

# Cross-temporal probabilistic forecast reconciliation

Daniele Girolimetto<sup>1</sup>

 [danigiro.github.io](https://danigiro.github.io)

@daniele.girolimetto@phd.unipd.it

George Athanasopoulos<sup>2</sup>, Tommaso Di Fonzo<sup>1</sup>, Rob J Hyndman<sup>2</sup>

<sup>1</sup> Department of Statistical Sciences, University of Padova

<sup>2</sup> Department of Econometrics and Business Statistics, Monash University

Prato – September 8<sup>th</sup> 2023

IIF Workshop on Forecast Reconciliation

# Outline

1. Introduction
2. Cross-sectional, temporal and cross-temporal framework
3. Point forecast reconciliation (cross-temporal framework)
4. Probabilistic forecast reconciliation (cross-temporal framework)
5. Forecasting the Australian Tourism Demand
6. Conclusions

# Introduction

## Forecast reconciliation

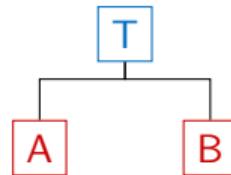
Post-forecasting process aimed to improve the quality of the base forecasts (however obtained) of a **linearly constrained** multiple time series by exploiting cross-sectional (e.g., spatial) and/or temporal constraints of the **target** forecasts

**cross-sectional** framework + **temporal** framework  $\Rightarrow$  **cross-temporal** framework

- Hot topic on forecasting methodology and practice:
  - ➔ several contributions starting from [Hyndman et al. \(2011\)](#)
- Looking for **statistically well-grounded**, **feasible** (practical implementation), **effective** (quality of the results) approaches
- Many forecasting applications: sales, production, tourism, energy demand, healthcare, real estate, supply chain, macroeconomics, ...
- **Point** and **probabilistic** forecast reconciliation

# Cross-sectional framework

Hyndman *et al.* (2011); Panagiotelis *et al.* (2021); Girolimetto and Di Fonzo (2023b)



2-level: bottom and upper ts

A cross-sectional hierarchical/grouped time series is a collection of  $n$  variables for which - at each time - **aggregation relationships** hold. It is a special case of **multiple time series** with exact **linear constraints**.

- Two equivalent representations (Girolimetto and Di Fonzo, 2023b)

## Zero-constrained

$$\mathcal{C}_{cs} \mathbf{y}_t = \mathbf{0}_{(n_a \times 1)}$$

$$\mathcal{C}_{cs} = [I \quad -A]$$

not unique

$$\begin{bmatrix} 1 & -1 & -1 \end{bmatrix}$$

## Structural

$$\mathbf{y}_t = S_{cs} \mathbf{b}_t$$

$$S_{cs} = \begin{bmatrix} A \\ I_{n_b} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 \\ I_2 \end{bmatrix}$$

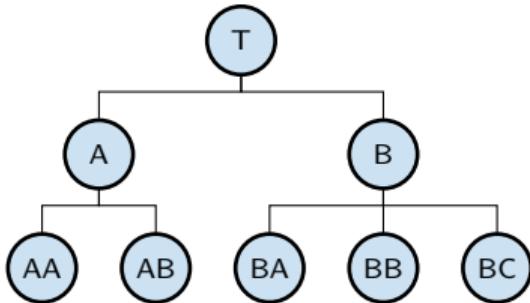
$$A : \mathbf{a}_t = A \mathbf{b}_t$$

Linear combination (or aggregation) matrix

$$\begin{bmatrix} 1 & 1 \end{bmatrix}$$

# Hierarchical, grouped and linearly constrained time series

*Genuine hierarchical time series*



*General linearly constrained time series*

**Constraints**

$$T = A + B$$

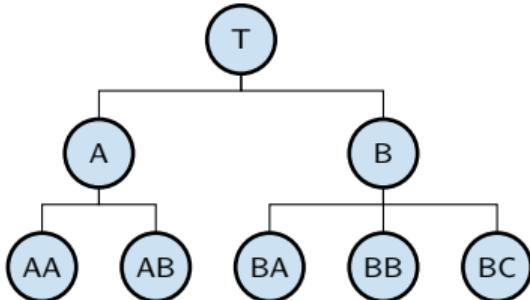
$$A = AA + AB$$

$$B = BA + BB + BC$$

*Grouped time series: two or more genuine hierarchies sharing the same top and bottom variables*

# Hierarchical, grouped and linearly constrained time series

Genuine hierarchical time series



Constraints

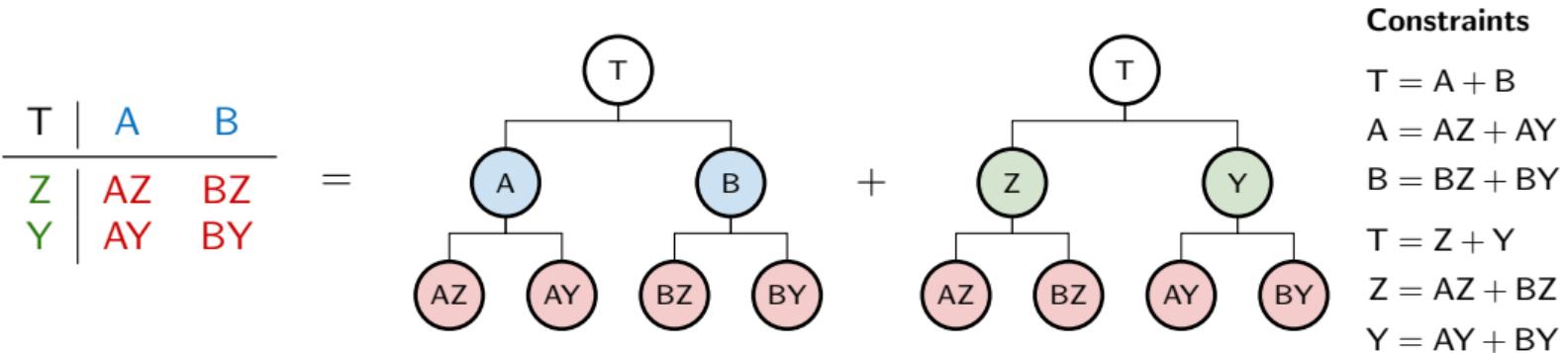
$$T = A + B$$

$$A = AA + AB$$

$$B = BA + BB + BC$$

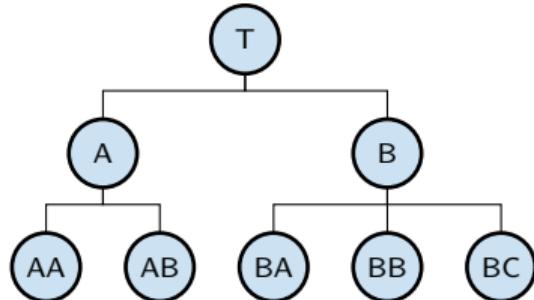
General linearly constrained time series

Grouped time series: two or more genuine hierarchies sharing the same top and bottom variables



# Hierarchical, grouped and linearly constrained time series

*Genuine hierarchical time series*



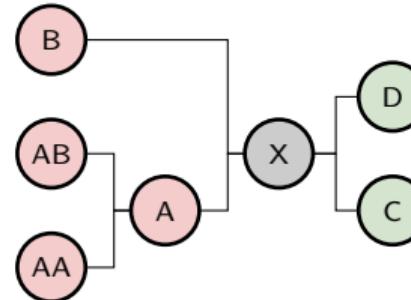
**Constraints**

$$T = A + B$$

$$A = AA + AB$$

$$B = BA + BB + BC$$

*General linearly constrained time series*



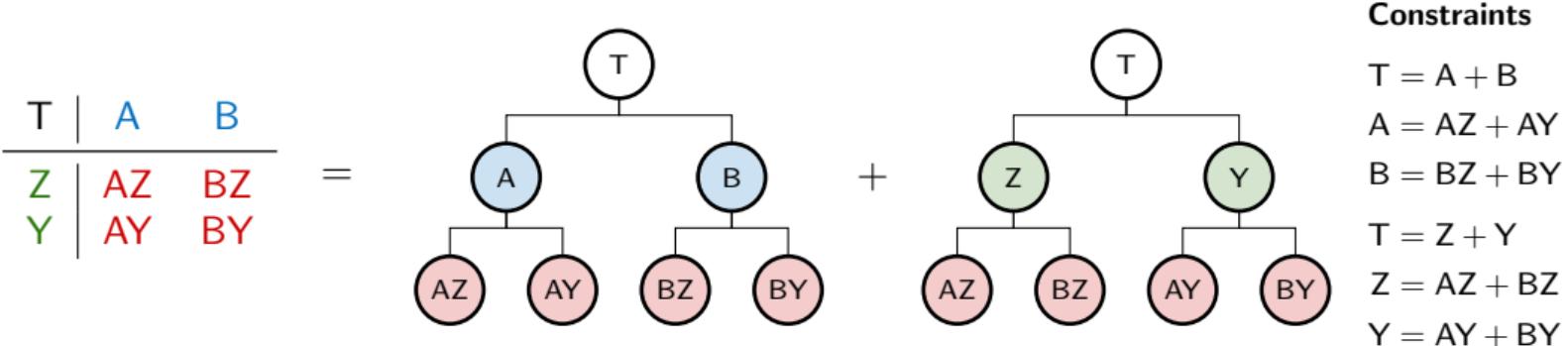
**Constraints**

$$X = A + B$$

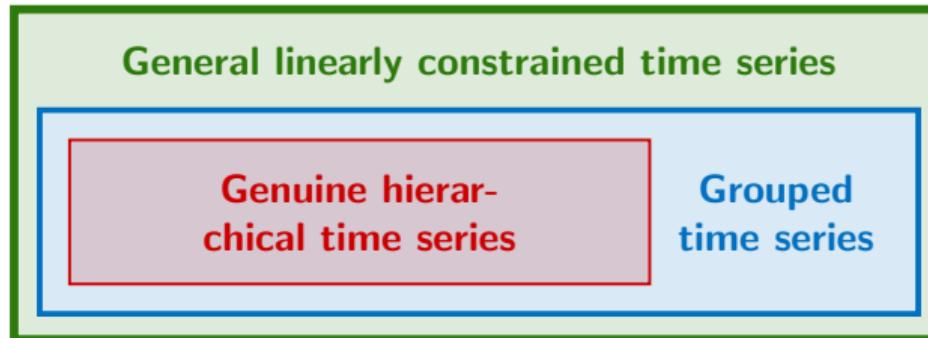
$$X = C + D$$

$$A = AA + AB$$

*Grouped time series: two or more genuine hierarchies sharing the same top and bottom variables*



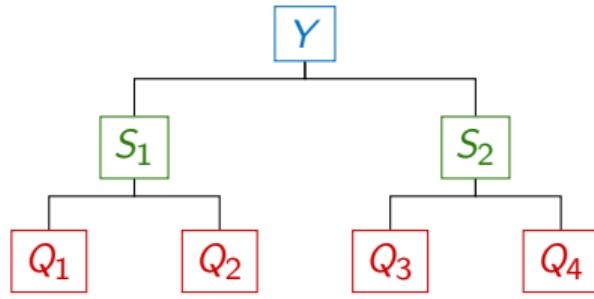
# General linearly constrained time series



- Forecast reconciliation may be **always** expressed according to a **zero-constrained** framework
- A **structural-like** reconciliation formula may be derived
  - ✗ ~~upper and bottom~~ time series
  - ✓ constrained and free (**unconstrained**) time series
- Girolimetto, D. and Di Fonzo, T. (2023) **Point and probabilistic forecast reconciliation for general linearly constrained multiple time series.** [arXiv:2305.05330](https://arxiv.org/abs/2305.05330)

# Temporal framework

Athanasiopoulos *et al.* (2017)



Quarterly hierarchy:  
quarterly, semi-annual and annual series

**Temporal hierarchy** → non-overlapping aggregation of the observations of a time series ( $y_i$ ) at regular intervals

$$x_{i,\tau}^{[k]} = \sum_{t=(\tau-1)k+1}^{\tau k} y_{i,t} \quad \text{for } \tau = 1, \dots, \lfloor T/k \rfloor$$

**NB:** For  $k = 1$ ,  $\tau = t = 1, \dots, T$  and  $x_{i,\tau}^{[1]} = y_{i,t}$

- $k \in \mathcal{K} = \{k_p, \dots, k_1\}$  denote the  $p$  factors of  $m$  in descending order, where  $k_1 = 1$  and  $k_p = m$
- $k^* = \sum_{k \in \mathcal{K} \setminus \{k_1\}} k$  is the number of upper time series of the temporal hierarchy
- Unlike cross-sectional hierarchies ( **$n$  variables at the same time index** are considered), in temporal hierarchies we have **one variable observed at different frequencies**

