Hindawi
Security and Communication Networks
Volume 2023, Article ID 9865347, 1 page
https://doi.org/10.1155/2023/9865347

Retraction

Retracted: Measurement of Economic Fluctuations Based on High-Frequency Financial Time Series

Security and Communication Networks

Received 8 August 2023; Accepted 8 August 2023; Published 9 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] H. Zhang, "Measurement of Economic Fluctuations Based on High-Frequency Financial Time Series," *Security and Communication Networks*, vol. 2022, Article ID 9310697, 18 pages, 2022.

Hindawi
Security and Communication Networks
Volume 2022, Article ID 9310697, 18 pages
https://doi.org/10.1155/2022/9310697

Research Article

Measurement of Economic Fluctuations Based on High-Frequency Financial Time Series

HuiLi Zhang

Yellow River Conservancy Technical Institute, KaiFeng, 475000, China

Correspondence should be addressed to HuiLi Zhang; zhanghuili@yrcti.edu.cn

Received 15 April 2022; Accepted 12 May 2022; Published 13 June 2022

Academic Editor: Fang Liu

In order to improve the analysis and forecasting effect of economic fluctuations, this paper combines the high-frequency financial sequence algorithm to conduct measured analysis of economic fluctuations. Under the continuous jump-diffusion price model, this paper considers the jump part and the continuous part in the process of asset pricing. Moreover, for the jump part, this paper uses the wavelet method to analyze the observation data, and obtains the estimator of the second-order jump covariation difference matrix and its convergence speed. For the continuous part, this paper adopts a two-scale realized volatility model under the continuous price model. In addition, this paper verifies the effect of the intelligent model proposed in this paper through simulation experiments. The simulation data shows that the economic fluctuation analysis system based on high-frequency financial time series proposed in this paper has good economic analysis and economic forecasting effects.

1. Introduction

When the economy suffers from monetary disruptions, such as an unexpected increase in the money supply, the increase in the money supply increases the aggregate consumer demand for commodities, that is, commodity market competition intensifies. At the same time, an expansion in aggregate demand leads to an increase in the general price level in the commodity market, but the relative price of each commodity does not actually change. Due to the incompleteness and asymmetry of information, a single manufacturer cannot observe the rise in the prices of all commodities, and mistakenly believes that the prices of the commodities it produces will rise. As a result, manufacturers will have wrong expectations about the demand for their own goods. When the price rises above the manufacturer's expected price, the manufacturer will increase investment and employment, and expand production. If all firms took such action, the level of output and employment would be higher than in competitive equilibrium. In such a case, the economy deviates upward from the potential level and the economy expands. Conversely, if the money supply falls unexpectedly, the general price level in the commodity market falls. When prices fall and fall below firms' expected prices, firms scale back production, output and employment fall below full employment levels, the economy deviates downward from potential, and the economy contracts.

Supply-side shocks, especially productivity shocks, are also one of the important factors affecting economic volatility. When a positive technological shock occurs in the economy, it will lead to an increase in the marginal product of labor. In the labor market, the demand for labor increases and real wages rise. If the economic person expects the technical shock to be temporary, the future real wages may fall, that is, the current leisure price will be higher, and the economic person will tend to replace the current leisure with future leisure. The result is that the economic man reduces the leisure time of the current period and provides more labor supply. As a result, current employment increases and total output increases. If the economic man predicts that the technological shock will be permanent, producers will invest more because of the higher marginal return on capital in the future. The increase in demand for investment goods raises the real interest rate. Consumers believe that the price of current leisure time will rise, and consumers will replace current leisure time with future leisure time. As a result, employment increased and output increased. When a negative technical shock occurs in the economy, if the technical shock is temporary, the demand for labor in the labor market will decrease, employment will decline, and output will decline. If the negative technology shock is permanent, producers will reduce investment. The decline in demand for investment goods leads to a decline in the real interest rate. Consumers believe that the price of current leisure time has fallen, so consumers will increase current leisure time and reduce labor supply. Reduced labor supply and reduced employment, resulting in lower output.

This paper combines the high-frequency financial sequence algorithm to conduct measured analysis of economic fluctuations, and builds an intelligent economic analysis and forecasting system to promote the analysis and forecasting effect of subsequent economic fluctuations.

2. Related Work

When a country's economy is very dependent on external demand, once the international economic environment deteriorates and external demand weakens, it is difficult to make up for the decline in external demand by increasing domestic demand in the short term. Therefore, higher trade openness means that changes in final demand will become more difficult. The higher the stability, the stronger the volatility of the macro economy [1]. Trade openness helps a country absorb advanced technologies from developed countries, improve resource allocation efficiency, and prompt the government to implement better macro-control policies, thereby improving the ability to cope with domestic shocks and helping to ease economic fluctuations [2]. The impact of trade opening on economic fluctuations is uncertain, which depends on whether potential external shocks can materialize and the magnitude of the shocks. According to the theory of real business cycles, technological shocks are the main source of economic fluctuations [3]. Technological progress is mainly divided into independent innovation and technology introduction. The latter has the characteristics of low cost and fast speed compared with the former. Therefore, developers often adopt the method of importing technology from developed countries to achieve economic catch-up [4]. For a long time, technology introduction has also been used as the main way of technological progress, which makes technological shocks mainly come from external. For developers, technological progress under technology introduction does not always bring about economic fluctuations [5]. Generally speaking, technology introduction improves the efficiency of resource allocation and the potential growth rate of the economy, can maintain a stable long-term economic growth path, and also strengthens the macroeconomy's ability to resist other shocks. However, in the period of economic transformation, the introduction of technologies supporting the old economic growth model will bring certain obstacles to the adjustment of economic structure [6]. The negative impact of technology introduction on economic development will gradually appear, and may cause economic fluctuations [7]. The impact of technology introduction on economic fluctuations mainly depends on the stage of economic development. Short-term international capital is highly speculative, and seeking arbitrage is the main incentive for its flow [8]. When the economy is prosperous, a large amount of short-term international capital flows to seek arbitrage opportunities, resulting in flooding of liquidity, pushing up asset prices, and causing economic overheating [9]. When the economy is depressed, there will be a massive outflow of short-term international capital, triggering market panic and falling asset prices, further aggravating the downward pressure on the economy. The pro-cyclical feature of short-term international capital makes it easy to become an amplifier of the economic cycle, so generally speaking, short-term international capital flows will bring about economic fluctuations [10].

There are a lot of unstable mechanisms in the economy, these unstable mechanisms may not only cause economic fluctuations; but also when the economy fluctuates, the unstable mechanisms will lead to the continuous expansion of economic fluctuations [11]. The research constructs a model of instability mechanism in Harrods economy from the perspective of investment demand. The current output or aggregate demand in this model is determined by the current investment through a multiplier effect. To define a capacity utilization rate, the current capacity utilization rate is equal to the ratio of the current total demand to the current capacity, where the current capacity is defined as the product of the output-capital ratio and the capital existing in the previous period [12]. The capital accumulation equation determines that the capital stock in the current period is equal to the capital stock in the previous period minus depreciation plus additional investment in the current period. At the end of the model, an investment equation is also constructed, and the investment is negatively correlated with the production capacity, that is to say, the investment depends on the capacity utilization rate of the previous period [13]. When examining the stability of the model, it is found that the system is unstable. The study found that enterprises will not only improve their own production capacity through investment behavior, but also bring about an increase in total social demand through the multiplier effect [14]. Generally speaking, the output-capital ratio will not be greater than 1, because the return on investment is often not realized in one period; on the other hand, the investment multiplier is often greater than 1. This means: when the economy is good, companies will increase investment. The total demand increment created by the investment behavior through the multiplier effect will be greater than the actual increase in the production capacity of the enterprise [15]. The result is that companies will continue to invest more, the economy will grow hotter, and the economy will move further and further away from equilibrium. Conversely, when there is a recession, companies reduce investment. The decline in aggregate demand will outweigh the decline in the productive capacity of firms, with the result that investment continues to decline and the economy becomes more sluggish. The price equation is constructed. Add price and monetary policy into the investment equation as stabilizing mechanisms, and test the stability of the model. Finally, it was found that when the economy fluctuated, price adjustment as a “weak stability mechanism” could not make the economy stable [16]. The expansion of the economy will be accompanied by the accumulation of debts. When the accumulation of debts reaches a certain level (such as the Minsky moment), a large amount of debt in the field of international lending exceeds the borrower’s own repayment ability, resulting in inability to repay the debt or the need to defer debt repayment. The phenomenon will inevitably bring huge pressure to the economy, and the crisis will break out [17]. In order to pursue GDP and stimulate economic growth, monetary authorities will implement loose monetary policies, mainly for two reasons: a large number of currency issuance and low interest rates. Stimulated by low interest rates, the manufacturing industry borrows money to buy equipment and build factories, real estate companies borrow money to buy land and build a large number of houses, and the government borrows money for infrastructure construction.

3. Estimation of market microstructure noise error under high frequency financial data

Due to the existence of market microstructure noise, there is a certain deviation between the high-frequency financial data we observe and collect and the unobservable real data. The time period [0, T] is divided into N points, that is, $0 = t_0 < t_1 < t_2 < \cdots t_N = T$, where $\Delta_{t_t} = t_1 - t_{i-1}$, $i = 1, 2, \ldots, N$.

Then, the observed value of the log price at time t_t is:

$$Y_{t_1} = X_{t_1} + \varepsilon_{t_t}, i = 0, 1, \ldots, N, t \in [0, T]. \tag{1}$$

Among them, ε_4 is the market microstructure noise. For unobservable real data X_t, we assume that the logarithmic price process obeys a stochastic differential formula, that is:

$$dX_t = \mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t. \tag{2}$$

Among them, μ_t is the drift coefficient, σ_t is the instantaneous volatility, and W_t is a standard Brownian motion.

In fact, it exists in a semi-martingale form as follows:

$$X_t = \int_0 \mu_s \mathrm{d}s + \int_0 \sigma_s \mathrm{d}W_s. \tag{3}$$

We assume that the asset price data is observed at time interval δ on [0, T], which is denoted as:

$$R^* (i, \delta) = \ln S_i - \ln S_{i-\delta} = X_t - X_{i-\delta}. \tag{4}$$

The above formula is the continuous compound interest rate of financial assets in the i-th time interval, that is, the logarithmic rate of return. In the financial environment of risk-free arbitrage, the logarithmic rate of return R° of financial assets obeys a special semi-martingale process. If (Ω, I, P) is a complete probability space, and the information filter $(I_t)_{t\in[0,T]} \subseteq I$ is an ascending σ subalgebraic sequence, I_t is P complete and right continuous. $S_t\,(t \in [0, T]$ is the same as the above definition, it represents the price process of financial native assets in this space, that is, S_t is in the information set I_t at time t.

The logarithmic rate of return of investing in a financial native asset in the Δ time period on day t is:

$$R^* (i, \Delta) = X_1 - X_{i-\Delta}. \tag{5}$$

Among them, t represents the t day, $\Delta > 0$.

When $\Delta = 1$, the realized volatility on day t is defined as follows:

$$[Y, Y]_i = \sum_{i=1}^{n} \left(Y_{t_1} - Y_{t_{t-1}}\right)^2. \tag{6}$$

The above formula is the realized volatility on day t.

The observed (logarithmic) rate of return is now defined as:

$$r_{i,N} = Y_{t_1} - Y_{t_{t-1}}, \quad i = 1, 2, \ldots, N. \tag{7}$$

Accordingly, the effective rate of return is:

$$r^*_{i,N} = X_{t_i} - X_{t_H}, \quad i = 1, 2, \ldots, N. \tag{8}$$

When $\Delta = 1$, the realized volatility on day t is defined as follows:

$$(\mathrm{RV})\mathrm{RV} = \sum_{i=1}^{N} r_{i,N}^2. \tag{9}$$

The noise series of returns is the difference between the observed return and the effective return and is defined as $e_{i,N} = r_{i,N} - r^*_{i,N}$, $i = 1, 2, \ldots, N$.

