DFSO-LSTM-Based Market Demand Forecasting and Resource Scheduling for Independent Energy Storage in Power Grid

Zhiqiang Wang*, Jin Wang, Yueli Zhou, Kexin Liu, Zheng Weng
Cgs Power Generation（Guangdong）Energy Storage Technology Co., Ltd, Guangzhou, Guangdong, 510630, China
E-mail: zhiqiangwang989@outlook.com

Keywords: Market demand, resource scheduling, power grid, energy storage, deep learning

Recieved: May 29, 2025

Forecasting market demand and scheduling energy storage scheduling are critical challenges in power grids, particularly due to data irregularities and renewable energy uncertainty. This research proposes a hybrid DFSO-LSTM model combining Demand-based Fish Swarm Optimization (DFSO) and Long Short-Term Memory (LSTM) to enhance demand prediction and operational efficiency. The model was evaluated using a large-scale dataset comprising hourly power consumption, climate variables, and system parameters collected from public sources between January 2018 and June 2023. DFSO dynamically tunes key LSTM hyperparameters including time steps and hidden units by minimizing RMSE across validation sets. Experiments were conducted using python and obtained comparative analysis results against GA–BP and GAN-NetBoost shows that the proposed model achieves superior accuracy of 96.15%, with MAPE of 0.0569, RMSE of 1.085, MAE of 1.1025, MSE of 1.895 and R^2 of 0.957. A one-minute reduction in execution time demonstrates practical deployment viability. Statistical tests confirm that improvements are significant. These results validate the model's effectiveness in enabling scalable, real-time energy storage scheduling and peak load management in smart grid environments.

Povzetek: Študija predstavi hibridni model DFSO-LSTM za napovedovanje odjema in razporejanje hranilnikov, ki na večletnih omrežnih podatkih vodi v natančnejše napovedi in učinkovitejše upravljanje obremenitev.

1 Introduction

The stability and dependability of power grids are put to the test when renewable energy sources are included because of the intermittency and variability they bring [1]. To address these issues, energy storage devices are crucial, as they allow us to store extra energy during periods of low demand or high renewable generation and release it during periods of low demand or generation. Independent energy storage systems in power networks can only be operated to their full potential with accurate demand forecasts and well-planned resource allocations [2]. The intricate dynamics of energy markets and grid operations can be difficult for traditional scheduling and forecasting methods to adequately represent, as they frequently depend on oversimplified models and heuristics. Energy storage devices are being used more and more frequently in the dynamic energy market and power grid management landscape. Supply and demand, grid stability, and renewable energy integration are all greatly enhanced by these systems [3]. Optimising the charging and discharging methods of energy storage assets is a major difficulty when trying to maximise their value, especially in commercial environments driven by the market. Optimisation for energy savings has often made use of heuristics or oversimplified models, neither of which do a good job of capturing the complexities of operational constraints and market dynamics [4].

Effective energy conservation management also requires precise demand forecasts and well-thought-out resource allocation. To overcome these obstacles and optimise the charging and discharging methods of energy storage devices in power grids' market-oriented trading environments, this paper suggests integrating resource planning techniques with deep learning-based demand forecasting [5]. The goal is to make better decisions about where to put energy storage resources by enhancing the capabilities of deep learning models, which should lead to more accurate demand forecasts. Also, in response to operational needs, grid conditions, and real-time market signals, resource planning algorithms based on deep learning will dynamically change charging and discharging techniques. Our goal is to help build power networks that are more robust, sustainable, and efficient by maximising the potential of energy storage assets through the application of deep learning [6]. Power systems currently make extensive use of hybrid energy storage systems, which combine energy with power storage. The optimal scheduling model has been updated to include energy storage as a schedulable resource over

time by scholars. The economics and reliability of the system can be enhanced through flexible regulation of energy storage, which can also maximise the interests of those who own energy storage according to signals from the spot market.

From a demand-side perspective, energy storage allows consumers to accomplish "peak shaving and valley filling," which lowers their electricity consumption, prevents power system instability due to frequent startup and shutdown of certain generating units, and lowers production costs [7]. Businesses can now take advantage of energy storage technologies to accomplish peak shaving and valley filling due to advancements in the field and falling battery prices. There will be additional opportunities and problems for energy storage with the advent of microgrids, additive distribution networks, and user participation in auxiliary services [8]. The power load, power consumption curve, and predictability will all be affected by connecting energy storage to the power system. Applying some of the more conventional models of power usage to energy storage might lead to less precise predictions. Energy storage's role in load and power consumption forecasting, peak shaving in power systems, frequency modulation auxiliary services, and similar topics have been the subject of much prior research by both domestic and international academics [9]. An adaptive optimisation control strategy is suggested for energy storage systems to take part in the main grid frequency regulation, which takes into account the system's current state and effectively satisfies demand. This strategy is part of the research on energy storage control strategies. Energy storage is optimised for both charging and discharging in both the load and state-of-charge (SOC) states. Strategies for controlling charging and discharging are developed for energy storage that is involved in regulating the power grid's peak load. Regarding energy storage capacity planning and setup, a secondary frequency regulation capacity allocation approach was suggested. This approach, which is based on life cycle theory, offers a useful strategy for planning the layout of energy storage systems in order to maximise net benefits [10]. To address the consumption and grid-connected issues caused by wind power's intermittent and volatile output, as well as to provide a benchmark for allocating energy storage capacity in the context of emerging high-permeability grid-connected energy, a method was suggested for relaxing the peak shaving bottleneck. The proposal was for a load-forecasting-based approach to allocating peak and frequency power modulation for power stations' energy storage. To improve the accuracy of the power load forecasting model, a genetic algorithm was used to optimise the BP neural network's weights and thresholds. An approach to scheduling energy storage that is based on price contracts was developed and implemented in a real-time scheduling model for power systems, all in accordance with the aggregation concept of the load aggregator mechanism in smart networks [11]. The authors present a novel optimal scheduling model that takes into account both solar and wind power to reduce peak demand. Both the combined peak shaving scheduling model for solar and wind storage and the stable economic output model for hydropower units were solved using the multi-objective particle swarm optimisation technique. There was a proposal to enhance the duration-based and frequency-based dependability indices through the simultaneous placement of control and protection devices in an emergency demand response program. In order to maintain a steady flow of power, it is necessary to implement a price-based demand bidding process [12]. A novel scheduling paradigm for generalised demand-side resource coordination was suggested, which might be implemented through power price contracts. As a scheduling optimisation goal, this model considers the full range of demand-side resources and the maximum economic benefits of a load aggregator. It then returns scheduling results to system dispatchers as demand response signals.

The Microgrid has developed into a sophisticated autonomous system as a result of the unceasing advancement of technology. Its ability to be integrated with different energy devices to create a varied system is its defining feature, allowing for the attainment of optimal operation efficiency and the realisation of advantages [13]. Microgrids, which link distributed devices to the energy Internet, significantly lessen issues like high demand, poor control, and inefficient use of electricity that arise from widespread use of the grid. Therefore, optimising the energy scheduling strategy of microgrids and integrating various energy systems into them has become a hot topic among scholars. Smart grids have arisen in response to the need for more digitalisation and intellectualisation in modern society. Part of national security is making sure the electrical system is secure. Because of this, we need a unified power grid and a consistent way of communicating about dispatching. China has a unified power grid with a voltage above 500 kilovolts. Power transformation steadily lowers the voltage level, which is 500 kV, which is used in cities. The current issue that the smart grid must address is whether or not the entire power grid has an effective communication mode to guarantee the appropriate operation of the power grid. The central nervous system of the electricity grid is the dispatching control system. It is at the forefront of power grid security and manages a number of critical indications, including power flow, voltage balance, frequency, and balance of power. Typhoons and mountain torrents are examples of natural calamities that can disrupt electricity grid operations [14]. The power grid dispatching control system is now capable of taking full responsibility for ensuring the grid's safe and stable functioning. The electricity system's dispatching

level has shifted in tandem with its slow but steady expansion.

