Machine learning based design optimization of centrifugal impellers

*Correspondence:
YL: yanliu@dlut.edu.cn

Ao Zhang[1], Yan Liu[1,*], Jinguang Yang[1], Zhi Li[1], Chuang Zhang[1], Yiwen Li[2]

[1]Dalian University of Technology, No. 2, Linggong Street, Dalian City, China
[2]Shenyang Blower Works Group, No. 2, Linggong Street, Dalian City, China

Abstract

Big data and machine learning are developing rapidly, and their applications in the aerodynamic design of centrifugal impellers and other turbomachinery have attracted wide attention. In this paper, centrifugal impellers with large flow coefficient (0.18−0.22) are taken as research objects. Firstly, through one-dimensional design and optimization, main one-dimensional geometric parameters of those centrifugal impellers are obtained. Subsequently, hundreds of samples of centrifugal impellers are obtained by using an in-house parameterization program and Latin hypercube sampling method. The NUMECA software is used for CFD calculations to build a sample library of centrifugal impellers. Then, applying the artificial neural network (ANN) to deal with the data in the sample library, a nonlinear model between the flow coefficients, the geometric parameters of these centrifugal impellers and the aerodynamic performance is constructed, which can replace CFD calculations. Lastly with the help of the multi-objective genetic algorithm, a global optimization is carried out to fulfill a rapid design optimization for centrifugal impellers with flow coefficients in the range of 0.18−0.22. Three examples provided in the paper show that the design and optimization method described above is faster and more reliable compared with the traditional design method. This method provides a new way for the rapid design of centrifugal impellers.

Introduction

Centrifugal compressors are widely used in industry. With the advance of industries, centrifugal compressors with great capacity and high performance are required. However this type of centrifugal compressors with large flow coefficient is difficult to design and optimize (Zhu et al., 2016) due to complicated internal flow in the impellers. The traditional aerodynamic design and optimization method of impeller generally follows one-dimensional design, two-dimensional meridian channel design, quasi-three-dimensional blade design and full three-dimensional CFD verification. It requires not only the repeated numerical simulations for flow field analysis, but also iterations of the impeller geometry for continuous modifications. Although the traditional aerodynamic design method can achieve reasonable results, it needs manual debugging repeatedly, and greatly relies on the design experience of the designer, which is time-consuming. Thus, the combination between traditional aerodynamic design method and machine learning that have emerged in recent years has currently become a new direction of research.

Machine learning is a general term for a class of algorithms. We can use algorithms to mine the hidden laws in the data that cannot be processed by the human brain. As a branch of artificial intelligence, the machine learning is committed to study how to use computing to improve the performance of the system (Zhou, 2016). Its origin can be traced back to the basic tools of machine learning such as the least squares method and Markov chain in the 17th century. Its development has evolved through the connectionism in the 1980s, statistical learning in the 1990s, and deep learning in the 21st century. To date, the machine learning has become a new subject that integrates the theoretical foundations of multiple disciplines such as mathematics, computers, biology, and neurology. The interdisciplinary research related to machine learning is emerging, and some research results have already been transformed into practical applications.

With the continuous improvement of computing power and data volume, the rapid development of surrogate models such as artificial neural network models (Hassoun, 1996), kriging models, convolutional neural networks and recurrent neural networks has provided us with better tools to describe the relationship of "number-number". In the design of centrifugal impeller, the relationship between the design parameters and aerodynamic parameters is mostly non-linear, which is difficult to be described by specific formulas. A new design method for centrifugal impellers is to use the surrogate model to construct a non-linear relationship and dig out the internal relationship.

In the process of centrifugal impeller design optimization, most optimization problems are not single-objective optimization problems, but multi-objective and multi-constrained global optimization problems. Therefore, gradient-based optimization methods are inherently unsuitable. The heuristic optimization algorithms such as evolutionary algorithm, particle swarm algorithm, annealing algorithm, ant colony algorithm, etc., have become important tools for solving such problems. These methods have developed rapidly in the field of machine learning in recent years. They are more suitable for solving multi-objective optimization problems, and better approaching the non-convex or discontinuous optimal front ends (Tang et al., 2004). A multi-objective genetic algorithm is a global algorithm that specifically solves multi-objective global optimization problems. It searches for the optimal solution by simulating the natural evolution process. In order to solve the problem of which the optimized sub-objectives may be contradictory, the concept of Pareto solution set is proposed. The optimization problems in this paper are all multi-objective global optimization, so the multi-objective genetic algorithm is selected (Deb et al., 2002).

Ibaeaki et al. (2015) and Verstraete et al. (2010) optimized a centrifugal impeller and diffuser by combining the surrogate model and advanced algorithms, and achieved good results. Guo et al. (2016) predicted the flow field of a centrifugal impeller based on the convolutional neural network. At the expense of partial accuracy, the computing speed was increased by two orders of magnitude. Duccio et al. (2002) optimized an impeller by combining the response surface and genetic algorithm, which greatly improved the aerodynamic performance of the impeller, but took too long time in the process. Kim et al. (2009) optimized the meridian flow channel of a centrifugal impeller through the radial basis function to improve its efficiency and total pressure ratio. These studies show the advantages and potential of machine learning methods for the design optimization of impellers machinery.

In this paper, the traditional design method of centrifugal impeller is combined with machine learning to design and analyze impellers with flow coefficients of 0.18, 0.2 and 0.22, which provides a new way for design of centrifugal impellers.

