Cold Regions Science and Technology 224 (2024) 104249

Contents lists available at ScienceDirect

Cold Regions Science and Technology

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

A metaheuristic-optimization-based neural network for icing prediction on transmission lines

Reda Snaiki [a,*], Abdeslam Jamali [b], Ahmed Rahem [b], Mehdi Shabani [a], Brian L. Barjenbruch [c]

[a] Department of Construction Engineering, École de Technologie Supérieure, University of Quebec, Montreal, QC, Canada
[b] Department of Applied Sciences, Université du Québec à Chicoutimi, Chicoutimi, QC, Canada
[c] NOAA/National Weather Service, Valley, NE, USA

ARTICLE INFO

Keywords:
Ice accretion
Ice-to-liquid ratio
Machine learning
Feature selection
Metaheuristic optimizer

ABSTRACT

Ice accretion on overhead transmission line systems is a leading cause of power outages and can lead to substantial economic losses in northern regions. Therefore, accurately and rapidly predicting ice accretion on power lines is crucial for ensuring the safe operation of the power grid. This study introduces a machine learning method for predicting the ice-to-liquid ratio (*ILR*), an important parameter for assessing ice accretion efficiency. While estimating *ILR* is vital for operational forecasting, many existing ice accretion models do not include this capability. A feedforward neural network (FFNN) trained with stochastic gradient descent and various metaheuristic optimizers - specifically particle swarm optimization, grey wolf optimizer, whale optimizer, and slime mold optimizer - is employed to forecast hourly *ILR*. Environmental data required for training and testing the FFNN model were obtained from the Automated Surface Observing System (ASOS). A global sensitivity analysis using the Sobol index, evaluated via the coefficients of a polynomial chaos expansion, was conducted to identify the most influential input parameters. The results indicate that only four input parameters significantly contribute to the variance in the response: precipitation, temperature, dew point temperature, and wind speed. Furthermore, the FFNN model trained with metaheuristic optimizers outperformed the stochastic gradient descent approach. With the predicted *ILR*, ice accumulation can be easily calculated as the product of *ILR* and the amount of liquid precipitation depth.

1. Introduction

Maintaining grid reliability and safety becomes a significant challenge for utilities operating in regions susceptible to freezing temperatures and winter storms due to ice accretion on transmission lines (Ruszczak and Tomaszewski, 2014; Tomaszewski et al., 2019). This ice buildup can lead to catastrophic events when combined with wind loads, causing power outages, structural damage, and substantial economic losses (Fikke et al., 2008). Furthermore, ice accumulation threatens the integrity of the distribution network, potentially exceeding infrastructure capacity and triggering cascading failures. Understanding the mechanisms of ice load formation is crucial, as it can lead to power line failures through two primary modes: structural overload and galloping. For instance, Ontario and Quebec experienced a major ice storm in 1998 that caused the collapse of several power transmission towers and nearly $1.7 billion in economic losses (Chang et al., 2007; Chang et al., 2012). Similarly, an ice storm in eastern Canada in 2013 resulted in $200 million in insured losses and extended power outages for over 1 million customers (Armenakis and Nirupama, 2014; Sheng et al., 2023). In the US alone, icing events are estimated to inflict an average annual cost of $313 million (Zarnani et al., 2012). With the increasing frequency of extreme weather events due to global warming, severe weather-induced icing incidents have become one of the most significant risks in power grid operations (Jeong et al., 2019; Chen et al., 2020). Therefore, accurate and rapid prediction of icing events and their severity is crucial for assisting impacted communities and electric power companies in preparing for ice events, guiding decision-making processes, designing preventive measures, and planning for recovery (DeGaetano et al., 2008).

Several models have been proposed in the literature to predict ice thickness on transmission lines systems by considering meteorological and environmental parameters (Lozowski and Makkonen, 2005). Generally, these models can be categorized as physics-based or data-driven (He et al., 2021). Physics-based models simulate ice accretion

* Corresponding author.
E-mail address: reda.snaiki@etsmtl.ca (R. Snaiki).

https://doi.org/10.1016/j.coldregions.2024.104249
Received 27 March 2024; Received in revised form 23 May 2024; Accepted 7 June 2024
Available online 8 June 2024

by integrating established physical processes such as heat transmission principles and fluid mechanics with simplified mathematical formulations which might be calibrated or validated using experimental data (Imai, 1953; Lenhard Jr., 1955; Goodwin et al., 1983; Makkonen, 1984; Makkonen, 1998; Zhang et al., 2012). However, since ice accretion is a complex, nonlinear process influenced by several environmental factors with numerous uncertainties, generalizing mathematical equations to accurately simulate this process is extremely challenging (He et al., 2021). On the other hand, data-driven models utilize data from different sources such as field measurements and controlled experiments (Saviz Naeini, S. and Snaiki, 2024). These models can then be derived using various machine learning techniques, ranging from relatively simple statistical methods like linear regression (Chaîné and Castonguay, 1974; Jones, 1998; Sanders and Barjenbruch, 2016) to complex algorithms like deep neural networks (Wu and Snaiki, 2022). The simplicity and computational efficiency of simplified autoregressive formulas have led to their widespread implementation in various engineering applications, such as ice maps generation (Jeong et al., 2018; Jarrett et al., 2019; Jeong et al., 2019; Sheng et al., 2023). However, the use of predefined basis functions and autoregressive formulas may not be suitable for accurately modeling such complex dynamic systems. Alternatively, machine learning approaches can be considered as effective tools for rapidly predicting ice accretion (Shabani et al., 2022; Snaiki, 2024). They can simulate nonlinear dynamic systems without relying on simplified assumptions or linearization. For example, artificial neural networks trained using the backpropagation algorithm have been developed for simulating ice accretion (Li et al., 2011; Huang and Li, 2012; Chen et al., 2012). Other techniques, such as genetic algorithms (Du et al., 2010), the continuous ant colony algorithm (Yin et al., 2012), and the fruit fly optimization algorithm (Niu et al., 2017) have also been used to enhance neural network performance in ice accretion prediction. Various versions of Support Vector Machine (SVM) models have also been employed to predict ice thickness based on available environmental factors (Zarnani et al., 2012; Dai et al., 2013; Huang et al., 2014; Ying and Su, 2014; Xu et al., 2015; Ma and Niu, 2016; Ma et al., 2016; Niu et al., 2017). Additionally, models like the extreme learning machine (Sun and Wang, 2019) and deep learning models (He et al., 2021) have been developed for ice accretion prediction. While several machine learning (ML) models have been developed for predicting ice accretion, a thorough investigation into feature selection to identify the best environmental parameters for the ML model among multiple variables has been lacking. Failure to select the most influential factors could result in inaccurate simulation results. Additionally, these models have often been developed based on the same icing processes or the same transmission lines, without considering varying environmental conditions or utilizing large, geographically diverse datasets.

Most ML applications have been developed to predict the equivalent radial ice thickness, which is notoriously challenging to measure directly. In addition, in operational weather forecasting, accurately predicting ice accretion thickness remains a challenge due to the inherent variability in icing efficiency (e.g., runoff rate). As a result, various assumptions (e.g., assumed ice accretion efficiency) are often employed to approximate it, introducing significant uncertainties into the prediction process. In contrast, the ice-to-liquid ratio (*ILR*), which represents the ratio of accumulated ice depth to accumulated liquid depth on a specific surface, is a crucial parameter used to assess the efficiency of ice accretion. By directly targeting *ILR*, which depends on readily predictable parameters including temperature, wind, and precipitation rate, it is possible to simplify the complex process of assessing icing efficiency, and achieve substantially more accurate ice accretion forecasts for upcoming weather events. However, despite the importance of *ILR* estimation, particularly for operational forecasting, most existing ice accretion models do not provide this information. Sanders and Barjenbruch (2016) investigated the effects of several environmental factors on *ILR*s using a large ASOS database. They developed the Freezing Rain Accumulation Model (FRAM), which predicts hourly *ILR*s based on three environmental factors: wind speed, precipitation rate, and wet-bulb temperature. With the predicted *ILR*, ice accumulation can be easily calculated by multiplying the *ILR* by the amount of liquid precipitation accumulation. The resulting ice depth can then be used to estimate ice accretion thickness on elevated objects, whether on a flat surface or as a radial ice thickness measurement, using simplified formulas (Ryerson and Ramsay, 2007). However, it is important to note that FRAM is a simplified autoregressive model and may not capture the nonlinearities within the system accurately.

