



基于强化学习的高超声速飞行器序贯融合建模方法

刘燕斌*, 王熙杰, 杨犇, 陈柏屹

南京航空航天大学航天学院, 南京 211106

* E-mail: liuyb@nuaa.edu.cn

收稿日期: 2025-12-04; 接受日期: 2026-01-29; 网络版发表日期: 2026-03-17

国家自然科学基金(批准号: 52272369, U2570206)和中央高校基本科研业务费专项资金(编号: NC2024009)资助项目

摘要 高超声速飞行器一体化设计阶段的性能分析需要大量的高精度模型数据以提高设计性能, 但高精度数据获取成本十分高昂. 如何在满足合理的精度指标前提下, 降低高精度数据量, 从而支持飞行器制导控制系统的迭代设计, 是亟待解决的重要技术领域. 为此, 本文提出一种高超声速飞行器基于强化学习的序贯融合建模方法, 以解决迭代设计过程的模型精度与建模成本难以平衡的问题. 首先, 应用偏差RBF核函数设计粗糙惩罚准则, 建立高超声速飞行器Co-RBF多可信度代理模型架构; 然后, 提出一种基于响应面极值点的主动探索策略, 对高超声速飞行器高可信度模型数据的样本空间进行采样; 进而, 基于强化学习策略设计一种模型数据自适应序贯迭代方法, 在节省成本的同时实现了高超声速飞行器多源模型数据的高效融合. 最后, 应用多个数值算例和仿真实例, 验证了所提出方法的有效性.

关键词 高超声速飞行器, 融合建模, 序贯采样, 强化学习

1 引言

人工智能的发展为工程设计提供了新思路, 在面向复杂工程系统的设计中, 真实物理系统的实验数据难以获得, 常常使用计算机仿真来对系统的响应信息进行获取. 在航空航天领域的工程设计中, 为了获取精确的气动响应, 常常需要进行真实的飞行试验^[1]. 飞行试验数据获取成本非常高昂, 获取的数据样本往往较少, 而且只分布在设计空间中某些分散的区域中, 具有小子样的特征. 因此工程设计人员常常使用计算

机仿真的方法对于高超声速飞行器系统的响应规律进行数值模拟^[2].

然而, 在高精度模型的仿真中, 高可信度响应信息获取成本较大. 例如在高超声速飞行器气动设计中的CFD仿真, 有时样本中单点响应信息的获取就要消耗数个小时, 这使得探索整个设计空间的计算成本变得相当高昂^[3]. 为了降低仿真计算成本, 研究者提出了建立代理模型来模拟系统的输入-输出响应的映射关系. 通过插值、拟合等手段, 实现对高精度模型的近似等效. 自Sacks等人^[4]发表关于计算机设计的革命性文章

引用格式: 刘燕斌, 王熙杰, 杨犇, 等. 基于强化学习的高超声速飞行器序贯融合建模方法. 中国科学: 技术科学, 2026, 56: 790–808
Liu Y B, Wang X J, Yang B, et al. A sequential fusion modelling approach for hypersonic vehicles based on reinforcement learning (in Chinese). Sci Sin Tech, 2026, 56: 790–808, doi: 10.1360/SST-2025-0394

以来,代理模型在过去的40年得到了快速发展.代理模型技术具有控制仿真计算成本的优势,在多学科优化设计中备受青睐,被广泛应用到飞行器设计与其他工程应用领域中^[5-8].当前常见的代理模型包括多项式响应面(polynomial response surface, PRS)^[9]、克里金(Kriging)^[10]、支持向量回归(support vector regression, SVR)^[11]、径向基函数(radial basis functions, RBF)^[12]、人工神经网络(artificial neural network, ANN)^[13]、以及物理信息神经网络(physics-informed neural network, PINN)^[14]等.

数据驱动代理建模的性能依赖于样本数据的精度与规模.高精度样本信息数量较少,要获取精度更高的模型,仍需要大规模调用高精度模型库^[15],从而增加了高精度仿真计算成本.另一方面,研究者可以通过使用成本更低、更高效的建模手段,获取可信度较低的响应信息.然而,仅仅依靠大量中低可信度的廉价数据,会影响模型的准确度与泛化性能.如何平衡建模成本与模型精度的矛盾,更有效地将多可信度的模型数据组合并获取最优系统响应,是数据驱动建模的关键问题.近二十年来,以高可信度数据为基础,结合大量低可信度数据的多可信度代理模型(multi-fidelity surrogate, MFS)方法,成为广大工程设计人员研究的热点^[16-18].

然而,多可信度代理模型的精度主要取决于样本点的位置与数量,这正是代理建模关键的一步:实验设计(design of experiments, DOE),如拉丁超立方采样^[19]、全因子设计^[20]等传统一次性DOE方法已经无法满足当今复杂的工程设计问题,基于序贯采样的建模方法具有用更少的样本点构建更精确代理模型的巨大潜力,因此越来越受到各领域研究者的关注^[21].序贯建模方法在机器学习领域通常被称为主动学习^[22],其核心在于采样准则的设计.

序贯采样指通过从当前的数据和代理模型中学习或挖掘有效信息,以指导下一个采样点的位置,这是与传统的空间填充采样的主要区别.序贯建模一般对代理模型没有过多要求,序贯建模的核心在于采样准则的设计.目前,期望提升(expected improvement, EI)^[23]准则、置信下界(lower confidence bounds, LCB)^[24]准则、均方误差(mean squared error, MSE)^[25]准则等得到了广泛的研究与应用.

全局代理建模的序贯采样策略主要可以分为基于

方差、基于交叉验证、基于梯度三类^[26].基于方差的策略利用模型提供的方差估计来选择采样点,代表性工作包括最大化均方误差、最大化互信息以及熵准则等^[25,27-29];基于交叉验证的策略不依赖于特定的模型类型,通过评估留一误差等指标寻找潜在的关键区域,如Li等人^[30]、Aute等人^[31]、Xu等人^[32]及Zhou等人^[33]提出的改进策略;基于梯度的策略则利用导数信息衡量局部的非线性程度,从而辨识函数变化剧烈的区域进行针对性采样^[34-37].

尽管现有序贯采样方法通过构造优化准则都在一定程度上实现了全局探索与局部开发的折中,但大多依赖静态启发式规则,难以捕捉建模中的动态特性.而高超声速飞行器气动特性复杂,仅依赖传统静态启发式规则无法保证多源数据的融合效率与建模精度.强化学习(reinforcement learning, RL)^[38]为克服这一局限提供了新的解决思路:通过将采样过程形式化为序列决策优化问题,它能够驱动智能体根据模型状态自适应地选择最优策略.

针对上述问题,本文在多可信度协同径向基函数(Co-RBF)模型的基础上,提出了一种基于强化学习指导的自适应多可信度序贯建模框架(RL based adaptive sampling cooperative RBF, RLASC-RBF).首先,通过引入粗糙惩罚准则构建改进的Co-RBF模型,增强了模型对低可信度数据偏差的修正能力与泛化性能;其次,充分挖掘低可信度数据的全局趋势信息,设计基于极值点主动探索(active exploration sampling, AES)的初始采样策略,有效引导高精度样本聚焦于关键非线性区域;进而,将序贯采样过程建模为强化学习的序列决策问题,通过智能体感知模型当前的局部曲率与稀疏度状态,实现采样策略的自适应调整.最后,将该方法应用于高超声速飞行器气动特性建模及燃油最省定动压爬升轨迹优化任务中.

2 问题描述

2.1 高超声速飞行器模型

本文研究对象为典型的吸气式高超声速巡航飞行器^[39].该模型采用细长体气动布局,以满足高超声速飞行条件下的高升阻比与低阻力需求,同时实现弹体与超燃冲压发动机的高效集成.该飞行器头部采用非轴对称的下颌式进气道设计.前体下表面构成压缩坡

面, 利用飞行时产生的激波对来流空气进行预压缩, 以满足燃烧室入口的气动参数要求, 并采用X型全动尾翼的气动舵面控制方式。

高超声速飞行器飞行包线宽阔, 气动环境复杂, 其纵向运动与横侧向运动之间存在强非线性耦合效应。本文旨在研究基于强化学习的序贯融合建模方法, 且主要聚焦于飞行器的纵向气动特性, 为构建适用的动力学模型, 对飞行器物理特性及飞行环境作如下假设:

- (1) 假设飞行器关于纵向平面具有几何对称性及质量分布对称性;
- (2) 仅考虑飞行器在纵向平面内的运动, 忽略横向侧滑角、滚转角及偏航角的影响;
- (3) 视飞行器为刚体, 忽略气动弹性效应及结构变形对气动力及力矩的非定常影响;
- (4) 忽略大气紊流扰动, 并忽略地球自转引起的科氏力及离心力对飞行轨迹的影响。

基于上述假设, 对高超声速飞行器的纵向运动与横侧向运动进行解耦, 基于拉格朗日法建立的高超声速飞行器纵向模型为^[40]

$$\begin{cases} \dot{v} = \frac{T \cos \alpha - D}{m} - g \sin(\theta - \alpha), \\ \dot{\alpha} = q + \frac{-T \sin \alpha - L}{mv} + \frac{g \cos(\theta - \alpha)}{v}, \\ \dot{h} = v \sin(\theta - \alpha), \\ \dot{\theta} = q, \\ \dot{q} = \frac{M}{I_y}, \\ \dot{m} = \frac{-T}{g_0 I_{sp}}, \end{cases} \quad (1)$$

其中, v 、 α 、 h 、 θ 、 q 和 m 分别表示飞行速度、攻角、高度、俯仰角、俯仰角速率以及飞行器的质量; I_{sp} 和 T 分别代表比冲和发动机推力, 与飞行过程中的燃油消耗有关; M 、 L 和 D 分别是飞行过程中产生的俯仰力矩、升力与阻力; g 和 I_y 分别代表当地重力加速度以及绕 y 轴的转动惯量。

2.2 气动特性估算

为了更好地验证本文所提方法, 需要对研究对象构建多可信度气动数据库。气动特性数据的获取来源一般有四类: 基于经验公式估算、CFD仿真计算、风洞试验和飞行试验。本节主要通过基于工程估算的机理建模与CFD仿真获取两种可信度气动数据。将基于

工程估算的机理建模作为低可信度数据, CFD仿真的数据作为高可信度数据。

在高速气动工程估算中, 首先需要明确每一个面元上的来流信息。面元的外法向单位矢量和来流速度分别为 $\mathbf{n}^T = [n_x, n_y, n_z]$, $\mathbf{V}_\infty^T = [V_x, V_y, V_z]$ 。则气流撞击角 δ_w 、物面倾角 θ_w 、径向角 φ_w 以及等价锥角 θ_{TC} 为

$$\begin{cases} \delta_w = \frac{\pi}{2} - \arccos\left(\frac{-\mathbf{n} \cdot \mathbf{V}}{\|\mathbf{n}\| \cdot \|\mathbf{V}\|}\right), \\ \theta_w = \frac{\pi}{2} - \arccos(n_x), \\ \varphi_w = \pm \arccos\left(\frac{n_y}{\sqrt{n_y^2 + n_z^2}}\right), \\ \theta_{TC} = \arcsin(\sin \beta + \cos \theta_w \sin \beta \sin \varphi). \end{cases} \quad (2)$$

确定来流信息后, 本文采用Dahlem-Buck法^[41]对面元压力系数 C_p 进行计算:

$$C_p = \begin{cases} \left[\frac{1}{\sin^{3/4}(4\delta_w)} + 1 \right] \sin^2 \delta_w, & \delta_w \leq 22.5^\circ, \\ K_w \sin^2 \delta_w, & \delta_w > 22.5^\circ. \end{cases} \quad (3)$$

