Intraday volatility analysis of CSI 300 index futures: a dependent functional data method

Danni Wang, Zhifang Su & Qifang Li

Published online: 06 Jun 2022.

Submit your article to this journal

Article views: 684

View related articles

View Crossmark data

Citing articles: 1 View citing articles

Routledge
Taylor & Francis Group

Intraday volatility analysis of CSI 300 index futures: a dependent functional data method

Danni Wang[a], Zhifang Su[a] and Qifang Li[b]

[a]School of Economics and Finance, Huaqiao University, Quanzhou, China; [b]School of Mathematics and Statistics, Minnan Normal University, Zhangzhou, China

ABSTRACT

This study introduces a new volatility model based on dependent functional data to investigate the intraday volatility characteristics of CSI 300 in the context of high-frequency data. The volatility curve is fitted and reconstructed using three methods: functional principal component analysis, Newey-West kernel, and truncation-free Bartlett kernel. We adopt a functional time series approach for short-term dynamic forecasting. The empirical results show that the proposed dependent functional volatility estimation model based on the long-term covariance of the truncated Bartlett kernel can accurately capture the intraday volatility trajectory and outperforms other models in terms of forecast accuracy and profitability. This study improves the volatility-related research methodology, which is conducive to discovering the price formation mechanism of the stock index futures market and improving risk management capabilities.

1. Introduction

Volatility, as an important indicator of financial market risk, can effectively reflect the process of asset price volatility and is useful for discovering the market's micro price formation mechanism (Poon & Granger, 2003). However, in the actual trading process, volatility exhibits certain deviations over time that cannot be measured directly, with the distribution of asset returns having a leverage effect on volatility (Bollerslev & Zhou, 2006). Therefore, it is of great theoretical significance to construct a model that can both characterise volatility and accurately predict future volatility trajectories to improve the trading mechanisms of financial markets, enhance the efficiency of market operations, and achieve the optimal allocation of capital.

To adequately describe the volatility process of asset returns, Engle (1982) and Bollerslev (1986) proposed the autoregressive conditionally heteroscedastic (ARCH) and generalised autoregressive conditionally heteroscedastic (GARCH) models, respectively, which can characterise volatility aggregation but can hardly explain the

CONTACT Zhifang Su ✉ suzufine@hqu.edu.cn

leverage effect of returns. Since then, Shimada et al. (2009), Wei (2012), Wu et al. (2018), Wang et al. (2021), Kim et al. (2021), and many others have further proposed GARCH-type and stochastic volatility (SV)-type models, which reflect the time-varying characteristics of volatility by giving it a certain structure and apply them to the dynamics of asset prices in financial markets. However, the above parametric models can only indirectly solve the problem of unpredictable volatility based on low-frequency return data, such as daily and monthly returns. In contrast, the realised volatility (RV) and heterogeneous autoregressive-realised volatility (HAR-RV) models proposed by Andersen et al. (2003) and Corsi (2008) are measures of a non-parametric nature, which consider the long memory characteristics of RV and are gradually becoming benchmark models for volatility modelling. In recent years, many studies have demonstrated the high fitting accuracy of RV models based on intraday high-frequency data, such as Kambouroudis et al. (2021), Wang et al. (2019), Li et al. (2021). However, such models can only capture daily frequency fluctuations and cannot accurately characterize volatility at the intraday level. Simultaneously, with the development of financial markets and the increase in big data processing power, investors have access to a larger volume of financial data with more complex data structures, making it difficult for traditional low-frequency models to accurately describe the underlying intraday stochastic processes of the actual observed data and meet the needs of financial market development (Madden, 2012).

The current forecasting methods for financial statistics can be divided into three main categories. The first is forecasting analysis based on microstructures such as investor sentiment and market information transmission efficiency (Nie & Li, 2019). The market measures of such analysis are somewhat subjective, and the applicability of the analysis results is limited. The second is forecasting using machine learning methods (Bao et al., 2017; Lin & Gong, 2017). Such methods generally assume that the samples are independent of each other; however, in practice, the samples of financial asset returns are somewhat correlated in the time series and do not satisfy the independence assumption. The third category includes methods that directly feed high-frequency data to traditional time series models for forecasting analysis. Nevertheless, such methods can lead to ineffective exclusion of market noise (Wang et al., 2018).

In contrast, the functional data analysis (FDA) method, proposed by Ramsay (1982), can describe the underlying stochastic process of asset movements from the perspective of a curve by treating frequency-mixed, unequally spaced discrete intraday high-frequency observations as a continuous smooth sample curve. Since then, it has been further refined by scholars such as Ramsay and Dalzell (1991), Ramsay and Silverman (2002), Huang et al. (2001) and He et al. (2003, 2004) to form a systematic theoretical framework. On this basis, Dauxois et al. (1982) and Besse (1992) used functional principle component analysis (FPCA) to transform infinite dimensional feature vectors into finite dimensional score vectors. In recent years, FDA has also proven to be more applicable in analysis and is widely used in meteorology, economics, biomedicine, and other fields (Cerovecki et al., 2019; Das et al., 2019; Tsay, 2016). In particular, Müller et al. (2011) proposed a functional volatility model by applying FDA to study asset return volatility and the dynamics of volatility over very short

time intervals. Shang et al. (2019) combined traditional time series forecasting methods with FDA to propose a functional time series that is more efficient than traditional time series models with higher forecasting accuracy.

However, the existing FDA literature generally assumes that functional data are subject to an independent identical distribution condition, whereas in practice, most financial statistics, such as stock index futures and stock indices, are inter-memorably correlated with each other. Thus, a direct conventional econometric analysis of such an interdependent series would lead to biased statistical conclusions. Therefore, Hörmann and Kokoszka (2010), Horvath and Kokoszka (2012), and Kokoszka and Young (2017) proposed a dependent functional data analysis that uses a long-term covariance function to replace the original independent homogeneous covariance function to obtain more accurate eigenvalues and eigenfunctions, thus correcting the estimation bias, reducing information loss, and making the analysis results more robust and credible. Although this method can neatly characterise price changes with significant stochasticity, the estimation of the long-term covariance function is subject to the problem of choosing the kernel function and window width, which can be too large or too small, leading to errors in the estimation.

Therefore, To avoid the selection error of kernel function and window width, this study innovatively proposes a truncated Bartlett kernel long-term covariance estimation statistic, and on this basis, introduces the framework of dependent functional data analysis into the estimation process of intraday volatility, constructs a new dependent functional volatility model, and empirically identifies the CSI 300 high-frequency intraday stock price volatility pattern. Specifically, the discrete intraday high-frequency observations are first converted into a dependent functional dataset, dimensionality reduction and function curve reconstruction are achieved through principal component decomposition, and the volatility process is further investigated based on principal component regression for short-term dynamic forecasting and returns. This study contributes to the existing literature by improving theortical volatility models. Furthermore, it has far-reaching practical impliactions, as it effectively predicts the stock market trends and broadens the application of functional data analysis in the fields of economics and finance.

The remainder of this paper is organised as follows. Section 2 introduces the model construction. Section 3 describes the forecasting process. Section 4 presents and discusses the empirical results. Finally, Section 5 summarises the main conclusions.