This study addresses the safety of power transmission/distribution infrastructure by focusing on improved ice accretion prediction on overhead transmission lines. Specifically, the hourly *ILR* (calculated using the hourly accumulated ice and liquid depths) will be predicted using a feedforward neural network (FFNN) trained with several metaheuristic optimizers. This approach aims to overcome the limitations of gradient-based algorithms, which are known for their relatively slow convergence rates and susceptibility to getting stuck in local minima. Metaheuristics utilize heuristics as informed guesses to guide the search process towards promising regions in the solution space. To accomplish this, the environmental data of freezing rain events over the United States was extracted from the Automated Surface Observing System (ASOS). Six ASOS-observed environmental parameters were selected, representing the mean values calculated over 60-min duration of continuous freezing rain precipitation. These parameters are air temperature, dew point temperature, sustained wind speed, wind gust speed, precipitation rate, and wet-bulb temperature. A global sensitivity analysis was then conducted to evaluate the impact of each input parameter on the *ILR*. This analysis utilized Sobol variance decomposition, which was evaluated analytically from the coefficients of a polynomial chaos expansion (PCE) metamodel and compared with the Monte Carlo technique. Additionally, the obtained Sobol's indices were compared with the total Kucherenko indices to further assess their significance. After identifying the most influential input parameters in terms of their variance contribution to the *ILR*, they will be used as inputs to the FFNN model. This model will be trained using four metaheuristic optimization algorithms: particle swarm optimization (PSO), grey wolf optimizer (GWO), whale optimizer (WOA), and slime mold optimizer (SMO), as well as the stochastic gradient descent approach. The simulation results will then be compared and discussed to evaluate the performance of each optimization method in predicting the *ILR*. It should be noted that accurate *ILR* predictions offer significant benefits to grid operators by enabling proactive management strategies. These strategies include pre-emptive de-icing of power lines or load reduction during periods of anticipated heavy ice accumulation, thereby mitigating the risk of line failures. Furthermore, real-time or near real-time *ILR* predictions enhance situational awareness, allowing grid operators to better anticipate and respond to weather events that could lead to outages. Finally, improved ice load forecasts inform targeted maintenance schedules, focusing resources on areas most susceptible to ice-related damage. This combination of proactive measures, enhanced awareness, and optimized maintenance ultimately contributes to a more resilient power grid infrastructure.

2. Feedforward neural network

Feedforward neural networks (FFNNs) are supervised neural networks used in a wide range of applications, including regression and classification. These algorithms excel at learning hidden patterns within data and providing dependable predictions. Two well-known architectures of FFNNs are the multi-layer perceptron and deep neural networks. Both consist of an input layer, one or more hidden layers, and an output layer. Each neuron within these layers performs a specific calculation to transform the data it receives. The output of each neuron can be calculated as follows:

$$y = \phi\left(\sum_{i=1}^{n} \omega_i x_i + b\right) \quad (1)$$

where ϕ = activation function; ω_i = weight associated with input x_i; and b = bias. The activation function is responsible for capturing the non-linearities present in the data and can be selected from a range of commonly used functions, such as hyperbolic tangent, sigmoid, ReLU, and others. To train FFNNs, suitable values for the weights and biases need to be determined to establish a desirable relationship between the predicted and expected outputs. This is typically achieved using an optimization algorithm. For example, the backpropagation algorithm is commonly employed to update the weights and biases of FFNNs. It does so by propagating the error (the difference between predicted and expected outputs) backward through the network and adjusting the weights and biases based on the gradient of the error. More specifically, with a predefined loss function E, the new weight or bias (θ_{new}) can be obtained based on the following formula:

$$\theta_{new} = \theta_{old} - \alpha \frac{\partial E}{\partial \theta} \quad (2)$$

where θ_{old} = current weight or bias; and α = learning rate. While the backpropagation technique has been widely used to train neural networks in various applications, it does have certain drawbacks. Specifically, this gradient-based algorithm is known for its relatively slow convergence rate and its susceptibility to getting stuck in local minima. Additionally, the backpropagation technique can suffer from the vanishing and exploding gradient problems, which can significantly affect the training of the neural network. To address these issues, advanced optimization techniques, such as those based on metaheuristic algorithms, have been developed. These techniques aim to overcome the limitations of traditional backpropagation by providing more efficient and robust training methods for neural networks. A schematic representation of a FFNN model trained with a metaheuristic algorithmic is depicted in Fig. 1.

3. Metaheuristic algorithms

A metaheuristic is a procedure that uses nature-inspired heuristics to effectively balance exploration and exploitation strategies, aiming to find an optimal solution. Unlike conventional methods that rely on continuous and differentiable objective functions (such as gradient-based techniques), metaheuristic algorithms have the capability to handle highly nonlinear and non-differentiable problems. Metaheuristic algorithms can be broadly classified into three categories (Ojha et al., 2017): 1. Single solution-based methods which focus on improving a single solution iteratively (Kirkpatrick et al., 1983; Mladenović and Hansen, 1997); 2. Population-based methods which manage a group of solutions, inspired by natural phenomena like evolution (genetic algorithms) or swarm behavior (particle swarm optimization) (Geem et al., 2001; Rashedi et al., 2009); 3. Hybrid methods which combine multiple metaheuristic models for enhanced optimization.

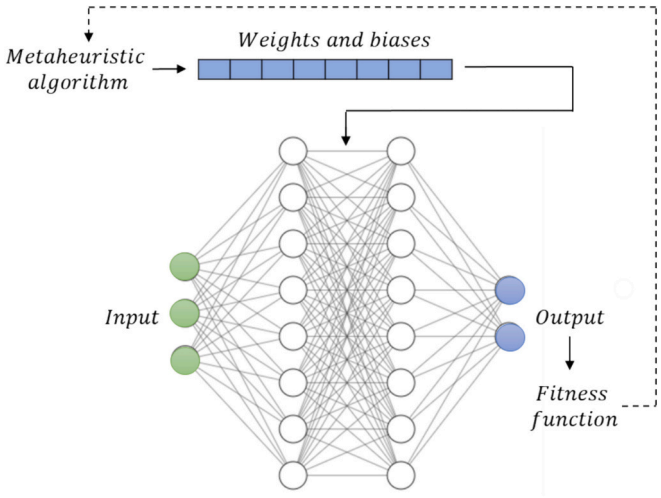

Fig. 1. Schematic of a FFNN model trained with a metaheuristic algorithm.

In this study, four state-of-the-art population-based algorithms will be employed and compared to optimize the weights and biases of FFNNs, namely the particle swarm optimization (PSO), the grey wolf optimizer (GWO), the whale optimizer (WOA) and the slime mold optimizer (SMA). A brief description of each algorithm will be provided in the following sections.

3.1. Particle swarm optimization

Particle Swarm Optimization (PSO) is a population-based optimization algorithm that emulates the behavior of bird flocks or fish schools. It is suitable for both single-objective and multi-objective optimization problems (Kennedy and Eberhart, 1995). In PSO, a population of potential candidate solutions, referred to as particles, explores a search space to find optimal or near-optimal solutions by iteratively updating their positions. Each particle updates its position by considering its own best-known position as well as the best-known position of the entire swarm. The updated velocity ($\vec{v}^{j+1}$) and position ($\vec{X}^{j+1}$) for each particle are derived based on the current velocity ($\vec{v}^{j}$) and position ($\vec{X}^{j}$) (indicated by j) as:

$$\vec{v}^{j+1} = \vec{v}^{j} + C_1 r_1 . \left[\vec{X}^{j}_{lopt} - \vec{X}^{j}\right] + C_2 r_2 . \left[\vec{X}^{j}_{gopt} - \vec{X}^{j}\right] \quad (3)$$

$$\vec{X}^{j+1} = \vec{X}^{j} + \vec{v}^{j+1} \quad (4)$$

where $\vec{X}$ = position vector; $\vec{v}$ = velocity vector; $\vec{X}_{lopt}$ = personal best position; $\vec{X}_{gopt}$ = best global position; C_1 = cognitive component; C_2 = social component; and (r_1, r_2) = random values sampled from an interval $[0, 1]$. It should be noted that for the PSO algorithm to effectively identify a global optimal solution, the hyperparameters, including the cognitive and social components, along with the inertia weight, should be carefully selected based on the specific optimization problem. Additionally, the termination criterion is typically linked to a maximum number of iterations or a desired level of solution quality. This careful selection of hyperparameters and termination criteria is crucial for the algorithm to converge efficiently and produce high-quality solutions.

3.2. Grey wolf optimizer

The Grey Wolf Optimizer (GWO) is a population-based optimization algorithm inspired by the social leadership and hunting behavior of grey wolves (Mirjalili et al., 2014). This optimization algorithm divides the population into four main groups known as alpha (α), beta (β), delta (δ), and omega (ω) wolves. The alpha, beta, and delta wolves, representing the three fittest individuals, assume leadership roles and guide the other wolves towards promising regions in the search space. During the optimization process, the wolves update their current positions (indicated by j) around the alpha, beta, or delta wolves using the following equations:

$$\vec{X}^{j+1} = \vec{X_p}^{j} - \vec{A} . \vec{D} \quad (5)$$

$$\vec{D} = \left| \vec{C} . \vec{X_p}^{j} - \vec{X}^{j} \right| \quad (6)$$

where $\vec{X_p}$ = position of the prey; $\vec{X}$ = position of a grey wolf; and ($\vec{A}$, $\vec{C}$) = coefficient vectors which help to explore different regions around

the best agent in the search space during the optimization process.

