arXiv:2301.09443v1 [cs.CE] 23 Jan 2023

用于提升数据驱动湍流建模泛化能力的概率机器学习

预印本

Joel Ho独立研究人员，英国牛津
joel.ho@hotmail.com

Nick Pepper艾伦·图灵研究所 大英图书馆 英国伦敦
npepper@turing.ac.uk

Tim Dodwell
计算机科学系 埃克塞特大学
英国埃克塞特

2023年1月24日

关键词 计算流体动力学；概率机器学习；场反演；湍流建模；高斯过程仿真器；深度集成

摘要

本文提出一种概率机器学习模型，用于增强 $k-\omega\ SST$ 湍流模型，以改进分离流的建模效果，并提升所学习修正的泛化能力。机器学习方法日益广泛地用于利用实验和高保真数据，从而提高工业领域常用的雷诺平均纳维–斯托克斯（RANS）湍流模型的准确性。这类方法面临的一项重大挑战，是能否泛化至未见几何形状和流动条件。此外，还必须高效处理包含实验数据和仿真数据的异构数据集。本文采用场反演和高斯过程仿真器（GPE）集成来应对这两项挑战。该集成模型应用于一系列基准测试案例，结果表明，对于具有逆压梯度的分离流案例，该模型改善了湍流建模效果；在这些案例中，RANS 模拟通常被认为并不可靠。更重要的是，在流场中呈现训练数据范围之外物理特性的区域，模拟会恢复使用未修正模型。

1 引言

计算流体力学（CFD）通过基于第一性原理对拟议几何形状周围的流动进行建模，使人们能够评估新的气动机械设计。与较简单的基于关联式的设计方法相比，这使设计人员能够探索现有设计范围之外的设计空间（1）。尽管这带来了新的机遇，但行业通常采用的雷诺平均纳维–斯托克斯（RANS）模型存在根本性缺陷，削弱了其作为设计工具的效用（2）。在 RANS 中，中间和较小的能量尺度是通过建模而非直接解析来处理的。与大涡模拟（LES）和直接数值模拟（DNS）相比，这使模型的计算成本较低；后两者能够解析更小的尺度，但模型的准确性也因此有所降低（例如，见（3））。

近年来，人们开始关注采用机器学习（ML）方法，利用来自大涡模拟（LES）、直接数值模拟（DNS）和实验的高保真数据，在仅增加极少计算成本的同时提高雷诺平均纳维–斯托克斯模型的准确性（4）。从许多方面来看，由于所需模型包含大量参数且具有非线性特征，这是一个非常适合应用机器学习的自然问题（5）。在这些研究中可以概括出两种主要策略：一种方法是直接从高保真数据中学习经过修正的雷诺应力张量，重点关注雷诺平均纳维–斯托克斯结果中的某些假设失效的区域。例如，可以通过神经网络实现这一点；Frey 等人（6）和 Ling（7）利用基于直接数值模拟数据训练的神经网络，对其中的涡黏性项进行修正。

布西内斯克方程。Weatheritt 和 Sandberg（8）采用了基因表达式编程（GEP），在该方法中，Pope（9）的显式代数雷诺应力（EASM）框架被用于执行符号回归，从而学习改进后的各向异性非线性表达式。该方法提供的显式表达式比神经网络更易于解释，但代价是模型的自由度有所降低。

另一种策略是修改 RANS 湍流方程，加入空间修正函数，使解能够更好地拟合参考平均流动解。如果整个流场的数据均可用，则可以采用冻结方法，将待拟合变量保持不变（例如速度和 k），同时求解其余 RANS 建模变量及所需的修正项。Schmelzer 等人（10）采用的就是这种方法。如果只有稀疏数据可用，则可以使用场反演或数据同化技术。通过最小化参考数据与 RANS 数据之间误差构成的目标函数来求解，通常采用伴随方法获得目标函数梯度。即使只有线状或平面数据（11）、表面压力系数或升力系数（12）可用，也能得到整个流场的修正函数。随后，使用监督学习算法根据局部流动变量学习修正函数。采用这种方法的例子包括 Parish 和 Duraisamy（13），以及较近的 Ho 和 West（14）。

同时采用这两种策略的研究取得了令人鼓舞的成果，但也引发了关于实际运用机器学习方法改进湍流建模的疑问（15；16；17）。本文探讨了其中三个挑战：

- 面对新的几何形状和流动条件时，我们应该在多大程度上信任机器学习算法的预测?
- 我们如何判断经过机器学习修正的湍流模型是否超出了训练数据集的范围进行外推?
- 该模型能否处理规模庞大、异构的数据集?

为应对这些挑战，本文将概率机器学习框架与场反演相结合。分别训练多个模型组成的集成，提供了一种模块化方案，能够自然地处理异构数据。通过分析机器学习修正的方差，我们表明，该框架能够提高与训练数据中几何形状相似的几何形状的 RANS 模拟精度；与此同时，也许更为重要的是，对于超出训练数据范围的几何形状，该框架会拒绝采用机器学习修正。
本文结构如下：第 2 节讨论场反演方法、无量纲流动特征的识别以及概率机器学习模型。第 3 节介绍用于生成异构数据的训练案例，并将流动反演结果与参考解进行比较。此外，还将讨论如何利用这些数据训练概率机器学习框架，以及作为其性能基线的方法。最后，我们展示该框架在绕流四种此前未见几何形状中的应用。

2 方法论

2.1 场反演

对 $k-\omega$ SST 模型的比耗散率 ω方程进行了修改，使RANS解能够适应参考解。将空间变化的修正因子 $\beta_c(\boldsymbol{x})$ 应用于生成项，得到修改后的 ω 方程：

$$\frac{D\omega}{Dt} = -\nabla\cdot(\rho D_\omega \nabla\omega) + \beta_c(\boldsymbol{x})\frac{\rho\gamma G}{\epsilon} - \frac{2}{3}\rho\gamma\omega(\nabla\cdot\boldsymbol{u}) - \rho\beta\omega^2 - \rho(F_1-1)CD_{k\omega}. \tag{1}$$

有关湍流模型各项的详细信息，可参见Menter（18）的文献。Matai和Durbin（19）研究了将修正因子置于 ω 方程不同部分的影响，并发现无论将 β_c 作为生成项还是耗散项的乘数，都能恢复精确的 ν_t 场。Duraisamy等人（20）指出，将修正因子作为乘数引入现有项，等同于向模型中添加一个新的修正项。将 β_c 作为乘数的优点在于它是无量纲的，且其优化值与未校正值1处于同一数量级。

众所周知，标准 RANS湍流模型在分离流中的表现较差，通常会过度预测分离区，并导致再附着延迟。这是因为模型低估了分离流剪切层中存在的较大涡，从而低估了湍流剪切应力，并降低了分离区中的湍流混合（例如，见文献 (21)）。将 β_c 引入 ω 方程，会改变湍流产生

与耗散在局部的平衡，使模型能够克服上述缺陷。这一点体现在第 3 节所示的反演结果中：对于所有案例，修正均主要作用于分离区的剪切层，其中大多是由 $\beta_c < 1$ 导致 k增加。

给定 $k-\omega$ SST 模型的未校正解以及高保真或实验参考解，我们希望求得场 $\beta_c(\boldsymbol{x})$ 的估计值，使两个流场之间的差异最小。所需的 β_c 场通过将流场离散为 n 个单元并求解以下优化问题来确定：

$$\begin{aligned} &\min_{\boldsymbol{\beta_c}} \mathcal{J}(\boldsymbol{\beta_c}, \boldsymbol{w}) \\ &\text{such that } \mathcal{N}(\boldsymbol{\beta_c}, \boldsymbol{w}) = 0, \end{aligned} \tag{2}$$

其中，$\boldsymbol{\beta_c} \in \mathbb{R}^n$ 收集每个单元的修正因子，$\boldsymbol{w}$ 表示守恒变量。$\mathcal{N}(\cdot)$ 是求解流动方程的非线性离散算子，而 $\mathcal{J}(\cdot)$ 是用于计算 n_a 个单元上的 RANS 解与参考解之间差异的目标函数（该单元数可以少于模拟中的总单元数，即 n）。目标函数定义如下：

$$\mathcal{J}(\boldsymbol{\beta_c}, \boldsymbol{w}) = \frac{1}{n_a}\sum_{i=1}^{n_a}\Big(\boldsymbol{u}_i(\boldsymbol{\beta_c}, \boldsymbol{w}) - \boldsymbol{u}_{\text{ref},i}\Big)^2 + \frac{\lambda}{n}\sum_{i=1}^{n}(\boldsymbol{\beta_{c,i}} - 1)^2, \tag{3}$$

其中，$\boldsymbol{u}_i(\boldsymbol{\beta_c}, \boldsymbol{w})$ 表示第 i^{th} 个单元处求得的 RANS 速度场，而 $\boldsymbol{u}_{\text{ref},\ i}$ 表示该单元处的参考解。第一项表示求得的 RANS 速度场与参考解之间的距离，第二项是修正因子的正则化项，其中 λ 为正则化常数。方程 (2) 和 (3) 中所述的优化问题采用梯度下降算法求解。优化结束后，得到一个在每个单元内分段常数的 β_c 场。采用离散伴随公式求得所需的梯度：

$$\begin{aligned} \left(\frac{\partial R}{\partial \boldsymbol{w}}\right)^{\top}\boldsymbol{\phi} &= \frac{\partial \mathcal{J}}{\partial \boldsymbol{w}}, \\ \frac{d\mathcal{J}}{d\boldsymbol{\beta_c}} &= \frac{\partial \mathcal{J}}{\partial \boldsymbol{\beta_c}} - \boldsymbol{\phi}^{\top}\frac{\partial R}{\partial \boldsymbol{\beta_c}}, \end{aligned} \tag{4}$$

其中，$R \in \mathbb{R}^{(n\times n_w)}$ 表示流动求解器驱使其趋于零的残差，$\boldsymbol{w}$ 表示守恒变量，n_w 表示守恒变量的数量，而 $\boldsymbol{\phi} \in \mathbb{R}^{(n\times n_w)}$ 是伴随解。请注意，$\boldsymbol{\beta_c} \in \mathbb{R}^n$ 指每个单元中的修正量。

本研究采用了具有二阶迎风离散的有限体积求解器。使用SIMPLE算法求解带有线性涡黏模型的不可压缩RANS方程。获得收敛的原始流场后，使用离散伴随求解器计算优化所需的梯度。

2.2 特征空间表示

完成反演后，将 β_c 回归到描述流动的无量纲特征空间，记为$\boldsymbol{\alpha} \in \mathbb{R}^{52}$。其中五个特征是人工设计的，具体如下：

流线压力梯度：

$$\frac{2}{\pi}\arccos\left(\frac{\nabla p \cdot \boldsymbol{u}}{\sqrt{||\nabla p||^2||\boldsymbol{u}||^2}}\right) - 1. \tag{5}$$

湍流时间尺度比：

$$2\left[\left(\frac{S}{S+\omega}\right)^{0.25} - 0.5\right]. \tag{6}$$

Q判据：

$$\frac{S^2 - \Omega^2}{S^2 + \Omega^2}. \tag{7}$$

湍流强度：

$$2\left[\left(\frac{k}{0.5||\boldsymbol{U}||^2+k}\right)^{0.25}-0.5\right]. \tag{8}$$

