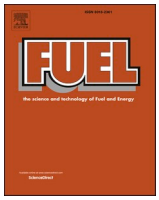
Fuel 340 (2023) 127569

Contents lists available at ScienceDirect

Fuel

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



Full Length Article

Thermo-economic optimization of an enhanced geothermal system (EGS) based on machine learning and differential evolution algorithms

Zhenqian Xue [a], Shuo Yao [b], Haoming Ma [a], Chi Zhang [a], Kai Zhang [a], Zhangxin Chen [a,*]

[a] Department of Chemical & Petroleum Engineering, University of Calgary, 2500 University Drive NW, Calgary, Alberta T2N 1N4, Canada
[b] PetroChina Changchun Oil Production Plant of Jilin Oilfield Company, Changchun 130000, China

ARTICLE INFO

Keywords:
Geothermal energy
Optimization
Artificial neural network
Differential evolution
LCOE

ABSTRACT

Hot dry rock (HDR) is considered as a promising low-carbon alternative to fossil fuels, but the remaining economic challenges are leading to its unsuccessful exploitation. Therefore, incorporating economic indicators into consideration is essential to optimize an enhanced geothermal system (EGS). However, the conventional optimization approaches based on numerical simulations are time-consuming and a global optimal operation strategy is hard to determine. In this study, an optimization framework based on Artificial Neural Network (ANN) and Differential Evolution (DE) is proposed with considering a levelized cost of electricity (LCOE) being an economic performance indicator to optimize a three-horizontal-well EGS in the Qiabuqia field. Specifically, four different ANN models are constructed to predict different geothermal productivities to substitute numerical models. Based on these ANN models, a DE optimization process is conducted to determine an optimal LCOE under two field operating constraints, followed by a performance comparison between the resulting optimal geothermal system and 2,150 randomly created cases using a numerical simulator. The results show that these ANN models all achieve a coefficient of determination R^2 higher than 0.996, demonstrating their predictive abilities and potential as surrogate models. The determined optimal parameters configuration brings a promising LCOE of 0.0376 $/kWh which is around 50 % of a local electricity cost, and this is the lowest LCOE among all random cases. Importantly, the proposed framework can significantly save operation time by 36,000 times compared with the numerical simulation method. The proposed method provides a valuable reference for the geothermal system studied, and it can also be effectively applied to other energy systems, thereby facilitating their optimal development.

1. Introduction

Fossil fuels are getting depleted due to rapidly increasing energy demand owing to population growth and industrial development [1] Besides, the consumption of fossil fuels directly emits a significant amount of CO_2 causing harmful effects on climate change [2–4]. Several international agreements such as the Paris Agreement and the United Nations Sustainable Development Goal 7 (SDG7) were raised to motivate the use of clean energy [5,6]. Nations have, therefore, gradually focused on the development of sustainable and environment-friendly renewable resources [7]. Geothermal energy can be a substitute of fossil fuels to meet the future energy demand owing to its sustainable and low-carbon features. It has been estimated that the geothermal resources under exploitation reach 30,000 MW, which can avoid the emissions of 46 million tons of CO_2 and save 0.352 billion barrels of oil annually [8–11].

Geothermal power can be extensively utilized in direct thermal energy harvesting and electricity generation. Rock formations can be divided into four types according to their underground temperatures ranging from 50C to 350C: (1) geopressured, (2) hydrothermal, (3) magma, and (4) dry hot rock [12,13]. Among them, hot dry rock (HDR) is the most potent option for long-term electrical power generation due to its high temperature and abundant reserves [14]. Its temperature can achieve 300C and energy reserves are estimated to reach approximately 1.3×10^{27} J which can provide global energy needs for 217 million years [14–16]. Technically, an enhanced geothermal system (EGS) technology is necessary since HDR normally exists with an ultra-low permeability porous medium, and it is utilized to create a highly conductive area for extracting heat [17]. Nowadays, HDR has not been widely exploited because of economic challenges [18]. Therefore, the optimization of a geothermal system based on its economic

* Corresponding author.
E-mail address: zhachen@ucalgary.ca (Z. Chen).

https://doi.org/10.1016/j.fuel.2023.127569
Received 30 August 2022; Received in revised form 8 January 2023; Accepted 19 January 2023

Nomenclature

EGS	Enhanced geothermal system
CMG	Computer Modelling Group
HDR	Hot dry rock
RMSE	Root mean squared error
ANN	Artificial Neural Network
MAE	Mean absolute error
DE	Differential Evolution
R^2	Coefficient of determination
TPE	Tree-structured Parzen Estimator
LCOE	Levelized cost of electricity
Symbols	
W_t	Cumulative electricity generation
P_{inj}	The injection pressure at the well bottom
T_{pro}	The temperature of produced water
W_a	Average electrical power generation

consideration is a significant task for a sustainable EGS construction [19].

The Chinese government announced to sufficiently utilize renewable energy to cut down CO_2 emissions in the China National Strategies responding to climate change by The State Council Information Office of the People's Republic of China [20]. In China, it has not successfully developed long-term geothermal electricity generation due to the failure in reservoir exploration until the discovery of borehole GR1 in the Qiabuqia geothermal field, which is located in the Gonghe basin, Qinghai province, China [21]. This field is broad with an approximate 246.9 km^2 area and abundant recoverable energy of HDR resources, which is equivalent to approximately 200 billion tons of standard coal [22]. Besides, the geothermal gradient in the Indosinian granitic formation is approximately 7.1C/100 m. The remarkably successful drilling well GR1 holds a 236C peak temperature at 3,705 m depth and the HDR section from 3,200 m to 3,705 m has an average temperature of 218C, implying the potential for constructing a long-term EGS power plant [23,24]. Therefore, this field is selected as the case study.

The economic analysis of a geothermal system is highly related to geothermal productivity [25]. The production of an EGS is influenced by many parameters, such as the properties of circulating fluid, well completion parameters and stimulation method design [1,18,22,23,26]. These parameters are, therefore, mainly considered in optimizing an EGS, and the numerical simulation method is widely used because of its accurate calculation ability [27]. It was used to operate parametric studies for searching optimal parameters related to the generated geothermal electricity, including a well placement and an injection temperature and rate of the circulating fluid [1,7,18,22,26,28–30]. Studies based on numerical simulations demonstrated that a number of hydraulic fractures was also optimized through sensitivity analysis regarding a production temperature and heat flux [31–33]. However, the relationship between these properties and geothermal productivity is highly nonlinear, resulting in that numerical solvers need to run many times during an optimization process [34].

An artificial neural network (ANN) provides an alternative for calculating an objective function due to its powerful function approximation capability and time-saving ability, which can analyze generated data comprehensively by evaluating their underlying relationships [35]. It has been extensively applied in the geothermal industry. Ishitsuka et al. [36] developed an ANN model to accurately estimate a temperature distribution in the Kakkonda geothermal field. Hu et al. [37] generated an ANN model to predict the hourly performance of a hybrid geothermal-solar system. Their results showed that ANN can provide high accuracy in estimation. Yilmaz et al. [25] calculated geothermal energy and hydrogen production of a geothermal-solar system by constructing ANN models, which achieved good performance in predicting both targets. Besides, Acar et al. [38] used ANN to create a surrogate model and optimize a geothermal energy powered Kalina cycle. Bassam et al. [39] estimated static formation temperatures in geothermal wells through the ANN. Pei et al. [40] utilized ANN to predict the long-term performance of energy pile designs of six typical thermal load distributions in China. Xiao et al. [41] developed an ANN model to estimate the outlet temperature of abandoned exploitation wells in geothermal power plants based on the data from a thermodynamic model.

Previous studies have proven the predictive ability of ANN in the geothermal industry. However, it was normally used to predict geothermal productivity. The optimization of a geothermal system was generally operated by manually created cases generated by numerical simulations or ANN. However, limited operation strategies can be considered in manually created cases, and it is difficult to determine the global optimal scenario due to the complex relationships between working factors and geothermal productivity. Therefore, it is necessary to implement a global search in different working factors. The Differential Evolution (DE) algorithm is a tool that can comprehensively search an optimal configuration for different properties in a range and achieve global convergence [42]. In addition, how to optimize an ANN model is rare in previous works, which is critical for a high-performance ANN model [43]. It is important to determine a method that can bring an optimal ANN model with comprehensive consideration of both prediction accuracy and efficiency.

In this work, an optimization framework based on ANN and DE is introduced to optimize a three-horizontal-well EGS in the Qiabuqia geothermal field. The ANN is utilized to generate surrogate models for substituting numerical simulations. A TPE (tree-structured parzen estimator) algorithm is developed in ANN models to determine optimal hyperparameters. The DE algorithm is operated to globally search for an optimal EGS with the lowest LCOE. The results provide a valuable reference for the optimization of a geothermal system, and the proposed ANN-based DE method can also be effectively applied in other renewable systems for determining their optimal operational strategies.

