



# Review Reinforcement Learning in Energy Finance: A Comprehensive Review

Spyros Giannelos D

Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, UK; s.giannelos@imperial.ac.uk

Abstract: The accelerating energy transition, coupled with increasing market volatility and computational advances, has created an urgent need for sophisticated decision-making tools that can address the unique challenges of energy finance—a gap that reinforcement learning methodologies are uniquely positioned to fill. This paper provides a comprehensive review of the application of reinforcement learning (RL) in energy finance, with a particular focus on option value and risk management. Energy markets present unique challenges due to their complex price dynamics, seasonality patterns, regulatory constraints, and the physical nature of energy commodities. Traditional financial modeling approaches often struggle to capture these intricacies adequately. Reinforcement learning, with its ability to learn optimal decision policies through interaction with complex environments, has emerged as a promising alternative methodology. This review examines the theoretical foundations of RL in financial applications, surveys recent literature on RL implementations in energy markets, and critically analyzes the strengths and limitations of these approaches. We explore applications ranging from electricity price forecasting and optimal trading strategies to option valuation, including real options and products common in energy markets. The paper concludes by identifying current challenges and promising directions for future research in this rapidly evolving field.

**Keywords:** reinforcement learning; energy finance; option value; stochastic optimization; machine learning; risk management

# 1. Introduction

# 1.1. Overview

The intersection of energy markets and financial engineering has allowed for the rise of a specialized field known as energy finance. This domain encompasses the valuation of energy commodities, derivatives pricing, risk management, and investment decisions in energy infrastructure. The complexity of energy markets stems from several distinctive characteristics: high volatility, significant seasonality, mean-reversion tendencies, extreme price spikes, regulatory influences, and the physical constraints of energy production and delivery systems [1]. These complexities make conventional financial modeling approaches insufficient for many applications in energy finance.

Simultaneously, recent advances in artificial intelligence, particularly reinforcement learning (RL), have opened new avenues for addressing complex decision-making problems under uncertainty. RL differs from other machine-learning paradigms in its focus on sequential decision-making and delayed rewards, making it particularly suitable for financial applications where decisions unfold over time and outcomes become apparent only in the future [2]. Unlike supervised learning, which requires labeled examples of optimal



Academic Editors: Peter V. Schaeffer and Dimitris Damigos

Received: 29 March 2025 Revised: 26 April 2025 Accepted: 21 May 2025 Published: 23 May 2025

Citation: Giannelos, S. Reinforcement Learning in Energy Finance: A Comprehensive Review. *Energies* 2025, *18*, 2712. https:// doi.org/10.3390/en18112712

Copyright: © 2025 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/ licenses/by/4.0/). decisions, RL algorithms can learn through interaction with an environment, gradually improving their decision policies based on the rewards received.

The application of RL to energy finance represents a convergence of these two complex domains. The inherent volatility and structural complexities of energy markets create an ideal testing ground for RL methodologies, while the limitations of traditional approaches in capturing these complexities create a clear need for more sophisticated techniques. This review paper aims to systematically analyze how RL has been applied to various problems in energy finance, with particular attention to derivatives valuation and trading strategies.

While both reinforcement learning and energy finance represent active research areas individually, the intersection of these fields warrants dedicated review due to their rapid evolution and the unique challenges that arise in this convergence. Several excellent reviews exist in adjacent areas. Specifically, Fischer (2018) [2] surveyed reinforcement learning applications in general financial markets, focusing primarily on stock trading and portfolio optimization. In addition, Weron (2014) [3] comprehensively examined forecasting methodologies in electricity markets without a specific focus on reinforcement learning. However, a comprehensive review focusing specifically on reinforcement learning applications in energy finance, particularly in derivatives valuation and risk management, is notably absent from the literature. This gap is significant due to the distinctive characteristics of energy markets that create unique challenges and opportunities for reinforcement learning methodologies.

The need for this review is particularly timely for several reasons. First, energy markets globally are undergoing fundamental transformation driven by decarbonization policies, technological advances in renewable generation, and the emergence of distributed energy resources. These changes have introduced new sources of uncertainty and complexity that traditional modeling approaches struggle to address adequately. Second, recent advances in reinforcement learning, particularly deep reinforcement learning and its variants, have demonstrated remarkable success in complex decision domains with characteristics similar to those found in energy markets. Third, energy derivatives and structured products continue to evolve in complexity, creating both challenges for valuation and opportunities for novel methodological approaches. The convergence of these trends creates a compelling need for systematic assessment of how reinforcement learning can address the distinctive challenges of energy finance.

This review makes several specific contributions to the literature. First, it provides a unified conceptual framework for understanding reinforcement learning applications in energy finance, establishing clear connections between RL methodologies, energy market characteristics, and financial applications. Second, it systematically analyzes the distinctive features of energy markets—such as extreme price dynamics, physical constraints, and market incompleteness—that make them particularly suitable for reinforcement learning approaches while challenging for traditional methodologies. Third, it comprehensively examines current reinforcement learning methodologies applied to energy finance problems, critically evaluating their strengths, limitations, and comparative advantages over conventional approaches. Fourth, it identifies significant research gaps and promising future directions, providing a roadmap for researchers and practitioners seeking to advance this emerging field.

This comprehensive review is targeted at several key audiences. First, researchers in financial engineering and machine learning will find a systematic overview of how reinforcement learning techniques are being adapted to the unique challenges of energy markets, highlighting methodological innovations and performance benchmarks [4]. Second, energy market practitioners—including traders, risk managers, and investment analysts—will gain insights into cutting-edge quantitative tools that may enhance decision-making in increasingly complex and volatile markets. Third, policy makers and regulators concerned with energy market design and systemic risk will benefit from understanding how advanced algorithmic approaches may influence market behavior and efficiency. Finally, graduate students and early-career researchers entering this interdisciplinary field will find this review provides essential background knowledge and identifies promising research directions. By bridging theoretical foundations with practical applications, this paper aims to foster collaboration between academic research and industry practice in advancing reinforcement learning solutions for energy finance challenges.

The remainder of this paper is organized as follows: Section 2 provides the theoretical foundations of RL and its relevance to financial applications. Section 3 examines the specific characteristics of energy markets that make them suitable candidates for RL approaches. Section 4 reviews the literature on RL applications in energy price forecasting and trading strategy optimization. Section 5 focuses on derivatives valuation in energy markets using RL, including options pricing and real options analysis. Section 6 discusses implementation challenges and methodological considerations when applying RL to energy finance problems. Section 7 discusses option value in power systems, particularly regarding smart grid technologies. Section 8 concludes the paper with a synthesis of key findings and perspectives on the future evolution of this field.

## 1.2. Illustration

Figure 1 presents the conceptual framework that organizes our comprehensive review of reinforcement learning applications in energy finance. The framework illustrates the three fundamental pillars of our analysis: reinforcement learning foundations, energy market characteristics, and application domains.



Figure 1. Conceptual framework of reinforcement learning applications in energy finance.

The first pillar, RL Foundations, encompasses the theoretical underpinnings of reinforcement learning methodologies relevant to energy finance. The Markov Decision Process (MDP) framework provides the mathematical structure for sequential decision-making under uncertainty, including states, actions, transition probabilities, reward functions, and discount factors that form the basis of RL algorithms. We review key RL algorithms applicable to energy finance problems, from classical methods like Q-learning to advanced approaches such as deep reinforcement learning. The framework also highlights comparisons between RL and traditional financial methods, emphasizing the distinctive advantages of RL in handling complex, non-linear dynamics. Implementation considerations address practical aspects of applying RL to energy finance, including data requirements, computational needs, state and action space design, and reward function formulation.

The second pillar, Energy Markets Characteristics, identifies the distinctive features that make energy markets particularly suitable for RL applications. These include complex price dynamics (volatility, mean-reversion, jumps), multi-layer seasonality patterns, regulatory and market structure factors, physical constraints of energy assets, and market incompleteness challenges. These characteristics create both challenges for traditional modeling approaches and opportunities for RL methodologies.

The third pillar organizes the application domains into three major categories. The first category, Forecasting and Trading, covers RL applications in energy price prediction and optimal trading strategy development, including risk management approaches. The second category, Derivatives Valuation, examines RL methods for pricing various energy options and analyzing real options embedded in physical assets. The third category, Option Value in Power Systems, focuses on applications specific to electricity systems, including VaR/CVaR approaches for system reliability and valuation of flexibility provided by smart grid technologies.

The framework concludes with Challenges and Future Research Directions, identifying current limitations in applying RL to energy finance and promising avenues for future research. This comprehensive structure provides a roadmap for understanding how reinforcement learning is transforming analysis and decision-making in energy finance while highlighting areas where further methodological development is needed.

#### 1.3. Literature Review and Research Gaps

This section systematically reviews the existing literature at the intersection of reinforcement learning and energy finance, identifying key research streams and critical gaps that motivate this paper.

#### 1.3.1. Reinforcement Learning in Financial Markets

Reinforcement learning has gained significant traction in financial applications over the past decade. Fischer (2018) [2] provided a comprehensive survey of reinforcement learning applications in general financial markets, focusing primarily on equity trading, portfolio optimization, and traditional asset classes. This work established RL's potential for sequential decision-making under uncertainty but concentrated predominantly on conventional securities markets rather than commodity or energy-specific applications.

Similarly, foundational work on reinforcement learning methodologies by Sutton and Barto (2018) [5] established theoretical frameworks applicable across domains but did not address the specific challenges of energy markets. The ability of RL algorithms to learn optimal policies through interaction with complex environments, as demonstrated by Mnih et al. (2015) [6] in other contexts, suggests particular promise for energy finance applications, though this connection remains underdeveloped in the existing literature.

## 1.3.2. Energy Finance Modeling Approaches

Traditional energy finance approaches have evolved to address sector-specific challenges but often struggle with the full complexity of modern energy markets. Eydeland and Wolyniec (2003) [1] developed foundational frameworks for energy and power risk management, identifying distinctive characteristics that include high volatility, significant seasonality, mean-reversion tendencies, extreme price spikes, regulatory influences, and physical constraints. However, their methodologies primarily relied on parametric models and conventional stochastic processes that face limitations in capturing the full complexity of energy market dynamics.

In the specific area of electricity price forecasting, Weron (2014) [3] comprehensively examined various methodologies, including time series models, artificial intelligence techniques [7–20], and fundamental approaches. While this work acknowledged the unique challenges of electricity markets, it did not specifically explore reinforcement learning's potential for addressing these challenges or connect forecasting to broader financial decision-making frameworks.

Energy derivatives valuation has received attention from researchers, including Carmona and Coulon (2014) [21], who examined structural models for electricity prices, and Benth et al. (2008) [22], who developed stochastic modeling approaches for electricity markets. These works established sophisticated mathematical frameworks but generally relied on closed-form solutions or Monte Carlo methods rather than learning-based approaches capable of handling market incompleteness and complex constraints.

#### 1.3.3. Energy System Operations with Learning-Based Methods

A separate research stream has focused on optimization and learning methods for energy system operations.

The operational challenges of power generation assets have been examined by researchers, including Conejo et al. (2010) [23], who developed decision-making frameworks under uncertainty, and Thompson et al. (2009) [24], who studied energy storage valuation and optimization. These works established the complex optimization problems inherent in energy systems but generally treated financial considerations as secondary to technical constraints and reliability objectives.

More recently, smart grid technologies have introduced new flexibility options into energy systems, as demonstrated by Konstantelos et al. (2017) [25], Giannelos et al. (2018) [26], and other works focused on option value and stochastic optimization [27]. While these studies incorporate uncertainty and flexibility valuation, they typically employ conventional stochastic optimization methods rather than reinforcement learning approaches.

#### 1.3.4. Research Gaps and Contributions

Based on this literature review, several critical research gaps emerge at the intersection of reinforcement learning and energy finance:

Gap 1: Lack of an integrated conceptual framework. While separate bodies of literature address reinforcement learning for finance and optimization methods for energy systems, a comprehensive framework connecting RL methodologies to the specific characteristics of energy markets is notably absent. This paper addresses this gap by establishing clear connections between RL methodologies, energy market characteristics, and financial applications, providing a unified conceptual structure (as illustrated in Figure 1) that bridges previously disconnected research streams.

Gap 2: Insufficient analysis of energy market features requiring specialized RL approaches. The existing literature has not systematically analyzed which distinctive features of energy markets—such as extreme price dynamics, physical constraints, and market

incompleteness—make them particularly suitable for reinforcement learning approaches. This paper provides this analysis in Section 3, establishing a foundation for understanding why conventional methods may fall short and how RL can address these limitations.

Gap 3: Limited comparative assessment of RL methodologies for energy finance applications. While various RL algorithms have been applied to isolated energy finance problems, a comprehensive assessment of their relative strengths and weaknesses across different application domains is missing. This paper addresses this gap in Section 2.2 and throughout application-specific sections, evaluating algorithm suitability for different energy finance challenges.

Gap 4: Absence of a comprehensive review of real options analysis with RL in energy systems. Despite the significant embedded optionality in energy assets and infrastructure, the existing literature lacks a comprehensive treatment of how RL can enhance real options valuation in this context. This paper fills this gap in Sections 5.3 and 6, connecting option theory with reinforcement learning to provide new perspectives on flexibility valuation.

Gap 5: Fragmented understanding of option value in power systems. The literature on option value in power systems, particularly regarding smart grid technologies, has developed separately from the RL literature, limiting cross-fertilization between these fields. This paper bridges this divide in Section 6, examining how RL methods can enhance option valuation for smart grid investments.

By addressing these gaps, this paper makes several novel contributions to the literature. First, it provides the first comprehensive review specifically focused on reinforcement learning applications in energy finance, creating a reference point for researchers and practitioners working across these domains. Second, it establishes a conceptual framework that organizes existing and future research, clarifying how different RL methodologies align with specific energy finance challenges. Third, it systematically evaluates the comparative advantages of RL approaches over traditional methods across multiple application domains. Finally, it identifies promising research directions and methodological improvements that could further advance this emerging field.

This review is particularly timely due the accelerating energy transition [1,2,4,21,23,26–89], which is creating new sources of uncertainty and complexity in energy markets that conventional modeling approaches struggle to address adequately. Simultaneously, recent advances in reinforcement learning, particularly deep reinforcement learning and its variants, have demonstrated remarkable success in complex sequential decision domains with characteristics similar to those found in energy markets.

# 2. Theoretical Foundations of Reinforcement Learning in Energy Finance

## 2.1. Reinforcement Learning Framework

A comprehensive review of reinforcement learning applications in energy finance is both timely and necessary for several compelling reasons. First, research at this intersection has grown exponentially in recent years but remains fragmented across multiple disciplines, including finance, computer science, energy systems, and operations research. This fragmentation creates barriers to knowledge transfer and impedes the identification of common methodological challenges and solutions. Second, the rapid evolution of both reinforcement learning techniques and energy market structures means that practitioners and researchers often lack awareness of the full spectrum of available approaches and their relative strengths for specific energy finance problems. Third, energy markets worldwide are undergoing fundamental transformation driven by decarbonization policies, technological change, and increasing penetration of renewable resources, creating new valuation and risk management challenges that traditional methods struggle to address. Finally, the practical implementation of reinforcement learning in energy finance requires interdisciplinary expertise that is rarely found within a single research group or company, highlighting the need for a synthesized review that bridges these knowledge domains.

By reviewing the theoretical underpinnings of reinforcement learning in a financial context, this section provides the foundational understanding necessary to appreciate the unique advantages RL offers for addressing the complex decision-making challenges in energy markets. The remainder of this section examines the core components of reinforcement learning and their specific relevance to financial applications.

Reinforcement learning is a computational method for learning how to make optimal decisions through interactions with an environment (Sutton and Barto, 2018) [51]. The core framework of RL is the Markov Decision Process (MDP), which includes:

- A set of states *S* representing the environment
- A set of actions *A* available to the agent
- Transition probabilities P(s'|s, a) defining how actions lead to new states
- A reward function *R*(*s*, *a*, *s*') providing feedback on action quality
- A discount factor γ determining the relative importance of immediate versus future rewards

In this framework, an agent learns a policy  $\pi$  that maps states to actions, with the goal of maximizing the expected cumulative discounted reward over time (Sutton and Barto, 2018) [51]. The value function  $V^{\pi}(s)$  represents the expected return starting from the state *s* and following policy  $\pi$  thereafter:

$$V^{\pi}(s) = \mathbb{E}_{\pi} \Big[ \sum_{t=0}^{\infty} \gamma^t R_{t+1} \big| S_0 = s \Big]$$

Similarly, the action–value function  $Q^{\pi}(s, a)$  represents the expected return starting from the state *s*, taking action *a*, and following policy  $\pi$  thereafter:

$$Q^{\pi}(s,a) = \mathbb{E}_{\pi} \left[ \sum_{t=0}^{\infty} \gamma^t R_{t+1} \big| S_0 = s, \quad A_0 = a \right]$$

The optimal policy  $\pi$  maximizes these value functions, yielding the optimal value function  $V^s$  and optimal action–value function  $Q^*(s, a)$ .

