



A Rapid Design Method for Centrifugal Pump Impellers Based on Machine Learning

Y. Chen, W. Li[†], Y. Luo, L. Ji, S. Li, and Y. Long

Jiangsu University, National Research Center of Pumps, Zhenjiang, JiangSu, 212013, China

[†]Corresponding Author Email: lwjiangda@ujs.edu.cn

ABSTRACT

Centrifugal pumps are widely used across various industries, and the design of high-efficiency centrifugal pumps is essential for energy savings and emission reductions. The development of centrifugal pump models primarily uses an iterative design approach combining direct and inverse problem-solving based on one-dimensional flow theory. However, this semi-empirical, semi-theoretical design process is time-consuming and costly. To reduce development time and costs, this paper proposes a rapid impeller design method focused on hydraulic performance, integrating traditional similarity design theory with machine learning. The proposed model uses neural networks to predict empirical coefficients, determine key dimensions such as the impeller's inlet diameter, outlet diameter, outlet width, and axial distance. Once these parameters are defined, the main dimensions of the impeller can be calculated. The blade profile is defined using a 5-point B'ezier curve. Variations in the cross-sectional area of the flow passage influence the internal flow state of the centrifugal pump, ultimately impacting its hydraulic efficiency. A genetic algorithm, guided by variations in the cross-sectional area of the flow passage, optimizes the blade profile, achieving an improved impeller flow path and completing the rapid design of the centrifuge. This method significantly shortens the development cycle and lowers design costs, making it a promising technique for future impeller designs.

Article History

*Received November 24, 2024
Revised February 11, 2025
Accepted February 13, 2025
Available online May 5, 2025*

Keywords:

*Centrifugal pump
Machine learning
Impeller design
Cross-sectional area
Neural network*

1. INTRODUCTION

A pump is a commonly used device for fluid transportation, utilizing an impeller to convert the motor's mechanical energy into the energy of the fluid medium in a conveying pipeline. Centrifugal pumps have a simple structure and are adaptable to diverse working conditions, making them the most widely used machinery in fields such as water conservancy, automobiles, aerospace, petroleum and energy (Wilson et al., 2006). According to relevant studies, centrifugal pumps consume approximately 10% of the total annual electricity (Li, et al. 2024). Centrifugal pumps are driven by the impeller, their core component, meaning that improving impeller efficiency directly reduces energy consumption (Tuzson, 2000; Zhou, et al. 2022). Designing high-efficiency centrifugal pumps is crucial for energy conservation and emission reduction. The optimization of centrifugal pump designs typically involves a complex cycle of design, simulation, testing, and iterative modifications based on test results. Recent studies propose innovative approaches

to optimize centrifugal pump designs. Ju et al. (2021) investigated the impact of manufacturing uncertainties on impeller performance, proposing methods to enhance performance robustness. Li et al. (2019a) performed aeromechanical multidisciplinary optimization to enhance isentropic efficiency and reduce stress in high-speed centrifugal impellers. Ma et al. (2024) applied a model-based optimization framework to fully open absorption heat pumps, achieving enhanced efficiency and economic benefits. Wang et al. (2017) used a combined energy loss model and computational fluid dynamics to optimize multistage centrifugal pumps, focusing on reducing interstage and volumetric leakage losses to improve efficiency. Despite these advancements, the optimization process remains resource-intensive, requiring significant computational resources, time-consuming simulations, and iterative testing to achieve desired performance improvements. These challenges highlight the necessity of more efficient optimization methods to streamline the design process while ensuring accuracy and performance.

Designing a new high-performance and structurally sound model requires substantial manpower and time. Therefore, improving design efficiency, shorten the design cycle, and reducing design difficulty have become key research focuses. The initial design of pump models has traditionally relied on Stepanoff's classic theory, a fundamentally semi-empirical approach (Stepanoff, 1943). This approach frequently employs similarity theory and coefficients derived from historical models. However, these empirical coefficients, derived from older models, have inherent limitations and cannot effectively address the diverse and complex application scenarios of modern systems. Many parameters in the design process depend on design's experience, and the accuracy of these values directly affects both the calculation time of iterative processes and the hydraulic performance of the impeller. As a result, this often leads to prolonged iterative cycles, and even after multiple iterations, achieving a high-performance hydraulic model remains difficult, leading to significant waste of manpower and time. For most designers, rapidly developing an initial model with appropriate values and good performance demands significant design experience. However, such experience can only be gained through long-term engineering practice, leading to low development efficiency for new centrifugal pump products and placing high demands on the experience of development personnel.

To address the limitations of traditional design methods relying on empirical and semi-empirical approaches, researchers have conducted extensive studies in recent years. Kim et al. (2022) developed a method for designing ultra-low specific speed centrifugal pumps for small liquid rocket engines, using empirical formulas to rapidly determine design variable. This approach enables the rapid derivation of design solutions in the early stages, offering an efficient framework to address the unique demands of pumps with very low specific speeds. Monteverde et al. (2023) critically analyzed and enhanced empirical models for predicting the performance of Electrical Submersible Pumps (ESPs) under viscous flow conditions (Verde et al. 2024). Their work emphasizes the application and refinement of empirical and semi-empirical models to improve the accuracy of performance predictions, which is critical for adapting pump designs to real-world conditions. Chernobrova et al. (2024) investigated the impact of volute casing design methods on pump performance, comparing the conservation of angular momentum approach with the constant-velocity approaches. Their study highlights the absence of clear guidelines for selecting design method and underscores the need for empirical analysis to optimize volute casing parameters for greater efficiency. These studies collectively demonstrate significant value of empirical methods in improving pump performance and streamlining design processes. However, exclusive reliance on previously summarized empirical formulas for pump design can result in substantial time and labor costs. Thus, integrating new technologies is essential to refine and optimize these empirical methods, enabling the rapid derivation of required design solutions in the early stages.

With the advancement in artificial intelligence, technologies such as machine learning and genetic algorithms have been widely adopted across various industries (Turing, 2009; Cioffi et al., 2020; Wadi et al., 2024). Artificial neural networks, known for their self-learning, adaptability, and robust nonlinear mapping capabilities, have been applied in image recognition, autonomous driving, big data prediction, and other fields. Machine learning models have become a prevalent method for replacing human judgment, learning from existing data, and making predictions based on accumulated experience (Rumelhart, et al. 1986). Additionally, various machine learning methods and optimization algorithms are extensively used to predict and optimize centrifugal pump performance. Joshi et al. (2024) developed machine learning models, including artificial neural networks (ANN), to predict power numbers for closed-clearance impellers, achieving high accuracy and improving process. Barsi et al. (2022) employed advanced simulation techniques, potentially augmented by machine learning, to analyze flow interactions in low-pressure turbine stages, providing insights relevant to centrifugal pump systems. Huang et al. (2020) proposed a hybrid neural network to efficiently predict the energy performance of centrifugal pumps, achieving significantly higher prediction accuracy compared to traditional methods.

Building on these advanced concepts, this article aims to refine and optimize the original centrifugal pump impeller design method. By leveraging machine learning techniques to extract insights from existing high-efficient hydraulic models, we aim to address the limitations of semi-empirical design methods. This approach aims to improve the hydraulic efficiency of centrifugal pump models, reduce energy consumption and increase design process. In doing so, it overcomes the limitations of traditional empirical coefficients and better meets the diverse and complex demands of modern applications.

2. METHOD

The traditional process of designing a centrifugal pump impeller involves determining its main dimensional parameters and structure, designing the impeller flow path profile (axial projection), and defining the blade geometry. This chapter focuses on the design of the impeller's main dimensions, the parameterization method and optimization strategy for the impeller flow path profile, and the determination of the main blade angles. The effectiveness and validation of this rapid design method will be discussed in the section on Rapid Design Method Validation section.

2.1 Main Dimensions Design of Centrifugal Pump Impeller

The velocity coefficient method is the most commonly used approach for determining the main dimensions of a centrifugal pump impeller. This method is based on the principle of similar design and is derived from the analysis of numerous high-performance models (Verhoeven, et al. 1988). The calculation of the impeller's main dimensions using the velocity coefficient method is as follows (Guan,

2011).

$$D_0 = k_0 \sqrt[3]{\frac{Q}{n}}, D_2 = k_D \sqrt[3]{\frac{Q}{n}}, b_2 = k_b \sqrt[3]{\frac{Q}{n}} \quad (1)$$

The above formulas are used to calculate the equivalent diameter of the inlet diameter D_0 , the impeller outlet width b_2 and the impeller outlet diameter D_2 . Here, Q represents the flow rate; n denotes the rotational speed; and k_0 , k_D and k_b are the velocity coefficients used in determining the main dimensions. The relationship between the impeller inlet diameter D_j , the equivalent diameter at the impeller inlet D_0 , and the impeller hub diameter d_h is given as follows.

$$D_j = \sqrt{D_0^2 + d_h^2} \quad (2)$$

The axial extension distance of the impeller Z_e and the curvature radius of the front cover plate R_{DS} are calculated based on empirical values, as expressed in the following formulas.

$$Z_e = (D_2 - d_h) \left(\frac{n_q}{20.3}\right)^{1.07} \quad (3)$$

In the traditional design process, the values of these coefficients are determined based on statistical data with different schemes applied depending on whether efficiency or cavitation is prioritized.

