AEM JOURNAL .ORG

Cost-Effectiveness Evaluation of Standby Power Configuration for Cluster Data Centers Based on Multi-Objective Optimization Algorithm Combined with Deep Learning

Qingsheng Li[1], Zhaofeng Zhang[1*], Jian Wang[1], Renren Du[1], Dongjunming Yang[1], Yu Zhang[1]
[1]Guizhou Power Grid Planning Center, Guiyang, 550003, China
Corresponding author: Zhaofeng Zhang (e-mail: 464620553@qq.com).

ABSTRACT Cluster data centers serve as the core carriers of the digital industry, and their standby power systems directly determine the power supply reliability and overall operational costs of data centers. Traditional standby power configuration methods mostly rely on empirical selection, which fail to simultaneously address multiple indicators including construction costs, operational energy consumption and power supply efficiency, nor adapt to the dynamic operational characteristics of clustered and large-scale data centers. This paper combines multi-objective optimization algorithms with deep learning technologies to establish an evaluation system for standby power configuration of cluster data centers, and conducts quantitative analysis centered on two core dimensions: cost and efficiency. First, this paper sorts out the operational characteristics and configuration requirements of standby power for cluster data centers. Second, a multi-objective optimization model is constructed, and the calculation rules for cost and efficiency indicators are defined. Third, deep learning algorithms are adopted to complete load forecasting and parameter fitting, and the optimization algorithm is combined to solve the optimal configuration scheme. Simulation verification and comparative analysis are carried out based on measured data, with four statistical tables presenting detailed data of different configuration schemes intuitively. The research results indicate that the evaluation method integrating the two technologies can effectively balance the investment cost and operational efficiency of standby power, providing practical references for the planning, type selection and operation & maintenance of standby power systems in cluster data centers.

INDEX TERMS Cluster Data Center; Standby Power Supply; Multi-Objective Optimization Algorithm; Deep Learning; Cost-Effectiveness Evaluation

I. INTRODUCTION

A. RESEARCH BACKGROUND

With the sustained rapid development of the digital economy, industries including internet services, cloud computing, big data and artificial intelligence keep expanding[1, 2]. The social demand for data storage, computing and network services has increased substantially, driving the construction of numerous large-scale data centers across the country[3]. Individual small-scale data centers have been gradually integrated and upgraded into cluster data centers[4]. Composed of multiple sub-data centers, cluster data centers feature large overall power consumption, frequent load fluctuations and non-stop operation, imposing extremely high requirements on the stability of power supply systems[5].

Mains electricity acts as the primary power supply for data centers. However, power outages, voltage fluctuations and line faults may occur to mains power[6]. Once the main power supply is interrupted, servers, storage devices and network equipment in data centers will stop operating immediately, resulting in data loss, service disruption and economic losses, and even cyber security incidents in severe cases[7, 8]. Therefore, standby power systems have become essential supporting facilities for cluster data centers[9]. At present, mainstream standby power equipment includes Uninterruptible Power Supply (UPS), diesel generator sets and energy storage batteries[10]. These devices differ significantly in procurement price, operating cost, power supply duration and response speed.

In practical engineering applications, two extreme situations frequently occur in standby power configuration. The first is over-configuration: excessive equipment quantity and redundant power capacity guarantee power supply safety but greatly increase procurement costs, site occupation costs and daily energy consumption[11]. The second is under-configuration with insufficient equipment capacity[12]. Although this method controls the initial investment, it fails to meet the full-load power demand during mains power failures, leading to substandard power supply efficiency. Traditional configuration methods depend on industrial experience and simple calculation formulas, which can only focus on either cost control or efficiency guarantee, and cannot achieve the coordinated optimization of multiple indicators[13, 14].

Meanwhile, the power load of cluster data centers changes dynamically, with obvious gaps between working days and holidays, as well as daytime and nighttime loads. Traditional static evaluation methods cannot accurately predict load changes, resulting in declining adaptability of configuration schemes. Deep learning boasts powerful capabilities in data mining and trend prediction, while multi-objective optimization algorithms can process multiple conflicting objectives simultaneously[15, 16]. Combining the two technologies for standby power configuration evaluation can address the drawbacks of traditional methods, which constitutes the main background of this research.