2. Methodology

2.1. Workflow

In this section, the detailed methodologies of the proposed ANN-based DE optimization framework are outlined as well as the application procedures to optimize a three-horizontal-well EGS in the Qiabuqia geothermal field.

Firstly, data samples are collected through numerical models using the Computer Modelling Group (CMG) software STARS. A base numerical model is first constructed to customize the Qiabuqia geothermal field. Depending on their effects on geothermal productivity, five operating parameters are considered the controlled operating parameters, including the rate and temperature of circulating water, well spacing, and the spacing and half-length of hydraulic fractures. Second, a total of 1,316 numerical models with different values of these factors are developed. The average electrical power generation (W_a) and cumulative electricity generation (W_t), the injection pressure at the bottom of a well (P_{inj}), and the temperature of produced water (T_{pro}) during a 20-year operation cycle are determined as the targets for calculating.

Secondly, based on the generated data points, four ANN models are created to predict four different targets. Before model training, a data preprocessing process is operated on these data samples, including data splitting and standardized transformation. During the model training, a *K*-fold cross-validation method and a TPE algorithm are applied to determine the best hyperparameters combination of different ANN models. Finally, the divided testing samples are transferred into the resulting optimal models to evaluate their predictive performance.

Lastly, regarding the determined ANN surrogate models, a DE algo-

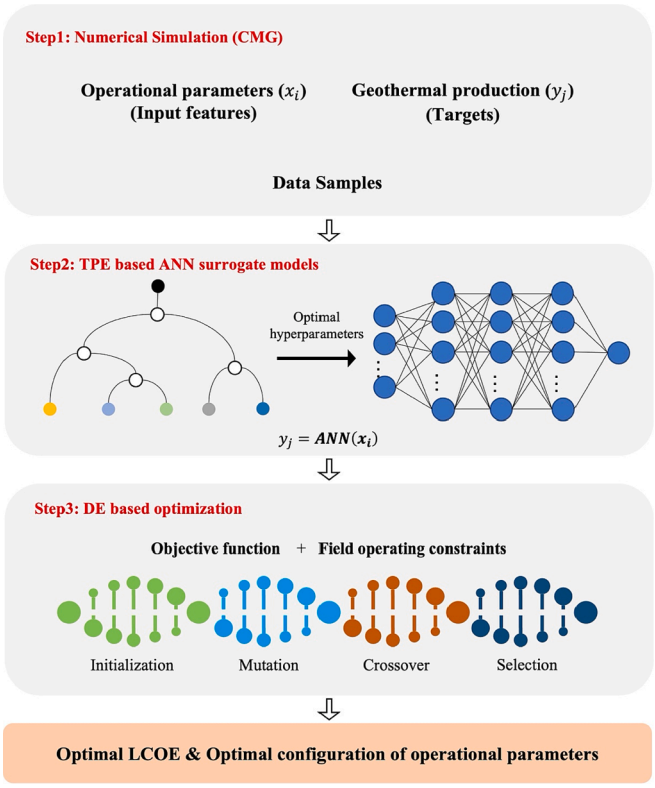


Fig. 1. Flowchart in this study.

rithm is utilized to optimize a geothermal system. An objective function is developed for calculating the LCOE. Furthermore, two field constraints are considered in the optimization process, containing the requirements on P_{inj} and T_{pro}. Consequently, to evaluate the feasibility of the proposed ANN-based DE optimization method, 2,150 randomly generated cases through numerical simulations are created for comprehensively considering different operational schemes of this geothermal field, and their LCOE values are compared with the LCOE of the resulting optimized geothermal system. The flowchart of this study is illustrated in Fig. 1.

2.2. Theories

2.2.1. Artificial neuron network (ANN)

An Artificial Neural Network (ANN) is biologically inspired by the sophisticated functionality of human brains [44,45]. In an ANN model, there are an input layer, one or two (or more than two) hidden layers and an output layer, and each of these layers contains different amounts of neurons [44]. Fig. 2 shows a typical architecture of an ANN model. When the inputs are transformed into an ANN model, they are calculated by the weights and biases layer by layer, and the output in the output layer can be expressed by Eq. (1) [44].

$$y_i = \sigma\left(\sum_{j=1}^{N} W_{ij}x_j + b\right) \tag{1}$$

where $\sigma()$ is an activation function, N is the number of input neurons, W_{ij} is a weight, x_j is an input value in the input neurons, and b is a bias which is considered as a special weight. The activation function is used to introduce the nonlinearity into a neural network and to bound the values of neurons for preventing the neural network from influencing by divergent neurons [46].

The flow proceeding from the first layer to the output layer is called a feed-forward process, and a back-propagation (BP) process operated from the output layer to the input layer is another significant step in an ANN model [21]. BP is utilized to minimize variance between predictions and targets [47]. It is achieved through altering the values of weights in the output layer by a small amount and updating the weights of each layer [48,49]. The results of a new generation are then obtained through another feed-forward process with updated weights. By comparing the values of the predictions and the true targets, the degree of a weight change in the next BP process is determined until a minimum overall error is reached.

2.2.2. Tree-structured parzen estimator (TPE) algorithm

A TPE algorithm is one of the sequential model-based optimization (SMBO) algorithms, which is also known as the Bayesian optimization

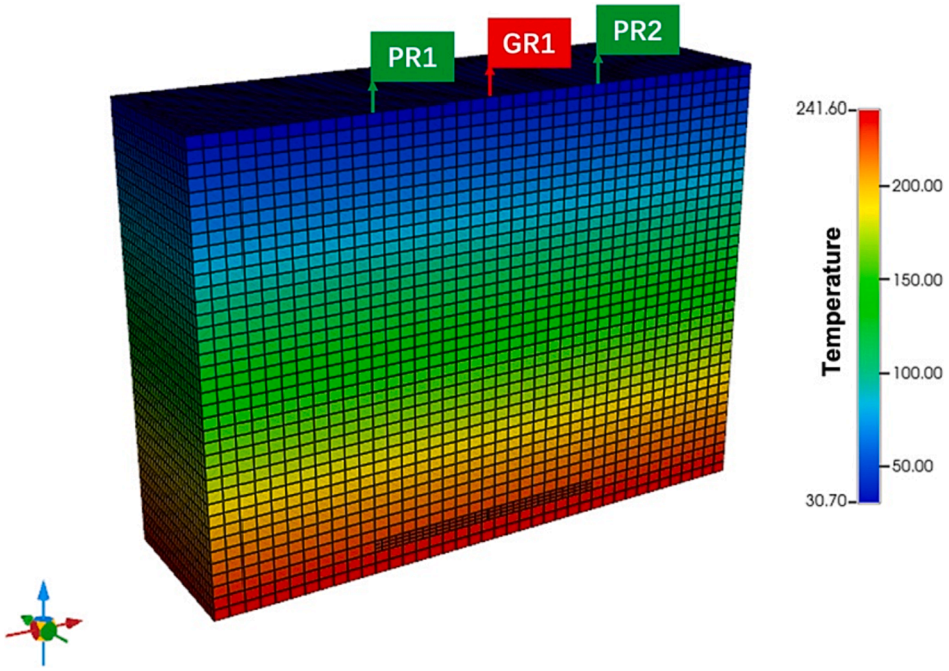


Fig. 2. Base numerical model.

algorithm [50,51]. A Gaussian Mixture Model is applied in a TPE optimization to find the values that can minimize an objective function [52]. It is highly performant and time-efficient compared with other typical algorithms, such as manual tuning, and grid and random search because its next generation of hyperparameters is determined according to its previous evaluation results [52,53]. In the TPE algorithm, a conditional probability from the Bayes theory is defined firstly, which can be written as Eq. (2) [53]

$$p(x|y) = \begin{cases} l(x) if y < y^* \\ g(x) if y \geq y^* \end{cases} \tag{2}$$

where x represents a hyperparameter; y denotes an objective function; $y^*$is a threshold for loss; $\ell(x)$is the density formed by using the observations such that the corresponding loss was less than the threshold; $g(x)$ is the density formed by using the remaining observations.

