Optimization of twisted blade of pump as turbine based on high dimensional surrogate model

Bingxiao Jiang1, Junhu Yang1,2, Xiaohui Wang1,2, Fengxia Shi1,2, and Xiaobang Bai1,3
1School of Energy and Power Engineering, Lanzhou University of Technology, Lanzhou, China
2Key Laboratory of Fluid Machinery and Systems, Lanzhou, China
3National Enterprise Technology Center, Chongqing Pump Industry Co., Ltd, Chongqing, China

Abstract
In order to improve the operation efficiency of the twisted blade pump as turbine (PAT), a medium specific speed PAT was selected as the research object. The variables of the twisted blade plane blade profile were defined, the twisted blade was transformed into three plane blade profiles, and the blade profiles were parameterized by MATLAB 9.7 software. MATLAB 9.7, CFturbo 2020 and Fluent 19.2 were used to build the support vector machine-high dimensional model representation (SVM-HDMR) surrogate model function for efficiency optimization of PAT. Genetic algorithm was run on MATLAB 9.7 to optimize the surrogate model function, and the optimized blade profiles were fed back. The optimization results were verified by numerical simulation and experiment. The results show that the simulation efficiency of the PAT after optimization at the design operating point is 3.51% higher than the efficiency of the PAT before optimization, and the output power is increased by 5.3%. The test efficiency of the PAT after optimization at the design operating point is 3.4% higher than the efficiency of the PAT before optimization, and the output power is increased by 5.1%.

Keywords
pump as turbine, twisted blade, blade profile, support vector machine, high dimensional model representation, surrogate model, genetic algorithm

Corresponding author:
Junhu Yang, School of Energy and Power Engineering, Lanzhou University of Technology, No.287 Langongping Road Qilihe District, Lanzhou City, Gansu, P.R.China.
Email: lzyangjh@lut.cn
Introduction

In the petrochemical industry, seawater desalination, iron and steel metallurgy and other industries, there are a lot of high-pressure liquid to be reused, and the selection of pump as turbine (PAT) to recover the energy of these high-pressure liquid is one of the most effective ways. Impeller blade is the main component of PAT to convert high-pressure liquid energy into mechanical energy. The degree of compliance between blade shape and internal flow of impeller has a significant impact on the hydraulic loss of impeller. Therefore, the blade shape optimization is of great significance to improve the efficiency of PAT. Since the PAT blade is designed according to the working condition of centrifugal pump, the PAT blade and the centrifugal pump blade have the same characteristics, which directly leads to the low efficiency of the PAT. According to the geometric similarity criterion, centrifugal pump can be divided into low, medium, and high specific speed. The low specific speed centrifugal pump usually adopts cylindrical blade. Under the premise of uniform distribution of axial velocity along the inlet of the centrifugal pump at medium and high specific speed, due to the forward extension of the inlet side of the blade, the circumferential velocity of each point on the inlet side is different, so that the inlet placement angle can no longer be equal. In order to meet the flow situation, twisted blade should be used. Therefore, the corresponding PAT blade can be divided into cylindrical blade and twisted blade.

At present, there are many models studied in the field of hydraulic machinery blade optimization, but few of them have been popularized and applied. The reasons mainly include two aspects: first, the traditional conformal transformation method cannot quickly and accurately transform the two-dimensional blade profile into three-dimensional twisted blade, which makes the blade research based on modern optimization design methods mainly focus on cylindrical blade, such as incomplete sensitivity method, artificial neural network (ANN) method, proper orthogonal decomposition (POD) method, response surface method (RSM), Kriging approximate model method, adjoint method, multidisciplinary feasibility (MDF) etc. Secondly, in the process of blade optimization from cylindrical blade to twisted blade, the high-dimensional problem caused by the increase of variables has not been well solved, which leads to the increase of the number of samples used in the training and learning of surrogate model. For example, the ANN surrogate model which is widely studied at present, however, has higher requirement for the selection of sample points in the experimental design, and the specific requirement is that the selected sample points should meet the requirement of “uniform dispersion – the test points are evenly distributed in the design space, so that each test point has sufficient representation”. Because the design points of Latin square test are generated by random combination, the distribution of points is no longer regular and the uniformity cannot be guaranteed. Then the experimental methods that can be combined with these two surrogate models are the full factorial design (FFD) and the central combination design (CCD). FFD method will calculate the combination of all levels of all factors, and calculate the response of all nodes after evenly dividing the design space, so the number of test points of this method is \( N_1 \times N_2 \times \cdots \times N_i \times \cdots \times N_n \) (where \( N_i \) is the level number of the \( i \)-th factor). When the design variable is 5 and each variable only takes 4 values, the number of test points needed for this method is \( 4^5 = 1024 \). Therefore, in the case of high dimension of design space, the number of test points needed for FFD will be very large. Based on the
two-level orthogonal table, the minimum number of experimental points in the CCD is \(2^n + 1 + 2n\). More importantly, with the increase of dimension, the amount of computation is growing exponentially. In reference,\(^{17}\) ANN combined with genetic algorithm is applied to the optimization of centrifugal impeller. In order to establish ANN model, even after learning 662 models, the accuracy of surrogate model cannot be fully guaranteed. For industries with easy access to data, such as face recognition and data monitoring in big data industry, the above two experimental design methods are well applied. For the hydraulic machinery industry, the calculation amount of single hydraulic model has been considerable, it is very difficult to complete more than 100 hydraulic models’ calculation. In summary, the optimization of hydraulic machinery for high-dimensional problem has not been well solved. Therefore, it is the key to solve the problem that the rapid modeling of twisted blade and the research of small sample surrogate model.

