

DOI: 10.7511/jslx20240612002

# 基于多可信度代理模型的导弹气动数据生成技术研究

付朋真<sup>1</sup>, 陈江涛<sup>2</sup>, 张培红<sup>\*2</sup>

(1. 西南科技大学 计算机科学与技术学院, 绵阳 621010;

2. 中国空气动力研究与发展中心 计算空气动力研究所, 绵阳 621000)

**摘要:**代理模型是气动数据生成的新研究方向,传统的代理模型方法依赖于大量高精度仿真模型的样本点及其响应值来确保模型的精度,多可信度代理模型通过融合多层的高可信度和低可信度模型,能在保持一定精度的同时降低计算成本,对降低导弹研发周期有着重要意义。本文以 Co-Kriging 模型为代表,对不同数量的低可信度样本点对多可信度模型的影响与最佳的高低可信度样本点比例展开研究,并提出一种适用于气动数据的多可信度代理模型采样方法,该方法在利用高低可信度气动数据一一映射关系,在低可信度气动数据上使用 K-means 聚类算法获得训练用的高可信度气动数据对应位置。应用到导弹气动数据预测中及三种多可信度代理模型构建中,其中 Co-Kriging 模型综合预测效果最优,推荐高低可信度样本数比例为 1:4 与 1:3 之间。

**关键词:** 气动数据;多可信代理模型;高低可信度样本点比例;K-means 聚类算法

**中图分类号:** O355 **文献标志码:** A **文章编号:** 1007-4708(2025)05-0714-08

## 1 引言

气动特性的预测是导弹气动设计的核心工作之一。导弹的气动特性数据库的建立通常需要开展规模庞大和复杂的风洞试验和计算流体力学 CFD(Computational fluid dynamics)计算,成为影响研制经费和进度的重要因素<sup>[1]</sup>。随着外界环境的显著变化,研究人员发现,为了提升气动数据库建立的效率,降低研发成本,并缩短整体研发周期,必须寻求新的解决方案。在这一背景下,代理模型技术因其独特优势,迅速成为飞行器设计领域中备受瞩目的研究方向。

传统的代理模型方法依赖于大量高精度仿真模型的样本点及其响应值来确保模型的精度<sup>[2]</sup>,这通常导致计算量依旧巨大。特别是在气动分析这样的复杂系统中,不同物理简化原理的模型(如线性势流模型或欧拉模型)和网格的精细程度都会显著影响仿真分析的精度和计算时间。多可信度代理模型通过融合多层的高可信度和低可信度模型,能在保持一定精度的同时降低计算成本。

多可信度代理模型根据构建方法可以分为 4 类,分别是基于标度函数的多可信度代理模型、基于空间映射的多可信度代理模型、多可信度神经网络代理模型、Co-Kriging 多可信度代理模型<sup>[3]</sup>。基于标度函数的变可信度代理模型主要依靠高精度代理模型与低精度代理模型之间的标度函数(加法标度函数、乘法标度函数或混合标度函数)与低精度代理模型构成。文献[4,5]提出使用乘法标度函数进行多可信度建模的方法,能有效修正低可信度模型的预测值,使之更接近高可信度模型的精度。Wang 等<sup>[6]</sup>提出的基于最小二乘法的加法标度函数构造方法,通过提高模型的泛化能力来优化修正效果。然而,标度函数方法需要高低可信度模型的样本点嵌套,并且获取差值或比值数据的过程可能涉及大量高可信度模型的计算,导致计算成本高昂。基于空间映射的多可信度模型构造方法由 Bandler 等<sup>[7]</sup>提出,该方法主要根据高/低精度代理模型间的映射转换函数构建,适用于高/低精度代理模型的设计空间维度不同的情况。多可信度神经网络代理模型遵循一个递进的方法。首先基于低可信度样本建立

收稿日期:2024-06-12; 修改稿收到日期:2024-07-29.

基金项目:国家数值风洞工程项目(NNW);国家自然科学基金(52175214);国家自然科学基金委员会-中国工程物理研究院 NSAF 联合项目(U2230208)资助。

作者简介:张培红\*(1980-),男,博士,硕士生导师,副研究员(E-mail:zph2s@sina.com).

引用本文:付朋真,陈江涛,张培红. 基于多可信度代理模型的导弹气动数据生成技术研究[J]. 计算力学学报,2025,42(5):714-721.  
FU Pengzhen, CHEN Jiangtao, ZHANG Peihong. Research on missile aerodynamic data generation technology based on multi-credibility proxy model[J]. Chinese Journal of Computational Mechanics, 2025, 42(5):714-721.

