MONASH University

MONASH BUSINESS SCHOOL

Forecasting: principles and practice

Rob J Hyndman

3.4 Practical issues

Outline

Models for different frequencies

Models for annual data

- ETS, ARIMA, Dynamic regression

Models for different frequencies

Models for annual data

- ETS, ARIMA, Dynamic regression

Models for quarterly data

- ETS, ARIMA/SARIMA, Dynamic regression, Dynamic harmonic regression, STL+ETS, STL+ARIMA

Models for different frequencies

Models for annual data

- ETS, ARIMA, Dynamic regression

Models for quarterly data

- ETS, ARIMA/SARIMA, Dynamic regression, Dynamic harmonic regression, STL+ETS, STL+ARIMA

Models for monthly data

- ETS, ARIMA/SARIMA, Dynamic regression, Dynamic harmonic regression, STL+ETS, STL+ARIMA

Models for different frequencies

Models for weekly data

- ARIMA/SARIMA, Dynamic regression, Dynamic harmonic regression, STL+ETS, STL+ARIMA, TBATS

Models for different frequencies

Models for weekly data

- ARIMA/SARIMA, Dynamic regression, Dynamic harmonic regression, STL+ETS, STL+ARIMA, TBATS

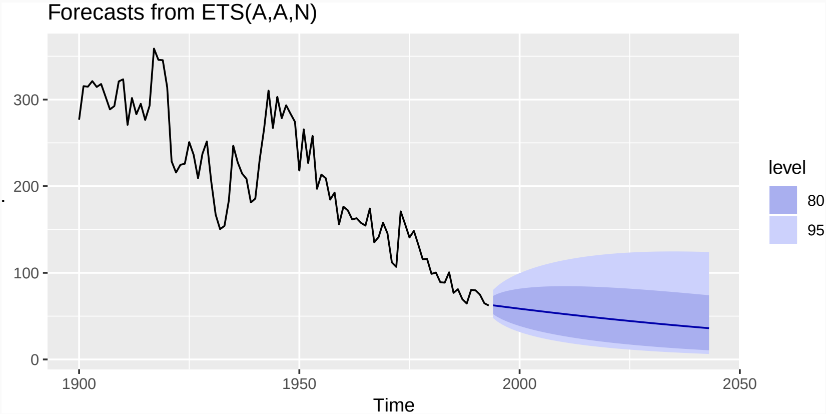
Models for daily, hourly and other sub-daily data

- ARIMA/SARIMA, Dynamic regression, Dynamic harmonic regression, STL+ETS, STL+ARIMA, TBATS

Outline

1. Models for different frequencies
2. **Ensuring forecasts stay within limits**
3. Forecast combinations
4. Missing values
5. Outliers

Positive forecasts



Outline

Forecast combinations

Clemen (1989)

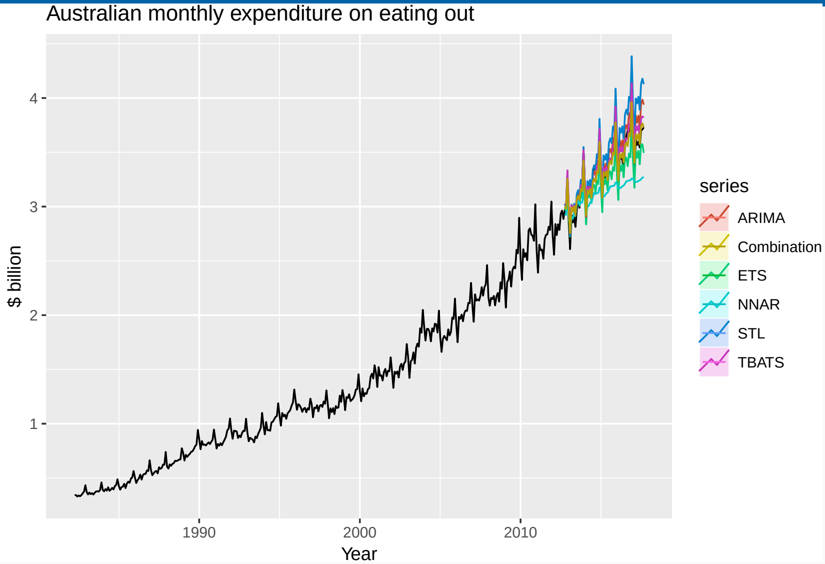
“The results have been virtually unanimous: combining multiple forecasts leads to increased forecast accuracy. …In many cases one can make dramatic performance improvements by simply averaging the forecasts.”

