

多输出模型确认中的混合矩指标*

赵录峰¹, 吕震宙¹, 张磊刚², 王新维³

(1. 西北工业大学 航空学院, 陕西 西安 710072; 2. 中国运载火箭技术研究院, 北京 100076;

3. 中国人民解放军 93363 部队, 辽宁 沈阳 110141)

摘要:在不确定性条件下,同时考虑到多维输出之间的相关关系和单输出的均值,构建由多输出数学期望列阵和协方差矩阵组成的多输出模型确认局部混合矩指标和全局混合矩指标。其中局部混合矩指标包括绝对指标(LA-3M)和相对指标(LR-3M),它们适合单点位置的多输出局部模型确认;全局混合矩指标也包括绝对指标(GA-3M)和相对指标(GR-3M),它们适合多点位置的多输出全局模型确认。通过数字算例和工程算例,所提指标可行有效,能够方便地度量计算模型和物理实验之间的差异程度。

关键词:模型确认;混合矩;多输出模型;相关性;不确定性

中图分类号:TB114.3 **文献标志码:**A **文章编号:**1001-2486(2015)06-061-08

Mixed moment validation metric for models with multivariate output

ZHAO Lufeng¹, LYU Zhenzhou¹, ZHANG Leigang², WANG Xinwei³

(1. School of Aeronautics, Northwestern Polytechnical University, Xi'an 710072, China;

2. China Academy of Launch Vehicle Technology, Beijing 100076, China; 3. The PLA Unit 93363, Shenyang 110141, China)

Abstract: Considering the relations among multi-outputs and the mean of single output, the mathematical expectation of single dimensional variable and covariance metric of multi-dimensional variables were introduced into the validation metrics for models. The new metrics of LA-3M and LR-3M were proposed for validating multi-responses at a single validation site, while the metrics of GA-3M and GR-3M were proposed to collect data of multiple responses observed at multiple validation sites. These metrics were examined through a numerical test case and an engineering example to illustrate their feasibility and effectiveness. Results show that the proposed metrics are efficient and they can easily measure the differential degree of multiple responses between calculation model and physical experiment.

Key words: model validation; mixed moment; models with multivariate output; correlation; uncertainty

随着计算机仿真技术的飞速发展和复杂产品工程实验费用的不断增加,很多复杂产品的物理实验逐渐被计算模型所替代。计算模型是对真实物理过程的一个抽象描述,但建模过程受各种不确定性因素的影响,模型预测与物理实验结果之间可能存在一定的差异。如何科学地定量度量它们之间的差异程度,是当前“模型确认”领域的重点研究内容,也是国内外学术界和工业界普遍关注和讨论的热点。模型确认被定义为:“从目标用途角度出发,客观地评估计算模型在多大程度上准确描述真实物理过程”^[1-3]。在模型确认过程中,首要的核心问题是构建一个“模型确认指标”,该指标要求能够科学有效地定量度量计算模型和实际物理过程之间的差异程度。模型确认指标被定义为“用来度量模型响应量与实验结果

之间差异程度的指标”^[4-5]。它是模型确认工作的基础和依据,在工程设计与分析领域,广泛用于多模型方案决策、模型预测能力评估和模型校准效果评价等工作之中。因此,开展模型确认指标研究,对推进模型确认工作的科学发展,提高计算机数字仿真技术的工程应用水平具有重要的理论价值和实践意义。

近几年国内外学术界对模型确认指标进行了大量研究,文献[6]对现有指标进行了系统分类和分析。根据模型的关键特性,模型确认指标可分为不同的类型。按不确定性类型,模型确认指标可分为确定型和不确定型指标^[7-8];按响应量的数目,可分为单输出和多输出指标;按输入可控变量的位置,可分为单点位置和多点位置指标。本文主要研究不确定性条件下的模型确认指标构

* 收稿日期:2015-01-09

基金项目:国家自然科学基金资助项目(51475370);高等学校博士学科点专项科研基金资助项目(20116102110003)

作者简介:赵录峰(1973—),男,陕西富平人,博士研究生,E-mail:zlf315611@126.com;