湍流黏度（ν_t）比：

$$2\left[\left(\frac{\nu_t}{100\nu+\nu_t}\right)-0.5\right]. \tag{9}$$

S 和 Ω 分别表示应变率张量和旋转率张量的弗罗贝尼乌斯范数（即 $S = \sqrt{\sum_i \sum_j S_{ij}S_{ij}}$）。 p 表示运动压力， k 表示湍动能， $\boldsymbol{U}$ 表示雷诺平均速度， ν 表示运动黏度。对湍流时间尺度比和湍流强度特征取 0.25 次幂，以改善其分布。需要指出的是，方程（5）和（8）不具备伽利略不变性；然而，本研究中的测试案例处于同一参考系，因此其影响并不显著。

其余 47 个特征基于张量集合 $\{S, \Omega A_p, A_k\}$ 的旋转不变性最小完整基构造，其中上方的帽号表示无量纲化。张量 A_p 和 A_k 分别是压力梯度和湍动能梯度，例如通过 $-I \times \nabla p$ 映射为反对称张量。关于这些特征的更多详情，参见 Wu等人（22）和 Ling等人（23）。

2.3 概率机器学习框架

从雷诺平均纳维-斯托克斯方程的未校正解中获得无量纲特征 $\boldsymbol{\alpha}$，并通过场反演获得校正因子 β_c 后，我们引入一个概率机器学习框架来学习映射 $\boldsymbol{\alpha} \rightarrow \beta_c$。考虑到实际应用中的参考数据可能来自异构数据源（例如粒子图像测速、直接数值模拟和大涡模拟）以及多种几何形状，因此这种映射很可能是一对多映射。在整个数据集上训练的机器学习模型可能无法学习各个数据集中的趋势，从而限制其预测能力（辛普森悖论（24））。为解决这一问题，我们引入一种基于高斯过程仿真器集成（GPE）的概率机器学习模型来表示该映射。

给定来自 n_m 个独立数据源的 n_t 个参数点 $X = [\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(n_t)}]$ 和 $Y = [\beta_c^{(1)}, \ldots, \beta^{(n_t)_c}]$的训练数据集，我们为每个数据源分别构建一个插值函数 Ψ，即：

$$\beta_c(\boldsymbol{\alpha}) = \Psi_k(\boldsymbol{\alpha}) + \boldsymbol{\epsilon}_k, \quad k = 1, \ldots, n_m, \tag{10}$$

其中， $\boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{\mu_0}, \text{diag}(\boldsymbol{\sigma}_\epsilon))$假设为方差为$\boldsymbol{\sigma}_\epsilon^2 \in \mathbb{R}^{n_m}$、均值为$\boldsymbol{\mu_0} \in \mathbb{R}^{n_m}$的独立同分布高斯噪声。 Ψ_k指的是 k^{th} 数据源的插值函数。假设这些函数服从高斯过程先验，因此可将特征空间中新参数点$\boldsymbol{\alpha}$*处的 β_c 后验密度表示为高斯分布：

$$\beta_c | X_k, Y_k, \boldsymbol{\alpha}^* \sim N(\mu_k, \sigma_k), \tag{11}$$

其中， X_k 和 Y_k 表示训练数据中源自 k^{th} 高保真数据集的子集，从而能够在框架中纳入不同的数据集，例如离散测量、高保真计算流体力学，以及可能的基于经验的规则。假设均值函数为常数 $\boldsymbol{\mu_0}$，则 k^{th} GPE的均值和标准差表示为：

$$\begin{aligned} \mu_k &= \boldsymbol{\mu}_{\mathbf{0},k} + K(\boldsymbol{\alpha}^*, X)^\top K(X,X)^{-1}(Y_k - \boldsymbol{\mu}_{\mathbf{0},k}), \\ \sigma_k^2 &= K(\boldsymbol{\alpha}^*, \boldsymbol{\alpha}^*) - K(\boldsymbol{\alpha}^*, X)^\top K(X,X)^{-1} K(\boldsymbol{\alpha}^*, X), \end{aligned} \tag{12}$$

其中， K 指协方差矩阵，其元素为：

$$K_{ij}(X,X) = k(\boldsymbol{\alpha}^{(i)}, \boldsymbol{\alpha}^{(j)}) + \sigma_{\epsilon_k}^2 \delta_{ij}. \tag{13}$$

k(。)表示协方差函数， δ 表示克罗内克δ函数。本研究采用平方指数核作为协方差函数。GPE 的超参数由数据推断，代价为 $O(n_t^3)$。为降低 n_t 较大时（例如 $n_t > 10^3$）的计算代价，我们对高斯过程的超参数采用随机变分回归。我们

感兴趣的读者可参阅 Blei 等人（25）和 Hensman 等人（26），了解这一方法的更多细节；有关高斯过程专家的总体介绍，可参阅 Rasmussen 和 Williams（27）以及 Santner 等人（28）的经典参考书。

获得训练好的高斯过程专家集成 $N(\mu_k(\boldsymbol{\alpha})\ \sigma_k(\boldsymbol{\alpha}))\ \ k=1,\ \ldots\ldots,\ n_m$ 后，我们讨论如何在概率机器学习框架中为各专家的贡献加权，以及依据何种标准接受此修正因子。对于特征空间中的未见参数点 $\boldsymbol{\alpha}^*$，通过公式 (12) 确定集成中每个高斯过程专家的均值和标准差，并通过加权求和得到集成的均值和方差：

$$\mu^*(\boldsymbol{\alpha}^*) = \sum_{k=1}^{n_m} w_k \mu_k(\boldsymbol{\alpha}^*), \tag{14}$$
$$\sigma^{*2}(\boldsymbol{\alpha}^*) = \sigma^{*2}_{\mu}(\boldsymbol{\alpha}^*) + \sigma^{*2}_{\sigma}(\boldsymbol{\alpha}^*),$$

其中，集成的标准差不仅包含各个高斯过程专家的方差 σ^*_{σ}，也包含其均值的离散程度 σ^*_{μ}，如 Lakshminarayanan 等人 (29) 所述，其中：

$$\sigma^{*2}_{\mu} = \sum_{k=1}^{n_m} w_k \mu_k^2(\boldsymbol{\alpha}^*) - \mu^{*2}(\boldsymbol{\alpha}^*), \tag{15}$$
$$\sigma^{*2}_{\sigma} = \sum_{k=1}^{n_m} w_k \sigma_k^2(\boldsymbol{\alpha}^*).$$

各个混合权重满足 $\sum_{k=1}^{n_m} w_k = 1$，并通过逆加权方差确定：

$$w_k = \frac{1/\sigma_k^2}{\sum_{j=1}^{n_m} 1/\sigma_j^2}. \tag{16}$$

因此，在该框架中，置信度较高、方差较低的 GPE 会被赋予更大的权重。采用概率机器学习的一个关键动机是，在预测修正因子时，能够以方差的形式同时给出不确定性度量。方差用于识别两种情形：一是该框架正在对训练数据范围之外的特征空间区域进行泛化；二是 $\boldsymbol{\alpha}^*$ 位于训练数据范围内，但各高斯过程专家的预测存在显著分歧。正如后文将会看到的，在这些情况下应用数据驱动修正可能会降低雷诺平均纳维–斯托克斯结果的准确性，相较于未校正的解尤为如此。因此，我们对该框架的预测采用以下接受准则：

$$\beta_c = \begin{cases} \mu^*(\boldsymbol{\alpha}^*) \text{ if } \sigma^*(\boldsymbol{\alpha}^*) \leq \bar{\sigma} \\ 1 \text{ otherwise} \end{cases}, \tag{17}$$

其中，$\bar{\sigma}$ 是由用户指定的容差，必须在修正幅度与保守性之间进行权衡；当该方法应用于训练数据范围之外的流动条件或几何形状时，尤需如此。除上述高斯过程集成外，我们还依据 Lakshminarayanan 等人（29）训练了一个独立的深度集成（DE）模型，并采用类似的接受准则，以便与所提出的方法进行比较。两个模型均使用从三个实验数据集和高保真数据集中汇集的训练数据集进行训练，这些数据集将在下一节介绍。与文献中常见的方法不同——该方法会在自由来流中最小化不确定性，并在雷诺应力各向异性较高的流动区域中最大化不确定性（例如，见（30））——我们转而在 β_c 处于激活状态的区域中最小化模型的预测不确定性。具体做法是从训练数据集中移除区间 $\beta_c \in [0.9, 1.1]$ 内的点；回顾一下，$\beta_c = 1$ 表示修正处于非激活状态。

模型训练完成后，将对未来案例的预测作为一次性校正，添加到已收敛的未校正流场中；随后再次求解，以获得最终流场。本文所述的概率机器学习框架并不局限于这种实现方式，也可应用于每次迭代都会调用的嵌入式机器学习模型，且不会损失任何通用性。

3 数据生成与模型训练

训练数据收集自一系列公开可用的实验和高保真度结果。所选的三个案例分别是：Greenblatt 等人（31）的美国国家航空航天局壁挂式凸起、Bentaleb 等人（32）的弯曲后向台阶，以及 Lehmkuhl 等人（33）的圆柱体。数据类型各不相同，依次包括粒子图像测速、大涡模拟和直接数值模拟。这些案例的共同特征是在逆压梯度下，流动从曲面发生分离。

3.1 NASA 驼峰训练案例

原始实验中选择了经过修改的格劳尔特驼峰几何形状，因为其分离点不由几何形状固定，能够代表空气动力学体。经过较长的前体后，流动到达一段相对较短的凹形斜坡，并在此处分离。据报道，在较大的雷诺数范围内（$2.4\times10^6 < Re < 26\times10^6$），分离流场对雷诺数不敏感。最初曾开展一系列在斜坡前进行主动吹气的实验，但本研究仅使用未进行流动控制的基线结果。与本研究使用的其他计算数据集相比，此处的雷诺数更高，为 $Re_c = 9.36 \times 10^5$。

