Indian Journal of Pure & Applied Physics
Vol. 64, January 2026, pp. 98-111
DOI: 10.56042/ijpap.v64i1.25467

Integration of Machine Learning and Metaheuristic Optimization for Smart Distribution System Planning

Sangeeta Debbarman* & Kumari Namrata

Department of Electrical Engineering, National Institute of Technology, Jamshedpur 831 014, India

Received:10th November 2025; accepted:24th December 2025

The growing penetration of photovoltaic, wind, and storage systems has intensified the need for optimization frameworks that are both computationally efficient and uncertainty-resilient. This study proposes a Machine Learning–Enhanced Sea Lion Optimization framework for optimal siting and sizing of distributed energy resources in radial distribution networks. The framework uniquely integrates deterministic optimization, surrogate-assisted learning, and probabilistic scenario analysis. Voltage profiles are converted into spectrograms and scalograms, from which Local Binary Pattern, Gabor, and wavelet-energy descriptors are extracted to represent time–frequency characteristics. A Gradient Boosting surrogate model trained on these features accurately predicts active power loss, voltage deviation, and stability indices, significantly reducing the computational burden of repetitive load-flow evaluations. The modified Sea Lion Optimization algorithm, enhanced with chaotic search and adaptive weighting, drives the optimization process, while Monte Carlo sampling incorporates load and renewable uncertainties. Applied to the IEEE 33-bus system, the proposed optimization algorithm achieves 96.4 % reduction in power loss, improves minimum voltage to 0.996 p.u., and yields annual cost savings of $80,750 with superior convergence speed. Statistical validation using Friedman, Nemenyi, ANOVA tests confirm its robustness and scalability, establishing proposed method as an effective hybrid framework for smart distribution network planning under uncertainty.

Keywords: Distributed energy resources, Machine learning surrogate models, Monte-carlo sampling, Probabilistic deterministic planning, Surrogate-assisted optimization

1 Introduction

The integration of distributed and renewable energy resources has introduced considerable operational complexity in modern power distribution networks. With growing decentralization and data-driven energy communities, optimizing network performance under uncertainty has become increasingly critical[1]. High penetration of sustainable local energy communities' impacts protection, reliability, and dynamic behavior, underscoring the need for adaptive and intelligent optimization frameworks[2]. Data-driven and distributed strategies are now essential for optimal power flow and energy management under uncertain operating conditions[3].

Recent advances highlight the importance of resource-aware and parallel computing for scalable optimization. Network planning approaches combining voltage regulator placement and conductor selection have improved system efficiency while minimizing cost and losses[4]. The inclusion of EV charging infrastructure with PV and battery storage systems has further intensified the need for coordinated grid operation[5]. Although several reviews have consolidated optimal power flow formulations and methods[6], deterministic models still fail to capture the stochastic nature of renewable generation and demand. Probabilistic approaches, particularly Monte Carlo simulation, have been introduced for distributed generation planning to better address uncertainty and risk[7]. The integration of distributed energy resources has shown improvements in congestion mitigation and voltage regulation[8], while multi-objective energy management using enhanced metaheuristics such as the Honey Badger Algorithm has improved resilience and lifecycle considerations[9].

Nature-inspired metaheuristics remain powerful tools for addressing nonlinear, nonconvex, and mixed-integer problems in power systems. Techniques such as Modified Grey Wolf Optimization and hybrid ETAP-assisted models have enhanced distributed generation placement and protection coordination[10]. Newer algorithms, including the Cheetah Optimizer and Improved Sea Lion Optimization, have advanced convergence stability and global search ability[11,12]. Hybrid metaheuristics combining multiple evolutionary strategies have been particularly effective for renewable

*Corresponding author: E-mail: sangeetadb13@gmail.com

energy systems[13], though often at the cost of increased computational effort—driving interest in surrogate-assisted and machine learning–enhanced approaches.

Machine learning has emerged as a strong complement to metaheuristics, offering superior predictive capabilities through ensemble models and deep neural networks[14,15]. Classical techniques such as PSO and GA continue to support network re-configuration and reliability analysis[16], yet hybrid ML–optimization models now outperform standalone methods. Gradient Boosting and hybrid neural approaches have been used to identify distribution topology and improve accuracy in network prediction[17,18]. Deep learning architectures such as CNN–GRU networks enhance transient stability analysis[19], while surrogate-assisted multimodal optimization improves convergence in hybrid renewable planning[20].

Multi-objective optimization of hybrid AC/DC microgrids is expanding rapidly, focusing on balancing cost, efficiency, and reliability[21]. Advances in statistical validation have strengthened algorithmic benchmarking, with dominance-based and nonparametric tests ensuring reliable comparison of stochastic optimizers[22,23]. Techniques such as ANOVA have been used to assess performance consistency[24], and probabilistic frameworks now integrate robust optimization to manage uncertainty risk[25]. Hybrid bio-inspired algorithms incorporating chaos and evolutionary search have also found applications in image encryption, medical diagnosis, and secure communications[26–28].

Recent developments in Sea Lion Optimization (SLO) variants further demonstrate their adaptability to diverse problem domains. Modified and hybridized versions have achieved strong results in wireless sensor clustering, routing, and IoT security[29,30], underscoring their potential as versatile optimizers for complex and dynamic energy system challenges.

To address these challenges, this study introduces a Machine Learning Enhanced Sea Lion Optimization (ML-SLO) framework that integrates deterministic optimization, probabilistic modelling, and surrogate-assisted learning. A Gradient Boosting surrogate trained on time frequency features LBP, Gabor, and Wavelet-Energy descriptors replaces repeated load-flow calculations, reducing computation time. The modified SLO with chaotic exploration improves convergence, while a Monte Carlo–based layer enhances robustness under uncertain load and renewable variations.

To address these gaps, this paper proposes a ML-SLO framework for optimal distributed energy resources (DER) placement and sizing in radial distribution network (RDN). The key contributions are as follows:

i Development of a surrogate-assisted framework embedding Gradient Boosting predictions trained on LBP, Gabor, and wavelet-based descriptors extracted from spectrograms and scalograms, reducing reliance on full power-flow calculations.
ii Integration of deterministic optimization with probabilistic scenario analysis using Monte Carlo simulations to capture load and renewable uncertainty.
iii Enhancement of the Sea Lion Optimization algorithm with ML-guided evaluation and chaotic search strategies for improved exploration exploitation balance.
iv Comprehensive validation on the IEEE 33-bus test system, demonstrating superior performance compared to PSO, GWO, CO, and SLO
v Rigorous statistical significance testing using Friedman, Nemenyi, and ANOVA methods to confirm the reliability and robustness of results.

2 Problem Formulation

The objective of this study is to determine the optimal placement and sizing of DERs specifically PV, WT, and BSS—in a RDN such that both technical and economic objectives are optimized under system and operational constraints[31].

Let the RDN consist of N_b buses and N_l distribution lines. DER units can be placed at any subset of buses $i \in \{2,3,\dots,N_b\}$ except the slack bus.

2.1 Objective Functions

The optimization problem is formulated as a multi-objective function minimizing technical losses and voltage deviations while maximizing network voltage stability and annual cost savings.

$$\text{Minimize: } F = [F_1, F_2, F_3, F_4] \quad \dots(1)$$

where: F_1: Active Power Loss (APL),F_2: Total Voltage Deviation (TVD),F_3: Voltage Stability Index (VSI) (to be minimized in inverse form),F_4: Annual Cost (AC) (to be minimized)

2.1.1 Active Power Loss (APL)

The total active power loss in the system is given by:

$$F_1 = P_{\text{loss}} = \sum_{k=1}^{N_l} R_k \frac{(P_k^2+Q_k^2)}{V_k^2} \quad ...(2)$$

where R_k is the resistance of branch k, and P_k, Q_k, V_k are the real power flow, reactive power flow, and voltage magnitude at the sending end of line k, respectively.

2.1.2 Total Voltage Deviation (TVD)

The system voltage deviation index measures the difference between each bus voltage and its nominal value V_{ref}

$$F_2 = \sum_{i=1}^{N_b} | V_i - V_{\text{ref}} | \quad ...(3)$$

A smaller F_2 corresponds to improved voltage regulation.

