Applied Energy 384 (2025) 125433

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Near real-time machine learning framework in distribution networks with low-carbon technologies using smart meter data

Emrah Dokur [a], Nuh Erdogan [b,*], Ibrahim Sengor [c], Ugur Yuzgec [d], Barry P. Hayes [e]

[a] Department of Electrical Electronics Engineering, Bilecik Seyh Edebali University, Bilecik, Turkiye
[b] Department of Engineering, School of Science and Technology, Nottingham Trent University, Nottingham, NG11 8NS, UK
[c] Department of Electrical and Electronic Engineering, Munster Technological University, Cork, T12 P928, Ireland
[d] Department of Computer Engineering, Bilecik Seyh Edebali University, Bilecik, Turkiye
[e] School of Engineering and Architecture, University College Cork, Cork, T12 K8AF, Ireland

ARTICLE INFO

Keywords:
Distribution networks
Low carbon technologies
Machine learning
Meta-heuristic
Single candidate optimizer
Smart meter
Voltage forecasting

ABSTRACT

The widespread adoption of low-carbon technologies, such as photovoltaics, electric vehicles, heat pumps, and energy storage units introduces challenges to distribution network congestion and power quality, particularly raising concerns about voltage stability. Enhanced voltage visibility in low-voltage networks is increasingly vital for active grid management, making efficient voltage forecasting tools essential. This study introduces a novel data-driven approach for forecasting node voltages in low-voltage networks with high penetration of low-carbon technologies. Using time series of power measurements from smart meter data, the study integrates an Extreme Learning Machine with the Single Candidate Optimizer to enhance computational efficiency and forecasting accuracy. The model is validated using smart meter datasets from two different low-voltage networks with low-carbon technologies and is compared with several established machine learning models. The results demonstrate that the optimization algorithm significantly improves the tuning of model parameters, achieving up to a 17-fold reduction in computation time compared to the fastest metaheuristic methods implemented. The proposed model demonstrated superior accuracy, with an average voltage deviation of 0.56%. Although the computation time per node achieved is not yet suitable for real time applications, the study shows that the optimization method significantly improves the performance of the forecasting tool.

1. Introduction

The global energy transition is accelerating with the increasing adoption of low-carbon technologies (LCTs) such as solar photovoltaic (PV) systems, plug-in electric vehicles (EVs), heat pumps (HPs) and energy storage units [1]. For example, the rooftop solar PV market increased by 50%, with installations reaching 118 GW, up from 79 GW the previous year [2]. Similarly, the home and public EV infrastructure expanded by 40% in 2023 compared to 2022 [3]. By 2030, HPs are projected to heat 40% of residential buildings and 65% of commercial buildings in Europe [4]. However, the higher integration of these technologies poses significant congestion and power quality challenges to low voltage (LV) distribution networks [5]. These challenges can cause voltage stability issues, resulting in inefficiencies and threats to network security [6]. Current network planning and operating decisions often rely on assumptions and estimates regarding voltage-related reinforcement without an automated real-time interface. An active grid management framework has been proposed to improve network hosting capacity and optimize the use of network assets [7]. This framework would allow distribution network operators to engage in proactive control room activities and address voltage issues in real time. To enable active grid management, there is a critical need for increased voltage visibility in LV networks. In this context, the use of voltage forecasting models becomes increasingly important in supporting decision-making for both distribution network planning and operational management.

Voltage calculations are fundamental to various functions within distribution networks. Emerging technologies rely on these calculations to ensure compliance with regulatory voltage thresholds during the operational phase. In the planning process, voltage calculations assess the impact of new connection requests and determine whether additional interventions are necessary [8]. Traditionally, power flow analyses are used for voltage calculations, which require detailed three-phase LV network models, including the number of consumers, consumer types, cable types, cable lengths, cable impedances, node points, and connected LCTs. However, detailed electrical models are often unavailable

* Corresponding author.
E-mail addresses: emrah.dokur@bilecik.edu.tr (E. Dokur), nuh.erdogan@ntu.ac.uk (N. Erdogan), ibrahim.sengor@mtu.ie (I. Sengor), ugur.yuzgec@bilecik.edu.tr (U. Yuzgec), barry.hayes@ucc.ie (B.P. Hayes).

https://doi.org/10.1016/j.apenergy.2025.125433
Received 17 September 2024; Received in revised form 13 January 2025; Accepted 23 January 2025
Available online 3 February 2025

Nomenclature

ϵ	Non-zero training error.
$\omega(t)$	Weight parameter.
ω_j	Input connection weight.
θ_j	Output layer connection weight.
b_j	Input connection bias.
i	The number of nodes.
L	The number of hidden layer samples.
N	The number of training samples.
P	Time series of active power measurements.
Q	Time series of reactive power measurements.
R^2	Coefficient of determination.
V	Time series of voltage measurements.
$X(k)$	Position of the candidate solution.
ABC	Artificial bee colony.
ANFIS	Adaptive neuro-fuzzy inference system.
ANN	Artificial neural network.
DE	Differential evolution.
ELM	Extreme learning machine.
EV	Plug-in electric vehicle.
GA	Genetic algorithm.
HP	Heat pump.
KELM	Kernel extreme learning machine.
LCT	Low-carbon technology.
LSTM	Long short-term memory.
LV	Low-voltage.
MAE	Mean absolute error.
ML	Machine learning.
ML-Meta-KELM	Multi-layer Meta-Kernel-ELM.
MLP	Multi-layer perceptron.
MSE	Mean squared error.
PSO	Particle swarm optimization
PV	Solar photovoltaic.
RMSD	Root-mean-square deviation.
RMSE	Root mean squared error.
SCO	Single Candidate Optimizer.
SLFN	Single Layer Feed-forward Networks.
SVM	Support Vector Machine.

for LV networks. Network topology, connectivity, phase grouping, and impedance information are often erroneous or incomplete, complicating this challenge. Furthermore, the rapid and unpredictable integration of LCTs into LV distribution networks exacerbates the challenge of accurately predicting voltage.

The increasing rollout of smart meters offers a potential model-free solution to this challenge by providing high-resolution data from LCTs. Smart meters can acquire significant amounts of relevant data with unique temporal precision, offering valuable insights into consumption patterns and aiding in decision-making processes related to distribution network planning and management thanks to forecasting methods. Using these extensive data, data-driven approaches are emerging as viable methods to improve voltage predictions in LCT-enriched LV distribution networks. In [9], smart meter data are used for short-term load forecasting at the LV level, demonstrating that simple load models derived from these data can achieve prediction accuracies similar to those of more complex forecasting approaches. In [10], smart meter data are utilized to develop energy forecasting and operational planning services, such as demand and voltage profile forecasting, for LV networks with significant distributed energy resources. The study shows that smart meter data can estimate voltages probabilistically in LV networks, providing early warnings of potential voltage issues.

The authors in [11] apply a data-driven approach to identify the network topology for LV distribution networks including the load phase connectivity from the smart meter data. Similarly, a multiple linear regression model is applied to smart meter data to estimate topology, line parameters, and customer and line phasing connections in LV networks in [12]. A 15-day data sample at a 60-m resolution (360 data points per meter) is shown to be cost-effective for data acquisition and accuracy, with the possibility of shortening the duration when using high-precision smart meters. In [13], a smart meter-driven model is proposed to assess the impact of distributed energy resources on voltage behavior. It is highlighted that while the accuracy of the voltage estimations depends on the accuracy of the estimated LV feeder impedances, which is influenced by the correlation among input variables (e.g., active and reactive power measurements), the resulting voltage calculations can be highly accurate. To address challenges with smart meter data, including limited resolution, small data volumes, and issues with uncorrelated or unclear information, there is a growing need for advanced machine learning (ML) models, such as neural networks.

