analytics

Article

System Inertia Cost Forecasting Using Machine Learning: A Data-Driven Approach for Grid Energy Trading in Great Britain

Maitreyee Dey [1,*], Soumya Prakash Rana [2] and Preeti Patel [1]

1 GENESIS Research Lab, Cyber Security Research Centre, London Metropolitan University, 166-220 Holloway Road, London N7 8DB, UK; p.patel@londonmet.ac.uk
2 School of Engineering, University of Greenwich, Medway Campus, Central Avenue, Chatham ME4 4TB, UK; s.rana@greenwich.ac.uk
* Correspondence: m.dey@londonmet.ac.uk

Abstract

As modern power systems integrate more renewable and decentralised generation, maintaining grid stability has become increasingly challenging. This study proposes a data-driven machine learning framework for forecasting system inertia service costs—a key yet underexplored variable influencing energy trading and frequency stability in Great Britain. Using eight years (2017–2024) of National Energy System Operator (NESO) data, four models—Long Short-Term Memory (LSTM), Residual LSTM, eXtreme Gradient Boosting (XGBoost), and Light Gradient-Boosting Machine (LightGBM)—are comparatively analysed. LSTM-based models capture temporal dependencies, while ensemble methods effectively handle nonlinear feature relationships. Results demonstrate that LightGBM achieves the highest predictive accuracy, offering a robust method for inertia cost estimation and market intelligence. The framework contributes to strategic procurement planning and supports market design for a more resilient, cost-effective grid.

Keywords: smart grid; frequency; inertia; market estimation; machine learning

check for updates

Academic Editor: Ping-Feng Pai

Received: 4 July 2025
Revised: 17 October 2025
Accepted: 21 October 2025
Published: 23 October 2025

Citation: Dey, M.; Rana, S.P.; Patel, P. System Inertia Cost Forecasting Using Machine Learning: A Data-Driven Approach for Grid Energy Trading in Great Britain. *Analytics* **2025**, *4*, 30. https://doi.org/10.3390/analytics4040030

1. Introduction

The National Energy System Operator (NESO) is responsible for the long-term planning and real-time operation of the electricity and gas systems in Great Britain [1,2]. It ensures a secure and reliable energy supply, promotes efficient energy transmission and distribution, and supports the government's efforts to achieve its net-zero target. NESO also plays a key role in market operations, system planning, and future-proofing the energy infrastructure.

NESO regularly publishes reports—such as system performance reports, balancing services performance monitoring reports, and the GB Electricity System Operator Daily Reports—that provide data on grid performance and operations [3]. In its operability strategy report published in December 2022, National Grid stated that its current policy is to maintain system inertia above 140 GJ. However, by 2025, it aims to maintain a minimum system inertia of 96 GJ [3]. To balance the system, address forecasted energy requirements, and ensure system security, NESO engages in energy trading. This includes trading with third parties to adjust supply and demand and to manage grid constraints.

Inertia in power systems refers to the energy stored in large rotating generators and certain industrial motors, which enables them to resist changes in rotational speed. This stored energy is especially valuable in the event of a large power plant failure, as it can temporarily compensate for the lost generation [4]. Monitoring system inertia and its associated costs is essential for maintaining grid balance and minimising balancing costs, which are ultimately passed on to consumers via energy bills [5].

In recent years, AI and machine learning (ML) have emerged as powerful tools in the operation and optimisation of modern energy systems. These technologies enable real-time data processing, anomaly detection, demand forecasting, and market modelling. In the context of system inertia, ML algorithms can uncover complex temporal patterns and nonlinear interactions that are often missed by traditional statistical models. By training models on historical cost and grid performance data, AI-based solutions can predict future cost fluctuations, identify high-risk scenarios, and support more agile decision-making [6]. Inertia forecasting in particular stands to benefit from sequence learning and gradient-boosted decision trees, which have shown success in related applications such as energy demand prediction, electricity price forecasting, and renewable output estimation.

Accurate forecasting of system inertia costs is crucial for operational planning and market optimisation in low-inertia power grids. While significant work has been conducted on energy price, demand, and renewable generation forecasting, inertia-related cost prediction remains largely unexplored. The decline in synchronous generation and growing dependence on inverter-based renewable sources have introduced variability in system inertia, driving the need for data-driven models capable of anticipating associated balancing costs.

This research addresses that gap by proposing a comparative machine learning framework for daily system inertia cost forecasting. The study systematically evaluates four models—Long Short-Term Memory (LSTM), Residual LSTM, eXtreme Gradient Boosting (XGBoost), and Light Gradient-Boosting Machine (LightGBM)—using operational data from the National Energy System Operator (NESO) between 2017 and 2024. These models are selected for their complementary strengths: LSTM architectures capture temporal dependencies, while ensemble learning methods like XGBoost and LightGBM efficiently model nonlinear feature interactions. The framework's performance is assessed using RMSE and MAE to identify the most suitable approach for operational forecasting and market participation.

This paper is organised as follows. Section 2 reviews related work on energy trading in smart grids, with a focus on incentive models, optimisation techniques, and recent applications of machine learning. Section 3 outlines the methodology, including data preprocessing, model architecture, and evaluation metrics. Section 4 presents the experimental results, highlighting cost trends and model performance comparisons. Section 5 discusses the implications of accurate inertia cost forecasting for energy trading and system planning. Finally, Section 6 concludes the paper with a summary of key findings and outlines future research directions.

2. Literature Review

Energy trading in smart grids has evolved significantly with the integration of distributed generation, storage systems, and advanced communication technologies [7,8]. While several studies and reports explore smart grid performance and design, comprehensive analyses based on real-world data are still scarce. One notable example is the UK Energy Research Centre's report Building a Resilient UK Energy System, which outlines key challenges and opportunities in smart energy systems [9].

A taxonomy for energy trading models in smart grids—adapted from Aggarwal et al. [10]—categorises research into Incentive-Based, Mathematical, and Simulation approaches, which often overlap in practice, shown in Figure 1. Incentive mechanisms frequently use mathematical formulations and simulation validation to optimise participation and efficiency in energy markets.

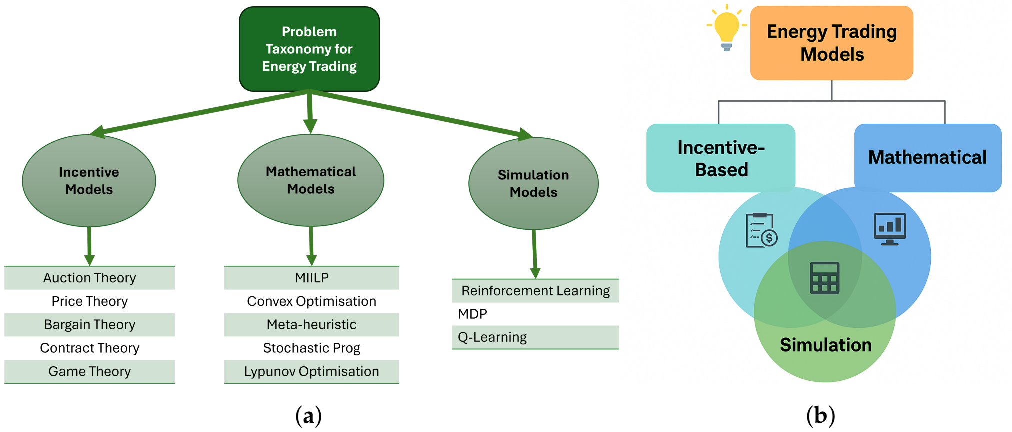


Figure 1. (**a**) Taxonomy of energy trading modelling approaches in smart grids, (**b**) illustrating the interaction between Incentive-Based, Mathematical, and Simulation frameworks.

