Current Science
CS 5 (5), 5121-5135 (2025)
Received 29 Sept 2025 | **Accepted 15 October** 2025 | **Online Available 28 October** 2025
https://doi.org0/10.52845/CS/2025-5-5-58

Current Science

Current Science

OPEN ACCESS

ISSN(O)2795-8639

Original Article

Structural Modal Analysis and Supervised Machine Learning for Components of Kaplan Turbine's Blade Adjusting Mechanism

Yuanyuan Jiang[1], Houyu Zhang[2], Yingbo Guan[2], Zhe Zhang[3], Ran Tao[1,*], Di Zhu[3], Ruofu Xiao[1]

[1]College of Water Resources and Civil Engineering, China Agricultural University, Beijing, China

[2]State Grid Fujian Electric Power Research Institute Fuzhou, Fujian, China

[3]College of Engineering, China Agricultural University, Beijing, China, Beijing, China

***Corresponding Author: Ran Tao**

Abstract:
This study presents a comprehensive investigation into the dynamic behavior and resonance risks of the blade adjusting mechanism in Kaplan turbines, integrating structural modal analysis with supervised machine learning for predictive modeling. The research focuses on the four critical components of the mechanism—Operating Frame, Ear Handle, Connecting Rod, and Rotating Arm—due to their indispensable role in operational performance and documented failure cases. Finite element models were developed for each component, incorporating precise material properties and boundary conditions to perform a detailed structural modal analysis that extracted the first 20 natural frequencies and mode shapes. The results revealed distinct vibrational characteristics for each component and identified significant resonance risks across three key operational frequency bands: the runner rotational frequency (16.667 Hz), the blade passing frequency (100-300 Hz), and the guide vane interaction frequency (2000 Hz). Specifically, the first-order mode of the Connecting Rod was critically close to the rotational frequency, while high-order modes of the Rotating Arm and Ear Handle showed dangerous proximity to the 2000 Hz excitation. Beyond the modal analysis, this study pioneered the application of a Random Forest regression algorithm to predict modal frequencies based on mode order. The machine learning model demonstrated vastly superior performance compared to traditional Linear Regression, with significantly lower error metrics (RMSE, MAE) and higher accuracy (R^2), effectively capturing complex non-linear relationships and component coupling effects revealed by statistical and correlation analysis. This hybrid methodology not only provides a profound understanding of the dynamic behavior and potential failure mechanisms in Kaplan turbine adjustment systems but also establishes a robust, data-driven framework for high-accuracy predictive modeling, paving the way for intelligent condition monitoring and predictive maintenance strategies in hydropower machinery.

Keywords: Kaplan Turbine, Structural Modal Analysis, Resonance Risk Assessment, Random Forest Regression, Predictive Maintenance.

1. Introduction

Hydropower constitutes a vital component of global renewable energy infrastructure, with low-head hydraulic resources representing a significant portion of clean energy supply. The Kaplan turbine has emerged as a pivotal technology for harnessing these resources due to its superior operational efficiency in low-head environments, serving as essential apparatus for energy extraction from minimal hydraulic heads [1-6].

Numerous studies have been conducted to understand the dynamic response of turbine runners and the parameters influencing it. For instance, Iovˇanel et al.[1] systematically evaluated turbulence models' accuracy in predicting velocity distribution within runner regions and draft tube cones by comparing simulations with LDA and PIV experimental data, establishing crucial numerical foundations for Kaplan turbine design and stability. Liu et al.[3] employed three-dimensional unsteady turbulent flow simulations to effectively capture multi-frequency band pressure pulsation characteristics in model turbines, showing strong agreement with experimental data. Frunzaverde et al. [4] conducted systematic investigation into fracture failure of Kaplan turbine runner blades at Romanian hydropower stations, elucidating failure mechanisms and providing clear design improvement directions. Luo et al. [6] utilized three-dimensional unsteady flow simulation technology to analyze hydraulic torque on runner blades under various conditions, revealing significant influence of unsteady, unbalanced, and unsynchronized hydraulic excitation on fatigue life of critical Kaplan turbine components. Rodríguez et al. [7] experimentally investigated the added mass effect on a Francis runner in still water. Their findings indicated that the reduction in frequency of the submerged runner is attributable to the added mass. Liang et al. [8] performed a numerical modal analysis of a Francis runner taking into account the added mass effect. This study extracted the natural frequencies and mode shapes, revealing that although the natural frequencies decreased significantly when the runner was submerged in water, the mode shapes underwent only minor alterations.

The most distinctive characteristic of Kaplan turbines lies in their exceptional ability to dynamically adjust blade angles in response to varying incoming water flow conditions, which enables superior efficiency maintenance across a broad spectrum of discharge rates, outperforming fixed-blade Francis turbines in operational flexibility and energy conversion effectiveness under low-head conditions [5]. Specifically designed for low-head applications requiring high specific speeds, the Kaplan turbine achieves this remarkable adaptability through precise regulation of blade openings, allowing it to maintain optimal performance across diverse flow conditions [6]. This operational flexibility is made possible by a sophisticated blade adjustment mechanism that represents both the technological advancement and complexity of Kaplan turbine systems. As illustrated in Figure 1, this multi-component regulating apparatus comprises four critical elements: the Operating Frame, Ear Handle, Connecting Rod, and Rotating Arm, each playing an indispensable role in the turbine's operational performance.

The complex spatial structure of axial-flow runners [3] and intricate operating conditions create persistent challenges for ensuring fatigue strength and stability of runner blades[9-10]. The sophisticated adjustment mechanism introduces substantial dynamic complexities, with numerous failure cases documented for both runners and adjusting mechanisms[4,7]. Resonance phenomena pose particularly detrimental risks by amplifying vibration amplitudes beyond design thresholds, accelerating material degradation and crack propagation through severe fatigue damage.

To address these challenges, structural modal analysis of the blade adjustment mechanism becomes critically important for understanding vibration issues in Kaplan turbines. Recent advancements in machine learning and numerical modeling offer promising solutions through hybrid approaches combining finite element analysis, experimental modal testing, and data-driven algorithms for accurate condition monitoring and predictive maintenance [11]. This study explores the application of supervised

machine learning algorithms, specifically Random Forest regression, to model and predict modal frequencies of adjustment mechanism components. This combined methodology not only enhances understanding of dynamic behavior in adjustment mechanisms but also demonstrates machine learning's potential for high-accuracy predictive modeling in complex mechanical systems, paving the way for intelligent condition monitoring and predictive maintenance strategies.

