Probabilistic Forecast Reconciliation For Emergency Services Demand

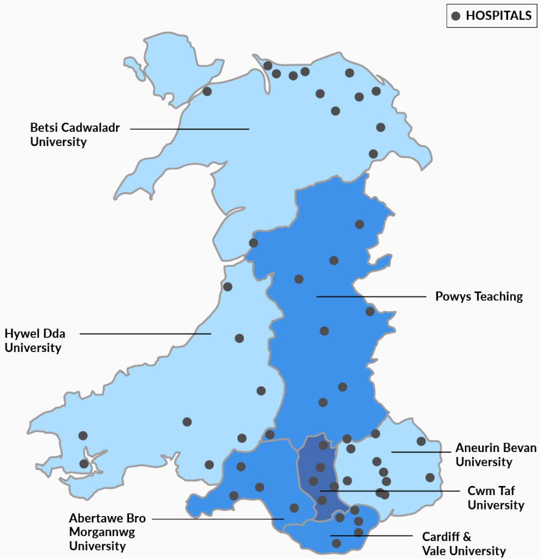
Rob J Hyndman
Bahman Rostami-Tabar

MONASH University

robjhyndman.com/asc2023

Photo by Graham Richardson on Wikimedia

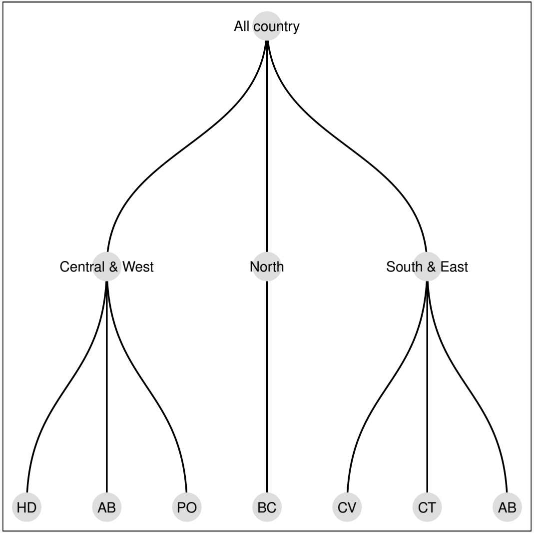
Wales Health Board Areas



Data

- Daily number of attended incidents:
  1 October 2015 – 31 July 2019
- Disaggregated by:
  - control area
  - health board
  - priority
  - nature of incidents
- 2,142,000 rows observations from 1,530 time series.

Data structure



*

Priority

Red
Amber
Green

*

Nature of incident

Chest pain
Stroke
Breathing problem
...
Abdominal pain

Data structure

Level	Number of series
All country	1
Control	3
Health board	7
Priority	3
Priority * Control	9
Priority * Health board	21
Nature of incident	35
Nature of incident * Control	105
Nature of incident * Health board	245
Priority * Nature of incident	104
Control * Priority * Nature of incident	306
Control * Health board * Priority * Nature of incident (Bottom level)	691
Total	1530