## 2.2. RL Algorithms Relevant to Financial Applications

The application of reinforcement learning to finance has been facilitated by several classes of algorithms, each with distinct characteristics that make them suitable for different financial problems.

Value-based methods focus on learning the value function or action-value function from which a policy is derived. Q-learning, introduced by Watkins and Dayan (1992) [90], and its neural network extension, Deep Q-Networks (DQN) [6], have found significant applications in financial domains. These approaches excel in environments with discrete action spaces, such as binary trading decisions or discrete investment choices, due to their ability to estimate the expected return of each possible action precisely.

Policy gradient methods, in contrast, directly parameterize and optimize the policy without explicitly computing a value function. This category includes algorithms such as REINFORCE [91], Trust Region Policy Optimization (TRPO) [92], and Proximal Policy Optimization (PPO) [93]. The strength of these methods lies in their ability to handle continuous action spaces effectively, making them particularly valuable for portfolio allocation, hedging decisions, and other financial applications requiring nuanced control.

Actor–critic methods represent a hybrid approach that maintains both a value function approximator (the critic) and a separate policy representation (the actor). Prominent examples include Advantage Actor–Critic (A2C) and Deep Deterministic Policy Gradient (DDPG) [11]. These methods have demonstrated effectiveness in complex financial environments by combining the stability of value-based methods with the capability to handle continuous actions. This dual structure allows for more efficient learning in the intricate and often non-stationary conditions characteristic of financial markets.

Model-based RL algorithms learn an explicit model of the environment's dynamics to facilitate planning and decision-making. Notable implementations include Dyna-Q [94] and more recent approaches such as Model-based Policy Optimization (MBPO) [84]. The data efficiency of these methods presents a significant advantage in financial applications, where data acquisition may be limited or costly. By learning to predict market behavior, these algorithms can simulate potential outcomes of different strategies without requiring actual market interaction, potentially reducing both risk and the data requirements for effective learning.

## 2.3. RL vs. Traditional Financial Modeling Approaches

Financial modeling and decision-making have historically relied on several wellestablished methodologies, each with distinct characteristics and limitations when applied to complex markets such as those in the energy sector.

Dynamic programming and stochastic control techniques, formalized through the Hamilton–Jacobi–Bellman equation, provide mathematically rigorous frameworks for sequential decision-making under uncertainty. Despite their theoretical elegance, these approaches typically require explicit specification of system dynamics and reward functions, rendering them computationally intractable for high-dimensional problems or systems with complex transition dynamics [95]. This limitation is particularly relevant in energy markets, where multiple interacting factors influence price dynamics.

Monte Carlo simulation methods have been widely employed to address uncertainty in financial modeling by generating numerous random scenarios to estimate expected outcomes. While effective for many applications, these techniques generally necessitate a predefined model of the underlying stochastic processes, potentially introducing model risk when the specified processes deviate from actual market behavior [68].

Parametric models, such as the Black–Scholes framework for option pricing or GARCH models for volatility forecasting, rely on specific assumptions about the underlying stochastic processes. Although these models offer computational efficiency and interpretability, their underlying assumptions—including normality of returns, constant volatility, or specific mean-reversion properties—often fail to capture the complex dynamics observed in energy markets [80].

Reinforcement learning presents several comparative advantages in addressing these limitations. First, the model-free nature of many RL algorithms enables learning optimal policies without requiring explicit specification of environmental dynamics, a valuable characteristic when these dynamics are complex, unknown, or difficult to parameterize. Second, RL approaches, particularly when implemented with deep neural networks, demonstrate superior capacity for capturing non-linear relationships that resist effective parametric modeling [96]. Third, RL frameworks inherently accommodate adaptability through continuous policy updates based on new observations, allowing them to respond to evolving market conditions. Finally, RL methodologies can naturally incorporate complex constraints and multiple objectives that may prove challenging to formulate within closed-form optimization problems.

However, these advantages must be weighed against certain trade-offs involving interpretability, data requirements, and computational complexity—considerations that will be examined in subsequent sections of this review.

# 3. Characteristics of Energy Markets Relevant to RL Applications

Energy markets possess several distinctive characteristics that make them both challenging for traditional modeling approaches and suitable candidates for RL applications.

#### 3.1. Price Dynamics and Volatility

Energy commodities (electricity, coal, natural gas, crude oil, etc.) exhibit distinctive price dynamics that differentiate them from conventional financial assets (stocks, bonds, currencies, etc.), presenting unique challenges for modeling and trading strategies. These dynamics can be characterized by several key features that make traditional financial models often inadequate.

Energy markets display exceptional volatility, particularly in electricity markets, where price amplitudes significantly exceed those observed in conventional securities markets. While typical financial assets may experience annual volatility of 20–30%, electricity prices can undergo fluctuations of several hundred percent within equivalent timeframes [3]. This extraordinary volatility is primarily attributable to the limited storability of electricity and the necessity for instantaneous balance between supply and demand. The 2021 Texas winter storm provides a striking illustration, with wholesale electricity prices reaching the market cap of USD 9000/MWh, representing an approximately 9000% increase from typical levels [97].

Unlike many financial assets that follow random walk processes, energy prices typically exhibit mean-reverting behavior. This tendency to return to fundamental equilibrium levels occurs because energy prices are intrinsically linked to production costs. When prices deviate significantly from these costs, market mechanisms induce corrective movements—excessive prices stimulate increased production, while depressed prices lead to supply contraction. The mean-reversion rate varies considerably across energy commodities, with electricity prices potentially reverting within hours, while natural gas might require months to return to equilibrium levels [98].

A distinctive characteristic of energy markets, particularly electricity, is the occurrence of extreme price spikes. These episodic events manifest as transient but dramatic price increases, potentially orders of magnitude above normal levels. Such spikes typically result from supply constraints, extreme weather events, or technical failures in generation or transmission infrastructure. During the aforementioned 2021 Texas winter storm, the confluence of increased heating demand and widespread generation outages produced price spikes reaching the market cap [97]. These non-normally distributed events present significant challenges for conventional modeling approaches while creating opportunities for adaptive algorithms capable of recognizing precursory patterns [83].

Energy price dynamics operate across multiple overlapping timescales, creating complex temporal structures. These include intraday patterns reflecting diurnal demand fluctuations, weekly cycles distinguishing between workdays and weekends, seasonal variations driven by weather-dependent consumption, and long-term trends reflecting technological and regulatory evolution [99]. This multi-layered temporal structure necessitates modeling approaches capable of simultaneously capturing short-term fluctuations and long-term evolutionary patterns.

These complex dynamics exceed the capabilities of traditional parametric models, which typically rely on simplifying assumptions inappropriate for energy markets. This limitation creates a compelling opportunity for reinforcement learning approaches, which can learn directly from empirical data without imposing restrictive structural assumptions.

These theoretical price dynamics are vividly illustrated by several historical events in electricity markets. During the 2021 Texas winter storm, wholesale electricity prices in ERCOT reached the market cap of USD 9000/MWh, representing an approximately

9000% increase from typical levels of around USD 20–30/MWh [50,100]. This extreme price spike reflected both the physical non-storability of electricity and supply–demand imbalance when approximately 48.6% of generation capacity was forced offline due to weather conditions, while heating demand simultaneously surged. Following this crisis, prices rapidly reverted to normal levels once generation facilities were restored and demand normalized, demonstrating the mean-reverting characteristic discussed above.

The Australian National Electricity Market provides another instructive example of complex price dynamics. Specifically, Higgs and Worthington (2008) [76] documented that this market exhibited mean-reverting behavior with both intraday and seasonal patterns, but it also experienced frequent extreme price spikes. Their analysis showed that these spikes followed distinct statistical distributions that conventional models struggled to capture, highlighting the challenge for traditional pricing approaches.

Natural gas markets demonstrate different temporal dynamics but similar complexity. In their comprehensive analysis, Nick and Thoenes (2014) [101] showed that European natural gas prices exhibit mean-reversion at multiple time scales—short-term reversions following supply disruptions or weather events, and longer-term reversions toward production costs. Their study documents how these dynamics interact with seasonality patterns and storage levels to create complex price behaviors that cannot be adequately modeled by conventional stochastic processes.

#### 3.2. Seasonality and Cyclicality

Energy markets exhibit pronounced temporal patterns across multiple timescales, creating complex cyclical structures in price formation. These patterns manifest through intraday fluctuations reflecting diurnal demand variations, with peak consumption hours typically commanding price premiums. Weekly cycles emerge from the distinct consumption profiles of weekdays versus weekends, while seasonal variations are predominantly driven by weather-dependent demand—heating requirements during winter months and cooling demand during summer periods in most regions. Certain energy commodities, particularly natural gas, display marked annual cyclicality attributable to storage injection—withdrawal cycles and seasonal consumption patterns.

The interaction of these temporal components creates a multi-layered structure that evolves dynamically in response to changing consumption behaviors, technological advancements, and regulatory modifications. Reinforcement learning methodologies offer the potential to capture these intricate temporal dependencies without necessitating explicit parameterization of individual cyclical components.

These theoretical patterns are clearly observable in empirical data across energy markets. Examining the PJM electricity market, Knittel and Roberts (2005) [102] documented pronounced diurnal patterns with peak/off-peak price differentials averaging 25–45%, depending on the season. Their analysis identified predictable load patterns driving these cycles, with price peaks typically occurring between 4–7 pm on weekdays. Beyond daily patterns, they found weekly cycles, with Sunday prices averaging 15–30% below Tuesday–Thursday levels due to lower commercial and industrial activity.

Natural gas markets provide a striking illustration of annual seasonality. Analyzing the Henry Hub benchmark, Suenaga et al. (2008) [103] documented how the injection–withdrawal cycle creates predictable price patterns, with late-summer-to-early-fall prices (during peak storage injection) historically averaging 10–15% below winter prices (during peak withdrawal). This seasonality interacts with storage inventory levels—Brown and Yücel (2008) [104] showed that when storage levels fall significantly below 5-year averages, winter price premiums can expand dramatically, sometimes exceeding 50% above summer prices.

European energy markets demonstrate how these cyclical patterns can evolve with changing consumption behaviors. Particularly, Paraschiv et al. (2015) [105] analyzed German electricity markets following substantial renewable integration, finding that traditional seasonality was increasingly overlaid with renewable generation cycles. Their research documented how solar generation created midday price depressions (sometimes resulting in negative prices) that altered the traditional peak/off-peak pattern, demonstrating how technological change can reshape fundamental market dynamics.

## 3.3. Regulatory and Market Structure Considerations

Energy markets operate within complex regulatory frameworks that substantially influence price formation mechanisms and market dynamics. Market design varies significantly across jurisdictions, ranging from fully liberalized structures to partially regulated environments, each with distinctive price formation processes [87]. Many electricity markets incorporate separate capacity mechanisms that provide supplementary revenue streams for generators, further complicating asset valuation and investment decision-making [40].

Renewable energy integration policies, including subsidies, feed-in tariffs, and priority dispatch provisions, significantly alter market dynamics and can precipitate negative price episodes [106]. Additionally, carbon pricing mechanisms and environmental regulations introduce further complexity to energy price formation [48]. These regulatory factors create regime-dependent dynamics that traditional modeling approaches struggle to accommodate. The adaptive learning capabilities of reinforcement learning algorithms are particularly suited to navigating these regulatory complexities.

The impact of regulatory frameworks on energy price formation is clearly illustrated by comparing market designs and policy impacts across different jurisdictions. Comparing the PJM and ERCOT electricity markets, Potomac Economics and Electric Reliability Council of Texas (2020) [107] documented how their structural differences created divergent price dynamics despite similar underlying fundamentals. While both markets use locational marginal pricing, PJM's capacity market provides generators with a separate revenue stream beyond energy prices, resulting in less extreme price volatility during scarcity conditions compared to ERCOT's energy-only design. During comparable reserve shortage events, maximum real-time prices reached significantly higher levels in ERCOT than in PJM, demonstrating how market design fundamentally shapes price behavior.

The impact of renewable energy policies is evident in Germany's electricity market transformation. Specifically, Ketterer (2014) [106] empirically analyzed how Germany's renewable integration policies, particularly solar subsidies and priority dispatch provisions, fundamentally altered market dynamics. Her econometric analysis documented a 36% reduction in average daily price levels between 2006–2012 attributable to solar and wind penetration, along with increased volatility. Furthermore, the study identified 40 negative price episodes during that period, a phenomenon virtually non-existent before these policies.

Carbon pricing mechanisms provide another example of regulatory impacts on energy markets. Analyzing the EU Emissions Trading System, Fabra and Reguant (2014) [48] found that power producers passed through approximately 80% of carbon prices to whole-sale electricity prices, demonstrating how environmental regulations directly influence price formation. Their study showed that pass-through rates varied significantly across market conditions and generator types, creating complex interactions between carbon and electricity price dynamics.

#### 3.4. Physical Constraints and Real Options

Energy assets are subject to substantial physical constraints that generate embedded optionality in their operation. Power generation facilities face operational limitations, including minimum and maximum output thresholds, ramp rate restrictions, and startup/shutdown costs that create complex optimization problems [23]. Energy storage facilities, such as natural gas storage or hydroelectric reservoirs, operate under capacity constraints, injection/withdrawal rate limitations, and cycle efficiency losses [24]. Additionally, transmission infrastructure constraints can induce locational price differentials and restrict arbitrage opportunities [77].

These physical constraints create real options that present significant valuation challenges for traditional methodologies. The sequential decision-making framework inherent in reinforcement learning approaches aligns naturally with the temporal exercise of these real options, offering advantages over conventional valuation techniques.

The operational constraints of power generation units create complex optimization challenges with significant economic implications. In their analysis of gas-fired power plants, Staffell and Green (2016) [108] documented how start-up costs ranged from GBP 3000–GBP 30,000 per start (approximately USD 4000–USD 40,000), depending on plant size and technology, while ramp rates limited output changes to 2–7% of capacity per minute. These constraints transformed the simple spark spread calculation into a complex optionality valuation problem that traditional methodologies struggle to accurately capture.

Transmission constraints frequently create locational price differentials that reflect embedded real options in the energy system. In a comprehensive analysis of the PJM market, Woo et al. (2011) [109] documented congestion-driven price differences exceeding USD 50/MWh between neighboring nodes during approximately 15% of hours studied. These differentials reflect the option value of transmission capacity, which varies with system conditions, demand patterns, and generation availability—a complexity well-suited to reinforcement learning approaches.

Energy storage facilities operate under multi-dimensional constraints that generate sophisticated optionality. Analyzing grid-scale battery storage, Staffell and Rustomji (2016) [110] detailed how cycle degradation (0.2–0.5% capacity loss per full cycle), depth-of-discharge limitations, and round-trip efficiency losses (15–25%) created complex trade-offs between short-term revenue opportunities and long-term asset value. These characteristics create real options that require advanced modeling techniques to value appropriately, particularly as storage technologies evolve and market services expand.

#### 3.5. Market Incompleteness and Liquidity Constraints

Energy markets exhibit characteristics of financial incompleteness that impede comprehensive risk management. Basis risk emerges when standardized trading instruments fail to match the temporal or locational specificity of physical energy exposures. This mismatch between available hedging instruments and actual physical positions creates inherent inefficiencies, necessitating risk premiums to compensate for unhedgeable exposures [1].

Liquidity constraints represent a related challenge, as many energy derivatives markets exhibit limited depth. Restricted participation results in widened bid–ask spreads, elevated transaction costs, and potential price impacts from large transactions. This thin trading environment undermines a fundamental assumption of risk-neutral valuation continuous portfolio rebalancing without price impact. The inability to establish and maintain cost-effective perfect hedges compromises the theoretical foundation of traditional pricing models, necessitating additional risk premiums that cause deviations from theoretical values. Furthermore, energy markets encompass heterogeneous participants with divergent objectives, operational constraints, and risk preferences. This diversity in market participation leads to complex price formation dynamics that may not conform to standard equilibrium assumptions [111]. These market imperfections challenge traditional valuation methodologies predicated on no-arbitrage principles and market completeness. Reinforcement learning approaches offer alternative methodologies capable of incorporating transaction costs, liquidity constraints, and participant heterogeneity.

Basis risk in energy hedging creates significant challenges that illustrate market incompleteness. Examining natural gas markets, Brinkmann and Rabinovitch (1995) [112] documented basis risk between Henry Hub futures and 28 delivery locations, finding correlations ranging from 0.42 to 0.96. This incomplete correlation meant that even "hedged" positions retained substantial exposure, with risk reduction ranging from 18% to 92% across locations. Their analysis demonstrated how geographical specificities prevented perfect hedging even with standardized instruments.

The liquidity constraints in energy derivatives markets significantly impact trading costs and risk management. In their analysis of electricity forward markets, Frestad et al. (2010) [51] found bid–ask spreads in the Nordic electricity market (Nord Pool) ranging from 0.5% for front-month contracts to over 4% for quarters beyond 1 year. These transaction costs materially impacted hedging effectiveness and implied that continuous portfolio rebalancing—a key assumption in many valuation models—was economically infeasible beyond short time horizons.

