基于物理的机器学习与分层贝叶斯建模技术的融合

Omid Sedehi[1], Antonina M. Kosikova[2], Costas Papadimitriou[3], Lambros S. Katafygioti s[4]

摘要
机器学习（ML）已广泛用于物理系统的建模与预测。这些技术具有很强的表达能力，并且在观测数据集范围内进行插值时泛化能力良好。然而，黑箱模型的缺点在于，在盲测条件下表现不佳，因为它们未融入任何物理知识。基于物理的机器学习旨在解决这一问题：在保留机器学习技术数学灵活性的同时融入物理知识。为此，本文提出将基于力学的模型嵌入高斯过程（GP）模型的均值函数，并通过核机器刻画潜在的偏差。本文着重介绍一类特定的核函数，该函数与基于物理的模型关于输入和参数的梯度相关，并且与线性动力系统的精确自协方差函数相似。核函数的谱属性使我们能够考虑由物理模型设定错误引起的主要周期过程。然而，在长数据集的序贯处理中，核函数的平稳性是一大难题，可通过分层贝叶斯技术加以解决。这种实现方式还有助于降低计算成本，缓解高斯过程处理序贯数据时的可扩展性问题。通过数值示例和实验示例，本文展示了所提方法在结构动力学反问题中的潜在应用。

关键词：基于物理的机器学习；高斯过程；核协方差函数；切线核函数；分层贝叶斯；

缩略语
BIC 贝叶斯信息准则 ML 机器学习 GP 高斯过程
GWN 高斯白噪声 MMTE 多模态三角指数 MPV 最
概然值 NN

神经网络 PSD 功率谱密度 SHM 结构
健康监测 SVD 奇异值分解 RMS 均方根

符号
α 阻尼系数

[1] 博士后研究员，土木与环境工程系，香港科技大学，香港，电子邮箱：osedehi@connect.ust.hk[2] 博士生，土木与环境工程系，香港科技大学，香港，电子邮箱：akosikova@connect.ust.hk[3] 教授，机械工程系，色萨利大学，希腊沃洛斯，电子邮箱：costasp@uth.gr[4] 教授，土木与环境工程系，香港科技大学，香港，电子邮箱：katafygiotis.lambros@gmail.com

β	阻尼系数狄拉克δ函数高斯过程的潜在算子
$\delta(.)$	结构模态频率余弦核函数振幅噪声方差时间
$\varphi(.)$	差频率余弦核函数频率欧几里得距离 和基于
$\hat{\varpi}$	物理模型的模拟响应核协方差函数基于物理
σ_k^2	的切核函数相关长度MMTE核中的余弦函数
σ_n^2	数量带参数的基于物理的模型MMTE核函数
τ	的功率谱密度超参数向量预测误差过程
ω	MMTE核参数辅助变量条件均值向量超均值
ω_k	超协方差矩阵条件协方差矩阵 基于物理的模
ζ ζ M_{θ}	型参数阻尼矩阵梯度矩阵刚度矩阵带参数的
	完整协方差矩阵质量矩阵随机游走模型参数
$k_{\varepsilon}()$	初始条件和输入条件实测时程响应
$k_{\theta}()$	
ℓ_k	
m	
M_{θ}	$\boldsymbol{\theta}$
$\boldsymbol{\delta}$	
$\boldsymbol{\varepsilon}_{\boldsymbol{\delta}_i}$	
ϕ	
$\boldsymbol{\zeta}$	
$\boldsymbol{\mu}$	
$\boldsymbol{\mu}_{\boldsymbol{\theta}}$	$\boldsymbol{\theta}$
$\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$	$\boldsymbol{\theta}$
$\boldsymbol{\Sigma}$	
$\boldsymbol{\theta}$	
$\mathbf{C}$	
$\mathbf{J}$	
$\mathbf{K}$	ϕ
$\mathbf{M}$	
$\mathbf{Q}_{\boldsymbol{\delta}}$	
$\mathbf{x}$	
$\mathbf{y}$	

1. 引言 第一性原理 m

物理系统模型构成了计算科学与工程学的骨干，并已得到极大发展，广泛应用于各种实际应用。这些模型的性能直接取决于开发人员的知识与专业能力，以及他们能否准确描述所关注的物理系统。因此，不同科学与工程学科的研究人员一直在持续探索基于物理的模型的确认与验证，将其视为一项必要且活跃的任务。尽管传统的数据驱动模型

识别方法已经存在，但仍迫切需要提出具有更强稳健性和准确性的系统化工具[1,2]。

基于物理的机器学习（ML）近年来通过深度神经网络（NN）和高斯过程（GP）在计算物理学和力学领域兴起[3,4]。这些模型看起来可能截然不同，但近期研究揭示了它们之间也存在相互关系与联系[5,6]。在性能方面，值得注意的是，深度网络具有中等的泛化能力和较强的表达能力。然而，由于网络架构复杂，不确定性往往难以处理。此外，深度网络缺乏物理可解释性，且计算成本高昂，尤其是在与精细的基于物理的模型结合使用时。相比之下，高斯过程模型可以将基于物理的模型嵌入其均值过程中，并通过核协方差机器表达相关性和不确定性。在这方面，基于物理的模型与核机器可以通过一组共同的不确定参数进行参数化，并可通过最大似然估计和贝叶斯推断来识别这些参数[7,8]。

物理信息高斯过程具有显著优势，因为它们能够补偿基于物理的模型与测量数据之间的差异[9]。在此背景下，模型差异可能源于控制物理方程、初始条件、边界方程和物理参数设定不当。只有建立适当的核协方差函数，高斯过程模型才能捕捉这类不确定性[10]。在机器学习文献中，核机器常用的核函数包括指数核、指数-正弦核、多项式核、Matern 核、神经网络核和有理协方差函数 [7]。这类核还可以通过基本代数运算（如求和和相乘）组合起来，以模拟更复杂的关系[7,11]。这些核函数以底层随机过程的平稳性为基础，用于描述数据点之间的相关性；这种相关性取决于数据点在辅助潜在空间中的投影距离 [12]。这类辅助参数空间可以利用现有输入信息来设定，也可以从数据中推断得到。无论采用哪种方式，核函数的平稳性都使维纳-辛钦定理能够给出谱实现，从而有助于更好地理解其理论性质，并指导用户进行选择[7]。

乍看之下，将上述协方差函数用于识别基于物理的模型似乎是合理的。例如，高斯过程模型通常与平方指数核函数相关联，并基于辅助输入空间 [13–15] 中的欧几里得距离来确定数据点之间的相关性。在大多数实际应用中，这类高斯过程（称为克里金法）能够对测量数据训练集内的数据进行良好的插值。然而，在预测留出数据集时，它们给出的预测往往失真，无法反映模型差异 [16] 中呈现的主导趋势。问题在于，这类核机器无法利用任何物理知识来捕捉主要的相关模式和依赖关系。除此之外，模型差异不太可能是广义平稳的，因为它可能随时空域而变化，而前述核函数则充分依赖平稳性。因此，基于核的高斯过程解仍存在研究空白：它们尚未纳入特定应用的情境物理知识，也未考虑非平稳性。

尽管存在这些不足，高斯过程模型在结构健康监测（SHM）领域仍取得了成功 [17–19]。Yuen 和 Katafygiotis 使用高斯过程模型描述贝叶斯模态更新问题中的预测误差相关性，并将精确协方差矩阵与样本自相关函数进行比较 [20,21]。Papadimitriou 和 Lombaert 使用带有指数核函数的高斯过程，考虑近间距传感器的空间相关性以及感知信息的冗余性，从而避免传感器在某些位置聚集 [22]。Simoen 等人研究了众多类别的核函数，并建立了一种贝叶斯方法，以在准确性与核函数复杂性之间进行权衡 [10]。Kosikova 等人提出了一种新的模型选择方案，该方案还将预测质量纳入核函数选择过程 [16]。Avendano-Valencia 等人使用高斯过程模型描述动力

学属性随运行和环境条件的变化 [23]。同样，Zhu 和 Au 使用高斯过程描述模态属性与载荷、预测误差谱属性之间的关系 [24]。高斯过程的状态空间表示已被用于描述和识别未知输入力 [25,26]。Jiang 等人应用基于高斯过程的模型更新平面闸门的有限元模型 [27]。Gardner 等人应用基于核的领域自适应开展基于总体的结构健康监测 [28]。Ramancha 等人确定了用于线性结构识别的精确核协方差函数 [29]。Cross 等人使用 Matérn协方差函数创建几何一致的高斯过程，以描述声发射映射 [30]。随着新的灵感和应用进入研究视野，这一持续发展的趋势似乎将在未来几年延续下去。

本研究通过推进一类新型核协方差函数并考虑非平稳效应，拓展了基于物理的机器学习。本文从理论上研究了该核函数，并证明它由多模态三角指数（MMTE）协方差函数和物理驱动的正切协方差结构构成。通过这种方式，基于物理的模型的梯度信息被纳入协方差函数，从而在面对参数设定错误时增强稳健性。本文通过层次贝叶斯建模技术考虑非平稳效应，例如 [31–34]，，其中通过将测量数据划分为多个子分区来考虑参数的时间变化，而这些子分区内的平稳假设仍然成立。在阐述理论发现后，本文给出数值示例和实验示例，以展示其潜在应用及相较于经典贝叶斯方法的优点。

本文接下来介绍第 2 节，该节将新公式形式化并引入新的核函数。第3节研究所提框架的理论性质，并第4节提出了一种计算算法，用于总结详细的步骤流程。第5节通过数值示例和实验示例展示了所提方法，第6节得出结论并规划了我们的未来工作。

2. 所提方法

2.1. 物理信息高斯过程设 表示测得的时程

$\mathbf{y}\in\mathbb{R}^{nN_o}$ 响应，该响应来自具有已知初始条件和输入条件的机械系统，其中 是测量量的数量，是输入空间的维度，而 是时间增量的数量。N_o 目标是执行监督学习任务，并利用输入-输出对更新基于物理的模型及其未知参数。为此，我们首先定义：给定输入和参数，基于物理的模型 提供显式的函数关系，以生成物理响应，如下所示：

$$M_\theta:\begin{cases}\mathbb{R}^{nN_x+N_\theta} \mapsto \mathbb{R}^{N_o}\\ (\mathbf{x},\boldsymbol{\theta}) \mapsto f(\mathbf{x},\boldsymbol{\theta})\end{cases} \tag{1}$$

其中 是基于物理的模型的模拟响应。由于建模误差和测量噪声，模型响应会偏离测量值。此外，结构模型参数的选择 可能会显著影响模拟响应。为解决这些问题，我们采用一类物理信息核机器来描述输出 与输入 之间的概率关系：

$$\mathbf{y}\,|\,\mathbf{x},\boldsymbol{\theta},\phi \sim GP\left(\mathbf{y}\;\middle|\;\underbrace{f(\mathbf{x},\boldsymbol{\theta})}_{Physics-based\ Model},\ \underbrace{k_\varepsilon(\zeta,\zeta';\phi)}_{Kernel\ Machines}\right) \tag{2}$$

其中 是一个具有GP均值和 协方差函数的高斯过程（GP）； 是一个由未知参数 和辅助变量 描述的核机器。核函数

刻画了一个潜在点积欧几里得空间，该空间将数据点投影到辅助参数的多维空间中 并根据辅助变量的接近程度描述数据点之间的相关性。这一规范实质上意味着核函数可以写成

$$k_\varepsilon : \begin{cases} \mathbb{R}^{N_\zeta \times N_\zeta} \mapsto \mathbb{R} \\ (\zeta, \zeta') \mapsto k_\varepsilon(\zeta, \zeta'; \phi) = \left(\varphi(\zeta)\right)^T \varphi(\zeta') \end{cases} \tag{3}$$

其中 是一个作用于 上、用于描述数据点之间相关性 的潜在算子。需要指出的是，本文假设的数学表达式已知，因此无需获取或显式指定该算子。

根据公式 (2)，对于形如 的观测向量，其似然函数可表示为

$$p(\mathbf{y} \mid \mathbf{x}, \boldsymbol{\theta}, \phi) = \mathcal{N}(\mathbf{y} \mid f(\mathbf{x}, \boldsymbol{\theta}), \mathbf{K}(\phi)) \tag{4}$$

其中， 是由已知参数参数化的预测误差协方差矩阵，稍后将对其进行说明。

2.2. 切线核机器-物理模型提供

在 与之间建立了一对应关系，前提是已知。由于建模误差，无论选择何种 ，都可能导致模型与测量响应之间存在一定的残余失配。这类不确定性与其他误差来源混杂，但可以通过允许 按照参数化高斯分布变化来建模，具体如下

$$\boldsymbol{\theta} \mid \boldsymbol{\mu}_{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}} \sim \mathcal{N}\left(\boldsymbol{\theta} \mid \boldsymbol{\mu}_{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}\right) \tag{5}$$

其中， 和 分别是基于物理的模型参数的超均值和协方差。这些超参数 也应根据数据进行推断。为此，使用贝叶斯法则得到所有未知参数的后验分布，如下所示

$$\begin{aligned} p(\phi, \boldsymbol{\theta}, \boldsymbol{\mu}_{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}} \mid \mathbf{x}, \mathbf{y}) &\propto p(\mathbf{y} \mid \mathbf{x}, \boldsymbol{\theta}, \phi)\, p(\boldsymbol{\theta} \mid \boldsymbol{\mu}_{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})\, p(\boldsymbol{\mu}_{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})\, p(\phi) \\ &\propto \mathcal{N}(\mathbf{y} \mid f(\mathbf{x}, \boldsymbol{\theta}), \mathbf{K}(\phi))\, \mathcal{N}(\boldsymbol{\theta} \mid \boldsymbol{\mu}_{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}) \end{aligned} \tag{6}$$

