25th International Congress on Modelling and Simulation, Darwin, NT, Australia, 9 to 14 July 2023
mssanz.org.au/modsim2023

Combining risk forecasts

S.J. Soltyk [a], F. Chan [a], A. Phatak [b] and E.L. Jackson [c]

[a] *School of Accounting, Economics and Finance, Curtin University, GPO BOX U1987, Perth, Western Australia*
[b] *ARC Centre for Transforming Maintenance through Data Science, School of EECMS, Curtin University, GPO BOX U1987, Perth, Western Australia*
[c] *School of Management and Marketing, Curtin University, GPO BOX U1987, Perth, Western Australia*
Email: sylvia.soltyk@postgrad.curtin.edu.au

Abstract: It is a common finding in the forecasting literature that combining forecasts of the mean typically results in improved forecast accuracy when compared to the forecasts from the individual models when measured by mean squared forecast error (MSFE). While the literature on forecast combination of the mean is extensive, it contains fewer studies on combining other aspects of the forecast distribution, particularly quantile forecast combination.

The aim of this study is to empirically investigate the performance of forecast combination of a quantile forecast known as Value-at-Risk (VaR), motivated by the fact that forecast accuracy is improved when combining forecasts of the mean. We empirically examined the forecast performance of combining VaR for the S&P500 and FTSE financial indices. To do so, we used four models: two GARCH models, a neural network, and a stochastic volatility model to forecast VaR. The forecast performance of these four individual models is compared to that of a combined VaR forecast combination comprised of all four models.

The results of this study indicate that VaR forecast combination may improve VaR forecast accuracy. We find that, overall, forecast combination performed better than almost all individual models for the 1, 2, and 5% quantiles for both indices. Unlike a continuous loss function such as MFSE or related loss functions which measure the distance between the forecast and the true observed value, however, we evaluated the forecast performance using the actual proportion of violations of VaR compared to the expected violations (the quantile), and ranked the forecast performance based on this. Therefore, while the results do show that the simple average forecast combination typically improves forecast accuracy based on proportion of violations of VaR, further research is necessary to investigate why, and under which conditions, combining forecasts of the quantile may improve forecast accuracy. Understanding the theoretical underpinning of quantile forecast combination has implications for risk management practices, as accurate risk forecasts are imperative for decision making.

Keywords: *Value-at-Risk, VaR, forecast combination*

1 INTRODUCTION

When risk forecasts from different models are available, it can be difficult to decide which forecast is most reliable and relevant for industry applications. An inaccurate forecast of risk can have serious consequences, as was evident in the global financial crisis of 2007-8, where the risk models underestimated risk. Overestimating risk can also have negative implications for capital holding requirements.

It is a common finding in the forecasting literature that combining forecasts of the mean results in improved forecast accuracy when compared to the forecasts from individual models as measured by a loss function such as mean squared forecast error (see Bates & Granger (1969) and Timmermann (2006)). While the literature on forecast combination of the mean is extensive, it contains fewer studies on combining other aspects of the forecast distribution, particularly quantile forecast combination (Taylor (2020)). Some of the limited empirical studies include Chan & James (2011), Halbleib & Pohlmeier (2012) Fuertes & Olmo (2013), Jeon & Taylor (2013), Huang & Lee (2013), McAleer et al. (2013*a*,*b*), Bayer (2018), and Taylor (2020).

The aim of this study is to empirically investigate the performance of forecast combination of a quantile forecast known as Value-at-Risk (VaR) using a simple average method, motivated by the fact that forecast accuracy is improved when combining forecasts of the mean, and the simple average often outperforms other weighting strategies (Chan et al. (2020)).

VaR is a method of estimating risk; it can be understood as the worst possible loss over a target horizon that does not exceed a given probability (Jorion (2006)). While VaR has its limitations, it is a widely used measure of risk and is used by, for example, the Basel Committee on Banking and Supervision (Basel Committee on Banking Supervision (2019)). There are many ways to forecast VaR. The literature does not indicate the superiority of one method over another (see Kuester et al. (2006); Boucher et al. (2014); Nieto & Ruiz (2016)). Thus, forecast combination may be a useful tool for improving VaR forecasts if it is found that combining forecasts of quantiles improve VaR forecast accuracy.

2 METHODOLOGY

All analysis was performed using the statistical programming language R (R Core Team (2022)).

