Predicting Hydropower Production Using Deep Learning CNN-ANN Hybridized with Gaussian Process Regression and Salp Algorithm

Mohammad Ehtearm[1] · Hossein Ghayoumi Zadeh[2] · Akram Seifi[3] · Ali Fayazi[2] · Majid Dehghani[4]

Abstract

The hydropower industry is one of the most important sources of clean energy. Predicting hydropower production is essential for the hydropower industry. This study introduces a hybrid deep learning model to predict hydropower production. Statistical methods are unsuitable for modeling hydropower production because their accuracy depends on seasonal and periodic fluctuations. For accurate predictions, deep learning models can capture daily, weekly, and monthly patterns. Since ANNs may not capture latent and nonlinear patterns, we use deep learning models to predict hydropower production. We used Convolutional Neural Network-Multilayer Perceptron-Gaussian Process Regression (CNNE-MUPE-GPRE) to extract key features and predict outcomes. The main advantages of the hybrid model are the quantification of production uncertainty, the accurate prediction of hydropower production, and the extraction of features from input data. We use a binary SSOA to select optimal input scenarios. The new model is benchmarked against the long short term memory neural network (LSTM), Bi directional LSTM (BI-LSTM), MUPE, GPRE, MUPE-GPRE, CNNE-GPRE, and CNNE-MUPE models. The models are used to predict 1-, 2-, and 3-day ahead power. The root mean square error values of CNNE-MUPE-GPRE, CNNE-MUPE, CNNE-GPRE, MUPE-GPRE, BI-LSTM, LSTM, CNNE, MUPE, GPRE were 578, 615, 832, 861, 914, 934, 1436, 1712, and 1954 KW at the 1-day prediction horizon. The RMSE of the CNNE-MUPE-GPRE was 595, 600, and 612 at the 1-day, 2-days, and 3-days prediction horizons. Extending the prediction horizon degrades accuracy. The uncertainty of the CNNE-MUPE-GPRE model was lower than that of the other models. The CNNE-MUPE-GPRE model is recommended for more accurate hydropower production predictions.

Keywords Hydropower · Deep learning models · Optimization algorithms · Power generation

✉ Mohammad Ehtearm
   mohammdehteram@semnan.ac.ir

Extended author information available on the last page of the article

1 Introduction

Sustainable development relies on energy resources (Ehteram et al. 2017). Energy is an important economic factor in industrial societies. The consumption of non-renewable energy leads to an increase in greenhouse gases that can change the global climate. Clean energy is energy that does not pollute the environment (Ehteram et al. 2018a). The hydropower industry is one of the most important sources of clean energy (Hou et al. 2021). Hydropower is an important renewable energy source due to its low cost. Climate, social, and economic factors can affect the hydropower system. The hydropower industry plays a key role in the development of electricity (Ehteram et al. 2018b). Hydropower plants play a key role in meeting the energy demand. An accurate prediction of power production is necessary for decision-makers in order to meet demand. Predicting hydropower production is necessary for managing energy resources. These predictions are needed for energy management. Furthermore, there are a number of uncertainties that may affect power production. Our research contributes to the development of energy prediction models. In addition, our study examines the impact of uncertainties on power generation.

The optimal operation of hydropower plants is a key topic in energy engineering. Accurate hydropower generation predictions can help the management of hydropower plants.

Hydropower production predictions can prevent energy shortages during droughts. Researchers have developed different models to predict hydropower production (Dehghani et al. 2019).

Machine learning models are widely used to predict hydropower production and energy demand. These models can find relationships between complex input and output data. The advantages of these models are fast calculation, easy implantation, and high precision. Guo et al. (2018) used support vector machine models (SVMs) to predict power production. The SVM model successfully predicted power generation. Dehghani et al. (2019) developed a neuro-fuzzy adaptive system (ANFIS) for hydropower production prediction. ANFIS parameters were adjusted using the grey wolf optimization (GWO) algorithm. It was reported that the ANFIS-GWO model performed better than the ANFIS model. Gao et al. (2019) proposed an ANN model for predicting one day-ahead power. They predicted power using a long short-term memory (LSTM) neural network. An LSTM is a recurrent neural network (RNN) that overcomes the vanishing gradient problem of conventional RNNs. An LSTM network uses memory cells to store. The meteorological data were used as the inputs to the models. The LSTM model predicted power successfully.

Rahman et al. (2021) developed LSTM, convolutional neural network (CNNE), and recurrent neural network models (RNN) to predict power energy. They stated that the different ANN models successfully predicted power energy. For electricity prediction, Zolfaghari and Golabi (2021) combined adaptive wavelet transforms (WT) with the LSTM model. They found that the WT-LSTM outperformed the LSTM model. They reported that the R^2 of wavelet- ANN and wavelet- LSTM models was 0.951 and 0.979, respectively. The root mean square error of the wavelet-ANN and wavelet LSTM was 8.65 and 6.73, respectively.

Barzola-Monteses et al. (2022) used artificial neural networks (ANNs) to predict hydropower production. The model parameters were set using a grid search algorithm. The developed model was a reliable tool for energy management. They considered two scenarios for their study. One step (one-month) and multi-step (12 months) were used to predict hydropower production. The average execution of the models was 1.48 and 1.37 min for the first and second scenarios. The RMSE of the MLP and LSTM models

was 195.1 and 177.68 for the first scenario. The RMSE of the MLP and LSTM models was 154.1 and 173.2 for the second scenario.

Hanoon et al. (2022) developed ANN, SVM, and auto regressive integrated moving average (ARIMA) models to predict power production of a reservoir in China. They reported that the ANN and SVM model successfully predicted power production. The correlation coefficient of the MLP model varied from 0.8761 to 0.8779. The correlation coefficient of the radial basis function neural network model varied from 0.8480 to 0.8710.

Studies have shown that ANN models can accurately predict power production, but they have some limitations. These models are unable to automatically extract the important features from time series data (Panahi et al. 2021). Setting model parameters requires robust algorithms. The preprocessing methods are required to determine the most appropriate input scenario (Panahi et al. 2021). Also, these models cannot automatically predict interval times. Using interval time prediction, we can quantify the uncertainty of a model. This paper aims to develop ANN models for predicting daily power production. This paper will use new techniques to fill research gaps. In recent years, many researchers have developed deep learning models to analyze complex data (Sharifzadeh et al. 2019). A deep learning model is becoming an increasingly popular alternative to traditional machine learning models for predicting hydropower generation. Deep learning models can capture complex relationships between the input parameters, and are more accurate than traditional models.

Studies have shown that hybrid ANN models outperform ANN models. Hybridizing ANNs and deep learning models can improve their performance (Sharifzadeh et al. 2019).

A Convolutional Neural Network model (CNNE) is a robust deep learning model because it can extract important information from time series data (Sinitsin et al. 2022). The CNNE model can be integrated with the ANN model to extract complex nonlinear patterns and important features (Sinitsin et al. 2022). An ANN model can be trained more efficiently if it receives relevant features as inputs. In this study, the CNNE model was coupled with an ANN to predict the daily power production of a hydropower plant. The CNNE-ANN model can provide better results because it combines the advantages of convolutional neural networks and multilayer perceptron layers. The CNNE-ANN model can be easily scaled to handle different sizes and complexities of input data. A convolutional neural network can efficiently extract features from input data, reducing the workload of the ANN layers. The architecture of the CNNE-ANN model allows it to identify features of the input data accurately. As the CNNE-ANN model uses pooling layers and regularization techniques, it is less prone to overfitting than ANN. For accurate hydro-power production predictions, the CNNE-ANN model can handle noisy data and outliers. As the CNNE-ANN model can be scaled up or down, it can be used for different applications and environments.

Since a CNNE-ANN model cannot capture uncertainty values, this model can be coupled with a Bayesian approach. This study also introduces an approach for determining the most appropriate input scenarios. The main contributions of the current study are as follows:

- The CNNE-ANN is introduced for predicting daily power production.
- We evaluate the accuracy of the new model against several ANN models, including ANN, CNNE, LSTM, and Bidirectional LSTM (BI-LSTM) models.
- The CNNE-ANN model is integrated with a Bayesian approach to quantify the uncertainty values.
- A new method is introduced to determine the appropriate input scenarios.

2 Materials and Methods

2.1 An Optimization Algorithm for Adjusting Model Parameters and Feature Selection

Since selection of the best model parameters is time-consuming and difficult, this study applied binary and continuous versions of an optimization algorithm to determine optimum values of model parameters, train different models, and select the best input combinations. For solving complex problems, the Salp swarm optimization algorithm (SSOA) is widely used. The SSOA is broadly applied in different fields such as feature selections (Faris et al. 2018), global optimization (Zhang et al. 2022), discounted knapsack problem (Dang and Truong 2022), training unreal network models (Panda and Majhi 2020), and training support vector machine models (Samantaray et al. 2022). The high speed and accuracy are the advantages of SSOA.

The salp chains are divided in two groups. The first group (leader) guides salps. The second group (remaining salps) follow leaders. In the search space, this swarm is looking for food sources. The leader location is updated based on the following equation

$$Salp_j^l = \begin{bmatrix} Food_j + \rho_1\left(\left(\left(upp_j - low_j\right)\rho_2 + low_j\right)\right) \leftarrow \rho_3 \geq 0.50 \\ Food_j - \rho_1\left(\left(\left(upp_j - low_j\right)\rho_2 + low_j\right)\right) \leftarrow \rho_3 < 0.50 \end{bmatrix} \tag{1}$$

where $Salp_j^l$: the leader location, $Food_j$: food source, ρ_1, ρ_2, and ρ_3: random parameters, upp_j: upper value of decision variable, and low_j: lower value of decision variable. Equation (2) is used to update the location of followers:

$$follower_j^i = \frac{1}{2}\left(follower_j^i + follower_j^{i-1}\right) \tag{2}$$

where $follower_j^i$: the ith follower at the jth dimension. A population of solutions is created before the optimization process begins. In the next step, an objective function is used to evaluate the generated solution. A food source is considered as the best solution. Equations (1) and (2) are used to update the location of followers and salps. When the stop creation is met, the process ends. The SSOA is a continuous optimization algorithm. A transfer function can convert the continuous SSA to a binary SSOA (BSSOA).

$$T\left(salp_j^i(t)\right) = \frac{1}{1 + exp^{-salp_j^i(t)}} \tag{3}$$

where $T\left(salp_j^i(t)\right)$: transfer function and $salp_j^i(t)$: i^{th} salp in j^{th} dimension. The final location of salps is computed based on the following equation:

$$salp_i^j = \begin{bmatrix} 1 \leftarrow if(\lambda) < T\left(salp_j^i\right) \\ 0 \leftarrow otherwise \end{bmatrix} \tag{4}$$

2.2 Structure of Convolutional Neural Network Model

A CNNE is a fusion of feature extraction and feature classification (Zou and Ergan 2023). CNNEs consist of convolutional layers and pooling layers followed by fully connected layers. A convolutional layer contains convolutional kernels (Zhao et al. 2022). Convolution layers consist of a finite number of filters (kernels) that are combined with input data to extract relevant features. Convolution kernels represent a kind of feature named a feature map (Tang et al. 2021). The pooling layer has two important tasks. The pooling layer can accelerate the network operation. The next convolution layer requires fewer calculations if feature maps are pooled. The pooling layer also enhances the performance of the CNNE (Liu et al. 2021). The CNNE will yield better results by selecting the most important features. The fully connected layer is similar to the traditional ANN models.

