Improving load forecasts in Italian bidding zones: A coherent combination approach

Daniele Girolimetto

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

Introduction

Many real-world forecasting scenarios involve data structures with **multiple variables** linked by **constraints** (Athanasopoulos *et al.*, 2024). In these cases, when **multiple forecasts** of the individual variables are available, a forecast combination-and-reconciliation approach, that we call *coherent forecast combination*, may be used to improve the accuracy of individual forecasts while ensuring that the final forecasts satisfy the constraints. For instance, when base, i.e., possibly incoherent multiple forecasts, are available for various components of a constrained time series, these forecasts can be combined and reconciled to produce coherent and more accurate predictions.

Single-task forecast combination $\longrightarrow$ *multiple experts*, `no coherence`
(Bates and Granger, 1969; Timmermann, 2006)

Local: one individual variable ↑

Global: n individual variables ↓

Forecast reconciliation $\longrightarrow$ *single expert* and `coherence`
(Stone *et al.*, 1942; Hyndman *et al.*, 2011)

Multi-task forecast combination $\longrightarrow$ *multiple experts*, `no coherence`
(On the top of Sun and Deng, 2004)

Sequential coherent combination
Optimal coherent combination
$\Big\}\rightarrow$ *multiple experts* and `coherence`

Main contribution: `arXiv.2412.03429` and `FoCo2`

Bates and Granger (1969): linear forecast combination

+

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

↓

Girolimetto and Di Fonzo (2024): optimal combined and coherent forecasts for multiple **linearly constrained** time series: new result that **unifies linear forecast reconciliation and combination** in a simultaneous and statistically justified way, improving accuracy and ensuring coherence of the forecasts

Optimal coherent forecast combination

Suppose we have base forecasts of the n individual variables of the target vector $\boldsymbol{y} \in \mathbb{R}^n$ s.t. $\boldsymbol{C}\boldsymbol{y} = \boldsymbol{0}_{(n_u\times 1)}$, made by $p \geq 2$ experts: $\widehat{\boldsymbol{y}}^1 \in \mathbb{R}^{n_1}, \;\dots,\; \widehat{\boldsymbol{y}}^p \in \mathbb{R}^{n_p}, 1 \leq n_j \leq n, j = 1, \dots, p$. Assuming unbiased base forecasts, we consider the model

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

Then, the linear relationship linking $\widehat{\boldsymbol{y}}$ and $\boldsymbol{y}$ can be expressed as

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

where $\boldsymbol{K} = \boldsymbol{L}(\boldsymbol{1}_p \otimes \boldsymbol{I}_n) \in \{0,1\}^{m\times np}$, $\boldsymbol{L} = \text{Diag}(\boldsymbol{L}_1, \dots, \boldsymbol{L}_j, \dots, \boldsymbol{L}_p)$ the $(m \times np)$ selection matrix, $\boldsymbol{L}_j \in \{0,1\}^{n_j\times n}$ selects the $n_j \leq n$ entries of $\boldsymbol{y}$ according to j-th expert, and $\boldsymbol{\varepsilon}$ is a zero-mean random vector with $(m \times m)$ covariance matrix $\boldsymbol{W} = E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^\top)$.

Theorem 1. Optimal linear coherent forecast combination $\widetilde{\boldsymbol{y}}^c$
The minimum mean square error (MMSE) linear coherent combined forecast vector $\widetilde{\boldsymbol{y}}^c$, obtained as solution to the linearly constrained quadratic program

$$\widetilde{\boldsymbol{y}}^c = \arg\min_{\boldsymbol{y}} (\widehat{\boldsymbol{y}} - \boldsymbol{K}\boldsymbol{y})^\top \boldsymbol{W}^{-1} (\widehat{\boldsymbol{y}} - \boldsymbol{K}\boldsymbol{y}) \qquad \text{s.t. } \boldsymbol{C}\boldsymbol{y} = \boldsymbol{0}_{(n_u\times 1)},$$

