UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Scienze Statistiche
Corso di Dottorato di Ricerca in Scienze Statistiche
Ciclo XXXVI

Forecast reconciliation: Methodological issues and applications

Coordinatore del Corso: Prof. Nicola Sartori

Supervisore: Prof. Tommaso Di Fonzo

Co-supervisori: Prof. George Athanasopoulos, Prof. Rob J Hyndman

Dottorando: Daniele Girolimetto

10 January 2024, Padova

Abstract

In recent years, there has been growing interest in the development of reconciliation techniques for forecasting multiple time series, both in academia and industries. Time series data are widely employed to predict the future behavior of phenomena such as retail sales, tourist flows, macroeconomic and climatic trends. However, forecasts of different series, linked by some constraints and derived from different sources, often exhibit discrepancies that compromise the reliability of the results. Forecast reconciliation is a post-forecasting process aimed at improving the quality of forecasts for a system of linearly constrained multiple time series. Despite notable progress, these approaches encounter various challenges, both methodological and practical. A significant issue lies in ensuring the non-negativity of reconciled forecasts, a critical consideration when forecasting variables that cannot assume negative values, such as sales or tourist flows. Another challenge involves extending reconciliation methods beyond genuine hierarchical/grouped time series, where the roles of variables are uniquely identified. This also includes more complex scenarios, combining the cross-sectional (involving multiple time series at the same frequency) and temporal (a single series at multiple frequencies) frameworks to obtain cross-temporally coherent forecasts. Additionally, dealing with large volumes of data and selecting efficient algorithms are crucial factors that must be addressed to ensure the effectiveness of reconciliation processes.

This thesis aims to provide solutions, both from a methodological and computational point of view. To achieve this goal, each chapter of the thesis will present one or two specific applications, illustrating the effectiveness of the proposed methodologies in real-world contexts. Chapter 1 introduces the concept of simultaneous reconciliation of cross-sectional and temporal forecasts. This chapter presents two novel contributions: an expression for cross-temporal point forecasts and an iterative procedure that alternates reconciliation along a single dimension until convergence is achieved. Chapter 2 focuses on the application of forecast reconciliation of hierarchical photovoltaic (PV)

power generation for a simulated PV dataset in California. Various aspects are investigated, including the non-negativity problem and forecast coherency through cross-temporal reconciliation approaches. Chapter 3 extends the approach based on forecast combination for reconciling a hierarchy, known as LCC. This method minimizes a quadratic loss function subject to an exogenous constraint imposed by the base forecast of the higher-level series within the hierarchy, whose value remains unchanged. An alternative approach is proposed, which modifies this constraint by incorporating endogenous constraints within the same context. Chapter 4 explores the reconciliation of generally constrained multiple time series, thereby extending the results obtained thus far, which have been limited to genuine hierarchical structures. Chapter 5 extends the cross-sectional probabilistic reconciliation approach to the cross-temporal framework. This chapter introduces both parametric and non-parametric approaches to draw base forecasts samples. Finally, in Chapter 6, the reconciliation of realized volatility forecasts using intra-day decompositions is addressed for the first time in a financial framework.

Sommario

Negli ultimi anni, si è registrato un crescente interesse, sia in ambito accademico che professionale, per lo sviluppo di tecniche di riconciliazione finalizzate alle previsioni di serie storiche multiple. Le serie storiche sono ampiamente impiegate per prevedere il comportamento di fenomeni come le vendite al dettaglio, i flussi turistici, eventi climatici e variabili macroeconomiche ed energetiche. Tuttavia, le previsioni provenienti da diverse serie, ottenute da fonti o metodi differenti, spesso presentano discrepanze, compromettendo l'affidabilità complessiva dei risultati. In questo contesto, la riconciliazione delle previsioni viene definito come un processo di post-previsione mirato a migliorare la qualità delle previsioni per sistemi di serie multiple linearmente vincolate. Nonostante il crescente studio di queste tecniche, molte sfide sia metodologiche che applicative devono ancora essere affrontate. Un problema chiave è rappresentato dalla necessità di garantire la non negatività delle previsioni riconciliate, particolarmente cruciale in contesti pratici come le previsioni di vendite o flussi turistici. Un'altra sfida è rappresentata dalla generalizzazione a scenari più complessi, non limitati a serie storiche strettamente gerarchiche o raggruppate dove il ruolo delle variabili è univoco. Ciò implica anche la necessità di unificare l'ambito cross-sezionale (più serie alla stessa frequenza) e temporale (una singola serie a più frequenze) per garantire previsioni coerenti rispetto ad entrambe le dimensioni (coerenza cross-temporale). Inoltre, l'elaborazione di grandi volumi di dati e la scelta di algoritmi efficienti sono tutti elementi critici da prendere in considerazione per assicurare l'efficacia del processo di riconciliazione.

Questa tesi mira a fornire soluzioni a tali problemi, sia dal punto di vista metodologico che empirico. Per raggiungere questo obiettivo, ogni capitolo della tesi presenterà una o due applicazioni specifiche, illustrando l'efficacia delle proposte metodologiche in contesti reali. Nel Capitolo 1 viene studiato il problema della riconciliazione simultanea cross-temporale. In particolare vengono forniti due nuovi risultati: un'espressione per le previsioni riconciliate puntuali, e una procedura iterativa che alterna la riconciliazione lungo una singola dimensione fino alla convergenza. Il Capitolo 2 presenta la

riconciliazione spazio-temporale delle previsioni dell'energia solare fotovoltaica prodotta da impianti facenti parte di una gerarchia simulata californiana. Vengono affrontate varie tematiche tra cui il problema della non negatività e della coerenza delle previsioni utilizzando diversi approcci di riconciliazione cross-temporale. Nel Capitolo 3 il problema della riconciliazione cross-sezionale viene reinterpretato in termini di combinazione delle previsioni per la riconciliazione di una gerarchia, dando origine ad un approccio di *Level Coherent Combination* (LCC). Questo metodo minimizza una funzione di perdita quadratica, con un vincolo esogeno dato dalla previsione di base della serie di livello superiore della gerarchia, il cui valore rimane inalterato. Viene inoltre fornita un'alternativa che nello stesso contesto "aggiusta" tale valore utilizzando vincoli endogeni. Nel Capitolo 4 viene studiata la riconciliazione delle previsioni per generiche serie storiche multiple linearmente vincolate, estendendo tutti i risultati ottenuti finora per le sole serie strettamente gerarchiche/raggruppate. Il Capitolo 5 estende l'approccio di riconciliazione probabilistica cross-sezionale al caso cross-temporale, secondo approcci parametrici e non per ottenere campioni dalla distribuzione delle previsioni di base. Infine, nel Capitolo 6 viene affrontata per la prima volta la riconciliazione delle previsioni giornaliere della volatilità realizzata in ambito finanziario, sfruttando varie decomposizioni intra-giornaliere.

Contents

List of Figures

List of Tables

Introduction

Overview

Forecasting is a fundamental tool for decision-making in many fields (Petropoulos *et al.*, 2022), including finance, economics, marketing, energy and supply chain management. Accurate predictions help organizations optimize operations, allocate resources effectively, and make strategic decisions, but it is often challenging due to the complexity of the data. In such a framework, mostly when we face a prediction problem involving data that follow some linear constraints (e.g., *linearly constrained multiple time series*), forecasting may benefit from the data organization structure. For example, the number of tourists who visit a particular country is made up of the tourists who visit each state within that country. This is the simplest structure of a linearly constrained multiple time series and it is called *hierarchical time series*. It is also possible to further break down this hierarchy by grouping tourists according to different variables, such as their reason for travel (e.g., visiting friends, business, or vacation), which is referred to as *grouped time series*.

In general, most approaches involve forecasting the individual components of the system, often failing to satisfy the constraints observed in the actual data. Forecast reconciliation is a post-forecasting process intended to improve the quality of (base) forecasts for a system of linearly constrained multiple time series. Base forecasts may be produced by separate organizational silos or using distinct models for each variable (Chase, 2013). Therefore, the reconciliation process involves combining base forecasts for related variables in a way that satisfies the constraints between them, thereby producing *coherent* and (possibly) reliable forecasts for the entire system. This approach has gained increasing attention in recent years due to its potential to improve forecast accuracy and reduce inconsistencies in the predictions of related variables, making it a crucial tool for decision-making. Temporal coherence as well (Athanasopoulos *et al.*, 2017) is another crucial aspect that may help organizations to align their forecasting efforts effectively. The temporal approach involves reconciling forecasts that are generated at different

time horizons, such as monthly, quarterly, or annually. For instance, a retail company may need to reconcile monthly sales forecasts with quarterly forecasts to guarantee that they are consistent.

Classical reconciliation methods aim to resolve the issue of incoherent forecasts in a cross-sectional hierarchy by forecasting only one level and utilizing these forecasts to generate predictions for the remaining series. One of the most widely used approaches is the bottom-up (Dunn *et al.*, 1976), which aggregates the forecasts for the most disaggregated level. On the other hand, the top-down approach (Gross and Sohl, 1990) forecasts the most aggregated level and disaggregates it. A hybrid approach, known as middle-out reconciliation (Athanasopoulos *et al.*, 2009), combines both methods by selecting an intermediate level and using the top-down approach for lower levels and the bottom-up for higher levels. However, it is worth noting that all these approaches ignore useful information available at other levels, as highlighted by Pennings and Van Dalen (2017). To overcome the limitations imposed by these approaches, in recent years the concept of regression-based reconciliation has been introduced.

Since the 1940s, least squares reconciliation has been implemented in several contexts outside of the forecasting domain. Stone *et al.* (1942) developed a constrained estimation method for balancing national accounts, which involves a weighted linear combination of initial estimates. Richard Stone won the 1984 Nobel Prize in Economics with works related to this result. Byron (1978, 1979) further formalized and extended Stone's work by using more computationally efficient techniques. Later, the same idea was applied in other national statistics offices (Dagum and Cholette, 2006, chp. 11). For example, dealing with a table of directly seasonally adjusted time series incoherence is usually observed, so some least squares-based solutions are used to solve this problem (Di Fonzo and Marini, 2011; Corona *et al.*, 2021; Quenneville and Fortier, 2012; Di Fonzo and Marini, 2015). Temporal reconciliation is also of interest to national statistics offices using indirect temporal disaggregation techniques to estimate quarterly or monthly aggregates coherent with their temporally aggregated counterparts (e.g, annual or quarterly). In this case, it is wished that monthly or quarterly estimates sum to the annual estimates (Chow and Lin, 1971; Chen *et al.*, 2018). In this context, Di Fonzo (1990) (see also Rossi, 1982) introduces simultaneous least squares reconciliation of time series estimates in both cross-sectional and temporal dimensions. In addition, this concept has been developed in other domains, e.g. chemical (Romagnoli and Sánchez, 2000) and engineering (Simon, 2006, 2010; Simon and Tien Li Chia, 2002).

One of the earliest works that applied ordinary least squares (OLS) to forecast reconciliation was the PhD thesis of Ahmed (2009), from which the seminal paper by

Hyndman *et al.* (2011) has been derived. However, the OLS approach weights all series equally, whether they are aggregates or disaggregates, and whether their base forecasts are "good" or "bad". To address this issue, Hyndman *et al.* (2016) propose a weighted least squares (WLS) solution, where they use the variances of the base forecasts for each serie as weighting factors. Wickramasuriya *et al.* (2019) provide theoretical insights into the problem by taking an optimization approach rather than a regression approach. They formulate the problem as minimizing the trace (MinT) of the covariance matrix of the reconciled forecast errors, and the solution can be seen as a feasible generalized least squares (GLS) estimation, with WLS and OLS as special cases. van Erven and Cugliari (2015) propose a game-theoretic approach and show that the solution to their minimax problem is equivalent to the GLS when a L_2 loss is used. Moreover, Panagiotelis *et al.* (2021) unify these results providing a geometric interpretation of the main forecast reconciliation results.

Focusing on temporal hierarchies, Kourentzes *et al.* (2014) propose the Multiple Aggregation Prediction Algorithm (MAPA), which combines independent models for a time series at multiple temporal aggregation levels using a state-space model (e.g., ETS by Hyndman, 2008). Athanasopoulos *et al.* (2017) exploit the MAPA's idea of multiple temporal aggregation levels and introduce temporal hierarchies. One significant advantage over MAPA is the absence of model constraints, allowing for the use of diverse forecasting models/methods at each level. Achieving cross-temporal coherence in forecasts across both the temporal and cross-sectional dimensions has been a challenge that has thus far been tackled by sequential approaches that focus on each dimension independently (Kourentzes and Athanasopoulos, 2019; Yagli *et al.*, 2019; Punia *et al.*, 2020; Spiliotis *et al.*, 2020).

Probabilistic forecasts differ from point forecasts in their ability to quantify prediction uncertainty, rendering them a crucial element in making informed decisions (Abramson and Clemen, 1995; Gneiting and Katzfuss, 2014). Thanks to their nature, they are widely employed in various fields, including economics (Zarnowitz and Lambros, 1987; Clements, 2004, 2018; Liu *et al.*, 2021), meteorology (Gneiting *et al.*, 2008; Leutbecher and Palmer, 2008; Pinson *et al.*, 2009; McLean Sloughter *et al.*, 2013; Scheuerer and Hamill, 2015; Leutbecher, 2019), energy (Jeon and Taylor, 2012; Wytock and Kolter, 2013; Hong *et al.*, 2016; Ben Taieb *et al.*, 2016), and retail (Kolassa, 2016; Böse *et al.*, 2017; Berry *et al.*, 2020). Shang and Hyndman (2017) make an initial effort towards probabilistic forecast by reconciling the quantiles, instead of point forecasts, to generate prediction intervals. Ben Taieb *et al.* (2017) present a bottom-up method that incorporates top-level information into the algorithm to adjust the mean of each series, which

was later extended in Ben Taieb *et al.* (2021). However, this method requires the size of the samples drawn from the forecast distribution and the training data to be equal, making it unsuitable for limited training data. To overcome this limitation, Panamtash and Zhou (2018) and Zhao *et al.* (2019) propose estimating predictive quantiles directly via quantile regression. Jeon *et al.* (2019) propose a two-step temporal reconciliation approach for producing probabilistically coherent forecasts, which involves drawing a sample from the forecast distribution of each series pre-multiplied by a projection matrix to obtain a sample from the coherent multivariate distribution. Panagiotelis *et al.* (2023) provide formal definitions for coherence and reconciliation in the cross-sectional framework, including parametric and non-parametric approaches. Moreover, Wickramasuriya (2024) thoroughly examines the assumption of Gaussianity for reconciliation. Corani *et al.* (2021) propose a cross-sectional Bayesian algorithm that focuses on probabilistic forecast reconciliation. Recently, this approach has also been extended to discrete data by Corani *et al.* (2023) and Zambon *et al.* (2024) in the cross-sectional and temporal framework, respectively.

For a comprehensive and in-depth review of the historical and contemporary developments in the field of forecast reconciliation, Athanasopoulos *et al.* (2023) provide a very recent and valuable resource.

Main contributions of the thesis

This thesis focuses on several aspects of point and probabilistic forecast reconciliation, filling several previously neglected or minimally explored features of both cross-sectional and temporal cases. By establishing a unified cross-temporal framework, it aims to comprehensively encompass and consolidate both aspects, while also standardizing the notation.The thesis consists of six main chapters.

In Chapter 1, the development of optimal combination cross-temporal forecast reconciliation is proposed as a novel approach. This method extends existing reconciliation techniques by considering both cross-sectional and temporal dimensions simultaneously. Moreover, an iterative cross-temporal forecast reconciliation procedure is introduced, building upon previous heuristic methods (Kourentzes and Athanasopoulos, 2019) to enhance accuracy and coherency. Finally, the proposed approach is evaluated using a forecasting experiment with the Australian National Accounts dataset, demonstrating superior performance compared to other reconciliation procedures in terms of accuracy and coherency, particularly in improving GDP forecasts.

Chapter 2 focuses on solar power data and presents several significant contributions. It introduces the concept of cross-temporal forecast reconciliation, which combines geographical and temporal hierarchies to generate accurate forecasts of photovoltaic (PV) generated power. Different reconciliation procedures (e.g., Yagli *et al.*, 2019 and Kourentzes and Athanasopoulos, 2019 among others) are discussed, providing new methodological and computational insights dealing with non-negativity and coherency problems. Additionally, a comparative performance analysis of different reconciliation approaches is conducted, helping practitioners select the most suitable method for their specific needs. These contributions advance the understanding and application of solar power forecasting, facilitating the integration of solar energy into electrical power grids and supporting the transition towards sustainable energy sources.

Chapter 3 provides a detailed mathematical derivation and analysis of the Level-l Conditional Coherent (L_lCC) forecast reconciliation formula proposed by Hollyman *et al.* (2021), as linearly constrained minimization of a quadratic loss function, with an exogenous constraint given by the base forecast of the top level series of the hierarchy, which is not revised. Then, the concept of endogenous constraints is introduced, which enables level conditional reconciliation by coherently revising both the top and bottom level time series forecasts. Furthermore, the L_lCC procedure does not guarantee the non-negativity of the reconciled forecasts, which in many situations is a natural property that forecasts should satisfy. Therefore, some solutions are proposed to address this issue in the concluding section. Finally, two forecasting experiments are considered to evaluate the performance of various cross-sectional forecast combination-based point forecast reconciliation procedures in a fair setting. In this framework, due to the crucial role played by the (possibly different) models used to compute the base forecasts, a reinterpretation of the combined conditional coherent reconciliation procedure (Hollyman *et al.*, 2021) as a forecast pooling approach is developed.

Chapter 4 presents several key contributions to unify different cross-sectional reconciliation representations. Most of the point and probabilistic regression-based forecast reconciliation results ground on the so called "structural representation" (Hyndman *et al.*, 2011; Panagiotelis *et al.*, 2021) and on the related unconstrained GLS reconciliation formula. However, the structural representation naturally applies to genuine hierarchical/grouped time series, where the top- and bottom-level variables are uniquely identified. When a general linearly constrained multiple time series is considered, the forecast reconciliation is naturally expressed according to a projection approach (Stone *et al.*, 1942; Byron, 1978, 1979; Weale, 1992; Smith *et al.*, 1998). While it is well known that the classic structural reconciliation formula is equivalent to its projection approach

counterpart, so far it is not completely understood if and how a structural-like reconciliation formula may be derived for a general linearly constrained multiple time series. Such an expression would permit to extend reconciliation definitions, theorems and results in a straightforward manner. The chapter shows that for general linearly constrained multiple time series it is possible to express the reconciliation formula according to a "structural-like" approach that keeps distinct free and constrained variables, instead of bottom and upper (aggregated), establish the probabilistic forecast reconciliation framework, and obtain fully reconciled point and probabilistic forecasts for the aggregates of the Australian GDP from income and expenditure sides, and for the European Area GDP disaggregated by income, expenditure and output sides and by 19 countries.

Chapter 5 extends the concept of cross-sectional probabilistic forecast reconciliation to the cross-temporal case. It presents a comprehensive framework for handling both cross-sectional and temporal dimensions simultaneously. Effective methods for generating samples from the base forecast distribution are proposed, considering both parametric and non-parametric approaches. The chapter addresses the challenge of capturing cross-temporal relationships by using multi-step residuals, and introduces overlapping residuals to handle high-dimensionality issues. New shrinkage procedures are proposed to identify an appropriate covariance matrix structure, both in the base and reconciled forecast distributions. Empirical applications demonstrate the effectiveness of the proposed approaches in improving forecast accuracy and coherence in real-world datasets. This chapter significantly extends probabilistic forecast reconciliation to the cross-temporal framework and provides new insights and practical tools for enhanced forecasting processes and decision-making.

In Chapter 6, the forecast reconciliation approach is applied to a classical financial forecasting problem, i.e. volatility forecasting. It proposes a method to improve daily realized volatility (RV) forecasts by utilizing intraday decompositions (Patton and Sheppard, 2015; Bollerslev *et al.*, 2022). The importance of volatility forecasting in various financial areas is emphasized, and different models for RV forecasting are reviewed. The concept of hierarchical forecasting is introduced (bottom-up and regression-based), where daily RV is considered as the top-level series with intraday components. A combined-aggregative forecasting method is presented, which involves segment-level forecasting and the combination of forecasts to enhance the daily RV. Furthermore, an extensive out-of-sample forecasting experiment on the Dow Jones Industrial Average index and 26 stocks demonstrates the improved forecast accuracy achieved by incorporating intraday RV decompositions through reconciliation. Finally, a robustness

analysis is conducted, examining different scenarios and variations to validate the reliability and stability of the proposed methodology. Overall, these contributions enhance the understanding and practice of volatility forecasting in financial econometrics, offering valuable insights for various financial applications such as risk management, asset allocation, hedging, and pricing.

Publications and Conferences

The content presented in each chapter has been published or submitted to several academic journals, as indicated in Table 0.1. In addition, preliminary/complete parts were presented in various conferences. Chapter 1 was presented at the 2021 ASA Joint Statistical Meeting, the 50th Meeting of the Italian Statistical Society, and the 40th IIF International Symposium on Forecasting. Chapters 2 and 3 were presented at the 42nd and 41st IIF International Symposium on Forecasting, respectively. Chapter 4 was presented at the 42nd IIF International Symposium on Forecasting, the 8th International Conference on Time Series and Forecasting, the 51st Meeting of the Italian Statistical Society (Di Fonzo and Girolimetto, 2022b), and the 16th International Conference on Computational and Financial Econometrics. Chapter 5 was presented at the 43rd IIF International Symposium on Forecasting[1] , the 9th International Conference on Time Series and Forecasting, and the 2023 IIF Workshop on Forecast Reconciliation. Finally, Chapter 6 was presented at the Quantitative Finance and Financial Econometrics 2023 conference.

Software and open data

Dedicated GitHub repositories have been set up to store the code used to produce each chapter of this thesis. These repositories contain both the necessary functions and documentation, making it possible to replicate the analyses presented in the thesis. The primary language used for this work was R (R Core Team, 2022).

All reconciliation methodologies used in this thesis were implemented using the R package `FoReco` (Girolimetto and Di Fonzo, 2023a, `https://CRAN.R-project.org/package=FoReco`). This package provides both classical (bottom-up and top-down) and modern (optimal and heuristic approaches) forecast reconciliation procedures for linearly constrained time series data across temporal, cross-sectional, and cross-temporal frameworks. The documentation and the manual of the package are available

[1]Winner of the best student presentation award.

Chapter	Title, Journal, Status and Co-author(s)	
1	**Title**	Cross-temporal forecast reconciliation: Optimal combination method and heuristic alternatives
	Journal	*International Journal of Forecasting*
	Status	Published (2023, volume 39, issue 1, p. 39–57)
	Co-author	Tommaso Di Fonzo
	DOI	`10.1016/j.ijforecast.2021.08.004`
2	**Title**	Spatio-temporal reconciliation of solar forecasts
	Journal	*Solar Energy*
	Status	Published (2023, volume 251, p. 13–29)
	Co-author	Tommaso Di Fonzo
	DOI	`10.1016/j.solener.2023.01.003`
3	**Title**	Forecast combination-based forecast reconciliation: Insights and extensions
	Journal	*International Journal of Forecasting*
	Status	Accepted (2022)
	Co-author	Tommaso Di Fonzo
	DOI	`10.1016/j.ijforecast.2022.07.001`
4	**Title**	Point and probabilistic forecast reconciliation for general linearly constrained multiple time series
	Journal	*Statistical Methods & Applications*
	Status	Accepted (2023)
	Co-author	Tommaso Di Fonzo
	DOI	`10.1007/s10260-023-00738-6`
5	**Title**	Cross-temporal probabilistic forecast reconciliation: methodological and practical issues
	Journal	*International Journal of Forecasting*
	Status	Accepted (2023)
	Co-authors	George Athanasopoulos, Tommaso Di Fonzo, Rob J. Hyndman
	DOI	`10.1016/j.ijforecast.2023.10.003`
6	**Title**	Exploiting intraday decompositions in realized volatility forecasting: a forecast reconciliation approach
	Pre-print	`arXiv:2306.02952`
	Status	Submitted (2023)
	Co-authors	Massimiliano Caporin, Tommaso Di Fonzo

Table 0.1: *Title, journal/pre-print, status (published, accepted or submitted), and co-authors of research articles extracted from each chapter.*

at `https://danigiro.github.io/FoReco/` and `https://cloud.r-project.org/web/packages/FoReco/FoReco.pdf`, respectively.

Publicly available GitHub repositories (`https://github.com/danigiro`) have been set up specifically for each chapter, which provide the code and documentation required to replicate the analyses presented:

- **Chapter 1**: `danigiro/AusGDP_IJF`
- **Chapter 2**: `danigiro/SE_PV324`
- **Chapter 3**: `danigiro/IJF_VN525` and `danigiro/IJF_AUSGDP`
- **Chapter 4**: `danigiro/mtsreco`
- **Chapter 5**: `danigiro/ctprob`

In addition, supplementary material (extended tables and figures) are provided for each chapter in a separate online file (`https://github.com/danigiro/phd-thesis`):

- **Chapter 1**: `online_appendix/oa_chp1.pdf`
- **Chapter 2**: `online_appendix/oa_chp2.pdf`
- **Chapter 3**: `online_appendix/oa_chp3.pdf`
- **Chapter 4**: `online_appendix/oa_chp4.pdf`
- **Chapter 5**: `online_appendix/oa_chp5.pdf`
- **Chapter 6**: `online_appendix/oa_chp6.pdf`

Chapter 1

Optimal and heuristic methods for cross-temporal reconciliation

1.1 Introduction

In several operational fields, decisions to be successful require the support of accurate, detailed but also coherent forecasts. Forecasts are coherent when the predicted values at the disaggregate and aggregate scales are equal when brought to the same level. For example, temporally coherent monthly predictions sum up to annual ones and similarly geographically coherent regional predictions add up to country level ones. Such a coherence is an important qualifier for forecasts, so as to support aligned decision making across different planning units and horizons, while avoiding that different decision making units plan on different views of the future. For example, Kourentzes and Athanasopoulos (2019) generate forecasts for Australian domestic tourism that are coherent across multiple geographical divisions (regions, zones, states, and whole country), but are also coherent across time (at monthly, bi-monthly, quarterly, 4-monthly, 6-monthly, and annual levels), i.e. for different planning horizons.

As a matter of fact, in the growing literature on hierarchical forecast reconciliation the cross-sectional (contemporaneous) and temporal coherency dimensions are mostly considered in separate ways: either 'time-by-time' cross-sectional forecast reconciliation for a n-dimensional time series (Hyndman *et al.*, 2011, 2022), or temporal coherency for the forecasts of a single variable for different time frequencies (Athanasopoulos *et al.*, 2017; Hyndman and Kourentzes, 2018). The issue of getting coherent forecasts along both cross-sectional and temporal dimensions (i.e., cross-temporal coherency) has been dealt with by Yagli *et al.* (2019) and Spiliotis *et al.* (2020). However, as far as we know, the procedure proposed by Kourentzes and Athanasopoulos (2019) is the only

one able to simultaneously fulfill both cross-sectional and temporal coherency in the final reconciled point forecasts at any considered aggregation dimension.

Whereas the most recent forecast reconciliation procedures for each single coherence dimension are based on some optimality criterion (van Erven and Cugliari, 2015; Wickramasuriya *et al.*, 2019), the cross-temporal reconciliation procedure by Kourentzes and Athanasopoulos (2019) can be considered as an heuristic with a simple and effective structure, i.e. an approach that employs a practical method that is not guaranteed to be optimal, but which is nevertheless sufficient for reaching an immediate goal, which in our case is the coherency along all the considered dimensions of the reconciled forecasts. This fact is probably because of the consideration that in a cross-temporal forecast reconciliation framework the complexity and dimensions of the problem grow very quickly along with the requested computational time and memory (Kourentzes and Athanasopoulos, 2019; Nystrup *et al.*, 2020). This is certainly true, but appropriate, thorough use of some well known linear algebra tools, and the expanding computation facilities, in terms of both calculation power and memory, makes it feasible to look for the optimal solution (in the least squares sense), expanding the field of application of the 'forecast reconciliation methodology' to simultaneously encompass both contemporaneous (cross-sectional) and temporal coherency dimensions.

Bottom-up and top-down approaches to forecast reconciliation are well known to both practitioners and researchers. In short, according to the bottom-up approach (Dunn *et al.*, 1976), the forecasts at the most disaggregated level are summed up to obtain the aggregates of interest. On the contrary, in the top-down approach (Gross and Sohl, 1990), the most aggregated series is forecasted first, and then is disaggregated to provide lower level predictions (Fliedner, 2001; Athanasopoulos *et al.*, 2009, and the references therein). A combination of these two approaches, known as middle-out (Athanasopoulos *et al.*, 2009), selects an intermediate level of (dis)aggregation, and moves downward in a top-down fashion, and onwards according to bottom-up. However, in the last decade there have been several contributions exploiting a regression based optimal combination approach (Hyndman *et al.*, 2011). These have proven convincing from a mathematical-statistical point of view, flexible enough to be adapted to both cross-sectional and temporal frameworks (Wickramasuriya *et al.*, 2019; Athanasopoulos *et al.*, 2017), and very effective in improving the base forecasts from many different application fields (economics, demography, energy, tourism, etc.).

This Chapter considers an optimal combination approach, which takes the base (incoherent and however obtained) forecasts of all nodes in the hierarchy and reconciles them. We show that all the linear constraints induced by the cross-temporal hierarchy

underlying the time series, may be used to reconcile the base forecasts through a simple projection in a well chosen linear space. At this end, we extend the optimal (in the least squares sense) solutions separately proposed for each coherency dimension, developing the optimal point forecasting procedure which - while exploiting both cross-sectional and temporal hierarchies - simultaneously fulfills both contemporaneous and temporal constraints. We refer to this as optimal combination cross-temporal forecast reconciliation. In addition, grounding on the existing literature on this topic (Wickramasuriya *et al.*, 2019; Athanasopoulos *et al.*, 2017; Nystrup *et al.*, 2020), we discuss some simple approximations of the covariance matrix to be used in the statistical point forecast reconciliation, focusing on those making use of the in-sample residuals (when available) of the models used to get the base forecasts. The strictly, and very important related issue of probabilistic forecast reconciliation (Panagiotelis *et al.*, 2023; Wickramasuriya, 2024) will be considered in Chapter 5.

The Chapter is organized as follows. We start by presenting the general framework of point forecast reconciliation according to Byron (1978, 1979) (see also van Erven and Cugliari, 2015, and Panagiotelis *et al.*, 2021), avoiding reference to cross-sectional or time indices (Section 1.2). Hierarchical and grouped systems of multivariate time series are introduced in Section 1.3. The temporal hierarchies which characterize a single time series are discussed in Section 1.4. The cross-sectional and temporal coherency dimensions are dealt with simultaneously in Section 1.5, and the optimal (in the least squares sense) solution to the cross-temporal forecast reconciliation problem is then developed in Section 1.6. An iterative cross-temporal forecast reconciliation procedure which extends the heuristics proposed by Kourentzes and Athanasopoulos (2019) is described in Section 1.7. The feasibility of all the proposed new procedures, along with the evaluation of their performance when compared to the most performing 'single dimension' (either cross-sectional or temporal) forecast reconciliation procedures, is studied in Section 1.8. This is completed through a forecasting experiment on the 95 quarterly time series of the Australian *GDP* from Income and Expenditure sides considered by Athanasopoulos *et al.* (2020) and Bisaglia *et al.* (2020).

1.2 Optimal point forecast reconciliation

Forecast reconciliation is a post-forecasting process aimed to improve the quality of the *base* forecasts for a system of hierarchical/grouped, and more generally linearly constrained, time series (Hyndman *et al.*, 2011; Panagiotelis *et al.*, 2021) by exploiting the constraints that the series in the system must fulfill, whereas in general, the base forecasts do not. Let $\boldsymbol{y}$ be a $(s \times 1)$ vector of target point forecasts that is needed to

Symbols	Description
n_b, n_a, n	Number of bottom, upper, and total time series ($n = n_a + n_b$)
m	Highest available sampling frequency per seasonal cycle (max. order of temporal aggregation)
$h \geq 1$	Forecast horizon for the lowest frequency time series
T	Number of high-frequency observations used in the forecasting models
N	Number of observations at the lowest frequency: $N = \frac{T}{m}$
$\mathcal{K}$	Set of factors of m in descending order: $\mathcal{K} = \{k_p, k_{p-1}, \dots, k_2, k_1\}$, $k_p = m$, $k_1 = 1$
k^*	$\sum_{j=2}^{p} \frac{m}{k_j}$
M_k	$\frac{m}{k}$, $k \in \mathcal{K}$
$\boldsymbol{b}_t \in \mathbb{R}^{n_b}$	vector containing the bottom time series (bts) at time t
$\boldsymbol{a}_t \in \mathbb{R}^{n_a}$	vector containing the upper time series (uts) at time t
$\boldsymbol{y}_t \in \mathbb{R}^{n}$	vector containing the time series $\boldsymbol{y}_t = [\boldsymbol{a}_t' \;\; \boldsymbol{b}_t']'$ at time t
$\boldsymbol{Y}, \widehat{\boldsymbol{Y}}, \widetilde{\boldsymbol{Y}} \in \mathbb{R}^{n \times h(k^*+m)}$	Matrices of target, base and cross-temporally reconciled forecasts
$\boldsymbol{Y}^{[k]}, k \in \mathcal{K}$	$(n \times hM_k)$ matrix containing the target forecasts of the level k temporally aggregated series. Component of matrix $\boldsymbol{Y} = \left[\boldsymbol{Y}^{[m]} \; \boldsymbol{Y}^{[k_{p-1}]} \; \dots \; \boldsymbol{Y}^{[k_2]} \; \boldsymbol{Y}^{[1]}\right]$
$\widehat{\boldsymbol{Y}}^{[k]}, k \in \mathcal{K}$	$(n \times hM_k)$ matrix containing the base forecasts of the level k temporally aggregated series. Component of matrix $\widehat{\boldsymbol{Y}} = \left[\widehat{\boldsymbol{Y}}^{[m]} \; \widehat{\boldsymbol{Y}}^{[k_{p-1}]} \; \dots \; \widehat{\boldsymbol{Y}}^{[k_2]} \; \widehat{\boldsymbol{Y}}^{[1]}\right]$
$\widetilde{\boldsymbol{Y}}^{[k]}, k \in \mathcal{K}$	$(n \times hM_k)$ matrix containing the cross-temporally reconciled forecasts of the level k temporally aggregated series. Component of matrix $\widetilde{\boldsymbol{Y}} = \left[\widetilde{\boldsymbol{Y}}^{[m]} \; \widetilde{\boldsymbol{Y}}^{[k_{p-1}]} \; \dots \; \widetilde{\boldsymbol{Y}}^{[k_2]} \; \widetilde{\boldsymbol{Y}}^{[1]}\right]$
$\boldsymbol{A}^{[k]}$, $\boldsymbol{B}^{[k]}$, $k \in \mathcal{K}$	$(n_a \times hM_k)$ and $(n_b \times hM_k)$ components of matrix $\boldsymbol{Y}^{[k]} = \begin{bmatrix}\boldsymbol{A}^{[k]} \\ \boldsymbol{B}^{[k]}\end{bmatrix}$
$\boldsymbol{y}, \hat{\boldsymbol{y}}, \tilde{\boldsymbol{y}} \in \mathbb{R}^{hn(k^*+m)}$	$\boldsymbol{y} = \text{vec}\,(\boldsymbol{Y}')$, $\hat{\boldsymbol{y}} = \text{vec}\left(\widehat{\boldsymbol{Y}}'\right)$, $\tilde{\boldsymbol{y}} = \text{vec}\left(\widetilde{\boldsymbol{Y}}'\right)$
$\boldsymbol{P} \in \mathbb{R}^{hn(k^*+m) \times hn(k^*+m)}$	Commutation matrix, such that $\boldsymbol{P}\,[\text{vec}\,(\boldsymbol{Y})] = \text{vec}\,(\boldsymbol{Y}')$
$\boldsymbol{C} \in \mathbb{R}^{n_a \times n_b}$	Cross-sectional (contemporaneous) aggregation matrix
$\boldsymbol{S} \in \mathbb{R}^{n \times n_b}$	Cross-sectional (contemporaneous) summing matrix
$\boldsymbol{U}' \in \mathbb{R}^{n_a \times n}$	Zero-constraints cross-sectional matrix: $\boldsymbol{U}'\boldsymbol{Y} = \boldsymbol{0}_{[n_a \times (k^*+m)]}$
$\boldsymbol{K}_h \in \mathbb{R}^{hk^* \times hm}$	Temporal aggregation matrix
$\boldsymbol{R}_h \in \mathbb{R}^{h(k^*+m) \times hm}$	Temporal summing matrix
$\boldsymbol{Z}_h' \in \mathbb{R}^{hk^* \times h(k^*+m)}$	Zero-constraints temporal matrix: $\boldsymbol{Z}_h'\boldsymbol{Y}' = \boldsymbol{0}_{[hk^* \times n]}$
$\boldsymbol{H}' \in \mathbb{R}^{hn_a^* \times nh(k^*+m)}$	Zero-constraints full row-rank cross-temporal matrix: $\boldsymbol{H}'\boldsymbol{y} = \boldsymbol{0}$

Table 1.1: *Cross-temporal forecast reconciliation: symbols and notation*

satisfy the system of linearly independent constraints

$$\boldsymbol{H}'\boldsymbol{y} = \boldsymbol{0}_{(r \times 1)}, \tag{1.1}$$

where $\boldsymbol{H}'$ is a $(r \times s)$ matrix, with $\text{rank}(\boldsymbol{H}') = r < s$, and $\boldsymbol{0}_{(r \times 1)}$ is a $(r \times 1)$ null vector. Let $\widehat{\boldsymbol{y}}$ be a $(s \times 1)$ vector of unbiased point forecasts, not fulfilling the linear constraints (1.1) (i.e., $\boldsymbol{H}'\widehat{\boldsymbol{y}} \neq \boldsymbol{0}$).

Drawing upon Stone *et al.* (1942) and Byron (1978, 1979), we consider a regression-based reconciliation method assuming that $\widehat{\boldsymbol{y}}$ is related to $\boldsymbol{y}$ by

$$\widehat{\boldsymbol{y}} = \boldsymbol{y} + \boldsymbol{\varepsilon}, \tag{1.2}$$

where $\boldsymbol{\varepsilon}$ is a $(s \times 1)$ vector of zero mean disturbances, with known p.d. covariance matrix $\boldsymbol{W}$. The reconciled forecasts $\widetilde{\boldsymbol{y}}$ are found by minimizing the generalized least squares (GLS) objective function $(\widehat{\boldsymbol{y}} - \boldsymbol{y})' \boldsymbol{W}^{-1} (\widehat{\boldsymbol{y}} - \boldsymbol{y})$ constrained by (1.1):

$$\widetilde{\boldsymbol{y}} = \arg\min_{\boldsymbol{y}} (\boldsymbol{y} - \widehat{\boldsymbol{y}})' \boldsymbol{W}^{-1} (\boldsymbol{y} - \widehat{\boldsymbol{y}}), \quad \text{s.t. } \boldsymbol{H}'\boldsymbol{y} = \boldsymbol{0}.$$

The solution is given by (Byron, 1978, p. 360):

$$\widetilde{\boldsymbol{y}} = \widehat{\boldsymbol{y}} - \boldsymbol{W}\boldsymbol{H} (\boldsymbol{H}'\boldsymbol{W}\boldsymbol{H})^{-1} \boldsymbol{H}'\widehat{\boldsymbol{y}} = \boldsymbol{M}\widehat{\boldsymbol{y}}, \tag{1.3}$$

where $\boldsymbol{M} = \boldsymbol{I}_s - \boldsymbol{W}\boldsymbol{H} (\boldsymbol{H}'\boldsymbol{W}\boldsymbol{H})^{-1} \boldsymbol{H}'$ is a $(s \times s)$ projection matrix[1]. Denoting with $\boldsymbol{d}_{\widehat{\boldsymbol{y}}} = \boldsymbol{0} - \boldsymbol{H}'\widehat{\boldsymbol{y}}$ the $(r \times 1)$ vector containing the base forecasts' 'coherency' errors, we can re-state expression (1.3) as $\widetilde{\boldsymbol{y}} = \widehat{\boldsymbol{y}} + \boldsymbol{W}\boldsymbol{H} (\boldsymbol{H}'\boldsymbol{W}\boldsymbol{H})^{-1} \boldsymbol{d}_{\widehat{\boldsymbol{y}}}$, which makes it clear that the reconciliation formula (1.3) simply 'adjusts' the original forecasts vector $\widehat{\boldsymbol{y}}$ with a linear combination – according to the smoothing matrix $\boldsymbol{W}\boldsymbol{H} (\boldsymbol{H}'\boldsymbol{W}\boldsymbol{H})^{-1}$ – of the coherency errors in the base forecasts. In addition, if the error term of model (1.2) is Gaussian, the reconciliation error $\widetilde{\varepsilon} = \widetilde{\boldsymbol{y}} - \boldsymbol{y}$ is a zero-mean Gaussian vector with covariance matrix

$$E (\widetilde{\boldsymbol{y}} - \boldsymbol{y}) (\widetilde{\boldsymbol{y}} - \boldsymbol{y})' = \boldsymbol{W} - \boldsymbol{W}\boldsymbol{H} (\boldsymbol{H}'\boldsymbol{W}\boldsymbol{H})^{-1} \boldsymbol{H}' = \boldsymbol{M}\boldsymbol{W}.$$

Hyndman *et al.* (2011) propose an alternative formulation as for the reconciled estimates, equivalent to expression (1.3) and obtained by GLS estimation of the model

$$\widehat{\boldsymbol{y}} = \boldsymbol{S}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

where $\boldsymbol{S}$ is a 'structural summation matrix' describing the aggregation relationships operating on $\boldsymbol{y}$, and $\boldsymbol{\beta}$ is a subset of $\boldsymbol{y}$ containing the target forecasts of the bottom level series, such that $\boldsymbol{y} = \boldsymbol{S}\boldsymbol{\beta}$ (see Section 1.3). Since the hypotheses on $\boldsymbol{\varepsilon}$ remain unchanged, it can be shown (Wickramasuriya *et al.*, 2019) that $\widetilde{\boldsymbol{\beta}} = (\boldsymbol{S}'\boldsymbol{W}^{-1}\boldsymbol{S})^{-1} \boldsymbol{S}'\boldsymbol{W}^{-1}\widehat{\boldsymbol{y}}$ is the best linear unbiased estimate of $\boldsymbol{\beta}$, and that the whole reconciled forecasts vector is given by $\widetilde{\boldsymbol{y}} = \boldsymbol{S}\widetilde{\boldsymbol{\beta}} = \boldsymbol{S}\boldsymbol{G}\widehat{\boldsymbol{y}}$, where $\boldsymbol{G} = (\boldsymbol{S}'\boldsymbol{W}^{-1}\boldsymbol{S})^{-1} \boldsymbol{S}'\boldsymbol{W}^{-1}$.

As witnessed by the extensive literature on adjusting preliminary data (as the base forecasts can be considered) in order to fulfill some externally imposed constraints (Stone *et al.*, 1942; Byron, 1978, 1979; Weale, 1992), the distinctive feature of the GLS reconciliation approach is that it can take into account the 'quality', however measured, of the preliminary estimates, through an appropriate choice of the covariance matrix $\boldsymbol{W}$.

[1] A geometric interpretation of the entire hierarchical forecasting problem is given by Panagiotelis *et al.* (2021).

However, for a long time, these procedures have depended on the assumption that this matrix (or any other indicators of the estimates' accuracy) of the figures to be reconciled was known. In many practical situation $\boldsymbol{W}$ is assumed to be diagonal, and the data are adjusted in the light of their relative variances to satisfy the linear restrictions. But another - perhaps more delicate - challenge arises when either any reliability measure is available, or it can hardly be deduced by the data. The solutions proposed in literature for this case are basically of two types, both consistent with the least-squares approach shown so far:

1. mathematical/mechanical solutions: the base forecasts are balanced by minimizing a penalty criterion which 'induces' a covariance matrix (which is simply a statistical artifact);
2. data-based solutions: the variability of the base forecasts to be reconciled is estimated through the models and the data used to produce the forecasts.

As for point forecast reconciliation, in the following, we will consider both approaches, with an explicit preference towards approximations of $\boldsymbol{W}$ based on the in-sample residuals (when available), which appear more convincing from a statistical point of view, and generally well-performing in practical applications. However, this topic deserves further attention (Jeon *et al.*, 2019, p. 368), and will be considered for future research.

1.3 Hierarchical and grouped Time Series

Following Panagiotelis *et al.* (2021), we define a linearly constrained time series $\boldsymbol{y}_t$ as a n-dimensional time series such that all observed values $\boldsymbol{y}_1 \dots \boldsymbol{y}_T$ and all future values $\boldsymbol{y}_{T+1}, \boldsymbol{y}_{T+2} \dots$ lie in the coherent linear subspace $\mathcal{S} \subset \mathbb{R}^n$, that is: $\boldsymbol{y}_t \in \mathcal{S}$, $\forall t$. In many situations, the time series are linked through summation constraints, inducing a hierarchy. Figure 1.1 (left) gives an example of a hierarchical system with eight variables and three levels: the top-variable at level 0, two variables (A and B) at level 1, and five variables at level 2 (AA, AB, BA, BB, BC). The temporal observations of these variables form a hierarchical time series, consisting of five bottom time series (bts) and three aggregated upper time series (uts).

Assuming that the relationship mapping the lower-level series in the hierarchy of Figure 1.1 (left) into the higher ones always be a simple summation, the bottom-level series can be thought of as building blocks that cannot be obtained as sum of other series in the hierarchy, while all the series at upper levels can be expressed by appropriately summing part or all of them. For all time periods $t = 1, \dots, T$, the link between the

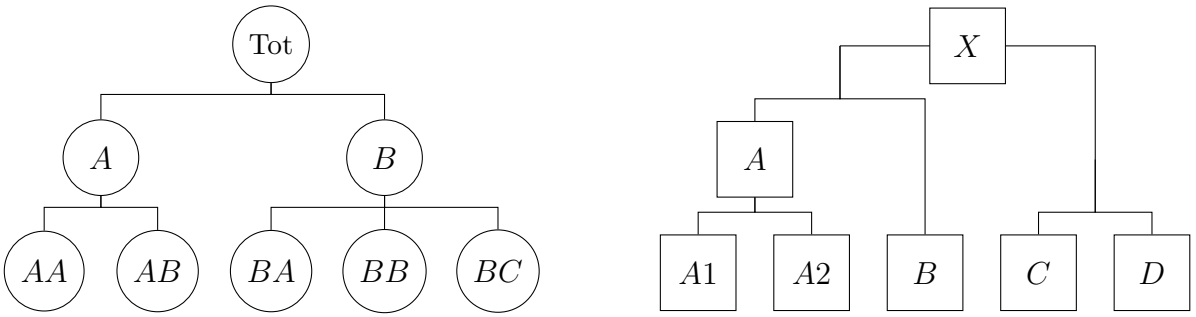


Figure 1.1: *Two examples of linearly constrained time series. Left: a simple three-level hierarchical structure. Right: two hierarchies sharing the same top-level series*

top level series y_t and the bottom level series is given by:

$$y_t = y_{AA,t} + y_{AB,t} + y_{BA,t} + y_{BB,t} + y_{BC,t}. \tag{1.4}$$

At the same time, the nodes at the intermediate level of the hierarchy satisfy the aggregation constraints:

$$\begin{array}{rcl} y_{A,t} & = & y_{AA,t} + y_{AB,t} \\ y_{B,t} & = & y_{BA,t} + y_{BB,t} + y_{BC,t} \end{array}. \tag{1.5}$$

Consider now the matrices $\boldsymbol{C}$, $\boldsymbol{S}$, and $\boldsymbol{U}'$, of dimension (3×5), (8×5), and (3×8), respectively:

$$\boldsymbol{C} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}, \quad \boldsymbol{S} = \begin{bmatrix} \boldsymbol{C} \\ \boldsymbol{I}_5 \end{bmatrix}, \quad \boldsymbol{U}' = [\boldsymbol{I}_3 \; -\boldsymbol{C}],$$

where matrix $\boldsymbol{U}'$ encodes each summation relationship in a row, with 1 at the associated node, and -1 at its leaves. Expressions (1.4) and (1.5) can be written in a more compact way if we define the vectors of *bottom level* ($\boldsymbol{b}_t$) and *upper level* ($\boldsymbol{a}_t$) time series at time t as, respectively, $\boldsymbol{b}_t = [y_{AA,t} \quad y_{AB,t} \quad y_{BA,t} \quad y_{BB,t} \quad y_{BC,t}]'$, $\boldsymbol{a}_t = [y_t \quad y_{A,t} \quad y_{B,t}]'$. Denoting by $\boldsymbol{y}_t$ the (8×1) vector $\boldsymbol{y}_t = [\boldsymbol{a}_t' \;\; \boldsymbol{b}_t']'$, the relationships linking bottom and upper time series can be equivalently expressed as:

$$\boldsymbol{a}_t = \boldsymbol{C}\boldsymbol{b}_t, \quad \boldsymbol{y}_t = \boldsymbol{S}\boldsymbol{b}_t, \quad \boldsymbol{U}'\boldsymbol{y}_t = \boldsymbol{0}_{(3\times 1)}, \quad t = 1, \dots, T. \tag{1.6}$$

Thus, for any time index t, $\boldsymbol{y}_t$ is in the kernel of $\boldsymbol{U}'$, also known as null-space of the linear transformation induced by matrix $\boldsymbol{U}'$, which is given by the set of vectors $\boldsymbol{v} \in \mathbb{R}^8$, such that $\boldsymbol{U}'\boldsymbol{v} = \boldsymbol{0}_{(3\times 1)}$ (Harville, 2008, p. 591). We call *structural representation* of series $\boldsymbol{y}_t$ the formulation $\boldsymbol{y}_t = \boldsymbol{S}\boldsymbol{b}_t$, $t = 1, \dots, T$, and *zero-constrained representation* of series $\boldsymbol{y}_t$ the equivalent expression $\boldsymbol{U}'\boldsymbol{y}_t = \boldsymbol{0}$, $t = 1, \dots, T$.

A linearly constrained time series formed by two or more hierarchical time series sharing the same top level series, and the same bottom level series, is called grouped time series (Hyndman *et al.*, 2011; Hyndman and Athanasopoulos, 2021). Provided matrix $\boldsymbol{C}$ is appropriately designed, the definitions of matrices $\boldsymbol{S}$ and $\boldsymbol{U}'$, depending solely on matrix $\boldsymbol{C}$, remain unchanged.

It should be noted that we can face linearly constrained time series for which the structural representation $\boldsymbol{y}_t = \boldsymbol{S}\boldsymbol{b}_t$ does not give a straightforward view of the links between bottom and upper level time series. Figure 1.1 (right) shows two very simple hierarchies, where the variables of each hierarchy contribute (from different sides) to the same top level variable X, and the bottom level series of the hierarchy on the left side ($A1$,$A2$,B) are independent of those on the right side (C,D). The aggregation relationships between the upper variables X and A, and the bottom ones $A1$, $A2$, B, C, and D are given by:

$$\begin{array}{rcl} X & = & A1 + A2 + B \\ X & = & C + D \\ A & = & A1 + A2 \end{array} \quad . \tag{1.7}$$

Expression (1.7) cannot be represented as a mapping from the bottom variables into (themselves, and) the upper variables. Nevertheless, it is possible to set up the constraints valid for all the component series in $\boldsymbol{y} = [X\ A\ A1\ A2\ B\ C\ D]'$ through the matrix

$$\check{\boldsymbol{U}}' = \begin{bmatrix} 1 & 0 & -1 & -1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & 1 & -1 & -1 & 0 & 0 & 0 \end{bmatrix},$$

such that $\check{\boldsymbol{U}}'\boldsymbol{y} = \boldsymbol{0}_{(3\times 1)}$. After simple operations on expression (1.7), it is found:

$$\begin{array}{rcl} X & = & C + D \\ A & = & -B + C + D \\ A1 & = & -A2 - B + C + D \end{array} \quad , \tag{1.8}$$

so we can write $\boldsymbol{U}'\boldsymbol{y} = \boldsymbol{0}_{(3\times 1)}$, with

$$\boldsymbol{U}' = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 1 & 1 & -1 & -1 \end{bmatrix} = [\boldsymbol{I}_3 \quad -\boldsymbol{C}]\,.$$

While there is no practical problem in working with such constraints, it is clear that they do not conform to the visual pattern of the linearly constrained time series in Figure 1.1 (right), where $A1$ appears as a 'bottom variable'. However, in (1.8) it is expressed as

linear combination of series $A2$, B, C, and D. We will come back on this important issue, and develop it in Chapter 4.

1.3.1 Alternative approximations of the covariance matrix for cross-sectional point forecast reconciliation

Suppose we have the $(n\times 1)$ vector $\widehat{\boldsymbol{y}}_h$ of unbiased base forecasts for the n variables of the linearly constrained series $\boldsymbol{y}_t$ for the forecast horizon h. If the base forecasts have been independently computed, generally they do not fulfill the cross-sectional aggregation constraints, i.e. $\boldsymbol{U}'\widehat{\boldsymbol{y}}_h \neq \boldsymbol{0}_{(n\times 1)}$. By adapting the general point forecast reconciliation formula (1.3), the vector of reconciled forecasts is given by:

$$\widetilde{\boldsymbol{y}}_h = \widehat{\boldsymbol{y}}_h - \boldsymbol{W}_{\text{cs}}\boldsymbol{U}\left(\boldsymbol{U}'\boldsymbol{W}_{\text{cs}}\boldsymbol{U}\right)^{-1}\boldsymbol{U}'\widehat{\boldsymbol{y}}_h, \tag{1.9}$$

where $\boldsymbol{W}_{\text{cs}}$ is a $(n \times n)$ p.d. matrix, assumed known, and suffix 'cs' stands for 'cross-sectional'. Alternative choices for $\boldsymbol{W}_{\text{cs}}$ proposed in literature are the following:

- identity (cs-ols): $\boldsymbol{W}_{\text{cs}} = \boldsymbol{I}_n$ (Hyndman *et al.*, 2011),
- structural (cs-struc): $\boldsymbol{W}_{\text{cs}} = \text{diag}\left(\boldsymbol{S}\boldsymbol{1}_{n_b}\right)$ (Athanasopoulos *et al.*, 2017),
- series variance (cs-wls): $\boldsymbol{W}_{\text{cs}} = \widehat{\boldsymbol{W}}_{\text{cs-var}} = \boldsymbol{I}_n \odot \widehat{\boldsymbol{W}}_1$ (Hyndman *et al.*, 2016),
- MinT-shr (cs-shr): $\boldsymbol{W}_{\text{cs}} = \widehat{\boldsymbol{W}}_{\text{cs-shr}} = \widehat{\lambda}\widehat{\boldsymbol{W}}_{\text{cs-var}} + (1-\widehat{\lambda})\widehat{\boldsymbol{W}}_1$ (Wickramasuriya *et al.*, 2019),
- MinT-sam (cs-sam): $\boldsymbol{W}_{\text{cs}} = \widehat{\boldsymbol{W}}_1$ (Wickramasuriya *et al.*, 2019),

where the symbol $\odot$ denotes the Hadamard product, $\widehat{\lambda}$ is an estimated shrinkage coefficient (Ledoit and Wolf, 2004), $\widehat{\boldsymbol{W}}_1$ is the covariance matrix of the one-step ahead in-sample forecast errors $\widehat{\boldsymbol{e}}_t = \boldsymbol{y}_t - \widehat{\boldsymbol{y}}_t$, $t = 1, \dots, T$:

$$\widehat{\boldsymbol{W}}_1 = \frac{1}{T}\sum_{t=1}^{T}\widehat{\boldsymbol{e}}_t\widehat{\boldsymbol{e}}_t'. \tag{1.10}$$

The first three matrices are diagonal, and in the first case, the projection is orthogonal. In contrast, the latter two ones (cs-shr and cs-sam) have been proposed within the minimum-trace point forecast reconciliation approach by Wickramasuriya *et al.* (2019). It should be noted that the quality of the estimate $\widehat{\boldsymbol{W}}_1$ crucially depends on the dimension of T. In particular, when $T < n$, matrix $\widehat{\boldsymbol{W}}_1$ is singular, which prevents the matrix inversion in expression (1.9). The shrunk version $\widehat{\boldsymbol{W}}_{\text{cs-shr}}$ is a feasible alternative, well performing in many practical situations (Wickramasuriya *et al.*, 2019).

1.3.2 Matrix representation of the cross-sectional constraints

Let us denote with

$$\boldsymbol{b}_t = \left[b_{1t} \quad \dots b_{it} \quad \dots \quad b_{n_b t}\right]', \qquad t = 1, \dots, T, \tag{1.11}$$

the T vectors, each of dimension $(n_b \times 1)$, containing the *high-frequency bottom-time series* (hf-bts), that is the bottom series of the hierarchy/group observed at the highest available temporal frequency. As we shall see in Section 1.5, in cross-temporal hierarchies of time series the hf-bts should be considered as the 'very' bottom series of the system, since they cannot be formed as either contemporaneous or temporal sum of other variables. Likewise, let us denote with

$$\boldsymbol{a}_t = [a_{1t} \quad \dots \quad a_{jt} \quad \dots \quad a_{n_a t}]', \qquad t = 1, \dots, T, \tag{1.12}$$

the T vectors, each of dimension $(n_a \times 1)$, containing the *high-frequency upper-time series* (hf-uts), which are the cross-sectionally aggregated series of the hierarchy/group, observed at the highest temporal frequency. At each time $t = 1, \dots, T$, the cross-sectional (contemporaneous) aggregation constraints that map the hf-bts into the hf-uts can be written as:

$$\boldsymbol{a}_t = \boldsymbol{C}\boldsymbol{b}_t, \quad t = 1, \dots, T, \tag{1.13}$$

where $\boldsymbol{C}$ is a $(n_a \times n_b)$ *contemporaneous aggregation matrix*. The structural representation of the linearly constrained time series $\boldsymbol{y}_t$ is in turn given by (Hyndman *et al.*, 2011) $\begin{bmatrix} \boldsymbol{a}_t \\ \boldsymbol{b}_t \end{bmatrix} = \begin{bmatrix} \boldsymbol{C} \\ \boldsymbol{I}_{n_b} \end{bmatrix} \boldsymbol{b}_t$, that is, $\boldsymbol{y}_t = \boldsymbol{S}\boldsymbol{b}_t$, $t = 1, \dots, T$, where $\boldsymbol{S} = \begin{bmatrix} \boldsymbol{C} \\ \boldsymbol{I}_{n_b} \end{bmatrix}$ is a $(n \times n_b)$ *contemporaneous summing matrix*, with $n = n_a + n_b$. The constraints valid for $\boldsymbol{y}_t$ can be expressed in zero-constrained form through the $(n_a \times n)$ *zero-constraints matrix* $\boldsymbol{U}' = [\boldsymbol{I}_{n_a} \; -\boldsymbol{C}]$, that is: $\boldsymbol{U}'\boldsymbol{y}_t = \boldsymbol{0}_{(n_a \times 1)}$, $t = 1, \dots, T$. Now, denote $\boldsymbol{B}$ the $(n_b \times T)$ matrix containing the T observations of the n_b hf-bts of the system:

$$\boldsymbol{B} = \begin{bmatrix} b_{11} & \dots & b_{1t} & \dots & b_{1T} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ b_{i1} & \dots & b_{it} & \dots & b_{iT} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ b_{n_b 1} & \dots & b_{n_b t} & \dots & b_{n_b T} \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1 & \dots & \boldsymbol{b}_t & \dots & \boldsymbol{b}_T \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1^{*\prime} \\ \vdots \\ \boldsymbol{b}_i^{*\prime} \\ \vdots \\ \boldsymbol{b}_{n_b}^{*\prime} \end{bmatrix},$$

where $\boldsymbol{b}_t$ has been defined by (1.11), and $\boldsymbol{b}_i^* = [b_{i1} \dots b_{it} \dots b_{iT}]'$, $i = 1, \dots, n_b$, is the $(T \times 1)$ vector containing all the observations of the i-th univariate hf-bts, where the asterisk in $\boldsymbol{b}_i^*$ is used to distinguish this vector, which combines b_{it} across all times for one series, from $\boldsymbol{b}_t$, which combines b_{it} across all series for one time. We consider the

$(n_a \times T)$ matrix $\boldsymbol{A}$ for the hf-uts as well:

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & \dots & a_{1t} & \dots & a_{1T} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{j1} & \dots & a_{jt} & \dots & a_{jT} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n_a 1} & \dots & a_{n_a t} & \dots & a_{n_a T} \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1 & \dots & \boldsymbol{a}_t & \dots & \boldsymbol{a}_T \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1^{*\prime} \\ \vdots \\ \boldsymbol{a}_j^{*\prime} \\ \vdots \\ \boldsymbol{a}_{n_a}^{*\prime} \end{bmatrix},$$

where $\boldsymbol{a}_t$ was defined by (1.12), and $\boldsymbol{a}_j^* = [a_{j1} \dots a_{jt} \dots a_{jT}]'$, $j = 1, \dots, n_a$, is the $(T \times 1)$ vector containing all the observations of the j-th univariate component hf-uts. The cross-sectional (contemporaneous) aggregation relationships (1.13) linking bottom and upper time series of $\boldsymbol{y}_t$ can thus be expressed in compact form, by simultaneously encompassing all T time periods, as

$$\boldsymbol{A} = \boldsymbol{C}\boldsymbol{B}, \tag{1.14}$$

which is equivalent to

$$\boldsymbol{U}'\boldsymbol{Y} = \boldsymbol{0}_{(n_a \times T)}, \tag{1.15}$$

where $\boldsymbol{Y} = [\boldsymbol{A}' \quad \boldsymbol{B}']'$ is the $(n \times T)$ matrix containing the observations of all n series. It is worth noting that the cross-sectional constraints (1.14) and (1.15) hold at any time observation index of any temporal frequency. This has to be considered when dealing with cross-temporal hierarchies (see Section 1.5).

Now, let us consider the vectorized forms of matrices $\boldsymbol{B}'$ and $\boldsymbol{A}'$, according to which the data is organized 'by-variable-first-and-then-by-time': $\boldsymbol{b} = \text{vec}\,(\boldsymbol{B}')$, $\boldsymbol{a} = \text{vec}\,(\boldsymbol{A}')$, where $\boldsymbol{b} = \left[\boldsymbol{b}_1^{*\prime} \dots \boldsymbol{b}_i^{*\prime} \dots \boldsymbol{b}_{n_b}^{*\prime}\right]'$, and $\boldsymbol{a} = \left[\boldsymbol{a}_1^{*\prime} \dots \boldsymbol{a}_j^{*\prime} \dots \boldsymbol{a}_{n_a}^{*\prime}\right]'$ have dimensions $(Tn_b \times 1)$, and $(Tn_a \times 1)$, respectively. The cross-sectional constraints (1.14) can thus be expressed as (Harville, 2008, p. 345):

$$\boldsymbol{a} = (\boldsymbol{C} \otimes \boldsymbol{I}_T)\,\boldsymbol{b}, \tag{1.16}$$

where the symbol $\otimes$ denotes the Kronecker product. Since $\boldsymbol{y} = \text{vec}\,(\boldsymbol{Y}') = [\boldsymbol{a}' \quad \boldsymbol{b}']'$, expression (1.16) can be also re-stated as:

$$(\boldsymbol{U}' \otimes \boldsymbol{I}_T)\,\boldsymbol{y} = \boldsymbol{0}_{(Tn_a \times 1)}. \tag{1.17}$$

1.4 Temporal hierarchies

Following Athanasopoulos et al. (2017), we consider a time series $\{x_t\}_{t=1}^T$ observed at the highest available sampling frequency per seasonal cycle, say m (e.g., month per year, $m = 12$, quarter per year, $m = 4$, hour per day, $m = 24$). Given a factor k of

m,[2] we can construct a temporally aggregated version of the time series x_t, through the non-overlapping sums of its k successive values, which has a seasonal period equal to $M_k = m/k$. To avoid ragged-edge data, we assume that the total number of observations of x_t involved in the non-overlapping aggregation is a multiple of m, i.e. $T = N \cdot m$, where N is the length of the most temporally aggregated version of the series, i.e. the series observed at the lowest available frequency. We denote with $\mathcal{K} = \{k_p, k_{p-1}, \ldots, k_2, k_1\}$ the set of the p factors of m, in descending order, where $k_p = m$ and $k_1 = 1$. The temporally aggregated series of order k can be written as

$$x_l^{[k]} = \sum_{t=(l-1)k+1}^{lk} x_t, \quad l = 1, \ldots, \frac{T}{k}, \quad k \in \mathcal{K}. \tag{1.18}$$

Expression (1.18) accounts also for the temporal aggregation transforming x_t in itself (i.e., $x_t \equiv x_l^{[1]}$, $l = t$). Since the observation index l in (1.18) varies with each aggregation level k, to express a common index for all levels, we define τ as the observation index of the most aggregated series, such that $l = \tau$ at that level, i.e. $x_\tau^{[m]}$, $\tau = 1, \ldots, N$. As for the other temporally aggregated series defined in expression (1.18), we stack the observations for each aggregation level below m in the $(M_k \times 1)$ column vectors

$$\boldsymbol{x}_\tau^{[k]} = \left[x_{M_k(\tau-1)+1}^{[k]} \; x_{M_k(\tau-1)+2}^{[k]} \; \cdots \; x_{M_k\tau}^{[k]}\right]', \tag{1.19}$$

with $\tau = 1, \ldots, N$ and $k \in \{k_{p-1}, \ldots, k_2, 1\}$. We may collect $x_\tau^{[m]}$ and the $p-1$ vectors defined by expression (1.19) in a single column vector, by keeping distinct the temporally aggregated data from the high-frequency one:

$$\boldsymbol{x}_\tau = \left[x_\tau^{[m]} \; \boldsymbol{x}_\tau^{[k_{p-1}]\prime} \; \ldots \; \boldsymbol{x}_\tau^{[k_2]\prime} \; \boldsymbol{x}_\tau^{[1]\prime}\right]' = \left[\boldsymbol{t}_{x_\tau}{}' \; \boldsymbol{x}_\tau^{[1]\prime}\right]', \quad \tau = 1, \ldots, N,$$

where $\boldsymbol{t}_{x_\tau} = \left[x_\tau^{[m]} \; \boldsymbol{x}_\tau^{[k_{p-1}]\prime} \; \ldots \; \boldsymbol{x}_\tau^{[k_2]\prime}\right]'$ is a $(k^* \times 1)$ vector, with $k^* = \sum_{j=2}^{p} \frac{m}{k_j}$, containing all the temporally aggregated series at the observation index τ, $\boldsymbol{x}_\tau^{[1]}$ is the $(m \times 1)$ vector of observations of the time series at the highest available frequency within the complete τ-th cycle, and thus each $\boldsymbol{x}_\tau$ has dimension $[(k^* + m) \times 1]$. The relationships linking the nodes in the hierarchy can be expressed as we did in (1.6) for the cross-sectional (contemporaneous) hierarchy case:

$$\boldsymbol{t}_{x_\tau} = \boldsymbol{K}_1 \boldsymbol{x}_\tau^{[1]}, \quad \boldsymbol{x}_\tau = \boldsymbol{R}_1 \boldsymbol{x}_\tau^{[1]}, \quad \boldsymbol{Z}_1' \boldsymbol{x}_\tau = \boldsymbol{0}_{(k^* \times 1)}, \quad \tau = 1, \ldots, N, \tag{1.20}$$

where $\boldsymbol{K}_1 = \left[\boldsymbol{1}_m \quad \boldsymbol{I}_{\frac{m}{k_{p-1}}} \otimes \boldsymbol{1}_{k_{p-1}} \quad \ldots \quad \boldsymbol{I}_{\frac{m}{k_2}} \otimes \boldsymbol{1}_{k_2}\right]'$ is the $(k^* \times m)$ temporal aggregation matrix converting the high-frequency observations into lower-frequency (temporally aggregated) ones, $\boldsymbol{R}_1 = [\boldsymbol{K}_1' \quad \boldsymbol{I}_m]'$ is the $[(k^* + m) \times m]$ *temporal summing* matrix, and

[2] If k is not a factor of m, then the seasonality of the aggregate series is a non-integer, and so forecasts of the aggregate are more difficult to compute.

$\boldsymbol{Z}_1' = [\boldsymbol{I}_{k^*} \ -\boldsymbol{K}_1]$ is the zero-constraints matrix valid for $\boldsymbol{x}_\tau$. For example, with quarterly data it is $m = 4$, $k^* = 3$, and $\boldsymbol{K}_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$.

The temporal aggregation relationships can be extended to the whole time span covered by series x_t. Denoting by $\boldsymbol{x} = (\boldsymbol{x}_1' \dots \boldsymbol{x}_\tau' \dots \boldsymbol{x}_N')'$ the $[N(k^* + m) \times 1]$ vector containing all the data of series X at any observed temporal frequency, the complete set of temporal aggregation constraints is given by

$$\boldsymbol{Z}_N' \boldsymbol{x} = \boldsymbol{0}_{(Nk^* \times 1)}, \tag{1.21}$$

where $\boldsymbol{Z}_N' = [\boldsymbol{I}_{Nk^*} \quad -\boldsymbol{K}_N]$, and $\boldsymbol{K}_N = \left[\boldsymbol{I}_N \otimes \boldsymbol{1}_m \quad \boldsymbol{I}_{N\frac{m}{k_{p-1}}} \otimes \boldsymbol{1}_{k_{p-1}} \quad \dots \quad \boldsymbol{I}_{N\frac{m}{k_2}} \otimes \boldsymbol{1}_{k_2}\right]'$.

1.4.1 Alternative approximations of the covariance matrix for temporal point forecast reconciliation

Suppose we have the $[(k^* + m) \times 1]$ vector $\widehat{\boldsymbol{x}}_h$ of unbiased base forecasts for the p temporal aggregates of a single time series X within a complete time cycle, i.e. at the forecast horizon $h = 1$ for the lowest (most aggregated) time frequency. If the base forecasts have been independently computed, generally they do not fulfill the temporal aggregation constraints, i.e. $\boldsymbol{Z}_1' \widehat{\boldsymbol{x}}_h \neq \boldsymbol{0}_{(k^* \times 1)}$. By adapting the general point forecast reconciliation formula (1.3), and not considering suffix h to simplify the notation, the vector of temporally reconciled forecasts is given by:

$$\widetilde{\boldsymbol{x}} = \widehat{\boldsymbol{x}} - \boldsymbol{\Omega} \boldsymbol{Z}_1 \left(\boldsymbol{Z}_1' \boldsymbol{\Omega} \boldsymbol{Z}_1\right)^{-1} \boldsymbol{Z}_1' \widehat{\boldsymbol{x}}, \tag{1.22}$$

where $\boldsymbol{\Omega}$ is a $[(k^* + m) \times (k^* + m)]$ p.d. matrix, assumed known.

To consider possible residual-based estimates of matrix $\boldsymbol{\Omega}$, denote

$$\widehat{\boldsymbol{e}}_\tau^{[k]} = \boldsymbol{x}_\tau^{[k]} - \widehat{\boldsymbol{x}}_\tau^{[k]}, \quad \tau = 1, \dots, N, \quad k \in \mathcal{K}, \tag{1.23}$$

the $(M_k \times 1)$ vectors of the in-sample residuals at time index τ for the models used to generate the base forecasts of the temporally aggregated series of order k. These vectors can be organized as

$$\widehat{\boldsymbol{E}}_x^{[k]} = \left[\widehat{\boldsymbol{e}}_1^{[k]} \quad \dots \quad \widehat{\boldsymbol{e}}_\tau^{[k]} \quad \dots \quad \widehat{\boldsymbol{e}}_N^{[k]}\right]', \quad k \in \mathcal{K}, \tag{1.24}$$

where each matrix $\widehat{\boldsymbol{E}}_x^{[k]}$ has dimension $(N \times M_k)$, and then grouped in the $[N \times (k^* + m)]$ matrix of in-sample residuals $\widehat{\boldsymbol{E}}_x = \left[\widehat{\boldsymbol{E}}_x^{[m]} \ \widehat{\boldsymbol{E}}_x^{[k_{p-1}]} \ \dots \ \widehat{\boldsymbol{E}}_x^{[k_2]} \ \widehat{\boldsymbol{E}}_x^{[1]}\right]$. Each column of this matrix contains the in-sample residuals pertaining to a specific node of the temporal hierarchy, thus the sample cross-covariance matrix of the $k^* + m$ nodes of the temporal

hierarchy is given by:

$$\widehat{\boldsymbol{\Omega}} = \frac{1}{N}\left(\widehat{\boldsymbol{E}}_x\right)' \widehat{\boldsymbol{E}}_x. \tag{1.25}$$

This matrix is well defined if $N > (k^*+m)$, otherwise there might be singularity issues which would prevent its use in expression (1.22) in place of matrix $\boldsymbol{\Omega}$. Athanasopoulos *et al.* (2017) and Hyndman and Kourentzes (2018) consider the following alternative choices for $\boldsymbol{\Omega}$ (the suffix 't' stands for 'temporal', to keep the 't'-procedures distinct from the 'cs'-ones shown in Section 1.3.1):

- identity (t-ols): $\boldsymbol{\Omega} = \boldsymbol{I}_{k^*+m}$,
- structural (t-struc): $\boldsymbol{\Omega} = \widehat{\boldsymbol{\Omega}}_{\text{t-struc}} = \text{diag}\left(\boldsymbol{R}_1\mathbf{1}_m\right)$
- hierarchy variance scaling (t-wlsh): $\boldsymbol{\Omega} = \widehat{\boldsymbol{\Omega}}_{\text{t-wlsh}} = \boldsymbol{I}_{k^*+m} \odot \widehat{\boldsymbol{\Omega}}$
- series variance scaling (t-wlsv): $\boldsymbol{\Omega} = \widehat{\boldsymbol{\Omega}}_{\text{t-wlsv}}$
- MinT-shr (t-shr): $\boldsymbol{\Omega} = \widehat{\boldsymbol{\Omega}}_{\text{t-shr}} = \widehat{\lambda}\widehat{\boldsymbol{\Omega}}_{\text{t-wlsh}} + (1-\widehat{\lambda})\widehat{\boldsymbol{\Omega}}$
- MinT-sam (t-sam): $\boldsymbol{\Omega} = \widehat{\boldsymbol{\Omega}}$

The series variance scaling matrix $\widehat{\boldsymbol{\Omega}}_{\text{t-wlsv}}$ 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), that requires a reduced number (p instead of k^*+m) of variances to be estimated as compared to the hierarchy variance scaling matrix $\widehat{\boldsymbol{\Omega}}_{\text{t-wlsh}}$. To include potential information in the residuals' autocorrelation, we also consider two matrices recently proposed by Nystrup *et al.* (2020):

- auto-covariance scaling (t-acov): $\boldsymbol{\Omega} = \widehat{\boldsymbol{\Omega}}_{\text{t-acov}}$
- series Markov (t-sar1): $\boldsymbol{\Omega} = \widehat{\boldsymbol{\Omega}}_{\text{t-sar1}}$

The auto-covariance scaling makes use of the estimates of the full autocovariance matrices within each aggregation level, while ignoring correlations between aggregation levels:

$$\widehat{\boldsymbol{\Omega}}_{\text{t-acov}} = \begin{bmatrix} \widehat{\boldsymbol{\Omega}}^{[m]} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \widehat{\boldsymbol{\Omega}}^{[1]} \end{bmatrix},$$

where the $(M_k \times M_k)$ matrices $\widehat{\boldsymbol{\Omega}}^{[k]}$ are given by:

$$\widehat{\boldsymbol{\Omega}}^{[k]} = \frac{1}{N}\sum_{\tau=1}^{N} \widehat{\boldsymbol{e}}_\tau^{[k]}(\widehat{\boldsymbol{e}}_\tau^{[k]})' = \frac{1}{N}\left(\widehat{\boldsymbol{E}}_x^{[k]}\right)' \widehat{\boldsymbol{E}}_x^{[k]}, \quad k \in \mathcal{K}, \tag{1.26}$$

with vector $\widehat{\boldsymbol{e}}_\tau^{[k]}$ and matrix $\widehat{\boldsymbol{E}}_x^{[k]}$ given by (1.23) and (1.24), respectively. Because it is sometimes difficult to estimate the covariance matrix within each aggregation level without assuming that it has some special form, Nystrup *et al.* (2020) consider the Toeplitz matrix for the estimated first-order autocorrelation coefficients of the in-sample

residuals for the $p-1$ levels $k = k_1, \dots, k_{p-1}$, of the series' temporal hierarchy. Denoting these autocorrelation coefficents with $\rho_{[k]}$, it is $\mathbf{\Gamma}^{[m]} = 1$,

$$\mathbf{\Gamma}^{[k]} = \begin{bmatrix} 1 & \cdots & \rho_{[k]}^{M_k-1} \\ \vdots & \ddots & \vdots \\ \rho_{[k]}^{M_k-1} & \cdots & 1 \end{bmatrix}, \quad k = k_1, \dots, k_{p-1},$$

where each matrix $\mathbf{\Gamma}^{[k]}$, $k \in \mathcal{K}$, has dimension $(M_k \times M_k)$. The p matrices are used to build the $[(k^* + m) \times (k^* + m)]$ matrix

$$\mathbf{\Gamma} = \begin{bmatrix} 1 & \mathbf{0}' & \cdots & \mathbf{0}' \\ \mathbf{0} & \mathbf{\Gamma}^{[k_{p-1}]} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{\Gamma}^{[1]} \end{bmatrix},$$

which can be used to estimate matrix $\mathbf{\Omega}$ as $\widehat{\mathbf{\Omega}}_{\text{t-sar1}} = \widehat{\mathbf{\Omega}}^{\frac{1}{2}}_{\text{t-wlsv}} \mathbf{\Gamma} \widehat{\mathbf{\Omega}}^{\frac{1}{2}}_{\text{t-wlsv}}$.

1.5 Cross-temporal reconciliation framework

1.5.1 Cross-temporal aggregation constraints

The cross sectional aggregation relationships (1.17), linking n series observed at the same time frequency over a T-length time span, and the equivalent temporal aggregation relationships (1.20) and (1.21), valid for an individual variable expressed at different time frequencies, can be simultaneously considered, by extending (i) the cross-sectional constraints to all observation frequencies, and (ii) the temporal aggregation relationships to all variables.

Considering contemporaneous and temporal dimensions in the same framework requires extending and adapting the notations used so far. To this end, define the p matrices $\boldsymbol{Y}^{[k]}$, each of dimension $(n \times NM_k)$, as $\boldsymbol{Y}^{[k]} = \left[\boldsymbol{A}^{[k]\prime} \quad \boldsymbol{B}^{[k]\prime}\right]'$, $k \in \mathcal{K}$, where $\boldsymbol{B}^{[k]} = \left[\boldsymbol{b}_1^{[k]} \quad \dots \quad \boldsymbol{b}_i^{[k]} \quad \dots \quad \boldsymbol{b}_{n_b}^{[k]}\right]$, and $\boldsymbol{A}^{[k]} = \left[\boldsymbol{a}_1^{[k]} \quad \dots \quad \boldsymbol{a}_j^{[k]} \quad \dots \quad \boldsymbol{a}_{n_a}^{[k]}\right]$, $k \in \mathcal{K}$, are the matrices containing the k-order temporal aggregates of the bts ($\boldsymbol{B}^{[k]}$) and uts ($\boldsymbol{A}^{[k]}$), of dimension $(n_b \times NM_k)$ and $(n_a \times NM_k)$, respectively. To be consistent with the notation so far, $\boldsymbol{Y}^{[1]}$, $\boldsymbol{B}^{[1]}$, and $\boldsymbol{A}^{[1]}$ denote the matrices containing data at the highest available sampling frequency, while $\boldsymbol{Y}$, $\boldsymbol{B}$, and $\boldsymbol{A}$ are used now to denote the matrices containing the data at any considered temporal frequency:

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

where $\boldsymbol{Y}$, $\boldsymbol{A}$, and $\boldsymbol{B}$ have n, n_a and n_b rows, respectively, and the same number of columns, $[N(k^*+m)]$.

Cross-sectional aggregation constraints

By exploiting the results shown in Section 1.3.2, the cross-sectional aggregation relationships operating along all the time observation indices can be expressed in compact form as $\boldsymbol{U}'\boldsymbol{Y} = \mathbf{0}_{[n_a \times N(k^*+m)]}$, or equivalently in vectorized form as

$$\left(\boldsymbol{U}' \otimes \boldsymbol{I}_{N(k^*+m)}\right) \boldsymbol{y} = \mathbf{0}_{[n_a N(k^*+m)\times 1]}. \tag{1.27}$$

Temporal aggregation constraints

The temporal aggregation relationships (1.20), valid for a single series, can be extended to each component of the time series $\boldsymbol{y}_t$ as follows:

$$\begin{bmatrix} \boldsymbol{A}^{[m]\prime} & \boldsymbol{B}^{[m]\prime} \\ \vdots & \vdots \\ \boldsymbol{A}^{[k_2]\prime} & \boldsymbol{B}^{[k_2]\prime} \end{bmatrix} = \boldsymbol{K}_N \begin{bmatrix} \boldsymbol{A}^{[1]\prime} & \boldsymbol{B}^{[1]\prime} \end{bmatrix}, \tag{1.28}$$

which can be equivalently written as $[\boldsymbol{I}_{Nk^*} \quad -\boldsymbol{K}_N]\,\boldsymbol{Y}' = \mathbf{0}_{(Nk^*\times n)}$, that is $\boldsymbol{Z}_N'\boldsymbol{Y}' = \mathbf{0}_{(Nk^*\times n)}$. This last expression can be re-stated in vectorized form as:

$$(\boldsymbol{I}_n \otimes \boldsymbol{Z}_N')\,\boldsymbol{y} = \mathbf{0}_{(nNk^*\times 1)}. \tag{1.29}$$

In summary, by considering expressions (1.27) and (1.29) together, the cross-temporal constraints working on the complete set of observations can be expressed as $\breve{\boldsymbol{H}}'\boldsymbol{y} = \mathbf{0}_{(n^*\times 1)}$, where $n^* = n_a N(k^*+m) + nNk^*$, and

$$\breve{\boldsymbol{H}}' = \begin{bmatrix} \boldsymbol{U}' \otimes \boldsymbol{I}_{N(k^*+m)} \\ \boldsymbol{I}_n \otimes \boldsymbol{Z}_N' \end{bmatrix}$$

is a $[n^* \times nN(k^*+m)]$ cross-temporal zero-constraints matrix. Due to the simultaneous consideration of temporal and cross-sectional relationships linking the various time series of the system, some rows of $\breve{\boldsymbol{H}}'$ are redundant, and can be eliminated in order to get a full row-rank zero-constraints matrix. In detail, matrix $\breve{\boldsymbol{H}}'$ consists in:

- Nn_ak^* rows defining the cross-sectional constraints operating on the lf-uts;
- Nn_am rows defining the cross-sectional constraints operating on the hf-bts;
- $Nn_a\,(k^*+m)$ rows defining the temporal constraints operating on both hf- and lf-uts;
- $Nn_b\,(k^*+m)$ rows defining the temporal constraints operating on both hf- and lf-bts.

Since the first set of Nn_ak^* constraints is linearly dependent from the other rows of matrix $\breve{\boldsymbol{H}}'$, a full row-rank cross-temporal zero-constraints matrix $\boldsymbol{H}'$ can be obtained by:

1. considering the $[Nn(k^*+m)\times Nn(k^*+m)]$ commutation matrix (Magnus and Neudecker, 2019, p. 54) $\boldsymbol{P}$, such that $\boldsymbol{P}\,[\mathrm{vec}\,(\boldsymbol{Y})] = \mathrm{vec}\,(\boldsymbol{Y}')$;

2. defining $\boldsymbol{U}^* = \left[\mathbf{0}_{(Nn_am\times Nnk^*)} \quad \boldsymbol{I}_{Nm}\otimes\boldsymbol{U}'\right]\boldsymbol{P}'$;
3. considering the $[N(n_am+nk^*)\times Nn(k^*+m)]$ matrix:

$$\boldsymbol{H}' = \begin{bmatrix} \boldsymbol{U}^* \\ \boldsymbol{I}_n\otimes\boldsymbol{Z}'_N \end{bmatrix}, \tag{1.30}$$

which has full row-rank equal to $N(n_am+nk^*) = n^* - Nn_ak^*$, and allows to re-state the complete cross-temporal constraints as:

$$\boldsymbol{H}'\boldsymbol{y} = \mathbf{0}. \tag{1.31}$$

1.5.2 Cross-temporal forecast reconciliation: introduction

Let us assume to have unbiased base forecasts for all the individual time series of a linearly constrained multiple time series, and for all levels of the temporal hierarchies built from the highest available sampling frequency. In addition, assume that the forecast horizon for the most temporally aggregated time series be $h = 1$,[3] and that the forecast horizons for the other temporally aggregated series cover the entire time cycle. This means that (i) the forecast horizon for the highest frequency time series is equal to m, and (ii) in general, the forecast horizon for a temporally aggregated time series of order k spans from 1 to M_k.

The base forecasts for each bottom time series of the system form the vectors $\widehat{\boldsymbol{b}}_i^{[k]}$, $i = 1,\dots,n_b$, $k\in\mathcal{K}$, where $\widehat{\boldsymbol{b}}_i^{[1]} = \left\{\widehat{b}_{il}^{[1]}\right\}_{l=1}^{m}$ is the $(m\times 1)$ vector containing the base forecasts for the i-th high-frequency bottom time series (hf-bts), which are the 'very' bottom time series in the cross-temporal framework, while the remaining $\widehat{\boldsymbol{b}}_i^{[k]}$'s (for $k\neq 1$) contain the M_k forecasts for the lower-frequency ones (lf-bts). The base forecasts for the upper time series can be similarly defined as $\widehat{\boldsymbol{a}}_j^{[k]}$, $j = 1,\dots,n_a$, $k\in\mathcal{K}$, where $\widehat{\boldsymbol{a}}_j^{[1]} = \left\{\widehat{a}_{jl}^{[1]}\right\}_{l=1}^{m}$ is the $(m\times 1)$ vector containing the base forecasts for the high-frequency j-th upper time series (hf-uts), and the $\widehat{\boldsymbol{a}}_j^{[k]}$'s (for $k\neq 1$) are $(M_k\times 1)$ vectors of low-frequency upper time series (lf-uts) forecasts. Let us collect these base forecasts in the $(n_b\times M_k)$ and $(n_a\times M_k)$, respectively, matrices $(k\in\mathcal{K})$

$$\widehat{\boldsymbol{B}}^{[k]} = \left[\widehat{\boldsymbol{b}}_1^{[k]} \;\dots\; \widehat{\boldsymbol{b}}_i^{[k]} \;\dots\; \widehat{\boldsymbol{b}}_{n_b}^{[k]}\right], \text{ and } \widehat{\boldsymbol{A}}^{[k]} = \left[\widehat{\boldsymbol{a}}_1^{[k]} \;\dots\; \widehat{\boldsymbol{a}}_j^{[k]} \;\dots\; \widehat{\boldsymbol{a}}_{n_a}^{[k]}\right]. \tag{1.32}$$

The matrix containing the base bts forecasts is given by $\widehat{\boldsymbol{B}} = \left[\widehat{\boldsymbol{B}}^{[m]} \;\; \widehat{\boldsymbol{B}}^{[k_{p-1}]} \;\dots\; \widehat{\boldsymbol{B}}^{[k_2]} \;\; \widehat{\boldsymbol{B}}^{[1]}\right]$, where $\widehat{\boldsymbol{B}}$ has dimension $[n_b\times(k^*+m)]$. The base uts forecasts can be similarly arranged in the $[n_a\times(k^*+m)]$ matrix $\widehat{\boldsymbol{A}} = \left[\widehat{\boldsymbol{A}}^{[m]} \;\; \widehat{\boldsymbol{A}}^{[k_{p-1}]} \;\dots\; \widehat{\boldsymbol{A}}^{[k_2]} \;\; \widehat{\boldsymbol{A}}^{[1]}\right]$. From expression (1.32) we can define the p matrices $\widehat{\boldsymbol{Y}}^{[k]}$, each of dimension $(n\times M_k)$, containing the base forecasts for the temporal

[3]The general case $h\geq 1$ can be dealt with straightforwardly.

aggregation level k of both uts and bts: $\widehat{\boldsymbol{Y}}^{[k]} = \left[\widehat{\boldsymbol{A}}^{[k]\prime} \quad \widehat{\boldsymbol{B}}^{[k]\prime}\right]'$, $k \in \mathcal{K}$. Finally, denoting with $\widehat{\boldsymbol{Y}}$ the $[n \times (k^* + m)]$ matrix containing the base forecasts of all series and for all temporal aggregation levels, it is $\widehat{\boldsymbol{Y}} = \left[\widehat{\boldsymbol{Y}}^{[m]}\ \widehat{\boldsymbol{Y}}^{[k_{p-1}]}\ \ldots\ \widehat{\boldsymbol{Y}}^{[k_2]}\ \widehat{\boldsymbol{Y}}^{[1]}\right] = \left[\widehat{\boldsymbol{A}}' \quad \widehat{\boldsymbol{B}}'\right]'$. In general, the base forecasts fulfill neither cross-sectional (contemporaneous) nor temporal aggregation constraints (i.e., respectively, $\boldsymbol{U}'\widehat{\boldsymbol{Y}} \neq \boldsymbol{0}_{[n_a\times(k^*+m)]}$, and $\boldsymbol{Z}_1'\widehat{\boldsymbol{Y}}' \neq \boldsymbol{0}_{(k^*\times n)}$). The cross-temporal point forecast reconciliation problem can thus be stated as follows. We are looking for a reconciled point forecast matrix, say $\widetilde{\boldsymbol{Y}}$, which is 'as-close-as-possible' (according to a pre-specified metric) to the base forecast matrix $\widehat{\boldsymbol{Y}}$, and simultaneously in line with the cross-sectional and temporal aggregation constraints:

$$\boldsymbol{U}'\widetilde{\boldsymbol{Y}} = \boldsymbol{0}_{n_a\times(k^*+m)} \qquad \text{and} \qquad \boldsymbol{Z}_1'\widetilde{\boldsymbol{Y}}' = \boldsymbol{0}_{(k^*\times n)}. \tag{1.33}$$

As we have previously shown, the relationships (1.33) can be expressed in vectorized form as $\boldsymbol{H}'\widetilde{\boldsymbol{y}} = \boldsymbol{0}$, where $\widetilde{\boldsymbol{y}} = \text{vec}\left(\widetilde{\boldsymbol{Y}}'\right)$ and, since $h = 1$, the full row-rank matrix $\boldsymbol{H}'$ in (1.30) becomes

$$\boldsymbol{H}' = \begin{bmatrix} \boldsymbol{U}^* \\ \boldsymbol{I}_n \otimes \boldsymbol{Z}_1' \end{bmatrix}, \tag{1.34}$$

with $\boldsymbol{U}^* = \left[\boldsymbol{0}_{(n_a m\times nk^*)} \quad \boldsymbol{I}_m \otimes \boldsymbol{U}'\right]\boldsymbol{P}'$, $\boldsymbol{P}$ being the $[n(k^* + m) \times n(k^* + m)]$ commutation matrix defined above.

1.6 Cross-temporal optimal forecast combination

Let us consider the multivariate regression model

$$\widehat{\boldsymbol{Y}} = \boldsymbol{Y} + \boldsymbol{E}, \tag{1.35}$$

where the involved matrices have each dimension $[n \times (k^* + m)]$ and contain, respectively, the base ($\widehat{\boldsymbol{Y}}$) and the target forecasts ($\boldsymbol{Y}$), and the coherency errors ($\boldsymbol{E}$) for the n component variables of the linearly constrained time series of interest. For each variable, $k^* + m$ base forecasts are available, pertaining to all aggregation levels of the temporal hierarchy for a complete cycle of high-frequency observation, m. Consider now two vectorized versions of model (1.35), by transforming the matrices either in original form or in transposed form, respectively:

$$\text{vec}\left(\widehat{\boldsymbol{Y}}\right) = \text{vec}\left(\boldsymbol{Y}\right) + \text{vec}\left(\boldsymbol{E}\right) \quad \Leftrightarrow \quad \widehat{\boldsymbol{\mathcal{Y}}} = \boldsymbol{\mathcal{Y}} + \boldsymbol{\varepsilon}, \tag{1.36}$$

$$\text{vec}\left(\widehat{\boldsymbol{Y}}'\right) = \text{vec}\left(\boldsymbol{Y}'\right) + \text{vec}\left(\boldsymbol{E}'\right) \quad \Leftrightarrow \quad \widehat{\boldsymbol{y}} = \boldsymbol{y} + \boldsymbol{\eta}. \tag{1.37}$$

The target forecasts must fulfill both the cross-sectional (contemporaneous) constraints $\boldsymbol{U}'\boldsymbol{Y} = \boldsymbol{0}_{[n_a\times(k^*+m)]}$, and the temporal aggregation constraints $\boldsymbol{Z}_1'\boldsymbol{Y}' = \boldsymbol{0}_{(k^*\times n)}$. That is,

in vectorized form:

$$(\boldsymbol{I}_{k^*+m} \otimes \boldsymbol{U}')\,\boldsymbol{\mathcal{Y}} = \boldsymbol{0}_{[n_a(k^*+m)\times 1]} \quad \Leftrightarrow \quad (\boldsymbol{U}' \otimes \boldsymbol{I}_{k^*+m})\,\boldsymbol{y} = \boldsymbol{0}_{[n_a(k^*+m)\times 1]}, \tag{1.38}$$

$$(\boldsymbol{Z}_1' \otimes \boldsymbol{I}_n)\,\boldsymbol{\mathcal{Y}} = \boldsymbol{0}_{(k^* n\times 1)} \quad \Leftrightarrow \quad (\boldsymbol{I}_n \otimes \boldsymbol{Z}_1')\,\boldsymbol{y} = \boldsymbol{0}_{(k^* n\times 1)}. \tag{1.39}$$

Denote with $\boldsymbol{P}$ the $[n(k^*+m)\times n(k^*+m)]$ commutation matrix such that $\boldsymbol{P}\text{vec}\,(\boldsymbol{Y}) = \text{vec}\,(\boldsymbol{Y}')$ (i.e. $\boldsymbol{P}\boldsymbol{\mathcal{Y}} = \boldsymbol{y}$). As a consequence, using the full row-rank matrix $\boldsymbol{H}'$ defined by expression (1.34), the constraints (1.38) and (1.39) can be re-stated as either $\boldsymbol{H}'\boldsymbol{y} = \boldsymbol{0}$, or $\boldsymbol{H}'\boldsymbol{P}\boldsymbol{\mathcal{Y}} = \boldsymbol{0}$. Now, denote $\boldsymbol{W} = E\,[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}']$, and $\boldsymbol{\Omega} = E\,[\boldsymbol{\eta}\boldsymbol{\eta}']$. $\boldsymbol{W}$ and $\boldsymbol{\Omega}$ are different parameterizations of the same statistical object, the covariance matrix of the random disturbances in the multivariate regression model (1.35), for which the following relationships hold: $\boldsymbol{\Omega} = \boldsymbol{P}\boldsymbol{W}\boldsymbol{P}'$, and $\boldsymbol{W} = \boldsymbol{P}'\boldsymbol{\Omega}\boldsymbol{P}$. Thus, to apply the general point forecast reconciliation formula (1.3) to a cross-temporal forecast reconciliation problem, we may consider either the expression $\widetilde{\boldsymbol{y}} = \widehat{\boldsymbol{y}} - \boldsymbol{\Omega}\boldsymbol{H}\left(\boldsymbol{H}'\boldsymbol{\Omega}\boldsymbol{H}\right)^{-1}\boldsymbol{H}'\widehat{\boldsymbol{y}}$, or equivalently re-state it as $\widetilde{\boldsymbol{\mathcal{Y}}} = \widehat{\boldsymbol{\mathcal{Y}}} - \boldsymbol{W}\boldsymbol{P}'\boldsymbol{H}\left(\boldsymbol{H}'\boldsymbol{P}\boldsymbol{W}\boldsymbol{P}'\boldsymbol{H}\right)^{-1}\boldsymbol{H}'\boldsymbol{P}\widehat{\boldsymbol{\mathcal{Y}}}$.