A hyperparameter configuration is determined according to the greatest Expected Improvement (EI). At each iteration in an optimization process, a new hyperparameter configuration is selected by EI and transferred back to calculate two new densities $\ell(x)$ and $g(x)$, and then the greatest EI in the next iteration is estimated until the number of iterations is reached [53,54]. Based on the parametrization of p (x,y) as $p(y)p(x|y)$ in the TPE algorithm, EI can be calculated by Eq. (3) [53].

$$EI_{y^*}(x) = \frac{\gamma y^* \ell(x) - \ell(x)\int_{-\infty}^{y^*} p(y)dy}{\gamma \ell(x) + (1-\gamma)g(x)} \propto \left(\gamma + \frac{g(x)}{\ell(x)}(1-\gamma)\right)^{-1} \tag{3}$$

$$\text{Subject to} \begin{cases} \gamma = p(y < y^*) \\ p(x) = \int p(x|y)p(y)dy = \gamma \ell(x) + (1-\gamma)g(x) \\ \int_{-\infty}^{y^*} (y^* - y)p(x|y)p(y)dy = \gamma y^* \ell(x) - \ell(x)\int_{-\infty}^{y^*} p(y)dy \end{cases}$$

2.2.3. Differential evolution (DE) algorithm

DE is a heuristic stochastic algorithm based on the ideas of a genetic algorithm (GA) [55]. It is achieved by simulating a biological evolutional process, and an individual that can adapt to an environment iteratively is determined [34]. Compared with GA, DE is adaptive for a continuous optimization problem and easier to determine its control parameters, and can converge globally and faster [42]. The DE based optimization is processed to compare an objective function. The population is installed in each iteration, and the process of mutation, crossover and selection is repeated until an optimum is located, or a prespecified termination criterion is satisfied [56]. The detailed descriptions of these steps are given as follows [42].

1. Population construction: Produce an initial population of N individuals in an n-dimensional search domain with upper and lower bounds for each parameter.
2. Initialization: Assign each parameter in each vector a value from within a prescribed range, which is written as Eq. (4) [42]

$$x_{j,i,0} = rand_j(0,1) \times \left(b_{j,U} - b_{j,L}\right) + b_{j,L} \tag{4}$$

where j and i represent the indices of the parameter and vector separately; $rand_j(0,1)$ returns a uniformly distributed random number from the range $[0,1)$; $b_{j,U}$ and $b_{j,L}$ indicate the upper and lower boundd representatively.

3. Mutation: create a new mutate vector by combining three randomly selected vectors, which is written as Eq. (5) [42].

$$V_{i,g} = X_{r_0,g} + F \times \left(X_{r_1,g} - X_{r_2,g}\right) \tag{5}$$

where $X_{r_0,g}$ is a base vector which is different from the target vector; $X_{r_1,g}$ and $X_{r_2,g}$ are two randomly selected vectors; F is a differential mutation factor.

Table 1
Reservoir properties of the base model.

Parameter	Value
Granite density,ρ	2623 kg/m^3
Original porosity,ϕ	2.49 %
Original permeability, ($kx = ky = 10 \times kz$)	0.26 mD
Granite heat conductivity,λ	3.0 $W/(m^{\circ}C)$
Granite specific heat,C_R	980 J /($kg^{\circ}C$)
Horizontal natural fracture spacing,L_{n1}	10 m
Vertical natural fracture spacing,L_{n2}	10 m
Initial pressure,p	$p = 1.01 \times 10^5 - 10000z$ (Pa)
Initial temperature,T	T = 25-0.057z (°C)

4. Crossover: Build trial vectors $U_{i,g}$ out of parameter values that have been copied from two different vectors and cross each vector with a mutant vector, which is written as Eq. (6) [42].

$$U_{i,g} = u_{j,i,g} = \begin{cases} v_{j,i,g} if \left(rand_j(0,1) \leq C_r or j = j_{rand}\right) \\ x_{j,i,g} otherwise \end{cases} \tag{6}$$

where C_r is a crossover probability to control the fractions of parameter values. It is used to determine a source that contributes to a given parameter by a comparison with the uniform random number r and j (0,1 If the chosen number is less than or equal to Cr the trial parameter is inherited from the mutant vector $V_{i,g}$; otherwise, a parameter in vector $X_{i,g}$ is selected.

5. Selection: Select a target vector in the next generation through a comparison between the trial vector $U_{i,g}$ and the target vector $X_{i,g}$. This process is written as Eq. (7) [42].

$$X_{i,g+1} = \begin{cases} U_{i,g} if \left(f\left(U_{i,g}\right) \leq f\left(X_{i,g}\right)\right) \\ X_{i,g} otherwise \end{cases} \tag{7}$$

3. TPE-based ANN surrogate models

3.1. Data preparation

Data samples for creating ANN models are collected from numerical models using the CMG STARS, which is a well-accepted numerical simulator for multi-dimensional fluid flow and heat transfer analyses in multi-phase, multi-component fluid mixtures in porous and fractured media [57]. The dual porosity method in CMG STARS can approximately model the fluid flow and heat transfer in the fractured porous media [58]. It models fractures based on the concept that fractures have large permeability and small porosity [59]. This simulator has been proven that it can provide high performance on different works of geothermal reservoirs [60–64]. Therefore, the CMG STARS is highly suitable for this study.

A 3D base model with a dimension of 2,000 × 1,000 × 3,800 m is customized to describe the Qiabuqia geothermal field, which is shown in Fig. 2. According to the previous study of the fracturing investigation, the height of the fractures created in the Qiabuqia granite formation is approximately 100 m [1]. Therefore, the enhanced reservoir in this study is from 3,500 m to 3,600 m. Three horizontal wells, including an injection well (GR1) and two production wells (PR1, PR2), are created with a horizontal section length of 600 m and a well spacing of 400 m in the enhanced reservoir. The spacing of natural fractures is set to 10 m according to the core analysis [22]. The modelling domain is discretized into 36,480 cells, of which the grid blocks of the enhanced reservoir are refined five times in the $\times$ and y directions, and three times in the z directions. Hydraulic fractures are then constructed in this enhanced reservoir domain with a fracture spacing of 50 m, a fracture half-length of 400 m, a height of 100 m, and an average conductivity of 10 mDm, which have been proven to be beneficial in the exploitation of the Qiabuqia geothermal field [1]. The detailed reservoir properties in this base model are shown in Table 1.

Table 2
Statistic values of the operational factors.

	Min value	Max value
Injection rate, (kg/s)	20	100
Injection temperature, (°C)	40	80
Hydraulic fracture half-length, (m)	250	600
Hydraulic fracture spacing, (m)	25	300
Well spacing, (m)	300	600

The initial temperature distribution of the modelling domain is based on the temperature logging curve of the GR1 well [1]. According to the logging data of well GR1, the initial reservoir pressure is assumed at $p = 1.01 \times 10^5 - 10000z(Pa)$ and the initial reservoir temperature is assumed at $T = 25 - 0.057z°C$ [22].

Finally, a steady stream of 60 Cwater is circularly injected into the enhanced reservoir at a rate of 40 kg/s and a production pressure of 37 MPa to obtain a stable production rate [22]. The geothermal productivities are calculated after a 20-year operating cycle, which can represent the potential for the extraction of the Qiabuqia geothermal field [1,22,26].

Based on the base model, 1,316 numerical models with different combinations of operating factors are constructed using the CMG CMOST software. Five parameters are determined as the properties to be optimized since they have effects on geothermal productivity, including the rate and temperature of circulating water, well spacing, hydraulic fracture spacing, and a half-length of hydraulic fractures [1,18,22,26,29]. The Latin hypercube method in CMOST randomly changes the values of these properties within different ranges. The data section ranges of these parameters are shown in Table 2. Four different geothermal productivities during a 20-year operation cycle are calculated, containing the average electrical power generation (W_a), the cumulative power generation (W_t), the injection pressure at the bottom of the well (P_{inj}), and the temperature of produced water (T_{pro}). The average electrical power generation (W_a) and the cumulative power generation (W_t) are calculated based on the geothermal electricity generation, which can be expressed as Eq. (8) [65].

$$W_e = 0.45Q\Delta H\left(1 - T_{rej}/T_{out}\right) \tag{8}$$

where Q(kg/s) is a water production rate of the whole system; ΔH(J/kg) represents an enthalpy change between the injected enthalpy and produced enthalpy; T_{rej}(K) is the rejection temperature that is set at 277.25 K according to the mean annual temperature in the Qiabuqia field [22] _ENREF_22; T_{out}(K) is the average temperature of the production wells.

In conclusion, a total of 1,316 data points are established, consisting of different values of the five working parameters that are treated as the input features in each ANN model, and the corresponding targets (W_a, W_t, p_{inj}, T_{pro}) which are the outputs in ANN models.

3.2. Data preprocessing

Data preprocessing is a critical step for achieving a high-performance ANN model [66]. In order to evaluate the generalized ability of ANN models, data samples are split into 70 % training samples and 30 % testing samples such that the training samples are used to train and determine the optimal ANN models and the testing samples are responsible for evaluating these models. In addition, the dimensions of the input features are quite different, which are shown in Table 2. This kind of dimension variance probably damages the predictive ability of an ANN model [67]. Therefore, a standardization operation is implemented in the data samples to transform features by scaling each feature to a given range between −1 and 1, which can be written as Eq. (9) [68].

$$X_{scaled} = X_{std} \times (Max - Min) + Min \tag{9}$$

$$subjecttoX_{std} = (X - X_{min})/(X_{max} - X_{min})$$

Table 3
Detailed descriptions of hyperparameters and their values range [68].