B. RESEARCH SIGNIFICANCE

Operators of cluster data centers need to control overall investment while ensuring the safe operation of equipment. The evaluation method proposed in this paper can quantify the cost and efficiency data of each standby power configuration scheme. Managers can select the optimal scheme based on evaluation results and reduce capital waste caused by blind type selection. This method is applicable to the power planning of newly-built cluster data centers as well as the renovation and upgrading of standby power systems in outdated data centers. After implementation, it can improve the power supply reliability of data centers and reduce the life-cycle operational costs, possessing high engineering practical value.

C. RESEARCH STATUS

Data center construction started earlier abroad, and relevant research has been carried out for a longer period. Some foreign scholars focus on energy consumption optimization of data centers and adopt intelligent algorithms to adjust the operation mode of power supplies to reduce electricity costs. Other researchers establish reliability evaluation models for standby power supplies, taking power supply duration and fault response speed as core indicators.

Driven by the vigorous development of computing infrastructure in China, the data center industry has entered a stage of rapid growth in recent years. Domestic scholars have conducted extensive research on energy conservation, power supply safety and equipment type selection for data centers. Many studies apply optimization algorithms to optimize the combination of UPS and energy storage equipment, and some researches combine big data technology to forecast power load.

D. RESEARCH CONTENTS AND TECHNICAL ROUTE

1) Research Contents

This paper consists of six core parts:
(1)Introduction: Elaborate on the research background, significance, domestic and overseas research status and overall framework.
(2)Basic theories: Explain the load characteristics of cluster data centers, types of standby power supplies, principles of multi-objective optimization algorithms and deep learning prediction models.
(3)Construction of cost-effectiveness evaluation indicator system: Classify cost and efficiency indicators and determine the calculation methods for each indicator.
(4)Establishment of integrated model: Substitute the load prediction results of deep learning into the multi-objective optimization model, set constraint conditions and clarify the solution process.
(5)Example analysis: Select an actual cluster data center as the research object, set up multiple groups of comparative schemes, present data via tables and conduct result analysis.
Conclusion and prospect: Summarize research achievements, point out the limitations of this study and propose directions for further improvement.

2) Technical Route

(1) Collect operational data of cluster data centers, parameter data of standby power equipment and relevant industrial standards and specifications.
(2) Sort out basic theories and determine evaluation indicators.
(3) Build a deep learning load prediction model to output accurate dynamic load data.
(4) Establish a multi-objective optimization mathematical model with the goals of minimum cost and optimal efficiency.
(5) Substitute predicted data into the model for solution and obtain multiple alternative configuration schemes.
(6) Collect statistics on cost and efficiency data of each scheme, draw comparative tables and carry out simulation analysis.
(7) Summarize research conclusions and put forward application suggestions.

E. RESEARCH INNOVATIONS

(1) Innovation in research object: This study takes cluster data centers as the research subject and considers the load characteristics of joint operation of multiple sites, differing from traditional researches focusing on single data centers. The proposed model is more in line with the current development status of the industry.
(2) Innovation in technical integration: The deep learning-based load prediction is taken as the pre-process to provide dynamic input data for the multi-objective optimization

model, realizing the full-process connection of "prediction - optimization - evaluation" and breaking the limitations of separate application of the two technologies.
(3) Innovation in evaluation system: An integrated two-way cost-effectiveness evaluation system is constructed to quantify both economic and operational indicators simultaneously, changing the previous single-dimensional evaluation mode and delivering more comprehensive evaluation results.

II. OVERVIEW OF BASIC THEORIES AND TECHNOLOGIES

A. OVERVIEW OF CLUSTER DATA CENTERS AND STANDBY POWER SUPPLIES

1) Operational Characteristics of Cluster Data Centers

A cluster data center is composed of multiple sub-data centers with adjacent geographical locations and unified scheduling and management, featuring three prominent characteristics: First, the total power load is large, and the equipment operates non-stop all year round, with an annual operating duration of nearly 8,760 hours. Second, the load is volatile: the load reaches the peak during business rush hours and drops significantly during off-peak periods, with prominent differences between day and night loads as well as peak and off-season loads. Third, the load is interactive: load can be allocated among sub-data centers. When a single site fails, other sites will share its services, leading to short-term sudden changes in the overall load.

The above characteristics determine that fixed standards cannot be adopted for standby power configuration in cluster data centers. Configuration schemes must be designed according to dynamic load changes to balance safety and cost.