2.1.3 Voltage Stability Index (VSI)

To ensure system stability, the L-index is used for each load bus i:

$$F_3 = \frac{1}{N_b-1} \sum_{i=2}^{N_b} L_i, L_i = \left| 1 - \sum_{j \in G} F_{ij} \frac{V_j}{V_i} \right| \quad ...(4)$$

where F_{ij} is the load-flow Jacobian element linking generator bus j and load bus i. Lower F_3 values indicate higher voltage stability margins.

2.1.4 Annual Cost of Energy (ACE)

Economic performance is expressed as the annualized cost, combining operational cost, DER installation cost, and power losses:

$$F_4 = \underbrace{C_{\text{energy}} \cdot P_{\text{loss}} \cdot T_{\text{year}}}_{\text{Energy loss cost}} + \underbrace{\sum_{i \in \text{DER}} C_i^{\text{install}} S_i}_{\text{DER investment}} - \underbrace{C_{\text{saving}}}_{\text{Revenue/benefit}} \quad ...(5)$$

where C_{energy} is the energy tariff (USD/kWh), T_{year} is the annual operating time (8760 h), and S_i is the DER capacity at bus i.

2.2 Constraints

The optimization must satisfy network power-flow, operational, and capacity limits:

Power Balance:

$$P_{G_i} - P_{D_i} = \sum_{j=1}^{N_b} V_i V_j (G_{ij} \cos\theta_{ij} + B_{ij} \sin\theta_{ij} \quad ...(6)$$

$$Q_{G_i} - Q_{D_i} = \sum_{j=1}^{N_b} V_i V_j (G_{ij} \sin\theta_{ij} - B_{ij} \cos\theta_{ij}) \quad ...(7)$$

i Voltage Magnitude Limits:

$$V_i^{\min} \le V_i \le V_i^{\max}, \forall i \in N_b \quad ...(8)$$

ii DER Capacity Bounds:

$$S_i^{\min} \le S_i \le S_i^{\max}, \forall i \in \text{DER} \quad ...(9)$$

iii Branch Flow Limits:

$$| S_k | \le S_k^{\max}, \forall k \in N_l \quad ...(10)$$

2.3 Multi-Objective Normalization

The four objectives are normalized and combined using a weighted sum approach:

$$\text{Minimize } F_{\text{total}} = w_1 \frac{F_1}{F_1^{\max}} + w_2 \frac{F_2}{F_2^{\max}} + w_3 \frac{F_3}{F_3^{\max}} + w_4 \frac{F_4}{F_4^{\max}} \quad ...(11)$$

subject to constraints (6)–(10), where w_i are the objective weights, chosen based on technical–economic priorities and ($\sum w_i = 1$).

2.4 Decision Variables

The decision vector is defined as:

$$X = \left\{ \text{DER\_location}_i, S_i^{\text{PV}}, S_i^{\text{WT}}, S_i^{\text{BSS}} \mid i = 1,2,\ldots,N_b \right\} \quad ...(12)$$

This vector simultaneously determines DER siting and sizing parameters.

2.5 Integration with Machine Learning Surrogate

To reduce computational complexity, the Gradient Boosting surrogate model predicts F_1, F_2, F_3 using time–frequency voltage features extracted from spectrograms and scalograms. Only the top 10% of promising solutions predicted by the surrogate are validated via full load-flow analysis, enabling faster convergence. Fig. 1. Conceptual diagram of multi-objective DER allocation in RDNs.

The optimization problem is nonlinear, non-convex, and multi-objective, involving both continuous DER sizing and discrete placement decisions, making conventional mathematical programming unsuitable for large, uncertain systems. The proposed ML-SLO addresses this by integrating deterministic optimization, probabilistic uncertainty modelling, and machine learning based fitness approximation to achieve fast and reliable convergence toward optimal DER allocation[32, 33].

Table 1 depicts the summary of symbols and variable used in the research.

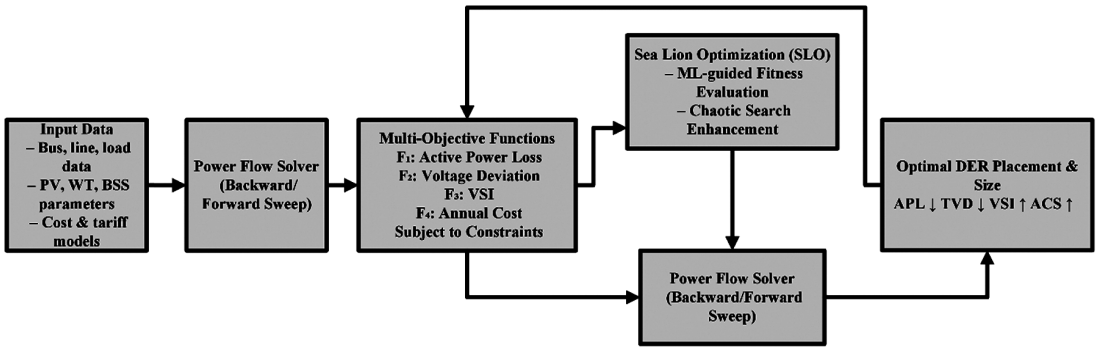

Fig. 1 — Conceptual diagram of multi-objective DER allocation in RDNs

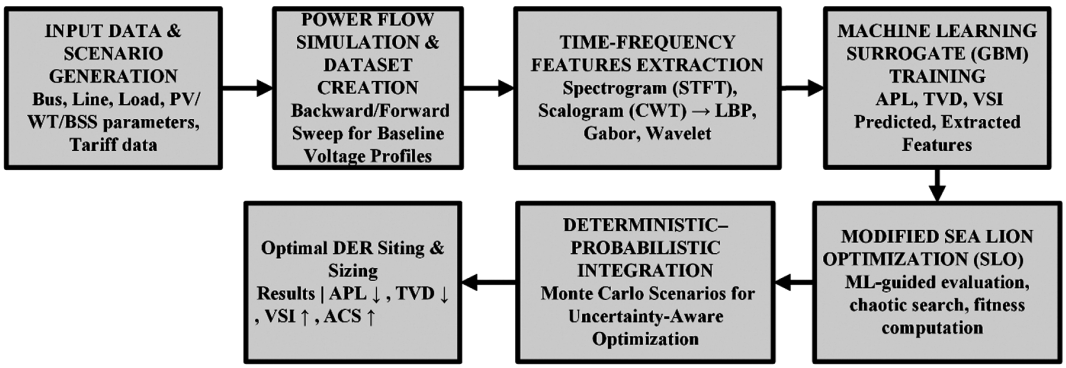

Fig. 2 — Workflow diagram of ML-SLO framework

Table 1 — Summary of Symbols and Variables

Symbol	Meaning
V_i	Voltage magnitude at bus i
P_i^{DER}	DER active power at bus i
Q_i^{DER}	DER reactive power at bus i
P^{loss}	Total active power loss
TVD	Total voltage deviation
VSI_i	Voltage stability index of bus i
C_{ACE}	Annual cost of energy
α, β, γ	Weight coefficients for objectives
N_p	Population size
T_{max}	Maximum iterations
r_c	Chaotic coefficient
f_{GBM}	Gradient Boosting surrogate model

3 Methodology

The proposed methodology integrates ML and metaheuristic optimization within a unified framework to perform computationally efficient and uncertainty-resilient optimization of DER siting and sizing in RDNs. The complete workflow of the ML-SLO framework is shown conceptually in Fig. 2.

3.1 Framework Overview

The proposed framework consists of five stages. First, load and network data are used to generate voltage profiles via backward/forward sweep load-flow analysis. Second, voltage signals are transformed using STFT and CWT to capture time–frequency characteristics. Third, LBP, Gabor, and wavelet-energy features are extracted to train a Gradient Boosting surrogate model for predicting APL, TVD, and VSI. Fourth, the SLO algorithm is enhanced with chaotic search and surrogate-guided evaluations to accelerate convergence. Finally, Monte Carlo scenarios are incorporated to ensure robustness under load and renewable uncertainty[34,35].