一个初步的神经网络模型;接着将这一低可信度模型的预测值与高可信度样本数据相结合,共同作为高可信度神经网络的输入数据;最后优化模型参数,依据高可信度神经网络输出的预测值与其对应的高可信度样本标签值之间的差异来定义损失函数,并通过迭代训练的方式逐步调整模型的参数,以达到更准确的预测效果<sup>[8]</sup>。协同克里金(Co-Kriging)由 Kennedy 等<sup>[9]</sup>以 Kriging 模型为基础提出,通过引入高斯随机过程来建立高低可信度模型响应数据的差值模型,该方法不再需要样本点嵌套,从而大大增强了其在实际应用中的灵活性。韩忠华等<sup>[10]</sup>对 Co-Kriging 代理模型方法的改进,特别是在协方差矩阵的构建上,为该方法在气动系数分析等领域的应用提供了有力支持。其中 Co-Kriging 在于能够有效地整合高低可信度样本点,通过构建协方差矩阵来提高预测能力,并优化计算效率,成为一种有效的数据处理和预测工具,在各种复杂工程和科学问题中发挥着重要作用。

代理模型构建与采样方法密不可分,目前采样方法分为一次性采样<sup>[11-14]</sup>和自适应采样<sup>[15-17]</sup>,多可信度代理模型因涉及不同层次样本数据与样本点的嵌套,许多针对 PRS 模型<sup>[18]</sup>、Kriging 模型<sup>[19]</sup>、RBF 模型<sup>[20]</sup>等传统代理模型的采样方法并不适用。

本文针对 Co-Kriging 双层可信度代理模型,在数学算例中对比不同高低可信度样本数比例构建的 Co-Kriging 模型精度与拟合优度差距,找寻合适的高低可信度样本比例,并基于 K-means 聚类算法通过气动数据特点构建双层可信度模型采样方法,将 Co-Kriging 模型与其余两种多可信度代理模型一起应用到翼型测试和导弹气动数据预测中,对比不同变可信度代理模型和高低样本点比例在工程中的适用性。

## 2 多可信度代理模型

### 2.1 Co-Kriging 模型

Co-Kriging 模型是克里金法(Kriging)的一种改进算法,其结合了不同可信度数据,实现了多可信度模型响应信息的有效整合。在特定的随机过程,如固有平稳过程中,克里金法能够给出最优线性无偏估计 BLUP(Best Linear Unbiased Prediction),因此在地统计学中也称为空间最优无偏估计器(spatial BLUP)。Co-Kriging 模型利用大量计算量小的低可信度样本点及少量高可信度样本点构建模型进行优化,具有较强的工程意义。如在气动优化设计领域,采用 Co-Kriging 代理模型进

行优化设计,可以明显提高优化效率,降低总耗时,并改善优化结果。与单一的高可信度样本构建的 Kriging 模型相比,Co-Kriging 模型能够更有效地整合和利用不同可信度的数据,从而在保持模型精度的同时,降低建模的计算成本。

此外,Co-Kriging 模型还可以与机器学习算法结合使用,实现更加复杂和优化的空间数据分析。然而,克里金法包括 Co-Kriging 模型在缺乏数据的区域的准确性会比较差,因此在使用时需要原始数据的质量进行严格的评估和筛选,以确保插值结果的准确性。

假设某个存在  $m$  个气动数据设计变量,对高低可信度样本分别抽样:

$$\begin{cases} S_1 = [x_1^{(1)} & x_1^{(2)} & \cdots & x_1^{(n_1)}] \in \mathbf{R}^{n_1 \times m} \\ S_2 = [x_2^{(1)} & x_2^{(2)} & \cdots & x_2^{(n_2)}] \in \mathbf{R}^{n_2 \times m} \end{cases} \quad (1)$$

式中  $S_1$  为高可信度样本点  $x_1$  组成的高可信度样本集,  $S_2$  为低可信度样本点  $x_2$  组成的低可信度样本集,  $n_1$  和  $n_2$  分别为高低可信度样本集中样本点个数。其对应的目标函数的响应值为

$$\begin{cases} y_1 = [y_1^{(1)} & y_1^{(2)} & \cdots & y_1^{(n_1)}]^T \in \mathbf{R}^{n_1} \\ y_2 = [y_2^{(1)} & y_2^{(2)} & \cdots & y_2^{(n_2)}]^T \in \mathbf{R}^{n_2} \end{cases} \quad (2)$$

Co-Kriging 代理模型的预估值定义为

$$\hat{y} = \lambda_1^T y_1 + \lambda_2^T y_2 \quad (3)$$

式中  $\lambda_1$  和  $\lambda_2$  分别是高低可信度响应值的加权系数。假设存在两个随机静态过程分别与  $y_1$  和  $y_2$  相对应:

$$\begin{cases} Y_1 = \beta_1 + Z_1(x) \\ Y_2 = \beta_2 + Z_2(x) \end{cases} \quad (4)$$

整个设计空间中两个不同位置的随机变量之间的方差和协方差为

$$\begin{cases} \text{Cov}(Z(x_1^{(i)}), Z(x_1^{(j)})) = \sigma_1^2 R^{(11)}(x_1^{(i)}, x_1^{(j)}) \\ \text{Cov}(Z(x_2^{(i)}), Z(x_2^{(j)})) = \sigma_2^2 R^{(22)}(x_2^{(i)}, x_2^{(j)}) \\ \text{Cov}(Z(x_1^{(i)}), Z(x_2^{(j)})) = \sigma_1 \sigma_2 R^{(12)}(x_1^{(i)}, x_2^{(j)}) \end{cases} \quad (5)$$