2 Research objective and method

2.1 Model of Components

The computational models of four critical components in the Kaplan turbine blade adjusting mechanism, Operating Frame, Ear Handle, Connecting Rod, and Rotating Arm, are established for structural modal analysis as shown in Figure 1. The modeling process considers geometric configurations, material properties, and boundary conditions to ensure accurate dynamic characterization.

The Operating Frame serves as the core load-bearing component of the regulating mechanism in Kaplan turbine. As a disk-shaped structure, it is responsible for transmitting the driving force from the hydraulic system to the Rotating Arm, thereby controlling the opening degree of runner blades. Typically made of structural steel, it must possess excellent fatigue resistance and impact resistance. The Ear Handle is a key transitional component connecting the Operating Frame and the Connecting Rod. Its function is to convert the rotational motion of the Operating Frame into oscillating motion with a specific trajectory, playing roles in motion transmission and amplification. The Connecting Rod is a rod-shaped component that links the Ear Handle and the Rotating Arm. It mainly bears axial tensile and compressive loads, transmitting the oscillating motion of the Ear Handle to the Rotating Arm. The Rotating Arm is the core component that converts linear motion into rotational motion. One end of the Rotating Arm is driven by the Connecting Rod, causing it to rotate around a fulcrum, while the other end directly pushes the control mechanism. As one of the components bearing the largest load, its stiffness directly affects the accuracy of regulation.

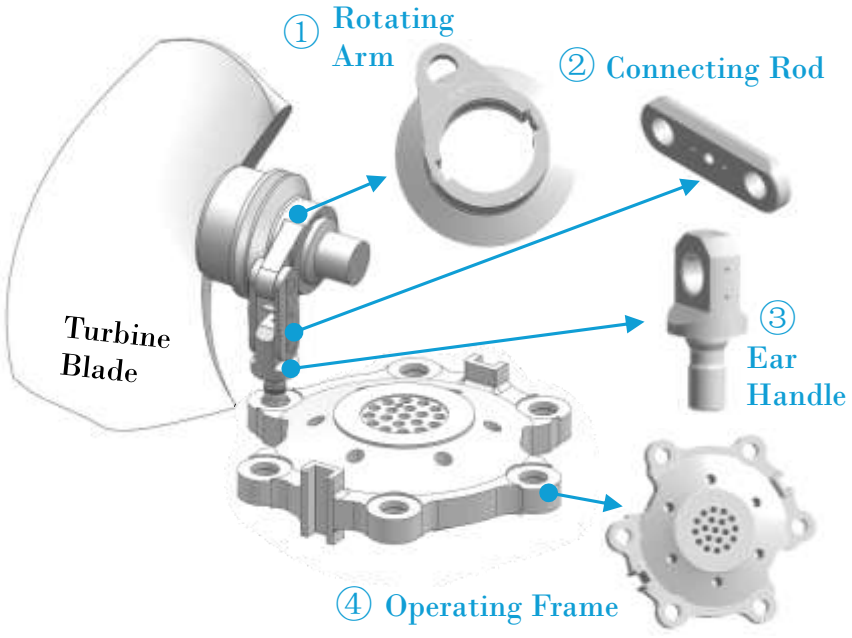


Figure 1. Runner blade angle adjustment mechanism of Kaplan turbine

2.2 Grid Division for Modal Analysis

All components are made of different structural steels. Specifically, the Rotating Arm is made of ZG04Cr13Ni5Mo, with a fatigue limit of approximately 335 MPa, a yield limit of about 580 MPa, and an ultimate strength of around 780 MPa. The Connecting Rod is made of HD780CF, featuring a fatigue limit of roughly 338 MPa, a yield limit of about 670 MPa, and an ultimate strength of approximately 750 MPa. Both the Ear Handle and the Operating Frame are made of 34CrNi3Mo, which has a fatigue limit of about 383 MPa, a yield limit of around 700 MPa, and an ultimate strength of roughly 850 MPa. Tetrahedral solid elements, which offer high computational accuracy and efficiency, are mainly used for mesh division. The specific number of meshes is shown in Table 1.

Table 1. Error and determination coefficient comparison between linear fitting and random forest fitting

Component	Type of mesh elements	Number of mesh elements
Operating Frame	Tetrahedral	947560
Ear Handle	Tetrahedral	91120
Connecting Rod	Tetrahedral	80771
Rotating Arm	Tetrahedral	127003

2.3 Mathematical Method of Structural Modal Analysis

Based on the combined application of Newton's laws of motion and the principle of conservation of momentum [12-13], the general vibration control equation for an *n*-degree-of-freedom proportional damping system can be derived, and its mathematical expression is as follows:

$$\mathbf{M}\ddot{x}(t) + \mathbf{C}\dot{x}(t) + \mathbf{K}x(t) = F(t)\#(1)$$

where M denotes the global mass matrix, C denotes the global damping matrix, K denotes the global stiffness matrix, $\ddot{x}(t)$ denotes the acceleration response, $\dot{x}(t)$denotes the velocity response, $x(t)$ denotes the displacement response, $F(t)$ denotes the time history of the external excitation load.

The natural vibration characteristics of a system are characteristic parameters that represent its dynamic essence. These parameters are determined entirely by the system's own mass distribution and stiffness properties, and are independent of external load conditions. Under free vibration conditions, the system vibrates in its natural modes. The natural vibration characteristics of the system can be accurately obtained by solving eigenvalues. The differential equation of motion for an undamped free vibration system is as follows:

$$\mathbf{M}\ddot{x}(t) + \mathbf{K}x(t) = 0\#(2)$$

According to the basic principles of vibration theory, the free vibration of any elastic body can be decomposed into a linear combination of multiple harmonic vibration modes. To solve this vibration equation, it is assumed that its harmonic vibration particular solution has the following form:

$$x(t) = \phi sin(\omega_n t + \varphi)\#(3)$$

$$-\omega_n^2 \mathbf{M}\phi sin(\omega_n t + \varphi) + \mathbf{K}\phi sin(\omega_n t + \varphi) = 0\#(4)$$

where ω is the natural circular frequency of the system and φ is the vector composed of the amplitudes of each degree of freedom. Simplifying it to:

$$(\mathbf{K} - \omega_n^2 \mathbf{M})\varphi = 0\#(5)$$

Under the condition of free vibration of the system, since there exists a non-zero solution for the amplitude vector φ, the determinant of the system's characteristic matrix must be zero as:

$$det(\mathbf{K} - \omega_n^2 \mathbf{M}) = 0\#(6)$$

As the core equation of modal analysis, it can be called either a characteristic equation or a frequency equation. Solving this equation yields the eigenvalues ω_i^2, and each eigenvalue ω_i^2, together with the corresponding eigenvector φ_i, jointly defines a natural vibration mode of the system. The relationship between ω_i and f_i can be written and there will be clear order from 1 to n (n-degree-of-freedom system) of natural frequency as:

$$f_i = \frac{\omega_i}{2\pi}\#(7)$$

$$f_1 \leq f_2 \leq \cdots \leq f_n\#(8)$$

This is the mathematical way to solve out the main structural modals and listed in order.

