Pumps-as-Turbines’ (PaTs) performance prediction improvement using evolutionary artificial neural networks

Pumps-as-Turbines' (PaTs) performance prediction improvement using evolutionary artificial neural networks

Akbar Telikani[a], Mosè Rossi[b,*], Naghmeh Khajehali[a], Massimiliano Renzi[c]

[a] *University of Wollongong, School of Computing and Information Technology, Northfields Ave, North Wollongong NSW 2500, Australia*
[b] *Marche Polytechnic University, Department of Industrial Engineering and Mathematical Sciences, Via Brecce Bianche 12, 60131, Ancona, Italy*
[c] *Free University of Bozen/Bolzano, Faculty of Science and Technologies, Piazza Università 5, 39100, Bolzano, Italy*

Abstract

Energy production from clean sources is mandatory to reduce pollutant emissions. Among different options of hidden hydropower potential, Pump-as-Turbine (PaT) represents a viable solution in pico- and micro-hydropower applications for its flexibility and low-cost. Pumps are widely available in the global market in terms of both sizes and spare parts. To date, there are several PaTs' performance prediction models in the literature, but very few of them use optimization algorithms and only for specific and limited prediction goals. The present work proposes evolutionary Artificial Neural Networks (ANNs) based on JADE, which is a typology of differential evolution algorithm, to forecast Best Efficiency Point (BEP) and performance curves of a PaT, starting from the pump operational data. In this model, JADE is employed as optimizer of basic ANNs to upgrade parameter values of the learning rate, weights, and biases. The accuracy of the proposed model is evaluated through experimental data available from the literature and compared to a basic ANN and two versions of the differential evolution algorithm. Results are also validated with experiments on a PaT, showing that the proposed method can achieve an average R^2-value of 0.97, which is 5% higher than the one obtained with a basic ANN.

*Corresponding author
Email address: mose.rossi@staff.univpm.it (Mosè Rossi)

Keywords: Pumps-as-Turbines; Performance forecast; Artificial Neural Network; Evolutionary algorithms; Adaptive Differential Evolution; Hidden Hydropower.

1. Introduction

Nowadays, fossil fuels such as petroleum, natural gas, and coal are the main sources of energy, being essential to carry on the daily life activities of all the human kind. It is estimated that the global primary energy demand will expand within the year 2030 by approximately 60% compared to the year 2002 with an average increase of 1.7% per year [1]. Due to the negative environmental impact caused by the use of traditional energy sources, the exploitation of renewables has become a mandatory target to meet present and future challenges [2, 3, 4]. Indeed, it has been predicted that the electricity generated by renewable energy sources will rise up to 39% by the year 2050 [5]. Fossil fuels will be replaced by renewable energy in the next fifty years; among the most known technologies that exploit clean sources, hydropower is one of the most used with a high penetration in the power generation sector [6]. Large-scale hydropower has been widely exploited in the past decades and the main water resources have been already exploited, especially in the developed countries. New installations are therefore limited and impose complex challenges, such as the need of collecting a large amount of water using dams, which are costly and impacting infrastructures, difficulty on producing electricity in remote and rural areas, and significant environmental impacts. Within this context micro-hydropower applications are gaining more and more interest. Micro-hydropower is considered an effective alternative that benefits the possibility to recover the residual energy potential from small applications with lower costs [7, 8], the so-called hidden hydropower potential.

In such a context, the Pump-as-Turbine (PaT) technology is considered one of the most viable and interesting alternatives in terms of both energy recovery and power generation. PaTs are common pumps that are used in reverse mode: the advantages of PaTs are related to their low capital and O&M costs, as well as their large availability in the market, in terms of sizes and spare parts, that make them a competitive solution [6]. Furthermore, PaTs can have a payback period five times shorter than conventional micro hydro-turbines [9]. PaTs are used for recovering energy in several applications like water supply systems, water distribution networks, wastewater plants, and irrigation systems. To fasten the wide-spread applications of PaTs, a reliable forecast model of their performance in turbine mode is anyway needed; this task has been a significant challenge for both choosing and applying this technology because pump manufactures provide only the specifications of the hydraulic machine in pump mode and not in turbine mode [7]. Indeed, without a reliable figure of the machine's performance in reverse mode, the designers cannot choose the proper machine to be used in a specific installation site and also an economic evaluation of the return of the investment is not feasible. The performance prediction of PaTs in turbine mode (BEP + off-design operating conditions) is a quite complex task due to the fluid flow phenomena that change depending on the single machine design and operating conditions.

Thanks to the effort of several scientists in the last years, different performance prediction models for PaTs are now available in the literature [10, 11, 12]. The studies presented in [10, 13] aimed at forecasting the performance of PaTs at their BEP through relationships between the specific speed (N_S) of the PaTs in pump and turbine modes, respectively. Huang et al. [11] used the relationships between the rotor and the volute to forecast both flow rate and head of PaTs. Then, the Euler equation was used for defining the turbomachinery and the velocity vector relations at the inlet and outlet of the rotor. Novara et al. [12] proposed a methodology based on fixed-coefficient polynomials by analysing and elaborating the performance of 113 PaTs. Compared to other prediction models, the accuracy of the head curves improved by 5%; furthermore, the average mechanical efficiency of PaTs with varying flow rate has been also compared to a Francis turbine, showing good results. Mitrovic et al. [14] used the classical hydraulic regulation scheme with the Nedler–Mead simplex direct search algorithm to find the best available centrifugal PaTs on the market. Up to now, there are still two main gaps on the performance prediction models of PaTs available in the literature that indicate room for improvement in the forecast capability:

- while most of the developed models provide now reliable results, they refer mostly on single-stage radial machines: this is a strong limitation since also other kinds of PaTs (e.g., axial ones) can be used in several applications; therefore, in such cases, the preliminary performance assessment can still be made only through Computational Fluid Dynamics (CFD) analyses that are complex and expensive. [15];

- so far, optimization models have not been implemented in the development of both BEP and off-design performance prediction models of PaTs, while they could further contribute to improve the accuracy of the forecasts.

This paper focuses on the prediction of both the BEP and off-design performance curves of PaTs using a newly developed Artificial Intelligence (AI)-based model. This model is the combination of evolutionary computation algorithms and Artificial Neural Networks (ANNs). Such models have shown to be more accurate and robust than traditional or simpler predictive models in several energy-related applications such as solar radiation prediction [5], wind power ramp events detection [16], short term load forecast [17], and dew point cooler prediction [18]. The purpose of using evolutionary algorithms in neural networks is to reduce the local optimization problem in optimization algorithms (e.g., Stochastic Gradient Descent (SGD) and Nesterov) for training neural networks, which are affected by network hyper-parameters (e.g., learning rate). In traditional neural networks, the hyper-parameters are adjusted by an expert based on prior knowledge, which is a very complex task. Evolutionary algorithms can be also useful in automatic hyper-parameters adjustment. In this paper, evolutionary algorithms are used both as a network training algorithm and as a hyper-parameter optimizer. In this regard, JADE approach [19] is used, being considered as a robust and reliable evolutionary algorithm. To the best of the authors' knowledge, there is no existing research work that employs evolutionary ANNs for PATs' performance prediction and hydraulic machines so far. The characteristics of these algorithms should help to increase the forecasting capability that still misses some accuracy, especially in off-design operating conditions and for some typology of PaTs. Performance of the designed evolutionary ANN is compared to other AI models, including regression, a basic ANN and two differential evolution-based ANN using the data collected from experimental data.

The paper is organized as follows: Section 2 briefly summarizes some works related to PaTs performance at both BEP and off-design operating conditions, and evolutionary ANNs available in the literature. In Section 3, the characteristics of the dataset used in this work are described. Section 4 proposes an evolutionary ANN that has been developed by employing JADE approach for the parameters optimization and topology adjusting to improve the PaTs' performance prediction. Section 5 evaluates the proposed algorithm based on different measures, including the comparison between the results obtained in [7] with those of the proposed ANN based on JADE and discusses the results. Finally, the conclusions are provided in Section 6.

