MONASH University

MONASH BUSINESS SCHOOL

Forecasting is not prophecy

Dealing with high-dimensional probabilistic forecasts in practice

Rob J Hyndman

What is it?

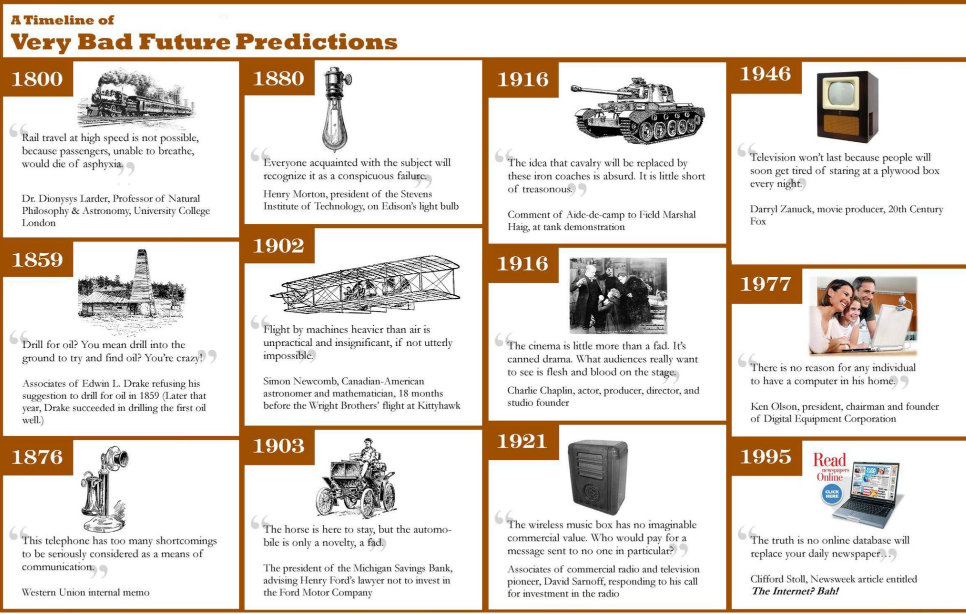
What is it?

Clay model of sheep's liver

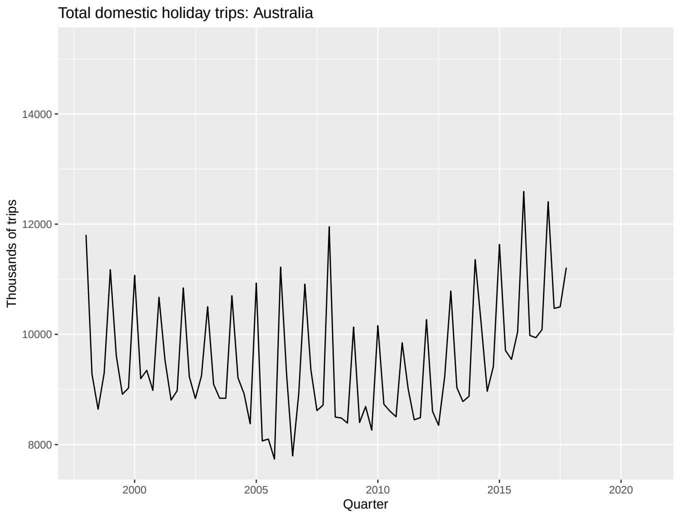
Used by Babylonian forecasters approximately 600 B.C.

Now in British Museum.