吕震宙(通信作者),女,教授,博士,博士生导师,E-mail:zhenzhou@nwpu.edu.cn

建问题。概括起来,现有的不确定性条件下的模型确认指标,主要分为四类:假设检验类、贝叶斯因子类、频率类和面积类。通过对比分析,这些指标都有各自的优缺点及适用范围。文献[6,9]分别运用假设检验和贝叶斯方法开展模型确认工作,这两种方法都是在一定置信水平下对模型预测与实验结果的一致性进行判断,但没有对它们之间差异程度进行定量度量,实质上是一种定性与定量相结合的方法,并不是严格意义上的模型确认过程。文献[1]提出的频率指标是对一定置信水平下模型预测与实验结果均值之间距离的度量,这一指标虽然是一种严格意义上的模型确认指标,且便于设计人员理解,但由于没有考虑响应量的方差和相关性,因此仅适合响应量分布中心发生偏离情况下的模型的确认。文献[1,8,10]运用模型输出响应量与实验结果经验分布函数之间的面积差异,提出了直接面积和 u-pooling 两种模型确认指标,适合于单维和多维独立响应量的模型确认,但由于该方法建立在求解响应量边缘分布面积差异的基础之上,没有考虑到多维响应量之间的相关性,因此不适合多维相关响应量的模型确认问题。为了解决这一问题,文献[11]在直接面积和 u-pooling 指标的基础上,应用多维概率积分转换方法,提出了概率积分转换(Probability Integral Transformation, PIT)和 t-pooling 模型确认指标,分别适合多维相关响应量在单点位置和多点位置的模型确认。虽然这一指标在多维相关响应量模型确认方面具有一定的优越性,但也存在一定的缺陷。一是该指标采用多维概率积分转换方法,压缩了大量信息,存在评估过度和评估不足的风险;二是该指标属于无量纲指标,物理意义不明确,不便于工程人员理解和使用;三是当实验数据不足时,无法正确地求解多输出响应量的联合分布函数,从而使得该指标较难应用。为此,本文针对多维相关响应量的模型确认问题,依据随机变量各阶矩的优良特性,提出了一种物理意义相对明确且便于使用的基于混合矩的多维相关模型确认指标。

1 混合矩模型确认指标构建

1.1 多输出模型确认特点分析

不确定性条件下,如果将多输出计算模型的输入变量或参数(如:载荷、几何尺寸或边界条件等)看作随机变量,那么它们的输出响应量(如:应力、应变、挠度或加速度等)也是多维随机变量;同理,物理实验数据也属于多维随机变

量。因此,依据概率论和数理统计相关理论,不确定性条件下多输出模型确认指标实质上就是对多维计算模型输出响应量的联合概率分布与实验结果所服从的联合概率分布之间的差异程度的度量^[9]。

然而,在物理实验和数学建模过程中,受随机因素和人的认知能力影响,多输出模型确认指标的构建比较复杂,面临多种情况的挑战^[1]。一是多输出响应量的物理意义可能不同,各自的量纲存在差异。如有的响应量用空间指标(应变、挠度)度量,有的用时空指标(速度、加速度)度量。二是多维响应量之间可能具有相关性。当多输出模型具有相同的输入变量,或输入变量之间存在很强的相关性时,这些输入变量之间的相关性就会传递到输出响应量,使得输出响应量之间存在很强的相关关系。三是多维输出响应量的输入变量可能是可控变量,也可能是随机变量。可控输入变量在一定时空范围变化^[12],当它在单点位置时,多维输出响应量服从某一联合概率分布;当它在多个位置变动时,多维输出响应量所服从的联合概率分布随之发生变化。

1.2 多输出模型确认指标构建的数学基础

随机变量的数字特征是由随机变量的分布确定的,能够描述随机变量某一方面的随机取值特征的常数。对多维随机变量来说,最重要的数字特征是数学期望(一阶中心矩)、方差(二阶中心矩)和协方差(二阶混合矩)。数学期望描述随机变量取值的平均大小,方差描述随机变量与自身的数学期望的偏离程度,协方差描述随机变量之间的相互关系。数学期望、方差和协方差虽然不像分布函数、分布律和概率密度函数那样完整地描述随机变量,但它能够描述随机变量的重要方面或人们最关心的重要特征。由于不确定性条件下的多输出模型确认指标的本质是对模型输出响应量与实验观察值所服从的两个联合分布函数之间差异程度的度量,然而在实际工程应用中,物理实验观察值的分布函数一般很难准确求得,因此,多维模型确认指标可以运用这两个联合分布函数的数字特征之间的差异程度来表示。

设物理实验的多输出响应量向量为 $Y^e(X)$, X 表示可控输入向量; $Y_i^e(X)$ 表示 $Y^e(X)$ 中第 i 维响应量, $i=1, \dots, d$; d 表示输出向量的维数。物理实验响应量的一阶中心矩、二阶混合中心矩和二阶混合中心矩形成的协方差矩阵分别表示为:

$$\mu_i^e = E(Y_i^e), \quad (i=1, \dots, d) \quad (1)$$

$$c_{ij}^e = Cov(Y_i^e, Y_j^e) = E\{[Y_i^e - E(Y_i^e)][Y_j^e - E(Y_j^e)]\} \quad (2)$$

$$(i, j = 1, \dots, d)$$

$$C^e = \begin{pmatrix} c_{11}^e & c_{12}^e & \dots & c_{1d}^e \\ c_{21}^e & c_{22}^e & \dots & c_{2d}^e \\ \vdots & \vdots & & \vdots \\ c_{d1}^e & c_{d2}^e & & c_{dd}^e \end{pmatrix} \quad (3)$$