Throughout the optimization process, it is assumed that the first three best solutions are represented by α, β and δ, respectively. The remaining wolves (ω) reposition themselves based on α, β and δ. Specifically, the following quantities which determine the distances between the current position and α, β and δ, respectively, are first determined:

$$\begin{aligned}\overrightarrow{D_\alpha} &= \left|\overrightarrow{C_1}.\overrightarrow{X_\alpha} - \overrightarrow{X}\right| \\ \overrightarrow{D_\beta} &= \left|\overrightarrow{C_2}.\overrightarrow{X_\beta} - \overrightarrow{X}\right| \\ \overrightarrow{D_\delta} &= \left|\overrightarrow{C_3}.\overrightarrow{X_\delta} - \overrightarrow{X}\right|\end{aligned} \tag{7}$$

where $\overrightarrow{X_k}$ = position of the k wolf (with $k = \alpha, \beta$ or δ); and ($\overrightarrow{C_1}$, $\overrightarrow{C_2}$, $\overrightarrow{C_3}$) = random vectors. Then the updated position $\overrightarrow{X}^{j+1}$ can be calculated as:

$$\overrightarrow{X}^{j+1} = \left[\left(\overrightarrow{X_\alpha} - \overrightarrow{A_1}.\overrightarrow{D_\alpha}\right) + \left(\overrightarrow{X_\beta} - \overrightarrow{A_2}.\overrightarrow{D_\beta}\right) + \left(\overrightarrow{X_\delta} - \overrightarrow{A_3}.\overrightarrow{D_\delta}\right)\right]/3 \tag{8}$$

where ($\overrightarrow{A_1}$, $\overrightarrow{A_2}$, $\overrightarrow{A_3}$) = random vectors. It should be noted that the boundary constraints are applied to the new positions to ensure that they remain within the defined problem boundaries. This is important to prevent solutions from going beyond the feasible region of the problem space.

3.3. Whale optimizer

The Whale Optimization Algorithm (WOA) is a population-based optimization algorithm that utilizes a population of search agents to iteratively enhance a set of random candidate solutions by emulating the bubble-net feeding behavior of humpback whales (Mirjalili and Lewis, 2016). Its objective is to identify the global optimum for the optimization problem. The search process is divided into two phases: exploration and exploitation. During the exploitation phase, represented by a random coefficient vector $\overrightarrow{A}$ where $\left|\overrightarrow{A}\right| \le 1$, the next position $\overrightarrow{X}^{j+1}$ is determined as follows:

$$\overrightarrow{X}^{j+1} = \begin{cases} \overrightarrow{X^*}^j - \overrightarrow{A}.\overrightarrow{D} \text{ if } p < 0.5 \\ \overrightarrow{D'}.exp(bl).cos(2\pi l) + \overrightarrow{X^*}^j \text{ if } p \ge 0.5 \end{cases} \tag{9}$$

where p = a random number in the range $[0, 1]$; $\overrightarrow{X^*}$ = position of the best solution; $\overrightarrow{D} = \left|\overrightarrow{C}.\overrightarrow{X^*}^j - \overrightarrow{X}^j\right|$; $\overrightarrow{D'} = \left|\overrightarrow{X^*}^j - \overrightarrow{X}^j\right|$; $\overrightarrow{C}$ = coefficient vector; b = a constant that determines the shape of spiral bubble-net; and l = a random number in the range $[-1, 1]$. On the other hand, the exploration phase (represented by a random coefficient vector $\left|\overrightarrow{A}\right| > 1$) can be formulated as follows:

$$\overrightarrow{X}^{j+1} = \overrightarrow{X}_{rand} - \overrightarrow{A}.\overrightarrow{D} \tag{10}$$

where $\overrightarrow{X}_{rand}$ = random position vector selected from the current whale optimization; and $\overrightarrow{D} = \left|\overrightarrow{C}.\overrightarrow{X}^j_{rand} - \overrightarrow{X}^j\right|$.

3.4. Slime mold optimizer

The Slime mold optimizer (SMO) is a population-based optimization algorithm that emulates the foraging behavior and morphological transformations observed in slime mold *Physarum Polycephalum* to explore the search space and find the optimal or near-optimal solutions (Li et al., 2020). This optimizer utilizes weight values to mimic the positive and negative feedback of the bio-oscillator during the foraging stage, resulting in the formation of a diverse feeding vein network thickness, while the slime mold's morphology undergoes changes through three distinct contraction patterns. The algorithm operates in two primary phases: the approach phase, where it searches for food, and the wrap phase, where it exploits the identified food. It iterates between these two phases in an oscillating manner. The mathematical formulation which mimics the slime mold behavior is given as:

$$\overrightarrow{X}^{j+1} = \begin{cases} rnd.(ub - lb) + lb \text{ if } rnd < z \\ \overrightarrow{X}^j_b + \overrightarrow{v_b}.\left(\overrightarrow{W}.\overrightarrow{X}^j_A - \overrightarrow{X}^j_B\right) \text{ if } r < p \\ \overrightarrow{v_c}.\overrightarrow{X}^j \text{ if } r \ge p \end{cases} \tag{11}$$

where ($\overrightarrow{X}_A$, $\overrightarrow{X}_B$) = two randomly selected individuals from the population; $\overrightarrow{X}_b$ = location of the individual with the optimal fitness; $\overrightarrow{W}$ = weight of the slime mold; $\overrightarrow{v_b}$ = a parameter within the range of $[-a, a]$; $\overrightarrow{v_c}$ = a parameter which decreases linearly from one to zero; lb and ub are the lower and upper boundaries of the search range; p = a selected value controlling the exploration and exploitation processes; (rnd, r) = random values in the range of $[0, 1]$; and z = a hyperparameter that controls the balance between exploration and exploitation. It's worth noting that the first term of the equation signifies the random exploration process, where the hyperparameter z can be selected based on the specific problem being studied. The second and third terms correspond to the exploration and exploitation phases, respectively.

3.5. Metaheuristic design for training FFNNs

Metaheuristic algorithms are suitable for various applications, including those related to Feedforward Neural Networks (FFNNs), due to their capability to provide near-optimal solutions for complex and non-differentiable problems. These algorithms can be utilized to optimize different aspects of FFNNs, such as weights and biases, architecture, activation functions, and other hyperparameters like learning rate. However, they have been primarily employed to optimize weights and biases for fixed architectures, aiming to minimize the overall error of FFNNs.

In this study, the four selected metaheuristic models, namely PSO, GWO, WOA, and SMO, are employed to train a FFNN with a single hidden layer. To achieve this, the weights and biases of a candidate neural network are first stored in a one-dimensional vector, with the vector's length corresponding to the total number of weights and biases in the FFNN. Next, an objective function (or fitness function) is defined for the metaheuristic algorithm, aiming to optimize the FFNN's ability to achieve the highest accuracy in either regression or classification applications. In this study, the mean square error (MSE) metric is chosen as the fitness function for the four metaheuristic algorithms and can be expressed as:

$$MSE = \frac{1}{n}\sum_{i-1}^{n}(y - \hat{y})^2 \tag{12}$$

where y = actual (true) value; $\hat{y}$ = predicted value generated by the FFNN model; and n = number of training samples. It should be noted that the correlation coefficient R will also be employed to assess the performance of the proposed model. The proposed procedure is illustrated in Fig. 2.