在这两个随机静态过程  $Y_1(x)$  和  $Y_2(x)$  中,其各自的方差分别为  $\sigma_1^2$  和  $\sigma_2^2$ 。预估值是基于这两个随机过程的特性,通过式(6)得出

$$\hat{y}_1(x) = \varphi^T \beta + r^T(x) R^{-1}(\tilde{y}_s - F\beta) \quad (6)$$

### 2.2 加法标度函数修正模型

加法标度函数是一种在数学和工程领域中常用的工具,主要用于调整或修正模型的输出值,以便更准确地反映实际数据或系统行为。在基于加法标度函数修正的变可信度融合模型(AS-MFS)中,这个函数起着至关重要的作用。

具体来说,加法标度函数通过向低可信度模型

的输出添加一个标度项,来引入高可信度数据的影响。这个标度项通常是根椐高可信度数据和低可信度数据之间的差异来计算,旨在弥补低可信度模型在精度上的不足。通过这种方式,加法标度函数能够更好地融合不同可信度的数据,从而构建出更加精确和可靠的模型。在实际应用中,加法标度函数的形式和参数通常需要根据具体问题和数据进行调整和优化。这种方法的核心在于加法标度函数,其具体的数学表达式为

$$\gamma(x) = y_{hf}(x) - y_{lf}(x) \quad (7)$$

式中  $y_{hf}(x)$  为高可信度样本构建的高可信度模型,  $y_{lf}(x)$  为低可信度样本点构建的低可信度模型,  $\gamma(x)$  通过高低可信度模型对高可信度样本点的预测差值获得。最终加法标度模型的预测值表示为

$$\hat{y}_{hf}(x) = \hat{\gamma}(x) + \hat{y}_{lf}(x) \quad (8)$$

### 2.3 多可信度神经网络

多可信度神经网络是一种特殊的神经网络结构,其结合了不同可信度的数据源或模型,以提高预测和决策的准确性和可靠性。这种网络结构在多个领域中都有广泛的应用,特别是在需要处理不同精度、不同来源数据的情况下,其优势尤为突出。

在航空航天领域,多可信度神经网络得到了广泛的关注和应用。航空航天领域的研究往往涉及大量的复杂数据,这些数据可能来自不同的传感器、模拟器或实验设备,其精度和可信度可能各不相同。多可信度神经网络能够有效地融合这些不同来源的数据,提取出有用的信息,从而提高航空航天系统的性能。

多可信度神经网络也成为变可信度代理模型中重要一支。

假设高低可信度数据之间的关系可以表示为

$$y_1 = f(x, y_2) \quad (9)$$

式中  $f(\cdot)$  代表数据从低可信度层向高可信度层的转换过程。基于普遍的观点,无论是在高可信度数据层还是低可信度数据层,数据之间都展现出线性和非线性的相关性。因此,映射函数  $f(\cdot)$  可以分解为两个部分:

$$f = f_{\text{linear}} + f_{\text{nonlinear}} \quad (10)$$

## 3 算例测试分析

分别选取一个二维数学测试函数和一个翼型气动数据测试算例,以经典的 Co-Kriging 模型为代表,对比不同数量低可信度样本点和多种高低可信度样本点比例构建的 Co-Kriging 模型精度与拟合优度,

总结不同数量的低可信度样本点对多可信度模型的影响与最佳的高低可信度样本点比例。

### 3.1 数学算例测试

二维数学测试函数公式为

$$y_1 = (1 - e^{-\frac{1}{2x_2}}) \left( \frac{1000a \times x_1^3 + 1900x_1^2 + 2092x_1 + 60}{1000c \times x_1^3 + 500x_1^2 + 4x_1 + 20} \right)$$

$$y_2 = (1 - e^{-\frac{1}{2x_2}}) \left( \frac{1000a \times x_1^3 + 1900x_1^2 + 2092x_1 + 60}{1000c \times x_1^3 + 500x_1^2 + 4x_1 + 20} \right) + \frac{5e^{-a}x_1^{b/2}}{x_2^{2+b} + 1}$$

$$(a = 0.2; b = 0.3; c = 0.1; x_1, x_2 \in [0, 1]) \quad (11)$$

对低可信度函数使用拉丁超立方采样分别获得 30 和 60 个样本点,以两种不同数量的低可信度样本点为基础,分别以约 1:6、1:5、1:4、1:3 的高低样本比例再次使用拉丁超立方采样获取高可信度样本点与低可信度样本点一起训练 Co-Kriging 模型,将训练好的模型使用 100 个高可信度样本点进行测试,循环 10 次取平均值的测试结果列入表 1 和表 2。

表 1 不同低可信度样本数下不同高低样本比例的均方根误差 RMSE 对比

Tab. 1 Comparison of root mean square error of different ratios of high and low samples under different low confidence sample sizes

样本比例	样本数	
	30	60
1:6	1.5292	0.7624
1:5	1.2703	0.6884
1:4	0.9215	0.6304
1:3	0.8648	0.5864

