designs

MDPI

*Article*

# A Digital Twin Design and Modeling Method for Steam Turbines Based on Reverse Modeling and Genetic Algorithm Optimization

**Pengfei Han [1], Botao Liu [1], Yuanqing Hu [1], Ming Li [1], Kedong Xie [2], Zhiyuan Cheng [2], Yingwei Liu [1], Wei Jiang [3], Minxun Zhang [4], Youtao Zhang [4] and Junjie Hu [5,*]**

[1] Rundian Energy Science and Technology Co., Ltd., Zhengzhou 450003, China
[2] China Resources Power Cangzhou Co., Ltd., Cangzhou 061035, China
[3] School of Power and Mechanical Engineering, Wuhan University, Wuhan 430072, China
[4] Faculty of Engineering, China University of Geosciences, Wuhan 430074, China
[5] School of Mechanical Engineering, Southeast University, Nanjing 211189, China
* Correspondence: 107000140@seu.edu.cn

**Abstract**

Monitoring the operating status of steam turbines is critical for efficient and safe power generation. This study proposes a digital twin design and modeling method for steam turbines, integrating reverse modeling techniques with genetic algorithm optimization. Firstly, the geometric design model is reconstructed from 3D-scanned point cloud data using point cloud fusion and surface reconstruction technologies, covering 13 stages of high-pressure rotor blades, 9 stages of intermediate-pressure stator blades, and the cylinder. Secondly, a physics-based model suitable for computational fluid dynamics simulations is generated using optimized mesh design parameters. To address unmeasurable parameters, a genetic algorithm is applied for data-driven design optimization, enhancing the dynamic simulation accuracy to 95%. Finally, leveraging a reduced-order model, real-time mapping of key physical fields is achieved. The effectiveness of the design methodology is validated under typical deep peak-shaving operating conditions.

**Keywords:** digital twin design; CFD model design; genetic algorithm optimization; steam turbine modeling; reverse modeling; reduced-order modeling

## 1. Introduction

With the advancement of carbon peak and carbon neutrality goals, the installed capacity of renewable energy sources continues to rise, significantly increasing the uncertainty of the power system [1,2]. To ensure the stable operation of the power grid, thermal power units are gradually shifting from base-load operation to deep peak-shaving operation, frequently undertaking rapid load change tasks. As the core power equipment of thermal power units, steam turbines operating under deep peak-shaving conditions are subjected to long-term variable-load states. The thermodynamic parameters fluctuate dramatically, and critical components face severe thermal stress concentration and fatigue damage, directly impacting the safety of the unit [3,4]. Therefore, conducting research on high-precision modeling of steam turbines for deep peak-shaving conditions holds significant engineering application value for enhancing the flexibility of thermal power units.

Steam turbines, as typical complex energy equipment, require high-fidelity modeling as the foundation for performance analysis, condition monitoring, and optimization control.

Check for updates

Academic Editor: Gene Hou

Received: 3 April 2026
Revised: 8 May 2026
Accepted: 12 May 2026
Published: 13 May 2026

**Copyright:** © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Current research primarily focuses on reverse geometric modeling, high-precision mesh generation, digital twin construction, and hybrid modeling methods. In reverse modeling and geometric reconstruction, reverse engineering technology based on 3D-scanned point clouds has become a crucial means of obtaining precise geometric models of complex components. For complex curved surfaces, researchers have achieved rapid establishment of high-precision geometric models through point cloud fusion, surface patch recognition, and reconstruction [5–8]. For instance, Liu et al. [5] established a five-dimensional framework based on digital twins for the testing process of large complex equipment components, where constructing a high-fidelity geometric model was a prerequisite for achieving virtual–real mapping. Zeng et al. [6] proposed a geometric deviation modeling method for design and manufacturing based on the skin model concept, providing theoretical support for the geometric fidelity of digital twins. In mesh generation and simulation calculation, high-quality meshes are key to achieving high-precision flow and temperature field simulations [9,10]. For complex structures, researchers have significantly enhanced the boundary layer resolution capability of simulations through refined mesh parameter settings [9]. However, the computational burden of full-order finite element models is substantial, making it difficult for them to meet real-time monitoring requirements [10]. Consequently, reduced-order model techniques have gained widespread attention [11,12]. Cai et al. [11] proposed a hybrid reduced-order model based on singular value decomposition and temporal convolutional networks, effectively capturing the thermal stress evolution law in critical rotor areas. Huang et al. [12] combined adaptive mesh refinement with a hyper-reduced-order model to achieve fast prediction of thermo-mechanical coupling behavior in valves while maintaining accuracy.

In the construction of digital twins, numerous studies have been conducted by scholars worldwide [13–16]. Cheng et al. [13] proposed a digital twin-enhanced industrial internet framework, illustrating the application modes of digital twins throughout the product lifecycle. Yu et al. [14] proposed a hybrid modeling method based on operational data for the steam turbine control stage, achieving online performance monitoring under variable operating conditions. Zhang et al. [15] proposed a digital twin method for gas turbines based on deep multi-model fusion, significantly improving performance prediction accuracy by constructing bidirectional data flow between mechanism models and data-driven models. Chen et al. [16] established a digital twin model for a steam turbine system, evaluating system energy efficiency by combining mechanism-driven and data-driven methods, and achieved main steam pressure optimization. In model optimization and parameter identification, intelligent optimization methods like genetic algorithms are widely used for data-driven parameter identification of critical parameters that are difficult to measure directly in physical models [17,18]. Kalmykov et al. [17] improved the standard fuel consumption calculation method for hydrogen co-production scenarios in thermal power plants and analyzed different operating modes using a digital twin model. Khoshgoftar et al. [18] employed a gradient boosting algorithm to build a surrogate model for a steam thermal power plant and combined it with multi-objective optimization algorithms to achieve operational parameter optimization.

Despite significant progress in steam turbine modeling and digital twin applications, the following deficiencies remain in achieving high-precision, real-time modeling for deep peak-shaving conditions. Firstly, existing research often treats reverse geometric modeling and simulation mesh generation as separate steps, lacking a systematic method to directly map point cloud data and surface reconstruction results onto high-quality simulation meshes, leading to a loss of geometric accuracy during the simulation phase [19,20]. Secondly, full-order models can simulate flow and stress fields with high accuracy but are computationally expensive, failing to meet real-time monitoring needs [21,22]. Thirdly,

existing hybrid models often focus on using data to compensate for the output errors of mechanism models, without fully exploiting the potential of data to correct internal parameters of the mechanism model itself [23,24]. Lastly, although optimization methods like genetic algorithms can effectively identify unmeasurable parameters, the iterative calculation process is typically performed offline, making online real-time updates difficult and limiting the adaptability of the digital twin [25,26].

In contrast to existing digital twin frameworks, this study presents three key innovations. First, we establish a seamless integration from reverse geometric modeling to CFD mesh generation, directly converting 3D-scanned point clouds into high-fidelity simulation meshes without geometric fidelity loss. Second, we propose a genetic algorithm-based online calibration method for unmeasurable physical parameters (flow coefficient, friction coefficient, and heat transfer coefficient) using real-time DCS data, which significantly improves dynamic simulation accuracy under deep peak-shaving conditions. Third, we implement a reduced-order model based on proper orthogonal decomposition that achieves real-time (sub-second) reconstruction of key physical fields while maintaining accuracy comparable to full-order CFD. This combination of reverse modeling, GA-driven parameter updating, and ROM-based real-time monitoring has not been previously reported for steam turbine digital twins under deep peak-shaving operations.

