



Rob J Hyndman

Forecasting using



4. White noise and time series decomposition

OTexts.com/fpp/2/6 OTexts.com/fpp/6/

Forecasting using R

Student award

- To students who make the most contribution to the class, as voted by their peers.
- Contributions can be on Piazza or during the webex online sessions.
- Contribute questions, answers, comments, code suggestions, etc.
- Looking for engaged learners, not experts.
- At end of course, all students to vote for best contributor.
- \$100 to the top voted student and \$50 to the second student (plus t-shirts).

Forecasting using R

Time zones

All sessions are at UTC 22:00 for the entire course.

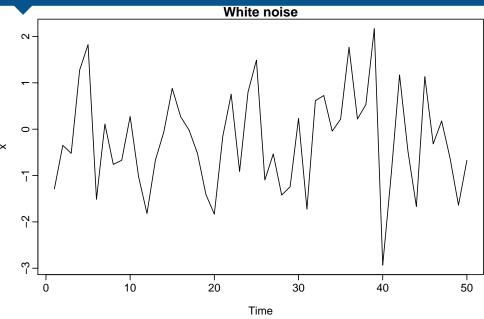
Be aware when your time zone changes.

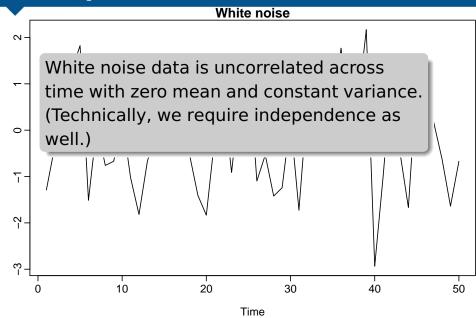
- Most of Europe changed last Sunday. So classes are now one hour earlier for most European residents.
- Most of USA and Canada change next Sunday. So classes will be one hour earlier for most North American residents from next week.

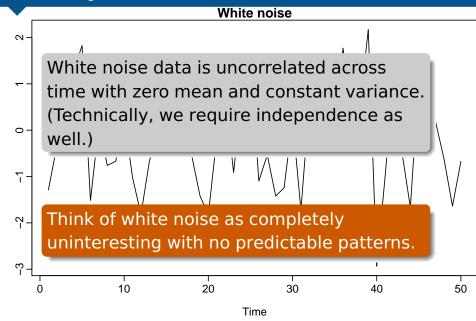
Forecasting using R

Outline

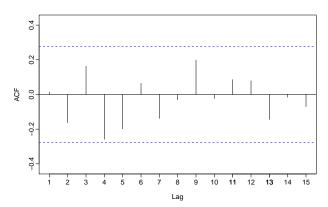
- 1 White noise
- 2 Time series decomposition
- 3 Seasonal adjustment
- 4 Forecasting and decomposition







$$r_1 = 0.013$$
 $r_2 = -0.163$
 $r_3 = 0.163$
 $r_4 = -0.259$
 $r_5 = -0.198$
 $r_6 = 0.064$
 $r_7 = -0.139$
 $r_8 = -0.032$
 $r_9 = 0.199$
 $r_{10} = -0.240$



Sample autocorrelations for white noise series. For uncorrelated data, we would expect each autocorrelation to be close to zero.

- 95% of all r_k for white noise must lie within $\pm 1.96 / \sqrt{T}$
- If this is not the case, the series is probably now WN.
 - m Common to plot lines at $\pm 1.96/\sqrt{T}$ when plotting ACF. These are the *critical value*

- 95% of all r_k for white noise must lie within $\pm 1.96/\sqrt{T}$.
- If this is not the case, the series is probably not WN.
- Common to plot lines at $\pm 1.96/\sqrt{T}$ when plotting ACF. These are the *critical values*.

- 95% of all r_k for white noise must lie within $\pm 1.96/\sqrt{T}$.
- If this is not the case, the series is probably not WN.
- Common to plot lines at $\pm 1.96/\sqrt{T}$ when plotting ACF. These are the *critical values*.

- 95% of all r_k for white noise must lie within $\pm 1.96/\sqrt{T}$.
- If this is not the case, the series is probably not WN.
- Common to plot lines at $\pm 1.96/\sqrt{T}$ when plotting ACF. These are the *critical values*.

