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

Vincent Su

1 Background

Risk managers and regulators use quantitative risk measures to quantify and manage potential losses in uncertain environments. Two familiar choices are Value-at-Risk (VaR)–a loss threshold exceeded with some probability–and Expected Shortfall (ES; also known as Conditional VaR)–the average loss beyond that threshold.

These tail risks are, however, not managed in isolation. Consider a hedge fund's investment portfolio returns, which compose of returns of different asset classes, and each asset class further contain returns of various individual securities. This is an example of a hierarchical structure, with multiple *aggregation levels*. Risk managers need to assess and forecast the returns or risk quantities at several levels, from the overall portfolio, down to individual securities. Ignoring this naturally formed hierarchical structure can lead to misaligned forecasts, for example, the predicted returns of individual securities are low while the overall portfolio returns are high. This results in misallocation of capital and suboptimal risk management strategies.

Forecasting data from such hierarchical structures also arises in many other decision-making contexts, from supply chains (Angam et al., 2025; Seaman & Bowman, 2022) and energy planning (Di Modica et al., 2021), to macroeconomics (El Gemayel et al., 2022; Li et al., 2019) and tourism analysis (Athanasopoulos et al., 2009). Stakeholders in these settings need forecasts at several aggregation levels to allocate resources and manage risk. These hierarchical structures naturally form *aggregation constraints* (child levels must sum up to parents). Taken together, a collection of time series organised in a hierarchy and subject to aggregation constraints is refered as *hierarchical time series*.

In practice, when forecasts are produced for all series (often called *base forecasts*), they typically violate the aggregation constraints observed in the data; such forecasts are *incoherent*. This can undermine downstream decisions that require internal consistency. Forecast reconciliation, a post-processing step, addresses this by adjusting the base forecasts to the final adjusted set of forecasts that are *coherent* with the aggregation structure. Reconciliation approaches such as OLS (Hyndman et al., 2011), WLS (Hyndman et al., 2016), MinT (Wickramasuriya et al., 2019), and more are developed and commonly used in practice. As well as guaranteeing coherence, reconciliation methods have been shown to improve forecast accuracy (Athanasopoulos et al., 2024). For instance, the International Monetary Fund reported improvements in their liquidity forecasts, generating economic benefits via more efficient monetary policy operations (El Gemayel et al., 2022).

2 Motivation

Modern risk management and operations need more than a typical mean or median forecast. As many decisions are tail-sensitive (e.g., extreme demand in retail settings or energy generation shortfalls), they need quantities that capture uncertainty and tail behaviour. While VaR is conceptually straightforward, it ignores the severity of losses beyond the VaR threshold and does not satisfy the subadditivity property[1] (Yamai & Yoshiba, 2005). Violating subadditivity is a critical problem, because it means that the risk measure of a combined portfolio can exceed the sum of stand-alone risks, which contradicts the principle of diversification. VaR therefore can mislead risk assessments across different levels of the hierarchy, and may underestimate the true risk exposure under market stress. ES, on the other hand, is a consistent metric that captures the average loss in the tail beyond the VaR threshold. This makes ES a more reliable measure for assessing and managing "tail risks", and it has been adopted as the preferred regulatory metric since the Basel III framework (Basel Committee on Banking Supervision, 2019). Although ES is the metric that risk managers actually use, there is a lack of research on forecasting ES within the reconciliation framework.

Most reconciliation methodology has been developed for mean forecasts and are optimised for point scoring functions such as mean squared error (MSE). Generalising these methods to reconcile probabilistic forecasts does not guarantee optimality for tail-risk measures such as VaR or ES. Recent works from Panagiotelis et al. (2023) that introduced score optimisation for reconciliation weights using proper scoring rules (energy score and variogram score) has shred lights on reconciling full multivariate predictive distributions. Building on this, I aim to construct methods that directly address the tail functionals. The challenge will be to *"construct forecast reconciliation methods that are explicitly optimised for tail-risk measures such as ES."*

A practical challenge is that ES is not elicitable on its own, meaning there is no scoring rule or loss function that can uniquely identify ES as the optimiser. However, ES is jointly elicitable with VaR (Fissler et al., 2015), and is amenable to convex optimisation via an auxiliary VaR variable (Rockafellar & Uryasev, 2000). This opens the door to train reconciliation specifically for ES.

As a result, the objective my PhD research project is to bring these theoretical advances into the reconciliation framework and develop methods that produce coherent probabilistic forecasts optimised for ES. The goal is a reliable, consistent, and decision-aligned risk-forecasting framework that helps national and international organisations improve ES backtesting, stabilise capital, and allocate resources more effectively across hierarchical levels.

[1] **Subadditivity.** A risk measure ρ on loss variables is subadditive if $\rho(X+Y) \leq \rho(X) + \rho(Y)$ for any losses X, Y. Expected Shortfall satisfies subadditivity for all $\alpha \in (0,1)$, whereas VaR can violate it. Thus, the portfolio ES cannot exceed the sums of individual security ESs.

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. Base forecasts $\hat{\boldsymbol{y}}$, which we generate for each series, are typically *incoherent*, while reconciled forecasts $\tilde{\boldsymbol{y}}$ are coherent by construction. 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. Ultimately, selecting an optimal $\boldsymbol{P}$ is the core design problem. In this project, we will focus solely on linear reconciliation.

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. Hereafter α is dropped for convenience. Thus, we can define $\widetilde{\mathrm{ES}}_k(\boldsymbol{P}) = \mathrm{ES}(\tilde{y}_k(\boldsymbol{P}))$ as the ES of the k-th series after reconciliation, where $\tilde{y}_k(\boldsymbol{P})$ is the k-th element of the reconciled forecasts $\tilde{\boldsymbol{y}}(\boldsymbol{P}) = \boldsymbol{P}\hat{\boldsymbol{y}}$; and similarly define $\tilde{q}_k(\boldsymbol{P}) = \mathrm{VaR}(\tilde{y}_k(\boldsymbol{P}))$ as the quantile of the k-th series after reconciliation.

Optimisation Framework

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:

$$L^{(\alpha)}(y,q,e) = (\mathbb{I}(y \leq q) - \alpha)(q-y) + \frac{1}{\alpha}\exp(e)\mathbb{I}(y \leq q)(q-y) + \exp(e)(e-q-1),$$

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): $\widetilde{\mathrm{ES}}_k(\boldsymbol{P}) = \max_q E_{\tilde{\boldsymbol{y}}}[H(\tilde{y}_k(\boldsymbol{P}), q)]$ and $\tilde{q}_k(\boldsymbol{P}) = \arg\max_q E_{\tilde{\boldsymbol{y}}}[H(\tilde{y}_k(\boldsymbol{P}), q)]$, where $H(y,q) = H^{(\alpha)}(y,q) = q - \frac{1}{\alpha}\max\{0, y-q\}$. Thus, we jointly reconcile the quantile and ES using the following bilevel optimisation problem:

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

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.

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