Methodology

Research object

According to the actual needs of the project, the object of this study is a series of closed centrifugal impellers with flow coefficients of 0.18–0.22, energy head coefficient of about 0.6, and impeller rotational Mach number of about 0.95. The fluid type is perfect gas. Impellers with different flow coefficients have the same design speed and impeller exit diameter, aiming to obtain impellers with the best aerodynamic performance under different flow coefficients. Table 1 shows the design parameters and requirements of the centrifugal impeller investigated in the paper.

Design and optimization process

This section introduces the design and optimization process. Figure 1 shows the design and optimization flow chart, and gives the detailed introduction to the important parts.

Table 1. Design parameters and requirements of centrifugal impeller.

Parameters	Value
Inlet total pressure, P_0	98100 Pa
Inlet total temperature, T_0	288.15 K
Impeller exit diameter, D_2	0.45 m
Speed, n	13140 r/min
Number of blades, Z	15Full
Total pressure ratio, ε	1.85
Isentropic efficiency, η	>90%
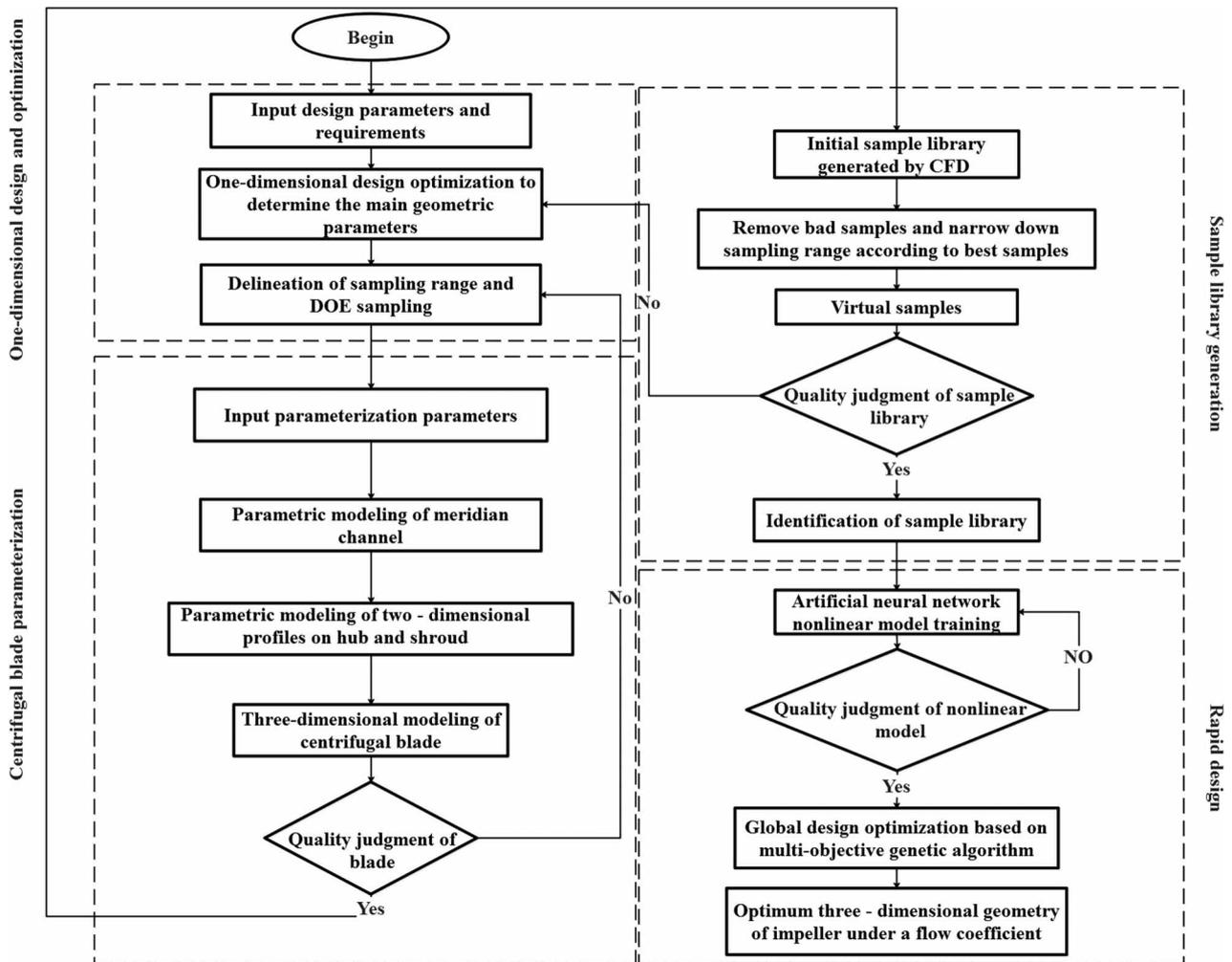
Figure 1. Design and optimization flow chart.

1. First, input the design parameters and design requirements of the impeller, mainly including P_0, T_0, D_2, ε, n, q_m, Z, the fluid type, etc;
2. Based on the one-dimensional analysis program of the two-zone model and combined with the optimization algorithm, the specific values of the main one-dimensional geometric parameters (G_i) of the centrifugal

impeller are obtained, including D_{1s}, D_2/D_{1h}, β_{1s}, β_{1h}, β_2, b_2, etc. The upper and lower bounds of each variable are determined according to the specific value in the initial sampling space;

3. Choose an appropriate design of experiment (DOE) to generate the initial sample parameters. In this study, the Latin superposition method was used for the initial sampling;

4. The meridian channel of the impeller blade, the β angle and the thickness of the blade on the shroud and hub are parameterized. Generate three-dimensional impellers according to the initial sampling parameters.