# Temporal matrices: quarterly data

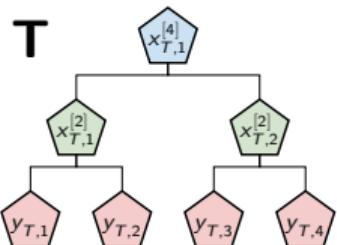
$$\mathbf{x}_{i,\tau} = \begin{bmatrix} \mathbf{x}_{i,\tau}^{[m]} \\ \mathbf{x}_{i,\tau}^{[k_{p-1}]} \\ \vdots \\ \mathbf{x}_{i,\tau}^{[k_2]} \\ \mathbf{x}_{i,\tau}^{[1]} = \mathbf{y}_{i,\tau} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{i,\tau}^{[4]} \\ \mathbf{x}_{i,2(\tau-1)+1}^{[2]} \\ \mathbf{x}_{i,2(\tau-1)+2}^{[2]} \\ \mathbf{y}_{i,4(\tau-1)+1} \\ \mathbf{y}_{i,4(\tau-1)+2} \\ \mathbf{y}_{i,4(\tau-1)+3} \\ \mathbf{y}_{i,4(\tau-1)+4} \end{bmatrix} \quad \mathbf{A}_{te} = \begin{bmatrix} \mathbf{I}_{\frac{m}{k_{p-1}}} \otimes \mathbf{1}'_{k_{p-1}} \\ \vdots \\ \mathbf{I}_{\frac{m}{k_2}} \otimes \mathbf{1}'_{k_2} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

- $\tau$  is the time index for the most aggregated series ( $\tau = 1, \dots, \lfloor T/m \rfloor$ )
- $i$  is the cross-sectional index (e.g.  $i = T, A, B$ )
- $M_k = m/k$  is the seasonal period of aggregated series
- Structural  $(\mathbf{x}_{i,\tau} = \mathbf{S}_{te}\mathbf{x}_{i,\tau}^{[1]})$  and zero-constrained  $(\mathbf{C}_{te}\mathbf{x}_{i,\tau} = \mathbf{0}_{(k^* \times 1)})$  representations still hold, and may be alternatively used for reconciliation

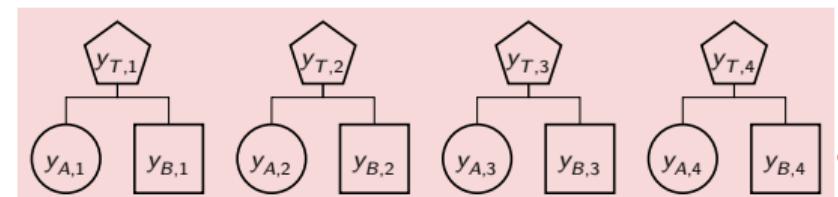
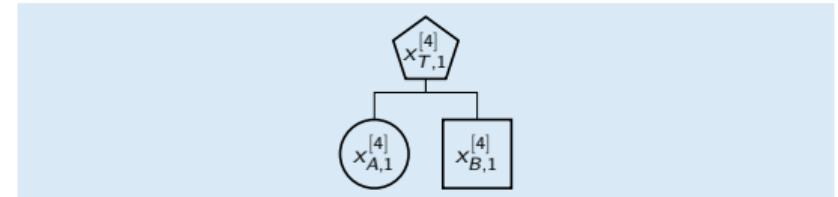
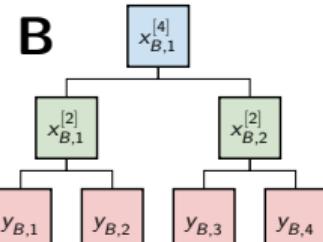
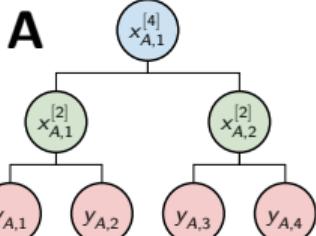
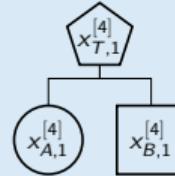
# Cross-sectional + Temporal = Cross-temporal

A cross-temporal hierarchy of three quarterly time series ( $T = A + B$ )

cross-sectional  $\longrightarrow$  temporal



temporal  $\longrightarrow$  cross-sectional



# Cross-temporal framework

Di Fonzo and Girolimetto (2023a)

$$\mathbf{X}_\tau = \begin{bmatrix} \mathbf{x}'_{1,\tau} \\ \vdots \\ \mathbf{x}'_{n,\tau} \end{bmatrix} = \begin{bmatrix} \mathbf{U}^{[4]} & \mathbf{U}^{[2]} & \mathbf{U}^{[1]} \\ x_{T,\tau}^{[4]} & x_{T,2\tau-1}^{[2]} & x_{T,2\tau}^{[2]} & y_{T,4\tau-3} & y_{T,4\tau-2} & y_{T,4\tau-1} & y_{T,4\tau} \\ x_{A,\tau}^{[4]} & x_{A,2\tau-1}^{[2]} & x_{A,2\tau}^{[2]} & y_{A,4\tau-3} & y_{A,4\tau-2} & y_{A,4\tau-1} & y_{A,4\tau} \\ x_{B,\tau}^{[4]} & x_{B,2\tau-1}^{[2]} & x_{B,2\tau}^{[2]} & y_{B,4\tau-3} & y_{B,4\tau-2} & y_{B,4\tau-1} & y_{B,4\tau} \\ \mathbf{B}^{[4]} & \mathbf{B}^{[2]} & \mathbf{B}^{[1]} \end{bmatrix}$$

- Any cross-temporal matrix may be constructed starting from the one-dimensional equivalents  $\mathbf{C}_{ct} \rightarrow$  easy to compute as a function of  $\mathbf{C}_{cs}$  and of the highest time frequency ( $m$ )

$$\mathbf{C}_{cs} \mathbf{X}_\tau \mathbf{C}_{te}' = \mathbf{0}_{(n_a \times k^*)} \iff \mathbf{C}_{ct} \mathbf{x}_\tau = \mathbf{0}_{[(n_a m + nk^*) \times 1]} \quad \text{where} \quad \mathbf{x}_\tau = \text{vec}(\mathbf{X}_\tau')$$

$\mathbf{S}_{ct} \rightarrow$  fast to compute as  $\mathbf{S}_{cs} \otimes \mathbf{S}_{te}$

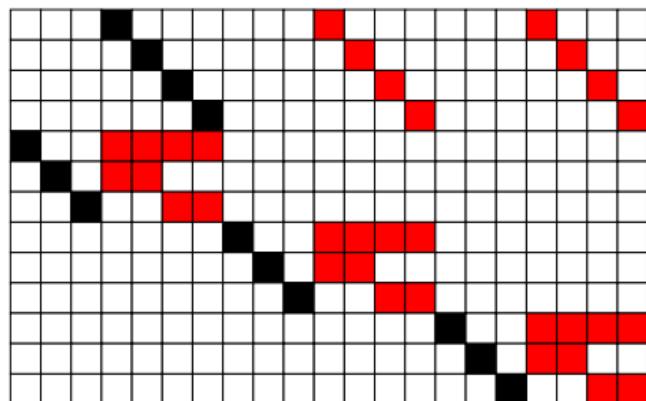
$$\mathbf{X}_\tau = \mathbf{S}_{cs} \mathbf{B}_\tau^{[1]} \mathbf{S}_{te}' \iff \mathbf{x}_\tau = \mathbf{S}_{ct} \mathbf{b}_\tau^{[1]} \quad \text{where} \quad \mathbf{b}_\tau^{[1]} = \text{vec}(\mathbf{B}_\tau^{[1]'}')$$

# Cross-temporal representations

Two dimensions (**spatio-temporal**) to capture the complete nature of a multiple time series

## Zero-constrained representation

$$\mathbf{C}_{ct} = \begin{bmatrix} (\mathbf{0}_{(n_a m \times nk^*)} \otimes \mathbf{I}_m \otimes \mathbf{C}_{cs}) \mathbf{P}' \\ \mathbf{I}_n \otimes \mathbf{C}_{te} \end{bmatrix}$$



$$\text{where } \mathbf{P} \text{vec}(\mathbf{X}_\tau) = \text{vec}(\mathbf{X}'_\tau)$$

## Structural representation

$$S_{cs} \otimes S_{te} = S_{ct}$$

Legend:

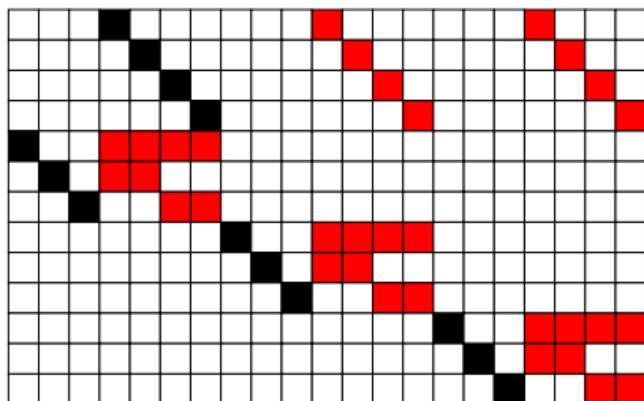
- White square = 0
- Black square = 1
- Red square = -1

## Cross-temporal representations

Two dimensions (**spatio-temporal**) to capture the complete nature of a multiple time series

## Zero-constrained representation

$$\mathbf{C}_{ct} = \begin{bmatrix} (\mathbf{0}_{(n_a m \times nk^*)} \mathbf{I}_m \otimes \mathbf{C}_{cs}) \mathbf{P}' \\ \mathbf{I}_n \otimes \mathbf{C}_{te} \end{bmatrix}$$



where  $\mathbf{P}\text{vec}(\mathbf{X}_\tau) = \text{vec}(\mathbf{X}'_\tau)$

## Structural representation

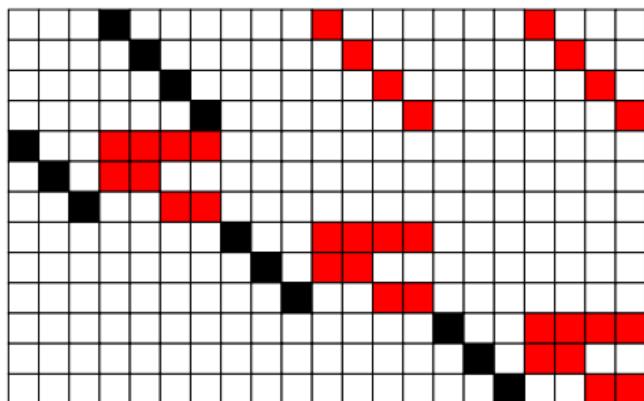
$$\begin{array}{c}
 S_{cs} \\
 \otimes \\
 = I_2
 \end{array}
 \quad
 \begin{array}{c}
 S_{te} \\
 \otimes \\
 = I_4
 \end{array}
 \quad
 \begin{array}{c}
 S_{ct} \\
 = \\
 \text{A 16x16 grid with black and white squares. The pattern is: white, black, white, black. This is followed by a vertical column of four black squares. Then the pattern repeats: white, black, white, black. This is followed by a vertical column of four black squares. The process continues until the end of the grid.}
 \end{array}$$

## Cross-temporal representations

Two dimensions (**spatio-temporal**) to capture the complete nature of a multiple time series

## Zero-constrained representation

$$\mathbf{C}_{ct} = \begin{bmatrix} (\mathbf{0}_{(n_a m \times nk^*)} \mathbf{I}_m \otimes \mathbf{C}_{cs}) \mathbf{P}' \\ \mathbf{I}_n \otimes \mathbf{C}_{te} \end{bmatrix}$$



where  $\mathbf{P}\text{vec}(\mathbf{X}_\tau) = \text{vec}(\mathbf{X}'_\tau)$

## Structural representation

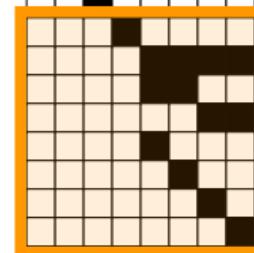
$$S_{cs} \otimes S_{te} = S_{ct}$$

$= I_2$

$\otimes$

$= I_4$

$$\begin{array}{c} \text{white square} \\ \text{black square} \\ \text{red square} \end{array} = \begin{array}{c} 0 \\ 1 \\ -1 \end{array}$$



$$\neq I_8$$

# Point forecast reconciliation

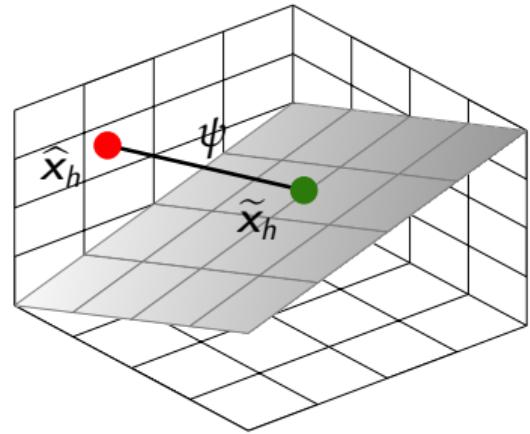
Wickramasuriya *et al.* (2019); Panagiotelis *et al.* (2021)

## Definition

Forecast reconciliation aims to adjust the base forecast  $\hat{\mathbf{x}}_h$  via a mapping  $\psi : \mathbb{R}^{n(m+k^*)} \rightarrow \mathfrak{s}$  such that

$$\tilde{\mathbf{x}}_h = \psi(\hat{\mathbf{x}}_h),$$

where  $\tilde{\mathbf{x}}_h \in \mathfrak{s}$  is the vector of the reconciled forecasts



# Point forecast reconciliation

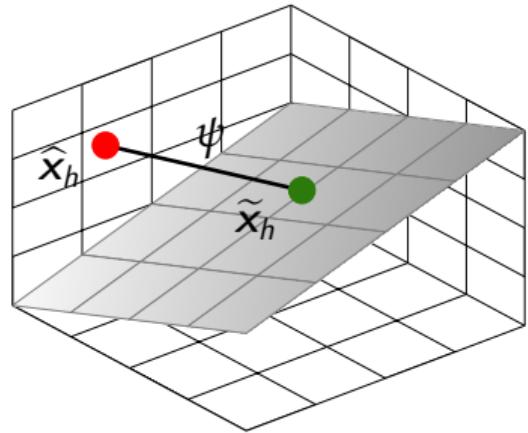
Wickramasuriya *et al.* (2019); Panagiotelis *et al.* (2021)