1.6.1 Simple approximations of the covariance matrix for cross-temporal point forecast reconciliation

Consider the column vectorized form of the multivariate regression (1.36), whose random disturbances can be written as $\boldsymbol{\varepsilon} = \left[\boldsymbol{\varepsilon}_1^{[m]}\ \boldsymbol{\varepsilon}_1^{[k_{p-1}]}\ \dots\ \boldsymbol{\varepsilon}_{\frac{m}{k_{p-1}}}^{[k_{p-1}]}\ \dots\ \boldsymbol{\varepsilon}_1^{[1]}\ \dots\ \boldsymbol{\varepsilon}_m^{[1]}\right]'$, where each $(n\times 1)$ vector $\boldsymbol{\varepsilon}_l^{[k]}$, $k\in\mathcal{K}$, $l = 1,\dots,M_k$, contains contemporaneous random disturbances, i.e. at the same observation index for a given temporal aggregation order. A simple generalization to the cross-temporal framework of the cross-sectional approach (see Section 1.3.1) consists of assuming that only the disturbances at the same time index of the same temporal aggregation level are correlated. No temporal dependence (either within the same series at different times, or between the n series) is admitted:

$$E\left[\boldsymbol{\varepsilon}_r^{[k_i]}\left(\boldsymbol{\varepsilon}_s^{[k_j]}\right)'\right] = \begin{cases} \boldsymbol{W}_l^{[k]} & \text{if } k_i = k_j = k,\ r = s = l \\ \boldsymbol{0} & \text{otherwise} \end{cases}, \quad \begin{array}{l} k\in\mathcal{K}, \\ l = 1,\dots,M_k. \end{array}$$

Furthermore, if it is assumed that within each temporal aggregation level the random disturbances follow a multivariate white noise, which means that the contemporaneous covariance matrices are constant in time (i.e., $\boldsymbol{W}_l^{[k]} = \boldsymbol{W}^{[k]}$, $k\in\mathcal{K}$, $l = 1,\dots,M_k$), the matrix $\boldsymbol{W}$ has the following block-diagonal pattern:

$$\boldsymbol{W} \;=\; \begin{bmatrix} \boldsymbol{W}^{[m]} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \left(\boldsymbol{I}_{\frac{m}{k_{p-1}}}\otimes\boldsymbol{W}^{[k_{p-1}]}\right) & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \left(\boldsymbol{I}_m\otimes\boldsymbol{W}^{[1]}\right) \end{bmatrix}. \tag{1.40}$$

From a practical point of view, each $(n \times n)$ matrix $\boldsymbol{W}^{[k]}$, $k \in \mathcal{K}$, may be approximated like in the cross-sectional forecast reconciliation case, possibly using the in-sample residuals (see Section 1.3.1). Expression (1.40) can thus be seen as a simple extension to the cross-temporal case of the approach developed in the cross-sectional framework, where no temporal dependence is accounted for both within and between the n series. We may similarly propose a simplified pattern of the disturbances covariance matrix of the multivariate regression model (1.35) by considering the row vectorization form (1.37). In this case, the random disturbances vector $\boldsymbol{\eta}$ can be written as $\boldsymbol{\eta} = \left[\boldsymbol{\eta}_1' \ \dots \ \boldsymbol{\eta}_i' \ \dots \boldsymbol{\eta}_n'\right]'$, where each $[(k^* + m) \times 1]$ vector $\boldsymbol{\eta}_i$, $i = 1, \dots, n$, contains the random disturbances at different observation indices of the various temporal aggregation levels for the same series i. If we assume that the n series are uncorrelated at any observation index for any temporal aggregation level (i.e. neither contemporaneous nor temporal correlation is admitted between the series) denoting with $\boldsymbol{\Omega}_{ii} = E\big(\boldsymbol{\eta}_i \boldsymbol{\eta}_i'\big)$, $i = 1, \dots, n$, the $[(k^* + m) \times (k^* + m)]$ covariance matrix of the coherency errors of the temporal hierarchies of series i, the complete matrix $\boldsymbol{\Omega}$ can be written as follows:

$$\boldsymbol{\Omega} = \begin{bmatrix} \boldsymbol{\Omega}_{11} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \boldsymbol{\Omega}_{nn} \end{bmatrix}, \tag{1.41}$$

where each matrix $\boldsymbol{\Omega}_{ii}$, $i = 1, \dots, n$, may be approximated as in the temporal forecast reconciliation case, possibly using the in-sample residuals (see Section 1.4.1). Thus expression (1.41) can be seen as a simple extension to the cross-temporal case of the approach developed in the temporal hierarchies framework, where no correlation is admitted between the random errors of the n series. The covariance patterns (1.40) and (1.41) (i) are placed at opposite ends of possible ways of dealing with cross-temporal variables, and (ii) should be considered as the first practical devices to make the optimal combination forecast approach feasible for the cross-temporal framework as well.

Residual-based estimates of the covariance matrix $\boldsymbol{W}$ (and its re-parameterized counterpart $\boldsymbol{\Omega}$) use of the in-sample residuals of the models used to forecast the n time series considered at any temporal aggregation level. Denote by $\widehat{\boldsymbol{E}}_l^{[k]}$,$k \in \mathcal{K}$, $l = 1, \dots, M_k$, the $(n \times N)$ matrix containing the in-sample residuals for a single node of the cross-temporal hierarchy (i.e., the i-th row contains the residuals for the N sub-periods l of the model used to forecast the temporal aggregate of order k of series i). For each temporal aggregation level $k \in \mathcal{K}$, the M_k matrices $\widehat{\boldsymbol{E}}_l^{[k]}$ can be grouped into the $(n \times NM_k)$ matrix $\widehat{\boldsymbol{E}}^{[k]} = \begin{bmatrix} \widehat{\boldsymbol{E}}_1^{[k]} & \dots & \widehat{\boldsymbol{E}}_l^{[k]} & \dots & \widehat{\boldsymbol{E}}_{M_k}^{[k]} \end{bmatrix}$, $k \in \mathcal{K}$. The $(n(k^* + m) \times N)$ matrix containing all the residuals at any time observation index and any temporal aggregation level can, in turn, be written as: $\widehat{\boldsymbol{E}} = \begin{bmatrix} \widehat{\boldsymbol{E}}_1^{[m]\prime} & \widehat{\boldsymbol{E}}_1^{[k_{p-1}]\prime} & \dots & \widehat{\boldsymbol{E}}_{\frac{m}{k_{p-1}}}^{[k_{p-1}]\prime} & \dots & \widehat{\boldsymbol{E}}_1^{[1]\prime} & \dots & \widehat{\boldsymbol{E}}_m^{[1]\prime} \end{bmatrix}' =$

$\begin{bmatrix}\widehat{\boldsymbol{e}}_1 & \dots & \widehat{\boldsymbol{e}}_\tau & \dots & \widehat{\boldsymbol{e}}_N\end{bmatrix}$, where each $[n(k^*+m)\times 1]$ vector $\widehat{\boldsymbol{e}}_\tau$, $\tau=1,\dots,N$, is given by

$$\widehat{\boldsymbol{e}}_\tau = \Big[\underbrace{\widehat{e}^{[m]}_{1,\tau} \quad \dots \quad \widehat{\boldsymbol{e}}^{[1]\prime}_{1,\tau}}_{k^*+m} \quad \cdots \quad \underbrace{\widehat{e}^{[m]}_{n,\tau} \quad \dots \quad \widehat{\boldsymbol{e}}^{[1]\prime}_{n,\tau}}_{k^*+m}\Big]'.$$

The sample residual covariance matrix can be calculated according to both parameterization as either $\widehat{\boldsymbol{\Omega}}_{\text{sam}} = \dfrac{1}{N}\displaystyle\sum_{\tau=1}^{N}\widehat{\boldsymbol{e}}_\tau\,(\widehat{\boldsymbol{e}}_\tau)' = \dfrac{1}{N}\widehat{\boldsymbol{E}}\widehat{\boldsymbol{E}}'$, or $\widehat{\boldsymbol{W}}_{\text{sam}} = \boldsymbol{P}'\widehat{\boldsymbol{\Omega}}_{\text{sam}}\boldsymbol{P}$. However, in many practical situations matrix $\widehat{\boldsymbol{E}}$ has a number of rows - which is equal to the number of nodes in the cross-temporal hierarchy - much larger than the number of columns, which is equal to $N=\dfrac{T}{m}$. Thus matrices $\widehat{\boldsymbol{\Omega}}_{\text{sam}}$ and $\widehat{\boldsymbol{W}}_{\text{sam}}$ might not have good properties (in particular, they are not p.d. if $N \le n(k^*+m)$), and simplified approximations must be found. Two feasible alternatives are given by either the diagonalization or the shrinkage of matrix $\widehat{\boldsymbol{W}}_{\text{sam}}$, that is, respectively, $\widehat{\boldsymbol{W}}_{\text{wlsh}} = \boldsymbol{I}_{n(k^*+m)}\odot\widehat{\boldsymbol{W}}_{\text{sam}}$, and $\widehat{\boldsymbol{W}}_{\text{shr}} = \widehat{\lambda}\widehat{\boldsymbol{W}}_{\text{wlsh}} + (1-\widehat{\lambda})\widehat{\boldsymbol{W}}_{\text{sam}}$, where $\widehat{\boldsymbol{W}}_{\text{wlsh}}$ is a diagonal matrix containing the estimates of the 'hierarchy variances' for each node of the cross-temporal hierarchy, $\widehat{\boldsymbol{W}}_{\text{shr}}$ is the matrix obtained by shrinkage of $\widehat{\boldsymbol{W}}_{\text{sam}}$ with target $\widehat{\boldsymbol{W}}_{\text{wlsh}}$, and $\widehat{\lambda}$ is an estimate of the coefficient of shrinkage intensity λ, $0\le\lambda\le 1$. Both $\widehat{\boldsymbol{W}}_{\text{sam}}$ and $\widehat{\boldsymbol{W}}_{\text{shr}}$ refer to all the $n(k^*+m)$ hierarchy nodes simultaneously taken, but unlike the former matrix, the latter should not suffer from possible singularity problems. An alternative choice is the block-diagonal matrix (1.40). Following Wickramasuriya *et al.* (2019), full and shrunk estimates of matrices $\boldsymbol{W}^{[k]}$, $k\in\mathcal{K}$, forming the blocks on the diagonal of this matrix, may be computed as:

$$\widehat{\boldsymbol{W}}^{[k]}_{\text{sam}} = \frac{1}{NM_k}\widehat{\boldsymbol{E}}^{[k]}(\widehat{\boldsymbol{E}}^{[k]})', \tag{1.42}$$

$$\widehat{\boldsymbol{W}}^{[k]}_{\text{shr}} = \widehat{\lambda}_k\left(\boldsymbol{I}_n\odot\widehat{\boldsymbol{W}}^{[k]}_{\text{sam}}\right) + (1-\widehat{\lambda}_k)\widehat{\boldsymbol{W}}^{[k]}_{\text{sam}}, \quad k\in\mathcal{K}, \tag{1.43}$$

and used to approximate $\boldsymbol{W}$ as follows ('BD' stands for 'Block-Diagonal'):

$$\widehat{\boldsymbol{W}}^{BD}_{\text{sam}} = \begin{bmatrix}\widehat{\boldsymbol{W}}^{[m]}_{\text{sam}} & \dots & \boldsymbol{0}\\ \vdots & \ddots & \vdots\\ \boldsymbol{0} & \dots & \boldsymbol{I}_m\otimes\widehat{\boldsymbol{W}}^{[1]}_{\text{sam}}\end{bmatrix}, \quad \widehat{\boldsymbol{W}}^{BD}_{\text{shr}} = \begin{bmatrix}\widehat{\boldsymbol{W}}^{[m]}_{\text{shr}} & \dots & \boldsymbol{0}\\ \vdots & \ddots & \vdots\\ \boldsymbol{0} & \dots & \boldsymbol{I}_m\otimes\widehat{\boldsymbol{W}}^{[1]}_{\text{shr}}\end{bmatrix}.$$

Most of the choices for $\boldsymbol{W}$ (or $\boldsymbol{\Omega}$) shown so far are simple extensions to the cross-temporal framework of the approximations for $\boldsymbol{W}$ (or $\boldsymbol{\Omega}$) considered either in cross-sectional or in temporal forecast reconciliation. We consider the following approximations ('oct' stands for 'optimal cross-temporal'):

- identity (oct-ols): $\boldsymbol{W}=\boldsymbol{\Omega}=\boldsymbol{I}_{n(k^*+m)}$
- hierarchy variance scaling (oct-wlsh): $\boldsymbol{W}=\widehat{\boldsymbol{W}}_{\text{wlsh}}$

- series variance scaling (oct-wlsv): $\boldsymbol{W} = \widehat{\boldsymbol{W}}_{\text{wlsv}} = \boldsymbol{P}'\widehat{\boldsymbol{\Omega}}_{\text{wlsv}}\boldsymbol{P}$, where $\widehat{\boldsymbol{\Omega}}_{\text{wlsv}}$ is a straightforward extension of $\widehat{\boldsymbol{\Omega}}_{\text{t-wlsv}}$ (see Section 1.4.1)
- block-diagonal shrunk cross-covariance scaling (oct-bdshr): $\boldsymbol{W} = \widehat{\boldsymbol{W}}^{BD}_{\text{shr}}$
- block-diagonal cross-covariance scaling (oct-bdsam): $\boldsymbol{W} = \widehat{\boldsymbol{W}}^{BD}_{\text{sam}}$
- auto-covariance scaling (acov): $\boldsymbol{W} = \widehat{\boldsymbol{W}}_{\text{acov}} = \boldsymbol{P}'\widehat{\boldsymbol{\Omega}}_{\text{acov}}\boldsymbol{P}$, where $\widehat{\boldsymbol{\Omega}}_{\text{acov}}$ is a straightforward extension of $\widehat{\boldsymbol{\Omega}}_{\text{t-acov}}$ (see Section 1.4.1)
- MinT-shr (oct-shr): $\boldsymbol{W} = \widehat{\boldsymbol{W}}_{\text{shr}}$
- MinT-sam (oct-sam): $\boldsymbol{W} = \widehat{\boldsymbol{W}}_{\text{sam}}$

1.7 Heuristic cross-temporal reconciliation

1.7.1 The KA procedure

Kourentzes and Athanasopoulos (2019), henceforth KA, have proposed a cross-temporal reconciliation method that can be viewed as an ensemble forecasting procedure that exploits the simple averaging of different forecasts. The procedure consists of the following steps (it is assumed $h = 1$):

Step 1 - For each individual variable, compute the temporally reconciled forecasts and collect them in the $[n \times (k^* + m)]$ matrix $\breve{\boldsymbol{Y}}$. This result can be obtained by applying the point forecast reconciliation formula (1.22) to each column of matrix

$$\widehat{\boldsymbol{Y}}' = \begin{bmatrix} \widehat{\boldsymbol{t}}_{a_1} & \cdots & \widehat{\boldsymbol{t}}_{a_{n_a}} & \widehat{\boldsymbol{t}}_{b_1} & \cdots & \widehat{\boldsymbol{t}}_{b_{n_b}} \\ \widehat{\boldsymbol{a}}^{[1]}_1 & \cdots & \widehat{\boldsymbol{a}}^{[1]}_{n_a} & \widehat{\boldsymbol{b}}^{[1]}_1 & \cdots & \widehat{\boldsymbol{b}}^{[1]}_{n_b} \end{bmatrix}.$$

These reconciled forecasts are in line with the temporal aggregation constraints, i.e. $\boldsymbol{Z}_1'\breve{\boldsymbol{Y}}' = \boldsymbol{0}_{(k^* \times n)}$, but in general they are not in line with the cross-sectional constraints, that is: $\boldsymbol{U}'\breve{\boldsymbol{Y}} \neq \boldsymbol{0}_{[n_a \times (k^*+m)]}$.

Step 2 - Transform $\breve{\boldsymbol{Y}}$ by computing time-by-time cross-sectional reconciled forecasts for all the temporal aggregation levels, and collect them in the $[n \times (k^* + m)]$ matrix $\breve{\boldsymbol{Y}} = \left[\breve{\boldsymbol{Y}}^{[m]}\ \breve{\boldsymbol{Y}}^{[k_{p-1}]} \dots \breve{\boldsymbol{Y}}^{[k_2]}\ \breve{\boldsymbol{Y}}^{[1]}\right]$, where $\breve{\boldsymbol{Y}}^{[k]}$, $k \in \mathcal{K}$, has dimension $(n \times M_k)$. Thus, the cross-sectionally reconciled forecasts can be computed by transforming each $\breve{\boldsymbol{Y}}^{[k]}$ as $\breve{\breve{\boldsymbol{Y}}}^{[k]} = \boldsymbol{M}^{[k]}\breve{\boldsymbol{Y}}^{[k]}$, $k \in \mathcal{K}$, where $\boldsymbol{M}^{[k]} = \boldsymbol{I}_n - \boldsymbol{W}^{[k]}\boldsymbol{U}\left(\boldsymbol{U}'\boldsymbol{W}^{[k]}\boldsymbol{U}\right)^{-1}\boldsymbol{U}'$ denotes the $(n \times n)$ projection matrix used to reconcile forecasts of k-level temporally aggregated time series, and $\boldsymbol{W}^{[k]}$ is a $(n \times n)$ known p.d. matrix. Since it is $\boldsymbol{U}'\boldsymbol{M}^{[k]} = \boldsymbol{0}_{(n_a \times n)}$, $k \in \mathcal{K}$, the reconciled forecasts are cross-sectionally coherent, i.e. $\boldsymbol{U}'\breve{\breve{\boldsymbol{Y}}} = \boldsymbol{0}_{[n_a \times (k^*+m)]}$, but not temporally: $\boldsymbol{Z}_1'\breve{\breve{\boldsymbol{Y}}}' \neq \boldsymbol{0}_{(k^* \times n)}$.

Step 3 - Transform again the step 1 forecasts $\breve{\boldsymbol{Y}}$, by computing time-by-time cross-sectional reconciled forecasts for all the temporal aggregation levels using the $(n \times n)$

matrix $\overline{\boldsymbol{M}}$, given by the average of the matrices $\boldsymbol{M}^{[k]}$ obtained at step 2. Matrix $\overline{\boldsymbol{M}}$ can be expressed as $\overline{\boldsymbol{M}} = \frac{1}{p}\sum_{k\in\mathcal{K}} \boldsymbol{M}^{[k]}$, and the final cross-temporal reconciled forecasts are given by:

$$\widetilde{\boldsymbol{Y}}^{KA} = \overline{\boldsymbol{M}}\breve{\boldsymbol{Y}}. \tag{1.44}$$

Since $\boldsymbol{U}'\overline{\boldsymbol{M}} = \frac{1}{p}\sum_{k\in\mathcal{K}} \boldsymbol{U}'\boldsymbol{M}^{[k]} = \boldsymbol{0}_{(n_a\times n)}$, and $\boldsymbol{Z}_1'\breve{\boldsymbol{Y}}' = \boldsymbol{0}_{(k^*\times n)}$, the reconciled forecasts (1.44) fulfill both cross-sectional and temporal aggregation constraints: $\boldsymbol{U}'\widetilde{\boldsymbol{Y}}^{KA} = \boldsymbol{U}'\overline{\boldsymbol{M}}\,\breve{\boldsymbol{Y}} = \boldsymbol{0}_{[n_a\times(k^*+m)]}$, and $\boldsymbol{Z}_1'\left(\widetilde{\boldsymbol{Y}}^{KA}\right)' = \boldsymbol{Z}_1'\,\breve{\boldsymbol{Y}}\,'\overline{\boldsymbol{M}}' = \boldsymbol{0}_{(k^*\times n)}$.

1.7.2 Some remarks

To perform step 1, KA consider either t-struc or t-wlsv (see Section 1.4.1) as forecast reconciliation procedures through temporal hierarchies, while in step 2 either cs-wls or cs-shr are used (see Section 1.3.1).

Remark 1 - These two steps can be seen as the successive applications of two distinct multivariate reconciliation procedures: in the first step it is solved a linearly constrained quadratic minimization problem, where only temporal aggregation constraints are considered:

$$\breve{\boldsymbol{y}} = \arg\min_{\boldsymbol{y}} \left(\boldsymbol{y}-\widehat{\boldsymbol{y}}\right)'\boldsymbol{\Omega}^{-1}\left(\boldsymbol{y}-\widehat{\boldsymbol{y}}\right), \quad \text{s.t. } \left(\boldsymbol{I}_n\otimes\boldsymbol{Z}_1'\right)\boldsymbol{y} = \boldsymbol{0},$$

where $\boldsymbol{\Omega}$ is the block-diagonal matrix in (1.41). The second step consists of another quadratic minimization problem, where only cross-sectional (contemporaneous) constraints are considered:

$$\breve{\boldsymbol{\mathcal{Y}}} = \arg\min_{\boldsymbol{\mathcal{Y}}} \left(\boldsymbol{\mathcal{Y}}-\breve{\boldsymbol{\mathcal{Y}}}\right)'\boldsymbol{W}^{-1}\left(\boldsymbol{\mathcal{Y}}-\breve{\boldsymbol{\mathcal{Y}}}\right), \quad \text{s.t. } \left(\boldsymbol{I}_{k^*+m}\otimes\boldsymbol{U}'\right)\boldsymbol{\mathcal{Y}} = \boldsymbol{0},$$

where $\boldsymbol{W}$ is the block-diagonal matrix in (1.40). It is worth noting that one may switch between the two data representations using the permutation matrix $\boldsymbol{P}$, that is $\breve{\boldsymbol{y}} = \boldsymbol{P}'\breve{\boldsymbol{\mathcal{Y}}}$.

Remark 2 - In general the final result of the reconciliation procedure would change if the user inverts the order of application of the two reconciliation steps. Since the differences between the reconciled point forecasts according to these two approaches could be not negligible (see Section 1.7.3), this seems to be a weakness of the procedure, and calls for a decision rule about the final reconciled forecasts to retain. A practical way of doing this could be choosing the reconciled forecasts, which are the 'closest' (according to a given metric) to the base forecasts.

Remark 3 - The calculation of the average matrix $\overline{\boldsymbol{M}}$ in the final step of the procedure, needed to recover the cross-temporal coherency across the point forecasts, requires the availability of the projection matrices used in the second step. This poses no problem

when closed form reconciliation formulae can be used. Unfortunately, this is not the case when non-negativity of the final reconciled estimates is desired, calling for appropriate numerical procedures (Wickramasuriya *et al.*, 2020; Kourentzes and Athanasopoulos, 2021).

1.7.3 An iterative heuristic cross-temporal reconciliation

Taking inspiration from the KA reconciliation procedure, we consider an iterative procedure that produces cross-temporal reconciled forecasts by alternating forecast reconciliation along one single dimension (cross-sectional or temporal). Each iteration consists of the first two steps of the KA procedure, so the forecasts are reconciled by alternating reconciliation through temporal hierarchies and cross-sectional reconciliation in a cyclic fashion. Starting from the base forecasts $\widehat{\boldsymbol{Y}}$, denote with d_{cs} and d_{te}, respectively, the cross-sectional and the temporal gross discrepancies, given by $d_{cs} = \|\boldsymbol{U}'\widehat{\boldsymbol{Y}}\|_1$, and $d_{te} = \|\boldsymbol{Z}_1'\widehat{\boldsymbol{Y}}'\|_1$, where $\|\boldsymbol{X}\|_1 = \sum_{i,j}|x_{i,j}|$. Since the base forecasts are not in line with either type of constraints, in general both d_{cs} and d_{te} are greater than zero. The iterative procedure can be described as follows:

1. Start the iterations by calculating the temporally reconciled forecasts $\widetilde{\boldsymbol{Y}}^{(1)}$, such that $\boldsymbol{Z}_1'\left(\widetilde{\boldsymbol{Y}}^{(1)}\right)' = \boldsymbol{0}$, and $d_{cs}^{(1)} = \|\boldsymbol{U}'\widetilde{\boldsymbol{Y}}^{(1)}\|_1 \geq 0$.
2. The point forecasts in matrix $\widetilde{\boldsymbol{Y}}^{(1)}$ are then cross-sectionally reconciled, obtaining $\widetilde{\boldsymbol{Y}}^{(2)}$, which is such that $\boldsymbol{U}'\widetilde{\boldsymbol{Y}}^{(2)} = \boldsymbol{0}$, and $d_{te}^{(1)} = \|\boldsymbol{Z}_1'\widetilde{\boldsymbol{Y}}^{(2)\prime}\|_1 \geq 0$.
3. The updates in steps 1. and 2. are performed at each iteration j, $j = 1, 2, \ldots$, until a convergence criterion is met, that is $d_{te}^{(j)} < \delta$, where δ is a positive tolerance value (e.g., $\delta = 10^{-6}$), and matrix $\widetilde{\boldsymbol{Y}}^{(2j)}$ contains the final cross-temporal reconciled forecasts.

The above procedure can be seen as an extension of the well known iterative proportional fitting procedure (Deming and Stephan, 1940), also known as the RAS method (Miller and Blair, 2009), to adjust the internal cell values of a two-dimensional matrix iteratively until they sum to some predetermined row and column totals. In that case, the adjustment follows a proportional adjustment scheme. In contrast, each adjustment step is made in the cross-temporal reconciliation framework according to the penalty function associated to the single-dimension reconciliation procedure adopted.

Indeed, the choice of dimension along which the first reconciliation step in each iteration is performed is up to the user. There is no particular reason to perform the temporal reconciliation first and the cross-sectional reconciliation second. Figure 1.2 shows the percentage discrepancies in the Australian *GDP* at current prices one-step-ahead forecasts for any temporal aggregation level (quarterly, semi-annual, annual, see

Figure 1.2: *Quarterly, semi-annual and annual Australian GDP one-step-ahead reconciled forecasts according to the Kourentzes and Athanasopoulos (2019) cross-temporal reconciliation approach (t-wlsv for the temporal step, cs-shr for the cross-sectional step) by alternating the constraint dimensions to be fulfilled: percentage differences between the reconciled forecasts obtained through (i) temporal-then-cross-sectional reconciliation, and (ii) cross-sectional-then-temporal reconciliation. The differences between the two reconciled forecasts are divided by their arithmetic mean.*

Section 1.8), when the cross-temporal reconciliation is performed according to either the KA approach, or to the analogous procedure where the cross-sectional constraints are considered first, and then the temporal dimension is accounted for (see Section G in the online appendix). Percentage differences in the reconciled forecasts for this single, important variable, are visually evident, though bounded within (-0.3% – +0.4%).

Figure 1.3 completes the results shown so far by considering the forecasts of the strictly positive 79 (out of 95) variables from both Income and Expenditure sides, cross-temporally reconciled according to the KA procedure and its iterative variant. The boxplots show the distributions of the percentage discrepancies between the reconciled forecasts obtained using temporal reconciliation first, and cross-sectional reconciliation then, *vis-à-vis* the results obtained by inverting the order of application of the two reconciliation procedures. It appears that the iterative variant of the original KA proposal produces less pronounced discrepancies.

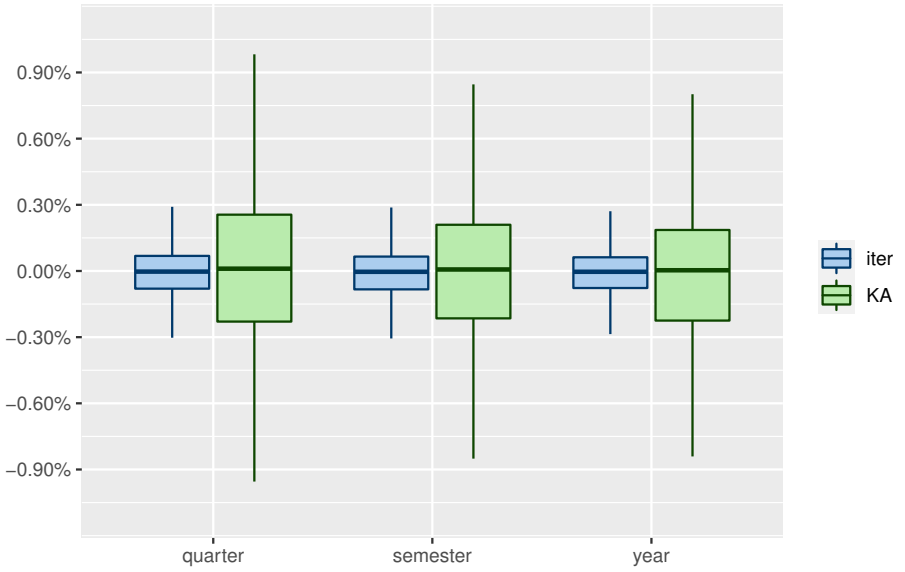


Figure 1.3: *Quarterly, semi-annual and annual one-step-ahead reconciled forecasts of 79 out of 95 times series of the Australian GDP from Income and Expenditure sides using both the original KA cross-temporal reconciliation procedure (t-wlsv for the temporal step, and cs-shr for the cross-sectional one), and its iterative variant: boxplots of the percentage differences between the reconciled forecasts obtained through (i) temporal-then-cross-sectional reconciliation, and (ii) cross-sectional-then-temporal reconciliation. The differences between each pair of reconciled forecasts are divided by their arithmetic mean.*

It must be said that the convergence speed of the iterative procedure does not seem to be affected by the choice of the first dimension to be fulfilled when the iteration starts. Figure 1.4 shows an example of the convergence speed of the iterative procedure either starting with the cross-sectional (right panel) or temporal (left panel) reconciliation procedure for the Australian *GDP* forecasts. In both cases, the convergence is achieved very quickly: fixing $\delta = 10^{-6}$, 15 (14) iterates are needed when starting from the temporal (cross-sectional) dimension. Furthermore, from the fourth iteration onwards the constraints are practically fulfilled in both cases. Nevertheless, since the final reconciled values depend on this choice, it would be useful having an ex-ante 'choice rule' between the two alternatives. For ease of presentation, in the following we maintain the original choice made by KA, performing temporal forecast reconciliation first and cross-sectional reconciliation then.

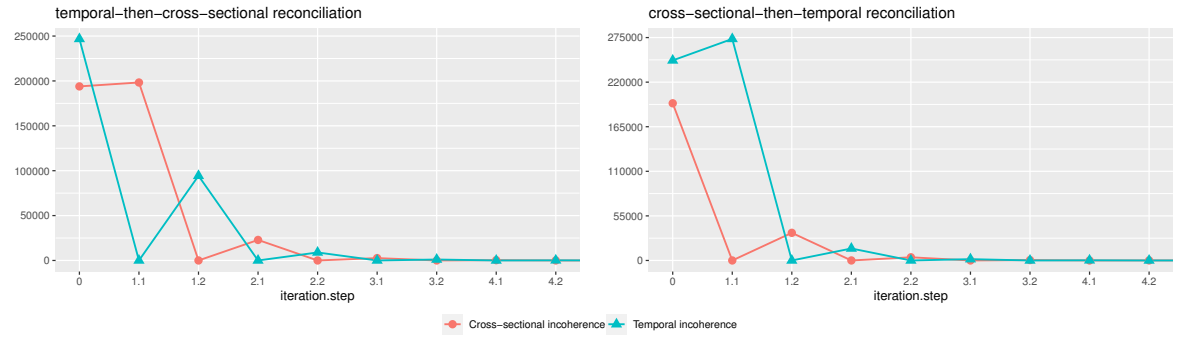


Figure 1.4: *Incoherence at each iteration step of the iterative cross-temporal forecast reconciliation procedure (t-wlsv + cs-shr) for the Australian GDP time series, at the first forecast origin 1994:Q3.*

1.8 Cross-temporal reconciliation of the Australian *GDP* forecasts

In a recent paper, Athanasopoulos *et al.* (2020) propose "the application of state-of-the-art forecast reconciliation methods to macroeconomic forecasting" to perform aligned decision making and to improve forecast accuracy. In their empirical study, they consider the cross-sectional forecast reconciliation for 95 Australian Quarterly National Accounts time series, describing the Gross Domestic Product (*GDP*) at current prices from Income and Expenditure sides, interpreted as two distinct hierarchical structures. In the former case (Income), *GDP* is on the top of 15 lower level aggregates, while in the latter (Expenditure), *GDP* is the top level aggregate of a hierarchy of 79 time series (for details, refer to Athanasopoulos *et al.*, 2020, pp. 702-705, and to figures 21.4-21.7 therein). By managing the complete set of 95 time series following the approach described in Section 1.3, Bisaglia *et al.* (2020) have extended the results of Athanasopoulos *et al.* (2020), showing that fully cross-sectional reconciled quarterly *GDP* forecasts, coherent with all the reconciled forecasts from both Expenditure and Income sides can be obtained through the general reconciliation approach described in Section 1.2. According to the notation adopted so far, the (33×95) zero-constraints matrix accounting for the cross-sectional constraints is given by:

$$\boldsymbol{U}' = \begin{bmatrix} 1 & \mathbf{0}'_{(5\times1)} & -\mathbf{1}'_{(10\times1)} & \mathbf{0}'_{(26\times1)} & \mathbf{0}'_{(53\times1)} \\ 1 & \mathbf{0}'_{(5\times1)} & \mathbf{0}'_{(10\times1)} & \mathbf{0}'_{(26\times1)} & -\mathbf{1}'_{(53\times1)} \\ \mathbf{0}_{(5\times1)} & \boldsymbol{I}_5 & -\boldsymbol{C}^I & \mathbf{0}_{(5\times26)} & \mathbf{0}_{(5\times53)} \\ \mathbf{0}_{(26\times1)} & \mathbf{0}_{(26\times5)} & \mathbf{0}_{(26\times10)} & \boldsymbol{I}_{26} & -\boldsymbol{C}^E \end{bmatrix},$$

where $\boldsymbol{C}^I$ and $\boldsymbol{C}^E$ are the contemporaneous aggregation matrix mapping the bts into the uts for the Income and the Expenditure sides, respectively (Bisaglia *et al.*, 2020).

By exploiting this result, cross-temporal forecast reconciliation is now applied within the same forecasting experiment designed by Athanasopoulos *et al.* (2020)[4] extended to consider semi-annual and annual forecasts as well. For the available time series span (1984:Q4 - 2018:Q1), quarterly base forecasts from 1 up to 4 quarters ahead have been obtained for the $n = 95$ separate time series through simple univariate ARIMA models selected using the `auto.arima` function of the `R`-package `forecast` (Hyndman *et al.*, 2023). The forecasting experiment uses a recursive training sample with expanding window length. The first training sample is set from 1984:Q4 to 1994:Q3, and the last ends on 2017:Q1, for a total of 91 forecast origins. Similarly, in the same automatic fashion, we have computed (i) one and two-step ahead forecasts for the time series obtained by temporal aggregation of two successive quarters, and (ii) one-step-ahead forecasts for the time series obtained by temporal aggregation of four successive quarters. All the base forecasts have been reconciled using the `R`-package `FoReco` (Girolimetto and Di Fonzo, 2023a).

1.8.1 Performance measures for multiple comparisons

We evaluate the performance of multiple (say, $J > 1$) forecast reconciliation procedures through forecast accuracy indices calculated on the forecast error

$$\widehat{e}^{[k],h}_{i,j,t} = y^{[k]}_{i,t+h} - \widehat{y}^{[k],h}_{i,j,t}, \qquad \begin{matrix} i = 1, \dots, 95, \\ j = 0, \dots, J, \end{matrix} \quad t = 1, \dots, 91, \quad \begin{matrix} k \in \mathcal{K}, \\ h = 1, \dots, h_k, \end{matrix}$$

where y and $\widehat{y}$ are the observed and forecasted values, respectively, i denotes the series ($i = 1, \dots, 32$, for the uts, $i = 33, \dots, 95$, for the bts), $j = 0$ denotes the base forecasts, t is the forecast origin ($t = 1$ corresponds to 1994:Q3), $\mathcal{K} = \{4, 2, 1\}$, and $h_4 = 1$, $h_2 = 2$, $h_1 = 4$, are the forecast horizons for annual, semi-annual, and quarterly time series, respectively. The accuracy is evaluated using the Average Relative Mean Square Error (AvgRelMSE, Davydenko and Fildes, 2013; Kourentzes and Athanasopoulos, 2019; Athanasopoulos and Kourentzes, 2022), obtained by transforming the MSE index, given by the average across all 91 forecasts origins of the squared forecast errors:

$$\text{MSE}^{[k],h}_{i,j} = \frac{1}{91} \sum_{t=1}^{91} \left(\widehat{e}^{[k],h}_{i,j,t} \right)^2, \qquad \begin{matrix} i = 1, \dots, 95, \quad k \in \mathcal{K}, \\ j = 0, \dots, J, \quad h = 1, \dots, h_k. \end{matrix} \tag{1.45}$$

The AvgRelMSE is the geometric mean across all 95 series of the MSE ratio[5] of a forecast over a benchmark given by the base, incoherent ARIMA forecasts, across all

[4] We did not change this first, crucial step in the forecast reconciliation workflow, since the focus is on the potential of cross-temporal forecast reconciliation.

[5] Davydenko and Fildes (2013) develop the Average Relative MAE, based on the Mean Absolute Error, and suggest that this formulation 'can also be extended to other measures of dispersion or loss functions', as the AvgRelMSE in (1.46).

evaluation samples, for a given horizon h:

$$\text{AvgRelMSE}_{j}^{[k],h} = \left(\prod_{i=1}^{95} \text{rMSE}_{i,j}^{[k],h}\right)^{\frac{1}{95}}, \qquad \begin{array}{l} j = 0, \dots, J, \\ k \in \mathcal{K}, \\ h = 1, \dots, h_k, \end{array} \tag{1.46}$$

where $\text{rMSE}_{i,j}^{[k],h} = \dfrac{\text{MSE}_{i,j}^{[k],h}}{\text{MSE}_{i,0}^{[k],h}}$ is the relative MSE. If a forecast outperforms the base forecasts, then the AvgRelMSE becomes smaller than one and vice-versa, and the percentage improvement in accuracy over the benchmark can be calculated as $\left(1 - \text{AvgRelMSE}_{j}^{[k],h}\right) \times 100$. Expression (1.46), which refers to all 95 time series, can be re-stated for (i) groups of variables (e.g., bts and uts), (ii) multiple forecast horizons (e.g., $h = 1 - 4$ for quarterly forecasts, $k = 1$; $h = 1 - 2$ for semi-annual forecasts, $k = 2$), (iii) different temporal aggregation levels over the whole forecast horizon (e.g., accuracy indices for the whole temporal hierarchy of each series). To give a complete picture of the evaluation results, in the following subsection, we show and discuss the MSE-based accuracy indices, at multiple timescales and forecast horizons, for a set of selected forecast reconciliation procedures. Furthermore, we use the non-parametric Friedman and the post-hoc 'Multiple Comparison with the Best' (MCB) Nemenyi tests (Koning *et al.*, 2005; Kourentzes and Athanasopoulos, 2019; Makridakis *et al.*, 2022) to establish if the forecasting performances of the considered techniques are significantly different.

1.8.2 The considered forecast reconciliation procedures

The empirical application mainly aims to evaluate the performance of the most convincing new cross-temporal reconciliation procedures, which are those using residual-based approximations of the covariance matrix, compared to the state-of-the-art point forecast reconciliation procedures. At this end, we consider five selected procedures recently proposed in the hierarchical forecasting literature, five (two-step and iterative) variants of the KA approach, and three optimal combination forecast procedures:

- cs-shr (Wickramasuriya *et al.*, 2019),
- t-wlsv (Athanasopoulos *et al.*, 2017),
- t-acov (Nystrup *et al.*, 2020),
- t-sar1 (Nystrup *et al.*, 2020),
- kah-wlsv-shr (Kourentzes and Athanasopoulos, 2019),
- tcs-acov-shr, i.e. two-step t-acov + cs-shr,
- tcs-sar1-shr, i.e. two-step t-sar1 + cs-shr,
- ite-wlsv-shr, i.e. iterative t-wlsv + cs-shr,

- ite-acov-shr, i.e. iterative t-acov + cs-shr,
- ite-sar1-shr, i.e. iterative t-sar1 + cs-shr,
- oct-wlsv, i.e. $\boldsymbol{W} = \widehat{\boldsymbol{W}}_{\text{wlsv}}$,
- oct-bdshr, i.e. $\boldsymbol{W} = \widehat{\boldsymbol{W}}^{BD}_{\text{shr}}$,
- oct-acov, i.e. $\boldsymbol{W} = \widehat{\boldsymbol{W}}_{\text{acov}}$.

The first five procedures have proven well performing in several empirical applications (Athanasopoulos *et al.*, 2020, 2017; Bisaglia *et al.*, 2020; Nystrup *et al.*, 2020). The one-dimension reconciliation procedures (cs-shr, t-wlsv, t-acov, and t-sar1) do not give fully coherent forecasts. Rather, as far as they are expected to improve the base forecasts, the best-practice one-dimension procedures should be viewed as stricter benchmarks for the cross-temporal forecast reconciliation techniques. These are requested to give accurate one-number-forecasts as well. The forecasting experiment was designed to evaluate the capability of the cross-temporal forecast reconciliation procedures to improve the forecast accuracy as compared to (i) the base forecasts and (ii) the most performing one-dimension forecast reconciliation methods. In addition, the experiment should help in assessing (iii) the performance of both KA-variants (two-step and iterative procedures) and optimal combination forecasts as compared to the original proposal by KA, and (iv) the feasibility and the accuracy of the optimal combination cross-temporal reconciliation procedures, which for the time being - even when they are computed using the in-sample residuals - are based on rather simple/unrealistic approximations of the covariance matrix (see Section 1.6.1). As for this last point, we investigate any significant difference between the reconciled forecasts produced by the most performing heuristic and optimal combination procedures.

1.8.3 Main results

Table 1.2 presents the AvgRelMSE's obtained for the forecasting techniques (base + 13 reconciliation procedures) listed in the previous sub-section. We provide results for all 95 component time series, and for the 32 upper-level and the 63 bottom-level time series separately. The results are shown by the level of temporal aggregation and forecast horizon. The lowest error is highlighted in red boldface at each column, while values greater than one, which mean that the reconciled forecasts are worse than the base ones, are highlighted in black boldface.

Most of the data in the table are represented in Figure 1.5, containing the graphs of the AvgRelMSE's for the considered procedures, across all forecast horizons, by temporal aggregation level of the forecasted series. The procedures have been put in the order given by the overall AvgRelMSE. This seems a good compromise to represent such a multiple comparison. Figure 1.6 shows the Multiple Comparison with the Best (MCB)

Reconciliation procedure	**Quarterly** 1	2	3	4	1-4	**Semi-annual** 1	2	1-2	**Annual** 1	**All**
					all 95 series					
base	1	1	1	1	1	1	1	1	1	1
cs-shr	0.958	0.970	0.976	0.982	0.972	0.953	0.978	0.965	0.966	0.969
t-wlsv	**1.002**	0.999	0.988	0.985	0.993	0.844	0.932	0.887	0.773	0.928
t-acov	0.978	0.991	0.999	0.989	0.989	0.825	0.935	0.879	0.769	0.923
t-sar1	**1.002**	0.999	0.988	0.985	0.994	0.845	0.932	0.887	0.773	0.928
kah-wlsv-shr	0.968	0.970	0.960	**0.960**	0.965	0.818	**0.909**	0.862	0.752	0.901
tcs-acov-shr	0.945	**0.958**	0.971	0.963	0.959	0.798	0.912	0.853	0.748	0.895
tcs-sar1-shr	0.968	0.970	0.960	0.960	0.965	0.818	0.909	0.862	0.752	0.901
ite-wlsv-shr	0.961	0.968	**0.959**	0.960	0.962	0.815	0.909	0.861	0.751	0.900
ite-acov-shr	**0.940**	0.958	0.971	0.965	**0.959**	**0.796**	0.913	**0.852**	**0.748**	**0.895**
ite-sar1-shr	0.961	0.968	0.959	0.961	0.962	0.815	0.909	0.861	0.751	0.900
oct-wlsv	0.969	0.972	0.962	0.963	0.967	0.820	0.913	0.865	0.756	0.904
oct-bdshr	0.984	0.980	0.962	0.967	0.973	0.830	0.914	0.871	0.757	0.910
oct-acov	0.955	0.965	0.977	0.971	0.967	0.801	0.919	0.858	0.753	0.902
					32 upper series					
base	1	1	1	1	1	1	1	1	1	1
cs-shr	**0.916**	**0.927**	0.930	0.932	**0.926**	0.917	0.939	0.928	0.923	0.926
t-wlsv	**1.006**	**1.009**	0.991	0.992	1.000	0.856	0.939	0.896	0.768	0.933
t-acov	**1.002**	**1.015**	0.992	0.993	**1.000**	0.854	0.938	0.895	0.768	0.9332
t-sar1	**1.007**	**1.009**	0.991	0.992	1.000	0.856	0.939	0.896	0.768	0.933
kah-wlsv-shr	0.940	0.947	0.928	0.930	0.936	0.800	0.877	0.837	0.715	0.873
tcs-acov-shr	0.941	0.944	0.931	0.933	0.937	0.796	0.878	0.836	0.715	0.873
tcs-sar1-shr	0.940	0.946	0.928	0.930	0.936	0.800	0.877	0.837	0.715	0.873
ite-wlsv-shr	0.925	0.942	0.922	0.927	0.929	0.793	0.874	0.833	0.711	**0.867**
ite-acov-shr	0.928	0.940	0.926	0.931	0.931	**0.789**	0.875	**0.831**	**0.711**	0.868
ite-sar1-shr	0.926	0.942	**0.922**	**0.927**	0.929	0.794	**0.874**	0.833	0.711	0.867
oct-wlsv	0.941	0.951	0.932	0.933	0.939	0.803	0.881	0.841	0.720	0.876
oct-bdshr	0.945	0.956	0.925	0.934	0.940	0.809	0.879	0.843	0.717	0.877
oct-acov	0.939	0.950	0.935	0.937	0.940	0.798	0.884	0.840	0.719	0.876
					63 bottom series					
base	1	1	1	1	1	1	1	1	1	1
cs-shr	0.981	0.993	1.000	**1.009**	0.996	0.971	0.999	0.985	0.988	0.991
t-wlsv	0.999	0.995	0.986	0.982	0.990	0.839	0.928	0.882	0.775	0.925
t-acov	0.966	0.980	**1.002**	0.986	0.983	0.811	0.934	0.870	0.770	0.917
t-sar1	0.999	0.994	0.986	0.982	0.990	0.839	0.928	0.882	0.775	0.925
kah-wlsv-shr	0.983	0.982	**0.976**	**0.976**	0.979	0.827	**0.925**	0.875	0.771	0.916
tcs-acov-shr	0.947	**0.966**	0.992	0.978	**0.971**	**0.799**	0.929	**0.862**	**0.766**	**0.907**
tcs-sar1-shr	0.983	0.982	0.976	0.976	0.979	0.827	0.925	0.875	0.771	0.916
ite-wlsv-shr	0.980	0.981	0.978	0.978	0.979	0.826	0.928	0.875	0.772	0.917
ite-acov-shr	**0.946**	0.968	0.995	0.983	0.973	0.799	0.932	0.863	0.767	0.909
ite-sar1-shr	0.980	0.982	0.978	0.978	0.979	0.826	0.928	0.876	0.773	0.917
oct-wlsv	0.984	0.983	0.978	0.979	0.981	0.829	0.929	0.878	0.775	0.919
oct-bdshr	**1.004**	0.992	0.981	0.984	0.990	0.840	0.933	0.885	0.778	0.927
oct-acov	0.964	0.973	0.998	0.988	0.981	0.803	0.936	0.867	0.771	0.915

Table 1.2: *AvgRelMSE at any temporal aggregation level and any forecast horizon. Bold entries identify the best performing approaches. Red entries identify the approaches worsening the automatic ETS base forecasts' accuracy.*

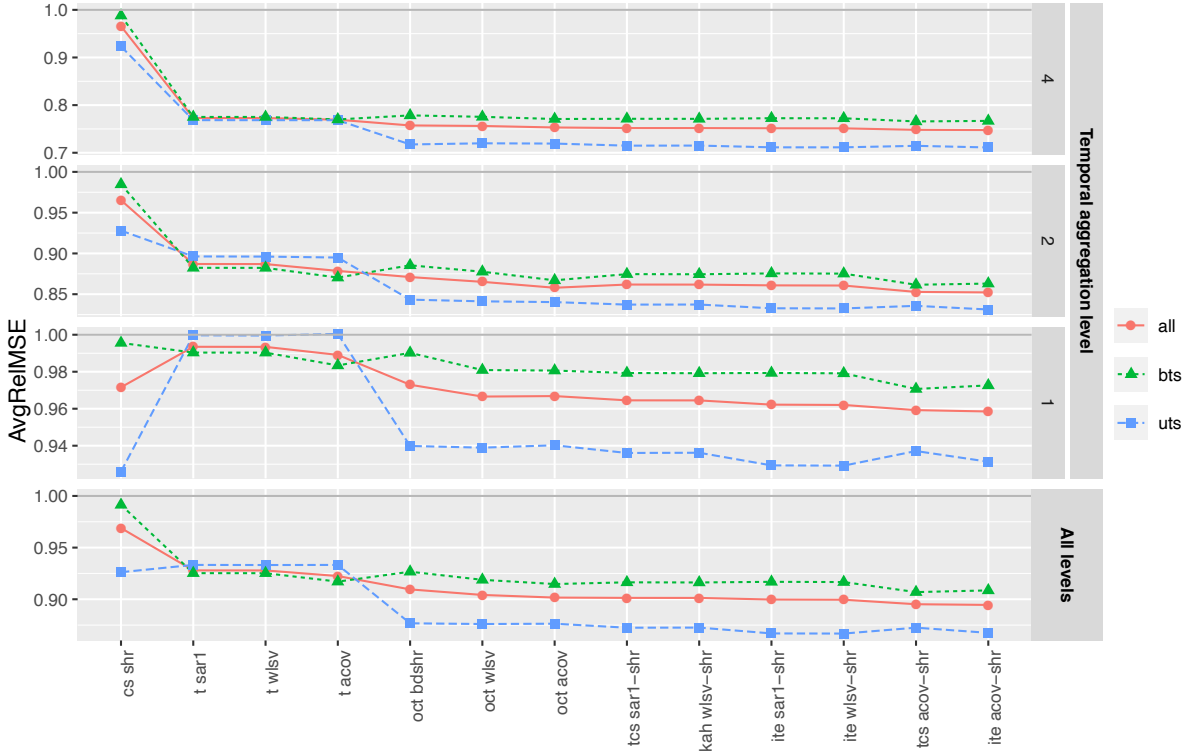


Figure 1.5: *Average Relative MSE across all series and forecast horizons, by temporal aggregation level.*

Nemenyi test. The Friedman test has shown that the forecasts given by the considered procedures are different both when all temporal aggregation levels and forecast horizons (left panel), and when only one-step-ahead quarterly forecasts (right panel), are considered. The main results found on this dataset can be summarized as follows:

- the cross-temporal reconciliation reduces the AvgRelMSE for the uts (the most important variables for the decision maker, e.g. *GDP*) at any temporal aggregation level and any forecast horizon;
- this accuracy improvement is less marked, though yet visually evident, for the bottom level series, as compared to the reconciled forecasts through temporal hierarchies alone, which however are cross-sectionally incoherent;
- each iterative procedure performs better than its two-step counterpart;
- within the cross-temporal procedures, the heuristic procedures provide better results than the optimal combination ones.

Looking at the performances of each procedure, it is worth noting that cs-shr scores first as for the quarterly forecasts of the uts, and almost always improves on the base forecasts' accuracy, regardless of series' group, temporal aggregation level and forecast horizon. The only exception is an AvgRelMSE greater than 1 (1.009) for the bts quarterly forecasts at horizon 4. In addition, from the right panel of Figure 1.6 we observe

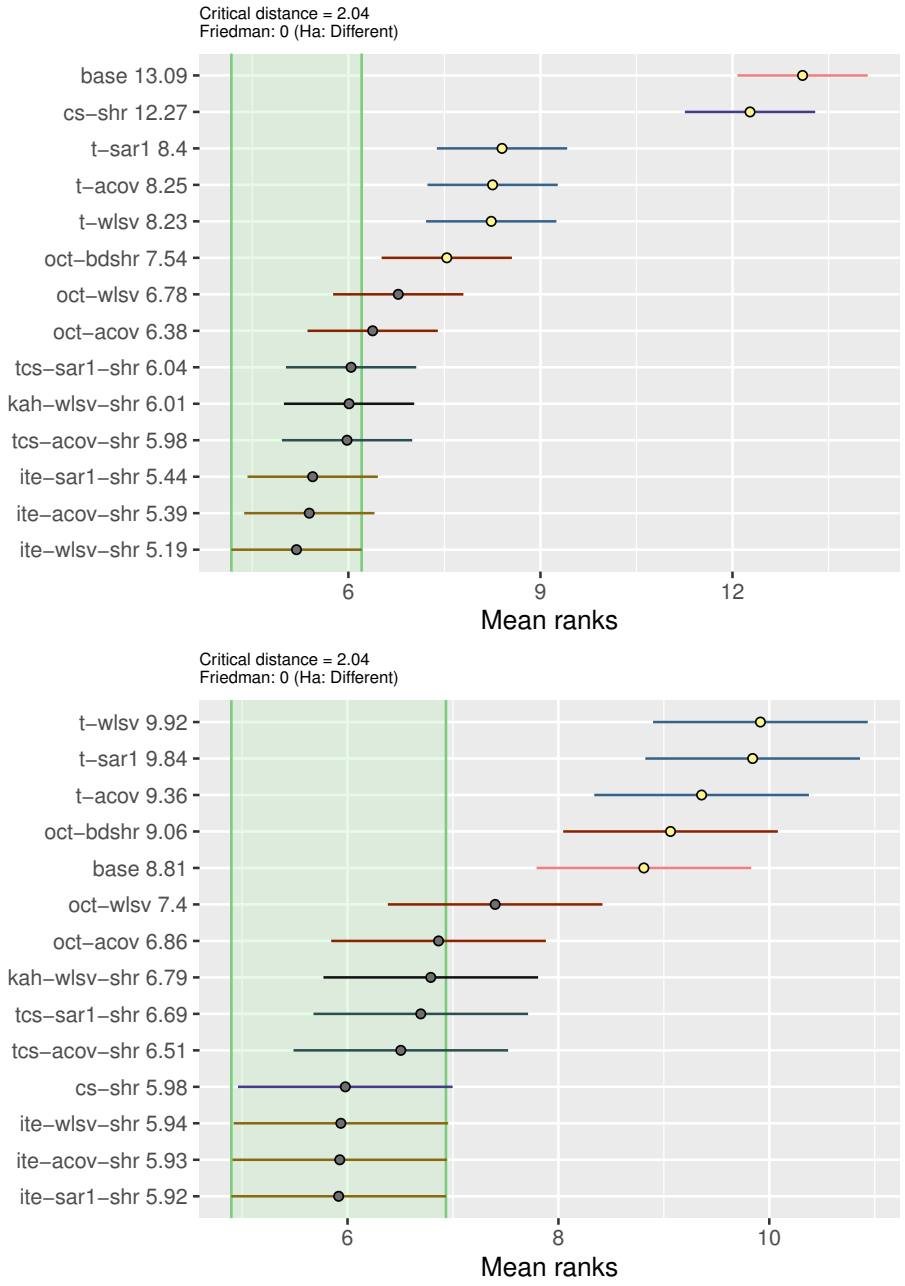


Figure 1.6: *MCB Nemenyi test results: average ranks and 95% confidence intervals. The reconciliation procedures are sorted vertically according to the MSE mean rank (i) across all time frequencies and forecast horizons (left), and (ii) for one-step-ahead quarterly forecasts (right). The mean rank of each method is displayed to the right of their names. If the intervals of two forecast reconciliation procedures do not overlap, this indicates a statistically different performance. Thus, methods that do not overlap with the green interval are considered significantly worse than the best, and vice-versa.*

that, when considered only on quarterly basis, the one-step-ahead forecasts for all series provided by cs-shr are (temporally incoherent and) not significantly different from those provided by the best procedure (which in this case is ite-acov-shr). However, since this reconciliation procedure does not account for the temporal dimension, its relative performance worsens (i.e., the cross-temporal procedures improve on the base forecasts more than cs-shr, as shown in the top panels in Figure 1.5) as the temporal aggregation level increases. Overall, ite-acov-shr always scores best for all series and all forecast horizons, and second-best for the bts series and all forecast horizons, while tcs-acov-shr scores second and first, in turn. However, ite-acov-shr shows good results for the uts forecasts as well. In this case, the best performances are given by ite-sar1-shr and ite-wlsv-shr. The differences in the forecasts produced by all the considered heuristic procedures are not statistically significant at any temporal aggregation level and forecast horizon. Furthermore, two optimal combination procedures (oct-acov and oct-wlsv) produce reconciled forecasts not significantly different from the best procedure according to the Nemenyi test, while oct-bdshr is significantly (worse and) different from the best forecast reconciliation procedure.

To conclude, the general improvement registered on average (last column of Table 1.2) by the cross-temporal reconciliation procedures may be considered a positive outcome. They combine an acceptable forecasting performance at the quarterly level with a good performance at the semi-annual and annual-levels, with the additional feature that the complete system of forecasts is internally and temporally coherent.

Chapter 2

Cross-temporal reconciliation of solar forecasts

2.1 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[1] 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 (Sengupta *et al.*, 2017). The prediction of solar irradiation despite its variability is particularly important as it

[1]The Paris Agreement is a legally binding international treaty on climate change. Its goal is to limit global warming to well below 2, preferably to 1.5 degrees Celsius, compared to pre-industrial levels. To achieve this long-term temperature goal, countries aim to reach global peaking of greenhouse gas emissions as soon as possible to achieve a climate neutral world (i.e., with net-zero greenhouse gas emissions) by mid-century.

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 (Antonanzas *et al.*, 2016). 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 cross-sectional and temporal resolution of the available forecasting models (Benavides Cesar *et al.*, 2022, and references therein) 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 fronteer 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 were carried out in cross-sectional and temporal domains separately. To further improve prediction accuracy, in the third case these two reconciliation approaches were sequentially applied. During the replication of the forecasting experiment[2], 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 both geographical (i.e., cross-sectional) and temporal hierarchies.

At a given time point, a cross-sectional hierarchy describes the accounting relationships linking the series of different levels of contemporaneous aggregation (in the simplest

[2]The forecasting experiment grounds on the documentation files and data made available by Yang *et al.* (2017b).

case, the total variable is equal to the sum of its most disaggregated component series). A temporal hierarchy describes the temporal aggregated time series of a single variable originally defined at a specific time frequency, obtained by non-overlapping sums of the high-frequency observations (e.g., a daily time series may be obtained as the sum of the twentyfour values in each single day of an hourly time series). 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 by Kourentzes and Athanasopoulos (2019), Yagli *et al.* (2019), Spiliotis *et al.* (2020), and Punia *et al.* (2020). In Chapter 1, we have shown 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[3].

In this chapter, 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 in Chapter 1, 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 normalized Root Mean Square Error and the normalized Mean Bias Error are used to measure forecasting accuracy, and a statistical multiple comparison procedure is performed to rank the approaches.

The chapter is organized as follows. The deterministic (point) cross-temporal forecast reconciliation framework is described in Section 2.2, and in Section 2.3 some useful connections between apparently different approaches are shown. The forecasting experiment of Yagli *et al.* (2019) is replicated and discussed in Section 2.4, and the performance of the proposed forecasting approaches is presented in Section 2.5.

[3]All the forecast reconciliation procedures considered in Chapter 1, and in this chapter, are available in the `R` package `FoReco` (Girolimetto and Di Fonzo, 2023a).

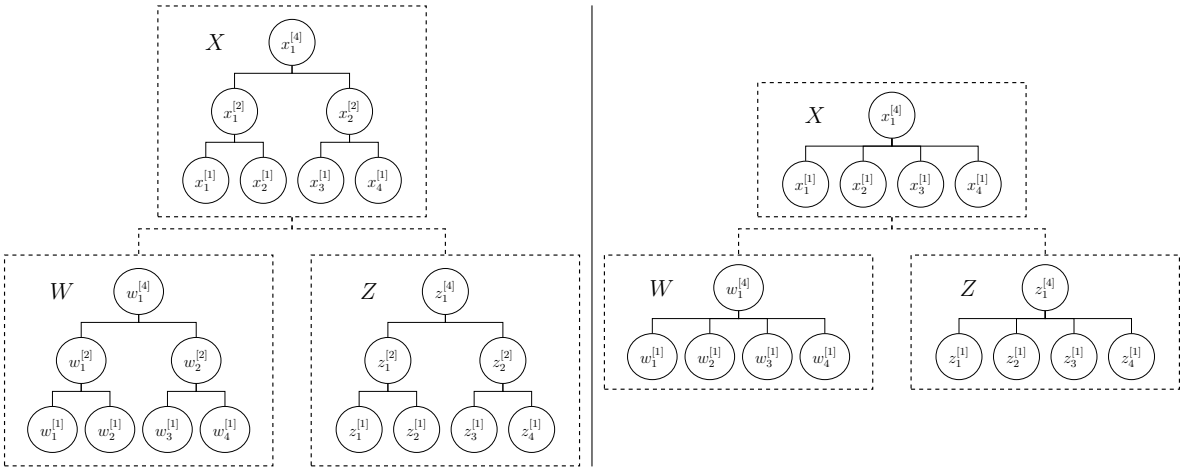


Figure 2.1: *Complete (left) and reduced (right) cross-temporal hierarchies for a quarterly two-level hierarchical time series.*

2.2 Cross-temporal point forecast reconciliation: a recap

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[4], 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 2.1 gives a visual representation of such cross-temporal hierarchy for a time cycle of 1 year. While the left panel of Figure 2.1 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 $k \in \mathcal{K} = \{4, 2, 1\}$ (i.e., $x_\tau^{[k]} = w_\tau^{[k]} + z_\tau^{[k]}$, $\tau = 1, \ldots, 4/k$). 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

[4]In this chapter, 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 introduced in Chapter 1 and discussed in Chapter 4.

as

$$\begin{aligned} v_1^{[4]} &= v_1^{[2]} + v_2^{[2]} && \text{the annual value is the sum of the two semi-annual values,} \\ v_1^{[2]} &= v_1^{[1]} + v_2^{[1]} && \text{the first half-year value is the sum of the first two quarters' values,} \\ v_2^{[2]} &= v_3^{[1]} + v_4^{[1]} && \text{the second half-year value is the sum of the last two quarters' values,} \end{aligned}$$

and thus

$$v_1^{[4]} = v_1^{[1]} + v_2^{[1]} + v_3^{[1]} + v_4^{[1]} \quad \text{the annual value is the sum of the four quarterly values.}$$

All the relationships so far can be expressed in more compact form, using matrix notation. Denote by $\boldsymbol{y}_\tau^{[k]} = \left[x_\tau^{[k]}\ w_\tau^{[k]}\ z_\tau^{[k]}\right]'$ the (3×1) vector of the observations for temporal granularity $k \in \mathcal{K}$ 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:

$$x_\tau^{[k]} = \boldsymbol{C} \begin{bmatrix} w_\tau^{[k]} \\ z_\tau^{[k]} \end{bmatrix}, \quad \boldsymbol{y}_\tau^{[k]} = \boldsymbol{S} \begin{bmatrix} w_\tau^{[k]} \\ z_\tau^{[k]} \end{bmatrix}, \quad \boldsymbol{U}'\boldsymbol{y}_\tau^{[k]} = 0, \qquad \tau = 1, \ldots, \frac{4}{k}, \quad k \in \mathcal{K},$$

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

$$\boldsymbol{C} = \begin{bmatrix} 1 & 1 \end{bmatrix}, \quad \boldsymbol{S} = \begin{bmatrix} \boldsymbol{C} \\ \boldsymbol{I}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{U}' = \begin{bmatrix} \boldsymbol{I}_1 & -\boldsymbol{C} \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 \end{bmatrix},$$

with $\boldsymbol{I}_l$ 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

$$\boldsymbol{K} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \quad \boldsymbol{R} = \begin{bmatrix} \boldsymbol{K} \\ \boldsymbol{I}_4 \end{bmatrix} \quad \text{and} \quad \boldsymbol{Z}' = \begin{bmatrix} \boldsymbol{I}_3 & -\boldsymbol{K} \end{bmatrix}$$

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

$$\begin{bmatrix} v_1^{[4]} \\ v_1^{[2]} \\ v_2^{[2]} \end{bmatrix} = \boldsymbol{K} \begin{bmatrix} v_1^{[1]} \\ v_2^{[1]} \\ v_3^{[1]} \\ v_4^{[1]} \end{bmatrix} = \boldsymbol{K}\boldsymbol{v}^{[1]}, \quad \boldsymbol{v} = \begin{bmatrix} v_1^{[4]} \\ v_1^{[2]} \\ v_2^{[2]} \\ v_1^{[1]} \\ \vdots \\ v_4^{[1]} \end{bmatrix} = \boldsymbol{R}\boldsymbol{v}^{[1]}, \quad \boldsymbol{Z}'\boldsymbol{v} = \boldsymbol{0}_{(3\times 1)}.$$

Reduced temporal hierarchies can be obtained by simply eliminating the appropriate rows[5] from matrix $\boldsymbol{K}$.

To simultaneously consider cross-sectional and temporal aggregation relationships, all the nodes of the complete cross-temporal hierarchy in Figure 2.1 can be expressed in terms of the quarterly time series $w_\tau^{[1]}$ and $z_\tau^{[1]}$, $\tau = 1, \dots, 4$, according to the *structural* representation:

$$\boldsymbol{y} = \boldsymbol{F}\boldsymbol{b}^{[1]}, \tag{2.1}$$

that is

$$\underbrace{\begin{bmatrix} x_1^{[4]} \\ x_1^{[2]} \\ x_2^{[2]} \\ x_1^{[1]} \\ x_2^{[1]} \\ x_3^{[1]} \\ x_4^{[1]} \\ w_1^{[4]} \\ w_1^{[2]} \\ w_2^{[2]} \\ w_1^{[1]} \\ w_2^{[1]} \\ w_3^{[1]} \\ w_4^{[1]} \\ z_1^{[4]} \\ z_1^{[2]} \\ z_2^{[2]} \\ z_1^{[1]} \\ z_2^{[1]} \\ z_3^{[1]} \\ z_4^{[1]} \end{bmatrix}}_{\boldsymbol{y}} = \underbrace{\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}}_{\boldsymbol{F}} \underbrace{\begin{bmatrix} w_1^{[1]} \\ w_2^{[1]} \\ w_3^{[1]} \\ w_4^{[1]} \\ z_1^{[1]} \\ z_2^{[1]} \\ z_3^{[1]} \\ z_4^{[1]} \end{bmatrix}}_{\boldsymbol{b}^{[1]}},$$

where $\boldsymbol{y} = \left[\boldsymbol{x}' \; \boldsymbol{w}' \; \boldsymbol{z}'\right]'$ is the vector containing the data for all variables at any temporal granularity, $\boldsymbol{b}^{[1]} = \left[\boldsymbol{w}^{[1]\prime} \; \boldsymbol{z}^{[1]\prime}\right]'$ is the vector of the high-frequency bottom time series, and $\boldsymbol{F}$ is the cross-temporal summing matrix mapping $\boldsymbol{b}^{[1]}$ into $\boldsymbol{y}$. Expression (2.1) is the natural extension of the cross-sectional structural representation firstly shown by

[5]This option is available in the `R` package `FoReco` (Girolimetto and Di Fonzo, 2023a).

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 (see Chapter 1).

Besides the number of variables forming the cross sectional hierarchy ($n = 3$ in the above example), two crucial aspects affecting the dimension of matrix $\boldsymbol{F}$ 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 $\boldsymbol{R}$ defining all infra-day temporal granularities ($\mathcal{K} = \{24, 12, 8, 6, 4, 3, 2, 1\}$) has dimension (60×24), and is equal to

$$\boldsymbol{R} = \Big[\mathbf{1}_{24} \quad \boldsymbol{I}_2 \otimes \mathbf{1}_{12} \quad \boldsymbol{I}_3 \otimes \mathbf{1}_8 \quad \boldsymbol{I}_4 \otimes \mathbf{1}_6 \quad \boldsymbol{I}_6 \otimes \mathbf{1}_4 \quad \boldsymbol{I}_8 \otimes \mathbf{1}_3 \quad \boldsymbol{I}_{12} \otimes \mathbf{1}_2 \quad \boldsymbol{I}_{24}\Big]'.$$

Matrix $\boldsymbol{F}$ is thus a large and sparse matrix[6] 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 chapter (see Section 2.4) is $n = 324$, with $n_b = 318$ bottom time series, thus matrix $\boldsymbol{F}$ has dimension (19,440 × 7,632). However, if the interest in forecasting at certain time granularities is low, this dimensonality issue may be mitigated by considering only part of the temporal granularities between the highest and lowest temporal frequencies[7]. For example, if one considers only hourly and daily forecasts, the reduced $\boldsymbol{F}$ is a (8,100 × 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.

[6]Sparse matrices require less memory than dense matrices, and allow some computations to be more efficient (Paige and Saunders, 1982; Davis, 2006; Bates *et al.*, 2023).

[7]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.

2.2.1 Notation

Suppose we want to forecast a n-variate high-frequency hierarchical time series $\left\{\boldsymbol{y}_t^{[1]}\right\}_{t=1}^{T}$, 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 $\boldsymbol{y}_t^{[1]}$, 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-overlaping aggregation is a multiple of m, and define N the number of the lowest-frequency series observations, i.e. $N = T/m$. Let $\mathcal{K}$ be the set of p factors of m, in descending order, $\mathcal{K} = \{k_p, k_{p-1}, \dots, k_2, k_1\}$, where $k_p = m$ and $k_1 = 1$, and define $k^* = \sum_{j=2}^{p} k_j$.

Following Chapter 1, denote $\boldsymbol{Y}_{N+h} \equiv \boldsymbol{Y}$ the $[n \times (k^* + m)]$ matrix of the target forecasts for any temporal granularity, with low-frequency temporal horizon h, given by:

$$\begin{aligned}\boldsymbol{Y} &= \left[\boldsymbol{Y}^{[m]}\ \boldsymbol{Y}^{[k_{p-1}]} \dots \boldsymbol{Y}^{[k_2]}\ \boldsymbol{Y}^{[1]}\right] \\ &= \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}^{[m]} & \boldsymbol{A}^{[k_{p-1}]} & \dots & \boldsymbol{A}^{[k_2]} & \boldsymbol{A}^{[1]} \\ \boldsymbol{B}^{[m]} & \boldsymbol{B}^{[k_{p-1}]} & \dots & \boldsymbol{B}^{[k_2]} & \boldsymbol{B}^{[1]} \end{bmatrix},\end{aligned}$$

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

$$\begin{aligned}\widehat{\boldsymbol{Y}} &= \left[\widehat{\boldsymbol{Y}}^{[m]}\ \widehat{\boldsymbol{Y}}^{[k_{p-1}]} \dots \widehat{\boldsymbol{Y}}^{[k_2]}\ \widehat{\boldsymbol{Y}}^{[1]}\right] \\ &= \begin{bmatrix} \widehat{\boldsymbol{A}}^{[m]} & \widehat{\boldsymbol{A}}^{[k_{p-1}]} & \dots & \widehat{\boldsymbol{A}}^{[k_2]} & \widehat{\boldsymbol{A}}^{[1]} \\ \widehat{\boldsymbol{B}}^{[m]} & \widehat{\boldsymbol{B}}^{[k_{p-1}]} & \dots & \widehat{\boldsymbol{B}}^{[k_2]} & \widehat{\boldsymbol{B}}^{[1]} \end{bmatrix}.\end{aligned}$$

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:

$$\boldsymbol{U}'\boldsymbol{Y} = \boldsymbol{0}_{[n_a \times (k^*+m)]} \quad \text{and} \quad \boldsymbol{Z}'\boldsymbol{Y}' = \boldsymbol{0}_{[k^* \times n]},$$

while

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

where $\boldsymbol{U}' = \Big[\boldsymbol{I}_{n_a} \; -\boldsymbol{C}\Big]$ and $\boldsymbol{Z}' = \Big[\boldsymbol{I}_{k^*} \; -\boldsymbol{K}\Big]$ 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

$$\boldsymbol{y}_{\tau}^{[k]} = \boldsymbol{S}\boldsymbol{b}_{\tau}^{[k]}, \quad \tau = 1, \dots, \frac{m}{k}, \quad \rightarrow \quad \boldsymbol{Y}^{[k]} = \boldsymbol{S}\boldsymbol{B}^{[k]} \qquad k \in \mathcal{K},$$

that is, in compact form

$$\boldsymbol{Y} = \boldsymbol{S}\boldsymbol{B} \quad \rightarrow \quad \boldsymbol{y} = (\boldsymbol{S} \otimes \boldsymbol{I}_{k^*+m})\, \boldsymbol{b},$$

where $\boldsymbol{y} = \text{vec}\,(\boldsymbol{Y}')$, and $\boldsymbol{b} = \text{vec}\,(\boldsymbol{B}')$.

Temporal hierarchies structural representation for n individual time series

$$\begin{array}{ll} \boldsymbol{a}_i = \boldsymbol{R}\boldsymbol{a}_i^{[1]}, & i = 1, \dots, n_a \\ \boldsymbol{b}_j = \boldsymbol{R}\boldsymbol{b}_j^{[1]}, & j = 1, \dots, n_b \end{array} \quad \rightarrow \quad \begin{array}{l} \boldsymbol{A} = \boldsymbol{R}\boldsymbol{A}^{[1]\prime} \\ \boldsymbol{B} = \boldsymbol{R}\boldsymbol{B}^{[1]\prime} \end{array},$$

that is

$$\boldsymbol{Y}' = \boldsymbol{R}\boldsymbol{Y}^{[1]\prime} \quad \rightarrow \quad \boldsymbol{y} = (\boldsymbol{I}_n \otimes \boldsymbol{R})\, \boldsymbol{y}^{[1]},$$

where $\boldsymbol{y}^{[1]} = \text{vec}\left(\boldsymbol{Y}^{[1]\prime}\right)$, and $\boldsymbol{R}$ is the temporal summing matrix (see Chapter 1).

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

$$\text{vec}\left(\widetilde{\boldsymbol{Y}}'_{\text{ct}(bu)}\right) = \boldsymbol{F}\text{vec}\left(\widehat{\boldsymbol{B}}^{[1]\prime}\right) \quad \leftrightarrow \quad \widetilde{\boldsymbol{y}}_{\text{ct}(bu)} = \boldsymbol{F}\widehat{\boldsymbol{b}}^{[1]}, \tag{2.2}$$

where $\boldsymbol{F} = (\boldsymbol{S} \otimes \boldsymbol{R})$ is the cross-temporal summing matrix, with $\otimes$ denoting the Kronecker product, and $\widehat{\boldsymbol{b}}^{[1]} = \text{vec}\left(\widehat{\boldsymbol{B}}^{[1]\prime}\right)$ is the vector containing the base forecasts of the high-frequency bottom time series. The cross-temporal bottom-up reconciliation can be tought 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, are 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($rec_{\text{te}}, bu_{\text{cs}}$), or ct($rec_{\text{cs}}, bu_{\text{te}}$), where '$rec_{\text{te}}$' and '$rec_{\text{cs}}$' denote a generic forecast reconciliation approach in time and in space, respectively.

2.2.3 Regression-based cross-temporal reconciliation

Let us consider the multivariate regression model

$$\widehat{\boldsymbol{Y}} = \boldsymbol{Y} + \boldsymbol{E},$$

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

$$\text{vec}\left(\widehat{\boldsymbol{Y}}'\right) = \text{vec}\left(\boldsymbol{Y}'\right) + \text{vec}\left(\boldsymbol{E}'\right) \quad \Leftrightarrow \quad \widehat{\boldsymbol{y}} = \boldsymbol{y} + \boldsymbol{\eta}, \tag{2.3}$$

where $\boldsymbol{\eta} = \text{vec}\left(\boldsymbol{E}'\right)$ is the cross-temporal reconciliation error with zero mean and p.d. covariance matrix $\boldsymbol{\Omega}_{ct}$. Assuming $\boldsymbol{\Omega}_{ct}$ known, the optimal combination reconciled forecasts $\widetilde{\boldsymbol{y}}_{\text{oct}} = \text{vec}\left(\widetilde{\boldsymbol{Y}}'_{\text{oct}}\right)$ are found by linearly constrained minimization of the generalized least squares (GLS) objective function

$$\left(\boldsymbol{y} - \widehat{\boldsymbol{y}}\right)' \boldsymbol{\Omega}_{ct}^{-1} \left(\boldsymbol{y} - \widehat{\boldsymbol{y}}\right) \qquad \text{s.t. } \boldsymbol{H}'\boldsymbol{y} = \boldsymbol{0}, \tag{2.4}$$

where

$$\boldsymbol{H}'\boldsymbol{y} = \begin{bmatrix} \boldsymbol{U}^* \\ \boldsymbol{I}_n \otimes \boldsymbol{Z} \end{bmatrix} \boldsymbol{y} = \boldsymbol{0}_{[n(k^*+m)\times 1]}$$

is a full row-rank cross-temporal zero-constraint matrix, with $\boldsymbol{U}^* = \left[\boldsymbol{0}_{(n_a m \times n k^*)}\ \boldsymbol{I}_m \otimes \boldsymbol{U}'\right] \boldsymbol{P}$, and $\boldsymbol{P}$ is the $[n(k^* + m) \times n(k^* + m)]$ commutation matrix, such that $\boldsymbol{P}\text{vec}(\boldsymbol{Y}) = \boldsymbol{y}$ (see Chapter 1).

The cross-temporally reconciled forecasts according to the projection approach solution (Byron, 1978, 1979; see also van Erven and Cugliari, 2015; Wickramasuriya *et al.*, 2019; Panagiotelis *et al.*, 2021 and Chapter 1) are given by

$$\widetilde{\boldsymbol{y}}_{\text{oct}} = \left[\boldsymbol{I}_{n(k^*+m)} - \boldsymbol{\Omega}_{ct}\boldsymbol{H}\left(\boldsymbol{H}'\boldsymbol{\Omega}_{ct}\boldsymbol{H}\right)^{-1}\boldsymbol{H}'\right]\widehat{\boldsymbol{y}} = \boldsymbol{M}_{ct}\widehat{\boldsymbol{y}}. \tag{2.5}$$

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

$$\widehat{\boldsymbol{y}} = \boldsymbol{F}\boldsymbol{\beta} + \boldsymbol{\eta},$$

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

forecast reconciliation formula is

$$\widetilde{\boldsymbol{y}}_{\text{oct}} = \boldsymbol{F}\left(\boldsymbol{F}'\boldsymbol{\Omega}_{ct}^{-1}\boldsymbol{F}\right)^{-1}\boldsymbol{F}'\boldsymbol{\Omega}_{ct}^{-1}\widehat{\boldsymbol{y}} = \boldsymbol{F}\boldsymbol{G}_{ct}\widehat{\boldsymbol{y}}, \tag{2.6}$$

with $\boldsymbol{G}_{ct} = \left(\boldsymbol{F}'\boldsymbol{\Omega}_{ct}^{-1}\boldsymbol{F}\right)^{-1}\boldsymbol{F}'\boldsymbol{\Omega}_{ct}^{-1}$. In Chapter 1, we considered the following approximations for the cross-temporal covariance matrix ('*oct*' stands for 'optimal cross-temporal'):

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

where the symbol $\odot$ denotes the Hadamard product, $\hat{\lambda}$ is an estimated shrinkage coefficient (Ledoit and Wolf, 2004), $\widehat{\boldsymbol{\Omega}}_{ct,D} = \boldsymbol{I}_{n(k^*+m)} \odot \widehat{\boldsymbol{\Omega}}_{ct}$, and $\widehat{\boldsymbol{\Omega}}_{ct}$ 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 $\boldsymbol{\Omega}_{ct} = \boldsymbol{W}$ is an $(n \times n)$ p.d. matrix):

$$\widetilde{\boldsymbol{y}} = \left[\boldsymbol{I}_n - \boldsymbol{W}\boldsymbol{U}\left(\boldsymbol{U}'\boldsymbol{W}\boldsymbol{U}\right)^{-1}\boldsymbol{U}'\right]\widehat{\boldsymbol{y}} = \boldsymbol{M}_{cs}\widehat{\boldsymbol{y}},$$

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

$$\widetilde{\boldsymbol{y}} = \boldsymbol{S}\left(\boldsymbol{S}'\boldsymbol{W}^{-1}\boldsymbol{S}\right)^{-1}\boldsymbol{S}'\boldsymbol{W}^{-1}\widehat{\boldsymbol{y}} = \boldsymbol{S}\boldsymbol{G}_{cs}\widehat{\boldsymbol{y}},$$

with $\boldsymbol{G}_{cs} = \left(\boldsymbol{S}'\boldsymbol{W}^{-1}\boldsymbol{S}\right)^{-1}\boldsymbol{S}'\boldsymbol{W}^{-1}$. 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 $\boldsymbol{\Omega}_{ct} = \boldsymbol{\Omega}$ is a $(k^* + m \times k^* + m)$ p.d. matrix):

$$\widetilde{\boldsymbol{y}} = \left[\boldsymbol{I}_{(k^*+m)} - \boldsymbol{\Omega}\boldsymbol{Z}\left(\boldsymbol{Z}'\boldsymbol{\Omega}\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}'\right]\widehat{\boldsymbol{y}} = \boldsymbol{M}_{te}\widehat{\boldsymbol{y}},$$

or, equivalently,

$$\widetilde{\boldsymbol{y}} = \boldsymbol{R}\left(\boldsymbol{R}'\boldsymbol{\Omega}^{-1}\boldsymbol{R}\right)^{-1}\boldsymbol{R}'\boldsymbol{\Omega}^{-1}\widehat{\boldsymbol{y}} = \boldsymbol{R}\boldsymbol{G}_{te}\widehat{\boldsymbol{y}},$$

with $\boldsymbol{G}_{te} = \left(\boldsymbol{R}'\boldsymbol{\Omega}^{-1}\boldsymbol{R}\right)^{-1}\boldsymbol{R}'\boldsymbol{\Omega}^{-1}$. Table 2.1 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 Chapter 1.

	Cross-sectional framework	**Temporal framework**
identity	cs(ols): $\boldsymbol{W} = \boldsymbol{I}_n$	te(ols): $\boldsymbol{\Omega} = \boldsymbol{I}_{k^*+m}$
structural	cs($struc$): $\boldsymbol{W} = \text{diag}(\boldsymbol{S}\boldsymbol{1}_{nb})$	te($struc$): $\boldsymbol{\Omega} = \text{diag}(\boldsymbol{R}\boldsymbol{1}_m)$
series variance	cs(wls): $\boldsymbol{W} = \widehat{\boldsymbol{W}}_D = \boldsymbol{I}_n \odot \widehat{\boldsymbol{W}}$	te($wlsv$): $\boldsymbol{\Omega} = \widehat{\boldsymbol{\Omega}}_{wlsv}$
MinT-shr	cs(shr): $\boldsymbol{W} = \hat{\lambda}\widehat{\boldsymbol{W}}_D + (1-\hat{\lambda})\widehat{\boldsymbol{W}}$	te(shr): $\boldsymbol{\Omega} = \hat{\lambda}\widehat{\boldsymbol{\Omega}}_D + (1-\hat{\lambda})\widehat{\boldsymbol{\Omega}}$
MinT-sam	cs(sam): $\boldsymbol{W} = \widehat{\boldsymbol{W}}$	te(sam): $\boldsymbol{\Omega} = \widehat{\boldsymbol{\Omega}}$

Table 2.1: *Approximations for the cross-sectional (Hyndman* et al.*, 2011, 2016; Wickramasuriya* et al.*, 2019; Chapter 1) and temporal (Athanasopoulos* et al.*, 2017; Chapter 1) covariance matrix to be used in a reconciliation approach.* $\widehat{\boldsymbol{W}}$ *(*$\widehat{\boldsymbol{\Omega}}$*) is the covariance matrix of the cross-sectional (temporal) one-step ahead in-sample forecast errors,* $\widehat{\boldsymbol{\Omega}}_{wlsv}$ *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* $\widehat{\boldsymbol{\Omega}}_D = \boldsymbol{I}_{k^*+m} \odot \widehat{\boldsymbol{\Omega}}$.

2.2.4 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 Chapter 1):

KA-Step 1 compute the temporally reconciled forecasts for each variable $i \in \{1, \dots, n\}$, and arrange them in the $[n \times (k+m)]$ matrix $\widetilde{\boldsymbol{Y}}_{\text{te}}$;

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

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

$$\widetilde{\boldsymbol{Y}}_{\text{KA}} = \left(\frac{1}{p}\sum_{k\in\mathcal{K}} \boldsymbol{M}_{\text{cs}}^{[k]}\right)\widetilde{\boldsymbol{Y}}_{\text{te}} = \overline{\boldsymbol{M}}\widetilde{\boldsymbol{Y}}_{\text{te}}. \tag{2.7}$$

In Chapter 1 we presented an iterative approach, that produces cross-temporal 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 ($\widetilde{\boldsymbol{Y}}_{\text{te}}^{(j)}$) for each variable $i \in \{1, \dots, n\}$ of $\widetilde{\boldsymbol{Y}}_{\text{cs}}^{(j-1)}$;

Step 2 compute the time-by-time cross-sectionally reconciled forecasts ($\widetilde{\boldsymbol{Y}}_{\text{cs}}^{(j)}$) for all the temporal aggregation levels of $\widetilde{\boldsymbol{Y}}_{\text{te}}^{(j)}$.

At $j = 0$, the starting values are given by $\widetilde{\boldsymbol{Y}}_{\text{cs}}^{(0)} = \widehat{\boldsymbol{Y}}$, and the iterates end when the entries of matrix $\boldsymbol{D}_{\text{te}} = \boldsymbol{Z}'\widetilde{\boldsymbol{Y}}_{\text{cs}}^{(j)\prime}$, containing all the temporal discrepancies, are small enough according to a suitable convergence criterion (Chapter 1). In the description above, *temporal-then-cross-sectional* reconciliation is iteratively performed (ite($rec_{\text{te}}, rec_{\text{cs}}$)), otherwise the order may be reversed, thus generating *cross-sectional-then-temporal* reconciliation (ite($rec_{\text{cs}}, rec_{\text{te}}$)).

2.3 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., $\widetilde{\boldsymbol{Y}}_{TSR} \neq \widetilde{\boldsymbol{Y}}_{TSR}$), and either cross-sectionally (TSR) or temporally (STR), but not cross-temporally, reconciled:

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

and

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

At this regard, we have found an interesting result, here shown as Theorem 2.1, 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 2.1. *Let* $\boldsymbol{W}_j^{[k]}$, $k \in \mathcal{K}$, $j = 1, \ldots, m/k$, *be the* $(n \times n)$ *cross-sectional hierarchy error covariance matrix, and* $\boldsymbol{\Omega}_i$, $i = 1, \ldots, n$, *the* $[(k^* + m) \times (k^* + m)]$ *error covariance matrix for the i-th series temporal hierarchy. If* $\boldsymbol{W}_j^{[k]} = \boldsymbol{W}$, $\forall k \in \mathcal{K}$ *and* $\forall j = 1, \ldots, m/k$, *and* $\boldsymbol{\Omega}_i = \boldsymbol{\Omega}$, $\forall i = 1, \ldots, n$, *then:*

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

$$\widetilde{\boldsymbol{Y}}_{tcs} = \widetilde{\boldsymbol{Y}}_{cst} = \widetilde{\boldsymbol{Y}}_{seq},$$

   where $\widetilde{\boldsymbol{Y}}_{tcs}$ *(*$\widetilde{\boldsymbol{Y}}_{cst}$*) is the* $[n \times (k^* + m)]$ *matrix of the temporal-then-cross-sectional (cross-sectional-then-temporal) reconciled forecasts;*
2. *Denoting* $\widetilde{\boldsymbol{Y}}_{oct}$ *the optimal (in least squares sense) combination cross-temporal reconciliation approach with cross-temporal covariance matrix* $\boldsymbol{\Omega}_{ct} = \boldsymbol{W} \otimes \boldsymbol{\Omega}$, *it is:*

$$\widetilde{\boldsymbol{Y}}_{seq} = \widetilde{\boldsymbol{Y}}_{oct}.$$

Proof. See Appendix A.1. □

It is worth noting that, when only constant error covariance matrices are involved in both unidimensional steps, the iterative cross-temporal reconciliation approach proposed in Chapter 1 reduces to the sequential procedure proposed by Yagli *et al.* (2019), and gives a unique result, $\widetilde{\boldsymbol{Y}}_{\text{seq}}$, which does not depend on the order of application of the reconciliation phases[8]. In addition, it does exist a simple simultaneous, optimal (in least-squares sense) reconciliation approach equivalent to the above sequential 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 A.1, after the proof of the theorem). All these results can be summarized as follows:

$$\begin{array}{l} \boldsymbol{W}^{[k]} = \boldsymbol{W}, \forall k \in \mathcal{K} \\ \boldsymbol{\Omega}_i = \boldsymbol{\Omega}, i = 1, \ldots, n \\ \boldsymbol{\Omega}_{ct} = \boldsymbol{W} \otimes \boldsymbol{\Omega} \end{array} \rightarrow \underbrace{\widetilde{\boldsymbol{Y}}_{\text{KA}(tcs)} = \widetilde{\boldsymbol{Y}}_{\text{KA}(cst)}}_{\substack{\text{Kourentzes and} \\ \text{Athanasopoulos (2019)}}} = \underbrace{\widetilde{\boldsymbol{Y}}_{\text{seq}}}_{\substack{\text{Yagli } \textit{et al.} \\ \text{(2019)}}} = \underbrace{\widetilde{\boldsymbol{Y}}_{\text{ite}(tcs)} = \widetilde{\boldsymbol{Y}}_{\text{ite}(cst)} = \widetilde{\boldsymbol{Y}}_{\text{oct}}}_{\text{Chapter 1}}$$

[8]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).

Using the covariance matrices in Table 2.1, we may thus establish some useful equivalences between the cross-temporal reconciliation approaches considered in Chapter 1, and identify new ones:

1. oct(ols) equivalent to seq/KA/ite(ols_{te}, ols_{cs}), and seq/KA/ite(ols_{cs}, ols_{te}), with $\boldsymbol{W} = \boldsymbol{I}_n$, $\boldsymbol{\Omega} = \boldsymbol{I}_{k^*+m}$ and $\boldsymbol{\Omega}_{\text{ct}} = \boldsymbol{I}_{n(k^*+m)}$.
2. oct($struc$) equivalent to seq/KA/ite($struc_{te}, struc_{cs}$), and seq/KA/ite($struc_{cs}, struc_{te}$), with $\boldsymbol{W} = \text{diag}\,(\boldsymbol{S}\mathbf{1}_{n_b})$, $\boldsymbol{\Omega} = \text{diag}\,(\boldsymbol{R}\mathbf{1}_m)$ and $\boldsymbol{\Omega}_{\text{ct}} = \text{diag}\,(\boldsymbol{F}\mathbf{1}_{mn_b})$.
3. oct($ols_{cs}, struc_{te}$) equivalent to seq/KA/ite($ols_{cs}, struc_{te}$), with $\boldsymbol{W} = \boldsymbol{I}_n$, $\boldsymbol{\Omega} = \text{diag}\,(\boldsymbol{R}\mathbf{1}_m)$ and $\boldsymbol{\Omega}_{ct} = \text{diag}\,[(\boldsymbol{I}_n \otimes \boldsymbol{R})\mathbf{1}_{mn_b}]$.
4. oct($struc_{cs}, ols_{te}$) equivalent to seq/KA/ite($struc_{cs}, ols_{te}$), with $\boldsymbol{W} = \text{diag}\,(\boldsymbol{S}\mathbf{1}_{n_b})$, $\boldsymbol{\Omega} = \boldsymbol{I}_{k^*+m}$ and $\boldsymbol{\Omega}_{ct} = \text{diag}\,[(\boldsymbol{S} \otimes \boldsymbol{I}_{k^*+m})\mathbf{1}_{mn_b}]$.

It is worth noting that all the procedures considered so far make use of very simple covariance matrices, not using the information coming from the in-sample forecast errors.

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($wlsv$)), 'converges' to the optimal combination reconciliation approach oct($wlsv$) (Chapter 1), regardless of unidimensional reconciliation steps' orders. In other terms,

$$\widetilde{\boldsymbol{Y}}_{\text{ite}(wlsv_{te},wls_{cs})} \simeq \widetilde{\boldsymbol{Y}}_{\text{ite}(wls_{cs},wlsv_{te})} \simeq \widetilde{\boldsymbol{Y}}_{\text{oct}(wlsv)},$$

where the quality of the approximation solely depends on the convergence criterion: the lower the tolerance value δ (see Section 2.2.4), the better the approximation will be. A visual, empirical support to this result is given by Figure 2.2, showing the Frobenius norm[9] of the difference between the matrices of the reconciled forecasts using ite($wlsv_{te}, wls_{cs}$) with decreasing tolerance values δ, and oct($wlsv$), in the 350 replications of the forecasting experiment described in Section 2.4. We generalise this result in Theorem 2.2.

Theorem 2.2. *Let $\boldsymbol{W}^{[k]}$, $k \in \mathcal{K}$, be the cross-sectional error covariance matrix with $\boldsymbol{W}_{ct} = diag(\boldsymbol{W}^{[m]}, \dots, \boldsymbol{W}^{[1]})$, and $\boldsymbol{\Omega}_i$, $i = 1, \dots, n$, the i-th series temporal error covariance matrix with $\boldsymbol{\Omega}_{ct} = diag(\boldsymbol{\Omega}_1, \dots, \boldsymbol{\Omega}_n)$. If $\boldsymbol{W}_{ct} = \boldsymbol{P}'\boldsymbol{\Omega}_{ct}\boldsymbol{P}$, then the iterative reconciled forecasts converge in norm to the optimal (in least squares sense) combination cross-temporal reconciled forecasts with cross-temporal covariance matrix $\boldsymbol{\Omega}_{ct}$.*

Proof. See Appendix A.2. □

[9]The Frobenius norm of a real-valued matrix $\boldsymbol{X}$ is defined as the square root of the sum of the squares of its elements (Golub and Van Loan, 1996): $||\boldsymbol{X}||_F = \sqrt{\sum_{i,j} x_{i,j}^2}$.