2) Types and Characteristics of Mainstream Standby Power Supplies

Currently, three types of standby power supplies are widely used in cluster data centers: Uninterruptible Power Supply (UPS), diesel generator sets and electrochemical energy storage batteries.
1.UPS: It features the fastest response speed and can switch to power supply instantly after mains power failure, which is suitable for protecting precision electronic equipment such as servers against short-term power outages. It has a medium procurement cost but high no-load energy consumption, resulting in poor economic efficiency for long-term continuous operation.
2.Diesel Generator Set: It boasts large power supply capacity and long continuous power supply duration, making it applicable to deal with prolonged mains power failures. However, it has a startup delay ranging from several seconds to tens of seconds, and its operation generates fuel costs, maintenance costs, noise and exhaust emissions.
3.Energy Storage Battery: It operates with low noise and energy consumption and supports flexible charging and discharging, which can be used for load regulation. Nevertheless, it requires high initial procurement costs and has a limited service life, with considerable replacement costs in the later stage.

In practical engineering, combined modes of "UPS + diesel generator set" or "UPS + energy storage battery" are generally adopted, since a single type of equipment cannot meet all operational requirements.

B. BASIC THEORIES OF MULTI-OBJECTIVE OPTIMIZATION ALGORITHM

1) Basic Theories of Multi-Objective Optimization

Single-objective optimization only pursues the optimal result of one target, such as minimizing costs. Multi-objective optimization aims to optimize two or more mutually restricted objectives simultaneously.

In the standby power configuration problem, cost and power supply efficiency form a typical pair of conflicting objectives. Reducing configuration costs usually leads to decreased power supply efficiency, while improving power supply efficiency inevitably increases investment costs.

The core function of the multi-objective optimization algorithm is not to achieve the theoretical optimal value for every objective, but to find equilibrium solutions among multiple objectives and obtain schemes with optimal overall performance.

This paper adopts the multi-objective particle swarm optimization algorithm for calculation. With a simple structure and fast computing speed, this algorithm is suitable for solving engineering parameters.

2) Algorithm Solution Process

The solution process of the multi-objective particle swarm optimization algorithm includes five steps:
1.Determine optimization variables, namely the quantity, single-unit power and combination ratio of various standby power supplies in this research.
2.Establish objective functions corresponding to cost and efficiency goals respectively.
3.Set constraint conditions, including the upper limit of equipment power, site restrictions, industrial power supply standards and load matching requirements.
4.Initialize the particle population and set basic parameters such as iteration times and learning factors.
5.Conduct iterative calculations continuously, update the position and velocity of particles, screen the optimal solution set, and finally output feasible configuration schemes.

C. BASIC THEORIES OF DEEP LEARNING

1) Principle of Deep Learning-Based Load Prediction

The historical power load data of cluster data centers are affected by multiple factors including time, season, business volume and holidays. Traditional linear prediction methods cannot capture complex non-linear variation rules. Deep learning models can mine hidden rules in historical data through multi-layer network structures and predict real-time power load in different periods based on historical load sequences.

This paper adopts the Long Short-Term Memory (LSTM) network as the prediction model. The LSTM network is good at processing time-series data and can memorize load variation characteristics over a long time span, achieving higher prediction accuracy than traditional prediction methods.

2) Operation Steps of the Prediction Model

The deep learning prediction model operates in four steps:
1.Collect historical load data, time labels and environmental data, conduct data cleaning and normalization on raw data, and eliminate abnormal data.
2.Divide the dataset into training set, verification set and test set, build the LSTM network structure and set parameters such as network layers and neuron quantity.
3.Train the model with the training set data, continuously revise model parameters and reduce prediction errors.
4.Adopt the well-trained model to output dynamic load predicted values for different periods, which serve as the basic input data of the subsequent multi-objective optimization model.

III. CONSTRUCTION OF COST-EFFECTIVENESS EVALUATION INDICATOR SYSTEM FOR STANDBY POWER SUPPLIES OF CLUSTER DATA CENTERS

A. PRINCIPLES FOR CONSTRUCTING THE INDICATOR SYSTEM

To ensure scientific, reasonable and applicable evaluation results, this paper follows four basic principles to build the indicator system:
1.Comprehensiveness: Indicators shall cover economic and technical operational aspects to fully reflect the overall performance of configuration schemes.
2.Quantifiability: All indicators can be calculated via formulas and data, without ambiguous qualitative indicators that cannot be statistically measured.
3.Targetedness: Indicators shall conform to the operational characteristics of cluster data centers and standby power supplies, instead of adopting general power evaluation indicators.
4.Independence: Repeated calculation among indicators shall be avoided with clear logical relationships.

In accordance with the above principles, the evaluation system is divided into two primary indicators: cost indicator and efficiency indicator. Each primary indicator is further subdivided into secondary indicators to form a complete evaluation framework.

