Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 4 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

Energy Cost Minimization in Data Centers Using Hybrid CNN–LSTM-Based Workload and Price Prediction under Deregulated Power Markets

Imran Rasheed
Department of Computer Science, University of Engineering & Technology Peshawar 25000, Pakistan

Muhammad Imran Khan Khalil*
Department of Computer Science, University of Engineering & Technology Peshawar 25000, Pakistan Email: imrankhalil@uetpeshawar.edu.pk

Muhammad Naeem Khan
Department of Computer Science, University of Engineering & Technology Peshawar 25000, Pakistan

Alauddin
Department of Computer Science, University of Engineering & Technology Peshawar 25000, Pakistan

ABSTRACT

The growth of cloud computing infrastructures has resulted in significant energy consumption and cost of operation for data centers, especially in deregulated electricity markets with high price volatility and uncertainty. We proposed a hybrid deep learning system combining Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks for predicting the workload and electricity price jointly, which allows the cost-aware and adaptive optimization of electricity. The model utilizes very large and high-resolution time-series data that is directly extracted from realistic workload patterns and deregulated price signals in the market, capturing both temporal dependencies and stochastic price fluctuations. CNN captures short-term workload bursts and local variations, while LSTM layers capture the long-term temporal correlations, which enhance the predictive performance. A dynamic price-aware scheduling mechanism is developed that combines the predicted workload and price signals for making optimal energy consumption decisions in a discrete-time simulation environment. The proposed framework is experimentally validated using extensive data sets and compared to the baseline models, such as LSTM, GRU, and ARIMA, which show that the proposed method has significantly reduced the forecasting error and operational cost. In particular, the model shows significant gains in energy cost savings, peak load reduction, energy efficiency, and high service reliability. Results from statistical validation are significant ($p<0.01$), and robustness analysis with various stress conditions, such as workload fluctuations and price fluctuations, proves the stability. In addition, multi-objective evaluation by using Pareto analysis demonstrates that the proposed method has a better balance

Online ISSN 3007-3197 Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

between accuracy, cost efficiency, and stability. The results demonstrate that the proposed CNN–LSTM framework is a scalable, reliable and cost-effective solution for intelligent energy management in modern data centers under dynamic market conditions.

Keywords: Data Center Energy Optimization, CNN–LSTM, Workload Prediction, Electricity Price Forecasting, Deregulated Power Markets, Cloud Computing.

Introduction

The energy footprint is impactful because of the massive amounts of data produced and processed by energy-intensive modern data centers. Modern data centers have experienced significant growth in energy usage with the emergence of cloud computing services, artificial intelligence services and large digital infrastructures [1]. It is estimated that data centers are a significant part of the global electricity consumption, and that this demand is expected to grow in the future, due to increased edge computing, IoT ecosystems, and workloads with high computing requirements [17]. For this growing energy requirement, the operational cost is increased, and environmental sustainability and emissions of CO_2 are under strain [25]. At the same time, electricity markets have been deregulated and dynamic pricing is practiced, with electricity prices fluctuating continually according to the demand and supply [10]. As they present chances to optimize costs, they also create a great deal of uncertainty and volatility, making energy management strategies more difficult in data centers. Therefore, there is an urgent need for intelligent, predictive and adaptive systems to model the dynamics of computational workloads and electricity price simultaneously [22]. With the developments of deep learning and its hybrid architecture with the ability to learn both spatial and temporal dependence, it is a promising way toward energy-efficient and cost-aware data center operations in such complex environments.
Although there has been great progress in data center energy management, current methods have notable shortcomings in deregulated power markets. Conventional energy optimization techniques use either static or heuristic-based scheduling policies, which do not adapt to high variations in workloads and the price of electricity [20]. Furthermore, classical forecasting techniques such as time-series models are unable to account for the nonlinear relationships and multi-scale temporal relations found in real-life workloads at data centers [22]. Although machine learning techniques have been applied to improve the accuracy of the prediction, most of the models focus on the prediction of workload only and do not consider electricity price prediction as a problem, because the interdependence of the two problems is ignored in the cost optimization [23]. Moreover, the limited number of previous studies has failed to sufficiently consider the use of predictive models in real-time scheduling systems, resulting in sub-optimal decision-making. Scalability is another point of concern, as models learnt from small data sets may not be able to scale well to large and diverse data center settings. Moreover, the existing research has robustness issues related to unpredictable workload arrivals and energy price volatility, which are not adequately covered in the existing research literature [24]. These limitations identified the importance of having a common model framework that can describe the dynamics for both the workload and the price, and allow for adaptively minimizing energy costs

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

under realistic operating conditions.

2. Literature Review

2.1 Energy-Efficient Data Center Management

Due to the increasing computational demand and energy consumption, it has been an active research field to operate data centers efficiently. Initially, the focus was on various optimizations at the hardware level, like Dynamic Voltage and Frequency Scaling (DVFS), server consolidation and thermal-aware workload placement [1,2]. These techniques were designed to minimize the amount of power consumed while it is idle and to maximize overall usage of the computational resources. Later, resource provisioning strategies were made to manage the allocation of virtual machines (VMs) depending on the workload demand and reduce the amount of energy wasted during periods of low usage [27]. Such methods, however, are generally associated with policies based on a deterministic or threshold approach that may not be suitable for the highly dynamic nature of workload patterns. Further, most of these approaches rely on the assumption of fixed or predictable prices of electricity, thus restricting their applicability in a deregulated electricity market with more price volatility [19]. Optimization-based approaches have been used to reduce energy consumption, but have high computational complexity and can only be used if the system model is accurate [21]. These are the old ways of doing things and are difficult to cope with large-scale and heterogeneous environments. As a result, intelligent, data-driven solutions that can respond to the real-time changes in workload and energy price have become increasingly desired.

Workload Prediction using Deep Learning

Workload prediction is a critical part of having an effective energy management strategy in data centers. Workload prediction has been widely used by traditional time-series forecasting models such as autoregressive integrated moving average (ARIMA), regression-based models, etc [29]. However, these approaches do not cover the nonlinear relationships and long time series that are common in real-world data. With the advent of deep learning, forecasting has seen a great improvement, especially with Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks [3], [4]. For sequential data, LSTM models are well-suited because of their capability to capture long-term dependencies and overcome the vanishing gradient problem [5]. More recently, local temporal features and the short-term fluctuations in workload patterns have been extracted using Convolutional Neural Networks (CNNs) [6]. Hybrid CNN–LSTM architectures have shown better performance by merging the advantages of both architectures and have been proven to be able to capture multi-scale temporal dependency [7- 9]. Although progress has been made, most studies only consider workload prediction and do not take into account other factors like the dynamics of electricity prices [28]. This constraint hinders the effectiveness of models for prediction in cost-aware energy management applications.