表 2 不同低可信度样本数下不同高低样本比例的决定系数  $R^2$  对比

Tab. 2 Comparison of the coefficient of determination of different high-low sample ratios under different low-confidence sample sizes

样本比例	样本数	
	30	60
1:6	0.7840	0.9433
1:5	0.8462	0.9561
1:4	0.9293	0.9658
1:3	0.9367	0.9708

在不同的低可信度样本点数量下,随着训练用的高低可信度样本点比例增加,构建的 Co-Kriging

模型误差都在不断降低,拟合度也逐步提升。对比相同高低可信度样本比例下,60个低可信度样本点为基础的 Co-Kriging 模型比 30个低可信度样本点为基础的 Co-Kriging 模型的预测精度普遍有 35%以上的提升,在高低样本点比例为 1:6 时,预测误差甚至降低超 50%,这说明多可信度代理模型对于低可信度样本点的数量有一定要求,本次测试的数学算例较为复杂,需要低可信度样本点的数量也随之提升。

当高低可信度样本比例从 1:6 增加到 1:3,误差下降约 42%,在高低样本点比例为 1:3 时,Co-Kriging 模型的拟合优度到达 97%以上,对比高低可信度样本点比例为 1:4 时,模型精度提升效果并不明显。

### 3.2 翼型算例测试

气动数据具有较高的非线性特性,测试函数并不能反映出这一特性,因此,为验证 Co-Kriging 模型在气动数据预测领域的适用性,本文采用 NACA0012 翼型算例进行数值模拟并对其进行气动数据预测,NACA0012 翼型是一个 12%厚度的对称翼型,计算状态是  $M_\infty = 0.15 \sim 0.7$  (间隔 0.05, 共计 12 个马赫数),  $\alpha = 0^\circ, 3^\circ, 5^\circ, 7^\circ, 10^\circ, 12^\circ, 14^\circ, 16^\circ, 18^\circ, 20^\circ$ ,  $T_\infty = 288.5\text{K}$ , 计算网格如图 3 所示。

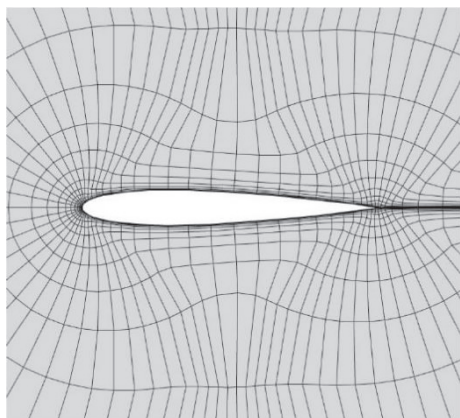


图1 NACA0012 翼型计算网格  
Fig.1 NACA0012 airfoil calculation grid

对该翼型分别采用密网格和中网格来计算高低可信度气动数据,其中密网格单元数约为中网格单元数的两倍,计算分别得到 120 个高可信度气动状态数据和 120 个低可信度气动状态数据。选取马赫数(Ma)、迎角( $\alpha$ )等 9 维变量作为模型输入,输出变量为俯仰力矩( $Cm_y$ ),使用均匀采样分别抽取 30 个和 60 个低可信度样本点,再以 1:6、1:5、1:4、1:3 的高低样本比例选取高可信度样本点与低可信度代理模型构建 Co-Kriging 可信度代理模型,测试结果列入表 3 和表 4。

表3 不同低可信度样本数下不同高低样本比例的均方根误差 RMSE 对比

Tab.3 Comparison of with different ratios of low and low confidence samples

样本比例	样本数	
	30	60
1:6	1.3779e-4	7.3335e-5
1:5	1.1179e-4	5.1062e-5
1:4	1.0179e-4	3.8062e-5
1:3	0.8568e-5	3.1593e-5

表4 不同低可信度样本数下不同高低样本比例的决定系数  $R^2$  对比

Tab.4 Comparison of the coefficient of determination of different high-low and low-sample ratios under different low-confidence sample sizes

样本比例	样本数	
	30	60
1:6	0.3220	0.7228
1:5	0.5036	0.8635
1:4	0.6280	0.9480
1:3	0.7651	0.9706

由于低可信度翼型气动数据由中密度网格计算得来,本身就具有较高精度,在 30 个低可信度样本点为基础训练得到的 Co-Kriging 模型就具有较高精度,但对比 30 个和 60 个低可信度样本点训练得到的模型精度,60 个低可信度样本点明显更优。

在 60 个低可信度样本点基础上,当高低可信度样本点为 1:6 提升到 1:3 时,Co-Kriging 的误差值下降了约 60%,拟合优度增加 24%,说明在一定量的低可信度样本点的基础上,增加少量的高可信度样本点,多可信度代理模型的精度就能获得较大提升。当高低可信度样本比例为 1:4 时,相对样本比例 1:3,拟合优度仅下降 3%,而高可信度样本点增加 33%,对于计算成本来说,高低可信度样本比例为 1:4 时更为划算。