2. Review of Pump-as-Turbine (PaT) studies and performance prediction models

In the scientific literature there are several studies related to the use of PaTs that are mainly applied to water supply systems/distribution networks and industrial plants for energy recovery purposes. Crespo Chacón et al. [20] presented a methodology for designing a micro-hydropower plant using PaTs and results were assessed considering a real case study in Southern Spain. The selection of the most suitable PaT has been carried out by using equations obtained by [21, 12, 14]; the forecast performance showed slightly difference with respect to the results measured in the real application. Indeed, 2258€ and 8.4 tCO_2 were achieved in the real case study, being close to the estimated ones with the model presented in [22]. Spedaletti et al. [23] investigated on the use of PaTs in a water supply system of a town located in the Center of Italy. The study concluded that the size of the PaT has an impact on the Payback Period (PBP) of the investment. Smaller machines $(1-2\ kW)$ normally present longer return of the investment (almost 11 years), while the PBP has been sensibly shortened considering PaT's power outputs higher than $3-4\ kW$ (almost 6 years). Bekker et al. [24] studied different technologies for exploiting the low-head hydropower potential available in wastewater plants. PaTs are also suitable for such applications since high specific speed machines can operate down to an available head range of $3-30$ m.

However, as stated by all the previously mentioned studies, the application of PaTs requires the accurate knowledge of their performance in reverse mode, and thus prediction models are required to assess the operation of these hydraulic machines. The analytical approach of Huang et al. [11] is a work for predicting the flow rate and head of PaTs in off-design operating conditions in pump or turbine mode. Even though a relationship between the rotor and the stator was used, detailed information about the geometrical factors of a PaT are required. Due to the difficulty of getting the geometrical parameters from pump manufacturers, some studies tried to develop techniques to predict the PaTs performance in turbine mode using operating data of PaTs in turbine mode obtained through experimental campaigns. A mechanism for predicting curves and rotational speed of a PaT through a Hermite polynomial chaos expansion has been developed in [25]. Novara et al. [12] designed a similar methodology based on fixed-coefficient polynomials to predict the performance curves of PaTs based on the BEP knowledge. A data-driven model based on a large set of data was presented by Rossi et al. [6] that was validated with three different centrifugal PaTs, showing good accuracy. The off-design performance of PaTs was predicted through correlations involving the non-dimensional analysis and the normalization method; however, this model requires the knowledge of both flow rate and head factors of the PaTs in turbine mode.

Since detailed geometrical information about pump data cannot be obtained easily from pump manufacturers, some papers formulated analytical equations to find correlations between data in pump mode and those in turbine mode to forecast the BEP of PaTs in turbine mode. Some studies focused on developing analytical methodologies to evaluate flow rate, head, and efficiency of PaTs in turbine mode [26]. However, this methodology is based on a limited number of PaTs that lead to good accuracy only for the machines belonging to a narrow range of specific speed N_S.

The prediction of both BEP and off-design operating conditions has attracted the attention of several scientists. In this regard, a statistical model has been developed by Barbarelli et al. [21] based on the experimental performance data of 12 PaTs. Recently, Rossi et al. [7] developed a PaTs' performance prediction model based on ANN for both the BEP and the PaTs' performance curves forecast. Their results led to a flow rate forecast error of 5.18% and a head forecast error of 1.85%. In addition, a physics-based simulation model has been developed by Venturini et al. [27] that used an optimization procedure to identify the non-available performance and geometrical data of the machine, evaluate the specific parameters, and finally predict the PaTs' performance curves.

3. Dataset description

In this study, a dataset that includes the performance of 32 PaTs in both pump and turbine modes has been used [7]. The physical performance values were converted into non-dimensional parameters to feed the ANNs; among them, the specific speed (N_S) is the non-dimensional parameter that better characterizes the functioning principle of a turbomachinery and its design as well (Eq. 1). In the case of the same N_S values, the machines have the similar fluid dynamic behaviour and belong to the same typology [7]. The flow coefficient (ϕ), the head coefficient (ψ), and the power coefficient (λ) are the other non-dimensional parameters included in the dataset. The previous mentioned values are obtained using Eqs. (2), (3), and (4), respectively.

$$N_S = \omega \left[\frac{rad}{s}\right] \frac{\sqrt{Q\,[m^3/s]}}{\sqrt[4]{\left(g\,[m/s^2] \cdot H\,[m]\right)^3}} \tag{1}$$

$$\phi = \frac{Q\,[m^3/s]}{\omega\,[rad/s] \cdot \left(D\,[m]\right)^3} \tag{2}$$

$$\psi = \frac{g\,[m/s^2] \cdot H\,[m]}{\left(\omega\,[rad/s]\right)^2 \cdot \left(D\,[m]\right)^2} \tag{3}$$

$$\lambda = \frac{P[W]}{\rho[kg/m^3] \cdot (\omega\,[rad/s])^3 \cdot \left(D\,[m]\right)^5} \tag{4}$$

Table 1 lists the minimum and the maximum values of the non-dimensionless parameters constituting the dataset, which includes nine variables: six of them are inputs, while the remaining three are outputs.

Table 1: Main non-dimensional parameters of the 32 PaTs used in this work

	Parameter	**Symbol**	**Min**	**Max**
Input	Flow coefficient	ϕ	0.001	0.127
	Head coefficient	ψ	0.007	0.146
	Mechanical efficiency	η	0.44	0.87
	Power coefficient	λ	0.0001	0.0102
	Specific speed	N_S	0.28	2.24
	Rotational speed	ω	78.54	256.06
Output	Output 1	ϕ_t/ϕ_{tBEP}	0.17	1.75
	Output 2	ψ_t/ψ_{tBEP}	0.29	2.60
	Output 3	η_t/η_{tBEP}	0.01	1.00

4. Evolutionary ANN for PaT's performance forecasting

In this section, an evolutionary ANN algorithm using JADE is proposed. JADE approach is considered as an optimizer for the ANN's hyper-parameters and ANN training algorithm. Fig. 1 shows a schematic flowchart of the proposed algorithm, named JADE4ANN.

4.1. Chromosome encoding

Encoding a real-world application in the form of a solution is considered an essential process in evolutionary algorithms that aims to effectively display the influencing factors related to an application. In this case, a chromosome represents three types of ANN parameters: learning rate, bias, and weights. The best final evolved chromosome is selected as the final model network parameters. Based on the optimization goals, a chromosome has three parts (Fig. 2): the first gene represents the learning rate, the next m bits comprise the bias values, and the ending n genes include weights of the ANN. The size of m is $(L \times h) + 1$, where L and h are the numbers of the hidden layers and neurons in each layer, respectively. The size of n is equal to $(I \times h) + ((L-1) \times h) + h$, where I is the number of input variables (i.e., the size of the input layer). The initial population is initialized with a random value between -1.0 and 1.0.

4.2. ADE operators

Once the population is initialized, JADE algorithm starts to improve the population from generation to generation using three evolutionary operators: mutation, crossover, and selection. The DE/rand/1 mutation strategy is used to mutate chromosomes based on the current parents in the population. After the mutation, the non-consecutive binomial crossover is applied to produce offspring from two parents (Fig. 3). There is a dedicated crossover probability per individual instead of a fixed value for all chromosomes. In the selection step, the best individual is selected from its parents and offspring based on their fitness. At each generation step of JADE, the control parameters are automatically optimized without prior parameter setting knowledge. A normal distribution with mean μ_{Cr} and the standard deviation $\sigma = 0.1$ is used to generate the crossover probability of each chromosome independently. The Cauchy distribution with location parameter μ_F and scale parameter 0.1 is employed to generate the mutation factor of each chromosome independently. Parameter adaption for both μ_{Cr} and μ_F is performed at the end of each generation.