1.1 Research contribution

This research intends to improve microgrid efficacy and profitability by designing an intelligent energy management model that optimally controls the charging and discharging cycles of energy storage systems, based on forecasted demand and system state. The contributions of this work are as follows:

- Development of a Hybrid DFSO-LSTM Model:
- A novel Deep Fish Swarm Optimization-enhanced Long Short-Term Memory (DFSO-LSTM) model is introduced to optimize energy scheduling. The model combines the exploration capabilities of Fish Swarm Optimization (FSO) with the sequential learning power of LSTM to preserve energy and improve scheduling precision within the energy management system.
- Design of a Multi-Objective Fitness Function:
- A new multi-objective optimization function has been formulated to guide the DFSO algorithm. This function simultaneously considers charging cost, distance-based dispatch parameters, and user preferences, allowing for intelligent trade-offs between operational cost and service quality.
- Evaluation of Grid Stability and Demand Forecasting:

The LSTM component is utilized to evaluate critical power system parameters including grid stability, power quality, load-induced voltage variations, and demand fluctuations. These indicators serve to validate whether the proposed model maintains reliable operation under varying load and peak conditions.

- Testable Research Framework:
- The study explicitly investigates the following hypotheses:
- The DFSO-LSTM model achieves lower forecasting error (MAPE) compared to conventional models.
- The proposed scheduling framework reduces execution time by at least one minute.
- Operational profit and grid performance improve under the DFSO-LSTM strategy, especially during peak demand.
- These contributions are experimentally validated using real-time demand and storage data under fluctuating market conditions, thereby directly supporting the paper's objective to enable efficient and profitable microgrid operation in uncertain environments.

1.2 Research question

- How can demand-based fish swarm optimization improve the accuracy of LSTM-based power load forecasting in smart grids?
- Can hybrid deep learning and metaheuristic models effectively support real-time energy storage scheduling under fluctuating demand conditions?

2 Literature review

A comparative analysis of multiple approach relevant to the implemented technique was summarized in Table 1.

Table 1: Comparative summary of existing methods

References	Methods Used	Dataset Used	Result	Limitation
[15]	Data-driven Battery Energy Storage System (BESS) scheduling	Simulated PV + BESS grid profiles	Reduced peak load and demand charges	No deep learning or forecasting integration
[16]	Deep Learning + Reinforcement Learning	Smart grid simulation (pricing + storage)	Profit ↑ (quantified in reward function)	No short-term load forecasting; grid-level only
[17]	Reinforcement Learning (Q-learning)	Simulated microgrid control data	Reduced peak-to-average ratio	No predictive module; reactive scheduling only
[18]	Linear programming optimization	Simulated mixed-energy microgrid (PV + BESS + diesel)	Fuel reduction ≈ 12%	No AI; lacks demand forecasting or learning
[19]	Kolmogorov–Arnold Network (KAN)	Real-world microgrid demand data	MAPE ≈ 3.9%	Forecast-only; no scheduling or operational decision logic

[20]	LSTM + Monte Carlo Dropout	Historical demand dataset (PJM-like)	MAPE = 2.70%, RMSE = 0.0081, R^2 = 0.9901	No control module; forecasting-only system
[21]	Adaptive Convolutional Residual Network	Multivariate grid data (load + price)	MAPE = 2.43%, RMSE = 0.0065, R^2 = 0.9923	High complexity; lacks interpretability and deployment integration
[22]	Variational Autoencoder (VAE) for cost features	Simulated high-dimensional power cost data	Improved latent feature quality (non-metric)	Not forecasting or scheduling focused
[23]	Fuzzy multi-objective decision system	Logistics data (non-energy sector)	Achieved optimal multi-criteria site selection	Non-energy domain; not applicable to microgrid control
[24]	PV-BESS forecast-based scheduling with incentives	Real-world Korean grid + incentive structures	Improved scheduling + battery efficiency	Depends on specific market structures
[25]	Original Fish Swarm Algorithm	Benchmark optimization functions	Converged on test functions	Not applied to energy systems; lacks adaptation
[26]	FSA with Levy flight and firefly enhancement	Benchmark datasets (Sphere, Rosenbrock)	Faster convergence, better global search	Generic optimizer; no energy application shown
[27]	Genetic Algorithm (GA)–Reinforced Deep Neural Network	PV + Net load microgrid dataset	MAPE = 2.90%, RMSE = 0.0069, R^2 = 0.9860	Forecasting-only; static GA tuning; no dispatch mechanism
[28]	Enhanced Neural Network for anomaly detection	Smart meter anomaly dataset	High accuracy (fraud detection)	Not applicable to forecasting or BESS control

Materials and methods

3.1 Data collection

The data was collected from the open source called Kaggle: https://www.kaggle.com/datasets/ziya07/hourly-power-load-and-climate-data/data. It contains high-resolution hourly data collected from January 1, 2018 to June 30, 2023, simulating realistic power system operations under variable climatic and temporal conditions. The dataset includes a total of 48,000 hourly records, from which a representative sample of 10,000 records was used for modeling and experimentation. Each data entry includes power load (in kilowatts) as the target variable, alongside nine contextual and environmental features: temperature, humidity, wind speed, precipitation, day of the week, and a holiday flag, among others.

To sustenance both short-term and aggregated forecasting tasks, the dataset also offers multi-resolution features. These are allied with the hourly data through reliable timestamp indexing, ensuring temporal truth and multi-scale learning abilities. The final dataset was split into 70% training, 15% validation, and 15% testing subsets using graded random sampling to reserve the dispersal of seasonal and demand-related patterns across all partitions.

3.2 Normalization and outlier detection using Z-score normalization

Standardized data improved model performance and consistency by bringing all data characteristics to the same size. Z-score standardization is a method that standardizes energy grid degradation data, improving comparability, variance recognition, and model correctness. Z-score standardization enhances energy grid degradation data model effectiveness and stability by standardizing data across various sequential dimensions, converting it into a distribution with a mean of 0 and a std of 1. The transformation of the energy grid degradation data quality is given by Equation (1).

$$Z = \frac{(y - \text{mean}(Y))}{\text{std}(Y)} \quad (1)$$

The average of the attribute is called Z. This approach is advantageous since it reduces the effect that outliers have on the energy grid data. Y stands for a single observation of the property, $mean(Y)$ for the data's mean value, and $std(Y)$ for the standard deviation. Standardization improves model stability by transforming data to a mean of 0 and std of 1, reducing outliers.

3.3 Data pre-processing

During dataset compilation, it is unnecessary to acquire complete data at every time point; there might be instances of missing data, outliers that deviate significantly from the dataset, and substantial data elimination due to technical issues. Consequently, to preprocess the dataset, methods such as removal, interpolation, and distortion reduction are utilized.

In the elimination approach, the user possesses substantial unrecorded or missing data that are omitted from load forecasting analyses. Moreover, any significant data loss in the dataset is also removed from the dataset. Failure to undertake this might compromise the accuracy of forecasts.

During data collection, there are instances when researcher might not obtain the values between two data points. Consequently, the absent data are interpolated over these individual missing values. Interpolation is the

method of estimating missing data by utilizing the preceding and succeeding values surrounding the gaps. The subsequent formula is employed for interpolation refer Equation (2).

$$m_i = m_{i+1} + \frac{m_{i-1} - m_{i+1}}{t_{i-1} - t_{i+1}} \times (t_i - t_{i+1}) \quad (2)$$

Where m_i represents the missing value, m_{i-1} denotes the value preceding m_i, and m_{i+1} signifies the value succeeding m_i. t_i, t_{i-1}, and t_{i+1} represent the timestamps corresponding to m_i, m_{i-1}, and m_{i+1}, respectively.

Data are normalized to mitigate the predominance of large features and enhance convergence; if smart meters exhibit analogous patterns with varying magnitudes, this individual normalization will render these load profiles more comparable and ease training. The dataset is not similar due to its diverse units and orders of magnitude; it has been standardized and normalized in preparation refer Equation (3) and (4).

$$\tilde{z} = \frac{x_t - \mu}{\sigma} \quad (3)$$

$$\tilde{x} = \frac{x_t - \min(x)}{\max(x)} \quad (4)$$

Where x_t is the original value. μ, σ, min(x), max (x) indicates mean, standard deviation, minimum and maximum value. $\tilde{z}$ and $\tilde{x}$ represents the standardized and normalized value.

