A Review of the Application of Neural Networks and Global Optimization Algorithms in Hydroturbine Optimization

Zihang Wang *

School of Energy and Power Engineering, Xi'an Jiaotong University, Xi’an, China
*Corresponding author: Zihang Wang

Abstract: Turbine optimization is a crucial step in enhancing the efficiency and stability of hydropower generation. In recent years, significant advancements have been made in the application of artificial neural network (ANN) and optimization algorithms in this field. This paper reviews the applications of neural networks in turbine performance modeling, fault detection, and real-time control, highlighting their advantages in handling complex nonlinear systems and big data analysis. Additionally, it introduces the practical implementations of optimization algorithms such as genetic algorithms, particle swarm optimization, and simulated annealing in turbine parameter optimization, structural design, and operation scheduling, emphasizing their effectiveness in improving efficiency, reducing energy consumption, and extending equipment lifespan. By combining the predictive capabilities of neural networks with the problem-solving strengths of optimization algorithms, new optimization strategies are proposed, providing important research directions and technical support for future turbine optimization.

Keywords: artificial neural network; optimization algorithms; hydroturbine

1. Artificial Neural Network

Artificial Neural Network (ANN) is a computational model that simulates the biological neural system, achieving the learning and recognition of complex data patterns through a hierarchical structure of input layers, one or more hidden layers, and an output layer. Its main features include adjustable weighted connections between nodes, handling of nonlinear relationships, utilization of activation functions to determine output values, and optimization of weights through algorithms like backpropagation. ANN possesses powerful parallel processing capabilities and adaptive learning characteristics, widely applied in fields such as image recognition, speech recognition, natural language processing, autonomous driving, and financial forecasting. Its development benefits from advancements in big data and high-performance computing resources, demonstrating outstanding performance in many tasks.

Artificial Neural Networks can be classified into various types based on their structure and purpose, including feedforward neural networks, convolutional neural networks, recurrent neural networks, and generative adversarial networks. According to learning methods, they are categorized into supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning neural networks. Based on their purposes, they can be classified into classification networks, regression networks, generative networks, and sequence-to-sequence networks. Additionally, there are graph neural networks and self-attention networks. These classifications can be intersected to adapt to different application scenarios.

Since the advent of neural networks, hundreds of types have emerged, with more than a dozen being well-known. Based on their topological structure and learning algorithms, neural networks can be divided into four categories: feedforward networks, competitive networks, feedback networks, and stochastic networks. Feedforward networks include MP model, perceptron neural network, adaptive linear neural network, BP neural network, and radial basis function neural network; Competitive networks include self-organizing competitive neural network, self-organizing feature mapping neural network, and backpropagation neural network; Feedback networks include adaptive resonance theory neural network, learning vector quantization neural network, Elman neural network, and Hopfield neural network; Stochastic networks include Boltzmann neural network. These neural network models have been widely applied and developed in various fields, like hydroturbine and wind turbine [1].

In 1943, McCulloch and Pitts first completed the construction of the artificial neural network model [2], this type of model was first successfully applied to optimization problems.

In 1986, D.E. Rumelhart and J.L. McClelland proposed a neural network model utilizing error backpropagation training algorithm, commonly referred to as the BP (Back propagation) network. It is a type of multi-layer feedforward network with hidden layers, systematically addressing the learning problem of hidden unit connection weights in multilayer networks [3]. The basic principle of the BP learning algorithm is gradient descent, with its central idea being to adjust weights to minimize the total error of the network. The network learning process involves propagating errors backward while simultaneously adjusting the weight coefficients.

In 1985, Powell introduced the radial basis function (RBF) method for multivariate interpolation. Subsequently, in 1988, Broomhead and Lowe were the first to apply RBF to neural network design, thus forming the RBF neural network [4]. The structure of the RBF network is similar to a multi-layer feedforward network, but it is a two-layer feedforward network with a single hidden layer. The input layer consists of nodes representing the signal sources, while the number of units in the hidden layer depends on the requirements of the problem being addressed. The output layer responds to the input. The radial basis function network is a type of local approximation network capable of approximating any continuous function with arbitrary precision. When there are many training vectors, the network is highly effective [5].

