Forecasting prison numbers: a grouped time series approach

George Athanasopoulos

with Tom Steel & Don Weatherburn
1 Hierarchical and grouped time series
2 BLUF: Best Linear Unbiased Forecasts
3 Forecasting Australian prison population
4 Other forecast reconciliation settings
5 References
Produce accurate but also detailed forecasts of prisoner numbers at the aggregate national level but also for multiple groupings based on attributes (and their interactions) that are of interest to a variety of policy makers and correctional administrators.

The level of detail and the coherent nature of the forecasts enables informed and importantly aligned decision making across multiple departments and at all levels of management: strategic, tactical and operational.
Aim

- Produce **accurate** but also **detailed** forecasts of prisoner numbers at the aggregate national level but also for multiple groupings based on attributes (and their interactions) that are of interest to a variety of policy makers and correctional administrators.

- The level of detail and the **coherent** nature of the forecasts enables **informed** and importantly **aligned** decision making across multiple departments and at all levels of management: strategic, tactical and operational.
Australian prison population
Demographics (243 series = 1 + 16 + 60 + 104 + 64):
- State (8)
- Sex (2)
- Legal Status (2)
- Indigenous Status (2)

ANZ Standard Offence Classification (243 series):
- Divisions (16) (Homicide, Sexual Assault, Robbery, Illicit drugs, etc.)
- Subdivisions (66) (Manslaughter and driving causing death, Murder, Attempted Murder, etc.)
- Groups (160) (Manslaughter, Driving causing death, etc.)
Forecasting prison population

- Demographics (243 series = 1 + 16 + 60 + 104 + 64):
  - State (8)
  - Sex (2)
  - Legal Status (2)
  - Indigenous Status (2)

- ANZ Standard Offence Classification (243 series):
  - Divisions (16) (Homicide, Sexual Assault, Robbery, Illicit drugs, etc.)
  - Subdivisions (66) (Manslaughter and driving causing death, Murder, Attempted Murder, etc.)
  - Groups (160) (Manslaughter, Driving causing death, etc)
Forecasting prison population

Demographics (243 series = 1 + 16 + 60 + 104 + 64):
- State (8)
- Sex (2)
- Legal Status (2)
- Indigenous Status (2)

ANZ Standard Offence Classification (243 series):
- Divisions (16) (Homicide, Sexual Assault, Robbery, Illicit drugs, etc.)
- Subdivisions (66) (Manslaughter and driving causing death, Murder, Attempted Murder, etc.)
- Groups (160) (Manslaughter, Driving causing death, etc)
Australian domestic tourism

Hierarchical:
- Australia (1)
- States (7)
- Zones (27)
- Regions (76)

Total: 111 series

Grouped:
- Grand total: 555 series

Purpose of Travel: \(\times 4\)
- Holiday
- VFR
- Business
- Other

Total: 444 series
Total number of Monash Students

- Faculty (8)
- Campus (2 + Other)
- Funding source (3)
- Course level (3)
- Commencing/returning (2)
- Courses (457)
- Units (5605) (not sure we will get here).
Forecasting student numbers

- Total number of Monash Students
  - Faculty (8)
  - Campus (2 + Other)
  - Funding source (3)
  - Course level (3)
  - Commencing/returning (2)
  - Courses (457)
  - Units (5605) (not sure we will get here).
  - Total: 152,289 time series.
Forecasting student numbers

Challenges:

- Large number of series to forecast.
- We want a flexible forecasting process using all information available.
- We want forecasts to be coherent (add up).

- Courses (457)
- Units (5605) (not sure we will get here).
Key idea

- Forecast all series at all levels of aggregation or groupings (in contrast to typical bottom-up, top-down or middle-out approaches).

- Reconcile the forecasts so they add up correctly using least squares optimization, i.e., find closest reconciled forecasts to the original forecasts.
Key idea

- Forecast all series at all levels of aggregation or groupings (in contrast to typical bottom-up, top-down or middle-out approaches).

- Reconcile the forecasts so they add up correctly using least squares optimization, i.e., find closest reconciled forecasts to the original forecasts.
Hierarchical time series

$y_{\text{Tot},t}$: observed aggregate of all series at time $t$.

$y_{X,t}$: observation on series $X$ at time $t$.  

