Received: 22 Sep. 2025 | Revised: 24 Sep. 2025 | Accepted: 26 Sep. 2025 | Published online: 15 Oct. 2025

RESEARCH ARTICLE

Journal of Global Humanities and Social Sciences
2025,Vol. 6(6)307-312
DOI: 10.61360/BoniGHSS252018960610

Forecasting oil futures price volatility with geopolitical risk: A SV-MIDAS model

Rongrong Tu[1,*]

[1] *School of Finance, Anhui University of Finance and Economics, China*

Abstract: This paper proposes the stochastic volatility model with mixed data sampling and geopolitical risk (SV-MIDAS-GPR) for modeling and forecasting the volatility of INE crude oil futures. The model is capable of capturing both the influence of geopolitical risk on the volatility of INE crude oil futures and the high continuity of volatility. We develop the maximum likelihood method based on continuous particle filters to estimate the model parameters. Our findings demonstrate that the SV-MIDAS-GPR model surpasses a variety of benchmark models excels in both within-sample accuracy and predictive power for volatility, including the GARCH model, the stochastic volatility (SV) model, and the SV-MIDAS model, thus highlighting the value including the addition of both the element volatility (MIDAS) framework and geopolitical risk into volatility modeling and forecasting.

Keywords: geopolitical risk, volatility forecasting, Chinese crude oil futures, SV-MIDAS

1. Introduction

Crude oil, as a critical strategic resource and energy commodity, significantly impacts financial markets and the development of the real economy. Crude oil futures play a pivotal role in the crude oil market, representing the most actively traded commodity derivatives, with price fluctuations reflecting investors' expectations of market uncertainty. Moreover, the fluctuations in oil futures prices are vital for energy trading, portfolio refinement, financial risk handling, options valuation, and strategic speculation planning. In recent years, the frequent occurrence of international geopolitical risk(GPR) events has severely affected the crude oil market. The occurrence of conflict events has led to disruptions in crude oil supply, altering trade volumes and subsequently impacting the price volatility of crude oil futures. Therefore, accurately measuring and modeling the volatility of crude oil futures from a geopolitical risk perspective is of great significance and has garnered widespread attention from both the academic and practical communities(zhang et al., 2024; Liu et al., 2025; Li et al., 2024).

Since the seminal work by Bollerslev (1986), the generalized autoregressive conditional heteroskedasticity (GARCH) models have been widely used to modeling and forecasting crude oil market volatility. However, the return-based GARCH models is treated as a deterministic function of a set of historical information. In contrast, stochastic volatility (SV) models assume that the dynamics of volatility are described by a first-order autoregressive process and are unobservable. The inclusion of a news process in the volatility equation of SV models renders them more flexible than GARCH models, thereby providing superior within-sample adaptation and predictive accuracy for future estimates (Yu, 2005).

However, the SV model is part of the single-component volatility models category, potentially inadequate for capturing the elevated continuity, or long memory, characteristic of financial volatility. To address this issue, Shang alongside Liu (2017) and Shang in collaboration with

Corresponding Author: Rongrong Tu
School of Finance, Anhui University of Finance and Economics, China

©The Author(s) 2025. Published by BONI FUTURE DIGITAL PUBLISHING CO.,LIMITED This is an open access article under the CC BY License(https://creativecommons.org/licenses/by/4.0/).

Zheng (2021) introduced the SV-MIDAS model. This framework divides volatility into a random (short-term) segment and a consistent (long-term) segment. The latter, which is driven by a low-frequency variable is modelled using the MIDAS structure. Shang and Liu (2017) along with Shang and Zheng (2021) demonstrate that the SV-MIDAS model effectively captures the fluctuating long-term element and surpasses the standard SV model regarding within-sample fit accuracy. Recently, Wang et al. (2024) propose an examination of the effects of economic policy uncertainty on the volatility of China's stock market, utilizing the SV-MIDAS model based on the t-distribution framework.