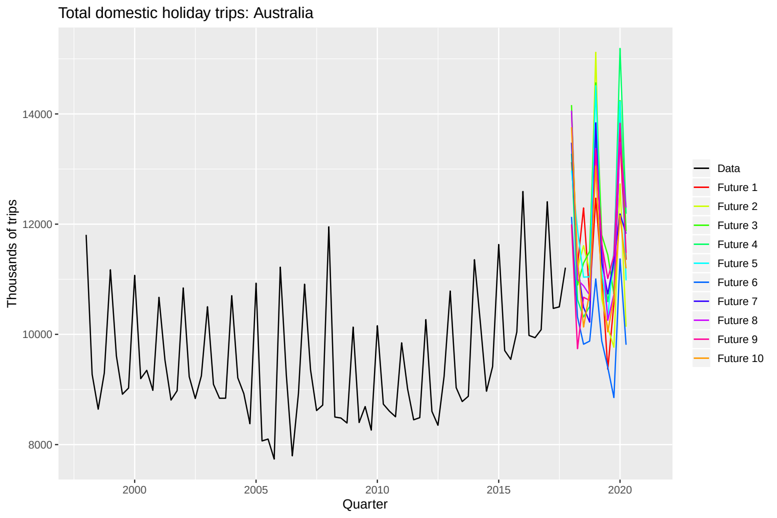
Modern prophecy



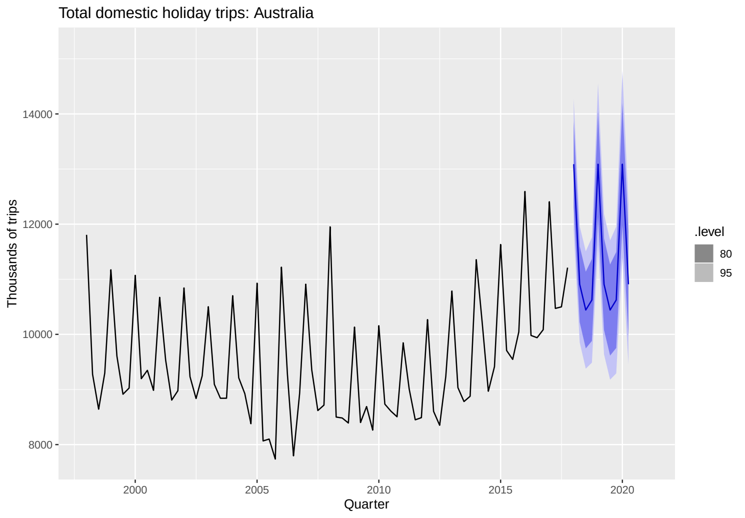
Probabilistic forecasting



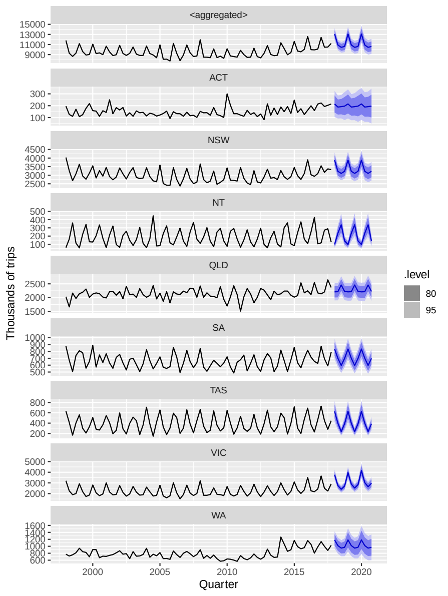
Probabilistic forecasting



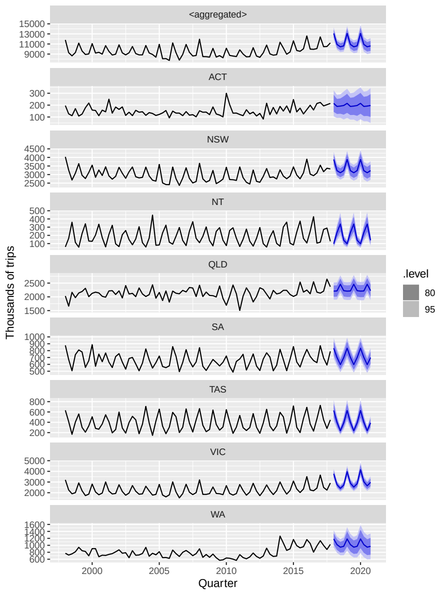
Probabilistic forecasting



Multivariate probabilistic forecasting



Multivariate probabilistic forecasting



These show the marginal distributions of the 9-dimensional distribution. In practice, we also want forecast for smaller geographic areas as well.

Who needs multivariate probabilistic forecasts?

- Tourism authorities forecasting visitor numbers to plan facilities.
- Manufacturing companies forecasting product demand to plan their supply chains.
- Call centres forecasting call volume to inform staff scheduling.
- Technology companies forecasting web traffic to maintain service levels.
- Energy companies forecasting electricity demand to prevent blackouts.