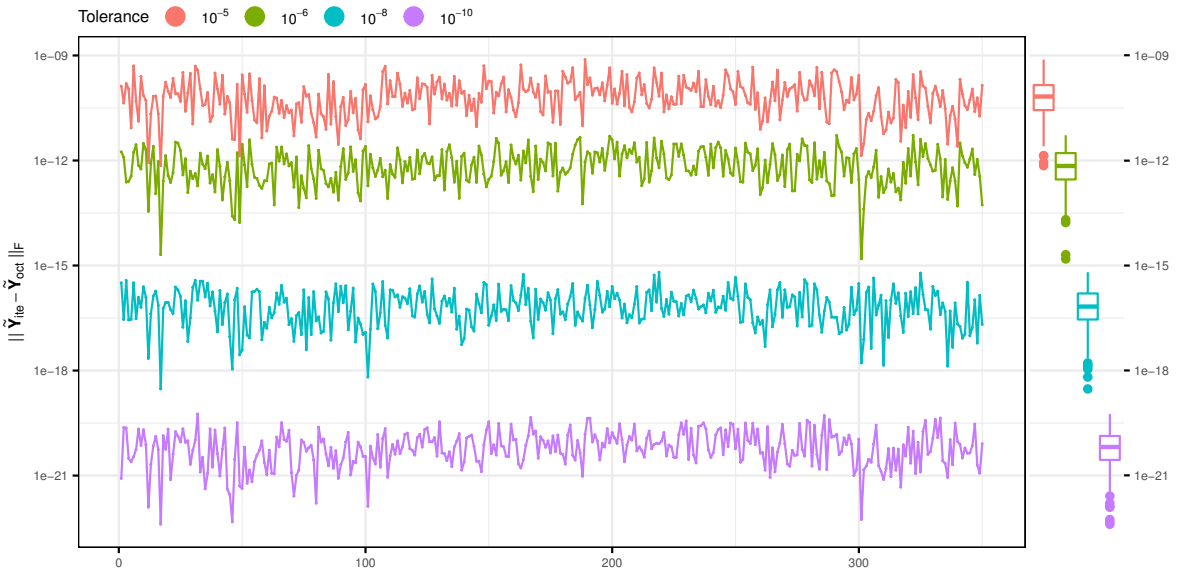


Figure 2.2: *Frobenius norm of the difference between the matrices of the reconciled forecasts using ite($wlsv_{te}, wls_{cs}$) and oct($wlsv$), with different tolerance value δ. 350 replications of the forecasting experiment described in Section 2.4.*

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).

2.4 Replication 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 2.3):

- $\mathcal{L}_0$: 1 time series for the Independent System Operator (ISO), given by the sum of the 318 plant series;
- $\mathcal{L}_1$: 5 time series for the Transmission Zones (TZ), each given by the sum of 27, 73, 101, 86, and 31 plants, respectively;
- $\mathcal{L}_2$: 318 bottom time series at plant level (P).

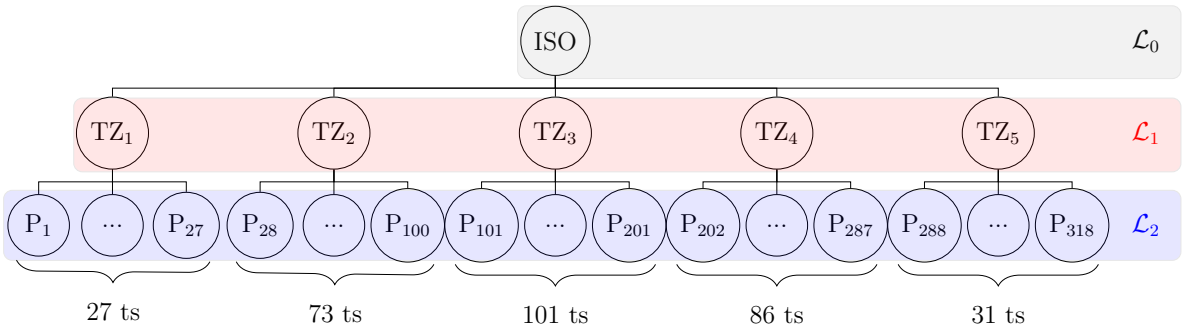


Figure 2.3: *PV324 hierarchy*

Following Yang *et al.* (2017b) and Yagli *et al.* (2019), we perform a forecasting experiment with 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 at plant level, numerical weather prediction (NWP) forecasts generated by 3TIER (3TIER, 2010) are used as base forecasts. All the remaining base forecasts, for the six $\mathcal{L}_0$ and $\mathcal{L}_1$ time series at any time granularity $k \in \mathcal{K}$, and for the $\mathcal{L}_2$ 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.*, 2023), not controlling for possible negative forecasts. Furthermore, following Yagli *et al.* (2019), day-ahead persistence is used as the reference model (PERS)[10]:

$$\widehat{y}^{[1]}_{T+h|T,\,PERS} = y^{[1]}_{T+h-48}. \tag{2.8}$$

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 cross-sectional hierarchy and for any temporal granularity are obtained through cross-temporal bottom-up of the 318 hourly bottom time series, that is:

$$\widetilde{\mathbf{y}}_{PERS_{bu}} = \mathbf{F}\widehat{\mathbf{b}}^{[1]}_{PERS} = \text{vec}\left(\widetilde{\mathbf{Y}}'_{PERS_{bu}}\right). \tag{2.9}$$

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

$$\widetilde{\mathbf{y}}_{3TIER_{bu}} = \mathbf{F}\ \widehat{\mathbf{b}}^{[1]}_{3TIER} = \text{vec}\left(\widetilde{\mathbf{Y}}'_{3TIER_{bu}}\right). \tag{2.10}$$

[10]On the choice of the standard reference forecasting method, see Yang (2019).

Approach	# rep	# series		values		# rep	# series		values	
	(350)	min	max	min	max	(350)	min	max	min	max
	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($ols_{\text{cs}}, struc_{\text{te}}$)	350	91	324	-17.961	-0.000	6	0	28	-8.915	-0.008
oct($struc_{\text{cs}}, ols_{\text{te}}$)	350	137	324	-20.739	-0.000	11	0	33	-10.363	-0.000
oct($struc$)	350	89	324	-14.292	-0.000	4	0	11	-1.020	-0.026
te(ols_2)+cs(ols)*	350	42	324	-12.993	-0.000	10	0	72	-15.037	-0.018
te($struc_2$)+cs(ols)*	350	36	324	-12.994	-0.000	10	0	71	-14.791	-0.005
te(ols_2)+cs($struc$)*	350	144	324	-7.526	-0.000	10	0	43	-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 cross-sectional and temporal incoherent forecasts.
* The approach produces temporal incoherent forecasts.

Table 2.2: *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 out of 324 (# 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). Hourly and daily forecasts, forecast horizon: operating day.*

According to Theorem 2.1, the STR and TSR sequential reconciliations proposed by Yagli *et al.* (2019) reduce to the following approaches: oct(ols), oct($ols_{\text{cs}}, struc_{\text{te}}$), oct($struc_{\text{cs}}, ols_{\text{te}}$), oct($struc$). In addition, Yagli *et al.* (2019) consider other four Temporal-then-Spatial-Reconciliation approaches, called $\text{TSR}_{\mathcal{L}_2}$, 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, the theorem so far no longer holds, so 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_2)+cs(ols), te($struc_2$)+cs(ols),te(ols_2)+cs($struc$), te($struc_2$)+cs($struc$), respectively.

2.4.1 Non-negativity and aggregation consistency issues

Standard forecast reconciliation, both in space and/or in time, may produce negative revised forecasts (Yang *et al.*, 2017a,b; Yagli *et al.*, 2019), unless specific non-negative reconciliation approaches are applied (Wickramasuriya *et al.*, 2020; Girolimetto and Di Fonzo, 2023a). Thus it is not surprising that, with the exception of PERS_{bu} and 3TIER_{bu}, the other approaches considered by Yagli *et al.* (2019) produce some negative reconciled forecasts. Details on this issue are shown in Table 2.2.

Negative hourly forecasts are obtained in all 350 replications of the forecasting experiment, and there are some cases (in a range within 4 and 11 replications out of 350)

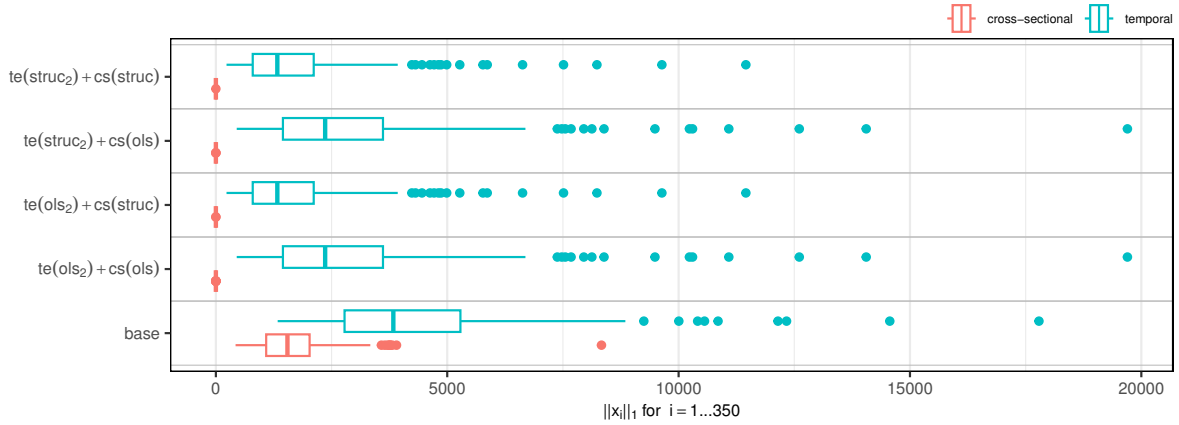


Figure 2.4: *Base forecasts and sequential TSR$_{\mathcal{L}_2}$ (Yagli* et al.*, 2019) reconciled forecasts. Boxplots of the distribution in the 350 replications of the forecasting experiment of the cross-sectional (in red) and temporal (in blue) gross discrepancies, as defined in (2.11). For cross-temporal reconciled forecasts, both discrepancies are expected to be zero.*

where negative daily forecasts are produced as well[11]. Furthermore, the base forecasts are incoherent both in space and time, and the sequential TSR$_{\mathcal{L}_2}$ approaches proposed by Yagli *et al.* (2019) are temporally incoherent. This can be visually appreciated from Figure 2.4, showing the boxplots from the distribution of the cross-sectional and temporal gross discrepancies registered in the 350 replications of the forecasting experiment, computed as (Chapter 1):

$$\begin{aligned} \text{Cross-sectional gross discrepancy:} \quad & d_{\text{cs}} = ||\boldsymbol{U}'\widehat{\boldsymbol{Y}}||_1 \\ \text{Temporal gross discrepancy:} \quad & d_{\text{te}} = ||\boldsymbol{Z}'\widehat{\boldsymbol{Y}}'||_1 \end{aligned}, \tag{2.11}$$

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

In this case, forecast reconciliation may thus generate physically unreasonable values. Furthermore, if coherency is desired, 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 2.5 we propose a simple operational strategy, able to generate fully reconciled non-negative forecasts.

[11]The relatively low number of negative hourly base forecasts is explained by 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 online appendix.

2.4.2 Forecast evaluation

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

$$\text{nRMSE}_{i,j}^{[k]} = \frac{\sqrt{\frac{1}{L}\sum_{l=1}^{L}\left(\widetilde{y}_{i,j,l}^{[k]} - y_{i,l}^{[k]}\right)^2}}{\frac{1}{L}\sum_{l=1}^{L} y_{i,l}^{[k]}}, \quad \text{nMBE}_{i,j}^{[k]} = \frac{\frac{1}{L}\sum_{l=1}^{L}\left(\widetilde{y}_{i,j,l}^{[k]} - y_{i,l}^{[k]}\right)}{\frac{1}{L}\sum_{l=1}^{L} y_{i,l}^{[k]}}, \tag{2.12}$$

$$\text{and} \qquad \text{FS}_{i,j}^{[k]} = 1 - \frac{\text{nRMSE}_{i,j}^{[k]}}{\text{nRMSE}_{i,0}^{[k]}},$$

where $i = 1, \ldots, n$, denotes the series, $k \in \mathcal{K}$, $j = 0, \ldots, J$, 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. Whereas nRMSE penalizes large errors, nMBE reveals over- and under-prediction through the sign of the metric (Yagli *et al.*, 2019, p. 394). Forecast Skill can be either negative (approach is worse than the reference model) or positive (approach is better than the reference model).

It should be noted that for cross-temporal coherent forecasts the index $\text{nMBE}_{i,j}^{[k]}$ does not depend on the temporal aggregation order $k \in \mathcal{K}$ (see Appendix A.3). This means that, if j denotes a fully coherent cross-temporal forecast reconciliation approach, the following identities hold:

$$\text{nMBE}_{i,j}^{[m]} = \text{nMBE}_{i,j}^{[k_{p-1}]} = \ldots = \text{nMBE}_{i,j}^{[k_2]} = \text{nMBE}_{i,j}^{[1]}, \quad i = 1, \ldots, n. \tag{2.13}$$

From Table 2.3 it appears that on average all the considered forecasting approaches improve on the benchmark PERS_{bu}, in a range between 9.2% (oct(ols) for the 5 $\mathcal{L}_1$ series' hourly forecasts) and 38.3% (3TIER_{bu} for the $\mathcal{L}_0$ total series' daily forecasts)[12].

The forecasting accuracy indices for each Transmission Zone forecasts are reported in Table 2.4. We observe that:

- 3TIER_{bu} is the best for the total series ($\mathcal{L}_0$) at any temporal granularity, the second best being te($struc_2$)+cs($struc$).
- The approach te($struc_2$)+cs($struc$) ranks first for 4 out of the 5 hourly upper series at level $\mathcal{L}_1$, whereas at daily level 3TIER_{bu} 'wins' again. The performance of oct($struc$) appears very close to that of te($struc_2$)+cs($struc$).

[12]The results for all temporal aggregation orders, $k \in \{24, 12, 8, 6, 4, 3, 2, 1\}$, are available in the online appendix.

Approach	$\mathcal{L}_0$		$\mathcal{L}_1$		$\mathcal{L}_2$	
	H	D	H	D	H	D
	nRMSE(%)					
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
base^{+}	27.85	18.17	34.24	20.94	53.46	25.82
oct(ols)	30.69	17.91	39.17	21.74	51.54	26.73
oct($ols_{\text{cs}}, struc_{\text{te}}$)	28.26	16.96	35.19	20.61	48.20	25.46
oct($struc_{\text{cs}}, ols_{\text{te}}$)	28.74	17.33	35.48	20.82	49.32	26.14
oct($struc$)	26.71	*16.24*	*33.11*	*19.64*	*46.74*	**24.73**
te(ols_2)+cs(ols)*	27.80	17.96	34.36	21.80	48.52	26.71
te($struc_2$)+cs(ols)*	27.80	17.95	34.34	21.79	47.77	26.36
te(ols_2)+cs($struc$)*	27.02	17.24	33.40	20.62	47.76	25.98
te($struc_2$)+cs($struc$)*	*26.17*	16.68	**32.44**	19.96	**46.25**	*25.03*
	nMBE(%)					
PERS_{bu}	**0.075**	**0.075**	*0.058*	**0.058**	**0.075**	**0.075**
3TIER_{bu}	-4.213	-4.213	-3.745	-3.745	-4.362	-4.362
base^{+}	-0.394	0.725	-0.131	1.000	-4.362	0.902
oct(ols)	0.593	0.593	0.557	0.557	0.590	0.590
oct($ols_{\text{cs}}, struc_{\text{te}}$)	0.359	*0.359*	0.339	*0.339*	0.354	*0.354*
oct($struc_{\text{cs}}, ols_{\text{te}}$)	0.735	0.735	0.744	0.744	0.733	0.733
oct($struc$)	0.385	0.385	0.426	0.426	0.374	0.374
te(ols_2)+cs(ols)*	-0.353	0.769	-0.369	0.718	-0.352	0.754
te($struc_2$)+cs(ols)*	-0.355	0.768	-0.369	0.718	-0.360	0.747
te(ols_2)+cs($struc$)*	0.100	0.858	0.119	0.863	*0.099*	0.850
te($struc_2$)+cs($struc$)*	*-0.093*	0.666	**-0.048**	0.695	-0.104	0.647
	forecast skill					
3TIER_{bu}	**0.248**	**0.383**	0.213	**0.318**	0.105	0.178
base^{+}	0.196	0.102	0.206	0.148	0.105	0.158
oct(ols)	0.113	0.115	0.092	0.115	0.137	0.128
oct($ols_{\text{cs}}, struc_{\text{te}}$)	0.184	0.162	0.184	0.161	0.193	0.169
oct($struc_{\text{cs}}, ols_{\text{te}}$)	0.170	0.143	0.178	0.152	0.175	0.147
oct($struc$)	0.228	*0.197*	*0.233*	*0.201*	*0.218*	**0.193**
te(ols_2)+cs(ols)*	0.197	0.112	0.204	0.113	0.188	0.129
te($struc_2$)+cs(ols)*	0.197	0.113	0.204	0.113	0.201	0.140
te(ols_2)+cs($struc$)*	0.219	0.148	0.226	0.161	0.201	0.152
te($struc_2$)+cs($struc$)*	*0.244*	0.176	**0.248**	0.188	**0.226**	*0.183*

$^{+}$ The approach produces cross-sectional and temporal incoherent forecasts.
* The approach produces temporal incoherent forecasts.

Table 2.3: *Forecast accuracies in terms of nRMSE(%), nMBE(%), and forecast skill over the PERS_{bu} benchmark of base forecasts and sequential reconciliation approaches. Unconstrained 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 and italic entries identify the best and the second best performing approaches, respectively.*

Approach	TZ_1		TZ_2		TZ_3		TZ_4		TZ_5	
	H	D	H	D	H	D	H	D	H	D
	nRMSE(%)									
$PERS_{bu}$	28.72	16.12	40.27	22.40	46.48	26.96	52.82	29.75	47.44	27.62
$3TIER_{bu}$	22.81	**10.43**	33.34	**16.72**	*34.01*	**17.06**	46.40	**22.80**	**33.16**	**16.75**
base^{+}	22.41	13.55	32.05	18.70	35.14	21.95	44.94	26.89	36.67	23.58
oct(ols)	29.27	14.89	34.58	19.25	39.55	22.91	47.30	26.50	45.16	25.15
oct($ols_{\mathrm{cs}}, struc_{\mathrm{te}}$)	24.26	13.93	32.32	18.47	36.69	21.78	43.89	25.30	38.81	23.59
oct($struc_{\mathrm{cs}}, ols_{\mathrm{te}}$)	23.44	13.50	32.89	18.83	37.42	22.33	44.76	25.97	38.90	23.49
oct($struc$)	*22.00*	*12.81*	*30.92*	*17.82*	34.79	*20.94*	*42.02*	*24.50*	35.83	*22.12*
te(ols_2)+cs(ols)*	22.95	15.46	32.06	19.23	35.12	22.38	44.93	26.91	36.73	25.02
te($struc_2$)+cs(ols)*	22.93	15.46	32.05	19.22	35.11	22.37	44.92	26.90	36.70	25.00
te(ols_2)+cs($struc$)*	22.11	13.36	31.40	18.70	34.70	21.93	42.97	26.06	35.82	23.06
te($struc_2$)+cs($struc$)*	**21.58**	13.00	**30.48**	18.09	**33.58**	21.18	**41.86**	25.21	*34.69*	22.32
	nMBE(%)									
$PERS_{bu}$	**-0.013**	**-0.013**	*0.099*	*0.099*	*0.127*	**0.127**	**0.030**	0.030	**0.050**	**0.050**
$3TIER_{bu}$	-1.967	-1.967	-3.215	-3.215	-3.636	-3.636	-7.364	-7.364	-2.542	-2.542
base^{+}	-0.101	0.889	-0.419	0.842	-0.134	1.613	-0.119	0.159	*0.118*	1.499
oct(ols)	0.250	0.250	0.307	0.307	1.006	1.006	0.345	0.345	0.876	0.876
oct($ols_{\mathrm{cs}}, struc_{\mathrm{te}}$)	0.078	*0.078*	**0.091**	**0.091**	0.577	0.577	0.326	0.326	0.625	*0.625*
oct($struc_{\mathrm{cs}}, ols_{\mathrm{te}}$)	0.660	0.660	0.478	0.478	0.968	0.968	0.591	0.591	1.026	1.026
oct($struc$)	0.442	0.442	0.203	0.203	0.541	*0.541*	0.185	0.185	0.759	0.759
te(ols_2)+cs(ols)*	-0.492	0.428	-0.577	0.653	-0.248	1.471	-0.268	**-0.007**	-0.258	1.047
te($struc_2$)+cs(ols)*	-0.490	0.429	-0.578	0.653	-0.249	1.470	-0.273	*-0.012*	-0.256	1.049
te(ols_2)+cs($struc$)*	0.135	0.794	-0.123	0.685	0.182	1.233	*0.114*	0.446	0.289	1.157
te($struc_2$)+cs($struc$)*	*0.047*	0.706	-0.249	0.559	**0.009**	1.060	-0.242	0.090	0.194	1.063
	forecast skill									
$3TIER_{bu}$	0.206	**0.353**	0.172	**0.254**	*0.268*	**0.367**	0.121	**0.234**	**0.301**	**0.394**
base^{+}	0.220	0.159	0.204	0.165	0.244	0.186	0.149	0.096	0.227	0.146
oct(ols)	-0.019	0.076	0.141	0.141	0.149	0.150	0.104	0.109	0.048	0.089
oct($ols_{\mathrm{cs}}, struc_{\mathrm{te}}$)	0.155	0.136	0.197	0.176	0.211	0.192	0.169	0.150	0.182	0.146
oct($struc_{\mathrm{cs}}, ols_{\mathrm{te}}$)	0.184	0.163	0.183	0.159	0.195	0.172	0.153	0.127	0.180	0.149
oct($struc$)	*0.234*	*0.205*	*0.232*	*0.205*	0.251	*0.223*	*0.204*	*0.177*	0.245	*0.199*
te(ols_2)+cs(ols)*	0.201	0.041	0.204	0.141	0.244	0.170	0.149	0.095	0.226	0.094
te($struc_2$)+cs(ols)*	0.202	0.041	0.204	0.142	0.245	0.170	0.150	0.096	0.227	0.095
te(ols_2)+cs($struc$)*	0.230	0.171	0.220	0.165	0.253	0.187	0.187	0.124	0.245	0.165
te($struc_2$)+cs($struc$)*	**0.249**	0.194	**0.243**	0.192	**0.277**	0.215	**0.208**	0.153	*0.269*	0.192

$^{+}$ The approach produces cross-sectional and temporal incoherent forecasts.

* The approach produces temporal incoherent forecasts.

Table 2.4: *Forecast accuracies in terms of nRMSE(%), nMBE(%), and forecast skill over the $PERS_{bu}$ benchmark of base forecasts and sequential reconciliation approaches for the series at $\mathcal{L}_1$ level (Transmission Zones). Unconstrained 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 and italic entries identify the best and the second best performing approaches, respectively.*

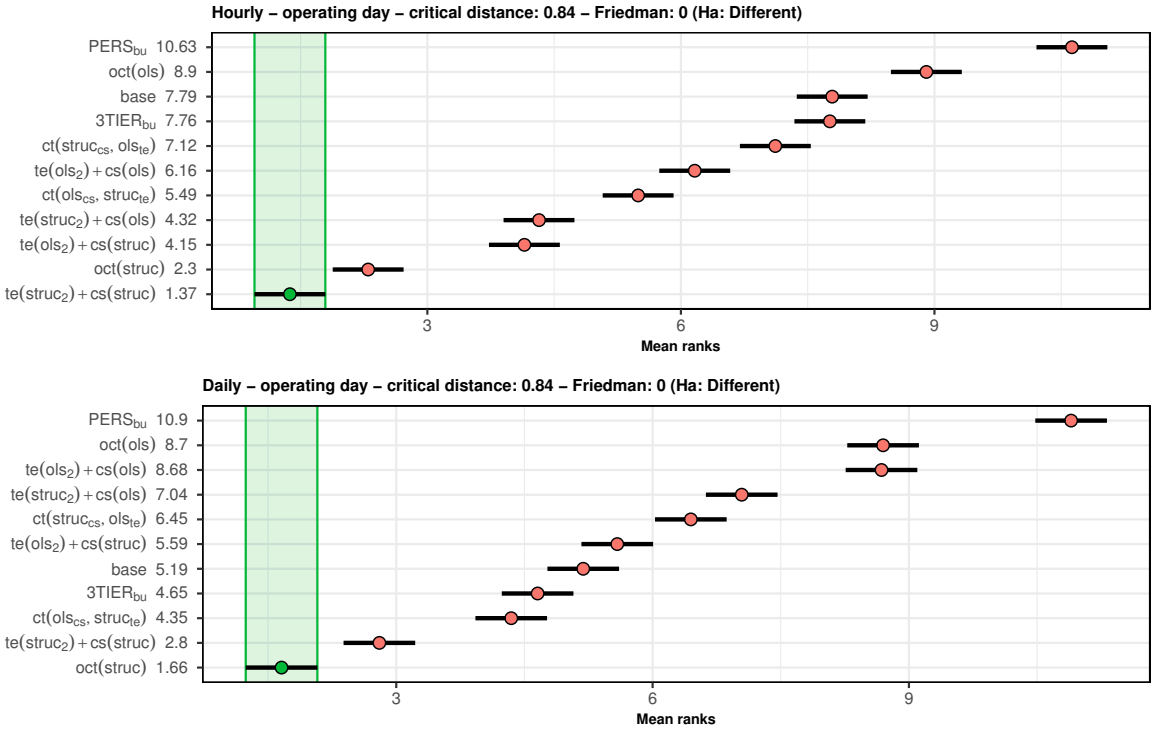


Figure 2.5: *MCB Nemenyi test results: average ranks and 95% confidence intervals. The unconstrained reconciliation approaches considered by Yagli* et al. *(2019) are sorted vertically according to the nRMSE(%) mean rank. Hourly (top panel) and Daily (bottom panel) forecasts for $\mathcal{L}_0, \mathcal{L}_1, \mathcal{L}_2$ 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.*

- te($struc_2$)+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 $k \geq 3$). It should be noted that, unlike te($struc_2$)+cs($struc$), oct($struc$) forecasts are cross-temporally coherent.
- Unlike te($struc_2$)+cs($struc$), 3TIER$_{bu}$ forecasts are always non-negative, and coherent both in space and time at any granularity.
- For these reasons, 3TIER$_{bu}$ should be considered as a challenging competitor in the evaluation of the new proposed procedures (Section 2.5).

To give a complete picture of the evaluation results for hourly and daily forecasts, in Figure 2.5 the Multiple Comparison with the Best (MCB) Nemenyi tests are shown (Koning *et al.*, 2005; Kourentzes and Athanasopoulos, 2019; Makridakis *et al.*, 2022). 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($struc_2$)+cs($struc$) at the second place. This result is reversed at hourly level: te($struc_2$)+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($struc_2$)+cs($struc$) are not temporally coherent, which means that the sum of the hourly forecasts does not match with the daily forecast.

Limiting ourselves to consider fully coherent reconciled forecasts, the scatter plots of the 324 couples of nRMSE(%) for oct($struc$) vs., respectively, PERS_{bu} and 3TIER_{bu} (Figure 2.6), show that the most performing regression-based cross-temporal reconciliation approach improves uniformly on the benchmark, and in the majority of cases on 3TIER_{bu}, particularly at hourly level ($k = 1$), where 89% of the variables are observed to show an improved 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.

2.5 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 cross-sectionally coherent. For this reason, among the approaches proposed by Yagli *et al.* (2019), we consider the reference benchmark PERS_{bu}, the NWP base forecasts 3TIER_{bu}, and oct($struc$). These approaches are then compared with the following 7 cross-temporal forecasting procedures:

- KA($wlsv_{\text{te}}, wls_{\text{cs}}$): the heuristic approach by Kourentzes and Athanasopoulos (2019), using te($wlsv$) in the first step, and cs(wls) in the second, respectively;
- ct($wlsv_{\text{te}}, bu_{\text{cs}}$), ct($struc_{\text{te}}, bu_{\text{cs}}$), ct($wls_{\text{cs}}, bu_{\text{te}}$), and ct($struc_{\text{cs}}, bu_{\text{te}}$): partly bottom-up cross-temporal reconciliation (see Section 2.2.2) according to, respectively, te($wlsv$), te($struc$), cs(wls), and cs($struc$);
- oct($wlsv$) and oct($bdshr$): optimal (in least squares sense) cross-temporal forecast reconciliation approaches using the in-sample forecast errors (Chapter 1).

2.5.1 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 (2.4), is not guaranteed to always produce non-negative reconciled forecasts. This fact may be an issue for the analyst, since in many practical situations negative forecasts could have no meaning, thus

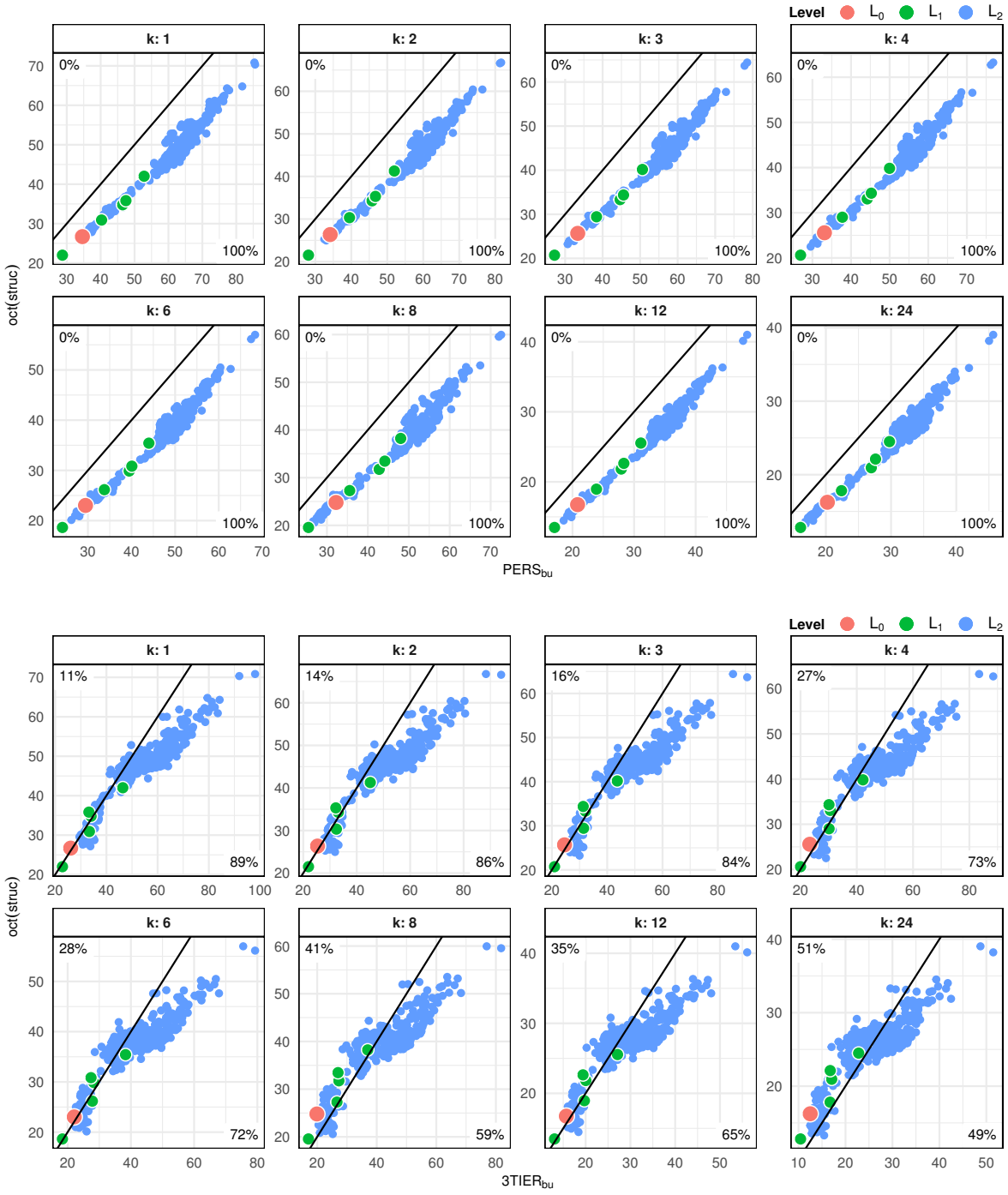


Figure 2.6: *Comparison of nRMSE(%) between $PERS_{bu}$ and oct(struc) (top panel), and between $3TIER_{bu}$ and oct(struc) (bottom panel). The black line represents the bisector, where the nRMSE's for both approaches are equal. On the top-left (bottom-right) corner of each graph, the percentage of points above (below) the bisector is reported.*

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[13]. More precisely, possible negative values of the unconstrained reconciled high-frequency bottom time series forecasts are set to zero. Denote $\widetilde{\boldsymbol{B}}_0^{[1]}$ 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 $\widetilde{\boldsymbol{b}}_0^{[1]} = \text{vec}\left(\widetilde{\boldsymbol{B}}_0^{[1]'}\right)$, that is, according to expression (2.2):

$$\widetilde{\boldsymbol{y}}_0 = \boldsymbol{F}\widetilde{\boldsymbol{b}}_0^{[1]}. \tag{2.14}$$

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 state-of-the-art numerical optimization procedure like `osqp` (Stellato *et al.*, 2020) implemented in `FoReco` (Girolimetto and Di Fonzo, 2023a), could be negligible.

For example, Figure 2.7 shows the graphs of one day of hourly forecasts computed by unconstrained oct(*struc*), oct(*struc*)$_{\text{osqp}}$ and oct(*struc*)$_{\text{sntz}}$, respectively, for the aggregated series (ISO and five Transmission Zones), and for two bottom variables at plant level (P_{295} and P_{315}) with a large number of negative reconciled forecasts. For each variable we consider the day with the highest number of negative forecasts. It appears that the physically feasible forecasts produced by the non-negative reconciliation approaches are about the same at the aggregate levels, whereas some difference is visible at Plant level.

Table 2.5 shows the indices nRMSE(%) of the reconciled forecasts produced by the oct(*struc*) approach according to unconstrained and non-negative (both sntz and osqp) variants. It is worth noting that the sntz heuristic always gives the lowest nRMSE(%), independently of the temporal granularity of the forecasts[14]. This result is visually

[13]Recent contributions on this topic in the hierarchical forecasting field are Wickramasuriya *et al.* (2020), Girolimetto and Di Fonzo (2023a).

[14]Similar results were found for all the considered reconciliation approaches.

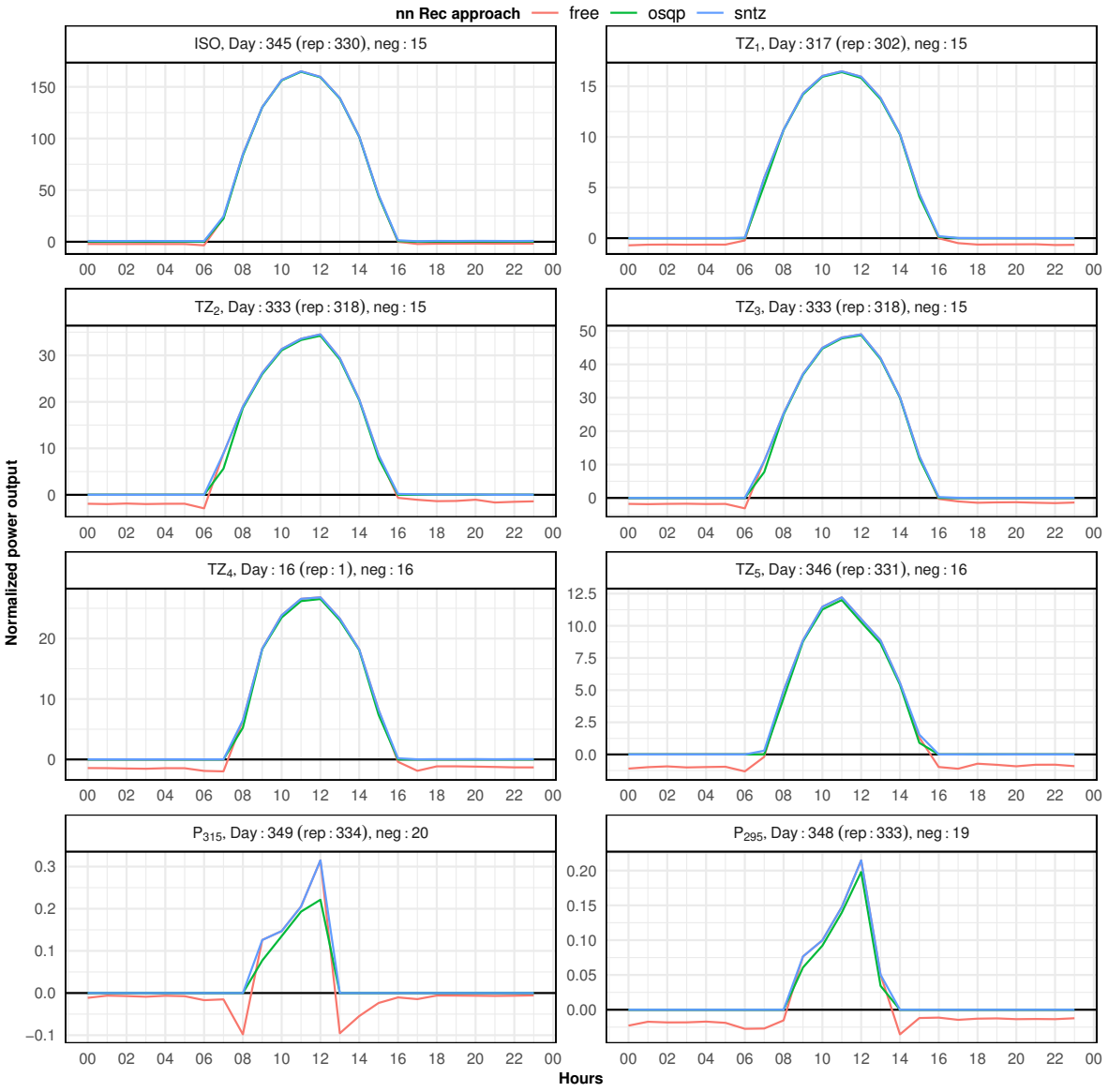


Figure 2.7: *One day of hourly reconciled forecasts for two of the 318 bottom variables (Plants P_{295} and P_{315}, component of TZ_5), and for the six upper time series (5 Transmission Zones and the Total ISO). For each series, it is shown the day with the highest number of negative forecasts produced by the reconciliation approach oct(struc) (in red). The non-negative forecasts are obtained by $oct(struc)_{osqp}$ (in green) and $oct(struc)_{sntz}$ (in blue).*

Level (series)	Non-neg. reconciliation	nRMSE(%) 1	2	3	4	6	8	12	24	nMBE(%) 1-24
$\mathcal{L}_0$ (1)	free	26.71	26.35	25.67	25.57	23.03	24.80	16.74	16.24	**0.385**
	sntz	**26.64**	**26.27**	**25.56**	**25.48**	**22.86**	**24.71**	**16.25**	**15.73**	-1.079
	osqp	26.74	26.36	25.67	25.55	23.01	24.77	16.36	15.85	-0.398
$\mathcal{L}_1$ (5)	free	33.11	32.54	31.60	31.37	28.19	30.06	20.49	19.64	0.426
	sntz	**32.99**	**32.41**	**31.45**	**31.23**	**27.96**	**29.91**	**19.88**	**19.00**	-0.973
	osqp	33.11	32.52	31.58	31.32	28.15	29.99	20.01	19.14	**-0.311**
$\mathcal{L}_2$ (318)	free	46.74	44.02	42.05	41.07	36.67	38.16	26.56	24.73	**0.374**
	sntz	**46.51**	**43.80**	**41.80**	**40.83**	**36.34**	**37.90**	**25.85**	**23.97**	-1.109
	osqp	46.63	43.93	41.94	40.93	36.53	37.99	25.97	24.10	-0.422

Table 2.5: *Forecast accuracy in terms of nRMSE(%) and nMBE(%) of unconstrained 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.*

confirmed by the graphs in Figure 2.8, showing the scatter plots of the 324 couples of nRMSE(%) for oct$(struc)_{\text{osqp}}$ *vs.* oct$(struc)_{\text{sntz}}$. For this dataset, oct$(struc)_{\text{sntz}}$ beats oct$(struc)_{\text{osqp}}$ in no less than 94% of the 324 series for any temporal granularity. However, looking at indices nMBE(%) in Table 2.5, it emerges also that the bias of the oct$(struc)_{\text{sntz}}$ forecasts is more pronounced than oct$(struc)_{\text{osqp}}$. This seems to be a price to pay for using such a simple heuristic.

2.5.2 Forecast accuracy of the selected approaches

Table 2.6 shows the indices nRMSE(%) and nMBE(%) 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 new 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$(wls_{\text{cs}}, bu_{\text{te}})$ and ct$(struc_{\text{cs}}, bu_{\text{te}})$) show a good performance for the six series at levels $\mathcal{L}_0$ and $\mathcal{L}_1$. For the 318 disaggregated series at the Plant level, ct$(wlsv_{\text{te}}, bu_{\text{cs}})$ and oct$(wlsv)$ rank first and second, respectively, and their accuracy indices are very close each other, whereas the nMBE(%) of oct$(wlsv)$ is lower than ct$(wlsv_{\text{te}}, bu_{\text{cs}})$. In addition, all the considered non-negative forecast reconciliation approaches, even using the sntz heuristic (see Section 2.5.1), have smaller bias than the NWP forecasts 3TIER.

From Table 2.7 it emerges that even for distinct $\mathcal{L}_1$ series (Transmission Zones), the new approaches perform better than all the approaches considered by Yagli *et al.* (2019),

Approach	$\mathcal{L}_0$		$\mathcal{L}_1$		$\mathcal{L}_2$	
	H	D	H	D	H	D
			nRMSE(%)			
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($struc$)	26.64	15.73	32.99	19.00	46.51	23.97
KA($wlsv_{\text{te}}, wls_{\text{cs}}$)	23.59	13.55	30.23	16.91	44.48	*22.21*
ct($struc_{\text{cs}}, bu_{\text{te}}$)	*21.99*	*12.38*	**28.80**	*15.82*	49.76	24.33
ct($wls_{\text{cs}}, bu_{\text{te}}$)	**21.23**	**10.82**	*28.97*	**15.01**	49.88	23.66
ct($struc_{\text{te}}, bu_{\text{cs}}$)	24.81	14.43	31.41	17.80	45.30	22.93
ct($wlsv_{\text{te}}, bu_{\text{cs}}$)	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
			nMBE(%)			
PERS$_{bu}$	**0.075**		**0.058**		**0.075**	
3TIER$_{bu}$	-4.213		-3.745		-4.362	
oct($struc$)	-1.079		-0.973		-1.109	
KA($wlsv_{\text{te}}, wls_{\text{cs}}$)	*-0.691*		-0.823		-1.399	
ct($struc_{\text{cs}}, bu_{\text{te}}$)	-1.834		-1.593		-1.935	
ct($wls_{\text{cs}}, bu_{\text{te}}$)	-2.810		-2.418		-2.910	
ct($struc_{\text{te}}, bu_{\text{cs}}$)	-1.435		-1.308		-1.473	
ct($wlsv_{\text{te}}, bu_{\text{cs}}$)	-1.429		-1.320		-1.454	
oct($wlsv$)	-1.312		-1.195		-1.335	
oct($bdshr$)	-0.734		*-0.734*		*-0.728*	
			forecast skill			
3TIER$_{bu}$	0.248	0.383	0.213	0.318	0.105	0.178
oct($struc$)	0.231	0.222	0.235	0.227	0.222	0.218
KA($wlsv_{\text{te}}, wls_{\text{cs}}$)	0.319	0.330	0.299	0.312	0.256	*0.275*
ct($struc_{\text{cs}}, bu_{\text{te}}$)	*0.365*	*0.388*	**0.333**	*0.356*	0.167	0.206
ct($wls_{\text{cs}}, bu_{\text{te}}$)	**0.387**	**0.465**	*0.329*	**0.389**	0.165	0.228
ct($struc_{\text{te}}, bu_{\text{cs}}$)	0.283	0.287	0.272	0.276	0.242	0.252
ct($wlsv_{\text{te}}, bu_{\text{cs}}$)	0.321	0.337	0.302	0.317	**0.257**	**0.278**
oct($wlsv$)	0.317	0.326	0.300	0.308	*0.256*	0.273
oct($bdshr$)	0.303	0.312	0.294	0.303	0.243	0.265

Table 2.6: *Forecast accuracy of selected non-negative cross-temporal reconciliation approaches and base forecasts in terms of nRMSE(%), nMBE(%), and forecast skill over the PERS$_{bu}$ benchmark. Hourly and daily forecasts, forecast horizon: operating day. Bold entries and italic entries identify the best and the second best performing approaches, respectively.*

Approach	TZ_1		TZ_2		TZ_3		TZ_4		TZ_5	
	H	D	H	D	H	D	H	D	H	D
					nRMSE(%)					
$PERS_{bu}$	28.72	16.12	40.27	22.40	46.48	26.96	52.82	29.75	47.44	27.62
$3TIER_{bu}$	22.81	*10.43*	33.34	16.72	34.01	17.06	46.40	22.80	33.16	*16.75*
oct($struc$)	21.92	12.41	30.84	17.30	34.67	20.21	41.84	23.84	35.66	21.23
KA($wlsv_{\text{te}}, wls_{\text{cs}}$)	20.23	11.10	28.57	15.43	30.98	17.45	39.72	22.09	31.64	18.51
ct($struc_{\text{cs}}, bu_{\text{te}}$)	*20.15*	10.92	**27.01**	*14.37*	**28.74**	*15.74*	**38.12**	*20.89*	*29.99*	17.17
ct($wls_{\text{cs}}, bu_{\text{te}}$)	**19.33**	**9.72**	*28.11*	**14.18**	*29.55*	**15.10**	*39.20*	**20.17**	**28.68**	**15.88**
ct($struc_{\text{te}}, bu_{\text{cs}}$)	21.04	11.63	*29.55*	16.24	32.57	18.73	40.51	22.67	33.37	19.75
ct($wlsv_{\text{te}}, bu_{\text{cs}}$)	20.20	10.98	28.53	15.35	30.89	17.37	39.69	22.01	31.32	18.19
oct($wlsv$)	20.18	11.18	28.53	15.50	30.97	17.53	39.56	22.16	31.84	18.69
oct($bdshr$)	21.00	11.74	29.25	16.02	31.63	17.85	39.73	22.27	30.65	17.74
					nMBE(%)					
$PERS_{bu}$	**-0.013**		**-0.099**		**0.127**		**0.030**		**0.050**	
$3TIER_{bu}$	-1.967		-3.215		-3.636		-7.364		-2.542	
oct($struc$)	*-0.422*		-1.082		-0.864		-1.707		-0.790	
KA($wlsv_{\text{te}}, wls_{\text{cs}}$)	-0.575		-0.959		-0.814		-1.348		*-0.417*	
ct($struc_{\text{cs}}, bu_{\text{te}}$)	-0.702		-1.434		-1.514		-3.390		-0.926	
ct($wls_{\text{cs}}, bu_{\text{te}}$)	-1.437		-2.237		-2.530		-4.937		-0.949	
ct($struc_{\text{te}}, bu_{\text{cs}}$)	-0.620		-1.300		-1.224		-2.243		-1.153	
ct($wlsv_{\text{te}}, bu_{\text{cs}}$)	-0.890		-1.417		-1.246		-2.039		-1.011	
oct($wlsv$)	-0.768		-1.330		-1.169		-1.865		-0.843	
oct($bdshr$)	-0.652		*-0.727*		*-0.676*		*-0.828*		-0.778	
					forecast skill					
$3TIER_{bu}$	0.206	*0.353*	0.172	0.254	0.268	0.367	0.121	0.234	0.301	*0.394*
oct($struc$)	0.237	0.230	0.234	0.228	0.254	0.250	0.208	0.199	0.248	0.231
KA($wlsv_{\text{te}}, wls_{\text{cs}}$)	0.296	0.312	0.291	0.311	0.334	0.353	0.248	0.258	0.333	0.330
ct($struc_{\text{cs}}, bu_{\text{te}}$)	*0.299*	0.323	**0.329**	*0.359*	**0.382**	*0.416*	**0.278**	*0.298*	*0.368*	0.378
ct($wls_{\text{cs}}, bu_{\text{te}}$)	**0.327**	**0.397**	*0.302*	**0.367**	*0.364*	**0.440**	*0.258*	**0.322**	**0.396**	**0.425**
ct($struc_{\text{te}}, bu_{\text{cs}}$)	0.268	0.279	0.266	0.275	0.299	0.305	0.233	0.238	0.297	0.285
ct($wlsv_{\text{te}}, bu_{\text{cs}}$)	0.297	0.319	0.292	0.315	0.335	0.356	0.249	0.260	0.340	0.341
oct($wlsv$)	0.297	0.307	0.292	0.308	0.334	0.350	0.251	0.255	0.329	0.323
oct($bdshr$)	0.269	0.272	0.274	0.285	0.320	0.338	0.248	0.251	0.354	0.358

Table 2.7: *Forecast accuracy of selected non-negative cross-temporal reconciliation approaches and base forecasts in terms of nRMSE(%), nMBE(%), and forecast skill over the $PERS_{bu}$ benchmark, for the series at $\mathcal{L}_1$ level (Transmission Zones). Hourly and daily forecasts, forecast horizon: operating day. Bold entries and italic entries identify the best and the second best performing approaches, respectively.*

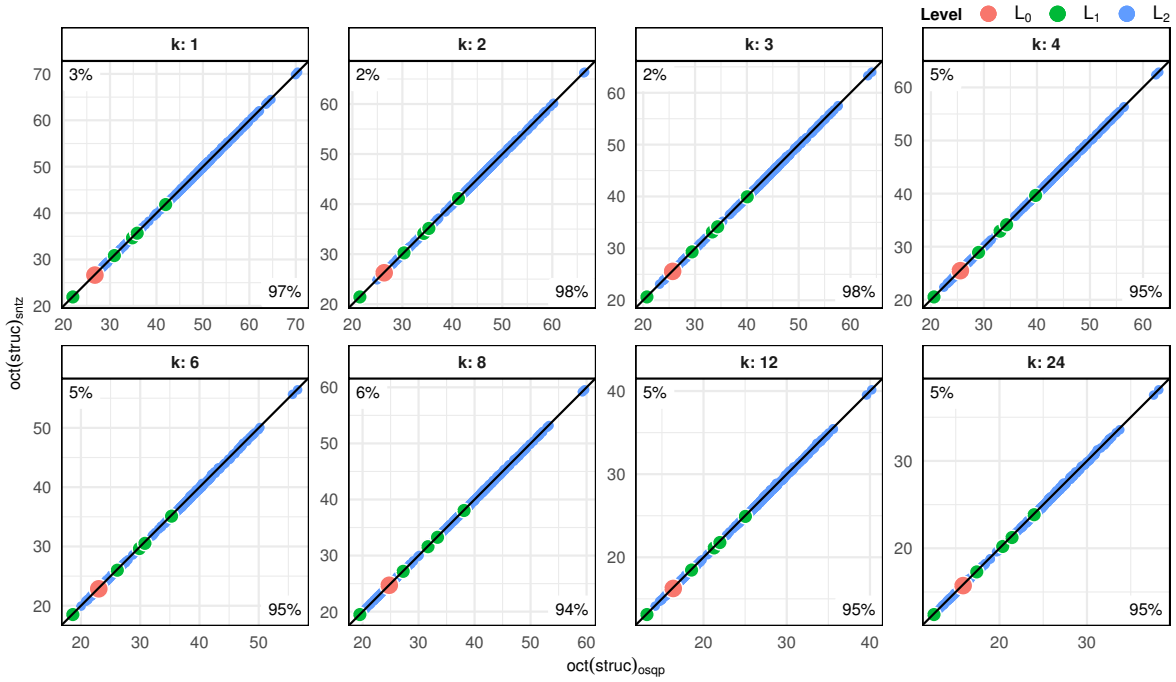


Figure 2.8: *Comparison of nRMSE(%) between sntz and osqp non-negative forecast reconciliation using the oct(struc) approach. The black line represents the bisector, where the nRMSE's for oct(struc)$_{osqp}$ and oct(struc)$_{sntz}$ are equal. On the top-left (bottom-right) corner of each graph, the percentage of points above (below) the bisector is reported.*

reported in Table 2.4. Furthermore, almost all the new approaches significantly outperform PERS$_{bu}$, 3TIER$_{bu}$, and oct(*struc*) for both hourly and daily forecasts (Figure 2.9).

For a more compelling comparison of different forecasting methods, in Table 2.8 are shown the forecast skills in terms of nRMSE of the new cross-temporal reconciliation approaches over the more challenging NWP 3TIER$_{bu}$ forecasts. All the forecasts skills are positive, and in a range between 0.047 and 0.184, for the hourly PV generated power forecasts at any cross-sectional level. In this case it turns out that the new forecast reconciliation approaches, in addition to assuring full coherence and non-negativity of the revised forecasts, have a better forecasting accuracy. For daily forecasts, however, the picture is less clear-cut. The new forecasting approaches always improve for the 318 disaggregate series at the Plant level ($\mathcal{L}_2$), whereas for ISO and Transmission Zones this happens only for two partly bottom-up approaches.

Finally, focusing on the two best performing approaches of the forecasting experiment (ct($wlsv_{\text{te}}, bu_{\text{cs}}$) and oct($wlsv$)), and looking at the nRMSEs of the individual series, it is worth noting that:

- ct($wlsv_{\text{te}}, bu_{\text{cs}}$) always produces more accurate forecasts than oct(*struc*), for any series and granularity (Figure 2.11);

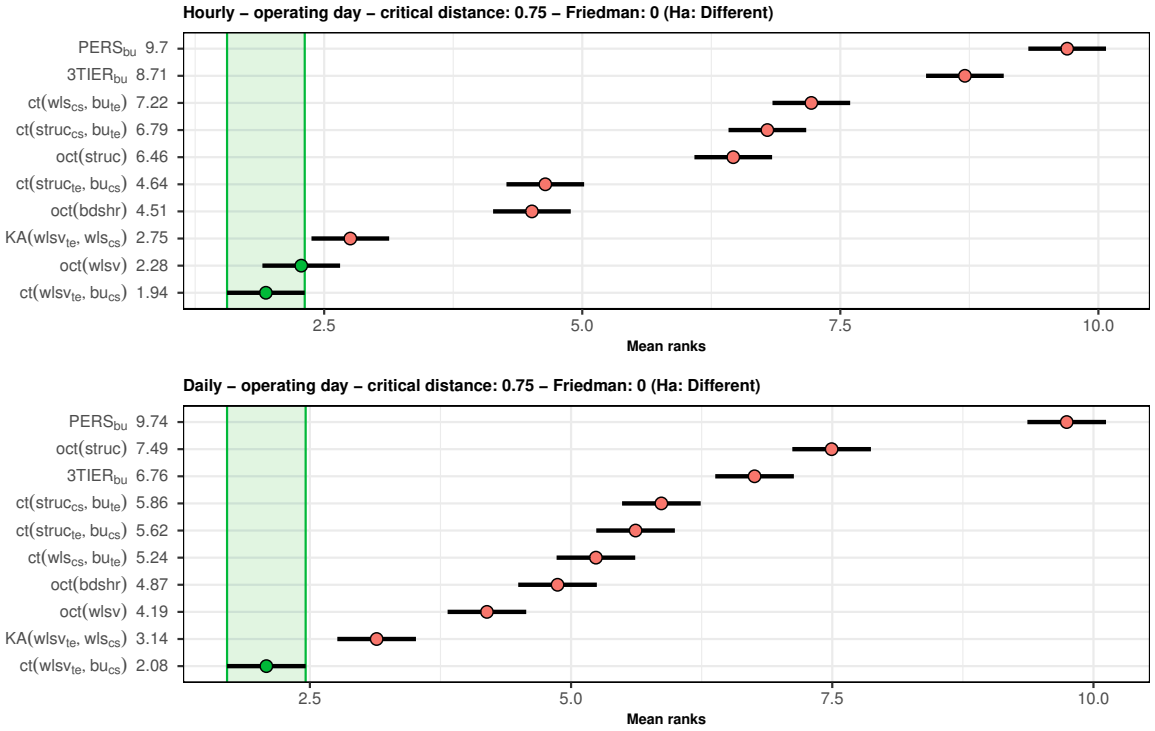


Figure 2.9: *MCB-Nemenyi test on selected non-negative cross-temporal reconciliation approaches with operating day forecast horizon. $\mathcal{L}_0, \mathcal{L}_1, \mathcal{L}_2$ levels (324 series). Top panel: hourly forecasts; Bottom panel: daily forecasts. The mean rank of each approach is displayed 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.*

Approach	Hourly forecasts			Daily forecasts		
	$\mathcal{L}_0$	$\mathcal{L}_1$	$\mathcal{L}_2$	$\mathcal{L}_0$	$\mathcal{L}_1$	$\mathcal{L}_2$
oct($struc$)	-0.023	0.028	0.130	-0.260	-0.134	0.048
KA($wlsv_{\text{te}}, wls_{\text{cs}}$)	0.094	0.110	0.168	-0.085	-0.010	0.118
ct($struc_{\text{cs}}, bu_{\text{te}}$)	0.155	0.152	0.069	0.008	0.056	0.034
ct($wls_{\text{cs}}, bu_{\text{te}}$)	0.184	0.147	0.067	0.133	0.104	0.061
ct($struc_{\text{te}}, bu_{\text{cs}}$)	0.047	0.075	0.153	-0.156	-0.063	0.089
ct($wlsv_{\text{te}}, bu_{\text{cs}}$)	0.097	0.113	0.169	-0.075	-0.002	0.122
oct($wlsv$)	0.092	0.110	0.169	-0.093	-0.016	0.116
oct($bdshr$)	0.073	0.103	0.154	-0.115	-0.022	0.105

Table 2.8: *Forecast skills over the NWP 3TIER$_{bu}$ forecasts of oct(struc) and of the new cross-temporal forecast reconciliation approaches. Red entries identify negative skills.*

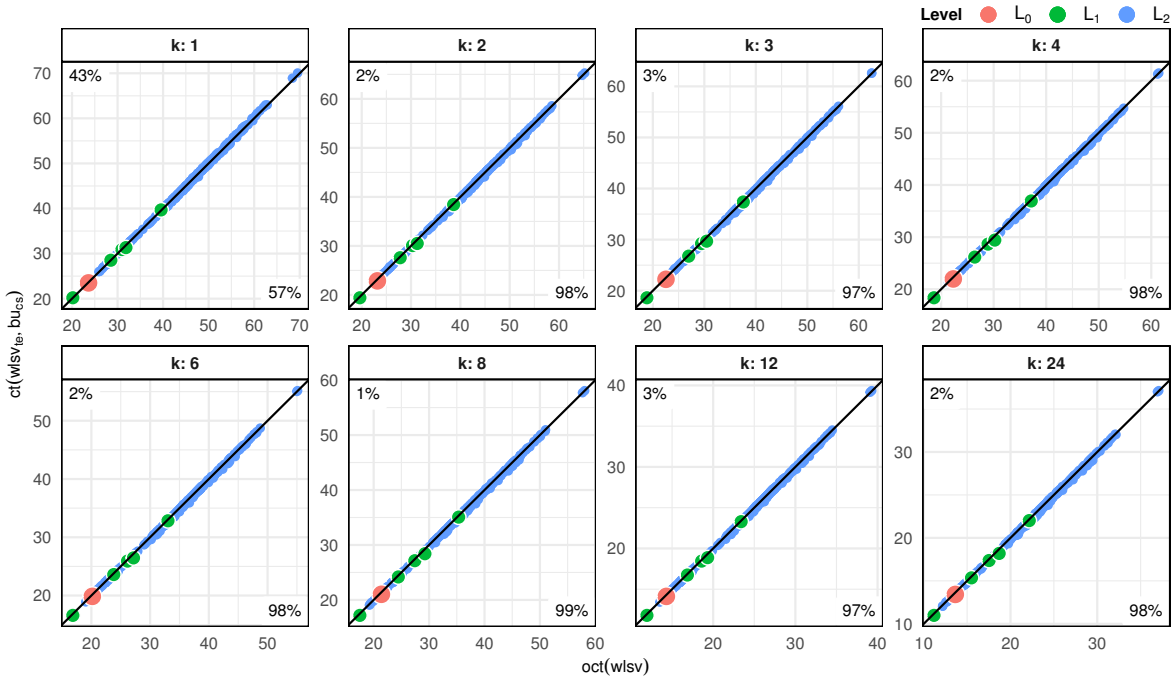


Figure 2.10: *Comparison of nRMSE(%) between non-negative reconciliation approaches: oct(wlsv) and ct($wlsv_{te}, bu_{cs}$). Forecast horizon: operating day. The black line represents the bisector, where the nRMSE's for both approaches are equal. On the top-left (bottom-right) corner of each graph, the percentage of points above (below) the bisector is reported.*

- the accuracy of ct($wlsv_{\text{te}}, bu_{\text{cs}}$) is practically indistinguishable from that of oct($wlsv$) (Figure 2.10);
- the accuracy increases of ct($wlsv_{\text{te}}, bu_{\text{cs}}$) over the NWP approach 3TIER$_{bu}$ are still clear (Figure 2.11), even though for daily forecasts 3TIER$_{bu}$ performs better in about 1 case out 4.

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 over 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 te($wlsv$) (first step of cs($wlsv_{\text{te}}, bu_{\text{cs}}$)), and oct($wlsv$), increases the forecasting accuracy at different time granularities.

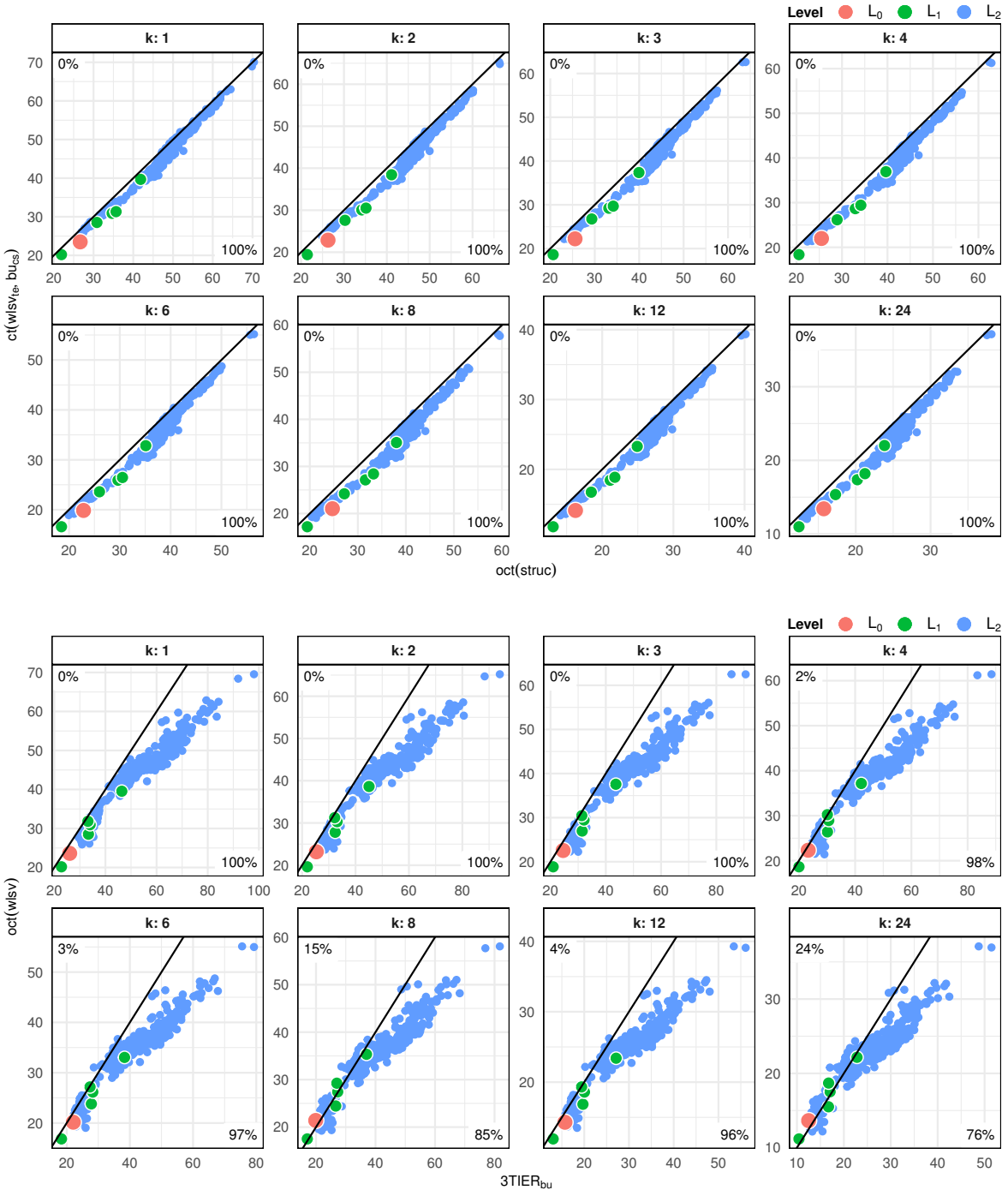


Figure 2.11: *Comparison of nRMSE(%) between non-negative reconciliation approaches: oct(struc) and $ct(wlsv_{te}, bu_{cs})$ (top), and $3TIER_{bu}$ and $ct(wlsv_{te}, bu_{cs})$ (bottom). Forecast horizon: operating day. The black line represents the bisector, where the nRMSE's for both approaches are equal. On the top-left (bottom-right) corner of each graph, the percentage of points above (below) the bisector is reported.*

Chapter 3

Forecast combination-based forecast reconciliation

3.1 Introduction

A hierarchical/grouped time series is a linearly constrained multiple time series consisting of a collection of time series that follows an aggregation structure (Panagiotelis *et al.*, 2021). As an example, sales data can be disaggregated by product categories, and then by product subcategories, down to Stock Keeping Unit (SKU). The use of hierarchical forecasting is rising in order to deliver relevant demand forecast information given various managerial levels within organizations, with the goal of providing coherent forecasts while enhancing their accuracy (Fliedner, 2001).

More generally, forecasting hierarchical/grouped time series is the process of generating coherent forecasts (or reconciling incoherent forecasts), preserving the relationships within the hierarchy/group (Hyndman and Athanasopoulos, 2021, ch. 11). It can be seen as a statistical device that can improve forecast accuracy through the use of forecast combinations, whose main motivation is to provide coherent forecasts for decisions at different levels of the hierarchy.

Classical approaches to hierarchical forecasting are bottom-up, top-down, and middle-out (Athanasopoulos *et al.*, 2009). Bottom-up forecasting (Dunn *et al.*, 1976) involves forecasting the most granular level of the hierarchy, then aggregating up to create estimates for the higher levels. The main advantage of this method is that, because forecasts are obtained at the lowest level of the hierarchy, no information is lost due to aggregation. However, it ignores the relationships between the series, and usually performs poorly on highly aggregated data. Furthermore, information at lower levels of

the hierarchy tends to be noisier, potentially resulting in a reduced overall forecast accuracy. In the top-down approach (Gross and Sohl, 1990), the top level of the hierarchy is first forecast, and then this forecast is split up to get estimates for the lower levels, typically using historical or forecasted proportions (Athanasopoulos *et al.*, 2009). The middle-out approach is a combination of the bottom-up and top-down approaches: the middle level (neither the most granular nor the most aggregated) variables are forecast. These values are then used to compute the higher levels' forecasts using the bottom-up approach, and the forecasts for lower levels with the top-down approach.

Modern least squares-based reconciliation techniques emerged in the cross-sectional framework (optimal combination approach, Hyndman *et al.*, 2011), and have then been extended both in temporal (Athanasopoulos *et al.*, 2017; Nystrup *et al.*, 2020), and cross-temporal (Kourentzes and Athanasopoulos, 2019, see also Chapter 1) frameworks, designed to align short- and long-term forecasts for consistency of different planning and budgeting purposes. This class of techniques is usually associated to a forecasting scheme where all the time series are independently forecast at all levels (cross-sectional and/or temporal), producing incoherent base forecasts, that are then transformed using all the information and relationships a hierarchy can offer in order to be coherent along the chosen dimensions (cross-sectional, temporal, or both). This result is obtained using a linear regression model, and the newly coherent forecasts are a weighted sum of the forecasts from all levels, with the weights found by solving a system of equations ensuring the natural relationships between the different levels of the hierarchy are satisfied. If the base forecasts are unbiased, Hyndman *et al.* (2011) show that the optimal combination approach provides unbiased reconciled forecasts at all levels with minimal loss of information, taking advantage of the relationships between time series to find patterns.

Hollyman *et al.* (2021) were the first to examine the forecast reconciliation problem from a forecast combination perspective (Bates and Granger, 1969; Timmermann, 2006). They have proposed a forecast combination based approach to the reconciliation of a simple hierarchy (called Level-l Conditional Coherent, L_lCC), showing how to generate unbiased top-down and middle-out forecasts for a hierarchy. In addition, it was shown that the simple average of a set of L_lCC and bottom-up reconciled forecasts results in good performance as compared to those obtained through the state-of-the-art cross-sectional reconciliation procedures (Wickramasuriya *et al.*, 2019). In this chapter, we build upon and extend this proposal along some new directions.

1. We shed light on the nature and the mathematical derivation of the L_lCC reconciliation formula for an elementary two-level hierarchy, showing that it is the

result of a constrained minimization of a quadratic loss function in the differences between the target and the base forecasts, with a diagonal associated matrix and exact linear *inhomogeneous* constraints. The constant part of the constraints is *exogenously* given by the base forecasts of the upper time series at the generic level l of the hierarchy, which is thus not revised, as it happens for the top-down approach (Athanasopoulos *et al.*, 2009).

2. In the same framework, exact linear *homogeneous* constraints may be considered as well, where no exogenous value is imposed, resulting in level conditional reconciled forecasts of all the involved series, where both the top and the bottom level time series are coherently revised. Similar to the previous case, we qualify this approach as *endogenous*, in the sense that no external value (i.e., base forecast) is part of the constraints. We show that the reconciled forecasts can be interpreted as the combination of direct (base) and indirect forecasts, the latter ones depending on the accounting relationships linking upper and bottom series.
3. We show that the L_lCC approach (i.e., with exogenous constraints, but the result holds in the endogenous case as well) does not guarantee the non-negativity of the reconciled forecasts, that can be an issue in cases where non-negativity is a natural attribute of the variables to be forecast (e.g., sales, tourism flows, etc.).
4. The new procedures, available in the `R` package `FoReco` (Girolimetto and Di Fonzo, 2023a), are used in two forecasting experiments on hierarchical datasets already used in the forecasting literature (Wickramasuriya *et al.*, 2019; Athanasopoulos *et al.*, 2020). In particular, the original results found by Hollyman *et al.* (2021) on the classical Australian Tourism Demand (Visitor Nights) dataset are re-assessed (i) using the relative accuracy indices for multiple comparisons recommended by Davydenko and Fildes (2013), and (ii) taking into account the non-negativity issues posed by the dataset at hand.
5. Finally, due to the crucial role played by the (possibly different) models used to compute the base forecasts, the CCC reconciliation approach proposed by Hollyman *et al.* (2021) (called CCC_H in this chapter) is interpreted as a forecast pooling strategy (Hendry and Clements, 2004; Marcellino, 2004; Geweke and Amisano, 2011; Kourentzes *et al.*, 2019), and we show that the intuition behind the CCC_H strategy can be further improved by adopting a somehow less arbitrary simple forecast averaging strategy.

The chapter is organized as follows. In Section 3.2 we set the notation, define the forecast reconciliation problem, and show the Level-l Conditional Coherent forecast reconciliation approach proposed by Hollyman *et al.* (2021). We re-interpret this procedure

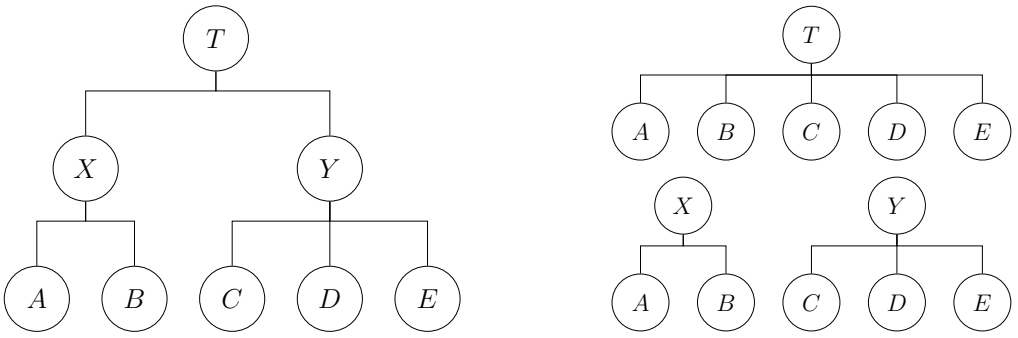


Figure 3.1: *A three-level hierarchy (left), and its elementary hierarchies (right)*

in terms of an optimization problem, with either exogenous (Section 3.3), or endogenous (Section 3.4) constraints, which encompasses the forecast reconciliation procedure so far as a particular case.

3.2 Problem definition and notation

Let us consider a linearly constrained multiple time series with a genuine hierarchical/-grouped structure consisting of $L > 0$ levels above the bottom level. When $L = 1$ we face an *elementary hierarchy*, formed by one top-level series and n_b component bottom time series (bts). In general, denote n_a the number of the upper time series (uts) in a hierarchy, and n_l the number of knots (series) of the generic level l of the series, $l = 1, \ldots, L$, with $n_1 = 1$, and $\sum_{l=1}^{L} n_l = n_a$.

Consider now the L hierarchical structures simply formed by the n_b bts and the n_l upper time series (uts) of level l, $l = 1, \ldots, L$. For example, the hierarchical series in the left panel of Figure 3.1 (see Appendix 1 in Hollyman *et al.*, 2021) consists of three levels: the total series T at level 1, series X and Y at the intermediate level 2, and the five bottom time series A, B, C, D, E, at the bottom level 3, with $T = X + Y = A + B + C + D + E$, $X = A + B$, and $Y = C + D + E$. Such a series may be represented in structural form as $\boldsymbol{y} = \boldsymbol{S}\boldsymbol{b}$, where

$$\boldsymbol{S} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ & & \boldsymbol{I}_5 & & \end{bmatrix}$$

is a (8×5) cross-sectional summing matrix, $\boldsymbol{y} = \Big[T\ X\ Y\ A\ B\ C\ D\ E\Big]'$, and $\boldsymbol{b} = \Big[A\ B\ C\ D\ E\Big]'$.

The main series $\boldsymbol{y}$ contains other two linearly constrained multiple time series, sharing the same bts: $\boldsymbol{y}_1 = \boldsymbol{S}_1\boldsymbol{b}$ and $\boldsymbol{y}_2 = \boldsymbol{S}_2\boldsymbol{b}$, with, respectively,

$$\boldsymbol{S}_1 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ & & \boldsymbol{I}_5 & & \end{bmatrix}, \quad \boldsymbol{y}_1 = \begin{bmatrix} T & A & B & C & D & E \end{bmatrix}'$$

and

$$\boldsymbol{S}_2 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ & & \boldsymbol{I}_5 & & \end{bmatrix}, \quad \boldsymbol{y}_2 = \begin{bmatrix} X & Y & A & B & C & D & E \end{bmatrix}'.$$

Notice that the matrix $\boldsymbol{S}_1$ describes the elementary hierarchy formed by the top-level and the bottom series, not considering the intermediate level $l = 2$, whereas $\boldsymbol{S}_2$ does not correspond to a *summing matrix* of a standard structural representation of a hierarchical/grouped series, because a unique top level series is not present. In this latter case, we may recognize two distinct elementary hierarchies (right panel of Figure 3.1): the former valid for series X, re-interpreted as a 'top level series', and the latter for series Y as well, respectively given by:

$$\begin{bmatrix} X \\ A \\ B \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \boldsymbol{I}_2 & \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix}, \quad \begin{bmatrix} Y \\ C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ & \boldsymbol{I}_3 & \end{bmatrix} \begin{bmatrix} C \\ D \\ E \end{bmatrix}.$$

Let $\boldsymbol{C}$ be the (3×5) cross-sectional (contemporaneous) aggregation matrix mapping the five bts into the three uts of the multiple series, which is linked to the summing matrix $\boldsymbol{S}$ by the relationship $\boldsymbol{S} = \begin{bmatrix} \boldsymbol{C}' & \boldsymbol{I}_5 \end{bmatrix}'$. Matrix $\boldsymbol{C}$ consists of two sub-matrices $\boldsymbol{C}_l$, $l = 1, 2$:

$$\boldsymbol{C} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix} \quad \boldsymbol{C}_1 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad \boldsymbol{C}_2 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}.$$

The level-l constrained multiple time series may be thus represented as $\boldsymbol{y}_l = \boldsymbol{S}_l\boldsymbol{b}$, $l = 1, 2$, where the vector $\boldsymbol{y}_l = \begin{bmatrix} \boldsymbol{a}_l \\ \boldsymbol{b} \end{bmatrix}$ has dimension $[(n_l + n_b) \times 1]$, n_l being the number of knots (series) in the $(n_l \times 1)$ vector of the level-l time series $\boldsymbol{a}_l$ (i.e., $n_1 = 1$, and $n_2 = 2$), and $\boldsymbol{S}_l = \begin{bmatrix} \boldsymbol{C}_l' & \boldsymbol{I}_{n_b} \end{bmatrix}'$, $l = 1, 2$, has dimension $[(n_l + 5) \times 5]$.

It should be added that, in order to develop the results that follow, the hierarchy has to be **balanced**, that is, each 'knot' (series) at an upper level wrt the bottom one, must have at least a 'child' series. A simple unbalanced three-level hierarchy is shown in the left panel of Figure 3.2, where variable C has no 'children', and thus is considered as

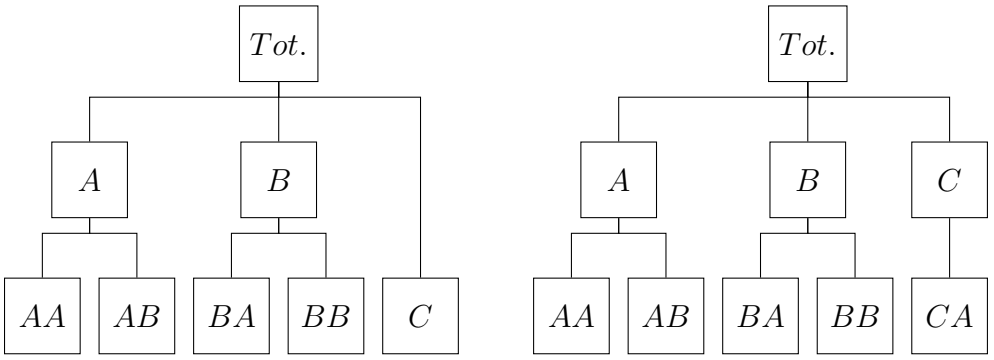


Figure 3.2: *A simple unbalanced hierarchy (left) and its balanced version (right)*

a bottom variable, at level three of the hierarchy. The right panel shows the 'balanced version' of the same hierarchy, where $CA = C$, and thus variable C is (duplicated and) present at both levels two and three. Possible duplication of some variables should be conveniently accounted for, e.g. when evaluating the reconciled forecasts' accuracy for all the series in the hierarchy.

3.2.1 Notation

Let l, $l = 1, \ldots, L$, the index associated to a generic level - above the bts level - of the hierarchy/grouping which characterizes the multiple time series whose base forecasts are wished to be reconciled, and assume that $l = 1$ denotes the top-level, consisting of the total of the whole hierarchical/grouped system. Let $\boldsymbol{y} = \begin{bmatrix}\boldsymbol{a}' & \boldsymbol{b}'\end{bmatrix}'$ be the $(n \times 1)$ vector of target forecasts, formed by the $(n_a \times 1)$ vector $\boldsymbol{a}$ of upper time series (uts), and by the $(n_b \times 1)$ vector $\boldsymbol{b}$ of bottom time series (bts). Denote the base forecasts vector $\widehat{\boldsymbol{y}} = \begin{bmatrix}\widehat{\boldsymbol{a}}' & \widehat{\boldsymbol{b}}'\end{bmatrix}'$. In addition, decompose vector $\widehat{\boldsymbol{a}}$ into the sub-vectors forming each of the upper L levels of the hierarchy/grouping: $\widehat{\boldsymbol{a}} = \begin{bmatrix}\widehat{a}_1 & \widehat{\boldsymbol{a}}_2' & \ldots & \widehat{\boldsymbol{a}}_L'\end{bmatrix}'$, where $\widehat{\boldsymbol{a}}_l$, $l = 1, \ldots, L$, has dimension $(n_l \times 1)$.

Denote $\boldsymbol{C}_l$ the $(n_l \times n_b)$ matrix mapping the bts into the level-l uts (i.e., $\boldsymbol{a}_l = \boldsymbol{C}_l\boldsymbol{b}$). The complete aggregation matrix $\boldsymbol{C}$, mapping all the bts into the uts of all levels $l = 1, \ldots, L$, may be written as

$$\boldsymbol{C} = \begin{bmatrix}\boldsymbol{C}_1 \\ \boldsymbol{C}_2 \\ \vdots \\ \boldsymbol{C}_L\end{bmatrix}, \tag{3.1}$$

where the generic matrix $\boldsymbol{C}_l$ is $(n_l \times n_b)$, $l = 1, \ldots, L$, and $\boldsymbol{C}_1 = \mathbf{1}_{n_b}'$ is a $(1 \times n_b)$ row vector of ones (sum vector).

3.2.2 Level-l Conditional Coherent (L_lCC) reconciliation

The central core of the proposal by Hollyman *et al.* (2021) is a *level conditional coherent* forecast reconciliation procedure which, for any level l of the hierarchy, transforms the vector $\widehat{\boldsymbol{b}}$ of unbiased base forecasts of the bts in unbiased reconciled forecasts $\widetilde{\boldsymbol{b}}^{(l)}$ coherent with the **base forecasts** of the series at that specific level of the hierarchy. The exponent $^{(l)}$ highlights that the base forecasts of the bts are transformed in such a way that they are coherent with the base forecasts of the series at level l (or, equivalently, that the reconciled forecasts are coherent conditional to the base forecasts of the series at level l). A set of L reconciled forecasts, each conditional to the n_l base forecasts of a specific level-l, may thus be computed. When $l = 1$, i.e. the only upper level series is the total aggregate at the top of the hierarchy, Hollyman *et al.* (2021) interpret this procedure as a *top down forecast reconciliation*, as the conditioning top-level forecast remains unchanged, just as it happens with the classical top-down reconciliation (Gross and Sohl, 1990; Athanasopoulos *et al.*, 2009). In the remaining cases ($1 < l \leq L$), in order to coherently adjust the whole vector of forecasts, the level conditional reconciled forecasts are transformed through a *middle-out* reconciliation procedure, by simply pre-multiplying the bts reconciled forecasts vector by the summing matrix $\boldsymbol{S}$.

More precisely, denoting $\widehat{a}_1$ and $\widehat{\boldsymbol{b}}$ the base forecasts of, respectively, the total (top level) and the n_b bottom time series for a fixed forecast horizon, for a given $(n_b \times 1)$ vector $\boldsymbol{p}$ of combination weights, i.e. $0 < p_i < 1$, $\sum_{i=1}^{n_b} p_i = 1$, Hollyman *et al.* (2021) show that the Level-1 Conditional Coherent (L_1CC) bts reconciled forecasts are given by

$$\widetilde{b}_i^{(1)} = \widehat{b}_i + p_i \left(\widehat{a}_1 - \sum_{j=1}^{n_b} \widehat{b}_j \right), \quad i = 1, \ldots, n_b, \tag{3.2}$$

and the complete reconciled forecasts vector is given by $\widetilde{\boldsymbol{y}}^{(1)} = \boldsymbol{S}\widetilde{\boldsymbol{b}}^{(1)}$. Furthermore, by re-stating expression (3.2) as

$$\widetilde{b}_i^{(1)} = (1 - p_i)\widehat{b}_i + p_i \left(\widehat{a}_1 - \sum_{j=1,\, j\neq i}^{n_b} \widehat{b}_j \right), \quad i = 1, \ldots, n_b, \tag{3.3}$$

the L_1CC reconciled forecast (3.3) can be seen as the forecast combination of the direct (base) forecast $\widehat{b}_i$ and of its indirect counterpart $\left(\widehat{a}_1 - \sum_{j=1,\, j\neq i}^{n_b} \widehat{b}_j \right)$, which is coherent with the accounting constraint linking the total and the bottom time series, with weights given by $(1 - p_i)$ and p_i, respectively.

Expression (3.2) may be extended to any level l, $1 \leq l \leq L$, as

$$\widetilde{\boldsymbol{b}}^{(l)} = \widehat{\boldsymbol{b}} + \boldsymbol{P}_l \left(\widehat{\boldsymbol{a}}_l - \boldsymbol{C}_l \widehat{\boldsymbol{b}} \right), \quad l = 1, \ldots, L, \tag{3.4}$$

where $\boldsymbol{P}_l$ is the $(n_b \times n_l)$ matrix containing the combination weights of each forecast in the n_l elementary hierarchies linking each of the n_l level-l series to their 'afferent' bts. More precisely,

$$\boldsymbol{P}_l = \begin{bmatrix} \boldsymbol{p}_{l,1} & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{p}_{l,n_l} \end{bmatrix}, \quad l = 1, \ldots, L, \tag{3.5}$$

where each $\boldsymbol{p}_{l,j}$, $l = 1, \ldots, L$, $j = 1, \ldots, n_l$, is a vector of weights between zero and one, which sum up to one, with a number of elements equal to the bts contributing to the upper level's counterparts. For $l = 1$ $\boldsymbol{P}_l$ reduces to vector $\boldsymbol{p}$, and it is easy to check that $\boldsymbol{C}_l \boldsymbol{P}_l = \boldsymbol{I}_{n_l}$, $l = 1, \ldots, L$. The algebra behind these results found by Hollyman *et al.* (2021) is shown in detail in Appendix B.1. In addition, Hollyman *et al.* (2021) show that, if the base forecasts are unbiased, the reconciled forecasts this way are unbiased as well.

It should be noted that expression (3.2) is not guaranteed to produce non-negative $\widetilde{b}_i^{(1)}$: if the quantity (discrepancy) $\left(\widehat{a}_1 - \sum_{j=1}^{n_b} \widehat{b}_j \right) < 0$, it may happen that the correction to $\widehat{b}_i$, equal to a share $0 < p_i < 1$ of the discrepancy, be larger in absolute value than the base forecast to be adjusted, thus producing a negative reconciled forecast. In particular, when the discrepancy is negative and $\widehat{b}_i \leq 0$, it is always $\widetilde{b}_i^{(1)} < 0$. This could be (counter-intuitive, and) 'annoying' when the variables of the hierarchy are intrinsically non-negative (e.g., sales, tourism flows, etc.), and should be kept in mind by the analyst. In the next two sections we provide a different, meaningful interpretation of expression (3.2), by finding the general solution to the problem of deriving reconciled forecasts according to a level conditional coherent procedure, with either exogenous or endogenous constraints, which contains the forecast reconciliation procedure by Hollyman *et al.* (2021) as a particular case.

3.3 L_lCC reconciliation with exogenous constraints

Given a generic level l, reconciled forecasts coherent with the base forecasts of that level can be obtained by solving the following linearly constrained quadratic minimization problem:

$$\widetilde{\boldsymbol{b}}^{(l)}_{\mathrm{L_lCC}} = \arg\min_{\boldsymbol{b}} \left(\boldsymbol{b} - \widehat{\boldsymbol{b}} \right)' \boldsymbol{W}_b^{-1} \left(\boldsymbol{b} - \widehat{\boldsymbol{b}} \right) \quad \text{s.t. } \boldsymbol{C}_l \boldsymbol{b} = \widehat{\boldsymbol{a}}_l, \quad l = 1, \ldots, L, \tag{3.6}$$

where $\boldsymbol{W}_b$ is a $(n_b \times n_b)$ p.d. matrix, and vector $\widehat{\boldsymbol{a}}_l$ defines the known constant part of the exact linear *inhomogeneous* constraints $\boldsymbol{C}_l\boldsymbol{b} = \widehat{\boldsymbol{a}}_l$, *exogenously* given by the the base forecasts of the upper time series at level l of the hierarchy. The solution is given by (Appendix B.2):

$$\widetilde{\boldsymbol{b}}^{(l)}_{\mathrm{L}_l\mathrm{CC}} = \widehat{\boldsymbol{b}} + \boldsymbol{W}_b\boldsymbol{C}_l' \left(\boldsymbol{C}_l\boldsymbol{W}_b\boldsymbol{C}_l'\right)^{-1} \left(\widehat{\boldsymbol{a}}_l - \boldsymbol{C}_l\widehat{\boldsymbol{b}}\right), \quad l = 1, \ldots, L, \tag{3.7}$$

that is $\widetilde{\boldsymbol{b}}^{(l)}_{\mathrm{L}_l\mathrm{CC}} = \boldsymbol{L}_l\widehat{\boldsymbol{a}}_l + \left(\boldsymbol{C}_l\boldsymbol{W}_b\boldsymbol{C}_l'\right)^{-1}\widehat{\boldsymbol{b}}$, with $\boldsymbol{L}_l = \boldsymbol{W}_b\boldsymbol{C}_l' \left(\boldsymbol{C}_l\boldsymbol{W}_b\boldsymbol{C}_l'\right)^{-1}$.

It is worth noting that formula (3.7) is a well known expression in the field of the least squares adjustment of noisy data while fulfilling an aggregation - either cross-sectional or temporal - constraint (Stone *et al.*, 1942; Denton, 1971; Chow and Lin, 1971; Byron, 1978). Put simply, the reconciliation formula (3.7) 'adjusts' the bts base forecasts with a linear combination - according to the smoothing matrix $\boldsymbol{L}_l$ - of the level-l coherency errors $\left(\widehat{\boldsymbol{a}}_l - \boldsymbol{C}_l\widehat{\boldsymbol{b}}\right)$.

Matrix $\boldsymbol{W}_b$ plays a crucial role in the L_lCC reconciliation formula. From a mathematical point of view, it is the associated matrix in the quadratic objective function of problem (3.6). From a statistical point of view, $\boldsymbol{W}_b$ is usually seen as the forecast error covariance matrix, whose evaluation/approximation is based either on simple assumptions of independence between different forecasts, which results in a diagonal $\boldsymbol{W}_b$ matrix, or on using the in-sample errors of the models used to compute the base forecasts (see Wickramasuriya *et al.*, 2019; Nystrup *et al.*, 2020 and Chapter 1, for alternative choices, respectively, in the cross-sectional, temporal, and cross-temporal forecast reconciliation frameworks).

Expression (3.7) is equivalent to formula (3.4) when we assume a diagonal bts forecast error covariance matrix, where the variances on the diagonal are equal to the reciprocal of the combination weights:

$$\boldsymbol{W}_b = \begin{bmatrix} \sigma_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{n_b}^2 \end{bmatrix} = \begin{bmatrix} \dfrac{1}{p_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \dfrac{1}{p_{n_b}} \end{bmatrix} = \left[\mathrm{diag}\left(\boldsymbol{p}\right)\right]^{-1}. \tag{3.8}$$

In other terms, the L_lCC reconciliation formula (3.4) by Hollyman *et al.* (2021) may be interpreted as the solution to a linearly constrained minimization of a quadratic form with the associated diagonal matrix (3.8). Hollyman *et al.* (2021) show that the diagonal pattern of matrix $\boldsymbol{W}_b$ is consistent with the usual practice in forecast combination of discarding possible covariances between the forecasts to be combined (Bates and Granger, 1969). It should be noted that nothing prevents us to consider a full, instead of diagonal, $\boldsymbol{W}_b$ matrix. Obviously, this would pose non trivial estimation

RECONCILIATION STEPS

Upper Level Conditional Coherent reconciliation					Bottom up
$\widehat{a}_1$	$\cdots$	$\widehat{\boldsymbol{a}}_l$	$\cdots$	$\widehat{\boldsymbol{a}}_L$	
$\downarrow$		$\downarrow$		$\downarrow$	
$\widehat{\boldsymbol{b}} \rightarrow \boxed{L_1CC}$	$\cdots$	$\widehat{\boldsymbol{b}} \rightarrow \boxed{L_lCC}$	$\cdots$	$\widehat{\boldsymbol{b}} \rightarrow \boxed{L_LCC}$	
$\downarrow$		$\downarrow$		$\downarrow$	
$\begin{bmatrix} \widehat{a}_1 \\ \widetilde{\boldsymbol{b}}^{(1)} \end{bmatrix}$	$\cdots$	$\begin{bmatrix} \widehat{\boldsymbol{a}}_l \\ \widetilde{\boldsymbol{b}}^{(l)} \end{bmatrix}$	$\cdots$	$\begin{bmatrix} \widehat{\boldsymbol{a}}_L \\ \widetilde{\boldsymbol{b}}^{(L)} \end{bmatrix}$	
$\downarrow$		$\downarrow$		$\downarrow$	
$\widetilde{\boldsymbol{y}}^{(1)} = \boldsymbol{S}\widetilde{\boldsymbol{b}}^{(1)}$	$\cdots$	$\widetilde{\boldsymbol{y}}^{(l)} = \boldsymbol{S}\widetilde{\boldsymbol{b}}^{(l)}$	$\cdots$	$\widetilde{\boldsymbol{y}}^{(L)} = \boldsymbol{S}\widetilde{\boldsymbol{b}}^{(L)}$	$\widetilde{\boldsymbol{y}}^{(L+1)} = \boldsymbol{S}\widehat{\boldsymbol{b}}$

$$\Downarrow$$

$$\widetilde{\boldsymbol{y}}_{CCC} = \frac{1}{L+1}\sum_{l=1}^{L+1}\widetilde{\boldsymbol{y}}^{(l)}$$

Figure 3.3: *CCC: Combined Conditional Coherent forecast reconciliation procedure with the same bts base forecasts* $\widehat{\boldsymbol{b}}$

issues, which may benefit of possibly available forecast error estimates, either from in-sample residuals or out-of-sample forecast errors in validation sets. In this chapter we adopt the choice by Hollyman *et al.* (2021) of using a simple formulation of $\boldsymbol{W}_b$, consisting of the diagonal matrix of the observed variability of the bottom time series in the training set used to estimate the base forecasts, but we think that this issue is worth considering in future research on this topic.