Autocorrelation

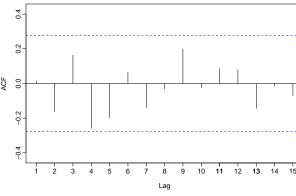
Example:

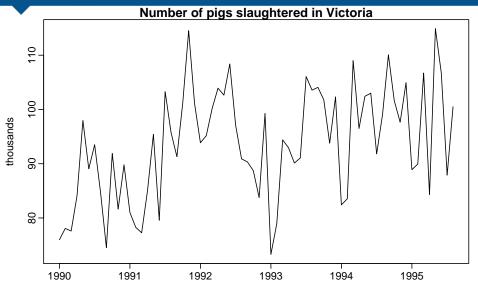
T=50 and so critical values at $\pm 1.96/\sqrt{50}=\pm 0.28$.

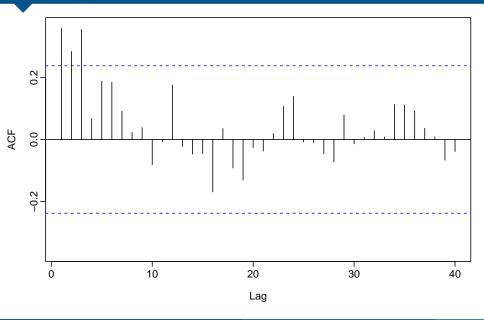
All autocorrelation coefficients lie within these limits, confirming that the data are white noise.

(More precisely, the data cannot be

distinguished from white noise.)







Monthly total number of pigs slaughtered in the state of Victoria, Australia, from January 1990 through August 1995. (Source: Australian Bureau of Statistics.)

- Difficult to detect pattern in time plot.
- ACF shows some significant autocorrelation at lags 1, 2, and 3.
- r_{12} relatively large although not significant.
- This may indicate some slight seasonality.

Monthly total number of pigs slaughtered in the state of Victoria, Australia, from January 1990 through August 1995. (Source: Australian Bureau of Statistics.)

- Difficult to detect pattern in time plot.
- ACF shows some significant autocorrelation at lags 1, 2, and 3.
- $Arr r_{12}$ relatively large although not significant. This may indicate some slight seasonality.

Monthly total number of pigs slaughtered in the state of Victoria, Australia, from January 1990 through August 1995. (Source: Australian Bureau of Statistics.)

- Difficult to detect pattern in time plot.
- ACF shows some significant autocorrelation at lags 1, 2, and 3.
- r_{12} relatively large although not significant. This may indicate some slight seasonality.

Monthly total number of pigs slaughtered in the state of Victoria, Australia, from January 1990 through August 1995. (Source: Australian Bureau of Statistics.)

- Difficult to detect pattern in time plot.
- ACF shows some significant autocorrelation at lags 1, 2, and 3.
- $Arr r_{12}$ relatively large although not significant. This may indicate some slight seasonality.

Monthly total number of pigs slaughtered in the state of Victoria, Australia, from January 1990 through August 1995. (Source: Australian Bureau of Statistics.)

- Difficult to detect pattern in time plot.
- ACF shows some significant autocorrelation at lags 1, 2, and 3.
- $Arr r_{12}$ relatively large although not significant. This may indicate some slight seasonality.

Monthly total number of pigs slaughtered in the state of Victoria, Australia, from January 1990 through August 1995. (Source: Australian Bureau of Statistics.)

- Difficult to detect pattern in time plot.
- ACF shows some significant autocorrelation at lags 1, 2, and 3.
- $Arr r_{12}$ relatively large although not significant. This may indicate some slight seasonality.

- We assume that the residuals are white noise (uncorrelated, mean zero, constant variance). If they aren't, then there is information left in the residuals that should be used in computing forecasts.
- So a standard residual diagnostic is to check the ACF of the residuals of a forecasting method.
- We expect these to look like white noise.

 $\hat{y}_{t|t-1} = y_{t-1}$

- We assume that the residuals are white noise (uncorrelated, mean zero, constant variance). If they aren't, then there is information left in the residuals that should be used in computing forecasts.
- So a standard residual diagnostic is to check the ACF of the residuals of a forecasting method.
- We *expect* these to look like white noise.

Dow-Jones naive forecasts revisited

 $\hat{y}_{t|t-1} = y_{t-1} \\ e_t = y_t - y_{t-1}$

- We assume that the residuals are white noise (uncorrelated, mean zero, constant variance). If they aren't, then there is information left in the residuals that should be used in computing forecasts.
- So a standard residual diagnostic is to check the ACF of the residuals of a forecasting method.
- We *expect* these to look like white noise.