再将飞行器划分为 N_w 个面元, 在获得第 i 个面元的压强系数后, 每个面元上的气动力 $\mathbf{F}_i$ 及相对指定参考点的气动力矩 $\mathbf{M}_i$ 为

$$\begin{cases} \mathbf{F}_i = -(C_{p,i} Q_\infty + p_\infty) \cdot S_i \cdot \mathbf{n}_i, \\ \mathbf{M}_i = -(C_{p,i} Q_\infty + p_\infty) \cdot S_i \cdot \mathbf{n}_i \times \mathbf{r}_i, i = 1, 2, \dots, N_w, \end{cases} \quad (4)$$

其中, $C_{p,i}$ 为面元压力系数; p_∞ 和 Q_∞ 分别为自由流静压和动压; S_i 为面元面积; $\mathbf{n}_i$ 为面元的外法向单位矢量; $\mathbf{r}_i$ 为面元相对于参考点的位移矢量。

同时, 使用Fluent软件对飞行器外形进行RANS CFD数值求解, 获取高可信度气动数据, 仿真整机网格数量约为 5.8×10^6 。

本文针对具有机体/推进一体化特征的吸气式高超声速飞行器^[39], 采用上述工程估算方法计算飞行器的升力系数 C_L 、阻力系数 C_D 和俯仰力矩系数 C_M , 并和CFD计算得到的数据进行对比。图1展示了在5~8马赫范围内两种可信度模型的拟合曲面对比情况。

本文选取 Ma 5~8作为高超声速飞行器建模与验证的核心速域, 主要是因为该区间涵盖了吸气式高超声速飞行器超燃冲压发动机稳定工作的典型巡航包

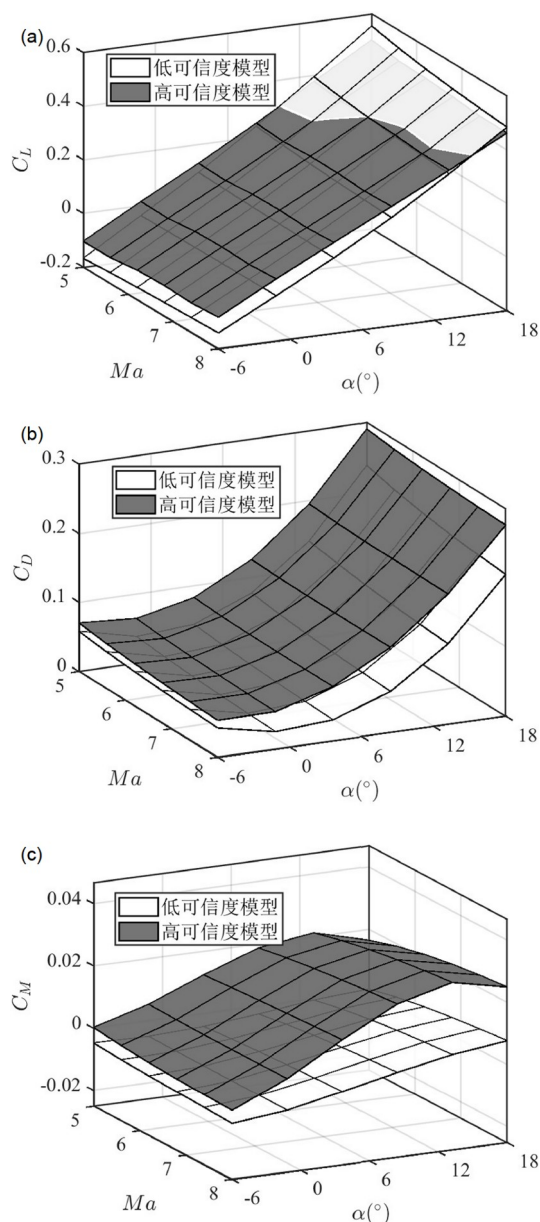


图 1 高低可信度模型拟合曲面对比。(a) 升力系数; (b) 阻力系数; (c) 俯仰力矩系数

Figure 1 Comparison of high and low fidelity model fitting curves. (a) Lift coefficient; (b) drag coefficient; (c) pitch moment coefficient.

线, 是飞行任务中耗时最长、对气动建模精度要求最为严苛的飞行阶段。理论上来说, 工程估算模型可覆盖更宽广的飞行包线, 但考虑到该速域内气动/推进耦合效应显著且流场非线性特征复杂, 针对此核心工况的研究更具代表性, 也更能有效验证本文序贯融合建模方法的精度优势。

由图1可知, 基于工程估算的机理建模方法虽然在整体趋势上符合规律, 但相较于CFD参考数据仍然存在不可忽略的误差, 若进一步对比精度更高的风洞试验数据与飞行试验数据, 则偏差更为显著。飞行器设计流程前期, 为满足设计方案的快速迭代, 仅依靠大样本的低精度数据或者高精度数据, 将导致精度难以保证或者经济和时间成本过高。因此需要构建更优的多源数据驱动建模策略, 以平衡建模精度与成本, 加快高超声速飞行器设计周期。

3 基于强化学习的自适应序贯建模方法

传统的序贯采样方法主要依赖于静态的启发式策略, 往往通过预设的线性加权机制在探索与利用之间寻求平衡^[42]。然而, 此类方法难以捕捉模型构建过程中的动态特性, 无法依据实时反馈灵活调整采样策略。引入强化学习框架可以有效克服上述静态启发式策略的局限性, 即通过将采样决策过程形式转化为序列决策优化问题, 该方法能够驱动算法依据当前模型的构建状态, 自适应地选择最优采样策略, 进而实现样本利用效率与模型全局精度的双重最大化。

鉴于强化学习在解决此类序列决策优化问题上的显著优势, 为突破传统静态采样策略的瓶颈, 本文提出一种基于强化学习的自适应序贯融合建模(RLASC-RBF)方法, 构建了如图2所示的算法框架。该框架由多可信度融合环境与强化学习采样智能体两大核心模块构成, 通过感知、决策和反馈的闭环交互机制实现模型的主动迭代更新。

多可信度融合环境作为物理场的重构载体, 通过内部的多源数据驱动机制实现代理模型的主动迭代更新。具体而言, 多源数据层负责整合高、低可信度样本集, 为Co-RBF融合模型提供实时更新的训练支撑; Co-RBF模型通过解析其Hessian矩阵与预测残差, 将物理场的几何特征与统计特征传递给状态计算模块, 从而生成描述当前建模阶段的高维状态向量 S_t 。

强化学习采样智能体作为系统的决策核心, 主要依据环境反馈的状态信息, 利用Q-learning策略网络在动作空间中执行策略映射。智能体在权衡全局探索与局部开发后, 输出最优动作 A_t , 指导系统在设计空间中筛选出最具价值的新增样本 x_{new} 。新增样本再被反

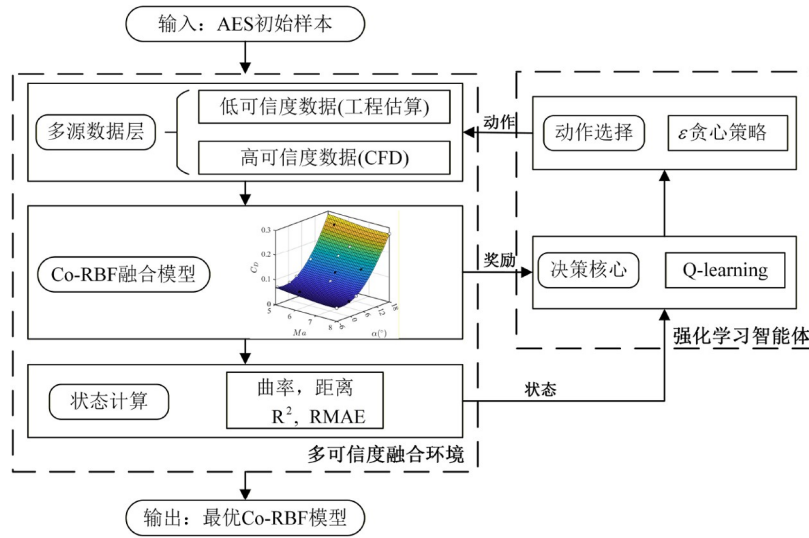


图2 (网络版彩图)基于强化学习的自适应序贯融合建模算法框架图

Figure 2 (Color online) Adaptive sequential fusion modelling algorithm framework based on reinforcement learning.

馈至多源数据层, 触发新一轮的模型训练与状态更新。同时, 环境依据模型更新前后的精度提升指标计算奖励信号并传递给智能体, 驱动其策略网络不断优化, 最终实现针对高超声速飞行器非线性气动特征的高效序贯融合建模。

综上所述, 多可置信度融合环境与强化学习采样智能体的深度耦合框架构建了数据流与决策流相互促进的动态闭环。从而在有限的计算资源下, 实现了高超声速飞行器气动建模精度与效率的全局最优。

3.1 基于偏差RBF设计的粗糙惩罚准则

3.1.1 RBF单可置信度模型

RBF模型以样本点作为径向基函数的中心, 通过对称基函数线性组合而成, 具有较高的模型构建效率和拟合精度, 能够实现对任意函数的逼近^[43]。对于 n 维输入变量 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]^T$, 与设计点对应的响应值为 $\mathbf{Y} = [y(\mathbf{x}_1), y(\mathbf{x}_2), \dots, y(\mathbf{x}_n)]^T$, 则RBF模型在未知点 $\mathbf{x}$ 处的预测响应值可以表示为

$$\hat{y}(\mathbf{x}) = \sum_{i=1}^n \varphi(r_i) \cdot \lambda_i, \quad (5)$$

其中, $r_i = \|\mathbf{x} - \mathbf{x}_i\|$ 表述未知点 $\mathbf{x}$ 和训练集样本 $\mathbf{x}_i$ 之间的欧氏距离; λ_i 为对应于基函数的权重系数; $\varphi(\cdot)$ 为径向基函数。

由于MQ(multi-quadric)函数插值精度高、稳定性较好, 便于进行大型矩阵运算, 因此本文采用MQ函数作为基函数。其定义如下:

$$\varphi(r_i, c) = \sqrt{r_i^2 + c^2}, \quad (6)$$

其中, c 为模型超参数形状因子。

将训练集 $\mathbf{X}$ 和其响应值 $\mathbf{Y}$ 代入式(5)中, 得到表达式:

$$\begin{bmatrix} y(\mathbf{x}_1) \\ y(\mathbf{x}_2) \\ \vdots \\ y(\mathbf{x}_n) \end{bmatrix} = \begin{bmatrix} \varphi(\|\mathbf{x}_1 - \mathbf{x}_1\|) & \cdots & \varphi(\|\mathbf{x}_1 - \mathbf{x}_n\|) \\ \varphi(\|\mathbf{x}_2 - \mathbf{x}_1\|) & \cdots & \varphi(\|\mathbf{x}_2 - \mathbf{x}_n\|) \\ \vdots & \ddots & \vdots \\ \varphi(\|\mathbf{x}_n - \mathbf{x}_1\|) & \cdots & \varphi(\|\mathbf{x}_n - \mathbf{x}_n\|) \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix}. \quad (7)$$