Incentive mechanisms are central to promoting active participation in decentralised energy systems, motivating prosumers to generate, store, and exchange electricity efficiently. Peer-to-peer (P2P) energy trading leverages these incentives by enabling direct transactions between energy producers and consumers—often households equipped with solar panels or battery storage. Recent studies have shown that integrating P2P trading with real-time pricing and demand response strategies can significantly reduce energy costs for participants by encouraging adaptive consumption and trading behaviour [11]. Furthermore, Islam et al. highlighted that auction-based and pricing-driven frameworks are essential for ensuring fairness and transparency in P2P markets, allowing prosumers to engage and benefit equitably from local energy exchanges [12].

Beyond P2P trading frameworks, incentive-based mechanisms have also been explored in broader market settings to address issues of information asymmetry, fairness, and participation efficiency. These approaches are particularly important in coordinating interactions between heterogeneous market actors, such as small-scale electricity suppliers and consumers, where trust, risk, and optimisation under uncertainty play a key role [13]. Zhang et al. [14] introduced a contract-theoretic trading model between small-scale electricity suppliers (SESs) and energy consumers (ECs) under asymmetric information. Their contract game formulation supported optimal strategies for both short-term and long-term market conditions, validated through simulation. Similarly, Wang et al. [15] proposed a green energy trading market for residential users with solar panels, employing incentive algorithms to address energy supply–demand mismatches and ensure fairness among non-cooperative participants. Timilsina and Silvestri [16] utilised reinforcement learning to develop automated pricing mechanisms for P2P energy trading, enabling sellers to optimise profits while accounting for user behaviour under uncertainty. Their approaches include a Q-learning-based model and a scalable deep Q-network (ProDQN) that incorporates risk sensitivity through a prospect theory-informed loss function.

Mathematical optimisation techniques have also been widely adopted in the energy trading literature. Wu et al. [17] presented a two-layered optimisation framework for local energy trading in smart microgrids, where both consumers and providers interact with a local trading manager (LTM). Their approach dynamically adjusted local trading decisions and pricing strategies, demonstrating effective benefit distribution and profitability through simulation. Zhong et al. [18] extended this line of research by designing auction-based mechanisms for multi-energy districts, incorporating electricity, gas, and heating. Their mechanisms optimised social welfare in both day-ahead and real-time markets and ensured economic efficiency, truthfulness, and individual rationality.

System-level challenges in energy trading have also been discussed in the context of data integration and platform interoperability. Silva et al. [19] reviewed key concerns related to Transmission System Operators (TSOs), Distribution System Operators (DSOs), and third-party aggregators, particularly in handling renewable integration and sensor-based data streams from IoT devices. Their work highlighted the need for robust architecture models to manage increasing system complexity.

With the rise of peer-to-peer (P2P) energy trading, game-theoretic and simulation-based models have gained traction. Yaagoubi et al. [20] proposed a regret matching algorithm that facilitates energy trading between residential users. Their approach enabled sellers to profit from stored renewable energy and allowed buyers to save on energy bills by accessing discounted, clean electricity from neighbours. In a similar vein, simulation-based reinforcement learning techniques have emerged as powerful tools for adaptive decision-making. Lu et al. [21] developed a deep reinforcement learning (RL) framework to enable microgrids (MGs) to autonomously determine trading policies based on predicted renewable output, power demand, and battery status. Their model, validated using real-world data, successfully reduced dependence on central power plants and enhanced overall utility.

Game-theoretic models have been widely applied to improve coordination and participation in P2P energy trading systems. Tushar et al. [22] used a cooperative Stackelberg game where the grid sets prices to encourage prosumer activity during peak hours, with user utility modelled through logarithmic functions. In [23], a coalition game approach supports stable group formation by evaluating the collective value of peer coalitions, promoting fair energy sharing. Separately, [24] introduced an auction-based framework in which participants submit bids, and optimal allocations are computed to balance efficiency and individual incentives.

Existing research on ancillary service cost forecasting primarily addresses frequency response, reserve, and balancing markets. However, relatively few studies examine inertia-specific costs associated with synthetic or fast frequency response services. For example, the game-theoretical optimisation framework proposed in [25] highlights the role of predictive intelligence in market pricing and operator decision-making. Our study extends this perspective to the domain of system inertia, providing a data-driven basis for anticipating future cost dynamics in evolving low-inertia grids.

While a wide range of approaches has been explored for energy trading in smart grids, few studies explicitly address the influence of system inertia on market dynamics. Maintaining grid frequency close to the national standard of 50 Hz is critical for operational stability and is strongly dependent on available inertia. Traditionally, this inertia was provided by large rotating generators; however, the increasing integration of renewable energy sources and decentralised technologies—such as vehicle-to-grid (V2G) systems—has significantly reduced these conventional sources.

This shift has introduced new challenges in maintaining grid balance, highlighting the growing importance of inertia forecasting [26]. Accurate prediction of system inertia is essential not only for stability and control but also for informing strategic decisions in energy trading and pricing. Despite its importance, this remains an underexplored area in the existing literature, particularly in terms of real-world data analysis [27].

Recent advancements in short-term forecasting highlight the growing convergence of traditional time-series and hybrid machine learning approaches. Ali et al. [28] demonstrated that integrating statistical and data-driven models can substantially improve short-term load prediction accuracy under renewable variability. Their findings reinforce the motivation for the present study, which similarly explores hybrid deep and ensemble learning frameworks to forecast inertia-related costs in a dynamic smart grid context. In this study, we address this gap by applying machine learning techniques to real-time system inertia cost data obtained from the National Energy System Operator (NESO). Our aim is to forecast inertia-related costs and uncover patterns that can support more informed and resilient energy trading strategies within future smart grids. Through this data-driven approach, we contribute valuable insights toward market optimisation in an increasingly complex and dynamic energy landscape.

3. Materials and Methods

Figure 2 illustrates the proposed methodological framework for system inertia cost forecasting. It begins with data acquisition from NESO, followed by preprocessing and feature engineering (including temporal and lag features). Machine learning models—LSTM, Residual LSTM, XGBoost, and LightGBM—are trained and validated using RMSE and MAE. The resulting forecasts are analysed to support strategic decision-making for market trading and grid stability management.