A fully connected neural network is employed to predict the coefficients in the above formula. The conceptual framework of the neural networks is illustrated in Fig. 2.

2.2 Database Structure and Composition

The high-quality hydraulic model database used to train the neural network in this study is primarily sourced from the “*Compilation of Centrifugal Pump Hydraulic Models*”, the “*Compilation of Hydraulic Model Data for the National Pump Industry*”, and several clear water

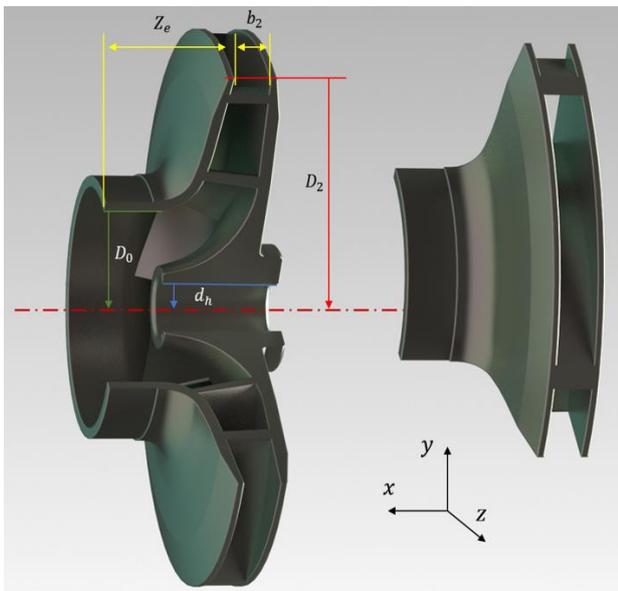


Fig. 1 Three-dimensional schematic diagram of the impeller's main dimensions

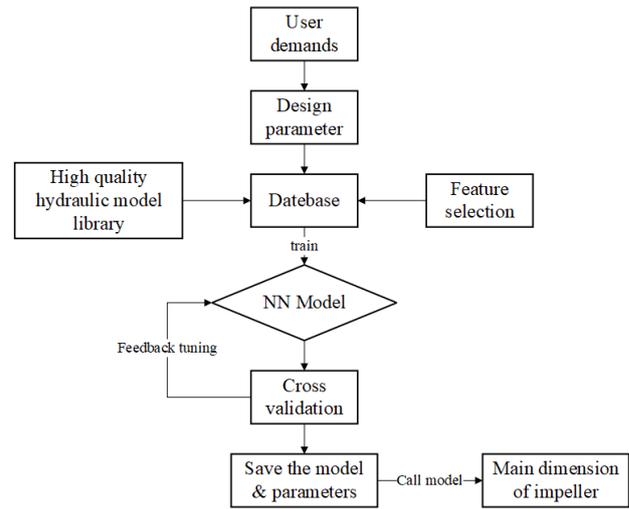


Fig. 2 Main dimension design neural network architecture

centrifugal pump models that have been successfully applied in engineering, totaling 128 datasets. The database includes all the centrifugal pump parameters required for subsequent training and validation, such as the flow rate Q , head H , speed n , number of blades N . It also contains the main geometric dimension parameters, including the hub diameter d_h , impeller inlet diameter D_j , impeller outlet diameter D_2 , impeller outlet width b_2 , blade wrap angle φ , blade inlet angle β_1 , and blade outlet angle β_2 . In addition to the above parameters, the database also includes the specific speed n_s and the coefficients k_0 , k_D , k_b obtained through reverse calculation. The specific speed is calculated uses the German formula, as shown below.

$$n_s = \frac{n\sqrt{Q}}{H^{3/4}} \quad (4)$$

The specific speed n_s is calculated, and Formula 1 is used to compute the coefficients k_0 , k_D and k_b through reverse calculation. These coefficients are then added to the database, and a portion of the database is displayed in Table 1. For the main dimension prediction neural network, the variables Q , n , H and n_s from the database serve as the input to the neural network, with a size of 4×1 . The coefficients k_0 , k_D and k_b , corresponding to the three dimensions D_j , D_2 , and b_2 , are used as the neural network output, with a size of 3×1 .

2.3 Neural Network for Impeller Main Dimensions

The structure of a fully connected neural network comprises two main modules: a forward propagation network for information flow and a backward propagation network for error correction (Li, et al. 2012). The network is typically organized into three layers: an input layer, a hidden layer, and an output layer, as illustrated in Fig. 3.

The number of neurons in the hidden layer and the learning rate during training significantly influence the prediction performance of the neural network. Therefore, it is necessary to explore different combinations (Li et al. 2019b). The specific combination scheme is presented in

Table 1 Partial pump database

$Q(m^3/s)$	$H(m)$	$n(r/min)$	$d_h(mm)$	$D_1(mm)$	$D_2(mm)$	$b_2(mm)$	N
0.00347	80.78	2900	0	52	242	4	4
0.00347	50.34	2900	0	48	200	6	5
0.00347	31.25	2900	0	44	160	5.6	5
0.01389	80.96	2900	0	80	252	6.5	5
0.00694	30.13	2900	0	62	160	9.5	5
0.05556	52.35	1450	19	150	399.5	21.5	5
0.00694	20.14	2900	0	65	130	9.5	6
0.01389	32.94	2900	23	76	168	13	6
0.02778	56.6	2900	0	90	210	16	6

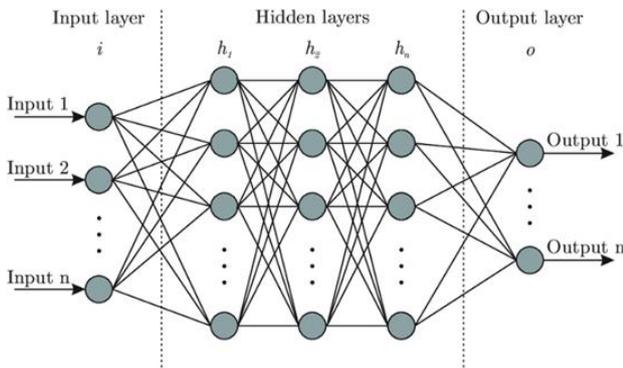


Fig. 3 Full connected neural network model

Table 2 Hidden layer node number and hyperparameter scheme

Nodes	4	5	6	7
Learning rate	1.00e-4	1.00e-3	1.00e-2	1.00e-1

Table 2, which outlines the ranges for the number of neurons and the learning rate. This results in a total of $4 \times 4 = 16$ combinations.

This study employs the grid search cross-validation method to evaluate the parameter schemes mentioned above. Grid search is an automatic parameter optimization technique. After defining the parameter combinations, the grid search evaluates each combination to identify the optimal scheme (Adnan et al. 2022). The grid search validates these hyperparameter schemes using a 10-fold cross-validation process, as illustrated in Fig. 4.

According to the grid search cross-validation results presented in Table 3, the minimum prediction error is 0.18777, achieved with 4 hidden layers and a learning rate of $1e-2$. The trained neural network model with the optimal parameter combination can be stored and utilized for predicting k_0 , k_D and k_b , enabling the rapid calculating the main dimensions of the centrifugal pump impeller. As an example, consider a case with a design flow rate of $Q = 50m^3/h$, a design head of $H = 34m$, a rated speed of $n = 2900r/min$. Using the trained neural network, the coefficients k_0 , k_D and k_b are predicted, leading to the determination of the equivalent inlet diameter $D_0 = 82.1mm$, the outlet diameter $D_2 = 169.7mm$, and the width $b_2 = 12.9mm$.