The output feature map is computed based on the following equation (Tang et al. 2021):

$$out_j^l = f\left(\sum_{i \in M_j} in_i^{l-1} * k_l^j + b_l^j \right) \tag{5}$$

where out_j^l: the new feature map, in_i^{l-1}: input, b_l^j: bias, f: activation function, $*$: convolution operation, j: number of feature maps, and l: number of layers. A pooling layer resizes feature maps based on the pooling operation:

$$\alpha_j = f\left(w_j^l down\left(M_j^{l-1} \right) + b_j^l \right) \tag{6}$$

where $down$: the pooling operation, M_j^{l-1}: the feature map of layer l-1, w_j^l: weight connection, and α_j: the new feature map after decreasing size. Figure 1 shows the structure of CNNE model.

2.3 Structure of BI-LSTLM

LSTM neural network is an artificial deep learning method based on recurrent neural network (RNN), which was presented by Hochreiter and Schmidhuber (1997). The LSTM overcomes the vanishing gradient problem of recurrent neural networks. The LSTM model has been widely used for time series data prediction and has also achieved excellent results (Azzouni and Pujolle 2017).
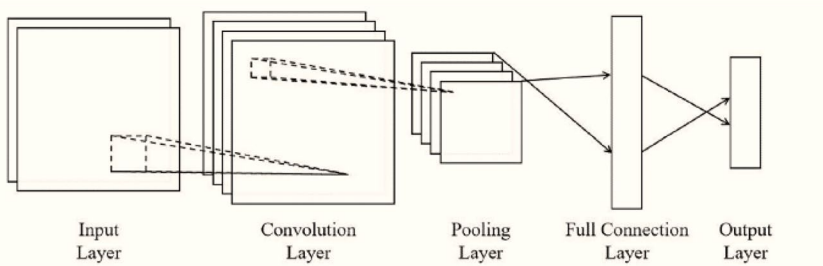

Fig. 1 The structure of CNN model

LSTM networks consist of memory blocks, memory cells, and gate units. A cell stat retains information. An LSTM model uses these gates to store and process the relevant information (Zha et al. 2022). The gates will learn what information can be retained and forgotten (Imrana et al. 2021). Input gates determine which information should be added to a cell state. The output gate provides outputs. A forget gate determines which information must be retained from a previous state (Jaseena and Kovoor 2021; Li et al. 2022; Jamei et al. 2022).

In this article, a BI-LSTM model based on a conventional LSTM neural network is developed to predict hydroelectric power based on multivariable inputs. The BI-LSTM considers past and future states to improve prediction accuracy. While ordinary LSTMs consider only historical observations, BI-LSTM considers future and previous observations. Reverse LSTMs use future information and forward LSTMs use past information. The BI-LSTM achieves better accuracy than LSTM because it utilizes both past and future information (He et al. 2022). Equations (7)–(12) mathematically describes the relationship between weighted inputs and outputs:

$$O_t = \sigma\left(\varphi_o z_{t-1} + V_o x_t + \kappa_o\right) \tag{7}$$

$$I_t = \sigma\left(\varphi_i z_{t-1} + V_i x_t + \kappa_b\right) \tag{8}$$

$$F_t = \sigma\left(\varphi_f z_{t-1} + V_f x_t + \kappa_f\right) \tag{9}$$

$$\widehat{S}_t = tanh\left(\varphi z_{t-1} + V x_t + b\right) \tag{10}$$

$$S_t = F_t \odot S_{t-1} + i_t \odot \widehat{S}_t \tag{11}$$

$$z_t = O_t \odot tanh\left(S_t\right) \tag{12}$$

where O_t, I_t, and F_t: output, input, and forget gates, x_t: input, z_t: the output state at time t, $\widehat{S}_t$: memory cell, $\widehat{S}_t$: new value of memory cell, φ_o, φ_i, and φ_f: weight matrices of hidden layer, and V_o, V_i, V_f: weights corresponding to input data, z_t: the output state a time t, and o, I, f: subscribes corresponding to output, input, and forget gate.

A BILSTM network consists of forward and backward LSTMs that can process data in both directions. In the forward LSTM layer, forward calculations are performed from time 1 to time t. The backward LSTM layer performs the backward calculation from time t to time 1. We obtain and save the output of the forward hidden states and backward hidden states. The BILSTM output is calculated by connecting the two hidden states.

Figure 2 shows a schematic diagram of a simple Bidirectional LSTM that has expanded over time (Zhou et al. 2016).

2.4 Structure of the ANN Models

One of the most commonly used models in hydrological modeling, feed-forward multi-layer perceptron (MUPEs) ANN, is used in this study. The MUPE models consist of an input layer, a number of hidden layers, and an output layer. Each layer of the MUPE model has weight connections that connect it to the next layer (Panahi et al. 2021). The
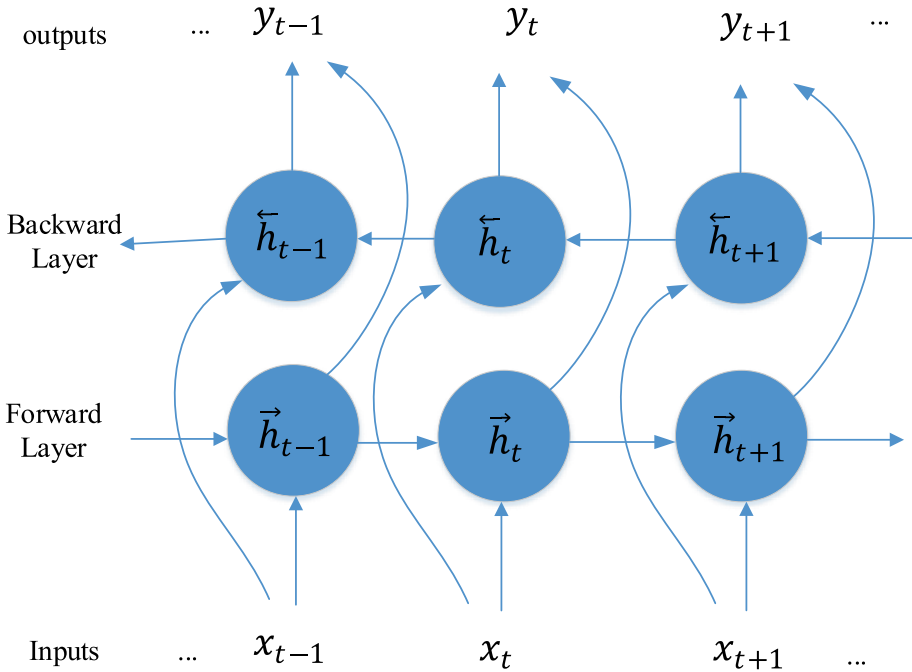
Fig. 2 Schematic diagram of a simple BI-LSTM

number of inputs determines the number of input neurons (Ehteram et al. 2023). Hidden layers receive weighted inputs from the input layer. Hidden and output layers can use linear or nonlinear activation functions. An activation function creates a relationship between weighted inputs and outputs. Figure 3 shows the structure of MUPE model. Unknown parameters of MUPE models include bias and weight. Parameter values are obtained through an optimization algorithm.

Fig. 3 Structure of the MLP model (Heidari et al. 2020)
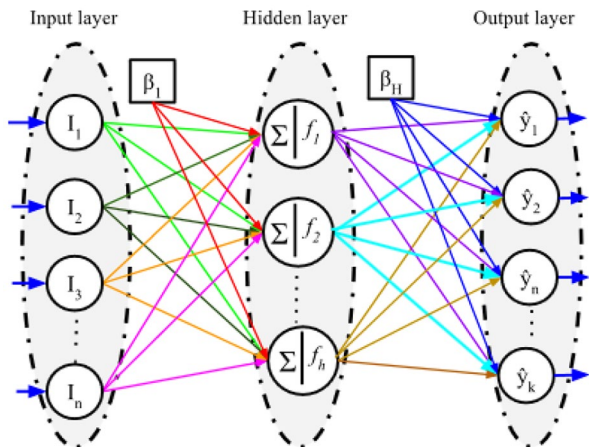
2.5 Evaluation of Models' Uncertainty Using Gaussian Progress Regression

Since ANN models cannot capture uncertainty values, this study proposes a Gaussian progress model (GPRE) for quantifying uncertainty. The GPRE model is widely used in the different fields such as groundwater quality monitoring (Shadrin et al. 2021), short term solar power forecasting (Wang et al. 2021), traffic load prediction (Wang et al. 2020a, b), wind speed forecasting (Huang et al. 2018), and short term-prediction of wind speed (Wang et al. 2020a, b). GPRE is a nonparametric model based on the Bayesian framework. The GPRE can prediction interval times. Thus, GPRE model can quantify the uncertainty values. The mathematical model of the GPRE model is defined based on the following equation:

$$Q = f(b_i) + \varepsilon \tag{13}$$

where, Q: observation, $f(\eta_i)$: an underlying function, b_i: input, and ε: noise.

$$\varepsilon \sim N(0, \sigma^2) \tag{14}$$

where σ^2: variance. The joint prior distribution of the observed data is computed based on the following equation (Sun et al. 2022):

$$Q \sim N(0, K(B,B) + \sigma_n^2 I_n) \tag{15}$$

$$\begin{bmatrix} Q \\ q \end{bmatrix} \sim N\left[0, \begin{bmatrix} K(B,B) + \sigma^2 & K(B,B_*) \\ K(B_*,B) & K(B_*,B_*) \end{bmatrix}\right] \tag{16}$$

where q: estimate value, $K(B,B)$: The covariance matrix of all input data, $K(B,B_*)$: the covariance matrix of test data point and all input data, B_*:test points, $K(B_*,B_*)$: the self-covariance of test points. The posterior distribution of estimated value is:

$$q|Q \sim N\left(\overleftarrow{q}, \sigma_q^2\right) \tag{17}$$

where $\overleftarrow{q}$: mean, and σ_q^2: variance. The mean and variance are computed based on the following equations:

$$\bar{q} = K(B_*,B)\left[K(B,B) + \sigma^2 I\right]^{-1} q \tag{18}$$

$$\sigma^2 = K(B_*,B_*) - K(B_*,B)\left[K(B,B) + \sigma^2 I\right]^{-1} K(B,B_*) \tag{19}$$

2.6 Hybrid Structure of CNNE-MUPE-GPRE

In this study, CNNE-ANN-GPRE is used to predict hydropower production. The model is created based on the following levels:

1. The 80% and 20% of data are used for training and testing levels because they provided the lowest error function values.

