

Coherent forecast combination for linearly constrained energy series

Daniele Girolimetto

Department of Statistical Sciences, University of Padova

🌐 danigiro.github.io

🔗 github.com/danigiro

@ daniele.girolimetto@unipd.it

Energy Finance Italia 11 – EFI 11

Padova, Italia – 02/02/2026



Today's contributions: arXiv.2412.03429

Joint work with Tommaso Di Fonzo

- **The challenge:** **Energy forecasting** relies on a delicate balance between **accuracy** (e.g. through expert combination) and **coherence** (the sum of parts must be equal to the total). Standard methodologies often treat these as **separate, unconnected tasks**
- **New result** that **unifies linear forecast reconciliation and combination** in a simultaneous and statistically justified way, **improving accuracy** and **ensuring coherence** of the forecasts

Bates and Granger (1969): **linear forecast combination**

+

Stone *et al.* (1942): **constrained multivariate least-squares adjustment**

↓

optimal combined and coherent forecasts
for multiple **linearly constrained** time series

Overview

Single-task forecast combination
(Bates and Granger, 1969; Timmermann, 2006)

→ *multiple* experts, *no coherence*

Local: *one* individual variable ↑

Global: *n* individual variables ↓

Forecast reconciliation
(Stone *et al.*, 1942; Hyndman *et al.*, 2011)

→ *single* expert and *coherence*

Overview

Single-task forecast combination
(Bates and Granger, 1969; Timmermann, 2006)

→ *multiple* experts, *no coherence*

Local: *one* individual variable ↑

Global: *n* individual variables ↓

Forecast reconciliation
(Stone *et al.*, 1942; Hyndman *et al.*, 2011)

→ *single* expert and *coherence*

Multi-task forecast combination
(On the top of Sun and Deng, 2004 and
Lavancier and Rochet, 2016)

→ *multiple* experts, *no coherence*

Sequential coherent combination
Optimal coherent combination

} → *multiple* experts and *coherence*

Linear forecast combination

Bates and Granger (1969); Timmermann (2006)

- Multiple forecasts of a single variable made by $p \geq 2$ different experts are combined to produce a new forecast
- Strengths and weaknesses of each expert (e.g., different models) are exploited to improve forecast accuracy

Linear forecast combination

Bates and Granger (1969); Timmermann (2006)

- Multiple forecasts of a single variable made by $p \geq 2$ different experts are combined to produce a new forecast
- Strengths and weaknesses of each expert (e.g., different models) are exploited to improve forecast accuracy
- Single-task linear forecast combination:

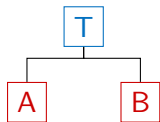
$$\hat{y}_i^c = \omega_i^\top \hat{\mathbf{y}}_i = \sum_{j=1}^p \omega_{ij} \hat{y}_i^j,$$

where $\hat{\mathbf{y}}_i$ is the vector of p forecasts and $\omega_i \in \mathbb{R}^p$ is the vector of combination weights:

- ew $\rightarrow$ equal weights
- ow_{var} $\rightarrow$ optimal weights inversely proportional to MSE (Bates and Granger, 1969)
- ow_{cov} $\rightarrow$ optimal weights in the unit simplex with MSE matrix (Conflitti *et al.*, 2015)

Forecast reconciliation

Hyndman *et al.* (2011); Wickramasuriya *et al.* (2019); Panagiotelis *et al.* (2021)



2-level: **bottom** and **upper**

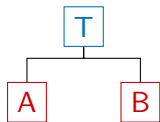
A cross-sectional hierarchical/grouped time series is a collection of n variables for which - at each time - **aggregation relationships** hold.



multiple time series with exact **linear constraints**.

Forecast reconciliation

Hyndman *et al.* (2011); Wickramasuriya *et al.* (2019); Panagiotelis *et al.* (2021)



2-level: **bottom** and **upper**

A cross-sectional hierarchical/grouped time series is a collection of n variables for which - at each time - **aggregation relationships** hold.



multiple time series with exact linear constraints.

■ Post-forecasting process:

1. Forecast **all series at all levels** of aggregation → **base forecasts** (*single expert*)
2. Make the base forecasts **coherent** → **reconciled forecasts**

Target
 $Cy = 0$

Base forecasts
 $C\hat{y} \neq 0$



Reconciled forecasts
 $C\tilde{y} = 0$

Optimal forecast reconciliation

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

■ Projection reconciliation approach (zero-constrained representation)

$$\begin{aligned}\hat{\mathbf{y}} = \mathbf{y} + \boldsymbol{\varepsilon} \text{ s.t. } \mathbf{C}\mathbf{y} = 0 &\Rightarrow \min_{\mathbf{y}} (\hat{\mathbf{y}} - \mathbf{y})^\top \mathbf{W}^{-1}(\hat{\mathbf{y}} - \mathbf{y}) \text{ s.t. } \mathbf{C}\mathbf{y} = 0 \\ &\Rightarrow \tilde{\mathbf{y}} = \left[\mathbf{I} - \mathbf{W}\mathbf{C}^\top (\mathbf{C}\mathbf{W}\mathbf{C}^\top)^{-1} \mathbf{C} \right] \hat{\mathbf{y}} = \mathbf{M}\hat{\mathbf{y}}\end{aligned}$$

- In practice, approximate forms of $\mathbf{W}$ are used, possibly using training set residuals
→ shrinkage approximation (Wickramasuriya *et al.*, 2019):

$$\mathbf{W} = \hat{\lambda} \widehat{\mathbf{W}}_D + (1 - \hat{\lambda}) \widehat{\mathbf{W}}_1$$

