Rob J Hyndman

Forecasting using R

9. Non-seasonal ARIMA models

OTexts.com/fpp/8/

Outline

Autoregressive models

Autoregressive (AR) models:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + e_t,$$

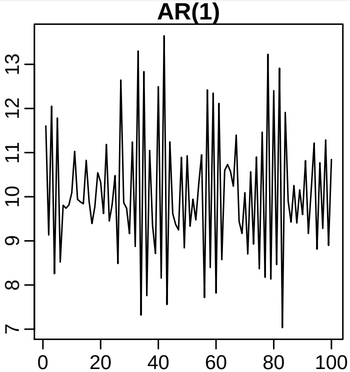
where e_t is white noise. This is a multiple regression with **lagged values** of y_t as predictors.

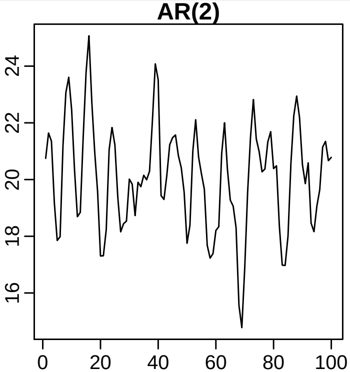
Autoregressive models

Autoregressive (AR) models:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + e_t,$$

where e_t is white noise. This is a multiple regression with **lagged values** of y_t as predictors.





AR(1) model

$$y_t = c + \phi_1 y_{t-1} + e_t$$

- When $\phi_1 = 0$, y_t is **equivalent to WN**
- When $\phi_1 = 1$ and $c = 0$, y_t is **equivalent to a RW**
- When $\phi_1 = 1$ and $c \neq 0$, y_t is **equivalent to a RW with drift**
- When $\phi_1 < 0$, y_t tends to **oscillate between positive and negative values**.

AR(1) model

$$y_t = c + \phi_1 y_{t-1} + e_t$$

- When $\phi_1 = 0$, y_t is **equivalent to WN**
- When $\phi_1 = 1$ and $c = 0$, y_t is **equivalent to a RW**
- When $\phi_1 = 1$ and $c \neq 0$, y_t is **equivalent to a RW with drift**
- When $\phi_1 < 0$, y_t tends to **oscillate between positive and negative values**.

AR(1) model

$$y_t = c + \phi_1 y_{t-1} + e_t$$

- When $\phi_1 = 0$, y_t is **equivalent to WN**
- When $\phi_1 = 1$ and $c = 0$, y_t is **equivalent to a RW**
- When $\phi_1 = 1$ and $c \neq 0$, y_t is **equivalent to a RW with drift**
- When $\phi_1 < 0$, y_t tends to **oscillate between positive and negative values**.

AR(1) model

$$y_t = c + \phi_1 y_{t-1} + e_t$$

- When $\phi_1 = 0$, y_t is **equivalent to WN**
- When $\phi_1 = 1$ and $c = 0$, y_t is **equivalent to a RW**
- When $\phi_1 = 1$ and $c \neq 0$, y_t is **equivalent to a RW with drift**
- When $\phi_1 < 0$, y_t tends to **oscillate between positive and negative values**.

Moving Average (MA) models

Moving Average (MA) models:

$$y_t = c + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_q e_{t-q},$$

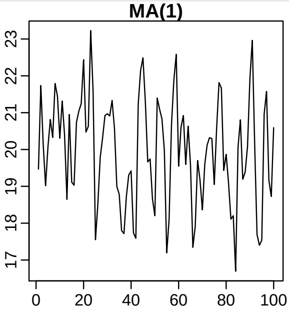
where e_t is white noise. This is a multiple regression with **past *errors*** as predictors. *Don't confuse this with moving average smoothing!*

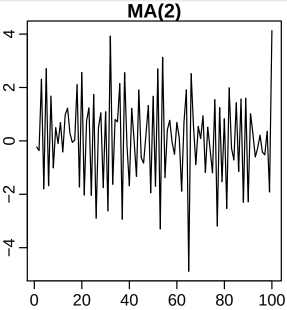
Moving Average (MA) models

Moving Average (MA) models:

$$y_t = c + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_q e_{t-q},$$

where e_t is white noise. This is a multiple regression with ***past errors*** as predictors. *Don't confuse this with moving average smoothing!*





ARIMA models

Autoregressive Moving Average models:

$$y_t = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} + e_t.$$

- Predictors include both **lagged values of y_t and lagged errors.**
- ARMA models can be used for a huge range of stationary time series.
- They model the short-term dynamics.
- An ARMA model applied to *differenced data* is an ARIMA model.