Market participant heterogeneity further contributes to energy market incompleteness. Analyzing the UK electricity market, Karakatsani and Bunn (2008) [113] documented how different categories of participants (generators, suppliers, financial traders) systematically valued forward contracts differently based on their physical positions and risk preferences. Generators consistently valued forward contracts at a 5–12% discount to expected spot prices, while suppliers paid a 3–8% premium, creating a persistent risk premium that violated risk-neutral pricing assumptions. This heterogeneity demonstrates how energy markets operate with multiple subjective valuations rather than the single risk-neutral measure assumed by complete market theory.

# 4. RL for Energy Price Forecasting and Trading Strategies

Before examining specific applications in detail, Table 1 provides a systematic classification of reinforcement learning implementations in energy trading and forecasting. This classification organizes the literature by algorithm type, application domain, data characteristics, and key findings, offering a structured framework for understanding the evolving landscape of RL applications in energy markets. The table highlights the progression from traditional RL algorithms toward more sophisticated approaches, including distributional and multi-agent frameworks, reflecting the increasing complexity of energy market challenges being addressed. This classification also reveals patterns in methodological choices for different problem types, with value-based methods predominating in forecasting applications and policy-based approaches showing particular strength in trading strategy development. As the subsequent sections elaborate on these applications, this classification serves as a reference point for identifying methodological trends and comparative performance across different market contexts. **Table 1.** Classification of Reinforcement Learning Applications in Energy Trading and Forecasting. This table categorizes key studies by algorithm type, application domain, data characteristics, and principal findings, illustrating the diverse approaches to implementing RL in energy market prediction and trading strategy development (2018–2024).

Study	RL Algorithm	Energy Application	Data Characteristics	Key Findings
Jiang and Powell (2018) [85]	Value iteration with function approximation	Ensemble forecasting of electricity prices	PJM hourly price data	RL-based ensembles adapt better to regime changes than static ensemble methods
Boukas et al. (2020) [114]	Proximal policy optimization	Intraday electricity trading	Nord Pool intraday market	RL strategy outperforms benchmark strategies by 15–28% in risk-adjusted returns
Du et al. (2021) [46]	Multi-agent DQN	Bidding strategy in day-ahead markets	ERCOT market data	Multi-agent approach effectively approximates Nash equilibrium solutions
Moti (2022) [115]	Q-learning	Electricity price prediction in blockchain-based grid	Simulated smart grid environment	RL framework mediates operator–consumer interactions for price prediction
Pannakkong et al. (2023) [116]	Double deep Q-network	Peak electricity demand forecasting	Thailand's electricity demand data	DDQN outperformed individual ML models by dynamically selecting optimal models
Guo and Wang (2020) [72]	Deep Q-network	Adaptive model selection for price forecasting	ISO New England data	RL framework reduced MAPE by 18% compared to the best individual model
Cao et al. (2023) [30]	Deep distributional RL	Options portfolio hedging	Simulated & empirical energy data	Outperformed delta hedging by 22–30% in managing non-linear risks
Mulliez (2021) [117]	Q-learning	Dynamic hedging with basis risk	Natural gas basis spreads	Adaptive hedging outperformed traditional approaches under time-varving risks
Chen et al. (2020) [34]	Hybrid RL + supervised learning	Energy portfolio hedging	Futures and spot price data	Cross-learning improved profit-risk tradeoffs vs. static hedging
Karimi Madahi et al. (2024) [118]	Distributional RL	Battery storage arbitrage	UK imbalance settlement prices	Captured asymmetric risk profiles better than expected value methods

## 4.1. Energy Price Forecasting with RL

Accurate price forecasting constitutes a fundamental component of energy trading and risk management. Traditional forecasting methodologies in energy markets encompass time series models (ARIMA, GARCH), fundamental models based on supply-demand balances, grey system theory models, and conventional machine-learning approaches, including neural networks and support vector machines [3]. The grey system theory, introduced by Deng (1982) [43], offers prediction methods particularly suited for systems with limited and uncertain information-characteristics often present in energy markets. Grey prediction models, notably GM (1, 1) and its variants, have demonstrated effectiveness in electricity price forecasting by requiring minimal historical data while maintaining reasonable accuracy [119,120]). These methods are especially valuable when dealing with non-stationary series and limited samples, complementing traditional statistical approaches. Grey models have been further enhanced through hybrid approaches, combining them with wavelets, neural networks, or residual correction mechanisms to improve forecasting performance across different time horizons (Zhao et al., 2015) [121]. Reinforcement learning offers a distinctive paradigm for addressing forecasting challenges by reformulating prediction as a sequential decision-making problem rather than a static estimation task.

Energy forecasting methodologies have evolved significantly in recent years, with numerous innovations emerging in 2024–2025. Specifically, Mystakidis et al. (2024) [122] provided a comprehensive review of energy forecasting techniques across different time horizons, highlighting how deep-learning approaches are increasingly outperforming

traditional methods in capturing complex patterns in energy data. Several noteworthy advances have emerged in neural network architectures optimized for energy forecasting. Also, Majeske et al. (2025) [17] introduced dynamic attention neural networks (A-RNN) for industrial energy forecasting, demonstrating how attention mechanisms can capture spatiotemporal dependencies in multi-device systems while providing interpretability through learned feature importance. This approach shares conceptual similarities with attention-based reinforcement learning algorithms, suggesting potential integration paths between these methodologies.

Hybrid decomposition techniques combined with deep learning have shown remarkable results in renewable energy forecasting. Boucetta et al. (2024) [123] developed a novel approach combining Variational Mode Decomposition (VMD) with CNN-LSTM architectures for photovoltaic power forecasting, significantly outperforming conventional deep-learning methods. Similarly, Famoso et al. (2024) [49] integrated Artificial Neural Networks with stochastic dependability modeling for wind power forecasting, achieving substantial accuracy improvements by accounting for operational uncertainties like turbine failures. These approaches demonstrate the value of incorporating domain-specific knowledge into forecasting models, a principle that reinforcement learning naturally extends by learning optimal forecasting policies through environment interaction.

Optimization of neural network architectures has emerged as another promising direction. Specifically, Hosseini et al. (2025) [78] proposed a hybrid GA-PSO approach that optimizes Deep Neural Networks for energy consumption and photovoltaic production forecasting, achieving up to 27% accuracy improvements over traditional methods. This evolutionary optimization shares conceptual similarities with policy optimization in reinforcement learning, suggesting potential cross-fertilization between these approaches.

In their survey of medium- and long-term energy forecasting methods, Rodrigues Dos Reis et al. (2025) [124] highlighted the progressive shift toward advanced computational intelligence that can handle increasingly complex data and environment interactions. This trend aligns with the reinforcement learning paradigm discussed in this review, which reformulates forecasting as a sequential decision-making process rather than a static prediction task. As these forecasting techniques continue to evolve, their integration with reinforcement learning frameworks presents a compelling research direction, potentially combining the predictive power of specialized neural architectures with the adaptive decision-making capabilities of RL agents.

Several research streams have emerged in the application of reinforcement learning to energy price forecasting. The direct forecasting approach employs RL algorithms to predict future prices by learning from historical patterns and market feedback. Specifically, Pannakkong et al. (2023) [116] developed a framework utilizing Double Deep Q-Networks (DQN) to dynamically select optimal ensembles of machine-learning models for peak electricity demand forecasting. Their approach demonstrates how reinforcement learning can integrate model selection directly into the forecasting process, thereby improving prediction accuracy while adapting to evolving demand patterns in real-time environments.

Adaptive forecasting strategies represent another promising application of RL methodologies. Specifically, Guo et al. (2020) [72] proposed an adaptive reinforcement learning framework for electricity price forecasting that dynamically selects from a portfolio of forecasting models based on prevailing market conditions. In their implementation, the state space incorporates recent price trajectories and exogenous variables, while the action space consists of candidate forecasting models. This meta-learning approach demonstrated superior performance compared to individual forecasting methodologies, particularly during regime transitions in market behavior. The application of reinforcement learning to feature selection and engineering has also yielded promising results. Specifically, Moti (2022) [115] introduced a novel framework employing Q-learning within blockchain-based smart grid environments for electricity price prediction. Their approach implements a Stackelberg game-theoretic framework mediating interactions between grid operators and consumers to optimize pricing mechanisms in real time, illustrating the potential of RL methodologies in dynamic, multi-agent environments.

Ensemble methods leveraging reinforcement learning have demonstrated particular efficacy in energy price forecasting. In particular, Jiang and Powell (2018) [85] developed an ensemble approach that combines multiple forecasting models with an RL algorithm determining optimal model weights based on recent performance metrics and prevailing market conditions. This methodology exhibited a remarkable capacity to adapt to regime changes in energy markets, consistently outperforming static ensembles and individual forecasting models during periods of market transition.

Traditional forecasting methodologies in energy markets encompass time series models (ARIMA, GARCH), fundamental models based on supply-demand balances, grey system theory models, and conventional machine-learning approaches, including neural networks and support vector machines [3]. The grey system theory, introduced by Deng (1982) [43], offers prediction methods particularly suited for systems with limited and uncertain information—characteristics often present in energy markets. Grey prediction models, notably GM (1, 1) and its variants, have demonstrated effectiveness in electricity price forecasting by requiring minimal historical data while maintaining reasonable accuracy [119]. These methods are especially valuable when dealing with non-stationary series and limited samples, complementing traditional statistical approaches. Reinforcement learning offers a distinctive paradigm for addressing forecasting challenges by reformulating prediction as a sequential decision-making problem rather than a static estimation task.

## 4.2. Optimal Trading Strategies

The implementation of effective trading strategies in energy markets requires addressing unique challenges, including high volatility, complex seasonality patterns, and distinctive price dynamics. Reinforcement learning methodologies have emerged as promising approaches for developing adaptive trading strategies capable of navigating these market complexities.

In day-ahead electricity markets, Du et al. (2021) [46] proposed a multi-agent deep reinforcement learning framework for optimizing bidding strategies. Their approach approximated Nash equilibrium solutions whereby market participants, represented as autonomous agents, learned optimal bidding policies through iterative interaction within simulated market environments. This methodology demonstrated particular effectiveness in capturing strategic behaviors in oligopolistic market structures where participants must consider competitors' potential actions.

Concerning shorter-term trading horizons, Boukas et al. (2020) [114] developed a reinforcement learning framework, specifically addressing intraday electricity trading challenges. Their implementation utilized the Proximal Policy Optimization algorithm with a comprehensive state representation incorporating recent price trajectories, order book information, and temporal features. Empirical evaluation demonstrated the framework's capacity to exploit intraday price patterns, particularly short-term price reversals and momentum effects. The authors noted that their approach exhibited superior adaptability to changing market conditions compared to traditional rule-based trading strategies.

More recently, Seyed Soroush Karimi Madahi et al. (2024) [118] advanced the application of reinforcement learning to energy storage arbitrage by implementing a distributional reinforcement learning framework. Their approach focused specifically on optimizing battery operation for profit maximization within imbalance settlement mechanisms in electricity markets. The authors' key contribution lies in their modeling of complete return distributions rather than merely expected values, enabling enhanced decision-making under uncertainty. This distributional perspective proved particularly valuable in capturing the asymmetric risk profiles characteristic of imbalance markets, where price volatility exhibits pronounced skewness and kurtosis [67].

These studies collectively demonstrate the evolution of reinforcement learning applications in energy trading, from strategic bidding in structured markets to tactical exploitation of short-term price dynamics and sophisticated risk management in storage optimization.

## 4.3. Risk Management Applications

The application of reinforcement learning to risk management in energy trading has garnered increasing attention, particularly in the domains of hedging, portfolio optimization, and risk measurement. Several recent studies have demonstrated the efficacy of RL-based approaches in addressing the unique challenges of energy markets.

Value at Risk (VaR) represents a fundamental cornerstone of financial risk management in energy markets, quantifying the maximum potential loss over a specified time horizon at a particular confidence level [88]. Traditional VaR methodologies in energy finance include historical simulation, parametric approaches, and Monte Carlo simulation, each with distinct advantages and limitations in capturing the complex dynamics of energy price behavior [125]. In this context, Halkos and Tsirivis (2019) [73] demonstrated the importance of employing advanced GARCH-type models for VaR estimation in energy portfolios, highlighting how these approaches can better quantify capital at risk due to the distinctive volatility patterns of energy commodities. The application of neural networks to VaR estimation has shown particular promise, with Wang, Liu, and Yao (2024) [126] developing an explainable quantile regression neural network (QRNN) method for VaR forecasting in energy markets that addresses both accuracy and interpretability concerns [127].

Conditional Value at Risk (CVaR), which measures the expected loss exceeding VaR, has found significant applications in power system reliability assessments. In particular, Zhang et al. (2023) [128] developed a CVaR-based reserve assessment model for power systems that explicitly accounts for primary energy supply risks, demonstrating how advanced risk metrics can enhance reliability in systems with high renewable penetration. These developments in VaR and CVaR methodologies provide important foundations for reinforcement learning applications in energy risk management, as RL frameworks can potentially learn to dynamically adjust risk measures based on evolving market conditions.

Specifically, Cao et al. (2023) [30] developed a deep distributional reinforcement learning framework for options portfolio hedging that has direct applicability to energy derivatives. Their methodology extends beyond conventional delta hedging by incorporating gamma and vega exposure management through quantile regression techniques. This approach enables more comprehensive risk-aware decision-making by modeling the entire distribution of potential outcomes rather than merely expected values. Empirical evaluations demonstrated that their framework consistently outperformed benchmark strategies in managing non-linear risks across diverse market scenarios, suggesting significant potential for application in the complex derivatives structures common in energy markets.

The challenge of basis risk—a persistent issue in energy hedging due to imperfect correlation between physical exposures and available hedging instruments—was addressed by Mulliez (2021) [117] through an innovative Q-learning framework. This study demonstrated how reinforcement learning methodologies can dynamically adapt hedging strategies in environments characterized by structural pricing mismatches. Comparative analysis revealed superior performance relative to traditional analytical hedging approaches, partic-

ularly in scenarios where the underlying and hedging instruments exhibit time-varying correlation structures—a common occurrence in regional energy markets with transmission constraints.

Particularly, Chen et al. (2020) [34] proposed a hybrid methodology combining reinforcement learning with supervised learning techniques specifically tailored to the dynamic nature of energy portfolios. Their cross-learning framework effectively captured temporal patterns in energy price dynamics while supporting adaptive strategy generation in response to changing market conditions. Performance evaluations indicated improved profit–risk trade-offs compared to both static hedging protocols and purely statistical methodologies, highlighting the advantages of RL's sequential decision-making paradigm in volatile energy markets.

Specifically, Trabelsi et al. (2025) [129] investigated tail risk transmission between crude oil and clean energy stock indices using a Time-Varying Parameter Vector Auto-Regressive model integrated with a Conditional Autoregressive Value-at-Risk approach. Their findings highlighted how crises like the COVID-19 pandemic intensified volatility spillovers between these markets, providing crucial insights for risk management strategies. These findings underscore the potential value of reinforcement learning approaches that can adapt to such regime shifts in market behavior, learning optimal risk management policies across different market states.

In examining broader portfolio construction considerations, Barrera–Rivera and Valencia–Herrera (2022) [130] developed an integrated framework combining machinelearning techniques with conditional risk measures for energy asset portfolios. Their research explored the construction of efficient frontiers under dynamic conditions and leveraged scenario simulations to enhance decision robustness. The implementation of machine-learning models provided particular advantages in forecasting non-linear dependencies characteristic of energy price behavior, enabling more effective hedging strategy formulation under uncertainty.

The integration of reinforcement learning with VaR and CVaR methodologies presents a promising frontier for energy risk management. The ability of RL algorithms to learn complex, non-linear relationships and adapt to changing market conditions aligns naturally with the challenges of risk quantification in volatile energy markets. As demonstrated by applications in related fields [69,131], the combination of advanced risk metrics with learning-based approaches offers potential to enhance both the accuracy and adaptability of risk management strategies in energy markets.

Collectively, these studies demonstrate reinforcement learning's significant potential in addressing the complex risk management challenges endemic to energy trading environments, where traditional models often prove inadequate due to market incompleteness, basis risk, and extreme price dynamics.

## 5. RL for Derivatives Valuation in Energy Markets

## 5.1. Option Pricing Fundamentals

Derivatives valuation in energy markets presents distinctive challenges that differentiate these instruments from their counterparts in conventional financial markets. These challenges arise from several fundamental characteristics of energy commodities and their corresponding markets.

The non-storability of certain energy commodities, particularly electricity, represents a significant departure from traditional financial assets. Unlike securities or even physical commodities such as precious metals, electricity cannot be economically stored in substantial quantities. This characteristic violates a fundamental assumption underpinning traditional arbitrage-based pricing models—the ability to construct replicating portfolios through buying and holding the underlying asset [44].

Energy price processes exhibit complex stochastic behavior that extends beyond the relatively simple dynamics assumed in standard financial models. These processes frequently incorporate features such as mean-reversion, reflecting the tendency of prices to return to production cost levels; discontinuous jumps, capturing sudden supply or demand shocks; and regime-switching, representing distinct market states with different underlying dynamics. These complex behaviors necessitate sophisticated stochastic modeling approaches that extend well beyond conventional geometric Brownian motion assumptions [21].