5. Qualified samples are calculated by CFD, the isentropic efficiency and total pressure ratio corresponding to each sample are obtained, and the sample library is initially generated. The samples are screened according to the sample with the best aerodynamic performance in the sample library. Re-determine the sampling range and repeat step 2 to screen and supplement the sample library;

6. The artificial neural network model is trained using the back propagation algorithm to construct a nonlinear model among the flow coefficient, geometric parameter control points, isentropic efficiency, and total pressure ratio.

7. The nonlinear model is used to replace the CFD calculation, and the multi-objective genetic algorithm is used for global optimization to achieve the rapid design of the centrifugal impeller;

8. Feedbacks are set up in the whole process to reduce the time of sample library construction and search, which is stated as follows:
   • Centrifugal blade parameterization part: check the requirements of the exit lean angle. The criterion is: exit lean angle <±20°;
   • Aerodynamic analysis part: check the isentropic efficiency, total pressure ratio and flow coefficient;
   • Non-linear model training part: check the value of the model correlation coefficient R. The closer R is to 1, the closer the sample value is to the 45 degree regression line. It shows that the higher the fitting accuracy, the better the regression effect. The criterion is R > 0.90.

One-dimensional analysis optimization

One-dimensional analysis simplifies the actual flow by considering the change of parameters in the direction of the main flow only. Although, due to the simplification, the accuracy of one-dimensional analysis is limited, it still occupies an important position in the preliminary design of the centrifugal impeller on account of its fast convergence rate of calculation and short time-consumption, and also the initial plan it can provide for a design.

The two-zone model divides the gas flow in the impeller into the main flow zone and secondary flow zone. The flow in the main flow zone is regarded as isentropic flow, and the flow in the secondary flow zone is regarded as non-isentropic flow (Japikse, 1985). Based on the two-zone model, researchers have proposed a large number of loss models to analyze and predict the aerodynamic performance of centrifugal impellers, and have achieved relatively good results.

This research has developed a one-dimensional analysis program, mainly based on the two-zone model. The program can output the characteristic curves of mass flow-isentropic efficiency and mass flow-total pressure ratio by inputting the design parameters and main geometric parameters of a centrifugal impeller. One-dimensional analysis program has the advantage that the time cost of a single run is only on the order of seconds. Combined with the multi-objective genetic algorithm, the main geometric parameters are optimized globally. The optimization objective is the maximum isentropic efficiency and the minimum difference between the total pressure ratio and the target pressure ratio under the design condition. The mathematical model is shown in Equation 1. Through the one-dimensional design optimization of impellers with flow coefficients of 0.18, 0.2 and 0.22 respectively, the upper and lower bounds of the best main one-dimensional parameters impellers with the specified flow coefficients are determined, and a three-dimensional sampling space is determined.

$$\begin{cases} \max : \eta_{des} \\ \min : |\varepsilon_{des} - \varepsilon_{obj}| \end{cases} \tag{1}$$
$$st. \min G_i \leq G_i \leq \max G_i$$

Parametric method of centrifugal impeller

A reasonable parameterization method of a centrifugal impeller should not be too complicated. While ensuring accuracy, it is better to use fewer parameter control points to reasonably describe the shape of blade. In order to

achieve better coupling with the one-dimensional analysis program, an in-house parameterization program for centrifugal impellers has been developed in this study.

In the meridian coordinate system (R, Z), the meridian flow channel is given, and in the dimensionless meridian direction (m', β), the β angle distribution is given. The β angle is converted into camber line through integration, and the thickness is superimposed on its normal direction to generate a two-dimensional blade shape. Finally, three-dimensional blades are generated by stacking the leading edges (Wang et al., 2019).

The flow condition of the main flow is determined by the meridian channel. As shown in Figure 2a, the meridian flow channel is composed of two meridian curves, both of which are constructed using 4th-order Bezier curves. The inlet and outlet geometry of the meridian curves is determined by the control points S_1, H_1, S_5 and H_5, and the remaining control points are used to adjust the shape of the meridian curves. Figure 2b shows the distribution of β angle at the hub, the β angle distribution of hub is constructed using 5th-order Bezier curves. The inlet and outlet β angle are respectively determined by the B_1 and B_6 control points, and the remaining control points are used to adjust the β angle distribution. Figure 2c shows the distribution of thickness at the hub, the thickness distribution of hub is constructed using 4th-order Bezier curves. T_2 and T_4 are used to control the shape of the leading and trailing edge, and their abscissas are the same as the control points of T_1 and T_5, respectively at the beginning and the end of the dimensionless meridian direction. T_3 is used to adjust the position and size of the maximum thickness. The parameterization method of β angle and thickness distribution of the shroud is the same as the hub. S_e (e = 1,…,5), H_f (f = 1,…,5), B_c (c = 1,…,12) and T_h (h = 1,…,10) respectively represent the control points and the corresponding coordinate values.