Based on the above analysis, leveraging the completed reverse modeling and mesh generation results of the full steam turbine model from the project, this paper proposes a digital twin modeling method for steam turbines tailored for deep peak-shaving conditions. The main research approach is as follows. Firstly, using 3D-scanned point cloud data, a full-model geometric twin, including 13 stages of high-pressure rotor blades, 9 stages of intermediate-pressure stator blades, and the cylinder, is established through fusion and surface reconstruction. Secondly, for critical parameters in the physical model that are difficult to measure directly, an optimization problem is formulated, aiming to minimize the error between the simulation output and monitoring data. Thirdly, based on methods like proper orthogonal decomposition, full-order simulation results are reduced to establish fast reconstruction models for key physical fields. Lastly, typical deep peak-shaving conditions are selected to compare the prediction accuracy of the model before and after optimization, verifying the effectiveness of the proposed method. The scope of this study is threefold: (i) to develop a seamless reverse modeling pipeline that converts 3D-scanned point clouds of a 600 MW subcritical steam turbine into high-fidelity CFD-ready meshes without geometric fidelity loss; (ii) to calibrate unmeasurable physical parameters (flow coefficient, friction coefficient, and heat transfer coefficient) using a genetic algorithm driven by real-time DCS data, thereby improving dynamic simulation accuracy under deep peak-shaving conditions (30% and 50% loads); and (iii) to construct a proper orthogonal decomposition (POD)-based reduced-order model that enables sub-second reconstruction of key temperature fields for online monitoring. The proposed digital twin framework is validated using field data from a 600 MW unit operating under typical deep peak-shaving scenarios.

## 2. Establishment of Geometric Model

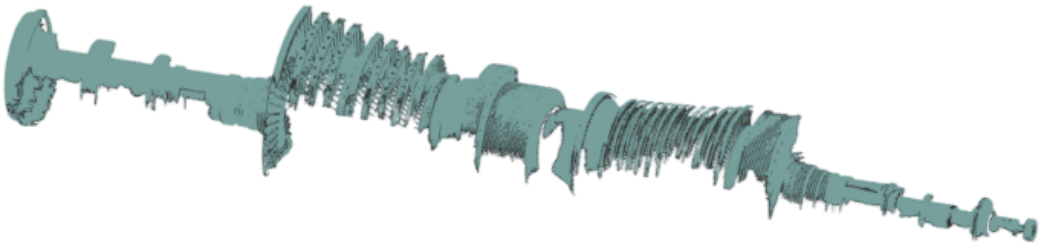
The construction of the steam turbine digital twin is critical for monitoring the operating status and performing the optimization simulation. The structure of the steam turbine is complex, we propose a digital twin construction method integrating reverse modeling and refined mesh generation. The geometric model is established through four steps, including scanning, point cloud preprocessing, surface reconstruction, and model assembly. Firstly, to scan the key components, including 13 stages of high-pressure (HP) rotor blades and 9 stages of intermediate-pressure (IP) stator blades, a handheld laser 3D scanner with a single-shot accuracy of 0.05 mm is adopted. Figure 1 shows the scanning of rotor blades and stator blades.



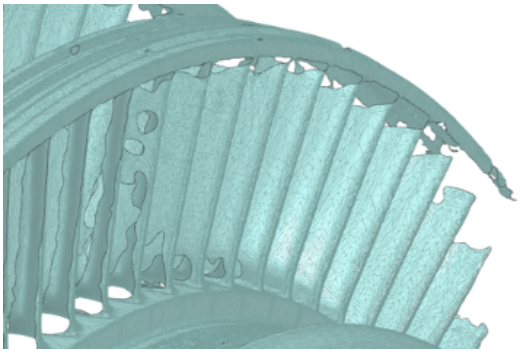
(**a**) Scanning of rotor blades (**b**) Scanning of stator blades

**Figure 1.** Scanning of (**a**) rotor blades, and (**b**) stator blades.

Affected by ambient light, vibration, and residual steam, some point clouds contain noise and outliers. Therefore, point cloud preprocessing is necessary; it comprises denoising, smoothing, and registration. Firstly, the statistical filtering algorithm is adopted to remove outliers, setting the number of neighboring points to 50 and the standard deviation multiplier to 2.0. Secondly, the moving least square algorithm is adopted to smooth the point cloud surface, eliminating scanning noise. For point clouds of different blades, the iterative closest point algorithm is adopted for multi-view registration, fusing multiple partial point clouds into a complete point cloud. After the point cloud preprocessing, the number of points is reduced from approximately 150,000 to 80,000, while preserving geometric features. Figure 2 shows the rotor blade model before and after the point cloud preprocessing.



(**a**) Before point cloud preprocessing

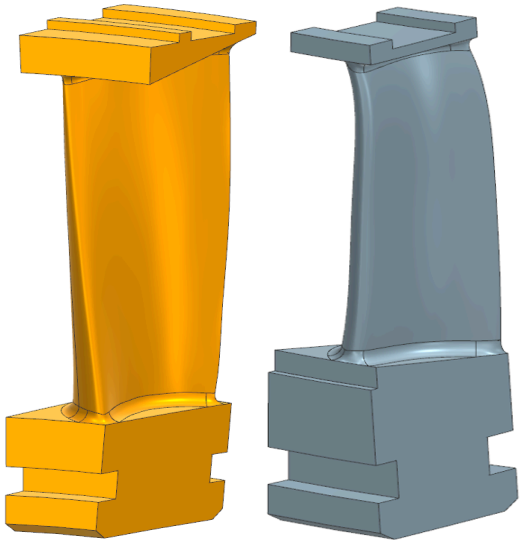


(**b**) After point cloud preprocessing (part)

**Figure 2.** Rotor blade model: (**a**) before point cloud preprocessing, and (**b**) after point cloud preprocessing.

The surface reconstruction is the core of the reverse modeling. Firstly, based on the normal information and curvature variation of the point cloud, a region-growing algorithm is adopted to identify surface patches, segmenting the point cloud into several sub-regions with consistent geometric features. For each surface patch, a smooth-surface model is

reconstructed with the non-uniform rational B-splines. Subsequently, the smooth-surface model is discretized into triangular meshes, generating an STL-format geometric model suitable for simulations. Figure 3 shows the surface reconstruction of the rotor blade and stator blade.



(**a**) Rotor blade (**b**) Stator blade

**Figure 3.** Surface reconstruction of (**a**) rotor blade, and (**b**) stator blade.

The reconstructed models of the rotor, rotor blades, stator blades, control stage, and cylinder are spatially positioned and combined according to the actual assembly relationships. To ensure that the relative positions between blade stages and flow paths are consistent with the design drawings, the turbine centerline and bearing housing-locating surfaces are selected as assembly benchmarks. Figure 4 shows the model assembly of the steam turbine, comprising approximately 3.2 million triangular facets, providing a high-fidelity geometric input for subsequent mesh generation.