It can be seen that $e_{i,N} = \varepsilon_{t_t} - \varepsilon_{i_{i-1}}$, $i = 1, 2, \ldots, N$, that is, the noise series of returns has MA(1). We assume that ε_t is an i.i.d. sequence that satisfies:

$$E\varepsilon_{t_1} = E\varepsilon = 0, E\varepsilon_{t_1}^2 = E\varepsilon^2. \tag{10}$$

Moreover, it is not related to the effective price X_t. The following formula is obtained:

$$\mathrm{RV} = 2\,\mathrm{NE}\,\varepsilon^2 + O_p\left(N^{1/2}\right). \tag{11}$$

In fact, their results are based on the following formula:

$$RV = \sum_{i=1}^{N} r_{i,N}^2 = \sum_{i=1}^{N} r_{i,N}{}^2 + 2\sum_{i=1}^{N} r^{\bullet}_{i,N} e_{i,N} + \sum_{i=1}^{N} e_{i,N}^2. \tag{12}$$

When the sampling frequency is very high, that is, $(N \longrightarrow \infty)$, and the sum of the squares of the noise constitutes the main part of the above formula as $(\mathrm{RV} \longrightarrow \sum_{i=1}^{N} e_{i,N}^2)$. Therefore, the RV of the highest frequency data can be used as an estimate of the noise sequence variance.

The Realized Volatility (RV) based on observational data is as follows:

$$\sum_{i=1}^{N} r_{i,N}^2 = \sum_{i=1}^{N} r^*_{i,N}{}^2 + 2\sum_{i=1}^{N} r^*_{i,N} e_{i,N} + \sum_{i=1}^{N} e_{i,N}^2. \tag{13}$$

Its conditional mean and variance are as follows:

$$E(\mathrm{RV}|X) = \mathrm{RV}_X + 2\,\mathrm{NE}\,\varepsilon^2,$$
$$\mathrm{Var}(\mathrm{RV}|X) = 4\,\mathrm{NE}\,\varepsilon^4 + O_p(1), \tag{14}$$

Among them, there are:

$$\mathrm{RV}_X = \sum_{i=1}^{N} {r_{i,N}}^2. \tag{15}$$

Therefore, for the efficient price process X, when $N \longrightarrow \infty$, we have its asymptotic normality as follows:

$$N^{-1/2}\left(\mathrm{RV} - 2\mathrm{NE}\,\varepsilon^2\right) \xrightarrow{d} 2\left(E\varepsilon^4\right)^{1/2} Z_{\mathrm{nolse}}. \tag{16}$$

From the above analysis, we can obtain the estimator of the variance of the noise series as follows:

$$\widehat{E}\varepsilon^2 = \frac{\mathrm{RV}}{2N}. \tag{17}$$

For a certain real data process X, when $N \longrightarrow \infty$, we have:

$$N^{1/}\left(\widehat{E}\varepsilon^2 - E\varepsilon^2\right) \xrightarrow{d} N\left(0, E\varepsilon^4\right). \tag{18}$$

Among them, the asymptotic variance $E\varepsilon^4$ can obtain a consistent estimator as follows:

$$\widehat{E}\varepsilon^4 = \frac{1}{2N}\sum_{i=1}^{N} r_{i,N}^4 - 3\left(\widehat{E}^2\right)^2. \tag{19}$$

We assume that the full set is $G, G = \{t_0, t_1, \ldots, t_N\}$, and it is divided into K disjoint subsets $G^{(k)}, k = 1, 2, \ldots, K$, that is, $G = \cup_{k=1}^{K} G^{(k)}$. Among them, when $k \neq l$, $G^{(k)} \cap G^{(l)} = \phi$.

The method for selecting the elements of $G^{(k)}$ in the k-th subset is to start with t_{k-1}, and then take every K steps as sampling points until T. That is to say, $G^{(k)} = \{t_{k-1}, t_{k-1+K}, t_{k-1+2K}, \ldots t_{k-1+n_k K}\}$.

Among them, $k = 1, 2, \ldots, K, n_k$ is an integer that makes $t_{k-1+n_k K}$ the last element in $G^{(k)}$. Thus, we obtain an estimate of volatility, that is:

$$\mathrm{RV}^{(\mathrm{avg})} = \frac{1}{K}\sum_{k=1}^{K} \mathrm{RV}^{(k)}. \tag{20}$$

Its mean and variance are:

$$E\left(\mathrm{RV}^{(\mathrm{arg})}|X\right) = \mathrm{RV}_X^{(\mathrm{ang})} + 2\overline{n}E\varepsilon^2,$$
$$\mathrm{Var}\left(\mathrm{RV}^{(\mathrm{avg})}|X\right) = \frac{1}{K^2}\sum_{k=1}^{K}\mathrm{Var}\left(RV^{(\mathrm{avg})}|X\right)$$
$$= 4\frac{\overline{n}}{K}E\varepsilon^4 + O_p\left(\frac{1}{K}\right). \tag{21}$$

Among them, $\overline{n} = 1/K\sum_{k=1}^{K} n_k = N - K + 1/K$.

From the estimator of the variance of the first noise series and 11, we have a revised estimator of volatility as follows:

$$\mathrm{RV}^{(adt)} = \mathrm{aRV}^{(avg)} - b\frac{\overline{n}}{N}\mathrm{RV}. \tag{22}$$

According to 7 and 12, there are:

$$E\left(\mathrm{RV}^{(\mathrm{at})}|X\right) = \mathrm{aE}\left(\mathrm{RV}^{\left(\mathrm{aq}_g\right)}|X\right) - b\frac{\overline{n}}{N}E(\mathrm{RV}|X)$$
$$= a\left(\mathrm{RV}_X^{(\mathrm{avg})} + 2\overline{n}E\varepsilon^2\right) - b\frac{\overline{n}}{N}\left(\mathrm{RV}_X + 2\mathrm{NE}\,\varepsilon^2\right)$$
$$= \mathrm{aRV}_X^{(\mathrm{ang})} - b\frac{\overline{n}}{N}\mathrm{RV}_X + 2(a-b)\overline{n}E\varepsilon^2. \tag{23}$$

In order to completely eliminate the effect of $E\varepsilon^2$, we reasonably take a = b, which is an unbiased estimate of volatility, so that we can get another estimator of $E\varepsilon^2$, as follows:

$$\widehat{E}\varepsilon^2_{(\mathrm{adj})} = \frac{1}{2}(N-\overline{n})^{-1}\left(\mathrm{RV} - \mathrm{RV}^{(\mathrm{arg})}\right). \tag{24}$$

Furthermore, it satisfies $E(\widehat{E}\varepsilon^2_{(\mathrm{adj})}|X) = E\varepsilon^2 + 1/2\,(N-\overline{n})^{-1}(\sqrt{\mathrm{RV}}_X - \mathrm{RV}_X^{(\mathrm{arg})})$., and it is higher-order unbiased.

For a certain real data process X, we assume that $N^{1/2}(\widehat{E}\varepsilon^2_{(\mathrm{adj})} - E\varepsilon^2)$. and $N^{1/2}(\widehat{E}\varepsilon^2 - E\varepsilon^2)$ have the same asymptotic distribution when $\max_i \Delta_{t_1} \longrightarrow 0\,(i = 1, 2, \ldots, N)$ and $N \longrightarrow \infty, N/K \longrightarrow \infty$.

When $K = O(N^{2/3})$, there are:

$$\widehat{E}\varepsilon^2_{(\mathrm{adj})} - E\varepsilon^2 = \left(\widehat{E}\varepsilon^2 - E\varepsilon^2\right)\left(1 + O\left(K^{-1}\right)\right) + O_p\left(KN^{-3/2}\right)$$
$$= \widehat{E}\varepsilon^2 - E\varepsilon^2 + O_p\left(N^{-1/2}K^{-1}\right) + O_p\left(\mathrm{K}N^{-3/2}\right)$$
$$= \widehat{E}\varepsilon^2 - E\varepsilon^2 + O_p\left(N^{-5/6}\right). \tag{25}$$

According to the previous definition, the noise sequence has a MA(A) structure. Therefore, under the condition, and the drift coefficient is not timed, $E(r_{i,N}r_{i+1,N}) = -E\varepsilon^2$, and the higher-order autocovariances are all 0. Thus, we have an estimator of $E\varepsilon^2$ as follows:

$$\widehat{E}\varepsilon^2_A = -\frac{1}{N-1}\sum_{i=1}^{N-1} r_{i,N}r_{i+1,N}. \tag{26}$$

We assume that Y_t and X_t satisfy 1 and 2, and it satisfies 4, X_t and ε_t are independent, we have:

$$\widehat{E}\varepsilon^2_A = -\frac{1}{N-1}\sum_{i=1}^{N-1} r_{i,N}r_{i+1,N} \xrightarrow{p} E\varepsilon^2,$$
$$N^{1/2}\left(\widehat{E}\varepsilon_1^2 - E\varepsilon^2\right) \xrightarrow{d} N\left(0, 5E\varepsilon^4\right). \tag{27}$$

The noise sequence ε_t is a covariance stationary process (weak stationary process) with mean 0, and its autocovariance function is defined as $\pi(s) = E(\varepsilon_t\varepsilon_{t+s})$.

For $s = 0, \pi(0) = E\varepsilon^2$ is the variance of the noise. Obviously, the noise sequence is a special case of it.

The estimator bias is $\mathrm{bias}(\mathrm{RV}) = E(\mathrm{RV} - \mathrm{RV}_X)$..

In fact, the full estimator bias is $E(\mathrm{RV} - \mathrm{IV}^*) = E(\mathrm{RV} - \mathrm{RV}_X) + E(\mathrm{RV}_X - \mathrm{IV}^*)$, where the former term is due to noise and the latter term is due to computational discretization. Without loss of generality, we examine the above definitions regarding the form of estimator bias arising from noise.

The correlation between noise and effective price is $\gamma_N^{(h)}$, and for $h = 0, \pm 1, \pm 2, \cdots$, we define $\gamma_N^{(h)} = E(r_{i,N} e_{i+h,N})$.

When $r_{i,N}^*$ and $e_{t,N}$ are independent, $\gamma_N^{(h)} = 0$, when $h = 0, \gamma_N^{(0)} = E(r_{i,N}^* e_{i,N})$.

We have:

$$\text{bias}(\text{RV}) = 2N\gamma_N^{(0)} + 2N(\pi(0) - \pi(\Delta_N)). \tag{28}$$

Among them, $\Delta_N = T/N$.

We assume that $\pi(\cdot)$ is derivable at 0 and $\lim_{N\longrightarrow\infty} \gamma_N^{(0)}/\Delta_N = \gamma_N$,, when $N \longrightarrow \infty$ and $\Delta_N \longrightarrow 0$, we have:

$$\begin{aligned}\lim_{N\longrightarrow\infty} \text{bias}(\text{RV}) &= 2N\Delta_N \lim_{N\longrightarrow\infty} \gamma_N^{(0)}/\Delta_N \\ &\quad + 2N\Delta_N \lim_{N\longrightarrow\infty} (\pi(0) - \pi(\Delta_N))/\Delta_N \\ &= 2T(\gamma_N - \pi'(0)).\end{aligned} \tag{29}$$

The noise sequence satisfies the following conditions:

(1) For the finite number $\rho_0 \geq 0$, if $s > \rho_0$, then $\pi(s) = 0$;

(2) For finite numbers $\rho_0 \geq 0$ and $t_1 < t_2$ 3, if $s \notin [t_1 - \rho_0, t_2 + \rho_0]$, then $E(\varepsilon_t(X_{t_1} - X_{t_2})) = 0$;

(3) For $i \neq j, E(r_{i,N}^* r_{j,N}^*) = 0$.

We can see that the above conditions (1) and (2) say that the correlation of the noisy series process is limited to a period not exceeding ρ_0, and the condition (3) states that the effective price increments (over an arbitrarily long period) are uncorrelated. A method for RV estimation by introducing q_N-term autocorrelation is proposed.

When there exists an integer q_N such that $q_N/N > \rho_0$, we have:

$$\text{bias}\left(\text{RV}_{\text{AC},q_N}\right) = 0. \tag{30}$$

Among them, $\text{RV}_{\text{AC},q_N} = \sum_{i=1}^N r_{i,N}^2 + 2\sum_{h=1}^{q_N} N/N - h \sum_{i=1}^{N-h} r_{i,N} r_{i+h,N}$.

For w, N and q_N, the relational expression $q_N = \text{ceil}\,(w/\text{TN})$ should be satisfied. When the window parameter w is determined, N is given, and q_N can be defined. For N, it can be determined under the optimal sampling problem.

In fact, q_N is a typical choice: when $N \longrightarrow \infty$, $q_N/N \longrightarrow 0$, for example $q_N = \text{int}(4(N/100)^{2/9})$. However, we believe that correlation is specific to time period, it is independent of N, and it is inappropriate here.