ARIMA models

Autoregressive Moving Average models:

$$y_t = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} + e_t.$$

- Predictors include both **lagged values of y_t and lagged errors.**
- ARMA models can be used for a huge range of stationary time series.
- They model the short-term dynamics.
- An ARMA model applied to *differenced* data is an **ARIMA model**.

ARIMA models

Autoregressive Moving Average models:

$$y_t = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} + e_t.$$

- Predictors include both **lagged values of** y_t **and lagged errors.**
- ARMA models can be used for a huge range of stationary time series.
- They model the short-term dynamics.
- An ARMA model applied to *differenced* data is an **ARIMA model**.

ARIMA models

Autoregressive Moving Average models:

$$y_t = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} + e_t.$$

- Predictors include both **lagged values of y_t and lagged errors.**
- ARMA models can be used for a huge range of stationary time series.
- They model the short-term dynamics.
- An ARMA model applied to *differenced* data is an **ARIMA model**.

ARIMA models

Autoregressive Moving Average models:

$$y_t = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} + e_t.$$

- Predictors include both **lagged values of y_t and lagged errors.**
- ARMA models can be used for a huge range of stationary time series.
- They model the short-term dynamics.
- An ARMA model applied to *differenced* data is an **ARIMA model**.

ARIMA models

Autoregressive Integrated Moving Average models

ARIMA(p, d, q) model

AR: $p =$ order of the autoregressive part
I: $d =$ degree of first differencing involved
MA: $q =$ order of the moving average part.

- White noise model: ARIMA(0,0,0)
- Random walk: ARIMA(0,1,0) with no constant
- Random walk with drift: ARIMA(0,1,0) with const.
- AR(p): ARIMA(p,0,0)
- MA(q): ARIMA(0,0,q)

ARIMA models

Autoregressive Integrated Moving Average models

ARIMA(p, d, q) model

AR: $p =$ order of the autoregressive part
I: $d =$ degree of first differencing involved
MA: $q =$ order of the moving average part.

- White noise model: ARIMA(0,0,0)
- Random walk: ARIMA(0,1,0) with no constant
- Random walk with drift: ARIMA(0,1,0) with const.
- AR(p): ARIMA(p,0,0)
- MA(q): ARIMA(0,0,q)

ARIMA models

Autoregressive Integrated Moving Average models

ARIMA(p, d, q) model

AR:	$p =$ order of the autoregressive part
I:	$d =$ degree of first differencing involved
MA:	$q =$ order of the moving average part.

- White noise model: ARIMA(0,0,0)
- Random walk: ARIMA(0,1,0) with no constant
- Random walk with drift: ARIMA(0,1,0) with const.
- AR(p): ARIMA(p,0,0)
- MA(q): ARIMA(0,0,q)

ARIMA models

Autoregressive Integrated Moving Average models

ARIMA(p, d, q) model

AR:	$p =$ order of the autoregressive part
I:	$d =$ degree of first differencing involved
MA:	$q =$ order of the moving average part.

- White noise model: ARIMA(0,0,0)
- Random walk: ARIMA(0,1,0) with no constant
- Random walk with drift: ARIMA(0,1,0) with const.
- AR(p): ARIMA(p,0,0)
- MA(q): ARIMA(0,0,q)