2.1 Value-at-Risk

The VaR can be defined as the lower quantile of the returns distribution for a given probability. In this study, we compare forecasts of 1, 2, and 5% VaR from $m = 1, \ldots, M$ models to a combination of those forecasts. Consider the stationary time series $r_t : \{r_1, \ldots, r_T\}$ for $t = 1, \ldots, T$. The forecast of VaR from the m^{th} model can be written as

$$V\hat{a}R_{t+1,m} = \hat{r}_{t+1,m} - |\kappa|\hat{\sigma}_{t+1,m}, \tag{1}$$

where κ is the critical value associated with the chosen quantile (in our case, 1, 2, or 5%), and $\hat{r}_{t+1,m}$ and $\hat{\sigma}_{t+1,m}$ are the one-step ahead forecasts of the series mean and standard deviation and are estimated using the following four volatility models, which gives $M = 4$ models for each quantile and each financial index:

(i) ARMA-GJRGARCH where the standardised residuals from the model are used to estimate κ assuming the skew-t distribution which can accommodate skewness and/or kurtosis,

(ii) and in addition, as an alternative method to estimate κ under the ARMA-GJRGARCH framework (above, model (i)), we use a bootstrap with replacement to estimate the empirical quantile,

(iii) a neural network with two hidden layers where κ is estimated by using a bootstrap of the original return subtract the estimated means from the neural network and dividing by the estimated standard deviations also from the neural network, (Eq. (4))

(iv) AR(1) stochastic volatility model where κ is estimated using the average of the empirical quantiles of the MCMC draws.

A simple average (SA) is used to combine forecasts as it is often the best weighting strategy when combining mean forecasts (Chan et al. (2020)):

$$V\hat{a}R^{SA}_{t+1,m} = \frac{1}{m}\sum_{m=1}^{M} V\hat{a}R_{t+1,m}. \tag{2}$$

Forecasts from individual models and the forecast combination are ranked according to their performance by comparing the actual number of violations to the expected number of violations (1, 2, or 5%). The forecast which resulted in the proportion of violations closest (either above or below) to the expected number of violations was deemed the best (ranked number 1), and so forth.

2.2 The Models

Model (i) and Model (ii) ARMA-GJRGARCH with skew-t and bootstrap for κ: The GJRGARCH(p, q) model of Jagannathan et al. (1993) takes the general form

$$r_t = \mu_t + \epsilon_t, \quad \epsilon_t = \sigma_t z_t, \quad z_t \sim iid(0,1)$$
$$\sigma_t^2 = \omega + \sum_{i=1}^{q}\alpha_i\epsilon^2_{t-i} + \sum_{i=1}^{q}\phi_i\epsilon^2_{t-i}I_{t-1} + \sum_{i=1}^{p}\beta_i\sigma^2_{t-i}$$

where μ_t and σ_t are the conditional mean (estimated as an ARMA process in this paper) and conditional standard deviation of r_t, respectively, ϵ_t are the residuals (errors), $(\omega, \alpha, \phi, \beta)$ are the parameters of the conditional volatility, σ_t^2 to be estimated, where $I_{t-1} = 0$ if $\epsilon_{t-1} \geq 0$ and $I_{t-1} = 1$ if $\epsilon_{t-1} < 0$. z_t are the standardized residual (shock) which are independently and identically distributed (iid) with zero mean and unit variance. In this paper, we specify a GARCH(1,1) and assume a normal distribution for z_t for the maximum likelihood estimation process, but to obtain values for κ, we fit a skew-t to the standardized residuals to estimate the quantiles (model i), in addition to we also use a bootstrap with replacement to estimate the empirical quantiles of the standardized residuals (model ii). The ARMA order is selected using the *auto.arima* function in the *forecast* package (Hyndman et al. (2023)), and the GJRGARCH is fitted using the *rugarch* package (Ghalanos (2022)) both available in R.

Model (iii) Neural Network: Three separate neural networks are fitted, one for the conditional mean, μ_t conditional variance, σ_t, and estimation of κ. The conditional mean is estimated as a lag of itself, $\hat{\mu}_t = E(r_t|r_{t-1})$, and the conditional variance can be estimated similarly using the second (raw) moment, $E(r_t^2|r_{t-1}^2)$, and by using

$$\hat{\sigma}_t^2 = E(r_t^2|r_{t-1}^2) - E(r_t|r_{t-1})^2. \tag{3}$$

The standardized residuals can be estimated using

$$\hat{z}_t = \frac{r_t - \hat{\mu}_t}{\hat{\sigma_t}} \tag{4}$$

and a bootstrap with replacement can be used to estimate κ. Each neural network has two hidden layers, is estimated using a resilient backpropagation algortithm, with a sigmoid activation function. The neural network was fitted using the *neuralnet* package in R.