B. SELECTION AND CALCULATION METHODS OF COST INDICATORS

The life-cycle cost of standby power supplies is the core economic indicator of this evaluation, including four secondary indicators: initial procurement cost, installation and construction cost, operational energy consumption cost and later maintenance & replacement cost.

1) Initial Procurement Cost

It refers to the expenses incurred in purchasing all standby power equipment, calculated as the product of the unit price of single equipment and equipment quantity. Costs of UPS, diesel generator sets and energy storage batteries are calculated separately and then summed up. As a one-time investment, it constitutes the main part of the total cost of a scheme.

2) Installation and Construction Cost

This cost covers equipment hoisting, line laying, computer room renovation, fire supporting facilities and grounding facilities. It is positively correlated with the total power and quantity of equipment; the larger the equipment scale, the higher the construction cost.

3) Operational Energy Consumption Cost

This cost consists of two parts: the standby power consumption of equipment operating in standby mode all year round, and the energy consumption and fuel expenses generated when standby power supplies are put into operation during mains power failures. The annual operational energy consumption cost is adopted for statistics in this paper.

4) Later Maintenance & Replacement Cost

It includes expenses for routine inspection, component maintenance and replacement of aging equipment. Energy storage batteries have a fixed service life and must be replaced entirely upon expiration, and such expenses are counted separately. The annual average maintenance cost is converted and included in the life-cycle cost statistics. The total cost is the sum of the four secondary cost indicators, which is set as the minimization objective in multi-objective optimization.

C. SELECTION AND CALCULATION METHODS OF COST INDICATORS

The efficiency of standby power supplies reflects power supply capacity, reliability and adaptability. Four secondary efficiency indicators are selected in this paper: power supply reliability, load matching degree, fault response speed and continuous power supply duration. A higher value of each efficiency indicator represents better performance of the configuration scheme.

1) Initial Procurement CostPower Supply Reliability

It refers to the probability that standby power supplies can be normally put into operation when mains power is interrupted. Calculated in accordance with industrial power reliability standards, this indicator increases with the equipment redundancy and directly determines the safety level of data center equipment.

2) Load Matching Degree

It represents the matching level between the total power of standby power supplies and the dynamic load of data centers. The power of standby power supplies should neither be far lower than the real-time load nor excessively redundant. A smaller gap between power and load means a higher load matching degree and better resource utilization efficiency.

3) Fault Response Speed

It is measured by the switching time of equipment, namely the time interval from mains power disconnection to normal power supply of standby equipment. Shorter switching time indicates faster response speed. For precision servers, the switching time must be controlled at the millisecond or second level.

4) Continuous Power Supply Duration

It refers to the maximum duration that standby power supplies can operate continuously under full load, which is used to cope with prolonged mains power failures. Longer duration means stronger emergency response capability.

D. DIMENSIONLESS PROCESSING OF INDICATORS

Cost indicators and efficiency indicators differ greatly in units and value ranges, making direct comparison and calculation impossible. This paper adopts the extremum method to conduct dimensionless processing on all indicators, converting all values into the range of 0 to 1, so as to facilitate multi-objective calculation and comprehensive evaluation of the model.

IV. CONSTRUCTION OF INTEGRATED MODEL COMBINING DEEP LEARNING AND MULTI-OBJECTIVE OPTIMIZATION

A. OVERALL ARCHITECTURE OF THE MODEL

The integrated model is divided into two layers: the upper layer is the deep learning load prediction module, and the lower layer is the multi-objective optimization and evaluation module. The two modules are connected sequentially with one-way data transmission.

1.First layer: Input historical load data of cluster data centers, and predict the maximum load, average load and fluctuating load in different periods via the LSTM deep learning model. The prediction results are transmitted as fixed parameters to the second layer.

2.Second layer: On the basis of predicted load, establish multi-objective optimization functions combined with cost and efficiency indicators, set various constraint conditions, and adopt the particle swarm algorithm to solve the optimal configuration scheme. The model finally outputs the type, quantity, power and combination mode of equipment, and calculates the cost and efficiency score of each scheme.

B. CONSTRUCTION OF DEEP LEARNING LOAD PREDICTION MODEL

1) Data Preprocessing

The research team collected hourly power load data of a cluster data center for 12 consecutive months, together with auxiliary data such as date, time period, weather and holidays. Abnormal data caused by power outages and equipment maintenance were eliminated first, and then all raw data were normalized to the range of 0 to 1 to reduce the calculation difficulty of the model. The dataset was divided into training set, verification set and test set at a ratio of 7:2:1.