ARIMA models

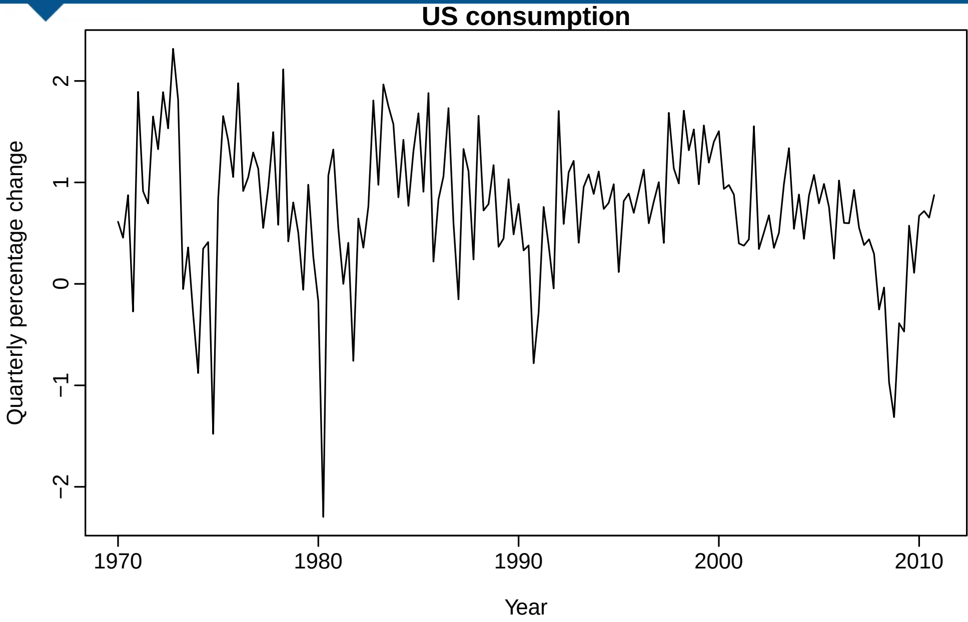
Autoregressive Integrated Moving Average models

ARIMA(p, d, q) model

AR:	$p =$ order of the autoregressive part
I:	$d =$ degree of first differencing involved
MA:	$q =$ order of the moving average part.

- White noise model: ARIMA(0,0,0)
- Random walk: ARIMA(0,1,0) with no constant
- Random walk with drift: ARIMA(0,1,0) with const.
- AR(p): ARIMA(p,0,0)
- MA(q): ARIMA(0,0,q)

US personal consumption



US personal consumption

```
> fit <- auto.arima(usconsumption[,1], seasonal=FALSE)

ARIMA(0,0,3) with non-zero mean
Coefficients:
         ma1     ma2     ma3  intercept
      0.2542  0.2260  0.2695     0.7562
s.e.  0.0767  0.0779  0.0692     0.0844

sigma^2 estimated as 0.3856:  log likelihood=-154.73
AIC=319.46   AICc=319.84   BIC=334.96
```

US personal consumption

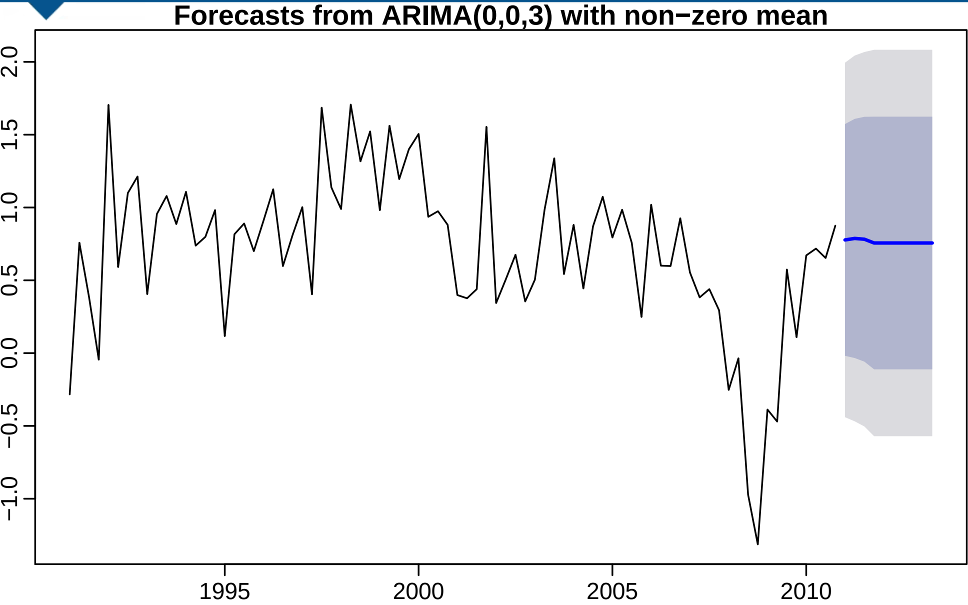
```
> fit <- auto.arima(usconsumption[,1], seasonal=FALSE)

ARIMA(0,0,3) with non-zero mean
Coefficients:
         ma1     ma2     ma3  intercept
      0.2542  0.2260  0.2695     0.7562
s.e.  0.0767  0.0779  0.0692     0.0844

sigma^2 estimated as 0.3856:  log likelihood=-154.73
AIC=319.46   AICc=319.84   BIC=334.96
```

