

# Probabilistic Forecast Reconciliation For Emergency Services Demand

Rob J Hyndman  
Bahman Rostami-Tabar



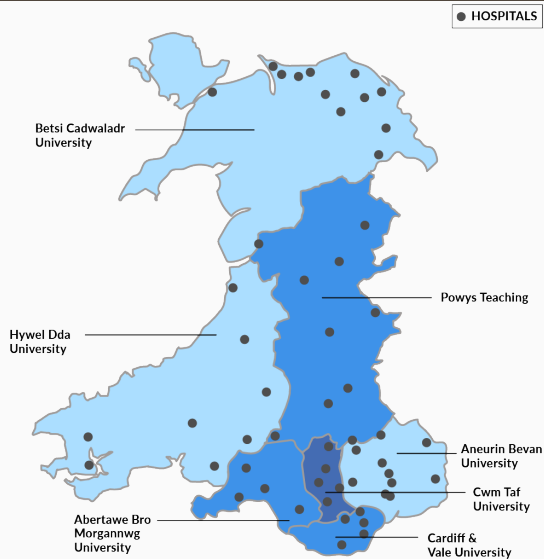
MONASH University

[robjhyndman.com/fem2023](http://robjhyndman.com/fem2023)



Photo by Graham Richardson on Wikimedia

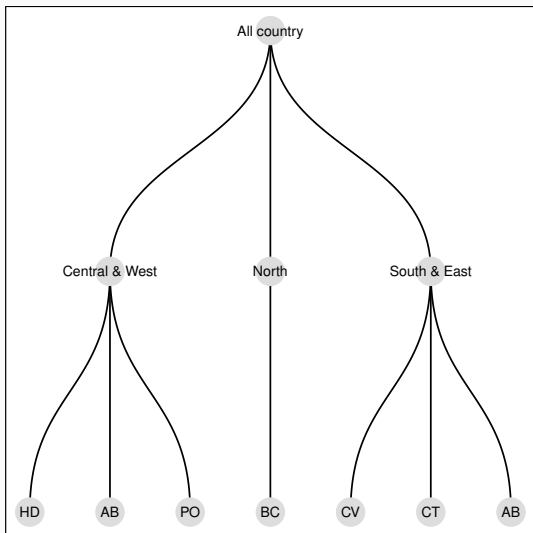
# Wales Health Board Areas



# 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

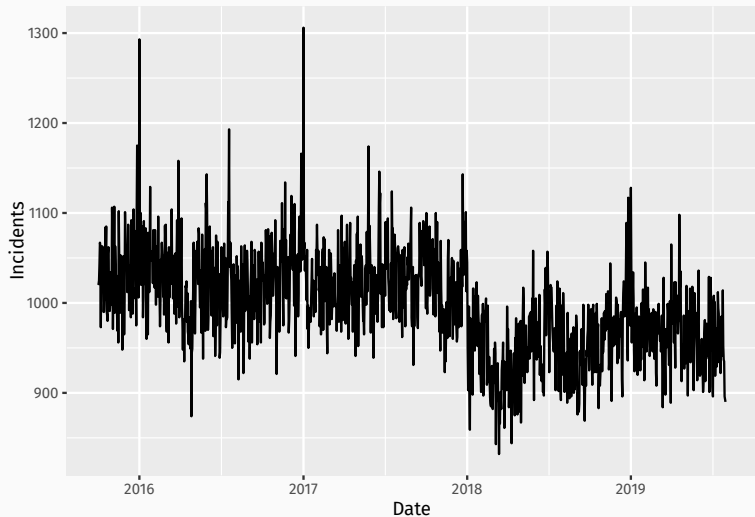
# Data

```
# 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





# Daily incidents by control area



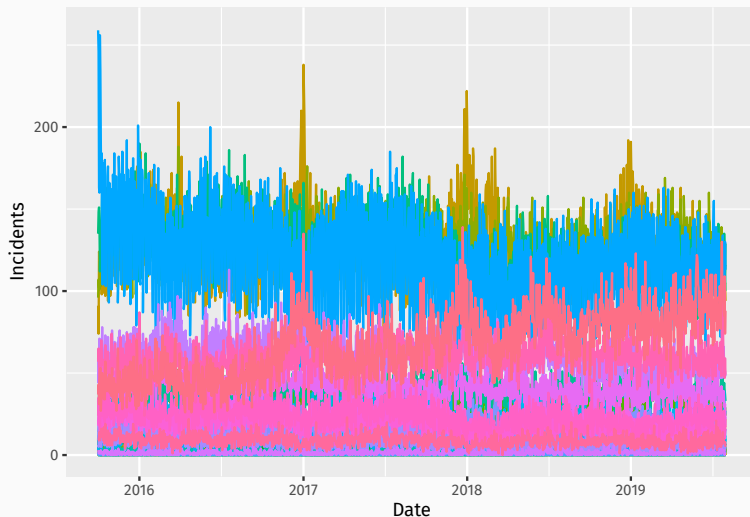