Generally, artificial neural networks consist of an input layer, one or more hidden layers, and an output layer, as shown in the Figure 1. To improve the fitting accuracy of artificial neural networks, the number of hidden layers can be increased. Compared to response surface methods, although constructing artificial neural networks requires much more time and larger datasets to achieve the desired prediction accuracy, their advantages over other surrogate models are more prominent. Firstly, they offer higher fitting accuracy for complex nonlinear relationships. Secondly, the modeling and training process is a "black box" process, offering greater flexibility and applicability to different fields. Additionally, artificial neural networks possess strong generalization capabilities and are more tolerant to sample space data, effectively mitigating the impact of "data noise" on the prediction accuracy of the network model.

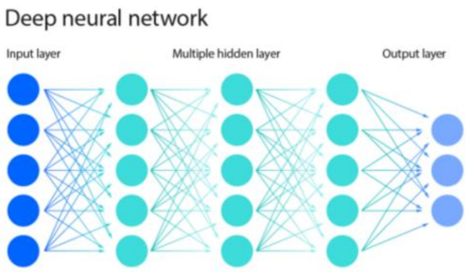


Figure 1. Structure of artificial neural network.

Chen proposes a long short-term memory artificial neural network (LSTMNN) with output feedback based on actual operational data to achieve real-time dynamic modeling of hydroturbines, which is shown in Figure 2. Subsequently, a feedback-based hydro turbine LSTMNN model is developed. Finally, the proposed modeling method is compared with a standard backpropagation neural network (F-BPNN) with output feedback to validate its effectiveness and applicability. The hydroturbine F-LSTMNN model is shown in Figure 3. The results indicate that by introducing output feedback, the accuracy of the nonlinear hydro turbine model can be improved, and the proposed modeling method outperforms F-BPNN [6].

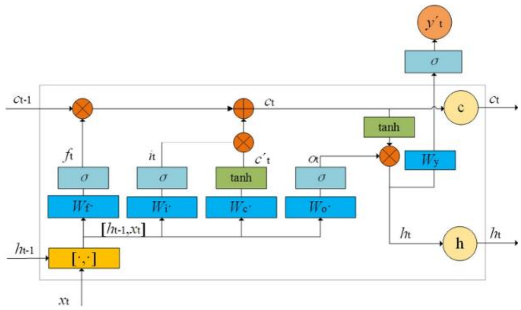
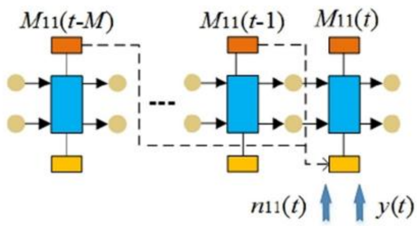
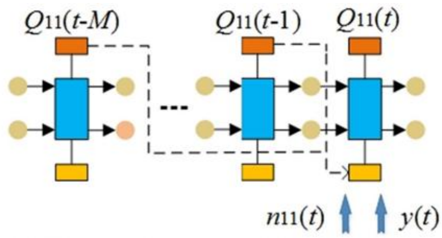


Figure 2. Long short-term memory unit structure [6].



(a) Hydro-turbine F-LSTMNN torque characteristic model.



(b) Hydro-turbine F-LSTMNN discharge characteristic model.

Figure 3. Hydroturbine F-LSTMNN model [6].

Wang introduces a novel deep learning method called Neural Controlled Differential Equations for modeling the unknown nonlinear dynamics of controlled continuous-time systems such as Francis turbines in hydraulic power systems. To address the overfitting issue during online training, meta-learning techniques are applied to pre-train meta-initial values for each parameter of the proposed NCDEs. The structure of f_{NN} is shown in Figure 4. Results demonstrate that utilizing meta-learning techniques can significantly reduce the prediction mean square error by over 60% [7].