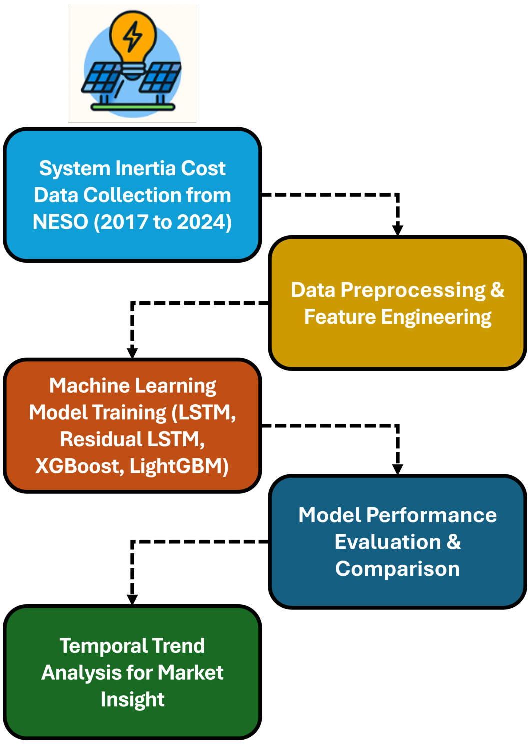


Figure 2. The proposed research framework for system inertia cost forecasting, comprising five stages: (1) data collection from NESO; (2) preprocessing; (3) model training; (4) model validation; and (5) temporal trend analysis and interpretation for operational and market insights.

3.1. Data Description and Analysis

This dataset was collected from the NESO website over 8 years. It provides estimated average daily prices (in GBP per GVA·s) for offers in the Balancing Mechanism (BM) that were tagged as System Inertia actions. The tagging system is intended to identify the most likely primary reason for each BM action, but it is not perfect, as some actions may serve multiple purposes simultaneously. The NESO tagging system occasionally assigns multiple operational purposes to a single balancing mechanism action. To minimise noise introduced by such overlap, this study aggregates all tagged transactions into daily averages of total cost per GVA·s. This aggregation smooths out mixed-purpose variations at the individual offer level. Days with missing or anomalous data were excluded to ensure consistent representation of inertia-related balancing activity.

The data are reported at daily intervals. For each day, the average price per GVA·s is calculated as follows:

- Sum the total cost of all BM offers tagged as System Inertia.
- Sum the total estimated inertia shortfall (in GVA·s) that these offers addressed that day.
- Divide the total cost by the total inertia volume to obtain the average daily cost per GVA·s.

Figure 3 presents the annual average trend in total System Inertia Cost in Great Britain, based on data obtained from NESO for the period 2017 to 2024. This cost represents the expenditure incurred to ensure adequate system inertia and frequency stability across the grid. The trend shows a steady rise in costs from 2017 to 2021, with a sharp increase observed between 2020 and 2022. This surge likely reflects growing reliance on flexible, inertia-providing ancillary services as traditional synchronous generators are displaced by renewable sources. The cost peaked in 2023, marking the highest expenditure on inertia services during the study period. However, a noticeable decline is seen in 2024, possibly due to improved system operability measures, changes in procurement strategies, or the adoption of more cost-efficient technologies.

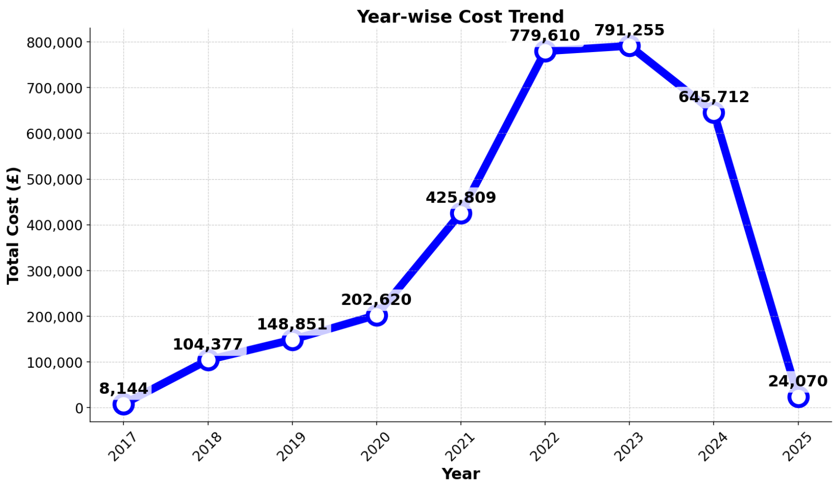


Figure 3. Yearly average inertia costing trends for the consecutive eight years.

3.2. Feature Engineering and Preparation

The raw dataset provided by the NESO contains daily estimates of system inertia cost (in GBP/GVA·s). To enable predictive modelling, several temporal and statistical features were derived from the time index and cost series, as implemented in the analysis scripts:

- **Temporal features:** day of week (0–6), month (1–12), and calendar day (1–31) were extracted from the settlement date to represent cyclic and seasonal variations.
- **Lag features:** for ensemble learning models (XGBoost and LightGBM), 7 lag variables (lag_1 *to* lag_7) were generated to capture persistence and short-term dependencies in daily inertia cost.
- **Rolling mean:** a 7-day moving average (rolling-mean-7) was included to smooth high-frequency fluctuations and capture weekly market trends.
- **Sequence windows:** for the LSTM-based models, sequential windows of 30 consecutive days were constructed, with the network trained to predict the next day's cost. Each sequence incorporated the four features [Cost, dayofweek, month, day].
- **Scaling:** all features were normalised to the range [0, 1] using MinMaxScaler to stabilise training and ensure comparability across feature dimensions.

- **Missing data handling:** occasional missing cost entries were interpolated before feature generation to ensure continuity in lagged and sequential inputs.

This preprocessing yielded a structured feature matrix used uniformly across models. Deep learning models utilised sequential (3D) data arrays, while ensemble models operated on flattened tabular features with 7-day lag and rolling components.

3.3. Machine Learning Models

This study employs a combination of deep learning and ensemble learning techniques to forecast system inertia service costs. The selected models are designed to capture both the sequential dependencies and the nonlinear feature interactions inherent in the cost time series data.

3.3.1. Deep Learning Models

LSTM with Temporal Features: Long Short-Term Memory (LSTM) networks are a class of recurrent neural networks specifically designed to model temporal dependencies. In this work, the LSTM input is augmented with engineered temporal features such as day of the week, month, and day of the month to capture periodic cost patterns.

The LSTM unit relies on a series of gating mechanisms to regulate the flow of information:

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (1)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (2)$$

$$\tilde{C}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (3)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (4)$$

$$h_t = o_t \odot \tanh(C_t) \quad (5)$$

The hidden state h_t is passed through a fully connected layer to predict the cost at each time step. The base LSTM architecture consists of two stacked LSTM layers with 64 hidden units each, ReLU activation, and a dropout rate of 0.2 to prevent overfitting, followed by a dense output layer. Training uses the Adam optimiser (learning rate = 0.001) and Mean Squared Error (MSE) as the loss function.