Although the SV-MIDAS model adeptly captures the persistent aspect of volatility, it fails to account for the influence of macroeconomic indicators, specifically geopolitical risk (GPR), on volatility. Over the long term, the measurement of geopolitical risk has been relatively challenging. Despite the widespread recognition of its impact on the macroeconomy and financial markets, there is a scarcity of empirical studies on this subject. Only after Caldara and Iacoviello (2022) developed the Geopolitical Risk (GPR) index, grounded on newspaper news that scholars were able to more readily investigate the effects of geopolitical risk.Considering geopolitical risk as an significant source of uncertainty and its more direct connection to oil supply, it is reasonable to infer that geopolitical risk may exert a great influence on the volatility of China's crude oil futures market. Therefore, incorporating geopolitical risk into the model may further enhance the precision and reliability of volatility forecasting in China's crude oil futures market.

Motivated by the aforementioned insights, this paper introduces the Stochastic Conditional Mixed Data Sampling (SV-MIDAS-GPR) model, which integrates perspectives from both the GARCH-MIDAS and SV models. A principal attribute of the suggested model is its capacity to account for the extended memory of daily volatility in INE crude oil futures, along with geopolitical risk (GPR). The principal contributions of this study can be summarized in three key aspects. Firstly, we develop the SV-MIDAS-GPR model to model the volatility of INE crude oil futures. It is important to note that the SV-MIDAS-GPR model is an extension of the SV model, inheriting its original advantages, particularly its ability to capture the random information flow that is difficult to observe directly in financial markets. More importantly, the SV-MIDAS-GPR model draws inspiration from the GARCH-MIDAS framework by Engle et al. (2013), which divides the volatility of INE crude oil futures into a short-term element and a long-term element through a multiplicative approach. The short-term element refers to assumed to follow a latent process, while the long-term component is modeled through the MIDAS regression approach. Consequently, our proposed SV-MIDAS-GPR model has the capacity to capture both the random information and the high persistence in the volatility of INE crude oil futures. In addition, the SV-MIDAS-GPR model takes into account the influence of macroeconomic variables (GPR) on the volatility of INE crude oil futures. Finally, we utilize the suggested SV-MIDAS-GPR model to predict the volatility of INE crude oil futures. The empirical results indicate that volatility of INE crude oil futures exhibits strong time-varying, clustering, and long-memory characteristics, and is also influenced by geopolitical risk (GPR). We compare the SV-MIDAS-GPR model with the GARCH, SV, and SV-MIDAS models as a standard. The findings indicate that the SV-MIDAS-GPR model enhances the accuracy of out-of-sample volatility predictions compared to the reference models.

The subsequent sections of this paper are structured accordingly. Section 2 offers an exhaustive overview of the SV-MIDAS-GPR model, including its theoretical foundation and the rationale for integrating geopolitical risk factors with the MIDAS framework. Section 3 presents the empirical analysis, where we apply the model to the INE crude oil futures market and compare its performance with other established models. Finally, Section 4 summarizes the key findings and concludes the paper, highlighting the implications for future research and practical applications.

2. The Model

2.1. The SV-MIDAS-GPR model

Shang and Liu (2017) propose the SV-MIDAS model. The SV-MIDAS model assumes that $r_{i,t}$ follows the following process

$$r_{j,t} = \exp(\xi_{j,t}/2)\varepsilon_{j,t}, \quad \xi_{j,t} = \tau_t + \lambda_{j,t} \quad (1)$$

$$\lambda_{j,t} = \beta\lambda_{j-1,t} + \eta_{j,t} \quad (2)$$

$$\tau_t = m + \theta\sum_{k=1}^{K}\varphi_k(\gamma)\log(RV_{t-k}) \quad (3)$$

where $r_{i.t}$ is the log-return on day j during month t. It is clear that in the SV-MIDAS model, the logarithm of volatility $\xi_{i.t}$ is separated into two segments: a daily transient stochastic segment $\lambda_{i.t}$ and a monthly enduring stable segment τ_t. The daily transient stochastic segment $\lambda_{i.t}$ follows a Gaussian AR(1) procedure. To ensure the stationarity of $\lambda_{i.t}$ process, we impose the restriction |β| < 1. It should be emphasized that the long-term stable segment τ_t is defined by smoothing the low-frequency variable, which is the logarithm of monthly realized volatility (RV), in accordance with the approach of MIDAS regression. The monthly RV is defined as

$$RV_t = \sum_{i=1}^{N_t} r_{j,t}^2 \quad (4)$$

where N_t the count of trading days during month t. φ_k is the Beta weighting formula, which is expressed as