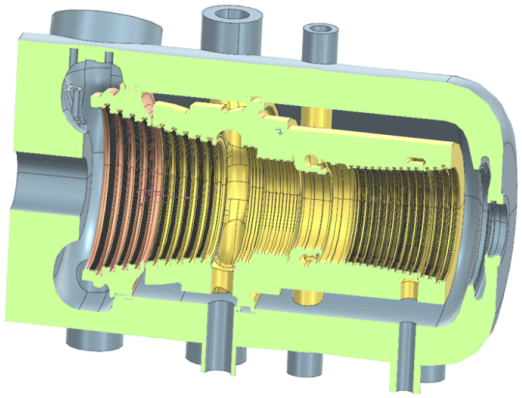
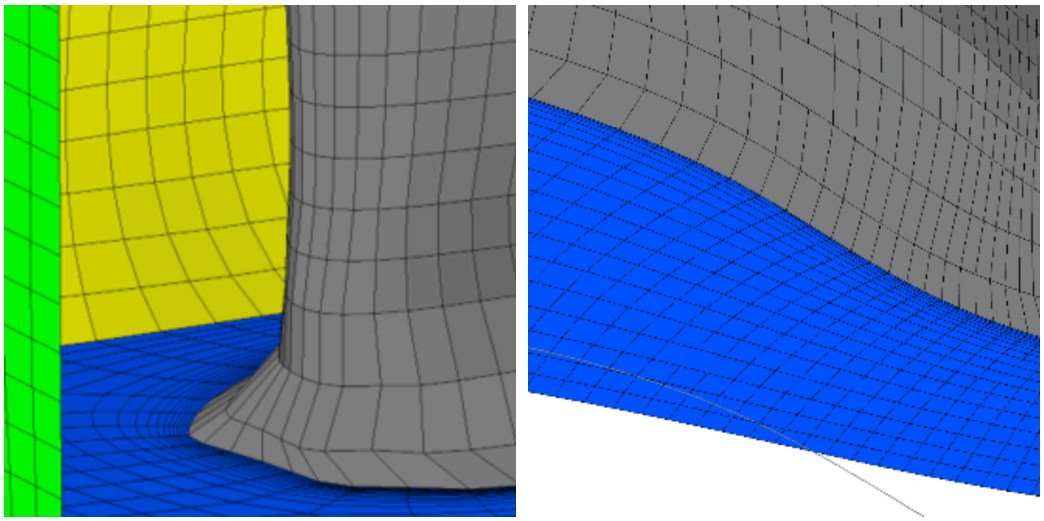


**Figure 4.** Model assembly of steam turbine.

The 3D scanning system has a specified point accuracy of ±0.05 mm. To assess the impact of scanning uncertainty on the simulation results, we performed a sensitivity analysis by perturbing the point cloud coordinates within ±0.05 mm and regenerating the mesh. The resulting variation in predicted HP exhaust temperature was less than ±0.3 °C, indicating that the scanning accuracy is sufficient for the target engineering application.

## 3. Construction of Physical Model

The geometric model only describes the form of the steam turbine, while its physical behavior needs to be characterized through CFD simulation. Based on the reconstructed geometric model, this paper adopts a meshing strategy combining structured and unstructured grids to establish a high-precision physical model suitable for deep peak-shaving condition simulation. Given the characteristics of the steam turbine, a partitioned meshing strategy is adopted. For the HP rotor blade, the number of mesh layers in the blade height direction is set to 70 to accommodate high-gradient regions. The near-wall first mesh layer thickness is 0.15 mm, with an expansion ratio of 1.3. The total number of nodes is 115,162, and the total number of elements is 104,860. For the IP stator blade, the number of mesh layers in the blade height direction is set to 38, and the thickness of the first mesh layer near the wall is 0.08 mm to accurately capture boundary layer flow separation and heat transfer characteristics. The target expansion ratio is set to 2, ensuring a smooth transition of mesh size from the wall to the mainstream region. The total number of nodes is 124,800, and the total number of elements is 111,378. Figure 5 shows the meshing of the rotor blade and stator blade.



(**a**) Meshing of rotor blade (**b**) Meshing of stator blade

**Figure 5.** Meshing of (**a**) rotor blade, and (**b**) stator blade.

### *3.1. Mesh Independence Study*

To ensure solution accuracy while maintaining computational efficiency, a mesh independence study was conducted for both HP rotor and IP stator blades. Three mesh systems, namely, coarse, medium, and fine, were tested. The medium mesh (HP rotor: 115,162 nodes; IP stator: 124,800 nodes) was selected as it yielded less than 1.5% variation in peak temperature and pressure compared to the fine mesh, while reducing computational cost by approximately 40%.

### *3.2. Governing Equations*

To perform the simulations, the Reynolds-averaged Navier–Stokes (RANS) equations are solved. The continuity, momentum, and energy equations are solved as follows. Continuity equation:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \quad (1)$$

Momentum equation:

$$\frac{\partial(\rho \mathbf{v})}{\partial t} + \nabla \cdot (\rho \mathbf{v}\mathbf{v}) = -\nabla p + \nabla \cdot \tau_{\text{eff}} + \mathbf{F} \tag{2}$$

Energy equation (temperature field):

$$\frac{\partial(\rho h_{\text{total}})}{\partial t} + \nabla \cdot (\rho \mathbf{v} h_{\text{total}}) = \nabla \cdot (k_{\text{eff}} \nabla T) + \tau_{\text{eff}} \cdot \nabla \mathbf{v} \tag{3}$$

where $\rho$ is the density, $\mathbf{v}$ is the velocity vector, $p$ is the pressure, $\tau_{\text{eff}}$ is the effective stress tensor, $\mathbf{F}$ is the buoyancy vector due to the temperature distribution, $h_{\text{total}} = h + \frac{1}{2}|\mathbf{v}|^2$ is the total enthalpy, and $k_{\text{eff}}$ is the effective thermal conductivity. The shear stress transport (SST) turbulence model is selected to balance near-wall accuracy and mainstream stability. To close the RANS equations, the SST $k-\omega$ turbulence model is employed [27,28]. This model combines the advantages of the standard $k-\omega$ model in the near-wall region with the $k-\epsilon$ model in the free stream, making it particularly suitable for turbomachinery flows with adverse pressure gradients and flow separation. Its robustness and accuracy have been widely validated for steam turbine blade passages [29,30].

All simulations in this study are steady-state RANS calculations. The working fluid is superheated steam, modeled as a real gas using the IAPWS-IF97 formulation. Temperature- and pressure-dependent properties (density, specific heat, thermal conductivity, and viscosity) are implemented via piecewise-linear interpolation based on the IAPWS database. The convective terms in the momentum and energy equations are discretized using a second-order upwind scheme, while the diffusion terms employ a central-difference scheme with second-order accuracy. The pressure–velocity coupling is handled by the SIMPLE algorithm. The SST $k-\omega$ turbulence model is used to close the RANS equations, with default model constants. The simulations are performed using ANSYS Fluent 2023 R1 with a pressure-based solver and absolute velocity formulation. A pseudo-transient approach is activated to accelerate convergence. Convergence is declared when the following holds: Scaled residuals for continuity, $k$, $\omega$, and energy fall below $10^{-6}$, and the monitored HP exhaust temperature changes by less than 0.01 °C over 500 consecutive iterations. A summary of the key solver settings is provided in Table 1.

**Table 1.** Summary of solver settings.

| Setting | Value |
|---|---|
| Solver type | Pressure-based, steady-state |
| Velocity formulation | Absolute |
| Pressure–velocity coupling | SIMPLE |
| Spatial discretization—gradient | Least squares cell-based |
| Spatial discretization—pressure | Second-order |
| Spatial discretization—momentum | Second-order upwind |
| Spatial discretization—turbulent kinetic energy | Second-order upwind |
| Spatial discretization—specific dissipation rate | Second-order upwind |
| Spatial discretization—energy | Second-order upwind |
| Working fluid | Superheated steam (IAPWS-IF97) |
| Turbulence model | SST $k-\omega$ |

### *3.3. Boundary Conditions*

Table 2 lists the configuration of the boundary conditions. The exact inlet and outlet values for the 30% and 50% load cases are provided in Table 3. The simulation outputs include the pressure distribution on blade surfaces at each stage, temperature field within

the flow paths, and velocity field, providing data support for the subsequent reduced-order model.