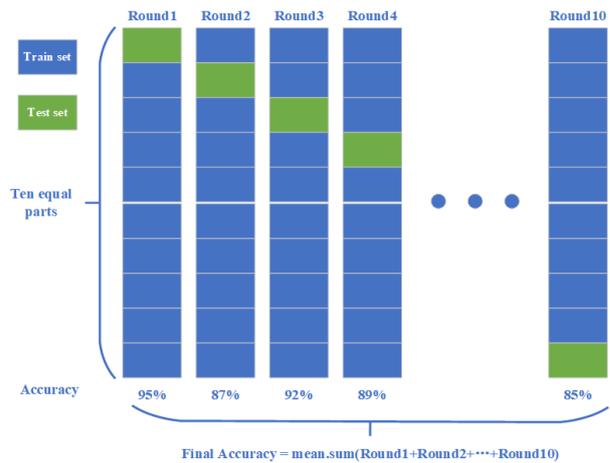


Fig. 4 10-fold Cross Validation

Table 3 Results of grid search & cross-validates

Nodes	Learning rate			
	1.00e-4	1.003e-4	1.00e-2	1.00e-1
4	1.21464	0.47592	0.18777	0.28104
5	1.07435	0.62799	0.19475	0.32165
6	1.25059	0.70541	0.20329	0.39863
7	1.38176	0.94747	0.07322	0.20671

3. DESIGN OF IMPELLER FLOW PATH FOR CENTRIFUGAL PUMP

The design of the impeller flow path plays a critical role in fluid movement across the width of the impeller blades. It directly affects internal pressure distribution, blade load distribution, turbulent dissipation losses, partial load performance, and the stability of the external characteristic curve. In this study, the parameterization of the flow path is achieved using a 5-point Bézier curve, with hydraulic efficiency defined as the primary objective and the cross-sectional area of the flow serving as the evaluation criterion. Constraints are applied to the control points, and a genetic algorithm is employed to automatically optimize these points, thereby determining the optimal flow path structure.

Table 4 Control point coordinates

Curve	Point	Value	Curve	Point	Value
Hub	P1	$(Z_e+b_2, 0)$	Shroud	P6	$(Z_e+b_2, D_j/2)$
	P2	Free: $(Z_e+b_2)/2, 0)$		P7	Free: $(Z_e/2)+b_2, D_j/2)$
	P3	Free: $(0, 0)$		P8	Free: $(b_2, D_j/2)$
	P4	Free: $(0, D_2/4)$		P9	Free: $(b_2, D_2/4-D_j/4)$
	P5	$(0, D_2/2)$		P10	$(b_2, D_2/2)$

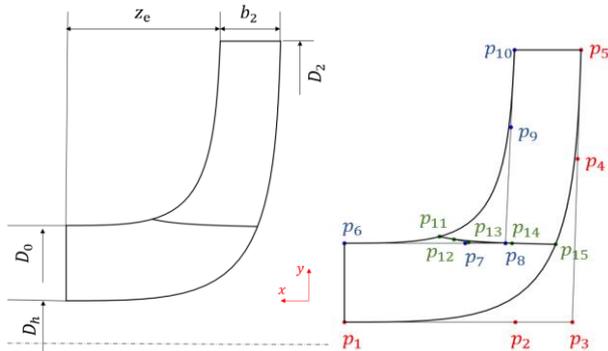


Fig. 5 Main dimensions and control points of shaft section

3.1 Parameterization of Impeller Flow Path Profile

Bézier curve is a mathematical curve tool widely used in two-dimensional graphic applications. It enables the generation of complex and smooth curves with only a few control points. The recursive calculation for an n -th-order Bézier curve is defined as follows (Shao, et al. 1996; Mineur, et al. 1998; Baydas, et al. 2019).

$$P_i^k \begin{cases} P_i, k = 0 \\ (1-t)P_i^{k-1} + tP_{i+1}^{k-1}, k = 1, 2, \dots, n; \\ i = 0, 1, \dots, n-k \end{cases} \quad (5)$$

In the design of the flow path, the blade inlet edge, sleeve, shroud, and the blade input edge are represented using a fifth-order Bézier curves, each defined by five points. As illustrated in Fig. 5, the design consists of a total of 15 control points.

The coordinates of each control point are listed in Table 4.

The control point on the blade’s inlet edge is established based on its proportional position along the respective profile lines at the intersects with the hub and shroud, as detailed in Table 5.

According to Formula 3, the coordinates of the 9 control points Z_e , P2-P4, P7-P9, and P12-P14 can be calculated. These points can be freely adjusted along their respective Bézier control lines. Once the coordinates of all 15 Bézier control points are determined, the shape of the impeller flow path can be parameterized and controlled.

3.2 Variation of Cross-Sectional Area of Impeller Passage

During pump operation, intense rotation and shear forces generate numerous vortex structures within the

Table 5 Inlet edge control point coordinates

Curve	Point	Value
Inlet Edge	P11	$(P_{shroud} *(Z_e+b_2), P_{shroud} *D_2)$
	P12	Free
	P13	Free
	P14	Free
	P15	$((P_{hub} * Z_e) + b_2, P_{hub} *(D_2/2))$

impeller flow path (Zhao, et al. 2018). The formation, evolution, and collapse of these vortex significantly impact the energy dissipation in the internal flow. To address this, variations in the cross-sectional area of the flow passage are analyzed using diagnostic methods from vortex dynamics theory. This approach prioritizes the control of internal flow separation and structural optimization when designing the main dimensions of the impeller flow path. In the field of rotating machinery, vortex dynamics diagnostics commonly employ two methods: cross-sectional flow diagnostics and boundary vorticity flow diagnostics. Among these, cross-sectional flow diagnostics evaluate quality monitoring changes in overall energy. Therefore, when designing the main dimensions of a centrifugal pump, integrating cross-sectional flow diagnostics to regulate the trend of overall energy changes within the flow passage can enhance the hydraulic performance of the initial design.

In the cross-sectional flow diagnostic method, the integral form of the momentum equation for a continuous fluid medium is given as follows (Wu, et al. 2007):

$$\int_V \rho \frac{D\mathbf{u}}{Dt} dV = \int_V \rho f dV + \oint_{\partial V} \tau dS \quad (6)$$

In the equation, V represents the control volume, ρ denotes the fluid density, D/Dt is the material derivative, f represents the body force, and τ denotes the spatial variable. For a surface element, the spatial variable τ is a function of the normal vector n and the time variable t .

$$\tau = \tau(x, y, n) \quad (7)$$

Given a second-order tensor $T(x, t)$, where $\tau = \tau(x, y, n) = n \cdot T(x, y, n)$, substituting equation (7) into equation (6) and taking the dot product with the velocity vector u transforms equation (6) into:

$$\rho \frac{D}{Dt} \left(\frac{1}{2} \|u\|^2 \right) = \rho f \cdot u + p \nabla \cdot u + \nabla \cdot (T \cdot u) - \Phi \quad (8)$$

In the equation Φ denotes the dissipation rate resulting from entropy increase. Considering the symmetry of the tensor T and applying the Reynolds Transport Theorem, we derive:

$$\frac{\partial K}{\partial t} + \int_{\partial V} \rho q^2 (u_n - u_{bn}) dS = \int_V \rho f dV + \int_{\partial V} t \cdot u dS - P - D \quad (9)$$

In the equation, u_{bn} denotes the velocity of the moving boundary of the control volume, K denotes the total kinetic energy of the control volume, and P and D respectively denote the work done by compression and the dissipation work within the entire control volume. Each of these quantities is represented by the following equations:

$$P \equiv - \int_V p \nabla \cdot u dV \quad (10)$$

$$D \equiv \mu \int_V \Phi dV \quad (11)$$

Neglecting the body force in the absolute coordinate system allows the dot product of the fluid surface stress and velocity to be converted into the shaft power transferred from the impeller to the fluid. Consequently, this relationship can be expressed as:

$$\int_{S_{b+h}} \hat{t} \cdot u dS = \Omega e_z \cdot \int_{S_{b+h}} r \times \hat{t} dS = \Omega \hat{M}_z \quad (12)$$

Substituting equation (12) into equation (9) transforms it into

$$\Omega \hat{M}_z = \frac{\partial K}{\partial t} + \int_W (\Pi^* u_l + \hat{t} \cdot u) dS - G + P + D \quad (13)$$

In the equation, W denotes the cross-sectional area of the flow passage. At high Reynolds numbers, where inertial forces dominate over viscous forces, it follows that $\Pi^* \gg |\hat{t}|$ and $p^* \gg |\mu \nabla \cdot u|$. Under these conditions, the equation can be simplified to:

$$\Omega \hat{M}_z = \frac{\partial K}{\partial t} + G + P + D \quad (14)$$

Here, $\Omega \hat{M}_z$ represents the shaft power transferred from the impeller to the fluid, and K is the total kinetic energy. The term G in equation (13) describes the energy increase as the fluid flows through the passage, which can be expressed as:

$$G \equiv \int_W p^* u_l dS - p_\infty^* U S_{in} \quad (15)$$

In the equation, the streamline velocity is denoted by u_l , U represents the axial velocity, S_{in} denotes the inlet cross-sectional area of the flow passage, and p^*, p_∞^* are respectively defined by the following equations:

$$p^* = p + \frac{1}{2} \rho \|u\|^2 \quad (16)$$

$$p_\infty^* = p_\infty + \frac{1}{2} \rho \|u\|^2 \quad (17)$$

In the equation, p_∞ represents the static pressure, and $P_u = \int_W p^* u_l dS$ denotes the total pressure flow integral. The term P_u objectively reflects the actual process of pressure increase along the flow passage, or the work done by the impeller on the fluid. It is associated with the local streamline density and is often referred to as the “total pressure flow integral” (Van, et al. 2007).