Sample library and training of nonlinear models

Design of experiment (DOE) plays an important role in the generation of the sample library, and a reasonable DOE method can make the sample distribution in the sample library more reasonable. Latin Hypercube is a stratified sampling DOE method. Compared with random sampling and full factorial design, it has a more effective space filling capacity on the basis of reducing sample requirements. For geometric parameters b_2, D_{1s} and the blade angle control points on the shroud and hub, Latin hypercube sampling was performed. The remaining parameters were selected based on the result of one-dimensional design optimization. Part of the control points of the β angle distribution and the meridian flow channel cannot be obtained through one-dimensional design and optimization. Therefore, in the initial sampling, a larger range is given, and the second sampling reduces the sampling range of the control points according to the optimal sample. After the sampling, three-dimensional samples are generated using the centrifugal impeller parameterization program.

CFD calculation was performed on the three-dimensional samples of the centrifugal impeller. The grid was generated using the AutoBlade module of the NUMECA software. In order to trade off between calculation accuracy and computing cost, the grid number was set to 550,000. Figure 3 is shows the grid on the blade surface and the hub. The Fine module of the NUMECA software was used for numerical simulation, and the Spalart-Allmaras (SA) turbulence model was used. The chain of grid generation and flow calculation was automated using an in-house python script, thus to reduce manpower and time costs.

According to the optimal sample in the initial calculation results, the sample space was reduced. Firstly, the samples were sorted based on the aerodynamic performance, and the non-convergent samples and the samples with poor aerodynamic performance were removed. Secondly, new ranges for the sample space were determined based on the optimal samples. Finally, new samples were filled. For the initial sampling, the number of samples in the sample library was 1,000. After removing samples and filling samples again, the total number of samples in the sample library became 463. Table 2 shows the second sampling range.
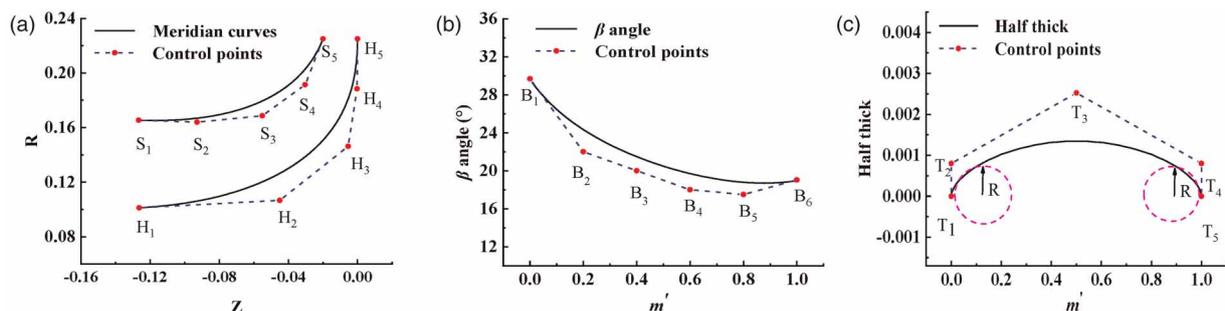
Figure 2. Parameterization model of meridional channel (a), β at the hub (b), and thickness at the hub (c).
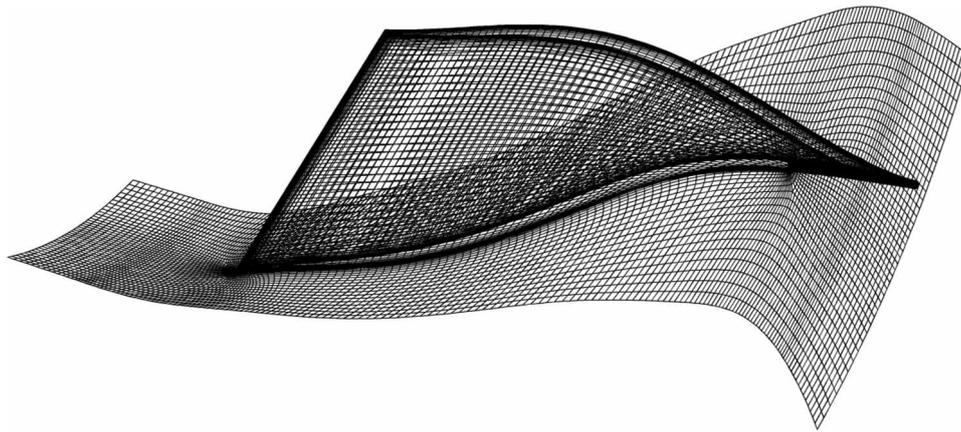
Figure 3. Surface grid of the impeller.

Table 2. The second sampling range.

Design Variable	Lower Bound	Upper Bound	Design Variable	Lower Bound	Upper Bound	Design Variable	Lower Bound	Upper Bound
S_1	0.039	0.043	B_4	5	8	B_9	50	58
S_5	0.016	0.019	B_5	12	21	B_{10}	54	91
B_1	45	57	B_6	28	39	B_{11}	27	45
B_2	48	80	B_7	42	51	B_{12}	20	27
B_3	10	17	B_8	25	42	φ	0.18	0.22

The data was normalized, and the min-max standardization method was used to linearly transform the original data, so that the result could be mapped to 0–1, as shown in Equation 2.

$$X_{\text{new}} = \frac{(X - X_{\min})}{(X_{\max} - X_{\min})}$$

(2)

The ratio of the sample training set and the prediction set in the sample library is 0.15:0.85. After repeating tests, this research selected a 4-layer artificial neural network. BP algorithm was used to train artificial neural network. Figures 4 and 5 respectively show the model training curve and the test set 45-degree regression line (Lei, 2020). It can be seen that when the training epoch is about 70, the convergence is stable, and the mean square error of the training curve and the test curve are both less than 10^{-2}, which meets the convergence requirements. Most of the test samples are distributed around the diagonal of the 45-degree regression line, and only a few points are far away from the regression line. In addition, the regression coefficient R = 0.95, which is close to 1, indicating that the model has high fitting accuracy and strong generalization ability. Although there are still small deviations between the predicted aerodynamic performance of the model and the actual aerodynamic performance, it is difficult to find the actual optimal solution for complex problems with a large number of design variables, and a lot of computing costs are required. Considering the reduction of computing cost and design optimization cycle, these deviations can be considered acceptable (Zhang et al., 2018).

