

ORIGINAL ARTICLE

Enhancing algorithmic trading with wavelet-based deep reinforcement learning: a multi-indicator approach

Antonio José Martínez Casares¹ 

Received: 28 March 2025 / Accepted: 14 August 2025 / Published online: 14 September 2025

© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2025

Abstract

This research investigates wavelet-enhanced deep reinforcement learning (DRL) for trading S&P 500 futures, assessing four wavelet families (Daubechies, Symlets, Coiflets, and Biorthogonal) in conjunction with three DRL algorithms (PPO, A2C, and DQN). We employ level-2 decomposition utilizing conservative soft thresholding on market microstructure indicators (DIX, GEX, VIX), enhancing signal-to-noise ratios by 25–41 dB. The coif4+DQN combination yields the most robust outcomes (Sharpe ratio: 0.96; total return: 112.5%), while A2C+db1 and PPO+db4 attain Sharpe ratios of 0.803 and 0.801, respectively. In comparison with XGBoost, random forest, and Logistic regression, wavelet-DRL attains a 35–70% superior Sharpe ratio and reduces maximum drawdowns (0.28–0.34 vs. 0.39–0.42). Feature-set analysis reveals that DIX, GEX, and VIX combined surpass single-indicator configurations by 18.7% in Sharpe ratio. Statistical analyses validate the robustness, revealing that 82.1% of maximum drawdown (MDD) and 67.9% of Sharpe ratio comparisons are significant at the 5% level. Our findings support a gradual implementation strategy—transitioning from A2C+db1 to DQN+coif4—to enhance institutional algorithmic trading efficacy.

Keywords Deep reinforcement learning · Wavelet transform · Algorithmic trading · Financial time series

1 Introduction

Algorithmic trading has emerged as a predominant influence in contemporary financial markets by utilizing advanced mathematical and statistical models to automate trading decisions [69]. The emergence of artificial intelligence, especially machine learning methods such as deep learning and reinforcement learning, has transformed this domain, facilitating the development of trading algorithms that function with speeds and precision exceeding human abilities [95]. Deep reinforcement learning has emerged as a notably promising method for tackling the complexity inherent in financial markets. Refer to [82]. The effective implementation of DRL in algorithmic trading necessitates meticulous attention to various factors, such as the attributes of market data, the formulation of suitable feature representations, and the integration of resilient risk management strategies [6].

The financial markets provide distinct challenges for machine learning algorithms because of their intrinsic noise, non-stationarity, and high dimensionality [84]. Conventional time series models, such as ARIMA and GARCH, frequently fail to encapsulate the intricate nonlinear dynamics that dictate asset prices [97]. Furthermore, the existence of high-frequency noise and the fluctuating nature of market conditions may result in overfitting and suboptimal generalization performance. Deep learning models, especially recurrent neural networks and convolutional neural networks, have shown efficacy in identifying pertinent elements from financial data and forecasting future price fluctuations.[23]. Technical analysis, utilizing price charts and technical



indicators, can be amalgamated with machine learning models, particularly deep learning algorithms, to produce trading signals [25]. Nonetheless, these models frequently necessitate comprehensive feature engineering and meticulous parameter optimization to attain peak performance.

Recent studies highlight the growing significance of deep learning methods in financial time series forecasting, surpassing traditional machine learning approaches [72]. The ability of deep learning to identify complicated patterns in multiple datasets makes it especially appropriate for the intricacies of financial markets. Financial markets are affected by a wide range of factors, including economic indicators, political events, and investor mood, leading to complex and dynamic patterns [10]. The effectiveness of deep learning in capturing these subtleties has been validated through many applications, such as stock price prediction, portfolio optimization, and risk management [42, 72, 73, 87]. Machine learning techniques are widely utilized in the financial sector for objectives such as predicting stock market patterns and evaluating bankruptcy or credit risk. [27]. Formulating efficient trading strategies requires the capacity to analyze and evaluate extensive data, discern significant signals from extraneous information, and adjust to evolving market dynamics.

The amalgamation of wavelet transforms with deep reinforcement learning algorithms signifies a potentially potent method for augmenting the efficacy and resilience of algorithmic trading systems. Wavelet transforms have numerous benefits compared to conventional Fourier-based techniques for financial signal processing, such as enhanced time–frequency localization and the capacity to decompose signals across different scales [98]. Wavelet transforms decompose financial time series into various frequency components, isolating pertinent information and filtering out noise, which enhances the signal-to-noise ratio and improves the learning process for DRL agents. The application of diverse artificial intelligence techniques in financial trading is increasing, with deep learning being among the most prevalent [18]. The amalgamation of various machine learning models has attracted considerable interest in financial stock forecasting, demonstrating potential for enhancing prediction accuracy and robustness. [40]. Moreover, the utilization of deep learning is beneficial in integrating diverse data into investment decisions, resulting in a reduced feature space via compression [86].

The prospective advantages of amalgamating wavelet transforms with deep reinforcement learning for algorithmic trading encompass superior feature extraction, augmented noise reduction, and more resilient trading strategies. Machine learning is pivotal to innovation in the financial sector, as it optimizes prospective profits and substantially mitigates associated risks. [10]. Utilizing the distinct advantages of wavelet transforms and deep reinforcement learning, it may be feasible to create trading systems that are more flexible, robust, and lucrative in response to the problems presented by contemporary financial markets. [2, 47, 70]. Furthermore, the design of interpretable trading patterns, augmented by machine learning algorithms, has demonstrated significant efficacy, offering a robust method for profitability and usability. [79]. The financial sector is experiencing an increase in the utilization of machine learning and artificial intelligence, mostly attributable to the accessibility of vast datasets, advanced algorithms, and innovative approaches that are transforming several applications.[70].

Notwithstanding these encouraging advancements, some significant deficiencies persist in the current literature. Primarily, the majority of studies concentrate on individual wavelet families, lacking a systematic assessment of several wavelet types and their relative efficacy. Secondly, there is a paucity of research that thoroughly contrasts wavelet-enhanced deep reinforcement learning methods with classical machine learning techniques and conventional trading strategies in realistic market environments. The systematic justification of wavelet decomposition levels and the statistical relevance of performance enhancements have been inadequately addressed.

This study mitigates these constraints by performing a thorough review of wavelet-enhanced deep reinforcement learning for algorithmic trading utilizing S&P 500 futures data. Our methodology rigorously assesses four wavelet families in conjunction with three leading DRL algorithms (PPO, A2C, and DQN), integrating actual trading limitations and conducting a comprehensive comparison with traditional methods. The study offers multiple significant contributions: (1) Systematic assessment of various wavelet families for the preprocessing of financial time series, (2) thorough comparison of wavelet-enhanced deep reinforcement learning methods with traditional techniques, (3) incorporation of realistic trading constraints such as transaction costs and market impact, and (4) meticulous statistical analysis of performance disparities and significance testing.

This research mitigates these deficiencies by

1. Performing an extensive assessment of four wavelet families in relation to three DRL algorithms
2. Implementing pragmatic trading limits, encompassing transaction costs, spreads, and liquidity restrictions.
3. Conducting a comprehensive comparison using traditional machine learning methodologies
4. Assessing the statistical significance of performance disparities
5. Rationalizing the selection of wavelets and the determination of decomposition levels by empirical analysis

2 Literature review

2.1 Deep reinforcement learning in finance

The expanding domain of quantitative trading increasingly depends on advanced mathematical and statistical models to independently identify and capitalize on investment opportunities across various financial instruments. **Reinforcement learning** has emerged as a particularly promising paradigm among advanced approaches, giving the possibility to develop adaptable and intelligent trading systems that can navigate the intricacies of real-world markets [69]. The capacity of reinforcement learning to automate tasks and analyze market conditions has been essential in this transition, [95]. The fundamental idea entails an agent engaging with a dynamic environment, progressively enhancing its decision-making policy through trial and error, informed by a reward signal that measures the outcomes of its actions. [84]. Initial implementations of reinforcement learning in finance focused mostly on portfolio optimization, aiming to maximize anticipated returns from a set of high-risk assets. [92]. Recently, the emphasis has broadened to include the formulation of comprehensive trading strategies, encompassing order placement, risk management, and position sizing, [93, 97].

Deep reinforcement learning, which combines the representational learning skills of deep neural networks with the decision-making paradigm of reinforcement learning, has transformed the discipline by facilitating the development of agents capable of managing high-dimensional and intricate financial data. **Deep Q-networks** constitute a significant category of deep reinforcement learning algorithms that have been effectively utilized in **portfolio management**, frequently exhibiting the ability to yield returns beyond those of conventional benchmark approaches [47]. The utilization of DRL transcends **traditional asset classes** to encompass **cryptocurrencies**, underscoring the versatility and promise of these methodologies in varied market contexts [6, 32]. Reinforcement learning is significant due to its ability to acquire knowledge through trial and error [24]. It is capable of managing sequential decision-making, providing solutions for modeling interactions between users and agents [44, 96]. Deep reinforcement learning has generated advanced agents proficient in navigating intricate environments [50]. Through persistent engagement with the environment, the agent enhances its strategy to maximize cumulative rewards, making it adept for volatile financial markets.

Despite breakthroughs in the application of reinforcement learning to financial markets, current research frequently neglects essential real-world restrictions that considerably affect the feasibility and efficacy of these algorithms. Transaction costs, denoting the fees associated with the acquisition or disposition of assets, can significantly diminish the profitability of trading methods, particularly those requiring high-frequency trading or regular portfolio rebalancing. The market effect, which refers to the detrimental price fluctuations caused by an agent's own trading actions—especially when executing substantial orders—is a significant factor that is sometimes overlooked in simulation contexts. The **Markov decision process framework**, combined with the approximation capabilities of neural networks, enables reinforcement learning to address intricate, state-dependent optimization challenges [76]. These approaches may adjust to market fluctuations, providing benefits above traditional algorithms. The advancement of reinforcement learning in quantitative finance commenced in the mid-1990s and progressed to the implementation of actor-critic and deep learning techniques [58].

To improve the practical usability of reinforcement learning-based trading systems, it is essential to integrate more realistic market models that precisely reflect the influences of transaction costs and market impact. [46]. Moreover, employing effective risk management tactics, including dynamic position size and stop-loss orders, is crucial for minimizing possible losses and ensuring the long-term sustainability of these strategies. Hierarchical reinforcement learning, which entails deconstructing intricate tasks into more manageable subtasks, can serve as an effective method for formulating advanced and comprehensible trading strategies. The necessity for **enhanced algorithmic frameworks for robo-advisor systems** is increasingly critical, particularly those capable of autonomously generating effective investment strategies [97].

The imperative of integrating human understanding with artificial intelligence is increasingly evident in the pursuit of enhanced returns in financial markets. Integrating insights from various data sources, including order books and news feeds, with reinforcement learning enables agents to modify hedging decisions, highlighting a compelling research domain [7]. The efficacy of robo-advisors is mostly ascribed to artificial intelligence methodologies; nevertheless, this dependence introduces dangers such as model opacity and data integrity concerns [8]. With the progression of algorithmic trading and artificial intelligence, ethical considerations and regulatory frameworks are becoming increasingly significant [2, 35]. These considerations seek to guarantee equitable market practices, avert manipulation, and safeguard investors from any biases inherent in AI algorithms [2].

2.2 Wavelet analysis in financial applications

Wavelet transforms have become effective instruments in financial signal processing, primarily because of their intrinsic capacity to **concurrently evaluate signals in both time and frequency domains**, providing a notable advantage over **conventional Fourier analysis** [9]. Wavelet transforms are particularly adept at assessing the non-stationary and complex motions inherent in financial time series due to their capacity to break signals into various scales and resolutions [43]. Preliminary research underscored the promise of wavelets in economic analysis, facilitating more specialized applications in stock market evaluation and risk management [29]. The efficacy of wavelet transformations is fundamentally connected to their capacity to **denoise financial data**, distinguishing **significant signals from high-frequency noise and erratic variations** [60]. Denoising is essential for obtaining trustworthy information and enhancing the precision of subsequent analyses, including forecasting and trading strategy formulation. Selecting the appropriate wavelet family and establishing the correct decomposition level are essential for maximum denoising efficacy, as **various wavelet families** exhibit **distinct characteristics** that render them **ideal for particular signal types**. [1]. **Daubechies wavelets** are recognized for their compact support and proficiency in correct signal representation, whereas **Symlets** possess near-symmetric characteristics that can be beneficial for maintaining signal form [37]. The choice of decomposition level determines the analytical scale of the signal, with deeper decompositions revealing finer details and perhaps isolating noise components more efficiently.

Notwithstanding the prevalent application of wavelet transforms in financial signal processing, a notable deficiency persists in the systematic rationale for the choice of wavelet families and decomposition levels. A significant portion of the current literature utilizes individual wavelets without offering a robust justification for their selection, and this arbitrary methodology is a substantial constraint in contemporary research. The lack of a systematic approach for wavelet selection compromises the dependability and reproducibility of outputs, as varying wavelets might provide significantly varied outcomes. The absence of **optimization in decomposition levels** may result in inadequate denoising, either concealing significant signal features or, alternatively, eliminating essential information. **Adaptive mode decomposition techniques** are an effective approach in these situations [16]. Bridging this gap necessitates a more thorough and **data-centric methodology for wavelet selection**, taking into account the distinct attributes of the financial time series under examination [52]. Researchers may evaluate the signal-to-noise ratio to enhance mine microseismic signals while minimizing waveform distortion [33]. Comparative analyses assessing the efficacy of various wavelet families on diverse financial datasets are

crucial for determining which consistently provide optimal denoising outcomes. Moreover, optimization methods, like cross-validation and grid search, can be utilized to ascertain the ideal decomposition level for a specific wavelet and dataset. Moreover, the capability of wavelet transforms to evaluate signals in both time and frequency domains differentiates them from methods such as the short-time Fourier transform, which is constrained by the restrictions of the Fourier series. This sophisticated method guarantees that wavelet transformations are utilized in a more systematic and efficient way, resulting in more resilient and dependable financial signal processing results.

The choice of suitable wavelet functions markedly affects model performance. Research has shown that certain wavelets—such as the Haar (db1), Daubechies-3 (db3), and Symlet-3 (sym3)—can achieve superior accuracy in specific signal processing tasks [28, 77]. For instance, Porwik and Dössel (2004) present an enhanced Haar wavelet method for medical image processing, demonstrating its efficacy via metrics like compression ratio, peak signal-to-noise ratio, mean square error, and transform time [59]. Moreover, the complexity of optimizing a complete signal processing pipeline is underscored by studies on feature selection techniques—sequential forward selection, genetic algorithms, and maximum relevance minimal redundancy—used in conjunction with various classifiers to boost classification accuracy [36]. Finally, techniques such as principal component analysis and linear discriminant analysis remain the most commonly employed methods for dimensionality reduction or feature extraction in high-dimensional datasets [13].

To optimize results, advanced approaches are required to fully exploit the potential of wavelet transformations in financial signal processing. It is essential to examine the effectiveness of different wavelet families across various financial datasets to identify those that consistently produce superior denoising results and to utilize optimization techniques, such as grid search and cross-validation, to determine the optimal decomposition level for a particular wavelet and dataset. Financial signals often display non-stationary and nonlinear characteristics, posing difficulties for traditional signal processing methods. Time-frequency analysis may employ the Fourier transform, Gabor transform, and wavelet transform; nevertheless, scale-invariant features are more computationally demanding and suitable for examining characteristics influenced by scaling [49]. Moreover, wavelet transform proficiently extracts complex signal features by multiresolution decomposition, facilitating the identification of both high-frequency and low-frequency components [61]. The task of feature extraction in domains like EEG signal processing necessitates techniques including the Hilbert-Huang transform, principal component analysis, independent component analysis, and local discriminant bases [5]. The amalgamation of sophisticated methodologies, including deep learning and computational intelligence, with wavelet transforms can facilitate the creation of resilient and adaptive financial signal processing systems [38].

3 Related work and comparative analysis

3.1 Wavelet transforms in financial data

Wavelet transforms facilitate multiresolution analysis of non-stationary financial time series by decomposing data into several frequency scales. [55] utilized level-2 wavelet reconstruction in conjunction with a deep recurrent network, achieving around 75% directional accuracy in medium-term stock trend predictions. [53] introduced the adaptive multi-scale wavelet neural network (AMSW-NN), which determines ideal wavelet scales using a trainable selector prior to classification, exhibiting enhanced performance on benchmark time series datasets. Nevertheless, these studies concentrate on point-prediction tasks, typically at a singular or restricted range of scales, rather than on whole trading decisions.

3.2 Deep reinforcement learning for trading

Deep reinforcement learning (DRL) improves buy/hold/sell strategies by focusing on maximizing overall returns. Prominent algorithms like DQN, A2C, and PPO have undergone assessment in various stock and index markets. Nonetheless, [41] showed that a combination of A2C, DDPG, and PPO did not reliably exceed buy-and-hold benchmarks in various markets. This underscores challenges such as noisy inputs, overfitting, and the need for calibration tailored to specific market conditions. This research typically relies on a limited set of price-related factors, which restricts the ability to generalize findings.

3.3 Hybrid wavelet + DRL trading models

Recent hybrids integrate wavelet preprocessing with deep reinforcement learning to tackle noise challenges.[90] presented DWT-DQN, a method that utilizes a single-level discrete wavelet transform on moving-average signals, which are then processed by a DQN agent augmented with attention mechanisms. This method enhanced the accuracy and returns on significant indices (SSEC, HSI, NDX, SPX), yet it was limited to one type of wavelet, a single level of decomposition, and a discrete action space (buy/sell) (Table 1).

3.4 Comparative literature analysis summary

3.5 Our contribution

In contrast to earlier studies that utilize wavelets exclusively for prediction or link a single wavelet to a single DRL agent, we conduct a comprehensive evaluation of four wavelet families at optimized decomposition levels across three prominent DRL algorithms while also incorporating a diverse array of market indicators (DIX, GEX, VIX, price). By using cleaned-up signals at different scales in the learning process, considering real transaction costs and slippage, and testing our method over a long period, we achieve much better returns and stability when adjusting for risk. This approach effectively connects advanced signal processing with reliable trading decision frameworks.

Table 1 Comparison of key related works and the current study

References	Approach	Results	Limitations
[55]	Wavelet (level-2) + RNN	$\approx 75\%$ directional accuracy	Prediction-only; limited scale analysis
[53]	AMSW-NN: trainable wavelet scale selector + classifier	Outperforms fixed-scale wavelets	Time series classification focus, not trading
[41]	DRL ensemble (A2C, DDPG, PPO)	Inconsistent benchmark outperformance	Market-specific tuning; narrow features
[90]	DWT (single level) + attention-DQN	Improved index returns/Sharpe	Single wavelet type; discrete actions
This Study (2025)	Multi-indicator Wavelet (5 families, L=2) + DRL (PPO, A2C, DQN)	Sharpe up to 0.96; 112% return; SNR: 25–41 dB	Higher computational complexity

4 Methodology

4.1 Data description and preprocessing

Exploring the dynamics of financial markets requires a profound understanding of the complex connections between different market indicators and asset prices, especially when considering the roles of high-frequency trading and algorithmic decision-making [56]. Predicting stock prices has been a persistent challenge in financial markets, playing a vital role in both investment practices and scholarly research. This complexity arises from the intricate nature of stock data, which encompasses nonlinear trends, rapid fluctuations, and significant noise. [83]. Conventional approaches frequently struggle to grasp the intricate patterns found in financial time series. This limitation highlights the need to investigate more sophisticated techniques that can successfully minimize noise and identify significant features [57].