Under the definition of estimation bias, there are:

$$\begin{aligned}\text{bias}\left(\text{RV}_{\text{NC},q_N}^{\widetilde{N}}\right) &= E\left(\text{RV}_{\text{NC},q_N}^{\widetilde{N}}\right) - E(\text{RV}_X) = 0 \\ \text{bias}\left(\text{RV}^{(N)}\right) &= E\left(\text{RV}^{(N)}\right) - E(\text{RV}_X) = 2N\gamma_N^{(0)} \\ &\quad + 2N(\pi(0) - \pi(\Delta_N)).\end{aligned} \tag{31}$$

Therefore, there is the following formula:

$$E\left(\text{RV}^{(N)}\right) - E\left(\text{RV}_{\text{AC},q_N}^{\widetilde{N}}\right) = 2N\gamma_N^{(0)} + 2N(\pi(0) - \pi(\Delta_N)). \tag{32}$$

Among them, N and $\widetilde{N}$ may be different or the same.

On the one hand, $\gamma_N^{(0)}$ represents the correlation between noise and effective price, and it remains unchanged over time. $\pi(\Delta_N)$ is the correlation of noise series in the time period, and the correlation between noise series is limited. If $\Delta_N > \rho_0$, then $\pi(\Delta_N) = 0$. Therefore, the appropriate frequency N_l is chosen such that $\Delta_N = T/N_l > \rho_0$, then $\pi(\Delta_{N_1}) = 0$. From formula (32), we have:

$$E\left(\text{RV}^{(N)}\right) - E\left(\text{RV}_{\text{AC},q_N}^{\widetilde{N}}\right) = 2N_l\left(\gamma_N^{(0)} + \pi(0)\right)\gamma_N^{(0)}. \tag{33}$$

On the other hand, when $N \longrightarrow \infty$, $\max \Delta_t \longrightarrow 0$ and thus $\pi(\Delta_N) \longrightarrow \pi(0)$. Therefore, we can choose the highest frequency N_h as possible, then by formula (32), we have:

$$E\left(\text{RV}^{(N_k)}\right) - E\left(\text{RV}_{\text{AC},q_R}^{\widetilde{N}}\right) \longrightarrow 2N_h\gamma_{N_k}^{(0)}. \tag{34}$$

From formula (33) and formula (34), the noise variance expression under the assumption of correlated series noise can be obtained, as follows:

$$\begin{aligned}\pi(0) &= \frac{E\left(\text{RV}^{(N_l)}\right) - E\left(\text{RV}_{\text{AC},q_N}^{\widetilde{N}}\right)}{2N_l} \\ &\quad - \lim_{N_h \longrightarrow \infty} \frac{E\left(\text{RV}^{(N_h)}\right) - E\left(\text{RV}_{\text{AC},q_N}^{\widetilde{N}}\right)}{2N_h}.\end{aligned} \tag{35}$$

Then, for a fairly large N_h and N_l and $\widetilde{N}$, the sample mean is used to replace the expected value, and the noise variance estimator under the assumption of correlated series noise can be obtained.

$$\widehat{\pi}(0) = \frac{\overline{\text{RV}}^{(N_l)} - \overline{\text{RV}}_{\text{AC},q_N}^{(\widetilde{N})}}{2N_l} - \frac{\overline{\text{RV}}^{(N_k)} - \overline{\text{RV}}_{\text{AC},q_N}^{(\widetilde{N})}}{2N_h}. \tag{36}$$

Among them, n represents the number of working days. For the frequency selection of N_h, N_l and $\widetilde{N}$, N_h is easy to determine, using the highest frequency available. N_l is the frequency at which noise correlation is removed. As long as N_l is taken to make $\Delta_{N_1} > \rho_0$, the influence of the correlation between noises on sampling estimation can be eliminated.

Due to the existence of microscopic noise in the market, there is a certain degree of deviation between the high-frequency financial data we observed and collected and the underlying real data, and the time period [0, 1] is divided as follows:

$$0 = t_0 < t_1 < \cdots < t_n = 1. \tag{37}$$

The observed value of the log price at time t_i is:

$$Y_{t_i} = X_{t_i} + \varepsilon_{t_1}, t_t = \frac{i}{n}, \quad i = 0, 1, \ldots, n.t \in [0, 1]. \tag{38}$$

Among them, ε_{t_t} is the microstructure noise, which satisfies $E\varepsilon_{t_1} = E\varepsilon = 0, E\varepsilon_{t_t}^2 = E\varepsilon^2$, and ε and the logarithmic price X are independent of each other.

$\Delta Y_t = Y_{t_i} - Y_{t_{i-1}}$, $(i = 1, 2, \ldots, n)$ is the logarithmic rate of return of financial assets over time period $[t_{i-1}, t_t]$ based on observational data.

The real logarithmic price X_t exists in a semi-martingale form:

$$X_t = \int_0 \mu_s ds + \int_0 \sigma_s dW_s + \sum_{l=1}^{N(t)} L_l, \quad t \in [0, 1]. \tag{39}$$

The three parts on the right side of the formula are the drift, diffusion and jump parts of X, respectively. For the diffusion part, W_t is a standard Brownian motion and the diffusion coefficient σ_t^2 is called the point volatility. For the jump part, $N(t)$ represents the number of jumps up to time t, and S represents the height of the l-th jump.

The logarithmic price process in the form of formula (39) has quadratic variation as follows:

$$[X, X] = \int_b \sigma_s^2 ds + \sum_{l=1}^{N(t)} L_l^2. \tag{40}$$

For two assets X_1, X_2, the observed value of the logarithmic price at time t_i is:

$$Y_{1t_i} = X_{1t_t} + \varepsilon_{1t_i}, Y_{2t_i} = X_{2t_i} + \varepsilon_{2t_i}, t_t = \frac{i}{n}, \quad i = 0, 1, \ldots, n. \tag{41}$$

Among them, ε_{k_i} is ii.d., satisfying $E\varepsilon_{k_i} = E\varepsilon_k = 0, E\varepsilon_{k_i}^{21} = E\varepsilon_k^2$, and ε_k and the corresponding logarithmic price X_k are independent of each other, $k = 1, 2$.

Its true logarithmic price X_{11} ` X_{21} also exists in semi-martingale form as formula (39), as follows:

$$\begin{aligned} X_{1t} &= \int_d \mu_{1s} ds + \int_b \sigma_{1s} dW_{1s} + \sum_{l=1}^{N_1(t)} L_{1l}, \quad t \in [0, 1], \\ X_{2t} &= \int_0 \mu_{2s} ds + \int_d \sigma_{2s} dW_{2s} + \sum_{i=1}^{N_2(t)} L_{2l}, \quad t \in [0, 1]. \end{aligned} \tag{42}$$

The corresponding logarithmic price process X_1 ` X_2 also has a quadratic covariance matrix as follows:

$$\begin{pmatrix} [X_1, X_1] & [X_1, X_2] \\ [X_2, X_1] & [X_2, X_2] \end{pmatrix} = \begin{pmatrix} \int_0^1 \sigma_{11s}^2 ds & \int_0^1 \sigma_{12s}^2 ds \\ \int_0^1 \sigma_{21s}^2 ds & \int_0^1 \sigma_{22s}^2 ds \end{pmatrix} + \begin{pmatrix} \sum_{l=1}^{N_1(t)} L_{1l}^2 & \sum_{l=1}^{N'(t)} L_{1l} L_{2l} \\ \sum_{l=1}^{N'(t)} L_{1l} L_{2l} & \sum_{l=1}^{N_2(t)} L_{2l}^2 \end{pmatrix}. \tag{43}$$

Among them, the set of corresponding jump points is denoted as A_1 ` $A_2, N'(t)$ represents the number of elements of $A_1 \cup A_2$. For the jump covariation matrix, we have:

$$\Gamma\Delta \begin{pmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{pmatrix} = \begin{pmatrix} \sum_{l=1}^{N_1(t)} L_{1l}^2 & \sum_{l=1}^{N'(t)} L_{ll} L_{2l} \\ \sum_{l=1}^{N'(t)} L_{1l} L_{2l} & \sum_{l=1}^{N_2(t)} L_{2l}^2 \end{pmatrix}. \tag{44}$$

Among them, the jump difference is $\xi_{kk} = \sum_{l=1}^{N_k(t)} L_{kl}{}^2$, $k = 1, 2$, and the jump covariance difference is $\xi_{12} = \xi_{21} = \sum_{l=1}^{N'(t)} L_{1l} L_{2l}$. For the volatility matrix, there are:

$$\theta = \begin{pmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{pmatrix} = \begin{pmatrix} \int_0 \sigma_{11s}^2 ds & \int_0 \sigma_{12s}^2 ds \\ \int_0 \sigma_{21s}^2 ds & \int_0 \sigma_{22s}^2 ds \end{pmatrix}. \tag{45}$$

Among them, $\int_0 \sigma_{12s}^2 ds = \int_0 \sigma_{21s}^2 ds$

The wavelet analysis method is a time-frequency localized analysis method with a fixed window size but its shape can be changed, and both the time window and the frequency window can be changed. That is, it has higher frequency resolution and lower time resolution in the low frequency part, and higher time resolution and lower frequency resolution in the high frequency part. Therefore, it is called a mathematical microscope. It not only inherits and develops the localization idea of short-time Fourier transform, but also has the characteristics of multi-resolution analysis. It uses multi-scale wavelet to analyze high-frequency data without sampling, there is no trouble of sampling, and it can also deal with mutation and noise problems at the same time. The wavelet transform is:

$$w_f(s, k) = \int_R f(x) g_{s,k}(x) dx. \tag{46}$$

Among them, there are:

$$g_{s,k}(x) = g\left(\frac{(x - k/s)}{\sqrt{s}}\right). \quad (47)$$

Among them, k is the time domain index, s is the scale. $K = 2^{-j}k, s = 2^{-j}$. If $g_{s,k}(x) = \varphi_{j,k}(x)$ or $g_{s,k}(x) = \psi_{j,k}(x)$, there are:

$$\varphi_{j,k} = 2^{j/2}\varphi\left(2^j x - k\right), \psi_{j,k} = 2^{J/2}\psi\left(2^j x - k\right). \quad (48)$$

Moreover, $\int_R \varphi_{j,k}(x) = 0, \int_R \psi_{j,k}(x) = 0, \varphi_{j,k}$ is called the parent wavelet, and $\psi_{j,k}$ is called the mother wavelet. For any function $f(x)$, we have:

$$f(x) = \sum_k^R c_k^{j_0}\varphi_{j_0,k}(x) + \sum_{j\geq j_0}\sum_k d_k^j \psi_{j,k}(x). \quad (49)$$

Among them, $c_k^{j_0}$ is the scale coefficient, d_k^j is the degree of detail coefficient. For the orthogonal wavelet family, there is $c_k^j = \int_R f(x)\varphi_{j,k}(x)\mathrm{d}x, d_k^j = \int_R f(x)\psi_{j,k}(x)\mathrm{d}x.$.

The estimates of the jump points τ_{1l}`$\tau_{2l}$ are $\hat{\tau}_{1l}$`$\hat{\tau}_{2l}$, and the number of elements $\hat{N}_k(t)$ of $\{\hat{\tau}_{kl}\}$ is recorded as the estimate of $N_k(t)$, k=1,2. Then, we use the estimated jump point to estimate the jump height of each jump, so that the jump variation $\hat{\xi}_{11}$`$\hat{\xi}_{22}$ and jump covariation difference $\hat{\xi}_{12}$`$\hat{\xi}_{21}$ can be estimated.

For $\delta_n > 0$, $\hat{I}_{k+} = [\hat{\tau}_{kl}, \hat{\tau}_{kl} + \delta_n), \hat{I}_{k-} = (\hat{\tau}_{kl} - \delta_n, \hat{\tau}_{kl}]k = 1, 2$, where m_{k+}, m_{k-} are the observation points of each asset in the interval $\hat{I}_{k+}, \hat{I}_{k-}$, k=1,2.

Therefore, the estimators of jump difference and jump covariation are:

$$\hat{\xi}_{kt} = \sum_{l=1}^{\hat{N}_k(t)} \hat{L}_{kl}^{\ 2}, \quad k = 1, 2\, \hat{\xi}_{12} = \hat{\xi}_{21} = \sum_{l=1}^{\hat{N}'(t)} \hat{L}_{1l}\hat{L}_{2l}. \quad (50)$$

Thus, the estimator of the second-order hop covariance matrix is obtained:

$$\hat{\Gamma} = \begin{pmatrix} \hat{\xi}_{11} & \hat{\xi}_{12} \\ \hat{\xi}_{21} & \hat{\xi}_{22} \end{pmatrix}. \quad (51)$$

Now, we consider the convergence rate of the jump covariation difference matrix.