Dow-Jones naive forecasts revisited

$$\hat{y}_{t|t-1} = y_{t-1} \\ e_t = y_t - y_{t-1}$$

- We assume that the residuals are white noise (uncorrelated, mean zero, constant variance). If they aren't, then there is information left in the residuals that should be used in computing forecasts.
- So a standard residual diagnostic is to check the ACF of the residuals of a forecasting method.
- We *expect* these to look like white noise.

Dow-Jones naive forecasts revisited

$$\hat{y}_{t|t-1} = y_{t-1}$$

 $e_t = v_t - v_{t-1}$

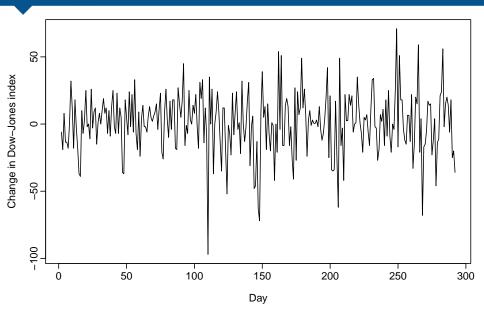
- We assume that the residuals are white noise (uncorrelated, mean zero, constant variance). If they aren't, then there is information left in the residuals that should be used in computing forecasts.
- So a standard residual diagnostic is to check the ACF of the residuals of a forecasting method.
- We *expect* these to look like white noise.

Dow-Jones naive forecasts revisited

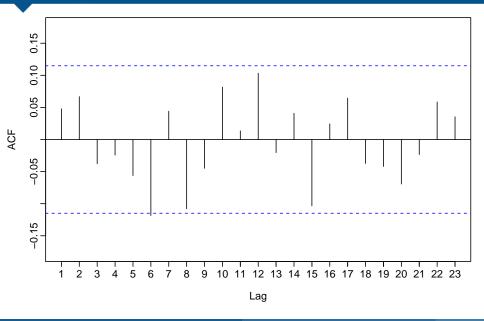
$$\hat{y}_{t|t-1} = y_{t-1}$$

 $e_t = y_t - y_{t-1}$

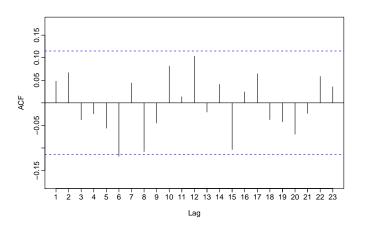
Forecasting Dow-Jones index



Forecasting Dow-Jones index

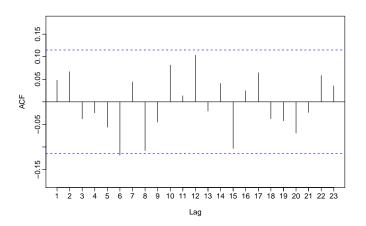


Example: Dow-Jones residuals



- These look like white noise.
- But the ACF is a multiple testing problem.

Example: Dow-Jones residuals



- These look like white noise.
- But the ACF is a multiple testing problem.

Consider a whole set of r_k values, and develop a test to see whether the set is significantly different from a zero set.

Consider a whole set of r_k values, and develop a test to see whether the set is significantly different from a zero set.

Box-Pierce test

$$Q = T \sum_{k=1}^{h} r_k^2$$

where h is max lag being considered and T is number of observations.

Consider a whole set of r_k values, and develop a test to see whether the set is significantly different from a zero set.

Box-Pierce test

$$Q = T \sum_{k=1}^{h} r_k^2$$

where h is max lag being considered and T is number of observations.

- My preferences: h = 10 for non-seasonal data, h = 2m for seasonal data.
- If each r_k close to zero, Q will be **small**.
- If some r_k values large (positive or negative), Q will be large.

Consider a whole set of r_k values, and develop a test to see whether the set is significantly different from a zero set.

Box-Pierce test

$$Q = T \sum_{k=1}^{h} r_k^2$$

where h is max lag being considered and T is number of observations.

- My preferences: h = 10 for non-seasonal data, h = 2m for seasonal data.
- If each r_k close to zero, Q will be **small**.
- If some r_k values large (positive or negative), Q will be large.

Consider a whole set of r_k values, and develop a test to see whether the set is significantly different from a zero set.

Box-Pierce test

$$Q = T \sum_{k=1}^{h} r_k^2$$

where h is max lag being considered and T is number of observations.

- My preferences: h = 10 for non-seasonal data, h = 2m for seasonal data.
- If each r_k close to zero, Q will be **small**.
- If some r_k values large (positive or negative), Q will be large.

Consider a whole set of r_k values, and develop a test to see whether the set is significantly different from a zero set.

