DOI: 10.46299/ISG.2025.MONO.TECH.2.3.3

3.3 Intelligent decision-making in energy finance: the role of reinforcement learning

The integration of energy markets with financial engineering has given rise to the specialized field of energy finance, which addresses valuation, risk management, and investment decisions in energy-related assets. Energy markets are uniquely complex due to high volatility, seasonality, regulatory impacts, physical constraints, and mean-reversion behaviors, rendering traditional financial models often inadequate. Recent advancements in artificial intelligence—particularly reinforcement learning (RL)—offer promising tools for managing such complexity. RL's strength lies in its focus on sequential decision-making and learning from interaction with uncertain environments, making it well-suited to financial domains where outcomes unfold over time. Unlike supervised learning, RL does not require pre-labeled data, enabling adaptive strategies based on evolving market feedback.

This review paper explores the intersection of RL and energy finance, with a particular emphasis on its applications in derivatives pricing, trading strategies, and risk management. While existing literature has reviewed RL in general finance and forecasting in energy markets separately, a comprehensive review focused on RL within energy finance remains lacking [132].

This gap is timely and significant. Global energy markets are undergoing rapid change due to decarbonization, renewable integration, and distributed energy resources—trends that introduce new uncertainty. Simultaneously, RL methods, especially deep reinforcement learning, are advancing rapidly and proving effective in similarly complex environments. The paper makes several contributions: it introduces a conceptual framework linking RL to energy finance, highlights the unique market characteristics suitable for RL, critically assesses current RL methodologies, and identifies open research questions. The review is intended for researchers in finance and machine learning, market practitioners, regulators, and students, providing both foundational knowledge and direction for future work.

This paper is structured around three pillars to explore the application of reinforcement learning (RL) in energy finance.

Pillar 1 – RL Foundations: This section outlines the theoretical basis of RL, focusing on the Markov Decision Process (MDP) framework and its relevance to sequential decision-making under uncertainty. It reviews key algorithms—from classical Q-learning to deep RL—and compares RL to traditional financial methods, highlighting RL's strengths in modeling complex, non-linear systems. Practical implementation issues, such as data needs and reward design, are also discussed.

Pillar 2 – Energy Market Characteristics: Energy markets exhibit features such as high volatility, seasonality, regulatory constraints, and asset-specific physical limitations. These complexities challenge traditional models but align well with RL's adaptive capabilities.

Pillar 3 – Application Domains: Three main application areas are identified:

1. **Forecasting and Trading** – RL for price prediction and strategy optimization.
2. **Derivatives Valuation** – RL methods for pricing energy options and real asset valuation.
3. **Power Systems Option Value** – RL for system reliability and flexibility valuation in smart grids.

Previous work on RL in finance has focused on traditional assets, with limited attention to energy-specific challenges. While foundational energy finance literature (e.g., Eydeland & Wolyniec; Weron) established key market dynamics, they primarily relied on parametric models. More recent research on smart grids and system optimization has largely excluded RL methods.

This paper addresses these gaps by linking RL theory to energy market complexities and demonstrating its potential in various financial applications. A comprehensive review of reinforcement learning (RL) in energy finance is both timely and necessary. Despite rapid growth in this field, research remains fragmented across finance, computer science, energy systems, and operations research, hindering cross-disciplinary knowledge sharing. Additionally, the fast evolution of RL methods and energy markets has outpaced practitioners' understanding of available tools and their suitability for specific applications. Structural changes in global energy systems—driven by decarbonization, technology, and renewables—further underscore the need for adaptive, data-driven approaches. Implementing RL in energy finance also demands interdisciplinary expertise, reinforcing the value of an integrated review.

This section provides foundational insights into RL’s theoretical principles in financial contexts, highlighting how RL supports complex decision-making under uncertainty. It then categorizes key RL algorithms based on their relevance to financial applications [133]:

- **Value-based methods** (e.g., Q-learning, DQN) estimate value functions and are well-suited for discrete decision problems such as trading strategies.
- **Policy gradient methods** (e.g., REINFORCE, TRPO, PPO) optimize policies directly and are effective in continuous control tasks like portfolio management.
- **Actor–critic methods** (e.g., A2C, DDPG) combine value estimation and policy learning, offering stability and flexibility for complex financial environments.
- **Model-based methods** (e.g., Dyna-Q, MBPO) build environment models for planning and simulation, improving data efficiency—an advantage in costly or data-scarce financial settings.

