INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN BUSINESS & SOCIAL SCIENCES
Vol 15, Issue 12, (2025) E-ISSN: 2222-6990

Price Dynamics and Arbitrage Opportunities in the Palm Oil Futures Markets: Insights from Malaysia and China Using High-Frequency Data

Li Yang[1,2], Razli Ramli[1], Darween Rozehan Shah Iskandar Shah[3], Almi Mahmod[1]
[1]Azman Hashim International Business School, University of Technology Malaysia, 54100 Kuala Lumpur, Malaysia, [2]School of Digital Economics, Nanning University, 530000, Nanning, China, [3]Institute for Advanced Studies, University of Malaya, 50603 Kuala Lumpur, Malaysia
*Corresponding Author Email: razli@utm.my

DOI Link: http://dx.doi.org/10.6007/IJARBSS/v15-i12/27157
Published Date: 24 December 2025

Abstract
This study investigates the price dynamics and arbitrage opportunities between Malaysia's Crude Palm Oil Futures (FCPO) and China's Palm Oil Futures (P.DCE), using high-frequency 5-minute intraday data from January 2019 to December 2023. Employing an integrated analytical framework that combines linear and nonlinear methods including Pearson correlation, E-G cointegration tests, Hurst exponent analysis, wavelet-based multiscale decomposition, threshold vector error correction modeling (TVECM), and transfer entropy we further enhance robustness by incorporating machine learning techniques, specifically Support Vector Regression (SVR) with a linear kernel. Our results reveal an exceptionally strong long-run co-movement between FCPO and P.DCE (Pearson's r = 0.9882) and confirm the presence of a stable cointegrating relationship from both linear and non-linear perspectives. However, short-term deviations from equilibrium persist, exhibiting mean-reverting behavior (H < 0.5), which creates statistically detectable, albeit transient, arbitrage windows. Nonlinear analyses indicate asymmetric adjustment dynamics and directional information flow, with FCPO maintaining a leading role in price discovery, though P.DCE's influence has grown significantly. These findings demonstrate that high-frequency statistical arbitrage strategies are feasible in this cross-market context. By bridging a critical gap in the literature using intraday data and advanced analytics, this study offers valuable insights for traders, risk managers, and policymakers, while contributing to the broader discourse on market efficiency and cross-border linkages in agricultural commodity futures.
Keywords: Crude Palm Oil Futures, High-Frequency Trading, Machine Learning, Market Efficiency

Introduction

The global palm oil market plays a vital role in the agricultural commodity sector, significantly influencing international trade, investment flows, and economic stability in both major producing and consuming countries (FAO), 2022; Malaysian Palm Oil Council, 2023b; World Bank, 2023). As one of the most widely consumed vegetable oils worldwide, palm oil has surpassed soybean oil in global consumption, with Malaysia and Indonesia emerging as the largest producers (Tandra et al., 2022). Malaysia, as the second-largest exporter, plays a pivotal role in global price formation, while China being the second-largest importer and the fourth-largest consumer exerts substantial influence on international demand and price volatility (M. Y. Ahmad et al., 2022). Given the strong economic interdependence between these two economies, understanding the price dynamics of palm oil futures markets is of critical importance to traders, investors, and policymakers (Ahmed et al., 2020).

Among the global futures markets, Malaysia's Crude Palm Oil Futures (FCPO), traded on Bursa Malaysia Derivatives Exchange (BMD), and China's Palm Oil Futures (P.DCE), listed on the Dalian Commodity Exchange (DCE), serve as key instruments for price discovery and risk management (Du, 2018; Malaysian Palm Oil Council, 2023a). Introduced in 1980 and re-listed in 1985, FCPO has long been recognised as a global benchmark for palm oil pricing (Bursa Malaysia, 2019). In contrast, P.DCE, launched in 2007, has gained increasing prominence in China's domestic futures market, reflecting the country's growing consumption and strategic importance in the global palm oil supply chain (Du, 2018). Despite the significance of both contracts, empirical evidence on the dynamic linkage between FCPO and P.DCE remains limited, particularly from a high-frequency trading (HFT) perspective.

From a theoretical standpoint, this study is grounded in the Efficient Market Hypothesis (EMH) and the law of one price. According to EMH, asset prices in efficient markets should fully and rapidly incorporate available information, thereby limiting the persistence of arbitrage opportunities (Sewell, 2012). Similarly, the law of one price posits that futures contracts written on the same underlying commodity should converge to a common value once price differentials, transaction costs, and exchange rate effects are accounted for. However, in practice, market frictions such as liquidity constraints, information asymmetry, regulatory differences, and trading delays often prevent instantaneous price convergence. These frictions may give rise to short-lived mispricing and arbitrage opportunities, particularly in globally interconnected commodity markets (Fassas, 2011);(Barboza Martignone et al., 2022). Importantly, price transmission across markets may be asymmetric and nonlinear, suggesting that deviations from equilibrium can persist long enough to be exploited by sophisticated traders (Kallandranis et al., 2024; WOODE, 2024).

Empirical studies on FCPO have predominantly relied on low-frequency data, focusing on daily or monthly price movements. For example, (Ahmed et al., 2020) analysed daily FCPO prices between 1998 and 2010 and found strong cointegration between futures and spot markets, indicating long-run equilibrium relationships. Subsequent research identified macroeconomic determinants including production levels, export demand, stock inventories, interest rates, and exchange rates as key drivers of FCPO price fluctuations (Ahmed et al., 2020; Go & Lau, 2024; Rizal et al., 2023).While these studies provide valuable insights into long-term price behaviour, their reliance on low-frequency data limits their ability to capture short-term market inefficiencies and intraday arbitrage dynamics.

Research on cross-market interactions between Malaysian and Chinese palm oil futures remains even more limited. Early work by Mahadi et al. (2014) documented cointegration between FCPO and P.DCE prices, reinforcing Malaysia's dominant role in global palm oil price discovery. Nevertheless, these studies largely employ daily data and therefore fail to reflect the realities of modern electronic markets, where algorithmic and high-frequency traders operate on intraday time horizons (Kenanga Futures, 2023). As algorithmic trading and HFT have become increasingly prevalent, arbitrage opportunities tend to emerge and disappear within minutes or even seconds, necessitating the use of high-frequency data for accurate detection and analysis (N. Ahmad, 2010).

Recent evidence from other financial and commodity markets highlights the advantages of high-frequency data in uncovering short-term arbitrage opportunities and improving trading strategies. Hu et al., (2024) demonstrate that high-frequency pairs-trading frameworks can effectively identify arbitrage opportunities in China's commodity futures markets. Similarly, Poutré et al., (2023) show that incorporating high-frequency data enhances arbitrage strategies in North American equity markets, while Miao, (2014) emphasises the superior ability of intraday data to generate trading signals and improve market efficiency analysis. Collectively, these findings underscore the need to extend high-frequency analytical frameworks to palm oil futures markets, where empirical evidence remains scarce.

Against this backdrop, the present study examines the dynamic relationship between Malaysian and Chinese palm oil futures prices using five-minute intraday data spanning the period from 2019 to 2023. By analysing linear interdependencies, long-run equilibrium relationships, and short-term price adjustments, this study aims to provide a comprehensive assessment of market efficiency and arbitrage potential between FCPO and P.DCE. Specifically, the study addresses the following research questions:
(1) To what extent do FCPO and P.DCE exhibit market efficiency when analysed using high-frequency data?
(2) What is the nature and strength of the price relationship between the two markets?
(3) Do statistically significant arbitrage opportunities exist between FCPO and P.DCE futures contracts?

By addressing these questions, this study contributes to the literature on cross-market price transmission and high-frequency arbitrage in commodity futures markets. The integration of advanced econometric techniques with machine learning and nonlinear methods enables a richer understanding of price dynamics, information flow, and market efficiency. The findings offer practical implications for traders seeking to design effective hedging and arbitrage strategies, as well as for policymakers aiming to enhance market transparency, stability, and regulatory effectiveness in the global palm oil industry.

Proposed Theoretical Framework Paragraph

From a theoretical perspective, this study is grounded in the Efficient Market Hypothesis (EMH), the law of one price, and the literature on information transmission and market microstructure. EMH posits that asset prices should rapidly and fully reflect available information, thereby limiting the persistence of arbitrage opportunities. In a similar vein, the law of one price suggests that futures contracts written on the same underlying commodity across different markets should converge to a common equilibrium once exchange rates, transaction costs, and market frictions are accounted for. However, empirical evidence indicates that institutional differences, liquidity constraints, trading hours, and informational asymmetries may impede instantaneous price adjustment, particularly in cross-border futures markets. These frictions give rise to short-lived mispricing and asymmetric adjustment dynamics, which are more likely to be observed at high frequencies. Consequently, this theoretical framework motivates the use of high-frequency data and a combination of linear and nonlinear analytical methods to examine market efficiency, equilibrium relationships, and transient arbitrage opportunities between FCPO and P.DCE.

Methodology

Data Collection

This study aimed to quantitatively analyse the data. This study was primarily based on secondary data sources. Secondary data were collected from publicly available historical data. The market publicly available historical data used in this study include the 5-minute intraday trading prices and volumes of the main continuous contract of P.DCE traded on Dalian Commodity Exchange in China and the main continuous contract of FCPO traded on Bursa Malaysia from January 4, 2019, to December 31, 2023. The wind database can provide high-frequency data for the P.DCE. The high-frequency FCPO data used in this study were obtained from the Portara CQG Data Factory.