Ljung-Box test

$$Q^* = T(T+2) \sum_{k=1}^{h} (T-k)^{-1} r_k^2$$

where h is max lag being considered and T is number of observations.

- My preferences: h = 10 for non-seasonal data, h = 2m for seasonal data.
- Better performance, especially in small samples.

Portmanteau tests

- If data are WN, Q^* has χ^2 distribution with (h K) degrees of freedom where K = no. parameters in model.
- When applied to raw data, set K = 0.
- For the Dow-Jones example,

```
# lag=h and fitdf=K
> Box.test(res, lag=10, fitdf=0)
Box-Pierce test
X-squared = 14.0451, df = 10, p-value = 0.1709
> Box.test(res, lag=10, fitdf=0, type="Lj")
Box-Ljung test
X-squared = 14.4615, df = 10, p-value = 0.153
```

Portmanteau tests

- If data are WN, Q^* has χ^2 distribution with (h K) degrees of freedom where K = no. parameters in model.
- When applied to raw data, set K = 0.
- For the Dow-Jones example,

```
# lag=h and fitdf=K
> Box.test(res, lag=10, fitdf=0)
Box-Pierce test
X-squared = 14.0451, df = 10, p-value = 0.1709
> Box.test(res, lag=10, fitdf=0, type="Lj")
Box-Ljung test
X-squared = 14.4615, df = 10, p-value = 0.153
```

Portmanteau tests

- If data are WN, Q^* has χ^2 distribution with (h K) degrees of freedom where K = no. parameters in model.
- When applied to raw data, set K = 0.
- For the Dow-Jones example,

```
res <- residuals(naive(dj))

# lag=h and fitdf=K
> Box.test(res, lag=10, fitdf=0)
Box-Pierce test
X-squared = 14.0451, df = 10, p-value = 0.1709
> Box.test(res, lag=10, fitdf=0, type="Lj")
Box-Ljung test
X-squared = 14.4615, df = 10, p-value = 0.153
```

Exercise

- Calculate the residuals from a seasonal naive forecast applied to the quarterly Australian beer production data from 1992.
- Test if the residuals are white noise.

Exercise

- Calculate the residuals from a seasonal naive forecast applied to the quarterly Australian beer production data from 1992.
- Test if the residuals are white noise.

```
beer <- window(ausbeer,start=1992)
fc <- snaive(beer)
res <- residuals(fc)
Acf(res)
Box.test(res, lag=8, fitdf=0, type="Lj")</pre>
```

Outline

- 1 White noise
- 2 Time series decomposition
- 3 Seasonal adjustment
- 4 Forecasting and decomposition

$$Y_t = f(S_t, T_t, E_t)$$

where $Y_t = \text{data at period } t$

 $S_t =$ seasonal component at period t

 $T_t = \text{trend-cycle component at period } t$

 $E_t = \text{remainder (or irregular or error)}$

component at period t

Additive decomposition: $Y_t = S_t + T_t + E_t$. Multiplicative decomposition: $Y_t = S_t \times T_t \times E_t$.

$$Y_t = f(S_t, T_t, E_t)$$

where $Y_t = \text{data at period } t$

 S_t = seasonal component at period t

 $T_t = \text{trend-cycle component at period } t$

 E_t = remainder (or irregular or error)

component at period t

Additive decomposition: $Y_t = S_t + T_t + E_t$.

Multiplicative decomposition: $Y_t = S_t \times T_t \times E_t$.

$$Y_t = f(S_t, T_t, E_t)$$

where $Y_t = \text{data at period } t$

 $S_t =$ seasonal component at period t

 $T_t = \text{trend-cycle component at period } t$

 E_t = remainder (or irregular or error)

component at period t

Additive decomposition: $Y_t = S_t + T_t + E_t$. Multiplicative decomposition: $Y_t = S_t \times T_t \times E_t$.

- Additive model appropriate if magnitude of seasonal fluctuations does not vary with level.
- If seasonal are proportional to level of series, then multiplicative model appropriate.
- Multiplicative decomposition more prevalent with economic series
- Logs turn multiplicative relationship into an additive relationship:

$$Y_t = S_t \times T_t \times E_t \quad \Rightarrow \quad \log Y_t = \log S_t + \log T_t + \log E_t$$

- Additive model appropriate if magnitude of seasonal fluctuations does not vary with level.
- If seasonal are proportional to level of series, then multiplicative model appropriate.
- Multiplicative decomposition more prevalent with economic series
- Logs turn multiplicative relationship into an additive relationship:

$$Y_t = S_t \times T_t \times E_t \quad \Rightarrow \quad \log Y_t = \log S_t + \log T_t + \log E_t.$$