2. A binary vector is created based on the names of the input variables. A binary vector is defined as an initial population of the BSSOA.
3. The CNNE parameters are considered as the initial population of the CSSOA.
4. I = I + 1 (I: Iteration number)
5. At the training level, the CNNE model is run using the training data.
6. The quality of the solutions is evaluated using an error function (Nash Sutcliff efficiency (NSE)).
7. The operators of CSSOA are applied to change the values of model parameters.
8. If I > maximum number of iterations and NSE > 0.90, the model goes to the next level; otherwise, go to step 6.
9. The testing data are used to run the CNNE model.
10. The outputs of CNNE models are flattened.
11. The flattened outputs of CNNE models are inserted into the MUPE model
12. The model parameters of MUPE are defined as the initial population of the CSSOA.
13. An error function (NSE) is used to evaluate the quality of the solutions.
14. The operators of CSSOA are used to update the values of model parameters.
15. If the stop creation is met, the MUPE model goes to the step 16; otherwise, it goes to the step 13.
16. The testing data are used to run MUPE model.
17. The GPRE receives the outputs of the MUPE model.
18. The GPRE model is run at the training and testing levels.

This study benchmarks CNNE-MUPE-GPRE against the LSTM, BI-LSTM, CNNE, GPR, CNNE-ANN, MLP-GPRE, and CNNE-GPRE models to compare the performance of models. Figure 4a illustrates the flowchart of proposed algorithms for predicting hydropower using deep neural networks. Figure 4b shows the mechanism of the modelling process.

2.7 Evaluation Criteria

In order to evaluate the performance of the developed models, several evaluation criteria are used in the current study.

Root mean square error

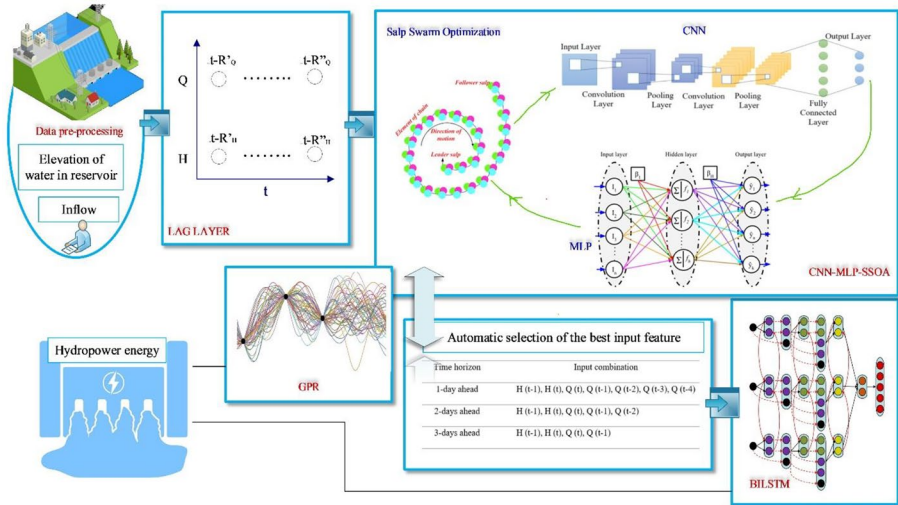$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(P_{esi} - P_{obi}\right)^2}$$

(20)

1. Index of agreement (IA)

$$IA = 1 - \frac{\sum_{i1}^{N}\left(P_{obi} - P_{esi}\right)^2}{\sum_{i=1}^{N}\left(\left|P_{obi} - \bar{P}\right| + \left|P_{esi} - \bar{P}\right|\right)^2}$$

(21)

2. Nash–Sutcliffe efficiency

$$NSE = 1 - \frac{\sum_{i=1}^{N}\left(P_{obi} - P_{esi}\right)^2}{\sum_{i=1}^{N}\left(P_{obi} - \bar{P}_{obi}\right)^2}$$

(22)

3. Kling–Gupta Efficiency (KGE)

a
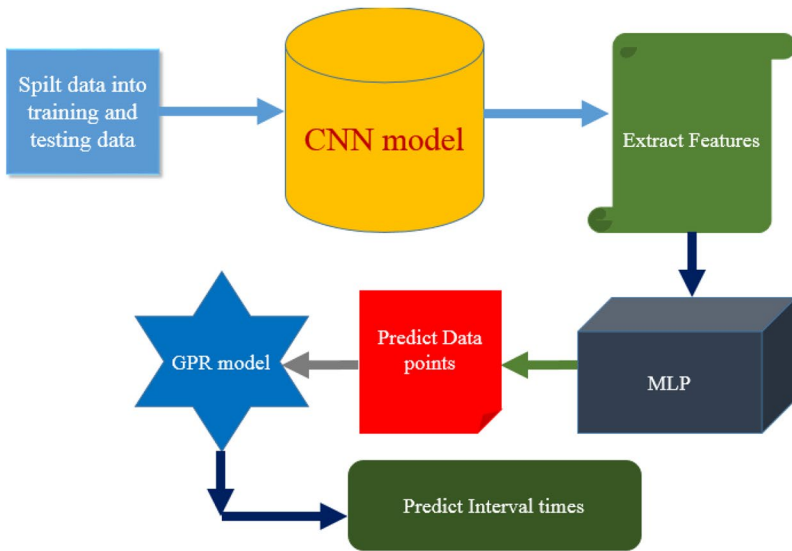
b

Fig. 4 a Proposed algorithm for predicting hydropower using deep neural network, **b** mechanism of modelling process

$$KGE = 1 - \sqrt{(1-r) + \left(\frac{(P_{obi})}{(P_{esi})} - 1\right)^2 + \left(\frac{CV_p}{CV_{em}}\right)^2} \qquad (23)$$

4. Prediction interval coverage probability

$$PICP = \frac{1}{N} \sum_{i=1}^{N} \mu_i \tag{24}$$

$$\mu_i = \begin{bmatrix} 1, \; if \left(P_{obi}\right) \in \left[L_i, U_i\right] \\ 0, f \left(P_{obi}\right) \notin \left[L_i, U_i\right] \end{bmatrix} \tag{25}$$

5. Prediction Interval Normalized Average Width (PINW)

$$PINW = \frac{1}{NR} \sum_{i=1}^{N} \left(U_i - L_i\right) \tag{26}$$

where P_{obi}: observed data, P_{esi}: predicted power, R: range of data points, P: average observed data, r: correlation coefficient, CV_p and CV_{em}: a coefficient of variation for predicted data, and observed data. High and low values of PICP and PINW show the best model. PICP and PINW are used to quantify uncertainty values of models.

2.8 Case Study

The Karun basin (Fig. 5a) is one of the largest basins in Iran, located in the southwest of Iran. In this basin, average annual precipitation ranges from 153 mm in the southern plains to more than 2000 mm in mountainous regions. The variations of daily temperature over the basin are from a minimum of -30.6 °C at Koohrang station to a maximum of 52.2 °C at Ahvaz station (Fallah et al. 2020). The Karun-III dam is one of the most important dams located on the Karun River. The annual mean flow of the dam is 300 m³/s. The storage volume of the Karun-III dam is 2,970,000,000 m³. The dam is designed to generate hydropower. Thus, the dam reservoir is used for power generation. In this study, the models are used to predict the 1-, 2-, and 3-day ahead power. The elevation of water in reservoir and inflow discharge are used to predict power generation. The lag times were from (t-1), …, (t-30). Table 1 shows details of input and output data. Figure 5b shows time series data. The data were collected from 2005 to 2017.

