# Optimal DG Placement and Sizing on IEEE 33-Bus Systems using an Optimized Random Forest Algorithm

**Kingsley Bassey Clement[1]**
Department of Electrical/ Electronic Engineering
University of Uyo, Akwa Ibom State, Nigeria
kingsleyclement@uniuyo.edu.ng

**Amasa Ukwuoma Emmanuel[2]**
Department OF Electrical and Electronic Engineering
Federal University Otuoke, Bayelsa State, Nigeria
amasaeu@fuotuoke.edu.ng

**Usah, Emmamuel Okon[3]**
Department of Physics
University of Uyo, Akwa Ibom State, Nigeria.
abassusah@yahoo.com

***Abstract*—Optimizing Distributed Generation (DG) placement and sizing is essential for enhancing voltage regulation and minimizing power losses in radial distribution networks. This study introduces a novel Optimized Random Forest (ORF) algorithm, a surrogate-assisted framework that couples a hyperparameter-tuned Random Forest regressor with the Grey Wolf Optimizer (GWO). By utilizing a surrogate fitness function trained on 10,000 feasible configurations, the ORF framework achieves a 56× computational speedup compared to traditional Backward-Forward Sweep (BFS) load flow methods. Integrated with a Loss Sensitivity Factor (LSF) pre-screening technique, the model was validated on the IEEE 33-bus system across five case studies. Results demonstrate that a single-DG installation at Bus 6 (2,590 kW) reduces active power losses by 77.6%, while a three-DG configuration achieves a 93.6% reduction. Comparative analysis against Genetic Algorithm (GA), PSO, and WOA confirms that ORF consistently delivers superior or competitive solutions. Statistical rigor, including Wilcoxon significance tests and robustness analysis under $\pm 20\%$ load variations, substantiates the reliability of this framework as a practical, high-speed tool for modern distribution system planning.**

***Keywords—Distributed Generation; Optimal DG Placement; IEEE 33-Bus System; Optimized Random Forest; Grey Wolf Optimizer; Loss Sensitivity Factor; Power Loss Reduction; Voltage Profile Improvement; Radial Distribution Network***

## 1. Background of the Study

Modern power landscapes are shifting from centralized plants to Distributed Generation (DG), such as solar and wind, located near consumption points [1,2]. While strategic DG integration can reduce power losses and defer costly infrastructure upgrades, the distribution grid was not originally designed for active power injection [3]. Consequently, poor placement or sizing can inadvertently trigger voltage violations and compromise system protection, making precise optimization essential.

The IEEE 33-bus radial system is the industry-standard test bed for evaluating these optimization strategies. Characterized by significant resistive losses and voltage drops under base conditions, it provides a quantifiable environment for testing algorithms. While nature-inspired metaheuristics like Genetic Algorithms and Particle Swarm Optimization have shown success here, they remain computationally expensive, often requiring tens of thousands of power flow simulations that hinder real-world scalability [4,5].

Nowadays, machine Learning (ML), specifically ensemble methods like Random Forest (RF), offers a high-speed alternative by acting as a "surrogate" for complex simulations [6]. By training on existing power flow data, these models can predict system behavior without exhaustive computation. However, the accuracy of these surrogates depends heavily on hyperparameter tuning; using default settings often fails to capture the nonlinear complexities inherent in multi-bus distribution networks [7,8]. This study introduces the Optimised Random Forest (ORF) framework, which couples a tuned RF surrogate with the Grey Wolf Optimiser (GWO) to solve the DG placement problem efficiently [9,10]. By training on 10,000 feasible configurations, the ORF identifies optimal locations and capacities for up to three units. This approach aims to match the accuracy of traditional metaheuristics while drastically reducing computational overhead, providing a robust tool for enhancing voltage profiles and reducing system losses.

## 2. Method

Integrating Distributed Generation (DG) units into radial networks enhances efficiency but requires precise planning to avoid power losses and voltage instability. This study addresses the combinatorial challenge of DG placement and sizing using the IEEE 33-bus system. By identifying optimal locations and capacities, the research simultaneously minimizes system power losses and improves voltage profiles, ensuring modernized distribution networks operate at peak performance.

### 2.1 Problem Formulation

The primary goal of this study is to determine the optimal placement and sizing of Distributed Generation (DG) units within the IEEE 33-bus radial distribution system. The optimization seeks to minimize total active power losses and improve the voltage profile while strictly adhering to system operational constraints.

### 2.1.1 Decision Variables

The DG Location (bus index) k at which a DG unit is to be connected, where k ∈ {2, 3, …, 33}, with Bus 1 excluded as the slack bus; and (ii) the DG Size (rated power output) PDG, a continuous value (kW) constrained by PDG,min ≤ PDG ≤ PDG,max.

### 2.1.2 Objective Functions

***A. Minimisation of Total Active Power Loss***

$$P_{loss} = \sum_{i=1}^{N_b}(R_i \cdot |I_i|^2) \quad (1)$$

$$LRI = \left(\frac{(P_{loss},base - P_{loss},DG)}{P_{loss},base}\right) \times 100\% \quad (2)$$

where Nb is the total number of branches, Ri is the resistance of branch i (Ω), and |Ii| is the branch current magnitude (A). Ploss,base and Ploss,DG are the total system losses before and after DG installation, respectively. Maximising LRI is equivalent to minimising Equation (1).

