Integrating High-Frequency Data In Volatility Prediction: A DCC-GARCH-MIDAS Approach

Shuang Jin[1*], Wei Chong Choo[1], Matemilola Bolaji Tunde[1], Wan Cheong Kin[2], and Pengrui Jin[3]

[1] School of Business and Economics, Universiti Putra Malaysia.
[2] Department of Economics and Corporate Administration, Faculty of Accountancy, Finance and Business, Tunku Abdul Rahman University of Management and Technology (TARUMT).
[3] University of Birmingham, Edgbaston, Birmingham.
[*] Corresponding author. E-mail: gs62875@student.upm.edu.my

Received: Jan. 29, 2023; Accepted: Apr. 14, 2024

Traditional econometric models are restricted in their capacity to examine same-frequency data, resulting in the loss of valuable information from high-frequency data. To address this problem, We propose the DCC-GARCH-MIDAS model by combining dynamic conditional correlation modeling with information from high-frequency data. For our empirical study, we utilised historical data from 2000 to 2019 as in-sample data and trained a model for predicting volatility. We applied the trained model to forecast data from 2020 to 2022, calculating the discrepancy between predicted volatility and actual observations, and comparing differences between the predicted and actual values. The research findings not only enhance comprehension of the correlation between macroeconomics and financial market instability but also propose a distinct strategy for resolving the problem of incongruent data frequencies.

Keywords: DCC, MIDAS, GARCH, Long-run correlation, Macroeconomic variables.

http://dx.doi.org/10.6180/jase.202505_28(5).0013

1. Introduction

Macroeconomic variables serve as pivotal indicators that encapsulate the economic status and performance of a specific region or country within a given timeframe[1]. Essential metrics, such as Gross Domestic Product (GDP), quantifying the total value of goods and services produced over a specified period, are integral components of these variables. For policymakers, economists, businesses, and individuals, a comprehensive understanding of these factors is crucial to interpret the current economic landscape, anticipate future developments, and formulate corresponding strategies or policies[2]. (Yu et al., 2023)[3]. The well-established correlation between macroeconomic variables and financial market volatility has been extensively examined in literature. Numerous studies have scrutinized the impact of factors such as interest rates, inflation, GDP growth, and fiscal policies on the volatility of financial assets (Wang et al., 2023)[4].

Engle et al. (2004)[5] propose a model that conceptualizes equity volatilities as a synthesis of macroeconomic effects and time series dynamics. This model enables long horizon forecasts of volatility to be contingent upon macroeconomic developments, providing estimates of the anticipated volatility in a newly opened market. However, a notable challenge arises in studying this relationship due to the inherent disparity between high36 frequency and low-frequency data. According to Diebold et al. (2008)[6], there exists a direct correlation between stock market volatility and macroeconomic fundamentals, with turbulent fundamentals precipitating volatile stock markets. (Malik et al., 2011)[7] propose likelihood-based inference of unknown parameters for a class of discrete-time stochastic volatility (SV) models. High-frequency data capture minute-by-minute or point-by-point fluctuations, while low-frequency data (usually provided on a daily, monthly or quarterly basis) provide aggregated values [8].

(Broadstock et. al., 2019)[9] investigate the determinants of correlation patterns between green and black bond markets. In order to determine the determinants of market correlations, a two-step sequential methodology was used to estimate both the time-varying correlations and determinants. First, dynamic conditional correlations (DCC) were extracted, and in the second stage, dynamic model averaging (DMA) was applied. (Ma et. al., 2019) [10] build three predictors of crude oil price volatility using the generalized dynamic factor model (Ma et al., 2019): a basic (physical) predictor, a financial predictor, and a macroeconomic uncertainty predictor. In 2019, the findings of (Akkoc et al.) [11] suggest that Turkey needs flexible macroeconomic policies to manage the consequences of volatile shocks in the aftermath of the world financial crisis.

In order to capture intraday patterns in stock market oscillations prediction, (Sun et al., 2019) [12] developed a machine learning method called ARMA-GARCH-NN. This confirms the efficacy of ARMA-GARCH-NN in identifying patterns in massive stock data without making strong assumptions about the distribution. The goal of (Wang et al., 2019) [13] is to argue that the volatility description is more explanatory and the volatility model is more rational. Whereas low-frequency data is not fully utilized, high-frequency data contains a variety of information. Based on high-frequency financial data, several new large volatility matrix estimate techniques have been created. (Song and others, 2019) [14] presents a unified approach to modelling high-frequency financial data that can accommodate both discrete-time realised GARCH models and continuous-time jump-diffusion models. These models are based on the combination of non-parametric volatility obtained from option data estimators and daily composite volatility estimated from the high-frequency data. The increased availability of high-frequency data has led to many new areas of research in statistics.Consistency and asymptotic normality results are derived for an alternative quasi-likelihood approach promoted by (Cerovecki et al., 2019)[15]. (Chen et al., 2019) [16]give a computer-intensive approach to multi-step forward forecasting of the volatility of financial return series under ARCH/GARCH models and in a model-free context, i.e., utilizing the NoVaS transformation. Under various data-generating procedures, the effectiveness of GARCH models and the model-free method for multi-step forward prediction was also contrasted. (Wang et al., 2020) [17]found that high volatility levels, negative crude oil RVs, and high-frequency data all represent the main predictive power of crude oil RVs.The hybrid technique proposed by Gao et al. (2021) [18] combines the nonlinear modeling capabilities of Long-Short Term Memory (LSTM) neural networks with the benefits of non-stationary parametric models like Generalized Autoregressive Conditional Heteroskedasticity (GARCH). The suggested architecture is also applicable to other applications where the primary features are high volatility and shortage of data. A two-step procedure was used in the empirical implementation of (Yao et al., 2022) [19]. Firstly, it looked into the possibility of a GARCH effect in the observed PM2.5 data. The study's findings suggest the appropriateness of using SVM-based models for PM2.5 prediction and the possibility of enhancing prediction accuracy with the integration of GARCH model.

Because of information asymmetry, (Yasar et al., 2020) [20] seek to validate and expand signaling theory. This models assume fixed and unilateral relationships between variables, making them unsuitable for exploring the dynamic connections between macroeconomic factors and market volatility. Furthermore, their applicability is limited to scenarios where data frequencies are identical, which can constrain their usefulness when high-frequency data is influential.The purpose of (Dufitinema, 2021) [21] is to evaluate the accuracy of several models in modeling and forecasting the volatility and returns of Finnish housing prices. (Hansen, 2021) [22] demonstrate how macroeconomic risk and the U.S. Treasury yield curve have a changing connection over time.

One notable area of research involves the application of advanced imaging techniques in the medical domain. Laghari and Yin (2022) addressed the complex challenges associated with collecting and interpreting medical images obtained in highly challenging environments, ranging from the nanoscale to hyperspectral imaging. Their work underscores the importance of robust data collection and interpretation methodologies in extracting meaningful insights from medical imaging data, thereby inspiring advancements in other fields such as financial econometrics.

Furthermore, Laghari et al. (2023a)[23] proposed a novel deep learning approach for atrial fibrillation detection based on a deep residual-dense network integrated with a bidirectional recurrent neural network (RNN). By leveraging the temporal dependencies within electrocardiogram (ECG) signals, their model achieved promising results in automated arrhythmia detection, highlighting the potential of deep learning techniques in medical signal analysis.In addition to medical imaging and signal processing, advancements in video streaming technologies have also contributed to the proliferation of high-frequency data sources. Laghari et al. (2023b)[24] conducted a comprehensive review of the state-of-the-art video streaming technologies, elucidating the underlying principles and emerging trends in this rapidly evolving field. Their insights into video streaming architectures and protocols provide valuable context for understanding the complexities of processing high-frequency data streams in real-time applications.Moreover, Karim et al. (2022)[25] conducted a thorough review of hyperspectral imaging techniques and their applications in medical diagnostics. By leveraging the rich spectral information provided by hyperspectral imaging, researchers can enhance tissue characterization and disease detection capabilities, paving the way for more accurate and reliable medical diagnoses.Furthermore, Das et al. (2023)[26] introduced Eldo-Care, a telehealthcare system that combines EEG with Kinect sensor technology to provide remote monitoring and assistance for disabled and elderly individuals. By integrating multimodal sensor data, Eldo-Care enables real-time monitoring of physiological signals and physical movements, thereby facilitating early intervention and personalized care delivery.Saeed et al. (2023) [27]proposed DeepLeukNet, a deep learning model based on convolutional neural networks (CNNs) for acute lymphoblastic leukemia (ALL) classification using microscopy images. Their model demonstrates the potential of CNN-based approaches in automating the diagnosis of hematologic malignancies, thereby reducing the burden on healthcare professionals and improving diagnostic accuracy.

