Chemosphere 362 (2024) 142683

Contents lists available at ScienceDirect

Chemosphere

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

Polychlorinated dibenzo-p-dioxins and polychlorinated dibenzofurans (PCDD/Fs) prediction model based on limited peat samples using an evolved artificial neural network

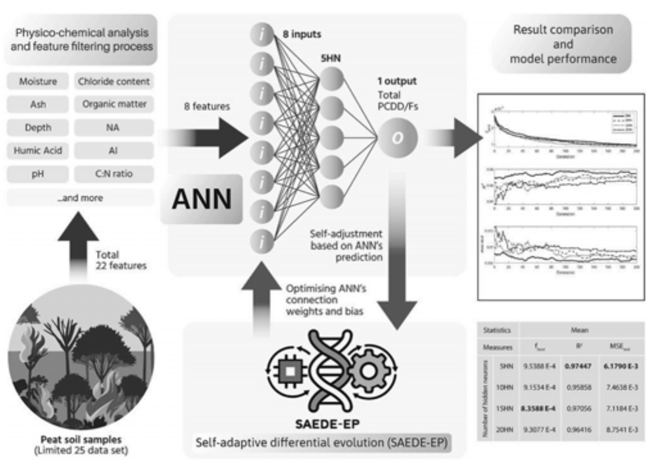
Shir Li Wang [a,b], Theam Foo Ng [c], Khairulmazidah Mohamed [d,e], Sumayyah Dzulkifly [a], Xiaodong Li [f], Yin-Hui Leong [d,*]

[a] Faculty of Computing and Meta-Technology, Universiti Pendidikan Sultan Idris, 35900, Tanjong Malim, Perak, Malaysia
[b] Data Intelligent and Knowledge Management (DILIGENT), Universiti Pendidikan Sultan Idris, 35900, Tanjong Malim, Perak, Malaysia
[c] Centre for Global Sustainability Studies, Universiti Sains Malaysia, 11800, Penang, Malaysia
[d] National Poison Centre, Universiti Sains Malaysia, 11800, Penang, Malaysia
[e] Faculty of Applied Sciences, Universiti Teknologi MARA, 40450, Shah Alam, Selangor, Malaysia
[f] State Key Laboratory of Clean Energy Utilization, Yuquan Campus of Zhejiang University, Hangzhou, Zhejiang Province, 310027, PR China

HIGHLIGHTS

- PCDD/Fs prediction model is built based on an evolutionary-optimised ANN.
- Adopted limited datasets of PCDD/Fs emissions and physico-chemical of peat samples.
- Differential evolutionary algorithm improved accuracy and error estimates.
- Cost-effective solution for pollution and environmental monitoring.

GRAPHICAL ABSTRACT



ARTICLE INFO

Handling editor: Michael Bank

Keywords:
Polychlorinated dibenzo-p-dioxins and polychlorinated dibenzofurans (PCDD/Fs)
Differential evolution
Artificial neural network

ABSTRACT

Polychlorinated dibenzo-p-dioxins and polychlorinated dibenzofurans (PCDD/Fs) are involuntary by-products of incomplete combustion and are highly toxic to humans and the environment. The Malaysian peat is often acidic or extremely acidic having high levels of chlorine and/or other organic acids that act as catalysts or precursors in PCDD/Fs formation. This study aims to predict PCDD/Fs emissions in peat soil using an artificial neural network (ANN) approach based on limited emission data and selected physico-chemical properties. The ANN's prediction performance is affected by uncertainties in its initial connection weights. To improve prediction performance, an

* Corresponding author.
E-mail address: yhleong@usm.my (Y.-H. Leong).

https://doi.org/10.1016/j.chemosphere.2024.142683
Received 12 March 2024; Received in revised form 20 May 2024; Accepted 20 June 2024
Available online 21 June 2024
0045-6535/© 2024 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

Peatland
Adaptive parameters

optimisation algorithm, termed differential evolution (DE), is used to optimise the ANN's initial connection weights and bias. The study adopts several ANNs with fixed architecture to predict PCDD/Fs emissions, each consisting of a multilayer perceptron (MLP) with a backpropagation algorithm. Eight input variables and one output variable were adopted to train and test various neural network architectures using real-world datasets. The model optimisation procedure was conducted to ascertain the network architecture with the best predictive accuracy. The evolved ANN based on 5 hidden neurons, with the assistance of self-adaptive ensemble-based differential evolution with enhanced population sizing (SAEDE-EP), successfully produced the lowest MSE_{test} (6.1790×10^{-3}) and highest R^2 (0.97447) based on the mean among the other HNs. An evolutionary-optimised ANN-based methodology is a viable solution to predict PCDD/Fs in peat soil. It is cost-effective for pollution control, environmental monitoring and capable of aiding authorities prevent PCDD/Fs exposure, e.g., during a fire.

1. Introduction

Peatlands play a vital role in the global ecosystem serving as carbon sinks and providing critical ecological services. However, the degradation of peatlands can lead to the release of harmful pollutants, including polychlorinated dibenzo-p-dioxins and polychlorinated dibenzofurans (PCDD/Fs), posing significant environmental and health risks. Dioxins are the collective name for 210 different polychlorinated dibenzo-p-dioxins (PCDDs) and polychlorinated dibenzofurans (PCDFs). The most potent dioxin congener is 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD), classified as a Class 1 human carcinogen by the International Agency for Research on Cancer (IARC, 1997).

Peat is primarily derived from plant components such as roots, leaves, stems, small insects and animals. It is composed of organic materials, including carbohydrates and lignins, which can serve as precursors for PCDD/Fs formation when burned. Any process that involves chlorine or chloride has the potential to generate PCDD/Fs. Extensive research has been conducted on PCDD/Fs emission from combustion sources such as medical and municipal solid waste (MSW) incinerators. Even small amounts of organic or inorganic chloride can produce low levels of PCDD/Fs emissions (approximately 10×9 to 10×12 g Nm^{-3}) during combustion processes (Uloth et al., 2005).

Peat soil in Malaysia covers approximately 7% or 2.5 million hectares of the country's land area, and peat fires are a recurrent issue. Malaysian tropical peat is typically highly acidic, with a pH range of 2.68–4.50 (Andriesse, 1988; Könönen et al., 2015; Sutejo et al., 2017). This suggests that the presence of high levels of chlorine and/or other organic acids could lead to PCDD/Fs emissions during combustion. Burning peat produces a smouldering fire that can occur on both the surface and underground layers of the peat. This burning process can last for an extended period, ranging from several days to even months. During this time, the temperature can reach between 450 and 700 °C, potentially creating the formation of PCDD/Fs through either the "homogeneous" or "heterogeneous" pathway (Rein, 2016). This condition has worsened over the past few decades after many tropical peatlands were converted to agricultural farms where herbicides are commonly used. The risks of PCDD/Fs formation at such sites are further amplified when fires occur.

Determining the level of PCDD/Fs emission through analytical methods is a tedious process requiring advanced instruments, expensive chemicals and standards, as well as skilled laboratory staff. As such, it involves higher costs and a significant amount of time. Hence, predicting PCDD/Fs emission concentrations using mathematical modelling and machine learning methods is a potentially cost-effective alternative. Several studies have reported the prediction of PCDD/Fs emission by using these approaches. A study by Blumenstock et al. (1999) investigated the estimation of PCDD/Fs emission in the fuel and stack gas of a hazardous waste incinerator. They utilised the principal component analysis (PCA) method and correlation coefficients through linear regression to establish the relationships between PCDD/Fs emission and potential indicator substances. In another study, Choi et al. (2007) employed multi-regression analysis to predict PCDD/Fs variations using collected sampling data sets from seven MSW incinerators in Korea. Bunsan et al. (2013) presented a PCDDs prediction model based on a back-propagation neural network (BPNN), a promising method in handling complex and non-linear data with the assistance of statistics in selecting useful variables for modelling. They found that an artificial neural network (ANN) architecture that consists of 5 input factors and 3 basic layers with 8 hidden nodes was suitable for PCDDs prediction (Bunsan et al., 2013).

