



A hybrid novel SVM model for predicting CO₂ emissions using Multiobjective Seagull Optimization

Mohammad Ehteram¹ · Saad Sh. Sammen² · Fatemeh Panahi³ · Lariyah Mohd Sidek⁴

Received: 7 April 2021 / Accepted: 27 June 2021 / Published online: 30 July 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

The agricultural sector is one of the most important sources of CO₂ emissions. Thus, the current study predicted CO₂ emissions based on data from the agricultural sectors of 25 provinces in Iran. The gross domestic product (GDP), the square of the GDP (GDP²), energy use, and income inequality (Gini index) were used as the inputs. The study used support vector machine (SVM) models to predict CO₂ emissions. Multiobjective algorithms (MOAs), such as the seagull optimization algorithm (MOSOA), salp swarm algorithm (MOSSA), bat algorithm (MOBA), and particle swarm optimization (MOPSO) algorithm, were used to perform three important tasks for improving the SVM models. Additionally, an inclusive multiple model (IMM) used the outputs of the MOSOA, MOSSA, MOBA, and MOPSO algorithms as the inputs for predicting CO₂ emissions. It was observed that the best kernel function based on the SVM-MOSOA was the radial function. Additionally, the best input combination used all the gross domestic product (GDP), squared GDP (GDP²), energy use, and income inequality (Gini index) inputs. The results indicated that the quality of the obtained Pareto front based on the MOSOA was better than those of the other algorithms. Regarding the obtained results, the IMM model decreased the mean absolute errors of the SVM-MOSOA, SVM-MOSSA, SVM-MOBA, and SVM-PSO models by 24, 31, 69, and 76%, respectively, during the training stage. The current study showed that the IMM model was the best model for predicting CO₂ emissions.

Keywords CO₂ emissions · SVM · Multiobjective algorithms · Environment management

Introduction

Environmental degradation is one of the main challenges faced by policymakers and decision-makers. The increasing emission rate of CO₂ into the atmosphere is one of the important causes of environmental degradation. Increasing industrialization and urbanization are the other causes of CO₂

emissions in developing countries (Zhao et al. 2018). The increasing levels of CO₂ emissions affect global warming and climate change. Additionally, ocean acidification and desertification are the other consequences of the emission of CO₂ into the atmosphere. Pollution levels and related diseases are increased by the emission of CO₂. Hence, CO₂ emissions affect human health. Thus, predicting CO₂ emissions is one of the most important issues for researchers. The agricultural sector is one of the most important sources of CO₂ emissions. In 2018, 9.9% of greenhouse gas emissions were related to the agricultural sector. Cows, agricultural soils, and rice production can increase CO₂ emissions. The agricultural sector causes 10–14% of global anthropogenic greenhouse gas emissions (Shabani et al. 2021). If decision-makers want to manage CO₂ emissions, it is necessary to accurately estimate the CO₂ emissions caused by different sectors, such as the agricultural sector. Predictive models utilize parameters that are relevant to CO₂ emissions to estimate the amounts of emitted CO₂ in different years. Shi et al. (2019a) stated that the CO₂ emission depends on different parameters such as industrial activity and energy intensity. Shi et al. (2019b) stated that the

Responsible Editor: Marcus Schulz

✉ Saad Sh. Sammen
Saad1234engineer@gmail.com

¹ Department of Water Engineering and Hydraulic Structures, Faculty of Civil 4 Engineering, Semnan University, Semnan, Iran

² Department of Civil Engineering, College of Engineering, University of Diyala, Baqubah, Diyala Governorate, Iraq

³ Faculty of Natural Resources and Earth Sciences, University of Kashan, Kashan, Iran

⁴ Institute of Energy Infrastructure, Universiti Tenaga Nasional, Jalan IKRAM-UNITEN, 43000 Kajang, Selangor, Malaysia

household sector of china is responsible for 12.6% of CO₂ emissions.

In recent years, machine learning algorithms (MLAs) have been widely used for predicting different variables, such as climate variables, pollutants, gas emissions, and hydrological variables (Banadkooki et al. 2020a). The advantages of MLAs include their handling of unlimited input data, fast processing speed, and accurate predictions. In the context of CO₂ emissions, researchers have used different MLAs to predict the CO₂ emission and greenhouse gases (GG). Table 1 reviews CO₂ emission forecasting approaches. As observed in Table 1, the MLAs have the high ability for predicting CO₂ emissions. SVM is a family member of MLAs that are widely used for predicting target variables. The SVM models have the high ability in high-dimensional space. The modelers can define different kernel functions depending upon their requirement. The SVM models provide good generalization capability. Also, they can reduce the computational complexity. Saleh et al. (2016) investigated the ability of SVM model in predicting CO₂ emission. The data of electrical energy and burning coal were used as the inputs to the models. They used trial and error to adjust the parameters of the SVM model. It was concluded that the SVM model was effective for predicting CO₂ emission. Sun and Liu (2016) exploited the least square SVM (LSSVM) model to predict different kinds of CO₂ emission. They concluded that the classification of CO₂ emission enhanced forecast accuracy. Ahmadi et al. (2019a) explored the use of least square SVM (LSSVM) in predicting CO₂ emission. They used evolutionary algorithms to train the SVM model. They coupled genetic algorithm (GA) with particle swarm optimization (PSO) to make a new hybrid algorithm for training SVM model. It was concluded that the LSSVM-PSO-GA had better efficiency in predicting CO₂ emission compared with LSSVM-GA and LSSVM-PSO models. Wu and Meng (2020) coupled LSSVM with bat algorithm (BA) for predicting CO₂ emission. It was observed that the LSSVM-BA outperformed the extreme learning model and backpropagation neural network.

Generally, MLAs have strong prediction abilities with respect to CO₂ emissions, but there are challenges:

- 1- The model parameters of MLAs need to be tuned based on powerful training algorithms such as advanced optimization algorithms (Ehteram et al. 2020; Banadkooki et al. 2021).
- 2- Some MLAs, such as the ANN, SVM, and ANFIS models, have different kinds of kernel functions and activation functions. Thus, the best function should be selected to predict the target variable based on the received input data (Darabi et al. 2021).

Table 1 Detailed paper reviewed

Author	Description	Results
Khoshnevisan et al. (2013)	Khoshnevisan et al. (2013) predicted GGs using artificial neural networks (ANNs) and energy inputs	The results indicated that the ANN model with a coefficient of determination of 0.99 had a strong ability to predict CO ₂ emissions
Wang et al. (2015)	Wang et al. (2015) used a principal component analysis-support vector machine (PCA-SVM) for predicting CO ₂ emissions	They reported that the PCA-SVM based on the radial basis function provided accurate results for predicting CO ₂ emissions
Sun and Liu (2016)	Sun and Liu (2016) used the least-squares support vector method (LSSVM), an ANN, and a gray model for predicting CO ₂ emissions	It was found that the LLSVM model outperformed the other models.
Sun et al. (2017)	Sun et al. (2017) used an extreme learning machine model (ELM) and a particle swarm optimization-ELM (PSO-ELM) for predicting CO ₂ emissions	They used factor analysis to identify the appropriate input parameters. The use of PSO as a training algorithm improved the accuracy of the ELM when predicting CO ₂ emissions
Li et al. (2017)	Li et al. (2017) applied the SVM genetic algorithm for predicting CO ₂ emissions	The results revealed that economic growth, resident population growth, and energy intensity are the most important parameters affecting CO ₂ emissions. The GA could also improve the accuracy of the standalone SVM model
Zhao et al. (2018)	Zhao et al. (2018) used the LLSVM-salp swarm algorithm (LLSVM-SSA) for predicting CO ₂ emissions. Different parameters, such as the population, energy consumption, energy intensity, and gross domestic product, were used as the inputs for the models	It was concluded that the LLSVM-SSA outperformed the LLSVM model in terms of predicting CO ₂ emissions
Ahmadi et al. (2019a)	Ahmadi et al. (2019b) used the group method of data handling (GMDH) to predict GG emissions. Iran, Kuwait, Qatar, Saudi Arabia, and the United Arab Emirates were considered in case studies for predicting	It was reported that GMDH was a powerful method for predicting CO ₂ emissions

Table 1 (continued)

	CO ₂ emissions.	
Xu et al. (2019)	Xu et al. (2019) applied a dynamic nonlinear ANN to predict CO ₂ emissions.	It was found that the dynamic nonlinear ANN had a strong ability to predict CO ₂ emissions.
Duan and Luo (2020)	Duan and Luo (2020) coupled the gray Verhulst model with particle swarm optimization (GVMPPO) for predicting CO ₂ emissions.	The results demonstrated that GVMPPO could predict CO ₂ emissions accurately for problems with a limited number of data

- 3- The selection of the best input scenario for each MLA requires preprocessing methods.
- 4- Previous studies compared various models and determined a superior model for predicting CO₂ emissions. In fact, competing models are rejected or accepted based on their accuracy, but the main question is how the synergy among multiple models can be used.

To address the abovementioned challenges, the current study uses a new hybrid framework for predicting CO₂ emissions. One of the most powerful models for predicting target variables such as CO₂ emissions is the SVM. The kernel functions of the SVM have the same parameters. The values of the SVM parameters can be obtained based on robust optimization algorithms. In this study, four multiobjective algorithms (MOAs), namely, the MO seagull optimization algorithm (MOSOA), MO salp swarm algorithm (MOSSA), MO bat algorithm (MOBA), and MO particle swarm optimization (MOPSO), are used to improve the performance of the SVM model and address the challenges mentioned above:

- 1- To obtain the values of the SVM parameters, an objective function, such as the root mean square error (RMSE), is used as the first objective function.
- 2- To choose the best kernel function, the names of kernel functions are inserted into the optimization algorithms as decision variables. A second objective function, such as the mean square error (MAE), is used to choose the best kernel function.
- 3- The names of the input variables are inserted into the algorithms as the decision variables. The Nash Sutcliffe efficiency is used as the third objective function for choosing the name each of input variables.
- 4- An inclusive multiple model is used to predict CO₂ emissions based on the synergy among the SVM-MOSOA, SVM-MOSSA, SVM-MOBA, and SVM-MOPSO algorithm.

Regarding the abovementioned points, the main novelties of the current paper are as follows:

- 1- The establishment of new hybrid SVM models that were not used in previous articles for predicting CO₂ emissions.
- 2- The creation of an inclusive multiple model to use the contributions of different SVM models for predicting CO₂ emissions.
- 3- The application of the current study is not limited to CO₂ emissions, and modelers can use the proposed models for predicting other variables, such as hydrological variables.
- 4- The presentation of an effective approach for finding the best values of the random parameters of the compared multiobjective algorithms.

To the best of the authors’ knowledge, no previous article has investigated the new hybrid SVM models proposed in the current study for predicting CO₂ emissions. In this study, new hybrid SVM models are used to predict CO₂ emissions in the agricultural sector of Iran based on data from 20 provinces (Fig. 1). Section 2 of the current study explains the structures of the compared models and the methods utilized. The case study is explained in Section 3. Section 4 presents a discussion and the experimental results. Section 5 explains the conclusions of the paper.

Materials and methods

Structure of the support vector machine

The first version of the SVM was introduced by Sain and Vapnik (1996). The SVM model uses a kernel function to find the relationships between the model inputs and outputs. The linear form of the SVM is as follows:

$$f(x) = \eta^{Tr} \cdot x + \beta \tag{1}$$

where x is the input, Tr denotes the transpose operation, η denotes the weighting coefficients of the input variables, β is the bias, and $f(x)$ is the variable predicted by the SVM.