3.3.1 Combined Conditional Coherent (CCC) reconciliation

For any fixed level l, the complete vector of L_lCC reconciled forecasts may be computed as

$$\widetilde{\boldsymbol{y}}^{(l)} = \boldsymbol{S}\widetilde{\boldsymbol{b}}^{(l)}, \quad l = 1, \ldots, L. \tag{3.9}$$

For $l = 1$, expression (3.9) returns reconciled forecasts obtained by summing-up the L_1CC bts reconciled forecasts $\widetilde{\boldsymbol{b}}^{(1)}$. In the remaining cases ($1 < l \leq L$), each complete vector of L_lCC reconciled forecasts is the result of a middle-out reconciliation procedure applied to the L_lCC bts reconciled forecasts. Finally, taking inspiration from Hollyman *et al.* (2021), the bottom-up reconciled forecasts may be considered as the reconciled forecasts coherent with the bottom level of the hierarchy ($l = L + 1$)[1]: $\widetilde{\boldsymbol{y}}^{(L+1)} = \boldsymbol{S}\widehat{\boldsymbol{b}}$. These $L + 1$ vectors contain different and coherent forecasts, so that any convex linear

[1]As we will discuss in Section 3.5.1, the empirical application of Hollyman *et al.* (2021) makes use of two different base forecasts when performing the L_lCC and the bottom-up reconciliation steps, respectively.

combination of the form

$$\widetilde{\boldsymbol{y}}_{\omega} = \sum_{l=1}^{L+1} \omega_l \widetilde{\boldsymbol{y}}^{(l)} = \boldsymbol{S} \sum_{l=1}^{L+1} \omega_l \widetilde{\boldsymbol{b}}^{(l)} = \boldsymbol{S}\widetilde{\boldsymbol{b}}_{\omega},$$

where $\widetilde{\boldsymbol{b}}_{\omega} = \sum_{l=1}^{L+1} \omega_l \widetilde{\boldsymbol{b}}^{(l)}$, with $0 \leq \omega_l \leq 1$, $\sum_{l=1}^{L+1} \omega_l = 1$, is coherent as well. Combined Conditional Coherent (CCC) forecasts are calculated by using equal weights, which means to combine the $L+1$ reconciled forecasts through the simple average (see Figure 3.3)

$$\widetilde{\boldsymbol{y}}_{CCC} = \frac{1}{L+1} \sum_{l=1}^{L+1} \widetilde{\boldsymbol{y}}^{(l)}. \tag{3.10}$$

From the scheme in Figure 3.3 it appears that the L_lCC steps are logically different from the bottom-up one. In the former case, base forecasts of both upper series (part of them at each step) and bottom series, are combined through the optimization mechanism described above, while in the latter no upper time series base forecasts are used. Thus, it seems rather sensible also considering a variant of the CCC procedure, called LCC, which discards the bottom-up forecasts from the average:

$$\widetilde{\boldsymbol{y}}_{LCC} = \frac{1}{L} \sum_{l=1}^{L} \widetilde{\boldsymbol{y}}^{(l)}. \tag{3.11}$$

3.3.2 Some examples

Level 1 Conditional Coherent (L_1CC) reconciliation

Consider $l = 1$, and let $\boldsymbol{W}_b$ be defined as in (3.8). Since $\boldsymbol{C}_1 = \mathbf{1}'_{n_b}$, it is immediately recognized that $\boldsymbol{W}_b\boldsymbol{C}_1' = \left[\sigma_1^2 \ \dots \ \sigma_i^2 \ \dots \ \sigma_{n_b}^2\right]'$, and $\boldsymbol{C}_1\boldsymbol{W}_b\boldsymbol{C}_1' = \sum_{j=1}^{n_b} \sigma_j^2$. The generic item of the vector obtained through the reconciliation formula (3.7) is thus given by:

$$\widetilde{b}_i^{(1)} = \widehat{b}_i + \frac{\sigma_i^2}{\sum_{j=1}^{n_b} \sigma_j^2} \left(\widehat{a}_1 - \sum_{j=1}^{n_b} \widehat{b}_j \right), \quad i = 1, \dots, n_b,$$

that is

$$\widetilde{b}_i^{(1)} = \left(\frac{\sum_{j=1,\, j\neq i}^{n_b} \sigma_j^2}{\sum_{j=1}^{n_b} \sigma_j^2} \right) \widehat{b}_i + \frac{\sigma_i^2}{\sum_{j=1}^{n_b} \sigma_j^2} \left(\widehat{a}_1 - \sum_{j=1,\, j\neq i}^{n_b} \widehat{b}_j \right), \quad i = 1, \dots, n_b,$$

which corresponds to equation (3.3).

CCC reconciliation in a three-level hierarchy

When a level $l > 1$ is considered, again with a diagonal $\boldsymbol{W}_b$, the scalar expressions of the L_lCC reconciled forecasts are more complex. Nevertheless, starting from the simple hierarchy in Figure 3.1 (which corresponds to the example 2 in Hollyman *et al.*, 2021), we can easily interpret the final result. In this case it is easy to check that

$$\boldsymbol{W}_b\boldsymbol{C}_2' = \begin{bmatrix} \sigma_A^2 & 0 \\ \sigma_B^2 & 0 \\ 0 & \sigma_C^2 \\ 0 & \sigma_D^2 \\ 0 & \sigma_E^2 \end{bmatrix} \quad \text{and} \quad \boldsymbol{C}_2\boldsymbol{W}_b\boldsymbol{C}_2' = \begin{bmatrix} \sigma_A^2+\sigma_B^2 & 0 \\ 0 & \sigma_C^2+\sigma_D^2+\sigma_E^2 \end{bmatrix},$$

from which we obtain

$$\widetilde{b}_i^{(2)} = \begin{cases} \widehat{b}_i + \dfrac{\sigma_i^2}{\sigma_A^2+\sigma_B^2}\left(\widehat{a}_X - \widehat{b}_A - \widehat{b}_B\right) & i = A, B \\ \widehat{b}_i + \dfrac{\sigma_i^2}{\sigma_C^2+\sigma_D^2+\sigma_E^2}\left(\widehat{a}_Y - \widehat{b}_C - \widehat{b}_D - \widehat{b}_E\right) & i = C, D, E \end{cases},$$

where $\widehat{a}_X$ and $\widehat{a}_Y$ denote the base forecasts of X and Y, respectively. In other words, level 2, consisting of series X and Y, is 'decomposed' in two elementary hierarchies, each consisting in a single aggregated series belonging to that level, and in the corresponding bts. Then, a L_1CC forecast reconciliation procedure is applied to each elementary hierarchy.

The whole reconciliation is obtained by summing-up the reconciled forecasts of the bts (and thus the final result should be viewed as a middle-out forecast reconciliation):

$$\widetilde{a}_T^{(2)} = \widehat{a}_X + \widehat{a}_Y = \widetilde{b}_A^{(2)} + \widetilde{b}_B^{(2)} + \widetilde{b}_C^{(2)} + \widetilde{b}_D^{(2)} + \widetilde{b}_E^{(2)}.$$

Notice that in general $\widetilde{a}_T^{(2)} \neq \widehat{a}_T$, and likewise all the reconciled forecasts this way are different from those obtained if we consider the reconciliation conditional to level 1 (that is, to the base forecast of the top-level series, $\widehat{a}_T$). As for the CCC reconciled forecasts, by using expression (3.10) we get:

$$\widetilde{a}_{T,CCC} = \frac{1}{3}\left(\widetilde{a}_T^{(1)} + \widetilde{a}_T^{(2)} + \widetilde{a}_T^{(3)}\right) = \frac{1}{3}\Big(\widehat{a}_T + \underbrace{\widehat{a}_X + \widehat{a}_Y}_{\widetilde{a}_T^{(2)}} + \underbrace{\widehat{b}_A + \widehat{b}_B + \widehat{b}_C + \widehat{b}_D + \widehat{b}_E}_{\widetilde{a}_T^{(3)}}\Big)$$

$$\widetilde{a}_{X,CCC} = \frac{1}{3}\left(\widetilde{a}_X^{(1)} + \widetilde{a}_X^{(2)} + \widetilde{a}_X^{(3)}\right) = \frac{1}{3}\Big(\underbrace{\widetilde{b}_A^{(1)} + \widetilde{b}_B^{(1)}}_{\widetilde{a}_X^{(1)}} + \widehat{a}_X + \underbrace{\widehat{b}_A + \widehat{b}_B}_{\widetilde{a}_X^{(3)}}\Big)$$

$$\widetilde{a}_{Y,CCC} = \frac{1}{3}\left(\widetilde{a}_Y^{(1)} + \widetilde{a}_Y^{(2)} + \widetilde{a}_Y^{(3)}\right) = \frac{1}{3}\Big(\underbrace{\widetilde{b}_C^{(1)} + \widetilde{b}_D^{(1)} + \widetilde{b}_E^{(1)}}_{\widetilde{a}_Y^{(1)}} + \widehat{a}_Y + \underbrace{\widehat{b}_C + \widehat{b}_D + \widehat{b}_E}_{\widetilde{a}_Y^{(3)}}\Big)$$

$$\widetilde{b}_{A,CCC} = \frac{1}{3}\left(\widetilde{b}_A^{(1)} + \widetilde{b}_A^{(2)} + \widehat{b}_A\right), \quad \widetilde{b}_{B,CCC} = \frac{1}{3}\left(\widetilde{b}_B^{(1)} + \widetilde{b}_B^{(2)} + \widehat{b}_B\right),$$

$$\widetilde{b}_{C,CCC} = \frac{1}{3}\left(\widetilde{b}_C^{(1)} + \widetilde{b}_C^{(2)} + \widehat{b}_C\right), \quad \widetilde{b}_{D,CCC} = \frac{1}{3}\left(\widetilde{b}_D^{(1)} + \widetilde{b}_D^{(2)} + \widehat{b}_D\right),$$
$$\widetilde{b}_{E,CCC} = \frac{1}{3}\left(\widetilde{b}_E^{(1)} + \widetilde{b}_E^{(2)} + \widehat{b}_E\right).$$

It is worth noting that the top-level CCC reconciled forecast, $\widetilde{a}_{T,CCC}$, is given by the simple average of the direct base forecast $\widehat{a}_T$, and of the two indirect forecasts obtained by summing the base forecasts at the intermediate ($\widetilde{a}_T^{(2)} = \widehat{a}_X + \widehat{a}_Y$), and at the bottom ($\widetilde{a}_T^{(3)} = \widehat{b}_A + \widehat{b}_B + \widehat{b}_C + \widehat{b}_D + \widehat{b}_E$) levels, respectively, without any use of the uncertainty associated to the bts base forecasts. This information is instead taken into account when computing the CCC reconciled forecasts of variables X and Y, through the L_1CC reconciled forecasts $\widetilde{\boldsymbol{b}}^{(1)} = \left[\widetilde{b}_A^{(1)} \;\; \widetilde{b}_B^{(1)} \;\; \widetilde{b}_C^{(1)} \;\; \widetilde{b}_D^{(1)} \;\; \widetilde{b}_E^{(1)}\right]'$.

3.4 L_lCC reconciliation with endogenous constraints

In summary, the approach by Hollyman *et al.* (2021) consists in (i) decomposing the hierarchical/grouped structure in a sequence of elementary hierarchies, and (ii) for each elementary hierarchy, the bts base forecasts are reconciled according to a *level conditional coherent* procedure with exogenous constraints. Furthermore, in the approach suggested by Hollyman *et al.* (2021), only weights given by the variances of the base forecasts are considered, discarding - as it is often found in the forecast combination literature (Bates and Granger, 1969) - the covariances between couples of forecasts.

In the following we broaden the perspective, by relaxing the assumption of exogenous constraints in the optimization program (3.6). We consider a level conditional reconciliation procedure with exact linear *homogeneous* constraints, which means that level-l reconciled forecasts are computed by transforming the base forecasts looking for internal coherency of the targets $\boldsymbol{a}_l$ and $\boldsymbol{b}$, without imposing the external constraint of the uts base forecasts, as in the approach described in Section 3.3. We thus qualify this approach as *endogenous*.

Denote with $\boldsymbol{y}_l = \left[\boldsymbol{a}_l' \quad \boldsymbol{b}'\right]'$ the $[(n_b + n_l) \times 1]$ vector of the target forecasts of the level-l series, $\boldsymbol{a}_l$, and of the bts, $\boldsymbol{b}$. The corresponding base forecasts form the vector $\widehat{\boldsymbol{y}}_l = \left[\widehat{\boldsymbol{a}}_l \quad \widehat{\boldsymbol{b}}\right]'$. Let $\boldsymbol{U}_l = [\boldsymbol{I}_{n_l} \quad -\boldsymbol{C}_l]$ be the $[n_l \times (n_b + n_l)]$ matrix of the homogeneous constraints valid for the level-l series:

$$\boldsymbol{U}_l'\boldsymbol{y}_l = \boldsymbol{0}_{n_l}, \quad l = 1, \ldots, L.$$

Notice that in this case, unlike the procedure by Hollyman *et al.* (2021), the uts forecasts $\boldsymbol{a}_l$ are themselves object of the reconciliation process. In order for this may to happen, we need to have weights for the base forecasts of the aggregated series. Thus, let $\boldsymbol{W}_l$ be the $[(n_l + n_b) \times (n_l + n_b))]$ matrix (assumed diagonal) containing the variances of

the $n_l + n_b$ base forecasts to be reconciled in coherence with the linear aggregation relationships between the bts and the level-l aggregated series. In this case the $n_l + n_b$ level-l reconciled forecasts are given by (Chapter 1):

$$\widetilde{\boldsymbol{y}}_l = \boldsymbol{M}_l \widehat{\boldsymbol{y}}, \quad \text{with} \quad \boldsymbol{M}_l = \boldsymbol{I}_{n_l+n_b} - \boldsymbol{W}_l \boldsymbol{U}_l \left(\boldsymbol{U}_l' \boldsymbol{W}_l \boldsymbol{U}_l\right)^{-1} \boldsymbol{U}_l', \quad l = 1, \ldots, L. \tag{3.12}$$

In order to express the complete $(n \times 1)$ vector of reconciled forecasts $\widetilde{\boldsymbol{y}}_{\text{en}}^{(l)}$, where the superscript $^{(l)}$ stands for the level at which the reconciliation is performed, and the subscript '$_{\text{en}}$' stands for 'endogenously constrained', it is sufficient to apply the structural sum matrix $\boldsymbol{S}$ to the vector formed by the bottom n_b items of vector $\widetilde{\boldsymbol{y}}_l$, that is the vector $\widetilde{\boldsymbol{b}}_{\text{en}}^{(l)}$ of the bts reconciled forecasts[2]: $\widetilde{\boldsymbol{y}}_{\text{en}}^{(l)} = \boldsymbol{S}\widetilde{\boldsymbol{b}}_{\text{en}}^{(l)}$, $l = 1, \ldots, L$. The CCC and LCC approaches with endogenous constraints can thus be expressed as with formulae (3.10) and (3.11), respectively:

$$\widetilde{\boldsymbol{y}}_{CCC_{en}} = \frac{1}{L+1} \sum_{l=1}^{L+1} \widetilde{\boldsymbol{y}}_{en}^{(l)}, \qquad \widetilde{\boldsymbol{y}}_{LCC_{en}} = \frac{1}{L} \sum_{l=1}^{L} \widetilde{\boldsymbol{y}}_{en}^{(l)}, \tag{3.13}$$

where $\widetilde{\boldsymbol{y}}_{en}^{(L+1)} = \boldsymbol{S}\widehat{\boldsymbol{b}}$ is the vector of the bottom-up forecasts[3].

It is instructive to consider what happens when reconciliating a very simple hierarchy of 3 series, with $T = X + Y$. In this case $l = 1$, $n_l = 1$, $n_b = 2$,

$$\boldsymbol{W}_1 = \begin{bmatrix} \sigma_T^2 & 0 & 0 \\ 0 & \sigma_X^2 & 0 \\ 0 & 0 & \sigma_Y^2 \end{bmatrix}, \quad \boldsymbol{U}_1' = \begin{bmatrix} 1 & -1 & -1 \end{bmatrix}, \quad \boldsymbol{W}_1\boldsymbol{U}_1 = \begin{bmatrix} \sigma_T^2 \\ -\sigma_X^2 \\ -\sigma_Y^2 \end{bmatrix},$$

and $\boldsymbol{U}_1'\boldsymbol{W}_1\boldsymbol{U}_1 = \sigma_T^2 + \sigma_X^2 + \sigma_Y^2$. After a bit of algebra, it is found that the reconciled forecasts are given by:

$$\begin{aligned} \widetilde{T}_{\text{en}} &= \widehat{T} - \frac{\sigma_T^2}{\sigma_T^2 + \sigma_X^2 + \sigma_Y^2} \left(\widehat{T} - \widehat{X} - \widehat{Y}\right), \\ \widetilde{X}_{\text{en}} &= \widehat{X} + \frac{\sigma_X^2}{\sigma_T^2 + \sigma_X^2 + \sigma_Y^2} \left(\widehat{T} - \widehat{X} - \widehat{Y}\right), \\ \widetilde{Y}_{\text{en}} &= \widehat{Y} + \frac{\sigma_Y^2}{\sigma_T^2 + \sigma_X^2 + \sigma_Y^2} \left(\widehat{T} - \widehat{X} - \widehat{Y}\right). \end{aligned}$$

This result has the usual 'linearly constrained least squares adjustment of noisy data' interpretation (Stone *et al.*, 1942; Byron, 1978), where the reconciled forecast is given by the algebraic sum of the base forecast and of a share of the discrepancy observed

[2]Formally, let $\boldsymbol{J}_l' = \begin{bmatrix} \boldsymbol{0}_{n_b \times n_l} & \boldsymbol{I}_{n_b} \end{bmatrix}$ be the $[n_b \times (n_l + n_b)]$ matrix that 'extracts' the last n_b rows from a $[(n_l + n_b) \times 1]$ vector, it is $\widetilde{\boldsymbol{b}}_{\text{en}}^{(l)} = \boldsymbol{J}_l'\widetilde{\boldsymbol{y}}_l$, $l = 1, \ldots, L$.

[3]We thank a reviewer for having pointed out that, in an elementary hierarchy, L_1CC_{en} and the optimal combination forecast reconciliation approach coincide, provided the same error covariance matrix is used. However, in a generic hierarchy, with intermediate levels between the top and bottom ones, optimal combination forecast reconciliation approaches give in general different reconciled forecasts from LCC_{en} and CCC_{en}.

in the base forecasts, where the share is proportional to the variance (uncertainty) of the forecast. Furthermore, in this case too it is possible to give the expressions above a forecast combination interpretation. For, each reconciled forecast, at both upper and bottom level, can be written as the combination of the 'direct' (i.e., base) forecast, and the 'indirect' (i.e., implicitly obtained using the accounting relationships) one:

$$\widetilde{T}_{\text{en}} = \left(\frac{\sigma_X^2+\sigma_Y^2}{\sigma_T^2+\sigma_X^2+\sigma_Y^2}\right)\widehat{T} + \left(\frac{\sigma_T^2}{\sigma_T^2+\sigma_X^2+\sigma_Y^2}\right)\left(\widehat{X}+\widehat{Y}\right),$$
$$\widetilde{X}_{\text{en}} = \left(\frac{\sigma_T^2+\sigma_Y^2}{\sigma_T^2+\sigma_X^2+\sigma_Y^2}\right)\widehat{X} + \left(\frac{\sigma_X^2}{\sigma_T^2+\sigma_X^2+\sigma_Y^2}\right)\left(\widehat{T}-\widehat{Y}\right),$$
$$\widetilde{Y}_{\text{en}} = \left(\frac{\sigma_T^2+\sigma_X^2}{\sigma_T^2+\sigma_X^2+\sigma_Y^2}\right)\widehat{Y} + \left(\frac{\sigma_Y^2}{\sigma_T^2+\sigma_X^2+\sigma_Y^2}\right)\left(\widehat{T}-\widehat{X}\right).$$

However, unlike what happens in the L_lCC forecast reconciliation with exogenous constraints, in this case all the base forecasts are 'revised' in view of all the variances, not only the bts ones. This can be appreciated by looking at the reconciliation formulae in the exogenous (L_1CC) case - that can be obtained also by rearranging expression (16) in Hollyman et al. (2021, p. 154) -, where σ_T^2 is not involved, and the top-level series base forecast $\hat{T}$ is not adjusted:

$$\widetilde{T}_{L_1CC} = \widehat{T},$$
$$\widetilde{X}_{L_1CC} = \frac{\sigma_Y^2}{\sigma_X^2+\sigma_Y^2}\widehat{X} + \frac{\sigma_X^2}{\sigma_X^2+\sigma_Y^2}\left(\widehat{T}-\widehat{Y}\right),$$
$$\widetilde{Y}_{L_1CC} = \frac{\sigma_X^2}{\sigma_X^2+\sigma_Y^2}\widehat{Y} + \frac{\sigma_Y^2}{\sigma_X^2+\sigma_Y^2}\left(\widehat{T}-\widehat{X}\right).$$

The development of an analogous result for the *toy example* considered in Figure 3.1, can be found in Appendix B.3 .

3.5 Empirical applications

To evaluate in a fair setting the performance of various cross-sectional forecast combination based point forecast reconciliation approaches *vis-à-vis* the state-of-the-art procedures, in this section we consider two forecasting experiments on datasets already used in the hierarchical forecasting literature:

- the monthly time series of the Australian Tourism Flows disaggregated by geographic divisions and purpose of travel, firstly studied by Wickramasuriya *et al.* (2019), and used by Hollyman *et al.* (2021) to evaluate their forecast combination based forecast reconciliation approaches;
- the quarterly time series of the Australian Gross Domestic Product (GDP) from Income and Expenditure sides (Athanasopoulos *et al.*, 2020).

Thanks to the data and codes made available by the previously cited authors, in both cases we have successfully replicated their results, and then we have re-analyzed the datasets using our newly proposed methods, in line with the need for reproducibility and replicability in forecasting research (Hyndman, 2010; Boylan *et al.*, 2015; Makridakis *et al.*, 2018) to ensure good science based on truthful claims, and as a driver of further discovery and innovation (Baker *et al.*, 2020).

3.5.1 Monthly Australian Tourism Demand reconciliation

We reprise and extend the forecasting experiment performed by Hollyman *et al.* (2021) on the Australian Tourism Demand dataset. We start by (i) re-assessing the results found by Hollyman *et al.* (2021) using the forecast accuracy evaluation approach recommended by Davydenko and Fildes (2013), and (ii) considering some non-negativity issues that emerge during both the base forecasting and the reconciliation phases of the analysis. On this latter point we note that, with the notable exceptions of Wickramasuriya *et al.* (2020) and Kourentzes and Athanasopoulos (2021), this issue is generally overlooked in the forecast reconciliation literature, even though it has not irrelevant implications as for the interpretation of the results (e.g., a negative forecast for tourism demand makes no sense). Possible, though not fully convincing, motivations for this are that, on the practical side, adopting non-negative forecast reconciliation procedures is perceived as computation burdensome, and on the theoretical side, assuring non-negativity does not preserve unbiasedness in the final non-negative reconciled forecasts (Ben Taieb and Koo, 2019; Wickramasuriya *et al.*, 2020; Wickramasuriya, 2021).

Performance measures for multiple comparisons

We evaluate the performance of multiple (say, $J > 1$) forecast reconciliation approaches through accuracy indices calculated on the forecast error

$$\hat{e}^{j}_{i,h,t} = y_{i,t+h} - \hat{y}^{j}_{i,h,t}, \quad i = 1, \dots, 525, \quad h = 1, \dots, 12, \quad j = 0, \dots, J, \quad t = 1, \dots, q_h,$$

where y and $\hat{y}$ are the observed and forecast values, respectively, i denotes the series ($i = 1, \dots, 221$, for the uts, $i = 222, \dots, 525$, for the bts), h denotes the forecast horizon, t is the forecast origin ($t = 1$ corresponds to 2005:12, $q_1 = 132, \dots, q_{12} = 121$), and $j = 0$ denotes the automatic ETS base forecasts. The accuracy across multiple series and forecast horizons is evaluated following Davydenko and Fildes (2013) (see also Athanasopoulos and Kourentzes, 2022), who recommend the use of a metric based on aggregating performance ratios across time series using the weighted geometric mean. We consider both the Average Relative Mean Absolute Error (AvgRelMAE), and the Average Relative Mean Square Error (AvgRelMSE), obtained by transforming MAE and MSE index, respectively. As the conclusions drawn from both indices are basically the

same, for space reason in the rest of the chapter we comment only on the AvgRelMSE index[4].

For a fixed series i, forecast origin h, and approach j, the MSE index is given by the average across all q_h forecast origins of the squared forecast errors:

$$\text{MSE}_{i,h}^{j} = \frac{1}{q_h}\sum_{t=1}^{q_h}\left(\hat{e}_{i,h,t}^{j}\right)^2, \quad i = 1,\dots,525, \quad h = 1,\dots,12, \quad j = 0,\dots,J. \tag{3.14}$$

The AvgRelMSE of a forecasting approach for a given horizon is the geometric mean across all 525 series of the MSE ratio over a benchmark given by the base, incoherent ETS forecasts:

$$\text{AvgRelMSE}_{h}^{j} = \left(\prod_{i=1}^{525}\text{rMSE}_{i,h}^{j}\right)^{\frac{1}{525}}, \; h = 1,\dots,12, \; j = 0,\dots,J, \tag{3.15}$$

where $\text{rMSE}_{i,h}^{j} = \dfrac{\text{MSE}_{i,h}^{j}}{\text{MSE}_{i,h}^{0}}$ is the relative MSE. If a forecast outperforms the base forecasts, then the AvgRelMSE becomes smaller than one and vice-versa, and the percentage improvement in accuracy over the benchmark can be calculated as $\left(1 - \text{AvgRelMSE}_{h}^{j}\right) \times 100$. Expression (3.15), which refers to all 525 time series, can be re-stated for (i) groups of variables (e.g., bts and uts), and (ii) multiple forecast horizons (e.g., h = 1:6, h = 1:12).

Furthermore, we use the non-parametric Friedman and the post-hoc 'Multiple Comparison with the Best' (MCB) Nemenyi tests (Koning *et al.*, 2005; Kourentzes and Athanasopoulos, 2019; Makridakis *et al.*, 2022) to establish if the forecasting performances of the considered approaches are significantly different.

The data and the original forecasting experiment of Hollyman *et al.* (2021)

We consider 228 monthly observations (Jan 1998 - Dec 2016) of the Australian touristic flows (*Visitor Nights*) measured by the public project "National Visitor Survey" (Wickramasuriya *et al.*, 2019). The time series dataset, called VN525, consists in a grouped time series obtained by combination of a hierarchy by geographical division (destination) with a classification by Purpose of Travel (PoT). The Australian total is thus disaggregated by States (7), Zones (27), and Regions (76). Nominally, the geographic hierarchy comprises 111 destinations. However, since 6 Zones consist of a single Region, the non-redundant knots of the hierarchy are 105 instead of 111. Thus, we face an 'unbalanced hierarchy' (see sec. 3.2). As PoT can assume 4 different values: holiday (Hol), visiting friends and relatives (Vis), business (Bus), and other (Oth), in the grouped series obtained by crossing geographic divisions and PoT, 24 knots (6 Zones

[4]The interested reader may find all the tables and graphs based on the AvgRelMAE index in the online appendix.

with a single region by each PoT category) are redundant. The 304 most disaggregated variables, when combined according to the considered classifications, and discarding the duplications, produce 221 upper time series. In summary, the dataset comprises 8 levels, with 304 bottom time series, and 221 upper time series (525 *unique* time series in all)[5].

Hollyman *et al.* (2021) have performed a rolling forecast experiment with fixed length (96 months) window, producing base forecasts with forecast horizons varying from 1 to 12 months. Exponential Smoothing models (ETS, Hyndman, 2008) have been used to compute the base forecasts of the 525 series according to the automatic default of the `R` package `forecast` (Hyndman *et al.*, 2023)[6]. Hollyman *et al.* (2021) claim that, in agreement with Wickramasuriya *et al.* (2019), ETS models produce "substantially more accurate base forecasts than ARIMA based models in this setting, therefore presenting a more challenging environment for forecast combination techniques we consider". We basically agree on this point, stressing however that the intrinsic non-negative nature of the variables under analysis would have been better taken into account by modeling the log-transformed data, possibly adopting the same strategy as in Wickramasuriya *et al.* (2020) for dealing with the null values present in the observations. Anyway, the focus of the chapter being on the insights of the proposal by Hollyman *et al.* (2021), and the potential of the new reconciliation approach, we decided to continue using base forecasts obtained by ETS models in the level of the variables, limiting ourselves to recognize and quantitatively assess the problem during the base forecasting phase of the experiment, postponing the non-negativity issues to the reconciliation phase, where possibly negative base forecasts will be reconciled through effective non-negative linearly constrained least squares procedures (Stellato *et al.*, 2023, 2020; Wickramasuriya *et al.*, 2020; Hyndman *et al.*, 2022; Girolimetto and Di Fonzo, 2023a).

Besides the (incoherent) base forecasts, Hollyman *et al.* (2021) have considered the reconciled forecasts produced by the following approaches[7]:

- Bottom Up (BU): forecasts obtained by simple summation of the automatic ETS base forecasts for the 304 most disaggregate series.
- Level 1 Coherent Combination (L_1CC): coherent forecasts with the base forecast of the top-level of the hierarchy (Total Australia), computed using expression (3.2).

[5]Details can be found in the online appendix.

[6]We acknowledge Ross Hollyman, who kindly made us available the Python scripts used in Hollyman *et al.* (2021), thus allowing us to fully reproduce their results through the `R` package `FoReco` (Girolimetto and Di Fonzo, 2023a).

[7]For homogeneity, we use the labels of the reconciliation approaches adopted in the rest of the chapter. The correspondence with those originally used by Hollyman *et al.* (2021) is as follows: $L_1CC \equiv$ TD, $wls \equiv$ WLSv, $shr \equiv$ MinTShrink, and $CCC_H \equiv$ CCC.

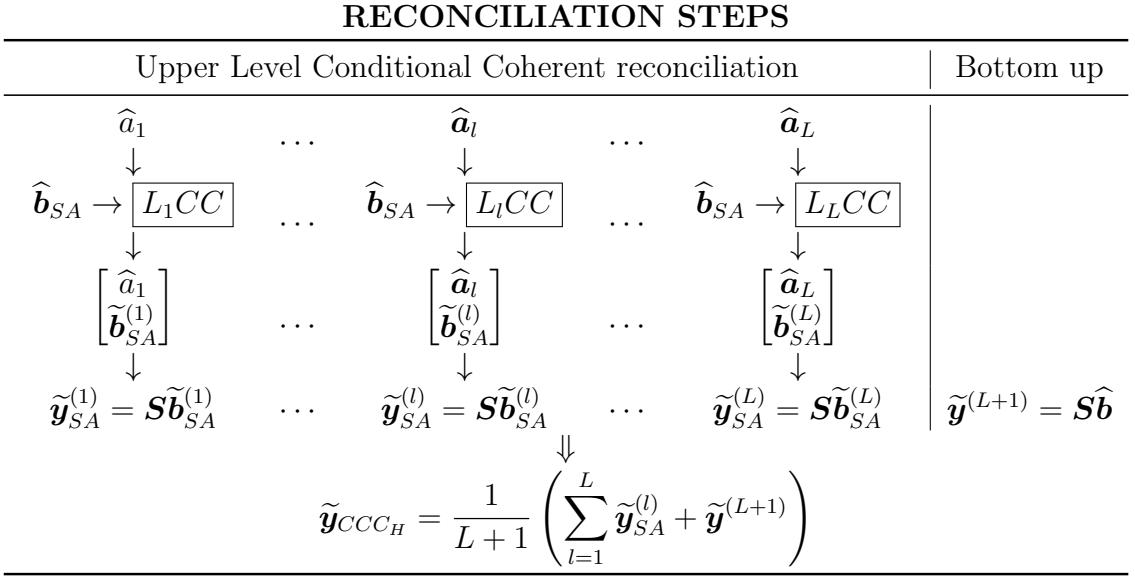


Figure 3.4: *CCC_H: Combined Conditional Coherent forecast reconciliation procedure according to Hollyman* et al. *(2021). In the Upper Level Conditional Coherent reconciliation steps the base forecasts $\widehat{\boldsymbol{b}}_{SA}$ are used, while in the bottom-up reconciliation $\widehat{\boldsymbol{b}}$ (automatic ETS) is used.*

- Top Down Historical Proportions (TDHP): forecasts obtained through a top-down reconciliation procedure using the "HP2" approach of Athanasopoulos *et al.* (2009), where the historical proportions disaggregation coefficients are computed on monthly basis.
- OLS: the original forecast reconciliation model of Hyndman *et al.* (2011), which assumes that the base forecasts are uncorrelated and identically distributed. It can be seen as a particular case of the Minimum Trace reconciliation approach of Wickramasuriya *et al.* (2019).
- *wls*: like OLS, with the error variances of each series taken into account; the forecast errors are assumed to be uncorrelated but heteroskedastic (Hyndman *et al.*, 2016).
- *shr*: the Minimum Trace optimal approach of Wickramasuriya *et al.* (2019) based on a shrinkage estimator of the covariance matrix of forecast errors.
- CCC_H: an equally weighted average of the reconciled forecasts derived from the 8 levels of the VN525 hierarchy. Unlike the *CCC* approach described in Figure 3.3, Hollyman *et al.* (2021) use different bottom time series base forecasts. The ETS forecasts ($\widehat{\boldsymbol{b}}$) are used in the bottom-up step, while seasonal averages of the observations ($\widehat{\boldsymbol{b}}_{SA}$) are used in the other reconciliation steps (see Figure 3.4, details in Hollyman *et al.*, 2021). A comment on this approach can be found in Section 3.5.1.

In most replications of both base forecasting and reconciliation phases of the experiment, negative forecasts have been produced. In 114 out of 132 replications of the experiment, and for a maximum of 8 series in the same replication, at least one automatic ETS base forecast was negative. The reconciliation phase seems to somehow worsen this issue: the OLS approach always produces a few negative reconciled forecasts (ranging from 15 to 119 series at each replication), while this phenomenon, though still not negligible, is less present in the remaining cases. It is worth noting that the L_1CC approach gives negative forecasts in more than 63% of the replications (84 out of 132), up to a maximum of about 9% of series (47 out of 525). As expected, negative forecasts are mostly present at the most disaggregate level (L8: Regions by PoT), with a less pronounced intensity of the phenomenon for the upper levels of the hierarchy (details can be found in the online appendix). In order to guarantee comparability with the results of Hollyman *et al.* (2021), Table 3.1 shows the AvgRelMSE's for the approaches considered in their paper, without any treatment of the negative values produced by the forecasting experiment.

Rather than considering all the 8 levels of the hierarchy[8], we present a more aggregated articulation of the results, by keeping distinct only all, upper, and bottom time series. The AvgRelMSE accuracy indices confirm the good performance of the CCC_H approach, as stated by Hollyman *et al.* (2021). However, it should be noted that CCC_H makes use of two different bts base forecasts (automatic ETS and Seasonal Averages), with the risk of an unfair comparison with *wls* and *shr*: we will come back on this point later. We add that the overview of the accuracy results does not change when non-negative reconciliation is performed. In fact, the results obtained by using the non-negative reconciliation facilities of the `R` package `FoReco` (Girolimetto and Di Fonzo, 2023a; see also Wickramasuriya *et al.*, 2020, and Hyndman *et al.*, 2022), not reported for space reasons but available in the online appendix, show that the overall accuracy slightly improves, mainly due to the benefit gained by the forecasts of the many intermittent series at the most disaggregated level, whose unconstrained forecasts often presented negative values. For this reason, and to deal with a more realistic operational context, the extended analysis in the following subsection will be performed by always considering the non-negative variant of the considered forecast reconciliation procedures.

In addition, due to their poor performance, and in order to set a more challenging empirical comparison, in the following we will not consider OLS and TDHP approaches

[8]The detailed results for all levels, here not presented for space reasons, are available in the online appendix.

	Forecast horizon						
Approach	1	2	3	6	12	1:6	1:12
			all (525 series)				
BU	0.9974	0.99342	0.9924	0.9976	1.0010	0.9956	0.9985
L_1CC	1.0032	1.0041	1.0031	0.9966	0.9797	1.0004	0.9935
TDHP	1.0055	1.0070	1.0059	0.9980	0.9785	1.0027	0.9944
OLS	1.0740	1.0748	1.0781	1.0730	1.0991	1.0757	1.0790
wls	0.9806	0.9805	0.9809	0.9816	0.9837	0.9809	0.9818
shr	**0.9745**	**0.9761**	0.9760	0.9758	0.9783	0.9757	0.9763
CCC_H	0.9764	0.9765	**0.9759**	**0.9726**	**0.9664**	**0.9743**	**0.9713**
			upper time series (221 series)				
BU	0.9939	0.9863	0.9821	0.9943	1.0024	0.9897	0.9965
L_1CC	0.9769	0.9790	0.9782	0.9753	0.9548	0.9765	0.9705
TDHP	0.9801	0.9826	0.9815	0.9768	0.9529	0.9792	0.9714
OLS	0.9997	0.9984	1.0025	0.9986	1.0141	1.0000	1.0015
wls	0.9614	0.9595	0.9595	0.9626	0.9676	0.9609	0.9633
shr	0.9537	0.9542	0.9534	0.9554	0.9609	0.9545	0.9563
CCC_H	**0.9476**	**0.9476**	**0.9470**	**0.9486**	**0.9451**	**0.9471**	**0.9468**
			bottom time series (304 series)				
BU	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
L_1CC	1.0227	1.0227	1.0217	1.0124	0.9982	1.0181	1.0105
TDHP	1.0244	1.0251	1.0240	1.0138	0.9976	1.0201	1.0115
OLS	1.1315	1.1339	1.1366	1.1306	1.1653	1.1343	1.1390
wls	0.9948	0.9960	0.9966	0.9956	0.9955	0.9957	0.9955
shr	**0.9899**	**0.9923**	**0.9927**	0.9909	0.9912	**0.9914**	0.9910
CCC_H	0.9978	0.9981	0.9974	**0.9905**	**0.9821**	0.9946	**0.9896**

Table 3.1: *Monthly forecasts reconciliation in the forecasting experiment on the Australian tourism dataset: AvgRelMSE of the approaches considered by Hollyman* et al. *(2021). Approach TDHP apart, some reconciled forecasts are negative. Bold entries identify the best performing approaches. Approaches performing worse than base forecasts are highlighted in red.*

any more, while the L_1CC approach will be only considered as one of the constituent parts of LCC and CCC approaches.

Forecast combination based forecast reconciliation using the same bts base forecasts

From Figure 3.4, which simply adapts the LCC and CCC reconciliation scheme in Figure 3.3, it appears that CCC_H approach is using two different bts base forecasts: in the Upper Level Conditional Coherent reconciliation steps, $\widehat{\boldsymbol{b}}_{SA}$ is used, while in the bottom-up reconciliation $\widehat{\boldsymbol{b}}$ (more precisely, $\widehat{\boldsymbol{b}}_{ETS}$) is used. It must be added that no particular saving of time is obtained by proceeding this way, since ETS base forecasts of all time series (both upper and bottom) are however calculated (and used). Rather, according to CCC_H the base forecasts to be used in each step are *de-facto* appropriately chosen.

In the light of above, we think that if one wishes to exploit both type of bts base forecasts, a sensible starting point would be considering the 'Seasonal-Averages-based'

LCC and CCC reconciled forecasts as well, obtained by simply substituting $\widehat{\boldsymbol{b}}$ with $\widehat{\boldsymbol{b}}_{SA}$ in the scheme of Figure 3.3. Denoting with $\widetilde{\boldsymbol{y}}_{SA}^{(l)} = \boldsymbol{S}\widetilde{\boldsymbol{b}}_{SA}^{(l)}$, $l = 1, \ldots, L$, the L_lCC reconciled forecasts using the seasonal averages as bts base forecasts, and with $\widetilde{\boldsymbol{y}}_{SA}^{(L+1)} = \boldsymbol{S}\widehat{\boldsymbol{b}}_{SA}$ the corresponding bottom-up reconciled forecasts[9], we obtain the 'SA' counterparts of the 'ETS-base-forecasts-based' reconciliation approaches described in Figure 3.3:

$$\widetilde{\boldsymbol{y}}_{LCC_{SA}} = \frac{1}{L}\sum_{l=1}^{L}\widetilde{\boldsymbol{y}}_{SA}^{(l)}, \qquad \widetilde{\boldsymbol{y}}_{CCC_{SA}} = \frac{1}{L+1}\sum_{l=1}^{L+1}\widetilde{\boldsymbol{y}}_{SA}^{(l)}. \tag{3.16}$$

In both cases, the benefit of the forecast combination based forecast reconciliation approaches LCC and CCC is clearly visible: on average across series and forecast horizons, LCC_{SA} reconciled forecasts are always the best performing ones when the seasonal averages are used as bts base forecasts (Table 3.2), while when using automatic ETS base forecasts, LCC_{ETS} scores best for all and uts series, while CCC_{ETS} 'wins' at the most disaggregated level (Table 3.3).

It must be added that overall the ETS-based LCC_{ETS} and CCC_{ETS} reconciliation perform better than using SA, particularly at the most disaggregated level (Region by PoT, i.e. bts), where the AvgRelMSE's are always less than one, unlike what happens for their SA-based counterparts, which show a stable decrease (AvgRelMSE > 1) of the forecast accuracy as compared to the benchmark.

To conclude on this point, we think that a 'fair' comparison of the LCC_{SA}/LCC_{ETS} and CCC_{SA}/CCC_{ETS} approaches with the optimal combination forecast reconciliation procedures wls and shr, may be established by considering the latter procedures when using either 'SA' or 'ETS' bts base forecasts. We call these specifications wls_{SA}/wls_{ETS}, and shr_{SA}/shr_{ETS}, respectively, which will be considered in the next sub-section.

Forecast averaging vs. pooling of reconciled forecasts

A simple, natural way to exploit the features of different bts base forecasts consists in considering a reconciled vector of forecasts obtained by averaging each approach using the bts base forecasts from both models (SA and ETS). This means to consider the vector of averaged forecasts (Abouarghoub *et al.*, 2018; Spiliotis *et al.*, 2019)

$$\widetilde{\boldsymbol{y}}_{\overline{\text{rec}}} = \frac{\widetilde{\boldsymbol{y}}_{\text{rec}_{SA}} + \widetilde{\boldsymbol{y}}_{\text{rec}_{ETS}}}{2}, \quad \text{rec} = wls, shr, CCC, LCC, CCC_{en}, LCC_{en}. \tag{3.17}$$

Coherently with this framework, the CCC_H approach can be seen as the forecast combination of the $2(L+1)$ forecasts $\widetilde{\boldsymbol{y}}_{ETS}^{(l)}$ and $\widetilde{\boldsymbol{y}}_{SA}^{(l)}$, $l = 1, \ldots, L+1$, where the L forecasts'

[9]When the seasonal averages of the training dataset are used as bts base forecasts, the bottom-up reconciled uts forecasts are equal to the seasonal averages of the uts in the training datasets. Put in other words, unlike the automatic ETS base forecasts, the SA base forecasts of all 525 series are trivially coherent, and need not to be reconciled.

	Forecast horizon						
Approach	1	2	3	6	12	1:6	1:12
			all (525 series)				
BU	1.0341	1.0345	1.0323	1.0249	1.0029	1.0295	1.0205
.....	L_lCC, *LCC* and *CCC* with exogenous constraints						.....
L_1CC	1.0025	1.0035	1.0026	0.9962	0.9791	0.9998	0.9930
L_2CC	1.0040	1.0021	1.0007	0.9922	0.9797	0.9982	0.9911
L_3CC	1.0145	1.0136	1.0132	1.0072	0.9958	1.0103	1.0047
L_4CC	1.0275	1.0296	1.0277	1.0259	1.0121	1.0270	1.0227
L_5CC	0.9931	0.9939	0.9932	0.9890	0.9769	0.9912	0.9865
L_6CC	1.0010	1.0017	1.0035	0.9976	1.0012	1.0003	0.9986
L_7CC	1.0097	1.0055	1.0063	1.0012	0.9946	1.0052	1.0017
LCC	**0.9857**	**0.9865**	**0.9861**	**0.9818**	**0.9736**	**0.9839**	**0.9801**
CCC	0.9881	0.9889	0.9884	0.9842	0.9753	0.9862	0.9823
.....	*LCC* and *CCC* with endogenous constraints						.....
LCC_{en}	1.0094	1.0101	1.0088	1.0033	0.9882	1.0064	1.0001
CCC_{en}	1.0121	1.0127	1.0114	1.0057	0.9898	1.0089	1.0024
			upper time series (221 series)				
BU	1.0349	1.0365	1.0331	1.0283	0.9980	1.0311	1.0213
.....	L_lCC, *LCC* and *CCC* with exogenous constraints						.....
L_1CC	0.9761	0.9783	0.9775	0.9749	0.9540	0.9758	0.9699
L_2CC	0.9794	0.9761	0.9741	0.9676	0.9561	0.9730	0.9668
L_3CC	0.9957	0.9934	0.9934	0.9923	0.9819	0.9917	0.9881
L_4CC	1.0085	1.0108	1.0091	1.0137	0.9988	1.0104	1.0085
L_5CC	0.9599	0.9616	0.9609	0.9621	0.9504	0.9607	0.9584
L_6CC	0.9742	0.9748	0.9782	0.9754	0.9928	0.9756	0.9783
L_7CC	0.9930	0.9857	0.9878	0.9865	0.9818	0.9884	0.9871
LCC	**0.9525**	**0.9535**	**0.9535**	**0.9539**	**0.9487**	**0.9527**	**0.9516**
CCC	0.9564	0.9577	0.9575	0.9581	0.9515	0.9567	0.9554
.....	*LCC* and *CCC* with endogenous constraints						.....
LCC_{en}	0.9931	0.9946	0.9929	0.9913	0.9734	0.9916	0.9864
CCC_{en}	0.9976	0.9992	0.9973	0.9954	0.9760	0.9959	0.9902
			bottom time series (304 series)				
BU	1.0335	1.0330	1.0318	1.0224	1.0065	1.0283	1.0200
.....	L_lCC, *LCC* and *CCC* with exogenous constraints						.....
L_1CC	1.0221	1.0223	1.0212	1.0120	0.9977	1.0177	1.0101
L_2CC	1.0223	1.0214	1.0205	1.0105	0.9972	1.0170	1.0092
L_3CC	1.0283	1.0286	1.0279	1.0181	1.0060	1.0240	1.0169
L_4CC	1.0416	1.0435	1.0414	1.0350	1.0218	1.0391	1.0331
L_5CC	1.0180	1.0181	1.0174	1.0090	0.9966	1.0140	1.0074
L_6CC	1.0209	1.0217	1.0223	1.0141	1.0074	1.0187	1.0137
L_7CC	1.0219	1.0202	1.0200	1.0120	1.0040	1.0176	1.0125
LCC	**1.0106**	**1.0112**	**1.0105**	**1.0026**	**0.9922**	**1.0072**	**1.0013**
CCC	1.0117	1.0122	1.0115	1.0036	0.9930	1.0083	1.0023
.....	*LCC* and *CCC* with endogenous constraints						.....
LCC_{en}	1.0215	1.0215	1.0206	1.0121	0.9991	1.0172	1.0102
CCC_{en}	1.0227	1.0227	1.0218	1.0132	0.9999	1.0184	1.0113

Table 3.2: *AvgRelMSE of LCC and CCC monthly forecast reconciliation approaches in the forecasting experiment on the Australian Tourism Demand dataset. LCC and CCC are defined by expressions (3.11) and (3.10), respectively. LCC_{en} and CCC_{en} are defined by expression (3.13). Seasonal averages of the training sets are used as bts base forecasts. BU identifies the bottom-up approach. Bold entries identify the best performing approaches. Approaches performing worse than base forecasts are highlighted in red.*

Approach	Forecast horizon 1	2	3	6	12	1:6	1:12
			all (525 series)				
BU	0.9972	0.9940	0.9923	0.9974	1.0008	0.9955	0.9983
.....	L_lCC, *LCC* and *CCC* with exogenous constraints						.....
L_1CC	0.9924	0.9910	0.9902	0.9939	0.9944	0.9922	0.9940
L_2CC	0.9920	0.9896	0.9885	0.9890	0.9934	0.9901	0.9912
L_3CC	0.9973	0.9950	0.9951	0.9958	0.9944	0.9954	0.9960
L_4CC	1.0107	1.0113	1.0112	1.0103	1.0048	1.0109	1.0095
L_5CC	0.9876	0.9877	0.9866	0.9908	0.9917	0.9886	0.9907
L_6CC	0.9957	0.9959	0.9984	0.9996	1.0084	0.9980	1.0004
L_7CC	0.9994	0.9951	0.9971	0.9953	0.9895	0.9967	0.9950
LCC	**0.9780**	**0.9772**	**0.9771**	**0.9785**	**0.9791**	**0.9777**	**0.9786**
CCC	0.9792	0.9781	0.9778	0.9797	0.9808	0.9787	0.9799
.....	*LCC* and *CCC* with endogenous constraints						.....
LCC_{en}	0.9868	0.9846	0.9835	0.9874	0.9905	0.9857	0.9880
CCC_{en}	0.9879	0.9855	0.9844	0.9884	0.9916	0.9866	0.9890
			upper time series (221 series)				
BU	0.9937	0.9861	0.9819	0.9941	1.0021	0.9894	0.9963
.....	L_lCC, *LCC* and *CCC* with exogenous constraints						.....
L_1CC	0.9823	0.9802	0.9783	0.9880	0.9911	0.9831	0.9883
L_2CC	0.9834	0.9787	0.9761	0.9794	0.9902	0.9802	0.9840
L_3CC	0.9900	0.9848	0.9848	0.9894	0.9912	0.9869	0.9897
L_4CC	1.0010	1.0012	1.0003	1.0033	0.9992	1.0019	1.0025
L_5CC	0.9738	0.9733	0.9707	0.9816	0.9859	0.9759	0.9817
L_6CC	0.9864	0.9860	0.9888	0.9930	1.0145	0.9895	0.9960
L_7CC	0.9947	0.9867	0.9893	0.9914	0.9866	0.9911	0.9912
LCC	**0.9605**	**0.9585**	**0.9579**	**0.9633**	**0.9684**	**0.9604**	**0.9640**
CCC	0.9626	0.9599	0.9590	0.9652	0.9708	0.9620	0.9661
.....	*LCC* and *CCC* with endogenous constraints						.....
LCC_{en}	0.9758	0.9703	0.9678	0.9771	0.9853	0.9731	0.9789
CCC_{en}	0.9777	0.9719	0.9692	0.9788	0.9871	0.9747	0.9807
			bottom time series (304 series)				
BU	0.9998	0.9998	0.9998	0.9999	0.9998	0.9998	0.9998
.....	L_lCC, *LCC* and *CCC* with exogenous constraints						.....
L_1CC	0.9998	0.9989	0.9989	0.9983	0.9967	0.9988	0.9981
L_2CC	0.9983	0.9975	0.9976	0.9961	0.9957	0.9973	0.9965
L_3CC	1.0026	1.0025	1.0026	1.0004	0.9968	1.0016	1.0005
L_4CC	1.0178	1.0187	1.0192	1.0154	1.0089	1.0175	1.0146
L_5CC	0.9978	0.9983	0.9983	0.9976	0.9960	0.9979	0.9973
L_6CC	1.0026	1.0032	1.0055	1.0045	1.0040	1.0043	1.0037
L_7CC	1.0028	1.0012	1.0029	0.9982	0.9916	1.0008	0.9977
LCC	**0.9910**	**0.9910**	**0.9913**	**0.9898**	**0.9871**	**0.9905**	**0.9893**
CCC	0.9915	0.9914	0.9917	0.9904	0.9881	0.9910	0.9900
.....	*LCC* and *CCC* with endogenous constraints						.....
LCC_{en}	0.9949	0.9951	0.9952	0.9949	0.9942	0.9949	0.9946
CCC_{en}	0.9954	0.9955	0.9956	0.9954	0.9948	0.9954	0.9951

Table 3.3: *AvgRelMSE of LCC and CCC monthly forecast reconciliation approaches in the forecasting experiment on the Australian Tourism Demand dataset. LCC and CCC are defined by expressions (3.11) and (3.10), respectively. LCC_{en} and CCC_{en} are defined by expression (3.13). Automatic ETS are used as bts base forecasts. BU identifies the bottom-up approach. Bold entries identify the best performing approaches. Approaches performing worse than base forecasts are highlighted in red.*

vectors $\widetilde{\boldsymbol{y}}_{ETS}^{(l)}$, $l = 1, \ldots, L$, and the single forecast vector $\widetilde{\boldsymbol{y}}_{SA}^{L+1}$, are given zero weights, while the remaining forecasts are equally weighted by $\dfrac{1}{L+1}$.

This choice may be seen as a forecast pooling (Hendry and Clements, 2004; Marcellino, 2004; Geweke and Amisano, 2011), a forecast combination procedure according to which from the complete set of forecasts, only a subset is deemed relevant to be combined. Kourentzes *et al.* (2019) investigate pooling for business forecasting, claiming "that the forecast selection criteria and the different approaches that are used to combine forecasts, can be considered as two independent types of operations, that follow pooling". In particular, "forecast selection and forecast combinations can be seen as two extremes of a spectrum that is defined by forecast pooling, combined with some selection/weighting operator"[10]. Aiolfi and Timmermann (2006) state that pooling can be beneficial, but recognise that the methods proposed depend on multiple subjective choices. We think that this is just the case of the CCC_H approach, which is grounded on a forecast pooling whose motivation might appear somehow subjective, therefore not immediately generalizable to different forecasting situations. Rather, simple complete combinations like either $\overline{CCC}$ or $\overline{LCC}$ seem to be very simple as well, able to exploit the diversity of forecasts coming from different models, and less prone to subjective choices. In addition, the comparison with the forecasting accuracy of $\overline{wls}$ and $\overline{shr}$ approaches seems to be logically well founded.

Table 3.4 sets out the AvgRelMSE for the reconciliation approaches described so far, distinct by all, upper and bottom variables, and grouped in such a way to highlight the common background in terms of the bts base forecasts used. For completeness, in Table 3.4 we consider the forecast accuracy (relative to the ETS base forecasts) of the seasonal averages as well, as these simple forecasts are at the same time cross-sectionally coherent, and part of the 'SA-based' forecast reconciliation approaches.

In summary, we find that:

- Forecast averaging of the LCC reconciled forecasts (i.e., $\overline{LCC}$) gives the best results for most forecast horizons: in 7 cases out 7 across all 525 series, and 3 out 7 across the 304 bottom series. In this last case, $\overline{CCC}$ ranks first, the differences with $\overline{LCC}$ being at the fourth decimal point.
- Forecast averaging of the optimal combination forecasts approaches ($\overline{shr}$ and $\overline{shr}$) always performs better than CCC_H at the most disaggregate level (304 bts series).

[10] An interesting Machine Learning approach to model selection in hierarchical forecasting has been recently proposed by Abolghasemi *et al.* (2022).

Approach**	Base forecasts* uts	bts	Forecast horizon 1	2	3	6	12	1:6	1:12
			all (525 series)						
SA	SA	SA	1.0341	1.0345	1.0323	1.0249	1.0029	1.0295	1.0205
wls_{SA}	ETS	SA	0.9892	0.9904	0.9902	0.9867	0.9809	0.9883	0.9853
shr_{SA}	ETS	SA	*0.9833*	*0.9855*	*0.9851*	*0.9812*	0.9754	*0.9831*	*0.9800*
CCC_{SA}	ETS	SA	0.9881	0.9889	0.9884	0.9842	0.9753	0.9862	0.9823
LCC_{SA}	ETS	SA	0.9857	0.9865	0.9861	0.9818	*0.9736*	0.9839	0.9801
$CCC_{en,SA}$	ETS	SA	1.0121	1.0127	1.0114	1.0057	0.9898	1.0089	1.0024
$LCC_{en,SA}$	ETS	SA	1.0094	1.0101	1.0088	1.0033	0.9882	1.0064	1.0001
wls_{ETS}	ETS	ETS	0.9801	0.9801	0.9802	0.9806	0.9816	0.9802	0.9806
shr_{ETS}	ETS	ETS	*0.9739*	*0.9757*	*0.9754*	*0.9748*	*0.9768*	*0.9750*	*0.9752*
CCC_{ETS}	ETS	ETS	0.9792	0.9781	0.9778	0.9797	0.9808	0.9787	0.9799
LCC_{ETS}	ETS	ETS	0.9780	0.9772	0.9771	0.9785	0.9791	0.9777	0.9786
$CCC_{en,ETS}$	ETS	ETS	0.9879	0.9855	0.9844	0.9884	0.9916	0.9866	0.9890
$LCC_{en,ETS}$	ETS	ETS	0.9868	0.9846	0.9835	0.9874	0.9905	0.9857	0.9880
$\overline{wls}$	ETS	SA & ETS	0.9715	0.9719	0.9719	0.9703	0.9685	0.9709	0.9697
$\overline{shr}$	ETS	SA & ETS	0.9698	0.9716	0.9711	0.9686	0.9671	0.9699	0.9684
CCC_H	ETS	SA & ETS	0.9757	0.9759	0.9753	0.9721	0.9659	0.9737	0.9708
$\overline{CCC}$	ETS	SA & ETS	0.9623	0.9622	0.9617	0.9608	0.9583	0.9612	0.9602
$\overline{LCC}$	ETS	SA & ETS	**0.9618**	**0.9618**	**0.9615**	**0.9602**	**0.9577**	**0.9608**	**0.9596**
$\overline{CCC}_{en}$	ETS	SA & ETS	0.9709	0.9702	0.9688	0.9688	0.9651	0.9690	0.9679
$\overline{LCC}_{en}$	ETS	SA & ETS	0.9700	0.9694	0.9681	0.9680	0.9646	0.9682	0.9671
			bottom time series (304 series)						
SA	SA	SA	1.0335	1.0330	1.0318	1.0224	1.0065	1.0283	1.0200
wls_{SA}	ETS	SA	1.0122	1.0128	1.0121	1.0049	0.9961	1.0093	1.0040
shr_{SA}	ETS	SA	*1.0094*	*1.0109*	1.0105	1.0034	0.9944	1.0075	1.0024
CCC_{SA}	ETS	SA	1.0117	1.0122	1.0115	1.0036	0.9930	1.0083	1.0023
LCC_{SA}	ETS	SA	1.0106	1.0112	*1.0105*	*1.0026*	*0.9922*	*1.0072*	*1.0013*
$CCC_{en,SA}$	ETS	SA	1.0227	1.0227	1.0218	1.0132	0.9999	1.0184	1.0113
$LCC_{en,SA}$	ETS	SA	1.0215	1.0215	1.0206	1.0121	0.9991	1.0172	1.0102
wls_{ETS}	ETS	ETS	0.9941	0.9953	0.9956	0.9937	0.9911	0.9945	0.9931
shr_{ETS}	ETS	ETS	*0.9892*	0.9917	0.9917	*0.9890*	0.9873	*0.9903*	0.9889
CCC_{ETS}	ETS	ETS	0.9915	0.9914	0.9917	0.9904	0.9881	0.9910	0.9900
LCC_{ETS}	ETS	ETS	0.9910	*0.9910*	*0.9913*	0.9898	*0.9871*	0.9905	*0.9893*
$CCC_{en,ETS}$	ETS	ETS	0.9954	0.9955	0.9956	0.9954	0.9948	0.9954	0.9951
$LCC_{en,ETS}$	ETS	ETS	0.9949	0.9951	0.9952	0.9949	0.9942	0.9949	0.9946
$\overline{wls}$	ETS	SA & ETS	0.9839	0.9845	0.9843	0.9797	0.9752	0.9824	0.9793
$\overline{shr}$	ETS	SA & ETS	0.9865	0.9880	0.9878	0.9827	0.9780	0.9856	0.9824
CCC_H	ETS	SA & ETS	0.9969	0.9972	0.9965	0.9897	0.9814	0.9938	0.9887
$\overline{CCC}$	ETS	SA & ETS	**0.9759**	**0.9761**	**0.9757**	0.9717	0.9669	**0.9740**	0.9711
$\overline{LCC}$	ETS	SA & ETS	0.9761	0.9762	0.9759	**0.9716**	**0.9667**	0.9741	**0.9710**
$\overline{CCC}_{en}$	ETS	SA & ETS	0.9784	0.9785	0.9777	0.9744	0.9701	0.9764	0.9737
$\overline{LCC}_{en}$	ETS	SA & ETS	0.9782	0.9782	0.9775	0.9741	0.9700	0.9762	0.9735

* SA: seasonal averages; ETS: automatic ETS forecasts.
** *LCC* and *CCC* are defined by expressions (3.11) and (3.10), respectively. LCC_{en} and CCC_{en} are defined by expression (3.13). $\overline{wls}, \overline{shr}, \overline{CCC}$, $\overline{LCC}$, $\overline{CCC}_{en}$, and $\overline{LCC}_{en}$ are defined according to expression (3.17). CCC_H: base forecasts as in Hollyman *et al.* (2021), see Figure 3.4.

Table 3.4: *AvgRelMSE of monthly reconciled forecasts in the forecasting experiment on the Australian Tourism Demand dataset. Optimal combination, LCC, and CCC reconciliation approaches, using seasonal averages and/or automatic ETS as bts base forecasts. Bold entries identify the best performing approaches independently of the base forecasts used. Italic entries identify the best performing approach using the same base forecasts (either SA or automatic ETS). Approaches performing worse than base forecasts are highlighted in red.*

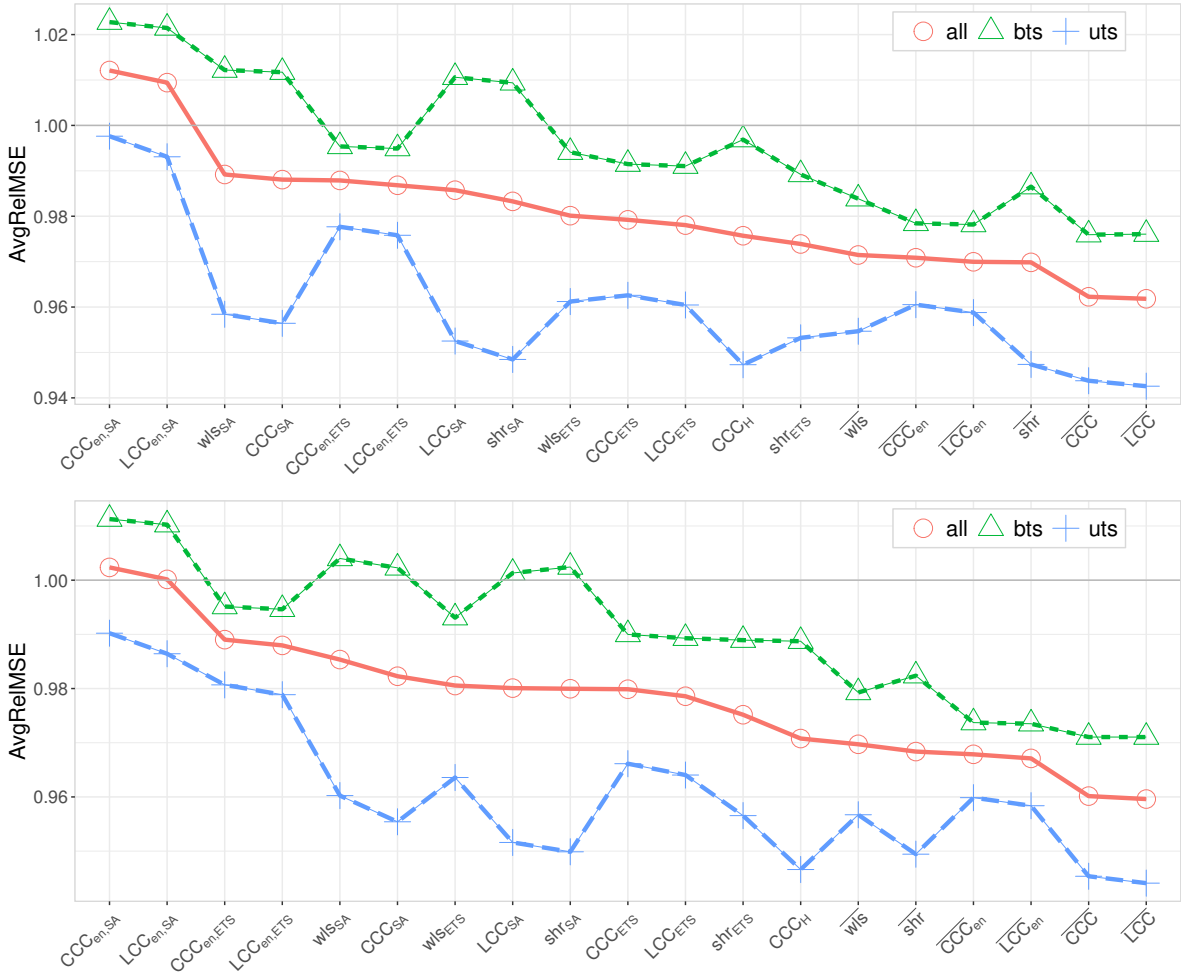
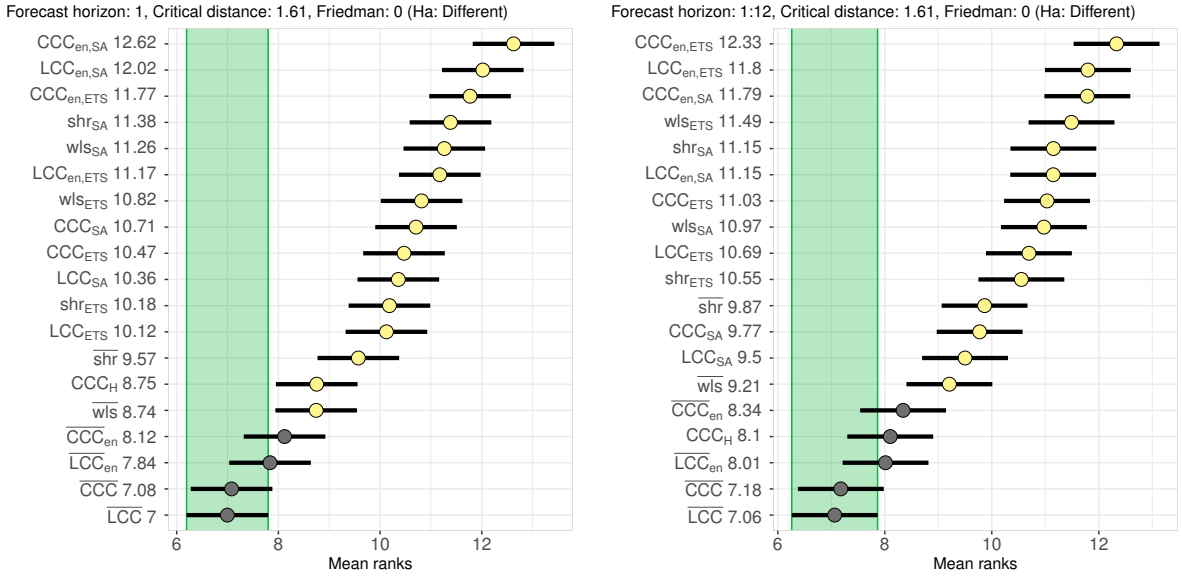


Figure 3.5: *Australian Tourism Demand dataset: AvgRelMSE of Optimal combination, LCC, and CCC reconciliation approaches, using seasonal averages and/or automatic ETS as bottom time series base forecasts. The acronyms for the considered approaches are described in footnote ** of Table 3.4. Top panel: forecast horizon h = 1. Bottom panel: forecast horizon h = 1:12.*

- When bts base forecasts from a single model (either SA or ETS) are used, *shr* performs on average best across all series. This result comes from a clear superiority for the most aggregated series in the upper levels of the hierarchy (details are reported in the online appendix), while at the bottom level the *shr* and *LCC* AvgRelMSE's look very similar.

Figure 3.5 gives an interesting visual summary of the results for two forecast horizons: $h = 1$ and $h = 1{:}12$. The 'SA-based' reconciliation approaches are located to the left part of the graphs, due to the low quality of the bts base forecasts. They are followed by the 'ETS-based' reconciled forecasts, which benefit from the higher quality of the bts base forecasts. The CCC_H approach clearly separates the single-model bts base forecasts from the multiple base forecasts combination based reconciliation, where the



(a) *forecast horizon $h = 1$* **(b)** *forecast horizon $h = 1{:}12$*

Figure 3.6: *Australian Tourism Demand dataset: MCB Nemenyi test results: average ranks and 95% confidence intervals for the 304 bts forecasts. The reconciliation approaches are sorted vertically according to the MSE mean rank for forecast horizon $h = 1$ (a), and forecast horizon $h = 1{:}12$ (b). The mean rank of each approach is displayed to the right of their names. The acronyms for the considered approaches are described in footnote ** of Table 3.4. If the intervals of two forecast reconciliation procedures do not overlap, this indicates a statistically different performance.*

forecast combination device of LCC and CCC approaches gives better results than the average of a couple of optimal combination reconciled forecasts.

Furthermore, it is confirmed and somehow reinforced the observation of Hollyman *et al.* (2021) that the gains from combination are particularly valuable when existing forecasts are of poor quality. Figure 3.6 shows the Multiple Comparison with the Best (MCB) Nemenyi test for the 304 bts, which form the noisiest section of the hierarchy. After that the Friedman test has shown that the considered forecasting approaches are different, $\overline{LCC}$, $\overline{CCC}$, $\overline{LCC}_{en}$, and $\overline{CCC}_{en}$ look (better than, and) significantly different from the other reconciliation forecasts at forecast horizon $h = 1$. These approaches still rank best for $h = 1{:}12$, with CCC_H forecasts resulting not significantly different from them.

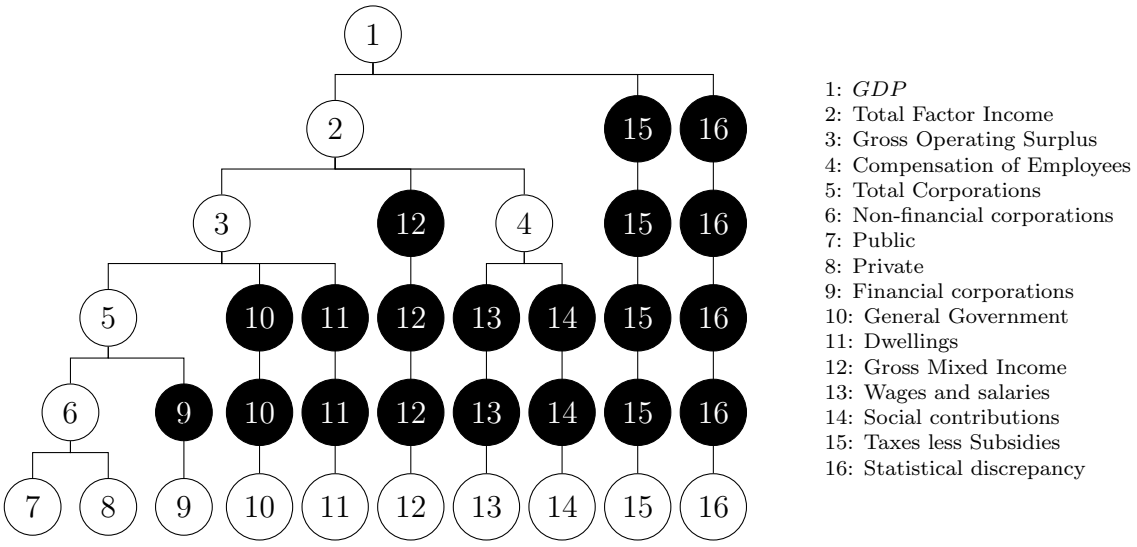


Figure 3.7: *Australian GDP from Income side: balanced hierarchical representations (duplicated series in black circles).*

3.5.2 Reconciliation of quarterly Australian *GDP* forecasts from Income and Expenditure sides

Athanasopoulos *et al.* (2020) consider 95 Australian Quarterly National Accounts time series, describing the Gross Domestic Product (*GDP*) - at current prices and not seasonally adjusted - from Income and Expenditure sides, interpreted as two distinct hierarchical structures. In the former case (Income), *GDP* is on the top of 15 lower level aggregates, while in the latter (Expenditure), *GDP* is the top level aggregate of a hierarchy of 79 time series (see figures 21.4-21.7 in Athanasopoulos *et al.* (2020), pp. 702-705). The available time series span over the period 1984:Q4 - 2018:Q1.

Our analysis is performed within the same forecasting experiment designed by Athanasopoulos *et al.* (2020). Base forecasts for the 95 separate time series have been obtained through simple univariate ARIMA models[11], selected using the `auto.arima` function of the `R`-package `forecast` (Hyndman *et al.*, 2023)[12]. Forecasts from $h = 1$ quarter ahead up to $h = 4$ quarters ahead are computed, using an *expanding* window, where the first training sample is set from 1984:Q4 to 1994:Q3, and forecasts are produced for 1994:Q4 to 1995:Q3. Then the training window is expanded by one quarter at a time, i.e. from 1984:Q4 to 2017:Q4, with the final forecasts produced by the last available observation in 2018:Q1. This leads to 94 1-step-ahead, 93 2-step-ahead, 92 3-step-ahead and 91 4-step-ahead forecasts available for evaluation.

[11]The `R` scripts, the data and the results of the paper by Athanasopoulos *et al.* (2020) are available in the github repository located at `https://github.com/PuwasalaG/Hierarchical-Book-Chapter`.

[12]Athanasopoulos *et al.* (2020) point out that this fast and flexible approach performs well in forecasting Australian *GDP* aggregates, even compared to other more complex methods.

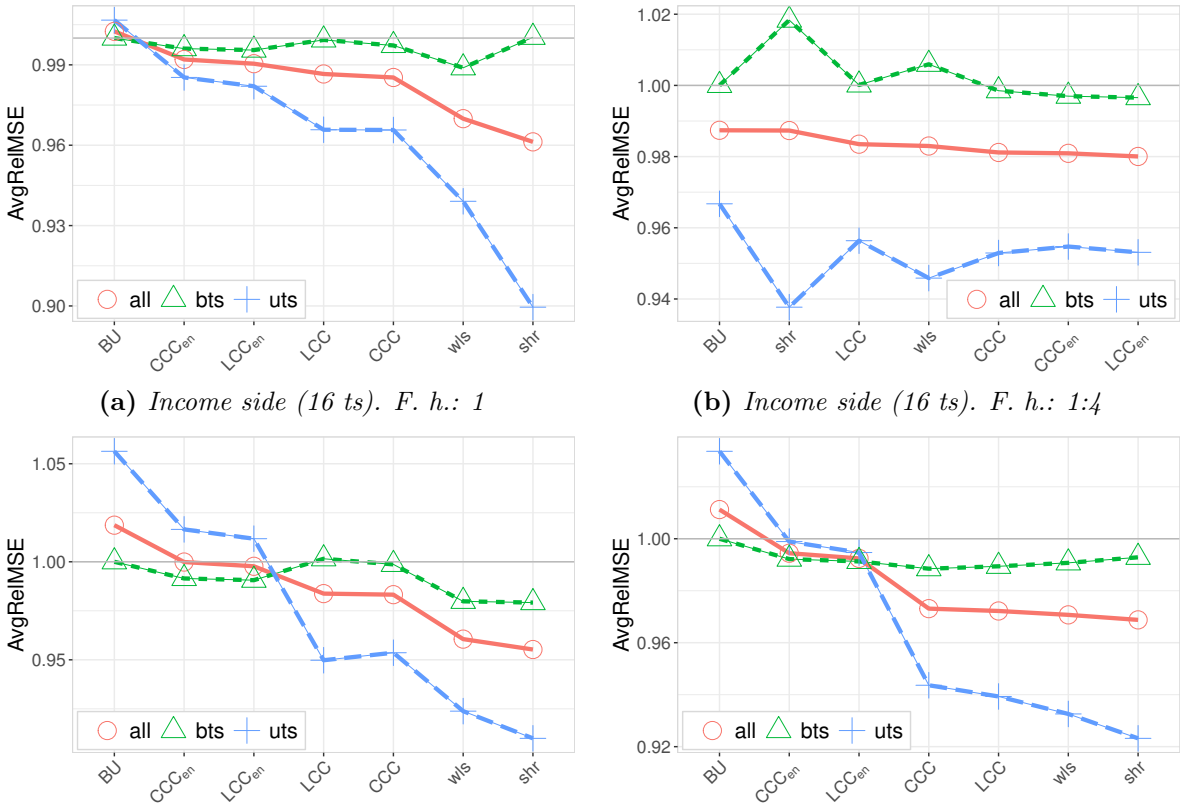
Approach	*Income side (16 series)* Forecast horizon 1	2	3	4	1:4	*Expenditure side (80 series)* Forecast horizon 1	2	3	4	1:4
Level-l Conditional Coherent reconciliation with exogenous constraints										
L_1CC	1.0532	1.0336	1.0396	1.0375	1.0409	1.0256	1.0270	1.0146	1.0223	1.0224
L_2CC	1.0681	1.0418	1.0596	1.0735	1.0607	1.0439	1.0307	1.0233	1.0228	1.0301
L_3CC	1.0323	1.0178	1.0040	1.0048	1.0146	1.0389	1.0051	0.9978	0.9969	1.0095
L_4CC	1.0263	0.9921	0.9614	0.9511	0.9823	1.0336	1.0194	0.9998	0.9973	1.0124
L_5CC	0.9651	0.9794	0.9836	0.9852	0.9783	1.0507	1.0253	1.0190	1.0109	1.0264
L_6CC						1.0166	1.0097	1.0069	1.0047	1.0095
L_7CC						1.0592	1.0356	1.0171	1.0106	1.0304
BU	1.0025	0.9796	0.9908	0.9769	0.9874	1.0187	1.0111	1.0097	1.0055	1.0112
LCC	0.9866	0.9799	0.9822	0.9853	0.9835	0.9838	0.9737	**0.9648**	**0.9667**	0.9722
CCC	0.9853	0.9767	**0.9810**	0.9817	0.9812	0.9832	0.9743	0.9664	0.9686	0.9731
LCC_{en}	0.9904	**0.9743**	0.9815	**0.9740**	**0.9800**	0.9977	0.9923	0.9900	0.9898	0.9924
CCC_{en}	0.9920	0.9748	0.9828	0.9742	0.9809	0.9999	0.9943	0.9921	0.9915	0.9944
wls	0.9699	0.9811	0.9883	0.9927	0.9830	0.9606	0.9691	0.9735	0.9798	0.9707
shr	**0.9613**	0.9832	1.0049	1.0005	0.9873	**0.9552**	**0.967**2	0.9715	0.9814	**0.9668**

Table 3.5: *AvgRelMSE of forecast combination based (LCC and CCC) and optimal combination (wls and shr) forecast reconciliation approaches in the forecasting experiment on the Australian GDP dataset. LCC and CCC are defined by expressions (3.11) and (3.10), respectively. LCC_{en} and CCC_{en} are defined by expression (3.13). Automatic ARIMA are used as bts base forecasts. BU identifies the bottom-up approach. Bold entries identify the best performing approaches within BU, LCC, CCC, LCC_{en}, CCC_{en}, wls and shr. Approaches performing worse than base forecasts are highlighted in red.*

The dataset and the forecasting experiment present some noteworthy features:

- the dataset consists of two distinct hierarchies, sharing only the top level series (GDP). The forecast combination based forecast reconciliation approach proposed by Hollyman *et al.* (2021) will thus produce two different reconciled forecasts of GDP, from Income and Expenditure sides, respectively[13];
- both Income and Expenditure hierarchies are unbalanced. To appropriately use LCC and CCC approaches, the two unbalanced hierarchies are transformed in a 6-level (Income, see Figure 3.7), and 8-level (Expenditure) balanced hierarchy, respectively;
- some of the 95 variables may take negative values (e.g., Statistical discrepancy), thus non-negativity of the reconciled forecasts is not a concern in this case;
- naive forecasts as seasonal averages or seasonal random walk have a very bad forecasting accuracy, that negatively affects the CCC_H results as well (details may be found in the online appendix). Thus in the following we focus only on the LCC and CCC approaches, with exogenous and endogenous constraints.

[13] Bisaglia *et al.* (2020) show how to obtain reconciled forecasts of GDP simultaneously coherent with all reconciled forecasts from both Expenditure and Income sides, by re-interpreting the two hierarchies as a more general linearly constrained multiple time series (Panagiotelis *et al.*, 2021).



(a) *Income side (16 ts). F. h.: 1*

(b) *Income side (16 ts). F. h.: 1:4*

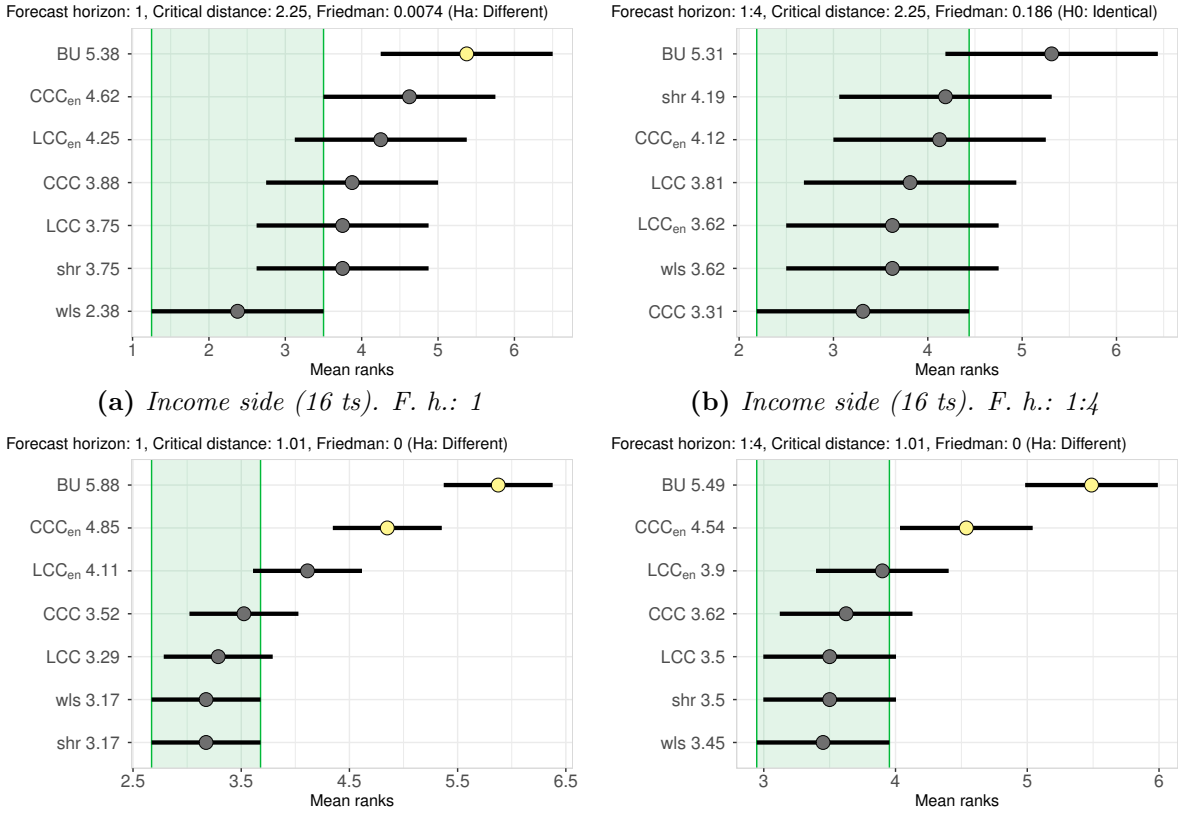
(c) *Expenditure side (80 ts). F. h.: 1*

(d) *Expenditure side (80 ts). F. h.: 1:4*

Figure 3.8: *Forecast reconciliation of quarterly Australian GDP. AvgRelMSE of LCC, CCC, and optimal combination reconciliation approaches, using automatic ARIMA as base forecasts.*

The base forecasts are reconciled using BU, LCC, CCC, LCC_{en}, CCC_{en}, wls and shr approaches. As already done in the previous empirical application, the forecasting accuracy is measured by the Average Relative Mean Squared Error (AvgRelMSE). The top panel of Table 3.5 reports also the AvgRelMSEs of the L_lCC reconciled forecasts, to evaluate the effect of forecast averaging on the quality of the LCC and CCC reconciled forecasts. The main findings can be summarized as follows:

- the accuracy gain of LCC and CCC over the level-l conditional reconciled forecasts is clearly visible for the Expenditure side series: 36 out of 40 AvgRelMSEs for L_lCC and BU are larger than 1, while their simple averages LCC and CCC always give accuracy indices less than 1. A less marked, however still evident, overall improvement of LCC and CCC over L_lCC and BU is registered for the Income side series as well. Similar results have been found for LCC_{en} and CCC_{en} approaches (see the online appendix);
- for both Income and Expenditure sides series, LCC, CCC, LCC_{en}, and CCC_{en} always improve the accuracy of the base forecasts (all AvgRelMSE < 1);



(a) *Income side (16 ts). F. h.: 1* **(b)** *Income side (16 ts). F. h.: 1:4*

(c) *Expenditure side (80 ts). F. h.: 1* **(d)** *Expenditure side (80 ts). F. h.: 1:4*

Figure 3.9: *MCB Nemenyi test results: average ranks and 95% confidence intervals for 1-step-ahead and 1:4-step-ahead quarterly Australian GDP forecasts. The reconciliation approaches are sorted vertically according to the MSE mean rank for Income (a) and Expenditure (b) sides, respectively. The mean rank of each approach is displayed to the right of their names. If the intervals of two forecast reconciliation procedures do not overlap, this indicates a statistically different performance.*

- the optimal combination reconciliation approach *shr* produces the smallest AvgRelMSEs at forecast horizon 1 for both Income and Expenditure sides series. However, in two cases (Income side, forecasting horizons 3 and 4), the *shr* reconciliation worsens the base forecasts' accuracy.

A visual appreciation of these results is given by Figure 3.8, where the forecasting accuracy indices (at forecast horizons 1 and 1:4) are distinct by all, upper and bottom time series, respectively. These findings are finally confirmed by the MCB Nemeny test (Figure 3.9). Despite the simplicity of the reconciliation formulae, *LCC* and *CCC* always show a forecasting accuracy not significantly different from the state-of-the-art forecast reconciliation approaches *wls* and *shr*.

Chapter 4

General linearly constrained multiple time series reconciliation

4.1 Introduction

Starting from Hyndman *et al.* (2011), regression-based forecast reconciliation has become an hot topic in the forecasting literature (van Erven and Cugliari, 2015; Wickramasuriya *et al.*, 2019; Wickramasuriya, 2021; Panagiotelis *et al.*, 2021; Jeon *et al.*, 2019; Ben Taieb and Koo, 2019; Ben Taieb *et al.*, 2021; Panagiotelis *et al.*, 2023; Wickramasuriya, 2024). By forecast reconciliation, we mean a post-forecasting procedure in which previously (and however) generated incoherent predictions (called base forecasts) for all the components of a multiple time series are adjusted in order to fulfill some externally given linear constraints. These reconciled forecasts are said to be coherent. In many real world applications, forecasts of a large collection of time series have a natural organization according to a hierarchical structure. More precisely, a system is classified as hierarchical when series are created by aggregating others in a tree shape. When the system is formed by a unique tree, then the collection is called "hierarchical time series" (Hyndman *et al.*, 2011). On the other side, when various hierarchies share the same series at both the most aggregated and disaggregated levels (top and bottom level, respectively), we face a "grouped time series" (Hyndman *et al.*, 2016).

In the field of forecast reconciliation, the bottom-up and top-down approaches are among the earliest and best known ones. Bottom-up forecasting (Dunn *et al.*, 1976) uses only forecasts at the bottom level to obtain all the reconciled forecasts. In contrast, top-down forecasting (Gross and Sohl, 1990) uses only the forecasts at the highest aggregated level. Having observed that neither forecasting approach uses all the available information, Hyndman *et al.* (2011) developed a regression-based reconciliation

approach consisting in (i) forecasting all the series with no regard for the constraints, and (ii) using then a regression model to optimally combining these (base) forecasts in order to produce coherent forecasts. This approach has witnessed a continuous growth of the related literature (Hyndman *et al.*, 2016; Athanasopoulos *et al.*, 2020; Wickramasuriya, 2021; Panagiotelis *et al.*, 2023), that in most cases grounds on the so-called structural representation of a hierarchical/grouped time series (Athanasopoulos *et al.*, 2009), in which the variables are classified either bottom if they belong to the most disaggregated level, or upper if they are obtained by summing the lower levels' variables. In mathematical terms, upper and bottom variables are linked by an aggregation matrix, which describes how the upper series are obtained from the bottom ones (Hyndman *et al.*, 2011). This representation is directly related to the hierarchical structure, where the series are naturally classifiable and an aggregation matrix may be obtained with little effort. Theoretical aspects for point and probabilistic forecast reconciliation have been developed using a structural representation by Panagiotelis *et al.* (2021) and Panagiotelis *et al.* (2023), respectively (see also Wickramasuriya, 2021, 2024).

However, it can be shown (van Erven and Cugliari, 2015; Wickramasuriya *et al.*, 2019; Bisaglia *et al.*, 2020, and Chapter 1) that reconciled forecasts may be obtained as the solution to a linearly constrained quadratic optimization problem[1], that does not require any "upper and bottom" classification of the involved variables. This approach grounds on a zero-constrained representation of the linearly constrained multiple time series (Chapter 1) describing the relationships linking all the individual series in the system. For a genuine hierarchical/grouped time series, where the top and bottom level variables are uniquely identified, it is easy to express the relationship between structural and zero-constrained representations. Wickramasuriya *et al.* (2019) show that the structural and the corresponding projection reconciliation approaches produce the same final reconciled forecasts. Unfortunately, given a zero-constrained representation of a linearly constrained multiple time series, finding the corresponding structural representation is not trivial, and one objective of this chapter is to fill this gap.

Most of the forecast reconciliation approaches proposed in the literature refer to genuine hierarchical/grouped time series, that do not take into account the full spectrum of possible cases encountered in real-life situations. As pointed out by Panagiotelis *et al.* (2021), "*concepts such as coherence and reconciliation (...) require the data to have only two important characteristics: the first is that they are multivariate, and the second is that they adhere to linear constraints*". Using a novel geometric interpretation,

[1]This approach dates back to the seminal paper by Stone *et al.* (1942) on the least squares adjustment of noisy data with accounting constraints (Byron, 1978, 1979). Recent applications to the reconciliation of systems of seasonally adjusted time series are provided by Di Fonzo and Marini (2011, 2015).

Panagiotelis *et al.* (2021) develop definitions and a formulation for linearly constrained multiple time series within a general framework and not just for simple summation hierarchical relationships. However, their results still ground on the structural representation valid only for genuine hierarchical/grouped time series, and this also holds for the probabilistic forecast reconciliation approach developed by Panagiotelis *et al.* (2023). Nevertheless, the point we wish to stress here is that when working with general linear constraints and many variables, the interchangeability between structural and zero-constrained representations, easy to recover for a genuine hierarchical/grouped time series, is no more always straightforward.

In this chapter we address a number of open issues in point and probabilistic cross-sectional forecast reconciliation for general linearly constrained multiple time series. First, we introduce a general linearly constrained multiple time series by exploiting its analogies with an homogeneous linear system. Second, we show that the classical hierarchical representation for a multiple time series is a simple special case of the general representation. Third, we show that if it is possible to express a general linearly constrained multiple time series according to a "structural-like" representation, we can easily achieve the formulation for point and probabilistic regression-based reconciled forecasts using a linear combination matrix, with elements in $\mathbb{R}$, that is the natural extension of the aggregation matrix used in the structural reconciliation approach, with elements only in $\{0, 1\}$. When the distinction between bottom and upper variables is no longer meaningful, we adopt a classification involving free and constrained variables, respectively, and show how to obtain a structural-like representation, possibly using well known linear algebra techniques, such as the Reduced Row Echelon Form and the QR decomposition (Golub and Van Loan, 1996, Meyer, 2000).

The remainder of the chapter is structured as follows. In Section 4.2, we present the notation and the main results about the point forecast reconciliation for a genuine hierarchical/grouped time series and in the general case. We define the structural-like representation for a general linearly constrained multiple time series by distinguishing between free and constrained variables, instead of bottom and upper, and use this result in the probabilistic forecast reconciliation framework set out by Panagiotelis *et al.* (2023). In Section 4.3, we show how to obtain the linear combination matrix for the structural-like representation in a computationally efficient way[2]. Two empirical applications are presented in Section 4.4. First, we extend the forecast reconciliation experiment for the Australian *GDP* originally developed by Athanasopoulos *et al.* (2020) and

[2]The procedures used in this chapter are implemented in the `R` package `FoReco` (Girolimetto and Di Fonzo, 2023a).

Bisaglia *et al.* (2020), to obtain both point and probabilistic GDP forecasts simultaneously coherent with their disaggregate counterpart forecasts from income and expenditure sides. Second, point and probabilistic forecasts are obtained for the European Area GDP from income, expenditure and output sides, geographically disaggregated by 19 component countries, where the large number of series and constraints makes a full row rank zero-constraints matrix difficult to build.

4.2 Reconciliation of a linearly constrained multiple time series

Let $\boldsymbol{x}_t$ be a n-dimensional linearly constrained multiple time series such that all the values for $t = 1, ...$ (either observed or not) lie in the *coherent* linear subspace $\mathcal{S} \in \mathbb{R}^n$, $\boldsymbol{x}_t \in \mathcal{S}\ \forall t > 0$ (Panagiotelis *et al.*, 2021), which means that all linear constraints are satisfied at time t. These constraints can be represented through linear equations and grouped as a (rectangular) linear system. Following Meyer (2000) and Leon (2015), let $x_{1,t}, \ldots, x_{n,t}$ be the observations of n individual time series at a given time $t = 1, ..., T$. An homogeneous linear system of p equations in the n variables present in $\boldsymbol{x}_t$ can be written as

$$\begin{cases} \gamma_{1,1}\, x_{1,t} & + & \gamma_{1,2}\, x_{2,t} & + & \ldots & + & \gamma_{1,n}\, x_{n,t} & = & 0 \\ \vdots & & \vdots & & \vdots & & \vdots & \vdots & \\ \gamma_{p,1}\, x_{1,t} & + & \gamma_{p,2}\, x_{2,t} & + & \ldots & + & \gamma_{p,n}\, x_{n,t} & = & 0 \end{cases},$$

where the γ_{ij}'s are real-valued coefficients. This system can be expressed in matrix form as

$$\boldsymbol{\Gamma}\boldsymbol{x}_t = \mathbf{0}_{(p\times 1)}, \tag{4.1}$$

where $\boldsymbol{\Gamma} \in \mathbb{R}^{(p\times n)}$ is the coefficient matrix

$$\boldsymbol{\Gamma} = \begin{bmatrix} \gamma_{1,1} & \gamma_{1,2} & \cdots & \gamma_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{p,1} & \gamma_{p,2} & \cdots & \gamma_{p,n} \end{bmatrix}.$$

Expression (4.1) is called "zero-constrained representation" of a linearly constrained multiple time series (see Chapter 1).