Several data-driven models, such as support vector machines (SVMs) and neural networks have been used to solve dioxins prediction problems (Wang et al., 2008; Xiao et al., 2017; Tang et al., 2019). Chang and Chen (2000) developed a PCDD/Fs emission model based on genetic programming and neural network to establish the unknown mapping relation among input features and PCDD/Fs. Bunsan et al. (2013) used correlation and PCA analyses to determine easily detectable process variables as input and construct a backpropagation neural network (BPNN) framework to model PCDDs effectively. However, these methods have their challenges due to the small sampling size, strong collinearity among input features and uncertainty in the MSW incineration process, which can potentially lead to overfitting, local minimums, and poor predictive performance stability (Bunsan et al., 2013). To address these issues, Xia et al. (2020) developed a dioxins prediction model with a hybrid integration method of random forest (RF) and gradient boosting decision tree. However, it was not effective in performing feature selection and analysis on MSW incineration process variables. Hence, Qiao et al. (2021) proposed a dioxins emission concentration detection mechanism with a multilayer feature selection strategy, while Tang et al. (2021a, 2021b) suggested an improved feature reduction theme and selective ensemble algorithm for this purpose. Ensemble learning appears to be a suitable strategy for modelling dioxins prediction, and Xia et al. (2022) have proposed a dioxins emission concentration prediction model based on an improved deep forest regression (ImDFR) approach. The ImDFR model outperforms the BPNN, RF, extreme gradient boosting, and deep forest regression models in terms of predictive performance and the lowest time cost (Xia et al., 2022). Machine learning has been applied widely in solving complex problems of solid waste-related issues, including forecasting PCDD/Fs emission during the MSW incineration process. However, to our knowledge, none of the research focuses on peatlands which are also potential PCDD/Fs sources.

The objective of this study is to address the pressing need for accurate and reliable PCDD/Fs prediction models for effective pollution control and environmental monitoring in peatlands. ANN is known to be able to capture complex nonlinear relationships, while evolutionary optimisation algorithms enable the finetuning of model parameters to maximise accuracy with limited available data. By integrating an ANN with evolutionary algorithms (EA), the evolved ANN model aims to improve the predictive capabilities while minimising the requirement for extensive datasets, making it cost-effective and practical for real-world applications.

The rest of this paper is organised as follows: Section 2 describes the studies related to PCDD/Fs data collection and feature selection. Section 3 presents the methodology of our evolved ANN model, and Section 4 describes the experimental setup. Section 5 discusses the experimental results through analysis. Section 6 presents the conclusions and future work.

2. Data collection and feature selection

The raw data of physico-chemical properties of peat soil samples contain 22 features, including contents of moisture, ash, chloride and humid acid. All the features are listed in supplementary material (Table S1). To ensure accurate predictions of PCDD/Fs emissions, it is essential to identify the relevant features that impact PCDD/Fs production in tropical peat soil. The inclusion of irrelevant features may affect the model's accuracy, prolong the training process, and increase computer memory usage. Thus, expert input and exploratory data analysis were performed to select the potential features that could affect PCDD/Fs formation or indicate high correlation with PCDD/Fs emissions. Consequently, 8 features were selected as the input variables from the 22 features to construct the PCDD/Fs prediction model (Table S2). This feature filtering process helped to streamline the network model and minimise training costs. The low correlation coefficients between the 8 input features and total PCDD/Fs concentration indicate a weak linear relationship (Table S3).

In this study, ANN-based PCDD/Fs prediction model was constructed based on the PCDD/Fs emission data and selected features: depth of peat soil, carbon-to-nitrogen ratio (C:N ratio), total humic acid (HA), phenol, 1,2,3-trichlorobenzene, chloride, Cu and Al contents. This is the first study adopting these variables for PCDD/Fs prediction model. A total of 25 peat samples were randomly collected from three states (Johor, Selangor and Terengganu), which have a high distribution of peat. These samples were collected from five different locations, including three plantations and two forest reserve areas. Sampling was conducted under different environmental and geographical conditions to represent the tropical peat of Peninsular Malaysia.

Sampling was done using an Auger core sampler by rotary drilling at different depths: surface, 1–50 cm, 51–100 cm, 101–150 cm, and 151–200 cm. Physico-chemical properties and total concentration of PCDD/Fs emission evolved with the different soil levels. The depths in ascending order were recorded into levels 0 to 4 to be used as one of the inputs to train the ANN model. C:N ratio determination was conducted to characterise the amount of carbon relative to the amount of nitrogen present in the peat samples. Analysis was performed using PerkinElmer 2400 Series II CHN Elemental Analyzer. Total HA in peat was analysed using a combination of ISO 5073:2013 method for the extraction procedure and Javanshah and Saidi's (2016) method using UV–Vis spectrophotometer. Two compounds with the most abundant (phenol and 1,2,3-trichlorobenzene) among the 16 chlorobenzenes and chlorophenols compounds were chosen for the PCDD/Fs modelling. Most of the samples were found to contain these two chemicals. In the chlorobenzenes and chlorophenols analysis, soil samples were first extracted using an Accelerated Solvent Extraction system, followed by quantification on Shimadzu Gas Chromatography Mass Spectrometer QP2010 Ultra equipped with Rtx-5MS with 30 m length, 0.25 mm ID and 0.25 μm df for separation. In addition, two cation elements (Cu and Al) and one anion (chloride) were selected for the modelling. The reason for selecting Cu and Al in this study is because of their expected ability to catalyse and retard PCDD/Fs formation respectively. Cations were analysed using an Inductive Coupled Plasma – Optical Emission Spectroscopy while anion was analysed using Metrohm Ion Chromatography instrumentation. The wet ashing method for sample extraction was applied for both analyses. Total PCDD/Fs emission was determined using High Resolution Gas Chromatography/High Resolution Mass Spectrometry (HRGC/HRMS) 6890 Series Gas Chromatograph (Agilent, USA) coupled to a JMS-800D mass spectrometer (JEOL, Japan) (Ying et al., 2023). The separation of PCDD/Fs congeners was achieved by a DB-5MS (60 m × 0.25 mm ID, 0.25 μm film thickness) column. Prior to PCDD/Fs quantification using HRGC/HRMS, the sample underwent pre-treatment procedures following the US EPA Method 1613 (USEPA, 1994).

3. Methodology

3.1. Artificial neural networks (ANNs)

ANNs have been widely used to solve solid waste-related issues, and their applications can be found in a review by Xu et al. (2021). Based on Xu's work, a single hidden layer is commonly sufficient to solve the prediction problems related to solid waste management. Xu's work also recommends researchers optimise the number of hidden layer nodes within at least 4–20 due to its maximum probability. It is suggested, therefore, that a shallow neural network instead of a deep neural network (DNN) be used to solve this problem. As such, a similar range of hidden neurons (HN) for the architecture of ANNs was used in this study. The performances of the ANNs are tested with different HN, i.e., 5, 10, 15 and 20. The ANNs are denoted as 5HN, 10HN, 15HN and 20HN, based on their number of HN.

Using ANNs does pose uncertainties: data set participation and ANNs' initial weight. A cross validation method can overcome the uncertainties derived from data set participation. Hence, we use similar approaches to reduce the uncertainties. Our model employs cross validation and optimisation of ANN's initial connection weights.

Based on the work performed by Xu et al. (2021), only five articles focus on using EA to optimise the ANNs' initial connection weights. Four articles adopted genetic algorithms (Bagheri et al., 2015; Lu et al., 2016; Oliveira et al., 2019; Soni et al., 2019), and one adopted particle swarm optimisation (PSO) (Ebrahimzade et al., 2020) to optimise the ANNs' initial connection weights. In contrast to the existing work, another EA, called differential evolution (DE), was used to optimise the ANNs' initial weight in our work. DE is widely used to solve optimisation problems owing to its simplicity, robustness, and computation efficacy. An adaptive DE called self-adaptive ensemble-based differential evolution with enhanced population sizing (SAEDE-EP) was used to optimise the ANNs' initial connection weights since the algorithm can adaptively adjust its parameters: scale factor, crossover rate, mutation strategy and population size (NP), with minimum reliance on the user to determine the parameter configurations.