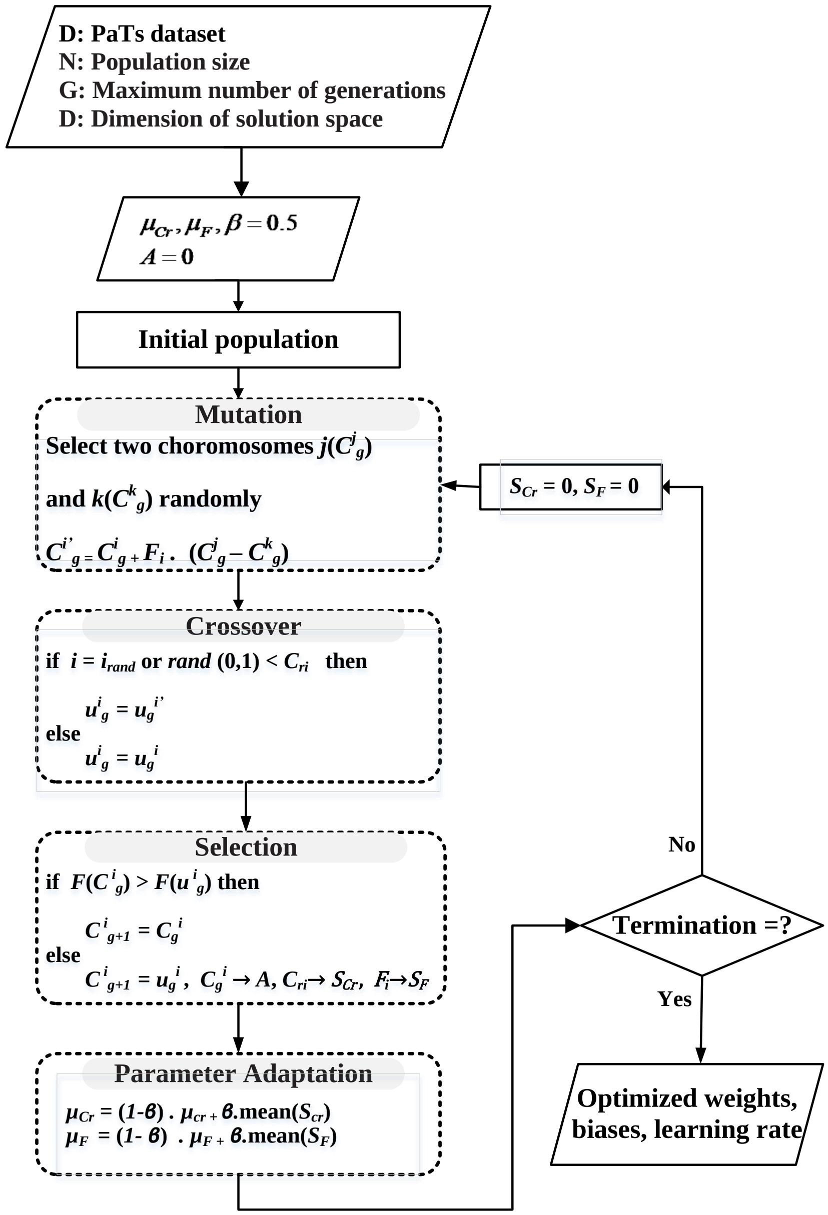


Figure 1: Flowchart of JADE for ANN (JADE4ANN)

4.3. Fitness function

To compute the quality of the individuals, the RMSE is used as a fitness function. To calculate the RMSE, the predicted value and its real value are

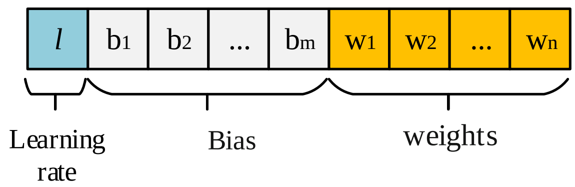


Figure 2: Chromosome representation

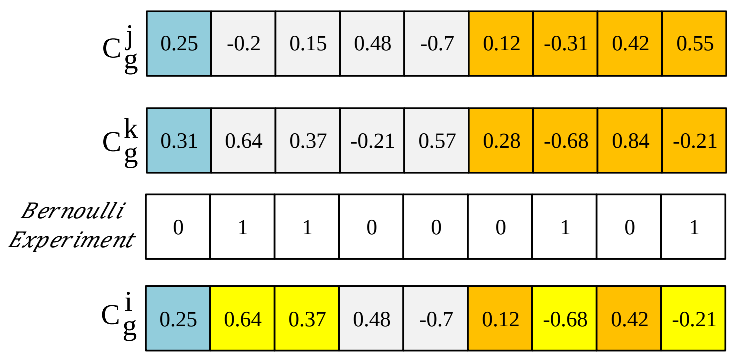


Figure 3: Non-consecutive binomial crossover

used as given in Eq. (5):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (y_i - \bar{y}_i)}{N}} \tag{5}$$

where y_i refers to the predicted, $\bar{y}_i$ is the actual value and N is the size of the dataset. The JADE technique tries to minimize the RMSE value.

4.4. Termination criteria

Evolutionary algorithms aim at providing a diverse and convergence population using an evolution process. A termination condition is set for an evolutionary algorithm, which is often the maximum number of generations. In the proposed approach, a convergence condition is defined in which the evolution process finishes when the best fitness value of the population does not change after the past 20 generations. Therefore, the evolution process finishes when either the maximum number of iterations or the convergence condition has been reached.

5. Results and discussion

This section provides the results obtained from experiments carried out with the JADE4ANN model in comparison with a regression predictor, the ANN used in [7], two evolutionary ANN models based on the basic DE [28] and Self-adaptive DE (SaDE) [29] and called these two evolutionary ANN models as DE4ANN and SaDE4ANN, respectively. To set the influencing parameters of the JADE algorithm (i.e., population size and β), a sensitivity analysis in Subsection 5.2 has been carried out. Based on the results obtained from the experiments in the sensitivity analysis, the best values for the population size and β are 100 and 0.5, respectively. Also, the value of the learning rate for the ANN algorithm is set to be equal to 0.1.

To define the optimal architecture of the ANN-based models, as recommended in [7], both trial and error approaches have been used to determine the number of hidden layers and neurons in each layer. Based on the findings, the best topology consists of two hidden layers and 12 neurons in each hidden layer.

5.1. Experimental framework

This section explains how the experiments have been carried out using both the ANN and the JADE4ANN predictive models. In the first step, data are normalized using the MinMax normalization technique. All the experiments were conducted using 80% of the dataset (i.e., 184 samples), which were randomly selected, as the training set. The remaining parameters of the dataset (i.e., 46 instances) were used as the testing set. Both the basic ANN and the evolutionary ANN models were built using the training set. Once the models have been generated, the testing set was used to evaluate the performance of the models using different measures.

Different performance parameters, including RMSE (Eq. 5), Scatter index (SI) (Eq. 6), MAE (Eq. 7), and R^2-value (Eq. 8) were used to evaluate the accuracy of the proposed algorithm.

$$SI = \frac{\sqrt{RMSE}}{y_i} \tag{6}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |y_i - \bar{y}_i| \tag{7}$$

$$R^2 = \frac{\left[\sum_{i=1}^{N} (y_i - y_{mean})^2\right] - \left[\sum_{i=1}^{N} (y_i - \bar{y}_i)^2\right]}{\left[\sum_{i=1}^{N} (y_i - y_{mean})^2\right]} \tag{8}$$

In the previous equations, N, y, and $\bar{y}$ are the number of samples, actual value, and predicted value, respectively.

5.2. Sensitivity analysis

In this section, two experiments were carried out to determine the optimal values of the population size and β. The population size is one of the most influencing parameters in evolutionary algorithms. In this experiment, four population sizes of 50, 100, 150, and 200 are considered. Fig. 4 shows the RMSE value of the JADE4ANN model when the number of chromosomes varies. It has been demonstrated that the best values are obtained when the population size are 100 and 150.

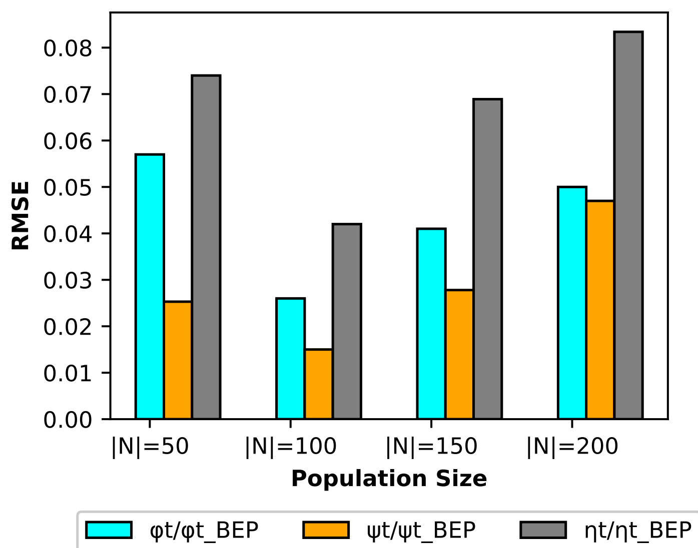


Figure 4: Performance of JADE4ANN model in predicting PaTs' performance based on different population sizes

Fig. 5 presents the impact of β that changes the PaTs' performance prediction when the value of β varies from 0.1 to 0.9. The best values are obtained for $\beta = 0.5$.