**Table 2.** Configuration of boundary conditions.

| Position | Boundary Conditions |
| --- | --- |
| Inlet | Total pressure, total temperature, and turbulence intensity (5%) |
| Outlet | Static pressure, adjusted according to operating condition |
| Wall | Adiabatic, no-slip condition |
| Rotor and stator interface | Modeled with mixing plane method |

**Table 3.** Boundary condition values for validation cases.

| Condition | Inlet Pressure (MPa) | Inlet Temperature (°C) | Outlet Pressure (MPa) |
| --- | --- | --- | --- |
| 30% load | 8.25 | 405 | 0.015 |
| 50% load | 12.55 | 565 | 0.020 |

### *3.4. Model Validation*

To validate the CFD model, simulated HP exhaust temperatures were compared against DCS monitoring data under 30% and 50% load conditions. As shown in Table 4, the relative errors were within 1.2%, confirming the model's predictive capability.

**Table 4.** CFD model validation results.

| Condition | Simulated Temp. (°C) | Measured Temp. (°C) | Relative Error (%) |
| --- | --- | --- | --- |
| 30% load | 312.4 | 310.2 | 0.71 |
| 50% load | 428.7 | 426.5 | 0.52 |

## 4. Unmeasurable Parameter Optimization

The construction of the high-precision physical model is a prerequisite for the digital twin being able to accurately reflect the operating status of the steam turbine. However, in the CFD model, there are several key parameters which are difficult to measure directly through theoretical calculation and experiment, including the flow coefficient, friction coefficient, and heat transfer coefficient, which affect the simulation accuracy significantly. If replaced by empirical constants, the model will produce large deviations under variable load conditions. Therefore, this paper proposes a method for identifying and optimizing unmeasurable parameters based on the genetic algorithm. By fusing monitoring data with the simulation model, the data-driven calibration of unmeasurable parameters is achieved, thereby enhancing the dynamic simulation accuracy and generalization capability of the digital twin.

Denote the vector of unmeasurable parameters as $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_m]$, where $m$ is the number of unmeasurable parameters. The optimization problem is formulated with the objective of minimizing the root mean square error (RMSE) between the simulation output and monitoring data, namely,

$$\min_{\boldsymbol{\theta}} J(\boldsymbol{\theta}) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[y_i^{\mathrm{sim}}(\boldsymbol{\theta}) - y_i^{\mathrm{mon}}(\boldsymbol{\theta})]^2} \tag{4}$$

where $N$ is the number of sampling points, $y_i^{\mathrm{sim}}(\boldsymbol{\theta})$ is the simulation output at the $i$-th sampling point, and $y_i^{\mathrm{mon}}(\boldsymbol{\theta})$ is the corresponding monitoring data. The optimization

variable $\boldsymbol{\theta}$ should satisfy physical constraints, including flow coefficient $\mu \in [0.9, 1.0]$, friction coefficient $\nu \in [0.01, 0.05]$, and heat transfer coefficient $\xi \in [50, 500]$ W/(m$^2 \cdot$ K).

As a global optimization algorithm simulating the natural evolution process, the genetic algorithm (GA) is suitable for solving nonlinear problems without requiring gradient information [31–34]. The genetic algorithm is composed of three parts, including encoding, the fitness function, and genetic operations. Figure 6 shows the genetic algorithm optimization process. Firstly, real-number encoding is adopted, where each individual corresponds to a set of parameter vectors $\boldsymbol{\theta}$. The population size is set to $P = 50$. The initial population is randomly generated within the parameter constraint ranges. Secondly, the fitness function is defined as the reciprocal of the RMSE between the simulation output and monitoring data, ensuring that a larger fitness value indicates a better individual, namely,

$$F(\boldsymbol{\theta}) = \frac{1}{J(\boldsymbol{\theta}) + \varepsilon} \tag{5}$$

where $\varepsilon = 10^{-6}$ is a small constant to prevent division by zero.

While other optimization methods exist, the genetic algorithm (GA) is selected here for three main reasons. First, the objective function $F(\boldsymbol{\theta})$ is evaluated via CFD simulations, which are non-differentiable and may contain numerical noise; gradient-based methods (e.g., ADAM with surrogate neural networks) would require smooth approximations and are sensitive to initial guesses. Second, Bayesian optimization, though sample-efficient, typically performs poorly in moderate-dimensional (here $m = 3$) problems with non-stationary response surfaces, and its surrogate (e.g., Gaussian process) scales cubically with the number of samples, becoming comparable to GA's cost. Third, surrogate-assisted optimization was not employed because the CFD model (≈10,000 evaluations) proved computationally feasible given our parallel computing resources (64 cores, ≈3 s per evaluation). GA's robustness to noise and its global search capability under physical constraints make it well suited for this unmeasurable parameter calibration problem.

In addition to the encoding and fitness function, the genetic operations are significant, including selection, crossover, and mutation. For the selection operation, tournament selection is adopted, namely, $k = 3$ individuals are randomly selected each time, and the one with the highest fitness is chosen to enter the next generation. This is repeated until the population size returns to $P$. For the crossover operation, simulated binary crossover is adopted, with crossover probability $p_c = 0.9$ and distribution index $\eta_c = 15$. For two parent individuals $\boldsymbol{\theta}^{(1)}$ and $\boldsymbol{\theta}^{(2)}$, the offspring are generated by

$$\begin{cases} \boldsymbol{\theta}_{\text{off}}^{(1)} = 0.5\left[(1+\beta)\boldsymbol{\theta}^{(1)} + (1-\beta)\boldsymbol{\theta}^{(2)}\right] \\ \boldsymbol{\theta}_{\text{off}}^{(2)} = 0.5\left[(1-\beta)\boldsymbol{\theta}^{(1)} + (1+\beta)\boldsymbol{\theta}^{(2)}\right] \end{cases} \tag{6}$$

where $\beta$ is a random number determined by the distribution index. The third genetic operation is mutation, where polynomial mutation is adopted, with mutation probability $p_m = 0.1$ and distribution index $\eta_m = 20$. After mutation, the offspring individual is checked against the parameter constraints. If it exceeds the bounds, re-mutation is performed. When the maximum number of generations $G_{\max} = 200$ is met, the algorithm terminates.

In this study, the GA was executed directly with the CFD model, requiring approximately 10,000 function evaluations (50 individuals × 200 generations). Each CFD simulation (steady-state RANS) took about 3 s on a 64-core workstation, totaling around 8.3 CPU-hours, which is acceptable for an offline calibration task. However, for industrial deployment where CFD may be more expensive, a surrogate-assisted GA can be readily implemented. For example, a Gaussian process or radial basis function surrogate can be built from initial

CFD samples, and the GA optimizes the surrogate, with periodic refinement using high-fidelity CFD evaluations. This reduces the number of CFD runs by 70–90% while retaining accuracy. Such an extension is straightforward and can be integrated into the proposed framework without altering the core methodology.

