Decision-Optimal Probabilistic Forecast Reconciliation for Expected Shortfall in Hierarchical Time Series

Vincent Su

1 Background

Many operational datasets are naturally hierarchical. For Starbucks, their total sales is the sum of all countries they are operating in, and each national sales is the sum for all cities, down to many outlets, forming an aggregation constraint. Forecasting such hierarchical time series arises in many decision-making contexts ranging from supply chains (Angam et al., 2025; Seaman & Bowman, 2022) and energy planning (Di Modica et al., 2021), to macroeconomic (El Gemayel et al., 2022; Li et al., 2019) and tourism analysis (Athanasopoulos et al., 2009), where stakeholders need forecasts at several aggregation levels to allocate resources and manage risk.

In practice, when base forecasts are produced independently for all series, they typically fail to satisfy the aggregation constraints. Such forecasts are *incoherent*, which can undermine downstream decisions that require internal consistency. Forecast reconciliation addresses this by adjusting the base forecasts so the final set is *coherent* with the hierarchy. Approaches such as OLS (Hyndman et al., 2011), WLS (Hyndman et al., 2016), and MinT (Wickramasuriya et al., 2019) implement this as linear mappings (often projections) so that the adjusted forecasts lie in the coherent subspace; see Athanasopoulos et al. (2024) for a comprehensive review.

2 Motivation

In the current literature, most reconciliation methodology has been developed for point forecasts, and its benefits for accuracy and coordination across organisation's departments or "silos" are well documented. However, in many applications the quantity of interest is probabilistic forecasts, not just a specific value. Decision makers care about the full distribution of future outcomes to quantify uncertainty, especially in risk management, where tail risks matter (e.g., extreme demand in newsvendor settings or energy generation shortfalls). This motivates moving from point reconciliation to probabilistic reconciliation, so that coherent forecasts also carry coherent uncertainty.

Existing reconciliation methods for point forecasts has recently been generalised to reconcile probabilistic forecasts. However, these approaches originally focus on point forecasts and are optimised for scoring rules such as mean squared error (MSE). There is a gap in the literature on how *optimal* this approach is with respect to the downstream decision. The current expansion of point reconciliation methods no longer guarantee optimality for probabilistic forecasts, such as quantiles or expected shortfall (ES). This leaves the question: *"How can probabilistic reconciliation be optimised with respect to different loss functions targeted at the downstream decision?"*

Although Panagiotelis et al. (2023) introduced an score optimisation for reconciliation weights using the energy score or variogram score, they are scoring rules for full multivariate predictive distributions and do not directly address the tail functionals. Among the risk measures, quantiles, also known as Value-at-Risk (VaR), is a common choice, but it has limitations. VaR ignores the severity of losses beyond that threshold and is not a coherent risk measure, since it does not satisfy the subadditivity property (Yamai & Yoshiba, 2005). This can lead to inconsistencies in risk assessments across different levels of the hierarchy.

Meanwhile, Expected Shortfall (ES) is a coherent risk measure that accounts for the severity of losses in the tail of the distribution (Yamai & Yoshiba, 2005). It has recently been adopted as the preferred risk measure in banking regulation since the Basel III regulatory framework (Basel Committee on Banking Supervision, 2019). It has ability to capture tail risks more effectively than VaR. ES is also applicable beyond finance, such as quantifying expected economic loss when electricity demand exceeds available capacity. Therefore, we aim to investigate and develop the optimal probabilistic forecast reconciliation under the Expected Shortfall criterion.

3 Methodology

Preliminaries

Let $\boldsymbol{y}_t \in \mathbb{R}^K$ denote a vector of K time series variables from the hierarchy, at time t. For simplicity, we suppress the subscript t and denote it as $\boldsymbol{y}$. These variables are subject to linear aggregation constraints represented by a matrix $\boldsymbol{C}$, such that $\boldsymbol{C}\boldsymbol{y} = 0$. For example, in a retail setting, the total sales at a warehouse (y_1) and two stores (y_2 and y_3) might satisfy the constraint $y_1 = y_2 + y_3$, leading to $\boldsymbol{C} = (1, -1, -1)$.