Model (iv) Stochastic Volatility: In this study, we implement the stochastic volatility model of Kastner & Frühwirth-Schnatter (2014). The mean model is estimated as an $AR(1)$ process, and each residual from the $AR(1)$ process, ϵ_t, is assumed to have its own contemporaneous variance, $e^{\sigma_t^2}$, where the volatility is estimated as follows

$$\epsilon|\sigma_t^2 \sim N(0, \exp(\sigma_t^2)), \quad \sigma_t^2|\sigma_{t-1}^2 \sim N(\mu + \phi(\sigma_{t-1}^2 - \mu), s_\eta^2), \quad \sigma_0^2 \sim N(\mu, s_\eta^2/(1-\phi^2)) \tag{5}$$

with $\theta = (\mu, \phi, s_\eta)^T$ as the vector of parameters: level of log-variance μ, the persistence of log-variance ϕ, and the volatility of log-variance s_η. A prior distribution for the parameter vector θ must be specified, with individual components, that is $p(\theta) = p(\mu)p(\phi)p(s_\eta)^1$. The initial state, σ_0^2 is distributed as the stationary distribution of an $AR(1)$ process. An MCMC algorithm is implemented to calculate draws from the posterior distribution. κ is estimated using the average of the empirical quantiles of the MCMC draws. The stochastic volatility model was fitted using the *stochvol* package in R (Kastner (2016)).

3 DATA

The data used in this study are the S&P500 (US) and FTSE (UK) stock indices, where the daily log returns are calculated using the expression $r_t = 100 \log(p_t/p_{t-1})$, where p_t is the price on day t, for the period 01/01/1990 – 18/08/2022 ($T = 8486$) obtained from Thompson Reuters DataStream. We split the data into the training and test sets, with an out-of-sample (test set) length of the 2000 most recent observations, leaving $T = 6486$ observations for the training set. Figure 1 shows the index values (top) and returns (bottom) for the S&P 500 and FTSE.