There are four reasons for importing high-frequency data in this study. First, high-frequency data allow for a more precise capture of short-term market movements and small price fluctuations, helping identify instant price deviations that provide timely trading signals and enhance arbitrage profit opportunities. Second, high-frequency data enable faster strategy responses to market changes, ensuring timely adjustments to trading decisions, reducing delays, and improving overall strategy execution. Furthermore, high-frequency data increase the number of sample points, which aids in the development of more complex predictive models, reduces the overfitting risk, and enhances model stability and accuracy. Finally, high-frequency data significantly improve the identification of arbitrage opportunities, as frequent market monitoring helps detect short-term price discrepancies, boosting the effectiveness and success of arbitrage strategies.

Data Cleaning

Table 1.1 presents a comparison of Malaysian FCPO and Dalian Commodity Exchange palm oil futures contracts. To guarantee data consistency, we standardised the trading time, transaction date, and pricing units of the sample data. There are some differences in the trading hours of the FCPO and P.DCE contracts. There was no jet lag between Malaysia and China. Therefore, the sample data should be adjusted to align with the trading times of both futures contracts. The adjusted co-trading time was from Monday to Friday, with morning sessions running from 10:30 AM to 11:30 AM, afternoon sessions running from 2:30 PM to 3:00 PM, and post-market sessions running from Monday to Thursday between 9:00 PM and 11:00 PM.

The sample data for both Malaysia and China's holidays were erased to prevent a scenario in which one market remained open while the other was closed. FCPO is expressed in Malaysian Ringgit (MYR), and P.DCE is expressed in Chinese Yuan Renminbi (CNY). Thus, this study uses the CNY-MYR exchange rate authorised by the China Foreign Exchange Trading Centre to consolidate the sample data for the two futures contract values. After cleaning 29, 866 observations were obtained. Data cleaning was performed using Python on a PyCharm CE.

Table 1.1
Comparison of Malaysian FCPO and Dalian Commodity Exchange Palm Oil Futures Contracts

Attribute/Commodity	**Malaysian Crude Palm Oil (FCPO)**	**Dalian Commodity Exchange Palm Oil**
Trading Unit	25 tons per lot	10 tons per lot
Quoting Unit	MYR/ton	CNY/ton
Price Limit Range	10% limit, 15% under specific conditions	4%
Contract Months	Spot month and the following 11 consecutive months. every other month for the next 36 months	January to December
Trading Hours	Monday to Friday: 10:30 am - 12:30 pm, 2:30 pm - 6:00 pm. Post-market： 9:00 pm - 11:30 pm from Monday to Thursday	Monday to Friday: 9:00 am - 11:30 am, 1:30 pm - 3:00 pm. Post-market： 9:00 pm - 11:00 pm from Monday to Thursday
Delivery Grade	Must meet the Malaysian Sustainable Palm Oil (MSPO) certification	Dalian Commodity Exchange Palm Oil Delivery Quality Standard
Delivery Location	Penang/North Butterworth, Port Klang, and Pasir Gudang ports, etc.	Designated palm oil delivery warehouses of Dalian Commodity Exchange
Delivery Method	Physical delivery	Physical delivery
Trading Code	FCPO	P
Exchange	Bursa Malaysia	Dalian Commodity Exchange
Attribute/Commodity	Malaysian Crude Palm Oil (FCPO)	Dalian Commodity Exchange Palm Oil

[Source**:** The author]

Data Analysis Procedure

To comprehensively investigate the dynamic interdependence between FCPO and P.DCE futures prices, we implement this study across four fundamental dimensions: (1) correlation structure, (2) long-term memory, (3) long-run equilibrium, and (4) causal linkage.

As a baseline, Pearson correlation and cross-correlation analysis provide initial insights into the strength and timing of short-term linear co-movements between the two series. The cointegration test assesses whether a stable long-run equilibrium exists, which is a prerequisite for statistical arbitrage (Bui & Ślepaczuk, 2022). Meanwhile, the Hurst index helps characterize the memory property of price deviations—distinguishing between random walk ($H \approx 0.5$), persistent trends ($H > 0.5$), and mean-reverting dynamics ($H < 0.5$). Finally, linear Granger causality examines the direction of predictive power under the assumption of symmetric and constant relationships.

However, financial markets often exhibit nonlinearities, regime shifts, and state-dependent dynamics that linear methods may overlook. To address this, we complement the above with advanced nonlinear techniques: wavelet-based correlation for multiscale dependence, threshold cointegration for asymmetric equilibrium adjustment, and transfer entropy for model-free information flow. This integrated framework allows us to rigorously test the arbitrage relationship between FCPO and P.DCE from linear and non-linear dimensions.

Linear Correlation

After a series of data cleaning, the sample data were determined to be the intraday 5-minute price series of FCPO and P.DCE after adjustment for the same trading time from 2019-01 to 2023-12. We first evaluated the linear correlation between the price series of FCPO and P.DCE.

The linear correlation coefficient is a statistical measure that quantifies the strength and direction of the linear relationship between two variables. The most used linear correlation coefficient is the Pearson Correlation Coefficient, often denoted by the letter "ρ." Its value ranges from -1 to +1， when ρ = +1, it indicates a perfect positive linear relationship between the two variables. When ρ = -1, it indicates a perfect negative linear relationship between the two variables. When ρ = 0, it indicates no linear relationship between the two variables. The formula for the Pearson Correlation Coefficient is shown in Equation (1)(Lee Rodgers & Nicewander, 1988):

$$\rho_{x,y} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}} \qquad (1)$$

Cross-Correlation

Considering only the Pearson correlation coefficient is insufficient to prove the existence of an arbitrage opportunity between the Malaysian FCPO and P.DCE. Testing the cross-correlation between the time series of these two futures prices at various lags is essential for determining the strength of the linkage between the two in different periods. In this study, we employ a sliding window cross-correlation method that can capture the short-term dynamic correlation between the returns of the two futures markets, thereby deeply analysing their short-term interaction characteristics in different periods.

$$\rho(x,y)_t = \frac{\sum_{i=\mathrm{t}}^{t+W-1}(x_i-\overline{x_W})(y_i-\bar{y}\mathrm{W})}{\sqrt{\sum_{i=\mathrm{t}}^{t+W-1}(x_i-\bar{x}\mathrm{W})^2}\sqrt{\sum_{i=\mathrm{t}}^{t+W-1}(y_i-\bar{y}_W)^2}} \qquad (2)$$

The formula for the Cross-Correlation Coefficient is shown in Equation (2), where W is the size of the sliding window, $\rho(x,y)_t$ is the Cross-correlation coefficient between X and Y within the sliding window starting at time t (Brockwell & Davis, 2002). The results of the sliding window cross-correlation typically appear as a time series of correlation coefficients, which reflect the correlation between these two-time series in each time. Positive values indicate a relationship in the same direction, negative values indicate a relationship in the opposite direction, and the closer the value is to 1 or -1, the stronger the relationship; close to 0, there is no obvious linear relationship. By observing the trends and changes in the results, we determined the dynamic correlation between the two and their changing patterns over time.

Stationary Test

To assess the stationarity of the time-series data, we used the Augmented Dickey-Fuller (ADF) test, a commonly employed unit root test. The null hypothesis of the ADF test posits that the time series possesses a unit root, indicating non-stationarity, whereas the alternative hypothesis asserts that the time series is stationary. The assessment entails estimating the subsequent regression in Equation (3)(Dickey & Fuller, 1979):

$$\Delta y_t = a + \beta t + \gamma y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta y_{t-i} + \epsilon_t \qquad (3)$$

The ADF test statistic is derived from the coefficient γ. The null hypothesis posits that γ=0, signifying that the series possesses a unit root and is non-stationary. The alternative hypothesis posits that γ<0, indicating that the series is stationary. The test statistic is compared with the critical values derived from the Dickey-Fuller distribution to determine the existence of a unit root. These critical values are contingent on the number of observations and the selected significance level. If the test statistic is below the critical value, the null hypothesis of a unit root is rejected, signifying that the series is stationary.

Hurst Index

The Hurst index serves as a metric for assessing the long-term memory of time-series data (Hurst, 1951a). It is intricately linked to autocorrelations within the series, and the rate at which these correlations diminish as the lag between the paired values increases. Researchers frequently employ two primary methods, R/S and DFA, to estimate the H-index (Peng et al., 1994). In this study, the input data series are the returns of FCPO and P.DCE, which were proven to be stationary in the last step. The primary advantage of DFA is its ability to remove trend components from time-series data. However, for data that are already stationary, the detrending step becomes redundant and can mislead the Hurst index estimation, causing the estimated value to deviate from the actual value. Therefore, we chose the R/S method to determine the Hurst Index for the two sets of data.

The core steps for calculating the Hurst Index using the R/S analysis method are as follows. First, the time series is divided into multiple non-overlapping sub-intervals, and the rescaled range R/S for each sub-interval is computed, where R is the range of the cumulative deviation series and S is the standard deviation of the sub-interval data. Subsequently, the logarithm is taken for different sub-interval lengths n and their corresponding average R/S values, yielding log(n) and $\log{(\mathrm{R}/S)_n}$. Finally, the Hurst Index H is obtained by fitting a linear regression between $\log{(\mathrm{R}/S)_n}$ and log(n) where the regression slope represents H. This formula is expressed by Equation (4)(Hurst, 1951b):

$$\mathrm{H} = \frac{COV(\log(n), \log(R/S)_n)}{var(\log(n))} \quad (4)$$

where COV denotes the covariance, measuring the linear association strength between log(n) and$\log(\mathrm{R}/S)_n$and var represents the variance, which is used to normalise the covariance. The value of the Hurst Index can be used to interpret the characteristics of the series: H=0.5 indicates a random walk, H>0.5 suggests persistence (trend reinforcement), and H<0.5 implies anti-persistence (mean reversion).

