Forecast reconciliation

4. Probabilistic forecast reconciliation

Rob J Hyndman

MONASH University

Photo by Edvard Alexander Rølvaag on

robjhyndman.com/fr2023

Outline

Notation reminder

- Data: $\boldsymbol{y}_t = \boldsymbol{S}\boldsymbol{b}_t$ where $\boldsymbol{S}$ is a summing matrix and $\boldsymbol{b}_t$ is a vector of disaggregated time series
- 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.

Probabilistic forecasts

- Gaussian
- Non-parametric
- Count

Probabilistic forecast reconciliation

Key papers

- Ben Taieb, Taylor, Hyndman (*ICML*, 2017)
- Jeon, Panagiotelis, Petropoulos (*EJOR*, 2019)
- Ben Taieb, Taylor, Hyndman (*JASA*, 2020)
- Panagiotelis, Gamakumara, Athanasopoulos, Hyndman (2020). robjhyndman.com/publications/coherentprob/

Probabilistic forecast reconciliation

Key papers

- Ben Taieb, Taylor, Hyndman (*ICML*, 2017)
- Jeon, Panagiotelis, Petropoulos (*EJOR*, 2019)
- Ben Taieb, Taylor, Hyndman (*JASA*, 2020)
- Panagiotelis, Gamakumara, Athanasopoulos, Hyndman (2020). robjhyndman.com/publications/coherentprob/

- The reconciled multivariate density must lie on the coherent subspace.
- The univariate density at each node is a convolution of the densities of its children.

Construction of reconciled distributions

Reconciled density of bottom-level

Density of bottom-level series under reconciled distribution is

$$\tilde{f}_{\boldsymbol{b}}(\boldsymbol{b}) = |\boldsymbol{G}^*| \int \hat{f}(\boldsymbol{G}^- \boldsymbol{b} + \boldsymbol{G}_\perp \boldsymbol{a}) d\boldsymbol{a}$$

- $\hat{f}$ is density of incoherent base probabilistic forecast
- $\boldsymbol{G}^-$ is $n \times m$ generalised inverse of $\boldsymbol{G}$ st $\boldsymbol{G}\boldsymbol{G}^- = \boldsymbol{I}$
- $\boldsymbol{G}_\perp$ is $n \times (n-m)$ orthogonal complement to $\boldsymbol{G}$ st $\boldsymbol{G}\boldsymbol{G}_\perp = \boldsymbol{0}$
- $\boldsymbol{G}^* = \left(\boldsymbol{G}^- \vdots \boldsymbol{G}_\perp\right)$, and $\boldsymbol{b}$ and $\boldsymbol{a}$ are obtained via

  the change of variables $\boldsymbol{y} = \boldsymbol{G}^* \begin{pmatrix} \boldsymbol{b} \\ \boldsymbol{a} \end{pmatrix}$

Construction of reconciled distributions

Reconciled density of full hierarchy

Density of full hierarchy under reconciled distribution is

$$\tilde{f}_{\boldsymbol{y}}(\boldsymbol{y}) = |\boldsymbol{S}^*|\,\tilde{f}_{\boldsymbol{b}}(\boldsymbol{S}^-\boldsymbol{y})\,\mathbb{1}\{\boldsymbol{y} \in \mathfrak{s}\}$$

- $\boldsymbol{S}^* = \begin{pmatrix}\boldsymbol{S}^{-\prime} & \boldsymbol{S}_\perp\end{pmatrix}'$
- $\boldsymbol{S}^-$ is $m \times n$ generalised inverse of $\boldsymbol{S}$ such that $\boldsymbol{S}^-\boldsymbol{S} = \boldsymbol{I}$,
- $\boldsymbol{S}_\perp$ is $n \times (n-m)$ orthogonal complement to $\boldsymbol{S}$ such that

Gaussian reconciliation

If the incoherent base forecasts are $\text{N}(\hat{\mu}, \hat{\Sigma})$, then the reconciled density is $\text{N}(\boldsymbol{SG}\hat{\mu}, \boldsymbol{SG}\hat{\Sigma}\boldsymbol{G}'\boldsymbol{S}')$.

Bootstrap reconciliation

Reconciling sample paths from incoherent distributions works.

Evaluating probabilistic forecasts

Proper scoring rule

optimized when true forecast distribution is used.

Evaluating probabilistic forecasts

Proper scoring rule

optimized when true forecast distribution is used.

