Optimization of pumps as turbines blades based on SVM-HDMR model and PSO algorithm

Bingxiao Jiang¹, Junhu Yang¹,², Xiaohui Wang¹,², Senchun Miao¹,² and Xiaobang Bai¹,³

Abstract
In view of the poor performance of pumps as turbines (PAT) operation, and the problem that the structural parameters cannot be optimized in the whole domain, the hybrid model of support vector machine (SVM) model and high-dimensional model representation (HDMR) method is applied to the optimization of PAT blade. Specifically, a PAT was selected, and the surrogate model for PAT blade optimization was constructed with MATLAB, Creo, and ANSYS software. The particle swarm optimization (PSO) algorithm was used to predict the performance data by global optimization. Finally, numerical prediction and experimental methods were used to verify the predicted data. These proved the applicability of the hybrid model in the optimization of fluid machinery. The numerical simulation results show that at the optimal operating point, the numerical simulation efficiency of the optimized PAT is 5.49% higher than that of the prototype PAT, and the output power is 7.2% higher. The test results show that the external characteristic curve of the numerical simulation PAT is basically consistent with the test results. At the optimal operating point, the test efficiency of the optimized PAT is 5.1% higher than that of the prototype PAT, and the output power is 6.9% higher.

Keywords
Pumps as turbines, support vector machine, high-dimensional model representation, surrogate model, particle swarm optimization, numerical simulation, test results

Date received: 16 March 2021; accepted: 23 June 2021
Handling Editor: James Baldwin

Introduction
There are a lot of high-pressure residual energy liquid in many process industries. Pumps as turbines (PAT) is the basic way to recover this part of energy because of its low cost, reliable operation, and convenient maintenance.¹² Due to the design reasons, the operation efficiency of PAT is generally low. As the main part of PAT energy conversion, impeller performance directly determines the hydraulic loss of impeller, and then determines the efficiency of PAT. Traditional optimization methods of fluid machinery include semi-empirical and semi-theoretical formula optimization method,³ experimental design method,⁴ and approximate model method.⁵ With the development of computer and the research of many computational surrogate models, such as support vector machine (SVM),⁶ artificial neural...
network (ANN),\(^7\) radial basis function (RBF),\(^8\) etc., the complex computational work is greatly reduced. Especially the combination of surrogate model and intelligent optimization algorithm, which lays the foundation of modern optimization design of fluid machinery, and provides an exploratory method for the optimization of PAT.

In recent years, the high-dimensional model representation (HDMR), which is built to solve nonlinear mathematical problems, has been applied in engineering. In Tang et al.,\(^9\) HDMR was applied to the optimization of the front longitudinal welding structure of the vehicle body to solve the problem of the deformation of the longitudinal section thickness. In Li et al.,\(^10\) HDMR was applied to the optimization of the cab of the mining dump truck to improve the reliability of the cab. SVM is a machine learning tool, which is widely used in big data research. Compared with the single machine learning method, SVM-HDMR model ensures the learning accuracy and reduces the sample size of learning.\(^11\)

A MH48-12.5 PAT was selected. The blade profile of PAT was taken as the optimization object and the efficiency of PAT was taken as the objective function, the efficiency adaptive function to blade profile parameters was established by using SVM-HDMR surrogate model. particle swarm optimization (PSO) algorithm was used to optimize the fitness function, and numerical simulation and experiment were used to verify the prediction data of SVM-HDMR surrogate model.

The structure of the article is as follows. First, the principle of SVM-HDMR surrogate model is introduced. Secondly, combined with the characteristics of PAT, the construction process of PAT SVM-HDMR surrogate model and optimization strategy of PAT are presented. Finally, the specific optimization process of the example and the internal flow field analysis and experimental verification of the optimization results are given.

### Optimization method and strategy

#### SVM-HDMR surrogate model theory

The idea of SVM-HDMR surrogate model can be simply expressed as: For a certain black box system, the output function \(f(x)\) of the system is expressed as a function of the input variables \(x = [x_1, x_2, \ldots, x_n]^T \in \mathbb{R}^n\) according to the HDMR theory. Secondly, the specific form of the function in each dimension and the specific form of the coupling function between dimensions need to be fitted out according to SVM theory. The specific theory of SVM-HDMR surrogate model is described as follows.

#### HDMR theory

HDMR is a method to accurately express the influence of multivariable coupling on the output of physical system, which is based on the idea of “divide and conquer,” proposed by Sobol,\(^12\) and improved by Rabitz Alis,\(^13\) Alis and Rabitz,\(^14\) Tunga,\(^15\) etc. According to HDMR theory, in n-dimensional space, the input variables \(x = [x_1, x_2, \ldots, x_n]^T \in \mathbb{R}^n\) and corresponding output function \(f(x)\) have the following mapping relationship:

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

where \(f_0\) is the zero-order function term, constant; \(f_i(x_i)\) is the first-order function term, which is the independent contribution of the variable \(x_i\) to \(f(x)\); \(f_{ij}(x_i, x_j)\) is the second-order function term, which is the contribution of the coupling of variables \(x_i\) and \(x_j\) to the function \(f(x)\); \(f_{12\ldots n}(x_1, x_2, \ldots, x_n)\) is the contribution of all input variables to the function \(f(x)\).