Cointegration Test

Linear Cointegration Test

To investigate the long-run equilibrium relationship between the two-time series, the Engle-Granger two-step cointegration test was applied (Engle & Granger, 1987). Traditionally, the first step, long-term relationship between y_t and x_t is assessed using OLS regression. In this study, we referred to the research of (Basak et al., 2007) and imported machine learning techniques. SVR is a robust machine learning algorithm that specializes in offering more robust and flexible estimates than conventional OLS models. SVR typically uses four types of kernels: linear, polynomial, radial basis function (RBF), and sigmoid (Basak et al., 2007b). In this study, to capture the linear relationships in time-series data, we adopted the SVR method with a linear kernel for regression analysis.

The objective of SVR is to perform a regression analysis on data points by creating a hyperplane in a high-dimensional space (Drucker et al., 1996a; Smola & Schölkopf, 2004a). For a given dataset $\{(x_1,\ y_1),\ \ (x_2,\ y_2), \ldots, (x_n,\ y_n)\}$, where $x_i \in R^d$ represents the feature vector and $x_i \in R$ represents the corresponding target value. SVR aims to find a function $f(x)$ to map the input data to the output values. Specifically, SVR optimises the following objective function for regression in Equation (5)(Vapnik, 2013a):

$$\min_{w,b,\xi_i,\xi_i^*} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) \quad (5)$$

where w is the weight vector of the hyperplane, b is the bias term, ξ_i and ξ_i^* are slack variables for handling regression errors.C is the regularisation parameter that balances the error and model complexity.

To capture the linear relationships in the data, SVR with a linear kernel is suitable for capturing the linear relationship between the two by minimising the regression error and seeking the best-fitting plane (Smola & Schölkopf, 2004b; Vapnik, 2013b). It is more adept at managing short-term market noise and crises and can autonomously filter essential data to enhance fitting outcomes, as opposed to OLS (Drucker et al., 1996b). SVR with a linear kernel is defined in Equation (6):

$$f(x) = (w\ x) + \mathrm{b} \quad (6)$$

where, w is the weight vector of the hyperplane, and b is the bias term. We minimised the regression error and fitted the data with appropriate parameters.

Threshold Cointegration Analysis

To investigate potential asymmetric adjustment dynamics in the long-run equilibrium between FCPO and P.DCE futures prices, we employ a two-regime threshold vector error-correction model (TVECM), following Hansen & Seo (2002). While traditional linear cointegration approaches assume a constant, symmetric speed of adjustment toward equilibrium, they may overlook economically meaningful nonlinearities. The TVECM explicitly accommodates such regime-dependent behavior by allowing the error-correction mechanism to switch based on a threshold in the cointegrating residual. This not only enhances model realism but also improves the detection of state-contingent arbitrage opportunities that linear frameworks would average out or miss entirely. Specifically, we assume a single cointegrating relationship $\beta^{'} X_t$ among the I(1) variables X_t and model the short-run dynamics as switching between two regimes based on whether the lagged cointegrating residual $z_{t-1} = \beta^{'} X_{t-1}$ exceeds an unknown threshold τ. The TVECM is specified as:

$$\Delta X_t = \begin{cases} a_1 z_{t-1} + \sum_{i=1}^{p-1} \Gamma_{1i} \Delta X_{t-i} + \varepsilon_t, if\ z_{t-1} < \tau \\ a_2 z_{t-1} + \sum_{i=1}^{p-1} \Gamma_{2i} \Delta X_{t-i} + \varepsilon_t, if\ z_{t-1} < \tau \end{cases} \quad (7)$$

where $a_1 \neq a_2$permits asymmetric error-correction speeds across regimes. The threshold τ and cointegrating vector β are estimated jointly by maximum likelihood, and the presence of a threshold effect is tested using the sup-LM statistic, with critical values obtained via simulation as the asymptotic distribution is nonstandard. This framework enables us to detect latent regime-dependent equilibria and assess whether economic adjustment intensifies under market stress or macroeconomic shocks.

Wavelet-Based Cointegration Analysis

To examine the co-movement between FCPO and P.DCE futures return across multiple time scales, A wavelet-based cointegration framework that integrates time-domain and frequency-domain analysis is introduced. Unlike traditional cointegration methods, wavelet analysis allows us to decompose the price series into distinct frequency bands, thereby capturing both transient short-term fluctuations and persistent long-term trends simultaneously.

Firstly, we implement the Maximal Overlap Discrete Wavelet Transform (MODWT) on the logarithmic price series of FCPO and P.DCE. MODWT has translation invariance and can deal with time series of any length without the need for down sampling. This decomposition produces wavelet coefficients across various resolution levels, with higher levels associated with lower frequencies (extended time scales) and lower levels reflecting high-frequency (short-term) dynamics. Each level isolates cyclical behavior within a specific time horizon, enabling scale-specific analysis of market linkages.

Second, at each wavelet scale j (j=0,1,2, ...,12), we calculate the Pearson correlation coefficient between the wavelet coefficients of FCPO and P.DCE:

$$\rho_j = Corr(W_j^{FCPO}, W_j^{P.DCE}) \quad (8)$$

where W_j^{FCPO} and $W_j^{P.DCE}$denote the wavelet coefficients of the two series at scale j. A high ρ_j indicates strong co-movement at that time scale, suggesting potential arbitrage opportunities or information transmission within that horizon.

Secondly, we conduct a wavelet energy distribution analysis to quantify the relative contribution of each scale to the total variance of the original series. The energy at scale j is defined as:

$$E_j = \sum_t (W_{j,t})^2 \quad (9)$$

the normalized energy share $\widetilde{E_j} = E_j / \sum_{k=0}^{12} E_k$, reveals which time scales dominate the overall price dynamics. This multiscale approach provides a nuanced view of the FCPO–P.DCE relationship: it identifies not only whether the markets are linked, but at which investment horizons (e.g., intraday, daily, or long-term) the linkage is strongest, thereby informing horizon-specific trading and hedging strategies.

Causality Test

Linear Method: Granger Causality Test

To detect causal relationships between the price series of FCPO and P.DCE, the Granger causality test should be conducted. However, traditional Granger causality tests usually use OLS regression to determine causal relationships between variables. In this study, we adopt conventional Granger Causality (Granger, 1969a) for testing causality between FCPO and P.DCE under a linear relationship.

The conventional Granger Causality Test aims to determine whether the past values of one variable can predict the future values of another variable beyond the information already contained in the past values of the latter (Granger, 1969b). In this test, a vector autoregression (VAR) model is fitted to both price series. The number of lags to be included in the model is determined using information criteria (Lütkepohl, 2005). The joint significance of the lags of one variable in predicting the other was then tested. If the p-value is below the chosen significance level (typically 0.05), the null hypothesis is rejected and it is concluded that there is a Granger causal relationship (Toda & Yamamoto, 1995). The F statistic is defined in Equation (8):

$$F = \frac{RSS_r - RSS_{ur}/q}{RSS_{ur}/(n-k)} \quad (10)$$

where the RSS_r is the residual sum of squares for the restricted model which includes only the lagged values of the independent variable. RSS_{ur} is the residual sum of squares for the unrestricted model which includes the lagged values of both the independent and dependent variables. q is the number of lags being tested in the model. k is the number of observations. k is the number of estimated parameters in the unrestricted model.

Non-Linear Method: Transfer Entropy

To further explore the nonlinear information flow patterns between FCPO and P.DCE, this work used transfer entropy, an information-theoretic tool for detecting causal relationships. Transfer entropy presents considerable advantages over conventional Granger causality tests: it not only detects linear causal relationships but also adeptly recognizes nonlinear information transfer between variables, aligning more closely with the intricate nature of financial time series. Let's assume the transfer entropy from FCPO returns (X) to P.DCE returns (Y), the basic definition of transfer entropy is as follows:

$$TE_{X\to Y} = \sum p(y_{t+1}, y_t^{(k)}, x_t^{(l)}) log \frac{p\left(y_{t+1}\middle|y_t^{(k)}, x_t^{(l)}\right)}{p\left(y_{t+1}\middle|y_t^{(k)}\right)} \quad (11)$$

Where $y_t^{(k)}$ and $x_t^{(l)}$ are the k and l dimensional embedding vectors of past values and the sum is over all joint states. TE intuitively measures the decrease of uncertainty about the future of Y given the past of X, beyond what is already explained by the past of Y itself. A positive TE value indicates a causal information flow from X to Y.

To implement this, the sample data are first discretized into three distinct states ("low", "medium", "high") using a quantile-based discretization strategy with KBins Discretizer, making the data suitable for information-theoretic computation. The analysis is then performed using the Multivariate TE class from the IDTxl toolbox, configured with a lag range of 1 to 5 periods to cover short-term price discovery mechanisms. To ensure the statistical robustness of the observed information flows, the significance of the calculated transfer entropy values is rigorously tested against a null hypothesis of no coupling using a permutation test with 1,000 surrogate datasets. This comprehensive methodology allows for a deeper investigation into the price discovery process and the nonlinear information transmission channels linking these crucial commodity futures. This approach is particularly well-suited for commodity markets, where price dynamics often exhibit nonlinear feedback, regime shifts, and asymmetric spillovers that linear models fail to capture.