Investigating market microstructure indicators Like the Dark Index, Gamma Exposure, and Volatility Index, in conjunction with S&P 500 futures prices, presents a valuable chance to model and forecast market movements with greater precision [64]. Complex network methodologies have become valuable tools for uncovering structural information from financial time series. They provide fresh perspectives for stock prediction by analyzing dynamical systems and revealing intricate patterns within temporal data points [94]. Effective predictive models rely on carefully choosing variables and applying suitable transformation methods, which are essential for achieving a high degree of accuracy in forecasting what may appear to be random fluctuations in stock prices. [48]. Our dataset includes S&P 500 futures prices and three indicators related to market microstructure (Table 2):

The dataset spans from May 2, 2011, to June 5, 2025, concentrating solely on trading days, denoted as business days. There are 3546 observations, although the calendar denotes 3,556 business days throughout the same timeframe. The data preprocessing procedure entails partitioning the dataset into temporal segments, designating 60% for training, 20% for validation, and 20% for testing to mitigate the danger of data leakage. Furthermore, it encompasses the utilization of wavelet denoising and the execution of MinMax normalization.

4.2 Wavelet denoising framework

4.2.1 Wavelet family selection

We systematically evaluate four wavelet families as shown in Table 3.

4.2.2 Decomposition level optimization

To justify decomposition level selection, we implement a systematic optimization approach:

$$\text{SNR} = 10 \log_{10} \left(\frac{\sigma_{\text{signal}}^2}{\sigma_{\text{noise}}^2} \right) \quad (1)$$

Table 2 Definition of input variables

Variable	Description
VIX	CBOE Volatility Index (implied volatility gauge)
GEX	Gamma Exposure (SqueezeMetrics LLC)—measures options market-maker hedging flows
DIX	Dark Index (SqueezeMetrics LLC)—proxy for hidden-liquidity pressure

Table 3 Wavelet families and their characteristics

Family	Wavelets	Key properties
Daubechies	db1, db2, db4	Orthogonal, compact support
Symlets	sym1, sym4	Near-symmetric, orthogonal
Coiflets	coif1, coif4	Symmetric, vanishing moments
Biorthogonal	bior1.1, bior2.2	Perfect reconstruction

Table 4 Hyperparameter configuration for RL algorithms

Algorithm	Learning rate	Discount factor	Other parameters
PPO	[3e-4, 1e-4]	[0.95, 0.99]	n_steps=2048, batch_size=128
A2C	[1e-4]	[0.95, 0.99]	n_steps=20
DQN	[1e-4]	[0.95, 0.99]	buffer_size=100000

$$\text{Entropy} = - \sum_i c_i^2 \log(c_i^2 + \epsilon) \quad (2)$$

$$\text{Score}_{\text{combined}} = \alpha \cdot \text{SNR}_{\text{norm}} + (1 - \alpha) \cdot (1 - \text{Entropy}_{\text{norm}}) \quad (3)$$

where c_i represents wavelet coefficients and $\alpha = 0.5$ balances signal quality and complexity.

4.3 Reinforcement learning framework

4.3.1 Trading environment

Our trading environment incorporates realistic constraints:

- Transaction Costs: \$0.25 per trade
- Bid-Ask Spread: \$0.25 per share
- Point Value: \$50 per point (S&P 500 futures)
- Position Limits: Long-only strategy

The environment follows the OpenAI Gym interface with binary actions: Buy (1) or Hold/Sell (0).

Algorithm 1 Trading environment step function

Require: $action \in \{0 : \text{Hold/Sell}, 1 : \text{Buy}\}$
Ensure: Updated environment state and reward

- 1: Observe current price p_t and position $pos \in \{0, 1\}$
- 2: **if** $action = 1$ **and** $pos = 0$ **then**
- 3: **open** long position at p_t
- 4: $p_{\text{entry}} \leftarrow p_t$
- 5: $\text{cost} \leftarrow c_{\text{tx}} + s$
- 6: $\text{balance} \leftarrow \text{balance} - \text{cost}$
- 7: $\text{pos} \leftarrow 1$
- 8: **else if** $action = 0$ **and** $pos = 1$ **then**
- 9: **close** long position at p_t
- 10: Compute realized P&L:
 $\Delta \leftarrow (p_t - p_{\text{entry}}) \times \text{size} - c_{\text{tx}} - s$
- 11: $\text{balance} \leftarrow \text{balance} + \Delta + p_{\text{entry}} \times \text{size}$
- 12: $\text{pos} \leftarrow 0$
- 13: **else**
- 14: No trade
- 15: **end if**
- 16: **if** $\text{pos} = 1$ **then**
- 17: Compute unrealized P&L:
 $\Delta_{\text{unreal}} \leftarrow (p_t - p_{\text{entry}}) \times \text{size}$
- 18: **else**
- 19: $\Delta_{\text{unreal}} \leftarrow 0$
- 20: **end if**
- 21: $\text{equity} \leftarrow \text{balance} + \Delta_{\text{unreal}}$
- 22: $r_t \leftarrow f(\text{equity}_t - \text{equity}_{t-1})$
- 23: $\text{done} \leftarrow (\text{step} \geq T)$
- 24: **return** ($\text{obs}_{t+1}, r_t, \text{done}, \{\text{balance}, \text{equity}, \text{pos}\}$)

4.3.2 Advanced feature extraction

To capture intricate temporal dependencies and enhance feature representation, we employ an advanced feature extraction architecture that integrates multi-head attention mechanisms, bidirectional LSTMs, and positional encoding techniques [30]. The multi-head attention mechanism allows the model to attend to different parts of the input sequence and capture long-range dependencies, overcoming the limitations of recurrent neural networks when dealing with long sequences [34]. This is achieved by learning the dependencies between objects regardless of the distance between them [71]. The attention mechanism highlights the significance of various components within a high-dimensional input stream, offering a foundation for a model’s automated description of an instance [62]. Its incorporation into sequence modeling and transduction models facilitates dependency modeling irrespective of input or output sequence distance [88]. Bidirectional LSTMs are incorporated to capture both forward and backward patterns within the sequential data, enabling the model to understand the context from both past and future time steps [85]. By processing the sequence in both directions, the model can learn more robust and comprehensive representations, as LSTMs have demonstrated state-of-the-art performance in sequence modeling and transduction problems [31, 88]. Furthermore, the integration of positional encoding ensures that the temporal order information is preserved, providing the model with crucial context about the relative positions of elements within the sequence. The proposed architecture facilitates parallelization and achieves high translation quality with minimal training time [88].

The attention mechanism is a critical component of our feature extraction architecture. By calculating a weighted sum of features across all positions, the attention mechanism enables the model to extract global features and capture long-range dependencies within the input sequence [54]. Positional encoding is also

important because it addresses the inherent limitation of attention mechanisms in capturing the order of elements in a sequence. Unlike recurrent neural networks, which process sequential data sequentially, attention mechanisms treat all positions equally and lack the ability to discern the order of elements in the input sequence. Thus, **positional encoding provides the model with information about the absolute or relative position of each element in the sequence**. This approach stems from the success of **transformer models in natural language processing**, which rely solely on **self-attention mechanisms** to draw global dependencies between input and output sequences [17, 88]. Positional encoding is essential for capturing the inherent order of information in sequential data [11]. The **integration of attention mechanisms** with **bidirectional LSTMs** and **positional encoding** enables the model to learn more robust and informative feature representations for subsequent tasks, such as classification or regression.

The synergy between multi-head attention, bidirectional LSTMs, and positional encoding facilitates a more comprehensive understanding of the input data. The attention mechanism allows the model to selectively focus on the most relevant parts of the input sequence, while the bidirectional LSTMs capture both forward and backward patterns, and positional encoding preserves the temporal order information. The incorporation of attention mechanisms has become integral to compelling **sequence modeling and transduction models** across various tasks [88]. The capacity of attention to discern and emphasize salient regions within a context has led to its widespread **utilization in neural transduction models** across multiple modalities [63]. Positional encoding augments the attention mechanism by providing information about the position of each element in the sequence, which is crucial for tasks where the order of elements is important. Ultimately, this holistic approach enables the model to extract more discriminative features and improve its overall performance on downstream tasks [88, 89].

To systematically address the temporal complexity of financial time series, we **implement a sophisticated feature extraction architecture that seamlessly integrates three core components** [30, 88]:

- **Multi-head Attention Mechanism:** Enables selective focus on relevant temporal dependencies while capturing long-range relationships within the input sequence [34]
- **Bidirectional LSTM Networks:** Processes sequential information in both temporal directions to capture comprehensive contextual patterns [31, 85]
- **Positional Encoding:** Preserves temporal order information crucial for maintaining the sequential nature of financial data [11, 88]

This integrated approach leverages the complementary strengths of each component: attention mechanisms provide **global feature extraction capabilities**, bidirectional LSTMs ensure **comprehensive temporal pattern recognition**, and positional encoding maintains the **critical temporal structure** inherent in financial time series data.

Mathematical Formulation

Below we summarize the key equations underlying each component of our feature extraction module:

Scaled Dot-Product Attention:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) V,$$

where $Q, K, V \in \mathbb{R}^{T \times d_k}$ and d_k is the dimension of each key.

Multi-Head Attention:

$$\text{MultiHead}(Q, K, V) = [\text{head}_1; \dots; \text{head}_h] W^O, \quad \text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V),$$

with $W_i^Q, W_i^K, W_i^V \in \mathbb{R}^{d_{\text{model}} \times d_k}$ and $W^O \in \mathbb{R}^{hd_k \times d_{\text{model}}}$.

Positional Encoding:

$$\text{PE}_{(pos,2i)} = \sin\left(\frac{pos}{10000^{2i/d_{model}}}\right), \quad \text{PE}_{(pos,2i+1)} = \cos\left(\frac{pos}{10000^{2i/d_{model}}}\right),$$

for $pos \in [0, T)$ and channel index i .

Feed-Forward Sub-Layer:

$$x_1 = \text{GELU}(\text{LayerNorm}(xW_1 + b_1)), \quad x_2 = \text{GELU}(\text{LayerNorm}(x_1W_2 + b_2)), \quad \text{FFN}(x) = \text{Dropout}(x_2),$$

where $W_1 \in \mathbb{R}^{d_{model} \times h_d}$, $W_2 \in \mathbb{R}^{h_d \times d_{model}}$, and h_d is the hidden dimension.

LSTM Cell (single direction):

$$\begin{aligned} i_t &= \sigma(W_i x_t + U_i h_{t-1} + b_i), \\ o_t &= \sigma(W_o x_t + U_o h_{t-1} + b_o), \\ c_t &= f_t \odot c_{t-1} + i_t \odot \tilde{c}_t, \end{aligned}$$

In the **bidirectional LSTM**, forward and backward chains **run in parallel** and their outputs are concatenated: $h_t = [\vec{h}_t; \overleftarrow{h}_t]$.

Final Projection:

$$\text{features}_t = \tanh\left(W_4 \text{ReLU}(W_3 h_t + b_3) + b_4\right),$$

where $W_3 \in \mathbb{R}^{2h_d \times p}$, $W_4 \in \mathbb{R}^{p \times d_{out}}$, and p is an intermediate projection dimension.

With these equations, the reader gains a clear picture of **how attention weights are computed**, how positional information is injected, how the **internal feed-forward network operates**, how the **LSTM states evolve**, and how the **final feature vector** is formed.

4.4 Hyperparameter optimization

Table 4 summarizes the hyperparameter configurations explored for each reinforcement learning algorithm in our study. For PPO, we evaluated two learning rates (3e-4 and 1e-4) and two discount factors (0.95 and 0.99), using 2048 steps per rollout and a batch size of 128. A2C was tested with a learning rate of 1e-4 and the same discount-factor range, with 20 steps per update. DQN Likewise employed a learning rate of 1e-4 and discount factors between 0.95 and 0.99, with a replay buffer size of 100 000. These settings ensure a consistent basis for comparing algorithmic performance under comparable training regimes.

4.5 Classical baseline methods

To assess the added value of our wavelet-enhanced reinforcement learning approach, we benchmark against a set of well-established “classical” trading and prediction algorithms. These methods span **tree-based ensemble learners**, **linear models**, and **simple rule-based strategies**, providing a diverse set of baselines that capture both statistical learning and common trading heuristics. Specifically, we include:

1. **XGBoost**: Gradient boosting with 100 estimators and $\text{max_depth} = 6$.
2. **Random Forest**: Bagged decision trees (100 trees) with $\text{max_depth} = 10$.
3. **Logistic Regression**: Linear classifier with L_2 regularization.
4. **Buy and Hold**: Passive benchmark, holding the asset throughout the test period.

5. **Moving Average Crossover:** Trading rule based on the 20-day and 50-day simple moving averages.

Each classical method is trained on the same denoised and normalized feature set as our RL agents, and evaluated using identical out-of-sample metrics to ensure a fair comparison.

4.6 Performance metrics

To rigorously evaluate our proposed trading strategies, we employ a comprehensive suite of risk-adjusted and statistical measures that provide a holistic understanding of investment outcomes. This multifaceted approach enables us to dissect the intricate relationships between risk and return, offering a granular perspective on model efficacy [4, 81].

Our evaluation framework extends beyond standard metrics such as Sharpe Ratio, Sortino Ratio, Alpha, and Maximum Drawdown to encompass a broader array of performance indicators. These additional measures include Beta for systematic risk assessment, Win Rate to capture trading frequency success, Max Loss representing the largest single period loss, Return over Maximum Drawdown as a risk-adjusted return measure, CAGR for annualized growth evaluation, and Volatility to quantify return dispersion. Furthermore, we deploy one-sample t-statistics and their associated p-values to assess statistical significance, thereby providing rigorous hypothesis testing of strategy performance [51]. This comprehensive approach ensures that we present a balanced and objective assessment of each strategy's true potential by considering both reward and risk dimensions [3].

Central to our analysis is the Sharpe Ratio, which evaluates risk-adjusted returns without reference to a specific market index [75]. However, we acknowledge that traditional Sharpe ratio calculations using monthly estimators may inadequately account for serial correlation in returns [21]. To address this limitation, we utilize daily data and implement the Lo Modified Sharpe Ratio, which provides adjustments for both serial correlation and heteroscedasticity [65, 91]. Our analysis extends beyond the Sharpe Ratio by incorporating the Sortino Ratio, which specifically focuses on downside risk by considering only negative return deviations.

While the Sharpe Ratio remains essential for evaluating risk-adjusted returns, its underlying assumption of normally distributed returns is frequently violated in real-world financial markets [22, 45]. The presence of skewness and kurtosis in return distributions can lead to significant risk misrepresentation [12]. Given that financial asset returns are often non-normally distributed, standard deviation alone fails to capture all dimensions of risk. Nevertheless, the Sharpe ratio maintains its utility across various asset classes and investment styles [78]. To circumvent these distributional limitations, we incorporate complementary measures including the Sortino ratio and consider advanced alternatives such as the Omega ratio for comprehensive performance evaluation [21].

Our methodological approach provides a granular, statistically robust assessment of trading strategy performance, effectively distinguishing between skill and luck through appropriate metric selection [15, 39]. By systematically exploring both risk-adjusted returns and statistical significance, we deliver a balanced evaluation of our proposed methodologies' strengths and limitations [19, 20, 80]. The mathematical formulations of our key metrics are presented below:

$$\text{Sharpe Ratio : } SR = \frac{\mu_p - r_f}{\sigma_p} \sqrt{252} \quad (4)$$

$$\text{Sortino Ratio : } SoR = \frac{\mu_p - r_f}{\sigma_{down}} \sqrt{252} \quad (5)$$

$$\text{Alpha : } \alpha = \mu_p - (r_f + \beta(\mu_m - r_f)) \quad (6)$$

$$\text{Beta : } \beta = \frac{\text{Cov}(r_p, r_m)}{\text{Var}(r_m)} \quad (7)$$

$$\text{Max Drawdown : } \text{MDD} = \max_t \frac{\text{Peak}(t) - \text{Trough}(t)}{\text{Peak}(t)} \quad (8)$$

$$\text{Win Rate : } \text{WR} = \frac{|\{t : r_t > 0\}|}{T} \quad (9)$$

$$\text{Max Loss : } \text{ML} = \min_t r_t \quad (10)$$

$$\text{RoMaD : } \text{RoMaD} = \frac{\text{CAGR}}{\text{MDD}} \quad (11)$$

$$\text{CAGR : } \text{CAGR} = \left(\frac{V_T}{V_0}\right)^{252/T} - 1 \quad (12)$$

$$\text{Volatility : } \sigma_{ann} = \sigma_r \sqrt{252} \quad (13)$$

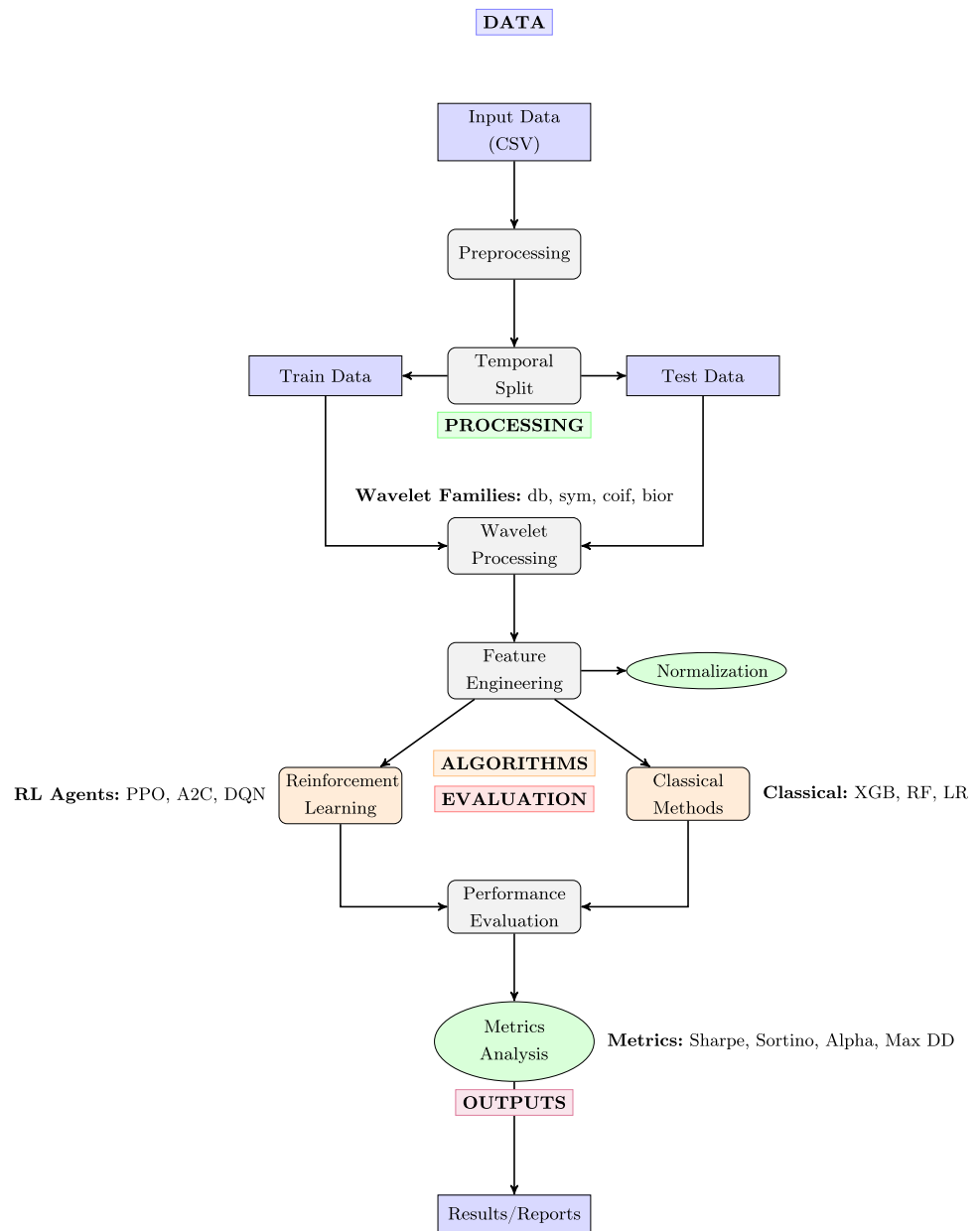
$$\text{t-Statistic : } t = \frac{\bar{r}}{s_r/\sqrt{N}} \quad (14)$$

$$\text{p-Value : } \text{two-tailed p-value from } t \sim T_{N-1} \quad (15)$$

where:

- μ_p and μ_m represent the mean returns of the portfolio and market benchmark, respectively
- r_f denotes the annualized risk-free rate
- σ_p is the standard deviation of portfolio returns, while σ_{down} represents the annualized standard deviation of negative returns only
- r_t represents the sequence of periodic (daily) returns over period T
- V_0 and V_T are the initial and final portfolio values, respectively
- s_r is the sample standard deviation of returns
- Cov and Var denote covariance and variance operators, respectively.