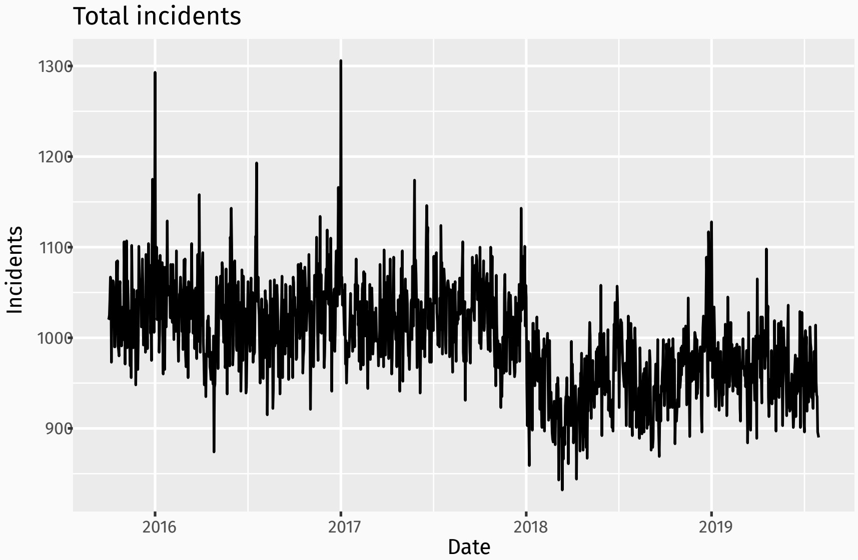
Data

```
# A tsibble: 2,142,000 x 6 [1D]
# Key:       region, category, nature, lhb [1,530]
   date       region       category     nature       lhb          incident
   <date>     <chr*>       <chr*>       <chr*>       <chr*>          <dbl>
 1 2015-10-01 <aggregated> <aggregated> <aggregated> <aggregated>     1020
 2 2015-10-02 <aggregated> <aggregated> <aggregated> <aggregated>     1021
 3 2015-10-03 <aggregated> <aggregated> <aggregated> <aggregated>     1025
 4 2015-10-04 <aggregated> <aggregated> <aggregated> <aggregated>     1043
 5 2015-10-05 <aggregated> <aggregated> <aggregated> <aggregated>     1067
 6 2015-10-06 <aggregated> <aggregated> <aggregated> <aggregated>     1063
 7 2015-10-07 <aggregated> <aggregated> <aggregated> <aggregated>      973
 8 2015-10-08 <aggregated> <aggregated> <aggregated> <aggregated>     1057
 9 2015-10-09 <aggregated> <aggregated> <aggregated> <aggregated>     1026
10 2015-10-10 <aggregated> <aggregated> <aggregated> <aggregated>     1063
# i 2,141,990 more rows
```

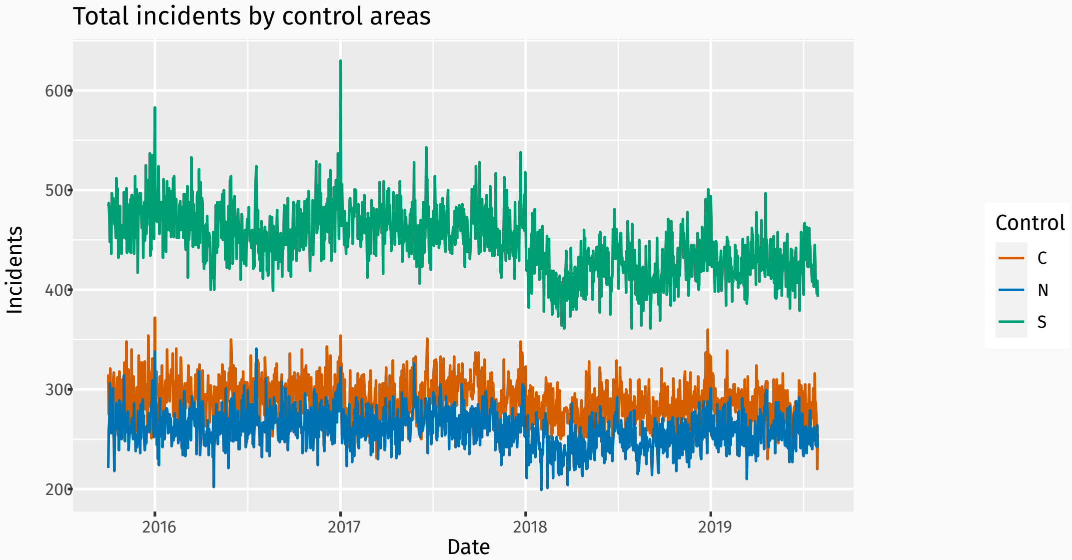
Data

```
# A tsibble: 2,142,000 x 6 [1D]
# Key:       region, category, nature, lhb [1,530]
   date       region category nature    lhb          incident
   <date>     <chr*> <chr*>   <chr*>    <chr*>          <dbl>
 1 2015-10-01 C      Amber    ABDOMINAL HD                  0
 2 2015-10-01 C      Amber    ABDOMINAL PO                  0
 3 2015-10-01 C      Amber    ABDOMINAL SB                  0
 4 2015-10-01 C      Amber    ABDOMINAL <aggregated>        0
 5 2015-10-01 C      Amber    ALLERGIES HD                  0
 6 2015-10-01 C      Amber    ALLERGIES PO                  1
 7 2015-10-01 C      Amber    ALLERGIES SB                  0
 8 2015-10-01 C      Amber    ALLERGIES <aggregated>        1
 9 2015-10-01 C      Amber    ANIMALBIT HD                  0
10 2015-10-01 C      Amber    ANIMALBIT PO                  0
# i 2,141,990 more rows
```