2) Parameter Setting of the LSTM Model

A three-layer LSTM network combined with a fully connected output layer is built in this paper.

1.Set basic parameters such as iteration epochs, batch size and learning rate before model training.

2.The verification set is used to monitor the model fitting status during training to prevent overfitting.

3.After training, the test set is used to verify the prediction accuracy and ensure that the error of load prediction results is within a reasonable range.

3) Output of Load Prediction Results

The well-trained model outputs three core types of load data: daytime peak load, nighttime off-peak load and full-period average load. These three groups of data represent the three most typical operating states of cluster data centers and serve as the core reference for standby power configuration.

C. CONSTRUCTION OF MULTI-OBJECTIVE OPTIMIZATION MATHEMATICAL MODEL

1) Setting of Objective Functions

Two independent objective functions are established in the model:

1.The first objective function: Minimize the total life-cycle cost. It integrates the four types of costs including procurement, installation, energy consumption and maintenance, with the optimization direction of minimizing the value.

2.The second objective function: Maximize the comprehensive efficiency score. It calculates the weighted sum of four efficiency indicators, with the optimization direction of maximizing the value.

3.The two objective functions jointly form the multi-objective optimization system, and the algorithm searches for equilibrium optimal solutions between the two objectives.

2) Setting of Constraint Conditions

Combined with engineering practices, five mandatory constraint conditions are set, and all solution results must meet the requirements:

1.Power constraint: The total output power of standby power supplies shall not be less than the maximum predicted load of the data center to ensure full load bearing capacity.

2.Quantity constraint: The quantity of each type of standby power equipment shall be a positive integer, complying with the rules of equipment procurement and deployment.

3.Site constraint: The total floor area occupied by equipment shall not exceed the reserved area of the power supply room in the data center.
4.Standard constraint: The fault switching time and continuous power supply duration shall comply with national industrial specifications for power supply of data centers.
5.Operation constraint: The equipment combination mode shall conform to safe operation logic, and the switching logic of different types of power supplies shall operate normally.

3) Model Solution Process

The multi-objective particle swarm optimization algorithm solves the model following the steps below:
1.Import load data, equipment parameters and constraint conditions output by deep learning.
2.Initialize the particle population and randomly generate multiple groups of initial standby power configuration schemes.
3.Calculate the total cost and comprehensive efficiency score of each scheme.
4.Compare the values of objective functions, update particle positions and iteratively optimize configuration schemes.
5.Stop calculation when the maximum iteration times are reached, and output the Pareto optimal solution set. The solution set contains multiple feasible and balanced configuration schemes for subsequent evaluation and comparison.

D. COST-EFFECTIVENESS COMPREHENSIVE EVALUATION MODEL

This paper adopts the weighted scoring method to establish the comprehensive evaluation model. Weights for cost indicators and each efficiency indicator are assigned based on industrial experience and expert opinions. After determining the weights, the dimensionless indicator values of each optimized scheme are substituted to calculate the comprehensive score of each scheme. A higher comprehensive score represents higher overall value of the configuration scheme.

V. EXAMPLE ANALYSIS AND RESULT VERIFICATION

A. BASIC CONDITIONS OF THE EXAMPLE

A large-scale cluster data center in China is selected as the research object, which consists of 4 sub-data centers with a total of 3,200 planned cabinets, belonging to a medium-sized computing node cluster. The data center operates non-stop all year round. Although the mains power access is stable, 3 to 5 short-term power outages occur in this area every year.

We collected basic data including existing equipment parameters, site area, equipment unit price and maintenance costs of the data center. The LSTM deep learning model was adopted for load prediction, and the results show that the maximum load of the cluster data center is 1860 kW, the average load is 1240 kW, and the off-peak load is 710 kW.

Four groups of different standby power configuration schemes are set for comparative analysis:
Scheme 1: Traditional empirical configuration (equipped with UPS only)
Scheme 2: Conventional combined scheme (UPS + low-power diesel generator set)
Scheme 3: Optimized combined scheme (UPS + high-power diesel generator set)
Scheme 4: Optimized scheme proposed in this paper (optimal combination obtained by deep learning and multi-objective optimization)

All four schemes meet the basic industrial power supply constraints. Relevant data are counted and compared through four tables below.

B. LOAD PREDICTION RESULTS AND ACCURACY ANALYSIS

Table 1 presents the load prediction results and error statistics of the deep learning model.