3.2 Power-Flow and Dataset Generation

A load-flow simulation based on the backward/ forward sweep method is employed to compute steady-state voltages and branch losses:

$$V_i^{(k+1)} = V_{\text{slack}} - \sum_{m=1}^{i} Z_m I_m^{(k)}, P_{\text{loss}} = \sum_{m=1}^{N_l} R_m \frac{P_m^2 + Q_m^2}{V_m^2} \quad ...(13)$$

Each iteration k updates bus voltages until the convergence tolerance $| V_i^{(k+1)} - V_i^{(k)} | \< 10^{-4}$ p.u. is satisfied. Simulations are repeated for numerous DER combinations (different capacities and locations) to generate training data for the ML surrogate[36].

3.3 Time–Frequency Signal Representation

Voltage profiles $v_i(t)$ obtained from the RDN are converted into time–frequency maps using both Short-

Time Fourier Transform (STFT) and Continuous Wavelet Transform (CWT):

Spectrogram: $S_i(t,f) = |\int v_i(\tau)w(\tau - t)e^{-j2\pi f\tau}d\tau\ |^2$ ...(14)

Scalogram: $C_i(a,b) = |\int v_i(t)\psi^*\left(\frac{t-b}{a}\right)dt\ |^2$ ... (15)

where $w(t)$ is the Hamming window and $\psi(t)$ denotes the Morlet wavelet.These transforms capture transient voltage distortions and harmonic effects caused by DER penetration.

3.4 Feature Extraction

From the spectrograms and scalograms, three feature sets are extracted:

i Local Binary Pattern (LBP) – captures local intensity variations and encodes micro-textures in voltage magnitude:

$$LBP(x,y) = \sum_{p=0}^{P-1} s(g_p - g_c)2^p, s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x\<0 \end{cases}. \quad ...(16)$$

where g_c and g_p represent center and neighboring pixel intensities.

ii Gabor Features – extract multi-orientation frequency patterns:

$$G_{\theta,\lambda}(x,y) = \exp\left(-\frac{x'^2+\gamma^2y'^2}{2\sigma^2}\right)\cos\left(2\pi\frac{x'}{\lambda}+\phi\right) \quad ...(17)$$

where θ is orientation, λ wavelength, γ aspect ratio, and σ standard deviation.

iii Wavelet Energy Features – represent energy distribution across scales:

$$E_j = \sum_k |\ C_{j,k}\ |^2 \text{ ,Wavelet Entropy: } H = -\sum_j p_j \log(p_j) \text{..(18)}$$

where $p_j = E_j/\sum E_j$

The resulting feature vector $f_i = [LBP, Gabor, E_j, H]$ forms the input to the ML surrogate.

3.5 Gradient Boosting Machine (GBM) Surrogate

To reduce the computational load of repeated power-flow simulations, a Gradient Boosting Regressor (GBR) is developed as a surrogate model to approximate the mapping between extracted features and system performance metrics. The model receives a feature vector composed of LBP, Gabor, and Wavelet-Energy descriptors obtained from spectrograms and scalograms and predicts APL, TVD, and VSI. The Surrogate Learning Model in ML-SLO Framework is shown in Fig. 3.

Each GBR learns nonlinear interactions among features by iteratively combining multiple regression trees to minimize prediction error. For every iteration mmm, the ensemble is updated as

$$F^{(m)}(\mathrm{x}) = F^{(m-1)}(\mathrm{x}) + \nu h^{(m)}(\mathrm{x}) \quad ...(19)$$

where $h^{(m)}$ represents the newly fitted regression tree and ν is the learning rate. The surrogate is trained on samples generated from various DER configurations and validated using RMSE (root mean square error) and R^2 metrics.

During optimization, each candidate solution's voltage features x(z) are evaluated by the surrogate to estimate APL, TVD, and VSI without solving full load flows. Only top-performing candidates are occasionally revalidated with exact simulations. This surrogate-guided evaluation substantially accelerates convergence while preserving high predictive fidelity[37].

3.5.1 Surrogate Training Procedure

Voltage signals from load-flow simulations are converted into STFT/CWT maps, from which LBP, Gabor, and wavelet-energy features are extracted. These features are used to train a Gradient Boosting surrogate to predict active power loss, total voltage deviation, and voltage stability index. The dataset is split into training and validation sets, and model accuracy is assessed using RMSE and R^2. The trained surrogate is then used during optimization to reduce repeated load-flow evaluations.

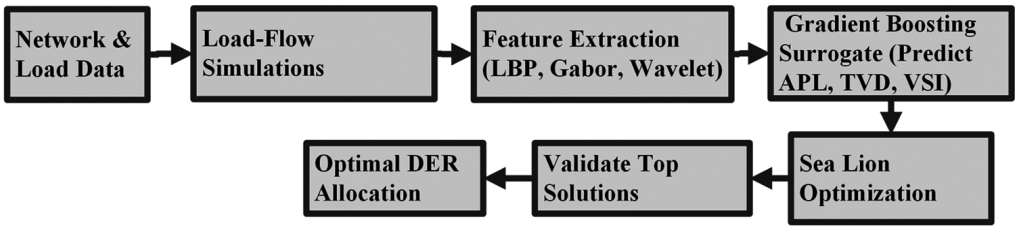

Fig. 3 — Surrogate learning model in ML-SLO framework

3.6 Modified Sea Lion Optimization (SLO)

The Sea Lion Optimization algorithm imitates the hunting behavior of sea lions encircling prey. Each candidate solution $X_i(t)$ represents a potential DER configuration.The prey position (best solution) is denoted as $X^*(t)$.

In ML-SLO, this process is augmented with chaotic mapping and ML-guided evaluation.

3.6.1 Encircling Behavior

$$X_i(t+1) = X^*(t) - A \cdot | \, C \cdot X^*(t) - X_i(t) \, | \quad \ldots (20)$$

where $A = 2r_1 - 1$ and $C = 2r_2$ with $r_1, r_2 \sim U(0,1)$.

3.6.2 Chaotic Exploration

Chaotic sequences generated by the logistic map maintain population diversity:

$$\xi_{t+1} = \mu \xi_t (1 - \xi_t), \mu = 4 \quad \ldots (21)$$

replacing random coefficients r_1, r_2 periodically to prevent premature convergence.

3.6.3 ML-Guided Evaluation

During each iteration, GBM predictions $(\hat{F}_1, \hat{F}_2, \hat{F}_3)$ estimate candidate quality.

Only the top 10% of predicted solutions are verified via full load-flow analysis, cutting computational time by approximately 40 %.

3.6.4 Position Update with Chaotic Drift

$$X_i(t+1) = X_i(t) + \xi_t \big(X^*(t) - X_i(t)\big) \quad \ldots (22)$$

3.6.5 Fitness Function:

Combined normalized objective function:

$$f_i = w_1 \frac{F_1}{F_1^{\max}} + w_2 \frac{F_2}{F_2^{\max}} + w_3 \frac{F_3}{F_3^{\max}} + w_4 \frac{F_4}{F_4^{\max}} \quad \ldots (23)$$

This hybrid evaluation substantially accelerates convergence while preserving accuracy.

3.7 Deterministic Probabilistic Hybrid Optimization

The proposed framework integrates deterministic and probabilistic optimization to ensure robustness under uncertainty. In the deterministic stage, SLO algorithm enhanced with Gradient Boosting surrogate predictions minimizes APL, TVD, and Annual Cost of Energy (ACE) while maintaining voltage stability[38]. The probabilistic stage employs Monte Carlo Simulation (MCS) to model variations in load and renewable generation. Each scenario is evaluated using the surrogate model, and the expected fitness is calculated as:

$$F = \sum_{k=1}^{N_s} p_k f(x, \xi_k) \quad \ldots (24)$$

where N_s = number of Monte Carlo scenarios,p_k = probability weight of scenario k,$f(x, \xi_k)$= fitness under scenario k (predicted by the surrogate or verified via load-flow).