LSTM with Residual Modelling: To improve predictive accuracy, a residual learning framework is employed. A base LSTM model first generates the primary prediction $\hat{y}_t^{\text{LSTM}}$, and the residual error is computed as

$$r_t = y_t - \hat{y}_t^{\text{LSTM}} \quad (6)$$

A second LSTM model is trained to learn the residual sequence, producing $\hat{r}_t$. The final prediction is obtained by combining both outputs:

$$\hat{y}_t = \hat{y}_t^{\text{LSTM}} + \hat{r}_t \quad (7)$$

The Residual LSTM employs three layers with 128 hidden units each, trained under the same conditions. The rationale for residual modelling is that inertia cost sequences exhibit irregular volatility and occasional step changes. The first LSTM captures long-term patterns, while the second network learns the residual error series—effectively modelling transient fluctuations that the base network underestimates. This two-stage approach improves tracking of dynamic cost variations without increasing network depth excessively.

The use of LSTM architectures is motivated by the temporal nature of system inertia costs, which exhibit sequential dependencies and recurring fluctuations. Empirical autocorrelation analysis of the NESO dataset revealed significant day-to-day persistence and weekly periodicity, suggesting that memory-based sequence models are well-suited for capturing these dynamics.

3.3.2. Ensemble Learning Models

XGBoost: XGBoost is a gradient boosting framework that constructs an ensemble of decision trees in a stage-wise manner. At each iteration t, the model minimises a regularised objective function:

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} \left[l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) \right] + \Omega(f_t) \tag{8}$$

Here, f_t represents the newly added tree, and $\Omega(f_t) = \gamma T + \frac{1}{2}\lambda \|w\|^2$ penalises model complexity based on the number of leaves T and the leaf weights w. The optimisation uses a second-order Taylor expansion of the loss function.

LightGBM: LightGBM is another gradient boosting framework optimised for speed and scalability. Unlike traditional level-wise tree growth, LightGBM employs a leaf-wise strategy and integrates two key techniques: (1) Gradient-based One-Side Sampling (GOSS), which prioritises samples with large gradients; (2) Exclusive Feature Bundling (EFB), which reduces feature dimensionality by combining mutually exclusive features.

The gain from splitting a node is calculated as

$$\text{Gain} = \frac{1}{2} \left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda} \right] - \gamma \tag{9}$$

where G and H denote the sum of gradients and Hessians for the left (L) and right (R) child nodes, respectively. This mechanism enables LightGBM to demonstrate high predictive performance while maintaining computational efficiency.

For LightGBM, hyperparameter tuning was performed using GridSearchCV with a parameter grid of $num_{leaves} = [20, 40, 60], max_{depth} = [5, 10, 15], learning_{rate} = [0.01, 0.05, 0.1]$, $n_{estimators} = [100, 200, 300], and\ min_{data-in-leaf} = [10, 20, 30]$. A TimeSeriesSplit with 5 folds was used instead of standard k-fold cross-validation to preserve temporal dependencies. Optimal parameters achieved balance between low bias and high stability.

3.4. Model Validation—Evaluation Metrics

To assess the predictive performance of the developed models, we employed two standard regression evaluation metrics: Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE).

Root Mean Squared Error (RMSE): The RMSE measures the square root of the average squared differences between predicted and actual values. It penalises larger errors more heavily and is given by

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{10}$$

where y_i is the actual value, $\hat{y}_i$ is the predicted value, and n is the number of data points. Lower RMSE indicates better model performance.

Mean Absolute Error (MAE): The MAE quantifies the average absolute deviation between predicted and actual values:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{11}$$

Unlike RMSE, MAE gives equal weight to all errors, making it more robust to outliers.

Each model was configured using architecture and training parameters consistent with its design principles and prior literature. Deep learning models employed moderate network sizes (64–128 units) and limited epochs to prevent overfitting given the dataset length, whereas ensemble models (XGBoost and LightGBM) were optimised via grid search to balance depth and regularisation. This ensures that each model operates near its empirically optimal configuration, allowing meaningful comparison of predictive accuracy rather than computational efficiency.

4. Results

4.1. Temporal Trends Analysis

This temporal analysis provides a critical foundation for the machine learning-based forecasting models developed in this study, facilitating the prediction of future inertia-related costs and supporting strategic decision-making in smart grid energy trading and system planning.

Figure 4 presents a year-wise distribution of daily inertia costs using box plots for 2017–2024. This visualisation captures both the central tendency (median) and spread (interquartile range and outliers) of daily cost data across each year. It clearly demonstrates the evolution of the system's operational cost burden over time. A distinct increase in both median costs and variability is observed from 2020 onwards, peaking in 2022, before tapering off slightly in 2023 and 2024. The high number of outliers in recent years underscores the frequent occurrence of extreme cost events, which reflect operational volatility and potentially strained grid conditions.

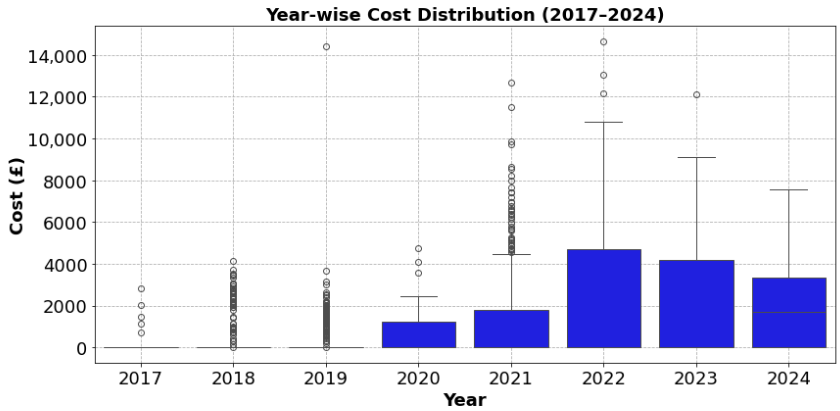


Figure 4. Year-wise distribution of daily inertia service costs from 2017 to 2024.

Figure 5 meanwhile provides a complementary temporal perspective by plotting the daily cost trends across the same period. This allows for a more granular view of how those outliers and cost escalations manifest over time—highlighting temporal clusters, seasonal spikes, or the impact of operational events that a box plot alone cannot reveal. For instance, the increased frequency and amplitude of spikes in 2021 and 2022 align with the widening spread seen in the corresponding box plots. Similarly, the visual flattening of cost spikes in 2024 corresponds to a narrowing of the box and fewer extreme outliers.

Figure 4 quantifies the statistical characteristics of cost distributions across years. Figure 5 contextualises these statistics by illustrating the timing, persistence, and temporal patterns of cost events. This combined analysis is important for system operators and policy-makers, as it enables both strategic planning (via trend observation) and risk management (via distribution analysis). It not only validates the potential impact of policy interventions and market mechanisms introduced in later years—evidenced by statistical compression and reduced spike volatility—but also highlights the dynamic and evolving nature of inertia costs. The temporal disaggregation reveals critical shifts in operational conditions and reinforces the need for predictive models. These evolving patterns underscore the complexity of maintaining frequency stability amid the increase in renewable integration and motivate the use of machine learning techniques for cost forecasting, anomaly detection, and smart grid optimisation.

