INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN BUSINESS & SOCIAL SCIENCES
Vol 14, Issue 12, (2024) E-ISSN: 2222-6990

Application of AO-GARCH-MIDAS Model Based on Volatility Effect in Stock Market Volatility Forecasting

Ting Liu[a], Weichong Choo[a*], Matemilola Bolaji Tunde[a], Han Xinping[b]
[a] School of Business and Economics, Universiti Putra Malaysia, Seri Kembangan, Malaysia,
[b]School of Economics and Trade, Henan Finance University, Zhengzhou, China
Email: wcchoo@upm.edu.my

To Link this Article: http://dx.doi.org/10.6007/IJARBSS/v14-i12/23888 DOI:10.6007/IJARBSS/v14-i12/23888
Published Date: 24 December 2024

Abstract
This paper investigates the performance of GARCH-MIDAS and AO-GARCH-MIDAS models in predicting stock market volatility based on considering the volatility effects of macroeconomic variables. The traditional GARCH-MIDAS model only considers the level effects of macroeconomic variables, while this paper, by introducing the volatility effects of macroeconomic variables, in particular, combines the realized volatility with macroeconomic variables (e.g., CPI, M2, UD). The results of the MCS test based on MAE and MSE show that the AO-GARCH-MIDAS family of models performs well in out-of-sample forecasting, especially the AO-GARCH-MIDAS-RV+UD model, which is the best performer under both assessment metrics and shows strong forecasting ability. In contrast, traditional GARCH-MIDAS models and combined models based on macroeconomic variables (e.g., GARCH-MIDAS-RV+UD and GARCH-MIDAS-UD) perform weakly in out-of-sample forecasts. Overall, this paper shows that the AO-GARCH-MIDAS model, which takes into account the volatility effects of macroeconomic variables, significantly improves forecasting accuracy when dealing with complex economic environments and abnormal volatility, providing more reliable stock market volatility forecasting results.
Keywords: Volatility, Stock Market, GARCH-MIDAS, Additive Outlier

Introduction
Stock markets not only provide investors with diverse savings and investment opportunities but also play a key role in reallocating funds across multiple sectors of the economy (Ahmad & Ramzan, 2016). Stock prices are the most direct reflection of the stock market and hold great importance for policymakers and economists, as future stock prices can indicate changes in long-term economic activity. However, stock prices only represent the amount investors are willing to pay at the time of trading and combine company fundamentals with

market conditions. In contrast, stock volatility measures the magnitude of stock price movements and reflects the market's sensitivity to risk. Its primary application lies in assessing market risk and developing risk management strategies, enabling investors to better understand market uncertainty. Volatility, as an explicit measure of risk, remains a central focus for financial economists (Park, 2002). Therefore, accurate measurement and forecasting of future volatility is crucial.

The level and volatility factors of the same macroeconomic variable have distinct impacts on the volatility of commodity futures markets (Mo et al., 2018). In comparison, the impact of volatility in macroeconomic variables is more significant than the impact of levels in both the Chinese and Indian markets, as confirmed. To test the impact of macroeconomic variables on long-run variance, Liu et al. (2020) introduce levels and variances of macro variables into the model. They similarly find that differences in macroeconomic variables (PPI and IP) have a greater impact on commodity futures market volatility, while level factors have a weaker impact. In contrast, Zhang et al. (2022), in examining the estimation of EPU on Treasury futures, show that the Economic policy uncertainty (EPU) level model contributes more than 17% of total volatility in China and the United States. In the variance model, the contribution to the U.S. decreases, while the contribution to the total volatility of the Chinese sample is only 1.41%. From the combined model of the two, the contribution is the largest, at 44% and 22%, respectively. This implies that the combination of the level and variance of EPU is an important source of volatility in Treasury futures.

The findings of Asgharian et al. (2013) suggest that including a long-run variance component of macroeconomic information in the model enhances its forecasting ability. Engle et al. (2013) consider the ability of levels and variances in the long-run component for forecasting. These levels and variances are examined not only separately but also in combination within the GARCH-MIDAS model. According to the long-run forecasting results, the model with a level long-run component driven by inflation and increases in industrial production is comparable to other models for out-of-sample forecasting at quarterly frequencies, yet outperforms purely time-series conventional models at longer horizons. Similarly, Yu and Huang (2021) explore the impact of economic policy uncertainty on stock volatility by considering both the level factor and the variance factor of the variable. The results indicate that both the level and variance of the indicator provide valuable information for estimating stock volatility. Moreover, the GARCH-MIDAS model, which incorporates realized volatility along with these two factors, demonstrates strong predictive power for forecasting.

