Hydropower Plant Available Energy Forecasting Using Artificial Neural Network and Particle Swarm Optimization

Suriya Kaewarsa * and Vanhkham Kongpaseuth

Posted Date: 25 July 2024

doi: 10.20944/preprints202407.2024.v1

Keywords: energy prediction; reservoir inflow forecasting; deep learning; artificial neural network; particle swarm optimization; PSO-ANN

 doi:10.20944/preprints202407.2024.v1

Article

Hydropower Plant Available Energy Forecasting Using Artificial Neural Network and Particle Swarm Optimization

Suriya Kaewarsa * and Vanhkham Kongpaseuth

Department of Electrical Engineering, Faculty of Industry and Technology, Rajamangala University of Technology Isan Sakon Nakhon Campus, Thailand; vanhkham.k@gmail.com

* Correspondence: suriya.ke@rmuti.ac.th; Tel.: +66-4277-2391

Abstract: Accurate forecasting of available energy portion that corresponds to the reservoir inflow of the month(s) ahead provides important decision support for the hydropower plants in energy production planning for revenue maximization, as well as for the environmental impact prevention, and flood control at the upstream and downstream of a basin. Therefore, a reliable forecasting tool or model is deemed necessary and crucial. Considering the fluctuation and nonlinearity of data which significantly influence the forecasting results, this study develops an effective hybrid model by integrating Artificial Neural Network (ANN) and Particle Swarm Optimization (PSO) called "PSO-ANN" model based on the hydrological and meteorological data pre-processed by cross-correlation function (CCF), autocorrelation function (AFC), and normalization techniques for predicting the available energy portion corresponding to the reservoir inflow mentioned above for a case study hydropower plant in Laos namely, Theun-Hinboun hydropower plant (THHP). The model was evaluated by using correlation coefficient (r), relative error (RE), root mean square error (RMSE), and Taylor diagram plots in comparison with the popular single algorithm approaches such as ANN, and NARX models. Results demonstrated the superiority of the proposed PSO-ANN approach over the other two models.

Keywords: energy prediction; reservoir inflow forecasting; deep learning; artificial neural network; particle swarm optimization; PSO-ANN

1. Introduction

At present, renewable energy (RE) plays a crucial role in a clean energy supply in which one of the significant sources of RE is hydropower [1]. In a hydropower plant, accurate forecasting of available energy or reservoir inflow is necessary for ensuring optimal operation of the reservoir as sufficient water needs to be maintained in the reservoir in the dry season while unplanned or unreasonable spillage needs to be avoided in the wet season. The optimal reservoir operation is directly associated with generation planning that supports revenue maximization, as well as environmental impact prevention, and flood control at the upstream and downstream of a basin [2]. Due to real hydrological and meteorological data having the characteristics of fluctuation and nonlinearity resulting from the extensive influence of human activity and natural factors [3]; several studies have given more attention to the methodology of high precision forecasting of which two fundamental techniques for forecasting, physically based and data-driven models, are widely implemented [4]. The results provided by the physically based models are more explainable compared to those of data-driven models [5]. The existing physically based models including, but not limited to CalSim, WEAP, and HEC ResSim are regarded as powerful tools for forecasting as they are based on the principles of mass and hydrological conservation, and the laws of physics [6]; however, these physical models had some limitations resulting from their requirements for a huge effort in terms of hydrological and climate data, and time-consuming in the learning process [7]. Therefore, the data-driven approaches are the alternative solutions as these approaches are highly

 doi:10.20944/preprints202407.2024.v1

adaptable and able to be constantly optimized and updated hence, promising accurate forecasting [8].

In the past decades, The data-driven methods with a single algorithm have gained a lot of attention as a result of their simplicity such as the Adaptive Neuro-Fuzzy Inference System (ANFIS) model utilized for monthly dam inflow prediction in which the model with three input variables i.e. monthly rainfall, dam inflow and rainfall forecast of next month provided a higher accuracy than the other models [9]; Recurrent Neural Network (RNN) based on radar rainfall data and reservoir inflow data for multi-step-ahead reservoir inflow forecasting outperformed the Back Propagation Neural Network (FFBPNN) [10]; multi-layer perceptron (MLP) artificial neural network model for one-month-ahead inflow forecasting for the Ubonratana reservoir situated in Thailand revealed the better performance of the Type F (with inflow known and assumed to be the forecast) model over the other models [11]; Support Vector Machine (SVM) models considered different monthly time lags for predicting inflow to Zayandehroud dam reservoir in Iran displayed more accurate results than those of the ANN models [12]; Bayesian Networks (BN) model for annual and monthly inflow forecasting for the Zayandehrud dam reservoir in Iran depicted a more reliability over the other limited research conducted in the study area [13]; Nonlinear Auto-Regressive with exogenous input (NARX) model, and Nonlinear Auto-Regressive (NAR) model provided good alternative approaches for inflow forecasting such as the forecast for a case study reservoir in Iran (Sefidruod dam reservoir) which depicted a capability of the NAR model over the NARX [14]; hydropower generation prediction using SVM, ANN and ARIMA for three Gorges Dam situated in China demonstrated the superiority of the ANN model over the other models with the correlation coefficient (r) between 0.9011 and 0.8883 [15]. After reviewing the literature, some limitations of the single technique were detected such as overfitting of ANN, parameter uncertainty of SVM, local minimization drawback of PSO, cognitive uncertainties of Fuzzy logic, and lack of memory in genetic programming [16] especially when employed for the fluctuation and nonstationary data of different hydrological and climatic conditions thus influenced the forecasting accuracy; hence many researchers developed the hybrid models, instead of the single algorithm models, with the aim of improving the accuracy in the forecasting [8].

A hybrid model of SARIMA–GEP integrated a Seasonal Autoregressive Integrated Moving Average (SARIMA) and the gene expression programming (GEP) carried out for monthly reservoir inflow forecasting provided more accurate forecasting over the single algorithm models like ANN, GEP, and SARIMA models [17]; a DEL-LSTM model combined a Decomposition-Ensemble Learning (DEL) and Long Short Term Memory (LSTM) for predicting daily reservoir inflow outperformed the standalone ARIMA, DNN, MDFL, and LSTM [18]; a Hybrid Deep Learning Inflow Prediction-Rolling Window called "HDeepLIP-RW models" provided superiority over several single algorithm models namely, ANN, RNN, MLR, GRU, and LSTM [19]; a fusion-based model combined several individual artificial neural networks (ANNs) with the moderate-resolution imaging spectroradiometer (MODIS) data increased the accuracy in forecasting monthly reservoir inflow [20]. As an alternative, this study develops effective hybrid models by integrating the Artificial Neural Network (ANN) and Particle Swarm Optimization (PSO) based on the hydrological and meteorological data, properly pre-processed by cross-correlation function (CCF), autocorrelation function (AFC), and normalization techniques, to forecast the available energy portion corresponding to the reservoir inflow mentioned above.

Previously, some researchers adopted a hybrid of Particle Swarm Optimization (PSO) and Artificial Neural Network (ANN) based-models known as "PSO-ANN models" in the forecasting works such as an electrical load prediction in which the PSO algorithm was used for solving the problem related to day-ahead load shifting and the cost saving with demand response (DR) program in combination with the multilayer feed-forward model [21]; a water level prediction by employing the PSO as an optimizer to search for the optimal parameter values for the ANN training process [22]; an application of PSO in optimizing the weights and biases of the ANN to enhance accuracy and performance of the prediction model for gas metering system [23]; a preference prediction in Multi-Criteria Recommender System (a powerful online tools that help to overcome problems of information overload) by training the feed-forward neural networks with the particle swarm

 doi:10.20944/preprints202407.2024.v1

optimization (PSO) algorithm and then used the neural network as an aggregation function for predicting the preferences of the users [24]; a behaviour prediction and structural optimization of lightweight sandwich composite heliostats by using the PSO-ANN approach [25]; a crack prediction in pipeline using PSO-ANN whereas the PSO algorithm was employed to enhance ANN training parameters (biases and weights) by minimizing the difference between actual and desired outputs and then using these parameters to generate the network [26]. Based on the aforementioned, we found that the hybrid of ANN and PSO has been employed in several fields of forecasting effectively, except in the reservoir inflow or available energy prediction area, hence this study aims to take advantage of the PSO-ANN approach to forecast the available energy portion corresponding to the Nam Gouang reservoir inflow of the case study hydropower plant in Laos (THHP) which constitutes a novelty approach for the forecasting on this field. To reflect the real application in the commercial hydropower plants and/or the case study hydropower plant as illustrated in Figure 3, this study converted the historical monthly reservoir inflow into the energy unit (available energy in GWh) before the forecasting. By forecasting the available energy portion corresponding to reservoir inflow, the case study hydropower plant can significantly improve its decision support for energy production planning that maximizes revenue. This forecasting also aids in environmental impact prevention and flood control upstream and downstream of the basin. In essence, the hydropower plant will consider the forecasted available energy portion alongside the actual available energy portion, that corresponds to existing water in the reservoir, to determine the amount of energy it can declare for sale in each month ahead. This decision aligns with annual supply targets and ensures sufficient available energy throughout the year. Additionally, this approach helps prevent unreasonable or unplanned reservoir spillage during the wet season, which could lead to lost generation opportunities.

This study has two main goals. The first is to develop the proposed PSO-ANN models in the MATLAB environment. The second is to evaluate the prediction performance of the proposed models using the testing data cluster (not included in the training data cluster), and statistical indicators including correlation coefficient (r), relative error (RE), and root mean square error (RMSE) in couple with Taylor diagrams.

The remainder of this paper is divided into three main Sections. Section 2 introduces the methodology. Section 3 describes the results and discussion. Finally, Section 4 presents the conclusions.

2. Methodology

In this Section, we introduce the research framework of the proposed methodology, which includes the overview of the PSO-ANN concept, comprehensive descriptions of the case study hydropower plant, and data preparation as well as those of the Artificial Neural Network (ANN) and Particle Swarm Optimization (PSO).