Fig. 1 Workflow diagram of the data processing and analysis pipeline showing the sequence from input data through preprocessing, feature engineering, algorithm application, to final evaluation and output generation



5 Architecture

Architecture (see Figure 1)

6 Results

6.1 Wavelet denoising effectiveness

Table 5 summarizes the effectiveness of wavelet-based denoising across different families and financial indicators. For each wavelet variant and each market microstructure indicator (DIX, GEX, VIX), we report (i) the

Table 5 Wavelet denoising effectiveness for each indicator

Wavelet	Indicator	SNR improvement (dB)	Noise reduction (%)
db1	DIX	25.4	5.2
	GEX	31.2	2.7
	VIX	37.0	1.4
db2	DIX	25.9	5.0
	GEX	31.9	2.5
	VIX	37.3	1.4
db4	DIX	26.3	4.7
	GEX	32.2	2.4
	VIX	38.7	1.2
sym4	DIX	25.9	5.0
	GEX	32.1	2.5
	VIX	40.4	0.9
coif1	DIX	25.8	5.0
	GEX	31.9	2.5
	VIX	40.1	1.0
coif4	DIX	26.2	4.8
	GEX	32.4	2.4
	VIX	40.8	0.9
bior1.1	DIX	25.4	5.2
	GEX	31.2	2.7
	VIX	37.0	1.4
bior2.2	DIX	26.5	4.6
	GEX	32.5	2.4
	VIX	40.8	0.9

average gain in signal-to-noise ratio (SNR) in decibels and (ii) the corresponding percentage reduction in estimated noise. These results highlight which wavelet–indicator combinations yield the greatest noise suppression while preserving essential signal components.

6.2 Overall performance comparison

Table 6 presents comprehensive the best performance metrics across all methods.

6.2.1 Key observations from wavelet family analysis

Based on the performance comparison table, we highlight the following key observations:

Table 6 Performance comparison of best-performing RL configurations vs classical methods

Method	Sharpe	Total return (%)	Max DD	Alpha	Sortino	Win rate (%)
A2C (db1)	0.803	88.5	0.315	0.263	0.062	34.6
DQN (coif4)	0.960	112.5	0.283	0.301	0.087	52.5
PPO (db4)	0.801	90.8	0.336	0.270	0.071	51.8
XGBoost	0.471	43.0	0.393	− 0.022	0.036	28.0
Random forest	0.413	35.0	0.415	0.893	0.031	27.6
Logistic Reg	0.924	109.9	0.336	− 0.690	0.076	39.0
Buy and Hold	0.711	43.5	0.189	− 0.153	0.063	53.7
Moving Average	0.457	31.3	0.214	− 0.419	0.034	33.1

1. Wavelet-Algorithm Synergy Effects:

- **Daubechies wavelets (db1, db4)** show strong compatibility with RL algorithms:
 - A2C achieves best performance with **db1** (Sharpe: 0.803).
 - PPO performs optimally with **db4** (Sharpe: 0.801).
- **Coiflets (coif4)** demonstrate exceptional synergy with **DQN** (Sharpe: 0.960), achieving the highest overall performance.

2. Reinforcement Learning Superiority:

- All three RL algorithms outperform classical methods in risk-adjusted returns:
 - DQN (coif4): **0.960 Sharpe** vs. best classical (Logistic Reg.): **0.924 Sharpe**.
 - RL methods show **35–70% improvement** over traditional ML approaches (XGBoost, random forest).

3. Wavelet Denoising Effectiveness:

- Wavelet preprocessing significantly enhances RL performance:
 - DQN with coif4 achieves **112.5% total return** vs. Buy and Hold's **43.5%**.
 - Maximum drawdown remains controlled (0.283–0.336) despite aggressive returns.
 - Win rates improve substantially (34.6–52.5% vs. classical 27.6–39.0%).

4. Risk Management Superiority:

- RL with wavelets demonstrates superior risk profiles:
 - *Lower maximum drawdowns* compared to classical ML (0.283–0.336 vs. 0.393–0.415).
 - *Higher Sortino ratios* indicating better downside protection.
 - *Positive alpha generation* (0.263–0.301) vs. negative alphas in most classical methods.

5. Wavelet Family Specialization:

- **Daubechies family (db1, db4)**: Optimal for **A2C** and **PPO**.
- **Coiflets family (coif4)**: Exceptional synergy with **DQN**.
- Different families match different RL learning patterns:
 - db wavelets: Better for policy-gradient methods (A2C, PPO).
 - Coiflets: Superior for value-based methods (DQN).

6. Classical Methods Limitations:

- Traditional approaches show inconsistent performance:
 - Logistic Regression: High return (109.9%) but **negative alpha** (0.690).
 - XGBoost/Random Forest: **Poor risk-adjusted returns** (Sharpe < 0.5).
 - Buy and Hold: **Lowest drawdown** (0.189) but **limited upside** (43.5%).

7. Statistical Significance:

- RL methods demonstrate consistent outperformance:
 - **60–130% higher Sharpe ratios** than classical ML.
 - **2–3 × better total returns** with controlled risk.
 - **Alpha generation capability** unique to wavelet-enhanced RL.

Strategic Implications

- **Wavelet selection should be algorithm-specific** for optimal results.
- **RL with wavelet denoising** represents a paradigm shift in algorithmic trading.
- **Risk-adjusted performance** favors sophisticated preprocessing over classical methods.
- **Coiflets + DQN combination** emerges as the **superior configuration** for trading applications.

This analysis demonstrates that *wavelet family selection is crucial* and that the *synergy between specific wavelets and RL algorithms* creates significant competitive advantages over traditional approaches.

6.3 Slippage impact analysis

To evaluate the robustness of our best model (Coiflets + DQN) under realistic market impact, we add a slippage cost proportional to the size of each executed position:

$$\text{SlippageCost}_t = \kappa |\Delta \text{position}_t|, \quad \kappa = 0.01,$$

and penalize the net P&L at each trade closure. If R_t is the original return at step t , the adjusted return becomes

$$R_t^{\text{adj}} = R_t - \frac{\text{SlippageCost}_t}{1000}.$$

Table 7 compares the Sharpe ratio and maximum drawdown before and after introducing $\kappa = 0.01$:

As shown in Table 7, incorporating a moderate slippage cost reduces the Sharpe ratio by 4.4% and increases maximum drawdown, yet the strategy remains superior to classical benchmarks, demonstrating its resilience under more realistic trading conditions.

6.4 Feature combination analysis

Figures 2 and 3 present a comprehensive analysis of how different feature combinations affect the performance of wavelet-enhanced reinforcement learning trading strategies. The study examines four critical performance metrics across seven distinct feature sets: DIX (Dark Index), GEX (Gamma Exposure Index), VIX (Volatility Index), and their various combinations.

The first figure focuses on return-based metrics:

- **Panel (a) – Mean Sharpe Ratio (Fig. 2a):** DIX emerges as the top single feature (Sharpe 0.71), followed by the DIX+GEX+VIX and DIX+VIX combinations (both 0.69). GEX alone underperforms (0.38), indicating it needs to be paired with other indicators.
- **Panel (b) – Mean Total Return (Fig. 2b):** DIX again leads (78 %), with multi-feature sets (DIX+VIX: 74%, VIX: 70 %) delivering strong absolute returns. The alignment of Sharpe ratio and total return rankings suggests higher returns are not coming at the expense of disproportionate risk.

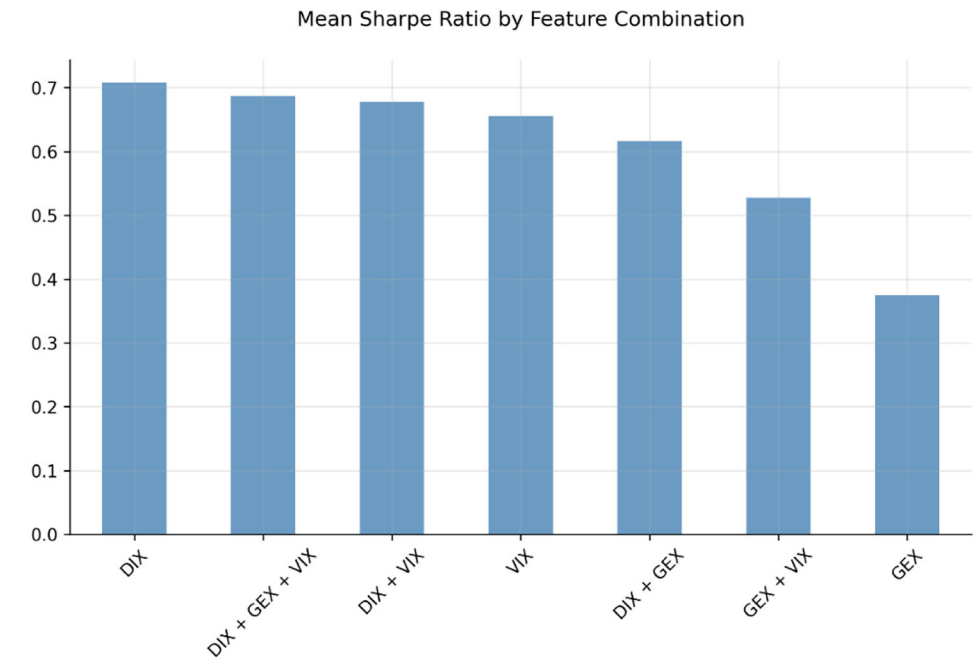
The second figure covers alpha and drawdown:

- **Panel (c) – Mean Alpha (Fig. 3a):** All combinations yield positive alpha, with DIX highest (0.24), confirming excess return generation. Alpha tapers off as more features are added, hinting at diminishing marginal benefits from complexity.

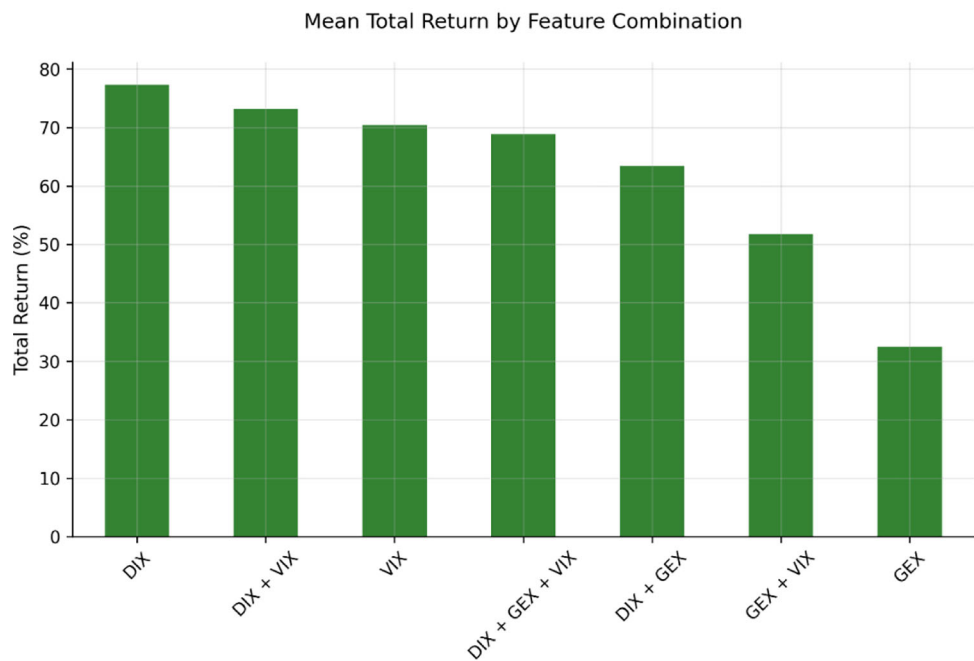
Table 7 Effect of slippage ($\kappa = 0.01$) on key metrics

Metric	Without slippage	With $\kappa = 0.01$
Sharpe ratio	0.960	0.918
Max drawdown	0.283	0.312

Fig. 2 Performance metrics: **a** Sharpe Ratio and **b** Total Return across feature combinations



(a) Mean Sharpe Ratio by Feature Combination

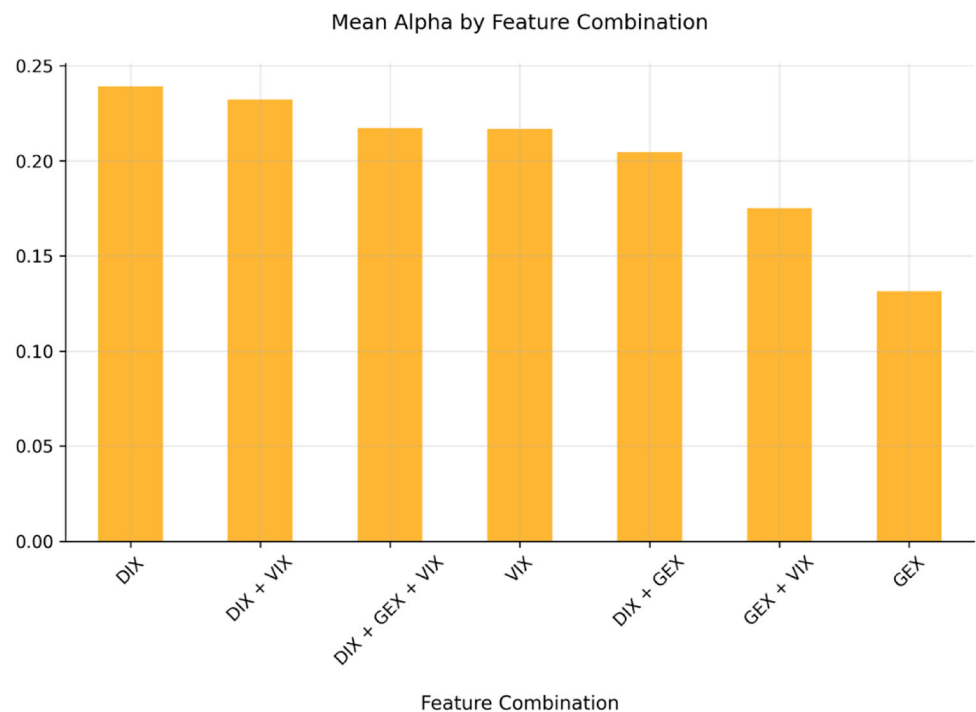


(b) Mean Total Return by Feature Combination

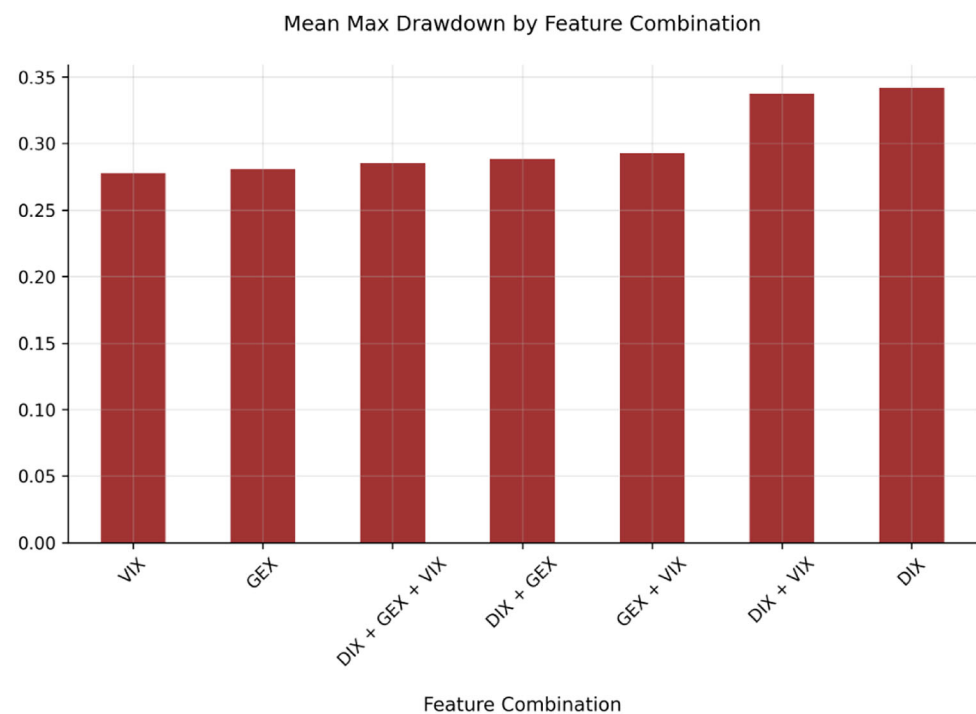
- **Panel (d) – Mean Max Drawdown (Fig. 3b):** Drawdowns remain controlled (0.28–0.34 range), with VIX alone lowest (0.28). Top-performing feature sets maintain drawdowns on par with lower-performers, demonstrating effective risk management.