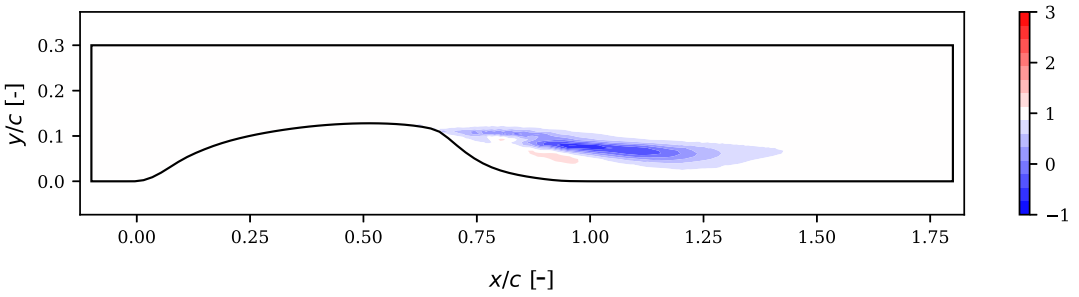
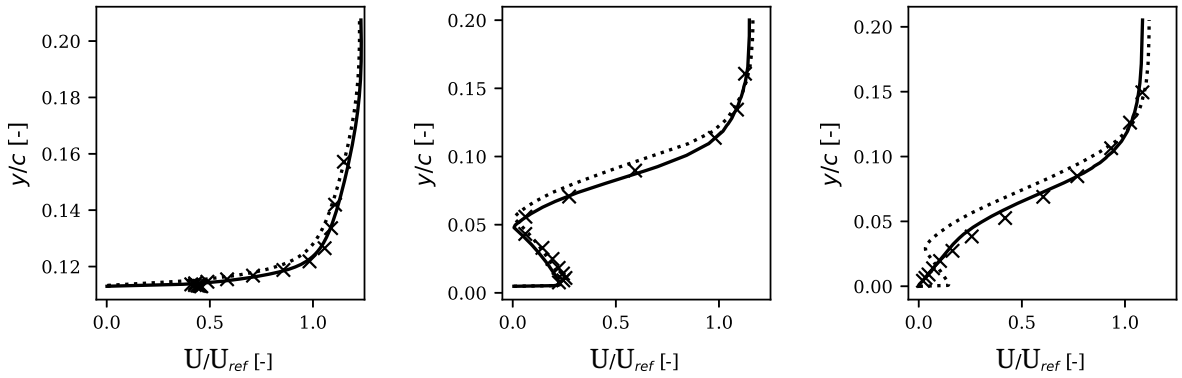


图 1：NASA 驼峰的反演 β_c 场。

将粒子图像测速数据作为数据线进行同化，数据线位于 $0.8 \leq x/c \leq 1.3$ 之间，间距为 0.1。还纳入了位于 $x/c = 0.65$ 和 0.66 的另外两条数据线，它们恰好位于凸起峰值之后。经过 63 次迭代后，目标函数下降了一个数量级并开始趋于平稳，因此停止了优化。最终的修正场在远离凸起的分离区最为活跃（图 1）。优化后的 β_c 大多小于 1，仅有一小片区域的 β_c 增大。图 2 显示，反演得到的速度场与测量得到的速度场吻合良好。



(a) $x/c = 0.66$。 (b) $x/c = 0.90$。 (c) $x/c = 1.10$。

图 2：速度 分布：倒置 NASA 驼峰案例。参考粒子图像测速（ ）×，未校正（ ）·,···倒置（ ） —— .

3.2 弯曲后向台阶训练案例

与 NASA 驼峰案例类似，弯曲后向台阶没有固定的分离点。这里的一个不同之处是流动没有上游加速，因此其流动分离发生在充分发展湍流边界层中。数据取自平均大涡模拟速度，原始模拟采用基于高度的雷诺数 $Re_h = 13700$。

数据在一个二维区域内进行了同化，其中 $0 < x/h < 8$ 和 $0 < y/h < 2$，当目标函数下降一个数量级后趋于平稳时，优化在经过 48 次迭代后停止。图 3 显示了反演的β_c 场，与其他反演案例相比，该场在边界层中的修正幅度更大。因此，在 $x/c = 0.66$处边界层中的动量增加，与图 4a 所示的大涡模拟速度分布更加吻合。分离区中的速度分布也与参考大涡模拟结果吻合良好（图 4）。

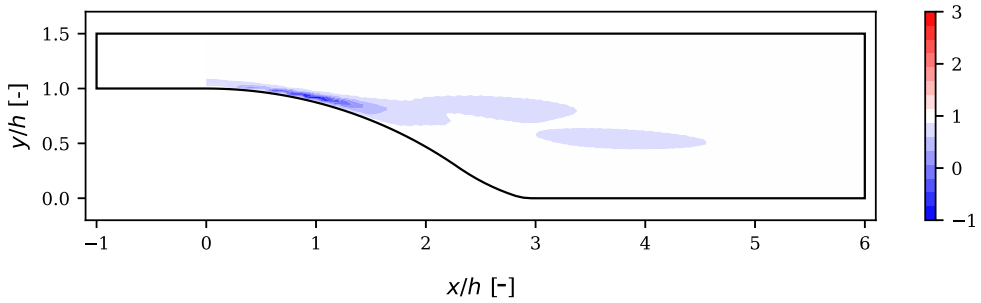
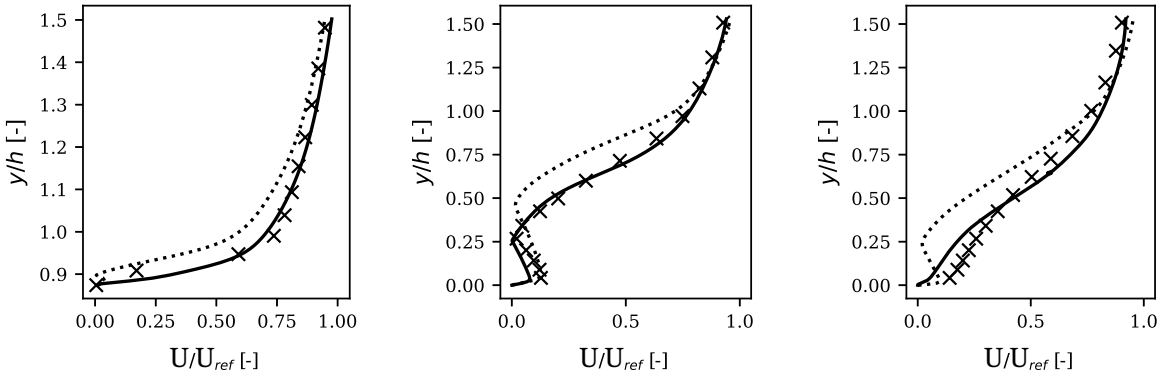


图 3：弯曲后向台阶的反转 β_c 场。



(a) $x/h = 1$。 (b) $x/h = 3$。 (c) $x/h = 5$。

图 4：速度 弯曲后向台阶的 y 剖面。参考大涡模拟（ ）、未校正（ ）、反演（ …… ——).

3.3 圆柱训练案例

该案例的参考数据基于一次相对较长的直接数值模拟，其研究重点是分离泡的低频非定常性。该案例包含分离剪切层及其与湍流尾流的相互作用。基于直径的雷诺数为 3900。

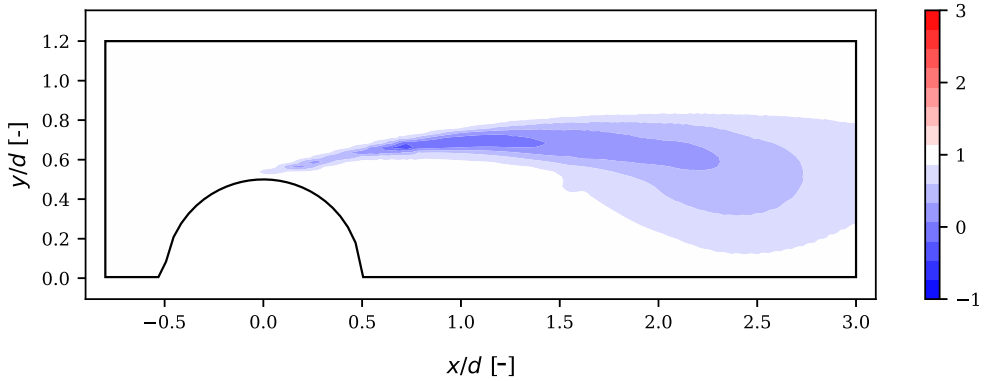
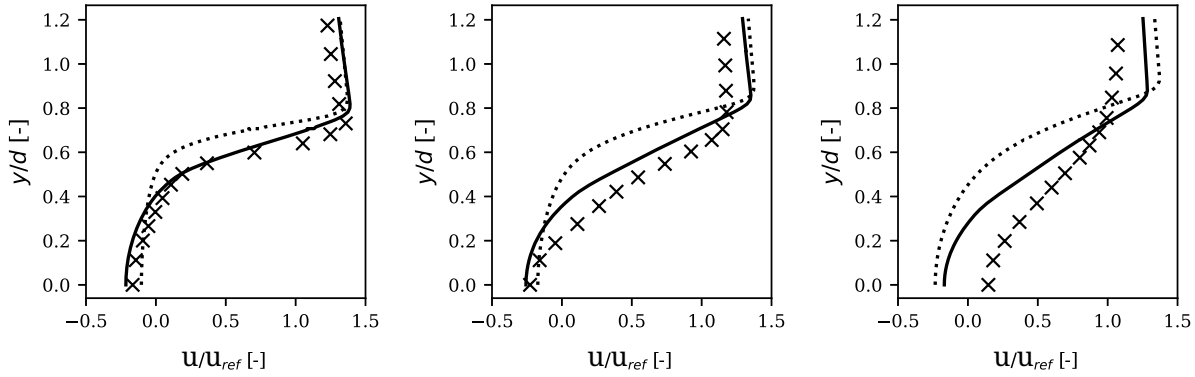


图5：圆柱体的反演 β_c 场。

对位于 $x/c = 1.06$、1,54 和 2.02 处线上的近尾流数据进行了同化。当目标函数开始趋于平稳时，优化在经过 80 次迭代后停止。与其他反演结果相比，分离泡末端附近尾流中的校正区域更大。这可能是由于未校正结果在分离后的松弛区预测较差，如图6中的速度分布所示。

反演流场与直接数值模拟流场在 $x/c = 1.06$ 和 $x/c = 1.54$ 处的速度分布吻合良好，且反演流场在 $x/c = 2.02$处有所改善。总体而言，反演流场中的尾流扩散有所增强。



(a) $x/d = 1.06$. (b) $x/d = 1.54$。 (c) $x/d = 2.02$。

图6： 横向来流中的圆柱的速度分布。参考直接数值模拟（）, 未校正（）, 反演（ ——).

