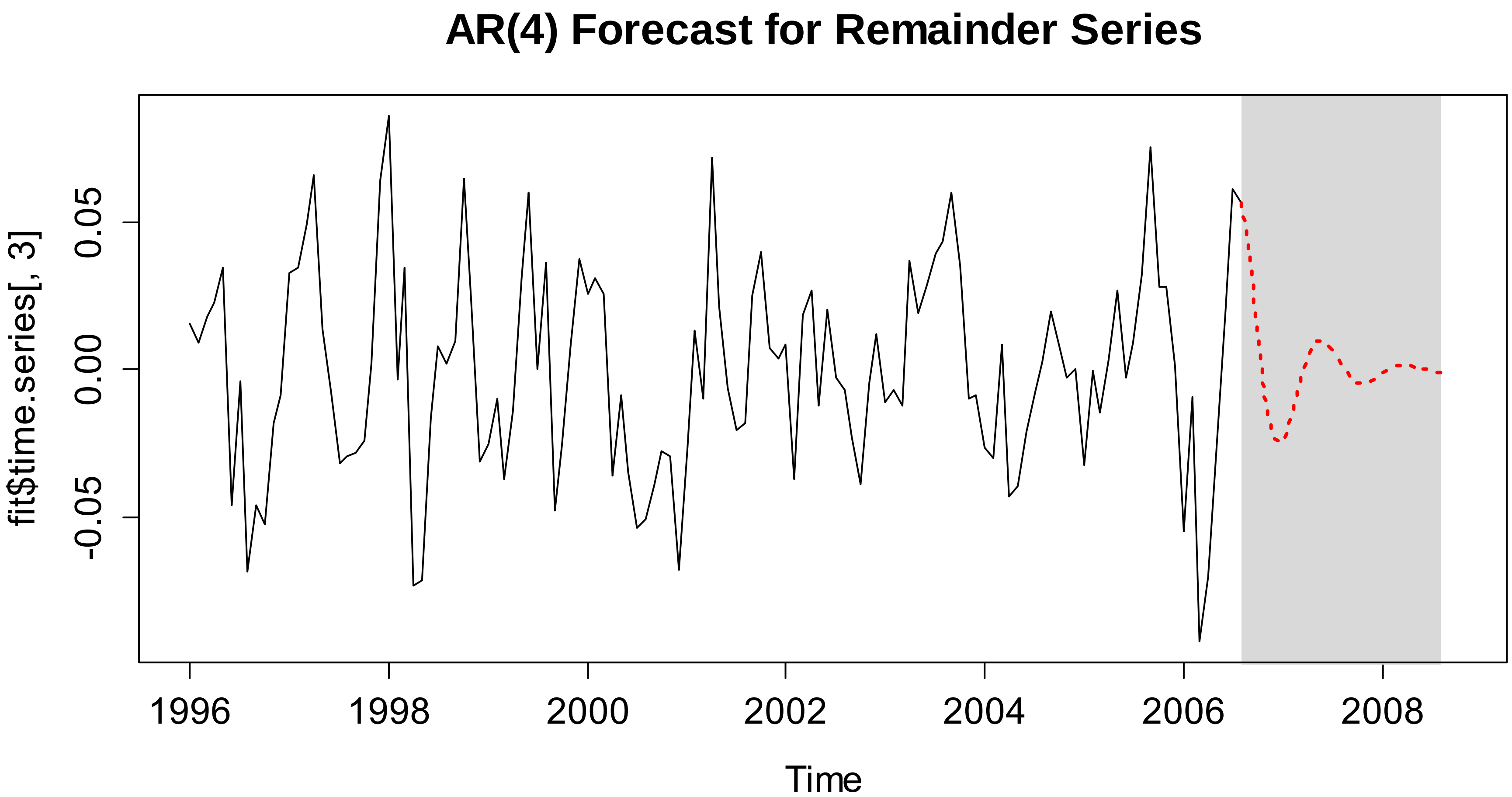
Forecasting Decomposed Series: Example



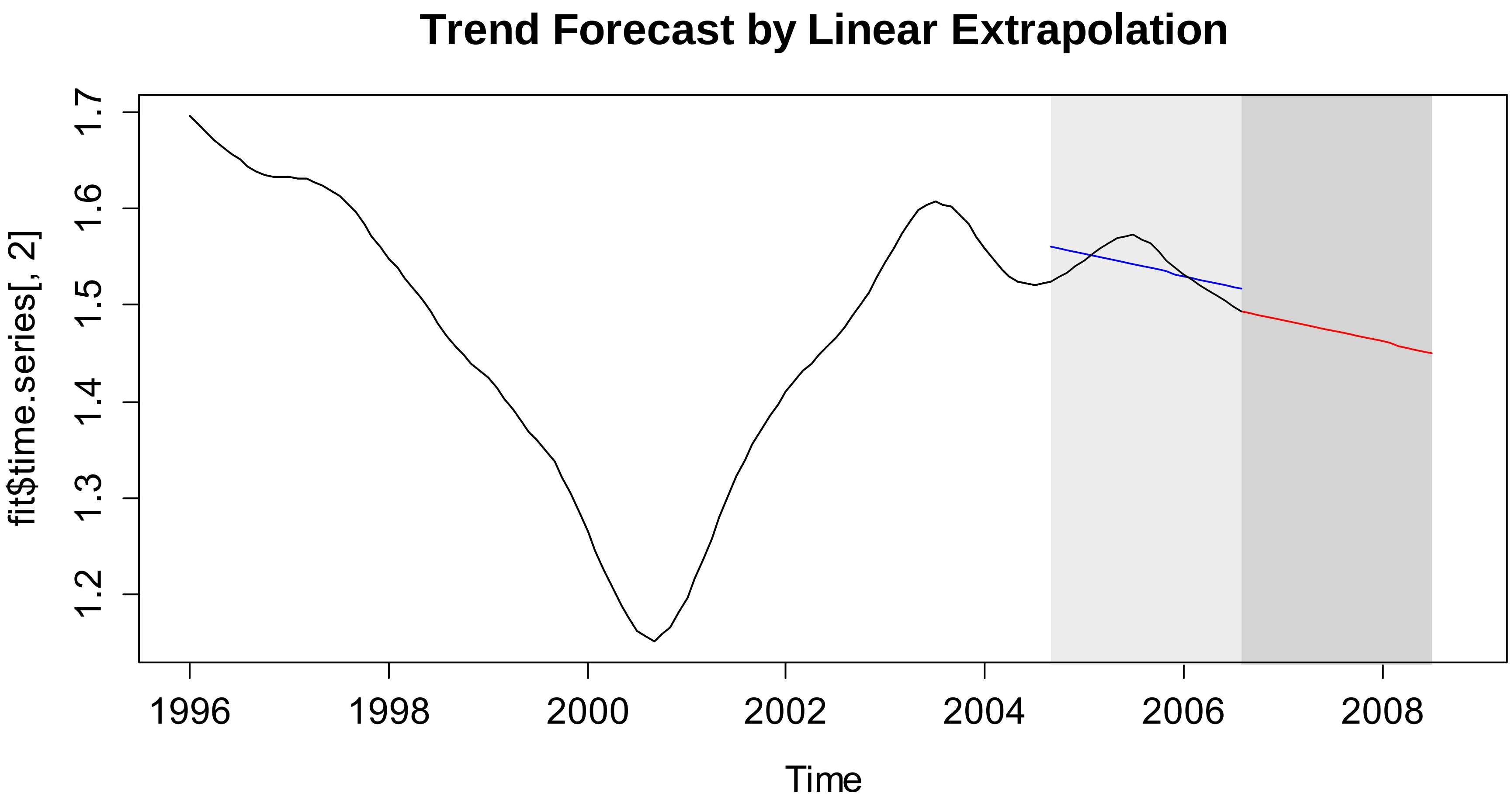
Forecasting Decomposed Series: Example



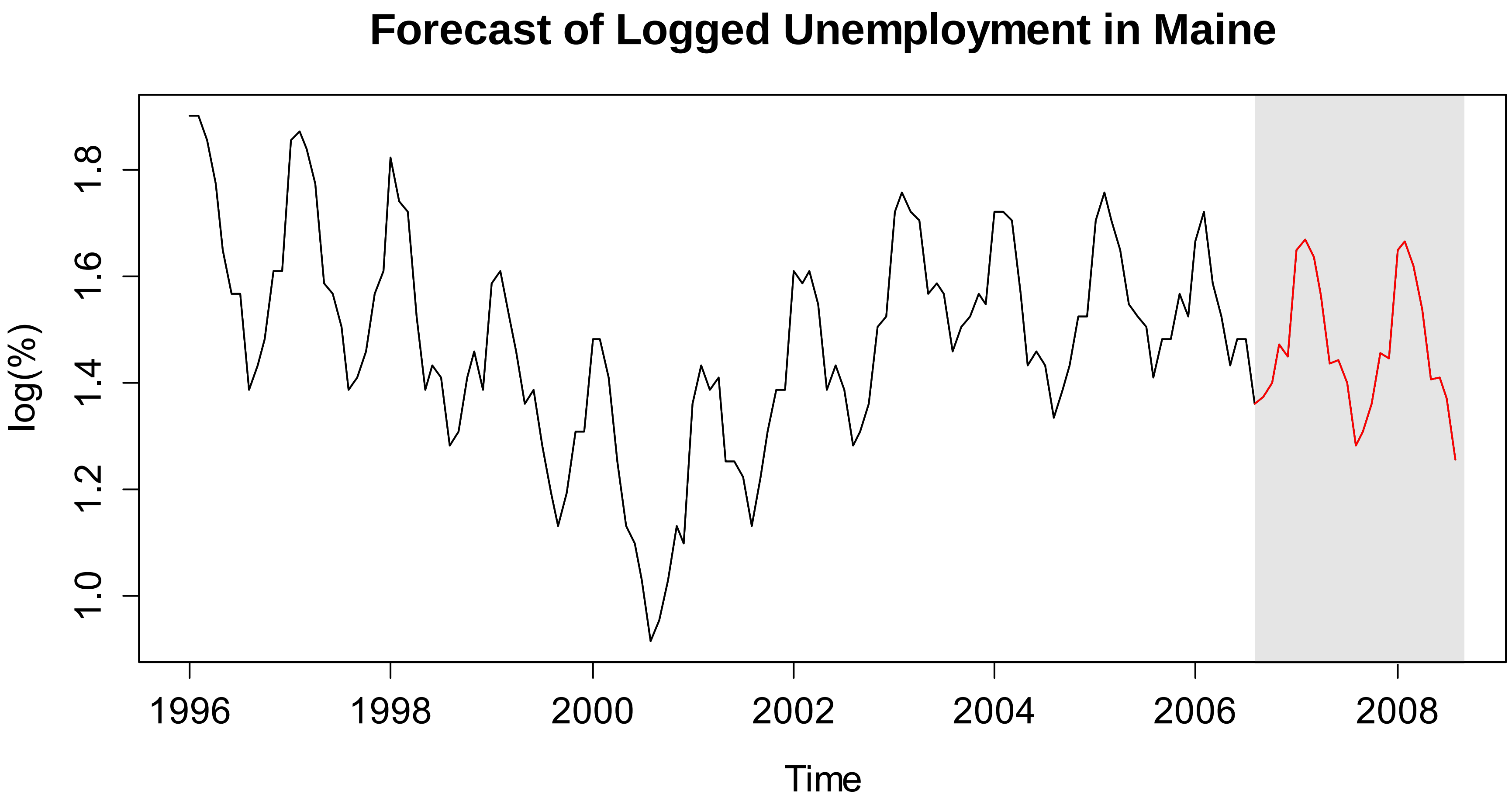
Forecasting Decomposed Series: Example



Forecasting Decomposed Series: Example



Forecasting Decomposed Series: Example



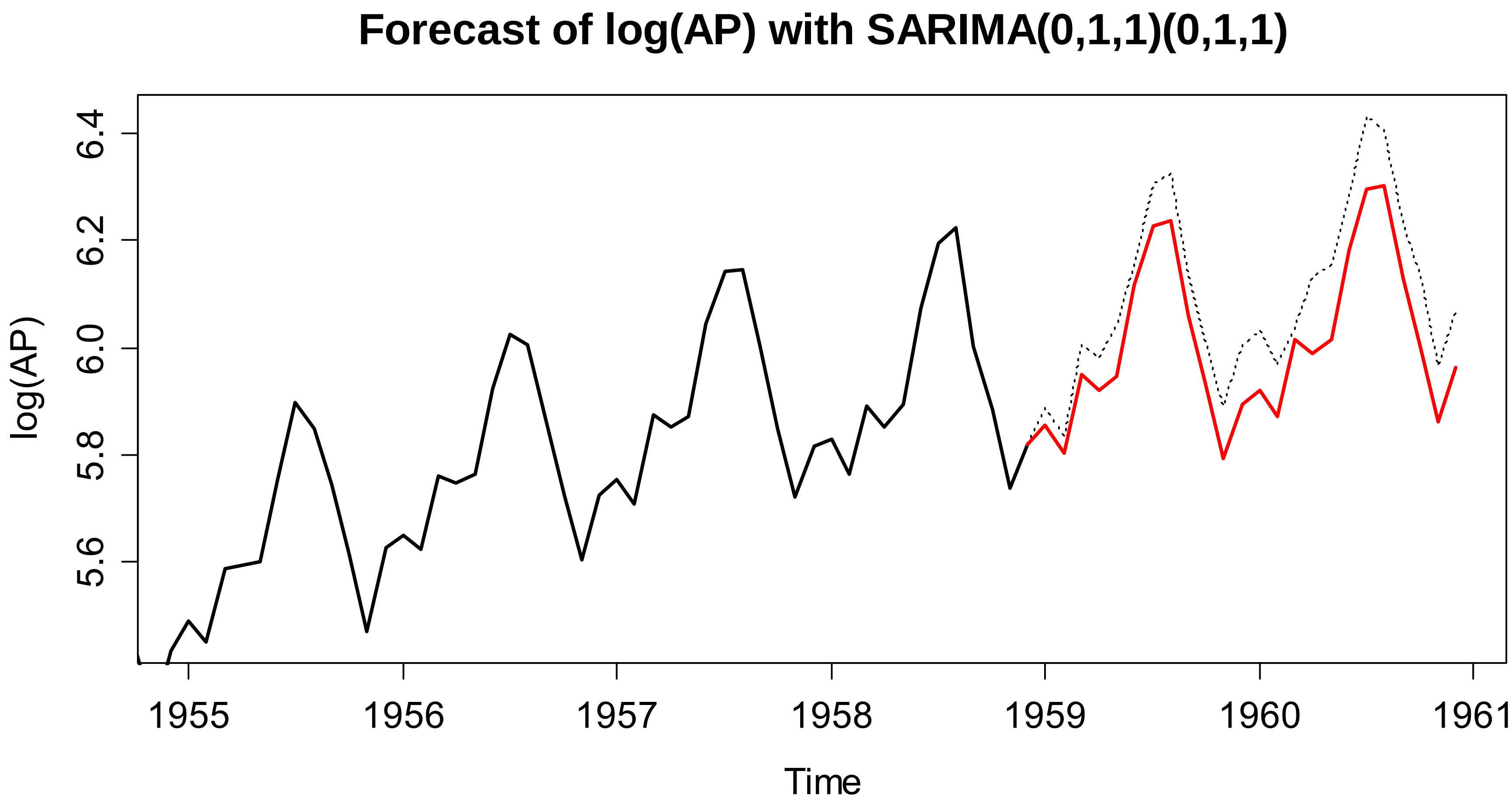
Forecasting with SARIMA

We have seen that forecasting decomposed series can be a somewhat laborious process. In R, it is easier and quicker to use a SARIMA model for forecasting season/trend-series.

- The SARIMA model is fitted in R as usual. Then, we can simply employ the `predict()` command and obtain the forecast plus a prediction interval.
- Technically, the forecast comes from the non-stationary ARMA(p,q)-formulation of the SARIMA model.
- The disadvantage of working with SARIMA forecasts is that it is much more of a **black box approach** than the one before!

Applied Time Series Analysis
SS 2013 – Week 11

Forecasting with SARIMA: Example



Exponential Smoothing

Simple exponential smoothing:

- works for stationary time series without trend & season
- is a heuristic, model-free approach
- further in the past -> less weight in the forecast

Turns out to yield these forecasts:

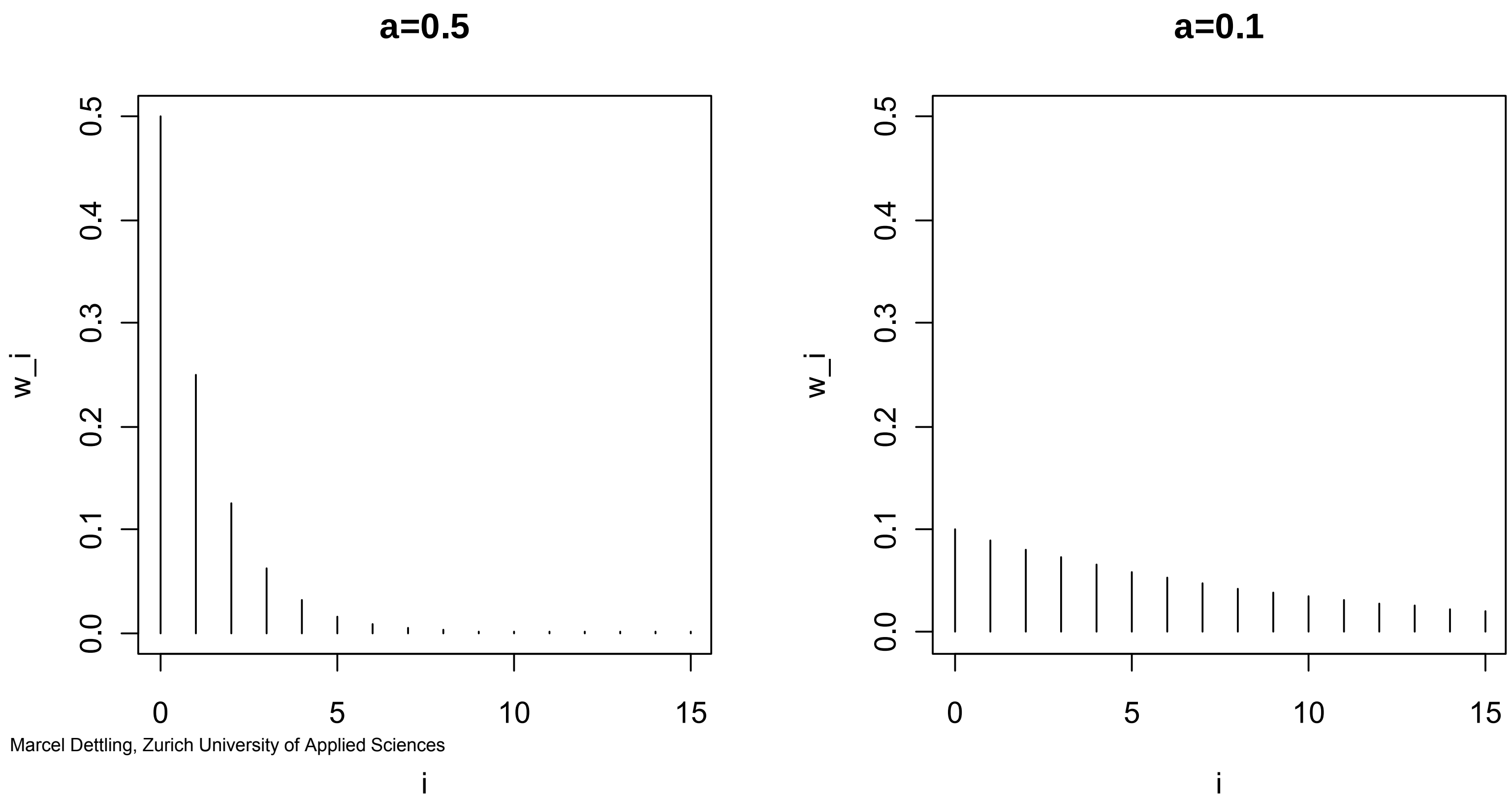
$$\hat{X}_{n+1,n} = \sum_{i=0}^{n-1} w_i x_{n-i} \quad \text{where } w_0 \geq w_1 \geq w_2 \geq \ldots \geq 0 \text{ and } \sum_{i=0}^{n-1} w_i = 1$$

→ **See the blackboard for the derivation...**

Choice of Weights

An usual choice are exponentially decaying weights:

$$w_i = a(1-a)^i \quad \text{where } a \in (0,1)$$



Forecasting with Exponential Smoothing

The 1-step forecast is:

$$\hat{X}_{n+1,n} = \underbrace{\sum_{i=0}^{n-1} a(1-a)^i \cdot x_{n-i}}_{\text{General Formula}} = \underbrace{ax_n + (1-a)\hat{X}_{n,n-1}}_{\text{“Update”-Formula}}$$

Remarks:

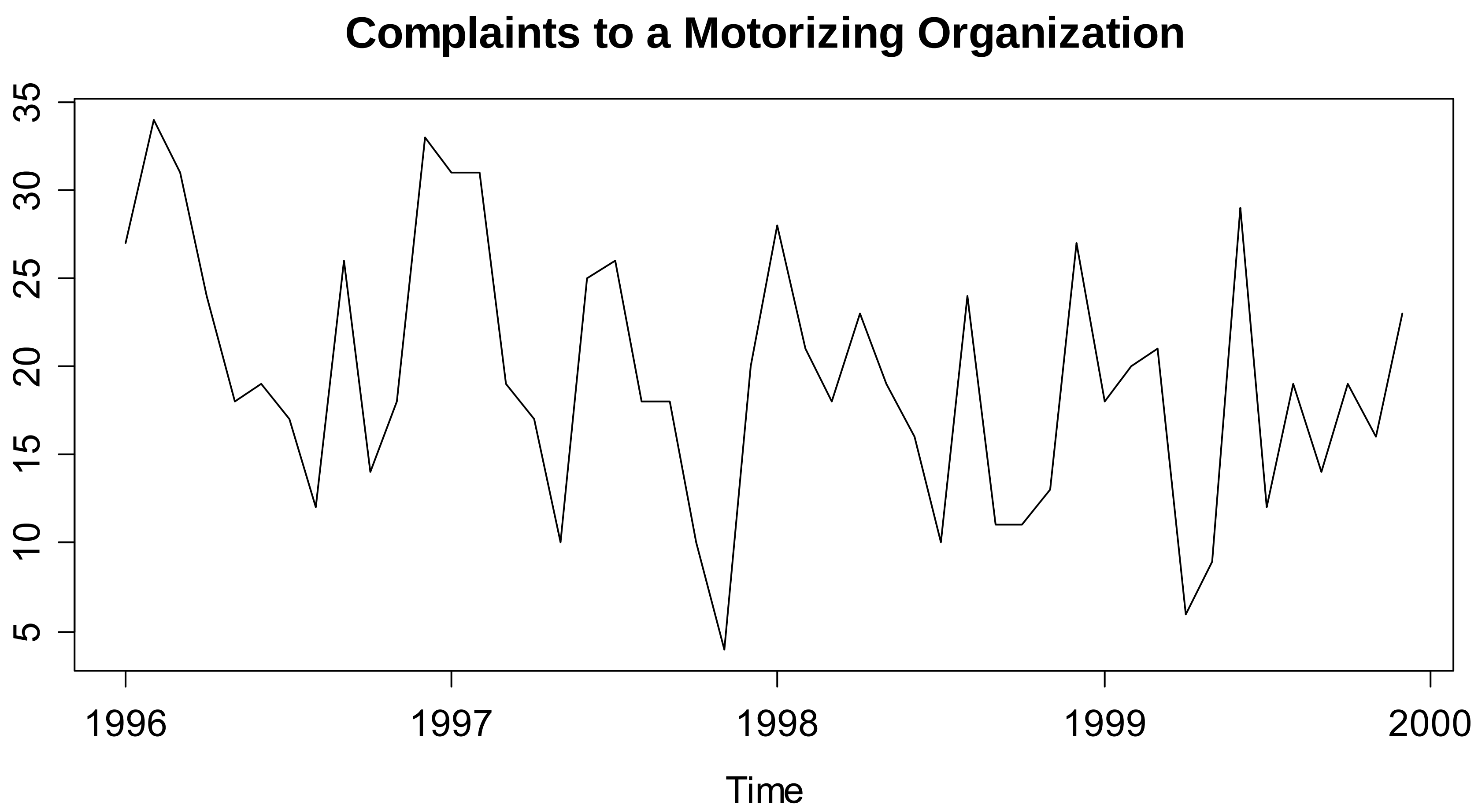
- in real applications (finite sum), the weights do not add to 1.
- the update-formula is useful if “new” observations appear.
- the k-step forecast is identical to the 1-step forecast.