- Total
  - A
  - B
  - C
Hierarchical time series

Key concept:

I can construct all time series in my collection if I know the aggregation structure and the bottom-level series.

\[ y_{\text{Tot},t} \] observed aggregate of all series at time \( t \).

\[ y_{X,t} \] observation on series \( X \) at time \( t \).
Hierarchical time series

\[ y_t = \begin{pmatrix} y_{Tot,t} \\ y_{A,t} \\ y_{B,t} \\ y_{C,t} \end{pmatrix} \]

- \( y_{Tot,t} \): observed aggregate of all series at time \( t \).
- \( y_{X,t} \): observation on series \( X \) at time \( t \).
Hierarchical time series

\[ y_{Tot,t} : \text{observed aggregate of all series at time } t. \]

\[ y_{X,t} : \text{observation on series } X \text{ at time } t. \]

\[
y_t = \begin{pmatrix}
y_{Tot,t} \\
y_{A,t} \\
y_{B,t} \\
y_{C,t}
\end{pmatrix} = \begin{pmatrix}
1 & 1 & 1 \\
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{pmatrix} \begin{pmatrix}
y_{A,t} \\
y_{B,t} \\
y_{C,t}
\end{pmatrix}
\]

\[ S \]
Hierarchical time series

\[ y_{Tot,t} : \text{observed aggregate of all series at time } t. \]

\[ y_{X,t} : \text{observation on series } X \text{ at time } t. \]

\[ b_t : \text{vector of all bottom-level series at time } t. \]

\[ y_t = \begin{pmatrix}
    y_{Tot,t} \\
    y_{A,t} \\
    y_{B,t} \\
    y_{C,t}
\end{pmatrix} = \begin{pmatrix}
    1 & 1 & 1 \\
    1 & 0 & 0 \\
    0 & 1 & 0 \\
    0 & 0 & 1
\end{pmatrix} \begin{pmatrix}
    y_{A,t} \\
    y_{B,t} \\
    y_{C,t}
\end{pmatrix} \]

\[ S \]

\[ b_t \]
Hierarchical time series

\[ y_{tot,t} : \text{observed aggregate of all series at time } t. \]

\[ y_{X,t} : \text{observation on series } X \text{ at time } t. \]

\[ b_t : \text{vector of all bottom-level series at time } t. \]

\[
y_t = \begin{pmatrix}
y_{Tot,t} \\
y_{A,t} \\
y_{B,t} \\
y_{C,t}
\end{pmatrix} = \begin{pmatrix}
1 & 1 & 1 \\
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{pmatrix} \begin{pmatrix}
y_{A,t} \\
y_{B,t} \\
y_{C,t}
\end{pmatrix}
\]

\[ y_t = Sb_t \]
Grouped time series

\[
y_t = \begin{pmatrix}
y_{\text{Tot},t} \\
y_{A,t} \\
y_{B,t} \\
y_{X,t} \\
y_{Y,t} \\
y_{AX,t} \\
y_{AY,t} \\
y_{BX,t} \\
y_{BY,t}
\end{pmatrix} = \begin{pmatrix}
1 & 1 & 1 & 1 \\
1 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 \\
1 & 0 & 1 & 0 \\
0 & 1 & 0 & 1 \\
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\]

\[
\begin{pmatrix}
y_{AX,t} \\
y_{AY,t} \\
y_{BX,t} \\
y_{BY,t}
\end{pmatrix} = b_t
\]

\[
S
\]
Grouped time series

\[
\mathbf{y}_t = \begin{pmatrix}
\mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\
\mathbf{1} & \mathbf{1} & 0 & 0 \\
0 & 0 & 1 & 1 \\
1 & 0 & 1 & 0 \\
0 & 1 & 0 & 1 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\]
Grouped time series

\[ y_t = S b_t \]

\[
\begin{pmatrix}
  y_{\text{Tot},t} \\
  y_{A,t} \\
  y_{B,t} \\
  y_{X,t} \\
  y_{Y,t} \\
  y_{AX,t} \\
  y_{AY,t} \\
  y_{BX,t} \\
  y_{BY,t}
\end{pmatrix} =
\begin{pmatrix}
  1 & 1 & 1 & 1 \\
  1 & 1 & 0 & 0 \\
  0 & 0 & 1 & 1 \\
  1 & 0 & 1 & 0 \\
  1 & 0 & 1 & 0 \\
  0 & 1 & 0 & 1 \\
  0 & 0 & 1 & 0 \\
  0 & 0 & 0 & 1
\end{pmatrix}
\]

Forecasting aggregation structures

Hierarchical and grouped time series

12
We can now deal effectively with both hierarchical and grouped aggregation structures.
Hierarchical and grouped time series

Every collection of time series with linear aggregation constraints can be written as:

\[ y_t = S b_t \]

where

- \( y_t \) is a vector of all series at time \( t \).
- \( S \) is a “summing matrix” containing the aggregation constraints.
- \( b_t \) is a vector of the most disaggregated series at time \( t \).
1 Hierarchical and grouped time series
2 BLUF: Best Linear Unbiased Forecasts
3 Forecasting Australian prison population
4 Other forecast reconciliation settings
5 References
Let $\hat{y}_T(h)$ be a vector of base (initial) $h$-step forecasts made at time $T$, stacked in same order as $y_t$.
(These will almost certainly never add up.)