Rabitz introduced the solution of Cut-HDMR based on the original HDMR, that is, every component function in HDMR can be represented by cutting point \(x_0\).

\[
f_0 = f(x_0) \quad (2)
\]

\[
f_i(x_i) = f(x_i, x_0) - f_0 \quad (3)
\]

\[
f_0(x_i, x_j) = f(x_i, x_j, x_0) - f(x_i) - f(x_j) - f_0 \quad (4)
\]

\[
f_{ijk}(x_i, x_j, x_k) = f_{ijk}(x_i, x_j, x_k, x_0) - f_{ij}(x_i, x_j) - f_{ik}(x_i, x_k) - f_{jk}(x_j, x_k) - f_{i}(x_i) - f_{j}(x_j) - f_{k}(x_k) - f_0 \quad (5)
\]

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

#### SVM theory

According to SVM theory,\(^16,17\) training set \(T = \{(x_1, y_1), (x_2, y_2), \ldots, (x_l, y_l)\}\) is known, in which \(x_i = \mathbb{R}^n, y_i = \mathbb{R}, i = 1, 2, \ldots, l\). When the training set is linear or approximately linear, there is an approximate linear regression function:

\[
\hat{f}(x) = y = (w \cdot x) + b \quad (6)
\]
Where \( w \) is the normal direction of the hyperplane and \( b \) is the translation distance of the hyperplane along the \( y \) axis.

When the training set is non-linear, the training set is mapped to the high-dimensional feature space according to the mapping relationship of \( \Phi : x \mapsto \Phi(x) \). Kernel function \( K(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle \) is introduced to make approximate linear regression for the training set in the high-dimensional feature space.

Hence, the SVM-HDMR surrogate model constructed by SVM theory and HDMR theory is as follows:

\[
\begin{align*}
f(x) &= f_0 + \sum_{i=1}^{n} f_i^*(x_i) + \sum_{1 \leq i < j \leq n} \tilde{f}_{ij}(x_i, x_j) \\
&\quad \cdots \quad \sum_{1 \leq i_1 < \cdots < i_k \leq n, 2 \leq k < n} \tilde{f}_{i_1i_2\cdots i_k}(x_i, x_{i_2}, \cdots, x_{i_k}) \\
&\quad \cdots + \tilde{f}_{12\cdots n}(x_1, x_2, \cdots, x_n)
\end{align*}
\]

(7)

Construction process of SVM-HDMR surrogate model for PAT

According to the SVM-HDMR theory, the construction process of the specific PAT SVM-HDMR surrogate model is described below.

Select cutting point. The initial blade profile parameter of the PAT is selected as the cutting point \( x_0 = [x_{10}, x_{20}, \cdots, x_{m0}]^T \). The efficiency of the PAT at the cutting center point is calculated by CFD, which is recorded as \( f_0 \).

Construct SVM model of first order function terms in HDMR. For the convenience of narration, the training points of blade shape as \( (x_{i0}, x_{i0}^0) \) and efficiency as \( \{ (x_{i0}, x_{i0}^0), f_i(x_{i0}) \} \) are defined as training points.

Generate training space. According to the value range of independent variable \( x_i \), the blade training space is generated, and the corresponding PAT efficiency value is calculated by CFD, that is, the corresponding PAT efficiency values \( f_i(x_{i0}, x_{i0}^0) \) and \( f_i(x_{i0}, x_{i0}^0) \) of points \( (x_{i0}, x_{i0}) \) and \( (x_{i0}, x_{i0}^0) \) are calculated by CFD.

Construct the one-dimensional linear SVM model function. According to the \( t \)th dimension training points \( \{ (x_{i0}, x_{i0}^0), f_i(x_{i0}, x_{i0}^0) - f_0 \} \) and \( \{ (x_{i0}, x_{i0}^0), f_i(x_{i0}, x_{i0}^0) - f_0 \} \), the corresponding SVM model function \( f_i(x_t) \) is determined.

Judge the linear degree of SVM model function \( \hat{f}_i(x_t) \). Substitute the \( t \)th dimension cutting point into \( \hat{f}_i(x_t) \), if the following formula is satisfied:

\[
\left| \frac{\hat{f}_i(x_{i1}) - \hat{f}_i(x_{i2})}{x_{i1} - x_{i2}} \right| \leq \varepsilon \quad (8)
\]

Where \( \varepsilon \) is the allowable error value. The model function is a linear function, otherwise, it is constructed based on kernel function.