Exponential Smoothing: Remarks

- the parameter a can be determined by evaluating forecasts that were generated from different a. We then choose the one resulting in the lowest sum of squared residuals.
- exponential smoothing is fundamentally different from AR(p)-forecasting. All past values are regarded for the 1-step forecast, but all k-step forecasts are identical to the 1-step.
- It can be shown that exponential smoothing can be optimal for MA(1)-models.
- there are double/triple exponential smoothing approaches that can deal with linear/quadratic trends.

Applied Time Series Analysis
SS 2013 – Week 11

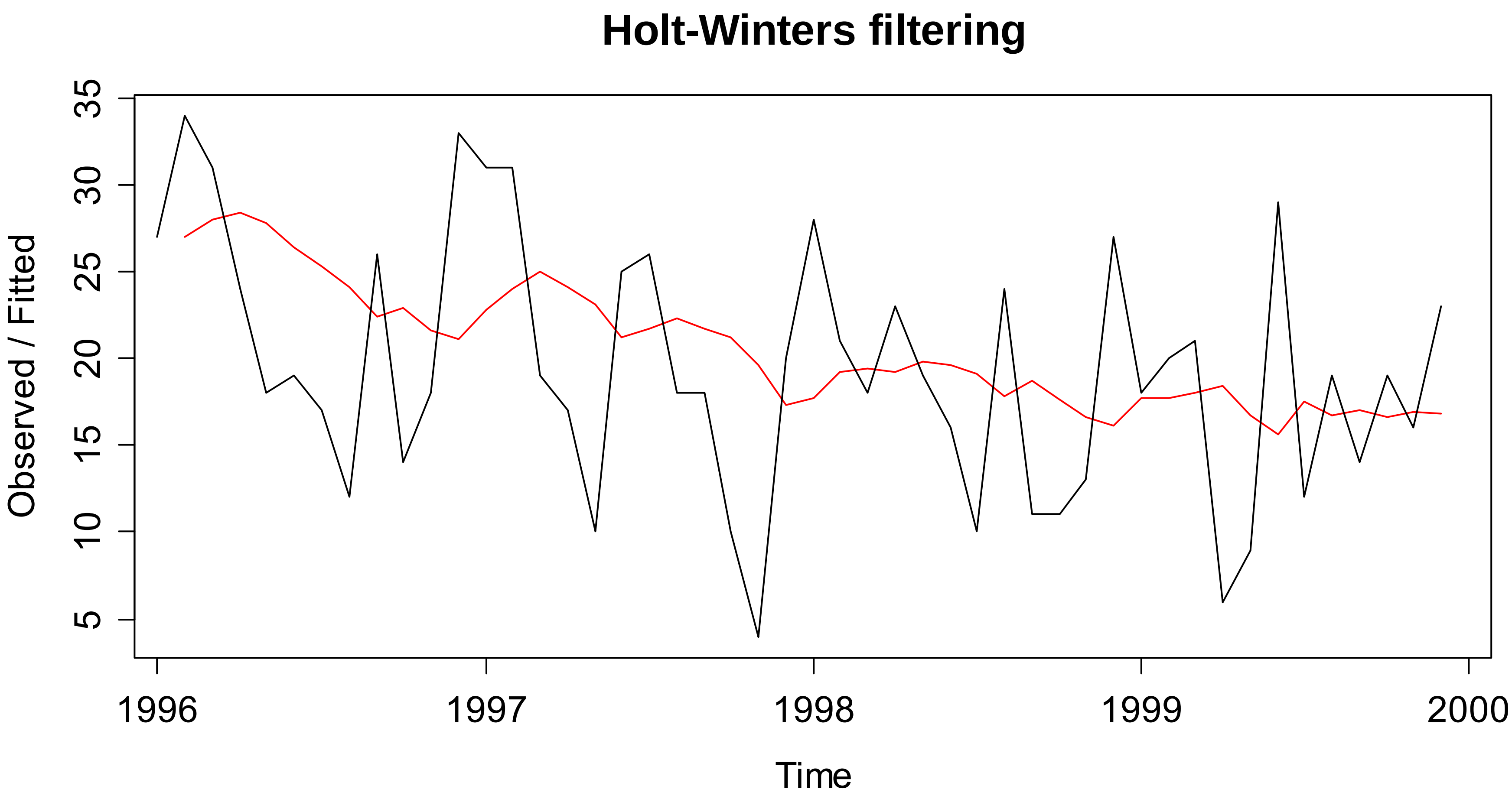
Exponential Smoothing: Example



Exponential Smoothing: Example

```
> fit  <- HoltWinters(cmpl, beta=F, gamma=F)

Holt-Winters exponential smoothing without trend
and without seasonal component.

Smoothing parameters:
 alpha:  0.1429622
 beta :  FALSE
 gamma:  FALSE

Coefficients:
      [,1]
a 17.70343
```