2. Methods

2.1. Functional volatility processes

With the development of functional data, Müller et al. (2011) transformed the characteristics of the intraday volatility trajectory mode of the return rate into the framework of functional data analysis and proposed a functional volatility model. This model assumes that the observed trajectory of volatility is realised by multiple repetitions of unknown random processes: $d \log X_i(t, \omega) = \mu_i(t, \omega)dt + \sigma_i(t, \omega)dW_i(t, \omega)$, where $\mu(t), \sigma(t)$ is a smooth non-stationary random process under the condition of

independent homogeneous distribution, and $W_i(t, \varpi)$ represents an independent standard Wiener process. As the random data of real volatility cannot be measured directly, and the impact of market microstructure noise leads to data jumping, it is necessary to discretise the data in the diffusion model. The scaled logarithmic return rate and the related diffusion terms are defined as follows:

$$Z_\Delta(t) = \sqrt{\Delta}^{-1} \log\left(\frac{X(t+\Delta)}{X(t)}\right)$$

$$W_\Delta(t) = \sqrt{\Delta}^{-1}[W(t+\Delta)-W(t)] \tag{1}$$

For high-frequency discrete data observations, the model can be further expressed as

$$Z_\Delta(t) = \sqrt{\Delta}^{-1}\left[\int_t^{t+\Delta} \mu(v)dv + \int_t^{t+\Delta} \sigma(v)dW(v)\right] \tag{2}$$

To represent the remaining term after discretisation, the above formula can be simplified as $Z_\Delta(t) = \mu(t)\sqrt{\Delta} + \sigma(t)W_\Delta(t) + R_1 + R_2$, where R_1 and R_2 denote the residual terms after discretisation as

$$R_1(t) = \frac{1}{\sqrt{\Delta}}\left[\int_t^{t+\Delta} \mu(v)dv - \mu(t)\sqrt{\Delta}\right]$$

$$R_2(t) = \frac{1}{\sqrt{\Delta}}\left[\int_t^{t+\Delta} \sigma(v)dW(v) - \sigma(t)W_\Delta(t)\right] \tag{3}$$

When the time interval of sample values is small enough and the observed samples are large enough, the volatility approaches $\Delta \to 0$, and the asymptotic hypothesis is approximately satisfied. Thereby, the approximate model can be obtained as $\log Z_\Delta^2(t) \approx \log \sigma^2(t) + \log W_\Delta^2(t)$.

Therefore, in actual observations, the smoothing process of functional volatility can be defined as $V(t)$, where the relationship between the functional volatility process and the actual observed value is

$$V(t) = \log \sigma^2(t)$$

$$\log Z_\Delta^2(t_j) - q_0 \approx Y_\Delta(t_j) = V(t_j) + U_\Delta(t_j) \tag{4}$$

Based on the independent increment property of the Wiener process $W_\Delta(t) \sim N(0, 1)$, further values can be obtained by q_0 and $q_0 = E[\log W_\Delta^2(t_j)] = \frac{4}{\sqrt{2\pi}}\int_0^\infty e^{-t^2/2}\log t \, dt \approx -1.27$.

Specific discrete time points $t_j = j\Delta, j = 1, 2, \ldots, J$ with equal intervals are expressed as:

$$\log Z_\Delta^2(t_j) - q_0 \approx Y_\Delta(t_j) = V(t_j) + U_\Delta(t_j)$$

$$U_\Delta(t) = \log W_\Delta^2(t) - q_0$$

$$Y_\Delta(t) = \log[\sigma(t)W_\Delta(t)]^2 - q_0 \tag{5}$$

The above formula satisfies both $EU_\Delta(t) = 0$ and $cov(U_\Delta(t), U_\Delta(s)) = 0, |t - s| > 0$. Therefore, $Y_\Delta(t)$ is a random process. It can be decomposed into a functional volatility process, $V(t_j)$, and a zero-mean independent increment process, $U_\Delta(t_j)$. Compared to the traditional volatility model, its advantage is that the stochastic process of volatility is smooth and does not depend on the trading time interval, Δ.

A functional data analysis method is introduced to extract the functional volatility process $V(t_j)$ in the above equation, where the mean function of the functional volatility process is denoted as $\mu_V(t) = EV(t)$, and the covariance function is denoted as $g_v(t, s) = cov(V(t), V(s))$. We define the linear integral covariance operator as:

$$G_v f(t) = \int_0^T g_v(t, s) f(s) ds \tag{6}$$

where $\phi_k(t)$ is the orthonormal characteristic function of the covariance operator, G_v. The corresponding eigenvalues λ_k are satisfied by $\lambda_1 \geq \lambda_1 \geq \ldots, \sum \lambda_k < \infty$. According to the K-L decomposition, a single random functional volatility trajectory $V(t)$ can be obtained, which can be expressed as:

$$V(t) = \mu_v(t) + \sum_{k=1}^{\infty} \xi_k \phi_k(t) \tag{7}$$

where $\phi_k(t)$ is the functional principal component, which reflects the fluctuation mode of the individual fluctuation trajectory; the mean function $\mu_v(t)$ reflects the overall fluctuation trend; ξ_k is the principal component score, which represents an irrelevant random variable, and satisfies $\xi_k = \int (V(t) - \mu_x(t)) \phi_k(t) dt, E\xi_k = 0$ $Var(\xi_k) = \lambda_k$.

2.2. A dependent functional data analysis method based on untruncated Bartlett kernel

The functional data analysis proposed by Ramsay (1982) has significant advantages in dealing with high-dimensional data; however, its model is mainly based on the assumption of an independent homogeneous distribution. However, high-frequency data in finance are not only naturally functional, but also have strong interdependencies. For dependent functional data that do not satisfy the independent identical distribution condition, this study uses a long-term covariance function instead of short-term covariance under the I.I.D condition to reconstruct the fit to the discrete data. Further, to avoid the artificiality of choosing kernel functions and optimal window widths, an innovative long-term covariance estimation statistic based on a truncation-free Bartlett kernel is proposed. The specific procedure is as follows.

If a random variable $X_{ij} = X_i(t_j)$ originates from a square integrable Hilbert space $H = L^2[0, T]$ and is a continuous random process, it can be expressed as a stationary functional time series:

$$X_{ij} = X_i(t_j) + \varepsilon_i(t_j), \quad 1 \le i \le N, \quad 1 \le j \le T. \tag{8}$$

When $X_n(t)$ is dependent on functional data, long-term covariance is used to replace covariance under the I.I.D. condition to correct the functional data. The long-term covariance function is defined as:

$$\Gamma(s,t) = \Gamma_0(s,t) + \sum_{h=1}^{\infty} [\Gamma_h(s,t) + \Gamma_{-h}(s,t)], \tag{9}$$

where $\Gamma_h(s,t)$ and $\Gamma_{-h}(s,t)$ are self-covariance functions of order H, which are defined as:

$$\begin{aligned}
\Gamma_h(s,t) &= \mathrm{cov}[X_i(s), X_{i+h}(t)] = E\{[X_i(s)-\mu(s)][X_{i+h}(t)-\mu(t)]\} \\
\Gamma_{-h}(s,t) &= \mathrm{cov}[X_i(s), X_{i-h}(t)] = E\{[X_i(s)-\mu(s)][X_{i-h}(t)-\mu(t)]\}.
\end{aligned} \tag{10}$$

Moreover, $h = 0, \Gamma_0(s,t)$ is the short-term covariance function under the I.I.D condition.

When discrete observation data are collected, the kernel function method can be used to estimate the long-term covariance function as:

$$\begin{aligned}
\hat{\Gamma}(s,t) &= \hat{\Gamma}_0(s,t) + \sum_{h=1}^{N-1} K\left(\frac{h}{q}\right)\left[\widehat{\Gamma}_h(s,t) - \widehat{\Gamma}_{-h}(s,t)\right] \\
\widehat{\Gamma}_h(s,t) &= (N-h)^{-1}\sum_{i=1}^{N-h}[X_i(s) - \bar{X}_N(s)][X_j(t) - \bar{X}_N(t)]
\end{aligned} \tag{11}$$

The optional kernel function is defines as $K(x) = 1-|x|, |x| \le 1$. Some scholars also choose the Newey-West kernel function to estimate the long-term covariance function (e.g. Kokoszka & Young, 2017), which is defined as:

$$\omega_m(h) = \begin{cases} 1-\dfrac{h}{m}, & h \le m \\ 0, & h > m \end{cases}, \quad m = 4 * \left(\frac{N}{100}\right)^{\frac{2}{9}}$$

However, all of the aforementioned methods for estimating long-term covariance functions face the problem of choosing the kernel function and the optimal window width, which could generate large errors if not properly chosen. Therefore, to avoid the above problems, this study proposes a long-term covariance estimation statistic without a truncated Bartlett kernel.