Empirical Analysis Result

Linear Correlation

Figure 1.1 shows the price series of FCPO and P.DCE between 2019 and 2023 at the 5-minut level. Both prices were exchanged for Malaysian Ringgit. As can be seen, the trends of the two curves are very similar, reflecting the similar price movements of the two commodities during this period. According to the calculation, the Pearson correlation coefficient between the two-price series is 0.9882, indicating a strong positive linear relationship. The p-value of 0.0 is below the 0.05 significance level, confirming that the correlation was statistically significant. This finding suggests a strong association between the two variables.

Figure 1.1 Price Trends of FCPO and P.DCE in 5-minut Level (2019-2023)
[Source: The author]

The strong positive correlation observed raises questions about the external factors that might be driving the price trends of FCPO and P.DCE in such a parallel manner. Exploring the role of global market dynamics, including fluctuations in crude oil prices, geopolitical tensions, and changes in global commodity demand, is valuable. For instance, Rizal et al. (2023) found that macroeconomic variables such as GDP, interest rates, and inflation significantly influence FCPO prices, underscoring how external factors shape pricing trends (Rizal et al., 2023).

Additionally, it would be beneficial to investigate whether seasonal or cyclical patterns contribute to the observed trends. For instance, certain months might see heightened demand for palm oil due to festivals or global production cycles, which might align with similar patterns in P.DCE pricing. Identifying such patterns could provide a deeper understanding of periodic market fluctuations. Policy changes, such as export restrictions or tariffs, could also have played a significant role in shaping the prices of FCPO and P.DCE. The study *of Technical Strategy on FCPO* supports this by highlighting how government policies and trade regulations have historically impacted FCPO pricing.

Although the high correlation suggests that both commodities move similarly, it is important to question whether this relationship extends beyond linearity. Non-linear models, such as those involving lag effects or volatility spillovers, could reveal more complex dynamics between FCPO and P.DCE prices. Exploring the robustness of the correlation using advanced statistical methods would provide additional credibility to these findings.

Comparisons with related commodities, such as soybean oil or crude oil, could also offer interesting perspectives. According to the KGI Securities, DCE RBD Palm Olein futures are closely connected to the palm oil physical market, and their prices exhibit strong correlations with FCPO prices. This implies that broader trends in global commodity markets might influence the price series of FCPO and P.DCE.

Finally, the strong correlation observed has practical implications for market participants. Investors and traders can use this information to develop hedging strategies or enhance their portfolio diversification. For example, recognising the synchronised movements of FCPO and P.DCE could help in predicting price behaviour during volatile periods, potentially leading to better decision-making in trading.

Cross-Correlation

To represent the short-term dynamic correlation between the returns of the two futures markets, it is necessary to consider cross-correlation. In this study, a sliding window method was used to evaluate the cross-correlation between the price series of FCPO and P.DCE. The window size is set to 50 (1.16 days) because there are up to 43 5-minute prices per day in the research sample. Thus, it is reasonable to capture short-term price fluctuations and dynamic correlation changes in high-frequency arbitrage.

Figure 1.2 illustrates the dynamic cross-correlation between the price series of the FCPO and DCE with a 50-sliding window. The correlation coefficient ranged from -1 to 1, with a mean of 0.5532, median of 0.5900, and mode of 0.63, indicating a mainly positive relationship. Although fluctuations occur, the correlation generally remains positive with periodic deviations from the usual tendency. These movements may indicate possible arbitrage opportunities and reveal instances of mispricing in both markets.

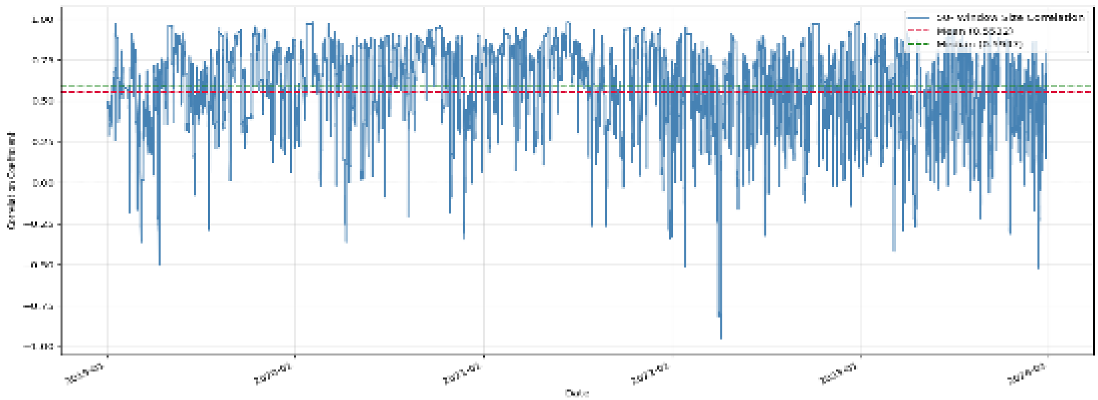

Figure 1.2 Dynamic Cross-Correlation Between FCPO and DCE Returns
[Source: The author]

Stationary Test

The stationarity of a time series is fundamental to time-series analysis. Hence, it is necessary to assess stationarity using a unit root test. This study employs the primary unit root testing methodology: the Augmented Dickey-Fuller (ADF) test to evaluate the price series and the first-order differenced price series of FCPO and P.DCE, respectively.

Table 1.2
ADF Test Result

variable	**ADF statistic**	**p-value**	**1% standard value**	**5% standard value**	**10% standard value**	**Judgement**
FCPO_CLOSE	-1.8184	0.3713	-3.431	-2.862	-2.567	Not Stationary
P.DCE_CLOSE	-1.767	0.3969	-3.431	-2.862	-2.567	Not Stationary
ΔFCPO_CLOSE	-31.1266	0.000	-3.431	-2.862	-2.567	Stationary
ΔP.DCE_CLOSE	-171.2892	0.000	-3.431	-2.862	-2.567	Stationary

[Source: The author]

Table 1.2 displays the outcomes of the ADF unit root test for the FCPO and P.DCE data. The ADF statistics for FCPO and P.DCE are -1.8184 and -1.7670, respectively, with p-values of 0.3713 and 0.3969, respectively. As both p-values exceed the standard significance thresholds (1%, 5%, or 10%), the null hypothesis cannot be rejected, suggesting that these two variables are nonstationary at the level. The ADF statistics for the differenced variables ΔFCPO and ΔP.DCE were – -31.1266 and -171.2892, respectively, with p-values of 0.000. The substantially lower p-values compared to the crucial values at the significance levels show that the differenced variables exhibit stationarity. In conclusion, the price series of FCPO and P.DCE are both I (1) sequences, which indicates that the data have a long-term trend component. Therefore, the next step is to further analyse the long-term dependence of the data using the Hurst Index.

Hurst Index

After the unit root test, it has been approved the differenced price series of FCPO and P.DCE were found to be stationary. Therefore, the R/S method was selected to calculate the

Hurst Index. For already stationary time-series data, the R/S method is generally more appropriate for calculating the Hurst exponent. This is because the DFA method is more suited to nonstationary data with trends because it can extract long-memory characteristics by detrending. However, its performance with stationary data is not superior to that of the R/S method.

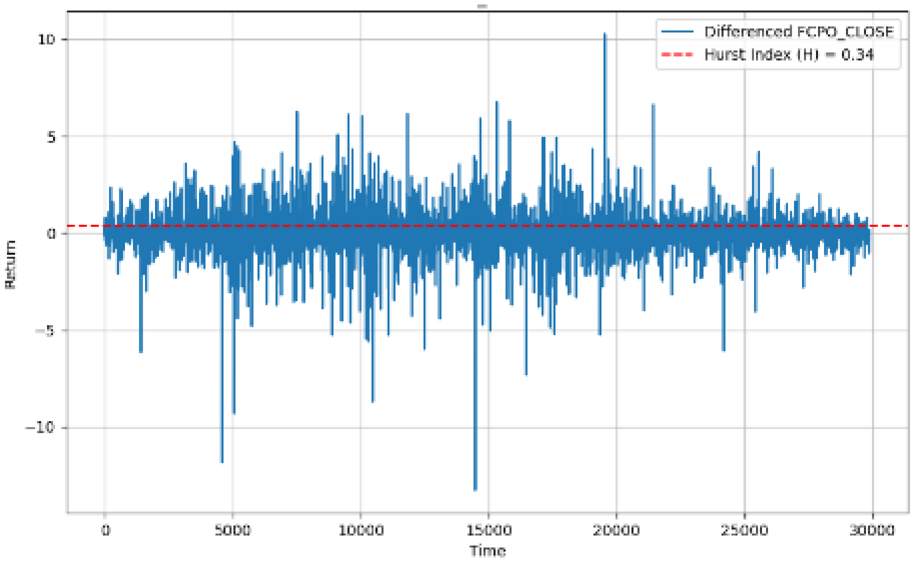

Figure 1.3 Differenced FCPO Price Series with Hurst Index
[Source: The author]

Figure 1.3 and 1.4 indicate that the Hurst indices of the differenced price series of FCPO and P.DCE are 0.34 and 0.35, respectively, both significantly lower than 0.5, indicating that the prices of these two markets exhibit long-term anti-persistence, which is consistent with the mean reversion characteristic. Specifically, in the short term, prices tend to revert to the mean, meaning that after large price fluctuations, the market is likely to return to its average level. Therefore, these markets may exhibit cyclical reversal patterns, making them suitable for mean reversion strategies. That is, when the price deviates significantly from the mean, it is expected to revert to equilibrium.