3 Results and Discussion

3.1 Feature Selection

In the modeling process, feature selection plays an important role. It is time consuming and difficult to manually determine the most appropriate input combination among $2^{60}-1$ input combinations. Correlation values and principal component analysis can identify significant inputs and lag times, but they cannot automatically determine the optimal inputs. A binary version of SSOA was used to determine the most suitable scenario for power generation prediction. Inputs were initialized as the initial population of salps. The salp location displays the names of input variables. Binary vectors contain 1 and 0 values that represent unselected and selected features, respectively. At each iteration, the SSOA updates input combinations using its operators. Table 2 lists the most appropriate scenario. The input scenarios are used to predict power generation. when a modeler encounters many data points and needs to estimate targets, the binary SSOA will be a useful tool.
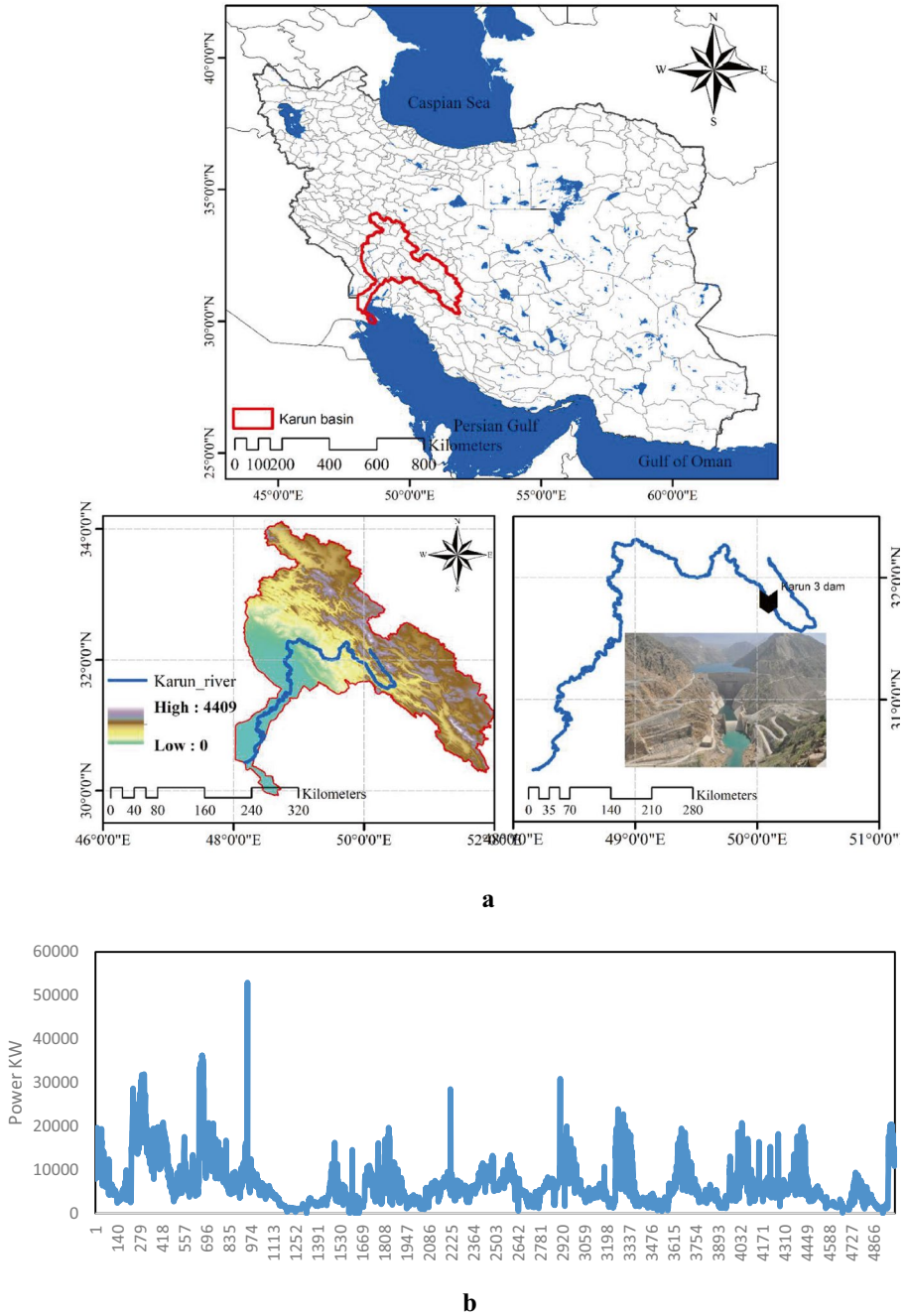
a

b

Fig. 5 **a** Location of the Karun III dam as case study, **b** Time series data

Table 1 Related information to the input data

Parameter	Average	Maximum	Minimum	Standard deviation
Inflow (m³/s) (Q)	242.81	5695.66	0	220.00
Elevation of water in reservoir (m) (H)	514.79	533.1	486.91	12.28
Power (KW)	12595	52942.2	612.2	5181.43

3.2 Determination of Random Parameters and Model Parameters

For determining model parameters and features, CSSOA is used. A CSSOA includes random parameters such as population size (PSI) and maximum number of iterations (MNITER). The root mean square error (RMSE) is used to determine the optimal values of random parameters. Table 3 shows the optimal values of PSI and MNITER. PSI values varied from 100 to 600. The PSI = 200 gave the lowest RMSE values. MNITER values varied from 50 to 300. The MNITER = 100 gave the lowest value of RMSE. Tables 4 shows the optimal values of CNN, MLP, and LSTM parameters.

3.3 Investigation of the Accuracy of Models

Figure 6 shows IA, NSE, and RMSE values of models for 1-day ahead power prediction. The training IA of the CNNE-MUPE-GPRE, CNNE-MUPE, CNNE-GPRE, MUPE-GPRE, BI-LSTM, LSTM, CNNE, MUPE, and GPRE models was 0.97, 0.94, 0.90, 0.85, 0.84, 0.81, 0.80, 0.75, and 0.70, respectively. The IA values of the CNNE-MUPE-GPRE, CNNE-MUPE, CNNE-GPRE, MUPE- GPRE, BI-LSTM, LSTM, CNNE, MUPE, and GPRE models were 0.95, 0.92, 0.86, 0.84, 0.82, 0.80, 0.76, 0.72, and 0.66 at the testing level. The NSE values of the CNNE-MUPE-GPRE, CNNE-MUPE, CNNE-GPRE, MUPE-GPRE, BI-LSTM, LSTM, CNNE, MUPE, and GPRE models were 0.94, 0.92, 0.87, 0.80, 0.75, 0.70, 0.68, 0.66, and 0.62 at the training level. The NSE values of the CNNE-MUPE-GPRE, CNNE-MUPE-, CNNE-GPRE, MUPE- GPRE, BI-LSTM, LSTM, CNNE, MUPE, and GPRE models were 0.93, 0.89, 0.85, 0.78, 0.74, 0.69, 0.67, 0.65, and 0.60 at the testing level. The RMSE values of CNNE-MUPE-GPRE, CNNE-MUPE, CNNE-GPRE, MUPE-GPRE, BI-LSTM, LSTM, CNNE, MUPE, and GPRE were obtained equal to 545, 612, 824, 855, 912, 914, 1400, 1700, and 1900 KW, respectively at training phase, and equal to 595, 723, 836, 897, 916, 1200, 1500, 1800, and 2000 KW, respectively at testing phase.

Figure 7 shows the accuracy of models for 2-day ahead power production prediction. The RMSE values of the CNNE-MUPE-GPRE, CNN-MUPE, CNN-GPRE, MUPE-GPRE, BI-LSTM, LSTM, CNNE, MUPE, and GPRE models were 578, 615, 832, 861,

Table 2 Optimal input scenarios for 1-day, 2-days, and 3-days ahead

Time horizon	Input combination
1-day ahead	H (t-1), H (t), Q (t), Q (t-1), Q (t-2), Q (t-3), Q (t-4)
2-days ahead	H (t-1), H (t), Q (t), Q (t-1), Q (t-2)
3-days ahead	H (t-1), H (t), Q (t), Q (t-1)

Table 3 Choice of random parameters of SSOA

PSI	Objective function value (RMSE:KW)	MNITER	Objective function value
100	678	50	612
200	545	100	543
300	612	150	624
400	723	200	721
500	824	250	844
600	941	300	912

914, 934, 1436, 1712, and 1954 KW, respectively at training level. The CNNE-MUPE-GPRE decreased RMSE values of the CNNE-MUPE, CNNE-GPRE, MUPE- GPRE, BI-LSTM, LSTM, CNNE, MUPE, and GPRE models by 18%, 28%, 50%, 36%, 52%, 61%, 66%, and 70%, respectively at the testing level.

At training phase, the NSE values of CNNE-MUPE-GPRE, CNNE-MUPE-, CNNE-GPRE, MUPE- GPRE, BI-LSTM, LSTM, CNNE, MUPE, and GPRE were 0.93, 0.90, 0.85, 0.78, 0.72, 0.69, 0.65, 0.60, and 0.58, respectively. The training IA values of CNN-MLP-GPR, CNN-MLP-, CNNE-GPRE, MUPE- GPRE, BI-LSTM, LSTM, CNNE, MUPE, and GPRE were 0.96, 0.91, 0.86, 0.81, 0.76, 0.73, 0.68, 0.65, and 0.63, respectively. The testing IA values of models were 0.94, 0.90, 0.82, 0.78, 0.74, 0.70, 0.67, 0.64, and 0.62, respectively.

Figure 8 shows the accuracy of models for 3-day ahead power production prediction. The CNNE-MUPE-GPRE, CNNE-MUPE, CNNE-GPRE, MUPE- GPRE, BI-LSTM, LSTM, CNNE, MUPE, and GPRE models had RMSE values of 582, 621, 839, 871, 924, 945, 1456, 1724, and 1831 KW at training level and 612, 731, 855, 923, 935, 1267, 1672, 1815, and 2045 KW at testing level, respectively. The NSE values of those models were 0.92, 0.89, 0.84, 0.77, 0.71, 0.65, 0.62, 0.58, and 0.54 at the training level, and 0.90, 0.86,

Table 4 Optimal values of CNNE, MUPE, and LSTM parameters

|	Hyper parameters	Range	Optimized value by CSSOA
CNN	Batch size	[24, 55]	52
	Max epoch	[30, 50]	32
	Kernel size of the first Convolution layer	[3,5]	3
	Kernel size of the second Convolution layer	[3,5]	3
	Kernel number of the first Convolution layer	[5, 15]	8
	Kernel number of the second Convolution layer	[5, 15]	8
MLP	Number of hidden layers		1
	Number of hidden neurons		6
	Learning rate		0.02
LSTM	Number of hidden layers	[2,8]	5
	Number of hidden neurons	[1, 30]	24
	Epoch	[5, 100]	45
	Batch size	[30, 60]	50
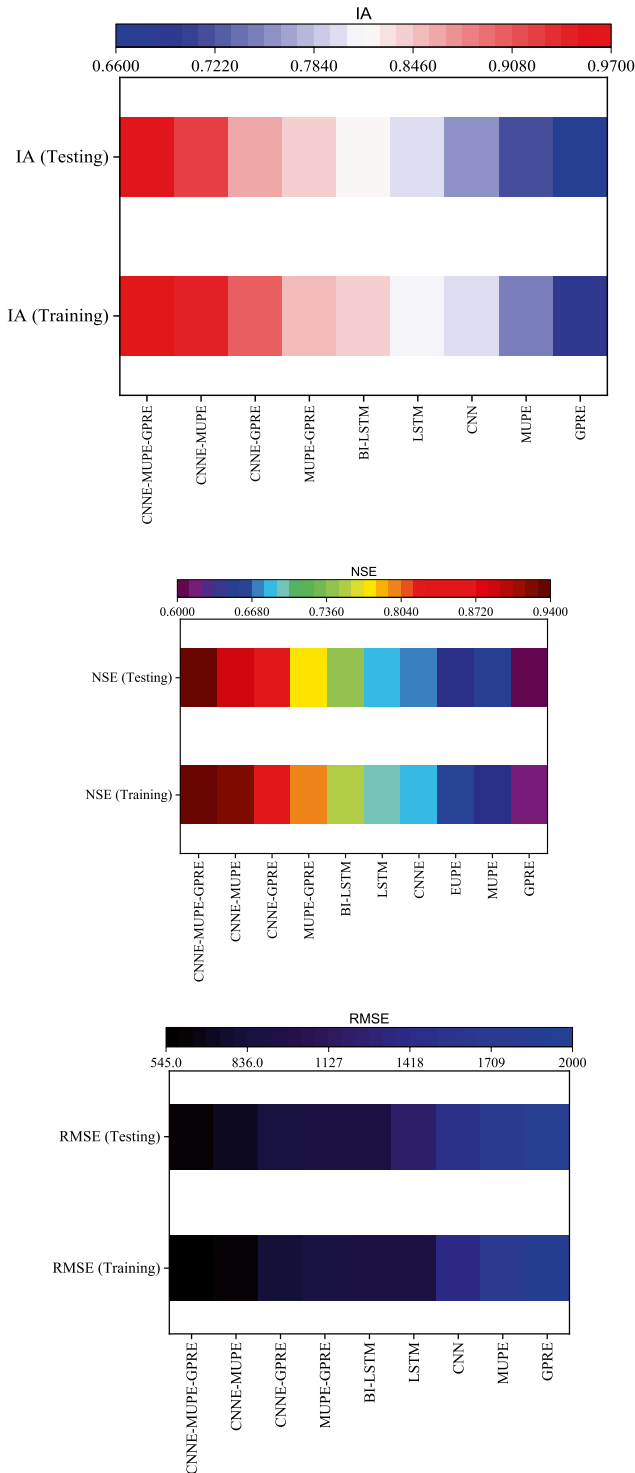
Fig. 6 Values of AI, NSE, and RMSE criteria for 1-day ahead power production

