



# Smart hydropower management: utilizing machine learning and deep learning method to enhance dam's energy generation efficiency

Muhammet Emin Sahin<sup>1</sup> · Mucella Ozbay Karakus<sup>1</sup>

Received: 24 April 2023 / Accepted: 19 February 2024  
© The Author(s) 2024

## Abstract

Renewable energy sources and hydroelectric power generation in large parts of the electricity market are crucial as environmental pollution worsens. Utilizing meteorological data from the region, where the Hirfanlı Dam is located, this study employs machine learning (ML) and introduces a novel hybrid Genetic Grey Wolf Optimizer (GGWO)-based Convolutional Neural Network/Recurrent Neural Network (CNN/RNN) regression technique to predict hydroelectric power production (PP). In the first section of the study, various ML techniques SVR (Support Vector Regression), ELM (Extreme Learning Machine), RFR (Random Forest Regression), ANN (Artificial Neural Networks) and WKNNR (Weighted K-Nearest Neighbor) are presented with the Principal Component Analysis (PCA) method and the minimum–maximum method in the normalization of the features. A novel GGWO and CNN/RNN model-Long Short-Term Memory (LSTM) regression technique is introduced in the second section. GGWO is used to select features, while the proposed CNN/RNN-LSTM model is employed for feature extraction and prediction of PP. The study demonstrates that the ELM algorithm in Method I outperforms other ML models, achieving a correlation coefficient ( $r$ ) of 0.977 and the mean absolute error (MAE) of 0.4 with the best feature subset. Additionally, the proposed CNN/RNN hybrid model in Method II yields even better results, with  $r$  and MAE values of 0.9802 and 0.314, respectively. The research contributes to the field of renewable energy prediction, and the results can aid in efficient decision making for electricity generation and resource management.

**Keywords** Deep learning · Green power · Hybrid model · Hydropower production prediction · Machine learning

## 1 Introduction

The combination of excessive population growth, uncontrolled depletion of natural resources and global warming, which contributes to climate change, has heightened the likelihood of uncertainty in the natural environment and disrupted ecological balance. The surge in population growth is directly linked to an escalating demand for food and energy. The same reason can be attributed to the uncontrolled removal of natural resources and the transformation of forest lands into agricultural or industrial land. Increasing imports in the energy sector also trigger the price increase per liter of gasoline, diesel, coal and fossil

fuels. For this reason; the search for an alternative energy source has become extremely important and extremely necessary to maintain the energy security of the population. Solar energy, wind energy, geothermal and hydro energy or flowing water are the most appropriate alternatives to the existing energy source. All these energy sources can be infinite and renewed. Hydroelectric is one of the main renewable energy sources due to its low cost and its potential to rapidly respond to the highest demand load on electricity and its potential [1]. In addition, hydroelectric is a very clean source of energy and uses only water. Water can also be used for other purposes after producing electrical energy. Hydroelectric is currently considered to be the largest renewable source of electricity in the world. Traditionally, it is a cheap and clean source of electricity. However, renewable energy often has various properties that cause uncertainties in renewable energy power systems. Therefore, the prediction of renewable energy is an

✉ Muhammet Emin Sahin  
emin.sahin@bozok.edu.tr

<sup>1</sup> Department of Computer Engineering, Yozgat Bozok University, Yozgat, Turkey

important way to cope with this problem [2]. Machine learning is a data-oriented process used to create a smart result. The basic steps of machine education include data collection and preprocessing, feature selection and extraction, model selection and model verification.

Machine learning (ML) techniques contribute to the use of these accumulated data and become increasingly important in various fields of science, finance and industry [1, 3]. Several studies were conducted based on the forecasting of the dam's inflow and the application of an operating reservoir model to calculate the generation of hydropower [4–7]. Dehghani et al. In their study, they combined the Grey Wolf Optimization (GWO) method with an adaptive neuro-fuzzy inference system (ANFIS) to predict hydropower generation. The results of this paper showed that the GWO-ANFIS method could satisfactorily predict hydropower generation, while ANFIS failed in nine input–output combinations [7]. Condemni et al. proposed a methodology for estimating hydropower capacity and feature summarization mechanisms (such as principal component analysis and feature grouping techniques) using machine learning (ML) regression techniques. They showed how ML regression techniques can be used to model the relationship between meteorological and climatic variables and hydroelectric generation using water tanks. In the study, MLP shows the best results in the problem with 0.2593 and 0.2128 TWH root mean square error and mean absolute error [8]. Another study was carried out by Al Rayess and Keskin, which presents the results of Machine Learning Techniques for short-term estimation of the amount of energy produced by General Circulation Models (GCMS) data by the Almus Dam and Hydroelectric Power Plant in Tokat, Turkey [9]. The results show that the correlation value of the gradient-enhanced tree model is equal to 0.717, indicating that the gradient-enhanced tree model is the most successful model for the available data. Li et al. in their study, they explore the potential of a GA-SVM (Genetic algorithms-support vector machine (SVM)) estimation model to accurately predict short-term power generation energy from small hydroelectric resources [10]. This article concludes that the GA-SVM model is a very potential candidate for estimating small hydroelectric (SHP) short-term power generation energy.

This article presents a study on the prediction of hydroelectric plant production using ML and a novel hybrid Genetic Grey Wolf Optimizer (GGWO)-based Convolutional Neural Network/Recurrent Neural Network (CNN/RNN)-Long Short-Term Memory (LSTM) regression techniques and an overview of the study is given Fig. 1. According to meteorological data, a hydroelectric power production (PP) is determined by the amount of electricity that can be produced while taking into account the level of water in its reservoir.

The article's remaining sections are organized as follows: Chapter 2 describes the data collection process, the dataset used, the two main methods of prediction used: Method I (Machine Learning Algorithms) and Method II (CNN/RNN Hybrid Method) and evaluation metrics. In Chapter 3, the study focuses on the application of machine learning algorithms to predict hydroelectric power production. The feature selection process using PCA (Principal Component Analysis) and the performance metrics of various machine learning models are discussed. In addition proposed CNN/RNN hybrid method for power production prediction in this section. The feature selection using the GGWO (Grey Wolf Optimizer) method is explained, and the parameters of the deep learning models (CNN and LSTM) are detailed. The findings of the study from both Method I and Method II are presented and analyzed in Chapter 3. The evaluation metrics for each prediction approach are compared, and the strengths and weaknesses of each method are discussed in Chapter 5. Chapter 6 provides a comprehensive conclusion, summarizing the key findings of the study.

The suggested methods include contributions below:

- The used dataset's feature which has time series of the feature is selected with the Principal Component Analysis (PCA) method by using the minimum–maximum method in the normalization of the features,
- In the study, the prediction process firstly ML techniques; SVR (Support Vector Regression), ELM (Extreme Learning Machine), RFR (Random Forest Regression), ANN (Artificial Neural Networks), WKNNR (Weighted K-Nearest Neighbor),
- In the study, the data are also predicted with the help of the new proposed hybrid Genetic Grey Wolf Optimizer (GGWO)-based Convolutional Neural Network/Recurrent Neural Network (CNN/RNN)-Long Short-Term Memory (LSTM) regression techniques,
- Within the scope of the study, correlation coefficient ( $r$ ), coefficient of determination (also known as  $R$ -squared or  $R^2$ ), mean square error (MSE) and its rooted variant (RMSE), mean absolute error (MAE) and their normalized equivalents (RRMSE, RMAE)) values are obtained,
- Comparative presentation of existing studies in the literature is provided.