is given by

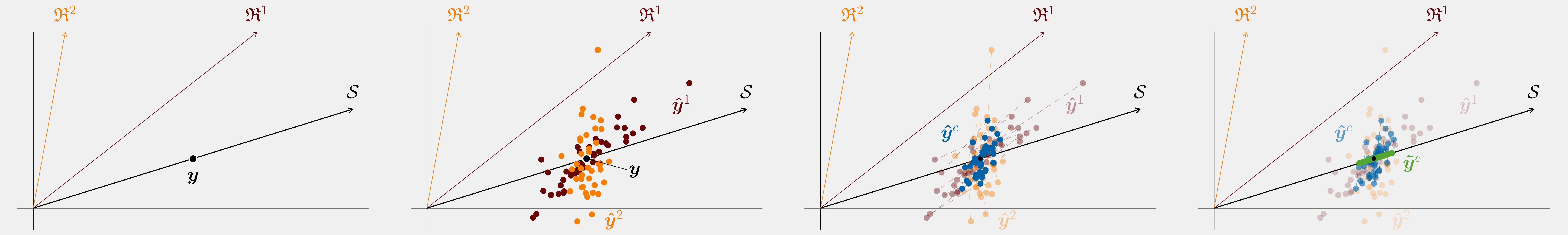
$$\widetilde{\boldsymbol{y}}^c = \boldsymbol{\Psi}^\top \widehat{\boldsymbol{y}} = \underbrace{\boldsymbol{M}}_{\text{projects } \widehat{\boldsymbol{y}}^c \to \mathcal{S} = \{\boldsymbol{y} \in \mathbb{R}^n \mid \boldsymbol{C}\boldsymbol{y} = \boldsymbol{0}_{(n_u\times 1)}\}} \overbrace{\boldsymbol{\Omega}^\top}^{\text{maps } \widehat{\boldsymbol{y}} \in \mathbb{R}^m \to \widehat{\boldsymbol{y}}^c \in \mathbb{R}^n} \widehat{\boldsymbol{y}},$$

with weight matrix $\boldsymbol{\Psi}^\top = \boldsymbol{M}\boldsymbol{\Omega}^\top \in \mathbb{R}^{n\times m}$, where

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

Corollary 1. Unbiasedness of $\widetilde{\boldsymbol{y}}^c$ and a property of its error covariance matrix
Denoting $\boldsymbol{\mu} = E(\boldsymbol{y})$, the coherent combined forecast vector $\widetilde{\boldsymbol{y}}^c$ is unbiased, i.e., $E(\widetilde{\boldsymbol{y}}^c) = \boldsymbol{\mu}$, with error covariance matrix $\widetilde{\boldsymbol{W}}_c = E\left[(\widetilde{\boldsymbol{y}}^c - \boldsymbol{y})(\widetilde{\boldsymbol{y}}^c - \boldsymbol{y})^\top\right] = \boldsymbol{M}\boldsymbol{W}_c$.
In addition, $\boldsymbol{L}_j\widetilde{\boldsymbol{W}}_c\boldsymbol{L}_j^\top \preceq \boldsymbol{L}_j\boldsymbol{W}_c\boldsymbol{L}_j^\top \preceq \boldsymbol{W}_j, \quad j = 1, \dots, p$.

Graphical visualisation of Theorem 1: 2 experts, n variables



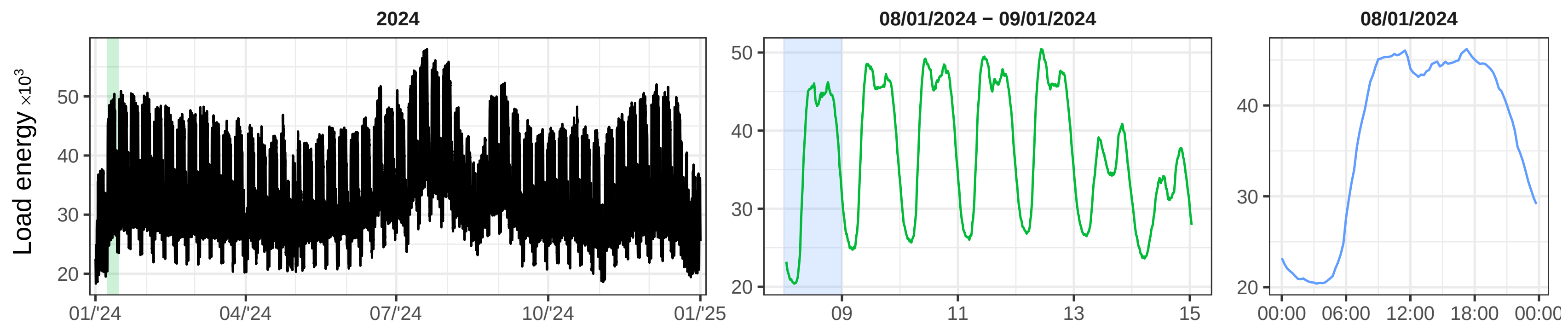
$\mathfrak{R}^1$ and $\mathfrak{R}^2$ show the most likely direction of deviations from the coherent subspace $\mathcal{S}$ for the 2 experts. The black dot $\boldsymbol{y}$ denotes the (unknown) target forecast.