- Additive model appropriate if magnitude of seasonal fluctuations does not vary with level.
- If seasonal are proportional to level of series, then multiplicative model appropriate.
- Multiplicative decomposition more prevalent with economic series
- Logs turn multiplicative relationship into an additive relationship:

$$Y_t = S_t \times T_t \times E_t \quad \Rightarrow \quad \log Y_t = \log S_t + \log T_t + \log E_t.$$

- Additive model appropriate if magnitude of seasonal fluctuations does not vary with level.
- If seasonal are proportional to level of series, then multiplicative model appropriate.
- Multiplicative decomposition more prevalent with economic series
- Logs turn multiplicative relationship into an additive relationship:

$$Y_t = S_t \times T_t \times E_t \quad \Rightarrow \quad \log Y_t = \log S_t + \log T_t + \log E_t.$$

- Classical method originated in 1920s. In R: decompose().
- Census II method introduced in 1957.
 Basis for modern X-12-ARIMA method
- STL method introduced in 1983. It only allows additive decomposition.
- TRAMO/SEATS introduced in 1990s.

- Classical method originated in 1920s. In R: decompose().
- Census II method introduced in 1957.
 Basis for modern X-12-ARIMA method.
- STL method introduced in 1983. It only allows additive decomposition.

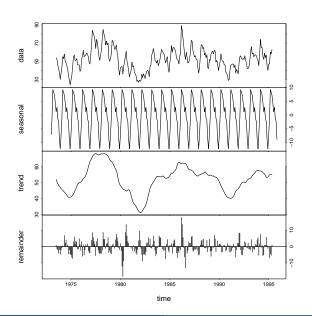
- Classical method originated in 1920s. In R: decompose().
- Census II method introduced in 1957. Basis for modern X-12-ARIMA method.
- STL method introduced in 1983. It only allows additive decomposition.
 - In R: stl().
- TRAMO/SEATS introduced in 1990s

- Classical method originated in 1920s. In R: decompose().
- Census II method introduced in 1957.Basis for modern X-12-ARIMA method.
- STL method introduced in 1983. It only allows additive decomposition.
 In R: stl().
- TRAMO/SEATS introduced in 1990s

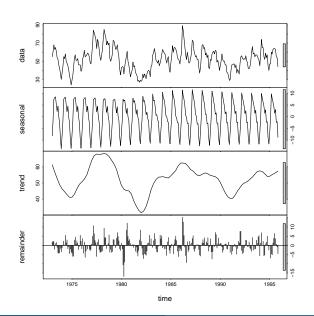
- Classical method originated in 1920s. In R: decompose().
- Census II method introduced in 1957.Basis for modern X-12-ARIMA method.
- STL method introduced in 1983. It only allows additive decomposition.
 - In R: stl().
- TRAMO/SEATS introduced in 1990s

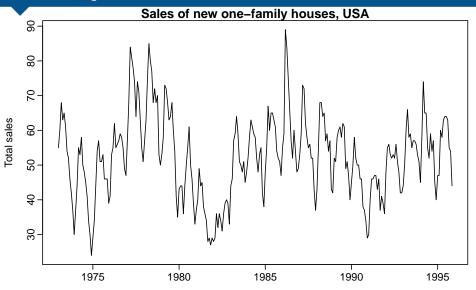
- Classical method originated in 1920s. In R: decompose().
- Census II method introduced in 1957.Basis for modern X-12-ARIMA method.
- STL method introduced in 1983. It only allows additive decomposition. In R: stl().
- TRAMO/SEATS introduced in 1990s.

