Integrating Macroeconomic Indicators and U.S. Commodity Futures Prices for Pair Trading on Chinese Commodity Futures

Xinyu Lin [1,2], Guanrou Deng [1,2], Li Zhang [1,2]*

[1]Institute of Finance Technology, University College London, London, United Kingdom
[2]Civil, Environmental and Geomatic Engineering, University College London, London, United Kingdom
{xinyu.lin.23, guanrou.deng.21, ucesl07}@ucl.ac.uk

Abstract

Traditional pair trading techniques, such as Z-Score model, classical machine learning models, alongside the reinforcement learning approach, primarily focus on statistical arbitrage without considering broader economic dynamics. This paper aims to address this limitation by incorporating macroeconomic indicators and U.S. commodity futures data to build a comprehensive framework and assist in creating trading signals for commodity futures pairs. Specifically, we introduce and analyze three distinct trading strategies: mean reversion (representing classical statistical methods), Support Vector Machine (SVM) (representing traditional machine learning), and reinforcement learning (Deep Q-Network, DQN), providing a holistic evaluation of the contributions of the integrated macroeconomic indicators and U.S. commodity futures data. To thoroughly assess the performance of these strategies, we employ annualized return, volatility, and the Sharpe Ratio as evaluation metrics. Experiment results highlight the potential of reinforcement learning, particularly the DQN model, which could enhance pair trading strategies by integrating broader economic factors.

Keywords: Pair Trading, Chinese Commodity Futures, Reinforcement learning, Deep Q-Network (DQN).

Introduction

The rapid growth of financial markets has advanced trading strategies, particularly in statistical arbitrage. Pair trading stands out as a significant strategy due to its market-neutral nature and ability to profit from relative price changes in related assets (Krauss 2015). China's commodity futures market has seen significant growth over the past two decades, and becomes a key global player due to its role as a major consumer and producer of metals, energy, and agricultural products (Roache 2012). Its high liquidity, diverse products, and unique regulatory environment create distinctive price dynamics, making it an ideal focus for specialized pair trading strategies.

The most widely recognized methods on pair trading include the distance approach and the cointegration approach (Girma and Paulson 1999; Do and Faff 2010; Krauss 2015). Recent research has demonstrated the utility of machine learning techniques, such as Support Vector Machines (SVM) (Huck 2009), reinforcement learning (Kim and Kim 2019), and deep learning based methods (Zhang 2021; Guo et al. 2024) to optimize pair selection and dynamic trading signals. Despite the growing prominence of China's commodity futures market, there remains relatively little research specifically focused on pair trading strategies within this domain. Most existing pair trading studies have centered on equity markets or more developed commodity markets such as those in the U.S. (Andrade, di Pietro, and Seasholes 2005; Gutierrez and Tse 2011).

Besides, the unique characteristics of Chinese commodity futures, influenced by both domestic economic conditions and China's expanding role in global trade, present a significant opportunity for further investigation. Current models often rely on price data to generate trading signals without incorporating the broader economic context, such as domain and domestic macroeconomic indicators, which are known to significantly influence commodity prices (Rad, Low, and Faff 2016; Perlin 2009). The lack of studies that integrate these indicators into the development of pair trading signals for Chinese commodity futures represents a critical shortfall.

This paper aims to address these gaps by developing new pair trading strategies that incorporate both domestic macroeconomic indicators and U.S. commodity futures prices.We begin by analyzing the correlations between Chinese commodity futures prices and both macroeconomic indicators and U.S. commodity prices. Next, we integrate the macroeconomic indicators and U.S. commodity prices with Chinese commodity futures, weighted by their respective correlation coefficients. This enables a deeper understanding of the interconnected factors driving Chinese commodity futures prices. Subsequently, we introduce three distinct trading strategies: mean reversion, Support Vector Machine (SVM), and reinforcement learning (Deep Q-Network, DQN), representing classical statistical methods, traditional machine learning techniques, and advanced machine learning approaches, respectively. The performance of these strategies is rigorously evaluated using key financial metrics, including annualized return, volatility, and the Sharpe Ratio. The results diverge: the Z-Score model's profitability declines, the SVM model experiences substantial instability, but the DQN model exhibits a dramatic improvement, yielding robust returns. The findings highlight the potential of reinforcement learning to capture complex mar-

*Corresponding author

ket dynamics, demonstrating the benefits of incorporating broader economic data into pair trading strategies in an emerging market context. Our key contributions can be summarized as follows:

- We investigate the influence of macroeconomic data and U.S. commodity futures on pair trading performance in the Chinese commodity futures market.
- We design and evaluate comprehensive pair trading strategies using three distinct methodologies in the Chinese commodity futures market.
- We conduct an in-depth analysis of the performance and effectiveness of various trading strategies.

Related Work

We first review pair trading in the Chinese market, identifying existing research gap. Secondly, we also summarize findings about the influence of macroeconomic indicators and U.S. futures on Chinese futures prices.

Recent studies highlight the profitability of pairs trading strategies in the Chinese market (Yang, Goncu, and Pantelous 2017; Chen et al. 2018; Lin and Qiu 2024). Despite simple rules such as the two-standard device approach, advanced methods, such as those that use machine learning (Luo and Chen 2013; Luo et al. 2017; Chen et al. 2018; Zhang 2021) and deep learning (Guo et al. 2024; Fernandez-Perez, Fuertes, and Miffre 2020), demonstrate high effectiveness and profitability. Furthermore, Guo et al. (2024) build a pairing trading strategy using the co-integration relationship and the deep reinforcement learning algorithm on coke futures, iron ore futures and deformed bar futures in the Chinese futures market. These methods only utilize historical price information, but lack comprehensive integration of broader macroeconomic and foreign market factors. This paper aims to address these gaps and push the frontier of pair trading strategies.