将上式简写为矩阵乘积形式:

$$\mathbf{Y} = \boldsymbol{\Phi} \cdot \boldsymbol{\lambda}. \quad (8)$$

当且仅当矩阵 $\varphi(\|\mathbf{x} - \mathbf{x}_i\|)$ 均为正定矩阵时, 上式存在唯一解, 则RBF模型的权重系数向量解算为

$$\boldsymbol{\lambda} = \boldsymbol{\Phi}^{-1} \cdot \mathbf{Y}. \quad (9)$$

除了权重系数向量以外, 还需确定超参数形状因子 c 。由于形状因子对RBF模型的拟合精度以及泛化误差影响较大, 因此本文采用留一交叉验证(leave-one-out cross validation, LOOCV)^[44]的误差平方和作为性能指标 J 构建如下超参数优化问题:

$$J = \min_c \sum_{k=1}^n (\text{LOOCV}_k)^2, \quad (10)$$

$$\text{s.t. } c_{lb} \leq c \leq c_{ub},$$

其中, LOOCV_k 表示第 k 个样本点处的留一交叉验证误差; c_{lb} 和 c_{ub} 分别为超参数的下、上边界。

由于留一交叉验证法需要反复构建 n 次模型, 耗时较长, 因此使用如下公式对留一交叉误差进行近似计算:

$$\text{LOOCV}_k = y(\mathbf{x}_k) - S^{(k)}(\mathbf{x}_k) = \frac{\lambda_k}{(\boldsymbol{\Phi}^{-1})_{kk}}, \quad (11)$$

其中, $S^{(k)}(\mathbf{x}_k)$ 表示缩减样本集在剔除样本点处的插值; $(\boldsymbol{\Phi}^{-1})_{kk}$ 表示矩阵的第 k 行第 k 列元素。

本文采用基于协方差矩阵传播的进化算法, 无需使用梯度信息, 即可完成对上述参数优化问题的求解。

3.1.2 Co-RBF 多可信度代理模型

RBF 模型凭借其优异的局部非线性逼近能力, 已成为高可信度气动数据建模的有力工具。然而, 由于本文旨在解决多源异构数据的序贯融合问题, 仅依赖单一高精度模型往往难以兼顾成本与效率。为此, 引入优化协同径向基函数多可信度代理模型, 以构建高、低可信度数据之间的关联映射。

假设系统状态变量 $\mathbf{x} \in \mathbf{R}^{m \times 1}$, 响应输出为 $y \in \mathbf{R}$, 定义高可信度模型包含 n_H 个已知的观测样本点 $\mathbf{X}_H = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n_H}]^T$ 及其对应的输出响应 $\mathbf{Y}_H = [y_H(\mathbf{x}_1), y_H(\mathbf{x}_2), \dots, y_H(\mathbf{x}_{n_H})]^T$; 低可信度模型包含 n_L 个已知的观测样本点 $\mathbf{X}_L = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n_L}]^T$ 及其对应的输出响应 $\mathbf{Y}_L = [y_L(\mathbf{x}_1), y_L(\mathbf{x}_2), \dots, y_L(\mathbf{x}_{n_L})]^T$ 。受限于高昂的计算成本, 高可信度样本往往较为稀疏, 高可信度样本数量一般小于低可信度样本数量, 即 $n_H < n_L$ 。基于 Kennedy-O'Hagan(KOH)自回归框架, 利用缩放系数 ρ 与偏差 RBF 模型可建立如下的多可信度 Co-RBF 融合模型以近似高可信度模型:

$$\hat{y}_H(\mathbf{x}) = \rho \hat{y}_L(\mathbf{x}) + \hat{d}(\mathbf{x}), \quad (12)$$

其中, $\hat{y}_L(\mathbf{x})$ 为近似低可信度模型, $\hat{d}(\mathbf{x})$ 为偏差 RBF 模型:

$$\hat{d}(\mathbf{x}) = \sum_{i=1}^{n_H} \varphi_d(\mathbf{x}) \cdot \lambda_d, \quad (13)$$

其中, $\varphi_d(\cdot)$ 为径向基函数, $\lambda_d^{(i)}$ 为对应的权重系数。偏差 RBF 模型实现了对缩放低可信度模型的校正, 从而允许不同可信度的模型具有不同的相关结构。

将式(12)展开写作

$$\begin{bmatrix} y_H(\mathbf{x}_1) \\ y_H(\mathbf{x}_2) \\ \vdots \\ y_H(\mathbf{x}_{n_H}) \end{bmatrix} = \begin{bmatrix} \hat{y}_L(\mathbf{x}_1) & \varphi_d(\|\mathbf{x}_1 - \mathbf{x}_1\|) & \cdots & \varphi_d(\|\mathbf{x}_1 - \mathbf{x}_{n_H}\|) \\ \hat{y}_L(\mathbf{x}_2) & \varphi_d(\|\mathbf{x}_2 - \mathbf{x}_1\|) & \cdots & \varphi_d(\|\mathbf{x}_2 - \mathbf{x}_{n_H}\|) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{y}_L(\mathbf{x}_{n_H}) & \varphi_d(\|\mathbf{x}_{n_H} - \mathbf{x}_1\|) & \cdots & \varphi_d(\|\mathbf{x}_{n_H} - \mathbf{x}_{n_H}\|) \end{bmatrix} \begin{bmatrix} \rho \\ \lambda_d^{(1)} \\ \lambda_d^{(2)} \\ \vdots \\ \lambda_d^{(n_H)} \end{bmatrix}. \quad (14)$$

进一步改写为矩阵符号形式:

$$\mathbf{Y}_H = [\hat{\mathbf{y}}_L(\mathbf{X}_H) \quad \boldsymbol{\Phi}_d] \begin{bmatrix} \rho \\ \lambda_d \end{bmatrix}^T. \quad (15)$$

显然, 当矩阵 $[\hat{\mathbf{y}}_L(\mathbf{X}_H) \quad \boldsymbol{\Phi}_d]$ 满秩时, 上式有唯一解:

$$\begin{bmatrix} \rho \\ \lambda_d \end{bmatrix}^T = [\hat{\mathbf{y}}_L(\mathbf{X}_H) \quad \boldsymbol{\Phi}_d]^T \times \left([\hat{\mathbf{y}}_L(\mathbf{X}_H) \quad \boldsymbol{\Phi}_d] [\hat{\mathbf{y}}_L(\mathbf{X}_H) \quad \boldsymbol{\Phi}_d]^T \right)^{-1} \mathbf{Y}_H. \quad (16)$$

为增强 Co-RBF 模型的泛化能力, 本文将偏差 RBF 模型响应面的粗糙惩罚项融入优化目标函数中, 建立了改进的 Co-RBF 多可信度代理模型架构。

在求解径向基函数的形状参数 c_d 时, 沿用与传统 RBF 模型一致的 LOOCV 策略。定义优化问题如下:

$$\min \sum_{k=1}^{n_H} (\text{LOOCV}_k)^2 + k_d \cdot \int_D \sum_{i=1}^d \sum_{j=1}^d \left(\frac{\partial^2 \hat{d}(\mathbf{x})}{\partial x_i \partial x_j} \right)^2 d\mathbf{x}, \quad (17)$$

$$\text{s.t. } c_d^{lb} \leq c_d \leq c_d^{ub},$$

其中, k_d 为粗糙惩罚项性能指标的权重。同理, 利用协方差矩阵自适应进化策略 (CMA-ES)^[45] 即可求解上述超参数优化问题。

在第 g 代, 算法首先基于当前的分布均值 $m^{(g)}$ 、全局步长 $\sigma^{(g)}$ 以及协方差矩阵 $C^{(g)}$, 通过多元正态分布采样生成 n 个形状因子的候选解:

$$c_k^{(g+1)} \sim m^{(g)} + \sigma^{(g)} \mathcal{N}(0, C^{(g)}), k = 1, \dots, n, \quad (18)$$

其中, $c_k^{(g+1)}$ 表示第 $g+1$ 代第 k 个候选个体的形状因子值. 随后, 将所有候选解代入式(15)计算适应度值, 并依据误差大小对个体进行排序. 选取适应度最好的前 μ 个, 通过加权线性组合计算下一代的分布均值:

$$m^{(g+1)} = \sum_{i=1}^{\mu} w_i c_{i:n}^{(g+1)}, \quad (19)$$

其中, w_i 为归一化权重系数, $c_{i:n}^{(g+1)}$ 代表排序后的第 i 个最优形状因子. 为了捕捉目标函数曲面的几何特征并加速收敛, 算法进一步利用进化路径机制对协方差矩阵进行自适应调整. 协方差进化路径 p_c 的更新公式为

$$p_c^{(g+1)} = (1 - c_c) p_c^{(g)} + h_{\sigma} \sqrt{c_c(2 - c_c)} \mu_{\text{eff}} \times \frac{m^{(g+1)} - m^{(g)}}{\sigma^{(g)}}. \quad (20)$$

进而结合更新策略, 协方差矩阵 C 迭代如下:

$$C^{(g+1)} = (1 - c_1 - c_{\mu}) C^{(g)} + c_p p_c^{(g+1)} (p_c^{(g+1)})^T + c_{\mu} \sum_{i=1}^{\mu} w_i y_{i:n}^{(g+1)} (y_{i:n}^{(g+1)})^T, \quad (21)$$

其中, $y_{i:n}^{(g+1)} = (c_{i:n}^{(g+1)} - m^{(g)}) / \sigma^{(g)}$ 为标准化后的变异向量. 通过上述迭代机制, CMA-ES 能够自适应地调整搜索分布的椭球形状, 从而高效地确定出使目标函数最小的最优形状因子 c .

由式(12)可知, 低可信度RBF模型是主要趋势项, 偏差RBF模型是修正项. 考虑到低可信度数样本量相对较多, 全局趋势易于拟合, 因此Co-RBF模型的精度主要取决于偏差修正项.

为验证优化Co-RBF多可信度代理建模的有效性, 选取一个一维Forrester函数进行测试, 其中高可信度函数表达式为 $y_H = (6x - 2)^2 \sin(12x - 4)$, 低可信度函数表达式为 $y_L = 0.5y_H + 10(x - 0.5) - 5$. 选择MQ函数, 此时缩放系数 $\rho = 1.949$, 形状因子 $c_d = 1.172$, Co-RBF模型具有良好预测性能, 预测曲线如图3所示.

3.2 基于响应面极值点的主动探索策略

初始采样作为代理模型构建的前置环节, 对模型全局精度及后续序贯更新的收敛效率起着重要作用. 针对单可信度建模问题, 拉丁超立方抽样及其优化方法^[46] 凭借其良好的空间填充特性, 已得到广泛应用.

而在多可信度建模场景下, 通常采用嵌套采样范式^[47], 即约束高可信度样本集为低可信度样本集的子集, 如图4所示.

尽管低可信度模型精度有限, 但其整体变化趋势与极值分布往往能够反映物理场的关键特征. 为此, 本文提出一种基于响应面极值点的主动探索策略(AES). 该策略旨在充分利用低可信度模型具有潜在价值的样本点, 以此为依据指导高可信度样本的增选. 具体探索步骤如下.

(1) 低可信度模型构建: 利用基于最大最小距离准则的优化拉丁超立方抽样方法获得低可信度初始样本集 $\mathbf{X}_L$, 并构建低可信度RBF模型 $\hat{y}_L$.