$$\varphi_k(\gamma) = \frac{(1-\frac{k}{K})^{\gamma-1}}{\sum_{j=1}^{K}(1-\frac{j}{K})^{\gamma-1}} \quad (5)$$

$$\varepsilon_{j,t} \sim i.i.d.N(0,1) \quad \eta_{j,t} \sim i.i.d.N(0,\sigma^2) \quad (6)$$

where K is the maximum MIDAS lag order and $0 \le \varphi_k(\gamma) < 1$. Within this study, we opt for K = 36, meaning that we employ three years' worth of MIDAS lags for the monthly realized volatility to refine the long-term stable element. The coefficient γ governs the decay rate of the weighting function. To ensure that the weighting function is monotonically decreasing, such that more recent observations assign greater weight, we impose the constraint $\gamma > 1$. $\varepsilon_{i.t}$ and $\eta_{i.t}$ are independently separate from each other, and $\varepsilon_{i.t}$ follows a standard exponential distribution (with unit mean). Although the conventional SV-MIDAS model is capable of depicting the fluctuating long-term aspect of volatility, it overlooks the driving effect of macroeconomic variables. In particular, a significant number of scholars have recently demonstrated that GPR has a notable impact on INE crude oil futures volatility. In light of this, to investigate the explanatory and predictive power of GPR on crude oil futures volatility, this paper further introduces GPR into the standard SV-MIDAS model to expand the long-term component process (7) as follows:

$$\tau_t = m + \theta_1\sum_{k=1}^{K}\varphi_k(\gamma_1)\log(RV_{t-k}) + \theta_2\sum_{k=1}^{K}\varphi_k(\gamma_2)\log(GPR_{t-k}) \quad (7)$$

The aforementioned enhancement of the traditional SV-MIDAS model further underscores the influence of GPR on the long-term volatility segment τ_t. It should be observed that under certain conditions $\theta_2 = 0$, the extended SV-MIDAS model (SV-MIDAS-GPR model) reverts to the standard SV-MIDAS model.

2.2. Estimation method

In the context of the SV-MIDAS-GPR model, the presence of latent state variables poses a significant challenge. Specifically, these latent variables introduce a level of complexity that prevents the likelihood function from being solved directly. As a result, the conventional maximum likelihood estimation method, which relies on the explicit form of the likelihood function, becomes inapplicable for estimating the model parameters. To surmount this hurdle, we have devised an innovative strategy. We introduce the continuous sampling importance resampling (CSIR) method, which is specifically designed to handle the intricacies associated with latent state variables. Through the use of the CSIR method, we successfully calculate the likelihood function of the SV-MIDAS-GPR model robustly and effectively. Once the likelihood function is successfully computed, we subsequently apply the maximum likelihood estimation to derive precise parameter estimates for the model. This two-step process not only addresses the issue of latent state variables but also ensures that the parameter estimation is both reliable and valid, thereby enhancing the overall performance and applicability of the SV-MIDAS-GPR model.

3. Empirical Analysis

3.1. Data

In conducting our data-driven study, we leverage information on daily closing prices for INE crude oil futures and infrequent monthly GPR index data. The sample period spans from March 26, 2018 to August 29, 2025. All data are obtained from the Wind Database of China. The daily return of INE crude oil future is calculated using $r_i = \log(P_i) - \log(P_{i-1})$ where P_i denotes the closing price on day t.

Table 1 presents the descriptive statistics for regarding the daily returns of INE crude oil futures and the monthly GPR. As shown in the table, the return distribution exhibits negative skewness and excess kurtosis, indicating a significant deviation from the Gaussian distribution. The JB test statistic is highly meaningful for the return series, which further confirms the non-normality of the series. The Ljung-Box statistic indicates pronounced serial correlation within the return series, suggest- ing a high persistence of volatility. The descriptive statistics of the GPR show that the GPR index has a high mean and large volatility, indicating frequent geopolitical adjustments.