## Definition

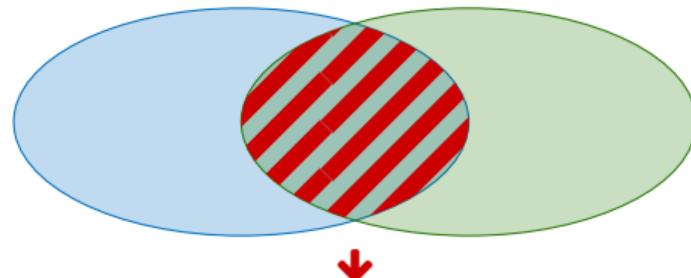
Forecast reconciliation aims to adjust the base forecast  $\hat{\mathbf{x}}_h$  via a mapping  $\psi : \mathbb{R}^{n(m+k^*)} \rightarrow \mathfrak{s}$  such that

$$\tilde{\mathbf{x}}_h = \psi(\hat{\mathbf{x}}_h),$$

where  $\tilde{\mathbf{x}}_h \in \mathfrak{s}$  is the vector of the reconciled forecasts



$\mathfrak{s}_{cs}$   
cross-sectional coherent  
subspace spanned by  
the columns of  $S_{cs}$



$\mathfrak{s}_{te}$   
temporal coherent  
subspace spanned by  
the columns of  $S_{te}$

$$\mathfrak{s} \text{ cross-temporal coherent subspace spanned by the columns of } S_{ct} = S_{cs} \otimes S_{te}$$

# Optimal forecast reconciliation

Target  
 $Cx_h = \mathbf{0}$

Base forecasts  
 $C\hat{x}_h \neq \mathbf{0}$

Reconciled forecasts  
 $C\tilde{x}_h = \mathbf{0}$

1. Forecast **all series at all levels** of aggregation (using e.g. ARIMA, ETS, VAR ...) → **base forecasts**

# Optimal forecast reconciliation

Target $\mathbf{C}x_h = \mathbf{0}$	Base forecasts $\mathbf{C}\hat{x}_h \neq \mathbf{0}$	Reconciled forecasts $\mathbf{C}\tilde{x}_h = \mathbf{0}$
--	---	--

1. Forecast **all series at all levels** of aggregation (using e.g. ARIMA, ETS, VAR ...) → **base forecasts**
2. Make the base forecasts **coherent** → **reconciled forecasts**

► **Projection approach** (Byron, 1978, 1979)

$$\begin{aligned} \arg \min_{x_h} & (\hat{x}_h - x_h)' \Omega_{ct}^{-1} (\hat{x}_h - x_h) \\ \text{s.t. } & \mathbf{C}_{ct} x_h = \mathbf{0} \end{aligned} \Rightarrow \tilde{x}_h = \left[ I - \Omega_{ct} \mathbf{C}'_{ct} (\mathbf{C}_{ct} \Omega_{ct} \mathbf{C}'_{ct})^{-1} \mathbf{C}_{ct} \right] \hat{x}_h = \psi(\hat{x}_h)$$

► **Structural approach** (Wickramasuriya *et al.*, 2019, cross-temporal extension)

$$\begin{aligned} \min_G & \text{tr} (\mathbf{S}_{ct} \mathbf{G} \Omega_{ct} \mathbf{G}' \mathbf{S}'_{ct}) \\ \text{s.t. } & \mathbf{S}_{ct} \mathbf{G} \mathbf{S}_{ct} = \mathbf{S}_{ct} \end{aligned} \Rightarrow \tilde{x}_h = \mathbf{S}_{ct} \underbrace{(\mathbf{S}'_{ct} \Omega_{ct}^{-1} \mathbf{S}_{ct})^{-1} \mathbf{S}'_{ct} \Omega_{ct}^{-1}}_G \hat{x}_h = \psi(\hat{x}_h)$$

# Optimal forecast reconciliation

Target $\mathbf{C}x_h = \mathbf{0}$	Base forecasts $\mathbf{C}\hat{x}_h \neq \mathbf{0}$	Reconciled forecasts $\mathbf{C}\tilde{x}_h = \mathbf{0}$
--	---	--

1. Forecast **all series at all levels** of aggregation (using e.g. ARIMA, ETS, VAR ...) → **base forecasts**
2. Make the base forecasts **coherent** → **reconciled forecasts**

► **Projection approach** (Byron, 1978, 1979)

$$\begin{aligned} \arg \min_{x_h} & (\hat{x}_h - x_h)' \Omega_{ct}^{-1} (\hat{x}_h - x_h) \\ \text{s.t. } & \mathbf{C}_{ct} x_h = \mathbf{0} \end{aligned} \Rightarrow \tilde{x}_h = \left[ I - \Omega_{ct} \mathbf{C}'_{ct} (\mathbf{C}_{ct} \Omega_{ct} \mathbf{C}'_{ct})^{-1} \mathbf{C}_{ct} \right] \hat{x}_h = \psi(\hat{x}_h)$$

► **Structural approach** (Wickramasuriya *et al.*, 2019, cross-temporal extension)

$$\begin{aligned} \min_G & \text{tr} (\mathbf{S}_{ct} \mathbf{G} \Omega_{ct} \mathbf{G}' \mathbf{S}'_{ct}) \\ \text{s.t. } & \mathbf{S}_{ct} \mathbf{G} \mathbf{S}_{ct} = \mathbf{S}_{ct} \end{aligned} \Rightarrow \tilde{x}_h = \mathbf{S}_{ct} \underbrace{(\mathbf{S}'_{ct} \Omega_{ct}^{-1} \mathbf{S}_{ct})^{-1} \mathbf{S}'_{ct} \Omega_{ct}^{-1}}_G \hat{x}_h = \psi(\hat{x}_h)$$

- ⚠ The formulation of  $\text{Var}(\hat{x}_h - x_h)$  is conceptually **complex**; in practice, approximate forms of  $\text{Var}(\hat{x}_h - x_h) \approx \Omega_{ct}$  are used, possibly using in-sample residuals

# Probabilistic forecast reconciliation

Starting point

Panagiotelis *et al.* (2023)  
definitions, theorems and approaches for the cross-sectional case



Coherence and probabilistic forecast reconciliation according to the more general cross-temporal framework

- ─ A generalized probabilistic cross-temporal framework for **count data** can be obtained following the cross-sectional work of Corani *et al.* (2023) → however, we only focus on the **continuous case**

# Probabilistic forecast reconciliation

Starting point

Panagiotelis *et al.* (2023)  
definitions, theorems and approaches for the cross-sectional case



Coherence and probabilistic forecast reconciliation according to the more general cross-temporal framework

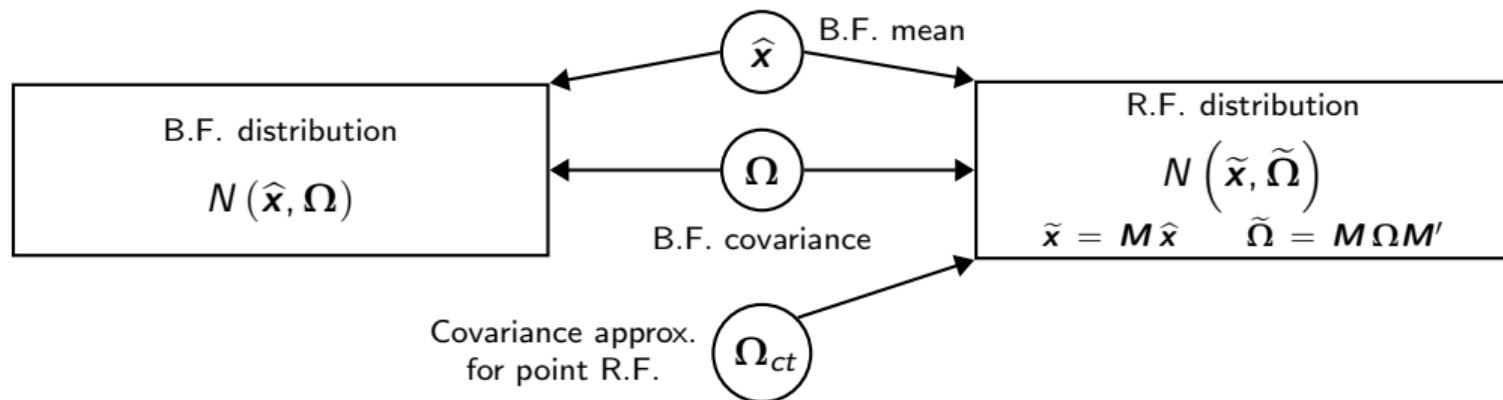
- A generalized probabilistic cross-temporal framework for **count data** can be obtained following the cross-sectional work of Corani *et al.* (2023) → however, we only focus on the **continuous case**

## Thm: Cross-temporal Reconciled Samples

$$\begin{array}{ccc} (\hat{x}_1, \dots, \hat{x}_L) & \xrightarrow{\tilde{x}^{[\ell]} = \psi(\hat{x}^{[\ell]})} & (\tilde{x}_1, \dots, \tilde{x}_L) \\ \text{sample drawn from an incoherent} & & \text{sample drawn from the reconciled} \\ \text{probability measure } \hat{\nu} & & \text{probability measure } \tilde{\nu} \end{array}$$

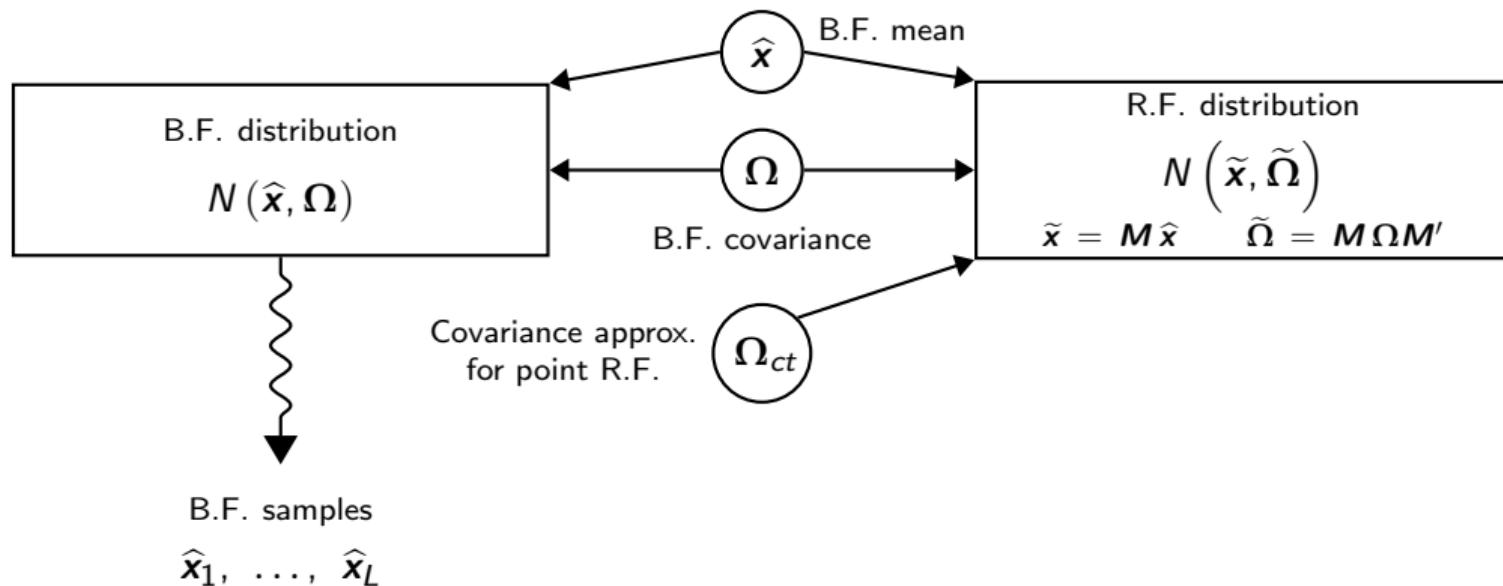
# A parametric gaussian approach

Base forecasts sample paths (Panagiotelis *et al.*, 2023; Wickramasuriya, 2023)