2.1. Research Framework

This study was carried out to develop the hybrid models for the forecasting of the month-ahead available energy portion corresponding to the reservoir inflow by integrating the ANN and PSO based on the available historical hydrological and meteorological data of the Nam Guouang reservoir of the Theun-Hinboun hydropower plant, which is the case study hydropower plant situated in Laos, hereinafter referred to as "THHP". Another goal of this study is to evaluate the performance of the proposed models in comparison with the single algorithm models i.e., ANN, and NARX models. The flowchart of the proposed approach is illustrated in Figure 1. In the data preparation and preprocessing phase, the hydrological and meteorological data including historical monthly rainfall, and inflow data (available in the case study hydropower plant) influencing the prediction results were collected, then the cross-correlation function (CCF) and autocorrelation function (AFC) were employed to determine the number of input lags and combinations for the proposed models. For the achievement of better results and faster learning, the datasets were normalized into numbers between 0 and 1 before plugging into the proposed PSO-ANN models. With respect to the modelling process,

 doi:10.20944/preprints202407.2024.v1

four main steps were carried out including (1) optimizing the ANN weights and biases using the PSO algorithm thus obtaining the optimal network, (2) training the models by the "trainlm" function, and using a training cluster (3) testing the trained models by using a separated testing cluster (not included in the training cluster), and (4) evaluating the model by calculating the correlation coefficient (r), relative error (RE), root mean square error (RMSE) in couple with Taylor diagrams.

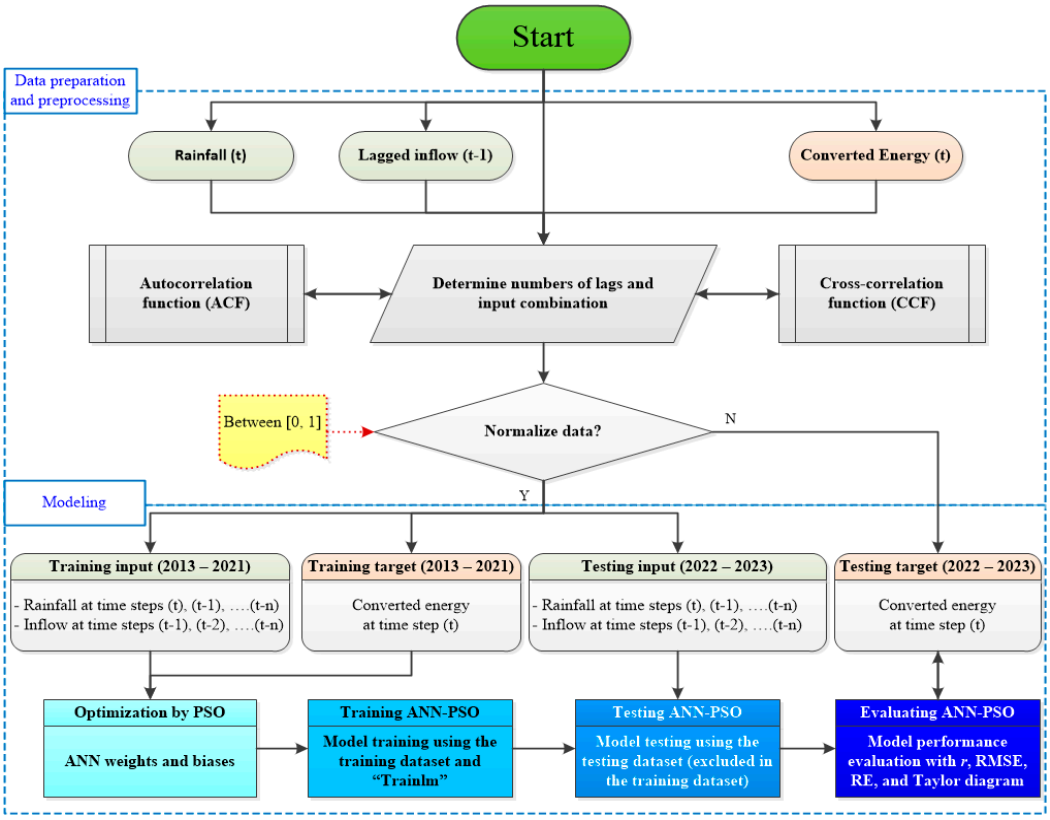

Figure 1. Research methodology flowchart.

2.2. Study Area

This study implemented the Energy Forecasting for the Theun-Hinboun Hydropower Plant (THHP) situated in the co-area of Khammouane and Bolikhamxay Provinces in Laos approximately 217 km to the east of Vientiane capital and owned by Theun Hinboun Power Company (THPC). Initially, a 220 MW run-of-river project, consisting of two spillway radial gates, a head pond with full supply level of 400 m asl, and generation capacity of 1,100 GWh per year, has been constructed of which 95% of electricity is purchased by the Electricity Generating Authority of Thailand (EGAT). The remaining 5% is for domestic supply to Electricité du Lao (EDL). After the operation for a decade, THPC has expanded the project to a total installed capacity of 500 MW by constructing a new reservoir on the Nam Gnouang River which is a tributary of the mainstream of the existing run-of-river project. The new Nam Gnouang Dam, hereinafter referred to as "NG", is constructed approximately 20 kilometers upstream of the existing Weir Dam in the form of a concrete gravity Dam with a structure of 480 meters wide and 65 meters high created the NG reservoir with storage capacity of up to 2,430 million cubic meters (MCM) at the full supply level of 455 masl. The NG Dam consists of five spillway gates operated only to discharge water during the wet season to avoid flood event(s).

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 July 2024 doi:10.20944/preprints202407.2024.v1

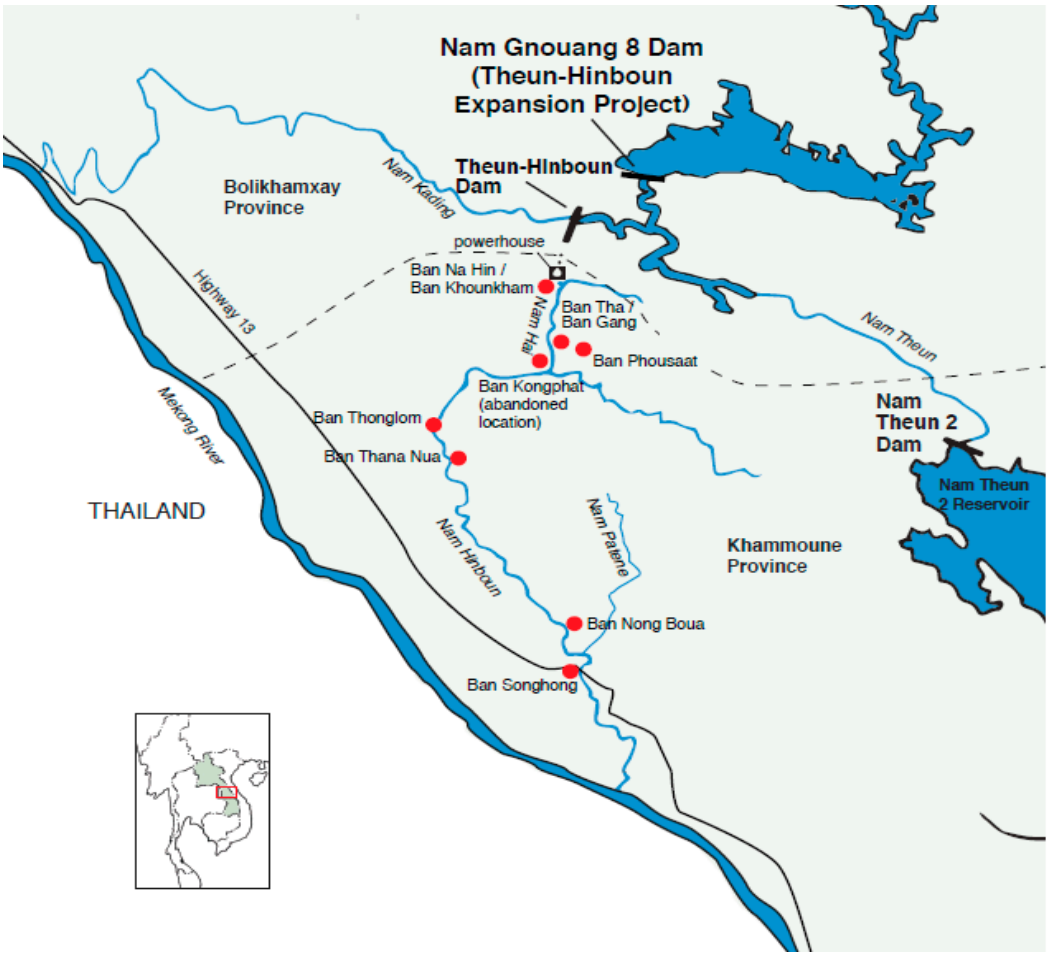

Figure 2. Map of Nam Gnouang reservoir of THHP.

During normal operation, the NG Dam releases water through the NG Powerhouse, which can generate power up to 60 MW for domestic supply to the EDL from the NG Powerhouse. After the NG reservoir, the water flows into the Theun River and the existing head pond above the original Weir Dam. The water finally flows through the new intake and tunnel system to the expanded Theun-Hinboun Powerhouse located approximately 240 meters below in the Na Hin town of Khounkham District. At this location, a new 220 MW Francis generator has been installed alongside the existing two (2) 110 MW generators. After the powerhouse, the water flows to a spillway and then a regulating pound before releasing into the Nam Hai River which is a tributary of the Hinboun River. The new NG reservoir enables THPC to export up to 440 MW to EGAT in which the power is transmitted to a designated delivery point at the Thakhek substation on the Lao PDR-Thailand border via a 230 kV double-circuit transmission line with a total length of approximately 86 kilometers while the remaining 60 MW are allocated for the domestic supply which purchased by the EDL.

Table 1. Basic design specifications of the case study hydropower plant.