$\widehat{\mathbf{W}}_1$ is the covariance matrix of the one-step ahead in-sample errors ($\hat{\mathbf{e}}_t = \mathbf{y}_t - \hat{\mathbf{y}}_t$),
 $\widehat{\mathbf{W}}_D = \mathbf{I}_n \odot \widehat{\mathbf{W}}_1$, where $\odot$ denotes the Hadamard product

Coherent forecast combination

Girolimetto and Di Fonzo (2024b): n variables, p experts

- Suppose we have base forecasts of the n individual variables of the target vector $\mathbf{y} \in \mathbb{R}^n$ s.t. $\mathbf{C}\mathbf{y} = \mathbf{0}_{(n_u \times 1)}$, made by $p \geq 2$ experts:

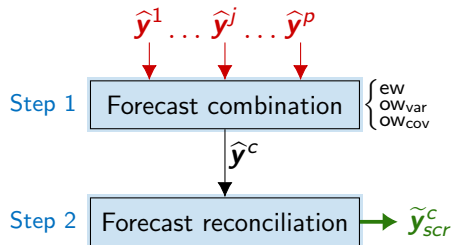
$$\hat{\mathbf{y}}^1 \in \mathbb{R}^{n_1}, \dots, \hat{\mathbf{y}}^p \in \mathbb{R}^{n_p} \quad 1 \leq n_j \leq n, \quad j = 1, \dots, p$$

- In general, we admit that the forecasts provided by each expert might refer to **different sets of individual variables** $\rightarrow$ **Unbalanced case** ($m = \sum_{j=1}^p n_j$)
- Denote $\mathbf{L} = \text{Diag}(\mathbf{L}_1, \dots, \mathbf{L}_p)$ the $(m \times np)$ selection matrix, where $\mathbf{L}_j \in \{0, 1\}^{n_j \times n}$ selects the $n_j \leq n$ entries of $\mathbf{y}$ for which base forecasts of the j -th expert are available
- **Balanced case:** $n_j = n \Rightarrow \mathbf{L} = \mathbf{I}_m, m = np$

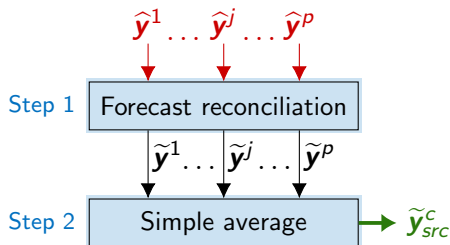
Sequential coherent forecast reconciliation

Girolimetto and Di Fonzo (2024b)

combination + reconciliation



reconciliation + simple average



✗ src approach is **limited to the balanced case** and does not apply to more general situations

■ In the following, we consider scr_{var} and scr_{cov} , with ow_{var} and ow_{cov} , respectively

■ In the working paper, we discuss also about scr_{ew} and src

Optimal coherent forecast combination

Girolimetto and Di Fonzo (2024b): n variables, p experts

- Assuming unbiased base forecasts, we consider the model

$$\hat{y}_i^j = y_i + \varepsilon_i^j, \quad i = 1, \dots, n, \quad j = 1, \dots, p$$

- The linear relationship linking $\hat{\mathbf{y}}$ and $\mathbf{y}$ can be expressed as

$$\hat{\mathbf{y}} = \begin{bmatrix} \hat{\mathbf{y}}^1 \\ \vdots \\ \hat{\mathbf{y}}^p \end{bmatrix} = \mathbf{K}\mathbf{y} + \boldsymbol{\varepsilon}, \quad \text{s.t.} \quad \mathbf{C}\mathbf{y} = \mathbf{0}_{(n_u \times 1)}$$

where $\mathbf{K} = \mathbf{L}(\mathbf{1}_p \otimes \mathbf{I}_n) \in \{0, 1\}^{m \times np}$, and $\boldsymbol{\varepsilon}$ is a zero-mean random vector with $(m \times m)$ covariance matrix $\mathbf{W} = E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^\top)$

- Linearly constrained quadratic program:

$$\tilde{\mathbf{y}}^c = \arg \min_{\mathbf{y}} (\hat{\mathbf{y}} - \mathbf{K}\mathbf{y})^\top \mathbf{W}^{-1} (\hat{\mathbf{y}} - \mathbf{K}\mathbf{y}) \quad \text{s.t.} \quad \mathbf{C}\mathbf{y} = \mathbf{0}_{(n_u \times 1)}$$

MMSE linear coherent combined forecasts

 arXiv.2412.03429

The **minimum mean square error** (MMSE) linear coherent combined forecast vector is given by

$$\tilde{\mathbf{y}}^c = \mathbf{M} \mathbf{\Omega}^T \hat{\mathbf{y}}$$

with weight matrix $\mathbf{\Psi}^T = \mathbf{M} \mathbf{\Omega}^T \in \mathbb{R}^{n \times m}$ and error covariance matrix $\widetilde{\mathbf{W}}_c = \mathbf{M} \mathbf{W}_c$, where

Property: $\tilde{\mathbf{y}}^c$ is unbiased and $\mathbf{L}_j \widetilde{\mathbf{W}}_c \mathbf{L}_j^T \preceq \mathbf{L}_j \mathbf{W}_c \mathbf{L}_j^T \preceq \mathbf{W}_j, j = 1, \dots, p$

MMSE linear coherent combined forecasts

 arXiv.2412.03429

The **minimum mean square error** (MMSE) linear coherent combined forecast vector is given by

$$\tilde{\mathbf{y}}^c = \mathbf{M} \mathbf{\Omega}^T \hat{\mathbf{y}} \quad \left\{ \begin{array}{l} \text{maps } \hat{\mathbf{y}} \in \mathbb{R}^m \rightarrow \hat{\mathbf{y}}^c \in \mathbb{R}^n \end{array} \right.$$

with weight matrix $\mathbf{\Psi}^T = \mathbf{M} \mathbf{\Omega}^T \in \mathbb{R}^{n \times m}$ and error covariance matrix $\widetilde{\mathbf{W}}_c = \mathbf{M} \mathbf{W}_c$, where

$$\mathbf{\Omega} = \mathbf{W}^{-1} \mathbf{K} \mathbf{W}_c, \quad \mathbf{W}_c = \left(\mathbf{K}^T \mathbf{W}^{-1} \mathbf{K} \right)^{-1}$$