From the total pressure flow integral formula, it is evident that the cross-sectional area S of the flow passage influences the variation of the total pressure flow integral. From a design perspective, achieving high hydraulic efficiency requires ensuring that the energy acquisition process of the fluid in the flow passage remains uniform, without localized sudden changes. Based on this theory, when designing the main dimensions of a centrifugal pump impeller, it is essential to maintain uniform changes in the cross-sectional area of each flow passage.

The traditional approach to drawing impeller flow passage diagrams involves selecting a reference impeller diagram with similar main dimensions and performance characteristics to the target impeller. During the drawing process, it is generally required that the hub and shroud remain parallel or exhibit symmetric changes at the inlet and outlet positions, and the curvature of the flow passage should not be excessively sharp. The radial dimension of the blade inlet edge at the intersection with the hub and shroud should be equal to or slightly greater than the inlet diameter of the impeller. Once the axial projection diagram is completed, it is necessary to verify the variation of the flow passage’s cross-sectional area. To prioritize hydraulic efficiency, the variation curve of the cross-sectional area should be smooth and nearly linear.

In the hydraulic models used in databases, some incorporate cavitation performance to meet application requirements, leading to fluctuations and changes in the first half of the curve representing the variation in the cross-sectional area of the flow passage. This occurs because cavitation in centrifugal pumps typically begins at the leading edge of the impeller blades. To improve cavitation performance, the diameter at the blade inlet is appropriately increased. The research presented in this paper focuses exclusively on hydraulic and aims to simplify the curve evaluation algorithm. Consequently, the variation curve of the flow passage cross-sectional area generally follows one of three trends: linear, cubic, and parabolic. All shown in Fig. 6, these curve functions ensure minimal variation in the cross-sectional area near the inlet and outlet of the impeller. This study investigates the variation in the flow passage cross-sectional area curve and employs a genetic algorithm to automatically optimize the 15 control points of the flow passage profile, thereby ensuring the quality of the design.

The cross-sectional area of the flow channel is a parabolic surface generated by rotating the cross-sectional shape around the axis, making the area calculation an approximation. As shown in Fig. 5, the tangent points of the inscribed circle in the flow channel are A and B, the center of the circle is O, and the perpendicular line OD is

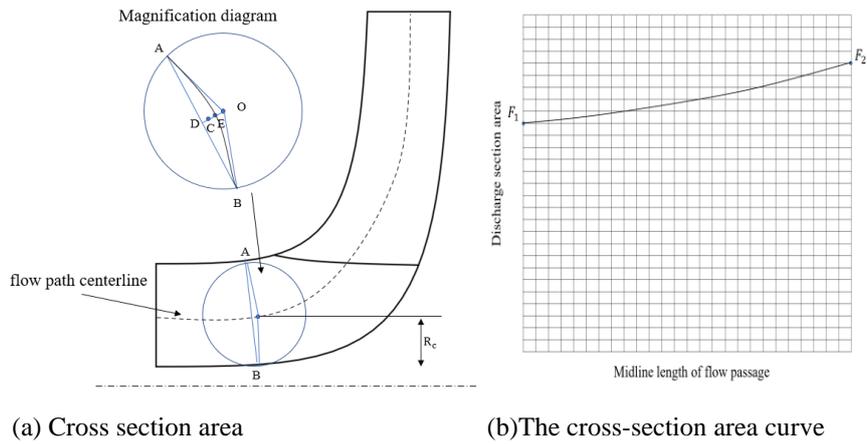


Fig. 6 Principle of flow section calculation

trisected. A circular arc c is drawn through the three points A, E and B, where the distance between the center O and the axis is R_c . For computational purposes, the length of the circular arc c is approximated as a straight line, represented as $OA+OB$. Thus, the cross-sectional area of the flow channel is calculated as follows:

$$F_p = 2\pi c R_c \approx 4\pi r R_c \quad (18)$$

3.3 Algorithm for Evaluating Cross-Sectional Area

Based on Stepanoff's classic theory and the analysis of the cross-sectional area change curve described above, this study focuses exclusively on hydraulic efficiency and temporarily disregards cavitation performance. The primary requirement is that the cross-sectional area change curve should be smooth and as close to a linear as possible. This ensures a relatively constant flow velocity inside the flow passage and minimize losses caused by the variations in flow velocity. To evaluate the linearity of the curve, this study employs the first and second derivatives. A smaller variation in the amplitude of the first derivative indicates that the cross-sectional area change curve is closer to a linear one. At the same time, the second derivative reflects the trend of changes in the first derivative. When the second derivative approaches zero, the variation in the amplitude of the first derivative decreases, and the curve becomes closer to a straight line. Based on this principle, the cross-sectional area change fL is discretized into a sufficient number of points. The slope of the curve between adjacent points is calculated, allowing the slope curve $f_1(x)$ of the area change curve fL to be plotted. Similarly, $f_1(x)$ is further discretized to calculate the second derivative function curve $f_2(x)$, which represents the second derivative of the area change curve fL . Subsequently, the number of extreme points E_x of the curve $f_1(x)$ and the absolute integral T of the curve $f_2(x)$ are computed. Typically, E_x is less than or equal to 2 (excluding endpoints). A smaller value of T indicates a more stable first derivative and a smoother variation of fL , as illustrated in Fig. 7.

In practical engineering, to account for anti-cavitation performance in the design of certain centrifugal pumps, the width b_1 at the blade inlet is increased. By increasing the cross-sectional area at the blade inlet, the inlet flow velocity is reduced, static pressure is raised, and dynamic