Reconciled (coherent) forecasts must be of the form:

$$\tilde{y}_T(h) = SP\hat{y}_T(h)$$

for some matrix $P$.

- $P$ extracts and combines base forecasts $\hat{y}_T(h)$ to get bottom-level forecasts, $P\hat{y}_T(h) = \hat{b}_T(h)$. E.g., $P = [0|I_m]$ for bottom-up, $P = [p|0_{n-1}]$ for top-down.
- $S$ adds them up, $\tilde{y}_T(h) = S\hat{b}_T(h)$. 

Forecasting aggregation structures

BLUF: Best Linear Unbiased Forecasts
Forecasting framework

Let \( \hat{y}_T(h) \) be a vector of base (initial) \( h \)-step forecasts made at time \( T \), stacked in same order as \( y_t \).
(These will almost certainly never add up.)

Reconciled (coherent) forecasts must be of the form:

\[
\tilde{y}_T(h) = SP\hat{y}_T(h)
\]

for some matrix \( P \).

- \( P \) extracts and combines base forecasts \( \hat{y}_T(h) \) to get bottom-level forecasts, \( P\hat{y}_T(h) = \hat{b}_T(h) \). E.g., \( P = [0|I_m] \) for bottom-up, \( P = [p|0_{n-1}] \) for top-down.

- \( S \) adds them up, \( \tilde{y}_T(h) = S\hat{b}_T(h) \).
Let $\hat{\mathbf{y}}_T(h)$ be a vector of base (initial) $h$-step forecasts made at time $T$, stacked in same order as $\mathbf{y}_t$. (These will almost certainly never add up.)

Reconciled (coherent) forecasts must be of the form:

$$\tilde{\mathbf{y}}_T(h) = S P \hat{\mathbf{y}}_T(h)$$

for some matrix $P$.

- $P$ extracts and combines base forecasts $\hat{\mathbf{y}}_T(h)$ to get bottom-level forecasts, $P \hat{\mathbf{y}}_T(h) = \hat{\mathbf{b}}_T(h)$. E.g., $P = [\mathbf{0} | I_m]$ for bottom-up, $P = [p | 0_{n-1}]$ for top-down.

- $S$ adds them up, $\tilde{\mathbf{y}}_T(h) = S \hat{\mathbf{b}}_T(h)$.  

BLUF: Best Linear Unbiased Forecasts
Let $\hat{y}_T(h)$ be a vector of base (initial) $h$-step forecasts made at time $T$, stacked in same order as $y_t$.  
(These will almost certainly never add up.)

Reconciled (coherent) forecasts must be of the form:

$$\tilde{y}_T(h) = SP\hat{y}_T(h)$$

for some matrix $P$.

- $P$ extracts and combines base forecasts $\hat{y}_T(h)$ to get bottom-level forecasts, $P\hat{y}_T(h) = \hat{b}_T(h)$. E.g., $P = [0|I_m]$ for bottom-up, $P = [p|0_{n-1}]$ for top-down.

- $S$ adds them up, $\tilde{y}_T(h) = S\hat{b}_T(h)$. 

Let $\hat{y}_T(h)$ be a vector of base (initial) $h$-step forecasts made at time $T$, stacked in same order as $y_t$. (These will almost certainly never add up.)

Reconciled (coherent) forecasts must be of the form:

$$\tilde{y}_T(h) = SP\hat{y}_T(h)$$

for some matrix $P$.

- $P$ extracts and combines base forecasts $\hat{y}_T(h)$ to get bottom-level forecasts, $P\hat{y}_T(h) = \hat{b}_T(h)$. E.g., $P = [0|l_m]$ for bottom-up, $P = [p|0_{n-1}]$ for top-down.

- $S$ adds them up, $\tilde{y}_T(h) = S\hat{b}_T(h)$. 