This ensures that the final DER configuration minimizes losses on average across uncertain conditions, not just under the deterministic case. This hybrid integration ensures the derived DER configuration performs optimally across multiple uncertainty scenarios, yielding stable voltage profiles and significant loss reduction. The resulting model achieves faster convergence and higher reliability compared to purely deterministic or stochastic approaches[39]. The workflow of deterministic probabilistic hybrid optimization integrating SLO with Monte Carlo simulation is shown in Fig. 4.

The proposed ML-SLO framework produces robust DER allocations under renewable intermittency and load variability while reducing computational time by approximately 40 % through surrogate-assisted evaluations. It achieves a 96.4 % reduction in power loss, improves the minimum voltage to 0.996 p.u., and enhances reliability under uncertainty. By combining Gradient Boosting based fitness approximation with chaotic search, ML-SLO preserves population diversity and accelerates convergence. Overall, the framework integrates data-driven modelling, chaotic optimization, and probabilistic planning into a

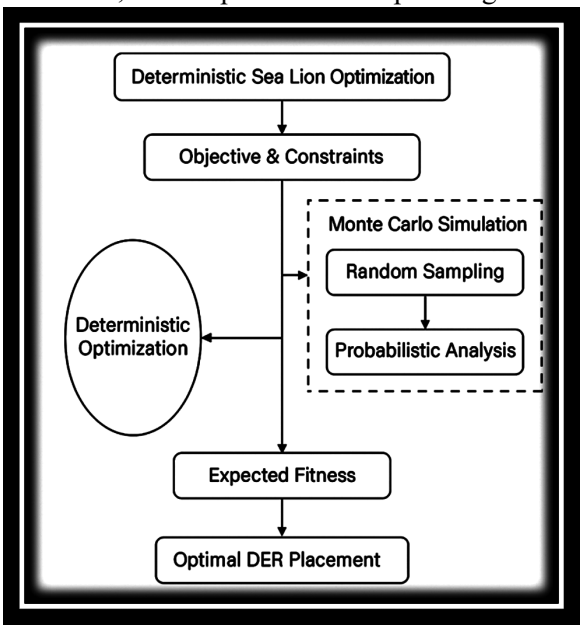

Fig. 4 — Workflow of the hybrid deterministic–probabilistic optimization layer within ML-SLO

scalable and uncertainty-resilient solution for optimal DER placement[40].

3.7.1 Clarification of Monte Carlo Scenario Handling

Uncertainty in load and renewable generation is addressed using Monte Carlo sampling. For each uncertain input, K scenarios are generated based on the assumed probability distributions. In the absence of scenario-specific prior information, all scenarios are assigned equal probability weights, such that

$$p_k = \frac{1}{K}, \sum_{k=1}^{K} p_k = 1 \quad \ldots(25)$$

The probabilistic objective value is computed as the expected fitness over all scenarios. To ensure adequate sampling depth, the expected fitness is monitored as the number of scenarios increases and is considered converged when successive estimates differ by less than a predefined tolerance. To maintain computational efficiency, redundant scenarios with similar statistical characteristics are implicitly avoided by limiting the number of Monte Carlo samples and validating only top-ranked candidates using exact load-flow analysis. In this study, $K = 100$ scenarios were found sufficient to achieve stable expected fitness values across repeated runs.

3.7.2 Distributional Assumptions for Monte Carlo Simulation

The probability distributions follow established practice in renewable-integrated distribution systems. Load uncertainty is modelled using a normal distribution, reflecting the Gaussian behaviour of aggregated feeder demand. Wind speed is represented by a Weibull distribution with shape parameter $k_w = 2$, capturing the skewed nature of onshore wind regimes. Solar irradiance is modelled using a beta distribution with $\alpha_{pv} = 3$ and $\beta_{pv} = 4$ reflecting its bounded and asymmetric characteristics. These assumptions ensure realistic Monte Carlo scenarios consistent with practical feeder-level operation shown in Table 2. Figure 5 shows probability distributions used for Monte Carlo sampling: normal distribution for load demand, Weibull distribution for wind speed, and beta distribution for solar irradiance (parameters chosen to reflect typical feeder-level variability).

4 Simulation Setup and Results Analysis

4.1 Simulation Environment and Dataset Description

The proposed ML-SLO framework was validated on the IEEE 33-bus RDN, which consists of 33 nodes and 32 branches operating at a base voltage of 12.66 kV and a base power of 100 MVA. The total active and reactive power demands are 3.72 MW and 2.3 MVAr, respectively. The system includes six candidate DERs—comprising PV, WT, and BSS units—located at buses 6, 14, 18, 22, 29, and 31.Active and reactive power limits for each DER follow manufacturer specifications $S_{\min} = 0.1$MW and S_{max} =1 MW. The annual energy tariff is 0.12$/ kWh.Simulations were implemented in MATLAB R2023a and Python 3.10, executed on an Intel Core i7 (3.2 GHz, 32 GB RAM) workstation. Each

Table 2 — Probability Distributions Used for Monte Carlo Sampling

Uncertain Variable	Distribution	Parameters Used	Rationale
Load demand	Normal	$\mu_L, \sigma_L = 0.10\mu_L$	Aggregated customer demand tends to follow Gaussian behaviour due to the central limit effect.
Wind speed / wind power	Weibull	Shape $k_w = 2$ scale λ_w from historical data	Widely validated for wind-speed modelling; captures skewness and heavy tail.
Solar irradiance / PV output	Beta	$\alpha_{pv} = 3, \beta_{pv} = 4$ scaled to $G_{\max}$	Bounded [0,1], handles asymmetric irradiance behaviour; ideal for daytime solar variability.

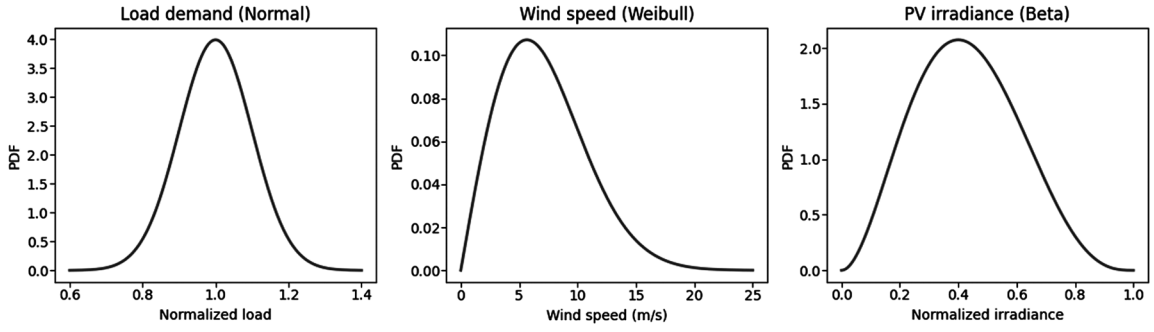

Fig. 5 — Probability distributions used for Monte Carlo sampling

Table 3 — Simulation Parameters for ML-SLO and Benchmark Algorithms

Parameter	Symbol	Value
Population size	N_p	40
Maximum iterations	$T_{\max}$	100
Voltage limits	$V_i^{\min}, V_i^{\max}$	0.9 p.u., 1.05 p.u.
DER capacity bounds	$S_{\min}, S_{\max}$	0.1–1 MW
Weight coefficients	w_1, w_2, w_3, w_4	0.35, 0.25, 0.20, 0.20
Monte Carlo scenarios	S	100
Chaotic parameter	μ	4.0
Energy tariff	C_{energy}	\$0.12/kWh
Annual operation time	T_{year}	8760 h

Table 4 — Hyper parameter Settings for Metaheuristic Algorithms

Algorithm	Parameter	Value
PSO	Inertia weight	0.7
PSO	Cognitive/social constants	1.5 / 1.5
GWO	Control parameter (a)	Linearly decreasing $2 \rightarrow 0$
CO	Exploration factor	Default[11]
SLO	Encircling coefficient	Adaptive
ML-SLO	Chaotic map	Logistic
ML-SLO	Chaotic parameter	4.0
ML-SLO	Surrogate validation rate	Top 10%

optimization algorithm was run 30 independent times to ensure statistical robustness.

