Editorial: Innovations in Hierarchical Forecasting

George Athanasopoulos[a,*], Rob J Hyndman[a], Nikolaos Kourentzes[b], Anastasios Panagiotelis[c]

[a] *Monash University, Australia,*

[b] *Skövde AI Lab, Sweden,*

[c] *University of Sydney, Australia,*

Work in hierarchical forecasting began by tackling a practical issue in demand planning and resource management within firms. In such scenarios, products and their associated materials are often organized in hierarchical structures across product categories, geographic regions, and so on. This organization aids in managing inventory, streamlining production planning, and enhancing procurement processes. The first forecasting solutions that ensured coherence between sub-aggregate and aggregate forecasts, were the widely implemented bottom-up and top-down forecasting approaches. In the bottom-up method, forecasts are made at the most disaggregated level of time series data, while the top-down approach starts by generating forecasts at the most aggregated level. These forecasts are then either summed up or proportionally allocated to generate forecasts for other levels within the hierarchy. A popular compromise between the bottom-up and top-down methods is the middle-out approach, which starts by generating forecasts at a middle level of the hierarchy, and then applies either top-down or bottom-up methods to obtain forecasts at the other levels.

Hierarchical forecasting significantly evolved with the introduction of forecast reconciliation by Athanasopoulos et al. (2009) and Hyndman et al. (2011). These papers highlighted the limitations of traditional bottom-up and top-down methods, and proposed a structured approach based on optimally combining forecasts from all levels of the hierarchy. Extensive follow-up research included advances in handling grouped structures, linking hierarchical forecasting with forecast combinations, a geometric understanding of the problem, and much more. These developments ignited further interest in the field, leading to new estimators, methods, and applications. Somewhat concurrently, Kourentzes et al. (2014) suggested modelling time series at various temporal aggregation levels, a concept expanded into temporal hierarchies by Athanasopoulos et al. (2017), opening additional new research avenues. Perhaps the most intriguing aspect has been to show the versatility of hierarchical forecasting beyond conventional cross-sectional data structures, demonstrating its applicability to

*Corresponding author

a wide range of modelling problems. This special issue aims to provide a comprehensive literature review, and highlight new contributions in this dynamic field.

Athanasopoulos, Hyndman, Panagiotelis, and Kourentzes offer a comprehensive review of the current state of hierarchical forecasting research. The review systematically describes and discusses the different aspects of the field, including its various forms, methodological advances, practical applications, and software developments. The subsequent papers that feature in this special issue advance the research further, proposing solutions to significant challenges in hierarchical forecasting, showcasing novel applications, and introducing fresh research questions, thereby contributing to the ongoing evolution and expansion of this area of study.

The first two papers focus on forecasting count data in the hierarchical context. *Corani, Azzimonti, and Rubattu* propose probabilistic reconciliation based on the method of virtual evidence, a generalization of Bayes' rule. The proposed reconciliation method is not restricted to count data and can be used with real-valued data as well. Nonetheless, the authors focus on the case of count data, given the lack of extant probabilistic reconciliation methods for this task. Intermittent demand and other low count time series are notoriously difficult to forecast. The paper demonstrates the benefits of the proposed method in a temporal hierarchy setting, in terms of both point and probabilistic forecasts. *Olivares, Meetei, Mab, Reddy, Cao, and Dicker* are mainly interested in generating coherent probabilistic forecasts. They propose the Deep Poisson Mixture Network (DPMN) which integrates neural networks with a statistical model based on Poisson mixtures for modeling the joint distribution of the hierarchical multivariate time series structure. The effectiveness of the DPMN is demonstrated on three datasets, evaluating both point and probabilistic predictions. Although the objective of the paper is to handle count data, clear extensions of the method beyond count data are discussed.

Next, we have a collection of papers that progress the understanding and modelling capabilities of hierarchical forecast reconciliation. *Di Fonzo and Girolimetto* revisit the level-l conditional coherent point forecast reconciliation procedure, and reveal many of its mathematical properties. This leads to several insightful extensions, drawing parallels from related bodies of work, such as forecast pooling. They provide useful simplifications of the underlying intuition behind the methods and further illustrate the effectiveness of the procedure and their proposed modifications. They conclude with a multitude of future avenues for research.

Møller, Nystrup, and Madsen focus on the issue of dimensionality reduction in estimating the variance-covariance matrix used for forecast reconciliation in the temporal setting. They propose an iterative, inference-based approach, asserting that this provides greater control over dimension reduction, unlike shrinkage methods which offer limited to no control. Their comparative analysis highlights specific areas where the relative advantages of the proposed method are shown. The application of hierarchical forecasting to diverse applications provides the setting for further methodological and modelling innovations.