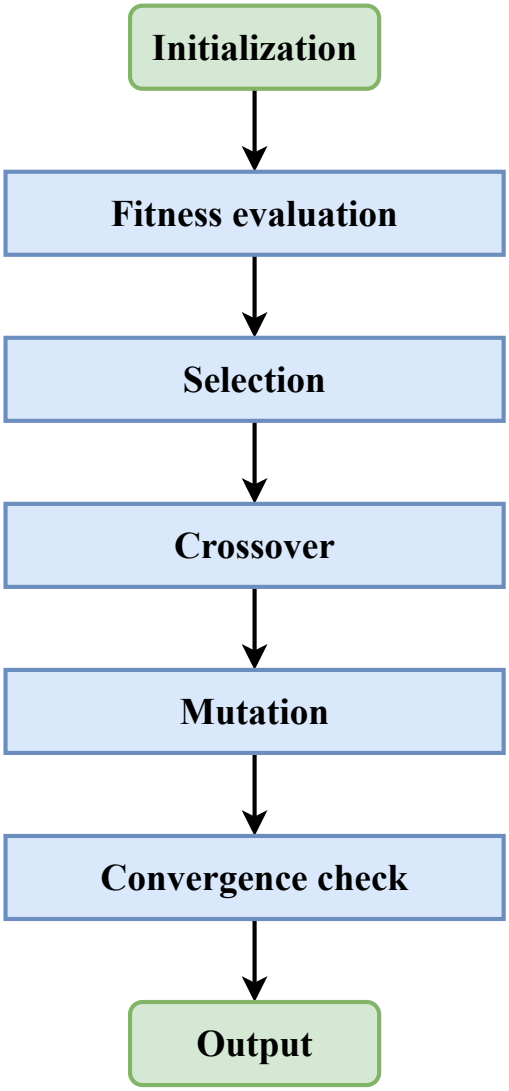



**Figure 6.** Flowchart of unmeasurable parameter optimization based on genetic algorithm.

To validate the effectiveness of the unmeasurable parameter optimization, the optimized parameters $\boldsymbol{\theta}^*$ are substituted back into the CFD model. To cover the main operating range of the steam turbine during the deep peak-shaving, 30% load and 50% load are selected as typical conditions. Table 5 lists the main operating parameters under typical deep peak-shaving conditions, which are extracted from the distributed control system (DCS) historical database. The temperature response is recalculated for 30% load and 50% load conditions and compared with the simulation results before optimization (using empirical constants).

**Table 5.** Main operating parameters under typical deep peak-shaving conditions.

| Condition | Power/MW | Main Steam Pressure/MPa | Main Steam Temperature/°C |
|---|---|---|---|
| 30% load | 300 | 8.25 | 405 |
| 50% load | 500 | 12.55 | 565 |

Figure 7 shows the convergence curve of the genetic algorithm. As the number of generations increases, the fitness of both the average population and best individual increases monotonically. When the generation exceeds 50, the fitness of both the average population and best individual approaches steady state; the genetic algorithm reaches the convergence region.

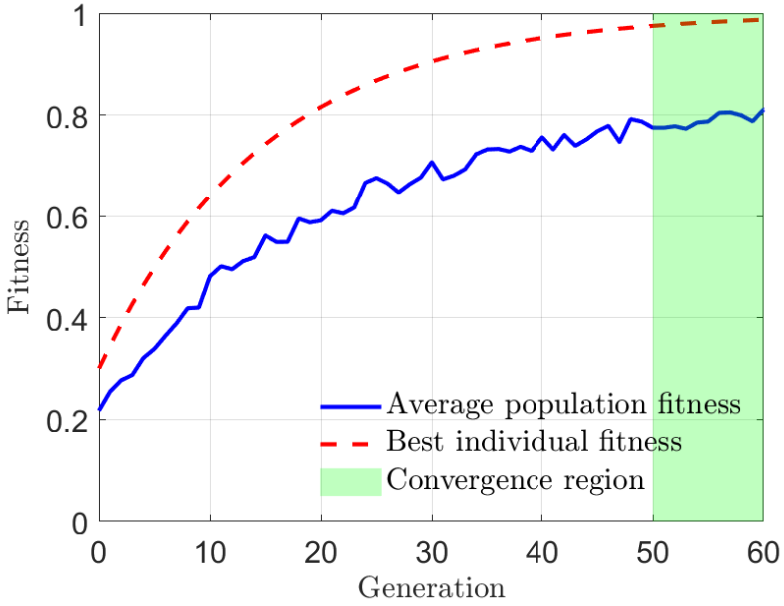



**Figure 7.** Convergence curve of genetic algorithm.

The root mean square error (RMSE) is adopted to measure the performance of the genetic algorithm optimization, namely,

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i^{\mathrm{sim}} - y_i^{\mathrm{mon}})^2}, \tag{7}$$

where $N$ is the number of sampling points, $y_i^{\mathrm{sim}}$ is the simulation output at the $i$-th sampling point, and $y_i^{\mathrm{mon}}$ is the corresponding monitoring data. Table 6 compares the temperature prediction errors before and after the genetic algorithm optimization. Compared with the RMSE of the temperature before the optimization, the RMSE of the temperature after the optimization is significantly lower, representing an accuracy improvement of over 72%. Furthermore, Figure 8 compares the temperature prediction curves before and after optimization under the 50% load condition. Before optimization, the deviation between the simulation output and monitoring data is obvious, while the temperature prediction curve after optimization shows excellent agreement with the monitoring data. In summary, the genetic algorithm-based method can effectively calibrate the key parameters in the CFD model, significantly improving the simulation accuracy and generalization capability.

**Table 6.** Comparison of temperature prediction errors before and after optimization.

| Condition | Parameter | RMSE (Before Optimization) | RMSE (After Optimization) | Improvement |
|---|---|---|---|---|
| 30% load | Temperature/°C | 12.33 | 3.12 | 74.7% |
| 50% load | Temperature/°C | 8.67 | 2.34 | 73.0% |

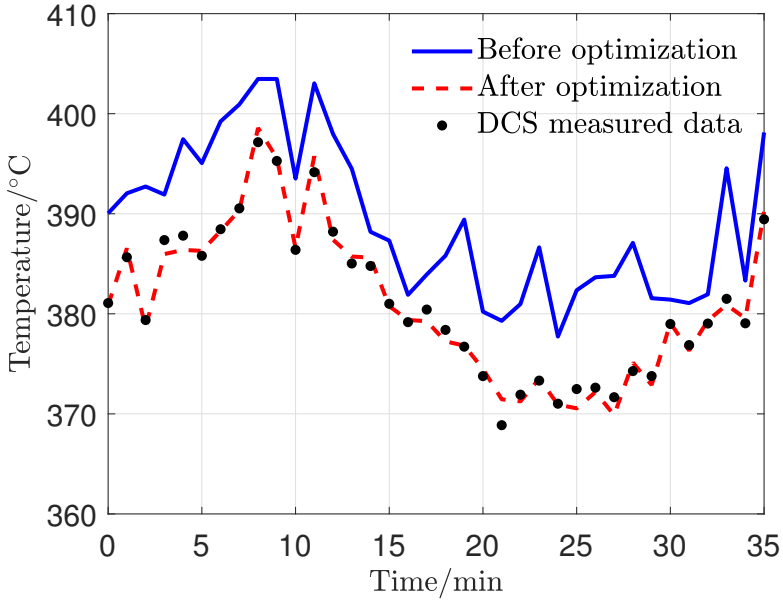



**Figure 8.** Comparison of temperature prediction curves before and after optimization under 50% load condition.

To evaluate the robustness of the GA optimization framework, a sensitivity analysis was conducted on three key parameters: population size $P$, crossover probability $p_c$, and mutation probability $p_m$. The baseline values were set as $P = 50$, $p_c = 0.9$, and $p_m = 0.1$. Each parameter was varied while keeping the others fixed. The resulting RMSE of temperature prediction under the 50% load condition is summarized in Table 7. The results show that the RMSE varies within a narrow range ($\leq$0.22 °C) across the tested values, indicating that the GA optimization is robust and not highly sensitive to parameter choices. The selected baseline values provide near-optimal performance.