# Data incidents by health board



# Data incidents by priority



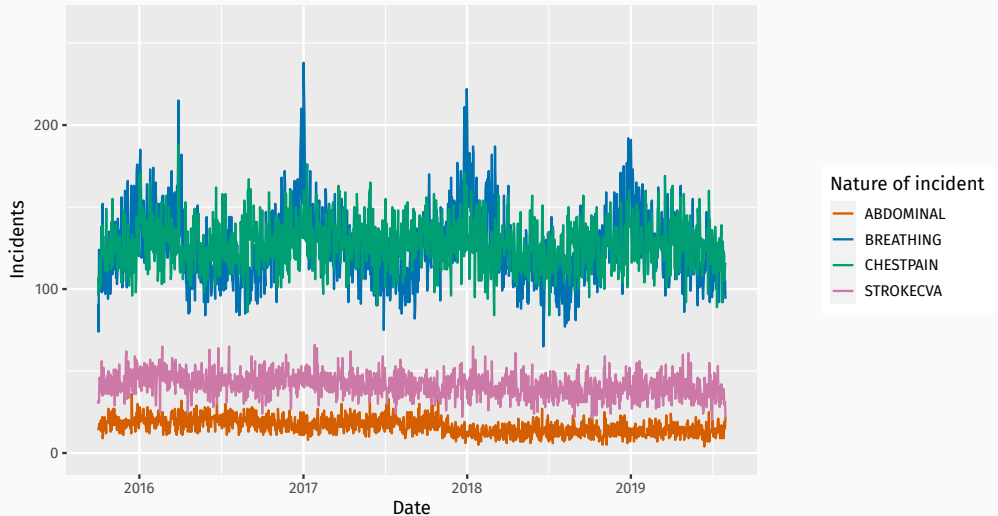
# Data incidents by nature of incident



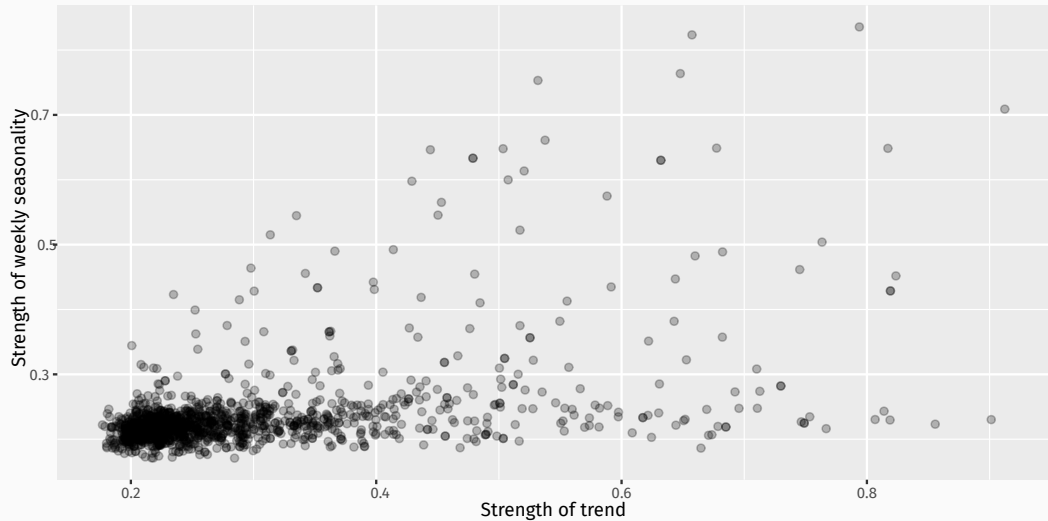
## Nature of incident

ABDOMINAL	HEATCOLDS
ALLERGIES	LIGHTNING
allothers	OVERDOSES
ANIMALBIT	PREGNANCY
BACKPAINS	PROFESSIO
BREATHING	PROQACOMP
BURNSSCAL	RESPIRATO
CARBONMON	SEXASSAUL
CHESTPAIN	SICKPERSO
CHOKINGSS	STABGUNSH
CONVULSIO	STROKECVA
DIABETICS	SUICIDEPS
DROWNINGS	TRAFFICAC
EYEPROBLE	TRAUMATIC
FALLSTHRO	UNCONSCIO
HAEMORRHA	UNKNOWNPR
HEADACHES	upgradess
HEARTPROB	

# Data incidents by nature of incident



# Data features