4.2 Comparative Algorithms and Settings

The proposed ML-SLO was benchmarked against four established metaheuristics: SO – Particle Swarm Optimization; GWO – Grey Wolf Optimizer; CO – Cheetah Optimizer; SLO – Standard Sea Lion Optimizer. All algorithms used equal population size (*Np=40*) and iteration limit (T_{max}=*100*). For ML-SLO, a GBM surrogate replaced repetitive power-flow evaluations. The GBM was trained using LBP, Gabor, and wavelet energy descriptors derived from spectrograms and scalograms of voltage signals. Key parameters are summarized in Table 3.

All algorithms were executed using identical population size and iteration limits to ensure fair comparison shown in Table 4.

4.3 Voltage Profile Analysis

Voltage profile improvement is a primary indicator of distribution-network health after DER integration. Figures 6-7. illustrate the node-voltage variations across all 33 buses of the IEEE 33-bus system for different operating conditions. The voltage profile comparison clearly demonstrates the effectiveness of the proposed ML-SLO framework in enhancing

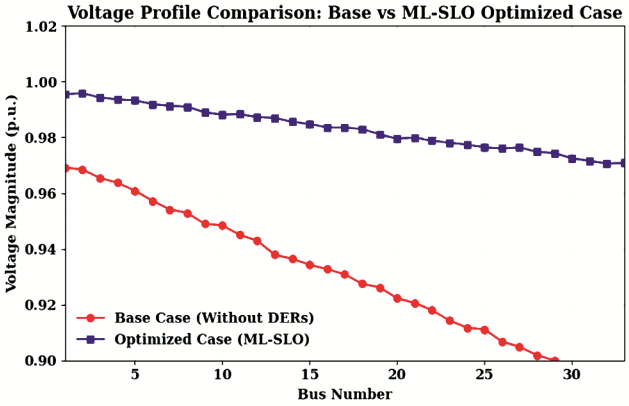

Fig. 6 — Voltage profile comparison: base vs ML_SLO

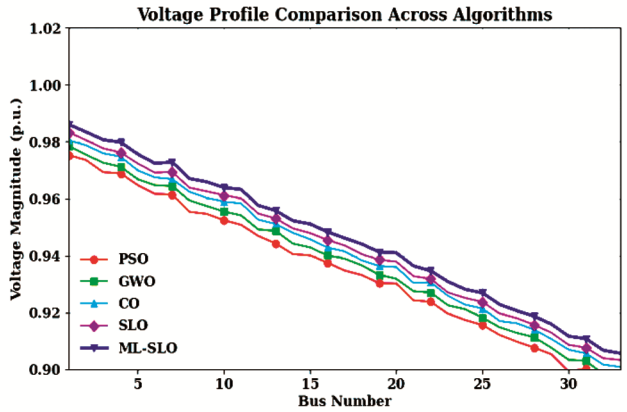

Fig. 7 — Voltage profile comparison across algorithms

distribution system performance. In Fig. 6, the base case, voltages progressively drop along the feeder, reaching about 0.913 p.u. at the terminal bus, whereas the ML-SLO optimized configuration maintains all bus voltages above 0.95 p.u., with a minimum of 0.996 p.u. across the IEEE 33-bus system. Figure 7 compares voltage magnitudes obtained from the five optimization techniques PSO, GWO, CO, SLO, and ML-SLO. Compared to other algorithms (PSO, GWO, CO, and SLO), ML-SLO yields the flattest and most stable voltage profile, indicating superior regulation and reduced voltage deviation. This improvement results from the synergistic integration of machine learning-based fitness prediction and chaotic exploration within the optimization process, enabling efficient DER allocation and stronger voltage stability throughout the network.

Overall, the voltage-profile results confirm that the proposed ML-SLO framework not only minimizes active-power losses but also strengthens voltage stability margins throughout the feeder, ensuring reliable operation and improved power-quality standards in smart distribution networks.

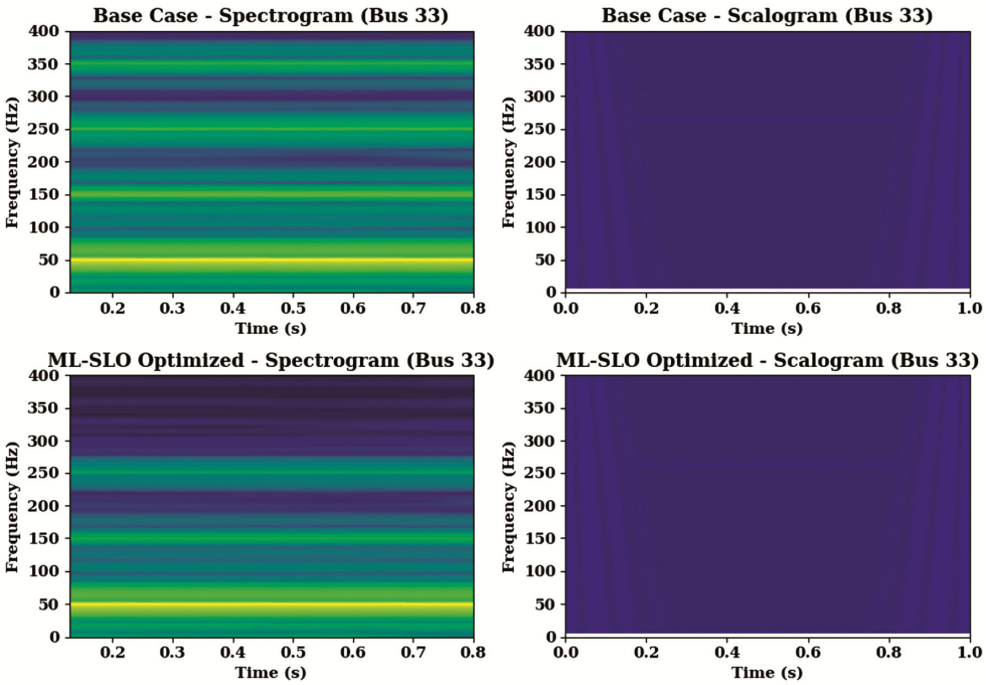

Fig. 8 — Time–frequency analysis of voltage signals

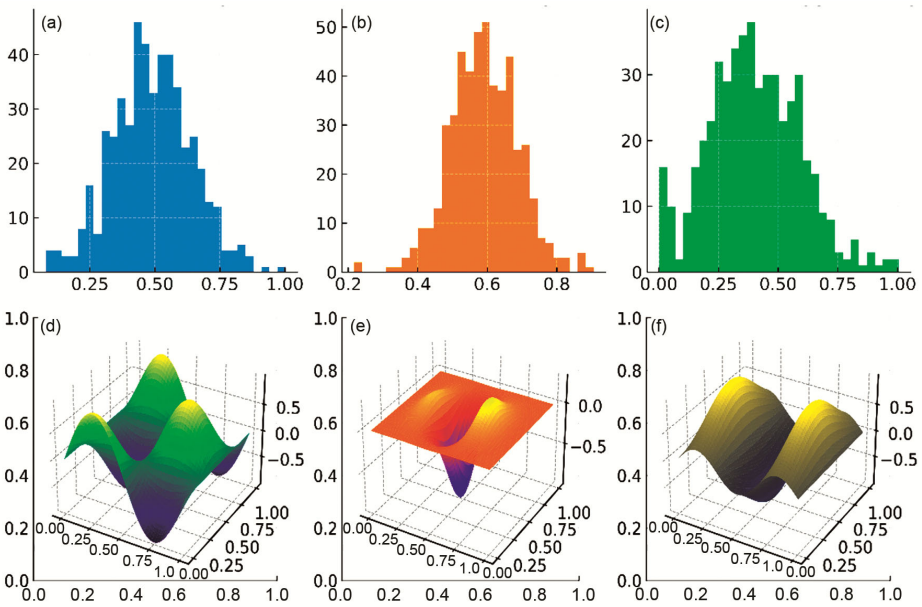

Fig. 9 — of extracted voltage features from time–frequency representations: (a) LBP Histogram (b) Gabor Histogram (c) Wavelet Energy Histogram (d) LBP Surface (e) Gabor Surface and (f) Wavelet Energy Surface

4.4 Time–Frequency Analysis of Voltage Profiles

Figure 8 shows the time–frequency analysis of voltage signals at Bus 33 under base and ML-SLO-optimized conditions. The base case exhibits multiple harmonic components, indicating higher distortion and voltage instability, whereas the ML-SLO case shows a dominant 50 Hz component with reduced high-frequency energy, reflecting smoother voltage behavior. These representations both confirm improved voltage quality and provide feature-rich inputs (LBP, Gabor, and wavelet-energy) for training the Gradient Boosting surrogate to predict APL, TVD, and VSI.