ARIMA(0,0,3) or MA(3) model:

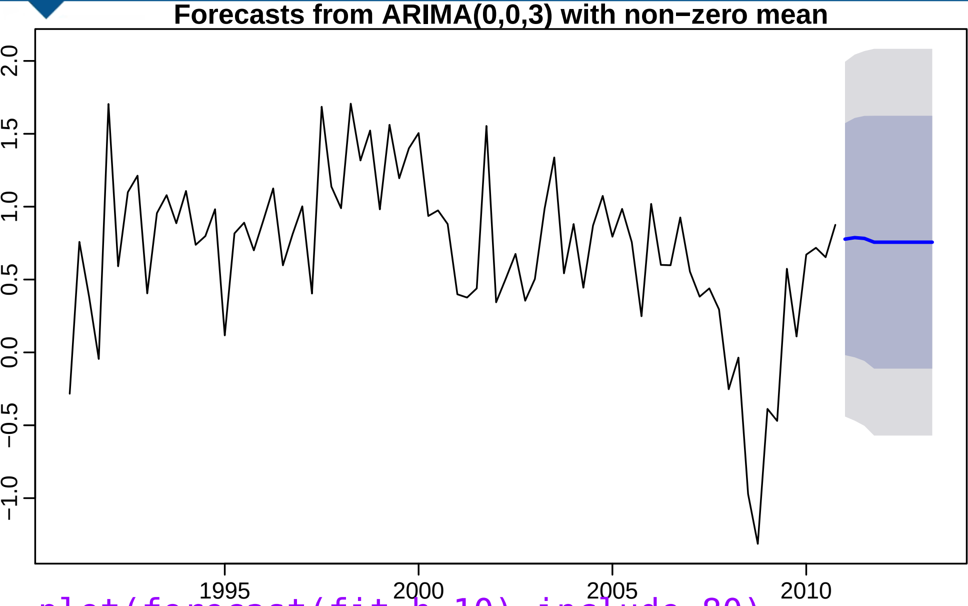
$$y_t = 0.756 + e_t + 0.254e_{t-1} + 0.226e_{t-2} + 0.269e_{t-3},$$

where e_t is white noise with standard deviation $0.62 = \sqrt{0.3856}$.

US personal consumption



US personal consumption




```
plot(forecast(fit,h=10),include=80)
```

Understanding ARIMA models

- If $c = 0$ and $d = 0$, the long-term forecasts will go to zero.
- If $c = 0$ and $d = 1$, the long-term forecasts will go to a non-zero constant.
- If $c = 0$ and $d = 2$, the long-term forecasts will follow a straight line.
- If $c \neq 0$ and $d = 0$, the long-term forecasts will go to the mean of the data.
- If $c \neq 0$ and $d = 1$, the long-term forecasts will follow a straight line.
- If $c \neq 0$ and $d = 2$, the long-term forecasts will follow a quadratic trend.

Understanding ARIMA models

Forecast variance and d

- The higher the value of d, the more rapidly the prediction intervals increase in size.
- For $d = 0$, the long-term forecast standard deviation will go to the standard deviation of the historical data.

Cyclic behaviour

- For cyclic forecasts, $p > 2$ and some restrictions on coefficients are required.

Understanding ARIMA models

Forecast variance and d

- The higher the value of d, the more rapidly the prediction intervals increase in size.
- For $d = 0$, the long-term forecast standard deviation will go to the standard deviation of the historical data.

Cyclic behaviour

- For cyclic forecasts, $p \geq 2$ and some restrictions on coefficients are required.
- If $p = 2$, we need $\phi_1^2 + 4\phi_2 < 0$. Then the average cycle is of length

$$(2\pi)/\left[\text{arc cos}(-\phi_1(1-\phi_2)/(4\phi_2))\right].$$

Understanding ARIMA models

Forecast variance and d

- The higher the value of d, the more rapidly the prediction intervals increase in size.
- For $d = 0$, the long-term forecast standard deviation will go to the standard deviation of the historical data.