Design Feature	Specification
***Original Weir Dam (run-of-river type)*:**	
Concrete Gravity Free Overflow Weir:	268 x 27 m (L x H)
Spillway radial gates	2
Head Pond Full Supply Level	400 m asl
Two concrete-lined headrace tunnels length	5,289 m and 5,496 m
Installed capacity	2 x 110 MW (+ 220 MW)
***NG Dam (storage type)*:**	

 doi:10.20944/preprints202407.2024.v1

Concrete Gravity Dam	480 x 65 m (L x H)
Reservoir Full Supply Level	455 m asl
Max Actual Storage Volume	2,430 MCM
Spillway radial gates	5
Installed capacity	2 x 30 MW

The power purchase agreement between the THPC and EGAT is in the form of the take-or-pay principle with a contract term of 25 years commencing from the commercial operation date (COD). To ensure optimal generation planning associated with revenue maximization as well as environmental impact prevention, and flood control at the upstream and downstream of the basin, the THPC needs to take into account two key energy portions before declaring the energy amount (availability) for selling including the actual available energy portion that corresponds to the existing water in the Nam Gnouang reservoir and head pound at the time of declaration, and the future energy (forecasted energy) that corresponds to the natural inflow (reservoir inflow) expected to flow into the reservoir and head pound in the days and/or months ahead as illustrated in Figure 3.

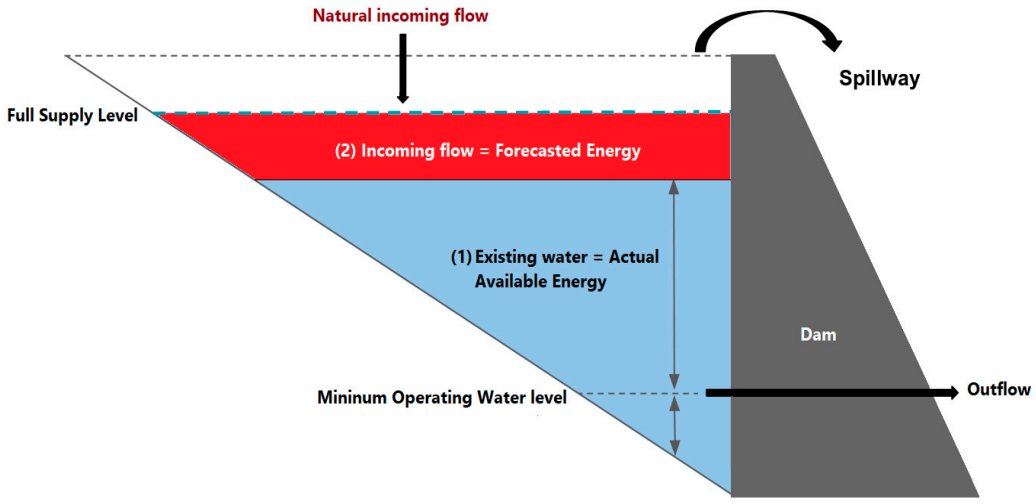

Figure 3. Hydropower plant energy schematic overview.

2.3. Data Preparation

Initially, the historical hydrological and meteorological data including rainfall, and reservoir inflow of the Nam Gnouang reservoir of the THHP were collected from the measurement stations, and some missing data were filled by using the backpropagation technique. The total duration of the data was 11 years (1 January 2013 to 31 December 2023). To reflect the real-world application in the commercial hydropower plants and/or the case study hydropower plant as illustrated in Figure 3, the historical reservoir inflow of the time step (t) has been converted into the energy unit (available energy in GWh) before applying to the models based on the THHP energy coefficient at 0.56117 kWh/m3. The collected data was then analyzed, normalized, and partitioned into four clusters including (i) training input cluster (rainfall and lagged inflows); (ii) training target cluster (converted available energy); (iii) testing input cluster (same parameters of the training input cluster); and (iv) testing target cluster (same parameter of the training target cluster) whereas the duration of the training input cluster and training target clusters were the same at 9 years (approximately 80%) ranging from 1 January 2013 to 31 December 2021, and those of the testing input cluster and testing target clusters at 2 years thereafter (approximately 20%) ranging from 1 January 2022 to 31 December 2023 in compliance with the partition ratio recommended by the previous study [4]. The training input and target clusters were collectively referred to as the "training dataset", and the testing input and target clusters were collectively referred to as the "testing dataset" in this study.

 doi:10.20944/preprints202407.2024.v1

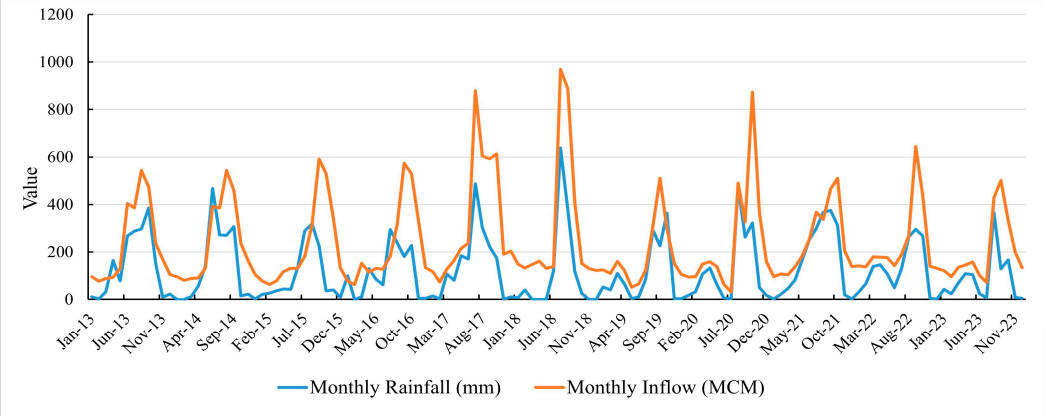

Figure 4. Original data of Nam Gnouang reservoir.

To determine the number of delays (input time lags) and input combinations for the proposed models, the correlation of every single parameter as well as between each input parameter, and the target parameter were analysed with the autocorrelation function (AFC), and cross-correlation function (CCF). In addition to the data analysis by ACF and CCF, all the data were also normalized into numbers between [0, 1] to ensure faster learning and achievement of better results [27] by using Equation (1):.

$$x'_k = \frac{(x_k - x_{\min})}{(x_{\max} - x_{\min})} \tag{1}$$

where x_k is the sampling data; $x_{\min}$ is the minimum value of the sampling data; and $x_{\max}$ is the maximum value of the sampling data.

2.4. Artificial Neural Networks (ANN)

Artificial neural network models were first introduced by McCulloch and Pitts in 1943 [28]. The ANNs are computer software or hardware models inspired by the structure and behavior of neurons in the human nervous system [29]. The ANN model is a part of the artificial intelligence modelling techniques, it can learn from the training data and find the data pattern to estimate the results [30]. Due to the fluctuation and nonlinearity behavior of the hydrological and meteorological time-series data, the ANN is widely used to develop forecasting models that produce results more accurately than statistical models. In general, the ANNs consist of three-layer architecture including the Input layer, Hidden layer, and Output layer, which is sufficient to solve a complex nonlinear problem like time-series forecasting. Figure 5 illustrates an ANN architecture with three layers.

Where $x_1, x_2, x_3, \ldots, x_n$ are the input signals; $w_{1,1}, w_{1,2}, w_{1,3}, \ldots, w_{n,n}$ are respective weights for the connection between the input layer and hidden layer of the network; $w_{1,O}, w_{2,O}, w_{3,O}, \ldots, w_{n,O}$ are the respective weights for the connection between the hidden layer and output layer of the network; $b_1, b_2, b_3, \ldots, b_n$ are the biases of the hidden nodes; b_O is the bias of the output node; $\varphi_1(\cdot), \varphi_2(\cdot), \varphi_3(\cdot), \ldots, \varphi_n(\cdot)$ are the activation functions of the hidden nodes; $\varphi_O(\cdot)$ is the activation function of the output node; and y_O is the output signal of the network. Mathematically, the ANN can be represented as follows:

$$u_1 = (w_{1,1}x_1 + w_{2,1}x_2 + w_{3,1}x_3 + \ldots + w_{n,1}x_n) \tag{2}$$

$$u_2 = (w_{1,2}x_1 + w_{2,2}x_2 + w_{3,2}x_3 + \ldots + w_{n,2}x_n) \tag{3}$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 July 2024 doi:10.20944/preprints202407.2024.v1

$$u_3 = (w_{1,3}x_1 + w_{2,3}x_2 + w_{3,3}x_3 + \ldots + w_{n,3}x_n) \quad (4)$$

$$u_n = (w_{1,n}x_1 + w_{2,n}x_2 + w_{3,n}x_3 + \ldots + w_{n,n}x_n) \quad (5)$$

$$h_1 = \varphi_1(u_1 + b_1) \quad (6)$$

$$h_2 = \varphi_2(u_2 + b_2) \quad (7)$$

$$h_3 = \varphi_3(u_3 + b_3) \quad (8)$$

$$h_4 = \varphi_4(u_4 + b_4) \quad (9)$$

$$h_n = \varphi_n(u_n + b_n) \quad (10)$$

$$u_O = (w_{1,O}h_1 + w_{2,O}h_2 + w_{3,O}h_3 + \ldots + w_{n,O}h_n) \quad (11)$$

$$y_O = \varphi_O(u_O + b_O) \quad (12)$$

where $u_1, u_2, u_3, \ldots, u_n$ are the linear combiners of the hidden nodes; $h_1, h_2, h_3, \ldots, h_n$ are the output signals of the hidden nodes; u_O is the linear combiner of the output node; and y_O is the output signal of the network. The optimal choice in selecting the activation functions can depend on the specific problem and/or dataset. However, this study selected the tangent sigmoid activation function based on its superior advantages in terms of balanced range that can be advantageous for PSO as particles don't get confined to a specific positive region, allowing for more balanced exploration during optimization, and its gradient consistency as it generally has smoother gradients compared to other activation functions like ReLU. This smoother gradient can be beneficial for PSO, as the particles rely on these gradients to update their positions and potentially lead to faster convergence during optimization.