Findings have shown the critical role of macroeconomic indicators in shaping market dynamics, and valuable tools for developing predictive models and trading strategies(Bailey and Chan 1993; Borensztein and Reinhart 1994). Macroeconomic announcements significantly impact commodity futures prices, with responses varying across economic cycles (Hess, Kräussl, and Ulrich 2008), and commodity returns are shown to predict GDP growth, positioning them as leading economic indicators (Ye et al. 2021). Moreover, commodity futures volatility is strongly influenced by macroeconomic uncertainty (Watugala 2015), and price uncertainty in sectors like agriculture and energy dampens U.S. economic activity (Triantafyllou, Bakas, and Ioakimidis 2023). These relationships extend globally, as evidenced in studies from countries like India (Sreenu, Rao, and D 2021) and Malaysia (Ahmed et al. 2020). However, a gap remains in leveraging macroeconomic indicators to develop trading signals for pairs trading in Chinese commodity futures markets. Our research aims to address this by integrating macroeconomic variables into pairs trading strategies for these markets.

Research highlights strong information flows and price interactions between U.S. and Chinese commodity futures markets. Studies show dominant U.S. information transmission in commodities like copper, soybeans, and aluminum (Fung, Leung, and Xu 2003; Fung, Liu, and Tse 2010), and a bidirectional relationship with stronger spillovers from the U.S. to China (Liu and Tang 2011). Structural changes since 2002 have further strengthened agricultural market linkages (Lee, Lin, and Liao 2013), and multiple structural breaks underline the integration of U.S. and Chinese soybean futures (Gao et al. 2024). Despite these findings, U.S. futures prices have not been used as variables for predicting trading signals in Chinese pair trading strategies. This motivates us to further incorporate actively traded U.S. commodity futures to enhance signal prediction in the Chinese futures market.

Code	Full Name	Code	Full Name
AU.SHF	Gold	CS.DCE	Polypropylene
AG.SHF	Silver	JD.DCE	Egg
CU.SHF	Copper	A.DCE	Soybean
AL.SHF	Aluminum	B.DCE	Soybean Meal
ZN.SHF	Zinc	M.DCE	Soybean Oil
PB.SHF	Lead	Y.DCE	Palm Oil
NI.SHF	Nickel	P.DCE	Corn Starch
SN.SHF	Tin	C.DCE	Corn
RU.SHF	Rubber	L.DCE	Polyethylene
FU.SHF	Fuel Oil	V.DCE	Polyvinyl Chloride
BU.SHF	Asphalt	PP.DCE	Polypropylene

Table 1: List of Commodities and Their Full Names

Methodology

This section begins by selecting Chinese commodity futures, macroeconomic indicators, and U.S. commodity futures as the dataset. Cointegration analysis is subsequently applied to identify pairs exhibiting stable long-term relationships, forming the basis for the development of trading strategies. The data is standardized and enriched with domestic data: macroeconomic indicators and international data: U.S. commodity futures, to build a comprehensive framework and assist in creating trading signals for commodity futures pairs.

Furthermore, we introduce three distinct trading strategies: mean reversion, Support Vector Machine (SVM), and reinforcement learning, representing classical statistical methods, traditional machine learning techniques, and advanced machine learning approaches, respectively. To thoroughly assess the performance of these strategies, we employ annualized return, volatility, and the Sharpe Ratio as evaluation metrics.

Commodity Futures Selection

The selection of Chinese commodity futures prioritizes contracts with high liquidity and trading volume, ensuring accurate market representation, minimal price manipulation, and efficient trade execution (Chen et al. 2017; Fernandez-Perez et al. 2020).

Gold and silver, being commodities with high liquidity (CME Group 2018), have prices that accurately reflect the true market value and are less susceptible to manipulation or sudden changes. Significant trading volume allows for convenient entry and exit, reducing the possibility of slippage. Metals such as copper and aluminium are essential in the manufacturing and construction sectors. Copper, often referred to as "Dr. Copper" due to its close relationship with global economic trends (Golding and Golding 2017), sees rising demand during periods of economic growth. Fuel oil and natural gas, as energy commodities, play a crucial role in the global economy by providing energy for industries, transportation, and households (Stevens 2016). Agricultural products are vital for the food supply chain and essential for ensuring food security. The selection, summarized in Table 1, covers a broad spectrum of industries, from precious metals to agricultural products and energy resources. This diversity allows for a comprehensive analysis that can capture different economic dynamics and sectoral shifts.

Macroeconomic and U.S. Futures Selection

China's economy is deeply integrated into the global supply chain, making its commodity futures markets sensitive to both domestic and international influences. Macroeconomic indicators reflect domestic economic conditions, while U.S. commodity futures provide insight into international market trends. Therefore, we consider the influence of both the Chinese macroeconomic indicator and U.S. commodity futures. The macroeconomic indicator included in this study includes Gross Domestic Product (GDP), Consumer Price Index (CPI), Chinese Commodity Price Index (CCPI). Loan Prime Rate for a 1-year period (LPR-1), and a major stock market index that tracks the top 300 stocks on the Shanghai and Shenzhen stock exchanges (000979.CSI). U.S. commodity futures include Gold futures traded on the COMEX exchange (GC.CMX), Silver futures traded on the COMEX exchange (SI.CMX), Copper futures traded on the COMEX exchange (HG.CMX), Platinum futures traded on the NYMEX exchange (PL.NYM), Palladium futures traded on the NYMEX exchange (PA.NYM).