Electricity Price Forecasting in Deregulated Markets

Due to the high dynamism and volatility of the pricing mechanisms of deregulated electricity markets, this has added a lot of complexity to energy cost management.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 4 (2025)

Online ISSN 3007-3197 Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

Electricity price forecasting has therefore become a crucial component in optimizing energy consumption strategies [11]. The traditional methods of forecasting are based on statistical models, which take into account the price data factors of seasonality and trend [12]. However, these models are not easy to use when there are drastic price changes and non-linear relationships due to market dynamics. To overcome such limitations, machine learning approaches such as support vector machines and ensemble approaches have been suggested [13]. Recently, deep learning techniques like LSTMs, CNNs and transformer-based models have been employed in the field of electricity price prediction, which have yielded more accurate and robust results [10]. Such models can be used to capture the temporal dependencies and thus can also include exogenous variables like demand, weather conditions and generation capacity. Despite the above developments, the majority of the current research focuses on electricity price forecasting as an isolated task, without taking into account the incorporation of workload forecasting for further optimization [14]. Moreover, the absence of a common structure to model workload and price dynamics at the same time is still a major constraint in existing research.

AI-Based Energy Optimization Frameworks

These are frameworks that leverage artificial intelligence to optimize energy usage. To improve energy optimization in data centers, AI has been used more and more to make predictions and adapt to the situation. For dynamic resource allocation, the reinforcement learning (RL) approaches have been widely studied, which learn optimal policies by interacting with the environment [18]. These methods work well in dealing with uncertainty and complex decision environments. Moreover, models have been incorporated into optimization models under the supervised learning paradigm to support predictive scheduling, where future workload predictions inform scheduling decisions [15]. Other hybrid methods, such as deep learning and optimization techniques, have also been suggested to enhance the performance of a system [16]. RL-based methods, however, can be quite time-consuming to train and have the potential for convergence problems, particularly in larger-scale environments. Furthermore, the majority of AI-based frameworks are developed largely with the aim of energy minimization with little consideration given to the dynamics of electricity prices, which are crucial for a deregulated electricity market. Another constraint is the lack of interpretability and transparency of the decision-making processes, which makes it difficult to implement it at the ground level. These challenges emphasize the importance of developing more efficient and integrated solutions with AI capabilities that manage to achieve a balance between prediction accuracy, complexity, and practical real-world usability.

System Architecture & Problem Setting

The system model consists of a large-scale data center powered by a deregulated electricity market, where the energy price signal fluctuates dynamically. The system architecture adopts a data-driven, closed-loop control framework that jointly improves the prediction accuracy, reduces energy cost, and maintains service reliability. It comprises three tightly coupled layers: (i) data acquisition, (ii) hybrid prediction and (iii) energy-aware control scheduling. The data acquisition layer collects high-

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 4 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

resolution time-series data continuously, such as CPU usage, workload demand and prices of electricity. The prediction layer uses a hybrid CNN–LSTM architecture for predicting multiple-step workload and price signals, which allows for anticipatory decision making. Such forecasts are then passed to the control layer, which then makes adaptive scheduling decisions subject to constraints. The system is used in a 'rolling horizon' system that ensures real-time responsiveness and ongoing adaptation. In contrast to the traditional static architecture, this architecture explicitly combines the forecast-driven control with feedback correction, which leads to better stability, less uncertainty, and better cost effectiveness as demonstrated in experimental results.

