Enhancements in cross-temporal forecast reconciliation, with an application to solar irradiance forecasts

Tommaso Di Fonzo\(^a\)\(^*,\) Daniele Girolimetto\(^a\)

\(^a\)Department of Statistical Sciences, University of Padua, Via C. Battisti 241, 35121 Padova (Italy)

Abstract

In recent works by Yang et al. (2017a,b), and Yagli et al. (2019), geographical, temporal, and sequential deterministic reconciliation of hierarchical photovoltaic (PV) power generation have been considered for a simulated PV dataset in California. In the first two cases, the reconciliations are carried out in spatial and temporal domains separately. To further improve forecasting accuracy, in the third case these two reconciliation approaches are sequentially applied. During the replication of the forecasting experiment, some issues emerged about non-negativity and coherence (in space and/or in time) of the sequentially reconciled forecasts. Furthermore, while the accuracy improvement of the considered approaches over the benchmark persistence forecasts is clearly visible at any data granularity, we argue that an even better performance may be obtained by a thorough exploitation of cross-temporal hierarchies. To this end, in this paper the cross-temporal point forecast reconciliation approach is applied to generate non-negative, fully coherent (both in space and time) forecasts. In particular, (i) some useful relationships between two-step, iterative and simultaneous cross-temporal reconciliation procedures are for the first time established, (ii) non-negativity issues of the final reconciled forecasts are discussed and correctly dealt with in a simple and effective way, and (iii) the most recent cross-temporal reconciliation approaches proposed in literature are adopted. The normalised Root Mean Square Error is used to measure forecasting accuracy, and a statistical multiple comparison procedure is performed to rank the approaches. Besides assuring full coherence, and non-negativity of the reconciled forecasts, the results show that for the considered dataset, cross-temporal forecast reconciliation significantly improves on the sequential procedures proposed by Yagli et al. (2019), at any cross-sectional level of the hierarchy and for any temporal granularity.

Keywords: Forecasting, Cross-temporal forecast reconciliation, Sequential and iterative approaches, Non-negative forecasts, Photovoltaic power generation

\(^*\)Corresponding author

Email address: difonzo@stat.unipd.it (Tommaso Di Fonzo)
1. Introduction

Traditional electricity relies heavily on fossil fuels such as coal and natural gas. Not only are they bad for the environment, but they are also limited resources. Net-zero emissions by 2050 are crucial to achieve the core Paris Agreement goals of a global average temperature rise of 1.5 degrees Celsius (United Nations, 2015), and this in turn can only be achieved if global greenhouse gas emissions are halved by the end of this decade (European Commission, 2019, United Nations, 2022). Solar power is one of the crucial production methods in the move to clean energy, and as economies of scale drive prices down, its importance will undoubtedly increase. The deployment of solar-power generation is causing total installed capacity to increase at a very high pace. Eurostat reports that the European Union added 18,224.8 MW of net capacity in 2020, compared to its 16,146.9 MW increase in 2019, registering a growth of 12.9%. At the end of 2020, the EU’s photovoltaic base stood at 136,136.6 MW, which is a 15% year-on-year increase (EurObserv'ER, 2022, p. 14).

Using solar resource as a stable source of energy is not an easy task. Estimating the solar energy potential is a key target to ensure its management in a reliable and efficient way for its integration into an electrical power grid. The prediction of solar irradiation despite its variability is particularly important as it is a precondition for (i) the management of the solar photovoltaic production through storage systems to reduce the impact of the intermittent nature of the solar resource, and (ii) the integration of the solar resource into a power grid in order to meet the local energy needs and to cope with the load fluctuations. Understanding of the need for short, mid or long term prediction (e.g., 1h, 6h or a day ahead forecasting) is growing as utilities and grid operators gain experience in dealing with solar-power sources. Increasing spatial and temporal resolution of the available forecasting models would enable grid operators to better forecast how much solar energy will be added to the grid. These efforts will improve the management of solar power’s variability and uncertainty, enabling its more reliable and cost-effective integration onto the grid.

Solar forecasting is a fast-growing sub-domain of energy forecasting (Yang et al., 2020, p. 20). We agree with the claim of (Yang et al., 2022, p. 7) that a “common misconception is that the novelty in solar forecasting should be solely revolved around forecasting methodology. Indeed, forecasting methodology is an important aspect, but it is never the only one”. Nevertheless, we think that a clear assessment of the available forecasting procedures may help in moving forward the frontier of knowledge of this fundamental topic, generating beneficial effects on the activities of practitioners.

A major goal for solar forecasting is to provide information on future photovoltaic (PV) power generation at different locations, time scales, and horizons to power system operators (Yang et al., 2022). In recent works by Yang et al. (2017a, b), and Yagli et al. (2019), geographical, temporal, and sequential deterministic reconciliation of hierarchical PV power generation have been considered for a simulated PV dataset in California. In the first two cases, the reconciliations are carried out in spatial and temporal domains separately. To further improve prediction accuracy, in the third case these two reconciliation approaches are sequentially applied. During the replication of the forecasting experiment some issues emerged about non-negativity and coherency (in space and/or in time) of the sequentially reconciled forecasts. Furthermore,
while the accuracy improvement of the considered approaches over the benchmark persistence forecasts is clearly visible at any data granularity, we think that an even better performance may be obtained by a thorough exploitation of cross-temporal hierarchies.

Cross-temporal forecast reconciliation is a rather recent sub-domain of the general theme ‘forecast reconciliation’ (Fliedner, 2001; Athanasopoulos et al., 2009; Hyndman et al., 2011). The idea of exploiting the aggregation relationships valid both in space (cross-sectional coherency), and for different time granularities (temporal coherency) to improve the forecast accuracy of base forecasts for a hierarchical time series, was discussed in the groundbreaking papers by Kourentzes and Athanasopoulos (2019), Yagli et al. (2019), Spiliotis et al. (2020), and Punia et al. (2020). The relevance of the topic has been further confirmed by (i) the ISF 2020 keynote speech by prof. Hyndman (2020), (ii) the entry ‘cross-temporal hierarchies’ in the encyclopedic review on the theory and practice of forecasting by Petropoulos et al. (2022), as well as (iii) the section entitled ‘cross-temporal hierarchies’ in a recent review paper on aggregation and hierarchical approaches for demand forecasting in supply chains (Babai et al., 2022). These important contributions paved the way for the findings of Di Fonzo and Girolimetto (2021), who showed the potentiality of a number of new cross-temporal reconciliation approaches, and discussed fundamental feasibility issues, offering insights to the practitioner wishing to evaluate the effort-to-benefit ratio of using this forecasting device.

In this paper, cross-temporal point forecast reconciliation is applied to generate non-negative, fully coherent (both in space and time) forecasts of PV generated power. In particular, (i) some useful relationships between two-step, iterative and simultaneous cross-temporal reconciliation procedures are for the first time established, (ii) non-negativity issues of the final reconciled forecasts are discussed and correctly dealt with in a simple and effective way, and (iii) the most recent cross-temporal reconciliation approaches proposed in literature are adopted. The iterative and simultaneous approaches by Di Fonzo and Girolimetto (2021), and the heuristic cross-temporal procedure proposed by Kourentzes and Athanasopoulos (2019) are applied to the base forecasts with forecast horizon of 1 day, of PV generated power at different time granularities (1 hour to 1 day), of a hierarchy consisting of 324 series along 3 levels. The normalised Root Mean Square Error is used to measure forecasting accuracy, and a statistical multiple comparison procedure is performed to rank the approaches.

The paper is organized as follows. The deterministic (point) cross-temporal forecast reconciliation framework is described in section 2 and in section 3 some useful connections between apparently different approaches are shown. The forecasting experiment of Yagli et al. (2019) is replicated and discussed in section 4 and the performance of the newly proposed forecasting approaches is presented in section 5. Conclusions follow in section 6.

---

3 The entry was written by N. Kourentzes.
4 All the forecast reconciliation procedures considered in Di Fonzo and Girolimetto (2021), and in this paper, are available in the R package FoReco (Girolimetto and Di Fonzo, 2022).
2. Cross-temporal point forecast reconciliation

2.1. Problem definition

To begin with, consider the very simple example of a two-level cross-sectional hierarchy, where the top variable \( X \) is equal to the sum of two bottom series \( W \) and \( Z \). Further, assume that the highest time frequency the variables are observed at is quarterly, which means that by simple non-overlapping temporal aggregation of quarterly time series, semi-annual and annual time series may be obtained as well. Figure 1 gives a visual representation of such cross-temporal hierarchy for a time cycle of 1 year.

While the left panel of figure 1 shows the complete cross-temporal hierarchy, consisting of all aggregated temporal granularities which may be defined starting from quarterly data (i.e., semi-annual, and annual), in the right panel is represented its reduced version, where only the highest (quarterly) and the lowest (annual) time granularities are respectively considered. The square boxes in the figure denote the nodes of a two-level cross-sectional (contemporaneous) hierarchy, while the circles denote the nodes of the temporal hierarchies.

In agreement with a standard notation in the temporal forecast reconciliation literature (Athanasopoulos et al., 2017, Yang et al., 2017b), the superscript \([k]\) denotes the temporal aggregation order for each time granularity, i.e. annual (\(k = 4\)), semi-annual (\(k = 2\)), and quarterly (\(k = 1\)). The cross-sectional hierarchy is described by the aggregation relationship \( X = W + Z \), which is valid for any temporal aggregation order \( k \in K = \{4, 2, 1\} \) (i.e., \( x^{[k]}_\tau = w^{[k]}_\tau + z^{[k]}_\tau \), \( \tau = 1, \ldots, 4/k \)). Assuming \( v \) alternatively equal to \( x, w, z \), the temporal hierarchies describing the relationships between different time granularities of a single time series, may be expressed as:

\[
\begin{align*}
v^{[4]}_1 &= v^{[2]}_1 + v^{[2]}_2 \quad &\text{the annual value is the sum of the two semi-annual values,} \\
v^{[2]}_1 &= v^{[1]}_1 + v^{[1]}_2 \quad &\text{the first half-year value is the sum of the first two quarters’ values,} \\
v^{[2]}_2 &= v^{[1]}_3 + v^{[1]}_4 \quad &\text{the second half-year value is the sum of the last two quarters’ values,} \\
\text{and thus} \\
v^{[1]}_1 &= v^{[1]}_1 + v^{[1]}_2 + v^{[1]}_3 + v^{[1]}_4 \quad &\text{the annual value is the sum of the four quarterly values.}
\end{align*}
\]

\[\text{Figure 1: Complete (left) and reduced (right) cross-temporal hierarchies for a quarterly two-level hierarchical time series.} \]

In this paper, we consider only genuine hierarchical/grouped time series, that share the same top- and bottom-level variables. The treatment of a general linearly constrained multiple time series is discussed in Di Fonzo and Girolimetto (2021).
All the relationships so far can be expressed in more compact form, using matrix notation. Denote by $y^{[k]}_\tau = \begin{bmatrix} x^{[k]}_\tau & w^{[k]}_\tau & z^{[k]}_\tau \end{bmatrix}'$ the $(3 \times 1)$ vector of the observations for temporal granularity $k \in K$ of the variables forming the cross-sectional hierarchy at time $\tau = 1, \ldots, 4/k$. The cross-sectional aggregation relationships can be described as follows:

$$x^{[k]}_\tau = C \begin{bmatrix} w^{[k]}_\tau \\ z^{[k]}_\tau \end{bmatrix}, \quad y^{[k]}_\tau = S \begin{bmatrix} w^{[k]}_\tau \\ z^{[k]}_\tau \end{bmatrix}, \quad U'y^{[k]}_\tau = 0, \quad \tau = 1, \ldots, \frac{4}{k}, \quad k \in K,$$

where $C$ is the cross-sectional aggregation matrix, $S$ is the cross-sectional summing matrix, and $U'$ is the zero-constraints matrix expressing the cross-sectional constraints in homogeneous form, respectively given by:

$$C = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad S = \begin{bmatrix} C \\ I_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad U' = [I_3 - C] = \begin{bmatrix} 1 & -1 & -1 \end{bmatrix},$$

with $I_l$ denoting the identity matrix of order $l$. The complete temporal aggregation relationships linking the values of a single variable (say $V$) at different time granularities, may in turn be expressed through matrices

$$K = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad R = \begin{bmatrix} K \\ I_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad Z' = [I_4 - K] = \begin{bmatrix} 1 & 0 & 0 & -1 & -1 & -1 \\ 0 & 1 & 0 & -1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \end{bmatrix},$$

where $K$ is the temporal aggregation matrix, $R$ is the temporal summing matrix, and $Z'$ is the temporal zero-constraints matrix expressing the temporal constraints in homogeneous form. It follows that:

$$\begin{bmatrix} v^{[4]}_1 \\ v^{[2]}_1 \\ v^{[1]}_1 \\ v^{[2]}_2 \\ v^{[1]}_2 \\ v^{[1]}_3 \\ v^{[4]}_4 \end{bmatrix} = K v^{[1]}, \quad v = \begin{bmatrix} v^{[1]}_1 \\ v^{[1]}_2 \\ v^{[1]}_3 \\ v^{[1]}_4 \end{bmatrix} = R v^{[1]}, \quad Z'v = 0_{(3 \times 1)}.$$

Reduced temporal hierarchies can be obtained by simply eliminating the appropriate row from matrix $K$.

To simultaneously consider cross-sectional and temporal aggregation relationships, all the nodes of the complete cross-temporal hierarchy in figure 1 can be expressed in terms of the quarterly time series $w^{[1]}_\tau$ and

---

5 This option is available in the R package FoReco (Girolimetto and Di Fonzo 2022).

6 This option is available in the R package FoReco (Girolimetto and Di Fonzo 2022).
$z_{\tau}^{[1]}$, $\tau = 1, \ldots, 4$, according to the structural representation:

$$y = Fb^{[1]},$$

that is

$$
\begin{bmatrix}
    x_1^{[1]} & \ldots & x_1^{[4]}
\end{bmatrix}
\begin{bmatrix}
    1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
    1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\
    0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\
    1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\
    0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
    0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
    0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
    0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
= \begin{bmatrix}
    w_1^{[1]} \\
    w_2^{[1]} \\
    w_3^{[1]} \\
    w_4^{[1]} \\
\end{bmatrix},
\begin{bmatrix}
    0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
\end{bmatrix}
\begin{bmatrix}
    1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
    1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\
    0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\
    0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\
\end{bmatrix}
\begin{bmatrix}
    x_1^{[1]} \\
    x_2^{[1]} \\
    x_3^{[1]} \\
\end{bmatrix},
$$

where $y = \begin{bmatrix} x' w' z' \end{bmatrix}'$ is the vector containing the data for all variables at any temporal granularity, $b^{[1]} = \begin{bmatrix} w^{[1]'} z^{[1]'} \end{bmatrix}'$ is the vector of the high-frequency bottom time series, and $F$ is the cross-temporal summing matrix mapping $b^{[1]}$ into $y$. Expression (1) is the natural extension of the cross-sectional structural representation firstly shown by Athanasopoulos et al. (2009). It relates the observations at the upper levels of both cross-sectional and temporal hierarchies, to the high-frequency bottom time series of the cross-sectional hierarchy, which are the ‘very’ bottom time series in a cross-temporal hierarchy (Di Fonzo and Girolimetto, 2021).