(2) 潜在极值点探索: 在探索空间随机生成 n 个寻

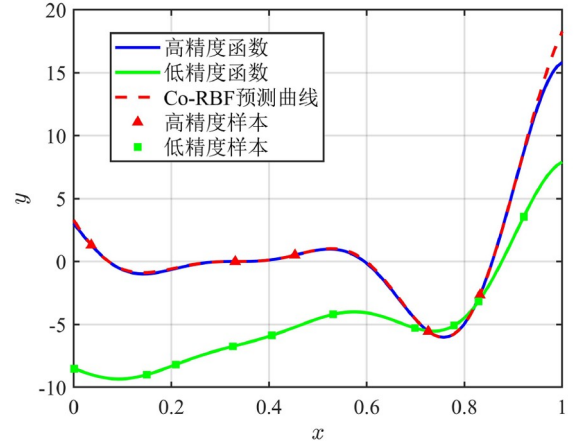


图3 (网络版彩图) 一维Forrester函数Co-RBF建模示意图
Figure 3 (Color online) Schematic diagram of one-dimensional Forrester function Co-RBF modelling.

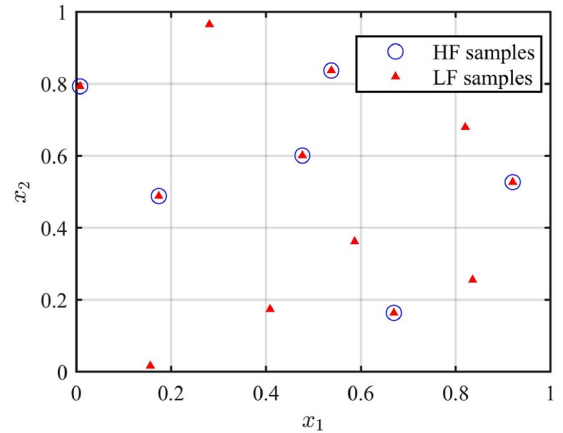


图4 (网络版彩图) 嵌套采样示意图
Figure 4 (Color online) Nested sampling diagram.

优初始点, 利用多起点^[48](Multistart)算法得到低可信度模型的 n 个局部极值点.

(3) 特征聚类: 通过聚类算法^[49]对上一步得到的局部极值点进行分类, 得到 n' 个具有相似属性的极值点集合 $\{\mathbf{X}_k, k = 1, 2, \dots, n'\}$.

(4) 高可信度样本筛选: 对于每一个集合 $\mathbf{X}_k$, 筛选出低可信度模型预测值 $\hat{y}_L$ 最小的极值点作为高可信度初始样本集 $\mathbf{X}_H = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$.

图5展示了二维Brannin函数基于极值点的主动探索步骤. 在使用多起点优化算法后, 得到低可信度模型的所有局部极值点, 如图5(a)所示, 然后再对极值点进行聚类汇总, 得到图5(b)所示的三个潜在探索点, 作为高精度模型的初始采样点.

本文提出的AES初始采样策略优势在于: 一方面, 增强了模型对强非线性区域的特征捕捉能力, 有效避免了计算资源在平坦区域的浪费, 从而更精准地重构物理场的复杂形态; 另一方面, 提升了工程优化问题的寻优效率与收敛精度, 这种启发式的初始化分布为后续的序贯更新提供了高质量的搜索空间, 在保证模型精度的前提下, 加速了极值搜索的收敛过程.

3.3 基于强化学习策略的自适应迭代方法

针对Co-RBF模型样本的迭代问题, 本文设计了一种基于强化学习(Q-learning^[50])策略的自适应迭代方法, 实现高超声速飞行器多源模型数据的高效融合. 该方法的核心思想是将样本点的选择过程视为一个序列

决策优化问题: 智能体根据当前模型的状态特征, 自适应地调整采样偏好, 在开发与探索之间寻找最优平衡, 从而实现模型精度的最大化.

定义状态空间 S 由四个核心维度构成: 局部曲率、欧氏距离、拟合优度 R^2 以及最大绝对误差 $RMAE$.

精度指标相关表达式为^[51]

$$\begin{cases} R^2 = 1 - \frac{\sum_i (\hat{y}_i - y_i)^2}{\sum_i (\bar{y}_i - y_i)^2}, \\ RMAE = \frac{\max(|y - \hat{y}|)}{\sigma(y)}, \end{cases} \quad (22)$$

其中, y_i 代表第 i 个样本点的真实响应值; $\hat{y}_i$ 代表代理模型在对应点的预测值; $\bar{y}_i$ 代表真实值的平均值; $\sigma(y)$ 代表真实值的标准差.

为了引导采样点向强非线性区域聚集, 引入Hessian矩阵的主曲率^[35]作为衡量标准. 对于一个 d 维的多元函数 $y(x_1, x_2, \dots, x_d)$, 其隐含函数的非线性信息的二阶偏导Hessian矩阵为

$$\mathbf{H}(y) = \frac{\partial^2 y}{\partial x^2} = \begin{bmatrix} \frac{\partial^2 y}{\partial x_1 \partial x_1} & \frac{\partial^2 y}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 y}{\partial x_1 \partial x_d} \\ \frac{\partial^2 y}{\partial x_2 \partial x_1} & \frac{\partial^2 y}{\partial x_2 \partial x_2} & \dots & \frac{\partial^2 y}{\partial x_2 \partial x_d} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 y}{\partial x_d \partial x_1} & \frac{\partial^2 y}{\partial x_d \partial x_2} & \dots & \frac{\partial^2 y}{\partial x_d \partial x_d} \end{bmatrix}. \quad (23)$$

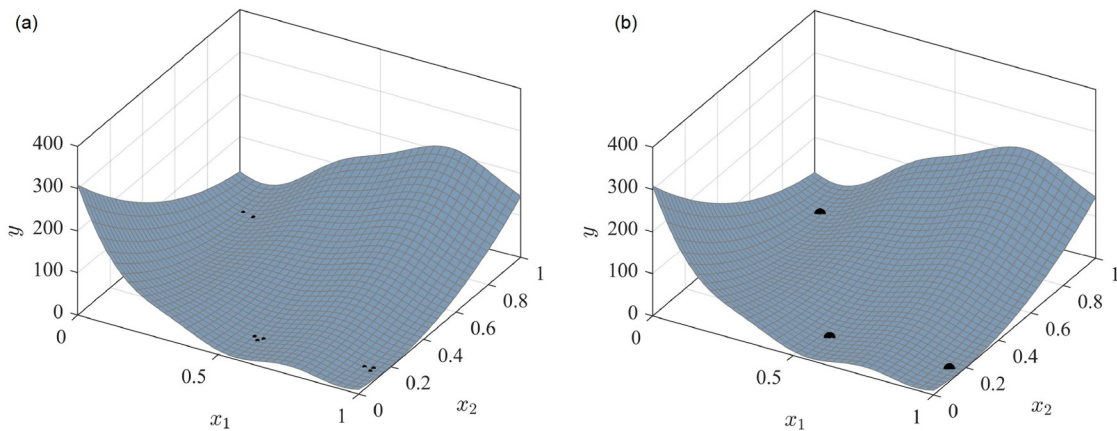


图5 (网络版彩图)二维Brannin函数极值点潜在探索示意图. (a) 局部极值点; (b) 极值点聚类

Figure 5 (Color online) Schematic diagram illustrating potential exploration of extrema points in two-dimensional Brannin functions. (a) Local extrema points; (b) clustering of extrema points.

根据Co-RBF的预测方程, 可得到任意点 $\mathbf{x}$ 处的预测值表达式:

$$\hat{y}(\mathbf{x}) = \rho \cdot \sum_{i=1}^{n_L} \lambda_i \varphi(\|\mathbf{x} - \mathbf{x}_{L,i}\|) + \sum_{i=1}^{n_H} \lambda_{d,i} \varphi(\|\mathbf{x} - \mathbf{x}_{H,i}\|). \quad (24)$$

该径向基函数表达式的二阶连续可微形式为

$$\begin{aligned} \mathbf{H}(\hat{y}) &= \frac{\partial^2 \hat{y}(\mathbf{x})}{\partial \mathbf{x}^2} \\ &= \rho \cdot \sum_{i=1}^{n_L} \frac{\lambda_i}{r_i(\mathbf{x})} \left[\varphi'(r_i) \cdot I + \left(\varphi''(r_i) - \frac{\varphi'(r_i)}{r_i(\mathbf{x})} \right) \right] \\ &\quad \times (\mathbf{x} - \mathbf{x}_{L,i})^T \\ &\quad + \sum_{i=1}^{n_H} \frac{\lambda_{d,i}}{r_i(\mathbf{x})} \left[\varphi'(r_i) \cdot I + \left(\varphi''(r_i) - \frac{\varphi'(r_i)}{r_i(\mathbf{x})} \right) \right] \\ &\quad \times (\mathbf{x} - \mathbf{x}_{H,i})^T, \end{aligned} \quad (25)$$

其中, $r_i(\mathbf{x}) = \|\mathbf{x} - \mathbf{x}_i\|$, 一阶偏导 $\varphi'(r_i) = \partial \varphi / \partial r_i$, 二阶偏导 $\varphi''(r_i) = \partial^2 \varphi / \partial r_i^2$. 由于本文采用的是MQ基函数 $\varphi(r_i) = (r_i^2 + c^2)^{1/2}$, 因此 $\varphi'(r_i) = r_i(r_i^2 + c^2)^{-1/2}$, $\varphi''(r_i) = c^2(r_i^2 + c^2)^{-3/2}$.

Hessian矩阵的特征值可以表征曲面模型的主曲率, 定义平均曲率指标:

$$K(\mathbf{x}) = \frac{1}{d} \sum_{i=1}^d |m_i(\mathbf{x})|, \quad (26)$$

其中, $m_1, m_2, \dots, m_d$ 为径向基函数模型的Hessian矩阵的 d 个特征值.

在获取模型的局部曲率信息后, 一种常用的策略是选取曲率最大点作为更新样本. 然而, 若单纯依赖曲率准则, 新增样本极易聚集于少数几个局部高非线性区域, 导致样本分布严重不均. 这种聚集现象不仅可能引发插值矩阵的病态问题, 影响数值稳定性, 还会因过分关注已知区域而造成对设计空间其他区域的探索不足, 降低模型的全局预测能力. 为克服上述局限, 引入样本间约束机制以增强空间填充特性. 计算候选点到现有样本集的最短欧氏距离, 对过度聚集行为施加惩罚, 以维持数值稳定性并确保对未知区域的有效探索.

在此基础上, 构建一个基于强化学习的自适应决策框架. 获取状态信息后, 智能体需选择具体采样的策略. 定义动作空间为5种曲率与距离的权重组合

$a = (w_c, w_d)$, 每一个动作代表一种特定的采样偏好策略. 随后, 系统基于 ε 贪心策略进行策略选择:

$$\text{action} = \begin{cases} \text{random}(1, N_a), & \varepsilon, \\ \arg\max_a Q(s, a), & 1 - \varepsilon, \end{cases} \quad (27)$$

其中, ε 代表探索率, 初始赋值0.8, 鼓励随机探索, 而随着学习过程的循环迭代, 探索率逐渐降低, 更多依赖学习到的知识. 这种探索与利用之间的平衡确保逐渐收敛到最优动作策略.