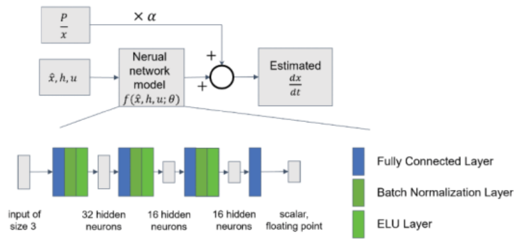


Figure 4. The structure of f_{NN} [7].

2. Optimization Algorithms

Currently, commonly used optimization algorithms can be roughly categorized into three main types: gradient-based optimization algorithms, direct search methods, and global optimization algorithms. Gradient-based optimization algorithms include general gradient descent methods and sequential quadratic programming methods. They can achieve rapid optimization in continuous and unimodal design spaces. However, this optimization method overly relies on gradient computation and is prone to getting trapped in local optima in design spaces with multiple peaks. Direct search methods no longer require gradient computation and use larger step sizes, enabling exploration of larger design spaces. However, they require solving a large number of cases, making them unsuitable for optimization problems with a large number of design variables or lengthy numerical computations or experimental tests. Similarly, there is also a possibility of getting trapped in local optima with this method. On the other hand, global optimization algorithms effectively address this issue. They do not rely on gradient computation, are generally applicable, and have stronger global search capabilities, enabling the search for global optimal solutions. However, these algorithms typically involve higher computational costs. Global optimization algorithms mainly include genetic algorithms, particle swarm optimization algorithms, and simulated annealing algorithms.

2.1. Genetic Algorithms

Genetic algorithms are stochastic adaptive search algorithms that simulate the mechanism of life evolution, developed based on Darwin's theory of biological evolution and Mendel's theory of genetics. They are an important branch of artificial intelligence [8]. The optimization process of genetic algorithms begins with a population of data points within the domain of the function, rather than a single data point. This characteristic indicates that GA is a method for global search. In fact, they can simultaneously approach any peak, reducing the probability of finding only local minima, which is a limitation of traditional optimization methods. The global nature, parallelism, speed, and robustness of GA algorithms have led to their widespread application in many fields, including function optimization, automatic control, image recognition, machine learning, artificial neural networks, and optimization scheduling. Currently, in many international conferences, fuzzy logic, neural networks, and genetic algorithms are collectively referred to as soft computing and are discussed as an important topic, indicating the tremendous development potential of the research field of GA [9].

Wang proposed a mathematical optimization model of the blade declination angle variation rule, utilizing the change in blade declination angle with position angle as the design parameter and turbine energy utilization rate as the objective function. Genetic algorithm and sequential quadratic programming method were combined for optimization calculations, resulting in the most suitable blade declination angle variation rule. Blade declination optimization was performed for cycloid turbines at different speed ratios, yielding optimized declination rules. The SQP-GA algorithm flowchart is shown in Figure 5. Results indicate that under the declination rule optimized by the hybrid genetic algorithm, the energy utilization rate of the hydro-turbine is improved compared to the original cycloid turbine [10].

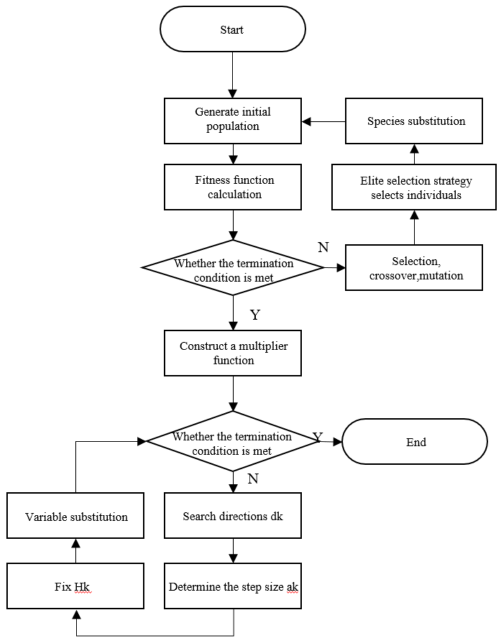


Figure 5. SQP-GA algorithm flow chart [10].