Challenges in multivariate probabilistic forecasting

1. How to produce high-dim probabilistic forecasts?
2. How to use and interpret high-dim probabilistic forecasts?
3. How to measure accuracy of high-dim probabilistic forecasts?

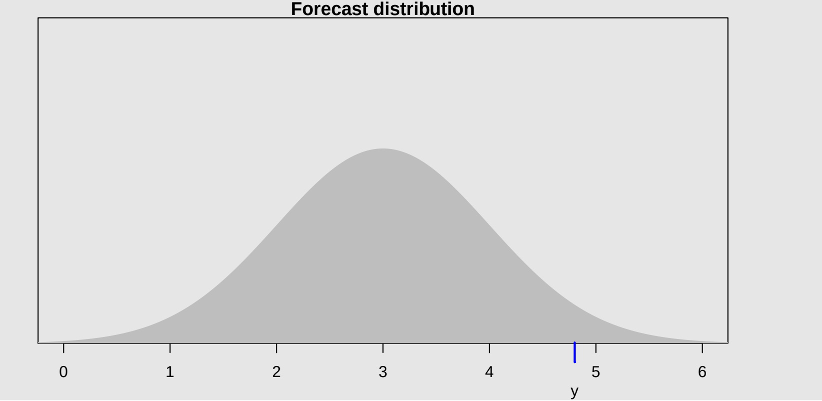
Challenges in multivariate probabilistic forecasting

1. How to produce high-dim probabilistic forecasts?
2. How to use and interpret high-dim probabilistic forecasts?
3. How to measure accuracy of high-dim probabilistic forecasts?

- Users care about marginal distributions, but we need multivariate distribution to compute them.
- Forecast reconciliation provides a way to efficiently compute marginal distributions.
- Probability scoring allows us to measure distributional forecast accuracy.

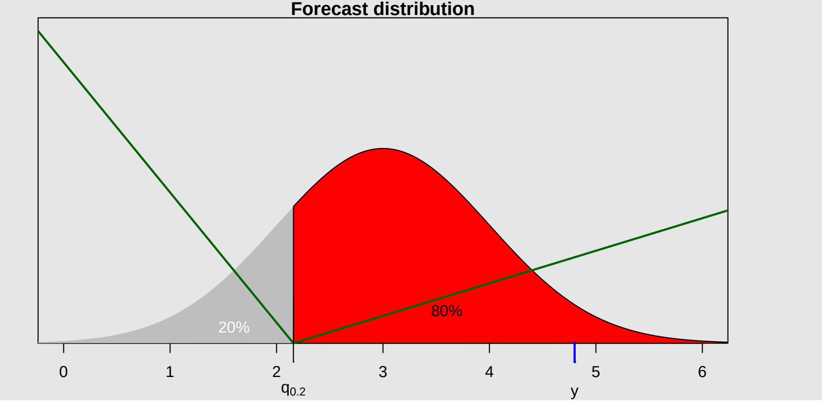
Probability scoring

- y = actual value
- F = forecast distribution
- $q_p = F^{-1}(p)$ = quantile with probability p.



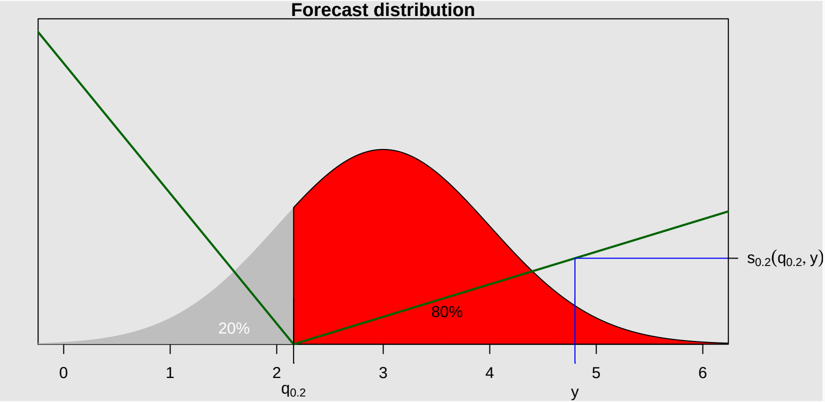
Probability scoring

- y = actual value
- F = forecast distribution
- $q_p = F^{-1}(p)$ = quantile with probability p.