Verify the accuracy of first order model function \( \hat{f}_i(x_t) \). According to the training points \( (x_{i0}, x_{i0}^0) \) and \( (x_{i1}, x_{i1}^0) \), the model function precision detection point \( (x_{i0} + \frac{x_{i1} - x_{i0}}{3}, x_{i0}^0) \) is generated. \( f_i(x_0 + \frac{x_{i1} - x_{i0}}{3}) \) is calculated by CFD simulation, if the following formula is satisfied:

\[
\left| f_i(x_0 + \frac{x_{i1} - x_{i0}}{3}) - f_0 - \hat{f}_i(x_0 + \frac{x_{i1} - x_{i0}}{3}) \right| \leq \varepsilon \quad (9)
\]

then the constructed model function \( \hat{f}_i(x_t) \) meets the precision requirement, and the construction of the first-order model functions are completed.

Construct SVM model of second order function terms in HDMR

Check the existence of second-order function terms. For each independent variable \( x_i \), randomly select a sample point \( (x_{i0}, x_{i0}^0) = [x_{i0}, x_{i0}, \cdots, x_{i0}, \cdots, x_{m0}]^T \), where \( x_{i0} \neq x_{i0} \), synthesize the sample points of each variable, generate the second-order function test sample point \( x_0 = [x_{i0}, x_{i0}, \cdots, x_{i0}, \cdots, x_{m0}]^T \), and calculate the performance function value \( f(x_0) \) corresponding to the blade profile at the test sample point through computational fluid dynamics (CFD), if it meets the following requirements:

\[
\left| f(x_0) - f_0 - \sum_{i=1}^{n} f_i(x_{i0}) \right| \leq \varepsilon \quad (10)
\]

shows that there are no second-order function terms, otherwise construct the second-order function terms.

Check whether any two variables have second-order function term. For any two variables \( x_i \) and \( x_j \), the combined sample point \( (x_{i0}, x_{i0}, x_{j0}, x_{j0}) \) is selected, and the performance function value \( f(x_{i0}, x_{i0}, x_{j0}, x_{j0}) \) corresponding to the blade profile at the combined sample point is calculated by CFD simulation. If satisfied:

\[
\left| f(x_{i0}, x_{i0}, x_{j0}, x_{j0}) - f_0 - f_i(x_{i0}) - f_j(x_{j0}) \right| \leq \varepsilon \quad (11)
\]

then there is no second-order function term between variables \( x_i \) and \( x_j \). Otherwise, the second-order function term between any two variables is constructed.
The performance function values \( f_i(x_i, x_j) \) are calculated from the value range of variables \( x_i \) and \( x_j \). The performance function values \( f_i(x_i, x_j, x'_0) \), \( f_i(x_i, x'_0, x'_0) \), and \( f(x_i, x_j, x'_0) \) corresponding to the blade profile at the combination training point are calculated by CFD, and the second-order function term \( f_0(x_i, x_j) \) is constructed by combining the training point \( \{ (x_i, x'_0), f(x_i, x_j, x'_0) - f_0 - f(x_i) - f(x_j) \} \) and the cutting center point.

Verify the accuracy of second-order function term \( f_0(x_i, x_j) \). According to the training points \( (x_i, x'_0, x'_0) \) and \( (x_i, x_j, x'_0) \), the second-order model function accuracy test point \( (x_i + \frac{x_i - x'_0}{3}, x_j + \frac{x_j - x'_0}{3}, x'_0) \) is generated, and the performance function value \( f(x_0 + \frac{x_i - x'_0}{3}, x_j + \frac{x_j - x'_0}{3}, x'_0) \) corresponding to the blade profile at the test sample point is calculated by CFD, if it meets the following requirement:

\[
|f(x_0 + \frac{x_i - x'_0}{3}, x_j + \frac{x_j - x'_0}{3}, x'_0) - f_0 - f(x_i) - f(x_j)| \leq \varepsilon
\]  

(12)

then the second-order model function \( f_0(x_i, x_j) \) satisfies the precision requirement, and the construction of the second-order model function is finished. Otherwise, the training points are added and the construction of \( f_0(x_i, x_j) \) is rebuilt.

In practical application, the low-order model function has met the accuracy requirements, so the SVM-HDMR surrogate model of PAT is no longer constructed with more than two-order model functions. Substitute the constructed model function into equation (7). The HDMR model based on SVM is completed.

**Optimization strategy flow of PAT**

The optimization strategy flow of PAT based on SVM-HDMR surrogate model is shown in Figure 1.