We first make the following assumptions.

(A_1) wavelet (φ, ψ) is differentiable in hop estimation.

$(A_2)\mu_{kt}, \sigma_{kt}^2$ is continuous almost everywhere with respect to t, and $E(\max_{0\leq s\leq 1}\mu_{kd}^2) < +\infty, E(\max_{0\leq s\leq 1}\sigma_{kt}^2) < +\infty, k = 1, 2$.

(A_3) is on the interval [0,1], $\hat{q}_k = \hat{N}_k(t)$.

$$q_k = N_k(t) < +\infty, \tau_{k1} < \tau_{k2} < \cdots < \tau_{\mathrm{kl}} < \cdots, 0 < |L_{k|}| < +\infty, \quad k = 1, 2. \quad (52)$$

$(A_4)(\mu_{\mathrm{kt}}, \sigma_{\mathrm{kt}}^2, W_{\mathrm{kt}})(N_k(t), L_{\mathrm{kl}})$ is independent of ε_{kt}, and $\varepsilon_{\mathrm{kt}}$ ii.$d.$, $E(\varepsilon_{\mathrm{kt}}^4) < +\infty, \quad k = 1, 2$.

The continuous part of $(A_5)X_k$ obeys the process, k=1,2.

We assume that $\delta_n \sim n^{-1/2}$, under the condition of $A_1 \sim A_4$, when $n \longrightarrow \infty$, there are:

$$\hat{\Gamma} - \Gamma = O_p\left(n^{-1/4}\right). \quad (53)$$

To prove the theorem, we need the following two lemmas.

Under the condition of $A_1 \sim A_4$, we have:
$\lim_{n\longrightarrow\infty} P(\hat{q}_k = q_k, \sum_{l=1}^{q_k} |\hat{\tau}_l - \tau_l| \leq n^{-1}\log^2 n|X) = 1, \quad k = 1, 2$ is established.

We assume that $\delta_n \sim n^{1/2}$, under the condition of $A_1 \sim A_4$, when $n \longrightarrow \infty$, we have:
$\Pi_k = \hat{L}_{\mathrm{kd}} - L_{\mathrm{kl}} = O_p(n^{-1/4}), \quad k = 1, 2$ is established.

The proof is as follows:

(1) The elements on the off-diagonal line are first considered:

$$\hat{\xi}_{12} - \xi_{12} = \sum_{l=1}^{\hat{N}'(t)} \hat{L}_{1l}\hat{L}_{21} - \sum_{l=1}^{N'(t)} L_{1l}L_{2l}. \quad (54)$$

For $\forall d > 0$, $\hat{N}'(t) = \hat{q}, N\prime(t) = q$, $\lim_{n\longrightarrow\infty} P(\hat{q} \neq q) = 0$, Therefore, there are:

$$\lim_{d\longrightarrow\infty n\longrightarrow\infty}\lim_{n\longrightarrow\infty} P\left(\left|\hat{\xi}_{12} - \xi_{12}\right| > \mathrm{d}n^{-1/4}\right) \leq \lim_{d\longrightarrow\infty}\lim_{n\longrightarrow\infty} P\left(\hat{q} = q, \left|\hat{\xi}_{12} - \xi_{12}\right| > \mathrm{d}n^{-1/4}\right). \quad (55)$$

When $\hat{q} = q$, then there are:

$$\begin{aligned}\hat{\xi}_{12} - \xi_{12} &= \sum_{l=1}^{q}\left(\hat{L}_{1l}\hat{L}_{2l} - L_{1l}L_{2l}\right) \\ &= \sum_{l=1}^{q}\left[\left(\hat{L}_{1l} - L_{1l}\right)\hat{L}_{2l} + \left(\hat{L}_{2l} - L_{2l}\right)L_{1l}\right] \\ &= \sum_{l=1}^{q}\left(\Pi_1\hat{L}_{2l} + \Pi_2 L_{1l}\right).\end{aligned} \quad (56)$$

Among them, there are:

$$\begin{aligned}\Pi_1 &= \hat{L}_{1l} - L_{1l}, \\ \Pi_2 &= \hat{L}_{2l} - L_{2l}, \\ \hat{L}_{1l} - L_{1l} &= O_p\left(n^{-1/4}\right), \\ \hat{L}_{2l} - L_{2l} &= O_P\left(n^{-1/4}\right),\end{aligned} \quad (57)$$

Moreover, because $\hat{L}_{2l} = O_p(1), L_{1l} = O_p(1)$,, so, there are:

$$\begin{aligned}\Pi_1\hat{L}_{2l} &= O_p\left(n^{1/4}\right), \\ \Pi_2 L_{1l} &= O_p\left(n^{-1/4}\right),\end{aligned} \quad (58)$$

The proof is as follows:

$$\sum_{l=1}^{q}\left|\Pi_1\widehat{L}_{2l}\right| = O_p\left(n^{1/4}\right), \sum_{l=1}^{q}\left|\Pi_2 L_{1l}\right| = O_p\left(n^{1/4}\right)$$

$$P\left(\sum_{l=1}^{q}\left|\Pi_1\widehat{L}_{2l}\right| > d_1 n^{1/4}\right) \leq P(N > m) + P\left(\sum_{l=1}^{m}\left|\Pi_1\widehat{L}_{2l}\right| > d_1 n^{1/4}, N \leq m\right)$$
$$\leq P(N > m) + P\left(\sum_{l=1}^{m}\left|\Pi_1\widehat{L}_{21}\right| > d_1 n^{1/4}\right) \leq P(N > m) + \sum_{l=1}^{m} p\left(\left|\Pi_1\widehat{L}_{21} > \frac{d_1}{m} n^{-1/4}\right|\right), \tag{59}$$

m is fixed, and because the following formula exists:

$$\Pi_1\widehat{L}_{2l} = O_p\left(n^{-1/4}\right). \tag{60}$$

Therefore, there are:

$$\lim_{d_1 \longrightarrow \infty}\lim_{n\longrightarrow\infty} P\left(\sum_{l=1}^{q}\left|\Pi_1\widehat{L}_{2l}\right| > d_1 n^{-1/4}\right) \leq P(N > m). \tag{61}$$

When $m \longrightarrow \infty$, there are:

$$\lim_{d_1 \longrightarrow \infty}\lim_{n\longrightarrow\infty} P\left(\sum_{l=1}^{q}\left|\Pi_1\widehat{L}_{2l}\right| > d_1 n^{-1/4}\right) \leq \lim_{m\longrightarrow\infty} P(N > m) = 0, \tag{62}$$

Thus, there are:

$$\sum_{l=1}^{q}\left|\Pi_1\widehat{L}_{2l}\right| = O_p\left(n^{1/4}\right). \tag{63}$$

Similarly, it can be proved that $\sum_{l=1}^{q}\left|\Pi_2 L_{ll}\right| = O_p\left(n^{-1/4}\right)$., as follows:

$$P\left(\widehat{q} = q, \left|\widehat{\xi}_{12} - \xi_{12}\right| > \mathrm{d}n^{-1/4}\right) \leq P\left(\sum_{l=1}^{q}\left\{\left|\Pi_1\widehat{L}_{2l}\right| + \left|\Pi_2 L_{11}\right|\right\} > \mathrm{d}n^{-1/4}\right)$$
$$\leq P\left(\sum_{l=1}^{q}\left|\Pi_1\widehat{L}_{2l}\right| > \mathrm{d}n^{-1/4}/2\right) + P\left(\sum_{l=1}^{q}\left|\Pi_2 L_{1l}\right| > \mathrm{d}n^{-1/4}/2\right) \longrightarrow 0. \tag{64}$$

Thus, there are:

$$\widehat{\xi}_{12} - \xi_{12} = O_p\left(n^{-1/4}\right). \tag{65}$$

(2) The elements on the diagonal are considered again

$$\widehat{\xi}_{\mathrm{kt}} - \xi_{\mathrm{kk}} = \sum_{l=1}^{\widehat{q}_k}\widehat{L}_{\mathrm{kl}}^{\ 2} - \sum_{l=1}^{q_k} L_{\mathrm{kl}}^{\ 2}, \quad k = 1, 2, \tag{66}$$

For $\forall d > 0$, there are:

$$P\left(\left|\widehat{\xi}_{\mathrm{kk}} - \xi_{\mathrm{kk}}\right| > \mathrm{d}n^{1/4}\right) \leq P\left(\widehat{q}_k = q_k, \left|\widehat{\xi}_{\mathrm{kk}} - \xi_{\mathrm{kk}}\right| > \mathrm{d}n^{-1/4}\right) + P\left(\widehat{q}_k \neq q_k\right),$$
$$\lim_{n\longrightarrow\infty} P\left(\widehat{q}_k \neq q_k\right) = 0, \tag{67}$$

Thus, there are:

$$\lim_{d\longrightarrow\infty n\longrightarrow\infty} P\left(\left|\widehat{\xi}_{\mathrm{kk}} - \xi_{\mathrm{kk}}\right| > \mathrm{d}n^{-1/4}\right) \leq \lim_{d\longrightarrow\infty}\lim_{n\longrightarrow\infty} P\left(\widehat{q}_k = q_k, \left|\widehat{\xi}_{\mathrm{kk}} - \xi_{\mathrm{kk}}\right| > \mathrm{d}n^{1/4}\right). \tag{68}$$

When $\widehat{q}_k = q_k$, then there are:

$$\widehat{\xi}_{\mathrm{kk}} - \xi_{\mathrm{kk}} = \sum_{l=1}^{q_k}\left(\widehat{L}_{\mathrm{kl}}^2 - L_{\mathrm{kl}}^2\right) = \sum_{l=1}^{q_k}\left(\Pi_{kl}^2 + 2\Pi_k L_{kl}\right),$$
$$\Pi_k = O_p\left(n^{-1/4}\right). \tag{69}$$

At the same time, it is easy to know that $L_{kl} = O_P(1)$. Therefore, there is $\Pi_k L_{kl} = O_p\left(n^{1/4}\right), \Pi^2{}_k = O_p\left(n^{-1/2}\right)$. The proof of formula 1 is easy to obtain:

$$\widehat{\xi}_{\mathrm{kk}} - \xi_{\mathrm{kk}} = O_P\left(n^{-1/4}\right), \quad k = 1, 2, \tag{70}$$

Combining 1 and 2, we can get:

$$\widehat{\Gamma}-\Gamma=O_p\left(n^{-1/4}\right). \quad (71)$$

We assume that in the estimation of the jump process, the wavelet method is used to obtain the estimators $\widehat{\tau}_{\text{kd}}$ and $\widehat{L}_{\text{kd}}$ of all jump points and jump heights of X_k, respectively. Therefore, the estimators of the jump parts of the counting process $N_k(t)$ and X_{kt} are respectively $\widehat{N}_k(t)=\sum_{l=1}^{\widehat{q}_k} I(\widehat{\tau}_{kl}\le t), \widehat{X}_{\text{kt}}^d=\sum_{l=1}^{\widehat{N}_k(t)}\widehat{L}_{\text{kt}}=\sum_{\widehat{t}_k\le t}\widehat{L}_{\text{kl}}$., where $I(\cdot)$ is an indicative function, k=1,2.

In order to eliminate the impact of the jumping process on the data, we can do the following: $Y_{\text{kt}_1}^*=Y_{\text{kk}_t}-\widehat{X}_{\text{kt}_i}^d=Y_{k_t}-\sum_{t_t\le t_i}\widehat{L}_{\text{kl}}, 4k=1,2$.

$$\begin{aligned}
\left[Y_k^*,Y_k^*\right]^{(M)}&=\frac{1}{M}\sum_{m=1}^{M}\sum_{j=1}^{n/M}\left(Y_{k_{m+\mu}}^*-Y_{k_{m+(-1)}^*}\right)^2=\frac{1}{M}\sum_{i=1}^{n-M}\left(Y_{k_{i+\mu}}^*-Y_{k_i}^*\right)^2, k=1,2\\
\left[Y_1^*,Y_2^*\right]^{(M)}&=\left[Y_2^*,Y_1^*\right]^{(M)}\\
&=\frac{1}{M}\sum_{m=1}^{M}\sum_{j=1}^{n/M}\left(Y_{1t_{m+\mu}}^*-Y_{1t_{m+(j-1)\mu}^*}^*\right)\left(Y_{2t_{m+\mu}}^*-Y_{2t_{m+((-1)\mu}^*}^*\right)=\frac{1}{M}\sum_{i=1}^{n-M}\left(Y_{1t_{i+\mu}}^*-Y_{1t_i}^*\right)\left(Y_{2t_{i+N}}^*-Y_{2t_i}^*\right).
\end{aligned} \quad (72)$$

Moreover, there is $\theta^*=\left(\begin{bmatrix}Y_1^*,Y_1^*\\Y_2^*,Y_1^*\end{bmatrix}^{(M)}\begin{bmatrix}(M)\\Y_1^*,Y_2^*\\Y_2^*,Y_2^*\end{bmatrix}^{(M)}\right)$.