TABLE 1. Load prediction results and error statistics of the deep learning model

Load Type	Actual Load (kW)	Predicted Load (kW)	Absolute Error (kW)	Relative Error (%)
Peak Load	1860	1852	8	0.43
Average Load	1240	1235	5	0.40
Off-Peak Load	710	706	4	0.56

It can be seen from the table that the relative errors of the model for the three types of loads are all below 1%. This proves that the LSTM deep learning model has high prediction accuracy, and the output load data is reliable enough to serve as the input of the subsequent optimization model. The error of traditional linear prediction methods is generally above 3%, so the adopted prediction model has obvious advantages.

C. EQUIPMENT PARAMETER STATISTICS OF EACH CONFIGURATION SCHEME

Table 2 counts the equipment combination form, power, quantity and floor area of the four schemes.

TABLE 2. Equipment parameter table of four standby power configuration schemes

Scheme No.	Equipment Combination	UPS Power (kW)	Diesel Generator Power (kW)	Total Equipment Quantity	Total Floor Area (m^2)
Scheme 1	UPS only	2000	0	16	86
Scheme 2	UPS + Low-power Generator	1200	800	12	72
Scheme 3	UPS + High-power Generator	800	1200	9	65
Scheme 4	Algorithm-Optimized Combination	1000	900	10	68

According to Table 2: Scheme 1 adopts UPS alone with excessive total power, the largest equipment quantity and floor area. Schemes 2 and 3 are equipped with diesel generator sets, with gradually reduced equipment quantity and floor area. Scheme 4 is the combination optimized by the algorithm, featuring balanced power ratio, moderate equipment quantity and floor area without over-configuration or under-configuration.

The total power of all schemes exceeds the maximum load of 1860 kW, complying with the power constraint requirements.

D. COST INDICATOR COMPARISON OF EACH SCHEME

Table 3 counts the life-cycle costs of the four schemes over a 10-year period.

TABLE 3. Life-cycle cost statistics of four schemes (unit: ten thousand yuan)

Scheme No.	Initial Procurement Cost	Installation & Construction Cost	10-Year Converted Operational Energy Consumption Cost	10-Year Converted Maintenance & Replacement Cost	Total Cost
Scheme 1	286	42	158	96	582
Scheme 2	212	31	126	78	447
Scheme 3	175	26	142	65	408
Scheme 4	193	28	119	71	411

It can be concluded from Table 3 that Scheme 1 has the highest total cost of 5.82 million yuan, due to the high procurement price, large standby energy consumption and frequent later maintenance and replacement of pure UPS equipment.

Scheme 3 simply pursues equipment simplification with the lowest procurement cost, but suffers from high energy consumption during long-term operation of generator sets. The total cost of Scheme 4 is 4.11 million yuan, only 30,000 yuan higher than the minimum value of Scheme 3, achieving excellent cost control. Compared with the traditional empirical scheme, the optimized scheme in this paper can save 1.71 million yuan over a 10-year cycle with prominent economic advantages.

E. EFFICIENCY INDICATOR AND COMPREHENSIVE SCORE COMPARISON OF EACH SCHEME

Table 4 shows four efficiency indicators and comprehensive scores, with the score ranging from 0 to 1. A higher score represents better efficiency.

TABLE 4. Efficiency indicators and comprehensive scores of four schemes

Scheme No.	Power Supply Reliability	Load Matching Degree	Fault Response Speed	Continuous Power Supply Duration	Comprehensive Efficiency Score
Scheme 1	0.96	0.72	0.98	0.61	0.82
Scheme 2	0.92	0.85	0.94	0.78	0.87
Scheme 3	0.88	0.91	0.86	0.89	0.88
Scheme 4	0.94	0.93	0.95	0.86	0.92

As shown in Table 4: Scheme 1 performs well in response speed and reliability but has poor load matching degree and continuous power supply duration, resulting in a low comprehensive score. Scheme 3 has good load matching degree and power supply duration but declines in response speed and reliability. All efficiency indicators of Scheme 4 remain at a high level without obvious weaknesses, and its comprehensive efficiency score reaches 0.92, ranking first among the four schemes.

F. COMPARISON OF TEN-YEAR COMPREHENSIVE BENEFITS OF EACH SCHEME

Stable operation of standby power supplies can prevent direct economic losses caused by mains power outages, including business suspension, data loss and customer compensation. Based on the historical power outage frequency and single downtime loss in this region, this paper calculates the outage losses and comprehensive benefits of each scheme over a ten-year period, as presented in the table5.

$$\text{Comprehensive Benefit} = \text{Avoided Outage Losses} - \text{Full Lifecycle Cost}.$$

A higher value indicates a greater overall value of the scheme.