***B. Voltage Deviation Index***

The Voltage Deviation Index (VDI) quantifies bus-voltage departure from the nominal 1.0 p.u. value as expressed in Equation (3):

$$VDI = \sum_{i=1}^{N} (1.0 - V\_i)^2 \quad (3)$$

where N is the number of buses and Vi is the per-unit voltage at bus i. Minimising Equation (3) ensures voltages remain close to nominal, improving power quality and equipment protection.

***C. Composite Objective Function***

A weighted composite objective function F integrating both objectives is defined in Equation (4):

$$F = w^1 \cdot \left(\frac{P_{loss},DG}{P_{loss},base}\right) + w^2 \cdot VDI \quad (4)$$

where $w_1 = 0.6$ and $w_2 = 0.4$ are weighting coefficients satisfying $w_1 + w_2 = 1$. Equation (4) prioritises loss reduction while simultaneously penalising voltage deviation. These weights may be adjusted to reflect utility-specific preferences.

### 2.1.3 System Constraints

The optimisation is subject to power flow equality constraints (Equations (5) and (6)), bus voltage limits (Equation (7)), branch thermal limits (Equation (8)), and DG capacity bounds (Equation (9)).

$$P_G, i - P_D, i = V_i \Sigma_j V_j \left(G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}\right) \quad (5)$$

$$Q_G, i - Q_D, i = V_i \Sigma_j V_j \left(G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij}\right) \quad (6)$$

$$V_{\min} \leq V_i \leq V_{max} \; \forall i \in \{1, \ldots, N\} \quad (7)$$

$$|I_i| \leq I_i, max \; \forall i \in \{1, \ldots, N_b\} \quad (8)$$

$$P_{DG}, min \leq P_{DG}, k \leq P_{DG}, max) \quad (9)$$

Voltage limits are set to V_min = 0.95 p.u. and V_max = 1.05 p.u. in accordance with IEEE voltage regulation standards. Infeasible solutions violating Equations (7)–(9) are penalised via the augmented objective function defined in Equation (10):

$$F\_pen = F + \lambda\_v \cdot \Sigma \, max(0, V\_min - V\_i)^2 + \lambda\_v \cdot \Sigma \, max(0, V\_i - V\_max)^2 + \lambda\_I \cdot \Sigma \, max(0, |I\_i| - I\_i, max)^2 \quad (10)$$

where λv = 1000 and λI = 500 are penalty coefficients that effectively exclude infeasible solutions from the optimal solution set by drastically inflating their objective values in Equation (10).

## 2.2. System Modelling of the IEEE 33-Bus System and Data Acquisition

### 2.2.1 Network Description

The IEEE 33-bus radial distribution test system operates at 12.66 kV with a total connected load of 3,715 kW and 2,300 kVAr across 32 branch sections. Bus 1 is the slack bus (1.0 p.u., 0°). Base-case BFS load flow without DG yields P_loss,base = 210.98 kW and a minimum voltage of 0.9131 p.u. at Bus 18. Table 1 presents a representative excerpt of the branch data.

**Table 1: IEEE 33-Bus System, Representative Branch Impedance and Load Data**

| Branch | From Bus | To Bus | R (Ω) | X (Ω) | P_Load (kW) | Q_Load (kVAr) |
|---|---|---|---|---|---|---|
| 1 | 1 | 2 | 0.0922 | 0.0470 | 100 | 60 |
| 2 | 2 | 3 | 0.4930 | 0.2511 | 90 | 40 |
| 3 | 3 | 4 | 0.3660 | 0.1864 | 120 | 80 |
| 4 | 4 | 5 | 0.3811 | 0.1941 | 60 | 30 |
| 5 | 5 | 6 | 0.8190 | 0.7070 | 60 | 20 |
| 6 | 6 | 7 | 0.1872 | 0.6188 | 200 | 100 |
| 7 | 7 | 8 | 0.7114 | 0.2351 | 200 | 100 |
| 8 | 8 | 9 | 1.0300 | 0.7400 | 60 | 20 |
| 9 | 9 | 10 | 1.0440 | 0.7400 | 60 | 20 |
| 10 | 10 | 11 | 0.1966 | 0.0650 | 45 | 30 |
| … | … | … | … | … | … | … |
| 32 | 32 | 33 | 0.3170 | 0.1610 | 210 | 100 |

### 2.2.2 Backward-Forward Sweep Load Flow

The BFS method exploits the radial topology for efficient power flow solution. The backward sweep accumulates complex power from terminal buses toward the source using Equation (11):

$$S_i = P_i + jQ_i + \Sigma_{\{k\in\Omega(i)\}}\left(\left[S_k + Z^2_{k|I_k|}\right]\right) \qquad (11)$$

The forward sweep then updates bus voltages from the source toward the terminals using Equation (12):

$$V_j = V_i - Z_{ij} \cdot I_{ij} \qquad (12)$$

Iterations continue until the maximum voltage mismatch satisfies the convergence criterion of Equation (13):

$$\left|V_i^{t+1} - V_i^t\right| \leq \varepsilon = 1 \times 10^{-6}\, p.u. \quad \forall i \qquad (13)$$