Hyperparameter	Description	Range
LR	The step size for weight updates	1e-5, 1e-4, 1e-3, 1e-2, 1e-1
AF	Activation function for the hidden layer	*'Sigmoid'*, *'ReLU'* and *'Tanh'*
OA	The solver for weight optimization	*'LBFGS'*, *'SGD'*, *'ADAM'*
N	Maximum number of iterations	500, 1000, 1500, 2000, 2500
HLS	The number of hidden layers and the corresponding neuron amounts.	(4,), (6,), …, (10,), (4, 2), (4, 4), …, (10, 20), (4, 2, 2), (4, 4, 2), …, (10, 20, 2)

where X denotes the data samples; X_{std} is the standard deviation of data samples; *Max* and *Min* represent the upper and lower values of a given range, which are 1 and −1 in this work; X_{max} and X_{min}, respectively, means the maximum and minimum values of data samples.

3.3. Evaluation matrix

Three statistical indicators are developed to quantitatively evaluate the performance of the ANN models, including a root mean squared error (*RMSE*), a mean absolute error (*MAE*), and a coefficient of determination (R^2). Among them, *RMSE* and *MAE* are positive numbers and their lower values demonstrate higher prediction precision, while R^2 is generally in the range of 0 and 1 and an ANN model with better performance can present a higher R^2 value. They are calculated as Eqs. 10–12 representatively [68].

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2} \tag{10}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \widehat{y_i}| \tag{11}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \widehat{y_i})^2}{\sum_{i=1}^{n}(y_i - \overline{y_i})^2} \tag{12}$$

where y_i represents a prediction result and $\widehat{y_i}$ denotes the corresponding actual value; $\overline{y_i}$ is the mean of the actual values; n is the number of samples.

3.4. TPE-based model design

Four basic ANN models are constructed to predict the average electrical power generation (W_a), total power generation (W_t), injection pressure (P_{inj}), and the temperature of produced water (T_{pro}). To update a basic ANN model to the one with the best performance, hyperparameter tuning is necessary to operate. Previous studies showed that a learning rate (*LR*), an activation function (*AF*), an optimizer algorithm (*OA*) and a structure of hidden layers (*HLS*) affect the performance of an ANN model [69–71]. Besides, a maximum iteration number (*N*) mainly controls the simulation time [72]. Therefore, a *K*-fold cross-validation integrated with the TPE algorithm is implemented in the training samples to determine the best collective choices of these five hyperparameters. The *K*-fold cross-validation method can fully utilize the training samples which are divided into *K* folds and each fold can be used to train and validate. Specifically, the training samples are divided into a training set and a validating set, and the optimal ANN model is determined under the TPE algorithm when its performance in the validating set is best.

In the TPE-based hyperparameter tuning process, the optimal ANN models are determined by iteratively selecting a configuration of hyperparameters within different ranges until 100 iterations are

Table 4
The generalization ability of ANN models under different optimization methods.

Methods	TPE			Grid search			Random search		
	R^2	*RMSE*	*MAE*	R^2	*RMSE*	*MAE*	R^2	*RMSE*	*MAE*
ANN_{W_a}	0.999	0.067 MW	0.037 MW	0.999	0.071 MW	0.037 MW	0.997	0.132 MW	0.095 MW
ANN_{W_t}	0.998	18.226GWh	11.883GWh	0.996	24.132GWh	1.491GWh	0.995	28.577GWh	18.475GWh
$ANN_{P_{inj}}$	0.996	0.315 MPa	0.212 MPa	0.993	0.409 MPa	0.311 MPa	0.994	0.395 MPa	0.297 MPa
$ANN_{T_{pro}}$	0.996	1.518C	0.890℃	0.996	1.569℃	0.887℃	0.993	2.055℃	1.308℃

reached, and then the testing samples are transferred into these optimal models to measure their predictive ability. The detailed setting of hyperparameters is shown in Table 3. Specifically, *HLS* can be explained through an example of (4, 6, 2), meaning that there are 3 hidden layers with 4 neurons in the first layer, 6 neurons in the second layer, and 2 neurons in the last layer.

4. DE-based optimization model

In a DE optimization process, an objective model is created based on the resulting ANN surrogate models to customize the levelized cost of electricity (LCOE). LOCE is a benchmark to measure the cost-effectiveness of geothermal energy generation by comparing the unit costs over an economic life [73–75]. It is calculated by dividing the total investments by the cumulative electricity generation [1,21].

For an EGS project, the total costs can be divided into a surface cost (C_{surf}) and a subsurface cost (C_{sub}). Specifically, C_{surf} includes the cost of geothermal field exploration and equipment installation, and C_{sub} contains the cost of drilling, reservoir development and reservoir maintenance [76].

According to the reported data, an investment in reservoir exploration of the Qiabuqia geothermal field is about 4.5 M$ [1]. Besides, the initial cost of equipment installation increases with the power capacity but the unit cost is inversely proportional to the capacity due to the economy of scale [76] and the unit capital cost is estimated to be 2000 $/kW [77]. Therefore, the cumulative surface investments can be expressed in Eq. (13) [77].

$$C_{surf} = 4.5 + \frac{2000 \times W_a}{10^3} \tag{13}$$

For the subsurface cost part, the drilling costs for a three-horizontal-well system can be calculated by Eq. (14), and the detailed drilling costs in the Qiabuqia field are expected: $H_v = 11,000$m, $H_h = 1,800$m, $P_v = \$600$, $P_h = \$1,500$ [1]. The reservoir development in this field can be split into a logging survey and a hydraulic fracturing treatment. It is estimated to be 0.45 M$ for the investment of high-precision logging at a depth of 3,700 m [1]. The cost of fracturing operation is evaluated to be 0.45 M$, which has been proven to undertake the operations of 300–600 m well spacing and 25–300 m fracture spacing [1]. The unit cost for the reservoir operation and maintenance is generally in an inverse ratio to the installed capacity, which is normally estimated in a trend from 20 $/MWh for a 5 MW plant to 14 $/MWh for a 150 MW plant [76,77]. Therefore, the unit cost of reservoir operation and maintenance can be expressed in Eq. (15) [76].

$$C_{drill} = (H_v \times P_v + H_h \times P_h) \times 10^{-6} \tag{14}$$

$$C_{O\&M} = 20 \times exp(-0.0025 \times (W_a - 5)) \tag{15}$$

Consequently, the total costs for the Qiabuqia EGS power plant can be calculated in Eq. (16) and the LCOE can be written in Eq. (17).

$$C_{total} = C_{surf} + C_{sub} = 15.21 + \frac{2000 \times W_a}{10^3} + \frac{W_t \times C_{O\&M}}{10^3} \tag{16}$$

$$LCOE = \left(15.21 + \frac{2000 \times W_a}{10^3} + \frac{W_t \times C_{O\&M}}{10^3}\right) \Big/ W_t \tag{17}$$

Furthermore, two constraints should be considered in an EGS operation. The first is that the temperature of the produced circulating water (T_{pro}) should be higher than 105.36C to satisfy the demand for converting heat energy into electricity effectively, which can be written in Eq. (18) [22]. The second is that the injection pressure at the bottom of a well (P_{inj}) should be less than the minimum principal stress (σ_h) to prevent slippage from occurring, which can be expressed in Eq. (19) [22].

$$T_{pro} > 105.36 \tag{18}$$

$$P_{inj} < \sigma_h = 60\text{MPa} \tag{19}$$

Based on the LCOE objective function and two constraint equations, the DE algorithm is operated to optimize the geothermal system, aiming to determine the optimal values of operating parameters which can bring the best economical performance. In the DE optimization, the population number of individuals is set to 200, the maximum number of generations is set to 1,000, and the values of a scaling factor and probability of crossover are 0.5 and 0.7 separately.