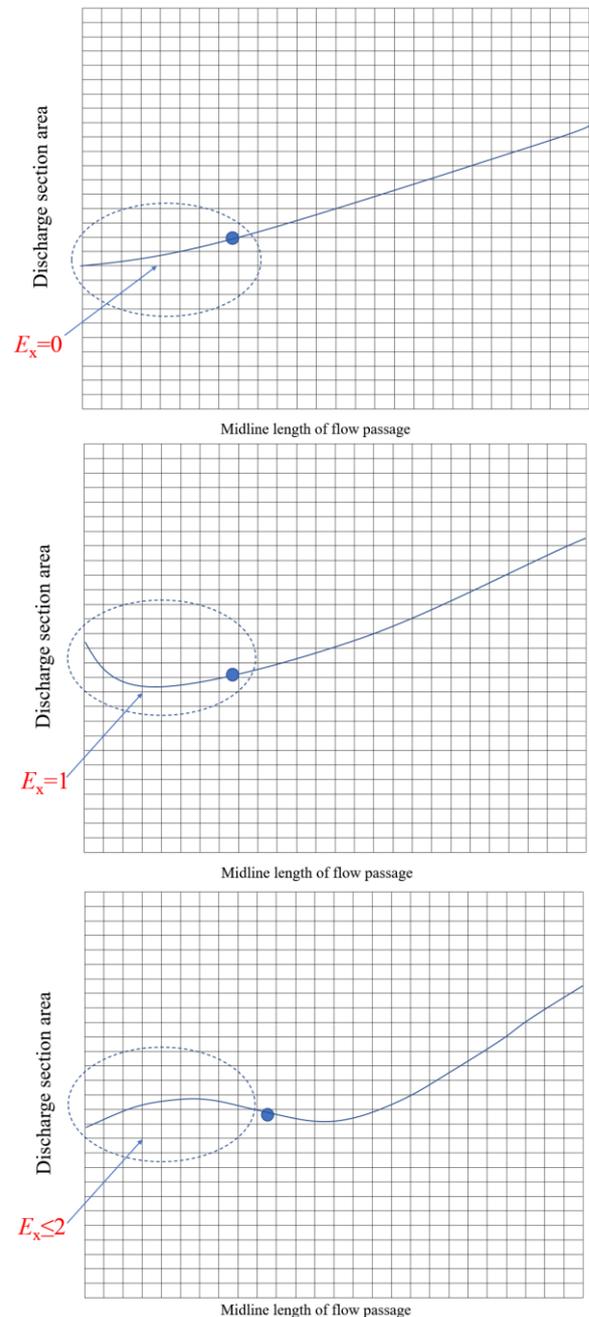


Fig. 7 Several main flow profiles

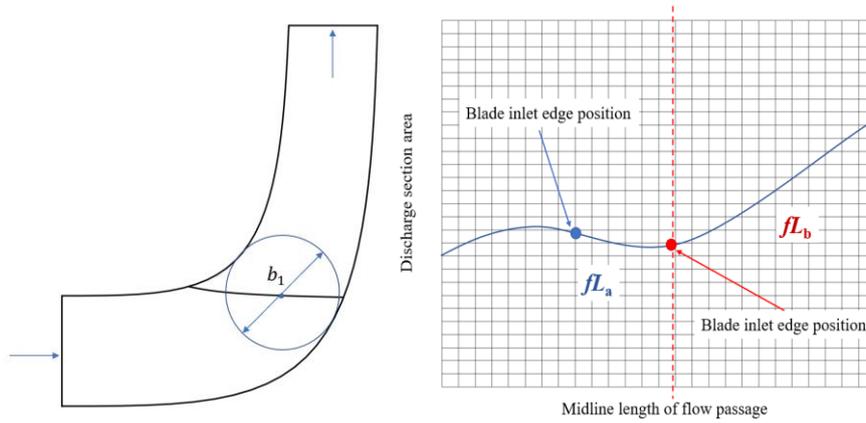


Fig. 8 Sectional evaluation diagram

Table 6 Evaluation of algorithm results

	<i>Efficient</i>	<i>T</i>		<i>Efficient</i>	<i>T</i>
Case1	0.723	0.389	Case10	0.819	0.079
Case2	0.764	1.054	Case11	0.695	1.786
Case3	0.801	0.438	Case12	0.742	0.863
Case4	0.638	1.184	Case13	0.785	0.058
Case5	0.823	1.032	Case14	0.682	1.585
Case6	0.656	0.372	Case15	0.822	1.342
Case7	0.731	0.052	Case16	0.653	1.876
Case8	0.719	0.863	Case17	0.672	0.574
Case9	0.747	1.287	Case18	0.736	0.495

pressure is lowered, thereby mitigating cavitation. As a result, the cross-section area curve at the inlet is no longer strictly linear but instead exhibits a convex region, as shown in Fig. 8. Thus, it is unreasonable to evaluate the T value and E_x of the entire cross-sectional area change curve as a whole. When using the evaluation index T , it is necessary to segment the cross-sectional area change curve. In this study, the midpoint of the centerline of the flow channel is used as the segmentation point. The first half of the curve fL_a , includes the convex region that may exist near the leading edge of the blades. For this segment, the E_x value should not exceed 2, and the T value is not particularly meaningful for evaluation. In contrast, the second half of the curve fL_b , is required to be nearly linear. For this segment, the E_x value should be less than or equal to 1, and the curve must not exhibit any additional fluctuations.

Once the evaluation method is established as the optimization objective, a genetic algorithm can be employed to optimize the coordinates of the impeller blade. However, prior to optimization, it is necessary to determine the approximate range of T to set reasonable termination conditions. To achieve this, 18 sets of centrifugal pump impeller flow passage area change curves with varying performance characteristics from the database are selected as test cases for the algorithm. The genetic algorithm is then applied to evaluate the impeller blade coordinates for these 18 centrifugal pump designs, and the evaluation results are presented in Table 6.

As shown in the table, none of the 18 cases exhibited negative values in the initial evaluation. Since this study focuses exclusively on hydraulic performance, it is sufficient to check whether the initial curve is positive. The data is visualized as a scatter plot in Fig. 8, where the horizontal axis represents the evaluation instance and the vertical axis corresponds to the T value of the evaluation function. In the scatter points plot, triangular points indicate models with efficiency below 70%, circular points represent models with efficiency between 70% and 80%, and quadrilateral points correspond to models with efficiency exceeding 80%. From Fig. 8, it can be observed that the evaluation results for models with efficiency between 70% and 80% are mostly within 1.1 or below. Meanwhile, models with efficiency greater than 80% predominantly fall below the average value of 0.79, which is represented by the dashed line labeled “Avg” in Fig. 9.

Once the evaluation method for the flow cross-sectional area has been established, the flow path can be assessed, and this method can serve as the objective function for optimizing the main dimensions of the impeller. In the previous chapter, a neural network for main dimension design was employed to predict the coefficients of the main dimensions, yielding the impeller D_j , D_2 , and b_2 . Based on Tab. 4, the coordinates of Bézier control points P1, P3, P5, P6, P8, and P10 for the shroud and hub can be calculated. Additional points, such as P2, are located on the segment P1P3; P4 is on the segment P3P5; P7 is on the segment P6P8; and P9 is on the segment

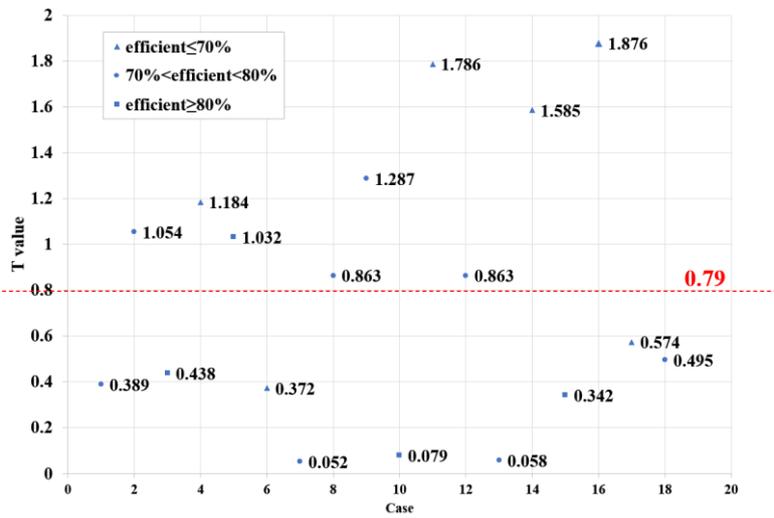


Fig. 9 Evaluation Result Chart

P8P10. The optimization range for the main dimensions that control the six points P1, P3, P5, P6, P8, and P10 is set to [95% of the initial value, 105% of the initial value]. Additionally, the four other Bézier curve control points P2, P4, P7, P9 can move freely along their respective line segments. If two points $A(x_1, y_1)$ and $B(x_2, y_2)$ are given, a point on the segment AB can be expressed parametrically using t , where $t \in [0, 1]$, as follows:

$$\begin{cases} x = (1 - t)x_1 + tx_2 \\ y = (1 - t)y_1 + ty_2 \end{cases} \quad (19)$$

After determining the optimization objects for the genetic algorithm, the algorithm's objective function is set as the flow cross-sectional evaluation method. During the optimization process, if the E_x value of the first half of the curve fL_a exceeds 2, the iteration is restarted. Conversely, when E_x is less than or equal to 2, the T value of the second half fL_b is calculated. The main dimension parameters corresponding to the minimum T value obtained through iterative calculation are identified as the final optimal impeller flow path. The basic process is illustrated in Fig. 10.

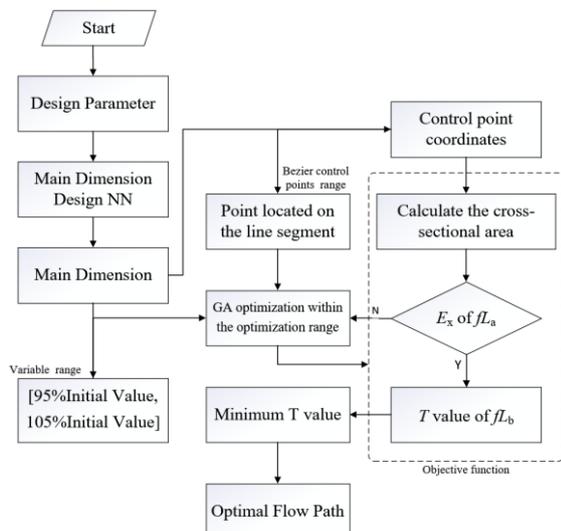


Fig. 10 Flow channel control point optimization process

When using a genetic algorithm for optimization, the error of the final optimization target is set to be within 0.05, or the iteration terminates after 200 rounds. The settings for genetic algorithm are as follows:

1. The T value is used as the optimization target.
2. The encoding length is 10 bits, with an encoding accuracy is 0.0029.
3. The population size is set to 50, and genetic operators such as proportional selection, single-point crossover, and single-point mutation are employed.
4. The crossover probability is 0.7, while the mutation probability is 0.1.
5. The maximum number of generation is set to 200, and elitism is applied to preserve the best solution.