Market incompleteness presents another substantial challenge for derivatives valuation in energy markets. The inability to construct perfect replicating portfolios—due to non-storability, limited market depth, or the absence of liquid trading in certain risk factors—undermines the theoretical foundation of risk-neutral valuation. This incompleteness introduces an unavoidable element of subjectivity in pricing, as different market participants may assign different values to non-hedgeable risks [22].

Many energy derivatives incorporate embedded optionality regarding delivery specifications, further complicating their valuation. These contingent features may include flexibility in delivery location, timing, or quantity. For instance, natural gas swing contracts permit buyers to vary daily consumption within specified limits, while power transmission rights grant optionality regarding the utilization of transmission capacity [1].

Despite these challenges, several methodological approaches have been developed to address energy derivatives pricing. Extended Black–Scholes frameworks modify the standard geometric Brownian motion assumption to incorporate mean-reversion, jumps, and other features characteristic of energy price dynamics [75,98]. Monte Carlo simulation techniques provide numerical solutions through the simulation of price paths based on specified stochastic processes, particularly valuable for path-dependent derivatives or contracts with complex exercise features [14]. Partial differential equation methods offer numerical solutions to the governing equations of derivative prices under specific assumptions about the underlying price process, though their application becomes increasingly challenging as the dimensionality of the problem increases [132].

Following our examination of option pricing fundamentals, Table 2 classifies significant research applying reinforcement learning methodologies to energy derivatives valuation. This classification organizes studies by valuation problem type, RL methodology, energy market focus, and key contributions. The table illustrates how different RL approaches address specific challenges in energy derivatives pricing and real options analysis, providing a framework for understanding the comparative advantages of these methods over traditional approaches.

**Table 2.** Summary of reinforcement learning applications in energy derivatives valuation, highlighting methodologies, energy domains, and key contributions.

Study	Valuation Problem	RL Methodology	<b>Energy Focus</b>	Key Contribution
Halperin (2019) [74]	General option pricing	Q-learning	Energy options	Model-free approach deriving pricing functions directly from empirical data
Buehler et al. (2019) [133]	Hedging under market frictions	Deep reinforcement learning	Options hedging	Framework accommodating transaction costs and market incompleteness
Becker et al. (2019) [134]	Optimal stopping	Deep Q-Network	American-style options	Direct learning of exercise policies without explicit continuation values

Study	Valuation Problem	RL Methodology	Energy Focus	Key Contribution
Becker et al. (2020) [135]	Gas swing option valuation	Deep reinforcement learning	Natural gas contracts	RL approach superior to LSMC for contracts with complex constraints
Marzban et al. (2023) [18]	Risk-aware option pricing	Actor-critic with risk measures	Energy derivatives	Incorporation of expectile risk measures for risk-averse valuation
Song (2022) [136]	Computationally efficient pricing	Deep RL with high-performance computing	Energy option pricing	Real-time pricing under dynamic market conditions
Carbonneau (2021) [29]	Equal risk pricing	Neural networks with RL	Energy derivatives	Pricing framework reflecting actual hedging costs and residual risks
Dalal et al. (2016) [41]	Generation asset valuation	Deep Deterministic Policy Gradient (DDPG)	Power generation	Operating policies maximizing value under technical constraints
Boogert and de Jong (2008) [137]	Gas storage valuation	Q-learning	Natural gas storage	Capturing complex intertemporal tradeoffs in storage operations
Lee et al. (2023) [10]	CCU investment valuation	RL with real options	Carbon capture	Framework for identifying optimal investment timing under uncertainty
Caputo and Cardin (2022) [31]	Waste-to-energy system valuation	Deep RL for flexibility analysis	Energy systems	DRL models improved economic outcomes by up to 69% vs. traditional approaches
Cheraghi et al. (2024) [38]	Energy transition investment	RL for sustainable planning	Renewable energy	Dynamic optimization considering environmental and regulatory uncertainty

#### Table 2. Cont.

## 5.2. RL Approaches to Option Pricing and Applications

Reinforcement learning methodologies offer promising alternatives to traditional derivatives valuation techniques, potentially addressing several limitations of conventional approaches. These methods have demonstrated particular utility in the context of energy derivatives, where complex market dynamics and incompleteness present significant modeling challenges.

Particularly, Halperin (2019) [74] developed a model-free approach to option pricing using reinforcement learning frameworks. This methodology enables an agent to learn pricing functions by directly interacting with a simulated market environment, circumventing the need for explicit specification of the underlying price process. The approach derives pricing functions directly from empirical data, thereby avoiding potentially restrictive parametric assumptions. When applied to energy options, this technique demonstrated notable efficacy, particularly for instruments with complex features that resist conventional parametric modeling. The flexibility of this approach proves especially valuable in energy markets, where price dynamics exhibit distinctive characteristics including extreme volatility, mean-reversion, and regime-switching behavior.

The deep hedging framework proposed by Buehler et al. (2019) [133] represents another significant application of reinforcement learning to derivatives pricing. This approach employs deep reinforcement learning to determine optimal hedging strategies that minimize hedging error under realistic market conditions. When extended to energy derivatives, this methodology naturally accommodates market frictions, including transaction costs, liquidity constraints, and market incompleteness—features that traditional risk-neutral pricing approaches struggle to incorporate. By optimizing hedging decisions directly, rather than deriving them from theoretical price processes, deep hedging provides more robust risk management strategies for complex energy derivatives. Specifically, Becker et al. (2019) [134] applied deep reinforcement learning techniques to optimal stopping problems, which have particular relevance for American-style options and swing options commonly traded in energy markets. Their methodology employed a Deep Q-Network (DQN) architecture to learn optimal exercise policies directly, eliminating the need for explicit modeling of continuation values that traditional approaches require. When applied to gas storage contracts with complex exercise constraints, their reinforcement learning approach demonstrated superior performance compared to conventional least-squares Monte Carlo methods. This performance advantage stems from the ability of reinforcement learning algorithms to capture complex exercise boundaries without restrictive functional form assumptions.

The valuation of energy derivatives presents unique challenges that traditional approaches struggle to address effectively. Electricity's non-storability, complex price dynamics, and the presence of operational constraints create an incomplete market where perfect hedging is impossible. RL offers a model-free framework that can learn optimal pricing and hedging strategies directly from market data without requiring explicit specification of the underlying stochastic processes.

Particularly, Marzban et al. (2023) [18] extend deterministic actor–critic reinforcement learning to incorporate time-consistent recursive expectile risk measures, addressing the risk-averse nature of energy market participants. Their approach accommodates complex hedging problems even when only historical asset data are available, generating nearly optimal hedging policies for energy derivatives without requiring full knowledge of asset dynamics. This is particularly valuable in electricity markets where price processes exhibit regime-switching behavior and extreme events that parametric models struggle to capture.

Also, Song (2022) [136] addresses the computational intensity of option pricing in energy markets by integrating high-performance computing with deep reinforcement learning. This approach enables real-time pricing of complex energy derivatives under dynamic market conditions, incorporating challenges like random interest rates and transaction costs that are prevalent in energy markets. The shift from analytical models to data-driven, computation-heavy frameworks aligns well with the realities of modern energy derivatives that often lack closed-form solutions.

The concept of equal risk pricing, explored in depth by Carbonneau (2021) [29], offers a promising framework for energy derivatives where market incompleteness makes traditional risk-neutral pricing problematic. By modeling hedging strategies as neural networks trained via deep reinforcement learning, this approach can generate fair prices for energy derivatives that reflect the actual hedging costs and residual risks faced by market participants. The ability to incorporate multiple hedging instruments is particularly relevant for energy markets where cross-commodity hedging is common practice.

The distinctive characteristics of energy markets have allowed for the rise of specialized option structures that present unique valuation challenges, prompting researchers to explore reinforcement learning approaches for these instruments.

Swing options represent a significant application domain in energy markets, particularly in natural gas and electricity contracts. These instruments afford holders the flexibility to determine both the timing and volume of delivery, subject to cumulative constraints over the contract period. This optionality is particularly valuable due to the volatile nature of energy prices and varying demand patterns. Also, Meinshausen and Hambly (2004) [19] pioneered the application of reinforcement learning to this valuation problem, implementing Q-learning algorithms to determine optimal exercise policies that respect both local and global constraints. Building upon this foundation, Becker et al. (2020) [135] employed deep reinforcement learning methodologies to value swing options in natural gas markets, demonstrating performance advantages over traditional least-squares Monte Carlo approaches. Their research revealed particularly significant improvements for contracts with complex constraint structures and during periods of high market volatility, where the adaptive nature of reinforcement learning offers distinct advantages.

Spread options, particularly spark spreads (electricity price minus gas price multiplied by heat rate) and dark spreads (electricity price minus coal price multiplied by heat rate), constitute fundamental instruments for managing generation asset exposure in energy portfolios. These instruments derive their value from the margin between output and input commodity prices, incorporating the efficiency factor of the conversion process. In addition, Carmona and Coulon (2014) [21] elucidated the challenges inherent in valuing these instruments using traditional methods, noting, particularly, the complex correlation structures and regime-switching behaviors that characterize the relationship between fuel and electricity prices. The application of reinforcement learning to spread option valuation remains an evolving research domain, with recent studies suggesting promising results in capturing the non-linear dependencies between underlying commodities. The multifactor nature of these options presents both challenges and opportunities for reinforcement learning approaches, as the high-dimensional state space benefits from the RL's capacity to learn complex value functions.

Locational spread options emerge from transmission constraints that create price differentials between geographic regions within interconnected energy networks. These instruments derive value from arbitraging price disparities between locations when transmission capacity permits. Particularly, Oren (2001) [138] articulated the difficulties in applying conventional valuation techniques to these instruments, highlighting the impact of network topology and congestion patterns on option values. Recent applications of reinforcement learning to this domain have demonstrated advantages in incorporating the complex network constraints and contingency scenarios that influence locational spread values. The stochastic nature of both congestion patterns and regional supply–demand dynamics creates a particularly suitable application for reinforcement learning methodologies, which can adapt to changing network conditions without requiring explicit modeling of all contingencies. Furthermore, the integration of increasing volumes of location-dependent renewable generation has enhanced the importance and complexity of these instruments, creating additional incentives for advanced valuation methodologies.

These specialized energy options illustrate the diversity of challenges in energy derivatives valuation and highlight the potential advantages of reinforcement learning approaches over traditional techniques, particularly when handling complex constraints, high-dimensional state spaces, and regime-dependent dynamics that characterize modern energy markets.

## 5.3. Real Options Analysis

Real options analysis provides a framework for valuing operational flexibility and managerial discretion embedded in physical assets and investment decisions in energy markets. Unlike traditional discounted cash flow methods, real options approaches explicitly account for the value of flexibility under uncertainty. Reinforcement learning methodologies offer a natural computational framework for addressing these complex sequential decision problems that often resist closed-form solutions.

The valuation of power generation assets represents a prominent application domain, as these facilities can be conceptualized as real options to transform fuel inputs into electricity when economically advantageous. Specifically, Tseng and Barz (2002) [139] articulated the limitations of conventional valuation methodologies in this context, particularly their inability to adequately incorporate technical constraints and operational characteristics. Advancing this research direction, Dalal et al. (2016) [41] implemented deep reinforcement learning techniques for generation asset valuation, developing operating policies that maximized economic value while respecting technical constraints. Their methodology employed a Deep Deterministic Policy Gradient algorithm with a multidimensional state space encompassing fuel prices, electricity prices, and plant operational status variables. This approach demonstrated superior performance compared to traditional methods, particularly in capturing the value of operational flexibility under complex market conditions.

Energy storage facilities, including natural gas storage, pumped hydroelectric systems, and battery installations, represent sophisticated real options with multi-dimensional constraints. Specifically, Boogert and de Jong (2008) [137] applied classical reinforcement learning techniques to natural gas storage valuation, demonstrating how these methods could capture the complex intertemporal trade-offs inherent in storage operation. Subsequent research has extended these approaches to incorporate additional constraints and market characteristics.

Investment timing decisions in energy infrastructure development involve embedded timing options that significantly impact project valuation. Also, Chronopoulos et al. (2016) [39] examined these timing options specifically within the context of renewable energy investments, highlighting how policy uncertainty affects optimal investment strategies. Reinforcement learning approaches offer particular advantages in this domain due to their ability to incorporate multiple uncertainties simultaneously, including regulatory changes, technological learning curves, and market condition evolution.

Recent literature has significantly advanced the integration of reinforcement learning with real options analysis in energy markets. Specifically, Nadarajah and Secomandi (2023) [140] provide a comprehensive review of real options applications across various energy domains, including electricity, natural gas, and crude oil, highlighting the evolving role of machine-learning techniques in capturing value under uncertainty. Their survey establishes a foundation for understanding how computational approaches are transforming valuation methodologies in energy finance, building upon the earlier works while identifying emerging research directions.

The application of reinforcement learning to environmentally significant energy technologies has emerged as a particularly active research area. Specifically, Lee et al. (2023) [10] developed a framework merging real options theory with reinforcement learning to evaluate the commercial viability of carbon capture and utilization (CCU) technologies. Their approach models uncertain factors such as market prices and policy shifts, identifying optimal investment timing and highlighting the value of flexibility in energy project deployment under deep uncertainty. Similarly, Alqubaisi (2023) [141] applied deep-learning methods to value real options in renewable energy projects, developing a framework that captures non-linear dependencies and offers a scalable approach to handle complex valuation under stochastic conditions.

Methodological advancements in applying reinforcement learning to real options have also progressed significantly. Furthermore, Caputo and Cardin (2022) [31] proposed a deep reinforcement learning approach to assess flexibility in engineering systems, tested on a waste-to-energy system. By comparing DRL models with traditional decision rule approaches, they demonstrated that DRL-enhanced models improved economic outcomes by up to 69%, generalizing across scenarios and supporting better-informed strategic design under uncertainty. Moreover, Lawryshyn (2023) [9] further emphasized the limitations of classical option pricing models in capturing multidimensional uncertainties and proposed RL-based solutions for adaptive decision-making, providing valuable insights into training RL agents for investment strategies with embedded flexibility.

Looking toward future energy systems, Cheraghi et al. (2024) [38] explored the use of reinforcement learning in planning and investment for sustainable energy transitions. Their

work introduces RL algorithms that dynamically optimize energy systems considering environmental and regulatory uncertainty, positioning RL as a crucial tool to unlock value creation in decarbonized energy strategies. This research direction highlights how reinforcement learning can address the complex, multi-objective decision problems inherent in

The application of reinforcement learning to real options valuation in energy markets remains an active research frontier with substantial opportunities for methodological advancement. Future research directions include the development of hybrid approaches that combine the interpretability of traditional real options models with the flexibility of reinforcement learning, and the extension of these methods to address multi-agent decision environments that better reflect competitive market dynamics.

# 6. Option Value in Power Systems

energy transition investments.

#### 6.1. Real Options Framework for Smart Grid Technologies

The concept of option value has found significant applications in power systems, particularly in economically evaluating smart grid investments under uncertainty [142]. In this context, the option value of deploying a smart grid technology represents the difference in total expected system costs between cases with and without the technology deployment [81]. This approach recognizes that smart grid investments create flexibility to deal with future sources of uncertainty, which has economic value beyond traditional deterministic cost–benefit analysis [143].

Several smart grid technologies have demonstrated significant option value when captured using a stochastic optimization framework [25]. Dynamic Line Rating (DLR) captures the variable transmission capacity based on real-time conditions rather than static conservative ratings, generating option value from deferred transmission investments and improved congestion management [144]. In this context, Giannelos et al. (2018) [63] quantified the option value of DLR in transmission systems, showing that DLR can lead to expensive network reinforcements being deferred or displaced.

Energy storage systems provide multiple services, including arbitrage, capacity deferral, ancillary services, and renewable integration. Their option value stems from the ability to charge/discharge load in response to price signals, system needs, and renewables output—all subject to significant uncertainty [145]. Also, Giannelos et al. (2020) [60] developed a methodology to quantify the contribution of energy storage to the security of supply through the F-Factor approach, capturing additional option value beyond traditional energy arbitrage applications.

Demand-Side-Response (DSR) programs create option value by providing network operators with flexibility to modify load profiles in response to supply constraints or price signals. This flexibility is particularly valuable during extreme events or unexpected system conditions, as shown by Papadaskalopoulos and Strbac, 2017 [146]. Also, Giannelos et al., 2018 [26] and Giannelos et al., 2018 [64]) analyzed the option value of the demand-side response under decision-dependent uncertainty, demonstrating the significant option value of DSR technology under endogenous uncertainty, using stochastic optimization.