Results and discussion

The trained nonlinear model was combined with the multi-objective genetic algorithm for global optimization, and the impellers with flow coefficients of 0.18, 0.2 and 0.22 were designed respectively. The problem is
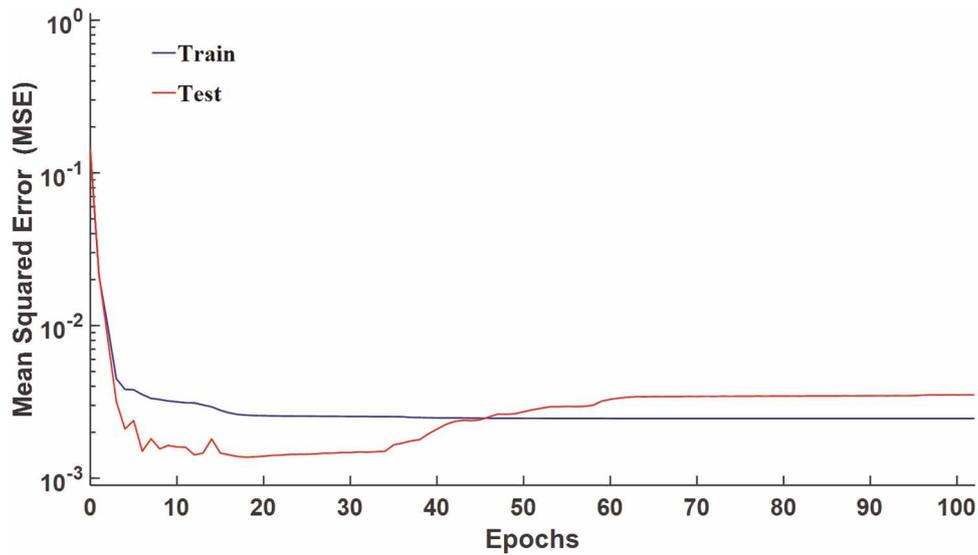
Figure 4. Model training curve.
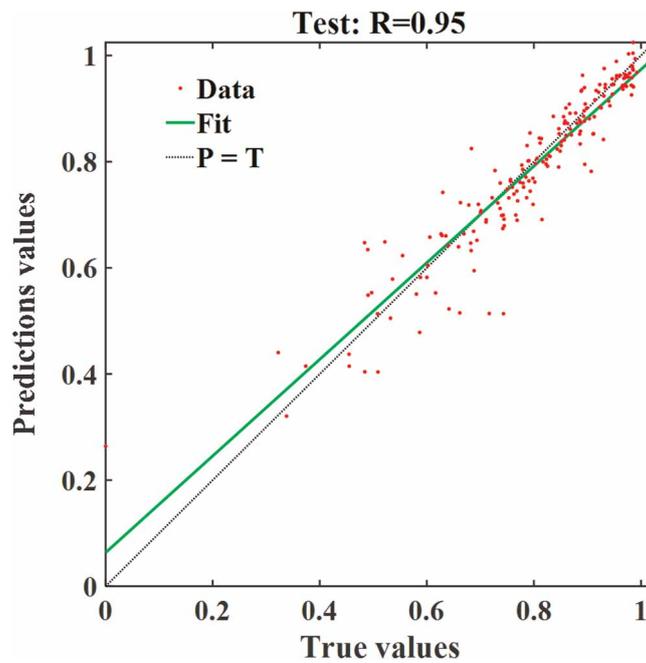
Figure 5. Test set 45 degree regression line.

mathematically formulated Equation 3. The basic parameters of the multi-objective genetic algorithm were set as follows: crossover probability 0.85, mutation probability 0.09, population size 150, and evolution generation 300 generations. After the convergence of algorithm, the best point was verified by CFD. The following is the analysis of results.

$$
\begin{cases}
\min : 1 - \eta_{\mathrm{des}} \\
\min : |\varepsilon_{\mathrm{des}} - \varepsilon_{\mathrm{obj}}| \\
st. \begin{cases}
\min G_i \leq G_i \leq \max G_i \\
\varphi = \varphi_{\mathrm{obj}}
\end{cases}
\end{cases}
\tag{3}
$$

The impellers designed by machine learning CFD (new impellers) results were compared with the CFD results of original impellers. The original impellers were designed by a traditional method, and it had gone through a much longer design and optimization cycle than the new impellers. Figure 6 compares the variable condition performance curve of the new impellers and the original impellers with flow coefficients of 0.18, 0.2 and 0.22. It can be seen that at the design conditions, under the premise that the total pressure ratio meets the design requirements, the isentropic efficiency of the new impellers exceeds the original impellers and also the design requirements. After the main parameters of the impeller are obtained after rapid design, the optimization can be carried out for non-design conditions.