The improved Bartlett kernel long-term covariance can be expressed as:

$$\Gamma = \Gamma_0 + \sum_{h=1}^{\infty} [\Gamma_k + \Gamma_{-k}]. \tag{12}$$

Among them, $\Gamma_{-k} = E(v_i v'_{i-h}), v_i = (v_{i1}v_{i2}\cdots v_{ik})', E(v_i) = 0$.

The long-term covariance estimation statistics are as follows:

$$\hat{\Gamma} = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \left(1 - \frac{|i-j|}{N}\right) v_i v'_j. \tag{13}$$

Only the sample long-term covariance between variables s and t at the time point is discussed, and the above equation is transformed into

$$\hat{\Gamma}(s,t) = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \left(1 - \frac{|i-j|}{N}\right) \left[X_i(s) - \bar{X}_N(s)\right] \left[X_j(t) - \bar{X}_N(t)\right]. \tag{14}$$

After estimating the long-term covariance of the sample, the principal component $\phi_k^\wedge$ and eigenvalue of the corresponding function $\hat{\lambda}_k$ can be calculated, and the principal component score of the function can be obtained as:

$$\xi_{ik}^\wedge = \int \left[X_i(t) - \hat{\mu}(t)\right] \phi_k^\wedge(t) dt. \tag{15}$$

The function expansion can be expressed as:

$$X_i(t) = \mu(t) + \sum_{k=1}^{\infty} \xi_{ik} \phi_k(t) \approx \hat{\mu}(t) + \sum_{k=1}^{K} \xi_{ik}^\wedge \phi_k^\wedge(t). \tag{16}$$

Theoretically, the Bartlett kernel long-term covariance estimation statistics proposed in this study do not need to manually select the kernel function and window width, which is simpler and more reasonable than the covariance estimation methods in the literature and more accurate than traditional functional data estimation methods.

2.3. Dependent function volatility model

Furthermore, the Bartlett kernel long-term covariance estimation statistics are substituted into the functional volatility estimation process to obtain a new dependent functional volatility process.

Assume N independent repeats of the price process $X(t)$, with equally spaced observed values. According to the relationship between RV and the scaled logarithmic return rate defined above, the following relationship can be obtained:

$$RV_i = \Delta \sum_{j=1}^{J} Z_{ij\Delta}^2$$

$$Z_{ij\Delta} = Z_{i\Delta}(t_j) = \frac{1}{\sqrt{\Delta}} r_{i\Delta}(t_j) = \frac{1}{\sqrt{\Delta}} \log\left(\frac{X_i(t_j + \Delta)}{X_i(t_j)}\right), \tag{17}$$

reducing $Y_{ij\Delta} = Y_{i\Delta}(t_j) = \log Z_{i\Delta}^2(t_j) - q_0 = \log r_{i\Delta}^2(t_j) - \log \Delta - q_0$. At this point, we can estimate the new principal component score as $\hat{\xi}_{ik} = \Delta \sum_{j=1}^{J}(Y_{i\Delta}(t_j - \hat{\mu}_V(t_j))\hat{\phi}_k(t_j)$.

According to K-L expansion, we can further obtain the volatility trajectory of dependence functional data as:

$$\hat{V}_i(t) = \hat{\mu}_V(t) + \sum_{k=1}^{K} \hat{\xi}_{ik}\hat{\phi}_k(t). \tag{18}$$

3. Forecasting methods

3.1. Forecasting process

According to Equation (18), it can be concluded that the random process of dependent functional volatility can be decomposed into mean function $\mu(t)$ and the cumulative product of principal component function and principal component score, namely

$$X_i(t) = \mu(t) + \sum_{k=1}^{\infty} \xi_{ik}\phi_k(t), \tag{19}$$

where ξ_{ik} represents the k principal component score of the i curve, which is the projection of $[V_i(t) - \mu(t)]$ onto the k principal component function, $\phi_k(t)$.

In practice, K principal component scores $(\xi_1, \xi_1, \ldots, \xi_k)$ are typically used to express the original data information, which can effectively achieve the effect of dimensionality reduction. That is,

$$X_i(t) = \mu(t) + \sum_{k=1}^{K} \xi_{ik}\phi_k(t) + e(t). \tag{20}$$

The functional time series dynamic predictive analysis method is based on Functional principal component analysis. Assuming that data are available for the first T_0 moments of the $n+1$ function curve, that is, $X_{n+1|n}(t_\varepsilon) = [X_{n+1}(t_1), \ldots, X_{n+1}(t_{T_0})]^T$, for forecasting data after T_0 moments, only the functional principal component score, which reconstructs the function series from the sample estimates of the mean and covariance statistics, is required to obtain the forward h-step forecast and is defined as

$$\hat{x}_{n+h}(t) = E[x_{n+h}(t)|\hat{\mu}(t), \hat{\Phi}(t), x(t)] = \hat{\mu}(t) + \sum_{k=1}^{K} \hat{\xi}_{n+h,k}\hat{\phi}_k(t). \tag{21}$$

Based on the above analysis, this study develops the following volatility prediction steps. First, data pre-processing is performed to estimate the mean function from a functional sample. Second, the principal component scores $\hat{\xi}_{1k}$ are obtained using covariance estimation under an independent identical distribution without considering sample dependence. Further, considering the sample dependence, the long-run

covariance estimation with the N-W kernel and Bartlett kernel without truncation are used to obtain the dependent function-based principal component scores $\hat{\xi}_{2k}$ and $\hat{\xi}_{3k}$. Finally, the corresponding eigenfunctions$\hat{\phi}_k(t)$ and eigenvalues $\hat{\lambda}_k$ are calculated by substituting the different functional principal component scores into the expansion equation, from which the predicted values of the function series $(\hat{x}_{n+1}(t), \hat{x}_{n+2}(t), \ldots, \hat{x}_{n+h}(t))$ can be obtained. k is determined by the proportion of varianceδ, $(\sum_{k=1}^{K} \hat{\lambda}_k / \sum_{k=1}^{T} \hat{\lambda}_k) \geq \delta$, which is chosen at 95% in this study.

3.2. Principal component score prediction

As mentioned above, the prediction of volatility is ultimately interpreted as a prediction of the principal component scores, and therefore, can be predicted using different time series methods. For example, Hyndman and Shang (2009) used a one-dimensional time series ARMA model.

However, in practice, there is no guarantee that the lagged autocovariance matrix of the principal component score vectors is fixed as a diagonal matrix; therefore, the one-dimensional time series model may lose the matrix, resulting in poor prediction accuracy. To address this problem, Aue et al. (2015) proposed a multivariate time series vector autoregressive (VAR) model. It is assumed that the principal component score series satisfy the VAR model, that is,

$$\xi_i = B'x_i + e_i, i = m + 1, \ldots, n, \tag{22}$$

where the independent variable takes the value of $X = (x_{m+1}, \ldots, x_n)'$, the dependent variable takes the value of $Y = (\xi_m, \ldots, \xi_n)'$, and the least-squares method can be used to obtain the parameter estimates$\hat{B}_{OLS} = (X'X)^{-1}X'Y$, from which the principal component scores can be obtained as the forward h-step prediction values $\hat{\xi}_{n+h|n} = \hat{B}_{OLS}X_{n+h}$.

The above time series VAR forecasts take the whole curve as a unit; however, in practice, the latest observations may not form a curve in time, and dynamic updates are needed to improve the forecast accuracy. Shen (2009) and Shang and Hyndman (2011) and Shang (2017) proposed several discrete-time dynamic forecasting methods, such as OLS, RR, and PLS to address this problem. Among them, Shang (2018) compared two dynamic forecasting methods (i.e. block moving (BM) and function-based linear regression (FLR)) with traditional methods, and found that BM and FLR predictions were better. The FLR model embodies the idea of treating the function as a whole, which makes its prediction better than discrete prediction. The FLR forecasting model was constructed as follows:

$$\hat{x}_{n+1}^l(t) = \hat{\mu}^l(t) + \int_{I_e} \left[x_{n+1}^e(s) - \mu^e(s) \right] a(s,t) ds + e_{n+1}^l(t), \tag{23}$$

where $I_e = [t_1, t_{m0}]$ and$I_l = (t_{m0}, t_p]$ correspond to the observation interval and the prediction interval, respectively.The mean function in the corresponding interval is denoted as $\mu^e(s), \mu^l(t)$; and the regression coefficient function and the regression error function are denoted by $a(s,t), e_{n+1}^l(t)$, respectively.