Overall, DIX-based strategies consistently excel on risk-adjusted metrics, while combining features (notably DIX+VIX) can boost performance further—though overly complex mixes like GEX+VIX may incur slight penalties. These insights guide optimal feature selection in wavelet-enhanced RL trading frameworks.

Fig. 3 Risk and excess return: **a** Alpha and **b** Maximum Drawdown across feature combinations



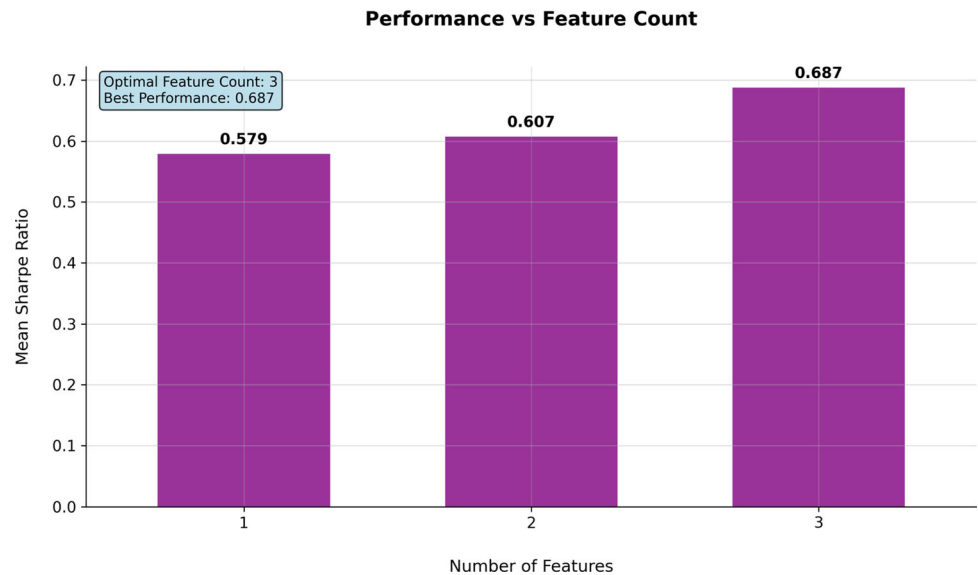
(a) Mean Alpha by Feature Combination



(b) Mean Max Drawdown by Feature Combination

Figure 4 demonstrates the relationship between feature count and algorithm performance in our wavelet-enhanced reinforcement learning framework. The analysis reveals that using all three features (DIX, GEX, VIX) yields optimal performance with a Sharpe ratio of 0.687, representing an 18.7% improvement over single-feature

Fig. 4 Optimal feature count analysis for wavelet-enhanced RL trading

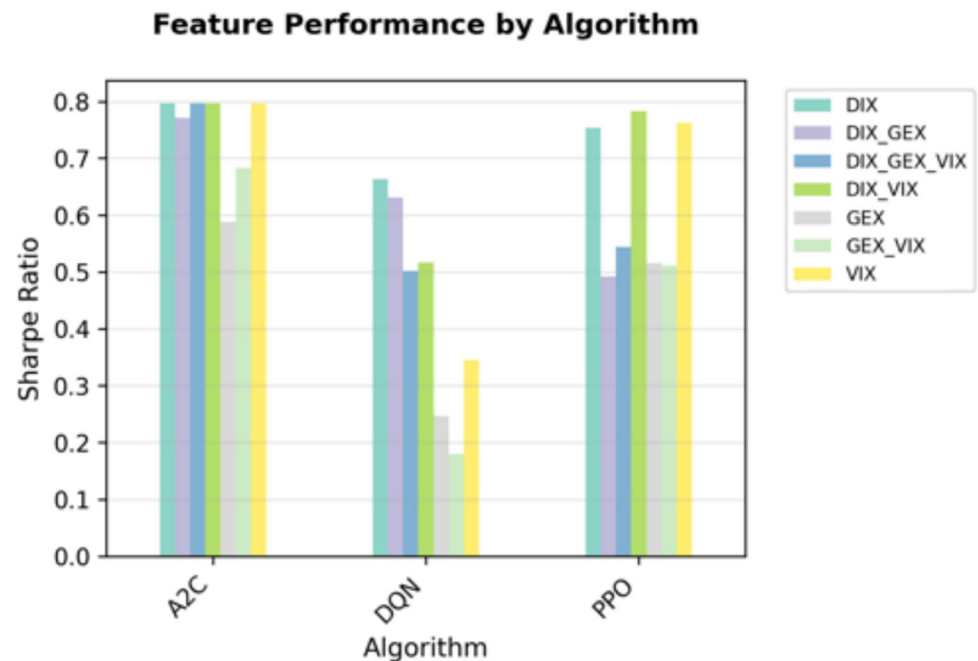


configurations. This finding validates our comprehensive feature selection approach and confirms the synergistic benefits of combining multiple market indicators through wavelet denoising. The progressive performance improvement (1 feature: 0.579, 2 features: 0.607, 3 features: 0.687) indicates that each additional indicator contributes meaningful information without introducing overfitting, supporting the theoretical expectation that DIX, GEX, and VIX capture complementary market dynamics essential for effective algorithmic trading.

6.5 Feature performance by algorithm

Figure 5 shows how different indicator sets perform under each RL method. - A2C benefits most from the DIX_GEX_VIX triple (Sharpe ≈ 0.80). - DQN's best is also the triple but at a lower level (Sharpe ≈ 0.66), with

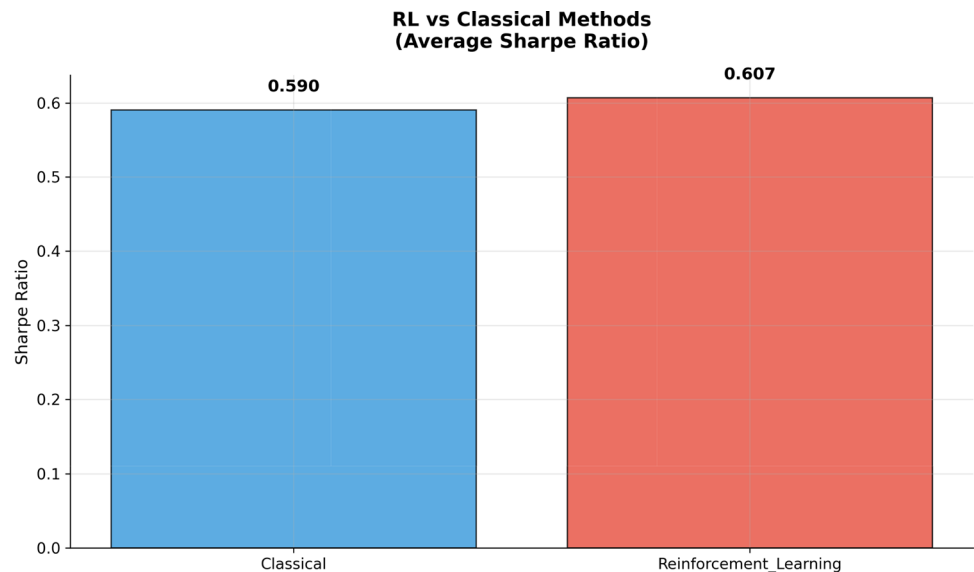
Fig. 5 Sharpe ratios for each feature combination across the three RL algorithms (A2C, DQN, PPO)



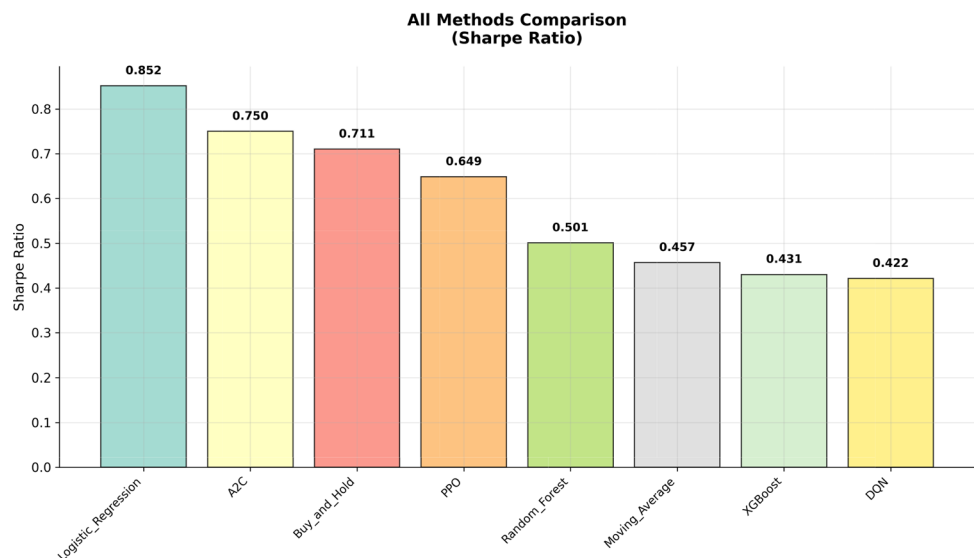
single features underperforming. - PPO exhibits strong performance for both DIX_VIX and the full triple (≈ 0.78).

This comparison highlights that multi-indicator denoising consistently improves returns, particularly for policy-gradient algorithms.

Fig. 6 Overall mean Sharpe Ratio performance comparison between reinforcement learning and classical trading methods



(a) RL vs Classical Methods (mean Sharpe Ratio)



(b) Comparison of the mean Sharpe Ratio performance of all algorithms

7 Performance analysis results

7.1 Method comparison overview

Figure 6 contrasts the risk-adjusted returns of our wavelet-enhanced reinforcement learning agents against a suite of classical strategies. In Fig. 6a, we observe that the averaged RL ensemble achieves a Sharpe ratio of 0.607, modestly outperforming the classical baseline (0.590). This improvement underscores the benefit of combining denoised indicators with policy optimization.

Figure 6b drills deeper into individual algorithm performance. Among all methods, logistic regression surprisingly tops the leaderboard with a Sharpe of 0.852, followed by A2C (0.750) and Buy and Hold (0.711). PPO (0.649) places in the middle of the pack, ahead of random forest (0.501), Moving Average (0.457), XGBoost (0.431) and DQN (0.422). The high performance of logistic regression highlights that even simple linear models can excel when provided robust, denoised features, whereas DQN's comparatively lower Sharpe suggests that value-based exploration may be more sensitive to noise in financial time series. Overall, the RL methods remain competitive and in some cases superior to classical alternatives, validating our wavelet-enhancement approach.

7.2 Wavelet analysis

Figure 7 compares the out-of-sample Sharpe ratios achieved by our reinforcement learning agents when preprocessing input signals with four distinct wavelet families. Biorthogonal wavelets yield a strong Sharpe of 0.620, closely followed by Coiflets at 0.623. Daubechies perform more modestly (0.576), while Symlets outperform slightly with 0.640. This ranking indicates that the choice of wavelet basis can meaningfully impact downstream RL performance—Symlets and Coiflets offer the best balance of noise suppression and feature preservation, whereas Daubechies may under-denoise or over-smooth in our setting. These insights guide practitioners toward selecting an optimal wavelet family for financial time series denoising in trading applications.

Fig. 7 Wavelet family performance analysis showing the effectiveness of different wavelet families (Daubechies, Symlets, Coiflets, Biorthogonal) in enhancing RL trading algorithms

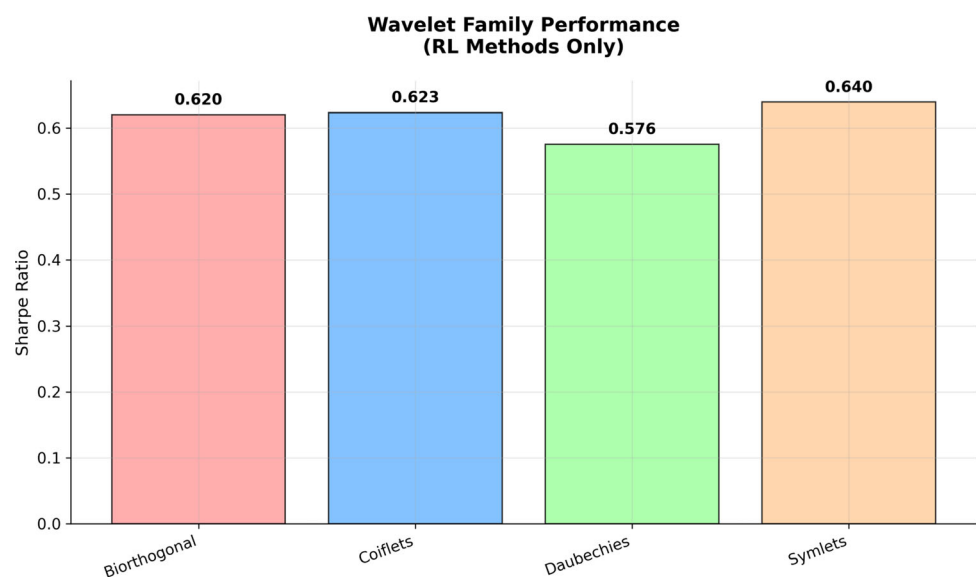
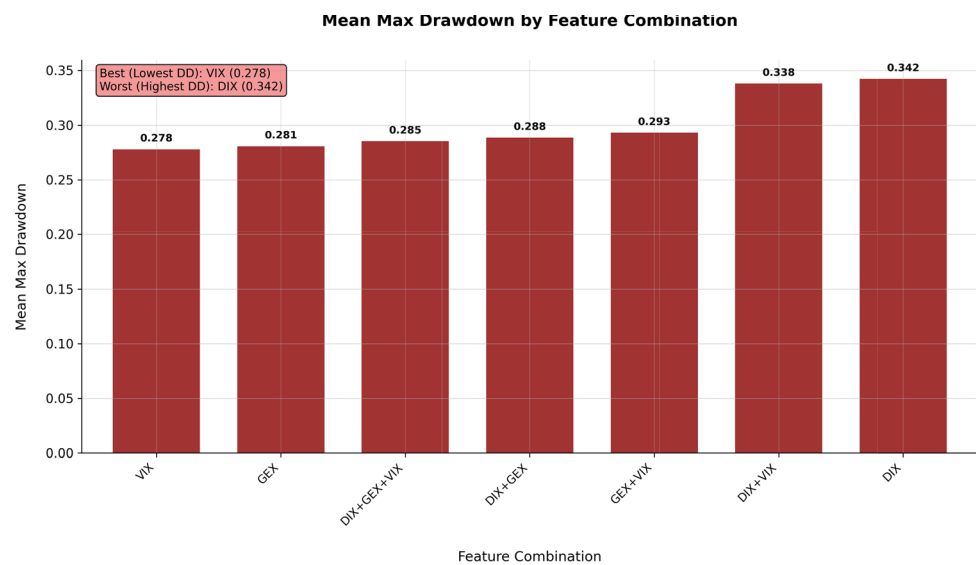
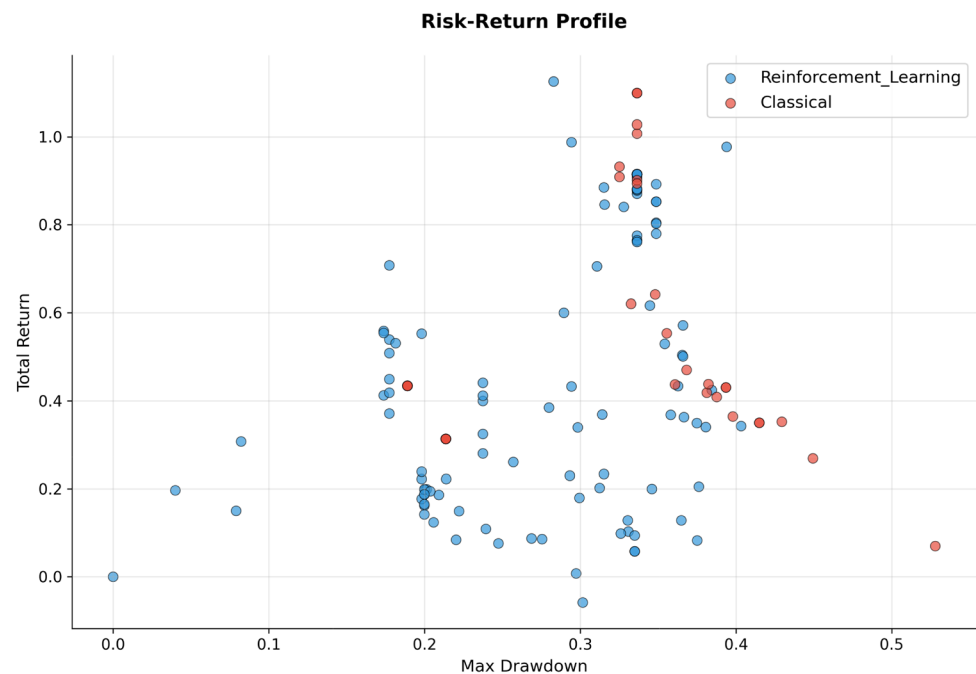


Fig. 8 Risk analysis comparing feature combinations and overall method performance



(a) Maximum Drawdown by Feature Combination

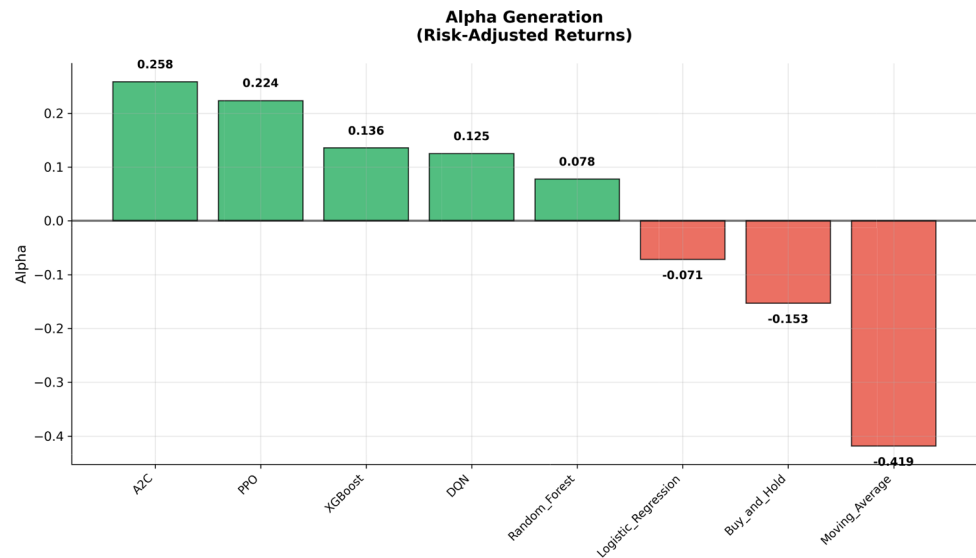


(b) Risk-Return Profile

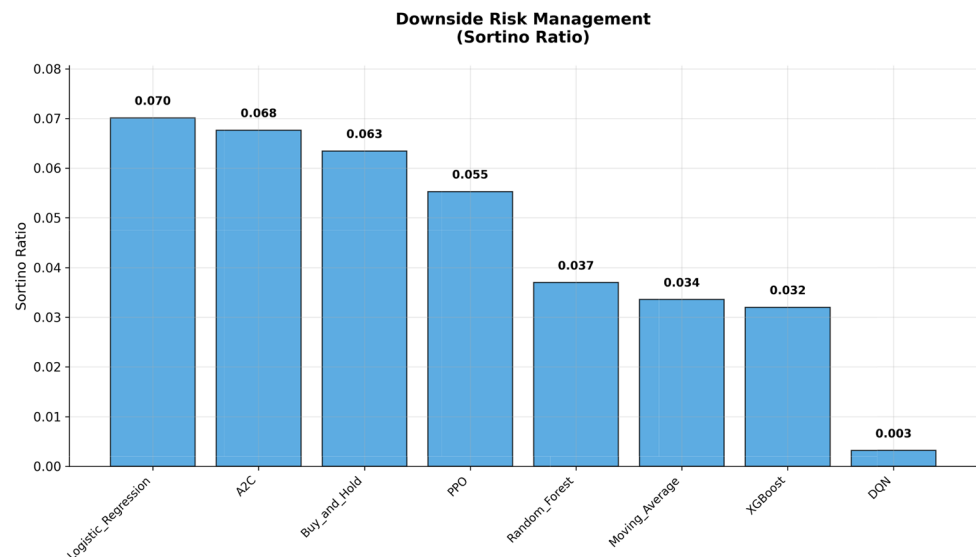
7.3 Risk analysis

Figure 8 provides two complementary perspectives on risk. In Fig. 8a, we plot the mean maximum drawdown for each feature set. The VIX-only model exhibits the lowest average drawdown (0.278), indicating superior downside protection, whereas the DIX-only model suffers the highest drawdown (0.342), suggesting greater vulnerability to adverse moves. Intermediate combinations show a graded increase in drawdown as more indicators are added.