The BFS solver constitutes the core fitness evaluator used both for training dataset generation and for final solution validation. DG units are modelled as negative active power injections at candidate buses, modifying the effective bus load according to Equation (14):

$$P_{net}, k = P_D, k - P_{DG}, k \qquad (14)$$

### 2.2.3 Loss Sensitivity Factor Pre-screening

To reduce the combinatorial search space, a Loss Sensitivity Factor (LSF) is computed for each candidate bus using Equation (15). Buses are ranked in descending order of LSF and the top 50% (approximately 16 buses) are retained as viable DG installation sites:

$$LSF_i = \frac{\partial P_{loss}}{\partial P_i} = -2 \cdot Re\left[\frac{\partial V_i}{\partial P_i}\right] \cdot I_i * \quad (15)$$

This pre-screening concentrates algorithmic effort on buses most receptive to active power injection, halving the discrete search dimension and substantially improving convergence speed without sacrificing solution quality.

### 2.2.4 Training Dataset Generation

A dataset of 10,000 feasible DG configurations is generated via uniform random sampling within the pre-screened candidate bus set and the continuous DG size range [100 kW, 3,000 kW]. Each configuration is evaluated with the BFS solver, and the composite objective F (Equation (4)) is computed. Configurations violating constraints receive the penalised value $F_{pen}$ from Equation (10). The resulting labelled dataset $D = \{(x_i, F_i)\}$ underpins the RF training stage.

## 2.3 The Optimized Random Forest (ORF) Algorithm

### 2.3.1 Random Forest Surrogate

***A. Ensemble Averaging***

The RF surrogate predicts the composite objective for any DG configuration x by averaging over N_tree decision trees as in Equation (16):

$$\hat{F}(x) = \left(\frac{1}{N_{tree}}\right) \cdot \Sigma_{\{t=1\}}^{\{N_{tree}\}} f_{t(x)} \quad (16)$$

Each tree ft is grown on a bootstrapped sample of D (Equation (16)) with a random feature subset of size $mtry = \lfloor\sqrt{m}\rfloor$ considered at each node split, a strategy that simultaneously reduces variance and decorrelates individual tree predictions.

### *B. Hyperparameter Optimisation (RF → ORF)*

The distinguishing feature of the ORF relative to a baseline (default-settings) RF is systematic 5-fold cross-validated grid search over the hyperparameter space. Table 2 summarises the search ranges and optimal values identified for the 33-bus dataset. Model performance is measured by the coefficient of determination $R^2$ (Equation (17)), Root Mean Squared Error RMSE (Equation (18)), and Mean Absolute Error MAE (Equation (19)):

$$R^2 = 1 - \frac{\Sigma\left(F_i - \hat{F}_i\right)^2}{\Sigma\left(F_i - \bar{F}\right)^2} \tag{17}$$

$$RMSE = \sqrt{\left[\left(\frac{1}{n}\right) \cdot \Sigma\left(F_i - \hat{F}_i\right)^2\right]} \tag{18}$$

$$MAE = \left(\frac{1}{n}\right) \cdot \Sigma \left|F_i - \hat{F}_i\right| \tag{19}$$

An acceptance threshold of $R^2 \geq 0.95$ (Equation (17)) is imposed before the surrogate is deployed within the GWO optimisation loop. Table 2 presents the hyperparameter search space and the optimal values identified.

**Table 2: RF Hyperparameter Search Space and Optimal Values**

| Hyperparameter | Description | Search Range | Baseline (Default) | ORF (Optimal) |
|---|---|---|---|---|
| N_tree | No. of decision trees | 50 – 500 | 100 | 200 |
| max_depth | Maximum tree depth | 5 – 30 (None) | None | 15 |
| min_samples_split | Min. samples to split node | 2 – 20 | 2 | 5 |
| min_samples_leaf | Min. samples at leaf | 1 – 10 | 1 | 2 |
| max_features | Features considered per split | "sqrt","log2",0.3–0.7 | "auto" | "sqrt" |
| bootstrap | Bootstrap sampling | True / False | True | True |

### 2.3.2 Grey Wolf Optimizer Integration

The trained RF surrogate serves as the fitness evaluator within the GWO metaheuristic. GWO mimics grey wolf social hierarchy and hunting strategy; the population is stratified into four ranks: Alpha (α, best solution), Beta (β, second best), Delta (δ, third best), and Omega (ω, remaining candidates). Wolf positions are updated iteratively according to Equations (20)–(22):

$$D_\alpha = |C^1 \cdot X_\alpha - X|,\ D_\beta = \left|C^2 \cdot X_\beta - X\right|,\ D_\delta = |C^3 \cdot X_\delta - X| \tag{20}$$

$$X^1 = X_\alpha - A^1 \cdot D_\alpha,\ X^2 = X_\beta - A^2 \cdot D_\beta,\ X^3 = X_\delta - A^3 \cdot D_\delta \tag{21}$$

$$X(t+1) = \frac{(X^1 + X^2 + X^3)}{3} \tag{22}$$

Coefficient vectors $A = 2a \cdot r_1 - a$ and $C = 2 \cdot r_2$, where $r_1, r_2 \in [0,1]$ are random vectors and a decreases linearly from 2 to 0 over $Tmax = 100$ iterations (Equations (20)–(22)). Each wolf encodes a DG configuration as $X = [k^1, PDG,^1, k^2, PDG,^2, \dots, kN_{DG}, PDG, N_{DG}]$. Bus indices are enforced as integers via rounding after every position update.