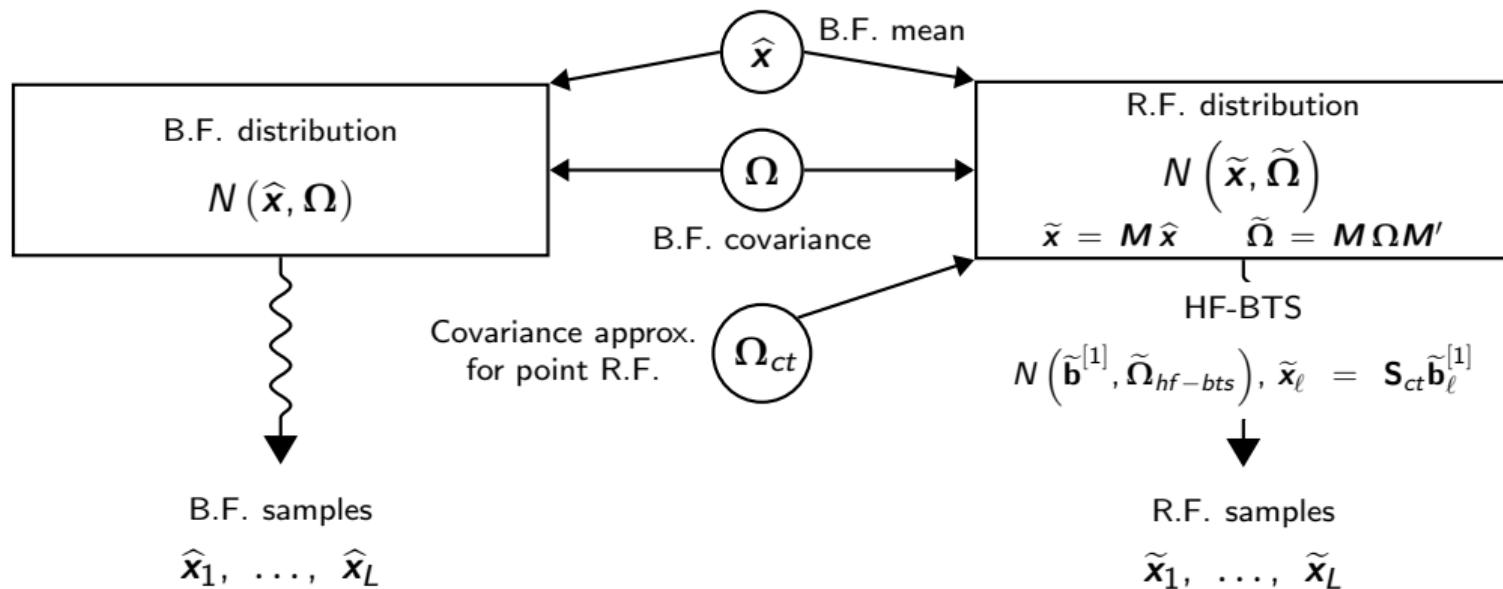
# A parametric gaussian approach

Base forecasts sample paths (Panagiotelis *et al.*, 2023; Wickramasuriya, 2023)



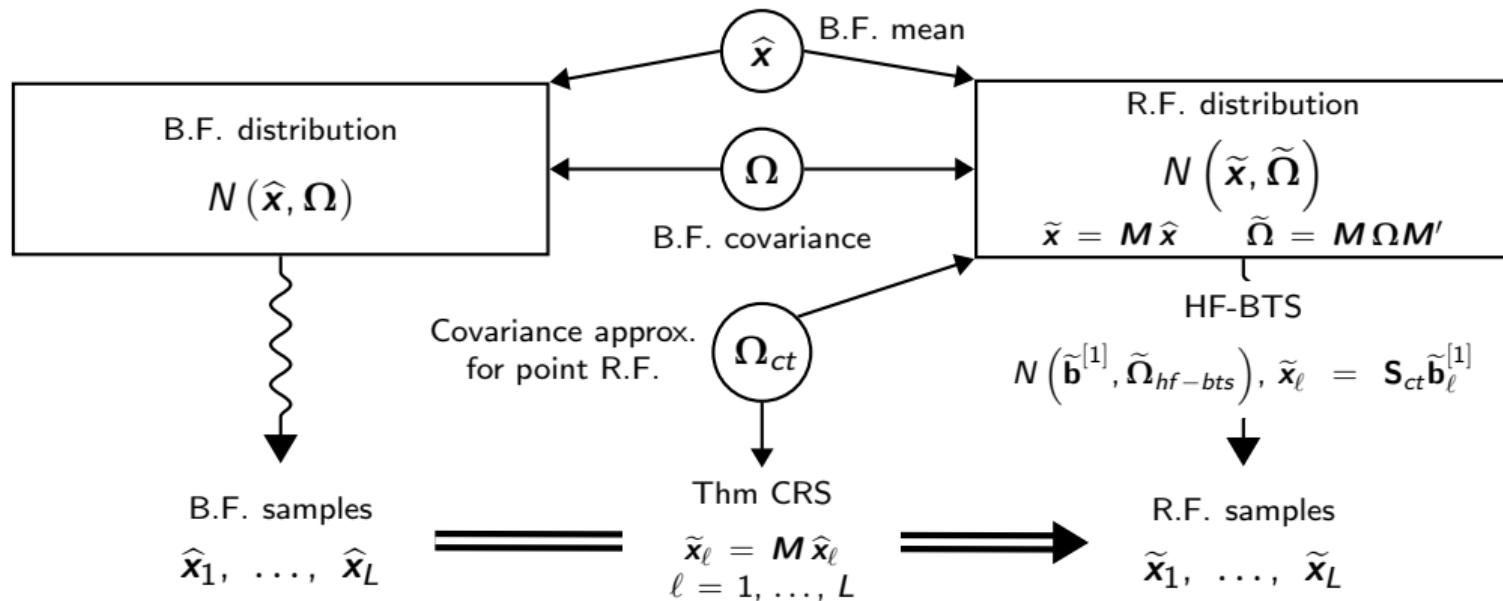
# A parametric gaussian approach

Base forecasts sample paths (Panagiotelis *et al.*, 2023; Wickramasuriya, 2023)



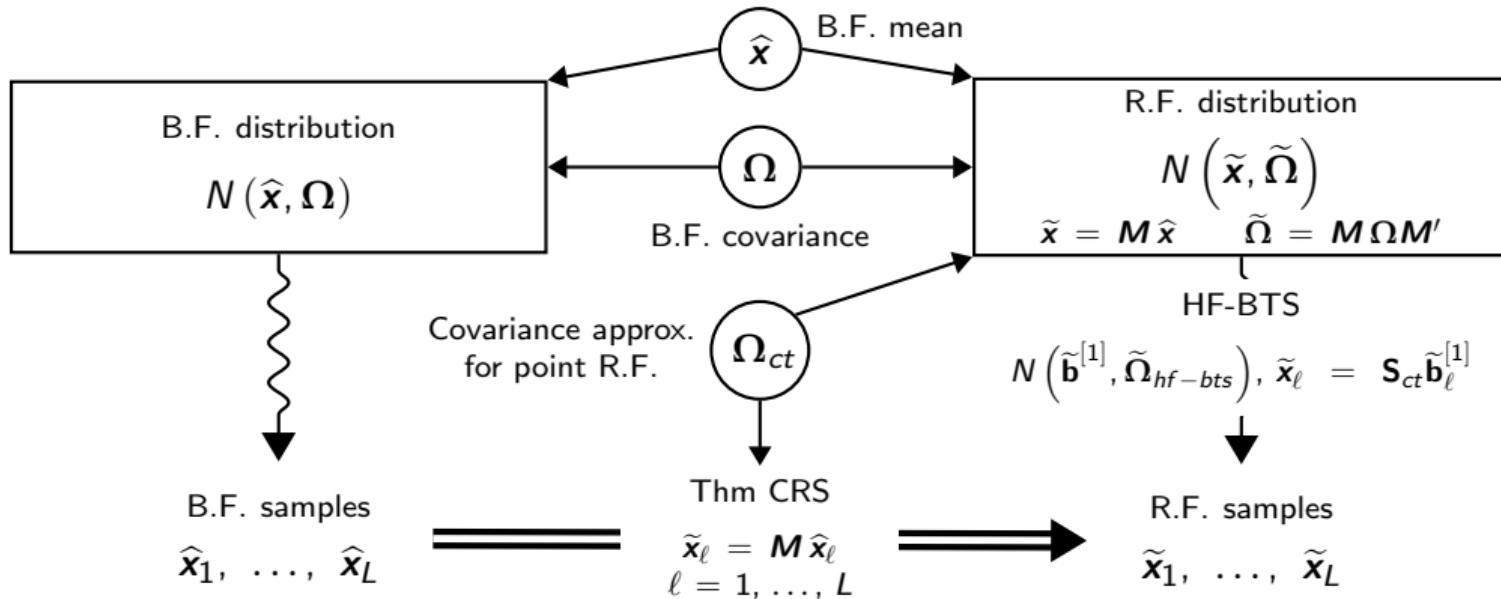
# A parametric gaussian approach

Base forecasts sample paths (Panagiotelis *et al.*, 2023; Wickramasuriya, 2023)



# A parametric gaussian approach

Base forecasts sample paths (Panagiotelis *et al.*, 2023; Wickramasuriya, 2023)



- $\Omega$  and  $\Omega_{ct}$  estimates may consider cross-sectional and/or temporal dimensions (G, H, B, HB)
- Using in-sample residuals is challenging → multi-step residuals

# A non-parametric bootstrap approach

Base forecasts sample paths (Panagiotelis *et al.*, 2023)

- Analytical expressions for the base and reconciled forecast distributions are challenging and parametric assumptions can be restrictive and unrealistic
- Joint (block) Bootstrap: simulate future base sample paths from all models using bootstrapped residuals, then reconcile them to obtain coherent sample paths

$$\hat{\mathbf{x}}_{i,\ell}^{[k]} = f_i(\mathcal{M}_i, \hat{\mathbf{e}}_{i,\tau}^{[k]})$$

- Need to generate samples that preserve cross-temporal relationships

# A non-parametric bootstrap approach

Base forecasts sample paths (Panagiotelis *et al.*, 2023)

- Analytical expressions for the base and reconciled forecast distributions are challenging and parametric assumptions can be restrictive and unrealistic
- Joint (block) Bootstrap: simulate future base sample paths from all models using bootstrapped residuals, then reconcile them to obtain coherent sample paths

$$\hat{\mathbf{x}}_{i,\ell}^{[k]} = f_i(\mathcal{M}_i, \hat{\mathbf{e}}_{i,\tau}^{[k]})$$

- Need to generate samples that preserve cross-temporal relationships

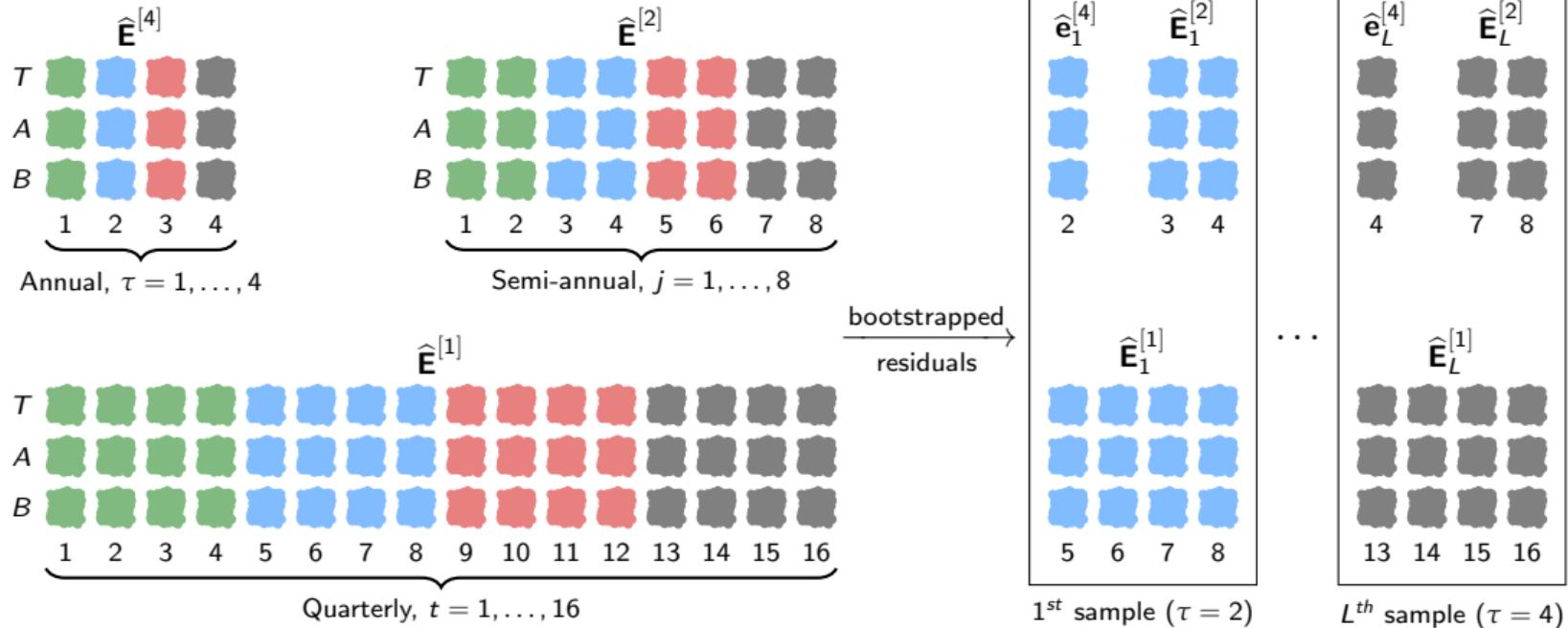
## 💡 Idea

1. sampling the index time  $\tau$  from the most temporally aggregated level
2. using it to determine the indices for all the other levels

- + Easily scalable in order to utilize multiple computing power simultaneously for each series
- + Take into account the dependence between the different levels of temporal aggregation and not only the cross-sectional dependencies

# Example of bootstrapped residuals

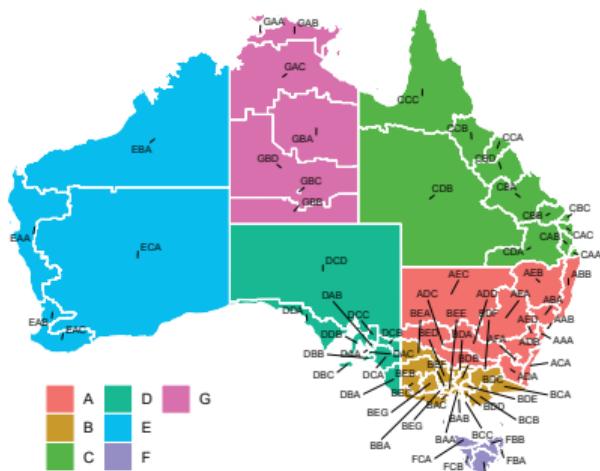
$T = A + B$ , 16 quarterly data: the green, blue, red and black colors correspond, respectively, to years 1, 2, 3 and 4.