## 2 Material and methods

This study predicts the electrical power production as well as the net head value representing the net water height of the Hirfanlı dam (this value is the primary component used for electricity production). In this part, dataset, data

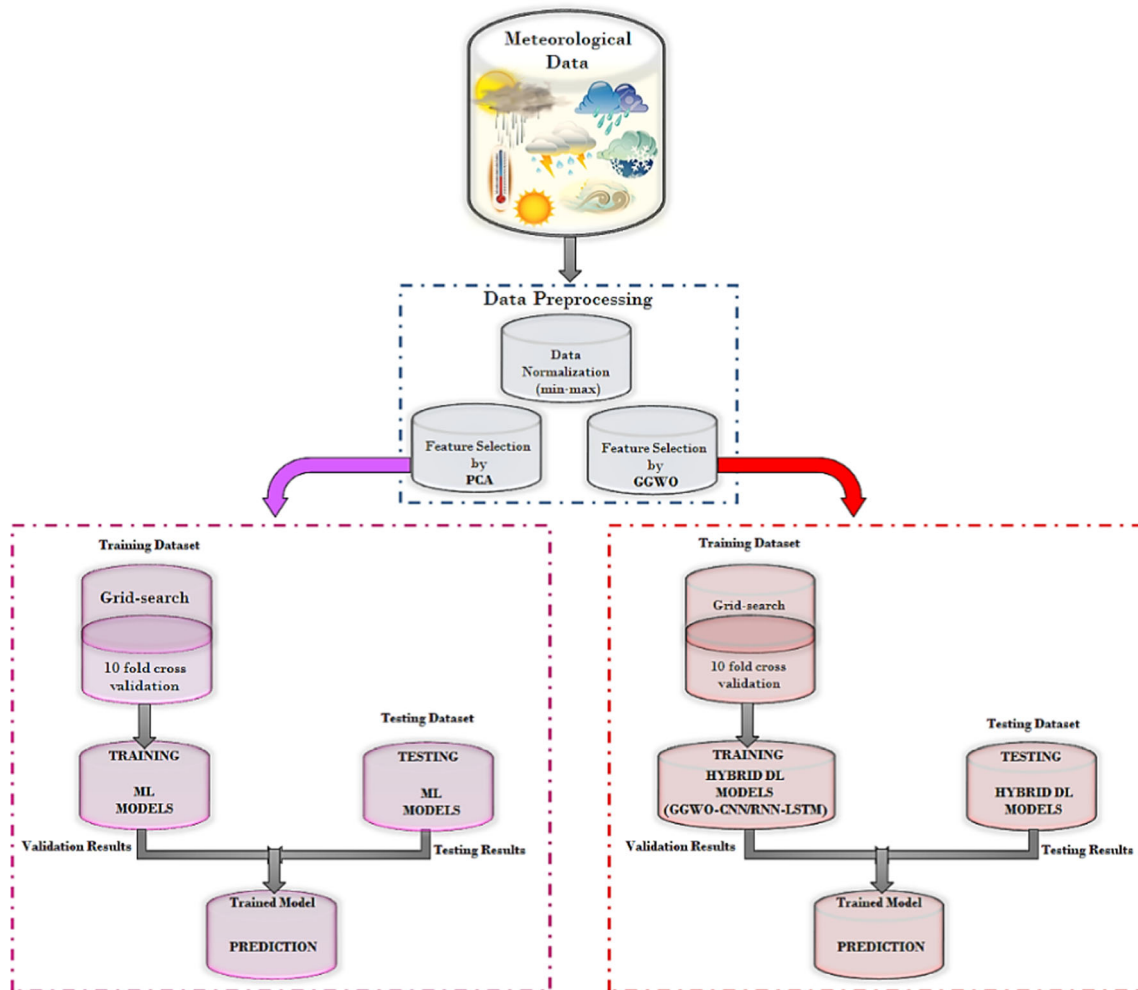


Fig. 1 An overview of the study

preprocessing, used algorithms and evaluation metrics are introduced.

## 2.1 Dataset

The final determinant of a region's hydropower potential is typically related to water scarcity, which is in turn affected by precipitation and/or snowmelt, temperature and solar radiation. Therefore, it is evident that meteorological events have the greatest impact on the electrical generation of the hydroelectric power production (HEPP). However, there is no direct correlation between these variations in meteorological conditions and HEPP capacity. This indicates that it takes some time for precipitation to turn into water that can be used in the HEPP after it occurs [8]. Furthermore, the impact of meteorological events over time might not be consistent.

The Hirfanlı Dam-HEPP, a dam constructed for flood control, energy generation and irrigation purposes between 1953 and 1959, is chosen as the study's sample plane. It is

situated on the Kızılırmak River in Kaman/Kırşehir, Turkey. Hirfanlı HEPP, whose image and location in Turkey are shown in Fig. 2, is the second-largest HEPP in Kırşehir and produces 400 GWh of energy annually. It is located in Türkiye and has an installed capacity of 128 MWe. The total electrical power generation capacity of a HEPP is determined by the amount of electricity that can be produced based on the amount of water that is available and, consequently, the information on its net head, and the amount of water inlet originates from natural sources.

Starting from this perspective, Kaman Meteorology Directorate and Hirfanlı HEPP collected meteorological data and data on power production over a 15-year period, from 1 January 2007 to 31 December 2021, and the author organized the data to produce an original dataset. There are many different working variables because all meteorological and electrical generation data are determined daily. Two government organizations (Kaman/Kırşehir Weather Directorate and Hirfanlı Hydropower Station) have provided and verified these historical data. The dataset used in



Fig. 2 An overview of the study

this study comprises valuable information from Table 1, encompassing essential meteorological features such as the highest, lowest and average air temperatures, as well as details regarding precipitation type, amount and intensity, spanning 15 years (365 days each year). In total, there are 64 variables captured daily over the entire period. This extensive dataset allows for a comprehensive analysis of factors influencing changes in water availability and electricity production in the region. By analyzing such a vast amount of data (350,400 data points in total), researchers gain a deeper understanding of the intricate relationships between meteorological variables and their impact on water resources and energy generation.

## 2.2 Data preprocessing

Data preprocessing tasks can significantly increase classification accuracy, depending on the domain and language, according to extensive experimental analyses. In addition to improving classification accuracy, it also has many other advantages like improved performance, smaller data sets and faster training cycles. When using data mining, it is not always appropriate to use raw data. When the value ranges of the variables in the data set differ from one another, normalization techniques are used. The pressure of variables with large mean and variance on other variables can be increased as a result of the variables' significantly different means and variances, which can have an impact on

accuracy and performance. To standardize the impact of each variable on the results, numerical data are normalized using normalization methods in data mining [11, 12]. Data are linearly normalized using the minimum-to-maximum approach. In other words, neither the magnitude nor the difference between the two values changes. The range of values after normalization is typically 0 to 1; minimum represents the lowest value that a data point can accept, and maximum represents the highest value that a data point can accept [13]. The minimum–maximum normalization procedure is carried out for the data used in this study during the data preprocessing stage. The values are derived from Eq. (1) for the minimum–maximum model.

$$\hat{x} = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

$\hat{x}$ : Normalized data,  $x_i$ : Input value,  $x_{\max}$ : Largest number in input,  $x_{\min}$ : Smallest number in the input.

## 2.3 Feature selection

Feature selection and Principal Component Analysis (PCA) are common techniques used in machine learning regression tasks to improve model performance and reduce the impact of overfitting. Feature selection involves selecting a subset of the most informative features from a dataset, while discarding redundant or irrelevant features. This can help to reduce the dimensionality of the data and improve

**Table 1** Descriptive information of daily meteorological and power production parameters for Hirfanlı HEPP

#	Feature	Unit	Range	Notation
<i>Meteorological Features</i>				
1	Highest Air Temperature	°C	(− 9.2) to (37.4)	HAT
2	Lowest Air Temperature	°C	(− 15.3) to (22.2)	LAT
3	Average Air Temperature	°C	(− 12.1) to (32.7)	AAT
4	Sunshine Duration	h	(0)–(13.5)	SD
5	Average Wind Direction	–	(9.7)–(359.6)	AWD
6	Average Wind Speed	m/s	(0.7)–(9.0)	AWS
7	Highest Wind Speed	m/s	(1.9)–(23.2)	HWS
8	Maximum Wind Time	h	(0)–(23.6)	MWT
<i>Precipitation Features</i>				
9	Precipitation Amount	mm	0–33.6	PA
10	Type of Precipitation	–	1–8	TOP
11	Snow Thickness	mm	(0)–(22.0)	ST
<i>Evaporation and Water Features</i>				
12	Open Surface Evaporation Amount	mm	(5)–(99)	OSE
13	Water Evaporating from the Plant	m <sup>3</sup>	(0)–(1.292.550)	WEP
14	Average Humidity	%	(5.0)–(99.0)	AH
15	Water for Drinking and Use	m <sup>3</sup>	(0)–(5000)	WDU
16	Leakage and Other Water Losses	m <sup>3</sup>	(86.400)–(1.527.289)	LOWL
17	Total Water Consumption	1000 m <sup>3</sup>	(91.4)–(16.928)	TWC
18	Total Incoming Water	m <sup>3</sup>	(35.455)–(8.930.447)	TIW
19	Flow Rate of Incoming Water	m <sup>3</sup> /s	(0.41)–(103.36)	FRIW
20	Lake Level at 00:00	m	(842.12)–(844.88)	LL
21	Volume Difference in the Lake (+/−)	1000 m <sup>3</sup>	(− 10.270) to (6.460)	VDL
<i>Pool Water Features</i>				
22	Pool Water Temperature	°C	(2.1)–(22.6)	PWT
23	Maximum Pool Water Temperature	°C	(12.4)–(34.8)	MXPT
24	Minimum Pool Water Temperature	°C	(1.1)–(18.2)	MNPT
25	Lake or Pool Level (Average)	m	(744.32)–(844.9)	LPL
26	Leaving Water Level (Average)	m	(783.1)–(786.35)	LWL
<i>HPP Features</i>				
27	Power Plant Internal Consumption	kWh	(745.1)–(746.1)	PPIC
28	Internal Consumption	kWh	(60)–(26.360)	IC
29	Ratio of Water to Energy	m <sup>3</sup> /kWh	(0)–(8.12)	RWE
30	Energy Consumed Water	m <sup>3</sup>	(0)–(15.857.879)	ECW
31	Net Head	m	(40.48)–(60.22)	NH
32	Power Production	MWh	(5000)–(2149 000)	PP
<i>Generator Performance Features (4 Generator)</i>				
33	Generator Working Time	h	(2.321)–(7.275)	GWT GR1–4
34	Uptime Since Last Revision	h	(80.08)–(298.92)	USLR GR1–4
35	Unit Fault Stop	h	(0)–(1)	UFS GR1–4
36	Switchboard Failed Stop	h	(0)–(1)	SWFS GR1–4
37	System Fault Stop	h	(0)–(0.03)	SFS GR1–4
38	Maintenance Downtime	h	(0)–(1)	MD GR1–4
39	Normal Downtime	h	(0)–(1.04)	ND GR1–4
40	Power Supplied From Output (4 Lines)	1000 MWh	(1008)–(35.488)	PSFO1–4
41	Output Feeders (11 Feeders)	1000 MWh	(1,130)–(725.630)	OF1–11
42	Power from the Feeders (5 Feeders)	1000 MWh	(19.992)–(763.756)	PFF1–5



