Sains Malaysiana 54(10)(2025): 2539-2551
http://doi.org/10.17576/jsm-2025-5410-16

Multiplicative Error Model Based on Robust Estimation: Evidence from High-Frequency Data in the Chinese Futures Market

(Model Ralat Pendaraban Berdasarkan Keteguhan Anggaran: Bukti daripada Data Frekuensi Tinggi dalam Pasaran Hadapan China)

TING LI* & SAIFUL IZZUAN HUSSAIN

School of Mathematical Sciences, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia

Received: 20 December 2024/Accepted: 28 August 2025

ABSTRACT

This study presents a robust estimation approach for the Multiplicative Error Model (MEM), developed for analyzing non-negative, high-frequency financial time series data. Although maximum likelihood estimation (MLE) is widely used, it is very sensitive to outliers and shows poor results for small sample sizes. To address this problem, we propose a self-weighted M-estimation method that accounts for infinite variance and weights outliers downwards, thereby improving the stability and robustness of the estimation. Simulation studies with four distributions confirm the superior performance of this method compared to MLE and LAD estimators. An empirical analysis using five-minute price spread data of eight major Chinese commodity futures - gold, petroleum asphalt, soybean, iron ore, soybean oil, corn, sugar, and rapeseed oil - demonstrates the practical advantages of this method. The results show a consistent improvement in model fit, which translates into lower AIC values and confirms the effectiveness of self-weighted M-estimation for noisy, high-frequency financial data.

Keywords: Empirical analysis; high-frequency data; Multiplicative Error Model; self-weighted M-estimation

ABSTRAK

Penyelidikan ini memperkenalkan pendekatan penganggaran teguh bagi Model Ralat Pendaraban (MEM) yang dibangunkan untuk menganalisis data siri masa kewangan frekuensi tinggi yang bukan negatif. Walaupun kaedah penganggaran kebolehjadian maksimum (MLE) digunakan secara meluas, namun ia sangat sensitif terhadap nilai terpencil dan menunjukkan prestasi yang lemah apabila saiz sampel kecil. Bagi mengatasi masalah ini, kajian ini mencadangkan kaedah penganggaran berpemberat-kendiri-M yang mengambil kira varian tak terhingga dan memberikan pemberat lebih rendah kepada nilai terpencil, sekali gus meningkatkan kestabilan dan keteguhan anggaran. Kajian simulasi yang melibatkan empat taburan menunjukkan prestasi kaedah ini lebih baik berbanding anggaran MLE dan LAD. Analisis empirik yang menggunakan data harga lima-minit bagi lapan kontrak niaga hadapan komoditi utama China - emas, asfalt petroleum, soya, bijih besi, minyak soya, jagung, gula dan minyak biji rapa - membuktikan kelebihan praktikal kaedah ini. Hasil kajian menunjukkan peningkatan tekal dalam kesesuaian model yang diterjemahkan kepada nilai AIC yang lebih rendah, sekali gus mengesahkan keberkesanan penganggaran berpemberat-kendiri-M bagi data kewangan frekuensi tinggi yang bising.

Kata kunci: Anggaran berpemberat-kendiri-M; analisis empirik; data frekuensi tinggi; Model Ralat Pendaraban

INTRODUCTION

China’s commodity futures market has experienced rapid growth over the past decade, playing a pivotal role in global commodity price discovery. With exchanges such as the Shanghai Futures Exchange (SHFE), Dalian Commodity Exchange (DCE), and Zhengzhou Commodity Exchange (ZCE), it offers deep liquidity and high-frequency trading volume, making it an ideal setting for testing robust modeling approaches under noisy and volatile conditions. With the gradual opening of the market and ongoing internationalization, the composition of market participants is becoming increasingly diverse. Zhang (2019) found in his research that the risk management function of the Chinese futures market has been significantly strengthened and that the socio-economic operation will continue to utilize this function. The development potential of China’s futures market is immeasurable, but it also faces new risk challenges. For example, Cai and Zhang (2021) have shown that there are many obstacles to the high-quality development of China’s futures market, including imperfect product structure, irrational investor structure, unclear boundary between government and market, weak futures society and low degree of internationalization.

With the development of the market, the use of high-frequency data has become an important trend in futures market analysis. High-frequency data, i.e., market information collected in extremely short time intervals, such as prices and trading volume per minute, allows researchers to observe the micro-dynamics of the market. Zhou (1996) showed in his early research that high-frequency data can show short-term market price fluctuations and the effects of liquidity. However, with the proliferation of high-frequency data applications, the problem of data noise has gradually gained attention. Data noise not only affects the accuracy of data analysis, but also has a negative impact on model predictions and decision making. Therefore, researching the problem of data noise in high-frequency data is of great theoretical and practical importance. Numerous studies have focused on how to effectively deal with data noise and minimize its impact on data analysis. For example, Ma and Yin (2012) introduced a threshold pre-averaging method to estimate high-frequency price fluctuations amid market microstructure noise and price jumps. This approach reduces volatility prediction errors and thus improves the precision of risk management.

Despite the increasing use of high-frequency data in commodity markets, many existing volatility and time series models, including GARCH-based approaches, are limited in their ability to handle non-negative, spiky and noisy data common in emerging markets such as China. In addition, most studies rely heavily on maximum likelihood estimation (MLE), which assumes finite variance and a well-specified distribution, making it highly susceptible to outliers. This poses a serious limitation for practical modeling of real-world financial time series, which often exhibit strong tails, skewness and non-Gaussian noise. Engle (2002) proposed a new non-negative financial time series model of MEM, which is based on ARCH models. MEM not only identifies the fluctuation characteristics of non-negative sequences, but also avoids the disadvantages of overlooking non-negative sequences, which is not the case with other models. Moreover, MEM solves the problem of combining low-frequency and high-frequency volatility components in both univariate and multivariate cases.

However, there is still a lack of robust estimation methods specifically designed to deal with infinite variance or non-normal data structures, which are common in high-frequency commodity price dynamics. Although the MEM offers a flexible framework for modeling non-negative financial time series, little has been done to improve its robustness under extreme data scenarios. This creates a gap between the theoretical models and their application in practice, especially in fast-moving markets such as the Chinese commodity futures exchanges.

The combination of a positive random variable and a time-varying scaling factor can represent the non-negative process that MEM investigates. Set $\{x_t, t\in N\}$ as a non-negative financial time series, $F_{t-1} = \sigma(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots)$ representing the information set available up to the time period $t-1$, then the standard MEM (p, q) can be represented as

$$x_t = \mu_t \varepsilon_t, \tag{1}$$

$$\mu_t = \omega + \sum_{i=1}^{p} \alpha_i x_{t-i} + \sum_{j=1}^{q} \beta_j \mu_{t-j}, \tag{2}$$

Among them, μ_t is the conditional expectation of x_t based on F_{t-1}, ε_t is a positive random error term with a mean of 1, and the coefficients in the model satisfy $\omega > 0$, $\alpha_i \geq 0$ ($i = 1,2,\ldots, p$) as well as $\beta_j \geq 0$ ($j = 1,2\ldots,q$), p and q are the lag order of the model.

