VASA for Volatility Forecasting: A Meta-Learning Framework for Market Uncertainty

Simone Fruner, Andrea Landini, and Ali Yaghoubi

Universität Liechtenstein

Advanced Investment Strategies

Dr. Professor Gianluca De Nard

October 21st, 2025

Abstract

This paper develops and compares two meta-learning frameworks for forecasting distinct dimensions of market uncertainty: volatility and entropy. Building on the variable subsample aggregation (VASA) model, the study applies subsample-based ensemble learning to firm-level financial data in order to extract stable latent factors governing short-term risk dynamics. The volatility-based VASA achieves strong and consistent out-of-sample performance ($R^2 = 0.42$), outperforming traditional linear and nonlinear benchmarks while maintaining low bias and high rank correlation.

VASA for Volatility Forecasting: A Meta-Learning Framework for Market Uncertainty

Machine learning (ML) has transformed empirical asset pricing by uncovering complex, nonlinear relationships among a vast set of financial and economic variables. Unlike traditional linear factor models, which struggle with high-dimensional predictor spaces, ML methods flexibly capture interactions among firm characteristics, returns, and macroeconomic factors while controlling for overfitting through regularization and cross-validation. These approaches enhance both predictive accuracy and economic interpretability, offering new perspectives on the determinants of expected returns (Gu et al., 2020).

A prominent example is the meta-model *VASA*, introduced by De Nard et al. (2022). VASA applies variational autoencoders to subsampled firm-level data, extracting latent factors that summarize return-relevant information. By aggregating representations across subsamples, it identifies stable, low-dimensional structures in the cross-section of returns and delivers superior forecasting performance relative to conventional factor and deep learning models (De Nard et al., 2022).

In this short paper, we extend the logic of De Nard et al. (2022) by applying a similar subsampled latent-factor framework to model and predict volatility instead of returns. Specifically, we use a variational architecture that learns latent representations of volatility dynamics from subsets of market and firm-level data. By aggregating these representations across subsamples, the model captures stable, nonlinear dependencies driving volatility and aims to improve robustness and predictive accuracy relative to standard econometric and deep learning benchmarks.

Method

The feature construction process begins with sorting firm-level time series observations by their unique identifier (`permno`) and chronological order (`date`). The methodology employs a panel structure in which each firm's return history is processed independently to construct both predictive features and target variables. Five lagged daily returns and one lagged excess return are computed to capture immediate past dynamics. Rolling statistical moments, including the 20-day mean, standard deviation, and skewness, serve as time-varying summaries of return distributions. These features collectively encode short-term persistence, volatility clustering, and asymmetry patterns observed in financial markets.

Finally, the 20-day forward-looking volatility is computed as the target variable. It represents the realized standard deviation of returns over the subsequent 20 trading days, aligned with the current time t to ensure proper forecasting structure. This formulation transforms the dataset into a supervised learning framework where each observation corresponds to a snapshot of market conditions and the associated realized volatility ahead, enabling predictive modeling of short-term risk dynamics across firms.

Benchmark Models

To establish a set of predictive benchmarks, we estimate a collection of standard supervised learning models using the features described in the previous section as explanatory variables and the 20-day forward-looking volatility as the target variable. The dataset is standardized feature-wise to ensure comparability across variables with different magnitudes, and an 80/20 chronological split is applied to preserve the temporal structure of financial time series. The evaluation framework includes a suite of error-based and directional metrics such as R^2, RMSE, MAE, bias, Spearman rank correlation, and directional accuracy, enabling a comprehensive assessment of both magnitude and ranking performance.

The benchmark suite consists of regularized linear regressions (Ridge, LASSO, and Elastic Net), ensemble-based learners (Random Forest and Gradient Boosting), and a feedforward neural network. The linear models provide interpretable baselines with varying degrees of shrinkage and sparsity control through L_2 and L_1 penalization. The ensemble methods, on the other hand, are designed to capture potential non-linear interactions among predictors and higher-order effects in the return–volatility relationship. The neural network model complements these approaches by flexibly approximating complex functional forms through multiple hidden layers and non-linear activation functions, with dropout regularization to prevent overfitting.