ANNs are computing units inspired by the biological complex neural structures to solve complex nonlinear relationships (Ebrahimzade et al., 2020). A commonly used ANN is a multilayer perceptron (MLP) neural network. Each neural network consists of multilayers of computing units with activation functions. Input and outputs are directly related to the input and output layers of an ANN. The hidden layers are located between the ANN's input and output layers. The information from one layer's neurons to the subsequent layer's neurons is transformed based on Eq. (1).

$$a = \zeta(b + wp) \tag{1}$$

where p is the input, b is the bias, w is the neurons' connection weights, ζ the activation or transfer function, and a is the neurons' output. Neurons can use any differentiable transfer function ζ to generate their output. The three commonly used transfer functions consist of logsig, tansig and purelin. The MLP training is carried out by iteratively weighing procedures from the inputs to the output neurons. After each iteration, the computed outputs are compared with the real outputs (targets) in the base of mean squared error (MSE) or other statistical measure criteria for a data set assigned as the training set. In this process, the connection weights and bias were modified according to the learning algorithm, such as backpropagation, to minimise the MSE of the training set in each iteration. The network's generalisation during the training phase is validated with another set of data called validation set. Following the training phase, the testing phase was performed with previously unused data in the training step and the unused data is called test set.

Each backpropagation training session begins with different initial connection weights and bias, and different data divisions into training, validation, and test sets. These different conditions can lead to different solutions for the same problem (Ebrahimzade et al., 2020). The performance of an ANN is affected by uncertainties from its initial connection weights and the data set partition (Xu et al., 2021). Considering the risk of affecting the modelling with suboptimal solutions, improving uncertainties derived from initial connection weights and data set partition are also highly recommended for modelling improvements.

ANN has three layers: an input layer, a hidden layer and an output layer. The neurons are named according to the layers: input (in), hidden (hn) and output (on) neurons. The architecture of an ANN is displayed in Fig. 1, where IW and LW refer to the connection weight matrices of hn × in and on × hn dimensions, respectively between the input-hidden and hidden-output layers. $\widehat{\rho}^1$ with in × 1 dimensions and $\widehat{\rho}^2$ with on × 1 dimensions refer to the vectors of bias units between the input-hidden and hidden-output layers. The number of input and output neurons is determined in quite a straightforward way, as they are based on the numbers of input and output variables. Hence, only the number of HN will be determined arbitrarily by a user.

An ANN with evolved initial connection weights learn faster and better than those with initial random connection weights (Nolfi and Floreano, 2000). Therefore, an adaptive EA is helpful in optimising the initial connection weights and bias between the layers of neurons. With the optimised initial connection weights and bias, the ANN was trained with fixed architecture and parameters in the evolved ANN model.

Each ANN in the population is evaluated iteratively for a number of simulation runs. At each time step in a simulation run, an individual ANN receives the following input and output variables in the PCDD/Fs modelling. Dividing data into training–validation–test sets is preferable in training and evaluating an ANN; the training set is supposed to account for 70% of the data set (Xu et al., 2021). Hence, this study applies the split ratio of 70:15:15, commonly applied for dividing data into training, validation, and test sets (Pandey et al., 2016). Since the sample size in our data is very limited, that is 25, we divided our data set based on the approximated ratio. The samples used for training and validation sets are 20, and the test set is 5. The fitness of each neural network is evaluated iteratively based on its performance in solving the modelling problem, that is, how low the MSE of the training and validation sets averaged over several repeated runs. Once the ANNs in the population were evaluated, SAEDE-EP was used to create the next generation of ANNs.

3.2. Differential evolution (DE)

Optimising initial connection weights and bias using other computational intelligence algorithms to solve solid waste-related problems has received research attention. However, there is a lack of studies on using evolutionary and PSO algorithms to optimise initial connection weights and bias (Xu et al., 2021). The hybrid of ANN and EA to solve a modelling problem requires the interested user to have knowledge and experience in both algorithms. Adaptation has been widely used in setting EA parameters to minimise the reliance on users and algorithms to solve optimisation problems. Therefore, an adaptive EA called SAEDE-EP (Budiman et al., 2020) was adopted in our research to evolve an ANN's initial connection weights and bias.

The connection weights and bias matrices are transformed into a row vector and encoded into a chromosome in SAEDE-EP. The row vector's length, D, is shown in Eq. (2).

$$D = (in+1)*hn + (hn+1)*on \tag{2}$$

Based on the PCDD/Fs modelling, the number of input neurons is based on the number of inputs: 8 input neurons and 1 output neuron. The rule of thumb for setting a reasonable NP lies in the dimensionality of problems, i.e., NP ∈ [5D, 10D], which is based on the suggestion by Storn and Price (1997), the authors of DE. Therefore, the initial population size, NP_{init} is set at approximately 5D.

Assuming that the number of HD is 10 and the variable D equals 501 (i.e., $(8+1)*10+(10+1)*1=501$), the $NP_{init}=510$, is used. The same procedure determines the NP_{init} for the ANNs with different numbers of HD except for 5HN. NP in SAEDE-EP is not constant but dynamically changes within [5D, 10D].

The variable length for the ANN of 5HN is 51, and the NP_{init} was supposed to be 255. Instead, $NP_{init}=500$ and NP ∈ [500, 600] were used to increase the population diversity. Large NP can diversify the direction and magnitude of the difference vector, which is conducive to finding more potential solutions (Li et al., 2023). Table S4 shows the range of NP denoted by [NP_{min}, NP_{max}] for the ANNs with different numbers of HN.

Given that ANN's architecture is fixed, IW and LW refer to the connection weights established between the input-hidden layers and hidden-output layers, respectively. The bias between the input-hidden layers and hidden-output layers are represented by $\widehat{\rho}^1$ and $\widehat{\rho}^2$. The solutions for IW, LW, $\widehat{\rho}^1$ and $\widehat{\rho}^2$ are represented by a population of individuals in SAEDE-EP.

SAEDE-EP begins with an NP_{init} of randomly generated individuals, yielding different connection weights and bias for a neural network. The network architecture and its relevant learning parameters are fixed and identical for all individuals. All networks in the population are then evaluated for several repeated runs. Once all individuals in the populations are evaluated, a new population is created through the differential mutation, crossover, and selection processes iteratively until a stopping criterion is achieved. The best individual after achieving the stopping criterion is used to form the ANN's initial connection weights and bias to solve the modelling problem. The fitness, f of individuals in the population is evaluated based on MSE of training (MSE_{train}) and validation ($MSE_{validate}$) sets (Eq. (3)).

$$f = MSE_{train} + MSE_{validate} \tag{3}$$

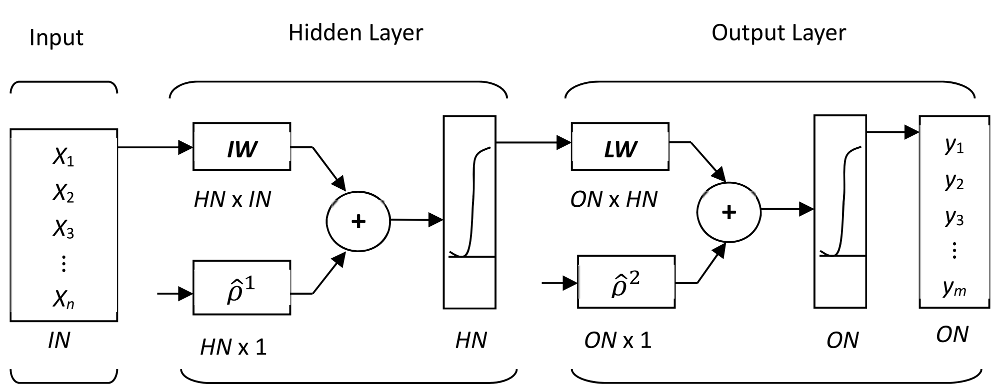


Fig. 1. The architecture of an ANN.