5. Results and discussion

5.1. Performance of ANN models

Four optimal ANN models are obtained after a TPE-based hyperparameter tunning. Among them, the hyperparameter selections of ANN_{W_a} and $ANN_{P_{inj}}$ are the same, where the learning rate (*LR*) is 1e-2, the activation function (*AF*) is *'Tanh'*, the optimizer algorithm (*OA*) is *'LBFGS'*, the maximum iteration number (*N*) is 1,500, and the structure of hidden layers (*HLS*) consists of three hidden layers with a total of 46 neurons. Meanwhile, ANN_{W_t} and $ANN_{T_{pro}}$ achieve the same *OA* and *N* values as ANN_{W_a} and $ANN_{P_{inj}}$. Differently, *LR* in ANN_{W_t} is 1e-4, *AF* is *'ReLU'* and *HLS* contains three hidden layers with a total of 50 neurons. $ANN_{T_{pro}}$ acquires an *LR* of 1e-2, an *AF* of *'Sigmoid'*, and an *HLS* of three hidden layers with 30 neurons. In order to evaluate the performance of the TPE-based optimization method, its performance is compared with that of two conventional approaches, the grid search and random search methods [78]. By comparing the generalization ability of the models (the model performance in a testing set) under these three optimization methods, which is shown in Table 4, the TPE-based ANN models all present the best predictive ability since they demonstrate the highest R^2 values and the lowest *RMSE* and *MAE* values. Furthermore, the optimization time based on TPE for each ANN model is around 60 s, while that of the grid search and random search is more than 5,400 s. Therefore, the TPE approach is highly recommended to optimize an ANN model since it can not only provide a hyperparameter configuration with better prediction performance but also can highly improve the efficiency of the optimization.

After determining the optimal models, testing samples are used to evaluate their generalized abilities, and the performance of these models in different sets is discussed, which is shown in Table 5.

Form the results, it can be estimated that these ANN models all present a strong predictive ability according to the results in the testing set. The R^2 values are all higher than 0.996, showing the high accuracy between the prediction results and the actual targets. In addition, their

Table 5
Statistical indicators values of four optimal ANN models in different data sets.

Evaluation indicators	ANN_{W_a}	ANN_{W_t}	$ANN_{P_{inj}}$	$ANN_{T_{pro}}$
Training set				
R^2	1.000	0.998	0.998	0.999
MAE	0.029 MW	12.468 GWh	0.173 MPa	0.601 ℃
RMSE	0.041 MW	17.813 GWh	0.235 MPa	0.846 ℃
Validating set				
R^2	0.998	0.996	0.996	0.992
MAE	0.046 MW	15.806 GWh	0.231 MPa	0.961 ℃
RMSE	0.091 MW	24.859 GWh	0.322 MPa	1.865 ℃
Testing set				
R^2	0.999	0.998	0.996	0.996
MAE	0.037 MW	11.883 GWh	0.212 MPa	0.890 ℃
RMSE	0.067 MW	18.226 GWh	0.315 MPa	1.518 ℃

RMSE and *MAE* values demonstrate a low variance between their prediction points and the actual samples. Compared with the values of the actual targets (the *y*-axes in Fig. 3), the variances demonstrated from ANN_{W_a} and ANN_{W_t} are less than 3.5 %, and are around 1 % in $ANN_{P_{inj}}$ and $ANN_{T_{pro}}$. The predictive ability of these models can also be evaluated from the scatter plots of the predictions of W_a, W_t, P_{inj} and T_{pro} and their corresponding actual points, which are shown in Fig. 3. The scatters are basically located at a 45-degree line, which means that they are highly consistent. On the other hand, these models bring strong stability since the performance differences between different sets are small. The value changes between the training set, the validating set and the testing set in the R2 value are less than 2.0 %, and those in *RMSE* and *MAE* values are less than 4.0 % and 3.5 % representatively. Importantly, the simulation time of each ANN model to predict 395 data points is less than 10 s, which is highly time-saving compared with around 165,900 s by the numerical simulation method. Therefore, the resulting ANN models can be the surrogate models to predict Wa Wt Pinj and Tpro.

5.2. Results of DE optimization

After a DE optimization, the optimal configurations of operational parameters are determined, where the temperature of the injected fluid is 53C, the rate of circulating water is 99Lg/s the well spacing is 550 m, and the spacing and half-length of hydraulic fractures are 25 m and 478 m separately. In order to evaluate the precision of the ANN models and the performance of the DE optimization, a total of 2,150 random cases are developed to compare with the optimal case. These random cases are generated with different randomly selected values of operating factors by the CMOST software and calculated through the STARS software, which can comprehensively consider different operating conditions in the Qiabuqia geothermal field. Fig. 4 shows the LCOE values of random cases and the determined optimal case under the corresponding cumulative electricity generation. The red scatter demonstrates the optimal

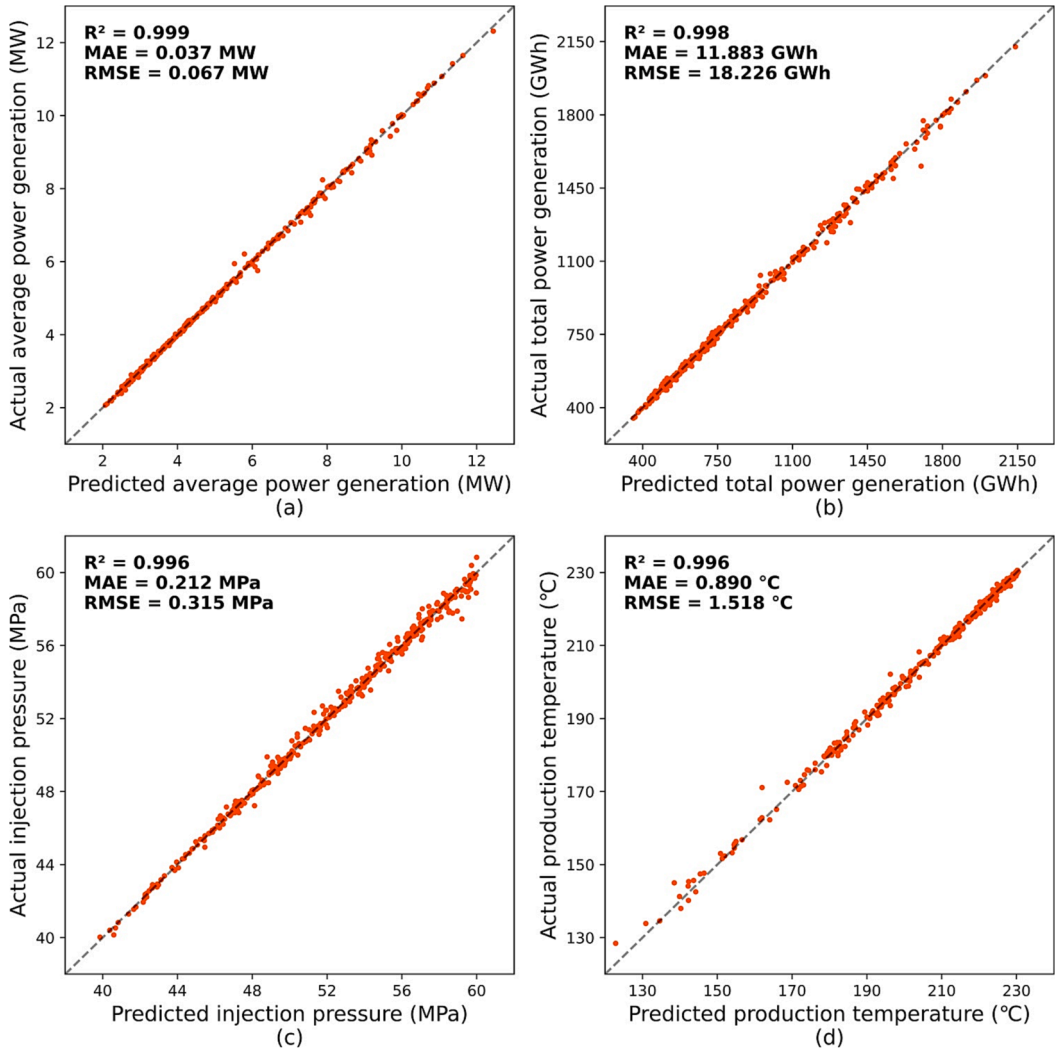


Fig. 3. Scatter plots of different targets and their corresponding predictions: (a) average electricity power generation, (b) total electricity power generation, (c) injection pressure at the bottom hole of the well, and (d) the temperature of the produced water.