TABLE 5. Comparison of ten-year comprehensive benefits of each scheme (unit: ten thousand yuan)

Scheme No.	Annual Average Outage Loss	Total Ten-Year Outage Loss	Total Ten-Year Cost	Ten-Year Comprehensive Benefit
Scheme 1	24.6	246.0	582	-336.0
Scheme 2	18.2	182.0	447	-265.0
Scheme 3	22.5	225.0	408	-183.0
Scheme 4	12.1	121.0	411	-290.0

As shown in Table 5:

Scheme 1 has deficiencies in power supply efficiency and incurs the largest economic losses from power outages. Coupled with high investment, it delivers the poorest comprehensive benefit.

Scheme 3 features the lowest upfront cost yet insufficient power supply stability, resulting in relatively high outage losses.

Adopting a balanced configuration, Scheme 4 minimizes economic losses from power outages. Though its total cost is slightly higher than that of Scheme 3, it boasts stronger overall risk resistance and ranks first in comprehensive benefit among the four schemes. For continuously operating clustered data centers, this scheme can effectively mitigate implicit economic losses.

G. COMPREHENSIVE ANALYSIS OF EXAMPLE RESULTS

Combined with the data of the four tables, the comprehensive analysis is carried out as follows:

1.From the perspective of load prediction: The LSTM deep learning model can accurately predict the dynamic load of cluster data centers with a prediction error below 1%. Accurate load data is the premise for reasonable standby

power configuration and avoids blind type selection from the source.
2.From the perspective of cost and efficiency balance: The scheme based on the integrated algorithm proposed in this paper (Scheme 4) achieves the optimal balance. It maintains a low total cost while ranking the first in comprehensive efficiency score. Compared with other schemes, it neither causes excessive capital investment nor reduces power supply guarantee standards, fully adapting to the long-term operation demands of cluster data centers.
3.From the perspective of engineering application: The technical route of "deep learning + multi-objective optimization" is verified feasible through this example. All indicators can be quantified with intuitive and clear results, which can be directly referenced by on-site staff for equipment type selection and deployment.

VI. CONCLUSIONS

A. MAIN RESEARCH CONCLUSIONS

Targeting the standby power configuration problem of cluster data centers, this paper combines multi-objective optimization algorithms with deep learning technologies to build an integrated cost-effectiveness evaluation system. Theoretical derivation, model construction and simulation examples are adopted to complete the whole research, and four main conclusions are drawn:
1.The load of cluster data centers is characterized by strong volatility and interactivity. The LSTM deep learning model can accurately predict dynamic load with the prediction error controlled within 1%, providing reliable data support for power configuration.
2.The cost indicators and efficiency indicators of standby power supplies are mutually restricted objectives. Simply pursuing low cost will reduce power supply reliability, while excessive pursuit of high efficiency will sharply increase operational investment. The multi-objective optimization algorithm can effectively coordinate the two types of objectives and find the comprehensively optimal configuration scheme.
3.The evaluation method integrating deep learning and multi-objective optimization outperforms traditional empirical configuration methods and single-algorithm optimization methods. The practical example verifies that the optimized scheme in this paper has significant improvements in life-cycle cost, equipment utilization rate and power supply efficiency, delivering outstanding economic and safety benefits.
4.The four evaluation tables constructed in this paper can intuitively present the equipment parameters, cost data and efficiency data of different schemes. The tabular evaluation form is easy to understand and conforms to the usage habits of engineering sites, possessing high promotion and application value.

B. LIMITATIONS

There are two limitations in this study:
1.The example adopted in this paper is a medium-sized cluster data center, and the model has not been verified on super-large and extra-large cluster data centers, so the application scope of the model can be further expanded.
2.This research mainly considers conventional mains power failure scenarios, without taking extreme working conditions such as extreme weather and simultaneous failures of multiple sites into account. The configuration optimization under extreme scenarios needs further supplementation.

C. FUTURE RESEARCH PROSPECTS

Combined with the limitations of this research and the development trend of the industry, follow-up research can be expanded in three directions:
1.Expand research samples, conduct simulation analysis on cluster data centers of different scales and in different regions, optimize model parameters and improve the universality of the model.
2.Introduce extreme operating conditions, add scenarios such as simultaneous failures of multiple sites and prolonged power outages, perfect constraint conditions and enhance the emergency response capability of configuration schemes.
3.Integrate new energy storage technologies, distributed photovoltaic power supply and other new energy forms to enrich the types of standby power supplies, build a more diversified power configuration optimization model and further reduce the comprehensive operational costs of data centers.