Pair Selection for Commodity Futures

The selection of commodity futures pairs is conducted using the cointegration methodology. Cointegration emphasizes the stationarity of the spread between assets, aiming to identify pairs of commodity futures that exhibit a stable long-term equilibrium relationship (Ogaki and Park 1997; Rad, Low, and Faff 2016). We use the price spread between each pair of commodity futures as input to test for cointegration.

$$P_{i,t} = \beta P_{j,t} + \varepsilon_t \tag{1}$$

where $P_{i,t}$ and $P_{j,t}$ are the two commodity futures prices being tested for cointegration. Then the residuals (ε_t) are tested for stationarity using the Augmented Dickey-Fuller (ADF) test (Cheung and Lai 1995).

Based on the resulting p-values from the ADF test, we retain the pairs with a p-value <0.05, indicating evidence against the null hypothesis that the series are not cointegrated. Therefore, the pairs that pass the ADF test are selected for further trading strategies.

Macroeconomic and Futures Data Merge

To factor in the potential impact of U.S. commodity futures and Chinese macroeconomic indicators on the price movements of Chinese commodity futures, this study merges the commodity prices with U.S. commodity prices and macroeconomic indicator values, based on the correlation matrix of the Chinese commodity futures with U.S. commodity futures and Chinese macroeconomic indicators.

The merging process combines $Q_{i,t}$, the standardised price of the i-th Chinese commodity futures at time t, with the additional features (U.S. commodity prices and macroeconomic indicator values), taking into account the correlations between them

$$X_{i,t} = Q_{i,t} + \sum_{j=1}^{N} r_{ij} \cdot F_{j,t} + \sum_{k=1}^{M} s_{ik} \cdot G_{k,t}, \tag{2}$$

where $X_{i,t}$ is the merged value for the i-th Chinese commodity futures at time t, r_{ij} is the Pearson correlation coefficient between the i-th Chinese commodity futures and the j-th U.S. commodity prices, s_{ik} is the Pearson correlation coefficient between the i-th Chinese commodity futures and the k-th macroeconomic indicators, $F_{j,t}$ is the value of the j-th U.S. commodity prices at time t, $G_{k,t}$ is the value of the k-th macroeconomic indicators at time t, N is the total number of U.S. commodity prices, M is the total number of macroeconomic indicators.

Trading Strategies

We discuss three strategies: (1) a mean reversion strategy using the z-score of the price ratio to generate signals, (2) an SVM-based strategy predicting spread movements and signaling long (+1), short (-1), or neutral (0) positions, and (3) a Deep Q-Network (DQN) strategy optimizing entry and exit thresholds for trading.

Mean Reversion The mean reversion strategy assumes that the price relationship between two assets reverts to the mean over time (Poterba and Summers 1988). Cointegrated pairs are used to trade in opposite positions, with the price ratio employed for signal generation.

The z-score is defined as:

$$Z_t = \frac{\mathrm{MA}_1(t) - \mathrm{MA}_2(t)}{\sigma_{\text{rolling}}(t)}, \tag{3}$$

where Z_t represents the z-score at time t. $\mathrm{MA}_1(t)$ is the short-term moving average (e.g., 5-day) at time t. $\mathrm{MA}_2(t)$ is the long-term moving average (e.g., 21-day) at time t,=.$\sigma_{\text{rolling}}(t)$ is the rolling standard deviation (e.g., 21-day) at time t.

The z-score measures how far the current ratio deviates from the mean. A z-score below -2 triggers a long on asset_1 and short on asset_2, while a z-score above 2 triggers the reverse. No trades occur in $[-2, 2]$ to avoid false signals (Gatev, Goetzmann, and Rouwenhorst 1999, 2006). Positions close when the z-score returns to $[-0.5, 0.5]$.

SVM We also apply the Support Vector Machine (SVM) method in this paper. The price spread S_t is calculated as the difference between the prices of the Chinese commodity futures,

$$S_t = Q_{i,t} - Q_{j,t}. \tag{4}$$

A feature matrix X_t, at time t, is constructed using lagged values of the spread over a 21-day lookback period

$$X_t = [S_{t-1}, S_{t-2}, \ldots, S_{t-21}]. \tag{5}$$

The target variable, Y, is the sign of the spread's first difference, with +1 indicating an increase and -1 a decrease. Any zeros are replaced with the last non-zero value for consistency.

The strategy starts with a neutral position. The SVM predicts the spread's movement (+1 or -1) based on the test set features, and trading actions follow: entering a long position for +1, or a short position for -1. Existing long or short positions are held unless the signal reverses, prompting the position to be closed. This prevents direct switching between positions, reducing whipsaw losses.

Reinforcement Learning For a comprehensive comparison, we further execute a pair trading strategy using reinforcement learning, specifically a Deep Q-Network (DQN). This approach allows the strategy to potentially capture more complex relationships and adapt to varying market conditions. Trades are recorded, including positions taken (long or short), quantities of each asset, costs, and profits. The method also manages trade exits based on the spread moving beyond predefined stop loss or take profit levels, updating the portfolio accordingly.

We utilise the list of cointegrated pairs selected through cointegration tests, and the state for the DQN model is constructed as normalised the spread between the chosen assets using min-max normalization. This state is passed to the DQN to predict the entry threshold, which determines when to enter a trade.