在此方程中，the likelihood function 取自式 (4)，条件先验分布取自式 (5)；其他先验分布，即 和 视为均匀分布。

尽管式（6）中的后验分布具有一般性，但由于参数相互依赖且涉及的参数数量众多，其计算可能难以处理。因此，从计算角度而言，将某些参数作为干扰参数进行边缘化是有利的。为此，我们首先近似函数关系

通过在附近进行一阶泰勒级数展开，得到模型响应

$$f(\mathbf{x}, \boldsymbol{\theta}) \approx f(\mathbf{x}, \boldsymbol{\mu}_{\boldsymbol{\theta}}) + \mathbf{J}(\boldsymbol{\theta} - \boldsymbol{\mu}_{\boldsymbol{\theta}}) \tag{7}$$

其中 表示在展开点处计算得到的梯度矩阵。将这一线性化结果代入公式（7）后，可以发现模型参数 能够被显式积分消去，从而得到

$$p\left(\phi, \boldsymbol{\mu}_{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}} \mid \mathbf{x}, \mathbf{y}\right) \propto \int_{\boldsymbol{\theta}} \mathcal{N}(\mathbf{y} \mid f(\mathbf{x}, \boldsymbol{\theta}), \mathbf{K}(\phi))\, \mathcal{N}(\boldsymbol{\theta} \mid \boldsymbol{\mu}_{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})\, d\boldsymbol{\theta} \approx \mathcal{N}\left(\mathbf{y} \mid f(\mathbf{x}, \boldsymbol{\mu}_{\boldsymbol{\theta}}), \mathbf{J}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}\mathbf{J}^T + \mathbf{K}(\phi)\right) \tag{8}$$

This result is a direct consequence of Theorem 1, discussed in Appendix (A), which sim

在此方程中，表达式 是切线协方差矩阵，通过对物理参数作进一步参数化而纳入其中。当矩阵 为对角矩阵时，表达式可以简化为

$$\mathbf{J}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}\mathbf{J}^{T}=\sum_{j=1}^{N_{\theta}}\sigma_{\theta_j}^{2}\left[\frac{\partial f(\mathbf{x},\boldsymbol{\theta})}{\partial\theta_j}\frac{\partial f^{T}(\mathbf{x},\boldsymbol{\theta})}{\partial\theta_j}\right]_{\boldsymbol{\theta}=\boldsymbol{\mu}_{\boldsymbol{\theta}}} \tag{9}$$

其中 表示 j^{th} 的对角元素。该协方差矩阵通过以下切线核函数刻画数据点之间的相关性：

$$k_{\theta}(x_p,x_q;\{\boldsymbol{\mu}_{\boldsymbol{\theta}},\boldsymbol{\Sigma}_{\boldsymbol{\theta}}\})=\sum_{j=1}^{N_{\theta}}\sigma_{\theta_j}^{2}\left[\frac{\partial f(x_p,\boldsymbol{\theta})}{\partial\theta_j}\frac{\partial f^{T}(x_q,\boldsymbol{\theta})}{\partial\theta_j}\right]_{\boldsymbol{\theta}=\boldsymbol{\mu}_{\boldsymbol{\theta}}} \tag{10}$$

其中 和 分别表示当所关注的物理系统受到增量输入时，基于物理的模型返回的模型响应和； 是基于物理的切核函数，用于描述实例与 之间的相关性。在此基础上，可以突出切线核函数的以下特征：

- 数据点 p 与 q 之间的相关性随着物理响应相对于 的敏感度增加而增强，表示为 和 。• 切线核 仅取决于响应敏感度和超参数。这意味着 p^{th} 数据点的方差与响应敏感度的平方成正比 。
- 根据公式（10），不同自由度响应之间的互相关与其对应响应敏感度的乘积成正比。这意味着，对于固定的参数协方差矩阵， 敏感度更高的响应具有更大的不确定性，反之亦然。

2.3. 三角核机器 本文中的 MMTE 核函数

式（[16], ）中提出的函数用于描述数据点之间的相关性。该核函数是平方指数函数与余弦函数的混合，并加上各向同性方差以考虑潜在的测量噪声，如下所示：

$$k_{\varepsilon}(\zeta,\zeta';\phi)=\left[\sum_{k=1}^{m}\sigma_k^{2}\exp\left(-\frac{d^{2}(\zeta,\zeta')}{\ell_k^{2}}\right)\cos\left(\omega_k d(\zeta,\zeta')\right)\right]+\sigma_n^{2}\delta\left(d(\zeta,\zeta')\right) \tag{11}$$

where, is the amplitude of the cos, is the noise variance 是 指数函数的相关长度，是 余弦函数的频率，是 与 之间的欧几里得距离，而 是狄拉克δ函数。该核函数的定义

这意味着核参数可以汇集到 。

该定义意味着，协方差矩阵 嵌入了这些元素。在本文中，辅助参数 和 被视为相应的时刻，而距离度量 可以简化为时间距离 。因此，MMTE函数可以重写为

$$k_\varepsilon(\tau;\phi)=\left[\sum_{k=1}^{m}\sigma_{k,i}^2\exp\left(-\frac{\tau^2}{\ell_k^2}\right)\cos(\omega_k\tau)\right]+\sigma_n^2\delta(\tau) \tag{12}$$

与高斯白噪声（GWN）输入下线性动力系统的自协方差函数相似[35]。这使得我们能够计算上述核函数的傅里叶变换，从而得到其功率谱密度（PSD），有助于理解相关性是如何建模的。通过这种方式， 的功率谱密度可表示为

$$S_\varepsilon(\omega;\phi)=\frac{1}{2}\sum_{k=1}^{m}\sqrt{\pi}\sigma_k^2\ell_k\left[\exp\left(-\frac{\ell_k^2}{4}(\omega+\omega_k)^2\right)+\exp\left(-\frac{\ell_k^2}{4}(\omega-\omega_k)^2\right)\right]+\sigma_n^2 \tag{13}$$

此时，我们考虑一个数值示例，以在频域中直观展示该核函数。为此，令参数为 $m=2$，$\sigma_1^2=2$，$\sigma_2^2=8$，$\ell_1^2=5$，$\ell_2^2=5$，$\omega_1=2$，$\omega_2=10$，$\sigma_n^2=0.5$。图 1(a) 展示了相应的核函数，说明 MMTE核函数由两个指数衰减余弦函数组成。图 1(b) 展示了由公式（13）给出的相应功率谱密度函数，其中在 2 rad/s 和 10 rad/s 频率处出现两个峰值。基于这一示例可以得出， 当预测误差过程由少数几个混合的主导正弦函数组成时，MMTE 适合用于建模时间校正。

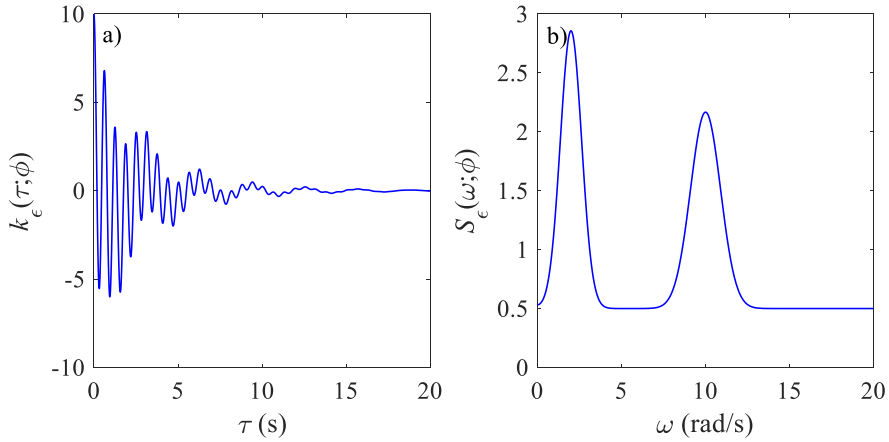

图 1。（a）MMTE 核协方差函数；（b）相应的功率谱密度函数

尽管该核函数具有理想的谱属性，但其内在的平稳性限制了其应用。下一节将通过分层贝叶斯技术解决这一问题。

3. 层次贝叶斯表述 3.1. 联合后验分布

设 $\mathbf{X}=[\mathbf{x}_1^T\ \mathbf{x}_2^T\ \dots\ \mathbf{x}_{N_D}^T]$ 和 $\mathbf{Y}=[\mathbf{y}_1^T\ \mathbf{y}_2^T\ \dots\ \mathbf{y}_{N_D}^T]$ 构成一个数据集，该数据集被划分为 N 个片段，其中 $\mathbf{x}_i$ 是已知输入， $\mathbf{y}_i$ 是测量得到的物理响应。目标与上一节相同，即通过这一多分区数据集推断基于物理的模型的未知参数，这是一项监督学习任务。前文提出的公式同样适用于解决这一问题；但是，当使用式（8）中的后验分布，根据每个单独的数据对 $(\mathbf{x}_i,\mathbf{y}_i)$，推断未知参数时，会发现未知参数及其分布似乎随分区而变化。这种变化与基于物理的模型的参数，以及 MMTE 核参数有关。因此，通过添加下标和 $\boldsymbol{\theta}_i$ 来引入参数向量，它们表示与 $\boldsymbol{\theta}_i$ 的对应关系

特定的一对。此外，参数 被视为相互独立，并按照与式（5）类似的方式通过以下先验分布进行参数化，从而得到

$$\boldsymbol{\theta}_i \mid \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0 \sim N(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0), \forall i = \{1,...,N_D\} \tag{14}$$

其中， 和 分别是分区特定参数的均值向量和协方差矩阵。

收集在 中的参数在不同数据分区之间有所变化。由于这种变异，除非考虑额外假设来约束这些参数在各分区间的变化，否则无法在分区之间进行知识迁移。为此，假设参数’s具有序列依赖性，并由一阶随机游走模型控制，表示为

$$\boldsymbol{\delta}_i = \boldsymbol{\delta}_{i-1} + \boldsymbol{\varepsilon}_{\boldsymbol{\delta}}, \quad \forall i = \{1,...,N_D\} \tag{15}$$

其中， 是均值为零、协方差矩阵为。因此，’s 的条件先验分布可写为

$$\boldsymbol{\delta}_i \mid \boldsymbol{\delta}_{i-1}, \mathbf{Q}_{\boldsymbol{\delta}} \sim N(\boldsymbol{\delta}_i \mid \boldsymbol{\delta}_{i-1}, \mathbf{Q}_{\boldsymbol{\delta}}), \ \forall i = \{1,...,N_D\} \tag{16}$$

Based on Eqs. (14), and (16), the joint prior

$$p(\{\boldsymbol{\theta}_i, \boldsymbol{\delta}_i\}_{i=1}^{N_D}, \mathbf{Q}_{\boldsymbol{\delta}}) = \left[\prod_{i=1}^{N_D} N(\boldsymbol{\theta}_i \mid \boldsymbol{\mu}_{\boldsymbol{\theta}_i}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}_i}) N(\boldsymbol{\delta}_i \mid \boldsymbol{\delta}_{i-1}, \mathbf{Q}_{\boldsymbol{\delta}})\right] p(\mathbf{Q}_{\boldsymbol{\delta}}) \tag{17}$$

其中， 是所有参数的联合先验分布， 是超参数的先验分布，而 可设为参数的先验估计。在作出这一假设后，根据贝叶斯法则，所有参数的联合后验分布如下：

$$p(\{\boldsymbol{\theta}_i, \boldsymbol{\delta}_i\}_{i=1}^{N_D}, \mathbf{Q}_{\boldsymbol{\delta}} \mid \mathbf{X}, \mathbf{Y}) \propto p(\mathbf{Y} \mid \{\boldsymbol{\theta}_i, \boldsymbol{\delta}_i\}_{i=1}^{N_D}, \mathbf{Q}_{\boldsymbol{\delta}}, \mathbf{X}) \, p(\{\boldsymbol{\theta}_i, \boldsymbol{\delta}_i\}_{i=1}^{N_D}, \mathbf{Q}_{\boldsymbol{\delta}}) \tag{18}$$

where the prior distribution 是包含所有数据分区的完整数据集的似然函数。根据公式（4），可表示如下：

$$p(\mathbf{Y} \mid \{\boldsymbol{\theta}_i, \boldsymbol{\delta}_i\}_{i=1}^{N_D}, \mathbf{Q}_{\boldsymbol{\delta}}, \mathbf{X}) = N\left(\mathbf{Y} \middle| \mathbf{f}(\mathbf{X}, \boldsymbol{\Theta}), \mathbf{K}(\boldsymbol{\Phi})\right) \tag{19}$$