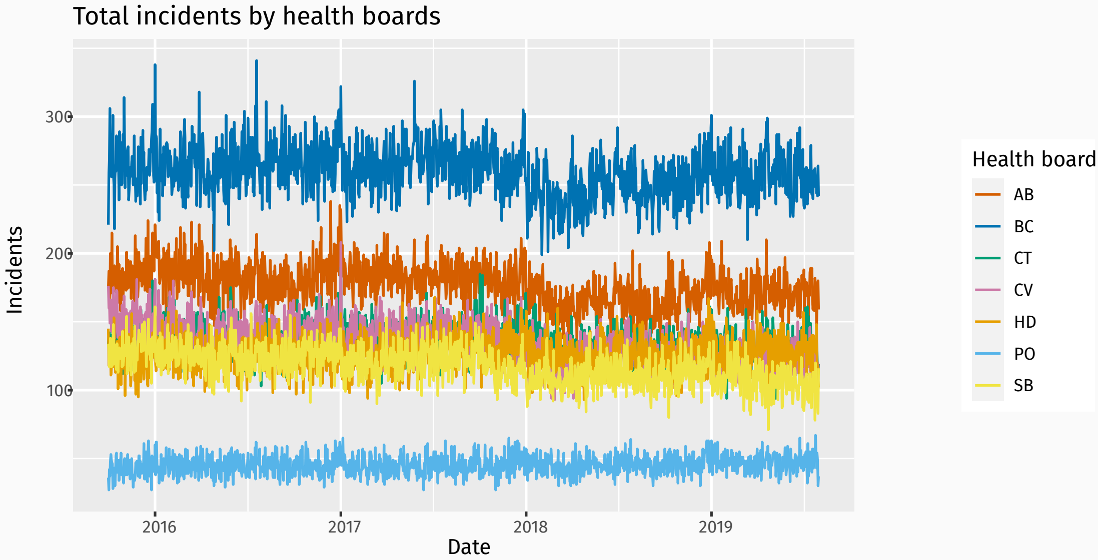
Aggregated daily incidents



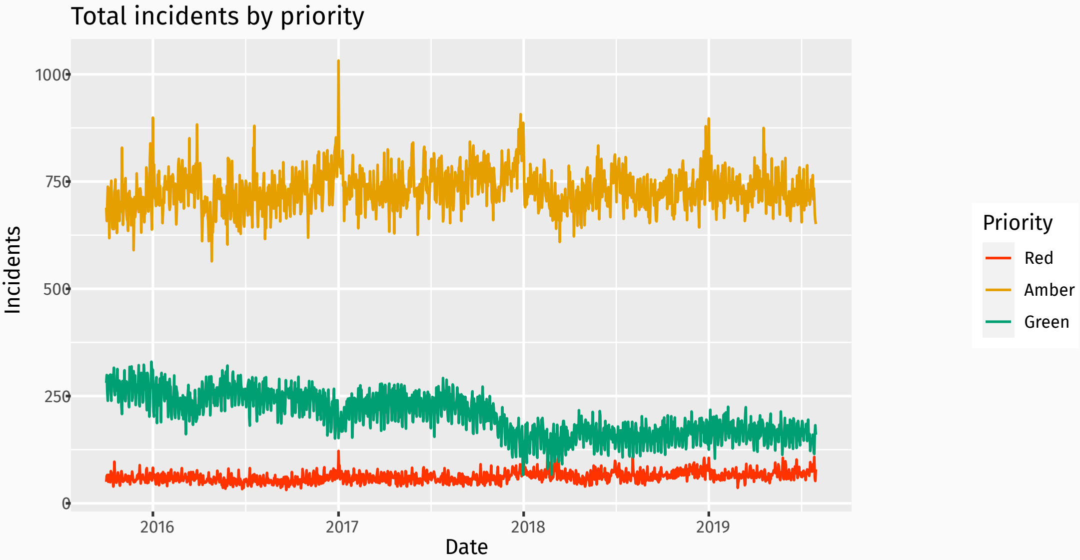
Aggregated daily incidents



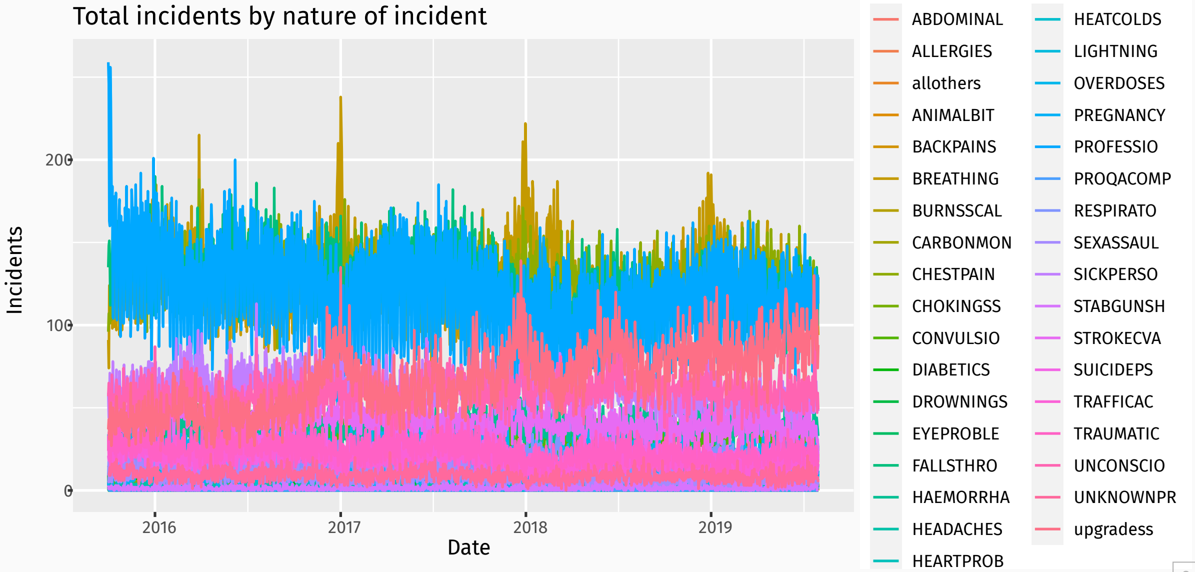
Aggregated daily incidents



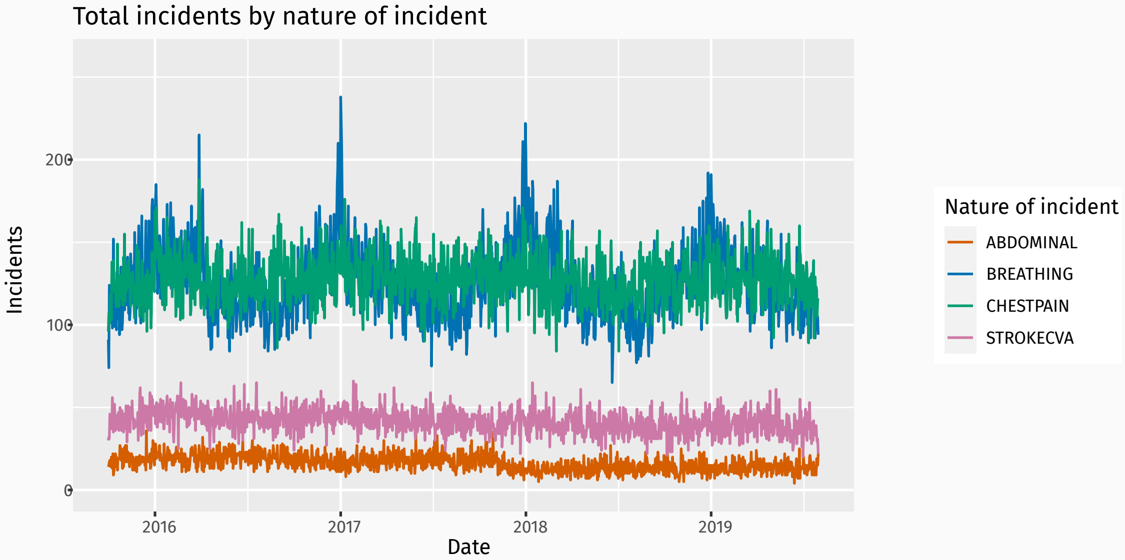
Aggregated daily incidents



Aggregated daily incidents



Aggregated daily incidents



Forecasting methods

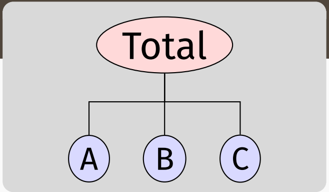
1. **Naïve**: Empirical distribution of past daily attended incidents.
2. **ETS**: Exponential Smoothing State Space models.
3. **GLM**: Poission Regression with spline trend, day of the week, annual Fourier seasonality, public holidays, school holidays, Christmas Day, New Year's Day.