Base forecasts $\hat{\boldsymbol{y}}$, which we generate independently for each series, are typically *incoherent* (i.e. $\boldsymbol{C}\hat{\boldsymbol{y}} \neq 0$), while reconciled forecasts $\tilde{\boldsymbol{y}}$ are coherent by construction (i.e. $\boldsymbol{C}\hat{\boldsymbol{y}} = 0$). We obtain coherence via a linear mapping $\tilde{\boldsymbol{y}} = \boldsymbol{P}\hat{\boldsymbol{y}}$, where $\boldsymbol{P}$ maps the full *incoherent* space $\mathbb{R}^K$ to the linear *coherent* subspace $\mathbb{S} \subset \mathbb{R}^K$, where aggregation constraints hold. This mapping $\boldsymbol{P}$ can be a projection operator $\boldsymbol{S}(\boldsymbol{S}'\boldsymbol{S})^{-1}\boldsymbol{S}'$ (Hyndman et al., 2011), where $\boldsymbol{S}$ is a matrix designed such that $\boldsymbol{S}\boldsymbol{C} = 0$. The mapping $\boldsymbol{P}$ is not unique, and selecting an optimal $\boldsymbol{P}$ is the core design problem. In this project, we will focus solely on linear reconciliation.

Talk about the probabilistic reconciliation and how it is extended from point reconciliation.

Optimisation Framework

For a random variable Y, the Expected Shortfall (ES) at level $\alpha \in (0,1)$ is defined as $\mathrm{ES}_\alpha(Y) = E[Y \mid Y < \mathrm{VaR}_\alpha(Y)]$, where $\mathrm{VaR}_\alpha(Y)$ is the quantile q at level α. The ES alone is not elicitable, meaning there is no scoring rule or loss function that can uniquely identify the ES as the optimal optimiser. However, Fissler et al. (2015) showed that the quantile function and ES are jointly elicitable, minimising the Fissler-Ziegel class of loss functions in expectation. One of which takes the form:

$$L_\alpha^{\mathrm{FZ}}(y,q,e) = (\mathbb{I}(y \le q) - \alpha)(q-y) + \frac{1}{\alpha}\exp(e)\mathbb{I}(y \le q)(q-y) + \exp(e)(e-q-1),$$

where y is the realisation of Y, q is the quantile, and e is the ES. This loss function is non-convex, thus greatly complicating the optimisation problem. Alternatively, we propose to use the Rockafellar-Uryasev joint characterisation of the quantile and ES (Rockafellar & Uryasev, 2000): $\widetilde{\mathrm{ES}}_k(\boldsymbol{P}) = \max_q \, E_{\tilde{\boldsymbol{y}}}[H_\alpha(\tilde{y}_k(\boldsymbol{P}),q)]$ and $\tilde{q}_k(\boldsymbol{P}) = \arg\max_q \, E_{\tilde{\boldsymbol{y}}}[H_\alpha(\tilde{y}_k(\boldsymbol{P}),q)]$, where $H_\alpha(y,q) = q - \frac{1}{\alpha}\max\{0, y-q\}$. Thus, we jointly reconcile the quantile and ES using the following bilevel optimisation problem:

$$\begin{aligned}
&\min_{\boldsymbol{P}} \sum_{k=1}^{K} E_{\boldsymbol{y}}\Big[L_\alpha^{\mathrm{FZ}}(y_k,\, \tilde{q}_k(\boldsymbol{P}),\, \widetilde{\mathrm{ES}}_k(\boldsymbol{P}))\Big] && (1)\\
&\text{s.t. } \tilde{q}_k(\boldsymbol{P}) = \arg\max_q \, E_{\tilde{\boldsymbol{y}}}\left[H_\alpha(\tilde{y}_k(\boldsymbol{P}),q)\right],\\
&\qquad \widetilde{\mathrm{ES}}_k(\boldsymbol{P}) = \max_q \, E_{\tilde{\boldsymbol{y}}}\left[H_\alpha(\tilde{y}_k(\boldsymbol{P}),q)\right], \qquad \forall\, k = 1,\dots,K.
\end{aligned}$$