其中，向量 和 分别是展开后的向量，包含所有分区对应的物理参数和核参数。注意，公式（19）中的高斯分布各元素表示为

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \mathbf{y}_3 \\ \vdots \\ \mathbf{y}_{N_D} \end{bmatrix}; \mathbf{f}(\mathbf{X}, \boldsymbol{\Theta}) = \begin{bmatrix} f(\mathbf{x}_1, \boldsymbol{\theta}_1) \\ f(\mathbf{x}_2, \boldsymbol{\theta}_2) \\ f(\mathbf{x}_3, \boldsymbol{\theta}_3) \\ \vdots \\ f(\mathbf{x}_{N_D}, \boldsymbol{\theta}_{N_D}) \end{bmatrix}; \mathbf{K}(\boldsymbol{\Phi}) = \begin{bmatrix} \mathbf{K}_1 & \mathbf{k}_{1,2} & \mathbf{k}_{1,3} & \cdots & \mathbf{k}_{1,N_D} \\ \mathbf{k}_{1,2}^T & \mathbf{K}_2 & \mathbf{k}_{2,3} & \cdots & \mathbf{k}_{2,N_D} \\ \mathbf{k}_{1,3}^T & \mathbf{k}_{2,3}^T & \mathbf{K}_3 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \mathbf{k}_{N_D-1,N_D} \\ \mathbf{k}_{1,N_D}^T & \mathbf{k}_{2,N_D}^T & \cdots & \mathbf{k}_{N_D-1,N_D}^T & \mathbf{K}_{N_D} \end{bmatrix} \tag{20}$$

In this equation, the elements of the covariance matrix are calculated based on the kernel func

化。这意味着元素 和 是根据核函数，其中核参数 会在时被替换

将公式（19）中的似然函数和公式（17）中的先验分布代入，可得联合后验分布：

$$
\begin{aligned}
p(\{\boldsymbol{\delta}_i,\boldsymbol{\theta}_i\}_{i=1}^{N_D},\mathbf{Q}_{\boldsymbol{\delta}}\,|\,\mathbf{X},\mathbf{Y}) &\propto N\left(\mathbf{Y}\middle|\mathbf{f}(\mathbf{X},\boldsymbol{\Theta}),\mathbf{K}(\boldsymbol{\Phi})\right)\left[\prod_{i=1}^{N_D}N(\boldsymbol{\theta}_i\,|\,\boldsymbol{\mu}_{\boldsymbol{\theta}_i},\boldsymbol{\Sigma}_{\boldsymbol{\theta}_i})N(\boldsymbol{\delta}_i\,|\,\boldsymbol{\delta}_{i-1},\mathbf{Q}_{\boldsymbol{\delta}})\right]p(\mathbf{Q}_{\boldsymbol{\delta}}) \\
&\propto N\left(\mathbf{Y}\middle|\mathbf{f}(\mathbf{X},\boldsymbol{\Theta}),\mathbf{K}(\boldsymbol{\Phi})\right)N(\boldsymbol{\Theta}\,|\,\boldsymbol{\mu}_{\boldsymbol{\Theta}},\boldsymbol{\Sigma}_{\boldsymbol{\Theta}})\left[\prod_{i=1}^{N_D}N(\boldsymbol{\delta}_i\,|\,\boldsymbol{\delta}_{i-1},\mathbf{Q}_{\boldsymbol{\delta}})\right]p(\mathbf{Q}_{\boldsymbol{\delta}})
\end{aligned}
\tag{21}
$$

其中，$\boldsymbol{\mu}_{\boldsymbol{\Theta}}=[\boldsymbol{\mu}_{\boldsymbol{\theta}_1}^T\ \cdots\ \boldsymbol{\mu}_{\boldsymbol{\theta}_{N_D}}^T]^T$ 以及 构成所有数据分区的超参数。与第2.2节中对单一分区表述类似，同时估计和 可能较为繁琐，因此对的从后验分布中进行边缘化是有益的，得到：

$$
p(\{\boldsymbol{\delta}_i\}_{i=1}^{N_D},\mathbf{Q}_{\boldsymbol{\delta}}\,|\,\mathbf{X},\mathbf{Y}) \propto p\left(\mathbf{Y}\,|\,\{\boldsymbol{\delta}_i\}_{i=1}^{N_D},\mathbf{Q}_{\boldsymbol{\delta}},\mathbf{X}\right)\left[\prod_{i=1}^{N_D}N(\boldsymbol{\delta}_i\,|\,\boldsymbol{\delta}_{i-1},\mathbf{Q}_{\boldsymbol{\delta}})\right]p(\mathbf{Q}_{\boldsymbol{\delta}}) \tag{22}
$$

其中，边缘似然函数定义为

$$
\begin{aligned}
p\left(\mathbf{Y}\,|\,\{\boldsymbol{\delta}_i\}_{i=1}^{N_D},\mathbf{Q}_{\boldsymbol{\delta}},\mathbf{X}\right) &= \int_{\boldsymbol{\Theta}} N\left(\mathbf{Y}\middle|\mathbf{f}(\mathbf{X},\boldsymbol{\Theta}),\mathbf{K}(\boldsymbol{\Phi})\right)N(\boldsymbol{\Theta}\,|\,\boldsymbol{\mu}_{\boldsymbol{\Theta}},\boldsymbol{\Sigma}_{\boldsymbol{\Theta}})d\boldsymbol{\Theta} \\
&\approx N\left(\mathbf{Y}\middle|\mathbf{f}(\mathbf{X},\boldsymbol{\mu}_{\boldsymbol{\Theta}}),\mathbf{K}(\boldsymbol{\Phi})+\mathbf{J}\boldsymbol{\Sigma}_{\boldsymbol{\Theta}}\mathbf{J}^T\right)
\end{aligned}
\tag{23}
$$

This integral is calculated using the results of Theorem 1 in the Appendix. 附录（A）中，对 周围的进行线性化。此外，需要注意的是，由 给出的 组成。在此基础上，边际似然函数可以表示为展开形式：

$$
\mathbf{J}_i = \left.\frac{\partial f(\boldsymbol{\theta}_i,\boldsymbol{\theta}_i)}{\partial \boldsymbol{\theta}_i^T}\right|_{\boldsymbol{\theta}_i=\boldsymbol{\mu}_{\boldsymbol{\theta}_i}}
$$

$$
p\left(\mathbf{Y}\,|\,\{\boldsymbol{\delta}_i\}_{i=1}^{N_D},\mathbf{Q}_{\boldsymbol{\delta}},\mathbf{X}\right)
= N\left(\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\\ \mathbf{y}_3\\ \vdots\\ \mathbf{y}_{N_D}\end{bmatrix}\middle|\begin{bmatrix}f(\mathbf{x}_1,\boldsymbol{\mu}_{\boldsymbol{\theta}_1})\\ f(\mathbf{x}_2,\boldsymbol{\mu}_{\boldsymbol{\theta}_2})\\ f(\mathbf{x}_3,\boldsymbol{\mu}_{\boldsymbol{\theta}_3})\\ \vdots\\ f(\mathbf{x}_{N_D},\boldsymbol{\mu}_{\boldsymbol{\theta}_{N_D}})\end{bmatrix},\begin{bmatrix}\mathbf{K}_1+\mathbf{J}_1\boldsymbol{\Sigma}_{\boldsymbol{\theta}_1}\mathbf{J}_1^T & \mathbf{k}_{1,2} & \mathbf{k}_{1,3} & \cdots & \mathbf{k}_{1,N_D}\\ \mathbf{k}_{1,2}^T & \mathbf{K}_2+\mathbf{J}_2\boldsymbol{\Sigma}_{\boldsymbol{\theta}_2}\mathbf{J}_2^T & \mathbf{k}_{2,3} & \cdots & \mathbf{k}_{2,N_D}\\ \mathbf{k}_{1,3}^T & \mathbf{k}_{2,3}^T & \mathbf{K}_3+\mathbf{J}_3\boldsymbol{\Sigma}_{\boldsymbol{\theta}_3}\mathbf{J}_3^T & \cdots & \vdots\\ \vdots & \vdots & \vdots & \ddots & \mathbf{k}_{N_D-1,N_D}\\ \mathbf{k}_{1,N_D}^T & \mathbf{k}_{2,N_D}^T & \cdots & \mathbf{k}_{N_D-1,N_D}^T & \mathbf{K}_{N_D}+\mathbf{J}_{N_D}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_{N_D}}\mathbf{J}_{N_D}^T\end{bmatrix}\right)
\tag{24}
$$

在该方程中， 数据分区的处理无法针对各个分区单独进行。由于未知参数的数量随数据分区数量线性增长，计算成本会逐渐变得更高。通过假设数据分区及其参数之间存在马尔可夫依赖来近似条件分布，可以解决这一问题，从而使序贯数据处理成为可能。这种简化意味着可以忽略相距较远的分区之间的相关性。对于 MMTE 核而言，这种简化是合理的，因为相关性会随距离呈指数衰减。基于这一假设，完整数据集似然可以简化为

$$p\left(\mathbf{Y}\,|\,\{\boldsymbol{\delta}_i\}_{i=1}^{N_D},\mathbf{Q}_{\boldsymbol{\delta}},\mathbf{X}\right)=\left[\prod_{k=2}^{N_D}p\left(\mathbf{y}_k\,|\,\{\mathbf{y}_i\}_{i=1}^{k-1},\{\boldsymbol{\delta}_i\}_{i=1}^{N_D},\mathbf{X}\right)\right]\times p\left(\mathbf{y}_1\,|\,\{\boldsymbol{\delta}_i\}_{i=1}^{N_D},\mathbf{X}\right)$$
$$\approx\left[\prod_{k=2}^{N_D}p\left(\mathbf{y}_k\,|\,\mathbf{y}_{k-1},\mathbf{x}_{k-1},\mathbf{x}_k,\boldsymbol{\delta}_{k-1},\boldsymbol{\delta}_k\right)\right]\times p\left(\mathbf{y}_1\,|\,\mathbf{x}_1,\boldsymbol{\delta}_1\right) \tag{25}$$
$$\approx\left[\prod_{k=2}^{N_D}\frac{p\left(\mathbf{y}_{k-1},\mathbf{y}_k\,|\,\mathbf{x}_{k-1},\mathbf{x}_k,\boldsymbol{\delta}_{k-1},\boldsymbol{\delta}_k\right)}{p\left(\mathbf{y}_{k-1}\,|\,\mathbf{x}_{k-1},\boldsymbol{\delta}_{k-1}\right)}\right]\times p\left(\mathbf{y}_1\,|\,\mathbf{x}_1,\boldsymbol{\delta}_1\right)$$

根据多元高斯分布的一般性质计算得出，从而得到 [26]。

From Eq. (24), the probability distributions required in Eq. (26) can be

$$p\left(\mathbf{y}_1\,|\,\mathbf{x}_1,\boldsymbol{\delta}_1\right)=N\left(\mathbf{y}_1\middle|\,f\left(\mathbf{x}_1,\boldsymbol{\mu}_{\boldsymbol{\theta}_1}\right),\mathbf{K}_1+\mathbf{J}_1\boldsymbol{\Sigma}_{\boldsymbol{\theta}_1}\mathbf{J}_1^T\right) \tag{26}$$
$$p\left(\mathbf{y}_{k-1}\,|\,\mathbf{x}_{k-1},\boldsymbol{\delta}_{k-1}\right)=N\left(\mathbf{y}_{k-1}\middle|\,f\left(\mathbf{x}_{k-1},\boldsymbol{\mu}_{\boldsymbol{\theta}_{k-1}}\right),\mathbf{K}_{k-1}+\mathbf{J}_{k-1}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_{k-1}}\mathbf{J}_{k-1}^T\right) \tag{27}$$
$$p\left(\mathbf{y}_{k-1},\mathbf{y}_k\,|\,\mathbf{x}_{k-1},\mathbf{x}_k,\boldsymbol{\delta}_{k-1},\boldsymbol{\delta}_k\right)=N\left(\begin{bmatrix}\mathbf{y}_{k-1}\\ \mathbf{y}_k\end{bmatrix}\middle|\begin{bmatrix}f\left(\mathbf{x}_{k-1},\boldsymbol{\mu}_{\boldsymbol{\theta}_{k-1}}\right)\\ f\left(\mathbf{x}_k,\boldsymbol{\mu}_{\boldsymbol{\theta}_k}\right)\end{bmatrix},\begin{bmatrix}\mathbf{K}_{k-1}+\mathbf{J}_{k-1}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_{k-1}}\mathbf{J}_{k-1}^T & \mathbf{k}_{k-1,k}\\ \mathbf{k}_{k-1,k}^T & \mathbf{K}_k+\mathbf{J}_k\boldsymbol{\Sigma}_{\boldsymbol{\theta}_k}\mathbf{J}_k^T\end{bmatrix}\right) \tag{28}$$

Based on the theorem 2 of Appendix (C), the likelihood function can be transformed into a Markov 形式，表示为

$$p\left(\mathbf{Y}\,|\,\{\boldsymbol{\delta}_i\}_{i=1}^{N_D},\mathbf{Q}_{\boldsymbol{\delta}},\mathbf{X}\right)=\prod_{i=1}^{N_D}N\left(\mathbf{y}_i\middle|\,\boldsymbol{\mu}_{i|i-1},\boldsymbol{\Sigma}_{i|i-1}\right) \tag{29}$$