4. Application

In this section, four selected metaheuristic algorithms, namely PSO, GWO, WOA, and SMO, will be employed to train a feedforward neural network (FFNN). The objective is to predict the ice-to-liquid ratio, which in turn determines the thickness of ice accretion on overhead transmission lines. This section will start will the ice accretion problem and

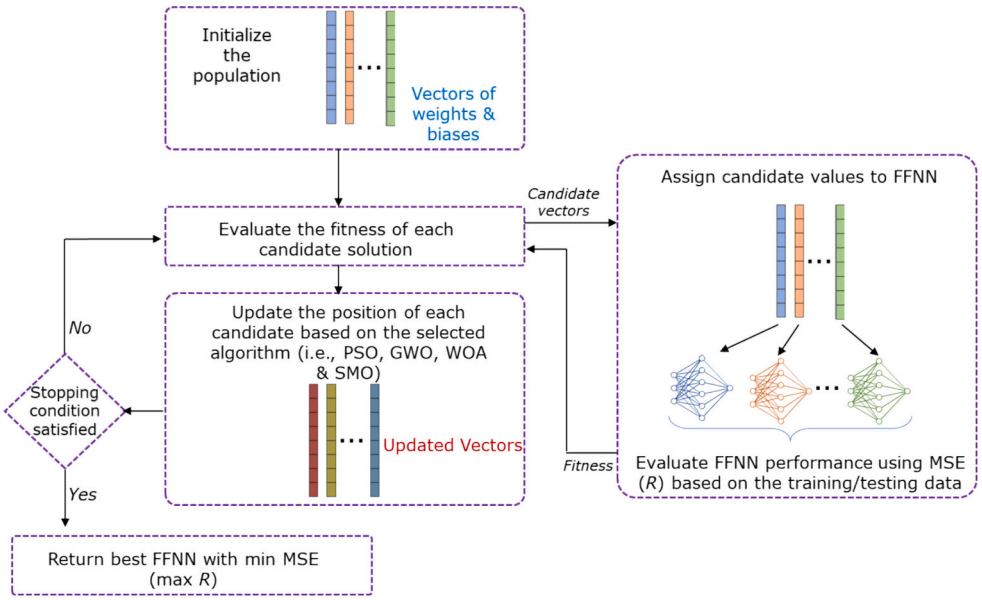

Fig. 2. General steps of the metaheuristic-based FFNN model.

datasets, followed by a discussion of feature selection. Finally, the training and testing results will be presented, along with an analysis and discussion of the findings.

4.1. Ice accretion problem & datasets

Ice accretion on overhead lines refers to the gradual buildup of ice on conductors, insulators, or other components of power transmission systems due to freezing precipitation. This accumulation can lead to failure due to the weight of the ice accretion, especially when combined with additional physical stress applied when the lines are exposed to wind, potentially resulting in power outages and even the collapse of the power transmission system. Therefore, accurate and rapid prediction of ice accretion on transmission lines is vital for ensuring system reliability, public safety, preventing power disruptions, and planning resource allocation for improved recovery time. In this study, a FFNN is proposed to rapidly predict the ice-to-liquid ratio (*ILR*) which can then be used to predict ice accretion on elevated surfaces including power lines.

The data used to train the neural networks were obtained from the Automated Surface Observing System (ASOS). ASOS is an advanced network of meteorological instruments and sensors operated by various agencies, including the National Weather Service. It collects real-time weather data such as wind speed and direction, temperature, atmospheric pressure, and precipitation. By analyzing ASOS data during freezing precipitation events, it is possible to accurately measure ice accretion on elevated surfaces. These ice accretion measurements are observed at the same time and location as the additional atmospheric variables, making it possible to correlate atmospheric conditions to *ILR*. For this study, ice accumulation data was acquired from the National Centers for Environmental Information (NCEI) for the period 2013–2017, encompassing 407 ASOS sites and providing geographically diverse coverage (Sanders and Barjenbruch, 2016). The data consisted of 60-min periods with continuous freezing rain, resulting in 2646 observations. To isolate impactful freezing rain events, further filtering criteria were applied. Events were excluded if they lacked any 60-min period with precipitation rates exceeding 0.50 mm/h (0.02 in/h) or ice accumulation rates below 0.25 mm/h (0.01 in/h), as these events were unlikely to cause significant infrastructure issues. Finally, a meticulous manual quality control process ensured data integrity by eliminating periods with missing or unreasonable meteorological data required for calculating mean sustained wind speed or wet-bulb temperature. This rigorous filtering resulted in a final dataset of 1355 high-quality observations. Six environmental parameters were evaluated as potential predictors of the *ILR*, with the mean of each parameter calculated over the 60-min duration of continuous freezing rain precipitation. These parameters are the air temperature ($T[^{\circ}C]$), dewpoint temperature ($DPT[^{\circ}C]$), sustained wind speed (U[kts]), wind gust speed (G[kts]), hourly liquid precipitation rate (P[in/h]), and wet-bulb temperature ($WBT[^{\circ}C]$). To ensure data quality, rigorous quality control procedures were implemented to remove any invalid or physically implausible data points.

It is important to highlight that the proposed ML model aims to predict *ILR*, a key parameter for operational forecasting. Directly targeting *ILR*, which depends on readily predictable parameters (e.g., temperature, wind, and precipitation rate), can help overcome the challenges associated with varying icing efficiency. This approach offers the potential for significantly more accurate ice accretion forecasts for upcoming weather events compared to directly predicting ice accretion thickness, which is challenging due to its inherent variability in icing efficiency. The *ILR* represents the ratio between the accumulated depth of ice on a specific surface and the accumulated depth of liquid, which, by definition, occurs on a horizontal surface. Given the *ILR*, the ice thickness on an elevated horizontal surface I_T can be obtained using the following formula (Sanders and Barjenbruch, 2016):

$$I_T = \sum_{0}^{h} ILR \times P \tag{13}$$

where h = number of hours over which the precipitation occurs. The equivalent radial ice thickness on an elevated surface (R_{eq}) can be then obtained using the following empirical formula (Ryerson and Ramsay, 2007):

$$R_{eq} = 0.394 I_T \tag{14}$$

4.2. Input selection

In order to select the best input variables for the FFNN model, this study employs the Sobol index. The Sobol index assesses the portion of the output variance that can be attributed to each input variable or combination of variables. The first-order Sobol's index quantifies the individual effect of each input variable on the output variability by measuring the proportion of the total variance that can be attributed to a

specific input variable independently of other variables. On the other hand, higher-order Sobol's indices quantify the combined effects of two or more variables on the output, measuring the extent to which interactions contribute to the total output variance. Similarly, the total Sobol index quantifies the proportion of the total output variance that can be attributed to a particular variable, considering all possible interactions with the other variables.

To calculate the Sobol's indices, the study employs an analytical approach based on the coefficients of a polynomial chaos expansion (PCE) metamodel, which is constructed using ASOS data. This involves characterizing random inputs and constructing the PCE to enable the computation of Sobol's indices. Each uncertain model parameter is represented by a random variable and a corresponding probability density function (PDF) in the form $X \sim f_X(x)$. The selection of appropriate probability distributions for the environmental parameters involved a two-step process. First, visual inspection of data histograms (Fig. 3) and knowledge of the physical characteristics of each parameter guided the identification of candidate distributions. The Kolmogorov-Smirnov (K—S) test then served as a final step to assess the goodness-of-fit between the candidate distributions and the empirical distributions observed in the dataset. This approach ensured that the chosen distributions closely resembled the actual data, enhancing the model's accuracy.

The histograms in Fig. 3 represent the marginal distributions for each variable, while the plots in the lower/upper diagonals depict the sampled points from the bi-dimensional histograms. The distributions found for the variables are as follows: a logistic distribution for *T*, Gumbel Min for *DPT*, Gumbel for *U*, Gumbel for *G*, lognormal for *P*, and Gumbel Min for *WBT*. It is important to note that the Gumbel distribution refers to the maximum Gumbel distribution or the extreme value distribution of type I, while Gumbel Min corresponds to the Gumbel minimum extreme value distribution, also known as the Smallest Extreme Value (Type I) distribution. For the obtained marginal distributions, a custom set of polynomials that are orthonormal to the non-standard distribution is used in this study. Specifically, the univariate orthonormal polynomials are computed numerically using the *Stieltjes* procedure (Stieltjes, 1884). Given these univariate orthonormal polynomials, a total-degree truncation scheme is defined, which includes all polynomials in the given input variables of a total degree less than or equal to the specified maximum polynomial degree p (in this case, $p = 12$). To compute the PCE coefficients, a sparse PCE method called Orthogonal Matching Pursuit (OMP) (Pati et al., 1993; Mallat and Zhang, 1993) is employed. The global sensitivity analysis is conducted using the PCE-based technique and compared to the Monte-Carlo-based technique, as shown in Fig. 4.