The aim of the SVM is to minimize the difference between the predicted values and observed values. Thus, an optimization problem is defined to minimize the error function, which is named the e-insensitive loss function. The SVM acts based on the following equations:

$$\text{Minimize } \frac{1}{2} \|\eta\|^2 + PE \sum_{i=1}^m (\psi_i^- + \psi_i^+) \tag{2}$$

$$\text{subject (to)} (\eta_i \cdot x_i + \beta) - z_i < \varepsilon + \psi_i^+, i = 1, 2, \dots, m z_i - (\eta_i \cdot x_i + \beta) \leq \varepsilon + \psi_i^- \tag{3}$$

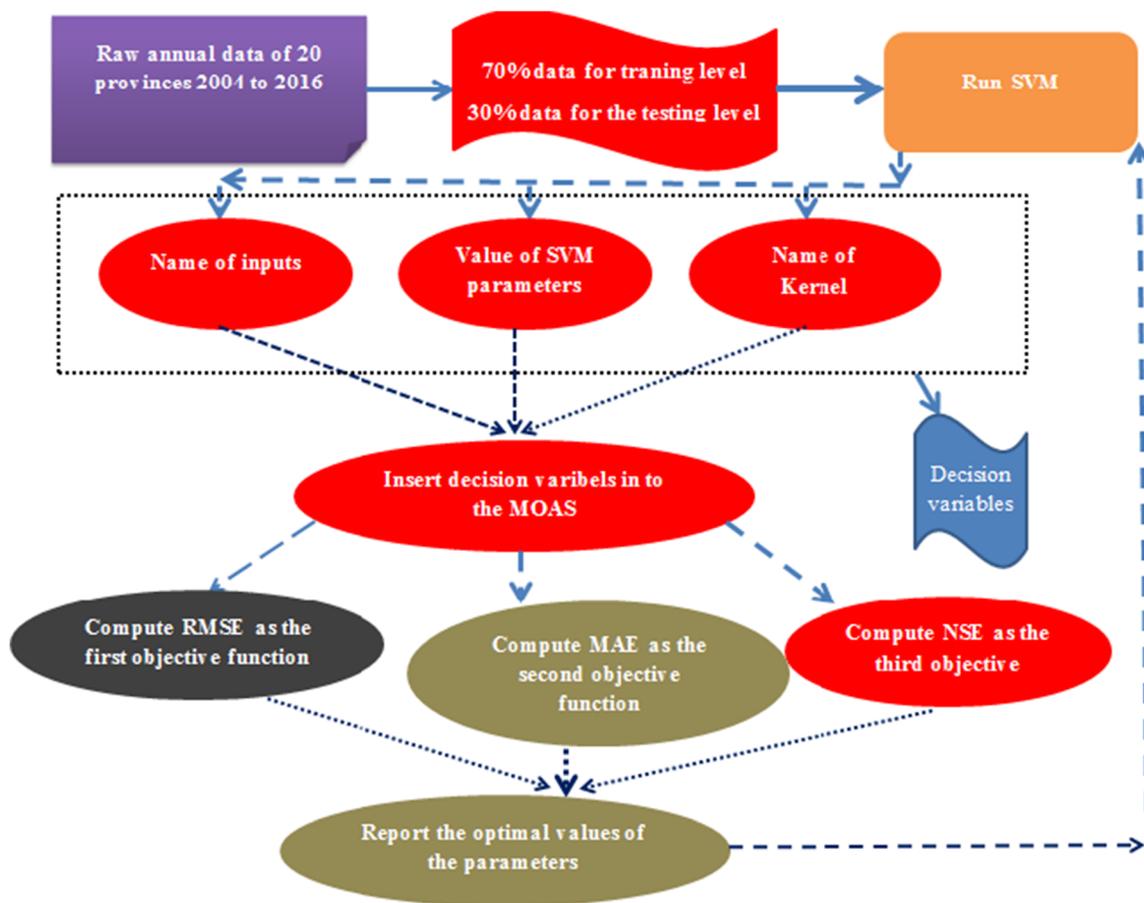


Fig. 1 Methodology flowchart

where PE is the penalty coefficient, m is the number of observed data, ψ_i^- and ψ_i^+ represent the violation of the i th training data ε : permitted error threshold, x_i is the input variable, z_i is the target variable, and η_i is the i th weight variable. The weight variable and bias are obtained based on Eqs. 2 and 3. Then, they are inserted into Eq. 1 to obtain $f(x)$. The SVM uses several kernel functions to map the dataset to the linearly separable space.

$$f(x) = \eta^{Tr} \cdot K(x, x_i) + \beta \tag{4}$$

Sigmoid function:

$$K(x, x_i) = \tanh(\gamma(x \cdot x_i) + r) \tag{5}$$

Radial basis function (RBF):

$$K(x, x_i) = \exp\left(-\gamma \|x - x_i\|^2 + PE\right) \tag{6}$$

Polynomial function:

$$K(x, x_i) = (\gamma(x \cdot x_i) + r)^d \tag{7}$$

where $K(x, x_i)$ is the kernel function and γ , r , PE , d , and ε are kernel parameters (the values of the kernel parameters are

obtained based on the corresponding MOAs).

Seagull optimization algorithm

The SOA has been widely applied in different fields, with applications such as multiobjective optimization (Dhiman et al. 2020; Dhiman and Kumar 2019), feature selection (Jia et al. 2019), experimental fuel cell modeling (Cao et al. 2019), and classification (Jiang et al. 2020). Few parameters, easy implementation, fast convergence, and the use of swarm experience are the advantages of the SOA. The migrating birds are the prey of the seagulls. The seagulls use natural spiral-shaped movement to attack migrating birds. Seagulls live in a group and travel towards the direction of the seagull that is most fit for survival. The mathematical model of the SOA is based on the migration phase and attack phase:

(1) Migration phase

When the seagulls move in the search space, collisions between neighbors should be avoided. Thus, the new locations of the search agents (seagulls) are computed as follows:

$$\vec{P}_s = \alpha \times \vec{L}_s(x) \tag{8}$$

where \vec{P}_s is the location of the search agent that prevents collisions with the other seagulls, $\vec{L}_s(x)$ is the location of the seagull during iteration (i), and α is the motion behavior of the seagull. α is computed as follows:

$$\alpha = f_c - \left(i \times \left(\frac{f_c}{\text{Max}(i)} \right) \right) \tag{9}$$

where f_c is the frequency control of α and i is the number of iterations. In the next phase, the seagulls move towards the direction of the best neighbor:

$$\vec{N}_e = \zeta \times (\vec{P}_b - \vec{L}_s) \tag{10}$$

where \vec{N}_e is the location of seagull \vec{L}_s (moving towards the best seagull \vec{P}_b) and ζ is a random value. A random value is obtained based on the following equation:

$$\zeta = 2 \times \alpha^2 \times RA \tag{11}$$

where RA is a random number. Finally, the seagulls update their locations as follows:

$$\vec{F}_s = |\vec{P}_s + \vec{N}_s| \tag{12}$$

where \vec{F}_s is the distance between the search agent and best seagull.

(2) Attack phase

During the attack, seagulls maintain their altitude using their wings and weight. The speed and angle of the seagull attack are changed continuously. They use spiral movements to attack their prey, as follows:

$$xx' = r \times \cos(\tau) \tag{13}$$

$$yy' = r \times \sin(\tau) \tag{14}$$

$$zz' = r \times \tau \tag{15}$$

$$r = u \times e^{\tau v} \tag{16}$$

where r is the radius of each spiral turn, xx' is the position of the seagull in plan x, yy' is the position of the seagull in plan y, zz' is the location of the seagull in plan z, u and v are constants, and τ is a random number. Finally, the seagulls change their locations as follows:

$$\vec{L}_s = (\vec{F}_s \times xx' \times yy' \times zz') + \vec{P}_b \tag{17}$$

where \vec{L}_s stores the best solution. Figure 2 shows the flow-chart of the SOA.

Salp swarm algorithm

The SSA is widely applied for optimization problems in different fields, such as feature selection (Tubishat et al. 2021), engineering optimization (Salgotra et al. 2021), optimal power flow calculation (Abd el-sattar et al. 2021), and ANN training (Kandiri et al. 2020). The SSA has a good balance between exploration and exploitation, as well as fast convergence. The positions of salps in an optimization problem signify a candidate solution. The first salp at the front of the salp chain is called the leader, and the other salps are called followers. Figure 3 shows a salp chain. The position of the leader is changed as follows:

$$\text{Salp}_{1,j} = \begin{cases} \text{Food}_j + \rho_1 \times \left((up_j - low_j) \times \rho_2 + low_j \right) & \leftarrow \rho_3 \geq 0.50 \\ \text{Food}_j - \rho_1 \times \left((up_j - low_j) \times \rho_2 + low_j \right) & \leftarrow \rho_3 < 0.50 \end{cases} \tag{18}$$

where $\text{Salp}_{1,j}$ is the location of leader, Food_j is the food source in the j th dimension, ρ_1 is a control parameter, and ρ_3 and ρ_2 are random parameters. ρ_1 is updated as follows:

$$\rho_1 = 2 \times e^{\left(\frac{-4 \times l}{L} \right)^2} \tag{19}$$

where L is the maximum number of iterations and l is the current number of iterations. The locations of the followers are updated as follows:

$$fol_{i,j} = \frac{1}{2} (fol_{i,j} + fol_{i-1,j}) \tag{21}$$

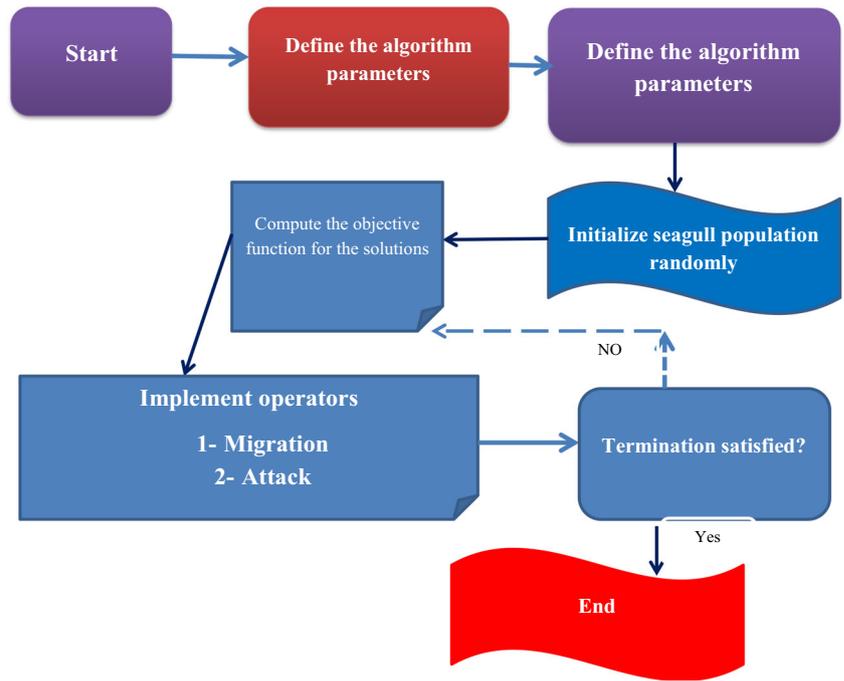
where $fol_{i,j}$ is the position of the i th follower in the j th dimension. Figure 4 shows the SSA flowchart.

Bat algorithm

The BA was inspired by the behavior of bats when finding food. The BA has been applied in different fields, such as image segmentation (Yue and Zhang 2020), parameter extraction for photovoltaic models (Deotti et al. 2020), MLA training (Dong et al. 2020), optimal reactive power dispatching (Mugemanyi et al. 2020), continuous optimization (Chakri et al. 2017), and numerical optimization (Wang et al. 2019). The bats use echolocation behavior to distinguish food from obstacles. A bat adjusts its distance from food based on its wavelength, frequency, and pulsation rate. The bats update their frequencies, velocities, and locations as follows:

$$fre_i = fre_{\min} + r_1 (fre_{\max} - fre_{\min}) \tag{22}$$

Fig. 2 The flowchart of the different optimization levels of SOA



$$ve_{t+1}^i = ve_t^i + f^i(X_t^i - X_t^{best}) \tag{23}$$

$$X_i^{t+1} = X_i^t + ve_i^{t+1} \tag{24}$$

where fre_i is the frequency of the i th bat, fre_{min} is the minimum frequency, r_1 is a random value, ve_{t+1}^i is the velocity of the i th bat at iteration $t+1$, X_t^i is the location of the i th bat at iteration t , and X_t^{best} is the location of the best bat (the best solution). The bats use a random walk to perform the local search operation:

$$X_i^{t+1} = X_i^t + \varphi A_i^{t+1} \tag{25}$$

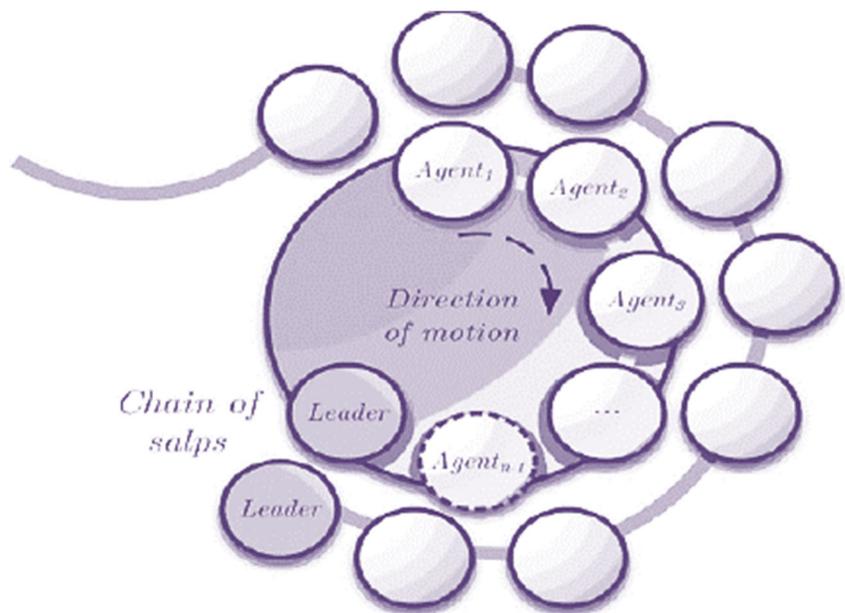
where φ is a random number and A_i^{t+1} is a loudness parameter. The pulsation rate and loudness are adjusted based on the following equations:

$$A_i^{t+1} = \mu A_i^t \tag{26}$$

$$r_i^{t+1} = r_i(0)[1 - \exp(-\gamma t)] \tag{27}$$

where μ and γ are constant values, $r_i(0)$ is the initial value of the pulsation rate, A_i^{t+1} is the value of loudness at iteration $t+1$, and r_i^{t+1} is the pulsation rate at iteration $t+1$. Figure 5 shows the BA flowchart.