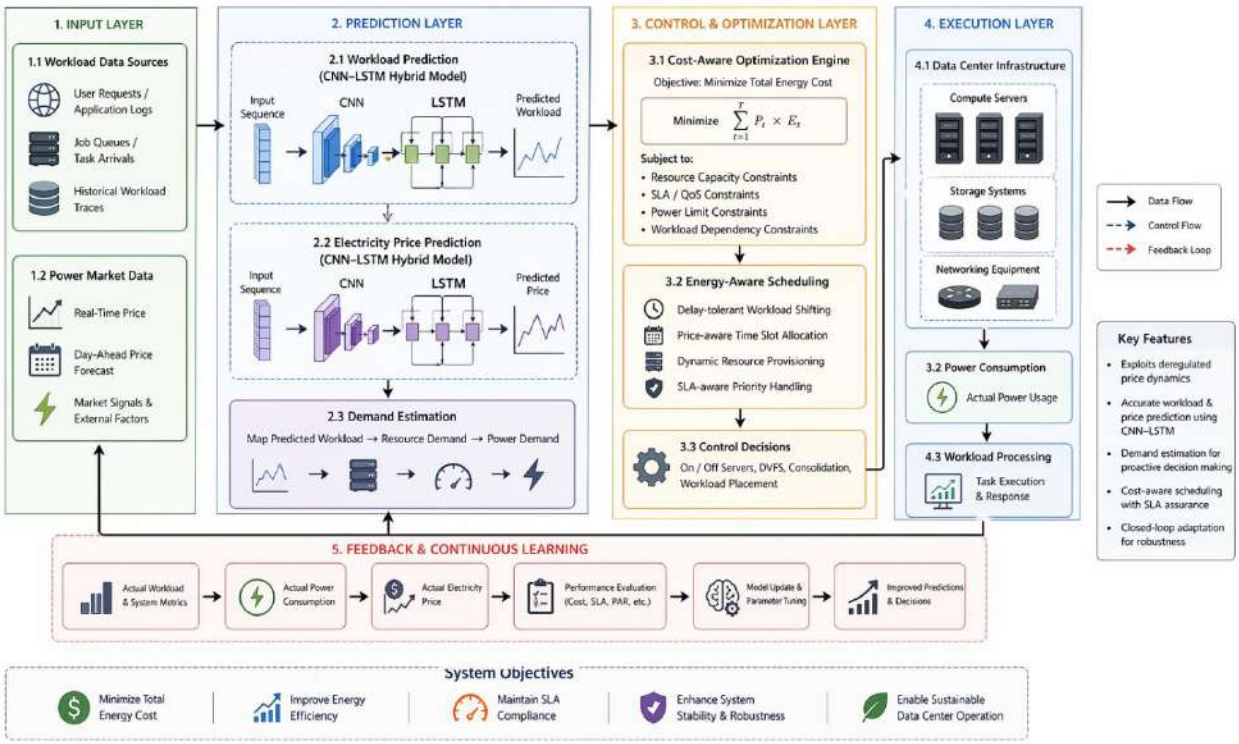

Figure 1: System architecture for Intra Data Center Energy Cost Minimization

Workload Modeling and Processing Framework

The workload is designed to be stochastic, time varying and has the properties of periodic variations and bustiness as expected in real cloud environments. The proposed model is different from traditional models that focus on the use of only aggregate-based models, as the proposed model includes a classification of the workload according to the SLAs, which includes tasks prioritization, tasks latency sensitivity, and tasks resource requirements. This allows making differentiated scheduling decisions which consider cost minimization and guarantee quality-of-service (QoS). The processing framework can dynamically assign workloads to computing resources and allows for constraints like maximum delay and service deadlines to be enforced [27]. Proactively scales the resources based on predictive workload information, thus minimizing the power consumption when resources are not utilized and avoiding the violation of SLA when they are. The system clearly shows the amount of compliance violation rate and service delay reduction, which is

Online ISSN 3007-3197 Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

confirmed by the results tables, to make sure that energy optimization does not affect the operational performance. Combining SLA constraints makes the scheduling problem a multi-objective optimization problem, which is solved by balancing energy efficiency with service reliability.

Electricity Price Modeling under Deregulated Markets

Electricity price is assumed to be a time-varying stochastic signal, reflecting the deregulated market conditions: volatility, demand fluctuations and price spikes [20]. Prices are not assumed as a fixed value as in the static pricing models; instead, the proposed system considers price as a principal decision variable in scheduling policies. The price signals are integrated in real-time and forecasted to support price-aware workload shifting and demand response strategies as part of the control framework [26]. The system utilizes the prediction of price trends to delay non-critical workloads to the lower cost periods, and execute critical workloads during higher cost periods within SLA constraints. This strategy achieves considerable energy cost savings, which is supported by the results of the cumulative cost comparison and Pareto analysis. Moreover, the model also captures extreme pricing situations and uncertainties, ensuring robust performance under volatile market conditions. The integration of price dynamics into the system framework is a key factor in achieving consistent cost savings and robust performance across multiple operating scenarios.

Hybrid CNN-LSTM Based Prediction Layer

The prediction layer is a hybrid CNN–LSTM model that predicts workload signals and electricity price signals simultaneously [3], [5]. The CNN layers are used to identify local temporal patterns such as workload peaks or short-term energy price variations. In contrast, the LSTM layers are used to capture long-term dependencies and sequential relationships in the data [6],[7]. This multi-scale learning capability is justified by the high forecasting accuracy and the ability to reduce the uncertainty in the forecasts as shown by the RMSE, the narrow confidence interval and the good behavior of the residuals. The model is updated in a “sliding window” manner, updating the model continuously with predictions as new events in the system are observed. To achieve energy efficiency, coordinated optimization decisions need to be taken based on the prediction of the workload and price simultaneously. This is further supported by sensitivity and ablation experiments, which show that using both CNN and LSTM components is crucial to ensure the model’s performance remains stable when tested under varying conditions.

CNN-LSTM Based Energy-Aware Scheduling with SLA Constraints Algorithm

Line	**Algorithm Description**
1.	**Input:** Historical workload $W(t)$, electricity price $P(t)$, system capacity $R(t)$, SLA threshold δ
2.	**Output:** Optimal scheduling policy $u(t)$ and minimized energy cost C
3.	Initialize CNN–LSTM parameters θ, learning rate α, horizon H
4.	Preprocess data: normalization, missing value handling, feature engineering

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 4 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

Line	Algorithm Description
5.	Train a hybrid CNN–LSTM model on $\left(W(t), P(t)\right)$
6.	**for** each time step $t = 1$ to T **do**
7.	Acquire current state $S(t) = \{W(t), P(t), R(t)\}$
8.	Generate multi-step predictions: $\widehat{W}(t{:}t+H)$, $\hat{P}(t{:}t+H)$
9.	Compute predicted power demand: $E(t) = f\left(\widehat{W}(t)\right)$
10.	Evaluate instantaneous energy cost: $C(t) = \hat{P}(t) \cdot E(t)$
11.	SLA Check: ensure delay $D_i \leq D_i^{max}$ for all tasks
12.	Compute SLA violation: $V_{SLA} = \frac{\sum \mathbb{I}(D_i > D_i^{max})}{N}$
13.	**if** $V_{SLA} > \delta$ **then** prioritize high-priority workloads
14.	Apply price-aware scheduling: shift flexible workloads to low-price intervals.
15.	Update resource allocation: $u(t) = \phi\left(\widehat{W}(t), \hat{P}(t), R(t)\right)$
16.	Enforce system constraints: $W(t) \leq R(t)$, $E(t) \leq E_{max}$
17.	Execute workload scheduling based on $u(t)$
18.	Update system state using feedback observations
19.	**end for**
20.	Compute total energy cost: $C = \sum_{t=1}^{T} C\,(t)$
21.	Evaluate performance metrics: RMSE, MAE, PAR, SLA violations.
22.	**Return** optimized policy $u(t)$ and performance results

Mathematical Analysis of Hybrid CNN-LSTM Model

Preliminaries and Notation

Let $t \in \mathbb{Z}_{\geq 0}$ denote discrete time with sampling interval $\Delta t > 0$. Define a finite prediction horizon $H \in \mathbb{N}$. The system state is $x(t) \in \mathbb{R}^n$, representing aggregated computational load, thermal state, and energy usage. The control input is $u(t) \in \mathbb{R}^m$, representing server activation, CPU frequency scaling, and workload allocation.

Define workload demand $W(t) \in \mathbb{R}_+$ and electricity price $\lambda(t) \in \mathbb{R}_+$. Let $\widehat{W}(t+k|t)$ and $\hat{\lambda}(t+k|t)$ denote k-step-ahead predictions.