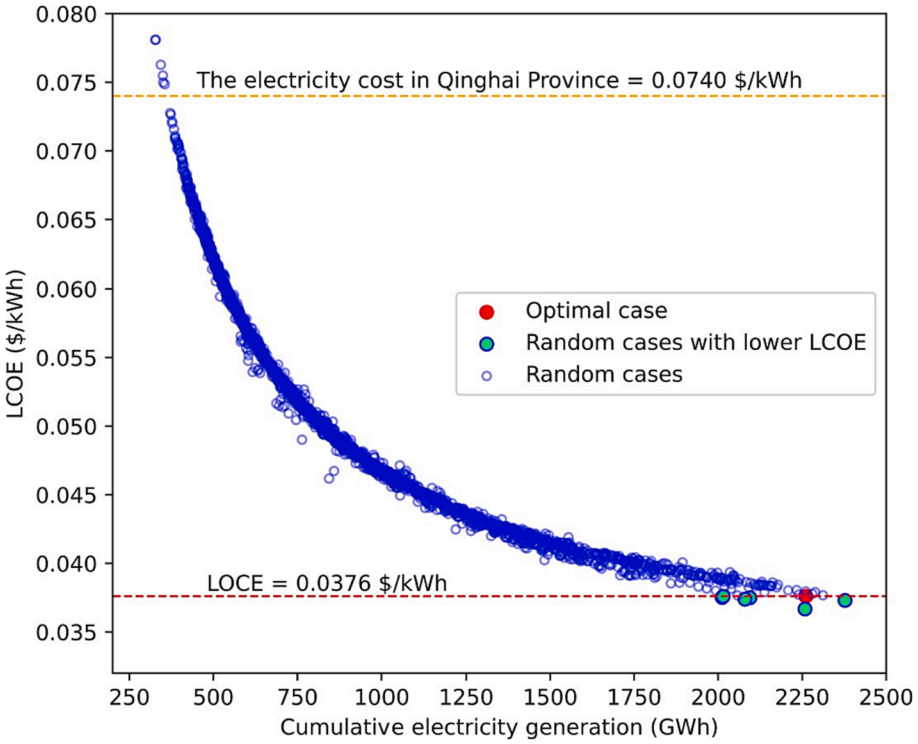


Fig. 4. LCOE values of random cases and the optimal case.

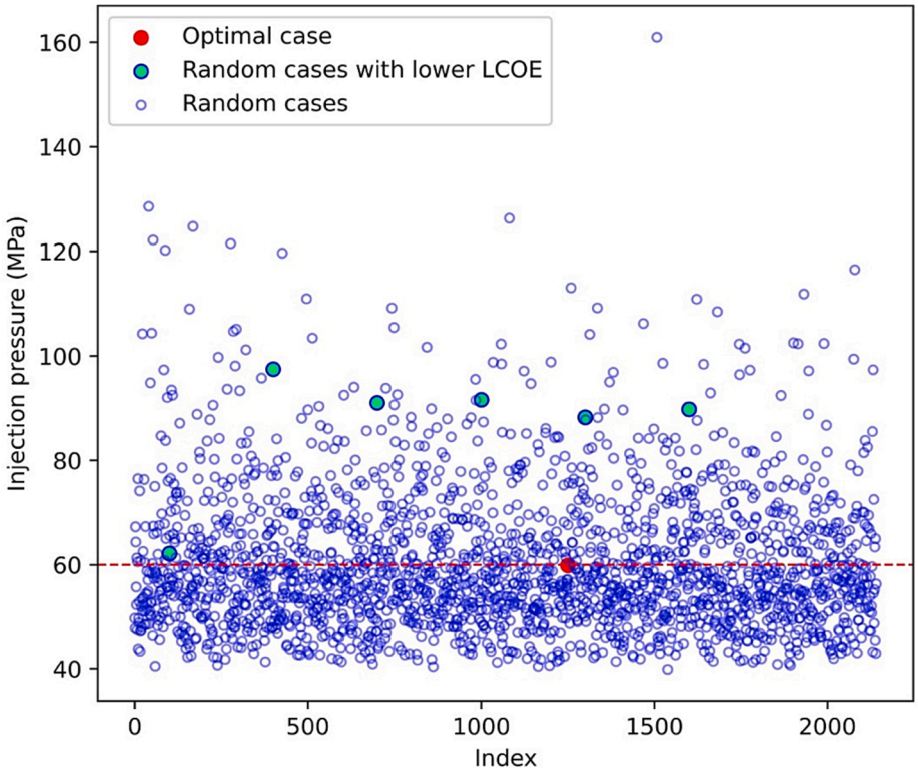


Fig. 5. Injection pressure of random cases and the optimal case.

case, the blue-hollow circles represent the random cases, and green-filled dots denote the random cases with a lower LCOE than the optimal case.

It is inferred that the LCOE generally decreases as the cumulative electricity generation increases. The resulting optimal case acquires an LCOE of 0.0376 $/kWh, which is well below the industrial electricity cost in Qinghai Province (0.074 $/kWh). Besides, this LCOE is lower than that of the random cases except for six cases of them. However, the operating parameters setting of these six cases is non-compliance with the constraint of field operation since they obtain a higher injection

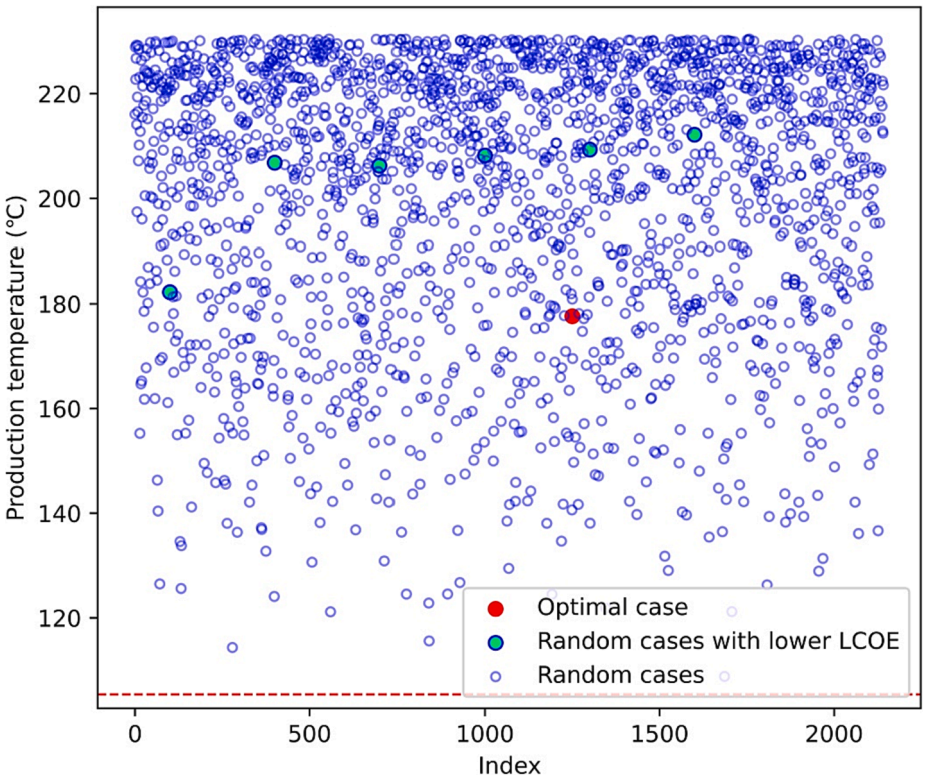


Fig. 6. Production temperature of random cases and the optimal case.

pressure at the well bottom hole than the minimum principal stress (60 MPa), which is shown in Fig. 5. Comparatively, the optimal case brings a reasonable injection pressure of 59.2 MPa. In addition, from Fig. 6, the optimal case can produce a steady stream of 178C water, which heavily exceeds the minimum requirement for generating electricity (105.36C), demonstrating the potential for long-term electricity generation. Therefore, the results show that the ANN-based DE optimization can effectively identify the optimal operating factors for economical geothermal energy extraction. Importantly, this optimization process only takes approximately 25 s, which is highly efficient compared with the numerical simulation method, which needs around 903,000 s for 2,150 numerical models. Therefore, the optimization efficiency can also be significantly promoted through the ANN-based DE optimization method.

5.3. Discussion

The proposed ANN models show their promising predictive ability to simulate the in-depth performance of a geothermal field. Although the ANN has been proven its ability in the geothermal field, another important part of an ANN model, the optimization process, is normally simplified in previous studies [25,36–38,40,79,80]. The huge consideration of a machine learning application is mainly due to its high efficiency, but it is not practical if the operation time of the optimization is not considered in the computation time. Therefore, the TPE algorithm is used in this study to design ANN models and its performance is compared with the conventional optimization methods, including the gird search and random search approaches. The results indicate that the TPE-based ANN models can improve the optimization efficiency compared to these conventional methods and can provide high prediction accuracy. More importantly, compared to the complex process of building a physical-driven model, the establishment of the proposed data-driven models is highly efficient and only takes around 240 s. Besides, it takes less than 10 s to predict 395 data points for each TPE-based ANN model, while the numerical simulation method needs about 165,900 s. Therefore, the proposed TPE-based ANN models also demonstrate a greatly effective computational process, showing its promising potential to simulate the performance of a geothermal system in a system design stage.