选定动作后, 系统并非直接采样, 而是遍历所有候选点, 执行基于当前权重下的采样动作计算, 对每个候选点计算其综合得分:

$$S(x) = w_c \times d_{\min}(x) + w_d \times C(x). \quad (28)$$

接着系统选择最优点 $x^* = \arg\max_x S(x)$ 进行高精度仿真, 并更新样本集与Co-RBF模型. 模型更新后, 智能体依据获得的奖励进行学习. 奖励函数的设计旨在全面评估所选的权重组合对模型整体精度提升的有效性. 本文的奖励函数计算基于三个关键精度预测指标在模型更新前后的相对改进率:

$$\begin{cases} I_{R^2} = \frac{R^2 - R'^2}{1 - R'^2}, \\ I_{\text{RMAE}} = \frac{\text{RMAE} - \text{RMAE}'}{\text{RMAE}}, \\ I_{\text{NRMSE}} = \frac{\text{NRMSE} - \text{NRMSE}'}{\text{NRMSE}}. \end{cases} \quad (29)$$

随后对改进率进行平滑归一化处理, 再通过加权融合形成一个综合奖励信号, 以此驱动智能体学习在不同建模阶段应采取何种采样偏好, 奖励函数形式如式(30)所示:

$$R(t) = k_{R^2} \cdot I_{R^2} + k_{\text{RMAE}} \cdot I_{\text{RMAE}} + k_{\text{NRMSE}} \cdot I_{\text{NRMSE}}. \quad (30)$$

奖励值综合反映了所选权重组合对单保真度和多保真度模型精度提升的综合效益, 为策略优化提供准确的反馈信号. 智能体随后进行Q表更新与优化, 依据如下规则:

$$\begin{aligned} Q(s, a) &\leftarrow Q(s, a) + \alpha \cdot R(s, a) \\ &\quad + \alpha \cdot \gamma \max_{a'} Q(s', a') - Q(s, a). \end{aligned} \quad (31)$$

此循环持续进行直至满足预设的收敛条件, 最终输出由最优采样策略引导构建的最优Co-RBF多可信度模型.

3.4 多可信度序贯建模流程

综合前文提出的AES初始采样策略与RL自适应序贯更新准则, 本节构建了基于强化学习指导的自适应多可信度序贯建模框架(RL based adaptive sampling cooperative RBF, RLASC-RBF). 该框架的核心思想在于协同利用不同可信度信息与计算资源: 首先, 充分发挥低可信度模型廉价、快速的优势, 通过智能特征提取完成对高可信度样本初始空间布局的全局引导; 随后在序贯更新阶段, 引入强化学习智能体感知模型状态, 自适应决策采样偏好, 从而驱动Co-RBF模型朝着全局精度提升的方向高效迭代优化. 具体实现步骤如下.

(1) 首先采用基于最大最小距离准则的最优超立方采样方法, 在设计空间内生成具有良好空间填充特性的低可信度样本集 $\mathbf{X}_L$.

(2) 基于样本集 $\mathbf{X}_L$ 运行低可信度模型, 获取系统响应值集合 $y_L(\mathbf{x}_j)$, 据此训练并构建低可信度RBF预测模型 $\hat{y}_L$. 该模型将作为探索设计空间、提取关键特征快速近似.

(3) 利用构建好的低可信度RBF预测模型进行潜在探索采样. 利用Multistart搜索策略识别一组分布于随机起始点附件的局部极值候选点, 并基于Mean Shift聚类算法对候选点集进行分类, 划为 k 个簇. 在每个簇内, 选取低可信度预测值最小的点作为该区域的代表性局部极值点, 从而获得包含 k 个局部极值点的特征集.

(4) 将步骤(3)中的 k 个局部极值点直接作为高可信度初始样本集. 若提供的极值点数量不满足最小样本量要求, 则依据最大最小距离准则在未被特征点覆盖的设计空间区域补充采样点, 以确保初始样本集的空间支撑能力.

(5) 整合高、低可信度数据集 $\mathbf{X}_H$ 和 $\mathbf{X}_L$, 分别调用对应精度的仿真模型计算响应值, 构建初始多可信度数据库 $\{\mathbf{X}_H, y_H, \mathbf{X}_L, y_L\}$.

(6) 基于融合数据集构建Co-RBF代理模型. 在此过程中, 以最小化LOOCV误差及偏差RBF的平滑性能为联合目标, 利用CMA-ES算法对模型缩放因子及形状参数等超参数进行寻优求解.

(7) 根据预设准则判定迭代是否终止. 收敛判据通常依据模型精度阈值设定; 在工程资源受限场景下, 也

常以最大允许采样数作为硬性约束. 若满足上述任一终止条件, 则跳转至步骤(9); 否则, 进入步骤(8)继续更新.

(8) 计算当前模型在全设计空间候选点集上的局部曲率场、欧氏距离及精度指标, 智能体基于当前状态通过Q-learning策略选择最优权重动作. 根据动作对应的曲率与距离权重, 计算所有候选点的综合得分, 选定得分最高的点作为本轮的新增样本点 x_{new} , 将其纳入高可信度样本集 $\mathbf{X}_H$ 后, 返回步骤(5)更新模型. 同时, 根据精度提升反馈计算奖励值, 在线更新Q表以优化后续决策;

(9) 当满足终值条件时, 循环结束. 输出最终满足精度与效率要求的Co-RBF多可信度预测模型, 该模型可替代昂贵的高可信度仿真, 用于后续的优化分析等任务.

整体建模流程如图6所示.

4 仿真实例

为验证前文所构建的基于强化学习指导的自适应多可信度序贯建模框架(RLASC-RBF)的有效性, 本文选取典型吸气式高超声速巡航飞行器作为仿真对象, 开展耦合建模与轨迹优化仿真分析.

高超声速飞行器设计高度依赖精确的气动力/热数据, 而高可信度数值仿真(如CFD)计算成本极其昂贵, 在高性能计算工作站(AMD Ryzen 9 5950X, 16核心)环境下, 单个状态点的CFD计算收敛耗时约20 h; 而基于机理的工程估算方法单点计算通常耗时1 s以内. 两者在计算效率上的巨大差异导致单纯依赖CFD数据进行多学科设计优化与全包线飞行性能分析难以实现. 因此, 在有限高精度样本下构建兼具高精度与高一致性的气动代理模型, 对提升仿真与优化效率具有重要意义.

仿真验证从两个层面展开: 首先, 通过静态建模仿真, 从数值逼近的角度出发, 重点评估序贯模型在宽飞行包线内对气动系数的拟合精度. 特别关注模型对由激波位置突变、流动分离等物理现象所引发的局部极值点及强非线性区域的捕捉能力, 从函数逼近角度量化模型的预测误差与空间适应性; 其次, 将生成的代理模型嵌入纵向运动方程进行动力学仿真, 构建用于轨迹优化的高保真动力学仿真环境, 开展典型的燃油

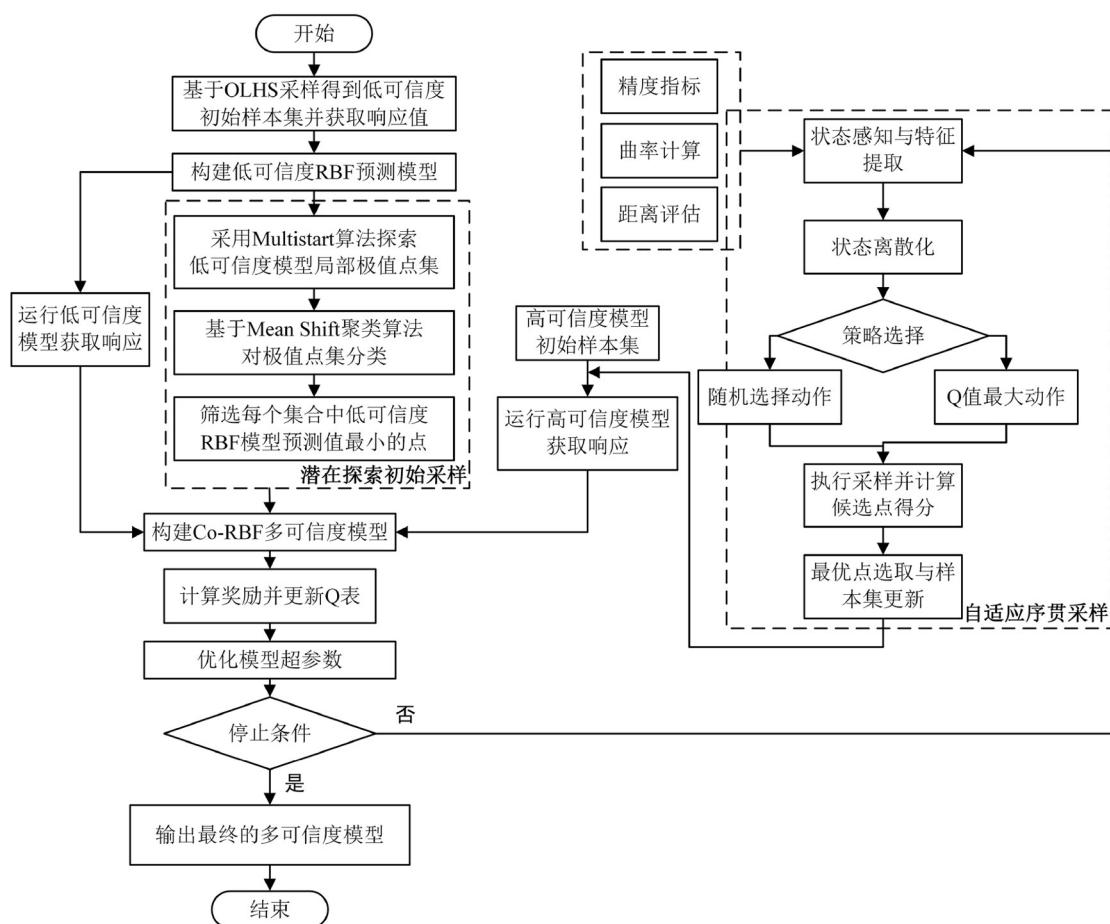


图6 基于强化学习的自适应序贯建模流程图

Figure 6 Adaptive sequential modeling flowchart based on reinforcement learning.

最省定动压爬升轨迹优化任务研究。采用高精度数值优化方法(如Radau伪谱法)分别求解不同模型精度下的最优飞行轨迹,通过对比分析燃油消耗量、爬升时间以及加速度状态量的演化规律,从系统任务性能与动态响应一致性的角度,综合评价模型在长时域、非线性动力学优化过程中的动静态一致性表现。

4.1 静态建模仿真实验设计

针对吸气式高超声速巡航飞行器纵向气动特性的序贯建模问题,开展仿真实验以验证RLASC-RBF框架的有效性。具体选取马赫数 Ma 和攻角 α 作为设计变量,分别构建纵向升力系数 C_L 、阻力系数 C_D 以及俯仰力矩系数 C_M 的代理模型。

仿真中,将设计空间定义为 $Ma \in [5, 8]$, $\alpha \in [-6^\circ, 10^\circ]$

的矩形区域,所有的初始与序贯采样点均处于这一矩形空间内。需要同时对三种气动系数进行建模,并且这三种气动非线性程度不一致,最终目的是尽可能使用较少的数据获得准确的高精度气动系数模型。

在采样策略与参数配置上,对三个气动系数统一使用相同的低可信度初始样本集,通过最优拉丁超立方采样(OLHS)生成50个样本,并选用Multi-quadric核函数构建初始低可信度RBF模型。由于高精度CFD分析计算耗时较长,并且考虑到工程实际中一般对实验总预算有限制,故设置初始样本数为5,最大增点样本数为10,即达到总成本为15次高精度CFD仿真时,RLASC-RBF序贯建模结束,以此模拟实际工程中的预算限制。在构建多可信度Co-RBF模型时,其偏差RBF部分同样选用MQ核函数进行校准,以保障模型的逼

近能力与数值稳定性.

4.2 静态建模结果与分析

4.2.1 建模精度与训练过程曲线

图7展示了基于RLASC-RBF方法构建的升力、阻力及俯仰力矩系数预测响应曲面, 并叠加显示了高可信度样本的空间分布.