Let $\hat{y}_T(h)$ be a vector of base (initial) $h$-step forecasts made at time $T$, stacked in same order as $y_t$. (These will almost certainly never add up.)

Reconciled (coherent) forecasts must be of the form:

$$\tilde{y}_T(h) = SP\hat{y}_T(h)$$

for some matrix $P$.

- $P$ extracts and combines base forecasts $\hat{y}_T(h)$ to get bottom-level forecasts, $P\hat{y}_T(h) = \hat{b}_T(h)$. E.g., $P = [0|I_m]$ for bottom-up, $P = [p|0_{n-1}]$ for top-down.

- $S$ adds them up, $\tilde{y}_T(h) = S\hat{b}_T(h)$. 
Let $\hat{y}_T(h)$ be a vector of base (initial) $h$-step forecasts made at time $T$, stacked in same order as $y_t$.
(These will almost certainly never add up.)

Reconciled (coherent) forecasts must be of the form:

$$\tilde{y}_T(h) = SP\hat{y}_T(h)$$

for some matrix $P$.

- $P$ extracts and combines base forecasts $\hat{y}_T(h)$ to get bottom-level forecasts, $P\hat{y}_T(h) = \hat{b}_T(h)$. E.g., $P = [0|I_m]$ for bottom-up, $P = [p|0_{n-1}]$ for top-down.

- $S$ adds them up, $\tilde{y}_T(h) = S\hat{b}_T(h)$.
Forecasting framework

Let $\hat{y}_T(h)$ be a vector of base (initial) $h$-step forecasts made at time $T$, stacked in same order as $y_t$. (These will almost certainly never add up.)

Reconciled (coherent) forecasts must be of the form:

$$\tilde{y}_T(h) = SP\hat{y}_T(h)$$

Key limitation:

 Traditional approaches use information only from a single level.

 Can we do better?

bottom-level forecasts, $P\hat{y}_T(h) = b_T(h)$. E.g., $P = [0|I_m]$ for bottom-up, $P = [p|0_{n-1}]$ for top-down.

$S$ adds them up, $\tilde{y}_T(h) = Sb_T(h)$. 
Optimal reconciliation approach

\[ \tilde{y}_T(h) = SP\hat{y}_T(h) \]

The error variance of the reconciled forecasts is

\[ \text{Var}(y_{T+h} - \tilde{y}_T(h)) = SPW_h P'S' \]

where \( W_h = \text{Var}(y_{T+h} - \hat{y}_T(h)) \), error variance of base forecasts.

Theorem: BLUF via trace minimisation (MinT)

For any \( P \) satisfying \( SPS = S \)

\[ \min_P \text{tr}[SPW_h P'S'] \]

has unique solution at \( P = (S'W_h^{-1}S)^{-1}S'W_h^{-1} \).

- Estimating \( W_h \) is challenging especially for \( h > 1 \).
Optimal reconciliation approach

\[ \tilde{y}_T(h) = SP\hat{y}_T(h) \]

The error variance of the reconciled forecasts is

\[ \text{Var}(y_{T+h} - \tilde{y}_T(h)) = SPW_hP'S' \]

where \( W_h = \text{Var}(y_{T+h} - \hat{y}_T(h)) \), error variance of base forecasts.

**Theorem: BLUF via trace minimisation (MinT)**

For any \( P \) satisfying \( SPS = S \)

\[ \min_P \text{tr}[SPW_hP'S'] \]

has unique solution at \( P = (S'W_h^{-1}S)^{-1}S'W_h^{-1} \).

- Estimating \( W_h \) is challenging especially for \( h > 1 \).
Optimal reconciliation forecasts

\[ \tilde{y}_T(h) = S(S'W_h^{-1}S)^{-1}S'W_h^{-1}\hat{y}_T(h) \]

Reconciled forecasts
Base forecasts
WLS Solution

- We assume that $W_h = k_h W_1$ and approximate $W_1$ by its diagonal using in-sample one-step ahead forecast errors.
- Easy to estimate, and places weight where we have best forecasts.