其中 和 分别是条件均值向量和协方差矩阵，给定 时如下：

$$i\geq 2$$
$$\boldsymbol{\mu}_{i|i-1}=f\left(\mathbf{x}_i;\boldsymbol{\mu}_{\boldsymbol{\theta}_i}\right)+\mathbf{k}_{i-1,i}^T\left(\mathbf{K}_{i-1}+\mathbf{J}_{i-1}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_{i-1}}\mathbf{J}_{i-1}^T\right)^{-1}\left(\mathbf{y}_{i-1}-f\left(\mathbf{x}_{i-1};\boldsymbol{\mu}_{\boldsymbol{\theta}_{i-1}}\right)\right) \tag{30}$$
$$\boldsymbol{\Sigma}_{i|i-1}=\left(\mathbf{K}_i+\mathbf{J}_i\boldsymbol{\Sigma}_{\boldsymbol{\theta}_i}\mathbf{J}_i^T\right)-\mathbf{k}_{i-1,i}^T\left(\mathbf{K}_{i-1}+\mathbf{J}_{i-1}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_{i-1}}\mathbf{J}_{i-1}^T\right)^{-1}\mathbf{k}_{i-1,i} \tag{31}$$

Based on this simplified likelihood function, the posterior分布可改写为

$$p\left(\{\boldsymbol{\delta}_i\}_{i=1}^{N_D},\mathbf{Q}_{\boldsymbol{\delta}}\,|\,\mathbf{X},\mathbf{Y}\right)=\left[\prod_{i=1}^{N_D}N(\mathbf{y}_i\middle|\,\boldsymbol{\mu}_{i|i-1},\boldsymbol{\Sigma}_{i|i-1})N(\boldsymbol{\delta}_i\,|\,\boldsymbol{\delta}_{i-1},\mathbf{Q}_{\boldsymbol{\delta}})\right]p(\mathbf{Q}_{\boldsymbol{\delta}}) \tag{32}$$

需要指出的是，由 和分别给出均值和协方差。在此，后验分布具有马尔可夫结构，因为每个分区的参数和数据仅取决于其前序对应项。该性质将有利于开发序贯计算算法。

3.2. 图形表示 th 的图形表示

所提出的概率模型采用板式记号展示于图2中 [37]。矩形板表示每个数据分区，并包含分区特定参数。被视为跨分区固定的参数，即 ，和 ，未纳入该板中，而是以紫色圆圈表示。白色圆圈表示未知参数，灰色圆圈表示各个数据分区。箭头表示参数之间的条件依赖关系，弯曲箭头表示不同分区间参数的序贯依赖关系。如图所示，数据分区和数据特定参数的马尔可夫结构，即 ，，和 ，依赖于其对应的 分区，如弯曲箭头所示。

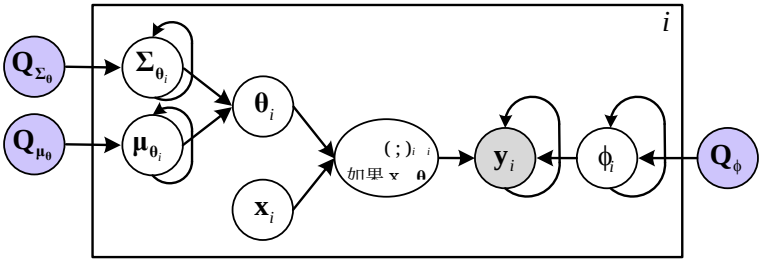

图2. 所提出的层次概率模型的图形表示

3.3. 后验预测分布 一旦后验分布计算完成

定后，可以预测未观测响应。要实现这一目标，需要利用观测数据预测未观测响应的统计属性。令表示新分区的参数，这些参数应从 个分区中推断得出。根据全概率原理，可以由以下内容计算 的后验预测分布

$$p(\boldsymbol{\delta}_{N_D+1}\,|\,\mathbf{X},\mathbf{Y}) = \int_{\mathbf{Q}_{\boldsymbol{\delta}}}\int_{\boldsymbol{\delta}_{N_D}} N(\boldsymbol{\delta}_{N_D+1}\,|\,\boldsymbol{\delta}_{N_D},\mathbf{Q}_{\boldsymbol{\delta}})\,p(\boldsymbol{\delta}_{N_D},\mathbf{Q}_{\boldsymbol{\delta}}\,|\,\mathbf{X},\mathbf{Y})\,d\boldsymbol{\delta}_{N_D}\,d\mathbf{Q}_{\boldsymbol{\delta}} \tag{33}$$

其中，分布 应通过适当的边缘化根据公式（32）进行计算，具体内容将在第4节中讨论。在该公式中， 和的不确定性会传播，以预测未知参数。随后，未观测响应可计算为

$$p(\mathbf{y}_{N_D+1}\,|\,\mathbf{x}_{N_D+1},\mathbf{X},\mathbf{Y}) = \int_{\boldsymbol{\delta}_{N_D+1}} p(\mathbf{y}_{N_D+1}\,|\,\mathbf{x}_{N_D+1},\boldsymbol{\delta}_{N_D+1})\,p(\boldsymbol{\delta}_{N_D+1}\,|\,\mathbf{X},\mathbf{Y})\,d\boldsymbol{\delta}_{N_D+1} \tag{34}$$

其中， 后验预测分布表示为，其中，均值和协方差由公式（30-31）给出。在这些公式中，对、和的多维空间进行积分的计算代价很高，因此将在第4节中提出具体策略。

4. 计算算法 在贝叶斯方

法中，后验分布的精确计算通常难以实现，因此有必要建立适当的近似方法。由公式（32）得到的后验分布同样存在这一问题，其中涉及大量参数，并且参数数量随分区数量线性增长。因此，如图2所示，由于参数之间存在复杂的依赖关系，采用抽样方法可能是一项繁琐的工作。

本文采用一种优化策略，通过忽略 的识别不确定性来近似后验分布。对于大多数实际情况，这种策略都是合理的，因为建模误差较大，使得分区间变异远大于识别不确定性 [38]。这样，式（32）中的后验分布便可

近似为

$$p\left(\{\boldsymbol{\delta}_i\}_{i=1}^{N_D},\mathbf{Q}_{\boldsymbol{\delta}}\,|\,\mathbf{X},\mathbf{Y}\right) \approx \left[\prod_{i=1}^{N_D}\delta(\boldsymbol{\delta}_i-\hat{\boldsymbol{\delta}}_i)N(\boldsymbol{\delta}_i\,|\,\boldsymbol{\delta}_{i-1},\mathbf{Q}_{\boldsymbol{\delta}})\right]p(\mathbf{Q}_{\boldsymbol{\delta}}) \tag{35}$$

其中 表示最概然值（MPV），而 是以 为中心的狄拉克δ函数。概率模型的马尔可夫结构提供了求得.....的最概然值的机会

$\boldsymbol{\delta}_i$ 依次最小化负对数似然函数，，对于每个，由下式给出 $\boldsymbol{\delta}$ =− $|\boldsymbol{\mu}_{|-1}\ \boldsymbol{\Sigma}_{|-1}\hat{}\ ()\ln(,)_{iiii}$ LNy

$\boldsymbol{\delta}_i$

$$L(\boldsymbol{\delta}_i)=\frac{1}{2}\ln\left|\hat{\boldsymbol{\Sigma}}_{i|i-1}\right|+\frac{1}{2}(\mathbf{y}_i-\hat{\boldsymbol{\mu}}_{i|i-1})^T\hat{\boldsymbol{\Sigma}}_{i|i-1}^{-1}(\mathbf{y}_i-\hat{\boldsymbol{\mu}}_{i|i-1})+cte. \tag{36}$$

其中 表示矩阵的行列式； 表示常数项；均值 和协方差矩阵 由公式（30-31）得出， 而参数设为其最可能值，这些值是在前一步分析中得到的。

$\boldsymbol{\delta}_{i-1}$

一旦获得 's 的最可能值（MPVs），对于 超参数 的后验分布可由式（35）计算得到：

$$p(\mathbf{Q}_{\boldsymbol{\delta}}\,|\,\mathbf{Y})\propto\left[\prod_{i=1}^{N_D}N(\hat{\boldsymbol{\delta}}_i\,|\,\hat{\boldsymbol{\delta}}_{i-1},\mathbf{Q}_{\boldsymbol{\delta}})\right]p(\mathbf{Q}_{\boldsymbol{\delta}}) \tag{37}$$

通过忽略 的识别不确定性即可得到该分布。考虑到 服从均匀分布，可由该方程通过最大化式（37）计算 的最可能值，适用于 ；也可以通过最小化其负对数来计算，具体如下：

$$L(\mathbf{Q}_{\boldsymbol{\delta}})=-\ln p(\mathbf{Q}_{\boldsymbol{\delta}}\,|\,\mathbf{Y})=\frac{N_D}{2}\ln|\mathbf{Q}_{\boldsymbol{\delta}}|+\frac{1}{2}\left[\sum_{i=1}^{N_D}(\hat{\boldsymbol{\delta}}_i-\hat{\boldsymbol{\delta}}_{i-1})^T\mathbf{Q}_{\boldsymbol{\delta}}^{-1}(\hat{\boldsymbol{\delta}}_i-\hat{\boldsymbol{\delta}}_{i-1})\right]+cte. \tag{38}$$

其中，由式（37）计算得到。对该表达式关于 求导，可得：

$$\frac{\partial L(\mathbf{Q}_{\boldsymbol{\delta}})}{\partial\mathbf{Q}_{\boldsymbol{\delta}}}=\frac{N_D}{2}\mathbf{Q}_{\boldsymbol{\delta}}^{-1}-\frac{1}{2}\mathbf{Q}_{\boldsymbol{\delta}}^{-1}\left[\sum_{i=1}^{N_D}(\hat{\boldsymbol{\delta}}_i-\hat{\boldsymbol{\delta}}_{i-1})(\hat{\boldsymbol{\delta}}_i-\hat{\boldsymbol{\delta}}_{i-1})^T\right]\mathbf{Q}_{\boldsymbol{\delta}}^{-1} \tag{39}$$

Setting this equation equal to zero provides:

$$\hat{\mathbf{Q}}_{\boldsymbol{\delta}}=\frac{1}{N_D}\sum_{i=1}^{N_D}(\hat{\boldsymbol{\delta}}_i-\hat{\boldsymbol{\delta}}_{i-1})(\hat{\boldsymbol{\delta}}_i-\hat{\boldsymbol{\delta}}_{i-1})^T \tag{40}$$

其中，是协方差矩阵 的最可能值。

与式（37）中引入的近似类似，式（33）的后验预测分布可写为由于该分布为高斯分布，因此可以直接从中抽样，以预测未来输入下的响应。为此，根据式（34），可写为：

$$p(\mathbf{y}_{N_D+1}\,|\,\mathbf{x}_{N_D+1},\mathbf{X},\mathbf{Y})\approx\int\int p(\mathbf{y}_{N_D+1}\,|\,\boldsymbol{\delta}_{N_D+1},\boldsymbol{\mu}_{N_D+1|N_D},\boldsymbol{\Sigma}_{N_D+1|N_D})\,p(\boldsymbol{\delta}_{N_D+1}\,|\,\boldsymbol{\delta}_{N_D},\mathbf{Q}_{\boldsymbol{\delta}})\,p(\boldsymbol{\delta}_{N_D},\mathbf{Q}_{\boldsymbol{\delta}}\,|\,\mathbf{X},\mathbf{Y})\,d\boldsymbol{\delta}_{N_D}\,d\mathbf{Q}_{\boldsymbol{\delta}}\,d\boldsymbol{\mu}_{N_D+1|N_D}\,d\boldsymbol{\Sigma}_{N_D+1|N_D}\,d\boldsymbol{\delta}_{N_D+1}$$

$$\approx\int\int N\left(\mathbf{y}_{N_D+1}\,\middle|\,\boldsymbol{\mu}_{N_D+1|N_D},\boldsymbol{\Sigma}_{N_D+1|N_D}\right)N(\boldsymbol{\delta}_{N_D+1}\,|\,\hat{\boldsymbol{\delta}}_{N_D},\hat{\mathbf{Q}}_{\boldsymbol{\delta}})\,d\boldsymbol{\delta}_{N_D+1}\,p(\boldsymbol{\delta}_{N_D},\mathbf{Q}_{\boldsymbol{\delta}}\,|\,\mathbf{X},\mathbf{Y})\,d\boldsymbol{\delta}_{N_D}\,d\mathbf{Q}_{\boldsymbol{\delta}}\,d\boldsymbol{\mu}_{N_D+1|N_D}\,d\boldsymbol{\Sigma}_{N_D+1|N_D} \tag{41}$$

$$p(\mathbf{y}_{N_D+1}\,|\,\mathbf{x}_{N_D+1},\mathbf{X},\mathbf{Y})=\int_{\boldsymbol{\delta}_{N_D+1}}N\left(\mathbf{y}_{N_D+1}\,\middle|\,\boldsymbol{\mu}_{N_D+1|N_D},\boldsymbol{\Sigma}_{N_D+1|N_D}\right)N(\boldsymbol{\delta}_{N_D+1}\,|\,\hat{\boldsymbol{\delta}}_{N_D},\hat{\mathbf{Q}}_{\boldsymbol{\delta}})\,d\boldsymbol{\delta}_{N_D+1}$$

$$\approx\frac{1}{N_s}\sum_{m=1}^{N_s}N\left(\mathbf{y}_{N_D+1}\,\middle|\,\boldsymbol{\mu}_{N_D+1|N_D}^{(m)},\boldsymbol{\Sigma}_{N_D+1|N_D}^{(m)}\right) \tag{42}$$