Gong et al. (2022) examine the effects of the level and volatility of macroeconomic variables on oil price volatility by analyzing these two forms of the variables separately in the model. Within the sample, the level and volatility of macro variables, except for the exchange rate, show different effects on oil price volatility. Additionally, the study indicates that exchange rate volatility is most closely related to the volatility of international oil prices and demonstrates the best empirical performance. Wu et al. (2020) analyze the performance of low-frequency economic indicator levels and volatility factors in predicting the volatility of crude oil futures prices. In terms of overall estimation performance, the volatility effect model outperforms the level effect model in predictive performance.

Asset return series generally exhibit several key characteristics, particularly notable volatility clustering, and conditional heteroskedasticity, which are crucial for constructing return-based volatility forecasting models. Given its ability to effectively capture the complex dynamics of volatility persistence and clustering, the family of GARCH models serves as an indispensable tool for volatility modeling in financial markets. However, studies based on GARCH-MIDAS models often overlook the impact of additive outliers on volatility forecasting. Franses and Ghijsels (1999) highlight that the presence of additive outliers in a time series can have dual detrimental effects: first, it introduces bias in in-sample parameter estimation, compromising the reliability of the model fit. Second, it significantly reduces the accuracy of out-of-sample forecasts by distorting the underlying structure of the data, thereby impairing the model's ability to generalize.

Building on the advantages of the GARCH-MIDAS model, the AO-GARCH-MIDAS model proposed by Liu et al. (2024) introduces the Additive Outlier (AO) model, which significantly enhances its ability to address abnormal fluctuations in macroeconomic data. By identifying and adjusting outliers, the model effectively mitigates the interference of abnormal fluctuations on volatility forecasts, improving both its robustness and forecasting accuracy. This advancement not only optimizes the ability to capture the relationship between macroeconomic variables and financial market volatility but also enhances the model's applicability and robustness in complex economic environments. However, the model primarily focuses on the impact of the level effects of macroeconomic variables on volatility forecasts and does not explore the role of volatility effects in depth. Given the important role that volatility effects play in modeling financial market dynamics—particularly in capturing the asymmetric and dynamic effects of volatility in macroeconomic variables on financial market risk—this study extends the AO-GARCH-MIDAS model to assess the dual mechanism of macroeconomic variables by incorporating volatility effects. This improvement aims to enhance the model's capacity to explain and predict changes in financial market volatility while providing a more accurate quantitative tool for risk management in complex economic environments.

Methodology

Ghysels et al. (2007) first propose the Mixed Data Sampling model (MIDAS), and Engle et al. (2013) use the MIDAS approach to relate macroeconomic variables to the long-run components of volatility. A GARCH-MIDAS model was constructed, which provides a new approach to revisit the relationship between stock market volatility and economic activity and volatility.

Assuming $r_{i,t}$ is the logarithmic rate of return on day i of month t, the return and volatility in this GARCH-MIDAS model are described as follows

$$\boldsymbol{r_{i,t} - E_{i-1,t}(r_{i,t}) = \sqrt{\tau_t g_{i,t}}\varepsilon_{i,t}, E_{i-1,t}(r_{i,t}) = \mu, \forall i = 1, 2, \cdots, N_t} \quad \textbf{Eq. 1}$$

$$\boldsymbol{\varepsilon_{i,t} \big| \psi_{i-1,t} \sim N(0,1), \sigma_{i,t}^2 = \tau_i g_{i,t}} \quad \textbf{Eq. 2}$$

Where, N_t is the number of days in month t, $E_{i-1,t}$ is the conditional expectation and $\psi_{i,t}$ is the information set of the i-1 day of the rate of return in month t. $\sigma_{i,t}^2$ is the conditional variance, $\varepsilon_{i,t}$ is the random disturbance term, assuming a standard normal distribution, assuming Eq. 1 also can be expressed as

$r_{i,t} = \mu + \sqrt{\tau_t g_{i,t}}\varepsilon_{i,t}, \forall i = 1, 2, \cdots, N_t$ **Eq. 3**

The volatility in Eq. 3 in the expression is decomposed into two parts: short-term volatility $g_{i,t}$, long-term volatility τ_t. Short-term volatility $g_{i,t}$ satisfies the GARCH (1, 1) model:

$\boldsymbol{g_{i,t} = (1 - \alpha - \beta) + \alpha \frac{(r_{i-1,t}-\mu)^2}{\tau_t} + \beta g_{i-1,t}}$ **Eq. 4**

Where $\alpha > 0, \beta > 0, \alpha - \beta < 1$. In addition, when the long-term trend τ_t is affected by the realized volatility (RV). This model is called a GARCH-MIDAS-X model of the form.