$MSE_{validate}$ is commonly used to monitor an ANN's generalisation. Generalisation refers to the ability of the machine learning algorithm to correctly apply the knowledge it learned to a new situation (Ferro et al., 2023). Generalisation can be achieved if a machine learning algorithm can avoid overfitting or overtraining. Overfitting in an ANN means the algorithms predict well in the training set but fail to reliably predict unseen data. Minimising the generalisation error is used to ensure an algorithm's generalisation. The accuracy of the algorithm on the validation set is used as the indicator of its generalisation error (LeCun et al., 1990). Hence, we used cross validation as the predictive basic to anticipate overfitting in the neural network's learning. A generalised neural network should perform appropriately well in training and validation sets. Thus, MSE_{train} and $MSE_{validate}$ are used as the fitness criteria to evaluate the modelling procedure and measure the effect of function on the DE-ANN optimisation. The model with the MSE value closer to zero performed better within the proposed models. Given that an ANN must have the minimum MSE of training and validation sets, the optimisation of the ANN's initial connection weights and bias is a minimisation problem.

However, the DE's performances are affected by the setting of its control parameters (Parouha and Verma, 2022) consisting of scale factor, crossover rate, NP and mutation strategy. There is no clear guidance or generalised rules when determining the parameters to ensure an acceptable performance (Yang, 2020). Most of the rules are either unclear, confusing, or contradicting. Therefore, an adaptive mechanism is commonly used to set DE's parameters. SAEDE-EP is one of the DE variants that operates based on an adaptive mechanism while having another two parameters, namely, the growth rate range of population size, R and stagnation threshold, T. $T = 30$ is used to control the stagnation level of SAEDE-EP. Besides T, SAEDE-EP's parameters can dynamically change within the options shown in Table S5. The stopping criterion is either the best-fitness, $f_{best} < 0.000001$ or the maximum generation, $G_{max} = 200$. NP in SAEDE-EP is limited within [5D, 10D], with D referring to the total length of connection weights and bias of an ANN.

3.3. Representation of connection weights and bias in chromosomes

The population of individuals represents the solutions for the connection weights and bias established between the input-hidden layers and hidden-output layers, respectively. Fig. 2 demonstrates the representation of an individual in the population. Therefore, IW, LW, $\widehat{\rho}^1$ and $\widehat{\rho}^2$ are mapped into an individual's vector.

$\widehat{W}_{IW}$ refers to a vector of connection weights and bias units between the input-hidden layers, mapped as $\widehat{W}_{IW} = \left\{\widehat{w}^1_{1,}, \widehat{w}^1_2, \cdots, \widehat{w}^1_{hn}, \cdots, \widehat{w}^1_{HN}\right\}$, as displayed in Fig. 3(a).

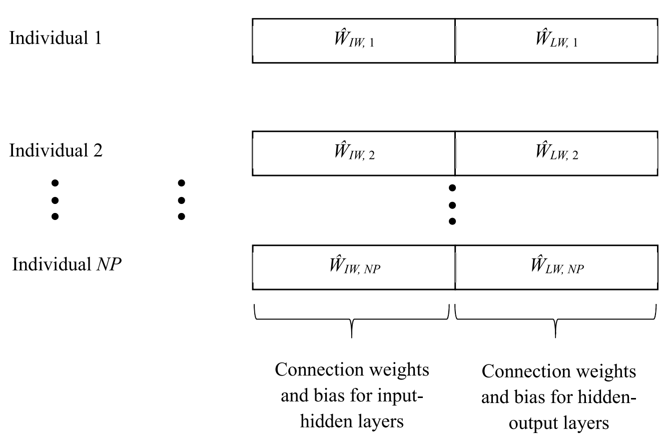


Fig. 2. Weights for the fixed architecture neural network in chromosome representation.

Similarly, $\widehat{W}_{LW}$ refers to a vector of connection weights and bias units between the hidden-output layers, mapped as $\widehat{W}_{LW} = \left\{\widehat{w}^2_{1,}, \widehat{w}^2_2, \cdots, \widehat{w}^2_{on}, \cdots, \widehat{w}^2_{ON}\right\}$, as shown in Fig. 3(b). The parameters in, hn and on with ranges $1 \leq \text{in} \leq \text{IN}$, $1 \leq \text{hn} \leq \text{HN}$, $1 \leq \text{on} \leq \text{ON}$ refer to the input, hidden and output layers' neurons.

A fixed architecture and parameters of the neural network are used to model the PCDD/Fs prediction. The networks consist of in, hn and on neurons. The architecture is illustrated in Fig. 1, where IW and LW refer to the connection weight matrices of hn $\times$ in and on $\times$ hn dimensions respectively, between the input-hidden and hidden-output layers. $\widehat{\rho}^1$ with in $\times$ 1 dimensions and $\widehat{\rho}^2$ with on $\times$ 1 dimensions refer to the vectors of bias units between the input-hidden and hidden-output layers.

In the experiments, the values of in and on are determined in a straightforward way as they are based on the numbers of input and output variables respectively. The values of hn and ANN's parameters are set based on the settings in Table S6. The adaptive SAEDE-EP is used to evolve the connection weights and bias between the input-hidden layers and hidden-output layers because it can adaptively adjust the settings of all parameters with minimum reliance on user-specified parameters.

A fixed architecture of ANN with its relevant parameters is initialised based on the settings in Table S6. An initial population of individuals representing the connection weights and bias is generated. Each individual solution is mapped into the ANN's initial connection weights and bias. With the initial connection weights and bias, the ANN is trained based on the k-fold validation for several runs. The ANN is evaluated based on MSE_{train} and $MSE_{validate}$. The fitness of each individual solution is represented by MSE_{train} and $MSE_{validate}$. The population of solutions underwent the iterative, evolutionary process until the stopping criterion was met. Fig. 4 summarises the flow of the optimisation of the ANN's connection weights and bias based on SAEDE-EP in the evolved ANN model.

4. Experimental setup

An ANN with the layers in, hn, and on neurons is used to predict PCDD/Fs emission. Each neural network consists of a MLP with a backpropagation algorithm. The ANNs are trained with a data set consisting of 8 input variables and 1 output variable. Details on the input and output variables are shown in Table S7. The input and output variables are in different units. Data for the PCDD/Fs emission prediction was normalised based on the variables' minimum and maximum in Table S7 and the normalised data was used to train the ANNs. The sample size of the PCDD/Fs emission data is limited, only 25 samples (Table S7).

Using the range of HN by Xu et al. (2021) as a reference, i.e., 4–20, the fixed architecture of ANNs that vary in HN are implemented in the experiments. The four architectures of ANNs consisting of 5, 10, 15, and 20 HN are used in the experiments. With one layer for the input, hidden and output neurons, the settings of the ANN's parameters in Table S6 are based on the work in (Ebrahimzade et al., 2020) as the work also adopted an EA, that is PSO to evolve the initial connection weights and it showed promising results. The learning rate for ANNs, η, is set as 0.1. Each connection weight of the neural network ranges in the interval [−1, 1]. The architectures and parameters of the ANNs in our experiments are summarised in Table S6.