**Table 1** (continued)

#	Feature	Unit	Range	Notation
43	Compensator Consumption (5 unit)	MWh	(73.830)–(293.930)	CC C1–5

the speed and accuracy of the model. PCA, on the other hand, involves transforming the original features into a new set of uncorrelated features, called principal components that capture the most variance in the data. This can also help to reduce the dimensionality of the data while retaining as much information as possible. A multivariate system with many variables that are related to one another is transformed by PCA into a system with new variables that, despite being less numerous and not related to one another as linear functions, can best explain the overall change of the prior system [14]. It is a method of multivariate statistical analysis. This technique transforms  $n$  features into  $m$  distinct features.

## 2.4 Algorithms

In essentially, classification is a technique, whereby certain judgments or predictions are made after learning and interpreting the information that is available, and the data are then divided into classes with labels based on those properties. There are many different fields, where the classification task is employed. Different algorithms are used by machine learning to solve issues. Various factors, including the nature of the problem, the quantity of variables and the kind of model, affect the algorithm that is employed [15–17]. The results of the study are obtained using 5 different machine learning algorithms, which are SVR (Support Vector Regression), ELM (Extreme Learning Machine), RFR (Random Forest Regression), ANN (Artificial Neural Networks), and WKNNR (Weighted K-Nearest Neighbor). The selected machine learning (ML) techniques (SVR, ELM, RFR, ANN, WKNNR) in the study are chosen based on their suitability for the prediction task of hydroelectric power production (PP) using meteorological data. Each technique has its strengths and advantages, making them valuable candidates for this specific regression problem.

## 2.5 Grid search

Machine learning models contain crucial parameters. Each method has its own set of parameters. These variables have a big impact on how well the model works. As a result, it is crucial to know the values of the parameters. The grid

search method is one way to determine which parameter has what value so that the best model can be created. This approach involves processing each possible value for each parameter one at a time, calculating success and identifying the parameter values that are most appropriate. Hyperparameter fine-tuning can be done in the form of manual search, grid search, random search [18] and Bayesian optimization [19, 20]. The results of the grid search used to perform a thorough search and fine-tuning of all possible combinations of the specified hyperparameters are presented in this section.

## 2.6 Cross-validation

Cross-validation is a statistical technique used to assess the performance and generalization capability of a machine learning model [20]. It involves dividing the available data into multiple subsets, or “folds,” and using these folds for training and testing the model in a rotation [21]. The basic idea behind cross-validation is to ensure that the model’s performance is evaluated on different subsets of data to obtain a more reliable estimate of its effectiveness. In our study, classification is carried out using the tenfold cross-validation technique, and the results are reported.

## 2.7 Evaluation methods

Various metrics are utilized to assess the performance of predicting operations. Error values that are high indicate poor prediction accuracy. Some of the evaluation criteria that are used to determine the prediction error rate and accuracy are discussed in this section.

$$R = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (2)$$

$R$  = correlation coefficient,  $x_i$  = quantities of the variable of  $x$ ,  $\bar{x}$  = mean of the quantities of  $x$ -variable,  $y_i$  = values of the variable  $y$ -variable,  $\bar{y}$  = mean of the quantities of  $y$ -variable

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (3)$$

RMSE = root mean squared error,  $n$ : number of data points,  $\hat{y}_i$  = the predicted values,  $y_i$  = the observed values

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (4)$$

MSE = mean squared error,  $n$  = number of data points,  $\hat{y}_i$  = the predicted values,  $y_i$  = the observed values

$$\text{MAE} = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (5)$$

MAE = mean absolute error,  $n$  = number of data points,  $\hat{y}_i$  = the predicted values,  $y_i$  = the observed values.

### 3 Experimental results

The dataset obtained in this study is applied to machine learning (Method I) and deep learning (Method II) algorithms. For the Method I study, feature selection is performed for the features used in the dataset and analyzed in five different machine learning models. In the Method II, a novel GGWO and CNN/RNN model)-Long Short-Term Memory (LSTM) regression technique is introduced. Feature selection is done by using GGWO, and the proposed CNN/RNN-LSTM model is used for feature extraction and prediction of power production (PP).

#### 3.1 Method I: prediction with machine learning algorithms

In this study, the net head value representing the net water height of the Hirfanlı dam (this value is the main component used for electricity generation) and the electrical power production are predicted. Within the scope of this study, the results are obtained by using machine learning

**Table 2** Selected features after using PCA method

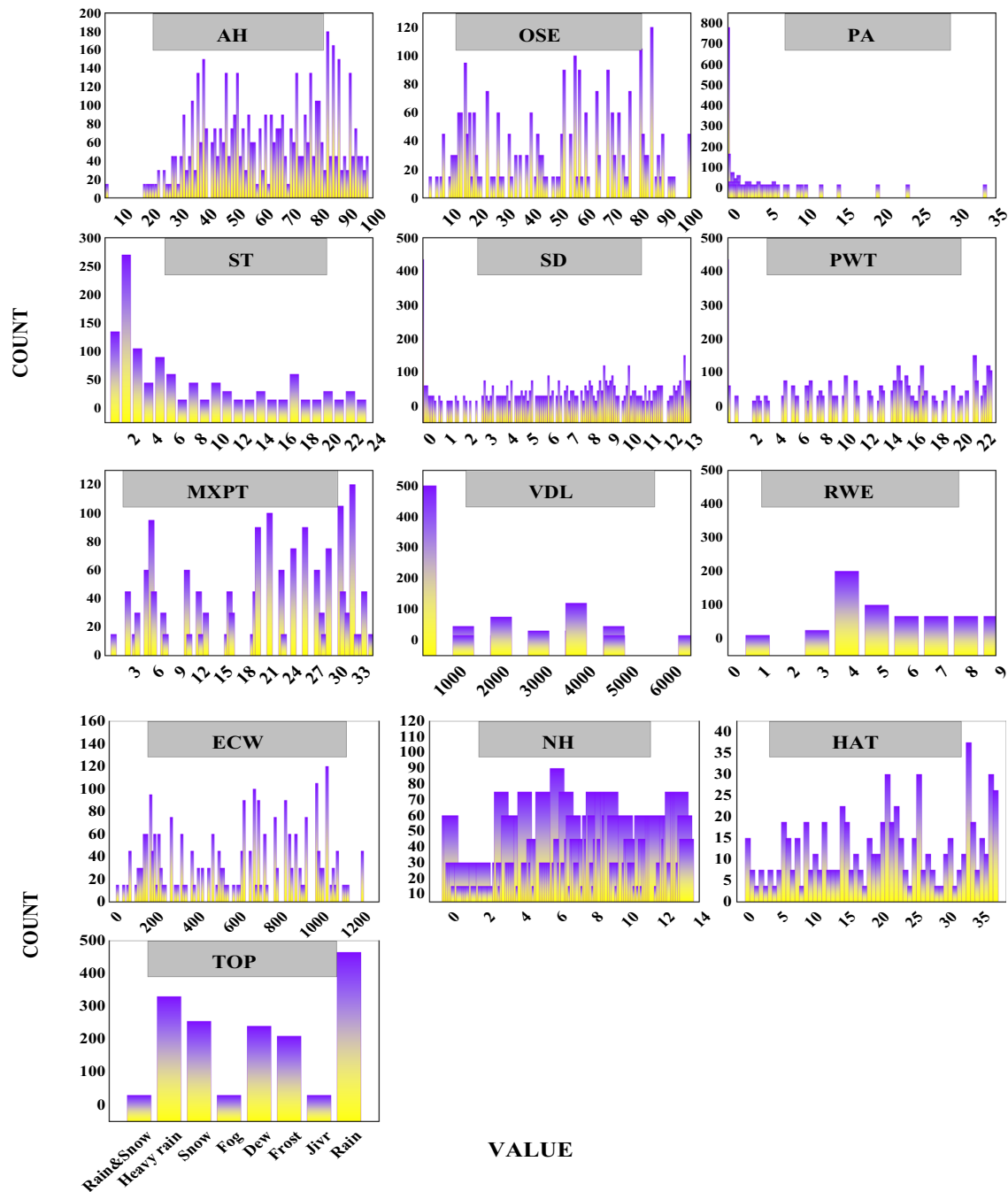
#	Feature	Notation
1	Highest air temperature	HAT
2	Volume difference in the lake (+/−)	VDL
3	Type of precipitation	TOP
4	Precipitation amount	PA
5	Open surface evaporation Amount	OSE
6	Sunshine duration	SD
7	Temperature of the pool water	PWT
8	Maximum temperature of the pool water	MXPT
9	Net head	NH
10	Ratio of water to energy	RWE
11	Energy consumed water	ECW
12	Snow thickness	ST
13	Average humidity	AH