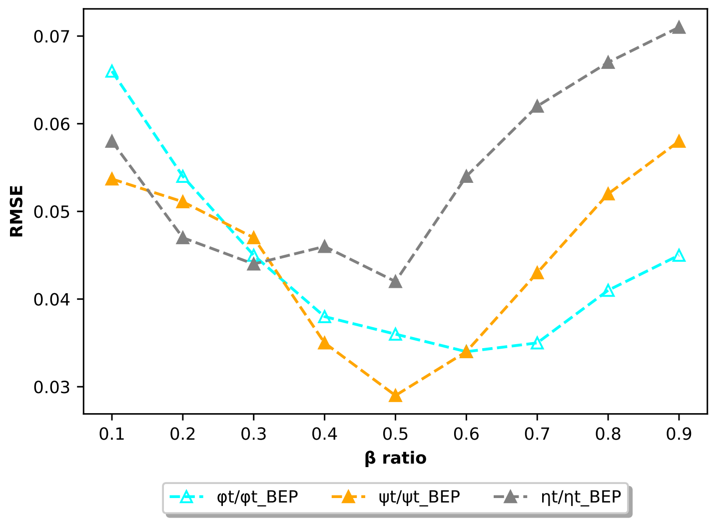


Figure 5: Performance of JADE4ANN model in predicting PaTs performance based on different β

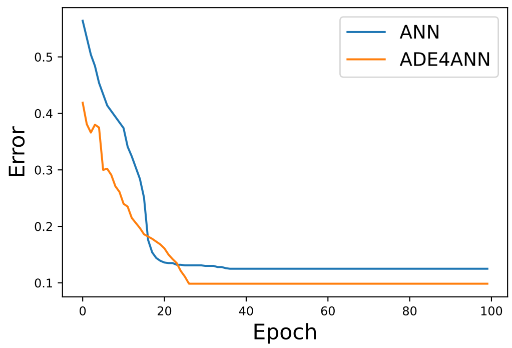
5.3. Performance curves of the PaTs in turbine mode

This section aims at carrying out some experiments to predict the non-dimensional parameters that correspond to the BEP operation of a PaT in turbine mode. Fig. 11 shows the average error obtained on the testing dataset in terms of the number of the epochs. This experiment has been performed with 100 epochs. As it can be noticed, the JADE4ANN model led to a lower error compared to the ANN and the regression models so that the average error for the JADE4ANN was about 0.03 and 0.1 lower than that of the ANN and the regression models, respectively.

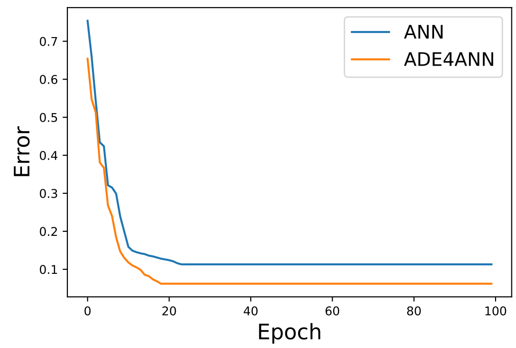
Table 2 shows the performance of the predictive models in terms of different measures, including RMSE, SI, MAE, and R^2-value. According to Table 2, the JADE4ANN algorithm led to a higher agreement between actual and estimated values, and the lowest error compared with other three predictors. The higher performance of the JADE4ANN model compared to the basic ANN has to be found in the capability of the evolutionary algorithms to optimize both hyper-parameters (i.e., learning rate) and connection weights (biases and weights) during the learning process according to a fitness function instead of by an expert's experiences and gradient descent algorithms. Thus, evolutionary algorithms are able to find the best parameters for different datasets and applications.

Table 2: Performance of the predictive models for different targets

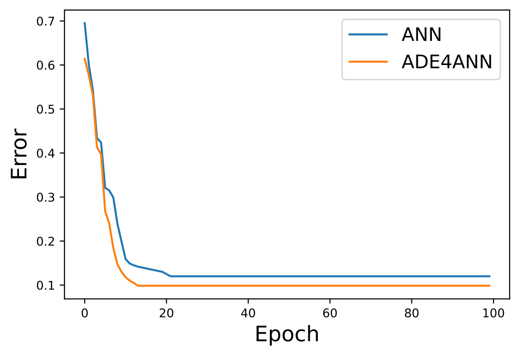
Model	RMSE		SI		MAE		R^2-value	
	Train	Test	Train	Test	Train	Test	Train	Test
Regression: ϕ_t/ϕ_{tBEP}	0.062	0.095	0.192	0.236	0.234	0.292	0.858	0.8
Regression: ψ_t/ψ_{tBEP}	0.036	0.038	0.19	0.28	0.177	0.183	0.879	0.882
Regression: η_t/η_{tBEP}	0.117	0.114	0.29	0.35	0.325	0.321	0.875	0.872
ANN of [7]: ϕ_t/ϕ_{tBEP}	0.049	0.057	0.27	0.285	0.205	0.222	0.896	0.889
ANN of [7]: ψ_t/ψ_{tBEP}	0.025	0.027	0.265	0.275	0.148	0.153	0.939	0.937
ANN of [7]: η_t/η_{tBEP}	0.074	0.078	0.36	0.284	0.253	0.216	0.945	0.938
JADE4ANN: ϕ_t/ϕ_{tBEP}	0.023	0.026	0.116	0.125	0.157	0.151	0.970	0.968
JADE4ANN: ψ_t/ψ_{tBEP}	0.016	0.015	0.16	0.18	0.120	0.115	0.975	0.979
JADE4ANN: η_t/η_{tBEP}	0.044	0.042	0.142	0.191	0.198	0.191	0.978	0.983
DE4ANN: ϕ_t/ϕ_{tBEP}	0.043	0.049	0.176	0.219	0.199	0.208	0.959	0.948
DE4ANN: ψ_t/ψ_{tBEP}	0.046	0.053	0.16	0.208	0.162	0.185	0.945	0.935
DE4ANN: η_t/η_{tBEP}	0.051	0.046	0.315	0.263	0.198	0.219	0.958	0.952
SaDE4ANN: ϕ_t/ϕ_{tBEP}	0.036	0.043	0.194	0.24	0.171	0.158	0.96	0.958
SaDE4ANN: ψ_t/ψ_{tBEP}	0.034	0.044	0.26	0.38	0.172	0.197	0.972	0.963
SaDE4ANN: η_t/η_{tBEP}	0.04	0.038	0.23	0.36	0.177	0.191	0.961	0.964



(a) ϕ_t/ϕ_{tBEP}



(b) ψ_t/ψ_{tBEP}

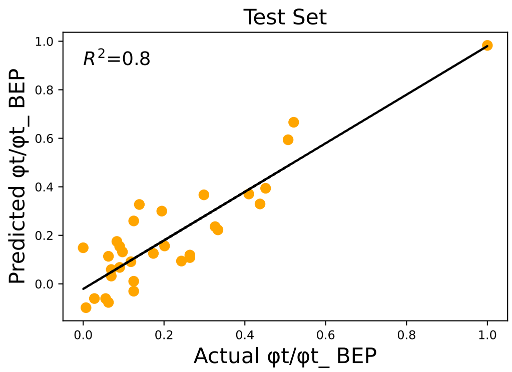


(c) η_t/η_{tBEP}

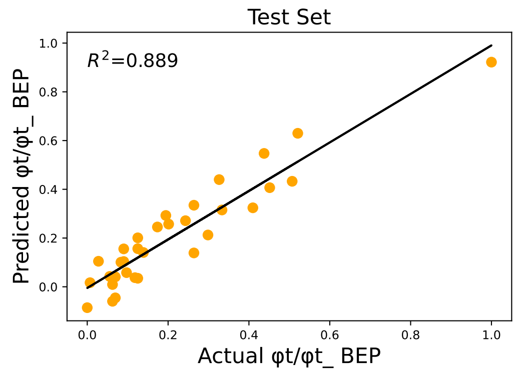
Figure 6: Average error in terms of the number of epochs.

Table 3 shows the average performance of the predictive models presented in Table 2. The JADE4ANN model led to the highest performance in all the analysed cases with an R^2-value of 0.977, followed by the SaDE4ANN, DE4ANN, and ANN of [7] models with 0.961, 0.945, and 0.921, respectively.

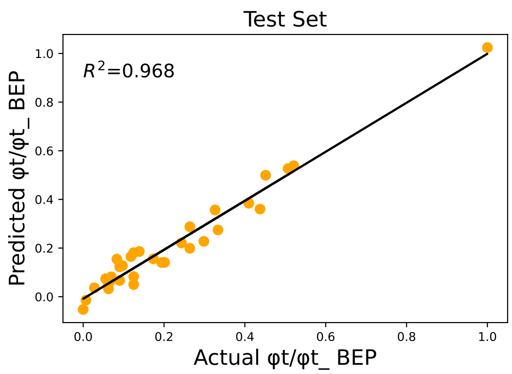
Figs. 7, 8 and 9 show the values of the performance curves in turbine mode in comparison to the values predicted by the predictors. According to these graphs, the JADE4ANN model with the R^2-value of 0.976 is, on average, the most reliable predictor to forecast PaTs' performance curves. Form these figures, it is demonstrated that the JADE4ANN model provides a relatively closer prediction as compared to the other AI models.