# Forecasting the Australian Tourism Demand

Classical dataset in the hierarchical forecasting literature

## Geographical division



## Purpose of travel

Holiday, Visiting friends and relatives,  
Business, Other

## ■ Grouped ts (geographical divisions × purpose of travel)

	AUS	States	Zones*	Regions	Tot
g.d.	1	7	21	76	105
p.o.t.	4	28	84	304	420
Tot	5	35	105	380	525

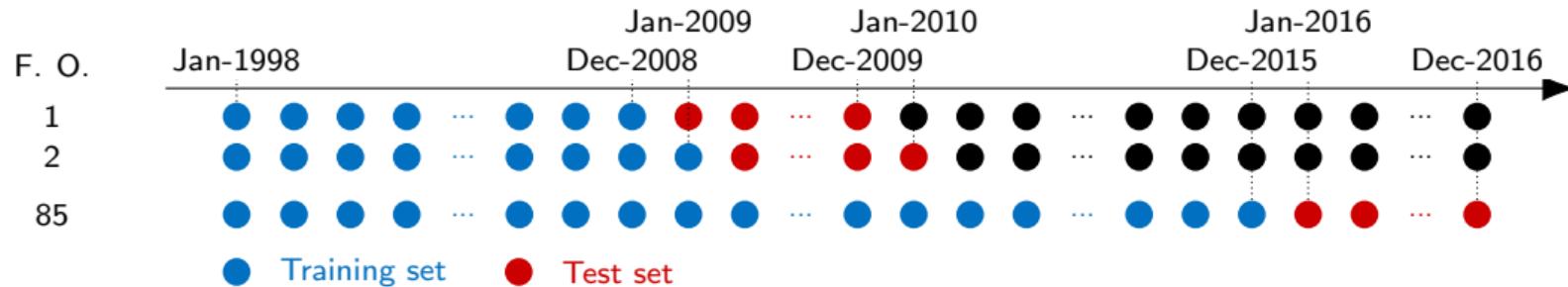
$$n_a = 221, n_b = 304, \text{ and } n = 525$$

- Unique time series, no redundancy (\*6 Zones with only one Region are included in the Regions)

## ■ Temporal framework, frequencies:

- Monthly
- Four-Monthly
- Bi-Monthly
- Semi-Annual
- Quarterly
- Annual

# The forecasting experiment



- **Monthly data:** expanding window, monthly step and 12-step ahead forecast horizons ( $h_1 = 12$ )
- For each training set, **temporally aggregated** series for any  $k \in \mathcal{K}$  are computed, and forecasts are produced up to  $h_2 = 6$ ,  $h_3 = 4$ ,  $h_4 = 3$ ,  $h_6 = 2$  and  $h_{12} = 1$  step ahead, respectively
- Automatic **ETS** forecasts on **log-transformed** data (Wickramasuriya *et al.*, 2020)
- **Accuracy indices** (Gneiting and Katzfuss, 2014)
  - Continuous Ranked Probability Score ([CRPS](#))
  - Energy Score ([ES](#))
- Dealing with **negativity issues**: [set-negative-to-zero](#) (Di Fonzo and Girolimetto, 2023b)

# Probabilistic forecasts sample paths

All the reconciliation procedures are available in FoReco

## Base forecasts sample paths

- Gaussian approach (4 variants)
- Cross-temporal Joint (block) Bootstrap (ctjb)

# Probabilistic forecasts sample paths

All the reconciliation procedures are available in FoReco

## Base forecasts sample paths

- Gaussian approach (4 variants)
- Cross-temporal Joint (block) Bootstrap (ctjb)

## Reconciliation approaches

- Cross-temporal **bottom-up** and partly bottom-up

$$ct(bu) \quad | \quad ct(shr_{cs}, bu_{te}) \quad | \quad ct(wlsv_{te}, bu_{cs})$$

- Optimal forecast reconciliation with one-step residuals (Di Fonzo and Girolimetto, 2023a)

$$oct(ols) \quad | \quad oct(struc) \quad | \quad oct(wlsv) \quad | \quad oct(bdshr)$$

- Optimal forecast reconciliation with multi-step residuals

$$oct_h(bshr) \quad | \quad oct_h(hshr) \quad | \quad oct_h(shr)$$

# RelCRPS for the Australian Tourism Demand dataset

Red: worse than the benchmark (ctjb). Bold: the best for each column. Blue: the overall lowest value

Reconciliation approach	ctjb	Generation of the base forecasts sample paths								
		Gaussian approach				ctjb	Gaussian approach			
		G	B	H	HB		G	B	H	HB
<b>All temporal level, <math>\forall k \in \{12, 6, 4, 3, 2, 1\}</math></b>								<b>Monthly level, <math>k = 1</math></b>		
base	<b>1.000</b>	0.971	0.971	0.973	0.973	<b>1.000</b>	0.972	0.972	0.972	0.972
ct(bu)	<b>1.321</b>	<b>1.011</b>	<b>1.011</b>	<b>1.011</b>	<b>1.011</b>	<b>1.077</b>	0.983	0.982	0.982	0.982
ct(shr <sub>cs</sub> , bu <sub>te</sub> )	<b>1.057</b>	0.974	0.969	0.974	0.969	0.976	0.963	0.962	0.963	0.962
ct(wlsv <sub>te</sub> , bu <sub>cs</sub> )	<b>1.062</b>	0.974	0.974	0.972	0.972	0.976	0.965	0.965	0.966	0.966
oct(ols)	0.989	0.989	0.989	0.987	0.987	0.982	0.986	0.988	0.986	0.989
oct(struc)	0.982	0.962	0.961	0.961	0.959	0.970	0.963	0.963	0.963	0.963
oct(wlsv)	0.987	0.959	0.959	0.958	0.957	0.952	0.957	0.957	0.957	0.957
oct(bdshr)	0.975	<b>0.956</b>	<b>0.953</b>	<b>0.952</b>	<b>0.951</b>	<b>0.949</b>	<b>0.955</b>	<b>0.953</b>	<b>0.954</b>	<b>0.954</b>
oct <sub>h</sub> (bshr)	0.994	<b>1.018</b>	<b>1.020</b>	<b>1.016</b>	<b>1.019</b>	0.988	<b>1.007</b>	<b>1.013</b>	<b>1.006</b>	<b>1.012</b>
oct <sub>h</sub> (hshr)	<b>0.969</b>	0.993	0.993	0.990	0.991	0.953	0.977	0.977	0.979	0.979
oct <sub>h</sub> (shr)	<b>1.007</b>	0.980	0.972	0.970	0.970	<b>1.000</b>	0.986	0.977	0.976	0.974

# RelCRPS for the Australian Tourism Demand dataset

Red: worse than the benchmark (ctjb). Bold: the best for each column. Blue: the overall lowest value

Reconciliation approach	ctjb	Generation of the base forecasts sample paths								
		Gaussian approach				ctjb	Gaussian approach			
		G	B	H	HB		G	B	H	HB
<b>All temporal level, <math>\forall k \in \{12, 6, 4, 3, 2, 1\}</math></b>								<b>Monthly level, <math>k = 1</math></b>		
base	<b>1.000</b>	0.971	0.971	0.973	0.973	<b>1.000</b>	0.972	0.972	0.972	0.972
ct(bu)	<b>1.321</b>	<b>1.011</b>	<b>1.011</b>	<b>1.011</b>	<b>1.011</b>	<b>1.077</b>	0.983	0.982	0.982	0.982
ct(shr <sub>cs</sub> , bu <sub>te</sub> )	<b>1.057</b>	0.974	0.969	0.974	0.969	0.976	0.963	0.962	0.963	0.962
ct(wlsv <sub>te</sub> , bu <sub>cs</sub> )	<b>1.062</b>	0.974	0.974	0.972	0.972	0.976	0.965	0.965	0.966	0.966
oct(ols)	0.989	0.989	0.989	0.987	0.987	0.982	0.986	0.988	0.986	0.989
oct(struc)	0.982	0.962	0.961	0.961	0.959	0.970	0.963	0.963	0.963	0.963
oct(wlsv)	0.987	0.959	0.959	0.958	0.957	0.952	0.957	0.957	0.957	0.957
oct(bdshr)	<b>0.975</b>	<b>0.956</b>	<b>0.953</b>	<b>0.952</b>	<b>0.951</b>	<b>0.949</b>	<b>0.955</b>	<b>0.953</b>	<b>0.954</b>	<b>0.954</b>
oct <sub>h</sub> (bshr)	0.994	<b>1.018</b>	<b>1.020</b>	<b>1.016</b>	<b>1.019</b>	0.988	<b>1.007</b>	<b>1.013</b>	<b>1.006</b>	<b>1.012</b>
oct <sub>h</sub> (hshr)	<b>0.969</b>	0.993	0.993	0.990	0.991	0.953	0.977	0.977	0.979	0.979
oct <sub>h</sub> (shr)	<b>1.007</b>	0.980	0.972	0.970	0.970	<b>1.000</b>	0.986	0.977	0.976	0.974

Overall, oct(bdshr) in terms of CRPS is almost always the best

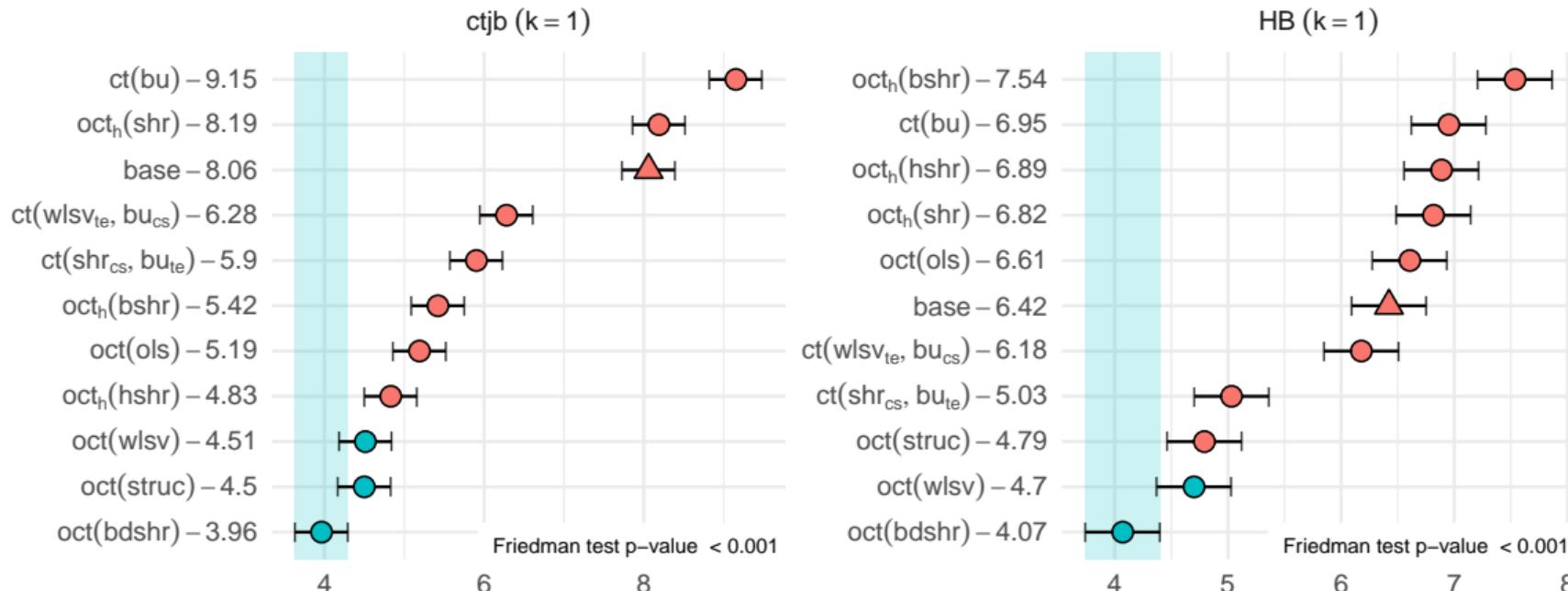
# RelCRPS for the Australian Tourism Demand dataset

Red: worse than the benchmark (ctjb). Bold: the best for each column. Blue: the overall lowest value