Besides the number of variables forming the cross sectional hierarchy ($n = 3$ in the above example), two crucial aspects affecting the dimension of matrix $F$ are (i) the temporal frequency of the highest-frequency granularity ($k = 1$), and (ii) the amount of temporal granularities taken into account in the temporal hierarchy. For example, if one is interested in coherently forecasting hourly time series within a day-cycle, the complete cross-temporal summing matrix $R$ defining all infra-day temporal granularities ($K = \{24, 12, 8, 6, 4, 3, 2, 1\}$) has dimension $(60 \times 24)$, and is equal to

$$R = \begin{bmatrix}
    I_{24} & I_2 \otimes I_{12} & I_3 \otimes I_8 & I_4 \otimes I_6 & I_6 \otimes I_4 & I_8 \otimes I_3 & I_{12} \otimes I_2 & I_{24}
\end{bmatrix}' .$$
Matrix $F$ is thus a large and sparse matrix\(^7\) of dimension $(60n \times 24n_b)$, where $n$ is the total number of series, and $n_b$ is the number of bottom time series in the cross-sectional hierarchy, respectively (in the above example, $n = 3$ and $n_b = 2$). Just to give an idea, the total number of variables in the dataset analyzed in this paper (see section \(\square\)) is $n = 324$, with $n_b = 318$ bottom time series, thus matrix $F$ has dimension $(19,440 \times 7,632)$. However, if the interest in forecasting at certain time granularities is low, this dimensionality issue may be mitigated by considering only part of the temporal granularities between the highest and lowest temporal frequencies.\(^8\) For example, if one considers only hourly and daily forecasts, the reduced $F$ is a $(8,100 \times 7,632)$ matrix, with a decrease of about 58% in the amount of matrix entries wrt its complete counterpart.

In this framework, by extending the seminal idea by Hyndman et al. (2011), a forecast reconciliation problem arises when, for the nodes of a cross-temporal hierarchy, a set of base forecasts - however obtained, and usually not aggregate consistent either in space and/or in time - are wished to be revised to fulfill the coherency relationships in space and time valid for the target data. The purpose is to improve the accuracy of the initial forecasts by combining forecasts at different aggregation levels in space and time, and by incorporating in the final forecasts the information given by cross-sectional and temporal constraints.

2.2. Notation

Suppose we want to forecast a $n$-variate high-frequency hierarchical time series $\{y_t^{[1]}\}_{t=1}^T$, with forecast horizon equal to the seasonal cycle $m$, (e.g., month per year, $m = 12$, quarter per year, $m = 4$, hour per day, $m = 24$), or a multiple thereof. Given a factor $k$ of $m$, we may consider a number of temporally aggregated versions of each component of $y_t^{[1]}$, given by the non-overlapping sums of $k$ successive values, each having seasonal period equal to $M_k = m/k$. To avoid ragged-edge data, we assume that the total number of observations involved in the non-overlapping aggregation is a multiple of $m$, and define $N$ the number of the lowest-frequency series observations, i.e. $N = T/m$. Let $K$ be the set of $p$ factors of $m$, in descending order, $K = \{k_p, k_{p-1}, \ldots, k_2, k_1\}$, where $k_p = m$ and $k_1 = 1$, and define $k^* = \sum_{j=2}^{p} k_j$.

Following Di Fonzo and Girolimetto (2021), denote $Y_{N+b} \equiv Y$ the $[n \times (k^* + m)]$ matrix of the target forecasts for any temporal granularity, with low-frequency temporal horizon $b$, given by:

$$Y = \begin{bmatrix}Y^{[m]} & Y^{[k_p-1]} & \cdots & Y^{[k_2]} & Y^{[k_1]} \end{bmatrix} = \begin{bmatrix}A & A^{[k_p-1]} & \cdots & A^{[k_2]} & A^{[k_1]} \\ B & B^{[k_p-1]} & \cdots & B^{[k_2]} & B^{[k_1]} \end{bmatrix},$$

where $m$ is the highest available sampling frequency per seasonal cycle (i.e., max. order of temporal aggregation). Each matrix $Y^{[k]} = \begin{bmatrix}A^{[k]} \\ B^{[k]} \end{bmatrix}$, $k \in K$, contains the order-$k$ temporal aggregates of the $n_a$ cross-sectional upper time series ($A^{[k]}$), and of the $n_b$ cross-sectional bottom time series ($B^{[k]}$), respectively,

\(\square\)Sparse matrices require less memory than dense matrices, and allow some computations to be more efficient (Paige and Sanders, 1982; Davis, 2006; Bates et al., 2022).

\(\square\)Possible losses in the forecasting accuracy of the reconciled forecasts according to reduced temporal hierarchies should however be evaluated. This issue is currently under study.
with \( n = n_a + n_b \). Accordingly, we define the matrix of base forecasts \( \hat{Y} \) as:

\[
\hat{Y} = \begin{bmatrix}
\hat{Y}^{[m]} & \hat{Y}^{[k_{p-1}]} & \ldots & \hat{Y}^{[k_2]} & \hat{Y}^{[1]}
\end{bmatrix} = \begin{bmatrix}
\hat{A}^{[m]} & \hat{A}^{[k_{p-1}]} & \ldots & \hat{A}^{[k_2]} & \hat{A}^{[1]}
\hat{B}^{[m]} & \hat{B}^{[k_{p-1}]} & \ldots & \hat{B}^{[k_2]} & \hat{B}^{[1]}
\end{bmatrix}.
\]

While the target forecasts are expected to be aggregate-consistent both in time and space, the base forecasts are in general cross-sectionally and/or temporally incoherent, that is:

\[
U'Y = 0_{[n_a \times (k^* + m)]} \quad \text{and} \quad Z'Y' = 0_{[k^* \times n]}, \quad \text{while} \quad U'\hat{Y} \neq 0_{[n_a \times (k^* + m)]} \quad \text{and/or} \quad Z'\hat{Y}' \neq 0_{[k^* \times n]},
\]

where \( U' = \begin{bmatrix} I_{n_a} - C \end{bmatrix} \) and \( Z' = \begin{bmatrix} I_{k^*} - K \end{bmatrix} \) are zero-constraints matrices associated to the cross-sectional and temporal constraints, respectively. Working with a single dimension (either sectional, or temporal), we may consider the multivariate generalization of the structural representations for cross-sectional (Athanasopoulos et al., 2009), and temporal (Athanasopoulos et al., 2017), hierarchies, respectively:

**Cross-sectional structural representation of \( k^* + m \) hierarchical time series**

\[
y^{[k]}_\tau = Sb^{[k]}_\tau, \quad \tau = 1, \ldots, \frac{m}{K}, \quad \rightarrow \quad Y^{[k]} = SB^{[k]} \quad k \in K,
\]

that is, in compact form

\[
Y = SB \quad \rightarrow \quad y = (S \otimes I_{k^*+m}) b,
\]

where \( y = \text{vec} \left( Y^{[1]} \right) \), and \( b = \text{vec} \left( B^{[1]} \right) \).

**Temporal hierarchies structural representation for \( n \) individual time series**

\[
a_i = RA^{[1]}_i, \quad i = 1, \ldots, n_a \quad \rightarrow \quad A = RA^{[1]}\nu,
\]

\[
b_j = RB^{[1]}_j, \quad j = 1, \ldots, n_b \quad \rightarrow \quad B = RB^{[1]}\nu,
\]

that is

\[
Y' = RY^{[1]}/ \quad \rightarrow \quad y = (I_n \otimes R) y^{[1]},
\]

where \( y^{[1]} = \text{vec} \left( Y^{[1]} \right) \), and \( R \) is the temporal summing matrix (Di Fonzo and Girolimetto, 2021).

2.3. *Cross-temporal bottom-up reconciliation*

Bottom-up is an old and classic approach in the forecast reconciliation literature (Dunn et al., 1976; Dangerfield and Morris, 1992). This approach simply consists in obtaining the upper-level series’ forecasts by summing-up the base forecasts of the bottom level series in the hierarchy. While its representation is rather straightforward when only a single dimension (either space or time) is involved, it is instructive considering how ‘bottom-up’ works with cross-temporal hierarchies. The cross-temporal bottom-up (ct(bu)) reconciliation of \( k^* + m \) hierarchical time series’ base forecasts at different time granularities, may be represented as:

\[
\text{vec} \left( \tilde{Y}_{ct(bu)} \right) = F \text{vec} \left( B^{[1]}/ \right) \leftrightarrow \tilde{y}_{ct(bu)} = F \tilde{b}^{[1]},
\]

where \( F = S \otimes R \) is the cross-temporal summing matrix, with \( \otimes \) denoting the Kronecker product, and \( \tilde{b}^{[1]} = \text{vec} \left( B^{[1]}/ \right) \) is the vector containing the base forecasts of the high-frequency bottom time series. In
Appendix A it is shown that the cross-temporal bottom-up reconciliation can be thought of as a two-step sequential reconciliation approach, where either cross-sectional reconciliation of the high-frequency bottom time series base forecasts is followed by temporal reconciliation, or vice-versa. This observation opens the way to ‘partly bottom-up’ cross-temporal reconciliation approaches, where forecasts of the \( n \) time series for different time granularities, and aggregation coherent only along a single dimension, are subsequently cross-temporally reconciled via simple bottom-up according to the other dimension. We call these cross-temporal forecast reconciliation approaches either \( ct(rec_{te}, bu_{cs}) \), or \( ct(rec_{cs}, bu_{te}) \), where ‘\( rec_{te} \)’ and ‘\( rec_{cs} \)’ denote a generic forecast reconciliation approach in time and in space, respectively.

2.4. Regression-based cross-temporal reconciliation

Let us consider the multivariate regression model

\[
\hat{\mathbf{Y}} = \mathbf{Y} + \mathbf{E},
\]

where the involved matrices have each dimension \( [n \times (k^* + m)] \) and contain, respectively, the base (\( \hat{\mathbf{Y}} \)) and the target forecasts (\( \mathbf{Y} \)), and the coherency errors (\( \mathbf{E} \)) for the \( n \) component variables of the hierarchical time series of interest. Consider now the vectorized version of the model, that is

\[
\text{vec}\left(\hat{\mathbf{Y}}'\right) = \text{vec}\left(\mathbf{Y}'\right) + \text{vec}\left(\mathbf{E}'\right) \iff \hat{\mathbf{y}} = \mathbf{y} + \eta,
\]

where \( \eta = \text{vec}\left(\mathbf{E}'\right) \) is the cross-temporal reconciliation error with zero mean and p.d. covariance matrix \( \Omega_{ct} \).

Assuming \( \Omega_{ct} \) known, the optimal combination reconciled forecasts \( \tilde{\mathbf{y}}_{oct} = \text{vec}\left(\hat{\mathbf{Y}}'_{oct}\right) \) are found by linearly constrained minimization of the generalized least squares (GLS) objective function

\[
(y - \hat{y})' \Omega_{ct}^{-1} (y - \hat{y}) \quad \text{s.t. } \mathbf{H}' \mathbf{y} = 0,
\]

where

\[
\mathbf{H}' \mathbf{y} = \begin{bmatrix} \mathbf{U}^* \\ \mathbf{I}_n \otimes \mathbf{Z} \end{bmatrix} \mathbf{y} = 0_{[n(k^*+m)x1]}
\]

is a full row-rank cross-temporal zero-constraint matrix, with \( \mathbf{U}^* = [0_{(n_k \times n_k^*)} \mathbf{I}_m \otimes \mathbf{U}' \mathbf{P}] \), and \( \mathbf{P} \) is the \([n(k^* + m) \times n(k^* + m)]\) commutation matrix, such that \( \mathbf{P}\text{vec}(\mathbf{Y}) = \mathbf{y} \) (Di Fonzo and Girolimetto 2021).

The cross-temporally reconciled forecasts according to the projection approach solution (Byron 1978; see also van Erven and Cugliari 2015; Wickramasuriya et al. 2019; Panagiotelis et al. 2021; Di Fonzo and Girolimetto 2021) are given by

\[
\tilde{\mathbf{y}}_{oct} = \left[\mathbf{I}_{n(k^*+m)} - \Omega_{ct} \mathbf{H}' (\mathbf{H} \Omega_{ct} \mathbf{H})^{-1} \mathbf{H}' \right] \hat{\mathbf{y}} = \mathbf{M}_{ct} \hat{\mathbf{y}}.
\]

Alternatively, they may be obtained according to the structural approach developed by Hyndman et al. (2011) for the cross-sectional framework:

\[
\hat{\mathbf{y}} = \mathbf{F} \hat{\mathbf{b}} + \eta,
\]

where \( \hat{\mathbf{b}} = E[b^{[1]}|\mathcal{I}_T] \) is the mean of the high-frequency bottom-level values conditional to \( \mathcal{I}_T \), the available
information up to time $T$, $b^{[1]} = \text{vec} \left( B^{[1]} \right)$. The structural approach forecast reconciliation formula is

$$\hat{y}_{\text{oct}} = F (F' \Omega_{ct}^{-1} F)^{-1} F' \Omega_{ct}^{-1} \hat{y} = FG_{ct}\hat{y},$$

(6)

with $G_{ct} = (F' \Omega_{ct}^{-1} F)^{-1} F' \Omega_{ct}^{-1}$. Di Fonzo and Girolimetto (2021) considered the following approximations for the cross-temporal covariance matrix (‘oct’ stands for ‘optimal cross-temporal’):

- $\text{oct}(\text{ols})$ - identity: $\Omega_{ct} = I_{n(k^*+m)}$
- $\text{oct}(\text{struc})$ - structural: $\Omega_{ct} = \text{diag}(F \Omega_{1n})$
- $\text{oct}(\text{wls})$ - series variance scaling: $\Omega_{ct} = \hat{\Omega}_{ct,\text{wls}}$, that is a straightforward extension of the series variance scaling matrix presented by Athanasopoulos et al. (2017) in the temporal framework
- $\text{oct}(\text{bdshr})$ - block-diagonal shrunk cross-covariance scaling: $\Omega_{ct} = P \hat{W}_{ct,\text{shr}} P'$
- $\text{oct}(\text{bdsm})$ - block-diagonal cross-covariance scaling: $\Omega_{ct} = P \hat{W}_{ct,\text{sam}} P'$
- $\text{oct}(\text{shr})$ - MinT-shr: $\Omega_{ct} = \lambda \hat{\Omega}_{ct, D} + (1 - \lambda) \hat{\Omega}_{ct}$
- $\text{oct}(\text{sam})$ - MinT-sam: $\Omega_{ct} = \hat{\Omega}_{ct}$

where the symbol $\odot$ denotes the Hadamard product, $\lambda$ is an estimated shrinkage coefficient (Ledoit and Wolf, 2004), $\hat{\Omega}_{ct, D} = I_{n(k^*+m)} \odot \Omega_{ct}$, and $\hat{\Omega}_{ct}$ is the covariance matrix of the cross-temporal one-step ahead in-sample forecast errors. The cross-sectional point forecast reconciliation formula is obtained by assuming $m = 1$ (which implies $k^* = 0$, and $\Omega_{ct} = W$ is an $(n \times n)$ p.d. matrix):

$$\hat{y} = \left[ I_n - WU (U'WU)^{-1} U' \right] \hat{y} = M_{cs}\hat{y},$$

or, equivalently (Hyndman et al., 2011, Hyndman et al., 2016, Wickramasuriya et al., 2019),

$$\hat{y} = S (S'W^{-1}S)^{-1} S'W^{-1}\hat{y} = SG_{cs}\hat{y},$$

with $G_{cs} = (S'W^{-1}S)^{-1} S'W^{-1}$. The reconciled forecasts through temporal hierarchies for a single time series (Athanasopoulos et al., 2017) are in turn obtained by setting $n = 1$ (i.e., $n_a = 0$ and $n_b = 1$, and $\Omega_{ct} = \Omega$ is a $(k^* + m \times k^* + m)$ p.d. matrix):

$$\hat{y} = \left[ I_{k^*+m} - \Omega Z (Z'\Omega Z)^{-1} Z' \right] \hat{y} = M_{te}\hat{y},$$

or, equivalently,

$$\hat{y} = R (R'\Omega^{-1}R)^{-1} R'\Omega^{-1}\hat{y} = RG_{te}\hat{y},$$

with $G_{te} = (R'\Omega^{-1}R)^{-1} R'\Omega^{-1}$. Table 1 presents some approximations for the cross-sectional and the temporal covariance matrices. Other alternatives for temporal reconciliation, exploiting possible information in the residuals’ autocorrelation, can be found in Nystrup et al. (2020) and Di Fonzo and Girolimetto (2021).