Together, these RL approaches offer promising solutions for the dynamic, uncertain, and data-constrained landscape of modern energy finance. Traditional financial modeling methods, while foundational, face significant limitations in addressing the complexity of energy markets.

Dynamic programming and stochastic control, grounded in the Hamilton–Jacobi–Bellman equation, offer rigorous solutions for sequential decision-making. However, they require fully specified system dynamics and reward structures, making them impractical for high-dimensional or complex environments like energy markets.

Monte Carlo simulations provide flexibility in handling uncertainty but depend on predefined stochastic models, which can introduce model risk if assumptions diverge from real-world behavior.

Parametric models, such as Black–Scholes or GARCH, offer interpretability and efficiency but rely on restrictive assumptions (e.g., normality, constant volatility) that often fail in the context of volatile and nonlinear energy markets.

Reinforcement learning (RL) addresses many of these shortcomings. Its model-free nature eliminates the need for explicit system modeling, and its integration with deep learning enables the capture of complex, non-linear relationships. RL's adaptability to new data allows for continuous policy updates in dynamic environments, and its flexible structure supports multiple objectives and constraints. Nevertheless, these benefits come with trade-offs, including increased data and computational demands, as well as challenges in interpretability—issues explored in later sections of this review. Energy commodities (e.g., electricity, natural gas, crude oil) exhibit unique price behaviors that distinguish them from traditional financial assets, posing distinct modeling challenges.

Extreme volatility is a defining feature—particularly in electricity markets—where prices can fluctuate by several hundred percent, far exceeding typical asset volatility. For example, during the 2021 Texas winter storm, wholesale electricity prices spiked to USD 9000/MWh, a 9000% increase, driven by surging demand and widespread generation outages.

Mean-reversion is another key characteristic. Unlike random walks observed in equities, energy prices tend to revert to levels anchored in production costs. Reversion speeds vary: electricity prices may normalize within hours, while commodities like natural gas may take months.

Price spikes are frequent and severe in energy markets, often triggered by weather events, infrastructure failures, or regulatory constraints. These transient but significant deviations pose challenges for traditional models, which struggle with non-normal distributions and sudden discontinuities.

Multi-scale temporal patterns further complicate modeling. Energy prices reflect intraday demand cycles, weekly and seasonal variations, and long-term shifts due to policy and technology changes. This layered structure requires methods capable of capturing both short-term volatility and long-term trends.

Traditional parametric models often fail to represent these complex dynamics due to restrictive assumptions. In contrast, **reinforcement learning (RL)** offers a flexible, data-driven approach that does not rely on predefined stochastic structures, enabling adaptive learning from empirical patterns.

Energy markets are governed by **complex temporal, physical, and financial dynamics** that challenge traditional financial modeling approaches. These complexities create opportunities for **reinforcement learning (RL)** methods, which offer greater adaptability and decision-making flexibility [134].

1. Temporal Structure and Cyclicality

Energy prices exhibit **multi-timescale temporal patterns**, including:

- **Intraday fluctuations**, with peak demand (typically 4–7 pm on weekdays) commanding price premiums.
- **Weekly cycles**, reflecting lower weekend consumption.
- **Seasonal variations**, driven by weather-related demand (e.g., heating in winter, cooling in summer).
- **Annual cycles** in commodities like natural gas, shaped by storage injection and withdrawal behaviors.

These overlapping temporal components form **dynamic, multi-layered price structures**. RL methods can model such dependencies **without requiring explicit parameterization** of cyclical components.

2. Physical Constraints and Embedded Optionality

Energy assets face **physical and operational constraints** that introduce **real optionality**:

- Power plants have **start-up costs**, **ramp-rate limits**, and **output thresholds**.
- Energy storage systems face **capacity limits**, **efficiency losses**, and **cycle degradation**.
- Transmission constraints lead to **locational price differentials**.

These operational realities generate **complex sequential decision problems**, which RL is well-suited to address due to its ability to **optimize over time under constraints**.

3. Market Incompleteness and Valuation Challenges

Energy markets exhibit **financial incompleteness**, creating valuation and risk management challenges:

- **Basis risk** arises when available hedging instruments do not match the specific temporal or locational exposure of physical assets.
- **Liquidity constraints** lead to wide bid–ask spreads, price impacts from large trades, and limit rebalancing—violating assumptions underpinning risk-neutral valuation.
- **Participant heterogeneity**, including diverse objectives and constraints, further complicates price formation and violates standard equilibrium assumptions.