First, according to the blade parameters of the prototype PAT, the corresponding blade profile is fitted by MATLAB software, and the training set of SVM is generated based on the variation range of the control points of the blade profile. Second, using Creo, ANSYS, and MATLAB software, according to the construction process of SVM-HDMR surrogate model of PAT, the fitness function of PAT efficiency is constructed. Finally, PSO algorithm is used to solve the fitness function, and numerical simulation and experimental methods are used to analyze the data.

**Optimization process of PAT**

**Structural parameters of PAT**

The performance parameters of selected PAT are: the flow rate is 27.5 m\(^3\)/h, the head is 49.2 m, and the rotational speed is 2900 rpm. Main structural parameters of impeller and volute (according to the description of pump structural parameters) are shown in Table 1.

| Name        | Parameter                                      | Value |
|-------------|------------------------------------------------|-------|
| Impeller    | Impeller inlet diameter \(D_1/(\text{mm})\)    | 48    |
|             | Impeller outlet diameter \(D_2/(\text{mm})\)    | 165   |
|             | Blade outlet width \(b_3/(\text{mm})\)         | 6     |
|             | Impeller inlet angle \(\beta_1/(^\circ)\)      | 32.5  |
|             | Impeller outlet angle \(\beta_2/(^\circ)\)      | 14    |
| Volute      | Volute base diameter \(D_3/(\text{mm})\)        | 170   |
|             | Volute outlet width \(b_4/(\text{mm})\)         | 16    |
|             | Volute outlet diameter \(D_4/(\text{mm})\)      | 32    |

**Flow field calculation domain.** The optimized object is the blade profile. According to the generated training set, many models need to be calculated. Considering the economy, in order to save computing resources, the flow field calculation domain of the PAT includes four parts: the import extension section, the volute, the impeller, and the export extension section. After the grid independence check, the final grid number is 1,521,070. The flow field calculation domain and grid division of PAT are shown in the Figure 2.
Turbulence model selection and boundary condition setting. At present, there are many models for turbulence numerical simulation, but none of them are universal. Considering the development of turbulence models, theoretical perspectives, computational economics, and practical considerations, the SST $k - \omega$ model considers the robustness of the $k - \omega$ model in the near wall region to the separation zone induced by the inverse pressure gradient, and the stability of the $k - \varepsilon$ model outside the boundary layer. Therefore, the SST $k - \omega$ model is selected for numerical simulation.

The numerical simulation boundary of the PAT includes: inlet boundary, wall boundary, and outlet boundary. Because the working condition of the PAT is known, the boundary conditions of the velocity inlet, free outflow, and no-slip wall boundary conditions are adopted.

The construction of PAT surrogate model

Parameterization of PAT blade profile. According to the construction process of the SVM-HDMR surrogate model of the PAT, firstly, the data of the prototype PAT blade was extracted by Creo software. Secondly, according to the original blade profile data and the programming function of MATLAB software, the program of Bezier spline curve was compiled. The blade of the PAT was parameterized by Bezier spline curve to generate the corresponding blade profile of the prototype PAT blade and the control point line with seven blade profile control points. According to the programming process of Bezier spline, the shape of blade profile will change with the position of control point. It is defined that the abscissa of blade profile is $z$ and the ordinate is $\chi$, as shown in Figure 3.

Figure 2. Computational domain and meshing of PAT flow field: (a) PAT computational domain of flow field and (b) computational domain meshing of PAT.

Figure 3. Parameterization of blade profile: (a) original blade profile and (b) parameterized blade profile. BP: blade profile; BPCPL: blade profile control point line.
The control point at the intersection of two adjacent control points’ lines was defined as an independent variable \( x_i \) of the adaptive function of the surrogate model (therefore, five independent variables could be determined from seven control points of the blade profile of the PAT in the calculation example), and the change direction of the included angle along the corresponding bisector was defined as the value change region of the independent variable. The change region of each variable is shown in Table 2.

### Table 2. Variable optimum interval of blade profile.

| Design variables | Meaning | Change interval |
|------------------|---------|-----------------|
| \( x_1 \) | Angle between two control points’ lines | \([-5^\circ, 5^\circ]\) |
| \( x_2 \) | \([-10^\circ, 10^\circ]\) |
| \( x_3 \) | \([-10^\circ, 10^\circ]\) |
| \( x_4 \) | \([-10^\circ, 10^\circ]\) |
| \( x_5 \) | \([-10^\circ, 10^\circ]\) |

The training space of surrogate model was generated according to the range of control points of blade profile. The first time to build the PAT SVM-HDMR surrogate model is taken as an example, the training points needed to build the first-order model and the corresponding PAT efficiency (\( \eta \)) calculated by CFD are shown in Table 3.