$\log(\tau_t) = m + \theta_0^l \sum_{k=1}^{k} \varphi_k(\omega_1, \omega_2) X_{t-k}^l + \theta_1^v \sum_{k=1}^{k} \varphi_k(\omega_1, \omega_2) X_{t-k}^v$ **Eq. 5**

$\boldsymbol{\log(\tau_t) = m + \theta_0 \sum_{k=1}^{k} \varphi_k(\omega_1, \omega_2) RV_{t-k} + \theta_1^l \sum_{k=1}^{k} \varphi_k(\omega_1, \omega_2) X_{t-k}^l + \theta_2^v \sum_{k=1}^{k} \varphi_k(\omega_1, \omega_2) X_{t-k}^v}$ **Eq. 6**

RV_t is calculated from the monthly sum of squared daily returns, $RV_t = \sum_{i=1}^{N_t} r_{i,t}^2$; K represents the maximum lag order of low-frequency variables selected by AIC and BIC information standards. $\varphi_k(\omega_1, \omega_2)$ is the weighting scheme of the Beta lag structure (Engle et al., 2013) because it is more flexible and more commonly used to accommodate various lag structures (Ghysels et al., 2007), the polynomial shown below

$\varphi_k(\omega_1, \omega_2) = \frac{(\kappa/K)^{\omega_1-1}(1-\kappa/(K+1))^{\omega_2-1}}{\sum_{j=1}^{K}(j/K)^{\omega_1-1}(1-j/K)^{\omega_2-1}}$ **Eq. 7**

Fix ω_1= 1 to ensure that the weight of the lag variable is in the form of attenuation. In other words, the closer the distance to the current period, the more significant the impact on the current period (Yaya et al., 2022). The coefficient determines the attenuation speed of the impact of low-frequency data on high-frequency data. Therefore, the polynomial can be simplified as

$\varphi_k(\omega_2) = \frac{(1-\kappa/(K+1))^{\omega_2-1}}{\sum_{j=1}^{K}(1-j/k)^{\omega_2-1}}$ **Eq. 8**

There are two approaches to estimating the GARCH-MIDAS model: the first one is a fixed window, and the other one is a rolling window. The results of Angelidis et al. (2004) and Degiannakis et al. (2008) show that a rolling window can capture changes in market activity more effectively because it allows the parameters to be re-estimated. Not only that, in this paper, we choose the 5-step rolling forecast so that each day updates the model and forecasts the volatility for the coming days.

In order to hybrid the additive outlier model and the GARCH-MIDAS model, the method introduced by Liu et al. (2024) is specified as follows

$\frac{r_{i,t}^2}{\tau_t} = g_{i,t} + z_{i,t}$ Eq. 9

$E_{t-1}(z_{i,t}) = 0$ Eq. 10

Based on the above formula rewrite the Short-term volatility $g_{i,t}$ as

$\frac{r_{i,t}^2}{\tau_t} - z_{i,t} = (1 - \alpha - \beta) + \alpha \frac{(r_{i-1,t}-\mu)^2}{\tau_t} + \beta(\frac{r_{i,t}^2}{\tau_t} - z_{i-1,t})$ Eq. 11

This formula corresponds to the paper of Franses and Ghijsels (1999) on GARCH (1, 1) model for $r_{i,t}^2$.

Let $f_{i,t} = \frac{r_{i,t}^2}{\tau_t}$

$(1-(\alpha+\beta)L)f_{i,t} = (1-\beta L)z_{i,t}$ Eq. 12

From this equation, $\phi(L)$ and $\theta(L)$ can be determined as below

$\phi(L) = 1-(\alpha+\beta)L$ Eq. 13

$\theta(L) = 1-\beta L$ Eq. 14

According to the equation $r_t^{*2} = \hat{z}_t^* + \hat{h}_t$ form Franses and Ghijsels (1999), the formula of ${r_{i,t}^*}^2$ can be constructed as follow

$r_{i,t}^{*2} = \tau_t(z_{i,t}^* + g_{i,t}) \quad t = \nu$ Eq. 15

Hence, the AO-corrected returns can be constructed

$r_{i,t}^* = r_{i,t} \quad t \neq \nu$ Eq. 16

$r_{i,t}^* = sign(r_{i,t}).({r_{i,t}^*}^2)^{1/2} \quad t = \nu$ Eq. 17

This expression shows that although $r_{i,t}$ is replaced, its sign is retained in $r_{i,t}^*$, when $t = \nu$. Based on Chen and Liu (1993) of AO-ARMA, the estimated residuals $\hat{\varepsilon}_t$ can be represented by