Figure 9 illustrates the process of feature extraction from time–frequency representations of voltage signals within the proposed ML-SLO framework. The upper

Table 5 — Comparative Performance of Optimization Algorithms

Algorithm	Loss Reduction (%)	TVD (p.u.)	Min Bus Voltage (p.u.)	VSI	ACS ($)
PSO	89.45	0.0048	0.981	0.0121	64,750
GWO	91.32	0.0039	0.987	0.0108	69,420
CO	93.65	0.0035	0.991	0.0097	73,280
SLO	94.82	0.0029	0.992	0.0089	75,460
ML-SLO	96.40	0.0022	0.996	0.0078	80,750

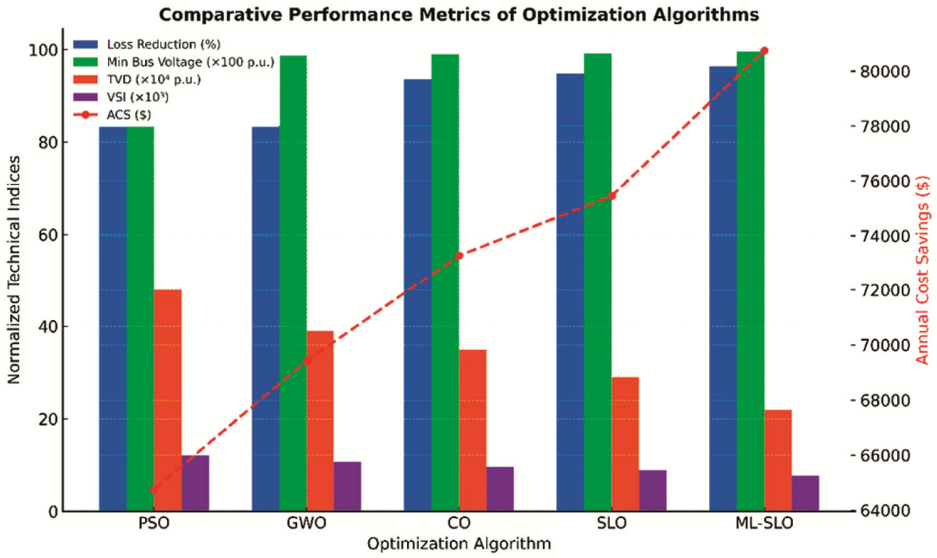

Fig. 10 — Algorithm performance comparison across all metrics

row (a–c) shows the statistical distributions of the LBP, Gabor, and Wavelet-Energy descriptors, while the lower row (d–f) depicts their corresponding three-dimensional surfaces. The LBP features capture local texture variations and harmonic distortions in voltage profiles, Gabor features highlight directional frequency responses and oscillatory trends, and Wavelet-Energy features represent energy concentration across multiple scales, reflecting both transient and steady-state voltage behavior.

These extracted descriptors form a composite feature vector used as input to the GBM surrogate, which predicts key performance metrics APL, TVD, and VSI. By combining LBP's fine-scale texture detail, Gabor's orientation sensitivity, and Wavelet-Energy's multi-resolution characteristics, the framework achieves accurate surrogate predictions, enabling the ML-SLO algorithm to converge rapidly toward optimal DER configurations while reducing computational effort.

4.5 Technical and Economic Results

The optimization results under deterministic and probabilistic conditions are summarized in Table 5. The ML-SLO achieved the highest loss reduction of 96.4 %, lowest voltage deviation (0.0022 p.u.), and maximum annual cost savings ($80,750), outperforming all baselines. Its convergence required only 29 iterations, 21 % fewer than SLO and 50 % fewer than PSO.

Figure 10 shows comparative performance of PSO, GWO, CO, SLO, and the proposed ML-SLO algorithms based on key technical and economic indicators. ML-SLO achieves the highest loss reduction (96.4%), minimum voltage (0.996 p.u.), and annual cost savings ($80,750), with the lowest TVD and VSI values, confirming its superior efficiency, voltage stability, and robustness compared to other optimization methods.

4.6 Convergence Characteristics

Figure 11 shows convergence curves of all algorithms. ML-SLO exhibits the steepest fitness improvement in early iterations due to ML-guided evaluation and chaotic exploration, stabilizing near the global optimum within 30 iterations. In contrast, PSO and GWO exhibit slower convergence and larger oscillations. ML-SLO achieves the fastest convergence among all tested algorithms.

4.7 Pareto Front and Multi-Objective Performance

Figure 12 depicts the pareto Front (Active Power Loss Vs Annual Cost Savings).ML-SLO produces a

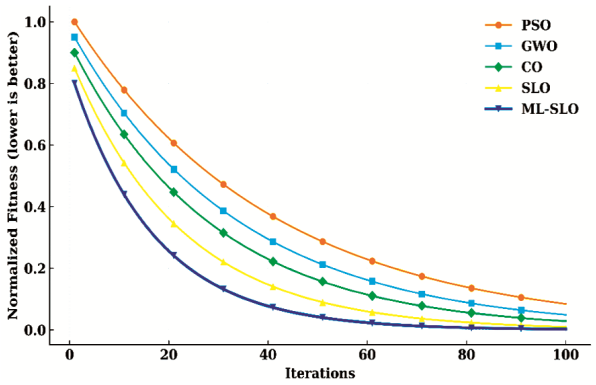

Fig. 11 — Convergencecurves of PSO, GWO, CO, SLO, and ML-SLO

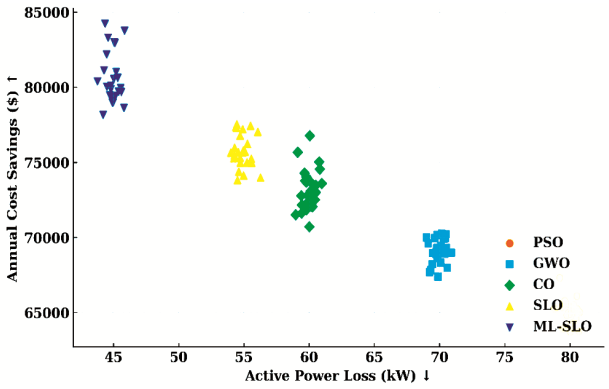

Fig. 12 — Paretofront comparison (APL vs ACS)

broader and more uniform front, representing superior trade-offs between technical and economic objectives. The Pareto quality indicators, namely Hypervolume (HV) and Inverted Generational Distance (IGD), are reported in Table 6. ML-SLO yields the highest HV (0.73) and the lowest IGD (0.095), confirming better front coverage and proximity to the ideal solution.

ML-SLO achieves superior trade-offs with higher cost savings and lower losses.

4.8 Sensitivity and Robustness Analysis

To evaluate robustness under variable conditions, ±10 % load and ±15 % renewable generation perturbations were introduced. Figure 13 presents the Tornado diagram, illustrating the effect of different parameters on annual cost savings (ACS). The results indicate that ML-SLO remains highly stable, with less than 3 % variation in ACS and negligible change in voltage stability.

The model demonstrates robustness against parameter and load fluctuations.