Property: $\tilde{\mathbf{y}}^c$ is unbiased and $\mathbf{L}_j \widetilde{\mathbf{W}}_c \mathbf{L}_j^T \preceq \mathbf{L}_j \mathbf{W}_c \mathbf{L}_j^T \preceq \mathbf{W}_j, j = 1, \dots, p$

MMSE linear coherent combined forecasts

 arXiv.2412.03429

The **minimum mean square error** (MMSE) linear coherent combined forecast vector is given by

$$\tilde{\mathbf{y}}^c = \underset{\substack{\text{projects } \hat{\mathbf{y}}^c \rightarrow \mathcal{S} = \{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{C}\mathbf{y} = \mathbf{0}_{(n_u \times 1)}\}}}{\mathbf{M}} \underset{\substack{\text{maps } \hat{\mathbf{y}} \in \mathbb{R}^m \rightarrow \hat{\mathbf{y}}^c \in \mathbb{R}^n}}{\mathbf{\Omega}^\top} \hat{\mathbf{y}}$$

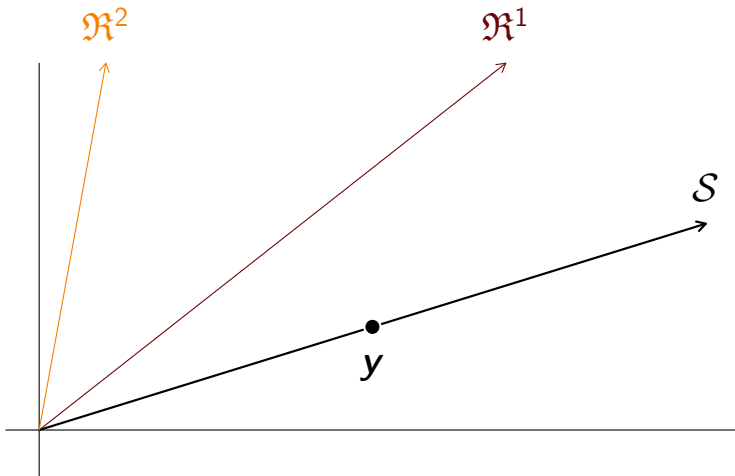
with weight matrix $\Psi^\top = \mathbf{M}\mathbf{\Omega}^\top \in \mathbb{R}^{n \times m}$ and error covariance matrix $\widetilde{\mathbf{W}}_c = \mathbf{M}\mathbf{W}_c$, where

$$\mathbf{\Omega} = \mathbf{W}^{-1}\mathbf{K}\mathbf{W}_c, \quad \mathbf{W}_c = \left(\mathbf{K}^\top \mathbf{W}^{-1} \mathbf{K}\right)^{-1} \quad \mathbf{M} = \left[\mathbf{I}_n - \mathbf{W}_c \mathbf{C}^\top \left(\mathbf{C}\mathbf{W}_c \mathbf{C}^\top\right)^{-1} \mathbf{C}\right]$$

Property: $\tilde{\mathbf{y}}^c$ is unbiased and $\mathbf{L}_j \widetilde{\mathbf{W}}_c \mathbf{L}_j^\top \preceq \mathbf{L}_j \mathbf{W}_c \mathbf{L}_j^\top \preceq \mathbf{W}_j, j = 1, \dots, p$

Graphical visualisation

2 experts, n variables

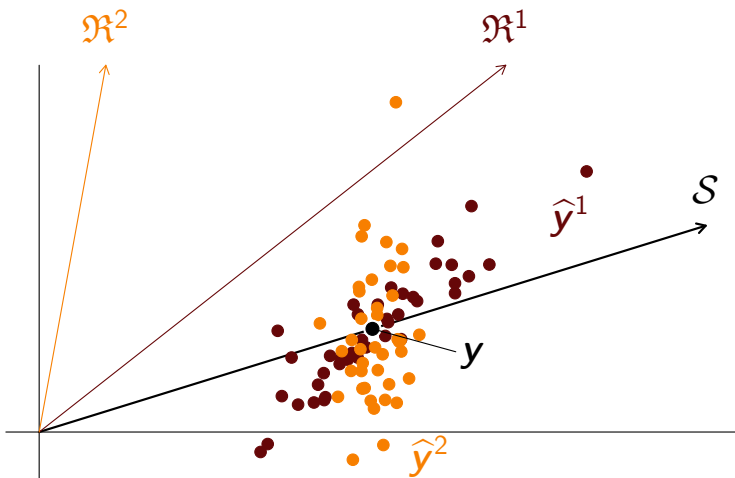


1. $\mathcal{R}^1$ and $\mathcal{R}^2$ show the most likely direction of deviations from the coherent subspace $\mathcal{S}$ for the 2 experts.
The black dot $\mathbf{y}$ denotes the (unknown) target forecast.