其中 和 分别是均值向量和协方差矩阵的样本，其中 替换为其最可能值，而 则替换为从 中抽取的样本。已知该高斯混合分布具有以下统计矩 [31]：

$$\mathbb{E}\left[\mathbf{y}_{N_D+1}\right]=\frac{1}{N_s}\sum_{m=1}^{N_s}\boldsymbol{\mu}_{N_D+1|N_D}^{(m)} \tag{43}$$

$$\mathbb{COV}\left[\mathbf{y}_{N_D+1}\right]=\frac{1}{N_s}\left[\sum_{m=1}^{N_s}\left(\boldsymbol{\mu}_{N_D+1|N_D}^{(m)}\boldsymbol{\mu}_{N_D+1|N_D}^{(m),T}+\boldsymbol{\Sigma}_{N_D+1|N_D}^{(m)}\right)\right]-\frac{1}{N_s^2}\left[\sum_{m=1}^{N_s}\boldsymbol{\mu}_{N_D+1|N_D}^{(m)}\right]\left[\sum_{m=1}^{N_s}\boldsymbol{\mu}_{N_D+1|N_D}^{(m)}\right]^T \tag{44}$$

其中 和 分别是未观测响应的均值和协方差，。

算法1给出了一套系统的方法，用于处理所提出概率模型的计算。该方法首先逐步求出每个数据分区的分区特定参数的最可能值（MPVs）。然后，继续计算和预测未测量结构响应的最可能值。在该算法中，计算行列式和矩阵乘法是第4行与第5行之间必不可少的关键步骤。由于在最大化的每一步都必须重复执行该过程，因此它占据了相当大的计算量。这个问题被称为高斯过程的“可扩展性”，意味着计算成本会随完整协方差矩阵的维度呈立方增长，即 [39]。然而，本研究采用的数据分区在很大程度上减轻了计算负担，因为较长时程数据可以按较短片段分别处理。这种实现方式实质上意味着计算成本可以从 降低至，从而显著节省计算资源。

此外，当每个分区中的数据点 数量 增加时，计算其行列式和逆矩阵可能会遇到病态性。为解决这一问题，可对进行奇异值分解（SVD）以截去冗余维度并降低协方差矩阵的维数。为确定合适的截断程度，可将奇异值衰减的斜率与最大奇异值的预定义容差（0.1–1%）进行比较。

算法1。计算未知参数的最可能值并预测未测量响应 1：设置、由、和组成的初始值。2：对于 （数据分区）{3：通过最小化由式（36）给出的对数似然函数，求得最可能值 4：}结束循环

5：通过表达式 近似求得的最可能值。6：从分布 中抽取样本

.7：使用公式（43-44）计算未观测响应的均值和协方差。

5. 示例

5.1. IASE-ASCE基准问题 结构健康监测基准结构用于

用于展示所提方法。该结构的尺寸和力学性能可在文献中查阅 [40]。如图3(a)所示，该结构为钢支撑框架，共四层，楼面铺设混凝土楼板。它受到作用于 4th 楼层沿 *x* 方向的集中力。动力荷载被视为均值为零、标准差为10kN的高斯白噪声过程。对于该结构，楼板的面外变形显著小于面内位移和转动。此外，与弯曲变形和剪切变形相比，柱的轴向伸长可以忽略。因此，可以忽略柱的轴向位移，并且

采用图3(b)所示的12自由度集中质量柱对结构进行模拟。这一简化意味着每个楼层均可沿x轴和y轴发生平动位移，同时还可沿z轴发生转动变形。基于该结构模型，计算了质量矩阵和刚度矩阵（和）的各元素。阻尼矩阵被视为与质量矩阵和刚度矩阵成比例，表示为，其中 和。该阻尼矩阵产生了约为0.1%–5%的模态阻尼比。

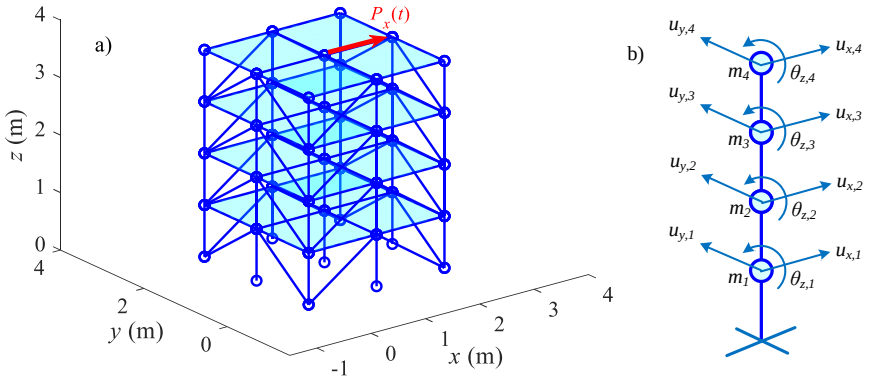

图3。（a）IASE-ASCE基准结构的三维视图；（b）用于推断的12自由度集中质量模型

在作出上述假设后，采用 采样间隔，对结构的合成响应进行生成并记录，记录时长足够长。将2nd 和3rd 楼层（ 和）的加速度响应作为测量量，并在其中加入零均值加性高斯白噪声，其标准差等于无噪声信号均方根（RMS）的5%。随后将这些模拟数据划分为 段，用于模型推断和响应预测。

假定用于模拟动力响应的基于物理的模型与原始结构具有相同的质量矩阵。刚度矩阵部分未知，其中的元素和 需要确定。这意味着应从数据中识别物理参数 。然而，阻尼矩阵被设定为与结构的阻尼矩阵存在偏差，具体采用 和。此外，假定物理参数在不同数据分区间变化，并从中抽取，其中。这种模型设定偏差旨在模拟描述观测动力响应时产生的系统误差。

MMTE核函数的精度和复杂度取决于所纳入余弦函数的数量（）。图4(a)显示了预测误差的功率谱密度，其中可以观察到频率为9、40和48Hz的三个主导模态。这初步表明可能是一个不错的选择。为了研究

进一步探讨这一问题时，在 将模态纳入核函数时，会计算贝叶斯信息准则（BIC）。如图4(b)所示，使用9、40和48Hz这三个模态时，获得的最大BIC得分最高。因此，从直观和理论角度来看，选择这三个模态是有依据的。

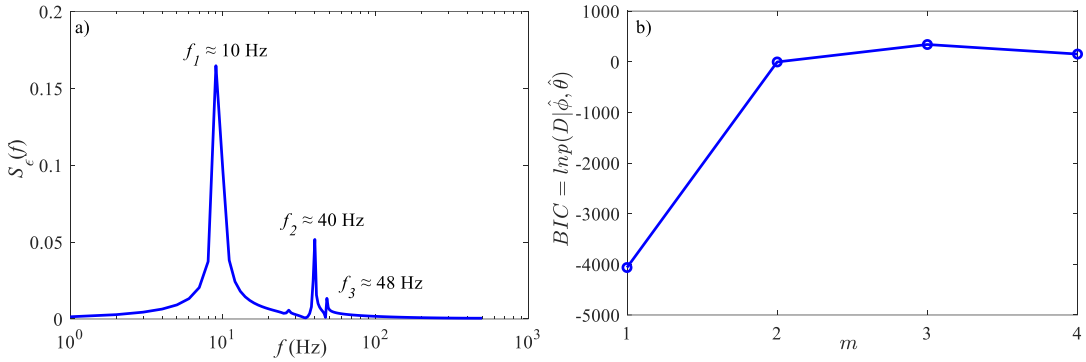

图 4. (a) PS 模型与实际加速度响应之间差异的 D，针对第 3rd 楼；(b) 与核模态数量相比的贝叶斯信息准则得分

模型和核模态。采用不同的数据点数量（）和分区数量（）时，核参数的估计均值如表1所示。对的估计值（其中）接近图4(a)中观测到的值，并且几乎不受数据点数量或分区数量的影响。相反，对、、和的估计值会随数据长度和分区数量而变化。这一观察是自然的，因为数据长度（或数据点数量）决定了核函数的时间特征。

n N_D σ_k^2 σ_n^2 $1/\ell^2$ n N_D

表1. 从后验预测分布中获得的未知参数估计均值

nN_D	n	N_D	σ_1^2	$1/\ell_1^2$	$\omega_1/(2\pi)$	σ_2^2	$1/\ell_2^2$	$\omega_2/(2\pi)$	σ_3^2	$1/\ell_3^2$	$\omega_3/(2\pi)$	σ_n^2
3000	500	6	0.88	3.95	9.84	0.99	2.57	39.39	0.87	2.98	47.79	3×10^{-6}
	1000	3	0.07	0.50	10.66	0.89	0.48	39.82	0.86	2.72	47.79	4×10^{-6}
9000	300	30	1.18	8.45	10.25	1.18	9.50	39.40	1.16	7.07	46.18	2×10^{-5}
	1000	9	0.45	1.24	10.83	0.99	1.03	39.80	1.04	1.43	47.90	7×10^{-6}

表2给出了在不同数据点数量和分区数量下识别得到的刚度参数的均值和标准差。估计均值与实际值（）非常接近，且估计标准差与实际值（）高度一致，验证了结果的有效性。此外，无论数据点数量和分区数量如何变化，刚度参数的均值和标准差都保持一致。这种渐近行为是理想的，因为它解决了经典贝叶斯方法的一个显著缺陷[1], ，即后验不确定性会随着数据点数量和分区数量的增加而降低。

表2. 刚度均值和标准差的估计值

基于后验预测分布获得的参数

nN_D	n	N_D	μ_{θ_1}	μ_{θ_2}	σ_{θ_1}	σ_{θ_2}
3000	500	6	0.998	1.020	0.0101	0.0099
	1000	3	1.002	0.997	0.0098	0.0082
9000	300	30	1.008	1.000	0.0103	0.0102
	1000	9	1.006	1.004	0.0092	0.0102

确定未知参数后，我们基于算法1传播不确定性，以预测动力响应。图5展示了4th 自由度（）的加速度响应。测量响应的前10秒片段被视为训练数据集，

接下来的 10 秒数据被视为留出数据集。为便于观察，提供了四个放大窗口。如图所示，估计平均响应与测量响应相符，阴影区域所示的不确定性界限合理地涵盖了潜在误差。在图6中7th 自由度 ( )也可以观察到类似的模式，这强调了( )所提出的方法的准确性。

