

Rob J Hyndman
George Athanasopoulos

FORECASTING

PRINCIPLES AND PRACTICE

A comprehensive introduction to the latest forecasting methods using R. Learn to improve your forecast accuracy using dozens of real data examples.



3RD EDITION

 **OTexts**
OPEN TEXTS FOR PRACTICE

5. The forecaster's toolbox

5.8 Evaluating point forecast accuracy

OTexts.org/fpp3/

Training and test sets



- A model which fits the training data well will not necessarily forecast well.
- A perfect fit can always be obtained by using a model with enough parameters.
- Over-fitting a model to data is just as bad as failing to identify a systematic pattern in the data.
- The test set must not be used for *any* aspect of model development or calculation of forecasts.
- Forecast accuracy is based only on the test set.

Forecast errors

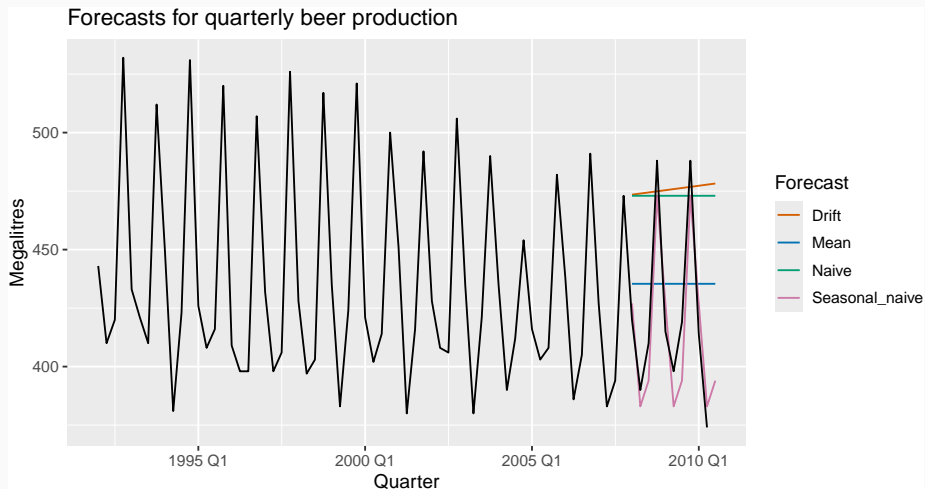
Forecast “error”: the difference between an observed value and its forecast.

$$e_{T+h} = y_{T+h} - \hat{y}_{T+h|T},$$

where the training data is given by $\{y_1, \dots, y_T\}$

- Unlike residuals, forecast errors on the test set involve multi-step forecasts.
- These are *true* forecast errors as the test data is not used in computing $\hat{y}_{T+h|T}$.

Measures of forecast accuracy



Measures of forecast accuracy

y_{T+h} = $(T + h)$ th observation, $h = 1, \dots, H$

$\hat{y}_{T+h|T}$ = its forecast based on data up to time T .

$$e_{T+h} = y_{T+h} - \hat{y}_{T+h|T}$$

$$\text{MAE} = \text{mean}(|e_{T+h}|)$$

$$\text{MSE} = \text{mean}(e_{T+h}^2)$$

$$\text{MAPE} = 100\text{mean}(|e_{T+h}|/|y_{T+h}|)$$

$$\text{RMSE} = \sqrt{\text{mean}(e_{T+h}^2)}$$

Measures of forecast accuracy

y_{T+h} = $(T + h)$ th observation, $h = 1, \dots, H$

$\hat{y}_{T+h|T}$ = its forecast based on data up to time T .

$$e_{T+h} = y_{T+h} - \hat{y}_{T+h|T}$$

$$\text{MAE} = \text{mean}(|e_{T+h}|)$$

$$\text{MSE} = \text{mean}(e_{T+h}^2)$$

$$\text{RMSE} = \sqrt{\text{mean}(e_{T+h}^2)}$$

$$\text{MAPE} = 100\text{mean}(|e_{T+h}|/|y_{T+h}|)$$

- MAE, MSE, RMSE are all scale dependent.
- MAPE is scale independent but is only sensible if $y_t \gg 0$ for all t , and y has a natural zero.

Scaled Errors

Proposed by Hyndman and Koehler (IJF, 2006).

- For non-seasonal time series, scale errors using naïve forecasts:

$$q_{T+h} = \frac{e_{T+h}}{\frac{1}{T-1} \sum_{t=2}^T |y_t - y_{t-1}|}.$$

Scaled Errors

Proposed by Hyndman and Koehler (IJF, 2006).