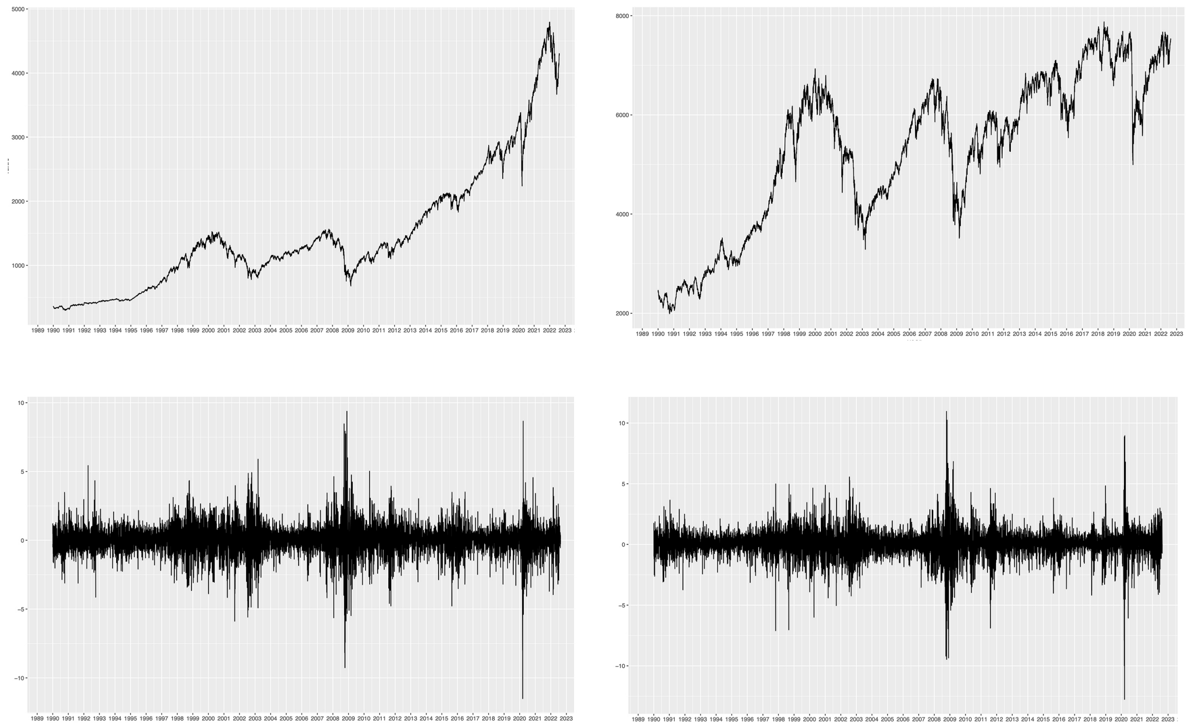

Figure 1. Daily S&P 500 (top left) and FTSE (top right) index value and returns (bottom left and right).

4 EMPIRICAL RESULTS

The results in Table 2 show that the stochastic volatility performed relatively poorly compared to the other models, while the neural network performed quite well, as it correctly estimated risk for 2% and 1% quantiles for the s&P500 and FTSE, respectively. Both the GARCH skew-t and stochastic volatility models underestimated risk for all quantiles of both indices. The GARCH with bootstrap underestimated risk across all quantiles for both indices with the exception of the S&P500 5% quantile where risk was overestimated. The neural network either correctly or overestimated risk. While the forecast combination either correctly estimated risk, or underestimated risk for both financial indices.

Table 2 shows the number of violations and proportion of violations (in brackets) of the VaR, and Table 2 shows the rank for each model and the forecast combination for both the S&P500 and FTSE financial indices

[1]Further details on specification of prior distributions are available upon request

for all quantiles, 1, 2, and 5%. We find that, overall, forecast combination performed well, ranking a mixture of first and second for both indices. Specifically, for the S&P 500, forecast combination ranked first, second and first, and for the FTSE, forecast combination ranked second, first, and first, for the 1, 2, and 5% quantiles, respectively.

	S&P500			FTSE		
Model	1%	2%	5%	1%	2%	5%
GJRGARCH skew-t	25 (0.0125)	49 (0.0245)	101 (0.0505)	28 (0.014)	50 (0.025)	113 (0.0565)
GJRGARCH bootstrap	27 (0.0135)	48 (0.024)	83 (0.0415)	31 (0.0155)	49 (0.0245)	103 (0.0515)
Neural Network	19 (0.0095)	40 (0.02)	99 (0.0495)	20 (0.01)	38 (0.019)	99 (0.0495)
Stochastic Volatility	46 (0.023)	73 (0.0365)	144 (0.072)	32 (0.016)	59 (0.0295)	105 (0.0525)
Forecast Combination	20 (0.01)	41 (0.0205)	101 (0.0505)	26 (0.013)	42 (0.021)	100 (0.05)

Table 1. Number and proportion (in brackets) of violations of VaR for each model and forecast combination for the S&P 500 (left) and FTSE (right) for the 1, 2, and 5% quantiles.

	S&P500			FTSE		
Model	1%	2%	5%	1%	2%	5%
GJRGARCH skew-t	3	4	1	3	3	5
GJRGARCH bootstrap	4	3	2	5	2	3
Neural Network	2	1	1	1	1	2
Stochastic Volatility	5	5	3	4	4	4
Forecast Combination	1	2	1	2	1	1

Table 2. Forecast performance ranking for each model and forecast combination for the S&P 500 (left) and FTSE (right) for the 1, 2, and 5% quantiles.

5 CONCLUSIONS

In this study, we empirically examined whether combining individual forecasts of VaR can improve forecast accuracy. The results of this study indicate that VaR forecast combination may improve VaR forecast accuracy. Specifically, we empirically examined the performance of VaR forecast combination for the S&P500 and FTSE stock indices using four different models and the combination of the forecasts from these models for the 1, 2, and 5% quantiles. We found that, overall, forecast combination performed well, ranking as the best forecast for almost all quantiles for both indices. However, unlike a continuous loss function such as MFSE or related loss functions which measure the distance between the forecast and the true observed value, we evaluated the forecast performance using the actual proportion of violations of VaR compared to the expected violations (the quantile), and ranked the forecast performance based on this. Therefore, while the results do show that the simple average forecast combination typically improves forecast accuracy based on proportion of violations of VaR, further research is necessary to investigate why, and under which conditions, forecast combination of the quantile may improve forecast accuracy. Understanding the theoretical underpinning of quantile forecast combination has implications for risk management practices, as accurate risk forecasts are imperative for decision making.