$$\tilde{\mathbf{y}}^c = \mathbf{M} \mathbf{\Omega}^\top \hat{\mathbf{y}}$$

Graphical visualisation

2 experts, n variables



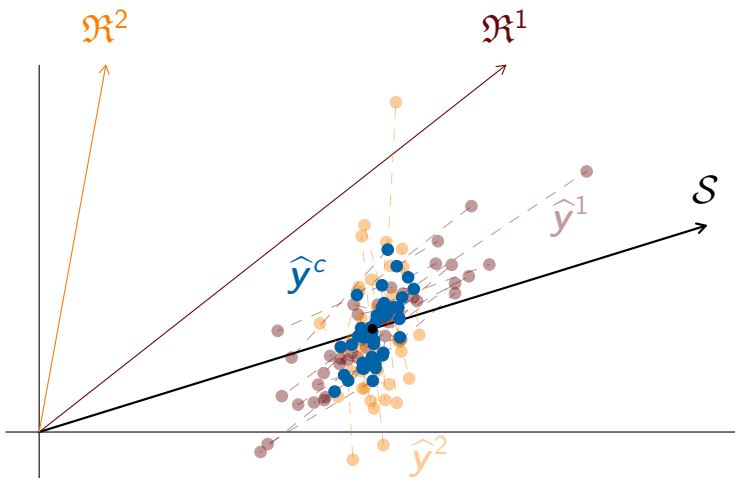
2. **Red** and **orange** points indicate the potential base forecasts for the 2 experts, $\hat{y}^1$ and $\hat{y}^2$, respectively

$$\Downarrow$$
$$\hat{y} = \begin{bmatrix} \hat{y}^1 \\ \hat{y}^2 \end{bmatrix}$$

$$\tilde{y}^c = M \Omega^\top \hat{y}$$

Graphical visualisation

2 experts, n variables

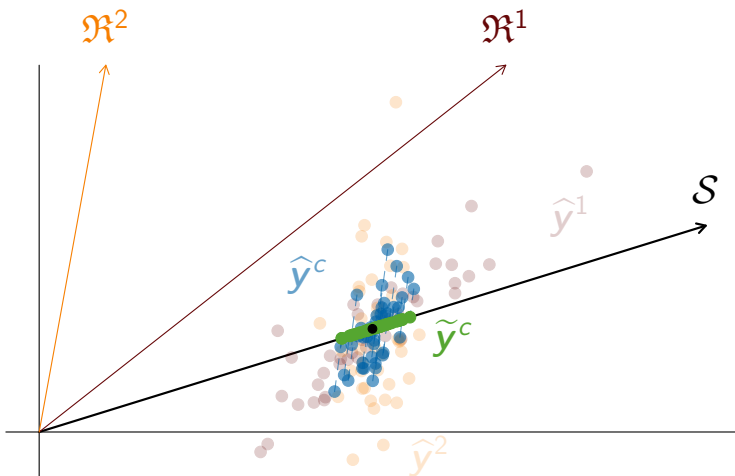


3. **Blue** points represent the unbiased MMSE linear multi-task combined forecasts, $\hat{\mathbf{y}}^c = \mathbf{\Omega}^\top \hat{\mathbf{y}}$.

$$\tilde{\mathbf{y}}^c = \mathbf{M} \mathbf{\Omega}^\top \hat{\mathbf{y}}$$

Graphical visualisation

2 experts, n variables



4. **Green** points represent the unbiased MMSE linear coherent combined forecasts, $\tilde{y}^c = M \Omega^\top \hat{y}$, as an oblique projection of $\hat{y}^c$ on S .

$$\tilde{y}^c = M \Omega^\top \hat{y}$$

About the covariance matrix

arXiv.2412.03429

- Matrix $\mathbf{W}$ determines **how the base forecasts are combined**, and then the **nature** of the coherent forecasts
- Special case: *by-expert* **block-diagonal** **shrunk** error covariance matrix (occ)

$$\widehat{\mathbf{W}}_{\text{bd-shr}} = \begin{bmatrix} \widehat{\mathbf{W}}_{1,\text{shr}} & \dots & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \widehat{\mathbf{W}}_{j,\text{shr}} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{0} & \dots & \widehat{\mathbf{W}}_{p,\text{shr}} \end{bmatrix} \xrightarrow{j=1,\dots,p} \begin{cases} \widehat{\mathbf{W}}_{j,\text{shr}} = \hat{\lambda}_j \left(\mathbf{I}_n \odot \widehat{\mathbf{W}}_j \right) + (1 - \hat{\lambda}_j) \widehat{\mathbf{W}}_j \\ \widehat{\mathbf{W}}_j = \frac{1}{T} \sum_{t=1}^T \hat{\boldsymbol{\epsilon}}_t^j \hat{\boldsymbol{\epsilon}}_t^{j\top} \\ \hat{\lambda}_j \rightarrow \text{Schäfer and Strimmer (2005)} \end{cases}$$

■ How to estimate $\widehat{\mathbf{W}}$:

1. **in-sample** errors (reconciliation) → **Australian electricity generation dataset** 📄
2. **validation** errors (combination) → **Italian energy load dataset by Terna** ⚙️

Australian electricity generation dataset

Daily electricity generation from various energy sources in Australia (AEMO, Panagiotelis *et al.*, 2023)

Linear combination matrix (8×15)

■ -1 □ 0 ■ 1

Total	+	+	+	-	+	-	+	+	+	+	+	+	+	+	+
non-Renew.									+	+	+	+	+	+	+
Renew.	+	+	+	-	+	-	+	+							
Batteries			+	-											
Hydro (-Pumps)					+	-									
Solar							+	+							
Coal									+	+					
Gas												+	+	+	+
	Wind	Biomass	Battery (Discharg.)	Battery (Charg.)	Hydro	Pumps	Solar (Rooftop)	Solar (Utility)	Distillate	Black Coal	Brown Coal	Gas (Recip.)	Gas (OCGT)	Gas (CCGT)	Gas (Steam)