and deep learning techniques separately. While 80% of the dataset is used for the training phase, 20% is used for the testing phase. For the test data, the data of 2011, 2016 and 2021 are chosen randomly. In this section, the feature matrix is reduced from the related variable in numbers 1 to 13 using the PCA method in Table 2. Input variables selected for PP prediction are given in Table 2. As histogram plots of the dataset, Fig. 3 also displays the frequency of some features.

In this study, the grid search method is employed for hyperparameter tuning, which is a systematic approach to finding the optimal combination of hyperparameters for a machine learning model. The goal of hyperparameter tuning is to enhance the model's performance by finding the parameter values that yield the best results on the validation set. Table 3 presents the specific parameter values used in the grid search. The grid search involves defining a set of possible values for each hyperparameter and then exhaustively searching through all possible combinations of these values. For each combination, the model is trained and evaluated using a predefined performance metric. By using the grid search method, researchers can efficiently explore a wide range of hyperparameter configurations to identify the optimal ones that lead to improved model performance.

Figure 4 displays the correlation between the features and the class label in the dataset. Upon examination, it becomes evident that there exists a strong correlation between the features NH, PWT and TOP, and the class label (PP). The high correlation observed between NH, PWT and TOP with the class label (PP) indicates that these features hold significant predictive power for determining the class label. In other words, variations in NH, PWT and TOP are closely associated with changes in the class label (PP), making them valuable indicators for the classification task. The results of ELM, which is the best classifier used in the study, are presented. The Power Production (GWh) vs Power Production prediction (GWh) graph for the ELM is given in Fig. 5. The importance of the features for the ELM model is shown in Fig. 6.

Within the scope of the study, correlation coefficient ( $r$ ), coefficient of determination (also known as R-squared or  $R^2$ ), mean square error (MSE) and its rooted variant (RMSE), mean absolute error (MAE) and their normalized equivalents (RRMSE, RMAE)) values are obtained and the values obtained with the PCA method and without using the PCA method are given in Table 4. According to the MAE, MSE, RMSE, MAPE and  $R^2$  metrics of the five machine learning algorithm models tested for the prediction of Hirfanlı hydroelectric power production, the most successful result is obtained with the ELM algorithm model. When the ELM algorithm is used in the estimation process using the PCA method, the  $r$ ,  $R^2$ , RMSE and MAE values are obtained as 0.977, 0.955, 0.56 and 0.40,



**Fig. 3** Distribution of the selected features of data for the input dataset as histogram plots ( $x$ -axis = VALUE  $\times$  1000)

respectively, and they are the values with the best results. When Table 4 is examined, RMSE and MAE values are obtained as 0.56 and 0.4, respectively. In the literature, MSE and MAE values below 0.4 TWh are considered very good values and predictions over 1 TWh are regarded as being very poor [18]. The study highlights the importance of evaluating machine learning models using multiple metrics and comparing the results with actual values. The study also suggests that combining feature selection techniques such as PCA with machine learning algorithms can

lead to improved prediction accuracy and reduced computational time.

## 3.2 Method II CNN/RNN hybrid method

### 3.2.1 Feature selection-based Grey Wolf Optimizer and genetic algorithm

Feature Selection (FS) is a popular and complex research area in various domains, such as machine learning, data



**Table 3** The hyperparameters of the used models

Model	Hyperparameters	Hyperparameter selection	Power production
SVR	Filter 1	[50, 80, 100, 200, 300, 400]	200
	Filter 2	[40, 50, 60, 70, 80]	60
	Filter 3	[20, 10, 30, 5]	10
	Batch size	[25, 50, 100, 400]	100
	Activation function	ReLU	
	Epochs	[1000, 1200, 300, 400, 700]	400
	Optimization algorithm	Adam	
	SVR Kernel	RBF	
ELM	Hidden Neuron	[20, 30, 40, 50]	40
	Activation Function	log, tanh	tanh
	Learning Rate	[0.005, 0.01, 0.1]	0.1
ANN	Activation Function	ReLU	
	Optimization Algorithm	Adam	
	Hidden Layer Structure	[1, 2, 3]	3
	Number of neurons	[N, 2N, 3N]	2N
	Batch Size	[1,10]	10
RFR	The tree's maximum depth	[5, 8, 10, 20, 25]	10
	The tree number of the forest	[50, 100, 150, 200, 400, 600, 800]	100
	The minimum sample number for separating an internal node	[2, 4, 6, 8, 10]	8
	The feature number of considering for the best separation	['auto', 'log2', 'sqrt']	auto
WKNNR	Number of neighbors	[5, 10, 20, 30, 50, 100]	20
	The algorithm that utilized for computing the nearest neighbor	['auto', 'kd_tree', 'ball_tree']	auto

mining and pattern recognition. It involves the identification of significant and informative attributes from a dataset, which is a crucial step in various applications, including predictive data analysis [22]. While FS is an attractive research topic, it also presents challenges that researchers must address to achieve accurate and efficient results. The authors have outlined three primary objectives of utilizing feature selection in classification [22]. Firstly, it aims to simplify the process of feature extraction, which can be a complex and time-consuming task. Secondly, it seeks to enhance the accuracy of classification models by selecting the most informative features that are relevant to the task at hand. Finally, feature selection is implemented to improve the reliability of performance estimation, ensuring that the classification model is producing consistent and dependable results. These objectives demonstrate the value and significance of feature selection in classification tasks across a range of domains.

*Grey Wolf Optimizer (GWO)*: The proposed Grey Wolf Optimization algorithm imitates the social and hunting behavior of grey wolves. Grey wolves are divided into *alpha*, *beta*, *delta* and *omega* social groups. Because the wolf group obeys its rules, the alpha group is the dominant species. Secondary grey known as *beta* class support *alpha*

in making decisions. The lowest ranking grey wolves are represented by Omega. A wolf is referred to as a delta if it does not fall under one of the aforementioned species. Along with their social interactions, grey wolves have an intriguing social behavior called group hunting. The phases of containing, hunting and attacking prey are the GWO's main components [23, 24].