The regression coefficient function$a(s, t)$ can be estimated by projecting the independent and dependent variables $x_{n+1}^e(s), x_{n+1}^l(t)$ onto the functional principal components as:

$$x_i^e(s) = \mu^e(s) + \sum_{k=1}^{\infty} \xi_{i,k} \phi_k^e(s) = \mu^e(s) + \sum_{k=1}^{K} \xi_{i,k} \phi_k^e(s) + u_i^e(s), s \in I_e$$

$$x_i^l(t) = \mu^l(s) + \sum_{m=1}^{\infty} \eta_{i,m} \psi_m^l(t) = \mu^l(t) + \sum_{m=1}^{M} \eta_{i,m} \psi_m^l(t) + v_i^l(t), \ s \in I_l. \tag{24}$$

In the above equation, the functional principal component is expressed as$\phi_k^e(s), \psi_m^l(t)$; the principal component score is expressed as $\xi_{i,k}, \eta_{i,m}$; the error function caused by the truncation is expressed as$u_i^e(s), v_i^l(t)$; the principal component score matrix of the independent and dependent variables is expressed as$\xi =, \xi_1, \ldots, \xi_K, , \eta = (\eta_1, \ldots, \eta_M)$; and the relationship between the two is $\eta = \xi\gamma$. γ can be estimated using OLS as $\hat{\gamma} =, \xi'\xi^{-1}, \xi'\eta$. The score matrices of the independent and dependent variables are modelled using the reciprocal covariance structure as $\int_{t \in I_l} \int_{s \in I_e} \phi_k(s)\psi_m(t)\text{cov}[x^e(s), x^l(t)]dsdt$ and

$$\hat{x}_{n+1}^l(t) = \hat{\mu}^l(t) + \sum_{m=1}^{\infty} \eta_{n+1,m} \psi_m^l(t) \approx \hat{\mu}^l(t) + \hat{\xi}_{n+1}\hat{\gamma}\hat{\psi}^l(t), \tag{25}$$

Similarly, using the estimated statistics based on the long-run covariance N-W kernel and the truncation-free Bartlett kernel, the predicted values of the dependent functional time series can be obtained as:

$$\hat{x}'_{n+h}(t) = E\left[x'_{n+h}(t)|\hat{\mu}'(t), \hat{\Phi}'(t), x'(t)\right] = \hat{\mu}'(t) + \sum_{k=1}^{K} \hat{\xi}'_{n+h,k}\hat{\phi}'_k(t). \tag{26}$$

4. Results

4.1. Intraday volatility trajectory estimation

This study uses the 5-minute closing price of the current-month contract (IF2012) of China's CSI 300 stock futures index from 2019-11-4 to 2020-10-29 as sample data. The data are obtained from the RESSET Financial High Frequency Database, with trading hours of 9:35-11:30 a.m. and 13:00-15:00 p.m. The entire contract period comprises 240 trading days. To avoid the 'overnight' and 'midday' effects on the intraday volatility of stock index futures, the first observation is discarded in the actual calculation, and 9:40 is taken as the first trading hour. The missing values in the original data of intraday volatility are filled in using k-nearest neighbors (KNN) interpolation.

First, intraday 5-minute raw closing price data are transformed into scaled log returns$Z_{ij\Delta} = Z_{i\Delta}(t_j)$, and then transformed into raw intraday volatility data by simplification of $Y_{ij\Delta} = Y_{i\Delta}(t_j) = \log Z^2_{ij\Delta} - q_0$. As shown in Figure 1[1], the solid line
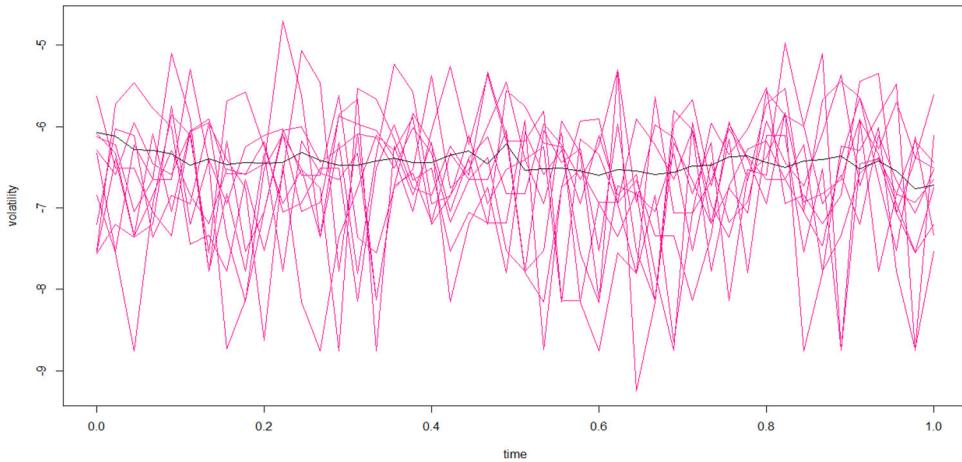
Figure 1. CSI 300 stock index futures 5 minutes high-frequency raw volatility.
Source: statistical results.

Table 1. Descriptive statistics for yield series.

Point	Mean	Mid	Min	Max	Upper	Lower
11280	−6.439	−6.325	−9.889	−3.269	−5.792	−6.977

Source: statistical results.

represents the raw volatility mean curve of the CSI 300 stock futures index. As shown in the graph, the original time series fluctuates around the mean, does not follow a normal distribution, and has obvious volatility aggregation characteristics. Descriptive statistics show that the original series is based on the median, with the upper and lower quartiles as the main ranges of volatility variation. From Figure 1 and Table 1, it can be seen that original volatility at this point is noisy, with the microstructure of the noise having a greater impact on the true volatility.

After obtaining the CSI 300 intraday volatility data with noise, the discrete volatility data with dependence are treated as a sample consisting of random curve-generated dependence functional dataset$X_i(t)$, according to the dependence functional volatility model constructed above. The data are first smoothed using a rough penalty method to estimate its mean function, long-run covariance function, principal component score, and other relevant elements used to represent the volatility trajectory. The Fourier basis is used as the basis function to consider overfitting problems. Considering the effect of noise, a rough penalty method based on generalised cross-validation (GCV) is used to estimate the characteristic function, which not only eliminates most of the noise but also provides the optimal number of bases and penalty parameter lambda. After obtaining an estimate of the principal components of the function, the function is fitted and reconstructed using the K-L expansion, which generates a function set consisting of 240 function curves.