4.2.1 Point reconciliation of a genuine hierarchical time series

In Figure 4.1 it is shown a simple, genuine hierarchical time series (Athanasopoulos *et al.*, 2009; Hyndman *et al.*, 2011; Athanasopoulos *et al.*, 2020; Hyndman and Athanasopoulos, 2021), that can be seen as a particular case of a linearly constrained multiple time series.

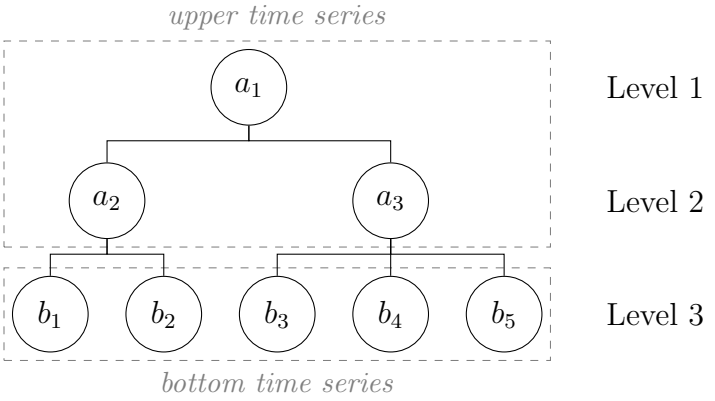


Figure 4.1: *A simple three-level hierarchical structure for a linearly constrained multiple time series*

This hierarchical structure is defined only by simple summation constraints,

$$\begin{array}{rcl} a_1 & = & b_1 + b_2 + b_3 + b_4 + b_5 \\ a_2 & = & b_1 + b_2 \\ a_3 & = & b_3 + b_4 + b_5 \end{array} , \tag{4.2}$$

that can be easily transformed into a zero-constrained representation $\boldsymbol{\Gamma x} = \mathbf{0}_{(8\times 1)}$, with $\boldsymbol{x} = [a_1\ a_2\ a_3\ b_1\ b_2\ b_3\ b_4\ b_5\ ]' = [\boldsymbol{a}'\quad \boldsymbol{b}']'$, and

$$\boldsymbol{\Gamma} = \begin{bmatrix} 1 & 0 & 0 & -1 & -1 & -1 & -1 & -1 \\ 0 & 1 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & -1 & -1 \end{bmatrix} = [\boldsymbol{I}_3 \quad -\boldsymbol{A}],$$

such as the $\gamma_{i,j}$'s coefficients are only -1, 0 and 1, and $\boldsymbol{a} = \boldsymbol{Ab}$.

In general, let $\boldsymbol{b}_t = \begin{bmatrix} b_{1,t} & \dots & b_{n_b,t} \end{bmatrix}' \in \mathbb{R}^{(n_b\times 1)}$ and $\boldsymbol{a}_t = \begin{bmatrix} a_{1,t} & \dots & a_{n_a,t} \end{bmatrix}' \in \mathbb{R}^{(n_a\times 1)}$, $t = 1, ..., T$, with $n = n_a + n_b$, be the T vectors containing the bottom and the aggregated series, respectively, of a hierarchy. All series can be collected in the T vectors

$$\boldsymbol{y}_t = \boldsymbol{P}\boldsymbol{x}_t = \begin{bmatrix} \boldsymbol{a}_t \\ \boldsymbol{b}_t \end{bmatrix} \in \mathbb{R}^{(n\times 1)},$$

where $\boldsymbol{P} \in \{0,1\}^{(n\times n)}$ is a permutation matrix used to appropriately re-order the original vector $\boldsymbol{x}_t$. If the classification as upper or bottom of the single time series in $\boldsymbol{x}_t$ is known in advance, we assume $\boldsymbol{x}_t = \boldsymbol{y}_t = [\boldsymbol{a}_t' \quad \boldsymbol{b}_t']'$, i.e. $\boldsymbol{P} = \boldsymbol{I}_n$ (no permutation of the original vector is needed). Moreover, also the linear combination (aggregation) matrix $\boldsymbol{A} \in \{0,1\}^{(n_a\times n_b)}$ describing the summation constraints linking the upper time series to the bottom ones, $\boldsymbol{a}_t = \boldsymbol{Ab}_t$, is assumed known in advance. Thus the "structural

representation" is simply given by (Athanasopoulos *et al.*, 2009)

$$\boldsymbol{y}_t = \boldsymbol{S}\boldsymbol{b}_t \qquad \text{with} \qquad \boldsymbol{S} = \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{I}_{n_b} \end{bmatrix}, \tag{4.3}$$

where $\boldsymbol{S} \in \{0,1\}^{(n \times n_a)}$ is the structural (summation) matrix.

Suppose now we have the vector $\widehat{\boldsymbol{y}}_h = \begin{bmatrix} \widehat{\boldsymbol{a}}_h' & \widehat{\boldsymbol{b}}_h' \end{bmatrix}' \in \mathbb{R}^{(n \times 1)}$ of unbiased and incoherent (i.e., $\widehat{\boldsymbol{y}}_h \neq \boldsymbol{S}\widehat{\boldsymbol{b}}_h$) base forecasts for the n variables of the linearly constrained series $\boldsymbol{y}_t$ for the forecast horizon h. Hyndman *et al.* (2011) use the structural representation (4.3) to obtain the reconciled forecasts $\widetilde{\boldsymbol{y}}_h$ as

$$\widetilde{\boldsymbol{y}}_h = \boldsymbol{S}\boldsymbol{G}\widehat{\boldsymbol{y}}_h, \quad \boldsymbol{G} = \left(\boldsymbol{S}'\boldsymbol{W}^{-1}\boldsymbol{S}\right)^{-1}\boldsymbol{S}'\boldsymbol{W}^{-1}, \tag{4.4}$$

where $\boldsymbol{W}$ is a $(n \times n)$ p.d. matrix assumed known and $\widehat{\boldsymbol{y}}_h$ ($\widetilde{\boldsymbol{y}}_h$) is the vector containing the base (reconciled) forecasts at forecast horizon h. Some alternative matrices $\boldsymbol{W}$ have been proposed in the literature for the cross-sectional forecast reconciliation case:

- identity (ols): $\widehat{\boldsymbol{W}}_{\text{ols}} = \boldsymbol{I}_n$ (Hyndman *et al.*, 2011),
- series variance (wls): $\widehat{\boldsymbol{W}}_{\text{wls}} = \boldsymbol{I}_n \odot \widehat{\boldsymbol{W}}_{\text{sam}}$ (Hyndman *et al.*, 2016),
- MinT-shr (shr): $\widehat{\boldsymbol{W}}_{\text{shr}} = \widehat{\lambda}\widehat{\boldsymbol{W}}_{\text{wls}} + (1- \widehat{\lambda})\widehat{\boldsymbol{W}}_{\text{sam}}$ (Wickramasuriya *et al.*, 2019),
- MinT-sam (sam): $\widehat{\boldsymbol{W}}_{\text{sam}} = \frac{1}{T}\sum_{t=1}^{T} \widehat{\boldsymbol{e}}_t\widehat{\boldsymbol{e}}_t'$ is the covariance matrix of the one-step ahead in-sample forecast errors $\widehat{\mathbf{e}}_t$ (Wickramasuriya *et al.*, 2019),

where the symbol $\odot$ denotes the Hadamard product, and $\widehat{\lambda}$ is an estimated shrinkage coefficient (Ledoit and Wolf, 2004).

The structural representation (4.3) may be transformed into the equivalent zero-constrained representation $\boldsymbol{a}_t - \boldsymbol{A}\boldsymbol{b}_t = \mathbf{0}_{(n_a \times 1)}$, that is (Wickramasuriya *et al.*, 2019)

$$\boldsymbol{C}\boldsymbol{y}_t = \mathbf{0}_{(n_a \times 1)} \qquad \text{with} \qquad \boldsymbol{C} = \begin{bmatrix} \boldsymbol{I}_{n_a} & -\boldsymbol{A} \end{bmatrix}. \tag{4.5}$$

$\boldsymbol{C} \in \{-1,0,1\}^{(n_a \times n)}$ is a full row-rank zero constraints matrix used to obtain the point reconciled forecasts according to the *projection approach* (van Erven and Cugliari, 2015; Wickramasuriya *et al.*, 2019, and see Chapter 1):

$$\widetilde{\boldsymbol{y}}_h = \boldsymbol{M}\widehat{\boldsymbol{y}}_h, \quad \boldsymbol{M} = \boldsymbol{I}_n - \boldsymbol{W}\boldsymbol{C}'\left(\boldsymbol{C}\boldsymbol{W}\boldsymbol{C}'\right)^{-1}\boldsymbol{C}. \tag{4.6}$$

Structural and zero-constrained representations can be interchangeable depending on which of the two is more convenient to use, allowing for greater flexibility in the calculation of the reconciled forecasts. For, the zero-constrained representation appears to be less computational intensive: in equation (4.4) two matrices must be inverted, one of size $(n \times n)$ and the other $(n_b \times n_b)$, whereas only the inversion of an $(n_a \times n_a)$ matrix is required in formula (4.6).

4.2.2 Zero-constrained and structural-like representations

In the hierarchical/grouped cross-sectional forecast reconciliation there is a natural distinction between upper and bottom time series that leads to the construction of the matrix $\boldsymbol{C}$ as in (4.5), where the first n_a columns refer to the upper and the remaining n_b $(= n - n_a)$ to the bottom variables, respectively. The time series in these two sets are categorized logically: all those related to the last level belong to the second group, the rest to the first. Most of the forecast reconciliation approaches proposed in the literature refer to genuine hierarchical/grouped time series and its structural representation. However, these do not take into account the full spectrum of possible cases encountered in real life situations. For a general linearly constrained multiple time series $(\boldsymbol{x}_t,\ t = 1, \dots, T)$, the classification between upper and bottom variables might not be meaningful, prompting us rather to look for two new sets: the **constrained** variables, denoted as $\boldsymbol{c}_t \in \mathbb{R}^{(n_c \times 1)}$, and the **free** (unconstrained) variables, denoted as $\boldsymbol{u}_t \in \mathbb{R}^{(n_u \times 1)}$, with $n = n_c + n_u$, such that $\boldsymbol{y}_t = \boldsymbol{P}\boldsymbol{x}_t = [\boldsymbol{c}_t' \quad \boldsymbol{u}_t']'$, and

$$\boldsymbol{c}_t = \boldsymbol{A}\boldsymbol{u}_t. \tag{4.7}$$

In this general framework, $\boldsymbol{A} \in \mathbb{R}^{(n_c \times n_u)}$ is the *linear combination* matrix associated to the linearly constrained multiple time series $\boldsymbol{x}_t = \boldsymbol{P}'\boldsymbol{y}_t$, that can be thus expressed *via* the "structural-like representation"

$$\boldsymbol{y}_t = \boldsymbol{P}\boldsymbol{x}_t = \begin{bmatrix} \boldsymbol{c}_t \\ \boldsymbol{u}_t \end{bmatrix} = \boldsymbol{S}\boldsymbol{u}_t \quad \text{with} \quad \boldsymbol{S} = \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{I}_{n_u} \end{bmatrix}, \tag{4.8}$$

where $\boldsymbol{S} \in \mathbb{R}^{(n_c \times n_u)}$ is the structural-like matrix[3]. It is worth noting that a full-rank zero-constraints matrix $\boldsymbol{C} \in \mathbb{R}^{(n_c \times n)}$ like in (4.5) may be easily obtained by (4.8) and used for the full-rank zero constrained representation $\boldsymbol{C}\boldsymbol{y}_t = \boldsymbol{C}\boldsymbol{P}\boldsymbol{x}_t = \boldsymbol{0}_{(n_c \times 1)}$. Therefore, even for a general, possibly not genuine hierarchical/grouped, linearly constrained multiple time series, either the structural (4.4) or the projection (4.6) approaches may be used to perform the point forecast reconciliation of incoherent base forecasts.

A simple example of a linearly constrained multiple time series that cannot be expressed as a genuine hierarchical/grouped structure is shown in Figure 4.2. In this case the variable X is at the top of two distinct hierarchies, that do not share the same bottom-level variables.

[3]Unlike the structural (summation) matrix in (4.3), that describes the simple summation relationships valid for genuine hierarchical/grouped time series, and has only elements in $\{0, 1\}$, in (4.8) $\boldsymbol{S}$ consists of real coefficients, appropriately organized to highlight the links between constrained and free variables.

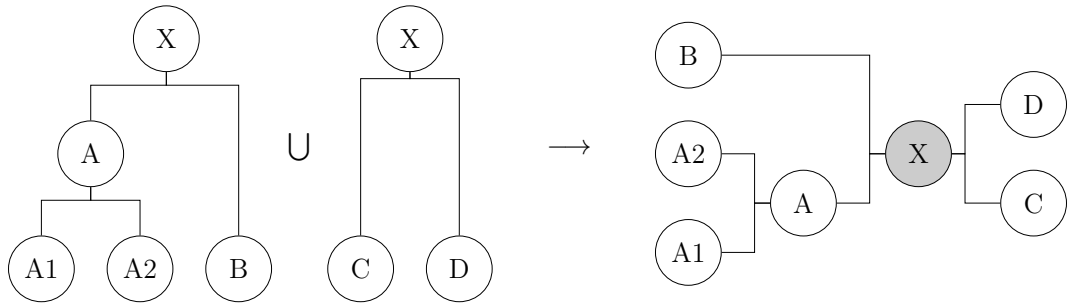


Figure 4.2: *A general linearly constrained structure: two hierarchies sharing only the same top-level series.*

The aggregation relationships between the upper variables X and A, and the bottom ones A1, A2, B, C, and D are given by:

$$\begin{aligned} X &= A1 + A2 + B \\ X &= C + D \\ A &= A1 + A2 \end{aligned} \quad . \tag{4.9}$$

In this case, both the zero-constrained and the structural-like representations are found in a rather straightforward manner. We consider $\{A2, B, C, D\}$ as free variables, such that $\boldsymbol{y}_t = \boldsymbol{x}_t$ (i.e., $\boldsymbol{P} = \boldsymbol{I}_4$), with $\boldsymbol{c}_t = \begin{bmatrix} x_{X,t} & x_{A,t} & x_{A1,t} \end{bmatrix}'$ and $\boldsymbol{u}_t = \begin{bmatrix} x_{A2,t} & x_{B,t} & x_{C,t} & x_{D,t} \end{bmatrix}'$. Thus, the coefficient matrix of the zero-constrained representation (4.1) is

$$\boldsymbol{\Gamma} = \begin{bmatrix} 1 & 0 & -1 & -1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & 1 & -1 & -1 & 0 & 0 & 0 \end{bmatrix}.$$

It is immediate to check that the system of linear constraints (4.9) may be re-written as

$$\begin{aligned} X &= C + D \\ A &= -B + C + D \\ A1 &= -A2 - B + C + D \end{aligned}$$

that is $\boldsymbol{C}\boldsymbol{x}_t = \boldsymbol{0}_{(3\times 1)}$, where

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{I}_3 & -\boldsymbol{A} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 1 & 1 & -1 & -1 \end{bmatrix} \quad \text{and} \quad \boldsymbol{A} = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}.$$

The structural-like representation of the general linearly constrained multiple time series in Figure 4.2 is then $\boldsymbol{y}_t = \boldsymbol{S}\boldsymbol{u}_t$, with $\boldsymbol{S} = [\boldsymbol{A}' \quad \boldsymbol{I}_3]'$.

It should be mentioned, however, that for medium/large systems (with many constraints and/or variables), manually operating on the constraints could be time-consuming and challenging. In Section 4.3 we will show a general technique to derive the linear combination matrix $\boldsymbol{A}$ from a general zero-constraints matrix $\boldsymbol{\Gamma}$.

4.2.3 Probabilistic reconciliation for general linearly constrained multiple time series

So far we have dealt with only point forecasting, but if one wishes to account for forecast uncertainty, probabilistic forecasts should be considered, as they - in the form of probability distributions over future quantities or events - measure the uncertainty in forecasts and are an important component of optimal decision making (Gneiting and Katzfuss, 2014).

Representation (4.8) states that $\boldsymbol{y}_t$ lies in an n-dimensional subspace of $\mathbb{R}^n$ spanned by the columns of $\boldsymbol{S}$, called "coherent subspace" and denoted by $\mathcal{S}$ (Panagiotelis *et al.*, 2023). Now, let $\mathscr{F}_{\mathbb{R}^{n_u}}$ be the Borel σ-algebra on $\mathbb{R}^{n_u}$, $(\mathbb{R}^{n_u}, \mathscr{F}_{\mathbb{R}^{n_u}}, \nu)$ a probability space for the free variables, and $s : \mathbb{R}^{n_u} \to \mathbb{R}^n$ a continuous mapping matrix. Then a σ-algebra $\mathscr{F}_{\mathcal{S}}$ can be constructed from the collection of sets $s(\mathcal{B})$ for all $\mathcal{B} \in \mathscr{F}_{\mathbb{R}^n}$.

Definition 4.1 (*Coherent probabilistic forecast for a linearly constrained multiple time series*)**.** Given the triple $(\mathbb{R}^{n_u}, \mathscr{F}_{\mathbb{R}^{n_u}}, \nu)$, we define a coherent probability triple $(\mathcal{S}, \mathscr{F}_{\mathcal{S}}, \breve{\nu})$ such that $\breve{\nu}(s(\mathcal{B})) = \nu(\mathcal{B})$, $\forall \mathcal{B} \in \mathscr{F}_{\mathbb{R}^{n_u}}$.

In order to extend forecast reconciliation to the probabilistic setting, let $(\mathbb{R}^n, \mathscr{F}_{\mathbb{R}^n}, \widehat{\nu})$ be a probability triple characterizing base (incoherent) probabilistic forecasts for all n series, and let $\psi : \mathbb{R}^{n_u} \to \mathbb{R}^n$ be a continuous mapping function defined by Panagiotelis *et al.* (2023) as the composition of two transformations, $s \circ g$, where $g : \mathbb{R}^n \to \mathbb{R}^{n_u}$ is a continuous function corresponding to matrix $\boldsymbol{G}$ in equation (4.4).

Definition 4.2 (*Probabilistic forecast reconciliation for a linearly constrained multiple time series*)**.** The reconciled probability measure of $\widehat{\nu}$ with respect to ψ is a probability measure $\widetilde{\nu}$ on $\mathcal{S}$ with σ-algebra $\mathscr{F}_{\mathcal{S}}$ such that

$$\widetilde{\nu}(\mathcal{A}) = \widehat{\nu}\left(\psi^{-1}(\mathcal{A})\right), \quad \forall \mathcal{A} \in \mathscr{F}_{\mathcal{S}}, \tag{4.10}$$

where $\psi^{-1}(\mathcal{A}) = \{\boldsymbol{x} \in \mathbb{R}^n : \psi(\boldsymbol{x}) \in \mathcal{A}\}$ is the pre-image of $\mathcal{A}$.

We consider two alternative approaches to deal with probabilistic forecast reconciliation according to the above definitions: a parametric framework, where probabilistic forecasts are produced under the assumption that the density function of the forecast errors is known, and a non-parametric framework, where no distributional assumption is made.

Parametric framework: Gaussian reconciliation

A reconciled probabilistic forecast may be obtained analytically for some parametric distributions, such as the multivariate normal (Yagli *et al.*, 2020; Corani *et al.*, 2021; Eckert *et al.*, 2021; Panagiotelis *et al.*, 2023; Wickramasuriya, 2024). In particular, if the base forecasts distribution is $\mathcal{N}(\widehat{\boldsymbol{y}}_h, \boldsymbol{W}_h)$, then the reconciled forecasts distribution is $\mathcal{N}(\widetilde{\boldsymbol{y}}_h, \widetilde{\boldsymbol{W}}_h)$, with

$$\widetilde{\boldsymbol{y}}_h = \boldsymbol{S}\,\boldsymbol{G}\widehat{\boldsymbol{y}}_h \quad \text{and} \quad \widetilde{\boldsymbol{W}}_h = \boldsymbol{S}\,\boldsymbol{G}\boldsymbol{W}_h\boldsymbol{G}'\,\boldsymbol{S}'.$$

The covariance matrix $\widetilde{\boldsymbol{W}}_h$ deserves special attention. In the simple case assumed by Wickramasuriya *et al.* (2019), $\boldsymbol{W}_h = k_h\boldsymbol{W}$, where k_h is a proportionality constant, and the reconciled covariance matrix reduces to (see Appendix C.1):

$$\widetilde{\boldsymbol{W}}_h = k_h\boldsymbol{S}\,\boldsymbol{G}\boldsymbol{W}. \tag{4.11}$$

However, the proportionality assumption along the forecast horizons h may be too restrictive, and computing k_h cannot be an easy task. Thus, three alternative formulations of $\boldsymbol{W}_h$, already shown in Section 4.2.1, have been proposed in the forecast reconciliation literature:

- $\boldsymbol{W}_h = \widehat{\boldsymbol{W}}_{\text{wls}}$ (Corani *et al.*, 2021; Panagiotelis *et al.*, 2023);
- $\boldsymbol{W}_h = \widehat{\boldsymbol{W}}_{\text{sam}}$ (Panagiotelis *et al.*, 2023);
- $\boldsymbol{W}_h = \widehat{\boldsymbol{W}}_{\text{shr}}$ (Athanasopoulos *et al.*, 2020).

Non-parametric framework: joint bootstrap-based reconciliation

When an analytical expression of the forecast distribution is either unavailable, or relies on unrealistic parametric assumptions, the empirical evaluation of the results may be grounded on reconciled samples (Jeon *et al.*, 2019; Yang, 2020; Panagiotelis *et al.*, 2023). At this end, we extend theorem 4.5 in Panagiotelis *et al.* (2023), originally formulated for genuine hierarchical/grouped time series, to the case of a general linearly constrained multiple time series. This theorem states that, if $\left(\widehat{\boldsymbol{y}}^{[1]}, \ldots, \widehat{\boldsymbol{y}}^{[L]}\right)$ is a sample drawn from an incoherent probability measure $\widehat{\nu}$, then $\left(\widetilde{\boldsymbol{y}}^{[1]}, \ldots, \widetilde{\boldsymbol{y}}^{[L]}\right)$, where $\widetilde{\boldsymbol{y}}^{[\ell]} := \psi\left(\widehat{\boldsymbol{y}}^{[\ell]}\right)$ for $\ell = 1, \ldots, L$, is a sample drawn from the reconciled probability measure $\widetilde{\nu}$ as defined in (4.10). According to this result, reconciling each member of a sample obtained from an incoherent distribution yields a sample from the reconciled distribution. As a consequence, coherent probabilistic forecasts may be developed through a post-forecasting mechanism analogous to the point forecast reconciliation setting. For this purpose, the bootstrap procedure by Athanasopoulos *et al.* (2020) is applied:

1. appropriate univariate models M_i for each series in the system are fitted based on the training data $\{y_{i,t}\}_{t=1}^{T}$, $i = 1, \ldots, n$, and the one-step-ahead in-sample forecast errors are stacked in an $(n \times T)$ matrix, $\widehat{\boldsymbol{E}} = \{\widehat{e}_{i,t}\}$;
2. $\widehat{\boldsymbol{y}}_{i,h}^{[l]} = f_i\left(M_i, \widehat{e}_{i,h}^{[l]}\right)$ is computed for $h = 1, \ldots, H$ and $l = 1, \ldots, L$, where $f(\cdot)$ is a function of the fitted univariate model and its associated error, $\widehat{\boldsymbol{y}}_{i,h}^{[l]}$ is a sample path simulated for the i-th series, and $\widehat{e}_{i,h}^{[l]}$ is the (i,h)-th element of an $(n \times H)$ block bootstrap matrix containing H consecutive columns randomly drawn from $\widehat{\boldsymbol{E}}$;
3. the optimal reconciliation formula, either according to the structural approach (4.4) or the projection approach (4.6), is applied to each $\widehat{\boldsymbol{y}}_{h}^{[l]}$.

4.3 Building the linear combination matrix A

In the previous section, we limited ourselves to introduce the linear combination matrix $\boldsymbol{A}$ in (4.7), in line with the novel classification distinguishing between free (unconstrained) and constrained variables. In this section we propose two ways of building such a matrix in practice.

First, consider the simple case where there are no redundant constraints ($n_c = p$) and the first n_c columns of $\boldsymbol{\Gamma}$ are linearly independent, so that $\boldsymbol{y}_t = \boldsymbol{x}_t = \begin{bmatrix}\boldsymbol{c}_t' & \boldsymbol{u}_t'\end{bmatrix}'$ and $\boldsymbol{\Gamma}\boldsymbol{y}_t = \boldsymbol{0}_{(n_c \times 1)}$. This homogeneous linear system can be written as

$$\boldsymbol{\Gamma}\boldsymbol{y}_t = \boldsymbol{0}_{(n_c \times 1)} \quad \Longleftrightarrow \quad \begin{bmatrix}\boldsymbol{\Gamma}_c & \boldsymbol{\Gamma}_u\end{bmatrix} \begin{bmatrix}\boldsymbol{c}_t \\ \boldsymbol{u}_t\end{bmatrix} = \boldsymbol{0}_{(n_c \times 1)},$$

where $\boldsymbol{\Gamma}_c \in \mathbb{R}^{(n_c \times n_c)}$ contains the coefficients for the constrained variables, and $\boldsymbol{\Gamma}_u \in \mathbb{R}^{(n_c \times n_u)}$ those for the free ones. Thanks to its non-singularity, $\boldsymbol{\Gamma}_c$ can be used to derive the equivalent zero-constrained representation:

$$\begin{bmatrix}\boldsymbol{I}_{n_c} & (\boldsymbol{\Gamma}_c)^{-1}\boldsymbol{\Gamma}_u\end{bmatrix} \begin{bmatrix}\boldsymbol{c}_t \\ \boldsymbol{u}_t\end{bmatrix} = \boldsymbol{C}\boldsymbol{y}_t = \boldsymbol{0}_{(n_c \times 1)}, \tag{4.12}$$

where

$$\boldsymbol{C} = \begin{bmatrix}\boldsymbol{I}_{n_c} & -\boldsymbol{A}\end{bmatrix} \qquad \text{and} \qquad \boldsymbol{A} = -(\boldsymbol{\Gamma}_c)^{-1}\boldsymbol{\Gamma}_u. \tag{4.13}$$

In practical situations, mostly if many variables and/or constraints are involved, categorizing variables as either constrained or free may be a challenging task[4]: the goal is to identify a valid set of free variables with invertible coefficient matrix $\boldsymbol{\Gamma}_c$.

[4]The issue of defining a valid set of free variables is studied by Zhang *et al.* (2023) in the framework of hierarchical/grouped reconciliation with immutable forecasts.

4.3.1 General (redundant) linear constraints framework

When n is large, it is not always immediate to find a set of non-redundant constraints, so the method shown in (4.12) may be hardly applied in several real-life situations. We propose to overcome these issues by employing standard linear algebra tools, like the Reduced Row Echelon Form or the QR decomposition (Golub and Van Loan, 1996, Meyer, 2000), that are able to deal with redundant constraints and do not request any a priori classification of the single variables entering the multiple time series.

Reduced Row Echelon Form (rref)

A matrix is said to be in rref (Meyer, 2000) if and only if the following three conditions hold:

- it is in row echelon form;
- the pivot in each row is 1;
- all entries above each pivot are 0.

The idea is then very simple: classify as constrained the variables corresponding to the pivot positions of the rref representation coefficient matrix, while the remaining columns form the linear combination matrix $\boldsymbol{A}$. Usually a rref form is obtained through a Gauss-Jordan elimination (more details in Meyer, 2000). So, let $\boldsymbol{Z} \in \mathbb{R}^{(n_c \times n)}$ be the rref of $\boldsymbol{\Gamma}$ deprived of any possible null rows, then a permutation matrix $\boldsymbol{P}$ can be obtained starting from the pivot columns of $\boldsymbol{Z}$, such that

$$\boldsymbol{y}_t = \boldsymbol{P}\boldsymbol{x}_t = \begin{bmatrix} x_{\pi_c(1),t} & \dots & x_{\pi_c(n_c),t} & x_{\pi_u(1),t} & \dots & x_{\pi_u(n_u),t} \end{bmatrix},$$

where $\pi_c(i)$, $i = 1, ..., n_c$, is the position of the i-th pivot column (i.e., one of the columuns that identify the constrained variables) and $\pi_u(j)$, $j = 1, ..., n_u$, is the position of the j-th no-pivot column (i.e., one of the columns associated to the free variables). Then, the linear combination matrix $\boldsymbol{A}$ can be extracted from the expression

$$\boldsymbol{C} = \boldsymbol{Z}\boldsymbol{P}' = \begin{bmatrix} \boldsymbol{I}_{n_c} & -\boldsymbol{A} \end{bmatrix}.$$

Additional examples can be found in the online appendix.

QR decomposition

Given the $(p \times n)$ coefficient matrix $\boldsymbol{\Gamma}$ of the zero-constrained representation (4.1), $\boldsymbol{\Gamma} = \boldsymbol{Q}\boldsymbol{R}\boldsymbol{P}$ is a QR decomposition with column pivoting (Lyche, 2020), where $\boldsymbol{Q} \in \mathbb{R}^{(p \times p)}$ is a square and orthonormal matrix ($\boldsymbol{Q}'\boldsymbol{Q} = \boldsymbol{Q}\boldsymbol{Q}' = \boldsymbol{I}_p$), $\boldsymbol{P} \in \{0, 1\}^{(n \times n)}$ is a permutation matrix, and $\boldsymbol{R} \in \mathbb{R}^{(p \times n)}$ is an upper trapezoidal matrix (Anderson *et al.*,

1992, 1999) such that

$$\boldsymbol{R} = \begin{cases} \begin{bmatrix} \boldsymbol{R}_c & \boldsymbol{R}_u \end{bmatrix} & \text{if } \boldsymbol{\Gamma} \text{ is full-rank} \\ \begin{bmatrix} \boldsymbol{R}_c & \boldsymbol{R}_u \\ \boldsymbol{0}_{[(p-n_c)\times n_c]} & \boldsymbol{0}_{[(p-n_c)\times n_u]} \end{bmatrix} & \text{if } \boldsymbol{\Gamma} \text{ is not full-rank} \end{cases},$$

where $\boldsymbol{R}_c \in \mathbb{R}^{(n_c \times n_c)}$ is upper triangular, and $\boldsymbol{R}_u \in \mathbb{R}^{(n_c \times n_u)}$ is nonsingular (Golub and Van Loan, 1996). Applying this decomposition to the homogenous system (4.1), we obtain

$$\boldsymbol{\Gamma x}_t = \boldsymbol{QRPx}_t = \boldsymbol{QRy}_t = \boldsymbol{0}_{(p\times 1)},$$

that is equivalent to (Meyer, 2000)

$$\begin{cases} \boldsymbol{Qz} = \boldsymbol{0}_{(p\times 1)} \\ \boldsymbol{Ry}_t = \boldsymbol{z} = \boldsymbol{0}_{(p\times 1)} \end{cases}.$$

Due to the non-singularity of $\boldsymbol{Q}$, $\boldsymbol{z} = \boldsymbol{0}_{(p\times 1)}$ is the unique solution to the homogenous system $\boldsymbol{Qx} = \boldsymbol{0}_{(p\times 1)}$ (Lyche, 2020). Then, the homogeneous system (4.1) representing a general linearly constrained time series may be re-written as[5] $\begin{bmatrix} \boldsymbol{R}_c & \boldsymbol{R}_u \end{bmatrix} \boldsymbol{y}_t = \boldsymbol{0}_{(n_c\times 1)}$. Finally, from (4.12) we obtain

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{I}_{n_c} & -\boldsymbol{A} \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{A} = -\boldsymbol{R}_c^{-1}\boldsymbol{R}_u,$$

where $\boldsymbol{R}_c$ is invertible by construction (Golub and Van Loan, 1996). It is worth noting that the Pivoted QR decomposition generates a permutation matrix $\boldsymbol{P}$ that "moves" the free variables in $\boldsymbol{x}_t$ to the bottom part of the re-ordered vector $\boldsymbol{y}_t$, that is $\boldsymbol{Px}_t = \boldsymbol{y}_t = \begin{bmatrix} \boldsymbol{c}_t' & \boldsymbol{u}_t' \end{bmatrix}'$.

It should be noted that in both cases (QR and rref), $\boldsymbol{P} = \boldsymbol{I}_n$ if the first n_c columns of $\boldsymbol{\Gamma}$ are linearly independent. This means that both algorithms start by assuming as constrained and free the variables as they appear in $\boldsymbol{x}$, whose ordering is then changed only if it is not feasible for the constraints operating on the multivariate time series in equation (4.1).

4.4 Empirical applications

In this section we present two macroeconomics applications involving general linearly constrained multiple time series which do not have a genuine hierarchical/grouped structure. In the first case, in the wake of Athanasopoulos *et al.* (2020), we forecast the Australian GDP from income and expenditure sides, for which Bisaglia *et al.* (2020)

[5]Possible null rows, present if $\boldsymbol{\Gamma}$ is not full-rank, are removed.

already provided a full row-rank $\boldsymbol{C}$ matrix. The second application concerns the European GDP disaggregated by three sides (income, expenditure and output) and 19 member countries. In this case, building a full row-rank zero constraints matrix is not an easy task, so we use a QR decomposition (Section 4.3.1). Detailed informations on the variables in either dataset are reported in the online appendix.

4.4.1 Reconciled probabilistic forecasts of the Australian GDP from income and expenditure sides

Athanasopoulos *et al.* (2020) first considered the reconciliation of point and probabilistic forecasts of the 95 Australian Quarterly National Accounts (QNA) variables that describe the Gross Domestic Product (GDP) at current prices from the income and expenditure sides, interpreted as two distinct hierarchies. In the former case (income), GDP is the top level aggregate of a hierarchy of 15 lower level aggregates with $n_a^I = 6$ and $n_b^I = 10$, whereas in the latter (expenditure), GDP is the top level aggregate of a hierarchy of 79 time series, with $n_a^E = 27$ and $n_b^E = 53$ (for details, Athanasopoulos *et al.*, 2020; Bisaglia *et al.*, 2020; Di Fonzo and Girolimetto, 2022b, and Chapter 1).

Considering these two hierarchies as distinct yields different GDP forecasts depending on the considered disaggregation (either by income or expenditure). The fact that the two hierarchical structures describing the National Accounts share only the same top-level series (GDP), prevents the adoption for the whole set of $n = 95$ distinct variables of the original structural reconciliation approach proposed by Hyndman *et al.* (2011). However, it is possible to use the results shown so far for a general linearly constrained multiple time series. The homogeneous constraints valid for the variables are described by the following (33×95) matrix $\boldsymbol{\Gamma}$ (Bisaglia *et al.*, 2020):

$$\boldsymbol{\Gamma} = \begin{bmatrix} 1 & \mathbf{0}_5' & -\mathbf{1}_{10}' & \mathbf{0}_{26}' & \mathbf{0}_{53}' \\ 1 & \mathbf{0}_5' & \mathbf{0}_{10}' & \mathbf{0}_{26}' & -\mathbf{1}_{53}' \\ \mathbf{0}_5 & \mathbf{I}_5 & -\boldsymbol{A}^I & \mathbf{0}_{5\times 26} & \mathbf{0}_{5\times 53} \\ \mathbf{0}_{26} & \mathbf{0}_{26\times 5} & \mathbf{0}_{26\times 10} & \mathbf{I}_{26} & -\boldsymbol{A}^E \end{bmatrix}, \tag{4.14}$$

where $\boldsymbol{A}^I \in \{0,1\}^{(5\times 10)}$ and $\boldsymbol{A}^E \in \{0,1\}^{(26\times 53)}$ are the aggregation matrices for the income and the expenditure sides, respectively, and $\boldsymbol{\Gamma}$ has already full row-rank. A structural-like representation of the multiple time series that incorporates both sides' accounting constraints may be obtained by transforming $\boldsymbol{\Gamma}$ through, for example, the QR technique described in Section 4.3.1. This operation results in a (33×95) matrix $\boldsymbol{C} = [\boldsymbol{I}_{33} \; -\boldsymbol{A}]$, where $\boldsymbol{A}$ is the (33×62) linear combination matrix shown in Figure 4.3, and $\boldsymbol{S} = [\boldsymbol{A}' \quad \boldsymbol{I}_{62}]'$ is the structural-like matrix (see Section 4.2.2).

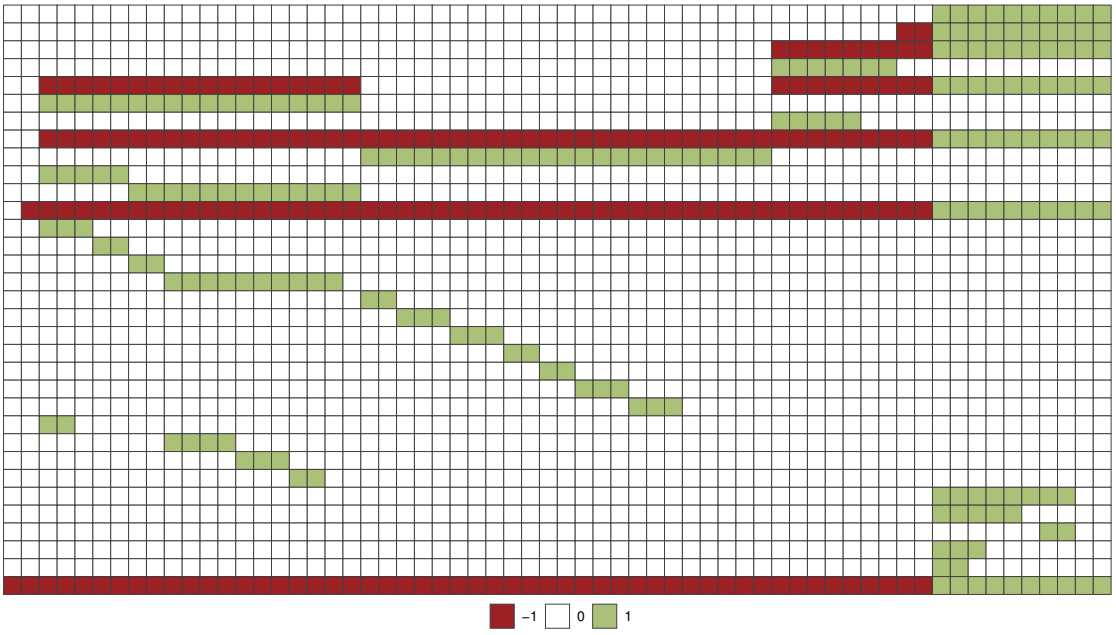


Figure 4.3: *Linear combination matrix* $\boldsymbol{A}$ *for the Australian GDP from income and expenditure sides.*

We perform a forecasting experiment as the one designed by Athanasopoulos *et al.* (2020). Base forecasts from $h = 1$ quarter ahead up to $h = 4$ quarters ahead for all the 95 separate time series have been obtained through simple univariate ARIMA models selected using the `auto.arima` function of the R-package `forecast` (Hyndman and Khandakar, 2008). The first training sample is set from 1984:Q4 to 1994:Q3, and a recursive training sample with expanding window length is used, for a total of 94 forecast origins. Finally the reconciled forecasts are obtained using three reconciliation approaches (ols, wls and shr, see Section 4.2.1) through the R package `Foreco` (Girolimetto and Di Fonzo, 2023a).

In Athanasopoulos *et al.* (2020) the probabilistic forecasts of the Australian quarterly GDP aggregates are separately reconciled from income $\left(\widetilde{X}^{I}_{GDP}\right)$ and expenditure $\left(\widetilde{X}^{E}_{GDP}\right)$ sides. This means that the empirical forecast distributions $\widetilde{X}^{I}_{GDP}$ and $\widetilde{X}^{E}_{GDP}$ are each coherent (see Section 4.2.3) within its own pertaining side with the other empirical forecast distributions, but in general $\widetilde{X}^{I}_{GDP} \neq \widetilde{X}^{E}_{GDP}$ at any forecast horizon. This circumstance could confuse the user, mostly when the difference between the empirical forecast distributions is not negligible, as shown in Figure 4.4, where the GDP empirical forecast distributions from income and expenditure sides for 2018:Q1 are presented

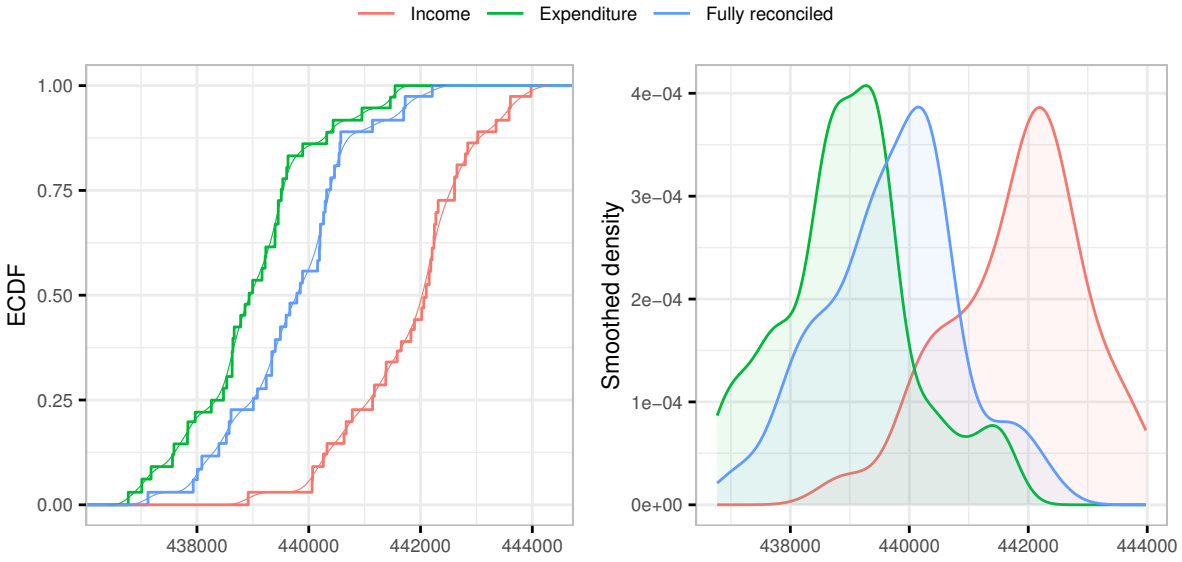


Figure 4.4: *Australian GDP empirical one-step-ahead forecast distributions for 2018:Q1, shr joint bootstrap-based reconciliation approach. Empirical Cumulative Distribution Function (left), and Smoothed density (right).*

along with their fully reconciled counterparts through the shr joint bootstrap-based reconciliation approach (see Section 4.2.3)[6].

Point and probabilistic forecasting accuracy

To evaluate the accuracy of the point forecasts we use the Mean Square Error (MSE)[7]:

$$MSE_{j,h} = \frac{1}{nT_h}\sum_{i=1}^{n}\sum_{t=1}^{T_h}\tilde{e}^{\,2}_{i,j,t+h},$$

where $\tilde{e}_{i,j,t+h} = (\tilde{y}_{i,j,t+h} - y_{i,j,t+h})$ is the the h-step-ahead forecast error using the approach j to forecast the i-th series, $j = 0$ denotes the base forecast (i.e., $\tilde{y}_{i,0,t+h} = \widehat{y}_{i,t+h}$), and t is the forecast origin. To assess any improvement in the reconciled forecasts compared to the base ones, we use the MSE-skill score:

$$\left(1 - \frac{MSE_{j,h}}{MSE_{0,h}}\right) \times 100.$$

The accuracy of the probabilistic forecasts is evaluated using the Cumulative Rank Probability Score (CRPS, Gneiting and Katzfuss, 2014):

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

[6]Note that the naive practice of averaging *GDP* forecasts from different sides yields a single forecast, that is though inconsistent with the component variables from both sides.

[7]The Mean Absolute Scaled Error (MASE) leads to the same conclusions (see the online appendix).

	Point forecasts MSE (%)			Probabilistic forecasts - CRPS(%) Bootstrap			Gaussian		
h	ols	wls	shr	ols	wls	shr	ols	wls	shr
Income									
1	1.63	1.07	**5.41**	**0.57**	-1.26	0.52	**0.13**	-1.87	-0.45
2	2.54	5.68	**6.10**	**1.46**	0.68	-0.50	1.11	**1.26**	-0.34
3	2.28	**7.81**	4.56	**0.18**	-0.64	-1.93	**0.14**	-0.24	-1.69
4	1.98	**9.33**	7.04	**-0.08**	-1.12	-1.35	**-0.23**	-1.00	-1.39
Expenditure									
1	**4.53**	0.07	2.48	**1.10**	0.78	0.08	**0.64**	-0.83	-0.70
2	**5.09**	3.90	1.72	**2.34**	1.25	-1.07	**1.68**	0.58	-1.95
3	6.96	**9.18**	6.24	**2.42**	1.50	-0.38	**1.69**	1.00	-1.02
4	8.01	**11.76**	8.34	**3.73**	2.73	0.72	**3.18**	2.49	0.20
Fully reconciled									
1	4.59	1.14	**4.77**	**1.13**	-0.75	-0.20	**-0.59**	-3.46	-2.85
2	5.76	**6.24**	4.76	**2.81**	1.07	-1.27	**1.30**	-0.33	-3.02
3	7.31	**10.94**	8.21	**1.99**	0.26	-1.46	**0.91**	-0.62	-2.80
4	7.90	**13.24**	10.81	**2.83**	0.75	-0.63	**1.86**	-0.43	-2.05

Table 4.1: *MSE and CRPS-skill scores (relative to base forecasts) for the point and probabilistic Australian GDP forecasts from alternative reconciliation approaches. Negative values are highlighted in red, the best for each row is marked in bold.*

where $\hat{P}_i(\omega) = \frac{1}{L}\sum_{l=1}^{L} \mathbf{1}\left(x_{i,l} \leq \omega\right)$, $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_L \in \mathbb{R}^n$ is a collection of L random draws taken from the predictive distribution and $\boldsymbol{z} \in \mathbb{R}^n$ is the observation vector. In addition, to evaluate the forecasting accuracy for the whole system, we employ the Energy Score (ES), that is the CRPS extension to the multivariate case[8]:

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

Results

Table 4.1 shows the MSE and CRPS-skill scores for the *GDP* point and probabilistic reconciled forecasts. Table 4.2 presents the MSE and ES-skill scores for all 95 Australian QNA variables from both income and expenditure sides. The 'Income' and 'Expenditure' panels, respectively, reproduce the results found by Athanasopoulos *et al.* (2020). The 'Fully reconciled' panels show the skill scores for the simultaneously reconciled forecasts. For the point forecasts, all the reconciliation approaches improve forecast accuracy compared to the base forecasts. In detail, shr is almost always the best approach for the one-step-ahead forecasts, whereas wls is competitive for $h \geq 2$. Looking at the probabilistic reconciliation results in Table 4.1, it is worth noting that for *GDP* ols outperforms both wls and shr, whatever side and framework (parametric or not)

[8] An alternative to the Energy Score is the Variogram Score (Scheuerer and Hamill, 2015), considered in the online appendix, that leads to similar conclusions.

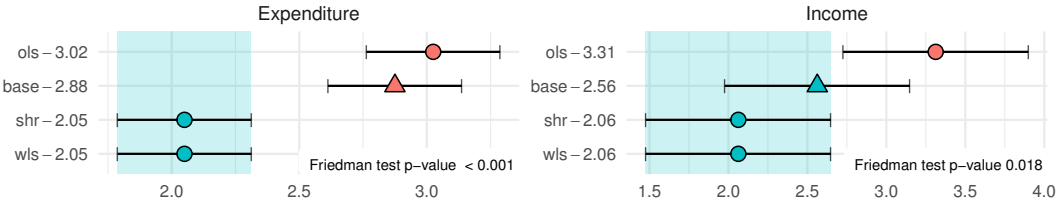
	Point forecasts MSE (%)			Probabilistic forecasts - ES(%) Bootstrap			Gaussian		
h	ols	wls	shr	ols	wls	shr	ols	wls	shr
Income									
1	3.16	6.32	**10.55**	2.30	4.24	**6.15**	1.87	3.47	**5.15**
2	2.58	6.07	**8.18**	2.14	4.08	**4.08**	1.62	**3.41**	3.08
3	2.18	**5.81**	4.09	1.78	**3.29**	3.19	1.28	**2.67**	2.35
4	2.18	**6.78**	5.51	1.86	3.73	**4.42**	1.42	3.10	**3.67**
Income - Fully reconciled									
1	3.78	7.57	**8.85**	2.77	4.58	**4.87**	1.30	1.94	**2.25**
2	2.91	6.12	**6.92**	2.59	**3.95**	3.40	1.20	**1.59**	0.87
3	2.67	**6.23**	5.57	2.24	**3.70**	3.17	0.92	**1.60**	0.79
4	2.87	**7.21**	6.07	2.65	**4.40**	3.95	1.44	**2.34**	1.59
Expenditure									
1	6.50	6.75	**8.78**	3.71	4.54	**4.94**	2.20	1.92	**2.58**
2	4.90	5.50	**5.52**	2.88	**3.04**	2.68	**1.28**	0.67	0.32
3	4.27	**6.08**	5.65	2.57	**2.94**	2.29	**0.88**	0.71	0.09
4	4.01	**6.69**	5.20	2.43	**2.65**	1.63	**0.77**	0.44	-0.65
Expenditure - Fully reconciled									
1	6.51	6.82	**9.08**	3.76	4.57	**4.99**	2.19	1.79	**2.44**
2	5.09	6.24	**6.54**	3.03	**3.48**	3.14	**1.18**	0.82	0.57
3	4.38	**6.75**	5.94	2.58	**3.11**	2.29	**0.81**	0.69	-0.12
4	3.98	**7.33**	5.82	2.36	**2.82**	1.81	**0.60**	0.32	-0.70

Table 4.2: *MSE and ES-skill scores (relative to base forecasts) for the point and probabilistic forecasts from alternative reconciliation approaches (all Australian QNA variables). Negative values are highlighted in red, the best for each row is marked in bold.*

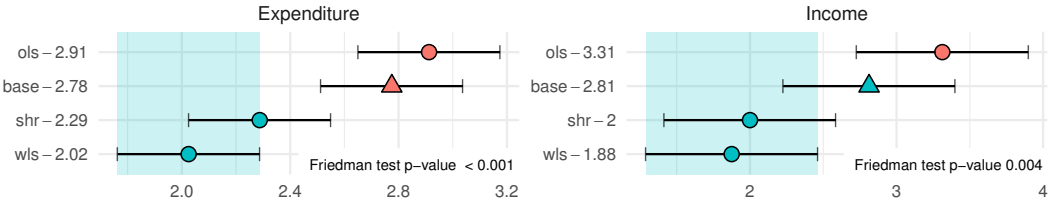
is considered. However, when all 95 variables are considered (Table 4.2), shr and wls approaches almost always show the best performance. In the Gaussian framework, these results are confirmed for the income side, whereas shr performs poorly when we look at the expenditure side (either fully reconciled or not).

In Figure 4.5 are shown the results obtained by the non-parametric Friedman test and the post hoc "Multiple Comparison with the Best" (MCB) Nemenyi test (Koning *et al.*, 2005; Kourentzes and Athanasopoulos, 2019; Makridakis *et al.*, 2022) to determine if the forecasting performances of the different techniques are significantly different from one another. In general, wls always falls in the set of the best performing approaches. For the bootstrap-based probabilistic reconciled forecasts of the expenditure side variables, wls and shr significantly improves in terms of MSE and CRPS compared to the base forecasts. This result is confirmed in the remaining cases as well.

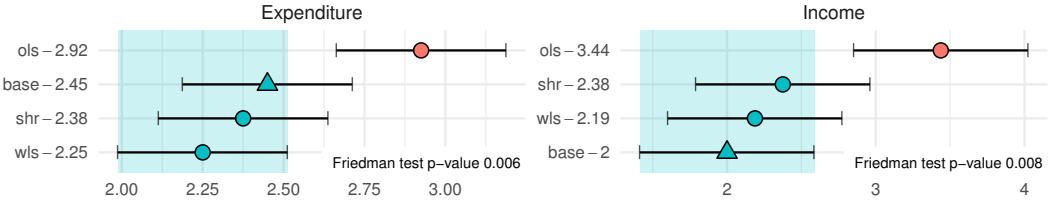
In conclusion, when income and expenditure sides are simultaneously considered for both point and probabilistic forecasts, forecast reconciliation succeeds in improving the base forecasts of GDP and its component aggregates, while preserving the full coherence with the National Accounts constraints.



(a) *Point forecasts (MSE)*



(b) *Non parametric joint bootstrap probabilistic forecasts (CRPS)*



(c) *Gaussian probabilistic forecasts (CRPS)*

Figure 4.5: *MCB Nemenyi test for the fully reconciled forecasts of the Australian QNA variables at any forecast horizon. In each panel, the Friedman test p-value is reported in the lower right corner. The mean rank of each approach is shown to the right of its name. Statistical differences in performance are indicated if the intervals of two forecast reconciliation approaches do not overlap. Thus, approaches that do not overlap with the blue interval are considered significantly worse than the best, and vice-versa.*

4.4.2 Reconciled probabilistic forecasts of the European Area GDP from output, income and expenditure sides

In this section, we consider the system of European QNA for the GDP at current prices (in euro), with time series spanning the period 2000:Q1-2019:Q4. This system has many variables linked by several, possibly redundant, accounting constraints, such that it is difficult to manually build a system of non-redundant constraints.

The National Accounts are a coherent and consistent set of macroeconomic indicators that are used mostly for economic research and forecasting, policy design, and coordination mechanisms. In this dataset, GDP is a key macroeconomic quantity that is measured using three main approaches, namely output (or production), income and expenditure. These parallel systems internally present a well-defined hierarchical structure of variables with relevant economic significance, such as Final consumption, on the

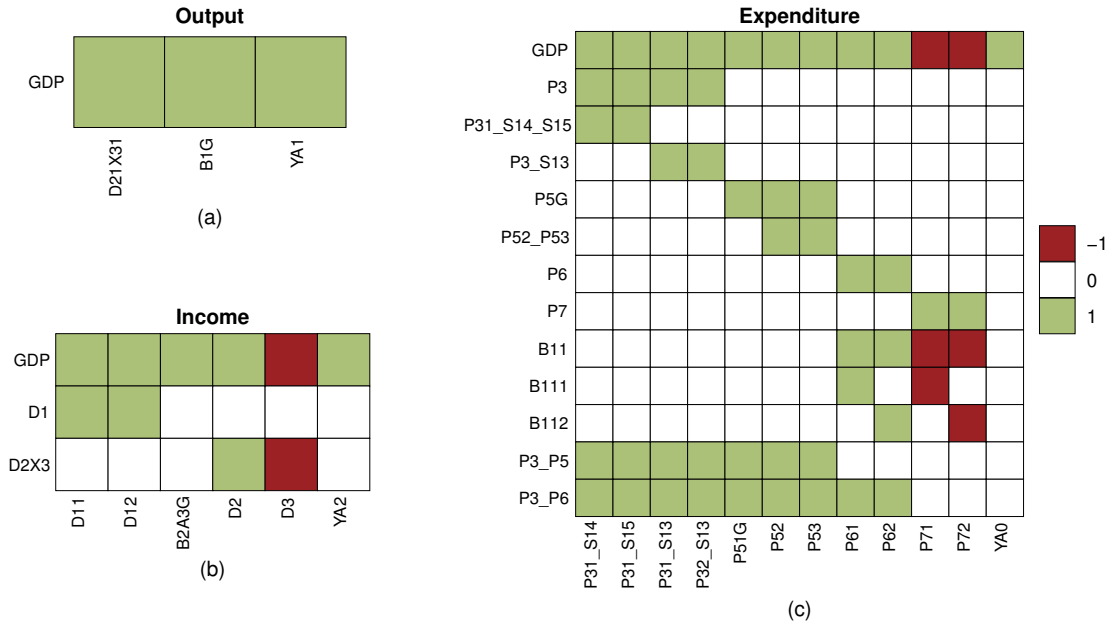


Figure 4.6: *Linear combination matrices $\boldsymbol{A}$ for the European Area GDP: output side in panel (a), income side in panel (b), expenditure side in panel (c).*

expenditure side, Gross operating surplus and mixed income on the income side, and Total gross value added on the output side. In the EU countries, the data is processed on the basis of the ESA 2010 classification and are released by Eurostat[9]. We consider the 19 Euro Area member countries (Austria, Belgium, Finland, France, Germany, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Greece, Slovenia, Cyprus, Malta, Slovakia, Estonia, Latvia, and Lithuania) that have been using the euro since 2015. In Figure 4.6 we have represented the aggregation matrices describing output, income, and expenditure constraints, respectively: in panel (a), matrix $\boldsymbol{A}_{\mathrm{O}}$ for the output side, in panel (b) matrix $\boldsymbol{A}_{\mathrm{I}}$ for the income side, and in panel (c) matrix $\boldsymbol{A}_{\mathrm{E}}$ for the expenditure side. The zero-constraints coefficient matrix describing the QNA variables for a single country can thus be written as

$$\boldsymbol{\Gamma}_{\mathrm{GDP}} = \begin{bmatrix} \boldsymbol{K}_{\mathrm{E}} & -\boldsymbol{A}_{\mathrm{E}} & \mathbf{0}_{(13\times 6)} & \mathbf{0}_{(13\times 3)} \\ \boldsymbol{K}_{\mathrm{I}} & \mathbf{0}_{(3\times 12)} & -\boldsymbol{A}_{\mathrm{I}} & \mathbf{0}_{(3\times 3)} \\ \boldsymbol{K}_{\mathrm{O}} & \mathbf{0}_{(1\times 12)} & \mathbf{0}_{(1\times 6)} & -\boldsymbol{A}_{\mathrm{O}} \end{bmatrix},$$

where $\boldsymbol{K}_{\mathrm{E}}$, $\boldsymbol{K}_{\mathrm{I}}$ and $\boldsymbol{K}_{\mathrm{O}}$, respectively, are the following (13×15), (3×15) and (1×15) matrices:

$$\boldsymbol{K}_{\mathrm{E}} = \begin{bmatrix} 1 & \mathbf{0}'_{(1\times 12)} & \mathbf{0}'_{(1\times 2)} \\ \mathbf{0}_{(12\times 1)} & \boldsymbol{I}_{12} & \mathbf{0}_{(12\times 2)} \end{bmatrix},$$

[9]Further information can be found at `https://ec.europa.eu/eurostat/esa2010/` and `https://ec.europa.eu/eurostat/web/national-accounts/data/database`.

$$\boldsymbol{K}_{\mathrm{I}} = \begin{bmatrix} 1 & \mathbf{0}'_{(1\times 12)} & \mathbf{0}'_{(1\times 2)} \\ \mathbf{0}_{(2\times 1)} & \mathbf{0}_{(2\times 12)} & \boldsymbol{I}_2 \end{bmatrix}, \quad \text{and} \quad \boldsymbol{K}_{\mathrm{O}} = \begin{bmatrix} 1 & \mathbf{0}'_{(1\times 14)} \end{bmatrix}.$$

This disaggregation is common for almost all European countries, the only differences being related to the presence/absence of an aggregate measuring the statistical discrepancy in each accounting side[10]. The (361×720) matrix describing the accounting relationships for the whole EA19 QNA by countries and accounting sides can be written as follows:

$$\boldsymbol{\Gamma} = \begin{bmatrix} \boldsymbol{\Gamma}_{\mathrm{GDP}} & \mathbf{0}_{(17\times 684)} \\ \begin{bmatrix} \mathbf{0}_{(21\times 15)} & \boldsymbol{I}_{21} \end{bmatrix} & -\mathbf{1}'_{19} \otimes \begin{bmatrix} \mathbf{0}_{(21\times 15)} & \boldsymbol{I}_{21} \end{bmatrix} \\ \mathbf{0}_{(323\times 36)} & \boldsymbol{I}_{19} \otimes \boldsymbol{\Gamma}_{\mathrm{GDP}} \end{bmatrix},$$

where $\otimes$ is the Kronecker product, and the top-left portion of $\boldsymbol{\Gamma}$ refers to the European Area aggregates as a whole. In order to proceed with the calculations, it is necessary to eliminate the columns related to null variables (e.g., the statistical discrepancy aggregate for the countries/sides where it is not contemplated, see footnote 10). Then, to eliminate possible remaining redundant constraints, we apply the QR decomposition (Section 4.3.1). Finally, we obtain the linear combination matrix $\boldsymbol{A}$, which refers to 311 free and 358 constrained time series, and the full rank zero-constraints matrix $\boldsymbol{C} = \begin{bmatrix} \boldsymbol{I}_{358} & -\boldsymbol{A} \end{bmatrix}$.

A rolling forecast experiment with expanding window is performed using ARIMA models to produce the individual series' base forecasts. The first training set is set from 2000:Q1 to 2009:Q4, which gives 40 one-step-ahead, 39 two-step-ahead, 38 three-step-ahead and 37 four-step-ahead ARIMA forecasts, respectively. The used reconciliation approaches are ols, wls and shr, and the forecast accuracy is evaluated through MSE, CRPS and ES indices, as described in Section 4.4.1.

Results

Table 4.3 shows the MSE indices for point forecasts, and the ES indices for probabilistic nonparametric (bootstrap) and parametric (Gaussian) forecasts. The rows of the table are divided into three parts: the first row shows the results for GDP, the second to fourth rows the National Accounts' divisions (income, expenditure, or output sides), while the

[10]For 14 countries (Belgium, France, Germany, Italy, Luxembourg, Netherlands, Spain, Greece, Slovenia, Cyprus, Malta, Slovakia, Latvia, and Lithuania), the expenditure, income and output statistical discrepancies are not present. A statistical discrepancy aggregate is present in the output QNA of Portugal and in the expenditure QNA of Finland, Estonia and Austria. Ireland is the only country where a statistical discrepancy aggregate is present in all accounting sides.

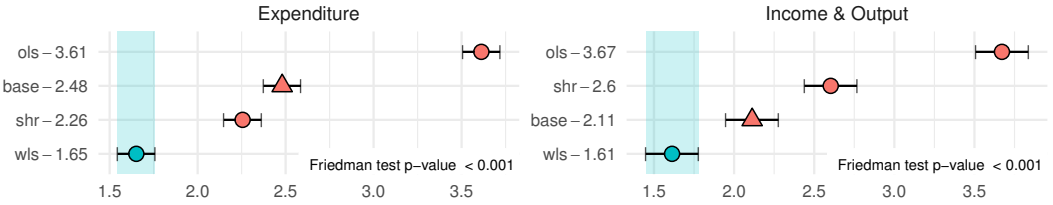
	Point forecasts MSE (%)			Probabilistic forecasts - ES(%) Bootstrap			Gaussian		
	ols	wls	shr	ols	wls	shr	ols	wls	shr
GDP	14.0	**41.1**	21.6	9.7	**25.8**	15.7	3.4	**20.8**	3.7
Sides									
Expenditure	15.8	**31.0**	28.3	9.1	**20.2**	16.3	6.8	**15.7**	8.7
Income	-2.2	**27.3**	9.2	-0.3	**15.3**	5.6	-2.7	**11.6**	-3.2
Output	0.3	**35.0**	15.0	1.2	**21.6**	12.9	-1.5	**17.1**	1.3
Countries									
EA19	17.6	**37.1**	31.6	10.1	**23.5**	18.1	7.8	**19.3**	10.0
Austria	<-30	18.5	**21.5**	-26.8	**12.1**	10.0	-21.6	**9.1**	5.4
Belgium	-5.9	13.5	**21.0**	-2.7	10.7	**10.8**	-3.1	**8.3**	5.9
Cyprus	<-30	**11.7**	8.6	<-30	**7.1**	1.7	<-30	**3.6**	-2.8
Estonia	<-30	20.9	**26.8**	<-30	**12.7**	11.2	<-30	**9.7**	6.7
Finland	<-30	**24.8**	14.7	<-30	**15.3**	9.5	<-30	**13.2**	3.4
France	4.4	**19.1**	-0.1	3.2	**11.7**	3.8	0.3	**7.9**	-6.8
Germany	15.0	20.6	**25.0**	13.5	**21.6**	18.6	7.5	**11.1**	6.8
Greece	<-30	**1.5**	<-30	-12.9	**3.9**	-7.1	-15.1	**0.8**	-14.1
Ireland	4.1	**6.6**	3.7	2.1	**5.9**	3.0	0.5	**2.9**	-0.7
Italy	9.8	12.6	**24.0**	8.4	10.8	**17.1**	4.4	4.9	**7.8**
Latvia	<-30	**26.2**	18.3	<-30	**13.0**	8.7	<-30	**13.7**	5.6
Lithuania	<-30	27.3	**28.5**	<-30	**15.0**	12.8	<-30	**12.1**	8.4
Luxembourg	<-30	**6.6**	-10.3	<-30	**6.1**	-3.0	<-30	**2.0**	-10.0
Malta	<-30	**5.4**	-9.9	<-30	**4.9**	-7.4	<-30	**0.8**	-12.0
Netherlands	7.4	**16.9**	11.9	6.1	**14.1**	7.3	3.5	**11.7**	4.4
Portugal	<-30	**11.4**	-8.1	<-30	**7.4**	-3.4	<-30	**4.1**	-10.1
Slovakia	<-30	**21.5**	16.4	<-30	**10.0**	6.9	<-30	**9.2**	2.0
Slovenia	<-30	**24.1**	15.5	<-30	**13.5**	8.3	<-30	**11.1**	1.8
Spain	6.5	**17.7**	0.9	5.3	**9.2**	1.4	1.7	**6.4**	-7.3

Table 4.3: *MSE and ES-skill scores (relative to base forecasts) for the point and probabilistic forecasts from alternative reconciliation approaches (European Area QNA). Negative values are highlighted in red, the best for each row is marked in bold.*

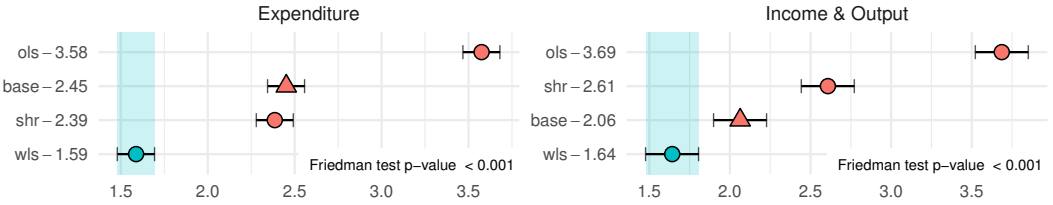
remaining rows correspond to the 19 countries and EA19. All forecast horizons are considered[11].

When only *GDP* is considered, any reconciliation approach consistently outperforms the base forecasts, both in the point and probabilistic cases. The wls approach confirms a good performance when we look at the income, expenditure, and output sides (second panel), while ols shows the worst performance. When the parametric framework is considered, shr is worse than the base forecast for the income side. At country level, ols is the approach that overall shows the worst performance, with many relative losses in the accuracy indices higher than 30%. It is worth noting that all reconciliation approaches always perform well for the whole Euro Area (EA19). Overall, in this forecasting experiment wls appears to be the most performing reconciliation approach, showing no negative skill score, and improvements higher than 20% and 10% for nonparametric and parametric probabilistic frameworks, respectively.

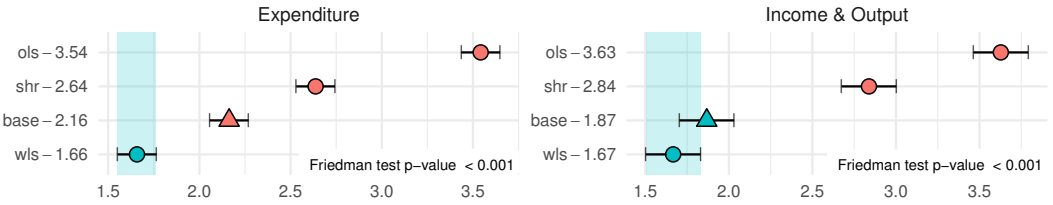
[11] A disaggregated analysis by forecast horizon is reported in the online appendix.



(a) *Point forecasts*



(b) *Non parametric joint bootstrap probabilistic forecasts*



(c) *Gaussian probabilistic forecasts*

Figure 4.7: *MCB Nemenyi test for the fully reconciled forecasts of the European Area QNA variables at any forecast horizon. In each panel, the Friedman test p-value is reported in the lower right corner. The mean rank of each approach is shown to the right of its name. Statistical differences in performance are indicated if the intervals of two forecast reconciliation approaches do not overlap. Thus, approaches that do not overlap with the blue interval are considered significantly worse than the best, and vice-versa.*

Figure 4.7 shows the MCB Nemenyi test at any forecast horizon, distinct by expenditure side and income and output sides, respectively. The results just seen are further confirmed by this alternative forecast assessment tool: it clearly appears that the wls approach almost always significantly improves compared to the base forecast, in terms of both point and probabilistic forecasts.

Finally, a visual evaluation of the accuracy improvement obtained through wls forecast reconciliation, although limited to a single forecast horizon, is offered by Figure 4.8, showing the European map with the CRPS skill scores for the one-step-ahead *GDP* non-parametric probabilistic reconciled forecasts. It is worth noting that only for two countries (Greece and Portugal) a decrease is registered (-3.3% and -4.9%, respectively). In all other cases, improvements in the forecasting accuracy are obtained, with

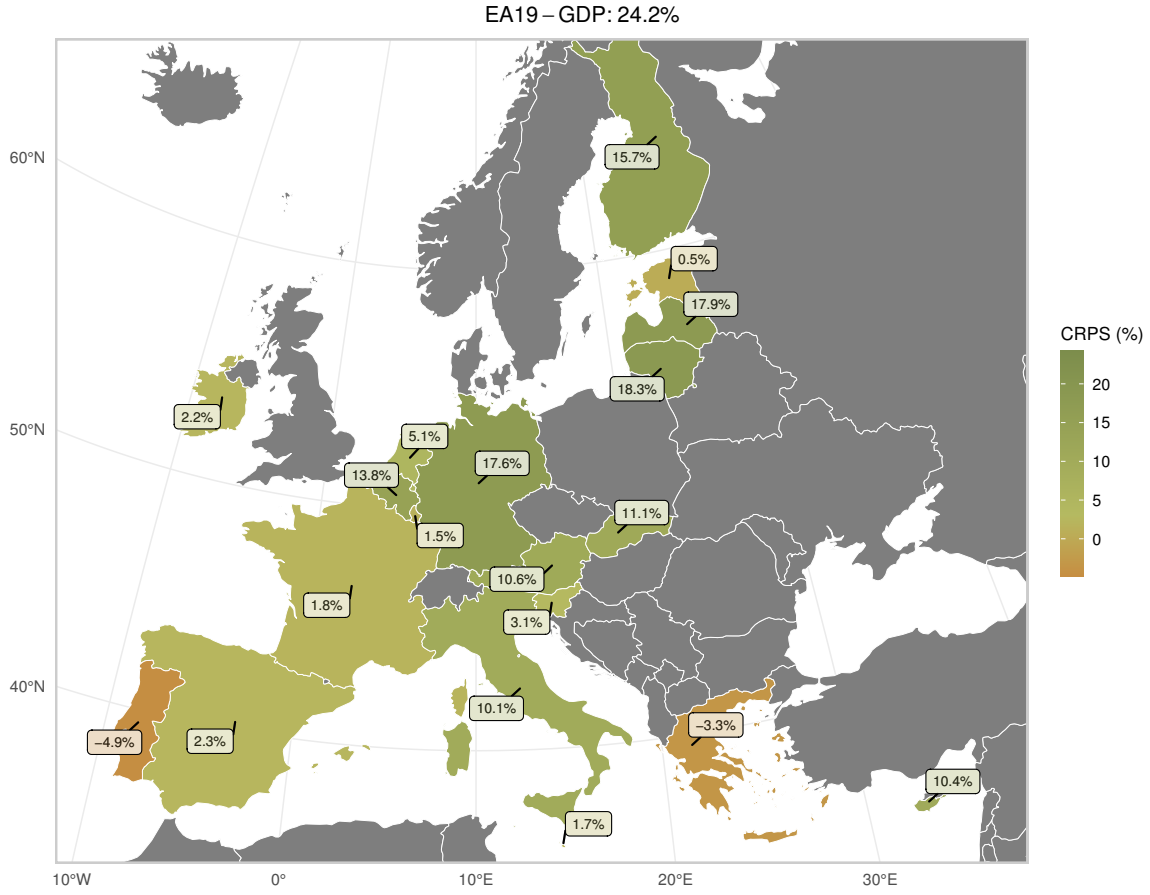


Figure 4.8: *CRPS-skill scores of the one-step-ahead GDP non-parametric joint bootstrap probabilistic reconciled (wls) forecasts for the 19 Euro Area countries.*

Germany and Lithuania leading the way with about 18%. Furthermore, the improvement in the forecasting accuracy for the $EA19-GDP$ is 24.2%, the highest throughout the whole Euro Area.

Chapter 5

Cross-temporal probabilistic forecast reconciliation

5.1 Introduction

Forecast reconciliation is a post-forecasting process intended to improve the quality of forecasts for a system of linearly constrained multiple time series (Hyndman *et al.*, 2011; Panagiotelis *et al.*, 2021). There are many fields where forecast reconciliation is useful, such as when forecasting demand in supply chains with product categories (Punia *et al.*, 2020; Kourentzes and Athanasopoulos, 2021), electricity demand and power generation (Spiliotis *et al.*, 2020; Ben Taieb *et al.*, 2021), GDP and its components (Athanasopoulos *et al.*, 2020), tourist flows across geographic regions and travel purpose (Kourentzes and Athanasopoulos, 2019), and more. Moreover, effective decision-making depends on the support of accurate and coherent forecasts, making the use of forecast reconciliation methods increasingly popular in recent years (Athanasopoulos *et al.*, 2023).

Temporal reconciliation is another important aspect of forecast reconciliation that can help organizations to better align their forecasting efforts. This approach consists in reconciling forecasts that are generated at different time horizons, such monthly, quarterly or annual. For example, a retail company may need to reconcile monthly forecasts of sales with quarterly forecasts of revenue to ensure that they are aligned and consistent.

Classical reconciliation approaches (bottom-up, top-down, middle-out, see Dunn *et al.*, 1976, Gross and Sohl, 1990, Athanasopoulos *et al.*, 2009, respectively) addressed the issue of incoherent forecasts in a cross-sectional hierarchy by forecasting only one level and using these to generate forecasts for the remaining series. All of these approaches ignore useful information available at other levels (Pennings and Van Dalen,

2017). Recently, hierarchical forecasting (Fliedner, 2001) has significantly evolved to include modern least squares-based reconciliation techniques in the cross-sectional framework (Hyndman *et al.*, 2011; Wickramasuriya *et al.*, 2019; Panagiotelis *et al.*, 2021), later extended to temporal hierarchies (Athanasopoulos *et al.*, 2017; Nystrup *et al.*, 2020). Obtaining coherent forecasts across both the cross-sectional and temporal dimensions (known as *cross-temporal coherence*) has been limited to sequential approaches that address each dimension separately (Kourentzes and Athanasopoulos, 2019; Yagli *et al.*, 2019; Punia *et al.*, 2020; Spiliotis *et al.*, 2020). In Chapter 1, we suggested a unified reconciliation step that takes into account both the cross-sectional and temporal dimensions, instead of dealing with them separately, utilizing the entire cross-temporal hierarchy.

However, these cross-temporal works focus on point forecasting, and do not consider distributional or probabilistic forecasts (Gneiting and Katzfuss, 2014). In the cross-sectional and temporal frameworks, there have been some developments towards probabilistic forecasting including Ben Taieb *et al.* (2017), Panamtash and Zhou (2018), Jeon *et al.* (2019), Yang (2020), Yagli *et al.* (2020), Ben Taieb *et al.* (2021), Corani *et al.* (2021), Zambon *et al.* (2024), Corani *et al.* (2023), and Wickramasuriya (2024). Panagiotelis *et al.* (2023) made a significant contribution by formalizing cross-sectional probabilistic reconciliation using the geometric framework for point forecast reconciliation of Panagiotelis *et al.* (2021). They show how a reconciled forecast can be constructed from an arbitrary base forecast when its density is available and when only a sample can be drawn. They also show that in the case of elliptical distributions, the correct predictive distribution can be recovered via linear reconciliation, regardless of the base forecast location and scale parameters, and derive conditions for this to hold in the special case of reconciliation via projection.

In this chapter, we extend cross-sectional probabilistic reconciliation to the cross-temporal case, working on issues related to the two-fold nature of this framework. First, we revise and develop the notation proposed in Chapter 1 to generalize the work of Panagiotelis *et al.* (2023). This allows us to move from cross-temporal point reconciliation to a probabilistic setting through the generalization of definitions and theorems well-established in the cross-sectional framework. Second, we propose solutions to draw a sample from the base forecast distribution according to either a parametric approach that assumes Gaussianity or a non-parametric approach that bootstraps the base model residuals. Third, we propose some solutions to specific problems that arise when combining the cross-sectional and temporal dimensions. We propose using multi-step residuals to estimate the relationships between different forecast horizons when we

deal with temporal levels, since one-step residuals are not suitable for this purpose. To solve high-dimensionality issues we introduce the idea of overlapping residuals and consider alternative forms for constructing the covariance matrix. Fourth, we propose new shrinkage procedures for reconciliation that aim to identify a feasible cross-temporal structure. The algorithms described in this chapter are implemented in the `FoReco` package (Girolimetto and Di Fonzo, 2023a) for R (R Core Team, 2022). Furthermore, the online appendix contains complementary materials on methodological and practical issues, and supplementary tables and graphs related to the empirical applications.

The remainder of the chapter is structured as follows. In Section 5.2, we provide a unified notation for the cross-sectional, temporal and cross-temporal point reconciliation. We generalize the cross-sectional definitions and theorems developed by Panagiotelis *et al.* (2023) in Section 5.3, and propose both a parametric Gaussian and a non-parametric bootstrap approach to draw a sample from the base forecast distribution. In Section 5.4, we analyze the structure of the cross-temporal covariance matrix, proposing four alternative forms, and propose shrinkage approaches for reconciliation. In addition, we explore cross-temporal residuals (overlapping and multi-step) looking at their advantages and limitations. Two empirical applications using the Australian GDP and the Australian Tourism Demand datasets are considered in Sections 5.5 and 5.6, respectively.

5.2 Notation and definitions

Let $\boldsymbol{y}_t = [y_{1,t}, \dots, y_{i,t}, \dots, y_{n,t}]'$ be an n-variate linearly constrained time series observed at the most temporally disaggregated level, with a seasonality of period m (e.g., $m = 12$ for monthly data, $m = 4$ for quarterly data, $m = 24$ for hourly data). Suppose that the constraints are expressed by linear equations such that (see Chapter 1)

$$\boldsymbol{C}_{cs}\boldsymbol{y}_t = \boldsymbol{0}_{(n_a \times 1)}, \qquad t = 1, \ \dots, \ T, \tag{5.1}$$

where $\boldsymbol{C}_{cs}$ is the $(n_a \times n)$ zero constraints cross-sectional matrix, that can be seen as the coefficient matrix of a linear system with n_a equations and n variables[1].

An example is a hierarchical time series where series at upper levels can be expressed by appropriately summing part or all of the series at the bottom level. Figure 5.1(a) shows the two-level hierarchical structure for three linearly constrained time series such that $y_{T,t} = y_{X,t} + y_{Y,t}$, $\forall t = 1, ..., T$. Now let $\boldsymbol{y}_t = [\boldsymbol{u}_t' \quad \boldsymbol{b}_t']'$, where $\boldsymbol{u}_t = [y_{1,t}, \dots, y_{n_a,t}]'$

[1]Hyndman (2022) and Chapter 4 show that this "zero-contrained representation" is more general and computationally efficient.

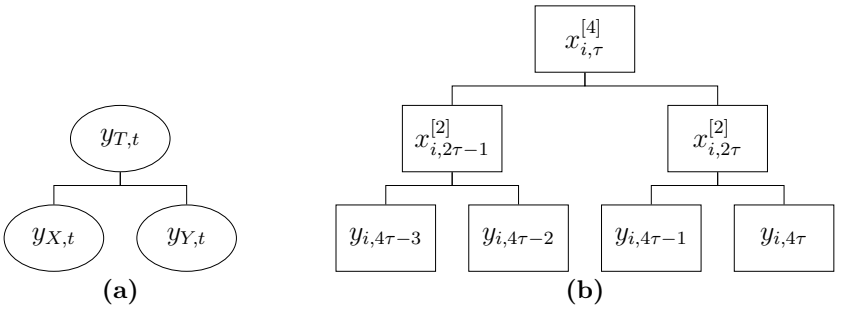


Figure 5.1: ***(a)*** *A simple two-level cross-sectional hierarchy for 3 time series with $n_a = 1$ and $n_b = 2$.* ***(b)*** *A temporal hierarchy for a quarterly series ($m = 4$ and $\mathcal{K} = \{4, 2, 1\}$).*

is the n_a-vector of upper levels time series and $\boldsymbol{b}_t = \begin{bmatrix} y_{(n_a+1),t} & \dots & y_{n,t} \end{bmatrix}'$ is the n_b-vector of bottom level time series with $n = n_a + n_b$. The upper and lower level time series are connected by the cross-sectional aggregation matrix $\boldsymbol{A}_{cs}$ such that $\boldsymbol{u}_t = \boldsymbol{A}_{cs}\boldsymbol{b}_t$. Following Chapter 4, we can always construct a zero-constraints cross-sectional matrix from the aggregation matrix, $\boldsymbol{C}_{cs} = [\boldsymbol{I}_{n_a} \quad -\boldsymbol{A}_{cs}]$, where $\boldsymbol{I}_{n_a}$ is an identity matrix of dimension n_a. Finally, the cross-sectional structural matrix is given by $\boldsymbol{S}_{cs} = \begin{bmatrix} \boldsymbol{A}_{cs} \\ \boldsymbol{I}_{n_b} \end{bmatrix}$, providing the structural representation (Hyndman *et al.*, 2011) $\boldsymbol{y}_t = \boldsymbol{S}_{cs}\boldsymbol{b}_t$. Considering the hierarchical example in Figure 5.1(a), we have

$$\boldsymbol{A}_{cs} = \begin{bmatrix} 1 & 1 \end{bmatrix}, \quad \boldsymbol{C}_{cs} = \begin{bmatrix} 1 & -1 & -1 \end{bmatrix} \quad \text{and} \quad \boldsymbol{S}_{cs} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

In general there is no reason for $\boldsymbol{u}_t$ to be restricted to simple sums of $\boldsymbol{b}_t$; therefore $\boldsymbol{A}_{cs} \in \mathbb{R}^{n_a \times n_b}$ may contain any real values, and not only 0s and 1s.

Considering now the temporal framework, we denote as $\mathcal{K} = \{k_p, k_{p-1}, \dots, k_2, k_1\}$ the set of p factors of m, in descending order, where $k_1 = 1$ and $k_p = m$ (Athanasopoulos *et al.*, 2017). For example, for quarterly time series $m = 4$, $p = 3$, and $\mathcal{K} = \{4, 2, 1\}$. Given a factor k of m, and assuming that $T = Nm$ (where N is the length of the most temporally aggregated version of the series), we can construct a temporally aggregated version of the time series of a single variable $\{\boldsymbol{y}_{i,t}\}_{t=1,\dots,T}$, through the non-overlapping sums of its k successive values, which has a seasonal period equal to $M_k = \dfrac{m}{k}$: $x^{[k]}_{i,j} = \displaystyle\sum_{t=(j-1)k+1}^{jk} y_{i,t}$, where $j = 1, \dots, N_k$, $i = 1, \dots, n$, $N_k = \dfrac{T}{k}$ and $x^{[1]}_{i,j} = y_{i,t}$. Define τ as the observation index of the most aggregate level k_p. For a fixed temporal aggregation order $k \in \mathcal{K}$, we stack the observations in the column vector $\boldsymbol{x}^{[k]}_{i,\tau} = \begin{bmatrix} x^{[k]}_{i,M_k(\tau-1)+1} & x^{[k]}_{i,M_k(\tau-1)+2} & \dots & x^{[k]}_{i,M_k\tau} \end{bmatrix}'$, and obtain the vector for all

the temporal aggregation orders $\boldsymbol{x}_{i,\tau} = \begin{bmatrix} x_{i,\tau}^{[k_p]} & \boldsymbol{x}_{i,\tau}^{[k_{p-1}]\prime} & \dots & \boldsymbol{x}_{i,\tau}^{[1]\prime} \end{bmatrix}', \tau = 1, \dots, N$. The structural representation of the temporal hierarchy (Athanasopoulos *et al.*, 2017) is then $\boldsymbol{x}_{i,\tau} = \boldsymbol{S}_{te}\boldsymbol{x}_{i,\tau}^{[1]}$, where $\boldsymbol{S}_{te} = \begin{bmatrix} \boldsymbol{A}_{te} \\ \boldsymbol{I}_m \end{bmatrix}$ is the $[(m + k^*) \times m]$ temporal structural matrix, $\boldsymbol{A}_{te} = \begin{bmatrix} \mathbf{1}_{k_p} & \boldsymbol{I}_{\frac{m}{k_{p-1}}} \otimes \mathbf{1}_{k_{p-1}} & \dots & \boldsymbol{I}_{\frac{m}{k_2}} \otimes \mathbf{1}_{k_2} \end{bmatrix}'$ is the $(k^* \times m)$ temporal aggregation matrix with $k^* = \sum_{k \in \mathcal{K}\setminus\{k_1\}} M_k$, the number of upper time series of the temporal hierarchy, $\mathbf{1}_{k_p}$ is a $(k_p \times 1)$ vector of all ones, and $\otimes$ is the Kronecker product. For each series $x_{i,\tau}$, $i = 1, \dots, n$, we have also the zero-constrained representation

$$\boldsymbol{C}_{te}\boldsymbol{x}_{i,\tau} = \mathbf{0}_{[k^* \times (m+k^*)]}, \qquad \tau = 1, \dots, N, \qquad i = 1, \dots, n \tag{5.2}$$

where $\boldsymbol{C}_{te} = [\boldsymbol{I}_{k^*} \;\; -\boldsymbol{A}_{te}]$ is the $[k^* \times (m + k^*)]$ zero constraints temporal matrix. Figure 5.1(b) shows the hierarchical representation of a quarterly time series, for which $m = 4$, $\mathcal{K} = \{4, 2, 1\}$ and

$$\boldsymbol{A}_{te} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \quad \boldsymbol{C}_{te} = \begin{bmatrix} 1 & 0 & 0 & -1 & -1 & -1 & -1 \\ 0 & 1 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & -1 \end{bmatrix} \quad \text{and} \quad \boldsymbol{S}_{te} = \begin{bmatrix} \boldsymbol{A}_{te} \\ \boldsymbol{I}_4 \end{bmatrix}.$$

When we temporally aggregate each series, the cross-sectional constraints for the most temporally disaggregated series (5.1) hold for all the temporal aggregation orders such that $\boldsymbol{C}_{cs}\boldsymbol{x}_j^{[k]} = \mathbf{0}_{(n_a \times 1)}$, for $k \in \mathcal{K}$ and $j = 1, \dots, N_k$, where $\boldsymbol{x}_j^{[k]} = \begin{bmatrix} \boldsymbol{u}_j^{[k]\prime} & \boldsymbol{b}_j^{[k]\prime} \end{bmatrix}'$ with $\boldsymbol{u}_j^{[k]} = \begin{bmatrix} x_{1,\,j}^{[k]} & \dots & x_{n_a,\,j}^{[k]} \end{bmatrix}'$ is the n_a-vector of upper time series and $\boldsymbol{b}_j^{[k]} = \begin{bmatrix} x_{(n_a+1),\,j}^{[k]} & \dots & x_{n,\,j}^{[k]} \end{bmatrix}'$ is the n_b-vector of bottom time series in the temporal hierarchy.

To include both cross-sectional and temporal constraints at the same time in a unified framework, we stack the series into a $[n \times (m + k^*)]$ matrix $\boldsymbol{X}_\tau$, where we recall that n, m, and k^* represent respectively the total number of time series, the seasonal period, and the number of upper time series of the temporal hierarchy. The rows and columns represent, respectively, the cross-sectional and the temporal dimension:

$$\boldsymbol{X}_\tau = \begin{bmatrix} \boldsymbol{x}'_{1,\tau} \\ \vdots \\ \boldsymbol{x}'_{n,\tau} \end{bmatrix} = \begin{bmatrix} \boldsymbol{U}_\tau^{[k_p]} & \boldsymbol{U}_\tau^{[k_p-1]} & \dots & \boldsymbol{U}_\tau^{[1]} \\ \boldsymbol{B}_\tau^{[k_p]} & \boldsymbol{B}_\tau^{[k_p-1]} & \dots & \boldsymbol{B}_\tau^{[1]} \end{bmatrix},$$

where for any fixed k, $\boldsymbol{U}_\tau^{[k]}$ is the $(n_a \times N_k)$ matrix grouping the upper time series, $\boldsymbol{B}_\tau^{[k]}$ is the $(n_b \times N_k)$ matrix grouping the bottom time series. For example, for the

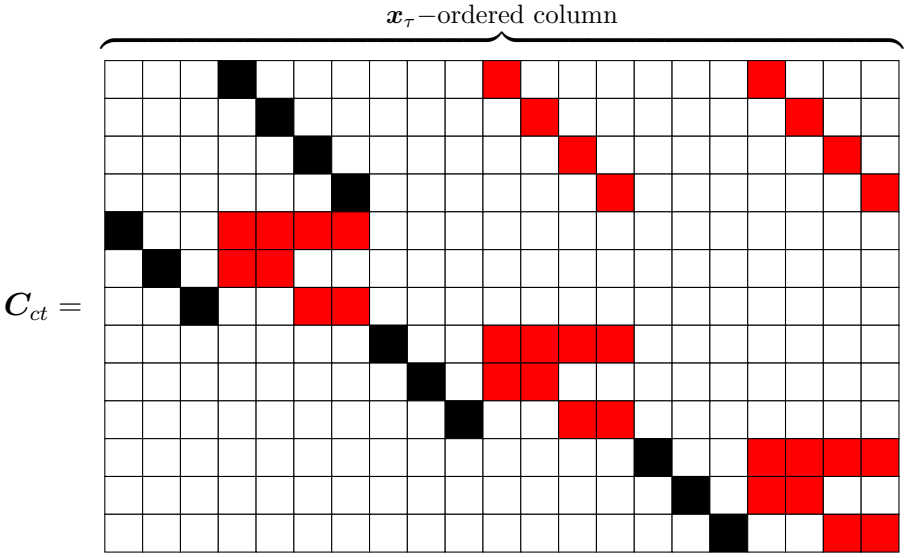


Figure 5.2: *Visual representation of the zero constraints cross-temporal matrix $\boldsymbol{C}_{ct}$ defined in (5.3) for a system of 3 linearly constrained quarterly time series (see Figure 5.1). The four upper rows describe the cross-sectional constraints (one for each quarter), the remaining rows the temporal constraints (one for each of the three time series). Colours legend: 0s in white, 1s in black, -1s in red.*

cross-temporal structure of Figure 5.1, we have

$$\boldsymbol{X}_\tau = \left[\begin{array}{c|cc|cccc} x^{[4]}_{T,\tau} & x^{[2]}_{T,2\tau-1} & x^{[2]}_{T,2\tau} & y_{T,4\tau-3} & y_{T,4\tau-2} & y_{T,4\tau-1} & y_{T,4\tau} \\ \hline x^{[4]}_{X,\tau} & x^{[2]}_{X,2\tau-1} & x^{[2]}_{X,2\tau} & y_{X,4\tau-3} & y_{X,4\tau-2} & y_{X,4\tau-1} & y_{X,4\tau} \\ x^{[4]}_{Y,\tau} & x^{[2]}_{Y,2\tau-1} & x^{[2]}_{Y,2\tau} & y_{Y,4\tau-3} & y_{Y,4\tau-2} & y_{Y,4\tau-1} & y_{Y,4\tau} \end{array}\right].$$

Further, $\boldsymbol{C}_{cs}\boldsymbol{X}_\tau = \boldsymbol{0}_{[n_a\times(m+k^*)]}$ and $\boldsymbol{C}_{te}\boldsymbol{X}'_\tau = \boldsymbol{0}_{(k^*\times n)}$. We can consider the cross-temporal framework as a generalization of the cross-sectional and temporal frameworks, that simultaneously takes into account both types of constraints. The cross-sectional reconciliation approach proposed by Hyndman *et al.* (2011) can be obtained by assuming $m = 1$, while the temporal one (Athanasopoulos *et al.*, 2017) is obtained when $n = 1$ (with $n_a = 0$ and $n_b = 1$).