Scoring Rule	Coherent v Incoherent	Coherent v Coherent
Log Score	Not proper	• Ordering preserved if compared using bottom-level only
Energy Score	Proper	• Full hierarchy should be used. • Rankings may

Score optimal reconciliation

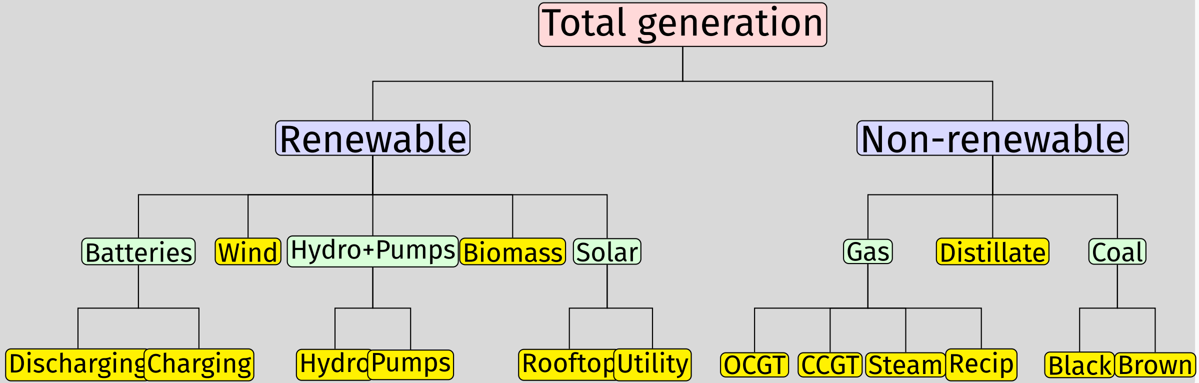
Algorithm proposed by Panagiotelis et al (2020) for optimizing $\boldsymbol{G}$ using stochastic gradient descent to optimize Energy Score.

1. Compute base forecasts over a test set.
2. Compute OLS reconciliation: $\boldsymbol{G} = (\boldsymbol{S}'\boldsymbol{S})^{-1}\boldsymbol{S}'$
3. Iteratively update $\boldsymbol{G}$ using SGD with Adam method and ES objective over a test set

Outline

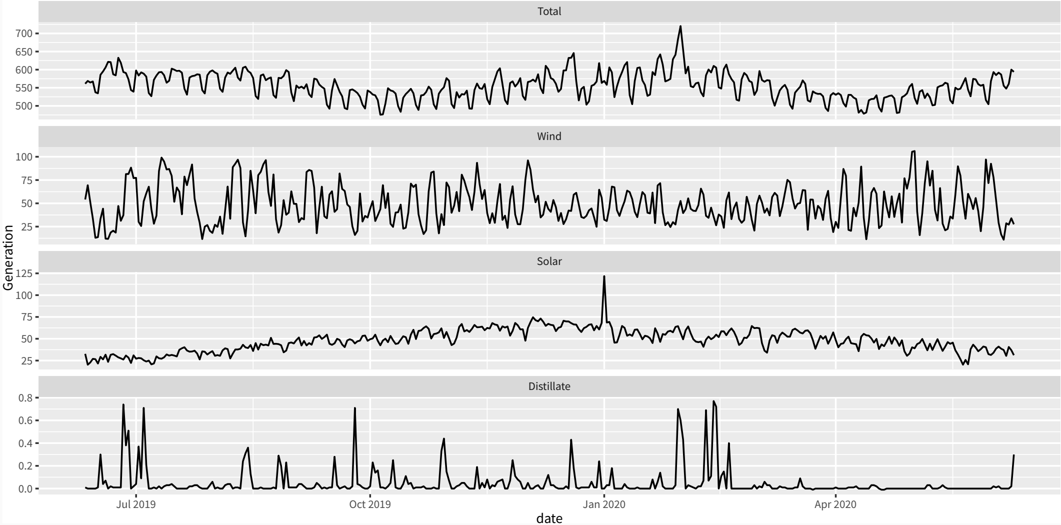
Example: Australian electricity generation