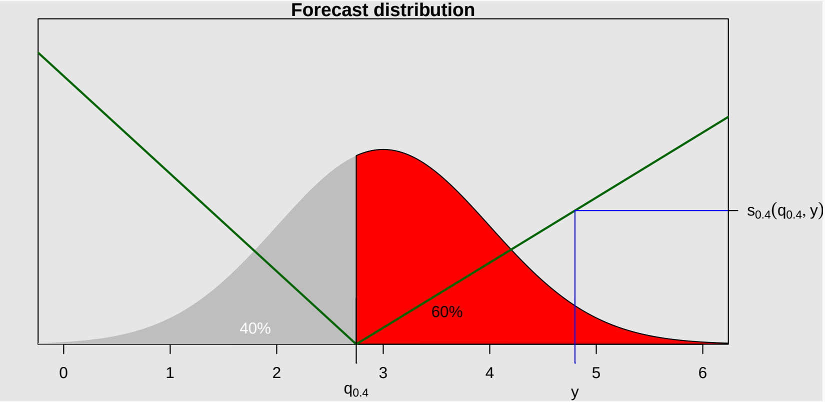
Probability scoring

- y = actual value
- F = forecast distribution
- $q_p = F^{-1}(p)$ = quantile with probability p.



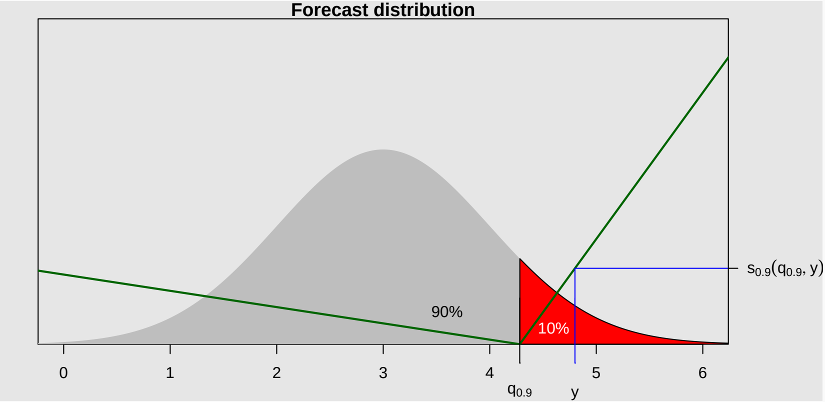
Probability scoring

- y = actual value
- F = forecast distribution
- $q_p = F^{-1}(p)$ = quantile with probability p.



Probability scoring

- y = actual value
- F = forecast distribution
- $q_p = F^{-1}(p)$ = quantile with probability p.



Probability scoring

- y = actual value
- F = forecast distribution
- $q_p = F^{-1}(p)$ = quantile with probability p.

Quantile Score (pinball loss)

$$s_p(q_p, y) = \begin{cases} (1-p)(q_p - y), & \text{if } y < q_p; \\ p(y - q_p), & \text{if } y \geq q_p. \end{cases}$$

Probability scoring

- y = actual value
- F = forecast distribution
- $q_p = F^{-1}(p)$ = quantile with probability p.

Quantile Score (pinball loss)

$$s_p(q_p, y) = \begin{cases} (1-p)(q_p - y), & \text{if } y < q_p; \\ p(y - q_p), & \text{if } y \geq q_p. \end{cases}$$

Continuous Rank Probability Score

$$\text{CRPS}(F, y) = 2\int_0^1 s_p(q_p, y)dp.$$

Probability scoring

- y = actual value
- F = forecast distribution
- $q_p = F^{-1}(p)$ = quantile with probability p.

Continuous ranked probability scores

$$\begin{aligned}\text{CRPS}(F, y) &= 2\int_0^1 s_p(q_p, y)dp. \\ &= \int_{-\infty}^{\infty} (F(x) - 1\{y \le x\})^2 dx\end{aligned}$$

Probability scoring

- y = actual value
- F = forecast distribution
- $q_p = F^{-1}(p)$ = quantile with probability p.