Figures 7–9 shows the relative Mach number contours and streamline at 10%, 50%, and 90% span of the impeller. It can be seen that the airflow is clean, and there is no obvious flow separation and shock wave phenomenon.

In order to study the uniformity of the airflow at the exit of the new impellers, Figures 10–12 respectively shows the absolute Mach number distribution, tangential velocity distribution and flow angle distribution of
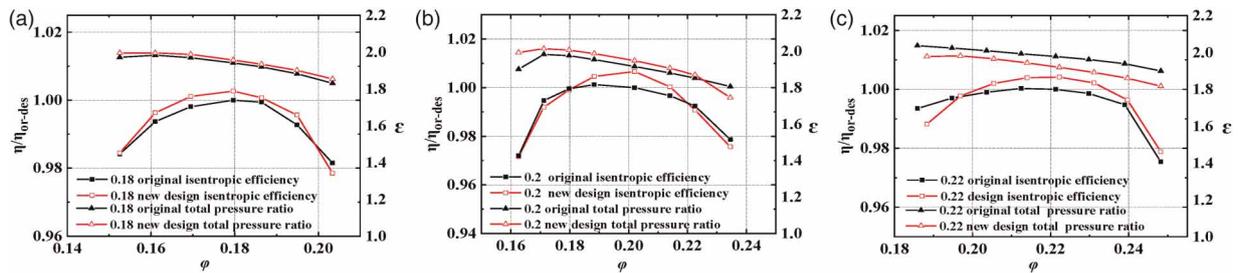
Figure 6. Comparison of performance curves for different impellers. (a) $\varphi = 0.18$. (b) $\varphi = 0.2$. (c) $\varphi = 0.22$.
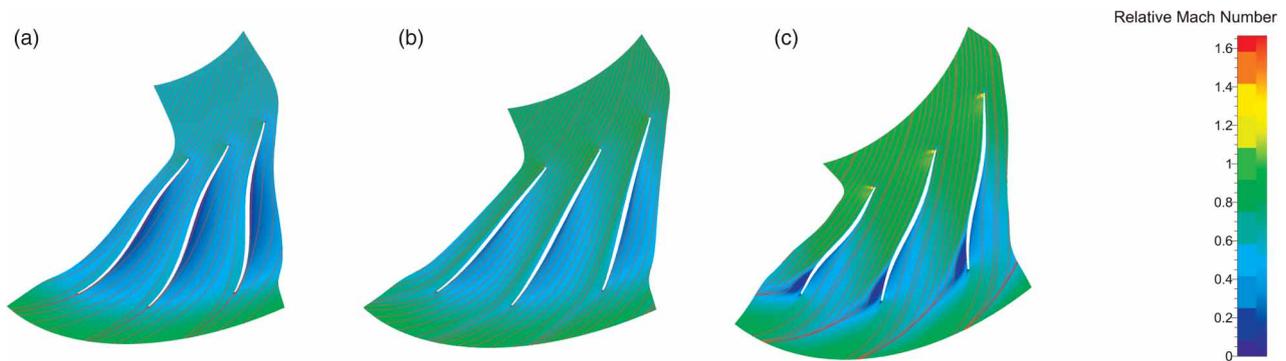
Figure 7. Relative Mach number contours and streamlines of the impelller with a flow coefficient of 0.18. (a) 10% span. (b) 50% span. (c) 90% span.
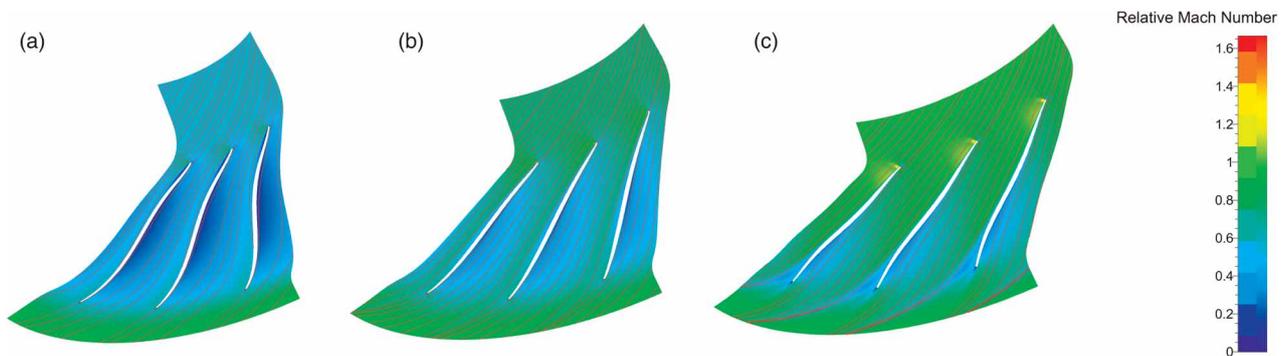
