



Department of Econometrics and Business Statistics

<http://monash.edu/business/ebs/research/publications>

Optimal forecast reconciliation with time series selection

Xiaoqian Wang, Rob J Hyndman, Shanika
L Wickramasuriya

September 2023

Working Paper no/yr



AACSB
ACCREDITED



Optimal forecast reconciliation with time series selection

Xiaoqian Wang

Monash University, VIC 3800, Australia

Email: xiaoqian.wang@monash.edu

Corresponding author

Rob J Hyndman

Monash University, VIC 3800, Australia

Email: rob.hyndman@monash.edu

Shanika L Wickramasuriya

Monash University, VIC 3145, Australia

20 September 2023

Optimal forecast reconciliation with time series selection

Abstract

Abstract

Keywords: Keyword 1, Keyword 2

1 Introduction

Hierarchical time series and forecast reconciliation.

Single-level approaches, least squares-based reconciliation approaches, geometric intuition, other extensions with constraints.

However... Two issues.

In this paper, our focus will be on...

The remainder of the paper is structured as follows.

2 Preliminaries

2.1 Notation

We denote the set $\{1, \dots, k\}$ by $[k]$ for any non-negative integer k . A *hierarchical time series* can be considered as an n -dimensional multivariate time series, $\{\mathbf{y}_t, t \in [T]\}$, that adheres to known linear constraints. Let $\mathbf{y}_t \in \mathbb{R}^n$ be a vector comprising observations of all time series in the hierarchy at time t , and $\mathbf{b}_t \in \mathbb{R}^{n_b}$ be a vector comprising observations of all bottom-level time series at time t . The full hierarchy at time t can be written as

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

where S is an $n \times n_b$ *summing matrix* that shows aggregation constraints present in the structure. We can write the summing matrix as $S = \begin{bmatrix} A \\ I_{n_b} \end{bmatrix}$, where A is an $n_a \times n_b$ *aggregation matrix* with $n = n_a + n_b$, and I_{n_b} is an n_b -dimensional identity matrix.

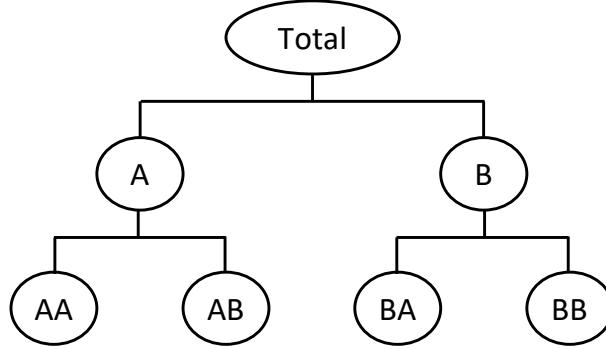


Figure 1: An example of a two-level hierarchical time series.

To clarify these notations, consider the example of the hierarchy in Figure 1. For this two-level hierarchy, $n = 7$, $n_b = 4$, $n_a = 3$, $\mathbf{y}_t = [y_{\text{Total},t}, y_{A,t}, y_{B,t}, y_{AA,t}, y_{AB,t}, y_{BA,t}, y_{BB,t}]'$, $\mathbf{b}_t = [y_{AA,t}, y_{AB,t}, y_{BA,t}, y_{BB,t}]'$, and

$$S = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ I_4 \end{bmatrix}.$$

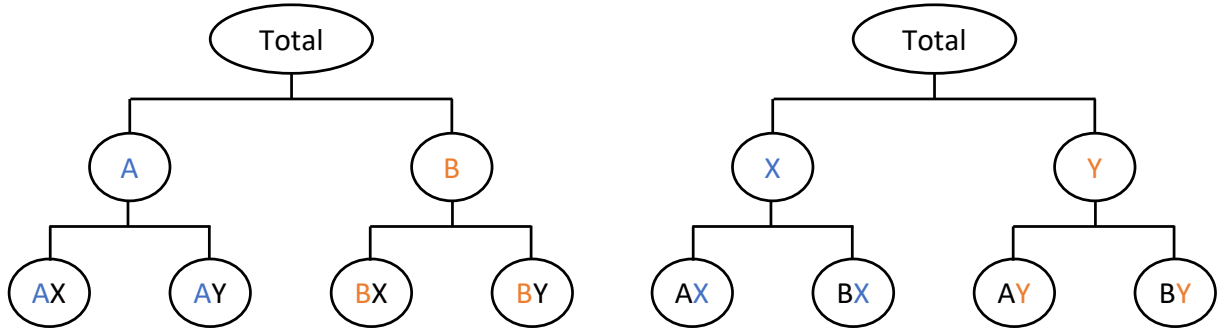


Figure 2: An example of a two level grouped time series.

When data structure does not naturally disaggregate in a unique hierarchical manner, we can combine these hierarchical structures to form a *grouped time series*. Thus, grouped time series can also be considered as hierarchical time series with more than one grouping structure. Figure 2 shows an example of a two level grouped time series with two alternative aggregation structures. For this example, $n = 9$, $n_b = 4$, $n_a = 5$, $\mathbf{y}_t = [y_{\text{Total},t}, y_{A,t}, y_{B,t}, y_{X,t}, y_{Y,t}, y_{AX,t}, y_{AY,t}, y_{BX,t}, y_{BY,t}]'$, $\mathbf{b}_t = [y_{AX,t}, y_{AY,t}, y_{BX,t}, y_{BY,t}]'$, and

$$S = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ & & & & I_4 \end{bmatrix}.$$

2.2 Linear forecast reconciliation

Let $\hat{\mathbf{y}}_{T+h|T} \in \mathbb{R}^n$ be a vector of h -step-ahead *base forecasts* for all time series in the hierarchy, given observations up to time T , and stacked in the same order as \mathbf{y}_t . We can use any method to generate these forecasts, but In general they will not add up especially when we forecast each series independently.

When forecasting hierarchical time series, we expect the forecasts to be *coherent* (i.e., aggregation constraints are satisfied). Let $\tilde{\mathbf{y}}_{T+h|T} \in \mathbb{R}^n$ denote a vector of h -step-ahead **reconciled forecasts** which are coherent by construction, ψ a mapping that reconciles base forecasts, $\hat{\mathbf{y}}_{T+h|T}$. Then we have *forecast reconciliation* $\tilde{\mathbf{y}}_{T+h|T} = \psi(\hat{\mathbf{y}}_{T+h|T})$, which is essentially a post-processing method. For the ease of exposition, we will use the subscript h rather than $T + h|T$. In this paper, we focus on linear forecast reconciliation given by

$$\tilde{\mathbf{y}}_h = S G_h \hat{\mathbf{y}}_h,$$

where

- G_h is an $n_b \times n$ weighting matrix that maps the base forecasts into the bottom level. In other words, it combines all base forecasts to form reconciled forecasts for bottom-level series.
- S is an $n \times n_b$ summing matrix that sums up bottom-level reconciled forecasts to produce coherent forecasts of all levels.

2.2.1 Minimum trace reconciliation

2.2.2 Reconciliation with regularization

3 Forecast reconciliation with time series selection

3.1 Constrained reconciliation with selection

Unbiasedness constraint

3.2 Unconstrained reconciliation with selection

4 Monte Carlo simulations

4.1 Model misspecification in a hierarchy

4.2 Exploring the effect of correlation

5 Applications

5.1 Forecasting Australian domestic tourism

5.2 Forecasting Australian prison population

6 Conclusion

Acknowledgement