For comparison, denoting the X_k and Y_k continuous parts with X_k^c and Y_k^c, there are:

$$X_{kt}^c=\int_0\mu_{\text{ks}}\text{d}s+\int_b\sigma_{\text{ks}}\text{d}W_{\text{ks}}, Y_{\text{kt}}^c=X_{\text{ka}}^c+\varepsilon_{\text{kt}}, \quad k=1,2. \quad (73)$$

The following formula is defined:

$$\begin{aligned}
\left[Y_k^c,Y_k^c\right]^{(M)}&=\frac{1}{M}\sum_{m=1}^{M}\sum_{j=1}^{n/M}\left(Y_{\text{kk}_{m+M}}^c-Y_{k_{m+(-1)}}^c\right)^2=\frac{1}{M}\sum_{i=1}^{n-M}\left(Y_{\text{kk}_{i+\mu}}^c-Y_{k_1}^c\right)^2, \quad k=1,2,\\
\left[Y_1^c,Y_2^c\right]^{(M)}&=\left[Y_2^c,Y_1^c\right]^{(M)}\\
&=\frac{1}{M}\sum_{m=1}^{M}\sum_{j=1}^{n/M}\left(Y_{1t_{++M}}^c-Y_{1t_{t+1-1\mu}}^c\right)\left(Y_{2t_{++\mu}}^c-Y_{2t_{i+c}(-1)M}^c\right)=\frac{1}{M}\sum_{i=1}^{n-M}\left(Y_{1t_{i+\mu}}^c-Y_{1t_i}^c\right)\left(Y_{2t_{i+\mu}}^c-Y_{2t_i}^c\right).
\end{aligned} \quad (74)$$

Moreover, there is $\theta^c=\left(\begin{bmatrix}Y_1^c,Y_1^c\\Y_2^c,Y_1^c\end{bmatrix}^{(M)}\begin{bmatrix}Y_1^c,Y_2^c\\Y_2^c,Y_2^c\end{bmatrix}^{(M)}\right)$.

It is worth noting that Y_k^c is not directly observable, it is only used as a theoretical comparison of Y_k^*, k=1,2.

Now, we generalize the dual-scale realized volatility to the two-asset jump-diffusion price model to obtain an estimator of θ as follows:

$$\widehat{\theta}_M=\begin{pmatrix}\widehat{\theta}_{11M} & \widehat{\theta}_{12M}\\ \widehat{\theta}_{21M} & \widehat{\theta}_{22M}\end{pmatrix}. \quad (75)$$

Among them, there are:

$$\begin{aligned}
\widehat{\theta}_{kkM}&=\left[Y_k^*,Y_k^*\right]^{(M)}-\frac{1}{M}\left[Y_k^*,Y_k^*\right]^{(1)}, k=1,2;\\
\widehat{\theta}_{12M}&=\widehat{\theta}_{21M}=\left[Y_1^*,Y_2^*\right]^{(M)}-\frac{1}{M}\left[Y_1^*,Y_2^*\right]^{(1)},
\end{aligned} \quad (76)$$

In the continuous diffusion price model, the best estimator of the integrated volatility θ is obtained by averaging the realized volatility under different optimal sampling data. However, under the jump-diffusion price model, if the estimates of jump points and jump heights are very accurate, the effect of the jump process on the average realized volatility will be asymptotically negligible. To prove it, some notations are now introduced. M is an integer, and the entire sample is sparse to obtain M sub-samples, which are defined as follows:

We assume that $M/n+\log^2 n/M\longrightarrow 0$, under the condition of $A_1\sim A_4$, $\theta^*-\theta^c=O_p(n^{-1/4}+M^{-1}\log^2 n)$ holds.

We assume that $M=O(n^{2/3})(c>0)$, under the condition of $A_1\sim A_5$, $\widehat{\theta}_M-\theta=O_p(n^{-1/6})$ is established.

Under the condition of $A_1\sim A_4$, $\lim_{n\longrightarrow\infty} P(\widehat{q}_k=q_k,\sum_{l=1}^{q_k}|\widehat{\tau}_{kt}-\tau_{kt}|\le n^{-1}\log^2 n|X)=1, \quad k=1,2.$ holds.

$$E\left[\max\left|X_t^c\right|\right] \le 2\int_0^1 E\left(\mu_s^2\right)\mathrm{d}s + 8E\left[\left|\int_0^1 \sigma_s \mathrm{d}W_s\right|^2\right] = \int_0^1 E\left(2\mu_s^2 + 8\sigma_s^2\right)\mathrm{d}s. \tag{77}$$

We assume that $X_n = O_p(a_n), Y_n = O_p(b_n)$, and a_n, b_n are greater than 0, n is a positive integer, there is $X_n + Y_n = O_p(a_n + b_n)$.

The proof is as follows: The following formula holds, as follows:

$$\left|\frac{X_n + Y_n}{a_n + b_n}\right| = \left|\frac{X_n}{a_n + b_n} + \frac{Y_n}{a_n + b_n}\right| \le \left|\frac{X_n}{a_n + b_n}\right| + \left|\frac{Y_n}{a_n + b_n}\right| < \left|\frac{X_n}{a_n}\right| + \left|\frac{Y_n}{b_n}\right|. \tag{78}$$

It is also known that the following formula exists:

$$\lim_{A\longrightarrow\infty} P\left(\left|\frac{X_n + Y_n}{a_n + b_n}\right| \ge A\right) < \lim_{A\longrightarrow\infty} P\left(\left|\frac{X_n}{a_n}\right| + \left|\frac{Y_n}{b_n}\right| \ge A\right) \le \lim_{A\longrightarrow\infty} P\left(\left|\frac{X_n}{a_n}\right| \ge \frac{A}{2}\right) + \lim_{A\longrightarrow\infty} P\left(\left|\left|\frac{Y_n}{b_n}\right|\right| \ge \frac{A}{2}\right) = 0. \tag{79}$$

Thus, there are:

$$X_n + Y_n = O_p(a_n + b_n). \tag{80}$$

The proof is as follows:

We assume that X_k has $q_k = N_k$ hops in τ_{kl}, which are denoted as the hop part of X_k, as follows:

$$X_{kt}^d = \sum_{l=1}^{N_k(t)} L_{\mathrm{kl}} = \sum_{T_t\ A3/4t} L_{\mathrm{kl}}, \quad k = 1, 2, \tag{81}$$

Because there is the following formula:

$$Y_{k_1} = X_{k_1} + \varepsilon_{k_1} = X_{k_1}^c + X_{k_1}^d + \varepsilon_{k_1} = Y_{k_1}^c + X_{k_1}^d$$
$$Y_{k_i}^{\bullet} = Y_{k_1} - \widehat{X}_{k_1}^d = Y_{k_1}^c + X_{k_1}^d - \widehat{X}_{k_1}^d. \tag{82}$$

Thus, there are:

$$Y_{\mathrm{kk}_{l+M}}^* - Y_{\mathrm{kt}_1}^* = Y_{\mathrm{kl}_{1+M}}^c - Y_{kt_1}^c + \sum_{t_1<\tau_k\le t_{l+\mu}} L_{\mathrm{kd}} - \sum_{t_i<i_{\mathrm{kl}}\le t_{i+M}} \widehat{L}_{\mathrm{kl}} \equiv Y_{\mathrm{kt}_{i+M}}^c - Y_{\mathrm{kt}_l}^c + \xi_{\mathrm{ki}}. \tag{83}$$

It is equivalent to:

$$\Omega_n = \left\{\min\left\{\tau_{kl} - \tau_{k(l-1)}, \quad l = 1, \ldots q_k\right\} > M/n\right\} \cap \left\{\left|\widehat{\tau}_{\mathrm{kl}} - \tau_{\mathrm{kl}}\right| \le n^{-1}\log^2 n, \quad l = 1, \ldots q_k = \widehat{q}_k\right\}. \tag{84}$$

From the condition (A_3), it can be known that:

$$\lim_{n\longrightarrow\infty} P(\Omega_n) = 1. \tag{85}$$

Therefore, it is only necessary to prove that $\theta^* - \theta^c = O_p(n^{-1/4} + M^{-1}\log^2 n)$ holds in Ω_n.

It is worth noting that in Ω_n, X_{kt} has at most one jump in the interval $[t_i, t_{i+M}]$ of length M/n, and since $M/n(n^{-1}\log^2 n) = M/\log^2 n \longrightarrow \infty$, the distance between τ_{kl} and $\widehat{\tau}_{kl}$ is very small. Therefore, there are the following three cases (k=1,2):

(i) $t_t, t_{l+M} < \tau_{kl}, \widehat{\tau}_{kl}$ or $t_t, t_{t+M} > \tau_{kl}, \widehat{\tau}_{kl}, a_{.s}$ is established;
(ii) $t_i < \tau_{kl}, \widehat{\tau}_{kl} < t_{i+M}$, is established, at this time, $n(M/n - |\tau_{kl} - \widehat{\tau}_{k|}|) \le M$;
(iii) We might as well set $\tau_{kl} \le \widehat{\tau}_{kl}, t_l \in (\tau_{kl}, \widehat{\tau}_{kl})$ or $t_{l+k} \in (\tau_{kl}, \widehat{\tau}_{kl})$, is established;

At this time, $2n|\tau_{kl} - \widehat{\tau}_{kl}| \le 2\log^2 n$.

In the above three cases, $\xi_{kt} = 0; L_{kt} - \widehat{L}_{k|}; L_{k|}$ or $\widehat{L}_{k\,d}$ holds.

Correspondingly, there are:

$$\left(Y_{1_{t+\mu}^*}^* - Y_{1t_t}^*\right)\left(Y_{2t_{t+\mu}}^* - Y_{2t_i}^*\right) - \left(Y_{1t_{t+\alpha}}^c - Y_{1t_t}^c\right)\left(Y_{2t_{t+\alpha}}^c - Y_{2t_i}^c\right)$$
$$= 0; O_p\left(\left|L_{1l} - \widehat{L}_{1l}\right|\right) + O_p\left(\left|L_{2l} - \widehat{L}_{2l}\right|\right)O_p\left(\left|L_{11}\right| + \left|\widehat{L}_{1l}\right|\right) + O_p\left(\left|L_{2l}\right| + \left|\widehat{L}_{2l}\right|\right)$$
$$\left(Y_{k_{t+\mu}}^* - Y_{k_t}^*\right)^2 - \left(Y_{k_{t+\mu}}^c - Y_{k_l}^c\right)^2 = 0; O_p\left(\left|L_k - \widehat{L}_{k|}\right|\right)O_p\left(\left|L_k\right| + \left|\widehat{L}_{k|}\right|\right). \tag{86}$$

(1) Off-diagonal elements are considered first:

$$[Y_1^*, Y_2^*]^{(M)} - [Y_1^c, Y_2^c]^{(M)} = \frac{1}{M}\sum_{i=1}^{n-M}\Big[\Big(Y^*_{i_t+\alpha} - Y^*_{1t_i}\Big)\Big(Y^*_{2t^*_{++\mu}} - Y^*_{2t_t}\Big) - \Big(Y^c_{1t_{t+\mu}} - Y^c_{1t_i}\Big)\Big(Y^c_{2t_{++\mu}} - Y^c_{2t_i}\Big)\Big]. \quad (87)$$

From the above discussion, it can be seen that the following formula holds:

$$[Y_1^*, Y_2^*]^{(M)} - [Y_1^c, Y_2^c]^{(M)} = O_p\big(H_{n,M}\big), \quad (88)$$

Among them, there are:

$$H_{n,M} = \sum_{l=1}^{q_1}\big|L_{1l} - \widehat{L}_{1l}\big| + \sum_{l=1}^{q_2}\big|L_{2l} - \widehat{L}_{2l}\big| + 2M^{-1}\log^2 n\sum_{l=1}^{q_1}\big(|L_{1l}| + |\widehat{L}_{1l}|\big) + 2M^{-1}\log^2 n\sum_{l=1}^{q_2}\big(|L_{2l}| + |\widehat{L}_{2l}|\big). \quad (89)$$