$\hat{\varepsilon}_t = \pi(L)y_t$ Eq. 18

When

$t < \nu, \quad x_t = 0$

$t = \nu, \quad x_t = 1$

$t = \nu + i(i > 0), \quad x_{t+i} = -\pi_i$

At time $t = \nu$, the impact ρ of AO can be estimated as

$\hat{\rho}(\nu) = \frac{\sum_{t=\nu}^{n} x_t\hat{\varepsilon}_t}{\sum_{t=\nu}^{n} x_t^2}$ Eq. 19

To test the significance of AO model, Chang et al. (1988) propose to standardize $\hat{\rho}(\nu)$. It requires an estimate of the variance of the residual process, this estimate should ideally not contain too much bias because of outliers. This study uses the method of Chen and Liu (1993) the so-called 'omit one' to estimate a robust error variance. Based on this approach, we can get a standardized statistic

$\hat{\nu} = \frac{\hat{\rho}(\nu)}{\hat{\sigma}_a}\sqrt{\sum_{t=\nu}^{n} x_t^2}$ Eq. 20

The influence of AO is significant when $\hat{\nu}$ exceeds the value C. As Franses and Ghijsels (1999) mentioned, $\hat{\nu}$ is asymptotically standard normal. As posited by Chen and Liu (1993), it is imperative to scrutinize the parameter C for values exceeding 3 when the dataset comprises more than 200 observations. Although other choices for C are viable, this study has identified superior outcomes when C equals 14. When the value of $\hat{\nu}$ exceeds the value C, and $t = \nu$, the observation y_t shall be substituted with AO-corrected y_t^*, derived from Eq. 19, and the additive outlier model $y_t = y_t^* + \rho I_t(\nu)$.

In the dataset, to avoid the existence of multiple AOs, these steps need to repeat unless $\hat{\nu}$ becomes insignificant. When there is no more additive outlier, the final step is to re-estimate the model parameters based on all observations, where some of them have been corrected by using AO model.

Data Description and Preliminary Analysis

This study focuses on the Standard and Poor's 500 index of the U.S. stock market, with data covering the period from October 1, 2009, to March 31, 2023, comprising a total of 3,406 trading days of daily closing prices obtained from Yahoo Finance (https://finance.yahoo.com). The selected macroeconomic variables include the monthly Consumer Price Index (CPI), and Money Supply (M2). The US Dollar (UD) Index is a representative variable for the foreign exchange market. To further analyze stock market volatility, the daily closing prices of the Standard and Poor's 500 index are converted into logarithmic returns, denoted as SP500, serving as a key measure of market volatility.

Table **1** presents the descriptive statistics of stock returns and various macroeconomic variables. The results indicate that after calculating the logarithmic return on the daily closing price of the stock, the mean value is 0.0173, and the UD index is the lowest at only 0.0682. Additionally, the stock market return is negatively skewed and leptokurtic, with a kurtosis value exceeding three. This indicates that it does not conform to a normal distribution and is characterized by spikes and thick trailing tails, shifted to the left. The kurtosis value of UD is less than three, and the p-value exceeds 10%, suggesting no significant difference from a normal distribution, which implies that this indicator conforms to or is close to a normal distribution. The other macroeconomic variables deviate from a normal distribution and are shifted to the right.

Table 1
Descriptive Statistics and Stationary Testing.=

	Mean	Median	Max	Min	Std. Dev	Skewness	Kurtosis	P(JB)
SP500	0.0173	0.0280	3.8949	-5.5439	0.4870	-0.7220	15.8770	0.0000
Macroeconomic data								
CPI	2.4833	1.9000	9.1000	-0.2000	2.0632	1.6236	5.0381	0.0000
M2	7.3846	6.1350	26.6400	-3.9200	5.4475	1.8315	6.6625	0.0000
UD	0.0682	0.0787	1.5751	-1.2026	0.5335	0.1675	2.8300	0.6210

Table **2** presents the results of the unit root test, Ljung-Box Q statistic, and ARCH test. The findings from all three stability tests confirm that the SP500 sequence is stable. Additionally, the results of the Ljung-Box Q statistic test for autocorrelation in the stock market return series show that the p-values for the SP500 return series are all significant, indicating significant autocorrelation in the return series. Consequently, the original hypothesis of no autocorrelation is rejected.

This result further indicates that the volatility of current returns is significantly influenced by prior period volatility, demonstrating strong time dependence in the return series. However, this autocorrelation can be effectively modeled and addressed using the GARCH model, which is well-suited for capturing the dynamic characteristics of market volatility.

Table 2
Heteroskedastic Test

	ADF	PP	KPSS	Ljung-Box Q-statistic (36)	ARCH
SP500	-19.5543*** (0.0000)	-66.3585*** (0.0001)	0.0301	306.2200*** (0.0000)	882.6692*** (0.0000)

Notes: ***Indicate rejections of the null hypothesis at the 1% significance level. The numbers in parentheses are the p-values of the tests.

