

采用 SHAP 的高压涡轮级高维设计空间数据挖掘

杨昭¹, 郭振东¹, 苏鹏飞², 汪祺能¹, 宋立明¹

(1. 西安交通大学能源与动力工程学院, 710049, 西安;

2. 东方汽轮机有限公司清洁高效透平动力装备全国重点实验室, 618000, 四川德阳)

摘要: 为厘清涡轮级高维设计空间中各变量对级性能的影响, 基于数据挖掘方法 SHAP 对 GE-E3 高压涡轮第一级的 93 个设计变量进行了知识挖掘。除常规数据挖掘工作中显著变量对总体性能的影响分析外, 发展了显著变量对涡轮级沿叶高方向性能分布影响的分析方法, 可视化表示了改善级总体性能的设计变量所影响的涡轮级的具体位置; 同时, 充分发挥 SHAP 局部解释的优势, 在设计空间中选取典型样本进行归因分析, 研究了各设计变量在样本性能指标变化中所发挥的作用。研究发现, 对于高压涡轮级, 影响级效率的显著变量包括有效出气角、静叶三维积叠参数、叶片吸力面前缘附近样条控制点等。基于 GE-E3 高压涡轮数据集进行数据挖掘归纳得到涡轮级设计准则: 减小动叶中间截面有效出气角、增大静叶中间截面有效出气角, 静叶三维积叠点周向顶部位置、周向中间位置向压力面偏移, 静叶中间截面、动叶中间截面吸力面前缘附近控制点均向叶片变薄方向移动。遵循设计准则得到的最终设计使级效率提高了 0.65%。

关键词: 高压涡轮级; 高维设计空间; 数据挖掘; SHAP 方法

中图分类号: TK471 **文献标志码:** A

DOI: 10.7652/xjtuxb202506015 **文章编号:** 0253-987X(2025)06-0144-11

Data Mining for High-Dimensional Design Space of High Pressure Turbine Stage Using SHAP

YANG Zhao¹, GUO Zhendong¹, SU Pengfei², WANG Qineng¹, SONG Liming¹

(1. School of Energy and Power Engineering, Xi'an Jiaotong University, Xi'an 710049, China;

2. State Key Laboratory of Clean and Efficient Turbomachinery Power Equipment, Dongfang Turbine Co., Ltd., Deyang, Sichuan 618000, China)

Abstract: In order to clarify the influence of various variables on the performance of turbine stage in high-dimensional design space, the knowledge mining of 93 design variables of GE-E3 high-pressure turbine first stage was carried out based on the data mining method SHAP. In addition to the analysis of the influence of significant variables on the overall performance in conventional data mining work, the analysis method of the influence of significant variables on the performance distribution along spanwise of turbine stages was developed, and the specific positions of turbine stage affected by the design variables to improve the overall performance of the stage were visualized. At the same time, SHAP's advantage in local explanations was fully leveraged, attribution analysis on typical samples in the design space was conducted, and the role of each design variable in the

收稿日期: 2024-11-22。 作者简介: 杨昭(2000—), 女, 硕士生; 宋立明(通信作者), 男, 教授, 博士生导师。

基金

项目: 国家科技重大专项资助项目(2019-II-0008-0028)。

网络出版时间: 2025-02-25

网络出版地址: <https://link.cnki.net/urlid/61.1069.T.20250225.0920.002>

change of performance indicators was studied. The results showed that for the high-pressure turbine stage, significant variables affecting stage efficiency included the effective output angle, three-dimensional stacking parameters of the stator, and the spline control points near the suction leading edge of blades. Based on GE-E3 high-pressure turbine dataset, data mining was conducted to obtain turbine stage design criteria: When the effective output angle of the middle section of the rotor is reduced and the effective output angle of the middle section of the stator is increased, the circumferential top and middle positions of the three-dimensional stack points of the stator will be shifted to the pressure surface, and the control points near the suction leading edge of the middle section of both the stator and rotor will move in the direction of blade thinning. The new design according to the design criteria has improved the stage efficiency by 0.65%.

Keywords: high-pressure turbine stage; high-dimensional design space; data mining; SHAP method

涡轮的气动设计一直是国内外学者研究的重点。随着现代航空发动机朝着高性能、低能耗方向的发展,通过对涡轮叶片气动形状的优化设计,降低叶栅损失、提高涡轮级效率极具挑战。同时,计算流体动力学(CFD)大大降低了涡轮叶片气动设计的成本,优化算法^[1-2]及深度学习技术^[3-4]在叶轮机械领域自动寻优、流场预测等方面的广泛应用为气动设计提供了有力的技术支持。

近年来,随着优化目标及设计变量的不断增加,样本数据也大幅增长^[5],多种数据挖掘技术被应用于涡轮叶片的气动设计之中。Li等^[6]回顾了 in 叶轮机械气动优化设计中使用的数据挖掘方法,包括总变差分析(ANOVA)、自组织映射(SOM)、本征正交分解(POD)等。利用这些数据挖掘方法可以完成显著变量识别,例如 ANOVA 通过计算各变量及其耦合效应对目标总方差的贡献来获得敏感性,同时还可以表征变量对总体性能指标的影响规律。Baert等^[7]利用 ANOVA 等数据挖掘方法于涡桨发动机叶片设计空间内进行知识挖掘,进而指导设计变量的取值,由此高效完成了包含 111 个设计变量的多变量叶片设计优化。Chiba等^[8]对某机翼数据集进行了总变差分析,发现变量之间的耦合效应是影响目标性能的重要组成部分。POD 将数据分解为一组正交基向量,从而提取其中的主要特征^[9]。段焰辉等^[10]在转子叶片优化研究中引入了 POD,从主要成分分析中归纳出几何知识,结果发现基于知识的设计明显优于参考设计。Oyama等^[11]使用 POD 从跨声速翼型的 Pareto 前沿解中获得设计信息,将优化后的翼型分为低阻力、高升阻比和高升力 3 类,并对每一类翼型给出了相应的设计指导。通过数据挖掘还可以识别多个设计目标之间的相关

性,SOM 通过非监督神经网络完成识别^[12]。Song等^[13]结合 SOM 及 ANOVA 等方法总结了设计变量对多个关键性能指标的影响,并通过对比 MBGO 优化算法得到的最优解造型验证了知识挖掘方法的准确性。Adjei等^[14]采用 SOM 获得了设计变量与两个风扇转子性能之间的相关性。除此之外,通过聚类等数据挖掘方法还可以归纳样本的共同特征。郭振东等^[15]联合 ANOVA 及平行坐标系等方法,提出了将来自不同数据挖掘方法的信息进行融合的设计空间知识挖掘框架,并由此给出了小展弦比叶栅端壁/叶身联合成型设计空间优化解特征。Qin等^[16]提出了一种元模型驱动数据挖掘方法(MDDM),结合统计性描述、SOM、 k -d 树和 Sobol 指数等数据挖掘技术进行统计、相关性、聚类 and 敏感性分析,最终完成了知识提取,并通过总结设计准则给定参数,提高了压气机的工作性能。Baert等^[17]将基于代理模型的动态优化与高效的数据挖掘方法相结合,应用于低压涡轮高维设计空间的优化指导,实现了近 0.5% 的整体效率提升。