### Table 3. Training point and PAT efficiency of first order SVM-HDMR model.

| No. | \( x_1 \) | \( x_2 \) | \( x_3 \) | \( x_4 \) | \( x_5 \) | \( \eta/\% \) |
|-----|----------|----------|----------|----------|----------|----------|
| 0   | 0        | 0        | 0        | 0        | 0        | 56.693   |
| 1   | 5^\circ  | 0        | 0        | 0        | 0        | 56.945   |
| 2   | -5^\circ | 0        | 0        | 0        | 0        | 55.724   |
| 3   | 0        | 10^\circ | 0        | 0        | 0        | 57.476   |
| 4   | 0        | -10^\circ| 0        | 0        | 0        | 56.358   |
| 5   | 0        | 0        | 10^\circ | 0        | 0        | 56.322   |
| 6   | 0        | 0        | -10^\circ| 0        | 0        | 56.865   |
| 7   | 0        | 0        | 0        | 10^\circ | 0        | 56.426   |
| 8   | 0        | 0        | 0        | -10^\circ| 0        | 56.381   |
| 9   | 0        | 0        | 0        | 0        | 10^\circ | 56.117   |
| 10  | 0        | 0        | 0        | 0        | -10^\circ| 56.869   |

\[
\hat{f}(x_4) = -9.5 \times 10^{-6}(10x_5 + 1)^2 + 9.298 \times 10^{-6} (1 - 10x_5)^2 + 2.001 \times 10^{-7} \quad (17)
\]

From equations (13)–(17), it can be concluded that the first-order SVM-HDMR functions constructed are all nonlinear functions.

The training points needed for the construction of the second order SVM-HDMR functions and the corresponding PAT efficiency calculated by CFD are shown in Table 4.

As the analytical expression of the second-order SVM-HDMR functions are not intuitionistic than that of the first-order SVM-HDMR functions, here we choose the way of image to express the second-order SVM-HDMR functions, as shown in Figure 4. Combining with the calculation data in Table 4, it can be seen that we have only constructed seven second-order SVM-HDMR functions such as \( f_{12}(x_1, x_2) \), \( f_{13}(x_1, x_3) \), \( f_{23}(x_2, x_3) \), \( f_{24}(x_2, x_4) \), \( f_{34}(x_3, x_4) \), \( f_{35}(x_3, x_5) \), \( f_{45}(x_4, x_5) \), and have not constructed second-order functions \( f_{14}(x_1, x_4) \), \( f_{15}(x_1, x_5) \), \( f_{25}(x_2, x_5) \). This is because: on the one hand, it is sufficient to determine that these second-order functions do not exist; on the other hand, from the geometric point of view (take \( f_{14}(x_1, x_4) \) as an example), when variable \( x_4 \) changes, because the distance between variables \( x_1 \) and \( x_4 \) is far enough, there is no or minimal impact on variable \( x_1 \), and there is no relevant second-order function term between them. The first order and second order SVM-HDMR functions were used to form the fitness function of the SVM-HDMR surrogate model of the PAT. The PSO algorithm was used to solve the fitness function and the optimized blade profile was obtained by feedback.

### Results and discussion

#### Optimization results of PAT surrogate model

Based on the optimized PAT blade profile, the SVM-HDMR surrogate model of PAT was built and solved.
Table 4. Training point and PAT efficiency of second order SVM-HDMR model.

| No. | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $\eta/\%$ | No. | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $\eta/\%$ |
|-----|-------|-------|-------|-------|-------|---------|-----|-------|-------|-------|-------|-------|---------|
| 1   | $5^\circ$ | $10^\circ$ | 0     | 0     | 0     | 57.919  | 15  | 0     | $-10^\circ$ | 0     | $10^\circ$ | 0     | 57.079  |
| 2   | $5^\circ$ | $-10^\circ$ | 0     | 0     | 0     | 55.543  | 16  | 0     | $-10^\circ$ | 0     | $-10^\circ$ | 0     | 56.757  |
| 3   | $-5^\circ$ | $10^\circ$ | 0     | 0     | 0     | 55.251  | 17  | 0     | 0     | 0     | $10^\circ$ | 0     | 56.844  |
| 4   | $-5^\circ$ | $-10^\circ$ | 0     | 0     | 0     | 56.486  | 18  | 0     | 0     | 0     | $10^\circ$ | 0     | 56.762  |
| 5   | $5^\circ$ | 0     | $10^\circ$ | 0     | 0     | 57.213  | 19  | 0     | 0     | 0     | 0     | 0     | 56.857  |
| 6   | $5^\circ$ | 0     | $-10^\circ$ | 0     | 0     | 57.005  | 20  | 0     | 0     | 0     | $-10^\circ$ | 0     | 56.937  |
| 7   | $-5^\circ$ | 0     | $10^\circ$ | 0     | 0     | 55.617  | 21  | 0     | 0     | 0     | $10^\circ$ | 0     | 55.609  |
| 8   | $-5^\circ$ | 0     | $-10^\circ$ | 0     | 0     | 55.111  | 22  | 0     | 0     | 0     | $10^\circ$ | 0     | 55.222  |
| 9   | 0     | $10^\circ$ | 0     | $10^\circ$ | 0     | 56.961  | 23  | 0     | 0     | 0     | $-10^\circ$ | 0     | 55.914  |
| 10  | 0     | $10^\circ$ | $-10^\circ$ | 0     | 0     | 56.242  | 24  | 0     | 0     | 0     | 0     | $10^\circ$ | 0     | 55.221  |
| 11  | 0     | $-10^\circ$ | $10^\circ$ | 0     | 0     | 55.185  | 25  | 0     | 0     | 0     | 0     | $10^\circ$ | 0     | 55.252  |
| 12  | 0     | $-10^\circ$ | $-10^\circ$ | 0     | 0     | 55.966  | 26  | 0     | 0     | 0     | 0     | $10^\circ$ | 0     | 55.471  |
| 13  | 0     | $10^\circ$ | 0     | $10^\circ$ | 0     | 57.162  | 27  | 0     | 0     | 0     | 0     | $-10^\circ$ | 0     | 57.411  |
| 14  | 0     | $10^\circ$ | 0     | $-10^\circ$ | 0     | 57.344  | 28  | 0     | 0     | 0     | 0     | $-10^\circ$ | 0     | 57.411  |