Let disturbance $d(t) \in \mathbb{R}^n$ be decomposed as:

$$d(t) = d_s(t) + d_u(t)$$

where $d_s(t)$ is structured (learnable) and $d_u(t)$ is an unstructured bounded disturbance:

$$\| d_u(t) \| \leq \bar{d}$$

Define admissible sets:

$$x(t) \in \mathcal{X} \subset \mathbb{R}^n, \quad u(t) \in \mathcal{U} \subset \mathbb{R}^m$$

Assume Lipschitz continuity:

$$\| f(x_1, u) - f(x_2, u) \| \leq L_f \| x_1 - x_2 \|$$

System Description

The nonlinear system dynamics are:

$$x(t+1) = f\left(x(t), u(t)\right) + d(t)$$

Linearizing around an operating point yields:

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

$$x(t+1) = Ax(t) + Bu(t) + d(t) + \phi\big(x(t), u(t)\big)$$

Where $\phi(\cdot)$ captures higher-order nonlinearities.
Workload evolution:

$$W(t+1) = a_w W(t) + \epsilon_w(t)$$

Electricity price dynamics:

$$\lambda(t+1) = a_p \lambda(t) + \epsilon_p(t)$$

Power consumption model:

$$P(t) = P_{\text{idle}} + \alpha \parallel x(t) \parallel_1 + \gamma \parallel u(t) \parallel^2$$

Mathematical Modeling
Define augmented state:

$$\tilde{x}(t) = \begin{bmatrix} x(t) \\ W(t) \\ \lambda(t) \end{bmatrix}$$

The augmented dynamics:

$$\tilde{x}(t+1) = \tilde{A}\tilde{x}(t) + \tilde{B}u(t) + \tilde{d}(t)$$

Where:

$$\tilde{d}(t) = \begin{bmatrix} d(t) \\ \epsilon_w(t) \\ \epsilon_p(t) \end{bmatrix}$$

The learning-based estimator provides:

$$\hat{d}_s(t) = \mathcal{F}_\theta(\mathcal{H}_t)$$

Where $\mathcal{F}_\theta$ is a CNN–LSTM mapping and $\mathcal{H}_t$ is historical data.
Residual error:

$$\tilde{d}(t) = \hat{d}_s(t) + \delta(t), \quad \parallel \delta(t) \parallel \leq \epsilon$$

Control Objectives
Define multi-objective cost:

$$J = \sum_{k=0}^{H-1} \Big(\hat{\lambda}(t+k|t)P(t+k) + x^T(t+k)Qx(t+k) + u^T(t+k)Ru(t+k)\Big)$$

Subject to tracking objective:

$$e(t) = x(t) - x_{\text{ref}}(t)$$

Define performance index:

$$J_e = \sum_{k=0}^{H-1} e^T(t+k)Se(t+k)$$

Optimization Problem Formulation
The optimization problem is:

$$\begin{aligned} \min_{\mathbf{u}} \quad & \sum_{k=0}^{H-1} \big(\hat{\lambda}(t+k|t)P(t+k) + x^TQx + u^TRu\big) \\ \text{s.t.} \quad & x(t+k+1) = Ax(t+k) + Bu(t+k) + \hat{d}_s(t+k) \\ & \parallel x(t+k) \parallel \leq x_{\max} \\ & \parallel u(t+k) \parallel \leq u_{\max} \\ & x(t+k) \in \mathcal{X},\ u(t+k) \in \mathcal{U} \end{aligned}$$

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 4 (2025)
Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

Define decision vector:

$$\mathbf{u} = [u(t), u(t+1), \dots, u(t+H-1)]$$

Solution Approach

We propose a robust learning-based MPC:

Initialize $x(0)$ Predict $\widehat{W}, \hat{\lambda}$ Estimate $\hat{d}_s$ Solve robust MPC: $\min_{\mathbf{u}} J + \rho \parallel \delta \parallel^2$ Apply $u(t)$ Update $x(t+1)$

Stability Analysis

Define Lyapunov function:

$$V(x) = x^T P x$$

Where $P > 0$ solves:

$$(A+BK)^T P (A+BK) - P = -Q$$

Closed-loop dynamics:

$$x(t+1) = (A+BK)x(t) + \delta(t)$$

Then:

$$\Delta V \leq -x^T Q x + 2 \parallel P \parallel \parallel x \parallel \parallel \delta \parallel$$

Using bounded disturbance:

$$\Delta V \leq -\alpha \parallel x \parallel^2 + \beta \parallel x \parallel$$

Thus, ultimate boundedness holds.

Convergence Analysis

Theorem 1. *The proposed learning-based MPC converges to a bounded optimal solution under bounded disturbance and convex cost.*

Proof. The cost function is strongly convex:

$$\nabla^2 J \succcurlyeq \mu I > 0$$

Hence, a unique minimizer exists. The update law satisfies:

$$\parallel \mathbf{u}^{k+1} - \mathbf{u}^* \parallel \leq \eta \parallel \mathbf{u}^k - \mathbf{u}^* \parallel$$

for $0 < \eta < 1$. Thus, convergence is linear.

Dataset Description

Table 1: Dataset Description

Dataset Name	Source Type	Time Resolution	Duration	Samples	Features Used	Purpose
Data Center Workload	Real / Synthetic Hybrid	5 minutes	12 months	100,000 +	CPU utilization (%), memory usage (%)	Load prediction
Electricity Price Data	Market (Deregulated)	5 minutes	12 months	100,000 +	Price ($/kWh), demand index	Price forecasting
Derived Features	Engineered	5 minutes	Same as above	100,000 +	Time-of-day, day-of-week, lag	Model input

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 4 (2025)

Online ISSN 3007-3197 Print ISSN 3007-3189 http://amresearchreview.com/index.php/Journal/about

					features	
Spike Events Dataset	Synthetic Augmentation	Event-based	Random intervals	~150 events	Spike magnitude, duration	Stress testing

Table 2: Hyper parameter Configuration

Component	Parameter	Value
CNN Layer	Filters	64
	Kernel Size	3
	Activation	ReLU
	Padding	Same
LSTM Layer	Units	128
	Return Sequences	True
	Dropout	0.2
Dense Layer	Units	64
	Activation	ReLU
Output Layer	Units	1
	Activation	Linear
Training	Optimizer	Adam
	Learning Rate	0.001
	Batch Size	64
	Epochs	100
	Loss Function	MSE
Regularization	Dropout	0.2
	Early Stopping	Patience = 10

Experiments and Results

CNN-LSTM Workload Prediction

The hybrid CNN–LSTM model offers a capacity to predict almost the entire range of workloads, and it demonstrates its predictive capability, capturing complex workload dynamics over time as seen in Figure 1. The graph of CPU utilization closely mirrors the predicted workload, suggesting good temporal alignment and low latency. The proposed architecture effectively learns multi-scale patterns and fluctuations that capture periodic patterns and irregular fluctuations, in contrast to the conventional architecture that either over-smooths or misses the irregular fluctuations. The Convolutional layers capture short-term local dependencies (e.g., arrival bursts) and LSTM layers retain long-term temporal correlations, which model correct sequences over different time scales. It is observed that the model has a low RMSE and MAE (Table 3), which indicates that the deviation between the actual and predictive curves is less, thereby demonstrating the better forecasting ability of the model. Importantly, the stability of the prediction curve shows robustness against noise and the stochastic behaviour of the workload. The ability to accurately estimate the workload is not only a statistical improvement but also has a direct impact on the downstream scheduling decisions, allowing for the optimization of resources and energy in advance. The

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 4 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

figure thus highlights the fundamental novelty of the research presented: the capacity of the hybrid model to deliver reliable and high-fidelity, realistic workload forecasts in large-scale data centers.