In optimizing a geothermal system, the ANN-based DE method is proposed and the results show that the LCOE value (0.0376 $/kWh) of the determined optimal geothermal system is half of the local electricity cost. To evaluate the feasibility of this result, 2,150 numerical models are created to widely consider different operation schemes, and their LCOE values are all higher than this 0.0376 $/kWh LCOE. For the Qiabuqia geothermal field, previous studies normally determined an optimal geothermal system by comparing the results between several manually created cases [1,26], and the 0.0376 $/kWh LCOE is also highly lower than their result (0.0666 $/kWh) [1] In addition, the optimization time of the proposed ANN-based DE approach is around 25 s, while the calculation time of a single numerical model is about 420 s. Therefore, the proposed optimization approach can more effectively design a geothermal system and can be also directly applied to other energy systems, thereby facilitating the in-depth development of optimization design.

Although the ANN-based DE optimization approach proposed in this study shows great advantages and potential in system operation and proactive design, it could be further improved in the following aspects: (1) to improve a machine learning algorithm to achieve accurate predictions with smaller datasets; (2) to be utilized in other geothermal fields or other energy systems in exploring their optimal operation strategies, and (3) to incorporate field uncertainties and quantify their impacts, such as the reservoir heterogeneity and permeability anisotropy.

6. Conclusions

This paper presents an ANN-based DE optimization framework to determine the optimal field operations in the Qiabuqia geothermal field. The developed ANN models are used to substitute the numerical models

to calculate the geothermal production of this field. Then the ANNs are integrated into a DE-based optimization procedure to determine the optimal system design over a 20-year life span.

Results prove the effectiveness and accuracy of the proposed data-driven method in predicting the off-design performance of the Qiabuqia geothermal field, which greatly reduces the calculation time by 16,590 times compared to the numerical simulation method while maintaining the accuracy within 0.8 %. Besides, compared to the conventional optimization methods (grid search and random search), the TPE algorithm significantly improves the optimization efficiency by 90 times. Therefore, the proposed TPE-based ANN approach makes it possible to consider the performance of the geothermal system in its system design stage.

Then the DE-based optimization is carried out for the design of the geothermal system considering two field operational constraints. Results indicate that, compared to the conventional numerical simulation method, the ANN-based DE optimization method can achieve a lower LCOE value of 0.0376 $/kWh and highly reduce the computational time of optimization by 36,000 times. Besides, the 0.0376 $/kWh LCOE value of the resulting optimal geothermal system is around 50 % of a local electricity cost, which can provide a promising profit in operating this field. Therefore, the proposed method shows its effectiveness and the huge application potential in the geothermal system and it can be also directly used in other energy systems to facilitate the in-depth development of their optimization design.

CRediT authorship contribution statement

Zhenqian Xue: Conceptualization, Methodology, Software, Validation, Investigation, Data curation, Writing – original draft, Writing – review & editing. **Shuo Yao:** Validation, Investigation, Writing – review & editing. **Haoming Ma:** Investigation, Writing – review & editing. **Chi Zhang:** Validation, Writing – review & editing. **Kai Zhang:** Data curation, Writing – review & editing. **Zhangxin Chen:** Conceptualization, Writing – review & editing, Supervision.

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.

Acknowledgments

This research has been made possible by contributions from the Natural Sciences and Engineering Research Council (NSERC)/Energy Simulation Industrial Research Chair in Reservoir Simulation, the Alberta Innovates (iCore) Chair in Reservoir Modeling, the China Scholarship Council (CSC), and the Energy Simulation/Frank and Sarah Meyer Collaboration Centre.

References

[1] Lei Z, et al. Electricity generation from a three-horizontal-well enhanced geothermal system in the Qiabuqia geothermal field, China: Slickwater fracturing treatments for different reservoir scenarios. Renew Energy 2020;145:65–83.
[2] Abanades JC, et al. On the climate change mitigation potential of CO2 conversion to fuels. Energ Environ Sci 2017;10(12):2491–9.
[3] Boden, T.A., G. Marland, and R.J. Andres, *Global, regional, and national fossil-fuel CO2 emissions.* Carbon dioxide information analysis center, Oak ridge national laboratory, US department of energy, Oak Ridge, Tenn., USA doi, 2009. **10**.
[4] Wu C, et al. Performance assessment and optimization of a novel geothermal combined cooling and power system integrating an organic flash cycle with an ammonia-water absorption refrigeration cycle. Energ Conver Manage 2021;227.
[5] C. Davenport, Nations Approve Landmark Climate Accord in Paris. 201Dec 12; Available from: https://www.nytimes.com/2015/12/13/world/europe/climate-changeaccord-paris.html.
[6] Independent Group of Scientists Appointed by the SecretaryGeneral. Global Sustainable Development Report 2019: the Future Is Now Science for Achieving Sustainable Development 2019; Available from: https://sustainabledevelopment.un.org/content/documents/24797GSDR_report_2019.pdf.
[7] Zhang S, et al. Well placement optimization for large-scale geothermal energy exploitation considering nature hydro-thermal processes in the Gonghe Basin, China. J Clean Prod 2021:317.
[8] Wang Z, et al. Hydrate deposition prediction model for deep-water gas wells under shut-in conditions. Fuel 2020;275:117944.
[9] Lund JW, Freeston DH, Boyd TL. Direct application of geothermal energy: 2005 worldwide review. Geothermics 2005;34(6):691–727.
[10] Cao W, Huang W, Jiang F. A novel thermal–hydraulic–mechanical model for the enhanced geothermal system heat extraction. Int J Heat Mass Transf 2016;100:661–71.
[11] Liao Y, et al. Geothermal exploitation and electricity generation from multibranch U-shaped well–enhanced geothermal system. Renew Energy 2021;163:2178–89.
[12] Mosaffa A, Zareei A. Proposal and thermoeconomic analysis of geothermal flash binary power plants utilizing different types of organic flash cycle. Geothermics 2018;72:47–63.
[13] Ashraf MA, et al. Proposal and comprehensive analysis of an innovative CCP plant based on an internal integration of double flash power system and ejector refrigeration cycle. Energ Conver Manage 2020;203:112232.
[14] Lu S-M. A global review of enhanced geothermal system (EGS). Renew Sustain Energy Rev 2018;81:2902–21.
[15] Lund JW, et al. Characteristics, development and utilization of geothermal resources–a Nordic perspective. Episodes J Int Geosci 2008;31(1):140–7.
[16] Rühl, C., *Bp statistical review of world energy, june 2008.* London: BP, 2008.
[17] Brown DW, Duchane DV. Scientific progress on the Fenton Hill HDR project since 1983. Geothermics 1999;28(4–5):591–601.
[18] Song G, et al. An integrated multi-objective optimization method to improve the performance of multilateral-well geothermal system. Renew Energy 2021;172:1233–49.
[19] Gudmundsdottir H, Horne RN. Prediction modeling for geothermal reservoirs using deep learning. 45th workshop on geothermal reservoir engineering. Stanford, California: Stanford University; 2020.
[20] China, T.S.C.I.O.o.t.P.s.R.o. Responding to Climate Change: China's Policies and Actions. 2021 October; Available from: http://www.scio.gov.cn/zfbps/32832/Document/1715506/1715506.htm#:~:text=In%202015%2C%20China%20set%20its,new%20NDC%20targets%20and%20measures.
[21] Xu T, et al. Prospects of power generation from an enhanced geothermal system by water circulation through two horizontal wells: A case study in the Gonghe Basin, Qinghai Province, China. Energy (Oxford) 2018;148:196–207.
[22] Lei Z, et al. Exploratory research into the enhanced geothermal system power generation project: The Qiabuqia geothermal field, Northwest China. Renew Energy 2019;139:52–70.
[23] Zhang C, et al. Parametric study of the production performance of an enhanced geothermal system: A case study at the Qiabuqia geothermal area, northeast Tibetan plateau. Renew Energy 2019;132:959–78.
[24] Lee, K., *Classification of geothermal resources-an engineering approach.* 1996, Geothermal Institute, The University of Auckland, Auckland, NZ.
[25] Yilmaz C, Sen O. Thermoeconomic analysis and artificial neural network based genetic algorithm optimization of geothermal and solar energy assisted hydrogen and power generation. Int J Hydrogen Energy 2022;47(37):16424–39.
[26] Song G, et al. Multi-objective Optimization of Geothermal Extraction from the Enhanced Geothermal System in Qiabuqia Geothermal Field, Gonghe Basin. Acta Geol Sin– English Edition 2021;95(6):1844–56.
[27] Gong F, et al. Evaluation of geothermal energy extraction in Enhanced Geothermal System (EGS) with multiple fracturing horizontal wells (MFHW). Renew Energy 2020;151:1339–51.
[28] Zhang L, et al. Well-Placement Optimization in an Enhanced Geothermal System Based on the Fracture Continuum Method and 0–1 Programming. Energies 2019;12(4).
[29] Guo L-L, et al. Parameter sensitivity analysis and optimization strategy research of enhanced geothermal system: A case study in Guide Basin, Northwestern China. Renew Energy 2020;153:813–31.
[30] Liu Y, et al. Numerical Simulation and Design Optimization of Large-Scale Geothermal Production Based on a Multiwell Layout in Xianxian Geothermal Field. Lithosphere 2022. 2021(Special 5.
[31] Gan Q, Elsworth D. Production optimization in fractured geothermal reservoirs by coupled discrete fracture network modeling. Geothermics 2016;62:131–42.
[32] Liu G, et al. Impacts of fracture network geometries on numerical simulation and performance prediction of enhanced geothermal systems. Renew Energy 2021;171:492–504.
[33] Zinsalo JM, Lamarche L, Raymond J. Design and optimization of multiple wells layout for electricity generation in a multi-fracture enhanced geothermal system. Sustainable Energy Technol Assess 2021;47.
[34] Wang, N., et al., *Deep learning based closed-loop optimization of geothermal reservoir production.* arXiv preprint arXiv:2204.08987, 2022.
[35] Huang Z, Chen Z. Comparison of different machine learning algorithms for predicting the SAGD production performance. J Pet Sci Eng 2021;202.
[36] Ishitsuka K, et al. Resistivity-Based Temperature Estimation of the Kakkonda Geothermal Field, Japan, Using a Neural Network and Neural Kriging. IEEE Geosci Remote Sens Lett 2018;15(8):1154–8.