Figure 4. Second order SVM-HDMR function images: (a) $\hat{f}_{12}(x_1, x_2)$, (b) $\hat{f}_{13}(x_1, x_3)$, (c) $\hat{f}_{23}(x_2, x_3)$, (d) $\hat{f}_{24}(x_2, x_4)$, (e) $\hat{f}_{34}(x_3, x_4)$, (f) $\hat{f}_{35}(x_3, x_5)$, and (g) $\hat{f}_{45}(x_4, x_5)$. 
with the PSO algorithm again. In this way, the flow field characteristics of 120 PAT models were studied, and the SVM-HDMR surrogate model of PAT was built and solved with the PSO algorithm three times. The three optimized PAT blade profile is compared with the original blade profile, as shown in Figure 5. From Figure 5(a), after the first optimization, the main change of blade profile is the curvature at the impeller outlet. From Figure 5(b), after the second optimization, the overall curvature of the blade profile becomes smaller based on the first optimization. From Figure 5(c), the third optimization is to change the curvature of the blade at the impeller inlet. Note that the first optimization and the second optimization are staggered and repeated at the impeller inlet, which indicates that there are two solutions for the first constructed model function.

In summary, after three iterations, the blade profile curvature of PAT becomes smaller than that of the original blade profile, and the curvature change at the impeller outlet is larger than that at the impeller inlet.

**Flow field analysis**

The optimized impeller in each stage and the original impeller in Figure 5 were simulated at the design operating point. The simulation results were processed by Tecplot software to get the spatial velocity streamline diagrams of PAT, as shown in Figure 6. The (1)–(4) in Figure 6 represent the three-dimensional streamline diagrams of the PAT corresponding to the impeller formed by the prototype blade profile and the blade profile optimized for the first time, the second time, and the third time. According to Figure 6(a) (1) to (c) (1), the low efficiency of the prototype PAT can be attributed to the existence of the impeller inlet flow separation zone (area A in Figure 6(b) (1)) and the impeller outlet disordered zone (area B in Figure 6(c) (1)).

![Figure 5. Comparison of blade profile before and after optimization: (a) comparison of the first optimized blade profile and the original blade profile, (b) comparison of the second optimized blade profile and the first optimized blade profile, (c) comparison of the third optimized blade profile and the second optimized blade profile, and (d) comparison of the third optimized blade profile and original blade profile. OBP: optimized blade profile; OBPCPL: optimized blade profile control point line.](image-url)
The separation zone is mainly formed by the positive impact of the fluid flowing from the volute into the impeller and the direction of the blade flow in the impeller is not consistent. Because the fluid in the separation zone can’t fit the blade well, the blade in some areas can’t do work. The existence of disordered zone increases the energy of turbulent dissipation. From Figure 6(a) (1)–(2) to (c) (1)–(2), after the blade profile is optimized for the first time, the flow separation zone at the impeller inlet decreases. From Figure 6(a) (2)–(4) to (c) (2)–(4), after the second and third optimization of blade profile, the disordered area at the impeller outlet is gradually eliminated. It is the reduction of the flow separation zone at the inlet of the optimized impeller and the elimination of the disordered zone at the outlet of the optimized impeller that improves the flow situation inside the impeller and makes the efficiency of the optimized PAT increase significantly.