Empirical Results

In-Sample Result

The in-sample results of this study are presented in

Table 3. First, the estimated coefficients α and β for the total daily volatility of the short-term component ($g_{i,t}$) of stock returns are significant at the 1% level in all models, with both parameters corresponding to the ARCH and GARCH terms for short-term components, respectively. Second, both parameters are positive, and their sum is close to one, which implies a strong volatility persistence effect in these five stock markets. Additionally, ω1 and

ω2 are the beta polynomial weights of the long-run component of the model, which are important for most of the variables. When the results are significant, it indicates that the low-frequency variables can predict long-run volatility. Finally, θ is the sum of the weighted rolling window realized volatility for each variable.

Table 3
In-*Sample Parameter Estimation Results*.

	GARCH-MIDAS-CPI	GARCH-MIDAS - M2	GARCH-MIDAS-UD	GARCH-MIDAS-RV+CPI	GARCH-MIDAS - RV+M2	GARCH-MIDAS-RV+UD	AO-GARCH-MIDAS-CPI	AO-GARCH-MIDAS - M2	AO-GARCH-MIDAS-UD	AO-GARCH-MIDAS-RV+CPI	AO-GARCH-MIDAS - RV+M2	AO-GARCH-MIDAS-RV+UD
μ	0.0346*** (0.0059)	0.0349*** (0.0060)	0.0342*** (0.0060)	0.0352*** (0.0059)	0.0352*** (0.0059)	0.0344*** (0.0060)	0.0345*** (0.0060)	0.0350*** (0.0061)	0.0331*** (0.0061)	0.0348*** (0.0059)		0.0337*** (0.0060)
α	0.2084*** (0.0299)	0.1969*** (0.0317)	0.2051*** (0.0283)	0.2049*** (0.0279)	0.2459*** (0.0355)	0.2088*** (0.0291)	0.2070*** (0.0303)	0.1945*** (0.0329)	0.1945*** (0.0286)	0.2025*** (0.0288)		0.1898*** (0.0266)
β	0.7218*** (0.0315)	0.7612*** (0.0270)	0.7407*** (0.0297)	0.7071*** (0.0316)	0.6759*** (0.0396)	0.7205*** (0.0333)	0.7211*** (0.0327)	0.7646*** (0.0305)	0.7544*** (0.0303)	0.7032*** (0.0337)		0.7372*** (0.0318)
m	-1.4362*** (0.2663)	- 1.6206*** (0.4010)	- 1.8925*** (0.2971)	- 2.3041*** (0.3221)	- 2.4663*** (0.4280)	- 2.4571*** (0.3851)	- 1.4558*** (0.2729)	- 2.6092*** (0.4328)	- 1.6984*** (0.2399)	- 2.3903*** (0.3544)		- 2.6033*** (0.3397)
θ_{RV}	/	/	/	0.4285*** (0.1442)	0.5618*** (0.1514)	0.3933* (0.2254)	/	/	/	0.4640*** (0.1765)		0.4530** (0.2057)
θ_{CPI^l}	-0.4113*** (0.1450)	/	/	-0.3321** (0.1305)	/	/	- 0.4055*** (0.1561)	/	/	-0.3232** (0.1468)	/	/
θ_{CPI^V}	0.0683*** (0.0158)	/	/	0.0528*** (0.0149)	/	/	0.0667*** (0.0166)	/	/	0.0510*** (0.0167)	/	/
θ_{M2^l}	/	-0.0468* (0.0263)	/	/	-0.0700* (0.0424)	/	/	0.1495*** (0.0414)	/	/		/
θ_{M2^V}	/	0.0057*** (0.0015)	/	/	0.0046** (0.0019)	/	/	-0.0009* (0.0005)	/	/		/
θ_{UD^l}	/	/	- 2.9942*** (1.0603)	/	/	-2.0009** (0.8185)	/	/	- 2.6310*** (0.9394)	/		-1.7648** (0.7211)