2.5. Heuristic and iterative cross-temporal reconciliation

Kourentzes and Athanasopoulos (2019) proposed an ensemble forecasting procedure (denoted KA), that exploits the simple averaging of different forecasts. It consists in the following steps (for further details, see
Step 1: compute the temporally reconciled forecasts (Di Fonzo and Girolimetto, 2021) and temporal covariance matrix (Athanasopoulos et al., 2017, Di Fonzo and Girolimetto, 2021).

\[
\begin{array}{|c|c|c|}
\hline
\text{Cross-sectional framework} & \text{Temporal framework} \\
\hline
\text{identity} & \text{te}(ols): \Omega = \mathbf{I}_{k^*+m} & \\
\text{structural} & \text{te}(struc): \Omega = \text{diag}(\mathbf{R}_m) \\
\text{series variance} & \text{te(wls):} \Omega = \tilde{\Omega}_{wls} \\
\text{MinT-shr} & \text{te(shr):} \Omega = \hat{\lambda} \tilde{\Omega} + (1 - \hat{\lambda}) \hat{\Omega} \\
\text{MinT-sam} & \text{te(sam):} \Omega = \hat{\Omega} \\
\hline
\end{array}
\]

* $\tilde{\Omega}$ is the covariance matrix of the cross-sectional (temporal) one-step ahead in-sample forecast errors, $\tilde{\Omega}_{wls}$ is a diagonal matrix which contains estimates of the in-sample one-step-ahead error variances across each level* (Athanasopoulos et al., 2017, p. 64), and $\hat{\Omega}_D = \mathbf{I}_{k^*+m} \odot \hat{\Omega}$.

**KA-Step 1**: compute the temporally reconciled forecasts for each variable $i \in \{1, \ldots, n\}$, and arrange them in the $[n \times (k + m)]$ matrix $\tilde{Y}_{te}$.

**KA-Step 2**: starting from $\tilde{Y}_{te}$, compute the time-by-time cross-sectional reconciled forecasts for all the temporal aggregation levels ($\tilde{Y}_{cs}$), and collect all the $(n \times n)$ projection matrices used to reconcile forecasts of $k$-level temporally aggregated time series, $\mathbf{M}^{[k]}$, $k \in K$;

**KA-Step 3**: transform the step 1 forecasts once more, by computing time-by-time cross-sectional reconciled forecasts for all temporal aggregation levels using the $(n \times n)$ matrix $\overline{\mathbf{M}}$, given by the average of the matrices $\mathbf{M}^{[k]}$.

\[
\tilde{Y}_{\text{KA}} = \left( \frac{1}{p} \sum_{k \in K} \mathbf{M}^{[k]} \right) \tilde{Y}_{te} = \overline{\mathbf{M}} \tilde{Y}_{te}.
\]

**Di Fonzo and Girolimetto (2021)** presented an iterative approach, that produces cross-temporally reconciled forecasts by alternating forecast reconciliation along one dimension (cross-sectional or temporal), based on the first two steps of the KA approach. The iteration $j \geq 1$ can be described as follows:

**Step 1**: compute the temporally reconciled forecasts ($\tilde{Y}_{cs}^{(j)}$) for each variable $i \in \{1, \ldots, n\}$ of $\tilde{Y}_{cs}^{(j-1)}$;

**Step 2**: compute the time-by-time cross-sectional reconciled forecasts ($\tilde{Y}_{te}^{(j)}$) for all the temporal aggregation levels of $\tilde{Y}_{te}^{(j)}$.

At $j = 0$, the starting values are given by $\tilde{Y}_{cs}^{(0)} = \tilde{Y}$, and the iterates end when the entries of matrix $\mathbf{D}_{te} = \mathbf{Z} \tilde{Y}_{cs}^{(j)}$, containing all the temporal discrepancies, are small enough according to a suitable convergence criterion.\footnote{Denoting $d_{te}$ a scalar measure of the gross temporal discrepancies contained in the matrix $\mathbf{D}_{te}$, the convergence is achieved when $d_{te} \leq \delta$, where $\delta$ is a positive tolerance value. Di Fonzo and Girolimetto (2021) propose to use $d_{te} = \|\mathbf{D}_{te}\|_1$, where $\|X\|_1 = \sum_{i,j} |x_{i,j}|$. When the number of series and the time granularities increase, this criterion may be too demanding, resulting in more iterations than necessary to converge. An alternative choice, available in ForReco (Girolimetto and Di Fonzo, 2022), and used in this paper, is setting $d_{te} = \|\mathbf{D}_{te}\|_\infty$, where $\|X\|_\infty = \max |x_{i,j}|$.}

The order of the dimensions to be reconciled in the two steps within each iteration is arbitrary (on this point, see Di Fonzo and Girolimetto (2021): in the description above, temporal-then-cross-sectional reconciliation is iteratively performed (ite($\text{rec}_{te}, \text{rec}_{cs}$)), otherwise the order may be reversed, getting cross-sectional-then-temporal reconciliation (ite($\text{rec}_{cs}, \text{rec}_{te}$)).
3. Cross-temporal coherency of sequential approaches and some remarkable equivalences

In order to exploit both cross-sectional and temporal hierarchies, Yagli et al. (2019) consider two sequential reconciliation approaches: Spatial-then-Temporal-Reconciliation (STR), and Temporal-then-Spatial-Reconciliation (TSR). In the former case, cross-sectional reconciliation of the base forecasts is performed first for any temporal granularity, followed by the temporal reconciliation of the individual series’ forecasts. In the latter case, the order of application of the two reconciliation approaches is reversed, temporal reconciliation being performed first, followed by cross-sectional reconciliation. In general, one would expect that the forecasts obtained this way are different (i.e., $\tilde{Y}_{TSR} \neq \tilde{Y}_{TSR}$), and either cross-sectionally (TSR) or temporally (STR), but not cross-temporally, reconciled:

$$U'\tilde{Y}_{TSR} = 0_{[n_a \times (k^* + m)]}, \quad Z'\tilde{Y}_{TSR} \neq 0_{(k^* \times n)}, \quad U'\tilde{Y}_{STR} \neq 0_{[n_a \times (k^* + m)]}, \quad Z'\tilde{Y}_{STR} = 0_{(k^* \times n)}.$$

At this regard, we have found an interesting result, here shown as Theorem 1, according to which, if both covariance matrices used in either steps are constant across levels and time granularities, the final result (i) does not depend on the order of application of the uni-dimensional reconciliation phases, and (ii) is equivalent to that obtained through an optimal combination approach using a separable covariance matrix, i.e. with a Kronecker product structure (Genton, 2007; Werner et al., 2008; Velu and Herman, 2017).

**Theorem 1.** Let $W[k], k \in K,$ be the cross-sectional hierarchy error covariance matrix, and $\Omega_i, i = 1, \ldots, n,$ the $i$-th series temporal hierarchy error covariance matrix. If $W[k] = W, \forall k \in K,$ and $\Omega_i = \Omega, \forall i = 1, \ldots, n,$ then:

1. the iterative procedure reduces to a single (two-step) iteration to obtain cross-temporal reconciled forecasts. Furthermore,

   $$\tilde{Y}_{tcs} = \tilde{Y}_{cst} = \tilde{Y}_{seq},$$

   where $\tilde{Y}_{tcs} (\tilde{Y}_{cst})$ is the $[n \times (k^* + m)]$ matrix of the temporal-then-cross-sectional (cross-sectional-then-temporal) reconciled forecasts;

2. Denoting $\tilde{Y}_{oct}$ the optimal (in least squares sense) combination cross-temporal reconciliation approach with cross-temporal covariance matrix $\Omega_{ct} = W \otimes \Omega,$ it is:

   $$\tilde{Y}_{seq} = \tilde{Y}_{oct}.$$

**Proof.** Appendix B

It is worth noting that, when only constant error covariance matrices are involved in both unidimensional steps, the iterative cross-temporal reconciliation approach proposed by Di Fonzo and Girolimetto (2021) reduces to the sequential procedure proposed by Yagli et al. (2019), and gives a unique result, $\hat{Y}_{seq}$, which does not depend on the order of application of the reconciliation phases\(^{10}\). In addition, it does exist a simple simultaneous, optimal (in least-squares sense) reconciliation approach equivalent to the above sequential

---

\(^{10}\)It is therefore surprising that in Yagli et al. (2019), where only constant matrices are used in the reconciliation approaches, the results for the STR procedures are different from those for the TSR counterparts. We guess this depends on the way the unidimensional reconciled forecasts are computed, as we deduced from the R scripts made available by Yang et al. (2017b).
forecast reconciliation procedure. Finally, it should be noted that under the constant covariance matrices assumption of the theorem, the heuristic cross-temporal reconciliation approach by Kourentzes and Athanasopoulos (2019) is equivalent to a sequential approach, since in this case the final averaging phase is not needed (more details can be found in Appendix B after the proof of the theorem). All these results can be summarized as follows:

\[
\begin{align*}
W[k] &= W, \forall k \in K \\
\Omega_t &= \Omega, i = 1, \ldots, n \\
\Omega_{ct} &= W \otimes \Omega
\end{align*}
\]

Using the covariance matrices in Table 1, we may thus establish some useful equivalences between the cross-temporal reconciliation approaches considered in Di Fonzo and Girolimetto (2021), and identify new ones:

1. oct(ols): \( \Omega_{ct} = I_{n(k^*+m)} \)
   equivalent to seq/KA/ite(ols_{te}, ols_{cs}), and seq/KA/ite(ols_{cs}, ols_{te}), with \( W = I_n, \Omega = I_{k^*+m} \).

2. oct(struct): \( \Omega_{ct} = \text{diag}(F1_{mn}) \)
   equivalent to seq/KA/ite(struct_{te}, struct_{cs}), and seq/KA/ite(struct_{cs}, struct_{te}), with \( W = \text{diag}(S1_{nk}), \Omega = \text{diag}(R1_m) \).

3. oct(ols_{cs}, struct_{te}): \( \Omega_{ct} = \text{diag}([I_n \otimes R1_{mn}] \cdot 1_{mn}) \)
   equivalent to seq/KA/ite(ols_{cs}, struct_{te}), with \( W = I_n, \Omega = \text{diag}(R1_m) \).

4. oct(struct_{cs}, ols_{te}): \( \Omega_{ct} = \text{diag}((S \otimes I_{k^*+m}) \cdot 1_{mn}) \)
   equivalent to seq/KA/ite(struct_{cs}, ols_{te}), with \( W = \text{diag}(S1_{nk}), \Omega = I_{k^*+m} \).

All the procedures considered so far make use of very simple covariance matrices, not using the information coming from the in-sample forecast errors. At this regard, another interesting result is that the iterative reconciliation procedure where both cross-sectional and temporal reconciliation are performed using diagonal covariance matrices, computed using the one-step-ahead in-sample forecast errors (i.e., cs(wls) and te(wls)), ‘converges’ to the optimal combination reconciliation approach oct(wls) \(^{[2]}\) \(^{12}\) Di Fonzo and Girolimetto (2021), irrespective of the order of application of the unidimensional reconciliation steps. In other terms:

\[
\tilde{Y}_{\text{ite}(wls_{te}, wls_{cs})} \simeq \tilde{Y}_{\text{ite}(wls_{cs}, wls_{te})} \simeq \tilde{Y}_{\text{oct}(wls)},
\]

where the quality of the approximation solely depends on the convergence criterion: the lower the tolerance value \( \delta \) (see section 2.5), the better the approximation will be. An empirical support to this result is given by figure 2, showing the Frobenius norm of the difference between the matrices of the reconciled forecasts using ite(wls_{te}, wls_{cs}) with decreasing tolerance values \( \delta \), and oct(wls), in the 350 replications of the forecasting experiment described in section 4.

\( ^{11}\) For the time being, we only have found strong empirical evidence of the correctness of this result, as shown in figure 2 and are currently involved in finding a general proof.

\( ^{12}\) The Frobenius norm of a real-valued matrix \( X \) is defined as the square root of the sum of the squares of its elements (Golub and Van Loan, 2013, p. 71): \( ||X||_F = \sqrt{\sum_{i,j} x_{i,j}^2} \)
In summary, when the considered diagonal covariance matrices are used in cross-sectional and temporal reconciliation phases, the iterative cross-temporal reconciliation approach is equivalent to a specific optimal combination cross-temporal approach. We argue that in practical application it is convenient to adopt the optimal combination version of a cross-temporal procedure, rather than its iterative counterpart, because the numerical results have better accuracy, and the computing time is often lower, mostly if the programming codes exploit sparse matrices tools (Davis 2006).

4. Replication and assessment of the forecasting experiment of Yagli et al. (2019)

The dataset used in this study, called PV324, is the same used by Yang et al. (2017a,b), and Yagli et al. (2019). It refers to 318 simulated PV plants in California, whose hourly irradiation data are organized in three levels (figure 3):

- \( L_0 \): 1 time series for the Independent System Operator (ISO), given by the sum of the 318 plant series;
- \( L_1 \): 5 time series for the Transmission Zones (TZ), each given by the sum of 27, 73, 101, 86, and 31 plants, respectively;
- \( L_2 \): 318 bottom time series at plant level (P).

Figure 3: PV324 hierarchy
Following Yang et al. (2017b) and Yagli et al. (2019), we perform a forecasting experiment with a fixed length window of 14 days (i.e., 336 hours), forecast horizon of two days, and forecasting evaluation taking into account only the day-2 forecasts. These settings are coherent with the forecast operational submission requirements of CAISO, the public corporation managing power grid operations in California (Makarov et al., 2011, Kleissl, 2013). For the 318 hourly time series, numerical weather prediction (NWP) forecasts generated by 3TIER (STIER, 2010) are used as base forecasts. All the remaining base forecasts, for the six $L_0$ and $L_1$ time series at any time granularity $k \in K$, and for the $L_2$ 2-3-4-6-8-12-24 hours time series, are computed using the automatic ETS forecasting procedure of the R-package forecast (Hyndman et al., 2021), not controlling for possible negative forecasts. Furthermore, day-ahead persistence is used as the reference model (PERS):

$$y^{[1]}_{T+h|T,PERS} = y^{[1]}_{T+h-48}.$$ 

Given the lead time of 48h, the day-ahead persistence takes the measurements made at day -2 as the forecasts for the operating day (Yagli et al., 2019, p. 394). Benchmark forecasts at any level of the spatial hierarchy and for any temporal granularity are obtained through cross-temporal bottom-up of the 318 hourly bottom time series, that is:

$$\tilde{y}_{PERS_{bn}} = F \hat{b}^{[1]}_{PERS} = \text{vec} \left( \tilde{Y}'_{PERS_{bn}} \right).$$

It is worth noting that the benchmark forecasts are always non-negative, and both spatially and temporally coherent. Furthermore, these important properties are valid also for the forecasts obtained by cross-temporal bottom-up reconciliation of the 318 hourly bottom time series’ NWP forecasts 3TIER:

$$\tilde{y}_{3TIER_{bn}} = F \hat{b}^{[1]}_{3TIER} = \text{vec} \left( \tilde{Y}'_{3TIER_{bn}} \right).$$

In light of the results provided in section 3, the eight STR and TSR sequential reconciliations proposed by Yagli et al. (2019) reduce to the following four approaches: oct($ols$), oct($ols_{cs}$, $struc_{te}$), oct($struc_{cs}$, $ols_{te}$), oct($struc$). In addition, Yagli et al. (2019) consider other four Temporal-then-Spatial-Reconciliation approaches, called TSR$_{L_2}$, where the temporal reconciliation is applied only to the 318 plant level series’ base forecasts. In this case, although constant matrices are used in either reconciliation steps, theorem 1 no longer holds, and the obtained forecasts are temporally incoherent, as we show in the following. In order to distinguish these approaches from the conventional sequential techniques, we call them te$_{ols} + cs_{ols}$, te$_{struc} + cs_{ols}$, te$_{ols_2} + cs_{struc}$, te$_{struc_2} + cs_{ols}$, respectively.