23 time series with 15 bottom-level series

- **Range:** 11 June 2019 – 10/06/2020 (1 year)
- **Forecasting experiment:** expanding window, daily step and 7-step ahead forecast horizons
- $p = 3$ base forecasts (R package forecast):
 - stlf Seasonal and Trend decomposition using Loess
 - arima AutoRegressive Integrated Moving Average
 - tbats Exponential smoothing with Box-Cox transformation, ARMA errors, Trend and Seasonality
- **Accuracy indices:** AvgRelMAE and AvgRelMSE + MCB, MCS and DM-test
- 📌 **Negativity issues:** numerical optimization with non-negativity and equality constraints

AvgRelMAE for the Australian electricity generation dataset

Red: worse than the benchmark (ew). Bold: the best approach. Italic: second best approach

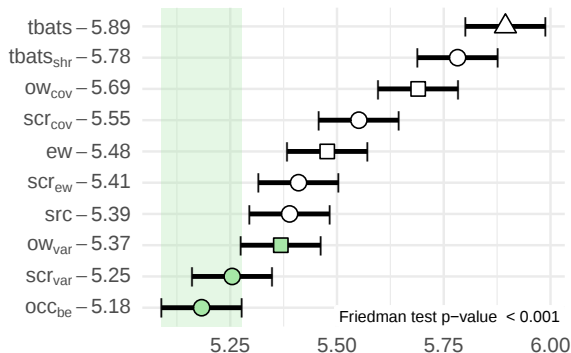
Approach	Forecast horizon							
	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 5$	$h = 6$	$h = 7$	1:7
<i>Base (incoherent forecasts) and single model reconciliation</i>								
tbats	1.0447	1.0515	1.0348	1.0266	1.0305	1.0288	1.0201	1.0331
tbats _{shr}	1.0320	1.0413	1.0231	1.0134	1.0212	1.0208	1.0188	1.0235
<i>Combination (incoherent forecasts)</i>								
ew	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
ow _{var}	0.9927	0.9921	0.9982	0.9983	0.9967	0.9967	0.9990	0.9965
ow _{cov}	1.0216	1.0208	1.0390	1.0423	1.0307	1.0250	1.0325	1.0309
<i>Coherent combination</i>								
src	0.9939	0.9941	0.9919	0.9895	0.9887	0.9908	0.9933	0.9915
scr _{ew}	0.9952	0.9959	0.9911	0.9908	0.9908	0.9932	0.9961	0.9930
scr _{var}	0.9819	0.9803	0.9869	0.9895	0.9887	0.9913	0.9972	0.9882
scr _{cov}	1.0081	1.0081	1.0270	1.0327	1.0245	1.0197	1.0250	1.0215
occ	0.9779	0.9745	0.9843	0.9852	0.9851	0.9880	0.9926	0.9843

- R package **FoCo2**
- Coherent forecast combination **outperforms incoherent** approaches
- scr_{var} is a **strong alternative** among **sequential** coherent combination procedures
- occ provides the **most accurate** forecasts for **all the horizons**

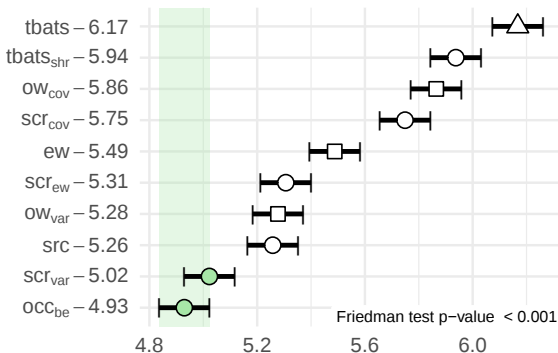
MCB Nemenyi test

R package `tsutils` (Kourentzes, 2023). The Friedman test p -value is reported in the lower right corner. The mean rank of each approach is shown to the right of its name. Statistical differences are indicated if the intervals of two forecast procedures do not overlap

MAE – All 23 ts – $h = 1$

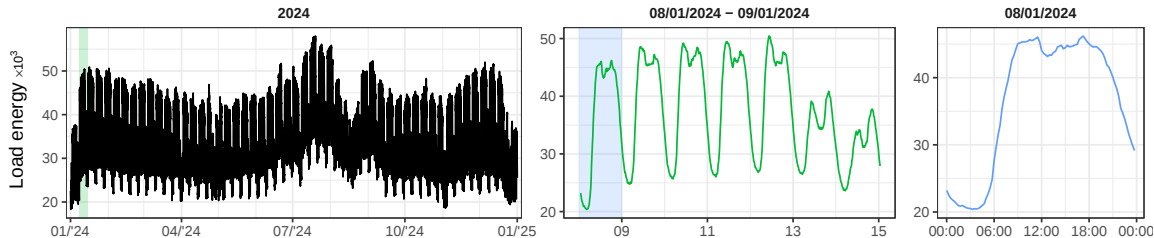


MAE – All 23 ts – $h = 1, \dots, 7$



Italian energy load forecasting by Terna

Terna is the Europe's largest independent electricity Transmission System Operator (TSO)



- Terna processes the official statistics of the **entire national electricity sector** and is responsible for **official communications** to international bodies such as **Eurostat, UN, ...**
- Among the various activities, Terna currently **publishes** on its data portal **very short-term load forecasts** for the next day, at **national level** and disaggregated by **7 bidding zones**
- **Historical 15-minutes time series** of observed and forecast load may be **easily downloaded**

The forecasting experiment

15-min data: rolling forecast experiment with daily iterations (2024) and 96-step ahead forecast horizons

Bidding zones



- 8 variables → Italy + 7 Bidding Zones (BZ)
- **Range:** 1/1/2023 – 31/12/2024 with 365 days as validation set to compute optimal weights and error covariance matrices
- **Test set:** all the 366 days of 2024
- **Accuracy evaluation:** AvgRelMAE (geometric Average Relative Mean Absolute Error) and DM-test
- **Coherency issue:** The aggregated forecasts for the 7 BZ must match the forecasts for Italy