Figure 8. Relative Mach number contours and streamlines of the impeller with a flow coefficient of 0.2. (a) 10% span. (b) 50% span. (c) 90% span.
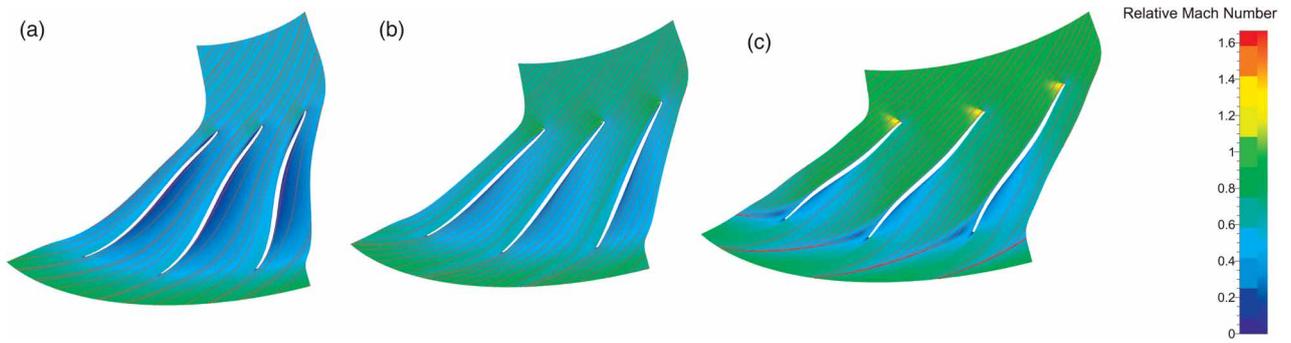
Figure 9. Relative Mach number contours and streamlines of the impeller with a flow coefficient of 0.22 (a) 10% span (b) 50% span (c) 90% span.
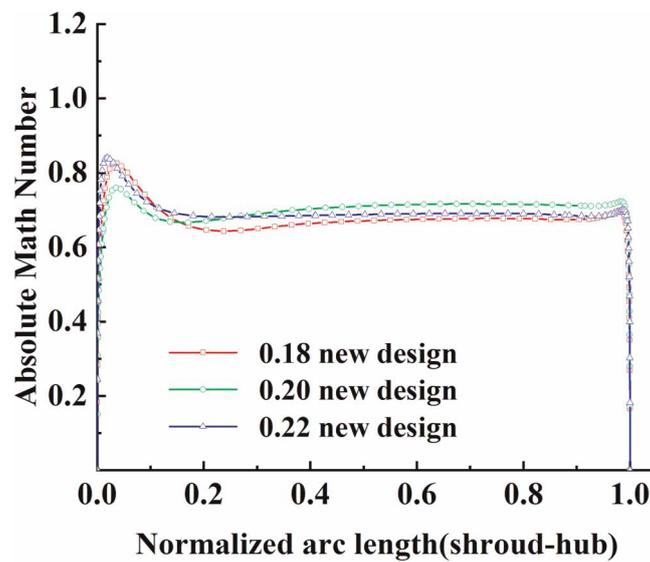
Figure 10. Distribution of absolute Mach number.
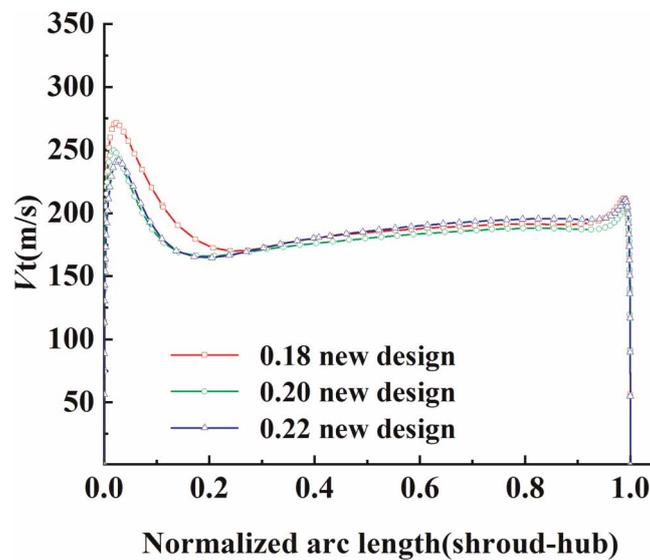
Figure 11. Distribution of tangential velocity.
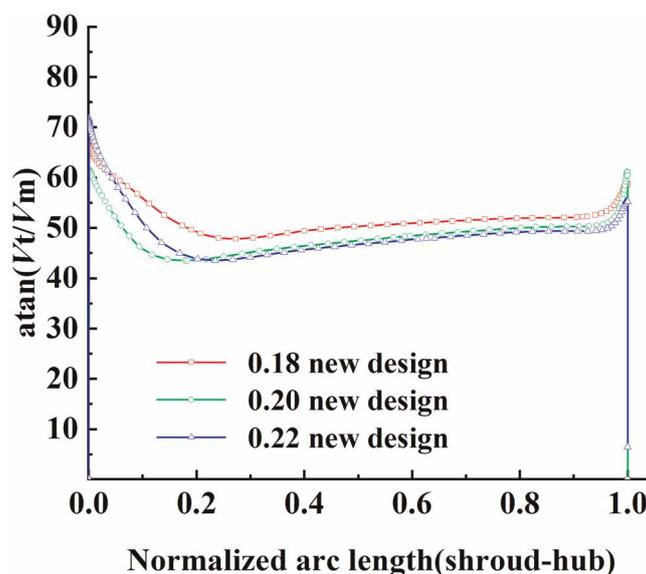
Figure 12. Distribution of flow angle.

three new impellers in the normalized blade exit section (from shroud to hub). It can be seen that the airflow at the exit of the blade is subsonic. The tangential velocity and flow angle fluctuates slightly on the shroud, and the overall distribution is uniform.