Chapter 1 shows that the cross-temporal constraints working on the complete set of observations corresponding to time period τ can be expressed in a zero-constrained representation through the full rank $[(n_a m + nk^*) \times n(m + k^*)]$ zero constraints cross-temporal matrix $\boldsymbol{C}_{ct}$ such that

$$\boldsymbol{C}_{ct} = \begin{bmatrix} \boldsymbol{C}_* \\ \boldsymbol{I}_n \otimes \boldsymbol{C}_{te} \end{bmatrix} \quad \Longrightarrow \quad \boldsymbol{C}_{ct}\boldsymbol{x}_\tau = \boldsymbol{0}_{[(n_a m+nk^*)\times 1]} \quad \text{for} \quad \tau = 1, \dots, N, \tag{5.3}$$

where $\boldsymbol{x}_\tau = \text{vec}(\boldsymbol{X}'_\tau) = [\boldsymbol{x}'_{1,\tau}, \ \dots, \ \boldsymbol{x}'_{n,\tau}]'$, $\boldsymbol{C}_* = [\boldsymbol{0}_{(n_a m\times nk^*)} \ \ \boldsymbol{I}_m \otimes \boldsymbol{C}_{cs}]\boldsymbol{P}'$, $\boldsymbol{P}$ is the commutation matrix (Magnus and Neudecker, 2019, p. 54) such that $\boldsymbol{P}\text{vec}(\boldsymbol{X}_\tau) = \text{vec}(\boldsymbol{X}'_\tau)$,

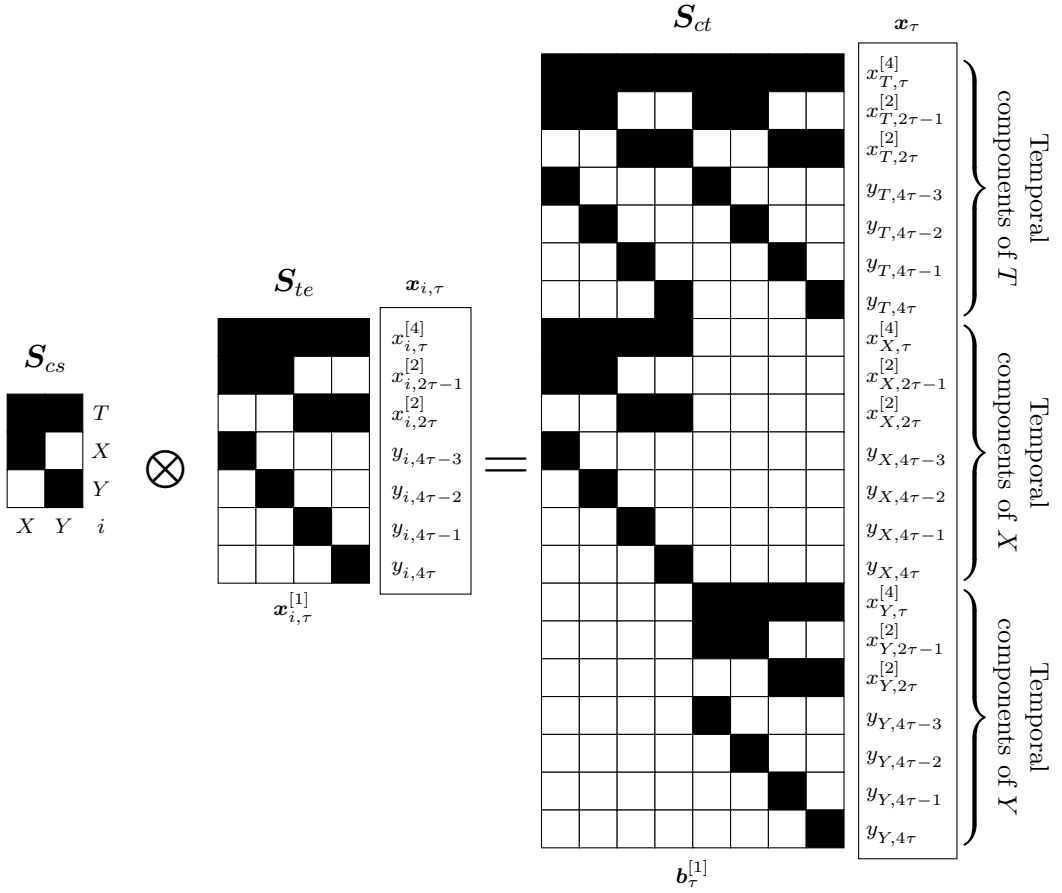


Figure 5.3: *Visual representation of the cross-temporal summation matrix* $\boldsymbol{S}_{ct} = \boldsymbol{S}_{cs} \otimes \boldsymbol{S}_{te}$ *defined in (5.4) for a system of 3 linearly constrained quarterly time series (see Figure 5.1). Colours legend: 0s in white, 1s in black.*

and the operator vec$(\cdot)$ converts a matrix into a vector. Figure 5.2 shows a visual example for the zero constraints cross-temporal matrix. A structural representation can be considered as well: $\boldsymbol{x}_\tau = \boldsymbol{S}_{ct}\boldsymbol{b}^{[1]}_\tau = s(\boldsymbol{b}^{[1]}_\tau)$, where

$$\boldsymbol{S}_{ct} = \boldsymbol{S}_{cs} \otimes \boldsymbol{S}_{te} \tag{5.4}$$

is the $[n(k^* + m) \times n_b m]$ cross-temporal summation matrix, $s : \mathbb{R}^{n_b m} \to \mathbb{R}^{n(m+k^*)}$ is the operator describing the pre-multiplication by $\boldsymbol{S}_{ct}$, and $\boldsymbol{b}^{[1]}_\tau = \text{vec}(\boldsymbol{B}^{[1]\prime}_\tau)$. In Figure 5.3, we have represented $\boldsymbol{S}_{ct}$ for a system of 3 linearly constrained quarterly time series (see Figure 5.1). In agreement with Panagiotelis *et al.* (2021), $\boldsymbol{x}_\tau$ lies in an $(n_b m)$-dimensional subspace $\mathfrak{s}_{ct}$ of $\mathbb{R}^{n(k^*+m)}$, which we refer to as the *cross-temporal coherent subspace*, spanned by the columns of $\boldsymbol{S}_{ct}$.

5.2.1 Optimal point forecast reconciliation

For $h = 1, \dots, H$, let

$$\widehat{\boldsymbol{X}}_h = \begin{bmatrix} \widehat{\boldsymbol{x}}'_{1,h} \\ \vdots \\ \widehat{\boldsymbol{x}}'_{n,h} \end{bmatrix} = \begin{bmatrix} \widehat{\boldsymbol{U}}_h^{[m]} & \dots & \widehat{\boldsymbol{U}}_h^{[k]} & \dots & \widehat{\boldsymbol{U}}_h^{[1]} \\ \widehat{\boldsymbol{B}}_h^{[m]} & \dots & \widehat{\boldsymbol{B}}_h^{[k]} & \dots & \widehat{\boldsymbol{B}}_h^{[1]} \end{bmatrix},$$

be the h-step ahead base forecasts, where $\widehat{\boldsymbol{U}}_h^{[k]}$ is the $(n_a \times M_k)$ matrix grouping the upper time series, $\widehat{\boldsymbol{B}}_h^{[k]}$ is the $(n_b \times M_k)$ matrix grouping the bottom time series for a given temporal aggregation order k and H is the forecast horizon for the most temporally aggregated time series. Based on the example in Figure 5.1 for $H = 1$, we have that

$$\widehat{\boldsymbol{X}}_1 = \left[\begin{array}{c|cc|cccc} \widehat{x}_{T,1}^{[4]} & \widehat{x}_{T,1}^{[2]} & \widehat{x}_{T,2}^{[2]} & \widehat{y}_{T,1} & \widehat{y}_{T,2} & \widehat{y}_{T,3} & \widehat{y}_{T,4} \\ \hline \widehat{x}_{X,1}^{[4]} & \widehat{x}_{X,1}^{[2]} & \widehat{x}_{X,2}^{[2]} & \widehat{y}_{X,1} & \widehat{y}_{X,2} & \widehat{y}_{X,3} & \widehat{y}_{X,4} \\ \widehat{x}_{Y,1}^{[4]} & \widehat{x}_{Y,1}^{[2]} & \widehat{x}_{Y,2}^{[2]} & \widehat{y}_{Y,1} & \widehat{y}_{Y,2} & \widehat{y}_{Y,3} & \widehat{y}_{Y,4} \end{array}\right].$$

The matrix $\widehat{\boldsymbol{X}}_h$, contains incoherent forecasts, such as $\boldsymbol{C}_{ct}\widehat{\boldsymbol{x}}_h \neq \boldsymbol{0}_{[(n_a m + n k^*)\times 1]}$ with $h = 1, \dots, H$ and $\widehat{\boldsymbol{x}}_h = \text{vec}(\widehat{\boldsymbol{X}}'_h)$. In this framework, the definition for forecast reconciliation in the cross-sectional framework given by Panagiotelis *et al.* (2021) can be generalized as follows.

Definition 5.1. Forecast reconciliation adjusts the base forecast $\widehat{\boldsymbol{x}}_h$ by finding a mapping $\psi : \mathbb{R}^{n(m+k^*)} \to \mathfrak{s}$ such that $\widetilde{\boldsymbol{x}}_h = \psi(\widehat{\boldsymbol{x}}_h)$, where $\widetilde{\boldsymbol{x}}_h \in \mathfrak{s}$ is the vector of the reconciled forecasts.

For a given forecast horizon $h = 1, \dots, H$, the mapping ψ may be defined as a projection onto $\mathfrak{s}$ given by (Panagiotelis *et al.*, 2021, and see also Chapter 1)

$$\widetilde{\boldsymbol{x}}_h = \psi(\widehat{\boldsymbol{x}}_h) = \boldsymbol{M}\widehat{\boldsymbol{x}}_h, \tag{5.5}$$

where $\boldsymbol{M} = \boldsymbol{I}_{n(m+k^*)} - \boldsymbol{\Omega}_{ct}\boldsymbol{C}'_{ct}\left(\boldsymbol{C}_{ct}\boldsymbol{\Omega}_{ct}\boldsymbol{C}'_{ct}\right)^{-1}\boldsymbol{C}_{ct}$, for a positive definite matrix $\boldsymbol{\Omega}_{ct}$, and $\widetilde{\boldsymbol{x}}_h = \text{vec}(\widetilde{\boldsymbol{X}}'_h)$. Wickramasuriya *et al.* (2019) showed that the minimum variance linear unbiased reconciled forecasts, satisfying the unbiasedness condition $\text{E}(\widetilde{\boldsymbol{x}}_h - \boldsymbol{x}_h) = 0$, has solution (5.5) when $\boldsymbol{\Omega}_{ct} = \text{Var}(\widehat{\boldsymbol{x}}_h - \boldsymbol{x}_h)$.

Alternatively, the cross-temporal reconciled forecasts $\widetilde{\boldsymbol{X}}_h$ may be found according to the structural approach proposed by Hyndman *et al.* (2011) for the cross-sectional framework, yielding $\widetilde{\boldsymbol{x}}_h = \boldsymbol{S}_{ct}\boldsymbol{G}\widehat{\boldsymbol{x}}_h$ for some matrix $\boldsymbol{G}$. Wickramasuriya *et al.* (2019) showed that this leads to a solution equivalent to the cross-temporally reconciled forecasts in (5.5), given by

$$\widetilde{\boldsymbol{x}}_h = \psi(\widehat{\boldsymbol{x}}_h) = (s \circ g)(\widehat{\boldsymbol{x}}_h) = \boldsymbol{S}_{ct}\boldsymbol{G}\widehat{\boldsymbol{x}}_h, \tag{5.6}$$

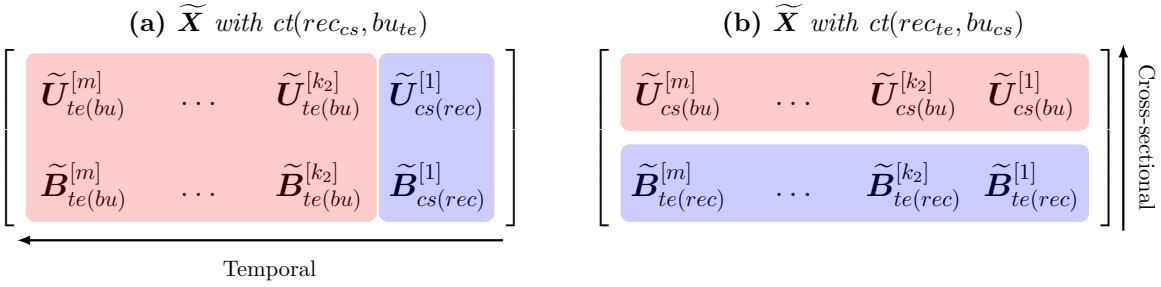


Figure 5.4: *A visual representation of partly bottom up starting from* (5.4a) *cross-sectionally reconciled forecasts for the temporal order 1 ($\widetilde{\boldsymbol{U}}^{[1]}$ and $\widetilde{\boldsymbol{B}}^{[1]}$) followed by temporal bottom-up, and* (5.4b) *temporally reconciled forecasts of the cross-sectional bottom time series ($\widetilde{\boldsymbol{B}}^{[k]}$, $k \in \mathcal{K}$) followed by cross-sectional bottom-up. The blue background indicates generating reconciled forecasts along one dimension, while the pink background indicates the forecasts obtained using bottom-up along the other.*

where $\boldsymbol{G} = (\boldsymbol{S}'_{ct}\boldsymbol{\Omega}^{-1}_{ct}\boldsymbol{S}_{ct})^{-1}\boldsymbol{S}'_{ct}\boldsymbol{\Omega}^{-1}_{ct}$, and $\boldsymbol{M} = \boldsymbol{S}_{ct}\boldsymbol{G}$. In this case, ψ is the composition of two transformations, say $s \circ g$, where $g : \mathbb{R}^{n(m+k^*)} \to \mathbb{R}^{n_b m}$ is a continuous function. In the online appendix A we report some cross-sectional, temporal and cross-temporal approximations for the covariance matrix to be used in (5.5) and (5.6).

5.2.2 Cross-temporal bottom-up forecast reconciliation

The classic bottom-up approach (Dunn *et al.*, 1976; Dangerfield and Morris, 1992) simply consists in summing-up the base forecasts of the most disaggregated level in the hierarchy to obtain forecasts of the upper-level series. To reduce the computational cost involved in optimal cross-temporal reconciliation, we may be interested in applying a reconciliation along only one dimension (cross-sectional or temporal) and reconstructing the cross-temporal structure using a partly bottom-up approach (Sanguri *et al.*, 2022, see also Chapter 2).

Figure 5.4 provides a visual representation of partly bottom-up in a two-step cross-temporal reconciliation approach. On the left (Figure 5.4a), we first compute the cross-sectionally reconciled forecasts at the highest frequency ($k = 1$), and then apply temporal bottom-up to obtain coherent cross-temporal forecasts. On the right (Figure 5.4b), we first compute temporally reconciled forecasts for the most disaggregated cross-sectional level, and then apply the cross-sectional bottom-up. We denote these two-step reconciliation approaches, respectively, as ct(rec_{te}, bu_{cs}), and ct(rec_{cs}, bu_{te}), where 'rec_{te}' and 'rec_{cs}' denote a forecast reconciliation approach in the temporal and

cross-sectional dimensions and, 'bu_{cs}' and 'bu_{te}' denote using bottom-up in the cross-sectional and temporal dimensions, respectively. It is worth noting that the simple cross-temporal bottom-up approach corresponds to $\mathrm{ct}(bu_{cs}, bu_{te}) = \mathrm{ct}(bu_{te}, bu_{cs}) = \mathrm{ct}(bu)$.

5.3 Probabilistic forecast reconciliation

To introduce the idea of coherence and probabilistic forecast reconciliation, we adapt the notations and the formal definitions introduced in Wickramasuriya (2024) and Panagiotelis *et al.* (2023) for the cross-sectional probabilistic case. These definitions can also be generalized to the cross-temporal framework by following the approach developed by Corani *et al.* (2023) for count data. However, in this chapter we only focus on the continuous case.

Our aim is to extend these definitions to *cross-temporal coherent probabilistic forecasts* and *cross-temporal probabilistic forecast reconciliation*. Let $(\mathbb{R}^{n_b m}, \mathcal{F}_{\mathbb{R}^{n_b m}}, \nu)$ be a probability space for the bottom time series $\boldsymbol{b}_\tau^{[1]}$, where $\mathcal{F}_{\mathbb{R}^{n_b m}}$ is the Borel σ-algebra on $\mathbb{R}^{n_b m}$. Then a σ-algebra $\mathcal{F}_{\mathfrak{s}}$ can be constructed from the collection of sets $s(\mathcal{B})$ for all $\mathcal{B} \in \mathcal{F}_{\mathbb{R}^{n_b m}}$.

Definition 5.2 (Cross-temporal coherent probabilistic forecasts)**.** Given the probability space $(\mathbb{R}^{n_b m}, \mathcal{F}_{\mathbb{R}^{n_b m}}, \nu)$, we define the coherent probability space as the triple $(\mathfrak{s}, \mathcal{F}_{\mathfrak{s}}, \breve{\nu})$ satisfying the following property: $\breve{\nu}(s(\mathcal{B})) = \nu(\mathcal{B})$, $\forall \mathcal{B} \in \mathcal{F}_{\mathbb{R}^{n_b m}}$.

Let $(\mathbb{R}^{n(m+k^*)}, \mathcal{F}_{\mathbb{R}^{n(m+k^*)}}, \hat{\nu})$ be a probability space referring to the incoherent probabilistic forecast $(\widehat{\boldsymbol{x}}_h)$ for all the n series in the system at any temporal aggregation order $k \in \mathcal{K}$.

Definition 5.3 (Cross-temporal probabilistic forecast reconciliation)**.** The reconciled probability measure of $\hat{\nu}$ with respect to ψ is a probability measure $\tilde{\nu}$ on $\mathfrak{s}$ with σ-algebra $\mathcal{F}_{\mathfrak{s}}$ satisfying

$$\tilde{\nu}(\mathcal{A}) = \hat{\nu}(\psi^{-1}(\mathcal{A})), \quad \forall \mathcal{A} \in \mathcal{F}_{\mathfrak{s}}, \tag{5.7}$$

where $\psi^{-1}(\mathcal{A}) = \{x \in \mathbb{R}^{n(m+k^*)} : \psi(x) \in \mathcal{A}\}$ denotes the pre-image of $\mathcal{A}$.

The map ψ may be obtained as the composition $s \circ g$, as for the cross-temporal point reconciliation (5.6).

Theorem 5.4 (Cross-temporal reconciled samples)**.** *Suppose that* $(\widehat{\boldsymbol{x}}_1, \dots, \widehat{\boldsymbol{x}}_L)$ *is a sample drawn from a (cross-temporal) incoherent probability measure* $\widehat{\nu}$*. Then* $(\widetilde{\boldsymbol{x}}_1, \dots, \widetilde{\boldsymbol{x}}_L)$*, where* $\widetilde{\boldsymbol{x}}_\ell = \psi(\widehat{\boldsymbol{x}}_\ell)$ *and* $\ell = 1, \dots, L$*, is a sample drawn from the (cross-temporal) reconciled probability measure* $\widetilde{\nu}$ *defined in* (5.7)*.*

Proof. See Theorem 4.5 in Panagiotelis *et al.* (2023) using Definition 5.3. □

Theorem 5.4 is the cross-temporal extension of Theorem 4.5 in Panagiotelis *et al.* (2023), valid only for the cross-sectional case. It means that a sample from the reconciled distribution can be obtained by reconciling each member of a sample from the incoherent distribution. With this result, we can separate the mechanism used to generate the base forecasts samples from the reconciliation phase.

5.3.1 Parametric framework: Gaussian reconciliation

It is possible to obtain a reconciled probabilistic forecast analytically for some parametric distributions, such as the multivariate normal (Corani *et al.*, 2021; Eckert *et al.*, 2021; Panagiotelis *et al.*, 2023; Wickramasuriya, 2024). In the cross-sectional framework, Panagiotelis *et al.* (2023) show that, starting from an elliptical distribution for the base forecasts, the reconciled forecast distribution is also elliptical. Using the results shown in Section 5.2, we extend[2] this results to the cross-temporal case. To obtain a reconciled forecast using the multivariate normal distribution, we start with a base forecast distributed as $\mathcal{N}(\widehat{\boldsymbol{x}}, \boldsymbol{\Omega})$, where $\widehat{\boldsymbol{x}}$ is the mean vector and $\boldsymbol{\Omega}$ is the covariance matrix of the base forecasts. Using standard results for the Gaussian case, the reconciled forecast distribution is given by $\mathcal{N}(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{\Omega}})$, where

$$\widetilde{\boldsymbol{x}} = \boldsymbol{M}\widehat{\boldsymbol{x}} \quad \text{and} \quad \widetilde{\boldsymbol{\Omega}} = \boldsymbol{M}\boldsymbol{\Omega}\boldsymbol{M}', \tag{5.8}$$

where $\boldsymbol{M}$ is the projection matrix defined in (5.5). Note that if we assume that $\boldsymbol{\Omega} = \boldsymbol{\Omega}_{ct}$ (see the projection matrices in (5.5) and (5.6)), then the covariance matrix in (5.8) simplifies to $\widetilde{\boldsymbol{\Omega}} = \boldsymbol{M}\boldsymbol{\Omega}_{ct}$. In the cross-temporal case, sensibly estimating the covariance matrix $\boldsymbol{\Omega}$ can be difficult because we need to simultaneously consider both the temporal and cross-sectional structures. This requires many parameters to be estimated, which can be challenging in practice. Additionally, naively using one-step residuals to estimate the cross-temporal correlation structure can lead to an inappropriate estimate of the covariance matrix[3]. These challenges will be explored in more depth in the following sections.

Focusing on the computational aspect[4], we can take several steps to reduce the time required to obtain simulations from the reconciled forecast distribution. For example when dealing with a genuine hierarchical structure, it is not necessary to simulate from a normal distribution with a defined covariance matrix for the entire structure. Instead,

[2] We assume $H = 1$ and simplify the notation by removing the h suffix without loss of generality

[3] In particular, some temporal covariances are fixed to zero (see the online appendix C for more details).

[4] We use two R packages to sample from a the base forecast Gaussian distribution: `MASS` (Venables and Ripley, 2002) and `Rfast` (Papadakis *et al.*, 2022) in Sections 5.5 and 5.6, respectively.

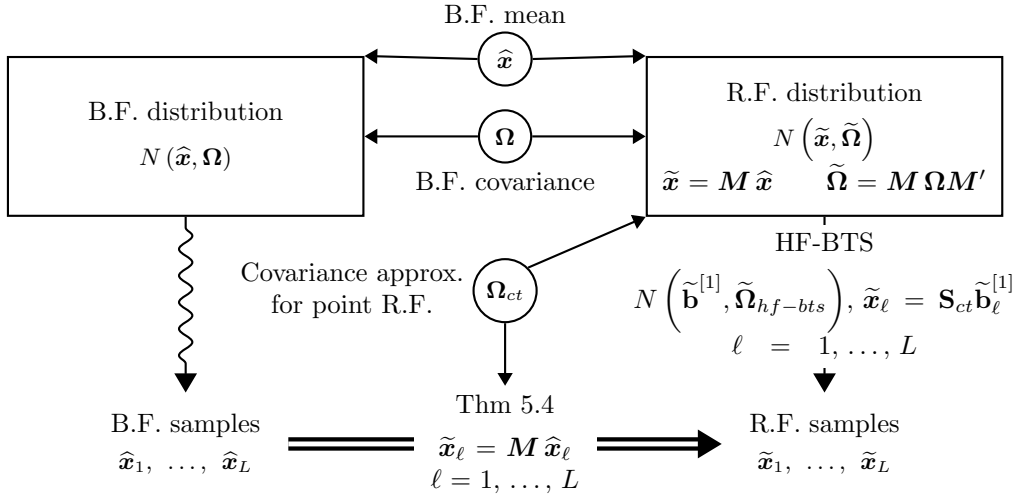


Figure 5.5: *Overview of cross-temporal forecast reconciliation in the Gaussian framework: two different but equivalent ways of obtaining reconciled forecast samples, as described in Section 5.3.1. The acronyms R.F and B.F. stand for Reconciled and Base Forecasts, respectively. HF-BTS stands for High Frequency Bottom Time Series.*

we can utilize the properties of elliptical distributions to simulate from the high frequency bottom time series and then obtain the complete simulation through the $\boldsymbol{S}_{ct}$ matrix. Furthermore, we do not need to calculate the reconciled mean and variance and generate a new sample if we already have a sample from the normal distribution of the base forecasts; we can simply apply the point forecast reconciliation (5.5) as outlined in Theorem 5.4. Figure 5.5 shows two different but equivalent ways of obtaining reconciled forecast samples: the former from the base distribution through the Theorem 5.4, and the latter from the reconciled distribution through the high frequency bottom time series forecasts $\widetilde{\boldsymbol{b}}^{[1]}$ only. The two rectangles represent the base and reconciled forecast distributions, respectively. Enclosed within circles are the distribution parameters involved in the point forecast reconciliation process, transforming $\widehat{\boldsymbol{x}}$ into $\widetilde{\boldsymbol{x}}$ and $\boldsymbol{\Omega}$ into $\widetilde{\boldsymbol{\Omega}}$. The wave-like arrows represent the simulation processes, generating both base and reconciled forecast samples. Finally, the bold double arrow "⇒" illustrates the generation of the reconciled forecast distributions as described in Theorem 5.4.

5.3.2 Non-parametric framework: bootstrap reconciliation

Analytical expressions for the base and reconciled forecast distributions are sometimes challenging to obtain. Furthermore parametric assumptions can be restrictive and unrealistic. We propose a procedure called *cross-temporal joint (block) bootstrap* (**ctjb**) to generate samples from the base forecast distributions that preserve cross-temporal relationships. This approach involves drawing samples of all series simultaneously from

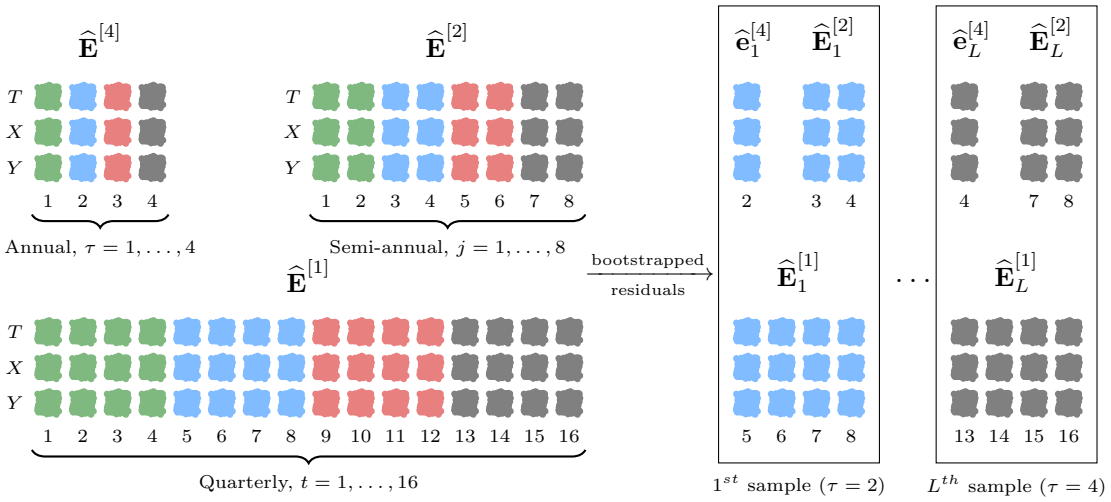


Figure 5.6: *Example of bootstrapped residuals for 3 linearly constrained quarterly time series (see Figure 5.1). On the left there are the residual matrices with 4 years of data ($N = 4$): the green, blue, red and black colors correspond, respectively, to years 1, 2, 3 and 4. On the right the bootstrapped residuals are represented.*

the most temporally aggregated level, and using the most temporally aggregated level to determine the corresponding time indices for the other levels.

Let $\widehat{\boldsymbol{E}}^{[k]}$ be the $(n \times N_k)$ matrix of the residuals for $k \in \mathcal{K}$. Figure 5.6 (on the left) provides a visualization of these matrices and how they are related to each other for the example in Figure 5.1. It is assumed that the residuals cover four years ($N = 4$): the green color corresponds to the first year, the blue to the second year, and so on. Further, let $\mathcal{M}_i$ be the model used to calculate the base forecasts and residuals for the i^{th} series. Assuming $H = 1$, τ is a random draw with replacement from $1, \ldots, N$ and the ℓ^{th} bootstrap incoherent sample is $\widehat{\boldsymbol{x}}_{i,\ell}^{[k]} = f_i(\mathcal{M}_i, \widehat{\boldsymbol{e}}_i^{[k]})$, where $f_i(\cdot)$ depends on the fitted model $\mathcal{M}_i$. That is, $\widehat{\boldsymbol{x}}_{i,l}^{[k]}$ is a sample path simulated for the i^{th} series with error approximated by the corresponding block bootstrapped sample residual $\widehat{\boldsymbol{e}}_i^{[k]}$, the i^{th} row of

$$\widehat{\boldsymbol{E}}_{\tau}^{[k]} = \begin{bmatrix} \widehat{e}_{1,M_k(\tau-1)+1}^{[k]} & \cdots & \widehat{e}_{1,M_k\tau}^{[k]} \\ \vdots & \ddots & \vdots \\ \widehat{e}_{n,M_k(\tau-1)+1}^{[k]} & \cdots & \widehat{e}_{n,M_k\tau}^{[k]} \end{bmatrix} \qquad k \in \mathcal{K}.$$

Figure 5.6 (on the right) shows $\widehat{\boldsymbol{E}}_{\tau}^{[k]}$ for the quarterly cross-temporal hierarchy in Figure 5.1.

One of the main advantages of the cross-temporal joint bootstrap is that it allows us to accurately account for the dependence between the different levels of temporal aggregation and not only the cross-sectional dependencies. By sampling residuals from the most temporally aggregated level and using it to determine the indices for the other

levels, we can ensure that the bootstrap sample reflects the underlying data distribution. Additionally, the cross-temporal joint bootstrap is easy to implement for many forecasting models, making it a practical and efficient tool. Furthermore, this approach is easily scalable in order to utilize multiple computing power simultaneously for each individual series. This can be especially useful when dealing with large datasets or when trying to speed up the analysis process.

5.4 Cross-temporal covariance matrix estimation

As the covariance matrix $\boldsymbol{\Omega}$ is unknown in practice, a natural estimate is the empirical sample covariance matrix of the base forecasts $\widehat{\boldsymbol{\Omega}}$. In this section, our focus will be exclusively on the cross-temporal framework., this means that we have to estimate $r = n(k^*+m)[n(k^*+m)-1]/2$ different parameters. A possible solution to estimating many parameters when we have fewer observations than r, is to construct a shrinkage estimator (Efron, 1975; Efron and Morris, 1975, 1977), using a convex combination of $\widehat{\boldsymbol{\Omega}}$ and a diagonal target matrix $\widehat{\boldsymbol{\Omega}}_D = \widehat{\boldsymbol{\Omega}} \odot \boldsymbol{I}_{n(k^*+m)}$, such that $\widehat{\boldsymbol{\Omega}}_G = \lambda\widehat{\boldsymbol{\Omega}}_D + (1-\lambda)\widehat{\boldsymbol{\Omega}}$, where $\lambda \in [0,1]$ is the shrinkage intensity parameter that can be estimate using the unbiased estimator proposed by Ledoit and Wolf (2004) (see Schäfer and Strimmer, 2005). The linear combination involving these two matrices is referred to as *Global shrinkage* (G), where all off-diagonal elements are shrunk towards zero. $\widehat{\boldsymbol{\Omega}}_G$ corresponds to the matrix used by the reconciliation approach oct(shr) (see Chapter 1). However, shrinking all off-diagonal elements to zero, when we know that the covariance matrix has a cross-sectional and/or temporal structure, results in information loss. Therefore, we propose to estimate a smaller matrix, and to use the cross-sectional and/or temporal structure to obtain a better estimator for the covariance matrix of the entire system. Given that $\boldsymbol{S}_{ct} = \boldsymbol{S}_{cs} \otimes \boldsymbol{S}_{te}$, it is possible to express the actual covariance matrix in terms of three smaller matrices such that

$$\begin{aligned}\widetilde{\boldsymbol{\Omega}} &= \boldsymbol{S}_{ct}\boldsymbol{\Omega}_{hf\text{-}bts}\boldsymbol{S}_{ct}' \\ &= (\boldsymbol{I}_n \otimes \boldsymbol{S}_{te})\,\boldsymbol{\Omega}_{hf}\,(\boldsymbol{I}_n \otimes \boldsymbol{S}_{te})' \\ &= (\boldsymbol{S}_{cs} \otimes \boldsymbol{I}_{m+k^*})\,\boldsymbol{\Omega}_{bts}\,(\boldsymbol{S}_{cs} \otimes \boldsymbol{I}_{m+k^*})',\end{aligned} \tag{5.9}$$

where $\boldsymbol{\Omega}_{hf\text{-}bts}$ is the $(n_b m \times n_b m)$ covariance matrix for the bottom time series at temporal aggregation level $k=1$ (highest frequency bottom time series), $\boldsymbol{\Omega}_{hf}$ is the $(nm \times nm)$ covariance matrix related to all the high frequency time series and $\boldsymbol{\Omega}_{bts}$ is the $[n_b(k^*+m) \times n_b(k^*+m)]$ covariance matrix related to bottom time series at any temporal aggregation. Equation (5.9) offers three decompositions of the covariance matrix $\widetilde{\boldsymbol{\Omega}}$, each characterized by well-defined structures: $\boldsymbol{S}_{ct}$ capturing cross-temporal, $\boldsymbol{I}_n \otimes \boldsymbol{S}_{te}$

Method	# of different parameters	GDP	Tourism
G	$r = \frac{n(k^*+m)[n(k^*+m)-1]}{2}$	221 445	108 052 350
B	$r_{HB} < \frac{n_b(k^*+m)[n_b(k^*+m)-1]}{2} < r$	94 395 (57%)	36 231 328 (66%)
H	$r_{HB} < \frac{nm[nm-1]}{2} < r$	72 390 (67%)	19 848 150 (82%)
HB	$r_{HB} = \frac{n_b m[n_b m-1]}{2} < r$	30 876 (86%)	6 655 776 (94%)

Table 5.1: *Number of different parameters that need to be estimated for the Australian GDP (see Section 5.5) and the Australian Tourism Demand (see Section 5.6) forecasting experiments. The percentage reductions in the number of parameters compared to the global approach G are reported in parentheses.*

temporal, and $\boldsymbol{S}_{cs} \otimes \boldsymbol{I}_{m+k^*}$ cross-sectional relationships. At the same time, each involves smaller covariance matrices as $\boldsymbol{\Omega}_{hf\text{-}bts}$, $\boldsymbol{\Omega}_{hf}$, and $\boldsymbol{\Omega}_{bts}$. Starting from these representations, we propose three different approaches (*HB*, *H*, and *B*, respectively) to approximate $\widetilde{\boldsymbol{\Omega}}$.

Therefore, we can apply the idea of "Stein-type shrinkage" (Efron and Morris, 1977) to $\boldsymbol{\Omega}_{hf\text{-}bts}$, $\boldsymbol{\Omega}_{hf}$ and $\boldsymbol{\Omega}_{bts}$ by using the corresponding empirical base forecasts residuals estimation. We obtain the following expressions (see the online appendix B for details):

- *High frequency Bottom time series shrinkage matrix* (*HB*):
  $\widehat{\boldsymbol{\Omega}}_{HB} = \lambda \boldsymbol{S}_{ct} \widehat{\boldsymbol{\Omega}}_{hf\text{-}bts,D} \boldsymbol{S}'_{ct} + (1-\lambda) \boldsymbol{S}_{ct} \widehat{\boldsymbol{\Omega}}_{hf\text{-}bts} \boldsymbol{S}'_{ct}$;
- *High frequency shrinkage matrix* (*H*):
  $\widehat{\boldsymbol{\Omega}}_{H} = \lambda (\boldsymbol{I}_n \otimes \boldsymbol{S}_{te}) \widehat{\boldsymbol{\Omega}}_{hf,D} (\boldsymbol{I}_n \otimes \boldsymbol{S}_{te})' + (1-\lambda)(\boldsymbol{I}_n \otimes \boldsymbol{S}_{te}) \widehat{\boldsymbol{\Omega}}_{hf} (\boldsymbol{I}_n \otimes \boldsymbol{S}_{te})'$;
- *Bottom time series shrinkage matrix* (*B*):
  $\widehat{\boldsymbol{\Omega}}_{B} = \lambda \left(\boldsymbol{S}_{cs} \otimes \boldsymbol{I}_{m+k^*}\right) \widehat{\boldsymbol{\Omega}}_{bts,D} \left(\boldsymbol{S}_{cs} \otimes \boldsymbol{I}_{m+k^*}\right)' + (1-\lambda) \left(\boldsymbol{S}_{cs} \otimes \boldsymbol{I}_{m+k^*}\right) \widehat{\boldsymbol{\Omega}}_{bts} \left(\boldsymbol{S}_{cs} \otimes \boldsymbol{I}_{m+k^*}\right)'$,

where $\widehat{\boldsymbol{\Omega}}_{l,D} = \boldsymbol{I}_{n_b m} \odot \widehat{\boldsymbol{\Omega}}_j$, $l = \{hf\text{-}bts,\ hf,\ bts\}$, and λ is the shrinkage parameter. These matrices are not full rank, meaning their inverses, needed to compute the projection to the coherent subspace, do not exist. To address this, a ridge regularization of the form $\widehat{\boldsymbol{\Omega}} + \omega \boldsymbol{I}$ was used (Marquardt, 1970), where ω is chosen to make the matrix invertible without introducing excessive bias. Figure 5.7 gives some visual insights on the covariance matrices obtainable with $\lambda = 0$ and $\lambda = 1$, respectively, for a simple cross-temporal hierarchical structure with 3 time series and $\mathcal{K} = \{4, 2, 1\}$ (see Figure 5.1).

Another important aspect is the number of parameters to be estimated through the residuals of the base forecasts. In Table 5.1 we report the number of different parameters for the two forecasting experiment: Australian GDP (see Section 5.5) and Australian

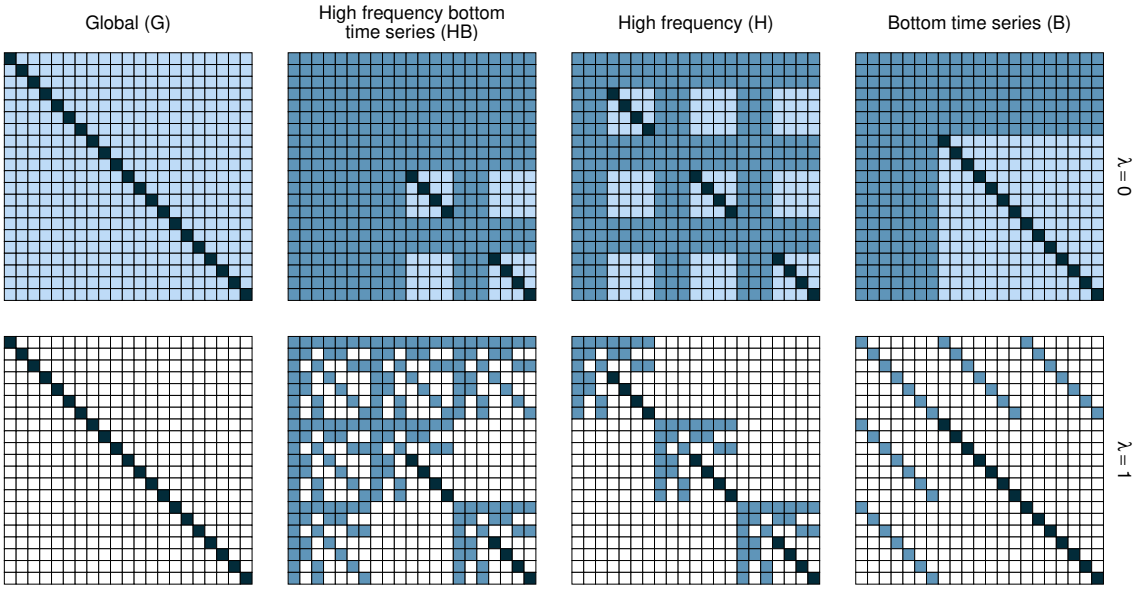


Figure 5.7: *Representation of four types of covariance matrices that can be obtained from the cross-temporal hierarchical structure (example based on the quarterly series of Figure 1) for two different values of* $\lambda \in \{0, 1\}$*, the shrinkage parameter. The entries in black are not modified by shrinkage, the entries in light blue are those actively involved in the shrinkage phase, while the entries in darker blue are derived directly from the cross-sectional and/or temporal structure and hence not estimated. For* $\lambda = 1$*, the white entries correspond to a zero value.*

Tourism Demand (see Section 5.6). In addition, we also calculate the percentage reductions in the number of parameters compared to the global approach. As we can see, G involves a considerably large number of parameters compared to other estimators. HB leads to the largest decrease of around 85%, whereas approaches H and B lie somewhere between G and HB. In general, as m and n increase, using H requires the estimation of less parameters than B.

It is worth noting that when using the HB covariance matrix, we make the assumption that the base error covariance matrix is coherent. This assumption is valid provided the base forecasts also approximately fulfil constraints (5.3), which is expected for any reasonable set of forecasts. In addition, with this covariance matrix, the computational complexity of the reconciliation phase is reduced. Specifically, Theorem 5.5 extends Theorem 1 in Hyndman *et al.* (2011), showing that reconciling using a coherent covariance matrix simplifies to the *ols* approach.

Theorem 5.5. *Let* $\widehat{\boldsymbol{\Omega}}_{hf-bts}$ *be a* $[(n_b m) \times (n_b m)]$ *positive-definite matrix. Then, using* $\boldsymbol{\Omega}_{ct} = \boldsymbol{S}_{ct}\widehat{\boldsymbol{\Omega}}_{hfbts}\boldsymbol{S}'_{ct}$ *in the reconciliation formulae (5.5) and (5.6) is equivalent to using* $\boldsymbol{\Omega}_{ct} = \boldsymbol{I}_{n(m+k^*)}$ *(ols approach).*

Proof. See the online appendix B. □

In the forecasting experiments that follow (and in the simulation in the online appendix C), we closely analyze these different constructions with a dual purpose. In particular, we use the full covariance matrix ($\lambda = 0$) of the base forecasts to obtain base forecast samples of the linearly constrained time series under Gaussianity. We also use the shrinkage versions as approximations of the covariance matrix to be used for reconciliation (excluding HB, see Theorem 5.5). This will allow us to better understand the properties and abilities of each parameterization.

5.4.1 Multi-step residuals

Model residuals may be used to estimate the covariance matrix in cross-temporal forecast reconciliation. In time series analysis, it is common to use residuals corresponding to one-step ahead forecasts. However, due to the temporal dimension in our setting, residuals corresponding to different forecast horizons are required. Thus, we define *multi-step residuals* as $e^{[k]}_{i,h,j} = x^{[k]}_{i,j+h} - \widehat{x}^{[k]}_{i,j+h|j}$, where $i = 1, \dots, n$, $j = 1, \dots, N_k$ and $\widehat{x}^{[k]}_{i,j+h|t}$ is the h-step fitted value, calculated as the h-step-ahead forecast using data up to time j. In general, these residuals will be autocorrelated except when $h = 1$.

Following Chapter 1, we use a matrix organization of the residuals similar to the one for the base forecasts in Section 5.2.1. Specifically, let N be the total number of observations for the most temporally aggregate time series. Then, the N_k-vectors of multi-step residuals for the temporal aggregation k and the series i, $\boldsymbol{e}^{[k]}_{i,h} = \begin{bmatrix} e^{[k]}_{i,h,1} & e^{[k]}_{i,h,2} & \dots & e^{[k]}_{i,h,N_k} \end{bmatrix}'$ with $h = 1, \dots, M_k$, can be organized in matrix form as

$$\boldsymbol{E}^{[k]}_{i} = \begin{bmatrix} e^{[k]}_{i,1,1} & e^{[k]}_{i,2,2} & \dots & e^{[k]}_{i,M_k,M_k} \\ \vdots & \vdots & & \vdots \\ e^{[k]}_{i,1,N_k-M_k+1} & e^{[k]}_{i,2,N_k-M_k+2} & \dots & e^{[k]}_{i,M_k,N_k} \end{bmatrix}.$$

Let $\boldsymbol{E}_i = \begin{bmatrix} \boldsymbol{E}^{[m]}_i & \boldsymbol{E}^{[k_p-1]}_i & \dots & \boldsymbol{E}^{[1]}_i \end{bmatrix}$. Then the $[N \times n(m+k^*)]$ cross-temporal residual matrix is given by $\boldsymbol{E} = \begin{bmatrix} \boldsymbol{E}_1 & \boldsymbol{E}_2 & \dots & \boldsymbol{E}_n \end{bmatrix}$.

To better understand the properties of the proposed alternatives, a simulation study was performed (the results are shown in the online appendix C). We have studied the effect of combining cross-sectional and temporal aggregations using a simple hierarchy, where the small size and nature of the data generating process make it possible to exactly calculate the true cross-temporal covariance structure, thus providing insights into the nature of the time series data involved in the forecast reconciliation process. We find that simulating base forecasts from multi-step residuals allows for a more accurate estimation of the covariance matrix and that reconciliation further improves the forecast accuracy.

5.4.2 Overlapping residuals

Another issue that arises in the case of cross-temporal reconciliation is the low number of available residuals, especially for the higher orders of temporal aggregation. A possible solution is to use residuals calculated using overlapping series by allowing the year to have a varying starting time. To better explain how to calculate overlapping residuals, assume we have a single series $\boldsymbol{y} = [y_1 \; y_2 \; y_3 \; \dots \; y_{T-1} \; y_T]'$. We can construct k non overlapping series such that $\boldsymbol{x}^{[k],s} = \left\{x_j^{[k],s}\right\}_{j=1}^{N_k-s}$ where

$$x_j^{[k],s} = \sum_{t=(j-1)k+s+1}^{jk-s} y_t,$$

with $s = 0, \dots, (k-1)$. For example, suppose we have a biannual series with $k = 2$ and $T = 6$, then we can construct two annual time series depending on which time is deemed the start of the year:

$$\boldsymbol{x}^{[2],0} = \begin{bmatrix} x_1^{[2],0} & x_2^{[2],0} & x_3^{[2],0} \end{bmatrix}' = \begin{bmatrix} y_1+y_2 & y_3+y_4 & y_5+y_6 \end{bmatrix}' \text{ and}$$
$$\boldsymbol{x}^{[2],1} = \begin{bmatrix} x_1^{[2],1} & x_2^{[2],1} \end{bmatrix}' = \begin{bmatrix} y_2+y_3 & y_4+y_5 \end{bmatrix}'.$$

To calculate overlapping residuals, we propose the following steps:

1. Fit a model to $\boldsymbol{x}^{[k],0}$ (i.e., select an appropriate model and estimate the model parameters using the available data) and calculate the residuals.
2. Apply the same model in step 1 to $\boldsymbol{x}^{[k],s}$ for $s = 1, \dots, k-1$, without re-estimating the parameters, and calculate the residuals.

The resulting residuals can be used to estimate the covariance matrix in cross-temporal forecast reconciliation. This increases the number of available residuals, particularly when working with higher frequency observations such as monthly or daily data. It is important to note that this approach assumes that the model used in step 1 is appropriate for all the different series $\boldsymbol{x}^{[k],s}$. Some seasonal models will not be appropriate as the seasonal pattern will be shifted for different values of s. However, this will not affect seasonal ARIMA models as the seasonality is defined in terms of lags which are unaffected by the value of s.

5.5 Forecasting Australian GDP

The Australian Quarterly National Accounts (QNA) dataset has been widely studied in the literature on forecast reconciliation (Athanasopoulos *et al.*, 2020; Bisaglia *et al.*, 2020). Building on these results, we now consider cross-temporally reconciled probabilistic forecasts.

Label	Description
$\text{ct}(shr_{cs}, bu_{te})$	Partly bottom-up (Section 5.2.2) starting from cross-sectional reconciled forecasts using the *shr* approach.
$\text{oct}(\cdot)$	Optimal cross-temporal reconciliation for the *struc*, *wlsv* and *bdshr* approaches. One-step residuals were used with *wlsv* and *bdshr*.
$\text{oct}_h(\cdot)$	Optimal cross-temporal reconciliation with multi-step residuals (see Section 5.4.1) for the approaches presented in Section 5.4: *hshr* for *High frequency shrinkage*, and *bshr* for *bottom time series shrinkage*.
$\text{oct}_o(\cdot)$	Optimal cross-temporal reconciliation with overlapping residuals (see Section 5.4.2) for the *wlsv* and *bdshr* approaches.
$\text{oct}_{oh}(hshr)$	Optimal cross-temporal reconciliation with overlapping and multi-step residuals (see Section 5.4.1 and 5.4.2) for the *hshr* (*High frequency shrinkage*) approach presented in Section 5.4.

Table 5.2: *Cross-temporal reconciliation approaches for the Australian GDP (see Section 5.5) and the Australian Tourism Demand (see Section 5.6) forecasting experiments. All the reconciliation procedures are available in* `FoReco` *(Girolimetto and Di Fonzo, 2023a).*

We use univariate ARIMA models[5] to obtain quarterly base forecasts for the $n = 95$ QNA time series, spanning the period 1984:Q4 – 2018:Q1, defining GDP from both the Income and Expenditure sides. We perform a rolling forecast experiment with an expanding window: the first training sample spans the period 1984:Q4 to 1994:Q3, and the last ends in 2017:Q1, for a total of 91 forecast origins. For the temporal aggregation dimension we aggregate the quarterly data to both semi-annual and annual. We obtain 4-step, 2-step and 1-step ahead base forecasts respectively from the quarterly, semi-annual and annual frequencies, i.e., $\mathcal{K} = \{4, 2, 1\}$.

The base forecast samples in the Gaussian case are obtained using the sample covariance matrices with the *Global* (G) and *High frequency* (H) parameterization (Section 5.4), since it is not possible to identify a unique representation for the other cases[6]. We compare the results obtained using multi-step residuals with and without overlapping, in order to measure the benefit of obtaining overlapping residuals. In the

[5]We use the `auto.arima` function from the R package `forecast` (Hyndman *et al.*, 2023).

[6]When simultaneously considering Income and Expenditure sides hierarchies, the result is a general linearly constrained time series, where bottom and upper time series are not uniquely defined, making unfeasible the cross-sectional bottom-up reconciliation approach (see Chapter 4).

Reconciliation approach	ctjb	G_h	H_h	G_{oh}	H_{oh}
base	29.01	2.21	2.17	2.14	2.18
$\text{ct}(shr_{cs}, bu_{te})$	+ 0.19	+ 0.13	+ 0.12	+ 0.30	+ 0.30
$\text{ct}(wls_{cs}, bu_{te})$	+ 0.21	+ 0.31	+ 0.31	+ 0.33	+ 0.35
$\text{oct}_o(wlsv)$	+ 0.25	+ 0.24	+ 0.22	+ 0.22	+ 0.22
$\text{oct}_o(bdshr)$	+ 0.48	+ 0.44	+ 0.45	+ 0.45	+ 0.45
$\text{oct}_{oh}(hshr)$	+ 0.64	+ 0.65	+ 0.64	+ 0.65	+ 0.64

Table 5.3: *Computational time (in seconds) for the first iteration of the Australian QNA forecasting experiment. The first row (base) reports the time to simulate 1000 samples, and the remaining rows the additional time to reconcile them with different approaches.*

non-parametric case, we use the cross-temporal joint bootstrap (ctjb) presented in Section 5.3.2. Finally, to reconcile the resulting (1000) base forecasts samples, we have applied the following techniques[7] (see Table 5.2): $\text{ct}(shr_{cs}, bu_{te})$, $\text{ct}(wls_{cs}, bu_{te})$, $\text{oct}_o(wlsv)$, $\text{oct}_o(bdshr)$, and $\text{oct}_{oh}(hshr)$.

The accuracy of the probabilistic forecasts is evaluated using the Continuous Ranked Probability Score (CRPS, Matheson and Winkler, 1976; Gneiting and Katzfuss, 2014), which is an index that considers the single series and provides us a marginal evaluation of the approaches. In addition, we employ the Energy Score (ES, Gneiting and Katzfuss, 2014), that is the CRPS extension to the multivariate case, to evaluate the forecasting accuracy for the whole system (Panagiotelis *et al.*, 2023; Wickramasuriya, 2024). In particular, we consider the geometric mean of the relative CRPS (Fleming and Wallace, 1986), and the relative ES:

$$\overline{\text{RelCRPS}}^{[k]}_{j,s} = \left(\prod_{i=1}^{n} \frac{CRPS^{[k]}_{i,j,s}}{CRPS^{[k]}_{i,0,0}}\right)^{\frac{1}{n}} \qquad \text{and} \qquad \text{RelES}^{[k]}_{j,s} = \frac{ES^{[k]}_{j,s}}{ES^{[k]}_{0,0}}, \tag{5.10}$$

where j denotes the reconciliation approach and s indicates the approach used to simulate the base forecasts. As a reference approach ($s = 0$ and $j = 0$), we consider the base forecasts produced by the Bootstrap approach. If we consider all the temporal aggregation orders (i.e. $\forall k \in \mathcal{K}$), the overall accuracy indices are given by, respectively,

$$\overline{\text{RelCRPS}}_{j,s} = \left(\prod_{\substack{i=1,\dots,n \\ k\in\mathcal{K}}} \frac{CRPS^{[k]}_{i,j,s}}{CRPS^{[k]}_{i,0,0}}\right)^{\frac{1}{n(k^*+m)}} \text{ and } \overline{\text{RelES}}_{j,s} = \left(\prod_{k\in\mathcal{K}} \frac{ES^{[k]}_{j,s}}{ES^{[k]}_{0,0}}\right)^{\frac{1}{(k^*+m)}}. \tag{5.11}$$

Tables 5.3 shows the runtime (in seconds) required for simulating 1000 samples (first row, base) and the additional time needed for reconciliation with various approaches in

[7]The results with shrunk covariance matrices are available in the online appendix D.2, where we also report the results obtained using other reconciliation approaches.

Reconciliation approach	**Base forecasts' sample approach**									
	ctjb	Gaussian approach*				ctjb	Gaussian approach*			
		G_h	H_h	G_{oh}	H_{oh}		G_h	H_h	G_{oh}	H_{oh}
					$\overline{RelCRPS}$					
		$\forall k \in \{4,2,1\}$					$k=1$			
base	1.000	0.979	0.995	0.968	0.976	1.000	0.988	0.988	0.971	0.971
ct(shr_{cs}, bu_{te})	0.937	0.956	0.956	0.976	0.976	0.992	1.008	1.008	1.029	1.029
ct(wls_{cs}, bu_{te})	0.930	0.917	0.917	0.898	0.898	0.986	0.974	0.975	0.956	0.956
oct$_o$($wlsv$)	**0.926**	**0.911**	**0.912**	**0.896**	**0.895**	**0.984**	**0.971**	**0.970**	**0.954**	**0.954**
oct$_o$($bdshr$)	0.978	0.964	0.946	0.952	0.930	1.034	1.016	1.003	1.005	0.989
oct$_{oh}$($hshr$)	1.006	0.983	1.009	0.974	1.001	1.068	1.046	1.059	1.034	1.061
		$k=2$					$k=4$			
base	1.000	0.984	0.993	0.968	0.976	1.000	0.966	1.004	0.964	0.981
ct(shr_{cs}, bu_{te})	0.949	0.966	0.966	0.987	0.987	0.874	0.896	0.896	0.914	0.914
ct(wls_{cs}, bu_{te})	0.942	0.928	0.928	0.909	0.909	0.866	0.853	0.853	0.834	0.834
oct$_o$($wlsv$)	**0.938**	**0.921**	**0.923**	**0.907**	**0.906**	**0.860**	**0.847**	**0.848**	**0.832**	**0.830**
oct$_o$($bdshr$)	0.991	0.974	0.957	0.964	0.942	0.914	0.905	0.883	0.892	0.865
oct$_{oh}$($hshr$)	1.021	0.996	1.021	0.987	1.016	0.934	0.912	0.951	0.904	0.931
					ES ratio indices					
		$\forall k \in \{4,2,1\}$					$k=1$			
base	1.000	0.970	0.988	0.960	0.970	1.000	0.977	0.977	0.965	0.965
ct(shr_{cs}, bu_{te})	0.897	0.944	0.944	0.973	0.973	0.964	1.001	1.001	1.033	1.033
ct(wls_{cs}, bu_{te})	**0.886**	0.880	**0.880**	**0.860**	**0.860**	**0.954**	**0.944**	**0.945**	**0.928**	**0.928**
oct$_o$($wlsv$)	0.891	**0.879**	0.881	0.864	0.864	0.958	0.945	0.945	0.931	0.931
oct$_o$($bdshr$)	0.940	0.928	0.910	0.918	0.895	1.004	0.986	0.971	0.980	0.961
oct$_{oh}$($hshr$)	0.986	0.968	0.999	0.959	0.992	1.053	1.034	1.049	1.024	1.055
		$k=2$					$k=4$			
base	1.000	0.972	0.985	0.959	0.969	1.000	0.959	1.000	0.957	0.976
ct(shr_{cs}, bu_{te})	0.915	0.961	0.960	0.991	0.991	0.818	0.874	0.874	0.899	0.900
ct(wls_{cs}, bu_{te})	**0.904**	0.896	**0.896**	**0.877**	**0.877**	**0.807**	0.805	**0.805**	**0.782**	**0.783**
oct$_o$($wlsv$)	0.908	**0.895**	0.898	0.881	0.882	0.812	**0.802**	0.806	0.786	0.786
oct$_o$($bdshr$)	0.960	0.947	0.929	0.938	0.915	0.860	0.856	0.836	0.841	0.816
oct$_{oh}$($hshr$)	1.007	0.988	1.017	0.979	1.014	0.904	0.888	0.934	0.881	0.913

*The Gaussian method employs a sample covariance matrix:
G_h and H_h use multi-step residuals and G_{oh} and H_{oh} use overlapping and multi-step residuals.

Table 5.4: *$\overline{RelCRPS}$ and ES ratio indices defined in* (5.10) *and* (5.11) *for the Australian QNA dataset. Approaches performing worse than the benchmark (bootstrap base forecasts, ctjb) are highlighted in red, the best for each column is marked in bold, and the overall lowest value is highlighted in blue. The reconciliation approaches are described in Table 5.2.*

the first iteration of the Australian QNA forecasting experiment. The system's hardware and software specifications are as follows:

- CPU: `Intel(R) Core(TM) i7-10700 CPU @ 2.90GHz 2.90 GHz`
- RAM size: `64 GB`
- R version: `R-4.2.1_2022-06-23 _ucrt`

5.5.1 Results

Forecasting accuracy indices based on CRPS and ES are presented in Table 5.4. As a benchmark approach, we use the base forecasts calculated using the bootstrap method.

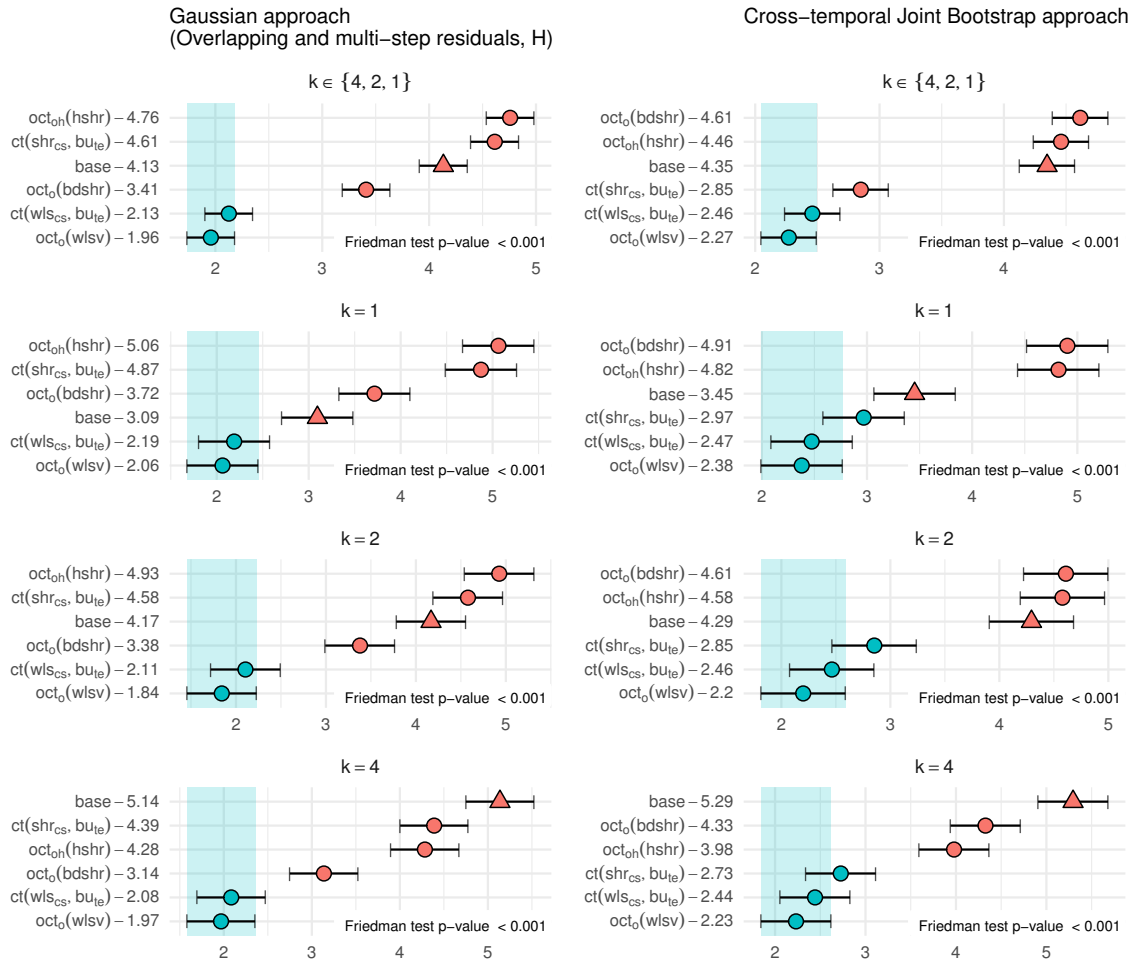


Figure 5.8: *MCB Nemenyi test for the Australian QNA dataset using CRPS at different temporal aggregation levels for the Gaussian (using overlapping and multi-step residuals, H) and the non-parametric bootstrap approaches. In each panel, the Friedman test p-value is reported in the lower right corner. The mean rank of each approach is shown to the right of its name. Statistically significant differences in performance are indicated if the intervals of two forecast reconciliation procedures do not overlap. Thus, approaches that do not overlap with the blue interval are considered significantly worse than the best, and vice-versa.*

For base forecasts, we find that using a parametric approach with the normal distribution performs better than the non-parametric bootstrap approach. This is likely due to the limited number of residuals available for bootstrapping, which does not allow for sufficient exploration of the data. Directly specifying diagonal covariance matrices seems to be more effective than shrinking to a target covariance matrix. Among all the procedures, $\text{ct}(wls_{cs}, bu_{te})$ and $\text{oct}_o(wlsv)$ show the greatest relative gains. In contrast, $oct_{oh}(hshr)$ does not show much improvement. Furthermore, the greatest improvements are observed for higher temporal aggregation levels.

We utilize the non-parametric Friedman test and the post hoc "Multiple Comparison with the Best" (MCB) Nemenyi test (Koning *et al.*, 2005; Kourentzes and Athanasopoulos, 2019; Makridakis *et al.*, 2022; Kourentzes, 2022) to determine if the forecasting performances of the different techniques are significantly different from one another. Figure 5.8 presents the MCB using the CRPS. The probabilistic forecasts from $\text{ct}(wls_{cs}, bu_{te})$ and $\text{oct}_o(wlsv)$ are significantly better than the base forecasts at any level of aggregation. Unlike the application on the Australian Tourism Demand (see Section 5.6), in this case one of the partly bottom-up approaches is not significantly worse than the most performing optimal approach.

Overall, we find that using overlapping residuals almost always leads to a greater improvement in terms of both ES and CRPS. Forecasts at the most aggregated level (year) seem to benefit the most from reconciliation, and using one-step overlapping residuals appears to be sufficient to improve forecasts if the generation of the base forecasts sample paths takes into account the multi-step structure.

5.6 Forecasting Australian Tourism Demand

The Australian Tourism Demand dataset (Wickramasuriya *et al.*, 2019) measures the number of nights Australians spent away from home. It includes 228 monthly observations of Visitor Nights (VN) from January 1998 to December 2016, and has a cross-sectional grouped structure based on a geographic hierarchy crossed by purpose of travel. The geographic hierarchy comprises seven states, 27 zones, and 76 regions, for a total of 111 nested geographic divisions. Six of these zones are each formed by a single region, resulting in a total of 105 unique nodes in the hierarchy. The purpose of travel comprises four categories: holiday, visiting friends and relatives, business, and other. To avoid redundancies (Di Fonzo and Girolimetto, 2022a), 24 nodes are not considered, resulting in an unbalanced hierarchy of 525 unique nodes instead of the theoretical 555 with duplicated nodes. The dataset includes the 304 bottom series, which are aggregated into 221 upper time series. Table 5.5 omits duplicated entries and updates the information in Table 7 from Wickramasuriya *et al.* (2019). This data can be temporally aggregated into 2, 3, 4, 6, or 12 months ($\mathcal{K} = \{12, 4, 3, 2, 1\}$).

We perform a rolling forecast experiment with an expanding window. The process begins by using the first 10 years, from January 1998 to December 2008, to generate forecasts for the entire following year (2009). Then, the training set is increased by one month. This process is repeated until the last training set is used (January 1998 to December 2015) with a total of 85 different test sets. For the temporal aggregation dimension we aggregate the monthly data up to annual data. We obtain 12-step, 6-step,

	Number of series		
	GD	**PT**	**Tot.**
Australia	1	4	5
States	7	28	35
Zones*	21	84	105
Regions	76	304	380
Total	**105**	**420**	**525**

Table 5.5: *Grouped time series for the Australian Tourism Demand dataset. 6 Zones with only one Region are included in Regions. GD: Geographic Division; PT: Purpose of Travel.*

4-step, 3-step, 2-step and 1-step ahead base forecasts respectively from the monthly data and the aggregation over 2, 3, 4, 6, and 12 months. ETS models selected by minimizing the AICc criterion (Hyndman *et al.*, 2023) are fitted to the log-transformed data, with the resulting base forecasts being back-transformed to produce non-negative forecasts (Wickramasuriya *et al.*, 2020).

The (1000) base forecast samples are obtained using the Gaussian approach with sample[8] covariance matrices (Section 5.4) using multi-step residuals[9] and the bootstrap approach (Section 5.3.2). For reconciliation, 6 different approaches have been adopted (see Table 5.2): ct(shr_{cs}, bu_{te}), oct($struc$), oct($wlsv$), oct($bdshr$), oct$_h$($bshr$), and oct$_h$($hshr$).

Negative forecasts may be produced during the reconciliation phase (Wickramasuriya *et al.*, 2020, see also Chapters 2 and 3) thus generating unreasonable values (e.g., a negative forecast for tourism demand makes no sense). To overcome this limitation, we applied the simple heuristic proposed in Chapters 2 and 3. Following Theorem 5.4, we are thus able to obtain reconciled samples respecting non-negativity constraints starting from an incoherent sample simulated from a Gaussian distribution. Finally, to evaluate the performance, we employ the Continuous Ranked Probability Score (CRPS), the Energy Score (ES), and the "Multiple Comparison with the Best" (MCB) Nemenyi test, introduced in Sections 5.5 and 5.5.1.

Tables 5.6 shows the runtime (in seconds) required for simulating 1000 samples (first row, base) and the additional time needed for reconciliation with various approaches in the first iteration of the Australian Tourism Demand forecasting experiment. The system's hardware and software specifications are the same as the first experiment.

[8]The results with shrunk covariance matrices are available in the online appendix E.2, where we also report the results obtained using other reconciliation approaches.

[9]We do not include overlapping, as we are unable to correctly determine the residuals for the overlapping series using ETS models (see Section 5.4.2).

Reconciliation approach	ctjb	G	B	H	HB
base	61.21	660.43	643.54	641.40	692.63
ct(shr_{cs}, bu_{te})	+ 3.79	+ 4.02	+ 3.79	+ 3.54	+ 4.18
oct($struc$)	+ 7.73	+ 7.10	+ 7.19	+ 7.11	+ 6.56
oct($wlsv$)	+ 8.24	+ 6.97	+ 6.99	+ 7.04	+ 6.46
oct($bdshr$)	+ 60.27	+ 52.65	+ 52.13	+ 51.92	+ 49.51
oct$_h$($bshr$)	+ 503.52	+ 426.20	+ 419.99	+ 418.94	+ 422.94
oct$_h$($hshr$)	+ 485.66	+ 482.13	+ 418.64	+ 418.82	+ 463.18

Table 5.6: *Computational time (in seconds) for the first iteration of the Australian Tourism Demand forecasting experiment. The first row (base) reports the time to simulate 1000 samples, and the remaining rows the additional time to reconcile them with different approaches.*

5.6.1 Results

The CRPS and ES indices are shown, respectively, in Table 5.7 for monthly, quarterly and annual forecasts. These tables are divided by different temporal levels and each column uses a different approach to calculate the base forecasts, referred to as "base". The bootstrap method is used as a benchmark to calculate the accuracy indices.

Base forecasts using a Gaussian approach are better in terms of both CRPS and ES compared to the bootstrap approach (the benchmark). Assumptions of truncated Gaussianity (Gaussian with negative values set to zero) may seem strict, but given the limited number of residuals, it can lead to improved forecasts in terms of CRPS and ES. Bootstrap forecasts suffer from the limited number of available residuals, leading in general to lower forecast accuracy. The Gaussian approach overcomes this limitation and provides better results. Regarding the different covariance matrix estimates for Gaussian base forecasts, there are no big differences. For this reason, using only the high frequency bottom time series can be useful to estimate fewer parameters and reduce the initial high dimensionality.

In the Gaussian case, partly bottom-up techniques like ct(shr_{cs}, bu_{te}) lead to better results than the benchmark (bootstrap base forecasts). However, it is not always guaranteed that the improvement is higher than the starting base forecasts (by comparing the value of each column). This is particularly true for higher levels of temporal aggregation. Overall, oct($bdshr$) in terms of CRPS is almost always the best. The shrinkage approach oct$_h$($hshr$) performs well in the bootstrap case: it is competitive with oct($bdshr$) at lower temporal frequency ($k \in \{2, 1\}$) and it is able to improve for $k \geq 3$. In terms of ES, oct($bdshr$) is still competitive, although it does not always show the best relative performance, like oct$_h$($bshr$). It is also worth noting that oct($struc$),

Reconciliation approach	**Base forecasts' sample approach**									
	ctjb	Gaussian approach*				ctjb	Gaussian approach*			
		G	B	H	HB		G	B	H	HB
					$\overline{RelCRPS}$					
		$\forall k \in \{12,6,4,3,2,1\}$					$k=1$			
base	1.000	0.971	0.971	0.973	0.973	1.000	0.972	0.972	0.972	0.972
ct(shr_{cs}, bu_{te})	1.057	0.974	0.969	0.974	0.969	0.976	0.963	0.962	0.963	0.962
oct($struc$)	0.982	0.962	0.961	0.961	0.959	0.970	0.963	0.963	0.963	0.963
oct($wlsv$)	0.987	0.959	0.959	0.958	0.957	0.952	0.957	0.957	0.957	0.957
oct($bdshr$)	0.975	**0.956**	**0.953**	**0.952**	**0.951**	**0.949**	**0.955**	**0.953**	**0.954**	**0.954**
oct$_h$($bshr$)	0.994	1.018	1.020	1.016	1.019	0.988	1.007	1.013	1.006	1.012
oct$_h$($hshr$)	**0.969**	0.993	0.993	0.990	0.991	0.953	0.977	0.977	0.979	0.979
		$k=3$					$k=12$			
base	1.000	0.971	0.971	0.972	0.973	1.000	0.968	0.967	0.969	0.969
ct(shr_{cs}, bu_{te})	1.041	0.977	0.974	0.977	0.974	1.163	0.977	0.965	0.977	0.965
oct($struc$)	0.986	0.967	0.966	0.966	0.965	0.982	0.951	0.949	0.947	0.943
oct($wlsv$)	0.983	0.963	0.962	0.962	0.962	1.025	0.954	0.953	0.949	0.947
oct($bdshr$)	0.972	**0.960**	**0.958**	**0.957**	**0.957**	1.002	**0.950**	**0.944**	**0.939**	**0.935**
oct$_h$($bshr$)	0.999	1.021	1.022	1.018	1.022	0.987	1.024	1.021	1.021	1.019
oct$_h$($hshr$)	**0.971**	0.994	0.994	0.992	0.993	**0.978**	1.003	1.005	0.996	0.997
					ES ratio indices					
		$\forall k \in \{12,6,4,3,2,1\}$					$k=1$			
base	1.000	0.956	0.955	0.958	0.951	1.000	0.952	0.950	0.952	0.950
ct(shr_{cs}, bu_{te})	1.243	0.886	0.879	0.886	0.879	1.098	0.929	0.928	0.930	0.927
oct($struc$)	1.085	0.917	0.915	0.916	0.912	1.027	0.943	0.942	0.943	0.942
oct($wlsv$)	1.132	0.933	0.929	0.931	0.927	1.050	0.951	0.949	0.950	0.949
oct($bdshr$)	1.047	0.904	0.897	0.897	0.891	1.009	0.936	0.933	0.934	0.931
oct$_h$($bshr$)	**0.931**	**0.867**	**0.866**	**0.863**	**0.860**	**0.965**	**0.927**	**0.927**	**0.925**	**0.923**
oct$_h$($hshr$)	1.081	0.935	0.931	0.935	0.927	1.028	0.952	0.951	0.952	0.950
		$k=3$					$k=12$			
base	1.000	0.961	0.958	0.960	0.955	1.000	0.942	0.947	0.951	0.937
ct(shr_{cs}, bu_{te})	1.245	0.911	0.904	0.911	0.904	1.326	0.779	0.767	0.777	0.766
oct($struc$)	1.096	0.939	0.936	0.938	0.933	1.077	0.826	0.822	0.823	0.818
oct($wlsv$)	1.142	0.953	0.949	0.951	0.946	1.149	0.851	0.845	0.847	0.840
oct($bdshr$)	1.060	0.926	0.920	0.921	0.915	1.021	0.808	0.796	0.796	0.787
oct$_h$($bshr$)	**0.954**	**0.895**	**0.895**	**0.892**	**0.887**	**0.833**	**0.741**	**0.741**	**0.737**	**0.735**
oct$_h$($hshr$)	1.093	0.955	0.951	0.956	0.949	1.066	0.851	0.846	0.848	0.838

*The Gaussian method employs a sample covariance matrix and includes four techniques (G, B, H, HB) with multi-step residuals..

Table 5.7: *$\overline{RelCRPS}$ and ES ratio indices defined in* (5.10) *and* (5.11) *for the Australian Tourism Demand dataset. Approaches performing worse than the benchmark (bootstrap base forecasts, ctjb) are highlighted in red, the best for each column is marked in bold, and the overall lowest value is highlighted in blue. The reconciliation approaches are described in Table 5.2.*

which does not make use of residuals, proves to be competitive by consistently improving on the base forecasts in terms of both CRPS and ES.

Figure 5.9 shows the MCB using the CRPS for the Gaussian approach using multi-step residuals (HB) and the non-parametric bootstrap approach. In general, the partly bottom-up procedure improves with respect to base forecasts at monthly level, but optimal cross-temporal procedures are always better. In the bootstrap framework, we

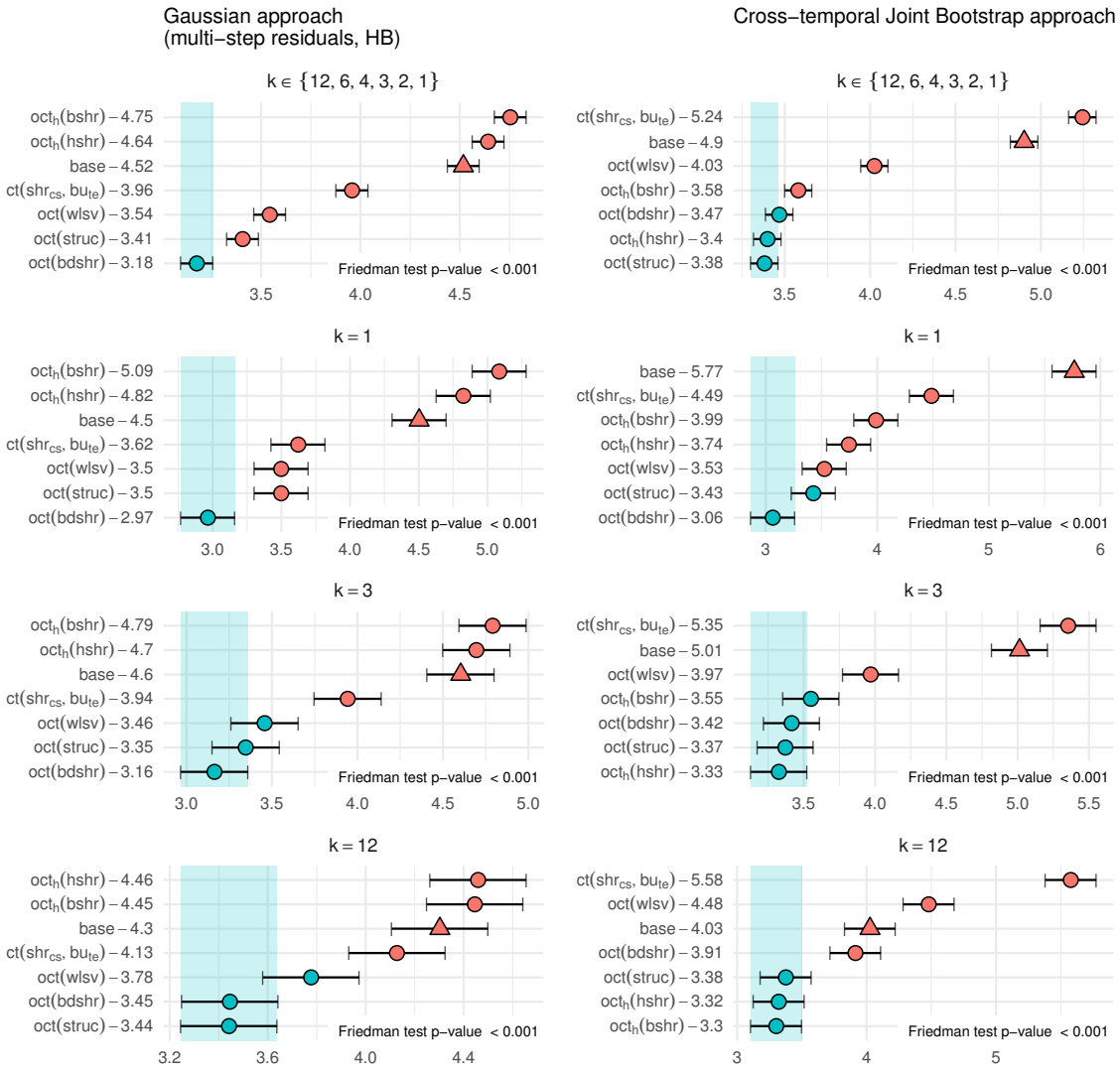


Figure 5.9: *MCB Nemenyi test for the Australian Tourism Demand dataset using CRPS at different temporal aggregation levels for the Gaussian (multi-step residuals, HB) and the non-parametric bootstrap approaches. In each panel, the Friedman test p-value is reported in the lower right corner. The mean rank of each approach is shown to the right of its name. Statistically significant differences in performance are indicated if the intervals of two forecast reconciliation procedures do not overlap. Thus, approaches that do not overlap with the blue interval are considered significantly worse than the best, and vice-versa.*

can identify a group of three procedures, oct(*bdshr*), oct(*hshr*) and oct(*struc*) that are almost always in the group of the best approaches (denoted by the blue dot). In the Gaussian framework, oct(*wlsv*), oct(*struc*), and oct(*bdshr*) are always significantly better than base forecasts and equivalent in terms of results for temporal aggregation orders greater than 2. For monthly series, oct(*bdshr*) is always significantly better than all other approaches.

Chapter 6

A reconciliation approach for the realized volatility

6.1 Introduction

Volatility forecasting has attracted a relevant amount of interest in the financial econometrics literature since the seminal contribution of Engle (1982). The reasons are well-known and ground on the importance of volatility in several areas, from risk management to asset allocation, from hedging to pricing. In the last two decades the interest has shifted from conditional variance models to the modeling and forecasting of Realized Variances, RV (Andersen *et al.*, 2001a,b, 2003). In this case, starting from the work of Corsi (2009), based on the introduction of a simple specification capable of capturing the strong serial correlation of RV sequences, several additional specifications have been introduced. These include models with price jumps in the variance dynamic (Andersen *et al.*, 2007a), controlling for residual heteroskedasticity (Corsi and Renò, 2012), dealing with measurement errors (Bollerslev *et al.*, 2016), disentangling the role of positive and negative returns (Patton and Sheppard, 2015), and considering a quantile-based intraday decomposition of RV (Bollerslev *et al.*, 2022). From a pure forecasting perspective, although all models might provide statistical and/or economic advantages compared to simpler specifications, there is no clear evidence that a model clearly superior to all competitors exists (Caporin, 2022).

A few contributions share a common feature from the modeling perspective: to predict the RV they extract information from a decomposition of lagged RV. This holds, in particular, when separating the continuous and discontinuous variance components, as in Andersen *et al.* (2007a), or when 'Good and Bad' volatilities are used, as in Patton and Sheppard (2015), or when the two approaches are combined (Caporin, 2022),

or finally when a more flexible decomposition according to conditional and time-varying intraday returns quantiles is considered (Bollerslev *et al.*, 2022). In all of these cases, the decomposition provides what is known as a *hierarchy* in the hierarchical forecasting literature (Hyndman *et al.*, 2011), that is a structure in which an aggregate series (e.g., daily RV) can be seen at the top over its constituents series (e.g., intraday RV decompositions).

Therefore, when dealing with intraday data, sometimes the observed returns may be grouped based on some criterion of similarity, such as the sign, or the occurrence in portions of the intraday returns density support. This decomposition may be employed in forecasting the daily RV, through segment-level forecasting within each segment. Challenges associated with successfully applying intraday decompositions include how to create segments and how to combine the segment-level RV forecasts to recover a daily RV forecast. The current paper proposes a method to exploit existing and to create new decompositions of the daily RV based on high-frequency intraday data, create segment-level forecasts, and then combine these forecasts to improve the daily RV forecasts. We present a combined-aggregative forecasting method for daily RV that allows to obtain a global prognosis by summing up/combining the forecasts of the compounding individual components. We detail a bottom-up (indirect) and a regression-based forecast reconciliation (Hyndman *et al.*, 2011, Wickramasuriya *et al.*, 2019) approach, and study their forecasting performance *vis-à-vis* the daily direct RV forecasts produced by the classic HAR model (Corsi, 2009), and two variants that take into account intraday RV decompositions (Patton and Sheppard, 2015, Bollerslev *et al.*, 2022). At this end, we have devised a forecasting experiment to evaluate the new proposed forecasting approaches on the high-frequency data of a few assets. The proposed method utilizes standard forecasting tools, but applies them in a unique combination that results in a higher level of daily RV forecast accuracy than other traditional methods.

A common technique used to forecast an aggregate involves bottom-up method. This procedure starts from forecasting all bottom-level components and then obtaining the top-level forecast by simply summing these bottom-level forecasts. By contrast, the direct approach simply produces forecasts at the top level. Nevertheless, realized volatility in different segments of the day usually have quite different patterns, hence the trivial approach of only forecasting the bottom-level series is unlikely to provide very accurate forecasts for the top-level series. In addition, the different behaviour of RV at different time periods, suggests to group the observed volatility according to different time intervals. This may be considered either alone or in conjunction with other grouping schemes related to the nature of the volatility itself (i.e., 'Good & Bad' volatility, Patton

and Sheppard, 2015). This gives raise to a hierarchical/grouped time series, where daily RV may be seen as the top-level series of a hierarchy, whose bottom level consists of the components obtained by crossing time periods and volatility decompositions. In this case, besides bottom-up, another (hopefully) more accurate method for hierarchical forecasting is to independently generate RV forecasts at all levels of the hierarchy. The advantage of independently generating the forecasts at each level is that each level can customize its forecasting model according to the varying characteristic of the RV at its own level. Thus, such approach could provide more accurate top-level forecasts than traditional direct or bottom-up approaches. However, these independently-made forecasts have the undesirable consequence that the lower-level forecasts cannot add up exactly to the higher-level forecasts. Thus, it is necessary to carry out some adjustments to ensure that hierarchical forecasts meet the constraints introduced by the hierarchical structure in the same way as their measurement data, i.e., in each day the sum of the lower-level intraday RV components forecasts should be equal to the higher-level daily RV forecast.

The methodology we put forward comprises two steps:

1. *Independent forecasting of daily RV and its components*, generating the so called "base forecasts". For the daily RV series and for all its intraday components from a specific decomposition we issue a forecast as accurate as possible, using three different HAR-based forecasting models proposed in the literature. In general, the base forecasts are not coherent with the additive decomposition law linking the observed daily RVs to its observed intraday components.
2. *Aggregation post-process.* We then combine these forecasting results to forecast the h-day-ahead RV, where $h \geq 1$ is the forecast horizon. In this paper, we consider a postprocess aggregation method suggested in the vast literature on regression-based cross-sectional and temporal forecast reconciliation (Hyndman *et al.*, 2011; Athanasopoulos *et al.*, 2017; Wickramasuriya *et al.*, 2019).

We compare the accuracy of the aggregate (direct) forecasting with the disaggregate (indirect) bottom-up and the regression-based forecast-reconciliation approach for the daily RV of the Dow Jones Industrial Average index and 26 of its constituents assets. Most of the existing studies on forecasting daily RV did not consider possible hierarchical structures deriving from intraday decompositions of the RV, and often missed the coherent relationships between individual components. An exception is Sohn and Lim (2007), who evaluated aggregate *vs.* disaggregate forecasting of 30 simulated coherent components of the DJIA index based on the AR(2)-GARCH(1,1) model. However, the results of this experiment were quite inconclusive, as it was found that the accuracy

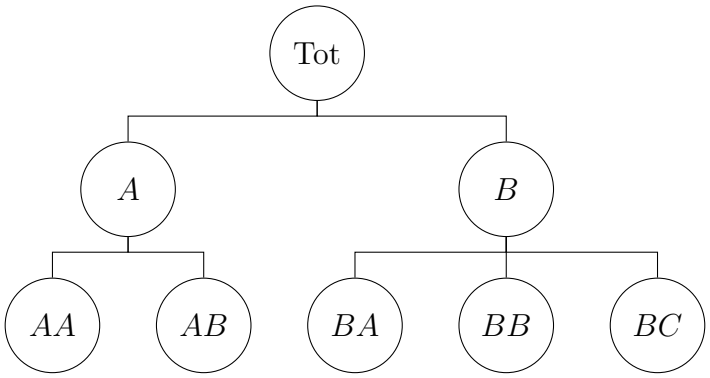


Figure 6.1: *A simple three-level hierarchical structure.*

of the indirect forecasting method varied depending on the correlation degree of the coherent components. Instead, and contrary to Sévi (2014), our results indicate that considering the various components of the realized variance do represent a significant improvement in an out-of-sample forecast evaluation framework.

The chapter proceeds as follows. The forecast reconciliation methodology is reviewed in Section 6.2. Section 6.3 briefly describes the volatility modeling and introduces the hierarchies. The empirical setup of the out-of-sample forecasting experiment is described in Section 6.4, and Section 6.5 shows the results. A robustness analysis is performed in Section 6.6.

6.2 Forecast reconciliation: a recap

Forecast reconciliation is a post-forecasting process aimed to improve the quality of the *base* forecasts for a system of hierarchical/grouped, and more generally linearly constrained, time series by exploiting the constraints that the series in the system must fulfil, whereas in general the base forecasts do not; see, among others, Hyndman *et al.* (2011), Chapter 3, and Chapter 4. Following Panagiotelis *et al.* (2021), a linearly constrained time series $\boldsymbol{y}_t$ is defined as a n-dimensional time series such that all observed values $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T$ and all future values $\boldsymbol{y}_{T+1}, \boldsymbol{y}_{T+2}, \ldots$ lie in the coherent linear subspace $\mathcal{S} \subset \mathbb{R}^n$, that is: $\boldsymbol{y}_t \in \mathcal{S}$, $\forall t$. In many cases, the linear constraints can be represented as a hierarchy, where the time series are linked through summation constraints. Figure 6.1 shows an example of a hierarchical time series with eight variables and three levels: the top-variable at level 0, two variables (A and B) at level 1, and five variables at level 2 (AA, AB, BA, BB, BB, BC). The 3 aggregated upper time series are linked to the

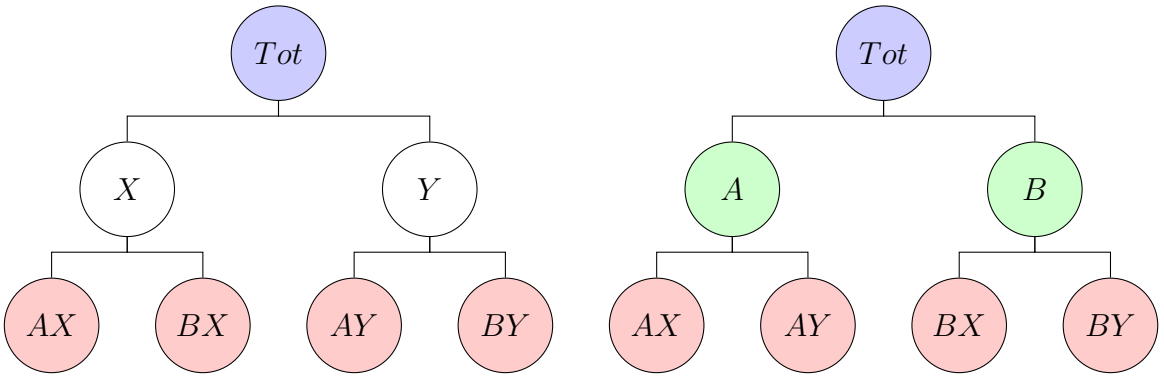


Figure 6.2: *A simple grouped structure.*

bottom-level variables through summation:

$$\begin{aligned} y_{Tot,t} &= y_{A,t} + y_{B,t} \\ y_{A,t} &= y_{AA,t} + y_{AB,t} \qquad\qquad \forall t = 1, \dots, T. \\ y_{B,t} &= y_{BA,t} + y_{BB,t} + y_{BC,t} \end{aligned}$$

The bottom-level series can be thought of as building blocks that cannot be obtained as sum of other series in the hierarchy, while all the series at upper levels can be expressed by appropriately summing part or all of them. In details, let $\boldsymbol{b}_t$ and $\boldsymbol{a}_t$ be the vectors of bottom level and upper level time series at time t, respectively. For example, $\boldsymbol{b}_t = [y_{AA,t} \quad y_{AB,t} \quad y_{BA,t} \quad y_{BB,t} \quad y_{BC,t}]'$, $\boldsymbol{a}_t = [y_{Tot,t} \quad y_{A,t} \quad y_{B,t}]'$. Denoting by $\boldsymbol{y}_t$ the vector $\boldsymbol{y}_t = [\boldsymbol{a}_t' \;\; \boldsymbol{b}_t']'$, the relationships linking bottom and upper time series can be equivalently expressed as:

$$\boldsymbol{a}_t = \boldsymbol{A}\boldsymbol{b}_t, \quad \boldsymbol{y}_t = \boldsymbol{S}\boldsymbol{b}_t, \quad \boldsymbol{C}\boldsymbol{y}_t = \boldsymbol{0}_{(n_a \times 1)}, \quad t = 1, \dots, T, \tag{6.1}$$

where $\boldsymbol{A}$ is the $(n_a \times n_b)$ aggregation matrix, $\boldsymbol{S} = \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{I}_{n_b} \end{bmatrix}$ is the $(n \times n_b)$ structural matrix and $\boldsymbol{C} = [\boldsymbol{I}_{n_a} \quad -\boldsymbol{A}]$ is the $(n_a \times n)$ zero constraints matrix. We call *structural representation* of series $\boldsymbol{y}_t$ the formulation $\boldsymbol{y}_t = \boldsymbol{S}\boldsymbol{b}_t$, $t = 1, \dots, T$, and *zero-constrained representation* of series $\boldsymbol{y}_t$ the equivalent expression $\boldsymbol{C}\boldsymbol{y}_t = \boldsymbol{0}$, $t = 1, \dots, T$.

A linearly constrained time series formed by two or more hierarchical time series sharing the same top level series, and the same bottom level series, is called *grouped time series* (Hyndman *et al.*, 2011). An example is shown in Figure 6.2, where the Total variable can be described as two different hierarchies with intermediate variables (X, Y) and (A, B), respectively, which share the same four bottom-level variables (AX, BX, AY, BY). Provided matrix $\boldsymbol{A}$ is appropriately designed, the definitions of matrices $\boldsymbol{S}$ and $\boldsymbol{C}$ remain unchanged.

Now, suppose we have the $(n \times 1)$ vector $\widehat{\boldsymbol{y}}_h$ of unbiased base forecasts for the n variables of the linearly constrained series $\boldsymbol{y}_t$ for the forecast horizon h. If the base forecasts have been independently computed, generally, they do not fulfil the cross-sectional aggregation constraints, that is, $\boldsymbol{C}\widehat{\boldsymbol{y}}_h \neq \boldsymbol{0}_{(n_a \times 1)}$. The aim of forecast reconciliation is to adjust the base forecast $\widehat{\boldsymbol{y}}_h$ by using a mapping $\psi : \mathbb{R}^n \to \mathcal{S}$ to obtain the reconciled forecast vector $\widetilde{\boldsymbol{y}}_h = \psi(\widehat{\boldsymbol{y}}_h)$, where $\widetilde{\boldsymbol{y}}_h \in \mathcal{S}$. The mapping ψ can be defined as a projection onto $\mathcal{S}$ (van Erven and Cugliari, 2015; Panagiotelis *et al.*, 2021):

$$\widetilde{\boldsymbol{y}}_h = \boldsymbol{M}\widehat{\boldsymbol{y}}_h, \tag{6.2}$$

where $\boldsymbol{M} = \boldsymbol{I}_n - \boldsymbol{W}\boldsymbol{C}'\left(\boldsymbol{C}\boldsymbol{W}\boldsymbol{C}'\right)^{-1}\boldsymbol{C}$, with $\boldsymbol{W}$ error covariance matrix of the base forecasts $\widehat{\boldsymbol{y}}_h$. Another way to obtain the reconciled forecasts is through the structural approach proposed by Hyndman *et al.* (2011), such that

$$\widetilde{\boldsymbol{y}}_h = \boldsymbol{S}\boldsymbol{G}\widehat{\boldsymbol{y}}_h, \tag{6.3}$$

where $\boldsymbol{G} = (\boldsymbol{S}'\boldsymbol{W}^{-1}\boldsymbol{S})^{-1}\boldsymbol{S}'\boldsymbol{W}^{-1}$, and it can be shown that $\boldsymbol{M} = \boldsymbol{S}\boldsymbol{G}$ (Wickramasuriya *et al.*, 2019). Several alternatives have been provided in the literature to approximate the covariance matrix $\boldsymbol{W}$ (Hyndman *et al.*, 2011, 2016; Wickramasuriya *et al.*, 2019). In this work, we consider the state of the art shrinkage covariance matrix approximation proposed by Wickramasuriya *et al.* (2019),

$$\boldsymbol{W} = \hat{\lambda}\widehat{\boldsymbol{W}}_D + (1-\hat{\lambda})\widehat{\boldsymbol{W}}, \tag{6.4}$$

where $\widehat{\boldsymbol{W}}$ is the covariance matrix of the in-sample errors, and $\widehat{\boldsymbol{W}}_D = \boldsymbol{I}_n \odot \widehat{\boldsymbol{W}}$, where $\odot$ denotes the Hadamard product. Usually, $\widehat{\boldsymbol{W}} = \dfrac{1}{T}\displaystyle\sum_{t=1}^{T}\hat{\boldsymbol{e}}_t\hat{\boldsymbol{e}}_t'$, that is the covariance matrix of the one-step ahead in-sample forecast errors $\widehat{\boldsymbol{e}}_t = \boldsymbol{y}_t - \widehat{\boldsymbol{y}}_t$, $t = 1, \ldots, T$. However, to address potential issues arising from the possible presence of heteroscedasticity and autocorrelation in the residuals, we employ also the Heteroscedasticity and Autocorrelation Consistent (HAC) covariance matrix proposed by Andrews (1991).