ML-based forecasting uses smart meter data to predict voltage without relying on network topology or parameters. This approach involves developing a learning model (e.g., a multi-layer perceptron, MLP) that correlates active and reactive power measurements to predict the voltage at each node in the distribution network [14]. In [15], autoregressive and feed-forward Artificial Neural Networks (ANNs) are applied to forecast voltage harmonic distortions. It is demonstrated that a 10% penetration of smart meters in the LV network achieves good harmonic estimation performance. In [16], a physical-model-free approach employing deep reinforcement learning is proposed to control rapid voltage fluctuations in real time within PV-dominated distribution networks. In [17], a deep ANN model captures the nonlinear relationships between power measurements and node voltages using single-phase smart meter data in PV-rich LV networks. Building on this, the authors in [8] introduce a new methodology to incorporate upstream medium-voltage network effects and test it on multiple LV feeders. In [18], the voltage regulation is formulated as a convex optimization problem and solved using a convex Neural Network (NN) model for rapid response. In these approaches, learning models minimize voltage forecasting errors by optimizing parameters using meta-heuristic algorithms like particle swarm optimization (PSO), gray-wolf optimization, or stochastic methods such as ADAM [19]. The Single Candidate Optimizer (SCO) has recently gained attention for its reduced computation costs and memory requirements [20]. It converges to optimal solutions faster than other algorithms [21], though its effectiveness for quasi-real-time applications still requires further investigation. Therefore, selecting an appropriate ML method is crucial for ensuring accuracy, scalability, and fast computation. The relationship established between input smart meter data and voltage outputs for a given network can be extended to other LV networks. The key characteristics of the ML approaches discussed above are synthesized in Table 1.

This paper presents a novel hybrid model for forecasting node voltages in LV distribution networks with significant LCTs such as PVs, HPs, and EVs, enabling near real-time voltage management. Our methodology combines the SCO with the Extreme Learning Machine (ELM) to improve forecasting accuracy while reducing model complexity and avoiding overfitting. This results in more efficient and robust voltage predictions. The proposed model is tested using high-resolution smart meter data from two LV networks with LCTs and validated by comparing it with well-established and hybrid ELM-based models, including MLP, Long short-term memory (LSTM), and Adaptive neuro-fuzzy inference system (ANFIS). The evaluation is performed through computation time and various accuracy metrics.

Table 1
Overview of ML methods for voltage forecasting and control.

Ref.	Approach	Analysis type	Advantages	Limitations
[8]	ANN	Voltage forecasting	Eliminates the need for network topology data	Accuracy highly depends on the quality of input data
[9]	MLP	Load forecasting	High prediction accuracy for small datasets	Limited scalability to larger networks
[10]	Nonlinear NN & regression models	Voltage and load profile analysis	Improves robustness with probabilistic estimates	Computationally intensive
[14]	ELM	Voltage forecasting	High computational efficiency; low prediction error	Real-time application remains a challenge
[15]	ANN	Harmonic distortion prediction	Effective with low data availability	Requires extensive preprocessing of harmonic data
[16]	Deep Reinforcement Learning	Voltage fluctuation control	Handles rapid voltage fluctuations in PV systems	Requires significant training data for effectiveness
[18]	Convex NN	Voltage regulation	Ensures global solutions with optimization techniques	Computationally expensive
[22]	ELM	Voltage stability margin prediction	High computational efficiency; robust under uncertainty	Requires high-resolution input data
[23]	ANN, SVM, DT, RF	Voltage stability prediction	Enabling the selection of the most appropriate ML method	Heuristic method, lacking theoretical guarantees; sensitive to misleading data
[24]	Reinforcement learning	Real-time voltage control	Adaptive self-learning model for grid operation	No use of real-world data
[25]	Regression-based ML models	Voltage stability margin prediction	Low prediction error	Computationally intensive for large datasets

The remainder of this paper is organized as follows. Section 2 details the proposed approach, and the development of the model. Section 3 presents smart meter data analysis and discusses the testing and validation results, including identified limitations. Finally, Section 4 presents concluding remarks and outlines future research directions.

2. Methodology

2.1. Approach

The main structure of the proposed model architecture is presented in Fig. 1. Our methodology combines the advantages of the SCO algorithm with the ELM approach to update the input weights and hidden biases of the ELM model, thereby optimizing voltage forecasting results. Herein, our focus is to accelerate the computation time while increasing the accuracy. The ELM achieves significant training speed by randomly assigning input weights and biases while utilizing the Moore–Penrose generalized inverse equation to analytically determine the output weights. The ELM may need a higher number of hidden neurons due to the random determination of the input weights and hidden biases. Due to the stochastic selection of input weights and biases, the ELM algorithm may occasionally result in the hidden layer output matrix H of Single Layer Feed-forward Networks (SLFN) not being of full column rank, thus diminishing the efficacy of the ELM. To overcome these issues, the proposed model structure uses the SCO algorithm to determine the initial parameters of the ELM model. The key advantage of using the SCO algorithm for optimizing ELM model parameters is its ability to operate with a single candidate solution, resulting in faster processing times compared to population based metaheuristic algorithms [20]. As a result, the proposed SCO-based ELM model maintains competitive forecasting accuracy while reducing model complexity and avoiding overfitting. The steps summarizing the prediction process of the ML-based framework proposed in this study are as follows:

1. Data collection: Historical smart meter data were utilized from a three-phase LV distribution network in Ireland. The details of the collected dataset are provided in Section 3.1.
2. Data preprocessing: Data were normalized using the z-score technique to ensure consistency. The data set was then split into training and testing subsets, with 70% of the data used for training and the remaining 30% for testing. Input features included active and reactive powers and past voltage values, which were incorporated to improve forecasting accuracy.
3. ELM model parameter optimization: To enhance the ELM model's performance, the initial layer weights and bias values of the ELM model, which are typically assigned randomly, were optimized using the SCO algorithm. The SCO algorithm takes advantage of the single candidate optimization structure to efficiently determine the optimal initial weights and biases without causing significant time costs. This process was executed over 100 iterations, resulting in the optimized parameters for the ELM model's initial layer.
4. ELM model training: The ELM model, with its initial layer weights and bias values set by the SCO algorithm, uses real-time active and reactive power data along with the previous time step voltage value as inputs during training. The training subset of the dataset is used to determine the ELMmodel's, output layer weights. This completes the ELM model's training, enabling accurate forecasting based on the optimized parameters and input features.
5. ELM model testing: At this stage, the pre-trained ELM model is evaluated using 30% of the data set reserved exclusively for testing and not used during training. This process assesses the model's ability to generalize and respond to previously unseen data.

2.2. Extreme learning machine and variants

The ELM is an ML algorithm employing a neural network with a single hidden layer. In this network, the input layer's weights and biases are initialized randomly, while the output weights are determined through solving a linear system. Among various neural network architectures, the ELM has emerged as a robust and highly effective methodology [26]. ELM is notable for its rapid learning speed, ease of implementation, and excellent generalization in forecasting. This section provides a detailed overview of the ELM architecture and its variants, including Meta-ELM and Multi-Layer Meta-Kernel-ELM (ML-MetaKELM), highlighting their core components.

The basic framework of the ELM has a SLFN composed of three layers: the input layer, the hidden layer, and the output layer. Unlike conventional neural networks, where the weights linking the input and hidden layers undergo iterative adjustments via backpropagation, ELM adopts an alternative approach. In ELM, these weights are randomly assigned and remain fixed, resulting in an exceptionally rapid learning process. The output of the SLFN is determined based on the input and