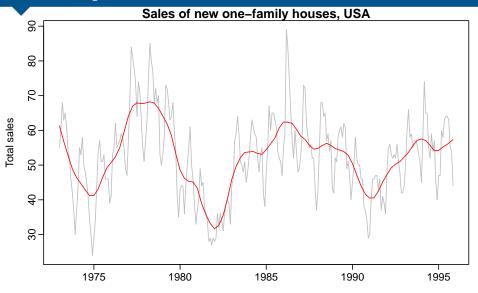
Classical decomposition

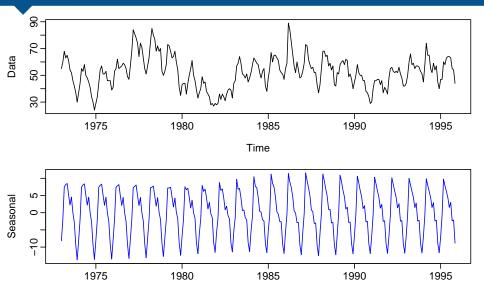


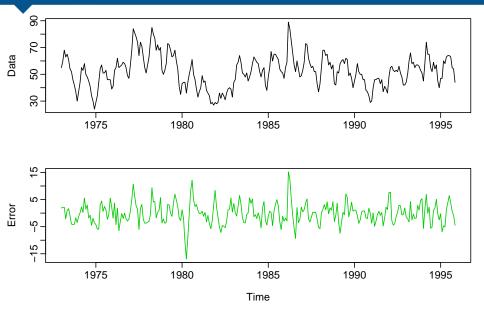
STL decomposition

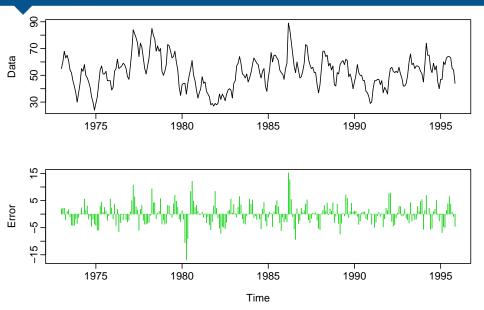


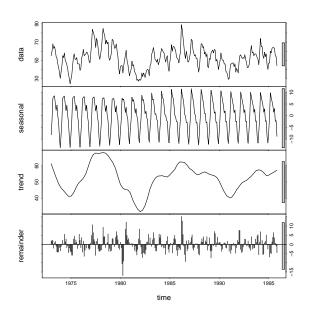




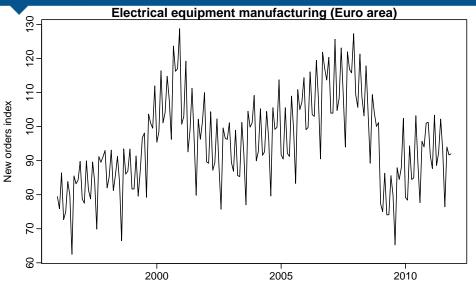


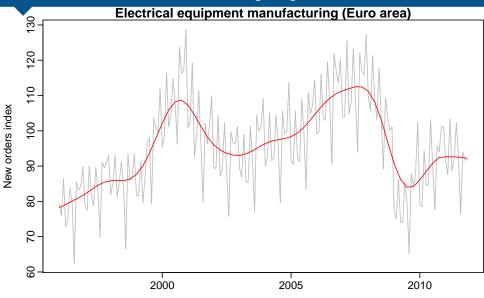


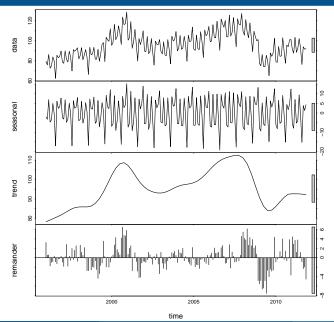


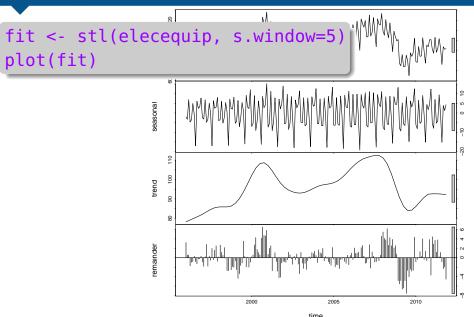


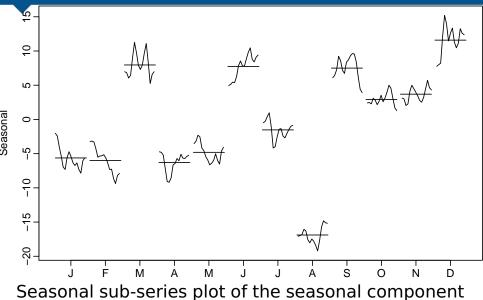
```
plot(decompose(hsales))
plot(stl(hsales,s.window="periodic"))
plot(stl(hsales,s.window=15))
```