Cyclic behaviour

- For cyclic forecasts, $p > 2$ and some restrictions on coefficients are required.
- If $p = 2$, we need $\phi_1^2 + 4\phi_2 < 0$. Then the average cycle is of length

$$(2\pi) / \left[\text{arc}\cos(-\phi_1(1 - \phi_2)/(4\phi_2))\right].$$

Understanding ARIMA models

Forecast variance and d

- The higher the value of d, the more rapidly the prediction intervals increase in size.
- For $d = 0$, the long-term forecast standard deviation will go to the standard deviation of the historical data.

Cyclic behaviour

- For cyclic forecasts, $p > 2$ and some restrictions on coefficients are required.
- If $p = 2$, we need $\phi_1^2 + 4\phi_2 < 0$. Then the average cycle is of length

$$(2\pi)/\left[\text{arc cos}(-\phi_1(1-\phi_2)/(4\phi_2))\right].$$

Understanding ARIMA models

Forecast variance and d

- The higher the value of d, the more rapidly the prediction intervals increase in size.
- For $d = 0$, the long-term forecast standard deviation will go to the standard deviation of the historical data.

Cyclic behaviour

- For cyclic forecasts, $p > 2$ and some restrictions on coefficients are required.
- If $p = 2$, we need $\phi_1^2 + 4\phi_2 < 0$. Then the average cycle is of length

$$(2\pi)/\left[\text{arc}\cos(-\phi_1(1-\phi_2)/(4\phi_2))\right].$$

Outline

1 Non-seasonal ARIMA models

2 Estimation and order selection

3 ARIMA modelling in R

Maximum likelihood estimation

Having identified the model order, we need to estimate the parameters c, $\phi_1, \dots, \phi_p$, $\theta_1, \dots, \theta_q$.

- MLE is very similar to least squares estimation obtained by minimizing

$$\sum_{t=1}^{T} e_t^2.$$

- Non-linear optimization must be used.
- Different software will give different estimates.

Maximum likelihood estimation

Having identified the model order, we need to estimate the parameters c, $\phi_1, \dots, \phi_p$, $\theta_1, \dots, \theta_q$.

- MLE is very similar to least squares estimation obtained by minimizing

$$\sum_{t-1}^{T} e_t^2.$$

- Non-linear optimization must be used.
- Different software will give different estimates.

Maximum likelihood estimation

Having identified the model order, we need to estimate the parameters c, $\phi_1, \dots, \phi_p$, $\theta_1, \dots, \theta_q$.

- MLE is very similar to least squares estimation obtained by minimizing

$$\sum_{t-1}^{T} e_t^2.$$

- Non-linear optimization must be used.
- Different software will give different estimates.

Maximum likelihood estimation

Having identified the model order, we need to estimate the parameters c, $\phi_1, \dots, \phi_p$, $\theta_1, \dots, \theta_q$.

- MLE is very similar to least squares estimation obtained by minimizing

$$\sum_{t-1}^{T} e_t^2.$$

- Non-linear optimization must be used.
- Different software will give different estimates.

Outline

1. Non-seasonal ARIMA models
2. Estimation and order selection
3. **ARIMA modelling in R**

How does auto.arima() work?

A non-seasonal ARIMA process

$$y'_t = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} + e_t.$$

where y'_t has been differenced d times. We need to select the appropriate orders: p, q, d

How does auto.arima() work?

A non-seasonal ARIMA process

$$y'_t = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} + e_t.$$

where y'_t has been differenced d times. We need to select the appropriate orders: p, q, d

Hyndman and Khandakar (JSS, 2008) algorithm:

- Select no. differences d via unit root tests.
- Select p, q by minimising AICc.
- Use stepwise search to traverse model space.

How does auto.arima() work?

Step 1: Select current model (with smallest AIC) from:
$\text{ARIMA}(2, d, 2)$
$\text{ARIMA}(0, d, 0)$
$\text{ARIMA}(1, d, 0)$
$\text{ARIMA}(0, d, 1)$

Step 2: Consider variations of current model:
- vary one of $p, q,$ from current model by ± 1
- p, q both vary from current model by ± 1
- Include/exclude c from current model

Model with lowest AICc becomes current model.

Repeat Step 2 until no lower AICc can be found.

How does auto.arima() work?