4. **TSGLM**: Poisson Regression with same covariates plus three autoregressive terms.
5. **Ensemble**: Mixture distribution of 1–4.

Notation

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

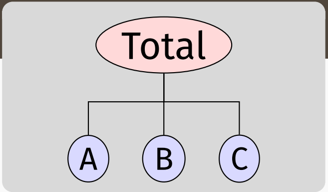
$$\boldsymbol{y}_t = \boldsymbol{S}\boldsymbol{b}_t$$

- $\boldsymbol{y}_t$ = vector of all series at time t
- $y_{\text{Total},t}$ = aggregate of all series at time t.
- $y_{X,t}$ = value of series X at time t.
- $\boldsymbol{b}_t$ = vector of most disaggregated series at time t
- $\boldsymbol{S}$ = "summing matrix" containing the linear constraints.



$$\boldsymbol{y}_t = \begin{pmatrix} y_{\text{Total},t} \\ y_{A,t} \\ y_{B,t} \\ y_{C,t} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\boldsymbol{S}} \underbrace{\begin{pmatrix} y_{A,t} \\ y_{B,t} \\ y_{C,t} \end{pmatrix}}_{\boldsymbol{b}_t}$$

Notation



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

$$\boldsymbol{y}_t = \boldsymbol{S}\boldsymbol{b}_t$$

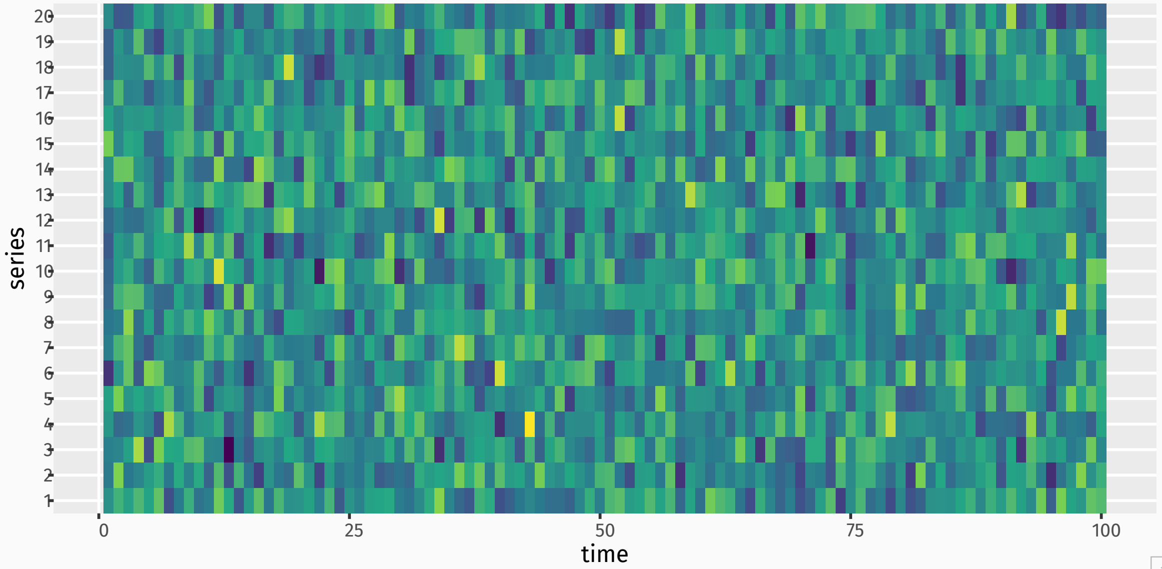
- $\boldsymbol{y}_t$ = vector of all series at time t
- $y_{\text{Total},t}$ = aggregate of all series at time t.
- $y_{X,t}$ = value of series X at time t.
- $\boldsymbol{b}_t$ = vector of most disaggregated series at time t
- $\boldsymbol{S}$ = "summing matrix" containing the linear constraints.