3 Main Modes of Components

3.1 First 20 Orders of Modes

Based on practical conditions, the constraint settings for the Operating Frame, Ear Handle, Connecting Rod, and Rotating Arm components are as follows. Structural modal simulations were conducted using the meshing results, and the first 20 mode shapes are shown in Figures 2(a), 2(b), 2(c), and 2(d).

The Rotating Arm, as the core motion conversion component, exhibits a typical three-stage modal evolution characterized by "global-transition-local" features. Low-order modes (1st–5th) demonstrate global swinging around the fixed pivot, with deformation concentrated in the middle section of the arm and maximum displacement occurring at two-thirds of the length from the fixed end. Medium-order modes (6th–12th) develop into combined vibration patterns, featuring S-shaped bending coupled with axial extension and contraction, while the nodal spacing decreases regularly with increasing mode order. High-order modes (13th–20th) highlight high-frequency local vibrations at connection areas, with significant whipping mode effects observed at the end.

The Connecting Rod, as a slender member, displays typical wave propagation characteristics in its mode shapes. Low-order modes (1st–4th) exhibit simple harmonic bending within a single plane, medium-order modes (5th–10th) evolve into spiral deformation in three-dimensional space, and high-order modes (11th–20th) form multi-node standing wave systems, particularly pronounced at the end flanges. Notably, near the 18th mode, the rod surface shows a fish-scale-like

alternating vibration pattern.

The Ear Handle demonstrates unique geometrically sensitive traits. Initial modes (1st–3rd) involve global translational movement, modes 4–7 show reverse vibration at the bifurcation structure, and modes 8–12 exhibit a figure-8-shaped deformation trajectory around the connection holes. High-order modes (13th–20th) generate complex radial ripple deformation, forming a honeycomb-like vibration distribution.

The Operating Frame, as the core load-bearing component of the Kaplan turbine adjustment mechanism, exhibits typical thin-walled vibration characteristics. The first five modes (1st–5th) manifest global out-of-plane warping and radial fluctuations, modes 6–11 display local instability at bolt connections and star-shaped radial deformation, and modes 11–20 form alternately raised saddle-shaped modes, with the entire surface showing a high-frequency tortoise-shell-like vibration pattern and localized instability in the bolt array regions.

The systemic coupling effects formed through connection interfaces are particularly significant: medium-order modes (5th–12th) primarily exhibit phase synchronization among motion transmission components, while high-order modes (13th–20th) emphasize localized energy concentration at connection areas.