Yang, Shang, and Raymer take up the task of fertility rate forecasting. The data are organized as functional time series, while exhibiting a hierarchical structure. This introduces a series of new challenges, requiring adaptations of the usual hierarchical methods. We draw attention to the innovative adaptation of the conventionally binary encoded summing matrix, which captures the hierarchical structure. This adaptation incorporates population ratios, demonstrating its potential as a valuable methodological extension.

From fertility to mortality. *Li and Chen* focus on forecasting mortality for solvency capital requirements in insurance. Beyond the novelty of the application itself, they propose a combination of hierarchical forecasting with extreme value theory to obtain coherent probabilistic forecasts. The extreme value theory comes into play to model the residuals of the time series models applied to the individual mortality series. They proceed to show that if extreme mortality risk is not considered, the estimated solvency capital requirements would be underestimated, leading to erroneous decisions.

Cengiz and Tekgüç are interested in improving the accuracy of causal effect estimates of different treatments, such as new policies. They address the challenge of estimating heterogeneous effect sizes on different subgroups within a hierarchical structure, proposing counterfactual reconciliation. The method incorporates aggregation constraints in hierarchical or grouped data, to reconcile the counterfactual estimates, leading to more precise estimates of causal effects. The authors recognize that outcome variables may often undergo transformations, such as nonlinear ones. This may be due to the theoretical context within which the effect size is modelled. They proceed to discuss hierarchical reconciliation and its properties in this case and demonstrate that it can lead to better estimates of the effect.

Ghelasi and Ziel focus on forecasting day-ahead electricity market aggregate supply and demand curves. They illustrate how these curves can be represented as hierarchical objects and discuss alternative representations. This enables the application of hierarchical forecasting methods. Their evaluation of both existing and novel reconciliation methods reveals that hierarchical forecasting enhances the accuracy of aggregate curve predictions.

Abolghasemi, Tarr, and Bergmeir examine the case of retail forecasting, focusing on the challenge of integrating promotional information into models. Although hierarchical forecasting methods do not preclude the use of explanatory variables, there is limited work in this area. They employ machine learning to disaggregate aggregate forecasts using explanatory variable information from all levels, allowing for tailored reconstruction of forecasts. The study highlights how these methods vary in effectiveness between promotional and non-promotional periods, offering new insights into retail forecasting dynamics.

A historical motivation for hierarchical forecasting stemmed from the incoherence between sub-aggregate and aggregate forecasts. Modern hierarchical approaches address this problem by refining sub-aggregate forecasts using the entire hierarchy, and summing these up as required. Typically, univariate forecasting methods are used, with the abundant cross-series information being implicitly incorporated in the reconciliation step. However, multivariate forecasting methods can also be used, as shown by *Sbrana and Pelagatti* who use multivariate exponential smoothing and bottom-up forecasts. A key contribution of the paper is addressing the major challenge of model parameter estimation, as the problem's dimensionality increases. Their findings show that in many instances, their proposed model outperforms or competes effectively with the usual hierarchical forecasting methods.

Also using multivariate models, *Koop, McIntyre, Mitchell, and Poon* investigate the use of hierarchical aggregation constraints for nowcasting regional economic aggregates. They begin from a high dimensional mixed frequency vector autoregressive model, and explore various ways of incorporating aggregation constraints. These constraints, being stochastic, are incorporated as additional measurement equations in their state space model. Their evaluation, focusing on both point and probabilistic forecasts, concludes that the inclusion of stochastic aggregation constraints is beneficial, offering new insights into regional economic nowcasting.

Antoniadis, Gaucher, and Goude consider transfer learning in the context of electricity load forecasting, leveraging hierarchical information. They examine scenarios where finer-scale data can enhance more aggregate forecasts, exploring both stable and changing data distributions, with the latter situation presenting additional challenges for transfer learning. The proposed methods go beyond standard hierarchical forecasting by focusing on transfer learning, introducing a novel dimension to hierarchical forecasting that has not previously been addressed.

The contributions in this special issue demonstrate that the research area has matured markedly over the last decade. While it once focused exclusively on point predictions, nearly half of the contributions in this issue incorporate probabilistic approaches. We also see the seamless integration of both statistical and machine learning methods as part of the hierarchical machinery. Perhaps more fascinating for the field, are the contributions which pioneer new research directions or solve other forecasting challenges, drawing inspiration from the hierarchical forecasting literature. Undoubtedly, this evolution underscores the field's growth beyond a niche area, opening up various exciting future research directions.

References

Athanasopoulos, G., Ahmed, R. A. and Hyndman, R. J. (2009), 'Hierarchical forecasts for Australian domestic tourism', *International Journal of Forecasting* **25**(1), 146–166.

Athanasopoulos, G., Hyndman, R. J., Kourentzes, N. and Petropoulos, F. (2017), 'Forecasting with temporal hierarchies', *European Journal of Operational Research* **262**(1), 60–74.

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.

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.