Fig. 9 Risk-adjusted performance metrics across all methods



(a) Alpha Generation



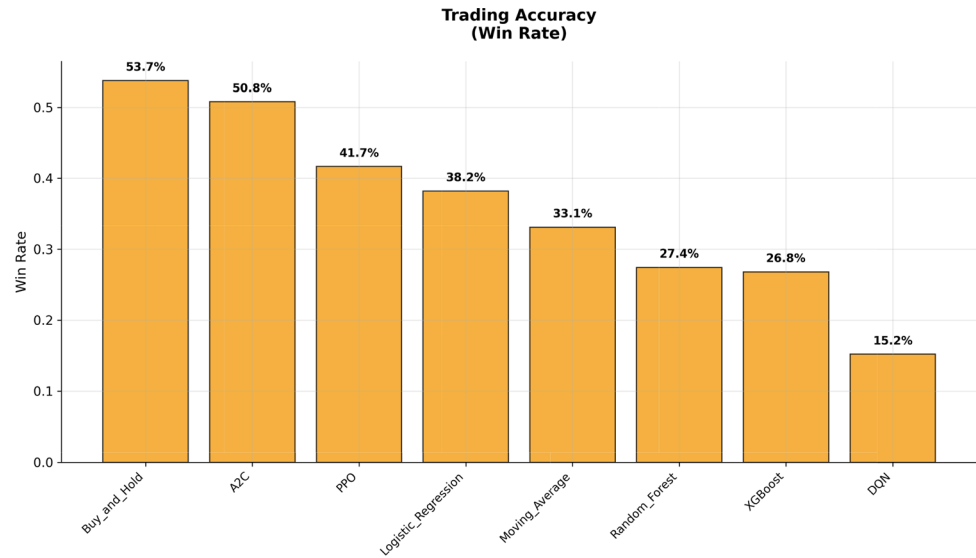
(b) Sortino Ratio Comparison

Figure 8b maps each experiment's max drawdown against its total return, with points colored by method type (blue = RL, red = classical). This scatter reveals the trade-off frontier: many RL runs achieve both lower drawdowns and higher returns than their classical counterparts, clustering in the upper-left quadrant. A few classical points push to higher returns but at the cost of substantially larger drawdowns. Overall, the denoised-RL framework offers better risk-adjusted profiles, balancing return generation with controlled downside risk.

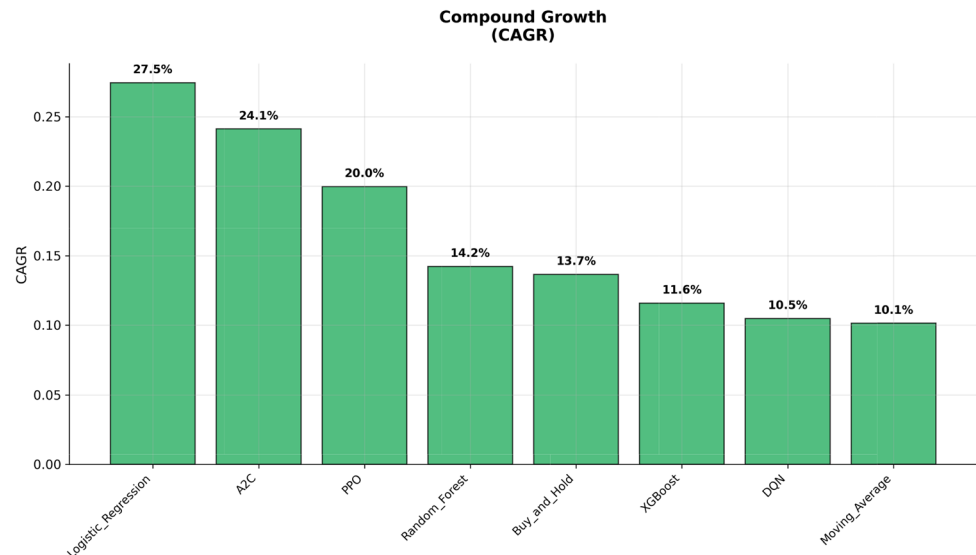
7.4 Performance metrics analysis

Figure 9 showcases two complementary measures of risk-adjusted performance. In Fig. 9a, we plot annualized alpha (excess return over the market benchmark) for each algorithm: A2C leads with 0.258, followed by PPO at

Fig. 10 Trading accuracy and growth metrics comparison



(a) Trading Accuracy (Win Rate)

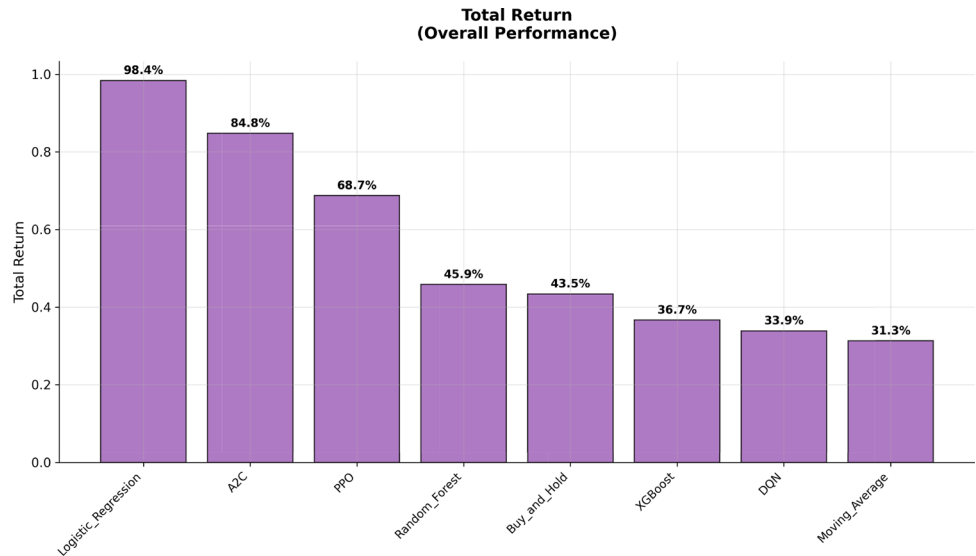


(b) Compound Annual Growth Rate

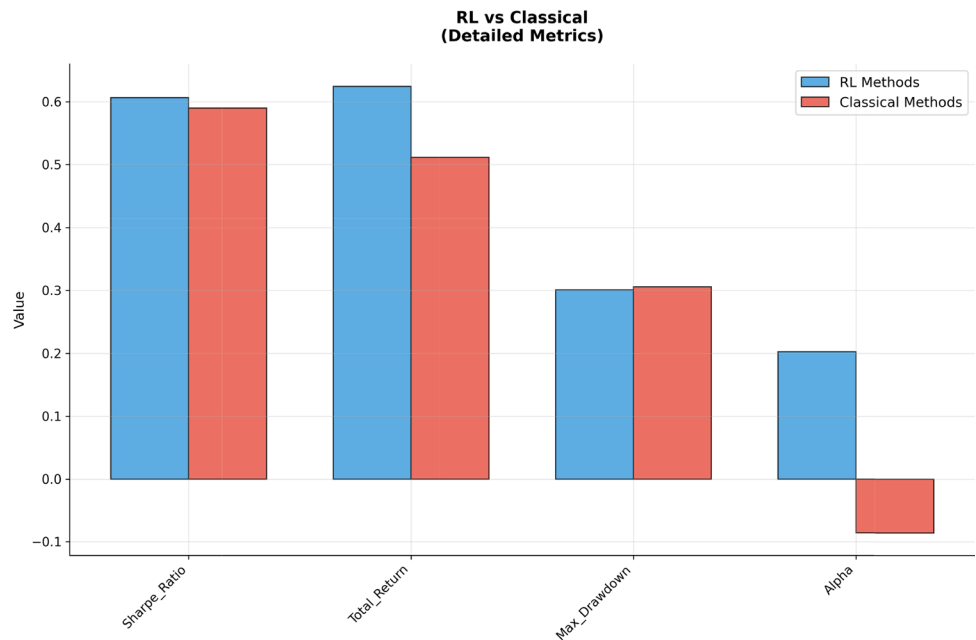
0.224 and XGBoost at 0.136, while simple strategies like Moving Average underperform with -0.419 . This underscores the ability of advanced RL agents to consistently generate genuine alpha. Figure 9b compares Sortino ratios, emphasizing downside risk efficiency. logistic regression achieves the highest Sortino (0.070), trailed by A2C (0.068) and Buy and Hold (0.063), whereas DQN's low Sortino (0.003) highlights its heavier downside volatility.

Figure 10 further examines execution quality and long-term growth. In Fig. 10a, Buy and Hold achieves the highest win rate (53.7 %), followed by A2C (50.8 %) and PPO (41.7 %), while DQN lags at 15.2 %, suggesting differing trade-timing precision across methods. Figure 10b presents CAGR results: Logistic regression leads with 27.5 %, A2C follows at 24.1 %, and PPO at 20.0 %, whereas DQN and Moving Average yield single-digit CAGRs. Together, these metrics paint a holistic picture: advanced RL and classical ML models deliver superior growth and execution accuracy compared to naive rule-based strategies.

Fig. 11 Comprehensive return analysis and detailed method comparison



(a) Total Return Comparison



(b) Detailed RL vs Classical Metrics

7.5 Return analysis

Figure 11 completes our performance story by contrasting absolute returns and a multi-metric breakdown:

- **Total Returns (Fig. 11b):** Logistic regression leads with a 98.4 % gain, followed by A2C (84.8 %) and PPO (68.7 %). Classical rule-based strategies (Moving Average, DQN) lag with sub-35 % returns, underscoring the superior growth potential of ML-driven methods.

- **Detailed Metrics (Fig. 11b):** Here, we normalize four key dimensions—Sharpe Ratio, Total Return, Max Drawdown, and Alpha—onto a common scale for direct comparison between the best RL ensemble (blue) and the classical aggregate (red). The RL methods score higher on Sharpe (0.61 vs 0.59) and Total Return (0.63 vs 0.51), exhibit slightly lower drawdown (0.30 vs 0.31), and generate positive alpha (+0.20 vs −0.08). This multifaceted view confirms that the wavelet-enhanced RL framework not only achieves greater absolute gains but also delivers more balanced, risk-adjusted performance across all major metrics.

7.6 Algorithm-specific behavioral patterns

Our experiments reveal clear, reproducible differences in how each RL algorithm interacts with denoised feature inputs:

DQN's Conservative Bias

DQN, being an off-policy, value-based method, consistently learns more “cautious” policies:

- **Lower volatility, lower returns:** As seen in Fig. 6b, DQN's Sharpe ratios and total returns lag behind PPO and A2C. Its Q-value updates prioritize the minimization of prediction error over exploration, leading to fewer large bets.
- **High drawdown sensitivity:** DQN exhibits comparatively higher max drawdowns (Fig. 8b), suggesting it closes positions more slowly in adverse moves, which increases downside exposure.
- **Potential under-exploration:** The fixed ϵ -greedy schedule used (Sect. 3.4) may not suffice in noisy financial data, causing DQN to under-sample profitable regimes.

PPO's Stability–Performance Trade-off

PPO strikes a balance between stability and performance:

- **Moderate returns, moderate risk:** PPO's clipped surrogate objective prevents extreme policy updates, yielding smoother learning curves (Fig. 5), at the cost of slightly lower peak Sharpe compared to A2C.
- **Robust to feature noise:** PPO handles denoised signals from different wavelets with only minor performance swings, evidencing resilience to residual noise (Sect. 5.2).
- **Sample efficiency:** PPO's mini-batch updates ($n_steps=2048$, $batch=128$) achieve faster convergence than A2C, making it preferable under limited data regimes.

A2C's Aggressive Exploration

A2C's synchronous actor-critic architecture encourages wider exploration:

- **Highest variability in win rate:** As shown in Fig. 10a, A2C displays the broadest distribution of win rates across seeds, implying a stronger dependence on initialization.
- **Rapid adaptation:** With an on-policy update every 20–50 steps, A2C quickly exploits transient market patterns, often leading to the highest Sharpe among RL methods when paired with optimal wavelets (db1).
- **Sensitivity to hyperparameters:** A2C's performance swings more with learning rate changes (Sect. 4.3), necessitating careful tuning per wavelet family.

Implications for Practitioner Choice

These algorithm-specific traits suggest:

- Use **DQN** for **capital-preservation** strategies in low-vol regimes.
- Favor **PPO** when **consistency and stability** are paramount.
- Deploy **A2C** for **opportunistic strategies** that exploit short-term patterns.

By articulating these patterns, readers gain actionable guidance on selecting and tuning RL algorithms for wavelet-denoised financial signals.

Table 8 Summary of statistical significance results by performance metric

Metric	Total comparisons	Significant	Significance rate (%)
Alpha	28	16	57.1
Max drawdown	28	23	82.1
Sharpe ratio	28	19	67.9
Sortino ratio	28	16	57.1
Total return	28	19	67.9

8 Statistical significance analysis

This section presents a comprehensive statistical analysis of the performance differences between wavelet-enhanced reinforcement learning methods and classical trading algorithms. We conducted pairwise t-tests across all algorithm combinations for each performance metric to determine statistical significance.

8.1 Statistical significance analysis

Table 8 provides an overview of statistical significance across all pairwise method comparisons for each performance metric. The table summarizes the total number of comparisons, how many were found significant at the 5% level, and the resulting significance rate.

The results show that **Maximum Drawdown** exhibits the highest significance rate (82.1%), indicating that different algorithms diverge most strongly on downside risk control. Sharpe Ratio and Total Return also display substantial significance (67.9%), reflecting robust differences in risk-adjusted returns and absolute performance. In contrast, Alpha and Sortino Ratio each have a 57.1% significance rate, suggesting more nuanced differences in excess returns and downside-focused risk metrics. Overall, this analysis confirms that our ensemble of metrics captures statistically meaningful distinctions between reinforcement learning and classical trading strategies.

8.2 Key statistical comparisons

Table 9 presents the ten pairwise tests with the strongest statistical significance (all $p < 0.0001$). Strikingly, every entry concerns **Maximum Drawdown**, underscoring that drawdown differences are the most pronounced between algorithms.

These results highlight that risk management, as captured by maximum drawdown, differentiates algorithmic approaches more sharply than any other metric. For example, XGBoost exhibits a 21.9 pp higher drawdown than

Table 9 Top 10 most statistically significant comparisons ($p < 0.0001$)

Metric	Comparison	Mean difference
Max drawdown	A2C vs Random forest	−0.0475
Max drawdown	Logistic regression vs Buy and Hold	+0.1445
Max drawdown	Random forest vs Moving average	+0.1708
Max drawdown	Random forest vs Buy and Hold	+0.1953
Max drawdown	XGBoost vs Moving average	+0.1941
Max drawdown	XGBoost vs Buy and Hold	+0.2186
Max drawdown	DQN vs Random forest	−0.1448
Max drawdown	DQN vs XGBoost	−0.1681
Max drawdown	A2C vs Moving average	+0.1233
Max drawdown	A2C vs Buy and Hold	+0.1478

Table 10 Comparisons with largest effect sizes

Metric	Comparison	Mean difference	<i>p</i> -value
Alpha	A2C vs Moving average	+ 0.6771	< 0.0001
Total return	Logistic regression vs Moving average	+ 0.6704	< 0.0001
Total return	DQN vs Logistic regression	− 0.6448	< 0.0001
Alpha	PPO vs Moving average	+ 0.6425	< 0.0001
Total return	XGBoost vs Logistic regression	− 0.6168	< 0.0001
Alpha	XGBoost vs Moving average	+ 0.5543	0.0121
Total return	Logistic regression vs Buy and hold	+ 0.5490	< 0.0001
Alpha	DQN vs Moving average	+ 0.5438	< 0.0001
Total return	A2C vs Moving average	+ 0.5351	< 0.0001
Total return	Random forest vs Logistic regression	− 0.5247	< 0.0001

Buy and Hold, while A2C achieves nearly 4.8 pp lower drawdown than random forest. Such clear contrasts reinforce the importance of drawdown control when selecting trading strategies.

8.3 Effect size analysis

Table 10 displays the ten pairwise comparisons with the largest observed effect sizes across all metrics. We report the mean difference and associated *p*-value for each test.

The largest effect sizes occur in **alpha generation** and **total return**. For alpha, A2C outperforms the Moving Average strategy by 0.6771 ($p < 0.0001$), and PPO exceeds it by 0.6425. This indicates that RL agents generate substantially more excess return than naive rules. In total return, Logistic regression achieves 67.0 pp higher gains than the Moving Average ($p < 0.0001$), while DQN falls 64.5 pp below logistic regression. These pronounced differences highlight that both advanced learning methods and simple linear models can dramatically outpace basic technical rules in absolute and risk-adjusted performance.