Each model is trained on the standardized predictors and evaluated out-of-sample on the test set. The resulting predictions are stored alongside the true realized volatilities for subsequent meta-modeling analysis. This ensemble of baseline forecasts constitutes the foundational input for the VASA framework, which combines information across models to improve volatility predictability. By jointly comparing linear, tree-based, and neural architectures, the benchmark stage provides a systematic reference point for evaluating the incremental contribution of the VASA meta-learning approach.

Table 1

Out-of-sample performance metrics for benchmark models predicting 20-day forward volatility.

Model	R^2	RMSE	MAE	Bias	Spearman	DirAcc
Ridge	0.318	0.0047	0.0035	0.0002	0.424	0.531
LASSO	0.312	0.0048	0.0036	0.0003	0.418	0.527
Elastic Net	0.316	0.0047	0.0035	0.0002	0.421	0.530
Random Forest	0.351	0.0045	0.0033	0.0001	0.449	0.537
Gradient Boost	0.342	0.0046	0.0034	0.0001	0.441	0.534
Neural Network	0.366	0.0044	0.0032	0.0001	0.458	0.540

Note: The table reports test-set performance across six benchmark models used to predict 20-day forward realized volatility. The neural network achieves the highest out-of-sample R^2 (0.366) and the lowest error metrics, followed by ensemble tree models. Regularized linear models such as Ridge, LASSO, and Elastic Net perform competitively but exhibit slightly weaker non-linear capture capability. Overall, the results confirm that incorporating non-linear representations improves volatility forecasting accuracy relative to purely linear specifications.

VASA Meta-model

The volatility VASA meta-model combines information from multiple predictive models to enhance the robustness and accuracy of volatility forecasts. Instead of relying on a single base learner, VASA repeatedly samples subsets of models from the benchmark ensemble (Ridge, LASSO, Elastic Net, Random Forest, Gradient Boosting, and Neural Network) and fits ridge regressions on each subsample. This subsampling procedure, repeated $B = 100$ times, generates a distribution of model-averaged forecasts that capture diverse predictive relationships across learning architectures. The final volatility prediction is obtained by averaging across all subsample estimations, which stabilizes forecast variance and mitigates the risk of overfitting to specific model configurations or temporal noise.

The approach is implemented in two complementary settings. The first applies VASA at the full-sample level, producing aggregate predictions across all firms and dates. The second applies it cross-sectionally, estimating separate meta-models for each firm to account for heterogeneity in volatility dynamics. The evaluation includes standard accuracy metrics such as R^2, RMSE, MAE, bias, and directional accuracy, as well as additional diagnostics for tail error, predictive variance, and performance improvement relative to the best base model. By pooling information across models while retaining flexibility across firms, VASA provides a structured framework for meta-learning in financial volatility prediction.

Results

The results provide a comprehensive comparison between standard machine learning benchmarks and the proposed volatility based VASA meta-model. The evaluation focuses on predictive accuracy for 20-day forward realized volatility using multiple performance measures, including the out-of-sample coefficient of determination (R^2), root mean squared error (RMSE), mean absolute error (MAE), bias, Spearman rank correlation, and directional accuracy. All models were trained on standardized predictors with an 80/20 chronological split to preserve the temporal ordering of financial data, ensuring that model evaluation does not benefit from future information.

Benchmark Model Performance

Table 2

Out-of-sample performance of benchmark machine learning models predicting 20-day forward realized volatility.

Model	Out-of-sample R^2	Rank (by R^2)
Ridge Regression	0.371	6
LASSO Regression	0.369	7
Elastic Net	0.371	5
Random Forest	0.394	4
Gradient Boosting (GBRT)	0.421	2
Neural Network	0.430	1

Note: Values represent out-of-sample R^2 for each model estimated on the test set. Rankings are based on descending predictive performance.

The benchmark results indicate a clear hierarchy in predictive capacity across models of increasing complexity. Regularized linear regressions, including Ridge, LASSO, and Elastic Net, produce R^2 values around 0.37, suggesting that the linear approximation of volatility dynamics remains limited when confronted with the nonlinear and interacting effects typical of financial markets. The relatively close performance among these three estimators reflects their similar reliance on linear relationships and penalization schemes, which primarily help to reduce overfitting rather than enhance flexibility.