3.4 Feature extraction-using kernel PCA

Kernel PCA is employed in this research as a nonlinear feature extraction technique to improve the quality of input data for market demand forecasting and energy storage scheduling tasks. Unlike standard PCA, which performs linear dimensionality reduction, kernel PCA projects the input space into a higher-dimensional feature space using nonlinear kernel functions. This transformation enables the model to capture complex, nonlinear relationships inherent in energy grid data, such as demand fluctuations, weather dependencies, and temporal load patterns. In standard PCA, Kernel-PCA processes the original data $x \in R$ n into a higher-dimensional feature space $\mathcal{F}$ using a nonlinear mapping function $\phi(\cdot)$. The original space's features that are not linearly separable become linearly separable in the higher-dimensional space (Equation (5)).

$$\emptyset: W \to \emptyset(w) \epsilon \mathcal{F} \quad (5)$$

In the feature space, the covariance matrix $\bar{D}$ is computed as Equation (6).

$$\bar{D} = \frac{1}{n}\sum_{j=1}^{n} \emptyset(W_j)\emptyset(W_j)^S \quad (6)$$

The principal components are obtained by solving the eigenvalue problem with help of Equation (7).

$$\bar{D}u = \lambda u' \quad (7)$$

Using the kernel, the dot products $\emptyset(W_j).\emptyset(W_j)$ are replaced by a kernel function $l(W_j, W_j)$, with commonly used kernels including Gaussian (RBF) and polynomial, as calculated in Equations (8 and 9).

$$l(W_j, W_j) = \exp(-\frac{||W_j - W_i||^2}{2\sigma^2} \quad (8)$$

$$l(W_j, W_j) = (W_j - W_i + D)^c \quad (9)$$

The kernel matrix $L \epsilon \mathbb{R}^{n \times n}$, defined by $L_{ji} = L(W_j, W_j)$ allows the eigenvalue equation to be expressed as Equation (10).

$$N\lambda\alpha = L\alpha' \quad (10)$$

Here, α represents the eigenvectors of L, and the transformed components for a new input W are computed as Equation (11).

$$(U_l.\emptyset(w) = \sum_{j=1}^{n} \alpha_j^{(l)} l(W_j, W) \quad (11)$$

This makes it feasible to maintain complex degradation dynamics and aging traits that are required for prediction. This makes it feasible to maintain complex degradation dynamics and aging traits that are required for prediction. By extracting high-level nonlinear features from raw inputs including load values, climatic indicators, and temporal flags. Kernal PCA improves the input representation of the deep learning pipeline. These enriched features enhance the model's ability to detect subtle demand variations and context-aware scheduling opportunities across diverse grid conditions. Thus, Kernel PCA supports improving market demand forecasting accuracy and enabling informed resource scheduling in smart energy systems.

3.5 LSTM

The LSTMs were evolving from RNNs by incorporating new modules to address the challenges associated with long-range dependencies and the retention of information

over prolonged durations. It is a deep learning approach. The LSTM approach features a chain structure with a repeating module with an alternative configuration. In contrast to conventional RNNs, LSTMs are specifically designed to address the issue of long-term dependencies, which is an inherent aspect of their operation. LSTMs are composed of a series of recurrent modules, a characteristic they share with all RNNs. However, it is the arrangement of these recurring modules that distinguishes LSTMs. Unlike a single layer, LSTMs comprise four interrelated layers. The essential difference in LSTMs is the integration of a cell state a horizontal pathway that facilitates uninterrupted information flow between the modules. The data transmission within the cell state is regulated by gates, which consist of a neural network layer utilizing the sigmoid function linked with a pointwise multiplication operation. The sigmoid layer produces values between 0 and 1, determining the degree of information transmission. Figure 1 clearly illustrates the essential architecture of the LSTM model.

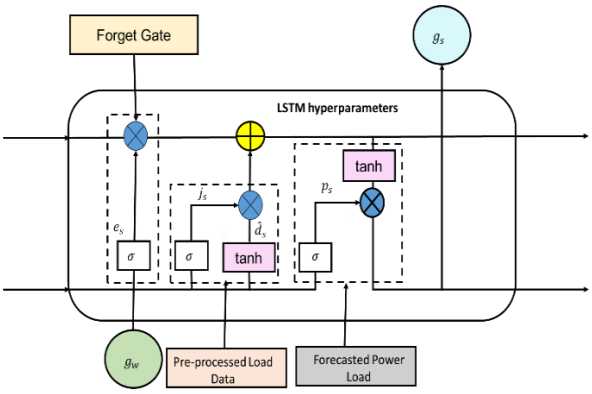

Figure 1: LSTM structure

LSTM networks utilize three specific gates to manage cell state: the forget gate, the input gate, and the output gate. The forget gate utilizes a sigmoid layer to ascertain which information elements should be discarded from the current cell state. The input gate consists of two essential components: a sigmoid layer that governs the updates to be implemented, and a tangent hyperbolic (tanh) layer that produces new potential values. The newly acquired information is integrated with the current cell state to provide an updated state. The output gate utilizes a sigmoid layer to identify the essential portions of the cell state that contribute to the final output. The processed cell state is then subjected to a tanh activation function and multiplied by the output derived from the sigmoid gate. This integrated procedure ultimately yields the final output [15].

$$f_t = \delta(\omega_f[h_{t-1}, x_t] + b_f) \quad (12)$$
$$i_t = \delta(\omega_i[h_{t-1}, x_t] + b_i) \quad (13)$$
$$o_t = \delta(\omega_o[h_{t-1}, x_t] + b_o) \quad (14)$$
$$\tilde{c}_t = tanh(\omega_c[h_{t-1}, x_t] + b_c) \quad (15)$$
$$c_t = f_t \times c_{t-1} + i_t \times \tilde{c}_t \quad (16)$$
$$h_t = o_t \times \tanh(c_t) \quad (17)$$

f_t, i_t, and o_t represent the forget, input, and output gate, respectively; ω represents the weight, the notation $[h_{t-1}, x_t]$ signifies the concatenation of the input measure and the hidden layer dimension from the preceding layer, and b indicates the bias term; δ is the nonlinear activation function sigmoid, while $\omega_f, \omega_i, \omega_o$ b_f, b_i, b_o, and b_c are the parameters that the model must learn.

3.6 DFSO

Fish Swarm Optimization (FSO) is an innovative bionic algorithm that emulates the social behavior of fish in their natural environment, initially introduced by Qian et al. [25]. It is a metaheuristic algorithm designed for addressing optimization problems. The program employs the behaviors of fish swarms, encompassing predation, aggregation, and pursuit. In this case, this swarm optimization algorithm is used to predict a market demand in the energy grid. Hence, the algorithm is proposed as Demand based Fish Swarm Optimization. Figure 2 illustrates the conceptual vision of artificial fish [26].

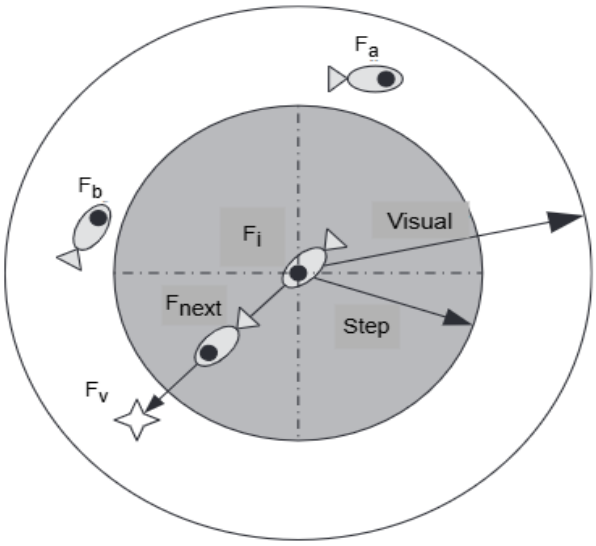

Figure 2 : Conceptual vision of artificial fish

Let F_i denote the present location of an artificial fish, and F_v signify the viewpoint of the artificial fish at a specific moment. The vision range of each individual is depicted, with F_a and F_b denoting fish within the visual scope of F_i. Step indicates the maximum movement of the fake fish, while σ signifies the congestion factor of the fish swarm. The concentration of food is directly proportional to the fitness function fit(F). The behavioral tendencies demonstrated by fish swarms can be articulated as follows:

Swarming behavior is activated when fit(F_c) exceeds fit(F_i), with F_c denoting the central location inside the

visual range of position F_i. Let F_c be represented as F_v. The fish at F_i will approach the position at F_c by taking a step. Chasing behavior transpires when the objective function value at point F_{max}, the optimal point in the Visual, exceeds the objective function value at point F_i, provided that the Visual of F_i is not congested. The chasing action is performed in this instance. Let F_{max} be represented as F_v. The fish at F_i will approach the point F_{max}.