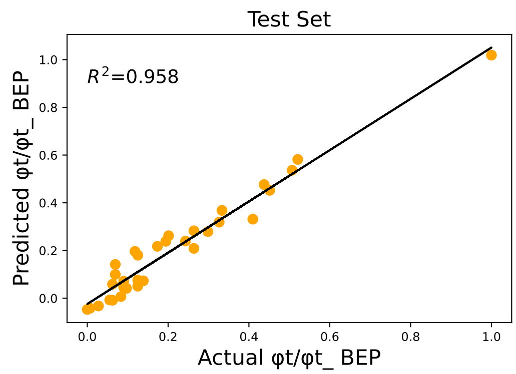
(a) Regression



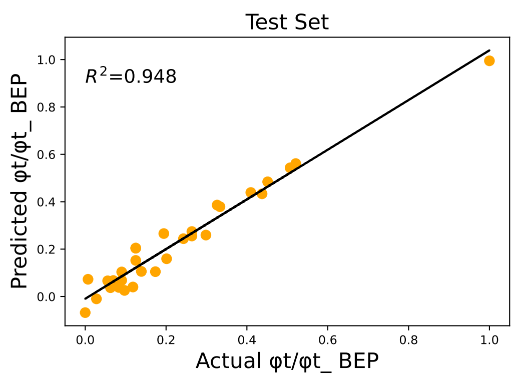
(b) ANN



(c) JADE4ANN

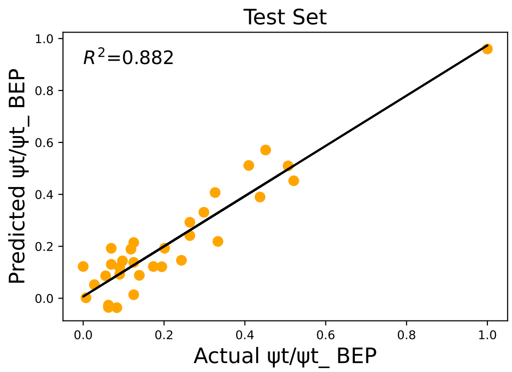


(d) SaDE4ANN

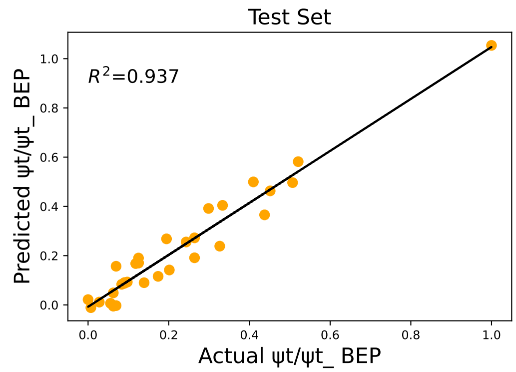


(e) DE4ANN

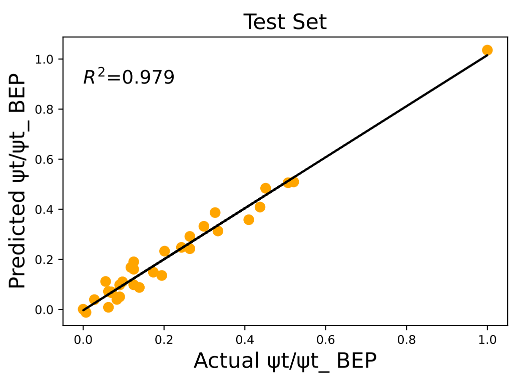
Figure 7: The performance of the predictive models in predicting ϕ_t/ϕ_{tBEP}



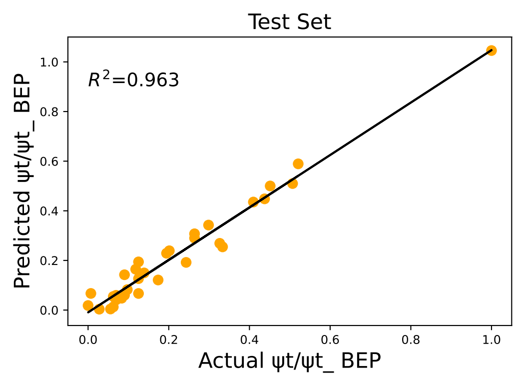
(a) Regression



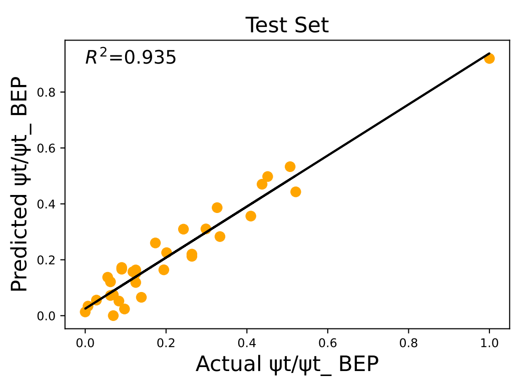
(b) ANN



(c) JADE4ANN

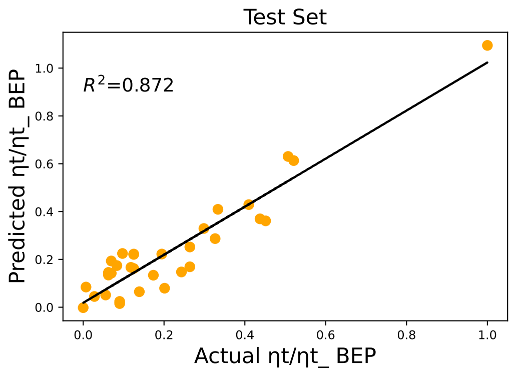


(d) SaDE4ANN

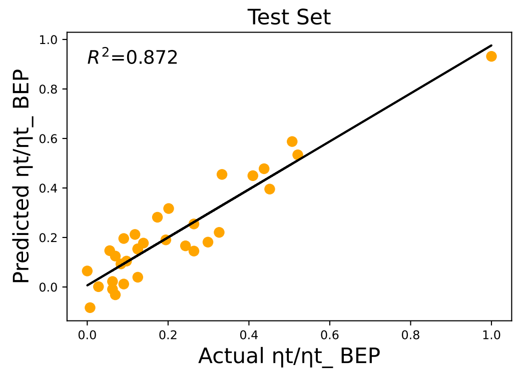


(e) DE4ANN

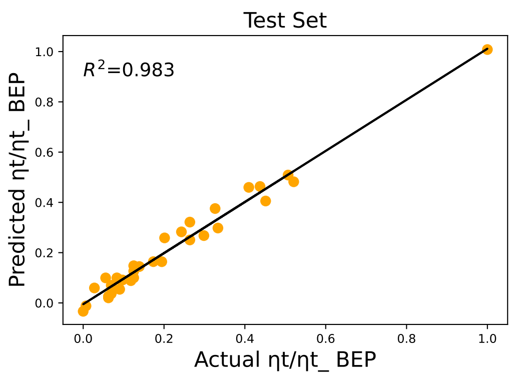
Figure 8: The performance of the predictive models in predicting ψ_t/ψ_{tBEP}