Bukhtoyarov proposed a method based on nonparametric Nadaraya-Watson kernel estimation for constructing regression models. On a test set, the methods of adjusting coefficients using standard genetic algorithms and hybrid genetic algorithms reduced modeling errors by 20% and 28%, respectively, compared to arbitrary selection. The proposed nonparametric method was found to reduce modeling errors by approximately 5% [11].

2.2. Particle Swarm Optimization Algorithm

Inspired by the foraging behavior of birds, Eberhart and Kennedy first proposed the Particle Swarm Optimization algorithm in 1995. In the PSO algorithm, each member of the swarm is treated as a particle representing a potential solution to the optimization problem. Each particle is assigned a velocity and position vector during the optimization process. The velocity vector of a particle is influenced by both its own search experience and the collective search experience of other particles in the swarm. Through the interaction of these two search experiences, particles adjust their search directions in real-time, gradually moving towards the global optimum solution of the optimization problem. In the PSO algorithm, the search experience of an individual particle is referred to as its personal cognitive experience, while the search experience of other particles in the swarm is known as the collective social experience. Due to its simplicity, ease of implementation, and fast search speed, the PSO algorithm has been widely used since its inception to solve various types of optimization problems, including trajectory planning, task allocation, and resource allocation problems, among others.

2.3. Simulated Annealing Algorithm

The Simulated Annealing algorithm is derived from the principle of solid annealing and is a probabilistic optimization algorithm. This algorithm mimics the process of heating a solid to a high temperature and then gradually cooling it down: when heated, the particles inside the solid become disordered, internal energy increases, and molecules and atoms are in an unstable state. During the cooling process, the particles gradually become ordered, energy decreases, and atoms tend to stabilize, ultimately reaching a ground state at room temperature where the internal energy is minimized. The Simulated Annealing algorithm starts from a relatively high initial temperature and, as the temperature continuously decreases, it combines probabilistic jumps to randomly search for the global optimum of the objective function within the solution space. Specifically, it uses probabilistic jumps with a time-varying probability that eventually approaches zero, effectively avoiding local minima and ultimately achieving global optimization. This characteristic enables the Simulated Annealing algorithm to be widely used in solving complex optimization problems such as the traveling salesman problem, scheduling problems, and network design problems.

Xiao and his team addressing the nonlinear characteristics of hydro-turbine generator set failure symptoms and types as well as the shortcomings of traditional wavelet neural network learning methods, designed a fault diagnosis model based on a simulated annealing algorithm wavelet neural network. This model was applied to hydro-turbine fault diagnosis. The results showed that, compared to wavelet neural networks and additional momentum BP neural networks, the designed model exhibited higher convergence accuracy and faster convergence speed. The simulated annealing algorithm wavelet neural network can be effectively applied to hydro-turbine fault diagnosis, providing a new approach for diagnosing hydro-turbine faults [12]. The structure of simulated annealing-wavelet neural network is shown in Figure 6.

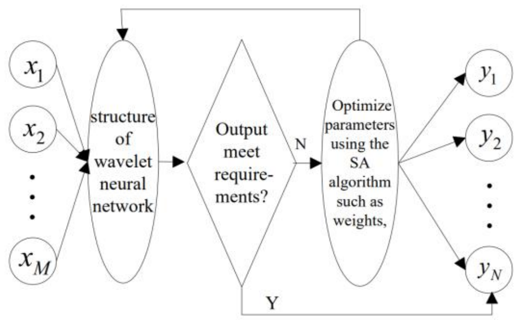


Figure 6. Structure of simulated annealing-wavelet neural network [12].