### 4.1. Non-negativity and aggregation consistency issues

Unlike PERS$_{bn}$ and 3TIER$_{bn}$, the other approaches considered by Yagli et al. (2019) produce some negative reconciled forecasts. Looking at table 2 we note that negative hourly forecasts have been produced in all 350 replications of the forecasting experiment, and there are some cases (in a range within 4 and 11 replications out of 350) where negative daily forecasts are produced as well. Furthermore, the base forecasts are generally cross-temporally incoherent (i.e., in space and/or in time), and the sequential TSR$_{L_2}$

---

13The relatively low number of negative hourly base forecasts, is due to the fact that the hourly base forecasts of the 318 bottom time series are NWP 3TIER forecasts, that are always non-negative. Negative values are instead present in the ETS base forecasts for the aggregated series. Details can be found in the on-line appendix.
approaches proposed by Yagli et al. (2019) are temporally incoherent. This can be visually appreciated from figure 4 showing the boxplots from the distribution of the gross cross-sectional and temporal discrepancies registered in the 350 replications of the forecasting experiment, computed as (Di Fonzo and Girolimetto 2021):

Cross-sectional gross discrepancy: \( d_{cs} = \| U' \hat{Y} \|_1 \)

Temporal gross discrepancy: \( d_{te} = \| Z' \hat{Y}' \|_1 \)

where \( \| X \|_1 = \sum_{i,j} |x_{i,j}| \). For truly cross-temporally reconciled forecasts, neither cross-sectional nor temporal discrepancies are present, that is \( d_{cs} = d_{te} = 0 \).

Table 2: Summary informations on the negative forecasts produced by the procedures considered by Yagli et al. (2019) in the forecasting experiment. Replications with at least a negative forecast (# rep), number of series (# series) with at least a negative forecast in a single replication (min and max), and min and max negative values found in all replications (values).

| Approach | Hourly forecasts | Daily forecasts |
|----------|-----------------|----------------|
| base\(^+\) | 350 4 6 -15.617 -0.000 | 11 0 35 -51.205 -0.006 |
| oct(ols) | 350 109 324 -69.165 -0.000 | 11 0 60 -29.615 -0.001 |
| oct(ols\(_c\), struc\(_c\)) | 350 137 324 -20.739 -0.000 | 11 0 33 -10.363 -0.000 |
| oct(struc) | 350 91 324 -17.961 -0.000 | 6 0 28 -8.915 -0.008 |
| te(ols\(_2\)) + cs(ols)\(^*\) | 350 89 324 -14.292 -0.000 | 4 0 11 -1.020 -0.026 |
| te(struc\(_2\)) + cs(ols)\(^*\) | 350 36 324 -12.994 -0.000 | 10 0 71 -14.791 -0.005 |
| te(struc\(_2\)) + cs(struc)\(^*\) | 350 144 324 -7.526 -0.000 | 10 0 45 -12.969 -0.005 |
| te(struc\(_2\)) + cs(struc)\(^*\) | 350 124 324 -7.642 -0.000 | 5 0 31 -7.042 -0.002 |

\(^*\) The approach produces spatial and temporal incoherent forecasts.

In this case, forecast reconciliation may thus generate physically unreasonable values. Furthermore, if coherency is wished, the apparently innocuous practice of setting possible negative forecasts to zero is not advisable, since incoherence in sectional and/or time dimensions would be produced. To overcome the above limitations, in section 5 we propose a simple operational strategy, able to generate fully reconciled non-negative forecasts.
4.2. Forecast evaluation

Following Yagli et al., 2019, the accuracy of the considered approaches is measured in terms of normalised Root Mean Square Error (nRMSE), and Forecast Skill score:

\[
\text{nRMSE}^{[k]}_{i,j} = \sqrt{\frac{1}{L} \sum_{l=1}^{L} \left( \hat{y}_{i,j,l}^{[k]} - y_{i,l}^{[k]} \right)^2}, \quad \text{Forecast Skill}^{[k]}_{i,j} = 1 - \frac{\text{nRMSE}^{[k]}_{i,j}}{\text{nRMSE}^{[0]}_{i,0}},
\]

where \( i = 1, \ldots, n \), denotes the series, \( k \in K \), \( j = 0, \ldots, J \), denotes the forecasting approach (\( j = 0 \) for the reference model PERS), and \( L = nrep \cdot \frac{m}{k} \), where \( nrep = 350 \) is the number of the forecasting experiment replications. Forecast Skill can be either negative (approach is worse than the reference model) or positive (approach is better than the reference model). From table 3 it appears that on average all the considered forecasting approaches improve on the benchmark PERS, in a range between 9.2% (oct(ols) for the \( L_1 \) level series’ hourly forecasts) and 38.3% (3TIER for the \( L_0 \) total series’ daily forecasts). The forecasting accuracy indices for each Transmission Zone forecasts are reported in table 4.

Table 3: Forecast accuracies (nRMSE %) and skills over the PERS benchmark of base forecasts and sequential reconciliation approaches. Free reconciliation procedures considered by Yagli et al. (2019), Tables 2, 3, p. 395. Hourly (H) and Daily (D) forecasts, forecast horizon: operating day. Bold entries identify the best performing approach.

| L_0 | L_1 | L_2 | L_0 | L_1 | L_2 |
|-----|-----|-----|-----|-----|-----|
|     | H   | D   | H   | D   | H   | D   | H   | D   | H   | D   |
| PERS | 34.62 | 20.23 | 43.15 | 24.57 | 59.75 | 30.65 | 0   | 0   | 0   | 0   |
| 3TIER | 26.03 | 12.48 | 33.95 | 16.75 | 53.46 | 25.19 | 0.248 | 0.383 | 0.213 | 0.318 |
| base | 27.85 | 18.17 | 34.24 | 20.94 | 53.46 | 25.82 | 0.196 | 0.102 | 0.206 | 0.148 |
| oct(ols) | 30.69 | 17.91 | 39.17 | 21.74 | 51.54 | 26.73 | 0.113 | 0.115 | 0.092 | 0.115 |
| oct(ols, struc) | 28.74 | 17.33 | 35.48 | 20.82 | 49.32 | 26.14 | 0.170 | 0.143 | 0.178 | 0.152 |
| oct(struc, ols) | 28.26 | 16.96 | 35.19 | 20.61 | 48.20 | 25.46 | 0.184 | 0.162 | 0.184 | 0.161 |
| oct(struc) | 26.71 | 16.24 | 33.11 | 19.64 | 46.74 | 24.73 | 0.228 | 0.197 | 0.233 | 0.201 |
| te(ols)+cs(ols)* | 27.80 | 17.96 | 34.36 | 21.80 | 48.52 | 26.71 | 0.197 | 0.112 | 0.204 | 0.113 |
| te(ols)+cs(struc)* | 27.50 | 17.95 | 34.34 | 21.79 | 47.77 | 26.36 | 0.197 | 0.113 | 0.204 | 0.113 |
| te(struc)+cs(ols)* | 27.02 | 17.24 | 33.40 | 20.62 | 47.76 | 25.98 | 0.219 | 0.148 | 0.226 | 0.161 |
| te(struc)+cs(struc)* | 26.17 | 16.68 | 32.44 | 19.96 | 46.25 | 25.03 | 0.244 | 0.176 | 0.248 | 0.188 |

* The approach produces spatial and temporal incoherent forecasts.

* The approach produces temporal incoherent forecasts.

14 The results for all temporal aggregation orders, \( k \in \{24, 12, 8, 6, 4, 3, 2, 1\} \), are available in the on-line appendix.
We observe that:

- **3TIER<sub>bu</sub>** gives the best forecasting performance for the total series (L<sub>0</sub>) at any temporal granularity, the second best being te(struc<sub>2</sub>) + cs(struc).

- The approach te(struc<sub>2</sub>) + cs(struc) ranks first for 4 out of the 5 hourly upper series at level L<sub>1</sub>, whereas at daily level 3TIER<sub>bu</sub> ‘wins’ again. The performance of oct(struc) results very similar to that of te(struc<sub>2</sub>) + cs(struc).

- te(struc<sub>2</sub>) + cs(struc) and oct(struc) show the best performance for the 318 bottom time series at any temporal granularity, with a slight prevalence of oct(struc) (for k ≥ 3), plus oct(struc) forecasts are cross-temporally coherent.

- Unlike te(struc<sub>2</sub>) + cs(struc), 3TIER<sub>bu</sub> forecasts are always non-negative, and coherent both in space and time with the forecasts at any granularity.

- For these reasons, 3TIER<sub>bu</sub> should be considered as a challenging competitor in the evaluation of the newly proposed procedures (section 5).

To give a complete picture of the evaluation results for hourly and daily forecasts, in figure<sup>5</sup> the Multiple Comparison with the Best (MCB) Nemenyi tests are shown. This allows to establish if the forecasting performances of the considered techniques are significantly different. At daily level, oct(struc) ranks first, and is significantly better than the other forecasting approaches, with te(struc<sub>2</sub>) + cs(struc) at the second place. This result is reversed at hourly level: te(struc<sub>2</sub>) + cs(struc) ranks first and is significantly better than all the other approaches, with oct(struc) at the second place. However, it should be recalled that the forecasts produced by te(struc<sub>2</sub>) + cs(struc) are not temporally coherent, which means that the sum of the hourly forecasts does

---

Table 4: Forecast accuracies (nRMSE %) and skills over the PERS benchmark of base forecasts and sequential reconciliation approaches for the series at L<sub>1</sub> level. Free reconciliation procedures considered by Yagli et al. (2019). Tables 2, 3, p. 395. Hourly (H) and Daily (D) forecasts, forecast horizon: operating day. Bold entries identify the best performing approach.

|          | H   | D   | H   | D   | H   | D   | H   | D   |
|----------|-----|-----|-----|-----|-----|-----|-----|-----|
| PERS<sub>bu</sub> | 28.72 | 16.12 | 40.27 | 22.49 | 46.48 | 26.96 | 52.82 | 29.75 |
| 3TIER<sub>bu</sub> | 22.81 | 10.43 | 33.34 | 16.72 | 34.01 | 17.06 | 46.40 | 22.80 |
| base<sup>+</sup> | 22.41 | 13.55 | 32.05 | 18.70 | 35.14 | 21.95 | 44.94 | 26.89 |
| oct(ds) | 29.27 | 14.89 | 34.58 | 19.25 | 39.55 | 22.91 | 47.30 | 26.50 |
| oct(ds<sub>sec</sub>, struct<sub>te</sub>) | 24.26 | 13.93 | 32.32 | 18.47 | 36.69 | 21.78 | 43.89 | 25.30 |
| oct(struct<sub>te</sub>, ds<sub>sec</sub>) | 23.44 | 13.50 | 32.89 | 18.83 | 37.42 | 22.33 | 44.76 | 25.97 |
| oct(struct) | 22.00 | 12.81 | 30.92 | 17.82 | 34.79 | 20.94 | 42.02 | 24.50 |
| te(ds<sub>2</sub>) + cs(ds<sup>+</sup>) | 22.95 | 15.46 | 32.06 | 19.23 | 35.12 | 22.38 | 44.93 | 26.91 |
| te(struct<sub>2</sub>) + cs(ds<sup>+</sup>) | 22.93 | 15.46 | 32.05 | 19.22 | 35.11 | 22.37 | 44.92 | 26.90 |
| te(struct<sub>2</sub>) + cs(struct) | 22.11 | 13.36 | 31.40 | 18.70 | 34.70 | 21.93 | 42.97 | 26.06 |
| te(struct<sub>2</sub>) + cs(struct<sup>+</sup>) | 21.58 | 13.00 | 30.48 | 18.09 | 33.58 | 21.18 | 41.86 | 25.21 |

The approach produces spatial and temporal incoherent forecasts.

* The approach produces temporal incoherent forecasts.
not match with the corresponding daily forecasts (see figure 4).

| Forecast Approach | Mean Rank |
|-------------------|-----------|
| oct(ois)          | 8.9       |
| base              | 7.79      |
| 3TIERbu           | 7.76      |
| cstruc = oct(struct) | 7.12    |
| te(ois) + cs(ois) | 6.16      |
| cstruc = oct(struct) | 5.49    |
| te(ois) + cs(struct) | 4.15    |
| ocl(struct)       | 2.3       |
| te(struct) + cs(struct) | 1.37   |

**Figure 5: MCB Nemenyi test results: average ranks and 95% confidence intervals.** The free reconciliation approaches considered by Yagli et al. (2019) are sorted vertically according to the nRMSE% mean rank. Hourly (top panel), and Daily (bottom panel) forecasts for L0, L1, L2 levels (324 series). Forecast horizon: operating day. The mean rank of each approach is displayed to the right of their names. If the intervals of two forecast reconciliation approaches do not overlap, this indicates a statistically different performance. Thus, approaches that do not overlap with the green interval are considered significantly worse than the best, and vice-versa.

Limiting ourselves to consider fully coherent reconciled forecasts, the scatter plots of the 324 couples of nRMSE(%) for oct(struct) vs., respectively, PERSbu and 3TIERbu (figure 6), show that the most performing regression-based cross-temporal reconciliation approach improves uniformly on the benchmark, and in the majority of cases on 3TIERbu, particularly at hourly level (k = 1), where the 89% of the variables register an improvement in nRMSE. However, it is worth noting that at daily level, this value decreases to 49%, which means that the NWP forecasts may still play a role at lower time granularity.
5. Extended analysis: non-negative cross-temporal reconciliation

In this section, we explore the performance of forecast reconciliation approaches able to produce non-negative PV forecasts, both temporally and spatially coherent. For this reason, among the approaches proposed by Yagli et al. (2019), we consider the reference benchmark PERS\textsubscript{bu}, the NWP base forecasts 3TIER\textsubscript{bu}, and oct\textsubscript{(struc)}. These approaches are then compared with the following 7 cross-temporal forecasting procedures:

- KA\textsubscript{(wlsv\texttt{te}, \texttt{wls\texttt{cs})}: the heuristic approach by Kourentzes and Athanasopoulos (2019), using \texttt{te(wlsv)} in the first step, and \texttt{cs(wls)} in the second, respectively;

Figure 6: Comparison of nRMSE(\%) between PERS\textsubscript{bu} and ct\textsubscript{(struc)} (top panel), and between 3TIER\textsubscript{bu} and ct\textsubscript{(struc)} (bottom panel). The black line represents the bisector, where the nRMSE's for both approaches are equal. On the top-left (bottom-right) corner of each graph, the percentage of points above (below) the bisector is reported.
• $ct(wls, bn, cs)$, $ct(struc, bn, cs)$, $ct(wls, bn, te)$, and $ct(struc, bn, te)$: partly bottom-up cross temporal reconciliation (see section 2.3) of the high frequency bottom time series’ reconciled forecasts according to, respectively, $te(wlsv)$, $te(struc)$, $cs(wls)$, and $cs(struc)$;

• $oct(wls)$ and $oct(bdshr)$: optimal (in least squares sense) cross-temporal forecast reconciliation approaches using the in-sample forecast errors [Di Fonzo and Girolimetto 2021].

5.1. Non negative forecast reconciliation: sntz

Each approach, when used in its ‘free’ version, i.e., without considering non-negative constraints in the linearly constrained quadratic program (4), is not guaranteed to always produce non-negative reconciled forecasts (see section 4.1). This fact may be an issue for the analyst, since in many practical situations negative forecasts could have no meaning, thus undermining the quality of the results found and the conclusions thereof. In what follows, we consider a simple heuristic strategy to avoid negative reconciled forecasts, without using any sophisticated, and time consuming, numerical optimization procedure. More precisely, possible negative values of the free reconciled high-frequency bottom time series forecasts are set to zero. Denote $\hat{B}_0^{[1]}$ the matrix containing the non-negative reconciled forecasts produced by a ‘free’ approach, and the ‘zeroed’ ones. The complete vector of non-negative cross-temporal reconciled forecasts is computed as the cross-temporal bottom-up aggregation of $\tilde{b}_0^{[1]} = \text{vec} \left( \left( \hat{B}_0^{[1]} \right) \right)'$, that is, according to expression (2):

$$\tilde{y}_0 = F \tilde{b}_0^{[1]}.$$

We call set-negative-to-zero (sntz) this simple, and quick device to obtain non negative reconciled forecasts. While it certainly increases the forecasting accuracy of the high-frequency bottom time series forecasts wrt the ‘free’ counterparts, this does not hold true in general for the upper level series forecasts. In addition, even if the originally reconciled forecasts are obtained according to an unbiased approach, the sntz-reconciled forecasts are no more unbiased, like the non-negative forecasts obtained through numerical optimization procedures [Wickramasuriya et al. 2020]. However, in practical situations the differences between the results produced by the sntz heuristic, and those obtained through a numerical optimization procedure could be negligible. For example, for the dataset and the forecasting experiment in hand, we have found that the forecasting performance of the non-negative reconciliation approaches through the non-linear optimization procedure osqp [Stellato et al. 2020] implemented in FoReco [Girolimetto and Di Fonzo 2022], is basically the same as the one obtained with the sntz heuristic. Table 5 shows the nRMSE(%) of the reconciled forecasts through $oct(struc)$ approach in both free and non-negative (sntz and osqp) variants. It is worth noting that the simple sntz heuristic always gives the lowest nRMSE(%), independently of the temporal granularity of the forecasts.