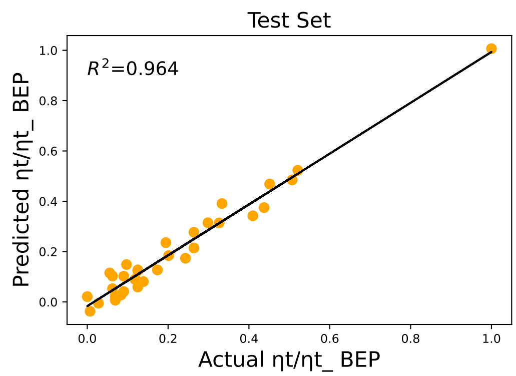
(a) Regression



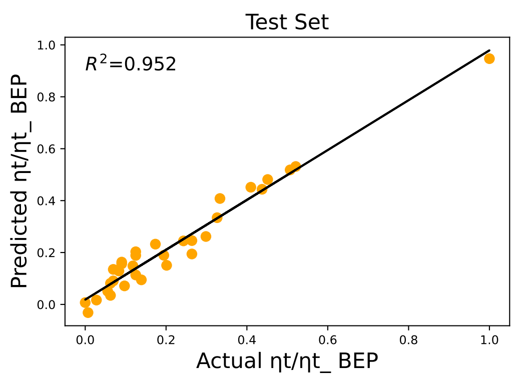
(b) ANN



(c) JADE4ANN



(d) SaDE4ANN



(e) DE4ANN

Figure 9: The performance of the predictive models in predicting η_t/η_{tBEP}

Table 3: Average performance of the predictive models

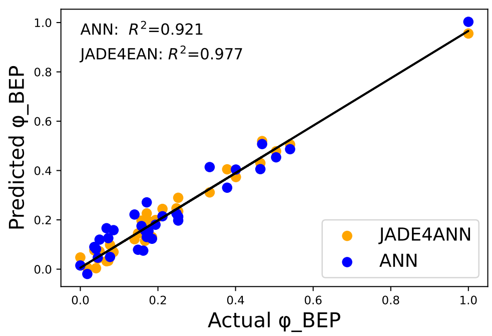
Model	RMSE		SI		MAE		R^2-value	
	Train	Test	Train	Test	Train	Test	Train	Test
Regression	0.0716	0.0823	0.336	0.288	0.245	0.265	0.87	0.85
ANN of [7]	0.049	0.054	0.298	0.281	0.202	0.197	0.926	0.921
JADE4ANN	**0.027**	**0.027**	**0.139**	**0.23**	**0.158**	**0.152**	**0.974**	**0.977**
DE4ANN	0.046	0.0493	0.217	0.23	0.186	0.204	0.954	0.945
SaDE4ANN	0.036	0.041	0.228	0.326	0.173	0.182	0.964	0.961

5.4. BEP prediction of PaTs in turbine mode

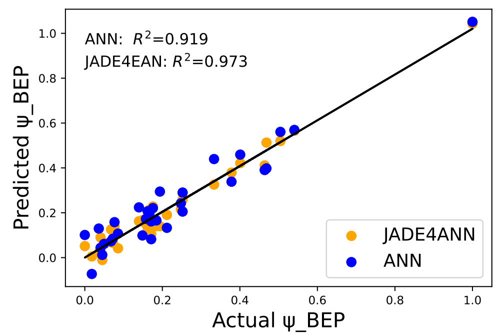
This section provides results obtained from intelligent systems (i.e., ANN of [7] and JADE4ANN models) in which the data of PaTs in pump mode feed the models for the performance prediction in turbine mode. Since manufacturers do not supply data of PaTs in turbine mode, their performance predictions can be useful to predict the behaviour of these machines in turbine mode. Fig. 10 presents the findings related to the experiments for the BEP prediction of PaTs in turbine mode. It can be noticed that the JADE4ANN model can provide more accurate predictions than the ANN one, so that the R^2-value value for the JADE4ANN is 0.975 against 0.928 of the ANN.

5.5. Comparison of the accuracy of PaTs prediction models

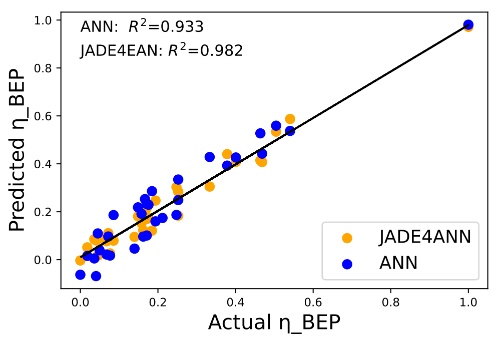
After the analysis performed so far, the accuracy of the presented models (i.e., JADE4ANN, DE4ANN, and SaDE4ANN) has been compared with the results obtained by the ANN model discussed in [7] and the optimization-based method proposed in [14]. Table 4 lists the values of the non-dimensional parameters related to the developed model here presented, the laboratory tests performed by Rossi et al. [7], and those obtained with the ANN of [7] and the optimization-based model of [14] along with the relative percentage errors with respect to the laboratory tests. As it can be noticed in Fig. 11, the developed DE-based ANN models can obtain lower error values than the ANN model of [7] and the optimization-based [14]; indeed, an average error of 1.35 has been obtained by the JADE4ANN against 3.93 of the ANN model [7]. The highest and the lowest errors were achieved in the evaluation



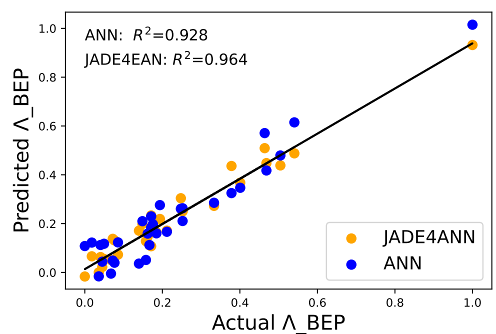
(a) Flow coefficient



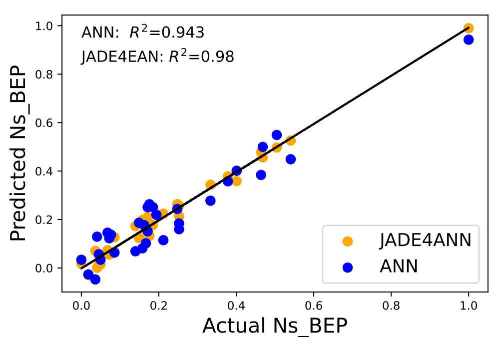
(b) Head coefficient



(c) Mechanical efficiency



(d) Power coefficient



(e) Specific speed

Figure 10: The performance of the predictive models in predicting the BEP in turbine mode

of the ϕ value (5.17%) and the ψ value (1.09%), respectively. Along the same line, Figs. 12 and 13 show the performance curves of the tested PaT in

Table 4: Average performance of the predictive models

Model	Laboratory tests of [7]	ANN of [7]	Optimization-Based [14]	JADE4ANN	DE4ANN	SaDE4ANN
ϕ_{tBEP}	0.0182	0.0191	0.0192	0.0180	0.0188	0.0179
ψ_{tBEP}	0.1559	0.1587	0.1632	0.1541	0.1577	0.1515
η_{tBEP}	0.7538	0.7176	0.7445	0.7665	0.7743	0.7311

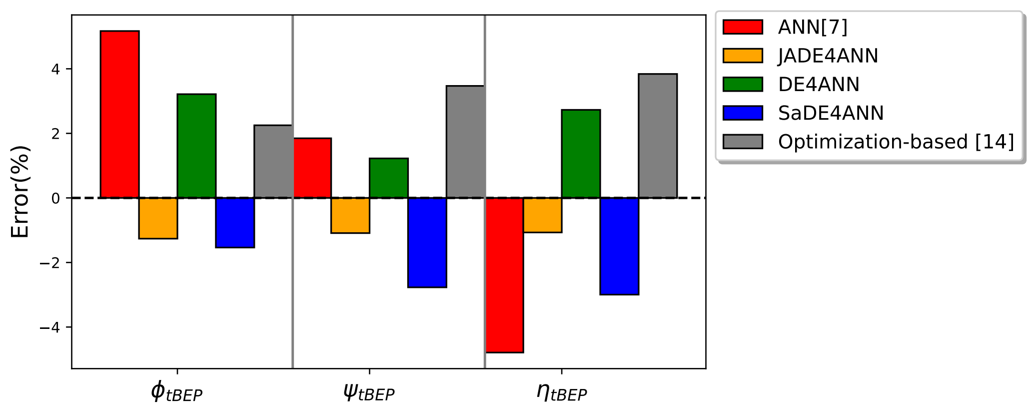


Figure 11: Comparison between the errors obtained by the present model and others available in the literature