Daily time series from opennem.org.au



n = 23 series m = 15 bottom-level series

Example: Australian electricity generation

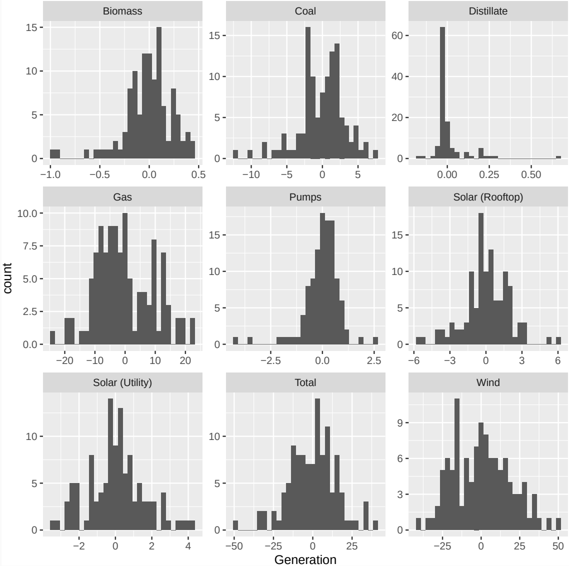


Example: Australian electricity generation

Forecast evaluation

- Rolling window of 140 days training data, and one-step-forecasts for 170 days test data.
- One-layer feed-forward neural network with up to 28 lags of target variable as inputs.
- Implemented using `NNETAR()` function in `fable` package.
- Model could be improved with temperature predictor.

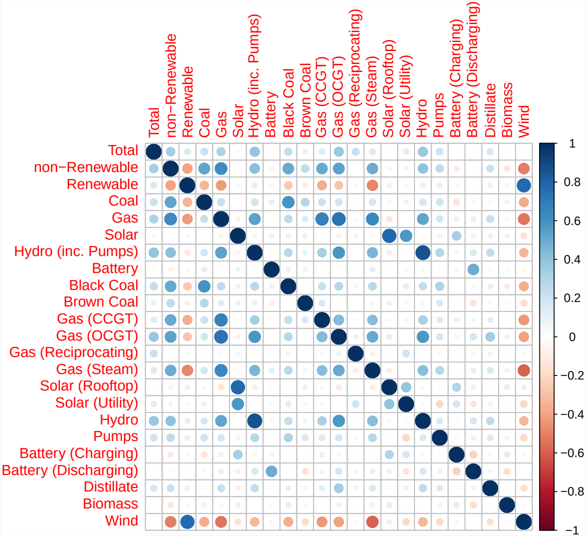
Example: Australian electricity generation



Histogram of residuals: 2 Oct 2019 – 21 Jan 2020

Clearly non-Gaussian

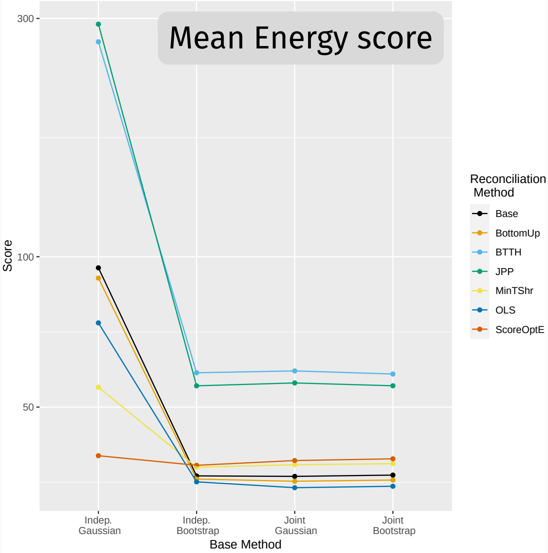
Example: Australian electricity generation



Correlations of residuals: 2 Oct 2019 – 21 Jan 2020

Blue = positive correlation.
Red = negative correlation.
Large = stronger correlations.

