Applied Time Series Analysis
SS 2013 – Week 10

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, April 29, 2013

Applied Time Series Analysis
SS 2013 – Week 10

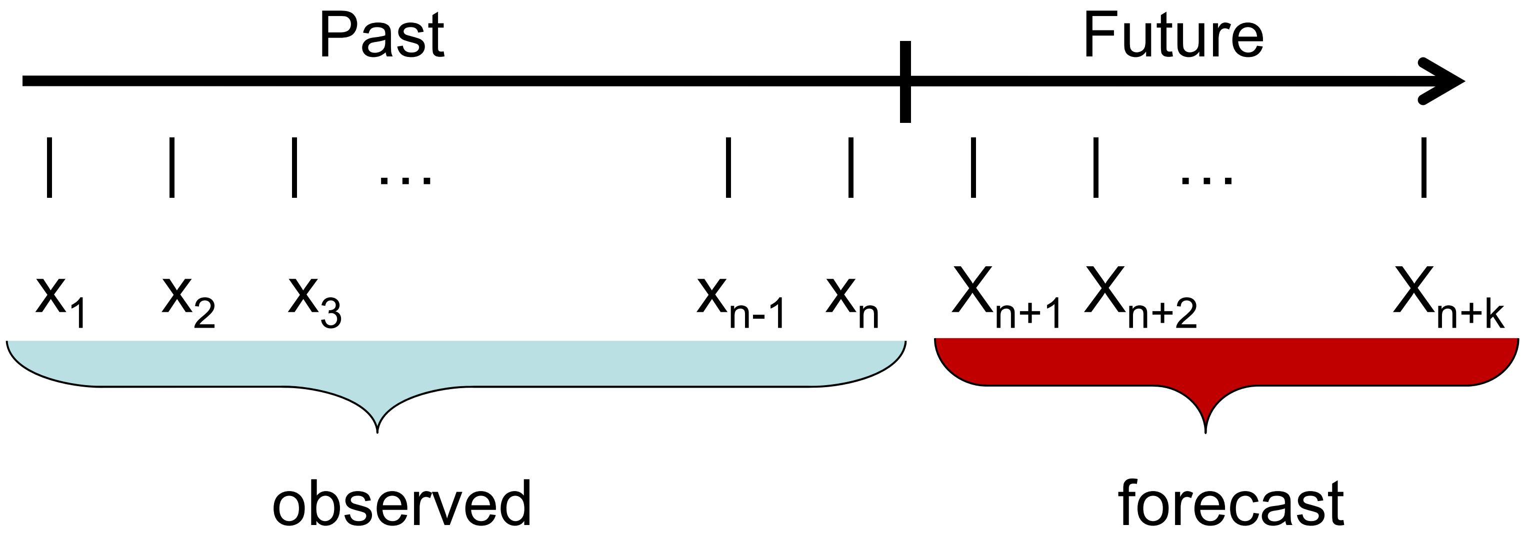
Forecasting with Time Series

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

Applied Time Series Analysis
SS 2013 – Week 10

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 10

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 $\{x_1, \ldots, x_n\}$ correctly chosen? What is the "true" order?

3) Are the parameters $\alpha_1, ..., \alpha_p, \sigma_E^2$ and μ 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 10

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

How to Apply the Principles?

- The principles provide a nice setup, but are only useful and practicable under additional assumptions.

- For stationary AR(1)-processes with normally distributed innovations, we can apply the principles and derive formulae

→ **see blackboard for the derivation!**

Applied Time Series Analysis
SS 2013 – Week 10

AR(1): 1-Step Forecast

The 1-step forecast for a shifted AR(1) process with mean m is:

$$\hat{X}_{n+1,n} = \alpha_1(x_n - m) + m$$

with prognosis interval

$$\hat{X}_{n+1,n} \pm 1.96 \cdot \sigma_E$$

Note that when $\hat{\alpha}_1, \hat{\mu}, \hat{\sigma}_E$ are plugged-in, this adds additional uncertainty which is not accounted for in the prognosis interval, i.e.

$$Var(\hat{X}_{n+1}) > Var(X_{n+1} \mid X_1^n)$$

Applied Time Series Analysis
SS 2013 – Week 10

Simulation Study

We have seen that the usual prognosis interval is too small. But by how much? A simulation study yields some insight:

Generated are 10‘000 1-step forecasts on a time series that was generated from an AR(1) process with $\alpha = 0.5$. The series length was variable.