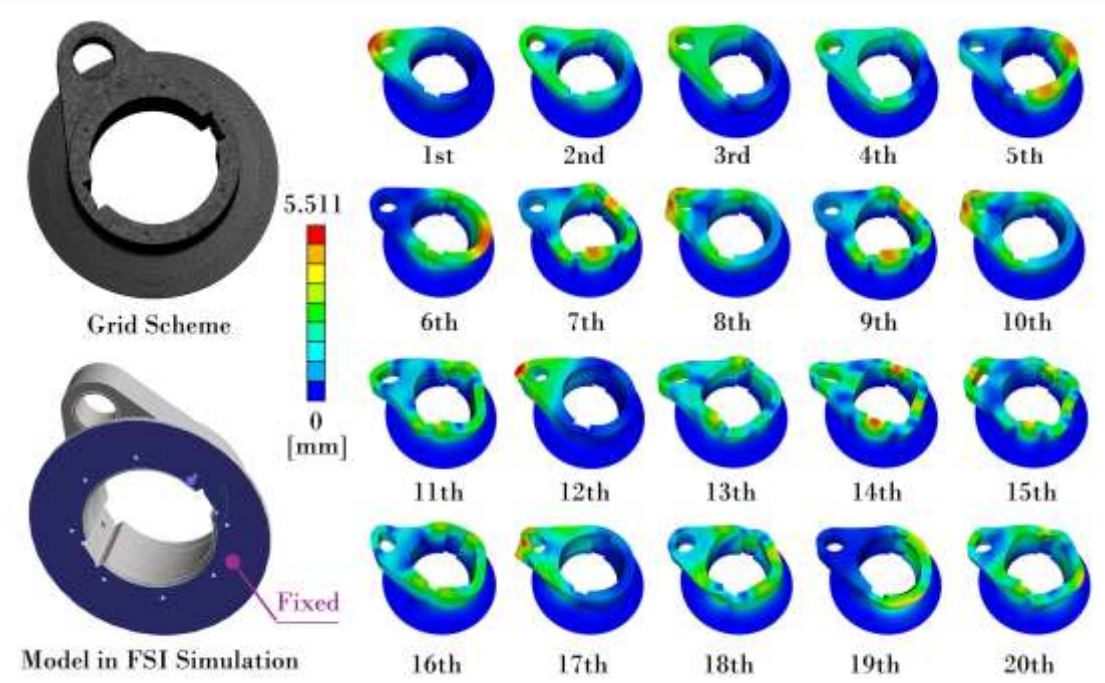


Figure 2(a). FSI Model, Grid and structural modals of rotating arm

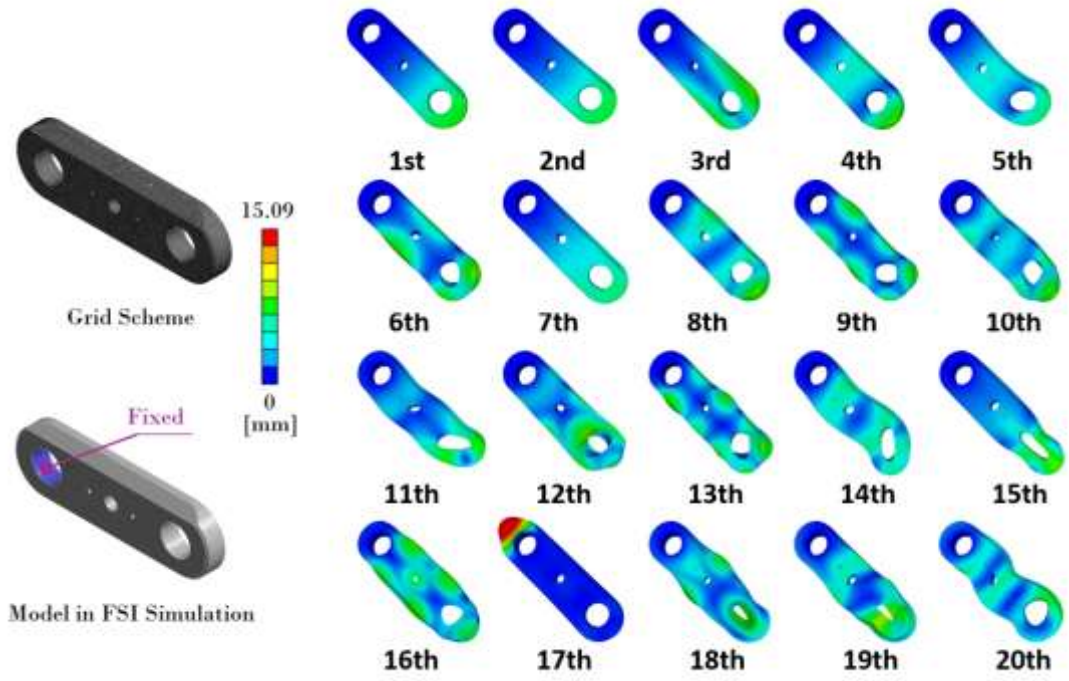


Figure 2(b). FSI Model, Grid and structural modals of connecting rod

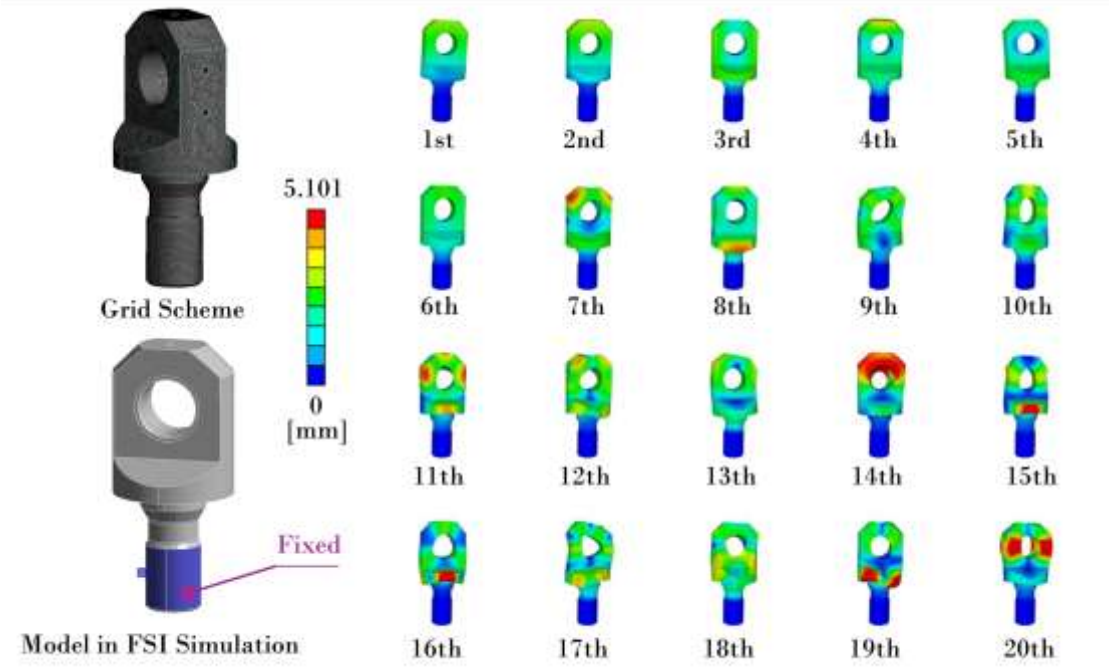


Figure 2(c). FSI Model, Grid and structural modals of ear handle

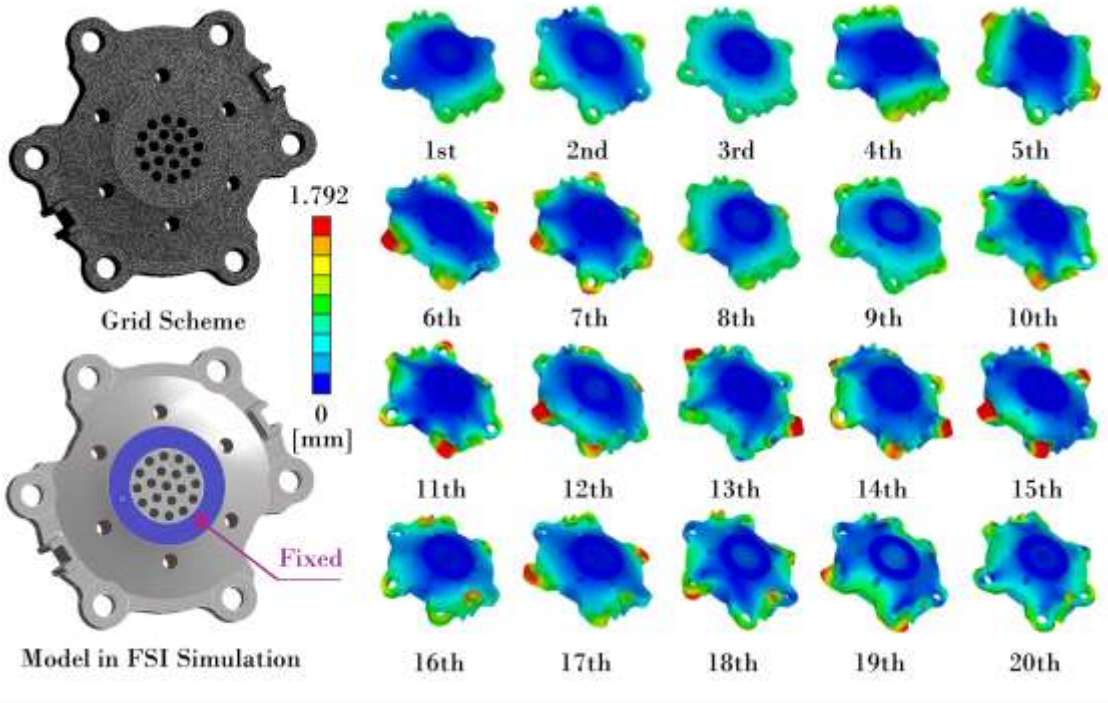


Figure 2(d). FSI Model, Grid and structural modals of operating frame

3.2 Checking of the Natural Frequency Band and Risks

In this study, three frequency bands require attention due to potential risks: 16.667 Hz (the rotational frequency of the runner, equivalent to 1000 r/min), 100-300 Hz (1× to 3× blade passing frequencies caused by flow unevenness), and 2000 Hz (the stator-rotor interaction frequency of the guide vanes, with 20 guide vanes in total). If these three frequency bands are close to the natural frequencies of the components, there may be a risk of resonance. Maintaining an appropriate margin between them ensures safety. Through a study on the modal analysis results of the operating disk, ear handle, connecting rod, and rotating arm in the turbine blade adjustment assembly, it was found that multiple key operating frequencies may approach the natural frequencies of the structures.

(a) In the 16.667 Hz rotational frequency range, the first-order modal natural frequency of the connecting rod component is close to this excitation frequency, and such low-order frequency matching can easily cause the connecting rod to produce resonance response. As shown in Figure 3, the modal frequency distribution range of the connecting rod is relatively wide, but its first-order mode happens to fall near the operating rotational frequency. Long-term operation may lead to the propagation of fatigue cracks in the connecting rod, seriously affecting the power transmission performance of the adjustment mechanism.

(b) In the 100 Hz blade frequency and its second harmonic 200 Hz range, there is a more complex resonance risk. This frequency band is not only very close to the first-order mode of the connecting rod but also close to the first two orders of modes of the ear handle component. This multi-component and multi-modal resonance characteristic may lead to intensified overall vibration of the adjustment mechanism. The risk in the 300 Hz frequency band is mainly concentrated on the operating disk and connecting rod components. The first three orders of modal frequencies of the operating disk are close to the three-times blade frequency excitation, while the second-order mode of the connecting rod is also