Ma, Guo and Zhao (2014) presented a method for measuring the volatility of high-frequency financial time series that incorporates the non-negative multiplicative error model. The analysis results showed that the MEM had the best prediction performance. Zhou and Zhang (2016) proposed a relative error estimation method for MEM based on the least squares criterion, and the simulation results showed that this method has certain advantages compared with other similar methods. Taylor and Xu (2017) developed a general logarithmic vector MEM (log-vMEM). The model is applied to high-frequency data associated with a set of NYSE-listed stocks. The results show that log-vMEM is a better fit for the data than the competing model.

The rapid development of the Chinese futures market has created new opportunities for the application of high-frequency data, but has also brought challenges for data processing. When analyzing high-frequency data, more noise and error problems need to be considered due to the fine-grained characteristics. The MEM as an effective method for error handling has shown promising prospects for improving data quality, optimizing market forecasts and risk management. In this article, the specific application of the MEM to the Chinese futures market is examined in more detail, its practical effectiveness in analyzing high-frequency data is evaluated, and new perspectives and directions for future research are pointed out. The following outline describes the further structure of this article: Next section provides an overview of the research methodology. Subsequent section includes the results and discussion through numerical simulations and empirical analysis. Finally, last section provides the conclusion.

To close this methodological and empirical gap, this paper proposes a self-weighted M-estimation method tailored to the MEM framework. The aim is to improve the robustness of the model and the stability of the estimation when analyzing high-frequency, non-negative data, which may contain extreme values and outliers. Using simulation studies and empirical analysis of eight major Chinese commodity futures contracts, this study demonstrates the superiority of the proposed method over traditional MLE and LAD estimators. This research contributes to both the advancement of robust econometric modeling and the practical understanding of volatility dynamics in one of the most important emerging financial markets.

MATERIALS AND METHODS

When using MEM for practical modeling, the traditional estimation method assumes that the error term follows a known distribution with bounded variance, and maximum likelihood estimation (MLE) is used to estimate its parameters. However, high-frequency data in the actual futures market often have strong spurs and contain many outliers, and the variance of these data may even be infinite. Lu, Wang and Gao (2020) proposed the M-estimation of the MEM to solve the problems with MLE. The results showed that M-estimation works well regardless of whether the data contain outliers or not. In this paper, self-weighted M estimation (SM estimation) is used, in which outliers are weighted differently depending on their size, further reducing the impact of outliers on the estimation results.

Let $\theta = (\omega, \alpha_1, ..., \alpha_p, \beta_1, ..., \beta_q)^T$ be the model parameter, and it is known from Equation (2) that it is related to μ_t and θ, so it can be recorded $\mu_t = \mu_t(\theta)$. By taking the logarithm at both ends of Equation (1), it can be obtained that

$$\ln x_t = \ln \mu_t(\theta) + \ln, \varepsilon_t, \tag{3}$$

Let $y_t = \ln x_t - c_0$, $\eta_t = \ln \varepsilon_t - c_0$, and $c_o = median\,(\ln \varepsilon_t)$, then

$$y_t = \ln \mu_t(\theta) + \eta_t, \tag{4}$$

Therefore, the SM-estimation of parameters in MEM (p, q) is

$$\tilde{\theta}_{SM} = \arg\min_{\theta\in\Theta} \sum_{t=v+1}^{n} w_t \rho\left(y_t - \ln \mu_t(\theta)\right), \tag{5}$$

where $v = \max(p, q)$, $\Theta \subset R^{p+q+1}$ represents the parameter space, $w_t = w_t(x_{t-1}, x_{t-2}, \cdots) > 0$ is an optional weight function, and $\rho(.)$ represents a non-negative loss function.

In SM-estimation, select

$$w_t = \begin{cases} 1, a_t = 0, \\ \dfrac{c^3}{a_t^3}, a_t \neq 0, \end{cases} \tag{6}$$

where $a_t = |x_{t-1}|\, I\left(|x_{t-1}| \geq c\right)$, C is the 90% quantile point of the sequence x_t.

Here, we choose the Huber function (1964) as the loss function in the SM-estimation, which is

$$\tilde{n}_k(x) = \begin{cases} \dfrac{1}{2}x^2, |x| \leq k, \\ k|x| - \dfrac{1}{2}k^2, |x| > k, \end{cases} \tag{7}$$

where k is to be determined.

RESULTS AND DISCUSSION

In this part, a finite sample simulation is used to investigate the SM estimation of the parameters. The main goal is to compare how well SM estimation works with MLE and M estimation when the error has some outliers. In M estimation, data outliers and normal points are weighted equally, which is somewhat inappropriate. In SM estimation, outliers can be weighted differently depending on their size, further reducing their impact on the estimation results. Let us first consider the M estimate when $w_t = 1$, and the loss function assumes an absolute value function. In this case, the M-estimate is the LAD-estimate. On the other hand, if we consider the SLAD estimate with the loss function as an absolute value function in the form of (6) for w_t. Comparing the results of SLAD estimation and LAD estimation, we find that the effectiveness of SLAD estimation has improved, but not significantly. For this reason, this paper proposes to use the form of (7) as the loss function in SM estimation, where SM estimation is S-Huber estimation.

Consider the standard MEM (1, 1), where the true value of è is taken (0.1, 0.3, 0.6) and the error ε_t follows four distributions with expected values of 1, Exp (1), Pareto (3, 1.9), Weibull (1, 1.1), Burr (3, 1.5, 2), respectively. We know that financial market transactions are frequent, and new information may emerge at any time, which can have an impact on market transactions. High-frequency data is more susceptible to the reflection of market information, and it frequently contains outliers. In order to better reflect the characteristics of high-frequency financial data, 10% of the data is randomly selected from the generated error sequence random numbers, and these data are added with three times the sample standard deviation of the distribution as outliers.

Generate observational data with a sample size of 500 and compare the mean bias and mean squared error (MSE) of the three estimation methods after 2000 replicates. The specific results are shown in Table 1. It can be seen from Table 1 that the results of S-Huber estimation are generally better than those of LAD and MLE for both light-tailed and heavy-tailed error distributions. This indicates that the results of SM-estimation are more robust and suitable for modeling financial data with heavy tails and outliers (similar simulation results can also be obtained for higher-order models such as MEM(1,2), MEM(2,1), and MEM(2,2), indicating that the results of SM-estimation are more robust. This section is limited to space and does not include numerical simulation results for higher-order models).

To demonstrate the effectiveness of the SM estimation method in practice, we selected eight large-scale, highly liquid and industry-representative five-minute high-frequency futures trading data from the Shanghai Futures Exchange, Dalian Commodity Exchange and Zhengzhou Commodity Exchange, including the price range of gold, petroleum asphalt, soybean, iron ore, soybean oil, corn, sugar and rapeseed oil from June 7, 2023 to June 7, 2023. June 7, 2023 to June 7, 2024, as samples for empirical analysis. Before analyzing the data, the raw data must be pre-processed: The price range is the highest daily price minus the lowest daily price. Table 2 shows the results of the descriptive statistical analysis of the price range of eight futures. Figures 1 to 8 respectively show the time series plot for eight futures. Figures 9 to 16 respectively show the boxplot for eight futures.