Support vector machine\(^{18–21}\)-high dimensional model representation (SVM-HDMR) model is a machine learning method based on high dimensional model representation method. It has good robustness and high accuracy in data mining with small sample size,\(^{22–25}\) and can effectively make up for the lack of learning sample size in the process of PAT blade optimization. In view of the difference between the three-dimensional modeling ideas of the twisted blade and the cylindrical blade of the PAT and the importance of the optimization of the twisted blade, the SVM-HDMR surrogate model was applied to the optimization process of the twisted blade of the PAT. A medium specific speed PAT was selected as the research object, and the effectiveness of SVM-HDMR model applied to the optimization process of the twisted blade of the PAT was verified through the optimization of the twisted blade of the PAT.

The structure of this paper is as follows. First, the main parameters setting of Fluent 19.2 software in PAT numerical simulation, the grid independence check of computational fluid domain, the accuracy verification of numerical simulation, the drawing principle of PAT twisted blade and the parameterization of blade profile are introduced. Secondly, the theory of SVM-HDMR surrogate model, the construction process of PAT surrogate model and the optimization strategy of PAT twisted blade are introduced. Finally, the optimization process of twisted blade of PAT, the comparison of twisted blade before and after optimization of PAT, the comparison of velocity flow field and cloud images of kinetic energy equation in fluid dynamics before and after optimization of impeller, and the comparison of external characteristic curves before and after optimization of PAT are introduced.

### Materials

Numerical simulation of PAT

**Computational fluid domain and parameter setting of numerical simulation.** The design parameters of the selected medium specific speed PAT under the condition of centrifugal pump are: the flow rate is \(Q = 25 \text{ m}^3/\text{h}\), the head is \(H = 20.14 \text{ m}\), and the rotational speed is \(n = 2900 \text{ rpm}\). The specific structural parameters are shown in Table 1.

In Table 1, \(\beta_{2RS}\) represents the blade outlet angle corresponding to the spatial blade profile formed by twisted blade and rear shroud (RS); \(\beta_{2M}\) represents the blade outlet
angle corresponding to the spatial blade profile formed by twisted blade and middle (M) flow surface; $\beta_{2FS}$ represents the blade outlet angle corresponding to the spatial blade profile formed by twisted blade and front shroud (FS). The specific twisted blade structure and the computational fluid domain of PAT determined by Table 1 are shown in Figure 1.

From Figure 1, the computational fluid domain includes import extension section, volute, impeller, and export extension section. As shown by the arrow, the liquid

| Name   | Parameter                                      | Value  |
|--------|------------------------------------------------|--------|
| Impeller | Impeller inlet diameter $D_1$ (mm)             | 130    |
| Impeller | Impeller outlet diameter $D_2$ (mm)            | 65     |
| Impeller | Impeller hub diameter $d_h$ (mm)              | 10     |
| Impeller | Impeller inlet width $b_1$ (mm)               | 9.5    |
| Impeller | Blade inlet angle $\beta_1$ ($^\circ$)        | 17     |
| Impeller | Blade outlet angle $\beta_{2RS}$ ($^\circ$)    | 21.8   |
| Impeller | Blade outlet angle $\beta_{2M}$ ($^\circ$)    | 18.7   |
| Impeller | Blade outlet angle $\beta_{2FS}$ ($^\circ$)   | 15.6   |
| Impeller | Blade wrap angle $\theta$ ($^\circ$)          | 100    |
| Impeller | Blade number $z$                               | 6      |
| Volute | Volute base diameter $D_0$ ($^\circ$)         | 140    |
| Volute | Volute outlet width $b_0$ ($^\circ$)          | 22     |
| Volute | Volute inlet diameter $D_t$ (mm)              | 50     |

**Figure 1.** The twisted blade structure and the computational fluid domain of PAT: (a) The blade angle of twisted blade; (b) computational fluid domain of PAT.
flows in through the inlet, then flows through the import extension, volute, impeller, export extension, and finally flows out through the outlet. In the numerical calculation, because the operating conditions of PAT are known, the boundary conditions of velocity inlet, free outflow and no-slip wall were adopted. The computational fluid domain includes static region and rotating region, the multiple reference frame (MRF) model was used to deal with this situation, the computational domain formed by the impeller was set as the rotating region, and the interfaces were used between the static-static and dynamic-static computational domains to transfer the computational data.

The SST $k-\omega$ model$^{26,27}$ was used to simulate the internal turbulence of PAT. The specific expression of the governing equation is as follows:

$$\frac{\partial k}{\partial t} + u_j \frac{\partial k}{\partial x_j} = \frac{1}{\rho} \tilde{P}_k - \beta * k \omega + \frac{\partial}{\partial x_j} \left[ (\nu + \sigma_k \nu_i) \frac{\partial k}{\partial x_j} \right]$$  

(1)

$$\frac{\partial \omega}{\partial t} + u_j \frac{\partial \omega}{\partial x_j} = \alpha S^2 - \beta \omega^2 + \frac{\partial}{\partial x_j} \left[ (\nu + \sigma_\omega \nu_i) \frac{\partial \omega}{\partial x_j} \right] + 2(1 - F_1) \frac{\sigma_{o2} k \omega}{\omega} \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_j}$$  

(2)

The definition of $\tilde{P}_k$ in formula (1) is as follows:

$$\tilde{P}_k = \min (P_k, 10 \beta^* \rho k \omega)$$  

(3)

where $P_k = \mu_t (\partial u_i / \partial x_j)((\partial u_i / \partial x_j) + (\partial u_j / \partial x_i))$. The definition of $F_1$ in formula (2) is as follows:

$$F_1 = \tanh \left\{ \min \left[ \max \left( \frac{\sqrt{k}}{\beta^* \omega y}, \frac{500 \nu}{y^2 \omega} \right), \frac{4 \rho \sigma_{o2} k}{CD_{k\omega} y^2} \right] \right\}$$  