- Base forecasts: $\hat{\boldsymbol{y}}_{T+h|T}$
- Reconciled forecasts: $\tilde{\boldsymbol{y}}_{T+h|T} = \boldsymbol{S}\boldsymbol{G}\hat{\boldsymbol{y}}_{T+h|T}$
- MinT: $\boldsymbol{G} = (\boldsymbol{S}'\boldsymbol{W}_h^{-1}\boldsymbol{S})^{-1}\boldsymbol{S}'\boldsymbol{W}_h^{-1}$ where $\boldsymbol{W}_h$ is covariance matrix of base forecast errors.

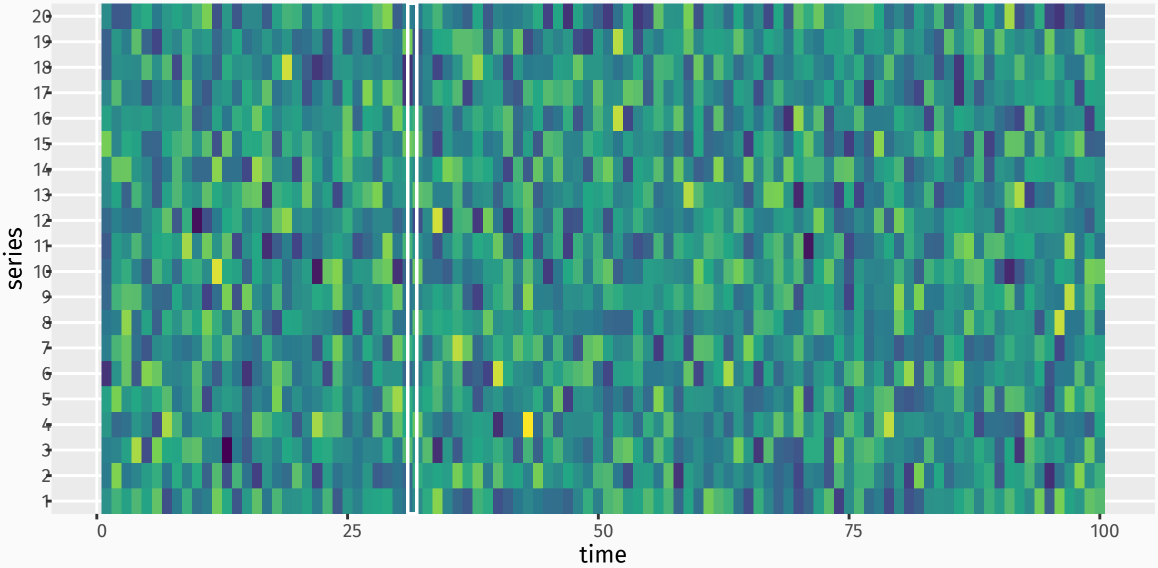
Nonparametric bootstrap reconciliation

- Fit model to all series and store the residuals as $\underset{\sim}{\varepsilon}_t$.
- These should be serially uncorrelated but cross-sectionally correlated.
- Draw iid samples from $\underset{\sim}{\varepsilon}_1, \dots, \underset{\sim}{\varepsilon}_T$ with replacement.
- Simulate future sample paths for model using the bootstrapped residuals.
- Reconcile each sample path using MinT.
- Combine the reconciled sample paths to form a mixture distribution at each forecast horizon.