θ_{UD^v}	/	/	1.8823** (0.7362)	/	/	0.8721 (0.7427)	/	/	0.9749** (0.4716)	/	0.6737 (0.7033)
ω_{RV}	/	/	/	1.000 (0.7360)	1.3375 (0.9592)	4.2813*** (1.5960)	/	/	/	6.0684** (3.0482)	4.1986*** (1.4374)
$\omega_2^{CPI^l}$	1.3074 (3.4203)	/	/	115.0159 (0.0127)	/	/	44.6143 (43.9606)	/	/	5.0889 (22.6683)	/
$\omega_2^{CPI^v}$	22.3309*** (7.9087)	/	/	44.0093 (0.3968)	/	/	1.2462 (1.7395)	/	/	9.1228 (7.5236)	/
$\omega_2^{M2^l}$	/	1.0001** (0.4108)	/	/	2.8756*** (0.6625)	/	/	1.0001*** (0.2980)	/	/	/
$\omega_2^{M2^v}$	/	1.0464*** (0.2762)	/	/	1.7783*** (0.4641)	/	/	1.0001 (0.6258)	/	/	/
$\omega_2^{UD^l}$	/	/	1.0000** (0.4942)	/	/	1.0000** (0.4353)	/	/	1.0000** (0.4383)	/	1.0000** (0.4217)
$\omega_2^{UD^v}$	/	/	4.0264*** (1.1896)	/	/	1.8791*** (0.5559)	/	/	1.0009 (1.6745)	/	2.3682*** (0.7334)
LLF/AIC	2219.1630	2315.8310	2237.7510	2206.4740	2222.0390	2231.082	2194.8980	2188.9500	2226.3780	2190.1490	2193.5520
LLF/BIC	2266.2060	2363.1240	2284.7940	2265.2790	2280.8430	2289.886	2241.9420	2235.9940	2273.4220	2248.9530	2252.3560

Notes: Notes: Standard errors in parentheses. *Indicate rejections of the null hypothesis at the 10% significance level. **Indicate rejections of the null hypothesis at the 5% significance level. ***Indicate rejections of the null hypothesis at the 1% significance level. The numbers in parentheses are the p-values of the tests.

Out-of-Sample Result

Market participants are more interested in a model's ability to predict future stock volatility rather than just its in-sample performance (Ma et al., 2019; Wang et al., 2020; Yu & Huang, 2021). Since most investors aim to uncover new investment opportunities from historical market information, models with higher efficiency are required to help mine this information effectively. Therefore, this section focuses on analyzing whether models with additive outliers can enhance predictive ability. All models in this study employ fixed-window forecasting. The stock return data is divided into two subgroups for forecasting: the in-sample group accounts for 80% of the total sample, while the out-of-sample group accounts for the remaining 20%, corresponding to 660 forecasting periods.

Three loss functions are selected in this study to evaluate the predictive performance of the models: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Median Absolute Error (MedAE). RMSE, the square root of the mean of the squared prediction errors, penalizes large errors more severely due to the squared term magnifying differences.

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|\sigma_t - \hat{\sigma}_t|,$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\sigma_t - \hat{\sigma}_t)^2},$$

$$MedAE = Median|\sigma_t - \hat{\sigma}_t|.$$

The results in

Table 4 show the differences in forecasting accuracy between different models and different combinations of macro variables, the two smallest values are highlighted in bold. For example, with CPI as a macroeconomic variable alone, the GARCH-MIDAS-CPI model has an MAE of 1.2490, an RMSE of 2.0532, and a MedAE of 0.6213, which exhibit weak forecasting accuracy. After adding volatility factors into the model, such as the GARCH-MIDAS-RV+CPI model, the prediction results deteriorate significantly, with an MAE of 1.6157 and a RMSE of 2.6225, suggesting that the volatility of the CPI variable poses a challenge to volatility prediction.

Among the combinations of other macroeconomic variables, the GARCH-MIDAS-M2 and GARCH-MIDAS-UD models perform relatively well, with the GARCH-MIDAS-M2 having an MAE of 0.9718, a RMSE of 1.4084, and a MedAE of 0.6093, which makes the forecasting results relatively more accurate. The models after adding realized volatility (e.g., GARCH-MIDAS-RV+M2 and GARCH-MIDAS-RV+UD) show more stable prediction accuracies, with GARCH-MIDAS-RV+UD having the lowest MAE of 0.8028, showing the positive effect of volatility on the prediction results.

In contrast, the AO-GARCH-MIDAS family of models outperforms the traditional GARCH-MIDAS model in most cases. For example, the AO-GARCH-MIDAS-RV+UD model has an MAE of 0.7759, an RMSE of 1.2076, and a MedAE of 0.4722, all of which are the lowest values among all the models, demonstrating strong predictive ability in dealing with the level and volatility of macroeconomic variables.

Overall, the results in the table indicate that combining the level and volatility of macroeconomic variables can effectively improve the accuracy of stock market volatility prediction, especially the AO-GARCH-MIDAS series of models show a superior prediction performance after considering the volatility effect.