This ensures that the lower-level problem is still convex, however the upper-level objective function with the Fissler-Ziegler loss is non-convex. Irrespective of this non-convexity, smoothing techniques combined with the implicit function and envelope theorems can be used to devise gradient-based methods for locally minimising.

4 Applications

5 Contributions

References

Angam, B., Beretta, A., De Poorter, E., Duvinage, M., & Peralta, D. (2025). Forecast reconciliation for vaccine supply chain optimization. In *Communications in computer and information science* (pp. 101–118). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-74650-5/_6

Athanasopoulos, G., Ahmed, R. A., & Hyndman, R. J. (2009). Hierarchical forecasts for australian domestic tourism. *International Journal of Forecasting*, *25*(1), 146–166. https://doi.org/10.1016/j.ijforecast.2008.07.004

Athanasopoulos, G., Hyndman, R. J., Kourentzes, N., & Panagiotelis, A. (2024). *Forecast reconciliation: A review*. *40*(2), 430–456. https://www.sciencedirect.com/science/article/pii/S0169207023001097

Basel Committee on Banking Supervision. (2019). *Explanatory note on the minimum capital requirements for market risk*. https://www.bis.org/bcbs/publ/d457_note.pdf

Di Modica, C., Pinson, P., & Ben Taieb, S. (2021). Online forecast reconciliation in wind power prediction. *Electric Power Systems Research*, *190*(106637), 106637. https://doi.org/10.1016/j.epsr.2020.106637

El Gemayel, J., Lafarguette, R., Itd, K. M., et al. (2022). *United arab emirates: Technical assistance reportliquidity management and forecasting*.

Fissler, T., Ziegel, J. F., & Gneiting, T. (2015). Expected shortfall is jointly elicitable with value at risk - implications for backtesting. *arXiv [q-Fin.RM]*. http://arxiv.org/abs/1507.00244

Hyndman, R. J., Ahmed, R. A., Athanasopoulos, G., & Shang, H. L. (2011). Optimal combination forecasts for hierarchical time series. *Computational Statistics & Data Analysis*, *55*(9), 2579–2589. https://doi.org/10.1016/j.csda.2011.03.006

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

Li, H., Li, H., Lu, Y., & Panagiotelis, A. (2019). A forecast reconciliation approach to cause-of-death mortality modeling. *Insurance, Mathematics & Economics*, *86*, 122–133. https://doi.org/10.1016/j.insmatheco.2019.02.011

Panagiotelis, A., Gamakumara, P., Athanasopoulos, G., & Hyndman, R. J. (2023). Probabilistic forecast reconciliation: Properties, evaluation and score optimisation. *European Journal of Operational Research*, *306*(2), 693–706. https://doi.org/10.1016/j.ejor.2022.07.040

Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, *2*, 21–42. https://www.risknet.de/fileadmin/eLibrary/Rockafellar-Conditional-Value-at-Risk.pdf

Seaman, B., & Bowman, J. (2022). Applicability of the M5 to forecasting at walmart. *International Journal of Forecasting*, *38*(4), 1468–1472. https://doi.org/10.1016/j.ijforecast.2021.06.002

Wickramasuriya, S. L., Athanasopoulos, G., & Hyndman, R. J. (2019). Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization. *Jour-*

nal of the American Statistical Association, *114*(526), 804–819. https://doi.org/10.1080/01621459.2018.1448825

Yamai, Y., & Yoshiba, T. (2005). Value-at-risk versus expected shortfall: A practical perspective. *Journal of Banking & Finance*, *29*(4), 997–1015. https://doi.org/10.1016/j.jbankfin.2004.08.010