除了传统的数据挖掘方法,为进一步厘清输入设计变量与输出性能指标之间的关系,近年来有研究者将基于合作博弈论的 SHAP 方法应用于叶轮机械领域,而该数据挖掘方法也展现出了对于单个样本局部解释方面的优势^[18-20]。Shen等^[21]以某航天器为例进行了多目标优化,利用 LightGBM 构建代理模型,利用 SHAP 解释器来量化变量的贡献,同时提出了 shap-shape map 方法,使 SHAP 值在几何表面上可见。Zhang等^[22]开发了一种基于元模型解释的数据挖掘框架,该框架采用 XGBoost 元模型结合 SHAP 解释器,应用于船用燃气轮机 3.5 级跨声速轴流压气机增效增稳设计准则的挖掘,结果

表明在部分转速下,压气机的失速余量在调整叶片倾斜和扭转后扩大到了 5.87%,附加调整可变动子后进一步扩大到 23.31%,设计转速下峰值绝热效率提高了 0.06%。Palar 等^[23-24]利用 SHAP 搭建起了可解释代理模型的框架,并与活动子空间和 Sobol 指数进行了比较,研究结果表明 SHAP 提供了更有价值的信息,特别是关于设计变量对目标的影响方向,同时还展示了 SHAP 如何有效地解开了多个目标之间的复杂权衡。Wang 等^[25]应用 SHAP 方法解释了神经网络行为,提供了单个几何变量的敏感性和多个几何变量之间的耦合效应等信息。

综上所述可知,目前公开文献中将数据挖掘新方法 SHAP 应用于涡轮级高维设计空间的研究较少,同时几乎所有分析都聚焦于研究对象的整体标量性能而非沿某一方向的矢量性能分布。鉴于此,本文基于数据挖掘方法 SHAP 对 GE-E3 高压涡轮第一级的 93 个设计变量进行了知识挖掘,同时发展了针对涡轮级沿叶高方向性能曲线的数据挖掘方法,还发挥 SHAP 局部解释的优势进行了典型样本的归因分析,构建起了基于 SHAP 的数据挖掘框架,以全面厘清高压涡轮级高维设计空间各变量及目标性能之间的交互作用关系。

1 基于 SHAP 的数据挖掘框架

1.1 XGBoost 驱动的 SHAP 方法

SHAP 可以通过对实例的特征进行归因从而解释设计变量对最终预测结果的影响,其整体工作流程如图 1 所示。基于所研究问题生成的数据集,通过模型构建进行数据预测,同时计算 SHAP 值对模型进行解释。本文采用 XGBoost 元模型与 SHAP 解释器相结合的方式来分析黑盒设计空间中各变量对目标性能的影响。

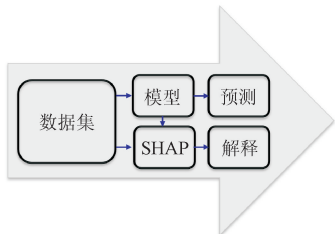


图 1 SHAP 工作流程

Fig. 1 SHAP workflow

1.1.1 XGBoost 元模型

XGBoost 通过执行顺序决策树的集成来提高模型的性能,该算法可以不断迭代以降低模型的预

测误差,并且在每一步都引入了正则化项来防止过拟合,从而提高了模型的泛化能力。其根本原理是通过构建多个决策树的组合从而形成一个强大的集成模型,以此预测所研究问题的输出值,公式表达为

$$\hat{y}_i^{(t)} = \sum_{k=1}^t f_k(x_i) = \hat{y}_i^{(t-1)} + f_t(x_i) \quad (1)$$

式中: k 为迭代数; $\hat{y}_i^{(t)}$ 为第 i 个样本第 t 轮的模型预测值; $\hat{y}_i^{(t-1)}$ 为第 i 个样本所保留的前 $t-1$ 轮的模型预测值; $f_t(x_i)$ 为第 i 个样本第 t 轮迭代所构建的新树模型对上次预测残差的拟合值。

在迭代过程中, XGBoost 的优化目标由损失函数和正则化项两部分组成,表示为

$$O_{bj} = \sum_i l(y_i, \hat{y}_i) + \Omega(f_k) \quad (2)$$

式中: $l(y_i, \hat{y}_i)$ 为损失函数,用于衡量真实值与预测值之间的差距,因此可以表示模型的准确度; $\Omega(f_k)$ 为正则化项,用来表示模型的复杂度。通过最小化目标函数 O_{bj} , 可以确定最新引入的树模型。

1.1.2 SHAP 解释器

SHAP 通过考虑所有可能的特征组合及其对预测结果的贡献,为每个特征分配“重要性”值,即 SHAP 值。假设所研究问题的第 i 个样本为 x_i , 第 i 个样本的第 j 个特征为 x_{ij} 、模型对该样本的预测值为 y_i 、整个模型的基准值(通常是所有样本的目标变量均值)为 y_{base} , 那么 SHAP 值服从

$$y_i = y_{base} + f(x_{i1}) + f(x_{i2}) + \cdots + f(x_{ik}) \quad (3)$$

式中: $f(x_{ij})$ 为 x_{ij} 的 SHAP 值。直观上看, $f(x_{i1})$ 就是第 i 个样本中第 1 个特征对最终预测值 y_i 的贡献值。当 $f(x_{i1}) > 0$ 时,说明该特征提升了预测值,即发挥正向作用;反之,说明该特征使得预测值降低,有反作用。

SHAP 是基于样本的局部解释,其最大优势是通过 SHAP 值能反映出每一个样本中每一个特征的影响力。求解 SHAP 值即计算每个特征的边际贡献,简单来讲就是某特征加入时的贡献和该特征加入之前的贡献之差,具体计算方法如下

$$\varphi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(M-|S|-1)!}{M!} [f_x(S \cup \{i\}) - f_x(S)] \quad (4)$$

式中: $N = \{1, 2, \dots, M\}$ 代表数据集中特征变量下标; M 是特征变量的总数; S 是集合 $\{1, 2, \dots, i-1, i+1, \dots, M\}$ 的子集; $|S|$ 是 S 中的元素的总数; $f_x(S \cup \{i\})$ 表示当只有 $S \cup \{i\}$ 中的特征时模型的预测值, $f_x(S)$ 表示只有 S 中的特征时模型的预测值,二者相

减即为第 i 个特征变量在子集 S 下的边际贡献。

1.2 基于 SHAP 的数据挖掘框架

基于数据挖掘方法 SHAP,建立起了本文的数据挖掘框架,如图 2 所示。在完成设计空间中的样本数据收集工作后,利用 SHAP 展开分析。首先,通过基于 SHAP 值构造的饼图进行显著变量识别;