## 4 基于 K-means 算法的多可信度代理模型采样算法

前文都是根据拉丁超立方采样采集样本点,本文考虑到同一气动状态下不同可信度气动数据的特点,提出使用 K-means 算法采样构建多可信度代理模型,文中的高低可信度气动数据是分别使用

稀疏网格计算得出,因此高低可信度气动数据属于一一映射关系,首先使用 K-means 算法在低可信度样本集中采样获得低可信度训练样本点,然后考虑高低可信度样本点嵌套与不嵌套两种情况。如果需要高低可信度样本点需要考虑嵌套,在刚获得的低可信度样本点中再次使用 K-means 算法得到高可信度样本点在低可信度样本集中的对应位置,利用位置对应关系,获得高可信度训练样本点。如果高低可信度样本点不需要嵌套,就再次使用 K-means 算法在低可信度样本集中获得高可信度训练样本点对应位置,根据位置对应关系得到高可信度训练样本点。

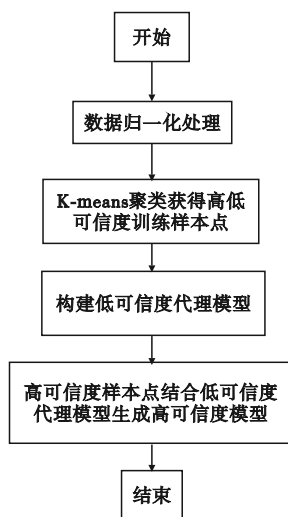


图2 K-means 算法构建多可信度模型流程

Fig. 2 Construction of a multi-confidence model based on the Gram-Mearns algorithm

本文使用的多可信度代理模型原理都是先构建出一个低可信度代理模型,再通过少量的高可信度样本点,使用不同的拟合方法,使低可信度代理模型不断趋近成为高可信度模型,主要差别就是拟合方法,通用具体步骤如下。

(1)对原始高低可信度输入数据采用归一化进行处理。

(2)对低可信度样本集使用 K-means 聚类采样,计算每个簇中距离聚类中心的距离,选择距离最近的作为低可信度训练样本点。

(3)使用低可信度训练样本点构建一个低可信度的基本代理模型。

(4)根据是否需要高低可信度嵌套使用 K-means 算法获得高可信度训练样本点。

(5)步骤(3)得到的低可信度代理模型构建高精度代理模型。

(6)输出高精度模型

多可信度模型的核心就是使用低可信度模型对高可信度样本进行拟合,高可信度样本点的选择是关键,本文使用 K-means 聚类来寻找能反映高可信度函数关键信息的采样点来保证 Cokriging 模型的准确性。

## 5 某导弹气动数据验证实例

本文针对长细比构型火箭弹展开研究,其具有小展弦比尾翼、高机动性和高性能等特性。导弹的升力系数 ( $C_l$ )、阻力系数 ( $C_d$ )、俯仰力矩系数 ( $M_z$ )是导弹主要气动数据,要计算导弹的升力系数与阻力系数等气动数据,需要进行计算流体力学 CFD (Computational Fluid Dynamics) 仿真,本文采用 NNW-FlowStar 计算软件计算,该软件经过大量标模和工程算例的考核,对于常规气动力的计算具有较高的可靠性。CFD 计算首先需要建立该导弹的模型,图 3 是模型流场。

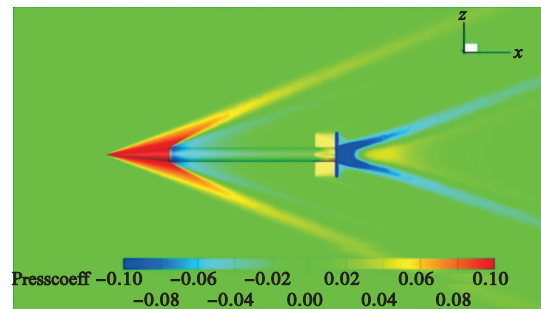


图3 导弹模型流场

Fig. 3 Flow field of the missile model

在攻角为 0 的情况下(导弹水平飞行状态下),马赫数为 0.4、2、6 的情况下描绘了阻力 ( $C_d$ )、俯仰力矩 ( $M_z$ )、横力 ( $C_t$ ) 的气动数据特征,如图 4 所示。

把密网格导弹外形经 CFD 计算得到 364 个样本点作为高可信度样本点,降低网格密度再经 CFD 计算得到 364 个低可信度样本点,共 7 个马赫数分别为 0.4、0.9、1.1、2、4、6、10,共 13 个攻角,分别为  $0^\circ$ 、 $3^\circ$ 、 $6^\circ$ 、 $9^\circ$ 、 $12^\circ$ 、 $15^\circ$ 、 $18^\circ$ 、 $21^\circ$ 、 $24^\circ$ 、 $27^\circ$ 、 $30^\circ$ 、 $33^\circ$ 、 $36^\circ$ ,共 4 个侧滑角,分别为  $0^\circ$ 、 $5^\circ$ 、 $10^\circ$ 、 $15^\circ$ ,选取马赫数 ( $Ma$ )、攻角 ( $\alpha$ )、侧滑角 ( $\beta$ ) 为 3 个设计变量,把升力系数、阻力系数、俯仰力矩系数作为预测指标。该导弹的一个高可信度样本点对比一个低可信度样本点的计算成本约为 4:1,本文为了计算方便,把一个低可信度样本点的计算成本当作 1,对应一个高可信度样本点的计算成本为 4,变可信度模型使用 80 个低可信度样本点不变,高可信度样本点分别为 10 个、20 个、30 个、40 个,其高低可信度样本点的比例分别为 1:8、2:8、3:8、4:8,对