Volatility prediction in financial markets has garnered substantial attention due to its crucial role in risk management, investment strategies, and policy formulation. In this context, our study focuses on advancing volatility prediction methodologies by integrating high-frequency data using a Dynamic Conditional Correlation-Generalized Autoregressive Conditional Heteroskedasticity-MIXING DATA Sampling (DCC-GARCH-MIDAS) approach. Our research builds upon existing literature and extends the scope of analysis by addressing key challenges in volatility prediction. The COVID-19 pandemic has presented unprecedented challenges to energy markets, disrupting traditional dynamics and necessitating innovative analytical frameworks to assess its impact. The paper by (Andreani,2021) [28] demonstrates the effects of the COVID-19 pandemic on energy commodities using a mixed-frequency approach that incorporates information from COVID-19-related data. This study highlights the significance of incorporating external factors, such as pandemic-related variables, in volatility modeling to capture nuanced market behaviors accurately. (Wu,2023) [29] further investigates the impact of recent crisis shocks on these dynamic correlations by studying the dynamic correlations of cross-assets through a multivariate GARCH framework. In addition, the analysis uncovers how macroeconomic factors affect the correlations of cross-assets and how they are linked to the corresponding crises. Moreover, (Wu,2023) [29] contributes to the understanding of dynamic correlation among cross-assets in the economy and finance, particularly during crisis periods. By employing corrected GARCH-DCC-MIDAS models, the study elucidates the impact of macroeconomic factors on cross-asset connectedness and explores the implications for risk management and investment strategies. This research underscores the importance of considering macroeconomic variables in volatility modeling to provide robust insights into market dynamics.

Furthermore, (Wang et al,2024) [30] investigate the asymmetric volatility of the Chinese stock market in response to economic uncertainty using an asymmetric GARCH-MIDAS model. Their findings reveal the differential impact of positive and negative news on market volatility and offer valuable insights for policymakers, investors, and regulators in mitigating systemic financial risks. This study underscores the relevance of asymmetric volatility modeling in capturing the complex interplay between economic uncertainty and stock market fluctuations.

Additionally, (Li,2024) [31]extends the GARCH-MIDAS framework to incorporate geopolitical risk uncertainty in volatility forecasting for the Shanghai Stock Exchange. By considering various extended models, the study evaluates the efficacy of incorporating economic policy uncertainty and geopolitical risk uncertainty in volatility prediction. The research findings provide practical implications for market participants in navigating volatile market conditions and implementing risk management strategies effectively.

In line with these scholarly endeavors, our paper proposes a novel DCC-GARCH-MIDAS approach to integrate high-frequency data into volatility prediction. By leveraging dynamic conditional correlation modeling alongside high-frequency data, our methodology aims to enhance the accuracy and informativeness of volatility forecasts. We believe that our research contributes to the advancement of volatility prediction methodologies and offers valuable insights for risk managers, investors, and policymakers in assessing and managing market risks effectively.

The fundamental purpose of our empirical investigation is to create a solid framework for predicting volatility.To achieve this, we utilize historical data covering the period from 2000 to 2019 as our in-sample dataset. Through careful training and calibration, we devise a model for forecasting volatility that can effectively capture the subtle correlation between macroeconomic indicators and financial market volatility. A thorough evaluation follows, wherein predicted volatility is compared to actual observations. By conducting a quantitative analysis of the disparities between our predictions and actual market volatilities, we intend to assess the practical usability and efficacy of our proposed paradigm. This study endeavors to not only enhance our comprehension of the complex correlations between macroeconomic aspects and financial market volatility but also introduces an innovative methodology that addresses the challenges posed by varying data frequencies.

2. Methodology

This section meticulously elucidates the theoretical underpinnings of the DCC-GARCH-MIDAS model, explaining the intricacies of the model construction process and detailing the methods used for training and forecasting. In doing so, it provides a thorough and nuanced understanding of the research methodology, shedding light on how the study was conducted and how the proposed model was skilfully applied to overcome the challenges posed by varying data frequencies. The construction of the DCC-GARCH-MIDAS model is a meticulous and systematic multi-stage process. The first stage involves the pre-processing of high frequency data, a critical step to capture relevant features and establish correlations. The DCC component is then used to estimate the conditional correlation matrix. This matrix serves as a reflective representation of the intricate relationships that exist between various macroeconomic variables and market returns. This is followed by the GARCH component, which carefully estimates the conditional variances associated with the financial market. In the final integration, the MIDAS component harmonises these meticulously derived elements. This holistic integration culminates in a deep and comprehensive understanding of how macroeconomic events exert their influence on market volatility across the spectrum of different data frequencies. The richness of detail provided by each component contributes to the robustness and effectiveness of the DCC-GARCH-MIDAS model in dealing with the complexities inherent in different data frequencies.

2.1. GARCH-MIDAS Model

Engle et al. (2013) proposed the GARCH-MIDAS model after applying the MIDAS model to the GARCH. The model divides financial asset volatility into short-term and long-term components. To investigate the impact of exogenous variables on the volatility of financial assets and increase the effectiveness of parameter estimation, the short-term component is composed of GARCH(1,1), and the long-term component is composed of realized volatility and exogenous variables. The frequency of data is erratic. The following are the specific contents of the GARCH-MIDAS model:

$$r_{i,t} = \mu_t + \sqrt{\tau_t}\cdot \quad (1)$$

$$\sigma^2_{i,t} = \tau_t \cdot g_{i,t}\cdot \quad (2)$$

r_i, t is the return rate on financial assets on the i day of the t month, quarter, or year. The conditional standard normal distribution is followed by the disturbed term i, t. Conditional volatility is decomposed into tau_t and g_i, t in the classic GARCH model, with tau_t representing the long-term component of volatility and g_i, t representing the short-term component of volatility, indicating the GARCH (1,1) process:

$$g_{i,t} = (1-\alpha-\beta) + \alpha\frac{(r_{i-1,t}-\mu_t)^2}{\tau_t} + \beta g_{i-1,t}\cdot \quad (3)$$

The long-term component z consists of Realised volatility (RV) and the exogenous variable x, The specific expression for the long-term component z is

$$\tau_t = m + \theta\sum_{k=1}^{K}\varphi_{1k}(w_{11}, w_{12})RV_{t-k} + \theta_2\sum_{k=1}^{K}\varphi_{2k}(w_{21}, w_{22})X_{t-k}\cdot \quad (4)$$

$$RV_{t-k} = \sum_{i=1}^{N_t} r^2_{i,t}\cdot \quad (5)$$

$theta_1$ and $theta_2$ illustrate how monthly realized volatility and economic uncertainty respectively affect the long-term component of volatility.K is the maximum lag order of the Beta weight function, $\varphi_k(w_1, w_2)$ is the weight function, in this paper the weight function takes the form of Beta weights, and $\varphi_k(w_1, w_2)$ takes the specific form of:

$$\varphi_k(w_1, w_2) = \frac{(k/K)^{w_1-1}(1-k/K)^{w_2-1}}{\sum_{j=1}^{K}(j/K)^{w_1-1}(1-j/K)^{w_2-1}}\cdot \quad (6)$$

In order to satisfy the law of information decay in financial markets, i.e., information from more recent periods has a greater impact on the current, here in this paper, we refer to Engle (2013) [32], where the first parameter of the two-parameter weight function is fixed to 1, $w_{11} = w_{21} = 1$, from which the weight function is in a decaying state. The long-term component's impact of exogenous variables can also be assessed using the coefficients $w_{12}andw_{22}$. The following is the simplified expression for the Beta weight function:

$$\varphi_k(w_2) = \frac{(1-k/K)^{w_2-1}}{\sum_{j=1}^{K}(1-j/K)^{w_2-1}}\cdot \quad (7)$$

The above Eqs. (1)-(5) jointly form the GARCH-MIDAS model. In this paper, the quasi-great likelihood estimation method to estimate the parameters of the model, and the quasi-great likelihood function is specifically expressed as:

$$\mathrm{LLF} = -\frac{1}{2}\sum_{t=1}^{T}\left[\log g_t(\Phi)\tau_t(\Phi) + \frac{(r_t-\mu)^2}{g_t(\Phi)\tau_t(\Phi)}\right]\cdot \quad (8)$$

However, the GARCH-MIDAS model is comparatively uncomplicated in modeling conditional correlations and typically does not encompass modeling volatility. Consequently, it inadequately depicts events such as volatility aggregation and abrupt fluctuations in financial markets.

2.2. DCC-GARCH-MIDAS Model

Colacito et al. (2011)[33] obtained the DCC-MIDAS model by taking the DCC model into consideration and applying it to the GARCHMIDAS regression setup, which makes use of the conditional covariance equation that

$$q_{o,c,t} = \bar{\rho}_{o,c,t}(1-a-b) + a\xi_{o,t-1}\xi_{c,t-1} + b\left(q_{o,c,t-1}\right)\cdot \quad (9)$$

where

$$\bar{\rho}_{o,c,t} = \sum_{k=1}^{K}\varphi_k(w_1,w_2)C_{o,c,t-1}\cdot \quad (10)$$

And

$$C_{o,c,t} = \sum_{k=t-N}^{t}(\xi_{o,k}\xi_{c,k})\left(\sqrt{\sum_{k=t-N}^{t}\xi_{o,k}^2}\cdot\sqrt{\sum_{k=t-N}^{t}\xi_{c,k}^2}\right)^{-1}\cdot \quad (11)$$

The DCC-GARCH-MIDAS model integrates elements from several key methodologies to address the challenge of different data frequencies effectively: a. Dynamic Conditional Correlation (DCC): The DCC component of the model is designed to capture the evolving correlation between high-frequency and low-frequency data over time. It allows for the modeling of time-varying conditional correlations, which is crucial for understanding the interplay between macroeconomic factors and market volatility. b. Generalized Autoregressive Conditional Heteroskedasticity (GARCH)[34]: The GARCH framework forms the basis for modeling the conditional variances of financial market returns. This component accounts for the volatility clustering observed in financial time series data. c. Mixed-Data Sampling (MIDAS)[35]: MIDAS extends the DCC-GARCH model to include both high- and low-frequency data. This inclusion permits the incorporation of timely, fine-grained information while maintaining the structural integrity of the model.