4. DESIGN OF IMPELLER BLADES FOR CENTRIFUGAL PUMPS

Once the design of the impeller passage in a centrifugal pump is completed, the blade design can proceed accordingly. The impeller passage has a complex geometry and serves as the location where the fluid medium acquires energy. The blade, as the core component, performs work on the medium within the passage. The traditional blade design method, known as the blade angle preservation transformation method, faces challenges in determining the blade streamline during the design process. It often requires repeated iterations to define parameters such as the inlet angle, outlet angle, and wrap angles φ . For the blade inlet angle β_1 , the blade inlet liquid flow angle β_1^* can be calculated using equation (19). In traditional design, determining the blade inlet angle is essentially a matter of selecting the impingement angle $\Delta\beta_1$, which is typically no more than 15° .

$$\tan\beta_1^* = \frac{v_{m1}}{u_1 - v_{u1}}, \beta_1 = \beta_1^* + \Delta\beta_1 \quad (20)$$

The blade exit angle β_2 generally ranges between 18° and 40° , while the blade wrap angle φ typically falls within 90° and 110° . Notably, the blade wrap angle φ has an inverse relationship with the specific speed n_s . Traditional

Table 7 PCA weight analysis results

	β_1	β_2	φ	ABSaverage
N	-0.4259	-0.5472	-0.6611	0.544733333
D_j	0.4526	0.5385	0.5954	0.528833333
n_s	0.4798	0.5316	0.5433	0.518233333
b_2	0.4793	0.4682	0.4384	0.461966667
H	-0.4543	-0.3511	-0.5634	0.456266667
D_2	0.3732	0.3467	0.2231	0.314333333
Q	0.3712	0.2316	0.2976	0.300133333
n	-0.1382	-0.1947	-0.1887	0.173866667

design processes are prone to errors, making blade modeling and subsequent performance analysis challenging, which in turn hinders improvements in centrifugal pump performance and design efficiency. To address these issues, this section employs machine learning methods to rapidly determine blade angle parameters and expedite blade modeling.

Similar to the calculation method for main dimensions, this section utilizes neural networks to efficiently compute blade angles. After the flow channel design in the previous section, the available parameters include flow rate Q , head H , speed n , specific speed n_s , as well as the predicted impeller outer diameter D_2 , impeller inlet diameter D_j , impeller outlet width b_2 , and number of blades N , for a total of 8 parameters. Given the large number of input parameters, feature selection is crucial to ensure design efficiency. Principal component analysis (PCA) is applied to determine the weight coefficients of these eight parameters in relation to the impeller inlet angle, impeller outlet angle, and blade wrap angle (Singh, et al. 2022; Zhao, et al. 2022). The results of the weight calculations are presented in Table 7. In the table, $ABS_{average}$ represents the average of the absolute values.

$$ABS_{average} = \frac{(|\text{Value of } \beta_1| + |\text{Value of } \beta_1\beta_2| + |\text{Value of } \beta_1\varphi|)}{3} \quad (21)$$

The results of the principal component analysis (PCA) indicate that, except for a weak correlation with the rotational speed, the other 7 parameters exhibit varying degrees of influence on the blade inlet angle, outlet angle, and wrap angle. Consequently, Q , H , n_s , D_j , D_2 , b_2 and N were selected as the input features for the neural network model. The neural network model consists of 7 input neurons, corresponding to the 7 selected features. The output layer comprises 3 neurons, representing the predicted blade inlet angle, blade outlet angle, and blade wrap angle. The model includes a single hidden layer, and its hyperparameters are optimized using cross-validation. The selection algorithms and their corresponding hyperparameter configurations are listed in Table 8.

Grid search cross-validation was performed using the proposed scheme, and the results are presented in Table 9.

Table 8 Hidden layer node number and hyperparameter scheme

Nodes	8	9	10	11
Learning rate	1.00e-4	1.00e-3	1.00e-2	1.00e-1

Table 9 Results of grid search & cross-validates

Nodes	Learning rate			
	1.00e-4	1.003e-4	1.00e-2	1.00e-1
8	2.38126	0.37485	0.93217	1.98537
9	1.31237	0.84721	1.23644	1.48392
10	2.39481	0.26328	0.49583	0.99487
11	2.24588	0.93131	1.38473	1.39558

The grid search cross-validation results indicate that the minimum prediction error is 0.26328 when the number of hidden layer neurons is 10 and the learning rate is 1e-4. The trained neural network model with the optimal parameter combination is saved and can be used to predict β_1 , β_2 and φ , enabling rapid calculation of the main blade angles. As a practical example, consider a case where the design flow rate $Q = 50\text{m}^3/\text{h}$, the design head is $H = 34\text{m}$, the specific speed is $n_s = 22.56$, the impeller inlet diameter is $D_j = 80.3\text{mm}$, the impeller outlet diameter is $D_2 = 170.9\text{mm}$, and the impeller outlet width is $b_2 = 12.9\text{mm}$. Using the trained neural network, the predicted blade angles are $\beta_1 = 18.2^\circ$, $\beta_2 = 22.7^\circ$, and $\varphi = 103^\circ$.

5. RAPID DESIGN METHOD VALIDATION

To verify the universality and robustness of the rapid design method for centrifugal pump impellers within the specific speed range of 10 to 40, a Latin hypercube sampling method was employed for uniform sampling (Iman, 2008):

- (1) The specific speed range of the centrifugal pump was divided into 10 equally probable sub-intervals spanning from 10 to 40.
- (2) For each sub-interval, a flow rate was randomly selected. The flow rate range for all sub-intervals was from 20 to 200, divided into 10 equal intervals.
- (3) The steps above were repeated until all representative flow rates were used, resulting in a sample set. The sample size for this study was 10, and the specific sampling results are presented in Table 10.

By applying Latin hypercube sampling to select samples, conducting a rapid design of the impeller model, and performing simulation calculations, the universality and robustness of the rapid design method were validated.

Table 10. Super Latin sampling results

Case	n_s	Q	Case	n_s	Q
1	21	152	6	30	190
2	24	50	7	35	172
3	25	58	8	26	36
4	36	133	9	40	93
5	29	132	10	33	104

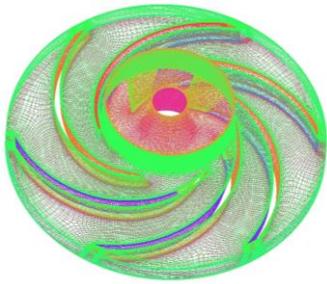


Fig. 11 Impeller mesh

Table 11 Grid independence analysis

Nodes Num	$H(m)$	$err(\%)$
1,584,348	28.7	-
1,905,674	29.5	2.71
2,538,454	30.3	2.64
3,038,057	30.5	0.67
3,630,583	30.7	0.61

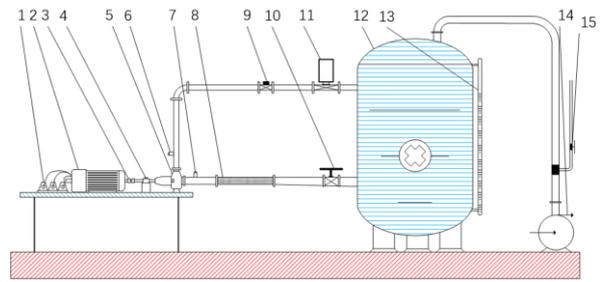
5.1 Experimental Design and Numerical Calculation Methods

ANSYS CFX was used to perform numerical simulations on the centrifugal pump impeller model. The working fluid water at 25°C, with a mass flow inlet applied at the impeller inlet. To eliminate the influence of volute on hydraulic efficiency across different designs, a free-flow outlet condition was applied at the impeller outlet. Steady-state calculations at the design operating point were conducted using the RNG $k-\epsilon$ turbulence model. The mesh was generated using TurboGrid, and the final mesh count was determined based on the results of mesh independence verification, as shown in Fig. 11 and Table 11. Additionally, the y^+ value of the boundary layer mesh was maintained at 16.7 or below.

Experimental verification of numerical calculation results was conducted using Case 2 from the sampled dataset. The experiment was performed on the test bench at the Research Center of Fluid Machinery Engineering and Technology, Jiangsu University. The test bench features a vertical closed-loop water system and supports the installation of models and devices such as centrifugal pumps and axial flow pumps. The primary objective of the experiment was to evaluate the external characteristics of centrifugal pump under normal operating conditions. The schematic diagram and physical image of the experimental setup are shown in Fig. 12 and Fig. 13.