ACKNOWLEDGEMENT

Sylvia Soltyk would like to thank the Food Agility CRC for financial support. Aloke Phatak was supported by the Australian Government through the Australian Research Council's Industrial Transformation Training Centres scheme (Project IC180100030).

References

Basel Committee on Banking Supervision (2019), Minimum capital requirements for market risk., Technical report, Basel Committee on Banking Supervision.

Bates, J. M. & Granger, C. W. J. (1969), 'The combination of forecasts', *Operational Research Quarterly* **20**, 451–468.

Bayer, S. (2018), 'Combining value-at-risk forecasts using penalized quantile regressions', *Econometrics and Statistics* **8**, 56–77.

Boucher, C. M., Daníelsson, J., Kouontchou, P. S. & Maillet, B. B. (2014), 'Risk models-at-risk', *Journal of Banking & Finance* **44**, 72–92.

Chan, F. & James, A. (2011), Application of forecast combination in volatility modelling, *in* F. Chan, D. Marinova & R. Anderssen, eds, 'MODSIM2011, 19th International Congress on Modelling and Simulation', Modelling and Simulation Society of Australia and New Zealand, pp. 1610–1616.

Chan, F., Pauwels, L. & Soltyk, S. (2020), Frequentist Averaging, *in* P. Fuleky, ed., 'Macroeconomic Forecasting in the Era of Big Data', Vol. 1, Springer, chapter 11, pp. 329–357.

Fuertes, A.-M. & Olmo, J. (2013), 'Optimally harnessing inter-day and intra-day information for daily value-at-risk prediction', *International Journal of Forecasting* **29**, 28–42.

Ghalanos, A. (2022), *rugarch: Univariate GARCH models.* R package version 1.4-9.

Halbleib, R. & Pohlmeier, W. (2012), 'Improving the value at risk forecasts: Theory and evidence from the financial crisis', *Journal of Economic Dynamics and Control* **36**, 1212–1228.

Huang, H. & Lee, T.-H. (2013), 'Forecasting value-at-risk using high-frequency information', *Econometrics* **1**, 127–140.

Hyndman, R., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O'Hara-Wild, M., Petropoulos, F., Razbash, S., Wang, E. & Yasmeen, F. (2023), *forecast: Forecasting functions for time series and linear models*. R package version 8.21.
URL: *https://pkg.robjhyndman.com/forecast/*

Jagannathan, R., Glosten, L. & Runkle, D. (1993), 'On the relation between the expected value and the volatility of the nominal excess return on stocks', *Journal of Finance* **48**, 1779–1801.

Jeon, J. & Taylor, J. (2013), 'Using caviar models with implied volatility for value-at-risk estimation', *Journal of Forecasting* **32**, 62–74.

Jorion, P. (2006), *Value at Risk: The New Benchmark for Managing Financial Risk*, 3rd edn, McGraw-Hill.

Kastner, G. (2016), 'Dealing with stochastic volatility in time series using the R package stochvol', *Journal of Statistical Software* **69**(5), 1–30.

Kastner, G. & Frühwirth-Schnatter, S. (2014), 'Ancillarity-sufficiency interweaving strategy (asis) for boosting mcmc estimation of stochastic volatility models', *Computational Statistics & Data Analysis* **76**, 408–423.

Kuester, K., Mittnik, S. & Paolella, M. (2006), 'Value-at-risk prediction: A comparison of alternative strategies', *Journal of Financial Econometrics* **4**, 53–89.

McAleer, M., Jimenez-Martin, J.-A. & Amaral, T. (2013*a*), 'Has the basel accord improved risk management during the global financial crisis?', *The North American Journal of Economics and Finance* **26**, 250–265.

McAleer, M., Jimenez-Martin, J.-A. & Amaral, T. (2013*b*), 'International evidence on gfc-robust forecasts for risk management under the basel accord', *Journal of Forecasting* **32**, 267–288.

Nieto, M. & Ruiz, E. (2016), 'Frontiers in var forecasting and backtesting', *International Journal of Forecasting* **32**, 475–501.

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/*

Taylor, J. (2020), 'Forecast combinations for value at risk and expected shortfall', *International Journal of Forecasting* **36**, 428–441.

Timmermann, A. (2006), Forecast Combinations, *in* 'Handbook of Economic Forecasting', Vol. 1, Elsevier, chapter 4, pp. 135–196.