Step 1: Select current model (with smallest AIC) from:
$\text{ARIMA}(2,d,2)$
$\text{ARIMA}(0,d,0)$
$\text{ARIMA}(1,d,0)$
$\text{ARIMA}(0,d,1)$

Step 2: Consider variations of current model:
- vary one of $p, q,$ from current model by ± 1
- p, q both vary from current model by ± 1
- Include/exclude c from current model

Model with lowest AICc becomes current model.

Repeat Step 2 until no lower AICc can be found.

How does auto.arima() work?

Step 1: Select current model (with smallest AIC) from:
$\text{ARIMA}(2, d, 2)$
$\text{ARIMA}(0, d, 0)$
$\text{ARIMA}(1, d, 0)$
$\text{ARIMA}(0, d, 1)$

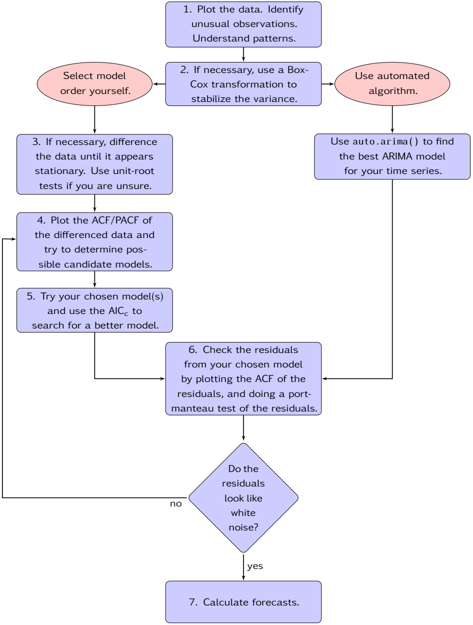
Step 2: Consider variations of current model:

- vary one of $p, q,$ from current model by ± 1
- p, q both vary from current model by ± 1
- Include/exclude c from current model

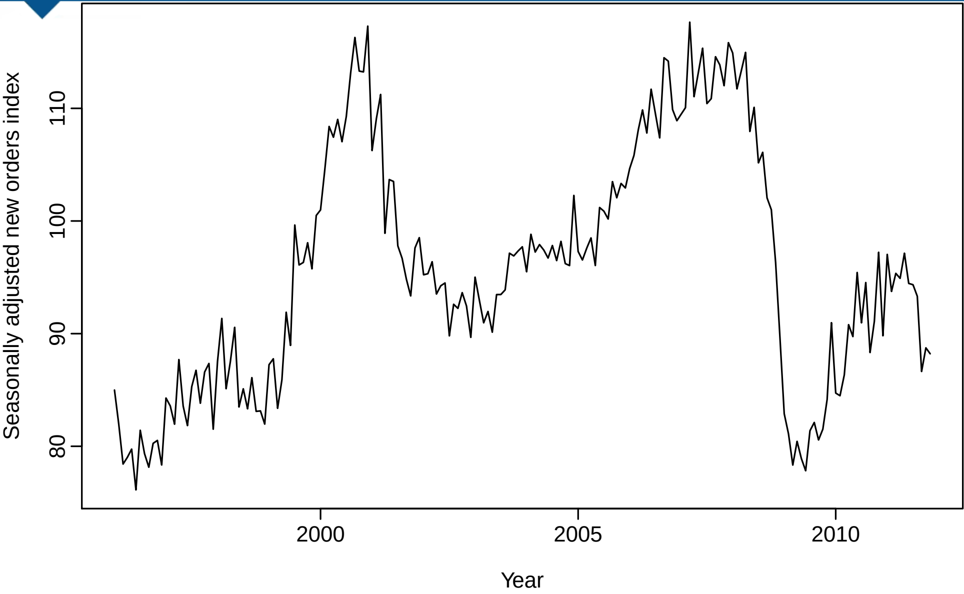
Model with lowest AICc becomes current model.

Repeat Step 2 until no lower AICc can be found.

Modelling procedure



Seasonally adjusted electrical equipment



Seasonally adjusted electrical equipment

1. Time plot shows sudden changes, particularly big drop in 2008/2009 due to global economic environment. Otherwise nothing unusual and no need for data adjustments.
2. No evidence of changing variance, so no Box-Cox transformation.
3. `auto.arima` suggests an ARIMA(3,1,1) model.