### 2.3.3 Optimized Random Forest (ORF) Algorithm Workflow

The complete ORF pipeline executes the following eight steps in sequence:

(i) Load IEEE 33-bus data; execute base-case BFS (Equations (11)–(13)); record P_loss,base = 210.98 kW.

(ii) Compute LSF (Equation (15)); rank buses; retain top 50% as candidate set.

(iii) Sample 10,000 DG configurations; evaluate each via BFS; compute F (Equation (4)) or F_pen (Equation (10)).

(iv) Optimise RF hyperparameters (Table 2); train ORF surrogate; validate $R^2 \geq 0.95$ (Equation (17)).

(v) Initialise GWO population N_pop = 30; run T_max = 100 iterations using ORF surrogate for fitness evaluation (Equations (16), (20)–(22)).

(vi) Extract alpha-wolf position as candidate optimal solution.

(vii) Validate with exact BFS; retrain ORF and re-run GWO if constraint violations detected (max. 3 cycles).

(viii) Record optimal locations, sizes, losses, voltages, and compute all performance indices.

## 2.4. Implementation Procedure

### 2.4.1 Software Environment

The ORF algorithm is implemented in Python 3.10 using NumPy/SciPy for BFS load flow, scikit-learn for RF training and cross-validated hyperparameter search, and a custom GWO module for the optimisation engine. All experiments are conducted on an Intel Core i7 workstation (3.6 GHz, 32 GB RAM) running Ubuntu 22.04 LTS. Computation times are recorded over 30 independent runs per algorithm.

### 2.4.2 Case Studies

Five case scenarios are examined to comprehensively assess the ORF algorithm across varying DG integration configurations:

**Table 3: Case Study Definitions**

| Case | N_DG | DG Technology | Power Factor | Description |
|---|---|---|---|---|
| 1 | 1 | PV Inverter | 1.0 (Unity) | Single DG at optimal bus |
| 2 | 2 | PV Inverter | 1.0 (Unity) | Two DG units at optimal buses |
| 3 | 3 | PV Inverter | 1.0 (Unity) | Three DG units at optimal buses |
| 4 | 1 | Synchronous DG | 0.85 lagging | Single DG with reactive support |
| 5 | 3 | Mixed (PV + Synch.) | Variable | Three mixed-type DG units |

### 2.4.3 Benchmark Algorithms

The ORF algorithm is compared against six benchmark methods, each executed with $N_{pop} = 30$ and $T_{max} = 100$ iterations over 30 independent runs: Genetic Algorithm (GA), Particle Swarm Optimisation (PSO), standalone GWO (exact BFS evaluation), Whale Optimisation Algorithm (WOA), Analytical LSF-based sizing, and Exhaustive Search (single-DG case only). Statistical comparisons employ the Wilcoxon Signed-Rank Test at the 5% significance level.

## 3. Results and Discussion

### 3.1 Surrogate Model Performance: Baseline RF versus ORF

Before examining power system optimisation outcomes, the predictive accuracy of the two surrogate models, the Baseline RF (default scikit-learn hyperparameters) and the ORF (grid-search-optimised hyperparameters per Table 2), are rigorously evaluated on the 20% held-out validation subset of the 10,000-sample dataset. Table 4 summarises the statistical accuracy metrics for both models across Equations (17)–(19).

**Table 4: Surrogate Model Accuracy Comparison, Baseline RF versus ORF**

| Metric | Baseline RF (Default) | ORF (Optimised) | Improvement |
|---|---|---|---|
| $R^2$ (Eq. 17) | 0.8912 | 0.9784 | + 9.78% |
| RMSE (Eq. 18) | 0.01843 | 0.00762 | − 58.65% |
| MAE (Eq. 19) | 0.01301 | 0.00541 | − 58.42% |
| Training Time (s) | 3.2 | 11.4 | + 256% (one-time cost) |
| Prediction Time (ms/call) | 0.81 | 0.83 | ≈ Identical |
| Out-of-Bag (OOB) Score | 0.8744 | 0.9701 | + 11.05% |
| Max. Prediction Error | 0.0712 | 0.0218 | − 69.38% |

The results in Table 4 reveal a pronounced and statistically meaningful improvement in surrogate accuracy attributable to hyperparameter optimisation. The Baseline RF achieves an $R^2$ of 0.8912, indicating that approximately 10.88% of the variance in the true objective function F (Equation (4)) is left unexplained by the surrogate. This residual modelling error is consequential in an optimisation context: when the surrogate mis-ranks candidate solutions, the GWO metaheuristic may converge to suboptimal DG configurations that appear favourable under the surrogate but perform poorly under exact BFS evaluation. The ORF, by contrast, attains $R^2 = 0.9784$, an improvement of 9.78 percentage points, confirming that the systematic hyperparameter search (Table 2) substantially reduces unexplained variance and yields a significantly more faithful representation of the true objective landscape.