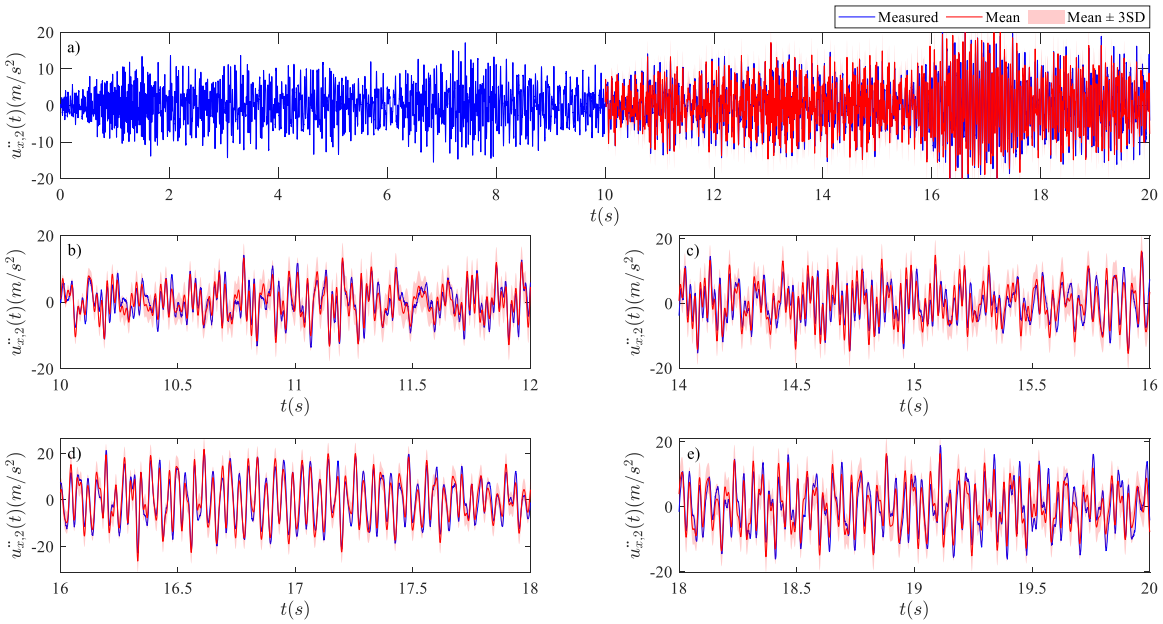

图5. 使用将 10 秒数据划分为多个2nd 段、每段包含 个时间采样点时，对 的楼层沿 x 轴的加速度响应进行预测

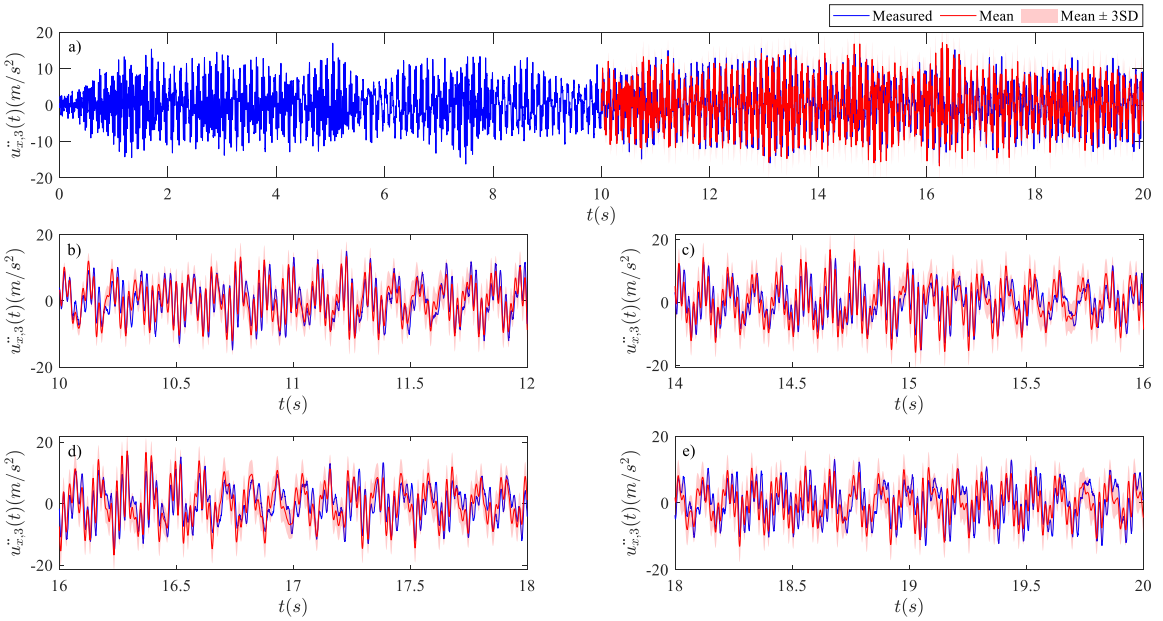

图6. 使用将 10 秒数据划分为多个分区、每个分区包含若干时间采样点时，对 3rd 楼层沿 x 轴的加速度响应进行预测

图7对比了 基于经典贝叶斯方法的预测（例如，[1]）与所提方法的预测。图7(a)表明，经典框架未考虑潜在的

估计均值与测量响应之间存在偏差，因为不确定性界限看起来比所需的更窄。相比之下，图7(b)显示，通过所提出的框架计算出的估计均值和不确定性界限（±3SD）能够充分反映模型设定错误导致的误差。

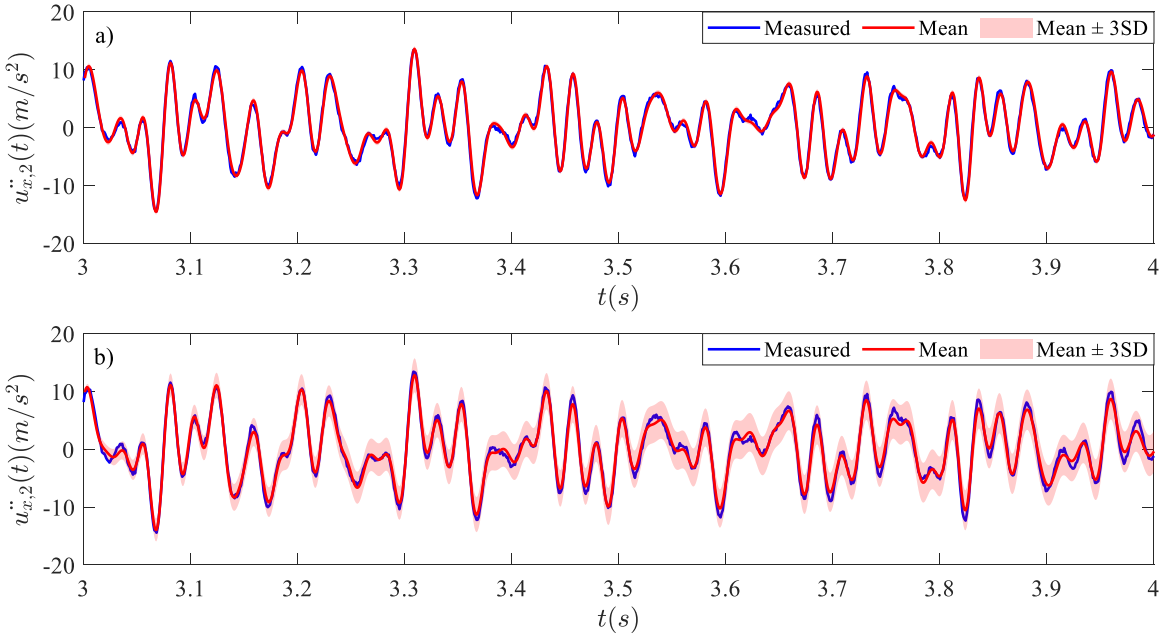

图7. 使用 (a) 经典贝叶斯框架和 (b) 所提出的框架预测沿 x轴方向的 2nd 楼层加速度时程响应

5.2. 实验示例 本节中，我们测试所提

出的框架，并使用实验数据进行验证。为此，我们进行了振动台试验，对图8(a)所示的小比例三层结构的基底施加激励。对该结构基底和三层楼沿 x轴方向的加速度响应进行测量，激励为沿 x轴方向的宽带高斯白噪声加速度。图8(b)所示的三层剪切框架可用于模拟该结构的动力学行为。施加的激励如图9(a)所示，3rd 楼层的实测加速度响应如图9(b)所示。三个楼层的加速度时程响应以 0.005s 的时间间隔记录。

图 9(b) 所示结构模型的质量被视为集中在各楼层上，这是由于楼板较厚所致。柱和梁的质量相对可以忽略。因此，质量矩阵为对角矩阵，其元素为 , , and . 阻尼矩阵 ( ) 被视为与质量矩阵成正比，表示为

$\hat{\mathbf{C}}$

$$\hat{\mathbf{C}} = \sum_{i=1}^{3} \left(2\hat{\varpi}_i \hat{\zeta}_i \frac{\hat{\mathbf{M}}\hat{\phi}_i \hat{\phi}_i^T \hat{\mathbf{M}}}{\hat{\phi}_i^T \hat{\mathbf{M}} \hat{\phi}_i} \right) \tag{45}$$

其中模态阻尼比由 [41,42], 考虑得到，， 以及； 为质量矩阵；，， 以及为标称模态频率；为对应于模态频率 i 的标称振型。与已知矩阵 和 不同，刚度矩阵被视为未知，表示为

$\hat{\varpi}_i$ $\hat{\mathbf{M}}$ $\hat{\mathbf{C}}$

$$\mathbf{K}(\boldsymbol{\theta}) = \begin{bmatrix} \theta_1 + \theta_2 & -\theta_2 & 0 \\ -\theta_2 & \theta_2 + \theta_3 & -\theta_3 \\ 0 & -\theta_3 & \theta_3 \end{bmatrix} \tag{46}$$

其中 是参数向量 包含每层楼的刚度。主要目标是利用楼层和楼层的加速度响应，采用所提方法识别未知参数；次要目标是预测未观测响应。

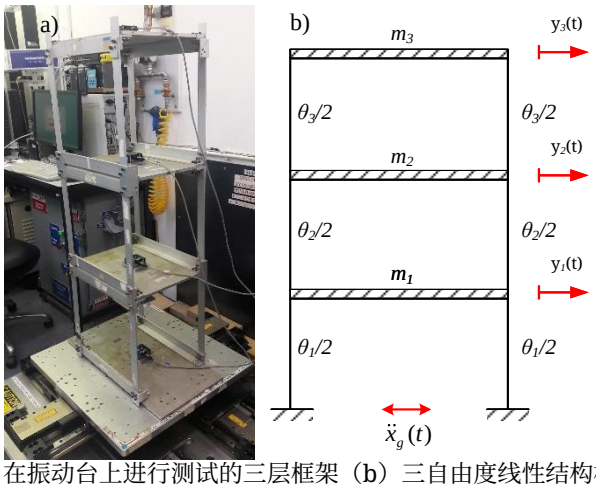

图 8。（a）在振动台上进行测试的三层框架（b）三自由度线性结构模型

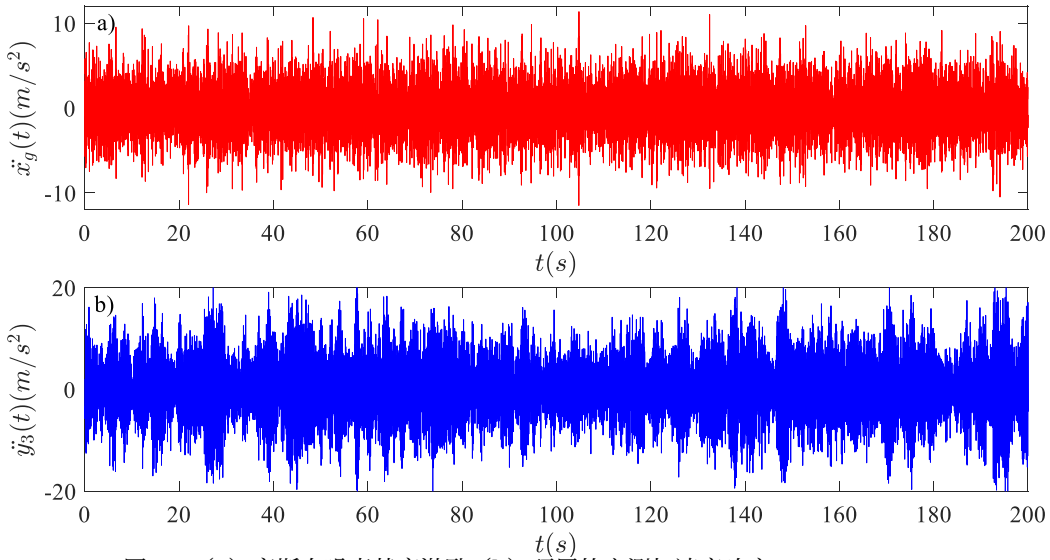

图 9。（a）高斯白噪声基底激励（b）顶层的实测加速度响应

基于算法1，过渡马尔可夫链蒙特卡洛采样方法[43–45] 用于抽取 样本，考虑到 和。图10展示了 基于马尔可夫样本的后验预测分布。红色虚线表示 [41], 中报告的参考值，这些参考值似乎接近计算分布的众数。此外，第一楼层的估计刚度明显小于

第二层和第三层。这种现象可能是由于首层柱与底板之间的连接比其他楼层松动所致。

图 11 显示了核函数参数的后验预测分布，。 核频率可近似为， 和。 这些值与前文提到的结构模态频率一一对应，分别为， 和 。 相关长度位于区间内，表明 时间相关性随时间快速衰减。

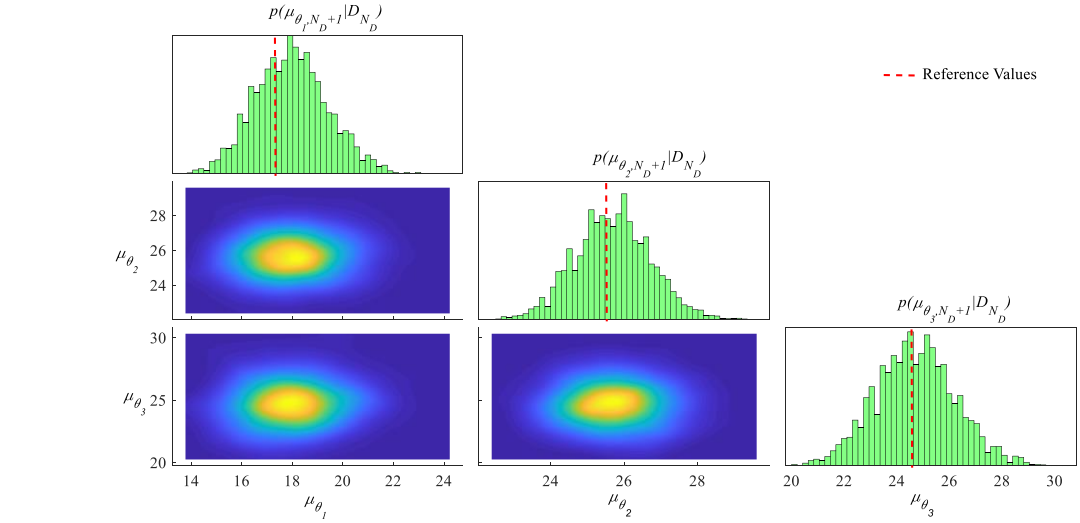

图 10。基于 20 组振动数据计算得出的后验预测分布

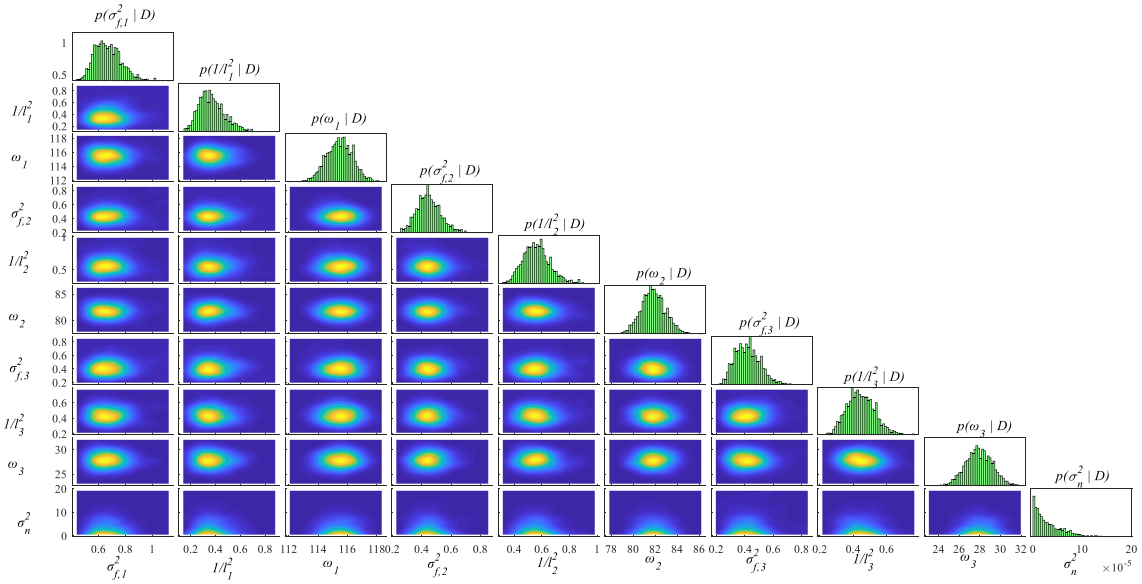

图 11. 基于 20 组振动数据计算得到的核参数后验预测分布

表3显示了在考虑不同数量的数据点（）和分区（）时识别出的结构参数的后验均值和标准差。可以看出，一楼的刚度识别结果小于另外两层，这可能源于基底螺栓松动。估计均值也似乎接近于先前研究中报告的数值[41]。此外，值得注意的是，无论和的取值如何，估计均值