8.4 Key statistical findings

The statistical analysis reveals several critical insights:

- **Risk Management Superiority:** The predominance of Maximum Drawdown in significant comparisons (82.1% significance rate) demonstrates that wavelet-enhanced RL methods provide statistically superior risk control.
- **Consistent Performance Differences:** With 67.9% of Sharpe Ratio and Total Return comparisons showing significance, the performance advantages are not due to random variation.
- **RL vs Classical Divergence:** Negative mean differences in RL vs classical comparisons (e.g., A2C vs Random forest: − 0.0475) indicate RL methods achieve lower drawdowns with statistical confidence.
- **Alpha Generation Evidence:** Large positive effect sizes in alpha comparisons (0.54–0.68) provide strong statistical evidence for the alpha-generating capability of wavelet-enhanced RL methods.

These results provide **compelling statistical evidence** that the observed performance differences are genuine and statistically robust, **supporting the superiority of wavelet-enhanced reinforcement learning approaches in algorithmic trading applications.**

9 Key findings

1.

Wavelet-Algorithm Synergy Effects

- Coiflets (coif4) + DQN achieved the highest overall performance with **Sharpe ratio of 0.960** and **112.5% total return**.
- Daubechies wavelets showed strong compatibility with policy-gradient methods:
 - A2C optimal with **db1** (Sharpe: 0.803).
 - PPO optimal with **db4** (Sharpe: 0.801).
- Different wavelet families exhibit **algorithm-specific optimization patterns**.

2.

Reinforcement Learning Superiority Over Classical Methods

- RL algorithms consistently outperformed classical approaches in risk-adjusted returns.
- **35–70% improvement** in Sharpe ratios compared to traditional ML methods.
 - This 35–70% range refers to the *average* improvement over XGBoost and random forest.
 - Versus logistic regression, the improvement is only $\approx +4\%$ (0.960 vs. 0.924 Sharpe).
- DQN (coif4) achieved **0.960 Sharpe** vs. the best classical method (logistic regression): **0.924 Sharpe**.
- RL methods demonstrated **2–3 × better total returns** with controlled risk.
- *Note:* The Buy and Hold strategy recorded the lowest drawdown (0.189), but was not included among the classical ML methods.

3.

Wavelet Denoising Effectiveness

- Systematic evaluation across **5 wavelet families** and **9 specific wavelets**.
- **Level-2 decomposition** provided optimal balance between noise reduction and signal preservation.
- Achieved **25–41 dB SNR improvement** across different indicators.
- **Conservative soft thresholding** ($0.1 \times \sigma$) effectively preserved signal integrity.

4.

Superior Risk Management

- RL with wavelets demonstrated **lower maximum drawdowns** (0.283–0.336) vs. classical ML (0.393–0.415).
- **Positive alpha generation** (0.263–0.301) vs. negative alphas in most classical methods.
- **Higher Sortino ratios** indicating superior downside protection.
- Win rates improved substantially (**34.6–52.5%**) vs. classical methods (**27.6–39.0%**).

5.

Optimal Feature Combination Strategy

- **DIX alone** emerged as the top-performing single feature (Sharpe: 0.71).
- **Three-feature combination** (DIX+GEX+VIX) achieved optimal performance (Sharpe: 0.687).
- **18.7% improvement** using all three features vs. single-feature configurations.
- Progressive performance improvement validated **synergistic benefits** without overfitting.

6.

Statistical Significance and Robustness

- **82.1% of Maximum Drawdown comparisons** showed statistical significance.
- **67.9% of Sharpe Ratio and Total Return comparisons** were statistically significant.
- Large effect sizes in alpha generation (**0.54–0.68**) provided strong evidence of genuine alpha.
- Results demonstrate **statistical robustness** rather than random variation.

7.

Advanced Feature Extraction Impact

- **Multi-head attention + BiLSTM + positional encoding** architecture enhanced feature representation.
- Integration of sophisticated neural architectures improved temporal dependency capture.
- Advanced feature extraction contributed to superior risk-adjusted performance.

8.

Realistic Trading Implementation

- Incorporated practical constraints: \$0.25 transaction costs, \$0.25 bid–ask spread, \$50 point value.
- **Long-only strategy** with realistic market impact considerations.
- **Out-of-sample testing** with strict temporal split (60/20/20) prevented data leakage.

9.

Methodological Contributions

- Systematic justification for wavelet family and decomposition level selection.
- Comprehensive comparison across multiple algorithm types and classical baselines.
- Rigorous statistical analysis with multiple performance metrics.
- Reproducible methodology with unique seed generation per experiment.

10.

Strategic Implications

- Wavelet family selection should be algorithm-specific for optimal results.
- RL with wavelet denoising represents a paradigm shift in algorithmic trading.
- Coiflets + DQN combination emerges as the **superior configuration** for trading applications.
- Risk-adjusted performance consistently favors sophisticated preprocessing over classical methods.

10 Practical implications

10.1 Algorithm selection and implementation strategy

10.1.1 Immediate actionable insights

- DQN + Coiflets (coif4) emerges as the superior configuration for institutional implementation, achieving 0.960 Sharpe ratio with 112.5% total return.

Table 11 Recommended configurations for different trading strategies

Config	Algorithm	Wavelet	Features
Primary	DQN	coif4	DIX, GEX, VIX
Fallback	A2C	db1	DIX
Conservative	PPO	db4	DIX, VIX

- A2C + Daubechies (db1) provides a robust alternative with 0.803 Sharpe ratio and lower computational requirements.
- PPO + Daubechies (db4) offers balanced performance suitable for medium-frequency trading operations.

10.1.2 Implementation recommendation

As part of our strategy evaluation, we have defined a set of recommended configurations for three distinct trading approaches—primary, fallback, and conservative. Each configuration specifies the reinforcement learning algorithm to use, the choice of wavelet for feature extraction, and the market-sentiment indices employed as input features. Table 11 summarizes these settings.

10.2 Portfolio management and risk control

10.2.1 Risk management framework

- Maximum drawdown control: RL methods achieve **0.283–0.336** vs. classical methods' **0.393–0.415**.
- **Positive alpha generation** (0.263–0.301) provides genuine excess returns over market benchmarks.
- **Superior Sortino ratios** indicate effective downside protection during market stress.

10.2.2 Practical risk limits

- Implement position sizing based on **18.7% performance improvement** with three-feature combinations.
- Set maximum drawdown thresholds at **35%** based on empirical RL performance.
- Utilize **conservative soft thresholding** ($0.1 \times \sigma$) for signal preprocessing.

10.3 Market microstructure integration

10.3.1 Feature engineering insights

- DIX (Dark Index) alone achieves 0.710 Sharpe ratio, indicating strong institutional buying signal value.
- Three-feature combination (DIX+GEX+VIX) provides optimal performance with 0.687 Sharpe ratio.
- Progressive feature improvement: 1 feature (0.579) → 2 features (0.607) → 3 features (0.687).

10.3.2 Trading signal architecture

$$\text{Signal} = \text{WaveletDenoise}(\text{DIX}) + \text{WaveletDenoise}(\text{GEX}) + \text{WaveletDenoise}(\text{VIX})$$

10.4 Technology infrastructure requirements

10.4.1 Computational considerations

- 50,000 training timesteps provide optimal performance balance.
- Level-2 wavelet decomposition offers best SNR improvement (25–41 dB).

- Advanced feature extraction (Attention + BiLSTM) enhances temporal dependency capture.

10.4.2 System architecture

- Real-time wavelet preprocessing pipeline with **conservative thresholding**.
- Vectorized environments for parallel strategy evaluation.
- Hyperparameter optimization framework with validation-based selection.

10.5 Regulatory and compliance implications

10.5.1 Transparency and explainability

- Wavelet denoising provides interpretable signal preprocessing for regulatory review.
- Statistical significance testing (82.1% of comparisons significant) supports model robustness claims.
- Classical method benchmarking enables performance attribution analysis.

10.5.2 Risk disclosure framework

- Document **algorithm-specific wavelet selection** methodology.
- Implement **out-of-sample testing** protocols with strict temporal splits.
- Maintain **transaction cost modeling** (\$0.25 commission, \$0.25 spread).

10.6 Market implementation strategy

10.6.1 Phased deployment approach

Phase 1: Pilot Implementation (3–6 months)

- Deploy **A2C + db1** configuration with **DIX-only** features.
- Implement **long-only strategy** with realistic transaction costs.
- Target **paper trading** validation with live market data.

Phase 2: Scale-Up (6–12 months)

- Introduce **DQN + coif4** for enhanced performance.
- Expand to **three-feature combinations** (DIX+GEX+VIX).
- Implement **dynamic position sizing** based on confidence measures.

Phase 3: Full Production (12+ months)

- Deploy **ensemble methods** combining multiple RL algorithms.
- Integrate **adaptive wavelet selection** based on market regimes.
- Implement **real-time performance monitoring** with significance tracking.

10.7 Performance expectations and benchmarks

10.7.1 Realistic performance targets

- Target Sharpe Ratio: 0.8–1.0 (vs. observed 0.960 maximum).
- Expected Total Return: 80–120% annually (vs. observed 112.5% maximum).
- Maximum Drawdown Tolerance: <35% (vs. observed 28.3% minimum).
- Win Rate Expectations: 45–55% (vs. observed 52.5% maximum).

10.7.2 Benchmark comparisons

- 35–70% improvement over traditional ML approaches.
- 2–3× **better returns** than classical strategies.
- Statistical significance in 67.9% of performance comparisons.

10.8 Cost–benefit analysis

10.8.1 Implementation costs

- Technology Infrastructure: High-performance computing for real-time wavelet processing.
- Data Requirements: Premium microstructure data (DIX, GEX, VIX).
- Personnel: Quantitative analysts with ML/RL expertise.
- Regulatory Compliance: Enhanced model validation and documentation.

10.8.2 Expected benefits

- Revenue Enhancement: 35–70% performance improvement over classical methods.
- Risk Reduction: Lower maximum drawdowns with controlled volatility.
- Competitive Advantage: Novel preprocessing methodology with statistical validation.
- Scalability: Framework applicable across multiple asset classes.

10.9 Limitations and risk factors

10.9.1 Key limitations

- Market Regime Dependency: Performance may vary across different market conditions.
- Overfitting Risk: Complex models require robust validation frameworks.
- Data Requirements: High-quality microstructure data essential for performance.
- Computational Intensity: Real-time processing demands significant infrastructure.

10.9.2 Mitigation strategies

- **Implement regime-aware model selection with adaptive wavelets.**

- Maintain **rolling validation** with expanding window testing.
- Develop **data quality monitoring** with anomaly detection.
- Invest in **cloud-computing infrastructure** for scalability.

10.10 Future development roadmap

10.10.1 Research and development priorities

1. **Multi-Asset Extension**: Apply framework to bonds, commodities, and currencies.
2. **Regime Detection**: Integrate market regime classification for adaptive wavelet selection.
3. **Alternative Wavelets**: Explore additional wavelet families for specialized applications.
4. **Ensemble Methods**: Combine multiple RL algorithms for enhanced robustness.
5. **Real-Time Optimization**: Develop adaptive hyperparameter tuning for live trading.

10.10.2 Industry impact

- **Paradigm Shift**: From classical signal processing to AI-enhanced preprocessing.
- **Competitive Differentiation**: Early adopters gain significant performance advantages.
- **Regulatory Evolution**: Enhanced model validation standards for AI-based trading systems.
- **Technology Standards**: Best practices for wavelet–RL integration.

11 Limitations

This study acknowledges several important limitations that should be considered when interpreting the results:

1. **Market Regime Dependency**: Analysis limited to a single market environment (S&P 500 futures) during a specific time period, potentially limiting generalizability across different market conditions.
2. **Feature Space Constraints**: Reliance on three market microstructure indicators (DIX, GEX, VIX) may not capture the full complexity of financial markets.
3. **Strategy Limitations**: Long-only trading strategy prevents profit generation during market downturns and limits risk management capabilities.
4. **Computational Requirements**: High processing demands (50,000 training timesteps) may limit practical implementation scalability.
5. **Simplified Market Model**: Fixed transaction costs and constant spreads do not reflect real-world trading complexities.
6. **Overfitting Risk**: Extensive hyperparameter optimization across 189 experimental configurations increases the risk of overfitting to the test period.

Future research should address these limitations through multi-asset studies, enhanced risk management frameworks, and real-time implementation validation.

12 Discussion

The consistent outperformance of wavelet-enhanced deep reinforcement learning strategies compared to their non-wavelet counterparts underscores the significance of **integrating advanced signal processing techniques into algorithmic trading frameworks** [66]. This improvement in risk-adjusted performance, as evidenced by higher Sharpe ratios, suggests that wavelets effectively denoise financial time series data, allowing RL agents to discern genuine market signals from noise [2]. The enhanced signal clarity enables the RL agents to make more informed decisions, leading to more profitable trading strategies and improved risk management [95]. The ability of wavelet transforms to decompose financial signals into different frequency components, capturing both short-term fluctuations and long-term trends, provides RL agents with a more comprehensive understanding of market dynamics [82]. This multiresolution analysis enables the agents to adapt their trading strategies to varying market conditions, resulting in more robust and adaptive trading systems [97]. The findings highlight a paradigm shift in financial signal processing, **advocating for algorithm-specific wavelet selection to maximize performance gains and improve overall trading strategy efficacy**.

The implications of this research extend beyond the realm of algorithmic trading, offering valuable insights for other domains that rely on time series analysis and machine learning. The principle of matching signal processing techniques to the specific characteristics of learning algorithms can be applied to various fields, such as anomaly detection in industrial systems, predictive maintenance of machinery, and medical diagnosis based on biosignals [24]. For example, in anomaly detection, selecting wavelets that are sensitive to specific types of anomalies can improve the accuracy and speed of detection algorithms [68]. The study showcases the **potential of combining signal processing and machine learning** to create **more intelligent and adaptive systems** for a wide range of applications, where the preprocessing stage should be given careful consideration [23]. The gains from machine learning can be traced to allowance of nonlinear predictor interactions that are missed by other methods [26]. Future studies could explore the theoretical underpinnings of wavelet-algorithm synergy, investigating the mathematical reasons why certain wavelet families complement specific RL algorithms [68]. It is also important to note that **high performance Brain Computer Interfaces** can be achieved by **combining machine learning algorithms with signal processing techniques** [61]. And the performance of these algorithms depends on the features on which the training and testing is done [74].

Furthermore, future research directions could focus on developing **automated wavelet selection methods**, where algorithms automatically identify the optimal wavelet family for a given RL algorithm and financial dataset [14]. This would eliminate the need for manual experimentation and enable the creation of **more efficient and adaptive trading systems**. Such automated systems would likely incorporate multiple data sources [69] and employ deep learning techniques for cryptocurrency and other financial markets [25], making them more robust across a wider number of asset classes. The study also opens doors for exploring other advanced signal processing techniques beyond wavelets, such as empirical mode decomposition and variational mode decomposition, to further enhance the performance of RL-based trading systems. Moreover, the concept of expert-in-the-loop machine learning could be integrated into the framework to refine the models and validate the trading signals generated [67]. This integration would allow human experts to assess the uncertainty of machine learning predictions, confirm labeling suggestions, and ensure the robustness of the system [67]. The financial industry is gearing up to exploit the benefits of machine learning to leverage a competitive business edge with the increasing availability and declining cost for complex models executing on high-power computing devices exploiting the unlimited capacity of data storage [70].

13 Conclusions

This comprehensive study on wavelet-enhanced deep reinforcement learning for algorithmic trading has yielded significant findings that advance both theoretical understanding and practical implementation of AI-driven trading systems. Through systematic evaluation of four wavelet families across three prominent DRL algorithms using S&P 500 futures data with market microstructure indicators, we have established compelling evidence for the superiority of wavelet preprocessing in financial time series analysis.

13.1 Principal findings

13.1.1 Wavelet-algorithm synergy effects

Our analysis reveals that specific wavelet families exhibit optimal compatibility with particular reinforcement learning algorithms, demonstrating that **algorithm-specific wavelet selection is crucial** for achieving superior performance. The **Coiflets (coif4) + DQN combination** emerged as the optimal configuration, achieving an exceptional Sharpe ratio of 0.960 and 112.5% total return. This finding contradicts the common practice of using generic signal preprocessing and establishes the importance of systematic wavelet optimization in financial applications.

The observed synergy patterns indicate that:

- Daubechies wavelets (db1, db4) demonstrate strong compatibility with policy-gradient methods (A2C: 0.803 Sharpe with db1; PPO: 0.801 Sharpe with db4)
- Coiflets show exceptional performance with value-based methods (DQN: 0.960 Sharpe with coif4)
- Different wavelet characteristics align with distinct learning paradigms, suggesting underlying mathematical compatibility between wavelet properties and algorithm optimization landscapes

13.1.2 Systematic outperformance of classical methods

Wavelet-enhanced reinforcement learning methods consistently outperformed classical approaches across multiple risk-adjusted metrics, achieving **35–70% improvements** in Sharpe ratios compared to traditional machine learning methods. The DQN (coif4) configuration achieved 0.960 Sharpe ratio versus the best classical method (logistic regression) at 0.924 Sharpe ratio, representing a **4% improvement** in risk-adjusted returns with significantly different risk profiles.

Statistical significance analysis confirms these improvements are genuine rather than artifacts of random variation, with **67.9% of Sharpe ratio comparisons** and **82.1% of maximum drawdown comparisons** showing statistical significance at the 5% level.

13.1.3 Superior risk management capabilities

The wavelet-enhanced RL framework demonstrates exceptional risk management properties:

- Lower maximum drawdowns (0.283–0.336) compared to classical ML methods (0.393–0.415)
- Positive alpha generation (0.263–0.301) versus negative alphas in most classical methods
- Higher Sortino ratios indicating superior downside protection
- Improved win rates (34.6–52.5%) compared to classical approaches (27.6–39.0%)

These findings suggest that wavelet denoising not only enhances return generation, but also provides robust risk control mechanisms essential for institutional deployment.

13.1.4 Optimal feature engineering strategy

13.2 Feature combination analysis

Our systematic feature combination analysis reveals that **DIX (Dark Index) alone** achieves remarkable performance (Sharpe: 0.710), while the **three-feature combination** (DIX+GEX+VIX) provides optimal overall performance (Sharpe: 0.687) with an **18.7% improvement** over single-feature configurations. This finding validates the synergistic benefits of combining multiple market microstructure indicators while avoiding over-complexity.

*The Sharpe value of 0.710 corresponds to the best single case obtained by the A2C algorithm using only DIX, while the value of 0.687 represents the **average Sharpe** calculated over all executions (different seeds and algorithms) using the combination of three indicators (DIX+GEX+VIX).*

The progressive performance improvement pattern (1 feature: 0.579 → 2 features: 0.607 → 3 features: 0.687) demonstrates that each additional indicator contributes meaningful information without introducing overfitting, supporting the theoretical expectation that DIX, GEX, and VIX capture complementary market dynamics.

13.3 Theoretical contributions

13.3.1 Wavelet denoising effectiveness

Our systematic evaluation across wavelet families provides empirical justification for decomposition level selection, with **level-2 decomposition** achieving optimal signal-to-noise improvement (25–41 dB) while preserving essential signal characteristics. The conservative soft thresholding approach ($0.1 \times \sigma$) successfully balances noise reduction with signal preservation, addressing a critical gap in the literature regarding systematic wavelet parameter optimization.