Figure 2 shows the fitted volatility curves for a random ten-day phase dependent functional type. As can be seen from the figure, the fitted volatilities are neither overly smooth nor over-fitted, achieving both goals of filtering noise and accurately capturing the volatility information contained in the data.
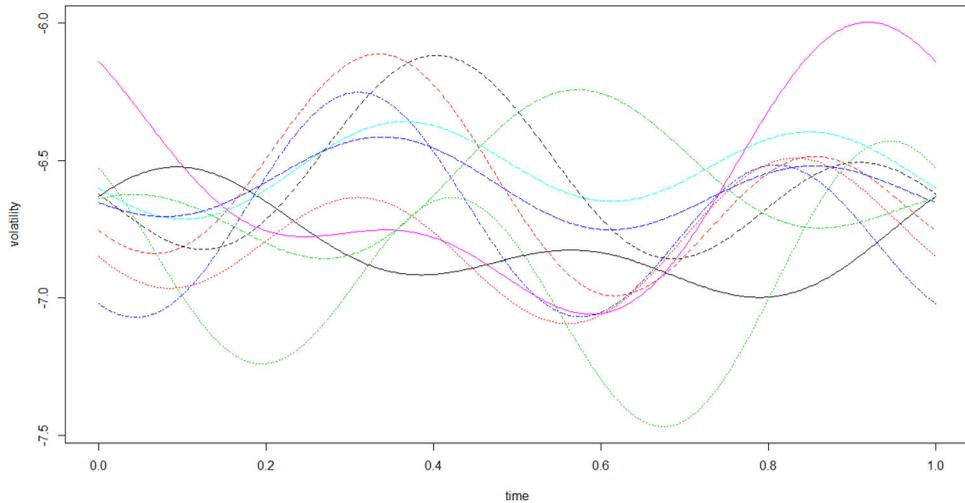
Figure 2. Volatility fitting curve (random 10 trading days).
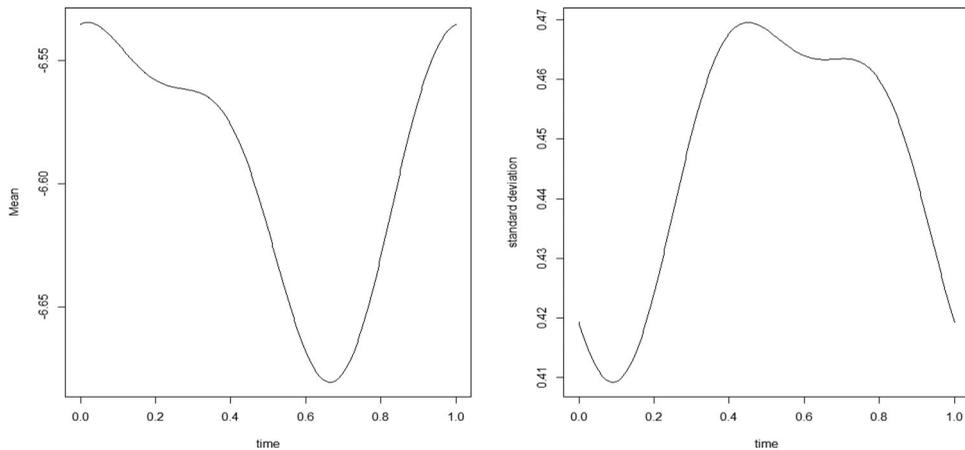Source: statistical results.

Figure 3. Mean and standard deviation functions.
Source: statistical results.

Descriptive research and analysis of the dependent functional data are carried out to obtain the mean function, derivative function, covariance function, and other statistics.

From Figure 3, the volatility trajectory curve generally shows a downward then an upward trend; however, there is a significant decline at noon, then a rebound process in the afternoon. In other words, the trajectory of intraday volatility shows a 'U-shaped' pattern of 'peaks' followed by 'troughs', and this pattern of volatility change is consistent with the 'calendar effects'. Andersen et al. (2003) suggested that the existence of a 'dip' in the intraday volatility curve at midday is due to a drop in return volatility caused by a drop in the trading volume. This theory also applies to the Chinese stock market, as the hour and a half long daily market closure between 11:30 and 13:00 is the main reason for the 'depression' in the intraday volatility curve
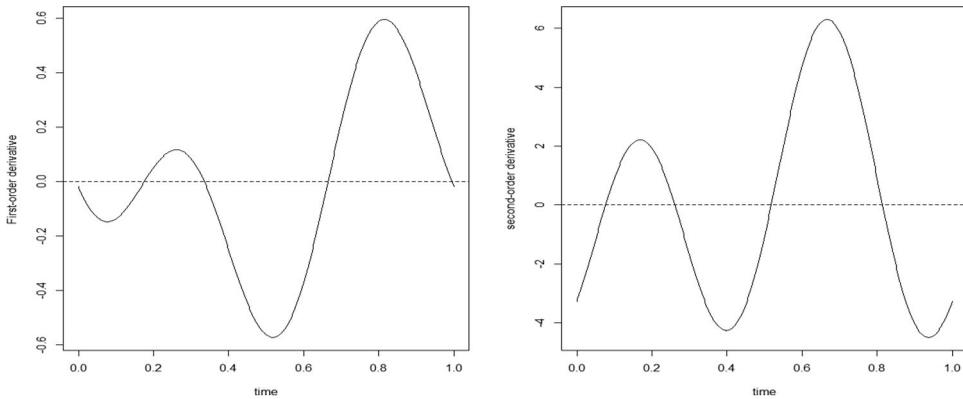
Figure 4. First-order and second-order derivatives of the volatility mean function.
Source: statistical results.

at midday. In addition, the nighttime closure of the stock index futures market also hinders the market price formation process, causing the volatility of the price series to be stronger at the beginning of the day than during other trading sessions and then to fall rapidly due to the interruption of trading during the lunchtime closure.

From Figure 4, the first- and second-order derivative functions of the mean volatility function reflect the speed of change and acceleration of the function, respectively. The fact that the derivative function is not constant at zero indicates that volatility is not constant but is always changing. The overall performance is as follows: volatility falls for a short period of time after the daily opening and then gradually starts to increase slowly; after a short break at lunchtime, volatility falls rapidly; before the market closes, volatility changes at an increased rate and gradually starts to rise. By analysing the characteristic curves of the mean and derivative functions of volatility after fitting the reconstruction, we find that the use of long-term covariance-based estimation methods as a theoretical basis can effectively reflect the pattern and speed of change of intraday volatility.

4.2. Principal component analysis of intraday volatility

After an initial exploratory analysis of the descriptive statistics, the covariance matrix is decomposed, and the most significant dependent functional principal components are extracted to infer the main patterns of intraday volatility. First, the number of principal components is determined. As shown in Figure 5, the eigenvalue fragmentation plot, the percentages of the first five eigenvalues can be judged to be 88.75%, 3.54%, 3.03%, 2.66%, and 2.01%, respectively, with the rate of decline levelling off from the sixth eigenvalue. From the cumulative variance explained ratio, it can be determined that the first three principal components explain approximately 95% of the variability in the original data. Considering the subsequent prediction of the principal component scores, the first three principal components are selected for the analysis in this study.

Figure 6 shows the principal and rotated principal component weight functions obtained using the maximum variance rotation method. The total variance explained
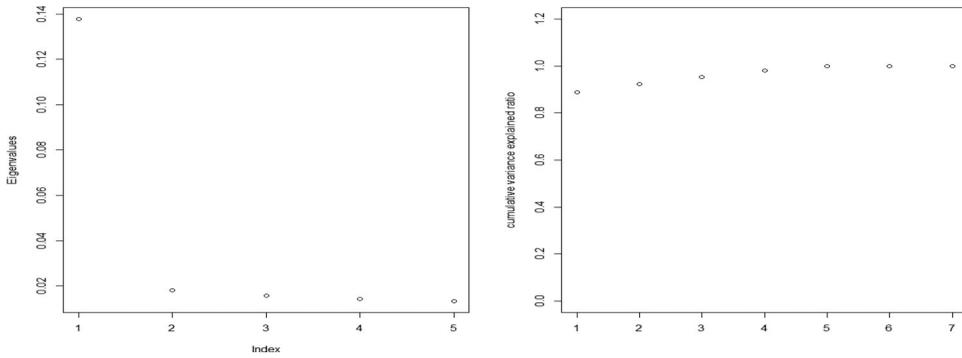
Figure 5. Eigenvalues and cumulative variance explained ratio.
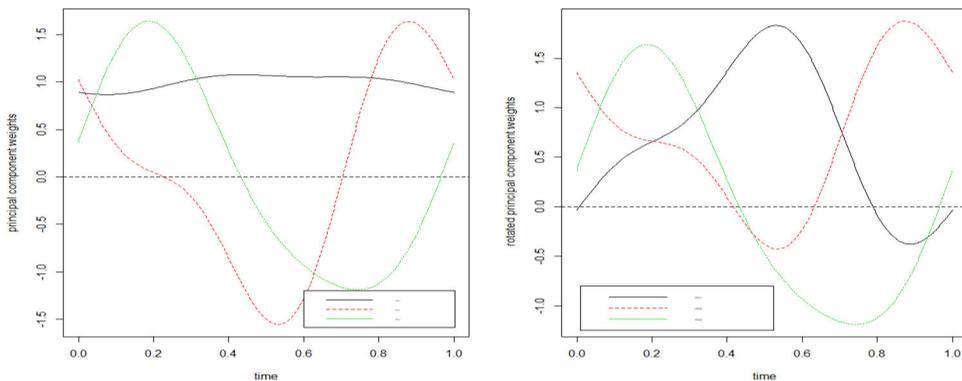Source: statistical results.