Fig. 3 The schematic structure of the salp chain (Zhang et al. 2021b)



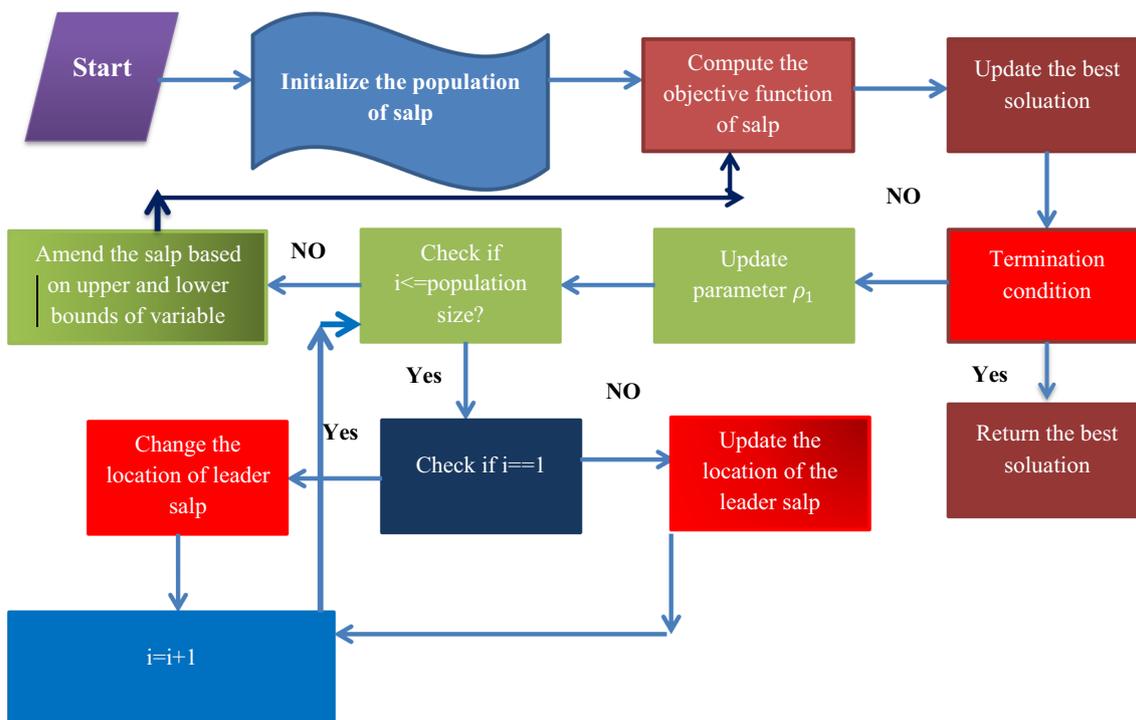


Fig. 4 The flowchart of the SSA

Particle swarm optimization

PSO is a powerful optimization algorithm that is widely used in different fields, such as for multiobjective optimization problems (Zhang et al. 2020), feature selection problems (El-Kenawy and Eid 2020), environmental economic dispatch problems (Xin-gang et al. 2020), green coal production

problems (Cui et al. 2020), and the training of ANN models (Darwish et al. 2020). The collaborative behavior of the swarm in the PSO algorithm is one of the advantages of PSO with regard to finding optimal solutions. In PSO, first, the random positions and velocities of the particles are initialized. Then, an objective function is computed for each particle so that the best particle position achieved by the population so

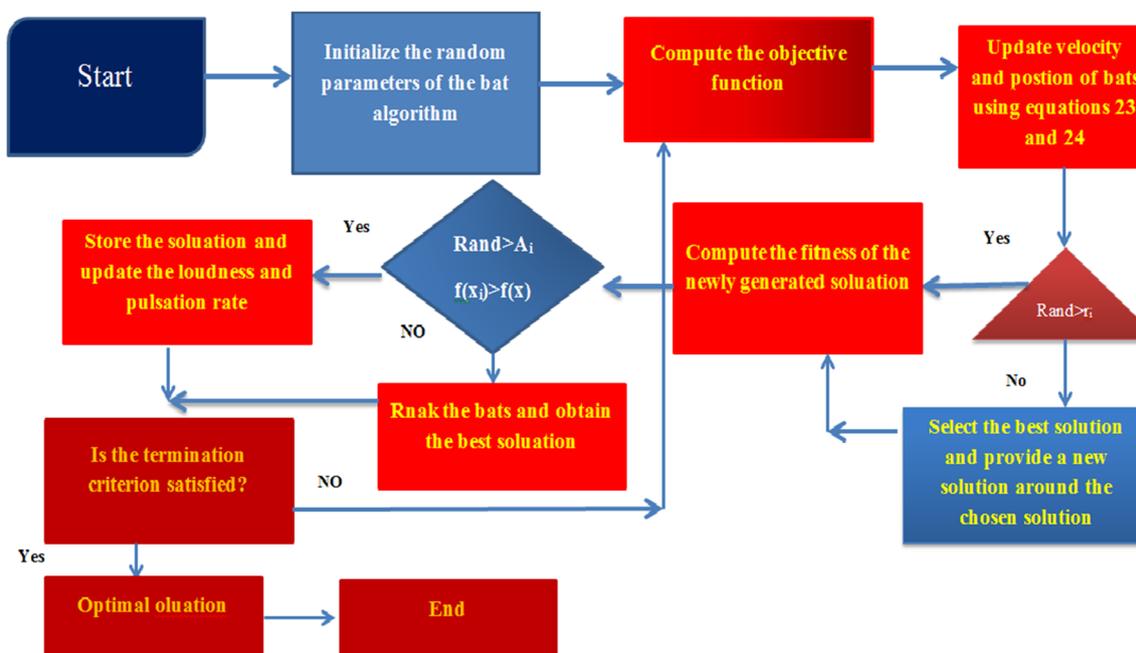


Fig. 5 The flowchart of the BA

far is identified. The velocities and positions of the particles are updated based on the following equations. The process continues until the stopping criterion is satisfied.

$$v_{i,d}(t+1) = \kappa v_{i,d}(t) + c_1 r_1 (p_{best,i,d} - x_{i,d}(t)) + c_2 r_2 (g_{best,d} - x_{i,d}(t)) \quad (28)$$

$$x_{i,d}(t+1) = x_{i,d}(t) + v_{i,d}(t) \quad (29)$$

where $v_{i,d}(t+1)$ is the velocity of the i th particle in the d th dimension during iteration $t+1$, c_1 and c_2 are acceleration coefficients, $p_{best,i,d}$ is the personal best position, $g_{best,d}$ is the global best position, $x_{i,d}(t+1)$ is the position of the i th particle during iteration $t+1$, and r_1 and r_2 are random numbers.

Multiobjective optimization problems

In a multiobjective optimization problem (MOOP), some objective functions can conflict with each other. A solution to an MOOP cannot be compared with other solutions based on relational operators. One of the most important conceptions in MOOPs is Pareto dominance. Solution A dominates solution B if it has equal values (with at least one better value) on all objectives. The circles in Fig. 6 are better than the squares in Fig. 6 based on the fact that the circles achieve lower objective function values of for a minimization problem with the aim of minimizing both objectives. While the circles dominate the squares, they do not dominate each other. Each MOOP has a set of best non-dominated solutions, namely, the Pareto optimal set. The projection of Pareto optimal solutions in the search space is known as the Pareto optimal front. In an MOOP, there is an external archive in which the non-dominated solutions of MOAs are stored. The next challenge is the selection of a target for each iteration of the algorithm. In fact, the targets are the best positions of the leader, bat, particle, and seagull for the MOSSA, MOBA, MOPSO algorithm,

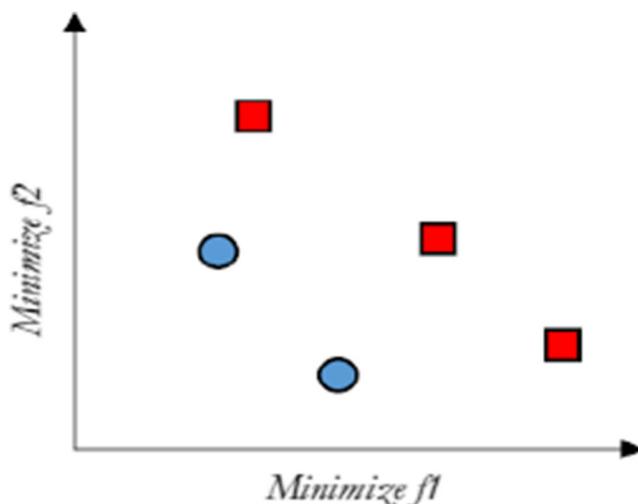


Fig. 6 Dominance conception

and MOSEOA, respectively. To find the target for updating the positions of other agents, the number of neighboring solutions (NSs) in the neighborhood of each solution is counted. During this phase, the target should be chosen from the set of non-dominated solutions with the least crowded neighborhood. To determine the populated neighborhood, NSs within a certain maximum distance are counted.

$$\vec{d} = \frac{\max \vec{x} - \min \vec{n}}{(\text{Archive})\text{size}} \quad (30)$$

\vec{d} is the crowding distance, $\max \vec{x}$ denotes maximum value for every objective, and $\min \vec{n}$ represents minimum value for every objective. Based on the computed distances, a rank is assigned to each solution. Then, a roulette wheel is used to choose the target. The solutions with more NSs have higher ranks, and thus, the target is chosen among the solutions with the lowest ranks. Setting the size of the archive is another challenge. Archives can store a limited number of non-dominated solutions. Thus, the solutions with the most crowded neighborhoods are chosen for removal from the archive. Following a similar process, the crowding distances are computed, and a rank is assigned to each solution. A roulette wheel is used for selecting the solutions with the highest ranks (the most crowded neighborhoods). To update the archive of non-dominated solutions, the following rules should be considered:

- 1- If one of the archive solutions dominates the external solutions, the external solution should be discarded.
- 2- If a solution dominates all non-dominated solutions in the archive, the external solution should be added to the archive.

The MOSOA, MOBA, MOPSO algorithm, and MOSSA are performed based on the following procedure:

- 1- The random parameters of the MOAs are defined for each algorithm.
- 2- The random positions (velocities) of agents (particles, bats, seagulls, and salps) are defined.
- 3- The objective function (OBF) is calculated for each agent.
- 4- The non-dominated solutions are determined based on the value of the OBF.
- 5- The size of the archive is checked, and if it is full, the solutions with the most NSs are removed from the archives.
- 6- The archive should be updated based on the rules mentioned above.
- 7- The target is selected from the solutions with the least crowded neighborhoods.

- 8- The process is repeated until the satisfaction of an end condition occurs.