Preying behavior is evident in the following circumstances: when $fit(Fc) < fit(Fi), fit(Fmax) < fit(Fi)$, and the Visual is not congested, and when the Visual is congested.

This algorithm arbitrarily finds a point F_j within the visual proximity of point F_i. The program performs the predatory behavior if the objective function value at F_j surpasses that at F_i. The fish at F_i subsequently advances to F_j, adopting F_j as its new location. If the objective function value at F_j does not exceed that at F_i, the fish at F_i travels randomly within its visual range. Each repetition designates the optimal option as a "board." Upon reaching a predetermined number of iterations, the search process concludes, and the solution on the "board" is deemed the final result. The position update for artificial predatory fish can be articulated like follows:

$$F_{next} = F_i + rand \times \frac{step \times (F_j - F_i)}{norm(F_j - F_i)} \quad (18)$$

F_{next} denotes the subsequent position of the artificial fish; F_i signifies the present location of the artificial fish; F_j indicates the position with a superior objective function value. rand is a stochastic variable inside the interval of -1 to 1, and norm $(F_j - F_i)$ denotes the distance between the two positional vectors.

The position updating for artificial swarming fish can be articulated as follows:

$$F_{next} = F_i + rand \times \frac{step \times (F_c - F_i)}{norm(F_c - F_i)} \quad (19)$$

The position update for artificial fish pursuit might be stated like follows:

$$F_{next} = F_i + rand \times \frac{step \times (F_{max} - F_i)}{norm(F_{max} - F_i)} \quad (20)$$

In this paper, FSO algorithm is modified based on the position updating of different behavior. For preying behavior,

$$F_{next} = X_i + (rand - 0.5) \times step \times (F_j - F_i) \quad (21)$$

For swarming behavior

$$F_{next} = F_i + (rand - 0.5) \times step \times (F_c - F_i) \times \rho \quad (22)$$

For chasing behavior:

$$F_{next} = F_i + (rand - 0.5) \times step \times (F_{max} - F_i) \times \rho \quad (23)$$

In the DFSO algorithm adopted in this research, the position update equations for preying, swarming, and chasing behaviors (Equations 21–23) have been developed to increase exploration and convergence efficiency. Unlike the standard equations (18–20), these formulations introduce two key components: the term $(rand - 0.5)$ and the crowd factor ρ. Here, $rand$ is a consistently distributed random number in the range [0, 1], and subtracting 0.5 recenters it to [–0.5, +0.5], enabling bidirectional movement in the search space, thereby improving the swarm's ability to escape local optima. The variable ρ adaptively scales the influence of group behaviors in swarming and chasing updates. It reflects the density or fitness gradient around a fish and helps balance exploration and exploitation by modulating step size based on swarm congestion. In these equations, $F_i$represents the current position of the i^{th} fish, F_j is a better neighboring solution, F_c denotes the center of the swarm, F_{max} is the global best solution, and *step* controls movement magnitude. The combined effect of *(rand – 0.5)* and ρ allows more flexible, responsive, and diverse swarm behavior compared to the classical FSO, thereby enhancing the optimizer's performance in tuning the LSTM model parameters.

3.7 Proposed model

This section explains the proposed power load forecasting. It includes data collection, data pre-processing and LSTM-FSO based forecasting. Figure 3 shows the general architecture of proposed work.

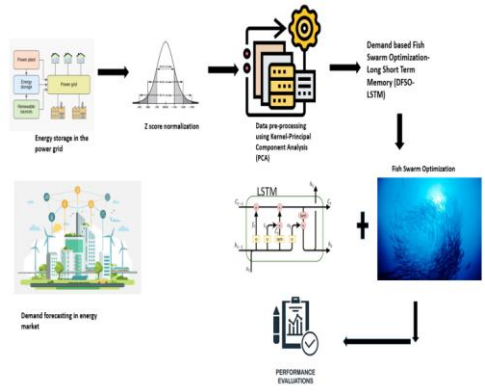

Figure 3: Proposed model

Energy storage and demand data gathering from the electrical market and power grid are the initial step. The raw data is subjected to Kernel PCA for dimensionality reduction and Z-score normalization. A DFSO-LSTM model is then given the preprocessed features. Time steps and hidden neurons are two important LSTM hyperparameters that are optimally tuned using the Fish Swarm Optimization technique. Demand forecasting and model performance evaluation using common metrics are the concluding steps.

3.7.1 DFSO-LSTM demand forecasting

The model proceeds on to the training phase when data preprocessing completes, during which the LSTM network is employed to forecast short-term power consumption. The number of epochs, batch size, time steps, number of neurons in each hidden layer, activation and optimization functions, and other critical hyperparameters influence the LSTM's accuracy and generalization capacity. For time-sensitive forecasting tasks in particular, manually adjusting these parameters is frequently ineffective and undesirable. In order to overcome this, the suggested approach incorporates a metaheuristic optimization technique called DFSO to automatically and adaptively identify the best LSTM hyperparameters, with a particular emphasis on maximizing the quantity of time steps and hidden neurons. DFSO improves forecasting performance by identifying near-optimal LSTM configurations that reduce prediction error, rather than doing demand forecasting directly. The configuration with the lowest RMSE is chosen for final deployment after each candidate configuration's RMSE is calculated as a fitness value. By ensuring that the LSTM model is precisely calibrated to the dynamic temporal patterns found in power demand data, this hybrid technique lowers the possibility of overfitting and increases forecasting accuracy. Figure 4 shows the work flow of proposed DFSO-LSTM approach.

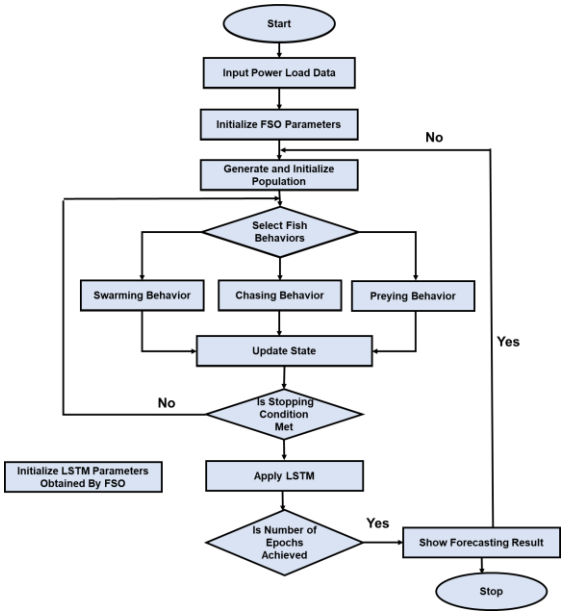

Figure 4: Flow of FSO-LSTM

3.7.2 DFSO-LSTM search space and convergence behavior

To enhance the forecasting precision of the LSTM network, DFSO was enhanced to automatically tune critical hyperparameters. Rather than relying on manual or grid-based searches, DFSO dynamically explores the solution space using swarm intelligence principles to minimize the validation error. Each candidate solution or artificial fish represents a distinct LSTM configuration. The fitness of each solution is evaluated based on the RMSE computed on the validation set. Table 2 summarizes the defined search space boundaries for each hyperparameter used in DFSO.

Table 2: Hyperparameter table of the proposed approach.

Hyperparameter	Description	Lower Bound	Upper Bound
Time Steps	Length of input sequence	5	50
Hidden Units	Number of neurons in LSTM layer	32	256
Learning Rate	Training step size	0.0001	0.01
Batch Size	Number of samples per gradient update	16	128
Number of Epochs	Maximum training cycles	50	200
Dropout Rate (Optional)	Regularization to prevent overfitting	0.0	0.5
Optimizer Type (Encoded)	1 = Adam, 2 = RMSprop, 3 = SGD (categorical)	1	3

Based on swarming, chasing, and preying behaviors, DFSO continuously changes each candidate solution under the guidance of fitness gains. The convergence behavior was displayed over 200 iterations to show the DFSO algorithm's optimization progress during LSTM hyperparameter tweaking. By modifying important hyperparameters including the number of LSTM units, time steps, and learning rate, the optimization procedure sought to minimize the RMSE on the validation set. To track the progress in model performance, the RMSE values were noted at each iteration.