(4)

where $CD_{k\omega} = \max (2 \rho \sigma_{o2} (1 / \omega)(\partial k / \partial x_j)(\partial \omega / \partial x_j), 10^{-10})$, and $y$ is the distance from the calculation point to the wall. The eddy viscosity coefficient $\nu_t$ in formula (1) and (2) is defined as follows:

$$\nu_t = \frac{a_1 k}{\max (a_1 \omega, SF_2)}$$

where $S$ is the absolute value of vorticity, $F_2$ is the second mixed function, and the definition is as follows:

$$F_2 = \tanh \left[ \left( \max \left( \frac{\sqrt{k}}{\beta^* \omega y}, \frac{500 \nu}{y^2 \omega} \right) \right)^2 \right]$$

The closure coefficients of equations (1) and (2) are as follows: $\beta^* = 0.09$, $\alpha_1 = 5 / 9$, $\beta_1 = 3 / 40$, $\sigma_{k1} = 0.85$, $\sigma_{o1} = 0.5$, $\alpha_2 = 0.44$, $\beta_2 = 0.0828$, $\sigma_{k2} = 1$, $\sigma_{o2} = 0.856$. When Fluent 19.2 software runs, the turbulence model $k-\omega$ and $k-\varepsilon$ are transformed according to the mixing function $a = \alpha_1 F_1 + \alpha_2 (1 - F_1)$. In the numerical simulation of PAT, other main settings of Fluent 19.2 software include: The discrete scheme of pressure term in the control equation was set as the standard discrete scheme, the discrete scheme of momentum, turbulent kinetic energy and turbulent dissipation rate term were set as the second order upwind scheme, and the algebraic equations after
discretization of the control equation were solved by SIMPLEC, and the convergence accuracy was set as $10^{-6}$.

**Grid independence check.** The computational fluid domain was divided into structured grid. The block, mesh generation, refinement and grid independence check of computational fluid domain are shown in Figure 2. From Figure 2(a), because of the existence of triangular area in the impeller inlet and outlet of a single impeller channel, it is necessary to carry out Y-type segmentation for the two blocks. It is worth noting that because the cross sections of the import extension, the export extension and the volute are circular or quasi circular, the fluid domain structure of these three parts are segmented by O-type to improve the mesh quality, as shown in Figure 2(b) and (c). Since the division method of import extension and export extension is the same, the meshes of import extension are not shown in Figure 2(c). Based on the meshing of computational fluid domain, the efficiency

![Figure 2. Computational domain meshing of PAT flow field: (a) The block and meshing of single impeller channel; (b) The block of volute; (c) computational fluid domain meshing and refinement of PAT; (d) grid independence check.](image)
of PAT was taken as the variable, and the number of grids was taken as the independent variable to check the grid independence. The result of grid independence check is shown in Figure 2(d). As can be seen from Figure 2(d), the efficiency of PAT increases with the increase of grid number. When the grid number changes from 913,058 to 1,029,816, the efficiency of PAT is basically at a constant value, so the final grid number is 1,029,816. In order to capture the flow state more clearly in the near wall area of the blade before and after optimization, the grids checked for independence were refined in the near wall area with \(d = 0.2\) mm (the distance from the first layer grid node to the wall is 0.2 mm), as shown in Figure 2(c).

**Accuracy verification of numerical simulation.** The structural diagram and the test site of the accuracy verification test system for the optimization results and numerical simulation of the PAT blade are shown in Figure 3. From Figure 3, the working principle of the PAT test system is as follows:

The water feed pump is driven by an electric motor to provide fluid with a certain pressure for the PAT test. An electromagnetic flowmeter is installed on the pipeline between the water feed pump and the PAT to measure the flow through the PAT. The rated working pressure is 2 MPa, the measuring range is \(0.35 - 84.78\) m\(^3\)/h, and the measuring accuracy is \(\pm 0.1\)%. The pressure transmitters are installed on the inlet and outlet pipes of the PAT to measure the pressure at the inlet and outlet of the PAT. The measuring range is \(0 - 2\) Mpa and the measuring accuracy is \(\pm 0.1\)%. The mechanical energy converted from the PAT is consumed by the energy consume pump, and a torque meter is installed between them to measure the output torque of PAT and the speed of the PAT. The rated torque is \(100\) N \(\cdot\) m, the speed range can be measured is \(0 - 6000\) r/min, the calibration coefficient is 8522, and the accuracy level is 0.2.

Numerical simulation and experiment were carried out on the prototype PAT to obtain the comparison diagrams of the external characteristic curves, as shown in Figure 4.

It can be seen from Figure 4 that the trend of the external characteristic curves obtained by numerical simulation are basically consistent with that obtained by experiment. From Figure 4(a), there is a big difference between the small flow area and the large flow area, which is mainly caused by the low accuracy of the turbulence model for the flow simulation with separation and high reverse pressure gradient. It is generally considered that the error controllable region (ECR) is \(0.8Q-1.2Q\) (the middle part of the two dotted lines in the Figure 4), and the error at the optimal operating point is the smallest.

**Parameterization of twisted blade of PAT**

The shape of twisted blade of PAT is at least formed by three blade profiles of FS, middle and RS. Therefore, the shape of blade profile is very important for improving the performance of impeller. This also attributes the optimization of twisted blade to the optimization of three blade profiles, and the forming principle from blade profile to twisted blade surface determines the speed of blade modeling. The spatial blade profiles corresponding to the twisted blade of PAT and the forming method of twisted blade are shown in Figure 5.
In the Figure 5(b), \( \theta \) is the wrap angle of spatial streamline, the unit is rad, and the change direction has been marked; \( z \) is the axial change direction of spatial streamline; \( r \) is the radial change direction of spatial streamline. The FS blade profile of twisted blade is taken as example, the spatial streamline in flow surface diagram is projected to the meridional plane and the plane perpendicular to the axis respectively to form the meridional plane projection and plane projection of spatial blade profile. From the meridional plane projection of the blade profile, it can be concluded that the length \( \Delta m \) of the blade profile along the streamline of the meridional plane, the length \( \Delta z \) of the blade profile along the axial direction, and the length \( \Delta r \) of the blade profile along the radial direction constitute a right triangle with the hypotenuse as a curve. When the curve triangle is small enough, it can be regarded as a right triangle with the hypotenuse as a