Exponential Smoothing: Example



Holt-Winters Method

Purpose:

- is for time series with deterministic trend and/or seasonality
- is still a heuristic, model-free approach
- again based on weighted averaging

Is based on these 3 formulae:

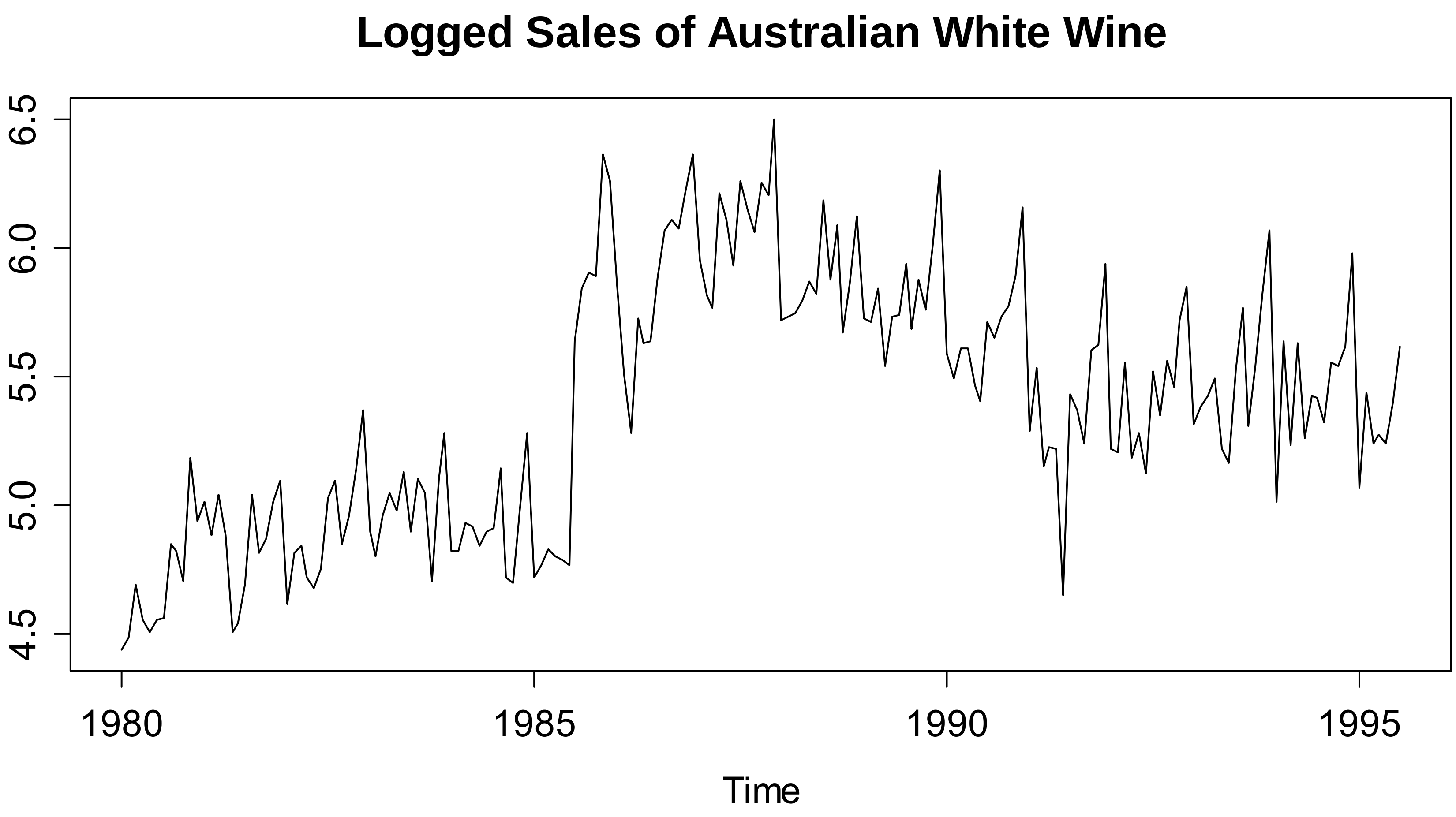
$$\begin{aligned} a_t &= \alpha(x_t - s_{t-p}) + (1-\alpha)(a_{t-1} + b_{t-1}) \\ b_t &= \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1} \\ s_t &= \gamma(x_t - a_t) + (1-\gamma)s_{t-p} \end{aligned}$$