Continuous ranked probability scores

$$\begin{aligned}\text{CRPS}(F, y) &= 2\int_0^1 s_p(q_p, y)dp.\\ &= \int_{-\infty}^{\infty}(F(x) - 1\{y \leq x\})^2 dx\end{aligned}$$

- CRPS in same units as observations
- Expected CRPS minimized when observation y from same distribution as F.
- Maximizes sharpness subject to calibration

Tidyverts R packages

tidyverts.org

tsibble

tsibbledata

feasts

fable

tsibble objects

```r
library(tsibble)
library(tsibbledata)
aus_holidays
```

```
## # A tsibble: 6,080 x 4 [1Q]
## # Key:       Region, State [76]
##    Quarter Region   State Trips
##      <qtr> <chr>    <chr> <dbl>
##  1 1998 Q1 Adelaide SA     224.
##  2 1998 Q2 Adelaide SA     130.
##  3 1998 Q3 Adelaide SA     156.
##  4 1998 Q4 Adelaide SA     182.
##  5 1999 Q1 Adelaide SA     185.
##  6 1999 Q2 Adelaide SA     135.
##  7 1999 Q3 Adelaide SA     136.
##  8 1999 Q4 Adelaide SA     169.
##  9 2000 Q1 Adelaide SA     184.
```

tsibble objects

```r
library(tsibble)
library(tsibbledata)
aus_holidays
```

```
## # A tsibble: 6,080 x 4 [1Q]
## # Key:       Region, State [76]
##     Index   Region    State Trips
##       <qtr> <chr>     <chr> <dbl>
##   1 1998 Q1 Adelaide SA      224.
##   2 1998 Q2 Adelaide SA      130.
##   3 1998 Q3 Adelaide SA      156.
##   4 1998 Q4 Adelaide SA      182.
##   5 1999 Q1 Adelaide SA      185.
##   6 1999 Q2 Adelaide SA      135.
##   7 1999 Q3 Adelaide SA      136.
##   8 1999 Q4 Adelaide SA      169.
##   9 2000 Q1 Adelaide SA      184.
```

tsibble objects

```r
library(tsibble)
library(tsibbledata)
aus_holidays
```

```
## # A tsibble: 6,080 x 4 [1Q]
## # Key:       Region, State [76]
##      Index   Keys            Trips
##        <qtr> <chr>     <chr> <dbl>
##   1 1998 Q1 Adelaide SA      224.
##   2 1998 Q2 Adelaide SA      130.
##   3 1998 Q3 Adelaide SA      156.
##   4 1998 Q4 Adelaide SA      182.
##   5 1999 Q1 Adelaide SA      185.
##   6 1999 Q2 Adelaide SA      135.
##   7 1999 Q3 Adelaide SA      136.
##   8 1999 Q4 Adelaide SA      169.
##   9 2000 Q1 Adelaide SA      184.
```

tsibble objects

```r
library(tsibble)
library(tsibbledata)
aus_holidays
```

```
## # A tsibble: 6,080 x 4 [1Q]
## # Key:       Region, State [76]
##      Index    Keys              Measure
##      <qtr> <chr>     <chr> <dbl>
##  1 1998 Q1 Adelaide  SA     224.
##  2 1998 Q2 Adelaide  SA     130.
##  3 1998 Q3 Adelaide  SA     156.
##  4 1998 Q4 Adelaide  SA     182.
##  5 1999 Q1 Adelaide  SA     185.
##  6 1999 Q2 Adelaide  SA     135.
##  7 1999 Q3 Adelaide  SA     136.
##  8 1999 Q4 Adelaide  SA     169.
##  9 2000 Q1 Adelaide  SA     184.
```

tsibble objects

```r
library(tsibble)
library(tsibbledata)
aus_holidays
```

```
## # A tsibble: 6,080 x 4 [1Q]
## # Key:       Region, State [76]
##
##        <qtr> <chr>    <chr> <dbl>
##  1 1998 Q1 Adelaide SA     224.
##  2 1998 Q2 Adelaide SA     130.
##  3 1998 Q3 Adelaide SA     156.
##  4 1998 Q4 Adelaide SA     182.
##  5 1999 Q1 Adelaide SA     185.
##  6 1999 Q2 Adelaide SA     135.
##  7 1999 Q3 Adelaide SA     136.
##  8 1999 Q4 Adelaide SA     169.
##  9 2000 Q1 Adelaide SA     184.
```

Index

Keys

Measure

Domestic overnight holiday trips in thousands by state/region.