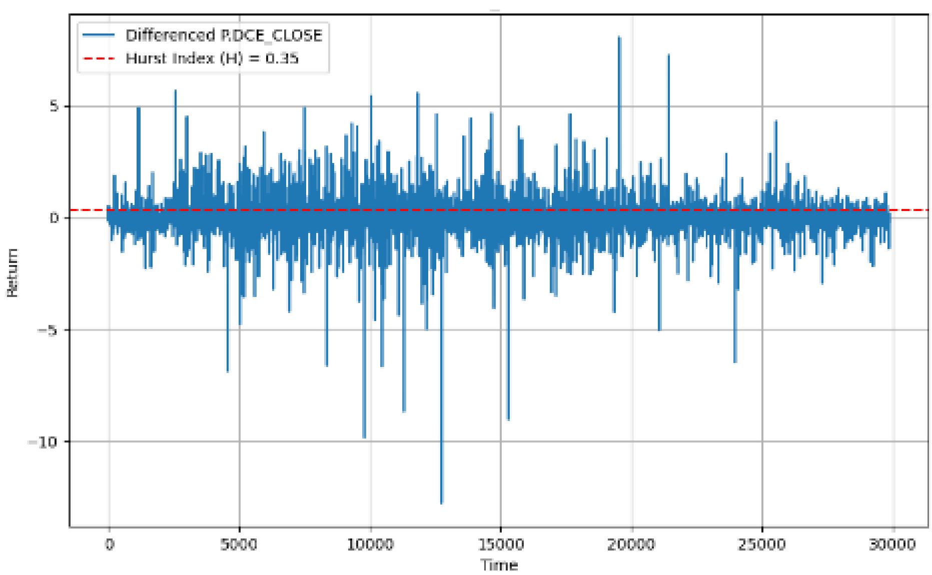

Figure 1.4 Differenced P.DCE Price Series with Hurst Index

[Source: The author]

Furthermore, owing to the relatively low and similar Hurst exponents, the price fluctuations in these two markets exhibit strong short-term anti-persistence with a lack of long-term trend continuation, suggesting higher market volatility. This volatility creates potential opportunities for cross-market arbitrage. When there are large short-term price fluctuations with mean reversion characteristics, arbitrageurs can exploit the short-term price differences between markets for arbitrage.

Cointegration Test

Linear Cointegration Test

The previous section demonstrated that the price series of FCPO and P.DCE are not stationary. Therefore, it is essential to conduct a cointegration test to determine whether a long-term equilibrium relationship exists between these two nonstationary series. Cointegration indicates that short-term price deviations are temporary and that prices will eventually revert to equilibrium. This is crucial for understanding the price transmission mechanism between these markets and confirming the existence of potential arbitrage opportunities based on the law of one-price theory. If the cointegration test confirms a significant relationship, it implies that cross-market arbitrage strategies could be profitable. Furthermore, it has challenged the EMH.

By conducting a cointegration test, the Engle-Granger two-step method based on a machine learning algorithm was applied in this study. In the first step, SVR with linear and RBF kernels is applied to perform a regression on these two-time series, estimating the long-term equilibrium relationship between them. SVR was chosen because of its robustness in handling both linear and nonlinear interactions that traditional methods, such as OLS, might miss. Once the model was trained, the residuals were computed to represent the difference between the actual values and predicted equilibrium relationship. In the second step, the ADF test was applied to the residuals to determine whether they were stationary.

In the first step, we set the price series of P.DCE and FCPO as x and y, then work out the logarithm of them. SVR with a linear kernel is imported first to fit the differenced log-transformed series, and the residuals are calculated. The dataset was divided into training (80%) and testing (20 %) sets. In the model setting section, we use Grid Search CV to perform 5-fold cross validation to find the optimal hyperparameters, which are demonstrated to be ‘C:100’, ‘epsilon:0.1’. The output of the SVR with a linear kernel model is as follows: the intercept is (-1.0169), and the regression coefficient is 1.09283519, which means that for every 1 unit increase in the log price of FCPO, the price will increase by approximately 1.0928 units. The R^2 score is 0.9817, which means that 98.17% of the variability in the log price of the FCPO can be explained by the model, indicating that the model fits the data very well. The MSE was 0.0017, which is relatively small, indicating that the prediction error of the model was very low and that the model fit the data well. Figure 1.5 and 1.6 show the regression results and residual distribution of the SVR model, respectively.

From Figure 1.5, it can be observed that the fit between the predicted and actual values was excellent. The red line closely followed the trend of the blue dots, indicating that the model predictions were accurate. As shown in Figure 1.6, the residuals were evenly distributed with no obvious patterns, suggesting that the model did not exhibit any systematic bias. In conclusion, SVR with the linear kernel model performed excellently, with a very good fit, minimal prediction error, and no systematic bias. The regression coefficient and residual analysis demonstrate that the log piece of P.DCE has a strong influence on predicting the log price of FCPO, and the overall model is stable and accurate.

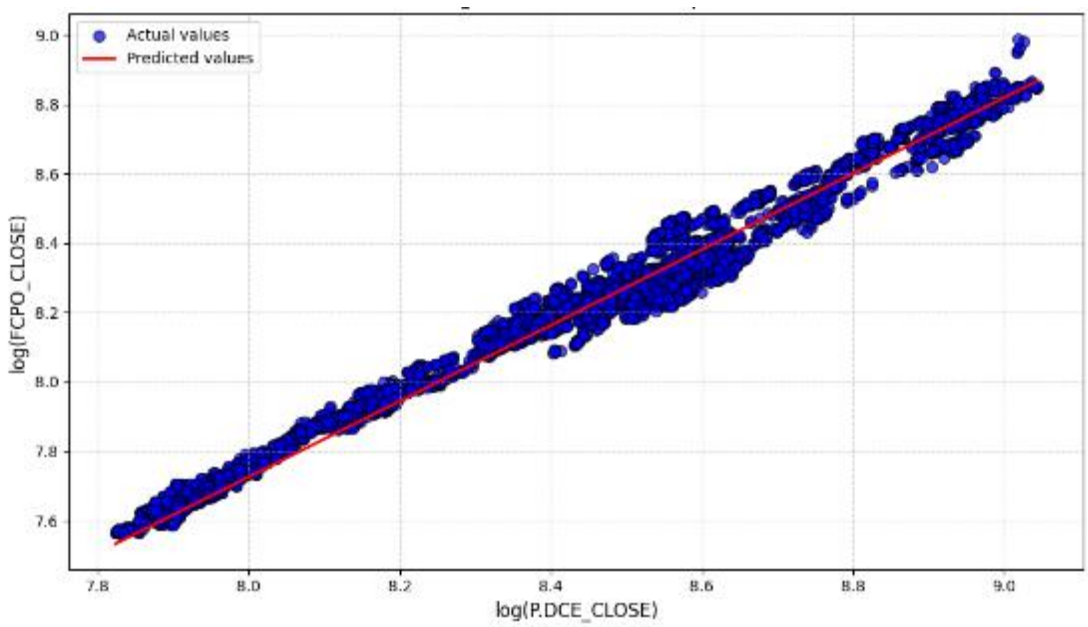

Figure 1.5 The Result of SVR with Linear Kernel Model
[Source: The author]

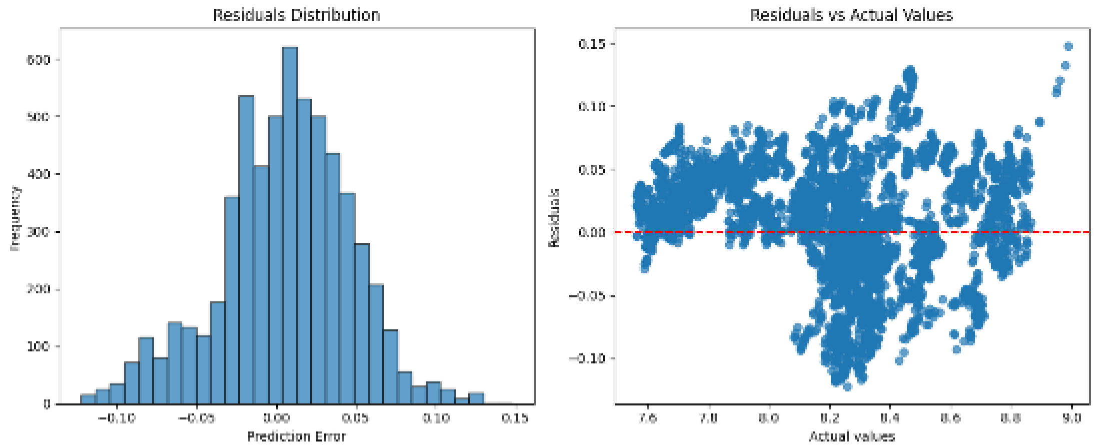

Figure 1.6 The Residual Analysis of SVR with Linear Kernel Model
[Source: The author]

Non-Linear Cointegration Test ：TVECM

To detect the non-linear cointegration relationship between FCPO and P.DCE, the TVECM framework was employed to investigate whether the error correction mechanism between FCPO and P.DCE futures return exhibits threshold effects. The long-term cointegration relationship was first established using the Johansen cointegration test, yielding a cointegration vector of beta [3.940111,1 -4.53803068], The error correction term (ECT) was constructed from this relationship, representing deviations from the long-run equilibrium.

The threshold value was identified as the median of the ECT ($\tau = 0$), dividing the data into two distinct regimes:
Regime 1 (Undervalued State) : $\tau \leq 0$, indicating prices are below equilibrium
Regime 2 (Overvalued State): $\tau \geq 0$, indicating prices are above equilibrium
OLS regression was conducted separately for each regime to estimate the speed and nature of adjustment back to equilibrium.