3.4 集成模型的训练

对上述讨论的三个案例分别进行场反演，得到训练数据集 $D = [D_1, D_2, D_3]$ D_i $\boldsymbol{\alpha}^{(j)}$ $\beta_c^{(j)}$ $j =$ 1 ... i^{th}，其中每个训练数据子集，收集了从相应案例中提取的参数点集合（, , n_i，其中 n_i 表示校正起作用的单元数量。针对每个数据子集训练一个GPE，从而构成由三个高斯过程专家组成的集成。训练结果如图7所示。图7a–7c展示了集成中各个高斯过程专家的训练结果，将GPE预测的 β_c 与反演结果进行比较。请注意，由于移除了校正不显著的参数点，训练数据中出现了一个空缺。图7d展示了整个集成的结果，图7e则绘制了集成误差（以 L_1 误差量化）与预测不确定性 σ^*之间的关系。

深度集成的训练结果 也显示在图8中。与 GPE 类似， L_1 误差随着模型预测不确定性的增加。

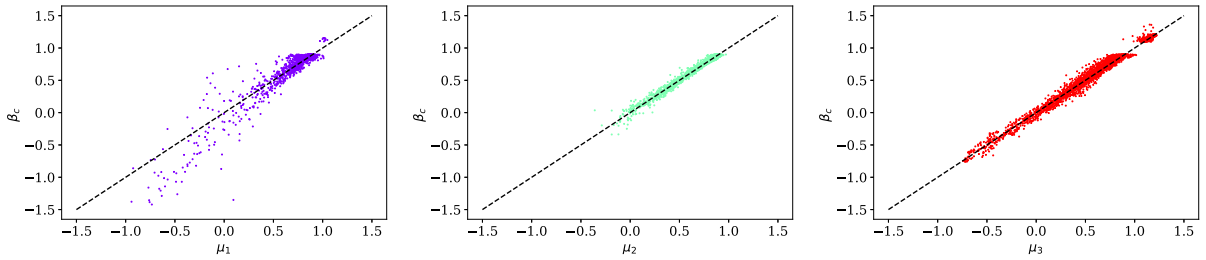
4 对未见测试案例的应用

选取了两类测试案例来评估概率机器学习框架的有效性。第一类是预期机器学习模型会在流动的特定区域发挥作用并改善结果的案例。针对这一类，测试了 Matai和Durbin (19) 的 h42 凸起，以及 Frohlich等人 (34) 的周期性丘陵算例。第二类案例不包含训练数据中修正所针对的流动物理。这类案例包括零压力梯度平板，以及不同攻角下的 NACA0012翼型；在这些案例中，经机器学习增强的湍流模型应回归至基线。

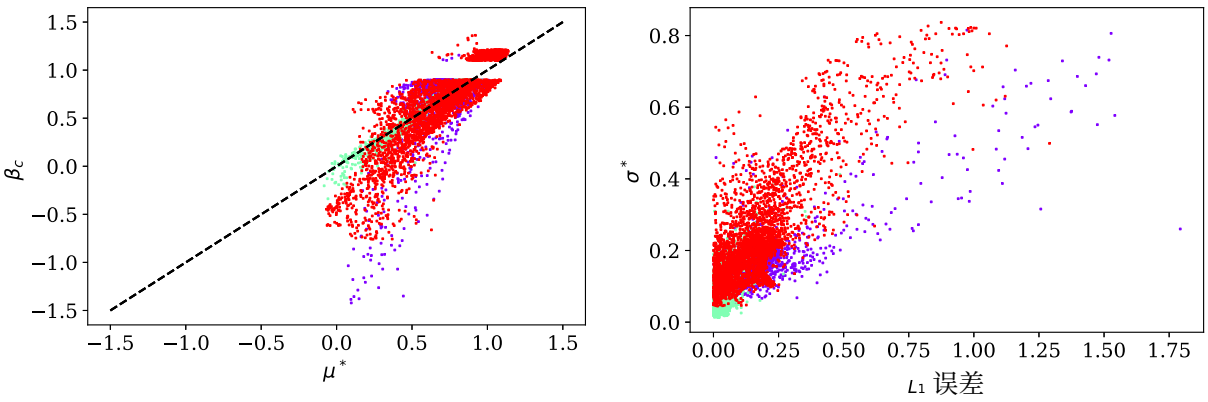
4.1 h42 凸起

Matai 和 Durbin（19）通过大涡模拟对一系列参数化凸起进行了模拟，以研究边界层和流动分离在逆压梯度逐渐增大时的系统性变化。最低的凸起上流动完全附着，而最高的凸起则出现了分离泡。McConkey（35）提供了论文中的数据。机器学习模型应用于最高的凸起（h42），并与未用于训练的场反演结果进行了比较。进行反演是为了提供一个参考，以评估机器学习模型的性能。

测试案例取自 McConkey（35），采用 412×175 网格，壁面上的最大 y^+ 为 0.85。最大台阶高度为 42 mm 时，基于高度的雷诺数为 $Re_h = 27850$。

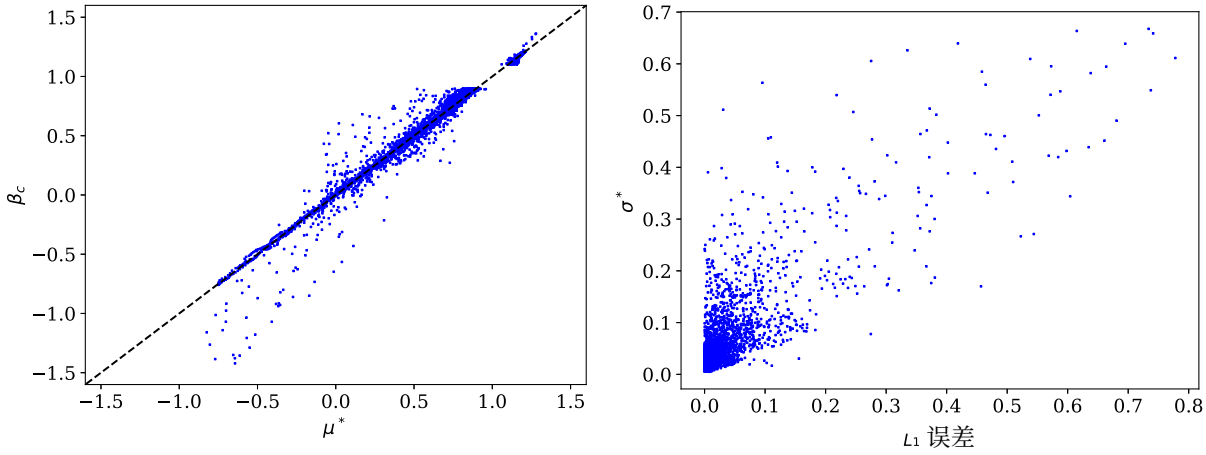


(a) 弯曲后向台阶 (b) 圆柱体 (c) NASA 驼峰



(d) 整个集成模型 (e) 集成模型 L_1 误差与 σ^*的关系

图7：在反演数据上训练的 GPE集成的结果：各个 GPE子模型 (a)–(c)；整个集成模型 (d)；以及集成模型 L_1 误差与 σ^* (e) 的关系。



(a) 深度集成 (b) 深度集成 L_1 误差与 σ^*的关系

图8：基于反演数据训练的深度集成模型的结果。

4.1.1 反演结果

首先同化了完整的大涡模拟流场。从图9可以看出，反演结果中的分离区与大涡模拟吻合良好。反演结果还显示，凸起迎风部分的壁面摩擦阻力有所降低，而后半部分则有所增加；这是由于流动分离前壁面附近的湍动能增加所致。

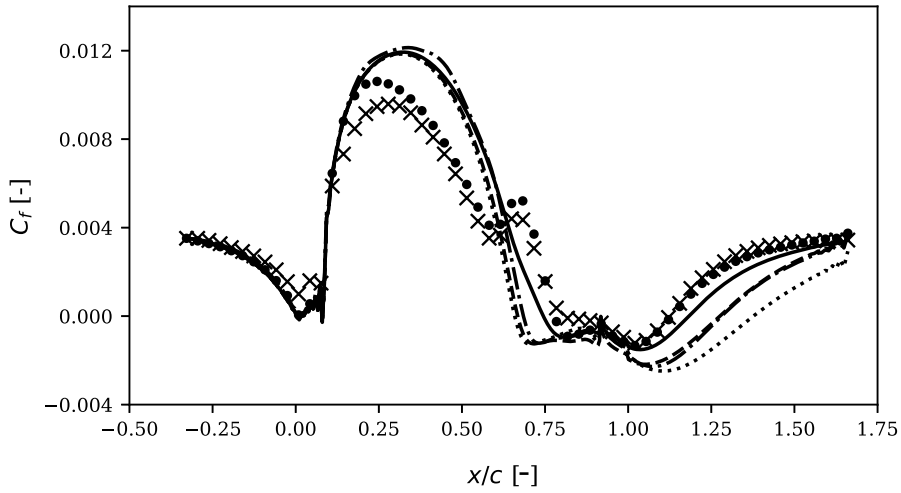


图9：凸起表面的壁面摩擦阻力。参考大涡模拟（）、反演（）、约束反演（）、未校正（）、高斯过程专家（）、深度集成（）。

反演结果的无量纲速度剖面与图10所示结果也吻合良好。与大涡模拟类似，同化结果在 $x/c = 0.5$ 和 $x/c = 0.6$处显示出更厚的黏性底层。在 $x/c = 1.5$处，流动恢复也更好。

雷诺平均纳维–斯托克斯结果高估了 $x/c < 0.6$ 附近的剪切应力，随后又低估了该位置之后的剪切应力。图11中的反演修正场显示，近壁区 $x < 0.6$ 附近的 β_c 有所增加，从而提高了比耗散率，进而降低了 k 和 ν_t。这使剪切应力降低。(19)中指出，流动并未重新层流化。未校正的雷诺平均纳维–斯托克斯剪切应力在 $x > 0.6$之后变为低估。此外，壁面外侧存在一片 $\beta_c < 1$ 区域，使 k 增加，从而缩小了分离泡。
由于训练数据主要由分离区内的 $\beta_c < 1$ 构成，因此预计机器学习模型无法预测凸起前部增加的 β_c。我们进行了约束场反演，其中仅允许 β_c 从 $x/c > 0.6$ 开始起作用，并同化了分离区内范围较小的一部分大涡模拟数据（即 $0.8 < x/C < 1.5$ 和 $0 < y/C < 0.16$）。这更能代表用于生成训练数据的反演过程，随后可用这些数据评估机器学习模型。约束反演结果仅在分离区内出现了ω 产量降低的区域（图11），近壁区的 k 则没有下降。

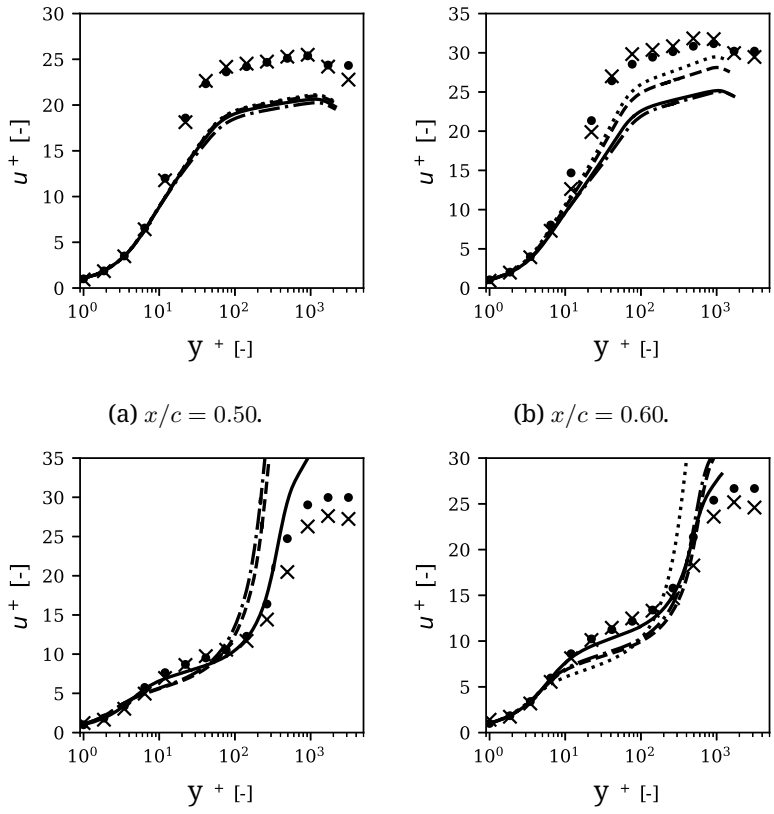
4.1.2 方差容差扫描

对两种模型进行了 $\bar{\sigma}$ 扫描，以评估调整式（17）中接受准则的影响。图12a显示，对于高斯过程专家，随着 $\bar{\sigma}$增加，壁面摩擦阻力剖面如何从未校正剖面逐渐向约束反演结果变化。当 $\bar{\sigma}$ 较低，为0.15时，再附着长度略有改善，且上游不受影响。 C_f 剖面始终处于约束反演结果与未校正结果之间，直至$\bar{\sigma} = 0.25$；此时再附着点与约束反演结果相符，但迎风面上的 C_f 从 $x/c = 0.25$ 开始超出界限。因此，所有测试案例中的高斯过程专家均选用0.2的容差。