# 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.

# Forecasting methods

- 1 **Naïve:** Empirical distribution of past daily attended incidents.

$$y_{T+h|T} \sim \text{Empirical}(y_1, \dots, y_T)$$



# Forecasting methods

- 1 **Naïve:** Empirical distribution of past daily attended incidents.

$$y_{T+h|T} \sim \text{Empirical}(y_1, \dots, y_T)$$

- 2 **ETS:** Exponential Smoothing State Space models.

$$y_{T+h|T} \sim \text{Normal}(\hat{y}_{T+h|T}, \hat{\sigma}_{T+h|T}^2)$$

# Forecasting methods

3

## GLM: Poission Regression

$$y_{T+h|T} \sim \text{Poisson}(\hat{y}_{T+h|T}) \quad \text{where} \quad \hat{y}_{T+h|T} = \exp(\mathbf{x}'_{T+h}\beta)$$

and  $\mathbf{x}_{T+h}$  is a vector of covariates including

- spline trend
- day of the week
- annual Fourier seasonality
- public holidays
- school holidays
- Christmas Day
- New Year's Day

# Forecasting methods

	Estimate	Std. Error	z value	Pr(> z )	
(Intercept)	6.998511	0.017412	401.93	< 2e-16	***
Spline_1	0.027859	0.004740	5.88	4.2e-09	***
Spline_2	-0.088244	0.006394	-13.80	< 2e-16	***
Spline_3	-0.075036	0.004784	-15.68	< 2e-16	***
Spline_4	-0.111854	0.010202	-10.96	< 2e-16	***
Spline_5	-0.043009	0.004462	-9.64	< 2e-16	***
Monday	0.019147	0.003174	6.03	1.6e-09	***
Tuesday	-0.016414	0.003180	-5.16	2.4e-07	***
Wednesday	-0.015479	0.003184	-4.86	1.2e-06	***
Thursday	-0.006804	0.003178	-2.14	0.03230	*
Friday	0.012235	0.003156	3.88	0.00011	***
Saturday	0.005293	0.003165	1.67	0.09438	.
Fourier_S1_365	0.005365	0.001294	4.15	3.4e-05	***
Fourier_C1_365	0.008263	0.001263	6.54	6.1e-11	***
Fourier_S2_365	0.004235	0.001271	3.33	0.00086	***
Fourier_C2_365	-0.010510	0.001216	-8.64	< 2e-16	***
Fourier_S3_365	-0.000556	0.001275	-0.44	0.66303	
Fourier_C3_365	0.002650	0.001243	2.13	0.03294	*
Public_holiday	0.033278	0.005697	5.84	5.2e-09	***
School_holiday	0.004857	0.002346	2.07	0.03843	*
Xmas	-0.051902	0.016772	-3.09	0.00197	**
New_years_day	0.120385	0.015573	7.73	1.1e-14	***

## Significance

***	$p < 0.001$
**	$p < 0.01$
*	$p < 0.05$
.	$p < 0.1$

# Forecasting methods

4

## **TSGLM:** Poisson Regression

$$y_{T+h|T} \sim \text{Poisson}(\hat{y}_{T+h|T})$$

where 
$$\hat{y}_{T+h|T} = \exp \left( \mathbf{x}'_{T+h} \beta + \sum_{k=1}^3 \alpha_k \log(y_{T+h-k} + 1) \right)$$

and  $\mathbf{x}_{T+h}$  is a vector of covariates including

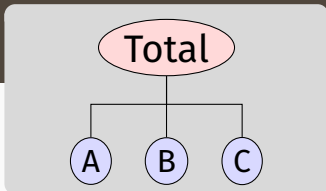
- spline trend
- day of the week
- annual Fourier seasonality
- public holidays
- school holidays
- Christmas Day
- New Year's Day

# Notation

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

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

- $\mathbf{y}_t$  = 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  $t$ .