From Table 2, it can be seen that the average price range of gold, petroleum asphalt, soybean, iron ore, soybean oil, corn, sugar and rapeseed oil is 0.33, 5.76, 6.36, 2.65, 14.91, 2.92, 8.14 and 17.58, respectively, with standard deviations of 0.36, 3.82, 4.14, 1.29, 7.65, 1.91, 5.14 and 9.97. From these two indicators, it can be seen that the volatility of the gold futures price is low and stable and that market expectations are relatively uniform. Trading activity could be relatively calm. The skewness is greater than 0 for all eight futures, i.e., the distribution curves of the price range are positively skewed for all eight futures and therefore well suited for modeling non-negative models. From the kurtosis point of view, the kurtosis of all eight futures is also greater than 3, indicating that the distribution curves of all eight futures have obvious characteristics of a thick peak end. In addition, it can be seen from the results of the JB statistics in the table that the P-values of all eight futures Jarque-Bra tests are less than 2.2e-16, which rejects the null hypothesis of normality and indicates that all eight futures are non-normal distributions.

Figures 1 to 8 show that the price spread between the eight futures contracts exhibits considerable data volatility. Data characterized by greater volatility tends to contain more market information, and data characterized by high volatility tends to contain more outliers and extreme scenarios. By analyzing this data, we can better compare different estimation methods and gain a deeper insight into the dynamic changes and underlying patterns of the market. From Figures 9 to 16, we can visually supplement the summary statistical data in Table 2 and provide visual evidence of potential outliers for analysis.

When performing MLE for models, a comprehensive comparison of the models under different distributions is performed to determine which empirical study is more appropriate for the price range. The use of MLE to determine BMEM for the price range of iron ore, soybean oil, sugar and rapeseed oil (compared to models such as EMEM, WMEM, PMEM. BMEM estimation is relatively effective), and the use of M-estimation and SM-estimation to estimate the MEM parameters. The specific results are shown in Table 3.

From Table 3, it can be observed that $\alpha + \beta$ of both futures is less than 1, meeting the requirements for model stationarity, and the values are both greater than 0.85, indicating a strong clustering effect. From the results of LogL, the LAD estimation results are relatively close to those of MLE. From the results of MLE to S-Huber estimation of AIC, iron ore decreases from 3.0258 to 2.9923, a decrease of 1.11%; soybean oil decreases from

TABLE 1. Simulation results under different distributions

Distributions	Estimations	*Bias* ($\grave{u}$)	*Bias* ($\acute{a}$)	*Bias* ($\hat{a}$)	*MSE* ($\grave{u}$)	*Bias* ($\acute{a}$)	*MSE* ($\hat{a}$)
Exp (1)	MLE	0.0520	0.0868	-0.0085	0.0053	0.0104	0.0024
	LAD	0.0233	-0.0011	-0.0121	0.0039	0.0035	0.0054
	S-Huber	0.0133	-0.0117	-0.0112	0.0023	0.0030	0.0040
Pareto (3,1.9)	MLE	0.0783	0.1246	-0.0097	0.2707	0.0255	0.0035
	LAD	0.0190	-0.0030	-0.0064	0.0088	0.0036	0.0047
	S-Huber	0.0164	-0.0005	-0.0068	0.0059	0.0042	0.0043
Weibull (1,1.1)	MLE	0.0757	0.1240	-0.0048	0.0108	0.0181	0.0017
	LAD	0.0308	-0.0046	-0.0049	0.0065	0.0027	0.0041
	S-Huber	0.0195	-0.0108	-0.0077	0.0038	0.0024	0.0031
Burr (3,1.5,2)	MLE	0.0694	0.1076	-0.0063	0.0098	0.0147	0.0022
	LAD	0.0242	-0.0032	-0.0051	0.0042	0.0021	0.0032
	S-Huber	0.0205	-0.0006	-0.0068	0.0030	0.0019	0.0026

3.4673 to 3.4337, a decrease of 0.97%; sugar decreases from 3.3599 to 3.3186, a decrease of 1.23%; rapeseed oil decreases from 3.6682 to 3.6278, a decrease of 1.10%. Overall, the S-Huber estimation results are better than MLE's.

Using MLE to establish WMEM for the price range of gold, petroleum asphalt, soybean, and corn (compared to models such as EMEM, PMEM, BMEM. WMEM estimation is relatively effective), and using M-estimation and SM-estimation to estimate MEM parameters. The specific results are shown in the Table 4.

From Table 4, it can be observed that $\alpha + \beta$ of both futures is less than 1, meeting the requirements for model stationarity, and the values are both greater than 0.85, indicating a strong clustering effect. From the results of LogL, the LAD estimation results are relatively close to those of MLE. From the results from MLE to S-Huber estimation of AIC, gold decreases from 3.4049 to 3.3477, a decrease of 1.68%; petroleum asphalt decreases from 3.3756 to 3.3352, a decrease of 1.20%; soybean decreases from 3.1610 to 3.1210, a decrease of 1.27%; corn decreases from 3.2053 to 3.1418, a decrease of 1.98%. Overall, the S-Huber estimation results are better than MLE's.

TABLE 2. Descriptive statistical analysis of futures price range

Futures name	Sample size	Mean	Std	Skewness	Kurtosis	JB Stat
Gold	26467	0.33	0.36	6.26	76.15	6568331 (< 2.2e-16)
Petroleum asphalt	16554	5.76	3.82	3.52	23.86	426991 (< 2.2e-16)
Soybean	16554	6.36	4.14	3.39	23.35	407809 (< 2.2e-16)
Iron ore	16554	2.65	1.29	2.21	10.93	95906 (< 2.2e-16)
Soybean oil	16554	14.91	7.65	2.89	21	327347 (< 2.2e-16)
Corn	16554	2.92	1.91	6.74	158.5	17457560 (< 2.2e-16)
Sugar	16554	8.14	5.14	8.89	284.37	56010575 (< 2.2e-16)
Rapeseed oil	16554	17.58	9.97	3.8	37.87	1029282 (< 2.2e-16)

TABLE 3. Comparisons of three estimations

Futures name	Estimations	*ù*	*á*	*â*	*LogL*	*AIC*
Iron ore	MLE	0.2984	0.2090	0.6808	-1.5127	3.0258
	LAD	0.2952	0.2103	0.6807	-1.5129	3.0262
	S-Huber	0.3255	0.2039	0.6807	-1.4960	2.9923
Soybean oil	MLE	0.8946	0.1568	0.7856	-1.7335	3.4673
	LAD	1.0192	0.1625	0.7716	-1.7335	3.4673
	S-Huber	0.9036	0.1397	0.8065	-1.7167	3.4337
Sugar	MLE	0.9814	0.2203	0.6615	-1.6797	3.3599
	LAD	1.0651	0.2233	0.6481	-1.6793	3.3590
	S-Huber	1.0694	0.2048	0.6735	-1.6591	3.3186
Rapeseed oil	MLE	0.8120	0.1637	0.7929	-1.8339	3.6682
	LAD	0.8258	0.1677	0.7884	-1.8338	3.6679
	S-Huber	0.8455	0.1437	0.8170	-1.8137	3.6278