Example: Australian electricity generation



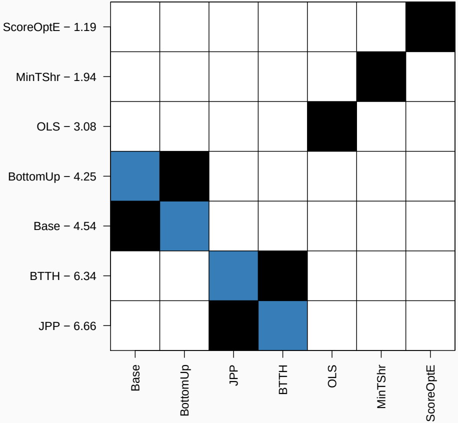
Base residual assumptions

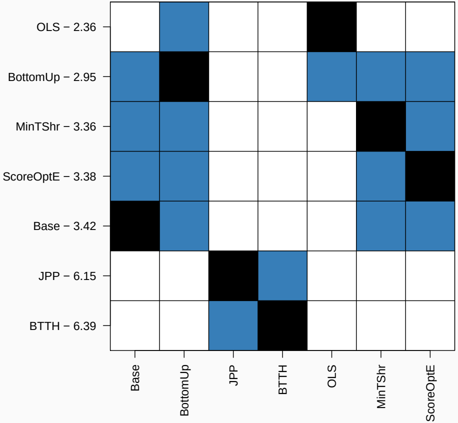
- Gaussian independent
- Gaussian dependent
- Non-Gaussian independent
- Non-Gaussian dependent

Reconciliation methods

- Base
- BottomUp
- BTTH: Ben Taieb, Taylor, Hyndman
- JPP: Jeon, Panagiotelis, Petropoulos
- OLS

Example: Australian electricity generation





Nemenyi test for different scores

Base forecasts are independent and

Nemenyi test for different scores

Base forecasts are obtained by jointly

Outline

Bayesian versions

Novak, McGarvie, and Garcia (2017)

Another strain of the literature brings a Bayesian approach to the regression model interpretation of forecast reconciliation. Novak, McGarvie, and Garcia (2017) recognise that the posterior of β_h can act as a probabilistic forecast for the bottom-level series. Using Markov chain Monte Carlo to obtain a sample from this posterior, and then aggregating, gives a probabilistic forecast for the entire hierarchy.

Bayesian versions

Eckert, Hyndman, and Panagiotelis (2021) also obtain a posterior on β_h, but their focus is on augmenting the reconciliation regression equation with a vector of intercepts that allow for base forecasts to be biased and evolve according to a state space representation.

Judgement can be incorporated via the prior, in the latter case via an explicit empirical example where prior information about a structural break in data classification can be exploited. Also, while both papers recognise the potential of Bayesian inference to obtain probabilistic forecasts, neither paper

Bayesian versions

Corani, Azzimonti, Augusto, and Zaffalon (2021) In particular, a prior is placed on the bottom-level series with the mean set to point forecasts obtained in the first step of forecast reconciliation and a variance given by the variance-covariance matrix of one-step ahead errors. This prior is updated using the top-level forecasts obtained in the first stage of forecast reconciliation via Bayes' rule. The method generalises MinT in the sense that the posterior mean is equivalent to the usual MinT approach. The necessary updates via Bayes' rule have parallels with the Kalman filter since the reconciliation problem is recast as a linear Gaussian model. The empirical

Thanks!

More information

- Slides and papers: **robjhyndman.com**
- Packages: **tidyverts.org**
- Forecasting textbook using fable package: **OTexts.com/fpp3**

Find me at ...

@robjhyndman

@robjhyndman

robjhyndman.com

References

- Ben Taieb, S, JW Taylor, and RJ Hyndman (2021). Hierarchical Probabilistic Forecasting of Electricity Demand with Smart Meter Data. *J American Statistical Association* **116**(533), 27–43.
- Corani, G, D Azzimonti, JPSC Augusto, and M Zaffalon (2021). Probabilistic reconciliation of hierarchical forecast via Bayes' rule. In: *Machine Learning and Knowledge Discovery in Databases*. Ed. by F Hutter, K Kersting, J Lijffijt, and I Valera. Cham: Springer International Publishing, pp.211–226.
- Corani, G, N Rubattu, D Azzimonti, and A Antonucci (2022). "Probabilistic reconciliation of count time series". https://arxiv.org/abs/2207.09322.
- Eckert, F, RJ Hyndman, and A Panagiotelis (2021). Forecasting Swiss exports using Bayesian forecast reconciliation. *European J Operational Research* **291**(2), 693–710.

References

- Girolimetto, D, G Athanasopoulos, T Di Fonzo, and RJ Hyndman (2023). "Cross-temporal probabilistic forecast reconciliation". robjhyndman.com/publications/ctprob.html.
- Novak, J, S McGarvie, and BE Garcia (2017). "A Bayesian model for forecasting hierarchically structured time series". https://arxiv.org/abs/1711.04738.
- Panagiotelis, A, P Gamakumara, G Athanasopoulos, and RJ Hyndman (2023). Probabilistic forecast reconciliation: properties, evaluation and score optimisation. *European J Operational Research* **306**(2), 693–706.