From the condition (A_3), it can be known that:

$$\sum_{l=1}^{q_k}\big|L_{\mathrm{kd}} - \widehat{L}_{\mathrm{kl}}\big| = O_p\big(n^{-1/4}\big), \sum_{l=1}^{q_l}\big(|L_{\mathrm{kl}}| + |\widehat{L}_{\mathrm{kl}}|\big) = O_p(1),$$
$$H_{n,M} = O_p\big(n^{-1/4} + M^{-1}\log^2 n\big). \quad (90)$$

Thus, there are:

$$[Y_1^*, Y_2^*]^{(M)} - [Y_1^c, Y_2^c]^{(M)} = O_p\big(n^{-1/4} + M^{-1}\log^2 n\big). \quad (91)$$

(2) The elements on the diagonal are now considered

$$[Y_k^*, Y_k^*]^{(M)} - [Y_k^c, Y_k^c]^{(M)} = \frac{1}{M}\sum_{i=1}^{n-M}\Big[\Big(Y_{k^*_{k+\alpha}} - Y^*_{k_i}\Big)^2 - \Big(Y^c_{k_{k+\mu}} - Y^c_{k_i}\Big)^2\Big], \quad k = 1, 2. \quad (92)$$

Likewise, we get:

$$\Big[Y_k^*, Y_k^{P_k^*}\Big]^{(M)} - [Y_k^c, Y_k^c]^{(M)} = O_p\big(J_{n,M}\big). \quad (93)$$

Among them, there are:

$$J_{n,\mu} = \sum_{i=1}^{q_k}\big|L_{kl} - \widehat{L}_k\big| + 2M^{-1}\log^2 n\sum_{t=1}^{q_k}\big(|L_k| + |\widehat{L}_{kl}|\big), \quad (94)$$

Because there is the following formula:

$$\sum_{l=1}^{q_k}\big|L_{\mathrm{kl}} - \widehat{L}_{\mathrm{kl}}\big| = O_p\big(n^{-1/4}\big),$$
$$\sum_{l=1}^{q_k}\big(|L_{\mathrm{kl}}| + |\widehat{L}_{\mathrm{kl}}|\big) = O_p(1), \quad (95)$$

Therefore, there are:

$$J_{n,M} = O_p\big(n^{-1/4} + M^{-1}\log^2 n\big). \quad (96)$$

Thus, there are:

$$[Y_k^*, Y_k^*]^{(M)} - [Y_k^c, Y_k^c]^{(M)} = O_p\big(n^{1/4} + M^{-1}\log^2 n\big). \quad (97)$$

Combining formulas (1) and (2), we can get:

$$\theta^* - \theta^c = O_p\big(n^{-1/4} + M^{-1}\log^2 n\big). \quad (98)$$

The proof is as follows:

$\widehat{\theta}^c_M = \begin{pmatrix} \widehat{\theta}^c_{11M} & \widehat{\theta}^c_{12M} \\ \widehat{\theta}^c_{21M} & \widehat{\theta}^c_{22M} \end{pmatrix}$ is the two-scale realized volatility estimator of θ under the continuous diffusion price model.

Among them, there are:

$$\widehat{\theta}^c_{kkM} = [Y_k^c, Y_k^c]^{(M)} - \frac{1}{M}[Y_k^c, Y_k^c]^{(1)}, k = 1, 2$$
$$\widehat{\theta}^c_{12M} = \widehat{\theta}^c_{21M} = [Y_1^c, Y_2^c](M) - \frac{1}{M}[Y_1^c, Y_2^c]^{(1)} \quad (99)$$
$$\widehat{\theta}^c_M - \theta = O_p\big(n^{-1/6}\big).$$

Therefore, it is now only necessary to prove that $\widehat{\theta}_M - \widehat{\theta}^c_M = O_p(n^{-1/6})$.

(1) The elements on the off-diagonal line are considered first

$$\widehat{\theta}^*_{12M} - \widehat{\theta}^c_{12M} = [Y_1^*, Y_2^*]^{(M)} - [Y_1^c, Y_2^c]^{(M)} + \frac{1}{M}\Big\{[Y_1^*, Y_2^*]^{(1)} - [Y_1^c, Y_2^c]^{(1)}\Big\}. \quad (100)$$

When $M = O(n^{2/3})$, then there are:

$$[Y_1^*, Y_2^*]^{(M)} - [Y_1^c, Y_2^c]^{(M)} = O_p\left(n^{-1/6}\right), \tag{101}$$

The following evidence shows that when $M = O(n^{2/3})$, there are:

$$\begin{aligned}&[Y_1^*, Y_2^*]^{(1)} - [Y_1^c, Y_2^c]^{(1)} = O_p\left(n^{1/4}\log n\right)\\ &Y_{k_{t+1}}^* - Y_{k_1}^* = Y_{k_{t+1}}^c - Y_{k_1}^c + \sum_{t_1<\tau_{kt}\le t_{t+1}} L_{\mathrm{kl}} - \sum_{t_1<\widehat{\tau}_{kl}\le s_{l+1}} \widehat{L}_{\mathrm{kd}} \equiv Y_{k_{l+1}}^c - Y_{k_i}^c + \zeta_{\mathrm{ki}}, \quad i = 1, \ldots n-1.\end{aligned} \tag{102}$$

In these (n-1) intervals $[t_i, t_{i+1}]$, at least $n-1-(\widehat{q}_k + q_k)$ does not contain $\tau_{kl}, \widehat{\tau}_{kl}$, at this time, $\zeta_{ki} = 0$. There are at most $(q_k + \widehat{q}_k)$ non-zero ζ_{ki} s, and $\zeta_{\mathrm{kd}} = L_{\mathrm{kl}} - \widehat{L}_{\mathrm{kl}}$; L_{kl} or $\widehat{L}_{\mathrm{kl}}$ at this time.

Therefore, there are:

$$\begin{aligned}\left|[Y_1^*, Y_2^*]^{(1)} - [Y_1^c, Y_2^c]^{(1)}\right| \le{}& \sum_{l=1}^{1} L_{1l}^2 + \sum_{l=1}^{q_1} \widehat{L}_{1\mathrm{l}}^2 + \sum_{l=1}^{q_2} L_{2l}^2 + \sum_{l=1}^{\widehat{q}_2} \widehat{L}_{2l}^2 + 2\max_{1\le i\le n}\left|Y_{2t_i}^c\right|\left\{\sum_{l=1}^{q_l}\left|L_{1l}\right| + \sum_{l=1}^{\widehat{q}_l}\left|\widehat{L}_{1l}\right|\right\}\\ &+ 2\max_{1\le i\le n}\left|Y_{1t_i}^c\right|\left\{\sum_{l=1}^{q_2}\left|L_{2l}\right| + \sum_{l=1}^{\widehat{q}_2}\left|\widehat{L}_{2l}\right|\right\}.\end{aligned} \tag{103}$$

First, there is the following formula:

$$\begin{aligned}&\sum_{l=1}^{q_k}\left|L_{kk}\right| = O_p(1), \sum_{l=1}^{q_k}\left|\widehat{L}_{\mathrm{kl}}\right| = O_p(1), \sum_{l=1}^{q_k} L_{\mathrm{kl}}^2\\ &= O_p(1), \sum_{l=1}^{q_k}\widehat{L}_{\mathrm{kl}}^2 = O_p(1).\end{aligned} \tag{104}$$

Second, there is the following formula:

$$\begin{aligned}\max_{1\le\le\le\le n}\left|Y_{k_i}^c\right| &\le \max_{1\le i\le 1\le n}\left\{X_{k_i}^c\right| + \left|\varepsilon_{k_i}\right|\}\\ &\le \max_{1\le\le\le\le n}\left|X_{ki}^c\right| + \max_{1\le\le\le n}\left|\varepsilon_{k_i}\right|.\end{aligned} \tag{105}$$

From the condition (A_2), it can be deduced that:

$$\max_{1\le\le\le 1}\left|X_{kt}^c\right| = O_p(1). \tag{106}$$

In addition, there is the following formula:

$$\begin{aligned}P\left(\max_{1\le i\le n}\left|\varepsilon_{k_1}\right| > \varepsilon^{1/4}\log n\right) &= 1 - \left[1 - P\left(\left|\varepsilon_{k_1}\right| > g^{1/4}\log n\right)\right]^n\\ &\le 1 - \left[1 - \frac{E\left(\varepsilon_{kk_1}^2\right)}{\varepsilon^2 n^{1/2}\log^2 n}\right]^n \longrightarrow 0\,(n \longrightarrow \infty).\end{aligned} \tag{107}$$

Therefore, there are:

$$\max_{1\le 1\le n}\left|\varepsilon_{k_i}\right| = O_p\left(n^{1/4}\log n\right). \tag{108}$$

In addition, there are:

$$[Y_1^*, Y_2^*]^{(1)} - [Y_1^c, Y_2^c]^{(1)} = O_p\left(n^{1/4}\log n\right). \tag{109}$$

When the above formula is established, there are:

$$\begin{aligned}&M = O\left(n^{2/3}\right)O\left(n^{2/3}\right)\hat{\mathrm{E}}\,\pm, \frac{1}{\mathrm{M}}\left\{\mathrm{Y}_1^*, \mathrm{Y}_2^*\right]^{(1)} - \left[\mathrm{Y}_1^c, \mathrm{Y}_2^c\right]^{(1)}\right\} = O_p\left(n^{-Y_6}\right)\\ &\widehat{\theta}_{12M}^* - \widehat{\theta}_{12M}^c = [Y_1^*, Y_2^*]^{(M)} - [Y_1^c, Y_2^c]^{(M)} + \frac{1}{M}\left\{Y_1^*, Y_2^*\right]^{(1)} - [Y_1^c, Y_2^c]^{(1)}\right\} = O_p\left(n^{-1/6}\right).\end{aligned} \tag{110}$$

(2) The elements on the diagonal are now considered

$$\widehat{\theta}^{*}_{kkM} - \widehat{\theta}^{c}_{kcM} = \left[Y^{*}_{k}, Y^{*}_{k}\right]^{(M)} - \left[Y^{c}_{k}, Y^{c}_{k}\right]^{(M)} + \frac{1}{M}\left\{\{\left[Y^{*}_{k}, Y^{*}_{k}\right]^{(1)} - \left[Y^{c}_{k}, Y^{c}_{k}\right]^{(1)}\right\}. \quad (111)$$

Similarly, it can be obtained that when $M = O(n^{2/3})$, there are:

$$\begin{aligned}&\left[Y^{*}_{k}, Y^{*}_{k}\right]^{(M)} - \left[Y^{c}_{k}, Y^{c}_{k}\right]^{(M)} = O_{p}\left(n^{-1/6}\right);\\ &\left|\left[Y^{*}_{k}, Y^{*}_{k}\right]^{\eta_i)} - \left[Y^{c}_{k}, Y^{c}_{k}\right]^{(1)}\right| \le 2\sum_{l=1}^{q_k} L^{2}_{kl} + 2\sum_{l=1}^{q_k} \widehat{L}^{2}_{\mathrm{kl}} + 4\max_{1\le\le\le n}\left|Y^{c}_{k_i}\right|\left\{\sum_{k=1}^{q_k}\left|L_k\right| + \sum_{i=1}^{q_k}\left|\widehat{L}_{\mathrm{kl}}\right|\right\}.\end{aligned} \quad (112)$$

Therefore, there are:

$$\frac{1}{M}\left\{Y^{*}_{k}, Y^{*}_{k}\right]^{(1)} - \left[Y^{c}_{k}, Y^{c}_{k}\right]^{(1)}\right\} = O_{p}\left(n^{-1/6}\right). \quad (113)$$

Thus, there are:

$$\widehat{\theta}^{*}_{\mathrm{kkM}} - \widehat{\theta}^{c}_{\mathrm{kkM}} = \left[Y^{*}_{k}, Y^{*}_{k}\right]^{(M)} - \left[Y^{c}_{k}, Y^{c}_{k}\right]^{(M)} + \frac{1}{M}\left\{Y^{*}_{k}, Y^{*}_{k}\right]^{(1)} - \left[Y^{c}_{k}, Y^{c}_{k}\right]^{(1)}\right\} = O_{p}\left(n^{-1/6}\right). \quad (114)$$

Combining 1 and 2, we can get $\widehat{\theta}_M - \widehat{\theta}^{c}_{M} = O_p(n^{-1/6})$, that is, $\widehat{\theta}_M - \theta = O_p(n^{-1/6})$.