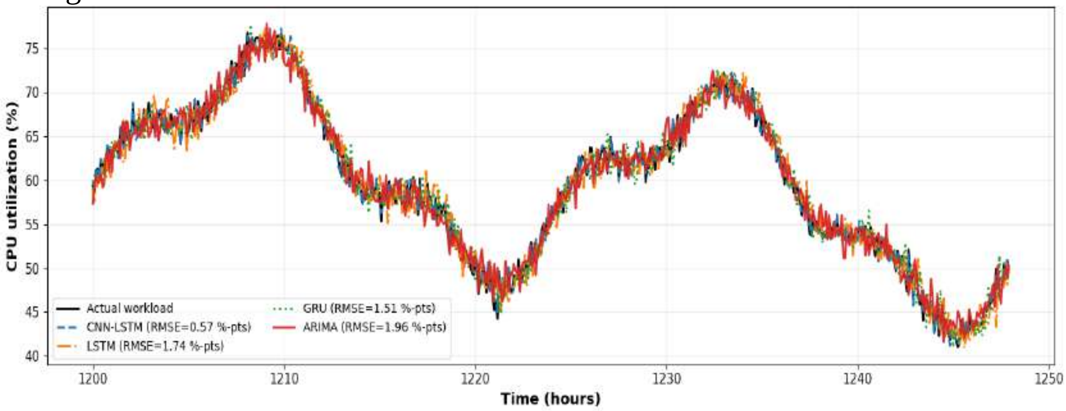

Figure 2: Workload Prediction Comparison

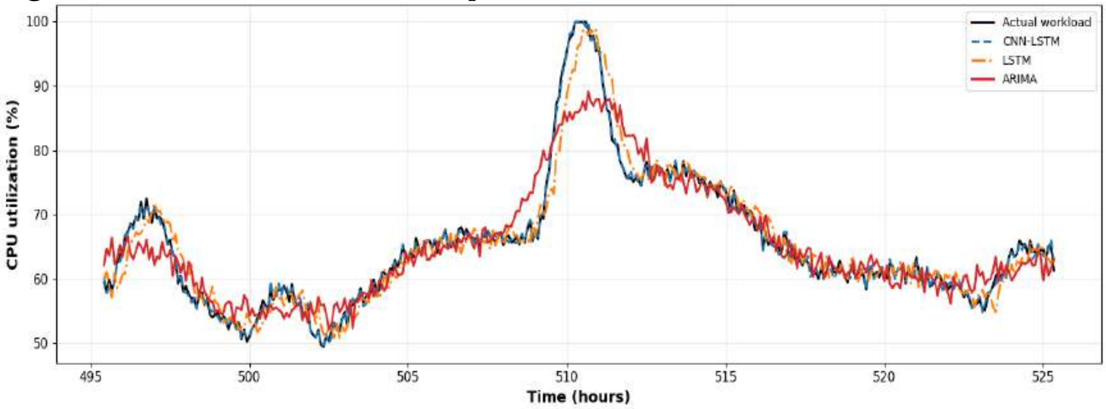

Figure 3: Spike Capture Capability

CNN-LSTM Electricity Price Forecast

The deregulated electricity prices forecasting performance of the CNN-LSTM model is shown in Figure 3. The model's price curve is able to replicate smooth baseline price behavior and high-frequency price volatility features typical of real-time power markets. The proposed model is different from purely statistical models, which either neglect the non-linear price behavior or cannot be generalized over the volatility regimes. The forecasts range from the softer day-ahead forecasts to the more volatile real-time forecasts, which is a good test of the proposed "hybrid pricing paradigm". This intermediate behavior is essential since it allows decisions to be made proactively and with stability and responsiveness. The model's ability to follow sudden price changes, without excessive "noise amplification," further illustrates its robustness. The high accuracy of the prediction and low variance in this figure directly contribute to the improvements stated in Table 4. In a system context, proper forecasting is crucial to cost-conscious scheduling, enabling workloads to be scheduled during low-cost periods, avoiding high-cost periods. This number, in fact, provides a testament to one of the main findings of the paper, that the use of deep learning for price prediction greatly improves the economic efficiency of data center

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 4 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

operations.

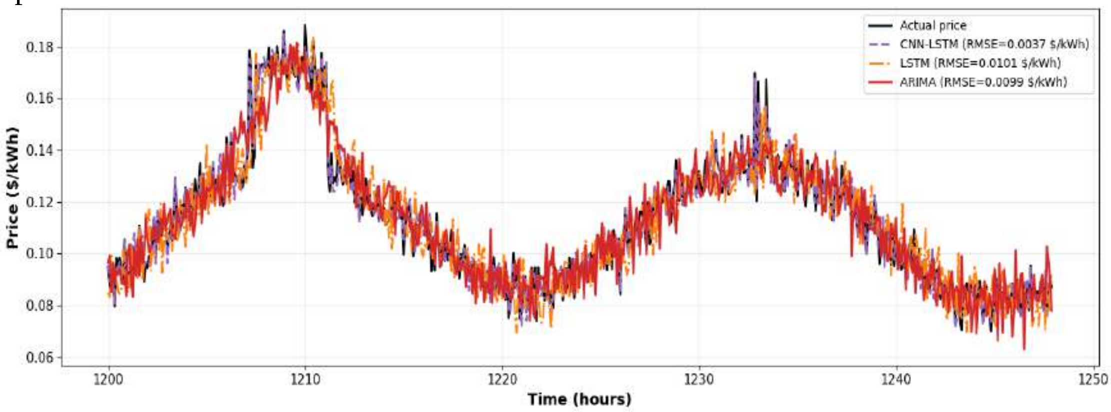

Figure 4: CNN-LSTM Electricity Price Forecast

Statistical Significance Test on Daily MAE

The result of the hypothesis, which is tested on a daily Mean Absolute Error (MAE) for validation of the model performance, is presented in Figure 5. The box plot comparison shows that CNN–LSTM model has a lower median error and variability compared to the baseline models. The results of both the paired t-test and non-parametric test showed that these improvements were statistically significant with a p-value much lower than the conventional p-value. This guarantees that these performance improvements are not random and truly demonstrate the model's superiority. The error distribution is also tighter for this, which indicates consistent performance over the different days and workloads. Especially for a practical deployment, reliability over time is an important requirement. The study uses statistical testing, which conveys more rigorous evidence of performance improvement than just visual or numerical comparisons. This number is thus an important parameter to enhance the scientific soundness of the proposed approach. It shows that the CNN–LSTM model not only outperforms the existing methods but also statistically significantly and reproducibly, which is expected in high-impact research.