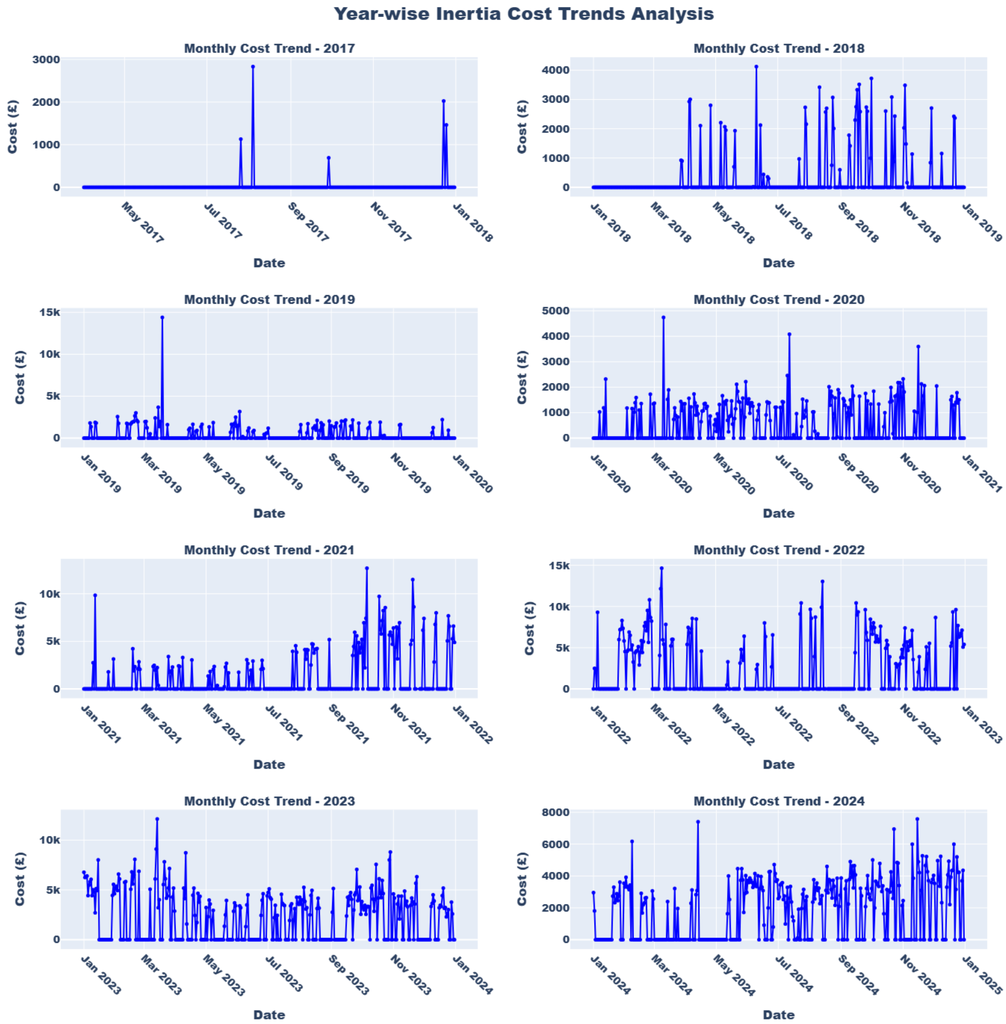
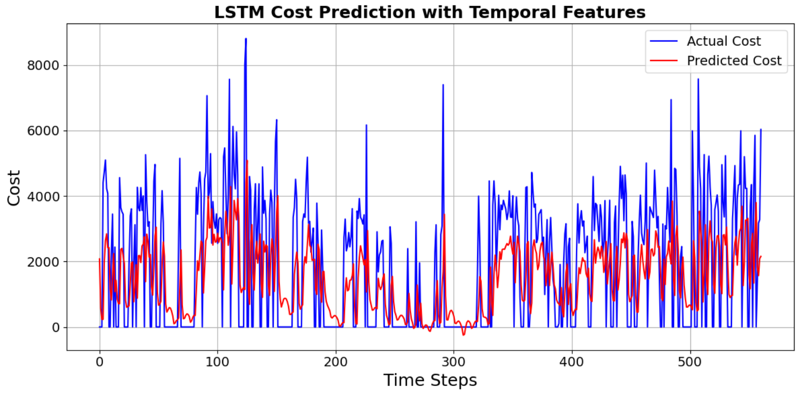


Figure 5. Monthly inertia service cost trends from 2017 to 2024 as recorded by the NESO. Each subplot represents a full calendar year, capturing daily cost variations associated with procuring inertia to support grid stability.

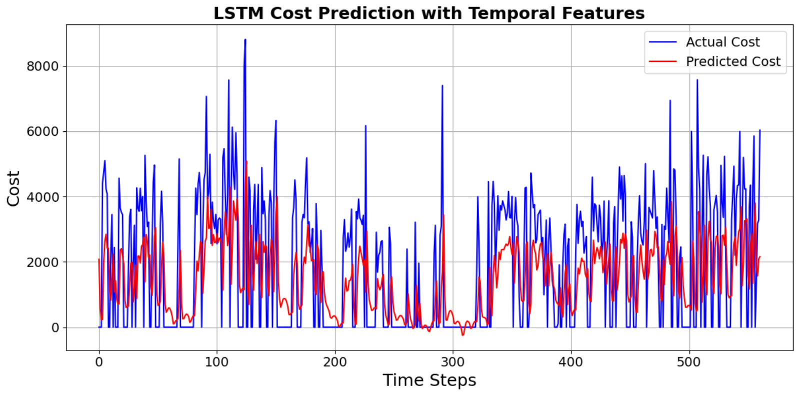
4.2. Model Comparison Analysis

The forecasting pipeline includes two deep learning models and two ensemble learning models, each contributing distinct advantages to cost prediction. Figure 6 presents the predicted versus actual system inertia cost trends for four forecasting models: (a) LSTM with temporal features, (b) LSTM with residual modelling, (c) XGBoost, and (d) LightGBM.

The LSTM with Temporal Features model uses two stacked LSTM layers (64 units each) with dropout to prevent overfitting. Trained for 20 epochs with a batch size of 32 on an 80/20 train–test split, it captures overall cost trends but tends to underestimate sharp spikes, producing smoother outputs than the actual data. Whereas to enhance accuracy, a Residual LSTM architecture is applied using a deeper LSTM with three layers of 128 units each. Also trained for 20 epochs, this two-stage setup captures more complex temporal patterns missed by the base model.