具体分析可知, 升力系数预测曲面清晰反映了飞行器在超声速至高超声速段的典型准线性特征, 且随马赫数增加逐渐呈现马赫数无关性; 由于该曲面整体拓扑平坦、曲率变化微弱, 序贯准则中的距离约束项占据主导地位, 致使新增样本在设计空间内呈现高度均匀的分布. 阻力系数模型准确捕捉了其随攻角变化的抛物线非线性特征及随马赫数增加的衰减趋势, 序贯样本显著聚集于高曲率的小攻角区域, 以精确修正局部非线性误差. 俯仰力矩系数的序贯样本主要集中于高马赫数区域, 尽管整体呈现上升态势, 但在高马赫数飞行条件下, 力矩系数随攻角变化的非线性程度相比低马赫数段更为显著, 促使RLASC-RBF算法自适应地在该区域加密采样, 以捕捉高速条件下更复杂的气动特性.

为了分析RLASC-RBF方法在序贯建模过程中的动态特性, 本文以阻力系数为例, 研究强化学习智能体的训练过程及模型精度的演化规律如图8所示.

图8(a)展示了智能体在序贯选点过程中的学习曲线. 图中左轴表示累积奖励值, 右轴表示探索率. 随着迭代次数的增加, 探索率按照预设的规律不断衰减, 智能体逐渐减少随机探索, 更多地利用Q表中的先验知识进行决策. 在此过程中, 累积奖励呈现出显著的单调上升趋势. 这表明智能体在每一次迭代中都能有效地识别出对模型精度提升贡献最大的采样点, 验证了采样策略的有效性与稳定性.

图8(b)量化了模型精度随迭代次数的收敛情况. 在初始的几次迭代中, 由于智能体优先探索了全局极值点及关键非线性区域, NRMSE出现了陡峭的下降, 从初始的0.12迅速降低至0.04以下. 随后模型进入精细修正阶段, 误差平稳下降. 最终在第10次迭代时, 模型达到了较高的预测精度, 最终归一化均方根误差收敛至0.0090, 同时决定系数提升至0.9990.

为了进一步验证RLASC-RBF方法的有效性, 使用决定系数 R^2 、归一化均方根误差NRMSE和最大绝对

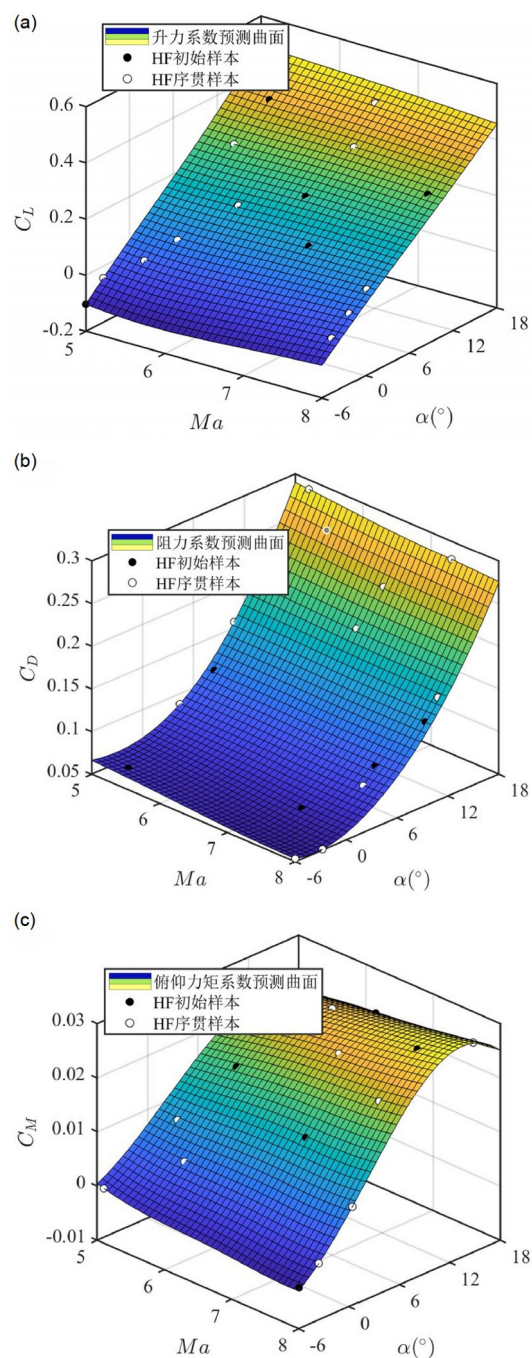


图7 (网络版彩图)气动系数预测曲面与自适应序贯样本分布视图. (a) 升力系数; (b) 阻力系数; (c) 俯仰力矩系数

Figure 7 (Color online) Predicted aerodynamic coefficient surfaces and adaptive sequential sample distribution views. (a) Lift coefficient; (b) drag coefficient; (c) pitch moment coefficient.

误差RMAE定量地描述气动系数的预测精度. 表1列出了不同建模方法在大量高精度验证样本上的三种气动

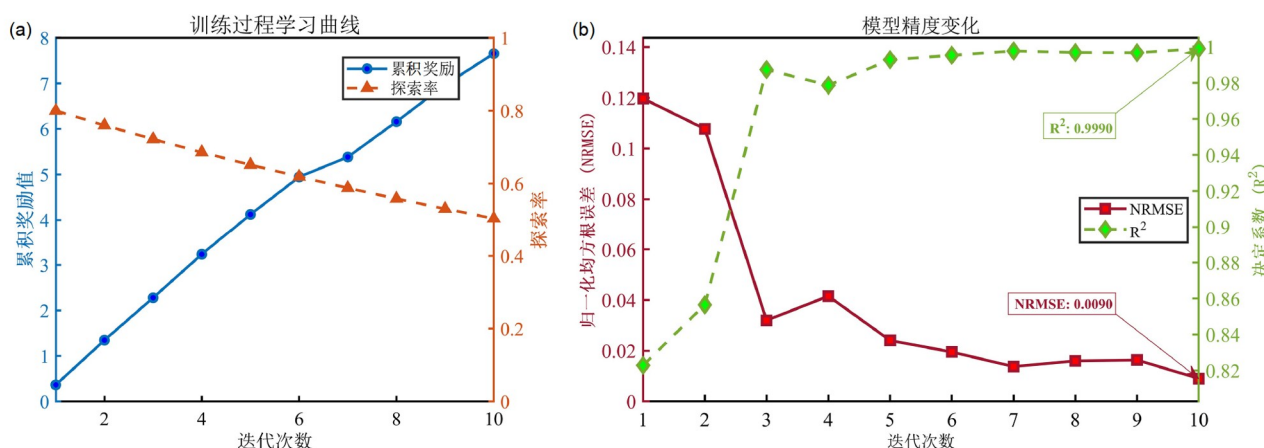


图8 (网络版彩图)智能体训练过程与模型精度变化曲线。(a) 训练过程学习曲线; (b) 模型精度曲线

Figure 8 (Color online) Training process of intelligent agents and model accuracy variation curve. (a) Learning curve during training process; (b) model accuracy curve.

系数预测精度。

其中RLASC-RBF为本文提出的基于强化学习的自适应序贯建模方法, HF-RBF为仅使用高精度样本的RBF建模方法, LF-RBF为仅使用工程估算数据的RBF建模方法。

精度对比验证了本文所提方法在三个气动系数预测问题上的优越性, 其全局与局部精度均为最优。首先, 所有系数预测结果的决定系数 R^2 均近乎为1, 展现了与真实数据的高度一致性; 其次, 在全局误差上, 升力系数与阻力系数的归一化均方根误差(NRMSE)均小于0.6%, 俯仰力矩系数小于1%, 预测性能较好。在衡量局部精度的相对最大绝对误差(RMAE)指标上, 三者均小于0.1, 在三种方法中性能表现最佳。

表1 不同建模方法下的气动系数预测精度

Table 1 Prediction accuracy of aerodynamic coefficients under different modelling approaches

气动系数	精度指标	RLASC-RBF	HF-RBF	LF-RBF
C_L	R^2	0.9989	0.9964	0.5624
	NRMSE	0.59%	0.96%	16.45%
	RMAE	0.056	0.158	1.489
C_D	R^2	0.9984	0.9932	0.7649
	NRMSE	0.48%	0.64%	13.87%
	RMAE	0.0658	0.1648	1.4895
C_M	R^2	0.9987	0.9969	0.7451
	NRMSE	0.96%	1.21%	13.21%
	RMAE	0.0844	0.1058	1.7864

4.2.2 消融实验分析

为了进一步验证本文提出的RLASC-RBF框架中Co-RBF多可信度融合机制与强化学习采样策略各自的贡献程度, 本节开展了消融实验分析。

实验设置了三组对照模型, 针对相同的高超声速飞行器气动数据进行建模测试, 每组模型均独立运行10次以消除随机性影响, 取均值作为最终评价指标。三组模型设置如下。

(1) RL-CoRBF: 本文提出的基于强化学习指导的多可信度序贯融合建模方法。

(2) RL-HFRBF: 仅利用强化学习策略对高可信度数据进行采样建模, 不引入低可信度数据辅助。

(3) ASC-CoRBF: 采用Co-RBF多可信度架构, 但序贯采样策略移除强化学习模块, 仅基于加权距离与曲率的传统自适应采样策略。

实验统计结果如表2所示, 记录了三种方法在同一测试集上的相关精度指标。

对比RL-CoRBF与RL-HFRBF的实验结果可知, 在同等高精度样本预算下, 引入低可信度数据进行偏差校正对模型精度的提升起到了决定性作用。相较于仅依赖高精度数据的RL-HFRBF模型, 采用融合架构的RL-CoRBF模型使得预测误差显著降低, NRMSE相对降低了37.7%; 同时模型拟合能力增强, 决定系数 R^2 相对提升了24.6%。这一结果证明了多可信度融合架构通过利用低成本数据的全局趋势信息, 有效弥补了高维空间中高精度样本稀疏的问题。

表 2 消融实验结果对比统计

Table 2 Statistical comparison of melting experiment results

模型架构	NRMSE	R^2
RL-CoRBF	0.0929	0.8770
RL-HFRBF	0.1492	0.7036
ASC-CoRBF	0.1401	0.7417

对比RL-CoRBF与ASC-CoRBF的实验结果可知,在相同的多可信度架构下,采样策略的智能化程度直接决定了样本的利用效率.相比于基于静态权重的传统ASC采样策略,本文提出的RL智能体策略能够动态感知模型状态并自适应调整采样偏好,从而进一步挖掘了模型性能潜力.分析可知,RL策略的加入使NRMSE相对降低了33.7%,决定系数 R^2 相对提升了18.2%.结果表明强化学习在序贯决策过程中,能够比传统启发式规则更高效地平衡探索与开发,从而实现更优的局部非线性捕捉.

综上所述,Co-RBF架构与RL策略在降低模型误差和提升模型精度方面均有显著贡献.在模型架构融合与策略优化的协同下,提升了数据利用率与模型收敛速度.

4.3 动力学仿真试验设计

为验证气动建模精度在动力学层面的传导效应,本节设计了定动压爬升任务场景.该方案旨在满足冲压发动机对进气道流场的高效稳定需求.综合考虑飞行器结构强度极限、热防护系统(TPS)承温能力以及燃烧室稳燃压力边界,设定飞行包线的动压约束走廊的上限为100 kPa,该约束主要由机体结构载荷极限与热防护系统(TPS)的承温能力共同决定,旨在规避结构过载与过热风险;动压下限设定为50 kPa,此约束与燃烧室稳定燃烧所需的最低压力水平相关,以防止因进气压缩不足导致的燃烧不稳定或熄火.