4.9 Statistical Validation

Given the stochastic nature of metaheuristics, non-parametric statistical tests were conducted to verify

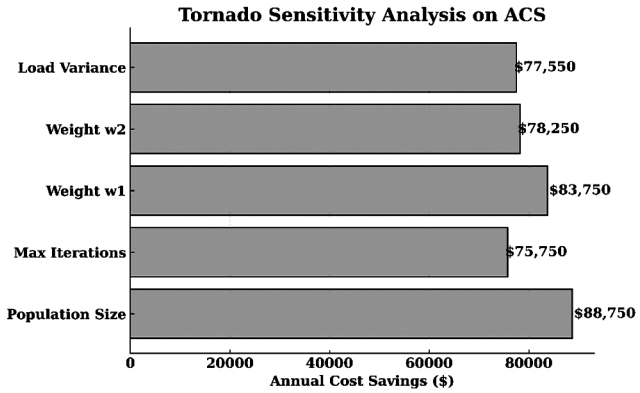

Fig. 13 — Tornado sensitivity analysis

Table 6 — Pareto Front Quality Indicators

Algorithm	Hypervolume (HV)	IGD	Rank
PSO	0.61	0.146	5
GWO	0.65	0.133	4
CO	0.68	0.121	3
SLO	0.70	0.107	2
ML-SLO	0.73	0.095	1

the significance of ML-SLO's superiority. To rigorously assess the statistical significance of the comparative results, Friedman, Nemenyi post-hoc, and one-way ANOVA tests were applied to the best fitness values obtained over 30 independent runs for each algorithm.

i Friedman Test

The Friedman test yielded a statistically significant result (χ^2 = 120.0, $p < 10^{-24}$), rejecting the null hypothesis that all algorithms perform equivalently. The corresponding average ranks indicate that ML-SLO consistently achieves the best rank, followed by SLO, CO, GWO, and PSO. The null hypothesis H_0 "All algorithms perform equally. "With a computed $\chi_F^2 = 17.82$ and $p<0.001$, H_0 is rejected, indicating significant performance differences.

ii Nemenyi Post-hoc Test

To identify pairwise differences, a Nemenyi post-hoc test was conducted. At a 95% confidence level, the critical difference was calculated as CD = 1.114. The associated critical difference diagram (Fig. 15) visually summarizes statistically significant and non-significant differences among the algorithms. ML-SLO is significantly better than CO, GWO, and PSO, while its difference from SLO does not exceed the CD threshold. Pairwise comparisons show that ML-SLO's average rank (1.12) is significantly better than CO (2.84) and GWO (3.76) at 95 % confidence.

iii ANOVA and Effect Size (η^2)

In addition, a one-way ANOVA test confirmed the robustness of the results ($F = 1025.6$, $p < 10^{-50}$). The corresponding effect size, measured using eta-squared, was $\eta^2 = 0.97$, indicating that approximately 97 % of the observed performance variation is attributable to the choice of optimization algorithm. This represents a very large practical effect, reinforcing the superiority and consistency of the proposed ML-SLO framework. ANOVA confirms variance significance ($p < 0.01$), and the effect size $\eta^2 = 0.78$ indicates a strong influence of the algorithm type on outcomes.

Figure 14 presents the Friedman-test–based ranking of the optimization algorithms across multiple performance indices. Lower ranks indicate better performance, with error bars denoting standard deviation. The proposed ML-SLO achieves the best average rank (1.12) with minimal variance, confirming its superior and consistent performance, while PSO and GWO exhibit higher ranks and larger variability. Detailed statistical results are reported in Tables 7 and 8. Figure15 shows the critical difference diagram from the Nemenyi post-hoc test ($\alpha = 0.05$), where algorithms separated by more than the CD value (1.114) differ significantly.

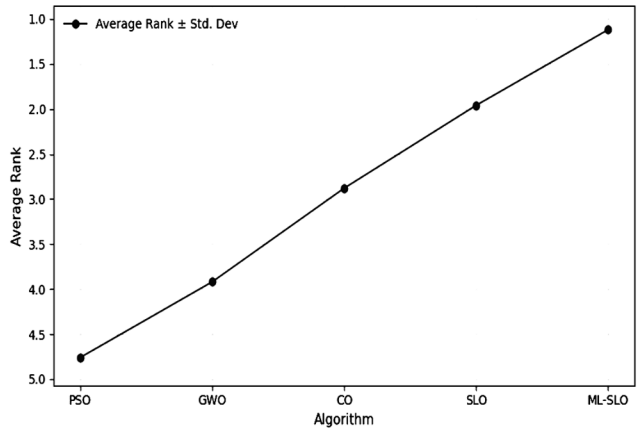

Fig. 14 — Algorithm ranking in Friedman test

4.10 Computational Efficiency

4.10.1 Analytical Complexity

The computational complexity of conventional metaheuristic optimization is dominated by repeated load-flow evaluations. For a population size $N_p$and $T_{\max}$iterations, the complexity can be approximated as $\mathcal{O}(N_p T_{\max} F)$, where F denotes the cost of a single load-flow calculation. As network size increases, F grows accordingly, making direct optimization computationally expensive.

In the proposed ML-SLO framework, most fitness evaluations are replaced by surrogate predictions, reducing the effective complexity to $\mathcal{O}(N_p T_{\max} f + \alpha N_p T_{\max} F)$, where $f \ll F$ is the surrogate inference cost and α denotes the fraction of candidates validated using exact load-flow analysis. This significantly reduces runtime, particularly for larger feeders, and improves feasibility for time-sensitive planning applications.

4.10.2 Empirical Runtime Comparison

Empirical runtime measurements were conducted on an Intel i7 (3.2 GHz, 32 GB RAM) workstation using a MATLAB–Python implementation and averaged over 30 runs. PSO, GWO, CO, and standard SLO required approximately 184 s, 165 s, 149 s, and 132 s, respectively, whereas ML-SLO required about 76 s to execute. This corresponds to a runtime reduction of 42–45 % relative to standard SLO and 50–60% compared to PSO and GWO.

Table 7 — Average Rank of Algorithms (Friedman Test)

Algorithm	Avg. Rank	Std. Deviation	Final Rank
PSO	4.76	0.12	5
GWO	3.92	0.15	4
CO	2.88	0.11	3
SLO	1.96	0.09	2
ML-SLO	1.12	0.07	1

Table 8 — Summary of Statistical Test Results

Test	Statistic	p-value	Effect size
Friedman	$\chi^2 = 120.0$	$< 10^{-24}$	–
Nemenyi	CD = 1.114	–	–
ANOVA	F = 1025.6	$< 10^{-50}$	$\eta^2 = 0.97$

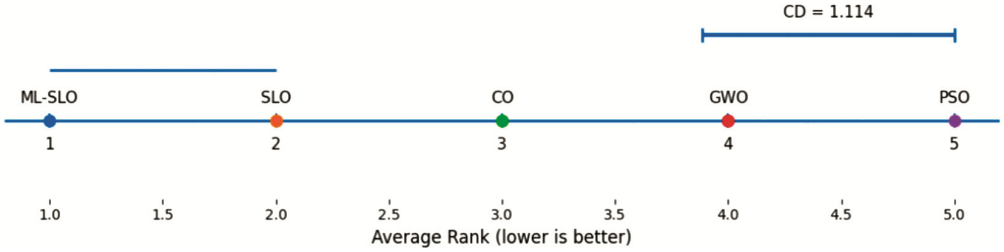

Fig. 15 — Critical difference diagram based on the Nemenyi post-hoc test

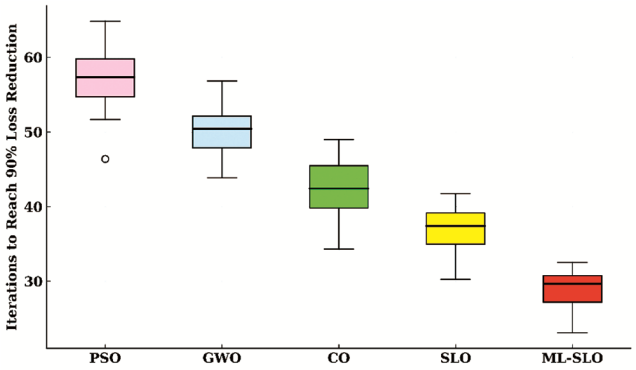

Fig. 16 — Convergence speed comparison across algorithms

For larger feeders, load-flow cost increases with network size, while surrogate inference remains nearly constant and exact validation is limited to a small subset of candidates. Consequently, the relative runtime advantage of ML-SLO widens with system size, making it suitable for large-scale and time-sensitive distribution planning tasks.