Case study

According to the Paris Agreement, Iran has agreed to decrease CO₂ emissions by 4% by 2030. Iran has stated that if there is international support without any risk of sanctions, a 12% reduction in CO₂ emissions is possible. However, there are challenges with respect to decreasing CO₂ emissions. During the 2000s, environmental issues had lower importance than economic and social issues. The developing industrial unit around Tehran (capital of Iran) is one of the main causes of increased CO₂ emissions and air pollutants. Iran is known as the 7th largest CO₂ emitter in the world. The agricultural sector of Iran is one of the important causes of increased CO₂ emissions. There are different parameters in the agricultural sector that affect CO₂ emissions. One of the most important parameters is the gross domestic product (GDP). Different studies have investigated the effect of the GDP on CO₂ emissions. Cowan et al. (2014) stated that there was a Granger causality between the GDP and CO₂ emissions in Brazil during the period from 1990 to 2000. Zubair et al. (2020) investigated the relationship between CO₂ emissions and the GDP in Nigeria. They stated that the CO₂ emissions during the long-term period (1980–2018) in Nigeria decreased with increasing GDP. Additionally, the literature has used the environmental Kuznets curve (Kuznets 1955) to find the relationship between economic growth and environmental quality. Grossman and Krueger (1991) used the Kuznets curve to explain the relationship between environmental degradation and economic growth. Based on the Kuznets curve, first, economic growth degraded the environment. After economic growth reached its maximum value, the growth improved the environment. Additionally, the square of the per capita GDP is one of the most effective parameters for predicting CO₂ emissions (Shabani et al. 2021; Hosseini et al. 2019). Another effective parameter for predicting CO₂ emissions is the Gini coefficient (Cheng et al. 2021; Shabani et al. 2021). The Gini index estimates income inequality in the agricultural sector of Iran. Energy consumption is yet another effective parameter for estimating CO₂ emissions. The amount of fossil fuels utilized in this sector is presented as the energy consumption variable. Ali et al. (2021) stated that increasing fossil fuel usage increases CO₂ emissions. Koengkan et al. (2019) stated that renewable energies should be used as alternatives to fossil fuels because fossil fuels increase CO₂ emissions.

Generally, based on the abovementioned discussions, the following inputs are used in the current study for predicting CO₂ emissions:

- 1- GDP
- 2- Square of the GDP
- 3- Energy use
- 4- GINI index

To obtain the real GDP, the nominal values are divided by the producer price index, and the total values are divided by the population to obtain the per capita GDP. However, different input combinations can be provided based on the inputs above. To find the best input combination, the SVM is coupled with the MOAs. Table 2 shows the statistical characteristic data. The annual data for 25 provinces of Iran from 1990 to 2018 are extracted to predict the CO₂ emissions caused by the agricultural sector of Iran. The website of the Statistics Center of Iran is used to collect the dataset.

Hybrid SVM and MOAs

In this section, the MOAs are used to improve the accuracy of the SVM based on the following procedure:

- 1- The input data are prepared for the SVM models. Seventy percent of the data are used for training, and 30% of the data are used for testing. These percentages are chosen because they provide the lowest error for the SVM model.
- 2- The SVM model uses the training data for predicting CO₂ emissions. If the stopping criterion is satisfied, it proceeds to the testing phase; otherwise, the MOAs are linked to the SVM model.
- 3- The MOAs are defined based on the rules in section 2.5. The names of the inputs, the names of the kernel functions, and the initial guesses of the kernel parameters are inserted into the MOAs as decision variables. Three objective functions are used to find the parameter values, the best kernel function, and the best input combination. The RMSE is used as the first objective function for finding kernel parameters. The MAE is used to find the best kernel function. The NSE is used to find the best input combination. In fact, the positions of the agents determine the values of the decision variables.
- 4- The Pareto front is created, and the nondominated solutions are placed on the Pareto front. Each solution includes three kinds of information: the best input combination, the values of the parameters, and the best kernel function.
- 5- A multicriteria decision process is used to choose the best solution from the Pareto front as the final solution.
- 6- The SVM model based on the obtained formations runs again, and the process is repeated.

Table 2 The statistical characteristic of input and output data

Parameter	Unit	Average	Maximum	Minimum	Standard deviation
Training					
CO ₂ (output)	Per capita CO2 Metrics kg	6.12	12.98	2.44	2.225
Gini	-	30.24	45.12	22.23	8.76
GDP	Real Per capita GDP (Billion Rial)	0.824	8.76	0.023	1.912
GDP ²	Square of GDP (Billion Rial)	612.44	898.23	9.25	124.56
EC	Energy use 1000 barrels of oil equivalent	1123.12	4567.87	567.24	85.12
Testing					
CO ₂ (output)	Per capita CO2 Metrics kg	7.44	12.25	1.98	2.34
Gini	-	32.25	47.12	24.25	9.12
GDP	Real Per capita GDP (Billion Rial)	0.912	9.12	0.045	1.998
GDP ²	Square of GDP (Billion Rial)	615.54	897.12	12.23	134.23
EC	Energy use 1000 barrels of oil equivalent	1112.2	4678.12	672.2	98.23

Inclusive multiple model

Previous studies used competitive models for predicting different target variables, but there are several concerns:

- 1- The final outputs of the previous studies were the selections of the worst model and the best model. The best model was suggested for subsequent studies, and the worst models were discarded.
- 2- The intercomparisons between the predictive models were never exhaustive.
- 3- There was no effort to provide more accurate results based on the synergy among all competitive models used.

In this study, first, SVM-MOSOA, SVM-MOPSO, SVM-MOSSA, and SVM-MOBA are used to predict CO₂ emissions. In the next stage, to increase the accuracy of the outputs, an inclusive multiple model is used to improve the accuracy of the final outputs based on the synergy among the SVM-MOSOA, SVM-MOPSO, SVM-MOSSA, and SVM-MOBA models as follows:

- 1- First, the SVM-MOSOA, SVM-MOPSO, SVM-MOSSA, and SVM-MOBA models are used to predict CO₂ emissions.
- 2- In the next stage, the outputs of SVM-MOSOA, SVM-MOPSO, SVM-MOSSA, and SVM-MOBA are used as the inputs for the ANN model to predict CO₂ emissions.

In fact, the outputs of the first stage are considered lower-order modeling results. The utilization of the inclusive multiple model causes the accuracy of the results to be increased based on the utilization of the advantages of all competitive models. The modeler ensures that the capacities of all models are used to extract the most accurate results possible.

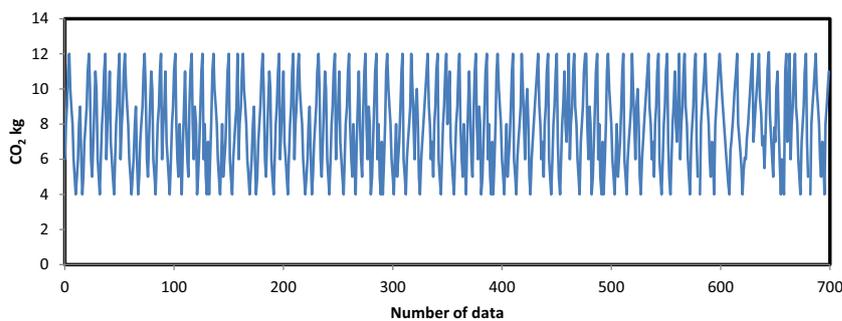
However, the ANN model used in the current study includes one input layer, one hidden layer, and one output layer, as observed in Fig. 7. The input layer receives the outputs of the competitive models as its inputs. The hidden layer processes the received inputs based on the chosen activation function. In this study, the sigmoid function is used as it was one of the successful activation functions used in previous studies (Banadkooki et al. 2020b, Ehteram et al. 2021). The ANN processes data based on the following equation:

$$Y = b_0 + \sum_{i=1}^{Nh} \omega_{ij} f \left(b_{oj} + \sum_{i=1}^{Nin} \omega_{ij} \mathbf{IN}_{ij} \right) \tag{31}$$

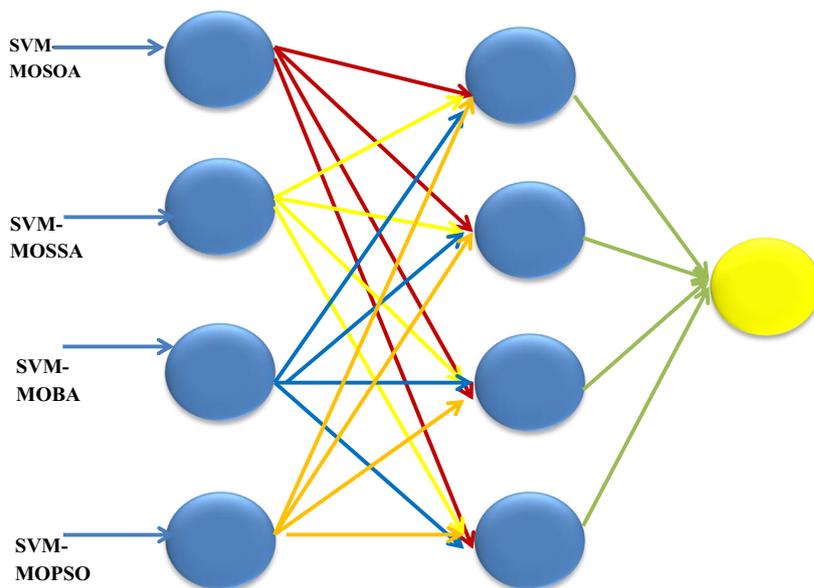
where b_0 and b_{oj} are the bias of the output and the hidden layer, respectively; ω_{ij} is the weight of the i th input in the j th hidden layer neuron; \mathbf{IN}_{ij} is the network input; f is the activation function; ω_j is the weight of the output from hidden neuron j ; Nh is the number of hidden neurons; Y is the output; and Nin is the number of inputs.

$$f(y) = \frac{1}{1 + e^{-y}} \tag{32}$$

Fig. 7 (a) The measured CO₂ for 700 data points and (b) the structure of the IMM model



(a)



(b)

$f(y)$ is the output of the activation function based on the input y ($y: b_{oj} + \sum_{i=1}^{Nin} \omega_{ij} IN_{ij}$). In this study, the back-propagation algorithm is used to train the ANN model.

Multicriteria decision model

A Pareto front includes multiple solutions. To choose the best solution, a multicriteria decision model is used. One of the powerful multicriteria decision models is the weighted aggregate sum product assessment (WASPAS) technique, which is widely used in different fields, such as solving solar wind power problems (Nie et al. 2017), fuel technology selection (Rani and Mishra 2020), and the development of smart cities (Khan et al. 2020). To use the WASPAS technique to choose

the best solution from the Pareto front, the following steps are considered:

- 1- A decision matrix is adjusted so that the values of each criterion (NSE, RMSE, and MAE) are inserted into the matrix for each solution.
- 2- The decision matrix is normalized based on the following equations:

For the NSE criterion:

$$z_{ij}^* = \frac{z_{ij}}{\max_i z_{ij}} \tag{33}$$

For the RMSE and MAE criteria:

$$z_{ij}^* = \frac{\min_i z_{ij}}{z_{ij}} \tag{34}$$

z_{ij}^* is a normalized value, and z_{ij} is the performance of the i th alternative with respect to the j th criterion.

3- In the next stage, the weighted sum model and weighted product model are computed as follows:

$$K_i = \sum_{j=1}^n \kappa_{ij} \cdot z_{ij} \quad (35)$$

$$L_i = \prod_{j=1}^n (z_{ij}^*)^{w_j} \quad (36)$$

where K_i is the weighted sum model, L_i is the weighted product model, κ_{ij} denotes the weight of each criterion, and n is number of criteria.

4- The aggregated measure is computed as follows:

$$S_i = \sigma K_i + (1-\sigma)L_i \quad (37)$$

where σ is a constant parameter ($\sigma:0.5$). The solution with the highest value of S_i is chosen as the best solution.