The traditional design method requires repeated CFD calculations on the impeller, and a single calculation requires several hours. A single performance prediction of the impeller using machine learning method costs less than 1 second, which greatly shortens the design optimization cycle. The above examples show that the aerodynamic performance of the new impellers designed based on the machine learning method can meet the design requirements and can reach or exceed the original impellers.

Conclusions

In this paper, through the one-dimensional design and optimization, the main geometric parameters of a centrifugal impeller were obtained in a relatively short time to provide a sampling space for the subsequent design. Using the three-dimensional samplings and numerical simulations, a data libarary for centrifugal impellers with large flow coefficient was built. Refering to the data in the libarary and based on the artificial neural networks, a nonlinear model between fgeometric parameters and aerodynamic parameters of an impeller was constructed. Combined with the heuristic optimization algorithm, the rapid design of the centrifugal impellers with flow coefficients of 0.18, 0.2 and 0.22 was carried out respectively. Following conclusions can be drawn:

The centrifugal impeller parameterization method can represent the geometric characteristics of a centrifugal impeller, realize the coupling with one-dimensional design and optimization, and generate three-dimensional samples.

The results of centrifugal impellers with flow coefficients of 0.18, 0.2 and 0.22 show that the non-linear model constructed can be combined with the multi-objective genetic algorithm to carry out the rapid design of centrifugal impellers with flow coefficients (from 0.18 to 0.22). A single performance prediction for the model takes seconds only, which greatly shortens the design optimization cycle.

The quality and sampling range of samples in the sample library determine the coefficient range of rapid design model. In theory, the design range can be broadened to other impellers with other flow coefficients by supplementing the samples, which is also verified by this research. The method explored in the paper shows a certain industrial practicability and development potential.

Nomenclature

q_m	mass flow
D_{1s}	inlet shroud diameter
D_{1h}	inlet hub diameter
β_{1s}	inlet hub blade turning angle

β_{1h}	inlet shroud blade turning angle
β_2	outlet blade turning angle
b_2	outlet blade height
η_{des}	isentropic efficiency under the design condition
$\eta_{or\text{-}des}$	isentropic efficiency of the original impeller under the design condition
ε_{des}	total pressure ratio under the design condition
ε_{obj}	target total pressure ratio
G_i	main geometric parameter values
φ_{obj}	target flow coefficient
R	regression coefficients
V_t	tangential velocity
V_m	meridian velocity

Acknowledgments

Competing interests

Ao Zhang declares that he has no conflict of interest. Yan Liu declares that she has no conflict of interest. Jinguang Yang declares that he has no conflict of interest. Zhi Li declares that he has no conflict of interest. Chuang Zhang declares that he has no conflict of interest.

Funding sources

References

Deb K., Pratap A., Agarwal S., and Meyarivan T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*. 6 (2): 182–197. https://doi.org/10.1109/4235.996017

Duccio B., Andrea A., and Mirco E. (2002). *Analysis and optimization of transonic centrifugal compressor impellers using the design of experiments technique*. In *ASME*. New York: ASME Press, pp. 1089–1098.

Guo X. X., Li W., and Iorio F. (2016). *Convolutional neural networks for steady flow approximation*. In *ACM SigKDD International Conference*. New York: Association for Computing Machinery, pp. 481–490.

Hassoun M. (1996). Fundamentals of artificial neural networks. *Proceedings of the IEEE*. 84 (6): 906–907. https://doi.org/10.1109/JPROC.1996.503146

Ibaeaki S., Tomita I., and Sugimoto K. (2015). Aerodynamic design optimization of centrifugal compressor impeller based on genetic algorithm and artificial neural network. *Mitsubishi Heavy*. 52 (1): 77–82.

Japikse D. (1985). *Assessment of single and two-zone modeling of centrifugal compressors, studies in component performance: Part 3*. In *ASME*. New York: ASME Press, p. 13.

Kim J. H., Choi J. H., Huasin A., and Kim K. Y. (2009). *Design optimization of a centrifugal compressor impeller by multi-objective genetic algorithm*. In *ASME*. New York: ASME Press, pp. 2–6.

Lei M. (2020). Machine learning principles, algorithms and applications. 1st ed. Beijing: Tsinghua University Press.

Tang H. W. and Qin X. Z. (2004). Practical methods of optimization. 4th ed. DaLian: Dalian University of Technology Press, pp. 332–363.

Verstraete T., Alsalihi Z., and Van den braembussche V. A. (2010). Multidisciplinary optimization of a radial compressor for microgas turbine applications. *Journal of Turbomachinery*. 132 (3): 7–14. https://doi.org/10.1115/1.3144162

Wang J. L., Yang J. G., Liu Y., and Zhang T. C. (2019). Multi-objective optimization design of a centrifugal impeller. *Compressor Blower & Fan Technology*. 61 (2): 26–31.

Zhang L. Z., Mi D., Yan C., and Tang F. M. (2018). Multidisciplinary design optimization for a centrifugal compressor based on proper orthogonal decomposition and an adaptive sampling method. *MDPI AG*. 8 (12): 2608.

Zhou Z. H. (2016). Machine learning. 4th ed. Beijing: Tsinghua University Press.

Zhu C. Y., Ju Y. P., and Zhang C. H. (2016). Aerodynamic design and analysis of impeller and diffuser of centrifugal compressor with large flow coefficient. *Compressor Blower & Fan Technology*. 2016 (5): 57–62.