The reward is defined as the total profit generated from the trading strategy as follows:

$$W = W_{\text{asset1}} + W_{\text{asset2}} - C \tag{6}$$

where W is the reward, which is the total profit, W_{asset1} and W_{asset2} are profits or losses from trading asset 1 and asset 2, C means any short fees and transaction costs. The target Q-value for the taken action is computed as:

$$Q_{\text{target}} = W + \gamma \max_{a \in A} Q(s_{t+1}, a) \tag{7}$$

where γ is the discount factor, s_{t+1} is the next state, $\max_{a \in A} Q(s_{t+1}, a)$ represents the maximum Q-value across all possible actions the agent can take in the next state, A is the set of possible actions (e.g., "long," "short," "hold").

The DQN is trained using the Mean Squared Error (MSE) loss function

$$L = \frac{1}{N} \sum_{i=1}^{N} \left(Q_{\text{predicted}} - Q_{\text{target}}\right)^2, \tag{8}$$

where $Q_{\text{predicted}}$ is the predicted Q-value, the trading threshold predicted by the DQN. By using this approach, the DQN dynamically adjusts the entry threshold based on the market's current state, ensuring that the threshold is always within a reasonable and interpretable range.

The DQN iteratively learns the optimal strategy by interacting with the market environment, adjusting its predictions ($Q_{\text{predicted}}$) based on the observed reward (W) and state (S). Stop-loss and take-profit levels are dynamically calculated at 15% of the spread's range during the trading period, reflecting the model's adaptive understanding of market conditions.

Initially, with no positions open (position $= 0$), a long signal triggers when the normalized spread exceeds the entry threshold but remains below the stop-loss. The strategy shorts asset_1 and goes long asset_2, adjusting quantities based on capital. Conversely, a short signal occurs when the spread falls below the negative entry threshold but stays above the negative stop-loss, prompting the reverse position. Positions are exited if the spread hits stop-loss, take-profit, or their negative counterparts. Profit is calculated considering price differences, transaction costs, and short position interest.

Performance matrices

The performance of each strategy is evaluated with annualized return, volatility, and the Sharpe Ratio.

Total return is determined by dividing the final portfolio value by the first nonzero portfolio value and then subtracting one. The annualized return is derived from the total return, and calculated based on the weighted portfolio value.

$$R_{\text{annual}} = (1 + R_{\text{total}})^{\frac{1}{T}} - 1, \tag{9}$$

where R_{annual} is the annualized return, R_{total} is the total return over the investment period, T is the number of years.

Volatility is calculated as the standard deviation of daily returns.

The Sharpe ratio SR is calculated as follows:

$$SR = \frac{\bar{r} - r_f}{\sigma}, \tag{10}$$

where $\bar{r}$ is the mean return, r_f is the risk-free rate, [1] σ is the standard deviation of returns (volatility). This measures the risk-adjusted return and indicates how much excess return is generated per unit of risk.

Experiments

This section is structured into three main parts: first, it outlines the dataset, which includes Chinese commodity futures, macroeconomic indicators, and U.S. commodity futures, spanning 2014–2024. The second part focuses on the selection results for cointegrated commodity pairs, and external factors (macroeconomic indicators and U.S. commodity futures) that are correlated with each Chinese commodity futures. Lastly, the section examines the results of strategy testing and performance evaluation.

[1] In this paper, the risk-free rate is set to be 2.16%, which is the current 10-year Chinese Government Bond Yield.

Commodity 1	Commodity 2	Commodity 1	Commodity 2
CU. SHF	AL. SHF	FU. SHF	L. DCE
AL. SHF	SN. SHF	BU. SHF	M. DCE
AL. SHF	M. DCE	BU. SHF	JD. DCE
AL. SHF	P. DCE	A. DCE	Y. DCE
NI. SHF	M. DCE	A. DCE	P. DCE
SN. SHF	P. DCE	A. DCE	CS. DCE
RU. SHF	FU. SHF	B. DCE	M. DCE
RU. SHF	BU. SHF	B. DCE	Y. DCE
RU. SHF	A. DCE	B. DCE	P. DCE
RU. SHF	B. DCE	M. DCE	Y. DCE
RU. SHF	M. DCE	M. DCE	P. DCE
RU. SHF	Y. DCE	M. DCE	C. DCE
RU. SHF	P. DCE	M. DCE	CS. DCE
RU. SHF	C. DCE	Y. DCE	P. DCE
RU. SHF	CS. DCE	C. DCE	CS. DCE
RU. SHF	JD. DCE	JD. DCE	L. DCE
RU. SHF	L. DCE	JD. DCE	V. DCE
RU. SHF	V. DCE	JD. DCE	PP. DCE
RU. SHF	PP. DCE		

Table 2: Table of Cointegrated Pairs.

Data

The dataset utilized in this study includes a diverse collection of financial and economic data obtained from Wind. This includes 22 Chinese commodity prices, covering a wide range of asset classes, 5 Chinese macroeconomic indicators and 5 U.S. commodity futures prices to account for international market interactions and cross-border dynamics, as we described in section: Methodology. The dataset spans from 2014 to 2024 with 2,434 daily observations, providing a long-term view that allows for the observation of trends, cycles, and significant events impacting these commodity futures prices. The training and testing splits are time-based, with 80% (data from 2014 to 2021, 1948 data points) of the data used as the training set and the remaining 20% (data from 2022 to 2023, 486 data points) as the testing set.