To ensure the accuracy of the data, we rely on a systematic uncertainty analysis of the test rig. We summarize the types of sensors we used and their systematic uncertainties:

- Inlet and outlet pressure: WIKA S-10 static pressure sensor, accuracy class 0.2, systematic uncertainty $\pm 0.2\%$
- Rotational speed: SGDN-50 dynamic torque meter,



1. Current sensor 2. Drive motor 3. Coupling 4. Torque meter 5. Centrifugal pump 6. Export pressure sensor 7. Imported pressure sensor 8. Bellows 9. Flowmeter 10. Ball valve 11. Export electromagnetic valve 12. Storage pressure vessel 13. Liquid level plate 14. Vacuum pump 15. Release valve

Fig. 12 Schematic of the experimental rig



Fig. 13 Scene photo of the experimental rig

accuracy class 0.3, systematic uncertainty $\pm 0.3\%$

- Torque: SGDN-50 dynamic torque meter, accuracy class 0.3, systematic uncertainty $\pm 0.3\%$
- Flow rate: LDG-SIN-CN65-Z2 electromagnetic flow meter, accuracy class 0.1, systematic uncertainty $\pm 0.1\%$

These uncertainty values reflect the high precision of our equipment and the reliability of the data. We implemented rigorous experimental control measures to ensure that the selected data accurately reflects the experimental results.

The parameters of the physical model of the centrifugal pump corresponding to the sample are shown in Table 12. External characteristic tests were conducted on the model pump, and the resulting external characteristic curves were compared with the numerical calculations, as shown in Fig. 14.

From the graph, the trends of the experimental values and the simulated calculation values are consistent. When the flow rate Q is less than 20 m³/h, the simulated efficiency closely matches the experimental efficiency, with a difference of approximately 2%. The maximum difference between the simulated and the experimental

Table 12 Design parameters of sample pump

Parameter	Unit	Value	Parameter	Unit	Value
Rated head	H_d (m)	34	Inlet diameter	D_1 (mm)	72
Rated speed	n (r/min)	2900	Outlet diameter	D_2 (mm)	169
Rated efficiency	η (%)	75	Outlet width	b_2 (mm)	11
Number of blades	Z	6	Blade inlet angle	β_1 (°)	15.5
Blade wrap angle	φ (°)	92	Blade outlet angle	β_2 (°)	20.2

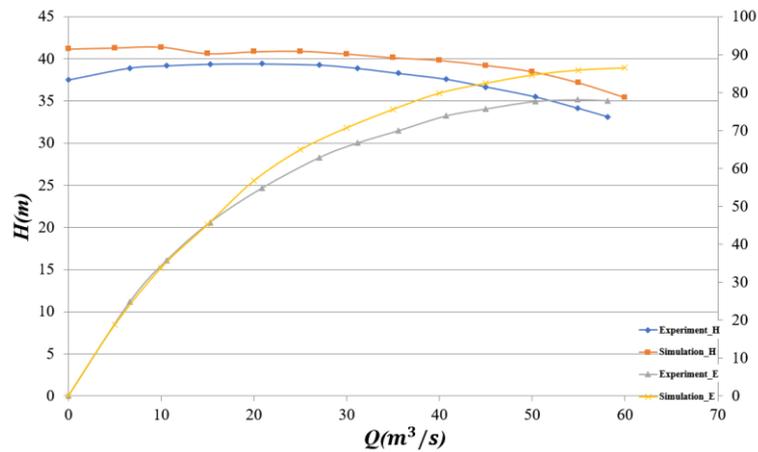


Fig. 14 Comparison of external characteristics between model pump test and numerical simulation

head value is about 4m. At the design operating point, the maximum difference between the simulated efficiency and the experimental efficiency is approximately 8%, while the simulated head value is about 3m higher than the experimental value. Overall, the simulated efficiency values are slightly higher than the experimental results. This discrepancy is likely due to processing errors in the model used for the simulation calculations, as well as the neglect of certain hydraulic losses, such as those from the inlet pipe and volute casing, during the simulation process. In conclusion, the numerical calculation results closely align with the experimental trends, demonstrating that the numerical calculation method meets the research requirements of this paper.

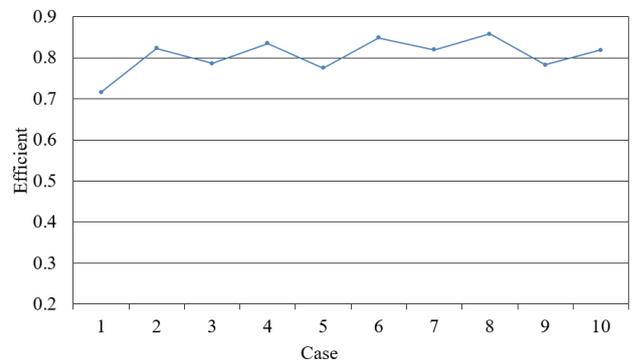


Fig. 15 The impeller efficiency designed within the range of specific speed

6. RESULT ANALYSIS

In terms of design effectiveness, centrifugal pump impellers were designed for 10 cases selected through Latin hypercube sampling using the proposed scheme, and numerical calculations were performed. The hydraulic efficiency results obtained from the numerical calculations are presented in Table 13 and Fig. 15.

Table 13 Numerical results.

Case	Efficient	Case	Efficient
Case1	74.6%	Case6	84.9%
Case2	82.3%	Case7	82.0%
Case3	78.6%	Case8	85.8%
Case4	83.5%	Case9	78.3%
Case5	77.5%	Case10	81.9%

As shown in the table, the impeller models designed using the rapid design method achieve a minimum efficiency of 74.6% and a maximum efficiency of 85.8% at the operating point, with an average efficiency of 80.94%. Figure 15 illustrates the impeller efficiency within the specific speed range, demonstrating that the rapid design method exhibits good universality and stable performance within the specific speed range of 10-40.

In traditional impeller design processes, parameters are often determined using the bisection method. For instance, when the empirical value range fluctuates by 5mm, the iteration error can be calculated using the bisection method formula. Given an initial interval $[a, b]$, where the initial length of the interval is $L = b - a$, the error after n iterations can be expressed as:

$$Error = \frac{b - a}{2^n} \quad (22)$$

To control the error within 0.5mm, at least four iterations are required. However, the machine learning approach employed in this study eliminated the need for such iterative calculations, directly providing suitable parameters and thereby accelerating the impeller design process.

In terms of design speed and efficiency, traditional centrifugal pump impeller design relies heavily on empirical values and iterative calculations. In contrast, the rapid design method based on machine learning quantifies the experience of existing high-efficiency models and utilizes trained neural networks to quickly determine various parameter value. This approach avoids the need to iteratively refine parameters based development time to less than 1/4 of the original duration.

7. CONCLUSION

This article proposes a rapid design method for centrifugal pump impellers based on machine learning. The method leverages neural networks to quantify the experience of existing high-efficiency models, enabling the quick prediction of key parameters during the impeller design process and reducing design time to less than 1/4 of that required by traditional methods. Numerical calculations performed on 10 cases using Latin hypercube sampling show that impellers designed within the specific speed range achieved a minimum efficiency of 74.6%, a maximum efficiency of 85.8%, and an average efficiency of 80.8%. These results demonstrate the method's good universality and stable design performance. The main conclusions of this study are as follows:

1. With sufficient high-quality hydraulic model data, machine learning can be effectively applied to design the structural parameters of centrifugal pump impellers, significantly improving the design speed for new models.

2. Parameterizing the flow passage control line using Bézier curves enables the calculation of the area variation curve of the flow passage cross-section. A flow passage shape assessment method based on vortex dynamics theory facilitate the rapid optimization of the flow passage shape.

3. Within the specific speed range of 10 to 40, Latin hypercube sampling was used to validate the robustness of the design method. Experimental validation of one cases confirmed the reliability of the design method.

It is also important to note that due to data acquisition challenges, the current dataset is limited in size. The hydraulic models used in this study were primarily sourced from existing industrial applications, which resulted in high costs and time requirements for data collection. However, the quality of the dataset is high, as the models have undergone rigorous engineering and market validation. Further research is being conducted to address existing shortcomings. Future work will focus on the following areas:

1. Expanding the hydraulic model dataset to include

different types of pumps, such as mixed-flow pumps and axial-flow pumps, to optimize and expand the proposed design method.

2. Incorporating cavitation performance into the optimization of main dimensions and blade angles, exploring the implicit relationships between geometric dimensions and cavitation performance, and developing appropriate algorithms to further refine this machine learning based impeller design method.

ACKNOWLEDGEMENTS

The authors are grateful for the financial support from the National Natural Science Foundation of China (52120105010). National Natural Science Foundation of China (52309112)

CONFLICT OF INTEREST

There is no conflict of interest to be declared in this work.