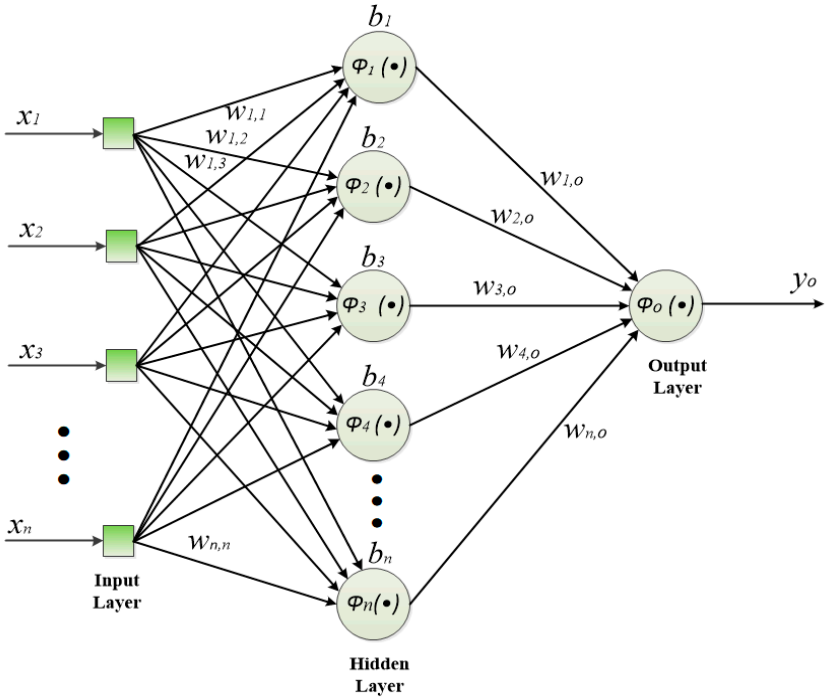

Figure 5. Feedforward Neural Network.

 doi:10.20944/preprints202407.2024.v1

2.5. Particle Swarm Optimization (PSO)

Particle Swarm Optimization, widely known as "PSO", was introduced by Kennedy and Eberhart in 1995, it is a computational swarm-based search method inspired by the flocking behavior of birds in searching for food. With its strong optimization ability and ease of implementation hence, the PSO has become a popular option for solving various optimization problems [31–32]. Like other evolutionary algorithms, the PSO algorithm begins with an initial random population of candidate solutions called "particles". These particles iteratively explore the search space, adjusting their positions based on their own flying experience and the best solution discovered by the entire swarm. This process is guided by a few key parameters to lead the swarm towards the optimal solution with minimal computational cost [33]. A swarm $P(t)$ of N particles herein initialized the $t=0$, a random position $x_i(t)$ was assigned to each particle $P_i(t)$, and the cost of each particle evaluated. During subsequent iterations, each particle performance was assessed at its current location to determine and update the personal best (P_{best}) and global best (G_{best}) experiences. Whereas the P_{best} refers to the best personal position that a particular particle has encountered, and G_{best} refers to the best position discovered by the entire swarm. The flight velocity and position of each particle were updated according to the following equations.

$$v_t(t) = wv_i(t-1) + \rho_1(P_{best} - x_i(t-1)) + \rho_2(G_{best} - x_i(t-1)) \quad (13)$$

$$x_i(t) = x_i(t-1) + v_i(t) \quad (14)$$

The random variables $\rho_1 = r_1C_1$ and $\rho_2 = r_2C_2$ whereas $r_1, r_2 \sim U$ [0, 1], C_1 and C_2 represent positive acceleration constants, and ω represents an inertia weight. The second and last portions of Equation (13) are referred to as the cognitive component and social component, respectively. Therefore, C_1 and C_2 herein are the cognitive and social acceleration constants, respectively.

While a larger PSO population explores a wider search space, potentially finding better solutions, however, it can be computationally expensive; adversely, a smaller population faster computation but might get stuck in local optima. Higher C_1 emphasizes individual learning, potentially leading to faster convergence towards individual bests but risking local optima in the meanwhile, higher C_2 emphasizes swarm learning, promoting exploration of the search space but potentially slowing down convergence thus adjusting C_1 and C_2 helps strike a balance between individual and swarm learning. With respect to the inertia weight (ω), the higher weight promotes exploration in the early stages, allowing particles to move further and explore a wider search space while the lower weight encourages exploitation in later stages, focusing particles closer to promising areas. Gradually reducing the weight over iterations helps with this transition. To ensure the PSO algorithm stability, the previous research [34] suggested $C_1 + C_2 \leq 4$. Considering the aforementioned, and based on the real forecasting results, this study finally configured C_1 and C_2 at 1.5 and 2, respectively. For the inertia weight ω, it was updated in each iteration according to Equation (14).

$$\omega = \omega_{damp} * \omega \quad (15)$$

where ω_{damp} is the inertia weight damping ratio finally set to 0.99 in this study.

2.6. Proposed PSO-ANN Hybrid Models

The ANNs are good for learning input-output relationships or patterns. However, they could face local minima and network paralysis problems due to rough weights assigned by learning algorithms such as backpropagation [35]. To overcome these drawbacks, we propose the PSO

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 July 2024 doi:10.20944/preprints202407.2024.v1

algorithm to optimize the weights and biases for the ANNs so that the results produced by the ANNs are more accurate as the prediction errors will be minimized. The descriptions of the ANN and PSO algorithms were clearly explained in Section 2.4 and Section 2.5 respectively. To enhance the prediction performance, this study also employed the autocorrelation function (ACF), cross-correlation function (CCF), and data normalization techniques during the data pre-processing phase as described in Section 2.3. For a better understanding, the proposed PSO-ANN approach can also be represented and described by the flowchart in Figure 6:

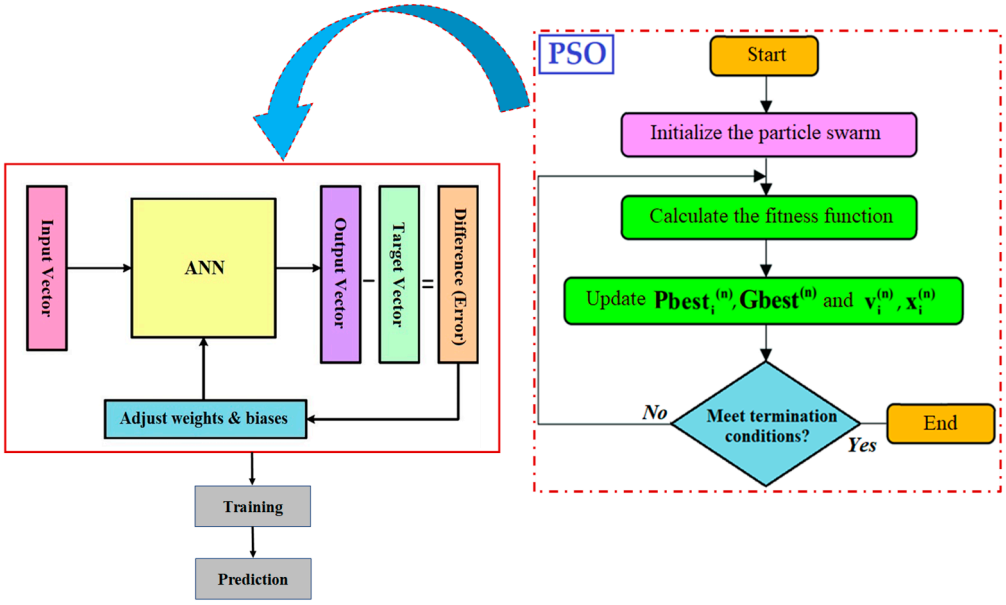

Figure 6. PSO-ANN approach.

In Figure 6, the output vector of the ANN network is compared with the target vector, and their difference called "error" is calculated in the form of normalized mean squared error (NMSE); If the average change in the global best (*Gbest*), corresponding to the NMSE, is greater than the options of the function tolerance (1e-15), the weights and biases of the network will be adjusted or optimized by using the PSO algorithm. The iteration continues until the average change is less than the options of the function tolerance. The proposed PSO-ANN models, for predicting the available energy portion corresponding to the reservoir inflow, can be summarized in the Table below:

Table 2. Summary of the various combination input designs.

Model Scenario	Input combinations	Output
SC1	R(t), I(t-1)	E(t)
SC2	R(t), R(t-1), I(t-1), I(t-2)	
SC3	R(t), R(t-1), I(t-1), I(t-2),I(t-3)	
SC4	R(t), R(t-1), R(t-2), I(t-1), I(t-2)	
SC5	R(t), R(t-1), R(t-2), I(t-1), I(t-2),I(t-3)	

* Notes:

R(t)	=	Rainfall of month (t) [mm].
R(t-1)	=	Rainfall of month (t-1) [mm]
R(t-2)	=	Rainfall of month (t-2) [mm]
I(t-1)	=	Inflow of month (t-1) [hm3]
I(t-2)	=	Inflow of month (t-2) [hm3]
I(t-3)	=	Inflow of month (t-3) [hm3]
E(t)	=	Energy of month (t) [GWh]

Based on the ACF and CCF analyses, the rainfall of the timestep (t) and the inflow of the timestep (t-1) have a good correlation with their own lagged values as well as with the target values i.e., with the converted available energy of the timestep (t) up to three timesteps hence, this study performed

 doi:10.20944/preprints202407.2024.v1

the simulation in several scenarios by alternating various input combinations, as summarized in Table 2 for seeking the best prediction model or scenario of the proposed approach to compare with the ANN and NARX approaches as mentioned above whereas the number of output variable for the models was determined to be equal to 1 corresponding to the forecasting objective, and the number of hidden nodes or neurons determined by using the trial-and-error method.

3. Prediction Error

To evaluate the performance of the proposed PSO-ANN prediction models, the three measurement or statistical indicators including root mean square error (RMSE), relative error (RE), and correlation coefficient (r) were calculated and analyzed using Equation (16)–(18) in couple with the Taylor diagrams.

$$r = \frac{\sum_{i=1}^{N}(E_{mi} - \overline{E}_m)\times(E_{ci} - \overline{E}_c)}{\sqrt{\left[\sum_{i=1}^{N}(E_{mi} - \overline{E}_m)^2 \times \sum_{i=1}^{N}(E_{ci} - \overline{E}_c)^2\right]}} \tag{16}$$

$$RMSE = \left(\frac{\sum_{i=1}^{N}(E_{mi} - E_{ci})^2}{N}\right)^{0.5} \tag{17}$$

$$RE = \frac{|E_{ci} - E_{mi}|}{E_{ci}}\times 100\,\% \tag{18}$$

where E_{mi} represents the target or observed available energy at the time i; $\overline{E}_m$ represents the average value of the target or observed available energy at the time i; E_{ci} represents the predicted available energy at the time i; $\overline{E}_{ci}$ represents the average value of predicted available energy at the time i; and N represents the number of data points.