The result of commodity futures pair selection based on the cointegration approach is provided in Table 2. For the correlation between macroeconomic indicators and commodity futures, GDP shows strong positive correlations with Gold, Copper, and Aluminum, this reflects its influence on industrial activity. CCPI strongly correlates with Aluminum, Soybean Oil, Gold, and Nickel, which links consumer price levels to commodity prices. LPR-1 has negative correlations with Aluminum, Zinc, and Gold, indicating higher borrowing costs may dampen demand. PPI moderately correlates with Aluminum and Zinc, while 000979.CSI reflects a positive link between stock market trends and commodity prices, particularly for Zinc and Gold.

In terms of the relationship between U.S. commodity futures and Chinese commodity futures, strong correlations in metals (e.g., Gold and Copper) reflect the globalized nature of these markets, enabling cross-market predictions. Weaker correlations in agriculture and energy suggest regional influences like weather and domestic policies.

	Z-score Model			SVM Model			DQN Model		
	AR	**Vol.**	**SR**	**AR**	**Vol.**	**SR**	**AR**	**Vol.**	**SR**
CCF	0.93	0.36	0.03	0.05	0.01	-1.86	-6.61	1.87	0.06
M_Price	0.86	1.61	-0.04	-0.35	30.15	-2.40	15.30	0.95	-0.02
Imp.	-0.06	1.25	-0.07	-0.41	30.14	-0.54	21.90	-0.92	-0.08

Table 3: Summary of Performance Matrices for each Model (Vol.: Volatility, Imp.: Improvement)

Results of Trading Strategies on CCF and Merged Data

The first row in Table 3 displays the performance metrics of three trading models — the Mean Reversion (Z-Score) Model, SVM, and DQN Model which are used to compare with only utilizing the CCF price and merged price(M_Price) that incorporates macroeconomic indicators and U.S. commodity futures.

Annualized Return The Z-Score Model shows a substantial positive annualized return of 0.93, indicating that this model is profitable in this context. In contrast, the SVM Model yields a much lower annualized return of 0.05, suggesting it generates far less profit than the Z-Score Model. This may indicate that the SVM model is less effective on this dataset or is overly conservative in identifying trading opportunities. The RL Model presents a negative annualized return of -6.61, which is significantly worse than both the Z-Score and SVM Models. This negative return suggests that the RL Model incurs considerable losses, possibly due to difficulties in finding profitable trades or being penalised by poor trades in this setting.

After incorporating macroeconomic indicators and U.S. commodity futures prices, the Z-Score Model shows a decreased annualized return of 0.86 compared to 0.93 when using only commodity futures prices. This reduction suggests that the inclusion of additional features does not enhance the model's profitability and may have introduced noise or complexity that the model struggled to handle. The SVM Model exhibits a significantly negative annualized return of -0.35, a sharp decline from the positive 0.05 return when using only commodity prices, indicating that the additional features substantially degraded its performance. In contrast, the RL Model shows a dramatic improvement, with an annualized return of 15.30, compared to -6.61 with only commodity prices. This suggests that the DQN Model effectively leveraged the additional features to generate significantly higher profits.

Volatility With a volatility measure of 0.86, the Z-Score Model exhibits moderate risk, reflecting some fluctuation in returns. However, given its positive return, this level of volatility may be acceptable depending on risk tolerance. The SVM Model shows exceptionally low volatility at 0.01, indicating highly stable returns, likely due to limited trading activity or overly cautious positions that fail to capture significant market movements. The RL Model shows high volatility at 1.87, more than twice that of the Z-Score Model. This high volatility, combined with negative returns, suggests that the RL Model is highly unstable, taking on significant risk without adequate compensation in terms of returns.

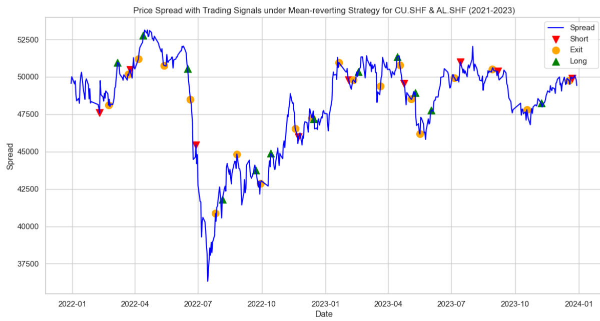

Figure 1: Trading Signal for the Z-score Model

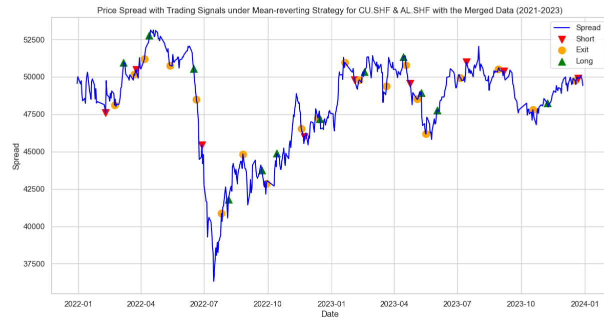

Figure 2: Trading Signal for the Z-score Model with merged data

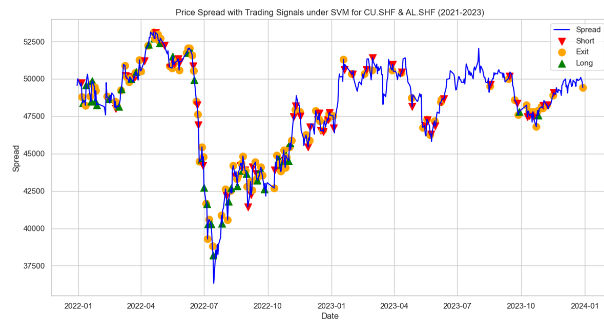

Figure 3: Trading Signal for the SVM Model

With the merged data, the volatility for the Z-Score Model increased to 1.61 from 0.86, indicating a higher risk level. Given the decline in returns, the Z-Score Model does not improve after incorporating macroeconomic indicators and U.S. commodity futures prices. The SVM Model's volatility skyrocketes to 30.15 from 0.01, reflecting extreme instability and likely erratic trading behavior. This sharp increase suggests that the model struggles to manage the complexity of the additional features. The DQN Model's volatility decreases to 0.95 from 1.87, indicating that while the model became more profitable, it also manages to reduce its risk, leading to more stable performance.