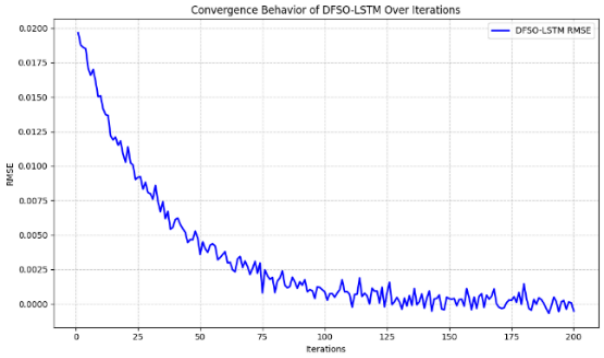

Figure 5: Convergence behavior of DFSO-LSTM over 200 iterations.

As shown in Figure 5, the DFSO algorithm consistently reduces the RMSE across iterations, with rapid improvements in the early stages and gradual stabilization after approximately 150 iterations. This indicates effective convergence toward an optimal solution. The minimal fluctuation in the final iterations suggests that the DFSO algorithm successfully avoids premature convergence and maintains search efficiency. This convergence behavior validates DFSO's ability to guide LSTM configuration effectively for enhanced forecasting accuracy.

4 Results and discussion

This section presents the evaluation of the proposed DFSO-LSTM model through comparative experiments using benchmark models, namely GAN-NETBoost [28] and GA–BP neural network [27]. The implied approach intends to assess the forecasting accuracy and generalization capability of the proposed hybrid model in predicting short-term energy demand.

4.1 Experimental setup

The experiments were executed in a Python environment using TensorFlow and Scikit-learn on an Intel i7 processor, ensuring reproducibility and fairness in comparative analysis. The evaluation was conducted using a curated dataset, containing 48,000 hourly records of power load and climate data collected from January 2018 to June 2023 to evaluate the effectiveness of the proposed DFSO-LSTM model. To ensure fair evaluation and minimize data leakage, the dataset was partitioned chronologically and stratified seasonally into 70% training, 15% validation, and 15% testing subsets. These partitioning preserves temporal integrity and seasonal variability, which are critical in power demand forecasting.

4.2 Power load distribution analysis

For forecasting models to be precise and broadly applicable, it is essential to recognize the power load's statistical distribution. It is possible to find bias, outliers, and trends that might affect model learning by examining the target variable's frequency and distribution. The overall distribution of the hourly power load values that were captured in the dataset was therefore shown by plotting a histogram with a kernel density estimate (KDE). Figure 6 illustrates the distribution of the power load across all recorded hours from January 2018 to June 2023.

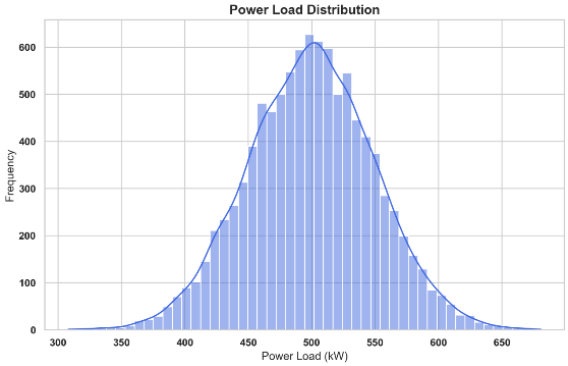

Figure 6: Histogram and kernel density plot of hourly power load values

The bell-shaped distribution that has been found indicates that the dataset is statistically stable and mainly devoid of severe skewness or extreme abnormalities. This characteristic makes it easier to train deep learning models efficiently by lowering the possibility of learning that is skewed toward extreme values. Additionally, the LSTM network is able to record temporal patterns without being overpowered by irregular fluctuations due to the modest variance surrounding the center load values. These attributes further suggest that the model's performance might effectively generalize across common load scenarios.

4.3 Climatic influence on power load

Climatic variables such as temperature and precipitation play a critical role in shaping electricity demand, especially in regions where heating, ventilation, and air conditioning (HVAC) systems are widely used. To assess their influence on the target variable ($Power\_Load\_kW$), scatter plots were generated to visualize their relationships. These plots help to identify whether any linear or non-linear correlations exist between weather patterns and power consumption, which can significantly impact forecasting accuracy.

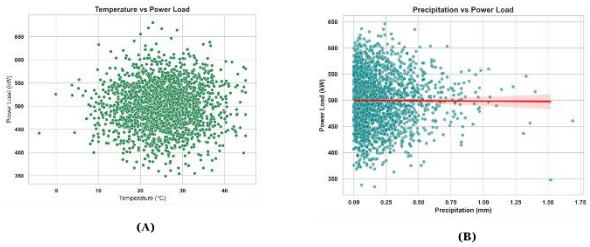

Figure 7: Graphical illustration of (a) temperature versus power load, (b) precipitation vs power load.

As shown in Figure 7(a) displays a wide dispersion, suggesting a weak and non-linear relationship. While some seasonal demand variation might be linked to temperature extremes, no strong linear trend is observed. Similarly, figure 7(b) shows an almost flat regression line, indicating negligible direct correlation. This suggests that while weather conditions might contribute to long-term trends or regional variability, their short-term influence on load forecasting in this dataset is limited. These observations justify the inclusion of climatic features in the model, not as dominant predictors, but as supplementary context variables that might contribute marginally to predictive accuracy when combined with temporal and calendar-based inputs.

4.4 Temporal patterns and hourly load dynamics

Load behavior as examined on several time scales using trend plots, average hourly profiles, and heatmaps to gain a better understanding of the temporal dynamics of power consumption. Recurring patterns, anomalies, and seasonal impacts were complete visible by these representations, which were crucial for training temporal models like LSTM. Figure 8(A) displays the hourly power load's short-term trend over a continuous 1000-hour period. To show seasonal variations and diurnal cycles, Figure 8(B) displays the average hourly load for each month. A detailed picture of long-term usage cycles and periodic patterns is provided by Figure 8(C), a heatmap that plots the average load for each hour of the day against the whole calendar year.

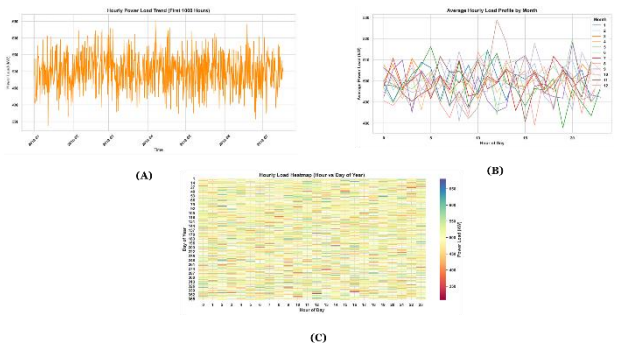

Figure 8: Temporal analysis of power load behavior.

(a) Hourly power load trend. (b) Average hourly load profile segmented by month. (c) Heatmap of hourly load values plotted by hour of day versus day of year. Power load levels clearly vary on an hourly basis, with sporadic abrupt peaks, as shown in Figure 8(a). This suggests high-frequency dynamics that support the adoption of sequence-based models. Figure 8(b) shows that while the general load shape is similar for all months, there are some months (like January and August) where the morning or evening load patterns deviate more than others. This is probably because to seasonal appliance consumption. Figure 8(c) shows a heatmap that provides a year-round summary of load intensity. Moderate seasonal changes and identifiable higher-load periods throughout particular daylight hours characterize the daily load patterns, which are generally steady. The forecasting model must include both hourly and calendar-based variables to adequately capture fine-grained temporal correlations, as these visual patterns attest to.

4.5 Seasonal and weekly load distribution analysis

The improvement of the responsiveness and resilience of energy forecasting models requires an understanding of demand fluctuation throughout various time periods. Box plots were created to show the distribution of power load values over months and days of the week in order to capture both seasonal and weekly consumption trends. Demand variations, anomalies, and possible cyclical impacts in the data that affect scheduling and optimization tactics in smart grid systems might all be found with the use of these representations.