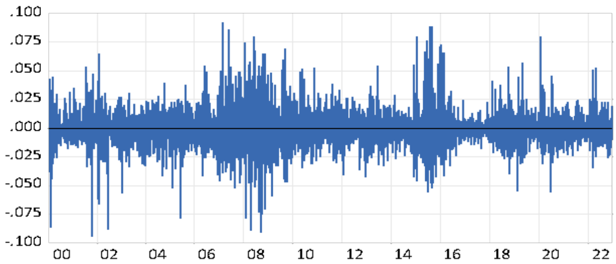

Fig. 1. Daily return on the Shanghai Composite Index

3. Empirical study

This section unfolds the empirical investigation conducted to assess the efficacy of the DCC-GARCH-MIDAS model in addressing the challenge of varying data frequencies, and concurrently, to scrutinize the impact of macroeconomics on financial market volatility.

To rigorously evaluate the model's performance, the dataset is meticulously partitioned into two distinct time periods: an in-sample phase spanning from 2000 to 2019, and an outof-sample phase spanning from 2020 to 2022. The in-sample data is instrumental in training the DCC-GARCH-MIDAS model. During this training process, model parameters are meticulously estimated, facilitating the capturing of dynamic relationships between macroeconomic variables and market volatility. Following the successful training, the model is adeptly applied to the out-of-sample data, allowing for the meticulous forecasting of market volatility throughout the specified period.

A crucial final step involves the computation of the error between the predicted volatility and the actual observations. This meticulous procedure serves as the linchpin in the evaluation of the precision and reliability of the model's forecasts. By shedding light on the model's efficacy in navigating the challenges posed by varying data frequencies, this analysis provides insightful information about the nuanced influence of macroeconomic factors on financial market volatility across the entire research duration. The results of this empirical investigation contribute valuable insights to the broader understanding of the model's applicability and its capacity to unravel the complexities associated with dynamic macroeconomic influences on financial markets.

3.1. Sample Period and Data

In this section, we leverage trading day statistics from the Shanghai Composite Index, Shenzhen Composite Index, and Hang Seng Index, spanning from January 4, 2000, to December 30, 2022, encompassing a total of 5574 observation samples. The data undergoes initial processing by computing the first-order logarithmic difference of prices to

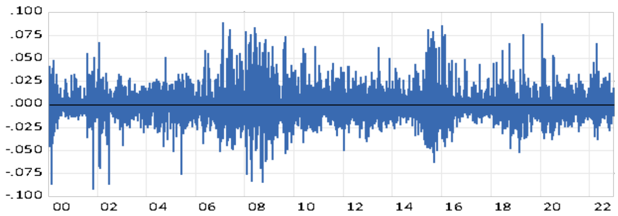

Fig. 2. Daily return on the Shanghai Composite Index

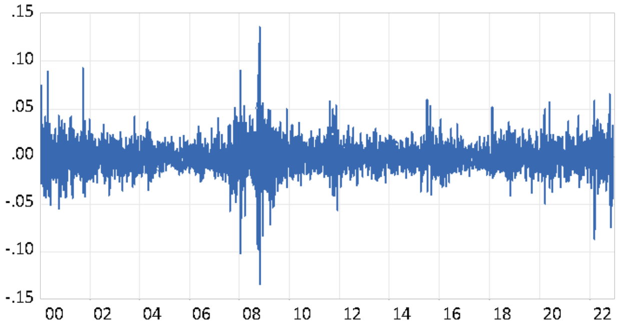

Fig. 3. Hang Seng Index Daily Rerurn

derive the daily returns. Figs. 1 to 3 visually depict the time series of daily returns. From our observations, it becomes evident that the Hang Seng Index exhibits comparatively lower volatility. This inference is visually corroborated by Figures 1 through 3, which highlight the relative stability of the Hang Seng Index's daily returns. To provide a quantitative understanding, descriptive statistics for the three stock indices are tabulated in Table 1. Our analysis of these statistics further reinforces the observation that the Hang Seng Index displays a lower standard deviation compared to the Shenzhen Composite Index and the Shanghai Composite Index. This implies a lower level of volatility in the Hang Seng Index, substantiating the notion that it is comparatively less prone to fluctuations when contrasted with its counterparts.

3.2. Analysis of Forecasting Results

The basic statistical description of each market return is further given in Table 1.The SSE, SZSE, and Hang Seng index market returns are all clearly different from the normal distribution at the 0.01 significant level, according to the Jarque-Bera statistic. Their kurtosis is larger than three, and their sharp peaks are easily noticeable. As a result, their distributions have thick tails and prominent peaks. As a result, the distribution clearly exhibits thick tails and high peaks.

We first estimate the dynamic correlation coefficients of the SSE, SZE, and HSI using the DCC-GARCH model. The results are shown in Table 2. The model's beta coefficients are substantial and successfully pass the significance test, indicating a persistent memory of index fluctuations. This implies that past price variations are linked to the magnitude of their ongoing long-term price changes. Integrality, or the continuous memory of the conditional variance fluctuations, suggests that the securities market responds to outside shocks more slowly and that it is challenging to completely eradicate the stock market once significant swings take place quickly.

Table 2 presents the dynamic correlation estimation findings using the DCC-GARCH model for the Shanghai Composite Index, Shenzhen | Composite Index, and Hang Seng Index. The observation demonstrates that the three indices' and GARCH calculated parameters are significant, and that the sum of both is less than 1. The alpha + beta in the GARCH equation indicates that the conditional variance series of the returns are smooth and the model is predictable. The sum of the parameters a and b in the DCC part is also less than 1, which is consistent with the theory.

$*, ^{**}, ^{***}$ indicate significant at the 10%, 5%, and 1% levels, respectively, with standard errors in parentheses.

Utilizing the derived parameters, we employ them to compute the dynamic correlation coefficient values for the three indices under consideration. The Dynamic Conditional Correlation (DCC) coefficient serves as a metric, gauging the extent of dynamic conditional correlation across diverse markets. Higher DCC values signify heightened inter-market linkages, indicating stronger connections, while lower values suggest weaker correlations. Simultaneously, Table 3 presents the descriptive statistics encapsulating the dynamic correlation coefficient values for the three indices. This tabulated data provides insights into the convergence or divergence of movements across multiple markets, offering a comprehensive perspective on the interplay of these indices.

The correlation coefficient plots of the three indices are given Figs. 4 to 6, and the following conclusions can be drawn:

1. With an average correlation value of 0.9108 and a standard deviation of 0.0632, there is a very significant correlation between the SSE (SSE Composite Index) and the SZE (SZE Composite Index). This implies that there is a strong correlation and near synchronization between the variations of the two indicators.

2. The correlation between the SSE (SSE Composite Index) and the Hang Seng Index is very weak, with an average correlation coefficient of 0.4220 and a standard deviation of 0.2206. this implies that the volatility rela-

Table 1. Descriptive statistics

Variable	Observation	Mean	Std.dev	Min	Median	Max	Skewness	Kurtosis	J-B
SSE	5574	−0.0001	0.0151	−0.094	−0.0006	0.0926	0.377	8.1876	6382.1740***
SZE	5574	−0.0003	0.017	−0.0924	−0.0013	0.0893	0.555	6.4793	3097.7420***
HSI	5668	0.0001	0.0147	−0.1341	−0.0003	0.1358	0.0261	10.6448	13145.3000***

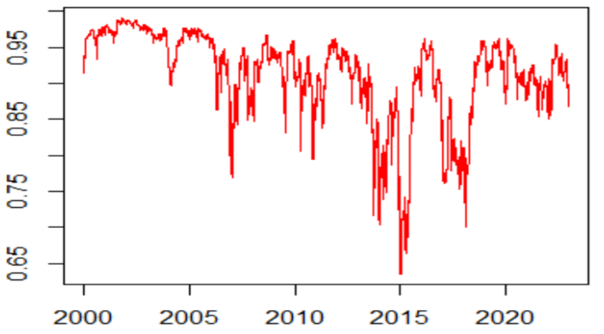

Fig. 4. Dynamic correlation coefficients of SSE and SZE indexes

tionship between the two indices is very close, almost negligible.