In order to analyze the internal energy transport and transformation process of the PAT before and after optimization, the turbulent viscosity field, turbulent kinetic energy field, and turbulent dissipation rate field of the PAT before and after optimization were obtained by further processing the numerical simulation results with Tecplot software, and the contour images of $z = 2.5\, \text{mm}$ plane (impeller outlet width is 6 mm) were selected as the analysis object, as shown in Figure 7. The (1)–(4) in Figure 7 represent the turbulent viscosity field, turbulent kinetic energy field, and turbulent dissipation rate field of the PAT corresponding to the impeller formed by the prototype blade profile and the first, second, and third optimized blade profile respectively. From Figure 7(a) (1)–(4), under the same conditions, the large turbulent viscosity area is mainly distributed in the impeller channel near the volute tongue. With the iteration of optimization times, the turbulent viscosity

![Figure 6. Velocity streamline diagrams of PAT: (a) spatial velocity streamline stereogram of PAT, (b) front view of spatial velocity streamline of PAT, and (c) back view of spatial velocity streamline of PAT. VS: velocity streamline.](image-url)
intensity in the impeller decreases. According to Boussinesq eddy viscosity hypothesis, the distribution of turbulent viscosity intensity corresponds to the distribution of energy generation term caused by Reynolds stress, which also determines the distribution characteristics of turbulent kinetic energy. From Figure 7(b) (1)–(4), under the same conditions, the high-energy area of turbulent kinetic energy distribution corresponds to the distribution of turbulent viscosity field. With the iteration of optimization times, the high-energy region of turbulent kinetic energy of impeller shrinks from the full flow channel to the inlet of impeller, indicating the decrease of turbulent kinetic energy transport intensity. According to the turbulent kinetic energy transport equation, the turbulent dissipation rate caused by turbulent kinetic energy transport will be reduced. From Figure 7(c) (1)–(4), under the same conditions, with the iteration of optimization times, the turbulent dissipation rate distribution of the impeller shrinks and the intensity decreases, indicating that the turbulent energy dissipation decreases. According to the cascade effect of turbulent eddies and Kolmogorov hypothesis, the energy generation term caused by Reynolds stress in Figure 7(a) and (b) decreases, which indicates that the energy transferred between different scale eddies in the impeller region decreases, and the dissipated energy also decreases when it reaches the dissipated scale eddies, which also illustrates the phenomenon in Figure 7(c).

In summary, Figures 6 and 7 show the change of fluid flow state in the impeller and the process of energy transport and transformation in the impeller respectively. It shows that the optimized impeller is more...
conducive to the conversion of liquid energy, and the efficiency of the optimized impeller will be higher than that of the prototype PAT impeller.

**Verification of optimization results**

**Numerical simulation verification of optimization results.** The external characteristic curve obtained by numerical simulation before and after optimization of PAT is shown in Figure 8.

From the Figure 8(a), at the optimal operating point, the efficiency of the numerical simulation of the PAT optimized by the surrogate model is higher than that of the prototype PAT 5.49%. From the Figure 8(b), the output power of the numerical simulation of the optimized PAT is higher than that of the prototype PAT 7.2%.

**Experimental verification of optimization results.** The structural diagram of the accuracy verification test system for the optimization results and numerical simulation of the PAT blades are shown in Figure 9.

From the Figure 9, the specific operation mode of the test rig is: The motor-driven feed pump provides high-pressure liquid for the PAT, and the consume pump forms a closed-loop energy conversion of the whole test system by consuming the mechanical energy converted by the PAT. An electromagnetic flow meter and pressure transmitter are installed between the feed pump and the PAT connecting pipe to measure the flow and pressure at the inlet of the PAT. A pressure transmitter is installed at the outlet pipeline of the PAT to measure the pressure at the outlet of the PAT. A torque meter is installed on the connecting shaft of the PAT and the consume pump to measure the output torque and speed of the PAT.

The test site corresponding to Figure 9 is shown in Figure 10. From the Figure 10, the specific parameters of the measuring instruments used in the test are as follows:

**Figure 8.** External characteristic curve of PAT numerical simulation: (a) flow-efficiency curve of PAT and (b) flow-power curve and flow-head curve of PAT.

BO: before optimization; AO: after optimization; PV: predictive value.

**Figure 9.** The structural diagram of the accuracy verification test system.
follows: The rated working pressure of the electromagnetic flow meter is 2 MPa, the measuring range is 0.3533 – 84.78 m³/h, and the measuring accuracy is ± 0.1%. The range of the pressure transmitter is 0 – 2 MPa, and the measuring accuracy is ± 0.1%. The rated torque of the torque meter is 100 N·m, the measurable speed range is 0 – 6000 r/min, the calibration coefficient is 8522, and the accuracy level is 0.2.

Therefore, the experimental data for drawing the external characteristic curve of the PAT were measured. The performance curve of the PAT obtained from the test is compared with that of the prototype PAT, as shown in Figure 11.