---

15 Recent contributions on this topic in the hierarchical forecasting field are [Wickramasuriya et al. 2020] and [Di Fonzo 2022].

16 Similar results have been found for all the considered reconciliation approaches. The details are available at request from the authors.
Table 5: Forecast accuracy (nRMSE%) of free and non-negative reconciled forecasts using the oct\((struc)\) approach. All temporal aggregation orders are considered, from hourly \((k = 1)\) to daily \((k = 24)\). Forecast horizon: operating day. Bold entries identify the best approach.

| non-negative reconciliation | \(k\): temporal aggregation order | \(L_0\) (1 series) | \(L_1\) (5 series) | \(L_2\) (318 series) |
|-----------------------------|-----------------------------------|---------------------|---------------------|---------------------|
| free                        | 1 2 3 4 6 8 12 24                 | 26.71 26.35 25.67 25.57 23.03 24.80 16.74 16.24 | 33.11 32.54 31.60 31.37 28.19 30.06 20.49 19.64 | 46.74 44.02 42.05 41.07 36.67 38.16 26.56 24.73 |
| sntz                        | 26.64 26.27 25.56 25.48 22.86 16.25 16.24 | 32.99 32.41 31.45 31.23 27.96 29.91 19.88 19.00 | 46.51 43.80 41.80 40.83 36.34 37.90 25.85 23.97 | 46.63 43.93 41.94 40.93 36.53 37.99 25.97 24.10 |
| osqp                        | 26.74 26.36 25.67 25.55 23.01 24.77 16.36 15.85 | 33.11 32.52 31.58 31.32 28.15 29.99 20.01 19.14 | 46.63 43.93 41.94 40.93 36.53 37.99 25.97 24.10 | 46.63 43.93 41.94 40.93 36.53 37.99 25.97 24.10 |

This result is visually confirmed by the graphs in figure 7, showing the scatter plots of the 324 couples of nRMSE(%) for oct\((struc)\)osqp vs. oct\((struc)\)sntz. For this dataset, we found that oct\((struc)\)sntz improves on oct\((struc)\)osqp in not less than the 94% of the 324 series for any temporal granularity.

Figure 7: Comparison of nRMSE(%) between sntz and osqp non-negative forecast reconciliation using the oct\((struc)\) approach. The black line represents the bisector, where the nRMSE’s for oct\((struc)\)osqp and oct\((struc)\)sntz are equal. On the top-left (bottom-right) corner of each graph, the percentage of points above (below) the bisector is reported.

5.2. Forecast accuracy of the selected approaches

Table 6 shows the nRMSE(%) of the considered non-negative cross-temporal forecast reconciliation approaches, using the sntz heuristic, and the corresponding forecast skills over the benchmark forecasts PERS\(_{bu}\). Overall, the accuracy improvements of the newly proposed approaches over the persistence model are in the range 16.5%–46.5%, whereas 3TIER\(_{bu}\) improvements are in the range 10.5%–38.3%. Partly bottom-up approaches, with cross-sectional reconciliation at the first step (i.e., ct\((wls_{cs}, bu_{te})\) and ct\((struc_{cs}, bu_{te})\)) show...
a good performance for the six series at levels $L_0$ and $L_1$. For the 318 most disaggregated series at the bottom level, $ct(wls_{vtc}, bu_{cs})$ and $oct(wlsv)$ rank first and second, respectively, and their accuracy indices are very close each other.

Table 6: Forecast accuracy of selected non-negative, cross-temporal reconciliation approaches and base forecasts in terms of nRMSE(%) for the series at $L_1$ level. Hourly and daily forecasts, forecast horizon: operating day. Bold entries and italic entries identify the best and the second best performing approaches, respectively.

|          | $L_0$ | $L_1$ | $L_2$ | $L_0$ | $L_1$ | $L_2$ |
|----------|-------|-------|-------|-------|-------|-------|
|          | H     | D     | H     | D     | H     | D     |
| PERS$_{bu}$ | 34.62 | 20.23 | 43.15 | 24.57 | 59.75 | 30.65 |
| 3TIER$_{bu}$ | 26.03 | 12.48 | 33.95 | 16.75 | 53.46 | 25.19 |
| oct(struct) | 26.64 | 15.73 | 32.99 | 19.00 | 46.51 | 23.97 |
| KA($wls_{vtc},wls_{cs}$) | 23.39 | 13.55 | 30.23 | 16.91 | 44.48 | 22.21 |
| ct(struct$_{cs},bu_{cs}$) | 21.99 | 12.38 | 29.80 | 15.82 | 49.76 | 24.33 |
| ct($wls_{vtc},bu_{cs}$) | 21.23 | 10.82 | 29.97 | 15.01 | 49.88 | 23.66 |
| $ct(struct_{cs},bu_{cs})$ | 24.81 | 14.43 | 31.41 | 17.80 | 45.30 | 22.93 |
| $ct(wls_{vtc},bu_{cs})$ | 23.50 | 13.42 | 30.13 | 16.78 | 44.42 | 22.12 |
| oct(wlsv) | 23.63 | 13.64 | 30.21 | 17.01 | 44.44 | 22.27 |
| oct(bdshr) | 24.13 | 13.92 | 30.45 | 17.12 | 45.24 | 22.54 |

From table 7, we see that even at distinct $L_1$ series (Transmission Zones), the newly proposed approaches perform better than all the approaches considered by Yagli et al. (2019), as reported in table 4.

Table 7: Forecast accuracy of selected non-negative, cross-temporal reconciliation approaches and base forecasts in terms of nRMSE(%) for the series at $L_1$ level. Hourly and daily forecasts, forecast horizon: operating day. Bold entries and italic entries identify the best and the second best performing approaches, respectively.

|          | $TZ_1$ | $TZ_2$ | $TZ_3$ | $TZ_4$ | $TZ_5$ |
|----------|--------|--------|--------|--------|--------|
|          | H     | D     | H     | D     | H     |
| PERS$_{bu}$ | 28.72 | 16.12 | 40.27 | 22.40 | 46.48 | 26.96 |
| 3TIER$_{bu}$ | 22.81 | 10.43 | 33.34 | 16.72 | 34.01 | 17.06 |
| oct(struct) | 21.92 | 12.41 | 30.84 | 17.30 | 34.67 | 20.21 |
| KA($wls_{vtc},wls_{cs}$) | 20.23 | 11.10 | 28.57 | 15.43 | 30.98 | 17.45 |
| ct(struct$_{cs},bu_{cs}$) | 19.45 | 12.92 | 27.01 | 14.37 | 31.77 | 15.74 |
| $ct(wls_{vtc},bu_{cs})$ | 19.33 | 9.72 | 28.11 | 14.18 | 29.55 | 15.10 |
| $ct(struct_{cs},bu_{cs})$ | 20.01 | 10.98 | 28.53 | 15.35 | 30.89 | 17.37 |
| $ct(wls_{vtc},bu_{cs})$ | 20.18 | 11.18 | 28.53 | 15.50 | 30.97 | 17.53 |
| oct(wlsv) | 21.00 | 11.74 | 29.25 | 16.02 | 31.63 | 17.85 |
| oct(bdshr) | 20.18 | 11.18 | 28.53 | 15.50 | 30.97 | 17.53 |

In addition, almost all the new approaches significantly outperform $PERS_{bu}$, $3TIER_{bu}$, and $oct(struct)$ for both hourly and daily forecasts (figure 8).
Finally, focusing on the two best performing approaches of the forecasting experiment ($\text{ct}(wlsv_{te}, bu_{cs})$ and $\text{oct}(wlsv)$), and looking at the nRMSEs of the individual series, it is worth noting that:

- $\text{ct}(wlsv_{te}, bu_{cs})$ always produces more accurate forecasts than $\text{oct}(\text{struct})$, for any series and granularity (figure 9),
- the accuracy increases of $\text{ct}(wlsv_{te}, bu_{cs})$ over the NWP approach $3\text{TIER}_{bu}$ are still clear (figure 10), even though for daily forecasts $3\text{TIER}_{bu}$ performs better in about 1 case out 4;
- the accuracy of $\text{ct}(wlsv_{te}, bu_{cs})$ and $\text{oct}(wlsv)$ are practically indistinguishable (figure 11).

![Figure 8: MCB-Nemenyi test on selected non-negative cross-temporal reconciliation approaches with operating day forecast horizon. $L_0, L_1, L_2$ levels (324 series). Top panel: hourly forecasts; Bottom panel: daily forecasts. The mean rank of each approach is displayed to the right of their names. If the intervals of two forecast reconciliation approaches do not overlap, this indicates a statistically different performance. Thus, approaches that do not overlap with the green interval are considered significantly worse than the best, and vice-versa.](image-url)
Figure 9: Comparison of nRMSE(%) between non-negative reconciliation approaches: oct\((\text{struct})\) and ct\((wls\text{vec}, bucs)\), forecast horizon: operating day. The black line represents the bisector, where the nRMSE’s for both approaches are equal. On the top-left (bottom-right) corner of each graph, the percentage of points above (below) the bisector is reported.

Figure 10: Comparison of nRMSE(%) between non-negative reconciliation approaches: 3TIER\textsubscript{bu} and ct\((wls\text{vec}, bucs)\). Forecast horizon: operating day. The black line represents the bisector, where the nRMSE’s for both approaches are equal. On the top-left (bottom-right) corner of each graph, the percentage of points above (below) the bisector is reported.
We may thus conclude that, for the PV324 dataset considered in this work, a thorough exploitation of cross-temporal hierarchies significantly improves the forecasting accuracy as compared to the approaches considered in Yagli et al. (2019). In particular, the use of the in-sample base forecast errors, even in the simple diagonal versions of \(\text{te}(w_{lsv})\) (first step of \(\text{ct}(w_{lsv_{te}}, bu_{cs})\)), and \(\text{oct}(w_{lsv})\), increases the forecasting accuracy at different time granularities.

6. Conclusion

Renewable energy is providing increasingly more energy to the grid all over the world. But grid operators must carefully manage the balance between the generation and consumption of energy to make the best use of abundant renewable energy. For an ISO, this provides greater grid stability, higher revenue and better use of what sun is available at any one time. Better short-term solar energy forecasts mean lower-emissions, cheaper energy and a more stable electricity grid. Solar forecasting is thus a key tool to achieve these results. In this paper, cross-temporal point forecast reconciliation has been applied to generate non-negative, fully coherent (both in space and time) forecasts of PV generated power. Both methodological and practical issues have been tackled, in order to develop effective and easy-to-handle cross-temporal forecasting approaches. Besides assuring both spatial and temporal coherence, and non-negativity of the reconciled forecasts, the results show that for the considered dataset, cross-temporal forecast reconciliation significantly improves on the sequential procedures proposed by Yagli et al. (2019), at any cross-sectional level of the hierarchy and for any temporal aggregation order. However, these findings should be considered neither conclusive, nor valid in general. Other forecasting experiments, through simulations and using other datasets, as well as in solar forecasting and in other application fields, should be performed to empirically robustify the results shown so far. In addition, other research queries could raise in the field of cross-temporal PV forecast
reconciliation, like using reduced temporal hierarchies, as well as forecasting through appropriate Machine Learning end-to-end approaches (Stratigakos et al., 2022), and probabilistic instead of deterministic (point) forecasting (Panamtash and Zhou, 2018; Jeon et al., 2019; Panagiotelis et al., 2022; Yang, 2020a, Yang, 2020b; Ben Taieb et al., 2021). All these topics are left for future research.

Appendix A. Cross-temporal bottom-up forecast reconciliation

A visual representation of cross-sectional, temporal, and cross-temporal bottom-up reconciliation is shown in figure A.1 where the ‘bottom time series’ base forecasts are highlighted on a purple background, and the ‘upper time series’ bottom-up reconciled forecasts are highlighted on a pink background. The arrows show the direction(s) along with bottom-up reconciliation is performed. The deponent abbreviations ‘cs’, ‘te’, ‘ct’, are alternatively used to denote reconciled forecasts according to cross-sectional, temporal, and cross-temporal bottom-up, respectively.

(a) \( \tilde{Y}_{cs(bu)} \): cross-sectional bottom-up

(b) \( \tilde{Y}_{te(bu)} \): temporal bottom-up

(c) \( \tilde{Y}_{ct(bu)} \): cross-temporal bottom-up

Figure A.1: A ‘visual’ representation of cross-sectional (\( \tilde{Y}_{cs(bu)} \)), temporal (\( \tilde{Y}_{te(bu)} \)), and cross-temporal (\( \tilde{Y}_{ct(bu)} \)) bottom-up reconciled forecasts.

Cross-sectional bottom-up reconciliation of \( k^* + m \) hierarchical time series’ base forecasts

The cross-sectional bottom-up (cs(bu)) reconciliation of hierarchical time series’ base forecasts at different temporal granularities (temporal aggregation levels) \( k \in \mathcal{K} \), may be represented as follows:

\[
\tilde{y}_{\tau,cs(bu)}^{[k]} = S \tilde{h}_{\tau}^{[k]}, \quad \tau = 1, \ldots, m \quad k \in \mathcal{K} \quad \rightarrow \quad \tilde{Y}_{cs(bu)}^{[k]} = S \tilde{B}^{[k]}, \quad k \in \mathcal{K}.
\]

In compact notation we have:

\[
\tilde{Y}_{cs(bu)} = S \tilde{B} \rightarrow \tilde{Y}_{cs(bu)}' = \tilde{B}' S' \rightarrow \tilde{Y}_{cs(bu)} = (S \otimes I_{k^* + m}) \tilde{b},
\]

where \( \tilde{Y}_{cs(bu)} = \text{vec} \left( \tilde{Y}_{cs(bu)}' \right) \), and \( \tilde{b} = \text{vec} \left( \tilde{B}' \right) \).