并且标准差几乎收敛到相同的数值。这种渐近行为是理想的，因为后验分布不受数据点数量和测量信息冗余的影响。

结构参数统计属性的收敛趋势如图12所示，适用于。均值（）和标准差（）由各数据分区估计，并以误差条表示。后验预测分布 针对计算，用以反映随着不同分区信息的逐渐累积，参数如何变化。基于该分布，刚度参数的估计均值和标准差分别通过图12中的红线和阴影区域表示。如图所示，不确定性界限能够涵盖不同数据分区中估计均值和标准差的变化。

表3. 针对不同数据点数量或分区选择计算得到的刚度参数后验均值和标准差

nN_D	n	N_D	$\hat{\mu}_{\theta_1}$	$\hat{\sigma}_{\theta_1}$	$\hat{\mu}_{\theta_2}$	$\hat{\sigma}_{\theta_2}$	$\hat{\mu}_{\theta_3}$	$\hat{\sigma}_{\theta_3}$
10000	1000	10	17.98	0.0867	25.58	0.0804	24.97	0.0938
	500	20	17.95	0.0945	25.55	0.1011	24.96	0.1060
25000	1000	25	17.99	0.0975	25.54	0.1041	24.99	0.1105
	500	50	17.96	0.1069	25.54	0.0896	24.96	0.1075

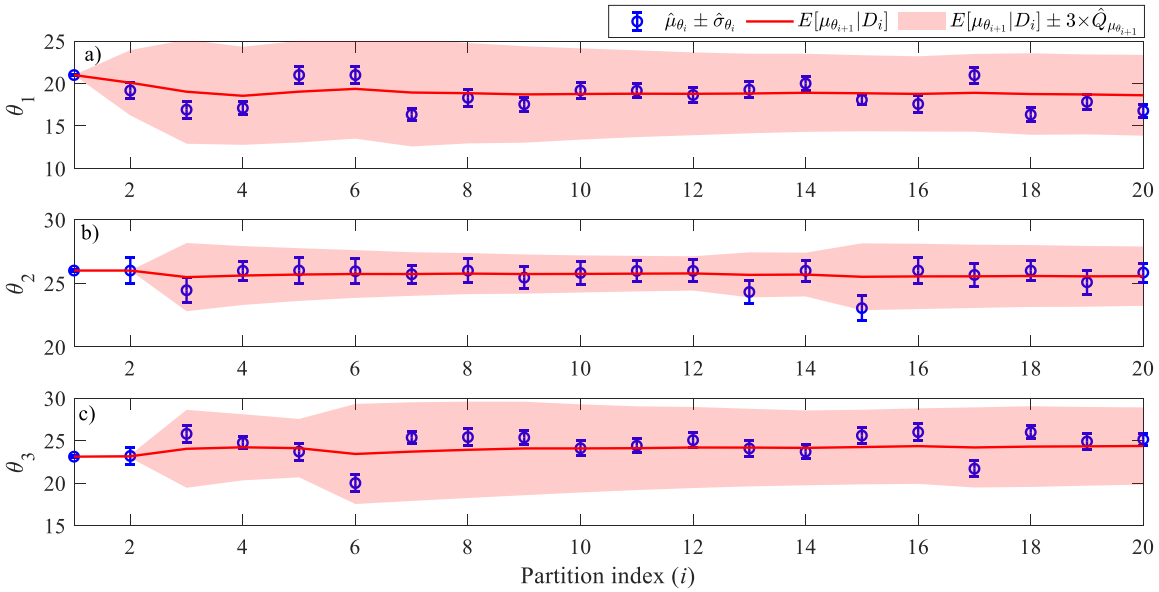

图12. 刚度参数的后验分布随分区数量的变化

图13绘出了1st 层和3rd 层的加速度响应预测结果。如图所示，估计平均响应与测量信号高度吻合。此外，表示99%置信区间（±3SD）的阴影区域较好地涵盖了测量响应与均值响应之间可能存在的偏差。

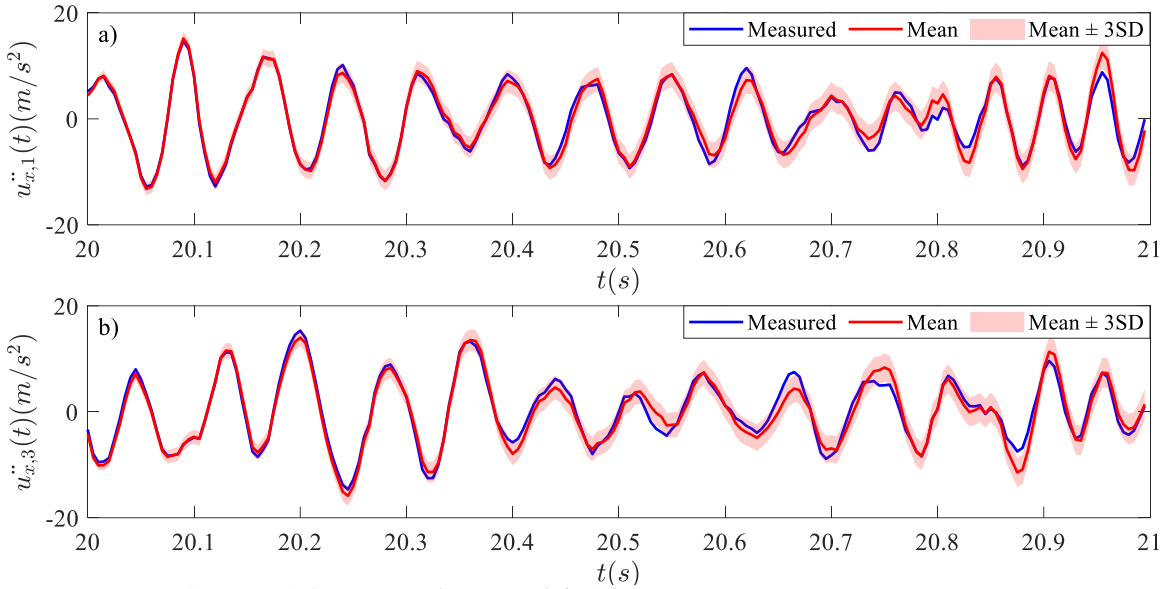

图13. 基于以下内容对1st 和3rd 楼层的加速度响应进行预测

6. 结论与未来工作 新型核协方差函数 a

本文在结构动力学反问题背景下引入了相关方法，以推动物理信息高斯过程模型的发展。这些协方差函数由MMTE核函数、物理驱动切线协方差函数和各向同性高斯噪声构成。对这一复合核函数进行理论研究后，我们发现，它将物理响应的梯度纳入协方差结构，从而在面对参数变化时具有更强的稳健性。此外，本文采用分层贝叶斯建模技术，通过将测量时程数据划分为多个子段来考虑时间变异性。尽管采用层次概率模型的主要动机是考虑测量响应的非平稳性，但事实证明，所提出的数据分区方法在计算方面也具有优势，能够缓解高斯过程的可扩展性问题。随后，本文提出一种计算算法，以便对测量数据进行序贯处理，并估计结构参数和未测量的动力响应。最后，本文给出两个示例，以验证所提出的方法，并据此得出以下具体结论：

- 模型差异的谱密度揭示了少数几个占主导地位的峰值。MMTE核函数适用于表征这种随机行为。• 通过贝叶斯模型选择方法确定MMTE核函数的阶数，结果表明，MMTE核的良好模型阶数可以是预测误差功率谱密度中出现的主导峰值数量。• 所提出的框架能够可靠地反映不同数据集之间结构模型参数的变异性，并为未测量的动力响应提供合理的后验分布。与其他非层次方法不同，计算得到的不确定性不会随着数据点和分区数量的增加而减小。• 预测响应与测量结果吻合良好，传播不确定性能够涵盖基于力学的模型可能存在的差异，并考虑潜在的建模误差。

尽管文中提供了简单示例来展示该方法，但本研究表明，将物理知识融入核协方差函数可以实现更高的准确率。不过，仍需开展更多研究，以进一步探索所提出方法的适用性

将该框架应用于其他类型的问题。从更广泛的角度来看，可以进一步探索所提出的框架，将其应用于非线性系统和非平稳类型的力激励。

附录（A）。式（8）中显式积分的推导

定理1。设 和 是由高斯分布 和 描述的两个随机变量向量，其中， ， 且均为已知量。将这两个高斯分布相乘后，对 在 上的积分为（其详细证明见[46]）

$$\int_{\boldsymbol{\mu}} N(\mathbf{X}\,|\,\mathbf{A}\boldsymbol{\mu},\boldsymbol{\Sigma})N(\boldsymbol{\mu}\,|\,\boldsymbol{\mu}_0,\boldsymbol{\Sigma}_0)d\boldsymbol{\mu} = N(\mathbf{X}\,|\,\mathbf{A}\boldsymbol{\mu}_0,\boldsymbol{\Sigma}+\mathbf{A}\boldsymbol{\Sigma}_0\mathbf{A}^T) \tag{A1}$$

推论1。根据定理1，式（8）中的积分可计算为

$$\begin{aligned}\int_{\boldsymbol{\theta}} N(\mathbf{y}\,|\,f(\mathbf{x},\boldsymbol{\theta}),\mathbf{K})N(\boldsymbol{\theta}\,|\,\boldsymbol{\mu}_0,\boldsymbol{\Sigma}_0)d\boldsymbol{\theta} &\approx \int_{\boldsymbol{\theta}} N(\mathbf{y}\,|\,f(\mathbf{x},\boldsymbol{\mu}_0)+\mathbf{J}(\boldsymbol{\theta}-\boldsymbol{\mu}_0),\mathbf{K})N(\boldsymbol{\theta}\,|\,\boldsymbol{\mu}_0,\boldsymbol{\Sigma}_0)d\boldsymbol{\theta} \\ &= N(\mathbf{y}\,|\,f(\mathbf{x},\boldsymbol{\mu}_0),\mathbf{J}\boldsymbol{\Sigma}_0\mathbf{J}^T+\mathbf{K})\end{aligned} \tag{A2}$$

此结 ding proves the result claimed earlier in Eq. (8).

附录（B）。式（13）的推导

根据定义，式（12）中给出的核协方差函数的功率谱密度可计算为

$$\begin{aligned}S_\varepsilon(\omega;\phi) &\triangleq \int_{-\infty}^{\infty}\exp(-j\omega\tau)k_\varepsilon(\tau;\phi)d\tau \\ &= \sum_{k=1}^{m}\left[\int_{-\infty}^{\infty}\sigma_k^2\exp(-j\omega\tau)\exp(-\tau^2/\ell_k^2)\cos(\omega_k\tau)d\tau\right]+\sigma_n^2\int_{-\infty}^{\infty}\exp(-j\omega\tau)\delta(\tau)d\tau\end{aligned} \tag{B1}$$

注意，积分 等于一，这是由于狄拉克δ函数的基本性质所致。此外，欧拉’公式允许将余弦函数写成。将此表达式代入式（B1）可得:

$$\begin{aligned}S_\varepsilon(\omega;\phi) &= \frac{1}{2}\sum_{k=1}^{m}\sigma_k^2\left[\int_{-\infty}^{\infty}\exp\left(-\frac{\tau^2}{\ell_k^2}-(j\omega+j\omega_k)\tau\right)+\exp\left(-\frac{\tau^2}{\ell_k^2}-(j\omega-j\omega_k)\tau\right)d\tau\right]+\sigma_n^2 \\ &= \frac{1}{2}\sum_{k=1}^{m}\sigma_k^2\exp\left(-\frac{\ell_k^2}{4}(\omega+\omega_k)^2\right)\left[\int_{-\infty}^{\infty}\exp\left(-\left(\frac{\tau}{\ell_k}+\frac{\ell_k}{2}(j\omega+j\omega_k)\right)^2\right)d\tau\right] \\ &+ \frac{1}{2}\sum_{k=1}^{m}\sigma_k^2\exp\left(-\frac{\ell_k^2}{4}(\omega-\omega_k)^2\right)\left[\int_{-\infty}^{\infty}\exp\left(-\left(\frac{\tau}{\ell_k}+\frac{\ell_k}{2}(j\omega-j\omega_k)\right)^2\right)d\tau\right]+\sigma_n^2\end{aligned} \tag{B2}$$

积分 积分 已知为，因此最新的积分可以简化为 o

$$S_\varepsilon(\omega;\phi) = \frac{1}{2}\sum_{k=1}^{m}\sqrt{\pi}\sigma_k^2\ell_k\left[\exp\left(-\frac{\ell_k^2}{4}(\omega+\omega_k)^2\right)+\exp\left(-\frac{\ell_k^2}{4}(\omega-\omega_k)^2\right)\right]+\sigma_n^2 \tag{B3}$$

This result proves Eq. (13), claimed earlier within the text.