→ **See the blackboard for the derivation...**

Holt-Winters: Example

Applied Time Series Analysis
SS 2013 – Week 11

Holt-Winters: Example



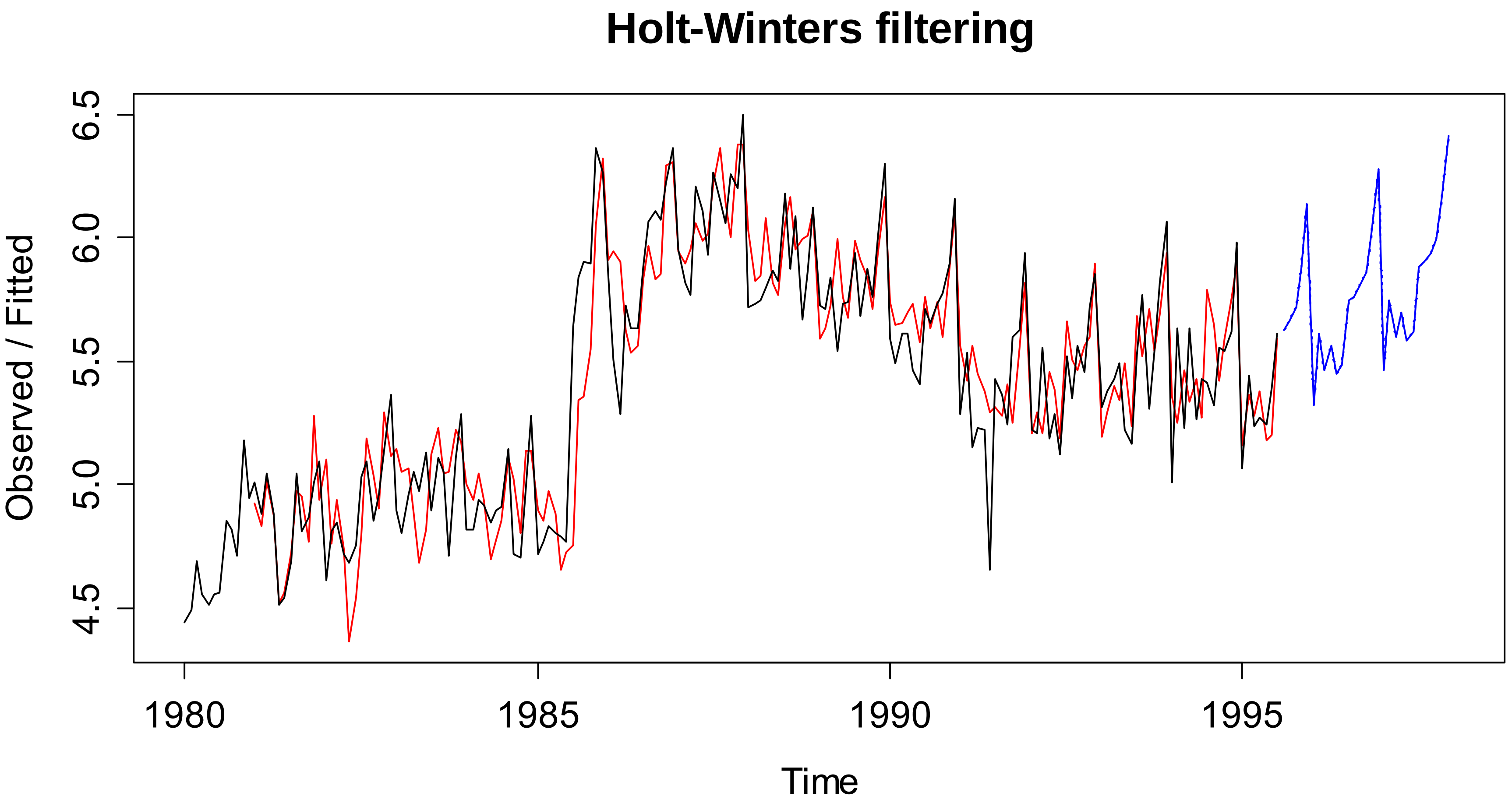
Applied Time Series Analysis
SS 2013 – Week 11

Holt-Winters: R-Code and Output

```
> HoltWinters(x = log(aww))

Holt-Winters exponential smoothing with trend and
additive seasonal component.

Smoothing parameters:
 alpha:  0.4148028; beta :  0; gamma:  0.4741967

Coefficients:
a     5.62591329; b    0.01148402
s1  -0.01230437; s2    0.01344762; s3   0.06000025
s4   0.20894897; s5    0.45515787; s6  -0.37315236
s7  -0.09709593; s8  -0.25718994; s9  -0.17107682
s10 -0.29304652; s11 -0.26986816; s12 -0.01984965
```

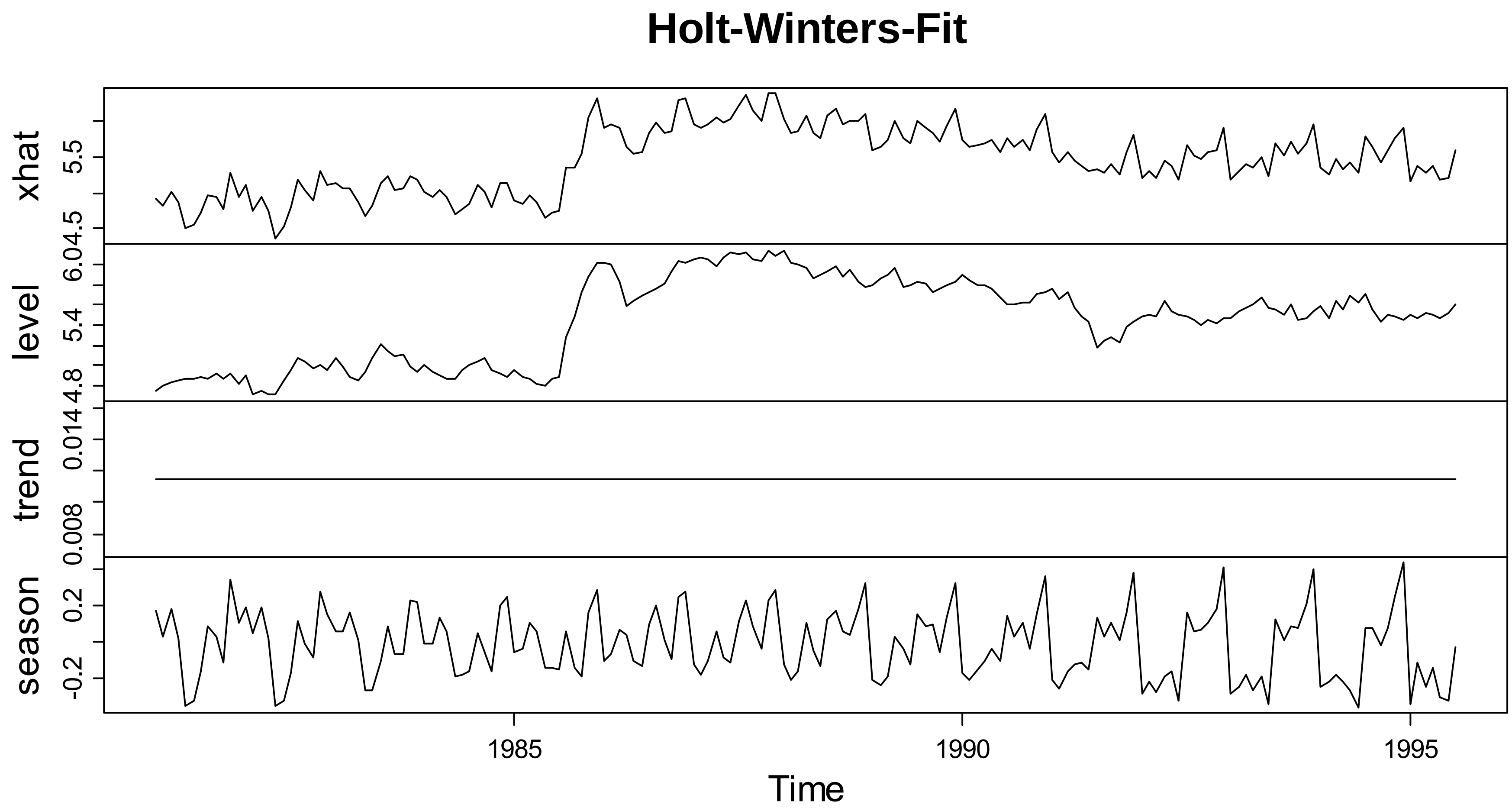
Applied Time Series Analysis
SS 2013 – Week 11