Sharpe Ratio The Sharpe Ratio for the Z-Score Model is 0.03, which is positive but relatively low, indicating that the model provides a modest return relative to its risk. Although profitable, the return may not sufficiently compensate for the risk taken. The SVM Model exhibits a negative Sharpe Ratio of -1.86, indicating poor performance when adjusted for risk. The negative ratio implies that the returns do not compensate for the risk, making the model underperform relative to a risk-free investment. The RL Model shows a slightly positive Sharpe Ratio of 0.06, but given its high volatility and negative returns, this suggests inefficient risk management. The slightly positive value indicates minimal compensation for the high risk, pointing to an ineffective strategy.

After including the macroeconomic indicators and U.S. commodity futures prices, the Sharpe Ratio for the Z-Score Model turns negative at -0.04, down from 0.03, suggesting underperformance on a risk-adjusted basis after including the additional features. This indicates that the returns do not sufficiently compensate for the increased risk, and the added features likely cause more harm than benefit. The SVM Model's Sharpe Ratio is worsened to -2.40 from -1.86, reflecting an even poorer risk-adjusted performance, as the model failed to generate profits and took on significantly more risk. The DQN Model's Sharpe Ratio remains slightly negative at -0.02, an improvement from its previous value, though still not optimal. However, the substantial increase in annualized return and the decrease in volatility suggest that the DQN Model is moving in the right direction, though further refinement is needed to fully optimise the risk-adjusted performance.

Trading Signal Figure 1 and Fig. 2 display the trading signals for the Z-Score model. Figure 1 shows signals generated based solely on commodity futures prices, while Fig. 2 illustrates the signals produced using the merged data. Upon closer inspection, there are no significant differences between the two charts. The price trends, the timing of buy and sell signals, and the number of signals are all very similar, if not identical. This indicates that incorporating additional features in the merged data did not significantly impact the trading signals, resulting in both charts reflecting the same trading strategy outcome.

Figure 3 and Fig. 4 show the trading signals for the SVM model. In Fig. 3, the model generates a high number of trading signals, with frequent entries and exits, even during minor fluctuations in the spread. This suggests that the SVM model is more sensitive to short-term market movements and may be attempting to capitalise on smaller, more frequent opportunities. The frequency of signals indicates potential overtrading, as the model appears to react to every minor movement. This could lead to higher transaction costs and reduced profitability, as many trades may not generate significant returns before the model exits. The timing of signals in Figure 3 often coincides with minor fluctuations, leading to a choppy trading pattern. This suggests that the SVM model might be overfitting to short-term noise in the data rather than capturing broader trends.

In Fig. 4, the SVM model with merged features generates far fewer signals, primarily short positions, and notably fewer than in the commodity-only version. The model appears overly cautious or perhaps overwhelmed by the com-

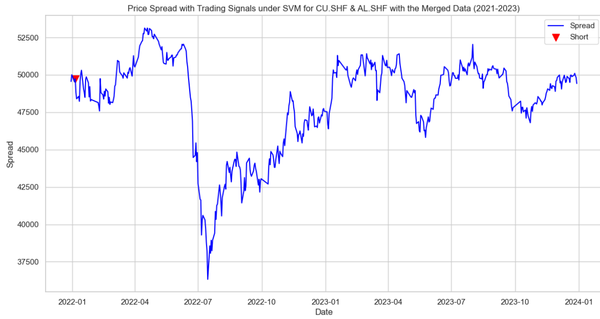

Figure 4: Trading Signal for the SVM Model with merged data

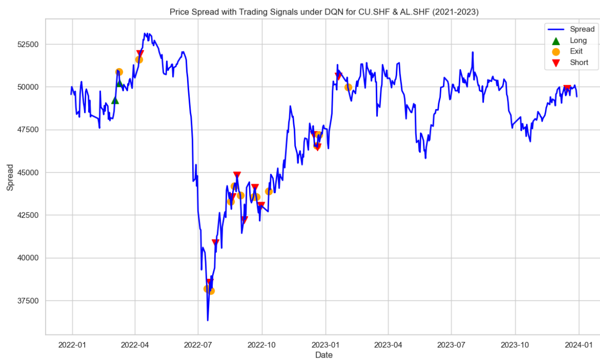

Figure 5: Trading Signal for the DQN Model

plexity introduced by the additional features. The lack of long signals and overall sparse trading activity suggests that the model may be too conservative or unable to identify profitable opportunities within the expanded feature space.

According to Fig.5, the DQN model generates relatively fewer trading signals compared to the SVM model on the left in Fig .3, adopting a more conservative approach to entering and exiting trades. This suggests that the DQN is more selective, likely aiming to avoid false signals by entering trades only when the predicted conditions are strongly favorable. The DQN signals seem to follow major trends, with entries and exits spaced out over longer periods. The model places trades at key inflection points, such as peaks and troughs in the spread, indicating a focus on capturing large movements rather than frequent small fluctuations. Exits occur relatively soon after entries, particularly during sharp market movements, showing that the DQN prioritises risk management, quickly locking in profits or cutting losses.