[37] Hu S, et al. Thermo-economic optimization of the hybrid geothermal-solar power system: A data-driven method based on lifetime off-design operation. Energ Conver Manage 2021;229.
[38] Senturk Acar M. Multi-stage artificial neural network structure-based optimization of geothermal energy powered Kalina cycle. J Therm Anal Calorim 2020;145(3):829–49.
[39] Bassam A, et al. Estimation of static formation temperatures in geothermal wells by using an artificial neural network approach. Comput Geosci 2010;36(9):1191–9.
[40] Pei H, et al. Long-term thermomechanical displacement prediction of energy piles using machine learning techniques. Renew Energy 2022;195:620–36.
[41] Xiao F, et al. Performance enhancement of horizontal extension and thermal energy storage to an abandoned exploitation well and satellite LNG station integrated ORC system. Appl Therm Eng 2022;214:118736.
[42] Price, K., R.M. Storn, and J.A. Lampinen, *Differential evolution: a practical approach to global optimization*. 2006: Springer Science & Business Media.
[43] Bardenet, R., et al. *Collaborative hyperparameter tuning*. in *International conference on machine learning*. 2013. PMLR.
[44] Wang, S.-C., *Artificial Neural Network*, in *Interdisciplinary Computing in Java Programming*, S.-C. Wang, Editor. 2003, Springer US: Boston, MA. p. 81-100.
[45] Agatonovic-Kustrin S, Beresford R. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. J Pharm Biomed Anal 2000;22(5):717–27.
[46] Ayyadevara VK. Pro machine learning algorithms. Berkeley, CA, USA: Apress; 2018.
[47] Lippmann R. An introduction to computing with neural nets. IEEE ASSP Mag 1987;4(2):4–22.
[48] Foroud T, Seifi A, AminShahidi B. Assisted history matching using artificial neural network based global optimization method – Applications to Brugge field and a fractured Iranian reservoir. J Pet Sci Eng 2014;123:46–61.
[49] Ayyadevara, V.K., *Pro Machine Learning Algorithms: A Hands-On Approach to Implementing Algorithms in Python and R*. 2018, Berkeley, CA: Berkeley, CA: Apress L. P.
[50] Bergstra J, et al. Algorithms for hyper-parameter optimization. Adv Neural Inf Process Syst 2011;24.
[51] Bergstra, J., D. Yamins, and D. Cox. *Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures*. in *International conference on machine learning*. 2013. PMLR.
[52] Sameen MI, Pradhan B, Lee S. Self-learning random forests model for mapping groundwater yield in data-scarce areas. Nat Resour Res 2019;28(3):757–75.
[53] Rong G, et al. Comparison of Tree-Structured Parzen Estimator Optimization in Three Typical Neural Network Models for Landslide Susceptibility Assessment. Remote Sens (Basel) 2021;13(22).
[54] Bergstra, J., D. Yamins, and D.D. Cox. Hyperopt: A python library for optimizing the hyperparameters of machine learning algorithms. In Proceedings of the 12th Python in science conference. 2013. Citeseer.
[55] Storn R, Price K. Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. J Glob Optim 1997;11(4):341–59.
[56] Kumar, B.V., D. Oliva, and P.N. Suganthan, *Differential evolution: from theory to practice*. Studies in Computational Intelligence; Volume 1009. 2022, Singapore: Springer.
[57] Asai P, et al. Effect of different flow schemes on heat recovery from Enhanced Geothermal Systems (EGS). Energy 2019;175:667–76.
[58] Warren J, Root PJ. The behavior of naturally fractured reservoirs. Soc Pet Eng J 1963;3(03):245–55.
[59] Torres, F., et al. Comparison of dual porosity dual permeability with embedded discrete fracture model for simulation fluid flow in naturally fractured reservoirs. in 54th US Rock Mechanics/Geomechanics Symposium. 2020. OnePetro.
[60] Hofmann H, et al. Potential for enhanced geothermal systems in Alberta, Canada. Energy 2014;69:578–91.
[61] Ansari E, Hughes R, White CD. Statistical modeling of geopressured geothermal reservoirs. Comput Geosci 2017;103:36–50.
[62] Ansari E, Hughes R, White CD. Modeling a new design for extracting energy from geopressured geothermal reservoirs. Geothermics 2018;71:339–56.
[63] Asai P, et al. Performance evaluation of enhanced geothermal system (EGS): Surrogate models, sensitivity study and ranking key parameters. Renew Energy 2018;122:184–95.
[64] Asai P, et al. Efficient workflow for simulation of multifractured enhanced geothermal systems (EGS). Renew Energy 2019;131:763–77.
[65] Zeng Y-C, et al. Numerical simulation of heat production potential from hot dry rock by water circulating through a novel single vertical fracture at Desert Peak geothermal field. Energy 2013;63:268–82.
[66] Zheng A, Casari A. Feature engineering for machine learning: principles and techniques for data scientists. O'Reilly Media, Inc.; 2018.
[67] Duboue P. The art of feature engineering: essentials for machine learning. Cambridge University Press; 2020.
[68] Pedregosa F, et al. Scikit-learn: Machine learning in Python. J Mach Learn Res 2011;12:2825–30.
[69] Negash BM, Yaw AD. Artificial neural network based production forecasting for a hydrocarbon reservoir under water injection. Pet Explor Dev 2020;47(2):383–92.
[70] He W, et al. Using of Artificial Neural Networks (ANNs) to predict the thermal conductivity of Zinc Oxide-Silver (50%–50%)/Water hybrid Newtonian nanofluid. Int Commun Heat Mass Transfer 2020;116:104645.
[71] Naqvi SR, et al. Pyrolysis of high-ash sewage sludge: Thermo-kinetic study using TGA and artificial neural networks. Fuel 2018;233:529–38.
[72] Tugcu A, Arslan O. Optimization of geothermal energy aided absorption refrigeration system—GAARS: A novel ANN-based approach. Geothermics 2017;65:210–21.
[73] Ouyang X, Lin B. Levelized cost of electricity (LCOE) of renewable energies and required subsidies in China. Energy Policy 2014;70:64–73.
[74] Branker K, Pathak MJM, Pearce JM. A review of solar photovoltaic levelized cost of electricity. Renew Sustain Energy Rev 2011;15(9):4470–82.
[75] OECD, N., *Projected Costs of Generating Electricity: 2010 Edition*. 2010, OECD/IEA.
[76] Sanyal, S.K. *Cost of geothermal power and factors that affect it*. in *Proceedings of World Geothermal Congress*. 2005.
[77] Chamorro CR, et al. World geothermal power production status: Energy, environmental and economic study of high enthalpy technologies. Energy 2012;42(1):10–8.
[78] Toal DJ, Bressloff NW, Keane AJ. Kriging hyperparameter tuning strategies. AIAA J 2008;46(5):1240–52.
[79] Yang F, et al. Artificial neural network based prediction of reservoir temperature: A case study of Lindian geothermal field, Songliao Basin, NE China. Geothermics 2022:106.
[80] Mehrenjani JR, Gharehghani A, Sangesaraki AG. Machine learning optimization of a novel geothermal driven system with LNG heat sink for hydrogen production and liquefaction. Energ Conver Manage 2022;254:115266.