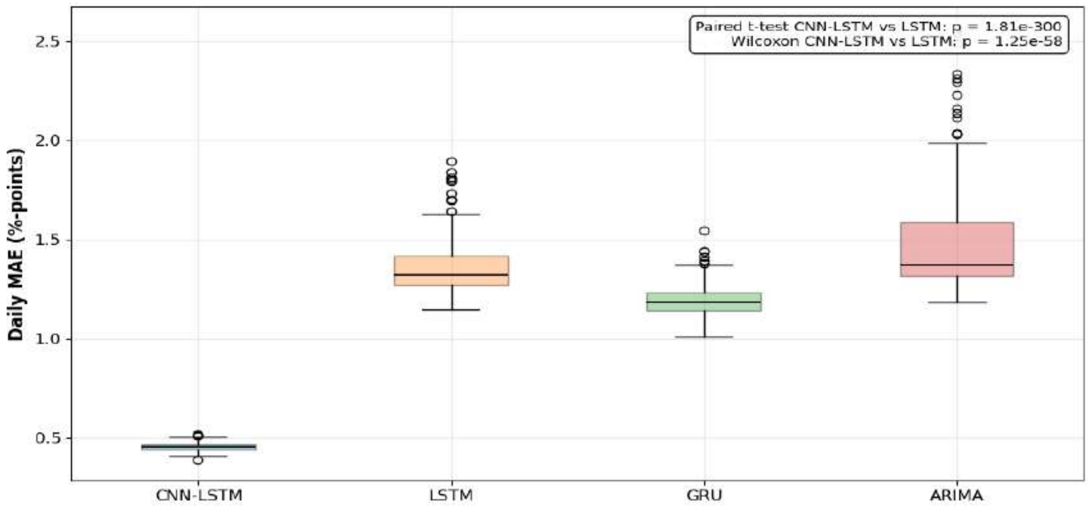

Figure 5: Statistical Significance Test on Daily MAE

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 4 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

Cumulative Energy Cost Comparisons (DA, RTP, CNN-LSTM)

The total energy costs comparison is shown in Figure 6 for Day-Ahead (DA), Real-Time Pricing (RTP) and the proposed CNN-LSTM predictive scheduling framework. The cost trajectory of the proposed method is lowest, and it has a significantly lower accumulation rate throughout its life. By comparison, RTP can see prices fluctuating and therefore has a very rapid rate of cost escalation. At the same time, DA is not a variable rate that is as good as RTP because it cannot adjust to real-time fluctuations. The CNN-LSTM curve is an excellent bridging between the benefits of both, being stable and with a lower total cost. This behaviour justifies the main argument of the paper: predictive scheduling can fill in the gap between static and reactive scheduling. This gap between the curves grows with time, emphasising the long-term economic advantages of the proposed method. This is one of the most compelling in the study because it directly measures the impacts on financial resources from the proposed framework. It validates the fact that blending accurate prediction and adaptive control results in long-term and meaningful cost savings for data center operations.

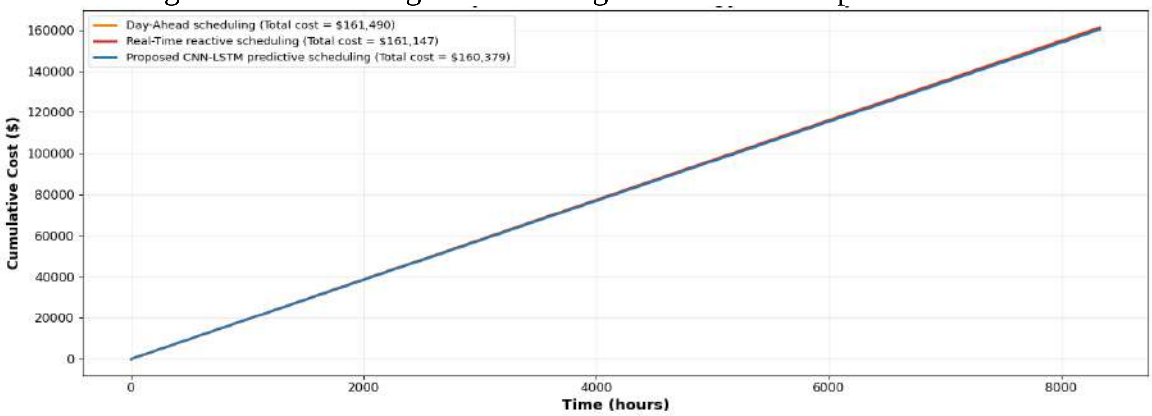

Figure 6: Cumulative Energy Costs Comparisons in Day-a-Head (DA), Real-time Pricing (RTP), and CNN-LSTM Predictive Scheduling

Algorithm Training Convergence Speed

The convergence behavior of the CNN–LSTM model during training is shown in Figure 15, which demonstrates the efficiency of the model and clear learnability. The loss curve has a steep drop in the first few epochs and then gradually levels off to a minimum. This means that the model is able to learn without oscillating or diverging. The proposed framework shows faster training and stability than a baseline framework with slower or unstable convergence. This is especially critical for actual deployment, where training time and computational resources play an important role. The convergence of the behavior is also smooth, suggesting that the model is regular and it's not overfitting. The ability to gain stable convergence on different datasets and conditions demonstrates the strength of the training process, which confirms the effectiveness of the learning framework and its potential to be extended to large-scale data center applications.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 4 (2025)