The 95%-prognosis interval was determined and it was checked whether it included the true value or not. The empirically estimated confidence levels were:

n=20	n=50	n=100	n=200
91.01%	93.18%	94.48%	94.73%

AR(1): k-Step Forecast

The k-step forecast for an AR(1) process is:

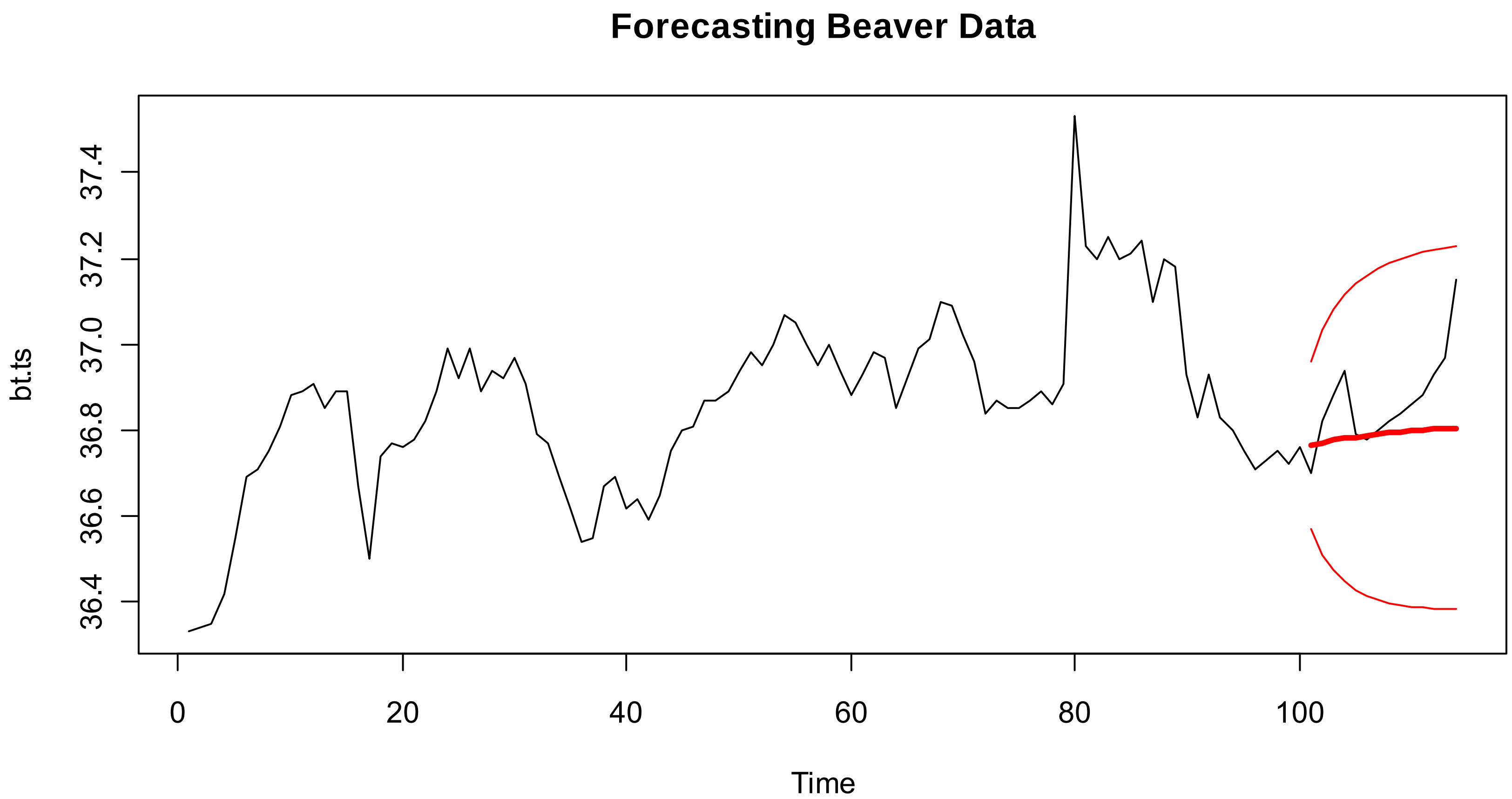
$$\hat{X}_{n+k,n} = \alpha_1^k (x_n - m) + m$$