设与物理实验对应的预测模型响应量向量为 $Y^m(X, \theta)$, $Y_i^m(X, \theta)$ 表示 $Y^m(X, \theta)$ 中第 i 维模型响应量, $i = 1, \dots, d$; d 为输出向量的维数; θ 为模型参数矢量。模型响应量的一阶中心矩、二阶混合中心矩和二阶混合中心矩形成的协方差矩阵分别表示为:

$$\mu_i^m = E(Y_i^m), \quad (i = 1, \dots, d) \quad (4)$$

$$c_{ij}^m = Cov(Y_i^m, Y_j^m) = E\{[Y_i^m - E(Y_i^m)][Y_j^m - E(Y_j^m)]\} \quad (5)$$

$$(i, j = 1, \dots, d)$$

$$C^m = \begin{pmatrix} c_{11}^m & c_{12}^m & \dots & c_{1d}^m \\ c_{21}^m & c_{22}^m & \dots & c_{2d}^m \\ \vdots & \vdots & & \vdots \\ c_{d1}^m & c_{d2}^m & & c_{dd}^m \end{pmatrix} \quad (6)$$

1.3 多输出模型确认指标的构建

依据多维随机变量各阶矩的性质,结合文献[1]和文献[10]提出的模型确认指标构建原则,分别建立了如式(7)和式(8)所示的局部混合矩模型确认绝对指标(Local Absolute Metric based on Mixed Moment for model validation, LA-3M)和全局混合矩模型确认绝对指标(Global Absolute Metric based on Mixed Moment for model validation, GA-3M)。当多输出模型的可控变量在单点位置时,可运用式(7)度量模型响应量与实验结果之间差异程度;当可控变量在多个位置变动时,则每一个单点位置对应的 $d_{LA}(Y^m, Y^e)$ 属于随机变量,因此,可运用 $d_{LA}(Y^m, Y^e)$ 的数学期望 $d_{GA}(Y^m, Y^e)$ 度量模型响应量与实验结果之间差异程度。

$$d_{LA}(Y^m, Y^e) = \sum_{i=1}^d |\mu_i^m - \mu_i^e| + \sum_{i=1}^d \sum_{j=i}^d |c_{ij}^m - c_{ij}^e|^{\frac{1}{2}} \quad (7)$$

$$d_{GA}(Y^m, Y^e) = E[d_{LA}(Y^m, Y^e)] \quad (8)$$

当 μ_i^e 和 c_{ij}^e 不为零时,采用式(9)和式(10)所示的无量纲局部混合矩模型确认相对指标(Local Relative Metric based on Mixed Moment for model validation, LR-3M)和全局混合矩模型确认相对指标(Global Relative Metric based on Mixed

Moment for model validation, GR-3M) 将更为合理。与绝对指标类似,式(9)适合可控变量在单点位置的多维输出模型确认,式(10)表示单点位置 $d_{LR}(Y^m, Y^e)$ 的数学期望,适合可控变量在多个位置的多维输出模型确认。

$$d_{LR}(Y^m, Y^e) = \sum_{i=1}^d \left| \frac{\mu_i^m - \mu_i^e}{\mu_i^e} \right| + \sum_{i=1}^d \sum_{j=i}^d \left| \frac{c_{ij}^m - c_{ij}^e}{c_{ij}^e} \right| \quad (9)$$

$$d_{GR}(Y^m, Y^e) = E(d_{LR}(Y^m, Y^e)) \quad (10)$$

上述确认指标不仅满足文献[1]和文献[10]所提的客观性、物理性、无界限、非负性、对称性、三角不等性和收敛性等基本要求,而且同时考虑了输出响应量的均值及二阶混合矩,合理地度量了计算模型与实验数据在均值、变异性及相关性等方面的差异程度,因此,该指标具有灵敏度高和便于工程应用等优良特性。

2 混合矩模型确认指标的求解

基于混合矩方法的多输出模型确认指标求解的核心,是求解模型和实验输出响应量的一阶中心矩和二阶混合中心矩形成的协方差矩阵。下面依据大数定理,运用 Monte Carlo 数字模拟法求解。