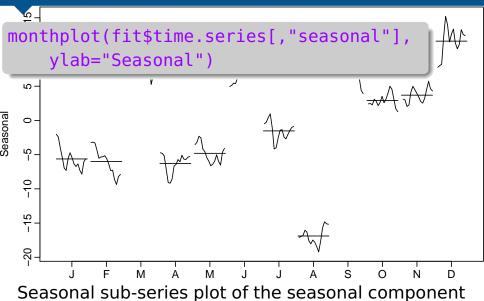


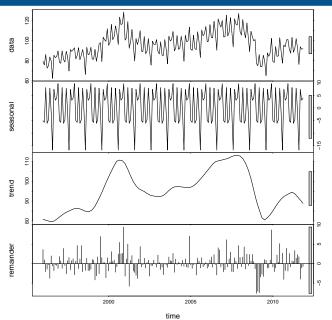


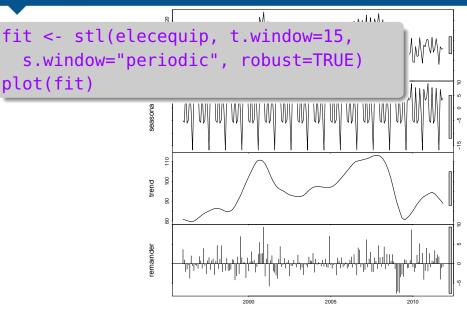




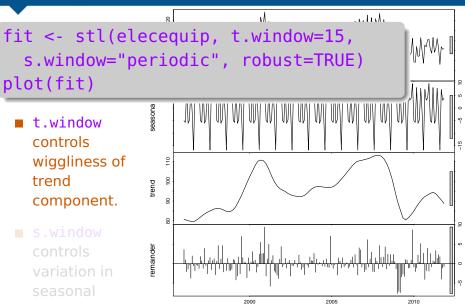




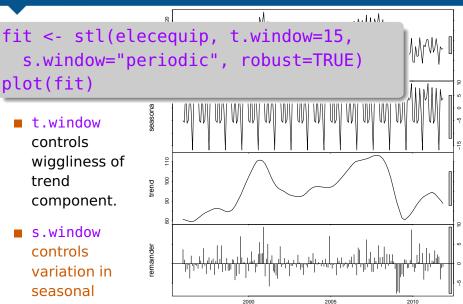




Euro electrical equipment



Euro electrical equipment



component.

Outline

- 1 White noise
- 2 Time series decomposition
- 3 Seasonal adjustment
- 4 Forecasting and decomposition

Seasonal adjustment

- Useful by-product of decomposition: an easy way to calculate seasonally adjusted data.
- Additive decomposition: seasonally adjusted data given by

$$Y_t - S_t = T_t + E_t$$

 Multiplicative decomposition: seasonally adjusted data given by

$$Y_t/S_t = T_t \times E_t$$

Seasonal adjustment

- Useful by-product of decomposition: an easy way to calculate seasonally adjusted data.
- Additive decomposition: seasonally adjusted data given by

$$Y_t - S_t = T_t + E_t$$

 Multiplicative decomposition: seasonally adjusted data given by

$$Y_t/S_t = T_t \times E_t$$

Seasonal adjustment

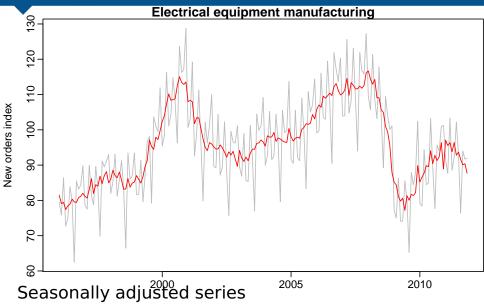
- Useful by-product of decomposition: an easy way to calculate seasonally adjusted data.
- Additive decomposition: seasonally adjusted data given by

$$Y_t - S_t = T_t + E_t$$

 Multiplicative decomposition: seasonally adjusted data given by

$$Y_t/S_t = T_t \times E_t$$

Euro electrical equipment



Seasonal adjustment in R

```
seasadj(obj)
where obj is the output from stl() or
decompose().
```

Example

```
plot(hsales,col="gray")
fit <- stl(hsales,s.window=15)
hsales.sa <- seasadj(fit)
lines(hsales.sa, col="red")</pre>
```

Seasonal adjustment in R

```
seasadj(obj)
where obj is the output from stl() or
decompose().
```

Example

```
plot(hsales,col="gray")
fit <- stl(hsales,s.window=15)
hsales.sa <- seasadj(fit)
lines(hsales.sa, col="red")</pre>
```

Outline

- 1 White noise
- 2 Time series decomposition
- 3 Seasonal adjustment
- 4 Forecasting and decomposition

- Forecast seasonal component by repeating the last year
- Forecast seasonally adjusted data using non-seasonal time series method. E.g.,
 - Holt's method next topic
 - Random walk with drift model
- Combine forecasts of seasonal component with forecasts of seasonally adjusted data to get forecasts of original data.
- Sometimes a decomposition is useful just for understanding the data before building a separate forecasting model.