Comparing with Figure 8, it can be concluded that the numerical simulation before and after optimization is basically consistent with the test performance curve. At the same time, it can be seen from the Figures 8(a) and 11(a) that the optimized flow efficiency curve of the PAT becomes flatter in the high efficiency region. From the Figure 11(a), at the optimal operating point, the efficiency of test of the optimized PAT is higher than that of the original PAT 5.1%. From the Figure 11(b), the output power of test of the optimized PAT is higher than that of the prototype PAT 6.9%.

Conclusions

The SVM-HDMR surrogate model combined with PSO algorithm was used to optimize the blade of PAT. The blade profile was parameterized by Bezier spline curve, and the correlation between various variables was analyzed. The model function of PAT efficiency on
blade profile parameters was constructed. Through the study of the optimization process and the verification and analysis of the optimization results, the following conclusions can be drawn:

(1) After three optimization of PAT blade, the efficiency of PAT at the optimal operating point is predicted to be 61.5% by PSO algorithm. The optimization results of different stages are analyzed from the point of view of the flow field in the impeller. The reasons for improving impeller efficiency after each optimization are explained.

(2) The predicted value of the surrogate model is verified by numerical simulation. The reason for the hydraulic loss of the impeller after optimization is analyzed from the flow field inside the impeller before and after optimization. The results show that the efficiency of the numerical simulation is increased by 5.49% and the output power is increased by 7.2%.

(3) The accuracy of the efficiency point predicted by the surrogate model and the numerical simulation results are verified through the experimental study. The results show that the experimental efficiency value is increased by 5.1% and the output power is increased by 6.9% compared with the prototype PAT.

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 research was funded by National Natural Science Foundation of China, grant number 51569013; Industrial Support Program of Colleges in Gansu Province, grant number 2020C-20; Open Fund of Key Laboratory of Fluid and Power Machinery, Ministry of Education, Xihua University, grant number szjj2019-016.

ORCID iDs
Bingxiao Jiang https://orcid.org/0000-0002-5938-0437
Senchun Miao https://orcid.org/0000-0002-8831-9723

References
1. Wang XH, Yang JH and Shi FX. Status and prospect of study on energy recovery hydraulic turbines. J Drain Irrig Mach Eng 2014; 32: 742–747.
2. Armando C, Shahram DH and Helena MR. Pumps as turbines fundamentals and applications. Berlin: 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. Gao XF, Shi WD, Zhang DS, et al. Optimization design and test of vortex pump based on CFD orthogonal test. Trans Chin Soc Agric Mach 2014; 45: 101–106.
5. Zhu YJ, Ju YP and Zhang CH. Adjoint optimization for direct and inverse design of cascade blade profiles. J Xi’an Jiaotong Univ 2019; 53: 100–106.
6. 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.
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 2019; 50: 167–172.
8. Chu F, Wang FL, Wang XG, et al. Hybrid model for multi-stage centrifugal compressor based on radial basis function neural network. Control Theory Appl 2012; 29: 1905–1910.
9. Tang L, Li GY and Wang H. Kriging-HDMR metamodeling technique for nonlinear problem. Chin J Theor Appl Mech 2011; 43: 781–784.
10. Li WP, Dou XD, Wang ZX, et al. BPNN-HDMR Nonlinear metamodeling technique and its application. J Human Univ 2014; 41: 32–38.
11. 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.
12. Sobol IM. Sensitivity estimates for nonlinear mathematical models. Math Model Comput Exp 1993; 1: 407–414.
13. Rabitz H and Alis ÖF. General foundations of high dimensional model representations. J Math Chem 1999; 25: 197–233.
14. Alis ÖF and Rabitz H. Efficient implementation of high dimensional model representations. J Math Chem 2001; 29: 127–142.
15. Tunga MA. An approximation method to model multivariate interpolation problems: indexing HDMR. Math Comput Model 2011; 53: 1970–1982.
16. Vapnik VN. Statistical learning theory. New York, NY: John Wiley & Sons, Inc, 1998.
17. Yang XW and Hao ZF. Algorithm design and analysis of support vector machine. Beijing: Science Press, 2013.
18. Menter FR. Improved two-equation $k$ – $\omega$ turbulence models for aerodynamic flows. NASA Tech Memo 1992; 103975: 1–31.
19. Menter FR, Kuntz M and Langtry R. Ten years of industrial experience with the SST turbulence model. Turbul Heat Transf 2003; 4: 625–632.
20. Zhang ZX and Dong ZN. Viscous fluid mechanics. 2nd ed. Beijing: Tsinghua University Press, 2011.
21. David CW. Turbulence modeling for CFD. 3rd ed. San Diego, CA: DCW Industries, 2006.