应的计算成本分别为 120、160、200、240,分别构建上文提到的三种变可信度代理模型,在相同计算成本下,测试结果列入表 5~表 10。

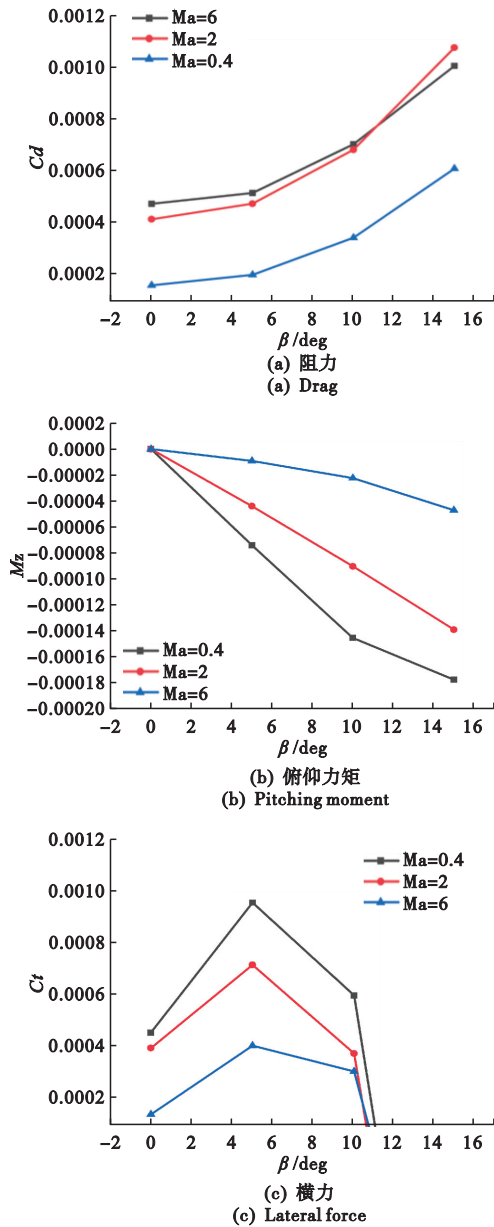


图 4 导弹气动数据特征

Fig. 4 Map of the aerodynamic data of the missile

表 5 针对升力系数(Cl)使用不同计算成本下的 RMSE 指标对比

Tab. 5 Comparison of RMSE indicators using different calculation costs for lift coefficient

模型	计算成本			
	120	160	200	240
Co-Kriging	0.1649	0.1266	0.1257	0.1225
AS-MFS	0.1751	0.1305	0.1252	0.1174
MFNN	0.3997	0.2271	0.1963	0.1832

表 6 针对升力系数(Cl)使用不同计算成本下的 R<sup>2</sup> 指标对比

Tab. 6 Comparison of R<sup>2</sup> indicators using different calculation costs for lift coefficient

模型	计算成本			
	120	160	200	240
Co-Kriging	0.9761	0.9859	0.9857	0.9861
AS-MFS	0.9722	0.9850	0.9857	0.9876
MFNN	0.5961	0.8588	0.9025	0.9536

表 7 针对阻力系数(Cd)使用不同计算成本下的 RMSE 指标对比

Tab. 7 Comparison of RMSE indicators using different computational costs for the resistance coefficient

模型	计算成本			
	120	160	200	240
Co-Kriging	0.0747	0.0553	0.0452	0.0437
AS-MFS	0.0939	0.0762	0.0696	0.0628
MFNN	0.1635	0.1265	0.1036	0.0864

表 8 针对阻力系数(Cd)使用不同计算成本下的 R<sup>2</sup> 指标对比

Tab. 8 Comparison of R<sup>2</sup> indicators using different computational costs for the resistance coefficient

模型	计算成本			
	120	160	200	240
Co-Kriging	0.9849	0.9915	0.9941	0.9944
AS-MFS	0.9754	0.9834	0.9860	0.9876
MFNN	0.9569	0.9631	0.9795	0.9835

表 9 针对俯仰力矩(Mz)使用不同计算成本下的 RMSE 指标对比

Tab. 9 Comparison of RMSE indicators using different computational costs for pitching moment

模型	计算成本			
	120	160	200	240
Co-Kriging	0.0917	0.0827	0.0859	0.0765
AS-MFS	0.1117	0.0966	0.0879	0.0861
MFNN	0.2261	0.1683	0.1382	0.1168