The number of neurons in the input and output layer is fixed, namely 8 input neurons and 1 output neuron, respectively. The HN in the hidden layers varies among the ANNs. For a fixed architecture of ANN, e.g., 8 input neurons, 5 hidden neurons and 1 output neuron, each individual solution for the connection weights and bias generated by SAEDE-EP is fed to the ANN. Using the fixed parameters in Table S6, the ANN was trained for 1000 iterations. The MSEs of the training and k-fold

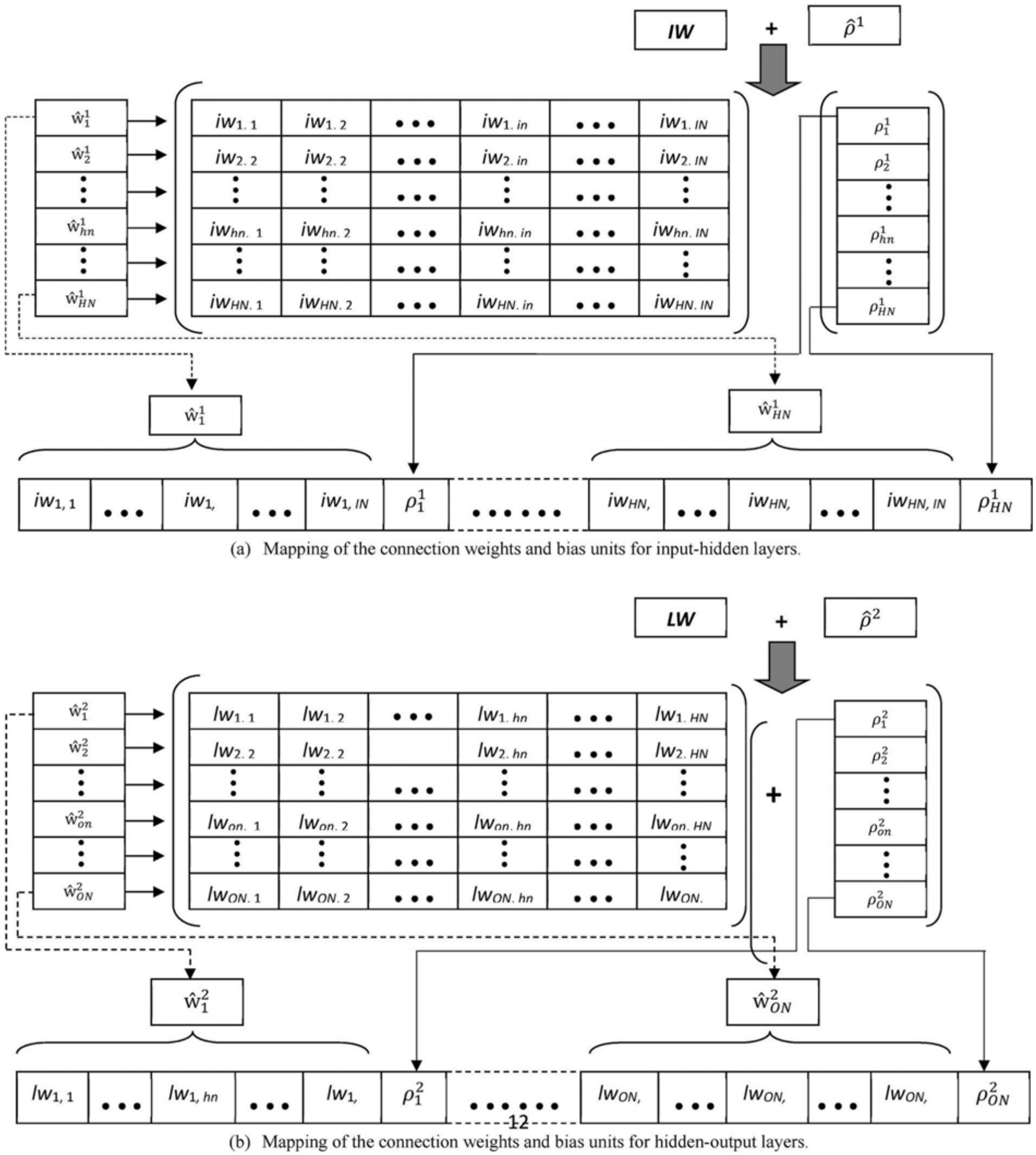


Fig. 3. Mapping of the connection weights and bias units for both input-hidden and hidden-output layers.

validation sets of the ANN for each run were evaluated. In this study, a 10-fold validation set is used. The MSEs of the training and 10-fold validation sets were used as the objective function for each individual solution in the population. The training and evaluation of the ANN are repeated for all individuals in the population. Hence, the fitness of each individual solution in the population is represented by MSE_{train} and $MSE_{validate}$. Based on the common ratio between training:validation:test, i.e., 70:15:15 (Pandey et al., 2016), the study divided 25 samples of data into training and validation sets consisting of 20 samples and the test set consisting of 5 samples. The 20 samples were used for the 10-fold cross validation. Thereafter, the ANN trained based on the 10-fold cross validation was tested on the test set. The population of solutions went through iterative mutation, crossover and selection until the stopping criterion was met. The optimisation by SAEDE-EP is to ascertain the optimal initial connection weights and bias for the ANN in the predictive stage.

The evolutions of SAEDE-EP to optimise each ANN's connection weights and bias are repeated by using the k-th seed numbers, $k = 1,2, \cdots,30$. The same 30 random seed numbers were used to evaluate the four ANN architectures in the experiments. Two stopping criteria were used in these experiments. First are the pre-specified threshold values (f_{target}) for comparison with f_{best}, and the other is the setting of maximum generation, G_{max}. The evolutionary processes are terminated if f_{best} is better than f_{target} (i.e., $f_{best} < f_{target}$). The f_{target} value is set as 0.000001 in these experiments. Or else, the evolutionary process continues until it reaches $G_{max} = 200$. The ANNs trained with the optimisation of SAEDE-EP are called evolved ANNs. Using the settings in Table S6 and the same 30 random seed numbers, the ANNs were trained without the