The cs(bu) reconciled forecasts are obviously cross-sectionally coherent, but in general they are not
temporally coherent:

\[
U'\tilde{Y}_{cs(bu)} = 0_{[n_a \times (k^* + m)]} ~ \leftrightarrow ~ (U' \otimes I_{k^* + m}) \tilde{Y}_{cs(bu)} = 0_{[n_a (k^* + m) \times 1]},
\]

\[
Z'\tilde{Y}_{cs(bu)} \neq 0_{[k^* \times n]} ~ \leftrightarrow ~ (I_n \otimes Z') \tilde{Y}_{cs(bu)} \neq 0_{[k^* \times n \times 1]}
\]

**Temporal bottom-up reconciliation of \( n \) individual time series’ base forecasts**

The temporal bottom-up (\( te(bu) \)) reconciliation of \( n \) individual time series’ base forecasts at different temporal granularities (temporal aggregation levels), may be represented as follows:

\[
\tilde{a}_{i,te(bu)} = RA_{i}^{[1]}, \quad i = 1, \ldots, n_a
\]

\[
\tilde{b}_{j,te(bu)} = RB_{j}^{[1]}, \quad j = 1, \ldots, n_b
\]

and, finally,

\[
\tilde{Y}'_{te(bu)} = R\tilde{Y}^{[1]}_{te(bu)} \rightarrow \tilde{y}_{te(bu)} = (I_n \otimes R)\tilde{y}^{[1]}_{te(bu)}
\]

The \( te(bu) \) reconciled forecasts are obviously temporally coherent, but in general they are not cross-sectionally coherent:

\[
Z'\tilde{Y}_{te(bu)} = 0_{[k^* \times n]} ~ \leftrightarrow ~ (I_n \otimes Z') \tilde{Y}_{te(bu)} = 0_{[k^* \times n \times 1]}
\]

\[
U'\tilde{Y}_{te(bu)} \neq 0_{[n_a \times (k^* + m)]} ~ \leftrightarrow ~ (U' \otimes I_{k^* + m}) \tilde{Y}_{te(bu)} \neq 0_{[n_a (k^* + m) \times 1]}
\]

**Cross-temporal bottom-up reconciliation of \( n \) individual time series’ base forecasts for different temporal granularities**

The cross-temporal bottom-up (\( ct(bu) \)) reconciliation of \( k^* + m \) hierarchical time series’ base forecasts at different temporal granularities, may be represented as follows:

\[
vec \left( \tilde{Y}'_{ct(bu)} \right) = (S \otimes R) vec \left( \tilde{B}^{[1]}_{ct(bu)} \right) ~ \leftrightarrow ~ \tilde{y}_{cs(bu)} = (S \otimes R)\tilde{b}^{[1]}_{ct(bu)}, \quad (A.1)
\]

where \( \tilde{b}^{[1]} = vec \left( \tilde{B}^{[1]} \right) \). Let us consider now the coherency checks for the \( ct(bu) \) reconciled forecasts:

- **cross-sectional coherency**

\[
(U' \otimes I_{k^* + m}) (S \otimes R)\tilde{b}^{[1]} = (U'S \otimes R)\tilde{b}^{[1]} = 0_{[n_a (k^* + m) \times 1]},
\]

since \( U'S = [I_{n_a} - C] \)

\[
\begin{bmatrix}
C \\
I_{n_a}
\end{bmatrix} = C - C = 0_{[n_a \times n_a]};
\]

- **temporal coherency**

\[
(I_n \otimes Z') (S \otimes R)\tilde{b}^{[1]} = (S \otimes Z'R)\tilde{b}^{[1]} = 0_{[k^* \times n \times 1]},
\]

since \( Z'R = [I_{k^*} - K] \)

\[
\begin{bmatrix}
K \\
I_m
\end{bmatrix} = K - K = 0_{[k^* \times m]}.
\]

In the end, cross temporal bottom-up reconciliation can be tought of as a two-step sequential cross temporal reconciliation approach, where either cross-sectional reconciliation of the high-frequency bottom time series base forecasts (the ‘very’ bottom series of a cross-temporal hierarchy), is followed by temporal reconciliation,
or *vice-versa*. For, looking at expression (A.1), and by exploiting the properties of the vec operator (Harville 2008, p. 345), it is:

\[
(S \otimes R) \text{vec}\left(\hat{\mathbf{B}}^{[1]r}\right) = \text{vec}\left( R \left( \hat{\mathbf{B}}^{[1]r} \right) S' \right),
\]

which may be interpreted in two different, equivalent ways:

- looking at the expression \( R \left( \hat{\mathbf{B}}^{[1]r} \right) S' \) from the right, one first computes the cross-sectional bottom-up reconciled forecasts of all the \( n \) high-frequency time series, \( \hat{\mathbf{Y}}_{cs(bu)}^{[1]} = S\hat{\mathbf{B}}^{[1]} \). These forecasts, that are not temporally coherent with the base forecasts of the same series at lower temporal granularity (i.e., \( \hat{\mathbf{Y}}^{[k]}, k \neq 1 \)), are then reconciled via temporal bottom-up, \( \hat{\mathbf{Y}}'_{ct(bu)} = R\hat{\mathbf{Y}}_{cs(bu)}^{[1]} \):

  \[
  \hat{\mathbf{B}}^{[1]} \xrightarrow{cs(bu)} \hat{\mathbf{Y}}_{cs(bu)}^{[1]} \xrightarrow{te(bu)} \hat{\mathbf{Y}}'_{ct(bu)};
  \]

- moving in turn from the left, one first computes the temporally reconciled forecasts of the \( n_b \) bottom time series of the hierarchy, \( \hat{\mathbf{B}}'_{te(bu)} = R \left( \hat{\mathbf{B}}^{[1]r} \right) \). These forecasts, that are not cross-sectionally coherent with the base forecasts of the \( n_u \) upper time series of the hierarchy, are then reconciled via cross-sectional bottom-up, \( \hat{\mathbf{Y}}_{ct(bu)} = S \left( \hat{\mathbf{B}}_{te(bu)} \right) \):

  \[
  \hat{\mathbf{B}}^{[1]} \xrightarrow{te(bu)} \hat{\mathbf{B}}_{te(bu)} \xrightarrow{cs(bu)} \hat{\mathbf{Y}}_{ct(bu)};
  \]

The two equivalent interpretations of the cross-temporal bottom-up approach are represented in Figure A.2, where \( \hat{\mathbf{B}}_{ct(bu)} \equiv \hat{\mathbf{B}}^{[1]}, \hat{\mathbf{A}}_{ct(bu)} \equiv \hat{\mathbf{A}}_{cs(bu)}, \) and \( \hat{\mathbf{B}}_{ct(bu)} \equiv \hat{\mathbf{B}}^{[k]}, k > 1 \).

![Figure A.2: A ‘visual’ interpretation of \( \hat{\mathbf{Y}}_{ct(bu)} \) in two different, equivalent ways.](image-url)
Appendix B. Proof of the theorem

1. Let $\mathbf{U}_{n \times n}$ and $\mathbf{Z}'_{m \times (k^*+m)}$ be the cross-sectional and temporal, respectively, zero constraints kernel matrix. Let

$$M_{cs}^{[k]} = M_{cs} = \left[ I_n - \mathbf{WU} (\mathbf{U}'\mathbf{WU})^{-1} \mathbf{U}' \right], \quad k \in \mathcal{K}$$

and

$$M_{te,i} = M_{te} = \left[ I_{k^*+m} - \Omega Z (Z'\Omega Z)^{-1} Z' \right], \quad i = 1, \ldots, n,$$

be the cross-sectional and temporal projection matrices, respectively, such that:

$$\tilde{Y}_{cs}^{[k]} = M_{cs} \hat{Y}^{[k]}, \quad k \in \mathcal{K},$$

$$\hat{y}_{te,i} = M_{te} \hat{y}_i, \quad i = 1, \ldots, n,$$

where $\hat{Y}^{[k]}$ and $\tilde{Y}_{cs}^{[k]}$ are the $(n \times \frac{m}{k})$ matrices with, respectively, the base and the cross-sectional reconciled forecasts of the $n$ series at time granularity $k$, and $\hat{y}_i$ and $\hat{y}_{te,i}$ are the $[(k^* + m) \times 1]$ vectors with, respectively, the base and the temporal reconciled forecasts at all time granularities for the $i$-th series. In compact matrix form we get:

$$\tilde{Y}_{cs} = \left[ \tilde{Y}_{cs}^{[m]} \tilde{Y}_{cs}^{[k_{p-1}]} \ldots \tilde{Y}_{cs}^{[k_2]} \tilde{Y}_{cs}^{[1]} \right] = M_{cs} \left[ \tilde{Y}^{[m]} \tilde{Y}^{[k_{p-1}]} \ldots \tilde{Y}^{[k_2]} \tilde{Y}^{[1]} \right] = M_{cs} \tilde{Y},$$

$$\hat{Y}_{te} = \left\{ M_{te} \hat{y}_1 \ldots \hat{y}_i \ldots \hat{y}_n \right\}' = \left\{ M_{te} \hat{Y} \right\}'.$$

Then, $\hat{Y}_{cs}$ is obtained by first operating a cross-sectional reconciliation (step 1, $\hat{Y}_{cs}$), and then reconciling via temporal hierarchies the forecasts obtained in the previous step (step 2, $\hat{Y}_{cs}$). In matrix terms:

step 1: $\tilde{Y}_{cs} = M_{cs} \hat{Y}$

step 2: $\tilde{Y}_{cs} = \left\{ M_{te} \hat{Y}' \right\}' = \tilde{Y}_{cs} M'_{te} = M_{cs} \tilde{Y} M'_{te}$

where $\hat{Y}$ is the $(n \times (k^* + m))$ base forecasts matrix. Vice versa, $\tilde{Y}_{cs}$ is obtained by first operating a temporal reconciliation (step 1, $\hat{Y}_{te}$), and finally a cross-sectional reconciliation is applied (step 2):

step 1: $\hat{Y}_{te} = \left\{ M_{te} \hat{Y} \right\}' = \hat{Y} M'_{te}$

step 2: $\tilde{Y}_{cs} = M_{cs} \hat{Y}_{te} = M_{cs} \hat{Y} M'_{te} = \tilde{Y}_{cs}$.

The forecasts $\tilde{Y}_{cs}$ (and $\tilde{Y}_{cs}$) (a single two-step iteration) are cross-temporally reconciled because:

$$U'\tilde{Y}_{cs} = U' M_{cs} \tilde{Y} M'_{te} = \left[ U' - U' \mathbf{WU} (U'\mathbf{WU})^{-1} U' \right] \tilde{Y} M'_{te} = [U' - U'] \tilde{Y} M'_{te} = 0_{n \times (k^*+m)}$$

and

$$Z'\tilde{Y}_{cs} = Z' M_{te} \tilde{Y} M'_{cs} = \left[ Z' - Z' \Omega Z (Z'\Omega Z)^{-1} Z' \right] M'_{cs} = [Z' - Z'] \tilde{Y} M'_{cs} = 0_{m \times n}.$$

2. Using matrix vectorization and structural representation for the cross-temporal reconciliation problem

30
we have:

\[ \tilde{y}_{\text{oct}} = F (F'\Omega_{ct}^{-1}F)^{-1}F'\tilde{y}, \]

where \( \tilde{y}_{\text{oct}} = \text{vec}\left(\hat{Y}_{\text{oct}}'\right) \), \( \tilde{y} = \text{vec}\left(\hat{Y}_{\text{oct}}'\right) \), and \( F = S \otimes R \) is the \([n(k^* + m) \times mn]\) cross-temporal summing matrix. Let \( \Omega_{ct} = W \otimes \Omega \) be the cross-temporal error covariance matrix, by exploiting the properties of the Kronecker product (Henderson and Searle 1981, Harville 2008), we obtain:

\[ \tilde{y}_{\text{oct}} = F (F'\Omega_{ct}^{-1}F)^{-1}F'\Omega_{ct}^{-1}\tilde{y} \]

\[ = (S \otimes R) [(S \otimes R)'(W \otimes \Omega)^{-1}(S \otimes R)]^{-1} (S \otimes R)'(W \otimes \Omega)^{-1}\tilde{y} \]

\[ = (S \otimes R) [(S' \otimes R')(W^{-1} \otimes \Omega^{-1})(S \otimes R)]^{-1} (S' \otimes R')(W^{-1} \otimes \Omega^{-1})\tilde{y} \]

\[ = (S \otimes R) (S'W^{-1}S \otimes R'\Omega^{-1}R)^{-1} (S'W^{-1} \otimes R'\Omega^{-1})\tilde{y} \]

\[ = \left[ S (S'W^{-1}S)^{-1}S'W^{-1} \otimes R (R'\Omega^{-1}R)^{-1} R'\Omega^{-1} \right] \tilde{y}. \]

Exploiting the structural notation also in the temporal then cross-sectional iterative procedure and the vectorization \( \tilde{y}_{\text{tcs}} = \text{vec}\left(\hat{Y}_{\text{tcs}}'\right) \), we obtain:

\[ \tilde{y}_{\text{tcs}} = P \left[ I_{k^*+m} \otimes S (S'W^{-1}S)^{-1}S'W^{-1} \right] P' \left[ I_n \otimes R (R'\Omega^{-1}R)^{-1} R'\Omega^{-1} \right] \tilde{y} \]

\[ = \left[ S (S'W^{-1}S)^{-1}S'W^{-1} \otimes I_{k^*+m} \right] \left[ I_n \otimes R (R'\Omega^{-1}R)^{-1} R'\Omega^{-1} \right] \tilde{y} \]

\[ = \left[ S (S'W^{-1}S)^{-1}S'W^{-1} \otimes R (R'\Omega^{-1}R)^{-1} R'\Omega^{-1} \right] \tilde{y} = \tilde{y}_{\text{oct}}. \]

Finally, \( \hat{Y}_{\text{tcs}} = \hat{Y}_{\text{oct}} \) and, following from point 1, \( \hat{Y}_{\text{tcs}} = \hat{Y}_{\text{ctst}} = \hat{Y}_{\text{oct}}. \)

\[ \square \]

The heuristic KA with constant cross-sectional and temporal error covariance matrices

In the Kourentzes and Athanasopoulos (2019) approach (section 2.5), the temporally reconciled predictions at step 1, \( \hat{Y}M'_{\text{te}} \), are cross-sectionally reconciled via premultiplication by the matrix \( M = \frac{1}{k} \sum_{k \in k} M^k \).

As \( W^k = W \), \( k \in k \), \( M^k_{\text{cs}} = M_{\text{cs}} \), and \( M = M_{\text{cs}} \). Then,

\[ \hat{Y}_{\text{KA}} = MM'_{\text{te}} = M_{\text{cs}} \hat{Y}M'_{\text{te}} = \hat{Y}_{\text{ctst}} = \hat{Y}_{\text{tcs}}. \]

References

3TIER, 2010. Development of Regional Wind Resource and Wind Plant Output Datasets: Final Subcontract Report. Electricity generation, Energy storage, Forecasting wind, Solar energy, Wind energy, Wind farms, Wind modeling. URL: https://digitalscholarship.unlv.edu/renew_pubs/4 (visited on May 10, 2022).

Athanasopoulos, G., Ahmed, R.A., Hyndman, R.J., 2009. Hierarchical forecasts for Australian domestic tourism. International Journal of Forecasting 25, 146–166. doi:10.1016/j.ijforecast.2008.07.004

Athanasopoulos, G., Hyndman, R.J., Kourentzes, N., Petropoulos, F., 2017. Forecasting with temporal hierarchies. European Journal of Operational Research 262, 60–74. doi:10.1016/j.ejor.2017.02.046

31
Babai, M.Z., Boylan, J.E., Rostami-Tabar, B., 2022. Demand forecasting in supply chains: a review of aggregation and hierarchical approaches. International Journal of Production Research 60, 324–348. doi:10.1080/00207543.2021.2005268

Bates, D., Maechler, M., Jagan, M., Davis, T.A., Oehlschlägel, J., Riedy, J., 2022. Package Matrix: Sparse and Dense Matrix Classes and Methods. Version 1.4-1. URL: https://cran.r-project.org/web/packages/Matrix/index.html

Ben Taieb, S., Taylor, J.W., Hyndman, R.J., 2021. Hierarchical Probabilistic Forecasting of Electricity Demand with Smart Meter Data. Journal of the American Statistical Association 116, 27–43. doi:https://doi.org/10.1080/01621459.2020.1736081

Byron, R.P., 1978. The Estimation of Large Social Account Matrices. Journal of the Royal Statistical Society. Series A 141, 359–367. doi:10.2307/2344807

Dangerfield, B., Morris, J., 1992. Top-down or bottom-up: Aggregate versus disaggregate extrapolations. International Journal of Forecasting 6, 233–241. doi:10.1016/0169-2070(92)90121-O

Davis, T.A., 2006. Direct Methods for Sparse Linear Systems. Society for Industrial and Applied Mathematics.