**Figure 3.** Accuracy verification test system for PAT numerical simulation: (a) The structural diagram of the accuracy verification test system; (b) The test site of the accuracy verification test system.
straight line. There is a geometric relation \( \Delta m^2 = \Delta z^2 + \Delta r^2 \). Similarly, the following relation \( \Delta l^2 = \Delta u^2 + \Delta r^2 \) can be obtained from the plane projection of blade profile, where \( \Delta l \) represents the length of blade profile along the plane streamline, \( \Delta u \) represents the change length of arc length when the wrap angle changes to \( \Delta \theta \). Combined with the meridional plane and plane projection of the spatial blade profile, the variable length \( \Delta s \) of the spatial blade profile can be determined, that is, \( \Delta s^2 = \Delta l^2 + \Delta z^2 = \Delta u^2 + \Delta m^2 \). Therefore, the setting angle \( \beta_i \) at any position of the blade profile has the following relationship with \( \Delta u_i \) and \( \Delta m_i \), \( \tan \beta_i = \Delta m_i / \Delta u_i = \Delta m_i / r_i \Delta \theta_i \). Define \( \theta = \sum \Delta \theta_i, m = \sum_{0}^{i} \Delta m_i / r_i \). \( \theta \) is taken as the abscissa of the blade profile change curve and \( m \) is taken as the ordinate of the blade profile change curve, and the corresponding PAT blade profile plane expansion diagram was drawn. The three blade profiles of twisted

**Figure 4.** Numerical simulation and experimental external characteristic curves of prototype PAT: (a) flow-efficiency curves of PAT; (b) flow-power curves and flow-head curves of PAT.

**Figure 5.** Drawing principle of twisted blade of PAT: (a) twisted blade and blade profiles; (b) drawing principle of twisted blade.
blade of PAT were parameterized by Bezier spline curve, and the intersection of adjacent control points was defined as the independent variable of fitness function of surrogate model, as shown in Figure 6.

From Figure 6(b), the shape of each blade profile is controlled by five control points, and three independent variables can be determined. It is specified that the independent variables change along the angle bisector of the included angle of the control line, and the included angle increases to positive and decreases to negative. The nine independent variables determined by the three twisted blade profiles and their value ranges are shown in Table 2.

**Optimization method and strategy**

**SVM-HDMR surrogate model theory**

The construction idea of SVM-HDMR surrogate model can be described as follows: First, according to the HDMR theory,\(^3\) the output \(f(x)\) for any system can be expressed by the input \(x = [x_1, x_2, \ldots, x_n]^T \in \mathbb{R}^n\) as follows:

\[
f(x) = f_0 + \sum_{i=1}^{n} f_i(x_i) + \sum_{1 \leq i < j \leq n} f_{ij}(x_i, x_j) + \cdots + \sum_{1 \leq i_1 < \cdots < i_k \leq n, 2 < k < n} f_{i_1i_2\cdots i_k}(x_{i_1}, x_{i_2}, \ldots, x_{i_k}) + \cdots + f_{12\cdots n}(x_1, x_2, \ldots, x_n)
\]

where \(f_0\) is the zeroth order function term; \(f_i(x_i)\) is the first order function term; \(f_{ij}(x_i, x_j)\) is the second order function term; \(f_{i_1i_2\cdots i_k}(x_{i_1}, x_{i_2}, \ldots, x_{i_k})\) is the higher order function term.

![Figure 6](image-url). Parameterization of PAT blade profiles: (a) expanded plan of PAT blade profiles; (b) parameterization of blade profiles. BP: Blade profile; BPCPL: Blade profile control point line.
the second order function term; \( f_{12\ldots n}(x_1, \ldots, x_n) \) is \( n \) order function term. Since the second order function term constructed can meet the accuracy requirements, omitting the function term above the second order, the original expression can be written as follows.\(^{34,35} \)

\[
 f(x) = f_0 + \sum_{i=1}^{n} f_i(x_i) + \sum_{1 \leq i < j \leq n} f_{ij}(x_i, x_j) \quad (8)
\]

Secondly, the solution method of cutting center point \( x_0 \)\(^{36} \) is introduced to express each term of output function \( f(x) \) as a function containing \( x_0 \), as follows:

\[
 f_0 = f(x_0) \quad (9)
\]

\[
 f_i(x_i) = f(x_i, x_i^0) - f_0 \quad (10)
\]

\[
 f_{ij}(x_i, x_j) = f(x_i, x_j, x_{ij}^0) - f_i(x_i) - f_j(x_j) - f_0 \quad (11)
\]

where \( (x_i, x_i^0) \) is the same as the value of \( x_0 \) except for the \( i \)-th dimension which can be sampled, that is, point \( (x_1^0, \ldots, x_i^{0-1}, x_i^0, x_i^{0+1}, \ldots, x_n^0) \).