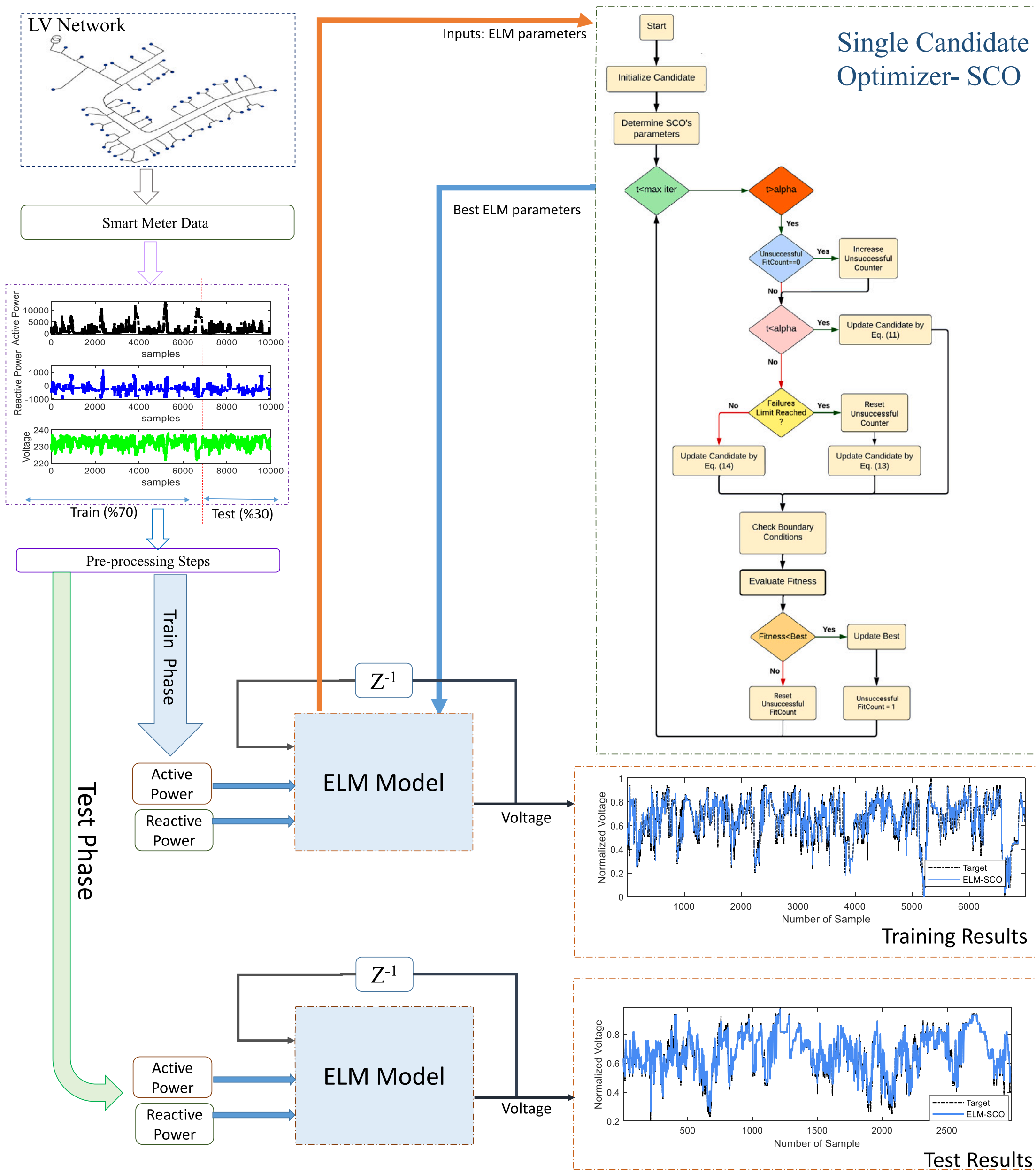

Fig. 1. Flowchart of the proposed integrated approach.

connections as follow:

$$f(x) = \sum_{j=1}^{N_h} \theta_j \cdot \varphi(x \cdot \omega_j + b_j). \tag{1}$$

Here, the input connection weights are denoted as ω_j, biases as b_j, and θ_j represents the weight values connecting the hidden layer outputs to the output layer. The number of hidden layers, denoted by N_h, along with the bias and input link weights, is randomly determined. The analytical computation of θ_k is carried out through the steps detailed below. In this study, input data $x \in \mathbb{R}^n$ is generated using the active power, reactive power and one-step previous voltage forecasting result. Here, n denotes the number of inputs.

With Eq. (1) covering N training samples, we can derive N equations accordingly. These equations can be clearly represented using matrix–vector notation, denoted as matrix H.

$$H = \begin{bmatrix} \varphi(x_1\omega_1 + b_1) & \cdots & \varphi(x_1\omega_{N_h} + b_{N_h}) \\ \vdots & \vdots & \vdots \\ \varphi(x_N\omega_1 + b_1) & \cdots & \varphi(x_N\omega_{N_h} + b_{N_h}) \end{bmatrix}_{NxN_h} \tag{2}$$

The equation below represents the output weights and their corresponding target values for each output.

$$T = H\gamma, \quad \gamma = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_{N_h} \end{bmatrix}, \quad T = \begin{bmatrix} t_1 \\ \vdots \\ t_N . \end{bmatrix} \tag{3}$$

The determination of the output connection weights requires computing the inverse of the Moore–Penrose matrix H.

$$\hat{\gamma} = H^{+}T. \tag{4}$$

The Meta-ELM network utilizes several ELMs, with each ELM trained on a distinct subset of the dataset. The output connection weights of the combined ELMs are established using the ELM learning rule, leveraging the entire dataset. Integrating kernel functions into the ELM can enhance stability and generalization. This upgraded iteration is known as the Kernel Extreme Learning Machine (KELM). Consequently, the Kernel theorem is formulated as follows:

$$K_{ELM} = HH^T, \text{and } K_{ELM}\left(x_i, x_j\right) = h(x_i)h(x_j), \tag{5}$$

$$f(x) = h(x)H^t(K_{ELM} + \frac{1}{C})^{-1}T, \tag{6}$$

Based on previous research findings [27,28], the regularization parameter C is set to 50 in this study.

Kernel functions satisfy the Mercer condition, making them suitable for practical applications. Various kernel functions are available, such as the Polynomial, Gaussian, Hyperbolic tangent (Sigmoid), and Wavelet kernels. In this paper, the Radial Basis function Kernel is analyzed. In the context of the Kernel-based Meta ELM, which employs the kernel function $K(x_i, x_j)$, the equation is formulated as follows:

$$\left\|K_{NxL}\beta - T\right\|^2 < \epsilon, \tag{7}$$

In this context, ϵ denotes a non-zero training error, while L represents the number of hidden layer samples, with N consistently exceeding L. To mitigate the error, norm minimization and sample number constraints are applied, resulting in the following formulation of the optimization problem:

$$Minimize \;\; \frac{1}{2}\|\beta\|^2 + \frac{1}{2}C\left\|\left(K_{NxL}\beta - T\right)\right\|^2. \tag{8}$$

The output vector can be represented as follows:

$$\beta = \left[\frac{1}{C} + K_{NxL}^T \; K_{NxL}\right]^{-1} K_{NxL}T. \tag{9}$$

The predicted output is derived as:

$$f(x) = \begin{bmatrix} K(x, x_1) \\ K(x, x_2) \\ \vdots \\ K(x, x_L) \end{bmatrix} \left[\frac{1}{C} + K_{NxL}^T \; K_{NxL}\right]^{-1} K_{NxL}T. \tag{10}$$

2.3. Single candidate optimizer based ELM model

The SCO algorithm is a unique optimization method that differs from traditional search algorithms by focusing solely on a single candidate solution during the optimization process. Unlike swarm-based algorithms, which involve the concurrent exploration of multiple particles, SCO focuses its efforts solely on refining one solution, thereby amplifying the efficacy of optimization outcomes [29]. Dividing the optimization process into two phases, SCO employs various strategies and mechanisms designed to enhance both exploration and exploitation capabilities. This unique combination of the single-candidate approach and the two-phase strategy makes SCO a powerful and versatile optimization algorithm [20]. In the first phase of SCO, the candidate solution's position is updated according to the following formula:

$$X(k) = \begin{cases} X_b(k) + w(t) \cdot |X_b(k)| & \text{if } r_1 < 0.5 \\ X_b(k) - w(t) \cdot |X_b(k)| & \text{otherwise,} \end{cases} \tag{11}$$

$$w(t) = e^{-(b.t/t_{max})^b}, \tag{12}$$

where, $X(k)$ denotes the position of the candidate solution, where k is the dimension, $w(t)$ is the weight parameter, $X_b(k)$ indicates the best candidate for each iteration, b is a constant parameter, t is the iteration, t_{max} refers to the maximum iteration, and r_1 is a random number between 0 and 1.

During the second phase of the SCO algorithm, a systematic exploration is initiated, starting with the immediate vicinity of the superior position identified in the initial stage. Unlike the first phase, this stage employs a counter that increments with each iteration where no improvement is observed in the cost value of the updated candidate solution. When the counter reaches a predefined threshold for unsuccessful attempts (e.g., m = 5), it is reset, and the candidate solution's position is updated using a specific mechanism. As the phase progresses, the search becomes increasingly focused, enabling a detailed evaluation of the most promising regions while ensuring comprehensive coverage of the search space. The process guiding these adjustments is systematically defined by the equation presented below:

$$X(k) = \begin{cases} X_b(k) + r_3 \cdot \left(u_b(k) - l_b(k)\right) & \text{if } r_2 < 0.5 \\ X_b(k) - r_4 \cdot \left(u_b(k) - l_b(k)\right) & \text{otherwise,} \end{cases} \tag{13}$$

where, $u_b(k)$ and $l_b(k)$ denote the boundary values, r_2, r_3, and r_4 represent random numbers. The Eq. (13) for updating the position is fundamental in facilitating the transition of the candidate solution from exploitation to exploration within the SCO algorithm. This transition is critical for assisting the algorithm in breaking away from local optima, where the candidate solution risks becoming trapped in sub-optimal regions of the search space. By implementing this adjustment to the position update mechanism, the SCO algorithm can effectively traverse multiple regions of the search space, thus minimizing the potential for sub-optimal solutions and improving the search for superior optimization outcomes. If, during this phase, the counter value remains below the predefined threshold m, the SCO algorithm determines the updated position of the candidate solution based on Eq. (14) as follows:

$$X(k) = \begin{cases} X_b(k) + w(t) \cdot r_6 \cdot \left(u_b(k) - l_b(k)\right) & \text{if } r_5 < 0.5 \\ X_b(k) - w(t) \cdot r_6 \cdot \left(u_b(k) - l_b(k)\right) & \text{otherwise.} \end{cases} \tag{14}$$

The SCO process can be used to optimize the parameters of forecast models. In the authors' previous study in [14], the SCO algorithm was employed to determine the optimal MLP parameters. Compared to well-known metaheuristic approaches including particle swarm optimization and differential evolution, the SCO-based model demonstrated superior performance. This scenario has encouraging promising outcomes regarding the SCO's integration with more innovative and faster algorithms like ELM.