Seasonally adjusted electrical equipment

1. Time plot shows sudden changes, particularly big drop in 2008/2009 due to global economic environment. Otherwise nothing unusual and no need for data adjustments.
2. No evidence of changing variance, so no Box-Cox transformation.
3. `auto.arima` suggests an ARIMA(3,1,1) model.

Seasonally adjusted electrical equipment

1. Time plot shows sudden changes, particularly big drop in 2008/2009 due to global economic environment. Otherwise nothing unusual and no need for data adjustments.
2. No evidence of changing variance, so no Box-Cox transformation.
3. `auto.arima` suggests an ARIMA(3,1,1) model.

Seasonally adjusted electrical equipment

```
> fit <- auto.arima(eeadj)
> summary(fit)
Series: eeadj
ARIMA(3,1,1)

Coefficients:
         ar1     ar2     ar3      ma1
      0.0519  0.1191  0.3730  -0.4542
s.e.  0.1840  0.0888  0.0679   0.1993

sigma^2 estimated as 9.532:  log likelihood=-484.08
AIC=978.17   AICc=978.49   BIC=994.4
```

Seasonally adjusted electrical equipment

6. ACF plot of residuals from ARIMA(3,1,1) model look like white noise.

```
Acf(residuals(fit))
Box.test(residuals(fit), lag=24,
  fitdf=4, type="Ljung")
```

Seasonally adjusted electrical equipment

6. ACF plot of residuals from ARIMA(3,1,1) model look like white noise.

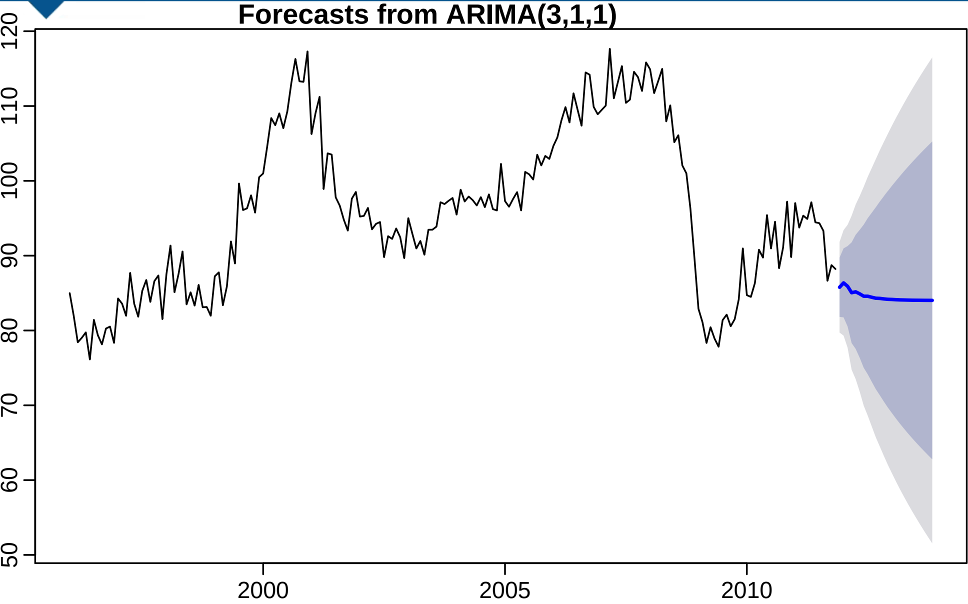
```
Acf(residuals(fit))
Box.test(residuals(fit), lag=24,
  fitdf=4, type="Ljung")
```