Zhou established an optimization model to address the challenging issue of coordinating the control of water hammer pressure and unit speed increase in hydropower stations. The Simulated Annealing algorithm was applied to this optimization model to find an optimal closing pattern that meets the regulatory requirements. The optimized closing pattern was then used for transition process calculations, all of which met the computational requirements and provided a substantial safety margin. The results indicate that the Simulated Annealing algorithm can quickly optimize the guide vane closing pattern and demonstrates good algorithm stability [13].

3. Conclusions

Neural networks and global optimization algorithms play a significant role in hydroturbine optimization. Neural networks utilize their powerful nonlinear modeling capabilities for performance prediction, design optimization, and intelligent control, thereby improving the operational efficiency and stability of hydroturbines. Global optimization algorithms, with their global search capabilities and multi-objective optimization features, address the optimization of complex design parameters, ensuring optimal performance under various operating conditions. The combination of these two technologies provides robust technical support for the design and operation of hydroturbines.

In the future, with the further development of deep learning and advanced optimization algorithms, neural networks are expected to train on larger datasets, enhancing prediction accuracy and generalization capabilities. Additionally, combined with IoT technology, the real-time monitoring and adaptive control of hydroturbines will become more intelligent and efficient. Global optimization algorithms will also deeply integrate with big data analysis and machine learning technologies, further improving optimization efficiency and solution precision. Moreover, new intelligent algorithms such as quantum computing and reinforcement learning are expected to play important roles in hydroturbine optimization, overcoming the limitations of traditional computing power and achieving faster and more efficient global optimization.

References

[1] Chen Y., Guo P., & Zhang D. Power improvement of a cluster of three Savonius wind turbines using the variable-speed control method. Renewable Energy, 2022; 193, 832-842.

[2] McCulloch, W. S., & Pitts, W. A logical calculus of the ideas immanent in nervous activity. The Bulletin of Mathematical Biophysics, 1943; 5(4), 115-133.

[3] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. Learning representations by back-propagating errors. Nature, 1986; 323(6088), 533-536.

[4] Broomhead, D., Lowe, D., & Casadevall, A. Multivariable functional interpolation and adaptive networks, Crisis in infectious diseases-time for a new paradigm. Complex Syst, 1996; 2, 312-355.

[5] Han, M., & Xi, J. Efficient clustering of radial basis perceptron neural network for pattern recognition. Pattern Recognition, 2004; 37(10), 2059-2067.

[6] Chen, J., Xiao, Z., Liu, D., Hu, X., Ren, G., & Zhang, H. Nonlinear modeling of hydroturbine dynamic characteristics using LSTM neural network with feedback. Energy Science & Engineering, 2021; 9(11), 1961-1972.

[7] Wang, H., Yin, Z., & Jiang, Z.-P. Real-time hybrid modeling of francis hydroturbine dynamics via a

neural controlled differential equation approach. IEEE ACCESS, 2023; 11, 139133-139146.
[8] Srinivas, M., & Patnaik, L. M. Genetic algorithms: A survey. Computer, 1994; 27(6), 17-26.
[9] Bala, J., & Wechsler, H. Shape analysis using hybrid learning. Pattern Recognition, 1996; 29(8), 1323–1333.
[10] Wang, Y., & Qian, Y. Optimization of blade deflection based on hybrid genetic algorithm-all database. Journal of Mechanical Strength, 2021; 43(4), 849-855.
[11] Bukhtoyarov, V. V., & Tynchenko, V. S. Design of Computational Models for Hydroturbine Units Based on a Nonparametric Regression Approach with Adaptation by Evolutionary Algorithms. Computation, 2021; 9(8), 83.
[12] Xiao, Z., Sun, Z., Song, L., Zhang, X., & Malik, O. P. Simulated annealing-wavelet neural network for vibration fault diagnosis of hydro-turbine generating unit. Journal of Optoelectronics and Advanced Materials, 2015; 17(5–6), 734–740.
[13] Zhou, T., &Zhang, J. Optimizing closure law of wicket gates in hydraulic turbine based on simulated annealing-All Databases. J Drain Irrig Mach Eng, 2018; 36(4), 320-326.