附录（C）。公式（30-31）的推导定理 2。设 为一个 ra 向量

由 描述的随机变量，其均值和协方差 均为已知。对于分块向量 ，其概率分布可写为

$$p(\mathbf{X}) = N\left(\begin{bmatrix}\mathbf{x}_1\\ \mathbf{x}_2\end{bmatrix}\middle|\begin{bmatrix}\boldsymbol{\mu}_1\\ \boldsymbol{\mu}_2\end{bmatrix}, \underbrace{\begin{bmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \boldsymbol{\Sigma}_{12}^T & \boldsymbol{\Sigma}_{22}\end{bmatrix}}_{\boldsymbol{\Sigma}}\right) \tag{C1}$$

Then, the probability distribution of $\mathbf{x}_2$ 条件分布于 可由[36]:

$$p(\mathbf{x}_2 \mid \mathbf{x}_1) = \frac{p(\mathbf{x}_1, \mathbf{x}_2)}{p(\mathbf{x}_1)} = N(\mathbf{x}_2 | \boldsymbol{\mu}_{2|1}, \boldsymbol{\Sigma}_{2|1}) \tag{C2}$$

$$\boldsymbol{\mu}_{2|1} = \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{12}^T \boldsymbol{\Sigma}_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1) \tag{C3}$$

$$\boldsymbol{\Sigma}_{2|1} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{12}^T \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12} \tag{C4}$$

致谢 财政支持来自

香港研究资助局资助的项目编号 16212918 和 16211019，谨此致谢。

参考文献[1] J.L. B

eck, L.S. Katafygiotis, 更新模型及其不确定性。I：贝叶斯统计框架，工程力学杂志 124 (1998) 455–461。https://doi.org/10.1061/(ASCE)0733- 9399(1998)124:4(455).[2] L.S. Katafygiotis, J.L. Beck, 更新模型及其不确定性。II：模型可辨识性，工程力学杂志 124 (1998) 463–467。https://doi.org/10.1061/(ASCE)0733- 9399(1998)124:4(463).[3] M. Raissi, P. Perdikaris, G.E. Karniadakis, 物理信息神经网络：用于求解涉及非线性偏微分方程的正向与反问题的深度学习框架，J. Comput. Phys. 378 (2019) 686–707。https://doi.org/10.1016/j.jcp.2018.10.045.[4] M. Raissi, P. Perdikaris, G.E. Karniadakis, 用于时变和非线性偏微分方程的数值高斯过程，SIAM J. Sci. Comput. 40 (2018) A172–A198。https://doi.org/10.1137/17M1120762.[5] R.M. Neal, 神经网络的贝叶斯学习，Springer New York，纽约，NY，1996。https://doi.org/10.1007/978-1-4612-0745-0.[6] A. Jacot, F. Gabriel, C. Hongler, 神经正切核：神经网络中的收敛性与泛化能力，(2018)。http://arxiv.org/abs/1806.07572.[7] C.E. Rasmussen, C.K.I. Williams, 机器学习的高斯过程，麻省理工学院出版社，2005.[8] K.P. Murphy, 机器学习：概率视角，麻省理工学院出版社，2012。https://doi.org/10.1007/978-94-011-3532-0_2.[9] M.C. Kennedy, A. O’ Hagan, 计算机模型的贝叶斯校准，J. R. Stat. Soc. Ser. B Stat. Methodol. 63 (2001) 425–464。https://doi.org/10.1111/1467-9868.00294.[10] E. Simoen, C. Papadimitriou, G. Lombaert, 贝叶斯模型更新中的预测误差相关性，J. Sound Vib. 332 (2013) 4136–4152。https://doi.org/10.1016/j.jsv.2013.03.019.[11] I. Bilionis, N. Zabaras, B.A. Konomi, G. Lin, 多输出可分离高斯过程：迈向一种高效、完全贝叶斯的不确定性量化范式，J. Comput.

Phys. 241 (2013) 212–239。https://doi.org/10.1016/j.jcp.2013.01.011.[12] C.M. Bishop, 模式识别与机器学习，纽约，2006.[13] G. Jia, A.A. Taflanidis, 用克里金元建模近似高维波浪与风暴潮响应，以实现实时风暴／飓风风险评估，Comput. Methods Appl. Mech. Eng. 261–262 (2013) 24–38。https://doi.org/10.1016/j.cma.2013.03.012.[14] P. Angelikopoulos, C. Papadimitriou, P. Koumoutsakos, X-TMCMC：用于贝叶斯反演建模的自适应克里金法，Comput. Methods Appl. Mech. Eng. 289 (2015) 409–428.[15] H.A. Jensen, C. Esse, V. Araya, C. Papadimitriou, 在时域中实现用于贝叶斯有限元模型更新的自适应元模型, Reliab. Eng. Syst. Saf. 160 (2017) 174–190。https://doi.org/10.1016/j.ress.2016.12.005.[16] A.M. Kosikova, O. Sedehi, L.S. Katafygiotis, 使用高斯过程回归进行贝叶斯模型更新，载于：J. Li, P.D. Spanos, J.B. Chen, Y.B. Peng (编)，第13届国际结构安全与可靠性会议（ICOSSAR 2021），中国上海，2021.[17] K.V. Yuen, 结构动力学与土木工程的贝叶斯方法，2010。https://doi.org/10.1002/9780470824566.[18] C.R. Farrar, K. Worden, 结构健康监测：机器学习视角，约翰·威立父子出版公司，2012.[19] S.-K. Au, 运行模态分析——建模、贝叶斯推断、不确定性定律，施普林格，新加坡，2017。https://doi.org/10.1007/978-981-10-4118-1.[20] K.-V. Yuen, L.S. Katafygiotis, 使用环境数据进行模态更新的贝叶斯时域方法，概率工程力学 16 (2001) 219–231。https://doi.org/10.1016/S0266- 8920(01)00004-2.[21] K.-V. Yuen, L.S. Katafygiotis, 使用完整输入与不完整响应噪声测量进行贝叶斯模态更新，工程力学杂志 128 (2002) 340–350。https://doi.org/10.1061/(ASCE)0733-9399(2002)128:3(340).[22] C. Papadimitriou, G. Lombaert, 预测误差相关性对结构动力学中最优传感器布置的影响，机械系统与信号处理 28 (2012) 105–127。https://doi.org/10.1016/j.ymssp.2011.05.019.[23] L.D. Avendaño-Valencia, E.N. Chatzi, D. Tcherniak, 用于减轻风力涡轮机结构健康监测中运行变异性的高斯过程模型，机械系统与信号处理 142 (2020) 106686。https://doi.org/10.1016/j.ymssp.2020.106686.[24] Y.-C. Zhu, S.-K. Au, 基于贝叶斯数据驱动模型，从运行模态分析中识别不确定模态特性，机械系统与信号处理 136 (2020) 106511。https://doi.org/10.1016/j.ymssp.2019.106511.[25] R. Nayek, S. Chakraborty, S. Narasimhan, 用于线性结构系统输入—状态联合估计的高斯过程潜在力模型，机械系统与信号处理 128 (2019) 497–530。https://doi.org/10.1016/j.ymssp.2019.03.048.[26] T.J. Rogers, K. Worden, E.J. Cross, 关于高斯过程潜在力模型在输入—状态—参数联合估计中的应用：兼论贝叶斯运行识别，机械系统与信号处理 140 (2020) 106580。https://doi.org/10.1016/j.ymssp.2019.106580.[27] C. Jiang, M.A. Vega, M.K. Ramancha, M.D. Todd, J.P. Conte, M. Parno, Z. Hu, 不可观测分布式响应下多层级模型的贝叶斯校准及其在平面闸门中的应用，机械系统与信号处理 170 (2022) 108852。https://doi.org/10.1016/j.ymssp.2022.108852.[28] P. Gardner, L.A. Bull, N. Dervilis, K. Worden, 关于核化贝叶斯迁移学习在基于群体的结构健康监测中的应用，机械系统与信号处理 167 (2022) 108519。https://doi.org/10.1016/j.ymssp.2021.108519.[29] M.K. Ramancha, J.P. Conte, M.D. Parno, 在线性动力系统的贝叶斯校准中考虑模型形式不确定性，机械系统与信号处理 171 (2022) 108871。https://doi.org/10.1016/j.ymssp.2022.108871。

[30] M.R. Jones, T.J. Rogers, E.J. Cross，基于物理信息的声发射映射中的高斯过程约束，机械系统与信号处理 188 (2023) 109984。https://doi.org/10.1016/j.ymssp.2022.109984。[31] O. Sedehi, C. Papadimitriou, L.S. Katafygiotis，时域模型更新与稳健预测的概率层次贝叶斯框架，机械系统与信号处理 123 (2019) 648–673。https://doi.org/10.1016/j.ymssp.2018.09.041。[32] M. Song, I. Behmanesh, B. Moaveni, C. Papadimitriou，通过层次贝叶斯模型更新方法考虑建模误差和固有结构变异性：综述，Sensors（瑞士）20 (2020) 1–27。https://doi.org/10.3390/s20143874。[33] O. Sedehi, C. Papadimitriou, L.S. Katafygiotis，利用模态统计信息对有限元模型进行层次贝叶斯不确定性量化，机械系统与信号处理 179 (2022) 109296。https://doi.org/10.1016/j.ymssp.2022.109296。[34] I. Behmanesh, B. Moaveni, G. Lombaert, C. Papadimitriou，结构识别的层次贝叶斯模型更新，机械系统与信号处理 64–65 (2015) 360–376。https://doi.org/10.1016/j.ymssp.2015.03.026。[35] K.-V. Yuen, L.S. Katafygiotis，利用环境数据进行模态更新的贝叶斯时域方法，概率工程力学 16 (2001) 219–231。https://doi.org/10.1016/S0266- 8920(01)00004-2。[36] K.B. Petersen, M.S. Pedersen，矩阵手册，丹麦技术大学，2012。[37] M.I. Jordan，图模型中的学习，Springer Science & Business Media，1998。[38] J.B. Nagel, B. Sudret，贝叶斯反问题中多层次不确定性量化的统一框架，概率工程力学 43 (2016) 68–84。https://doi.org/10.1016/j.probengmech.2015.09.007。[39] J. Quiñonero-Candela, C.E. Rasmussen，稀疏近似高斯过程回归的统一观点，机器学习研究期刊 6 (2005) 1939–1959。[40] E.A. Johnson, H.F. Lam, L.S. Katafygiotis, J.L. Beck，采用模拟数据的IASC-ASCE结构健康监测基准问题第一阶段，工程力学杂志 130 (2004) 3–15。https://doi.org/10.1061/(ASCE)0733-9399(2004)130:1(3)。[41] O. Sedehi, C. Papadimitriou, D. Teymouri, L.S. Katafygiotis，利用仅输出测量对动力系统中的状态和输入进行序贯贝叶斯估计，机械系统与信号处理 131 (2019) 659–688。https://doi.org/10.1016/j.ymssp.2019.06.007。[42] D. Teymouri, O. Sedehi, L.S. Katafygiotis, C. Papadimitriou，动力系统中的输入-状态-参数-噪声识别与虚拟传感：贝叶斯期望最大化（BEM）视角，机械系统与信号处理 185 (2023) 109758。https://doi.org/10.1016/j.ymssp.2022.109758。[43] J. Ching, Y.-C. Chen，用于贝叶斯模型更新、模型类别选择和模型平均的过渡马尔可夫链蒙特卡罗方法，工程力学杂志 133 (2007) 816–832。https://doi.org/10.1061/(ASCE)0733-9399(2007)133:7(816)。[44] S. Wu, P. Angelikopoulos, C. Papadimitriou, P. Koumoutsakos，贝叶斯退火序贯重要性抽样：过渡马尔可夫链蒙特卡罗的一种无偏版本，ASCE-ASME风险与不确定性工程系统杂志 B辑：机械工程 4 (2017) 011008。https://doi.org/10.1115/1.4037450。[45] W. Betz, I. Papaioannou, D. Straub，过渡马尔可夫链蒙特卡罗：观察与改进，工程力学杂志 142 (2016) 04016016。https://doi.org/10.1061/(ASCE)EM.1943-7889.0001066。[46] O. Sedehi, L.S. Katafygiotis, C. Papadimitriou，层次贝叶斯运行模态分析：理论与计算，机械系统与信号处理 140 (2020)。https://doi.org/10.1016/j.ymssp.2020.106663。