在升力系数预测指标中,Co-Kriging 模型的预测效果与 AS-MFS 模型的预测效果基本相同,在阻力系数与俯仰力矩为预测指标中,Co-Kriging 模

表 10 针对俯仰力矩( $M_z$ )使用不同计算成本下的  $R^2$  指标对比

Tab.10 Comparison of  $R^2$  indicators using different computational costs for pitching moment

模型	计算成本			
	120	160	200	240
Co-Kriging	0.9845	0.9870	0.9854	0.9882
AS-MFS	0.9763	0.9822	0.9852	0.9855
MFNN	0.8853	0.9102	0.9349	0.9516

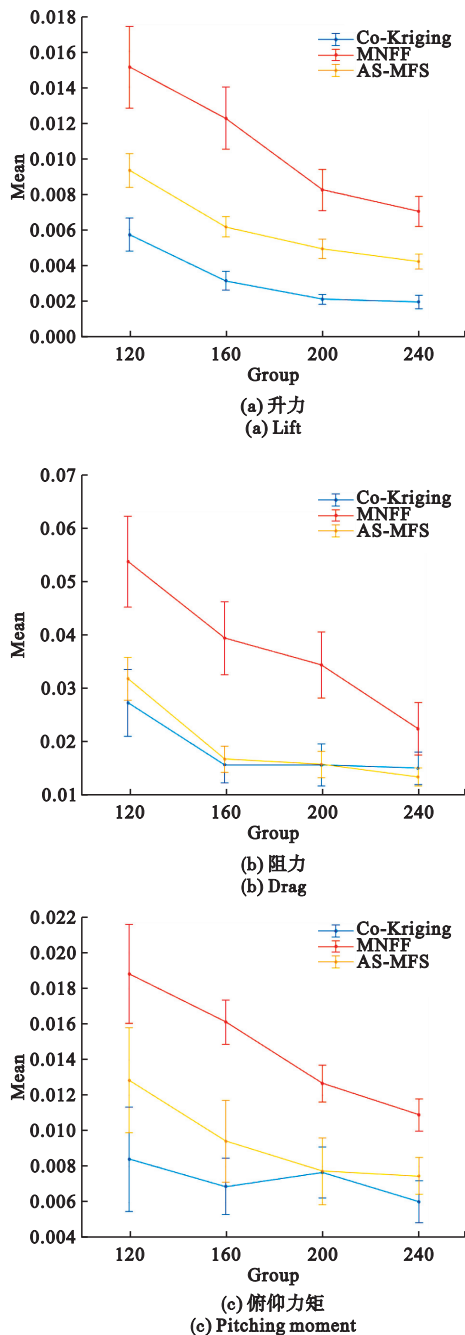


图 5 95%置信区间下三种模型的预测误差平均值

Fig. 5 Average prediction errors of three models within a 95% confidence interval

型在三种变可信度模型中模型精度最高,当计算成本为 160 时,再增加高可信度样本点对 Co-Kriging 模型和 AS-MFS 模型精度提升效果不明显,甚至会出现误差变大的情况,这是由于当高可信度样本点足够时,再增加高可信度样本点会对基础的低可信度代理模型造成反向拟合,所以在变可信度代理模型中,对高可信度样本点的选取需要适量。在导弹气动数据实验中,四种高低可信度样本数比例中最佳比例为 3:8,与第三节翼型测试结果相似。在三个测试指标中,变可信度神经网络模型在不同计算成本下精度都低于其余两种模型。综上所述,Co-Kriging 模型综合预测效果最优。

## 6 结 论

本文首先针对不同低可信度样本点与不同的高低可信度样本点比例两方面构建的 Co-Kriging 多可信度代理模型在数学算例上预测效果进行对比,表明低可信度样本点数量对于多可信度代理模型的预测效果影响较大,面对不同的问题要选择合适的低可信度样本点数,高低样本点比例为 1:4 到 1:3 之间适合构建多可信度代理模型。并在翼型与导弹气动数据预测中采用相同计算成本的样本点分别构建三种变可信度对比,结果表明 Co-Kriging 模型在精度和拟合度上都更具优势,但对于变可信度代理模型,高精度样本点的选择要适量。

## 参考文献(References):

- [1] 杜涛,许晨舟,王国辉,等.人工智能气动特性预测技术在火箭子级落区控制项目的应用[J].宇航学报,2021,42(1):61-73. (DU Tao, XU Chenzhou, WANG Guohui, et al. The application of aerodynamic coefficients prediction technique via artificial intelligence method to rocket first stage landing area control project[J]. *Journal of Astronautics*, 2021, 42(1): 61-73. (in Chinese))
- [2] 刘怡欣.基于多可信度建模的优化设计[D].北京理工大学,2017. (LIU Yixin. Optimization Design Using Multi-fidelity Metamodel Method[D]. Beijing Institute of Technology, 2017. (in Chinese))
- [3] 程吉,魏海鹏,林毅,等.基于并行变可信度置信下界算法的航行体天线罩结构优化设计[J].中国舰船研究,2021,16(4):53-62. (CHENG Ji, WEI Haipeng, LIN Yi, et al. Optimization design of antenna cover of deeply submerged body based on parallel variable-fidelity LCB algorithm[J]. *Chinese Journal of Ship Research*, 2021, 16(4): 53-62. (in Chinese))
- [4] Haftka R T. Combining global and local approximations[J]. *AIAA Journal*, 1991, 29(9): 1523-1525.