Reconciliation approach	ctjb	Generation of the base forecasts sample paths								
		Gaussian approach				ctjb	Gaussian approach			
		G	B	H	HB		G	B	H	HB
<b>All temporal level, <math>\forall k \in \{12, 6, 4, 3, 2, 1\}</math></b>								<b>Monthly level, <math>k = 1</math></b>		
base	1.000	0.971	0.971	0.973	0.973	1.000	0.972	0.972	0.972	0.972
ct(bu)	<b>1.321</b>	<b>1.011</b>	<b>1.011</b>	<b>1.011</b>	<b>1.011</b>	<b>1.077</b>	0.983	0.982	0.982	0.982
ct(shr <sub>cs</sub> , bu <sub>te</sub> )	<b>1.057</b>	0.974	0.969	0.974	0.969	0.976	0.963	0.962	0.963	0.962
ct(wlsv <sub>te</sub> , bu <sub>cs</sub> )	<b>1.062</b>	0.974	0.974	0.972	0.972	0.976	0.965	0.965	0.966	0.966
oct(ols)	0.989	0.989	0.989	0.987	0.987	0.982	0.986	0.988	0.986	0.989
oct(struc)	0.982	0.962	0.961	0.961	0.959	0.970	0.963	0.963	0.963	0.963
oct(wlsv)	0.987	0.959	0.959	0.958	0.957	0.952	0.957	0.957	0.957	0.957
oct(bdshr)	0.975	<b>0.956</b>	<b>0.953</b>	<b>0.952</b>	<b>0.951</b>	<b>0.949</b>	<b>0.955</b>	<b>0.953</b>	<b>0.954</b>	<b>0.954</b>
oct <sub>h</sub> (bshr)	0.994	<b>1.018</b>	<b>1.020</b>	<b>1.016</b>	<b>1.019</b>	0.988	<b>1.007</b>	<b>1.013</b>	<b>1.006</b>	<b>1.012</b>
oct <sub>h</sub> (hshr)	<b>0.969</b>	0.993	0.993	0.990	0.991	0.953	0.977	0.977	0.979	0.979
oct <sub>h</sub> (shr)	<b>1.007</b>	0.980	0.972	0.970	0.970	<b>1.000</b>	0.986	0.977	0.976	0.974

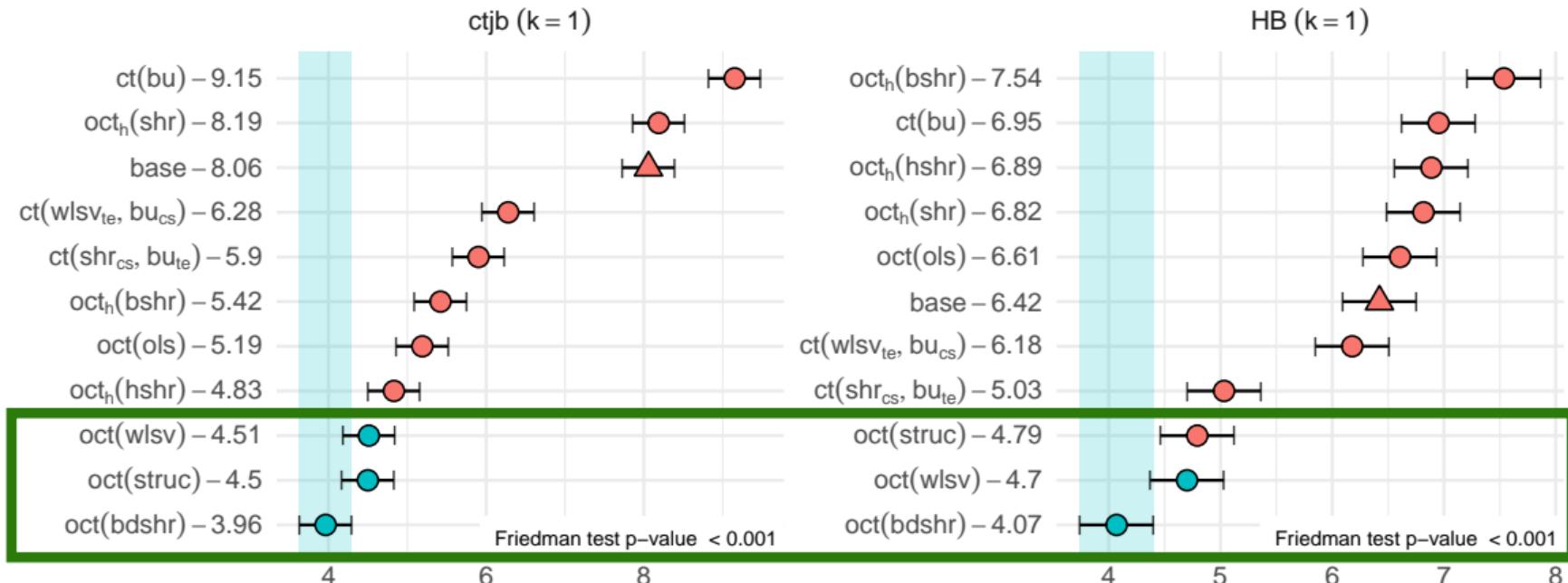
# MCB Nemenyi test

R package `tsutils` (Kourentzes, 2022). The mean rank of each approach is shown to the right of its name. Statistical differences are indicated if the intervals of two forecast reconciliation procedures do not overlap



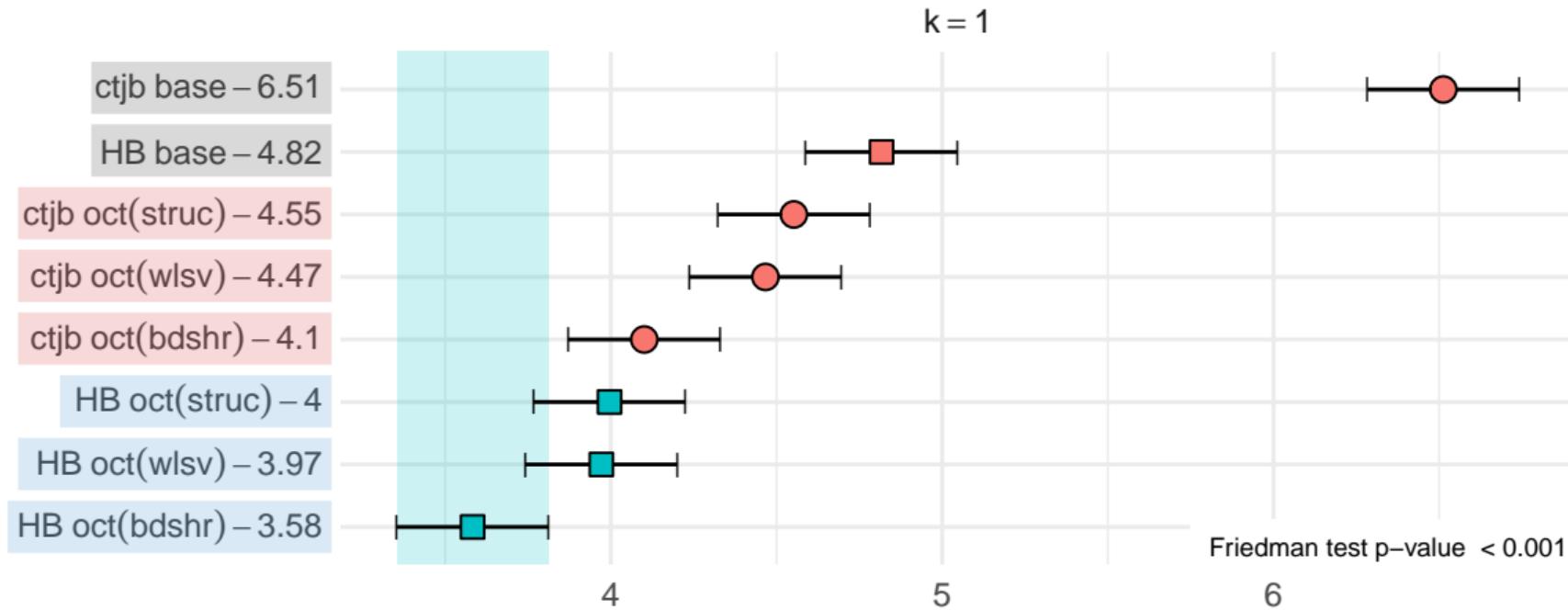
# MCB Nemenyi test

R package `tsutils` (Kourentzes, 2022). The mean rank of each approach is shown to the right of its name. Statistical differences are indicated if the intervals of two forecast reconciliation procedures do not overlap



# Gaussian vs bootstrap MCB Nemenyi test

Comparison between base and the best reconciled forecasts using gaussian □ and bootstrap ○ approaches



# Conclusions

- Notation and theorems from the cross-sectional probabilistic reconciliation setting are reinterpreted and extended for the cross-temporal case
- Parametric and non-parametric approaches are considered to simulate base forecasts that consider both cross-sectional and temporal relationships
- Issues related to the high-dimensionality are explored
  - ➔ multi-step residuals, different cross-sectional and/or temporal covariance matrix structures

# Conclusions

- Notation and theorems from the cross-sectional probabilistic reconciliation setting are reinterpreted and extended for the cross-temporal case
- Parametric and non-parametric approaches are considered to simulate base forecasts that consider both cross-sectional and temporal relationships
- Issues related to the high-dimensionality are explored
  - multi-step residuals, different cross-sectional and/or temporal covariance matrix structures
- Optimal cross-temporal reconciliation approaches significantly improve on base forecasts
- Comparison with bottom-up techniques shows that simultaneously exploiting both dimensions produces better results, especially at higher levels of temporal aggregation

## Take Home Message

Cross-temporal reconciliation have the potential to **improve forecast accuracy** and produce **coherent forecasts** also in a probabilistic framework

Prato – September 8<sup>th</sup> 2023

## IIF Workshop on Forecast Reconciliation



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA

# THANK YOU!



FoReco

**Daniele Girolimetto**  
Department of Statistical Sciences  
University of Padova (Italy)

@ danigiro  
danigiro.github.io

@ danielle.girolimetto@phd.unipd.it



arXiv:2303.17277

# References I

- Athanasiopoulos, G., Hyndman, R. J., Kourentzes, N. and Petropoulos, F. (2017) Forecasting with temporal hierarchies. *European Journal of Operational Research* **262**(1), 60–74
- Byron, R. P. (1978) The Estimation of Large Social Account Matrices. *Journal of the Royal Statistical Society. Series A (General)* **141**(3), 359–367
- Byron, R. P. (1979) Corrigenda: The Estimation of Large Social Account Matrices. *Journal of the Royal Statistical Society. Series A (General)* **142**(3), 405
- Corani, G., Azzimonti, D. and Rubattu, N. (2023) Probabilistic Reconciliation of Count Time Series. *International Journal of Forecasting, in press*
- Di Fonzo, T. and Girolimetto, D. (2022) Enhancements in cross-temporal forecast reconciliation, with an application to solar irradiance forecasts. *arXiv:2209.07146*
- Di Fonzo, T. and Girolimetto, D. (2023a) Cross-temporal forecast reconciliation: Optimal combination method and heuristic alternatives. *International Journal of Forecasting* **39**(1), 39–57
- Di Fonzo, T. and Girolimetto, D. (2023b) Spatio-temporal reconciliation of solar forecasts. *Solar Energy* **251**, 13–29
- Girolimetto, D. and Di Fonzo, T. (2023a) *FoReco: Point Forecast Reconciliation*. R package v0.2.6 **URL:** <https://danigiro.github.io/FoReco/>
- Girolimetto, D. and Di Fonzo, T. (2023b) Point and probabilistic forecast reconciliation for general linearly constrained multiple time series. *arXiv:2305.05330*
- Gneiting, T. and Katzfuss, M. (2014) Probabilistic Forecasting. *Annual Review of Statistics and Its Application* **1**(1), 125–151
- Hyndman, R. J., Ahmed, R. A., Athanasiopoulos, G. and Shang, H. L. (2011) Optimal combination forecasts for hierarchical time series. *Computational Statistics & Data Analysis* **55**(9), 2579–2589

## References II

- Kourentzes, N. (2022) *tsutils: Time Series Exploration, Modelling and Forecasting*. R package version 0.9.3 **URL:** <https://CRAN.R-project.org/package=tsutils>
- Ledoit, O. and Wolf, M. (2004) A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis* **88**(2), 365–411
- Panagiotelis, A., Athanasopoulos, G., Gamakumara, P. and Hyndman, R. J. (2021) Forecast reconciliation: A geometric view with new insights on bias correction. *International Journal of Forecasting* **37**(1), 343–359
- Panagiotelis, A., Gamakumara, P., Athanasopoulos, G. and Hyndman, R. J. (2023) Probabilistic forecast reconciliation: Properties, evaluation and score optimisation. *European Journal of Operational Research* **306**(2), 693–706
- Stellato, B., Banjac, G., Goulart, P., Bemporad, A. and Boyd, S. (2020) OSQP: An operator splitting solver for quadratic programs. *Mathematical Programming Computation* **12**(4), 637–672
- Wickramasuriya, S. L. (2023) Probabilistic forecast reconciliation under the gaussian framework. *Journal of Business & Economic Statistics, in press*
- Wickramasuriya, S. L., Athanasopoulos, G. and Hyndman, R. J. (2019) Optimal Forecast Reconciliation for Hierarchical and Grouped Time Series Through Trace Minimization. *Journal of the American Statistical Association* **114**(526), 804–819
- Wickramasuriya, S. L., Turlach, B. A. and Hyndman, R. J. (2020) Optimal non-negative forecast reconciliation. *Statistics and Computing* **30**(5), 1167–1182



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA



**George Athanasopoulos**



**Tommaso Di Fonzo**



**Rob J Hyndman**



MONASH  
University



# Projection approach

Di Fonzo and Girolimetto (2023a)