4. Economic fluctuation analysis based on high frequency time series

Changes in financial structure have enriched financial products. The pricing of financial products is different from that of ordinary commodities, and the pricing of financial products depends to a large extent on future expectations. Due to the complexity of the economic system, people cannot make accurate expectations. When people's expectations cannot be realized, it will lead to chaos or even collapse of the financial system. Moreover, when there is a problem in the financial system, it will have a profound impact on the real economy. The irrational expectations of investors will lead to violent fluctuations in the prices of financial assets, and the violent fluctuations in the prices of financial assets will be accompanied by a sudden change in the flow of a large number of funds, which will destroy the stable value relationship of the economic system and affect the continuous production capacity of the real economy. bad influence. In addition, asset price movements are influenced by speculative activities, supply and demand, and political factors. Figure 1 shows the transmission diagram of the mechanism of financial structure changes affecting economic fluctuations.

According to the "income-consumption effect", as more investment brings more jobs, more and more people regain their jobs, and income increases and consumption increases with it. Consumption is the driving force and purpose of production. When companies predict an increase in spending power, they will further expand production. As production resumes and expands, companies hire more workers. In this cycle, the effective social demand increases from AD1 to AD2, the price level rises from P1 to P0, and the expected profit increases. At the same time, the actual output returns to the potential output level Y0, and the economic growth rate gradually returns to the potential economic growth rate, thus slowing down economic fluctuations, as shown in Figure 2.

When the economy is booming, actual output (Y1) is higher than potential output (Y0), and the labor market is in short supply. In addition to the influence of factors such as rising living costs, the price of labor factors continues to rise, and the price of capital is underestimated, so the price of labor relative to capital is higher. According to the "price effect", in order to reduce the unit product cost, the enterprise will increase the degree of technological progress biased towards labor and reduce the degree of bias towards capital, as shown in Figure 3.

The relationship between the fluctuation of technological progress factor bias and the economic fluctuation is shown in Figure 4. It can be seen from the figure that the two are negatively correlated.

With a moderate degree of financial leverage, financial innovation can help companies transfer financial risks, reduce transaction costs, and improve the liquidity of financial resources. This is conducive to improving the efficiency of resource allocation, thereby promoting economic

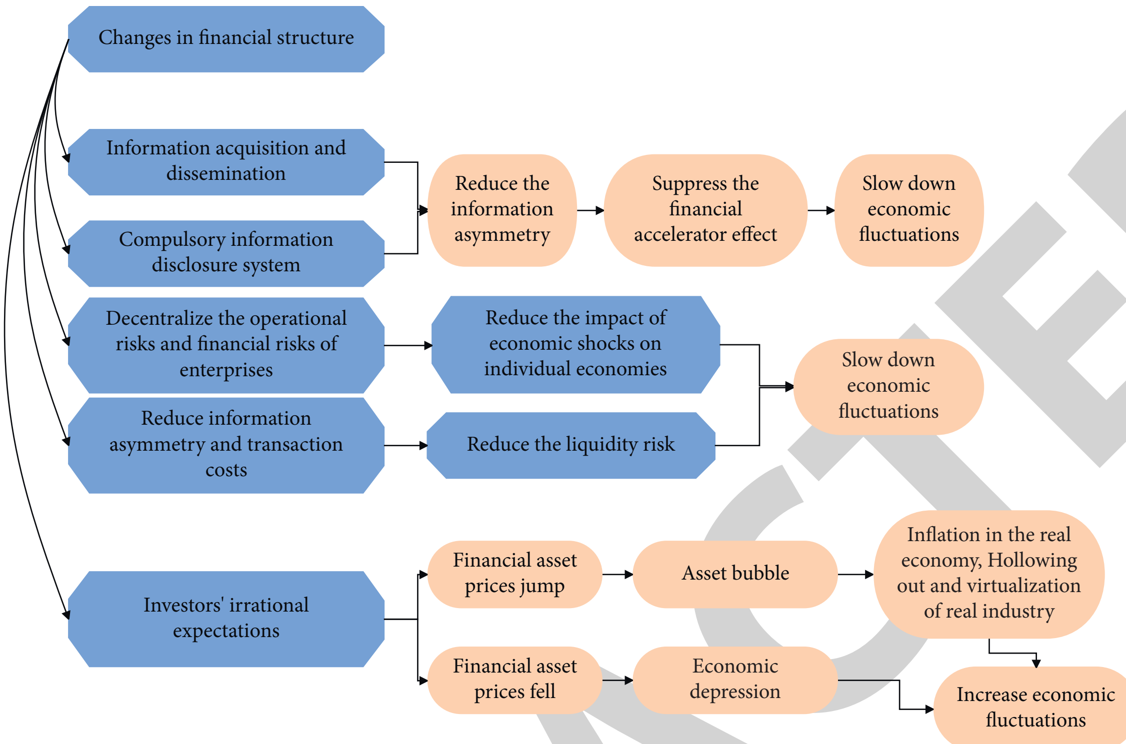

Figure 1: Transmission diagram of the mechanism of financial structure changes affecting economic fluctuations.

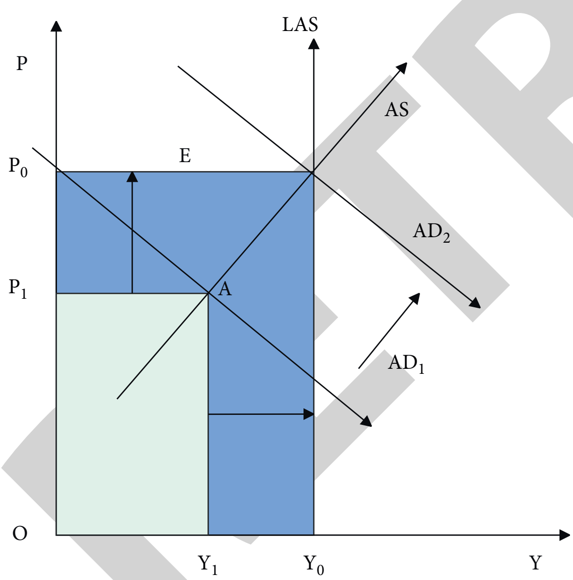

Figure 2: Biased capital during depression.

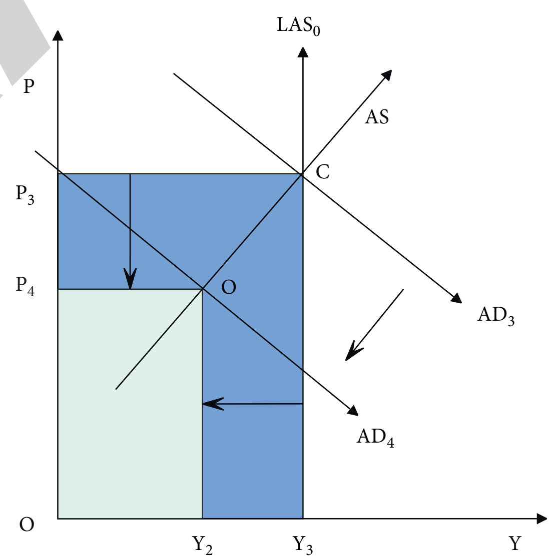

Figure 3: Biased labor in boom times.

development. Figure 5 shows the transmission mechanism model of financial leverage affecting economic fluctuations.

Figure 6 shows the RNN structure after unrolling in time series. From the perspective of time series expansion, each node in the figure from left to right represents the current layer of the recurrent neural network at each moment. From the perspective of the information processing structure of the network, the recurrent neural network can be roughly divided into three layers from top to bottom. The input layer represents the input state at the current moment, and it continues to progress as time unfolds. The hidden layer belongs to the memory structure, and the neurons in the hidden layer are also called memory units, which can store the information of the hidden layer at the previous moment. The output layer is the output state at the current moment. Figure 7 is the structure diagram of the LSTM deep network model.

The basic structure of the combined model is shown in Figure 8 below. First, it collects the original data of the

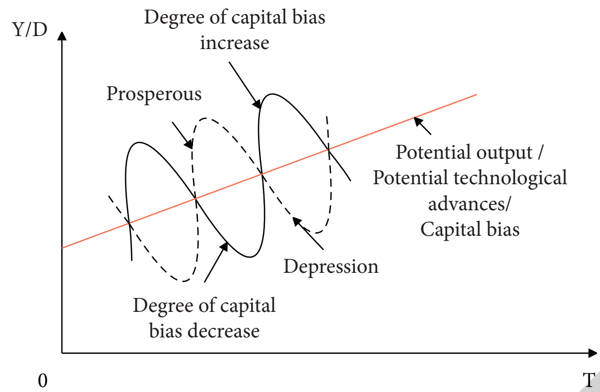

FIGURE 4: The relationship between the fluctuation of the degree of technological progress favoring capital and the fluctuation of the economy.

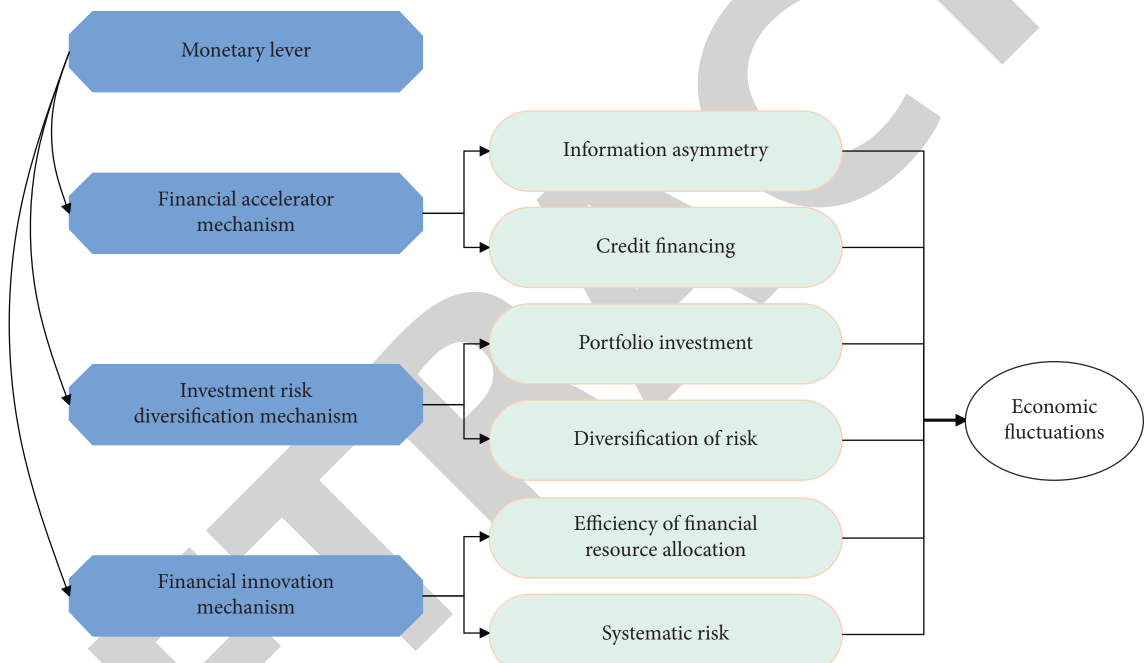

FIGURE 5: The transmission mechanism model of financial leverage affecting economic fluctuations.

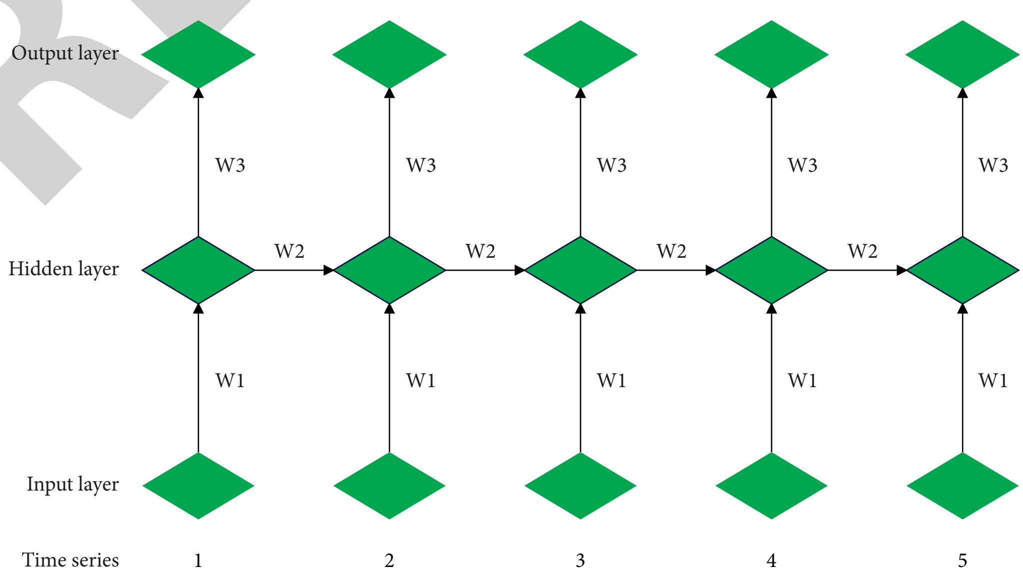

FIGURE 6: RNN network structure diagram after time series expansion.