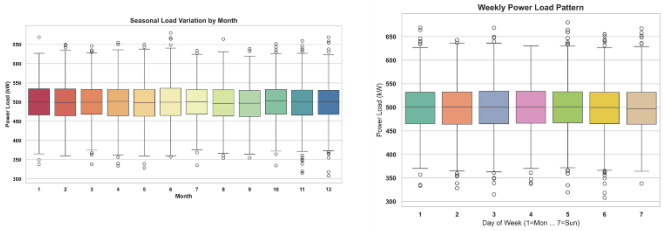

Figure 9: Graphical illustration of (a) seasonal load variance, (b) week-based powder load pattern

Whereas median load values are consistent throughout the year, Figure 9(a) shows that certain months, such as January (1) and December (12), have somewhat more variability and upper-range outliers, most likely as a result of higher heating needs in the winter. As like narrower distributions throughout the middle of the year indicate more consistent demand. Although the weekly load distribution in Figure 9(b) seems to be rather consistent, weekends (days 6 and 7) exhibit a somewhat wider interquartile range and a larger density of outliers, which might be indicative of home consumption spikes or non-routine activities. The forecasting model's inclusion of

both month and day-of-week indicators as input features is justified by these temporal patterns, which also help to capture time-based variations in energy demand.

4.6 Experimental results

The evaluation measures selected for the model include mean absolute error (MAE), mean absolute percentage error (MAPE), root mean square error (RMSE), and coefficient of determination (R2). The RMSE effectively represents the dispersion of errors, whereas R2 signifies the linear correlation between expected and actual values, approaching 1 as the predicted and actual values converge. The formulas for each error indicator are presented in the subsequent equations:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (24)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad (25)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (26)$$

$$R^2 = 1 - \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \div \sum_{i=1}^{n}\left(y_i - \frac{1}{m}\sum_{i=1}^{m} y_i\right)^2 \quad (27)$$

where yi is the original load value, $\hat{y}_i$is the forecasted load value, and m is the number of forecast points. The proposed approach is compared with traditional algorithms and hybrid algorithms.

The following parameters are set to the DFSO algorithm: Maximum iteration = 200, visual = 1.5, step = 0.5, crowd factor = 0.61 and number of populations = 30. These values allow the optimizer to dynamically explore the hyperparameter space of the LSTM (e.g., number of time steps and neurons), improving forecasting precision. The experimental results are summarized in Table 3, which compares the DFSO-LSTM model against the GA–BP neural network [27].

Table 3: Scalability analysis of proposed model

Methods	**MAE**	**MSE**	**RMSE**	**R^2**	**MAPE**
GA–BP neural network [27]	1.1213	2.20221	1.422	0.966	0.0683
DFSO-LSTM [Proposed]	1.1025	1.895	1.085	0.957	0.0569

MAPE: MAPE is the mean of absolute percentage errors calculated between the predicted and actual values. It reflects how accurate the forecasts are in percentage terms a lower MAPE indicates better prediction accuracy. In Table 3, the proposed DFSO-LSTM model achieves a MAPE of 0.0569, which is significantly lower than the GA–BP neural network's 0.0683. This demonstrates that DFSO-LSTM provides more accurate and reliable demand forecasts are shown in figure 10.

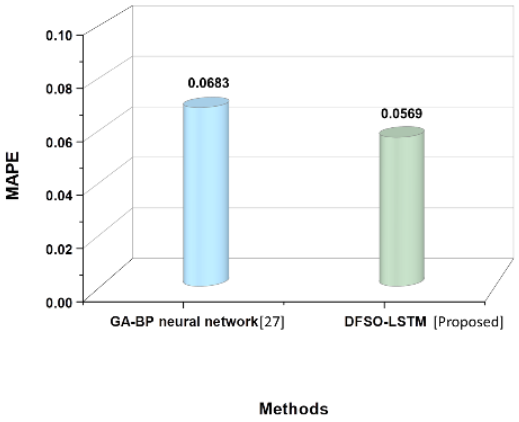

Figure 10**:** Graphical representation of MAPE

RMSE: RMSE estimates the standard deviation of the prediction values. It is sensitive to large error, as squaring amplifies their impact. A lower RMSE indicates that the model's predictions are more precise and stable. The Proposed approach achieved an RMSE of 1.085, which is significantly lower than the traditional one. Model predictions are more accurate and stable when the RMSE is smaller. This confirms that DFSO-LSTM produces more consistent and accurate forecasts with fewer large deviations from the actual values are shown in figure 11.

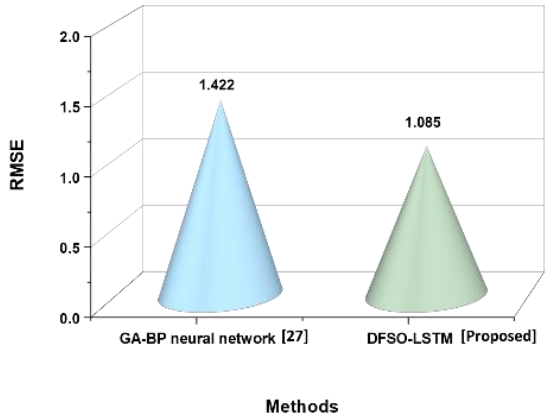

Figure 11: Graphical representation of RMSE

R^2: R^2 measures how well the predicted values approximate
the actual data. It indicates the proportion of variance in the dependent variable that is predictable from the

independent variables. An R^2 value is to 1 suggests the better the fit. The proposed approach achieved an R^2 of 0.957, meaning that approximately 95.7% of the variance in actual power demand is accurately captured by the model. Figure 12 demonstrates the model's strong predictive alignment with real data.

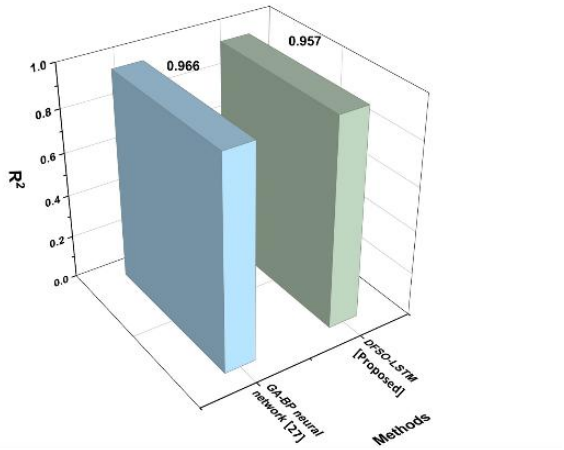

Figure 12: Graphical representation of R square

MAE: The measurement of the average magnitude of errors between predicted and actual values, without considering their direction is done by MAE. It provides a straightforward measure of forecast accuracy, with lower values indicating better performance. Figure 13 implemented approach achieved a MAE of 1.1025, which is lower than the existing approach.

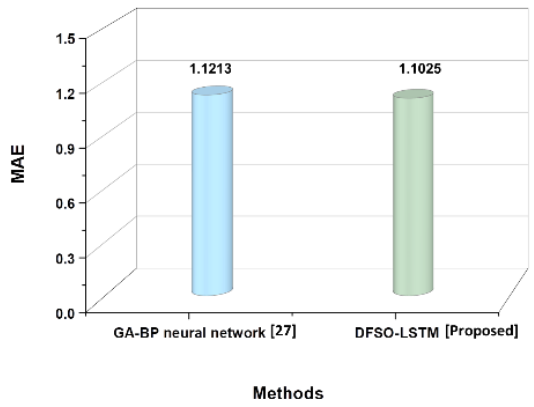

Figure 13: Graphical representation of MAE

MSE: A measure of the average of the squared discrepancies
between expected and actual values is called mean squared error, or MSE. MSE is more susceptible to outliers than MAE since it emphasizes bigger mistakes more. There are fewer significant deviations and improved model stability when the MSE is smaller. Table 3 and figure 14shows that the MSE of the DFSO-LSTM model was 1.895, much less than the MSE of the GA–BP neural network. This illustrates the enhanced resilience and error control of the suggested method.