3. The correlation between the SZE Index (SZECI) and the Hang Seng Index is also very weak, with an average correlation coefficient of 0.3457 and a standard deviation of 0.1910. This is similar to that between the SSE Index and the Hang Seng Index, which suggests that the volatility of these two indices is not closely related.

   To summarize, the volatility relationship between SSE and SZSE is very strong, while the volatility relationship between SSE and SZSE and Hang Seng is weak.

The utilization of mixed-frequency data models is increasingly prevalent, aiming to harness the wealth of information embedded in high-frequency data and enhance the validity of econometric model estimation and forecasting accuracy. Among these models, the GARCH-MIDAS model stands out as a notable representative of mixed-frequency data modeling. In the context of this research, we extend the DCC-GARCH-MIDAS model, delving into its predictive accuracy concerning mainland and Hong Kong stock market volatility. To facilitate a comprehensive evaluation, the entire dataset is divided into in-sample and out-of-sample segments. The model parameters are then meticulously estimated using the in-sample data, with a specific focus on investigating whether the incorporation of macroeconomic variables into the DCC-GARCH-MIDAS framework can augment the predictive performance of stock volatility. The construction of the DCC-GARCH-MIDAS model involves the daily returns of stocks from Shanghai, Shenzhen, and Hong Kong, alongside relevant macroeconomic indicators.

Subsequently, this study delves into exploring the impact of macroeconomic variables on the long-term dynamic correlation using the DCC-GARCH-MIDAS model, employing stock market volatility as a benchmark. To facilitate a clearer comparison, the parameter estimation results of the DCC-GARCH-MIDAS model are presented in two distinct steps. First, the DCC component parameter estimates are outlined, followed by the GARCH-MIDAS model parameter estimates pertaining to the macro variables.This multifaceted approach allows for a nuanced assessment of the extended DCC-GARCH-MIDAS model's efficacy in capturing the complexities of stock market volatility, shedding light on the potential enhancements brought about by the inclusion of macroeconomic variables.

Based on the estimation outcomes presented in the aforementioned table, the sensitivity of the index return to information is encapsulated by the value of 'a,' specifically 0.0491. Conversely, the memory contribution to volatility, denoted by 'b,' assumes a more substantial role, with a value of 0.8681 . Notably, the sum of ' a ' and ' b ' which approximates 1, signifies the sustained development of return volatility. This observation not only indicates a superior model fit but also underscores a significant volatility link among the indices. The DCC-MIDAS estimation results for the three index returns are detailed below. Within the correlation coefficient series plot, the overarching trend illustrates a coherent pattern across the entire series. It is possible to study the resulting graph in terms of correlation and volatility. In terms of volatility, it is evident that the three indices experienced more intense fluctuations in 2001, 2008, and 2015. In contrast, the Hang Seng Index's long-term return volatility is lower than the SSE and SZE's, while both indexes' long-term return volatility is higher. These differences suggest that the Hang Seng Index is more resilient to crises than the SSE and SZE.

Macro variables such as M2, CPI, Mx, Scg and Fx represent different economic indicators or factors as shown in table15. M2 refers to a measurement of the money supply containing cash, checking deposits and easily convertible near money. The Consumer Price Index, or CPI for short,

Table 2. DCC-GARCH model estimation results

Panel A: GARCH (1,1)						
Stocks	Parameter	Estimate	Std.Error	T-value	$\Pr(>\lvert t\rvert)$	Sig.
SSE	μ	0.0185	0.0161	1.1505	0.25	
	ω	0.02	0.0071	2.8181	0.0048	$***$
	α	0.0835	0.0147	5.6948	0	$***$
	β	0.9115	0.0145	62.6771	0	$***$
SZE	μ	0.0168	0.0195	0.8577	0.3911	
	ω	0.0449	0.0137	3.2796	0.001	$***$
	α	0.0807	0.0134	6.0137	0	$***$
	β	0.9051	0.0157	57.7467	0	$***$
HSI	μ	0.0266	0.0155	1.7165	0.0861	*
	ω	0.0189	0.006	3.132	0.0017	$***$
	α	0.0663	0.0099	6.6878	0	$***$
	β	0.9243	0.0116	79.8487	0	$***$
Panel B: DCC (1,1)						
	Parameter	Estimate	Std.Error	T-value	$\Pr(>\lvert t\rvert)$	Sig.
	a	0.0283	0.0033	8.6788	0	$***$
evaluate	b	0.9698	0.0037	265.1872	0	$***$
	Obs.	LLH	AIC	BIC		
	5398	-21759.36	8.0683	8.0891		

Table 3. Descriptive statistics of dynamic correlation coefficients

DCC	Observation	Mean	Std. dev	Max	Min	Median
SSE & SZE	5398	0.9108	0.0632	0.9908	0.6357	0.9268
SSE & HSI	5398	0.4220	0.2206	0.8032	−0.1918	0.4845
SZE & HSI	5398	0.3457	0.1910	0.7593	−0.1612	0.3663

tracks how prices for a variety of products and services have changed on average over time. Scg refers to the overall revenue generated from the sale of goods to individual consumers within a specific market or region. Fx is an acronym for Foreign Exchange, frequently employed by the financial industry to denote the worldwide market for the purchase and sale of currencies. Mx refers to the money supply of a region. Table 4 to 14 utilize the DCC-GARCH-MIDAS model for introducing monthly low-frequency data on macro variables (M2 CPI Mx Scg Fx) into daily high-frequency index return data for forming mixed-frequency volatility. The tables indicate that the majority of parameters are statistically significant and the a + b values are close to 1, demonstrating that the model performs well and the inter-index return volatility linkage is significant upon inclusion of the macro variables.

The GARCH-MIDAS model is one type of mixed-frequency data model that is being used more and more to fully utilize the rich information found in high-frequency data, as well as to increase the validity of econometric model estimation and forecasting accuracy. The DCC-GARCH-MIDAS model is extended in this work, which further examines the model's predictive accuracy for mainland and Hong Kong stock market volatility. The entire sample data is split into in-sample and out-of-sample data in order to determine whether adding macroeconomic variables to the DCC-GARCH-MIDAS model can enhance the performance of stock volatility prediction. The in-sample data is then used to estimate the model's parameters. Initially, this study will construct the DCC-GARCH-MIDAS model and estimate the associated outcomes using the daily returns of the stocks of Shanghai, Shenzhen, and Hong Kong as well as each macroeconomic indicator. Next, the influence of macroeconomic variables on the long-term realized correlation is investigated using the DCC-GARCH-MIDAS model with realized correlation (RC) as a benchmark. To make the comparison easier, we present the DCC-GARCH-MIDAS models parameter estimation results in two steps: the GARCH-MIDAS models parameter estimates for the macro variables are provided in the first step, and the DCC-MIDAS parts parameter estimates are provided in the second.

1. GARCH-MIDAS model parameter estimates for the macro variables.

2. DCC-MIDAS component parameter estimation results are presented in the second step

With this paragraph, the Dumping Forecast Assessment Indicator is now ready to assess the predictive effect of

Table 4. Parameter estimation results of the DCC-GARCH-MIDAS model:GARCH-MIDAS

Panel A: GARCH-MIDAS						
Stocks	Parameter	Estimate	Std. Error	T-value	Pr(> $\|t\|$)	Sig
SSE	μ	0.017	0.0161	1.0564	0.2908	
	α	0.0937	0.0173	5.4251	0	* * *
	β	0.8686	0.0347	25.0177	0	* * *
	m	-1.596	0.5483	-2.9107	0.0036	* * *
	θ	0.6976	0.1407	4.9587	0	* * *
	ω_2	2.5081	1.2697	1.9754	0.0482	**
SZE	μ	0.0207	0.0199	1.0405	0.2981	
	α	0.0955	0.0169	5.6587	0	* * *
	β	0.8476	0.0314	26.9786	0	* * *
	m	-1.3616	0.5219	-2.6089	0.0091	* * *
	θ	0.6335	0.1373	4.6139	0	* * *
	ω_2	1.9869	0.636	3.124	0.0018	* * *
HSI	μ	0.0282	0.0155	1.8219	0.0686	*
	α	0.0705	0.009	7.8519	0	* * *
	β	0.901	0.0159	56.7117	0	* * *
	m	-1.3284	0.4335	-3.0643	0.0022	* * *
	θ	0.5673	0.1229	4.6146	0	* * *
	ω_2	1.5637	0.4619	3.3857	0.0007	* * *

Table 5. Parameter estimation results of the DCC-GARCH-MIDAS model:DCC-MIDAS

Panel B: DCC-MIDAS						
	Parameter	Estimate	Std. Error	T-value	Pr(> $\|t\|$)	Sig.
	a	0.0491	0.0047	10.3583	0	***
	b	0.8681	0.0195	44.4944	0	***
		1.9028	0.5034	3.78	0.0002	* * *
evaluate	ω_2					
	Obs.	LLH	AIC	BIC		
	5167	−1913.066	3832.132	3851.782		

Table 6. Parameter estimates for the DCC-GARCH-MIDAS model introduced by M2: GARCH-MIDAS