3.2. Estimation results

We employ the maximum likelihood method based on CSIR with M=500 particles for determining the parameters of the SV-MIDAS-GPR model. To implement the SV-MIDAS-GPR model, we set the lag order K=36. Conrad and Loch (2015) show that a sufficiently large K allows the lagged variables to automati- cally adjust their weights, thereby enhancing the robustness of the estimation. Therefore, this paper selects K=36, corresponding to 36 monthly lags of the low-frequency variable (historical information from the past 3 years), to estimate the long-term volatility. The parameter estimation outcomes are detailed in Table 2. For comparison, we also estimate the GARCH, SV, and SV-MIDAS models. Table 2 shows that the GARCH and SV models have β estimates close to one, indicating high volatility persistence, which aligns with Ljung-Box statis-tics. The SV-MIDAS and SV-MIDAS-GPR models have significant positive $\theta > 1$ estimates, suggesting a long-term MIDAS component. Notably, these models have lower β estimates than the GARCH and SV models, due to their ability to capture the fluctuating long-term component. The positive and significant $\theta > 1$ estimates imply that higher monthly RV predicts a higher long-term compo- nent, while $\gamma > 1$ estimates greater than one indicate that the effect of monthly RV on the long-term element decreases as the lag increases.

Table 1. Descriptive statistics of daily returns and INE crude oil futures

	Mean	Min.	Max.	Std.	Skewness	Kurtosis	Jarque-Bera	Q(10)
Return	0.0001	-0.2618	0.1011	0.0236	-1.0027	13.5998	8733.1754	14.8957
GPR	113.4549	58.4200	318.9500	40.6524	1.9028	9.2835	202.3665	112.7736

Table 2. Parameter estimation results

	GARCH	SV	.SV-MIDAS	SV-MIDAS-GPR
β	0.9321	0.9308	0.8310	0.8397
	(0.0137)	(0.0064)	(0.0151)	(0.0117)
σ2 (α)	0.1009	0.0929	0.1726	0.1704
	(0.0110)	(0.0049)	(0.0134)	(0.0129)
c(ω)	0.0000	-7.8400	-	-
	(0.0185)	(0.0633)	-	-
m	-	-	-7.8400	-4.3172
	-	-	(0.1017)	(0.0561)
θ1	-	-	0.4541	0.6987
	-	-	(0.0155)	(0.0114)
γ1	-	-	103.2768	40.0981
	-	-	(1.5781)	(0.06181)
θ2	-	-	-	0.0390
	-	-	-	(0.0139)
γ2	-	-	-	124.4556
	-	-	-	(2.2921)
Log-lik	4268.6310	4393.1360	4401.0155	4402.0637

Table 3: Results of out-of-sample volatility forecasting evaluation.

	GARCH	SV	SV-MIDAS	SV-MIDAS-GPR
MSE	6.6086e-05	7.3753e-04	7.2619e-04	**7.1046e-04**
MAE	2.2484e-02	2.1312e-02	2.1031e-02	**2.0823e-02**
QLIKE	-2.9605	-3.9884	-4.0150	**-4.0246**

Note: Entries in bold highlight the model with the smallest loss value (in each row).

In the SV-MIDAS-GPR model, the positive θ_2 stimate indicates that low- frequency monthly GPR increases long-term volatility, reflecting geopolitical risk's impact on oil supply/demand uncertainty and its propagation through financial networks. The λ_2 estimate is significantly lower than λ_1 , indicating that the influence of monthly GPR on long-term volatility persists longer than that of RV. The SV-MIDAS and SV-MIDAS-GPR models also have higher log-likelihood values than the GARCH and SV models, suggesting that the MIDAS structure improves volatility modeling. The SV-MIDAS-GPR model further enhances model fit, as shown by its higher log-likelihood compared to the SV-MIDAS model.

3.3. Out-of-sample results

In empirical applications, investors tend to prioritize the out-of-sample predictive accuracy of models over their in-sample fit, as the former is more indicative of the model's performance in real-world, future scenarios. In view of this, we split the entire sample into two subsamples: a training period from March 26, 2018 to December 29, 2023 (consisting of the first 1400 observations) and an out-of-sample evaluation period from January 2, 2024, to August 29, 2025 (consisting of the remaining 402 observations). This period allows us to develop and refine our model based on historical data. We conduct the out-of-sample forecast procedure using the rolling window approach in which the estimation period is rolled forward daily.