Finally, according to the SVM theory,\(^{37,38} \) when the model training set is linear or approximately linear, the regression function can be constructed as follows:

\[
 \hat{f}(x) = y = (w \cdot x) + b \quad (12)
\]

where \( w \) is the normal direction of the hyperplane, and \( b \) is the translation distance of the hyperplane along the \( y \) axis. When the model training set is nonlinear, it is necessary to map the training set to the high dimensional feature space by mapping \( \Phi: x \mapsto \Phi(x) \), and introduce the kernel function \( K(x_i, x_j) = (\Phi(x_i), \Phi(x_j)) \), and then make an approximate linear regression on the training set. In conclusion, the SVM-HDMR surrogate model of PAT is as follows:

\[
 f(x) = f_0 + \sum_{i=1}^{n} \hat{f}_i(x_i) + \sum_{1 \leq i < j \leq n} \hat{f}_{ij}(x_i, x_j) \quad (13)
\]

The construction process of PAT surrogate model is shown in Figure 7.
**Objective function and optimization strategy**

Through the analysis of the PAT surrogate model function, the optimization problem of the maximum value of the objective function can be described as follows:

$$\max f(x)$$

s.t.  

$$x = [x_1, x_2, \ldots, x_n]^T \in \mathbb{R}^n$$

(14)

The objective function is globally searched in the domain of definition, and the maximum value is obtained by combining with genetic algorithm.\(^{39}\) The specific optimization strategy of twisted blade of PAT is described as follows:

Step 1. According to the coordinates of each point of twisted blade of PAT, the twisted blade is represented by three blade profiles of FS, Middle and RS by Creo 3.0 and CFturbo 2020 software.

Step 2. Three blade profiles are parameterized by Bezier curve with MATLAB 9.7 software programming function, and the training set of SVM-HDMR surrogate model is separated.

Step 3. According to the separated training set of surrogate model, and the construction process of PAT SVM-HDMR surrogate model shown in Figure 7, the surrogate model function based on PAT efficiency is constructed by comprehensively using MATLAB 9.7, CFturbo 2020 and Fluent 19.2 software.

![Figure 7](image)

**Figure 7.** The construction process of PAT surrogate model. CFD: Computational fluid dynamics.
Step 4. Genetic algorithm is used to search the surrogate model function globally in the optimization interval on the MATLAB 9.7 software platform, and the parameters of three blade profiles are fed back according to the optimization results.

When the surrogate model function obtains the optimal solution in the optimization interval, the optimization process ends. The three optimized blade profiles are modeled by CFturbo 2020 software, and the optimization results of the surrogate model are verified by numerical simulation. Experiment is conducted to verify the effectiveness of the optimization results and numerical simulation. When the surrogate model function obtains the optimal solution at the boundary of the optimization interval, it is necessary to repeat steps (3)–(4) based on the optimization result until the conditions of the end of the optimization process are satisfied.

Process, results, and discussion

The construction process of PAT surrogate model

According to the function characteristics of SVM-HDMR surrogate model, at least three function values are needed to determine a first order SVM-HDMR function besides the prototype model, that is, two efficiency values of PAT hydraulic model with single independent variable change are needed. To determine a second order SVM-HDMR function, at least five function values are needed, that is, four efficiency values of PAT hydraulic model with two independent variables changing each other are needed.

The first order and second order SVM-HDMR function terms constructed when the blade profile of the FS changes is taken as examples. MATLAB 9.7 software was used to generate Bezier curve of blade profile corresponding to the change of independent variables $x_1$, $x_2$ and $x_3$, and the Bezier curve was imported into CFturbo 2020 software to form the computational fluid domain of impeller. The parameters of Fluent 19.2 software were set to calculate the head and power of the corresponding PAT, and the efficiency of the PAT hydraulic model was solved. The calculated training points of surrogate model and the corresponding efficiency values of PAT hydraulic model are shown in Table 3.

The model number 0 represents the prototype model, and the model number 1-18 is the corresponding PAT model formed when the independent variables $x_1$, $x_2$ and $x_3$ change (the other six independent variables do not change). According to the data of model 0-6 in the Table 3, three first order surrogate model function terms $f_1(x_1), f_2(x_2)$ and $f_3(x_3)$ could be constructed for the FS blade profile control variables. Three second order surrogate model function terms $f_{12}(x_1, x_2), f_{13}(x_1, x_3)$ and $f_{23}(x_2, x_3)$ could be constructed from model 0 and model 7-18 in the Table 3. The images of six SVM-HDMR surrogate model functions constructed are shown in Figure 8. From Figure 8(a)–(c), the first order function terms constructed are all nonlinear functions.

Optimization results of PAT surrogate model

The function terms of the middle and the RS blade profile were constructed in the same way. In this way, after learning 55 PAT hydraulic models, the SVM-HDMR surrogate model function $f(x)$ based on PAT efficiency was formed according to formula (13).
The genetic algorithm was used to find the optimal solution in the whole domain, and whether the optimization process is terminated is determined according to “whether the optimal solution is obtained in the optimization interval”. As the SVM-HDMR

Table 3. Training point of surrogate model and efficiency of PAT.

| No. | $x_1$ | $x_2$ | $x_3$ | $\eta / \%$ | No. | $x_1$ | $x_2$ | $x_3$ | $\eta / \%$ |
|-----|-------|-------|-------|-------------|-----|-------|-------|-------|-------------|
| 0   | 0     | 0     | 0     | 77.16       | 10  | $-10^\circ$ | $-10^\circ$ | 0     | 77.859      |
| 1   | $10^\circ$ | 0     | 0     | 77.555      | 11  | $10^\circ$ | 0     | $10^\circ$ | 77.826      |
| 2   | $-10^\circ$ | 0     | 0     | 77.821      | 12  | $10^\circ$ | 0     | $-10^\circ$ | 77.562      |
| 3   | 0     | $10^\circ$ | 0     | 77.672      | 13  | $-10^\circ$ | 0     | $10^\circ$ | 78.05       |
| 4   | 0     | $-10^\circ$ | 0     | 77.699      | 14  | $-10^\circ$ | 0     | $-10^\circ$ | 77.824      |
| 5   | 0     | 0     | $10^\circ$ | 77.703      | 15  | 0     | $10^\circ$ | $10^\circ$ | 77.498      |
| 6   | 0     | 0     | $-10^\circ$ | 77.425      | 16  | 0     | 0     | $-10^\circ$ | 77.432      |
| 7   | $10^\circ$ | $10^\circ$ | 0     | 77.187      | 17  | 0     | $-10^\circ$ | $10^\circ$ | 77.926      |
| 8   | $10^\circ$ | $-10^\circ$ | 0     | 77.761      | 18  | 0     | $-10^\circ$ | $-10^\circ$ | 77.712      |
| 9   | $-10^\circ$ | $10^\circ$ | 0     | 77.773      |     |       |       |       |             |

Figure 8. The first order and second order SVM-HDMR function terms determined by FS: (a) $\hat{f}_1(x_1)$; (b) $\hat{f}_2(x_2)$; (c) $\hat{f}_3(x_3)$; (d) $\hat{f}_{12}(x_1, x_2)$; (e) $\hat{f}_{13}(x_1, x_3)$; (f) $\hat{f}_{23}(x_2, x_3)$.  