Table 4
Out-of-Sample Prediction Results

Model	MAE	RMSE	MedAE	Mean value
GARCH-MIDAS-CPI	1.2490	2.0532	0.6213	1.3078
GARCH-MIDAS-M2	0.9718	1.4084	0.6093	0.9965
GARCH-MIDAS-UD	0.8222	1.2982	0.4887	0.8697
GARCH-MIDAS-RV+CPI	1.6157	2.6225	0.7697	1.6693
GARCH-MIDAS-RV+M2	1.3026	1.9575	0.8179	1.3593
GARCH-MIDAS-RV+UD	0.8028	**1.2433**	0.4959	0.8473
AO-GARCH-MIDAS-CPI	1.0935	1.6619	0.6317	1.1290
AO-GARCH-MIDAS-M2	0.8222	1.2693	0.4932	0.8616
AO-GARCH-MIDAS-UD	**0.7926**	1.2532	**0.4555**	**0.8338**

AO-GARCH-MIDAS-RV+CPI	1.2648	1.9922	0.6645	1.3072
AO-GARCH-MIDAS-RV+M2	1.1278	1.6356	0.6764	1.1466
AO-GARCH-MIDAS-RV+UD	**0.7759**	**1.2076**	**0.4722**	**0.8186**

According to the results of the MCS test based on MAE and MSE in

Table **5** and

Table **6**, the AO-GARCH-MIDAS-RV+UD model performs well under both metrics, with a Tmax,M value of 1.0000 and Rank_M of 1 and 1, respectively, indicating that it has the strongest out-of-sample prediction accuracy. Secondly, the AO-GARCH-MIDAS-UD model ranks second under the MAE metric with a Tmax,M of 1.0000, showing good predictive ability. The GARCH-MIDAS-RV+UD model follows in second place under the MSE indicator with a Tmax,M of 1.0000, which also performs better. However, the GARCH-MIDAS-UD model performs poorly under both MAE and MSE metrics, with Tmax,M values of 0.2576 and 0.3504, and Rank_M is 5, showing its weak predictive ability. Overall, the AO-GARCH-MIDAS family of models, especially the AO-GARCH-MIDAS-RV+UD model, demonstrates significant advantages in dealing with out-of-sample forecasting, providing more accurate and reliable forecasting results, while the traditional GARCH-MIDAS family of models performs weakly, especially in models that consider macroeconomic variables.

Table 5
Results of MAE-based MCS Test

Model	Tmax,M	Rank_M
GARCH-MIDAS-CPI	/	/
GARCH-MIDAS-M2	/	/
GARCH-MIDAS-UD	0.2576	5
GARCH-MIDAS-RV+CPI	/	/
GARCH-MIDAS-RV+M2	/	/
GARCH-MIDAS-RV+UD	0.9944	3
AO-GARCH-MIDAS-CPI	/	/
AO-GARCH-MIDAS-M2	0.8304	4
AO-GARCH-MIDAS-UD	1.0000	2
AO-GARCH-MIDAS-RV+CPI	/	/
AO-GARCH-MIDAS-RV+M2	/	/
AO-GARCH-MIDAS-RV+UD	1.0000	1

Table 6
MSE-based MCS test results

Model	Tmax,M	Rank_M
GARCH-MIDAS-CPI	/	/
GARCH-MIDAS-M2	/	/
GARCH-MIDAS-UD	0.3504	5
GARCH-MIDAS-RV+CPI	/	/
GARCH-MIDAS-RV+M2	/	/
GARCH-MIDAS-RV+UD	1.0000	2

AO-GARCH-MIDAS-CPI	/	/
AO-GARCH-MIDAS-M2	0.9672	4
AO-GARCH-MIDAS-UD	1.0000	3
AO-GARCH-MIDAS-RV+CPI	/	/
AO-GARCH-MIDAS-RV+M2	/	/
AO-GARCH-MIDAS-RV+UD	1.0000	1

Conclusion

This study explores the role of macroeconomic variables and realized volatility in forecasting stock market volatility by comparing the out-of-sample forecasting performance of a variety of GARCH-MIDAS and AO-GARCH-MIDAS-based models. The experimental results show that the AO-GARCH-MIDAS family of models significantly outperforms the traditional GARCH-MIDAS model in terms of forecasting accuracy, especially when dealing with abnormal volatility and complex economic environments, exhibiting stronger forecasting capabilities.
Under the two error measures, MAE and MSE, the AO-GARCH-MIDAS models perform well overall, with the AO-GARCH-MIDAS-RV+UD model performing the best under both evaluation metrics, achieving an MAE of 0.7759, an RMSE of 1.2076, a MedAE of 0.4722, and a mean value of 0.8186. This suggests that the AO-GARCH-MIDAS model is effective in improving forecasting accuracy when combining realized volatility and economic variables, especially after accounting for volatility effects.

In contrast, traditional GARCH-MIDAS models and combined models based on macroeconomic variables (e.g., GARCH-MIDAS-RV+UD and GARCH-MIDAS-UD) are weaker in terms of forecasting accuracy, particularly under the MSE metric.