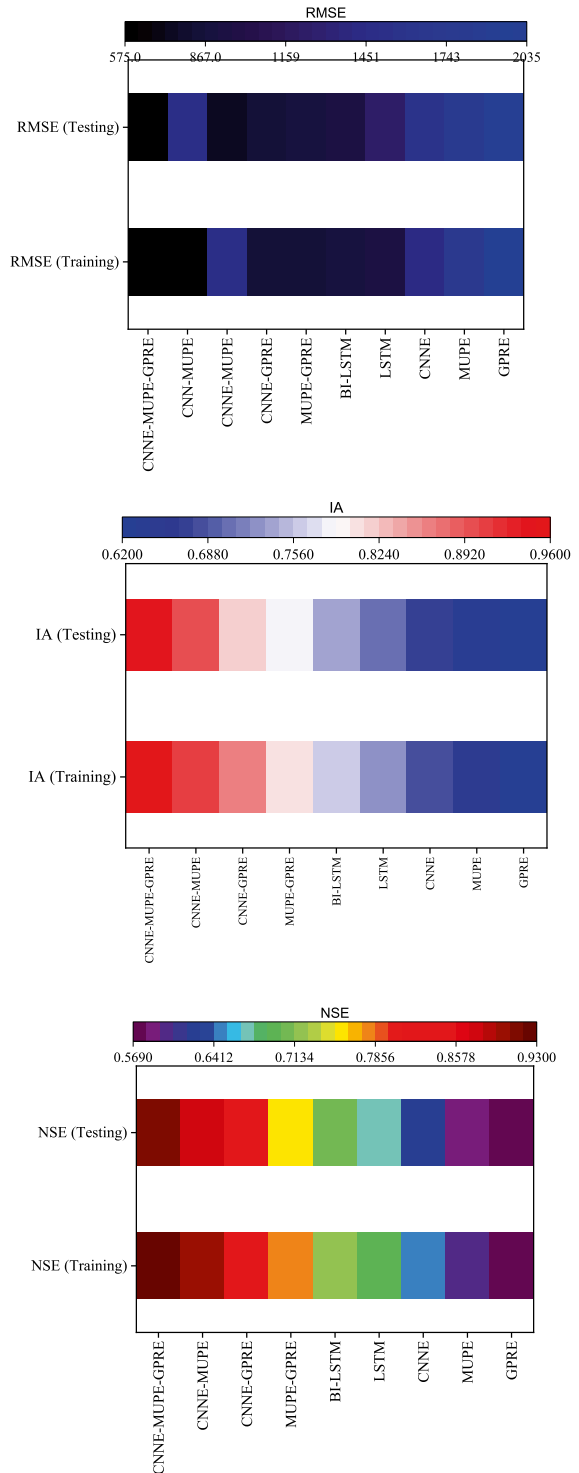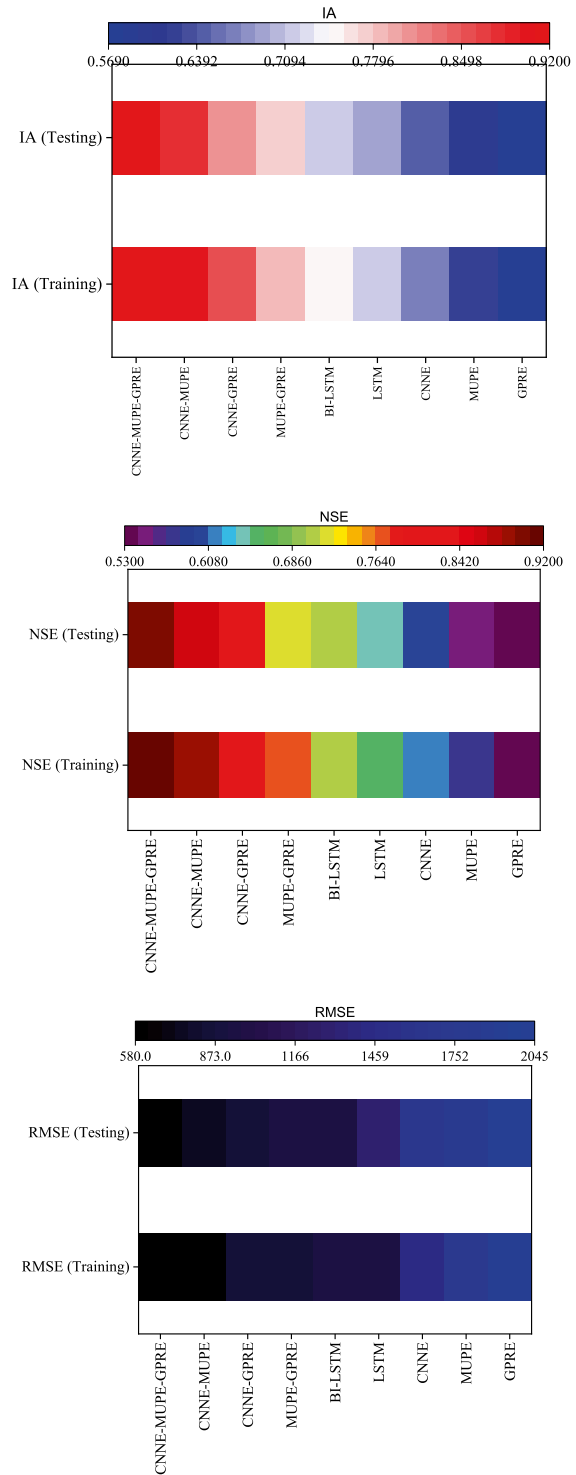**Fig. 7** Values of AI, NSE, and RMSE criteria for 2-day ahead power production

Fig. 8 Values of AI, NSE, and RMSE criteria for 3-days ahead power production

0.81, 0.72, 0.70, 0.64, 0.60, 0.55, and 0.63 at the testing level, respectively. The IA values for CNNE-MUPE-GPRE, CNNE-MUPE, CNNE-GPRE, MUPE- GPRE, BI-LSTM, LSTM, CNN, MUPE, and GPRE models were 0.92, 0.90, 0.85, 0.78, 0.75, 0.71, 0.67, 0.62, and 0.58 at the training level and 0.91, 0.88, 0.81, 0.77, 0.72, 0.69, 0.65, 0.60, and 0.57 at the testing level.

The main findings of this section are:

1. A CNNE-MUPE-GPRE model has the best precision since it combines the advantages of CNNE, MUPE, and GPRE models. The CNNE model extracted important features. An MUPE received the extracted features from a CNNE model. Based on the outputs of the MUPE models, the GPR predicted the outputs. Models can deeply learn complex and nonlinear patterns through this process.
2. The CNNE-MUPE-GPRE model combines the feature extraction capabilities of CNNs with the flexibility of GPRE. In addition, the CNNE-MUPE-GPRE model can handle high-dimensional data that the MUPE cannot handle. The CNNE-MUPE-GPRE model also accounts for uncertainty in its predictions.
3. The CNNE-MUPE-GPRE decreased RMSE of the CNNE-MUPE, CNNE-GPRE, MUPE-GPRE by 10%, 33%, and 36% at the 1-day prediction horizon. The CNNE-MUPE and CNN-GPRE outperformed the MUPE-GPR. The CNN-MUPE and CNNE-GPRE models performed better than the MUPE-GPRE because they took advantage of the CNNE model in the modeling process. Thus, the CNNE model had a key role in the modeling process.
4. The BI-LSTM model outperformed the LSTM model because it used the past and feature data. The BI-LSTM decreased RMSEs of LSTM by 24%, 26%, and 27% at the periods of 1-day, 2-day, and 3-day.
5. The MUPE-CNNE-GPRE model had RMSE of 595, 600 and 612, NSE of 0.93, 0.91, 0.90, and AI of 0.95, 0.94, and 0.91 at the periods of 1-day, 2-days, and 3-days. The accuracy deteriorated with an extension of the prediction horizon.
6. The hybrid models outperformed the CNEE, MUPE, and GPRE models. The GPRE model had the worst performance among other models.
7. A bidirectional Long Short Term Memory (BILSTM) neural network has several advantages over a traditional Long Short Term Memory (LSTM) neural network. The network can incorporate past and future information into its decisions. Thus, bidirectional Long Short Term Memory can maintain long-term memories and deal with complex temporal relationships. In contrast to traditional neural networks, Bidirectional LSTM networks avoid vanishing or exploding gradients.