然后,通过对涡轮级总体性能(0 维数据)的响应归纳显著变量对总体性能的影响规律;之后,通过对涡轮级沿叶高方向性能(1 维数据)的响应,可视化表示设计变量通过影响涡轮级哪一位置进而改善了总体性能;最后,发挥 SHAP 局部解释的优势,在设计空间中选取典型样本进行归因分析。

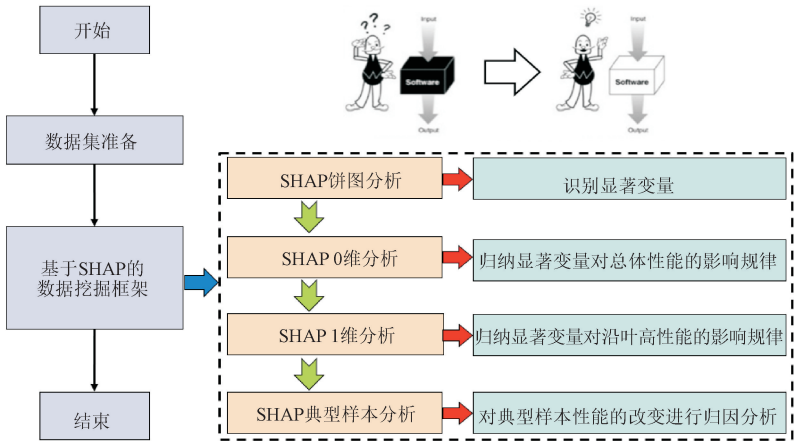


图 2 基于 SHAP 的数据挖掘框架
Fig. 2 SHAP based data mining framework

2 研究对象及设计空间

2.1 GE-E3 高压涡轮第一级数值预测模型

图 3 给出了本文研究对象 GE-E3 高压涡轮第一级的几何形状(即为参考设计),包括静叶、动叶、轮毂面及机匣面 4 部分。其中,静叶叶片数为 46,动叶叶片数为 76,动叶叶顶间隙为 0.3 mm。

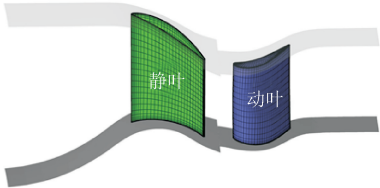


图 3 GE-E3 高压涡轮第一级几何示意图
Fig. 3 Geometric diagram of first stage of GE-E3 high pressure turbine

本文应用课题组成熟的全景式高压涡轮流场预测方法代替耗时的 CFD 计算,该方法利用基于 Transformer 的神经算子对涡轮级计算中构成 N-S 方程的 6 个物理量(温度 t 、压力 p 、密度 ρ 、速度 v_x 、 v_y 、 v_z)对应的基础物理场进行预测,并在此基础上得到其他衍生物理场及性能参数的预测结果。该方法在 GE-E3 高压涡轮第一级上的预测精度已得到验证^[26],本文不再详细阐述。涡轮级运行工况取进

口总温 719 K,进口总压 344.74 kPa,出口静压 139.473 kPa,动叶旋转速度 $8\,279\text{ r}\cdot\text{min}^{-1}$ 。以总效率(后文简写为效率)作为衡量涡轮级气动性能的关键指标,其表达式为

$$\eta = \frac{1 - T_2^*/T_0^*}{1 - (p_2^*/p_0^*)^{(\kappa-1)/\kappa}} \tag{5}$$

式中:下标 0、2 分别指代涡轮级进口和出口; T^* 、 p^* 分别表示总温、总压; κ 为比热容比。考虑到涡轮级设计时不能偏离设计工况,因此除效率外同时关注涡轮级质量流量的变化。

2.2 几何参数法及设计空间

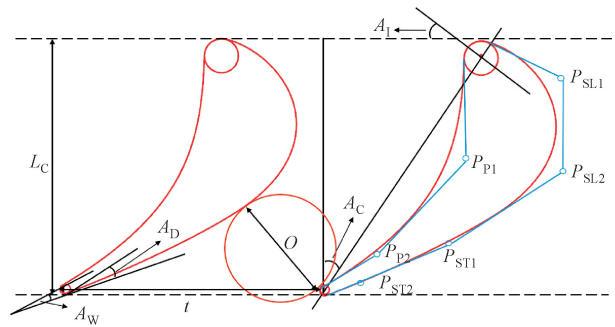
叶片造型使用几何参数法,对于静叶叶片,用 45 个控制参数来生成叶根、中间和叶顶 3 个截面型线,每个截面 15 个参数,而后使用 3 个参数调整叶片的三维积叠形状,参数具体定义见表 1。各控制参数如图 4 所示,其中图 4(a)对截面参数进行了标注,图 4(b)为三维积叠参数的变化方式,径向积叠线在叶根处固定不动,在叶中、叶顶设置控制点,使叶片积叠线在叶中位置可以沿周向移动,在叶顶位置可以同时进行周向、轴向平移,以此完成叶片的三维弯、掠造型。对于动叶,考虑其旋转过程中叶片的结构强度问题,因而只通过 45 个截面型线参数进行调整而不改变三维积叠形状。基于这一方法,本文使用 93 个参数精确描述两列叶片的几何形状。为

了便于区分,使用 S_1 、 S_2 、 S_3 分别表示静叶叶根截面、中间截面、叶顶截面,使用 R_1 、 R_2 、 R_3 分别表示动叶叶根截面、中间截面、叶顶截面,后文中以下标的形式出现在各设计变量名中。使用 D_1 、 D_2 、 D_3 分别表示静叶三维积叠参数中的周向中间平移量、周向顶部平移量、轴向顶部平移量。

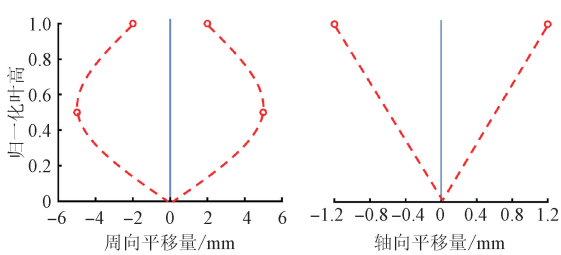
在设计空间内给定各变量的取值范围,随后采用拉丁超立方采样(LHS)方法生成样本,以叶片原型为基准将设计变量的值归一化到区间(0,1)范围内,0.5 表示基准值。

表 1 几何参数法中的具体参数定义
Table 1 Specific parameter definition of geometric parameter method