Holt-Winters: Fitted Values & Predictions



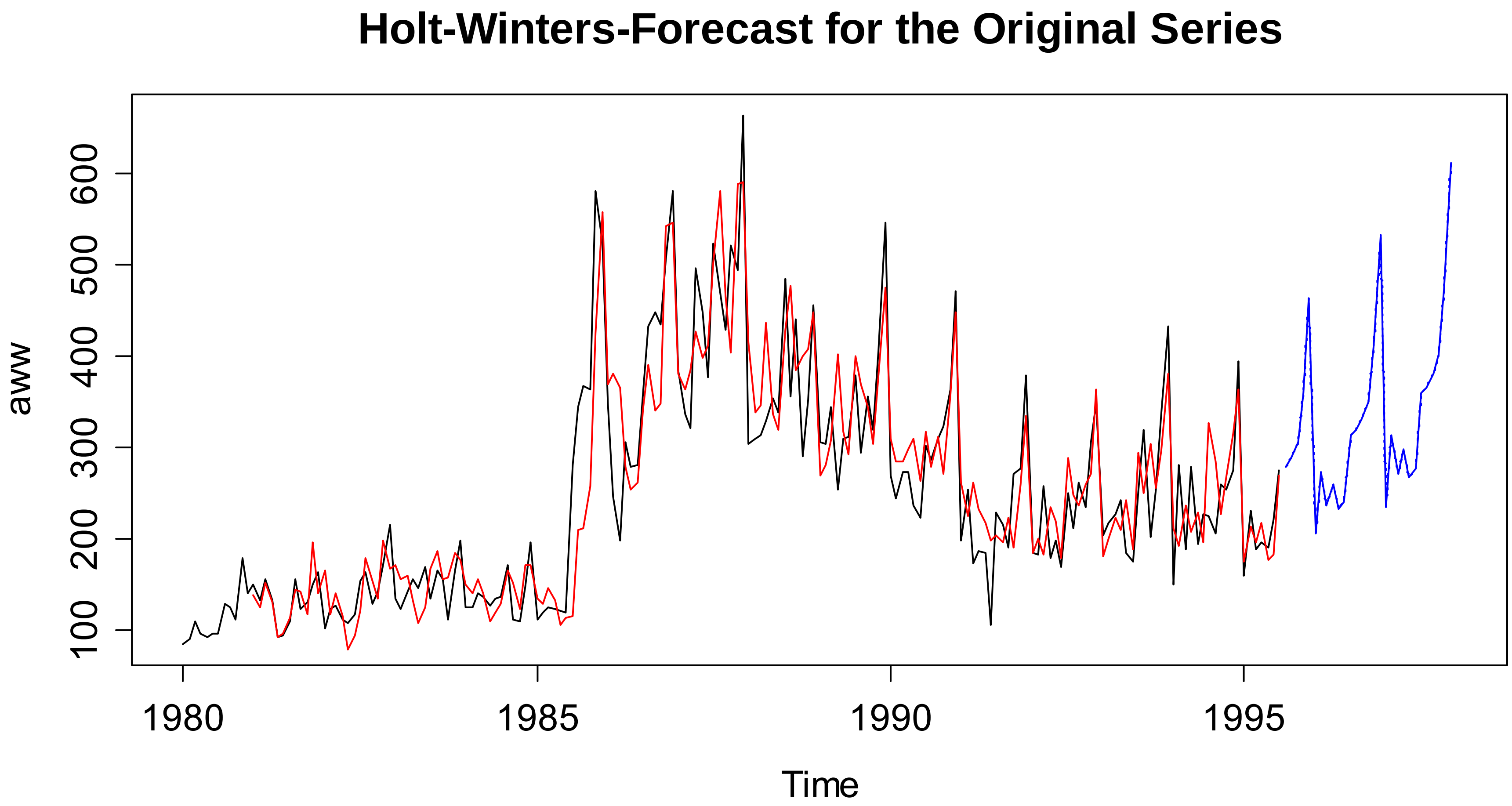
Applied Time Series Analysis
SS 2013 – Week 11

Holt-Winters: In-Sample Analysis



Applied Time Series Analysis
SS 2013 – Week 11

Holt-Winters: Predictions on Original Scale



Applied Time Series Analysis
SS 2013 – Week 11

Exercise

Data:

→ use the Australian white wine sales data...
→ ... or any other dataset you like

Goal:

- Find a good model describing these data
- Evaluate which model yields the best predictions
- Generate a 29-month forecast from this model

Method:

→ Remove the last 29 observations and mimic oos-forecasting