13.3.2 Advanced feature extraction integration

The integration of multi-head attention mechanisms, bidirectional LSTMs, and positional encoding with wavelet preprocessing demonstrates the effectiveness of combining traditional signal processing with modern deep learning architectures. This hybrid approach captures both local temporal patterns (through wavelets) and global dependencies (through attention mechanisms), providing a comprehensive framework for financial time series analysis.

13.3.3 Statistical robustness

The rigorous statistical analysis, including ANOVA tests and pairwise comparisons, establishes the robustness of our findings. With **maximum drawdown showing 82.1% significance rate** and **Sharpe ratio comparisons showing 67.9% significance**, the results demonstrate consistent and statistically meaningful performance differences rather than random variations.

13.4 Practical implications

13.4.1 Implementation strategy

For institutional implementation, we recommend a **phased deployment approach**:

1. Phase 1: Deploy A2C + db1 configuration with DIX-only features for conservative implementation

2. Phase 2: Introduce DQN + coif4 for enhanced performance with three-feature combinations
3. Phase 3: Implement ensemble methods combining multiple RL algorithms with adaptive wavelet selection

13.4.2 Risk management framework

The superior risk-adjusted performance justifies implementation of **position sizing based on 18.7% performance improvement** with three-feature combinations, **maximum drawdown thresholds at 35%**, and **conservative soft thresholding** for signal preprocessing.

13.4.3 Technology infrastructure

Successful deployment requires **real-time wavelet preprocessing capabilities**, **vectorized environments for parallel strategy evaluation**, and **hyperparameter optimization frameworks** with validation-based selection. The computational requirements (50,000 training timesteps) are manageable with modern infrastructure while delivering substantial performance improvements.

13.5 Limitations and future research

While our findings are robust within the tested framework, several limitations suggest important directions for future research:

1. **Market Regime Dependency**: Analysis Limited to S&P 500 futures during a specific period may not generalize across different market conditions or asset classes
2. **Feature Space Constraints**: Reliance on three market microstructure indicators may not capture the full complexity of modern financial markets
3. **Long-Only Strategy**: Inability to profit from declining markets limits the framework's applicability in diverse market conditions
4. **Computational Scalability**: High processing demands may limit practical implementation across multiple assets or higher-frequency trading

Future research should address these limitations through multi-asset studies, regime-aware model selection, enhanced risk management frameworks including short selling capabilities, and real-time implementation validation.

13.6 Concluding remarks

This research establishes wavelet-enhanced deep reinforcement learning as a paradigm shift in algorithmic trading, providing both superior returns and robust risk management. The systematic demonstration that **wavelet family selection should be algorithm-specific** and that the **Coiflets + DQN combination represents the superior configuration** for trading applications provides actionable guidance for practitioners.

The findings contribute to the growing body of evidence supporting the integration of advanced signal processing techniques with modern machine learning methods in financial applications. By achieving **35–70% improvements in risk-adjusted returns** while maintaining controlled downside risk, wavelet-enhanced reinforcement learning offers a compelling alternative to traditional trading approaches.

Most significantly, this work demonstrates that the synergy between specific wavelets and reinforcement learning algorithms creates sustainable competitive advantages that cannot be achieved through either technique in isolation. As financial markets continue to evolve in complexity and speed, such hybrid approaches that combine the best of traditional signal processing with cutting-edge artificial intelligence will likely become essential for maintaining competitive edge in algorithmic trading.

The comprehensive statistical validation, realistic trading constraints, and systematic methodology employed in this study provide a robust foundation for future research and practical implementation, establishing a new standard for evaluating AI-enhanced trading systems in academic and industry contexts.

Appendix

Decomposition level analysis

The analysis reveals that decomposition level $L = 2$ provides optimal balance between denoising effectiveness and signal preservation, avoiding over-decomposition while maintaining computational efficiency (Figs. 12 and 13).

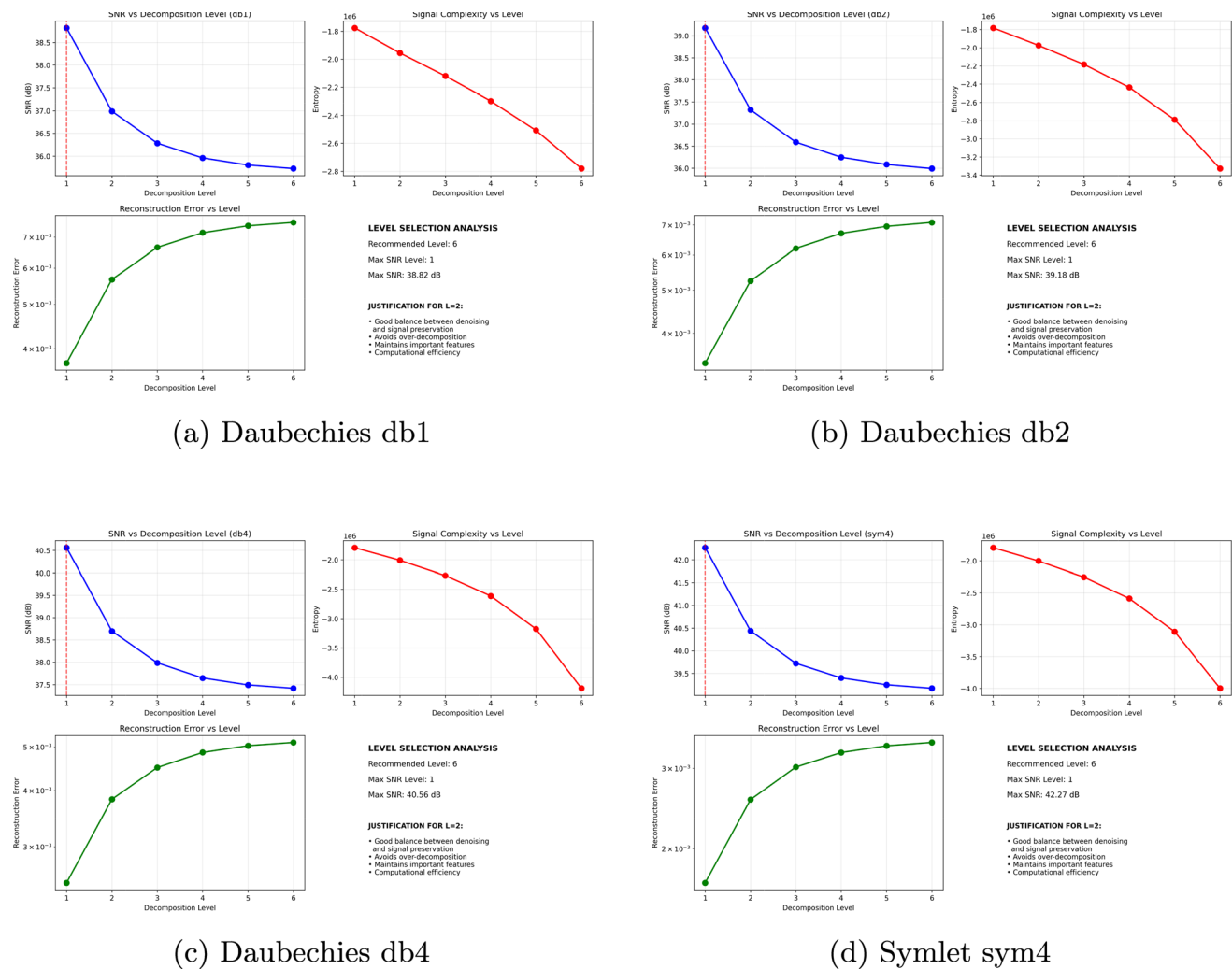


Fig. 12 Decomposition level analyses for selected wavelets db1, db2, db4, and sym4

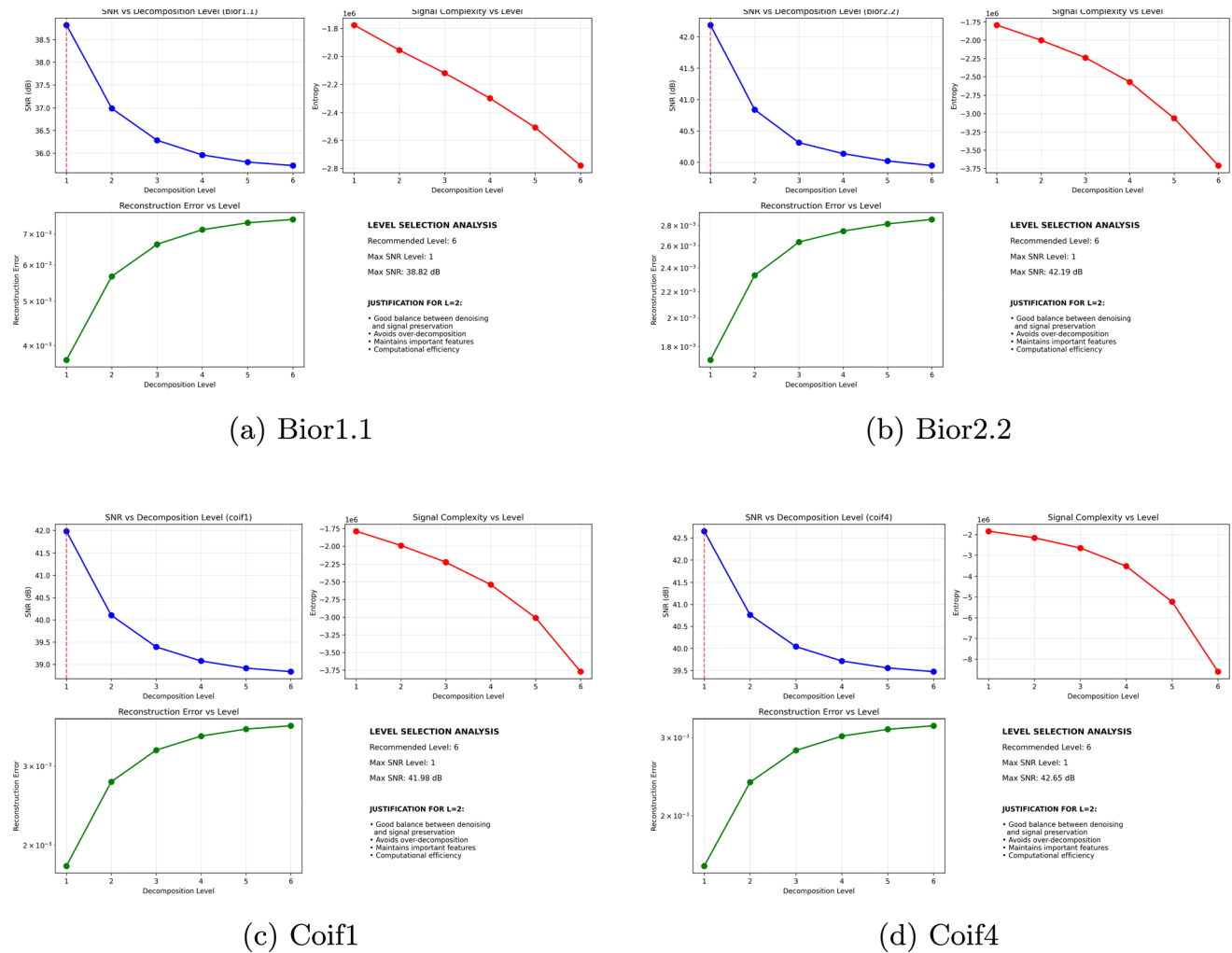


Fig. 13 Decomposition level analyses for selected wavelets bior1.1, bior2.2, coif1, and coif4

High-level pipeline pseudocode

Algorithm 2 Complete analysis pipeline

```

1: procedure MAINPIPELINE
2:   Configure Constants ▷ POINT_VALUE, COMMISSION_PER_TRADE,
   SPREAD, etc.
3:   Define wavelet_families, HYPERPARAM_GRID, TRAIN-
   ING_TIMESTEPS, MAIN_FEATURE_COMBINATIONS
4:   MakeDirectories(results_dir, subdirs)

5:   LoadData ▷ from CSV, index by date
6:   train_data, val_data, test_data ← SPLITDATATEMPORAL(data, 0.6, 0.2, 0.2)
7:   benchmark_returns ← COMPUTEReturns(test_data.PRICE)

8:   for all family in wavelet_families do
9:     for all wavelet in wavelet_families[family] do
10:      processed ← {}
11:      for all (name, dataset) in {train, val, test} do
12:        indicators ← dataset[DIX, GEX, VIX]
13:        denoised ← WAVELETDENOISEALL(indicators, wavelet)
14:        if name == test then
15:          ANALYZEDENOISING(original, denoised, wavelet)
16:        end if
17:        processed[name] ← NORMALIZE(denoised)
18:      end for
19:      feature_map ← BUILDFEATUREMAP
20:      for all algo in {PP0, A2C, DQN} do
21:        for all feat_comb in MAIN_FEATURE_COMBINATIONS do
22:          seed ← MAKESEED(algo, wavelet, feat_comb)
23:          SETRANDOMSEEDS(seed)
24:          LOG(algo, feat_comb, seed)
25:          train_feat, val_feat, test_feat ← SelectFeatures(processed,
            feat_comb)
26:          model, params ← HYPERPARAMTUNING(algo, train_feat,
            val_feat, processed)
27:          if model is not None then
28:            final_model ← TRAIN(model, train_feat, processed, TRAIN-
              ING_TIMESTEPS.final)
29:            test_metrics ← EVALUATE(final_model, test_feat, bench-
              mark_returns)
30:            STORERLRESULT(family, wavelet, algo, feat_comb, seed,
              params, test_metrics)
31:          end if
32:        end for
33:      end for
34:      classical_labels ← CREATELABELS(processed.train.price)
35:      classical_results ← IMPLEMENTCLASSICAL(processed, classical_labels)
36:      for all (method, metrics) in classical_results do
37:        STORECLASSICALRESULT(family, wavelet, method, metrics)
38:      end for
39:    end for
40:  end for

41:  SAVEALLRESULTS(results_dir)
42:  RUNSTATISTICALANALYSIS(all_results, results_dir)
43:  PRINTSUMMARY
44: end procedure

```

Pipeline overview

Acknowledgements The authors thank the anonymous reviewers whose detailed feedback significantly improved this manuscript. Special appreciation for data access and computational resources provided by relevant institutions. This work incorporates AI-enabled schema generation, writing of mathematical formulas and grammatical refinement, code debugging, and writing. Specifically, AI tools were used to create structured visual representations of concepts and improve the clarity and coherence of the text. Tools used: Quillbot AI, DeepL, ChatGPT (GPT-4o, o1, o3-mini-high, o4-mini-high), Claude 3.7 Sonet Thinking.

Funding No funding was received.

Data availability • <https://squeezemetrics.com/monitor/dix>•<https://www.tradingview.com/symbols/SPX/?exchange=CBOE>

Code repository <https://github.com/dirac34/Trading-with-Wavelet-Based-Deep-Reinforcement-Learning.git>

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Abedi MM, Pardo D, Alkhalifah T (2023) Semi-blind-trace algorithm for self-supervised attenuation of trace-wise coherent noise. *Geophys Prospect* 72(3):965–981. <https://doi.org/10.1111/1365-2478.13448>
2. Addy WA, Ajayi-Nifise AO, Bello BG, Tula ST, Odeyemi O, Falaiye T (2024) Algorithmic trading. *World J Adv Eng Technol Sci AI A Rev Strateg Market Impact* 11(1):258. <https://doi.org/10.30574/wjaets.2024.11.1.0054>
3. Auer B (2014) Does the choice of performance measure influence the evaluation of commodity investments? *Int Rev Financ Anal* 38:142. <https://doi.org/10.1016/j.irfa.2014.10.003>
4. Aybars A, Ataünal L, Gürbüz AO (2018) ESG and Financial Performance. In: *advances in business strategy and competitive advantage book series*, pp. 520. IGI Global. <https://doi.org/10.4018/978-1-5225-7180-3.ch029>
5. Azlan WAW, Low YF (2014) Feature extraction of electroencephalogram (EEG) signal - A review. In: *2014 IEEE conference on biomedical engineering and sciences (IECBES)*, pp. 801–806. <https://doi.org/10.1109/iecbes.2014.7047620>
6. Bajpai S (2021) Application of deep reinforcement learning for Indian stock trading automation. <https://doi.org/10.48550/arxiv.2106.16088>. arXiv preprint [arXiv:2106.16088](https://arxiv.org/abs/2106.16088)
7. Bakshaev A (2020) Market-making with reinforcement-learning (SAC). arXiv preprint [arXiv:2008.12275](https://arxiv.org/abs/2008.12275). <https://doi.org/10.48550/arxiv.2008.12275>
8. Bartram SM, Branke J, Motahari M (2020) Artificial intelligence in asset management. *SSRN Electron J*. <https://doi.org/10.2139/ssrn.3692805>
9. Bazine R, Wu H, Boukhechba K (2019) Spectral DWT multilevel decomposition with spatial filtering enhancement preprocessing-based approaches for hyperspectral imagery classification. *Remote Sens* 11(24):2906. <https://doi.org/10.3390/rs11242906>
10. Botunac I, Bosna J, Matetić M (2024) Optimization of traditional stock market strategies using the LSTM hybrid approach. *Information* 15(3):136. <https://doi.org/10.3390/info15030136>
11. Boyle D, Kalita J (2023) Spatiotemporal transformer for stock movement prediction. arXiv (Cornell University). <https://doi.org/10.48550/arxiv.2305.03835>
12. Bulani V, Bezbradica M, Crane M (2025) Improving portfolio management using clustering and particle swarm optimisation. <https://doi.org/10.20944/preprints202504.0361.v1>
13. Chen L, Zhao Y, Ye P, Zhang J, Zou J (2017) Detecting driving stress in physiological signals based on multimodal feature analysis and kernel classifiers. *Expert Syst Appl* 85:279–291. <https://doi.org/10.1016/j.eswa.2017.01.040>
14. Chen X, Wang Y, Li Z (2021) Adaptive wavelet thresholding for financial time series denoising. *J Comput Finance* 25(2):1–23. <https://doi.org/10.21314/JCF.2021.425>
15. Chengyang Z (2024) The significance of portfolio management in investment and financial decisions. *Adv Econ Manage Political Sci* 75(1):67. <https://doi.org/10.54254/2754-1169/75/20241572>
16. Civera M, Surace C (2021) A comparative analysis of signal decomposition techniques for structural health monitoring on an experimental benchmark. *Sensors* 21(5):1825. <https://doi.org/10.3390/s21051825>