基于上述物理约束,本文构建了如图9所示的飞行高度-马赫数标称剖面.

现有轨迹优化理论表明,吸气式高超声速飞行器沿最大允许动压边界爬升有利于减少重力损失,从而获得最优燃油经济性.为了在最优性能与飞行安全之间达到平衡,并考虑到模型不确定性、大气扰动等因素,本文在动压上限处预留5%的安全裕度,选取95 kPa作为标称定动压爬升条件.设定高超声速飞行

器初始质量为280 t,起始马赫数为4,终端马赫数为6,高度由选取的动压确定,为保证爬升段与巡航段的平滑衔接,起始航迹角和终端航迹角均设置为0.仿真中,轨迹优化的最省燃料性能指标为 $\min J = -m_f$,相关变量的约束条件如表3所示.

4.4 动力学仿真结果与分析

基于前文构建的三种气动数据驱动模型—低精度模型(low-fidelity, LF)、高精度模型(high-fidelity, HF)以及融合模型(multi-fidelity, MF),在定动压约束下进行燃料最省的定动压爬升轨迹优化,Radau伪谱法求解所得燃油消耗量与爬升时间结果如表4所示.

从表4可以看出,LF模型在任务性能评估上表现出显著偏差,LF模型的轨迹具有更短的爬升时间与更小的燃油消耗量,这与定动压条件下LF模型所获加速度大于HF模型有关,从而使优化算法得到一条在物理

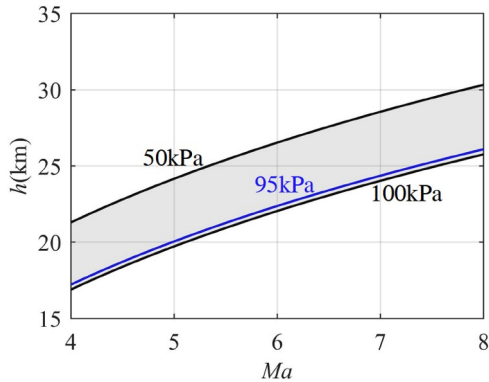


图 9 (网络版彩图)定动压爬升剖面
Figure 9 (Color online) Static and dynamic pressure climb profile.

表 3 相关变量的约束条件

Table 3 Constraints on relevant variables

变量	起始值	终端值	路径约束
高度 h (m)	17224	22479	[17224, 22479]
速度 v (m/s)	1180	1780	[1180, 1780]
航迹倾角 γ (°)	0	0	[-5, 5]
质量 m (t)	280	N/A	[180, 280]
攻角 α (°)	N/A	N/A	[-2, 8]
攻角变化率 $\dot{\alpha}$ (°/s)	N/A	N/A	[-0.01, 0.01]
燃油当量比 ϕ (A.U.)	N/A	N/A	[0, 1]
动压 q (kPa)	95	95	95

上难以实现的最优轨迹. 相比之下, MF模型的轨迹优化结果有效地修正了LF模型的偏差, 并且与HF模型高度吻合, 两者在燃油消耗和爬升时间上的相对差异仅为0.1%, 这验证了数据驱动方法的准确性. 通过伪谱法计算得到不同模型的最优爬升轨迹状态量随时间变化情况如图10所示.

对比发现, MF模型的最优轨迹状态量十分接近HF

表 4 不同数据驱动模型的最优爬升轨迹燃油消耗量和爬升时间

Table 4 Optimal climb trajectories for different data-driven models fuel consumption and climb time

数据驱动模型	燃油消耗量 (t)	时间 (s)
LF	13.50	129.3
HF	17.85	167.1
MF	17.87	166.9

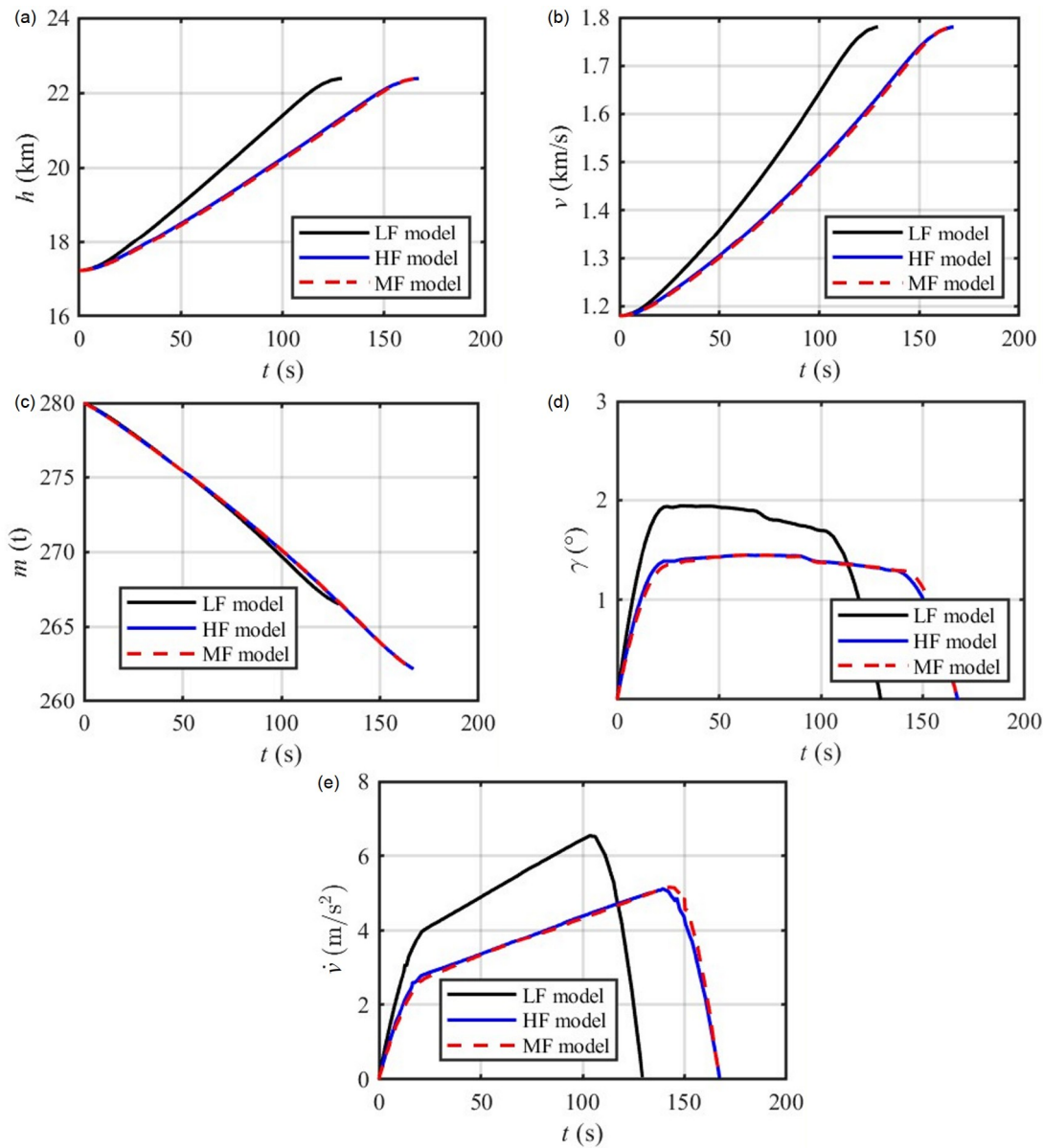


图 10 (网络版彩图)不同数据驱动模型的最优爬升轨迹状态量随时间变化关系. (a) 高度-时间; (b) 速度-时间; (c) 质量-时间; (d) 航迹倾角-时间; (e) 加速度-时间

Figure 10 (Color online) Optimal climb trajectory state variables for different data-driven models over time. (a) Altitude versus time; (b) velocity versus time; (c) mass versus time; (d) trajectory inclination versus time; (e) acceleration versus time.

表 5 不同建模方法的计算成本与精度对比

Table 5 Optimal climb trajectories for different data-driven models fuel consumption and climb time

建模方法	建模耗时 (s)	决策耗时 (s)	总耗时 (s)	NRMSE	R^2 (%)
ASC-RBF	0.5179	2.7125	35.1769	0.0125	0.9896
RLASC-RBF	0.4267	2.9031	33.3024	0.0090	0.9990

模型, 验证了多可信度建模方法的有效性. 从图10(e)加速度变化曲线可以看出, LF模型在爬升前中期获得了更大的加速度, 这使得LF模型具有比HF和MF模型更短的爬升时间, 从而产生了更小的燃油消耗.

在整个爬升任务周期内, MF模型生成的速度、高度及航迹角等状态变量曲线与高精度基准模型实现了高度重合, 精确复现了飞行器在复杂机动过程中的动态响应特征. 得益于RLASC-RBF框架对关键非线性区域的精准校准, MF模型能够利用极少量的高可信度样本, 有效修正低可信度模型的动力学趋势偏差, 从而在大幅降低计算成本的同时, 保证了对飞行器动态性能预测的高保真度.

4.5 算法性能与成本效率分析

为了量化评估强化学习模块引入的计算开销及其带来的性能增益, 本节采用了基于Co-RBF平均曲率度量的自适应多可信度序贯建模(ASC-RBF)作为对照基准. 该基准方法采用与本文相同的Co-RBF建模架构, 但在序贯采样阶段仅依据最大曲率与最大距离的静态加权准则进行增点.

在相同的计算环境下, 两种方法均从相同的初始样本集出发进行迭代, 直至达到最大预设的高精度样本数, 测试数据如表5所示.

由表5可以看出, 引入RL模块后, 单步决策仅增加了0.19 s的额外时间代价. 这表明本文采用的基于Q表的轻量化强化学习策略并未显著增加计算负担, 保证了算法的实时性. 由于RL智能体能够动态平衡探索与利用, 避免了在低价值区域的无效采样, 使得模型收敛

更加高效.

实验结果表明, 本文方法的总算法耗时反而比基准方法减少了1.16 s. 在耗时更短的前提下, RLASC-RBF方法的NRMSE相较于ASC-RBF降低了28%, 决定系数 R^2 得到提升. 这说明强化学习模块成功引导了高精度样本分布在气动非线性最显著的区域, 具有更强的数据挖掘能力.

5 结论与讨论

本文针对高超声速飞行器设计中高精度数据获取成本高昂、多可信度数据融合效率低的问题, 提出了一种基于强化学习的序贯融合建模方法(RLASC-RBF). 首先建立多可信度代理模型架构, 设计了一种主动探索策略, 创新性地提出了一种模型数据自适应序贯迭代方法, 实现了在有限高可信度样本下的高超声速飞行器多源模型数据的高效融合.

理论分析与仿真实例表明: 偏差Co-RBF模型通过偏差校正与粗糙惩罚准则, 提升了多源数据的融合精度与稳定性; 初始采样策略有效利用低可信度数据的全局趋势, 引导高精度样本聚焦于关键非线性区域; 强化学习框架能够依据模型实时状态动态平衡探索与利用, 在提升全局拟合精度的同时加速局部收敛, 且能满足策略的实时性调整. 所提出的方法在高超声速飞行器气动数据库预测中取得了优于传统方法的精度, 在燃油最省爬升轨迹优化中表现出与高精度模型高度一致的动静态一致性, 验证了融合建模方法的有效性与可靠性.