In the basic ELM, the parameters are typically initialized randomly or stochastically, a practice that can result in inefficient ELM performance, including potential overfitting issues. To address this challenge, this paper uses the SCO algorithm to determine the initial parameters of the ELM model. In Eq. (1), the weights and biases are input to the SCO process, which outputs the best optimal parameters based on the single candidate solution described above. During the training process, the weights are adjusted to minimize the error between the forecasted output and the actual output. The SCO is adapted to tune the weight and bias values of the ELM parameters in this paper. This proposed hybrid model is called ELM-SCO. In this paper, the ELM-SCO relates active power (P), reactive power (Q) measurements from smart meter data along with one-step-ahead forecasted voltage results (V) in the input layer to voltage measurements in the output layer. Our main aim is to predict voltage deviations for each node in a real LV network. Active power, reactive power and voltage measurements in one-minute resolution for each node collected for a rural LV network with heavily low-carbon loads are used. The determination of input is of great importance for the accuracy of the model. By utilizing historical time series measurements of active power $P_i = [P_{i,1}, P_{i,2}, \ldots, P_{i,T}]$, reactive power $Q_i = [Q_{i,1}, Q_{i,2}, \ldots, Q_{i,T}]$, and node voltage, $V_{i-1} = [V_{i-1,1}, V_{i-1,2}, \ldots, V_{i-1,T}]$, the ELM-SCO model is first trained to capture the nonlinear relationships between the inputs (P_i, Q_i and V_{i-1}) and the corresponding outputs (V_{i+1}) for the $i+1$th node, given by Eq. (15). Both w_{jk} and b_j are adjusted to improve the performance of the model. The voltage forecasting for each node is represented as follows:

$$V_{i+1} = f_{ELM-SCO}\left(P_i, Q_i, V_{i-1}, w_{jk,i}, b_{j,i}\right) \tag{15}$$

Fig. 2 shows the flow diagram of the operation of the SCO-based ELM model.

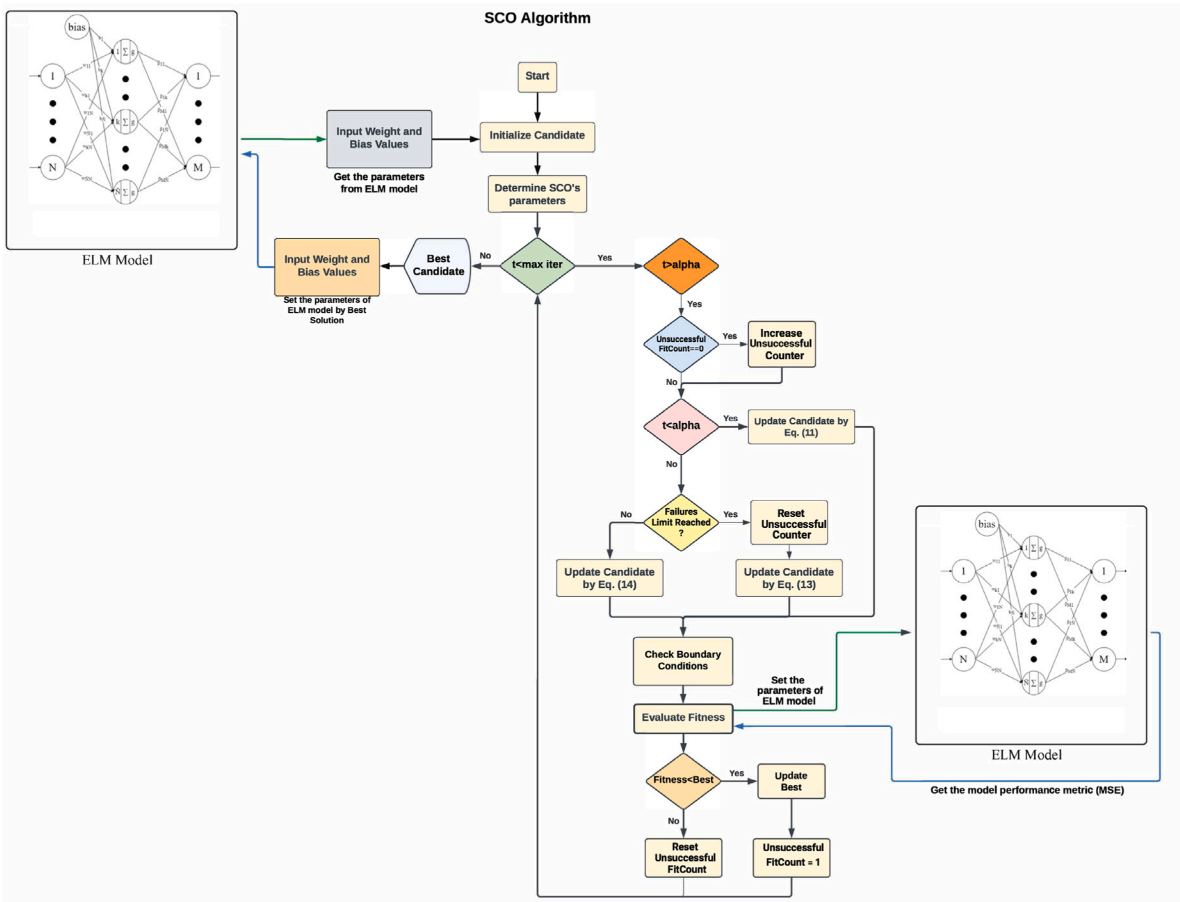

Fig. 2. Flowchart of the proposed SCO based ELM model (ELM-SCO).

3. Experimental results

3.1. Smart meter data presentation and analysis

In this study, the historical smart meter data extracted from a three-phase LV distribution network in Ireland are used, where several single-phase customers are equipped with various LCTs such as HPs, EVs, rooftop PVs, and battery energy storage units. The LV distribution network considered in this study reflects the characteristics of typical Irish networks, comprising a mixture of overhead and underground conductors. It predominantly serves residential customers in rural and coastal areas, with some industrial and commercial users included. This study builds on the framework established in our previous work [30], which provides a detailed analysis of similar LV networks and their operational characteristics. 37 dwellings were equipped with detailed data logging metering infrastructure from September to January. Net active and reactive power exchange with the grid, node voltage variation, active and reactive power coming from rooftop PVs, HP power consumption, and EV charging power values, if available, were obtained. Fig. 3 shows the active and reactive power, and voltage profiles for a dwelling over seven consecutive days. As seen from the box plots, although the median values of both active and reactive power display a trend, both datasets contain many outliers. These highly volatile power variations make forecasting challenging.

A subset of 10,000 data points, as shown in Fig. 4, was utilized for the analysis of active and reactive powers and node voltages. As shown in the figure, the node voltages fall within the range of the nominal line-to-neutral voltage $U_n \pm 10\%$ (253 V; 207 V), in compliance with the specifications outlined in the European Standard EN50160.

Table 2
The computation time (s) results of hybrid MLP algorithms using meta-heuristic algorithms.