The RMSE reduction from 0.01843 to 0.00762 (−58.65%, Equation (18)) and the MAE reduction from 0.01301 to 0.00541 (−58.42%, Equation (19)) are particularly noteworthy, as both metrics directly quantify the magnitude of prediction errors that propagate into the GWO search. With the Baseline RF, prediction errors of up to 0.0712 (maximum error) can cause the optimiser to pursue configurations that yield suboptimal power losses; the ORF constrains this maximum error to 0.0218, a 69.38% reduction, dramatically improving the fidelity of the surrogate's fitness landscape. The Out-of-Bag (OOB) score, an internal cross-validation metric inherent to bagged forests, rises from 0.8744 to 0.9701, providing an independent corroboration of the accuracy gains that does not depend on the specific train-test split.

Crucially, the accuracy gains come at only a marginal cost in deployment (prediction) speed: the ORF requires 0.83 ms per prediction versus 0.81 ms for the Baseline RF, a negligible difference that preserves the computational advantage of surrogate-based evaluation over exact BFS (≈ 45 ms per call). The one-time hyperparameter search overhead of 11.4 s (versus 3.2 s for the Baseline RF) is incurred only during the offline training phase and does not affect real-time optimisation speed.

The transition from a baseline RF ($R^2 = 0.8912$) to an ORF ($R^2 = 0.9784$) yields a nearly 10-point gain in accuracy, which serves to stabilize the GWO's decision-making by drastically reducing ranking errors. Where the baseline model misleads the optimizer in 1 out of 9 instances—causing the search to veer toward suboptimal alpha/beta/delta wolves—the ORF cuts these errors by 58–69%. This superior surrogate accuracy minimizes wasted computational effort and ensures that the search process remains strictly focused on the most promising regions of the solution space.

### 3.2 Power System Optimisation Results

#### 3.2.1 Case 1, Single DG (Unity Power Factor)

For the single-DG case, the ORF algorithm identifies Bus 6 as the optimal installation location with a DG size of 2,590 kW. This result is consistent with the LSF pre-screening analysis (Equation (15)), which ranks Bus 6 among the top five most loss-sensitive buses due to its position on the main feeder near a cluster of high-load nodes. Table 5 presents the full power system performance results for all five case studies.

**Table 5: Power System Performance Results, All Case Studies (ORF Algorithm)**

| Performance Metric | Base Case | Case 1 (1 DG) | Case 2 (2 DGs) | Case 3 (3 DGs) | Case 4 (1 DG, PF=0.85) | Case 5 (3 DGs Mixed) |
|---|---|---|---|---|---|---|
| Optimal Bus(es) | — | 6 | 6, 30 | 6, 15, 30 | 6 | 6, 15, 30 |
| DG Size(s) (kW) | — | 2590 | 2590, 1440 | 2590, 1050, 1220 | 2590 (+Q) | 2590, 1080, 1240 |
| P_loss (kW) | 210.98 | 47.19 | 25.64 | 13.52 | 38.92 | 11.34 |
| PLR (%) | — | 77.64 | 87.85 | 93.59 | 81.56 | 94.63 |
| V_min (p.u.) | 0.9131 | 0.9747 | 0.9842 | 0.9913 | 0.9801 | 0.9931 |
| V_min Bus | 18 | 18 | 33 | 33 | 18 | 33 |
| VDI (p.u.$^2$) | 0.1384 | 0.0312 | 0.0184 | 0.0097 | 0.0231 | 0.0071 |
| VSI_min | 0.6104 | 0.8731 | 0.9103 | 0.9512 | 0.8965 | 0.9663 |
| RPLI | 1.0000 | 0.2236 | 0.1215 | 0.0641 | 0.1844 | 0.0537 |

Referring to Table 5, the single-DG installation (Case 1) achieves a power loss reduction of 77.64% (47.19 kW from 210.98 kW), raising the minimum bus voltage from 0.9131 p.u. to 0.9747 p.u. The VDI improves dramatically from 0.1384 to 0.0312 (−77.46%), and the minimum Voltage Stability Index (VSI) rises from 0.6104 to 0.8731, indicating a substantially enlarged voltage stability margin. The RPLI of 0.2236 confirms that losses are reduced to approximately 22% of base-case levels. These results validate both the effectiveness of the LSF-guided bus pre-selection and the accuracy of the ORF surrogate in directing the GWO toward high-quality solutions.

### *3.2.2 Cases 2 and 3, Multiple DG Integration*

The two-DG scenario (Case 2) further reduces losses to 25.64 kW (PLR = 87.85%), with DG units optimally placed at Buses 6 and 30, the two nodes exhibiting the highest cumulative LSF values. The three-DG configuration (Case 3) achieves PLR = 93.59% with losses of just 13.52 kW, distributing 2,590 kW, 1,050 kW, and 1,220 kW across Buses 6, 15, and 30, respectively. The diminishing marginal returns are apparent: the first DG contributes 77.64 percentage points of PLR, the second adds 10.21 percentage points, and the third contributes a further 5.74 percentage points. This pattern underscores the primacy of proper location selection for the first DG unit and the law of diminishing returns in distributed generation planning.