Advanced Network Management Systems that enable greater observability and controllability of distribution networks create option value by allowing network operators to defer traditional reinforcement costs while maintaining reliability under uncertain load growth and distributed generation adoption [47]. In their comprehensive framework, Giannelos, Borozan, Konstantelos et al. (2024) [57] quantified the option value, investment costs, and deployment levels for various smart grid technologies, providing a systematic approach to evaluating these investments. Electric Vehicle Smart Charging creates significant option value by coordinating charging patterns to support grid requirements [147]. Specifically, Giannelos, S., Borozan, S., and Strbac, G. (2022) [55] developed a backwards induction framework that quantifies both the option value of smart charging and the risk of stranded assets under uncertainty. Their analysis demonstrated how smart charging strategies can significantly reduce network investment costs while accommodating growing EV penetration. Additionally, in two studies, Borozan, Giannelos, and Strbac [148,149]) integrated EV smart charging as an investment option within strategic network expansion planning, demonstrating substantial option value through deferred network reinforcements. Similarly, Giannelos, S., Borozan, S., Strbac, G. et al. (2024) [58] focused on vehicle-to-grid applications, quantifying their contribution to security of supply using the F-Factor methodology, capturing additional option value beyond traditional V2G applications. Moreover, the option value of smart charging portfolios is presented in Giannelos, Borozan, Moreira et al. (2023) [59].

Soft Open Points, power electronic devices that enable flexible network reconfiguration in distribution systems, have also demonstrated significant option value [52,86]. The paper by Lu et al. (2019) [16] proposes a multi-layer planning model for deploying Soft Open Points in active distribution networks, integrating a demand response to enhance grid flexibility and reduce operational costs. SOPs are used to control power flow and mitigate issues arising from high penetration of distributed generation, while demand response is modeled using time-of-use pricing to shift loads cost-effectively. The model, solved using an improved particle swarm optimization algorithm, shows through IEEE-33 node simulations that this combined planning approach improves both economic performance and operational feasibility.

A similar analysis is presented in Giannelos, Konstantelos et al. (2019), where soft open point technology is combined with energy storage [150].

#### 6.2. Stochastic Optimization Approach to Quantifying Option Value

The quantification of option value for smart grid technologies typically employs stochastic optimization methods in a two-stage process. In the first stage, a stochastic optimization problem is solved without the smart grid technology under future uncertainty. The objective function typically minimizes total system costs, including operational costs and investment costs. The result is the expected system cost without the technology. In the second stage, the same stochastic optimization problem is solved with the smart grid technology considered for deployment in the system [71]. This yields the expected system cost with the technology. A list of smart technologies and their formulations is presented in Giannelos, Borozan, and Aunedi (2023) [56].

The difference between these expected costs represents the option value of the technology. A positive difference indicates that the technology creates value by providing flexibility to deal with uncertainty. Key uncertainties typically modeled in such stochastic optimization models include renewable generation output, load growth and profiles, fuel prices, equipment outages, policy and regulatory changes, and technology costs. These uncertainties represent the primary factors that affect the valuation of smart grid technologies and must be adequately characterized to produce meaningful option value estimates.

Mathematically, this can be formulated as:

$$OV = E[C_{without}] - E[C_{with}]$$

where *OV* is the option value,  $E[C_{without}]$  is the expected cost without the technology, and  $E[C_{with}]$  is the expected cost with the technology, each quantified via stochastic optimization.

Specifically, Giannelos, Konstantelos et al. (2017) [62] developed a new class of planning models for option valuation of storage technologies under decision-dependent innovation uncertainty. Their approach incorporated how deployment decisions affect future technology costs and capabilities, showing that traditional models undervalue early adoption of promising technologies.

In particular, Zhao et al. (2015) [121] applied this approach to evaluate energy storage systems in a system with high renewable penetration. Their analysis showed that traditional deterministic approaches significantly undervalued storage by failing to capture its ability to mitigate the uncertainty of renewable generation.

Also, Giannelos, Jain, and Borozan (2021) [61] apply stochastic optimization to longterm expansion planning of the transmission network in India under multi-dimensional uncertainty. Their framework captures complex interactions between different sources of uncertainty, demonstrating how flexible investment strategies create substantial option value in rapidly evolving power systems. Building on this work, Giannelos, Zhang et al. (2024) [66] developed methods for Pareto frontier sensitivity analysis in power systems, enabling decision-makers to understand how economic value changes across different configurations.

Moreover, Dong, Zhang, et al. (2024) [45] examined how coordinated control of space heating across multiple buildings can enhance urban energy system flexibility. Using optimization models that incorporate building thermal dynamics and occupant comfort constraints, they demonstrated substantial option value from heating flexibility.

#### 6.3. Methodological Considerations for RL-Based Option Valuation

Applying RL to quantify the option value of smart grid technologies involves several methodological considerations. To accurately determine option value, the baseline (without technology) scenario must use the same RL algorithm and reward structure, differing only in the availability of the technology. This ensures a fair comparison of expected costs. The scenarios used for training and evaluation should properly represent the underlying uncertainties. Techniques like importance sampling, scenario reduction, and generative models can help create representative scenario sets without excessive computational burden.

The reward function should accurately reflect the total system costs, including operation and investment costs. Improperly specified reward functions can lead to biased option value estimates. Smart grid investments typically have multi-decade lifespans. RL approaches must properly account for long-term effects, either through appropriate discount factors or explicit modeling of long-term scenarios. RL-based valuation typically requires significant computational resources for training.

Techniques like transfer learning, model-based RL, and distributed computing can help address this challenge, similar to decomposition methodologies applied to stochastic optimization frameworks as in Borozan, Giannelos et al. (2024) [151], which present a machine-learning-enhanced Benders decomposition approach for multi-stage stochastic transmission expansion planning that significantly improves computational efficiency while capturing the full option value of smart grid investments [35].

These models could support more accurate scenario generation for RL-based option valuation. Similarly, Giannelos, Zhang, Pudjianto et al. (2024) [65] compared strategic versus incremental planning approaches in electricity distribution grids, providing insights into how different planning horizons affect option values—a key consideration when designing RL reward functions for long-term investment decisions.

#### 6.4. Future Research Directions

Several promising research directions could further enhance RL-based option valuation in power systems. Most studies focus on individual technologies, but real power systems deploy portfolios of complementary smart grid technologies. Research on RL-based valuation of technology portfolios could identify synergistic combinations (Charousset-Brignol et al., 2021) [33] that create greater option value than the sum of individual technologies. Rather than learning expected values, distributional RL learns the full distribution of returns. This approach could provide more comprehensive information about the option value distribution under extreme events and rare scenarios. Different stakeholders have different risk preferences when evaluating smart grid investments. Research on risk-sensitive RL could provide option valuations tailored to specific risk preferences, such as risk-averse utility regulators. Smart grid deployments can influence market prices and regulatory decisions, creating feedback effects not captured in current models. Research on RL approaches that model these feedback effects could provide more accurate option valuations. Regulatory approval of smart grid investments typically requires transparent justification. Research on explainable RL could make option valuation results more interpretable and trustworthy for regulators and other stakeholders.

# 7. Conclusions: Limitations, Future Directions, and Policy Limitations

This review has examined how reinforcement learning can be applied to energy finance, highlighting applications ranging from price forecasting and trading strategies to derivatives valuation and option value assessment in power systems. As this field continues to evolve, it is important to recognize current limitations, identify promising research directions, and consider policy implications.

## 7.1. Current Limitations of RL in Energy Finance

Despite demonstrating significant potential, the application of reinforcement learning methodologies to energy finance encounters several substantial challenges that merit careful consideration.

Interpretability remains a primary concern in the deployment of RL techniques within financial contexts. Deep reinforcement learning models frequently operate as "black boxes", with decision-making processes that resist straightforward human interpretation. This opacity presents significant impediments for risk management protocols and regulatory compliance frameworks that typically require transparent justification of trading strategies and investment decisions. While recent research has explored techniques such as attention mechanisms and feature importance analysis to enhance model interpretability [152], these approaches have yielded only incremental improvements. The development of inherently interpretable RL architectures that maintain competitive performance represents a critical avenue for future research, particularly as financial regulators increasingly scrutinize algorithmic decision-making systems.

Sample complexity constitutes another significant limitation. RL algorithms characteristically require extensive data for effective policy learning, a requirement that proves problematic in energy finance applications where historical data may be insufficient, particularly for rare events or emerging market structures. This limitation becomes especially pronounced when modeling extreme price events or evaluating strategies under novel regulatory frameworks. Current approaches addressing this constraint include modelbased reinforcement learning, which leverages environment models to reduce data requirements; transfer-learning techniques that apply knowledge from related domains; and synthetic data augmentation [153]. However, the effectiveness of these methods remains constrained when the target domain exhibits substantial structural differences from available training data. The integration of domain knowledge and physics-informed constraints into RL frameworks offers a promising direction for improving sample efficiency in energy applications. The challenge of generalization across distinct market regimes is particularly salient in energy finance. Reinforcement learning agents trained under specific market conditions frequently struggle to maintain performance when confronted with regulatory changes, technological disruptions, or structural market shifts. This limitation is particularly relevant in energy markets, which regularly experience significant policy interventions and infrastructure evolution. While meta-learning approaches have demonstrated promise for adapting to changing environments [13], these techniques remain in the nascent stages of development for financial applications. Robust evaluation methodologies that specifically assess RL algorithm performance across regime changes could provide valuable insights for practitioners implementing these systems in dynamic energy markets.

Computational requirements present practical implementation barriers, particularly for smaller market participants. Contemporary deep reinforcement learning methods typically demand substantial computational resources during both training and, in some cases, inference phases. This resource intensity creates asymmetric advantages for larger institutions with greater technological capabilities. Although algorithmic improvements and optimization techniques have somewhat mitigated these requirements, computational efficiency remains a significant concern [154]. The development of more efficient algorithmic formulations and hardware-specific optimizations could democratize access to advanced RL techniques across a broader spectrum of energy market participants. Additionally, federated learning approaches may offer pathways to collaborative model development while maintaining data privacy, potentially addressing both computational and data scarcity challenges simultaneously.

Finally, the field suffers from benchmarking difficulties that impede systematic evaluation and comparison of different methodologies. The absence of standardized benchmarks and evaluation protocols makes objective assessment of competing RL approaches, complicating both research progress and practical implementation decisions. The development of common benchmarks specifically designed for energy finance applications, incorporating realistic market constraints and evaluation metrics aligned with practitioner objectives, represents an important direction for future research [155]. These benchmarks should ideally capture the multifaceted nature of energy markets, including physical constraints, regulatory frameworks, and the multiple time scales characteristic of energy price dynamics discussed in earlier sections.

#### 7.2. Promising Research Directions

The intersection of reinforcement learning and energy finance presents fertile ground for innovative research. Below, we identify key research directions organized by methodological advancements, application domains, and comparative studies that warrant further exploration in this rapidly evolving field.

#### 7.2.1. Methodological Advancements for Energy Finance

Explainable RL for Energy Investment Decisions: Developing interpretable reinforcement learning models represents a critical research priority. For smart grid investments, transparency in decision processes, in particular, is essential for regulatory approval and stakeholder acceptance. Future research should focus on methods that balance performance with comprehensibility, perhaps through attention mechanisms that highlight which factors most influence investment timing and operational decisions in energy assets [79].

Multi-agent Reinforcement Learning Frameworks: As energy systems become increasingly decentralized, understanding strategic interactions becomes essential. Multi-agent reinforcement learning presents a compelling direction, particularly for optimizing distributed energy resources across smart grid ecosystems [156]. Research in this area could address how decentralized decision-making impacts financial returns across different stakeholders and potentially reveal emergent properties that centralized valuation methods might miss.

Transfer Learning and Domain Adaptation: Energy markets exhibit significant regional variations in regulations, market structures, and resource availability. Research on transferring knowledge from data-rich to data-sparse markets could enable more efficient application of RL techniques. This is particularly relevant for emerging energy technologies where historical data may be limited, but analogous applications exist in other domains.

#### 7.2.2. Enhanced Valuation of Energy Assets and Flexibility

Valuation of Operational Flexibility: Traditional real options valuation often struggles to capture the complex interdependencies between multiple flexibility options in smart grid technologies. Reinforcement learning algorithms, with their ability to learn optimal policies through interaction with dynamic environments, offer promising mechanisms to value these interconnected flexibilities more accurately. Such research could bridge the gap between theoretical option value and practical implementation challenges in smart grid investments.

Long-duration Energy Storage Valuation: Deep reinforcement learning techniques show promise for addressing the complex valuation of long-duration energy storage assets within smart grid systems. These assets present particular challenges in balancing shortterm operational decisions with long-term strategic value creation. Research applying deep reinforcement learning methods to capture these temporal dependencies could overcome limitations of traditional valuation methods that often oversimplify the strategic dimension of energy storage.

Renewable Integration Flexibility: Methods for accurately valuing flexibility in highrenewable energy systems remain underdeveloped. Research applying distributional reinforcement learning to capture the full range of outcomes under renewable uncertainty could provide more accurate valuations of flexible assets like batteries, demand response, and dispatchable generation. This research is particularly relevant as energy systems worldwide transition toward higher renewable penetration.

#### 7.2.3. Uncertainty Modeling and Risk Assessment

Regulatory Uncertainty: Energy markets face continuous regulatory evolution, creating significant uncertainty for investors. Research developing reinforcement learning algorithms that explicitly model and adapt to regulatory changes could help energy investors maintain option value in unstable policy environments. This is particularly relevant for smart grid technologies, which often rely on evolving market structures and incentive mechanisms.

Climate Risk Integration: Future research should explore integrating climate risk factors into reinforcement learning frameworks for energy investments. This includes modeling physical risks (extreme weather impacts on infrastructure) and transition risks (policy and technology shifts) within RL environments. Models that capture these complex, interacting uncertainties could significantly improve long-term energy investment decisions under climate change scenarios [56].

Cross-market Risk Dependencies: Energy markets exhibit strong interdependencies with other commodities and financial markets. Developing reinforcement learning approaches that capture these cross-market dynamics represents a promising research direction with significant implications for comprehensive risk management in energy investment portfolios.

## 7.2.4. Comparative Analysis of Decision-Making Frameworks

Comparative analysis of decision-making frameworks represents a critical research opportunity. By systematically evaluating reinforcement learning, stochastic optimization, and least-worst-regret approaches [36] for energy investment decisions, researchers can establish clearer guidelines for when each methodology excels. Smart grid technologies often involve multiple uncertainties across varying time horizons, making the selection of appropriate decision frameworks crucial. This comparative research could yield practical decision roadmaps for energy finance practitioners facing complex investment choices.

Each methodology offers distinct advantages in handling uncertainty, computational requirements, and interpretability. By advancing understanding of their relative strengths in the energy finance context, researchers can develop more robust decision-support tools for smart grid investment and operation. Standardized benchmarks and case studies would further enhance the practical utility of such comparative analyses.

The advancement of open-source implementations, standardized problem formulations, and common datasets would accelerate progress across these research directions. Addressing these complex challenges at the intersection of reinforcement learning and energy finance will require interdisciplinary collaboration between financial engineering, computer science, energy systems, and policy research communities.

## 7.2.5. Sustainable Communities and Energy Equity

The application of reinforcement learning to energy finance extends beyond technical optimization and economic efficiency to address pressing social challenges. There is growing recognition that energy systems must support sustainable communities by combating energy poverty [54,157–159] and ensuring equitable distribution of benefits [160]. This section examines how reinforcement learning methodologies can be leveraged to promote pragmatic solutions that balance technical, economic, and social dimensions of energy transitions [8].

Social Innovation in Community Energy Transitions: Energy transitions are increasingly viewed through the lens of social innovation and community participation rather than purely technological change. Alaize Dall–Orsoletta et al. (2022) [161] conducted a systematic review of how social innovations promote community-driven energy transitions, identifying major themes, including citizen participation, institutional support, and the role of cooperatives in renewable energy deployment. The authors highlighted practical examples of successful transitions facilitated through collective action, providing a foundation for understanding how reinforcement learning frameworks could be designed to support community-based decision processes [162,163].

Building on this foundation, Pillan et al. (2023) [164] proposed conceptual frameworks to help communities better understand and contribute to sustainable energy transitions, emphasizing the role of education and participatory design in fostering local energy initiatives. Their work suggests that reinforcement learning models could be developed to incorporate community preferences and knowledge, creating more robust and socially accepted energy optimization strategies. Similarly, Neij et al. (2025) [165] reviewed experiences of energy communities across Europe, identifying key success factors, including strong local engagement, supportive regulations, and diversified revenue streams—factors that could be parameterized within RL frameworks to better reflect community priorities.

Energy Poverty Assessment and Alleviation: Energy poverty—inadequate access to affordable, reliable energy services—remains a critical challenge globally. Recent advances in machine-learning applications for energy poverty have opened new avenues for addressing this issue. López–Vargas et al. (2022) [15] examined how AI methods are being applied to energy poverty contexts, noting that relatively few studies have explored AI solutions specifically for energy poverty and suggesting future directions for AI-based detection, prediction, and policy design.

More concretely, Gawusu et al. (2024) [53] used spatial data and predictive modeling to identify energy poverty hotspots and inform targeted policy measures, demonstrating how spatial analytical techniques could enhance the precision of interventions. Abbas et al. (2022) [166] employed machine learning to measure and predict extreme forms of energy poverty based on multiple socio-economic factors, identifying critical determinants such as income, education, and geographic variables. These advancements in prediction and classification provide foundations that reinforcement learning frameworks could build upon to optimize dynamic resource allocation for energy poverty alleviation programs [167].