Figure 6. Principal component weight functions: unrotated (left) rotated (right).
Source: statistical results.

ratio remains unchanged after the rotation. It can be judged that the first principal component weight function (PC1) reflects the pattern of volatility variation at the afternoon opening after the 'midday effect'; the second principal component (PC2) is the most variable during the afternoon trading session; the third principal component (PC3) is the opposite of the second principal component and mainly captures the pattern of volatility variation during the morning session. The three principal component weight functions are orthogonal to each other, with almost no overlap in the variability information, cumulatively explaining 95.4% of the variability in the original data. After determining the number of principal components, we can plot the deviation of the principal component weight functions from the mean function, as shown in Figure 7, with '+' and '-' indicating the appropriate multiples of the principal component weight functions added to and subtracted from the mean function, respectively. A more visual representation of the volatility variation characteristics is provided by the first three principal component weight functions.

In summary, descriptive statistics and principal component analysis of the volatility profile further confirm that CSI 300 volatility has a distinctive 'intraday effect', that is, higher volatility during the opening, closing, and nearby time periods, showing a
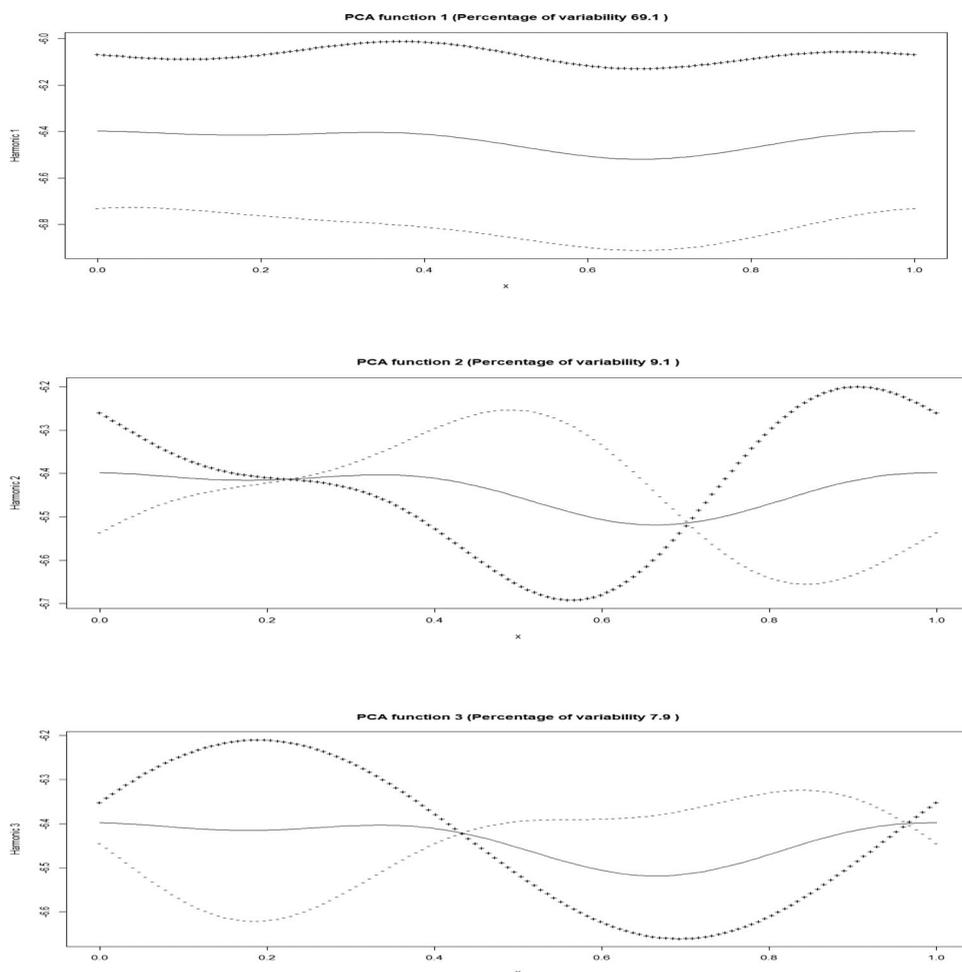
Figure 7. Function principal component weight function multiplier effect.
Source: statistical results.

'U-' or 'V-shaped characteristics'. In general, the first principal component controls volatility and is close to explaining 88.75% of the volatility variation. It can be confirmed that a theoretical framework based on dependent functional data provides more information on the pattern of intraday volatility variation.

4.3. Empirical evidence of intraday volatility forecasting

In the previous section, dependent functional data are analysed by obtaining noisy volatility trajectories from the variation in the return series, and then isolating smooth volatility trajectories through filtering the noise using non-parametric statistical methods. These smooth volatility trajectories are viewed as functional time series and are further investigated. The smoothness test is the basis of time series analysis. If the data are nonstationary, false conclusions from the pseudo-regression may be drawn. Therefore, the volatility series is first tested for smoothness, obtaining

$P = 0.3336$. The original hypothesis is that the series is smooth and can be subjected to subsequent predictive modelling analysis.

4.3.1. Comparison of prediction accuracy of principal component scores

In this study, three methods[2] (i.e. FPC, N-W, and WTB) are used to fit the intraday volatility principal component scores. Then, the estimated samples are used to make short-term forecasts of the principal component scores using the FLR method. Finally, the forecasts are compared with the results of the true volatility principal component scores. The data are divided into estimation and forecast (validation) samples. The model forecasting accuracy is evaluated using the mean absolute forecast error (MAFE) and the mean squared forecast error (MSFE), which are defined as

$$MAFE_j = \frac{1}{Jm} \sum_{j=1}^{J} \sum_{i=1}^{m} |x_{n+i}(t_j) - \hat{x}_{n+i|n+i-1}(t_j)|$$

$$MSFE_j = \frac{1}{Jm} \sum_{j=1}^{J} \sum_{i=1}^{m} [x_{n+i}(t_j) - \hat{x}_{n+i|n+i-1}(t_j)]^2.$$

In practice, new data are continually generated over time; however, this is not sufficient to form a complete functional curve. At this point, volatility data from the latest morning half-day can be considered based on FLR forecasts to predict volatility data for the remaining trading time periods of the same trading day. The data are divided into an estimation sample for the first 21 trading hours (morning period) and a forecast (validation) sample for the last 25 trading hours (afternoon period). Specifically, the volatility data from day 1 to the morning of day 201 are used for the first time to fit and forecast the afternoon curve of day 201, and so on, until the morning volatility data of day 240 are finally used to fit and forecast the afternoon curve of day 240. The point forecast accuracy is calculated for each validation sample by date averaging, with the horizontal and vertical axes representing time and forecast error, respectively.

Figure 8 shows the afternoon half-day of the trading day (corresponding to the positions of horizontal coordinates 25–47) obtained using the FLR method. From the error results, the principal component results for the FPC method, which does not consider sample dependence, have the highest error, alternating with oscillations up and down over time, with a tendency to increase the magnitude of error over time. The principal component forecasting results of the N-W and WTB estimation methods, which consider sample data dependence, generally remain at a lower level, with an average error smaller than that of the FPC, with the lowest level of error curve obtained from the WTB fitting proposed in this study, indicating a more accurate volatility fit. This again validates that the calendar effect at the intraday level causes volatility to change considerably closer to the opening and closing, leading to higher errors in forecasting.