The results indicate that only four stochastic dimensions significantly contribute to the response's variance. Precipitation (*P*), mean temperature (*T*), dew point temperature (*DPT*), and mean wind speed (*U*) are the most dominant environmental factors, with precipitation having the most significant effect on ice accretion. As a result, gust wind (*G*) and wet-bulb temperature (*WBT*) will be excluded from subsequent simulations, and only the four dominant environmental factors will be used as inputs for the FFNN model. Furthermore, the results from the Monte Carlo technique closely align with those from the PCE-based technique. To validate the findings, the total Kucherenko indices are also depicted in Fig. 4, providing additional insight into dependent input variables and supporting similar conclusions. It is worth noting that while wet bulb temperature is generally considered an important factor in predicting precipitation type, the data used in this study specifically focused on freezing rain events and did not show a significant relationship with

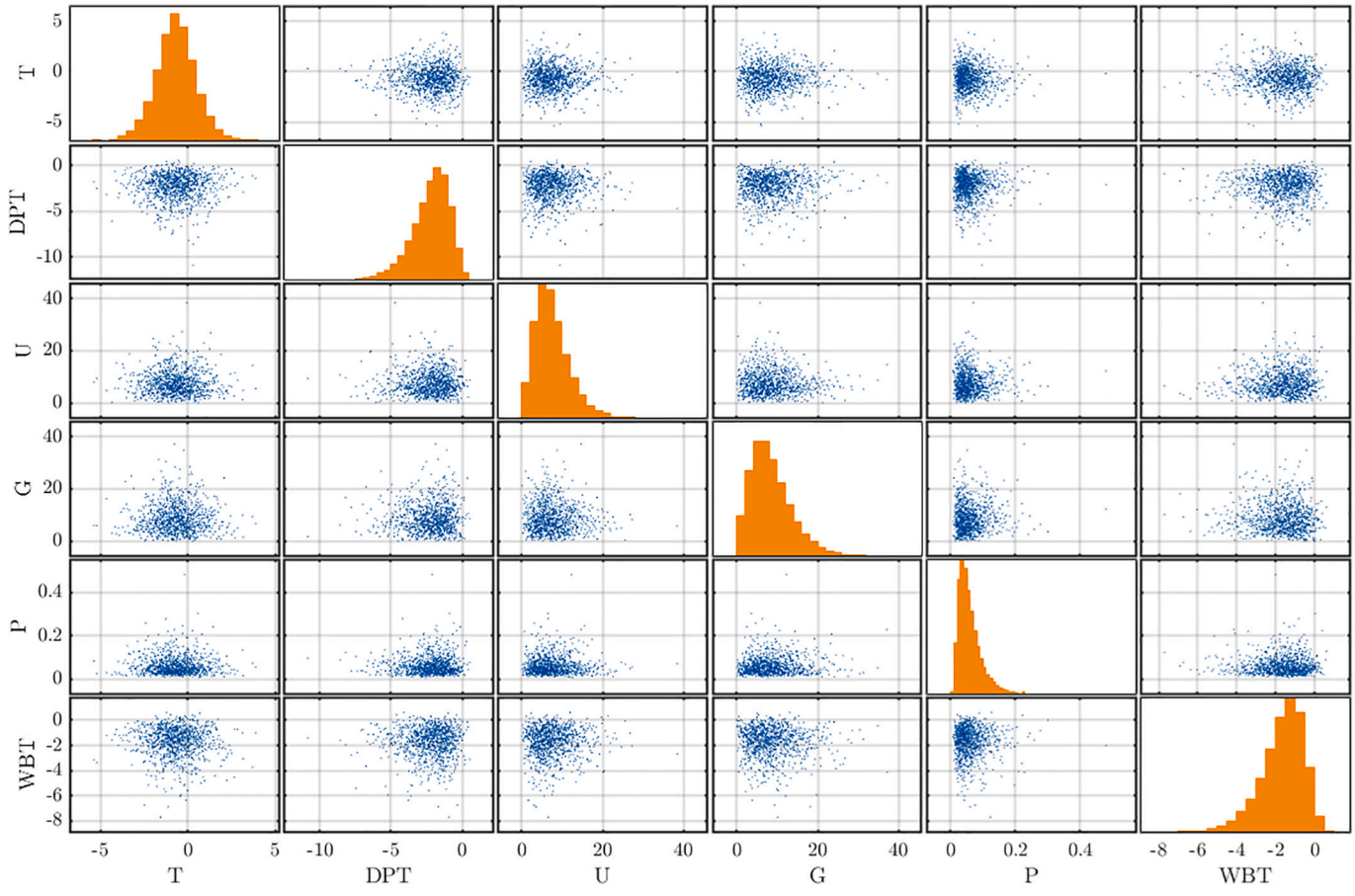

Fig. 3. Drawing samples from the obtained marginal PDF corresponding to the environmental parameters.

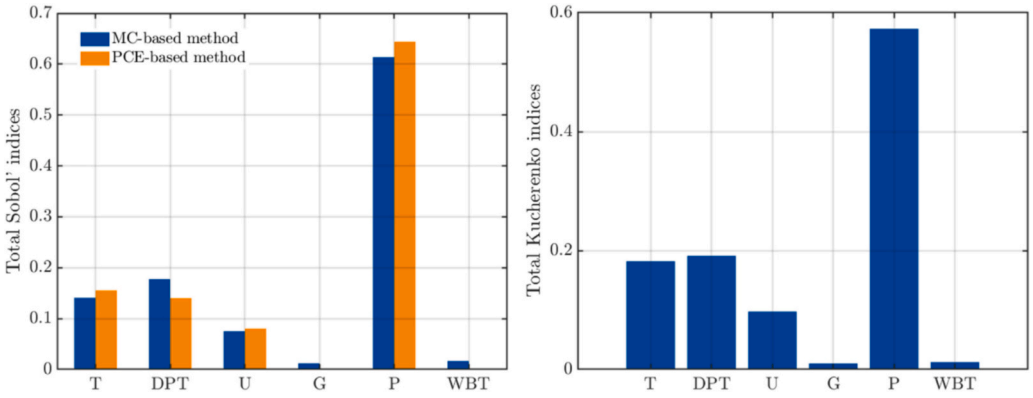

Fig. 4. Total Sobol indices evaluated using the MC and PCE approaches (left) and total Kucherenko indices (right).

this predictor. While not a primary objective of this study, an inclusion of non-freezing rain events (where the *ILR* is 0) could reveal a more prominent role for wet bulb temperature as a predictor.

4.3. Results and discussion

The architecture of the selected FFNN model includes: 1. An input layer comprising the mean air temperature (*T*), mean dewpoint temperature (*DPT*), mean sustained wind speed (*U*), and precipitation rate (*P*); 2. One hidden layer containing 21 neurons, determined through trial-and-error; and 3. An output layer with a single variable representing the ice-to-liquid ratio (*ILR*) as shown in Fig. 5. The selected activation function for the hidden layer is ReLU, while a linear activation function is chosen for the output layer. Before training, all input datasets are normalized using the min-max normalization technique. Seventy percent of the data is randomly assigned to the training set, with the remaining 30 % reserved for testing.

The FFNN model was trained using the four selected metaheuristic models—PSO, GWO, WOA, and SMO—following the approach outlined in Sect. 3. Additionally, ten different runs were executed for each metaheuristic algorithm to statistically characterize the results. A population size of 70 was chosen for all four metaheuristic models, determined through a trial-and-error approach. The convergence curves for all metaheuristic optimizers, depicting MSE over 200 iterations for a single independent run, are shown in Fig. 6. The figure illustrates that all selected metaheuristic optimizers exhibit good optimization efficiency and fast convergence, typically requiring fewer than 50 iterations. Furthermore, GWO demonstrates the fastest convergence speed for the ice accretion dataset.

To evaluate the performance of the metaheuristic optimizers, they were compared against the standard stochastic gradient descent (SGD) algorithm. Table 1 presents the average, median, standard deviation (STD), and best MSE across all training samples from ten independent runs.

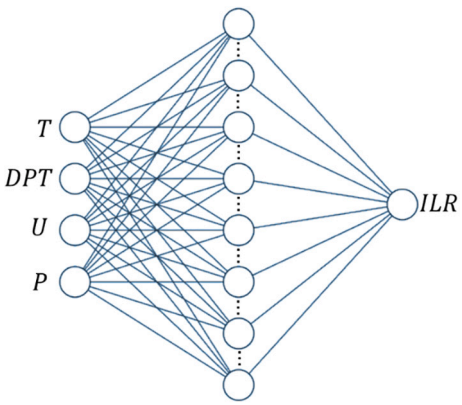

Fig. 5. FFNN architecture.

From Table 1, it is evident that FFNN models trained using metaheuristic algorithms outperform the model trained with stochastic gradient descent in terms of average, median, and standard deviation of MSE. This indicates their superior ability to avoid local minima in this context. Among the metaheuristic optimizers, PSO, WOA, and GWO stand out with average MSE values of 2.19E-02, 2.21E-02 and 2.24E-02, respectively. Additionally, all four metaheuristic models exhibit low standard deviation, suggesting robustness and stability compared to SGD. Fig. 7 displays Boxplots for 10 runs of SGD, PSO, GWO, WOA, and SMO. The Boxplots provide a clear view of the optimizer variability in terms of MSE values across all runs. These plots further confirm the superior performance of metaheuristic optimizers over the classical SGD algorithm. Notably, WOA and PSO demonstrate the best performance, followed by GWO, as evidenced by their low average MSE values and compact boxes, indicating the stability of the proposed training algorithms.

Table 2 summarizes the MSE and correlation coefficient (*R*) achieved by the different models on the testing set. PSO, GWO, and WOA displayed superior performance, consistent with the training set observations. Notably, these algorithms achieved comparable MSE values (around 2.29E-02) and high correlation coefficients (above 0.80).