In this study, to evaluate the performances of the tested models, the following indexes are used:

- Root mean square error (ideal values are close to zero)

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (CO_{2es} - CO_{2ob})^2} \quad (38)$$

• Scatter index (SI<0.10: excellent performance, SI:0.10 <SI< 0.20: good performance, SI:0.20 <SI< 0.30: fair performance, and SI>0.30: poor performance (Li et al. 2013):

$$SI = \frac{RMSE}{C\bar{O}_{2ob}} \quad (39)$$

- Mean absolute error (MAE):

$$MAE = \frac{1}{m} \sum_{i=1}^m |CO_{2es} - CO_{2ob}| \quad (40)$$

• Nash Sutcliffe efficiency (NSE) (values close to 1 are ideal)

$$NSE = 1 - \frac{\sum_{i=1}^m (CO_{2es} - CO_{2ob})^2}{\sum_{i=1}^m (CO_{2es} - C\bar{O}_{2es})^2} \quad (41)$$

- Uncertainty with a 95% confidence level (lowest values are ideal)

$$U_{95} = 1.96 \sqrt{(SD^2 + RMSE^2)} \quad (42)$$

- Percentage bias

$$PBIAS = 100 * \frac{\left[\sum_{i=1}^n (CO_{2es} - CO_{2ob})^2 \right]}{\sum_{i=1}^n CO_{2ob}} \quad (43)$$

where SD is the standard deviation of the residual, m is the number of data points, CO_{2ob} denotes the observed CO_2 levels, and $C\bar{O}_{2es}$ is the average estimated $C\bar{O}_{2es}$.

Additionally, to evaluate the quality of the Pareto fronts obtained by the different MOAs, the following indices are used:

1- Spacing index: this shows the spread of the computed Pareto front.

$$SP = \sqrt{\frac{1}{np-1} \sum_{i=1}^{np} (\bar{d} - d_i)^2} \quad (44)$$

where np is the number of Pareto solutions, d_i is the Euclidean distance between the two consecutive in the Pareto front, and \bar{d} is the average distance between solutions. The lowest SP value corresponds to the best algorithm.

Maximum spread (MS): the MS shows the distance between the boundary conditions. The highest MS value corresponds to the best algorithm:

$$MS = \sqrt{\sum_{i=1}^{nf} \max(d(r_i, t_i))} \quad (45)$$

where d is a function used to compute the Euclidean distance, r_i is the maximum value in the i -th objective function, t_i is the minimum value in the i th objective function, and nf is the number of objective functions.

Discussion and results

Determination of random algorithmic parameters

The MOAs used have random parameters, and their accurate values should be determined because the accuracies of MOAs depend on the random parameters chosen. One of the most

robust methods for designing parameters and experiments is the Taguchi model. The Taguchi model is widely used in different fields, such as for optimizing the tuning parameters of models (Dutta and Kumar Reddy Narala 2021), optimizing thermal conductivity (Aswad et al. 2020), optimizing air distributor channels (Feng et al. 2020), optimizing soil erosion (Zhang et al. 2021a), and optimizing runoff water quality (Liu et al. 2021). To determine the values of the random parameters, the following steps are considered:

- 1- The levels of the parameters and the number of parameters should be determined. For example, the population size and the maximum number of iterations should be determined for the MOSOA. Thus, there are two parameters. As observed in Tables 3a and 4 values are assigned to the population size and the maximum number of iterations to find the best values of these random parameters.
- 2- Regarding the number of parameters and their levels, an orthogonal array is chosen from the Taguchi table. The Taguchi model uses the orthogonal array to decrease the required number of experiments. Regarding the two parameters with four levels each in the MOSOA, L_{16} covers the required experiments for the parameters of the MOSOA. As observed in Table 3b, there are 16 experiments for two parameters with four levels.
- 3- In the next stage, the signal (S)-to-noise (N) ratio is computed for each experiment:

$$\frac{S}{N} \text{ratio} = -10 \log_{10}(\text{objective}(\text{function}))^2 \tag{44}$$

High values of the S/N are ideal.

- 4- The mean of the S/N values is computed for factors at the different levels and computed as follows:

$$\begin{aligned} (\text{Mean})_{\text{factor}} &= l^{\text{level}} = i = \frac{1}{n_{ij}} \\ &\times \sum_{j=1}^{n_{ij}} \left[\left(\frac{S}{N} \right)_{\text{factor}} = l = i \right] \end{aligned} \tag{45}$$

where *Mean* S/N denotes the mean signal-to-noise ratio.

Table 3b shows the S/N values for 16 experiments. The average S/N for each of the levels of the parameters is computed in Table 3c. The highest mean S/N values correspond to the best parameter values. As observed in Table 3c, the population size and the maximum number of iterations in level 1 have the highest mean S/Ns. Thus, the best values for the

population size and the maximum number of iterations are 14.26 and 15.33, respectively. Obtained via a similar process, the optimal values of the other parameters for the other compared MOAs are reported in Table 3d.

The Pareto fronts of different models

Figure 8 shows the Pareto fronts obtained for different models. As observed in Fig. 8, the RMSE, MAE, and NSE of SVM-MOSOA vary from 0.25 to 0.75 (kg), from 0.12 to 0.62 (kg), and from 0.92 to 0.98, respectively. The RMSE, MAE, and NSE of SVM-SSA vary from 0.35 to 10.05 (kg), from 0.24 to 0.72, and from 0.91 to 0.97, respectively. The RMSE and MAE of SVM-MOSOA for the non-dominated solutions are lower than those of the other models. The red circle shows the best solution obtained by the multicriteria decision process. As observed in Table 4, the best solution determines the best input scenario, the values of the SVM parameters, and the best kernel function. Table 4a shows the best kernel functions for the different SVM models. The best kernel function for SVM-MOSOA is the RBF, and it is the sigmoid function for SVM-MOSSA, SVM-MOBA, and SVM-MOPSO. The values of the SVM parameters are shown in Table 4b, and the best input combinations are shown in Table 4c. Figure 9 compares the quality of the Pareto fronts obtained by the different models. As observed in Fig. 9, MOSOA provides the lowest SP and the highest MS among the models. The MOPSO provides the lowest MS and the highest SP. Thus, MOPSO and MOSOA provide the worst and the best Pareto fronts among the tested models, respectively.

Comparison of model accuracies with respect to predicting CO₂ emissions

Table 5 compares the performances of different models. The RMSE of SVM-MOSOA during the training stage is 16, 55, 57, and 60% lower than those of the SVM-MOSSA, SVM-MOBA, and SVM-MOPSO models. The MAE obtained by SVM-MOSOA is 0.29, and they are 0.32, 0.69, and 0.91 for the SVM-MOSSA, SVM-MOBA, and SVM-MOPSO models, respectively. SVM-MOSOA obtains the highest NSE and the lowest PBIAS during the training stage among the different SVM-MOA models. If the accuracies of the SVM-MOA models are compared with those when the IMM model is added, it is clear that the IMM model increases the accuracy of each of the SVM-MOA models during the training stage. The IMM model decreases the MAEs of the SVM-SOA, SVM-MOSSA, SVM-MOBA, and SVM-PSO models by 24, 31, 68, and 76%, respectively. The model performance comparison indicates that the IMM and SVM-MOBA have the highest and lowest NSEs, respectively, among the tested models during the training phase. The performances of the models indicate that SVM-MOSOA has the best accuracy

Table 3 (a) The level and factors of MOSOA, (b) the S/N value of 16 experiments, (c) the mean S/N value for different parameters, and (d) the optimal value of other parameters of other MOAs

a			
Population size	Maximum number of iterations		
Level 1=50	Level 1=100		
Level 2=100	Level 2=200		
Level 3=150	Level 3=300		
Level 4=200	Level 4=400		
b			
Run	Population size	Maximum number of iterations	S/N
1	1	1	15.67
2	1	2	16.72
3	1	3	12.89
4	1	4	11.78
5	2	1	14.12
6	2	2	12.23
7	2	3	14.87
8	2	4	12.98
9	3	1	15.76
10	3	2	14.76
11	3	3	12.25
12	3	4	11.12
13	4	1	15.78
14	4	2	12.78
15	4	3	11.25
16	4	4	10.98
c			
Population size	S/N	Maximum number of iterations	S/N
Level 1=50	14.26	Level 1=100	15.33
Level 2=100	12.30	Level 2=200	14.12
Level 3=150	11.21	Level 3=300	12.75
Level 4=200	13.42	Level 4=400	11.99
d			
MOBA: population size: 200, the maximum number of iterations: 100, maximum frequency:5 Hz, minimum frequency:1, maximum loudness:0.70, and minimum loudness:0.20			
MOPSO: population size:100, the maximum number of iterations:100, acceleration coefficient :2.0, and inertia weight :0.5			
MOSSA: population size:200, the maximum number of iterations:100, $\rho_2=0.4$ and $\rho_3=0.8$			

among the SVM-MOA models. The NSE of SVM-MOSOA is 0.93, and they are 0.90, 0.89, and 0.84 for the SVM-MOSSA, SVM-MOBA, and SVM-MOPSO models, respectively. The performances of the models indicate that the IMM model outperforms all SVM-MOA models in the testing phase. The model accuracy comparison indicates that the IMM and SVM-MOPSO have the lowest and highest PBIAS among the models, respectively. As observed in Table 5, the IMM model decreases the RMSEs of the SVM-MOSSA, SVM-MOBA, and SVM-MOPSO models by 20, 36, 47, and 50%, respectively. Figure 10a compares the performances of the models based on the IS index. As observed

in this figure, the SI values of the IMM model are 0.04 and 0.060 during the training and testing stages, respectively, and thus, the performance of the IMM model is excellent. The performances of SVM-MOSOA, SVM-SSA, SVM-BA, and SVM-PSO are good during the training and testing phases. Figure 10b compares the performance of the models based on the U95% measure. As observed in Fig. 10b, it is clear that the IMM and SVM-MOSOA provide the lowest uncertainty levels in comparison with the other models. Figure 11 shows the scatterplots obtained using testing data. Additionally, the R^2 values obtained on the training data are mentioned for each model. The R^2 of the IMM is 0.995, while they are 0.9899,

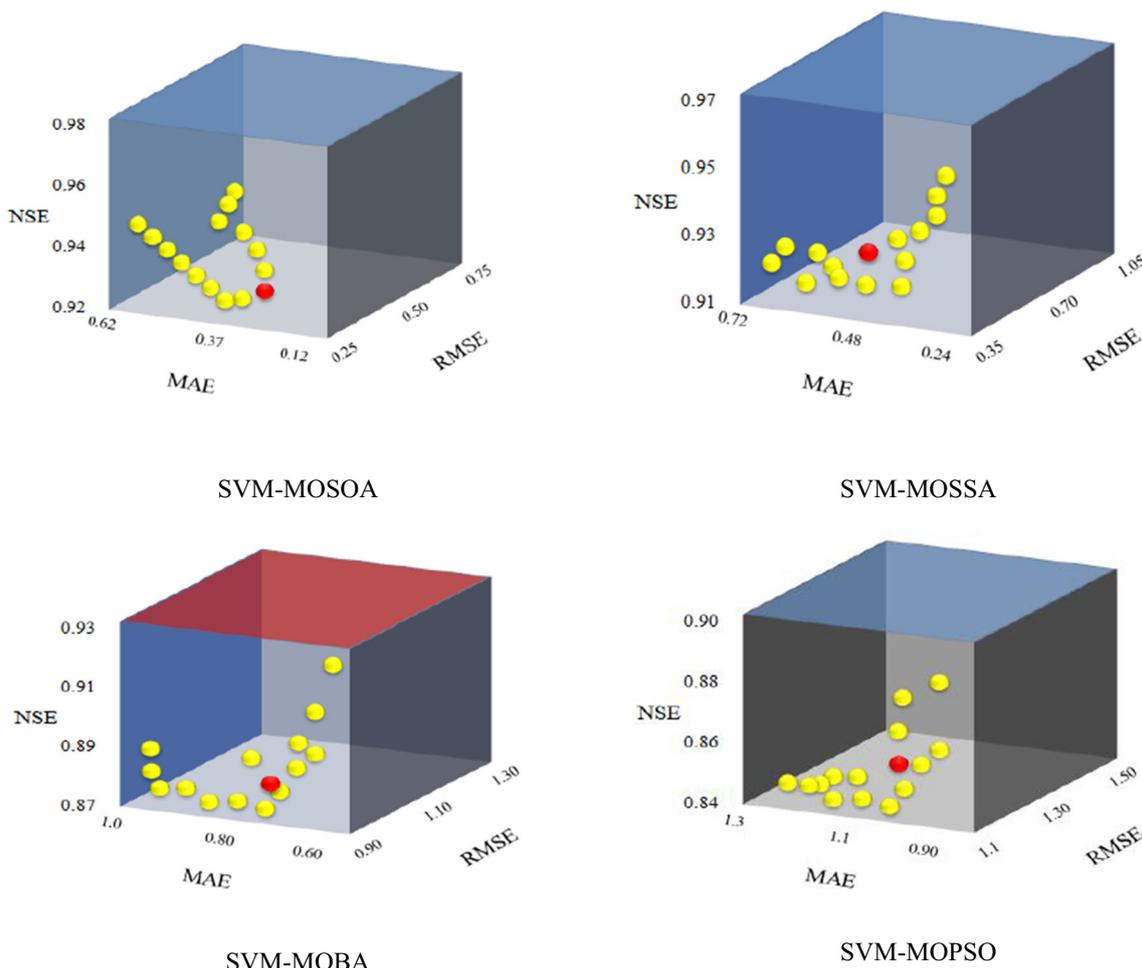


Fig. 8 Obtained Pareto fronts for different models

0.9830, 0.9798, and 0.9752 for the SVM-MOSOA, SVM-MOSSA, SVM-MOBA, and SVM-MOPSO models, respectively. Thus, the IMM has the highest R^2 values. Taylor diagrams are effective diagrams for comparing models. A Taylor diagram uses three evaluation criteria, namely, the standard

deviation, root mean square error, and correlation coefficient, to evaluate the accuracies of the compared models. A model has a better performance than those of other models if its point is closest to the reference point (observed data). As observed in Fig. 12, the IMM and SVM-MOSOA have better