- For non-seasonal time series, scale errors using naïve forecasts:

$$q_{T+h} = \frac{e_{T+h}}{\frac{1}{T-1} \sum_{t=2}^T |y_t - y_{t-1}|}.$$

- For seasonal time series, scale forecast errors using seasonal naïve forecasts:

$$q_{T+h} = \frac{e_{T+h}}{\frac{1}{T-m} \sum_{t=m+1}^T |y_t - y_{t-m}|}.$$

Scaled errors

Mean Absolute Scaled Error

$$\text{MASE} = \text{mean}(|q_{T+h}|)$$

Scaled errors

Mean Absolute Scaled Error

$$\text{MASE} = \text{mean}(|q_{T+h}|)$$

Root Mean Squared Scaled Error

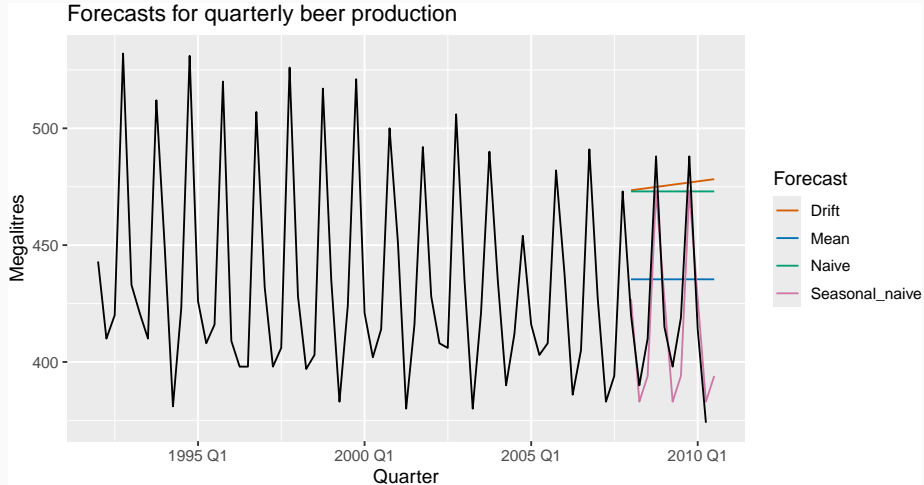
$$\text{RMSSE} = \sqrt{\text{mean}(q_{T+h}^2)}$$

where

$$q_{T+h}^2 = \frac{e_{T+h}^2}{\frac{1}{T-m} \sum_{t=m+1}^T (y_t - y_{t-m})^2},$$

and we set $m = 1$ for non-seasonal data.

Measures of forecast accuracy



Measures of forecast accuracy

```
recent_production <- aus_production |>
  filter(year(Quarter) >= 1992)
train <- recent_production |>
  filter(year(Quarter) <= 2007)
beer_fit <- train |>
  model(
    Mean = MEAN(Beer),
    Naive = NAIVE(Beer),
    Seasonal_naive = SNAIVE(Beer),
    Drift = RW(Beer ~ drift())
  )
beer_fc <- beer_fit |>
  forecast(h = 10)
```

Measures of forecast accuracy

```
accuracy(beer_fit) |>  
  arrange(.model) |>  
  select(.model, .type, RMSE, MAE, MAPE, MASE, RMSSE)
```

```
## # A tibble: 4 x 7
```

##	.model	.type	RMSE	MAE	MAPE	MASE	RMSSE
##	<chr>	<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
## 1	Drift	Training	65.3	54.8	12.2	3.83	3.89
## 2	Mean	Training	43.6	35.2	7.89	2.46	2.60
## 3	Naive	Training	65.3	54.7	12.2	3.83	3.89
## 4	Seasonal_naive	Training	16.8	14.3	3.31	1	1

Measures of forecast accuracy

```
accuracy(beer_fc, recent_production) |>  
  arrange(.model) |>  
  select(.model, .type, RMSE, MAE, MAPE, MASE, RMSSE)
```

```
## # A tibble: 4 x 7
```

##	.model	.type	RMSE	MAE	MAPE	MASE	RMSSE
##	<chr>	<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
## 1	Drift	Test	64.9	58.9	14.6	4.12	3.87
## 2	Mean	Test	38.4	34.8	8.28	2.44	2.29
## 3	Naive	Test	62.7	57.4	14.2	4.01	3.74
## 4	Seasonal_naive	Test	14.3	13.4	3.17	0.937	0.853