Based on this comparative study, it can be concluded that the metaheuristic optimizers outperformed the standard gradient descent approach. Compared to SGD, heuristic optimization algorithms offer several advantages that contribute to their superior performance in this study. SGD's reliance on gradient descent can lead it to get trapped in local optima, while heuristic optimizers excel at balancing exploration and exploitation. This exploration ability potentially helps them escape shallow local minima and find better solutions. Additionally, SGD performance is highly sensitive to learning rate selection, whereas heuristic algorithms typically require less hyperparameter tuning and can identify suitable configurations more efficiently. Finally, real-world problems often have complex error surfaces with multiple local minima. In essence, heuristic optimization's ability to balance exploration and exploitation, combined with less sensitivity to hyperparameter tuning, empowers it to navigate complex error surfaces and potentially outperform SGD in this specific application.

The performance of the proposed metaheuristic approach is further evaluated on 20 randomly chosen icing events from the testing set. These events' ILRs are predicted using both the metaheuristic optimizers and the standard SGD technique. The resulting predictions are then compared against the actual ASOS-based ILR values. Fig. 8 visualizes these comparisons.

Examination of the figure reveals that the metaheuristic approach generally outperforms the SGD technique. This is evident by the closer proximity of most points to the perfect-fit line ($y = x$) for the

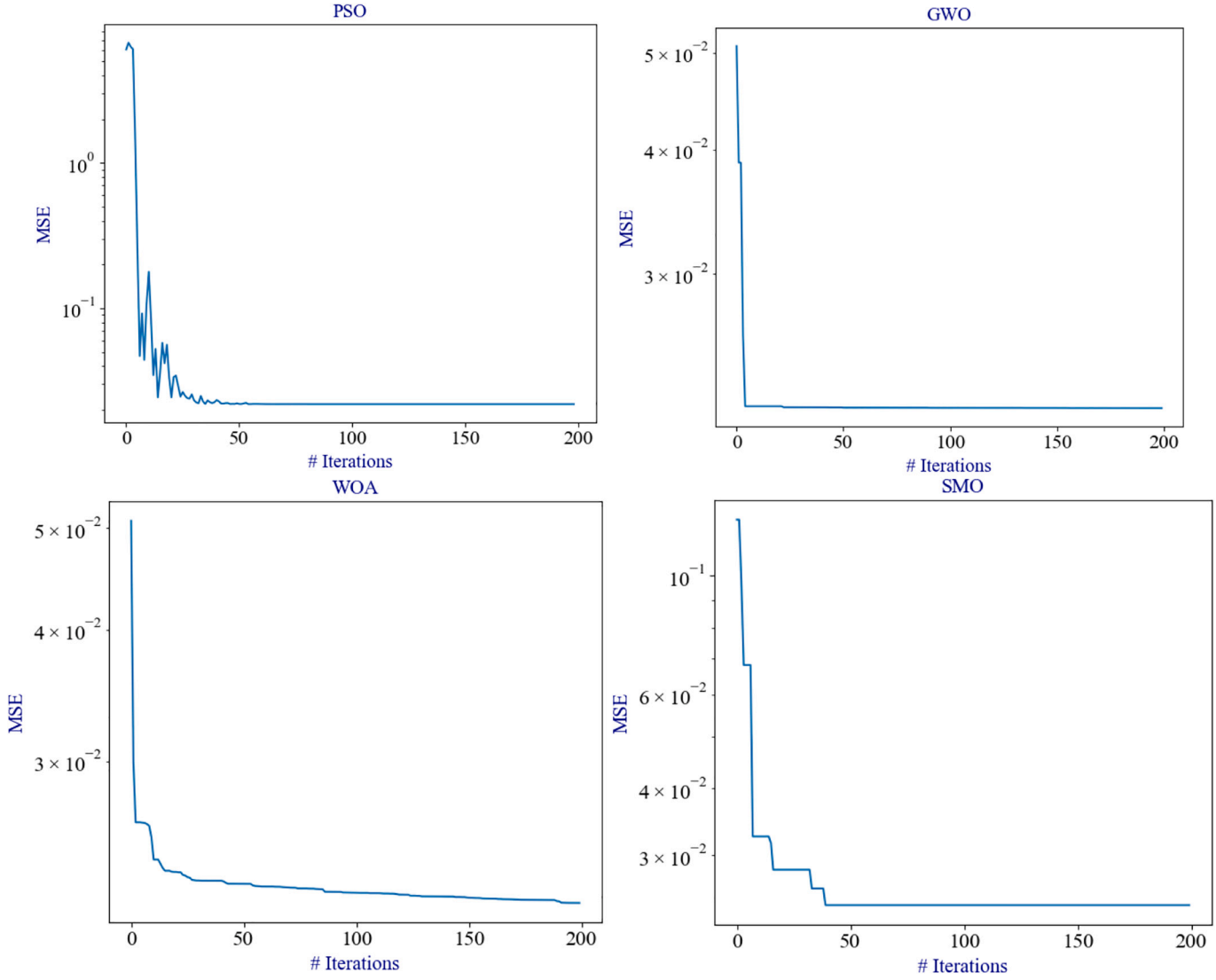

Fig. 6. MSE convergence curves using the PSO, GWO, WOA and SMO optimizers.

Table 1
Average, median, standard deviation, and best of MSE for all training samples over 10 independent runs using several optimization algorithms.

Algorithm	Average MSE	Median MSE	STD MSE	Best MSE
SGD	4.30E-02	4.24E-02	3.12E-03	3.91E-02
PSO	2.19E-02	2.19E-02	9.94E-05	2.17E-02
GWO	2.24E-02	2.22E-02	5.22E-04	2.19E-02
WOA	2.21E-02	2.21E-02	1.32E-04	2.19E-02
SMO	2.42E-02	2.44E-02	5.40E-04	2.31E-02

metaheuristic results. Additionally, the SGD model exhibits a tendency to overestimate ILRs below 0.65 and underestimate them for higher values.

To comprehensively understand the influence of various environmental parameters on the predictive performance of the neural network models, a systematic investigation was conducted. The initial step involved optimizing the neural network while incorporating all six initial environmental parameters. Subsequently, the performance of various four-parameter combinations was explored. Finally, a three-parameter model was tested using the three parameters most commonly employed in similar studies. The model trained with all six input parameters achieved comparable results in terms of prediction

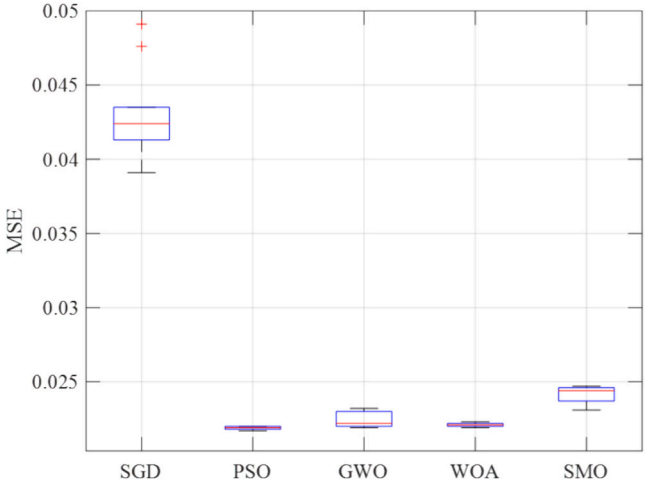

Fig. 7. Boxplot representation of the MSE for SGD, PSO, GWO, WOA and SMO on the ice accretion dataset.

Table 2
Mean square error (MSE) and correlation coefficient (R) of the test set using several optimization algorithms.