Table 4 The obtained information by the best solution of Pareto fronts for (a) the selection of kernel function, (b) the values of SVM parameters, and (c) the best input combination

a				
Kernel	SVM-MOSOA	SVM-MOSSA	SVM-MOBA	SVM-MOBA
Sigmoid function		✓	✓	✓
RBF	✓			
Polynomial				
b				
Parameters of sigmoid for SVM-MOSOA: PE:3, γ :2, and ϵ :0.168				
Parameters of RBF for SVM-MOSSA: PE:2, γ :0.509, r:2, and ϵ :0.168				
Parameters of RBF for SVM-MOBA: PE:2, γ :0.712 r:2, and ϵ :0.268				
Parameters of RBF for SVM-MOPSO: PE:2, γ :0.514 r:2, and ϵ :0.112				

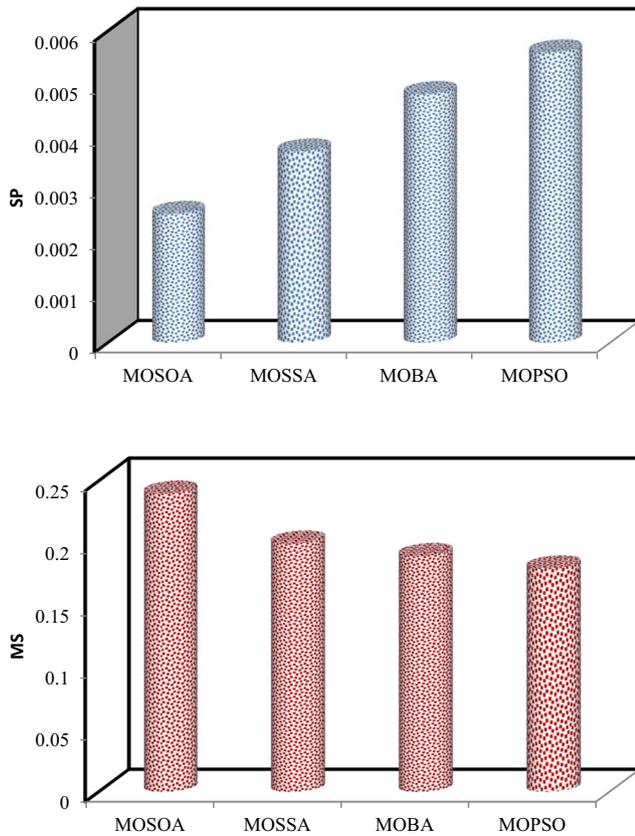
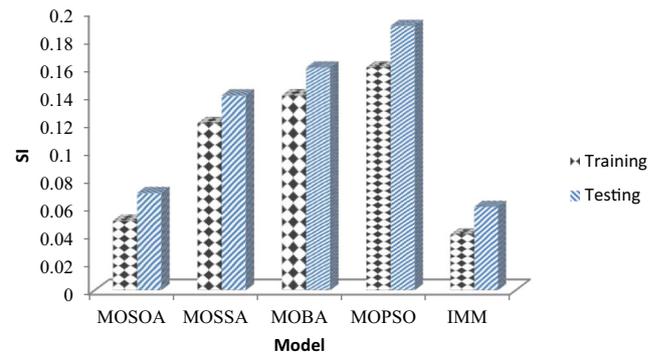


Fig. 9 The SP and MS values for comparing different models

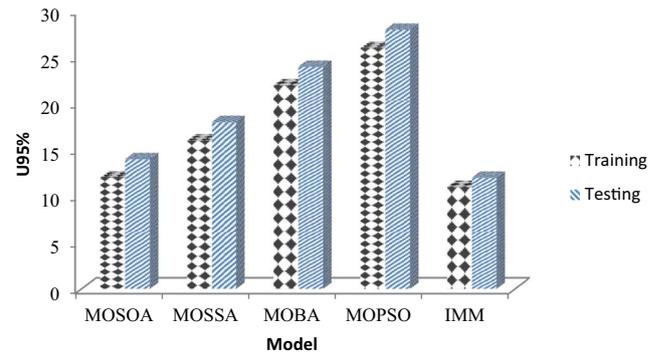
performances than the other models. Figure 13 shows the boxplots of the models. As observed in this figure, the IMM and SVM-MOSOA have good agreement with the measured data. The relative error of each data point for the different models is computed based on the heat map in Fig. 14. As observed in Fig. 14, the relative error induced by the IMM model varies from 10 to 20%, and this model has the lowest relative errors among the tested models. The relative error of

Table 5 The comparison of the accuracy of different models

Model	RMSE	MAE	NSE	PBIAS
Traning				
IMM	0.33	0.22	0.98	0.16
SVM-MOSOA	0.35	0.29	0.97	0.12
SVM-MOSSA	0.42	0.32	0.94	0.15
SVM-MOBA	0.78	0.69	0.90	0.19
SVM-MOPSO	0.89	0.91	0.86	0.21
Testing				
IMM	0.36	0.24	0.95	0.18
SVM-MOSOA	0.45	0.32	0.93	0.16
SVM-MOSSA	0.56	0.45	0.90	0.19
SVM-MOBA	0.97	0.93	0.89	0.21
SVM-MOPSO	0.99	0.96	0.84	0.23



(a)



(b)

Fig. 10 Comparison of the accuracy of different models based on a: SI index and b: U95%

SVM-MOBA varies from 40 to 60% and 50 to 60% for the SVM-MOPSO model.

Concluding discussion

Regarding the obtained results, the following points should be considered:

- 1- One of the main advantages of the current study is that it finds the best input combination without preprocessing data, such as through the use of the gamma test and principal component analysis. Each multiobjective algorithm can automatically find the best input combination with respect to a defined objective function.
- 5- The findings of the current research confirm the results from previous article by Ahmadi et al. (2019b), who showed the optimization algorithm could improve the accuracy of SVM and LSSVM models for predicting CO₂ emission. They stated that the optimization algorithms such as PSO and GA had the high abilities for improving the accuracy of soft computing models such as SVM and LSSVM model. Sun and Liu (2016) used the LSSVM as a family member of SVM model to predict CO₂ emission. They reported that the LSSVM model

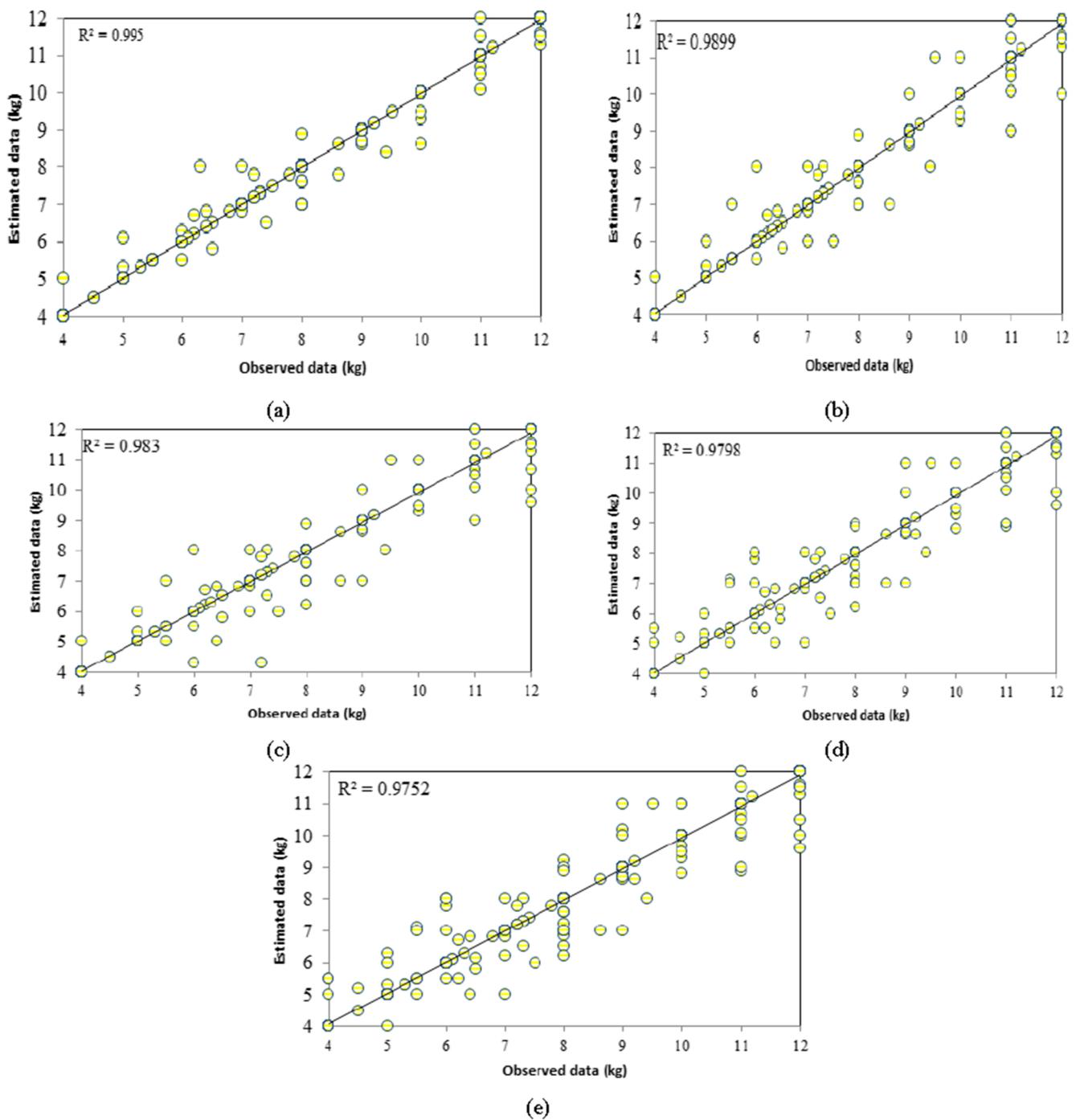
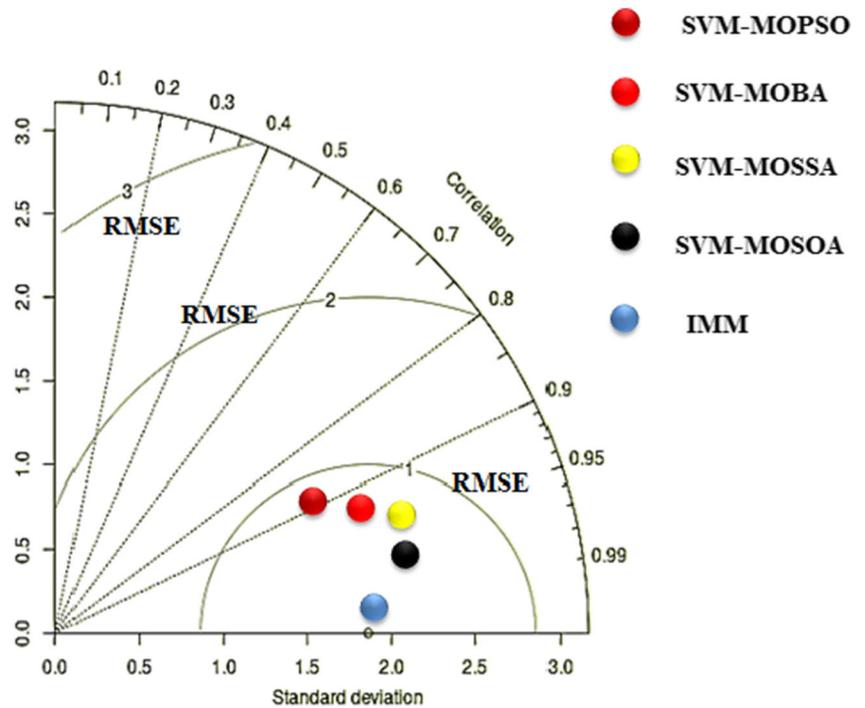


Fig. 11 The scatterplots for the (a) IMM (R^2 (training:0.997)), (b) SVM-MOSA (R^2 (training:0.9912)), (c) SVM-MOSSA (R^2 (training:0.9876)), (d) SVM-BA (R^2 (training:0.9812)), and (e) SVM-MOPSO (R^2 (training:0.9798))

outperformed the backpropagation neural network and grey models. Shabani et al. (2021) confirmed that the square of the GDP, GDP, energy use and GINI index were the effective inputs for predicting CO₂ emission. A comparison of the results obtained by the current study and Shabani et al. (2021) revealed accurate predictions with least errors for IMM model.