## ■ Multivariate constrained regression model

$$\begin{aligned} \widehat{\mathbf{X}}_h &= \mathbf{X}_h + \mathbf{E}_h \\ \text{s.t. } \mathbf{C}_{cs}\mathbf{X}_h &= \mathbf{0}_{(n_a^* \times 1)} \text{ and } \mathbf{C}_{te}\mathbf{X}'_h = \mathbf{0}_{(k^* \times 1)} \end{aligned} \iff \widehat{\mathbf{x}}_h = \mathbf{x}_h + \boldsymbol{\eta}_h \quad \text{s.t.} \quad \mathbf{C}_{ct}\mathbf{x}_h = \mathbf{0}_{[(n_a m + n k^*) \times 1]}$$

where  $\widehat{\mathbf{x}}_h = \text{vec}(\widehat{\mathbf{X}}'_h)$ ,  $\mathbf{x}_h = \text{vec}(\mathbf{X}'_h)$  and  $\boldsymbol{\eta}_h = \text{vec}(\mathbf{E}'_h)$  such that  $E[\boldsymbol{\eta}_h] = \mathbf{0}_{[n(m+k^*) \times 1]}$  and  $\text{Var}(\boldsymbol{\eta}_h) = \boldsymbol{\Omega}_{ct}$

## ■ Solution (Byron, 1978, 1979, cross-temporal extension)

$$\begin{aligned} \min_{\mathbf{y}_h} \quad & (\widehat{\mathbf{x}}_h - \mathbf{x}_h)' \boldsymbol{\Omega}_{ct}^{-1} (\widehat{\mathbf{x}}_h - \mathbf{x}_h) \quad \text{s.t.} \quad \mathbf{C}_{ct}\mathbf{x}_h = \mathbf{0}_{[(n_a m + n k^*) \times 1]} \\ \Rightarrow \quad & \widetilde{\mathbf{x}}_h = \left[ \mathbf{I} - \boldsymbol{\Omega}_{ct} \mathbf{C}'_{ct} (\mathbf{C}_{ct} \boldsymbol{\Omega}_{ct} \mathbf{C}'_{ct})^{-1} \mathbf{C}_{ct} \right] \widehat{\mathbf{x}}_h = \psi(\widehat{\mathbf{x}}_h) \end{aligned}$$

# Structural approach

Hyndman *et al.* (2011); Wickramasuriya *et al.* (2019); Di Fonzo and Girolimetto (2023a)

## ■ Multivariate regression model

$$\widehat{\mathbf{X}}_h = \mathbf{S}_{cs} \mathbf{B}_h^{[1]} \mathbf{S}'_{te} + \mathbf{E}_h \iff \widehat{\mathbf{x}}_h = \mathbf{S}_{ct} \mathbf{b}_h^{[1]} + \boldsymbol{\eta}_h$$

where  $\widehat{\mathbf{x}}_h = \text{vec}(\widehat{\mathbf{X}}'_h)$ ,  $\mathbf{b}_h^{[1]} = \text{vec}(\mathbf{B}_h^{[1]})'$  and  $\boldsymbol{\eta}_h = \text{vec}(\mathbf{E}'_h)$  such that  $E[\boldsymbol{\eta}_h] = \mathbf{0}_{[n(m+k^*) \times 1]}$  and  $\text{Var}(\boldsymbol{\eta}_h) = \boldsymbol{\Omega}_{ct}$

## ■ Solution (Wickramasuriya *et al.*, 2019, cross-temporal extension)

$$\begin{aligned} \min_{\mathbf{G}} \text{tr}(\mathbf{S}_{ct} \mathbf{G} \boldsymbol{\Omega}_{ct} \mathbf{G}' \mathbf{S}'_{ct}) \quad & \text{s.t.} \quad \mathbf{S}_{ct} \mathbf{G} \mathbf{S}_{ct} = \mathbf{S}_{ct} \\ \Rightarrow \quad \widetilde{\mathbf{x}}_h = \mathbf{S}_{ct} \underbrace{(\mathbf{S}'_{ct} \boldsymbol{\Omega}_{ct}^{-1} \mathbf{S}_{ct})^{-1} \mathbf{S}'_{ct} \boldsymbol{\Omega}_{ct}^{-1}}_{\mathbf{G}} \widehat{\mathbf{x}}_h = \psi(\widehat{\mathbf{x}}_h) \end{aligned}$$

with  $\text{Var}(\mathbf{x}_h - \widetilde{\mathbf{x}}_h) = \mathbf{S}_{ct} \mathbf{G} \boldsymbol{\Omega}_{ct} \mathbf{G}' \mathbf{S}'_{ct}$  if  $\widetilde{\mathbf{x}}_h = \mathbf{S}_{ct} \mathbf{G} \widehat{\mathbf{x}}_h$

# Accuracy indices for probabilistic forecasts

Gneiting and Katzfuss (2014)

## Continuous Ranked Probability Score

$$\text{CRPS}(\hat{P}_i, z_i) = \frac{1}{L} \sum_{l=1}^L |x_{i,l} - z_i| - \frac{1}{2L^2} \sum_{l=1}^L \sum_{j=1}^L |x_{i,l} - x_{i,j}|$$

## Energy Score

$$\text{ES}(\hat{P}, \mathbf{z}) = \frac{1}{L} \sum_{l=1}^L \|\mathbf{x}_l - \mathbf{z}\|_2 - \frac{1}{2(L-1)} \sum_{i=1}^{L-1} \|\mathbf{x}_i - \mathbf{x}_{i+1}\|_2$$

- $\hat{P}_i(\omega) = \frac{1}{L} \sum_{l=1}^L \mathbf{1}(x_{i,l} \leq \omega)$
- $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_L$  is a collection of  $L$  random draws taken from the predictive distribution
- $\mathbf{z} \in \mathbb{R}^n$  is the observation vector
- $i = 1, \dots, n$  denotes a single variable

# What is FoReco?

R package (Girolimetto and Di Fonzo, 2023a)

- FoReco offers classical (bottom-up and top-down), and modern (optimal and heuristic combination) **forecast reconciliation procedures** for **cross-sectional**, **temporal**, and **cross-temporal** **linearly constrained multiple time series**.

- **Matrix-based package**, exploiting the very sparse nature of the involved matrices

- **Links:**

 [cran.r-project.org/package=FoReco](https://cran.r-project.org/package=FoReco)  
 [github.com/daniGiro/FoReco](https://github.com/daniGiro/FoReco)  
 [danigiro.github.io/FoReco](https://danigiro.github.io/FoReco)



Available on 

**First release:** 01/10/2020  
**Last release:** 16/05/2023

+12k 

# Cross-temporal covariance matrix estimation

- As  $\Omega_{ct}$  (and  $\Omega$ ) is **unknown** in practice  $\rightarrow$  empirical sample covariance of the base forecasts  $\hat{\Omega}$

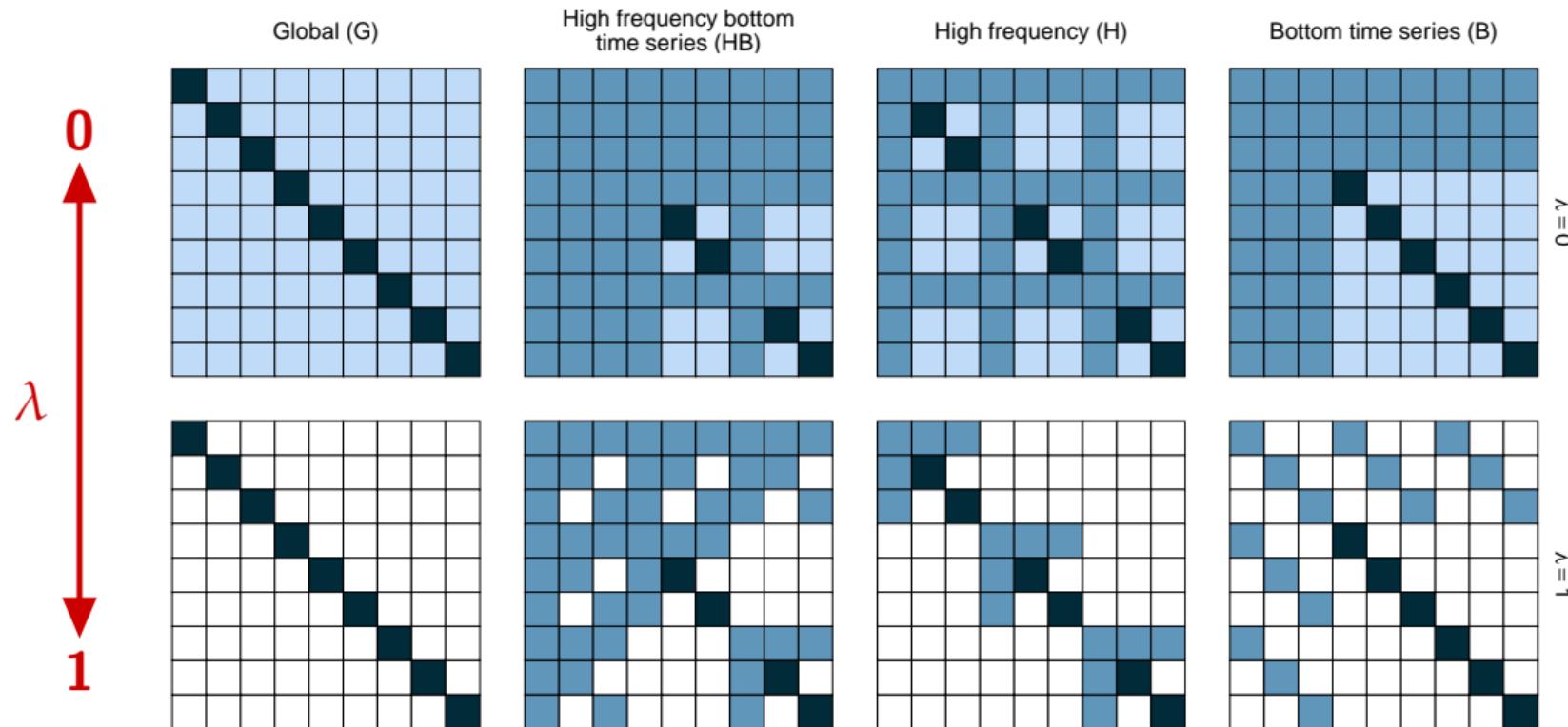
$$\frac{n(k^* + m)[n(k^* + m) - 1]}{2} \quad \text{different parameters}$$

- A possible solution to estimating many parameters, is to construct a **shrinkage estimator** using a convex combination of  $\hat{\Omega}$  and a diagonal target matrix  $\hat{\Omega}_D = \hat{\Omega} \odot I_{n(k^*+m)}$
- Shrinking all off-diagonal elements to zero, when we know that the **covariance matrix has a cross-sectional and/or temporal structure**, results in **information loss**
- Use the cross-sectional and/or temporal structure to obtain a **better estimator** for the covariance matrix of the entire system

$$\tilde{\Omega}_{ct} = \underbrace{\mathbf{S}_{ct} \Omega_{hf-bts} \mathbf{S}'_{ct}}_{\text{Cross-temporal structure}} = \underbrace{(\mathbf{I}_n \otimes \mathbf{S}_{te}) \Omega_{hf} (\mathbf{I}_n \otimes \mathbf{S}_{te})'}_{\text{Temporal structure}} = \underbrace{(\mathbf{S}_{cs} \otimes \mathbf{I}_{m+k^*}) \Omega_{bts} (\mathbf{S}_{cs} \otimes \mathbf{I}_{m+k^*})'}_{\text{Cross-sectional structure}}$$

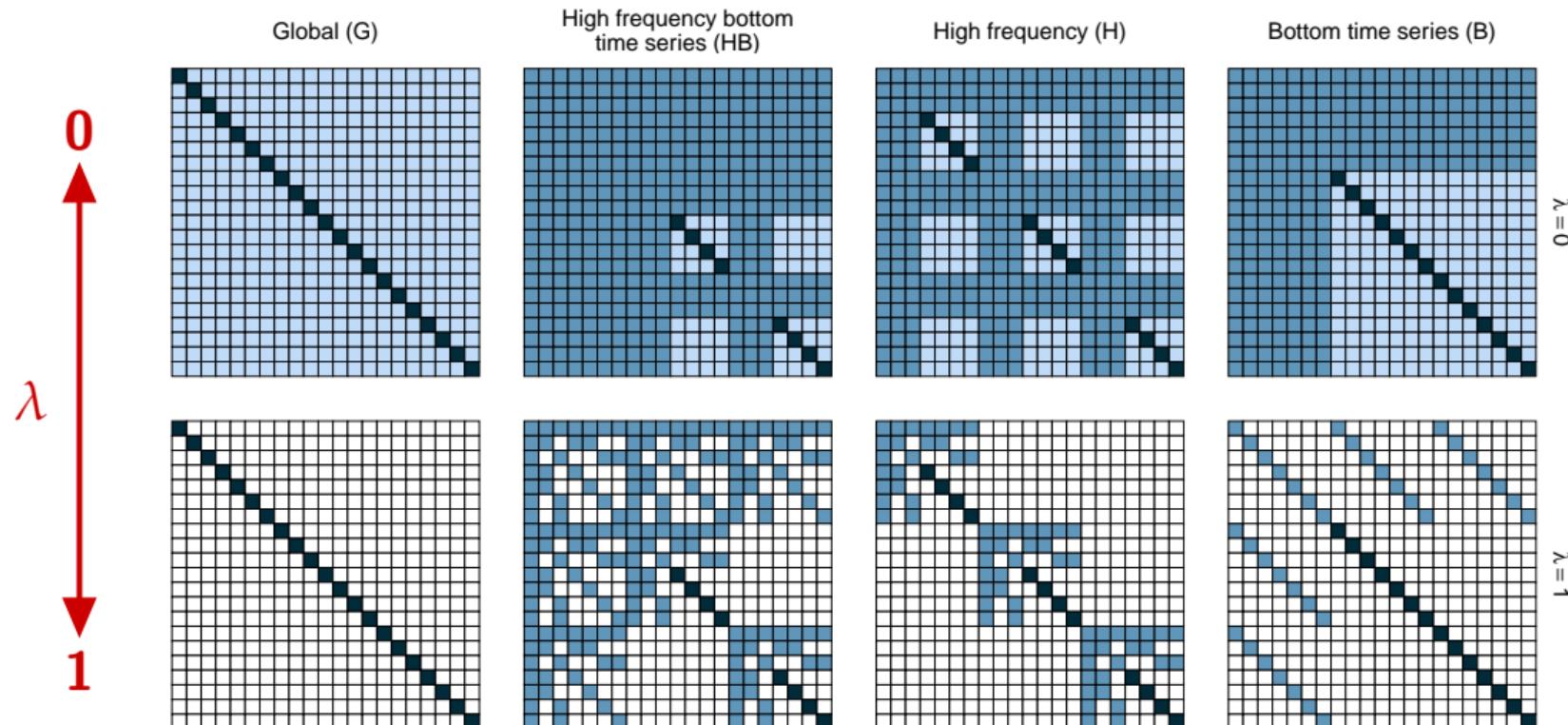
# Shrinking covariance matrix representation