	Algorithm				
	SGD	PSO	GWO	WOA	SMO
MSE	4.98E-02	2.29E-02	2.29E-02	2.29E-02	2.57E-02
R	0.58	0.81	0.80	0.81	0.79

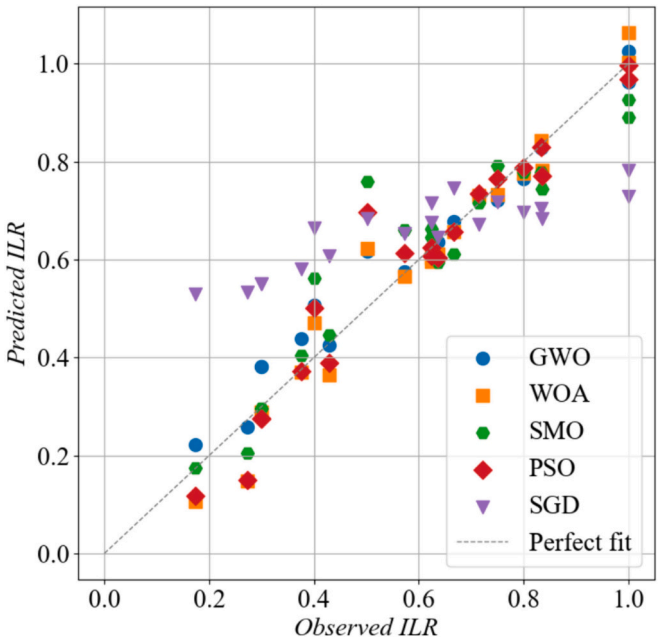

Fig. 8. Comparison of observed (ASOS) and predicted hourly ILR for 20 randomly selected icing events using metaheuristic and SGD optimization technique.

accuracy. However, this approach came at the expense of significantly increased training time. This rise in training time can be attributed to the model's increased complexity, requiring a more intricate optimization process with a larger number of tuning parameters. Next, the performance of various combinations of four input parameters was investigated. This analysis revealed that certain combinations, particularly those that included the core set identified by the Sobol index analysis (detailed in Table 3), achieved acceptable results. The best MSE values obtained during ten independent training runs for each combination are reported in Table 3. Finally, the three-parameter model incorporating the most commonly used parameters exhibited acceptable performance. However, the four-parameter set identified through the Sobol index analysis consistently yielded the best overall results. This four-parameter model achieved a superior balance between prediction accuracy and training efficiency. These findings strongly support the effectiveness of the global sensitivity analysis (Sobol index) in accurately identifying the most influential combination of input parameters for this specific neural network application.

Table 3
Best MSE of the test set using several input parameter combinations.

Combination	Algorithm			
	PSO	GWO	WOA	SMO
(T, U, P, WBT)	2.24E-02	2.75E-02	2.44E-02	2.56E-02
(T, DPT, G, P)	2.29E-02	3.36E-02	3.27E-02	2.60E-02
(DPT, U, P, WBT)	2.21E-02	2.34E-02	3.90E-02	2.54E-02
(DPT, G, P, WBT)	2.27E-02	2.74E-02	3.88E-02	2.62E-02
(T, U, P)	2.28E-02	2.34E-02	2.52E-02	2.44E-02

To further explore the individual impact of the four key environmental parameters identified by the Sobol index (precipitation, dew point temperature, temperature, and wind speed) on the predicted ILR, a case study analyzed how increasing each parameter by 10%, 20%, and 30% in a specific test case ($T = -1.4^{\circ}$C, $DPT = -4.12^{\circ}$C, $P = 0.08$ in/h and $U = 5.58$ kts) affects the predicted ILR. As expected and consistent with the global sensitivity analysis (Section 4.2), precipitation exhibited the strongest influence, followed by dew point temperature, temperature, and wind speed (results in Table 4). For example, a 30% increase in precipitation led to a 14.58% decrease in predicted ILR, while the same increase for wind speed only resulted in a 4.37% increase in predicted ILR. This suggests that higher precipitation rates lead to lower ILRs, likely due to increased runoff as the liquid doesn't freeze before more liquid is added to the ice surface (Sanders and Barjenbruch, 2016).

It's worth noting that the dataset used for training the model was drawn from a broad, geographically diverse range across the United States. This suggests that investigating the effects of site-specific factors could be valuable, potentially enhancing the model's predictive capabilities. Including such factors as additional inputs in the model might lead to further improvements in simulation accuracy. In addition, the ASOS data, including ice data and other meteorological variables, is collected at 1.5 m above ground level. Therefore, environmental data used with this model should ideally be given at the same height (1.5 m) for optimal accuracy. Wind speed is particularly sensitive to sensor height and requires conversion if measured at different heights (e.g., the standard 10-m level). Established methods like log or power law profiles, or other calibrated approaches, can be used for wind speed conversion. However, the model's intended use with operational forecast models, which typically predict ASOS-measured environmental variables (e.g., precipitation, wind speed, temperature, dewpoint), allows for flexibility in data height. If 10-m data is available, the model can be retrained on that configuration, and users should then provide inputs at the standard 10-m height. Future research exploring vertical wind profiles and scaling methods holds promise for further refinement. In addition, while the current approach identifies suitable hyperparameters (e.g., number of nodes in the hidden layer), a more systematic exploration could potentially lead to further improvements. Employing Bayesian optimization to explore the hyperparameter space offers a promising avenue for identifying configurations that optimize both model performance and training efficiency. Moreover, incorporating techniques like knowledge-enhanced neural networks (Karniadakis et al., 2021; Snaiki and Wu, 2019; Snaiki and Wu, 2022; Saviz Naeini and Snaiki, 2024) presents a promising avenue to potentially improve model performance. These networks leverage both data and underlying physical principles, offering the potential for more accurate and interpretable ILR predictions.

5. Concluding remarks

In this study, the *ILR* was predicted using a feedforward neural network (FFNN) trained with various metaheuristic optimizers to overcome the limitations of gradient-based algorithms. The environmental data for freezing rain events were sourced from the Automated Surface Observing System (ASOS), with analysis focusing on six key parameters. A Sobol index-based global sensitivity analysis, utilizing polynomial chaos expansion coefficients, was conducted to assess the significance of these input parameters. The results revealed that only four input parameters - precipitation, temperature, dew point temperature, and wind speed - substantially influenced the response variance. Five optimization algorithms were evaluated for training the FFNN model: particle swarm optimization (PSO), grey wolf optimizer (GWO), whale optimizer (WOA), slime mold optimizer (SMO), and the standard stochastic gradient descent (SGD). The metaheuristic optimizers consistently outperformed SGD. Training set MSE values for PSO, GWO, WOA, SMO, and SGD were 2.19E-02, 2.24E-02, 2.21E-02, 2.42E-02, and 4.30E-02, respectively. This trend continued for testing performance,

Table 4
Impact of individual parameter variations on predicted ILR.

% parameter increase	*P*		*DPT*		*T*		*U*	
	ILR	% *ILR* change	*ILR*	% *ILR* change	*ILR*	% *ILR* change	*ILR*	% *ILR* change
10%	0.569	−4.86	0.618	3.26	0.588	−1.66	0.607	1.46
20%	0.540	−9.72	0.637	6.52	0.579	−3.33	0.616	2.91
30%	0.511	−14.58	0.657	9.79	0.569	−4.99	0.625	4.37

with PSO, GWO, and WOA achieving comparable MSE values (around 2.29E-02) and high correlation coefficients (above 0.80). This superior performance is attributed to the metaheuristic optimizers' high exploratory and exploitative behaviors, which enable them to avoid local optima effectively and converge rapidly towards the global optimum. Overall, the proposed model has the potential to enhance the meteorological community's ability to predict freezing rain accumulation.

CRediT authorship contribution statement

Reda Snaiki: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Abdeslam Jamali:** Writing – original draft, Software, Data curation, Methodology. **Ahmed Rahem:** Writing – review & editing, Supervision, Investigation, Formal analysis, Conceptualization. **Mehdi Shabani:** Investigation, Software. **Brian L. Barjenbruch:** Writing – review & editing, Investigation, Formal analysis, Data curation.

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.

Data availability

Data will be made available on request.

Acknowledgement

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) [grant number CRSNG RGPIN 2022-03492].