AUTHORS CONTRIBUTION

Conceptualisation: **Yunfei Chen**; Methodology: **Yunfei Chen** and **Yin Luo**; Software: **Yunfei Chen**; Validation: **Yunfei Chen**, **Wei Li**, and **Leilei Ji**; Formal analysis: **Yunfei Chen** and **Shuo Li**; Investigation: **Yu Long**; Resources: **Wei Li**; Data curation: **Ying Luo** and **Yunfei Chen**; Writing—original draft preparation: **Yunfei Chen**; Writing—review and editing: **Yunfei Chen** and **Wei Li**; Visualisation: **Yunfei Chen** and **Yu Long**; Supervision: **Wei Li**; Project administration: **Shuo Li**; Funding acquisition: **Wei Li**.

REFERENCES

- Adnan, M., Alarood, A. A. S., Uddin, M. I., & ur Rehman, I. (2022). Utilizing grid search cross-validation with adaptive boosting for augmenting performance of machine learning models. *Peer J Computer Science*, 8, e803. <https://doi.org/10.7717/peerj-cs.803>
- Barsi, D., Lengani, D., & Simoni, D. (2022). Analysis of the loss production mechanism due to cavity–main flow interaction in a low-pressure turbine stage. *Journal of Turbomachinery*, 144(9), 091004. <https://doi.org/10.1115/1.4053745>
- Baydas, S., & Karakas, B. (2019). Defining a curve as a Bezier curve. *Journal of Taibah University for Science*, 13(1), 522–528. <https://doi.org/10.1080/16583655.2019.1601913>
- Chernobrova, A., Moloshnyi, O., & Szulc, P. (2024). Influence of Volute Casing Design Methods and Changes in Geometric Parameters on Pump Operation. *Energies*, 17(18), 4590. <https://doi.org/10.3390/en17184590>
- Cioffi, R., Travaglioni, M., Piscitelli, G., Petrillo, A., & De Felice, F. (2020). Artificial intelligence and machine learning applications in smart production: Progress, trends, and directions. *Sustainability*, 12(2),

492. <https://doi.org/10.3390/su12020492>
- Guan, X. F. (2011). *Modern Pump Theory and Design*. Beijing: China Machine Press.
- Huang, R., Zhang, Z., Zhang, W., Mou, J., Zhou, P., & Wang, Y. (2020). Energy performance prediction of the centrifugal pumps by using a hybrid neural network. *Energy*, 213, 119005. <https://doi.org/10.1016/j.energy.2020.119005>
- Joshi, S. S., Dalvi, V. H., Vitankar, V. S., J. B. Joshi, & A. J. Joshi (2024). Development of new correlation for the prediction of power number for closed clearance impellers using machine learning methods trained on literature data. *The Canadian Journal of Chemical Engineering*. <https://doi.org/10.1002/cjce.25385>
- Ju, Y., Liu, Y., Jiang, W., & Zhang, C. (2021). Aerodynamic analysis and design optimization of a centrifugal compressor impeller considering realistic manufacturing uncertainties. *Aerospace Science and Technology*, 115, 106787. <https://doi.org/10.1016/j.ast.2021.106787>
- Kim, H. I., Roh, T. S., Huh, H., & Lee, H. J. (2022). Development of ultra-low specific speed centrifugal pumps design method for small liquid rocket engines. *Aerospace*, 9(9), 477. <https://doi.org/10.3390/aerospace9090477>
- Li, C., Wang, J., Guo, Z., Song, L., & Li, J. (2019a). Aero-mechanical multidisciplinary optimization of a high-speed centrifugal impeller. *Aerospace Science and Technology*, 95, 105452. <https://doi.org/10.1016/j.ast.2019.105452>
- Li, J., Cheng, J. H., Shi, J. Y., & Huang, F. (2012). *Brief introduction of back propagation (BP) neural network algorithm and its improvement*. Advances in Computer Science and Information Engineering: Volume 2 (pp. 553–558). Springer. https://doi.org/10.1007/978-3-642-30223-7_87
- Li, W., Yang, Q., Yang, Y., Ji, L., Shi, W., & Agarwal, R. (2024). Optimization of pump transient energy characteristics based on response surface optimization model and computational fluid dynamics. *Applied Energy*, 362, 123038. <https://doi.org/10.1016/j.apenergy.2024.123038>
- Li, Y., Wei, C., & Ma, T. (2019b). Towards explaining the regularization effect of initial large learning rate in training neural networks. *Advances in Neural Information Processing Systems*, 32. <https://doi.org/10.48550/arXiv.1907.04595>
- Ma, Y., Gao, E., Zhang, X., & Huang, S. (2024). Parametric analysis and design optimization of a fully open absorption heat pump for heat and water recovery of flue gas. *Applied Energy*, 375, 124144. <https://doi.org/10.1016/j.apenergy.2024.124144>
- Mineur, Y., Lichah, T., Castelain, J. M., & Giaume, H. (1998). A shape controlled fitting method for Bézier curves. *Computer Aided Geometric Design*, 15(9), 879–891. [https://doi.org/10.1016/S0167-8396\(98\)00025-9](https://doi.org/10.1016/S0167-8396(98)00025-9)
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533–536. <https://doi.org/10.1038/323533a0>
- Shao, L., & Zhou, H. (1996). Curve fitting with Bezier cubics. *Graphical Models and Image Processing*, 58(3), 223–232. <https://doi.org/10.1006/gmip.1996.0019>
- Singh, D., & Singh, B. (2022). Feature wise normalization: An effective way of normalizing data. *Pattern Recognition*, 122, 108307. <https://doi.org/10.1016/j.patcog.2021.108307>
- Stepanoff, A. (1943). Centrifugal-pump performance as a function of specific speed. *Transactions of the American Society of Mechanical Engineers*, 65(6), 629–636. <https://doi.org/10.1115/1.4018866>
- Turing, A. M. (2009). *Computing machinery and intelligence*. Springer. https://doi.org/10.1007/978-1-4020-6710-5_3
- Tuzson, J. (2000). *Centrifugal pump design*. John Wiley & Sons.
- Van Oudheusden, B. W., Scarano, F., Roosenboom, E. W., Casimiri, E. W., & Souverein, L. J. (2007). Evaluation of integral forces and pressure fields from planar velocimetry data for incompressible and compressible flows. *Experiments in Fluids*, 43, 153–162. <https://doi.org/10.1007/s00348-007-0261-y>
- Verde, W. M., Kindermann, E., Bulgarelli, N. A. V., Pastre, L. F., Foresti, B., & Bannwart, A. C. (2024). A critical analysis and improvements of empirical models for predicting the performance of Electrical Submersible Pumps under viscous flow. *Geoenergy Science and Engineering*, 238, 212871. <https://doi.org/10.1016/j.geoen.2024.212871>
- Verhoeven, J. J. (1988). *Rotordynamic considerations in the design of high-speed centrifugal pumps*. Proceedings of the 5th International Pump Users Symposium. Turbomachinery Laboratories, Department of Mechanical Engineering, Texas A & M. <https://hdl.handle.net/1969.1/164295>
- Wadi Al-Fatlawi, A., Hashemi, J., Hossain, S., & M. El Haj Assad (2024). Applying machine learning in CFD to study the impact of thermal characteristics on the aerodynamic characteristics of an airfoil. *Journal of Applied Fluid Mechanics*, 17(4). <https://doi.org/10.47176/jafm.17.4.2276>
- Wang, C., Shi, W., Wang, X., Jiang, X., Yang, Y., Li, W., & Zhou, L. (2017). Optimal design of multistage centrifugal pump based on the combined energy loss model and computational fluid dynamics. *Applied Energy*, 187, 10–26. <https://doi.org/10.1016/j.apenergy.2016.11.046>
- Wilson, K. C., Addie, G. R., Sellgren, A., & Clift, R. (2006). *Centrifugal pumps*. Springer. <https://doi.org/10.1007/b101079>
- Wu, J. Z., Ma, H. Y., & Zhou, J. Z. (2007). *Vorticity and vortex dynamics*. Springer NetLibrary, Inc.

Zhao, B., Dong, X., Guo, Y., Jia, X., & Huang, Y. (2022). PCA dimensionality reduction method for image classification. *Neural Processing Letters*. <https://doi.org/10.1007/s11063-021-10632-5>

Zhao, X., Xiao, Y., Wang, Z., Luo, Y., & Cao, L. (2018). Unsteady flow and pressure pulsation characteristics analysis of rotating stall in centrifugal pumps under off-design conditions. *Journal of Fluids Engineering*,

140(2), 021105. <https://doi.org/10.1115/1.4037973>

Zhou, L., Hang, J., Bai, L., Krzemianowski, Z., El-Emam, M. A., Yasser, E., & Agarwal, R. (2022). Application of entropy production theory for energy losses and other investigation in pumps and turbines: A review. *Applied Energy*, 318, 119211. <https://doi.org/10.1016/j.apenergy.2022.119211>