with prognosis interval based on

$$Var(X_{n+k,n} \mid X_1^n) = \left(1 + \sum_{j=1}^{k-1} \alpha^{2j}\right) \cdot \sigma_E^2$$

It is important to note that for $k \to \infty$, the expected value and the variance from above go to μ and σ_X^2 respectively.

Forecasting the Beaver Data



Applied Time Series Analysis
SS 2013 – Week 10

Forecasting AR(p)

The principle is the same, forecast and prognosis interval are:

$$E[X_{n+k} \mid X_1^n] \text{ and } Var(X_{n+k} \mid X_1^n)$$

The computations are more complicated, but do not yield any further insight. We are thus doing without.

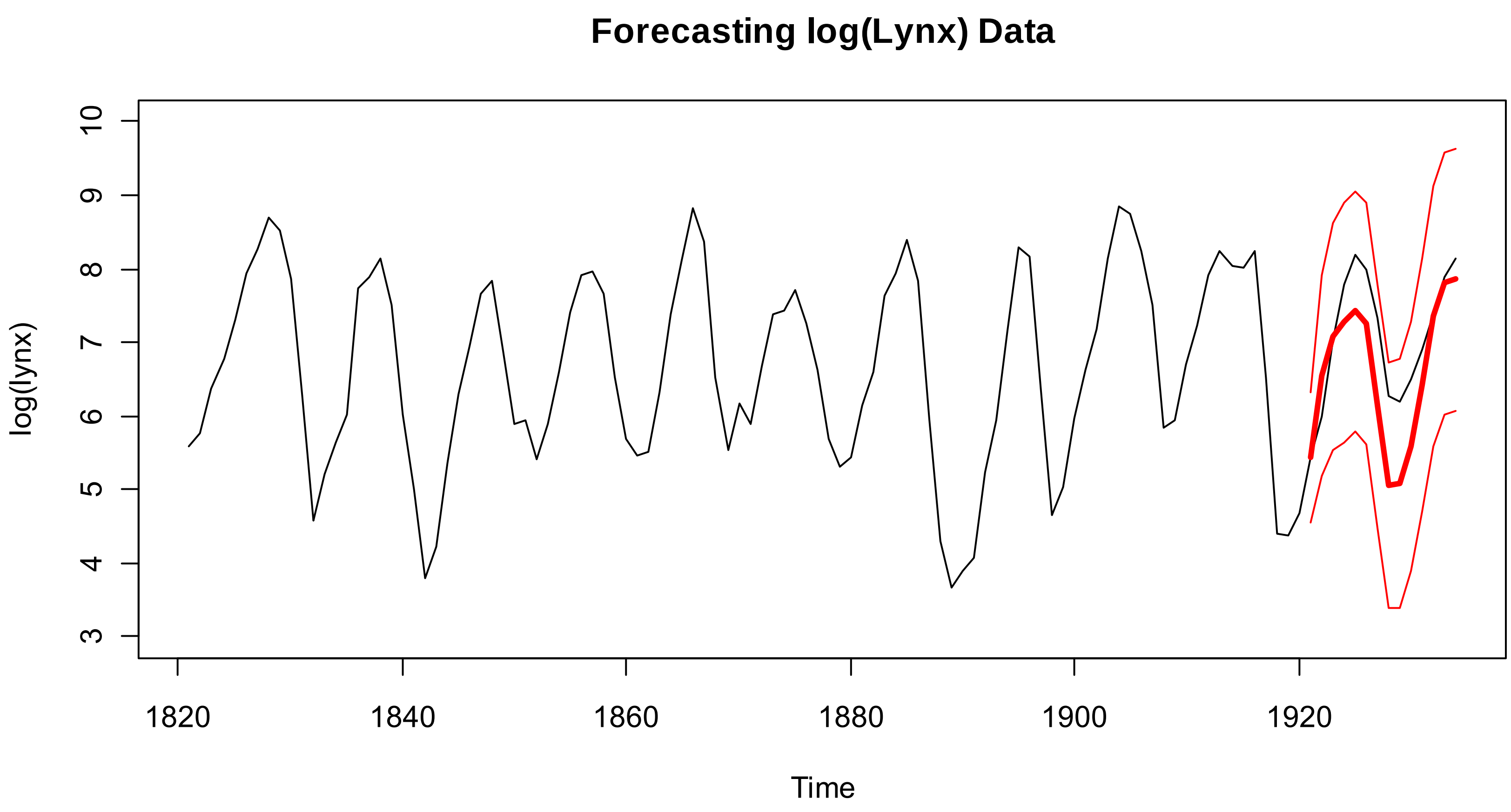
1-step-forecast: $\hat{X}_{n+1,n} = \alpha_1(x_n - m) + \ldots + \alpha_p(x_{n+1-p} - m) + m$