The feature selection problem involves selecting a subset of features to be used in a model [22]. To achieve this, researchers often represent potential subsets as binary strings of the same length as the total number of features. Each bit in the string indicates whether a corresponding feature is included or excluded from the subset. A value of '1' indicates that the feature is selected, whereas a '0' indicates exclusion. To model the social hierarchy of wolves, researchers typically use a system in which the most dominant wolves are considered *alpha* ( $\alpha$ ) and responsible for making decisions. *Beta* ( $\beta$ ) wolves assist *alpha* in decision making and provide feedback. *Delta* ( $\delta$ ) wolves act as scouts and control the actions of *omega* ( $\omega$ ) wolves, which must obey all other wolves. In the GWO (Grey Wolf Optimizer) algorithm, *alpha*, *beta* and *delta* wolves lead the hunting process while *omega* wolves

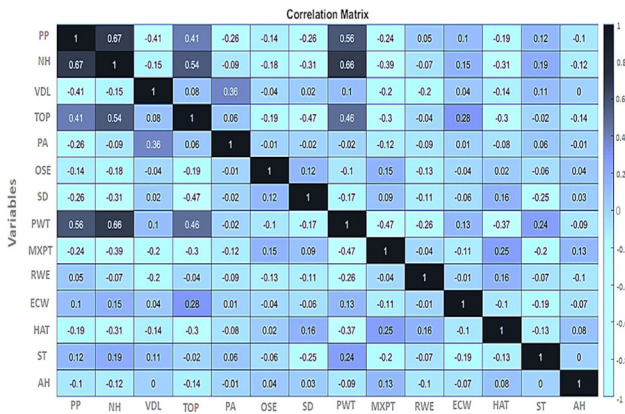


Fig. 4 The correlation coefficient matrix of the predictor parameters for the daily meteorological data (The abbreviations of the parameter names are outlined in Table 2)

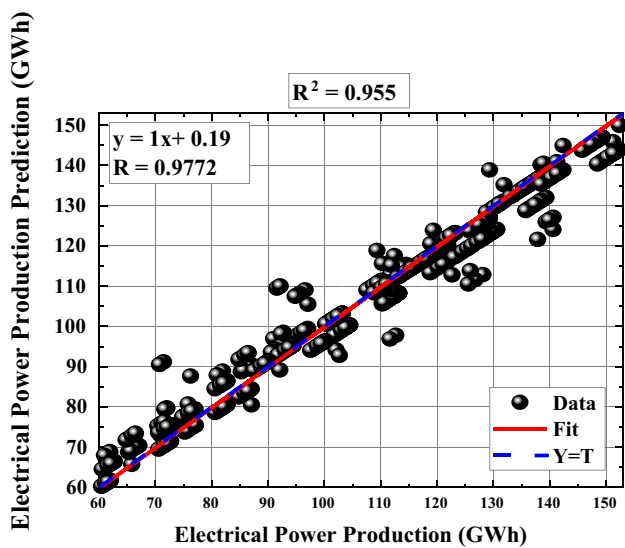


Fig. 5 Power Production (GWh) vs Power Production prediction (GWh) graph for the ELM model (inset: red line indicates the least square regression ( $y = mx + c$ ), where  $y$  is the prediction and  $x$  is the real data, additionally  $r$  is the coefficient of the correlation)

follow their lead [22]. The distance separating each wolf from the prey Eq. (6):

$$\vec{D} = \left| \vec{C} \cdot \vec{X}_p(t) - \vec{X}(t) \right| \tag{6}$$

$$\vec{C} = 2\vec{a} \cdot \vec{r}_2 \tag{7}$$

where  $\vec{X}$  and  $\vec{X}_p$  denote the position vectors of the prey and a wolf, respectively, and  $t$  represents the current iteration. The coefficient vector  $\vec{C}$  is computed following the Eq. (7), and  $\vec{r}_2$  is a random vector with values between 0 and 1 [25].

$$\vec{X}(t+1) = \vec{X}_p(t) - \vec{A} \cdot \vec{D} \tag{8}$$

$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a} \tag{9}$$

Within the optimization process,  $\vec{A}$  represents a coefficient vector, and its elements, denoted as  $\vec{a}$  gradually decrease linearly from 2 to 0. Additionally, the vector  $\vec{r}_1$  is composed of random values falling within the interval [0, 1] and they are given in Eqs. (8) and (9). In the GWO algorithm, the alpha, beta and delta wolves act as leaders, guiding the omega wolves to move toward the optimal position [25]. To calculate the new position of a wolf mathematically, a specific formula is used.

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \tag{10}$$

At a given iteration  $t$  in the GWO algorithm, the positions of the three best wolves in the swarm are identified and denoted as  $\vec{X}_1$ ,  $\vec{X}_2$  and  $\vec{X}_3$ . These wolves are used in a specific formula to determine the new position of each wolf in the pack. The movement pattern of wolves when  $|A|$  is less than 1 or greater than 1 is visually illustrated in Fig. 7.

The Genetic Algorithm considers the life adventure of living things. The good generations keep their lives, but the bad generations cannot save their lives and perish. The genetic algorithm acts within the framework of these principles. It is used in solving constrained problems, where there is no clear solution where the modeling process cannot be done mathematically [26]. In recent times, certain researchers have focused on solving the feature selection problem by employing genetic algorithms in a wrapper model approach. This method involves using genetic algorithms to search through all potential subsets of features in order to determine a set that can optimize the

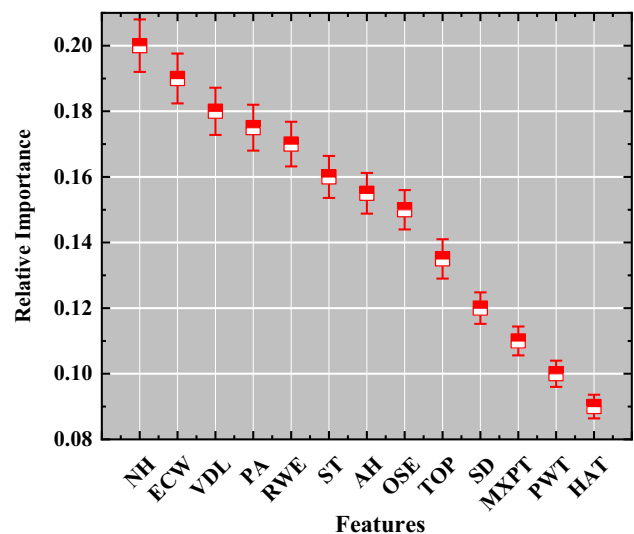


Fig. 6 Importance of the features for ELM model prediction

**Table 4** The testing performance metrics of the used ML algorithms

	SVR		RFR		ANN		ELM		WKNNR	
	PCA enable	PCA disable	PCA enable	PCA disable	PCA enable	PCA disable	PCA enable	PCA disable	PCA enable	PCA disable
<i>r</i>	0.965	0.973	0.973	0.968	0.966	0.969	<b>0.977</b>	0.957	0.966	0.935
<i>R</i> <sup>2</sup>	0.931	0.947	0.947	0.937	0.933	0.939	<b>0.955</b>	0.916	0.933	0.874
RMSE	1.427	0.976	0.988	1.154	1.286	1.025	<b>0.56</b>	1.598	1.325	2.328
MAE	1.019	0.697	0.706	0.824	0.919	0.732	<b>0.400</b>	1.141	0.946	1.663
RRMSE	30.00	37.60	3850	28.28	47.14	25.14	<b>23.02</b>	25.19	44.60	47.27
RMAE	34.80	43.62	44.660	32.81	54.68	29.16	<b>26.70</b>	29.22	51.74	54.83
Fit time (s)	3.19	18.09	0.81	0.77	0.12	1.162	2.67	18.93	11.81	0.84
Mean fit time (s)	361.77	153.28	0.79	0.84	1.53	1.04	15.50	18.65	0.17	0.12
Test time	0.008	0.005	0.020	0.154	0.005	0.005	0.030	0.024	0.029	0.044

accuracy of a given learning algorithm for prediction [22]. By exploring the entire feature space, the wrapper model can identify the most informative and relevant features, leading to enhanced predictive performance. The use of genetic algorithms in feature selection highlights the importance of advanced techniques in data analysis and the potential for significant improvements in predictive modeling [22, 27].