The surrogate model function $f(x)$ of the first construction achieved the maximum value, some of the independent variables were at the boundary of the optimization interval, which showed that the blade profile had the optimized space. It was necessary to construct the SVM-HDMR surrogate model function based on the first optimized blade profile. After the algorithm learning 109 PAT hydraulic models, the SVM-HDMR surrogate model based on PAT efficiency was constructed twice. The genetic algorithm was run twice to optimize the surrogate model function in the whole domain. The optimal efficiency value of PAT predicted by the surrogate model after one iteration was 80.52%, and the control variable value of feedback blade profiles could get the comparison diagrams of blade profiles and the corresponding three-dimensional comparison diagrams of twisted blade before and after optimization, as shown in Figure 9.

Figure 9(a)–(c) are the comparison diagrams of blade profiles changes of FS, middle and RS before and after optimization. Figure 9(d) is the comparison diagram of three-dimensional PAT twisted blade spatial blade profile before and after optimization. Figure 9(e) is the comparison diagram of three-dimensional PAT twisted blade before and after optimization (FS is included in the diagram for comparison). Figure 9(f) is the comparison diagram of the corresponding PAT twisted blade impeller for test before and after optimization. From Figure 9(a), the optimized PAT blade profile on the FS side has no obvious change at the blade inlet (the outlet of the centrifugal pump blade), which is shown in the comparison diagram of the three-dimensional blade in Figure 9(e), and there is a staggered zone at the inlet of the PAT blade on the FS side (area A in the Figure 9(e)). From Figure 9(a)–(c), the bending degree of the optimized twisted blade is smaller on the FS side, and the variation of blade profile from FS to RS increases gradually. From Figure 9(d), the variation law of the three-dimensional spatial blade profile of the twisted blade is consistent with that shown in Figure 9(a)–(c). From Figure 9(e), the offset of optimized PAT twisted blade on RS side is more than 1.5 blade thickness (area B in the Figure 9(e)).

**Flow field analysis**

In order to analyze the change of fluid flow state in the impeller before and after blade optimization, the surface streamline diagrams of $Z=0$, $Z=6$ mm and $Z=17$ mm before and after blade optimization were drawn, as shown in Figure 10.

The (1)–(3) in Figure 10 respectively represent the positions of each streamline plane in the impeller, the streamline diagrams of each surface before blade optimization and the streamline diagrams of each surface after blade optimization. From Figure 10(a) (2)–(3), the changes of flow conditions in the impeller before and after blade optimization mainly occur at the blade inlet (area C in the Figure 10(a) (2)) and blade outlet (area D in the Figure 10(a) (2)). Although the flow after optimization is more orderly, it is not obvious. From Figure 10(b) (2)–(3), the impeller before blade optimization has a flow separation zone at the blade outlet (area E in the Figure 10(b) (2)). After the fluid flows out of the blade area, there is an obvious velocity boundary (area F in the Figure 10(b) (2)). The existence of the separation zone weakens the work ability of the blade, and the existence of the velocity boundary increases the flow loss. From
Figure 10(c) (2)–(3), when the fluid flows to the outlet of the impeller, there are obvious differences in the velocity size at the same radius, and there are different directions in the velocity at different radii of the same axial plane. The differences have disappeared in the impeller after blade optimization.
The twisted blade impeller before and after optimization was numerically simulated at the design operating point, and the process of fluid energy conversion and transportation in the impeller was analyzed by combining with the kinetic energy equation of viscous fluid. The kinetic energy equation of steady incompressible fluid can be written as follows:

\[
\frac{1}{2} \frac{\partial (u_i u_j)}{\partial x_j} = u_i F_i + 2 \nu \frac{\partial (u_i s_{ij})}{\partial x_j} - \frac{1}{\rho} \frac{\partial (u_i p)}{\partial x_i} - 2 \nu s_{ij} s_{ij}
\]  

Here, \( F_i \) is the body force; \( \nu \) is the kinematic viscosity coefficient; \( p \) is the pressure; \( s_{ij} \) is the strain rate tensor.

Figure 10. Comparison of surface velocity streamline diagrams before and after impeller optimization: (a) \( Z = 0 \) plane; (b) \( Z = 6 \) mm plane; (c) \( Z = 17 \) mm plane. VS: Velocity streamline.
The first term at the right end of the formula is the work done by the body force to the fluid of unit mass in unit time. The body force includes gravity, electromagnetic force, etc., which will not be discussed here. The cloud images of the other terms at the right end of the formula and the left term of the formula at the plane with the axial coordinate of $Z = -2.1$ mm (the PAT impeller inlet width is 9.5 mm, PAT blade outlet position is $Z = -2.4$ mm, and the cloud forming plane is close to the RS) are shown in Figure 11.