- 2- The findings of the current study confirm the results from previous studies by Ehteram et al. (2021) and Seifi et al. (2020), who showed the high potential of multiobjective algorithms for improving the performances of MLAs.
- 3- The results of the current study are similar to those of Shabani et al. (2021) and Khatibi et al. (2017), who

Fig. 12 Taylor diagram for comparing different models



showed that the IMM-based models provide high accuracy for the prediction of target variables.

- 4- Future studies can investigate the effects of uncertain inputs and model parameters on the accuracies of the resulting models.
- 5- All of the evaluation criteria show that the performances of the IMM and SVM-MOSOA are the best, but future studies can use the multicriteria model to assign a rank to each model and then select the one with the best performance.
- 6- The performances of the models used in the current study are not limited to theoretical aspects; policy-makers and decision-makers can utilize the tools used here to identify the effective parameters of CO₂ emissions and then find real relationships between the inputs and outputs. Thus,

the models are suitable for environmental management and provide good insights for predicting CO₂ emissions.

- 7- However, it should be noted that there are challenges with improving MLA models based on MOAs, such as converting a single objective optimization algorithm to multiobjective optimization algorithms, defining various objective functions, and selecting the best solution on the given Pareto front.
- 8- Depending on the kinds of MLAs used in a given study, objective functions can be defined. For example, the number of hidden layers, the ANN parameters (weight and bias), the selection of the activation function, and the best input combination selection can be defined as the decision variables for an ANN model. Thus, four objective functions can be defined for the model under study.

Fig. 13 The box plot of models for predicting CO₂ emission

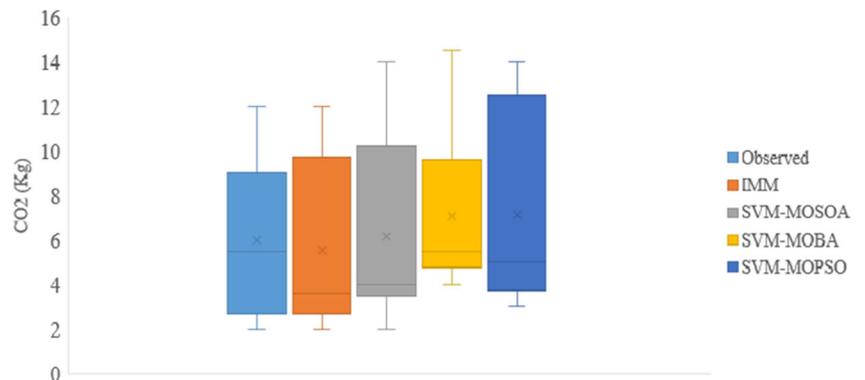
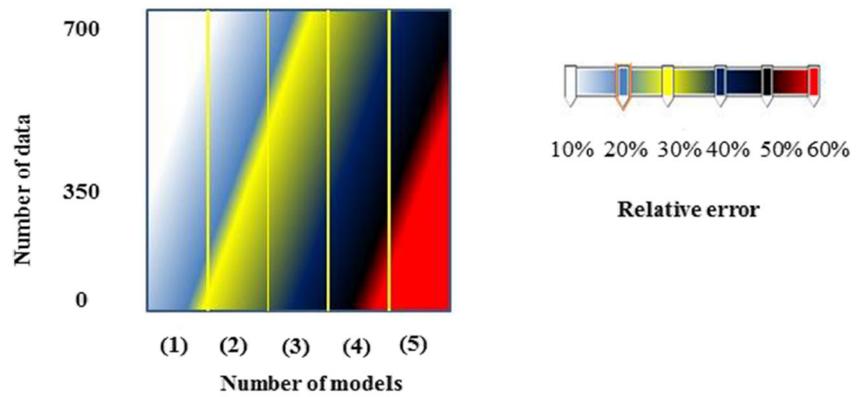


Fig. 14 The heat map for comparing models based on relative error



(1):IMM (3):SVM-MOSSA (5):SVM-MOPSO
 (2):SVM-MOSOA (4):SVM-MOBA

- 9- The increasing number of inputs and model parameters may increase the computational complexities and time consumptions of the models.
- 10- Another ability of the introduced models is that they provide spatial maps of CO₂ emissions so that the regions with high and low CO₂ emissions are classified accurately.
- 11- An ensemble of different kinds of kernel functions for SVMs can also be one of the strategies for increasing the accuracy of the SVM model because the model can use the advantages of different kernel functions.

Conclusion

Predicting CO₂ emissions is a real challenge for policy-makers and decision-makers with respect to managing the environment. CO₂ emissions rely on different parameters, and thus, finding accurate relationships between model inputs and outputs is an important issue. The agricultural sector is one of the important sources of CO₂ production. Thus, the current study used SVM models for predicting CO₂ emissions based on data from the agricultural sector. In this study, some MOAs were used to improve the accuracies of several SVM models based on finding the best input combinations, model parameters and kernel functions. Then, an IMM model used the outputs of the models as its inputs to increase their accuracies. The best input combination for all models was the use of the Gini index, GDP, GDP², and EU. The SVM-MOA models selected different kernel functions for mapping complex relationships between the inputs and outputs. The results indicated that the SOA provided the best Pareto front among the tested MOAs. It was observed that the IMM model

decreased the MAEs of the SVM-MOSOA, SVM-MOSSA, SVM-MOBA, and SVM-PSO models by 24, 31, 75, and 76%, respectively. Additionally, SVM-MOSOA achieved the highest accuracy among the SVM-MOA models. The MAE obtained by SVM-MOSOA was 0.29, and they were 0.32, 0.69, and 0.91 for the SVM-MOSSA, SVM-MOBA, and SVM-MOPSO models, respectively. The general results indicated that the IMM model could significantly increase the accuracies of the SVM models. The results of the current study could be useful for decreasing pollutants in the environments of different countries. The next papers can cover the deficiencies of the current paper. They can test the models for predicting CO₂ in the different regions of the world. Also, they can add the other socioeconomic to the inputs to consider the effect of different inputs on the accuracy of the models.

Author’ contributions Conceptualization: Mohammad ehteram, Fatemeh Panahi; Saad Sh. Sammen Methodology: Saad Sh. Sammen, Mohammad Ehteram; Formal analysis and investigation: Mohammad Ehteram, Lariyah Mohd Sidek Writing original draft preparation: Mohammad Ehteram, Lariyah Mohd Sidek

Funding 'Not applicable

Data availability The datasets used and/ or analyzed during the current study are available from the corresponding author on reasonable request

Declarations

Ethical approval Not applicable

Competing interests The authors declare that they have no conflict interest.

Research involving human participants and/ or animals Not applicable

Consent to participate All authors have given consent to their contribution.

Consent to publish All authors have agreed with the content and all have given explicit consent to publish.

References

- Abd el-sattar S, Kamel S, Ebeed M, Jurado F (2021) An improved version of salp swarm algorithm for solving the optimal power flow problem. *Soft Comput* 25:4027–4052. <https://doi.org/10.1007/s00500-020-05431-4>
- Ahmadi MH, Jashnani H, Chau KW, Kumar R, & Rosen MA (2019a) Carbon dioxide emissions prediction of five Middle Eastern countries using artificial neural networks. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 1–13. <https://doi.org/10.1080/15567036.2019.1679914>
- Ahmadi MH, Madvar MD, Sadeghzadeh M, Rezaei MH, Herrera M, & Shamshirband, S. (2019b) Current status investigation and predicting carbon dioxide emission in Latin American countries by connectionist models. *Energies*, 12(10), 1916. <https://doi.org/10.3390/en12101916>
- Ali MU, Gong Z, Ali MU, Wu X, & Yao C (2021) Fossil energy consumption, economic development, inward FDI impact on CO2 emissions in Pakistan: Testing EKC hypothesis through ARDL model. *International Journal of Finance & Economics*, 26(3), 3210–3221. <https://doi.org/10.1002/ijfe.1958>
- Aswad MA, Saud AN, Ahmed MA (2020) Thermal conductivity optimization of porous alumina ceramics via Taguchi model. *Mater Sci Forum* 1002:125–131. <https://doi.org/10.4028/www.scientific.net/MSF.1002.125>
- Banadkooki FB, Ehteram M, Ahmed AN, Teo FY, Ebrahimi M, Fai CM, Huang YF, El-Shafie A (2020a) Suspended sediment load prediction using artificial neural network and ant lion optimization algorithm. *Environ Sci Pollut Res* 27:38094–38116. <https://doi.org/10.1007/s11356-020-09876-w>
- Banadkooki FB, Ehteram M, Panahi F, Sh. Sammen S, Othman FB, EL-Shafie A (2020b) Estimation of total dissolved solids (TDS) using new hybrid machine learning models. *J Hydrol* 587:124989. <https://doi.org/10.1016/j.jhydrol.2020.124989>
- Banadkooki FB, Singh VP, Ehteram M (2021) Multi-timescale drought prediction using new hybrid artificial neural network models. *Nat Hazards* 106:2461–2478. <https://doi.org/10.1007/s11069-021-04550-x>
- Cao Y, Li Y, Zhang G, Jermstittiparsert K, Razmjoooy N (2019) Experimental modeling of PEM fuel cells using a new improved seagull optimization algorithm. *Energy Rep* 5:1616–1625. <https://doi.org/10.1016/j.egy.2019.11.013>
- Chakri A, Khelif R, Benouaret M, Yang XS (2017) New directional bat algorithm for continuous optimization problems. *Expert Syst Appl* 69:159–175. <https://doi.org/10.1016/j.eswa.2016.10.050>
- Cheng S, Fan W, Zhang J, Wang N, Meng F, Liu G (2021) Multi-sectoral determinants of carbon emission inequality in Chinese clustering cities. *Energy* 214:118944. <https://doi.org/10.1016/j.energy.2020.118944>
- Cowan WN, Chang T, Inglesi-Lotz R, Gupta R (2014) The nexus of electricity consumption, economic growth and CO2 emissions in the BRICS countries. *Energy Policy* 66:359–368. <https://doi.org/10.1016/j.enpol.2013.10.081>
- Cui Z, Zhang J, Wu D, Cai X, Wang H, Zhang W, Chen J (2020) Hybrid many-objective particle swarm optimization algorithm for green coal production problem. *Inf Sci* 518:256–271. <https://doi.org/10.1016/j.ins.2020.01.018>
- Darabi H, Mohamadi S, Karimidastenaei Z, Kisi O, Ehteram M, ELShafie A, Torabi Haghighi A (2021) Prediction of daily suspended sediment load (SSL) using new optimization algorithms and soft computing models. *Soft Comput* 25:7609–7626. <https://doi.org/10.1007/s00500-021-05721-5>
- Darwish A, Ezzat D, Hassanien AE (2020) An optimized model based on convolutional neural networks and orthogonal learning particle swarm optimization algorithm for plant diseases diagnosis. *Swarm Evol Comput* 52:100616. <https://doi.org/10.1016/j.swevo.2019.100616>
- Deotti LMP, Pereira JLR, da Silva Júnior IC (2020) Parameter extraction of photovoltaic models using an enhanced Lévy flight bat algorithm. *Energy Convers Manag* 221:113114. <https://doi.org/10.1016/j.enconman.2020.113114>
- Dhiman G, Kumar V (2019) Seagull optimization algorithm: Theory and its applications for large-scale industrial engineering problems. *Knowl-Based Syst* 165:169–196. <https://doi.org/10.1016/j.knosys.2018.11.024>
- Dhiman G, Singh KK, Soni M, Nagar A, Dehghani M, Slowik A, Kaur A, Sharma A, Houssein EH, Cengiz K (2020) MOSOA: A new multi-objective seagull optimization algorithm. *Expert Syst Appl* 167:114150. <https://doi.org/10.1016/j.eswa.2020.114150>
- Dong J, Wu L, Liu X, Li Z, Gao Y, Zhang Y, Yang Q (2020) Estimation of daily dew point temperature by using bat algorithm optimization based extreme learning machine. *Appl Therm Eng* 165:114569. <https://doi.org/10.1016/j.applthermaleng.2019.114569>
- Duan H, Luo X (2020) Grey optimization Verhulst model and its application in forecasting coal-related CO2 emissions. *Environ Sci Pollut Res* 27:43884–43905. <https://doi.org/10.1007/s11356-020-09572-9>
- Dutta S, & Kumar Reddy Narala S (2021). Optimizing turning parameters in the machining of AM alloy using Taguchi methodology. *Measurement*, 169, 108340. <https://doi.org/10.1016/j.measurement.2020.108340>
- Ehteram M, Salih SQ, Yaseen ZM (2020) Efficiency evaluation of reverse osmosis desalination plant using hybridized multilayer perceptron with particle swarm optimization. *Environ Sci Pollut Res* 27:15278–15291. <https://doi.org/10.1007/s11356-020-08023-9>
- Ehteram M, Ahmed AN, Latif SD, Huang YF, Alizamir M, Kisi O, Mert C, El-Shafie A (2021) Design of a hybrid ANN multi-objective whale algorithm for suspended sediment load prediction. *Environ Sci Pollut Res* 28:1596–1611. <https://doi.org/10.1007/s11356-020-10421-y>
- El-Kenawy ES, & Eid M (2020). Hybrid gray wolf and particle swarm optimization for feature selection. *Int J Innov Comput Inf Control*. doi: <https://doi.org/10.24507/ijicic.16.03.831>
- Feng G, Lei S, Guo Y, Shi D, Shen JB (2020) Optimisation of air-distributor channel structural parameters based on Taguchi orthogonal design. *Case Stud Thermal Eng* 21:100685. <https://doi.org/10.1016/j.csite.2020.100685>
- Grossman G, & Krueger A (1991). Environmental Impacts of a North American Free Trade Agreement (No. 3914). National Bureau of Economic Research, Inc. <https://doi.org/10.3386/w3914>
- Hosseini SM, Saifoddin A, Shirmohammadi R, Aslani A (2019) Forecasting of CO2 emissions in Iran based on time series and regression analysis. *Energy Rep* 5:619–631. <https://doi.org/10.1016/j.egy.2019.05.004>
- Jia H, Xing Z, Song W (2019) A new hybrid seagull optimization algorithm for feature selection. *IEEE Access* 7:49614–49631. <https://doi.org/10.1109/ACCESS.2019.2909945>
- Jiang H, Yang Y, Ping W, Dong Y (2020) A novel hybrid classification method based on the oppositionbased seagull optimization