TABLE 4. Comparisons of three estimations

Futures name	Estimations	*ù*	*á*	*â*	*LogL*	*AIC*
	MLE	0.0183	0.2418	0.7031	-1.7023	3.4049
Gold	LAD	0.0069	0.2508	0.7379	-1.7172	3.4345
	S-Huber	0.0082	0.2481	0.7454	-1.6738	3.3477
	MLE	1.2230	0.3750	0.4182	-1.6876	3.3756
Petroleum asphalt	LAD	0.5543	0.2571	0.6491	-1.6929	3.3861
	S-Huber	0.6700	0.2949	0.6046	-1.6674	3.3352
	MLE	1.3885	0.3970	0.3896	-1.5803	3.1610
Soybean	LAD	0.6878	0.2693	0.6235	-1.5879	3.1762
	S-Huber	0.7869	0.3037	0.5859	-1.5603	3.1210
	MLE	0.9408	0.3118	0.3654	-1.6025	3.2053
Corn	LAD	0.5459	0.2646	0.5542	-1.5952	3.1908
	S-Huber	0.5786	0.2913	0.5298	-1.5707	3.1418

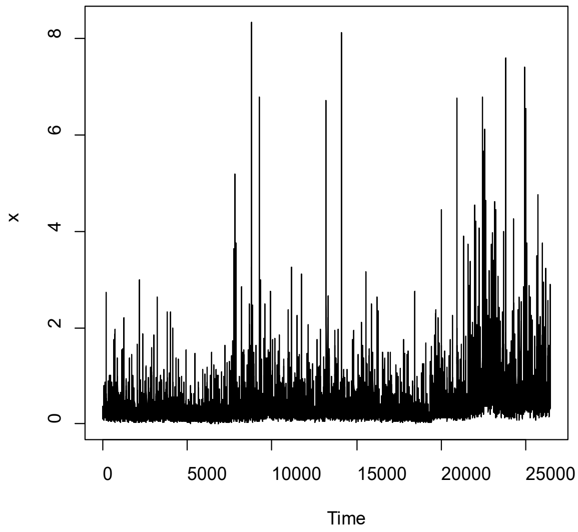

FIGURE 1. The time series plot for gold

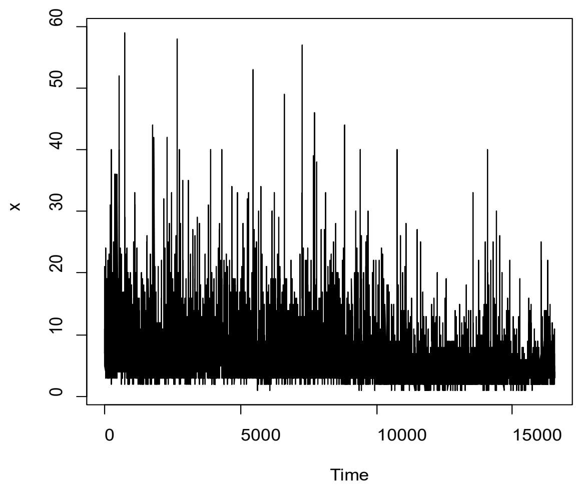

FIGURE 2. The time series plot for petroleum asphalt

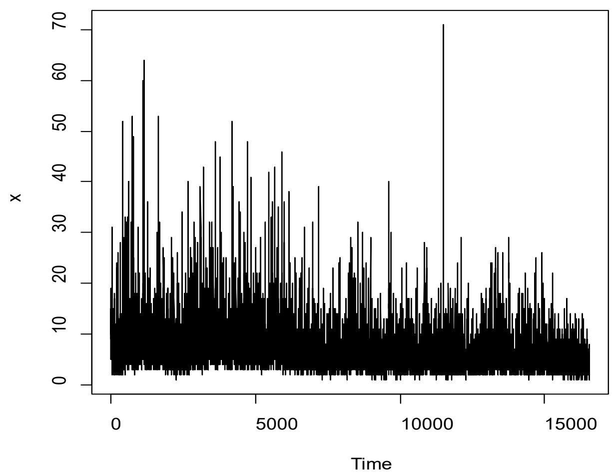

FIGURE 3. The time series plot for soybean

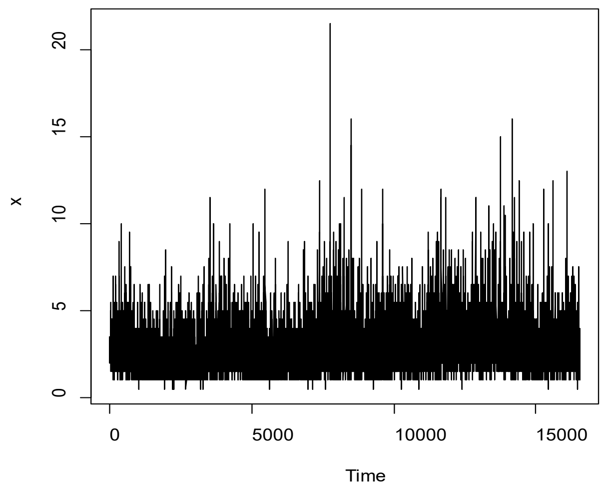

FIGURE 4. The time series plot for iron ore

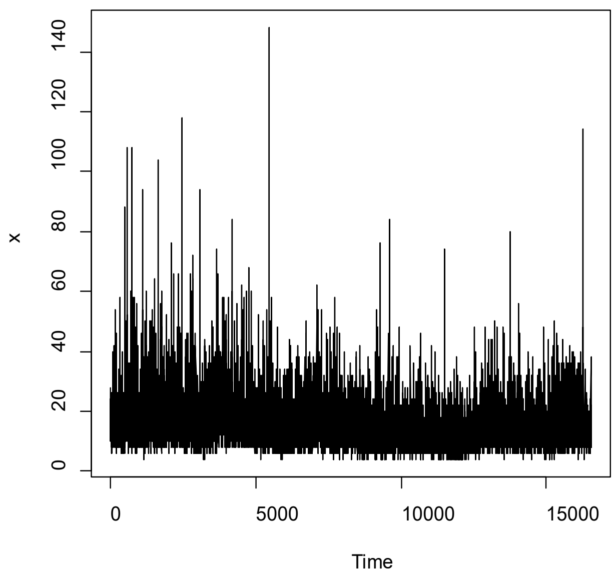

FIGURE 5. The time series plot for soybean oil

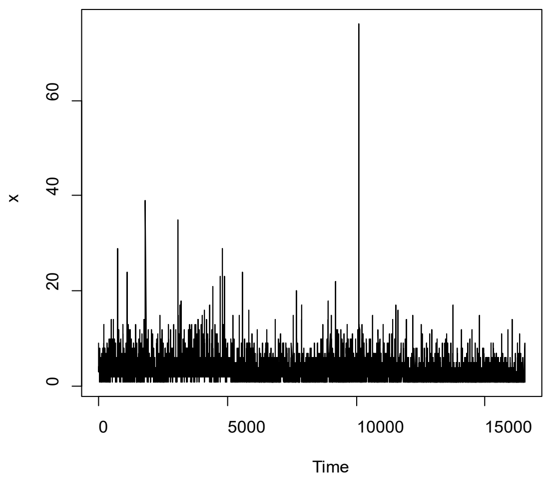

FIGURE 6. The time series plot for corn

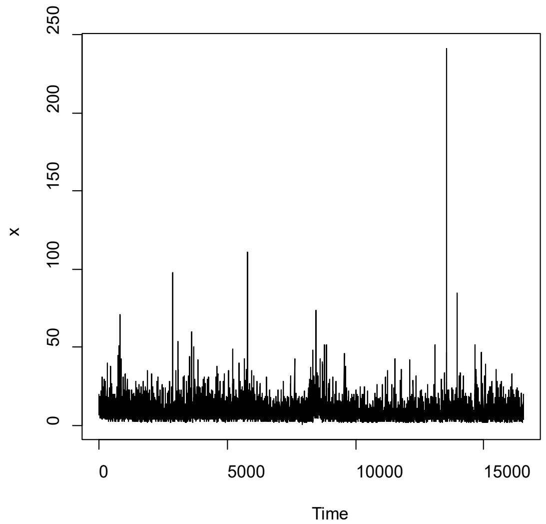

FIGURE 7. The time series plot for sugar

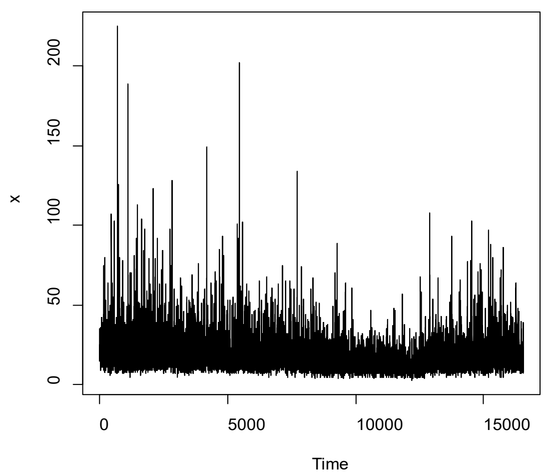

FIGURE 8. The time series plot for rapeseed oil

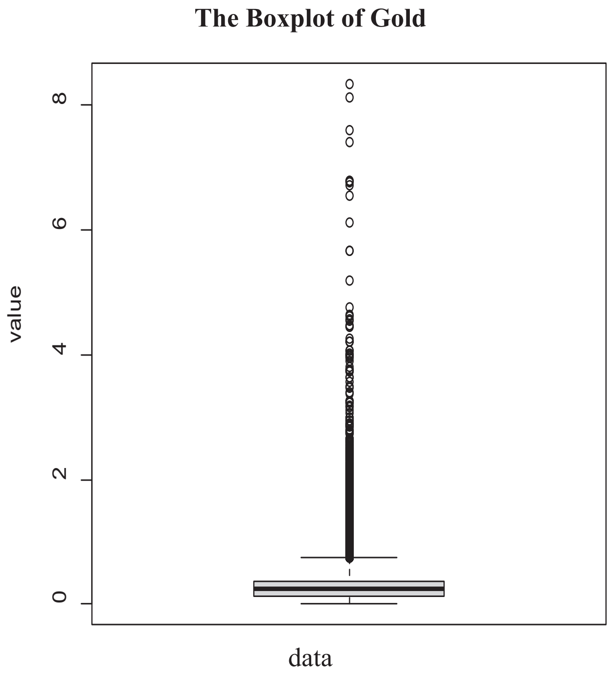

FIGURE 9. The boxplot for gold

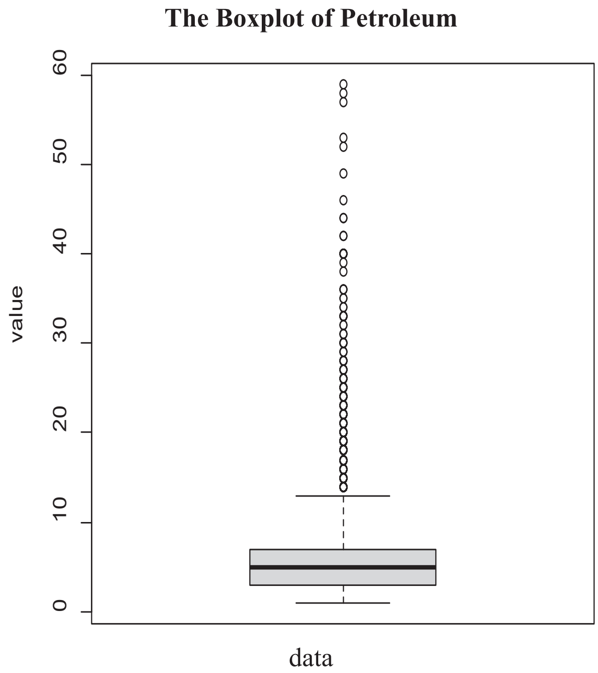