序号	参数定义	符号
1	前缘半径;进口圆半径	R_L
2	轴向弦长;轴向最大距离	L_C
3	圆心连线角;圆心连线与轴向夹角	A_C
4	进口几何角;进口方向与额线夹角	A_I
5	前缘修正系数;对吸力面前缘进行修正的参数	C_L
6	有效出气角; $\arcsin(o/t)$,其中 o 为喉部圆直径, t 为栅距	A_E
7	出口偏转角;出口圆切线与吸力面尾缘切线夹角	A_D
8	关联系数;控制喉部出口无遮盖段曲线	C_C
9	出口楔角;吸力面尾缘切线与压力面尾缘切线夹角	A_W
10	吸力面第一段曲线 Bezier 样条控制点 1	P_{SL1}
11	吸力面第一段曲线 Bezier 样条控制点 2	P_{SL2}
12	吸力面第二段曲线 Bezier 样条控制点 1	P_{ST1}
13	吸力面第二段曲线 Bezier 样条控制点 2	P_{ST2}
14	压力面曲线 Bezier 样条控制点 1	P_{P1}
15	压力面曲线 Bezier 样条控制点 2	P_{P2}
16	三维积叠:静叶周向中间平移量、周向顶部平移量、轴向顶部平移量	$D_1 \sim D_3$



(a)截面参数



(b)三维积叠参数

图 4 控制参数示意图

Fig. 4 Control parameters diagram

3 高压涡轮级数据挖掘及机理分析

3.1 数据集准备

基于 LHS 技术,对上文所述的参数化设计空间进行采样,获得 1 000 个样本进行初步探索。图 5 利用箱线图对输出数据效率及质量流量的分布进行描述。可以看出,二者各自的均值和中值都非常接近,其分布符合正态分布。对于效率,近 2/3 样本的效率值大于参考设计的基准值,这说明在设计空间内效率具有较大的提升能力,采样得到的最大效率值为 91.263%。对于流量,大部分样本的流量值大于参考设计的基准值,即设计空间内流量易朝着增大的方向发展,最大最小流量差为 1.111 kg/s,变化幅度达参考设计的 9.6%,变化范围过大。

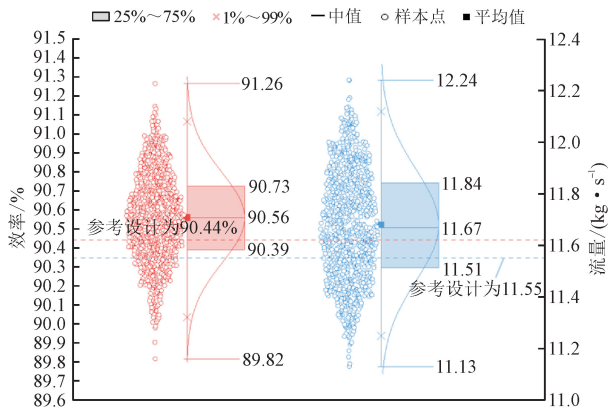


图 5 关键性能指标的箱线图

Fig. 5 Box plot of key performance indicators

基于以上统计性分析结果,为保证在涡轮级设计过程中级流量不发生大幅度增加、涡轮级不偏离设计工况,增加约束条件筛选保留流量变化不超过 $\pm 1\%$ 的样本,裁剪构成样本数为 352 的数据集进行后续分析。

3.2 显著变量分析

基于 352 个设计空间样本,进行敏感性分析。

利用 XGBoost 完成模型建立,模型精度如图 6 所示。可以看出,效率预测值与真实值所组成的坐标点紧密地分布在 45°线上,表明所构建的效率近似模型具有较高的预测精度。

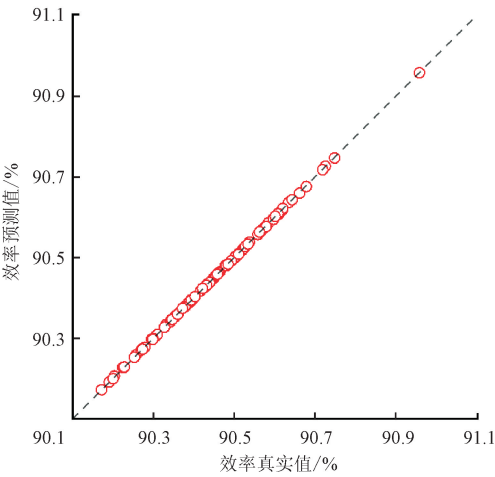


图 6 效率近似模型精度校验

Fig. 6 Precision check of approximate model of efficiency

利用 SHAP 解释器计算 SHAP 值,进行显著变量识别,并将属于同种几何参数的设计变量进行合并排序,保留贡献度大于 1% 的变量标签,结果如图 7 所示。可以看出,对于级效率而言,贡献度排名前 3 位的影响参数分别为叶片有效出气角(A_E)、静叶三维积叠参数($D_1 \sim D_3$)、叶片吸力面前缘附近样条控制点(P_{SL2}),三者之和占比近 45%。除此之外,在动、静叶共计 6 个截面中,动叶中间截面(R_2)、静叶中间截面(S_2)对性能指标的影响程度要高于其他截面。

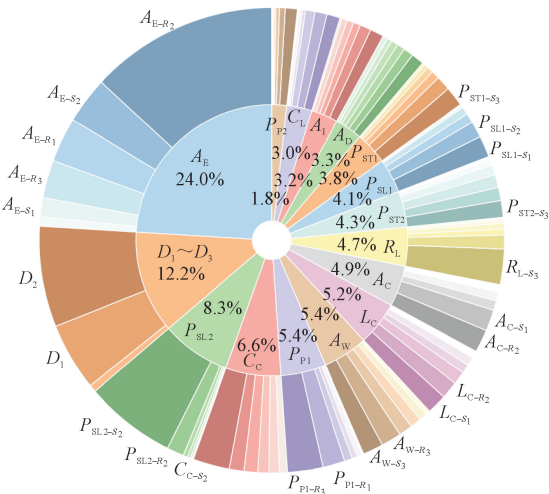


图 7 SHAP 饼图显著变量识别

Fig. 7 SHAP pie chart significant variable recognition

完成显著变量识别后,利用基于 SHAP 的 0 维、1 维分析方法对排名前 3 位的几何参数做进一步分析,探索各变量对涡轮级总体性能及沿叶高方向性能分布的影响规律。

3.2.1 有效出气角

有效出气角作为对级效率影响最大的几何参数,主要通过改变通道喉部构造发挥作用。图 8 给出了动叶中间截面有效出气角(A_{E-R_2})、静叶中间截面有效出气角(A_{E-S_2})对涡轮级效率的影响方向。SHAP 值为正时表示性能指标受输入参数影响朝正向变化,SHAP 值为负时表示性能指标受输入参数影响朝负向变化。由图可知:动叶中间截面有效出气角越小,对输出性能呈正向作用,即效率越高;静叶中间截面有效出气角越小,对输出性能则呈负向作用,即效率越低。换言之,为了提高涡轮级效率,应该减小动叶中间截面有效出气角、增大静叶中间截面有效出气角。