Figure 9 shows box plots of different models for 1-day, 2-day, and 3-day ahead. Based on the Fig. 9a, the median values were 12000, 12700, 12900, 13000, 13400, 13600, 14000, 14100, 14200, and 14500, while the minimum values were 612, 612, 615, 623, 626, 632, 656, 676, 679, and 681 for observed data, CNN-,MUPE-GPRE, CNNE-MUPE, CNNE-GPRE, MUPE-GPRE, BI-LSTM, LSTM, CNNE, MUPE, and GPRE models, respectively. Figure 9b shows the box plots of models for two-day-ahead prediction.
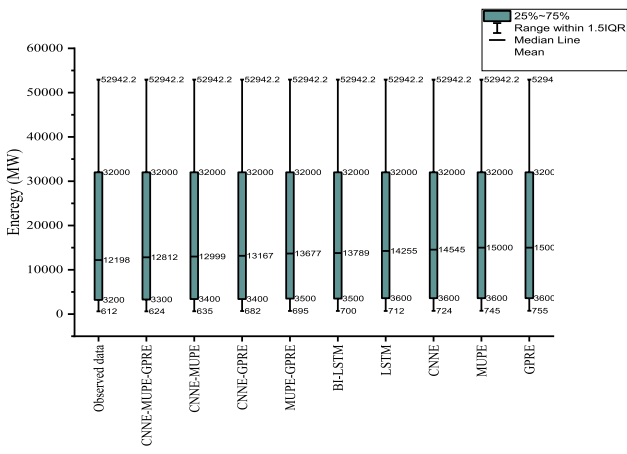
The median values of observed data, CNNE-MUPE-GPRE, CNNE-MUPE, CNNE-GPRE, MUPE-GPRE, BI-LSTM, LSTM, CNNE, MUPE, and GPRE models were 12100, 12719, 12954, 13050, 13467, 13589, 14050, 14198, 14700, and 14655 KW, respectively. The minimum values of observed data and those models were 612, 615, 615, 623, 645, 655, 667, 682, 690, and 699, respectively, Fig. 9c shows the box plots of models for three-day-ahead prediction.
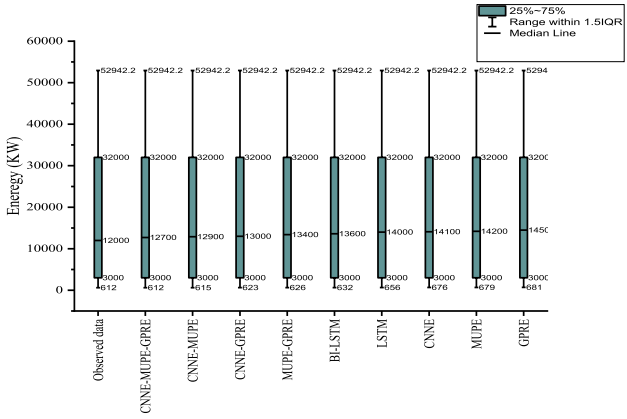
(a)
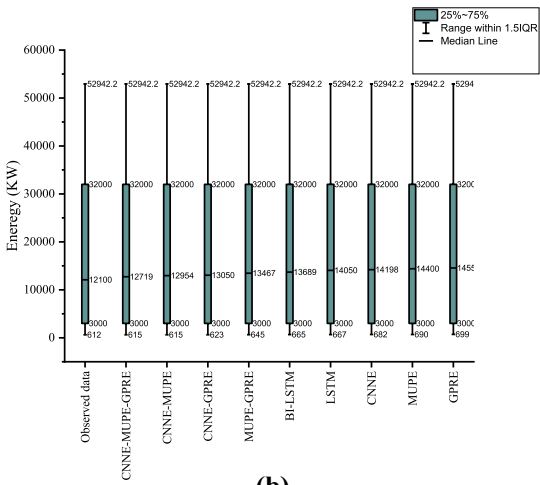
(b)

(c)

Fig. 9 Box plots of models for 1-day, 2-days, 3-days ahead predictions

The median values for 3-day ahead of observed data, CNNE-MUPE-GPRE, CNNE-MUPE, CNNE-GPRE, MUPE-GPRE, BI-LSTM, LSTM, CNNE, MUPE, and GPRE models were 12198, 12812, 12999, 13167, 13677, 13789, 14255, 14545, 15000, and 15050, respectively. Also, the minimum values of observed data and those models were 612, 624, 635, 682, 695, 700, 712, 724, 745, and 755, respectively.

Figure 10 shows KGE values of different models. The KGE of CNNE-MUPE-GPRE, CNNE-MUPE, CNNE-GPRE, MUPE, GPRE, BI-LSTM, LSTM, CNNE, MUPE, and GPRE was 0.97, 0.93, 0.91, 0.90, 0.86, 0.83, 0.78, 0.77, and 0.76, at the 1-day prediction horizon. The KGE of the CNN-MUPE-GPRE was 0.97, 0.96, and 0.94 at the 1-day, 2-day, and 3-day prediction horizons. The KGE of CNNE-MUPE-GPRE, CNNE-MUPE, CNNE-GPRE, MUPE, GPRE, BI-LSTM, LSTM, CNNE, MUPE, and GPRE models was 0.96, 0.92, 0.90, 0.88, 0.87, 0.85, 0.82, 0.76, 0.74, and 0.72 at the 2-day prediction horizon.

3.4 Investigation Ion of the Uncertainty of Models

For quantifying uncertainty values, the CNNE-MUPE model was coupled with the GPRE model. The CNNE-MUPE-GPRE model, MUPE-GPRE model, CNNE-GPRE model, and GPRE model can capture uncertainty values. Figure 11 shows 95% confidence for 1-day ahead prediction. The testing data points were used to draw these figures. The results show that more than 95% of data are bracketed by uncertainty bounds. The PICP values of the CNNE-MUPE-GPRE, CNNE-GPRE, MUPE-GPRE, and GPRE were calculated as 0.99, 0.98, 0.96, and 0.95, respectively. A high value of PICP demonstrated low uncertainty and instability and showed high accuracy of the model (Seifi et al. 2022). The PINW values of the CNNE-MUPE-GPRE, CNNE-GPRE, MUPE-GPRE, and GPRE models were 0.05, 0.10, 0.12, and 0.14, respectively. The results of PINW showed that the variability of GPR predictions was higher than hybrid models of CNNE-MUPE-GPRE, CNNE-GPRE, and MUPE-GPRE. The GPRE model had the highest uncertainty among other models. The CNN-MUPE-GPRE had the lowest uncertainty because it took advantage of three models.

Figure 12 shows PICP and PINW values for 2-days and 3-day ahead predictions. The PICPs of the CNNE-MUPE-GPRE, CNNE-GPRE, MUPE-GPRE, and GPRE model were 0.95, 0.94, 0.92, and 0.92 at the 2-day prediction horizon. The values of PICP of the CNNE-MUPE-GPRE, CNNE-GPRE, MUPE-GPRE, and GPRE models were 0.94, 0.90, 0.87, and 0.87 at the 3-day prediction horizon, respectively. As prediction horizons increased, uncertainty of predictions increased. As the 2-day and 3-day prediction horizons may include irregular and nonlinear patterns, the uncertainty may increase.
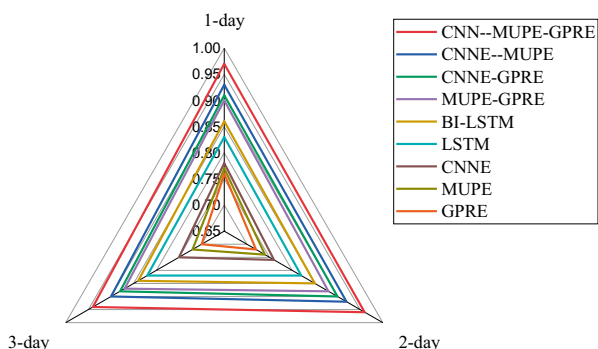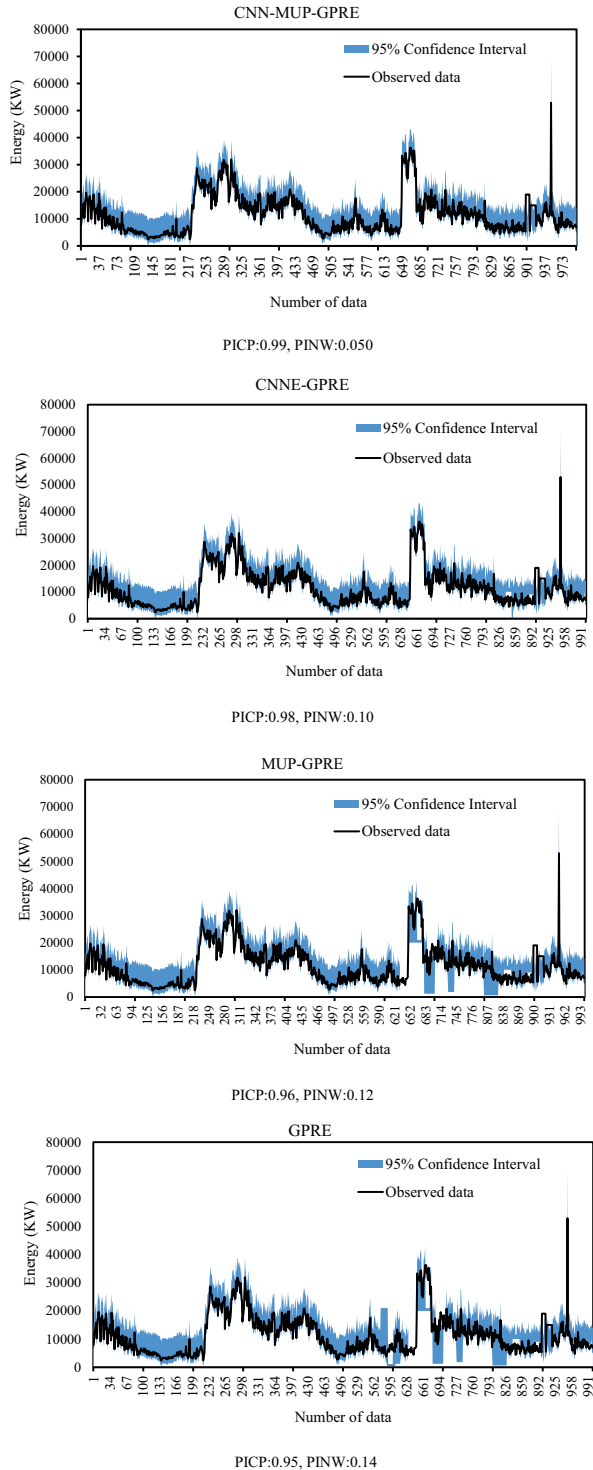