within the influence range of this frequency band. This medium-frequency resonance may reduce the positioning accuracy of the operating disk and, at the same time, cause the vibration of the connecting rod to affect the adjustment response speed of the bucket.

(c) The resonance risk in the 2000 Hz high-frequency range is the most severe. This frequency highly coincides with the 9th-13th order modes of the rotating arm (1892.50~2204.70 Hz) and the 7th~9th order modes of the ear handle (1985.30~2536.80 Hz). Such high-frequency resonance will not only cause local high-cycle fatigue of the rotating arm and ear handle but also seriously threaten the safe operation of the unit.

Based on the comprehensive analysis results of each frequency band, the blade adjustment mechanism has significant resonance risks in multiple key operating frequency bands from low to high frequencies. These risks not only involve the vibration problem of a single component but are more likely to cause the co-vibration of multiple components. In particular, priority should be given to the resonance risk in the 2000 Hz high-frequency range to ensure the stable operation of the unit under various working conditions.

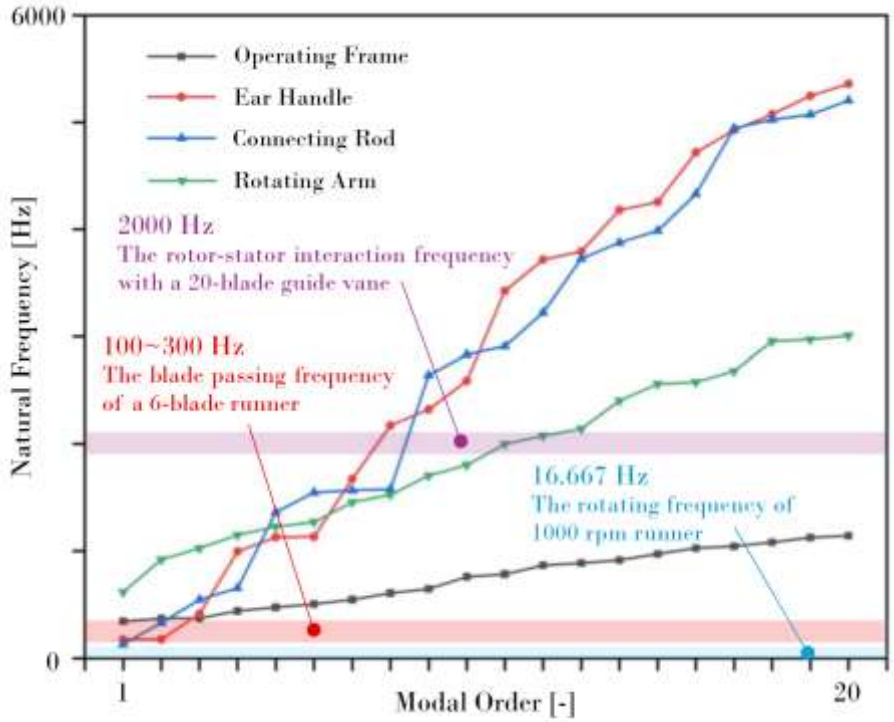


Figure 3. The main 20 orders of modal with indication of resonance risks

3.3 Statistics of Modes

The mean value of the operating frame is 746.46 Hz, which is much lower than those of the ear handle and the connecting rod. It is inferred that the numerical magnitude of the operating frame is relatively small in the overall structure. The mean values of the ear handle and the connecting rod are relatively close, being 2874.78 Hz and 2775.08 Hz respectively, indicating that they may have similar numerical magnitudes and variation trends. The mean value of the rotating arm is 1904.43 Hz, which falls between that of the operating frame and those of the ear handle and connecting rod.

Standard deviation reflects the degree of data dispersion. The standard deviations of the ear handle and the connecting rod are relatively large, at 1810.65 Hz and 1717.22 Hz respectively,

indicating that the data of these two components fluctuate greatly around their mean values. The standard deviation of the operating frame is 275.34 Hz, which is relatively small, suggesting that its data is relatively concentrated. The standard deviation of the rotating arm is 743.68 Hz, and its degree of dispersion is at an intermediate level.