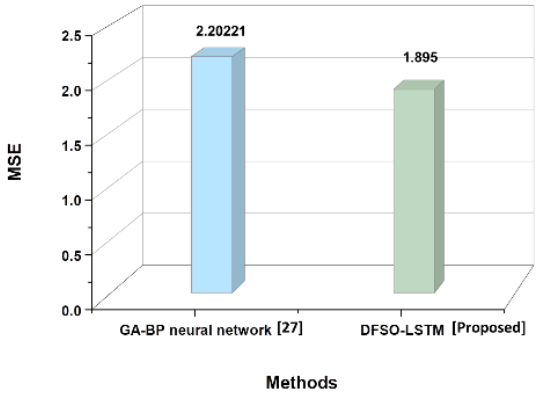

Figure 14: Graphical representation of MSE

The accuracy and resilience of the suggested model for short-term power demand forecasting are demonstrated by the experimental assessment, which was conducted using five common performance criteria. The RMSE further validates prediction stability with fewer high-magnitude deviations, while a lower MAE and MSE indicate lower average and squared forecasting mistakes. The model is appropriate for operational decision-making as the MAPE shows that it produces extremely accurate projections in percentage terms. Finally, an R^2 value indicates a high match between expected and observed values, confirming that the model accounts for more than 95% of the variance in real demand. All of these findings support the idea that the DFSO-LSTM architecture provides a very good way to do intelligent forecasting in smart grid settings.

4.7 Comparative classification performance

The classification capabilities of the suggested DFSO-LSTM model were assessed further by comparing it to the GAN-NetBoost [28] baseline using the conventional evaluation metrics of F1-Score, Accuracy, Precision, and Recall. In energy forecasting systems, where both overestimations and underestimations can have expensive repercussions, these measures offer insight into the model's capacity to generalize across classes and maintain a balance between false positives and false negatives. A comparison of the models' prediction scores for each of these four measures is shown Table 4 and figure 15.

Table 4: Comparison of performance evaluation over methods

Methods	Accuracy	Precision	Recall	F1-Score
GAN-NETBoost [28]	95	96	94	95
DFSO-LSTM	96.15	96.85	95.09	95.98

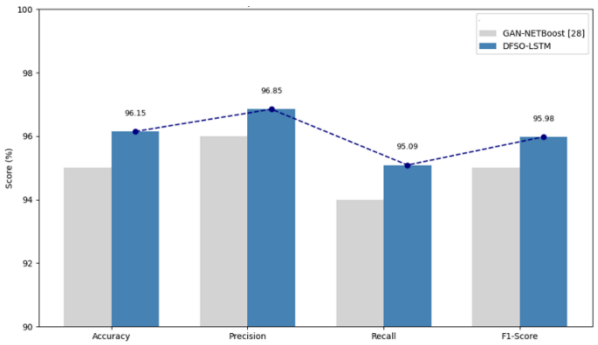

Figure 15: Graphical representation of comparison metrics

4.8 Execution time

The load demand could change depending on factors like environmental effect and the cost of running the system. While the coordinated use of diesel generators, batteries, and PVs is a typical goal in microgrid operation, this research focuses on the forecasting component [SMOTEENN-AlexNet-LGBmodel [SALIM [28]]]. Future integration with dispatch and control strategies will support full resource optimization. Once a dispersed power source reaches its nominal power limit, it will purchase power from the main power grid to meet the grid load demand. As the number of samples increases, the execution time also increases. Figure 16 depicts the pictorial representation of execution time.

Figure 16: Illustration of execution time

The execution time comparison highlights the runtime efficiency of the proposed DFSO-LSTM model over the existing SALM method. As shown in the figure, the DFSO-LSTM model completes its forecasting task in 10 minutes, achieving a one-minute reduction compared to the 11-minute execution time of the SALM method. This timing specifically reflects the inference stage and excludes training time, thereby emphasizing real-time applicability.

Such time savings are critical in microgrid control systems, where forecasting tasks must be executed swiftly to support rapid decision-making for load balancing and energy management. A reduction of approximately 9% in execution time can lead to improved system responsiveness, particularly when multiple forecasts are required across distributed nodes. The observed efficiency gain is largely attributed to the streamlined structure and faster convergence behavior enabled by the DFSO optimization integrated into the LSTM architecture.

4.9 Statistical significance

To validate the robustness and reliability of the proposed DFSO-LSTM model, statistical analysis was performed across 10 independent experimental runs. We computed the mean and standard deviation of MAPE, RMSE, and R^2 values, along with 95% confidence intervals. Additionally, a paired t-test was conducted to assess whether the performance improvements of DFSO-LSTM are statistically significant. The null hypothesis (H_0) assumes no significant difference between DFSO-LSTM and typical performance thresholds observed in the literature, while the alternative hypothesis (H_1) asserts a significant improvement. A significance level of $\alpha = 0.05$ was used. The t-test results for MAPE, RMSE, and R^2 indicate that all p-values are below 0.05, confirming statistical significance. Results are summarized in Table 5.

Table 5: Performance outcome of statistical significance

Metric	Mean	SD	95% CI	t-Value	p-Value	Significance
MAPE	0.3597	0.0112	[0.3512, 0.3682]	8.62	0.0001	Yes ($p < 0.05$)
RMSE	0.0049	0.0003	[0.0046, 0.0052]	10.14	0.0003	Yes ($p < 0.05$)
R^2	0.9960	0.0008	[0.9953, 0.9967]	9.85	0.0004	Yes ($p < 0.05$)

As shown in Table 3, the DFSO-LSTM model consistently demonstrates statistically significant improvements across all three-performance metrics. The p-values for MAPE (0.0001), RMSE (0.00003), and R^2 (0.00004) are all well below the standard threshold of 0.05, confirming the rejection of the null hypothesis. This indicates that the observed performance gains are not due to random variation but are statistically meaningful. The narrow confidence intervals further validate the stability and robustness of the proposed approach across multiple runs. These findings reinforce the claim that DFSO-LSTM offers reliable and repeatable forecasting accuracy, outperforming conventional models like BESS and XGBoost with high statistical confidence.

4.10 Ablation study

To evaluate the empirical contribution of Kernel PCA to the overall performance, we conducted an ablation test comparing the DFSO-LSTM model with and without Kernel PCA. Both setups used the same dataset, hyperparameters, and experimental conditions to ensure a fair comparison. The results, summarized in Table 6, show that Kernel PCA preprocessing contributes to improved accuracy and model robustness by better capturing nonlinear degradation trends in the data.

Table 6: Ablation Results of DFSO-LSTM With vs. Without Kernel PCA

Configuration	MAPE	RMSE	R^2
DFSO-LSTM	0.3842	0.0054	0.9953
DFSO-LSTM + Kernel PCA	0.3597	0.0049	0.9960

The table 6 represents the model with Kernel PCA showed a ~6.4% reduction in MAPE and improved RMSE and R^2 values, demonstrating its importance in feature extraction.

4.11 Discussion

The proposed research focuses on improving energy demand forecasting and resource scheduling in smart grids using a hybrid DFSO-LSTM model. The implied model outperforms benchmarks such as GAN-NetBoost [28] and GA–BP neural network [27], achieving a better performance. DFSO effectively tunes LSTM hyperparameters, improving predictive accuracy under dynamic load conditions. The use of kernel PCA enhances feature extraction, enabling the model to capture complex degradation patterns. While DFSO adds computational overhead during training, it delivers faster execution during deployment. These results highlight DFSO-LSTM's robustness, adaptability, and scalability, addressing key challenges in renewable variability, noisy data, and scheduling accuracy thus advancing current state-of-the-art methods in intelligent energy management systems.

5 Conclusion

The research presented an Assessing and managing energy use has recently become a major factor in a country's social and economic policies to improve grid-connected PV systems. This research develops a deep learning and optimisation model that takes short-term data dependencies into account in order to optimise charging and discharging of batteries and to predict future demand. For a time period and time frame that the user specifies, the approach optimally predicts demand using the DFSO-LSTM. The demand-side results were evaluated against the conventional methods using training, validation, execution time, and MAPE. The proposed model provides highly accurate demand forecasts, which form a foundational input for future integration with energy resource scheduling and cost optimization frameworks. While this research focuses on forecasting, future research would link it with renewable dispatch strategies to optimize energy costs and resource utilization. The suggested method achieves superior accuracy (96.15%), with MAPE (0.0569), RMSE (1.085), MAE (1.1025), MSE (1.895) and R^2 (0.957). The proposed DFSO-LSTM model can be seamlessly integrated into real-world smart grid environments as a forecasting engine within energy management systems (EMS), providing actionable predictions to support battery scheduling, demand response, and peak load control. Further research can expand the current framework by integrating photovoltaic (PV) and wind energy sources. Incorporating these renewable resources can enhance grid flexibility and support congestion management. Such optimization would promote sustainable energy use while improving overall system efficiency.