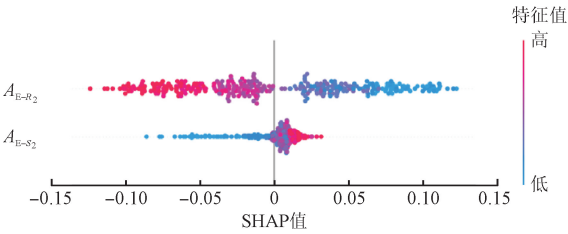


图 8 有效出气角对涡轮级总体性能的影响

Fig. 8 Effect of A_E on overall performance of turbine stage

以往的数据挖掘往往停留在涡轮级总体性能(0 维数据)的分析,本文在此基础上发展了对于涡轮级沿叶高方向性能分布(1 维数据)的分析方法,可以更加直观地展示设计变量是通过影响涡轮级哪一位位置进而改善了其总体性能。图 9(a)为动叶中间截面有效出气角对沿叶高方向效率分布的影响,其中纵轴表示无量纲化的叶高,横轴 SHAP 值表示对效率的影响方向。凭借 1 维 SHAP 可视化分析能够清楚地看到 A_{E-R_2} 减小,中间叶高大范围内的效率提高,总体性能得到改善,然而 A_{E-R_2} 的减小同时也会导致上端壁区域范围内效率的降低。对于动叶而言,有效出气角的减小增大了动叶反动度,可以减小型面损失,使涡轮级效率提升。图 9(b) 为静叶中间截面有效出气角对沿叶高方向效率分布的影响,静叶对涡轮级流量影响较大, A_{E-S_2} 增大可以在约束允许范围内增大级流量,使沿叶高大部分范围内效率有所提升,但随着流量的增加端区损失也会增强,上下端壁区域效率略有降低。

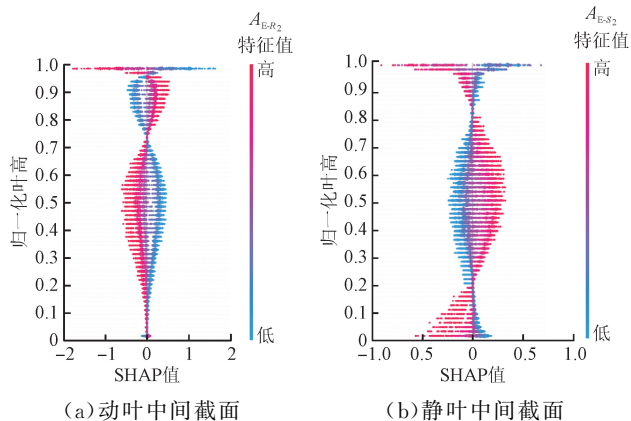
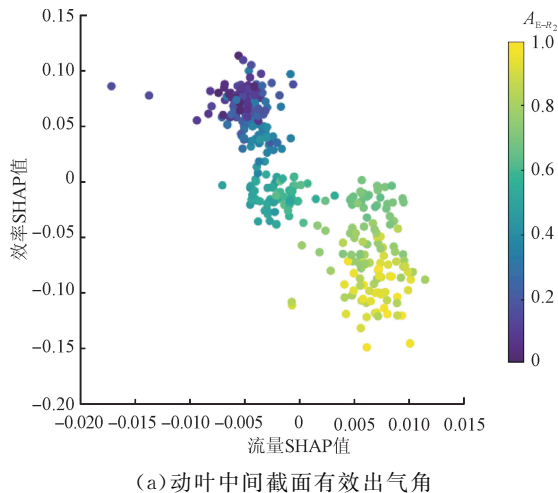


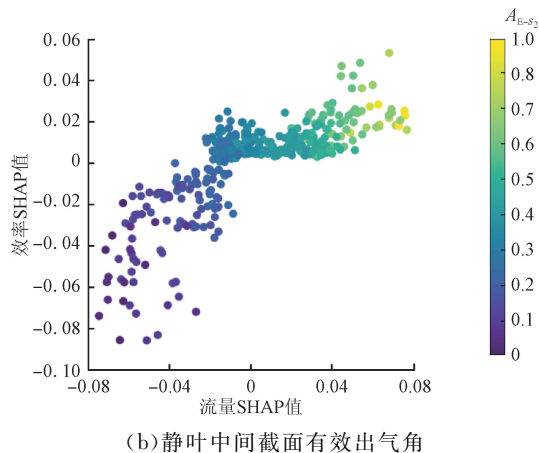
图 9 有效出气角对涡轮沿叶高方向性能的影响

Fig. 9 Effect of A_E on turbine performance along spanwise

由于有效出气角会改变喉部造型进而影响涡轮的通流能力,因此进一步通过 SHAP 散点图同时展示改变有效出气角对涡轮级效率及流量的影响方向,探索两个性能指标之间的相关性,如图 10 所示。



(a) 动叶中间截面有效出气角



(b) 静叶中间截面有效出气角

图 10 涡轮级效率及流量相关性

Fig. 10 Turbine stage efficiency and mass-flow correlation

图 10(a)为动叶中间截面有效出气角的影响趋势。可以看出,随着 A_{E-R_2} 的减小,涡轮级效率增大

而流量降低。图 10(b) 为静叶中间截面有效出气角的影响趋势。可以看出,随着 A_{E-S_2} 的增大,涡轮级效率增大,流量增大。在进行涡轮级设计时,往往希望在设计工况不改变,即流量点不发生偏移的情况下尽可能提高效率,而在动叶中间截面有效出气角减小、静叶中间截面有效出气角增大的综合作用下,可以在提升效率的同时控制流量的变化幅度,是一种“1+1>2”的造型方式。

3.2.2 静叶三维积叠参数

静叶三维积叠参数也是影响涡轮级效率的重要因素,叶片弯掠等三维造型技术可以通过改变叶栅径向压力梯度进而改善级效率。图 11 给出了三维积叠参数中对效率影响较大的周向顶部平移量 (D_2)、周向中间平移量 (D_1) 两个变量的作用效果。可以看出,二者均取负值时有利于涡轮级效率的提高。

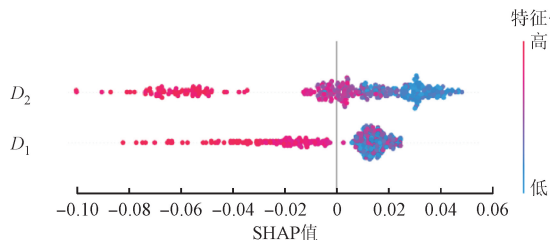
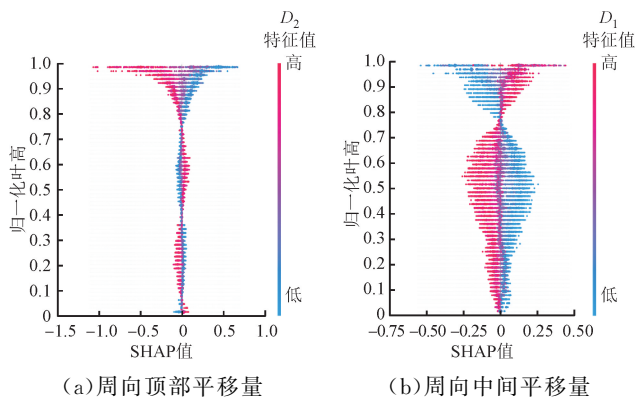