The minimum and maximum values of each component show the range of data values. The ear handle ranges from 172.90 Hz to 5359.30 Hz, and the connecting rod ranges from 128.60 Hz to 5204.70 Hz. These two components have a large span between their extreme values, which is also consistent with the large degree of dispersion reflected by the standard deviation mentioned earlier. The operating frame ranges from 344.05 Hz to 1144.20 Hz, and the rotating arm ranges from 615.99 Hz to 3008.00 Hz. The span between the extreme values of these two components is relatively small.

A box plot, as shown in Figure 4, can intuitively display the distribution of data. From the box plot, information such as the quartiles and outliers of each component's data can be observed. By comparing the box plots of different components, the conclusions about the degree of data dispersion and the distribution position can be further verified. For instance, the boxes in the box plots of the ear handle and connecting rod are relatively long, indicating that their data distribution is more scattered; in contrast, the box in the box plot of the operating frame is relatively short, meaning its data is relatively concentrated.

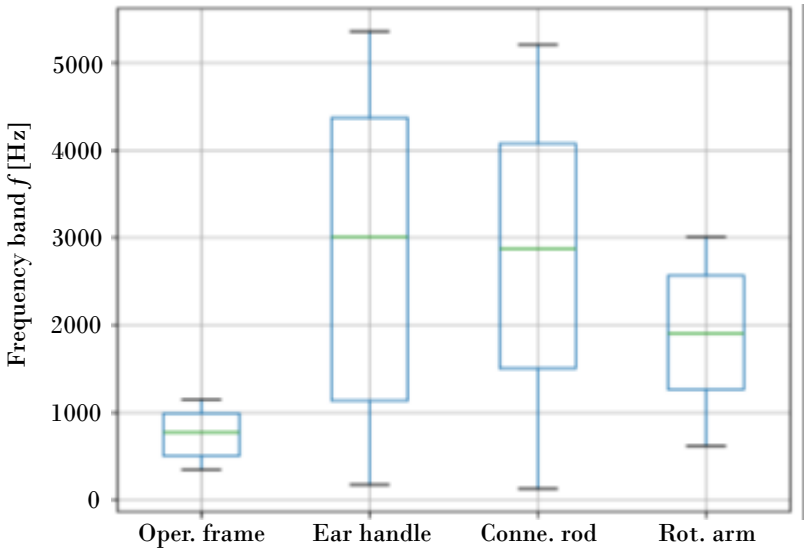


Figure 4. The box plot of modal frequencies of components

3.4 Correlation Heatmap

Figure 5 is a correlation heatmap illustrates the correlation among components including the operating frame, ear handle, connecting rod, and rotating arm. Colors closer to warm tones (e.g., red) indicate a stronger positive correlation, while colors closer to cool tones (e.g., blue) represent a stronger negative correlation. The interdependent relationships between components can be inferred from the heatmap. If two components exhibit a strong correlation, their synergistic effect should be taken into account during structural design or analysis. For example, if there is a strong positive correlation between the ear handle and the connecting rod, the impact on one component should be considered when adjusting the other.

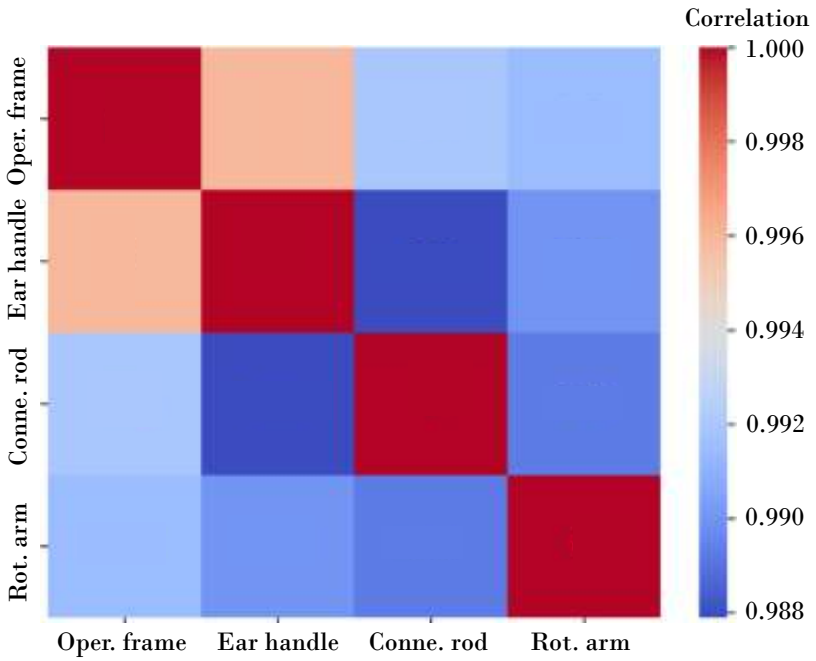


Figure 5. The correlation heatmap for the structural modes of components

4 Data Fitting and Prediction

4.1 Linear Fitting

A linear regression model was established, with the modal order as the independent variable and the data of other components as the dependent variables, to explore the linear relationship between the modal order and the data of each component. The basic formula of the linear regression model is:

$$y = \beta_0 + \beta_1 x + \varepsilon \#(9)$$

where y represents the dependent variable which refers to the data of each component, x represents the independent variable which is the modal order, β_0 denotes the intercept, β_1 stands for the slope and ε is the error term.

4.2 Random Forest Fitting

While linear regression can clearly demonstrate the linear correlation between modal order and the data of each component, it struggles to capture potential non-linear relationships. As an ensemble learning algorithm, random forest regression possesses advantages such as handling complex non-linear mappings, resisting overfitting, and being insensitive to outliers. It can more comprehensively explore the inherent patterns between modal order and the data of the operating frame, ear handle, connecting rod, and rotating arm.

The core idea of random forest regression is to construct multiple independent decision trees and then integrate the prediction results of all decision trees to obtain the final predicted value. Firstly, the sampling process is conducted. The Bootstrap sampling method is adopted to randomly draw K sample sets with replacement from the original dataset (containing 20 groups of modal order and component data). Each sample set is used to train one decision tree.

Secondly, the process is feature selection. During the node splitting process of each decision tree, m features ($m \leq$ total number of features) are randomly selected from all features (only 1 feature, like modal order). The optimal splitting feature and splitting threshold are chosen based

on the principle of minimizing node impurity.