The DQN model with merged features, shown on Fig. 6, generates fewer signals than the Z-Score and SVM models. The signals are more concentrated around critical points, such as peaks and troughs in the spread, suggesting that the DQN model is more selective and better at identifying significant trading opportunities with the merged data. Exits tend to follow shortly after entries, especially during volatile periods, highlighting the model's emphasis on risk management and profit-taking at crucial moments.

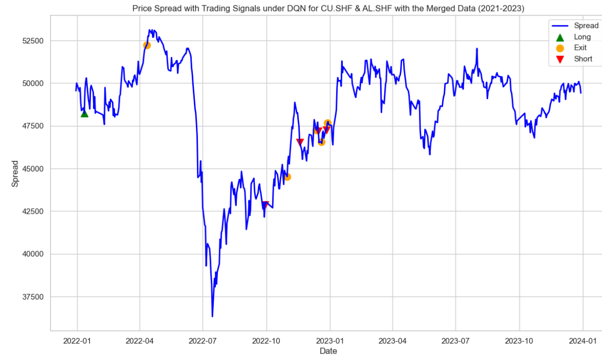

Figure 6: Trading Signal for the DQN Model with merged data

Discussion

The Z-Score model's performance declined with added features, indicating its preference for simpler datasets. Similarly, the SVM model's performance collapsed under the complexity of the merged dataset, generating sparse signals and missing opportunities, as shown in Figures 3 and 4. In contrast, the DQN model significantly benefits from the merged features, shifting from losses to high profitability by leveraging more informative signals for better decision-making.

These results highlight the DQN model's adaptability to complexity, while the Z-Score model favors simpler data, and the SVM model struggles with merged features. While additional features improved the DQN model, they hindered the Z-Score and SVM models, suggesting these are better suited to less complex datasets.

Conclusion

This paper explores the integration of macroeconomic indicators and U.S. commodity futures data into pair trading strategies, addressing a critical gap in traditional approaches that primarily rely on statistical arbitrage without considering broader economic dynamics. By analyzing three distinct trading strategies—mean reversion, Support Vector Machine (SVM), and Deep Q-Network (DQN), the research provides a comprehensive evaluation of the impact of incorporating macroeconomic data. Among price-only strategies, the Z-Score model was the most effective, achieving a moderate annualized return of 0.93, while the SVM model underperformed with a minimal return of 0.05, and the DQN model incurred significant losses with a negative return of -6.61. However, when macroeconomic indicators and U.S. futures data were incorporated, the Z-Score and SVM models struggled to adapt, with the Z-Score's performance slightly declining to 0.86 and the SVM suffering a negative return of -0.35. In contrast, the DQN model excelled with a robust annualized return of 15.30, demonstrating its ability to leverage the additional data effectively. These findings highlight the potential of reinforcement learning, particularly the DQN model, to enhance pair trading strategies by integrating broader economic factors.

References

Ahmed, K.; Zain, M.; Noor, A. M.; Alsaadi, A. R.; Milhem, M.; and 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: 2515–8260.

Andrade, S. C.; di Pietro, V.; and Seasholes, M. S. 2005. Understanding the Profitability of Pairs Trading. *UC Berkeley*. Working Paper.

Bailey, W.; and Chan, K. C. 1993. Macroeconomic Influences and the Variability of the Commodity Futures Basis. *The Journal of Finance*, 48(2): 555–573.

Borensztein, E.; and Reinhart, C. M. 1994. The Macroeconomic Determinants of Commodity Prices. *IMF Staff Papers*, 41(2): 236–261.

Chen, D.; Cui, J.; Gao, Y.; and Wu, L. 2017. Pairs trading in Chinese commodity futures markets: an adaptive cointegration approach. *Accounting & Finance*, 57(5): 1237–1264.

Chen, D.; Cui, J.; Gao, Y.; and Wu, L. 2018. Pairs trading in Chinese commodity futures markets: An adaptive cointegration approach. *Accounting & Finance*, 57(5): 1237–1264.

Cheung, Y.-W.; and Lai, K. S. 1995. Lag order and critical values of the augmented Dickey–Fuller test. *Journal of Business & Economic Statistics*, 13(3): 277–280.

CME Group. 2018. Trading COMEX Gold and Silver. Accessed: 2024-11-29.

Do, B.; and Faff, R. 2010. Does Simple Pairs Trading Still Work? *Financial Analysts Journal*, 66(4): 83–95.

Fernandez-Perez, A.; Frijns, B.; Indriawan, I.; and Tse, Y. 2020. Pairs trading of Chinese and international commodities. *Applied Economics*, 52(48): 5203–5217.

Fernandez-Perez, A.; Fuertes, A.-M.; and Miffre, J. 2020. Commodity markets, long-horizon predictability, and intertemporal relations: A comprehensive review and future perspectives. *Journal of Commodity Markets*, 19: 100118.

Fung, H.-G.; Leung, W. K.; and Xu, X. E. 2003. Information Flows Between the U.S. and China Commodity Futures Trading. *Review of Quantitative Finance and Accounting*, 21: 267–285.

Fung, H.-G.; Liu, Q.; and Tse, Y. 2010. The Information Flow and Market Efficiency Between the U.S. and Chinese Aluminum and Copper Futures Markets. *Journal of Futures Markets*, 30(12): 1192–1209.