图12b显示，深度集成的表现也类似。然而，它比高斯过程专家更早过冲 C_f，只有 $\bar{\sigma} = 0.05$ 结果仍保持有界。在该容差下，再附着长度几乎没有改善。为了使再附着长度的改善幅度达到高斯过程专家在其规定容差下的水平，为深度集成设定了 $\bar{\sigma}$值 0.1，允许其在凸起处出现一定过冲。本研究的所有测试案例均采用这些恒定值。

4.1.3 机器学习预测

GPEs 和深度集成都预测了分离区内 β_c 减小的区域，弥补了第 2 节所述湍流模型的不足。两种模型都改善了再附着和分离后的流动恢复。与深度集成相比，GPEs 的预测在距壁面更远处仍然有效，其 C_f 在 $x/c = 1.1$ 之前与未校正结果重叠，之后则优于未校正结果。与约束反演结果不同，这两种机器学习



(a) $x/c = 0.50$. (b) $x/c = 0.60$.

(c) $x/c = 1.33$. (d) $x/c = 1.50$.

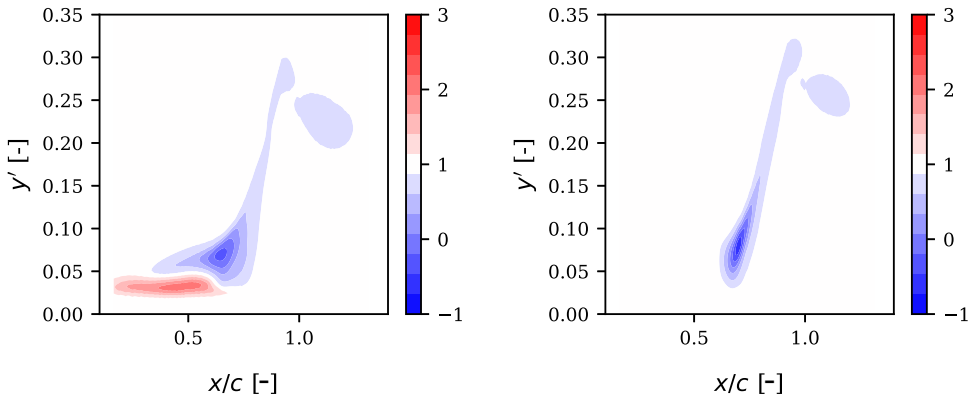
图10：无量纲速度。图例见图9。

模型均未预测到延迟分离。GPEs 模型的 C_f 剖面介于未校正结果和约束反演结果之间。由于深度集成在近壁区的激活程度更高，其结果在约 $0.25 < x/c < 0.6$ 处偏离了这一范围。

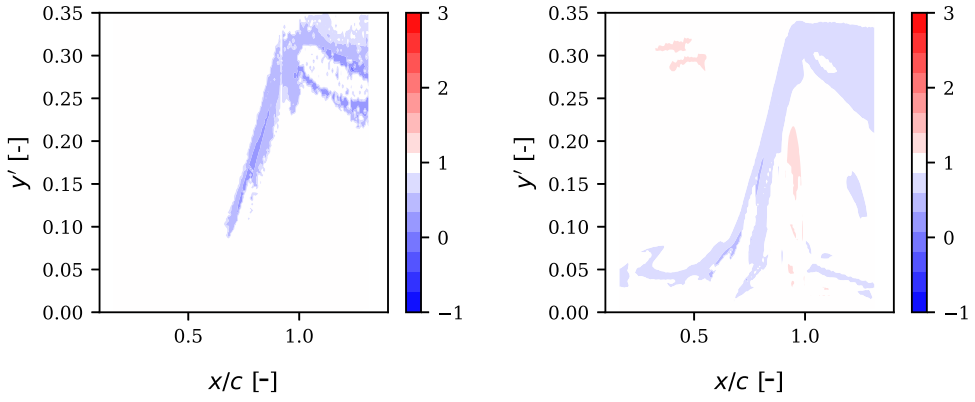
图13展示了模型预测的方差。红色等值线表示机器学习模型低于指定阈值、因而处于激活状态的区域。将方差与通过局部离群因子（LOF）进行新颖性检测分析所得的结果进行了比较；较低的分数表示点更接近训练数据集（36）。在远离壁面且 $y' > 0.1$的低 LOF 区域中，高斯过程专家和深度集成均处于激活状态。高斯过程专家的方差与该区域的 LOF 吻合良好：在 $x/c > 0.75$周围的低 LOF 带内，方差较低。低 LOF 带对应于分离泡上方的剪切层。

在 $y' < 0.1$ LOF 较低的区域，高斯过程专家并未预测出较低的方差。一个值得关注的点是$x/c = 0.69,\ y' = 0.06$, 该点的 LOF 较低，但高斯过程专家和深度集成的方差都较高。图13将方差分解为 σ_σ^* 和 σ_μ^*。高斯过程专家的 σ_σ^* 较低，这表明可以对基于各个独立案例训练的子模型所作的预测赋予较高置信度。然而， σ_μ^* 较高，这表明各训练案例预测的 β_c 值彼此不一致。深度集成在该点的 σ_σ^*较高，这表明训练 β_c 值在该点高度分散，进一步说明不同训练案例中该点的 β_c 存在一对多映射。图14a证实了这一点：高斯过程专家的各个子模型给出的预测置信度很高，但分散程度也很大，因此两个集成模型的置信度较低。图14b展示了另一个置信度较高的点的分布。在该点，高斯过程专家子模型预测的β_c 聚集在一起，因此两个集成模型的方差都较低。

图13d显示了基于各个 GPE 的先验长度尺度所得的 σ_σ^* 渐近值（即，对于远离训练数据集的测试点， σ_σ^* 会趋近于一个最大常数值）。此时， σ^* 的任何进一步增加都源于 β_c的一对多映射。因此，渐近 σ_σ^* 值为容差 $\bar{\sigma}$ 提供了一个上界，也是调节 $\bar{\sigma}$的良好起点。



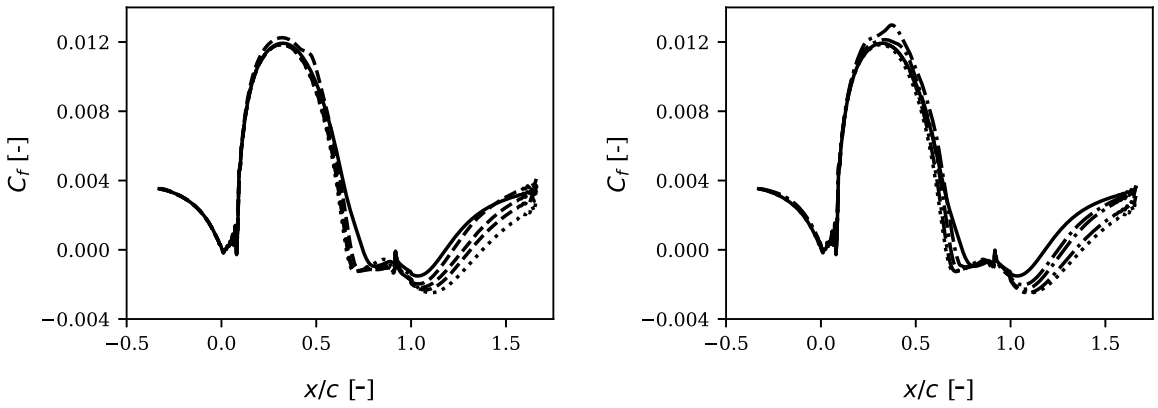
(a) 反演。 (b) 约束反演。



(c) 高斯过程专家。 (d) 深度集成。

图11：反演和预测的 β_c 场。使用缩放后的 y 坐标，其中 $y' = \sqrt{(y - y_{wall})}/c$。

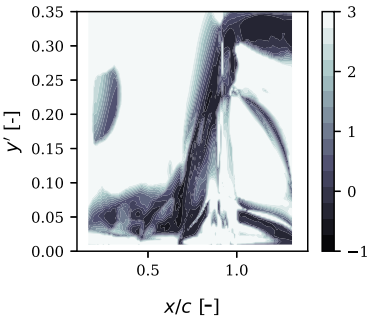
图13g显示，在存在数据的区域，深度集成能够学习训练数据中 β_c 的分布范围。然而，在局部离群因子较高、模型进行外推的区域， σ_σ^* 会降至较低值。这些区域中的总方差很大程度上取决于 σ_μ^*，而 σ_μ^*又取决于模型中初始权重的随机化。



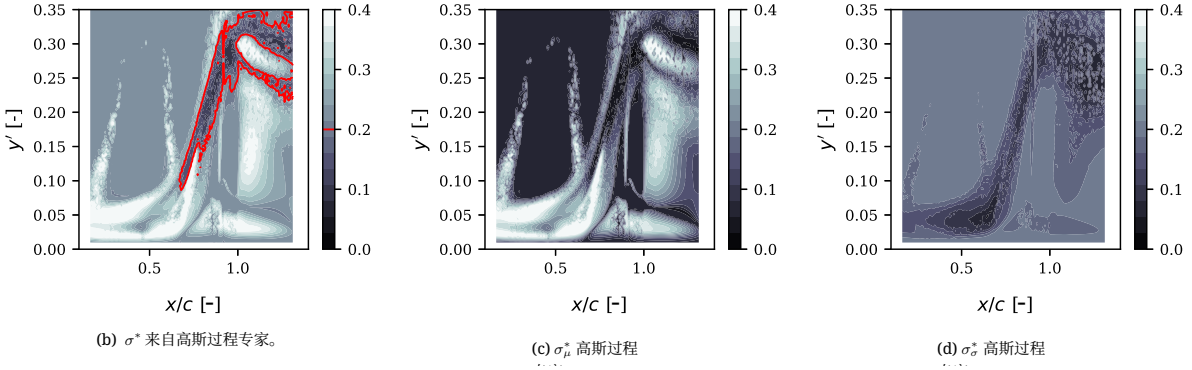
(a) $\bar{\sigma} = [0.15, 0.2, 0.25]$ 的高斯过程专家结果。