These features undermine traditional no-arbitrage valuation models, requiring tools that can handle transaction costs, partial hedging, and market frictions—domains in which RL excels.

The inherent complexities of energy markets—including layered temporal patterns, embedded physical optionality, and market incompleteness—necessitate advanced modeling frameworks. Reinforcement learning provides a compelling alternative by offering data-driven, adaptive approaches capable of navigating the intricate, dynamic environments of energy systems. Its capacity to learn optimal policies under uncertainty, accommodate constraints, and adapt to structural change positions it as a transformative tool for modern energy finance.

Accurate price forecasting remains a cornerstone of energy trading and risk management. Traditional forecasting methodologies in energy markets include time series models such as ARIMA and GARCH, fundamental models based on supply–demand equilibrium, grey system theory, and classical machine learning algorithms including neural networks and support vector machines. Among these, grey system theory—originating from Deng (1982) [135]—has proven particularly useful in settings characterized by data scarcity and uncertainty, traits commonly associated with energy markets. Grey models such as GM(1,1) and its variants require minimal historical data and have shown competitive performance in electricity price forecasting. These models are particularly effective when dealing with non-stationary series or limited datasets and have been enhanced further through hybridization with wavelet transforms, neural networks, and residual correction strategies to improve forecasting accuracy across multiple horizons.

Reinforcement learning (RL) presents a fundamentally different paradigm for addressing forecasting problems by framing them as sequential decision-making tasks rather than static prediction exercises. This shift enables learning agents to adaptively update forecasting strategies in response to dynamic market environments. Rodrigues Dos Reis et al. (2025) [136] identified a clear trend toward incorporating advanced computational intelligence in medium- and long-term energy forecasting, emphasizing the value of dynamic approaches like RL that are capable of capturing evolving market structures and nonlinear dependencies.

Several streams of research have applied RL directly to energy price forecasting. The *direct forecasting approach* employs RL algorithms to predict future prices by learning from historical patterns and market feedback. For example, Pannakkong et al. (2023) [137] implemented a Double Deep Q-Network (DDQN) to dynamically select the optimal ensemble of machine learning models for peak electricity demand forecasting. This adaptive ensemble selection improved accuracy and demonstrated RL's capacity for real-time adaptation in volatile environments.

Adaptive forecasting is another promising area. An adaptive RL framework that selects forecasting models from a portfolio based on prevailing market conditions. In this meta-learning approach, the RL agent navigates a state space incorporating recent prices and exogenous features, while its action space consists of alternative forecasting models. Their results indicated superior performance, especially during regime shifts—an advantage over static forecasting strategies.

Energy forecasting has also benefited from advances in deep learning and hybrid modeling techniques. Hybrid decomposition techniques have further enhanced forecasting accuracy. A method combining Variational Mode Decomposition (VMD) with CNN-LSTM architectures for photovoltaic output forecasting, outperforming conventional models. These domain-specific enhancements underscore the importance of embedding physical and operational knowledge in forecasting models—an area where RL also excels, given its interactive learning framework.

The optimization of neural network architectures has emerged as a complementary research direction. A hybrid genetic algorithm–particle swarm optimization (GA-PSO) technique for tuning deep neural networks used in photovoltaic generation and energy consumption forecasting. Their method improved accuracy by up to 27% over standard models. This evolutionary optimization paradigm shares conceptual similarities with policy optimization in RL, suggesting opportunities for methodological cross-fertilization.

Reinforcement learning has also demonstrated value in feature selection and engineering. Q-learning within a blockchain-enabled smart grid context for electricity price forecasting approach for modeled grid operator–consumer interactions using a Stackelberg game framework, enabling real-time pricing optimization in multi-agent systems—an increasingly relevant scenario in decentralized and prosumer-driven energy markets.

Ensemble forecasting using RL has also gained traction. This method outperformed both individual models and static ensembles, particularly during market regime transitions. Such adaptability is critical in energy markets, where price dynamics are shaped by a confluence of stochastic demand, intermittent generation, and regulatory change.

Beyond forecasting, RL methodologies are increasingly employed in trading strategy development. The challenges of high volatility, non-stationarity, and intricate seasonality in energy prices require strategies that can adapt dynamically. Multy-agent deep reinforcement learning framework for bidding in day-ahead electricity markets, where agents approximated Nash equilibria through iterative learning in simulated environments. This model effectively captured strategic behavior in oligopolistic settings.