FIGURE 10. The boxplot for petroleum asphalt

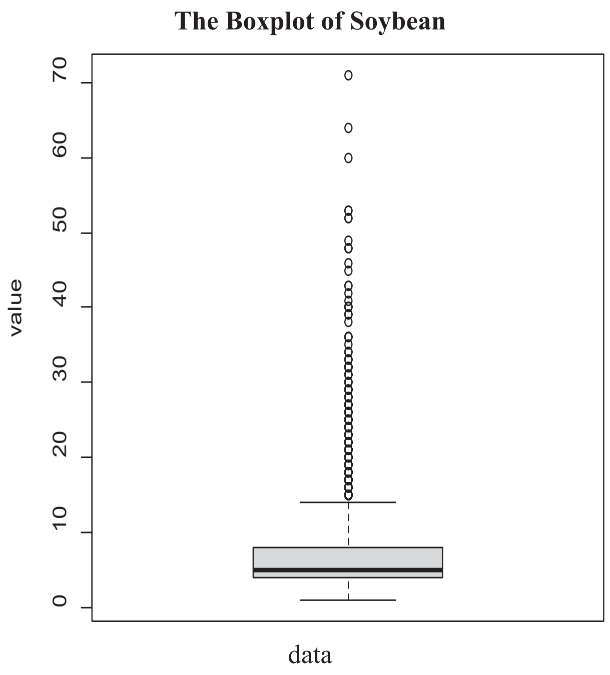

FIGURE 11. The boxplot for soybean

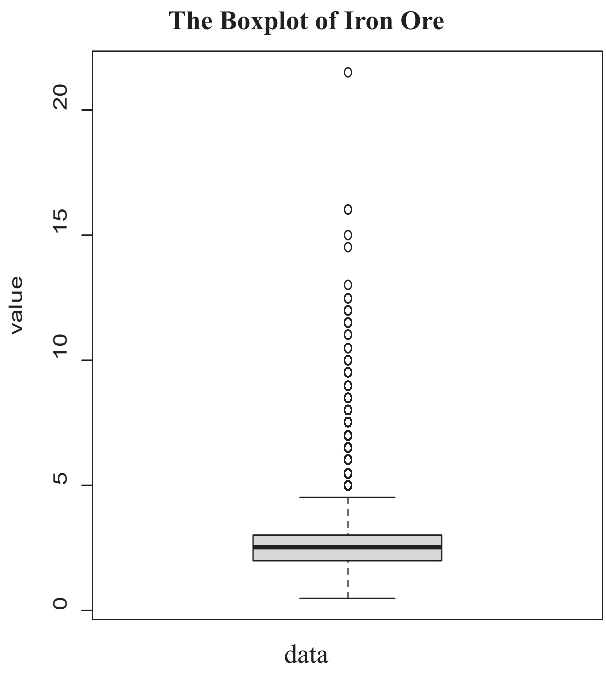

FIGURE 12. The boxplot for iron ore

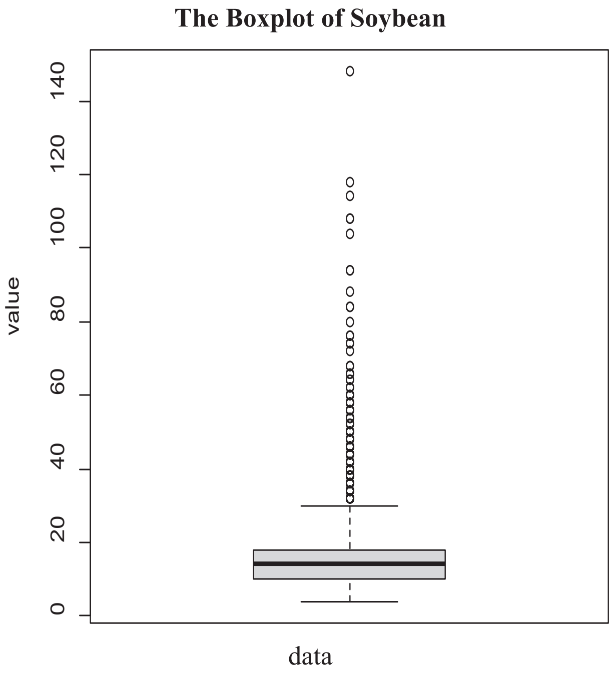

FIGURE 13. The boxplot for soybean oil

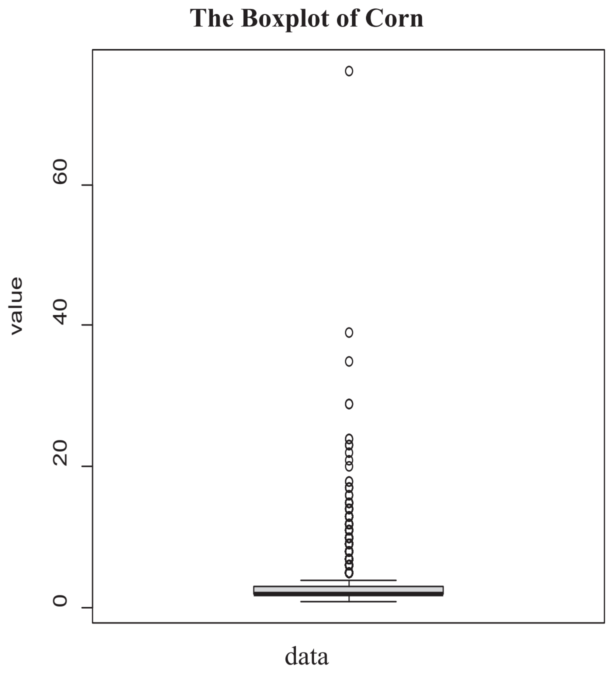

FIGURE 14. The boxplot for corn

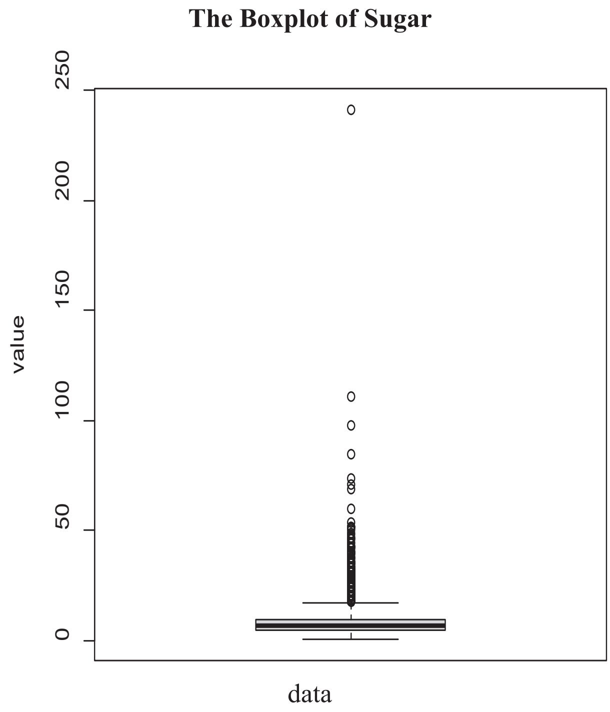

FIGURE 15. The boxplot for sugar

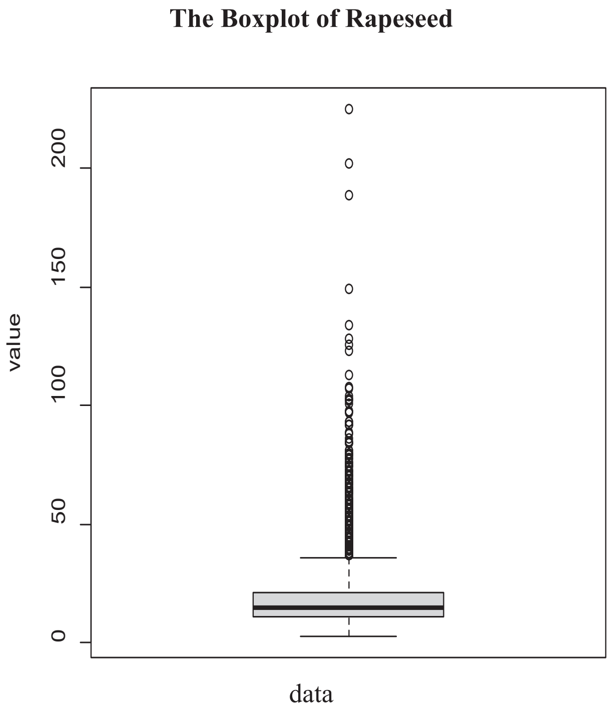

FIGURE 16. The boxplot for rapeseed oil

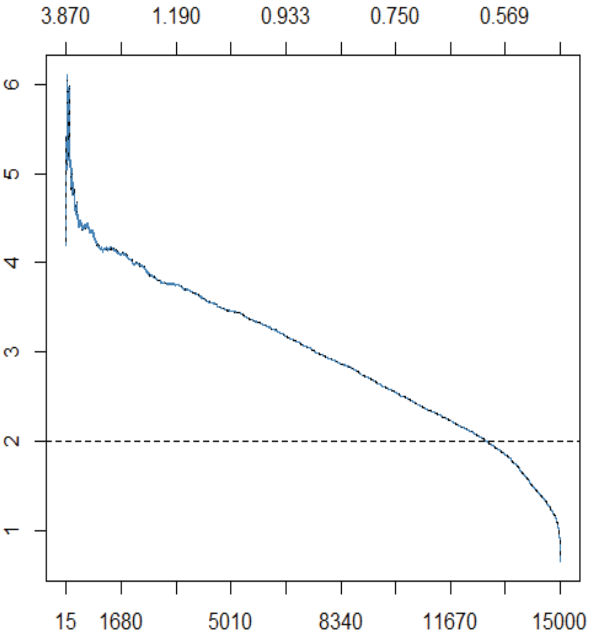

FIGURE 17. Hill estimation of iron ore price range

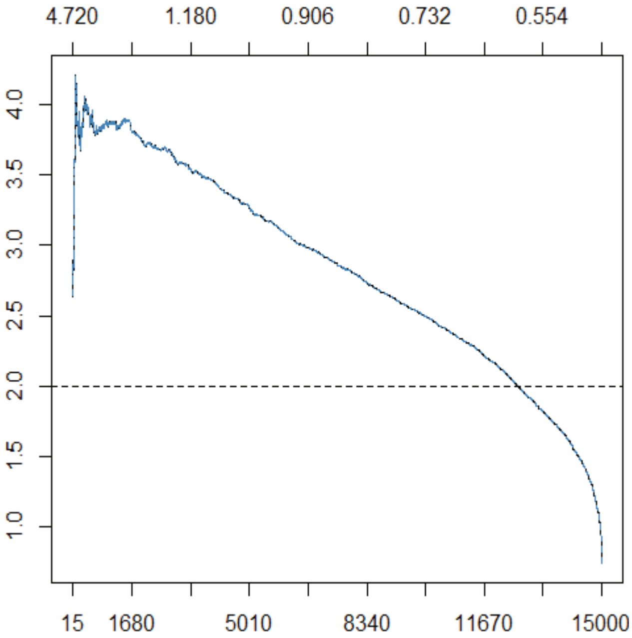

FIGURE 18. Hill estimation of soybean oil price range

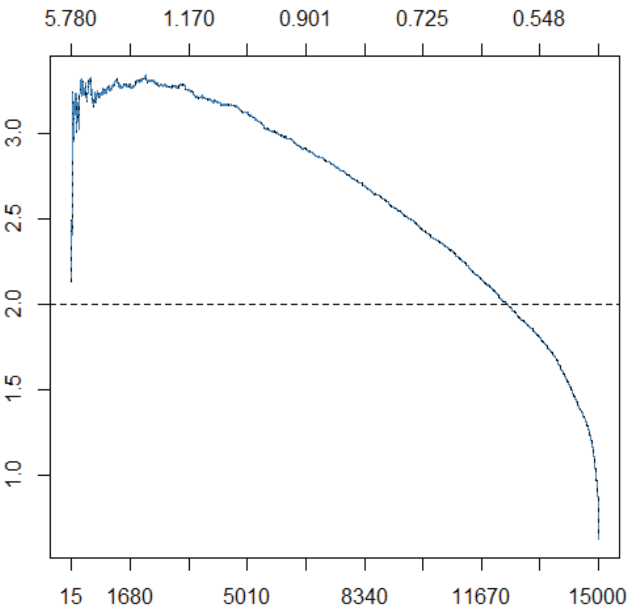

FIGURE 19. Hill estimation of sugar price range

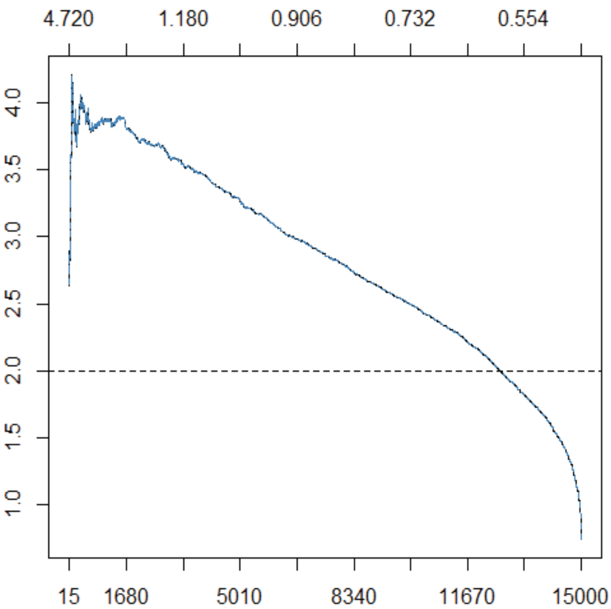

FIGURE 20. Hill estimation of rapeseed oil price range

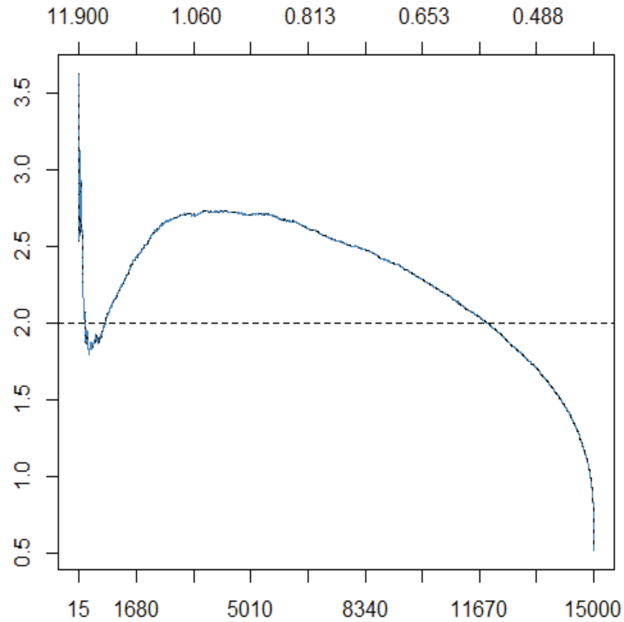

FIGURE 21. Hill estimation of gold price range

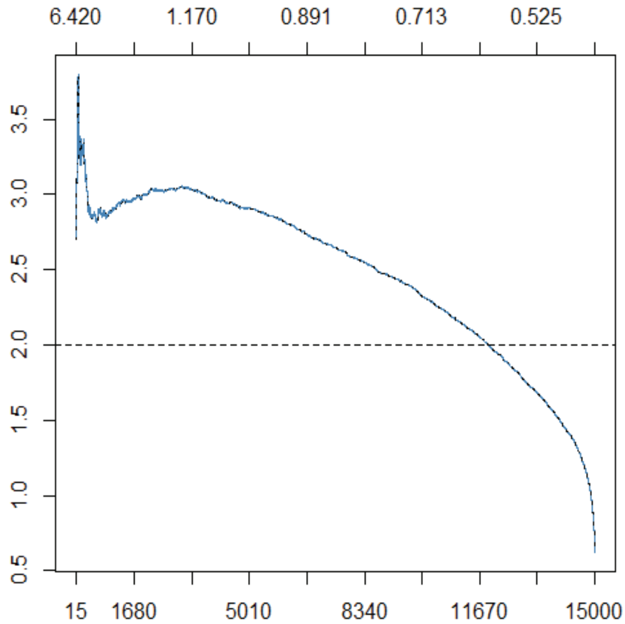

FIGURE 22. Hill estimation of petroleum asphalt price range

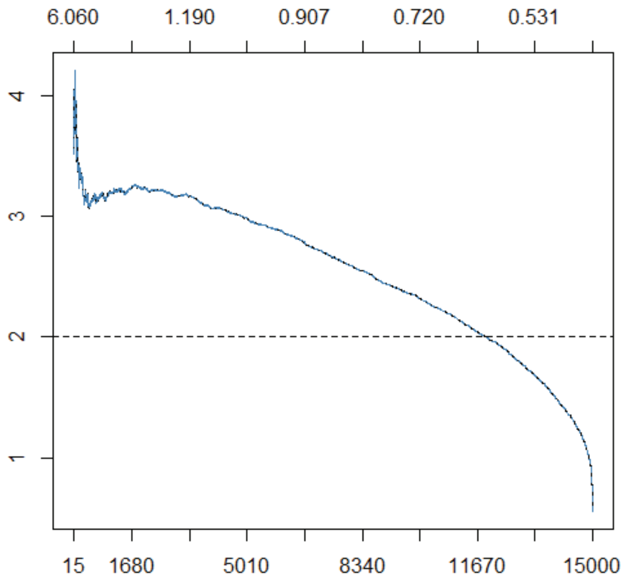

FIGURE 23. Hill estimation of soybean price range

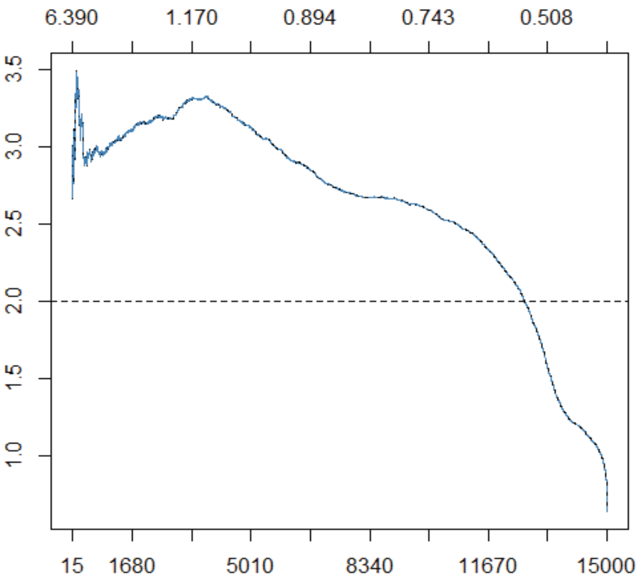

FIGURE 24. Hill estimation of corn price range