In the study, the GGWO method suggested by Kihel and Chouraqui is used for feature selection [22]. This approach extends to metaheuristics combination, where various bio-inspired hybrid methods are utilized to solve feature selection problems. One such method is the GGWO, which combines the capabilities of GA and GWO algorithms to achieve optimal results. In the GGWO approach, the population of grey wolves is randomly initialized with either a bit 1 or 0, and each wolf’s fitness is evaluated using K-NN. The best, second-best and third-best solutions are designated as alpha, beta and delta, respectively. The position of the wolf is then updated by applying the crossover between  $X_1$ ,  $X_2$  and  $X_3$ , resulting in improved feature selection performance. This hybridization approach demonstrates the potential for combining multiple techniques to enhance the efficiency and effectiveness of feature selection algorithms.

$$X(t + 1) = \text{Crossover}(X_1, X_2, X_3) \tag{11}$$

After initializing the population of grey wolves, their fitness is evaluated, and the positions of *alpha*, *beta* and *delta* are updated. The algorithm iterates until a stopping criterion is reached, which is typically the maximum number of iterations. Once the iterations are complete, the optimal feature subset is selected as the alpha solution. This process involves evaluating the fitness of each wolf and updating the positions of the alpha, beta and delta solutions in each iteration. By repeating this process, the

**Table 5** Selected features after using GGWO

#	Feature	Notation
1	Highest air temperature	HAT
2	Volume difference in the Lake (+/-)	VDL
3	Type of precipitation	TOP
4	Precipitation amount	PA
5	Open surface evaporation amount	OSE
6	Sunshine duration	SD
7	Temperature of the pool water	PWT
8	Net head	NH
9	Snow thickness	ST

algorithm can identify the most effective feature subset for the given problem. The use of the alpha solution as the optimal feature subset highlights the importance of selecting the best performing solution to achieve superior feature selection results. GGWO reduces the feature matrix from the closely related variable in numbers 1 to 9 in this section, and the features selected after using GGWO are given in Table 5. Table 6 shows the parameters which are selected by grid search methods.

*Convolutional Neural Networks (CNN):* A CNN is a type of model that is used for image recognition, and is characterized by its ability to detect complex patterns in video frame images through nonlinearity [28]. By adjusting various hyperparameters, it becomes possible to represent complex patterns in a high-dimensional space. CNNs consist of several key components, including convolutional layers, pooling layers, a dense network and a softmax layer which generates the output [29]. Each of these components plays an important role in enabling the CNN to effectively

**Table 6** The hyperparameters of deep learning models

Model	Hyperparameters	Hyperparameter selection	Power production
CNN	Filters	16	16
	Kernel size	2	2
	Activation function	log, tanh	tanh
	Hidden layer neuron number		16
	Hidden layer activation		tanh
	Output layer neuron number		1
RNN	Output layer activation function		linear
	Loss function		MSE
	Optimizers		Adam
	Learning rate		0.001
	Maximum iteration number		2000

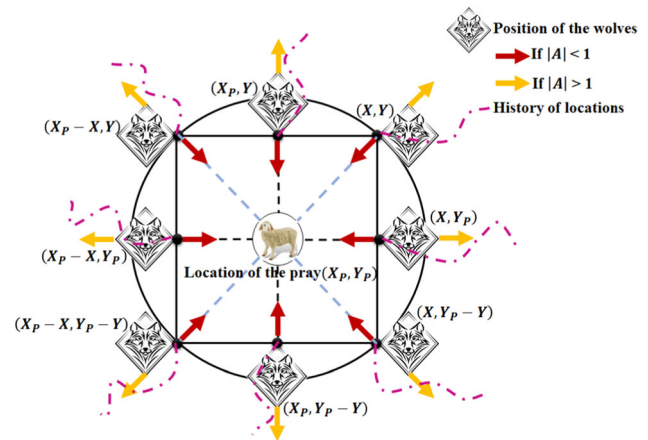
**Table 7** CNN parameters

Layer type	Output shape	Parameter
Conv1D_1	(None,111, 16)	1568
MaxPooling1D_1	(None,55, 16)	0
Conv1D_2	(None,55, 32)	17,073
MaxPooling1D_2	(None,26, 32)	0
Conv1D_3	(None,26, 64)	46,254
MaxPooling1D_3	(None,13, 64)	0
Conv1D_4	(None,13, 128)	114,675
MaxPooling1D_4	(None,7, 128)	0
Flatten	(None,896)	0
Dropout	(None,896)	0

Total parameters: 254,182

recognize patterns within images. The proposed CNN parameters are detailed in Table 7. Input data in the form of time series of 2 dimensions of 16 cores, then maximum pooling is made. In the second, 32 core and maximum pools are made of 3-dimensional output tensors, the second evolution layer. Then the third evolution layer was 4 dimensions of 64 cores and maximum pool layer, then 128 core and maximum pool in the last evolution 5 dimensions. All maximum pooling layers received a maximum value with 2 dimensions. The data from CNN are applied to RNN in the form of a one-dimensional time series.

**Recurrent Neural Network (RNN):** A Recurrent Neural Network (RNN) is a type of neural network that processes data sequentially, relying on its ability to retain a memory of the preceding sequence in order to learn effectively [30]. The name “recurrent” stems from the fact that the output from each time step is used as input for the subsequent step, building upon the output of the previous step by remembering it. By doing so, RNNs can capture long-term dependencies that exist within the training data.



**Fig. 7** The movement vectors' behavior is influenced by the value of |A|. [25]

**Long Short-Term Memory (LSTM):** LSTM, which is a variation of RNN, is utilized for generating captions from a range of video frames and images [28]. This is because LSTM is a highly effective sequence translation model that consists of numerous memory blocks known as cells. These cells are responsible for manipulating and retaining sequences using various categories of gates. One of the key benefits of using LSTM is that it helps overcome the vanishing gradient problem that can hinder the training of RNN models. As a result, LSTM has proven to be a highly useful tool for generating captions from visual data.

Figure 8 depicts the hybrid CNN/RNN model approach that has been put forward. This approach consists of four main components, as outlined in the diagram. Firstly, the feature vectors are generated by extracting features from video/images using a CNN and converting them into input dimensions. Secondly, a model is generated and used to provide a sequence through an RNN. Thirdly, semantic captioning is performed on the generated sequences. The evolution layer, LSTM recurrent layer and dense layer are the three types that make up the core layer. In order to

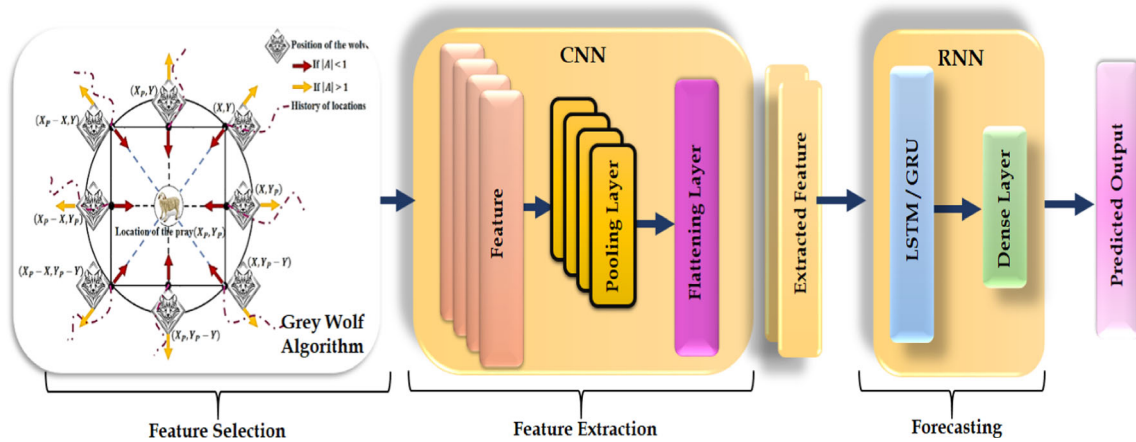


Fig. 8 Structure of the hybrid CNN/RNN model

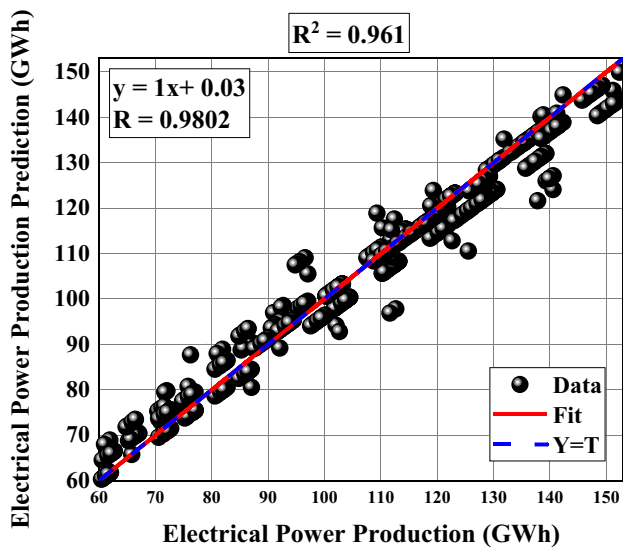


Fig. 9 Power Production (GWh) vs Power Production prediction (GWh) graph for the CNN/RNN hybrid model (inset: red line indicates the least square regression ( $y = mx + c$ ), where  $y$  is the prediction and  $x$  is the real data, additionally  $r$  is the coefficient of the correlation)

extract properties, the input data sequence is first used as input into two subsequent evolution layers. The LSTM layer is then introduced to the resulting attribute map, and the three active LSTM gates are used to thoroughly learn about complex additions. The gate of forgetting

specifically removes irrelevant or unnecessary information from earlier cell conditions. The entrance gate stores the useful new information during the entrance.