RL is also making inroads into energy risk management. Traditional risk metrics such as Value at Risk (VaR) and Conditional VaR (CVaR) have long underpinned energy finance. VaR quantifies the maximum expected loss over a given horizon at a specific confidence level, while CVaR captures tail risk beyond the VaR threshold. Advanced implementations of these metrics increasingly rely on machine learning. These developments provide a natural entry point for RL-based risk management solutions. By interacting with stochastic environments, RL agents can learn to dynamically adjust hedging strategies and risk exposure, potentially improving robustness in the face of market uncertainty.

This review has explored the diverse applications of reinforcement learning (RL) in the domain of energy finance, with a focus on key areas such as price forecasting, trading strategies, derivatives valuation, and the assessment of option value in power systems. The unique characteristics of energy markets—marked by extreme price volatility, pronounced seasonality, regulatory complexity, and the physical constraints of assets such as power plants and storage facilities—pose substantial modeling challenges that traditional financial approaches often struggle to address.

Reinforcement learning offers several critical advantages in tackling these challenges. Unlike classical methods that typically require simplifying assumptions about price dynamics or system behavior, RL learns directly from data through trial-and-error interaction with the environment. This data-driven approach enables RL to capture complex, nonlinear relationships that are pervasive in energy systems. Moreover, RL methods inherently adapt to non-stationary and evolving market conditions—an essential capability given the ongoing transformations in energy systems driven by decarbonization, decentralization, and digitalization. Another strength of RL lies in its ability to incorporate diverse operational, regulatory, and physical constraints directly into the learning and decision-making process, which is often difficult to achieve with traditional optimization models.

One of the most compelling applications of RL in energy finance lies in the quantification and capture of option value—the economic value derived from operational flexibility under uncertainty. Emerging technologies such as smart grids, energy storage systems, and demand-side response introduce significant real options. These options reflect the difference in expected system costs with and without access to flexible resources. Accurate valuation of this flexibility is essential for optimal investment planning and system operation.

Traditional approaches to option valuation in energy systems have relied heavily on stochastic optimization methods. While effective in many cases, these methods often become computationally prohibitive in high-dimensional settings and may struggle to model non-linear dependencies and sequential decision processes. In contrast, RL is particularly well-suited for this task due to its capacity for policy learning over long horizons, its ability to handle uncertainty in a dynamic and adaptive fashion, and its flexibility in modeling multi-stage decision problems [138].

Despite these advantages, several important challenges remain in the practical application of RL to energy finance. Interpretability remains a central concern, as the opaque nature of many RL algorithms limits their adoption in settings where transparency and regulatory compliance are paramount. Data scarcity can also hinder performance, particularly in settings where high-resolution market or operational data are limited. Generalization—the ability of an RL model trained in one context to perform effectively in another—remains a nontrivial hurdle, especially given the heterogeneity of energy markets across geographies and time. Furthermore, the computational demands of RL can be significant, particularly when modeling large-scale systems or incorporating high-frequency data. Finally, the absence of standardized benchmarks and evaluation protocols makes it difficult to compare methodologies and assess real-world readiness.

To address these limitations, several research directions warrant further exploration. First, developing interpretable and explainable RL models tailored for energy applications could enhance trust and facilitate adoption by industry stakeholders. Second, designing robust RL algorithms capable of performing reliably under extreme or previously unseen market conditions is crucial, especially in the context of climate change and geopolitical disruptions. Third, multi-agent reinforcement learning frameworks offer promise for modeling strategic interactions among heterogeneous market participants such as generators, consumers, aggregators, and regulators. The integration of sector coupling is another emerging area that will benefit from the flexible modeling capacity of RL.

In addition, hybrid approaches that combine RL with traditional methods such as stochastic optimization represent a particularly promising avenue for future research. In such frameworks, stochastic optimization can be used to define the scenario tree, uncertainty structure, and high-level planning boundaries, while RL algorithms are employed to discover fine-grained, adaptive operational policies within these structures. This integration could yield models that combine the rigor of optimization with the adaptability and learning capability of RL.

In conclusion, reinforcement learning represents a powerful and versatile methodology for addressing the unique challenges of energy finance. Its strengths in handling complexity, learning from data, adapting to change, and capturing the value of operational flexibility make it especially well-suited for the evolving landscape of modern energy systems. While several technical and practical hurdles remain, the growing body of research and early industrial applications suggest that RL is poised to significantly reshape how energy assets, contracts, and markets are valued, traded, and managed in the years ahead.