Iteration	100	500	3000
PSO-MLP	17.30	48.33	318.51
DE-MLP	12.26	47.24	289.80
GA-MLP	12.77	55.38	334.80
ABC-MLP	24.3	99.8	558.28
SCO-MLP	2.15	3.58	16.68

3.2. Testing and validation

We first evaluate the performance of the SCO approach in terms of computation time, comparing it to several well-known metaheuristic methods. Specifically, we apply the SCO to optimize the MLP model and compare it with optimization techniques including PSO, Differential evolution (DE), Genetic Algorithm (GA), and Artificial Bee Colony (ABC). All experiments are implemented in Matlab R2020b on Windows 10 with a 2.5-GHz Intel Core i5 7200U processor and a 64-bit operating system with 8 GB of RAM. Computation times for each method are reported in Table 2. Comparisons are made across 100, 500, and 3000 iterations. As shown, the SCO-MLP approach demonstrates significantly faster computation times. Notably, with 3000 iterations, the SCO algorithm outperforms the fastest DE approach by over 17-fold. Despite the time disadvantage of hybrid metaheuristic approaches, the SCO-MLP approach proves to be substantially more efficient.

Secondly, we assess the performance of the proposed ELM-SCO model during the training phase. To validate the model, we compared it against three alternative ELM-based models: ELM, Meta-ELM, and Multi-layer Meta-Kernel-ELM (ML-Meta-KELM). For a comprehensive

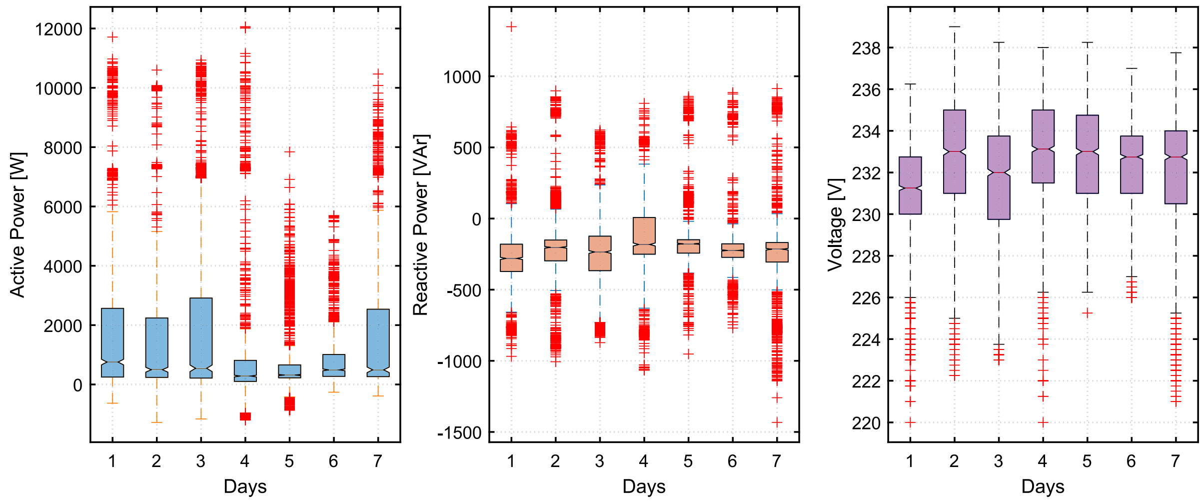

Fig. 3. Active and reactive power, and voltage profiles.

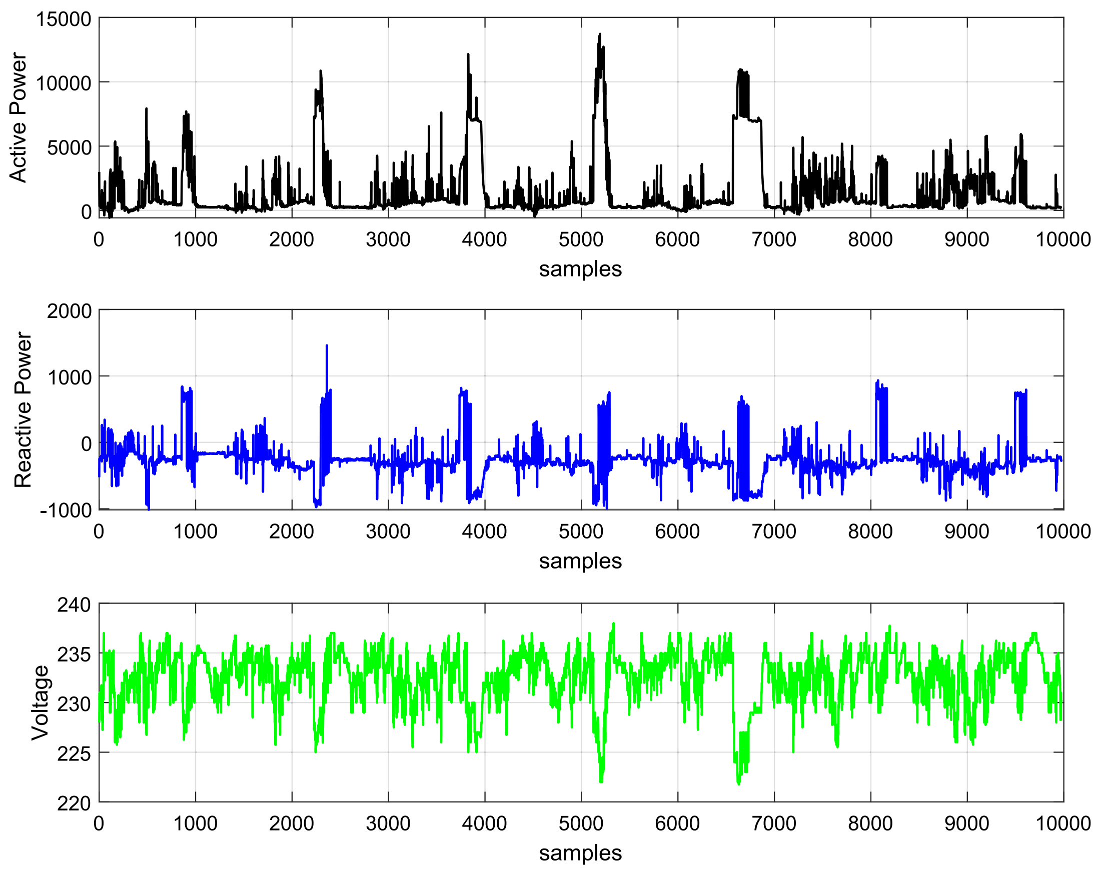

Fig. 4. Time series of smart meter data selected: active power (top), reactive power (middle), and node voltage (bottom).

evaluation, we also included three widely recognized ML models: MLP, LSTM, and ANFIS. These models were selected for their established capabilities in time-series forecasting. While they have not previously been applied specifically to voltage forecasting in the literature, their architectures and learning mechanisms are well-suited to address challenges similar to those encountered in voltage forecasting, such as capturing temporal dependencies (LSTM) and modeling complex nonlinear relationships (ANFIS). MLP is a traditional feed-forward NN, widely used for regression and forecasting tasks.

The SCO algorithm was used to optimize the initial parameters of all implemented models. Their performance was evaluated using Mean Squared error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Coefficient of determination (R^2) metrics. Table 3 summarizes the results, demonstrating that the ELM-SCO model achieved the lowest MSE of 0.0018, outperforming both ELM-based and traditional models. It also achieved the lowest MAE of 0.0279 and the highest R^2 of 0.9343, reflecting superior accuracy, fewer prediction errors, and better explanatory power. Among ELM-based models, ELM-SCO outperformed ELM (MSE: 0.0020, R^2: 0.9310), Meta-ELM (MSE: 0.0024, R^2: 0.9152), and ML-Meta-KELM (MSE: 0.0022, R^2: 0.9235). Similarly, it surpassed traditional ML models: MLP (MSE: 0.0022, R^2: 0.9169), LSTM (MSE: 0.0023, R^2: 0.9210), and ANFIS