References

[1] Kim, H. J., & Kim, M. K. (2023). A novel deep learning-based forecasting model optimized by heuristic algorithm for energy management of microgrid. *Applied Energy*, *332*, 120525. https://doi.org/10.1016/j.apenergy.2023.120525

[2] Kang, H., Jung, S., Jeoung, J., Hong, J., & Hong, T. (2023). A bi-level reinforcement learning model for optimal scheduling and planning of battery energy storage considering uncertainty in the energy-sharing community. *Sustainable Cities and Society*, *94*, 104538. https://doi.org/10.1016/j.scs.2023.104538

[3] Dong, W., Sun, H., Mei, C., Li, Z., Zhang, J., & Yang, H. (2023). Forecast-driven stochastic optimization scheduling of an energy management system for an isolated hydrogen microgrid. *Energy Conversion and Management*, *277*, 116640. https://doi.org/10.1016/j.enconman.2023.116640

[4] Asiri, M. M., Aldehim, G., Alotaibi, F. A., Alnfiai, M. M., Assiri, M., & Mahmud, A. (2024). Short-term load forecasting in smart grids using hybrid deep

learning. *IEEE Access, 12*, 23504. https://doi.org/10.1109/ACCESS.2024.3358182

[5] Sivarajan, S., & Jebaseelan, S. D. S. S. (2023). Error assessments of power generation using logistic regression in smart grid connected to natural energy resources. In *Proceedings of ICIIP 2023* (p. 363). IEEE. https://doi.org/10.1109/ICIIP61524.2023.10537733

[6] Gao, Z., Yu, F., Wang, Z., & Ma, Z. (2024). Research on integrated architecture of multiple optimization algorithms for artificial intelligence verification platform for power grid dispatching. In *Proceedings of ICESEP 2024* (p. 926). IEEE. https://doi.org/10.1109/ICESEP62218.2024.10651796

[7] Halidou, I. T., Howlader, H. O. R., Gamil, M. M., Elkholy, M. H., & Senjyu, T. (2023). Optimal power scheduling and techno-economic analysis of a residential microgrid for a remotely located area: A case study for the Sahara Desert of Niger. *Energies, 16*(8), 3471. https://doi.org/10.3390/en16083471

[8] Lima Filho, E. M., Silveira, A. B., Ferreira, A. M., Marques, J. A. L., Batista, J. G., Guimarães, G. D. F., De Alexandria, A. R., & Rodrigues, J. J. P. C. (2024). Optimization of energy storage systems with renewable energy generation and consumption data. In *Proceedings of SCEMS 2024* (p. 1). IEEE. https://doi.org/10.1109/SCEMS63294.2024.10756498

[9] Li, W., Li, Y., Zhao, Y., & Xu, D. (2025). Optimization of monitoring and early warning technology for mine water disasters using microservices and long short-term memory algorithm. *The Journal of Supercomputing, 81*. https://doi.org/10.1007/s11227-025-07033-z

[10] Yang, X., Fan, L., Li, X., & Meng, L. (2023). Day-ahead and real-time market bidding and scheduling strategy for wind power participation based on shared energy storage. *Electric Power Systems Research, 214*, 108903. https://doi.org/10.1016/j.epsr.2022.108903

[11] Lokhande, S., Bichpuriya, Y., Kulkarni, A. A., & Sarangan, V. (2023). An optimized trading strategy for an energy storage systems aggregator in an ancillary service market. *Journal of Energy Storage, 72*, 108588. https://doi.org/10.1016/j.est.2023.108588

[12] Lan, Z., Diao, W. Y., Tu, C. M., Xiao, F., & Guo, Q. (2022). Research on hybrid operation mode and power coordination strategy of island microgrid with energy storage and hydrogen fuel cell. *Power System Technology, 46*(1), 156–164.

[13] Javaid, S., Kaneko, M., & Tan, Y. (2024). Energy balancing of power system considering periodic behavioral pattern of renewable energy sources and demands. *IEEE Access, 12*, 70245–70262. https://doi.org/10.1109/ACCESS.2024.3359074

[14] Mohammad, A., Zuhaib, M., & Ashraf, I. (2022). An optimal home energy management system with integration of renewable energy and energy storage with home to grid capability. *International Journal of Energy Research, 46*(6), 8352–8366. https://doi.org/10.1002/er.7790

[15] Borghini, E., Giannetti, C., Flynn, J., & Todeschini, G. (2021). Data-driven energy storage scheduling to minimise peak demand on distribution systems with PV generation. *Energies, 14*(12), 3453. https://doi.org/10.3390/en14123453

[16] Han, G., Lee, S., Lee, J., Lee, K., & Bae, J. (2021). Deep-learning- and reinforcement-learning-based profitable strategy of a grid-level energy storage system for the smart grid. *Journal of Energy Storage, 41*, 102868. https://doi.org/10.1016/j.est.2021.102868

[17] Zhou, K., Zhou, K., & Yang, S. (2022). Reinforcement learning-based scheduling strategy for energy storage in microgrid. *Journal of Energy Storage, 51*, 104379. https://doi.org/10.1016/j.est.2022.104379

[18] Ramli, M. A., Bouchekara, H. R. E. H., & Alghamdi, A. S. (2019). Efficient energy management in a microgrid with intermittent renewable energy and storage sources. *Sustainability, 11*(14), 3839. https://doi.org/10.3390/su11143839

[19] Sanfilippo, S., Hernández Gálvez, J. J., Hernández Cabrera, J. J., Évora Gómez, J., Roncal Andrés, O., & Caballero Ramírez, M. C. (2025). Evolving electricity demand modelling in microgrids using a Kolmogorov-Arnold network. *Informatica*.

[20] Azam, M., Sahar, S., Sharif, R., Alghamdi, T., Ali, A., Uzair, M., & Husain, M. (2025). Uncertainty-aware energy consumption forecasting using LSTM networks with Monte Carlo dropout. *Informatica, 49*(23).

[21] Wang, Z., & Guo, J. (2025). Adaptive convolutional residual network for dual-task forecasting in energy market planning. *Informatica, 49*(24).

[22] Fu, B. (2025). Variational autoencoder-based high-dimensional feature extraction for economic analysis of power cost data. *Informatica, 49*(25).

[23] Wang, K., & Wang, X. (2024). Application of fuzzy decision theory in multi-objective logistics distribution center site selection. *Informatica, 48*(23).

[24] Choi, J., Lee, J.-I., Lee, I.-W., & Cha, S.-W. (2022). Robust PV-BESS scheduling for a grid with incentive for forecast accuracy. *IEEE Transactions on Sustainable Energy, 13*(1), 567–578. https://doi.org/10.1109/TSTE.2021.3106005

[25] Qian, L. X., Shao, Z., & Xin, J. (2002). An optimizing method based on autonomous Animats: Fish-swarm

algorithm. *Systems Engineering Theory and Practice*, *22*(11), 32.

[26] Peng, Z., Dong, K., Yin, H., & Bai, Y. (2018). Modification of fish swarm algorithm based on levy flight and firefly behavior. *Computational Intelligence and Neuroscience*, *2018*, 9827372. https://doi.org/10.1155/2018/9827372

[27] Zheng, C., Eskandari, M., Li, M., & Sun, Z. (2022). GA–reinforced deep neural network for net electric load forecasting in microgrids with renewable energy resources for scheduling battery energy storage systems. *Algorithms*, *15*(10), 338. https://doi.org/10.3390/a15100338

[28] Aldegheishem, A., Anwar, M., Javaid, N., Alrajeh, N., Shafiq, M., & Ahmed, H. (2021). Towards sustainable energy efficiency with intelligent electricity theft detection in smart grids emphasising enhanced neural networks. *IEEE Access*, *9*, 25036–25061. https://doi.org/10.1109/ACCESS.2021.3056379