Forecast and combination approaches

Base forecasts:

- ➔ **Terna**, exploiting a comprehensive set of influencing factors, including meteorological data, climate trends, and socio-economic variables
- ➔ the **daily random walk** (drw): $\hat{y}_{i,t+h|t} = y_{i,t-96+h}$

Local-single-task combination procedures

- ➔ Equal weights (ew)
- ➔ **Optimal single-task combination** (ow_{var} and ow_{cov})

Global-multi-task combination procedures

- ➔ **Sequential local-combination-then-reconciliation** (scr_{var} and scr_{cov})
- ➔ **Optimal multi-task combination** (occ) using a *by-expert* **block-diagonal** **shrunk** error covariance matrix

Legend: **incoherent** / **coherent** forecasts

AvgRelMAE

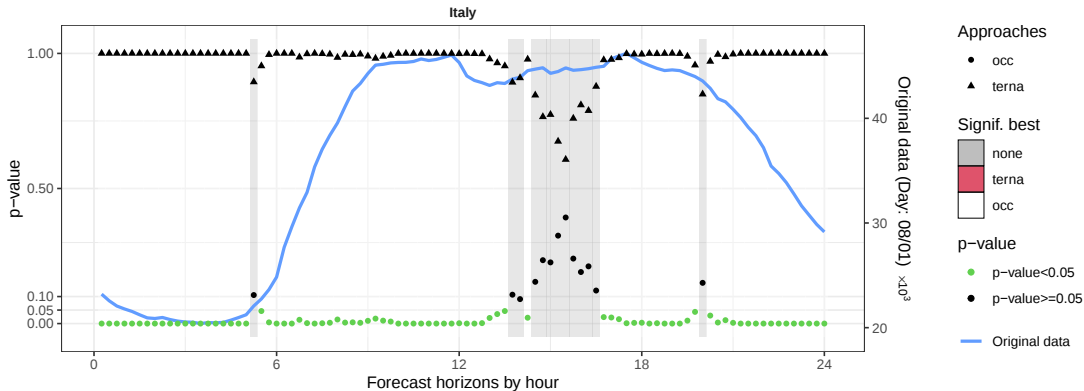
Bold entries identify the best approach. Red denotes approaches worse then Terna (benchmark)

App.	Country and 7 bidding zones								BTS	All
	Italy	North	C-North	C-South	South	Calabria	Sicily	Sardinia		
drw	4.6781	5.7847	5.1689	4.4872	6.0555	4.5870	3.1265	2.2250	4.2710	4.3199
ew	2.5376	3.0746	2.7368	2.3780	3.1048	2.4001	1.7056	1.2877	2.2872	2.3171
ow _{var}	0.9930	0.9980	0.9909	0.9897	0.9943	0.9879	0.9663	0.9282	0.9791	0.9808
ow _{cov}	0.9863	0.9905	0.9847	0.9847	0.9930	0.9854	0.9676	0.9312	0.9765	0.9777
scr _{var}	0.9863	0.9977	0.9889	0.9881	0.9927	0.9867	0.9648	0.9267	0.9777	0.9787
scr _{cov}	0.9827	0.9911	0.9841	0.9844	0.9926	0.9848	0.9674	0.9309	0.9763	0.9771
occ	0.8973	0.8997	0.8969	0.8966	0.8952	0.8947	0.8936	0.8936	0.8958	0.8960

- When using the **global approaches**, either two-step or optimal, **more accurate forecasts** are obtained
- **occ** approach consistently **outperforms Terna** and **all the other combinations**

Diebold-Mariano tests for each 15-min forecast horizon

Terna vs occ forecasts – 96 different forecast horizons – absolute loss – Italy



- occ **significantly** outperforms Terna in $\sim 86\%$ of the cases, with **no improvements** between 1 – 4 pm
- Terna **never significantly improves** with respect to occ



- **Problem:** The absence of a unified framework for **forecast combination and reconciliation** leads to **incoherent** and **less accurate** forecasts, **failing** to leverage the available information.
- **Solution:** We propose a novel approach for **combining forecasts** from **multiple** experts in **linearly constrained** time series, delivering **coherent** forecasts by construction.
- The working paper extends today's presentation through two **simulation** studies, including scenarios with **unbalanced structures**.
- **Future research:** Examine the roles of $\mathbf{M}$, $\mathbf{\Omega}$, and $\mathbf{W}$ in the MMSE formulation, and extend the framework to probabilistic forecasting.

Coherent forecast combination for linearly constrained multiple time series

Danièle Girolinetta*, Tommaso Di Fonzo

Department of Statistical Sciences, University of Padua, Padova 35121, Italy

Abstract

Linearly constrained multiple time series may be encountered in many practical contexts, such as the National Accounts (e.g., GDP disaggregated by Income, Expenditure and Output), and multilevel frameworks where the variables are organized according to hierarchies or groupings, like the total energy consumption of a country disaggregated by region and energy sources. In these cases, when multiple incoherent base forecasts for each individual variable are available, a forecast combination-and-reconciliation approach, that we call *coherent forecast combination*, may be used to improve the accuracy of the base forecasts and achieve coherence in the final result. In this paper, we develop an optimization-based technique that combines multiple unbiased base forecasts while assuring the constraints valid for the series. We present closed form expressions for the coherent combined forecast vector and its error covariance matrix in the general case where a different number of forecasts is available for each variable. We also discuss practical issues related to the covariance matrix that is part of the optimal solution. Through simulations and a forecasting experiment on the daily Australian electricity generation hierarchical time series, we show that the proposed methodology, in addition to adhering to sound statistical principles, may yield in significant improvement on base forecasts, single-task combination and single-expert reconciliation approaches as well.