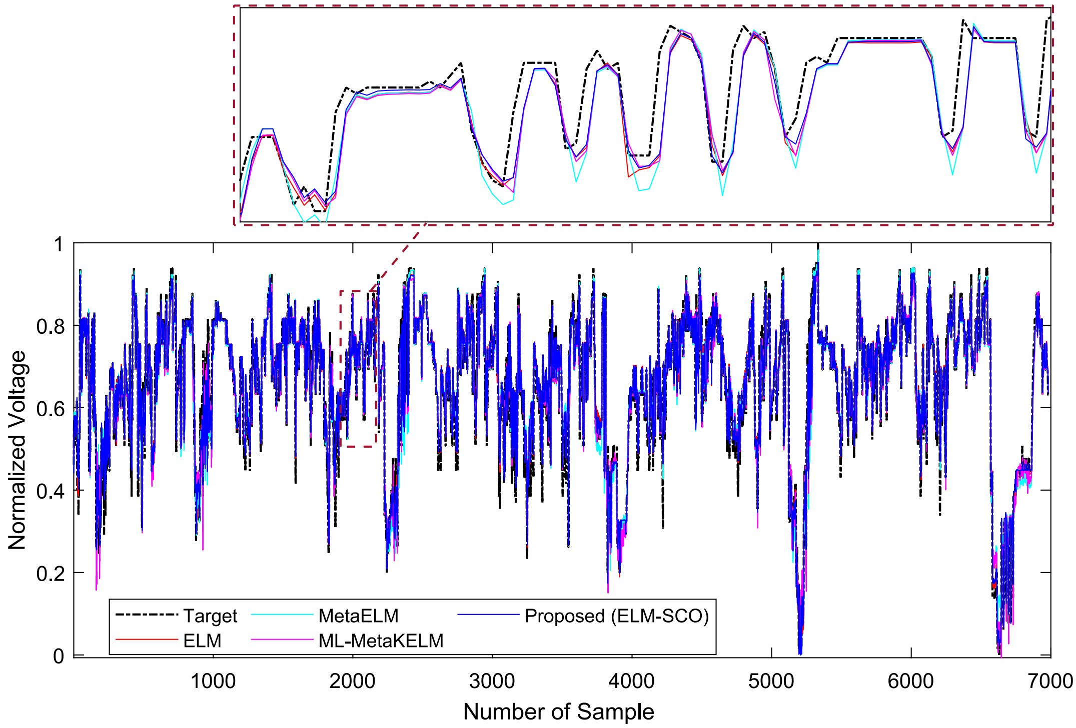

Fig. 5. Training results of the implemented ML models.

(MSE: 0.0019, R^2: 0.9337). While ANFIS achieved a comparable R^2, ELM-SCO consistently demonstrated better overall performance across all metrics.

Fig. 5 demonstrates the training results of the implemented models, showing the normalized voltage over the number of samples. The proposed ELM-SCO model closely follows the target values, indicating its superior accuracy during the training phase. Compared to the ELM, Meta-ELM, and ML-Meta-KELM, the ELM-SCO model demonstrates a better fit, with fewer deviations from the target values. This figure underscores the quantitative findings in Table 4, highlighting the robustness of the ELM-SCO model in the training phase.

Finally, Table 4 presents a comparison of the testing performance for the implemented ML models. The results show that the ELM-SCO model consistently outperforms all other models across all metrics. It achieves the lowest MSE (0.0023), RMSE (0.0487), and MAE (0.0321), demonstrating superior accuracy and predictive reliability. It has also been observed that it outperforms, including MLP (MSE: 0.0022, RMSE: 0.0469, MAE: 0.0309), LSTM (MSE: 0.0023, RMSE: 0.0475, MAE: 0.0325), and ANFIS (MSE: 0.0019, RMSE: 0.0522, MAE: 0.0327). Furthermore, the ELM-SCO model achieves the highest R^2 value of 0.8834, demonstrating its superior ability to explain the variance in the data. While ANFIS achieves the highest R^2 value of 0.9337, among traditional ML models, followed by LSTM (0.9210) and MLP (0.9169), all remain below the performance of the ELM-SCO model. These findings highlight the ELM-SCO model's effectiveness in improving predictive accuracy and reliability across diverse ML approaches.

Fig. 6 presents the test results of the implemented models, highlighting the normalized voltage over the number of samples. The ELM-SCO model consistently outperforms the other implemented models, closely aligning with the target values. This consistency between training and testing phases underscores the proposed model's reliability and generalization capabilities, making it a strong candidate for practical voltage forecasting applications.

Fig. 7 displays the R^2 values for the implemented models, highlighting the proportion of variance explained during both the training and testing phases. The ELM-SCO model achieves the highest R^2 value, underscoring its superior ability to capture underlying data patterns.

Table 3
Train forecasting performance results of the implemented ML models.

Models	MSE	RMSE	MAE	R^2
MLP	0.0022	0.0469	0.0309	0.9169
LSTM	0.0023	0.0475	0.0325	0.9210
ANFIS	0.0019	0.0522	0.0327	0.9337
ELM	0.0019	0.0443	0.0289	0.9310
Meta-ELM	0.0024	0.0491	0.0330	0.9152
ML-Meta-KELM	0.0021	0.0467	0.0309	0.9235
Proposed (ELM-SCO)	**0.0018**	**0.0433**	**0.0279**	**0.9343**

Table 4
Test forecasting performance results of the implemented ML models.

Models	MSE	RMSE	MAE	R^2
MLP	0.0026	0.0510	0.0344	0.8724
LSTM	0.0027	0.0515	0.0362	0.8697
ANFIS	0.0027	0.0488	0.0327	0.8663
ELM	0.0024	0.0491	0.0327	0.8816
Meta-ELM	0.0027	0.0521	0.0355	0.8663
ML-Meta-KELM	0.0026	0.0510	0.0342	0.8718
Proposed (ELM-SCO)	**0.0023**	**0.0487**	**0.0321**	**0.8834**

This metric is crucial for understanding how well the model captures the underlying patterns in the data, and the ELM-SCO model's high R^2 value signifies its effectiveness in this regard. Fig. 8 presents a Taylor diagram that summarizes model performance in terms of standard deviation, correlation coefficient, and root-mean-square deviation (RMSD). The ELM-SCO model is the closest to the ideal point, demonstrating superior performance across these metrics. This indicates that the ELM-SCO model achieves a strong balance between high correlation with the target values, low standard deviation, and minimal RMSD, making it a robust and effective forecasting tool.

To validate the generalizability and performance of the proposed ELM-SCO model, we utilized one second resolution residential power consumption and voltage smart meter dataset collected from a typical household in Alabama [31]. The results demonstrate that the ELM-SCO model consistently outperformed other models across all evaluation

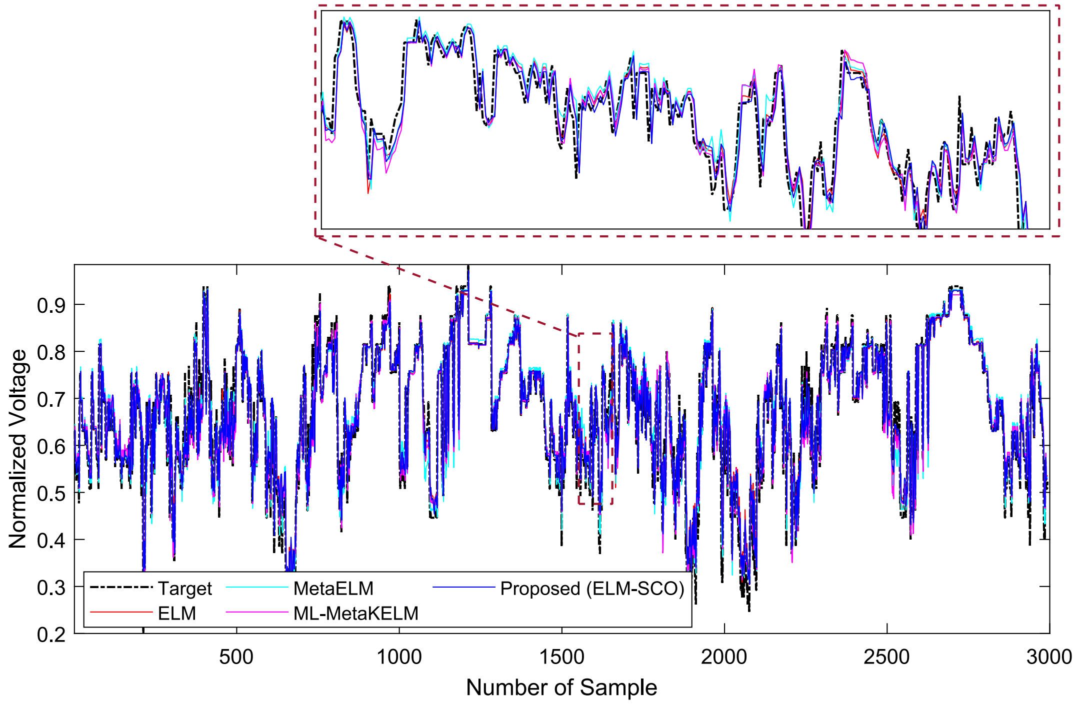

Fig. 6. Test results of the implemented ML models.