The root mean square error (RMSE) and relative error (RE) were used to evaluate the performance of the models whilst the correlation coefficient (r) is used to assess the relations between the target and predicted values as well as to measure the strength in terms of the relationship between the two variables [15].

4. Result and Discussion

A Hybrid model has been developed in this study to predict the available energy portion corresponding to the reservoir incoming flow of Nam Gouang reservoir of the THHP based on the combination of ANN and PSO algorithms described in Section 2.6 together with the ACF, CCF, and data normalization techniques. The historical hydrological and meteorological data of the case study hydropower plant were initially collected ranging from 2013–2023 and then some missing data were filled with the backpropagation technique before preprocessing by proper analysis and normalization. The normalized data was finally partitioned into training and testing clusters at a ratio of approximately 80:20 as described in Section 2.3. In addition to the proposed model, the same datasets were also applied for the ANN and NARX models for comparison purposes.

4.1. Hyperparameter Determination

Input combinations for the feedforward neural network of the proposed approach were determined based on the correlation analysis results in which five scenarios of input combinations were selected and tested as summarized in Table 2. The number of hidden neurons for the neural network was determined by the trial-and-error method and finally selected 30 which is the number that provided the best prediction results. Lastly, the number of output variables was determined to be equal to 1 corresponding to the forecasting objective. With respect to the PSO algorithm, all the

 doi:10.20944/preprints202407.2024.v1

hyperparameters were determined by the trial-and-error method and finally obtained the C_1 and C_2 of 1.5 and 2, respectively, ω_{damp} (inertia weight damping ratio) of 0.99, and a population (swarm) size of 200.

4.2. Results of the Proposed Approach

Tables 3 and 4 revealed that the proposed PSO-ANN model of the third scenario (PA3) with the structure (5,30,1) has given the best prediction results with the statistical indicators r = 0.973, RMSE = 22.994 and RE = 1.038 % for the single year of 2022; and r = 0.966, RMSE = 24.846 and RE = 2.853 % for 2023 which were superior over those of the other scenarios i.e., the RMSE and RE of the proposed scenario, both single years, were less than those of other scenarios indicating more accuracy in the forecasting; on the other hand, the correlation coefficient (r) of the proposed scenario was closer to 1 than those of the other scenarios (both years) demonstrating a stronger relationship between predicted and the target values. Hence, the third scenario (PA3) was selected to be the best structure for the proposed approach and for later comparison with the other prediction approaches in this study.

Table 3. Results by statistical indicators of the PSO-ANN models 2022.

Model Scenarios	Model Input combinations	Model Output	Different Models	Model Structures	*r*	RMSE	RE
PA1	R(t), I(t-1)	E(t)	PSO-ANN	(2,30,1)	0.951	38.168	8.261
PA2	R(t), R(t-1), I(t-1), I(t-2)	E(t)	PSO-ANN	(4,30,1)	0.968	33.261	18.459
PA3	R(t), R(t-1), I(t-1), I(t-2),I(t-3)	E(t)	PSO-ANN	(5,30,1)	0.973	22.994	1.038
PA4	R(t), R(t-1), R(t-2), I(t-1), I(t-2)	E(t)	PSO-ANN	(5,30,1)	0.942	32.924	7.847
PA5	R(t), R(t-1), R(t-2), I(t-1), I(t-2),I(t-3)	E(t)	PSO-ANN	(6,30,1)	0.930	34.354	3.662

Table 4. Results by statistical indicators of the PSO-ANN models 2023.

Model Scenarios	Model Input combinations	Model Output	Different Models	Model Structures	*r*	RMSE	RE
PA1	R(t), I(t-1)	E(t)	PSO-ANN	(2,30,1)	0.905	44.925	15.059
PA2	R(t), R(t-1), I(t-1), I(t-2)	E(t)	PSO-ANN	(4,30,1)	0.965	30.668	5.832
PA3	R(t), R(t-1), I(t-1), I(t-2),I(t-3)	E(t)	PSO-ANN	(5,30,1)	0.966	24.846	2.853
PA4	R(t), R(t-1), R(t-2), I(t-1), I(t-2)	E(t)	PSO-ANN	(5,30,1)	0.930	36.757	7.928
PA5	R(t), R(t-1), R(t-2), I(t-1), I(t-2),I(t-3)	E(t)	PSO-ANN	(6,30,1)	0.956	27.934	3.001

4.3. Training Results

The proposed approach started by adjusting or optimizing the weights and biases of the feedforward neural networks whereas the particle swarm optimization (PSO) algorithm was employed based on the training dataset. The weights and biases were then updated a little bit more with the default function of the network namely, "Trainlm" (a network training function that updates weight and bias values according to Levenberg-Marquardt optimization) during the training process. Several models with different structures have been trained and tested, one best model was then selected based on the results of the statistical indicators, of which the proposed PSO-ANN model of the third scenario (PA3) with the structure (5,30,1) has given the best prediction results.

A cost function applied in the model is the Normalised Mean Square Error (NMSE). The cost function returns an output value, called the "Cost", which is a numerical value representing the deviation, or degree of error, between the model prediction and the target data; the lower the cost, the lower the deviation (error); hence, the optimal prediction model would have a cost close to 0. Figure 7 depicted the convergence of the best cost in PSO versus the updating iteration for the selected scenario (PA3) in which the best cost value started from 0.61479 in the first iteration and gradually declined to 0.10206 in iteration 842 regarded as the last iteration of the PSO running process.

 doi:10.20944/preprints202407.2024.v1

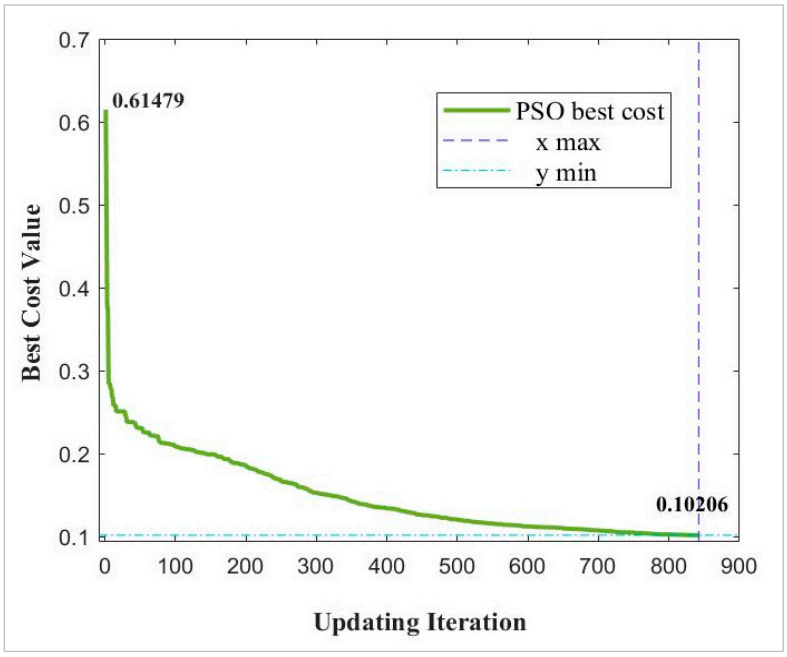

Figure 7. Best cost of the selected scenario (PA3).

Error histogram represents the histogram of the errors between target values and predicted values after training a feedforward neural network. As these error values indicate how predicted values are different from the target values, hence these can be negative or positive. The graph of the error histogram in Figure 8 visualized errors between the proposed model output and target values during the training process in which the vertical line represents the error at the zero point. In this graph, most of the errors are close to zero points, which demonstrates minor errors in all data clusters, i.e., the training, testing, and validation clusters.

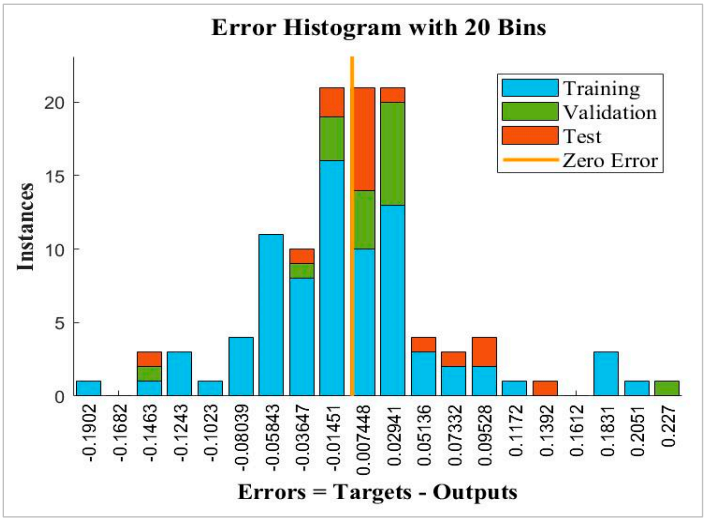

Figure 8. Training error histogram (PA3).