Figure 11 (1) represent the internal flow cloud chart of the prototype PAT impeller, and Figure 11 (2) represent the internal flow cloud chart of the optimized impeller. The second term at the right end of the formula is the mechanical energy transported by the viscous force to the moving unit mass fluid in unit time, which represents the ability of the viscous force to change the spatial distribution of energy in the impeller. From Figure 11(a), under the same conditions, the internal transport intensity of the optimized impeller is weakened. The third term at the right end of the formula is the work done by the pressure per unit time to the fluid per unit mass, that is, the flow work. From Figure 11(b) (1)–(2), this term mainly acts on the blade inlet. From Figure 11(b), under the same conditions, the high intensity area of the optimized impeller cloud distribution is significantly reduced. The fourth term at the right end of the formula is the deformation work done by the viscous force per unit time, which is essentially different from other terms. By expanding this term, it is always positive, which means that at any
position in the impeller space, some of the energy of fluid motion is converted into heat energy and dissipated, also known as the dissipation term. From Figure 11(c), under the same conditions, the area of high dissipation intensity of the optimized impeller is obviously reduced, which shows that the loss of the optimized impeller in the process of energy conversion is small. The left term of the formula is the convective transport term of kinetic energy. From Figure 11(d), under the same conditions, the transport intensity is obviously weakened, and the high-intensity region of transport is obviously reduced. It indicates that the amount of energy converted from kinetic energy to pressure energy will decrease, which indirectly indicates that the final viscous dissipation caused by transport in the impeller will decrease, represents the enhancement of the impeller’s ability to do work on liquid.

Figure 12. Performance curves of PAT: (a) flow-efficiency comparison diagram obtained by numerical simulation; (b) flow-power and flow-head comparison diagram obtained by numerical simulation; (c) flow-efficiency comparison diagram obtained by test; (d) flow-power and flow-head comparison diagram obtained by test BO: before optimization; AO: after optimization.
In summary, the analysis of Figure 11(a), (b) and (d) show that the energy transport strength of the twisted blade impeller is weakened after optimization, which confirms the reduction of energy dissipation in the middle wheel of Figure 11(c). It indicates that the twisted blade impeller is more favorable for the transformation of fluid energy.

**Verification of optimization results**

The optimized PAT was numerically simulated and verified on the experimental equipment in Figure 3. The flow-efficiency, flow-head, and flow-power performance curves of the PAT before and after optimization were drawn, as shown in Figure 12.

Figure 12(a) and (b) are the comparison diagrams of performance curves of PAT obtained by numerical simulation. Figure 12(c) and (d) are the comparison diagrams of performance curves of PAT obtained through test. From Figure 12(a)–(c), the numerical simulation results are basically consistent with the experimental results. From Figure 12(a), the efficiency of the optimized PAT at the design operating point is 3.51% higher than that before optimization. From Figure 12(b), under design condition, the shaft power of the optimized PAT is increased by 5.3% compared with that before optimization. From Figure 12(c), the test efficiency of the optimized PAT at the design operating point is 3.4% higher than that before optimization. From Figure 12(d), the test output power of the optimized PAT at the design operating point is increased by 5.1% compared with that before optimization.

**Conclusions**

By redefining the variable of twisted blade profile to meet the condition of rapid modeling, the optimization of twisted blade of PAT was completed by using SVM-HDMR surrogate model combined with genetic algorithm. The optimization results were analyzed by using the velocity streamline diagram and the cloud diagram formed by kinetic energy equation. The optimization results were verified by numerical simulation and experiment. The following conclusions are drawn:

1. Comparing the velocity streamlines of the impeller before and after optimization, the flow field of the optimized impeller has obviously improved at the outlet of the blade. From the view of kinetic energy equation, the energy transport intensity of the optimized impeller is weakened, which indicates that the optimized impeller is more conducive to energy conversion.

2. The optimization results are verified by numerical simulation. The results show that the simulation efficiency of the PAT after optimization at the design operating point is 3.51% higher than the efficiency of the PAT before optimization, and the output power is increased by 5.3%.

3. The optimization results are verified by experiment. The results show that the results of numerical simulation and experiment are basically consistent. The test efficiency of the PAT after optimization at the design operating point is 3.4% higher than the efficiency of the PAT before optimization, and the output power is increased by 5.1%.
Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Natural Science Foundation of China, Grant NO. 52169019; Industrial Support Program of Colleges in Gansu Province, Grant NO. 2020C-20; Gansu Outstanding Youth Fund, Grant NO. 20JR10RA203; Open Fund of Key Laboratory of Fluid and Power Machinery, Ministry of Education, Xihua University, Grant NO. szjj2019-016.

ORCID iDs

Bingxiao Jiang https://orcid.org/0000-0002-5938-0437
Fengxia Shi https://orcid.org/0000-0003-4170-3450