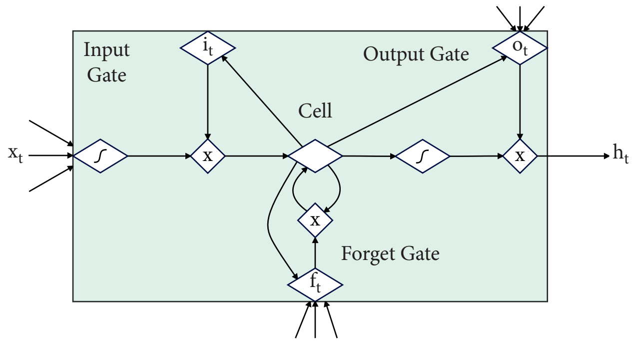

Figure 7: Structure diagram of the LSTM deep network model.

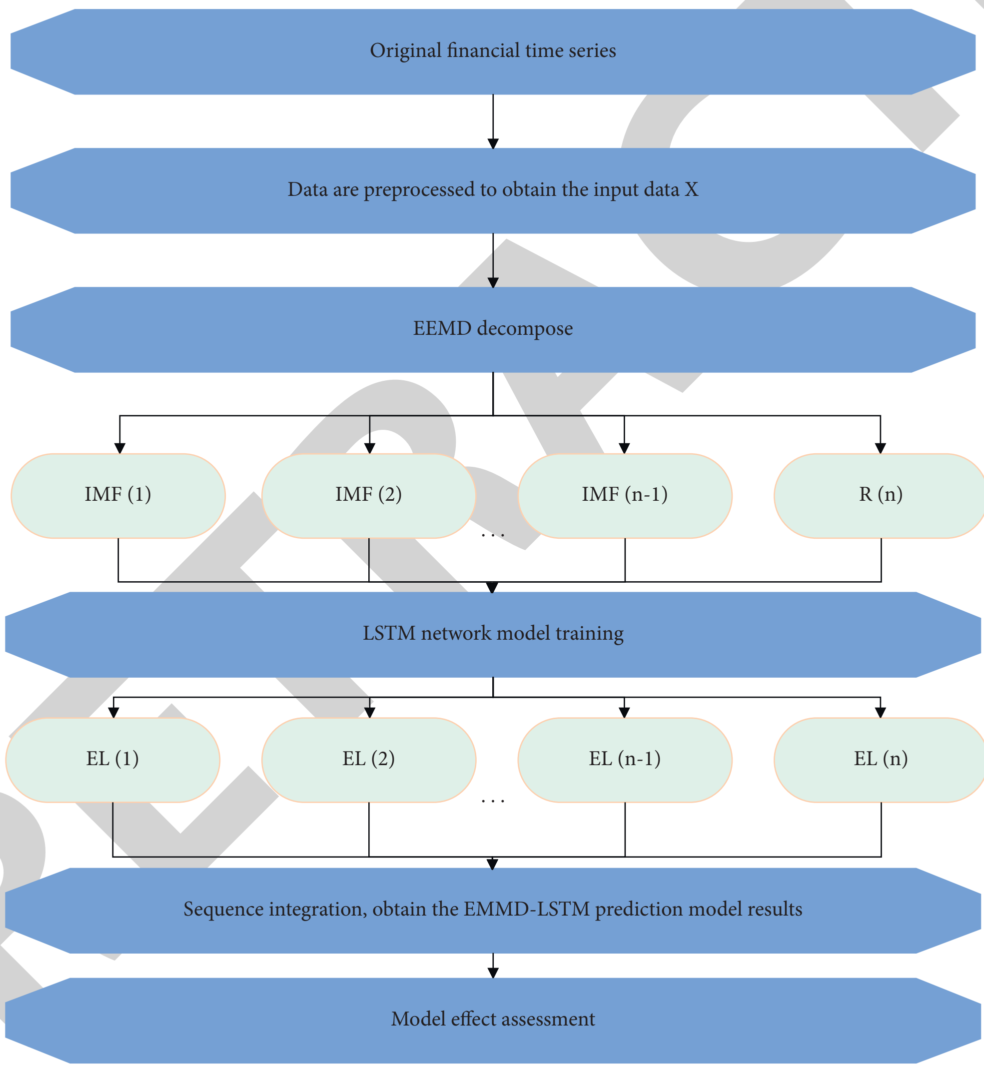

Figure 8: Flowchart of the EEMD-LSTM model.

financial practice sequence, preprocesses the input data, decomposes the preprocessed data through the EEMD algorithm, and decomposes it to obtain subsequences with different scale characteristics. After that, LSTM network is used to train and predict the decomposed sequences, and the final prediction result is obtained by comprehensive superposition. Finally, a comparative experiment is designed to evaluate the effect of the model.

On the basis of the above research, this paper combines Matlab to verify the effect of the model proposed in this paper, and the results shown in Table 1 and Table 2 are obtained.

From the above simulation data, it can be seen that the economic fluctuation analysis system based on high-frequency financial time series proposed in this paper has good economic analysis and economic forecasting effects.

Table 1: Economic analysis effect of economic fluctuation analysis system based on high-frequency financial time series.

Number	Economic Analysis	Number	Economic Analysis	Number	Economic Analysis
1	71.29	16	70.12	31	82.64
2	78.79	17	72.08	32	78.96
3	67.57	18	67.58	33	81.15
4	80.00	19	80.68	34	70.62
5	67.74	20	79.20	35	71.88
6	76.06	21	73.23	36	68.55
7	69.52	22	71.57	37	72.33
8	78.73	23	82.14	38	75.81
9	82.65	24	70.81	39	82.54
10	72.46	25	72.98	40	83.37
11	78.85	26	80.59	41	82.52
12	68.45	27	73.88	42	76.29
13	80.93	28	67.50	43	74.70
14	73.74	29	73.40	44	67.64
15	76.09	30	79.02	45	72.48

Table 2: Economic forecasting effect of economic fluctuation analysis system based on high-frequency financial time series.

Number	Economic forecast	Number	Economic forecast	Number	Economic forecast
1	70.85	16	66.89	31	62.88
2	70.09	17	63.28	32	65.51
3	75.32	18	70.19	33	73.46
4	70.07	19	76.13	34	69.29
5	75.58	20	65.14	35	76.39
6	72.69	21	78.53	36	69.17
7	70.37	22	73.26	37	75.00
8	77.31	23	66.43	38	77.01
9	67.36	24	70.95	39	65.81
10	74.14	25	65.89	40	61.83
11	66.97	26	73.50	41	76.37
12	80.26	27	64.69	42	63.72
13	64.69	28	68.96	43	74.47
14	71.89	29	69.74	44	78.03
15	77.05	30	73.99	45	70.94

5. Conclusion

In the field of financial economics, with the rapid development of science and technology, the collection and storage of high-frequency financial data has become easier and easier, which is of great significance for understanding market derivatives. Volatility provides important information to describe the dynamic evolution of market derivatives and their intrinsic properties, and is also an important basis for market derivatives pricing. Therefore, the description and prediction of the estimation model of volatility has always been an important subject in the field of financial statistics. This paper combines the high-frequency financial sequence algorithm to analyze the actual measurement of economic fluctuations, and builds an intelligent economic analysis and forecasting system. Moreover, this paper verifies the effect of the intelligent model proposed in this paper through simulation experiments. The simulation data shows that the economic fluctuation analysis system based on high-frequency financial time series proposed in this paper has good economic analysis and economic forecasting effects. [18–20]

Data Availability

The labeled dataset used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declare no competing interests.

Acknowledgments

This study is sponsored by Yellow River Conservancy Technical Institute.

References

[1] J. Pipek and S. Nagy, "An economic prediction of refinement coefficients in wavelet-based adaptive methods for electron structure calculations," *Journal of Computational Chemistry*, vol. 34, no. 6, pp. 460–465, 2013.

[2] S. Nagy and J. Pipek, "An economic prediction of the finer resolution level wavelet coefficients in electronic structure

calculations," *Physical Chemistry Chemical Physics*, vol. 17, no. 47, pp. 31558–31565, 2015.

[3] W. Yu and W. Huafeng, "Quantitative analysis of regional economic indicators prediction based on grey relevance degree and fuzzy mathematical model[J]," *Journal of Intelligent and Fuzzy Systems*, vol. 37, no. 2, pp. 1–14, 2019.

[4] P. Karanikić, I. Mladenović, S. Sokolov-Mladenović, and M. Alizamir, "Retraction Note: prediction of economic growth by extreme learning approach based on science and technology transfer," *Quality & Quantity*, vol. 53, no. 2, pp. 1095-1096, 2019.

[5] K. Ataka, "Prediction of election result and economic indicator," *resuscitation*, vol. 96, no. 6, p. 84, 2014.

[6] S. Barde, "Back to the future: economic self-organisation and maximum entropy prediction," *Computational Economics*, vol. 45, no. 2, pp. 337–358, 2015.

[7] A. Ferramosca, D. Limon, and E. F. Camacho, "Economic MPC for a changing economic criterion for linear systems," *IEEE Transactions on Automatic Control*, vol. 59, no. 10, pp. 2657–2667, 2014.

[8] L. Zhou, K. K. Lai, and J. Yen, "Bankruptcy prediction using SVM models with a new approach to combine features selection and parameter optimisation[J]," *International Journal of Systems Science*, vol. 45, no. 1-3, pp. 241–253, 2014.

[9] D. Bhattacharya, J. Mukhoti, and A. Konar, "Learning regularity in an economic time-series for structure prediction," *Applied Soft Computing*, vol. 76, no. 2, pp. 31–44, 2019.

[10] Y. Geng, Z. Wei, H. Zhang, and M. Maimaituerxun, "Analysis and prediction of the coupling coordination relationship between tourism and air environment: yangtze river economic zone in China as example," *Discrete Dynamics in Nature and Society*, vol. 2020, no. 10, 15 pages, Article ID 1406978, 2020.

[11] H. L. Vu, K. T. W. Ng, and D. Bolingbroke, "Time-lagged effects of weekly climatic and socio-economic factors on ANN municipal yard waste prediction models," *Waste Management*, vol. 84, no. 2, pp. 129–140, 2019.

[12] C. Teljeur, M. O'Neill, P. Moran, L. Murphy, and P. Harrington, "Using prediction intervals from random-effects meta-analyses IN an economic model[j]," *International Journal of Technology Assessment in Health Care*, vol. 30, no. 01, pp. 44–49, 2014.

[13] P. Rajsic, A. Weersink, A. Navabi, and P. Pauls, "Economics of genomic selection: the role of prediction accuracy and relative genotyping costs[J]," *Euphytica*, vol. 210, no. 2, pp. 1–18, 2016.

[14] F. Jahedpari, T. Rahwan, S. Hashemi et al., "Online prediction via continuous artificial prediction markets," *IEEE Intelligent Systems*, vol. 32, no. 1, pp. 61–68, 2017.

[15] V. Daksiya, H. T. Su, Y. H. Chang, and E. Y. M. Lo, "Incorporating socio-economic effects and uncertain rainfall in flood mitigation decision using MCDA," *Natural Hazards*, vol. 87, no. 1, pp. 515–531, 2017.

[16] S. Lahmiri, "A variational mode decompoisition approach for analysis and forecasting of economic and financial time series," *Expert Systems with Applications*, vol. 55, no. 8, pp. 268–273, 2016.

[17] N. Gordini, "A genetic algorithm approach for SMEs bankruptcy prediction: empirical evidence from Italy," *Expert Systems with Applications*, vol. 41, no. 14, pp. 6433–6445, 2014.