- algorithm. *IEEE Access*, 8, 100778–100790.(1) <https://doi.org/10.1109/ACCESS.2020.2997791>.
- Kandiri A, Mohammadi Golafshani E, Behnood A (2020) Estimation of the compressive strength of concretes containing ground granulated blast furnace slag using hybridized multi-objective ANN and salp swarm algorithm. *Constr Build Mater* 248:118676. <https://doi.org/10.1016/j.conbuildmat.2020.118676>
- Khan S, Haleem A, & Khan MI (2020). Analysing Challenges Towards Development of Smart City Using WASPAS. *Smart Cities—Opportunities and Challenges*, 463–474. https://doi.org/10.1007/978-981-15-2545-2_39
- Khatibi R, Ghorbani MA, Pourhosseini FA (2017) Stream flow predictions using nature-inspired Firefly Algorithms and a Multiple Model strategy – Directions of innovation towards next generation practices. *Adv Eng Inform* 34:80–89. <https://doi.org/10.1016/j.aei.2017.10.002>
- Khoshnevisan B, Rafiee S, Omid M, Yousefi M, Movahedi M (2013) Modeling of energy consumption and GHG (greenhouse gas) emissions in wheat production in Esfahan province of Iran using artificial neural networks. *Energy* 52:333–338. <https://doi.org/10.1016/j.energy.2013.01.028>
- Koengkan M, Santiago R, Fuinhas JA (2019) The impact of public capital stock on energy consumption: Empirical evidence from Latin America and the Caribbean region. *Int Econ* 160:43–55. <https://doi.org/10.1016/j.inteco.2019.09.001>
- Kuznets S (1955) Economic growth and income inequality. *The American economic review*, 45(1), 1–28
- Li MF, Tang XP, Wu W, Liu HB (2013) General models for estimating daily global solar radiation for different solar radiation zones in mainland China. *Energy Convers Manag* 70:139–148. <https://doi.org/10.1016/j.enconman.2013.03.004>
- Li J, Zhang B, & Shi J (2017) Combining a genetic algorithm and support vector machine to study the factors influencing CO2 emissions in Beijing with scenario analysis. *Energies*, 10(10), 1520. <https://doi.org/10.3390/en10101520>
- Liu W, Engel BA, Chen W, Wei W, Wang Y, Feng Q (2021) Quantifying the contributions of structural factors on runoff water quality from green roofs and optimizing assembled combinations using Taguchi method. *J Hydrol* 593:125864. <https://doi.org/10.1016/j.jhydrol.2020.125864>
- Mugemanyi S, Qu Z, Rugema FX, Dong Y, Bananeza C, Wang L (2020) Optimal Reactive Power Dispatch Using Chaotic Bat Algorithm. *IEEE Access* 8:65830–65867. <https://doi.org/10.1109/ACCESS.2020.2982988>
- Nie R xin, Wang J qiang, & Zhang H yu. (2017). Solving solar-wind power station location problem using an extended weighted aggregated sum product assessment (WASPAS) technique with interval neutrosophic sets. *Symmetry*, 9(7), 106. <https://doi.org/10.3390/sym9070106>
- Rani P, Mishra AR (2020) Multi-criteria weighted aggregated sum product assessment framework for fuel technology selection using q-rung orthopair fuzzy sets. *Sustain Prod Consum* 24:90–104. <https://doi.org/10.1016/j.spc.2020.06.015>
- Sain SR, Vapnik VN (1996) The Nature of Statistical Learning Theory. *Technometrics* 38:409. <https://doi.org/10.2307/1271324>
- Saleh C, Dzakiyullah NR, & Nugroho JB (2016, February) Carbon dioxide emission prediction using support vector machine. In *IOP Conference Series: Materials Science and Engineering* (Vol. 114, No. 1, p. 012148). IOP Publishing. <https://doi.org/10.1088/1757-899X/114/1/012148>
- Salgotra R, Singh U, Singh S, Singh G, Mittal N (2021) Self-adaptive salp swarm algorithm for engineering optimization problems. *Appl Math Model* 89:188–207. <https://doi.org/10.1016/j.apm.2020.08.014>
- Seifi A, Ehteram M, Soroush F (2020) Uncertainties of instantaneous influent flow predictions by intelligence models hybridized with multi-objective shark smell optimization algorithm. *J Hydrol* 587: 124977. <https://doi.org/10.1016/j.jhydrol.2020.124977>
- Shabani E, Hayati B, Pishbahar E, Ghorbani MA, Ghahremanzadeh M (2021) A novel approach to predict CO2 emission in the agriculture sector of Iran based on Inclusive Multiple Model. *J Clean Prod* 279: 123708. <https://doi.org/10.1016/j.jclepro.2020.123708>
- Shi Y, Han B, Zafar MW, Wei Z (2019a) Uncovering the driving forces of carbon dioxide emissions in Chinese manufacturing industry: An intersectoral analysis. *Environ Sci Pollut Res* 26:31434–31448. <https://doi.org/10.1007/s11356-019-06303-7>
- Shi Y, Han B, Han L, Wei Z (2019b) Uncovering the national and regional household carbon emissions in China using temporal and spatial decomposition analysis models. *J Clean Prod* 232:966–979. <https://doi.org/10.1016/j.jclepro.2019.05.302>
- Sun W, Liu M (2016) Prediction and analysis of the three major industries and residential consumption CO2 emissions based on least squares support vector machine in China. *J Clean Prod* 122:144–153. <https://doi.org/10.1016/j.jclepro.2016.02.053>
- Sun W, Wang C, Zhang C (2017) Factor analysis and forecasting of CO2 emissions in Hebei, using extreme learning machine based on particle swarm optimization. *J Clean Prod* 162:1095–1101. <https://doi.org/10.1016/j.jclepro.2017.06.016>
- Tubishat M, Ja'afar S, Alswaiti M, Mirjalili S, Idris N, Ismail MA, Omar MS (2021) Dynamic Salp swarm algorithm for feature selection. *Expert Syst Appl* 164:113873. <https://doi.org/10.1016/j.eswa.2020.113873>
- Wang Y, Ding G, & Liu L (2015, October) A regression forecasting model of carbon dioxide concentrations based-on principal component analysis-support vector machine. In *International Conference on Geo-Informatics in Resource Management and Sustainable Ecosystem* (pp. 447–457). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-45737-5_45
- Wang Y, Wang P, Zhang J, Cui Z, Cai X, Zhang W, & Chen J (2019). A novel bat algorithm with multiple strategies coupling for numerical optimization. *Mathematics*, 7(2), 135. <https://doi.org/10.3390/math7020135>
- Wu Q, Meng F (2020) Prediction of energy-related CO2 emissions in multiple scenarios using a least square support vector machine optimized by improved bat algorithm: a case study of China. *Greenh. Gases Sci. Technol*, 10(1), 160–175. <https://doi.org/10.1002/ghg.1939>
- Xin-gang Z, Ji L, Jin M, Ying Z (2020) An improved quantum particle swarm optimization algorithm for environmental economic dispatch. *Expert Syst Appl* 152:113370. <https://doi.org/10.1016/j.eswa.2020.113370>
- Xu G, Schwarz P, Yang H (2019) Determining China's CO2 emissions peak with a dynamic nonlinear artificial neural network approach and scenario analysis. *Energy Policy* 128:752–762. <https://doi.org/10.1016/j.enpol.2019.01.058>
- Yue X, Zhang H (2020) Modified hybrid bat algorithm with genetic crossover operation and smart inertia weight for multilevel image segmentation. *Appl Soft Comput J* 90:106157. <https://doi.org/10.1016/j.asoc.2020.106157>
- Zhang XW, Liu H, Tu LP (2020) A modified particle swarm optimization for multimodal multi-objective optimization. *Eng Appl Artif Intell* 95:103905. <https://doi.org/10.1016/j.engappai.2020.103905>
- Zhang H, Wang Z, Chen W, Heidari AA, Wang M, Zhao X, Liang G, Chen H, Zhang X (2021a) Ensemble mutation-driven salp swarm

- algorithm with restart mechanism: Framework and fundamental analysis. *Expert Syst Appl* 165:113897. <https://doi.org/10.1016/j.eswa.2020.113897>
- Zhang F, Wang M, Yang M (2021b) Successful application of the Taguchi method to simulated soil erosion experiments at the slope scale under various conditions. *Catena* 196:104835. <https://doi.org/10.1016/j.catena.2020.104835>
- Zhao H, Huang G, & Yan N (2018) Forecasting energy-related CO2 emissions employing a novel SSA-LSSVM model: considering structural factors in China. *Energies*, 11(4), 781. <https://doi.org/10.3390/en11040781>
- Zubair AO, Abdul Samad A-R, Dankumo AM (2020) Does gross domestic income, trade integration, FDI inflows, GDP, and capital reduces CO2 emissions? An empirical evidence from Nigeria. *Environ. Sustain* 2:100009. <https://doi.org/10.1016/j.crsust.2020.100009>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.