图 11 静叶三维积叠参数对涡轮级总体性能的影响

Fig. 11 Effect of 3D stacking parameters of stator on overall performance of turbine stage

通过 1 维 SHAP 分析对上述现象进行解释。图 12 为静叶三维积叠参数对沿叶高方向效率分布影响的示意图。可以看出: D_2 取负值,即周向顶部位置向叶片压力面偏移,可以提高叶顶区域的效率; D_1 取负值,即周向中间位置向压力面偏移,可以提高中间叶高区域的效率;二者综合作用形成正倾斜叶片,改善了叶栅通道中的径向压力梯度,有利于级效率的提高。



(a) 周向顶部平移量

(b) 周向中间平移量

图 12 静叶三维积叠参数对涡轮沿叶高方向性能的影响

Fig. 12 Effect of 3D stacking parameters of stator on turbine performance along spanwise

3.2.3 叶片吸力面前缘样条控制点

静叶、动叶中间截面吸力面前缘附近的控制点 P_{SL2} 也会对效率产生影响,如图 13 所示。可以看出,同时减小 P_{SL2-S_2} 、 P_{SL2-R_2} 有利于涡轮级效率的提升。

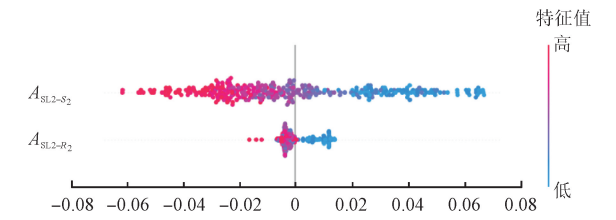


图 13 P_{SL2} 对涡轮级总体性能的影响

Fig. 13 Effect of P_{SL2} on overall performance of turbine stage

图 14 可以对相关机理作进一步解释,叶片中间截面吸力面前缘附近控制点取值减小,即向叶片变薄方向移动,有利于减小叶片前缘加速区的压力梯度,降低不可逆流动损失,从而提高涡轮级沿叶高方向大部分范围内的效率。对比图 14(a)和图 14(b)可以发现,相较于静叶中间截面 P_{SL2} 的影响,动叶中间截面 P_{SL2} 的作用相对较小。此外,由于叶顶间隙的存在,动叶顶部效率也会受到一定的影响。

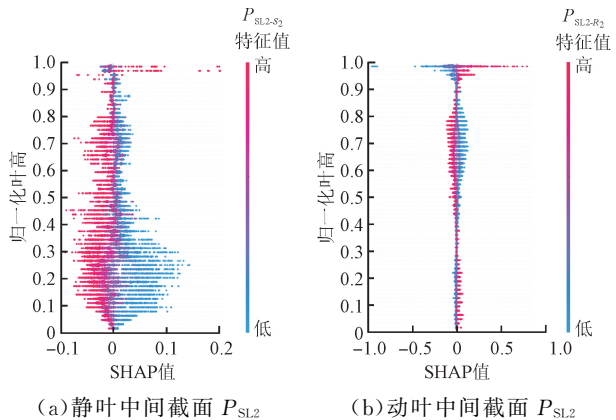


图 14 P_{SL2} 对涡轮级沿叶高方向性能的影响

Fig. 14 Effect of P_{SL2} on turbine performance along spanwise

3.3 典型样本归因分析

除了从 0 维数据到 1 维数据的分析,SHAP 与其他基于全局解释的数据挖掘方法相比的一大优势还体现在局部解释能力上,可以对单个样本进行归因分析,可视化各设计变量在样本性能指标变化过程中所发挥的作用。

选取裁剪后数据集中效率最大及最小的个体作为典型样本进行分析,结果如图 15 所示。图 15(a)为效率最大样本。可以看出,相较于数据集中所有样本效率的均值 90.456%,该样本效率提升为

90.956%,在这个过程中动叶中间截面有效出气角 (A_{E-R_2})、动叶叶根截面轴向弦长 (L_{C-R_1})、静叶中间截面吸力面前缘附近控制点 (P_{SL2-S_2}) 等变量发挥了重要作用。 A_{E-R_2} 取值为 0.112,小于基准值 0.5,使效率提高了 0.11%; L_{C-R_1} 取值为 0.007,使效率提高了 0.09%; P_{SL2-S_2} 取值为 0.029,使效率提高了 0.005%。图 15(b)为效率最小样本,相较于基准值 90.456%,该样本效率降低为 90.007%。动叶中间截面有效出气角 A_{E-R_2} 取值为 0.822,大于基准值 0.5,使效率降低了 0.09%,静叶叶顶截面轴向弦长 L_{C-S_3} 取值为 0.007,使效率降低了 0.08%,静叶周向中间平移量 D_1 取值为 0.911,使效率降低了 0.07%。

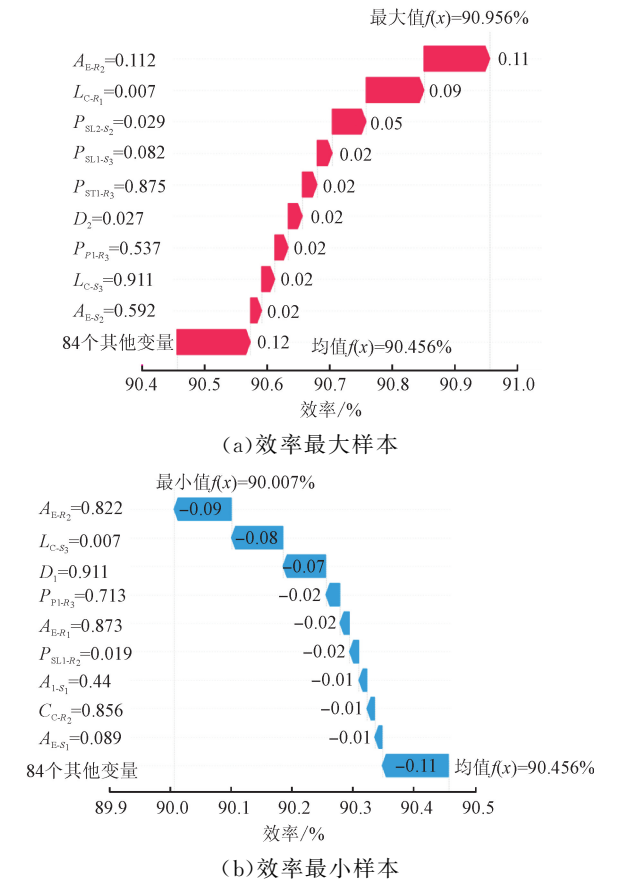


图 15 典型样本归因分析

Fig. 15 Typical sample attribution analysis

通过数据集内效率最大最小样本的归因分析可以发现,影响样本性能的重要参数包括有效出气角、叶片吸力面前缘控制点、静叶三维积叠、轴向弦长等,这与上文所识别的显著变量相同,其影响方向也与数据挖掘结果相一致。