In the next section, after briefly reviewing the estimation of RV, we link the hierarchical forecasting literature to the RV modeling one.

6.3 RV modeling: a hierarchical perspective

The measurement of daily RV builds on the availability of data at a frequency higher than the day. If we denote by $t = 1, \ldots, T$, the daily time index, and by $i = 1, 2, \ldots N$, the intraday time index, the prices of a financial instrument observed in high frequency are denoted by $P_{i,t}$. From the prices we move to log-returns $r_{i,t}$ and to the estimation

of RV in a given day as follows:

$$RV_t = \sum_{i=1}^{N} \left[\log\left(P_{i,t}\right) - \log\left(P_{i-1,t}\right)\right]^2 = \sum_{i=1}^{N} r_{i,t}^2, \tag{6.5}$$

where prices at the intraday level are assumed to be observed on an equally spaced time grid (e.g., every minute), and for $i = 1$ the lagged price corresponds to the opening price of the day, thus excluding the overnight return from the evaluation. The financial econometrics literature has extensively discussed the issue of estimation of RV in the presence of microstructure noise and of price jumps; see, among many others, Aït-Sahalia and Jacod (2014) and therein cited references. In this work, we refer to the simplest approach reported above. Moreover, as our final purpose is to adopt hierarchical forecast reconciliation approaches starting from the forecast of bottom time series, we do not consider the decomposition of RV into its continuous and discontinuous component, since it is known that the discontinuous component is not predictable by means of relatively simple linear models; among the possible approaches, see Andersen *et al.* (2011) and Aït-Sahalia *et al.* (2015).

As mentioned in the introduction, several authors have focused on decompositions of RV. The most known example is given by the use of signed variations, as in Patton and Sheppard (2015), such that $RV_t = SV_t^+ + SV_t^-$, with

$$SV_t^+ = \sum_{i=1}^{N} r_{i,t}^2 I\left(r_{i,t} \geq 0\right) \quad \text{and} \quad SV_t^- = \sum_{i=1}^{N} r_{i,t}^2 I\left(r_{i,t} < 0\right),$$

where $I\left(a\right)$ is an indicator function taking unit value when condition a is true and zero otherwise. The signed variations are also known as *Semi-Variances* (Barndorff-Nielsen *et al.*, 2010), or as *Good* (SV_t^+) and *Bad* (SV_t^-) volatility, respectively, and separate the contribution to RV coming from upside and downside price movements.

More recently, Bollerslev *et al.* (2022) introduced a quantile-based decomposition, generalizing the signed variation approach, $RV_t = \sum_{l=1}^{p} PV_t^{(l)}$, with

$$PV_t^{(l)} = \sum_{i=1}^{N} r_{i,t}^2 I\left(\mathcal{Q}_{r,t}\left(\alpha_{l-1}\right) < r_{i,t} \leq \mathcal{Q}_{r,t}\left(\alpha_l\right)\right),$$

where $\mathcal{Q}_{r,t}\left(\tau\right) = \sqrt{N^{-1}RV_t}\mathcal{Q}_{z,t}\left(\tau\right)$, $z_{i,t} = \dfrac{r_{i,t}}{\sqrt{N^{-1}RV_t}}$ is the standardized intraday return, $\mathcal{Q}_{z,t}\left(\tau\right)$ is the empirical τ-quantile of the intraday standardized returns distribution in day t, $\mathcal{Q}_{r,t}\left(\alpha_0\right) = -\infty$ and $\mathcal{Q}_{r,t}\left(\alpha_p\right) = +\infty$, and $0 < \alpha_1 < \ldots < \alpha_{p-1} < 1$ is a sequence of probabilities, with $\alpha_0 = 0$ and $\alpha_1 = 1$. The $PV_t^{(l)}$ components, also called *Partial Variances*, allow separating the contribution to RV coming from intraday returns according to both their sign and size.

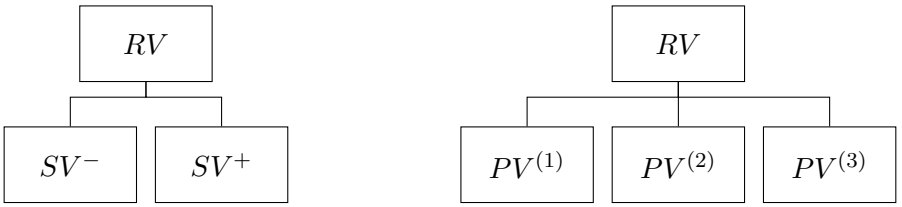


Figure 6.3: *Hierarchical representations of the Bad and Good (left, Patton and Sheppard, 2015) and PV*(3) *(right, Bollerslev, 2022) decompositions of daily RV.*

The concept of hierarchical time series in the framework of forecasting daily RV using its intraday decompositions may be illustrated by considering the two simple hierarchies deriving from the intraday RV decompositions reported above, i.e. Patton and Sheppard (2015) and Bollerslev *et al.* (2022), the last in the simple case with $p = 2$. Figure 6.3 provides a graphical representation of the two hierarchies.

To illustrate the advantage and flexibility of forecast reconciliation approaches in the RV context, we provide a more general form of temporal and threshold-based decomposition of daily RV. The use of quantiles computed from intraday returns is, in practice, a special case of a grouping of returns according to pre-defined, possibly time-varying thresholds, with general representation given by:

$$z_{l,t} = \sum_{t=1}^{N} r_{i,t}^2 I\left(c_{t,l-1} < r_{i,t} \leq c_{t,l}\right), \quad l = 1, \ldots, p, \tag{6.6}$$

where $c_{t,0} = -\infty$ and $c_{t,p} = +\infty$.

Differently, by exploiting the availability of information distributed over a range of minutes within a given day, we might group the intraday returns according to a temporal scheme:

$$w_{k,t} = \sum_{i=m(k-1)+1}^{mk} r_{i,t}^2, \quad k = 1, \ldots, \frac{N}{m}, \tag{6.7}$$

which is equivalent to

$$w_{k,t} = \sum_{t=1}^{N} r_{i,t}^2 I\left(m(k-1) < i \leq mk\right), \quad k = 1, \ldots, \frac{N}{m}. \tag{6.8}$$

The threshold- and time-based decompositions might be combined giving rise to the most disaggregated (bottom level) time series, defined as

$$x_{l,k,t} = \sum_{t=1}^{N} r_{i,t}^2 I\left[\left(c_{t,l-1} < r_{i,t} \leq c_{t,l}\right) \cap \left(m(k-1) < i \leq mk\right)\right], \quad \begin{matrix} l = 1, \ldots, p \\ k = 1, \ldots, \frac{N}{m} \end{matrix}, \tag{6.9}$$

or

$$x_{l,k,t} = \sum_{i=m(k-1)+1}^{mk} r_{i,t}^2 I\left(c_{t,l-1} < r_{i,t} \leq c_{t,l}\right), \quad \begin{matrix} l = 1, \ldots, p \\ k = 1, \ldots, \frac{N}{m} \end{matrix}. \tag{6.10}$$

Temporal decomposition	Semi-Variances $r_{t,i} < 0$ SV^-	Semi-Variances $r_{t,i} \geq 0$ SV^+	Quantile-based decomposition $r_{t,i} \leq Q(\tau_1)$ $PV^{(1)}$	Quantile-based decomposition $Q(\tau_1) < r_{t,i} \leq Q(\tau_2)$ $PV^{(2)}$	Quantile-based decomposition $r_{t,i} > Q(\tau_2)$ $PV^{(3)}$
T1: minutes 1-78	T1SV^-	T1SV^+	T1$PV^{(1)}$	T1$PV^{(2)}$	T1$PV^{(3)}$
T2: minutes 79-156	T2SV^-	T2SV^+	T2$PV^{(1)}$	T2$PV^{(2)}$	T2$PV^{(3)}$
T3: minutes 157-234	T3SV^-	T3SV^+	T3$PV^{(1)}$	T3$PV^{(2)}$	T3$PV^{(3)}$
T4: minutes 235-312	T4SV^-	T4SV^+	T4$PV^{(1)}$	T4$PV^{(2)}$	T4$PV^{(3)}$
T5: minutes 313-390	T5SV^-	T5SV^+	T5$PV^{(1)}$	T5$PV^{(2)}$	T5$PV^{(3)}$

Table 6.1: *Left panel: The ten bottom variables from the time-by-'Good & Bad' volatility decompositions. Right panel: The fifteen bottom variables from the time-by-quantile daily decompositions according to* $PV(3)$.

For example, assuming a day consisting of 6.5 hours, with data available at the 1-minute frequency ($N = 390$), setting $p = 3$ and $m = 78$, we have the following decompositions of RV:

RV	w_1	w_2	w_3	w_4	w_5
z_1	$x_{1,1}$	$x_{1,2}$	$x_{1,3}$	$x_{1,4}$	$x_{1,5}$
z_2	$x_{2,1}$	$x_{2,2}$	$x_{2,3}$	$x_{2,4}$	$x_{2,5}$
z_3	$x_{3,1}$	$x_{3,2}$	$x_{3,3}$	$x_{3,4}$	$x_{3,5}$

where the threshold-based $z_l = \sum_{k=1}^{5} x_{l,k}$, $l = 1, \ldots, 3$, and temporal-based $w_k = \sum_{l=1}^{3} x_{l,k}$, $k = 1, \ldots, 5$, decompositions represent the marginals of a combined and richer decomposition. We also note that $RV = \sum_{k=1}^{5} w_k = \sum_{l=1}^{3} z_l = \sum_{l=1}^{3}\sum_{k=1}^{5} x_{l,k}$.

In the following empirical analyses, we will make use of the hierarchies/groupings generated by crossing a temporal decomposition in five non-overlapping intervals of 78 minutes each, with either (ii) the dichotomous intraday decomposition in 'Good and Bad' volatility (Patton and Sheppard, 2015, see Table 6.1), or (ii) a quantile-based decomposition with $p = 3$, with quantiles set to τ_1 and τ_2, respectively (see Table 6.1). The online appendix reports the graphical and structural representations of the hierarchies.

6.4 The empirical setup

6.4.1 Data description and analysis

We evaluate the impact of forecast reconciliation in forecasting daily RV of individual stocks included in the Dow Jones Industrial Average (DJIA) index, from the beginning of January 2003 to the end of June 2022. We use price data at the 1-minute frequency, adjusted for splits and dividends. We consider prices recorded from 9:00 AM to 3:59 PM

(time identifies the start of each intraday interval), obtaining 390 observations per day. Our dataset includes 4,908 full days excluding weekends, holidays and closed market days. The data have been recovered from Kibot.com.[1] We consider the 26 stocks whose data are available to us for the entire sample, denoted by the following tickers:[2] AAPL, AMGN, AXP, BA, CAT, CSCO, CVX, DIS, GS, HD, HON, IBM, INTC, JNJ, JPM, KO, MCD, MMM, MRK, MSFT, NKE, PG, UNH, VZ, WBA, and WMT. The use of DJIA constituents is in line with the choice made by Bollerslev *et al.* (2022) and allows dealing with the possible presence of large amount of zeros at the intraday level. In fact, for those highly liquid stocks, the presence of zeros is extremely limited.

Starting from the 1-minute data, we estimate the daily RV, and the decompositions (hierarchies) we previously mentioned. First of all, we decompose the RV into the Good and the Bad components, following Patton and Sheppard (2015); this gives a hierarchy with two bottom series. Second, following Bollerslev *et al.* (2022), we decompose RV into the Partial Variances $PV^{(g)}$. Differently from the authors, we do not optimally select the quantiles used for the decomposition, nor we allow for a time-change in the quantile. On the contrary, building on the evidence in Bollerslev *et al.* (2022), and in particular on the values in Table 2 of their paper, we select the $PV(3)$ decomposition with fixed quantiles set at the 10% and 75% thresholds.[3] This gives a hierarchy with three bottom series, and allows us to simplify the treatment and the following analyses. Third, we apply both the Good and Bad and $PV(3)$ decompositions on sub-samples of the day. We first divide the entire day in 5 sub-intervals of length equal to 78 observations (minutes), and in each sub-sample we apply either the Good and Bad or the $PV(3)$ decompositions. We note that this gives, overall, 10 bottom series in the former case, and 15 in the latter. In addition, we do have two possible intermediate aggregations, by temporal sub-sample, or by volatility (either SV or $PV(3)$) components. The hierarchies/groupings in this case are thus much richer than if only volatility-based decompositions are used (see the online appendix for additional details and results).

[1] The quality of data from Kibot.com is comparable to that of NYSE TAQ data. A comparison on a selected equity is available from authors upon request.

[2] The online appendix contains a summary description of the data used in the forecasting experiment.

[3] Bollerslev *et al.* (2022) optimally select both the number of partial variances as well as the quantiles. Their results show evidence of heterogeneity across the analyzed stocks the DJIA constituents. The quantiles we select correspond to the median values in the cross-section of stocks reported in their Table 2.

6.4.2 Base forecasts: direct forecasts from benchmark models and intraday components' forecasts

In the past years, the subject of the comparison of the forecast accuracy of aggregating disaggregate forecasts versus forecasts based on aggregated data has received attention in different fields, such as macro-economic (Marcellino *et al.*, 2003, Frale *et al.*, 2011, Poncela and García-Ferrer, 2014, Grassi *et al.*, 2015), demand (Petropoulos *et al.*, 2014, Mircetic *et al.*, 2022), and energy (Silva *et al.*, 2018, Wang *et al.*, 2021) forecasting. However, as far as we know, a detailed comparison of direct, indirect (bottom-up) and combination (forecast reconciliation) procedures for daily RV forecasting has not been provided yet.

At this end, we conduct an out-of-sample forecasting experiment where we compare direct daily-RV forecasts with three reference models proposed, respectively, by Corsi (2009), Patton and Sheppard (2015), and Bollerslev *et al.* (2022), two indirect forecasts obtained through simple bottom-up of the HAR forecasts of either semi-variances (Barndorff-Nielsen *et al.*, 2010) or partial-variances (Bollerslev *et al.*, 2022) components, and finally the daily forecasts of RV obtained through forecast reconciliation of both the aggregate (daily RV) and disaggregate (corresponding components of the daily RV) forecasts.

The modeling strategy we adopt for each of the bottom and top time series, and that we use to produce direct and base forecasts, is very simple. As our purpose is to introduce the use of forecast reconciliation tools in the prediction of RV, and not to identify the best forecasting univariate model, we fit on all (possibly disaggregate) time series the HAR model of Corsi (2009). Let x_t be a generic time series, that is, either RV_t or one of the bottom series according to one of the hierarchies previously introduced. We model x_t as follows:

$$x_t = \beta_0 + \beta_D x_{t-1} + \beta_W x_{t-1:t-5} + \beta_M x_{t-1:t-22} + \varepsilon_t, \tag{6.11}$$

where $x_{t-1:t-m} = \dfrac{1}{m}\displaystyle\sum_{j=1}^{m} x_{t-m}$, and ε_t is an innovation term. Parameters refer to the intercept (β_0) and to the daily, weekly and monthly effects (β_D, β_W, and β_M, respectively). Parameter estimation is based on least squares and we adopt robust standard errors to to be coherent with the *volatility-of-volatility* effect (Corsi *et al.*, 2008). Further, we do not consider the modelling of logarithms of RV_t, leaving to future researches the generalization of our approach along this line.[4]

[4]We stress the use of the logarithmic transformation of RV sensibly impacts on the aggregation constraints, with the need of moving toward probabilistic hierarchical forecasting approaches (Panagiotelis *et al.*, 2023).

For the top-level variable (i.e., daily RV) forecasts, we consider other two reference models proposed by Patton and Sheppard (2015) and Bollerslev *et al.* (2022) to account for, respectively, Good and Bad volatility and Partial Variances. In this last case, for simplicity, we consider an *ex-ante* choice of the two quantiles defining the three partial-variances decomposition (see Table 6.2).

We obtain daily forecasts for the time series of the daily semi-variances according to the 'Good & Bad' (Barndorff-Nielsen *et al.*, 2010, Patton and Sheppard, 2015) and to the $PV(3)$ decompositions (Bollerslev *et al.*, 2022), then we apply a simple bottom-up procedure to compute indirect forecasts of the daily RV. Finally, the forecasts obtained in the two previous steps are combined through the forecast reconciliation approach proposed by Wickramasuriya *et al.* (2019) (see also Hyndman *et al.*, 2011), which is a regression-based forecast combination approach exploiting the simple hierarchical structure of the two considered decomposition settings. The competing forecasting approaches, and the corresponding acronyms, are the following ones (reported for a one-period forecast horizon):

Direct forecasting procedures

- HAR: $\widehat{RV}_{t+1}^{HAR}$ (Corsi, 2009);
- SV: $\widehat{RV}_{t+1}^{SV}$ (Patton and Sheppard, 2015);
- $PV(3)$: $\widehat{RV}_{t+1}^{PV(3)}$ (Bollerslev *et al.*, 2022);

Indirect forecasting procedures (bottom-up)

- SV_{bu}: $\widehat{RV}_{t+1}^{SV_{bu}} = \widehat{SV}_{t+1}^{+} + \widehat{SV}_{t+1}^{-}$;
- $PV(3)_{bu}$: $\widehat{RV}_{t+1}^{PV(3)_{bu}} = \widehat{PV}_{t+1}^{(1)} + \widehat{PV}_{t+1}^{(2)} + \widehat{PV}_{t+1}^{(3)}$;

Forecast reconciliation procedures

- SV_{shr}/SV_{hac}: $\widehat{RV}_{t+1}^{SV_{shr}} = f(\widehat{RV}_{t+1}^{SV}, \widehat{SV}_{t+1}^{+}, \widehat{SV}_{t+1}^{-})$;
- $PV(3)_{shr}/PV(3)_{hac}$: $\widehat{RV}_{t+1}^{PV(3)_{shr}} = f(\widehat{RV}_{t+1}^{PV_3}, \widehat{PV^{(1)}}_{t+1}, \widehat{PV^{(2)}}_{t+1}, \widehat{PV^{(3)}}_{t+1})$;

where $\widehat{\boldsymbol{W}}$ in (6.4) is the in-sample residuals (Wickramasuriya *et al.*, 2019) and the HAC (Andrews, 1991) covariance matrix for *shr* and *hac*, respectively.

We will always use the above-reported acronyms independently from the forecast horizon we consider. Coherently with the common practice, see, for instance, (Patton and Sheppard, 2015), when the forecast horizon differs from 1 and becomes h, we set the dependent variable of our model to the $h-$period average cumulative value.[5]

[5]In this case, equation 6.11 becomes $x_{t+h-1:t} = \beta_0 + \beta_D x_{t-1} + \beta_W x_{t-1:t-5} + \beta_M x_{t-1:t-22} + \varepsilon_t$, with $x_{t+h-1:t}$ being the average of x_{t+i} for $i = 0, 1, \dots h-1$.

HAR: *Heterogeneous AutoRegressive model*

$$RV_t = \alpha_0 + \alpha_D RV_{t-1} + \alpha_W RV_{t-1:t-5} + \alpha_M RV_{t-1:t-22} + \varepsilon_t^{HAR}$$

SV: *Semi-Variances Heterogeneous AutoRegressive model*

$$RV_t = \beta_0 + \beta_D^+ SV_{t-1}^+ + \beta_D^- SV_{t-1}^- + \beta_W RV_{t-1:t-5} + \beta_M RV_{t-1:t-22} + \varepsilon_t^{SHAR}$$

PV(3): *Partial-Variances Heterogeneous AutoRegressive model*

$$RV_t = \gamma_0 + \sum_{l=1}^{3} \gamma_D^{(j)} PV_{t-1}^{(l)} + \gamma_W RV_{t-1:t-5} + \gamma_M RV_{t-1:t-22} + \varepsilon_t^{PV_3}$$

Semi-variances decomposition

$$SV_t^+ = \delta_0^+ + \delta_D^+ SV_{t-1}^+ + \delta_W^+ SV_{t-1:t-5}^+ + \delta_M^+ SV_{t-1:t-22}^+ + \varepsilon_t^{SHAR^+}$$
$$SV_t^- = \delta_0^- + \delta_D^- SV_{t-1}^- + \delta_W^- SV_{t-1:t-5}^- + \delta_M^- SV_{t-1:t-22}^- + \varepsilon_t^{SHAR^-}$$

Partial-variances decomposition $(l = 1, 2, 3)$

$$PV_t^{(l)} = \theta_0^{(l)} + \theta_D^{(l)} PV_{t-1}^{(1)} + \theta_W^{(l)} PV_{t-1:t-5}^{(l)} + \theta_M^{(l)} PV_{t-1:t-22}^{(l)} + \varepsilon_t^{PV(l)}$$

Time-variances decomposition $(j = 1, 2, 3, 4, 5)$

$$T_{j,t} = \eta_0^j + \eta_1^j T_{j,t-1} + \eta_2^j T_{j,t-1:t-5} + \eta_3^j T_{j,t-1:t-22} + \varepsilon_t^{T(j)}$$

Time- and semi-variances decomposition $(j = 1, 2, 3, 4, 5)$

$$T_j SV_t^- = \lambda_0^{j-} + \lambda_1^{j-} T_j SV_{t-1}^- + \lambda_2^{j-} T_j SV_{t-1:t-5}^- + \lambda_3^{j-} T_j SV_{t-1:t-22}^- + \varepsilon_t^{T(j)SV^-}$$
$$T_j SV_t^+ = \lambda_0^{j+} + \lambda_1^{j+} T_j SV_{t-1}^+ + \lambda_2^{j+} T_j SV_{t-1:t-5}^+ + \lambda_3^{j+} T_j SV_{t-1:t-22}^+ + \varepsilon_t^{T(j)SV^+}$$

Time- and partial-variances decomposition $(j = 1, 2, 3, 4, 5,\ l = 1, 2, 3)$

$$T_j PV_t^{(l)} = \eta_0^{j,l} + \eta_1^{j,l} T_j PV_{t-1}^{(l)} + \eta_2^{j,l} T_j PV_{t-1:t-5}^{(l)} + \eta_3^{j,l} T_j PV_{t-1:t-22}^{(l)} + \varepsilon_t^{T(j)PV^{(l)}}$$

Table 6.2: *Models used to produce daily forecasts for RV and its components: the first panel includes models for direct and base forecasts of RV while the second panel includes the specifications adopted for the intraday RV decompositions according to either semi-variances or partial-variances, alone or with time-groupings of non-overlapping 78 consecutive minutes intervals. See Table 1 for the definitions of the various quantities used as dependent variables. All equations take a HAR-type dynamic including as explanatory variables weekly and monthly moving averages (see equation 6.11). Error terms superscripts are set coherently with the dependent variable.*

6.4.3 Out-of-sample forecast evaluation

We perform a fixed length rolling window forecasting experiment on the DJIA series and 26 individual stocks previously mentioned. The first training set spans the period January 2, 2003 - December 29, 2006 (1,007 days). From each training set three direct multistep forecasts for, respectively, one-, five- and twenty-two-steps (day) ahead are computed, and this is done for all the time series components of the various hierarchies defined by the time-and/or-quantile-based RV_t intraday decompositions.

The base forecasts of the top-level series in each hierarchy (RV_t) are obtained according to the HAR, SV and $PV(3)$ models, respectively. The base forecasts of either the semi- or partial-variances series forming each hierarchy are obtained using appropriately adapted HAR models. The base forecasts are then reconciled through the MinT-shr approach (Wickramasuriya *et al.*, 2019) using the `R` package `FoReco` (Girolimetto and Di Fonzo, 2023a). The point forecast accuracy of daily RV_t is evaluated using the Mean Square Error (MSE), and the $QLIKE$ index (Patton, 2011b):[6]

$$MSE = \frac{1}{|\mathcal{S}|}\sum_{t=1}^{|\mathcal{S}|}\left(\widehat{RV}_t - RV_t\right)^2 \text{ and } QLIKE = \frac{1}{|\mathcal{S}|}\sum_{t=1}^{|\mathcal{S}|}\left(\frac{RV_t}{\widehat{RV}_t} - \frac{\log RV_t}{\log \widehat{RV}_t} - 1\right),$$

where $\widehat{RV}_t$ and $|\mathcal{S}|$ denote the one-step-ahead forecast and the number of days in the test set, respectively. Both MSE and $QLIKE$ belong to the family of loss functions of Patton (2011b), that are robust to the noise in the volatility proxy. We consider the MSE and $QLIKE$ ratios, defined respectively as

$$rMSE = \frac{MSE_i}{MSE_{HAR}} \qquad rQLIKE = \frac{QLIKE_i}{QLIKE_{HAR}}, \tag{6.12}$$

where MSE_i ($QLIKE_i$) is defined as the forecast MSE ($QLIKE$) over the out-of-sample period of any our competing models, and MSE_{HAR} ($QLIKE_{HAR}$) is the respective value of the HAR benchmark model. The values less than 1 are associated with the superior forecast ability of the proposed model, and vice versa.

In order to examine the advantages of the individual HAR models and the HAR models with forecast reconciliation methods over the HAR benchmark model, we then employ the Diebold and Mariano (1995) test (DM test) to investigate the null hypothesis of equal predictive accuracy (EPA) where the HAR model is used as a benchmark.

[6]The $QLIKE$ index is computed as average of a simple modification of the familiar Gaussian log-likelihood loss function, which belongs to the family of robust and homogeneous loss functions defined by Patton (2011b), with parameter $b = -2$. The modification is such that the index amounts to zero when $\widehat{RV}_t = RV_t$, that is, the daily observed RV is forecast without error (Patton and Sheppard, 2009, Patton, 2011a). It gives asymmetric weights to the forecast errors, so that underestimating the RV is more important than overestimating.

Finally, we utilize the Model Confidence Set (MCS) approach developed by Hansen *et al.* (2011) to compare the point forecast accuracy between the direct daily forecasts and the reconciliation-based forecasts using intraday decompositions of RV.

6.5 Does reconciliation help in RV forecasting?

Our final purpose is to answer the following question: when 'volatility-based' decompositions of the daily realized volatility are available, does considering forecast reconciliation significantly improve the forecast accuracy of daily RV compared to the benchmark HAR-type models by Corsi (2009), Patton and Sheppard (2015) and Bollerslev *et al.* (2022)?

We start by evaluating the direct forecasts accuracy for the DJIA using the MSE and $QLIKE$ ratios (Table 6.3, Panel A). First of all, it appears that the benchmark HAR model is almost always outperformed by both SV and $PV(3)$ in terms of $QLIKE$, the only exception being for $h = 1$. On the contrary, $h = 1$ and $PV(3)$ model is the only combination forecast horizon/model at which the HAR model is outperformed in terms of MSE. Second, the $QLIKE$ indices of the forecast reconciliation-based approaches, either indirect (bu) or regression-based (shr and hac), improve on both the HAR benchmark model (apart SV_{bu} at $h = 1$) and their direct counterparts. Again, this picture is not confirmed by the MSE indices, because of the different view at the forecasting accuracy offered by these two indices.

The individual stocks' forecast performance analysis (Table 6.3, Panel B) provides more compelling findings. We note that considering very simple intraday decompositions of RV_t in a forecast-reconciliation framework, either indirect, or regression-based, always improves on the forecasting accuracy of the HAR benchmark model: both MSE and $QLIKE$ are less than one at any forecast horizon. In addition, regression-based approaches always improve on their direct forecasting approaches counterparts, at any forecast horizon and in terms of both MSE and $QLIKE$ indices, with the most notable results being offered by $PV(3)_{shr}$ and $PV(3)_{hac}$, which stably give the best accuracy indices (highlighted in bold in Table 6.3, Panel B).

Following Hansen *et al.* (2011), we implement the MCS procedure using the block bootstrap of Politis and Romano (1994) (see Hansen, 2005), in which blocks have length of 22 days, and results are based on 10,000 resamples. We choose both the MSE and the $QLIKE$ loss functions, and use the test statistic T_{max} to test the null hypothesis of no difference between the forecast accuracy of the considered model. The results for the forecast horizons $h = 1, 5, 22$ are shown in Tables 6.4, 6.5 and 6.6, respectively.

	rMSE			rQLIKE		
	$h=1$	$h=5$	$h=22$	$h=1$	$h=5$	$h=22$
Panel A: DJIA index						
SV	1.028	1.017	1.003	0.973	0.997	0.928
SV_{bu}	0.960	0.972	1.005	1.010	0.237	0.748
SV_{shr}	0.976	0.991	1.002	0.983	0.233	0.728
SV_{hac}	0.976	0.979	**1.002**	0.969	0.231	0.719
$PV(3)$	0.816	1.077	1.025	2.272	0.977	0.962
$PV(3)_{bu}$	0.924	**0.957**	1.009	1.001	0.235	0.596
$PV(3)_{shr}$	0.833	1.015	1.015	0.945	0.225	0.596
$PV(3)_{hac}$	**0.796**	0.975	1.012	**0.940**	**0.225**	**0.590**
Panel B: Individual stocks						
SV	1.016	1.008	1.001	1.092	0.998	0.993
SV_{bu}	0.964	0.992	0.998	0.921	0.977	0.971
SV_{shr}	0.980	0.996	0.998	0.899	0.947	0.960
SV_{hac}	0.979	0.995	0.998	0.898	0.951	0.961
$PV(3)$	0.899	0.976	1.010	1.364	1.025	1.055
$PV(3)_{bu}$	0.943	0.982	0.995	0.900	0.895	0.945
$PV(3)_{shr}$	**0.896**	**0.964**	0.994	**0.833**	**0.824**	**0.916**
$PV(3)_{hac}$	0.899	0.966	**0.990**	0.842	0.840	0.930

Table 6.3: *Forecast accuracy at forecast horizons $h = 1, 5, 22$. $rMSE$ and $rQLIKE$ indices defined in (6.12) for the DJIA index (panel A), and geometric means of $rMSE$ and $rQLIKE$ for individual stocks (panel B). Values larger than one are highlighted in red. The best index value in each column is highlighted in bold.*

For the DJIA index, $PV(3)_{hac}$ is the best model at $h = 1$ and $h = 5$, even if almost all models are equivalent according to the threshold 0.2. For the Diebold-Mariano test, under MSE only few cases lead to a rejection of the null hypothesis, while with $QLIKE$ for $h = 5$ and $h = 22$ the models with forecast reconciliation improves over the benchmark models in a statistically significant way in most cases. Moving to the individual stocks, we stress that Tables 6.4, 6.5 and 6.6 report aggregated results, thus providing an overall evaluation in the cross-section of the 26 stocks. We highlight that even in this case performances differ between MSE and $QLIKE$: for the former, improvements are limited and only in few cases we do have rejections of the null for the Diebold-Mariano test, or models excluded from the confidence set; for the latter, the use of forecast reconciliation leads to a clear improvement, with $PV(3)_{shr}$ and $PV(3)_{hac}$ providing, overall, better performances.

These evidences are confirmed and enriched by Figure 6.4, which shows the results of the Multiple Comparison with the Best (MCB) Nemenyi test, a non-parametric multiple comparison procedure frequently adopted in the forecasting literature (see **?**, **?**, and **?**, among others). In particular, the single model $PV(3)$ does not significantly improve on the benchmark HAR (the corresponding lines in the 'Multiple Comparison with the Best' graphs for both MSE and $QLIKE$ are overlapping). Further, the forecasts

	RV	SV	$PV(3)$	SV_{bu}	$PV(3)_{bu}$	SV_{shr}	$PV(3)_{shr}$	SV_{hac}	$PV(3)_{hac}$
Panel A: DJIA index									
MSE	6.352	6.527	5.184	6.099	5.867	6.202	5.293	6.200	*5.057*
p-value dm_{RV}	–	0.785	0.168	0.142	0.075	0.119	0.066	0.108	0.073
p-value dm_{SV}	–	–	0.131	0.148	0.079	0.115	**0.044**	0.111	0.051
p-value dm_{PV}	–	–	–	0.800	0.759	0.821	0.579	0.820	0.374
p-value MCS	**0.241**	0.199	**0.717**	**0.354**	**0.472**	**0.278**	**0.717**	**0.297**	**1.000**
$QLIKE$	0.216	0.210	0.491	0.218	0.217	0.212	0.204	0.210	*0.203*
p-value dm_{RV}	–	**0.004**	0.908	0.692	0.528	0.176	**0.008**	0.058	**0.006**
p-value dm_{SV}	–	–	0.911	0.988	0.956	0.796	**0.031**	0.358	**0.020**
p-value dm_{PV}	–	–	–	0.098	0.097	0.093	0.087	0.091	0.086
p-value MCS	**0.206**	**0.206**	**0.206**	**0.206**	**0.206**	**0.206**	**0.206**	**0.206**	**1.000**
Panel B: Individual stocks									
$\overline{MSE}$	39.595	40.993	36.374	37.805	36.947	38.970	*35.553*	38.765	35.815
p-value dm_{RV}	–	0	0	1	2	1	2	1	1
p-value dm_{SV}	–	–	0	0	1	5	2	3	2
p-value dm_{PV}	–	–	–	1	1	2	1	0	2
p-value MCS	23	21	26	24	26	25	26	24	26
$\overline{QLIKE}$	0.185	0.214	0.269	0.170	0.165	0.165	*0.151*	0.165	0.152
p-value dm_{RV}	–	2	1	2	4	7	19	8	15
p-value dm_{SV}	–	–	1	6	7	10	20	10	15
p-value dm_{PV}	–	–	–	7	8	8	10	8	10
p-value MCS	19	21	18	20	21	21	26	21	26

Table 6.4: *One-day-ahead forecasting performance: 2007-2022 (3,880 days)*

Note: The table reports the **1-step ahead** forecasting performance of the different models. The top panel shows the results for the DJIA index, while the bottom panel refers to individual stocks. MSE and $QLIKE$ refer to the loss function value for a given model (top panel) or average loss function across individual stocks (bottom panel). The one-sided tests between each forecasting model against HAR, SV, and $PV(3)$ are denoted by dm_{HAR}, dm_{SV}, and $dm_{PV(3)}$, respectively. The top panel includes p-values while the bottom panel reports 5% rejection frequencies. MCS denotes the p-value of that model being in the Model Confidence Set (top panel), or the number of times that model is in the 80% Model Confidence Set (lower panel). $PV(3)$ uses three intraday decompositions defined by two thresholds at 10% and 75%. In the upper panel we highlight in bold p-values < 0.05 for Diebold-Mariano and p-values > 0.2 for MCS. In both panels, we highlight in italic the minimum (average) loss function.

	RV	SV	$PV(3)$	SV_{bu}	$PV(3)_{bu}$	SV_{shr}	$PV(3)_{shr}$	SV_{hac}	$PV(3)_{hac}$
Panel A: DJIA index									
MSE	5.109	5.195	5.503	4.966	*4.888*	5.064	5.185	4.999	4.982
p-value dm_{RV}	–	0.774	0.833	**0.021**	**0.007**	0.235	0.613	**0.029**	0.227
p-value dm_{SV}	–	–	0.783	0.054	**0.023**	**0.049**	0.485	**0.017**	0.110
p-value dm_{PV}	–	–	–	0.105	0.077	0.139	**0.027**	0.110	**0.026**
p-value MCS	**0.403**	**0.403**	**0.403**	**0.646**	**1.000**	**0.470**	**0.470**	**0.470**	**0.646**
$QLIKE$	0.930	0.927	0.908	0.220	0.219	0.217	0.209	0.214	*0.209*
p-value dm_{RV}	–	**0.000**	0.441	**0.008**	**0.008**	**0.008**	**0.007**	**0.007**	**0.007**
p-value dm_{SV}	–	–	0.449	**0.008**	**0.008**	**0.008**	**0.007**	**0.008**	**0.007**
p-value dm_{PV}	–	–	–	**0.007**	**0.007**	**0.007**	**0.006**	**0.007**	**0.006**
p-value MCS	**0.429**	**0.484**	0.140	**0.484**	**0.484**	**0.484**	**0.717**	**0.484**	**1.000**
Panel B: Individual stocks									
$\overline{MSE}$	25.109	25.371	22.936	24.804	24.433	24.923	*22.753*	24.813	23.551
p-value dm_{RV}	–	0	1	2	9	4	2	4	5
p-value dm_{SV}	–	–	1	2	4	4	3	3	4
p-value dm_{PV}	–	–	–	0	0	0	0	0	0
p-value MCS	26	22	26	26	26	25	26	24	26
$\overline{QLIKE}$	0.186	0.187	0.201	0.181	0.161	0.174	*0.147*	0.174	0.149
p-value dm_{RV}	–	6	7	4	8	9	24	6	20
p-value dm_{SV}	–	–	7	2	4	3	21	1	18
p-value dm_{PV}	–	–	–	2	4	2	8	2	7
p-value MCS	15	19	23	14	18	19	25	18	22

Table 6.5: *Five-day-ahead forecasting performance: 2007-2022 (3,880 days)*

Note: The table reports the **5-step ahead** forecasting performance of the different models. The top panel shows the results for the DJIA index, while the bottom panel refers to individual stocks. MSE and $QLIKE$ refer to the loss function value for a given model (top panel) or average loss function across individual stocks (bottom panel). The one-sided tests between each forecasting model against HAR, SV, and $PV(3)$ are denoted by dm_{HAR}, dm_{SV}, and $dm_{PV(3)}$, respectively. The top panel includes p-values while the bottom panel reports 5% rejection frequencies. MCS denotes the p-value of that model being in the Model Confidence Set (top panel), or the number of times that model is in the 80% Model Confidence Set (lower panel). $PV(3)$ uses three intraday decompositions defined by two thresholds at 10% and 75%. In the upper panel we highlight in bold p-values < 0.05 for Diebold-Mariano and p-values > 0.2 for MCS. In both panels, we highlight in italic the minimum (average) loss function.

	RV	SV	$PV(3)$	SV_{bu}	$PV(3)_{bu}$	SV_{shr}	$PV(3)_{shr}$	SV_{hac}	$PV(3)_{hac}$
Panel A: DJIA index									
MSE	*4.424*	4.436	4.533	4.448	4.463	4.434	4.488	4.431	4.475
p-value dm_{RV}	–	0.613	0.851	0.631	0.683	0.569	0.780	0.555	0.774
p-value dm_{SV}	–	–	0.843	0.575	0.646	0.478	0.765	0.456	0.750
p-value dm_{PV}	–	–	–	0.209	0.257	0.158	0.225	0.148	0.190
p-value MCS	**1.000**	**0.903**	**0.743**	**0.842**	**0.842**	**0.903**	**0.809**	**0.903**	**0.842**
$QLIKE$	0.973	0.902	0.936	0.727	0.580	0.708	0.580	0.699	*0.574*
p-value dm_{RV}	–	0.160	0.150	**0.027**	**0.002**	**0.016**	**0.002**	**0.014**	**0.002**
p-value dm_{SV}	–	–	0.830	**0.049**	**0.004**	**0.028**	**0.003**	**0.023**	**0.003**
p-value dm_{PV}	–	–	–	**0.031**	**0.002**	**0.017**	**0.002**	**0.014**	**0.002**
p-value MCS	**0.265**	**0.314**	**0.314**	**0.314**	**0.555**	**0.314**	**0.555**	**0.373**	**1.000**
Panel B: Individual stocks									
$\overline{MSE}$	14.243	14.277	14.316	14.205	14.231	14.207	*14.122*	14.242	14.166
p-value dm_{RV}	–	1	2	2	4	3	3	4	6
p-value dm_{SV}	–	–	2	1	2	2	2	1	5
p-value dm_{PV}	–	–	–	3	2	3	6	3	7
p-value MCS	24	24	25	24	25	24	24	24	24
$\overline{QLIKE}$	0.269	0.268	0.292	0.262	0.255	0.259	*0.247*	0.259	0.250
p-value dm_{RV}	–	5	6	4	10	6	22	6	19
p-value dm_{SV}	–	–	5	3	6	4	17	5	16
p-value dm_{PV}	–	–	–	8	10	9	14	9	13
p-value MCS	20	22	24	19	20	21	24	19	24

Table 6.6: *Twenty-two-day-ahead forecasting performance: 2007-2022 (3,880 days)*

Note: The table reports the **22-step ahead** forecasting performance of the different models. The top panel shows the results for the DJIA index, while the bottom panel refers to individual stocks. MSE and $QLIKE$ refer to the loss function value for a given model (top panel) or average loss function across individual stocks (bottom panel). The one-sided tests between each forecasting model against HAR, SV, and $PV(3)$ are denoted by dm_{HAR}, dm_{SV}, and $dm_{PV(3)}$, respectively. The top panel includes p-values while the bottom panel reports 5% rejection frequencies. MCS denotes the p-value of that model being in the Model Confidence Set (top panel), or the number of times that model is in the 80% Model Confidence Set (lower panel). $PV(3)$ uses three intraday decompositions defined by two thresholds at 10% and 75%. In the upper panel we highlight in bold p-values < 0.05 for Diebold-Mariano and p-values > 0.2 for MCS. In both panels, we highlight in italic the minimum (average) loss function.

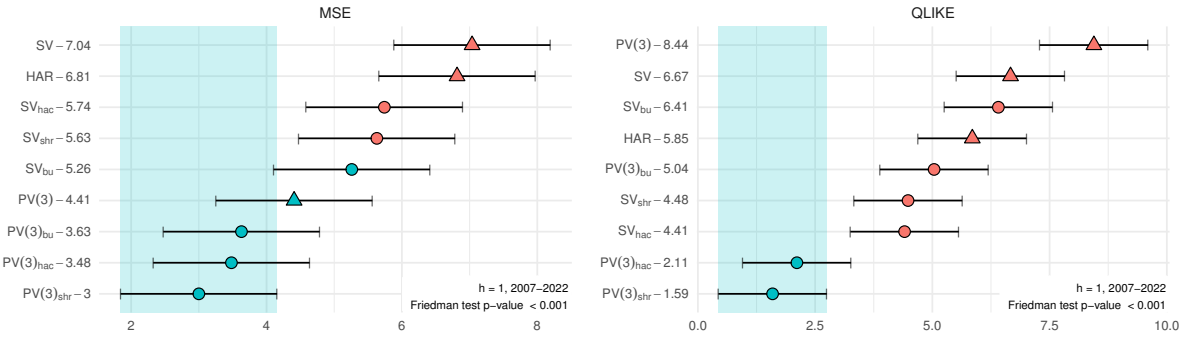


Figure 6.4: *MCB Nemenyi test results: average ranks and 95% confidence intervals for the* ***one-step ahead*** *RV forecasts of the DJIA index and 26 individual stocks. Direct daily RV forecasts from HAR, SV and PV(3) models, and from their extensions with the bottom-up (bu) and the MinT-shr (shr) forecast reconciliation-based approaches according to the corresponding intraday RV decomposition. The forecasting approaches are sorted vertically according to the MSE mean rank (left panel) and the QLIKE mean rank (right panel). The mean rank of each method is displayed to the right of their names. If the intervals of two forecasting models do not overlap, this indicates a statistically different performance. Thus, methods that do not overlap with the light blue interval are considered significantly worse than the best and vice versa.*

produced by $PV(3)_{shr}$ and $PV(3)_{hac}$ are significantly better than the benchmarks HAR and SV, both in terms of MSE and $QLIKE$, while direct $PV(3)$ forecasts appear significantly worse if the $QLIKE$ loss function is used to evaluate the forecast accuracy. Finally, overall $PV(3)_{shr}$ shows the best forecast accuracy: it is not significantly worse than $PV(3)_{bu}$, the most performing approach in terms of MSE, and ranks first in terms of $QLIKE$, while $PV(3)_{hac}$ is the only approach with a statistically equivalent forecasting accuracy.

Further insights are offered in a simple, but effective descriptive view, by Figure 6.5, where the scatter plots of the 27 couples of, respectively, MSE and $QLIKE$ indices obtained using the benchmark HAR model and $PV(3)_{shr}$ are represented. It emerges that $PV(3)_{shr}$ outperforms the benchmark HAR in the majority of cases (23 out of 27) in terms of MSE, and always if $QLIKE$ is used to evaluate the forecasting accuracy. This consideration is somehow extended, and further supported, by the results shown in Figure 6.6, which contains a summary view of the number of times each forecasting approach provided better forecasting accuracy than the other procedures considered in the comparison. Summarizing, $PV(3)_{shr}$ registers a better prediction performance than all other approaches, with success rates in terms of $QLIKE$ ranging from 96.3% (26 of 27) to 100% (27 of 27), and from 55.6% (15 of 27) to 88.9% (24 of 27) if MSE is used.

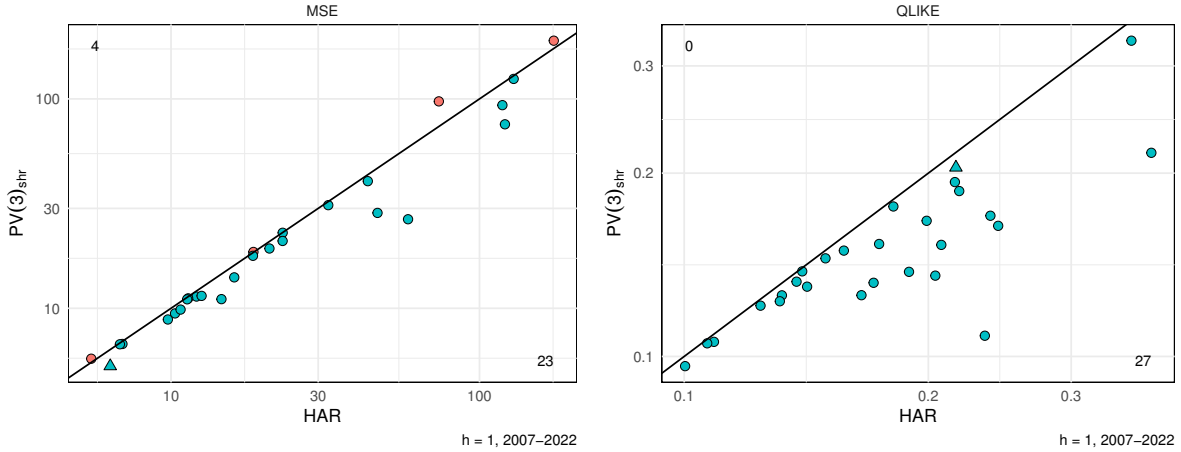


Figure 6.5: *Accuracy of the* ***one-step ahead*** *daily RV forecasts for the DJIA index (triangle) and 26 individual stocks (circle) in terms of MSE (left panel) and QLIKE (right panel) indices. Comparison between HAR direct and* $PV(3)_{shr}$ *reconciliation-based forecasts. The black line represents the bisector, where either MSE's or QLIKE's for both approaches are equal. On the top-left (bottom-right) corner of each graph, the number of points above (below) the bisector is reported.*

6.6 Additional results

6.6.1 Sub-sample analysis

The interpretation of the results so far can be further detailed and specified by considering 4 different time windows of the complete 2007-2022 interval previously analysed, namely: 2006-2010, 2011-2014, 2015-2019, 2020-2022 (see the online appendix). We observe that the predictive performance of models that make use of intraday decompositions in a reconciliation framework is constantly better in periods of high market volatility (2007-2010 and 2020-2022), while they do not worsen the predictive accuracy of the benchmark in the period 2015-2019, which is characterized by lower variances. The remaining period 2011-2014 offers different indications depending on whether one considers $QLIKE$ or MSE. In the latter case, the $PV(3)$ model, both in the single version and in the version that makes use of forecast reconciliation, shows a better overall accuracy, although not significantly to the HAR benchmark.

6.6.2 Grouped series

Here we consider the forecasting performance of reconciliation approaches applied to grouped time series defined by intraday RV decompositions based on time and returns' characteristics:

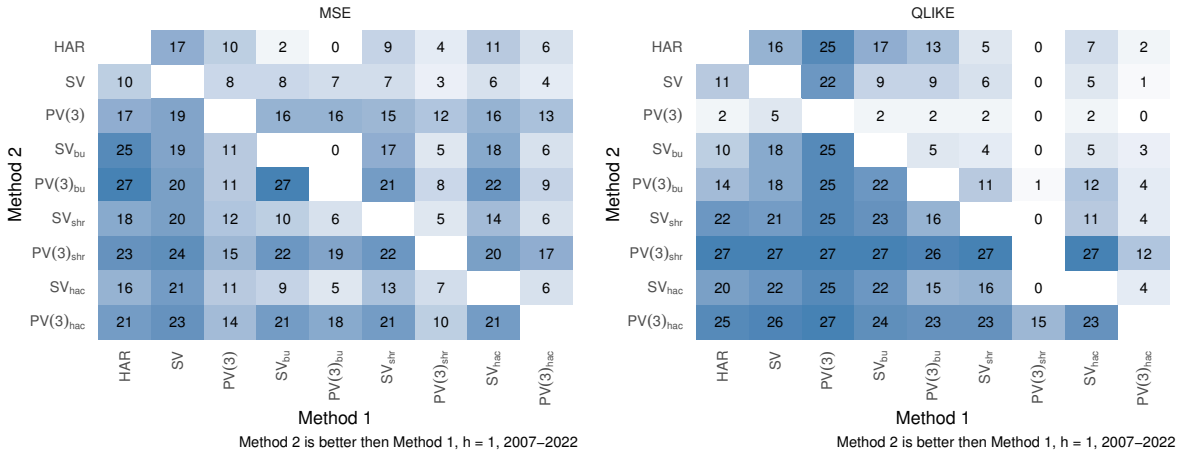


Figure 6.6: *Qualitative evaluation of the* ***one-step ahead*** *forecasting accuracy. Each cell reports the number of times the forecasting model in the row outperforms the model in the column.*

- TSV_{bu}: indirect (bottom-up) daily RV forecasts from ten bottom time series cross-classified by time interval and semi-variances;
- TSV_{shr}: two hierarchies sharing the ten bottom variables above, with eight upper variables;
- $TPV(3)_{bu}$: indirect (bottom-up) daily RV forecasts from fifteen bottom time series cross-classified by time interval and partial variances;
- TPV_{shr}: two hierarchies sharing the fifteen bottom variables above, with nine upper variables.

From Table 6.7 it appears that the regression-based reconciliation approach $TPV(3)_{shr}$ largely benefits from the adoption of a time decomposition (all indices are less than one). However, looking at Figure 6.7, $TPV(3)_{shr}$ gives results largely similar to those of the simpler $PV(3)_{shr}$ approach, that does not make use of any temporal decomposition.

6.6.3 An alternative PV decomposition

In the $PV(3)$ decomposition, the use of RV_t in the standardization of intraday returns exposes the estimators of standardized returns quantile to the influence of price jumps (i.e., extreme positive or negative returns) and of the intraday volatility pattern.

We consider here a slight modification in the construction of the decomposition by employing the contribution of Boudt *et al.* (2011) and exploiting a result of Andersen *et al.* (2007b). The former provide a methodology for estimating and filtering out the intraday variance pattern while the latter shows that, in large sample, the intraday

	rMSE			rQLIKE		
	$h=1$	$h=5$	$h=22$	$h=1$	$h=5$	$h=22$
Panel A: DJIA index						
TSV_{bu}	0.824	0.801	1.099	1.085	0.249	0.385
TSV_{shr}	0.869	0.878	0.965	1.011	0.236	0.379
$TPV(3)_{bu}$	0.822	**0.795**	1.103	1.090	0.250	0.386
$TPV(3)_{shr}$	**0.808**	0.925	**0.939**	**0.971**	**0.227**	**0.374**
Panel B: Individual stocks						
TSV_{bu}	0.991	1.054	1.019	0.893	0.908	0.875
TSV_{shr}	0.987	1.007	0.984	0.850	0.861	0.853
$TPV(3)_{bu}$	0.982	1.039	1.012	0.894	0.902	0.868
$TPV(3)_{shr}$	**0.924**	**0.967**	**0.965**	**0.805**	**0.793**	**0.824**

Table 6.7: *Forecast accuracy at forecast horizons $h = 1, 5, 22$. MSE and $QLIKE$ ratios over the benchmark HAR model for the DJIA index (panel A), and geometric means of the MSE and $QLIKE$ ratios for individual stocks (panel B). Values larger than one are highlighted in red. The best index value in each column is highlighted in bold.*

returns standardized by the Bi-power variation follow an asymptotically normal distribution. We combine these two elements as follows. First, we compute returns filtered from the intraday volatility pattern by adopting the *WDS* approach of Boudt *et al.* (2011). The filtered returns are defined as $r^{\star}_{i,t} = \dfrac{r_{i,t}}{\sigma_i}$ where σ_i is the intraday volatility at interval i, estimated with the *WDS* approach. Then, we evaluate the Bi-power variation, a jump-robust estimator of the integrated volatility, on the filtered returns: $\widetilde{BPV_t} = \dfrac{\pi}{2}\dfrac{N}{N-1}\sum_{i=2}^{M}|r^{\star}_{i,t}||r^{\star}_{i,t}|$. Finally, we assume the following distribution, for the filtered and standardized returns

$$\frac{r_{i,t}}{\sigma_i\sqrt{N^{-1}\widetilde{BPV_t}}} \sim \mathcal{N}(0,1), \tag{6.13}$$

This quantity has already been used in Boudt *et al.* (2011) for jump detection. Going back to the PV framework, the empirical quantiles of standardized returns might be replaced by theoretical quantiles of the asymptotic distribution leading to

$$\tilde{c}_{i,t,j} = \sigma_i\sqrt{N^{-1}\widetilde{BPV_t}}\Phi^{-1}(\tau_j), \quad j = 1, 2, \dots p-1. \tag{6.14}$$

Finally, we note that thresholds are both day-specific and intra-day-interval-specific. Similarly to the $PV(3)$, we now define a $PV(3)^{\star}$ decomposition where we employ the previously defined quantiles with $\tau_1 = 0.1$ and $\tau_2 = 0.75$.

In Figure 6.8 the MSE and $QLIKE$ indices for the bottom-up reconciliations in the two cases are shown. It clearly appears that no meaningful accuracy improvement is obtained for the considered assets, and this turns out to be confirmed by Figure 6.9, where the regression-based reconciliation approaches in the two cases are considered.

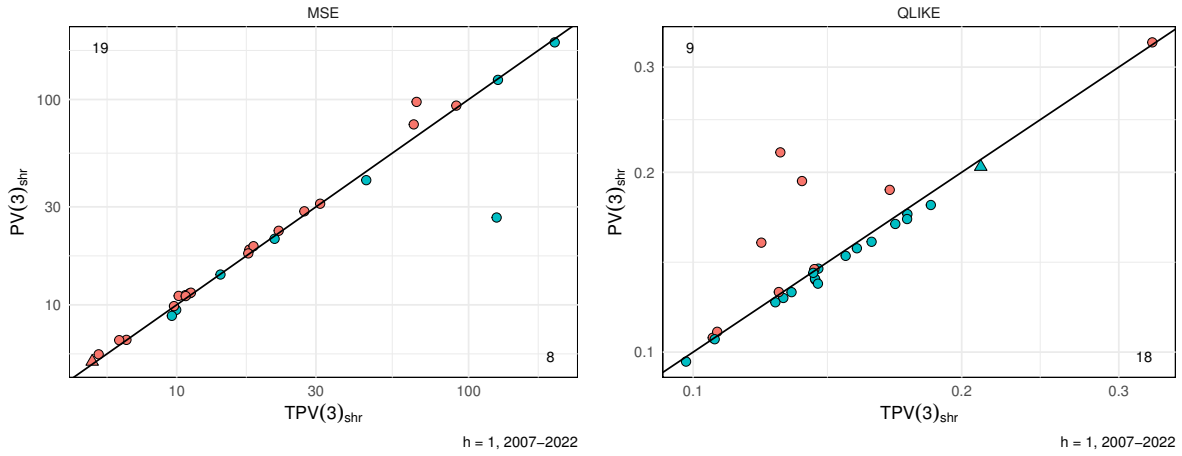


Figure 6.7: *Accuracy of the* ***one-step ahead*** *daily RV forecasts for the DJIA index (triangle) and 26 individual stocks (circle) in terms of MSE (left panel) and QLIKE (right panel) indices. Comparison between HAR direct and* $PV(3)_{shr}$ *reconciliation-based forecasts. The black line represents the bisector, where either MSE's or QLIKE's for both approaches are equal. On the top-left (bottom-right) corner of each graph, the number of points above (below) the bisector is reported.*

6.6.4 Decomposition optimality

We close this section by focusing on an issue intimately related to our approach: when is forecast reconciliation optimal? Or, in other words, can we identify an optimal *RV* decomposition? Bollerslev *et al.* (2022) provide an extensive analysis on the choice of the number of thresholds and of the thresholds value when focusing on partial variances. They show heterogeneity in the outcome across stocks, with $PV(3)$ being the best choice, with median threshold values at 0.1 and 0.75, the values we used in our empirical analysis. We try here to complement their findings with two examples based on a single series, the DJIA index.

In the first example, we focus on the optimal choice of quantiles. We compare the performances of HAR, $PV(3)$ and two alternative prediction approaches based on forecast reconciliation using $PV(3)$ components: one adopting the shrinkage approach, and the other one using the HAC forecast error covariance. We report results in terms of MSE ratios taking the HAR as the benchmark. We set the $PV(3)$ thresholds at two quantiles, q_1 and $q_2 = 1 - q_1$, with q_1 varying from 0.01 to 0.49. In Figure 6.10 we report the results. Notably, all approaches provide a reduction in MSE compared to HAR: for $PV(3)$, this is coherent with the evidence in Bollerslev *et al.* (2022), while for the forecast reconciliation cases, results are in line with the evidence in our work. In addition, the use of robust covariance is always improving the forecasts. We also note that, for the best model, the selected quantiles are 0.45 and 0.55, thus separating the central part

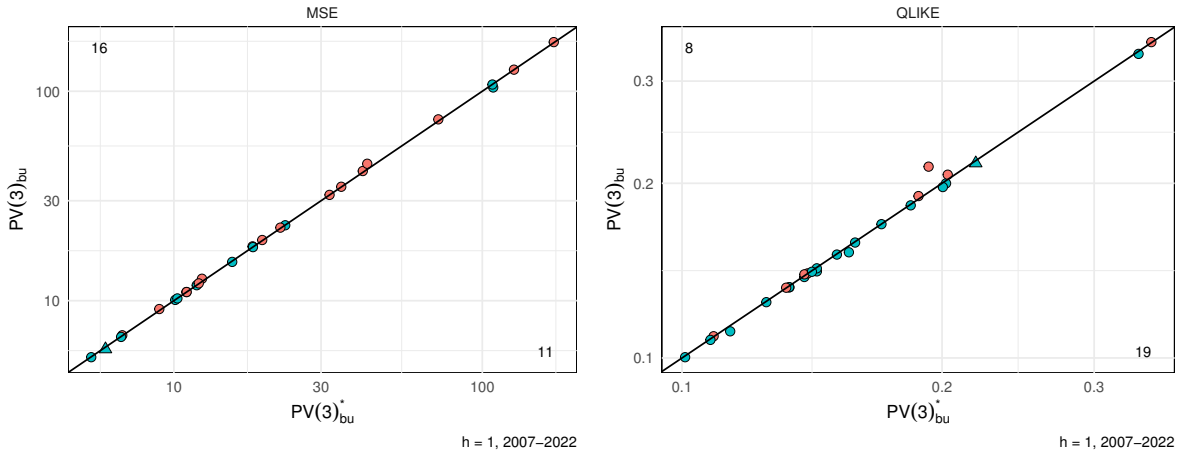


Figure 6.8: *Accuracy of the* ***one-step ahead*** *daily RV forecasts for the DJIA index (triangle) and 26 individual stocks (circle) in terms of MSE (left panel) and QLIKE (right panel) indices. Comparison between HAR direct and* $PV(3)_{shr}$ *reconciliation-based forecasts. The black line represents the bisector, where either MSE's or QLIKE's for both approaches are equal. On the top-left (bottom-right) corner of each graph, the number of points above (below) the bisector is reported.*

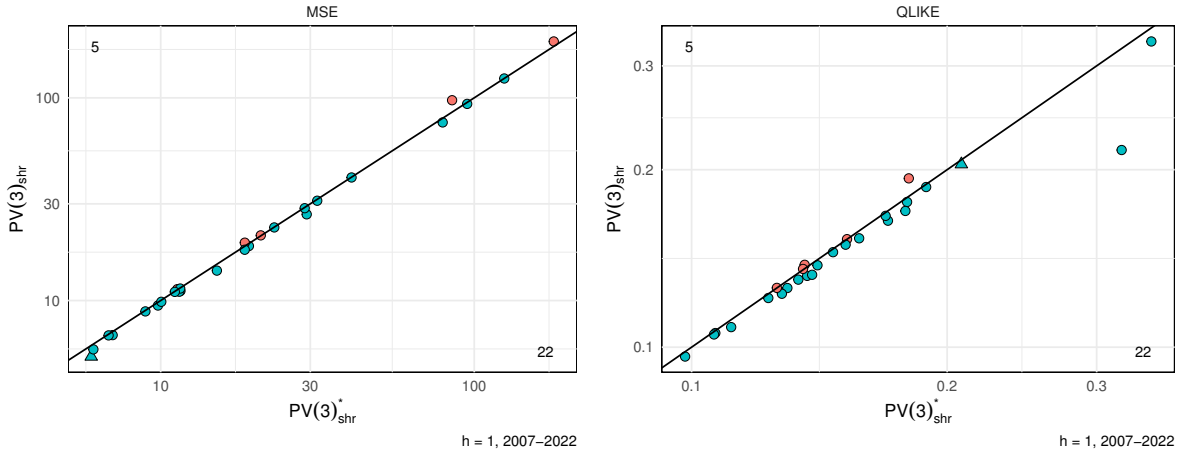


Figure 6.9: *Accuracy of the* ***one-step ahead*** *daily RV forecasts for the DJIA index (triangle) and 26 individual stocks (circle) in terms of MSE (left panel) and QLIKE (right panel) indices. Comparison between HAR direct and* $PV(3)_{shr}$ *reconciliation-based forecasts. The black line represents the bisector, where either MSE's or QLIKE's for both approaches are equal. On the top-left (bottom-right) corner of each graph, the number of points above (below) the bisector is reported.*

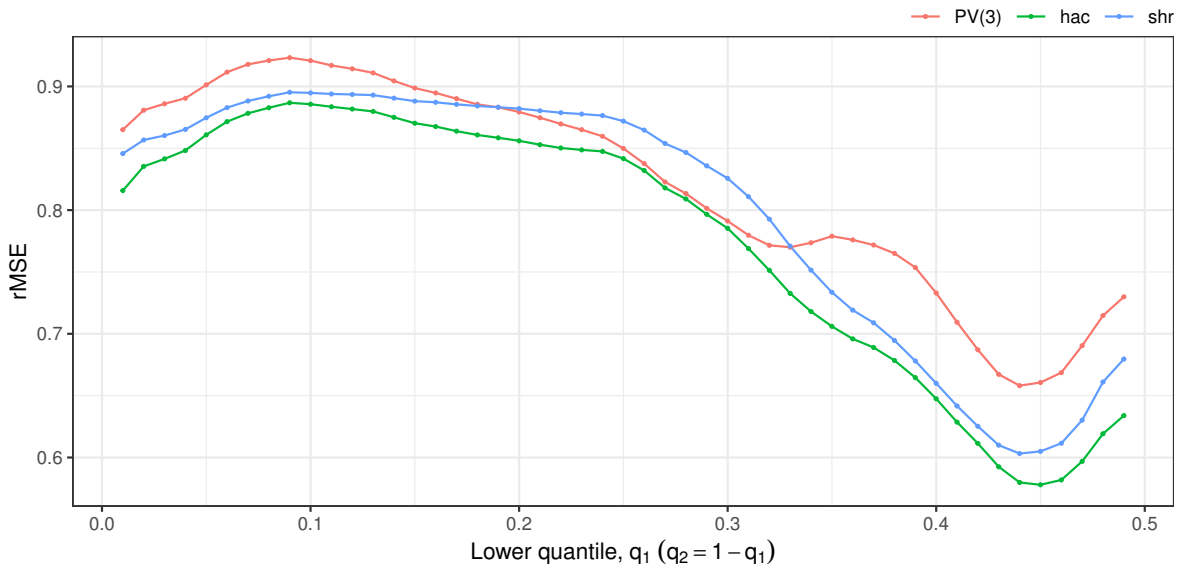


Figure 6.10: *Relative MSE with respect to the HAR model as benchmark (lower is better); the $PV(3)$ model is based on the q_1 and $q_2 = 1 - q_1$ quantiles; shr denotes shrinkage and hac the robust covariance approach; the minimum value is $q_1 = 0.45$ for hac, and $q_1 = 0.44$ for $PV(3)$ and shr.*

of the returns distribution (where returns are really close to zero), from the positive and negative returns (above a relatively small threshold). These quantiles cannot be generalized to all series due to the heterogeneity already noticed by Bollerslev *et al.* (2022). Nevertheless, such an approach might be used to identify optimal thresholds in a forecast reconciliation framework.

In the second example, we focus again on the HAR model as benchmark, but now we change the number of components in the alternative specifications, that is, we move to $PV(g)$, with $g = 2, 3, \dots 15$. Quantiles are here chosen to create an equally sized decomposition, that is, for $g = 2$ we set $q_1 = 0.5$, for $g = 3$ we use $q_1 = \frac{1}{3}$ and $q_2 = \frac{2}{3}$, and so on. This allows us to verify if there is an optimal number of components to include in the forecast reconciliation procedure. Results for the DJIA index are reported in Figure 6.11.

We observe that the minimum is at $g = 8$ for the best performing approach, the use of forecast reconciliation with a robust forecast errors covariance estimator. All models improve over the HAR for $2 < g < 10$. Across models, the introduction of a robust covariance improves results for $g < 11$. We believe these results are reasonable as by increasing the number of quantiles we identify decompositions being based on smaller number of data points, and most likely some partial variances will have a value close to zero (i.e., for the components around the center of the returns distribution, where returns are almost equal to zero) and likely add just noise to the entire procedure. Of

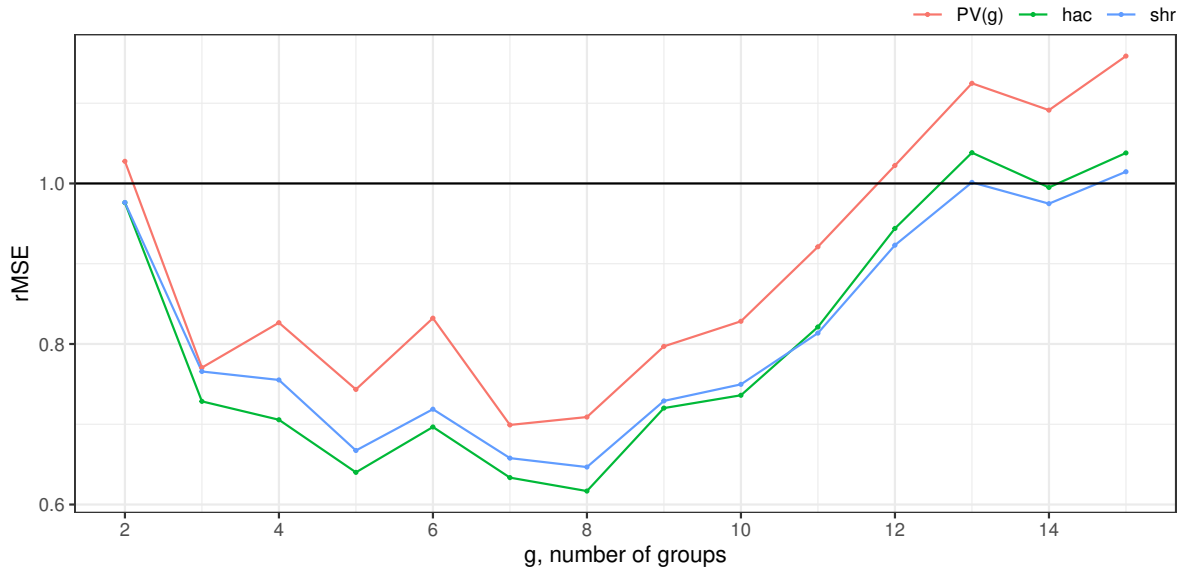


Figure 6.11: *Relative MSE with the HAR model as benchmark (lower is better); the 390 observations are splitted according to equally spaced quantiles; shr denotes shrinkage and hac the robust covariance approach; the minimum value is $g = 8$ for hac and shr, and $g = 7$ for PV.*

course, improvements can be made by an optimal selection of quantiles, combining them with an optimal temporal decomposition. In addition, the choice of the optimal decomposition might also be affected by assets liquidity. In fact, we might expect that less liquid assets would be associated with a smaller number of components in the optimal specification. Again, by contrasting results across competing RV decompositions, we show that improvements are clearly present and an optimal number of components can be identified on an asset-specific case.

Conclusions

Discussion and future research directions

This thesis has focused on various themes related to cross-sectional, temporal, and cross-temporal forecast reconciliation of time series. Methodological, computational, and applicative aspects have been explored, aiming to find solutions that are both statistically well-grounded and effective in achieving improved forecasts with real data. The results obtained demonstrate promising outcomes, indicating the potential for advancing the field of forecast reconciliation.

Optimal and heuristic methods for cross-temporal reconciliation

The hierarchical framework is currently considered an effective way to improve the accuracy of forecasts in many different fields of application. In this Chapter, we give contributions and extensions to a topic that has been widely studied in the last decade by connecting it to the widespread literature on least-squares adjustment of preliminary data (Stone *et al.*, 1942; Byron, 1978), with a focus on a projection approach for linearly constrained multiple time series. This *de facto* encompasses and extends the modeling framework by Hyndman *et al.* (2011) (see also Panagiotelis *et al.*, 2021, and Wickramasuriya *et al.*, 2019). However, we do agree with Jeon *et al.* (2019), p. 368, that a "shortcoming of many of the approaches above (...) is that the weights (...) are a function of in-sample errors and are not directly determined with reference to an objective function ultimately used to assess forecast quality". This problem, yet present for cross-temporal hierarchies, is added to the dimensionality issues that characterize these structures. The number of nodes is considerably larger than the relevant single-dimension hierarchies and calls for alternative estimation strategies, based for example, on cross validation, as proposed by Jeon *et al.* (2019), or - when enough data is available - on Machine Learning techniques (Mancuso *et al.*, 2021; Spiliotis *et al.*, 2021).

Nevertheless, cross-temporal point forecast reconciliation seems to be a promising theme worth considering for future research. Some topics in this field are (i) simulation experiments to better understand behavior, potentiality, and possible shortcomings of

the proposed procedures, (ii) more realistic (and hopefully effective) approximations of the covariance matrices for cross-temporal reconciliation; (iii) extending the cross-temporal framework to the reconciliation of probabilistic forecasts (Jeon *et al.*, 2019; Ben Taieb *et al.*, 2021; Panagiotelis *et al.*, 2023), and for Bayesian (Eckert *et al.*, 2021; Corani *et al.*, 2021) and fast (Ashouri *et al.*, 2022) forecast reconciliation procedures; (iv) extending the cross-temporal optimal combination approach to the case of intermittent demand forecasts (Petropoulos and Kourentzes, 2015), with the related non-negativity issues (Wickramasuriya *et al.*, 2020; Kourentzes and Athanasopoulos, 2021), and possible consideration of 'soft' constraints (Danilov and Magnus, 2008).

Cross-temporal reconciliation of solar forecasts

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 chapter, 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. In addition to assuring both cross-sectional 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 geographical level of the hierarchy and for any temporal aggregation order. It is worth noting that for the hourly PV generated power at any cross-sectional level, all forecasts skills are positive and range between 4.7% and 18.4% when the NWP 3TIER is used as reference model.

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; Yang, 2020; Yagli *et al.*, 2020; Ben Taieb *et al.*, 2021; Panagiotelis *et al.*, 2023).

Forecast combination-based forecast reconciliation

The Level-l Conditional Coherent cross-sectional forecast reconciliation approach recently proposed by Hollyman *et al.* (2021) has been re-visited and extended, showing that it grounds on the solution to a quadratic minimization problem with exact linear inhomogeneous constraints, whose known constant part is exogenously given by the upper time series base forecasts. We have also provided the expressions valid for a level conditional coherent reconciliation with endogenous constraints, where the upper time series forecasts are no longer considered as binding constraints, but are admitted to be revised in view of their variability.

The forecasting experiment on the Australian Tourism Demand dataset by Hollyman *et al.* (2021) has been extended accordingly, in order to compare the accuracy performance of the state-of-the-art optimal combination forecast reconciliation procedures (Wickramasuriya *et al.*, 2019) with those offered by simple averaging of the L_lCC reconciled forecasts. Results have been found (i) considering the non-negativity issues posed by the data, (ii) using the relative accuracy metrics recommended by Davydenko and Fildes (2013), and (iii) taking into account the role played by the bts base forecasts in the LCC and CCC forecast combination, which allows to establish a 'fair' comparison with the optimal combination forecasts reconciliation procedure. The learned lesson from this dataset is that pooling reconciled forecasts may play a positive and important role in forecast reconciliation. Applying simple pooling techniques, while fulfilling all the aggregation constraints, improves the quality of the single constituent forecasts (Abouarghoub *et al.*, 2018), in agreement with the vast amount of empirical evidence of the last five decades in the field of forecast combination (Bates and Granger, 1969; Clemen, 1989; Timmermann, 2006). Furthermore, the original intuition by Hollyman *et al.* (2021) of combining bottom time series base forecasts from different models (e.g., seasonal averages and automatic ETS) has been conveniently re-stated (and reinforced) by considering a forecast averaging strategy involving all the available forecasts.

The empirical application has considered another dataset (quarterly Australian GDP from both Income and Expenditure sides), already used in the hierarchical forecasting literature (Athanasopoulos *et al.*, 2020). Rather interesting, unlike the Australian Tourism dataset, in this case the CCC_H approach proposed by Hollyman *et al.* (2021), adopted by considering either seasonal averages or seasonal random walk as naive bottom time series base forecasts, does not perform well. It is confirmed that LCC and CCC, both the exogenous and the endogenous variants, improve on the base forecasts, and

have good forecasting accuracy, though not superior to the state-of-the-art approaches *wls* and *shr*. These findings suggest that *LCC* and *CCC* approaches are useful forecasting tools, simple to design and easy to implement and use, worth to be considered by the practitioners along with the standard reconciliation approaches available in literature.

In this chapter, only simple forecast averaging has been considered, in line with the idea that it is generally not worse (and often it is better) than more sophisticated weighting schemes (Genre *et al.*, 2013). Nevertheless, we think it would be interesting, and potentially fruitful, to consider alternative forecast pooling methods (Kourentzes *et al.*, 2019; Lichtendahl and Winkler, 2020). In addition, in order to exploit information differences and mitigate model uncertainty, the *LCC* approach could be easily extended to combine forecasts from multiple temporal aggregation levels (Athanasopoulos *et al.*, 2017), provided a sensible forecast error covariance matrix be considered (Nystrup *et al.*, 2020, 2021). By continuing on this path, since leveraging both cross-sectional and temporal hierarchies using cross-temporal reconciliation approaches has shown to be effective in order to forecast a linearly constrained multiple time series as shown in Chapter 1, forecast combination based forecast reconciliation could be adapted to this challenging framework as well, in order to gain predictive accuracy. Finally, as in practical applications a thorough forecast accuracy evaluation needs to deal with predictive distributions rather than point forecasts, the forecast combination approach to probabilistic forecast reconciliation (Jeon *et al.*, 2019; Panagiotelis *et al.*, 2023; Yang, 2020; Yagli *et al.*, 2020; Ben Taieb *et al.*, 2021; Wickramasuriya, 2024) is a valuable topic worth considering for future research.

General linearly constrained multiple time series reconciliation

Producing and using coherent information, disaggregated by different characteristics useful for different decision levels, is an important task for any practitioner and quantitative-based decision process. At this end, in this article we aimed to generalize the results valid for the forecast reconciliation of a genuine hierarchical/grouped time series to the case of a general linearly constrained time series, where the distinction between upper and bottom variables, which is typical in the hierarchical setting, is no longer meaningful.

Two motivating examples have been considered, both coming from the National Accounts field, namely the forecast reconciliation of quarterly *GDP* of (i) Australia, disaggregated by income and expenditure variables, and (ii) Euro Area 19, disaggregated by the income, output and expenditure side variables of 19 component countries. In both cases, the structure of the time series involved cannot be represented according

to a genuinely hierarchical/grouped scheme, so the standard forecast reconciliation techniques fail in producing a "unique" GDP forecast, either point or probabilistic, making it necessary to solve this annoying issue.

We have shown that using well known linear algebra tools, it is always possible to establish a formal connection between the unconstrained least squares structural approach originally developed by Hyndman *et al.* (2011), and the projection approach dating back to the work by Stone *et al.* (1942), and then applied to solve different reconciliation problems (Di Fonzo and Marini, 2011; van Erven and Cugliari, 2015; Wickramasuriya *et al.*, 2019, see also Chapter 1). We propose a new classification of the variables forming the multiple time series as free and constrained, respectively, that can be seen as a generalization of the standard bottom/upper variables classification used in the hierarchical setting. Furthermore, we show techniques for deriving a linear combination matrix describing the relationships between these variables, starting from the coefficient matrix summarizing the (possible redundant) constraints linking the series.

The application of these findings to both point and probabilistic reconciliation techniques proved to be easy to implement and powerful, resulting in significant improvements in the forecasting accuracy of GDP and its components in both forecasting experiments.

Cross-temporal probabilistic forecast reconciliation

In this chapter, we extend the probabilistic reconciliation setting developed by Panagiotelis *et al.* (2023) for the cross-sectional case to the cross-temporal framework. Through appropriate notation, we show how theorems and definitions valid for the cross-sectional case can be reinterpreted and extended. The general notation proposed can help investigate extensions following different probabilistic approaches, such as those in Jeon *et al.* (2019), Ben Taieb *et al.* (2021) and Corani *et al.* (2023). We propose a Gaussian and a bootstrap approach to simulate the base forecasts able to take into account both cross-sectional and temporal relationships simultaneously, opening the way for further research into cross-temporal probabilistic forecasting.

Moreover, we analyze the use of residuals, showing that one-step residuals fail to capture the temporal structure, and propose multi-step residuals that can fully capture the cross-temporal relationships. Due to the high-dimensionality of the cross-temporal setting when dealing with covariance matrices, we propose four alternative forms to reduce the number of parameters to be estimated, showing that the overlapping residuals may reduce the high-dimensionality burden by increasing the number of available residuals. These ideas are worth requiring further investigation in future works.

Finally, we perform empirical applications on two datasets commonly used in forecast reconciliation research: Australian GDP from Income and Expenditure sides and Australian Tourism Demand. We find that in both cases optimal cross-temporal reconciliation approaches significantly improve on base forecasts. We also compare these with partly bottom-up techniques that use uni-dimensional reconciliation (either cross-sectional or temporal) and confirm that simultaneously exploiting both dimensions in reconciliation produces better results, especially at higher levels of temporal aggregation. This is more evident in the Australian Tourism Demand application, where the involved temporal hierarchies are richer, allowing the regression-based forecast reconciliation approaches to capture and exploit more features of the data through the available temporal aggregation levels (Kourentzes *et al.*, 2014; Kourentzes and Petropoulos, 2016; Kourentzes *et al.*, 2017) compared to partly bottom-up. In these two datasets, oct($wlsv$) and oct($bdshr$) appear as the two best performing approaches, both in terms of improving forecast accuracy and computational efficiency (see the online appendix), thus corroborating the results of Chapter 1 for point forecast reconciliation.

In conclusion, cross-temporal forecast reconciliation is an important tool to improve the accuracy of forecasts while simultaneously ensuring their coherency both in space and time. Furthermore, these techniques can also be customized to suit the specific needs of an organization, allowing for the incorporation of relevant domain-specific knowledge (e.g., non negative constraints) and expertise, ensuring that the resulting forecasts are not only accurate but also coherent and more reliable for decision-making purposes.

Further developments will focus on the generalization of other cross-sectional and temporal probabilistic approaches (Jeon and Taylor, 2012; Ben Taieb *et al.*, 2021; Zambon *et al.*, 2024; Corani *et al.*, 2023). In addition, the exploration of machine learning techniques holds significant promise for future advancements in this field. Approaches proposed by Olivares *et al.* (2023b, 2022, 2023a) present exciting opportunities to introduce machine learning algorithms in cross-temporal probabilistic forecast reconciliation.

A reconciliation approach for the realized volatility

In this chapter, we address whether using the disaggregate components of the daily realized volatility or combining them with the daily realized volatility forecasts improves the forecasting accuracy compared to using the daily realized volatility series alone. To this end, we investigate alternative ways of leveraging intraday RV decompositions in forecasting daily RV. The key question of this study is whether it is beneficial to model and forecast daily RV at the sub-component level, thus exploiting the informative content (from a forecasting point of view) of bottom time series, or whether a direct

strategy should be preferred. The latter refers to the prediction of RV by directly modeling it, even when the explanatory variables include a decomposition of the RV itself. Differently, indirect (bottom-up) forecasting is based on the aggregation of models fit on the bottom series, whose forecasts are then aggregated to recover the prediction of the top series. Forecast reconciliation adds a further element, by restoring the aggregation constraints linking the bottom, possibly the intermediate, and the top-level time series. The idea is that an appropriate 'imposition' to the forecasts of the same constraints valid for the observed data should improve the overall forecasting accuracy. We obtain a forecast of different sub-components individually, and then combine them to estimate the forecast of the aggregated series in an indirect (bottom-up) and a reconciliation forecasting framework.

Our main results can be summarized as follows. Through a simple out-of-sample forecasting experiment, we show that both bottom-up and regression-based reconciliation procedures (Wickramasuriya *et al.*, 2019) perform relatively well compared to the benchmark direct forecasting models by Corsi (2009), Patton and Sheppard (2015), and Bollerslev *et al.* (2022), mostly when the $PV(3)$ model by Bollerslev *et al.* (2022) is used to produce base forecasts of the daily RV top-level series. We find substantial and significant reductions in forecast errors when using the new proposed indirect/reconciliation approaches. The disaggregated procedures do quite well in forecasting RV, and the reconciliation approach is generally more promising. For the HAR model, the differences between the bottom-up and the direct approach are clearly visible, and the regression-based reconciliation approach offers some additional improvements. This is somehow confirmed for both SV and $PV(3)$ forecasts, where generally the combination scheme of single component models provides smaller forecast errors than both the indirect and direct forecasting approaches.

Appendix A

Cross-temporal reconciliation of solar forecasts

A.1 Proof of the Theorem 2.1

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

$$\boldsymbol{M}^{[k]}_{cs,j} = \boldsymbol{M}_{cs} = \left[\boldsymbol{I}_n - \boldsymbol{W}\boldsymbol{U}\left(\boldsymbol{U}'\boldsymbol{W}\boldsymbol{U}\right)^{-1}\boldsymbol{U}'\right], \quad k\in\mathcal{K}, \quad j=1,...\frac{m}{k},$$

and

$$\boldsymbol{M}_{te,i} = \boldsymbol{M}_{te} = \left[\boldsymbol{I}_{k^*+m} - \boldsymbol{\Omega}\boldsymbol{Z}\left(\boldsymbol{Z}'\boldsymbol{\Omega}\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}'\right], \quad i=1,\dots,n,$$

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

$$\begin{aligned}\widetilde{\boldsymbol{Y}}^{[k]}_{cs} &= \boldsymbol{M}_{cs}\widehat{\boldsymbol{Y}}^{[k]}, \quad k\in\mathcal{K},\\ \widetilde{\boldsymbol{y}}_{te,i} &= \boldsymbol{M}_{te}\widehat{\boldsymbol{y}}_i, \qquad i=1,\dots,n,\end{aligned}$$

where $\widehat{\boldsymbol{Y}}^{[k]}$ and $\widetilde{\boldsymbol{Y}}^{[k]}_{cs}$ are the $\left(n\times\frac{m}{k}\right)$ matrices with, respectively, the base and the cross-sectional reconciled forecasts of the n series at time granularity k, and $\widehat{\boldsymbol{y}}_i$ and $\widetilde{\boldsymbol{y}}_{te,i}$ 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:

$$\begin{aligned}\widetilde{\boldsymbol{Y}}_{cs} &= \left[\widetilde{\boldsymbol{Y}}^{[m]}_{cs}\ \widetilde{\boldsymbol{Y}}^{[k_{p-1}]}_{cs}\dots\widetilde{\boldsymbol{Y}}^{[k_2]}_{cs}\ \widetilde{\boldsymbol{Y}}^{[1]}_{cs}\right]\\ &= \boldsymbol{M}_{cs}\left[\widehat{\boldsymbol{Y}}^{[m]}\ \widehat{\boldsymbol{Y}}^{[k_{p-1}]}\dots\widehat{\boldsymbol{Y}}^{[k_2]}\ \widehat{\boldsymbol{Y}}^{[1]}\right]\\ &= \boldsymbol{M}_{cs}\widehat{\boldsymbol{Y}},\end{aligned}$$

$$\begin{aligned}\widetilde{\boldsymbol{Y}}_{te} &= \left\{\boldsymbol{M}_{te}\left[\widehat{\boldsymbol{y}}_1\dots\widehat{\boldsymbol{y}}_i\dots\widehat{\boldsymbol{y}}_n\right]\right\}'\\ &= \left\{\boldsymbol{M}_{te}\widehat{\boldsymbol{Y}}'\right\}' = \widehat{\boldsymbol{Y}}\boldsymbol{M}'_{te}.\end{aligned}$$

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

$$\begin{aligned}
&\text{step 1} \qquad \widetilde{\boldsymbol{Y}}_{cs} = \boldsymbol{M}_{cs}\widehat{\boldsymbol{Y}} \\
&\text{step 2} \qquad \widetilde{\boldsymbol{Y}}_{cst} = \widetilde{\boldsymbol{Y}}_{cs}\boldsymbol{M}_{te}' = \boldsymbol{M}_{cs}\widehat{\boldsymbol{Y}}\boldsymbol{M}_{te}'
\end{aligned}$$

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

$$\begin{aligned}
&\text{step 1} \qquad \widetilde{\boldsymbol{Y}}_{te} = \widehat{\boldsymbol{Y}}\boldsymbol{M}_{te}' \\
&\text{step 2} \qquad \widetilde{\boldsymbol{Y}}_{tcs} = \boldsymbol{M}_{cs}\widetilde{\boldsymbol{Y}}_{te} = \boldsymbol{M}_{cs}\widehat{\boldsymbol{Y}}\boldsymbol{M}_{te}' = \widetilde{\boldsymbol{Y}}_{cst}.
\end{aligned}$$

The forecasts $\widetilde{\boldsymbol{Y}}_{tcs}$ (and $\widetilde{\boldsymbol{Y}}_{cst}$) resulting from a single two-step iteration, are cross-temporally reconciled because:

$$\begin{aligned}
\boldsymbol{U}'\widetilde{\boldsymbol{Y}}_{tcs} &= \boldsymbol{U}'\boldsymbol{M}_{cs}\widehat{\boldsymbol{Y}}\boldsymbol{M}_{te}' \\
&= \left[\boldsymbol{U}' - \boldsymbol{U}'\boldsymbol{W}\boldsymbol{U}\left(\boldsymbol{U}'\boldsymbol{W}\boldsymbol{U}\right)^{-1}\boldsymbol{U}'\right]\widehat{\boldsymbol{Y}}\boldsymbol{M}_{te}' \\
&= \left[\boldsymbol{U}' - \boldsymbol{U}'\right]\widehat{\boldsymbol{Y}}\boldsymbol{M}_{te}' \\
&= \boldsymbol{0}_{n_a\times(k^*+m)}
\end{aligned}$$

and

$$\begin{aligned}
\boldsymbol{Z}'\widetilde{\boldsymbol{Y}}_{tcs}' &= \boldsymbol{Z}'\boldsymbol{M}_{te}\widehat{\boldsymbol{Y}}'\widehat{\boldsymbol{Y}}'\boldsymbol{M}_{cs}' \\
&= \left[\boldsymbol{Z}' - \boldsymbol{Z}'\boldsymbol{\Omega}\boldsymbol{Z}\left(\boldsymbol{Z}'\boldsymbol{\Omega}\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}'\right]\boldsymbol{M}_{cs}' \\
&= \left[\boldsymbol{Z}' - \boldsymbol{Z}'\right]\widehat{\boldsymbol{Y}}'\boldsymbol{M}_{cs}' \\
&= \boldsymbol{0}_{m\times n}.
\end{aligned}$$

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

$$\widetilde{\boldsymbol{y}}_{oct} = \boldsymbol{F}\left(\boldsymbol{F}'\boldsymbol{\Omega}_{ct}^{-1}\boldsymbol{F}\right)^{-1}\boldsymbol{F}'\boldsymbol{\Omega}_{ct}^{-1}\widehat{\boldsymbol{y}},$$

where $\widetilde{\boldsymbol{y}}_{oct} = \text{vec}\left(\widetilde{\boldsymbol{Y}}_{oct}'\right)$, $\widehat{\boldsymbol{y}} = \text{vec}\left(\widehat{\boldsymbol{Y}}'\right)$, and $\boldsymbol{F} = \boldsymbol{S}\otimes\boldsymbol{R}$ is the $[n(k^*+m)\times mn_b]$ cross-temporal summing matrix. Let

$$\boldsymbol{\Omega}_{ct} = \boldsymbol{W}\otimes\boldsymbol{\Omega}$$

be the cross-temporal error covariance matrix, by exploiting the properties of the Kronecker product (Henderson and Searle, 1981; Harville, 2008), we obtain:

$$\begin{aligned}\widetilde{\boldsymbol{y}}_{oct} &= \boldsymbol{F}\left(\boldsymbol{F}'\boldsymbol{\Omega}_{ct}^{-1}\boldsymbol{F}\right)^{-1}\boldsymbol{F}'\boldsymbol{\Omega}_{ct}^{-1}\widehat{\boldsymbol{y}}\\ &= (\boldsymbol{S}\otimes\boldsymbol{R})\left[(\boldsymbol{S}\otimes\boldsymbol{R})'(\boldsymbol{W}\otimes\boldsymbol{\Omega})^{-1}(\boldsymbol{S}\otimes\boldsymbol{R})\right]^{-1}(\boldsymbol{S}\otimes\boldsymbol{R})'(\boldsymbol{W}\otimes\boldsymbol{\Omega})^{-1}\widehat{\boldsymbol{y}}\\ &= (\boldsymbol{S}\otimes\boldsymbol{R})\left[(\boldsymbol{S}'\otimes\boldsymbol{R}')(\boldsymbol{W}^{-1}\otimes\boldsymbol{\Omega}^{-1})(\boldsymbol{S}\otimes\boldsymbol{R})\right]^{-1}(\boldsymbol{S}'\otimes\boldsymbol{R}')(\boldsymbol{W}^{-1}\otimes\boldsymbol{\Omega}^{-1})\widehat{\boldsymbol{y}}\\ &= (\boldsymbol{S}\otimes\boldsymbol{R})\left(\boldsymbol{S}'\boldsymbol{W}^{-1}\boldsymbol{S}\otimes\boldsymbol{R}'\boldsymbol{\Omega}^{-1}\boldsymbol{R}\right)^{-1}(\boldsymbol{S}'\boldsymbol{W}^{-1}\otimes\boldsymbol{R}'\boldsymbol{\Omega}^{-1})\widehat{\boldsymbol{y}}\\ &= \left[\boldsymbol{S}\left(\boldsymbol{S}'\boldsymbol{W}^{-1}\boldsymbol{S}\right)^{-1}\boldsymbol{S}'\boldsymbol{W}^{-1}\otimes\boldsymbol{R}\left(\boldsymbol{R}'\boldsymbol{\Omega}^{-1}\boldsymbol{R}\right)^{-1}\boldsymbol{R}'\boldsymbol{\Omega}^{-1}\right]\widehat{\boldsymbol{y}}.\end{aligned}$$

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

$$\begin{aligned}\widetilde{\boldsymbol{y}}_{tcs} &= \boldsymbol{P}\left[\boldsymbol{I}_{k^*+m}\otimes\boldsymbol{S}\left(\boldsymbol{S}'\boldsymbol{W}^{-1}\boldsymbol{S}\right)^{-1}\boldsymbol{S}'\boldsymbol{W}^{-1}\right]\boldsymbol{P}'\left[\boldsymbol{I}_n\otimes\boldsymbol{R}\left(\boldsymbol{R}'\boldsymbol{\Omega}^{-1}\boldsymbol{R}\right)^{-1}\boldsymbol{R}'\boldsymbol{\Omega}^{-1}\right]\widehat{\boldsymbol{y}}\\ &= \left[\boldsymbol{S}\left(\boldsymbol{S}'\boldsymbol{W}^{-1}\boldsymbol{S}\right)^{-1}\boldsymbol{S}'\boldsymbol{W}^{-1}\otimes\boldsymbol{I}_{k^*+m}\right]\left[\boldsymbol{I}_n\otimes\boldsymbol{R}\left(\boldsymbol{R}'\boldsymbol{\Omega}^{-1}\boldsymbol{R}\right)^{-1}\boldsymbol{R}'\boldsymbol{\Omega}^{-1}\right]\widehat{\boldsymbol{y}}\\ &= \left[\boldsymbol{S}\left(\boldsymbol{S}'\boldsymbol{W}^{-1}\boldsymbol{S}\right)^{-1}\boldsymbol{S}'\boldsymbol{W}^{-1}\otimes\boldsymbol{R}\left(\boldsymbol{R}'\boldsymbol{\Omega}^{-1}\boldsymbol{R}\right)^{-1}\boldsymbol{R}'\boldsymbol{\Omega}^{-1}\right]\widehat{\boldsymbol{y}} = \widetilde{\boldsymbol{y}}_{oct}.\end{aligned}$$

Finally, $\widetilde{\boldsymbol{Y}}_{tcs} = \widetilde{\boldsymbol{Y}}_{oct}$ and, following from point 1,

$$\widetilde{\boldsymbol{Y}}_{tcs} = \widetilde{\boldsymbol{Y}}_{cst} = \widetilde{\boldsymbol{Y}}_{oct}.$$

□

KA with constant cross-sectional and temporal covariance matrices

In the Kourentzes and Athanasopoulos (2019) approach (Section 2.2.4), the temporally reconciled predictions at step 1, $\widehat{\boldsymbol{Y}}\boldsymbol{M}'_{te}$, are cross-sectionally reconciled via premultiplication by the matrix $\overline{\boldsymbol{M}} = \frac{1}{p}\sum_{k\in\mathcal{K}}\boldsymbol{M}_{cs}^{[k]}$. As $\boldsymbol{W}^{[k]} = \boldsymbol{W}$, $k\in\mathcal{K}$, $\boldsymbol{M}_{cs}^{[k]} = \boldsymbol{M}_{cs}$, and $\overline{\boldsymbol{M}} = \boldsymbol{M}_{cs}$. Then,

$$\widetilde{\boldsymbol{Y}}_{KA} = \overline{\boldsymbol{M}}\widehat{\boldsymbol{Y}}\boldsymbol{M}'_{te} = \boldsymbol{M}_{cs}\widehat{\boldsymbol{Y}}\boldsymbol{M}'_{te} = \widetilde{\boldsymbol{Y}}_{cst} = \widetilde{\boldsymbol{Y}}_{tcs}.$$

A.2 Proof of the Theorem 2.2

Let $\widetilde{\boldsymbol{Y}}_{tcs}$ and $\widetilde{\boldsymbol{Y}}_{cst}$ be the $[n\times(k^*+m)]$ matrix of the *temporal-then-cross-sectional* and *cross-sectional-then-temporal* iterative reconciled forecasts using $\boldsymbol{W}^{[k]}$ ($k\in\mathcal{K}$) and $\boldsymbol{\Omega}_i$ ($i=1,\ldots,n$) as the cross-sectional and temporal covariance matrices, respectively. The iterative solution obtained at a given complete recursion $J\geq 1$, may be written as the result of alternating oblique projections, obtained through generalized least squares, such that

$$\widetilde{\boldsymbol{y}}_{tcs} = (\boldsymbol{M}_{cs}\boldsymbol{M}_{te})^J\widehat{\boldsymbol{y}},$$

with

$$\begin{aligned} \boldsymbol{M}_{cs} &= \boldsymbol{K}_{cs}\left[\boldsymbol{K}_{cs}'(\boldsymbol{P}\boldsymbol{W}\boldsymbol{P}')^{-1}\boldsymbol{K}_{cs}\right]^{-1}\boldsymbol{K}_{cs}'(\boldsymbol{P}\boldsymbol{W}\boldsymbol{P}')^{-1} \\ \boldsymbol{M}_{te} &= \boldsymbol{K}_{te}\left[\boldsymbol{K}_{te}'\boldsymbol{\Omega}^{-1}\boldsymbol{K}_{te}\right]^{-1}\boldsymbol{K}_{te}'\boldsymbol{\Omega}^{-1} \end{aligned}, \tag{A.1}$$

where $\widetilde{\boldsymbol{y}}_{tcs} = \text{vec}\left(\widetilde{\boldsymbol{Y}}_{tcs}'\right)$, $\widehat{\boldsymbol{Y}}$ is the $[n \times (k^* + m)]$ base forecast matrix, $\widehat{\boldsymbol{y}} = \text{vec}\left(\widehat{\boldsymbol{Y}}'\right)$, $\boldsymbol{P}$ is the commutation matrix operating on matrices of dimension $[n \times (k^* + m)]$ (e.g., $\boldsymbol{Y}, \widehat{\boldsymbol{Y}}, \widetilde{\boldsymbol{Y}}$), such that $\boldsymbol{P}\text{vec}\left(\widehat{\boldsymbol{Y}}\right) = \text{vec}\left(\widehat{\boldsymbol{Y}}'\right)$ (see Chapter 1), and $\boldsymbol{K}_{cs}$ and $\boldsymbol{K}_{te}$ are extensions of, respectively, the cross-sectional summing matrix $\boldsymbol{S}$ and the temporal summing matrix $\boldsymbol{R}$:

$$\boldsymbol{K}_{cs} = \boldsymbol{S} \otimes \boldsymbol{I}_{(k^*+m)} \quad \text{and} \quad \boldsymbol{K}_{te} = \boldsymbol{I}_n \otimes \boldsymbol{R}.$$

More precisely, the cross-sectional structural representation valid for all the k^*+m time periods of all the temporal aggregates of the n variables, can be written as

$$\text{vec}\left(\boldsymbol{Y}'\right) = \boldsymbol{K}_{cs}\text{vec}\left(\boldsymbol{B}'\right),$$

whereas the temporal structural representation valid for all the n variables, can be expressed as

$$\text{vec}\left(\boldsymbol{Y}'\right) = \boldsymbol{K}_{te}\text{vec}\left(\boldsymbol{Y}^{[1]\prime}\right).$$

In addition, the cross-temporal optimal (in least squares sense) combination solution, which satisfies both cross-sectional and temporal constraints, can be written as $\widetilde{\boldsymbol{y}}_{oct} = \boldsymbol{M}_{ct}\widehat{\boldsymbol{y}}$, with

$$\boldsymbol{M}_{ct} = \boldsymbol{F}\left[\boldsymbol{F}'\boldsymbol{\Omega}^{-1}\boldsymbol{F}\right]^{-1}\boldsymbol{F}'\boldsymbol{\Omega}^{-1}, \tag{A.2}$$

where

$$\boldsymbol{F} = \boldsymbol{S} \otimes \boldsymbol{R} = \boldsymbol{K}_{cs}\boldsymbol{K}_{te} \tag{A.3}$$

is the cross-temporal summing matrix, such that the cross-temporal structural representation can be expressed as

$$\text{vec}\left(\boldsymbol{Y}'\right) = \boldsymbol{F}\text{vec}\left(\boldsymbol{B}^{[1]\prime}\right).$$

In other terms, $\widetilde{\boldsymbol{y}}_{cs}$, $\widetilde{\boldsymbol{y}}_{te}$, and $\widetilde{\boldsymbol{y}}_{ct}$ are obtained through **oblique** (i.e., non-orthogonal) projections of the base forecasts $\widehat{\boldsymbol{y}}$ onto the linear sub-spaces spanned by the columns of, respectively, $\boldsymbol{S} \otimes \boldsymbol{I}_{(k^*+m)}$, $\boldsymbol{I}_n \otimes \boldsymbol{R}$, and $\boldsymbol{S} \otimes \boldsymbol{R}$. It is worth noting that, according to expression (A.3), the sub-space spanned by the columns of $\boldsymbol{S} \otimes \boldsymbol{R}$ is the intersection of the sub-spaces spanned by the columns of $\boldsymbol{S} \otimes \boldsymbol{I}_{(k^*+m)}$ and $\boldsymbol{I}_n \otimes \boldsymbol{R}$, respectively.
We want to prove

$$\|\widetilde{\boldsymbol{y}}_{tcs} - \widetilde{\boldsymbol{y}}_{oct}\| = \|(\boldsymbol{M}_{cs}\boldsymbol{M}_{te})^J\widehat{\boldsymbol{y}} - \boldsymbol{M}_{ct}\widehat{\boldsymbol{y}}\| \xrightarrow{J\to+\infty} 0.$$

Since $\boldsymbol{\Omega}$ is p.d., we can use the Cholesky decomposition (Harville, 2008, p. 232) to express its inverse as $\boldsymbol{\Omega}^{-1} = \boldsymbol{Q}'\boldsymbol{Q}$, where $\boldsymbol{Q}$ is a (unique) upper triangular matrix with positive diagonal elements. Thus, denoting $\overline{\boldsymbol{y}} = \boldsymbol{Q}\widehat{\boldsymbol{y}}$, it can be shown that

$$\widetilde{\boldsymbol{y}}_{cs} = \boldsymbol{Q}^{-1}\overline{\boldsymbol{M}}_{cs}\overline{\boldsymbol{y}}, \qquad \widetilde{\boldsymbol{y}}_{te} = \boldsymbol{Q}^{-1}\overline{\boldsymbol{M}}_{te}\overline{\boldsymbol{y}}, \qquad \widetilde{\boldsymbol{y}}_{oct} = \boldsymbol{Q}^{-1}\overline{\boldsymbol{M}}_{ct}\overline{\boldsymbol{y}},$$

where $\overline{\boldsymbol{M}}_{rec} = \overline{\boldsymbol{K}}_{rec}\left(\overline{\boldsymbol{K}}'_{rec}\overline{\boldsymbol{K}}_{rec}\right)^{-1}\overline{\boldsymbol{K}}'_{rec}$, with $\overline{\boldsymbol{K}}_{rec} = \boldsymbol{Q}\boldsymbol{K}_{rec}$, $rec \in \{cs,\ te\}$, and $\overline{\boldsymbol{M}}_{ct} = \overline{\boldsymbol{F}}\left(\overline{\boldsymbol{F}}'\overline{\boldsymbol{F}}\right)^{-1}\overline{\boldsymbol{F}}'$, with $\overline{\boldsymbol{F}} = \boldsymbol{Q}\boldsymbol{F}$, are **orthogonal** projection matrices onto the linear sub-spaces spanned by the columns of, respectively, $\boldsymbol{Q}\boldsymbol{F}$, $\boldsymbol{Q}\boldsymbol{K}_{cs}$, and $\boldsymbol{Q}\boldsymbol{K}_{te}$.