Additionally, signals from the cell status are filtered before being passed through the output passage and sent to the following state. The processed temporal data are used as input layers for non-linear transformation. The results of the predictions are then generated after the information obtained is reflected in the output space. Finally, the proposed model is evaluated using a variety of metrics, while also taking into account the data distribution and cross-validation. Figure 9 demonstrates Power Production (GWh) vs Power Production prediction (GWh) graph for the CNN/RNN hybrid model.

According to the MAE, MSE, RMSE, MAPE and  $R^2$  metrics of proposed deep learning model tested for prediction of Hirfanlı hydroelectric power production, the most successful result is obtained with proposed deep learning model when comparing as Table 8. When proposed deep learning model is used in the prediction process and the  $r$ ,  $R^2$  and MAE values are obtained as 0.9802, 0.961 and 0.314, respectively, and they are the values with the best results. Below are some examples of the literature studies. These studies in Table 9 are compared to the evaluation of the parameters compared to the models that we proposed better results than other methods.



**Table 8** The testing performance metrics of the used ML and hybrid deep learning techniques

	SVR		RFR		ANN		ELM		WKNNR		CNN/ RNN
	PCA enable	PCA disable	PCA enable	PCA disable	PCA enable	PCA disable	PCA enable	PCA disable	PCA enable	PCA disable	
r	0.965	0.973	0.973	0.968	0.966	0.969	<b>0.977</b>	0.957	0.966	0.935	<b>0.9802</b>
R <sup>2</sup>	0.931	0.947	0.947	0.937	0.933	0.939	<b>0.955</b>	0.916	0.933	0.874	0.961
RMSE	1.427	0.976	0.988	1.154	1.286	1.025	<b>0.56</b>	1.598	1.325	2.328	0.42
MAE	1.019	0.697	0.706	0.824	0.919	0.732	<b>0.400</b>	1.141	0.946	1.663	0.314
RRMSE	30.00	37.60	3850	28.28	47.14	25.14	<b>23.02</b>	25.19	44.60	47.27	21.51
RMAE	34.80	43.62	44.660	32.81	54.68	29.16	<b>26.70</b>	29.22	51.74	54.83	22.2
Fit time (s)	3.19	18.09	0.81	0.77	0.12	1.162	2.67	18.93	11.81	0.84	0.11
Mean fit time (s)	361.77	153.28	0.79	0.84	1.53	1.04	15.50	18.65	0.17	0.12	0.10
Testing time	0.008	0.005	0.020	0.154	0.005	0.005	0.030	0.024	0.029	0.044	0.004

**Table 9** The testing performance metrics of the used ML and hybrid deep learning techniques

REF	Algorithm/method	Performance of the models		
[8]	SVR	R = 0.6886	RMSE = 1.0982	MAE = 0.9028
	ELM	R = 0.6983	RMSE = 1.2165	MAE = 0.9519
	REG	R = 0.3803	RMSE = 4.234	MAE = 2.9688
	MLP	R = 0.9735	RMSE = 0.2593	MAE = 0.2128
[31]	CNN-SVR	0.968 (PP)	RMSE => 0.378	
[32]	ANN (BR)	R = 0.73		
	MSE = >1.43 × 10 <sup>4</sup>			
[33]	ANN (SCG)	R = 0.76	MSE = 7.42 × 10 <sup>5</sup>	
	ANN	R = 0.89		
[34]	ANN (BR)	R = >0.94	MAPE = 12.41%	MAE = 450
	ANN (LM)	R = >0.91	MAPE = 17>.05%	MAE = 593
[35]	Stepwise multiple regression	R <sup>2</sup> = 0.753		
[36]	Deep Feed Forward Neural Network	R = 0.507		
[37]	SVR	R = 0.87	MAPE = 9.7%	
	GPR	R = 0.92	MAPE = 4.5%	
	MLR	R = 0.60	MAPE = 46.1%	
	PR	R = 0.67	MAPE = 35.7%	
[38]	Random forest	R = 0.90	NMAE = 0.17	NRMSE = 0.2
[39]	ANN	R = 0.77		
[40]	BP-ANN	0.87	RMSE = 17.06	MAE = 10.23
	RBF-ANN	0.91	13.60	7.92
	LSTM	0.92	12.14	7.01
	ELM	0.91	12.75	7.59
	ABC-ELM	0.95	9.99	5.75
Our study	ELM	R = 0.977	RMSE = 0.56	MAE = 0.4
	CNN-LSTM	R = 0.9802	RMSE = 0.42	MAE = 0.314

## 4 Discussion

The objective of this study is to develop a predictive model for hydropower production using various machine learning and deep learning techniques. The research also delves into methodologies for constructing predictive models, underscoring the significance of feature selection approaches. Feature selection is crucial as it reduces the number of features, thereby minimizing learning time and improving prediction accuracy. The study utilizes meteorological data from the years 2007 to 2021, adjusting and normalizing outliers for estimating hydroelectric power plant production. The dataset is then split into 80% for training and 20% for testing. Model performance is assessed using metrics such as MAE, MSE, RMSE, MAPE and  $R^2$ , with the PCA method employed in the analysis of algorithm results. According to metrics including MAE, MSE, RMSE, MAPE and  $R^2$ , the ELM algorithm produced the best results among five different machine learning techniques for estimating hydroelectric capacity. Additionally, a novel technique is introduced, combining the Genetic Grey Wolf Optimizer (GGWO) with a Convolutional Neural Network/Recurrent Neural Network (CNN/RNN) model using Long Short-Term Memory (LSTM) for regression analysis. This approach involves utilizing GGWO for feature selection to reduce dataset dimensionality, followed by applying the CNN/RNN-LSTM model for feature extraction and prediction. The authors demonstrate the effectiveness of their proposed technique through experiments on real-world datasets, highlighting superior performance compared to other state-of-the-art techniques. Table 8 summarizes the results across all metrics for both machine learning algorithms and deep learning models. Future research avenues for this technique include exploring alternative optimization algorithms for feature selection besides GGWO. Additionally, investigating the impact of dataset size on the technique's performance and exploring different architectures for the CNN/RNN-LSTM model could provide valuable insights. Moreover, extending the technique to incorporate other data types, such as time series or image data, and applying it to diverse domains like finance or healthcare would contribute to evaluating its performance in various applications.

## 5 Conclusion

This study focuses on predicting power production at the Hirfanlı dam using machine learning (ML) and deep learning (DL) techniques. Two distinct methods were

employed: Method I utilized various ML algorithms with feature selection through the PCA method, while Method II introduced a CNN/RNN-LSTM hybrid model with feature selection based on the Grey Wolf Optimizer (GGWO) and Genetic Algorithm (GA). The study incorporated the ten-fold cross-validation method and employed grid search for hyperparameter adjustments. In Method I, five ML models are utilized, with the Extreme Learning Machine (ELM) algorithm emerging as the most successful. It achieved a correlation coefficient ( $r$ ) of 0.977 and a coefficient of determination ( $R^2$ ) of 0.955 with the best feature subset. Additionally, ELM demonstrated a mean absolute error (MAE) of 0.4, indicating high precision in power production prediction. Method II introduced the CNN/RNN-LSTM hybrid model, surpassing all ML models from Method I. This model demonstrated exceptional prediction accuracy with a correlation coefficient ( $r$ ) of 0.9802 and a coefficient of determination ( $R^2$ ) of 0.961. Moreover, the CNN/RNN-LSTM model achieved a MAE of 0.314, highlighting its robustness and reliability in power production forecasting. In conclusion, this research provides valuable insights into renewable energy prediction, particularly in hydroelectric power production. ML and DL techniques, coupled with advanced feature selection methods, offer a robust framework for accurate and efficient power production forecasting. Future work should address factors such as data availability, model generalization across different dam systems, and the potential impact of external factors on power production. Additionally, exploring the integration of real-time data and considering the dynamic nature of environmental conditions could enhance the model's predictive capabilities. These considerations are crucial for refining the models and ensuring their applicability in diverse operational scenarios. The outcomes of this study lay the foundation for future research in renewable energy prediction, inspiring the development of more sophisticated models for other energy sources and addressing the identified limitations.