Gao, X.; Insuwan, A.; Li, Z.; and Tian, S. 2024. The dynamics of price discovery between the U.S. and Chinese soybean market: A wavelet approach to understanding the effects of Sino-US trade conflict and COVID-19 pandemic. *Data Science and Management*, 7: 35–46.

Gatev, E.; Goetzmann, W. N.; and Rouwenhorst, K. G. 1999. Pairs Trading: Performance of a Relative Value Arbitrage Rule. *Review of Financial Studies*, 19(3): 797–827.

Gatev, E.; Goetzmann, W. N.; and Rouwenhorst, K. G. 2006. Pairs trading: Performance of a relative-value arbitrage rule. *Review of Financial Studies*, 19(3): 797–827.

Girma, P.; and Paulson, D. 1999. Cointegration and error correction models in econometrics. *Econometric Reviews*, 18(4): 385–410.

Golding, B.; and Golding, S. D. 2017. Doctor Copper and King Coal. In *Metals, Energy and Sustainability*, 1–20. Springer.

Guo, M.; Liu, J.; Luo, Z.; and Han, X. 2024. Deep reinforcement learning for pairs trading: Evidence from China black series futures. *International Review of Economics & Finance*, 93: 981–993.

Gutierrez, J. A.; and Tse, Y. 2011. Illuminating the Profitability of Pairs Trading: A Test of the Relative Pricing Efficiency of Markets for Water Utility Stocks. *The Journal of Trading*, 6(2): 50–64. Spring 2011.

Hess, M. K.; Kräussl, R.; and Ulrich, M. 2008. Commodity Futures and Global Equity Markets: Evidence from the U.S., U.K., and Germany. *Journal of Futures Markets*, 28(4): 359–388.

Huck, N. 2009. Pairs selection and outranking: An application to the S&P 100 index. *European Journal of Operational Research*, 196(2): 819–825.

Kim, I.; and Kim, J. 2019. Pairs Trading in Cryptocurrency Markets. *Finance Research Letters*, 30: 234–240.

Krauss, C. 2015. Statistical arbitrage pairs trading strategies: Review and outlook. Working paper, Friedrich-Alexander-Universität Erlangen-Nürnberg, Institute for Economics.

Lee, K.-C.; Lin, C.-C.; and Liao, T. L. 2013. The Effect of Structural Change on Information Flow Between the U.S. and Chinese Agricultural Futures Markets. *China Economic Journal*, 6(1): 25–48.

Lin, F.; and Qiu, Z. 2024. Pairs Trading Strategy and Connected Stocks: Evidence from China. Technical report, School of Business, Ludong University and Renmin University of China. SSRN Working Paper.

Liu, J.; and Tang, D. 2011. Commodity Derivatives Pricing with Stochastic Convenience Yields and Interest Rates. *Journal of Futures Markets*, 31(5): 392–421.

Luo, L.; and Chen, X. 2013. Integrating piecewise linear representation and weighted support vector machine for stock trading signal prediction. *Applied Soft Computing*, 13(2): 806–816.

Luo, L.; You, S.; Xu, Y.; and Peng, H. 2017. Improving the integration of piece wise linear representation and weighted support vector machine for stock trading signal prediction. *Applied Soft Computing*, 56: 199–216.

Ogaki, M.; and Park, J. Y. 1997. A cointegration approach to estimating preference parameters. *Journal of Econometrics*, 82(1): 107–134.

Perlin, M. S. 2009. Evaluation of pairs-trading strategy at the Brazilian financial market. *Journal of Derivatives & Hedge Funds*, 15(2): 122–136.

Poterba, J. M.; and Summers, L. H. 1988. Mean reversion in stock prices: Evidence and implications. *Journal of financial economics*, 22(1): 27–59.

Rad, H.; Low, R. K. Y.; and Faff, R. 2016. The profitability of pairs trading strategies: distance, cointegration and copula methods. *Quantitative Finance*, 16(10): 1541–1558.

Roache, S. K. 2012. China's Impact on World Commodity Markets. IMF Working Paper WP/12/115, International Monetary Fund. Prepared by Shaun K. Roache and authorized for distribution by Thomas Helbling.

Sreenu, N.; Rao, K.; and D, K. 2021. The macroeconomic variables impact on commodity futures volatility: A study on Indian markets. *Cogent Business & Management*, 8: 1–18.

Stevens, P. 2016. The Role of Oil and Gas in the Development of the Global Economy. WIDER Working Paper 2016/175, World Institute for Development Economics Research (UNU-WIDER).

Triantafyllou, A.; Bakas, D.; and Ioakimidis, M. 2023. Commodity price uncertainty as a leading indicator of economic activity. *International Journal of Finance & Economics*, 28(4): 4194–4219.

Watugala, S. W. 2015. Economic Uncertainty and Commodity Futures Volatility. Working Paper 15-14, Office of Financial Research and University of Oxford.

Yang, Y.; Goncu, A.; and Pantelous, A. 2017. Pairs trading with commodity futures: Evidence from the Chinese market. *China Finance Review International*, 7(3): 304–325.

Ye, W.; Guo, R.; Deschamps, B.; Jiang, Y.; and Liu, X. 2021. Macroeconomic forecasts and commodity futures volatility. *Economic Modelling*, 94: 981–994.

Zhang, L. 2021. Pair trading with machine learning strategy in China stock market. In *Proceedings of the 2021 2nd International Conference on Artificial Intelligence and Information Systems (ICAIIS)*, 1–6. ACM.