FUNDING

Southern Power Grid Company Science and Technology Project (Project No.: GZKJXM20250030)

REFERENCES

[1] L. Xia, S. Baghaie, and S. Mohammad Sajadi, "The digital economy: Challenges and opportunities in the new era of technology and electronic communications," Ain Shams Eng. J., vol. 15, no. 2, p. 102411, 2024, doi: 10.1016/j.asej.2023.102411.
[2] K. Raina et al., "Artificial intelligence-driven management: Bridging innovation, knowledge creation, and sustainable business practices," J. Innov. Knowl., vol. 11, p. 100860, 2026, doi: 10.1016/j.jik.2025.100860.
[3] J. Kim et al., "Reviewing the socio-technical dynamics of AI, data centers and digitalization on energy and the environment," Renew. Sustain. Energy Rev., vol. 234, p. 116861, 2026, doi: 10.1016/j.rser.2026.116861.
[4] A. Kumar and N. M. Pindoriya, "Toward sustainable data center operation: A review on existing infrastructures, integrated smart energy management frameworks, and future perspectives," Renew. Sustain. Energy Rev., vol. 230, p. 116664, 2026, doi: 10.1016/j.rser.2025.116664.
[5] M. T. Takci et al., "Data centres as a source of flexibility for power systems," Energy Rep., vol. 13, pp. 3661–3671, 2025, doi: 10.1016/j.egyr.2025.03.020.
[6] A. Khosravi et al., "Review of energy efficiency and technological advancements in data center power systems," Energy Build., vol. 323, p. 114834, 2024, doi: 10.1016/j.enbuild.2024.114834.
[7] Z. Xu et al., "Reliability-based power supply capability evaluation of power distribution networks considering interruptible loads after contingencies," Sustain. Energy Technol. Assess., vol. 69, p. 103876, 2024, doi: 10.1016/j.seta.2024.103876.

[8] L. Bader, E. Wagner, and M. Serror, “PowerRange: An immersive cyber range for power grid operators,” Int. J. Crit. Infrastruct. Prot., vol. 53, p. 100845, 2026, doi: 10.1016/j.ijcip.2026.100845.
[9] S. Ghosh and J. Damodara, “Scalable data centers – Power generation and delivery challenges and solutions,” Energy Rep., vol. 15, p. 109178, 2026, doi: 10.1016/j.egyr.2026.109178.
[10] R. P. Narasipuram et al., “Supercapacitors in Modern Energy Systems: A Critical Review of Materials, Architectures, Digital Twins, AI Integration, and Applications,” Energy Eng., vol. 123, no. 5, 2026, doi: 10.32604/ee.2026.076542.
[11] L. Guo, W. Li, and M. Zhang, “Optimal capacity configuration and operation strategy of typical industry load with energy storage in fast frequency regulation,” Energy, vol. 308, p. 132874, 2024, doi: 10.1016/j.energy.2024.132874.
[12] X. Liu, “Optimization method for capacity configuration and power allocation of electrolyzer array in off-grid integrated energy system,” Heliyon, vol. 10, no. 11, p. e32312, 2024, doi: 10.1016/j.heliyon.2024.e32312.
[13] A. Torayev et al., “Optimal Manufacturing Configuration Selection: Sequential Decision Making and Optimization using Reinforcement Learning,” Procedia CIRP, vol. 120, pp. 986–991, 2023, doi: 10.1016/j.procir.2023.09.112.
[14] H. Sakata et al., “Production system configuration and layout planning for efficient manufacturing system design: An industrial case study,” CIRP J. Manuf. Sci. Technol., vol. 66, pp. 56–69, 2026, doi: 10.1016/j.cirpj.2026.02.001.
[15] M. Hasan et al., “A state-of-the-art comparative review of load forecasting methods: Characteristics, perspectives, and applications,” Energy Convers. Manag. X, vol. 26, p. 100922, 2025, doi: 10.1016/j.ecmx.2025.100922.
[16] T. Liu et al., “Intelligent load forecasting technologies for diverse scenarios in the new power systems,” Renew. Sustain. Energy Rev., vol. 226, p. 116366, 2026, doi: 10.1016/j.rser.2025.116366.