Fig. 10 KGE values for different models

Fig. 11 The 95% confidence interavl for 1-day ahead predction

CNN-MUP-GPRE

PICP:0.99, PINW:0.050

CNNE-GPRE

PICP:0.98, PINW:0.10

MUP-GPRE

PICP:0.96, PINW:0.12

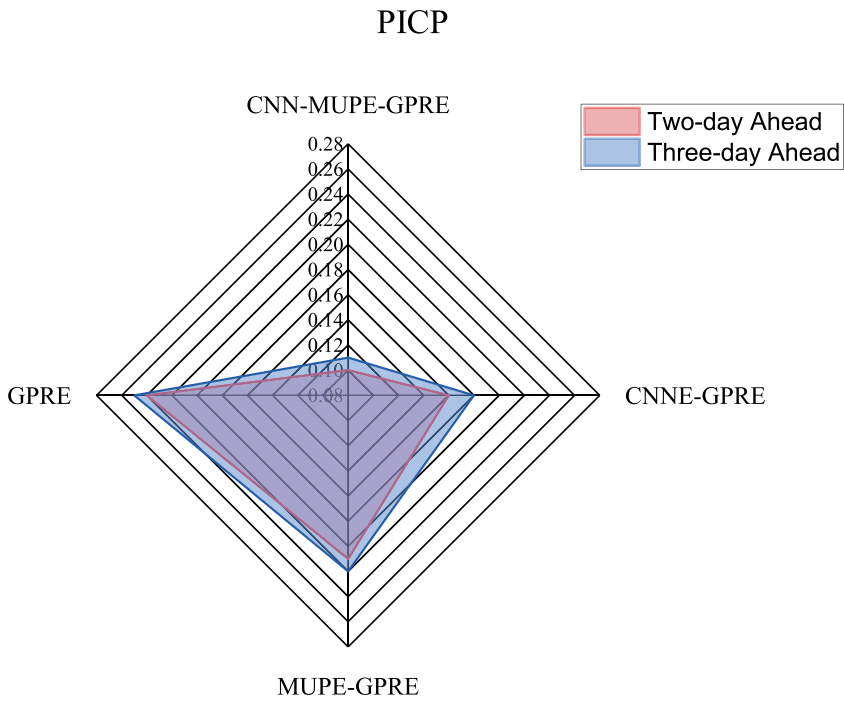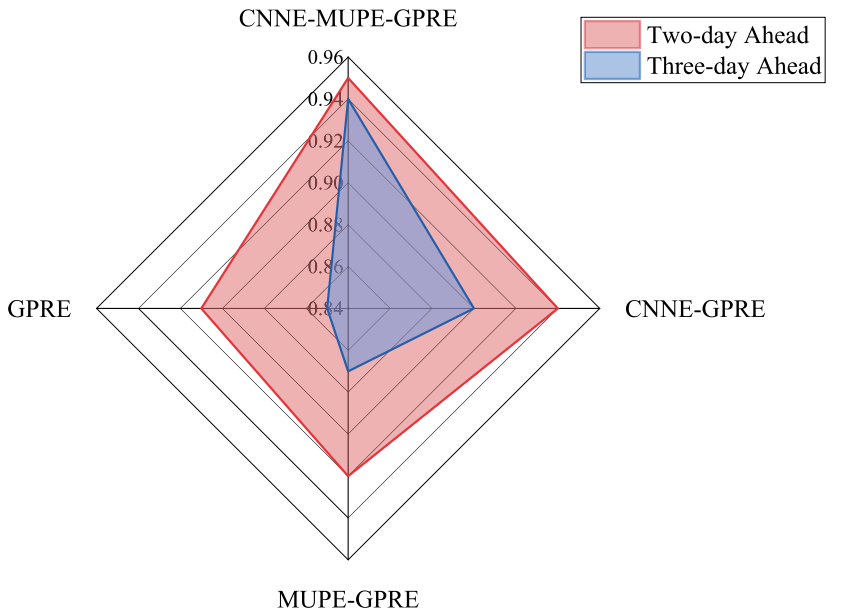GPRE

PICP:0.95, PINW:0.14

PICP

PINW values

Fig. 12 The PICP and PINW values of different models in uncertainty analysis

CNNE-MUPE-GPRE, CNNE-GPRE, MUPE-GPRE, and GPRE models had PINW values of 0.10, 0.16, 0.21, and 0.24 at the 2-day prediction horizon, respectively. The values of PINW of the CNNE-MUPE-GPRE, CNNE-GPRE, MUPE-GPRE, and GPRE models were 0.11, 0.18, 0.22, and 0.25 at the 3-day prediction horizon, respectively.

3.5 Main Findings of Paper

This paper used the CNNE-MUPE-GPRE for predicting hydrpower production. The main findings of paper are as follows:

1. A continuous and binary version of SSA was developed for adjusting model parameters and selecting inputs. Previous studies randomly selected the best input scenarios. The correlation method and principal component analysis were also suggested as methods for choosing the optimal input scenario. The binary SSA is superior to other methods because it automatically selects the best input scenario. Therefore, the current study fills a research gap between input selection and predictive models.
2. We used CNNE-MUPE-GPRE to predict different prediction horizons. The model successfully predicted 1-day, 2-day and 3-day ahead. The model can be used for both short-term and long-term predictions.
3. The CNNE-MUPE-GPRE can be used for long-term predictions because CNN helps extract features.
4. Since CNNE-MUPE-GPRE had the lowest uncertainty, it was a reliable tool. Due to the combination of three models, CNNE-MUPE-GPRE outperformed CNNE, MUPE, and GPRE.
5. Next studies can combine CNNE and BI-LSTM models since BI-LSTM uses past and feature data to predict outcomes. BI-LSTM uses backward and forward processes for deep learning.
6. For predicting spatial–temporal data, CNNE can be coupled with the LSTM model since both models can extract spatiotemporal patterns.
7. Furthermore, this study contributes to the development of energy engineering systems. This study develops energy monitoring systems that can be utilized in buildings, factories, and hydropower plants. Energy managers are looking for new technologies to monitor energy resources. Our models can be used to create sensors for energy monitoring. Also, these sensors can predict energy consumption. Our models will be useful for modifying patterns of energy consumption.
8. Different sciences can also use our models as early warning systems. These models can be used to monitor droughts and floods. In advanced engineering informatics, handling large data sets and quantifying uncertainty values are also important. The paper introduces a novel deep learning model and a GPRE model to achieve these aims. The CNNNE-MUPE-GPRE could quantify uncertainty values. Also, it had the lowest uncertainty among other models. Thus, this model can handle large data sets and quantify the uncertainty in advanced engineering informatics systems.
9. This study found that series hybridization improved the performance of standalone models. The hybrid structure allows the models to share their information with each other.
10. Hydropower plants can be successfully managed if their power generation capacity can be accurately predicted. When a drought reduces inflow to a hydropower plant, a power company can produce more electricity from alternative sources. In addition, our

Table 5 The accuracy of the CNN-MUPE-GPR based on new input combination

	RMSE	RMSE	NSE	NSE	IA	IA
best input combinations (H (t-1), H (t), Q (t), Q (t-1), Q (t-2), Q (t-3), Q (t-4) for one-month ahead)						
best input combinations (H (t-1), H (t), Q (t), Q (t-1), Q (t-2) for two-month ahead)						
best input combinations (H (t-1), H (t), Q (t), Q (t-1) for three-month ahead)						
One-Month ahead	545	595	0.94	0.93	0.97	0.95
Two-month ahead	578	600	0.93	0.91	0.96	0.94
Three-month ahead	582	612	0.92	0.90	0.92	0.91
Precipitation + evaporation + best input combinations						
One-Month ahead	542	589	0.95	0.94	0.98	0.96
Two-month ahead	576	592	0.94	0.92	0.97	0.95
Three-month ahead	580	602	0.93	0.92	0.93	0.93

results can contribute to grid stability. Energy companies can better manage energy demand and supply on the grid when they can predict hydropower power production.

11. There are many factors that influence hydropower production, including location, weather patterns, and hydropower plant design. Precipitation can directly affect the water level of the reservoir. Thus, it can affect hydropower production. When evaporation increases, available water decreases, which results in a decrease in hydropower production. Evaporation can significantly affect hydropower production in areas with high evaporation. Hydropower production is particularly vulnerable to drought conditions and other environmental factors that reduce water availability. The relationship between relative humidity and hydropower production is complex and depends on multiple factors such as wind speed, air temperature, and location of the hydropower plant. Understanding the effect of meteorological variables on hydroelectric power plants is necessary to improve their efficiency.

12. Based on new input combinations, Table 5 compares the accuracy of models. Precipitation and evaporation have been added to the previous input combinations. The RMSE of the CNNE-MUPE-GPRE model was 589, 592, and 602 for one, two, and three-month ahead. The results revealed that the new input combination does not significantly change the accuracy of the CNN-MUPE-GPRE model.

4 Conclusion

Prediction of hydropower production is essential for planning and managing water resources. Predictions of electricity production are used to make strategic decisions. This study is developed a new hybrid deep-learning model for daily hydropower production prediction. A CNNE-MUPE-GPRE is introduced to predict hydropower production. A binary SSA was used to select the best input combination. A CNNE-MUPE-GPRE model predicts data and selects outputs simultaneously. The results indicated that the CNNE-MUPE-GPRE was the best model among other models. The training IA of CNNE-MUPE-GPRE, CNNE-MUPE, CNNE-GPRE, MUPE, GPRE, BI-LSTM, LSTM, CNNE, MUPE, GPRE

was 0.97, 0.94, 0.90, 0.85, 0.84, 0.81, 0.80, 0.75, and 0.70 at the 1-day prediction horizon. The RMSE of the CNN-MUPE-GPRE was 595, 600 and 612 at the 1-day, 2-days, and 3-days prediction horizons. The accuracy deteriorated with an extension of the prediction horizon. The CNNE-MUPE-GPRE had the lowest uncertainty among other models. The results revealed that the new hybrid models outperformed the MLP, LSTM, GPRE, BI-LSTM, and CNNE models. We combined a CNN-MLP model with a GPR model to reflect uncertainty. The values of PINW of the CNN-MUPE-GPRE, CNNE-GPRE, MUPE-GPRE, and GPR were 0.10, 0.16, 0.21, and 0.24 at the 2-days prediction horizon. The GPR model had the highest uncertainty among other models. Future studies can use the metrological parameters to predict hydropower production. This study suggests combining CNNE models with LSTM and BI-LSTM to improve accuracy.

Author Contribution Formal analysis: Hossein Ghayoumi Zadeh, Moahammad Ehteram, writing, review, and editing: Ali Fayazi, Mohammad Ehteram, Akram Seifi, Majid Dehghani.