- $\mathbf{b}_t$  = vector of most disaggregated series at time  $t$
- $\mathbf{S}$  = “summing matrix” containing the linear constraints.



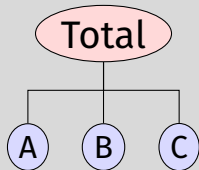
$$\mathbf{y}_t = \begin{pmatrix} y_{\text{Total},t} \\ y_{A,t} \\ y_{B,t} \\ y_{C,t} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\mathbf{S}} \underbrace{\begin{pmatrix} y_{A,t} \\ y_{B,t} \\ y_{C,t} \end{pmatrix}}_{\mathbf{b}_t}$$

# Notation

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

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

- $\mathbf{y}_t$  = 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  $t$ .
- $\mathbf{b}_t$  = vector of most disaggregated series at time  $t$
- $\mathbf{S}$  = “summing matrix” containing the linear constraints.



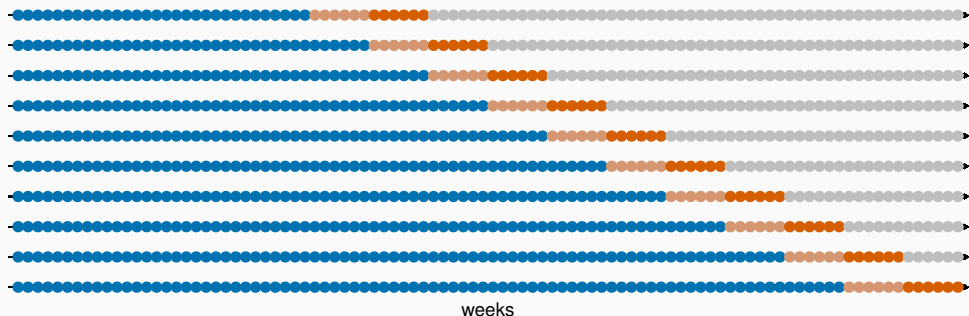
- Base forecasts:  $\hat{\mathbf{y}}_{T+h|T}$
- Reconciled forecasts:  $\tilde{\mathbf{y}}_{T+h|T} = \mathbf{S}\mathbf{G}\hat{\mathbf{y}}_{T+h|T}$
- MinT:  
 $\mathbf{G} = (\mathbf{S}'\mathbf{W}_h^{-1}\mathbf{S})^{-1}\mathbf{S}'\mathbf{W}_h^{-1}$   
where  $\mathbf{W}_h$  is covariance matrix of base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

# Nonparametric bootstrap reconciliation

- Fit model to all series and store the residuals as  $\varepsilon_t$ .
- These should be serially uncorrelated but cross-sectionally correlated.
- Draw iid samples from  $\varepsilon_1, \dots, \varepsilon_T$  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.