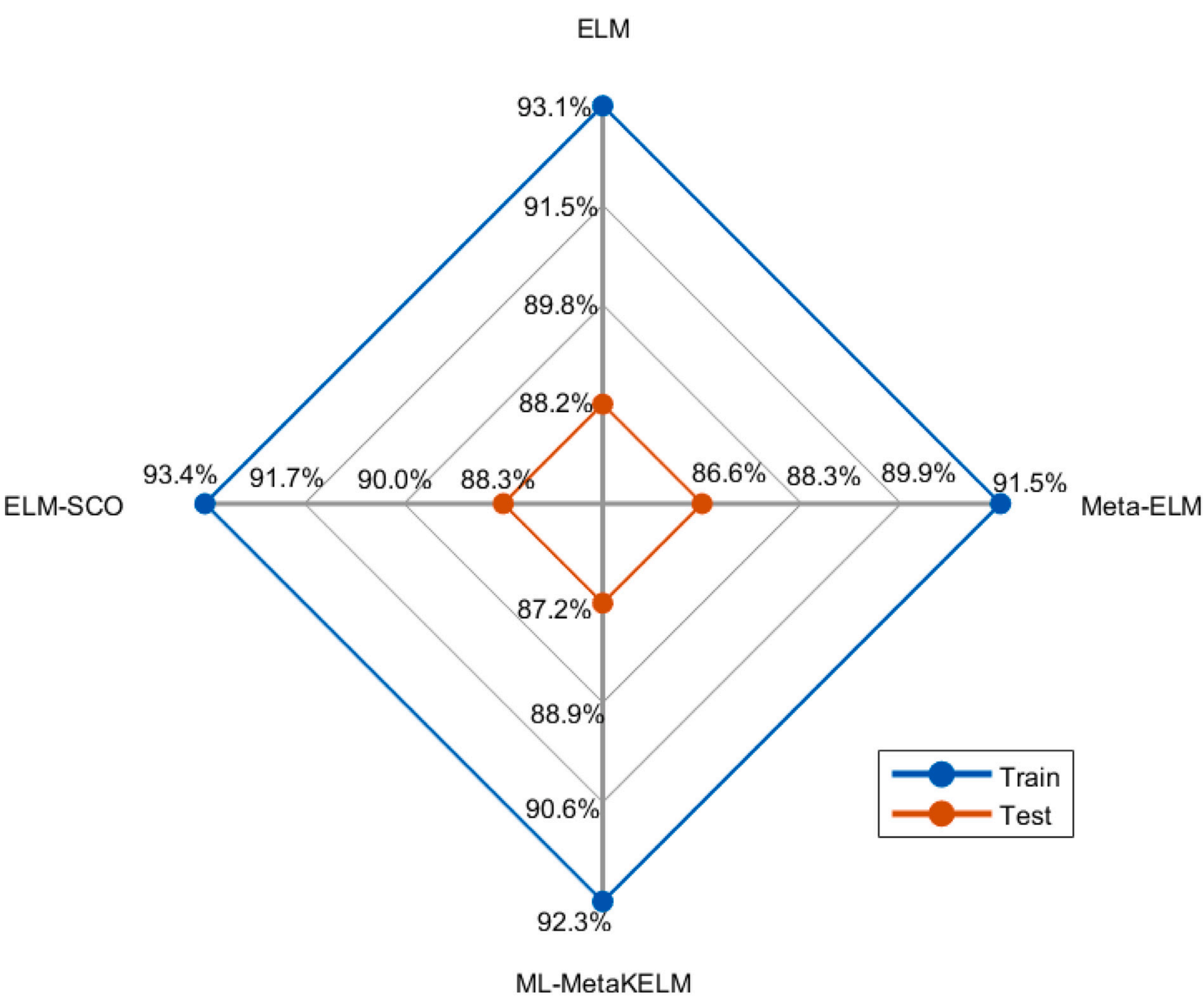

Fig. 7. R^2 results of the implemented models.

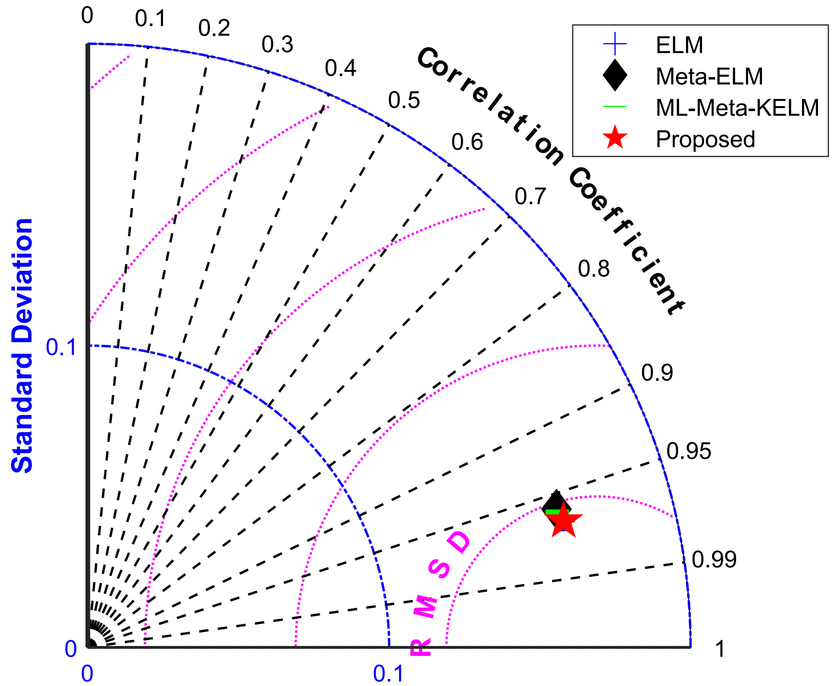

Fig. 8. Taylor diagram for standard deviation, correlation coefficient, and RMSD results of the implemented models.

metrics. During the training phase, the ELM-SCO model achieved the lowest MSE (0.000948), RMSE (0.0308), and MAE (0.0218), along with the highest R^2 value (0.9589). In comparison, other models performed lower performance: ELM achieved an MSE of 0.000958 and R^2 of 0.9584, while ANFIS recorded an MSE of 0.000954 and R^2 of 0.9586. Models such as MLP and LSTM demonstrated significantly higher error rates. For the testing phase, the ELM-SCO model consistently outperformed all alternatives, achieving the lowest MSE (0.0016), RMSE (0.0397), and MAE (0.028), with the highest R^2 value (0.9234). By contrast, ELM and ANFIS recorded MSE values of 0.0024 and 0.0018, and R^2 values of 0.8827 and 0.9148, respectively. Meta-ELM, ML-Meta-KELM, MLP, and LSTM exhibited significantly higher error rates, highlighting their limitations in effectively capturing the underlying patterns in the data. These results further demonstrate the robustness and effectiveness of the ELM-SCO model in accurately forecasting node voltages and capturing voltage variations.

3.3. Discussion and limitations

The results demonstrate that the proposed ELM-SCO model consistently outperforms traditional ELM-based models and other ML approaches. Unlike traditional ELM models that rely on random initialization, the SCO algorithm optimizes the initial weights and biases, significantly improving convergence, accuracy, and robustness. The ELM-SCO model's single-layer architecture ensures faster training times and lower computational costs compared to deep learning models like LSTM, which require extensive hyperparameter tuning and higher computational resources. Additionally, ELM-SCO demonstrates robustness against variability and outliers in the data, a limitation of models like ANFIS and MLP when applied to similar forecasting tasks. Consequently, the ELM-SCO model consistently achieves higher R^2 values across both training and testing phases, indicating its ability to generalize effectively across diverse datasets.

While the proposed approach shows significant advantages, there are limitations that warrant further study and model enhancement. First, the current study evaluates the ELM-SCO model using relatively small-scale datasets. Additional validation on larger, more diverse networks is essential to further assess the model's scalability and generalizability. Second, handling large volumes of smart meter data, particularly identifying uncorrelated or ambiguous information, remains a challenge. Employing advanced methods, such as large language models, could help efficiently manage and classify this data, improving data quality and enabling more reliable forecasting results.

4. Conclusions

This study has proposed a data-driven model for forecasting node voltages within LV distribution networks with LCTs by exploiting active and power measurements from smart meter data. To achieve computational efficiency for potential voltage control in an active grid management framework, we optimized ELM model parameters using the SCO metaheuristic method. The model was trained and tested on smart meter data from two datasets: a realistic LCT-enriched LV network and an additional residential dataset with different load behavior. Its performance was evaluated against several well-established ML models, several ELM-based models, including MLP, LSTM, and ANFIS, to ensure a comprehensive benchmark analysis.