Seasonally adjusted electrical equipment

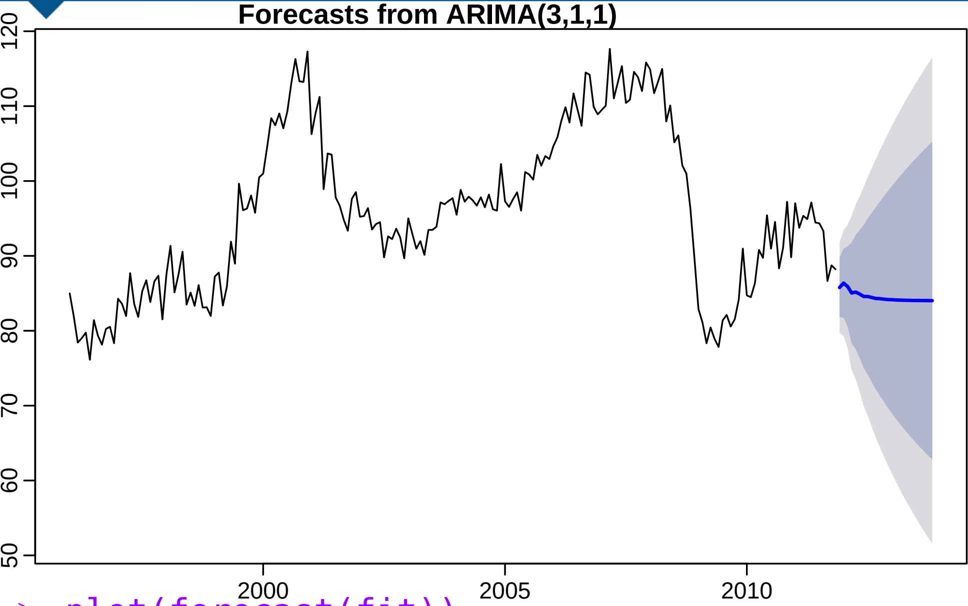
6. ACF plot of residuals from ARIMA(3,1,1) model look like white noise.

```
Acf(residuals(fit))
Box.test(residuals(fit), lag=24,
  fitdf=4, type="Ljung")
```

Seasonally adjusted electrical equipment



Seasonally adjusted electrical equipment




```
> plot(forecast(fit))
```

Prediction intervals

- Prediction intervals **increase in size with forecast horizon.**
- Prediction intervals can be difficult to calculate by hand
- Calculations assume residuals are **uncorrelated** and **normally distributed**.
- Prediction intervals tend to be too narrow.

Prediction intervals

- Prediction intervals **increase in size with forecast horizon.**
- Prediction intervals can be difficult to calculate by hand
- Calculations assume residuals are **uncorrelated** and **normally distributed**.
- Prediction intervals tend to be too narrow.
  - the uncertainty in the parameter estimates has not been accounted for.
  - the ARIMA model assumes historical patterns will not change during the forecast period.
  - the ARIMA model assumes uncorrelated future errors

Prediction intervals

- Prediction intervals **increase in size with forecast horizon.**
- Prediction intervals can be difficult to calculate by hand
- Calculations assume residuals are **uncorrelated** and **normally distributed**.
- Prediction intervals tend to be too narrow.
  - the uncertainty in the parameter estimates has not been accounted for.
  - the ARIMA model assumes historical patterns will not change during the forecast period.
  - the ARIMA model assumes uncorrelated future errors

Prediction intervals

- Prediction intervals **increase in size with forecast horizon.**
- Prediction intervals can be difficult to calculate by hand
- Calculations assume residuals are **uncorrelated** and **normally distributed**.
- Prediction intervals tend to be too narrow.
  - the uncertainty in the parameter estimates has not been accounted for.
  - the ARIMA model assumes historical patterns will not change during the forecast period.
  - the ARIMA model assumes uncorrelated future errors

Prediction intervals

- Prediction intervals **increase in size with forecast horizon.**
- Prediction intervals can be difficult to calculate by hand
- Calculations assume residuals are **uncorrelated** and **normally distributed**.
- Prediction intervals tend to be too narrow.
  - the uncertainty in the parameter estimates has not been accounted for.
  - the ARIMA model assumes historical patterns will not change during the forecast period.
  - the ARIMA model assumes uncorrelated future errors

Prediction intervals

- Prediction intervals **increase in size with forecast horizon.**
- Prediction intervals can be difficult to calculate by hand
- Calculations assume residuals are **uncorrelated** and **normally distributed**.
- Prediction intervals tend to be too narrow.
  - the uncertainty in the parameter estimates has not been accounted for.
  - the ARIMA model assumes historical patterns will not change during the forecast period.
  - the ARIMA model assumes uncorrelated future errors

Prediction intervals

- Prediction intervals **increase in size with forecast horizon.**
- Prediction intervals can be difficult to calculate by hand
- Calculations assume residuals are **uncorrelated** and **normally distributed**.
- Prediction intervals tend to be too narrow.
    - the uncertainty in the parameter estimates has not been accounted for.
    - the ARIMA model assumes historical patterns will not change during the forecast period.
    - the ARIMA model assumes uncorrelated future errors