Training regression graphs in Figure 9 representing the correlation between the model output and target values of each data cluster during the process of training (training, validation, testing, and all data clusters) indicated that all the calculated regression coefficients (R) having the directions close to 1 i.e., 0.94182, 0.95648, 0.96596 and 0.94786 for training, validation, test, and all data clusters respectively which mean that the output values of the proposed model (PA3) are having a very good relationship with the target values.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 July 2024 doi:10.20944/preprints202407.2024.v1

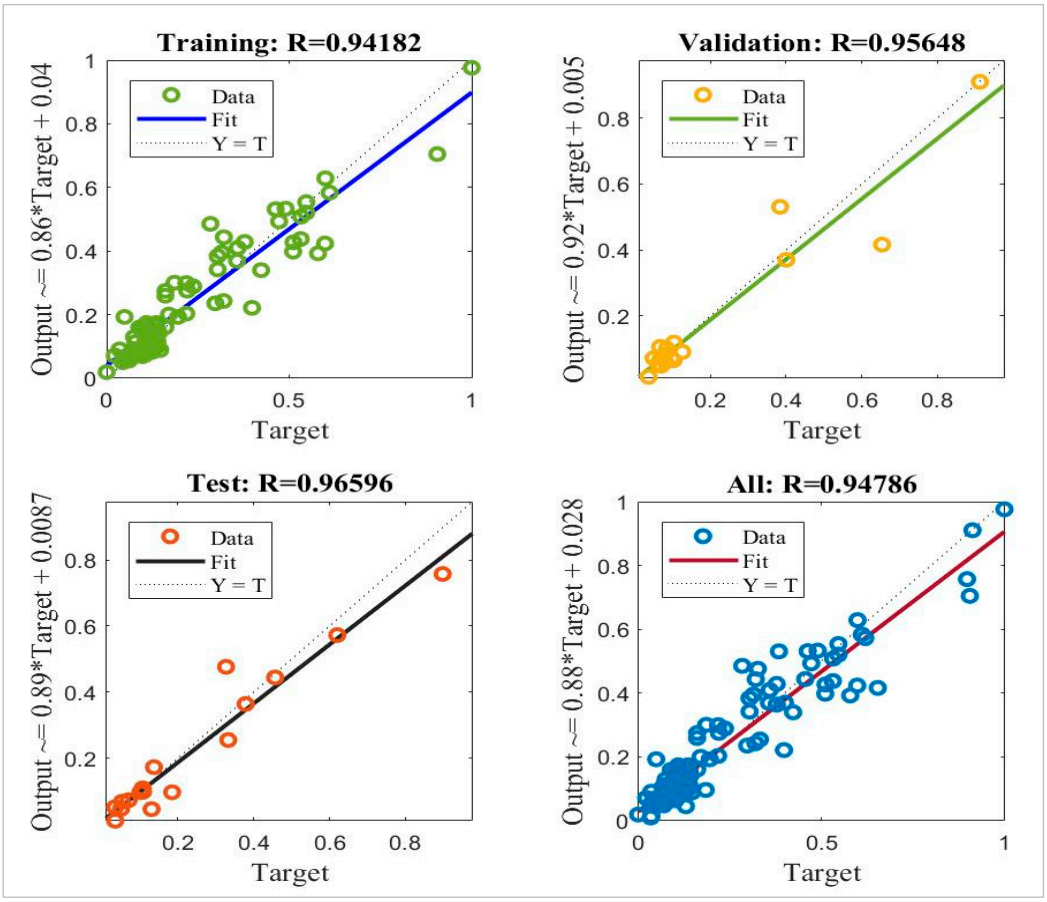

Figure 9. Training regression (PA3).

Figure 10 depicts the prediction results of the third scenario (PA3) of the proposed PSO-ANN approach compared to the target values with a total duration of two years i.e., from 2022 to 2023. We observed that the prediction values, represented by the blue curve, agree quite well with the target values, represented by the red curve, all along the testing duration. However, there are still some deviations visualized from the graph both, in the dry period and wet/peak period of the years in which the significant deviation mainly resulted from the high fluctuation and nonstationary of data in the periods of extreme weather conditions.

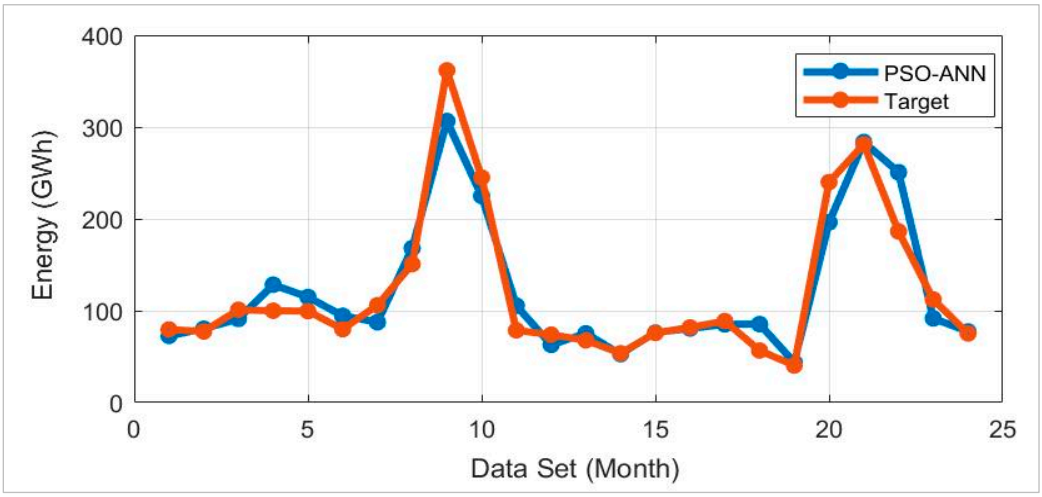

Figure 10. PSO-ANN prediction vs. target values (PA3).

 doi:10.20944/preprints202407.2024.v1

4.4. Comparison of PSO-ANN, ANN and NARX models

To compare the performance of the proposed PSO-ANN approach of which the third scenario (PA3) has been proved and selected to be the best scenario as described in Section 4.2, the same datasets i.e., the training and testing datasets, comprising of rainfall and lagged inflow, were also applied for the ANN and NARX prediction models and then the best scenario from each approach was selected and compared among the three approaches. In addition to the datasets, the same numbers of input combinations and hidden neurons were also applied for both, ANN and NARX models. The comparison among the best models, of the three approaches, has been made based on the statistical indicators in couple with the Taylor diagrams.

Tables 5 and 6 revealed that the third scenario of each prediction approach, with the structure (5,30,1) i.e., PA3, A3, and N3 for PSO-ANN, ANN, and NARX models respectively, has given the superior prediction results. We also observed that the third scenario (PA3) of the proposed approach with r = 0.973, RMSE = 22.994 and RE = 1.038 % for the single year of 2022, and r = 0.966, RMSE = 24.846 and RE = 2.853 % for 2023; outperformed those of the ANN and NARX models of which the scenario A3 of the ANN given r = 0.905, RMSE = 36.851 and RE = 7.872 % for the single year of 2022, and r = 0.942, RMSE = 37.238 and RE = 3.619 % for 2023; and the scenario N3 of the NARX given r = 0.939, RMSE = 30.315 and RE = 3.491 % for the single year of 2022, and r = 0.960, RMSE = 28.320 and RE = 3.548 % for 2023 i.e., the root mean square error (RMSE), and relative error (RE) of the proposed PSO-ANN model, both single years, were less than those of the ANN, and NARX models indicating more accuracy in the forecasting; on the other hand, the correlation coefficient (r) of the proposed model was closer to 1 than those of the other models (both years) depicted a stronger correlation between the predicted and target values.

Table 5. Results by statistical indicators of all models 2022.

Model Scenarios	Model Input combinations	Model Output	Different Models	Model Structures	*r*	RMSE	RE
PA1	R(t), I(t-1)	E(t)	PSO-ANN	(2,30,1)	0.951	38.168	8.261
PA2	R(t), R(t-1), I(t-1), I(t-2)	E(t)	PSO-ANN	(4,30,1)	0.968	33.261	18.459
PA3	R(t), R(t-1), I(t-1), I(t-2),I(t-3)	E(t)	PSO-ANN	(5,30,1)	**0.973**	**22.994**	**1.038**
PA4	R(t), R(t-1), R(t-2), I(t-1), I(t-2)	E(t)	PSO-ANN	(5,30,1)	0.942	32.924	7.847
PA5	R(t), R(t-1), R(t-2), I(t-1), I(t-2),I(t-3)	E(t)	PSO-ANN	(6,30,1)	0.930	34.354	3.662
A1	R(t), I(t-1)	E(t)	ANN	(2,30,1)	0.894	43.444	18.178
A2	R(t), R(t-1), I(t-1), I(t-2)	E(t)	ANN	(4,30,1)	0.900	55.512	41.521
A3	R(t), R(t-1), I(t-1), I(t-2),I(t-3)	E(t)	ANN	(5,30,1)	**0.905**	**36.851**	**7.872**
A4	R(t), R(t-1), R(t-2), I(t-1), I(t-2)	E(t)	ANN	(5,30,1)	0.821	76.032	27.723
A5	R(t), R(t-1), R(t-2), I(t-1), I(t-2),I(t-3)	E(t)	ANN	(6,30,1)	0.775	55.106	10.051
N1	R(t), I(t-1)	E(t)	NARX	(2,30,1)	0.894	40.182	3.030
N2	R(t), R(t-1), I(t-1), I(t-2)	E(t)	NARX	(4,30,1)	0.832	55.317	16.462
N3	R(t), R(t-1), I(t-1), I(t-2),I(t-3)	E(t)	NARX	(5,30,1)	**0.939**	**30.315**	**3.491**
N4	R(t), R(t-1), R(t-2), I(t-1), I(t-2)	E(t)	NARX	(5,30,1)	0.796	51.605	7.338
N5	R(t), R(t-1), R(t-2), I(t-1), I(t-2),I(t-3)	E(t)	NARX	(6,30,1)	0.844	45.577	4.250

Table 6. Results by statistical indicators of all models 2023.