Panel A: GARCH-MIDAS						
Stocks	Parameter	Estimate	Std. Error	T-value	Pr(> $\|t\|$)	Sig.
SSE	μ	0.0164	0.0163	0.8951	0.3707	
	α	0.0805	0.0145	5.538	0	* * *
	β	0.9087	0.0162	56.1385	0	* * *
	m	0.4357	0.3614	1.2056	0.228	
	θ	0.026	0.0121	2.1523	0.0314	* *
	ω_2	1	1.7675	0.5658	0.5716	
SZE	μ	0.0197	0.0215	0.9154	0.36	
	α	0.0652	0.0123	5.2955	0	* * *
	β	0.9204	0.0155	59.4496	0	* * *
	m	0.7602	0.2464	3.0854	0.002	* * *
	θ	0.0165	0.0086	1.9275	0.0539	*
	ω_2	1.0015	0.8101	1.2363	0.2164	
HSI	μ	0.0356	0.0159	2.2373	0.0253	* *
	α	0.0657	0.0086	7.6351	0	* * *
	β	0.9221	0.0106	87.2285	0	* * *
	m	0.8712	0.3114	2.7976	0.0051	* * *
	θ	-0.0175	0.0143	-1.2289	0.2191	
	ω_2	1.0002	2.1601	0.463	0.6433	

Table 7. Parameter estimates for the DCC-GARCH-MIDAS model introduced by M2: DCC-MIDAS

Panel B: DCC-MIDAS						
	Parameter	Estimate	Std. Error	T-value	$Pr(>\|t\|)$	Sig.
	a	0.0423	0.0001	0.051	0	* * *
	b	0.9239	0.0001	0.0001	0	* * *
	$\omega_2 \cdot RC$	1.0595	0.0823	0.1287	0	* * *
	$\omega_2 \cdot M2$	4.3361	0.0252	0.0172	0	* * *
	$m[1,2]$	-1.7394	0.1145	-0.1519	0	* * *
	$m[1,3]$	0.3964	0	0	0	* * *
	$m[2,3]$	0.2702	0.0063	0.4292	0	* * *
	$\theta \cdot$ RC$[1,2]$	3.4841	0.1249	0.2789	0	* * *
	$\theta \cdot$ RC$[1,3]$	0.5073	0	0	0	* * *
	$\theta \cdot$ RC$[2,3]$	0.3364	0.0149	0.2252	0	* * *
	$\theta \cdot$ M2$[1,2]$	0.091	0.0005	0.0197	0	* * *
	$\theta \cdot$ M2$[1,3]$	-0.2383	0	0	0	* * *
	$\theta \cdot$ M2$[2,3]$	0.1236	0.0004	0.0313	0	* * *
evaluate	Obs.	LLH	AIC		BIC	
	4711	-66035.109	13244.218		13310.268	

Table 8. Parameter estimates for the DCC-GARCH-MIDAS model introduced by CPI: GARCH-MIDAS

Panel A: GARCH-MIDAS						
Stocks	Parameter	Estimate	Std. Error	T-value	$Pr(>\|t\|)$	Sig.
SSE	μ	0.0146	0.0163	0.8951	0.3707	
	α	0.0805	0.0145	5.538	0	* * *
	β	0.9087	0.0162	56.1385	0	* * *
	m	0.4357	0.3614	1.2056	0.228	
	θ	0.026	0.0121	2.1523	0.0314	**
	ω_2	1	1.7675	0.5658	0.5716	
SZE	μ	0.0197	0.0215	0.9154	0.36	
	α	0.0652	0.0123	5.2955	0	* * *
	β	0.9204	0.0155	59.4496	0	* * *
	m	0.7602	0.2464	3.0854	0.002	* * *
	θ	0.0165	0.0086	1.9275	0.0539	*
	ω_2	1.0015	0.8101	1.2363	0.2164	
HSI	μ	0.0356	0.0159	2.2373	0.0253	**
	α	0.0657	0.0086	7.6351	0	* * *
	β	0.9221	0.0106	87.2285	0	* * *
	m	0.8712	0.3114	2.7976	0.0051	* * *
	θ	-0.0175	0.0143	-1.2289	0.2191	
	ω_2	1.0002	2.1602	0.463	0.6433	

Table 9. Parameter estimates for the DCC-GARCH-MIDAS model introduced by CPI: GDCC-MIDAS

Panel B: DCC-MIDAS						
	Parameter	Estimate	Std. Error	T-value	$Pr(>\|t\|)$	Sig.
	a	0.0074	0	0.0025	0	* * *
	b	0.9561	0	0	0	* * *
	$\omega_2 \cdot RC$	2.0023	0.0196	0.0102	0	* * *
	$\omega_2 \cdot M2$	2.015	0.0311	0.6482	0	* * *
	$m[1,2]$	0.2661	0.0606	4.3908	0	* * *
	$m[1,3]$	0.4553	0.1022	4.4536	0	* * *
	$m[2,3]$	0.6016	0.1275	4.7194	0	* * *
	$\theta \cdot$ RC$[1,2]$	0.2904	0.0723	4.017	0.0001	* * *
	$\theta \cdot$ RC$[1,3]$	0.3269	0.2133	1.5327	0.1253	
	θ.RC$[2,3]$	0.3544	0.3432	1.0325	0.3019	
	θ.CPI$[1,2]$	-0.425	0	-0.001	0	* * *
	$\theta \cdot$ CPI$[1,3]$	0.0051	0.0001	0.3475	0	* * *
	$\theta \cdot$ CPI$[2,3]$	0.0928	0.0048	0.1921	0	* * *
evaluate	Obs.	LLH	AIC		BIC	
	4711	-31123.053	62220.106		62136.156	

Table 10. Parameter estimates for the DCC-GARCH-MIDAS model introduced by Mx: GARCH-MIDAS

Panel A: GARCH-MIDAS						
Stocks	Parameter	Estimate	Std. Error	T-value	$Pr(>\|t\|)$	Sig.
SSE	μ	0.0146	0.0163	0.8951	0.3707	
	α	0.0805	0.0145	5.538	0	* * *
	β	0.9087	0.0162	56.1385	0	* * *
	m	0.4357	0.3614	1.2056	0.228	
	θ	0.026	0.0121	2.1523	0.0314	**
	ω_2	1	1.7675	0.5658	0.5716	
SZE	μ	0.0197	0.0215	0.9154	0.36	
	α	0.0652	0.0123	5.2955	0	* * *
	β	0.9204	0.0155	59.4496	0	* * *
	m	0.7602	0.2464	3.0854	0.002	* * *
	θ	0.0165	0.0086	1.9275	0.0539	*
	ω_2	1.0015	0.8101	1.2363	0.2164	
HSI	μ	0.0356	0.0159	2.2373	0.0253	**
	α	0.0657	0.0086	7.6351	0	* * *
	β	0.9221	0.0106	87.2285	0	* * *
	m	0.8712	0.3114	2.7976	0.0051	* * *
	θ	-0.0175	0.0143	-1.2289	0.2191	
	ω_2	1.0002	2.1601	0.463	0.6433	
evaluate	Obs.	LLH	AIC		BIC	
	4711	-2007.192	4040.384		4124.334	

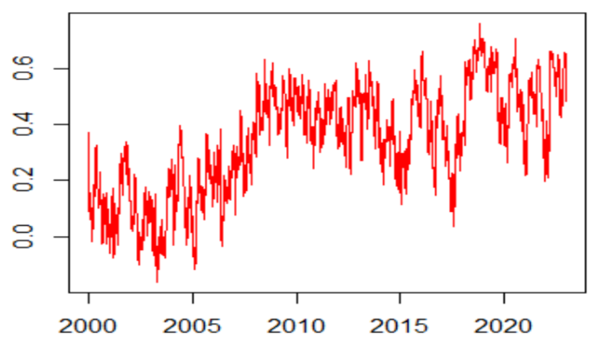

Fig. 5. Dynamic correlation coefficients of SSE and HSI

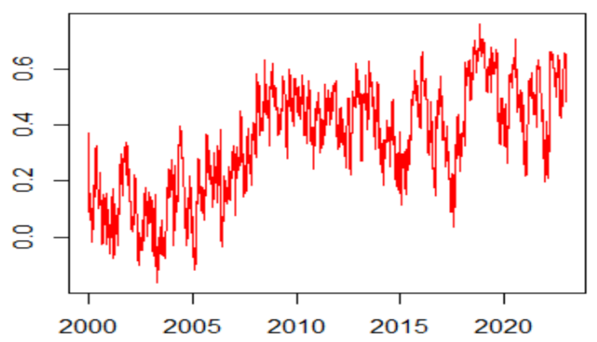

Fig. 6. Dynamic correlation coefficients of SZE and HSI

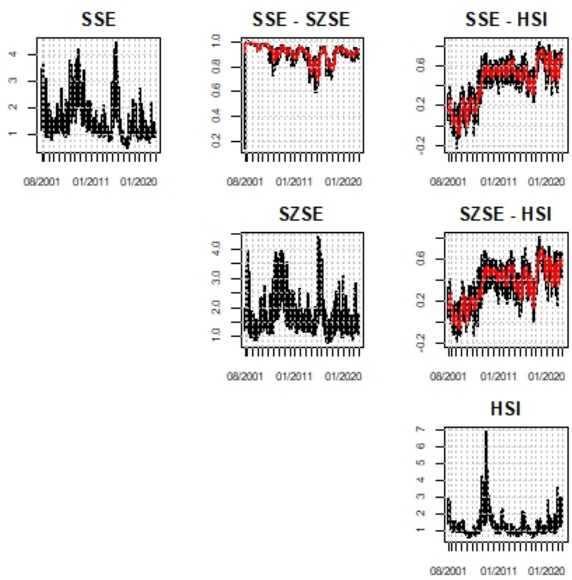

Fig. 7. Dynamic correlation coefficients of DCC-GARCH-MIDAS models

macro-variable volatility on the volatility of daily returns. This study takes a fixed window In this study, a fixed window is adopted for forecasting, and the volatility forecasting performance of a particular model is assessed by comparing the actual loss function with the previously calculated loss function to confirm whether the volatility generated by the introduction of the macro variables in the DCC-GARCH-MIDAS model improves the performance