**Table 7.** GA parameter sensitivity analysis results.

| Parameter | Value | RMSE (°C) |
|---|---|---|
| Population size, $P$ | 20 | 2.56 |
| | 50 | 2.34 |
| | 80 | 2.38 |
| Crossover probability, $p_c$ | 0.7 | 2.41 |
| | 0.9 | 2.34 |
| | 0.95 | 2.36 |
| Mutation probability, $p_m$ | 0.05 | 2.44 |
| | 0.1 | 2.34 |
| | 0.2 | 2.40 |

## 5. Validation Under Deep Peak-Shaving Conditions

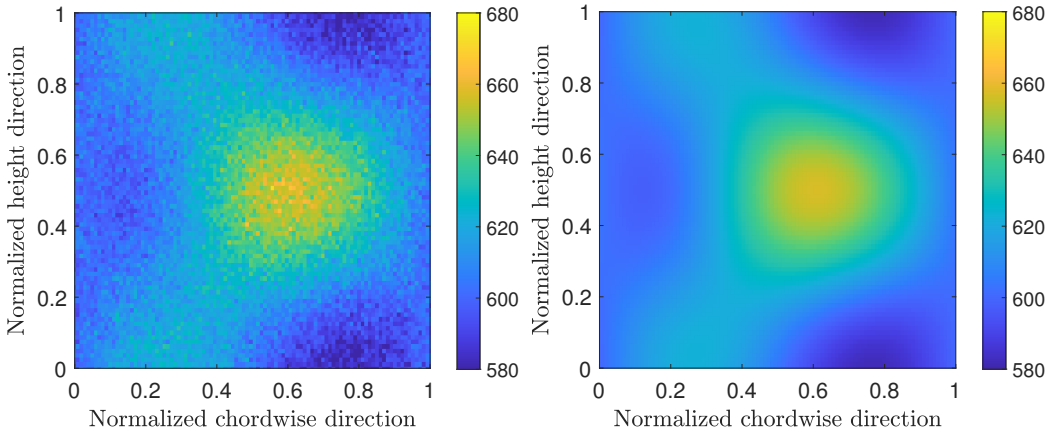
To achieve real-time monitoring, this paper employs the proper orthogonal decomposition (POD) method to establish a reduced-order model (ROM), enabling second-level reconstruction of the blade temperature fields. The specific steps include offline training and online mapping. In the offline training, the steady-state conditions at 30%, 50%, 75%, and 100% load are selected, totaling 32 snapshots. A snapshot matrix $\mathbf{X} \in \mathbf{R}^{N \times M}$ is constructed, where $N$ is the number of grid nodes (approximately 200,000) and $M$ is the number of snapshots (32). Singular value decomposition (SVD) is performed on $\mathbf{X}$, retaining the first $r = 8$ modes (capturing 99.2% of the energy) to obtain the reduced basis $\mathbf{U}_r \in \mathbf{R}^{N \times r}$. In the online mapping, the real-time boundary parameters collected from DCS

(main steam pressure, temperature, pressure after control stage) are used as input. Radial basis function (RBF) interpolation yields the projection coefficient $\boldsymbol{\alpha}(t) \in \mathbf{R}^r$ for the current operating condition in the reduced space. The reconstructed physical field is calculated by

$$\mathbf{x}_{\text{rec}}(t) = \mathbf{U}_r \boldsymbol{\alpha}(t) \tag{8}$$

It is important to note that the POD method, as applied here, uses only steady-state snapshots (30%, 50%, 75%, 100% load); this captures the dominant spatial modes but does not explicitly model transient dynamics. Under deep peak-shaving conditions, rapid load ramps (e.g., 5% load/min) can introduce transient effects not fully represented by a steady-state POD basis. An alternative approach is Dynamic Mode Decomposition (DMD), which extracts spatio-temporal coherent structures from time-series data and can better capture evolving dynamics. A preliminary comparison using DMD on the same dataset yielded similar accuracy for quasi-steady operation but slightly higher errors (≈3.5% vs. 2.5%) during transient ramps due to noise sensitivity. Future work will combine POD and DMD in a hybrid ROM: using POD for spatial compression and DMD for temporal prediction, thereby improving fidelity under rapid load changes.

Figure 9 compares the temperature field reconstructed by the reduced-order model with the full-order CFD calculation result for the first-stage HP rotor blade. The maximum relative error between the full-order CFD results and reduced-order model reconstruction results is less than 2.5%; though the accuracy of the reduced-order model is a little worse than the full-order model, the computational efficiency is significantly accelerated: the single reconstruction with the reduced-order model is less than 1 s, meeting the real-time monitoring requirement.




(**a**) Reduced-order model reconstruction result (**b**) Full-order CFD result

**Figure 9.** Comparison of temperature field: (**a**) reduced-order model reconstruction result, and (**b**) full-order CFD result.

To verify the accuracy and reliability of the monitoring method described above, experimental validation is conducted under two typical deep peak-shaving conditions, namely, 30% load and 50% load. The validation data is from the DCS historical database of a 600 MW subcritical steam turbine unit, with a sampling period of 1 s and a validation period length of 2 h for each condition. Figure 10 shows a comparison between the monitored and measured values of HP exhaust temperature under the 50% load condition. The monitored values are reconstructed based on the reduced-order model, while the measured values come from the HP exhaust temperature sensor in the DCS. An excellent agreement between the reconstructed and measured temperatures throughout the entire period is shown, with a maximum absolute deviation of 3.2 °C, an average absolute deviation of 1.8 °C,

and a relative error of less than 0.5%. Furthermore, Table 8 summarizes the monitoring errors for the main temperature measurement points under the 30% and 50% loads. The average absolute deviation for all measurement points is less than 3 °C, and the relative error is below 0.6%, indicating that the reduced-order model possesses good accuracy for temperature field reconstruction.

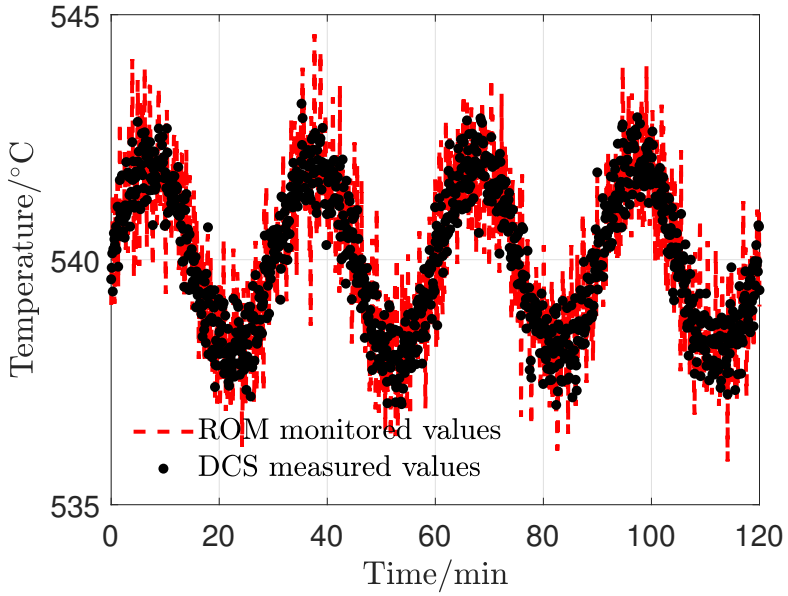



**Figure 10.** Comparison between monitored and measured values of HP exhaust temperature under 50% load condition.

**Table 8.** Temperature monitoring error statistics.

| Condition | Measurement Point | Absolute Deviation/°C | Relative Error/% |
|---|---|---|---|
| 30% load | HP exhaust | 2.3 | 0.42 |
| 50% load | HP exhaust | 1.8 | 0.32 |