3.4 设计准则归纳及气动校核

基于以上分析,可以总结出提高 GE-E3 高压涡轮第一级级效率的设计准则如下。

- (1)有效出气角。减小动叶中间截面有效出气角,增大静叶中间截面有效出气角。
- (2)静叶三维积叠参数。周向顶部位置、周向中间位置平移量均取负,即向压力面偏移。
- (3)叶片吸力面前缘控制点。静叶中间截面、动叶中间截面吸力面前缘附近控制点均向叶片变薄方向移动。

遵循上述设计准则,朝着有利于涡轮级效率提高的方向进行几何造型。首先,分别改变有效出气角、静叶三维积叠参数、叶片吸力面前缘控制点得到新设计。随后,同时改变以上所有参数,得到最终设计。不同设计之间效率的对比如表 2 所示。

表 2 不同设计的效率对比

Table 2 Different design efficiency comparison

几何设计	效率/%	较参考设计 增量/%
参考设计	90.44	
改变有效出气角的新设计	90.96	0.52
改变静叶三维积叠参数的新设计	90.58	0.14
改变叶片吸力面前缘控制点的新设计	90.48	0.04
最终设计	91.09	0.65

分析表 2 可知,基于设计准则得到的 4 种几何都使涡轮级效率有所提升。在只改变一个参数的情况下,有效出气角使效率提升最多,随后是静叶三维积叠参数,最后是叶片吸力面前缘控制点,这与前文 SHAP 饼图分析中所得到的显著变量贡献度占比相一致。完全遵循设计准则得到的最终设计的涡轮级效率提升至 91.09%,大于裁剪数据集内的最大值 90.96%,充分证明了基于 SHAP 的数据挖掘方法的有效性。

对参考设计及新设计进行 CFD 计算,进一步对比涡轮级气动性能的变化。利用子午面熵产来描述流动过程中产生的不可逆损失,其定义为流体由于黏性效应和速度梯度的相互作用产生的熵,公式为

$$\Phi = \mu \left[2 \left(\frac{\partial u}{\partial x} \right)^2 + 2 \left(\frac{\partial v}{\partial y} \right)^2 + 2 \left(\frac{\partial w}{\partial z} \right)^2 + \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x} \right)^2 + \left(\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y} \right)^2 \right] \quad (6)$$

式中: μ 为动力黏度; u, v, w 分别为流体在 x, y, z 方向上的速度, $2 \left(\frac{\partial u}{\partial x} \right)^2, 2 \left(\frac{\partial v}{\partial y} \right)^2, 2 \left(\frac{\partial w}{\partial z} \right)^2$ 能够表示流体在每个方向上的拉伸或压缩效应, $\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right)^2, \left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x} \right)^2, \left(\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y} \right)^2$ 能够表示速度梯度之间的

剪切效应。

图 16 为参考设计及最终设计子午面熵产分布云图。可以看出,相较于参考设计,最终设计子午面熵产明显减小。图中:区域 A 为静叶前缘,最终设计上下端壁处熵产值较高的位置均有所后移;区域 B 为动、静叶交界面,最终设计低熵产区域扩大;区域 C、D 为动叶出口上下端壁通道涡高熵产区,最终设计的下通道涡损失显著降低。

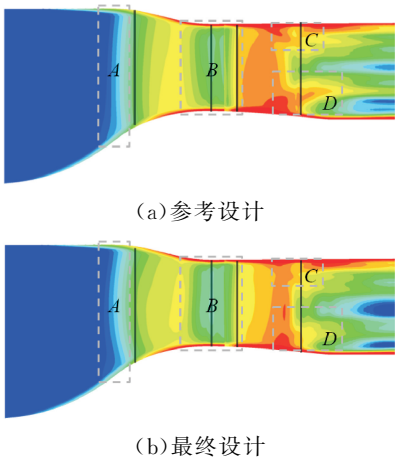


图 16 参考及最终设计子午面熵产分布云图
Fig. 16 Distribution of entropy production on the meridian plane of reference and final design

图 17 给出了参考设计及最终设计效率沿叶高方向的分布曲线。可以看出,相较于参考设计,最终设计沿叶高大部分范围内效率均有所提升,尤其是上端壁及中间叶高区域。

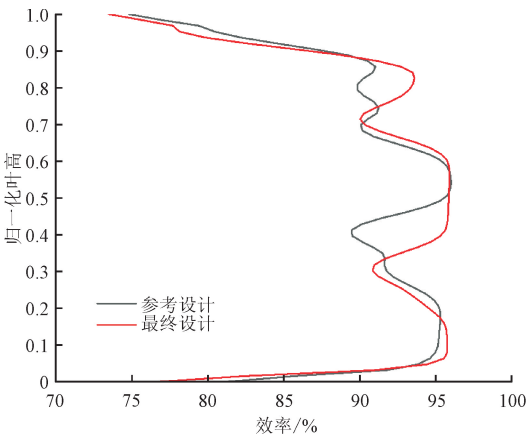


图 17 参考及最终设计效率沿叶高方向分布曲线
Fig. 17 Efficiency distribution curves along spanwise of reference and final design

4 结 论

本文构建起了基于 SHAP 的数据挖掘框架,在

此基础上完成了 GE-E3 高压涡轮第一级设计空间总体性能、沿叶高方向性能分布、典型样本的知识挖掘,由此归纳出了涡轮级的设计准则并进行了气动校核,主要获得以下结论。

(1)对于高压涡轮级效率,排名前3位的影响参数分别为叶片有效出气角、静叶三维积叠参数、叶片吸力面前缘样条控制点。动叶中间截面有效出气角减小、静叶中间截面有效出气角增大,中间叶高效率提升,二者综合作用可以在提升效率的同时控制流量的变化幅度。

(2)静叶三维积叠参数中,周向顶部位置向压力面偏移可以提高叶顶区域的效率,周向中间位置向压力面偏移可以提高中间叶高区域的效率,二者综合作用形成正倾斜叶片,改善了叶栅通道中的径向压力梯度,有利于涡轮级效率的提高。

(3)静、动叶中间截面吸力面前缘附近样条控制点向叶片变薄方向移动,有利于减小前缘加速区的压力梯度,提高沿叶高大部分范围内的效率。

(4)由设计准则得到的最终设计效率为91.09%,相较于参考设计提高了0.65%,大于裁剪数据集内的最大值,验证了 SHAP 分析的有效性。

本文所提出的基于 SHAP 的数据挖掘框架具有较强的普适性,可拓展用于各类叶轮机械多学科设计空间知识挖掘,为类似设计提供借鉴。

参考文献:

- [1] BENINI E. Three-dimensional multi-objective design optimization of a transonic compressor rotor [J]. *Journal of Propulsion and Power*, 2004, 20(3): 559-565.
- [2] 汪祺能,宋立明,郭振东,等.采用动态交互作用分析的叶轮机械优化算法研究[J].*西安交通大学学报*, 2023, 57(7): 139-150.
WANG Qineng, SONG Liming, GUO Zhendong, et al. Study on turbomachinery optimization algorithm based on dynamic interaction analysis [J]. *Journal of Xi'an Jiaotong University*, 2023, 57(7): 139-150.
- [3] VIANA F A C, SIMPSON T W, BALABANOV V, et al. Special section on multidisciplinary design optimization: metamodeling in multidisciplinary design optimization: how far have we really come? [J]. *AIAA Journal*, 2014, 52(4): 670-690.
- [4] WANG Qi, YANG Li, RAO Yu. Establishment of a generalizable model on a small-scale dataset to predict the surface pressure distribution of gas turbine blades [J]. *Energy*, 2021, 214: 118878.
- [5] JU Yaping, QIN Ruihong, KIPOUROS T, et al. A high-dimensional design optimisation method for centrifugal impellers [J]. *Proceedings of the Institution of Mechanical Engineers: Part A Journal of Power and Energy*, 2016, 230(3): 272-288.
- [6] LI Zhihui, ZHENG Xinqian. Review of design optimization methods for turbomachinery aerodynamics [J]. *Progress in Aerospace Sciences*, 2017, 93: 1-23.
- [7] BAERT L, DUMONT C, BEAUTHIER C, et al. Multidisciplinary design of a low-noise propeller: part II efficient aero-acoustic-mechanical design methodology exploiting surrogate models in an adaptive design space [C]//ASME Turbo Expo 2020: Turbomachinery Technical Conference and Exposition. New York, NY, USA: ASME, 2020: V02DT38A023.
- [8] CHIBA K, JEONG S, OBAYASHI S, et al. Data mining for multidisciplinary design space of regional-jet wing [C]//2005 IEEE Congress on Evolutionary Computation. Piscataway, NJ, USA: IEEE, 2005: 2333-2340.
- [9] IULIANO E, QUAGLIARELLA D. Proper orthogonal decomposition, surrogate modelling and evolutionary optimization in aerodynamic design [J]. *Computers & Fluids*, 2013, 84: 327-350.
- [10] 段焰辉,吴文华,范召林,等.基于本征正交分解的气动优化设计外形数据挖掘[J].*物理学报*, 2017, 66(22): 220203.
DUAN Yanhui, WU Wenhua, FAN Zhaolin, et al. Proper orthogonal decomposition-based data mining of aerodynamic shape for design optimization [J]. *Acta Physica Sinica*, 2017, 66(22): 220203.
- [11] OYAMA A, NONOMURA T, FUJII K. Data mining of Pareto-optimal transonic airfoil shapes using proper orthogonal decomposition [J]. *Journal of Aircraft*, 2010, 47(5): 1756-1762.
- [12] SHI Yayun, LAN Xiayu, WANG Yiwen, et al. Data mining-based design space exploration and optimization for tandem airfoils [J]. *International Journal of Aerospace Engineering*, 2023, 2023(1): 3513507.
- [13] SONG Liming, GUO Zhendong, LI Jun, et al. Research on metamodel-based global design optimization and data mining methods [J]. *Journal of Engineering for Gas Turbines and Power*, 2016, 138(9): 092604.
- [14] ADJEI R A, FAN Chengwei, WANG Weizhe, et al. Multidisciplinary design optimization for performance improvement of an axial flow fan using free-form deformation [J]. *Journal of Turbomachinery*, 2021, 143(1): 011003.

- [15] 郭振东, 王杰, 陈云, 等. 融合多源数据挖掘信息的非轴对称端壁/叶身联合成型设计空间知识挖掘 [J]. 西安交通大学学报, 2024, 58(1): 30-41.
GUO Zhendong, WANG Jie, CHEN Yun, et al. Exploration of the design space for non-axisymmetric end-wall/blade through integration of multi-source data mining techniques [J]. Journal of Xi'an Jiaotong University, 2024, 58(1): 30-41.
- [16] QIN Ruihong, JU Yaping, SPENCE S, et al. Meta-model-driven data mining model to support three-dimensional design of centrifugal compressor stage [J]. Journal of Turbomachinery, 2021, 143(12): 121013.
- [17] BAERT L, CHÉRIÈRE E, SAINVITU C, et al. Aerodynamic optimization of the low-pressure turbine module: exploiting surrogate models in a high-dimensional design space [J]. Journal of Turbomachinery, 2020, 142(3): 031005.
- [18] LUNDBERG S M, ERION G, CHEN H, et al. From local explanations to global understanding with explainable AI for trees [J]. Nature Machine Intelligence, 2020, 2(1): 56-67.
- [19] LUNDBERG S M, LEE S I. A unified approach to interpreting model predictions [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc., 2017: 4768-4777.
- [20] PALAR P S, ZUHAL L R, SHIMOYAMA K, et al. Shapley additive explanations for knowledge discovery via surrogate models [C]//AIAA SCITECH 2023 Forum. Reston, VA, USA: AIAA, 2023: AIAA 2023-0332.
- [21] SHEN Yang, HUANG Wei, YAN Li, et al. An automatic visible explainer of geometric knowledge for aeroshape design optimization based on SHAP [J]. Aerospace Science and Technology, 2022, 131(Part A): 107993.
- [22] ZHANG Xiawen, JU Yaping, LI Zhen, et al. Meta-model-interpreted data mining for stability and efficiency enhancement of multistage axial-flow compressors [J]. Journal of Turbomachinery, 2023, 145(4): 041001.
- [23] PALAR P S, DWIANTO Y B, ZUHAL L R, et al. Multi-objective design space exploration using explainable surrogate models [J]. Structural and Multidisciplinary Optimization, 2024, 67(3): 38.
- [24] PALAR P S, SHIMOYAMA K, OBAYASHI S. Polynomial-based Shapley additive explanations for design exploration in turbomachinery [C]//AIAA SCITECH 2024 Forum. Reston, VA, USA: AIAA, 2024: AIAA 2024-1226.
- [25] WANG Junying, HE Xiao, WANG Baotong, et al. Shapley additive explanations of multigeometrical variable coupling effect in transonic compressor [J]. Journal of Engineering for Gas Turbines and Power, 2022, 144(4): 041015.
- [26] WANG Qing, GUO Zhendong, SONG Liming, et al. A panoramic aerodynamic performance prediction method for turbomachinery cascades using transformer-enhanced neural operator [J/OL]. Chinese Journal of Aeronautics [2024-12-20]. <https://doi.org/10.1016/j.cja.2025.103473>.

(编辑 陶晴)