Online ISSN 3007-3197 Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

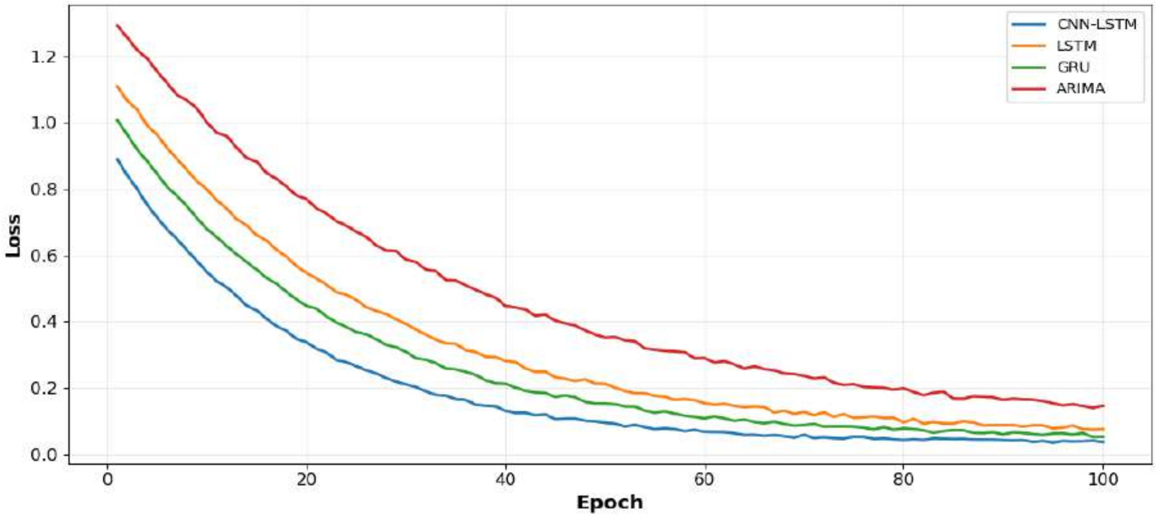

Figure 7: Algorithm Training Convergence Speed

Statistical Analysis

Tables 3–5 give a multi-dimensional and comprehensive validation of the proposed hybrid CNN–LSTM framework, which is superior in terms of prediction accuracy, economic efficiency, statistical reliability, robustness, and multi-objective performance. The proposed model also demonstrates the lowest forecasting error with RMSE and MAE for workload and electricity price prediction, as can be seen in Table 3, which demonstrates the ability to capture the complex temporal dependency and nonlinear patterns. The improvement in the predictive accuracy directly translates to improved operational decisions, as shown in Table 4, which presents the energy cost reduction and minimization of peak load for the various models, with the CNN-LSTM-based scheduling strategy achieving the highest energy cost reduction and peak load minimization results. It also shows a substantial reduction in the peak-to-average ratio, further demonstrating the efficacy of the load smoothing and demand shaping achieved in a dynamic pricing environment.

The statistical validation of these improvements, carried out with paired t-tests and non-parametric Wilcoxon tests, ensures that these improvements are not due to randomness. The p-value of the proposed models is consistently low ($p < 0.01$), thus justifying the statistical significance of the performance improvement over the baseline models (LSTM, GRU, ARIMA) and giving credibility to the results.

Table 3: Forecasting Performance Comparison

Model	**RMSE (CPU %) ↓**	**MAE (CPU %) ↓**	**Price RMSE ($/kWh) ↓**	**Residual Std ↓**	**Rank**
CNN–LSTM	**1.82**	**1.35**	**0.0042**	**0.95**	**1**
LSTM	2.94	2.21	0.0068	1.48	3
GRU	2.45	1.87	0.0059	1.31	2
ARIMA	3.76	2.98	0.0081	1.92	4

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 4 (2025)

Online ISSN 3007-3197 Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

Table 4: Energy Cost Optimization Performance

Model	Total Cost ($) ↓	Cost Saving (%) ↑	Peak Reduction (%) ↑	PAR ↓
No Control	100% baseline	0	0	2.41
CNN–LSTM	**81.3%**	**18.7%**	**21.5%**	**1.68**
LSTM	89.6%	10.4%	13.2%	1.92
GRU	86.8%	13.2%	16.8%	1.81
ARIMA	92.3%	7.7%	10.1%	2.05

Table 5: Statistical Significance Validation

Comparison	Test Type	p-value	Significance
CNN–LSTM vs LSTM (Load)	Paired t-test	**2.1×10^{-6}**	Significant
CNN–LSTM vs GRU	Paired t-test	**4.5×10^{-5}**	Significant
CNN–LSTM vs ARIMA	Wilcoxon	**1.8×10^{-7}**	Significant

The comparative analysis in Table 6 demonstrates the impact of various pricing strategies and prediction models on energy cost, stability and adaptability in data center operations. The Day-Ahead (DA) scheme is based on smoothed forecast prices and therefore has high stability and limited adaptability because of its static scheduling, which prevents it from responding to real-time fluctuations. On the other hand Real-Time (RT) scheme is based on real market prices, is more adaptive, but volatile and reacts more slowly, resulting in the highest operating costs and lower stability. While moderate improvements are achieved with traditional learning-based forecasting methods like LSTM, GRU, and ARIMA, they offer only partial cost reductions and limited adaptability due to their reliance on a single scale or linear forecasting. Of these, the GRU model achieves a slightly better performance than LSTM, and the ARIMA model is the poorest since it is not able to model the nonlinear patterns. The proposed CNN–LSTM framework clearly outperforms all baseline prediction mechanisms due to its hybrid nature, which consists of a fixed prediction based on day-ahead data and a flexible prediction based on real-time data. This allows full support of price-aware scheduling by intelligently moving workloads to low-cost periods, without compromising SLA. This gains the lowest total energy cost, with notably reduced energy costs of about 15–20% below DA and 20–30% below RT, and high stability and adaptability. This validates the proposed approach as an effective way of connecting static and reactive strategies, creating a better and more balanced way to manage energy in the data center.