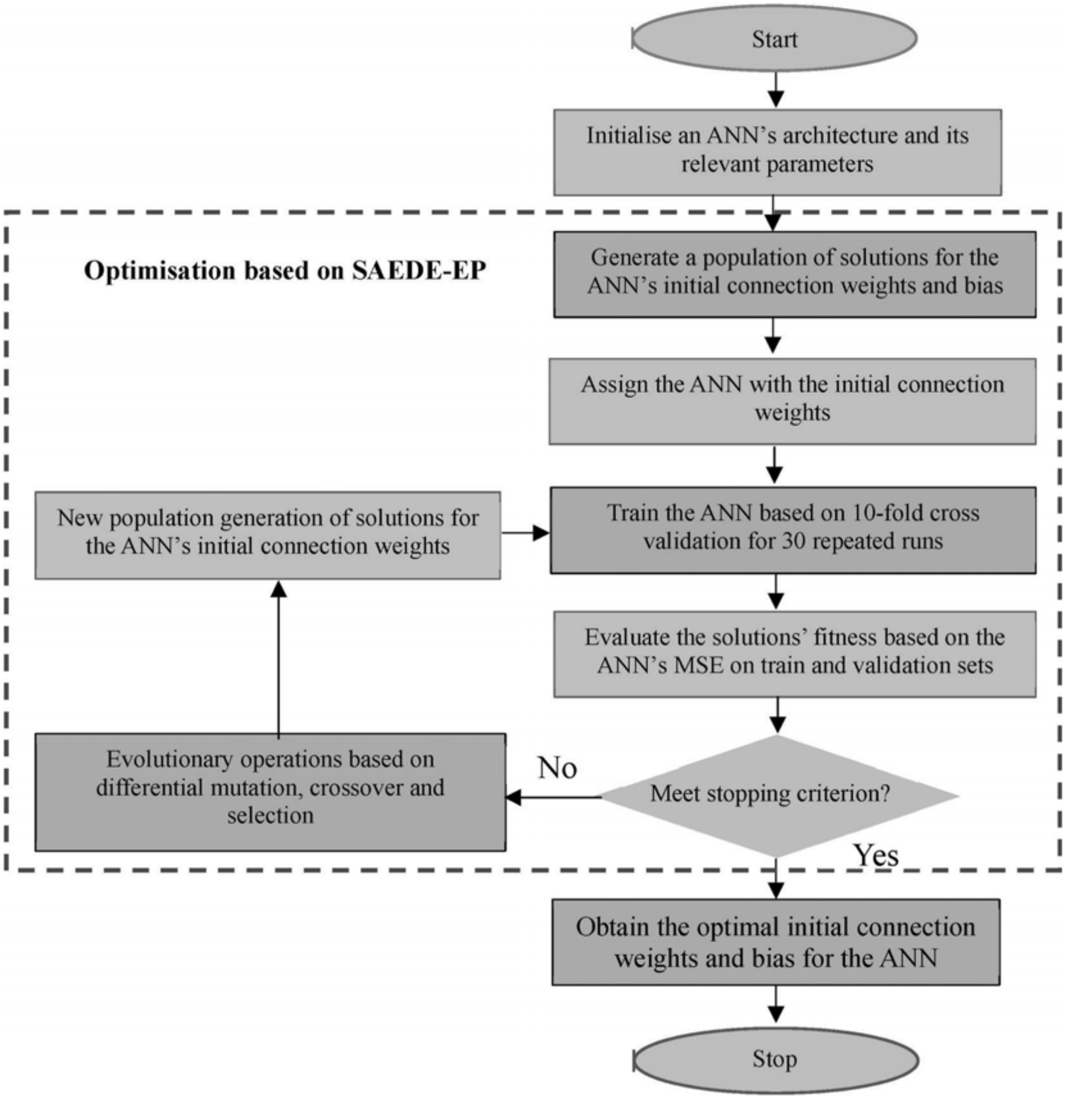


Fig. 4. Optimisation of an ANN's connection weights and bias based on SAEDE-EP in the evolved ANN model.

optimisation of SAEDE-EP and they are known as normal ANNs.

5. Result and analysis

The f_{best} of the SAEDE-EP for each architecture of ANN refers to the solutions producing the lowest MSE_{train} and $MSE_{validate}$. Comparisons of the five ANN architectures optimised by SAEDE-EP are evaluated based on f_{best} and the correlation coefficient, R^2. The calculation of R^2 is based on Eq. (4). R^2 shows how well the data fit the predicted model.

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(\psi_i^{target} - \psi_i^{predict}\right)^2}{\sum_{i=1}^{N}\left(\psi_i^{predict}\right)^2} \tag{4}$$

which, ψ_i^{target} is the target value, $\psi_i^{predict}$ is the model's predicted value and N is the data set number in the training and testing stages of each model.

Fig. 5 demonstrates the advancement of modelling by the SAEDE-EP algorithm in the training phase of ANN in the base of layer weights and bias and their performance measure trends. The performance measure trends are based on the f_{best}, R^2 and MSE_{test} trends.

The average of f_{best} and R^2 for each ANN model across 30 runs have been presented in the first two subplots in Fig. 5 to evaluate the modelling procedure and measure the effect of the connection weights and bias on SAEDE-EP-ANN optimisation. The ANN model with f_{best} closer to zero and R^2 closer to one has better performance within the proposed models. Fig. 5 shows that the ANNs with 15HN have the lowest f_{best}, $\leq 8.359 \times 10^{-4}$. On the other hand, the ANNs with 5HN, 10HN and 20HN have similar f_{best} upon the completion of evolution, 9.153×10^{-4} – 9.539×10^{-4}, as shown in the first row in Table 1. However, the difference of f_{best} between the three ANNs is very small, $\leq 3.853 \times 10^{-5}$.

The R^2 plot shows that the ANN with 5HN has the highest value ($R^2 = 0.974$) followed by 15HN ($R^2 = 0.971$), 10HN ($R^2 = 0.969$) and 20HN ($R^2 = 0.964$). The average MSE_{test} of the ANNs in Fig. 5 shows similar trends as R^2, with 5HN has the lowest value, followed by 15HN, 10HN and 20HN. The average measures show an interesting finding, which is the ANN with the lowest f_{best} does not necessarily guarantee the best result for MSE_{test} and R^2. The ANN of 15HN is associated with the lowest f_{best}, but it does not have the lowest MSE_{test} or highest R^2.

To understand the effect of the SAEDE-EP's optimisation on the ANNs, the best, worst and standard deviation based on f_{best}, R^2 and MSE_{test} for the ANNs across 30 runs are evaluated. The results based on the performance measures are summarised in Table 1 and the lowest f_{best} is highlighted in bold. The results in the table show that 15HN has the lowest f_{best} based on the mean, median and worst measurements, but it does not have the best R^2 and MSE_{test}. On the other hand, the ANN of

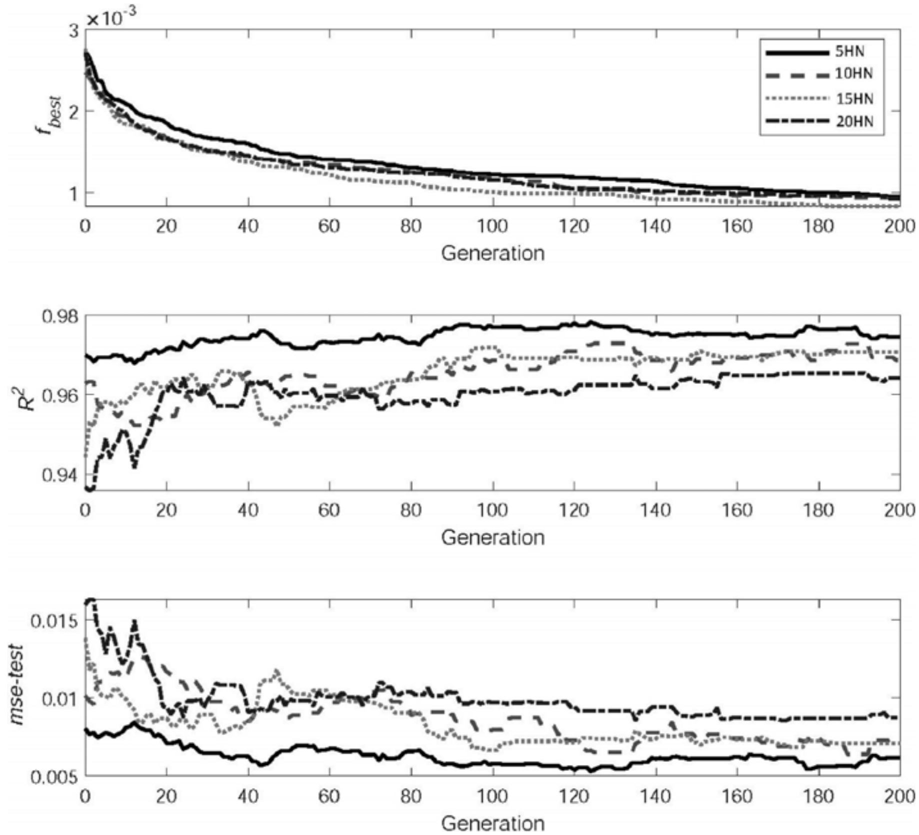


Fig. 5. The f_{best}, R^2 and MSE_{test} values of training procedures at different numbers of neurons.

Table 1
The statistics of f_{best}, R^2 and MSE_{test} for the evolved ANNs.

Statistics	Measures	Number of hidden neurons			
		5HN	10HN	15HN	20HN
Mean	f_{best}	9.5388 E−4	9.1534 E−4	**8.3588 E-4**	9.3077 E−4
	R^2	**0.97447**	0.95858	0.97056	0.96416
	MSE_{test}	**6.1790 E-3**	7.4638 E−3	7.1184 E−3	8.7541 E−3
Best	f_{best}	7.2681 E−4	6.1643 E−4	6.0598 E−4	**4.4683 E-4**
	R^2	0.99376	0.99275	0.98721	**0.99530**
	MSE_{test}	**8.5375 E-4**	1.3974 E−3	2.0929 E−3	1.1048 E−3
Median	f_{best}	9.7036 E−4	8.9204 E−4	**8.2563 E-4**	9.6100 E−4
	R^2	0.97688	**0.97794**	0.97709	0.96962
	MSE_{test}	**5.0682 E-3**	5.5349 E−3	5.7066 E−3	7.4753 E−3
Worst	f_{best}	1.3642 E−3	1.3880 E−3	**1.1129 E-3**	1.2547 E−3
	R^2	**0.91429**	0.87628	0.89743	0.90949
	MSE_{test}	**1.9399 E-2**	3.1526 E−2	2.8797 E−2	2.2074 E−2
Standard deviation	f_{best}	**1.5243 E-4**	1.8225 E−4	1.5736 E−4	1.7193 E−4
	R^2	**0.01463**	0.02455	0.01898	0.02151
	MSE_{test}	**3.7896 E-3**	6.4741 E−3	5.2479 E−3	5.3879 E−3

5HN has the best results for R^2 and MSE_{test} overall. Results in Table 1 consistently show that the ANN of 5HN produces the lowest MSE_{test} based on the measurements of mean, best, median, worst and standard deviation. In addition, the ANN of 5HN is associated with the highest R^2 based on the mean, worst and standard deviation. Based on the performance measures, it can be concluded that the ANN of 5HN is a better prediction model among the ANNs.