In order to assess the precision of the out-of-sample volatility forecasts generated by various models, this study employs three widely recognized loss functions that are commonly utilized in the field of financial econometrics. These include the Mean Absolute Error (MAE), which gauges the average size of the errors in a set of forecasts, irrespective of their direction; the Mean Square Error (MSE), which assesses the average of the squares of the errors—that is, the average squared difference between the estimated values and the actual value, giving more weight to larger errors; and the Quasi-Likelihood Information Criterion (QLIKE), which is a measure of the goodness of fit that is particularly useful for models that do not belong to the exponential family of distributions. By incorporating these three metrics, we aim to provide a comprehensive evaluation of the forecasting models' performance.The results further underscore the significant improvement in the forecasting ability of the original single-component SV model for volatility with the incorporation of the MIDAS structure within both the SV-MIDAS and SV-MIDAS-GPR models. This also highlights the value of utilizing the MIDAS structure to capture the long-term components of volatility for forecasting INE crude oil futures volatility.. In particular, the introduction of GPR into the SV-MIDAS model further en- hances the precision of the model's volatility forecasts, indicating that GPR contains valuable information that plays a crucial role in predicting volatility. In summary, the SV-MIDAS-GPR model, which integrates both MIDAS and GPR, demonstrates the best volatility forecasting capability.

4. Conclusion

This paper introduces the Stochastic Volatility-Mixed Data Sampling with Geopolitical Risk (SV-MIDAS-GPR) model, a novel approach that integrates geopolitical risk factors with the Mixed Data Sampling (MIDAS) structure to forecast the volatility about INE crude oil futures. The model coefficients are determined via the maximum likelihood estimation technique, which is renowned for its efficiency and accuracy in statistical modeling. The empirical analysis conducted in the INE crude oil futures market demonstrates the effectiveness of incorporating both MIDAS and geopolitical risk factors into volatility modeling and forecasting. Compared to other models such as the GARCH, the SV, and the standard SV-MIDAS models, the SV-MIDAS-GPR model exhibits excellent performance regarding both in-sample fit and out-of-sample forecasting precision. The consistent outperformance of the SV-MIDAS-GPR model suggests its potential for broader financial applications. Future research could further explore its applications in areas such as option pricing and risk assessment, which may provide additional insights into the practical implications of the model in financial decision-making processes.

Conflict of Interest

The author declares that she has no conflicts of interest to this work.

Acknowledgement

This research was funded by: Innovative Research Project for Graduates of Anhui University of Finance and Economics under Grant No. ACYC2023203

References

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedastic- ity. *Journal of econometrics 31*(3), 307-327.

Caldara, D., Iacoviello, M., 2022. Measuring geopolitical risk. *American economic review 112*(4), 1194-1225.

Conrad, C., Loch, K., 2015. Anticipating long-term stock market volatility. *Journal of Applied Econometrics 30*(7), 1090-1114.

Liu, L., Li, L., Li, D., Tang, L., 2025. Analysis of geopolitical risk impacts on crude oil volatility with an explainable machine learning approach: China versus the USA. *Applied Economics*, 1-18.

Li, X., Ye, C., Bhuiyan, M. A., Huang, S., 2024. Volatility forecasting with an extended GARCH-MIDAS approach. *Journal of Forecasting 43*(1), 24-39.

Shang, Y., Liu, L., 2017. An extension of stochastic volatility model with mixed frequency information. *Economics Letters 155*, 144-148.

Shang, Y., Zheng, T., 2021. Mixed-frequency SV model for stock volatility and macroeconomics. *Economic Modelling 95*, 462-472.

Wang, N., Yin, J., Li, Y., 2024. Economic policy uncertainty and stock market volatility in China:Evidence from SV-MIDAS-t model. *International Review of Financial Analysis 92*, Article 103090.

Yu Jun. On leverage in a stochastic volatility model. *Journal of Econo-metrics, 2005, 127*(02): 165-178.

Zhang, J., Xiang, Y., Zou, Y. and Guo, S., 2024. Volatility forecasting of Chinese energy market: Which uncertainty have better performance? *International Review of Financial Analysis 91*, Article 102952.

How to Cite: Tu, R. (2025). Forecasting oil futures price volatility with geopolitical risk: A SV-MIDAS model. *Journal of Global Humanities and Social Sciences*, *6*(6), 307-312
https://doi.org/ 10.61360/BoniGHSS252018960610