Table 11. Parameter estimates for the DCC-GARCH-MIDAS model introduced by Mx: DCC-MIDAS

Panel B: DCC-MIDAS						
	Parameter	Estimate	Std. Error	T-value	$Pr(>\|t\|)$	Sig.
	a	0.0407	0.0045	9.0009	0	$***$
	b	0.9016	0.0168	53.6535	0	$***$
	$\omega_2 \cdot RC$	4.2228	2.4447	1.7274	0.0841	*
	$\omega_2 \cdot M2$	1.001	1.0623	0.9423	0.346	
	$m[1,2]$	-1.2433	0.5132	-2.4226	0.0154	**
	$m[1,3]$	0.2352	0.097	2.4236	0.0154	$**$
	$m[2,3]$	0.2486	0.0947	2.624	0.0087	$***$
	$\theta \cdot \mathrm{RC}[1,2]$	3.1169	0.5699	5.4694	0	$**$
	$\theta \cdot \mathrm{RC}[1,3]$	0.6401	0.1984	3.2262	0.0013	$**$
	$\theta \cdot \mathrm{RC}[2,3]$	0.4755	0.2294	2.0728	0.0382	**
	$\theta \cdot \mathrm{MX}[1,2]$	0.0933	0.0327	2.8497	0.0044	***
	$\theta \cdot \mathrm{MX}[1,3]$	-0.0326	0.0361	-0.9037	0.3662	
	$\theta \cdot \mathrm{MX}[2,3]$	-0.0269	0.0321	-0.8391	0.4014	
evaluate	Obs.	LLH	AIC		BIC	
	4711	-2007.192	4040.384		4124.334	

Table 12. Parameter estimation results of the DCC-GARCH-MIDAS model introduced by Scg: GARCH-MIDAS

Panel A: GARCH-MIDAS						
Stocks	Parameter	Estimate	Std. Error	T-value	$Pr(>\|t\|)$	Sig
SSE	μ	0.0146	0.0163	0.8951	0.3707	
	α	0.0805	0.0145	5.538	0	$***$
	β	0.9087	0.0162	56.1385	0	$***$
	m	0.4357	0.3614	1.2056	0.228	
	θ	0.026	0.0121	2.1523	0.0314	**
	ω_2	1	1.7675	0.5658	0.5716	
SZE	μ	0.0197	0.0215	0.9154	0.36	
	α	0.0652	0.0123	5.2955	0	$***$
	β	0.9204	0.0155	59.4496	0	$***$
	m	0.7602	0.2464	3.0854	0.002	$***$
	θ	0.0165	0.0086	1.9275	0.0539	*
	ω_2	1.0015	0.8101	1.2363	0.2164	
HSI	μ	0.0356	0.0159	2.2373	0.0253	**
	α	0.0657	0.0086	7.6351	0	$***$
	β	0.9221	0.0106	87.2285	0	$***$
	m	0.8712	0.3114	2.7976	0.0051	$***$
	θ	-0.0175	0.0143	-1.2289	0.2191	
	ω_2	1.0002	2.1601	0.463	0.6433	

Table 13. Parameter estimation results of the DCC-GARCH-MIDAS model introduced by Scg: DCC-MIDAS

Pane1 B: DCC-MIDAS						
	Parameter	Estimate	Std. Error	T-value	$Pr(>\|t\|)$	Sig.
	a	0.0286	0	0.0001	0	***
	b	0.9246	0	0	0	$***$
	$\omega_2 \cdot RC$	1.0801	0.0017	0.0624	0	$***$
	$\omega_2 \cdot M2$	1.0659	0.0262	0.407	0	$***$
	$m[1,2]$	-0.7181	0.0173	-0.4151	0	$***$
	$m[1,3]$	0.7293	0.0035	0.021	0	$***$
	$m[2,3]$	0.4374	0	0.0006	0	***
	$\theta \cdot \mathrm{RC}[1,2]$	2.8211	0.0072	0.0289	0	***
	$\theta \cdot \mathrm{RC}[1,3]$	0.834	0.0415	0.2009	0	***
	$\theta \cdot \mathrm{RC}[2,3]$	0.4095	0.0315	0.1299	0	$***$
	$\theta.\mathrm{SCG}[1,2]$	0.3424	0.0239	0.1433	0	$***$
	$\theta.\mathrm{SCG}[1,3]$	0.2927	0.0023	0.0187	0	$***$
	$\theta.\mathrm{SCG}[2,3]$	-0.1862	0.0051	-0.3681	0	***
evaluate	Obs.	LLH	AIC		BIC	
	4711	-22331.159	44636.318		44552.368	

Table 14. Parameter estimates of the DCC-GARCH-MIDAS model introduced by Fx: GARCH-MIDAS

Panel A: GARCH-MIDAS						
Stocks	Parameter	Estimate	Std. Error	T-value	$\Pr(>\|t\|)$	Sig.
SSE	μ	0.0146	0.0163	0.8951	0.3707	
	α	0.0805	0.0145	5.538	0	$***$
	β	0.9087	0.0162	56.1385	0	$***$
	m	0.4357	0.3614	1.2056	0.228	
	θ	0.026	0.0121	2.1523	0.0314	**
	ω_2	1	1.7675	0.5658	0.5716	
SZE	μ	0.0197	0.0215	0.9154	0.36	
	α	0.0652	0.0123	5.2955	0	$***$
	β	0.9204	0.0155	59.4496	0	$***$
	m	0.7602	0.2464	3.0854	0.002	$***$
	θ	0.0165	0.0086	1.9275	0.0539	*
	ω_2	1.0015	0.8101	1.2363	0.2164	
HSI	μ	0.0356	0.0159	2.2373	0.0253	$**$
	α	0.0657	0.0086	7.6351	0	$***$
	β	0.9221	0.0106	87.2285	0	$***$
	m	0.8712	0.3114	2.7976	0.0051	$***$
	θ	-0.0175	0.0143	-1.2289	0.2191	
	ω_2	1.0002	2.1601	0.463	0.6433	
evaluate	Obs.	LLH	AIC		BIC	
	4711	-21225.724	42525.448		42541.498	

Table 15. Parameter estimates of the DCC-GARCH-MIDAS model introduced by Fx: DCCMIDAS

Panel B: DCC-MIDAS						
	Parameter	Estimate	Std. Error	T-value	$\Pr(>\|t\|)$	Sig
	a	0.0075	0	0.0016	0	$***$
	b	0.9579	0	0	0	$***$
	$\omega_2 \cdot RC$	1.997	0.0003	0.0067	0	$***$
	$\omega_2 \cdot M2$	1.9987	0	0.0056	0	$***$
	$m[1,2]$	0.2929	0	0.0007	0	$***$
	$m[1,3]$	0.3517	0	0	0	$***$
	$m[2,3]$	0.6657	0	0	0	$***$
	$\theta \cdot \mathrm{RC}[1,2]$	0.275	0	0.0004	0	$***$
	$\theta \cdot \mathrm{RC}[1,3]$	0.2469	0	0	0	$***$
	$\theta \cdot \mathrm{RC}[2,3]$	0.3932	0	0	0	$***$
	$\theta.\mathrm{FX}[1,2]$	-0.6184	0	0	0	$***$
	$\theta \cdot \mathrm{FX}[1,3]$	0.205	0	0	0	$***$
	$\theta.\mathrm{FX}[2,3]$	-0.3641	0	0	0	***
evaluate	Obs.	LLH	AIC		BIC	
	4711	-21225.724	42525.448		42541.498	

Table 16. Summary of macroeconomic variables and indicators for all models in this paper

Variable	Variable abbreviations	Significance of variables
Money Supply	M2	the currency circulation in the country
Consumer Price Index	CPI	the health of the economic developmen
Total imports and exports growth	MX	the countrys level of international trade development
Total retails sales of social consumer goods	SCG	the level of demand in the country
Fixed assets investment completion	FX	the level of investment in the country

of the model.