k-step-forecast: $\hat{X}_{n+k,n} = \alpha_1(\hat{X}_{n+k-1,n} - m) + \ldots + \alpha_p(\hat{X}_{n+k-p,n} - m) + m$

If an observed value is available, we plug it in. Else, the forecast is determined in a recursive manner.

Applied Time Series Analysis
SS 2013 – Week 10

Forecasting the Lynx Data



Forecasting: Remarks

- AR(p) processes have a Markov property. Given the model parameters, we only need the last p observations to compute the forecast.

- The prognosis intervals are not simultaneous prognosis intervals, and they are generally too small. However, simulation studies show that this is not excessively so.

- Retaining the final part of the series, and predicting it with several competing models may give hints which one yields the best forecasts. This can be an alternative approach for choosing the model order p.

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 ARMA(p,q)-model we fitted to the data $\{x_1, \ldots, x_n\}$ correctly chosen? What is the "true" order?

3) Are the parameters $\alpha, \beta,$ σ_E^2 and μ 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 10

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

How to Apply the Principles?

- The principles provide a nice setup, but are only useful and practicable under additional assumptions.
- Whereas for AR(p), knowing the last p observations is sufficient for coming up with a forecast, ARMA(p,q) models require knowledge about the infinite past.
- In practice, one is using recursive formulae