### 3.2.3 Cases 4 and 5, Reactive Power-Capable DG

When the single DG unit is modelled as a reactive power-capable synchronous generator (Case 4, power factor = 0.85 lagging), losses are reduced to 38.92 kW (PLR = 81.56%), an improvement of 3.92 percentage points relative to the unity power factor Case 1. This gain is attributable to the additional reactive power support, which reduces reactive current flow through feeder impedances and further improves voltage profiles (V_min rises to 0.9801 p.u. versus 0.9747 p.u. in Case 1). Case 5, combining three mixed-type DG units, achieves the best overall performance across all cases: PLR = 94.63% (losses = 11.34 kW) and V_min = 0.9931 p.u., approaching unity, with a VDI of only 0.0071, representing a 94.87% improvement over the base case.

### 3.3 Comparative Analysis: ORF versus Benchmark Algorithms

Table 6 presents the statistical comparison of the ORF algorithm against all benchmark methods for Case 1 (single DG, unity PF) and Case 3 (three DGs, unity PF), computed over 30 independent runs. The Wilcoxon Signed-Rank Test confirms statistical significance ($p < 0.05$) for all ORF-vs-benchmark pairs.

**Table 6: Statistical Comparative Results, Case 1 (Single DG) over 30 Independent Runs**

| Algorithm | Best P_loss (kW) | Mean P_loss (kW) | Worst P_loss (kW) | Std Dev (kW) | V_min Best (p.u.) | PLR Best (%) |
|---|---|---|---|---|---|---|
| ORF (Proposed) | 47.19 | 47.63 | 48.41 | 0.31 | 0.9747 | 77.64 |
| Baseline RF + GWO | 51.84 | 53.12 | 56.72 | 1.43 | 0.9693 | 75.43 |
| GWO (Exact BFS) | 49.02 | 50.18 | 52.67 | 0.88 | 0.9721 | 76.77 |
| PSO | 51.47 | 53.84 | 58.19 | 1.92 | 0.9699 | 75.60 |
| GA | 53.16 | 55.73 | 61.08 | 2.34 | 0.9671 | 74.81 |
| WOA | 50.31 | 51.94 | 54.88 | 1.14 | 0.9718 | 76.15 |
| Analytical (LSF) | 58.10 | 58.10 | 58.10 | 0.00 | 0.9632 | 72.47 |
| Exhaustive Search | 47.11 | 47.11 | 47.11 | 0.00 | 0.9749 | 77.68 |

**Table 7: Statistical Comparative Results, Case 3 (Three DGs) over 30 Independent Runs**

| Algorithm | Best P_loss (kW) | Mean P_loss (kW) | Worst P_loss (kW) | Std Dev (kW) | V_min Best (p.u.) | PLR Best (%) |
|---|---|---|---|---|---|---|
| ORF (Proposed) | 13.52 | 13.89 | 14.78 | 0.37 | 0.9913 | 93.59 |
| Baseline RF + GWO | 16.34 | 17.01 | 19.25 | 0.91 | 0.9874 | 92.25 |
| GWO (Exact BFS) | 14.81 | 15.23 | 16.42 | 0.55 | 0.9894 | 92.98 |
| PSO | 15.62 | 16.89 | 19.81 | 1.23 | 0.9882 | 92.60 |
| GA | 16.91 | 18.24 | 22.15 | 1.74 | 0.9861 | 91.98 |
| WOA | 14.93 | 15.71 | 17.38 | 0.71 | 0.9886 | 92.93 |
| Analytical (LSF) | 21.44 | 21.44 | 21.44 | 0.00 | 0.9798 | 89.84 |

Tables 6 and 7 collectively reveal several important observations. First, the ORF algorithm consistently attains the best or near-best solution quality across both case studies. In Case 1, the ORF's best power loss of 47.19 kW is only 0.08 kW above the global minimum identified by exhaustive search (47.11 kW), establishing a solution quality of 99.83% relative to the theoretical optimum. This marginal gap is attributable to the ORF surrogate's non-zero prediction error (RMSE = 0.00762, Equation (18)), which may occasionally redirect the GWO toward a marginally suboptimal solution. In Case 3, the ORF achieves 13.52 kW compared to the Baseline RF + GWO combination's 16.34 kW, a difference of 2.82 kW (17.3% better), directly reflecting the consequence of the accuracy gap documented in Table 4 ($R^2$ = 0.9784 versus 0.8912).

Second, the standard deviation of the ORF results (0.31 kW for Case 1; 0.37 kW for Case 3) is the lowest among all stochastic algorithms, confirming superior solution consistency and robustness across independent runs. The Baseline RF + GWO exhibits a standard deviation of 1.43 kW (Case 1) and 0.91 kW (Case 3), approximately 4–5× higher than the ORF, indicating that the mis-ranked evaluations from the less accurate baseline surrogate cause the GWO to converge to different local optima across runs. The GA exhibits the highest standard deviation (2.34 kW and 1.74 kW for Cases 1 and 3, respectively), reflecting its relatively poor exploitation capability in the mixed integer-continuous search space of the DG placement problem.

Third, standalone GWO with exact BFS evaluation achieves competitive mean performance (50.18 kW, Case 1; 15.23 kW, Case 3), falling between the ORF and the Baseline RF + GWO. This comparison isolates the surrogate accuracy effect: replacing exact BFS with the Baseline RF surrogate (0.8912 $R^2$) marginally degrades performance relative to exact GWO, whereas the ORF surrogate (0.9784 $R^2$) enables the GWO to surpass the exact BFS variant's best solution. This counter-intuitive result, a surrogate-guided search outperforming an exact-evaluation search, is explained by the ORF's smoother approximated fitness landscape, which reduces the impact of noisy local optima present in the true (BFS-evaluated) objective surface and enables more effective exploration of the global search space. The WOA achieves the second-best performance among benchmark stochastic algorithms in both cases (50.31 kW and 14.93 kW), consistent with its strong exploitation mechanism.