1. MSE: Mean Squared Error (MSE) [36] is one of the most widely adopted statistical loss functions to evaluate the performance of forecasting models. MSE is constructed as follows

$$MSE = \frac{1}{n}\sum_{t=1}^{n}\left(\sigma_t^2 - \widehat{\sigma}_t^2\right)^2 \tag{12}$$

$\hat{\sigma}_t^2$ is predicted volatility; σ_t^2 is actual volatility. The value of MSE is always non-negative and better to be close to zero. Literally, the more it is close to zero, the higher the accuracy of prediction it provides. The disadvantage of MSE is that it heavily weights the outliers by taking the square of the differential between the predicted value and the proxy of actual value . This undesirable property is avoided by Mean Absolute Error (MAE).

2. MAE: Mean Absolute Error (MAE) [37]is another widely used loss function to measure the errors between predicted value and actual value. Compared to MSE, MAE is conceptually easier to interpret since MAE is obtained by taking the absolute value of errors rather than squared errors. In this respect, MAE is more reliable than MSE in the existence of outliers. The formulation of MAE is expressed as follows:

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|\sigma_t^2 - \widehat{\sigma}_t^2\right|\Bigg| \tag{13}$$

3. HMSE & HMAE : Heteroscedasticity Adjusted Mean Square Error (HMSE) and Heteroscedasticity Adjusted Mean Absolute Error (HMAE) are the heteroscedasticity adjusted MSE and MAE [38]. The specification of HMSE and HMAE is described as follows:

$$HMSE = \frac{1}{n}\sum_{t=1}^{n}\left(1 - \sigma_t^2/\widehat{\sigma}_t^2\right)^2 \tag{14}$$

and

$$HMAE = \frac{1}{n}\sum_{t=1}^{n}\left|1 - \sigma_t^2/\hat{\sigma}_t^2\right| \tag{15}$$

4. QLIKE Gaussian Quasi-likelihood (QLIKE) is an asymmetric loss function and highly recommended by Patton & Sheppard (2009) and Patton (2011). QLIKE penalizes the under-prediction more heavily than over-prediction, thus it is widely used in risk management since underprediction is more costly than over-prediction [39]. Patton (2011) shows that QLIKE is robust to the noise even in the existence of imperfect proxy of actual volatility. Hence, it is widely adopted in volatility forecasting field. The specification of QLIKE is expressed as follows:

$$\text{QLIKE} = \frac{1}{n}\sum_{t=1}^{n}\left(\ln\left(\hat{\sigma}_t^2\right) + \sigma_t^2/\hat{\sigma}_t^2\right)^2 \tag{16}$$

5. R^2LOG

$$R^2LOG = \frac{1}{n}\sum_{t=1}^{n}\left(\left(\ln\left(\sigma_t^2/\hat{\sigma}_t^2\right)\right)^2\right. \tag{17}$$

Evaluation of in-sample estimation across GARCH-MIDAS models is based on a few goodness of fit statistics, comprise of R^2 LOG, MAE, RMSE, Log L (Log-likelihood), SIC (Schwarzs Information Criteria), AIC (Akaikes Information Criterion) and statistically significant parameters estimated in each GARCH-MIDAS models. Smaller value of MSE, MAE, HMSE, HMAE, QLIKE, AIC and larger value of R^2 LOG and Log L signify a better fitting estimated volatility model. Nevertheless, a best fitted in-sample estimated volatility model need not necessary be the best performing out-sample volatility forecasting model (Choo et al., 1999)[40]. Therefore, post-sample forecasting performance evaluation is the key emphasis of this study since the ultimate objective of this study is to develop volatility forecasting models to measure risk in stock markets investment.

Tables 17-19 report the results of the three index return volatility forecasts for the out-of-sample period 2020-2022. First, we use historical data for the in-sample period 2000-2019 as in-sample data and use these data to train the volatility forecasting models. Second, the trained volatility forecasting model is used to fixed-window 1-day forecast the data from 2020 to 2022. To examine the difference between the true values and the predicted results, the error between the predicted volatility and the actual observed values is finally determined. This is used to compute the volatility loss and evaluate the models predicting accuracy. The findings demonstrate that, when macro variables are introduced, the prediction loss of DCC-GARCH-MIDAS decreases in all three of the index returns, as compared to the volatility of the DCC-GARCH model returns without the introduction of macro variables. The predictive ability of DCC-GARCH-MIDAS model with the introduction of macro variables outperform that of DCC-GARCH model. This indicates a significant

Table 17. SSE-SZE stock markets correlation 1-day loss functions forecast

	Evaluation indicators					
Model	MSE	MAE	RMSE	RMAE	R^2LOG	QLIKE
DCC-GARCH-MIDAS-RC(M2)	0.0063	0.0506	0.7968	0.0932	0.0084	0.8943
DCC-GARCH-MIDAS-RC(CPI)	0.0092	0.0595	0.9370	0.1085	0.0119	0.8960
DCC-GARCH-MIDAS-RC(MX)	0.0063	0.0492	0.8204	0.0924	0.0083	0.8943
DCC-GARCH-MIDAS-RC(SCG)	0.0102	0.0676	0.1102	0.8976	0.0147	0.8976
DCC-GARCH-MIDAS-RC(FX)	0.0080	0.0571	0.7881	0.1018	0.0107	0.8953
DCC-GARCH-MIDAS-RC	0.0061	0.0483	0.7919	0.0908	0.0081	0.8942
DCC-GARCH	0.0106	0.0707	0.8153	0.1182	0.0137	0.8968

improvement with in macro variables DCC-GARCH-MIDASs ability to predict. The introduction of macro variables FX has better predictive ability for the volatility of three stock markets. SSE and SZE stock markets are more correlated than SSE and HSI, and SZE and HSI. this is due to historical reasons, as the economic systems of Mainland China and Hong Kong are different, and the stock markets are very different in terms of system and maturity.

4. Conclusions

This study has conducted a comprehensive examination of the intricate interplay between macroeconomics and financial market volatility, specifically tackling the challenge posed by varying data frequencies. Through the utilization of the innovative DCC-GARCH-MIDAS model, we have effectively integrated high-frequency and low-frequency data to dynamically simulate the relationships between macroeconomic variables and financial market volatility. The model's adept management of data frequency disparities has proven to be a viable solution, facilitating a thorough analysis of market dynamics. Significantly, the forecasting capabilities demonstrated by the DCC-GARCH-MIDAS model during the outof-sample period underscore its efficacy in capturing the subtle impact of macroeconomic events on financial market volatility. The model's practical importance is further emphasized by its precise predictive abilities and adaptable nature, enabling it to navigate the complexities of economic fluctuations with accuracy. This research significantly contributes to advancing our comprehension of the nuanced relationship between financial market volatility and macroeconomics. The model's unique ability to accommodate diverse data frequencies enhances the granularity of our understanding, offering a more detailed and insightful perspective on how macroeconomic events intricately shape market behavior. As we progress, this research establishes a firm groundwork for future studies aimed at refining our understanding of the intricate interplay between macroeconomic forces and financial market dynamics.

References

[1] F. A. Anindyntha and M. K. Fuddin, (2023) *"How do macroeconomic variables and financial inclusion affect financial stability in Indonesia?"* **Jurnal Perspektif Pembiayaan dan Pembangunan Daerah *11***(5): 359–370.

[2] B. K. Ababio and W. Lu, (2023) *"Barriers and enablers of circular economy in construction: a multi-system perspective towards the development of a practical framework"* **Construction Management and Economics *41***(1): 3–21. DOI: 10.1080/01446193.2022.2135750.

[3] M. Yu, M. Umair, Y. Oskenbayev, and Z. Karabayeva, (2023) *"Exploring the nexus between monetary uncertainty and volatility in global crude oil: a contemporary approach of regime-switching"* **Resources Policy *85***: 103886.

[4] R. Wang, B. S. Mohammed, and R. Parthasarathy, (2023) *"Analysis on Risk Awareness Model and Economic Growth of Finance Industry"* **Journal of Internet Technology *24***(3): 697–707.

[5] R. F. Engle and J. G. Rangel, (2008) *"The spline-GARCH model for low-frequency volatility and its global macroeconomic causes"* **The review of financial studies *21***(3): 1187–1222. DOI: 10.1093/rfs/hhn004.

[6] F. X. Diebold and K. Yilmaz. *Macroeconomic volatility and stock market volatility, worldwide*. Tech. rep. National Bureau of Economic Research, 2008.

[7] S. Malik and M. K. Pitt, (2011) *"Modelling stochastic volatility with leverage and jumps: A simulated maximum likelihood approach via particle filtering"* **Available at SSRN 1763783**:

[8] D. Buono, G. L. Mazzi, G. Kapetanios, M. Marcellino, and F. Papailias, (2017) *"Big data types for macroeconomic nowcasting"* **Eurostat Review on National Accounts and Macroeconomic Indicators *1***(2017): 93–145.