(b) $\bar{\sigma} = [0.05, 0.1, 0.15]$ 的深度集成结果。

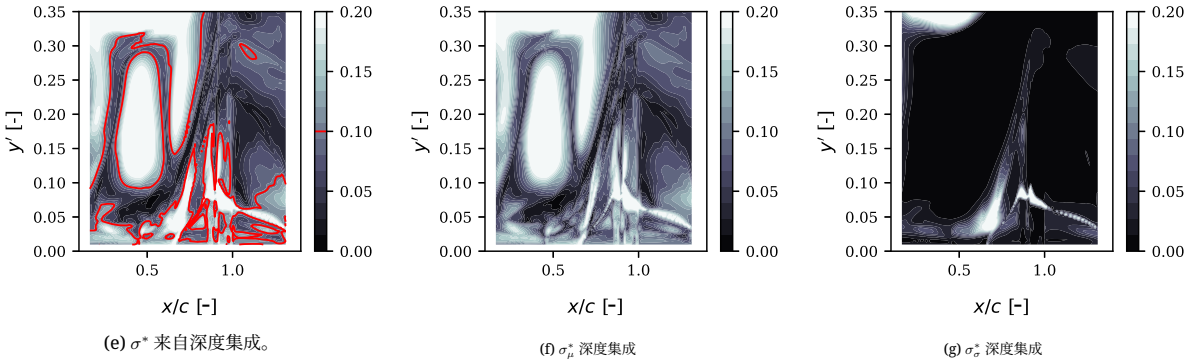
图12：壁面摩擦系数。约束反演（ ）、未校正（ ）、高斯过程专家（ ）、深度集成（ ）。$\bar{\sigma}$ 值较低时，结果更接近未校正结果。增大 $\bar{\sigma}$ 会使剖面更接近约束反演结果，但超过一定值后也会产生过冲。



(a) 局部离群因子。

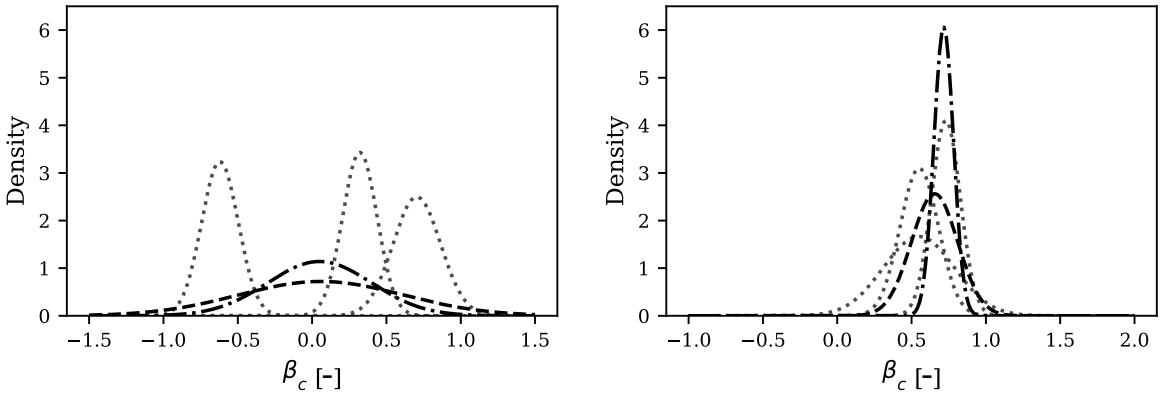


(b) σ^* 来自高斯过程专家。 (c) σ^*_μ 高斯过程 (d) σ^*_σ 高斯过程



(e) σ^* 来自深度集成。 (f) σ^*_μ 深度集成 (g) σ^*_σ 深度集成

图13：机器学习模型的外推检测与方差。(c) σ^*_μ 来自高斯过程专家，表示 β_c 在不同反演案例中的分布范围。(d) σ^*_σ 来自高斯过程专家，表示每个反演案例中的平均方差。数值较低表示测试特征接近训练特征，而数值较高则表明模型正在进行外推。(f) σ^*_μ 来自深度集成，表示子模型均值的分布范围，指示模型外推的程度。(g) σ^*_σ 来自深度集成，表示在所有反演案例中收集的训练数据中 β_c 的分布范围。

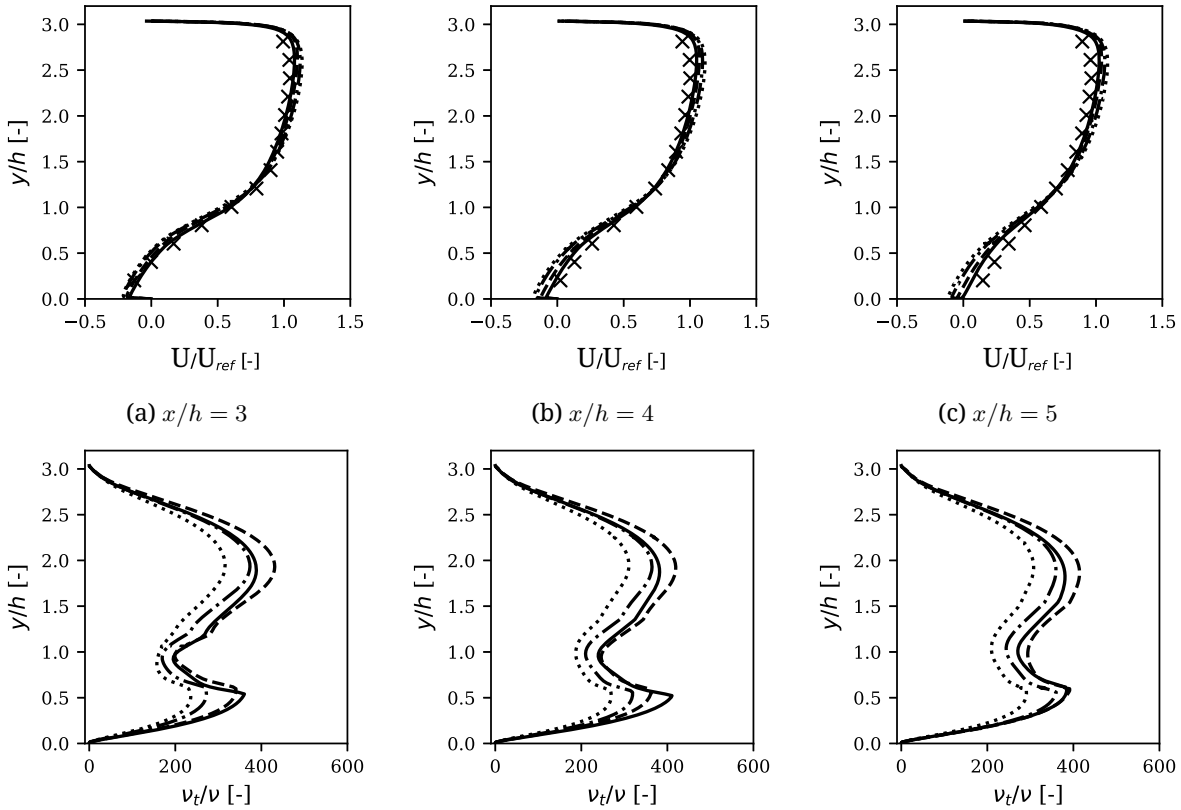


(a) $x/c = 0.69,\ \ y' = 0.06$。 (b) $x/c = 0.90,\ \ y' = 0.23$。

图 14：高斯输出。GPE 子模型输出（ ）、高斯过程专家集成（ ）、深度集成（ ）。

4.2 周期性山丘

选择了文献 (34) 中的周期性丘陵算例，因为它包含一个曲面上的大规模分离区域，与训练案例相似。因此，预期机器学习模型能够带来改进。与 h42 凸起算例类似，我们进行了场反演，以生成一组用于评估机器学习模型的参考结果（未用于训练）。模拟在山丘顶部周围设置了周期性边界条件，其中 $Re_h = 10595$基于山丘高度和顶部的体积平均速度。采用了一个由 300×180 构成的结构化块网格，壁面处的最大 y^+ 为 0.6。我们进行了无约束反演，以生成参考结果，并将文献 (34) 中大涡模拟的数据同化到区域 $0.46 < x/h < 4.6$ 和 $0 < y/h < 1.78$中。



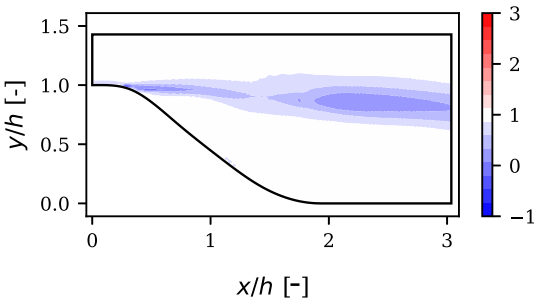
(a) $x/h = 3$ (b) $x/h = 4$ (c) $x/h = 5$

(d) $x/h = 3$ (e) $x/h = 4$ (f) $x/h = 5$

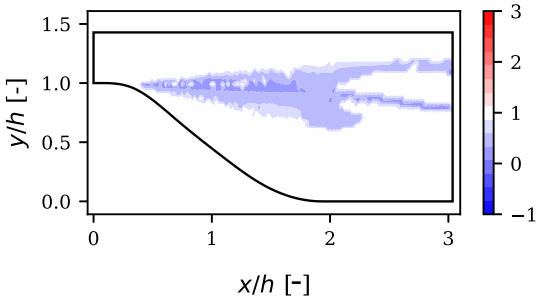
图15：周期性山丘的速度剖面和 ν_t 比率。大涡模拟（ ）、反演（ ）、未校正（ ）、高斯过程专家（ ）、深度集成（ ）。 ν_t 比率与反演结果进行比较，而非与大涡模拟结果比较。

在图16a的反演结果中，可以看到一个 β_c 降低区域，该区域从流动的分离点开始，沿剪切层向下游延伸。随后 k 的增加促进了剪切层的混合，这一点从速度剖面中可以看出：图15中的速度剖面分布更为分散。 ν_t 剖面呈现两个峰值，第一个峰值位于山丘顶部下方约 $x/h = 0.5$ 处，第二个峰值位于山丘上方区域约 $x/h = 2$ 处。正如图15所示，反演使第一个峰值更加明显。

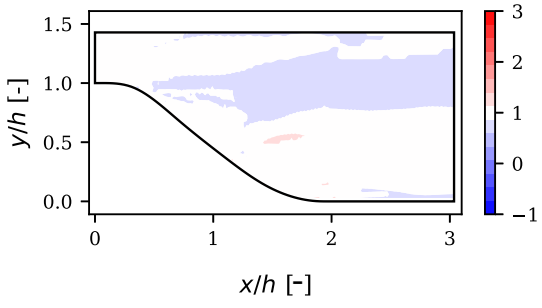
GPEs在反演场中的相似区域内处于激活状态，且修正源自距离壁面更远的位置。DEs的激活区域更大，大多数预测的 β_c 值约为1。DEs的 β_c 场中实际存在修正的区域与GPEs相似，但幅值更低。GPEs很好地预测了 ν_t 第一峰值的增加，而DEs的表现也相当不错。因此，与DEs相比，GPEs预测出了更好的速度场，其速度剖面接近反演结果。注意，在图16d中，自由来流区域的高斯过程集合方差相对较高，其中 $\mu^* \approx 1$，即模型预测 β_c处于非激活状态。鉴于训练数据的准备方式，这是预期现象；然而，在图16e中，DEs预测不确定性中的这一效应并不明显。