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

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

Green points represent the unbiased MMSE linear coherent combined forecasts, $\widetilde{\boldsymbol{y}}^c = \boldsymbol{M}\boldsymbol{\Omega}^\top\widehat{\boldsymbol{y}}$, as an oblique projection of $\widehat{\boldsymbol{y}}^c$ on $\mathcal{S}$.

Italian energy load forecasting by Terna



- Among the various activities, Terna (Europe's largest independent electricity Transmission System Operator) currently **publishes** on its data portal very short-term load forecasts for the next day, at national level and disaggregated by 7 bidding zones
- **Forecasting experiment**: rolling forecast experiment with daily iterations (365 days as validation set and as test set) and 96-step ahead forecast horizons (15-min data)
- Coherency issue: The aggregated forecasts for the 7 Bidding Zones must match the forecasts for Italy

References

Athanasopoulos, G., Hyndman, R. J., Kourentzes, N. and Panagiotelis, A. (2024) Forecast reconciliation: A review. *International Journal of Forecasting* **40**(2), 430–456. **doi:** `10.1016/j.ijforecast.2023.10.010`

Bates, J. and Granger, C. W. (1969) The combination of forecasts. *Operations Research Quarterly* **20**(4), 451–468. **doi:** `10.1057/jors.1969.103`

Girolimetto, D. and Di Fonzo, T. (2024) Coherent forecast combination for linearly constrained multiple time series. *arXiv* pp. 1–61. **doi:** `10.48550/arXiv.2412.03429`

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. **doi:** `10.1016/j.csda.2011.03.006`

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. **doi:** `10.2307/2967664`

Sun, S.-L. and Deng, Z.-L. (2004) Multi-sensor optimal information fusion Kalman filter. *Automatica* **40**(6), 1017–1023. **doi:** `10.1016/j.automatica.2004.01.014`

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. **doi:** `10.1016/S1574-0706(05)01004-9`

Forecasting performance

Forecast and combination approaches (incoherent / coherent forecasts)

- *Base forecasts:* Terna and daily random walk (drw, $\widehat{y}_{i,t+h|t} = y_{i,t-96+h}$)
- Equal weights (ew)
- **Optimal single-task combination** (ow$_{\text{var}}$ and ow$_{\text{cov}}$)
- **Optimal combination** (occ) with *by-expert* **block-diag** shrunk error cov matrix

Results

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

	Country and 7 bidding zones									
App.	Italy	North	C-North	C-South	South	Calabria	Sicily	Sardinia	BTS	All
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$_{\text{var}}$	0.9930	0.9980	0.9909	0.9897	0.9943	0.9879	0.9663	0.9282	0.9791	0.9808
ow$_{\text{cov}}$	0.9863	0.9905	0.9847	0.9847	0.9930	0.9854	0.9676	0.9312	0.9765	0.9777
occ	**0.8973**	**0.8997**	**0.8969**	**0.8966**	**0.8952**	**0.8947**	**0.8936**	**0.8936**	**0.8958**	**0.8960**

- **AvgRelMAE**: When using the global approaches, either two-step or optimal, more accurate forecasts are obtained
- **Diebold-Mariano tests** for each 15-min forecast horizon: occ **significantly** outperforms Terna in $\sim 86\%$ of the cases; Terna never significantly improve w.r.t. occ.

R pkg availabe on CRAN!

FoCo²

WP arXiv

Contacts

danigiro.github.io

github.com/danigiro

@ daniele.girolimetto@unipd.it

`https://linktr.ee/danigir0`

Department of Statistical Sciences – Retreat, *19 September 2025*