**Authors' contributions** M.E.S provided software and was involved in conceptualization, methodology, validation, formal analysis, investigation, data curation, writing—original draft preparation, supervision, visualization, writing—review and editing and investigation. M.O.K contributed to validation, writing—original draft preparation, supervision, visualization, writing—review and editing and formal analysis.

**Funding** Open access funding provided by the Scientific and Technological Research Council of Türkiye (TÜBİTAK). The authors did not receive support from any organization for the submitted work.

**Availability of data and materials** Data in the work are available and collected in the literature which is referenced in the manuscript.

## Declarations

**Ethical approval** This article does not contain any studies with human participants or animals performed by any of the authors.

**Conflict of interest** All Authors declare that they have no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Hanoon MS, Ahmed AN, Razzaq A, Oudah AY, Alkhayyat A, Huang YF, El-Shafie A (2023) Prediction of hydropower generation via machine learning algorithms at three Gorges Dam, China. *Ain Shams Eng J* 14(4):101919
- Lai JP, Chang YM, Chen CH, Pai PF (2020) A survey of machine learning models in renewable energy predictions. *Appl Sci* 10(17):5975
- Mehdinejadi B, Fathi P, Khodaverdiloo H (2022) An inverse model-based Bees algorithm for estimating ratio of hydraulic conductivity to drainable porosity. *J Hydrol* 608:127673
- Zhou H, Tang G, Li N, Wang F, Wang Y, Jian D (2011) Evaluation of precipitation forecasts from NOAA global forecast system in hydropower operation. *J Hydroinf* 13(1):81–95
- Rheinheimer DE, Bales RC, Oroza CA, Lund JR, Viers JH (2016) Valuing year-to-go hydrologic forecast improvements for a peaking hydropower system in the Sierra Nevada. *Water Resour Res* 52(5):3815–3828
- Peng Y, Xu W, Liu B (2017) Considering precipitation forecasts for real-time decision-making in hydropower operations. *Int J Water Resour Dev* 33(6):987–1002
- 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* 12(2):289
- Condemi C, Casillas-Perez D, Mastroeni L, Jiménez-Fernández S, Salcedo-Sanz S (2021) Hydro-power production capacity prediction based on machine learning regression techniques. *Knowl Based Syst* 222:107012
- Al Rayess HM, Keskin AÜ (2021) Forecasting the hydroelectric power generation of GCMs using machine learning techniques and deep learning (Almus Dam, Turkey). *G eofizika* 38(1):1–14
- Li, G., Sun, Y., He, Y., Li, X., Tu Q (2014) Short-term power generation energy forecasting model for small hydropower stations using GA-SVM. *Math Prob Eng* 2014:1–9
- Larose DT, Larose CD (2014) *Discovering knowledge in data: an introduction to data mining*, vol 4. Wiley, New York
- Özkan Y (2013) *Veri Madenciliği Yöntemleri*. 2. bs. İstanbul: Papatya Yayıncılık.
- Cihan P, Kalipsiz O, Gökçe E (2017) Hayvan Hastalığı Teşhisinde Normalizasyon Tekniklerinin Yapay Sinir Ağı Ve Özellik Seçim Performansına Etkisi. *Electron Turk Stud* 12(11):59–70
- Abdi H, Williams LJ (2010) Principal component analysis. *Wiley Interdiscip Rev Comput Stat* 2(4):433–459
- Bonaccorso G (2017) *Machine learning algorithms*. Packt Publishing Ltd, Birmingham
- Sahin, M. E. (2023). Image processing and machine learning-based bone fracture detection and classification using X-ray images. *Int J Imag Syst Technol* 33(3):853–865
- Sahin ME (2023) Real-time driver drowsiness detection and classification on embedded systems using machine learning algorithms. *Traitement du Signal* 40(3):847
- Bergstra J, Bengio Y (2012) Random search for hyper-parameter optimization. *J Mach Learn Res* 13(10):281–305
- Klein A, Falkner S, Bartels S, Hennig P, Hutter F (2017) Fast bayesian optimization of machine learning hyperparameters on large datasets. In: *Artificial intelligence and statistics*. PMLR, pp 528–536
- Ulutas H, Sahin ME, Karakus MO (2023) Application of a novel deep learning technique using CT images for COVID-19 diagnosis on embedded systems. *Alex Eng J* 74:345–358
- Ortatas, F. N., Ozkaya, U., Sahin, M. E., & Ulutas, H. (2024). Sugar beet farming goes high-tech: a method for automated weed detection using machine learning and deep learning in precision agriculture. *Neural Comput Appl* 36(9):4603–4622
- Kihel BK, Chouraqui S (2020) A novel genetic grey wolf optimizer for global optimization and feature selection. In: *2020 second international conference on embedded and distributed systems (EDiS)*. IEEE, pp 82–86
- Mirjalili SMSM, Mirjalili SM, Lewis A (2014) Grey Wolf optimizer. *Adv Eng Softw* 69:46–61
- Koc I, Baykan OK, Babaoglu I (2018) Gri kurt optimizasyon algoritmasına dayanan çok seviyeli imge eşik seçimi. *Politeknik Dergisi* 21(4):841–847
- Chantar H, Mafarja M, Alsawalqah H, Heidari AA, Aljarah I, Faris H (2020) Feature selection using binary grey wolf optimizer with elite-based crossover for Arabic text classification. *Neural Comput Appl* 32:12201–12220
- Özger Y, Akpinar M, Musayev Z, Yaz M (2019) Electrical load forecasting using genetic algorithm based holt-winters exponential smoothing method. *Sakarya Univ J Comput Inf Sci* 3(2):108–123
- Al-Ani A (2007) Ant colony optimization for feature subset selection world academy of science, engineering and technology. *Int J Comput Electr Autom Control Inf Eng* 1(4):999–1002
- Khamparia A, Pandey B, Tiwari S, Gupta D, Khanna A, Rodrigues JJ (2020) An integrated hybrid CNN–RNN model for visual description and generation of captions. *Circuits Syst Signal Process* 39:776–788
- Sahin ME (2022) Deep learning-based approach for detecting COVID-19 in chest X-rays. *Biomed Signal Process Control* 78:103977
- Nasir JA, Khan OS, Varlamis I (2021) Fake news detection: a hybrid CNN-RNN based deep learning approach. *Int J Inf Manag Data Insights* 1(1):100007
- Özbay Karakuş, M (2023). Impact of climatic factors on the prediction of hydroelectric power generation: a deep CNN-SVR approach. *Geocarto Int* 38(1):2253203.

32. Perera A, Rathnayake U (2020) Relationships between hydropower generation and rainfall-gauged and ungauged catchments from Sri Lanka. *Math Probl Eng* 2020:1–8
33. Abdulkadir TS, Salami AW, Anwar AR, Kareem AG (2013) Modelling of hydropower reservoir variables for energy generation: neural network approach. *Ethiop J Environ Stud Manag* 6(3):310–316
34. Lopes MNG, da Rocha BRP, Vieira AC, de Sá JAS, Rolim PAM, da Silva AG (2019) Artificial neural networks approaches for predicting the potential for hydropower generation: a case study for Amazon region. *J Intell Fuzzy Syst* 36(6):5757–5772
35. Boadi SA, Owusu K (2019) Impact of climate change and variability on hydropower in Ghana. *Afr Geogr Rev* 38(1):19–31
36. Drakaki KK, Sakki GK, Tsoukalas I, Kossieris P, Efstratiadis A (2022) Day-ahead energy production in small hydropower plants: uncertainty-aware forecasts through effective coupling of knowledge and data. *Adv Geosci* 56:155–162
37. Ekanayake P, Wickramasinghe L, Jayasinghe JJW, Rathnayake U (2021) Regression-based prediction of power generation at samanalawewa hydropower plant in Sri Lanka using machine learning. *Math Probl Eng* 2021:1–12
38. Sessa V, Assoumou E, Bossy M, Simões SG (2021) Analyzing the applicability of random forest-based models for the forecast of run-of-river hydropower generation. *Clean Technologies* 3(4):858–880
39. Jung J, Han H, Kim K, Kim HS (2021) Machine learning-based small hydropower potential prediction under climate change. *Energies* 14(12):3643
40. Yildiz C, Açikgöz H (2021) Forecasting diversion type hydropower plant generations using an artificial bee colony based extreme learning machine method. *Energy Sour Part B* 16(2):216–234

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