- Forecast seasonal component by repeating the last year
- Forecast seasonally adjusted data using non-seasonal time series method. E.g.,
 - Holt's method next topic
 - Random walk with drift model
- Combine forecasts of seasonal component with forecasts of seasonally adjusted data to get forecasts of original data.
- Sometimes a decomposition is useful just for understanding the data before building a separate forecasting model.

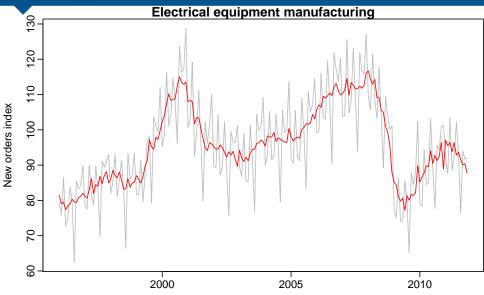
- Forecast seasonal component by repeating the last year
- Forecast seasonally adjusted data using non-seasonal time series method. E.g.,
 - Holt's method next topic
 - Random walk with drift model
- Combine forecasts of seasonal component with forecasts of seasonally adjusted data to get forecasts of original data.
- Sometimes a decomposition is useful just for understanding the data before building a separate forecasting model.

- Forecast seasonal component by repeating the last year
- Forecast seasonally adjusted data using non-seasonal time series method. E.g.,
 - Holt's method next topic
 - Random walk with drift model
- Combine forecasts of seasonal component with forecasts of seasonally adjusted data to get forecasts of original data.
- Sometimes a decomposition is useful just for understanding the data before building a separate forecasting model.

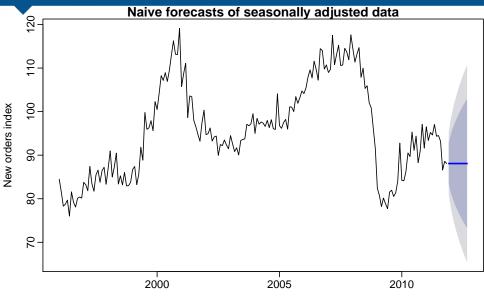
- Forecast seasonal component by repeating the last year
- Forecast seasonally adjusted data using non-seasonal time series method. E.g.,
 - Holt's method next topic
 - Random walk with drift model
- Combine forecasts of seasonal component with forecasts of seasonally adjusted data to get forecasts of original data.
- Sometimes a decomposition is useful just for understanding the data before building a separate forecasting model.

- Forecast seasonal component by repeating the last year
- Forecast seasonally adjusted data using non-seasonal time series method. E.g.,
 - Holt's method next topic
 - Random walk with drift model
- Combine forecasts of seasonal component with forecasts of seasonally adjusted data to get forecasts of original data.
- Sometimes a decomposition is useful just for understanding the data before building a separate forecasting model.

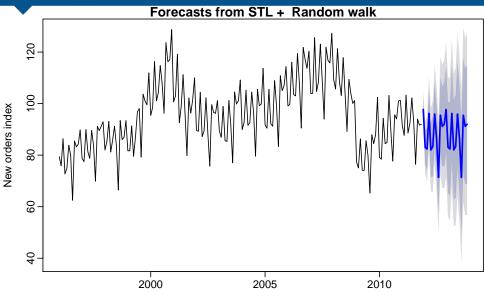
Seas adj elec equipment



Seas adj elec equipment



Seas adj elec equipment



How to do this in R

```
fit <- stl(elecequip, t.window=15,
  s.window="periodic", robust=TRUE)
eeadj <- seasadj(fit)</pre>
plot(naive(eeadj), xlab="New orders index")
fcast <- forecast(fit, method="naive")</pre>
plot(fcast, ylab="New orders index")
```

Decomposition and prediction intervals

- It is common to take the prediction intervals from the seasonally adjusted forecasts and modify them with the seasonal component.
- This ignores the uncertainty in the seasonal component estimate.
- It also ignores the uncertainty in the future seasonal pattern.

Decomposition and prediction intervals

- It is common to take the prediction intervals from the seasonally adjusted forecasts and modify them with the seasonal component.
- This ignores the uncertainty in the seasonal component estimate.
- It also ignores the uncertainty in the future seasonal pattern.

Decomposition and prediction intervals

- It is common to take the prediction intervals from the seasonally adjusted forecasts and modify them with the seasonal component.
- This ignores the uncertainty in the seasonal component estimate.
- It also ignores the uncertainty in the future seasonal pattern.