The best-trained ANNs with and without the SAEDE-EP optimisation are compared for different HN based on $MSE_{train\ +\ validate}$, MSE_{test} and R^2. The ANNs optimised by SAEDE are known as evolved ANNs, while those without SAEDE-EP optimisation are known as normal ANNs. All normal ANNs have the lower $MSE_{train\ +\ validate}$ than the evolved ANNs regardless of their HN but it does not guarantee a better MSE_{test}. Instead, the evolved ANNs produce lower MSE_{test} than the normal ANNs. The use of SAEDE-EP as the optimisation algorithm helped to enhance the ANN's generalisation ability even though it is trained with a limited sample size. As a result, the evolved ANNs have higher R^2 than the normal ANNs. The better measurements between the evolved and ANNs are highlighted in bold (Table S8).

The matrices of the initial connection weights and bias are extracted from the best of the evolved ANN of 5HN because it produces the lowest MSE_{test} (i.e., 8.53755×10^{-4}) and the second highest R^2 (i.e., 0.99376). The matrices of connection weights and bias of the ANN of 5HN are shown in Fig. 6(a). Based on the initial connection weights and bias, the ANN of 5HN was trained based on the parameters in Table S6. The optimal connection weights and bias of the feedforward ANN of 5HN upon the completion of the training are extracted and shown in Fig. 6(b).

6. Conclusion and future work

A hybrid machine learning algorithm comprising a combination of ANN and SADEE-EP is employed in this study to predict the PCDD/Fs emissions in peatlands using 25 data. The depth of peat soil, C:N ratio, chloride content, HA, phenol, 1,2,3-trichlorobenzene, Al and Cu as inputs and total concentration of PCDD/Fs as output were used by the evolved ANN to model PCDD/Fs emission. The PCDD/Fs modelling uses ANNs of different HN. The strength of the evolved ANN model lies in the adaptiveness of SAEDE-EP in setting its parameters corresponding to the

$$\underbrace{\begin{bmatrix} -0.807524 & -0.494569 & -0.249573 & -0.557478 & -0.058431 & -0.207669 & -0.213731 & -0.673466 \\ 0.707422 & -0.641411 & -0.200388 & -0.632463 & -0.643944 & 0.342955 & 0.138990 & 0.388621 \\ -0.520304 & -0.428086 & -0.582160 & -0.290244 & 0.174329 & -0.248413 & 0.869528 & -0.197394 \\ 0.996495 & 0.918593 & 0.845336 & 0.246928 & 0.947812 & -0.980271 & 0.533241 & -0.837455 \\ 0.329804 & 0.277834 & 0.484086 & -0.284412 & 0.610931 & -0.727427 & 0.390251 & -0.864486 \end{bmatrix}}_{IW} + \underbrace{\begin{bmatrix} -0.592069 \\ -0.065085 \\ 0.796647 \\ -0.226214 \\ 0.689490 \end{bmatrix}}_{\hat{\rho}^1}$$

Initial connection weights and biases for the input-hidden layer

$$\underbrace{\begin{bmatrix} -0.981077 & 0.823095 & 0.373286 & -0.799968 & 0.226525 \end{bmatrix}}_{LW} + \underbrace{\begin{bmatrix} 0.443994 \end{bmatrix}}_{\hat{\rho}^2}$$

Initial connection weights and bias for the hidden-output layer

(a) The matrices of the initial connection weights and bias of the evolved ANN of 5HN.

$$\underbrace{\begin{bmatrix} -0.641862906 & -0.410468929 & -0.501419388 & -0.738156502 & -0.232221600 & -0.099369686 & 0.146996613 & -0.655024639 \\ 0.459686024 & -0.606114533 & -0.382536783 & -0.634296708 & -0.698758186 & 0.276156583 & -0.045673562 & 0.102110496 \\ -0.630175968 & -0.426462634 & -0.558145396 & -0.095239183 & 0.103690064 & -0.373840285 & 0.715600317 & -0.296352347 \\ 1.915234349 & 1.079713726 & 1.013249579 & 0.564488371 & 1.178155635 & -1.411161685 & 1.202393992 & -0.620158535 \\ 0.300177637 & 0.322210897 & 0.497031277 & 0.001368343 & 0.555980298 & -0.784961885 & 0.256904673 & -0.919310157 \end{bmatrix}}_{IW} + \underbrace{\begin{bmatrix} -0.372468564 \\ -0.501889655 \\ 0.781341986 \\ 0.238199754 \\ 0.666824487 \end{bmatrix}}_{\hat{\rho}^1}$$

Connection weights and bias for the input-hidden layers

$$\underbrace{\begin{bmatrix} -1.058695545 & 0.113234855 & 0.093566860 & -1.461823246 & 0.201574408 \end{bmatrix}}_{LW} + \underbrace{\begin{bmatrix} 0.409684562 \end{bmatrix}}_{\hat{\rho}^2}$$

Connection weights and bias for the hidden-output layer

(b) The matrices of the optimal connection weights and bias of the feedforward ANN of 5HN.

Fig. 6. The matrices of the optimal connection weights and bias of the feedforward ANN of 5HN.

optimisation of the ANNs' connection weights and bias. With the assistance of SAEDE-EP to optimise the initial connection weights and bias, the evolved 5HN-ANN produces the lowest f_{best} based on the mean, worst and standard deviation, i.e., 0.97447, 0.91429 and 0.01463. It also has the lowest MSE_{test} based on the statistical measures. The comparison between the best evolved ANN and normal ANN of 5HN indicates that the evolved 5HN-ANN produces better MSE_{test} and R^2. In conclusion, the evolved ANN can assist in enhancing the prediction of PCDD/Fs emission with limited data samples and achieve better generalisation. The limitations of this work could be the accuracy and performance of the model are constrained by the narrow range of inputs used in its development. To improve its predictive capability, future research should focus on collecting more data on PCDD/Fs emissions from other peat samples, including those from temperate regions. Accordingly, the model for predicting PCDD/Fs levels in peat can be updated and re-trained to achieve more reliable results and a wider scope of application. Moreover, it would be worthwhile to compare the effectiveness of the existing methods with other models, such as SVM or DNN.

CRediT authorship contribution statement

Shir Li Wang: Writing – original draft, Software, Methodology, Conceptualization. **Theam Foo Ng:** Writing – review & editing, Software. **Khairulmazidah Mohamed:** Visualization, Investigation. **Sumayyah Dzulkifly:** Validation, Software. **Xiaodong Li:** Writing – review & editing, Resources. **Yin-Hui Leong:** Writing – review & editing, Writing – original draft, Project administration, 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.

Data availability

Data will be made available on request.

Acknowledgement

This project (IF1019J1112) is supported by the International Collaboration Fund (205.CRACUN.614201) from the Ministry of Science, Technology and Innovation (MOSTI) and Fundamental Research Grant Scheme (FRGS/1/2022/ICT02/UPSI/02/1) under the Ministry of Higher Education. The authors would like to thank Universiti Sains Malaysia and Universiti Pendidikan Sultan Idris, who assisted with managing this grant.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chemosphere.2024.142683.