The Hill estimation method proposed by Resnick (1997) is used to estimate the tail index of residual sequences. Figures 17 to 24 show the estimation of the tail index of residual sequences after modeling the price range of eight futures using the S-Huber estimation. It cannot be concluded from the figure that the variance of residual sequences is finite, which means that the results of SM estimation are more reliable compared to other estimation methods.

CONCLUSIONS

This study set out to address a critical challenge in modeling non-negative, high-frequency financial time series data - namely, the sensitivity of conventional estimation techniques such as maximum likelihood estimation (MLE) and least absolute deviation (LAD) to outliers and non-normality, especially with small sample sizes and infinite variance. Building on the MEM, we introduced a self-weighted M-estimation approach to improve the robustness and stability of parameter estimation under these challenging data conditions. The motivation for this research stems from the increasing complexity and noise in financial markets, especially in high-frequency environments. The Chinese commodity futures market, one of the most dynamic and liquid in the world, provides a rich empirical environment for testing advanced statistical models. Given the importance of Chinese exchanges such as SHFE, DCE, and ZCE to global commodity trading, the ability to extract meaningful insights from high-frequency data in this context is of great practical and scientific interest.

Through extensive simulation studies across multiple distributions, our proposed method consistently outperformed traditional estimators in terms of efficiency and resilience to outliers. Empirical validation using five-minute price-voltage data of eight actively traded Chinese commodity futures further demonstrated the applicability of the model in practice. The results not only confirmed improved model fit and lower AIC values, but also underscored the practical utility of self-weighted M-estimation in volatile and data-intensive environments. These results have important implications for market participants and policy makers. By improving the estimation of MEM in the context of Chinese high-frequency markets, our approach enables more accurate volatility modeling, better risk management and more informed trading strategies. It also sets the stage for future research on robust econometric methods suitable for emerging markets where structural shifts and data irregularities are common. To summarize, this study contributes to the literature on financial econometrics both methodologically and empirically. It offers a statistically sound and practically relevant solution for modeling noisy, non-negative time series data that is directly applicable to one of the fastest growing financial markets in the world.

REFERENCES

Cai, S.X. & Zhang, B. 2021. High quality development of China's futures market: Connotation, characteristics, and implementation path. *Southwest Finance* (12): 80-91.

Engle, R. 2002. New frontiers for arch models. *Journal of Applied Econometrics* 17(5): 425-446.