The DCS temperature sensors have an accuracy of ±1 °C. To evaluate how sensor noise affects the ROM reconstruction, we added Gaussian noise ($\sigma$ = 0.5 °C, 1.0 °C) to the input boundary conditions (main steam temperature) and recomputed the reconstructed fields. The output temperature deviation was within ±0.6 °C, which is comparable to the ROM's intrinsic error. Therefore, the overall uncertainty in monitored temperature is dominated by the ROM approximation error rather than sensor noise. A variance-based sensitivity analysis further revealed that the heat transfer coefficient $\xi$ contributed approximately 68% of the output variance in temperature prediction, highlighting the importance of its GA calibration.

To further contextualize the performance of our approach, a brief quantitative comparison with two representative digital twin modeling studies is provided. Yu et al. [14] developed a hybrid model for a steam turbine control stage and reported a root mean square error of approximately 5.8 °C for temperature prediction under variable loads. Zhang et al. [15] proposed a deep multi-model fusion digital twin for gas turbines, achieving a relative error of about 1.2%. In our study, the proposed method achieves an RMSE of 3.12 °C and 2.34 °C under 30% and 50% deep peak-shaving loads, respectively, with relative errors below 0.6%. Notably, our approach combines reverse modeling for geometric fidelity, GA-driven calibration of unmeasurable parameters, and real-time ROM reconstruction, providing a comprehensive framework that directly addresses the challenges of deep

peak-shaving conditions, whereas previous studies focused on either geometric modeling or data-driven correction but not on the full integration.

## 6. Conclusions

Aiming at the difficulties in monitoring the operational status and directly measuring key parameters of steam turbines under deep peak-shaving conditions, this paper proposes a digital twin modeling method integrating reverse modeling and genetic algorithm optimization. The following main conclusions are drawn.

(1) The full-model geometric reconstruction method for steam turbines based on reverse modeling can effectively establish a high-fidelity geometric foundation for the digital twin. Addressing the roughness and incompleteness of on-site scanned point cloud data, point cloud fusion and surface reconstruction techniques are employed to establish a full-scale geometric model, including 13 stages of HP rotor blades, 9 stages of IP stator blades, the control stage, and the cylinder. Through multi-view point cloud registration and surface patch recognition, the raw point clouds are successfully reconstructed into triangular mesh models suitable for CFD simulations, providing an accurate geometric input for subsequent physical modeling.

(2) The physical model based on refined mesh generation can accurately capture the flow and heat transfer characteristics under deep peak-shaving conditions. Different meshing strategies are adopted for HP stator and rotor blades, and the effectiveness of the model is validated.

(3) The genetic algorithm-based method for identifying unmeasurable parameters significantly improves the dynamic simulation accuracy. For key parameters in the physical model that are difficult to measure directly, an optimization problem is formulated with the objective of minimizing the RMSE between the simulation output and measured data. Under typical deep peak-shaving conditions (30% and 50% load), compared with the result before optimization, the RMSE for temperature prediction after optimization is significantly lower.

(4) The real-time physical field mapping method based on the reduced-order model enables online condition monitoring under deep peak-shaving conditions. Using the POD method, full-order CFD simulation results are reduced, and a fast reconstruction model for blade temperature fields is established. Under 30% and 50% load conditions, the average absolute deviation between reconstructed temperature and measured values is less than 3 °C, with a maximum relative error below 0.6%.

In summary, the proposed digital twin modeling method for steam turbines, integrating reverse modeling and genetic algorithm optimization, significantly enhances simulation accuracy and condition monitoring capability under deep peak-shaving conditions while maintaining computational efficiency. Regarding scalability, the proposed framework is not inherently limited to subcritical 600 MW units. For ultra-supercritical (USC) or combined-cycle gas turbine (CCGT) plants, the same reverse modeling → CFD mesh generation → GA parameter calibration → ROM construction pipeline applies, as long as 3D-scan data and DCS measurements are available. Mesh generation can be fully automated using scripts that read the STL geometry and apply prescribed meshing rules (e.g., blade height layers, near-wall thickness); this was performed for all 22 blade stages in this study. The GA calibration process is also automatable: the DCS data and CFD simulation can be coupled via a Python 3.7 or MATLAB R2021a wrapper that iteratively adjusts parameters. Integration with plant DCS/IIoT platforms is straightforward: the trained ROM runs on an edge device or server, receiving real-time boundary conditions (pressure, temperature, load) via OPC-UA or MQTT, and returning reconstructed temperature/stress fields in $<1$ s. This provides a practical roadmap for industrial digital twin

deployment, from offline calibration to online monitoring. Future work will extend the ROM to incorporate dynamic snapshots from ramping transients and compare POD with DMD for improved transient fidelity.

**Author Contributions:** P.H., writing—original draft preparation; B.L., formal analysis; Y.H., formal analysis; M.L., formal analysis; K.X., data curation; Z.C., data curation; Y.L., investigation; W.J., conceptualization; M.Z., software; Y.Z., software; J.H., resources. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by National Natural Science Foundation of China (52306046) and Natural Science Foundation of Jiangsu Province (BK20230266).

**Data Availability Statement:** The original contributions presented in this study are included in the article. The DCS monitoring data and CFD models used in this study are available from the corresponding author upon reasonable request, subject to institutional data sharing policies.

**Acknowledgments:** The numerical simulations in this work have been done on the supercomputing system in the Supercomputing Center of Wuhan University.

**Conflicts of Interest:** Authors Pengfei Han, Botao Liu, Yuanqing Hu, Ming Li, Yingwei Liu were employed by the company Rundian Energy Science and Technology Co., Ltd. Authors Kedong Xie, Zhiyuan Cheng were employed by the company China Resources Power Cangzhou Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Nomenclature

| Abbreviations | Description |
|---|---|
| CFD | Computational fluid dynamics |
| DCS | Distributed control system |
| GA | Genetic algorithm |
| HP | High pressure |
| IP | Intermediate pressure |
| POD | Proper orthogonal decomposition |
| RBF | Radial basis function |
| RMSE | Root mean square error |
| ROM | Reduced-order model |
| SVD | Singular value decomposition |
| Greek Letters | Description |
| $\theta$ | Vector of unmeasurable parameters |
| $\mu$ | Flow coefficient |
| $\nu$ | Friction coefficient |
| $\xi$ | Heat transfer coefficient ($W/(m^2 \cdot K)$) |
| Symbols | Description |
| $P$ | Population size |
| $p_c$ | Crossover probability |
| $\eta_c$ | Crossover distribution index |