References

Armenakis, C., Nirupama, N., 2014. Urban impacts of ice storms: Toronto December 2013. Nat. Hazards 74, 1291–1298.
Chaîné, P.M., Castonguay, G., 1974. New approach to radial ice thickness concept applied to bundle-like conductors. In: Industrial Meteorology Study, 4. AES Environment Canada, Toronto, ON, Canada, pp. 22–pp.
Chang, S.E., Mcdaniels, T.L., Mikawoz, J., Peterson, K., 2007. Infrastructure failure interdependencies in extreme events: power outage consequences in the 1998 Ice Storm. Nat. Hazards 41, 337–358.
Chang, S.E., Mcdaniels, T.L., Mikawoz, J., Peterson, K., 2012. Infrastructure Failure Interdependencies in Extreme Events: The 1998 Ice Storm. In: Disaster Risk and Vulnerability: Mitigation Through Mobilizing Communities and Partnerships, p. 275.
Chen, S., Dai, D., Huang, X., Sun, M., 2012. Short-term Prediction for Transmission Lines Icing Based on bp Neural Network. IEEE, pp. 1–5.
Chen, Y., Li, P., Ren, W., Shen, X., Cao, M., 2020. Field data–driven online prediction model for icing load on power transmission lines. Measure. Control 53, 126–140.
Dai, D., Huang, X.-T., Dai, Z., Hao, Y.P., Li, L.C., Fu, C., 2013. Regression model for transmission lines icing based on support vector machine. High Voltage Eng. 39, 2822–2828.
Degaetano, A.T., Belcher, B.N., Spier, P.L., 2008. Short-term ice accretion forecasts for electric utilities using the weather research and forecasting model and a modified precipitation-type algorithm. Weather Forecast. 23, 838–853.
Du, X., Zheng, Z., Tan, S., Wang, J., 2010. The Study on the Prediction Method of Ice Thickness of Transmission Line Based on the Combination of GA and BP Neural Network. IEEE, pp. 1–4.
Fikke, S.M., Kristjánsson, J.E., Nygaard, B.E.K., 2008. Modern meteorology and atmospheric icing. In: Farzaneh, M. (Ed.), Atmospheric Icing of Power Networks. Springer.
Geem, Z.W., Kim, J.H., Loganathan, G.V., 2001. A new heuristic optimization algorithm: harmony search. simulation 76, 60–68.
Goodwin, E.J., Mozer, J.D., Digioia, A.M., Power, B.A., 1983. Predicting ice and snow loads for transmission line design. Predict. Ice Snow Loads Trans. Line Design 267–273.
He, L., Luo, J., Zhou, X., 2021. A Novel Deep Learning Model for Transmission Line Icing Thickness Prediction. IEEE, pp. 733–738.
Huang, X., Li, J., 2012. Icing Thickness Prediction Model Using BP Neural Network. IEEE, pp. 758–760.
Huang, X., Xu, J., Yang, C.S., Wang, J., Xie, J.J., 2014. Transmission line icing prediction based on data driven algorithm and LS-SVM. Autom. Electr. Power Syst 38, 81–86.
Imai, I., 1953. Studies of ice accretion. Res. Snow Ice 1, 35–44.
Jarrett, P., Yau, K.-H., Morris, R., 2019. Update of the reference ice thickness amounts due to freezing rain for Canadian codes and standards, 2019, 23–28.
Jeong, D.I., Sushama, L., Vieira, M.J.F., Koenig, K.A., 2018. Projected changes to extreme ice loads for overhead transmission lines across Canada. Sustain. Cities Soc. 39, 639–649.
Jeong, D.I., Cannon, A.J., Zhang, X., 2019. Projected changes to extreme freezing precipitation and design ice loads over North America based on a large ensemble of Canadian regional climate model simulations. Nat. Hazards Earth Syst. Sci. 19, 857–872.
Jones, K.F., 1998. A simple model for freezing rain ice loads. Atmos. Res. 46, 87–97.
Karniadakis, G.E., Kevrekidis, I.G., Lu, L., Perdikaris, P., Wang, S., Yang, L., 2021. Physics-informed machine learning. Nat. Rev. Phys. 3 (6), 422–440.
Kennedy, J., Eberhart, R., 1995. Particle Swarm Optimization. IEEE, pp. 1942–1948.
Kirkpatrick, S., Gelatt, J.R., C. D, Vecchi, M.P., 1983. Optimization by simulated annealing. science 220, 671–680.
Lenhard, J.R., R. W, 1955. An indirect method for estimating the weight of glaze on wires. Bull. Am. Meteorol. Soc. 36, 1–5.
Li, S., Chen, H., Wang, M., Heidari, A.A., Mirjalili, S., 2020. Slime mould algorithm: A new method for stochastic optimization. Future generation computer systems 111, 300–323.
Li, P., Li, Q., Cao, M., Gao, S., Huang, H., 2011. Time Series Prediction for Icing Process of Overhead Power Transmission Line Based on BP Neural Networks. IEEE, pp. 5315–5318.
Lozowski, E.P., Makkonen, L., 2005. Fifty Years of Progress in Modelling the Accumulation of Atmospheric Ice of Power Network Equipment.
Ma, T., Niu, D., 2016. Icing forecasting of high voltage transmission line using weighted least square support vector machine with fireworks algorithm for feature selection. Appl. Sci. 6, 438.
Ma, T., Niu, D., Fu, M., 2016. Icing forecasting for power transmission lines based on a wavelet support vector machine optimized by a quantum fireworks algorithm. Appl. Sci. 6, 54.
Makkonen, L., 1984. Modeling of ice accretion on wires. J. Appl. Meteorol. Climatol. 23, 929–939.
Makkonen, L., 1998. Modeling power line icing in freezing precipitation. Atmos. Res. 46, 131–142.
Mallat, S.G., Zhang, Z., 1993. Matching pursuits with time-frequency dictionaries. IEEE Trans. Signal Process. 41, 3397–3415.
Mirjalili, S., Lewis, A., 2016. The whale optimization algorithm. Advances in engineering software 95, 51–67.
Mirjalili, S., Mirjalili, S.M., Lewis, A., 2014. Grey wolf optimizer. Advances in engineering software 69, 46–61.
Mladenović, N., Hansen, P., 1997. Variable neighborhood search. Comput. Oper. Res. 24, 1097–1100.
Niu, D., Wang, H., Chen, H., Liang, Y., 2017. The general regression neural network based on the fruit fly optimization algorithm and the data inconsistency rate for transmission line icing prediction. Energies 10, 2066.
Ojha, V.K., Abraham, A., Snášel, V., 2017. Metaheuristic design of feedforward neural networks: a review of two decades of research. Eng. Appl. Artif. Intell. 60, 97–116.
Pati, Y.C., Rezaiifar, R., Krishnaprasad, P.S., 1993. Orthogonal Matching Pursuit: Recursive Function Approximation with Applications to Wavelet Decomposition. IEEE, pp. 40–44.
Rashedi, E., Nezamabadi-Pour, H., Saryazdi, S., 2009. GSA: a gravitational search algorithm. Inf. Sci. 179, 2232–2248.
Ruszczak, B., Tomaszewski, M., 2014. Extreme value analysis of wet snow loads on power lines. IEEE Trans. Power Syst. 30, 457–462.
Ryerson, C.C., Ramsay, A.C., 2007. Quantitative ice accretion information from the Automated Surface Observing System. J. Appl. Meteorol. Climatol. 46, 1423–1437.

Sanders, K.J., Barjenbruch, B.L., 2016. Analysis of ice-to-liquid ratios during freezing rain and the development of an ice accumulation model. Weather Forecast. 31, 1041–1060.
Saviz Naeini, S., Snaiki, R., 2024. A novel hybrid machine learning model for rapid assessment of wave and storm surge responses over an extended coastal region. Coastal Engineering 190.
Saviz Naeini, S., Snaiki, R., 2024. A physics-informed machine learning model for time-dependent wave runup prediction. Ocean Eng. 295, 116986.
Shabani, M., Jamali, A., Snaiki, R., Rahem, A., 2022. Prediction of Ice Accretion on Transmission Lines Using Hybrid Particle Swarm Optimization-Based Artificial Neural Networks.
Sheng, C., Tang, Q., Hong, H.P., 2023. Estimating and mapping extreme ice accretion hazard and load due to freezing rain at Canadian Sites. Int. J. Disaster Risk Sci. 14, 127–142.
Snaiki, R., 2024. Performance-based ice engineering framework: A data-driven multi-scale approach. Cold Regions Science and Technology 104247.
Snaiki, R., Wu, T., 2019a. Knowledge-enhanced deep learning for simulation of tropical cyclone boundary-layer winds. J. Wind Eng. Ind. Aerodyn. 194, 103983.
Snaiki, R., Wu, T., 2022. Knowledge-enhanced deep learning for simulation of extratropical cyclone wind risk. Atmosphere 13 (5), 757.
Stieltjes, T.J., 1884. Quelques recherches sur la théorie des quadratures dites mécaniques, pp. 409–426.
Sun, W., Wang, C., 2019. Staged icing forecasting of power transmission lines based on icing cycle and improved extreme learning machine. J. Clean. Prod. 208, 1384–1392.
Tomaszewski, M., Ruszczak, B., Michalski, P., Zator, S., 2019. The study of weather conditions favourable to the accretion of icing that pose a threat to transmission power lines. Int. J. Crit. Infrastruct. Prot. 25, 139–151.
Wu, T., Snaiki, R., 2022. Applications of machine learning to wind engineering. Front. Built Environ. 8, 811460.
Xu, X., Niu, D., Wang, P., Lu, Y., Xia, H., 2015. The weighted support vector machine based on hybrid swarm intelligence optimization for icing prediction of transmission line. Math. Probl. Eng. 2015.
Yin, S., Lu, Y., Wang, Z., Li, P., Xu, K., 2012. Icing thickness forecasting of overhead transmission line under rough weather based on CACA-WNN. Electr. Power Sci. Eng 11.
Ying, Z.-R., Su, X.-L., 2014. Icing thickness forecasting of transmission line based on particle swarm algorithm to optimize SVM. J. Electr. Power 29, 6–9.
Zarnani, A., Musilek, P., Shi, X., Ke, X., He, H., Greiner, R., 2012. Learning to predict ice accretion on electric power lines. Eng. Appl. Artif. Intell. 25, 609–617.
Zhang, Z., Huang, H., Jiang, X., Hu, J., Sun, C., 2012. Analysis of ice growth on different type insulators based on fluid dynamics. Diangong Jishu Xuebao (Transact. China Electrotech. Soc.) 27, 35–43.