References

Andriesse, J., 1988. Chemical properties of peat materials. Nature and management of tropical peat soils. FAO Soils Bull. 59 (Rome).

Bagheri, M., Mirbagheri, S.A., Bagheri, Z., Kamarkhani, A.M., 2015. Modeling and optimization of activated sludge bulking for a real wastewater treatment plant using hybrid artificial neural networks genetic algorithm approach. Process Saf. Environ. Protect. 95, 12–25.

Blumenstock, M., Zimmermann, R., Schramm, K.-W., Kaune, A., Nikolai, U., Lenoir, D., Kettrup, A., 1999. Estimation of the dioxin emission (PCDD/F I-TEQ) from the concentration of low chlorinated aromatic compounds in the flue and stack gas of a hazardous waste incinerator. J. Anal. Appl. Pyrolysis 49, 179–190.

Budiman, H., Wang, S.L., Morsidi, F., Ng, T.F., Neoh, S.C., 2020. Self-adaptive ensemble-based differential evolution with enhanced population sizing. In: 2020 2nd International Conference on Cybernetics and Intelligent System (ICORIS), pp. 1–6. https://doi.org/10.1109/ICORIS50180.2020.9320767.

Bunsan, S., Chen, W.-Y., Chen, H.-W., Chuang, Y.H., Grisdanurak, N., 2013. Modeling the dioxin emission of a municipal solid waste incinerator using neural networks. Chemosphere 92, 258–264.

Chang, N.-B., Chen, W., 2000. Prediction of PCDDs/PCDFs emissions from municipal incinerators by genetic programming and neural network modeling. Waste Manag. Res. 18, 341–351.

Choi, K.-I., Lee, D.-H., Osako, M., Kim, S.-C., 2007. The prediction of PCDD/DF levels in wet scrubbers associated with waste incinerators. Chemosphere 66, 1131–1137.

Ebrahimzade, H., Khayati, G.R., Schaffie, M., 2020. PSO–ANN-based prediction of cobalt leaching rate from waste lithium-ion batteries. J. Mater. Cycles Waste Manag. 22, 228–239.

Ferro, M.V., Mosquera, Y.D., Pena, F.J.R., Bilbao, V.M.D., 2023. Early stopping by correlating online indicators in neural networks. Neural Network. 159, 109–124.

IARC (International Agency for Research on Cancer), 1997. Polychlorinated dibenzo-paradioxins and polychlorinated dibenzofurans. IARC Monogr. Eval. Carcinog. Risks Hum. 69. Available from: https://www.ncbi.nlm.nih.gov/books/NBK409980/. (Accessed 21 February 2024).

Javanshah, A., Saidi, A., 2016. Determination of humic acid by spectrophotometric analysis in the soils. Int. J. Adv. Biotechnol. Res. 7, 19–23.

Könönen, M., Jauhiainen, J., Laiho, R., Kusin, K., Vasander, H., 2015. Physical and chemical properties of tropical peat under stabilised land uses. Mires Peat 16, 1–13.

LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R., Hubbard, W.E., Jackel, L., Touretzky, D., 1990. Advances in Neural Information Processing Systems. Morgan Kaufmann Publishers Inc, San Francisco, CA, USA, pp. 396–404.

Li, Y., Wang, S., Yang, B., Chen, H., Wu, Z., Yang, H., 2023. Population reduction with individual similarity for differential evolution. Artif. Intell. Rev. 56, 3887–3949.

Lu, W., Peng, Y., Chen, X., Skitmore, M., Zhang, X., 2016. The S-curve for forecasting waste generation in construction projects. Waste Manag. 56, 23–34.

Nolfi, S., Floreano, D., 2000. Evolutionary Robotics: the Biology, Intelligence, and Technology of Self-Organizing Machines. MIT Press.

Oliveira, V., Sousa, V., Dias-Ferreira, C., 2019. Artificial neural network modelling of the amount of separately-collected household packaging waste. J. Clean. Prod. 210, 401–409.

Pandey, D.S., Das, S., Pan, I., Leahy, J.J., Kwapinski, W., 2016. Artificial neural network based modelling approach for municipal solid waste gasification in a fluidized bed reactor. Waste Manag. 58, 202–213.

Parouha, R.P., Verma, P., 2022. A systematic overview of developments in differential evolution and particle swarm optimization with their advanced suggestion. Appl. Intell. 1–45.

Qiao, J., Guo, Z., Tang, J., 2021. Soft sensing of dioxin emission concentration in solid waste incineration process based on multi-layer feature selection. Inf. Control 50, 75–87.

Rein, G., 2016. Smoldering Combustion. Springer, New York, New York, pp. 581–603.

Soni, U., Roy, A., Verma, A., Jain, V., 2019. Forecasting municipal solid waste generation using artificial intelligence models—a case study in India. SN Appl. Sci. 1, 1–10.

Storn, R., Price, K., 1997. Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. J. Global Optim. 11, 341–359.

Sutejo, Y., Saggaff, A., Rahayu, W., Hanafiah, H., 2017. Physical and chemical characteristics of fibrous peat. AIP Conf. Proc. 1903 (AIP Publishing).

Tang, J., Guo, Z., Qiao, J., Xu, Z., 2019. Dioxin emission concentration forecasting approach based on latent feature extraction and selection for municipal solid waste incineration. In: 2019 Chinese Control Conference (CCC), pp. 6845–6850. https://doi.org/10.23919/ChiCC.2019.8865796.

Tang, J., Qiao, J., Xu, Z., Guo, Z., 2021a. Soft measuring approach of dioxin emission concentration in municipal solid waste incineration process based on feature reduction and selective ensemble algorithm. Control Theory & Appl. 38, 110–120.

Tang, J., Xia, H., Zhang, J., Qiao, J., Yu, W., 2021b. Deep forest regression based on cross-layer full connection. Neural Comput. Appl. 33, 9307–9328.

Uloth, V., Duo, W., Leclerc, D., Karidio, I., Kish, J., Singbeil, D., 2005. Investigations into the variability and control of dioxins formation and emissions from coastal power boilers. In: Proceedings of the TAPPI Pulping, Engineering and Environmental Conference (Philadelphia, PA).

USEPA (United States Environmental Protection Agency), 1994. Tetra-through Octa-Chlorinated Dioxins and Furans by Isotope Dilution HRGC/HRMS.

Wang, H., Zhang, Y., Wang, H., 2008. A study of GA-BP based prediction model of dioxin emission from MSW incinerator. Microcomput. Inf. 24, 222–224.

Xia, H., Tang, J., Aljerf, L., 2022. Dioxin emission prediction based on improved deep forest regression for municipal solid waste incineration process. Chemosphere 294, 133716.

Xia, H., Tang, J., Qiao, J., Yan, A., Guo, Z., 2020. Soft measuring method of dioxin emission concentration for MSWI process based on RF and GBDT. In: 2020 32nd Chinese Control and Decision Conference (CCDC). IEEE, pp. 2173–2178. https://doi.org/10.1109/CCDC49329.2020.9164125.

Xiao, X., Lu, J., Hai, J., 2017. Support vector regression prediction of dioxins concentration in waste incineration flue gas. Renew. Energy 35, 1107–1114.

Xu, A., Chang, H., Xu, Y., Li, R., Li, X., Zhao, Y., 2021. Applying artificial neural networks (ANNs) to solve solid waste-related issues: a critical review. Waste Manag. 124, 385–402.

Yang, X.-S., 2020. Nature-inspired optimization algorithms: challenges and open problems. J. Comput. Sci. 46, 101104.

Ying, Y., Ma, Y., Wang, X., Yu, H., Wu, A., Lu, S., Li, X., Lin, X., Liu, L., Yin, W., 2023. Emission, partition, and formation pathway of polychlorinated dibenzo-p-dioxins and dibenzofurans during co-disposal of industrial waste with municipal solid waste. J. Environ. Chem. Eng. 11, 109242.