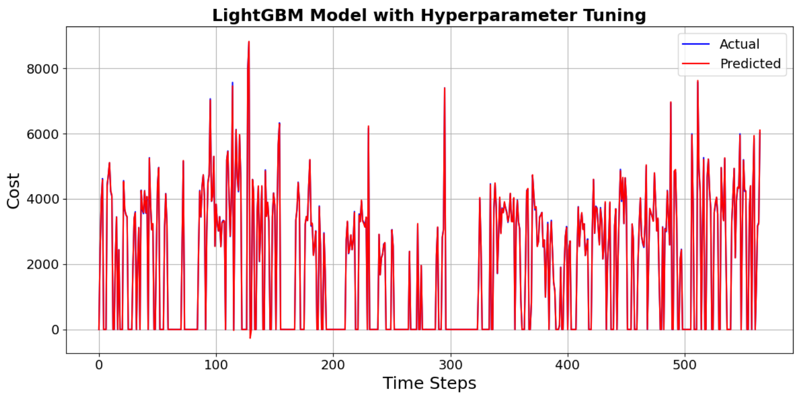


(**a**)



(**b**)

(**c**)



(**d**)

Figure 6. The actual vs prediction outcome for inertia cost forecasting analysis and comparison for four different models: (**a**) LSTM with temporal features. (**b**) LSTM with residuals. (**c**) XGBoost with lag features. (**d**) LightGBM with parameter hypertunning.

XGBoost, using lag features (1–7 days) and a rolling mean, is the fastest to train. It is configured with 100 estimators, a max depth of 5, and a learning rate of 0.01, demonstrating efficient results with moderate accuracy. LightGBM, with the same feature set, undergoes extensive hyperparameter tuning via GridSearchCV. It outperforms all other models in predictive accuracy, benefiting from optimised boosting and regularisation.

This combination of deep learning for capturing temporal dynamics and ensemble models for fast and robust prediction is a crucial aspect of the forecasting strategy, enabling both accuracy and efficiency in modelling highly variable cost data. Residual LSTM improves upon the base LSTM by learning from its prediction errors, offering better tracking of mid-range fluctuations, though it still underestimates extreme cost spikes. XGBoost performs well in stable regions but lacks responsiveness to sharp changes, often under-predicting peaks. LightGBM shows the closest alignment with actual values, effectively capturing both baseline trends and volatility, leading to the best overall performance. While LSTM models leverage temporal patterns, LightGBM emerges as the most reliable model for forecasting system inertia costs in this study.

Figure 7 compares the performance of four models—LSTM, Residual LSTM, XGBoost, and LightGBM—in predicting system inertia costs using key metrics: Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). Both LSTM models yield the highest RMSE and MAE, indicating a relatively poor fit to the data. XGBoost significantly reduces prediction error, while LightGBM achieves the lowest RMSE and MAE, highlighting its superior accuracy.

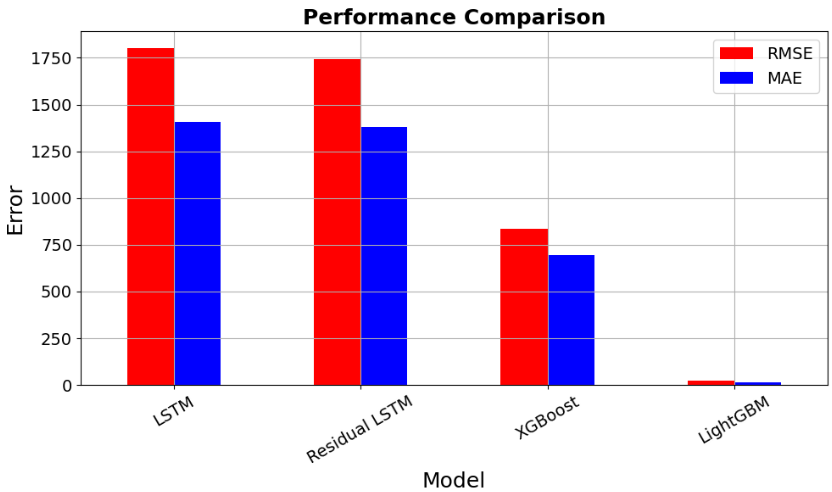


Figure 7. Performance matrices comparison for four different models.

All models were trained using the same 80/20 train–test split, loss function, and evaluation metrics to ensure consistent benchmarking. The residual LSTM employed a larger architecture than the base model to capture secondary temporal dependencies rather than to increase computational capacity. Although this setup allows fair comparison of forecasting performance, it remains sensitive to the chosen data partition; future work will address this by averaging results across multiple runs to improve statistical confidence. The underperformance of LSTM models during extreme cost spikes is attributed to their smooth activation functions (tanh and sigmoid), which reduce sensitivity to abrupt cost changes. Ensemble methods, by contrast, better capture such discontinuities through non-linear decision boundaries.

Additionally, to improve interpretability of the ensemble learning approach, feature importance analysis was conducted on the optimised LightGBM model. Figure 8 presents the top ten features ranked by their contribution to model performance. The results show that recent cost values, short-term lag features, and weekly rolling averages are dominant predictors of system inertia cost. Temporal calendar variables such as day and month also exhibited notable influence, indicating the presence of periodic and seasonal effects. These findings confirm that short-term dynamics and cyclical market behaviour jointly determine the fluctuations in system inertia service costs.

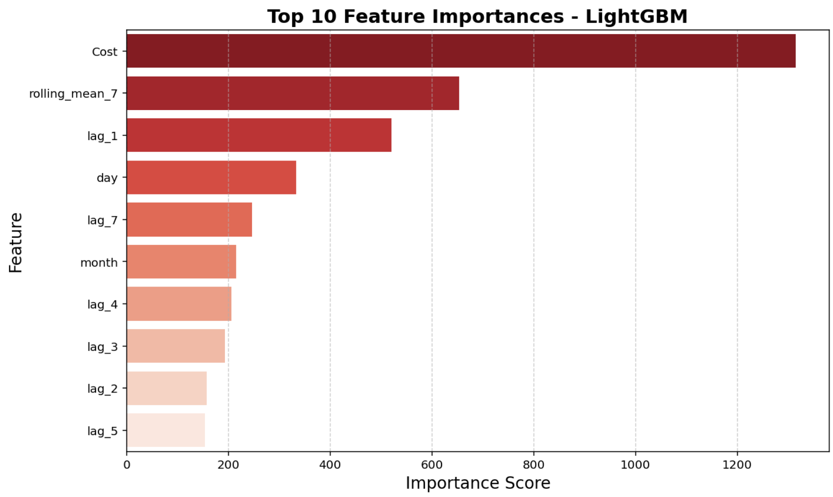


Figure 8. Top 10 feature importance scores derived from the LightGBM model.

5. Discussion

The proposed framework contributes beyond existing energy forecasting literature by specifically addressing system inertia cost prediction—a variable rarely modelled despite its rising impact on balancing markets. Unlike prior studies focusing solely on frequency or demand prediction, this work integrates temporal (LSTM-based) and ensemble (boosting-based) models within a unified evaluation framework. This comparative approach demonstrates that ensemble learning, particularly LightGBM, can provide interpretable, accurate, and computationally efficient forecasts for operational use in real-time grid management.