Compute all aggregates

```
aus_holidays_agg <- aus_holidays %>%
  aggregate_key(State/Region, Trips=sum(Trips))
aus_holidays_agg
```

```
## # A tsibble: 6,800 x 4 [1Q]
## # Key:       State, Region [85]
##    State        Region       Quarter  Trips
##    <chr>        <chr>          <qtr>  <dbl>
##  1 <aggregated> <aggregated> 1998 Q1 11806.
##  2 <aggregated> <aggregated> 1998 Q2  9276.
##  3 <aggregated> <aggregated> 1998 Q3  8642.
##  4 <aggregated> <aggregated> 1998 Q4  9300.
##  5 <aggregated> <aggregated> 1999 Q1 11172.
##  6 <aggregated> <aggregated> 1999 Q2  9608.
##  7 <aggregated> <aggregated> 1999 Q3  8914.
##  8 <aggregated> <aggregated> 1999 Q4  9026.
##  9 <aggregated> <aggregated> 2000 Q1 11071.
```

High-dimensional time series data

```r
aus_holidays_agg %>%
  filter(Quarter <= yearquarter("2015 Q4")) # 2 year test set
```

```
## # A tsibble: 6,120 x 4 [1Q]
## # Key:       State, Region [85]
##    State        Region       Quarter  Trips
##    <chr>        <chr>          <qtr>  <dbl>
##  1 <aggregated> <aggregated> 1998 Q1 11806.
##  2 <aggregated> <aggregated> 1998 Q2  9276.
##  3 <aggregated> <aggregated> 1998 Q3  8642.
##  4 <aggregated> <aggregated> 1998 Q4  9300.
##  5 <aggregated> <aggregated> 1999 Q1 11172.
##  6 <aggregated> <aggregated> 1999 Q2  9608.
##  7 <aggregated> <aggregated> 1999 Q3  8914.
##  8 <aggregated> <aggregated> 1999 Q4  9026.
##  9 <aggregated> <aggregated> 2000 Q1 11071.
## 10 <aggregated> <aggregated> 2000 Q2  9196.
```

Fit univariate models

```r
aus_holidays_agg %>%
  filter(Quarter <= yearquarter("2015 Q4")) %>%
  model(mean = MEAN(Trips), ets = ETS(Trips))
```

```
## # A mable: 85 x 4
## # Key:     State, Region [85]
##    State Region                 mean    ets
##    <chr> <chr>                  <model> <model>
##  1 ACT   Canberra               <MEAN>  <ETS(M,N,A)>
##  2 ACT   <aggregated>           <MEAN>  <ETS(M,N,A)>
##  3 NSW   Blue Mountains         <MEAN>  <ETS(M,N,M)>
##  4 NSW   Capital Country        <MEAN>  <ETS(A,N,N)>
##  5 NSW   Central Coast          <MEAN>  <ETS(M,N,M)>
##  6 NSW   Central NSW            <MEAN>  <ETS(M,N,A)>
##  7 NSW   Hunter                 <MEAN>  <ETS(A,N,A)>
##  8 NSW   New England North West <MEAN>  <ETS(M,N,N)>
##  9 NSW   North Coast NSW        <MEAN>  <ETS(A,N,A)>
```

Reconcile results

```r
aus_holidays_agg %>%
  filter(Quarter <= yearquarter("2015 Q4")) %>%
  model(mean = MEAN(Trips), ets = ETS(Trips)) %>%
  reconcile(ets = min_trace(ets))
```

```
## # A mable: 85 x 4
## # Key:     State, Region [85]
##    State Region                    mean    ets
##    <chr> <chr>                    <model> <model>
##  1 ACT   Canberra                 <MEAN>  <ETS(M,N,A)>
##  2 ACT   <aggregated>             <MEAN>  <ETS(M,N,A)>
##  3 NSW   Blue Mountains           <MEAN>  <ETS(M,N,M)>
##  4 NSW   Capital Country          <MEAN>  <ETS(A,N,N)>
##  5 NSW   Central Coast            <MEAN>  <ETS(M,N,M)>
##  6 NSW   Central NSW              <MEAN>  <ETS(M,N,A)>
##  7 NSW   Hunter                   <MEAN>  <ETS(A,N,A)>
##  8 NSW   New England North West   <MEAN>  <ETS(M,N,N)>
```