Outline

1. Hierarchical and grouped time series
2. BLUF: Best Linear Unbiased Forecasts
3. Forecasting Australian prison population
4. Other forecast reconciliation settings
5. References
### Australian Prison Population

<table>
<thead>
<tr>
<th>Location:</th>
<th>Sex:</th>
<th>Indigenous status:</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSW</td>
<td>Male</td>
<td>ATSI</td>
</tr>
<tr>
<td>QLD</td>
<td>Female</td>
<td>Non-ATSI</td>
</tr>
<tr>
<td>SA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VIC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remanded</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sentenced</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Australian Prison Population

Location:
- NSW
- QLD
- SA
- TAS
- VIC
- WA
- ACT
- NT

Sex:
- Male
- Female

Legal status:
- Remanded
- Sentenced

Indigenous status:
- ATSI
- Non-ATSI

Grouped Total:
243 series
Australian Prison Population

Forecast evaluation setup

- All adult prisoners in Australia: 2005Q1-2016Q4. (ABS corrective services database).
- 36 obs as training set and generate base forecasts with `auto.arima()` and `ets()` for $h = 1$ to 8-steps ahead.
- Obtain coherent forecasts using optimal reconciliation (WLS), and bottom-up.
- Use a rolling window: 12 1-step, 11 2-steps,....,4 8-steps ahead forecasts for evaluation.
Forecast evaluation - MAPE

Forecast performance by Forecast Horizons

Forecast method:

- ARIMA Base
- ARIMA WLS
- ETS Base
- ETS WLS
- ARIMA Bottom-up
- ETS Bottom-up

Forecasting aggregation structures

Forecasting Australian prison population
Forecasting evaluation - RMSE

Forecast performance by Forecast Horizons

- Forecast method:
  - ARIMA Base
  - ARIMA WLS
  - ETS Base
  - ETS WLS

Forecasting aggregation structures
Forecasting Australian prison population
### Forecast evaluation - Levels

**ARIMA (Forecast Horizons: 1–4)**

<table>
<thead>
<tr>
<th>Series</th>
<th>MAPE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State/NSW</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State/QLD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State/VIC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State/SA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State/TAS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State/WA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State/NT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State/ACT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sex/m</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sex/f</td>
<td></td>
<td></td>
</tr>
<tr>
<td>leg/r</td>
<td></td>
<td></td>
</tr>
<tr>
<td>leg/s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ind/I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ind/NI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State x Legal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State x Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State x Indigenous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex x Indigenous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex x Legal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indigenous x Legal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State x Legal x Indigenous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State x Sex x Indigenous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State x Sex x Legal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex x Legal x Indigenous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State x Sex x Legal x Indigenous</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

% difference of forecast errors

**ARIMA (Forecast Horizons: 5–8)**

<table>
<thead>
<tr>
<th>Series</th>
<th>MAPE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State/NSW</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State/QLD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State/VIC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State/SA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State/TAS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State/WA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State/NT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State/ACT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sex/m</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sex/f</td>
<td></td>
<td></td>
</tr>
<tr>
<td>leg/r</td>
<td></td>
<td></td>
</tr>
<tr>
<td>leg/s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ind/I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ind/NI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State x Legal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State x Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State x Indigenous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex x Indigenous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex x Legal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indigenous x Legal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State x Legal x Indigenous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State x Sex x Indigenous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State x Sex x Legal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex x Legal x Indigenous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State x Sex x Legal x Indigenous</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

% difference of forecast errors

---

Forecasting aggregation structures | Forecasting Australian prison population | 22
Forecast evaluation - Levels

ETS (Forecast Horizons: 1–4)

ETS (Forecast Horizons: 5–8)

Forecasting aggregation structures

Forecasting Australian prison population
1. Hierarchical and grouped time series

2. BLUF: Best Linear Unbiased Forecasts

3. Forecasting Australian prison population

4. Other forecast reconciliation settings

5. References
Figure: A simple two-level cross-sectional hierarchy.
Temporal reconciliation

**Figure:** A simple two-level cross-sectional hierarchy.

**Figure:** A temporal hierarchy for quarterly data.

Cross-temporal reconciliation
Cross-temporal reconciliation

Reconciliation (especially cross-temporal) offers a single/aligned view of the future to all decision makers, removing any organisational friction from misaligned decisions.

More crucially, it offers a data driven way to break within and between organisations information silos.
Reconciliation (especially cross-temporal) offers a single/aligned view of the future to all decision makers, removing any organisational friction from misaligned decisions.

More crucially, it offers a data driven way to break within and between organisations information silos.
References


Rob J Hyndman and Nikolaos Kourentzes (2016). *thief: Temporal Hierarchical Forecasting*. R package v0.2 on CRAN.


Rob J Hyndman and Nikolaos Kourentzes (2016). *thief: Temporal Hierarchical Forecasting*. R package v0.2 on CRAN.

Thank you!
Total Emergency Admissions via A&E

Forecasting aggregation structures

References 30