Hierarchical structure with 3 time series and  $m = 4$  with two different values of  $\lambda \in \{0, 1\}$ , the shrinkage parameter

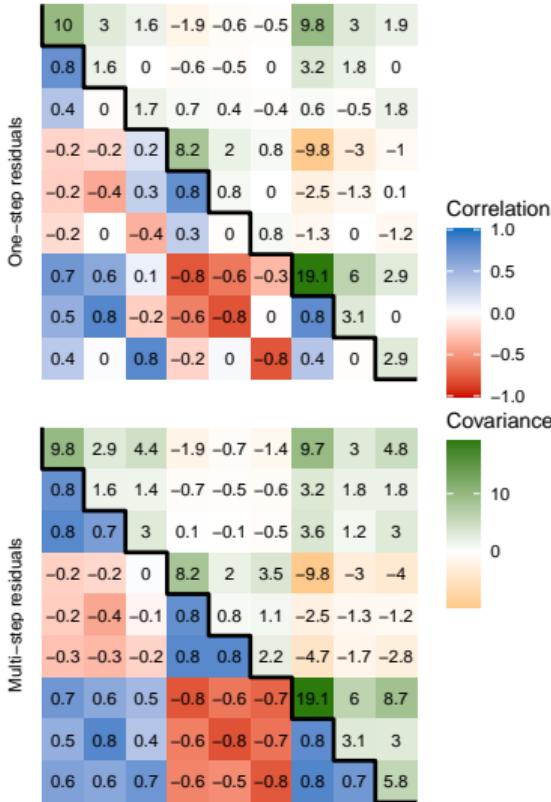


# Shrinking covariance matrix representation

Hierarchical structure with 3 time series and  $m = 2$  with two different values of  $\lambda \in \{0, 1\}$ , the shrinkage parameter



# From one- to multi-step residuals



- Model residuals may be used to estimate the covariance matrix for the base forecasts  $\Omega$  or for the reconciled formula  $\Omega_{ct}$
- In time series analysis, it is common to use residuals corresponding to **one-step ahead** forecasts
- Due to the temporal dimension, residuals corresponding to **different forecast horizons** are required
- **Multi-step residuals**

$$e_{i,h,j}^{[k]} = x_{i,j+h}^{[k]} - \hat{x}_{i,j+h|j}^{[k]}$$

where  $\hat{x}_{i,j+h|t}^{[k]}$  is the  $h$ -step fitted value.

- In general, these residuals will be **autocorrelated** except when  $h = 1$  (one-step residuals)

# Overlapping residuals

- High orders of temporal aggregation → low number of available residuals
- 💡 Use residuals calculated using overlapping series by allowing the year to have a varying starting time
- For example, suppose we have a biannual series with  $k = 2$  and  $T = 6$

$$\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_6 \end{bmatrix} \Rightarrow \mathbf{x}^{[2],0} = \begin{bmatrix} x_1^{[2],0} \\ x_2^{[2],0} \\ x_3^{[2],0} \end{bmatrix} = \begin{bmatrix} y_1 + y_2 \\ y_3 + y_4 \\ y_5 + y_6 \end{bmatrix} \text{ and } \mathbf{x}^{[2],1} = \begin{bmatrix} x_1^{[2],1} \\ x_2^{[2],1} \end{bmatrix} = \begin{bmatrix} y_2 + y_3 \\ y_4 + y_5 \end{bmatrix}$$

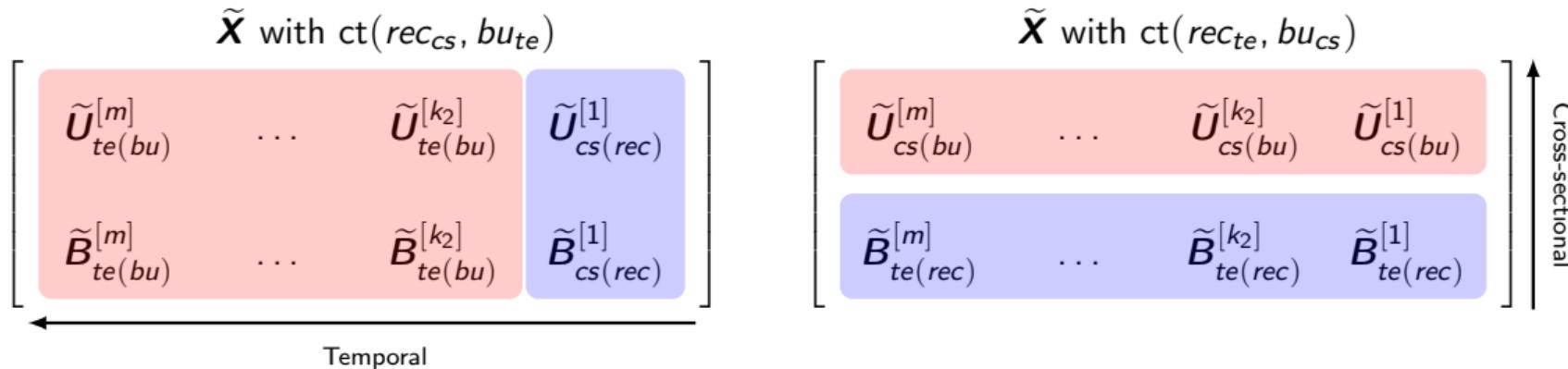
- Then
  1. Fit a model to  $\mathbf{x}^{[k],0}$  and calculate the residuals
  2. Apply the same model in step 1 to  $\mathbf{x}^{[k],s}$  for  $s = 1, \dots, k - 1$ , without re-estimating the parameters, and calculate the residuals
- Some seasonal models will not be appropriate as the seasonal pattern will be shifted for different values of  $s$

# ES ratio indices for the Australian Tourism Demand dataset

Red: worse than the benchmark (ctjb). Bold: the best for each column. Blue: the overall lowest value.

Reconciliation approach	ctjb	Generation of the base forecasts sample paths							
		Gaussian approach				ctjb Gaussian approach			
		G	B	H	HB	G	B	H	HB
		$\forall k \in \{12, 6, 4, 3, 2, 1\}$							
base	1.000	0.956	0.955	0.958	0.951	1.000	0.952	0.950	0.952
ct( <i>bu</i> )	<b>2.427</b>	0.983	0.983	0.983	0.983	<b>1.759</b>	0.982	0.982	0.982
ct( <i>shr<sub>cs</sub></i> , <i>bute</i> )	<b>1.243</b>	0.886	0.879	0.886	0.879	<b>1.098</b>	0.929	0.928	0.930
ct( <i>wlsv<sub>te</sub></i> , <i>buc<sub>s</sub></i> )	<b>1.499</b>	0.977	0.977	0.971	0.972	<b>1.241</b>	0.975	0.975	0.973
oct( <i>ols</i> )	0.955	0.893	0.891	0.893	0.888	0.975	0.937	0.936	0.935
oct( <i>struc</i> )	<b>1.085</b>	0.917	0.915	0.916	0.912	<b>1.027</b>	0.943	0.942	0.943
oct( <i>wlsv</i> )	<b>1.132</b>	0.933	0.929	0.931	0.927	<b>1.050</b>	0.951	0.949	0.950
oct( <i>bdshr</i> )	<b>1.047</b>	0.904	0.897	0.897	0.891	<b>1.009</b>	0.936	0.933	0.934
oct <sub>h</sub> ( <i>hbshr</i> )	0.956	0.889	0.886	0.888	0.884	0.975	0.937	0.936	0.937
oct <sub>h</sub> ( <i>bshr</i> )	<b>0.931</b>	<b>0.867</b>	<b>0.866</b>	<b>0.863</b>	<b>0.860</b>	<b>0.965</b>	<b>0.927</b>	0.927	0.925
oct <sub>h</sub> ( <i>hshr</i> )	<b>1.081</b>	0.935	0.931	0.935	0.927	<b>1.028</b>	0.952	0.951	0.952
oct <sub>h</sub> ( <i>shr</i> )	<b>1.068</b>	0.899	0.878	0.875	0.864	<b>1.023</b>	0.935	<b>0.923</b>	<b>0.921</b>
		$k = 1$							

# Partly bottom up



- The blue background indicates generating reconciled forecasts along one dimension, while the pink background indicates the forecasts obtained using bottom-up along the other
- L Cross-sectionally reconciled forecasts for  $k = 1$  ( $\tilde{\mathbf{U}}^{[1]}$  and  $\tilde{\mathbf{B}}^{[1]}$ ) followed by temporal bottom-up
- R Temporally reconciled forecasts of the cross-sectional bottom time series ( $\tilde{\mathbf{B}}^{[k]}, k \in \mathcal{K}$ ) followed by cross-sectional bottom-up

# Dealing with negativity issues: set-negative-to-zero

- One issue in working with time series data is the presence of negative values, which can cause difficulties for certain types of models or analyses
- Cross-sectional non negative reconciliation: Wickramasuriya *et al.* (2020)
- Non negative cross-temporal reconciliation proposed by Di Fonzo and Girolimetto (2022, 2023b)
  - State-of-the-art numerical optimization procedure (osqp, Stellato *et al.*, 2020)
  - Simple heuristic strategy: set-negative-to-zero (sntz)
    1. Negative high-frequency bottom ts reconciled forecasts are set to zero →  $\tilde{\mathbf{b}}_0^{[1]}$
    2. Apply cross-temporal bottom-up to obtain fully coherent non negative forecasts

$$\begin{array}{ccc} \hat{\mathbf{x}} & \rightarrow & \tilde{\mathbf{x}} \\ \text{Incoherent} & & \text{Coherent} \end{array} \quad \rightarrow \quad \tilde{\mathbf{x}}_0 = \mathbf{S}_{ct} \tilde{\mathbf{b}}_0^{[1]} \quad \text{Not negative and coherent}$$

- + Sntz requires much less time and computational effort than optimization
- + In the empirical application, both sntz and osqp produce similar quality forecasts

# Cross-temporal covariance approximations

Di Fonzo and Girolimetto (2023a)



oct(ols) - identity:  $\Omega_{ct} = I_{n(k^*+m)}$

oct(struc) - structural:  $\Omega_{ct} = \text{diag}(\mathbf{S}_{ct} \mathbf{1}_{mn_b})$

oct(wlsv) - series variance scaling:  $\Omega_{ct} = \hat{\Omega}_{ct,wlsv}$ , a straightforward extension of the series variance scaling matrix presented by Athanasopoulos *et al.* (2017)

oct(bdshr) - block-diagonal shrunk cross-covariance scaling:  $\Omega_{ct} = \mathbf{P} \hat{\mathbf{W}}_{ct,shr}^{BD} \mathbf{P}'$ , with  $\hat{\mathbf{W}}_{ct,shr}^{BD}$  a block diagonal matrix where each  $k$ -block ( $k = m, k_{p-1}, \dots, 1$ ) is  $I_{M_k} \otimes \hat{\mathbf{W}}_{shr}^{[k]}$ ,  $\hat{\mathbf{W}}_{shr}^{[k]}$  is the shrunk estimate of the cross-sectional covariance matrix proposed by Wickramasuriya *et al.* (2019), and  $\mathbf{P}$  is the commutation matrix such that  $\mathbf{P}\text{vec}(\mathbf{Y}_\tau) = \text{vec}(\mathbf{Y}'_\tau)$

oct(shr) - MinT-shr:  $\Omega_{ct} = \hat{\lambda} \hat{\Omega}_{ct,D} + (1 - \hat{\lambda}) \hat{\Omega}_{ct}$ , where  $\hat{\lambda}$  is an estimated shrinkage coefficient (Ledoit and Wolf, 2004),  $\hat{\Omega}_{ct,D} = I_{n(k^*+m)} \odot \hat{\Omega}_{ct}$  with  $\odot$  denoting the Hadamard product, and  $\hat{\Omega}_{ct}$  is the covariance matrix of the cross-temporal one-step ahead in-sample forecast errors