Produce probabilistic forecasts

```
aus_holidays_agg %>%
  filter(Quarter <= yearquarter("2015 Q4")) %>%
  model(mean = MEAN(Trips), ets = ETS(Trips)) %>%
  reconcile(ets = min_trace(ets)) %>%
  forecast(h = "2 years")
```

```
## # A fable: 1,360 x 6 [1Q]
## # Key:     State, Region, .model [170]
##    State Region        .model Quarter Trips .distribution
##    <chr> <chr>         <chr>    <qtr> <dbl> <dist>
##  1 ACT   Canberra      mean   2016 Q1  146. N(146, 1521)
##  2 ACT   Canberra      mean   2016 Q2  146. N(146, 1521)
##  3 ACT   Canberra      mean   2016 Q3  146. N(146, 1521)
##  4 ACT   Canberra      mean   2016 Q4  146. N(146, 1521)
##  5 ACT   Canberra      mean   2017 Q1  146. N(146, 1521)
##  6 ACT   Canberra      mean   2017 Q2  146. N(146, 1521)
##  7 ACT   Canberra      mean   2017 Q3  146. N(146, 1521)
```

Compare against test set

```
aus_holidays_agg %>%
  filter(Quarter <= yearquarter("2015 Q4")) %>%
  aggregate_key(State/Region, Trips=sum(Trips)) %>%
  model(mean = MEAN(Trips), ets = ETS(Trips)) %>%
  reconcile(coherent = min_trace(ets)) %>%
  forecast(h = "2 years") %>%
  accuracy(aus_holidays_agg, measures = list(CRPS = CRPS))
```

```
## # A tibble: 170 x 5
##    .model State Region                    .type  CRPS
##    <chr>  <chr> <chr>                     <chr> <dbl>
##  1 ets    ACT   Canberra                  Test   36.0
##  2 ets    ACT   <aggregated>              Test   36.0
##  3 ets    NSW   Blue Mountains            Test   15.2
##  4 ets    NSW   Capital Country           Test   15.0
##  5 ets    NSW   Central Coast             Test   11.5
##  6 ets    NSW   Central NSW               Test   15.9
```

Probability scoring

```r
aus_holidays_agg %>%
  filter(Quarter <= yearquarter("2015 Q4")) %>%
  aggregate_key(State/Region, Trips=sum(Trips)) %>%
  model(mean = MEAN(Trips), ets = ETS(Trips)) %>%
  reconcile(coherent = min_trace(ets)) %>%
  forecast(h = "2 years") %>%
  accuracy(aus_holidays_agg, measures = list(CRPS = CRPS)) %>%
  spread(key=.model, value=CRPS)
```

```
## # A tibble: 85 x 5
##     State Region                    .type   ets  mean
##     <chr> <chr>                     <chr> <dbl> <dbl>
##   1 ACT   Canberra                  Test   36.0  34.4
##   2 ACT   <aggregated>              Test   36.0  34.4
##   3 NSW   Blue Mountains            Test   15.2  35.8
##   4 NSW   Capital Country           Test   15.0  16.2
##   5 NSW   Central Coast             Test   11.5  23.3
```

Probability scoring

```r
aus_holidays_agg %>%
  filter(Quarter <= yearquarter("2015 Q4")) %>%
  aggregate_key(State/Region, Trips=sum(Trips)) %>%
  model(mean = MEAN(Trips), ets = ETS(Trips)) %>%
  reconcile(coherent = min_trace(ets)) %>%
  forecast(h = "2 years") %>%
  accuracy(aus_holidays_agg, measures = list(CRPS = CRPS)) %>%
  spread(key=.model, value=CRPS) %>%
  mutate(RelCRPS = ets/mean)
```