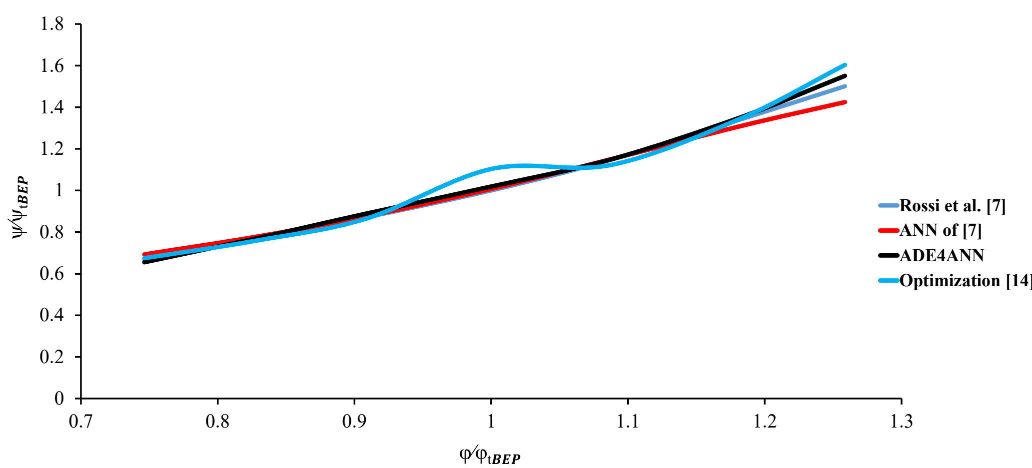


Figure 12: Non-dimensional characteristic curves comparison

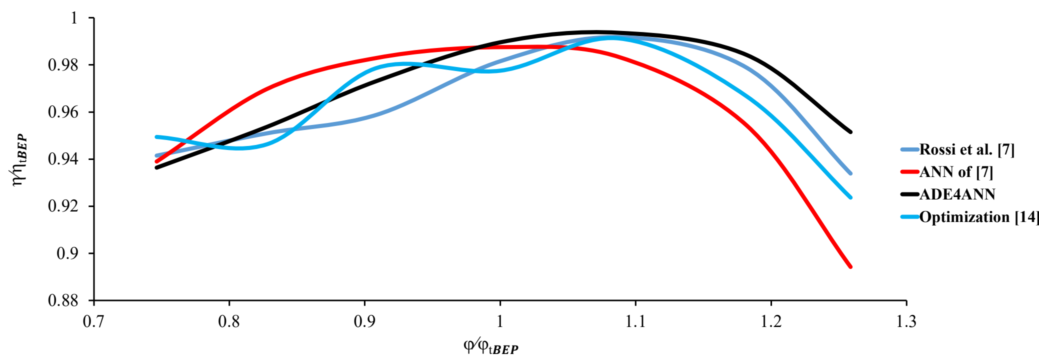


Figure 13: Efficiency curves comparison

[7] together with the predictions obtained with the ANN of [7], ADE4ANN, and optimization-based model [14]. Fig. 12 shows that the models can provide predictions with a lower error at strong part-load operating conditions of the PaT, while, increasing the flow rate, the prediction errors increases. However, the proposed model could yield more accurate predictions than the simple ANN model. Fig. 13 also demonstrates that the mechanical efficiency remains relatively high when the flow rate is between 80% and 120% of the BEP value, while the efficiency is reduced dramatically when operating with high flow rate. It is also worth to notice that the trends obtained with the methodology proposed in [14] present wavy trends, which are not realistic for the operation of a fluid machine. On the other hand the optimization model proposed in this work shows better prediction accuracy but also a more reliable trend.

6. Conclusions

In this study, an ANN algorithm based on the JADE algorithm has been developed to optimize all the design parameters of a basic ANN and minimize the prediction error of the PaTs' performance in turbine mode. The proposed model was used to predict both the BEP and performance curves of PaTs operating in turbine mode. Data of 32 different PaTs were collected from the literature [7] and they represent 230 operating samples in both pump and turbine modes with N_S values ranging between 0.28 and 2.24. The new JADE4ANN algorithm was compared to the basic ANN algorithm in terms of various criteria on a pump mode dataset. The JADE4ANN algorithm reduces the average errors by about 6%. The JADE4ANN achieves an approximate

R^2-value of 0.98 for both the BEP and performance curves predictions, while a standard ANN model based on the same data obtains an R^2-value of 0.92 as reported in [7]. In particular, the proposed model led to the best result in terms of non-dimensional magnitudes, achieving an error of 2.91% compared to the laboratory tests of a PaT performed by [7] and thus being 1.3% lower than the value obtained by the ANN of [7]. The same results have been obtained when comparing the performance curves predicted by both the ANN of [7] and the proposed ANN based on JADE; to be precise, 3.06% for the ANN against 2.26% for the JADE4ANN model has been obtained. The success of the JADE4ANN on forecasting PaTs' performance demonstrates that it can be applied to address the behaviour of other turbomachines and energy domains as well. As a next research development, the improvement of the performance parameters through their optimization with the *k*-fold cross validation will be considered, where an ensemble learning on different ANN models can be built.

References

[1] K. Solangi, M. Islam, R. Saidur, N. Rahim, H. Fayaz, A review on global solar energy policy, Renewable and sustainable energy reviews 15 (4) (2011) 2149–2163.

[2] O. Ogunmodede, K. Anderson, D. Cutler, A. Newman, Optimizing design and dispatch of a renewable energy system, Applied Energy 287 (2021) 116527.

[3] Y. Wu, T. Zhang, R. Gao, C. Wu, Portfolio planning of renewable energy with energy storage technologies for different applications from electricity grid, Applied Energy 287 (2021) 116562.

[4] K. Koponen, E. Le Net, Towards robust renewable energy investment decisions at the territorial level, Applied Energy 287 (2021) 116552.

[5] D. Guijo-Rubio, A. Durán-Rosal, P. Gutiérrez, A. Gómez-Orellana, C. Casanova-Mateo, J. Sanz-Justo, S. Salcedo-Sanz, C. Hervás-Martínez, Evolutionary artificial neural networks for accurate solar radiation prediction, Energy 210 (2020) 118374.

[6] M. Rossi, A. Nigro, M. Renzi, Experimental and numerical assessment of a methodology for performance prediction of pumps-as-turbines (pats) operating in off-design conditions, Applied Energy 248 (2019) 555–566.

[7] M. Rossi, M. Renzi, A general methodology for performance prediction of pumps-as-turbines using artificial neural networks, Renewable Energy 128 (2018) 265–274.

[8] F. Moazeni, J. Khazaei, Optimal energy management of water-energy networks via optimal placement of pumps-as-turbines and demand response through water storage tanks, Applied Energy 283 (2021) 116335.

[9] A. Morabito, P. Hendrick, Pump as turbine applied to micro energy storage and smart water grids: A case study, Applied energy 241 (2019) 567–579.

[10] M. Renzi, A. Nigro, M. Rossi, A methodology to forecast the main non-dimensional performance parameters of pumps-as-turbines (pats) operating at best efficiency point (bep), Renewable Energy 160 (2020) 16–25.

[11] S. Huang, G. Qiu, X. Su, J. Chen, W. Zou, Performance prediction of a centrifugal pump as turbine using rotor-volute matching principle, Renewable Energy 108 (2017) 64–71.

[12] D. Novara, A. McNabola, A model for the extrapolation of the characteristic curves of pumps as turbines from a datum best efficiency point, Energy Conversion and Management 174 (2018) 1–7.

[13] M. Stefanizzi, T. Capurso, G. Balacco, M. Binetti, S. M. Camporeale, M. Torresi, Selection, control and techno-economic feasibility of pumps as turbines in water distribution networks, Renewable Energy 162 (2020) 1292–1306.

[14] D. Mitrovic, J. G. Morillo, J. A. Rodríguez Díaz, A. Mc Nabola, Optimization-based methodology for selection of pump-as-turbine in water distribution networks: Effects of different objectives and machine operation limits on best efficiency point, Journal of Water Resources Planning and Management 147 (5) (2021) 04021019.

[15] M. Renzi, P. Rudolf, D. Štefan, A. Nigro, M. Rossi, Installation of an axial pump-as-turbine (pat) in a wastewater sewer of an oil refinery: A case study, Applied Energy 250 (2019) 665–676.

[16] L. Cornejo-Bueno, C. Camacho-Gómez, A. Aybar-Ruíz, L. Prieto, A. Barea-Ropero, S. Salcedo-Sanz, Wind power ramp event detection with a hybrid neuro-evolutionary approach, Neural Computing and Applications 32 (2) (2020) 391–402.

[17] P. Singh, P. Dwivedi, Integration of new evolutionary approach with artificial neural network for solving short term load forecast problem, Applied energy 217 (2018) 537–549.

[18] Y. Golizadeh Akhlaghi, K. Aslansefat, X. Zhao, S. Sadati, A. Badiei, X. Xiao, S. Shittu, Y. Fan, X. Ma, et al., Hourly performance forecast of a dew point cooler using explainable artificial intelligence and evolutionary optimisations by 2050, Applied Energy 281 (C) (2021).

[19] J. Zhang, A. C. Sanderson, Jade: adaptive differential evolution with optional external archive, IEEE Transactions on evolutionary computation 13 (5) (2009) 945–958.

[20] M. Crespo Chacón, J. A. Rodríguez Díaz, J. García Morillo, A. McNabola, Evaluation of the design and performance of a micro hydropower plant in a pressurised irrigation network: Real world application at farm-level in southern spain, Renewable Energy 169 (2021) 1106–1120.

[21] S. Barbarelli, M. Amelio, G. Florio, Experimental activity at test rig validating correlations to select pumps running as turbines in microhydro plants, Energy Conversion and Management 149 (2017) 781–797.