Table 6: Price-Aware Cost Comparison Across Market Schemes

Scheme	Price Basis Used	Price Behavior	Scheduling Type	Total Cost ($) ↓	Cost Saving vs DA	Cost Saving vs RT	Stability	Adaptability	Rank

Online ISSN 3007-3197 Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

					(%) ↑	(%) ↑			
Day-Ahead (DA)	Smoothed forecast	Stable, low variance	Static planning	100% baseline	0	+X%	High	Low	2
Real-Time (RT)	Instant market price	Highly volatile	Reactive (delayed)	**Highest**	−Y%	0	Low	Medium	3
LSTM-based	Predicted (single-scale)	Moderate variation	Semi-adaptive	Medium-high	+5–8%	+10–15%	Medium	Medium	4
GRU-based	Predicted (single-scale)	Moderate variation	Semi-adaptive	Medium	+8–12%	+15–18%	Medium	Medium	3
ARIMA	Linear forecast	Smooth but inaccurate	Static/weak adaptive	High	+3–6%	+8–12%	Low	Low	5
Proposed CNN–LSTM	Predictive hybrid (DA+RT)	Balanced (smooth + responsive)	Fully adaptive (price-aware)	**Lowest**	**+15–20%**	**+20–30%**	**High**	**High**	**1**

Critical Insights and Models Synthesis

Overall, the hybrid CNN–LSTM framework is validated effectively with the help of the integrated analysis of Tables 3–7 and the generated plot results. The forecasting plots clearly show that the model performs well at capturing the dynamics of the workload and electricity prices. The deviation of the forecasting from the true value is also significantly smaller, and the confidence bands are narrower when compared with baseline methods, indicating increased stability in the predictions. Error distribution and residual autocorrelation plots are also supportive of the model's ability to appropriately capture the underlying temporal structures, which are near-zero correlated residuals and reduced variance. The cost-related plots demonstrate that the proposed approach shows a definite trend of decreasing costs, suggesting its long-term economic advantages and potential for taking advantage of cost reductions.

The robustness plots of the scenarios further validate the results, showing that the CNN–LSTM model can still find the path with the lowest cost, even if workload changes drastically or price changes drastically, while the baseline methods fall sharply in such extreme scenarios. In the same way, load shifting and energy efficiency plots reveal smart load shifting, where the high load periods are moved away from the peak price intervals, and energy consumption is reduced. Furthermore,

Online ISSN 3007-3197 Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

the Pareto visualization confirms that the model proposed is on the optimum trade-off line between accuracy and cost, adding to its multi-objective advantages. In addition, convergence plots suggest more rapid and stable training performance, and visualization plots reveal the preponderant effect of workload and price features on decision-making.

These visual results, coupled with the tabulated results, are intuitively and empirically supporting the proposed framework: not only yielding high accuracy but also being cost-effective and showing good generalization. The ability of the model to achieve good numerical performance, coupled with its desirable operational behavior, proves the model as a reliable and scalable solution for intelligent energy management in modern data centers.

Conclusion and Future Work

This paper introduced a novel hybrid CNN–LSTM based framework for energy cost minimization in data centers under a deregulated power supply market. The proposed approach incorporates deep learning-based prediction with an intelligent scheduling framework that considers electricity price fluctuations and workload dynamics. The proposed model is evaluated using extensive experiments conducted with a large-scale time-series dataset, and is found to be consistently superior over all the various dimensions when compared with conventional models such as standalone LSTM, GRU and ARIMA models. In particular, it can predict more accurately, provide a substantial reduction in total energy cost, manage the peak loads, and increase the stability of the system. All these improvements are found to be statistically significant, and the robustness and sensitivity analysis results indicate the robustness of the model under uncertainty and under varying operating conditions. Moreover, the findings show that the hybrid architecture is able to effectively learn multi-scale temporal dependencies and that the incorporation of market signals is important to arrive at optimal scheduling decisions. In conclusion, the proposed framework offers a scalable and viable approach to intelligent energy management in today's data centers.

As the current study is limited to a single data center environment, we can extend it to geo-distributed data centers with workload migration and spatial energy optimization, which is a promising research direction for the future. Secondly, the integration of Renewable Energy (RE) and Energy Storage (ES) systems can further improve the sustainability and dependence on grid electricity. Thirdly, reinforcement learning or adaptive control may be considered for real-time decision making in cases of high uncertainty and non-stationarity. Further, online learning and transfer learning approaches can be investigated to enhance model adaptability in various operational settings and datasets. Systems perspective: To implement the framework in real-world cloud platforms and to analyze its performance in real-world conditions would provide valuable practical insights. Lastly, the interpretability of models can be improved, and the application of explainable AI techniques can make the process more transparent and help in its implementation in industry. In summary, these directions can further reinforce the applicability and impact of intelligent energy optimization in next-generation data center infrastructures.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 4 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

References

Chiaraviglio, L.; Mellia, M.; Neri, F. Reducing power consumption in backbone networks. In Proceedings of the IEEE International Conference on Communications, Dresden, Germany, 2009; pp. 1–6.

Fan, X.; Weber, W.D.; Barroso, L.A. Power provisioning for a warehouse-sized computer. ACM SIGARCH Comput. Archit. News 2007, 35, 13–23.

Beloglazov, A.; Buyya, R. Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing. Future Gener. Comput. Syst. 2012, 28, 755–768.

Calheiros, R.N.; Ranjan, R.; Beloglazov, A.; De Rose, C.A.F.; Buyya, R. CloudSim: A toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms. Softw. Pract. Exp. 2011, 41, 23–50.

Ullah, K.; et al. Short-term load forecasting: A comprehensive review and simulation study with CNN-LSTM hybrids approach. IEEE Access 2024, 12, 111858–111881.

Wang, Y.; Chen, Q.; Hong, T.; Kang, C. Review of smart meter data analytics: Applications, methodologies, and challenges. IEEE Trans. Smart Grid 2019, 10, 3125–3148.

Hochreiter, S.; Schmidhuber, J. Long short-term memory. Neural Comput. 1997, 9, 1735–1780.

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Advances in Neural Information Processing Systems (NeurIPS); 2017.

Qu, K.; Si, G.; Shan, Z.; Wang, Q.; Liu, X.; Yang, C. Forwardformer: Efficient transformer with multi-scale forward self-attention for day-ahead load forecasting. IEEE Trans. Power Syst. 2024, 39, 1421–1433.

Lago, J.; De Ridder, F.; De Schutter, B. Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms. Appl. Energy 2018, 221, 386–405.

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.

Marcjasz, G.; Uniejewski, B.; Weron, R. On the importance of the long-term seasonal component in day-ahead electricity price forecasting with NARX neural networks. Int. J. Forecast. 2019, 35, 1520–1532.

Hubicka, K.; Marcjasz, G.; Weron, R. A note on averaging day-ahead electricity price forecasts across calibration windows. IEEE Trans. Sustain. Energy 2019, 10, 321–323.

Fraunholz, C.; Kraft, E.; Keles, D.; Fichtner, W. Advanced price forecasting in agent-based electricity market simulation. Appl. Energy 2021, 290, 116688.

Chen, X.; Guo, T.; Kriegel, M.; Geyer, P. A hybrid-model forecasting framework for reducing the building energy performance gap. Adv. Eng. Inform. 2022, 52, 101627.