Forecast combinations

```
train <- window(auscafe, end=c(2012,9))
h <- length(auscafe) - length(train)
ETS <- forecast(ets(train), h=h)
ARIMA <- forecast(auto.arima(train, lambda=0, biasadj=TRUE),
  h=h)
STL <- stlf(train, lambda=0, h=h, biasadj=TRUE)
NNAR <- forecast(nnetar(train), h=h)
TBATS <- forecast(tbats(train, biasadj=TRUE), h=h)
Combination <- (ETS[["mean"]] + ARIMA[["mean"]] +
  STL[["mean"]] + NNAR[["mean"]] + TBATS[["mean"]])/5

autoplot(auscafe) +
  autolayer(ETS, series="ETS", PI=FALSE) +
  autolayer(ARIMA, series="ARIMA", PI=FALSE) +
  autolayer(STL, series="STL", PI=FALSE) +
  autolayer(NNAR, series="NNAR", PI=FALSE) +
  autolayer(TBATS, series="TBATS", PI=FALSE) +
  autolayer(Combination, series="Combination") +
  xlab("Year") + ylab("$ billion") +
  ggtitle("Australian monthly expenditure on eating out")
```

Forecast combinations



Forecast combinations

```
##          ETS       ARIMA     STL-ETS        NNAR         TBA
##   0.13699696  0.12146220  0.21446157  0.28932948  0.094060
```

Outline

Missing values

Functions which can handle missing values

- `auto.arima()`, `Arima()`
- `tslm()`
- `nnetar()`

Models which cannot handle missing values

- `ets()`
- `stl()`
- `stlf()`
- `tbats()`

Missing values

Functions which can handle missing values

- `auto.arima()`, `Arima()`
- `tslm()`
- `nnetar()`

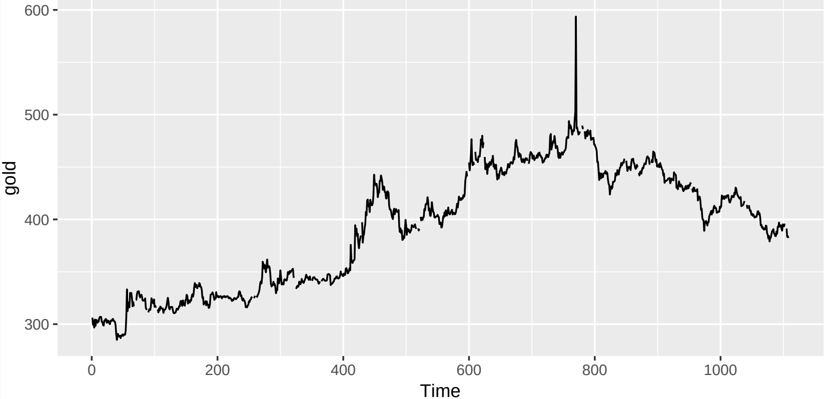
Models which cannot handle missing values

- `ets()`
- `stl()`
- `stlf()`
- `tbats()`

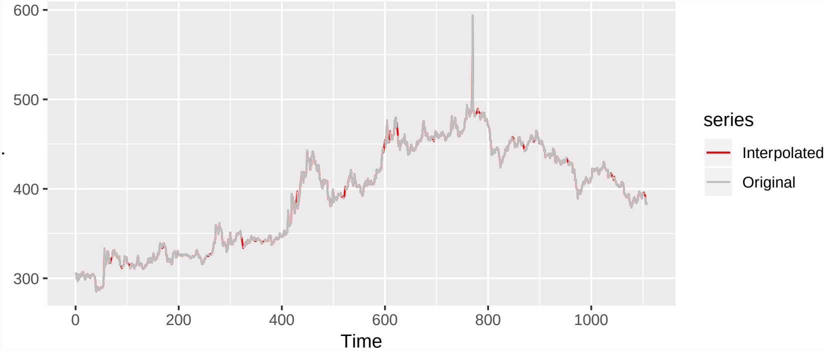
What to do?

1. Model section of data after last missing value.

Missing values

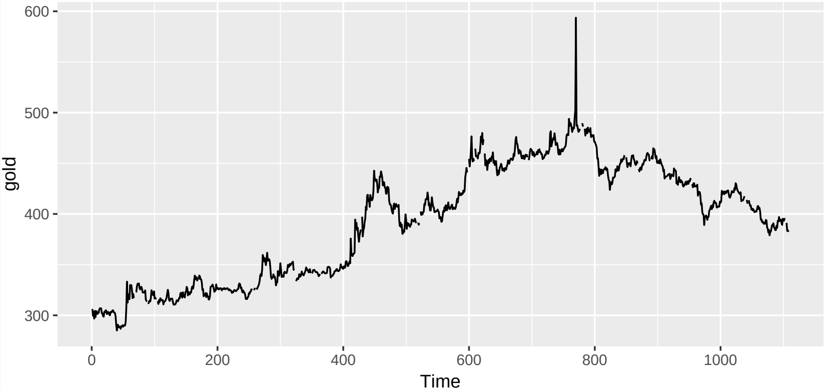


Missing values



Outline

Outliers



Outliers

```
## $index
## [1] 770
##
## $replacements
## [1] 494.9
```

Outliers