Che et al. (2021) [37] proposed an integrated evaluation framework for global energy poverty, stressing availability and affordability of energy as primary obstacles to alleviation. Their emphasis on regional disparities as barriers for global policy coordination highlights the need for adaptive solutions that reinforcement learning is well-positioned to provide. Complementing this work, Lippert and Sareen (2023) [12] explored how transitioning to low-carbon energy infrastructures can help reduce energy poverty, using big data analytics to identify systemic changes needed in infrastructure and agency behavior. Their finding that mere technological fixes are insufficient without systemic policy shifts aligns with the need for reinforcement learning approaches that can navigate both technical and institutional complexities.

Democratized Energy Markets and Community Participation: Reinforcement learning shows particular promise in enabling more inclusive participation in energy markets. Piras et al. (2024) [168] presented an open-source AI/ML-based tool designed to facilitate the automated creation of renewable energy communities, demonstrating that AI can directly enhance social coordination in energy system development. By integrating advanced energy modeling and citizen participation frameworks, such tools support a decentralized and democratic energy transition—an application area where reinforcement learning's sequential decision-making capabilities could prove particularly valuable.

The concept of a "just energy transition" has gained prominence in policy discussions, with del Guayo and Cuesta (2022) [42] critically examining this concept within European policy frameworks. Their analysis of the Just Transition Fund highlighted its emphasis on supporting coal-dependent regions while critiquing its narrow scope. The authors argued that energy justice challenges extend beyond coal closures to issues like lithium mining, rural environmental impacts, and growing energy poverty—complex trade-offs that reinforcement learning methodologies could help navigate by incorporating multiple objectives and constraints.

Equity-Aware Reinforcement Learning Frameworks: A critical challenge in applying RL to energy finance is ensuring that optimization objectives incorporate equity considerations. Chen et al. (2024) [36] addressed this challenge directly, focusing on how bias in ML models can exacerbate existing inequities in energy systems. They proposed technical and governance frameworks to mitigate biases and promote fairness across energy distribution networks, providing a crucial blueprint for ensuring AI-driven energy systems uphold principles of energy justice. These insights could inform the development of fairness-aware reward functions in reinforcement learning models for energy systems.

Kaur et al. (2024) [169] explored how AI, particularly machine learning and data analytics, can improve the sustainability and resilience of energy systems while emphasizing stakeholder engagement, such as involving local communities in solar energy initiatives. Their argument that AI must be socially inclusive to fully realize sustainable energy transitions suggests the need for reinforcement learning frameworks that explicitly incorporate distributional impacts and fairness constraints, similar to how risk constraints are integrated into financial optimization models.

Ethical Considerations in AI-Driven Energy Systems: The ethical dimensions of AI deployment in energy systems have received increasing attention. Chauhan et al. (2024) [32] reflected on ethical concerns surrounding AI and ML deployment in clean energy systems, particularly regarding fairness and social impact. Their discussion of the tension between rapid technological advancement and ensuring equitable outcomes highlights the need for careful design of reinforcement learning objectives and constraints in energy applications. Similarly, Jain and Mitra (2025) [82] advocated for human-centered AI systems that prioritize marginalized groups when supporting sustainable development goals, including energy access.

Nalli et al. (2025) [170] proposed frameworks for energy equity through intelligent system design, highlighting AI's role in enabling inclusive energy transitions at the community level. Their work on optimizing energy systems while ensuring equitable access to affordable power provides conceptual groundwork for reinforcement learning applications that balance efficiency with equity considerations.

Research Directions and Implementation Challenges: Despite growing interest in integrating social considerations into energy system optimization, implementing equity-aware reinforcement learning faces several challenges: defining appropriate fairness metrics, obtaining representative data across diverse communities, and balancing potentially competing objectives of efficiency and equity.

Alturif et al. (2024) [171] discussed using AI tools for poverty prediction and strategic alleviation, reviewing various machine-learning models and their policy applications. While focused broadly on poverty, their work highlights the transformative potential of AI in identifying at-risk populations—a capability that could be enhanced through reinforcement learning's ability to optimize interventions across time.

Future research should focus on developing reinforcement learning frameworks that explicitly incorporate community values and preferences, equity metrics, and distributional impacts. Multi-objective reinforcement learning approaches that simultaneously optimize for technical efficiency, economic viability, and social equity represent a particularly promising direction. As energy systems continue to evolve toward greater decentralization and complexity, reinforcement learning approaches that can navigate these multidimensional trade-offs will become increasingly valuable for creating truly sustainable energy futures.

## 7.3. Policy Recommendations

Based on the findings of this review, several policy recommendations emerge for regulatory bodies, market operators, and industry participants seeking to harness the potential of reinforcement learning in energy finance.

Regulatory Framework for Algorithmic Trading: As reinforcement learning adoption increases in energy markets, regulatory bodies should develop frameworks specifically addressing algorithmic trading that balance innovation with market stability. These frameworks should include disclosure requirements for trading entities using RL systems, stress testing protocols for extreme market scenarios, and circuit breaker mechanisms designed to prevent cascading algorithmic reactions during market stress. Importantly, regulations should be technology-neutral, focusing on outcomes and risk profiles rather than specific algorithmic approaches.

Market Design Considerations: Market operators should evaluate how current market rules and structures might be impacted by widespread RL adoption. Auction mechanisms, price formation processes, and market clearing rules may require reconsideration to ensure they remain robust in environments with significant algorithmic participation. In particular, operators should consider how information disclosure policies influence RL-based strategies and whether current market structures provide sufficient incentives for beneficial flexibility provision while discouraging manipulative behavior.

Transparency and Interpretability Standards: Industry associations and regulators should collaborate to develop standards for transparency and interpretability of RL systems in energy markets. These standards could include requirements for documenting model limitations, reporting backtest methodologies, and providing simplified explanations of decision processes for significant trading actions. Such standards would enhance stakeholder trust while providing a framework for responsible innovation.

Public Research Infrastructure: Government agencies and academic institutions should invest in creating public research infrastructure for energy finance RL applications. This infrastructure could include anonymized market data repositories, standardized simulation environments reflecting realistic market conditions, and benchmark problem sets that enable objective comparison of different approaches. Such resources would democratize research opportunities, accelerate methodological advances, and support more robust model evaluation.

Workforce Development: Educational institutions and industry stakeholders should prioritize developing interdisciplinary training programs that combine energy systems knowledge, financial engineering, and machine-learning expertise. The complexity of RL applications in energy finance requires professionals who understand both the technical nuances of advanced algorithms and the distinctive characteristics of energy markets. Targeted educational initiatives would help address the talent gap in this emerging field.

## 7.4. Synthesis and Outlook

While significant work remains to address the limitations outlined in this review, the promising results to date suggest that RL will increasingly transform how we value, trade, and manage energy assets and contracts in the coming years. The ongoing energy transition—characterized by increasing renewable penetration, storage deployment, and market decentralization—will likely accelerate this transformation by creating greater complexity and optionality that traditional approaches struggle to capture. By combining the adaptive learning capabilities of reinforcement learning with domain-specific knowledge of energy systems, researchers and practitioners can develop more sophisticated tools for navigating the evolving energy finance landscape.

## 8. Conclusions

This review has examined how reinforcement learning can be applied to energy finance. We have highlighted applications ranging from price forecasting and trading strategies to derivatives valuation and option value assessment in power systems.

Energy markets have unique features that make them challenging to model. These markets show extreme price swings, seasonal patterns, and complex regulations. Energy assets like power plants and storage facilities also have physical limitations that create special types of options. Traditional financial models often struggle with these complexities.

Reinforcement learning offers several important advantages for addressing these challenges. First, RL can learn directly from data without needing simplified assumptions about price behavior. Second, RL handles non-linear relationships well, which are common in energy markets. Third, RL adapts to changing market conditions, a crucial feature in evolving energy systems. Fourth, RL naturally incorporates complex constraints that are difficult to include in traditional models.

The ability of RL to capture option value is particularly important. Smart grid technologies, energy storage systems, and demand response, all create significant option value—the net economic benefit of having flexibility under uncertainty. This option value represents the difference in expected system costs between the cases with and without the deployment of the flexible smart grid technology. RL methods are well-suited to quantify this value because they can learn optimal decision policies across many possible future scenarios and capture the sequential nature of energy system operations. Current approaches to quantifying option value have relied primarily on stochastic optimization methods, which, while effective, often become computationally intractable for high-dimensional problems with many uncertainty sources and struggle to capture complex, non-linear relationships and sequential decision processes. In this context, RL provides several advantages over stochastic optimization for option valuation.

Despite these advantages, several challenges remain in applying RL to energy finance. Interpretability concerns make it difficult for decision-makers to trust complex RL models. Data limitations can be problematic since RL algorithms typically need large training datasets. Generalization across different market conditions remains difficult. Computational requirements can be extensive, especially for complex energy systems. Finally, the lack of standardized benchmarks makes it hard to compare different approaches objectively.

Future research directions include developing more explainable RL methods for energy applications, creating robust approaches that perform well under extreme market conditions, exploring multi-agent frameworks that capture strategic interactions among market participants, sector coupling (Goyal et al., 2024) [70], and integrating RL with traditional models. The integration of RL with traditional stochastic optimization methods represents a particularly promising direction as well. Hybrid approaches could combine stochastic optimization with the RL's ability to discover complex non-linear policies. For example, stochastic optimization could define scenario structures and boundary conditions, while RL determines detailed operational policies within these frameworks.

In conclusion, reinforcement learning represents a powerful approach for addressing the unique challenges of energy finance, particularly in capturing the option value created by flexible technologies and operating strategies [172]. While significant work remains to make these methods fully practical for industry applications, the promising results to date suggest that RL will increasingly transform how we value, trade, and manage energy assets and contracts in the coming years.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflicts of interest.

## References

- Eydeland, A.; Wolyniec, K. Energy and Power Risk Management: New Developments in Modeling, Pricing, and Hedging; John Wiley & Sons: Hoboken, NJ, USA, 2003.
- Fischer, T.G. Reinforcement Learning in Financial Markets—A Survey; FAU Discussion Papers in Economics No. 12/2018; Friedrich-Alexander-Universität Erlangen Nürnberg: Erlangen, Germany, 2018.
- 3. Weron, R. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *Int. J. Forecast.* **2014**, *30*, 1030–1081. [CrossRef]
- Giannelos, S.; Moreira, A.; Papadaskalopoulos, D.; Borozan, S.; Pudjianto, D.; Konstantelos, I.; Sun, M.; Strbac, G. A Machine Learning Approach for Generating and Evaluating Forecasts on the Environmental Impact of the Buildings Sector. *Energies* 2023, 16, 2915. [CrossRef]
- 5. Sutton, R.S.; Barto, A.G. Reinforcement Learning: An Introduction, 2nd ed.; MIT Press: Cambridge, MA, USA, 2018.
- Mnih, V.; Kavukcuoglu, K.; Silver, D.; Rusu, A.A.; Veness, J.; Bellemare, M.G.; Graves, A.; Riedmiller, M.; Fidjeland, A.K.; Ostrovski, G.; et al. Human-level control through deep reinforcement learning. *Nature* 2015, 518, 529–533. [CrossRef] [PubMed]
- 7. Kyriakarakos, G. Artificial Intelligence and the Energy Transition. Sustainability 2025, 17, 1140. [CrossRef]

- 8. Most, D.; Giannelos, S.; Yueksel-Erguen, I.; Beulertz, D.; Haus, U.-U.; Charousset-Brignol, S.; Frangioni, A. A Novel Modular Optimization Framework for Modelling Investment and Operation of Energy Systems at European Level; ZIB-Report—20–08; Zuse Institute: Berlin, Germany, 2020.
- 9. Lawryshyn, Y. Using Reinforcement Learning in Applied Real Options Modelling. J. Risk Financ. Manag. 2023, 16, 320.
- 10. Lee, J.S.; Chun, W.; Roh, K.; Heo, S.; Lee, J.H. Applying real options with reinforcement learning to assess commercial CCU deployment. *J. CO*<sub>2</sub> *Util.* **2023**, *77*, 102613. [CrossRef]
- 11. Lillicrap, T.P.; Hunt, J.J.; Pritzel, A.; Heess, N.; Erez, T.; Tassa, Y.; Silver, D.; Wierstra, D. Continuous control with deep reinforcement learning. *arXiv* 2015, arXiv:1509.02971.
- Lippert, I.; Sareen, S. Alleviation of energy poverty through transitions to low-carbon energy infrastructure. *Energy Res. Soc. Sci.* 2023, 100, 103087. [CrossRef]
- Liu, X.; Wu, J.; Chen, S. Efficient Hyperparameters optimization Through Model-based Reinforcement Learning and Meta-Learning. In Proceedings of the IEEE 6th International Conference on Data Science and Systems (HPCC/SmartCity/DSS), Virtual, 14–16 December 2020; pp. 1036–1041.
- 14. Longstaff, F.A.; Schwartz, E.S. Valuing American options by simulation: A simple least-squares approach. *Rev. Financ. Stud.* 2001, 14, 113–147. [CrossRef]
- 15. López-Vargas, A.; Ledezma-Espino, A.; Sanchis-De-Miguel, A. Methods, data sources and applications of the Artificial Intelligence in the Energy Poverty context: A review. *Energy Build.* **2022**, *268*, 112233. [CrossRef]
- 16. Lu, J.; Yang, H.; Wei, Y.; Huang, J. Planning of soft open point considering demand response. In Proceedings of the 2019 IEEE Sustainable Power and Energy Conference (iSPEC), Beijing, China, 21–23 November 2019; pp. 246–251.
- 17. Majeske, N.; Vaidya, S.S.; Roy, R.; Rehman, A.; Sohrabpoor, H.; Miller, T.; Li, W.; Fiddyment, C.R.; Gumennik, A.; Acharya, R.; et al. Industrial energy forecasting using dynamic attention neural networks. *Energy AI* **2025**, *20*, 100504. [CrossRef]
- Marzban, S.; Delage, E.; Li, J.Y.-M. Deep reinforcement learning for option pricing and hedging under dynamic expectile risk measures. *Quant. Financ.* 2023, 23, 1411–1430. [CrossRef]
- 19. Meinshausen, N.; Hambly, B.M. Monte Carlo methods for the valuation of multiple-exercise options. *Math. Financ.* **2004**, *14*, 557–583. [CrossRef]
- 20. Mhlanga, D. Artificial intelligence in the Industry 4.0, and its impact on poverty, innovation, infrastructure development, and the Sustainable Development Goals: Lessons from emerging economies? *Sustainability* **2021**, *13*, 5788. [CrossRef]
- Carmona, R.; Coulon, M. A Survey of commodity markets and structural models for electricity prices. In *Quantitative Energy Finance*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 41–83.
- 22. Benth, F.E.; Šaltytė Benth, J.; Koekebakker, S. Stochastic Modelling of Electricity and Related Markets; World Scientific: Singapore, 2008.
- 23. Conejo, A.J.; Carrión, M.; Morales, J.M. *Decision Making Under Uncertainty in Electricity Markets*; Springer Nature: Dordrecht, The Netherlands, 2010.
- 24. Thompson, M.; Davison, M.; Rasmussen, H. Natural gas storage valuation and optimization: A real options application. *Nav. Res. Logist. NRL* **2009**, *56*, 226–238. [CrossRef]
- 25. Konstantelos, I.; Giannelos, S.; Strbac, G. Strategic valuation of smart grid technology options in distribution networks. *IEEE Trans. Power Syst.* 2017, 32, 1293–1303. [CrossRef]
- 26. Giannelos, S.; Konstantelos, I.; Strbac, G. Option Value of Demand-Side Response Schemes Under Decision-Dependent Uncertainty. *IEEE Trans. Power Syst.* 2018, 33, 5103–5113. [CrossRef]
- 27. Giannelos, S.; Konstantelos, I.; Strbac, G. Stochastic optimisation-based valuation of smart grid options under firm DG contracts. In Proceedings of the 2016 IEEE International Energy Conference (ENERGYCON), Leuven, Belgium, 4–8 April 2016; pp. 1–7.
- 28. Vanegas Cantarero, M.M. Of renewable energy, energy democracy, and sustainable development: A roadmap to accelerate the energy transition in developing countries. *Energy Res. Soc. Sci.* **2020**, *70*, 101716. [CrossRef]
- 29. Carbonneau, A. Pricing and Hedging Financial Derivatives with Reinforcement Learning Methods. Ph.D. Thesis, Concordia University, Montreal, QC, Canada, 2021.
- 30. Cao, J.; Chen, J.; Farghadani, S.; Hull, J.; Poulos, Z.; Wang, Z.; Yuan, J. Gamma and vega hedging using deep distributional reinforcement learning. *Front. Artif. Intell.* **2023**, *6*, 1129370. [CrossRef]
- 31. Caputo, C.; Cardin, M.-A. Analyzing real options and flexibility in engineering systems design using decision rules and deep reinforcement learning. *J. Mech. Des.* **2022**, *144*, 021705. [CrossRef]
- Chauhan, V.S.; Sharma, R.; Shah, H. Exploring sustainability through clean energy, artificial intelligence, and machine learning: Ethical perspectives. In *AI Applications for Clean Energy and Sustainability*; Riswandi, B., Singh, B., Kaunert, C., Vig, K., Eds.; IGI Global Scientific Publishing: Hershey, PA, USA, 2024; pp. 119–138. [CrossRef]
- 33. Charousset-Brignol, S.; van Ackooij, W.; Oudjane, N.; Daniel, D.; Noceir, S.; Haus, U.-U.; Lazzaro, A.; Frangioni, A.; Lobato, R.; Ghezelsofla, A.; et al. Synergistic approach of multi-energy models for a European optimal energy system management tool. *Proj. Repos. J.* 2021, 9, 113–116.