Keywords: Forecasting; Linearly constrained multiple time series; Coherent forecasts; Forecast combination; Forecast reconciliation; Australian electricity generation

*Corresponding author
Email address: daniela.girolinetta@unipd.it (Danièle Girolinetta)

December 6, 2024

THANK YOU!

 arXiv.2412.03429

 danigiro/FoCo2

 CRAN/FoCo2



Energy Finance Italia 11 – EFI 11

Padova, Italy – 02/02/2026

References I

- Bates, J. and Granger, C. W. (1969) The combination of forecasts. *Operations Research Quarterly* **20**(4), 451–468
- Conflitti, C., De Mol, C. and Giannone, D. (2015) Optimal combination of survey forecasts. *International Journal of Forecasting* **31**(4), 1096–1103
- Girolimetto, D. and Di Fonzo, T. (2024a) Foco2: Coherent forecast combination. *R package* **URL:** <https://github.com/danigiro/FoCo2>
- Girolimetto, D. and Di Fonzo, T. (2024b) Coherent forecast combination for linearly constrained multiple time series. *arXiv* pp. 1–61
- Hyndman, R. J., Ahmed, R. A., Athanasopoulos, G. and Shang, H. L. (2011) Optimal combination forecasts for hierarchical time series. *Computational Statistics and Data Analysis* **55**(9), 2579–2589
- Kourentzes, N. (2023) tsutils: Time series exploration, modelling and forecasting. *R package version 0.9.4* **URL:** <https://CRAN.R-project.org/package=tsutils>
- Lavancier, F. and Rochet, P. (2016) A general procedure to combine estimators. *Computational Statistics and Data Analysis* **94**, 175–192
- Panagiotelis, A., Gamakumara, P., Athanasopoulos, G. and Hyndman, R. J. (2021) Forecast reconciliation: A geometric view with new insights on bias correction. *International Journal of Forecasting* **37**(1), 343–359

References II

- Panagiotelis, A., Gamakumara, P., Athanasopoulos, G. and Hyndman, R. J. (2023) Probabilistic forecast reconciliation: Properties, evaluation and score optimisation. *European Journal of Operational Research* **36**(2), 693–706
- Schäfer, J. and Strimmer, K. (2005) A Shrinkage Approach to Large-Scale Covariance Matrix Estimation and Implications for Functional Genomics. *Statistical Applications in Genetics and Molecular Biology* **4**(1)
- Stone, R., Champernowne, D. G. and Meade, J. E. (1942) The precision of national income estimates. *The Review of Economic Studies* **9**(2), 111–125
- Sun, S.-L. and Deng, Z.-L. (2004) Multi-sensor optimal information fusion Kalman filter. *Automatica* **40**(6), 1017–1023
- Timmermann, A. (2006) Forecast combinations. In *Handbook of Economic Forecasting*, eds G. Elliott, C. C. Granger and A. Timmermann, pp. 135–196. New York: Elsevier
- 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

AvgRelMSE for the Australian electricity generation dataset

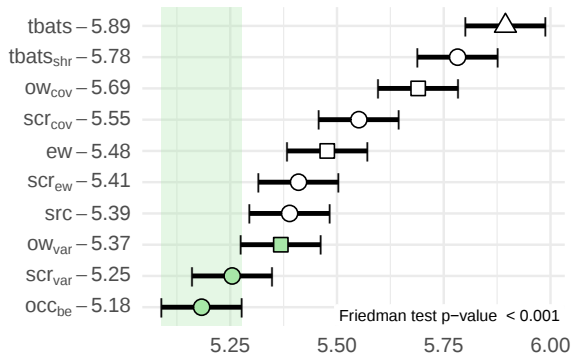
Red: worse than the benchmark (ew). Bold: the best approach. Italic: second best approach

Approach	Forecast horizon							
	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 5$	$h = 6$	$h = 7$	1:7
<i>Base (incoherent forecasts) and single model reconciliation</i>								
tbats	1.0796	1.0780	1.0445	1.0270	1.0322	1.0288	1.0142	1.0393
tbats _{shr}	1.0478	1.0577	1.0304	1.0108	1.0219	1.0213	1.0116	1.0257
<i>Combination (incoherent forecasts)</i>								
ew	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
ow _{var}	0.9840	0.9881	0.9995	1.0032	1.0020	1.0028	1.0054	0.9995
ow _{cov}	1.0279	1.0494	1.0972	1.1103	1.1009	1.0993	1.1055	1.0908
<i>Coherent combination</i>								
src	0.9827	0.9855	0.9863	0.9833	0.9852	0.9873	0.9911	0.9859
scr _{ew}	0.9875	0.9898	0.9859	0.9859	0.9885	0.9905	0.9962	0.9890
scr _{var}	0.9586	0.9683	0.9838	0.9942	0.9982	1.0017	1.0114	0.9910
scr _{cov}	1.0026	1.0287	1.0795	1.0972	1.0942	1.0913	1.0981	1.0773
occ	0.9481	0.9560	0.9754	0.9831	0.9891	0.9939	0.9993	0.9808

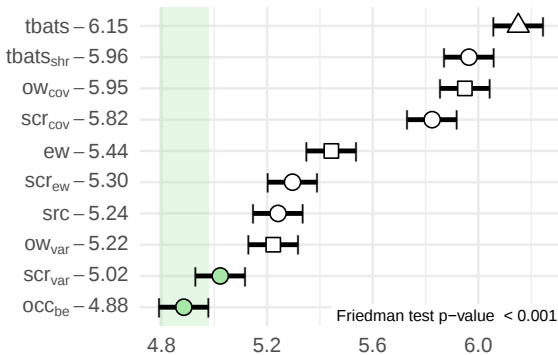
MCB Nemenyi test

R package `tsutils` (Kourentzes, 2023). The Friedman test p-value is reported in the lower right corner. The mean rank of each approach is shown to the right of its name. Statistical differences are indicated if the intervals of two forecast procedures do not overlap