References

1. Lin T, Zhu ZC, Li XJ, et al. Performance prediction of centrifugal pump as turbines-a review and prospect. J Drain Irrig Mach Eng 2021; 39: 562–568.
2. Armando C, Shahram DH and Helena MR Pumps as turbines fundamentals and applications. Berlin, Germany: Springer, 2018.
3. Yang JH, Ma QH, Luo YT, et al. Influence of geometric parameters on head-flow curve of pump as turbine. Trans Chin Soc Agric Mach 2019; 50: 167–172.
4. Liu M, Tan L and Cao SL. Theoretical model of energy performance prediction and BEP determination for centrifugal pump as turbine. Energy 2019;172:712–732.
5. Guan XF. Modern pumps theory and design. Beijing, Chinese: China Astronautic Publishing House, 2011.
6. Zhang RH, Guo GQ, Yang JH, et al. Optimization research on low specific speed centrifugal impeller based on incomplete sensitivities method. J Mech Eng 2014; 50: 162–166.
7. Miao SC, Yang JH, Wang XH, et al. Blade pattern optimization of the hydraulic turbine based on neural network and genetic algorithm. J Aerosp Power 2015; 30: 1918–1925.
8. Wang CL, Feng YM and Ye J. Multi-objective parameters optimization of centrifugal slurry pump based on RBF neural network and NSGA-II genetic algorithm. Trans Chin Soc Agric Eng 2017; 33: 109–115.
9. Wang CL, Hu BB, Feng YM, et al. Multi-objective optimization of double vane pump based on radial basis neural network and particle swarm. Trans Chin Soc Agric Eng 2019; 35: 25–32.
10. Zhang RH, Guo R, Yang JH, et al. Inverse method of centrifugal pump impeller based on proper orthogonal decomposition (POD) method. J Mech Eng 2017; 30: 1025–1031.
11. Zhang RH, Chen XB, Guo GQ, et al. Reconstruction and modal analysis for flow field of low specific speed centrifugal pump impeller. Trans Chin Soc Agric Mach 2018; 49: 143–149.
12. Guo GQ, Zhang RH and Chen XB. Optimization method for low specific speed centrifugal impeller based on POD surrogate model. J Huazhong Univ Sci & Tech (Nat Sci Ed) 2019; 47: 50–55.
13. Wang CL, Peng HB, Ding J, et al. Optimization for S-type blade of fire pump based on response surface method. Trans Chin Soc Agric Eng 2013; 49: 170–177.
14. Yuan SQ, Wang WJ, Pei J, et al. Multi-objective optimization of low-specific-speed centrifugal pump. Opti Trans Chin Soc Agric Eng 2015; 31: 46–52.
15. Zhang RH, Guo M, Yang JH, et al. Blade shape optimization of centrifugal pump impeller with adjoint method. J Drain Irrig Mach Eng 2014; 32: 943–954.
16. Miao SC, Zhang HB, Wang TT, et al. Optimal design of blade in pump as turbine based on multidisciplinary feasible method. Sci Progress 2020; 103: 1–16.
17. Chahine C, Seume JR and Verstraete T. The influence of metamodeling techniques on the multidisciplinary design optimization of a radial compressor impeller. In: Proc ASME Turbo Expo 2012, Copenhagen, Denmark, 2012, paper no. GT 2012–68358.
18. Rostami O and Kaveh M. Optimal feature selection for SAR image classification using biogeography-based optimization (BBO), artificial bee colony (ABC) and support vector machine (SVM): a combined approach of optimization and machine learning. Comput Geosci 2021; 25: 911–930.
19. Ali L, Wajahat J, Golilarz NA, et al. LDA–GA–SVM: improved hepatocellular carcinoma prediction through dimensionality reduction and genetically optimized support vector machine. Neural Comput & Applic 2021; 33: 2783–2792.
20. Le DN, Parvathy VS, Gupta D, et al. Iot enabled depthwise separable convolution neural network with deep support vector machine for COVID-19 diagnosis and classification. Int J Mach Learn & Cyber 2021; 12: 3235–3248.
21. Andres E, Salcedo S, Monge F, et al. Efficient aerodynamic design through evolutionary programming and support vector regression algorithms. Expert Syst Appl 2012; 39: 10700–10708.
22. Tang L, Li GY and Wang H. Kriging-HDMR metamodeling technique for nonlinear problem. Chin J Theor Appl Mech 2011; 43: 781–784.
23. Li WP, Dou XD, Wang ZX, et al. BPNN-HDMR Nonlinear metamodeling technique and its application. J Hunan Univ (Nat Sci) 2014; 41: 32–38.
24. Chen LM, Wang H, Ye H, et al. Comparative study of HDMRs and other popular metamodeling techniques for high dimensional problems. Struct Multidiscipl Optim 2019; 59: 21–42.
25. Jiang BX, Yang JH, Bai XB, et al. Optimization of centrifugal pump blade based on high-dimensional hybrid model and genetic algorithm. J Huazhong Univ Sci & Tech (Nat Sci Ed) 2020; 48: 128–132.
26. Menter FR. Improved two-equation turbulence models for aerodynamic flows. NASA Tech Memo 1992; 103975: 1–31.
27. Menter FR, Kuntz M and Langtry R. Ten years of industrial experience with the SST turbulence model. Turbul Heat Transf 2003; 4: 625–632.
28. Li XJ, Yuan SQ, Pan ZY, et al. Realization and application evaluation of near-wall mesh in centrifugal pumps. Trans Chin Soc Agric Eng 2012; 28: 67–72.
29. David CW. Turbulence modeling for CFD. 3rd ed. San Diego, USA: DCW Industries, 2006.
30. Johann FG. Centrifugal pumps. 4th ed. Berlin, Germany: Springer, 2020.
31. Sobol IM. Sensitivity estimates for nonlinear mathematical models. Math Model Comput Exp 1993; 1: 407–414.
32. Rabitz H and Alis ÖF. General foundations of high dimensional model representations. J Math Chem 1999; 25: 197–233.
33. Alis ÖF and Rabitz H. Efficient implementation of high dimensional model representations. J Math Chem 2001; 29: 127–142.
34. Xiong HD, Chen ZZ, Qiu HB, et al. Adaptive SVR-HDMR metamodeling technique for high dimensional problems. In: 2012 Int Con Model ID Control, AASRI Procedia. Amsterdam: Elsevier, 2012, pp.95–100.
35. Li L and Sun T. SVM-HDMR approximation model construction method for high dimensional nonlinear problems. Comput Eng Appl 2013; 49: 6–9.
36. Tunga MA. An approximation method to model multivariate interpolation problems: indexing HDMR. *Math Comput Model* 2011; 53: 1970–1982.

37. Vapnik VN. *Statistical learning theory*. New York, USA: John Wiley & Sons, Inc, 1998.

38. Yang XW and Hao ZF. *Algorithm design and analysis of support vector machine*. Beijing, Chinese: Science Press, 2013.

39. Chiho K, Rohit B, Chen LH, et al. Polymer design using genetic algorithm and machine learning. *Comput Mater Sci* 2021; 186: 110067.

40. Zhang ZX and Dong ZN. *Viscous fluid mechanics*. 2nd ed. Beijing, Chinese: Tsinghua University Press, 2011.