Data Availability The data sets are available based on reasonable request.

Declarations

Ethics Approval Informed constant.

Consent to Participate The author consents to participate.

Consent to Publish The author consents to publish.

Competing Interests The author states that she has not competing interest.

References

Azzouni A, Pujolle G (2017) A long short-term memory recurrent neural network framework for network traffic matrix prediction. arXiv preprint arXiv:1705.05690

Barzola-Monteses J, Gómez-Romero J, Espinoza-Andaluz M, Fajardo W (2022) Hydropower production prediction using artificial neural networks: an Ecuadorian application case. Neural Comput Appl. https://doi.org/10.1007/s00521-021-06746-5

Dang BT, Truong TK (2022) Binary salp swarm algorithm for discounted 0–1 knapsack problem. PLoS ONE 17(4):e0266537

Dehghani M, Riahi-Madvar H, Hooshyaripor F, Mosavi A, Shamshirband S, Zavadskas EK, Chau KW (2019) Prediction of hydropower generation using Grey wolf optimization adaptive neuro-fuzzy inference system. Energies. https://doi.org/10.3390/en12020289

Ehteram M, Mousavi SF, Karami H, Farzin S, Emami M, Othman FB, El-Shafie A (2017) Fast convergence optimization model for single and multi-purposes reservoirs using hybrid algorithm. Adv Eng Inform 32:287–298

Ehteram M, Karami H, Farzin S (2018a) Reservoir optimization for energy production using a new evolutionary algorithm based on multi-criteria decision-making models. Water Resour Manage 32(7):2539–2560

Ehteram M, Karami H, Farzin S (2018b) Reducing irrigation deficiencies based optimizing model for multi-reservoir systems utilizing spider monkey algorithm. Water Resour Manage 32(7):2315–2334

Ehteram M, Khozani ZS, Soltani-Mohammadi S, Abbaszadeh M (2023) Structure of Different Kinds of ANN Models. In Estimating Ore Grade Using Evolutionary Machine Learning Models (pp. 13–26). Singapore: Springer Nature Singapore

Fallah A, Rakhshandehroo GR, Berg POS, Orth R (2020) Evaluation of precipitation datasets against local observations in southwestern Iran. Int J Climatol 40(9):4102–4116

Faris H, Mafarja MM, Heidari AA, Aljarah I, Al-Zoubi AM, Mirjalili S, Fujita H (2018) An efficient binary Salp Swarm Algorithm with crossover scheme for feature selection problems. Knowl-Based Syst. https://doi.org/10.1016/j.knosys.2018.05.009

Gao M, Li J, Hong F, Long D (2019) Day-ahead power forecasting in a large-scale photovoltaic plant based on weather classification using LSTM. Energy. https://doi.org/10.1016/j.energy.2019.07.168

Guo L, Chen J, Wu F, Wang M (2018) An electric power generation forecasting method using support vector machine. Syst Sci Control Eng. https://doi.org/10.1080/21642583.2018.1544947

Hanoon MS, Ahmed AN, Razzaq A, Oudah AY, Alkhayyat A, Huang YF, El-Shafie A (2022) Prediction of hydropower generation via machine learning algorithms at three Gorges Dam, China. Ain Shams Eng J 101919

He YL, Chen L, Gao Y, Ma JH, Xu Y, Zhu QX (2022) Novel double-layer bidirectional LSTM network with improved attention mechanism for predicting energy consumption. ISA Trans. https://doi.org/10.1016/j.isatra.2021.08.030

Heidari AA, Yin Y, Mafarja M, Jalali SMJ, Dong JS, Mirjalili S (2020). Efficient moth-flame-based neuroevolution models. https://doi.org/10.1007/978-981-32-9990-0_4

Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780

Hou R, Li S, Wu M, Ren G, Gao W, Khayatnezhad M, gholinia, F. (2021) Assessing of impact climate parameters on the gap between hydropower supply and electricity demand by RCPs scenarios and optimized ANN by the improved Pathfinder (IPF) algorithm. Energy. https://doi.org/10.1016/j.energy.2021.121621

Huang Y, Liu S, Yang L (2018) Wind speed forecasting method using EEMD and the combination forecasting method based on GPR and LSTM. Sustainability (Switzerland). https://doi.org/10.3390/su10103693

Imrana Y, Xiang Y, Ali L, Abdul-Rauf Z (2021) A bidirectional LSTM deep learning approach for intrusion detection. Expert Syst Appl. https://doi.org/10.1016/j.eswa.2021.115524

Jamei M, Ali M, Malik A, Prasad R, Abdulla S, Yaseen ZM (2022) Forecasting daily flood water level using hybrid advanced machine learning based time-varying filtered empirical mode decomposition approach. Water Resour Manage 36(12):4637–4676

Jaseena KU, Kovoor BC (2021) Decomposition-based hybrid wind speed forecasting model using deep bidirectional LSTM networks. Energy Convers Manage. https://doi.org/10.1016/j.enconman.2021.113944

Li BJ, Sun GL, Liu Y, Wang WC, Huang XD (2022) Monthly runoff forecasting using variational mode decomposition coupled with gray wolf optimizer-based long short-term memory neural networks. Water Resour Manage 36(6):2095–2115

Liu Y, Pu H, Sun DW (2021) Efficient extraction of deep image features using convolutional neural network (CNN) for applications in detecting and analysing complex food matrices. Trends Food Sci Technol. https://doi.org/10.1016/j.tifs.2021.04.042

Panahi F, Ehteram M, Ahmed AN, Huang YF, Mosavi A, El-Shafie A (2021) Streamflow prediction with large climate indices using several hybrid multilayer perceptrons and copula Bayesian model averaging. Ecol Ind 133:108285

Panda N, Majhi SK (2020) Improved salp swarm algorithm with space transformation search for training neural network. Arab J Sci Eng. https://doi.org/10.1007/s13369-019-04132-x

Rahman MM, Shakeri M, Tiong SK, Khatun F, Amin N, Pasupuleti J, Hasan MK (2021) Prospective methodologies in hybrid renewable energy systems for energy prediction using artificial neural networks. Sustainability (Switzerland). https://doi.org/10.3390/su13042393

Samantaray S, Sawan Das S, Sahoo A, Prakash Satapathy D (2022) Monthly runoff prediction at Baitarani river basin by support vector machine based on Salp swarm algorithm. Ain Shams Eng J. https://doi.org/10.1016/j.asej.2022.101732

Seifi A, Ehteram M, Soroush F, Haghighi AT (2022) Multi-model ensemble prediction of pan evaporation based on the Copula Bayesian Model Averaging approach. Eng Appl Artif Intell 114:105124

Shadrin D, Nikitin A, Tregubova P, Terekhova V, Jana R, Matveev S, Pukalchik M (2021) An automated approach to groundwater quality monitoring-geospatial mapping based on combined application of gaussian process regression and bayesian information criterion. Water (Switzerland). https://doi.org/10.3390/w13040400

Sharifzadeh F, Akbarizadeh G, Seifi Kavian Y (2019) Ship classification in SAR images using a new hybrid CNN–MLP classifier. J Indian Soc Remote Sens 47(4):551–562

Sinitsin V, Ibryaeva O, Sakovskaya V, Eremeeva V (2022) Intelligent bearing fault diagnosis method combining mixed input and hybrid CNN-MLP model. Mech Syst Signal Process 180:109454

Sun Q, Tang Z, Gao J, Zhang G (2022) Short-term ship motion attitude prediction based on LSTM and GPR. Appl Ocean Res. https://doi.org/10.1016/j.apor.2021.102927

Tang S, Zhu Y, Yuan S (2021) An improved convolutional neural network with an adaptable learning rate towards multi-signal fault diagnosis of hydraulic piston pump. Adv Eng Inform. https://doi.org/10.1016/j.aei.2021.101406

Wang H, Zhang YM, Mao JX, Wan HP (2020a) A probabilistic approach for short-term prediction of wind gust speed using ensemble learning. J Wind Eng Ind Aerodyn. https://doi.org/10.1016/j.jweia.2020.104198

Wang W, Zhou C, He H, Wu W, Zhuang W, Shen XS (2020b) Cellular traffic load prediction with LSTM and Gaussian process regression. IEEE International Conference on Communications. https://doi.org/10.1109/ICC40277.2020.9148738

Wang Y, Feng B, Hua QS, Sun L (2021) Short-term solar power forecasting: a combined long short-term memory and gaussian process regression method. Sustainability (Switzerland). https://doi.org/10.3390/su13073665

Zha W, Liu Y, Wan Y, Luo R, Li D, Yang S, Xu Y (2022) Forecasting monthly gas field production based on the CNN-LSTM model. Energy 124889

Zhang H, Cai Z, Ye X, Wang M, Kuang F, Chen H, Li C, Li Y (2022) A multi-strategy enhanced salp swarm algorithm for global optimization. Eng Comput. https://doi.org/10.1007/s00366-020-01099-4

Zhao M, Fu X, Zhang Y, Meng L, Tang B (2022) Highly imbalanced fault diagnosis of mechanical systems based on wavelet packet distortion and convolutional neural networks. Adv Eng Inform 51:101535

Zhou P, Shi W, Tian J, Qi Z, Li B, Hao H, Xu B (2016). Attention-based bidirectional long short-term memory networks for relation classification. In Proceedings of the 54th annual meeting of the association for computational linguistics (volume 2: Short papers), pp. 207–212

Zou Z, Ergan S (2023) Towards emotionally intelligent buildings: a Convolutional neural network based approach to classify human emotional experience in virtual built environments. Adv Eng Inform 55:101868

Zolfaghari M, Golabi MR (2021) Modeling and predicting the electricity production in hydropower using conjunction of wavelet transform, long short-term memory and random forest models. Renew Energy 170:1367–1381

Authors and Affiliations

Mohammad Ehtearm[1] · Hossein Ghayoumi Zadeh[2] · Akram Seifi[3] · Ali Fayazi[2] · Majid Dehghani[4]

[1] Department of Water Engineering, Semnan University, Semnan, Iran

[2] Department of Electrical Engineering, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran

[3] Department of Water Science & Engineering, College of Agriculture, Vali-e-Asr University of Rafsanjan, P.O. Box 518, Rafsanjan, Iran

[4] Department of Technical and Engineering, Faculty of Civil Engineering, Vali-e-Asr University of Rafsanjan, P.O. Box 518, Rafsanjan, Iran