Huber, P.J. 1964. Robust estimation of a location parameter. *The Annals of Mathematical Statistics* 35: 73-101.

Lu, W., Wang, Y. & Gao, Y. 2020. M-estimates for the multiplicative error model. *Journal of Statistical Computation and Simulation* 90(1): 1-27.

Ma, D. & Yin, Y. 2012. Noise, jumping, and high frequency price fluctuations: Analysis of fluctuations based on threshold pre-averaging. *Financial Research* (04): 124-139.

Ma, Y., Guo, P. & Zhao, Y. 2014. The empirical research on volatility measurement model based multiplicative error model. *Seventh International Joint Conference on Computational Sciences and Optimization,* Beijing, China. pp. 455-458. doi: 10.1109/CSO.2014.156

Resnick, S.I. 1997. Heavy tail modelling and teletraffic data. *Annals of Statistics* 25(5): 1805-1869.

Taylor, N. & Xu, Y. 2017. The logarithmic vector multiplicative error model: An application to high frequency NYSE stock data. *Quantitative Finance* 17(7): 1021-1035.

Zhang, H.W. 2019. Research on the development and risk management of China's futures market. *Business News* 000 (052): 237.

Zhou, B. 1996. High-frequency data and volatility in foreign-exchange rates. *Journal of Business & Economic Statistics* (14): 45-52.

Zhou, S.B. & Zhang, B. 2016. Least squares relative error estimation for multiplicative error model. *Statistics & Decision* (20): 9-12.

*Corresponding author; email: p129842@siswa.ukm.edu.my