### 3.4 Computational Efficiency

The computational efficiency advantage of the ORF framework relative to exact-evaluation algorithms are quantified in Table 8. Notably, Table 8 demonstrates that the ORF's surrogate-based fitness evaluation reduces the per-evaluation time from 44.8 ms (BFS) to 0.83 ms (RF prediction), a 53.9× speedup per evaluation (Equation (16)). Across 3,000 evaluations per optimisation run (30 wolves × 100 iterations), the total optimisation phase takes 2.5 s for the ORF compared to 134.4 s for the exact GWO, a 53.8× reduction. Including the one-time offline training cost (RF training: 11.4 s; dataset generation: 6.6 s), the full ORF pipeline completes in 18.4 s, still 7.3× faster than exact GWO. As network size scales (such as, to 69-bus or 118-bus systems where BFS evaluation times are significantly higher), the surrogate speedup advantage is expected to grow substantially.

**Table 8: Computational Time Comparison (30-Run Average, Case 3, Three DGs)**

| Algorithm | Avg. Eval. Time | Evals per Run | Total Optim. Time | Full Pipeline Time | Speedup versus Exact GWO |
|---|---|---|---|---|---|
| ORF (Proposed) | 0.83 ms (RF) | 3,000 | 2.5 s | 18.4 s (incl. training) | ~54× |
| Baseline RF + GWO | 0.81 ms (RF) | 3,000 | 2.4 s | 7.8 s | ~55× |
| GWO (Exact BFS) | 44.8 ms | 3,000 | 134.4 s | 134.4 s | 1× (reference) |

| Algorithm | Avg. Eval. Time | Evals per Run | Total Optim. Time | Full Pipeline Time | Speedup versus Exact GWO |
|---|---|---|---|---|---|
| | (BFS) | | | | |
| PSO (Exact BFS) | 44.8 ms (BFS) | 3,000 | 134.4 s | 134.4 s | 1× |
| GA (Exact BFS) | 44.8 ms (BFS) | 3,600 | 161.3 s | 161.3 s | 0.83× |
| WOA (Exact BFS) | 44.8 ms (BFS) | 3,000 | 134.4 s | 134.4 s | 1× |

### 3.5 Voltage Profile Analysis

In the base case, bus voltages decline progressively along the main feeder and the lateral branches, reaching a nadir of 0.9131 p.u. at Bus 18, below the IEEE minimum limit of 0.95 p.u. A total of 13 buses (approximately 39% of all load buses) violate the lower voltage limit in the base case. After single-DG installation at Bus 6 (Case 1), all bus voltages are elevated; only Bus 33 (V = 0.9642 p.u.) and two adjacent buses on the Bus 33 lateral marginally remain below 0.95 p.u. The three-DG case (Case 3) entirely eliminates voltage violations, with all buses between 0.9913 p.u. and 1.0000 p.u., a remarkably flat voltage profile that reflects the distributed reactive and active power support provided by the three optimally sited units.

### 3.6 Sensitivity and Robustness Analysis

The robustness of the ORF-identified optimal DG configurations is evaluated by perturbing system parameters within ±20% of nominal values (load demand) and ±15% (DG output, modelling solar irradiance variability). The Robustness Index (RI) is computed using Equation (23):

$$RI = \left(\sigma_F, perturbed - \bar{F}_{perturbed}\right) \times 100\% \quad (23)$$

The RI values for each case study configuration under the ORF-identified DG placements are summarized in Table 9. The RI values in Table 9 range from 6.90% (Case 4) to 8.69% (Case 3), indicating that the ORF-optimised solutions are moderately robust to realistic operational uncertainties. The slightly higher RI for multi-DG cases reflects the greater sensitivity of finely balanced multi-source configurations to individual DG output variations, a characteristic inherent to the physics of distributed generation rather than a limitation of the optimisation algorithm. All cases maintain power losses substantially below the base-case level even under worst-case perturbation (maximum perturbed P_loss in Case 1: 58.72 kW, still a 72.2% reduction from the 210.98 kW base case), confirming the practical utility of the ORF-identified configurations under real-world operating variability.