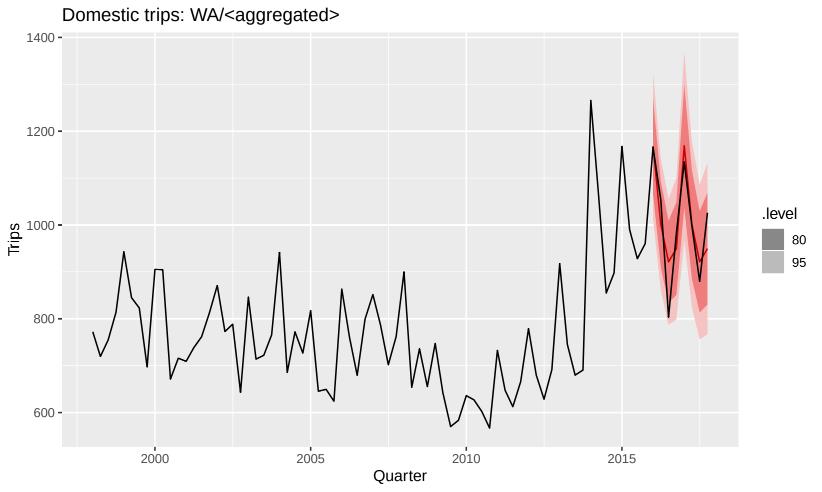
```
## # A tibble: 85 x 6
##    State Region                     .type   ets  mean RelCRPS
##    <chr> <chr>                      <chr> <dbl> <dbl>   <dbl>
##  1 ACT   Canberra                   Test   36.0  34.4   1.05
##  2 ACT   <aggregated>               Test   36.0  34.4   1.05
##  3 NSW   Blue Mountains             Test   15.2  35.8   0.424
##  4 NSW   Capital Country            Test   15.0  16.2   0.925
```

Probability scoring

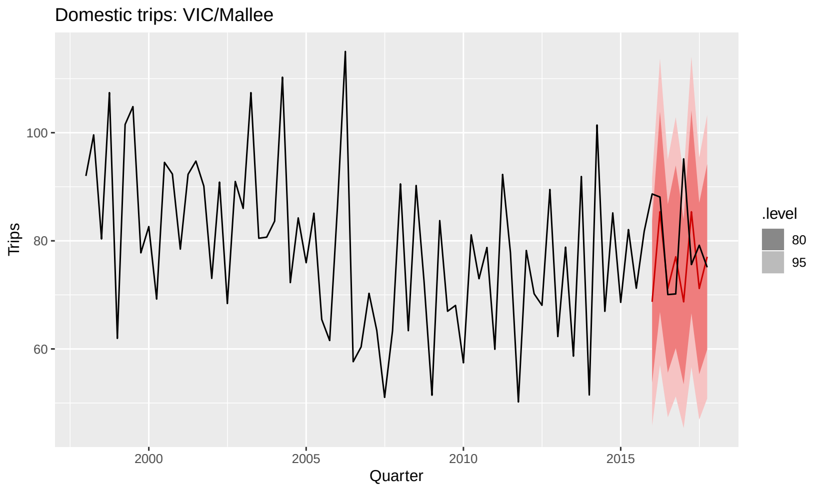
```r
aus_holidays_agg %>%
  filter(Quarter <= yearquarter("2015 Q4")) %>%
  aggregate_key(State/Region, Trips=sum(Trips)) %>%
  model(mean = MEAN(Trips), ets = ETS(Trips)) %>%
  reconcile(coherent = min_trace(ets)) %>%
  forecast(h = "2 years") %>%
  accuracy(aus_holidays_agg, measures = list(CRPS = CRPS)) %>%
  spread(key=.model, value=CRPS) %>%
  mutate(RelCRPS = ets/mean) %>%
  arrange(RelCRPS)
```

```
## # A tibble: 85 x 6
##    State Region                        .type   ets  mean RelCRPS
##    <chr> <chr>                         <chr> <dbl> <dbl>   <dbl>
##  1 WA    <aggregated>                  Test  34.8  177.    0.196
##  2 NSW   South Coast                   Test  26.2  101.    0.261
##  3 WA    Australia's South West        Test  27.1  102.    0.265
```

Best forecast (RelCRPS)



Worst forecast (RelCRPS)



Challenges in multivariate probabilistic forecasting

1. The fable package produces high-dim probabilistic forecasts?
2. We can focus on marginal distributions for interpretation, while needing the full distribution for calculation.
3. Forecast reconciliation provides a way to efficiently compute marginal distributions.
4. Probability scoring allows us to measure distributional forecast accuracy.

Acknowledgements

Mitchell O'Hara-Wild

Earo Wang

tidyverts.org

robjhyndman.com