# 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.





# Performance evaluation

$$\text{MASE} = \text{mean}(|q_j|)$$

$$q_j = \frac{e_j}{\frac{1}{T-m} \sum_{t=m+1}^T |y_t - y_{t-m}|}$$

- $y_t$  = observation for period  $t$
- $e_j$  = forecast error for forecast horizon  $j$
- $T$  = size of training set
- $m = 7$

# Performance evaluation

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

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

- $y_t$  = observation for period  $t$
- $e_j$  = forecast error for forecast horizon  $j$
- $T$  = size of training set
- $m = 7$

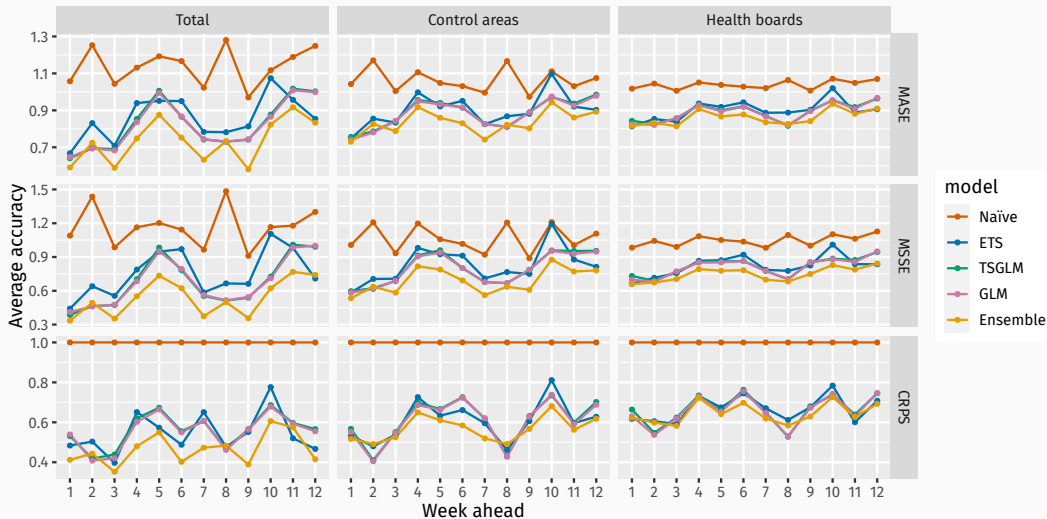
# Performance evaluation

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

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

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

# Forecast accuracy



# Forecast accuracy: 43–84 days ahead

Method	Model	MSSE			
		Total	Control areas	Health boards	Bottom
Base	Naïve	1.169	1.056	1.062	1.031
Base	ETS	0.979	0.875	0.816	<b>0.975</b>
Base	GLM	0.813	0.897	0.875	1.009
Base	TSGLM	0.822	0.901	0.875	1.050
Base	Ensemble	0.599	0.729	0.774	0.993
MinT	Naïve	1.168	1.057	1.062	2.095
MinT	ETS	0.785	0.852	0.845	0.994
MinT	GLM	0.720	0.827	0.837	1.803
MinT	TSGLM	0.722	0.833	0.839	1.851
MinT	Ensemble	<b>0.560</b>	<b>0.706</b>	<b>0.765</b>	1.557

# Forecast accuracy: 43–84 days ahead

Method	Model	MASE			
		Total	Control areas	Health boards	Bottom
Base	Naïve	1.139	1.059	1.047	1.019
Base	ETS	0.963	0.930	0.899	1.038
Base	GLM	0.910	0.940	0.923	<b>1.002</b>
Base	TSGLM	0.911	0.939	0.924	1.005
Base	Ensemble	0.782	0.856	0.876	1.008
MinT	Naïve	1.138	1.059	1.047	2.651
MinT	ETS	0.877	0.916	0.915	1.289
MinT	GLM	0.848	0.901	0.902	2.493
MinT	TSGLM	0.852	0.903	0.903	2.513
MinT	Ensemble	<b>0.753</b>	<b>0.844</b>	<b>0.872</b>	2.260

# Forecast accuracy: 43–84 days ahead

Method	Model	CRPS			
		Total	Control areas	Health boards	Bottom
Base	Naïve	30.387	10.882	5.500	0.302
Base	ETS	14.309	6.074	3.476	0.244
Base	GLM	15.396	6.253	3.576	0.244
Base	TSGLM	15.316	6.227	3.575	0.245
Base	Ensemble	12.978	<b>5.727</b>	3.430	0.243
MinT	Naïve	30.368	10.902	5.498	0.313
MinT	ETS	13.515	5.967	3.547	<b>0.243</b>
MinT	GLM	13.839	5.917	3.453	0.246
MinT	TSGLM	14.000	5.947	3.455	0.248
MinT	Ensemble	<b>12.585</b>	5.728	<b>3.426</b>	0.247

# 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.