Table 18. SSE-HSI stock markets correlation 1-day loss functions forecast

Model	Evaluation indicators					
	MSE	MAE	RMSE	RMAE	R^2LOG	QLIKE
DCC-GARCH-MIDAS-RC(M2)	0.4219	0.6197	14334.32	24.9822	5.4145	-0.5070
DCC-GARCH-MIDAS-RC(CPI)	0.3911	0.5974	14664.57	24.5260	5.2723	-0.539
DCC-GARCH-MIDAS-RC(MX)	0.2921	0.4748	10190.41	18.9111	4.5247	-0.0586
DCC-GARCH-MIDAS-RC(SCG)	0.4657	0.6319	11842.24	23.6409	5.3553	−0.8042
DCC-GARCH-MIDAS-RC(FX)	0.3102	0.5144	7618.762	18.7357	4.6249	-0.7469
DCC-GARCH-MIDAS-RC	0.4005	0.6081	14817.89	24.7832	5.3478	-0.5187
DCC-GARCH	0.5777	0.7403	27731.10	31.1049	6.1469	-0.3140

Table 19. SZE-HSI stock markets correlation 1-day loss functions forecast

	Evaluation indicators					
Model	MSE	MAE	RMSE	RMAE	R^2LOG	QLIKE
DCC-GARCH-MIDAS-RC(M2)	0.2814	0.4953	21016.58	26.5420	4.8995	-0.7461
DCC-GARCH-MIDAS-RC(CPI)	0.2459	0.4410	10179.01	21.7894	4.4639	-0.9879
DCC-GARCH-MIDAS-RC(MX)	0.3023	0.4792	11392.16	22.8784	4.6425	-1.4011
DCC-GARCH-MIDAS-RC(SCG)	0.1717	0.3446	7629.188	16.2366	3.4560	-1.4558
DCC-GARCH-MIDAS-RC(FX)	0.1994	0.3927	7387.44	18.4553	4.1114	-1.0828
DCC-GARCH-MIDAS-RC	0.2814	0.4994	19534.84	26.9867	4.9203	-0.7289
DCC-GARCH	0.3076	0.5170	22274.78	27.4334	5.0486	-0.7107

[9] D. C. Broadstock and L. T. Cheng, (2019) *"Time-varying relation between black and green bond price benchmarks: Macroeconomic determinants for the first decade"* **Finance research letters** ***29***: 17–22. DOI: 10.1016/j.frl.2019.02.006.

[10] Y.-r. Ma, Q. Ji, and J. Pan, (2019) *"Oil financialization and volatility forecast: Evidence from multidimensional predictors"* **Journal of Forecasting** ***38***(6): 564–581. DOI: 10.1002/for.2577.

[11] U. Akkoc and I. Civcir, (2019) *"Dynamic linkages between strategic commodities and stock market in Turkey: Evidence from SVAR-DCC-GARCH model"* **Resources Policy** ***62***: 231–239. DOI: 10.1016/j.resourpol.2019.03.017.

[12] J. Sun, K. Xiao, C. Liu, W. Zhou, and H. Xiong, (2019) *"Exploiting intra-day patterns for market shock prediction: A machine learning approach"* **Expert Systems with Applications** ***127***: 272–281. DOI: 10.1016/j.eswa.2019.03.006.

[13] J. Wang, Y. Huang, F. Ma, and J. Chevallier, (2020) *"Does high-frequency crude oil futures data contain useful information for predicting volatility in the US stock market? New evidence"* **Energy Economics** ***91***: 104897.

[14] X. Song, D. Kim, H. Yuan, X. Cui, Z. Lu, Y. Zhou, and Y. Wang, (2021) *"Volatility analysis with realized GARCH-Itô models"* **Journal of Econometrics** ***222***(1): 393–410.

[15] C. Cerovecki, C. Francq, S. Hörmann, and J.-M. Zakoıan, (2019) *"Functional GARCH models: The quasi-likelihood approach and its applications"* **Journal of econometrics** ***209***(2): 353–375.

[16] J. Chen and D. N. Politis, (2019) *"Optimal multi-step-ahead prediction of ARCH/GARCH models and NoVaS transformation"* **Econometrics** ***7***(3): 34.

[17] X. Wang, X. Wang, B. Li, and Z. Bai, (2020) *"The nonlinear characteristics of Chinese stock index futures yield volatility: Based on the high frequency data of CSI300 stock index futures"* **China Finance Review International** ***10***(2): 175–196. DOI: 10.1108/CFRI-07-2018-0069.

[18] Z. Gao, Y. He, and E. E. Kuruoglu. "A Hybrid Model Integrating LSTM and Garch for Bitcoin Price Prediction". In: *2021 IEEE 31st International Workshop on Machine Learning for Signal Processing (MLSP)*. IEEE. 2021, 1–6.

[19] K.-C. Yao, H.-W. Hsueh, M.-H. Huang, and T.-C. Wu, (2022) *"The role of GARCH effect on the prediction of air pollution"* **Sustainability** ***14***(8): 4459.

[20] B. Yasar, T. Martin, and T. Kiessling, (2020) *"An empirical test of signalling theory"* **Management Research Review** ***43***(11): 1309–1335. DOI: 10.1108/MRR-08-2019-0338.

[21] J. Dufitinema, (2021) *"Forecasting the Finnish house price returns and volatility: a comparison of time series models"* **International Journal of Housing Markets**

and Analysis ***15***(1): 165–187. DOI: 10.1108/IJHMA-12-2020-0145.

[22] A. L. Hansen, (2021) *"Yield curve volatility and macroeconomic risk"* **Available at SSRN 3943978**:

[23] A. A. Laghari, S. Yin, et al., (2022) *"How to collect and interpret medical pictures captured in highly challenging environments that range from nanoscale to hyperspectral imaging"* **Current Medical Imaging** ***54***(36582065): 1.

[24] A. A. Laghari, Y. Sun, M. Alhussein, K. Aurangzeb, M. S. Anwar, and M. Rashid, (2023) *"Deep residual-dense network based on bidirectional recurrent neural network for atrial fibrillation detection"* **Scientific Reports** ***13***(1): 15109.

[25] S. Karim, A. Qadir, U. Farooq, M. Shakir, and A. A. Laghari, (2023) *"Hyperspectral imaging: a review and trends towards medical imaging"* **Current medical imaging** ***19***(5): 417–427.

[26] S. Das, A. Adhikary, A. A. Laghari, and S. Mitra, (2023) *"Eldo-care: Eeg with kinect sensor based telehealthcare for the disabled and the elderly"* **Neuroscience Informatics**: 100130.

[27] A. A. Laghari, S. Shahid, R. Yadav, S. Karim, A. Khan, H. Li, and Y. Shoulin, (2023) *"The state of art and review on video streaming"* **Journal of High Speed Networks** (Preprint): 1–26.

[28] M. Andreani, V. Candila, G. Morelli, and L. Petrella, (2021) *"Multivariate Analysis of Energy Commodities during the COVID-19 Pandemic: Evidence from a Mixed-Frequency Approach"* **Risks** ***9***(8): 144.

[29] J. Wu. *"Corrected GARCH-DCC-MIDAS models in economics and finance"*. (phdthesis). Brunel University London, 2023.

[30] Z. Wang, T. Xing, and X. Wang, (2024) *"Economic uncertainty and stock market asymmetric volatility: analysis based on the asymmetric GARCH-MIDAS model"* **International Journal of Emerging Markets**:

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

[32] R. F. Engle, E. Ghysels, and B. Sohn, (2013) *"Stock market volatility and macroeconomic fundamentals"* **Review of Economics and Statistics** ***95***(3): 776–797. DOI: 10.1162/REST_a_00300.

[33] R. Colacito, R. F. Engle, and E. Ghysels, (2011) *"A component model for dynamic correlations"* **Journal of Econometrics** ***164***(1): 45–59.

[34] A. W. Ayele, E. Gabreyohannes, and H. Edmealem, (2020) *"Generalized autoregressive conditional heteroskedastic model to examine silver price volatility and its macroeconomic determinant in Ethiopia market"* **Journal of Probability and Statistics** ***2020***: 1–10.

[35] L. Wen, C. Liu, H. Song, and H. Liu, (2021) *"Forecasting tourism demand with an improved mixed data sampling model"* **Journal of Travel Research** ***60***(2): 336–353.

[36] M. S. Error, (2010) *"Mean squared error"* **MA: Springer US**: 653–653.

[37] T. Chai and R. R. Draxler, (2014) *"Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature"* **Geoscientific model development** ***7***(3): 1247–1250. DOI: 10.5194/gmd-7-1247-2014.

[38] C.-H. Chen, W.-C. Yu, and E. Zivot, (2012) *"Predicting stock volatility using after-hours information: Evidence from the NASDAQ actively traded stocks"* **International Journal of Forecasting** ***28***(2): 366–383. DOI: 10.1016/j.ijforecast.2011.04.005.

[39] R. W. Wedderburn, (1974) *"Quasi-likelihood functions, generalized linear models, and the Gauss—Newton method"* **Biometrika** ***61***(3): 439–447. DOI: 10.1093/biomet/61.3.439.

[40] C. W. Chong, M. I. Ahmad, and M. Y. Abdullah, (1999) *"Performance of GARCH models in forecasting stock market volatility"* **Journal of forecasting** ***18***(5): 333–343.