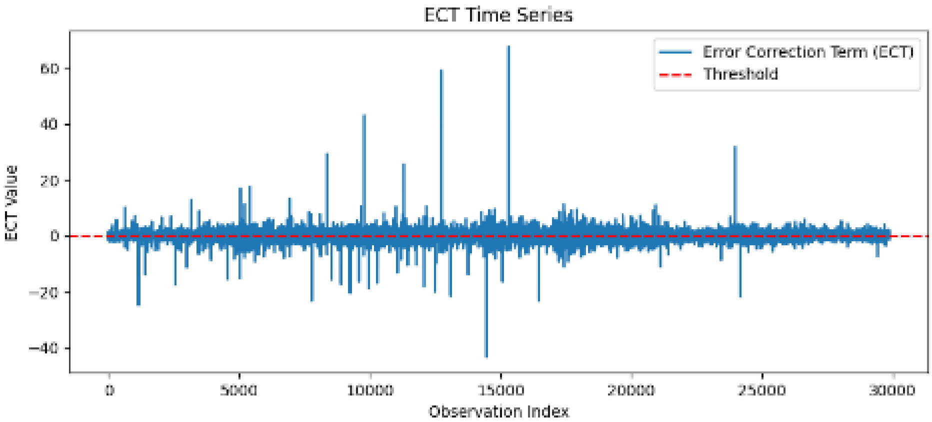

Figure 1.7 Asymmetric Adjustment Dynamics Around the Cointegration Equilibrium

Figure 1.7 depicts the blue curve, which signifies the error correction term (ECT) in the cointegration relationship, assessing the deviation of FCPO_return and P.DCE_return from their long-term equilibrium. The red dashed line denotes our chosen threshold (Threshold = 0), signifying the arbitrage frontier or the transition point. The ECT fluctuates around the threshold, confirming frequent mean-reverting behavior. Extreme deviations are occasional, supporting the stability of the identified cointegration relationship.

Table 1.3
Estimation Results of the TVECM

State	Condition	sample size	Constant Term	Lag Coefficient	FCPO Lag P.DCE Coefficient	R^2
Regime 1（undervalued）	ECT ≤ 0	15,117	-0.1001	-0.9227	-0.0621	0.480
Regime 2（overvalued）	ECT > 0	14,747	+0.1131	-0.9402	-0.0369	0.483

[Source: The author]

As can be seen from Table 1.3, the regression results reveal a statistically significant asymmetric adjustment mechanism. In regime 1 (undervaluation), the regression model yields a significantly negative constant term (−0.1001) and a strong negative autoregressive coefficient for FCPO returns (−0.9227), indicating rapid downward correction and pronounced responsiveness to its own past movements. Moreover, lagged P.DCE returns exert a statistically significant negative influence (−0.0621), suggesting active cross-market feedback during periods when FCPO is relatively cheap. Conversely, in Regime 2 (Overvaluation), the constant term becomes positive (+0.1131), reflecting persistent upward price momentum despite deviation from equilibrium. Although the autoregressive coefficient remains negative (−0.9402), the impact of P.DCE returns weakens (−0.0369), implying reduced sensitivity to the

Chinese market when FCPO is overpriced. These findings demonstrate that the market corrections are faster and more arbitrage-driven when FCPO is undervalued, whereas overvaluation tends to be sustained—potentially due to speculative behavior or structural frictions.

In sum, the relationship between FCPO and P.DCE is characterized by a nonlinear cointegration structure with state-dependent dynamics. This has important implications for arbitrageurs, risk managers, and policymakers: effective strategies must account for the prevailing market regime, as the speed and direction of price convergence are contingent on whether the market is in a state of overvaluation or undervaluation.

Non-Linear Cointegration Test: Wavelet-Based Cointegration Analysis

To better understand the time-scale-specific relationship between FCPO and P.DCE futures, we apply the MODWT to decompose the log price of this pair of futures into 13 scales (Level 0–12) and compute wavelet correlation and energy at each scale. The findings are exhibited in Table 1.4.

Table 1.4
Across Scales Wavelet Correlation Coefficients

Level	**Correlation**	**Level**	**Correlation**
Lv. 0	0.8373	Lv. 7	0.7746
Lv. 1	0.3782	Lv. 8	0.7902
Lv. 2	0.7405	Lv. 9	0.7890
Lv. 3	0.8646	Lv. 10	0.7409
Lv. 4	0.7884	Lv. 11	0.6454
Lv. 5	0.7708	Lv. 12	0.5160
Lv. 6	0.8090	-	-

The correlation coefficients across wavelet scales reveal a pronounced heterogeneity in market co-movement depending on the investment horizon. Short-term scales (Levels 0–3) show very high correlation, peaking at 0.8646 (Level 3), indicating tight intraday co-movement. Medium-term scales (Levels 4–9) maintain stable correlation (>0.77). This stable co-movement suggests that fundamental information is efficiently transmitted between the two markets over these horizons. Long-term scales (Levels 10–12) exhibit declining correlation, falling to 0.5160 at Level 12. This attenuation implies that long-run price drivers differ structurally between the Malaysian and Chinese palm oil markets.

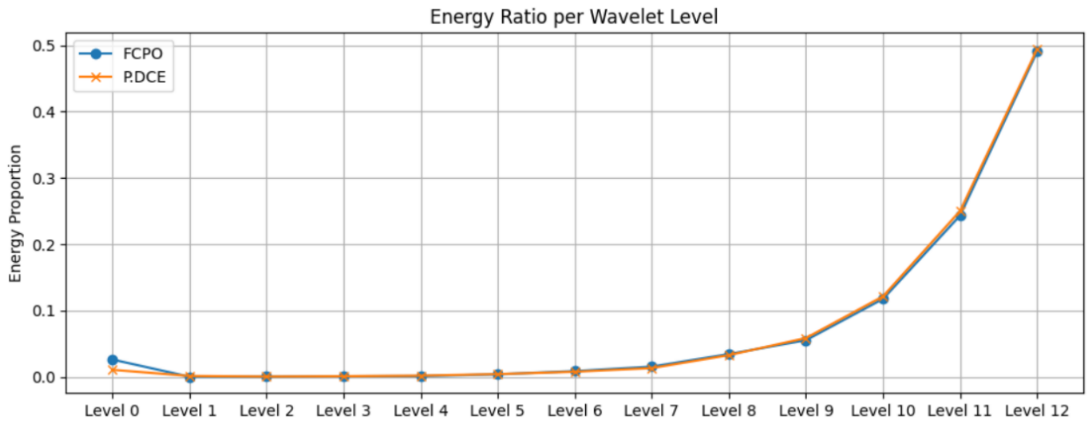

Figure 1.8 Wavelet Energy Distribution Between FCPO and P.DCE

Complementing the correlation analysis, the wavelet energy distribution sheds light on where the dominant price fluctuations originate. As shown in Figure 1.7, the majority of total energy is concentrated in the lowest-frequency components: nearly 50% of total wavelet energy is concentrated in Level 12, with Levels 10–12 dominating overall. In contrast, high-frequency components (Levels 1–5) contribute little energy. This pattern indicates that long-term trends constitute the primary source of price variation in both markets, characterizing them as “trend-driven” rather than “noise-driven.” Despite this, the relatively weaker correlation at these dominant long-term scales (Levels 10–12) reveals a critical insight: although both markets are shaped by persistent trends, the nature of these trends is not identical.

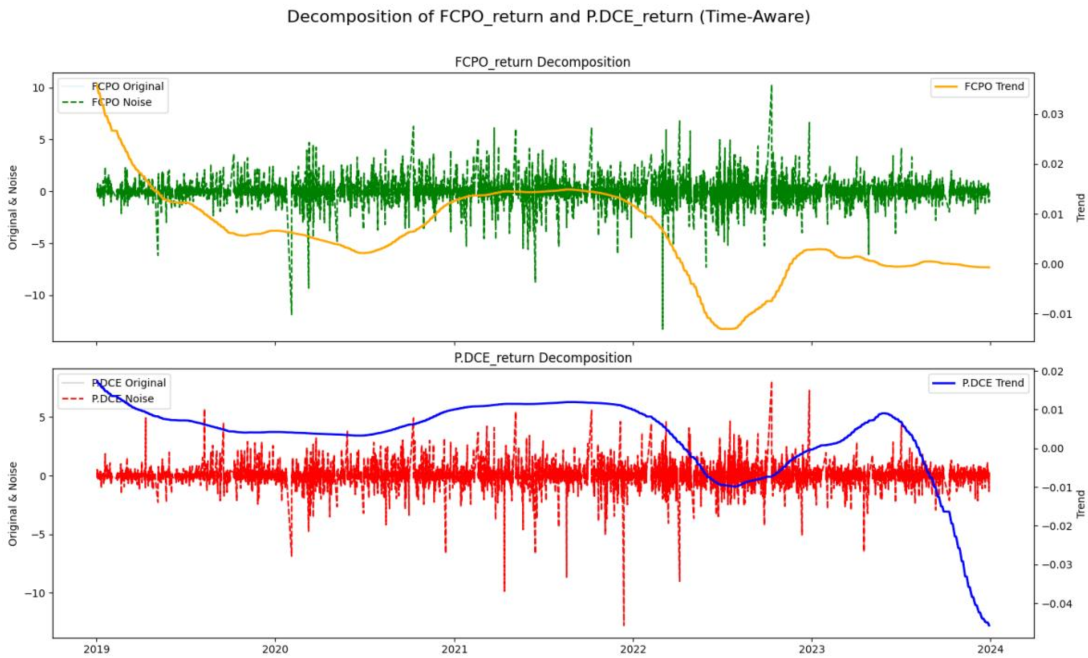

Figure 1.9 Trend and Noise Components of FCPO and P.DCE Returns

Figure 1.8 shows the long-term trends based on the return of FCPO and P.DCE. Comparing trends, the two markets exhibit different long-term trends, particularly in the amplitude and frequency of the noise component. This may indicate opportunities for cross-market arbitrage, especially during periods of high short-term volatility.