The results demonstrated a significant improvement in computation time, with the proposed SCO-optimized ELM achieving up to a 17-fold reduction compared to the fastest alternative metaheuristic method implemented. While the computation time per node achieved is not yet sufficient for real-time voltage control, the SCO demonstrated notable improvements in the ELM model's convergence to optimal values. Factors such as data size and computational resources also influence computation time. This highlights the importance of selecting appropriate smart datasets to accurately model non-linear relationships between power measurements and node voltages.

In terms of voltage forecasting, the proposed model outperformed the other ML models, achieving the lowest average voltage deviation of 0.56% compared to measured values. When tested on an additional smart meter dataset, the ELM-SCO model consistently demonstrated superior accuracy across all metrics, further validating its robustness and generalizability. This high level of accuracy is particularly beneficial for distribution network operators as they manage the increasing integration of LCTs.

The proposed model's predictive capabilities could serve as a foundation for future smart grid applications, such as assessing hosting capacity and enabling proactive voltage regulation, both of which are crucial for improving network reliability and reducing energy losses. Using smart meter data for voltage forecasting with advanced ML methods like the one proposed offers network operators a powerful tool for managing voltage stability and improving grid efficiency.

Future work will proceed in two key directions: First, employing large language models to efficiently manage and classify smart meter data, especially when handling large datasets and identifying uncorrelated or unclear information. This approach will be crucial for enhancing data quality and improving the model's input features, leading to more reliable forecasting results. Second, further improving the accuracy of the voltage forecasting model by integrating advanced signal processing techniques and optimizing dataset selection.

CRediT authorship contribution statement

Emrah Dokur: Writing – review & editing, Writing – original draft, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Nuh Erdogan:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Ibrahim Sengor:** Writing – review & editing, Writing – original draft, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Ugur Yuzgec:** Writing – review & editing, Writing – original draft, Validation, Methodology, Conceptualization. **Barry P. Hayes:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The work of Barry P. Hayes has been supported by Taighde Éireann (Research Ireland) award numbers 22/FFP-A/10455 and 12/RC/2302 P2.

Data availability

The authors do not have permission to share data.

References

[1] Guerra K, Gutiérrez-Alvarez R, Guerra OJ, Haro P. Opportunities for low-carbon generation and storage technologies to decarbonise the future power system. Appl Energy 2023;336:120828.
[2] Hemetsberger W, Schmela M, Cruz-Capellan T. Global market outlook for solar power 2023–2027. Brussels, Belgium: SolarPower Europe; 2023.
[3] IEA. Global EV outlook 2024, moving towards increased affordability. 2024, https://www.iea.org/reports/global-ev-outlook-2024.
[4] Commission E. Powering a climate-neutral economy: An EU strategy for energy system integration. 2020, URL https://energy.ec.europa.eu/index_en.
[5] Ahmed F, Al Kez D, McLoone S, Best RJ, Cameron C, Foley A. Dynamic grid stability in low carbon power systems with minimum inertia. Renew Energy 2023;210:486–506.
[6] Meunier S, Protopapadaki C, Baetens R, Saelens D. Impact of residential low-carbon technologies on low-voltage grid reinforcements. Appl Energy 2021;297:117057.
[7] Lazo J, Watts D. Stochastic model for active distribution networks planning: An analysis of the combination of active network management schemes. Renew Sustain Energy Rev 2024;191:114156.
[8] Bassi V, Ochoa LF, Alpcan T, Leckie C. Electrical model-free voltage calculations using neural networks and smart meter data. IEEE Trans Smart Grid 2023;14:3271–82.
[9] Hayes B, Gruber J, Prodanovic M. Short-term load forecasting at the local level using smart meter data. In: 2015 IEEE eindhoven powerTech. IEEE; 2015, p. 1–6.
[10] Hayes BP, Prodanovic M. State forecasting and operational planning for distribution network energy management systems. IEEE Trans Smart Grid 2015;7:1002–11.
[11] Pappu SJ, Bhatt N, Pasumarthy R, Rajeswaran A. Identifying topology of low voltage distribution networks based on smart meter data. IEEE Trans Smart Grid 2017;9:5113–22.
[12] Cunha VC, Freitas W, Trindade FC, Santoso S. Automated determination of topology and line parameters in low voltage systems using smart meters measurements. IEEE Trans Smart Grid 2020;11:5028–38.
[13] Wang Y, Liu MZ, Ochoa LF. Assessing the effects of der on voltages using a smart meter-driven three-phase lv feeder model. Electr Power Syst Res 2020;189:106705.
[14] Dokur E, Sengor I, Erdogan N, Yuzgec U, Hayes BP. Smart meter data-driven voltage forecasting model for a real distribution network based on sco-mlp. In: 2023 IEEE PES innovative smart grid technologies europe (ISGT EUROPE). IEEE; 2023, p. 1–5.
[15] Rodríguez-Pajarón P, Bayo AH, Milanović JV. Forecasting voltage harmonic distortion in residential distribution networks using smart meter data. Int J Electr Power Energy Syst 2022;136:107653.
[16] Cao D, Zhao J, Hu W, Ding F, Yu N, Huang Q, Chen Z. Model-free voltage control of active distribution system with pvs using surrogate model-based deep reinforcement learning. Appl Energy 2022;306:117982.
[17] Bassi V, Ochoa L, Alpcan T. Model-free voltage calculations for pv-rich lv networks: Smart meter data and deep neural networks. In: 2021 IEEE madrid powerTech. IEEE; 2021, p. 1–6.
[18] Chen Y, Shi Y, Zhang B. Data-driven optimal voltage regulation using input convex neural networks. Electr Power Syst Res 2020;189:106741.
[19] İnaç T, Dokur E, Yüzgeç U. A multi-strategy random weighted gray wolf optimizer-based multi-layer perceptron model for short-term wind speed forecasting. Neural Comput Appl 2022;34:14627–57.
[20] Shami TM, Grace D, Burr A, Mitchell PD. Single candidate optimizer: a novel optimization algorithm. Evol Intell 2024;17:863–87.
[21] Yuan X, Karbasforoushha MA, Syah RB, Khajehzadeh M, Keawsawasvong S, Nehdi ML. An effective metaheuristic approach for building energy optimization problems. Buildings 2022;13:80.
[22] Su Heng-Yi, Hong Hsu-Hui. An intelligent data-driven learning approach to enhance online probabilistic voltage stability margin prediction. IEEE Trans Power Syst 2021;36:3790–3.
[23] Malbasa Vuk, Zheng Ce, Chen Po-Chen, Popovic Tomo, Kezunovic Mladen. Voltage stability prediction using active machine learning. IEEE Trans Smart Grid 2017;8:3117–24.
[24] Diao Ruisheng, Wang Zhiwei, Di Shi, Chang Qianyun, Duan Jiajun, Zhang Xiaohu. Autonomous voltage control for grid operation using deep reinforcement learning. 2019, URL https://arxiv.org/abs/1904.10597. arXiv:1904.10597.
[25] Adhikari Ananta, Naetiladdanon Sumate, Sagswang Anawach, Gurung Samundra. Comparison of voltage stability assessment using different machine learning algorithms. In: 2020 IEEE 4th conference on energy internet and energy system integration (EI2). 2020, p. 2023–6.
[26] Liu Y, Wang J. Transfer learning based multi-layer extreme learning machine for probabilistic wind power forecasting. Appl Energy 2022;312:118729.
[27] Kamruzzaman J, Sarker RA, Ahmad I. Svm based models for predicting foreign currency exchange rates. In: Third IEEE international conference on data mining. IEEE; 2003, p. 557–60.
[28] Gui J, Sun Z, Cheng J, Ji S, Wu X. How to estimate the regularization parameter for spectral regression discriminant analysis and its kernel version? IEEE Trans Circuits Syst Video Technol 2013;24:211–23.
[29] Amala Manuela A, Gnana Saravanan A. Energy management for smart gds with hybrid ac/dc microgrid and renewable energy: Sco-gbdt approach. IETE J Res 2024;1–10.
[30] Mehigan Laura, Zehir Mustafa Alparslan, Cuenca Juan J, Sengor Ibrahim, Geaney Ciaran, Hayes Barry P. Synergies between low carbon technologies in a large-scale mv/lv distribution system. IEEE Access 2022;10:88655–66.
[31] Ucer Emin, Rahman Shahinur, McDonald Andrew, Kisacikoglu Mithat C. Residential active/reactive power consumption, voltage, and frequency data for a house in Alabama. 2019, URL http://ir.ua.edu/handle/123456789/6346.