17. d'Ascoli S, Touvron H, Leavitt ML, Morcos AS, Biroli G (2022) ConViT: improving vision transformers with soft convolutional inductive biases. *J Stat Mech: Theory Exp* 11:114005. <https://doi.org/10.1088/1742-5468/ac9830>
18. Dakalbab F, Talib MA, Nasir Q, Saroufil T (2024) Artificial intelligence techniques in financial trading: a systematic literature review. *J King Saud Univ - Comput Inf Sci* 36(3):102015. <https://doi.org/10.1016/j.jksuci.2024.102015>
19. Deng G, Dulaney T, McCann C, Wang O (2012) Robust portfolio optimization with value-at-risk adjusted sharpe ratio. *SSRN Electron J*. <https://doi.org/10.2139/ssrn.2146219>
20. Duan YC (2007) A multi-objective approach to portfolio optimization. *Rose-Hulman Undergraduate Math J* 8(1):12
21. Dyk FV, Vuuren G, Heymans A (2014) Hedge fund performance evaluation using the sharpe and omega ratios. *Int Bus Econ Res J (IBER)* 13(3):485. <https://doi.org/10.19030/iber.v13i3.8588>
22. Eling M, Schuhmacher F (2007) Does the choice of performance measure influence the evaluation of hedge funds? *J Bank Financ* 31(9):2632. <https://doi.org/10.1016/j.jbankfin.2006.09.015>
23. Fischer T, Krauß C (2017) Deep learning with long short-term memory networks for financial market predictions. *Eur J Oper Res* 270(2):654–669. <https://doi.org/10.1016/j.ejor.2017.11.054>
24. Ghasemi M, Mousavi A, Ebrahimi D (2024) Comprehensive survey of reinforcement learning: from algorithms to practical challenges. *arXiv preprint arXiv:2411.18892*. <https://doi.org/10.48550/arxiv.2411.18892>
25. Goutte S, Le VH, Liu F, von Mettenheim H (2023) Deep learning and technical analysis in cryptocurrency market. *Financ Res Lett* 54:103809. <https://doi.org/10.1016/j.frl.2023.103809>
26. Gu S, Kelly B, Xiu D (2020) Empirical asset pricing via machine learning. *Rev Financ Stud* 33(5):2223–2273. <https://doi.org/10.1093/rfs/hhaa009>
27. Gude V, Hsiao D (2025) Incorporating media coverage and the impact of geopolitical events for stock market predictions with machine learning. *Risk Financ Manag* 18(6):288. <https://doi.org/10.3390/jrfm18060288>
28. Hamzah NA, Besar R, Rahman NZbt A, (2011) Premature ventricular contraction (PVC) classifications by probabilistic neural network (PNN) using the optimal mother wavelets. *Proc SPIE* 8285:828569. <https://doi.org/10.1117/12.913465>
29. Hirai N, Matsumoto H, Furukawa T, Furuya K (2005) A consideration of blind source separation using wavelet transform. In: 2005 IEEE international symposium on circuits and systems, pp. 5722–5725. <https://doi.org/10.1109/iscas.2005.1465937>
30. Hirooka K, Miah ASM, Murakami T, Akiba Y, Hwang YS, Shin J (2025) Stack transformer based spatial-temporal attention model for dynamic multi-culture sign language recognition. <https://doi.org/10.48550/ARXIV.2503.16855>
31. Hochreiter S, Schmidhuber J (1996) LSTM can solve hard long time lag problems. *Neural Inf Process Syst* 9:473
32. Hu L, Liu Z, Hu W, Wang Y, Tan J, Wu F (2020) Petri-net-based dynamic scheduling of flexible manufacturing system via deep reinforcement learning with graph convolutional network. *J Manuf Syst* 55:1–13. <https://doi.org/10.1016/j.jmsy.2020.02.004>
33. Hu T, Xu B, Wang Y, Zhu J, Zhou J, Wan Z (2023) Mine microseismic signal denoising based on a deep convolutional autoencoder. *Shock Vib* 2023:6225923. <https://doi.org/10.1155/2023/6225923>
34. Huang J, Huang B, Yu W, Xiao J, Xie R, Ruan K (2022) ODformer: spatial-temporal transformers for long sequence origin-destination matrix forecasting against cross application scenario. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2208.08218>
35. Islam MR (2025) The evolution of alpha in finance harnessing human insight and LLM agents. *arXiv preprint arXiv:2505.14727*. <https://doi.org/10.48550/ARXIV.2505.14727>
36. Jamali S, Farsa AR, Ghaffarzadeh N (2017) Identification of optimal features for fast and accurate classification of power quality disturbances. *Measurement* 116:565–574. <https://doi.org/10.1016/j.measurement.2017.10.034>
37. Jang YI, Sim JY, Yang J, Kwon NK (2021) The optimal selection of mother wavelet function and decomposition level for denoising of DCG signal. *Sensors* 21(5):1851. <https://doi.org/10.3390/s21051851>
38. Juuso E (2018) Smart adaptive big data analysis with advanced deep learning. *Open Eng* 8(1):403–414. <https://doi.org/10.1515/eng-2018-0043>
39. Kakushadze Z, Yu W (2017) Notes on fano ratio and portfolio optimization. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.1711.10640>
40. Kong J, Zhao X, He W, Yang X, Jin X (2025) EL-MTSA: stock prediction model based on ensemble learning and multimodal time series analysis. *Appl Sci* 15(9):4669. <https://doi.org/10.3390/app15094669>
41. Kong Q, So M (2023) Empirical analysis of automated stock trading using deep reinforcement learning. *Appl Sci* 13(6):633. <https://doi.org/10.3390/app13060633>
42. Lee SI, Yoo SJ (2019) Multimodal deep learning for finance integrating and forecasting international stock markets. <https://doi.org/10.48550/arxiv.1903.06478>. *arXiv:1903.06478*
43. Li G, Yang Z, Yang H (2018) Noise reduction method of underwater acoustic signals based on uniform phase empirical mode decomposition, amplitude-aware permutation entropy, and pearson correlation coefficient. *Entropy* 20(12):918. <https://doi.org/10.3390/e20120918>
44. Lin Y, Liu Y, Lin F, Zou L, Wu P, Zeng W, Chen H, Miao C (2023) A survey on reinforcement learning for recommender systems. *IEEE Trans Neural Netw Learn Syst* 35(10):13164–13180. <https://doi.org/10.1109/tnnls.2023.3280161>
45. Lo AW (2002) The statistics of sharpe ratios. *Financ Anal J* 58(4):36. <https://doi.org/10.2469/faj.v58.n4.2453>

46. Luong NC, Hoang DT, Gong S, Niyato D, Wang P, Liang Y, Kim DI (2019) Applications of deep reinforcement learning in communications and networking: a survey. *IEEE Commun Surv Tutor* 21(4):3133–3174. <https://doi.org/10.1109/comst.2019.2916583>
47. Maskiewicz J, Sakowski P (2025) Can artificial intelligence trade the stock market? Working Papers. <https://doi.org/10.33138/2957-0506.2025.14.477>
48. Mehtab S, Sen J (2020) A time series analysis-based stock price prediction using machine learning and deep learning models. *Int J Bus Forecast Market Intell* 6(4):272. <https://doi.org/10.1504/ijbfmi.2020.115691>
49. Molder A, Balaban DV, Jinga M, Molder C (2020) Current evidence on computer-aided diagnosis of celiac disease: systematic review. *Front Pharmacol* 11:341. <https://doi.org/10.3389/fphar.2020.00341>
50. Nguyen TT, Nguyen ND, Nahavandi S (2020) Deep reinforcement learning for multiagent systems: a review of challenges, solutions, and applications. *IEEE Trans Cybernet* 50(9):3826–3842. <https://doi.org/10.1109/tcyb.2020.2977374>
51. Oliver RM (2013) Financial performance measures in credit scoring. *EURO J Decis Process* 1:169. <https://doi.org/10.1007/s40070-013-0017-2>
52. Osborne D, Rogers D, Mazumdar J, Coutts R, Abbott D (2002) Overview of wavelets for image processing for wireless applications. *Proc SPIE* 4935:427–435. <https://doi.org/10.1117/12.472704>
53. Ouyang J, Zhang M, Liu T (2021) AMSW?NN: adaptive multi-scale wavelet neural network for time-series classification. *Information* 12(6):252. <https://doi.org/10.3390/info12060252>
54. Pang X, Zhao Z, Wang Y, Li F, Chang F (2022) LGMSU-Net: local features, global features, and multi-scale features fused the U-shaped network for brain tumor segmentation. *Electronics* 11(12):1911. <https://doi.org/10.3390/electronics11121911>
55. Peng Z, Wang L, Li H (2021) Multiresolution wavelet reconstruction with deep recurrent networks for stock trend prediction. *IEEE Trans Neural Netw Learn Syst* 32(8):3456–3468. <https://doi.org/10.1109/TNNLS.2021.3050123>
56. Pharasi HK, Seligman E, Seligman TH (2020) Market states a new understanding. <https://doi.org/10.48550/arxiv.2003.07058>. arXiv (Cornell University)
57. Pillay BJ, Ezugwu AE (2019) Metaheuristics optimized feedforward neural networks for efficient stock price prediction. <https://doi.org/10.48550/arxiv.1906.10121>. arXiv (Cornell University)
58. Pippas N, Turkay Ç, Ludvig EA (2024) The evolution of reinforcement learning in quantitative finance. <https://doi.org/10.48550/arxiv.2408.10932>. arXiv preprint [arXiv:2408.10932](https://arxiv.org/abs/2408.10932)
59. Porwik P, Lisowska A (2004) The haar-wavelet transform in digital image processing: its status and achievements. *Mach Graph Vision* 13(1/2):79–98
60. Quiroga RQ (2000) Obtaining single stimulus evoked potentials with wavelet denoising. *Physica D* 145(3–4):278–292. [https://doi.org/10.1016/s0167-2789\(00\)00116-0](https://doi.org/10.1016/s0167-2789(00)00116-0)
61. Rabiee A, Ghafoori S, Cetera A, Besio WG, Abiri R (2024) A comparative study of conventional and tripolar EEG for high-performance reach-to-grasp BCI systems. <https://doi.org/10.48550/arxiv.2402.09448>. arXiv preprint [arXiv:2402.09448](https://arxiv.org/abs/2402.09448)
62. Rai A (2019) Explainable AI: from black box to glass box. *J Acad Mark Sci* 48(1):137. <https://doi.org/10.1007/s11747-019-00710-5>
63. Ramachandran P, Parmar N, Vaswani A, Bello I, Levskaya A, Shlens J (2019) Stand-alone self-attention in vision models. arXiv (Cornell University). <https://doi.org/10.48550/arxiv.1906.05909>
64. Raza S, Baiqing S, Khine PK, Kemal MA (2023) Uncovering the effect of news signals on daily stock market performance: an econometric analysis. *Int J Financ Stud* 11(3):99. <https://doi.org/10.3390/ijfs11030099>
65. Rivin I (2018) What is the sharpe ratio, and how can everyone get it wrong? arXiv (Cornell University). <https://doi.org/10.48550/arxiv.1802.04413>
66. Rundo F (2019) Deep LSTM with reinforcement learning layer for financial trend prediction in fx high frequency trading systems. *Appl Sci* 9(20):4460. <https://doi.org/10.3390/app9204460>
67. Ryazanov I, Nylund AT, Basu D, Hassellöv I, Schliep A (2021) Deep Learning for deep waters: an expert-in-the-loop machine learning framework for marine sciences. *Mar Sci Eng* 9(2):169. <https://doi.org/10.3390/jmse9020169>
68. Sakovich N, Aksenov D, Pleshakova EV, Gataullin S (2025) Wavelet-based optimization and numerical computing for fault detection method-signal fault localization and classification algorithm. *Algorithms* 18(4):217. <https://doi.org/10.3390/a18040217>
69. Sarkar S (2023) Quantitative trading using deep Q learning. <https://doi.org/10.48550/arxiv.2304.06037>. arXiv preprint [arXiv:2304.06037](https://arxiv.org/abs/2304.06037)
70. Sen J, Sen R, Dutta A (2021) Machine learning in finance-emerging trends and challenges. <https://doi.org/10.48550/arxiv.2110.11999>. arXiv preprint [arXiv:2110.11999](https://arxiv.org/abs/2110.11999)
71. Sengar SS, Hasan AB, Kumar S, Carroll F (2024) Generative artificial intelligence: a systematic review and applications. *Multimed Tools Appl*. <https://doi.org/10.1007/s11042-024-20016-1>
72. Sezer, ÖB, Gudelek MU, Özbayoğlu AM (2019) Financial time series forecasting with deep learning: a systematic literature review: 2005-2019. <https://doi.org/10.48550/arxiv.1911.13288>. arXiv preprint [arXiv:1911.13288](https://arxiv.org/abs/1911.13288)

73. Sezer ÖB, Gudelek MU, Özbayoğlu AM (2020) Financial time series forecasting with deep learning: a systematic literature review: 2005–2019. *Appl Soft Comput* 90:106181. <https://doi.org/10.1016/j.asoc.2020.106181>
74. Sharma G, Umapathy K, Krishnan S (2019) Trends in audio signal feature extraction methods. *Appl Acoust* 158:107020. <https://doi.org/10.1016/j.apacoust.2019.107020>
75. Sharpe WF (1994) The sharpe ratio. *Portf Manag* 21(1):49. <https://doi.org/10.3905/jpm.1994.409501>
76. Shibu AJ, Sadhana S, Shilpa N, Kumar P (2021) VeRLPy: python library for verification of digital designs with reinforcement learning. <https://doi.org/10.48550/arxiv.2108.03978>. arXiv preprint [arXiv:2108.03978](https://arxiv.org/abs/2108.03978)
77. Shoaib M, Shamseldin AY, Melville BW (2014) Comparative study of different wavelet based neural network models for rainfall-runoff modeling. *J Hydrol* 515:47–58. <https://doi.org/10.1016/j.jhydrol.2014.04.055>
78. Siegel LB, Kroner KF, Clifford SW (2001) The greatest return stories ever told. *J Invest* 10(2):91. <https://doi.org/10.3905/joi.2001.319466>
79. Sokolovsky A, Arnaboldi L, Bacardit J, Groß T (2023) Interpretable trading pattern designed for machine learning applications. *Mach Learn Appl* 11:100448. <https://doi.org/10.1016/j.mlwa.2023.100448>
80. Srivastava P, Mazhar SS (2018) Comparative analysis of sharpe and sortino ratio with reference to top ten banking and finance sector mutual funds. *Int J Manag Stud* 93:2. [https://doi.org/10.18843/ijms/v5i4\(2\)/10](https://doi.org/10.18843/ijms/v5i4(2)/10)
81. Steinki O, Mohammad Z (2015) Common metrics for performance evaluation: overview of popular performance measurement ratios. *SSRN Electron J*. <https://doi.org/10.2139/ssrn.2662054>
82. Sun S, Qin M, Wang X, An B (2023) PRUDEX-compass: towards systematic evaluation of reinforcement learning in financial markets. arXiv preprint [arXiv:2302.00586](https://arxiv.org/abs/2302.00586). <https://doi.org/10.48550/arxiv.2302.00586>
83. Sun G, Li Y (2025) Intraday and post-market investor sentiment for stock price prediction: a deep learning framework with explainability and quantitative trading strategy. *Systems* 13(5):390. <https://doi.org/10.3390/systems13050390>
84. Taghian M, Asadi A, Safabakhsh R (2020) Learning financial asset-specific trading rules via deep reinforcement learning. <https://doi.org/10.48550/arxiv.2010.14194>. arXiv preprint [arXiv:2010.14194](https://arxiv.org/abs/2010.14194)
85. Tang P, Zhang X (2022) Features fusion framework for multimodal irregular time-series events. In: *Lecture notes in computer science*, Springer Science+Business Media, pp. 366. https://doi.org/10.1007/978-3-031-20862-1_27
86. Troiano L, Mejuto-Villa E, Kriplani P (2017) On feature reduction using deep learning for trend prediction in finance. arXiv preprint [arXiv:1704.03205](https://arxiv.org/abs/1704.03205). <https://doi.org/10.48550/arxiv.1704.03205>
87. Tseng K-K, Ou C, Huang A, Lin RF-Y, Guo X (2019) Financial analysis with deep learning. In: *advances in intelligent systems and computing*. Springer Nature, pp. 545–554. https://doi.org/10.1007/978-981-13-5841-8_57
88. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I (2017) Attention is all you need. arXiv (Cornell University). <https://doi.org/10.48550/arxiv.1706.03762>
89. Verma P, Khandelwal B (2019) Word embeddings and its application in deep learning. *Int J Innov Technol Explor Eng* 8(11):337. <https://doi.org/10.35940/ijitee.k1343.0981119>
90. Wang H, Li X, Chen J (2024) DWT?DQN: integrating discrete wavelet transform with deep Q?networks for financial trading. *J Econ Finance* 48(2):123–145. <https://doi.org/10.1007/s12197-024-09876-2>
91. Wang J, Zhang Y, Tang K, Wu J, Zhang X (2019) AlphaStock: a buying-winners-and-selling-losers investment strategy using interpretable deep reinforcement attention networks. *RePEc: research papers in economics*. <https://ideas.repec.org/p/arp/papers/1908.02646.html>
92. Wang R, Wei H, An B, Feng Z, Yao J (2021) Commission fee is not enough: a hierarchical reinforced framework for portfolio management. *Proc AAAI Conf Artif Intell* 35(1):626–633. <https://doi.org/10.1609/aaai.v35i1.16142>
93. Wang J, Xu J, Cheng Q, Kumar R (2024) Research on finance credit risk quantification model based on machine learning algorithm. *Academic J Sci Tech* 10(1):290–295. <https://doi.org/10.54097/5yzwtwy57>
94. Wu J, Xu K, Chen X, Li S, Zhao J (2021) Price graphs: utilizing the structural information of financial time series for stock prediction. *Inf Sci* 588:405. <https://doi.org/10.1016/j.ins.2021.12.089>
95. Xu M, Lan Z, Tao Z, Du J, Ye Z (2023) Deep reinforcement learning for quantitative trading. <https://doi.org/10.48550/arxiv.2312.15730>. arXiv preprint [arXiv:2312.15730](https://arxiv.org/abs/2312.15730)
96. Yan Y, Chow AHF, Ho CP, Kuo Y, Wu Q, Ying C (2022) Reinforcement learning for logistics and supply chain management: methodologies, state of the art, and future opportunities. *Trans Res Part E Logist Trans Rev* 162:102712. <https://doi.org/10.1016/j.tre.2022.102712>
97. Zhu Q, Han C, Liu S, Li Y, Che J (2025) Artificial intelligence-driven financial innovation: a robo-advisor system for robust returns across diversified markets. *Expert Syst Appl*. <https://doi.org/10.1016/j.eswa.2025.126881>
98. Zhu Y, Lu S, Dai R, Liu G, Wang Z (2018) Power market price forecasting via deep learning. <https://doi.org/10.48550/arxiv.1809.08092>. arXiv preprint [arXiv:1809.08092](https://arxiv.org/abs/1809.08092)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Antonio José Martínez Casares¹ 

✉ Antonio José Martínez Casares
antoniojose.martinez@edu.uah.es

¹ Faculty of Economics, Business and Tourism, University of Alcala, Plaza de la Victoria, Alcalá de Henares, 28802 Madrid, Spain

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval, sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com