In Table 2, we calculate and compare the prediction accuracy of the fitting of the principal component scores using different methods. In terms of both MAFE and MSFE, the N-W and WTB estimation methods, which consider the dependence of
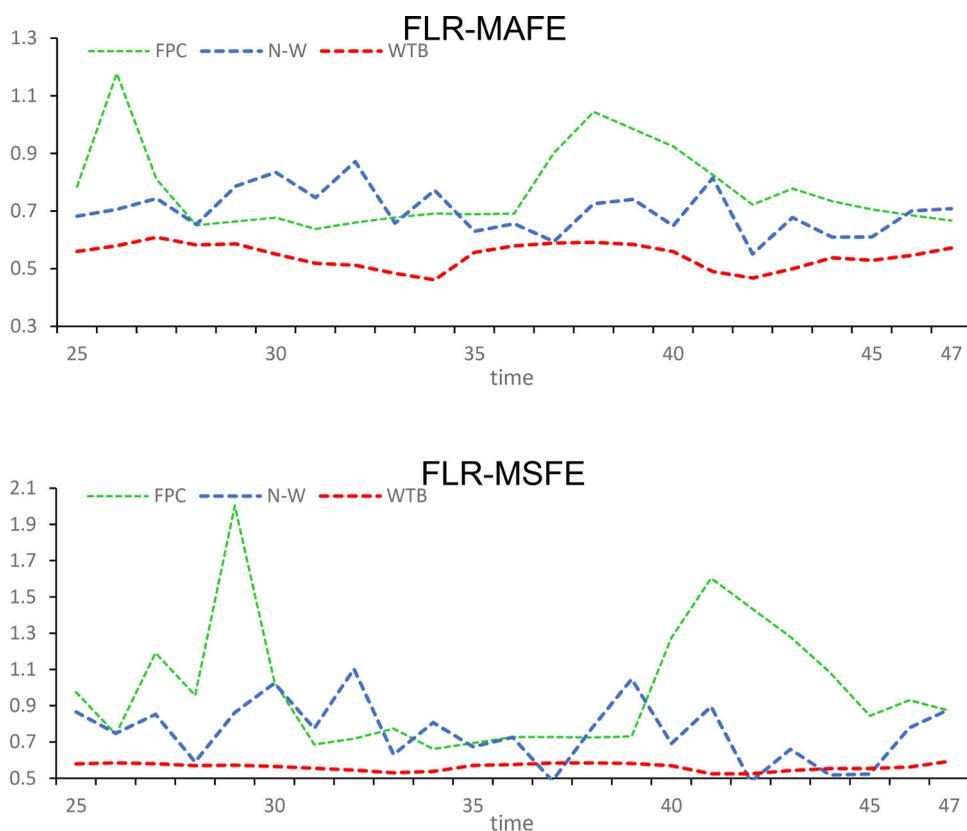
Figure 8. Prediction error analysis.
Source: statistical results.

Table 2. Descriptive statistics for FLR prediction accuracy.

	MAFE			MSFE		
FLR	FPC	N-W	WTB	FPC	N-W	WTB
minimum	0.6381	0.5236	0.4615	0.6619	0.4728	0.4952
lower quartile	0.6810	0.6396	0.5141	0.7075	0.6208	0.5551
median	0.7168	0.6873	0.5492	0.8243	0.7613	0.5739
average value	0.7736	0.6876	0.5457	0.8110	0.7639	0.5628
upper quartile	0.7416	0.74	0.5837	0.8940	0.8685	0.5924
maximum	1.1718	0.8945	0.6092	2.0811	1.3475	0.5933

Source: statistical results.

the sample data, have smaller errors than the FPC estimation methods that do not consider the dependence of the sample. Specifically, the WTB estimation method proposed in this study has the smallest error. Combined with the above figures, we find that the volatility model proposed in this study has the best prediction accuracy when fitting volatility at the intraday level. Additionally, the model treats the data as an implementation of a continuous stochastic process, which has unparalleled advantages over traditional methods.

Table 3. Profitability.

Profitability	FPC	N-W	WTB
Minimum	0.09848	0.582729	0.580343
average	0.334414	0.649151	0.692889
maximum	0.482973	0.694532	0.774226
Sum	7.022695	13.63217	14.55068

Source: statistical results.

4.3.2. Profitability test

This subsection constructs identical trading strategies to compare profitability under different models. We refer to the buy-and-sell trading strategy introduced by Bao et al. (2017). This strategy suggests that investors buy futures contracts when the forecast value at the point of trading is higher than the actual value and, conversely, choose to sell futures contracts. The total returns are obtained by accumulating the return on each trade $R = (P_{buy} + P_{sell}) * 100$.

$$P_{buy} = \sum_{t-1}^{b} \frac{(\hat{x}_{n+i|n+i-1}(t_j) - x_{n+i}(t_j) + (x_{n+i}(t_j) * B + \hat{x}_{n+i|n+i-1}(t_j) * S)}{x_{n+i}(t_j)},$$
$$P_{sell} = \sum_{t-1}^{s} \frac{(x_{n+i}(t_j) - \hat{x}_{n+i|n+i-1}(t_j) + (\hat{x}_{n+i|n+i-1}(t_j) * B + x_{n+i}(t_j) * S)}{\hat{x}_{n+i|n+i-1}(t_j)},$$

To account for the impact of transaction costs on profits, B and S [3] are set as the bid and ask transaction costs, respectively, in the above equation. To ensure the consistency of the test, in this section, we perform the empirical test of returns using the sample data of the last 40 days of forecasts in the same time period as in the previous section to compare the trading returns under the three models, FPC, N-W, and WTB.

Table 3 lists the profitability of the different models using the forecast sample. Other influencing factors, such as transaction costs and fees are the same under all three models. The combined average and total returns show that FPC performs unsatisfactorily overall, with the lowest level of returns. Both N-W and WTB achieve closer levels of returns. However, the WTB method achieves the highest profit returns compared to the other two models, with a daily return of 0.69% and a total return of 14.5%. Thus, in terms of both forecasting accuracy and profitability, our results demonstrate that the WTB long-term covariance-based volatility model outperforms other volatility models.

5. Conclusions

Modelling and forecasting volatility is of great practical importance for understanding the short-term volatility patterns of financial markets, optimising investment strategies, and avoiding market risks. In view of this, this study considers the dependence and functional characteristics of high-frequency data, uses long-term covariance estimation statistics based on the Bartlett kernel without truncation, modifies functional data analysis method under independent homogeneous distribution conditions,

constructs a new dependence functional volatility model, and thoroughly analyzes the intraday volatility variation pattern of CSI 300.

The empirical results show that the CSI 300 intraday volatility trajectory has typical intraday characteristics and is similar to the common 'U-shaped' intraday volatility pattern. In addition, compared to traditional functional data analysis methods under the assumption of an independent homogeneous distribution, the dependent function volatility model proposed in this paper, which considers the dependence of sample data, has non-parametric characteristics while satisfying the stochastic characteristics of volatility. Moreover, our proposed model can more accurately portray the dynamic regularity of the intraday volatility curve and outperforms other volatility models in terms of forecasting accuracy and return performance.

Our proposed dependent functional data analysis method greatly enriches the existing data analysis methodology literature and can effectively correct estimation bias and reduce information loss. To explore the dynamics of intraday volatility and improve the accuracy of volatility estimation, future research can further use functional data analysis methods to decompose the jump components of volatility at the intraday level.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes

1. The horizontal coordinates in the chart refer to the trading time period. That is, the horizontal 0.0 point corresponds to the first volatility point at 9:40 a.m. opening and the horizontal 1.0 point corresponds to the last volatility point at 15:00 p.m. closing.
2. FPC represents the principal component estimation method based on functions under the I.I.D condition, N-W represents the long-term covariance estimation method based on Newey-West estimation formula, and WTB represents the long-term covariance estimation method based on untruncated Bartlett kernel.
3. Bao, W., Yue, J., &Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory. Set the cost of the spot market at 0.25% buying and 0.45% selling.