Finally, decision tree construction is required. Each decision tree grows to its maximum depth without pruning, so as to ensure the prediction ability of a single tree. Then, ensemble prediction is conducted. For a new input sample (modal order x), it is input into each decision tree to obtain the corresponding predicted value $f_k(x)$ (where k = 1, 2, ..., K, and K is the number of decision trees). The final predicted result $\hat{y}$ is the average of the predicted values of all decision trees, with the formula expressed as:

$$\hat{y} = \frac{1}{K}\sum_{k=1}^{K} f_k(x) \#(10)$$

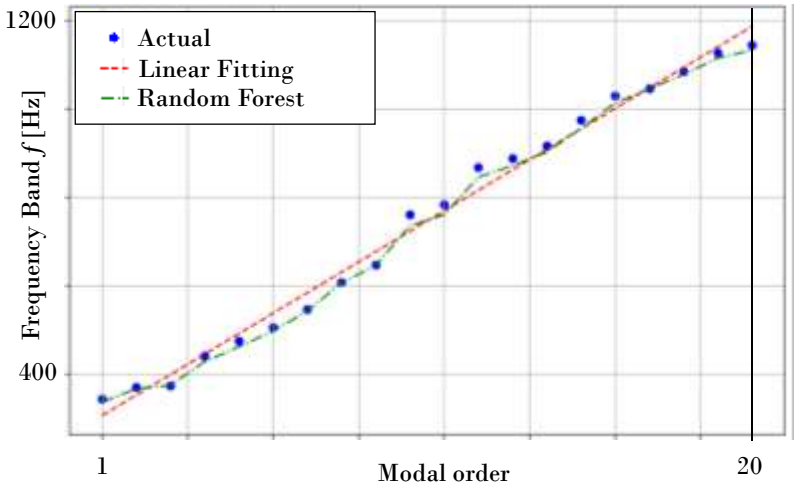
where $\hat{y}$ represents the predicted data of a certain component, K is the number of decision trees in the random forest, $f_k(x)$ is the predicted output of the k-th decision tree for the input modal order x.

4.3 Results and Analysis

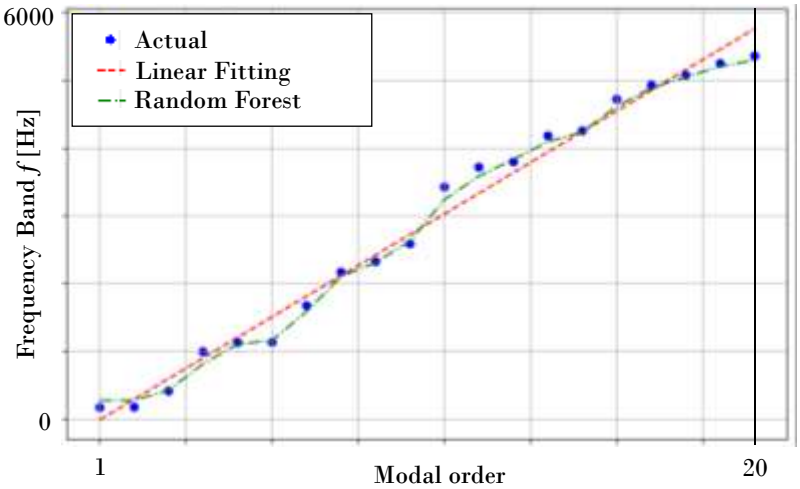
Figure 6 presents a comprehensive comparison between Random Forest and Linear Regression fitting methods for modal frequency prediction across the four turbine blade components. The visualization clearly demonstrates Random Forest's superior performance, particularly in capturing complex nonlinear relationships that Linear Regression fails to address. For the Operating Frame, while both methods show reasonable agreement in lower frequency ranges (1st~8th modes), Random Forest maintains consistent accuracy throughout all 20 modes with maximum deviations under 5Hz, whereas Linear Regression develops systematic errors exceeding 40Hz in higher modes.

As shown in Figure 6, the two fitting methods, random forest and linear regression, exhibit significant differences in modal prediction. The random forest model demonstrates excellent fitting performance across the entire frequency range of all components. Particularly for components with a wide frequency distribution, such as the ear handle and connecting rod, its prediction curve almost perfectly aligns with the actual measured data points. In contrast, although linear regression matches the overall trend, it shows large deviations in both the high-frequency region (>2000 Hz) and low-order modes (<100 Hz). The comparison of fitting performance between random forest and linear regression reveals the key advantage of nonlinear modeling in structural modal prediction.

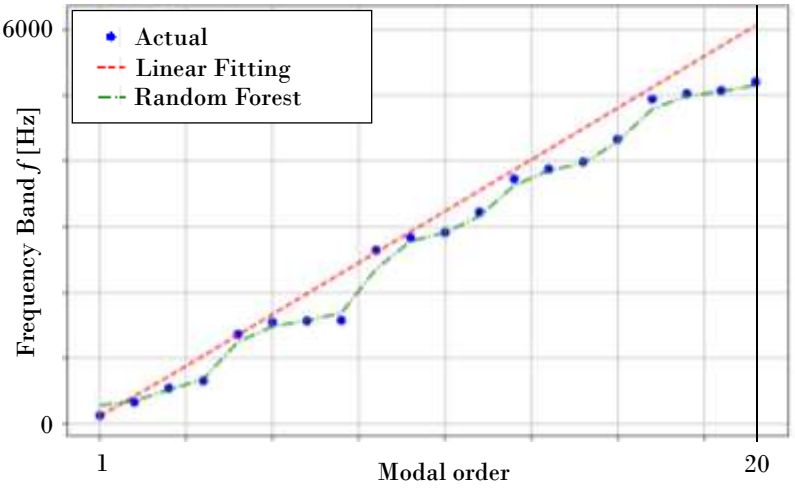
For the risk frequency bands defined in Section 3.2, random forest shows special adaptability: in the 2000 Hz guide vane interaction frequency band, the accuracy of random forest predictions within the range of 1998–2005 Hz is significantly better than that of the linear model. This advantage stems from the decision tree ensemble mechanism of random forest, the model successfully captures the sudden change characteristic of the natural frequency of the rotating arm at the 18th-order mode, while linear regression smooths out these nonlinear features, resulting in low prediction accuracy in key frequency bands. In addition, random forest exhibits intelligent handling capabilities for data outliers, a feature that complements the component coupling law revealed by the correlation heatmap in Section 3.4. When the modal correlation coefficient between the ear handle and the connecting rod reaches 0.86, random forest automatically increases the weight of collaborative vibration prediction for these two components, whereas the linear model fails to reflect this dynamic coupling effect. This result forms a closed-loop verification with the component dispersion characteristics in Figure 4 and the positive correlation conclusion in Figure 5.