Di Fonzo, T., Girolimetto, D., 2021. Cross-temporal forecast reconciliation: Optimal combination method and heuristic alternatives. International Journal of Forecasting (in press). doi:10.1016/j.ijforecast.2021.08.004

Dunn, D.M., Williams, W.H., Dechaine, T.L., 1976. Aggregate versus subaggregate models in local area forecasting. Journal of the American Statistical Association 71, 68–71. doi:10.1080/01621459.1976.10481478

van Erven, T., Cugliari, J., 2015. Game-Theoretically Optimal Reconciliation of Contemporaneous Hierarchical Time Series Forecasts, in: Antoniadis, A., Poggi, J.M., Brossat, X. (Eds.), Modeling and Stochastic Learning for Forecasting in High Dimensions. Springer Science and Business Media, LLC. volume 217, pp. 297–317. doi:10.1007/978-3-319-18732-7_15

EurObserv’ER, 2022. The state of renewable energies in Europe. 20th EurObserv’ER Report. Edition 2021 URL: https://www.eurobserv-er.org/pdf/20th-annual-overview-barometer/ (visited on September 8, 2022).

European Commission, 2019. Going climate-neutral by 2050: a strategic long-term vision for a prosperous, modern, competitive and climate-neutral EU economy. Directorate-General for Climate Action doi:10.2834/02074

Fliedner, G., 2001. Hierarchical forecasting: Issues and use guidelines. Industrial Management & Data Systems 101, 5–12. doi:10.1108/02635570110365952

Genton, M., 2007. Separable approximations of space-time covariance matrices. Environmetrics 18, 681–695. doi:10.1002/env.854

Girolimetto, D., Di Fonzo, T., 2022. Package FoReco: Point Forecast Reconciliation. Version 0.2.5. URL: https://cran.r-project.org/package=FoReco

Golub, G.H., Van Loan, C.F., 2013. Matrix computations. Johns Hopkins University Press.
Harville, D.A., 2008. Matrix Algebra From a Statistician’s Perspective. Springer-Verlag, New York.

Henderson, H.V., Searle, S.R., 1981. The vec-permutation matrix, the vec Operator and Kronecker products: a review. Linear and Multilinear Algebra 9, 271–288. URL: http://www.tandfonline.com/doi/abs/10.1080/03081088108817379 doi:10.1080/03081088108817379

Hyndman, R.J., 2020. Ten years of forecast reconciliation. 40th International Symposium on Forecasting 2020 virtual. Keynote speech URL: https://www.youtube.com/watch?v=5jB09R-sKDC (visited on April 2, 2022).

Hyndman, R.J., Ahmed, R.A., Athanasopoulos, G., Shang, H.L., 2011. Optimal combination forecasts for hierarchical time series. Computational Statistics and Data Analysis 55, 2579–2589. doi:10.1016/j.csda.2011.03.006

Hyndman, R.J., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O’Hara-Wild, M., Petropoulos, F., Razbash, S., Wang, E., Yasmeen, F., 2021. Package forecast: Forecasting functions for time series and linear models. Version 8.15. URL: https://pkg.robjhyndman.com/forecast/

Hyndman, R.J., Lee, A.J., Wang, E., 2016. Fast computation of reconciled forecasts for hierarchical and grouped time series. Computational Statistics and Data Analysis 97, 16–32. doi:10.1016/j.csda.2015.11.007

Jeon, J., Panagiotelis, A., Petropoulos, F., 2019. Probabilistic forecast reconciliation with applications to wind power and electric load. European Journal of Operational Research 279, 364–379. doi:10.1016/j.ejor.2019.05.020

Kleissl, J., 2013. Solar energy forecasting and resource assessment. Academic Press.

Koning, A.J., Franses, P.H., Hibon, M., Stekler, H.O., 2005. The M3 competition: Statistical tests of the results. International Journal of Forecasting 21, 397–409. doi:10.1016/j.ijforecast.2004.10.003

Kourentzes, N., Athanasopoulos, G., 2019. Cross-temporal coherent forecasts for Australian tourism. Annals of Tourism Research 75, 393–409. doi:10.1016/j.annals.2019.02.001

Ledoit, O., Wolf, M., 2004. A well-conditioned estimator for large-dimensional covariance matrices. Journal of Multivariate Analysis 88, 365–411. doi:10.1016/S0047-259X(03)00096-4

Makarov, Y.V., Etingov, P.V., Ma, J., Huang, Z., Subbarao, K., 2011. Incorporating Uncertainty of Wind Power Generation Forecast Into Power System Operation, Dispatch, and Unit Commitment Procedures. IEEE Transactions on Sustainable Energy 4, 433–442. doi:10.1109/TSTE.2011.2159254

Makridakis, S., Spiliotis, E., Assimakopoulos, V., 2022. The M5 Accuracy competition: Results, findings and conclusions. International Journal of Forecasting (in press). doi:10.1016/j.ijforecast.2021.11.013

Nystrup, P., Lindström, E., Pinson, P., Madsen, H., 2020. Temporal hierarchies with autocorrelation for load forecasting. European Journal of Operational Research 280, 876–888. doi:10.1016/j.ejor.2019.07.061

Paige, C.C., Sanders, M.A., 1982. LSQR: An Algorithm for Sparse Linear Equations and Sparse Least Squares. ACM Transactions on Mathematical Software 8, 43–71. doi:10.1145/355984.355989

Panagiotelis, A., Athanasopoulos, G., Gamakumara, P., Hyndman, R.J., 2021. Forecast reconciliation: A geometric view with new insights on bias correction. International Journal of Forecasting 37, 343–359. doi:10.1016/j.ijforecast.2020.06.004
Panagiotelis, A., Gamakumara, P., Athanasopoulos, G., Hyndman, R.J., 2022. Probabilistic Forecast Reconciliation: Properties, Evaluation and Score Optimisation. European Journal of Operational Research in press. doi:10.1016/j.ejor.2022.07.040

Panamtash, H., Zhou, Q., 2018. Coherent Probabilistic Solar Power Forecasting. 2018 IEEE International Conference on Probabilistic Methods Applied to Power Systems (PMAPS) doi:10.1109/PMAPS.2018.8440483

Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M., Barrow, D., Bergmeir, C., Bessa, R., Boylan, J., Browell, J., Carnevale, C., Castle, J., Cirillo, P., Clements, M., Cordeiro, C., Oliveira, F., De Baets, S., Dokumentov, A., Fiszeder, P., Franses, P., Gilliland, M., Gonul, M., Goodwin, P., Grossi, L., Grushka-Cockayne, Y., Guidolin, M., Guidolin, M., Gunter, U., Guo, X., Guseo, R., Harvey, N., Hendry, D., Hollyman, R., Januschowski, T., Jeon, J., Jose, V., Kang, Y., Kochler, A., Kolassa, S., Kourentzes, N., Leva, S., Li, F., Litsiou, K., Makridakis, S., Martinez, A., Meeran, S., Modis, T., Nikolopoulos, K., Onkal, D., Pacagnini, A., Panapakidis, I., Pavia, J., Pedio, M., Pedregal, D., Pinson, P., Ramos, P., Rapach, D., Reade, J., Rostami-Tabar, B., Rubaszek, M., Serpininis, G., Shang, H., Spiliotis, E., Syntetos, A., Talagala, P., Talagala, T., Tashman, L., Thomakos, D., Thorarinsdottir, T., Todini, E., Trapero Arenas, J., Wang, X., Winkler, R., Yusupova, A., Ziel, F., 2022. Forecasting: theory and practice. International Journal of Forecasting 38, 705–871. doi:10.1016/j.ijforecast.2021.11.001

Punia, S., Singh, S.P., Madaan, J.K., 2020. A cross-temporal hierarchical framework and deep learning for supply chain forecasting. Computers & Industrial Engineering 149, 106796. doi:10.1016/j.cie.2020.106796

Spiliotis, E., Petropoulos, F., Kourentzes, N., Assimakopoulos, V., 2020b. Cross-temporal aggregation: Improving the forecast accuracy of hierarchical electricity consumption. Applied Energy 261, 114339. doi:10.1016/j.apenergy.2019.114339

Stellato, B., Banjac, Goran, P., Bemporad, A., Boyd, S., 2020. OSQP: an operator splitting solver for quadratic programs. Mathematical Programming Computation 12, 637–672. doi:10.1007/s12532-020-00179-2

Stratigakos, A., van der Meer, D., Camal, S., Kariniotakis, G., 2022. End-to-end Learning for Hierarchical Forecasting of Renewable Energy Production with Missing Values. 17th International Conference on Probabilistic Methods Applied to Power Systems, PMAPS 2022, Jun 2022, Manchester - Online, United Kingdom doi:10.1109/PMAPS53380.2022.9810610

United Nations, 2015. Paris Agreement. URL: https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement (visited on September 9, 2022).

United Nations, 2022. The race to zero emissions accelerates in Asia. UN Climate Change News, 27 April 2022 URL: https://unfccc.int/news/the-race-to-zero-emissions-accelerates-in-asia (visited on September 8, 2022).

Velu, R., Herman, K., 2017. Separable covariance matrices and Kronecker approximation. Procedia Computer Science 108, 1019–1029. doi:10.1016/j.procs.2017.05.184

Werner, K., Jansson, M., Stoica, P., 2008. On estimation of covariance matrices with Kronecker product structure. IEEE Transactions on Signal Processing 56, 478–491. doi:10.1109/TSP.2007.907834
Wickramasuriya, S.L., Athanasopoulos, G., Hyndman, R.J., 2019. Optimal Forecast Reconciliation for Hierarchical and Grouped Time Series Through Trace Minimization. Journal of the American Statistical Association 114, 804–819. doi:10.1080/01621459.2018.1448825.

Wickramasuriya, S.L., Turlach, B.A., Hyndman, R.J., 2020. Optimal non-negative forecast reconciliation. Statistics and Computing 30, 1167–1182. doi:10.1007/s11222-020-09930-0.

Yagli, G.M., Yang, D., Srinivasan, D., 2019. Reconciling solar forecasts: Sequential reconciliation. Solar Energy 179, 391–397. doi:10.1016/j.solener.2018.12.075.

Yang, D., 2020a. Reconciling solar forecasts: Probabilistic forecast reconciliation in a nonparametric framework. Solar Energy 210, 49–58. doi:10.1016/j.solener.2020.03.095.

Yang, D., 2020b. Reconciling solar forecasts: Probabilistic forecasting with homoscedastic Gaussian errors on a geographical hierarchy. Solar Energy 210, 59–67. doi:10.1016/j.solener.2020.06.005.

Yang, D., Alessandrini, S., Antonanzas, J., Antonanzas-Torres, F., Badescu, V., Beyer, H., Blaga, R., Boland, J., Bright, J., Coimbra, C., David, M., Frimane, A., Gueymard, C., Hong, T., Kay, M., Killinger, S., Kleissl, J., Lauret, P., Lorenz, E., van der Meer, D., Pulescu, M., Perez, R., Perpiñán Lamigueiro, O., Peters, I., Reikard, G., Renné, D., Saint-Drenan, Y., Shuai, Y., Urraca, R., Verbois, H., Vignola, F., Voyant, C., Zhang, J., 2020. Verification of deterministic solar forecasts. Solar Energy 210, 20–37. doi:10.1016/j.solener.2020.04.019.

Yang, D., Quan, H., Disfani, V.R., Liu, L., 2017a. Reconciling solar forecasts: Geographical hierarchy. Solar Energy 146, 276–286. doi:10.1016/j.solener.2017.02.010.

Yang, D., Quan, H., Disfani, V.R., Rodríguez-Gallegos, C.D., 2017b. Reconciling solar forecasts: Temporal hierarchy. Solar Energy 158, 332–346. doi:10.1016/j.solener.2017.09.055.

Yang, D., Wang, W., Xia, X., 2022. A concise overview on solar resource assessment and forecasting. Advances in Atmospheric Sciences doi:10.1007/s00376-021-1372-8.

Acknowledgements

*Parts of this research were carried in the frame of the PRIN2017 project “HiDEA: Advanced Econometrics for High-frequency Data”, 2017RSMPZZ.*
On-line appendix: supplementary tables and graphs

Table OA.1: Number of series with at least one negative reconciled forecast. All temporal aggregation orders, forecast horizon: operating day.

| k | comb | n   | min | max | fmin | fmax |
|---|------|-----|-----|-----|------|------|
| 1 | base | 350 | 4   | 6   | -15.617 | 0.000 |
| 1 | oct\( (\text{ols}) \) | 350 | 109 | 324 | -69.165 | 0.000 |
| 1 | ite\( (\text{struc}, \text{ols}) \) | 350 | 137 | 324 | -20.739 | 0.000 |
| 1 | ite\( (\text{ols}, \text{struc}) \) | 350 | 91  | 324 | -17.961 | 0.000 |
| 1 | oct\( (\text{struc}) \) | 350 | 89  | 324 | -14.292 | 0.000 |
| 1 | te\( (\text{ols}) \) + cs\( (\text{ols}) \) | 350 | 42  | 324 | -12.993 | 0.000 |
| 1 | te\( (\text{struc}) \) + cs\( (\text{ols}) \) | 350 | 36  | 324 | -12.994 | 0.000 |

Continued on next page
| k | comb | n  | min | max  | fmin   | fmax   |
|---|------|----|-----|------|--------|--------|
| 6 | base | 350| 19  | 321  | -146.702 | 0      |
| 6 | oct(ols) | 350| 3   | 324  | -61.415 | 0      |
| 6 | ite(struc<sub>cs</sub>, ols<sub>te</sub>) | 350| 4   | 323  | -67.327 | 0      |
| 6 | ite(ols<sub>cs</sub>, struc<sub>te</sub>) | 350| 17  | 323  | -53.857 | 0      |
| 6 | oct(struc) | 350| 16  | 322  | -63.281 | 0      |
| 6 | te(ols<sub>2</sub>) + cs(ols) | 347| 0   | 324  | -133.572 | 0      |
| 6 | te(struc<sub>2</sub>) + cs(struc) | 347| 0   | 324  | -89.955 | 0      |
| 6 | te(ols<sub>2</sub>) + cs(struc) | 349| 0   | 324  | -88.924 | 0      |
| 6 | oct(struc) | 300| 0   | 305  | -78.762 | 0.000  |
| 8 | base | 212| 0   | 238  | -866.818 | 0.000  |
| 8 | oct(ols) | 300| 0   | 305  | -78.762 | 0.000  |
| 8 | oct(struc) | 276| 0   | 324  | -51.806 | 0.000  |
| 8 | ite(struc<sub>cs</sub>, ols<sub>te</sub>) | 296| 0   | 322  | -53.688 | 0.000  |
| 8 | ite(ols<sub>cs</sub>, struc<sub>te</sub>) | 274| 0   | 319  | -42.734 | 0.000  |
| 8 | te(ols<sub>2</sub>) + cs(ols) | 307| 0   | 320  | -680.766 | 0.000  |
| 8 | te(struc<sub>2</sub>) + cs(ols) | 285| 0   | 324  | -680.488 | 0.000  |
| 8 | te(ols<sub>2</sub>) + cs(struc) | 300| 0   | 241  | -56.011 | 0.000  |
| 8 | te(struc<sub>2</sub>) + cs(struc) | 284| 0   | 268  | -52.341 | 0.000  |
| 12 | base | 46 | 0   | 57   | -59.405 | -0.001 |
| 12 | oct(ols) | 14 | 0   | 78   | -15.328 | -0.001 |
| 12 | ite(struc<sub>cs</sub>, ols<sub>te</sub>) | 13 | 0   | 45   | -7.949  | -0.004 |
| 12 | ite(ols<sub>cs</sub>, struc<sub>te</sub>) | 8  | 0   | 31   | -6.672  | -0.001 |
| 12 | oct(struc) | 6  | 0   | 18   | -0.537  | -0.009 |
| 12 | te(ols<sub>2</sub>) + cs(ols) | 35 | 0   | 95   | -32.259 | 0.000  |
| 12 | te(struc<sub>2</sub>) + cs(ols) | 26 | 0   | 100  | -32.191 | 0.000  |
| 12 | te(ols<sub>2</sub>) + cs(struc) | 19 | 0   | 75   | -9.799  | 0.000  |
| 12 | te(struc<sub>2</sub>) + cs(struc) | 12 | 0   | 69   | -6.544  | -0.002 |
| 24 | base | 11 | 0   | 35   | -51.205 | -0.006 |
| 24 | oct(ols) | 11 | 0   | 60   | -29.615 | -0.001 |
| 24 | ite(struc<sub>cs</sub>, ols<sub>te</sub>) | 11 | 0   | 33   | -10.363 | 0.000  |
| 24 | ite(ols<sub>cs</sub>, struc<sub>te</sub>) | 6  | 0   | 28   | -8.915  | -0.008 |
| 24 | oct(struc) | 4  | 0   | 11   | -1.020  | -0.026 |
| 24 | te(ols<sub>2</sub>) + cs(ols) | 10 | 0   | 72   | -15.037 | -0.018 |
| 24 | te(struc<sub>2</sub>) + cs(ols) | 10 | 0   | 71   | -14.791 | -0.005 |
| 24 | te(ols<sub>2</sub>) + cs(struc) | 10 | 0   | 43   | -12.969 | -0.005 |
| 24 | te(struc<sub>2</sub>) + cs(struc) | 5  | 0   | 31   | -7.042  | -0.002 |
Table OA.2: Number of replications with at least one negative reconciled forecast. All temporal aggregation orders, forecast horizon: operating day.