- 34. Chen, A.-S.; Leung, M.T.; Pan, S.; Chou, C.-Y. Financial hedging in energy market by cross-learning machines. *Neural Comput. Appl.* **2020**, *32*, 10321–10335. [CrossRef]
- 35. Chen, B.; Wang, J.; Wang, L.; He, Y.; Wang, Z. Robust Optimization for Transmission Expansion Planning: Minimax Cost vs. Minimax Regret. *IEEE Trans. Power Syst.* 2014, 29, 3069–3077. [CrossRef]
- Chen, C.-F.; Napolitano, R.; Hu, Y.; Kar, B.; Yao, B. Addressing machine learning bias to foster energy justice. *Energy Res. Soc. Sci.* 2024, 116, 103653. [CrossRef]
- 37. Che, X.; Zhu, B.; Wang, P. Assessing global energy poverty: An integrated approach. Energy Policy 2021, 149, 112099. [CrossRef]
- Cheraghi, Y.; Bratvold, R.B.; Muhammad, R.B. Value Creation in Sustainable Energy Transition Using Reinforcement Learning. Energies 2024, 17, 854.
- Chronopoulos, M.; Hagspiel, V.; Fleten, S.-E. Stepwise green investment under policy uncertainty. *Energy J.* 2016, 37, 87–108. [CrossRef]
- 40. Cramton, P.; Ockenfels, A.; Stoft, S. Capacity market fundamentals. Econ. Energy Environ. Policy 2013, 2, 27-46. [CrossRef]
- Dalal, G.; Gilboa, E.; Mannor, S. Hierarchical decision making in electricity grid management. In Proceedings of the International Conference on Machine Learning, New York, NY, USA, 20–22 June 2016; pp. 1153–1162.
- del Guayo, Í.; Cuesta, Á. Towards a just energy transition: A critical analysis of the existing policies and regulations in Europe. J. World Energy Law Bus. 2022, 15, 212–222. [CrossRef]
- 43. Deng, J.L. Control problems of grey systems. Syst. Control Lett. 1982, 1, 288–294.
- 44. Deng, S.; Oren, S. Electricity derivatives and risk management. Energy 2006, 31, 940–953. [CrossRef]
- 45. Dong, Z.; Zhang, X.; Zhang, L.; Giannelos, S.; Strbac, G. Flexibility enhancement of urban energy systems through coordinated space heating aggregation of numerous buildings. *Appl. Energy* **2024**, *374*, 123971. [CrossRef]
- Du, Y.; Li, F.; Zandi, H.; Xue, Y. Approximating Nash Equilibrium in Day-ahead Electricity Market Bidding with Multi-agent Deep Reinforcement Learning. J. Mod. Power Syst. Clean Energy 2021, 9, 534–544. [CrossRef]
- 47. Ersen, H.Y.; Tas, O.; Ugurlu, U. Solar energy investment valuation with intuitionistic fuzzy trinomial lattice real option model. *IEEE Trans. Eng. Manag.* **2023**, *70*, 2584–2593. [CrossRef]
- 48. Fabra, N.; Reguant, M. Pass-through of emissions costs in electricity markets. Am. Econ. Rev. 2014, 104, 2872–2899. [CrossRef]
- 49. Famoso, F.; Oliveri, L.M.; Brusca, S.; Chiacchio, F. A Dependability Neural Network Approach for Short-Term Production Estimation of a Wind Power Plant. *Energies* **2024**, *17*, 1627. [CrossRef]
- 50. FERC (Federal Energy Regulatory Commission). *The February 2021 Cold Weather Outages in Texas and the South Central United States;* FERC, NERC and Regional Entity Staff Report; FERC: Washington, DC, USA, 2021.
- 51. Frestad, D.; Benth, F.E.; Koekebakker, S. Modeling term structure dynamics in the Nordic electricity swap market. *Energy J.* 2010, *31*, 53–86. [CrossRef]
- 52. Fuad, K.S.; Hafezi, H.; Kauhaniemi, K.; Laaksonen, H. Soft open point in distribution networks. *IEEE Access* 2020, *8*, 210550–210565. [CrossRef]
- 53. Gawusu, S.; Jamatutu, S.A.; Ahmed, A. Predictive modeling of energy poverty with machine learning ensembles: Strategic insights from socioeconomic determinants for effective policy implementation. *Int. J. Energy Res.* 2024, 2024, 9411326. [CrossRef]
- 54. Gawusu, S.; Jamatutu, S.A.; Zhang, X.; Moomin, S.T.; Ahmed, A.; Mensah, R.A.; Das, O.; Ackah, I. Spatial analysis and predictive modeling of energy poverty: Insights for policy implementation. *Environ. Dev. Sustain.* **2024**, 1–48. [CrossRef]
- 55. Giannelos, S.; Borozan, S.; Strbac, G. A Backwards Induction Framework for Quantifying the Option Value of Smart Charging of Electric Vehicles and the Risk of Stranded Assets under Uncertainty. *Energies* **2022**, *15*, 3334. [CrossRef]
- 56. Giannelos, S.; Borozan, S.; Aunedi, M.; Zhang, X.; Ameli, H.; Pudjianto, D.; Konstantelos, I.; Strbac, G. Modelling Smart Grid Technologies in Optimisation Problems for Electricity Grids. *Energies* **2023**, *16*, 5088. [CrossRef]
- 57. Giannelos, S.; Borozan, S.; Konstantelos, I.; Strbac, G. Option value, investment costs and deployment levels of smart grid technologies. *Sustain. Energy Res.* 2024, 11, 47. [CrossRef]
- 58. Giannelos, S.; Borozan, S.; Strbac, G.; Zhang, T.; Kong, W. Vehicle-to-Grid: Quantification of its contribution to security of supply through the F-Factor methodology. *Sustain. Energy Res.* **2024**, *11*, 32. [CrossRef]
- Giannelos, S.; Borozan, S.; Moreira, A.; Strbac, G. Techno-Economic Analysis of Smart EV Charging for Expansion Planning Under Uncertainty. In Proceedings of the 2023 IEEE Belgrade PowerTech, Belgrade, Serbia, 25–29 June 2023; pp. 1–7.
- 60. Giannelos, S.; Djapic, P.; Pudjianto, D.; Strbac, G. Quantification of the Energy Storage Contribution to Security of Supply through the F-Factor Methodology. *Energies* 2020, *13*, 826. [CrossRef]
- 61. Giannelos, S.; Jain, A.; Borozan, S.; Falugi, P.; Moreira, A.; Bhakar, R.; Mathur, J.; Strbac, G. Long-Term Expansion Planning of the Transmission Network in India under Multi-Dimensional Uncertainty. *Energies* **2021**, *14*, 7813. [CrossRef]
- Giannelos, S.; Konstantelos, I.; Strbac, G. A new class of planning models for option valuation of storage technologies under decision-dependent innovation uncertainty. In Proceedings of the 2017 IEEE Manchester PowerTech, Manchester, UK, 18–22 June 2017; pp. 1–6.

- 63. Giannelos, S.; Konstantelos, I.; Strbac, G. Option value of dynamic line rating and storage. In Proceedings of the IEEE International Energy Conference (ENERGYCON), Limassol, Cyprus, 3–7 June 2018.
- 64. Giannelos, S.; Konstantelos, I.; Strbac, G. Endogenously stochastic demand side response participation on transmission system level. In Proceedings of the IEEE International Energy Conference (ENERGYCON), Limassol, Cyprus, 3–7 June 2018.
- 65. Giannelos, S.; Zhang, T.; Pudjianto, D.; Konstantelos, I.; Strbac, G. Investments in Electricity Distribution Grids: Strategic versus Incremental Planning. *Energies* **2024**, *17*, 2724. [CrossRef]
- 66. Giannelos, S.; Zhang, X.; Zhang, T.; Strbac, G. Multi-Objective Optimization for Pareto Frontier Sensitivity Analysis in Power Systems. *Sustainability* **2024**, *16*, 5854. [CrossRef]
- 67. Giannelos, S.; Pudjianto, D.; Zhang, T.; Strbac, G. Energy Hub Operation Under Uncertainty: Monte Carlo Risk Assessment Using Gaussian and KDE-Based Data. *Energies* 2025, *18*, 1712. [CrossRef]
- 68. Glasserman, P. Monte Carlo Methods in Financial Engineering; Springer: Berlin/Heidelberg, Germany, 2003.
- 69. Gomez, A.A.; Consigli, G.; Liu, J. Multi-period portfolio selection with interval-based conditional value-at-risk. *Ann. Oper. Res.* **2024**, 1–39. [CrossRef]
- Goyal, A.; Bhattacharya, K. Optimal design of a decarbonized sector-coupled microgrid: Electricity-heat-hydrogen-transport sectors. *IEEE Access* 2024, 12, 38399–38409. [CrossRef]
- 71. Greenwood, D.M.; Djapic, P.; Sarantakos, I.; Giannelos, S.; Strbac, G.; Creighton, A. Pragmatic method for assessing the security of supply in future smart distribution networks. *Cired—Open Access Proc. J.* 2020, 2020, 221–224. [CrossRef]
- 72. Guo, Y.; Wang, N.; Xu, Z.-Y.; Wu, K. The internet of things-based decision support system for information processing in intelligent manufacturing using data mining technology. *Mech. Syst. Signal Process.* **2020**, *142*, 106630. [CrossRef]
- 73. Halkos, G.E.; Tsirivis, A.S. Value-at-risk methodologies for effective energy portfolio risk management. *Econ. Anal. Policy* **2019**, *62*, 197–212. [CrossRef]
- 74. Halperin, I. QLBS: Q-Learner in the Black-Scholes (-Merton) worlds. J. Deriv. 2019, 26, 99–123.
- 75. Hilliard, J.E.; Reis, J. Valuation of commodity futures and options under stochastic convenience yields, interest rates, and jump diffusions in the spot. *J. Financ. Quant. Anal.* **1998**, *33*, 61. [CrossRef]
- 76. Higgs, H.; Worthington, A. Stochastic price modeling of high volatility, mean-reverting, spike-prone commodities: The Australian wholesale spot electricity market. *Energy Econ.* **2008**, *30*, 3172–3185. [CrossRef]
- 77. Hogan, W.W. Contract networks for electric power transmission. J. Regul. Econ. 1992, 4, 211–242. [CrossRef]
- 78. Hosseini, E.; Saeedpour, B.; Banaei, M.; Ebrahimy, R. Optimized deep neural network architectures for energy consumption and PV production forecasting. *Energy Strategy Rev.* **2025**, *59*, 101704. [CrossRef]
- 79. Holttinen, H.; Kiviluoma, J.; Helistö, N.; Levy, T.; Menemenlis, N.; Jun, L.; Cutululis, N.; Koivisto, M.; Das, K.; Orths, A.; et al. Design and Operation of Energy Systems with Large Amounts of Variable Generation: Final Summary Report, IEA Wind TCP Task 25; VTT Technical Research Centre of Finland, VTT Technology: Espoo, Finland, 2021. [CrossRef]
- 80. Hull, J.C. Options, Futures, and Other Derivatives, 10th ed.; Pearson: New York, NY, USA, 2017.
- Ilo, A.; Prata, R.; Strbac, G.; Giannelos, S.; Bissell, G.R.; Kulmala, A.; Constantinescu, N.; Samovich, N.; Iliceto, A. White Paper ETIP SNET—Holistic Architectures for Power Systems. 2019. Available online: http://hdl.handle.net/20.500.12708/39729 (accessed on 6 February 2025).
- Jain, V.; Mitra, A. Artificial intelligence and machine learning for sustainable development: Enhancing health, equity, and environmental sustainability. In *Machine and Deep Learning Solutions for Achieving the Sustainable Development Goals*; Ruiz-Vanoye, J., Díaz-Parra, O., Eds.; IGI Global Scientific Publishing: Hershey, PA, USA, 2025; pp. 107–124. [CrossRef]
- 83. Janczura, J.; Trück, S.; Weron, R.; Wolff, R.C. Identifying spikes and seasonal components in electricity spot price data: A guide to robust modeling. *Energy Econ.* 2013, *38*, 96–110. [CrossRef]
- 84. Janner, M.; Fu, J.; Zhang, M.; Levine, S. When to trust your model: Model-based policy optimization. arXiv 2019, arXiv:1906.08253.
- Jiang, D.R.; Powell, W.B. Risk-averse approximate dynamic programming with quantile-based risk measures. *Math. Oper. Res.* 2018, 43, 554–579. [CrossRef]
- Jiang, X.; Zhou, Y.; Ming, W.; Yang, P.; Wu, J. An overview of soft open points in electricity distribution networks. *IEEE Trans. Smart Grid* 2022, 13, 1899–1910. [CrossRef]
- 87. Joskow, P.L. Lessons learned from electricity market liberalization. Energy J. 2008, 29, 9–42. [CrossRef]
- 88. Jorion, P. Value at Risk: The New Benchmark for Managing Financial Risk, 3rd ed.; McGraw-Hill: Columbus, OH, USA, 2007.
- 89. Judson, E.; Fitch-Roy, O.; Soutar, I. Energy democracy: A digital future? Energy Res. Soc. Sci. 2022, 91, 102732. [CrossRef]
- 90. Watkins, C.J.C.H.; Dayan, P. Q-learning. Mach. Learn. 1992, 8, 279-292. [CrossRef]
- 91. Williams, R.J. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Mach. Learn.* **1992**, *8*, 229–256. [CrossRef]