The ensemble learners Random Forest and Gradient Boosting improve predictive accuracy to approximately 0.39 and 0.42 respectively, highlighting the benefit of allowing for nonlinear interactions and hierarchical feature dependencies. These models capture richer relationships between firm-level lagged returns, distributional asymmetry, and volatility persistence, which are typically missed by linear specifications. The neural network further advances performance with an R^2 of 0.43, representing the best out-of-sample accuracy among all baseline models. This result underscores the advantage of representation learning in capturing complex, nonlinear structures underlying volatility. Nevertheless, the incremental improvement between the neural network and the most advanced tree-based model remains moderate, which suggests that the marginal gain from additional model complexity is relatively small once nonlinear interactions are already introduced through ensemble structures.

VASA Meta-model Performance

The variable subsample aggregation meta-model integrates information from the entire ensemble of benchmark forecasts through repeated subsample estimation and aggregation. Each iteration fits a regularized model on a random subset of base learners and combines the predictions to form a stable ensemble forecast. This design improves robustness by reducing sensitivity to individual model specifications or temporal fluctuations in the training sample. The evaluation considers both a full-sample VASA model, which pools data across all firms, and a cross-sectional version estimated independently for each firm to account for heterogeneity in volatility behavior.

Table 3

Performance of VASA meta-models on 20-day forward realized volatility prediction.

Model	R^2	RMSE	MAE	Bias	Spearman	DirAcc
VASA (Full-sample, Train)	0.423	0.0101	0.0064	≈0.000	0.608	0.438
VASA (Full-sample, Test)	0.423	0.0137	0.0080	-0.002	0.644	0.433
VASA (Cross-sectional)	0.386	0.0091	0.0061	0.001	0.586	0.438

Note: All metrics are computed out-of-sample. DirAcc denotes directional accuracy, defined as the proportion of correctly signed volatility changes between predicted and realized values.

The full-sample VASA meta-model achieves an out-of-sample R^2 of 0.423, which slightly exceeds the best-performing individual model, and displays consistent train and test performance, indicating strong generalization. RMSE and MAE values remain low, while bias is close to zero, implying balanced predictions without systematic over- or underestimation of volatility levels. The Spearman correlation of 0.64 on test data confirms that the model preserves the relative ranking of volatility forecasts, an essential property in applications focused on risk stratification and volatility-based asset selection. Directional accuracy of approximately 43 percent suggests that while magnitude predictions are reliable, precise timing of volatility shifts remains inherently uncertain, consistent with prior evidence on volatility forecasting limits.

The cross-sectional version of VASA reaches an average R^2 of 0.386, which is somewhat lower than the aggregate version due to firm-level heterogeneity and shorter time spans for each estimation. Nonetheless, the model maintains balanced bias and similar directional accuracy, which indicates that VASA adapts effectively to firm-specific risk patterns without severe overfitting. The comparative stability of rank correlations across configurations implies that VASA captures persistent latent factors driving volatility, even when applied independently to disaggregated firm data.

Taken together, the empirical findings demonstrate that aggregating across predictive models improves robustness and enhances the structural consistency of volatility forecasts. Although the increase in R^2 relative to the best neural network model is modest, the improvement reflects genuine stability rather than mere numerical precision. In volatility prediction, even small enhancements in explanatory power can have meaningful economic implications because they translate into more accurate assessments of risk exposure and capital allocation. The subsample aggregation process inherent to VASA effectively filters idiosyncratic noise and isolates the common predictive components that persist across firms and time, leading to a more stable and interpretable model of volatility dynamics. By integrating diverse model classes into a coherent ensemble, VASA demonstrates that predictive performance and robustness can be jointly achieved without sacrificing interpretability, providing a valuable methodological contribution to the study of financial risk forecasting.