Boostrapping residuals

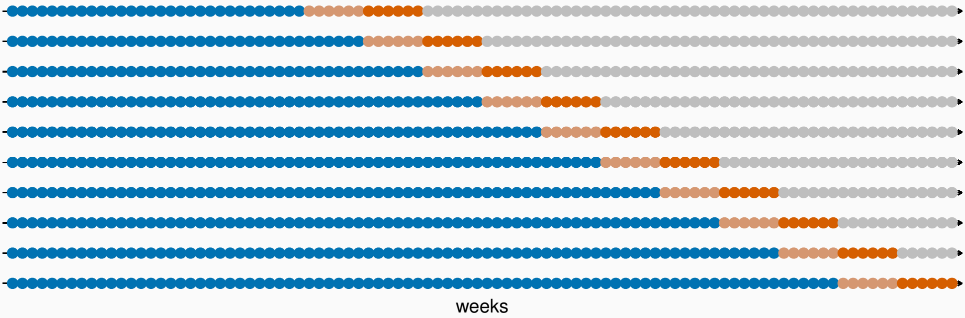


Boostrapping residuals



Performance evaluation

- Ten-fold time series cross-validation
- Forecast horizon of 1–84 days
- Each training set contains an additional 42 days.
- Forecasts at 43–84 days correspond to planning horizon.



Forecast accuracy: 43–84 days ahead

		MSSE		
Method	Model	Total	Control areas	Health boards
Base	Naïve	1.169	1.056	1.062
Base	ETS	0.979	0.875	0.816
Base	GLM	0.813	0.897	0.875
Base	TSGLM	0.822	0.901	0.875
Base	Ensemble	0.599	0.729	0.774
MinT	Naïve	1.168	1.057	1.062
MinT	ETS	0.785	0.852	0.845
MinT	GLM	0.720	0.827	0.837
MinT	TSGLM	0.722	0.833	0.839
MinT	Ensemble	**0.560**	**0.706**	**0.765**

$$\text{MSSE} = \text{mean}(q_j^2)$$

$$q_j^2 = \frac{e_j^2}{\frac{1}{T-7}\sum_{t=8}^{T}(y_t - y_{t-7})^2}$$

- Observations: $y_1, \dots, y_T$.
- Forecast errors: $e_j = y_{T+j} - \hat{y}_{T+j|T}$.

Forecast accuracy: 43–84 days ahead

		CRPS		
Method	Model	Total	Control areas	Health boards
Base	Naïve	30.387	10.882	5.500
Base	ETS	14.309	6.074	3.476
Base	GLM	15.396	6.253	3.576
Base	TSGLM	15.316	6.227	3.575
Base	Ensemble	12.978	**5.727**	3.430
MinT	Naïve	30.368	10.902	5.498
MinT	ETS	13.515	5.967	3.547
MinT	GLM	13.839	5.917	3.453
MinT	TSGLM	14.000	5.947	3.455
MinT	Ensemble	**12.585**	5.728	**3.426**

$$\text{CRPS} = \text{mean}(p_j)$$

$$p_j = \int_{-\infty}^{\infty} \left(G_j(x) - F_j(x)\right)^2 dx,$$

- $G_j(x)$ = forecast distribution for forecast horizon j
- $F_j(x)$ = empirical distribution for same period

Conclusions

- Ensemble mixture distributions give better forecasts than any component methods.
- Forecast reconciliation improves forecast accuracy, even when some component methods are quite poor.
- The ensemble without the Naïve method was worse.
- Forecast reconciliation allows coordinated planning and resource allocation.

Conclusions

- Ensemble mixture distributions give better forecasts than any component methods.
- Forecast reconciliation improves forecast accuracy, even when some component methods are quite poor.
- The ensemble without the Naïve method was worse.
- Forecast reconciliation allows coordinated planning and resource allocation.

robjhyndman.com/asc2023
aus.social/@robjhyndman
robjhyndman.com/fems
rob.hyndman@monash.edu