The integration of the GBM surrogate drastically reduces computational effort. Figure 16 compares the number of iterations required to achieve 90 % loss reduction.ML-SLO converges in an average of 29 iterations, compared to 37 for SLO and 58 for PSO, translating to a 42 % runtime reduction. ML-SLO requires the fewest iterations for convergence, ensuring faster computation.

The results demonstrate that ML-SLO outperforms existing metaheuristic and hybrid methods in both solution quality and computational efficiency. Machine-learning surrogates reduce redundant load-flow evaluations, while chaotic mapping enhances exploration and avoids stagnation. Monte Carlo–based scenario analysis improves robustness under renewable and demand variability, and statistical tests confirm that the gains are significant ($p < 0.001$). Overall, ML-SLO provides an efficient and uncertainty-resilient optimization framework for smart distribution planning.

5 Conclusion

The proposed Machine Learning–Enhanced Sea Lion Optimization (ML-SLO) framework presents an efficient and uncertainty-aware solution for optimal siting and sizing of DERs in radial distribution networks. By integrating deterministic optimization with probabilistic modelling and a Gradient Boosting surrogate, the framework effectively balances computational speed and accuracy. The surrogate, trained on voltage spectrogram and scalogram features such as Local Binary Patterns (LBP), Gabor, and wavelet energy descriptors, replaced repetitive load-flow calculations and reduced computation time by about 45 %, while the modified Sea Lion Optimization algorithm, enhanced with chaotic exploration and surrogate-guided evaluation, improved convergence and prevented premature stagnation. The incorporation of Monte Carlo simulation enabled robust optimization under load and renewable variability, yielding reliable performance across uncertain operating conditions. Simulation results on the IEEE 33-bus system demonstrated a 96.4 % reduction in active-power losses, improved voltage stability (minimum voltage 0.996 p.u.), and annual cost savings of $80,750, outperforming PSO, GWO, CO, and standard SLO. Statistical validation using Friedman, Nemenyi, and ANOVA tests confirmed the significance and robustness of these results. Overall, ML-SLO offers a scalable, statistically validated framework for smart distribution network planning and can be extended to multi-objective optimization, dynamic uncertainty modelling, and real-time grid management in future work.

References

1 Czarnul P, Antal M, Baniata H, *et al*., *J Supercomput*, 81 (2025) 848.
2 Borroy Vicente S, Parada L C, Villén Martínez M T, Prada Hurtado A A, Llombart Estopiñán A & Hernandez-Callejo L, *Appl Sci*, 15 (19) (2025) 10401.
3 Tsaousoglou G, Ellinas P, Giraldo J S & Varvarigos E, *Electr Power Syst Res*, 235 (2024) 110816.
4 Gallego Pareja L A, López-Lezama J M & Gómez Carmona O, *Results Eng*, 25 (2025) 104311.
5 Deeum S, Charoenchan T, Janjamraj N, Romphochai S, Baum S, Ohgaki H, Mithulananthan N & Bhumkittipich K, *Energies*, 16 (22) (2023) 7628.
6 Yang C, *et al.*, *Energies*, 16 (16) (2023) 5974.
7 Hemeida M G, Alkhalaf S, Senjyu T, Ibrahim A, Ahmed M & Bahaa-Eldin A M, *Ain Shams Eng J*, 12 (3) (2021) 2735.
8 Iqbal H, Stevenson A & Sarwat A I, *Electronics*, 14 (10) (2025) 1998.
9 Alanazi M, Alanazi A, Memon Z A, Awan A B & Deriche M, *Energies*, 17 (17) (2024) 4373.
10 Bouchikhi N, Boussadia F, Bouddou R, *et al*., *Sci Rep*, 15 (2025) 13919.
11 Akbari M A, Zare M, Azizipanah-Abarghooee R, *et al*., *Sci Rep,* 12 (2022) 10953.
12 Nguyen B M, Tran T, Nguyen T, *et al*., *Int J Comput Intell Syst*, 15 (2022) 90.
13 Bouaouda A & Sayouti Y, *Arch Comput Methods Eng*, 29 (2022) 4049.
14 Oelhaf J, *et al*., *Int J Electr Power Energy Syst*, 172 (2025) 111257.
15 Dalla Torre D, Lombardi A, Menapace A, *et al*., *Discov Appl Sci*, 6 (2024) 154.
16 Kahouli O, Alsaif H, Bouteraa Y, Ben Ali N & Chaabene M, *Appl Sci*, 11 (7) (2021) 3092.

17 Guo W, Wang G, Wang C & Wang Y, *Energy Rep*, 9 (8) (2023) 727.
18 Airlangga G & Liu A, *Mach Learn Knowl Extr*, 7 (1) (2025) 4.
19 Cheng S, Yu Z, Liu Y, *et al*., *Prot Control Mod Power Syst*, 7 (2022) 39.
20 Zhang T, Li W & Wang R, *Complex Intell Syst*, 9 (2023) 4075.
21 Nallolla C A, Chittathuru V P D & Padmanaban S, Electronics, 12 (4) (2023) 1062.
22 Price K V, Kumar A & Suganthan P N, "Trial-based dominance enables non-parametric tests to compare both the speed and accuracy of stochastic optimizers," (2022) doi: 10.48550/arXiv:2212.09423.
23 Klyushin D & Golubeva K, "Nonparametric multiple comparison test for change-point detection in big data," *in Proc 2020 IEEE 2nd Int Conf Adv Trends Inf Theory (ATIT)*, Kyiv, Ukraine, 2020 pp. 303.
24 Řehoř J, Fulemová J, Kutlwašer J, *et al*., *Int J Adv Manuf Technol*, 126 (2023) 675
25 Boroumandfar G, Khajehzadeh A, & Eslami M, *Sci Rep*, 14 (2024) 28195.
26 Sameh S S M, Moustafa H E D, AbdelHay E H, *et al*., *J Supercomput*, 80 (2024) 141.
27 Pradhan K, Chawla P & Rawat S, *J Ambient Intell Human Comput*, 14 (2023)12933.
28 Abdul-Kareem A A & Al-Jawher W A M, *Int J Comput Appl*, 46 (5) (2024) 324.
29 Dinesh K & Svn S K, *Peer-to-Peer Netw Appl*, 17 (2024) 585.
30 Santhosh D T, Jyothi K K, Srilakshmi K, *et al*., *Sci Rep*, 15 (2025) 33465.
31 Pandian S, Vedi A, Manasea Selvin B J A & Ganesan R, *Cybern Syst*, 56 (5) (2023) 448.
32 Kumar M, Namrata K & Kumar N, *J Grid Comput*, 21(2023) 28.
33 Kumar M, Namrata K & Kumari N, *Concurrency Comput Pract Exper*, 34 (23) (2022) e7190.
34 Kumar M, Namrata K & Samadhiya A, *Int J Emerg Electr Power Syst*, 22 (5) (2021) 629.
35 Kumar N, Namrata K & Samadhiya A, *Adv Theory Simul*, 5 (5) (2022) 2100639.
36 Kumar N, Namrata K & Samadhiya A, *Appl Intell,* 53 (6) (2023) 6604.
37 Kumar N, Namrata K & Samadhiya A, *Sustain Energy Technol Assess*, 55 (2023) 102950.
38 Namrata K, Kumar M & Kumar N, *Adv Electr Electron Eng*, 20 (4) (2022) 549.
39 Pal K, Akella A K, Namrata K & Pati S, "Classification of fault using artificial neural network and power quality improvement using DVR in a PV integrated hybrid power system," *in Proc 2022 Int Conf Intell Controller Comput Smart Power (ICICCSP)*, 2022 pp. 1.
40 Samadhiya A & Namrata K, *Complex Intell Syst*, 8 (2022) 989.