MSE – All 23 ts – $h = 1$

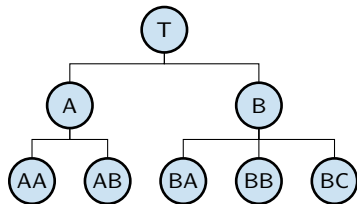


MSE – All 23 ts – $h = 1, \dots, 7$



Hierarchical, grouped and linearly constrained time series

Genuine hierarchical time series



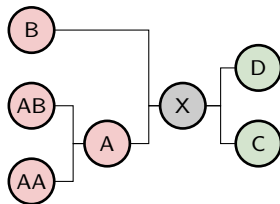
Constraints

$$T = A + B$$

$$A = AA + AB$$

$$B = BA + BB + BC$$

General linearly constrained time series



Constraints

$$X = A + B$$

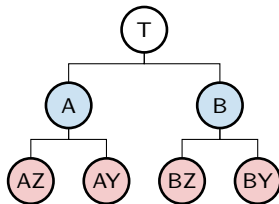
$$X = C + D$$

$$A = AA + AB$$

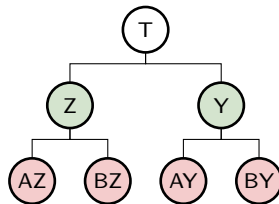
Grouped time series: two or more genuine hierarchies sharing the same top and bottom variables

T	A	B
Z	AZ	BZ
Y	AY	BY

=



+



Constraints

$$T = A + B$$

$$A = AZ + AY$$

$$B = BZ + BY$$

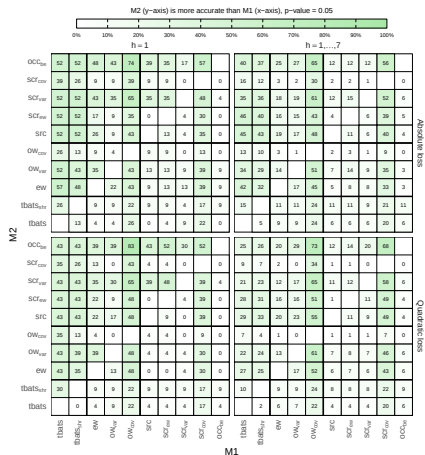
$$T = Z + Y$$

$$Z = AZ + BZ$$

$$Y = AY + BY$$

Diebold and Mariano (1995) test

Pairwise DM-test results evaluated using absolute loss (top panels) and quadratic loss (bottom panel) across different forecast horizons. The left panel corresponds to forecast horizon $h = 1$, while the right panel is for $h = 1, \dots, 7$. Each cell reports the percentage of series for which the p -value of the DM-test is below 0.05



Model Confidence Set

MCS results (10^4 bootstrap sample) evaluated using absolute loss (top panels) and quadratic loss (bottom panel) across different forecast horizons ($h = 1$ and $h = 1, \dots, 7$). Each cell reports the percentage of series for which that approach is in the Model Confidence Set across different thresholds ($\delta \in \{95\%, 90\%, 80\%\}$)

Approach	$h = 1$			$h = 1 : 7$		
	$\delta = 95\%$	$\delta = 90\%$	$\delta = 80\%$	$\delta = 95\%$	$\delta = 90\%$	$\delta = 80\%$
Absolute loss - All 23 time series						
<i>Base (incoherent forecasts) and single model reconciliation</i>						
tbats	56.5	56.5	52.2	78.3	69.6	56.5
tbats _{shr}	78.3	73.9	60.9	87.0	82.6	69.6
<i>Combination (incoherent forecasts)</i>						
ew	87.0	87.0	78.3	95.7	91.3	78.3
ow _{var}	95.7	95.7	82.6	95.7	91.3	82.6
ow _{cov}	73.9	69.6	60.9	73.9	65.2	43.5
<i>Coherent combination</i>						
src	91.3	91.3	87.0	95.7	95.7	87.0
scr _{ew}	91.3	91.3	87.0	95.7	91.3	78.3
scr _{var}	100.0	100.0	91.3	91.3	91.3	87.0
scr _{cov}	82.6	78.3	73.9	78.3	69.6	65.2
occ	100.0	100.0	95.7	95.7	95.7	87.0

Approach	$h = 1$			$h = 1 : 7$		
	$\delta = 95\%$	$\delta = 90\%$	$\delta = 80\%$	$\delta = 95\%$	$\delta = 90\%$	$\delta = 80\%$
Quadratic loss - All 23 time series						
<i>Base (incoherent forecasts) and single model reconciliation</i>						
tbats	65.2	65.2	60.9	91.3	73.9	73.9
tbats _{shr}	73.9	69.6	65.2	95.7	82.6	69.6
<i>Combination (incoherent forecasts)</i>						
ew	87.0	78.3	60.9	95.7	87.0	82.6
ow _{var}	100.0	82.6	78.3	91.3	87.0	82.6
ow _{cov}	78.3	69.6	52.2	78.3	60.9	34.8
<i>Coherent combination</i>						
src	95.7	91.3	82.6	91.3	91.3	91.3
scr _{ew}	95.7	87.0	73.9	91.3	91.3	91.3
scr _{var}	95.7	95.7	95.7	91.3	91.3	91.3
scr _{cov}	82.6	82.6	73.9	73.9	60.9	56.5
occ	100.0	95.7	95.7	91.3	91.3	91.3