From a financial perspective, the combination of high energy but low correlation at long horizons versus lower energy but high correlation at short-to-medium horizons provides important implications: the strong co-movement at Levels 0 - 9 supports the viability of short-term arbitrage, indicating frequent opportunities particularly during periods of high volatility. While the decoupling observed at Levels 10–12 points to structural long-term divergence driven by fundamental market; furthermore, market efficiency appears to be scale-dependent, as information is rapidly incorporated at short horizons (evidenced by high correlation), yet idiosyncratic long-term factors dominate the equilibrium relationship, collectively underscoring the limitations of single-scale analysis conventional cointegration tests may infer a “strong linkage” based on aggregate correlation but fail to capture the critical nuance that the nature of the FCPO–P.DCE relationship evolves dramatically across time scales.

Causality Test

Linear Causality: Conventional Granger Causality Test

To detect the linear causal relationships between the price series of FCPO and P.DCE, this study applies the conventional Granger Causality test to detect causal relationships. For the conventional Granger Causality test, the maximum lag period is set to 12 because we try to explore the short-term causality between the two futures. Since our sample data are at the 5-minute’ level, the past 1-hour trading time is tested for influence on the current period. We iterate through lag periods from 1 to 12 to determine the lag with the smallest p-value and set 0.05 as the significance threshold to suggest a significant causal relationship between the two.

Figure 1.8 shows the test results. For direction, FCPO> P. DCE, the p-values remain close to 0 and are well below 0.05, across all 12 lags. This indicates that the price of FCPO is a strong Granger cause of P.DCE, where the price of FCPO significantly influences the price of P.DCE.

For the direction of P.DCE > FCPO, the p-values significantly exceed 0.05 in most lags, and only lag 2 gives a low p-value. This indicates that P.DCE is generally not a Granger cause of the FCPO. In conclusion, the FCPO futures market may lead the P.DCE futures market. While the price of P.DCE may have some short-term influence on the price of FCPO (at lag 2), its overall impact is weak.

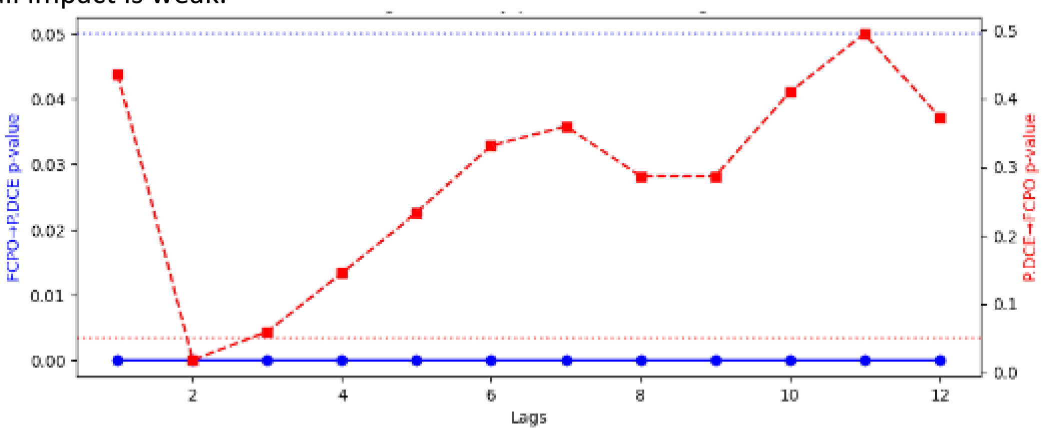

Figure 1.10 Conventional Granger Causality Test: p-values Across Different Lags
[Source: The author]

Non-Linear Causality: Transfer Entropy

To uncover the direction and strength of nonlinear information transmission between the Malaysian and Chinese palm oil futures markets, we estimate the transfer entropy (TE) in both directions from FCPO returns to P.DCE returns and vice versa and assess statistical significance via a permutation test with 1,000 replications.

The results reveal a highly asymmetric causal relationship. The transfer entropy from FCPO to P.DCE is 0.1790, and this value is statistically significant at the 5% level under the permutation test. This indicates that past returns in the Malaysian FCPO market contain substantial predictive information about future movements in the Chinese P.DCE market, beyond what can be inferred from P.DCE's own history. In other words, FCPO exhibits strong leading power over P.DCE, suggesting that price discovery in the global palm oil complex is primarily driven by the Malaysian benchmark. In stark contrast, the transfer entropy in the reverse direction from P.DCE to FCPO is only 0.0012, not statistically significant. This implies that Chinese market returns provide virtually no incremental predictive power for future FCPO movements. Despite the growing size and liquidity of the Dalian palm oil futures market, it appears to function largely as a price follower rather than a price leader in the bilateral information flow.

This asymmetry underscores are same with the linear causality result, therefore dominant role of the Malaysian market in the palm oil pricing ecosystem in both linear and non-linear perspectives. The larger magnitude of TE from FCPO to P.DCE reflects a stronger and more persistent causal information flow, consistent with Malaysia's status as the world's second-largest palm oil producer and the FCPO contract's long-standing role as the global reference price. Meanwhile, the negligible reverse TE suggests limited feedback from China possibly due to capital controls, trading restrictions, or the domestic orientation of Chinese speculative activity.

Significance of the Study

Understanding the relationship between Malaysian and Chinese palm oil futures is crucial for market participants, including traders, investors, and policymakers. Palm oil is one of the most widely traded commodities globally, and price movements in major markets, such as Malaysia and China, significantly impact global supply chains. As Malaysia is the world's second-largest palm oil producer and China is one of the largest importers and consumers, the linkage between these two markets directly influences price stability, trade flows, and risk management strategies. By examining this relationship using high-frequency data, this study provides a more detailed and accurate analysis of price transmission mechanisms, helping market participants make informed decisions.

For traders and institutional investors, identifying arbitrage opportunities between Malaysian and Chinese palm oil futures enhances trading strategies and profitability. HFT and algorithmic trading rely on real-time data to exploit price inefficiencies that exist only for brief periods. Understanding how prices move between BMD and DCE can provide insights into when and where arbitrage opportunities arise in the market. By leveraging high-frequency data, traders can execute precise and timely trades, reducing their exposure to market risk, while maximizing returns.

For investors and hedgers, such as plantation companies and food manufacturers, understanding the price dynamics between the Malaysian and Chinese futures markets is essential for effective risk management. As price fluctuations in one market can influence the other, companies exposed to palm oil price volatility can use futures contracts to hedge against adverse price movements. A better understanding of market efficiency and price transmission between these markets allows businesses to develop more effective hedging strategies, protecting them from sudden price swings that could affect their profitability and cost structure.

Policymakers and regulators can also benefit from this research by gaining insights into market efficiency, stability, and transparency. If price movements in one market consistently lead to the other, this suggests that the information flow is asymmetrical, which may indicate the need for better market regulations to ensure fair trading conditions. Additionally, understanding cross-market interactions helps policymakers anticipate the effects of trade policies, import/export regulations, and macroeconomic factors on palm oil prices. This is particularly important for Malaysia and China, as both countries play crucial roles in shaping the global palm oil industry.

From a broader economic and strategic perspective, palm oil is a vital agricultural commodity that contributes significantly to the economies of Malaysia and China. A well-functioning and efficient futures market supports economic stability by ensuring fair price discovery and reducing uncertainty for market participants. By shedding light on how these two major markets interact, this study provides valuable knowledge that can improve market transparency, enhance price stability, and support sustainable economic growth in the palm oil industry.

Conclusion

This study investigates the price dynamics and arbitrage opportunities between Malaysia's Crude Palm Oil Futures (FCPO) and China's Palm Oil Futures (P.DCE) listed on the Dalian Commodity Exchange, using high-frequency 5-minute intraday data from 2019 to 2023. By integrating traditional econometric techniques with machine learning and nonlinear information-theoretic approaches, our analysis provides a multifaceted understanding of market efficiency, interdependence, and cross-market arbitrage potential in one of the world's most strategically important agricultural commodity markets.

Our empirical results reveal a remarkably strong linear correlation (Pearson's $r = 0.9882$) between FCPO and P.DCE return series, indicating a high degree of co-movement even at the intraday level. This suggests that information is rapidly transmitted between the two markets, consistent with the law of one price in an increasingly integrated global commodity landscape. However, the presence of short-term deviations from this equilibrium captured through cointegration residuals, wavelet-based multiscale analysis, and threshold error-correction dynamics confirms the existence of transient inefficiencies. These inefficiencies, though fleeting, create statistically detectable arbitrage windows that are particularly exploitable in a high-frequency trading environment. The Hurst index analysis further indicates that both markets exhibit mean-reverting behavior ($H < 0.5$) in their return series, implying that price deviations tend to correct themselves over short horizons—a favorable condition for statistical arbitrage strategies such as pairs trading. Moreover, the application of Support

Vector Regression (SVR) with a linear kernel enhances the robustness of the long-run equilibrium estimation, demonstrating the added value of machine learning in modeling noisy financial time series. Nonlinear analyses, including transfer entropy and threshold cointegration, uncover asymmetric adjustment mechanisms and directional information flows, with evidence suggesting that FCPO continues to play a leading role in price discovery, although P.DCE's influence has grown notably in recent years.