The results show that the AO-GARCH-MIDAS model considering the volatility effect significantly outperforms the traditional GARCH-MIDAS model in terms of forecasting accuracy. This not only demonstrates the adaptability and reliability of the AO-GARCH-MIDAS model in complex economic environments and abnormal fluctuations but also reflects the key role of volatility effects in enhancing forecasting performance. Future research could explore the introduction of other potentially relevant macroeconomic variables, such as monetary policy indicators, consumer confidence indices, or global economic uncertainty indices, to further enhance the model's predictive power. Although the AO-GARCH-MIDAS model has demonstrated significant advantages, further research on optimization algorithms (e.g., hyper-parameter tuning or introduction of machine learning methods) can be conducted in the future to cope with more complex and variable market conditions.

References

Ahmad, N., & Ramzan, M. (2016). Stock market volatility and macroeconomic factor volatility. *International Journal of Research in Business Studies and Management*, *3*(7).

Angelidis, T., Benos, A., & Degiannakis, S. (2004). The use of GARCH models in VaR estimation. *Statistical methodology*, *1*(1-2), 105-128.

Asgharian, H., Hou, A. J., & Javed, F. (2013). The importance of the macroeconomic variables in forecasting stock return variance: A GARCH-MIDAS approach. *Journal of Forecasting*, *32*(7), 600-612.

Chang, I., Tiao, G. C., & Chen, C. (1988). Estimation of time series parameters in the presence of outliers. *Technometrics*, *30*(2), 193-204.

Chen, C., & Liu, L.-M. (1993). Joint estimation of model parameters and outlier effects in time series. *Journal of the American Statistical Association*, *88*(421), 284-297.

Degiannakis, S., Livada, A., & Panas, E. (2008). Rolling-sampled parameters of ARCH and Levy-stable models. *Applied Economics*, *40*(23), 3051-3067.

Engle, R. F., Ghysels, E., & Sohn, B. (2013). Stock market volatility and macroeconomic fundamentals. *Review of Economics and Statistics*, *95*(3), 776-797.

Franses, P. H., & Ghijsels, H. (1999). Additive outliers, GARCH and forecasting volatility. *International Journal of forecasting*, *15*(1), 1-9.

Ghysels, E., Sinko, A., & Valkanov, R. (2007). MIDAS regressions: Further results and new directions. *Econometric reviews*, *26*(1), 53-90.

Gong, X., Wang, M., & Shao, L. (2022). The impact of macro economy on the oil price volatility from the perspective of mixing frequency. *International Journal of Finance & Economics*, *27*(4), 4487-4514.

Liu, R., Yang, J., & Ruan, C. (2020). The Macroeconomic Influence of China Futures Market: A GARCH-MIDAS Approach. Advances in Harmony Search, Soft Computing and Applications 15,

Liu, T., Choo, W., Tunde, M. B., Wan, C., & Liang, Y. (2024). Enhancing stock volatility prediction with the AO-GARCH-MIDAS model. *PloS one*, *19*(6), e0305420.

Ma, Y. r., Ji, Q., & Pan, J. (2019). Oil financialization and volatility forecast: Evidence from multidimensional predictors. *Journal of Forecasting*, *38*(6), 564-581.

Mo, D., Gupta, R., Li, B., & Singh, T. (2018). The macroeconomic determinants of commodity futures volatility: Evidence from Chinese and Indian markets. *Economic Modelling*, *70*, 543-560.

Park, B. J. (2002). An outlier robust GARCH model and forecasting volatility of exchange rate returns. *Journal of Forecasting*, *21*(5), 381-393.

Wang, L., Ma, F., Liu, J., & Yang, L. (2020). Forecasting stock price volatility: New evidence from the GARCH-MIDAS model. *International Journal of forecasting*, *36*(2), 684-694.

Wu, C., Wang, X., Luo, S., Shan, J., & Wang, F. (2020). Influencing factors analysis of crude oil futures price volatility based on mixed-frequency data. *Applied Sciences*, *10*(23), 8393.

Yaya, O. S., Ogbonna, A. E., & Vo, X. V. (2022). Oil shocks and volatility of green investments: GARCH-MIDAS analyses. *Resources Policy*, *78*, 102789.

Yu, X., & Huang, Y. (2021). The impact of economic policy uncertainty on stock volatility: Evidence from GARCH–MIDAS approach. *Physica A: Statistical Mechanics and its Applications*, *570*, 125794.

Zhang, M., Zhao, Y., & Nan, J. (2022). Economic policy uncertainty and volatility of treasury futures. *Review of Derivatives Research*, *25*(1), 93-107.