Funding

References

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, *71*(2), 579–625. https://doi.org/10.1111/1468-0262.00418

Aue, A., Norinho, D. D., & Hörmann, S. (2015). On the prediction of stationary functional time series. *Journal of the American Statistical Association*, *110*(509), 378–392. https://doi.org/10.1080/01621459.2014.909317

Bao, W., Yue, J., & Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PloS ONE*, *12*(7), e0180944. https://doi.org/10.1371/journal.pone.0180944

Besse, P. (1992). PCA stability and choice of dimensionality. *Statistics & Probability Letters*, *13*(5), 405–410. https://doi.org/10.1016/0167-7152(92)90115-L

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327. https://doi.org/10.1016/0304-4076(86)90063-1

Bollerslev, T., & Zhou, H. (2006). Volatility puzzles: A simple framework for gauging return-volatility regressions. *Journal of Econometrics*, *131*(1–2), 123–150. https://doi.org/10.1016/j.jeconom.2005.01.006

Cerovecki, C., Francq, C., Hörmann, S., & Zakoian, J. M. (2019). Functional GARCH models: The quasi-likelihood approach and its applications. *Journal of Econometrics*, *209*(2), 353–375. https://doi.org/10.1016/j.jeconom.2019.01.006

Corsi, F. (2008). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, *7*(2), 174–196. https://doi.org/10.1093/jjfinec/nbp001

Das, S., Demirer, R., Gupta, R., & Mangisa, S. (2019). The effect of global crises on stock market correlations: Evidence from scalar regressions via functional data analysis. *Structural Change and Economic Dynamics*, *50*, 132–147. https://doi.org/10.1016/j.strueco.2019.05.007

Dauxois, J., Pousse, A., & Romain, Y. (1982). Asymptotic theory for the principal component analysis of a vector random function: Some applications to statistical inference. *Journal of Multivariate Analysis*, *12*(1), 136–154. https://doi.org/10.1016/0047-259X(82)90088-4

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, *50*(4), 987–1007. https://doi.org/10.2307/1912773

He, G., Müller, H. G., & Wang, J. L. (2003). Functional canonical analysis for square integrable stochastic processes. *Journal of Multivariate Analysis*, *85*(1), 54–77. https://doi.org/10.1016/S0047-259X(02)00056-8

He, G., Müller, H. G., & Wang, J. L. (2004). Methods of canonical analysis for functional data. *Journal of Statistical Planning and Inference*, *122*(1–2), 141–159. https://doi.org/10.1016/j.jspi.2003.06.003

Hörmann, S., & Kokoszka, P. (2010). Weakly dependent functional data. *Annals of Statistics*, *38*(3), 1845–1884.

Horvath, L., & Kokoszka, P. (2012). *Inference for functional data with applications.* Springer.

Huang, S. P., Quek, S. T., & Phoon, K. K. (2001). Convergence study of the truncated Karhunen-Loeve expansion for simulation of stochastic processes. *International Journal for Numerical Methods in Engineering*, *52*(9), 1029–1043. https://doi.org/10.1002/nme.255

Hyndman, R. J., & Shang, H. L. (2009). Forecasting functional time series. Forecasting Functional Time Series. *Journal of the Korean Statistical Society*, *38*(3), 199–211. https://doi.org/10.1016/j.jkss.2009.06.002

Kambouroudis, D. S., McMillan, D. G., & Tsakou, K. (2021). Forecasting realized volatility: The role of implied volatility, leverage effect, overnight returns, and volatility of realized volatility. *Journal of Futures Markets*, *41*(10), 1618–1639. https://doi.org/10.1002/fut.22241

Kim, J. M., Jun, C., & Lee, J. (2021). Forecasting the volatility of the cryptocurrency market by GARCH and stochastic volatility. *Mathematics*, *9*(14), 1614. https://doi.org/10.3390/math9141614

Kokoszka, P., & Young, G. (2017). Testing trend stationarity of functional time series with application to yield and daily price curves. *Statistics and Its Interface*, *10*(1), 81–92. https://doi.org/10.4310/SII.2017.v10.n1.a8

Li, X., Li, D., Zhang, X., Wei, G., Bai, L., & Wei, Y. (2021). Forecasting regular and extreme gold price volatility: The roles of asymmetry, extreme event, and jump. *Journal of Forecasting*, *40*(8), 1501–1523. https://doi.org/10.1002/for.2781

Lin, J., & Gong, Z. (2017). A research on forecasting of Shanghai zinc futures price based on artificial neural network. *Theory and Practice of Finance and Economics*, *38*(2), 54–57.

Madden, S. (2012). From databases to big data. *IEEE Internet Computing*, *16*(3), 4–6. https://doi.org/10.1109/MIC.2012.50

Müller, H. G., Sen, R., & Stadtmüller, U. (2011). Functional data analysis for volatility. *Journal of Econometrics*, *165*(2), 233–245. https://doi.org/10.1016/j.jeconom.2011.08.002

Nie, S. Y., & Li, M. H. (2019). Causal analysis of overnight and lunch break information for Chinese stock indexes and futures. *Journal of Applied Statistics and Management*, *38*(04), 719–731.

Poon, S. H., & Granger, C. W. J. (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, *41*(2), 478–539. https://doi.org/10.1257/.41.2.478

Ramsay, J. O. (1982). When the data are functions. *Psychometrika*, *47*(4), 379–396. https://doi.org/10.1007/BF02293704

Ramsay, J. O., & Dalzell, C. J. (1991). Some tools for functional data analysis. *Journal of the Royal Statistical Society: Series B*, *53*(3), 539–561.

Ramsay, J. O., & Silverman, B. W. (2002). *Applied functional data analysis: Methods and case studies* (2nd ed.). Springer.

Shang, H. L. (2017). Forecasting intraday S&P 500 index returns: A functional time series approach. *Journal of Forecasting*, *36*(7), 741–755. https://doi.org/10.1002/for.2467

Shang, H. L. (2018). Bootstrap methods for stationary functional time series. *Statistics and Computing*, *28*(1), 1–10. https://doi.org/10.1007/s11222-016-9712-8

Shang, H. L., & Hyndman, R. J. (2011). Nonparametric time series forecasting with dynamic updating. *Mathematics and Computers in Simulation*, *81*(7), 1310–1324. https://doi.org/10.1016/j.matcom.2010.04.027

Shang, H. L., Yang, Y., & Kearney, F. (2019). Intraday forecasts of a volatility index: Functional time series methods with dynamic updating. *Annals of Operations Research*, *282*(1–2), 331–354. https://doi.org/10.1007/s10479-018-3108-4

Shen, H. (2009). On Modeling and forecasting time series of smooth curves. *Technometrics*, *51*(3), 227–238. https://doi.org/10.1198/tech.2009.08100

Shimada, J., Tsukuda, Y., & Miyakoshi, T. (2009). Asymmetric international transmission in the conditional mean and volatility to the Japanese market from the US: EGARCH versus SV models. *The Singapore Economic Review*, *54*(01), 123–134. https://doi.org/10.1142/S0217590809003227

Tsay, R. S. (2016). Some methods for analyzing big dependent data. *Journal of Business & Economic Statistics*, *34*(4), 673–688. https://doi.org/10.1080/07350015.2016.1148040

Wang, S. S., Wang, J. B., & Guang-Lu, L. I. (2018). Research and prediction of high frequency intraday yield of the CSI 300 index futures based on ARMA model. *Journal of North China Electric Power University (Social Sciences)* , 3, 71–79.

Wang, X., Wang, X., Li, B., & Bai, Z. (2019). 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. https://doi.org/10.1108/CFRI-07-2018-0069

Wang, Y., Xiang, Y., Lei, X., & Zhou, Y. (2021). Volatility analysis based on GARCH-type models: Evidence from the Chinese stock market. *Economic Research-Ekonomska Istraživanja*, 1–25. https://doi.org/10.1080/1331677X.2021.1967771

Wei, Y. (2012). Forecasting volatility of fuel oil futures in China: GARCH-type, SV or realized volatility models? *Physica A: Statistical Mechanics and Its Applications*, *391*(22), 5546–5556. https://doi.org/10.1016/j.physa.2011.08.071

Wu, X., Zhou, H., & Wang, S. (2018). Estimation of market prices of risks in the GARCH diffusion model. *Economic Research-Ekonomska Istraživanja*, *31*(1), 15–36. https://doi.org/10.1080/1331677X.2017.1421989