参考文献

- 1 Simmons B M, Gresham J L, Woolsey C A. Nonlinear dynamic modeling for aircraft with unknown mass properties using flight data. *J Aircraft*, 2023, 60: 968–980
- 2 Snyder S, Wignall T J, Green J S, et al. AeroFusion: Data fusion and uncertainty quantification for entry vehicles. In: Proceedings of the AIAA Science and Technology Forum and Exposition. National Harbor, 2023. 1182

- 3 Dias Ribeiro M, Stradtner M, Bekemeyer P. Unsteady reduced order model with neural networks and flight-physics-based regularization for aerodynamic applications. *Comput Fluids*, 2023, 264: 105949
- 4 Sacks J, Welch W J, Mitchell T J, et al. Design and analysis of computer experiments. *Stat Sci*, 1989, 4: 409–423
- 5 Forrester A I J, Keane A J. Recent advances in surrogate-based optimization. *Prog Aerosp Sci*, 2009, 45: 50–79
- 6 Jiang P, Zhou Q, Shao X. Surrogate Model-Based Engineering Design and Optimization. Singapore: Springer, 2019
- 7 Zhang M, Bartoli N, Jungo A, et al. Enhancing the handling qualities analysis by collaborative aerodynamics surrogate modelling and aero-data fusion. *Prog Aerosp Sci*, 2020, 119: 100647
- 8 Rataczak J A, Boyd I D, McMahon J W. Surrogate models for hypersonic aerothermodynamics and aerodynamics using Gaussian process regression. In: Proceedings of the AIAA Science and Technology Forum and Exposition (AIAA SciTech Forum). Orlando, 2024. 0461
- 9 Balabanov V, Grossman B, Watson L, et al. Multifidelity response surface model for HSCT wing bending material weight. In: Proceedings of the 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization. St. Louis, 1998. 4804
- 10 Kleijnen J P C. Kriging metamodeling in simulation: A review. *Eur J Oper Res*, 2009, 192: 707–716
- 11 Cortes C, Vapnik V. Support-vector networks. *Mach Learn*, 1995, 20: 273–297
- 12 Akhtar T, Shoemaker C A. Multi objective optimization of computationally expensive multi-modal functions with RBF surrogates and multi-rule selection. *J Glob Optim*, 2016, 64: 17–32
- 13 Mengistu T, Ghaly W. Aerodynamic optimization of turbomachinery blades using evolutionary methods and ANN-based surrogate models. *Optim Eng*, 2008, 9: 239–255
- 14 Raissi M, Perdikaris P, Karniadakis G E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J Comput Phys*, 2019, 378: 686–707
- 15 Reuther J, Jameson A, Farmer J, et al. Aerodynamic shape optimization of complex aircraft configurations via an adjoint formulation. In: Proceedings of the 34th Aerospace Sciences Meeting and Exhibit. Reno, 1996. 94
- 16 Zanjani Foumani Z, Bostanabad R. Cost-aware Bayesian optimization with automatic stop condition under multi-fidelity constraints and data. *J Mech Des*, 2026, 148: 021711
- 17 Zhang Y, Gong Z, Zhou W, et al. Multi-fidelity surrogate modeling for temperature field prediction using deep convolution neural network. *Eng Appl Artif Intell*, 2023, 123: 106354
- 18 Peherstorfer B, Willcox K, Gunzburger M. Survey of multifidelity methods in uncertainty propagation, inference, and optimization. *SIAM Rev*, 2018, 60: 550–591
- 19 Helton J C, Davis F J. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliab Eng Syst Saf*, 2003, 81: 23–69
- 20 Bouzid T, Grich A, Naboulsi A, et al. Optimizing eriochrome black T adsorption through *in-situ* polymerization of poly(aniline-co-formaldehyde) on biochar: Multivariate approach using full factorial design, box-behnken, artificial intelligence, and DFT. *Sep Purif Tech*, 2024, 351: 128107
- 21 Liu H, Ong Y S, Cai J. A survey of adaptive sampling for global metamodeling in support of simulation-based complex engineering design. *Struct Multidisc Optim*, 2018, 57: 393–416
- 22 Mahmood S, Sun H, Ali Alhussan A, et al. Active learning-based machine learning approach for enhancing environmental sustainability in green building energy consumption. *Sci Rep*, 2024, 14: 19894
- 23 Jones D R, Schonlau M, Welch W J. Efficient global optimization of expensive black-box functions. *J Glob Optim*, 1998, 13: 455–492
- 24 Cheng J, Jiang P, Zhou Q, et al. A lower confidence bounding approach based on the coefficient of variation for expensive global design optimization. *Eng Comput*, 2019, 36: 830–849
- 25 Zhang Y, Neelakantan A, Park C, et al. Adaptive sampling with varying sampling cost for design space exploration. *AIAA J*, 2019, 57: 1032–1043
- 26 Zhang Y. Research on Kriging-based aerospace vehicle multidisciplinary surrogate modeling (in Chinese). Dissertation for Doctoral Degree. Changsha: National University of Defense Technology, 2019 [张熠. 基于Kriging的飞行器多学科近似建模方法研究. 博士学位论文. 长沙: 国防科技大学, 2019]
- 27 Beck J, Guillas S. Sequential design with mutual information for computer experiments (MICE): Emulation of a tsunami model. *SIAM ASA J Uncertain Quantif*, 2016, 4: 739–766
- 28 Shewry M C, Wynn H P. Maximum entropy sampling. *J Appl Stat*, 1987, 14: 165–170

- 29 Liu H, Xu S, Ma Y, et al. An adaptive Bayesian sequential sampling approach for global metamodeling. *J Mech Des*, 2016, 138: 011404
- 30 Li G, Aute V, Azarm S. An accumulative error based adaptive design of experiments for offline metamodeling. *Struct Multidisc Optim*, 2010, 40: 137–155
- 31 Aute V, Saleh K, Abdelaziz O, et al. Cross-validation based single response adaptive design of experiments for Kriging metamodeling of deterministic computer simulations. *Struct Multidisc Optim*, 2013, 48: 581–605
- 32 Xu S, Liu H, Wang X, et al. A robust error-pursuing sequential sampling approach for global metamodeling based on Voronoi diagram and cross validation. *J Mech Des*, 2014, 136: 071009
- 33 Zhou Q, Wang Y, Choi S K, et al. A sequential multi-fidelity metamodeling approach for data regression. *Knowl-Based Syst*, 2017, 134: 199–212
- 34 Yao W, Chen X, Luo W. A gradient-based sequential radial basis function neural network modeling method. *Neural Comput Applic*, 2009, 18: 477–484
- 35 Wei X, Wu Y Z, Chen L P. A new sequential optimal sampling method for radial basis functions. *Appl Math Comput*, 2012, 218: 9635–9646
- 36 Mackman T J, Allen C B. Investigation of an adaptive sampling method for data interpolation using radial basis functions. *Numer Meth Eng*, 2010, 83: 915–938
- 37 Lovison A, Rigoni E. Adaptive sampling with a lipschitz criterion for accurate metamodeling. *Commun Appl Ind Math*, 2011, 1: 110–126
- 38 Wang J R, Huang J H, Tang Y. Swarm intelligence capture-the-flag game with imperfect information based on deep reinforcement learning (in Chinese). *Sci Sin Tech*, 2023, 53: 405–416 [王健瑞, 黄家豪, 唐漾. 基于深度强化学习的不完美信息群智夺旗博弈. 中国科学: 技术科学, 2023, 53: 405–416]
- 39 Yang B, Chen B, Liu Y, et al. Gaussian process fusion method for multi-fidelity data with heterogeneity distribution in aerospace vehicle flight dynamics. *Eng Appl Artif Intell*, 2024, 138: 109228
- 40 Bolender M A, Doman D B. Nonlinear longitudinal dynamical model of an air-breathing hypersonic vehicle. *J Spacecraft R ockets*, 2007, 44: 374–387
- 41 Maughmer M, Ozoroski L, Straussfogel D, et al. Validation of engineering methods for predicting hypersonic vehicle control forces and moments. *J Guid Control Dyn*, 1993, 16: 762–769
- 42 Chen Z X, Chen W, Li C K, et al. Sequential detection based quickest Markov decision processes: Theory, algorithms, and applications (in Chinese). *Signal Process*, 2025, 41: 448–471 [陈祖旭, 陈巍, 李长坤, 等. 基于序贯检测的快速马尔可夫决策: 理论、方法及应用. 信号处理, 2025, 41: 448–471]
- 43 Gutmann H M. A radial basis function method for global optimization. *J Glob Optim*, 2001, 19: 201–227
- 44 Li Y F, Ren Q W, Wang Q M, et al. Fast inversion method of structural parameters based on PCE surrogate model and Bayesian optimization (in Chinese). *Sci Sin Tech*, 2022, 52: 928–940 [李翼飞, 任青文, 王启明, 等. 基于PCE代理模型和贝叶斯优化的结构参数快速反演方法. 中国科学: 技术科学, 2022, 52: 928–940]
- 45 Hansen N, Kern S. Evaluating the CMA evolution strategy on multimodal test functions. In: *Proceedings of the 8th International Conference on Parallel Problem Solving from Nature (PPSN VIII)*. Birmingham, 2004. 282–291
- 46 Morris M D, Mitchell T J. Exploratory designs for computational experiments. *J Stat Plan Infer*, 1995, 43: 381–402
- 47 Qian P Z G, Tang B, Wu C F J. Nested space-filling designs for computer experiments with two levels of accuracy. *Stat Sin*, 2009, 19: 287–300
- 48 Wu S H, Li Z X, Zhang F, et al. Algorithm for locating all local optimum based on fitness landscape analysis (in Chinese). *Syst Eng Electron*, 2020, 42: 346–355 [吴诗辉, 李正欣, 张发, 等. 基于适应值曲面分析的全部局部极值搜索算法. 系统工程与电子技术, 2020, 42: 346–355]
- 49 Wu K L, Yang M S. Mean shift-based clustering. *Pattern Recognit*, 2007, 40: 3035–3052
- 50 Clifton J, Laber E. Q-learning: Theory and applications. *Annu Rev Stat Appl*, 2020, 7: 279–301
- 51 Forrester A I J, Sobester A, Keane A J. *Engineering Design via Surrogate Modelling: A Practical Guide*. Hoboken: John Wiley & Sons, 2008

A sequential fusion modelling approach for hypersonic vehicles based on reinforcement learning

LIU YanBin, WANG XiJie, YANG Ben & CHEN BoYi

College of Astronautics, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

During the integrated design phase of hypersonic vehicles, performance analysis necessitates substantial high-fidelity model data to enhance design capabilities. However, acquiring such data proves prohibitively costly. A critical technical challenge lies in reducing the volume of high-fidelity data while maintaining reasonable accuracy metrics, thereby supporting iterative design of the vehicle's guidance and control systems. To address this, this paper proposes a reinforcement learning-based sequential fusion modelling approach for hypersonic vehicles, aiming to resolve the conflict between model accuracy and modeling costs during the iterative design process. Firstly, a Co-RBF multi-fidelity surrogate model architecture is established by employing a biased RBF kernel to design a roughness penalty criterion. Subsequently, an active exploration strategy based on response surface extrema points is proposed to sample the high-fidelity model data space of the hypersonic vehicle. Furthermore, based on reinforcement learning strategies, we design an adaptive sequential iterative method for model data, which achieves efficient fusion of multi-source model data for hypersonic vehicles while reducing computational costs. Finally, we demonstrate the effectiveness of the proposed methodology through multiple numerical and simulation examples.

hypersonic vehicles, hybrid modelling, sequential sampling, reinforcement learning

doi: [10.1360/SST-2025-0394](https://doi.org/10.1360/SST-2025-0394)