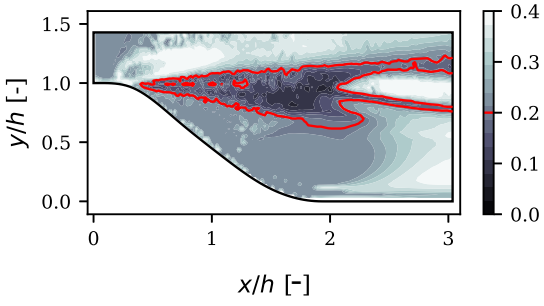
(a) 反转的 β_c 场。



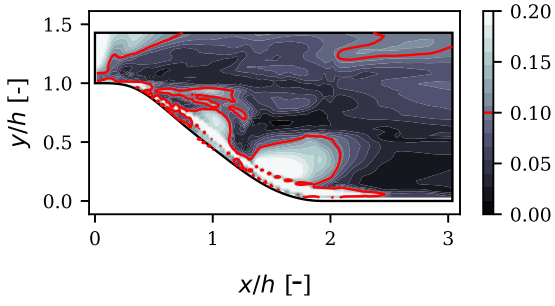
(b) 高斯过程专家 β_c 场。



(c) 深度集成 β_c 场。



(d) 高斯过程专家 σ^*。



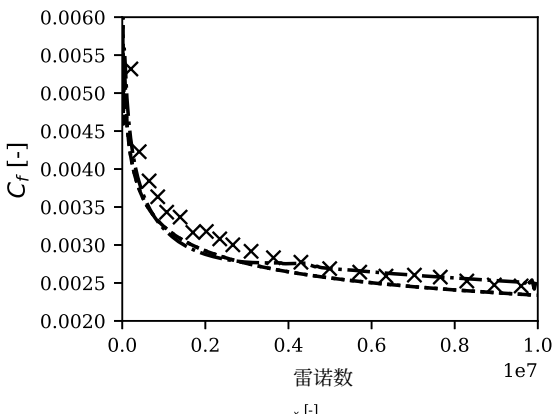
(e) 深度集成 σ^*。

图 16：周期性山丘预测。

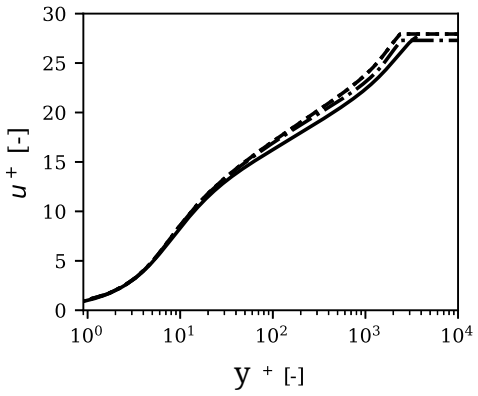
4.3 零压梯度平板

在此案例中，计算域从 $x = -0.33m$ 延伸至 $x = 2m$ ，壁面起始于 $x = 0m$。单位长度雷诺数为 5×10^6，在 $x = 0.97m$ 处测得的速度剖面对应于$Re_\theta = 1 \times 10^5$。所使用的网格为 545×770 ，在第一个单元处得到的最大 y^+ 为0.9。在图17a中，将壁面摩擦阻力与Wieghardt（37）的实验结果进行了比较。预计速度剖面应遵循壁面定律，并在图17b中使用Coles关系（38）作为理论参考解。大多数设计良好的湍流模型预计都能在此案例中取得良好结果；Rumsey等人（17）指出，在逆压梯度平板案例上训练的神经网络增强SA模型，在零压力梯度案例中预测出了虚假的 β_c ，导致壁面摩擦阻力剖面出现不期望的偏差。与训练案例相比，此案例不存在压力梯度和流线曲率，因此理想情况下机器学习模型不应处于激活状态。

高斯过程专家预测未发生激活，其壁面摩擦阻力和速度剖面与图 17 中的未修正模型重合。图 18 显示，深度集成预测沿壁面整个边界层中的 β_c 值，并且从 $Re_x = 0.3$ 开始， C_f 相应增加。如图 17b 所示，在对数律区，速度剖面也偏离了未修正模型。尽管深度集成的无量纲速度剖面趋向于理论关系，但深度集成的整体表现并不理想，因为它改变了壁面摩擦阻力剖面。

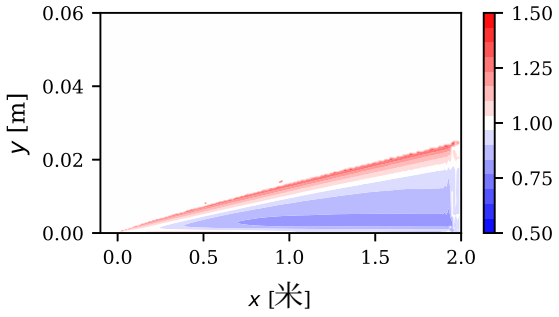


(a) 壁面摩擦阻力。

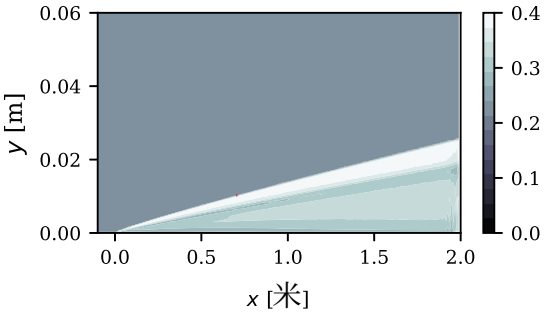


(b) $Re_x = 4.85e6$处的速度剖面。

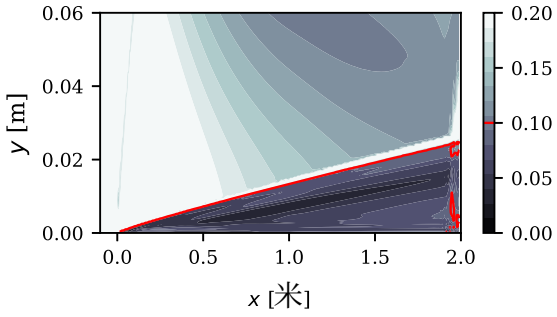
图 17：零压力梯度平板的结果。实验（）、理论（）、未校正（）、高斯过程专家（）、深度集成（）。未校正和高斯过程专家的结果重叠。



(a) 深度集成 β_c 场。



(b) 高斯过程专家 σ^*



(c) 深度集成 σ^*

图 18：零压力梯度平板案例中机器学习模型预测的变量。由于流场处处未校正，因此未显示来自 GPE 的 β_c 和红色等值线。

4.4 NACA 0012 翼型

该机器学习模型应用于附着流状态下的 NACA 0012 翼型，迎角为 $0°$、$10°$ 和 $15°$。雷诺数为 $Re_c = 6 \times 10^6$，最大网格为 $y^+ = 0.3$。表 1 显示了升力和阻力值。未修正模型能够很好地预测升力系数，所有 C_L 值与 Ladson（39）的实验数据相差均在 1.5% 以内。未修正模型的 C_D 值与实验值相差在 5% 以内，在零迎角时误差为 0.2%。因此，机器学习模型在此处不进行任何修正更为理想。

	实验		未校正		GPEs		DEs	
$\alpha[^\circ]$	C_D	C_L	C_D	C_L	C_D	C_L	C_D	C_L
0	0.00810	0	0.00809	0	0.00809	0	0.0084	0
10	0.01190	1.0586	0.01129	1.0739	0.01131	1.0774	0.0117	1.0687
15	0.01830	1.4938	0.01918	1.5099	0.01900	1.5320	0.0199	1.5030

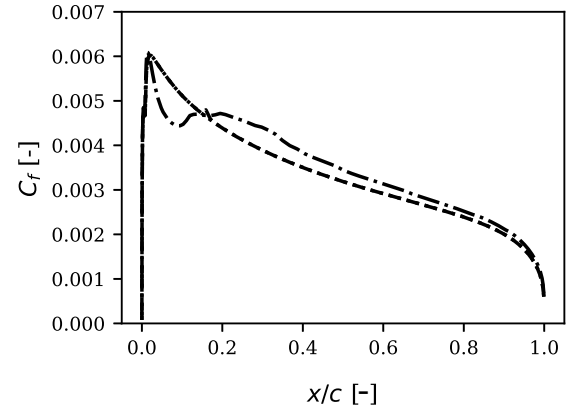
表 1：NACA0012翼型的升力和阻力系数与实验数据的比较。

	GPEs		DEs	
$\alpha[^\circ]$	ΔC_D[%]	ΔC_L[%]	ΔC_D[%]	ΔC_L[%]
0	0.08	–	3.34	–
10	0.12	0.33	3.38	-0.48
15	-0.91	1.47	3.71	-0.45

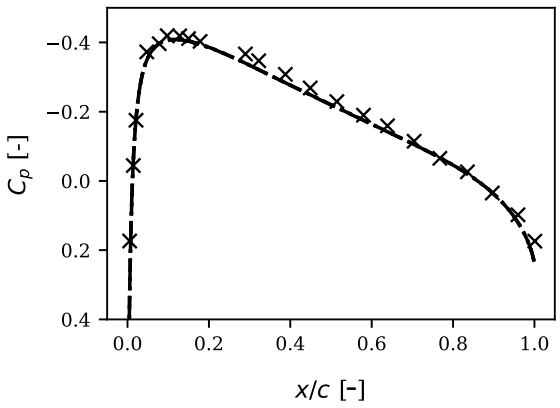
表 2：升 与未修正模型相比，机器学习模型的升力系数和阻力系数百分比差异 .