Model Scenarios	Model Input combinations	Model Output	Different Models	Model Structures	*r*	RMSE	RE
PA1	R(t), I(t-1)	E(t)	PSO-ANN	(2,30,1)	0.905	44.925	15.059
PA2	R(t), R(t-1), I(t-1), I(t-2)	E(t)	PSO-ANN	(4,30,1)	0.965	30.668	5.832
PA3	R(t), R(t-1), I(t-1), I(t-2),I(t-3)	E(t)	PSO-ANN	(5,30,1)	**0.966**	**24.846**	**2.853**
PA4	R(t), R(t-1), R(t-2), I(t-1), I(t-2)	E(t)	PSO-ANN	(5,30,1)	0.930	36.757	7.928

 doi:10.20944/preprints202407.2024.v1

PA5	R(t), R(t-1), R(t-2), I(t-1), I(t-2),I(t-3)	E(t)	PSO-ANN	(6,30,1)	0.956	27.934	3.001
A1	R(t), I(t-1)	E(t)	ANN	(2,30,1)	0.909	44.268	24.058
A2	R(t), R(t-1), I(t-1), I(t-2)	E(t)	ANN	(4,30,1)	0.873	51.048	28.574
A3	R(t), R(t-1), I(t-1), I(t-2),I(t-3)	E(t)	ANN	(5,30,1)	**0.942**	**37.238**	**3.619**
A4	R(t), R(t-1), R(t-2), I(t-1), I(t-2)	E(t)	ANN	(5,30,1)	0.910	55.710	25.550
A5	R(t), R(t-1), R(t-2), I(t-1), I(t-2),I(t-3)	E(t)	ANN	(6,30,1)	0.821	57.602	11.466
N1	R(t), I(t-1)	E(t)	NARX	(2,30,1)	0.905	43.566	14.521
N2	R(t), R(t-1), I(t-1), I(t-2)	E(t)	NARX	(4,30,1)	0.922	39.719	5.724
N3	R(t), R(t-1), I(t-1), I(t-2),I(t-3)	E(t)	NARX	(5,30,1)	**0.960**	**28.320**	**3.548**
N4	R(t), R(t-1), R(t-2), I(t-1), I(t-2)	E(t)	NARX	(5,30,1)	0.943	34.179	6.101
N5	R(t), R(t-1), R(t-2), I(t-1), I(t-2),I(t-3)	E(t)	NARX	(6,30,1)	0.911	39.604	7.106

In addition to the clear comparison in Tables 5 and 6, Figure 11 depicts the predicted values of the three prediction models (scenarios PA3, A3, and N3) versus the target values. We observed that all three prediction curves can track changes of the target (red curve) throughout the testing duration whereas the curve of the proposed PSO-ANN model has a better coincidence than the others, especially during the peak periods of the years indicating superiority in the forecasting. However, the ANN, and NARX models, in some few months during the dry season, performed slightly better than the proposed model, this depicted the advantages and limitations of each approach.

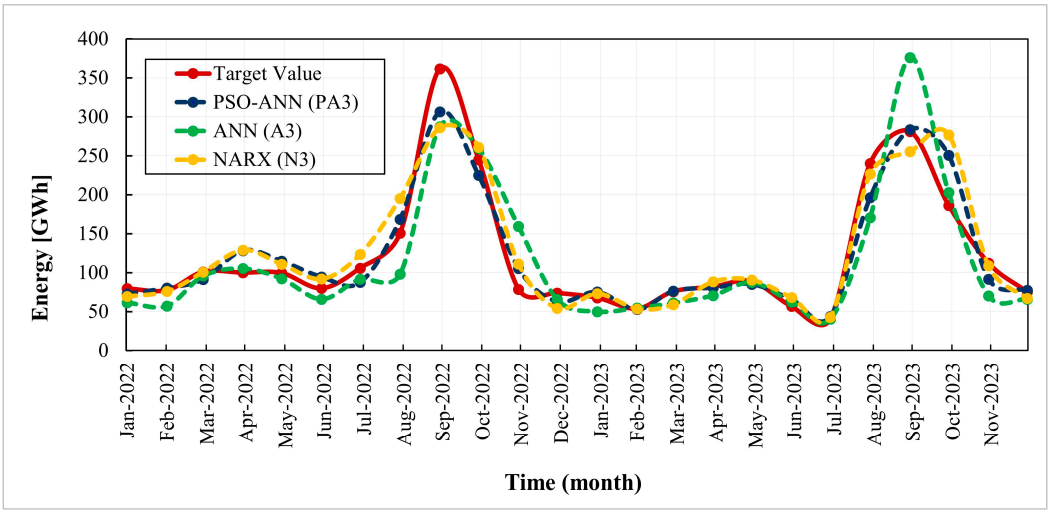

Figure 11. Comparison among the three models with the target values.

Testing regression graphs with a testing duration of 2 years in Figure 12, illustrate the correlation between each model prediction and target values of which the calculated regression coefficients (R) of the proposed PSO-ANN model at 0.95758 was superior to those of the ANN and NARX models (0.93290, and 0.90303 respectively). In general, Correlation coefficients with a magnitude between 0.9 and 1.0 indicate variables that can be considered very highly correlated. Correlation coefficients with a magnitude between 0.7 and 0.9 indicate variables that can be considered highly correlated. Correlation coefficients with a magnitude between 0.5 and 0.7 indicate variables that can be considered moderately correlated. Whilst the correlation coefficients with a magnitude between 0.3 and 0.5 indicate variables that have a low correlation, and the correlation coefficients with a magnitude less than 0.3 have little if any (linear) correlation.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 July 2024 doi:10.20944/preprints202407.2024.v1

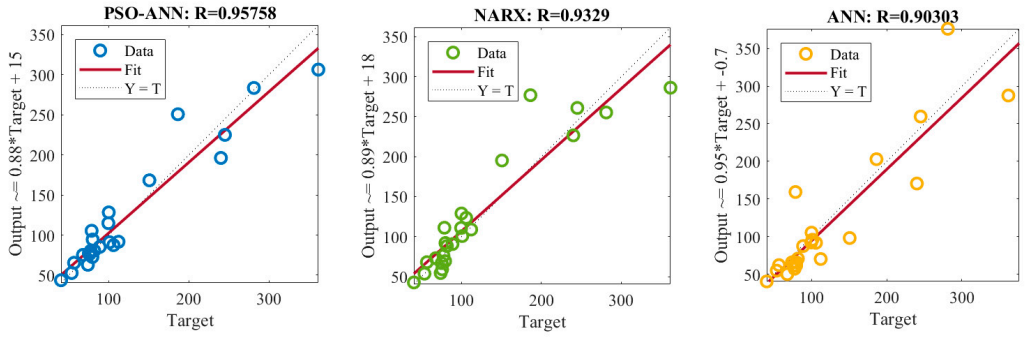

Figure 12. Regression comparison among the three models (testing duration of 2 years).

Taylor diagram in Figure 13 illustrates that the root mean square deviation (RMSD), and correlation coefficient (r) of the proposed PSO-ANN model have more vicinity to the target location compared to the ANN and NARX models. With respect to the standard deviation, the ANN and NARX models depicted slightly closer to the target location; however, the standard deviation just indicating how far each value lies from the mean or expected value (A high standard deviation means that values are generally far from the mean, while a low standard deviation indicates that values are clustered close to the means) thus the size of the standard deviation does not reflect a better or worse prediction performance whilst those of the proposed PSO-ANN model herein depicted less than those of the target as well as of the ANN and NARX models. Based on the comparison results illustrated by the Talor diagram in combination with the results of statistical indicators summarized in Tables 5 and 6, we can conclude that the proposed PSO-ANN approach is more robust and efficient for the available energy prediction for the case study hydropower plant.

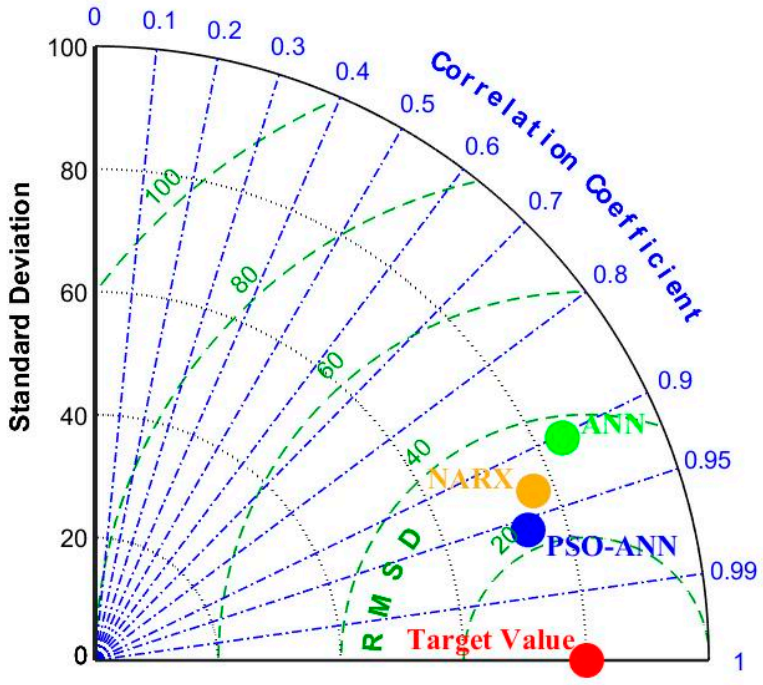

Figure 13. Taylor diagram of the three models vs. the target.

5. Conclusion

To forecast the available energy portion that corresponds to the reservoir inflow of the month(s) ahead for ensuring optimal generation planning associated with the revenue maximization as well as the environmental impact prevention, and flood control at the upstream and downstream of the basin, this study has developed the effective hybrid model by integrating Artificial Neural Network

 doi:10.20944/preprints202407.2024.v1

(ANN) and Particle Swarm Optimization (PSO) based on hydrological and meteorological data pre-processed by autocorrelation function (AFC), cross-correlation function (CCF), and normalization techniques whereas the Nam Gnouang reservoir of THHP in Laos was selected to be the case study hydropower plant. The model was evaluated using statistical indicators such as correlation coefficient (r), relative error (RE), root mean square error (RMSE), and Taylor diagrams as described in Section 3. Based on the results summarized in Tables 5 and 6, the third scenario of each prediction approach, with the structure (5,30,1) i.e., PA3, A3, and N3 for PSO-ANN, ANN, and NARX models respectively, has given the best prediction results whereas the proposed PSO-ANN prediction model given r = 0.973, RMSE = 22.994 and RE = 1.038 % for the single year of 2022, and r = 0.966, RMSE = 24.846 and RE = 2.853 % for 2023 demonstrated the superiority over the other models i.e., ANN, and NARX models. Taylor's diagram in Figure 12 also indicated the same conclusion. These comparison results proved that the proposed approach is more robust and efficient for the available energy forecasting. However, this study has some limitations such as the limited duration of the data, some missing data, limited numbers of hydrological and meteorological factors, and highly fluctuated and nonstationary data significantly affecting the prediction results. For future studies, we suggest considering using the data of longer periods and other relevant hydrological and meteorological variables as well as to seek for an optimal technique for handling climate variability and nonstationary data. In addition, incorporating data from multiple sources by considering using data from satellites, reanalysis products, and ground observations to capture the spatial and temporal variability of climate variables is also recommended.