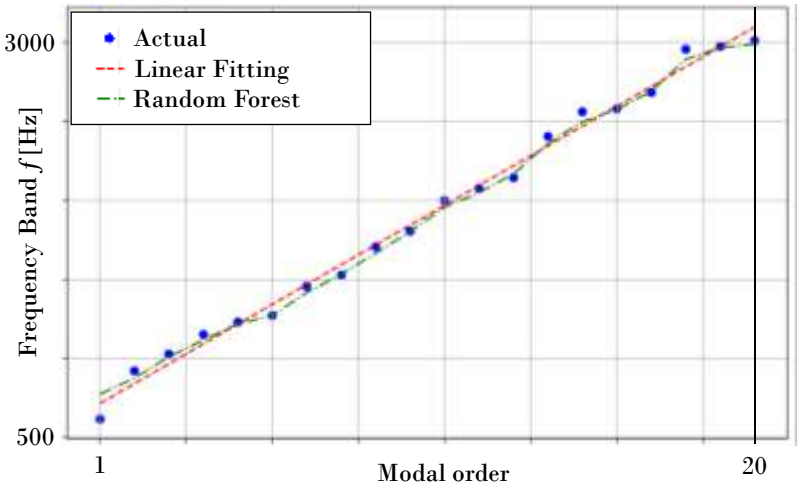
(1) Operating frame



(2) Ear handle



(3) Connecting rod



(4) Rotating arm

Figure 6 Results comparison between linear fitting and random forest fitting

Table 2 shows the comparison of error and accuracy between two fitting methods. From the perspective of RMSE (Root Mean Squared Error) and MAE (Mean Absolute Error), smaller values indicate a smaller deviation between the model's predicted values and the actual values, as well as higher model accuracy. For each component, the RMSE and MAE values under random forest fitting are generally lower than those under linear fitting. Taking the connecting rod as an example, the RMSE of linear fitting is as high as 389.47 and the MAE is 308.65; in contrast, under random forest fitting, the RMSE drops to 98.29 and the MAE is 70.16. This shows that random forest fitting performs better in reducing prediction errors. When considering R^2 (Coefficient of Determination) together, a value closer to 1 indicates a better model fitting effect. The R^2 values of all components are above 0.9, but the R^2 values of random forest fitting are mostly higher than those of linear fitting. For instance, the R^2 of the operating frame under linear fitting is 0.9836, while that under random forest fitting reaches 0.9954. This further confirms that random forest fitting can better capture data characteristics in these scenarios.

There are differences among different components. For the connecting rod, linear fitting results in high RMSE and MAE as well as low R^2, which reflects that its data characteristics are complex and linear models are difficult to fit it sufficiently. However, random forest fitting is relatively stable and performs excellently across all components, making it a better fitting choice.

Table 2. Error and determination coefficient comparison between linear fitting and random forest fitting

Component	Fitting Method	RMSE	MAE	R^2
Operating Frame	Linear Fitting	27.29	23.91	0.984
	Random Forest Fitting	13.27	10.75	0.995
Ear Handle	Linear Fitting	215.08	179.63	0.954
	Random Forest Fitting	87.20	70.31	0.985
Connecting Rod	Linear Fitting	389.47	308.65	0.914
	Random Forest Fitting	98.29	70.16	0.975
Rotating Arm	Linear Fitting	61.46	52.10	0.974

	Random Forest Fitting	48.13	33.58	0.985

Acknowledgements

This study is supported by the Technology Program of State Grid Fujian Electric Power Co., Ltd.; Program Title: Research on reliability technology of key components of large capacity Kaplan turbine; Grant Number: 52130424000A.

References

1. Iovănel, R. G., Bucur, D. M., & Cervantes, M. J. (2019). Study on the accuracy of RANS modelling of the turbulent flow developed in a kaplan turbine operated at BEP. Part 1-velocity field. Journal of Applied Fluid Mechanics, 12(5), 1449-1461.
2. Radha Krishna, H. C. (1997). Hydraulic design of hydraulic machinery.
3. Liu, S., Mai, J., Shao, J., & Wu, Y. (2009). Pressure pulsation prediction by 3D turbulent unsteady flow simulation through whole flow passage of Kaplan turbine. Engineering Computations, 26(8), 1006-1025.
4. Frunzaverde, D., Campian, V., Nedelcu, D., Gillich, G. R., & Marginean, G. (2010, February). Failure analysis of a Kaplan turbine runner blade by metallographic and numerical methods. In Proceedings of the 7th WSEAS International Conference on FLUID MECHANICS (FLUIDS'10), University of Cambridge, UK (pp. 60-67).
5. Zhang, M., & Chen, Q. G. (2021). Numerical model on the dynamic behavior of a prototype kaplan turbine runner. Mathematical Problems in Engineering, 2021(1),4421340.
6. Luo, Y., Wang, Z., Zeng, J., & Lin, J. (2010). Fatigue of piston rod caused by unsteady, unbalanced, unsynchronized blade torques in a Kaplan turbine. Engineering Failure Analysis, 17(1), 192-199.
7. Rodriguez, C. G., Egusquiza, E., Escaler, X., Liang, Q. W., & Avellan, F. (2006). Experimental investigation of added mass effects on a Francis turbine runner in still water. Journal of Fluids and Structures, 22 (5), 699-712.
8. Liang, Q. W., Rodríguez, C. G., Egusquiza, E., Escaler, X., & Avellan, F. (2005). Numerical modal analysis on a Francis turbine runner with finite element method considering added mass effect.
9. Liu, X., Luo, Y., & Wang, Z. (2016). A review on fatigue damage mechanism in hydro turbines. Renewable and Sustainable Energy Reviews, 54, 1-14.
10. Firth, D., Briggs, N., Geoghegan, B., & Dumbleton, E. (2004). Fitness for purpose assessment after an over pressure incident at Arapuni Power Station. *International journal of pressure vessels and piping*, *81*(6),471-479.
11. Vashishtha, G., Chauhan, S., Sehri, M., Hebda-Sobkowicz, J., Zimroz, R., Dumond, P., & Kumar, R. (2024). Advancing machine fault diagnosis: A detailed examination of convolutional neural networks. *Measurement Science and Technology*, *36*(2), 022001.
12. Dörfler, P., Sick, M., & Coutu, A. (2013). Flow-induced pulsation and vibration in hydroelectric machinery: engineer's guidebook for planning, design and trouble-shooting (pp. 163-168). London: Springer.
13. De Silva, C. W. (2006). Vibration: fundamentals and practice. CRC press.