[22] D. Novara, A. Carravetta, A. McNabola, H. M. Ramos, Cost model for pumps as turbines in run-of-river and in-pipe microhydropower applications, Journal of Water Resources Planning and Management 145 (5) (2019).

[23] S. Spedaletti, M. Rossi, G. Comodi, D. Salvi, M. Renzi, Energy recovery in gravity adduction pipelines of a water supply system (wss) for urban areas using pumps-as-turbines (pats), Sustainable Energy Technologies and Assessments 45 (2021) 101040.

[24] A. Bekker, M. Van Dijk, C. Niebuhr, A review of low head hydropower at wastewater treatment works and development of an evaluation framework for south africa, Renewable and Sustainable Energy Reviews 159 (2022) 112216.

[25] J. Delgado, L. Andolfatto, D. Covas, F. Avellan, Hill chart modelling using the hermite polynomial chaos expansion for the performance prediction of pumps running as turbines, Energy Conversion and Management 187 (2019) 578–592.

[26] C. Alatorre-Frenk, Cost minimisation in micro-hydro systems using pumps-as-turbines, Ph.D. thesis, University of Warwick (1994).

[27] M. Venturini, L. Manservigi, S. Alvisi, S. Simani, Development of a physics-based model to predict the performance of pumps as turbines, Applied Energy 231 (2018) 343–354.

[28] R. Storn, K. Price, Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces, Journal of global optimization 11 (4) (1997) 341–359.

[29] A. K. Qin, P. N. Suganthan, Self-adaptive differential evolution algorithm for numerical optimization, in: 2005 IEEE congress on evolutionary computation, Vol. 2, IEEE, 2005, pp. 1785–1791.

[30] A. Telikani, A. H. Gandomi, A. Shahbahrami, A survey of evolutionary computation for association rule mining, Information Sciences 524 (2020) 318–352.

Appendix A. Artificial Neural Networks (ANN)

The origin of the ANN backs to 1943 when the first computational model based on neural networks was proposed. In this model, the sensory input was represented by the neuronal activation. In this network, different layers of features were extracted, in which their combinations in the previous layer make the features in the next layer. A standard neural network includes many connected processors called neurons. Input neurons receive values from the environment, while other neurons receive activated values by weighted connections from the previously active neurons. Different generations of neural networks have been proposed so far. The perception was the first generation, while the second generation was introduced by using more hidden layers in the topology of the network and to back-propagate the error for learning.

Generally, an ANN consists of three main components: i) architecture, ii) transfer function, and ii) optimizer. A topology of an ANN includes three layers: (1) input layer, which is usually equal to the number of variables in the dataset, (2) one or more hidden layers, and (3) an output layer that can consist of one or multiple neurons depending on prediction or classification tasks. The features measured per each training instance correspond to the input layer. Then, the input passes through the hidden layer(s). The output layer receives the weighted outputs from the last hidden layer to make prediction/classification of given samples. The procedure to compute the output of each neuron in the hidden and output layers is shown in Fig. A.14. Firstly, the input is computed (Eq. A.1) by multiplying the corresponding weight (i.e., w_{ij}) by the output of the unit from the previous layer (i.e., O_i). Then, the summed value is added by the bias value (i.e., θ_j). Finally, an activation function (e.g., sigmoid) is applied to use the output for the neuron.

$$I_i = \sum_i w_{ij} O_i + \theta_j \tag{A.1}$$

In the supervised learning process of the weights, the main focus is to find an absolute optimal or close-enough optimal set of connection weights for a fixed size network, in terms of the number of layers and neurons at each layer, using back-propagation. In the field of neural networks, the dominant method for training neural networks is back-propagation, which is an efficient algorithm for calculating the loss function gradient. Each neural network weight can be modified by the back-propagation to greedily reduce loss. A

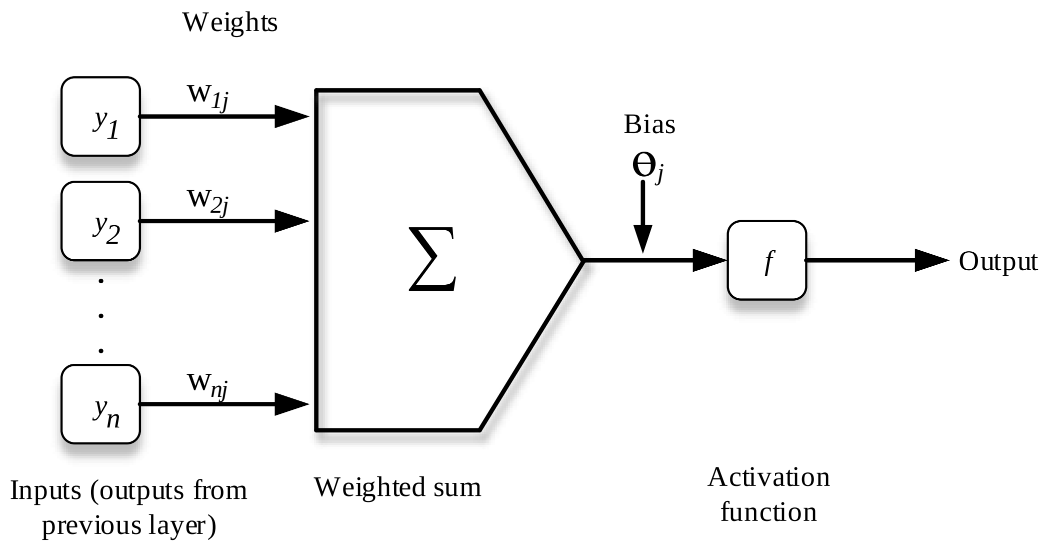


Figure A.14: Procedure to compute the weighted output for each neuron

back-propagation algorithm, represented in 1, is a generalization of the delta rule for training neural networks that updates the weights of the network at several successive iterations by minimizing the cost function of the error.

Algorithm 1 Back Propagation

Input: Data $D = \{(x_k, y_k)\}_{k=1}^{n}$, learning rate η, cost function E
Output: optimal weights w
Randomly initialize all weights and threshold
while *Stopping criteria is not met* **do**
    for *all* $(x^{(i)}, y^{(i)} \in D$ **do**
        Compute $w_{ij} = w_{ij} - \eta \partial E / \partial w_{ij}$
    end
end

Appendix B. JADE: An Adaptive Differential Evolution

Evolutionary approaches have been inspired by natural evolution principles. These approaches encode a problem in terms of individuals to be evolved with the aim of improving the quality of solutions. Evolutionary algorithms explore the search space using an iterative heuristic procedure to obtain gradually better solutions. The JADE [19] is a new evolutionary algorithm that is an improvement of Differential Evolution (DE) algorithm [28]. The DE uses the three genetic operators mutation (Eq. B.1), crossover (Eq. B.2), and selection (Eq. B.3) for evolution [30].

$$v_i^G = x_{S1}^G + F.(x_{S2}^G - x_{S3}^G) \tag{B.1}$$

$$u_{i,j}^G = \begin{cases} x_{i,j}^G & if \quad rand_{i,j}[0,1] \leq CR \\ \\ v_{i,j}^G & otherwise \end{cases} \tag{B.2}$$

$$u_i^{G+1} = \begin{cases} u_i^G & if \quad f(u_i^G) < f(x_i^G) \\ \\ x_i^G & otherwise \end{cases} \tag{B.3}$$

In the DE, these three operators of mutation, crossover, and selection are performed in a row. In each generation i, the mutation operator first generates a mutation vector (v_i^G) from three randomly selected individuals x_{S1}^G, x_{S2}^G, and x_{S3}^G from the population (Eq. B.1). The notation $F \in$[0,1] is a parameter that determines the exploration length ($x_{S2} - x_{S3}$). Then, the crossover operator generates the trial individual (u_i^G) by crossing the target individual (x_i^G) with its mutant counterpart (v_i^G) (Eq. B.2). The term j is a random integer between 1 and the length of an individual. The CR variable is the crossover probability in the range of [0,1]. Finally, the target individual (x_i^G) is replaced by (u_i^G) if its fitness value is higher than u_i^G (in our problem). Otherwise, the target individual is reserved for the next generation (Eq. B.3).

The first difference between the JADE and the DE is that the JADE uses "DE/current-to-pbest/1" for mutation (Eq.B.4), unlike the DE, which uses "DE/rand/1" (Eq. B.1). The second difference is that at each generation step of the JADE, the control parameters are automatically optimized without knowing the parameter settings beforehand.

$$v_i^G = x_i^G + F.(x_{best}^G - x_i^G) + F.(x_{S1}^G - x_{S2}^G) \tag{B.4}$$