| $k$ | comb | $\#L_1$ | $\#L_2$ | $\#L_3$ | $\#\text{tot}$ | $\%\text{tot}$ |
|-----|------|---------|---------|---------|----------------|----------------|
| 1   | base | 1       | 5       | 0       | 6              | 100.00%        |
|     | oct(ols) | 1      | 5       | 318     | 324            | 100.00%        |
|     | ite(ols, struc) | 1      | 5       | 318     | 324            | 100.00%        |
|     | ite(ols,ols) | 1      | 5       | 318     | 324            | 100.00%        |
|     | oct(struc) | 1      | 5       | 318     | 324            | 100.00%        |
|     | te(ols)+cs(ols) | 1      | 5       | 318     | 324            | 100.00%        |
|     | te(ols)+cs(struc) | 1      | 5       | 318     | 324            | 100.00%        |
|     | te(ols)+cs(struc) | 1      | 5       | 318     | 324            | 100.00%        |
|     | te(ols)+cs(struc) | 1      | 5       | 318     | 324            | 100.00%        |
| 2   | base | 1       | 5       | 318     | 324            | 100.00%        |
|     | oct(ols) | 1      | 5       | 318     | 324            | 100.00%        |
|     | ite(struc,ols) | 1      | 5       | 318     | 324            | 100.00%        |
|     | ite(ols,struc) | 1      | 5       | 318     | 324            | 100.00%        |
|     | oct(struc) | 1      | 5       | 318     | 324            | 100.00%        |
|     | te(ols)+cs(ols) | 1      | 5       | 318     | 324            | 100.00%        |
|     | te(ols)+cs(struc) | 1      | 5       | 318     | 324            | 100.00%        |
|     | te(ols)+cs(struc) | 1      | 5       | 318     | 324            | 100.00%        |
| 3   | base | 1       | 5       | 318     | 324            | 100.00%        |
|     | oct(ols) | 1      | 5       | 318     | 324            | 100.00%        |
|     | ite(struc,ols) | 1      | 5       | 318     | 324            | 100.00%        |
|     | ite(ols,struc) | 1      | 5       | 318     | 324            | 100.00%        |
|     | oct(struc) | 1      | 5       | 318     | 324            | 100.00%        |
|     | te(ols)+cs(ols) | 1      | 5       | 318     | 324            | 100.00%        |
|     | te(ols)+cs(struc) | 1      | 5       | 318     | 324            | 100.00%        |
|     | te(ols)+cs(struc) | 1      | 5       | 318     | 324            | 100.00%        |
| 4   | base | 1       | 5       | 318     | 324            | 100.00%        |
|     | oct(ols) | 1      | 5       | 318     | 324            | 100.00%        |
|     | ite(struc,ols) | 1      | 5       | 318     | 324            | 100.00%        |
|     | ite(ols,struc) | 1      | 5       | 318     | 324            | 100.00%        |
|     | oct(struc) | 1      | 5       | 318     | 324            | 100.00%        |
|     | te(ols)+cs(ols) | 1      | 5       | 318     | 324            | 100.00%        |
|     | te(ols)+cs(struc) | 1      | 5       | 318     | 324            | 100.00%        |
|     | te(ols)+cs(struc) | 1      | 5       | 318     | 324            | 100.00%        |

Continued on next page
| $k$ | comb          | $\#L_1$ | $\#L_2$ | $\#L_3$ | $\#\text{tot}$ | $\%L_1$ | $\%L_2$ | $\%L_3$ | $\%\text{tot}$ |
|-----|--------------|---------|---------|---------|----------------|---------|---------|---------|----------------|
| 6   | base         | 1       | 5       | 318     | 324           | 100.00% | 100.00% | 100.00% | 100.00%        |
|     | oct(ols)     | 1       | 5       | 318     | 324           | 100.00% | 100.00% | 100.00% | 100.00%        |
|     | ite(structcs, ols) | 1     | 5       | 318     | 324           | 100.00% | 100.00% | 100.00% | 100.00%        |
|     | ite(olscs, structcs) | 1   | 5       | 318     | 324           | 100.00% | 100.00% | 100.00% | 100.00%        |
|     | oct(struct)  | 1       | 5       | 318     | 324           | 100.00% | 100.00% | 100.00% | 100.00%        |
|     | te(ols2)+cs(ols) | 1     | 5       | 318     | 324           | 100.00% | 100.00% | 100.00% | 100.00%        |
|     | te(struct2)+cs(ols) | 1   | 5       | 318     | 324           | 100.00% | 100.00% | 100.00% | 100.00%        |
|     | te(ols2)+cs(struct) | 1   | 5       | 318     | 324           | 100.00% | 100.00% | 100.00% | 100.00%        |
| 8   | base         | 1       | 3       | 318     | 322           | 100.00% | 60.00%  | 100.00% | 99.38%        |
|     | oct(ols)     | 1       | 5       | 318     | 324           | 100.00% | 100.00% | 100.00% | 100.00%        |
|     | ite(structcs, ols) | 1     | 5       | 318     | 324           | 100.00% | 100.00% | 100.00% | 100.00%        |
|     | ite(olscs, structcs) | 1   | 5       | 318     | 324           | 100.00% | 100.00% | 100.00% | 100.00%        |
|     | oct(struct)  | 1       | 5       | 318     | 324           | 100.00% | 100.00% | 100.00% | 100.00%        |
|     | te(ols2)+cs(ols) | 1     | 5       | 318     | 324           | 100.00% | 100.00% | 100.00% | 100.00%        |
|     | te(struct2)+cs(ols) | 1   | 5       | 318     | 324           | 100.00% | 100.00% | 100.00% | 100.00%        |
|     | te(ols2)+cs(struct) | 1   | 5       | 318     | 324           | 100.00% | 100.00% | 100.00% | 100.00%        |
|     | te(struct2)+cs(struct) | 1 | 5     | 318     | 324           | 100.00% | 100.00% | 100.00% | 100.00%        |
| 12  | base         | 0       | 2       | 135     | 137           | 0.00%   | 40.00%  | 42.45%  | 42.28%        |
|     | oct(ols)     | 0       | 1       | 162     | 163           | 0.00%   | 20.00%  | 50.94%  | 50.31%        |
|     | ite(structcs, ols) | 0     | 1       | 99      | 100           | 0.00%   | 20.00%  | 31.13%  | 30.86%        |
|     | ite(olscs, structcs) | 0   | 1       | 67      | 68            | 0.00%   | 20.00%  | 21.07%  | 20.99%        |
|     | oct(struct)  | 0       | 0       | 28      | 28            | 0.00%   | 0.00%   | 8.81%   | 8.64%         |
|     | te(ols2)+cs(ols) | 0     | 2       | 172     | 174           | 0.00%   | 40.00%  | 54.09%  | 53.70%        |
|     | te(struct2)+cs(ols) | 0   | 2       | 156     | 158           | 0.00%   | 40.00%  | 49.06%  | 48.77%        |
|     | te(ols2)+cs(struct) | 0   | 1       | 126     | 127           | 0.00%   | 20.00%  | 39.62%  | 39.20%        |
|     | te(struct2)+cs(struct) | 0 | 1      | 105     | 106           | 0.00%   | 20.00%  | 33.02%  | 32.72%        |
| 24  | base         | 0       | 2       | 58      | 60            | 0.00%   | 40.00%  | 18.24%  | 18.52%        |
|     | oct(ols)     | 0       | 1       | 134     | 135           | 0.00%   | 20.00%  | 42.14%  | 41.67%        |
|     | ite(structcs, ols) | 0     | 1       | 72      | 73            | 0.00%   | 20.00%  | 22.64%  | 22.53%        |
|     | ite(olscs, structcs) | 0   | 1       | 47      | 48            | 0.00%   | 20.00%  | 14.78%  | 14.81%        |
|     | oct(struct)  | 0       | 0       | 17      | 17            | 0.00%   | 0.00%   | 5.35%   | 5.25%         |
|     | te(ols2)+cs(ols) | 0     | 1       | 121     | 122           | 0.00%   | 20.00%  | 38.05%  | 37.65%        |
|     | te(struct2)+cs(ols) | 0   | 1       | 102     | 103           | 0.00%   | 20.00%  | 32.08%  | 31.79%        |
|     | te(ols2)+cs(struct) | 0   | 1       | 80      | 81            | 0.00%   | 20.00%  | 25.16%  | 25.00%        |
|     | te(struct2)+cs(struct) | 0 | 1      | 41      | 42            | 0.00%   | 20.00%  | 12.89%  | 12.96%        |
Figure OA.1: Comparison of nRMSE(%) between PERS$^{BU}$ and not negative oct-struc (top panel), and between 3TIER$^{BU}$ and not negative oct-struc (bottom panel). The black line represents the bisector, where the nRMSE’s for both approaches are equal. On the top-left (bottom-right) corner of each graph, the percentage of points above (below) the bisector is reported.
A general overview of the accuracy of the considered procedures at any temporal granularity is given by table OA.3 while a visual confirmation of the performance of ct-struc is shown by the MCB graphs in figure 5 in the main paper. The ct-struc daily reconciled forecasts show the best performance, significantly different from all the other procedures, while at hourly level it ranks second, the best being t-struc$_2$ + cs-struc. However, t-struc$_2$ + cs-struc produces temporally incoherent forecasts, and a punctual assessment of the performance of these two procedures in terms of nRMSE (figure OA.2) confirms that their forecasting accuracy is about the same.

Table OA.3: Forecast accuracies of sequential reconciliation methods and base forecasts in terms of nRMSE(%). Free reconciliation approaches considered by Yagli et al. (2019), Tables 2, 3, p. 395. All temporal aggregation orders, forecast horizon: operating day.

| $\mathcal{L}$ | comb  | 1 | 2 | 3 | 4 | 6 | 8 | 12 | 24 |
|--------------|-------|---|---|---|---|---|---|----|----|
| 1 PERS$_bu$  | 34.62 | 34.14 | 33.35 | 33.11 | 29.45 | 32.23 | 20.83 | 20.23 |
| 1 3TIER$_bu$ | 26.03 | 25.34 | 24.56 | 23.31 | 22.09 | 19.89 | 15.70 | 12.48 |
| 1 base       | 27.85 | 27.55 | 28.31 | 29.69 | 28.01 | 35.32 | 21.04 | 18.17 |
| 1 oct(ols)   | 30.63 | 29.96 | 28.29 | 28.85 | 25.26 | 28.12 | 18.39 | 17.91 |
| 1 ite(struct,ols) | 28.74 | 28.25 | 27.14 | 27.38 | 24.34 | 26.64 | 17.82 | 17.33 |
| 1 ite(ols,struct) | 28.26 | 27.84 | 27.09 | 27.00 | 24.17 | 26.24 | 17.46 | 16.96 |
| 1 oct(struct) | 26.71 | 26.35 | 25.67 | 25.57 | 23.03 | 24.80 | 16.74 | 16.24 |
| 1 te(ols)+cs(ols) | 27.80 | 27.41 | 27.91 | 29.37 | 27.39 | 34.11 | 20.80 | 17.96 |
| 1 te(struc)$_2$+cs(struct) | 26.17 | 25.86 | 25.83 | 26.75 | 25.13 | 28.93 | 19.11 | 17.24 |
| 1 te(struc)$_2$+cs(struc) | 26.17 | 25.86 | 25.83 | 26.75 | 25.13 | 28.93 | 19.11 | 17.24 |
| 2 PERS$_bu$  | 43.15 | 42.42 | 41.27 | 41.27 | 36.10 | 38.97 | 25.66 | 24.57 |
| 2 3TIER$_bu$ | 33.95 | 32.94 | 31.90 | 30.62 | 28.02 | 26.99 | 20.80 | 17.95 |
| 2 base       | 34.24 | 34.04 | 33.79 | 35.01 | 33.11 | 40.51 | 25.16 | 20.94 |
| 2 oct(ols)   | 39.17 | 37.97 | 35.24 | 36.12 | 31.27 | 34.80 | 22.58 | 21.74 |
| 2 ite(struct,ols) | 35.48 | 34.74 | 33.24 | 34.44 | 29.65 | 32.16 | 21.66 | 20.82 |
| 2 ite(ols,struct) | 35.19 | 34.57 | 33.48 | 33.28 | 29.74 | 31.95 | 21.47 | 20.61 |
| 2 oct(struct) | 33.11 | 32.54 | 31.60 | 31.37 | 28.19 | 30.06 | 20.49 | 19.64 |
| 2 te(ols)$_2$+cs(ols) | 33.40 | 33.07 | 32.82 | 32.64 | 30.70 | 35.26 | 23.20 | 20.62 |
| 2 te(struc)$_2$+cs(struc) | 32.44 | 32.07 | 31.80 | 32.54 | 30.81 | 34.14 | 22.59 | 19.96 |
| 3 PERS$_bu$  | 59.75 | 56.81 | 54.23 | 52.87 | 46.82 | 49.07 | 33.12 | 30.65 |
| 3 3TIER$_bu$ | 53.46 | 50.57 | 48.33 | 46.19 | 41.36 | 40.72 | 29.38 | 25.19 |
| 3 base       | 53.46 | 47.22 | 44.87 | 44.30 | 39.68 | 42.66 | 30.92 | 25.82 |
| 3 oct(ols)   | 51.54 | 48.48 | 45.06 | 45.17 | 39.41 | 42.39 | 28.50 | 26.73 |
| 3 ite(struct,ols) | 49.32 | 46.47 | 43.80 | 43.45 | 38.38 | 40.64 | 27.92 | 26.14 |
| 3 ite(ols,struct) | 48.20 | 45.50 | 43.45 | 42.51 | 37.83 | 39.64 | 27.28 | 25.46 |
| 3 oct(struct) | 46.74 | 44.02 | 42.05 | 41.07 | 36.07 | 38.16 | 26.56 | 24.73 |
| 3 te(ols)$_2$+cs(ols) | 48.52 | 46.20 | 45.02 | 45.58 | 42.49 | 50.21 | 31.50 | 26.71 |
| 3 te(struc)$_2$+cs(ols) | 47.77 | 45.52 | 44.64 | 44.96 | 42.07 | 49.67 | 31.19 | 26.36 |
| 3 te(ols)$_2$+cs(struct) | 47.76 | 45.19 | 43.49 | 43.72 | 39.27 | 43.19 | 29.26 | 25.98 |
| 3 te(struc)$_2$+cs(struct) | 46.25 | 43.69 | 42.24 | 42.12 | 38.05 | 41.33 | 28.38 | 25.03 |
Figure OA.2: Comparison of nRMSE(%) between t-struc + cs-struc and ct-struc (top panel), and the not negative version (bottom pannel). The black line represents the bisector, where the nRMSE's for both approaches are equal.
Figure OA.3: MCB-Nemenyi test on Yang cross-temporal reconciliation approaches with operating day forecast horizon, $L_0, L_1, L_2$ levels (324 series). Top panel: hourly forecasts; Bottom panel: daily forecasts. The mean rank of each approach is displayed to the right of their names. If the intervals of two forecast reconciliation approaches do not overlap, this indicates a statistically different performance. Thus, approaches that do not overlap with the green interval are considered significantly worse than the best, and vice-versa.