- 92. Schulman, J.; Levine, S.; Abbeel, P.; Jordan, M.; Moritz, P. Trust region policy optimization. In Proceedings of the 32nd International Conference on Machine Learning (ICML), Lille, France, 6–11 July 2015; pp. 1889–1897.
- 93. Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; Klimov, O. Proximal policy optimization algorithms. arXiv 2017. [CrossRef]
- 94. Sutton, R.S. Integrated architectures for learning, planning, and reacting based on approximating dynamic programming. In Proceedings of the Seventh International Conference on Machine Learning, Austin, TX, USA, 21–23 June 1990; pp. 216–224.
- 95. Bertsekas, D.P. *Dynamic Programming and Optimal Control*; Approximate dynamic programming; Athena Scientific: Nashua, NH, USA, 2012; Volume II.
- 96. Wang, W.; Huang, Y.; Wang, S. First-order adversarial vulnerability of neural networks and input dimension. In Proceedings of the 36th International Conference on Machine Learning, Long Beach, CA, USA, 9–15 June 2019; pp. 6543–6552.
- 97. Prete, C.L.; Blumsack, S. Enhancing the reliability of bulk power systems against the threat of extreme weather: Lessons from the 2021 Texas electricity crisis. *Econ. Energy Environ. Policy* **2023**, *12*, 31–48. [CrossRef]
- 98. Schwartz, E.S. The stochastic behavior of commodity prices: Implications for valuation and hedging. *J. Financ.* **1997**, *52*, 923–973. [CrossRef]
- Zareipour, H.; Bhattacharya, K.; Canizares, C. Forecasting the hourly Ontario energy price by multivariate adaptive regression splines. In Proceedings of the 2006 IEEE Power Engineering Society General Meeting, Montreal, QC, Canada, 18–22 June 2006.
- 100. Busby, J.W.; Baker, K.; Bazilian, M.D.; Gilbert, A.Q.; Grubert, E.; Rai, V.; Rhodes, J.D.; Shidore, S.; Smith, C.A.; Webber, M.E. Cascading risks: Understanding the 2021 winter blackout in Texas. *Energy Res. Soc. Sci.* 2021, 77, 102106. [CrossRef]
- 101. Nick, S.; Thoenes, S. What drives natural gas prices?—A structural VAR approach. Energy Econ. 2014, 45, 517–527. [CrossRef]
- 102. Knittel, C.R.; Roberts, M.R. An empirical examination of restructured electricity prices. Energy Econ. 2005, 27, 791-817. [CrossRef]
- 103. Suenaga, H.; Smith, A.; Williams, J. Volatility dynamics of NYMEX natural gas futures prices. J. Futur. Mark. 2008, 28, 438–463. [CrossRef]
- 104. Brown, S.P.A.; Yücel, M.K. What drives natural gas prices? Energy J. 2008, 29, 45–60. [CrossRef]
- 105. Paraschiv, F.; Erni, D.; Pietsch, R. The impact of renewable energies on EEX day-ahead electricity prices. *Energy Policy* **2015**, *73*, 196–210. [CrossRef]
- 106. Ketterer, J.C. The impact of wind power generation on the electricity price in Germany. Energy Econ. 2014, 44, 270-280. [CrossRef]
- 107. Potomac Economics; Electric Reliability Council of Texas. 2019 State of the Market Report for the ERCOT Electricity Markets. May 2020. Available online: https://www.potomaceconomics.com/wp-content/uploads/2020/06/2019-State-of-the-Market-Report.pdf (accessed on 3 February 2025).
- 108. Staffell, I.; Green, R. Is there still merit in the merit order stack? The impact of dynamic constraints on optimal plant mix. *IEEE Trans. Power Syst.* **2015**, *31*, 43–53. [CrossRef]
- 109. Woo, C.; Horowitz, I.; Moore, J.; Pacheco, A. The impact of wind generation on the electricity spot-market price level and variance: The Texas experience. *Energy Policy* **2011**, *39*, 3939–3944. [CrossRef]
- 110. Staffell, I.; Rustomji, M. Maximising the value of electricity storage. J. Energy Storage 2016, 8, 212–225. [CrossRef]
- 111. Bunn, D.W.; Oliveira, F.S. Agent-based analysis of technological diversification and specialization in electricity markets. *Eur. J. Oper. Res.* **2007**, *181*, 1265–1278. [CrossRef]
- 112. Brinkmann, E.J.; Rabinovitch, R. Regional limitations on the hedging effectiveness of natural gas futures. *Energy J.* **1995**, *16*, 113–124. [CrossRef]
- 113. Karakatsani, N.V.; Bunn, D.W. Forecasting electricity prices: The impact of fundamentals and time-varying coefficients. *Int. J. Forecast.* 2008, 24, 764–785. [CrossRef]
- 114. Boukas, I.; Ernst, D.; Théate, T.; Bolland, A.; Huynen, A.; Buchwald, M.; Wynants, C.; Cornélusse, B. A deep reinforcement learning framework for continuous intraday market bidding. *Mach. Learn.* **2021**, *110*, 2335–2387. [CrossRef]
- 115. Al Moti, M.M.; Uddin, R.S.; Hai, A.; Bin Saleh, T.; Alam, G.R.; Hassan, M.M.; Hassan, R. Blockchain Based Smart-Grid Stackelberg Model for Electricity Trading and Price Forecasting Using Reinforcement Learning. *Appl. Sci.* **2022**, *12*, 5144. [CrossRef]
- 116. Pannakkong, W.; Vinh, V.T.; Tuyen, N.N.M.; Buddhakulsomsiri, J. A reinforcement learning approach for ensemble machine learning models in peak electricity forecasting. *Energies* **2023**, *16*, 5099. [CrossRef]
- 117. Mulliez, M. Dynamic Hedging in the Presence of Basis Risk: A Reinforcement Learning Approach. Master's Thesis, Imperial College London, London, UK, 2021.
- 118. Madahi, S.S.K.; Claessens, B.; Develder, C. Distributional reinforcement learning-based energy arbitrage strategies in imbalance settlement mechanism. *J. Energy Storage* 2024, 104, 114377. [CrossRef]
- 119. Bahrami, S.; Hooshmand, R.-A.; Parastegari, M. Short term electric load forecasting by wavelet transform and grey model improved by PSO (particle swarm optimization) algorithm. *Energy* **2014**, *72*, 434–442. [CrossRef]
- 120. Wu, L.; Liu, S.; Yao, L.; Yan, S. The effect of sample size on the grey system prediction. *Appl. Math. Model.* **2013**, *37*, 6577–6583. [CrossRef]

- 121. Zhao, H.; Wu, Q.; Hu, S.; Xu, H.; Rasmussen, C.N. Review of energy storage system for wind power integration support. *Appl. Energy* **2015**, *137*, 545–553. [CrossRef]
- 122. Mystakidis, A.; Koukaras, P.; Tsalikidis, N.; Ioannidis, D.; Tjortjis, C. Energy Forecasting: A Comprehensive Review of Techniques and Technologies. *Energies* 2024, 17, 1662. [CrossRef]
- 123. Boucetta, L.N.; Amrane, Y.; Chouder, A.; Arezki, S.; Kichou, S. Enhanced Forecasting Accuracy of a Grid-Connected Photovoltaic Power Plant: A Novel Approach Using Hybrid Variational Mode Decomposition and a CNN-LSTM Model. *Energies* 2024, 17, 1781. [CrossRef]
- 124. Dos Reis, J.R.; Tabora, J.M.; de Lima, M.C.; Monteiro, F.P.; Monteiro, S.C.D.A.; Bezerra, U.H.; Tostes, M.E.D.L. Medium and long term energy forecasting methods: A literature review. *IEEE Access* 2025, *13*, 29305–29326. [CrossRef]
- 125. Sadeghi, M.; Shavvalpour, S. Energy risk management and value at risk modeling. Energy Policy 2006, 34, 3367–3373. [CrossRef]
- 126. Wang, X.; Liu, H.; Yao, Y. Value-at-Risk forecasting for the Chinese new energy stock market: An explainable quantile regression neural network method. *Procedia Comput. Sci.* 2024, 242, 1096–1103. [CrossRef]
- Abdullah, B.U.D.; Khanday, S.A.; Islam, N.U.; Lata, S.; Fatima, H.; Nengroo, S.H. Comparative Analysis Using Multiple Regression Models for Forecasting Photovoltaic Power Generation. *Energies* 2024, 17, 1564. [CrossRef]
- 128. Zhang, S.; Tu, L.; Duan, Q.; Chao, Z.; Tang, X.; Wanyan, X.; Chen, X. Conditional Value at Risk Model of New Power System Reserve Assessment Considering Primary Energy Supply Risk. In *Energy Power and Automation Engineering*; ICEPAE Lecture Notes in Electrical Engineering; Springer: Singapore, 2023; Volume 1118.
- 129. Trabelsi, N.; Tiwari, A.K.; Ghallabi, F.; Khemakhem, I. Nexus of crude oil and clean energy stock indices: Evidence from time-vector-auto-regression in conjunction with conditional-autoregressive-value-at-risk. *Heliyon* **2025**, *11*, e40970. [CrossRef]
- Barrera-Rivera, R.R.; Valencia-Herrera, H. Hedging and optimization of energy asset portfolios. In Artificial Intelligence and Soft Computing for Energy Systems; Springer: Berlin/Heidelberg, Germany, 2022; pp. 113–126.
- 131. Syalsabila, A.; Prastyo, D.D.; Akbar, M.S.; Rahayu, S.P.; Deivanayagampillai, N. Conditional Value-At-Risk Modelling Using Hybrid LASSO-QRNN to Quantify the Market Risk Dependence on Oil and Gas Companies' Stock in Indonesia. In Advances in Manufacturing Processes and Smart Manufacturing Systems, GCMM 2023; Lecture Notes in Networks and Systems; Springer: Cham, Switzerland, 2024; Volume 1215. [CrossRef]
- 132. Wilmott, P.; Dewynne, J.; Howison, S. Option Pricing: Mathematical Models and Computation; Oxford Financial Press: Oxford, UK, 1995.
- 133. Buehler, H.; Gonon, L.; Teichmann, J.; Wood, B. Deep hedging. Quant. Financ. 2019, 19, 1271–1291. [CrossRef]
- 134. Becker, S.; Cheridito, P.; Jentzen, A. Deep optimal stopping. J. Mach. Learn. Res. 2019, 20, 1–25.
- Becker, S.; Cheridito, P.; Jentzen, A. Pricing and hedging American-style options with deep learning. J. Risk Financ. Manag. 2020, 13, 158. [CrossRef]
- 136. Song, C. Design and application of financial market option pricing system based on high-performance computing and deep reinforcement learning. *Sci. Program.* **2022**, 2022, 8525361. [CrossRef]
- 137. Boogert, A.; de Jong, C. Gas storage valuation using a Monte Carlo method. J. Deriv. 2008, 15, 81–98. [CrossRef]
- Oren, S.S. Integrating real and financial options in demand-side electricity contracts. *Decis. Support Syst.* 2001, 30, 279–288.
  [CrossRef]
- 139. Tseng, C.-L.; Barz, G. Short-Term Generation Asset Valuation: A Real Options Approach. Oper. Res. 2002, 50, 297–310. [CrossRef]
- Nadarajah, S.; Secomandi, N. A review of the operations literature on real options in energy. *Eur. J. Oper. Res.* 2022, 309, 469–487.
  [CrossRef]
- 141. Alqubaisi, A. Deep Real Options: Valuation of Real Options on Green Energy using Deep Learning Methods. *Energy Econ.* **2023**, *120*, 106553.
- Beulertz, D.; Charousset, S.; Most, D.; Giannelos, S.; Yueksel-Erguen, I. Development of a modular framework for future energy system analysis. In Proceedings of the 2019 54th International Universities Power Engineering Conference (UPEC), Bucharest, Romania, 3–6 September 2019; pp. 1–6.
- Siddiqui, A.S.; Maribu, K. Investment and upgrade in distributed generation under uncertainty. *Energy Econ.* 2009, 31, 25–37.
  [CrossRef]
- 144. Nick, M.; Cherkaoui, R.; Paolone, M. Optimal planning of distributed energy storage systems in active distribution networks embedding grid reconfiguration. *IEEE Trans. Power Syst.* **2017**, *33*, 1577–1590. [CrossRef]
- Wogrin, S.; Galbally, D.; Reneses, J. Optimizing storage operations in medium- and long-term power system models. *IEEE Trans. Power Syst.* 2015, *31*, 3129–3138. [CrossRef]
- 146. Papadaskalopoulos, D.; Strbac, G. Nonlinear and sequential pricing models for active demand response. *IEEE Trans. Smart Grid* **2017**, *8*, 1349–1359.

- 147. Amann, G.; Escobedo Bermúdez, V.R.; Boskov-Kovacs, E.; Gallego Amores, S.; Giannelos, S.; Iliceto, A.; Ilo, A.; Chavarro, J.R.; Samovich, N.; Schmitt, L.; et al. *E-Mobility Deployment and Impact on Grids: Impact of EV and Charging Infrastructure on European T&D Grids: Innovation Needs*; Gallego Amores, S., Ed.; MJ-09-22-246-EN-N; Publications Office of the European Union: Luxembourg, 2022. [CrossRef]
- 148. Borozan, S.; Giannelos, S.; Strbac, G. Strategic network expansion planning with electric vehicle smart charging concepts as investment options. *Adv. Appl. Energy* **2022**, *5*, 100077. [CrossRef]
- Borozan, S.; Giannelos, S.; Aunedi, M.; Strbac, G. Option value of EV smart charging concepts in transmission expansion planning under uncertainty. In Proceedings of the 2022 IEEE 21st Mediterranean Electrotechnical Conference (MELECON), Palermo, Italy, 14–16 June 2022; pp. 63–68.
- 150. Giannelos, S.; Konstantelos, I.; Strbac, G. Investment Model for Cost-effective Integration of Solar PV Capacity under Uncertainty using a Portfolio of Energy Storage and Soft Open Points. In Proceedings of the 2019 IEEE Milan PowerTech, Milan, Italy; 2019; pp. 1–6. [CrossRef]
- Borozan, S.; Giannelos, S.; Falugi, P.; Moreira, A.; Strbac, G. Machine Learning-Enhanced Benders Decomposition Approach for the Multi-Stage Stochastic Transmission Expansion Planning Problem. *Electr. Power Syst. Res.* 2024, 237, 110985. [CrossRef]
- 152. Khurana, U.; Samulowitz, H.; Turaga, D. Feature Engineering for Predictive Modeling Using Reinforcement Learning. *Proc. AAAI Conf. Artif. Intell.* 2018, 32, 3407–3414. [CrossRef]
- 153. Moran, M.; Gordon, G. Deep Curious Feature Selection: A Recurrent, Intrinsic-Reward Reinforcement Learning Approach to Feature Selection. *IEEE Trans. Artif. Intell.* 2023, *5*, 1174–1184. [CrossRef]
- Kaloev, M.; Krastev, G. Experiments Focused on Exploration in Deep Reinforcement Learning. In Proceedings of the 2021 5th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, Turkey, 21–23 October 2021; pp. 351–355.
- 155. Tittaferrante, A.; Yassine, A. Benchmarking Offline Reinforcement Learning. In Proceedings of the 2022 21st IEEE International Conference on Machine Learning and Applications (ICMLA), Nassau, Bahamas, 12–14 December 2022; pp. 259–263.
- 156. Kamruzzaman; Duan, J.; Shi, D.; Benidris, M. A Deep Reinforcement Learning-Based Multi-Agent Framework to Enhance Power System Resilience Using Shunt Resources. *IEEE Trans. Power Syst.* 2021, *36*, 5525–5536. [CrossRef]
- 157. Nkurunziza, F.; Kabanda, R.; McSharry, P. Enhancing poverty classification in developing countries through machine learning: A case study of household consumption prediction in Rwanda. *Cogent Econ. Financ.* **2024**, *13*, 2444374. [CrossRef]
- 158. Raghavendra, A.H.; Majhi, S.G.; Mukherjee, A.; Bala, P.K. Role of artificial intelligence (AI) in poverty alleviation: A bibliometric analysis. *VINE J. Inf. Knowl. Manag. Syst.* 2023, 55, 710–729. [CrossRef]
- 159. Satria, D.; Permani, R.; Winarno, K.; Kaluge, D.; Indraswari, C.R.; Handrito, R.P. An exploratory study of high-educated poverty through machine learning approach: A case study of East Java, Indonesia. *J. Bus. Manag. Econ. Eng.* 2025, 23, 92–107. [CrossRef]
- Kwilinski, A.; Lyulyov, O.; Pimonenko, T. Energy Poverty and Democratic Values: A European Perspective. *Energies* 2024, 17, 2837.
  [CrossRef]
- 161. Dall-Orsoletta, A.; Cunha, J.; Araújo, M.; Ferreira, P. A systematic review of social innovation and community energy transitions. *Energy Res. Soc. Sci.* **2022**, *88*, 102625. [CrossRef]
- Palma, G.; Guiducci, L.; Stentati, M.; Rizzo, A.; Paoletti, S. Reinforcement Learning for Energy Community Management: A European-Scale Study. *Energies* 2024, 17, 1249. [CrossRef]
- 163. Ponse, K.; Kleuker, F.; Fejér, M.; Serra-Gómez, Á.; Plaat, A.; Moerland, T. Reinforcement learning for sustainable energy: A survey. *arXiv* 2024. [CrossRef]
- 164. Pillan, M.; Costa, F.; Caiola, V. How could people and communities contribute to the energy transition? Conceptual maps to inform, orient, and inspire design actions and education. *Sustainability* **2023**, *15*, 14600. [CrossRef]
- 165. Neij, L.; Palm, J.; Busch, H.; Bauwens, T.; Becker, S.; Bergek, A.; Buzogány, A.; Candelise, C.; Coenen, F.; Devine-Wright, P.; et al. Energy communities—Lessons learnt, challenges, and policy recommendations. *Oxf. Open Energy* **2025**, *4*, oiaf002. [CrossRef]
- 166. Abbas, K.; Butt, K.M.; Xu, D.; Ali, M.; Baz, K.; Kharl, S.H.; Ahmed, M. Measurements and determinants of extreme multidimensional energy poverty using machine learning. *Energy* **2022**, 251, 123977. [CrossRef]
- 167. Alimi, O.A.; Ouahada, K.; Abu-Mahfouz, A.M. A review of machine learning approaches to power system security and stability. *IEEE Access* **2020**, *8*, 113512–113531. [CrossRef]
- 168. Piras, G.; Muzi, F.; Ziran, Z. Open tool for automated development of renewable energy communities: Artificial intelligence and machine learning techniques for methodological approach. *Energies* **2024**, *17*, 5726. [CrossRef]
- Kaur, S.; Kumar, R.; Singh, K.; Huang, Y. Leveraging Artificial Intelligence for Enhanced Sustainable Energy Management. J. Sustain. Energy 2024, 3, 1–20. [CrossRef]
- 170. Nalli, P.K.; Manikandan, K.P.; Padmapriya, G.; Bhatt, D.; Talukdar, N.; Premkumar, R. Optimizing energy systems using machine learning and artificial intelligence. In *Integrating Artificial Intelligence Into the Energy Sector*; Derbali, A., Ed.; IGI Global Scientific Publishing: Hershey, PA, USA, 2025; pp. 493–514. [CrossRef]

- 171. Alturif, G.; Saleh, W.; El-Bary, A.A.; Osman, R.A. Using artificial intelligence tools to predict and alleviate poverty. *Entrep. Sustain. Issues* **2024**, *12*, 400–413. [CrossRef]
- 172. Bose, S.; Kremers, E.; Mengelkamp, E.M.; Eberbach, J.; Weinhardt, C. Reinforcement learning in local energy markets. *Energy Inform.* **2021**, *4*, 7. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.