5.1. Model Interpretability and Forecasting Effectiveness

Among the forecasting models evaluated, LightGBM consistently outperforms others due to its efficient tree-building approach and ability to model complex, nonlinear data patterns. Unlike XGBoost's level-wise tree growth, LightGBM employs a leaf-wise strategy with depth constraints, enabling it to prioritise the most informative feature splits and capture sharp cost fluctuations more effectively. Its use of histogram-based training and native handling of missing values further enhances computational efficiency, while built-in regularisation mechanisms help mitigate overfitting. These characteristics make LightGBM particularly well-suited for forecasting high-variance time series such as system inertia costs.

The empirical evidence of cost autocorrelation supports the use of temporal models, as inertia procurement costs exhibit strong sequential dependencies driven by daily operational patterns and renewable output variability. LSTM-based models effectively capture these structured temporal dynamics, while ensemble methods such as LightGBM complement them by modelling complex nonlinear feature interactions. The resulting feature importance analysis demonstrates that LightGBM identifies interpretable and physically meaningful relationships within the data—where short-term historical behaviour and temporal patterns are dominant predictors. Together, these findings confirm that machine learning models can provide both accurate forecasts and transparent insights into the cost drivers influencing grid stability and market operations.

More broadly, the findings highlight the strength of machine learning models—particularly ensemble methods like LightGBM—in capturing temporal and non-linear dynamics in inertia cost data. Their balance of predictive accuracy, computational efficiency, and interpretability positions them as practical tools for integration into near-real-time operational forecasting pipelines. Additionally, models such as the residual-enhanced LSTM, while less accurate in extreme scenarios, demonstrate utility in capturing mid-range cost fluctuations, which are critical during transitional grid conditions. These complementary strengths underscore the value of hybrid modelling strategies in managing the growing complexity of power systems.

5.2. Operational Implications for Grid Management

System inertia remains a cornerstone of frequency stability in modern power grids. As traditional synchronous generators are phased out and replaced by renewable and inverter-based sources, the natural inertia of the grid diminishes. To maintain operability, system operators now rely on ancillary services to procure synthetic or fast-responding inertia, often at a substantial and variable cost. These expenses are increasingly visible in reserve and balancing markets, affecting the cost-efficiency and responsiveness of the grid.

The findings of this study demonstrate that accurate forecasting of inertia-related costs can directly enhance grid management. For system operators, the ability to anticipate cost fluctuations supports proactive procurement strategies, better reserve planning, and optimal scheduling of inertia-contributing resources. It also allows for more targeted use of flexible technologies like battery storage or demand-side response, particularly when high-frequency volatility is expected.

Ultimately, integrating inertia forecasting into grid control frameworks can lead to more informed operational decisions, contributing to both cost savings and improved system reliability.

Forecasting accuracy directly influences operational and economic outcomes. Underestimating future inertia costs may result in inadequate service procurement, compromising frequency stability margins. Conversely, overestimation can lead to inflated expenditure on ancillary services. Thus, even moderate improvements in forecasting precision translate into tangible savings and reduced system risk for the National Energy System Operator.

5.3. Strategic Benefits for Market Participation and Policy

In addition to its operational advantages, inertia cost forecasting provides meaningful strategic value for market participants and policymakers. For energy traders, aggregators, and service providers, predictive insights into inertia pricing enable more effective bidding in co-optimised energy and ancillary markets. By anticipating volatility, participants can adjust trading strategies, hedge against cost spikes, and align portfolios with system conditions—improving both competitiveness and profitability.

From a market design perspective, the ability to forecast inertia-related costs could inform the development of more dynamic, responsive pricing schemes. Such mechanisms would reflect real-time system needs more accurately and incentivise distributed energy resources (DERs) to participate in frequency support services. This creates new value streams for flexible assets and helps ensure that grid-supportive behaviours are appropriately compensated.

Furthermore, these insights have policy implications. As the energy system continues to decentralise and decarbonise, regulators and system planners must consider how to maintain stability in the absence of conventional inertia. Incorporating predictive tools into planning and procurement strategies can help bridge this gap. Forecast-informed procurement frameworks could be used to prioritise cost-effective inertia services, reduce over-reliance on fossil-based reserves, and support investment in scalable, non-synchronous solutions.

While this study shows promising results, future work should address challenges around model explainability, data reliability, and integration with existing control and trading platforms. Robust validation processes and uncertainty quantification will be essential for deploying ML-based forecasting tools in high-stakes operational contexts.

6. Conclusions

This study introduced a data-driven framework for forecasting system inertia costs in Great Britain's electricity grid using machine learning. Four models—LSTM, Residual LSTM, XGBoost, and LightGBM—were evaluated on NESO's 2017–2024 dataset to identify the most effective forecasting approach. Results show that LightGBM achieved the lowest RMSE and MAE, outperforming both deep learning and alternative ensemble methods.

The research contributes a clear comparative methodology that connects temporal sequence learning with gradient-boosted ensemble modelling, offering a practical tool for system operators and market participants. Accurate inertia cost forecasting enables proactive procurement planning, enhances frequency control, and supports market strategies for ancillary services.

Accurate inertia cost forecasting is not just a technical task; it provides strategic benefits across multiple layers of the energy system. For system operators, these predictions can improve procurement decisions, enhance frequency control planning, and reduce dependence on costly reserve services. For market participants, such forecasts support smarter bidding, hedging strategies, and real-time portfolio optimisation in co-optimised energy and ancillary markets. Moreover, from a policy perspective, predictive tools like these can inform market design by enabling inertia-aware pricing schemes and by creating new value streams for non-synchronous resources such as batteries and inverter-based generation.

As electricity grids continue to decarbonise and decentralise, the ability to anticipate system dynamics—particularly around frequency stability and inertia—will be critical for maintaining operational reliability and market efficiency. Machine learning tools, when properly validated and integrated, have the potential to become key enablers in this transition.

Limitations and Future Work

While this study provides important insights, several limitations should be addressed in future research. The models rely solely on historical cost data, which may not fully capture future market structures or regulatory interventions. Expanding the framework to higher temporal resolutions—such as intra-day or real-time forecasting—and incorporating exogenous variables like grid frequency deviations, weather conditions, and interconnection flows could enhance predictive accuracy and operational relevance.

Future work should also integrate uncertainty quantification to support risk-aware decision-making and validate the proposed models across different national or regional energy systems to test generalisability. Collaboration with system operators will be essential to ensure that model outputs align with practical operational needs and can be embedded within real-world control and market dispatch systems.

Author Contributions: Conceptualization, M.D.; methodology, M.D.; software, M.D. and S.P.R.; validation, M.D. and P.P.; formal analysis, M.D., S.P.R. and P.P.; investigation, M.D.; resources, M.D. and S.P.R.; data curation, M.D.; writing—original draft preparation, M.D.; writing—review and editing, S.P.R. and P.P.; visualization, M.D., S.P.R. and P.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data used for this study are publicly available at https://www.neso.energy/data-portal (accessed on 4 June 2025) [29].