在 0° 和 10° 处，高斯过程专家的升力系数和阻力系数与未修正模型的数值非常接近（如表 2 所示，差异为 < 0.5%）。在 15° 处，其预测的 C_L 数值高出 1.5%。绘制该角度下的 C_p 剖面后，可以从图 19f 看出，这些剖面与 Gregory 和 O'Reilly (40) 的实验数据吻合良好。高斯过程专家的 C_L 增加了 1.5%，这是由吸力面上的 C_p 发生了极其轻微的偏移所致，在图 19f 中几乎看不出来。图 19 还展示了吸力面上的壁面摩擦阻力。正如表 2 所示，高斯过程专家的 C_f 剖面与未修正模型的剖面重合。

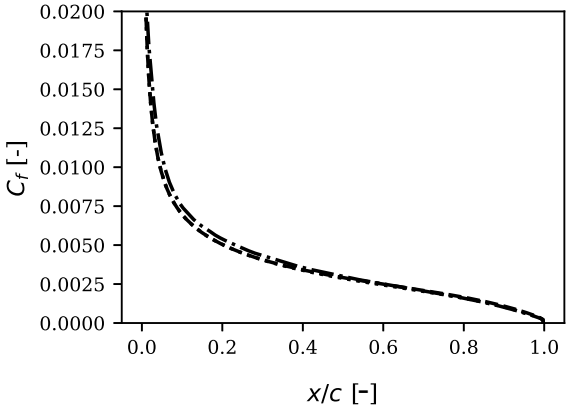
深度集成预测出的 C_f 剖面与未修正模型不同。在零迎角时，吸力峰附近的壁面摩擦阻力骤降，随后出现过冲。在 10° 和 15°时，深度集成预测的 C_f 从前缘到 $x/c = 0.3$都更高。不同的 C_f 剖面会导致深度集成的阻力系数相较于未修正模型最多变化约 4%（表 2）。



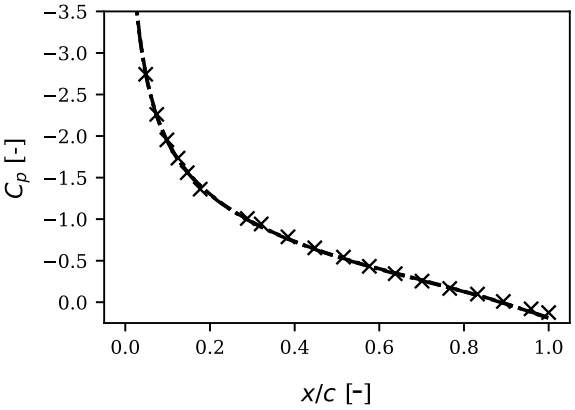
(a) C_f 在 0° 迎角下。



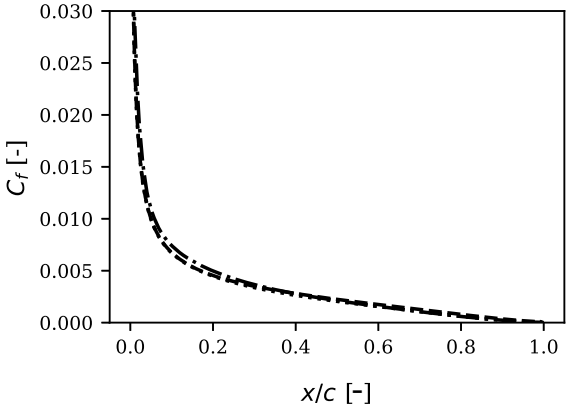
(b) C_p 在 0° 迎角下。



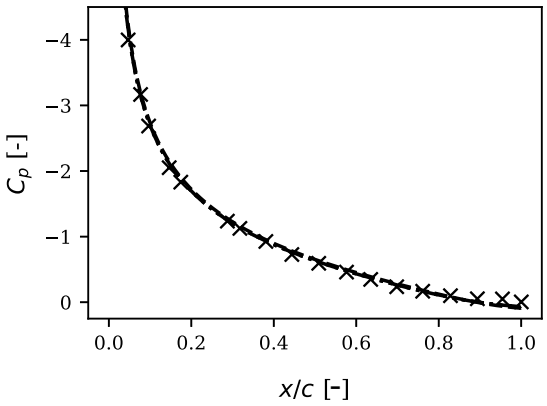
(c) C_f 在 10° 迎角下。



(d) C_p 在 10° 迎角下。



(e) C_f 在 15° 迎角下。



(f) C_p 在 15° 迎角下。

图 19：不同攻角下翼型吸力面的表面变量。实验（），未校正（），高斯过程专家（），深度集成（）。除 C_p 在 15° 迎角下的图外，所有图中高斯过程专家的结果均与未校正结果重叠。

5 结论

本研究采用了一种概率机器学习方法，结果表明，该方法能够减少不必要的修正，从而提高预测的稳健性。其实现方式是仅接受模型置信度高于特定阈值的预测。对于与训练数据集相似的案例（即 h42 凸起和周期性山丘），高斯过程专家和深度集成都能够识别流场中需要修正的区域。对于不包含相似流动物理特征的案例（即零压力梯度平板和 NACA 0012 翼型），高斯过程专家更能成功地使湍流模型恢复至基线状态。

高斯过程专家之所以成功，得益于其背后更丰富的概率理论。通过高斯过程集成，可以轻松结合异质来源，并从方差的不同组成部分中提取有用信息： σ_σ^*表示模型正在进行外推， σ_μ^* 表示训练数据集中的不确定性。

GPE集成的模块化方法使模型易于扩展。可以加入涵盖更多流动现象的训练案例，例如旋转流和壁面约束流。然而，当训练案例的数量超过一定规模（即数十个或数百个）时，如何将案例合理分配到数量有限的子模型中，以及如何构建这些子模型，就变得十分重要。如第 4.1 节所示，即使对于流动物理相似的案例，也存在非唯一映射。一个尚待探讨的问题是：当训练集涵盖范围更广的流动物理时，能否得到足够唯一的映射。这种行为部分取决于所选的特征集，而这也是值得深入研究的方向，有助于改进映射。

致谢

我们谨向 Vittorio Michelassi 致谢，感谢他提出的宝贵意见。我们也感谢 Christopher Rumsey 及其团队的讨论与反馈。

参考文献

[1] Michelassi V, Ling J. 涡轮机械应用中人工智能与高保真模拟的挑战与机遇：一种观点。《全球动力与推进学会期刊》。2021年5月；(May):1-14。[2] Hammond J, Pepper N, Montomoli F, Michelassi V. 涡轮机械计算流体力学中的机器学习方法：综述。《国际涡轮机械、推进与动力期刊》。2022；7(2)。[3] Chaouat B. 湍流模拟中混合雷诺平均纳维-斯托克斯/大涡模拟建模的最新进展。《流动、湍流与燃烧》。2017；99(2):279-327。[4] Duraisamy K, Iaccarino G, Xiao H. 数据时代的湍流建模。《流体力学年度评论》。2019；51(1)。[5]Brunton SL. 将机器学习应用于流体力学研究。《力学学报》。2021年12月；37(12):1718-26。[6] Frey Marioni Y, Ortiz EAdT, Cassinelli A, Montomoli F, Adami P, Vazquez R. 利用神经网络从直接数值模拟数据改进湍流建模的机器学习方法。《国际涡轮机械、推进与动力期刊》。2021；6(2)。[7] Ling J, Kurzawski A, Templeton J. 使用嵌入不变性的深度神经网络进行雷诺平均湍流建模。《流体力学杂志》。2016；807:155-66。[8] Weatheritt J, Sandberg RD. 将一种新型进化算法应用于雷诺平均纳维-斯托克斯应力-应变关系的代数修正。《计算物理学杂志》。2016；325:22-37。[9] Pope SB. 一种更通用的有效黏度假设。《流体力学杂志》。1975；72:331-40。[10] Schmelzer M, Dwight RP, Cinnella P. 利用稀疏符号回归发现代数雷诺应力模型。《流动、湍流与燃烧》。2020；104:579-603。[11] He C, Liu Y, Gan L. 基于连续伴随公式的湍流数据同化模型。《流体物理》。2018；30(10):105108。[12] Singh AP, Medida S, Duraisamy K. 机器学习增强的翼型湍流分离流预测建模。《美国航空航天学会期刊》。2017；55(7):2215-27。[13] Parish EJ, Duraisamy K. 一种利用场反演和机器学习进行数据驱动预测建模的范式。《计算物理学杂志》。2016；305:758-74。

[14] Ho J, West A. 场反演与机器学习在三维分离流湍流建模中的应用。载于：AIAA 2021 航空论坛；2021。第2903页。[15] McConkey R, Yee E, Lien FS. 机器学习辅助的各向异性映射在预测湍流建模中的泛化能力。《arXiv》；2022。[16] Sandberg RD, Zhao Y. 湍流与热通量模型开发中的机器学习：针对不同物理现象所面临挑战及当前进展的综述。《国际传热与流体流动杂志》。2022；95:108983。[17] Rumsey CL, Coleman GN, Wang L. 寻求数据驱动方法改进用于分离流的雷诺平均纳维–斯托克斯模型。；。[18] Menter F, Kuntz M, Langtry R. 剪切应力输运湍流模型的十年工业应用经验。《传热与传质》。2003年1月；4。[19] Matai R, Durbin P. 参数化凸起组上湍流的大涡模拟。《流体力学杂志》。2019；866:503–525。[20] Duraisamy K, Zhang ZJ, Singh AP. 利用数据驱动技术进行湍流与转捩建模的新方法。载于：第53届AIAA航空航天科学会议。佛罗里达州基西米；2015。。[21] Rumsey C. 流动控制模拟的成功与挑战。载于：第4届流动控制会议；2008。。[22] Wu JL, Xiao H, Paterson E. 用于增强湍流模型的物理信息机器学习方法：综合框架。《物理评论：流体》。2018；7(3):1-28。[23] Ling J, Jones R, Templeton J. 具有不变性属性的系统的机器学习策略。《计算物理学杂志》。2016；318:22-35。[24] Murphy KP. 概率机器学习：导论。麻省理工学院出版社；2022。[25] Blei DM, Kucukelbir A, McAuliffe JD. 变分推断：统计学家综述。《美国统计学会杂志》。2017年4月；112(518):859-77。[26] Hensman J, Matthews A, Ghahramani Z. 可扩展的变分高斯过程分类。arXiv；2014。[27] Ras mussen CE, Williams CKI.《机器学习中的高斯过程》。麻省理工学院出版社；2006。[28] Santner TJ, Williams BJ, Notz WI. 计算机试验的设计与分析。施普林格；2003。[29] Lakshminarayanan B, Pritzel A, Blundell C. 使用深度集成进行简单且可扩展的预测不确定性估计。arXiv；2016。[30] Ling J, Templeton J. 评估机器学习算法对雷诺平均纳维–斯托克斯不确定性较高区域的预测能力。《流体物理》。2015；27(8):085103。[31] Greenblatt D, Paschal KB, Yao CS, Harris J, Schaeffler NW, Washburn AE. 分离控制实验研究，第1部分：基线与稳态抽吸。《AIAA杂志》。2006；44(12):2820-30。[32] Bentaleb Y, Lardeau S, Leschziner MA. 圆角台阶处湍流边界层分离的大涡模拟。《湍流杂志》。2012；13:N4。[33] Lehmkuhl O, Rodr´ıguez I, Borrell R, P´erez-Segarra CD, Oliva A. Re = 3900时圆柱体尾流中的低频变化。《物理学杂志：会议系列》。2011年12月；318(4):042038。[34] Fr¨ohlich J, Mellen CP, Rodi W, Temmerman L, Leschziner MA. 对通道内具有流向周期性收缩段的分离流进行高分辨率大涡模拟。《流体力学杂志》。2005；526:19–66。[35] McConkey R, Yee E, Lien FS. 用于数据驱动湍流建模的精选数据集。《科学数据》。2021年9月；8(1):255。[36] Breunig MM, Kriegel HP, Ng RT, Sander J. 局部离群因子：识别基于密度的局部离群点。《SIGMOD记录》。2000年5月；29(2):93–104。[37] Wieghardt K, Tillmann W. 压力升高条件下的湍流摩擦层。美国国家航空咨询委员会；1951。NASA-TM- 1314。[38] Coles D. 湍流边界层中的尾流定律。《流体力学杂志》。1956；1(2):191–226。[39] Ladson CL. 马赫数和雷诺数独立变化对NACA 0012翼型剖面低速空气动力特性的影响。NASA兰利研究中心；1988。NASA-TM-4074。[40] Gregory N, O'Reilly C. NACA 0012翼型剖面的低速空气动力特性，包括模拟白霜的上表面粗糙度影响。航空研究委员会；1973。R&M 3726。