→ **see blackboard for the derivation in the MA(1) case!**

Applied Time Series Analysis
SS 2013 – Week 10

MA(1) Forecasting: Summary

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

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

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

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 10

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

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

where …

Applied Time Series Analysis
SS 2013 – Week 10

ARMA(p,q) Forecasting

…where

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

and

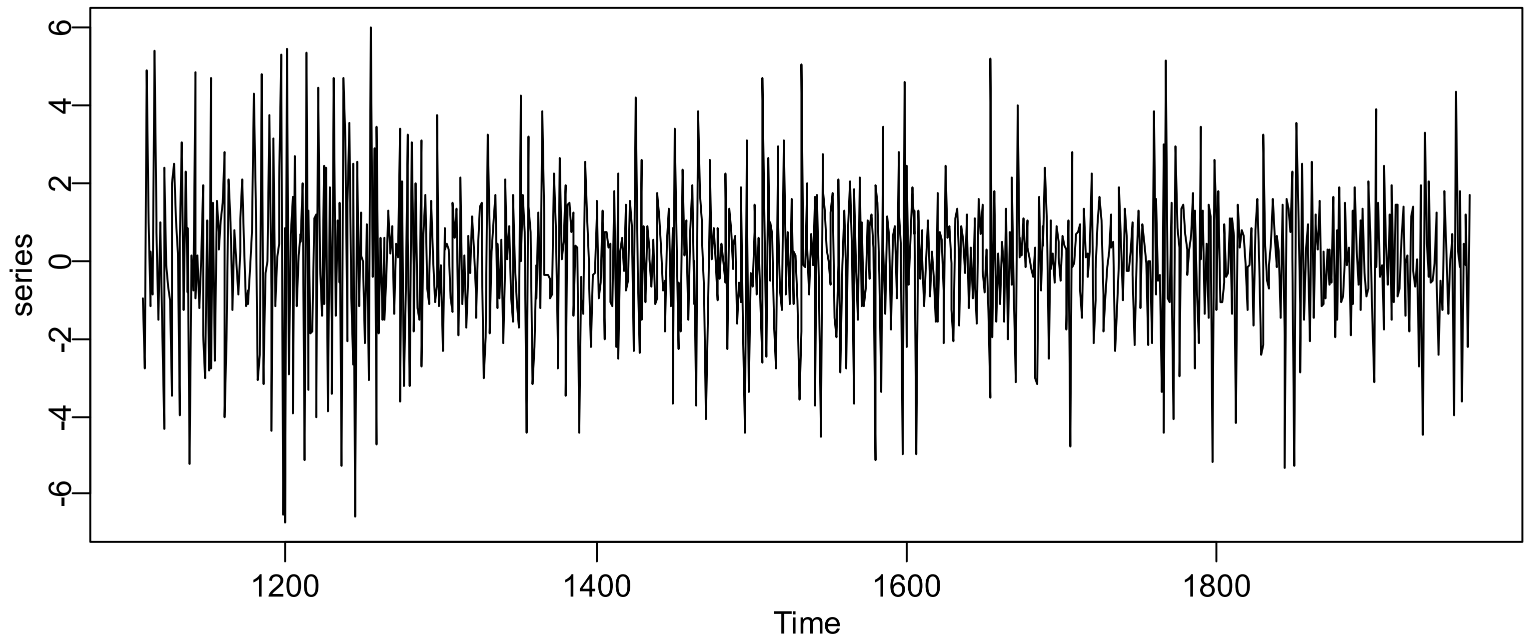
$$E[E_t \mid X_{-\infty}^{n}] = \begin{cases} e_t & \text{if } t \leq n \\ 0 & \text{if } t > n \end{cases}$$

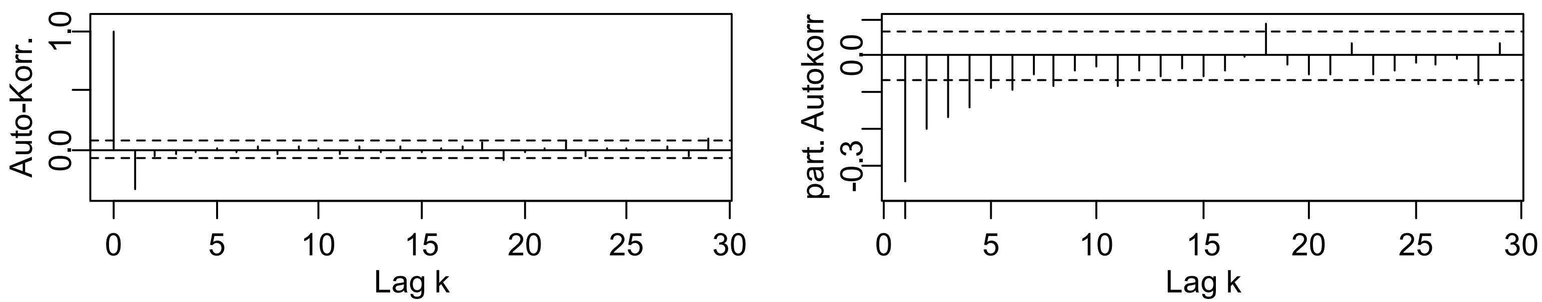
with

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

Applied Time Series Analysis
SS 2013 – Week 10

ARMA(p,q) Forecasting: Douglas Fir





Applied Time Series Analysis
SS 2013 – Week 10

ARMA(p,q) Forecasting: Example