From a practical standpoint, these findings offer actionable insights for traders, risk managers, and algorithmic trading firms. The identified arbitrage opportunities though narrow in duration can be systematically captured using high-frequency data and adaptive trading algorithms. For policymakers, the strong cross-market linkage underscores the need for coordinated regulatory oversight to mitigate systemic risks arising from speculative activity or liquidity shocks in either market. Furthermore, the integration of machine learning and high-frequency analytics into commodity market monitoring could enhance early-warning systems for market manipulation or excessive volatility. This study also contributes to the academic literature by bridging a critical gap in cross-market palm oil futures research. While prior work relied predominantly on low-frequency daily data, our use of 5-minute observations enables a granular examination of intraday arbitrage dynamics, aligning with the realities of modern electronic trading. Future research could extend this framework to include other related commodities, incorporate macroeconomic news events or ESG-related shocks, or explore the impact of exchange rate volatility and trade policy changes on cross-market efficiency.

In sum, while the FCPO and P.DCE markets are largely efficient in the long run evidenced by stable cointegrating relationship and rapid error correction short-term inefficiencies persist due to structural differences in trading mechanisms, liquidity, and regulatory environments. These findings invite a nuanced interpretation of the Efficient Market Hypothesis (EMH) in a cross-border context. Although each market may individually satisfy weak-form efficiency over extended horizons, the predictable, mean-reverting deviations between them reveal a form of joint market inefficiency at high frequencies. This cross-market anomaly arises not from irrational behavior, but from institutional frictions that impede instantaneous arbitrage, thereby allowing transient mispricing to emerge. Consequently, the presence of statistically detectable arbitrage windows does not necessarily contradict EMH in its strictest form, but rather highlights its limitations when applied to interconnected yet institutionally heterogeneous markets. These inefficiencies, when detected with high-frequency tools and advanced analytics, present viable arbitrage opportunities, reinforcing the importance of technological sophistication in today's global commodity trading landscape.

Building on these findings, future research will focus on developing and optimizing cross-

market quantitative arbitrage strategies. Emphasis will be placed on designing dynamic threshold statistical arbitrage models, leveraging deep learning architectures to capture nonlinear patterns in high-frequency price sequences, and incorporating asymmetric adjustment mechanisms based on transfer entropy. Concurrently, the study will validate strategy performance through historical high-frequency data back-testing, quantify profitability using transaction cost models.

Declaration of Competing Interest
The authors declare that they have no conflict of interest.

Acknowledgement
The authors would like to acknowledge the University Teknologi Malaysia for the study

Credit Authorship Contribution Statement
All authors critically revised the intellectual content, read, and approved the final manuscript, and agreed to be accountable for all aspects of the work, ensuring that questions related to the accuracy or integrity of any part of the article were appropriately investigated and resolved.

Funding
This research did not receive any specific grants from funding agencies, including public, commercial, or not-for-profit organisation.

Informend Consent
Informed consent is not applicable

Ethical Approval
The research conducted is not related either human and animal use.

References
Ahmad, M. Y., Wong, K. K. S., & Kamarulzaman, N. H. (2022). Selected Factors Influencing China's Palm Oil Import Demand from Malaysia. *AGRARIS: Journal of Agribusiness and Rural Development Research*, *8*(2), 261–274.

Ahmad, N. (2010). Impact of Automated Trading in the Crude Palm Oil Futures Market on its Underlying Spot Market. *International Research Journal of Finance and Economics*, *36*, 46–58.

Ahmed, K., Zain, M., Noor, A. M., Alsaadi, A. R., Milhem, M., & Ahmad, A. (2020). The impact of macroeconomic variables on commodity futures prices: An evidence from Malaysian crude palm oil futures. *European Journal of Molecular & Clinical Medicine*, *7*(06), 2095–2105.

Barboza Martignone, G., Behrendt, K., & Paparas, D. (2022). Price transmission analysis of the international soybean market in a trade war context. *Economies*, *10*(8), 203.

Basak, D., Pal, S., & Patranabis, D. C. (2007a). Support vector regression. *Neural Information Processing-Letters and Reviews*, *11*(10), 203–224.

Basak, D., Pal, S., & Patranabis, D. C. (2007b). Support vector regression. *Neural Information Processing-Letters and Reviews*, *11*(10), 203–224.

Brockwell, P. J., & Davis, R. A. (2002). *Introduction to time series and forecasting*. Springer.

Bui, Q., & Ślepaczuk, R. (2022). Applying Hurst Exponent in pair trading strategies on Nasdaq 100 index. *Physica A: Statistical Mechanics and Its Applications*, *592*, 126784.

Bursa Malaysia. (2019). *Crude palm oil futures (FCPO) frequently asked questions*. https://www.bursamalaysia.com/sites/5d809dcf39fba22790cad230/assets/5e0aefb539fba27b8a3efba2/FCPO_Physical_Delivery_FAQ_Dec2019.pdf

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, *74*(366a), 427–431.

Drucker, H., Burges, C. J., Kaufman, L., Smola, A., & Vapnik, V. (1996a). Support vector regression machines. *Advances in Neural Information Processing Systems*, *9*.

Drucker, H., Burges, C. J., Kaufman, L., Smola, A., & Vapnik, V. (1996b). Support vector regression machines. *Advances in Neural Information Processing Systems*, *9*.

Du, J. (2018). Empirical differences between the overnight and day trading hour returns: Evidence from the Chinese commodity futures. *China Finance Review International*, *8*(3), 315–331.

Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: Journal of the Econometric Society*, 251–276.

(FAO), F. and A. O. of the U. N. (2022). *World Agriculture: Towards 2030/2050*.

Fassas, A. P. (2011). Mispricing in stock index futures markets-The case of Greece. *Investment Management and Financial Innovations*, *8, Iss. 2*, 103–109.

Go, Y.-H., & Lau, W.-Y. (2024). Terms of trade or market power? Further evidence from dynamic spillovers in return and volatility between Malaysian crude palm oil and foreign exchange markets. *The North American Journal of Economics and Finance*, *73*, 102178.

Granger, C. W. J. (1969a). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, 424–438.

Granger, C. W. J. (1969b). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, 424–438.

Hansen, B. E., & Seo, B. (2002). Testing for two-regime threshold cointegration in vector error-correction models. *Journal of Econometrics*, *110*(2), 293–318. https://doi.org/10.1016/S0304-4076(02)00097-0

Hu, X., Zhu, B., Zhang, B., & Zeng, L. (2024). Extreme risk spillovers between US and Chinese agricultural futures markets in crises: A dependence-switching copula-CoVaR model. *Plos One*, *19*(3), e0299237.

Hurst, H. E. (1951a). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, *116*(1), 770–799.

Hurst, H. E. (1951b). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, *116*(1), 770–799.

Kallandranis, C., Dimitriou, D., & Tsioutsios, A. (2024). On the contagion effect between crude oil and agricultural commodity markets: a dynamic conditional correlation and spectral analysis. *Journal of Energy Markets*.

Kenanga Futures. (2023). *Leading the Charge*.

Lee Rodgers, J., & Nicewander, W. A. (1988). Thirteen ways to look at the correlation coefficient. *The American Statistician*, *42*(1), 59–66.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer Science & Business Media.

Malaysian Palm Oil Council. (2023a). *BMD on defensive mode*. https://www.mpoc.org.my/bmd-on-defensive-mode/

Malaysian Palm Oil Council. (2023b). *Palm Oil Industry in Malaysia*. www.mpoc.org.my

Miao, G. J. (2014). High frequency and dynamic pairs trading based on statistical arbitrage using a two-stage correlation and cointegration approach. *International Journal of Economics and Finance*, *6*(3), 96–110.

Peng, C.-K., Buldyrev, S. V, Havlin, S., Simons, M., Stanley, H. E., & Goldberger, A. L. (1994). Mosaic organization of DNA nucleotides. *Physical Review e*, *49*(2), 1685.

Poutré, C., Dionne, G., & Yergeau, G. (2023). International high-frequency arbitrage for cross-listed stocks. *International Review of Financial Analysis*, *89*, 102777.

Rizal, N., Saifuddin, S. A., Abd Rahim, S. H. A., Mohd Nazri, N., Ab Aziz, M. S., & Zainoddin, A. I. (2023). The Macroeconomic Factors on Malaysia's Future Crude Palm Oil (FCPO). *International Journal of Academic Research in Business and Social Sciences*, *13*(3).

Sewell, M. (2012). The efficient market hypothesis: Empirical evidence. *International Journal of Statistics and Probability*, *1*(2), 164–178.

Smola, A. J., & Schölkopf, B. (2004a). A tutorial on support vector regression. *Statistics and Computing*, *14*, 199–222.

Smola, A. J., & Schölkopf, B. (2004b). A tutorial on support vector regression. *Statistics and Computing*, *14*, 199–222.

Tandra, H., Suroso, A. I., Syaukat, Y., & Najib, M. (2022). The Determinants of Competitiveness in Global Palm Oil Trade. In *Economies* (Vol. 10, Issue 6). https://doi.org/10.3390/economies10060132

Toda, H. Y., & Yamamoto, T. (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics*, *66*(1–2), 225–250.

Vapnik, V. (2013a). *The nature of statistical learning theory*. Springer science & business media.

Vapnik, V. (2013b). *The nature of statistical learning theory*. Springer science & business media.

WOODE, J. K. (2024). *Commodities, Exchange Rates, And Equity Markets in Commodity-Dependent Sub-Saharan African Countries*. University of Cape Coast.

World Bank. (2023). *The World Bank and Palm Oil*. World Bank Publications.