The oblique projection matrices (A.1) can thus be written using the orthogonal projection matrices $\overline{\boldsymbol{M}}_{rec}$ such that $\boldsymbol{M}_{rec} = \boldsymbol{Q}^{-1}\overline{\boldsymbol{M}}_{rec}$., $rec \in \{cs,\ te\}$. Then,

$$\begin{aligned}
(\boldsymbol{M}_{cs}\boldsymbol{M}_{te})^{J}\widehat{\boldsymbol{y}} &= \boldsymbol{M}_{cs}\boldsymbol{M}_{te}\dots\boldsymbol{M}_{cs}\boldsymbol{M}_{te}\widehat{\boldsymbol{y}} \\
&= \boldsymbol{Q}^{-1}\boldsymbol{Q}\boldsymbol{M}_{cs}\boldsymbol{M}_{te}\dots\boldsymbol{M}_{cs}\boldsymbol{M}_{te}\widehat{\boldsymbol{y}} \\
&= \boldsymbol{Q}^{-1}\boldsymbol{Q}\boldsymbol{K}_{cs}\left[\boldsymbol{K}'_{cs}\boldsymbol{\Omega}^{-1}\boldsymbol{K}_{cs}\right]^{-1}\boldsymbol{K}'_{cs}\boldsymbol{\Omega}^{-1}\boldsymbol{K}_{te}\left[\boldsymbol{K}'_{te}\boldsymbol{\Omega}^{-1}\boldsymbol{K}_{te}\right]^{-1}\boldsymbol{K}'_{te}\boldsymbol{\Omega}^{-1}\dots \\
&\quad\dots\boldsymbol{K}_{cs}\left[\boldsymbol{K}'_{cs}\boldsymbol{\Omega}^{-1}\boldsymbol{K}_{cs}\right]^{-1}\boldsymbol{K}'_{cs}\boldsymbol{\Omega}^{-1}\boldsymbol{K}_{te}\left[\boldsymbol{K}'_{te}\boldsymbol{\Omega}^{-1}\boldsymbol{K}_{te}\right]^{-1}\boldsymbol{K}'_{te}\boldsymbol{\Omega}^{-1}\widehat{\boldsymbol{y}} \\
&= \boldsymbol{Q}^{-1}\underbrace{\boldsymbol{Q}\boldsymbol{K}_{cs}\left[\boldsymbol{K}'_{cs}\boldsymbol{Q}\boldsymbol{Q}'\boldsymbol{K}_{cs}\right]^{-1}\boldsymbol{K}'_{cs}\boldsymbol{Q}}_{\overline{\boldsymbol{M}}_{cs}}\underbrace{\boldsymbol{Q}'\boldsymbol{K}_{te}\left[\boldsymbol{K}'_{te}\boldsymbol{Q}\boldsymbol{Q}'\boldsymbol{K}_{te}\right]^{-1}\boldsymbol{K}'_{te}\boldsymbol{Q}}_{\overline{\boldsymbol{M}}_{te}}\boldsymbol{Q}'\dots \\
&\quad\dots\underbrace{\boldsymbol{Q}\boldsymbol{K}_{cs}\left[\boldsymbol{K}'_{cs}\boldsymbol{Q}\boldsymbol{Q}'\boldsymbol{K}_{cs}\right]^{-1}\boldsymbol{K}'_{cs}\boldsymbol{Q}}_{\overline{\boldsymbol{M}}_{cs}}\underbrace{\boldsymbol{Q}'\boldsymbol{K}_{te}\left[\boldsymbol{K}'_{te}\boldsymbol{Q}\boldsymbol{Q}'\boldsymbol{K}_{te}\right]^{-1}\boldsymbol{K}'_{te}\boldsymbol{Q}}_{\overline{\boldsymbol{M}}_{te}}\boldsymbol{Q}'\widehat{\boldsymbol{y}} \\
&= \boldsymbol{Q}^{-1}(\overline{\boldsymbol{M}}_{cs}\overline{\boldsymbol{M}}_{te})^{J}\overline{\boldsymbol{y}}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\widetilde{\boldsymbol{y}}_{tcs} - \widetilde{\boldsymbol{y}}_{oct} &= \boldsymbol{Q}^{-1}(\overline{\boldsymbol{M}}_{cs}\overline{\boldsymbol{M}}_{te})^{J}\overline{\boldsymbol{y}} - \boldsymbol{Q}^{-1}\overline{\boldsymbol{M}}_{ct}\overline{\boldsymbol{y}} \\
&= \boldsymbol{Q}^{-1}\left[(\overline{\boldsymbol{M}}_{cs}\overline{\boldsymbol{M}}_{te})^{J} - \overline{\boldsymbol{M}}_{ct}\right]\boldsymbol{Q}\widehat{\boldsymbol{y}}.
\end{aligned}$$

According to Von Neumann (1949) and Ginat (2018), the orthogonal alternating projection onto two subspaces converges in norm to the projection onto the intersection of the two subspaces[1], that is

$$\|(\overline{\boldsymbol{M}}_{cs}\overline{\boldsymbol{M}}_{te})^{J} - \overline{\boldsymbol{M}}_{ct}\|\boldsymbol{Q}\widehat{\boldsymbol{y}} \xrightarrow{J\to+\infty} 0$$

and, finally,

$$\|\widetilde{\boldsymbol{y}}_{tcs} - \widetilde{\boldsymbol{y}}_{oct}\| \xrightarrow{J\to+\infty} 0.$$

The convergence of $\widetilde{\boldsymbol{y}}_{cst}$ to $\widetilde{\boldsymbol{y}}_{oct}$ is easily obtained by inverting the order in which matrices $\boldsymbol{M}_{cs}$ and $\boldsymbol{M}_{te}$ are applied in the previous proof. □

[1]Thanks to Prof. Tommaso Proietti for suggesting considering alternating projections for this proof.

A.3 nMBE and fully coherent forecasts

Let $\widetilde{y}_t^{[k]}$, $t = 1, \ldots, N_k$, $k \in \mathcal{K}$, be the cross-temporal reconciled forecasts for a single time series in a complete time cycle (see Section 2.2). The temporal coherency implies that

$$\widetilde{y}_t^{[k]} = \sum_{j=k(t-1)+1}^{kt} \widetilde{y}_j^{[1]}, \quad \forall k \in \mathcal{K},$$

where $\widetilde{y}_j^{[1]}$, $j = 1, \ldots, N_1$, denotes the forecasts for the series observed at the highest time frequency. As the realizations of the time series are by definition temporally coherent (i.e., $y_t^{[k]} = \sum_{j=k(t-1)+1}^{kt} y_j^{[1]}$), the normalized Mean Bias Error defined in (2.12) can be written as

$$\begin{aligned}
\text{nMBE}^{[k]} &= \frac{\frac{1}{N_k}\sum_{t=1}^{N_k}\left(\widetilde{y}_t^{[k]} - y_t^{[k]}\right)}{\frac{1}{N_k}\sum_{t=1}^{N} y_t^{[k]}} = \frac{\sum_{t=1}^{N_k}\left(\sum_{j=k(t-1)+1}^{kt}\widetilde{y}_j^{[1]} - \sum_{j=k(t-1)+1}^{kt} y_j^{[1]}\right)}{\sum_{t=1}^{N_k}\sum_{j=k(t-1)+1}^{kt} y_j^{[1]}} \\
&= \frac{\sum_{j=1}^{N_1}\left(\widetilde{y}_t^{[1]} - y_t^{[1]}\right)}{\sum_{j=1}^{N_1} y_j^{[1]}} = \text{nMBE}^{[1]}, \quad k \in \{m, k_{p-1}, \ldots, k_2\}.
\end{aligned}$$

Appendix B

Forecast combination-based forecast reconciliation

B.1 The formulation by Hollyman *et al.* (2021)

B.1.1 Level-1 Conditional Coherent forecast reconciliation

For $l = 1$, in order to transform the bts base forecasts $\widehat{\boldsymbol{b}}$ in reconciled forecasts $\widetilde{\boldsymbol{b}}^{(1)}$ conditional to $\widehat{y}_1$, Hollyman *et al.* (2021) consider the $[(n_b + 1) \times (n_b + 1)]$ matrix $\boldsymbol{A}_1$:

$$\boldsymbol{A}_1 = \begin{bmatrix} 0 & \mathbf{1}'_{n_b} \\ \boldsymbol{p} & \boldsymbol{I}_{n_b} \end{bmatrix},$$

where $\boldsymbol{p}$ is a $(n_b \times 1)$ vector of *combination weights* p_i, $i = 1, \ldots, n_b$, $0 < p_i < 1$, $\sum_{i=1}^{n_b} p_i = 1$. As it is immediately recognized, such a matrix may be obtained by 'augmenting' the level-1 structural summation matrix $\boldsymbol{S}_1$, putting the vector $[0\ \boldsymbol{p}']'$ on its left side. Let $\bar{\boldsymbol{G}}_1$ be a $[(n_b + 1) \times (n_b + 1)]$ matrix linked to $\boldsymbol{A}_1$ by the relationship

$$\bar{\boldsymbol{G}}_1 \boldsymbol{A}_1 = \boldsymbol{I}_{(n_b+1)}.$$

By solving the previous relationship wrt $\bar{\boldsymbol{G}}_1$, i.e. $\bar{\boldsymbol{G}}_1 = \boldsymbol{A}_1^{-1}$, it is possible to get the weights to be used to combine the base forecasts of the bts, and the n_b lower rows of matrix $\bar{\boldsymbol{G}}_1$ compose the $[n_b \times (n_b + 1)]$ matrix $\boldsymbol{G}_1$, transforming the bts base forecasts in coherent forecasts. It can be easily checked that

$$\bar{\boldsymbol{G}}_1 = \boldsymbol{A}_1^{-1} = \begin{bmatrix} -1 & \mathbf{1}'_{n_b} \\ \boldsymbol{p} & \left(\boldsymbol{I}_{n_b} - \boldsymbol{p}\mathbf{1}'_{n_b}\right) \end{bmatrix},$$

and thus we can write $\widetilde{\boldsymbol{b}}^{(1)} = \boldsymbol{G}_1 \widehat{\boldsymbol{y}}_1$, where $\boldsymbol{G}_1 = \begin{bmatrix} \boldsymbol{p} & \left(\boldsymbol{I}_{n_b} - \boldsymbol{p}\mathbf{1}'_{n_b}\right) \end{bmatrix}$ is a $[n_b \times (n_b + 1)]$ matrix obtained by removing the first row of $\bar{\boldsymbol{G}}_1$.

All the reconciled forecasts are thus given by $\widetilde{\boldsymbol{y}}^{(1)} = \boldsymbol{S}\widetilde{\boldsymbol{b}}^{(1)} = \boldsymbol{S}\boldsymbol{G}_1\widehat{\boldsymbol{y}}_1$. In order to express the vector $\widetilde{\boldsymbol{y}}^{(1)}$ as a transformation of all the base forecasts, not only the ones for the top-level series and the bts, $n_a - 1$ zero columns have to be inserted between the first and the second column of matrix $\boldsymbol{G}_1$, thus obtaining the $(n_b \times n)$ matrix $\boldsymbol{G}^{(1)}$, such that:

$$\widetilde{\boldsymbol{y}}^{(1)} = \boldsymbol{G}^{(1)}\widehat{\boldsymbol{y}}.$$

For the simple hierarchical time series considered as an example in Figure 3.1, we would have the (5×8) matrix

$$\boldsymbol{G}^{(1)} = [\boldsymbol{p} \;\; \boldsymbol{0}_{5\times 2} \;\; (\boldsymbol{I}_5 - \boldsymbol{p}\boldsymbol{1}_5')] \,.$$

It is worth noting that, since $\boldsymbol{G}^{(1)}\boldsymbol{S} = \boldsymbol{I}_5$, this matrix satisfies the unbiasedness condition for the reconciled forecasts (Athanasopoulos et al., 2009).

B.1.2 Extension for $l > 1$

Let's focus now on a generic level l, $l = 1, \dots L$, of the hierarchical/grouped time series, and denote with $\widetilde{\boldsymbol{y}}_l$ the $[(n_l + n_b) \times 1]$ vector containing the reconciled forecasts conditional to the level-l uts base forecasts. In other words, the uts reconciled forecasts in $\widetilde{\boldsymbol{y}}_l$ are equal to the base forecasts of the corresponding uts in $\widehat{\boldsymbol{y}}_l$ (i.e., $\widetilde{\boldsymbol{a}}_l = \widehat{\boldsymbol{a}}_l$). Then, denote $\widetilde{\boldsymbol{b}}^{(l)}$, $l = 1, \dots, L$, the bts reconciled forecasts conditional to the base forecasts of the level-l series. In analogy to the case $l = 1$, Hollyman *et al.* (2021) propose to compute

$$\widetilde{\boldsymbol{b}}^{(l)} = \boldsymbol{G}_l\widehat{\boldsymbol{y}}_l,$$

where $\boldsymbol{G}_l$ is a $[n_b \times (n_l + n_b)]$ matrix built as follows.

Let $\boldsymbol{P}_l$ be the $(n_b \times n_l)$ matrix (3.5) containing the weights of each forecasts in the n_l elementary hierarchies linking each of the n_l level-l series to their 'afferent' bts, and define the $[(n_l + n_b) \times (n_l + n_b)]$ matrix $\boldsymbol{A}_l$ as:

$$\boldsymbol{A}_l = \begin{bmatrix} \boldsymbol{0}_{n_l \times n_l} & \boldsymbol{C}_l \\ \boldsymbol{P}_l & \boldsymbol{I}_{n_b} \end{bmatrix},$$

where $\boldsymbol{C}_l$ is the cross-sectional (contemporaneous) aggregation matrix, with dimension $(n_l \times n_b)$, mapping the n_b bts into the n_l level-l aggregated series, such that $\boldsymbol{C}_l\boldsymbol{P}_l = \boldsymbol{I}_{n_l}$. $\boldsymbol{A}_l$ is a 2×2 block-partitioned matrix, with a null upper-left block, whose inverse is given by (Lu and Shiou, 2002):

$$\boldsymbol{A}_l^{-1} = \begin{bmatrix} -(\boldsymbol{C}_l\boldsymbol{P}_l)^{-1} & (\boldsymbol{C}_l\boldsymbol{P}_l)^{-1}\boldsymbol{C}_l \\ \boldsymbol{P}_l(\boldsymbol{C}_l\boldsymbol{P}_l)^{-1} & \left[\boldsymbol{I}_{n_b} - \boldsymbol{P}_l(\boldsymbol{C}_l\boldsymbol{P}_l)^{-1}\boldsymbol{C}_l\right] \end{bmatrix} = \begin{bmatrix} -\boldsymbol{I}_{n_l} & \boldsymbol{C}_l \\ \boldsymbol{P}_l & (\boldsymbol{I}_{n_b} - \boldsymbol{P}_l\boldsymbol{C}_l) \end{bmatrix},$$

a very simple expression to compute, which does not need any matrix inversion.

The desired transformation matrix $\boldsymbol{G}_l$, which has dimension $[n_b \times (n_b + n_l)]$, allowing to calculate bts reconciled forecasts in line with the n_l base forecasts of the level-l series, can be obtained by simply discarding the top n_l rows of matrix $\boldsymbol{A}_l^{-1}$: $\boldsymbol{G}_l = \begin{bmatrix} \boldsymbol{P}_l & (\boldsymbol{I}_{n_b} - \boldsymbol{P}_l\boldsymbol{C}_l) \end{bmatrix}$. The *level-l conditional bts reconciled forecasts* are thus given by:

$$\begin{aligned} \widetilde{\boldsymbol{b}}^{(l)} &= \boldsymbol{P}_l\widehat{\boldsymbol{a}}_l + (\boldsymbol{I}_{n_b} - \boldsymbol{P}_l\boldsymbol{C}_l)\,\widehat{\boldsymbol{b}} \\ &= \widehat{\boldsymbol{b}} + \boldsymbol{P}_l\left(\widehat{\boldsymbol{a}}_l - \boldsymbol{C}_l\widehat{\boldsymbol{b}}\right) \end{aligned}, \quad l = 1, \ldots, L,$$

and the vector $\widetilde{\boldsymbol{y}}^{(l)}$ of all the reconciled forecasts conditional to the level-l base forecasts is easily obtained as:

$$\widetilde{\boldsymbol{y}}^{(l)} = \boldsymbol{S}\widetilde{\boldsymbol{b}}^{(l)}, \quad l = 1, \ldots, L.$$

B.2 Derivation of $\widetilde{\boldsymbol{b}}^{(l)}_{\mathbf{LCC}}$

Consider the lagrangean function

$$L\left(\boldsymbol{b}, \boldsymbol{\lambda}\right) = \boldsymbol{b}'\boldsymbol{W}_b^{-1}\boldsymbol{b} - 2\widehat{\boldsymbol{b}}'\boldsymbol{W}_b^{-1}\boldsymbol{b} + \widehat{\boldsymbol{b}}'\boldsymbol{W}_b^{-1}\widehat{\boldsymbol{b}} + 2\boldsymbol{\lambda}'\left(\boldsymbol{C}_l\boldsymbol{b} - \boldsymbol{a}_l\right).$$

The first order conditions are given by:

$$\begin{aligned} \frac{\partial L}{\partial \boldsymbol{b}} &= \boldsymbol{W}_b^{-1}\boldsymbol{b} - 2\boldsymbol{W}_b^{-1}\widehat{\boldsymbol{b}} + 2\boldsymbol{C}_l'\boldsymbol{\lambda} = \boldsymbol{0}, \\ \frac{\partial L}{\partial \boldsymbol{\lambda}} &= \boldsymbol{C}_l\boldsymbol{b} - \boldsymbol{a}_l = \boldsymbol{0}. \end{aligned}$$

After simplification and re-arrangement of the known terms on the right side of the expression, it is:

$$\begin{aligned} \boldsymbol{W}_b^{-1}\boldsymbol{b} + \boldsymbol{C}_l'\boldsymbol{\lambda} &= \boldsymbol{W}_b^{-1}\widehat{\boldsymbol{b}} \\ \boldsymbol{C}_l\boldsymbol{b} &= \boldsymbol{a}_l \end{aligned} \quad \longrightarrow \quad \begin{bmatrix} \boldsymbol{W}_b^{-1} & \boldsymbol{C}_l' \\ \boldsymbol{C}_l & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{b} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \boldsymbol{W}_b^{-1}\widehat{\boldsymbol{b}} \\ \boldsymbol{a}_l \end{bmatrix}.$$

The solution to the system is thus given by

$$\begin{bmatrix} \boldsymbol{b} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \boldsymbol{W}_b^{-1} & \boldsymbol{C}_l' \\ \boldsymbol{C}_l & \boldsymbol{0} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{W}_b^{-1}\widehat{\boldsymbol{b}} \\ \boldsymbol{a}_l \end{bmatrix}.$$

The inverse matrix is given by (Lu and Shiou, 2002):

$$\begin{bmatrix} \boldsymbol{W}_b^{-1} & \boldsymbol{C}_l' \\ \boldsymbol{C}_l & \boldsymbol{0} \end{bmatrix}^{-1} = \begin{bmatrix} \left(\boldsymbol{W}_b - \boldsymbol{W}_b\boldsymbol{C}_l'\left(\boldsymbol{C}_l\boldsymbol{W}_b\boldsymbol{C}_l'\right)^{-1}\boldsymbol{C}_l\boldsymbol{W}_b\right) & \boldsymbol{W}_b\boldsymbol{C}_l'\left(\boldsymbol{C}_l\boldsymbol{W}_b\boldsymbol{C}_l'\right)^{-1} \\ \left(\boldsymbol{C}_l\boldsymbol{W}_b\boldsymbol{C}_l'\right)^{-1}\boldsymbol{C}_l\boldsymbol{W}_b & -\left(\boldsymbol{C}_l\boldsymbol{W}_b\boldsymbol{C}_l'\right)^{-1} \end{bmatrix},$$

and after some algebra we obtain that

$$\widetilde{\boldsymbol{b}}^{(l)}_{\mathrm{LCC}} = \widehat{\boldsymbol{b}} + \boldsymbol{W}_b\boldsymbol{C}_l'\left(\boldsymbol{C}_l\boldsymbol{W}_b\boldsymbol{C}_l'\right)^{-1}\left(\widehat{\boldsymbol{a}}_l - \boldsymbol{C}_l\widehat{\boldsymbol{b}}\right), \quad l = 1, \ldots, L.$$

The previous expression can be re-stated as follows:

$$\widetilde{\boldsymbol{b}}^{(l)}_{\mathrm{LCC}} = \boldsymbol{L}_l\widehat{\boldsymbol{a}}_l + \boldsymbol{M}_l\widehat{\boldsymbol{b}},$$

with

$$\boldsymbol{L}_l = \boldsymbol{W}_b \boldsymbol{C}_l' \left(\boldsymbol{C}_l \boldsymbol{W}_b \boldsymbol{C}_l'\right)^{-1}$$

e

$$\boldsymbol{M}_l = \left[\boldsymbol{I}_{n_b} - \boldsymbol{W}_b \boldsymbol{C}_l' \left(\boldsymbol{C}_l \boldsymbol{W}_b \boldsymbol{C}_l'\right)^{-1} \boldsymbol{C}_l\right] = \left(\boldsymbol{I}_{n_b} - \boldsymbol{L}_l \boldsymbol{C}_l\right).$$

It is thus:

$$\begin{bmatrix} \widehat{\boldsymbol{a}}_l \\ \widetilde{\boldsymbol{b}}^{(l)}_{\text{LCC}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}_{n_l} & \boldsymbol{0}_{n_l \times n_b} \\ \boldsymbol{L}_l & \boldsymbol{M}_l \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{a}}_l \\ \widehat{\boldsymbol{b}} \end{bmatrix} \quad \rightarrow \quad \widetilde{\boldsymbol{y}}_{l,\text{LCC}} = \overline{\boldsymbol{M}}_l \widehat{\boldsymbol{y}}_l, \quad \text{with} \quad \overline{\boldsymbol{M}}_l = \begin{bmatrix} \boldsymbol{I}_{n_l} & \boldsymbol{0}_{n_l \times n_b} \\ \boldsymbol{L}_l & \boldsymbol{M}_l \end{bmatrix}.$$

Since, as can be easily checked, it is $\boldsymbol{M}_l \boldsymbol{L}_l = \boldsymbol{0}_{n_b \times n_b}$, the matrix $\overline{\boldsymbol{M}}_l$ is idempotent ($\overline{\boldsymbol{M}}_l \overline{\boldsymbol{M}}_l = \overline{\boldsymbol{M}}_l$). For, it is a projection matrix in a linear sub-space of $\mathcal{R}^{n_l+n_b}$ spanned by the relationship

$$\begin{bmatrix} \boldsymbol{I}_{n_l} & \boldsymbol{0}_{n_l \times n_b} \\ \boldsymbol{I}_{n_l} & -\boldsymbol{C}_l \end{bmatrix} \begin{bmatrix} \boldsymbol{a}_l \\ \boldsymbol{b} \end{bmatrix} = \begin{bmatrix} \widehat{\boldsymbol{a}}_l \\ \boldsymbol{0}_{n_b} \end{bmatrix}.$$

B.3 L_lCC with endogenous constraints for the toy example of Figure 3.1

In this case the relationships linking the variable forming the hierarchy are:

$$\begin{aligned} T &= X + Y = A + B + C + D + E \\ X &= A + B \\ Y &= C + D + E \end{aligned}$$

Thus, there are two upper levels ($L = 2$) for which it is possible to apply the L_lCC reconciliation procedure with endogenous constraints:

- $\underline{l = 1 \ (n_1 = 1,\, n_b = 5)}$

$$\boldsymbol{W}_1 = \text{diag}\left(\sigma_T^2, \sigma_A^2, \sigma_B^2, \sigma_C^2, \sigma_D^2, \sigma_E^2\right), \quad \boldsymbol{U}_1' = \begin{bmatrix} 1 & -1 & -1 & -1 & -1 & -1 \end{bmatrix}$$

$$\Rightarrow \boldsymbol{W}_1 \boldsymbol{U}_1 = \begin{bmatrix} \sigma_T^2 \\ -\sigma_A^2 \\ -\sigma_B^2 \\ -\sigma_C^2 \\ -\sigma_D^2 \\ -\sigma_E^2 \end{bmatrix}, \quad \boldsymbol{U}_1' \boldsymbol{W}_1 \boldsymbol{U}_1 = \sigma_T^2 + \sigma_A^2 + \sigma_B^2 + \sigma_C^2 + \sigma_D^2 + \sigma_E^2.$$

After a bit of algebra, it is found that the reconciled forecasts are given by:

$$\widetilde{T}^{(1)}_{\text{en}} = \widehat{T} - \frac{\sigma_T^2}{\sigma_T^2 + \sigma_A^2 + \sigma_B^2 + \sigma_C^2 + \sigma_D^2 + \sigma_E^2} \left(\widehat{T} - \widehat{A} - \widehat{B} - \widehat{C} - \widehat{D} - \widehat{E}\right)$$

$$\widetilde{A}^{(1)}_{\text{en}} = \widehat{A} + \frac{\sigma^2_A}{\sigma^2_T+\sigma^2_A+\sigma^2_B+\sigma^2_C+\sigma^2_D+\sigma^2_E}\left(\widehat{T}-\widehat{A}-\widehat{B}-\widehat{C}-\widehat{D}-\widehat{E}\right)$$
$$\widetilde{B}^{(1)}_{\text{en}} = \widehat{B} + \frac{\sigma^2_B}{\sigma^2_T+\sigma^2_A+\sigma^2_B+\sigma^2_C+\sigma^2_D+\sigma^2_E}\left(\widehat{T}-\widehat{A}-\widehat{B}-\widehat{C}-\widehat{D}-\widehat{E}\right)$$
$$\widetilde{C}^{(1)}_{\text{en}} = \widehat{C} + \frac{\sigma^2_C}{\sigma^2_T+\sigma^2_A+\sigma^2_B+\sigma^2_C+\sigma^2_D+\sigma^2_E}\left(\widehat{T}-\widehat{A}-\widehat{B}-\widehat{C}-\widehat{D}-\widehat{E}\right)$$
$$\widetilde{D}^{(1)}_{\text{en}} = \widehat{D} + \frac{\sigma^2_D}{\sigma^2_T+\sigma^2_A+\sigma^2_B+\sigma^2_C+\sigma^2_D+\sigma^2_E}\left(\widehat{T}-\widehat{A}-\widehat{B}-\widehat{C}-\widehat{D}-\widehat{E}\right)$$
$$\widetilde{E}^{(1)}_{\text{en}} = \widehat{E} + \frac{\sigma^2_E}{\sigma^2_T+\sigma^2_A+\sigma^2_B+\sigma^2_C+\sigma^2_D+\sigma^2_E}\left(\widehat{T}-\widehat{A}-\widehat{B}-\widehat{C}-\widehat{D}-\widehat{E}\right).$$

Expressing each reconciled forecasts as a combination of the 'direct' (base) forecast, and of the 'implicit' forecast, obtained by applying the constraints to the base forecasts, gives:

$$\widetilde{T}^{(1)}_{\text{en}} = \left(\frac{\sigma^2_A+\sigma^2_B+\sigma^2_C+\sigma^2_D+\sigma^2_E}{\sigma^2_T+\sigma^2_A+\sigma^2_B+\sigma^2_C+\sigma^2_D+\sigma^2_E}\right)\widehat{T}+ \left(\frac{\sigma^2_T}{\sigma^2_T+\sigma^2_A+\sigma^2_B+\sigma^2_C+\sigma^2_D+\sigma^2_E}\right)\left(\widehat{A}+\widehat{B}+\widehat{C}+\widehat{D}+\widehat{E}\right)$$
$$\widetilde{A}^{(1)}_{\text{en}} = \left(\frac{\sigma^2_T+\sigma^2_B+\sigma^2_C+\sigma^2_D+\sigma^2_E}{\sigma^2_T+\sigma^2_A+\sigma^2_B+\sigma^2_C+\sigma^2_D+\sigma^2_E}\right)\widehat{A}+ \left(\frac{\sigma^2_A}{\sigma^2_T+\sigma^2_A+\sigma^2_B+\sigma^2_C+\sigma^2_D+\sigma^2_E}\right)\left(\widehat{T}-\widehat{B}-\widehat{C}-\widehat{D}-\widehat{E}\right)$$
$$\widetilde{B}^{(1)}_{\text{en}} = \left(\frac{\sigma^2_T+\sigma^2_A+\sigma^2_C+\sigma^2_D+\sigma^2_E}{\sigma^2_T+\sigma^2_A+\sigma^2_B+\sigma^2_C+\sigma^2_D+\sigma^2_E}\right)\widehat{B}+ \left(\frac{\sigma^2_B}{\sigma^2_T+\sigma^2_A+\sigma^2_B+\sigma^2_C+\sigma^2_D+\sigma^2_E}\right)\left(\widehat{T}-\widehat{A}-\widehat{C}-\widehat{D}-\widehat{E}\right)$$
$$\widetilde{C}^{(1)}_{\text{en}} = \left(\frac{\sigma^2_T+\sigma^2_A+\sigma^2_B+\sigma^2_D+\sigma^2_E}{\sigma^2_T+\sigma^2_A+\sigma^2_B+\sigma^2_C+\sigma^2_D+\sigma^2_E}\right)\widehat{C}+ \left(\frac{\sigma^2_C}{\sigma^2_T+\sigma^2_A+\sigma^2_B+\sigma^2_C+\sigma^2_D+\sigma^2_E}\right)\left(\widehat{T}-\widehat{A}-\widehat{B}-\widehat{D}-\widehat{E}\right)$$
$$\widetilde{D}^{(1)}_{\text{en}} = \left(\frac{\sigma^2_T+\sigma^2_A+\sigma^2_B+\sigma^2_C+\sigma^2_E}{\sigma^2_T+\sigma^2_A+\sigma^2_B+\sigma^2_C+\sigma^2_D+\sigma^2_E}\right)\widehat{D}+ \left(\frac{\sigma^2_D}{\sigma^2_T+\sigma^2_A+\sigma^2_B+\sigma^2_C+\sigma^2_D+\sigma^2_E}\right)\left(\widehat{T}-\widehat{A}-\widehat{B}-\widehat{C}-\widehat{E}\right)$$
$$\widetilde{E}^{(1)}_{\text{en}} = \left(\frac{\sigma^2_T+\sigma^2_A+\sigma^2_B+\sigma^2_C+\sigma^2_D}{\sigma^2_T+\sigma^2_A+\sigma^2_B+\sigma^2_C+\sigma^2_D+\sigma^2_E}\right)\widehat{E}+ \left(\frac{\sigma^2_E}{\sigma^2_T+\sigma^2_A+\sigma^2_B+\sigma^2_C+\sigma^2_D+\sigma^2_E}\right)\left(\widehat{T}-\widehat{A}-\widehat{B}-\widehat{C}-\widehat{D}\right)$$

- $\underline{l = 2\ (n_2 = 2,\ n_b = 5)}$

$$\boldsymbol{W}_2 = \text{diag}\left(\sigma^2_X, \sigma^2_Y, \sigma^2_A, \sigma^2_B, \sigma^2_C, \sigma^2_D, \sigma^2_E\right), \quad \boldsymbol{U}_2' = \begin{bmatrix} 1 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & -1 & -1 \end{bmatrix}$$

$$\Rightarrow \boldsymbol{W}_2\boldsymbol{U}_2 = \begin{bmatrix} \sigma_X^2 & 0 \\ 0 & \sigma_Y^2 \\ -\sigma_A^2 & 0 \\ -\sigma_B^2 & 0 \\ 0 & -\sigma_C^2 \\ 0 & -\sigma_D^2 \\ 0 & -\sigma_E^2 \end{bmatrix}, \quad \boldsymbol{U}_2'\boldsymbol{W}_2\boldsymbol{U}_2 = \begin{bmatrix} \sigma_X^2 + \sigma_A^2 + \sigma_B^2 & 0 \\ 0 & \sigma_Y^2 + \sigma_C^2 + \sigma_D^2 + \sigma_E^2 \end{bmatrix}.$$

After a bit of algebra, we find that the reconciled forecasts are equal to

$$\begin{aligned}
\widetilde{X}_{\text{en}}^{(2)} &= \widehat{X} - \frac{\sigma_X^2}{\sigma_X^2 + \sigma_A^2 + \sigma_B^2}\left(\widehat{X} - \widehat{A} - \widehat{B}\right) \\
\widetilde{Y}_{\text{en}}^{(2)} &= \widehat{Y} - \frac{\sigma_Y^2}{\sigma_Y^2 + \sigma_C^2 + \sigma_D^2 + \sigma_E^2}\left(\widehat{Y} - \widehat{C} - \widehat{D} - \widehat{E}\right) \\
\widetilde{A}_{\text{en}}^{(2)} &= \widehat{A} + \frac{\sigma_X^2}{\sigma_X^2 + \sigma_A^2 + \sigma_B^2}\left(\widehat{X} - \widehat{B}\right) \\
\widetilde{B}_{\text{en}}^{(2)} &= \widehat{B} + \frac{\sigma_X^2}{\sigma_X^2 + \sigma_A^2 + \sigma_B^2}\left(\widehat{X} - \widehat{A}\right) \\
\widetilde{C}_{\text{en}}^{(2)} &= \widehat{C} + \frac{\sigma_Y^2}{\sigma_Y^2 + \sigma_C^2 + \sigma_D^2 + \sigma_E^2}\left(\widehat{Y} - \widehat{D} - \widehat{E}\right) \\
\widetilde{D}_{\text{en}}^{(2)} &= \widehat{D} + \frac{\sigma_Y^2}{\sigma_Y^2 + \sigma_C^2 + \sigma_D^2 + \sigma_E^2}\left(\widehat{Y} - \widehat{C} - \widehat{E}\right) \\
\widetilde{E}_{\text{en}}^{(2)} &= \widehat{E} + \frac{\sigma_Y^2}{\sigma_Y^2 + \sigma_C^2 + \sigma_D^2 + \sigma_E^2}\left(\widehat{Y} - \widehat{C} - \widehat{D}\right)
\end{aligned}$$

or equivalently:

$$\begin{aligned}
\widetilde{X}_{\text{en}}^{(2)} &= \left(\frac{\sigma_A^2 + \sigma_B^2}{\sigma_X^2 + \sigma_A^2 + \sigma_B^2}\right)\widehat{X} + \left(\frac{\sigma_X^2}{\sigma_X^2 + \sigma_A^2 + \sigma_B^2}\right)\left(\widehat{A} + \widehat{B}\right) \\
\widetilde{Y}_{\text{en}}^{(2)} &= \left(\frac{\sigma_C^2 + \sigma_D^2 + \sigma_E^2}{\sigma_Y^2 + \sigma_C^2 + \sigma_D^2 + \sigma_E^2}\right)\widehat{Y} + \left(\frac{\sigma_Y^2}{\sigma_Y^2 + \sigma_C^2 + \sigma_D^2 + \sigma_E^2}\right)\left(\widehat{C} + \widehat{D} + \widehat{E}\right) \\
\widetilde{A}_{\text{en}}^{(2)} &= \left(\frac{\sigma_X^2 + \sigma_B^2}{\sigma_X^2 + \sigma_A^2 + \sigma_B^2}\right)\widehat{A} + \left(\frac{\sigma_A^2}{\sigma_X^2 + \sigma_A^2 + \sigma_B^2}\right)\left(\widehat{X} - \widehat{B}\right) \\
\widetilde{B}_{\text{en}}^{(2)} &= \left(\frac{\sigma_X^2 + \sigma_A^2}{\sigma_X^2 + \sigma_A^2 + \sigma_B^2}\right)\widehat{B} + \left(\frac{\sigma_B^2}{\sigma_X^2 + \sigma_A^2 + \sigma_B^2}\right)\left(\widehat{X} - \widehat{A}\right) \\
\widetilde{C}_{\text{en}}^{(2)} &= \left(\frac{\sigma_Y^2 + \sigma_D^2 + \sigma_E^2}{\sigma_Y^2 + \sigma_C^2 + \sigma_D^2 + \sigma_E^2}\right)\widehat{C} + \left(\frac{\sigma_C^2}{\sigma_Y^2 + \sigma_C^2 + \sigma_D^2 + \sigma_E^2}\right)\left(\widehat{Y} - \widehat{D} - \widehat{E}\right) \\
\widetilde{D}_{\text{en}}^{(2)} &= \left(\frac{\sigma_Y^2 + \sigma_C^2 + \sigma_E^2}{\sigma_Y^2 + \sigma_C^2 + \sigma_D^2 + \sigma_E^2}\right)\widehat{D} + \left(\frac{\sigma_D^2}{\sigma_Y^2 + \sigma_C^2 + \sigma_D^2 + \sigma_E^2}\right)\left(\widehat{Y} - \widehat{C} - \widehat{E}\right) \\
\widetilde{E}_{\text{en}}^{(2)} &= \left(\frac{\sigma_Y^2 + \sigma_C^2 + \sigma_D^2}{\sigma_Y^2 + \sigma_C^2 + \sigma_D^2 + \sigma_E^2}\right)\widehat{E} + \left(\frac{\sigma_E^2}{\sigma_Y^2 + \sigma_C^2 + \sigma_D^2 + \sigma_E^2}\right)\left(\widehat{Y} - \widehat{C} - \widehat{D}\right).
\end{aligned}$$

Appendix C

General linearly constrained multiple time series reconciliation

C.1 Derivation of equation (4.11)

Denote $\boldsymbol{W}_h = k_h \boldsymbol{W}$, where k_h is a proportionality constant and $\boldsymbol{W}$ is the p.d. covariance matrix used in the point forecast reconciliation formula (4.4). Then:

$$\begin{aligned}
\widetilde{\boldsymbol{W}}_h &= \boldsymbol{S}\boldsymbol{G}\boldsymbol{W}_h\boldsymbol{G}'\boldsymbol{S}' \\
&= k_h \boldsymbol{S}\boldsymbol{G}\boldsymbol{W}\boldsymbol{G}'\boldsymbol{S}' \\
&= k_h \boldsymbol{S}\left(\boldsymbol{S}'\boldsymbol{W}^{-1}\boldsymbol{S}\right)^{-1} \underbrace{\boldsymbol{S}'\boldsymbol{W}^{-1}\boldsymbol{W}\boldsymbol{W}^{-1}\boldsymbol{S}\left(\boldsymbol{S}'\boldsymbol{W}^{-1}\boldsymbol{S}\right)^{-1}}_{\boldsymbol{I}_n} \boldsymbol{S}' \\
&= k_h \boldsymbol{S}\left(\boldsymbol{S}'\boldsymbol{W}^{-1}\boldsymbol{S}\right)^{-1}\boldsymbol{S}' \\
&= k_h \boldsymbol{S}\underbrace{\left(\boldsymbol{S}'\boldsymbol{W}^{-1}\boldsymbol{S}\right)^{-1}\boldsymbol{S}'\boldsymbol{W}^{-1}}_{\boldsymbol{G}}\boldsymbol{W} = k_h \boldsymbol{S}\boldsymbol{G}\boldsymbol{W}.
\end{aligned}$$

Bibliography

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*. (visited on September 8, 2022) **URL:** https://digitalscholarship.unlv.edu/renew_pubs/4

Abolghasemi, M., Hyndman, R. J., Spiliotis, E. and Bergmeir, C. (2022) Model selection in reconciling hierarchical time series. *Machine Learning* **111**(2), 739–789. **DOI:** 10.1007/s10994-021-06126-z

Abouarghoub, W., Nomikos, N. K. and Petropoulos, F. (2018) On reconciling macro and micro energy transport forecasts for strategic decision making in the tanker industry. *Transportation Research Part E: Logistics and Transportation Review* **113**, 225–238. **DOI:** 10.1016/j.tre.2017.10.012

Abramson, B. and Clemen, R. (1995) Probability forecasting. *International Journal of Forecasting* **11**(1), 1–4. **DOI:** 10.1016/0169-2070(94)02000-F

Ahmed, R. A. (2009) *Forecasting Hierarchical Time Series*. Ph.D. thesis, Monash University

Aiolfi, M. and Timmermann, A. (2006) Persistence in forecasting performance and conditional combination strategies. *Journal of Econometrics* **135**(1-2), 31–53. **DOI:** 10.1016/j.jeconom.2005.07.015

Aït-Sahalia, Y., Cacho-Diaz, J. and Laeven, R. J. (2015) Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics* **117**(3), 585–606. **DOI:** 10.1016/j.jfineco.2015.03.002

Aït-Sahalia, Y. and Jacod, J. (2014) *High-frequency financial econometrics*. Princeton: Princeton University Press. ISBN 9780691161433

Andersen, T., Bollerslev, T., Diebold, F. and Ebens, H. (2001a) The distribution of realized stock return volatility. *Journal of Financial Economics* **61**(1), 43–76. **DOI:** 10.1016/S0304-405X(01)00055-1

Andersen, T. G., Bollerslev, T. and Diebold, F. (2007a) Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility.

Review of Economics and Statistics **89**(4), 701–720. **DOI:** 10.1162/rest.89.4.701

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2001b) The distribution of realized exchange rate volatility. *Journal of the American Statistical Association* **96**(453), 42–55. **DOI:** 10.1198/016214501750332965

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2003) Modeling and forecasting realized volatility. *Econometrica* **71**(2), 579–625. **DOI:** 10.1111/1468-0262.00418

Andersen, T. G., Bollerslev, T. and Dobrev, D. (2007b) No-arbitrage semi-martingale restrictions for continuous-time volatility models subject to leverage effects, jumps and i.i.d. noise: Theory and testable distributional implications. *Journal of Econometrics* **138**(1), 125–180. **DOI:** 10.1016/j.jeconom.2006.05.018

Andersen, T. G., Bollerslev, T. and Huang, X. (2011) A reduced form framework for modeling volatility of speculative prices based on realized variation measures. *Journal of Econometrics* **160**(1), 176–189. **DOI:** 10.1016/j.jeconom.2010.03.029

Anderson, E., Bai, Z., Bischof, C., Blackford, S., Demmel, J., Dongarra, J., Du Croz, J., Greenbaum, A., Hammerling, S., McKenney, A. and Sorensen, D. (1999) *LAPACK users' guide: Third Edition*. Software, Environments, and Tools. Society for Industrial and Applied Mathematics. ISBN 978-0-89871-447-0

Anderson, E., Bai, Z. and Dongarra, J. (1992) Generalized QR factorization and its applications. *Linear Algebra and its Applications* **162–164**, 243–271. **DOI:** 10.1016/0024-3795(92)90379-O

Andrews, D. W. K. (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* **59**(3), 817–858. **DOI:** 10.2307/2938229

Antonanzas, J., Osorio, N., Escobar, R., Urraca, R., Martinez-de-Pison, F. and Antonanzas-Torres, F. (2016) Review of photovoltaic power forecasting. *Solar Energy* **136**, 78–111. **DOI:** 10.1016/j.solener.2016.06.069

Ashouri, M., Hyndman, R. J. and Shmueli, G. (2022) Fast forecast reconciliation using linear models. *Journal of Computational and Graphical Statistics* **31**(1), 263–282. **DOI:** 10.1080/10618600.2021.1939038

Athanasopoulos, G., Ahmed, R. A. and Hyndman, R. J. (2009) Hierarchical forecasts for Australian domestic tourism. *International Journal of Forecasting* **25**(1), 146–166. **DOI:** 10.1016/j.ijforecast.2008.07.004

Athanasopoulos, G., Gamakumara, P., Panagiotelis, A., Hyndman, R. J. and Affan, M. (2020) Hierarchical forecasting. In *Macroeconomic forecasting in the era of big data*, ed. P. Fuleky, volume 52, pp. 689–719. Cham: Springer International Publishing. **DOI:** 10.1007/978-3-030-31150-6_21

Athanasopoulos, G., Hyndman, R. J., Kourentzes, N. and Panagiotelis, A. (2023) Forecast reconciliation: A review. *International Journal of Forecasting* (in press). **URL:** https://www.monash.edu/business/ebs/research/publications/ebs/2023/wp08-2023.pdf

Athanasopoulos, G., Hyndman, R. J., Kourentzes, N. and Petropoulos, F. (2017) Forecasting with temporal hierarchies. *European Journal of Operational Research* **262**(1), 60–74. **DOI:** 10.1016/j.ejor.2017.02.046

Athanasopoulos, G. and Kourentzes, N. (2022) On the evaluation of hierarchical forecasts. *International Journal of Forecasting* **39**(4), 1502–1511. **DOI:** 10.1016/j.ijforecast.2022.08.003

Baker, L., Cristea, I., Errington, T., Jaśko, K., Lusoli, W., MacCallum, C., Parry, V., Pérignon, C., Šimko, T. and Winchester, C. (2020) *Reproducibility of scientific results in the EU*. European Commission and Directorate-General for Research and Innovation. Publications Office. **URL:** https://data.europa.eu/doi/10.2777/341654

Barndorff-Nielsen, O. E., Kinnebrouk, S. and Shephard, N. (2010) Measuring downside risk: Realised semivariance. In *Volatility and time series econometrics: Essays in honor of Robert F. Engle*, eds T. Bollerslev, J. Russell and M. Watson, pp. 117–136. Oxford University Press

Bates, D., Maechler, M., Jagan, M., Davis, T. A., Oehlschlägel, J. and Riedy, J. (2023) *Package Matrix: Sparse and dense matrix classes and methods*. R package v1.5-4. **URL:** https://CRAN.R-project.org/package=Matrix

Bates, J. M. and Granger, C. W. J. (1969) The combination of forecasts. *Journal of the Operational Research Society* **20**(4), 451–468. **DOI:** 10.1057/jors.1969.103

Ben Taieb, S., Huser, R., Hyndman, R. J. and Genton, M. G. (2016) Forecasting uncertainty in electricity smart meter data by boosting additive quantile regression. *IEEE Transactions on Smart Grid* **7**(5), 2448–2455. **DOI:** 10.1109/TSG.2016.2527820

Ben Taieb, S. and Koo, B. (2019) Regularized regression for hierarchical forecasting without unbiasedness conditions. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 1337–1347. Anchorage AK USA: ACM. **DOI:** 10.1145/3292500.3330976

Ben Taieb, S., Taylor, J. W. and Hyndman, R. J. (2017) Coherent probabilistic forecasts for hierarchical time series. In *Proceedings of the 34th International Conference on Machine Learning*, pp. 3348–3357. PMLR

Ben Taieb, S., Taylor, J. W. and Hyndman, R. J. (2021) Hierarchical probabilistic forecasting of electricity demand with smart meter data. *Journal of the American Statistical Association* **116**(533), 27–43. **DOI:** 10.1080/01621459.2020.1736081

Benavides Cesar, L., Amaro e Silva, R., Manso Callejo, M. Á. and Cira, C.-I. (2022) Review on spatio-temporal solar forecasting methods driven by in situ measurements or their combination with satellite and numerical weather prediction (NWP) estimates. *Energies* **15**(12), 4341. **DOI:** 10.3390/en15124341

Berry, L. R., Helman, P. and West, M. (2020) Probabilistic forecasting of heterogeneous consumer transaction–sales time series. *International Journal of Forecasting* **36**(2), 552–569. **DOI:** 10.1016/j.ijforecast.2019.07.007

Bisaglia, L., Fonzo, T. D. and Girolimetto, D. (2020) Fully reconciled GDP forecasts from Income and Expenditure sides. In *Book of short papers SIS 2020*, eds A. Pollice, N. Salvati and F. Schirripa Spagnolo, pp. 951–956. Pearson. ISBN 978-88-919-1077-6

Bollerslev, T. (2022) Realized semi(co)variation: Signs that all volatilities are not created equal. *Journal of Financial Econometrics* **20**(2), 219–252. **DOI:** 10.1093/jjfinec/nbab025

Bollerslev, T., Medeiros, M. C., Patton, A. J. and Quaedvlieg, R. (2022) From zero to hero: Realized partial (co)variances. *Journal of Econometrics* **231**(2), 348–360. **DOI:** 10.1016/j.jeconom.2021.04.013

Bollerslev, T., Patton, A. J. and Quaedvlieg, R. (2016) Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Econometrics* **192**(1), 1–18. **DOI:** 10.1016/j.jeconom.2015.10.007

Böse, J.-H., Flunkert, V., Gasthaus, J., Januschowski, T., Lange, D., Salinas, D., Schelter, S., Seeger, M. and Wang, Y. B. (2017) Probabilistic demand forecasting at scale. *Proceedings of the VLDB Endowment* **10**(12), 1694–1705

Boudt, K., Croux, C. and Laurent, S. (2011) Robust estimation of intraweek periodicity in volatility and jump detection. *Journal of Empirical Finance* **18**(2), 353–367. **DOI:** 10.1016/j.jempfin.2010.11.005

Boylan, J. E., Goodwin, P., Mohammadipour, M. and Syntetos, A. A. (2015) Reproducibility in forecasting research. *International Journal of Forecasting* **31**(1), 79–90. **DOI:** 10.1016/j.ijforecast.2014.05.008

Byron, R. P. (1978) The estimation of large social account matrices. *Journal of the Royal Statistical Society. Series A (General)* **141**(3), 359. **DOI:** 10.2307/2344807

Byron, R. P. (1979) Corrigenda: The estimation of large social account matrices. *Journal of the Royal Statistical Society. Series A (General)* **142**(3), 405. **DOI:** 10.2307/2982515

Caporin, M. (2022) The role of jumps in realized volatility modeling and forecasting. *Journal of Financial Econometrics* **21**(4), 1143–1168. **DOI:** 10.1093/jjfinec/nbab030

Caporin, M., Di Fonzo, T. and Girolimetto, D. (2023) Exploiting intraday decompositions in realized volatility forecasting: A forecast reconciliation approach. *arXiv* **DOI:** 10.48550/arXiv.2306.02952

Chase, C. W. (2013) Using big data to enhance demand-driven forecasting and planning. *The Journal of Business Forecasting* **32**(2), 27–32

Chen, B., Di Fonzo, T. and Mushkudiani, N. (2018) Benchmarking, temporal disaggregation, and reconciliation of systems of time series: Introduction. *Statistica Neerlandica* **72**(4), 402–405. **DOI:** 10.1111/stan.12157

Chow, G. C. and Lin, A.-l. (1971) Best linear unbiased interpolation, distribution, and extrapolation of time series by related series. *The Review of Economics and Statistics* **53**(4), 372. **DOI:** 10.2307/1928739

Clemen, R. T. (1989) Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting* **5**(4), 559–583. **DOI:** 10.1016/0169-2070(89)90012-5

Clements, M. P. (2004) Evaluating the Bank of England density forecasts of inflation. *The Economic Journal* **114**(498), 844–866

Clements, M. P. (2018) Are macroeconomic density forecasts informative? *International Journal of Forecasting* **34**(2), 181–198. **DOI:** 10.1016/j.ijforecast.2017.10.004

Corani, G., Azzimonti, D., Augusto, J. P. S. C. and Zaffalon, M. (2021) Probabilistic reconciliation of hierarchical forecast via bayes' rule. In *Machine Learning and Knowledge Discovery in Databases*, eds F. Hutter, K. Kersting, J. Lijffijt and I. Valera, volume 12459, pp. 211–226. Cham: Springer International Publishing. ISBN 978-3-030-67663-6 978-3-030-67664-3. **DOI:** 10.1007/978-3-030-67664-3_13

Corani, G., Azzimonti, D. and Rubattu, N. (2023) Probabilistic reconciliation of count time series. *International Journal of Forecasting* (in press). **DOI:** 10.1016/j.ijforecast.2023.04.003

Corona, F., Guerrero, V. M. and López-Peréz, J. (2021) Optimal reconciliation of seasonally adjusted disaggregates taking into account the difference between direct and indirect adjustment of the aggregate. *Journal of Official Statistics* **37**(1), 31–51. **DOI:** 10.2478/jos-2021-0002

Corsi, F. (2009) A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics* **7**(2), 174–196. **DOI:** 10.1093/jjfinec/nbp001

Corsi, F., Mittnik, S., Pigorsch, C. and Pigorsch, U. (2008) The volatility of realized volatility. *Econometric Reviews* **27**(1-3), 46–78. **DOI:** 10.1080/07474930701853616

Corsi, F. and Renò, R. (2012) Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. *Journal of Business & Economic Statistics* **30**(3), 368–380. **DOI:** 10.1080/07350015.2012.663261

Dagum, E. B. and Cholette, P. A. (2006) *Benchmarking, Temporal Distribution, and Reconciliation Methods for Time Series*. Volume 186 of *Lecture Notes in Statistics*. Springer New York. ISBN 978-0-387-31102-9. **DOI:** 10.1007/0-387-35439-5

Dangerfield, B. J. and Morris, J. S. (1992) Top-down or bottom-up: Aggregate versus disaggregate extrapolations. *International Journal of Forecasting* **8**(2), 233–241. **DOI:** 10.1016/0169-2070(92)90121-0

Danilov, D. and Magnus, J. R. (2008) On the estimation of a large sparse Bayesian system: The Snaer program. *Computational Statistics & Data Analysis* **52**(9), 4203–4224. **DOI:** 10.1016/j.csda.2008.02.019

Davis, T. A. (2006) *Direct Methods for Sparse Linear Systems*. Fundamentals of Algorithms. Philadelphia: Society for Industrial and Applied Mathematics. ISBN 978-0-89871-613-9

Davydenko, A. and Fildes, R. (2013) Measuring forecasting accuracy: The case of judgmental adjustments to SKU-level demand forecasts. *International Journal of Forecasting* **29**(3), 510–522. **DOI:** 10.1016/j.ijforecast.2012.09.002

Deming, W. E. and Stephan, F. F. (1940) On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *The Annals of Mathematical Statistics* **11**(4), 427–444. **DOI:** 10.1214/aoms/1177731829

Denton, F. T. (1971) Adjustment of monthly or quarterly series to annual totals: An approach based on quadratic minimization. *Journal of the American Statistical Association* **66**(333), 99–102. **DOI:** 10.1080/01621459.1971.10482227

Di Fonzo, T. (1990) The estimation of M disaggregate time series when contemporaneous and temporal aggregates are known. *The Review of Economics and Statistics* **72**(1), 178. **DOI:** 10.2307/2109758

Di Fonzo, T. and Girolimetto, D. (2022a) Enhancements in cross-temporal forecast reconciliation, with an application to solar irradiance forecasts. *arXiv* **DOI:** 10.48550/arXiv.2209.07146

Di Fonzo, T. and Girolimetto, D. (2022b) Fully reconciled probabilistic GDP forecasts from Income and Expenditure sides. In *Book of short papers SIS 2022*, eds A. Balzanella, M. Bini, C. Cavicchia and R. Verde, pp. 1376–1381. Pearson. ISBN 978-88-919-3231-0

Di Fonzo, T. and Girolimetto, D. (2022c) Forecast combination-based forecast reconciliation: Insights and extensions. *International Journal of Forecasting* (in press). **DOI:** 10.1016/j.ijforecast.2022.07.001

Di Fonzo, T. and Girolimetto, D. (2023a) Cross-temporal forecast reconciliation: Optimal combination method and heuristic alternatives. *International Journal of Forecasting* **39**(1), 39–57. **DOI:** 10.1016/j.ijforecast.2021.08.004

Di Fonzo, T. and Girolimetto, D. (2023b) Spatio-temporal reconciliation of solar forecasts. *Solar Energy* **251**, 13–29. **DOI:** 10.1016/j.solener.2023.01.003

Di Fonzo, T. and Marini, M. (2011) Simultaneous and two-step reconciliation of systems of time series: Methodological and practical issues. *Journal of the Royal Statistical Society Series C: Applied Statistics* **60**(2), 143–164. **DOI:** 10.1111/j.1467-9876.2010.00733.x

Di Fonzo, T. and Marini, M. (2015) Reconciliation of systems of time series according to a growth rates preservation principle. *Statistical Methods & Applications* **24**(4), 651–669. **DOI:** 10.1007/s10260-015-0322-y

Diebold, F. and Mariano, R. (1995) Comparing predictive accuracy. *Journal of Business & Economic Statistics* **13**(3), 253–63

Dunn, D. M., Williams, W. H. and Dechaine, T. L. (1976) Aggregate versus subaggregate models in local area forecasting. *Journal of the American Statistical Association* **71**(353), 68–71. **DOI:** 10.1080/01621459.1976.10481478

Eckert, F., Hyndman, R. J. and Panagiotelis, A. (2021) Forecasting Swiss exports using Bayesian forecast reconciliation. *European Journal of Operational Research* **291**(2), 693–710. **DOI:** 10.1016/j.ejor.2020.09.046

Efron, B. (1975) Biased versus unbiased estimation. *Advances in Mathematics* **16**(3), 259–277. **DOI:** 10.1016/0001-8708(75)90114-0

Efron, B. and Morris, C. (1975) Data analysis using Stein's estimator and its generalizations. *Journal of the American Statistical Association* **70**(350), 311–319. **DOI:** 10.1080/01621459.1975.10479864

Efron, B. and Morris, C. (1977) Stein's paradox in statistics. *Scientific American* **236**(5), 119–127

Engle, R. F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50**(4), 987. **DOI:** 10.2307/1912773

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

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

Fleming, P. J. and Wallace, J. J. (1986) How not to lie with statistics: The correct way to summarize benchmark results. *Communications of the ACM* **29**(3), 218–221. **DOI:** 10.1145/5666.5673

Fliedner, G. (2001) Hierarchical forecasting: Issues and use guidelines. *Industrial Management & Data Systems* **101**(1), 5–12. **DOI:** 10.1108/02635570110365952

Frale, C., Marcellino, M., Mazzi, G. L. and Proietti, T. (2011) EUROMIND: A monthly indicator of the euro area economic conditions. *Journal of the Royal Statistical Society Series A: Statistics in Society* **174**(2), 439–470. **DOI:** 10.1111/j.1467-985X.2010.00675.x

Genre, V., Kenny, G., Meyler, A. and Timmermann, A. (2013) Combining expert forecasts: Can anything beat the simple average? *International Journal of Forecasting* **29**(1), 108–121. **DOI:** 10.1016/j.ijforecast.2012.06.004

Genton, M. G. (2007) Separable approximations of space-time covariance matrices. *Environmetrics* **18**(7), 681–695. **DOI:** 10.1002/env.854

Geweke, J. and Amisano, G. (2011) Optimal prediction pools. *Journal of Econometrics* **164**(1), 130–141. **DOI:** 10.1016/j.jeconom.2011.02.017

Girolimetto, D., Athanasopoulos, G., Di Fonzo, T. and Hyndman, R. J. (2023) Cross-temporal probabilistic forecast reconciliation: Methodological and practical issues. *International Journal of Forecasting* (in press). **DOI:** 10.1016/j.ijforecast.2023.10.003

Girolimetto, D. and Di Fonzo, T. (2023a) *FoReco: Point forecast reconciliation.* R package v0.2.6. **URL:** https://danigiro.github.io/FoReco/

Girolimetto, D. and Di Fonzo, T. (2023b) Point and probabilistic forecast reconciliation for general linearly constrained multiple time series. *Statistical Methods & Applications* (in press). **DOI:** 10.1007/s10260-023-00738-6

Gneiting, T. and Katzfuss, M. (2014) Probabilistic forecasting. *Annual Review of Statistics and Its Application* **1**(1), 125–151. **DOI:** 10.1146/annurev-statistics-062713-085831

Gneiting, T., Stanberry, L. I., Grimit, E. P., Held, L. and Johnson, N. A. (2008) Assessing probabilistic forecasts of multivariate quantities, with an application to ensemble predictions of surface winds. *TEST* **17**(2), 211–235. **DOI:** 10.1007/s11749-008-0114-x

Golub, G. H. and Van Loan, C. F. (1996) *Matrix computations*. Baltimore: Johns Hopkins University Press.

Grassi, S., Proietti, T., Frale, C., Marcellino, M. and Mazzi, G. (2015) EuroMInd-C: A disaggregate monthly indicator of economic activity for the euro area and member countries. *International Journal of Forecasting* **31**(3), 712–738. **DOI:** 10.1016/j.ijforecast.2014.08.015

Gross, C. W. and Sohl, J. E. (1990) Disaggregation methods to expedite product line forecasting. *Journal of Forecasting* **9**(3), 233–254. **DOI:** 10.1002/for.3980090304

Hansen, P. R. (2005) A Test for Superior Predictive Ability. *Journal of Business & Economic Statistics* **23**(4), 365–380. **DOI:** 10.1198/073500105000000063

Hansen, P. R., Lunde, A. and Nason, J. M. (2011) The model confidence set. *Econometrica* **79**(2), 453–497. **DOI:** 10.3982/ECTA5771

Harville, D. A. (2008) *Matrix algebra from a statistician's perspective*. New York: Springer. ISBN 978-0-387-78356-7 978-0-387-22677-4

Henderson, H. V. and Searle, S. R. (1981) The vec-permutation matrix, the vec operator and Kronecker products: A review. *Linear and Multilinear Algebra* **9**(4), 271–288. **DOI:** 10.1080/03081088108817379

Hendry, D. F. and Clements, M. P. (2004) Pooling of forecasts. *The Econometrics Journal* **7**(1), 1–31. **DOI:** 10.1111/j.1368-423X.2004.00119.x

Hollyman, R., Petropoulos, F. and Tipping, M. E. (2021) Understanding forecast reconciliation. *European Journal of Operational Research* **294**(1), 149–160. **DOI:** 10.1016/j.ejor.2021.01.017

Hong, T., Pinson, P., Fan, S., Zareipour, H., Troccoli, A. and Hyndman, R. J. (2016) Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond. *International Journal of Forecasting* **32**(3), 896–913. **DOI:** 10.1016/j.ijforecast.2016.02.001

Hyndman, R. J. (2008) *Forecasting with exponential smoothing: The state space approach*. Springer Series in Statistics. Berlin: Springer. ISBN 978-3-540-71918-2

Hyndman, R. J. (2010) Encouraging replication and reproducible research. *International Journal of Forecasting* **26**(1), 2–3. **DOI:** 10.1016/j.ijforecast.2009.12.003

Hyndman, R. J. (2022) Notation for forecast reconciliation. *https://robjhyndman.com/hyndsight/reconciliation-notation.html* pp. Accessed: 2023–07–20. **URL:** https://robjhyndman.com/hyndsight/reconciliation-notation.html

Hyndman, R. J., Ahmed, R. A., Athanasopoulos, G. and Shang, H. L. (2011) Optimal combination forecasts for hierarchical time series. *Computational Statistics & Data Analysis* **55**(9), 2579–2589. **DOI:** 10.1016/j.csda.2011.03.006

Hyndman, R. J. and Athanasopoulos, G. (2021) *Forecasting: Principles and Practice*. Third edition. Melbourne: OTexts.

Hyndman, R. J., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., Kuroptev, K., O'Hara-Wild, M., Petropoulos, F., Razbash, S., Wang, E., Yasmeen, F., Garza, F., Girolimetto, D., Ihaka, R., R Core Team, Reid, D., Shaub, D., Tang, Y., Wang, X. and Zhou, Z. (2023) *forecast: Forecasting functions for time series and linear models*. R package v8.21. **URL:** `https://pkg.robjhyndman.com/forecast/`

Hyndman, R. J. and Khandakar, Y. (2008) Automatic time series forecasting: The Forecast package for R. *Journal of Statistical Software* **27**(3). **DOI:** `10.18637/jss.v027.i03`

Hyndman, R. J. and Kourentzes, N. (2018) *Package thief: Temporal HIErarchical Forecasting*. R package v0.3. **URL:** `http://pkg.robjhyndman.com/thief/`

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

Hyndman, R. J., Lee, A. J., Wang, E. and Wickramasuriya, S. L. (2022) *Package hts: Hierarchical and grouped time series*. R package v6.0.2. **URL:** `https://pkg.earo.me/hts/`

Jeon, J., Panagiotelis, A. and Petropoulos, F. (2019) Probabilistic forecast reconciliation with applications to wind power and electric load. *European Journal of Operational Research* **279**(2), 364–379. **DOI:** `10.1016/j.ejor.2019.05.020`

Jeon, J. and Taylor, J. W. (2012) Using conditional kernel density estimation for wind power density forecasting. *Journal of the American Statistical Association* **107**(497), 66–79. **DOI:** `10.1080/01621459.2011.643745`

Kleissl, J. (ed.) (2013) *Solar energy forecasting and resource assessment*. Oxford: Elsevier Ltd. ISBN 978-0-12-397772-4

Kolassa, S. (2016) Evaluating predictive count data distributions in retail sales forecasting. *International Journal of Forecasting* **32**(3), 788–803. **DOI:** `10.1016/j.ijforecast.2015.12.004`

Koning, A. J., Franses, P. H., Hibon, M. and Stekler, H. (2005) The M3 competition: Statistical tests of the results. *International Journal of Forecasting* **21**(3), 397–409. **DOI:** `10.1016/j.ijforecast.2004.10.003`

Kourentzes, N. (2022) *tsutils: Time series exploration, modelling and forecasting*. R package v0.9.3. **URL:** `https://CRAN.R-project.org/package=tsutils`

Kourentzes, N. and 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

Kourentzes, N. and Athanasopoulos, G. (2021) Elucidate structure in intermittent demand series. *European Journal of Operational Research* **288**(1), 141–152. **DOI:** 10.1016/j.ejor.2020.05.046

Kourentzes, N., Barrow, D. and Petropoulos, F. (2019) Another look at forecast selection and combination: Evidence from forecast pooling. *International Journal of Production Economics* **209**, 226–235. **DOI:** 10.1016/j.ijpe.2018.05.019

Kourentzes, N. and Petropoulos, F. (2016) Forecasting with multivariate temporal aggregation: The case of promotional modelling. *International Journal of Production Economics* **181**, 145–153. **DOI:** 10.1016/j.ijpe.2015.09.011

Kourentzes, N., Petropoulos, F. and Trapero, J. R. (2014) Improving forecasting by estimating time series structural components across multiple frequencies. *International Journal of Forecasting* **30**(2), 291–302. **DOI:** 10.1016/j.ijforecast.2013.09.006

Kourentzes, N., Rostami-Tabar, B. and Barrow, D. K. (2017) Demand forecasting by temporal aggregation: Using optimal or multiple aggregation levels? *Journal of Business Research* **78**, 1–9. **DOI:** 10.1016/j.jbusres.2017.04.016

Ledoit, O. and Wolf, M. (2004) A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis* **88**(2), 365–411. **DOI:** 10.1016/S0047-259X(03)00096-4

Leon, S. J. (2015) *Linear algebra with applications*. 9th edition. Boston: Pearson. ISBN 978-0-321-96221-8

Leutbecher, M. (2019) Ensemble size: How suboptimal is less than infinity? *Quarterly Journal of the Royal Meteorological Society* **145**(S1), 107–128. **DOI:** 10.1002/qj.3387

Leutbecher, M. and Palmer, T. (2008) Ensemble forecasting. *Journal of Computational Physics* **227**(7), 3515–3539. **DOI:** 10.1016/j.jcp.2007.02.014

Lichtendahl, K. C. and Winkler, R. L. (2020) Why do some combinations perform better than others? *International Journal of Forecasting* **36**(1), 142–149. **DOI:** 10.1016/j.ijforecast.2019.03.027

Liu, L., Moon, H. R. and Schorfheide, F. (2021) Panel forecasts of country-level Covid-19 infections. *Journal of Econometrics* **220**(1), 2–22. **DOI:** 10.1016/j.jeconom.2020.08.010

Lu, T.-T. and Shiou, S.-H. (2002) Inverses of 2 × 2 block matrices. *Computers & Mathematics with Applications* **43**(1-2), 119–129. **DOI:** 10.1016/S0898-1221(01)00278-4

Lyche, T. (2020) *Numerical linear algebra and matrix factorizations*. New York: Springer. ISBN 978-3-030-36468-7

Magnus, J. R. and Neudecker, H. (2019) *Matrix differential calculus with applications in statistics and econometrics*. Third edition. New York: Wiley. ISBN 978-0-471-98632-4

Makarov, Y. V., Etingov, P. V., Ma, J., Huang, Z. and Subbarao, K. (2011) Incorporating uncertainty of wind power generation forecast into power system operation, dispatch, and unit commitment procedures. *IEEE Transactions on Sustainable Energy* **2**(4), 433–442. **DOI:** 10.1109/TSTE.2011.2159254

Makridakis, S., Assimakopoulos, V. and Spiliotis, E. (2018) Objectivity, reproducibility and replicability in forecasting research. *International Journal of Forecasting* **34**(4), 835–838. **DOI:** 10.1016/j.ijforecast.2018.05.001

Makridakis, S., Spiliotis, E. and Assimakopoulos, V. (2022) M5 accuracy competition: Results, findings, and conclusions. *International Journal of Forecasting* **38**(4), 1346–1364. **DOI:** 10.1016/j.ijforecast.2021.11.013

Mancuso, P., Piccialli, V. and Sudoso, A. M. (2021) A machine learning approach for forecasting hierarchical time series. *Expert Systems with Applications* **182**, 115102. **DOI:** 10.1016/j.eswa.2021.115102

Marcellino, M. (2004) Forecast pooling for European macroeconomic variables. *Oxford Bulletin of Economics and Statistics* **66**(1), 91–112. **DOI:** 10.1111/j.1468-0084.2004.00071.x

Marcellino, M., Stock, J. H. and Watson, M. W. (2003) Macroeconomic forecasting in the Euro area: Country specific versus area-wide information. *European Economic Review* **47**(1), 1–18. **DOI:** 10.1016/S0014-2921(02)00206-4

Marquardt, D. W. (1970) Generalized Inverses, Ridge Regression, Biased Linear Estimation, and Nonlinear Estimation. *Technometrics* **12**(3), 591–612. **DOI:** 10.2307/1267205

Matheson, J. E. and Winkler, R. L. (1976) Scoring rules for continuous probability distributions. *Management Science* **22**(10), 1087–1096

McLean Sloughter, J., Gneiting, T. and Raftery, A. E. (2013) Probabilistic wind vector forecasting using ensembles and Bayesian model averaging. *Monthly Weather Review* **141**(6), 2107–2119. **DOI:** 10.1175/MWR-D-12-00002.1

Meyer, C. D. (2000) *Matrix analysis and applied linear algebra*. Philadelphia: Society for Industrial and Applied Mathematics. ISBN 978-0-89871-454-8

Miller, R. E. and Blair, P. D. (2009) *Input-output analysis: Foundations and extensions*. Second edition. Cambridge University Press. ISBN 978-0-521-73902-3 978-0-521-51713-3 978-0-511-62698-2. **DOI:** 10.1017/CBO9780511626982

Mircetic, D., Rostami-Tabar, B., Nikolicic, S. and Maslaric, M. (2022) Forecasting hierarchical time series in supply chains: An empirical investigation. *International Journal of Production Research* **60**(8), 2514–2533. **DOI:** 10.1080/00207543.2021.1896817

Nystrup, P., Lindström, E., Møller, J. K. and Madsen, H. (2021) Dimensionality reduction in forecasting with temporal hierarchies. *International Journal of Forecasting* **37**(3), 1127–1146. **DOI:** 10.1016/j.ijforecast.2020.12.003

Nystrup, P., Lindström, E., Pinson, P. and Madsen, H. (2020) Temporal hierarchies with autocorrelation for load forecasting. *European Journal of Operational Research* **280**(3), 876–888. **DOI:** 10.1016/j.ejor.2019.07.061

Olivares, K. G., Garza, F., Luo, D., Challú, C., Mergenthaler, M., Taieb, S. B., Wickramasuriya, S. L. and Dubrawski, A. (2022) HierarchicalForecast: A reference framework for hierarchical forecasting in Python. *arXiv* **DOI:** 10.48550/arXiv.2207.03517

Olivares, K. G., Luo, D., Challu, C., La Vattiata, S., Mergenthaler, M. and Dubrawski, A. (2023a) HINT: Hierarchical mixture networks for coherent probabilistic forecasting. *arXiv* **DOI:** 10.48550/arXiv.2305.07089

Olivares, K. G., Meetei, O. N., Ma, R., Reddy, R., Cao, M. and Dicker, L. (2023b) Probabilistic hierarchical forecasting with deep poisson mixtures. *International Journal of Forecasting* (in press). **DOI:** 10.1016/j.ijforecast.2023.04.007

Paige, C. C. and Saunders, M. A. (1982) LSQR: An algorithm for sparse linear equations and sparse least squares. *ACM Transactions on Mathematical Software* **8**(1), 43–71. **DOI:** 10.1145/355984.355989

Panagiotelis, A., Athanasopoulos, G., Gamakumara, P. and Hyndman, R. J. (2021) Forecast reconciliation: A geometric view with new insights on bias correction. *International Journal of Forecasting* **37**(1), 343–359. **DOI:** 10.1016/j.ijforecast.2020.06.004

Panagiotelis, A., Gamakumara, P., Athanasopoulos, G. and Hyndman, R. J. (2023) Probabilistic forecast reconciliation: Properties, evaluation and score optimisation. *European Journal of Operational Research* **306**(2), 693–706. **DOI:** 10.1016/j.ejor.2022.07.040

Panamtash, H. and Zhou, Q. (2018) Coherent probabilistic solar power forecasting. In *2018 IEEE International Conference on Probabilistic Methods Applied to Power*

Systems (PMAPS), pp. 1–6. Boise, ID: IEEE. ISBN 978-1-5386-3596-4. **DOI:** 10.1109/PMAPS.2018.8440483

Papadakis, M., Tsagris, M., Dimitriadis, M., Fafalios, S., Tsamardinos, I., Fasiolo, M., Borboudakis, G., Burkardt, J., Zou, C., Lakiotaki, K. and Chatzipantsiou, C. (2022) *Rfast: A collection of efficient and extremely fast R functions*. R package v2.0.6. **URL:** https://CRAN.R-project.org/package=Rfast

Patton, A. J. (2011a) Data-based ranking of realised volatility estimators. *Journal of Econometrics* **161**(2), 284–303. **DOI:** 10.1016/j.jeconom.2010.12.010

Patton, A. J. (2011b) Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics* **160**(1), 246–256. **DOI:** 10.1016/j.jeconom.2010.03.034

Patton, A. J. and Sheppard, K. (2009) Optimal combinations of realised volatility estimators. *International Journal of Forecasting* **25**(2), 218–238. **DOI:** 10.1016/j.ijforecast.2009.01.011

Patton, A. J. and Sheppard, K. (2015) Good volatility, bad volatility: Signed jumps and the persistence of volatility. *Review of Economics and Statistics* **97**(3), 683–697. **DOI:** 10.1162/REST_a_00503

Pennings, C. L. and Van Dalen, J. (2017) Integrated hierarchical forecasting. *European Journal of Operational Research* **263**(2), 412–418. **DOI:** 10.1016/j.ejor.2017.04.047

Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M. Z., Barrow, D. K., Ben Taieb, S., Bergmeir, C., Bessa, R. J., Bijak, J., Boylan, J. E., Browell, J., Carnevale, C., Castle, J. L., Cirillo, P., Clements, M. P., Cordeiro, C., Cyrino Oliveira, F. L., De Baets, S., Dokumentov, A., Ellison, J., Fiszeder, P., Franses, P. H., Frazier, D. T., Gilliland, M., Gönül, M. S., Goodwin, P., Grossi, L., Grushka-Cockayne, Y., Guidolin, M., Guidolin, M., Gunter, U., Guo, X., Guseo, R., Harvey, N., Hendry, D. F., Hollyman, R., Januschowski, T., Jeon, J., Jose, V. R. R., Kang, Y., Koehler, A. B., Kolassa, S., Kourentzes, N., Leva, S., Li, F., Litsiou, K., Makridakis, S., Martin, G. M., Martinez, A. B., Meeran, S., Modis, T., Nikolopoulos, K., Önkal, D., Paccagnini, A., Panagiotelis, A., Panapakidis, I., Pavía, J. M., Pedio, M., Pedregal, D. J., Pinson, P., Ramos, P., Rapach, D. E., Reade, J. J., Rostami-Tabar, B., Rubaszek, M., Sermpinis, G., Shang, H. L., Spiliotis, E., Syntetos, A. A., Talagala, P. D., Talagala, T. S., Tashman, L., Thomakos, D., Thorarinsdottir, T., Todini, E., Trapero Arenas, J. R., Wang, X., Winkler, R. L., Yusupova, A. and Ziel, F. (2022) Forecasting: Theory and practice. *International Journal of Forecasting* **38**(3), 705–871. **DOI:** 10.1016/j.ijforecast.2021.11.001

Petropoulos, F. and Kourentzes, N. (2015) Forecast combinations for intermittent demand. *Journal of the Operational Research Society* **66**(6), 914–924. **DOI:** 10.1057/jors.2014.62

Petropoulos, F., Makridakis, S., Assimakopoulos, V. and Nikolopoulos, K. (2014) 'Horses for Courses' in demand forecasting. *European Journal of Operational Research* **237**(1), 152–163. **DOI:** 10.1016/j.ejor.2014.02.036

Pinson, P., Madsen, H., Nielsen, H. A., Papaefthymiou, G. and Klöckl, B. (2009) From probabilistic forecasts to statistical scenarios of short-term wind power production. *Wind Energy* **12**(1), 51–62. **DOI:** 10.1002/we.284

Politis, D. N. and Romano, J. P. (1994) The stationary bootstrap. *Journal of the American Statistical Association* **89**(428), 1303–1313. **DOI:** 10.1080/01621459.1994.10476870

Poncela, P. and García-Ferrer, A. (2014) The effects of disaggregation on forecasting nonstationary time series. *Journal of Forecasting* **33**(4), 300–314. **DOI:** 10.1002/for.2291

Punia, S., Singh, S. P. and 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

Quenneville, B. and Fortier, S. (2012) Restoring accounting constraints in time series: Methods and software for a statistical agency. In *Economic Time Series: Modeling and Seasonality*, eds W. R. Bell, S. H. Holan and T. S. McElroy, pp. 31–253. Taylor & Francis

R Core Team (2022) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. **URL:** https://www.R-project.org/

Romagnoli, J. A. and Sánchez, M. C. (2000) *Data processing and reconciliation for chemical process operations*. San Diego: Academic Press. ISBN 978-0-08-053027-7

Rossi, N. (1982) A note on the estimation of disaggregate time series when the aggregate is known. *The Review of Economics and Statistics* **64**(4), 695. **DOI:** 10.2307/1923955

Sanguri, K., Shankar, S., Punia, S. and Patra, S. (2022) Hierarchical container throughput forecasting: The value of coherent forecasts in the management of ports operations. *Computers & Industrial Engineering* **173**, 108651. **DOI:** 10.1016/j.cie.2022.108651

Schäfer, J. and Strimmer, K. (2005) A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Statistical Applications in Genetics and Molecular Biology* **4**(1). **DOI:** 10.2202/1544-6115.1175

Scheuerer, M. and Hamill, T. M. (2015) Variogram-based proper scoring rules for probabilistic forecasts of multivariate quantities. *Monthly Weather Review* **143**(4), 1321–1334. **DOI:** 10.1175/MWR-D-14-00269.1

Sengupta, M., Habte, A., Gueymard, C., Wilbert, S. and Renne, D. (2017) Best practices handbook for the collection and use of solar resource data for solar energy applications: Second edition. Technical Report NREL/TP-5D00-68886, 1411856. **DOI:** 10.2172/1411856

Sévi, B. (2014) Forecasting the volatility of crude oil futures using intraday data. *European Journal of Operational Research* **235**(3), 643–659. **DOI:** 10.1016/j.ejor.2014.01.019

Shang, H. L. and Hyndman, R. J. (2017) Grouped functional time series forecasting: An application to age-specific mortality rates. *Journal of Computational and Graphical Statistics* **26**(2), 330–343. **DOI:** 10.1080/10618600.2016.1237877

Silva, F. L., Souza, R. C., Cyrino Oliveira, F. L., Lourenco, P. M. and Calili, R. F. (2018) A bottom-up methodology for long term electricity consumption forecasting of an industrial sector - Application to pulp and paper sector in Brazil. *Energy* **144**, 1107–1118. **DOI:** 10.1016/j.energy.2017.12.078

Simon, D. (2006) *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. Hoboken, N.J: Wiley-Interscience. ISBN 978-0-471-70858-2

Simon, D. (2010) Kalman filtering with state constraints: A survey of linear and nonlinear algorithms. *IET Control Theory & Applications* **4**(8), 1303–1318. **DOI:** 10.1049/iet-cta.2009.0032

Simon, D. and Tien Li Chia (2002) Kalman filtering with state equality constraints. *IEEE Transactions on Aerospace and Electronic Systems* **38**(1), 128–136. **DOI:** 10.1109/7.993234

Smith, R. J., Weale, M. R. and Satchell, S. E. (1998) Measurement error with accounting constraints: Point and interval estimation for latent data with an application to U.K. Gross Domestic Product. *Review of Economic Studies* **65**(1), 109–134. **DOI:** 10.1111/1467-937X.00037

Sohn, S. Y. and Lim, M. (2007) Hierarchical forecasting based on AR-GARCH model in a coherent structure. *European Journal of Operational Research* **176**(2), 1033–1040. **DOI:** 10.1016/j.ejor.2005.08.019

Spiliotis, E., Abolghasemi, M., Hyndman, R. J., Petropoulos, F. and Assimakopoulos, V. (2021) Hierarchical forecast reconciliation with machine learning. *Applied Soft Computing* **112**, 107756. **DOI:** 10.1016/j.asoc.2021.107756

Spiliotis, E., Petropoulos, F. and Assimakopoulos, V. (2019) Improving the forecasting performance of temporal hierarchies. *PLOS ONE* **14**(10), e0223422. **DOI:** 10.1371/journal.pone.0223422

Spiliotis, E., Petropoulos, F., Kourentzes, N. and Assimakopoulos, V. (2020) 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, G., Goulart, P., Bemporad, A. and Boyd, S. (2020) OSQP: An operator splitting solver for quadratic programs. *Mathematical Programming Computation* **12**(4), 637–672. **DOI:** 10.1007/s12532-020-00179-2

Stellato, B., Banjac, G., Goulart, P., Boyd, S., Anderson, E., Bansal, V. and Balasubramanian, N. (2023) *Package OSQP: Quadratic Programming Solver using the 'OSQP' Library. Version 0.6.0.8.* R package v0.6.0.8. **URL:** osqp.org

Stone, R., Champernowne, D. G. and Meade, J. E. (1942) The precision of National Income estimates. *The Review of Economic Studies* **9**(2), 111–135. **DOI:** 10.2307/2967664

Stratigakos, A., van der Meer, D., Camal, S. and 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* pp. 1–6. **DOI:** 10.1109/PMAPS53380.2022.9810610

Timmermann, A. (2006) Forecast combinations. In *Handbook of economic forecasting*, volume 1, pp. 135–196. Elsevier. ISBN 978-0-444-51395-3. **DOI:** 10.1016/S1574-0706(05)01004-9

United Nations (2015) Paris Agreement. (visited on September 8, 2022) **URL:** https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement

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

van Erven, T. and Cugliari, J. (2015) Game-theoretically optimal reconciliation of contemporaneous hierarchical time series forecasts. In *Modeling and stochastic learning for forecasting in high dimensions*, eds A. Antoniadis, J.-M. Poggi and X. Brossat, volume 217, pp. 297–317. Cham: Springer International Publishing. ISBN 978-3-319-18731-0 978-3-319-18732-7. **DOI:** 10.1007/978-3-319-18732-7_15

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

Venables, W. N. and Ripley, B. D. (2002) *Modern Applied Statistics with S*. Fourth edition. New York: Springer. **URL:** https://www.stats.ox.ac.uk/pub/MASS4/, ISBN 0-387-95457-0

Wang, S., Deng, X., Chen, H., Shi, Q. and Xu, D. (2021) A bottom-up short-term residential load forecasting approach based on appliance characteristic analysis and multi-task learning. *Electric Power Systems Research* **196**, 107233. **DOI:** 10.1016/j.epsr.2021.107233

Weale, M. (1992) Estimation of data measured with error and subject to linear restrictions. *Journal of Applied Econometrics* **7**(2), 167–174

Werner, K., Jansson, M. and Stoica, P. (2008) On estimation of covariance matrices with Kronecker product structure. *IEEE Transactions on Signal Processing* **56**(2), 478–491. **DOI:** 10.1109/TSP.2007.907834

Wickramasuriya, S. L. (2021) Properties of point forecast reconciliation approaches. *arXiv* **DOI:** 10.48550/arXiv.2103.11129

Wickramasuriya, S. L. (2024) Probabilistic forecast reconciliation under the gaussian framework. *Journal of Business & Economic Statistics* **42**(1), 272–285. **DOI:** 10.1080/07350015.2023.2181176

Wickramasuriya, S. L., Athanasopoulos, G. and Hyndman, R. J. (2019) Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization. *Journal of the American Statistical Association* **114**(526), 804–819. **DOI:** 10.1080/01621459.2018.1448825

Wickramasuriya, S. L., Turlach, B. A. and Hyndman, R. J. (2020) Optimal non-negative forecast reconciliation. *Statistics and Computing* **30**(5), 1167–1182. **DOI:** 10.1007/s11222-020-09930-0

Wytock, M. and Kolter, J. Z. (2013) Large-scale probabilistic forecasting in energy systems using sparse Gaussian conditional random fields. In *52nd IEEE Conference on Decision and Control*, pp. 1019–1024. Firenze: IEEE. ISBN 978-1-4673-5717-3 978-1-4673-5714-2 978-1-4799-1381-7. **DOI:** 10.1109/CDC.2013.6760016

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

Yagli, G. M., Yang, D. and Srinivasan, D. (2020) 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. (2019) Standard of reference in operational day-ahead deterministic solar forecasting. *Journal of Renewable and Sustainable Energy* **11**(5), 053702. **DOI:** 10.1063/1.5114985

Yang, D. (2020) 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., Alessandrini, S., Antonanzas, J., Antonanzas-Torres, F., Badescu, V., Beyer, H. G., Blaga, R., Boland, J., Bright, J. M., Coimbra, C. F., David, M., Frimane, Â., Gueymard, C. A., Hong, T., Kay, M. J., Killinger, S., Kleissl, J., Lauret, P., Lorenz, E., van der Meer, D., Paulescu, M., Perez, R., Perpiñán-Lamigueiro, O., Peters, I. M., Reikard, G., Renné, D., Saint-Drenan, Y.-M., Shuai, Y., Urraca, R., Verbois, H., Vignola, F., Voyant, C. and 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. and 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. and 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. and Xia, X. (2022) A concise overview on solar resource assessment and forecasting. *Advances in Atmospheric Sciences* **39**(8), 1239–1251. **DOI:** 10.1007/s00376-021-1372-8

Zambon, L., Azzimonti, D. and Corani, G. (2024) Efficient probabilistic reconciliation of forecasts for real-valued and count time series. *Statistics and Computing* **34**(1), 21. **DOI:** 10.1007/s11222-023-10343-y

Zarnowitz, V. and Lambros, L. A. (1987) Consensus and uncertainty in economic prediction. *Journal of Political Economy* **95**(3), 591–621. **DOI:** 10.1086/261473

Zhang, B., Kang, Y., Panagiotelis, A. and Li, F. (2023) Optimal reconciliation with immutable forecasts. *European Journal of Operational Research* **308**(2), 650–660. **DOI:** 10.1016/j.ejor.2022.11.035

Zhao, T., Wang, J. and Zhang, Y. (2019) Day-ahead hierarchical probabilistic load forecasting with linear quantile regression and empirical copulas. *IEEE Access* **7**, 80969–80979. **DOI:** 10.1109/ACCESS.2019.2922744