Author Contributions: Conceptualization, S.K. and V.K.; methodology, S.K. and V.K.; software, S.K. and V.K.; validation, S.K.; formal analysis, S.K.; investigation, S.K.; resources, V.K.; data curation, V.K.; writing—original draft preparation, V.K.; writing—review and editing, S.K.; visualization, S.K.; supervision, S.K.; project administration, S.K.; funding acquisition, S.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Fang Y., Ahmadianfar, I., Koucheksaraee, A. S., Azarsa R., Scholz, M., & Yaseen, Z. M. An accelerated gradient-based optimization development for multi-reservoir hydropower systems optimization. Energy Reports **2021**, 7(1474), 7854–7877. https://doi.org/10.1016/j.egyr.2021.11.010.
2. Long, Y., Hui, W., Changbo, J., & Shang, L. Seasonal inflow forecasts using gridded precipitation and soil moisture information: Implications for reservoir operation. Water Resources **2019**, 33, 3743–3757. https://doi.org/10.1007/s11269-019-02330-8.
3. Fugang, L., Guangwen, M., Shijun, C., & Weibin, H. An ensemble modeling approach to forecast daily reservoir inflow using bidirectional long and short-term memory (Bi-LSTM), variational mode decomposition (VMD), and energy entropy method. Water Resource Management **2021**, 35, 2941–2963. https://doi.org/10.1007/s11269-021-02879-3.
4. Rajesh, M., Sachdeva, A., Pansari, S. V., Srivastav, A., & Rehana, S. Improving short-range reservoir inflow forecasts with machine learning model combination. Water Management **2022**, 37, 75–90. https://doi.org/10.1007/s11269-022-03356-1.
5. Yu, J., Chang, W. S., & Dong, Y. Building energy prediction models and related uncertainties. A Review **2022**, 12, 1284. https://doi.org/10.3390/buildings12081284.
6. Khorram, S., & Jehbez, N. A hybrid CNN-LSTM approach for monthly reservoir inflow forecasting. Water Resource Management **2023**, 37, 4097–4121. https://doi.org/10.1007/s11269-023-03541-w.
7. Saab, S. M., Othman, F., Tan, C. G., Allawi, M. F., Sherif, M., & El-Shafie, A. Utilizing deep learning machine for inflow forecasting in two different environment regions: A case study of a tropical and semi-arid region. Applied Water Science **2022**, 12, 272. https://doi.org/10.1007/s13201-022-01798-x.
8. Latif, S. D., & Ahmed, A. N. A review of deep learning and machine learning techniques for hydrological inflow forecasting. Environmental Development and Sustainability **2023**, 25, 12189–12216. https://doi.org/10.1007/s10668-023-03131-1.

9. Awan, J. A., & Bae, D. H. Improving ANFIS based model for long-term dam inflow prediction by incorporating monthly rainfall forecasts. Water Resource Management **2014**, 28, 1185–1199. https://doi.org/10.1007/s11269-014-0512-7.
10. Chang, F. J., Lo, Y. C., Chen, P. A., Chang, L. C., & Shieh, M. C. Multi-step-ahead reservoir inflow forecasting by artificial intelligence techniques. Smart Innovation System and Technology **2015**, 30, 235–249. https://doi.org/10.1007/978-3-319-13545-8_14.
11. Chiamsathit, C., Adeloye, A. J., & Soundharajan, B. S. Inflow forecasting using artificial neural networks for reservoir operation. Hydrology Science **2019**, 373, 209–214. https://doi.org/10.5194/piahs-373-209-2016.
12. Babaei, M., Moeini, R., & Ehsanzadeh, E. Artificial neural network and support vector machine models for inflow prediction of dam reservoir (Case study: Zayandehroud dam reservoir). Water Resource Management **2019**, 33, 2203–2218. https://doi.org/10.1007/s11269-019-02252-5.
13. Noorbeh, P., Roozbahani, A., & Moghaddam, H. K. Annual and monthly dam inflow prediction using Bayesian networks. Water Resource Management **2020**, 34, 2933–2951. https://doi.org/10.1007/s11269-020-02591-8.
14. Hadiyan, P. P., Moeini, R., Ehsanzadeh, E., & Karvanpour, M. Trend analysis of water inflow into the dam reservoirs under future conditions predicted by dynamic NAR and NARX models. Water Resource Management **2022**, 36, 2703-2723. https://doi.org/10.1007/s11269-022-03170-9.
15. Hanoon, M. S., Ahmed, A. N., Razzaq, A., Oudah, A. Y., Alkhayyat, A., Huang, Y. F., Kumar, P., & El-Shafi, A. Prediction of hydropower generation via machine learning algorithms at three Gorges dam, China. Hydropower Generation **2022**, 14, 1019-1914. https://doi.org/10.1016/j.asej.2022.101919.
16. Fallah, S. N., Deo, R. C., Shojafar, M., Conti, M., & Shamshirband, S. Computational intelligence approaches for energy load forecasting in smart energy management grids: State of the art, future challenges, and research directions. Smart Grid **2018**, 11, 596. https://doi.org/10.1016/j.asej.2022.101919.
17. Moeeni, H., Bonakdari, H., & Ebtehaj, I. Monthly reservoir inflow forecasting using a new hybrid SARIMA genetic programming approach. Earth System Science **2017**, 18, 126. https://doi.org/10.1007/s12040-017-0798-y.
18. Qi, Y., Zhou, Z., Yang, L., Quan, Y., & Miao, Q. A decomposition-ensemble learning model based on LSTM neural network for daily reservoir inflow forecasting. Water Resource Management **2019**, 33, 4123–4139. https://doi.org/10.1007/s11269-019-02345-1.
19. Feizi, H., Apaydin, H., Sattari, M. T., Colak, M. S., & Sibtain, M. Improving reservoir inflow prediction via the rolling window and deep learning-based multi-model approach: A case study from Ermenek dam, Turkey. Stochastic Environment Reservation and Risk Assessment **2022**, 36, 3149–3169. https://doi.org/10.1007/s00477-022-02185-3.
20. Ghazali, M., Honar, T., & Nikoo, M. R. A fusion-based neural network methodology for monthly reservoir inflow prediction using MODIS products. Hydrology Science **2019**, 63, 2076-2096. https://doi.org/10.1080/02626667.2018.1558365.
21. Bayat, N., & Park, J. H. Particle swarm optimization-based demand response using artificial neural network-based load prediction. North American Power Symposium (NAPS) **2023**, https://doi.org/10.1109/NAPS56150.2022.10012263.
22. Pornnapa, P., Paskorn, C., & Chuchoke, A. Water level prediction using artificial neural network with particle swarm optimization model. Natural Disaster Management **2017**, https://doi.org/10.1109/ICoICT.2017.8074670.
23. Rosli, N.S., Ibrahim, R., & Ismail, I. Neural network model with particle swarm optimization for prediction in gas metering system. Intelligent and Advanced System **2017**, https://doi.org/10.1109/ICIAS.2016.7824049.
24. Hamada, M., & Hassan, M. Artificial neural networks and particle swarm optimization algorithms for preference prediction in multi-criteria recommender systems. Informatics **2018**, https://doi.org/10.3390/informatics5020025.
25. Fadlallah, S. O., Anderson, T. N., & Nates, R. J. Artificial neural network–particle swarm optimization (ANN-PSO) approach for behaviour prediction and structural optimization of lightweight sandwich composite heliostats. Informatics **2021**, 46, 12721–12742. https://doi.org/10.1007/s13369-021-06126-0.
26. Seguini, M., Boutchicha, D., & Nedjar, D. Crack prediction in pipeline using ANN-PSO based on numerical and experimental modal analysis. Smart Structures and System **2021**, 27(3), 507–523. http://dx.doi.org/10.12989/sss.2021.27.3.507.
27. Lu, H., Wu, J., Ruan, Y., Qian, F., Meng, H., Gao, Y., & Xu, T. A multi-source transfer learning model based on LSTM and domain adaptation for building energy prediction. International Journal of Electrical Power & Energy Systems **2023**, 149, 109024. https://doi.org/10.1016/j.ijepes.2023.109024.
28. Moayed, F. A., & Shell, R. L. Application of artificial neural network models in occupational safety and health utilizing ordinal variables. Occupational Hygiene **2021**, 55(2), 132–142. https://doi.org/10.1093/annhyg/meq079.

29. Tzeng, F. Y., & Ma, K. L. Opening the black box | data driven visualization of neural networks. Computer Science **2005**, http://dx.doi.org/10.1109/VISUAL.2005.1532820.
30. Bhadeshia, H. K. D. H. Neural networks in materials science. ISIJ International **1999**, 39(10), 966–979.
31. Shami, T. M., Alswaitti, M., El-Saleh, A. A., & Al-Tashi, Q. Particle swarm optimization: A comprehensive survey. IEEE Access (Electronic ISSN: 2169-3536) **2022**, 10, 10031–10061. http://dx.doi.org/10.1109/ACCESS.2022.3142859.
32. Wang, R., Hao, K., Chen, L., Wang, T., & Jiang, C. A novel hybrid particle swarm optimization using adaptive strategy. Information science **2021**, 579, 231–250. https://doi.org/10.1016/j.ins.2021.07.093.
33. Semero, Y. K., Zheng, D., & Zhang, J. A PSO-ANFIS based hybrid approach for short term PV power prediction in microgrids. Electrical Power Components and Systems **2018**, 46, 95–103. https://doi.org/10.1080/15325008.2018.1433733.
34. Kennedy, J. The behavior of particles. Volitionary Program **2005**, 1447, 4104–4109. https://doi.org/10.1007/BFb0040809.
35. Sharma, D. K., Hota, H. S., Brown, K., & Handa, R. Integration of genetic algorithm with artificial neural network for stock market forecasting. Engineering and Management **2021**, 13, 828–841. http://dx.doi.org/10.1007/s13198-021-01209-5.