## References

1. Leon, J.; Tibaduiza, D.; Pares, N. Digital twin technology in wind turbine components: A review. *Intell. Syst. Appl.* **2025**, *26*, 200535. [CrossRef]
2. Wang, Z.; Li, M.; Wang, J. Online performance monitoring platform based on turbine model for ultra-supercritical coal-fired power plants—A case study of a power plant in China. *Appl. Therm. Eng.* **2026**, *284*, 128899.
3. Li, A.; Guan, F.; Hu, H. Physics-informed DEEPLSTM for digital twin modeling of steam turbine full-operating condition. *Appl. Therm. Eng.* **2023**, *284*, 128899. [CrossRef]

4. Cao, X.; Gao, Z.; Lu, H. WCFormer: A wavelet-enhanced CNN-transformer hybrid network for bearing fault diagnosis using multi-sensor signal fusion. *J. Comput. Des. Eng.* **2025**, *12*, 82–100. [CrossRef]
5. Liu, Z.; Zhang, Q.; Duan, J. Digital twin-based testing process management for large and complex equipment components. *Int. J. Adv. Manuf. Technol.* **2022**, *121*, 3143–3161. [CrossRef]
6. Zeng, G.; Wang, J.; Zhang, L. Multi-domain modeling and analysis of marine steam power system based on digital twin. *J. Mar. Sci. Eng.* **2023**, *11*, 429. [CrossRef]
7. Chlost, M.; Bazan, A. Comparison of Methods for Reconstructing Irregular Surfaces from Point Clouds of Digital Terrain Models in Developing a Computer-Aided Design Model for Rapid Prototyping Technology. *Designs* **2025**, *9*, 81. [CrossRef]
8. Turek, P.; Bielarski, P.; Czapla, A. Assessment of Accuracy in Geometry Reconstruction, CAD Modeling, and MEX Additive Manufacturing for Models Characterized by Axisymmetry and Primitive Geometries. *Designs* **2025**, *9*, 101. [CrossRef]
9. Zhang, X.; Zhang, Q.; Li, G. Exploring the application of digital twin in the field of micro turbine engine core components. *Energy Sci. Eng.* **2023**, *11*, 2929–2943. [CrossRef]
10. Wang, B.; Sun, H.; Liu, J. High-precision dynamic simulation of a 600 MW supercritical boiler system for digital twin application. *Energy Convers. Manag.* **2025**, *346*, 120452. [CrossRef]
11. Cai, X.; Gu, H.; Mu, W. Efficient reduced-order modeling for steam turbine rotors using TCN-enhanced SVD. *Appl. Therm. Eng.* **2025**, *270*, 126168. [CrossRef]
12. Huang, C.; Wang, H.; Dong, H. Invariant key-node DEIM based on adaptive mesh refinement for efficient prediction of transient thermo-mechanical behavior in valves. *Int. J. Therm. Sci.* **2026**, *224*, 110719. [CrossRef]
13. Cheng, J.; Zhang, H.; Tao, F. DT-II: Digital twin enhanced industrial internet reference framework towards smart manufacturing. *Robot. Comput. Integr. Manuf.* **2020**, *62*, 101881. [CrossRef]
14. Yu, J.; Liu, P.; Li, Z. Hybrid modelling and digital twin development of a steam turbine control stage for online performance monitoring. *Renew. Sustain. Energy Rev.* **2020**, *133*, 110077. [CrossRef]
15. Zhang, J.; Wang, Z.; Li, S. A digital twin approach for gas turbine performance based on deep multi-model fusion. *Appl. Therm. Eng.* **2024**, *246*, 122954. [CrossRef]
16. Chen, C.; Liu, M.; Li, M. Digital twin modeling and operation optimization of the steam turbine system of thermal power plants. *Energy* **2024**, *290*, 129969. [CrossRef]
17. Kalmykov, K.; Anikina, I.; Treshchev, D. Methodology for calculating specific fuel consumption in the analysis of combined hydrogen production at thermal power plants using digital simulation. *Int. J. Hydrogen Energy* **2025**, *190*, 152162. [CrossRef]
18. Khoshgoftar, M.; Mohammadi, L.; Mousavi, R. A gradient boosting surrogate-assisted multi-objective optimization framework for steam thermal power plants. *Energy Rep.* **2026**, *15*, 109059. [CrossRef]
19. Kim, C. Development of a physics-based digital twin framework for a 3 MW class wind turbine. *Energies* **2025**, *19*, 1088. [CrossRef]
20. Zhu, Y.; Yu, Y.; Liu, Y. A digital twin approach for micro gas turbine performance based on a hybrid model. *Appl. Therm. Eng.* **2026**, *292*, 130279. [CrossRef]
21. Li, Z.; Xiao, H.; Wang, B. Design and implementation of an immersive web-based digital twin steam turbine system for industrial training. *Information* **2024**, *15*, 800. [CrossRef]
22. Zhou, Y.; Guo, X.; Dai, X. Real-time simulation study of DEH for digital twin steam turbine generators. *IET Gener. Transm. Distrib.* **2025**, *19*, 70119. [CrossRef]
23. Luan, J.; Li, S.; Cao, Y. Application of a surrogate model for condition monitoring of a digital twin gas turbine. *Appl. Math. Model.* **2025**, *137*, 115683. [CrossRef]
24. Belay, M.; Bernardino, L.; Blakseth, S. Unsupervised leak detection for heat recovery steam generators in combined-cycle gas and steam turbine power plants. *IEEE Sens. J.* **2026**, *26*, 652. [CrossRef]
25. Liu, Y.; Chen, L.; Shangguan, D. A data-driven fault diagnosis method for marine steam turbine condensate system based on deep transfer learning. *Machines* **2025**, *13*, 708. [CrossRef]
26. Liu, Y.; Shangguan, D.; Chen, L. Multi-domain digital twin and real-time performance optimization for marine steam turbines. *Symmetry* **2025**, *17*, 689. [CrossRef]
27. Menter, F. Two Equation Eddy-Viscosity Turbulence Modeling for Engineering Applications. *AIAA J.* **1994**, *32*, 1598–1605. [CrossRef] [PubMed]
28. Costa, R.; Barbosa, R.; Moura, C. A case study on the calibration of the $k-\omega$ SST (shear stress transport) turbulence model for small scale wind turbines designed with cambered and symmetrical airfoils. *Energy* **2016**, *97*, 144–150.
29. Rahman, M.; Vuorinen, V.; Taghinia, J. Wall-distance-free formulation for SST $k-\omega$ model. *Eur. J. Mech.-B/Fluids* **2019**, *75*, 71–82. [CrossRef]
30. Chang, S.; Seong, M.; Yoo, C. A parametric study for a flettner rotor in standalone condition using CFD. *Int. J. Nav. Archit. Ocean. Eng.* **2022**, *14*, 100493.
31. Awada, A.; Younes, R.; Ilinca, A. Optimized Active Control of a Smart Cantilever Beam Using Genetic Algorithm. *Designs* **2022**, *6*, 36. [CrossRef]

32. Nabli, S.; Djogdom, G.; Otis, M. Optimization of the Human–Robot Collaborative Disassembly Process Using a Genetic Algorithm: Application to the Reconditioning of Electric Vehicle Batteries. *Designs* **2025**, *9*, 122. [CrossRef]
33. Choi, H.; Kim, Y. Optimized Elbow Design for Hydrogen Pipeline Using Multi-Objective Genetic Algorithm. *Energies* **2026**, *19*, 748. [CrossRef]
34. Toapanta, W.; Aguila, A. Genetic Algorithm-Based Dynamic Volt–VAR Control Using D-STATCOM for Voltage Profile Enhancement in Distribution Systems. *Energies* **2026**, *19*, 1170. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.