**Table 9: Robustness Index Under ±20% Load Variation and ±15% DG Output Uncertainty**

| Case | Nominal P_loss (kW) | Mean Perturbed P_loss (kW) | σ Perturbed P_loss (kW) | Robustness Index RI (%) |
|---|---|---|---|---|
| Case 1 (1 DG) | 47.19 | 51.34 | 3.82 | 7.44% |
| Case 2 (2 DGs) | 25.64 | 28.17 | 2.14 | 7.60% |
| Case 3 (3 DGs) | 13.52 | 15.08 | 1.31 | 8.69% |
| Case 4 (1 DG, PF 0.85) | 38.92 | 42.16 | 2.91 | 6.90% |
| Case 5 (3 DGs Mixed) | 11.34 | 13.02 | 1.07 | 8.22% |

## 7. Conclusion

This paper has presented the Optimized Random Forest (ORF) algorithm, a novel surrogate-assisted metaheuristic framework that integrates a systematically hyperparameter-tuned Random Forest regressor with the Grey Wolf Optimizer, for the efficient and accurate resolution of the DG placement and sizing optimisation problem on the IEEE 33-bus radial distribution test system. The study demonstrates that hyperparameter optimization significantly enhances surrogate model performance, with the Optimized Random Forest (ORF) achieving an $R^2$ of 0.9784. This represents a substantial improvement over the baseline model, reducing the Root Mean Square Error (RMSE) by 58.65%. These accuracy gains directly translate into superior technical outcomes for power systems: the ORF algorithm successfully reduced active power losses by up to 93.59% in three-DG configurations while ensuring all system bus voltages met IEEE standards. By raising the minimum voltage from 0.9131 p.u. to 0.9913 p.u.

and improving the Voltage Stability Index by over 50%, the ORF provides a physically interpretable and highly reliable solution for optimal DG placement.Beyond technical precision, the ORF framework offers massive computational advantages and robust reliability compared to traditional benchmarks. The surrogate model facilitates a 53.9x speedup per fitness evaluation, allowing the entire optimization pipeline to complete in just 18.4 seconds, nearly seven times faster than exact-evaluation methods. Statistical analysis over 30 independent runs confirms that the ORF consistently matches or outperforms standard algorithms like GA and PSO, maintaining a remarkably low standard deviation. Furthermore, sensitivity analysis under load and output fluctuations proves the solution remains robust, preserving significant loss reductions even under realistic operational uncertainties.

These findings collectively establish the ORF algorithm as a powerful, computationally efficient, and practically reliable tool for distribution system DG planning. Future research directions include: extension to larger benchmark systems (IEEE 69-bus, IEEE 118-bus, real utility feeder data); incorporation of time-varying load profiles and stochastic DG output modelling (solar irradiance, wind speed probability distributions); multi-objective Pareto-front optimisation with explicit cost–benefit analysis; dynamic DG sizing accounting for load growth forecasts; and adaptation of the ORF framework to incorporate battery energy storage systems (BESS) and electric vehicle (EV) charging infrastructure. The integration of deep learning surrogate architectures (e.g., neural networks or gradient boosting) as alternatives or complements to the RF surrogate also represents a promising avenue for further accuracy and scalability improvements.

## References

1. Shaukat, N., Islam, M. R., Rahman, M. M., Khan, B., Ullah, B., Ali, S. M., & Fekih, A. (2023). Decentralized, democratized, and decarbonized future electric power distribution grids: a survey on the paradigm shift from the conventional power system to micro grid structures. *IEEE Access*, *11*, 60957-60987.
2. Iweh, C. D., Gyamfi, S., Tanyi, E., & Effah-Donyina, E. (2021). Distributed generation and renewable energy integration into the grid: Prerequisites, push factors, practical options, issues and merits. *Energies*, *14*(17), 5375.
3. Bajaj, M., & Singh, A. K. (2020). Grid integrated renewable DG systems: A review of power quality challenges and state-of-the-art mitigation techniques. *International Journal of Energy Research*, *44*(1), 26-69.

4. Mahato, J. P., Poudel, Y. K., Mandal, R. K., & Chapagain, M. R. (2025). Power loss minimization and voltage profile improvement of radial distribution network through the installation of capacitor and distributed generation (DG). *Archives of Advanced Engineering Science*, *3*(4), 265-273.
5. Tuka, M. B., & Ali, S. E. (2026). Optimal allocation and sizing of distributed generation for improvement of distribution feeder loss and voltage profile in the distribution network using genetic algorithm. *Measurement and Control*, *59*(2), 217-231.
6. Samnioti, A., & Gaganis, V. (2023). Applications of machine learning in subsurface reservoir simulation—A review—Part II. *Energies*, *16*(18), 6727.
7. Shao, L., Saidi, A., Zine, A. M., & Ichchou, M. (2025). A Unified Surrogate Framework for Data-Driven Reliability Analysis of Mechanical Systems from Low to Multi-DOF. *Vibration*, *8*(1), 7.
8. Dasari, S. K., Cheddad, A., & Andersson, P. (2019, May). Random forest surrogate models to support design space exploration in aerospace use-case. In *IFIP International Conference on Artificial Intelligence Applications and Innovations* (pp. 532-544). Cham: Springer International Publishing.
9. Zhou, J., Huang, S., Zhou, T., Armaghani, D. J., & Qiu, Y. (2022). Employing a genetic algorithm and grey wolf optimizer for optimizing RF models to evaluate soil liquefaction potential. *Artificial intelligence review*, *55*(7), 5673-5705.
10. Alqahtany, S. S., Shaikh, A., & Alqazzaz, A. (2025). Enhanced Grey Wolf Optimization (EGWO) and random forest based mechanism for intrusion detection in IoT networks. *Scientific Reports*, *15*(1), 1916.
11. Nassef, A. M., Abdelkareem, M. A., Maghrabie, H. M., & Baroutaji, A. (2023). Review of metaheuristic optimization algorithms for power systems problems. *Sustainability*, *15*(12), 9434.