- [5] Chang K J, Haftka R T, Giles G L, et al. Sensitivity-based scaling for approximating structural response [J]. *Journal of Aircraft*, 1993, **30**(2): 283-288.
- [6] Wang S, Liu Y, Zhou Q, et al. A multi-fidelity surrogate model based on moving least squares: Fusing different fidelity data for engineering design [J]. *Structural and Multidisciplinary Optimization*, 2021, **64**(6): 3637-3652.
- [7] Bandler J W, Biernacki R M, Chen S H, et al. Space mapping technique for electromagnetic optimization [J]. *IEEE Transactions on Microwave Theory and Techniques*, 1994, **42**(12): 2536-2544.
- [8] 许晨舟, 杜涛, 韩忠华, 等. 机器学习数据融合方法在火箭子级栅格舵气动特性建模应用中的比较研究 [J]. *实验流体力学*, 2022, **36**(3): 79-92. (XU Chenzhou, DU Tao, HAN Zhonghua, et al. Comparison of machine learning data fusion methods applied to aerodynamic modeling of rocket first stage with grid fins [J]. *Journal of Experiments in Fluid Mechanics*, 2022, **36**(3): 79-92. (in Chinese))
- [9] Kennedy M C, O'Hagan A. Predicting the output from a complex computer code when fast approximations are available [J]. *Biometrika*, 2000, **87**(1): 1-13.
- [10] Han Z H, Grtz S. Alternative cokriging method for variable-fidelity surrogate modeling [J]. *AIAA Journal*, 2012, **50**(5): 1205-1210.
- [11] McKay M D, Beckman R J, Conover W J. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code [J]. *Technometrics*, 2000, **42**(1): 55-61.
- [12] Queipo N V, Haftka R T, Shyy W, et al. Surrogate-based analysis and optimization [J]. *Progress in Aerospace Sciences*, 2005, **41**(1): 1-28.
- [13] Johnson M E, Moore L M, Ylvisaker D. Minimax and maximin distance designs [J]. *Journal of Statistical Planning and Inference*, 1990, **26**(2): 131-148.
- [14] Villagran A, Huerta G, Vannucci M, et al. Non-parametric sampling approximation via voronoi tessellations [J]. *Communications in Statistics-Simulation and Computation*, 2016, **45**(2): 717-736.
- [15] Kushner H J. A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise [J]. *Journal of Basic Engineering*, 1964, **86**(1): 97-106.
- [16] Schonlau M, Welch W J, Jones D R. Global versus local search in constrained optimization of computer models [J]. *Lecture Notes-Monograph Series*, 1998, **34**: 11-25.
- [17] Söbester A, Leary S J, Keane A J. On the design of optimization strategies based on global response surface approximation models [J]. *Journal of Global Optimization*, 2005, **33**(1): 31-59.
- [18] Box G E P, Wilson K B. *On the Experimental Attainment of Optimum Conditions* [M]. NY: Springer New York, 1992.
- [19] Matheron G. Principles of geostatistics [J]. *Economic Geology*, 1963, **58**(8): 1246-1266.
- [20] Hardy R L. Multiquadric equations of topography and other irregular surfaces [J]. *Journal of Geophysical Research (1896-1977)*, 1971, **76**(8): 1905-1915.

## Research on missile aerodynamic data generation technology based on multi-credibility proxy model

FU Pengzhen<sup>1</sup>, CHEN Jiangtao<sup>2</sup>, ZHANG Peihong<sup>\*2</sup>

(1. School of Computer Science and Technology, Southwest University of Science and Technology, Mianyang 621010, China; 2. Computational Aerodynamic Institute of China Aerodynamic Research & Development Center, Mianyang 621000, China)

**Abstract:** A surrogate model is a new research idea direction of aerodynamic data generation. A traditional surrogate model relies on a large number of high-precision simulation sample points and their responses values to ensure the accuracy of the model. A multi-credibility surrogate model can reduce the computational cost while maintaining a certain accuracy by integrating multi-layer high and low credibility models, which is of great significance for reducing a missile development cycle. In this paper, the influence of different numbers of low-confidence sample points on the multi-confidence model and the optimal ratio of high-low-confidence sample points are studied, and a multi-confidence sampling method suitable for aerodynamics data is proposed. It is applied to the construction of three multi-credibility surrogate models in the prediction of missile aerodynamic data, among which the Co-Kriging model has the best comprehensive prediction effect. The recommended ratio of high and low confidence sample size is between 1:4 and 1:3.

**Key words:** erodynamic data; multi-trusted proxy model; the proportion of high and low confidence sample points; K-means clustering algorithm