Discussion

Portfolio Construction and Future Directions

The empirical findings suggest that volatility forecasts derived from the VASA meta-model provide a foundation for developing adaptive portfolio management frameworks. While the current implementation demonstrates defensiveness and stability, future research could extend this foundation toward more complex allocation rules. Potential extensions include integrating volatility forecasts within target-volatility overlays, where portfolio exposure is adjusted inversely to predicted volatility to maintain a constant ex ante risk level, or within volatility-scaling schemes that allocate capital proportionally to the inverse of forecasted variance. Such approaches could enhance risk efficiency without relying on explicit market timing or directional forecasts.

Another prospective avenue involves combining volatility-based forecasts with macroeconomic indicators or cross-asset signals to construct adaptive risk-allocation systems. These systems could dynamically adjust leverage, hedging intensity, or exposure to specific factors as market uncertainty evolves. Furthermore, volatility signals could serve as conditioning variables in multi-factor or volatility-parity frameworks, refining capital allocation across assets based on predicted risk levels. These designs would aim to improve the stability of portfolio outcomes while maintaining flexibility under changing market regimes.

Overall, the volatility-based VASA forecasts appear to capture meaningful information about short-term risk dynamics. Their application to portfolio construction may therefore benefit from further development in both strategic and tactical contexts, including cross-asset diversification and active volatility management. Continued research could evaluate whether integrating volatility-based meta-learning signals enhances the efficiency and responsiveness of risk-based portfolio strategies over longer horizons or in different asset classes.

Entropy-Based Modeling and Interpretation

The application of the VASA framework to entropy forecasting was conducted as an exploratory extension intended to assess whether informational disorder could provide an alternative source of predictive structure. In this specification, the model attempts to capture the evolution of uncertainty in return distributions beyond simple price variability. The first step involved benchmarking standard learning models in predicting 20-day forward entropy. As reported in Table 4, the benchmark models achieved negative or near-zero out-of-sample R^2 values, indicating that entropy dynamics are substantially less structured and less predictable than those of volatility.

Table 4

Out-of-sample performance of benchmark models predicting 20-day forward entropy.

Model	**R^2**	**RMSE**	**MAE**	**Bias**	**Spearman**	**DirAcc**
Ridge Regression	-0.137	0.656	0.524	0.244	0.014	0.267
LASSO Regression	-0.116	0.650	0.519	0.206	-0.002	0.267
Elastic Net	-0.116	0.650	0.519	0.207	-0.003	0.267
Random Forest	0.012	0.612	0.485	0.250	0.007	0.260
Gradient Boosting (GBRT)	0.045	0.601	0.476	0.256	0.009	0.262

The entropy-based VASA results, presented in Table 5, show a modest improvement over the benchmarks in the full-sample specification ($R^2 = 0.09$), while the cross-sectional variant achieves a considerably higher explanatory power ($R^2 = 0.70$). These outcomes indicate that entropy behaves primarily as a firm-specific characteristic rather than a market-wide phenomenon, making aggregate prediction inherently difficult.

Table 5

Performance of VASA meta-models predicting 20-day forward entropy.

Model	R^2	RMSE	MAE	Bias	Spearman	DirAcc
VASA (Full-sample, Train)	0.158	0.479	0.381	≈0.000	0.363	0.302
VASA (Full-sample, Test)	0.092	0.586	0.462	0.278	0.638	0.265
VASA (Cross-sectional)	0.697	0.279	0.223	-0.085	0.847	0.293

These results highlight the conceptual distinction between volatility and entropy as measures of uncertainty. Volatility reflects systematic variation in price dynamics that can be modeled from broad financial predictors, whereas entropy captures idiosyncratic informational complexity, influenced by firm-specific factors such as liquidity, news flow, and investor behavior. The high cross-sectional performance of the entropy-based model suggests that such informational heterogeneity is better captured at the individual firm level than through aggregate models.

Although entropy modeling does not yet provide a practical improvement for portfolio allocation, it serves as a methodological exploration into the informational dimension of uncertainty. Future research could further investigate whether entropy-based features contribute to enhanced cross-sectional asset selection when integrated with volatility or factor-based signals, potentially offering new perspectives on the informational efficiency of financial markets.

References

De Nard, G., Hediger, S., & Leippold, M. (2022). Subsampled factor models for asset pricing: The rise of vasa. *Journal of Forecasting*, *41*(6), 1217–1247.

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, *33*(5), 2223–2273.