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Pollitt, M.G. Lessons from the history of independent system operators in the energy sector. *Energy Policy* **2012**, *47*, 32–48. [CrossRef]
2. Dey, M.; Rana, S.P.; Wylie, J.; Simmons, C.V.; Dudley, S. Detecting power grid frequency events from µPMU voltage phasor data using machine learning. *IET Conf. Proc.* **2022**, *2022*, 125–129. [CrossRef]
3. National Energy System Operator. *Operability Strategy Report 2022* ; National Energy System Operator: London, UK, 2022.
4. Mehigan, L.; Al Kez, D.; Collins, S.; Foley, A.; Ó'Gallachóir, B.; Deane, P. Renewables in the European power system and the impact on system rotational inertia. *Energy* **2020**, *203*, 117776. [CrossRef]
5. Anaya, K.; Pollitt, M. *Regulating the Electricity System Operator: Lessons for Great Britain from Around the World*; EPRG Working Paper 1718; Energy Policy Research Group, University of Cambridge: Cambridge, UK, 2017.
6. Zhou, L.; Wang, H.; Wang, Y.; Zhang, H.; Li, W.; Li, R. Evaluation of the impact of inertia on system operation cost. *Front. Energy Res.* **2023**, *11*, 1118349. [CrossRef]
7. Cao, J.; Yang, M. Energy internet–towards smart grid 2.0. In Proceedings of the 2013 Fourth International Conference on Networking and Distributed Computing, Los Angeles, CA, USA, 21–24 December 2013; IEEE: Piscataway, NJ, USA, 2013; pp. 105–110.
8. Diggikar, S.; Patil, A.; Katkar, S.S.; Samad, K. Machine learning-based inertia estimation in power systems: A review of methods and challenges. *Energy Inform.* **2025**, *8*, 57. [CrossRef]
9. Chaudry, M.; Ekins, P.; Ramachandran, K.; Shakoor, A.; Skea, J.; Strbac, G.; Wang, X.; Whitaker, J. *Building a Resilient UK Energy System: Research Report*; UKERC: London, UK, 2011.
10. Aggarwal, S.; Kumar, N.; Tanwar, S.; Alazab, M. A survey on energy trading in the smart grid: Taxonomy, research challenges and solutions. *IEEE Access* **2021**, *9*, 116231–116253. [CrossRef]
11. Shan, S.; Yang, S.; Becerra, V.; Deng, J.; Li, H. A case study of existing peer-to-peer energy trading platforms: Calling for integrated platform features. *Sustainability* **2023**, *15*, 16284. [CrossRef]
12. Islam, S.N. A review of peer-to-peer energy trading markets: Enabling models and technologies. *Energies* **2024**, *17*, 1702. [CrossRef]
13. Kashyap, S.; Schaffer, C.; Fischer, T.; Zauner, M.; Fischer, F.; Kurz, M.; Hartner, G.; Grünberger, S.; Hödl, O. Empowering Energy Communities and P2P Energy Sharing: A Novel End-to-End Ecosystem for Planning, Deployment, and Operation. *ACM SIGEnergy Energy Inform. Rev.* **2025**, *4*, 238–244. [CrossRef]
14. Zhang, B.; Jiang, C.; Yu, J.L.; Han, Z. A contract game for direct energy trading in smart grid. *IEEE Trans. Smart Grid* **2016**, *9*, 2873–2884. [CrossRef]
15. Wang, H.; Zhang, J.X.; Li, F. Incentive mechanisms to enable fair renewable energy trade in smart grids. In Proceedings of the 2015 Sixth International Green and Sustainable Computing Conference (IGSC), Las Vegas, NV, USA, 14–16 December 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 1–6.
16. Timilsina, A.; Silvestri, S. P2p energy trading through prospect theory, differential evolution, and reinforcement learning. *ACM Trans. Evol. Learn.* **2023**, *3*, 11. [CrossRef]
17. Wu, Y.; Tan, X.; Qian, L.; Tsang, D.H. Optimal management of local energy trading in future smart microgrid via pricing. In Proceedings of the 2015 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Hong Kong, China, 26 April–1 May 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 570–575.
18. Zhong, W.; Xie, K.; Liu, Y.; Yang, C.; Xie, S. Auction mechanisms for energy trading in multi-energy systems. *IEEE Trans. Ind. Inform.* **2017**, *14*, 1511–1521. [CrossRef]
19. Silva, F.; O'Leidhin, E.; Tahir, F.; Mould, K.; O'Regan, B. System integration and data models to support smart grids energy trading. In Proceedings of the European Conference on Renewable Energy Systems (ECRES), Istanbul, Turkey, 21–23 April 2021.
20. Yaagoubi, N.; Mouftah, H.T. Energy Trading In the smart grid: A game theoretic approach. In Proceedings of the 2015 IEEE International Conference on Smart Energy Grid Engineering (SEGE), Oshawa, ON, Canada, 17–19 August 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 1–6.
21. Lu, X.; Xiao, X.; Xiao, L.; Dai, C.; Peng, M.; Poor, H.V. Reinforcement learning-based microgrid energy trading with a reduced power plant schedule. *IEEE Internet Things J.* **2019**, *6*, 10728–10737. [CrossRef]
22. Tushar, W.; Saha, T.K.; Yuen, C.; Morstyn, T.; Poor, H.V.; Bean, R. Grid influenced peer-to-peer energy trading. *IEEE Trans. Smart Grid* **2019**, *11*, 1407–1418. [CrossRef]

23. Tushar, W.; Saha, T.K.; Yuen, C.; Liddell, P.; Bean, R.; Poor, H.V. Peer-to-peer energy trading with sustainable user participation: A game theoretic approach. *IEEE Access* **2018**, *6*, 62932–62943. [CrossRef]
24. Tushar, W.; Yuen, C.; Mohsenian-Rad, H.; Saha, T.; Poor, H.V.; Wood, K.L. Transforming energy networks via peer-to-peer energy trading: The potential of game-theoretic approaches. *IEEE Signal Process. Mag.* **2018**, *35*, 90–111. [CrossRef]
25. Cheng, L.; Huang, P.; Zhang, M.; Yang, R.; Wang, Y. Optimizing electricity markets through game-theoretical methods: Strategic and policy implications for power purchasing and generation enterprises. *Mathematics* **2025**, *13*, 373. [CrossRef]
26. Dey, M.; Rana, S.P.; Simmons, C.V.; Dudley, S. Solar farm voltage anomaly detection using high-resolution μPMU data-driven unsupervised machine learning. *Appl. Energy* **2021**, *303*, 117656. [CrossRef]
27. Maitreyee, D.; Rana, S.P. High-Resolution Electrical Measurement Data Processing. G.B. Patent Application No. GB2016025.5A, 21 December 2022.
28. Ali, S.; Bogarra, S.; Riaz, M.N.; Phyo, P.P.; Flynn, D.; Taha, A. From time-series to hybrid models: Advancements in short-term load forecasting embracing smart grid paradigm. *Appl. Sci.* **2024**, *14*, 4442. [CrossRef]
29. National Energy System Operator (NESO). Available online: https://www.neso.energy/data-portal (accessed on 4 June 2025).