2.1 单点位置指标求解

基于混合矩方法的单点位置多输出模型确认指标的求解流程如图1所示,求解过程可概括为4个步骤。

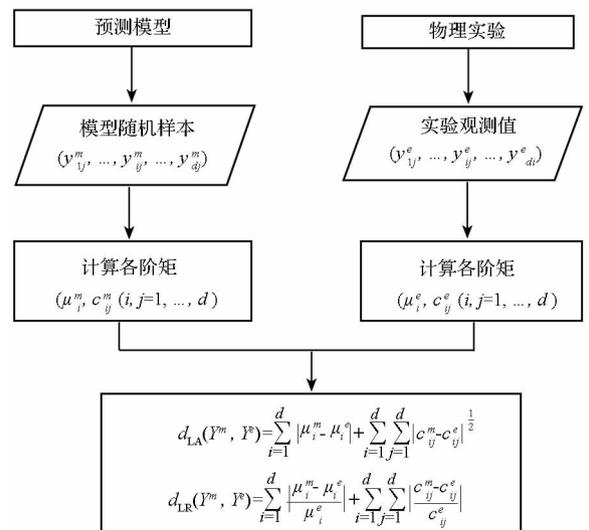


图1 单点位置多响应量模型确认指标计算流程

Fig. 1 Flowchart of validation metric for models with multivariate output at a single validation site

步骤1:收集物理实验观测值。流程图的右边,通过物理实验得到在单点位置矢量 x^* 处的 m

组 d 维实验观测值: $(y_{1j}^e, \dots, y_{ij}^e, \dots, y_{dj}^e), j = 1, \dots, m$ 。

步骤 2: 获取模型响应量随机样本。流程图的左边,应用相关变量随机数生成函数,得到在单点位置矢量 \mathbf{x}^* 处的 n 组 d 维模型响应量随机样本: $(y_{1j}^m, \dots, y_{ij}^m, \dots, y_{dj}^m), j = 1, \dots, n$ 。

步骤 3: 计算模型响应量和实验观测值的各阶矩。采用式(11) ~ (14) 计算模型响应量和实验观测值的各阶矩。

$$\mu_i^m = E(Y_i^m) \approx \frac{1}{n} \sum_{j=1}^n (y_{ij}^m) \quad (11)$$

$$\mu_i^e = E(Y_i^e) \approx \frac{1}{m} \sum_{j=1}^m (y_{ij}^e) \quad (12)$$

$$c_{ik}^m = Cov(Y_i^m, Y_k^m) \approx \frac{1}{n-1} \sum_{j=1}^n [y_{ij}^m - E(Y_i^m)][y_{kj}^m - E(Y_k^m)] \quad (13)$$

$$c_{ik}^e = Cov(Y_i^e, Y_k^e) \approx \frac{1}{m-1} \sum_{j=1}^m [y_{ij}^e - E(Y_i^e)][y_{kj}^e - E(Y_k^e)] \quad (14)$$

步骤 4: 求解模型确认指标。将求解所得的各阶矩带入式(7) 和式(9),可得到单点位置多维响应量模型确认指标 $d_{LA}(\mathbf{Y}^m, \mathbf{Y}^e)$ 和 $d_{LR}(\mathbf{Y}^m, \mathbf{Y}^e)$ 。

2.2 多点位置指标求解

多点位置指标求解建立在单点位置指标求解的基础之上,主要包括 3 个步骤。

步骤 1: 根据实验计划,确定 k 个不同位置的可控输入矢量 $\mathbf{x}^h, h = 1, \dots, k$ 。

步骤 2: 按照 2.1 节单点位置指标求解步骤,依次求解 k 个不同位置的可控输入矢量 \mathbf{x}^h 处的多维响应量模型确认指标 $d_{LA}^h(\mathbf{Y}^m, \mathbf{Y}^e)$ 和 $d_{LR}^h(\mathbf{Y}^m, \mathbf{Y}^e), h = 1, \dots, k$ 。

步骤 3: 根据式(8) 和式(10) 求解多点位置的模型确认指标。其中 $d_{GA}(\mathbf{Y}^m, \mathbf{Y}^e)$ 和 $d_{GR}(\mathbf{Y}^m, \mathbf{Y}^e)$ 分别由式(15) 和式(16) 求解。

$$d_{GA}(\mathbf{Y}^m, \mathbf{Y}^e) = E[d_{LA}(\mathbf{Y}^m, \mathbf{Y}^e)] = \frac{1}{k} \sum_{h=1}^k d_{LA}^h(\mathbf{Y}^m, \mathbf{Y}^e) \quad (15)$$

$$d_{GR}(\mathbf{Y}^m, \mathbf{Y}^e) = E[d_{LR}(\mathbf{Y}^m, \mathbf{Y}^e)] = \frac{1}{k} \sum_{h=1}^k d_{LR}^h(\mathbf{Y}^m, \mathbf{Y}^e) \quad (16)$$

3 算例分析

下面针对基于混合矩的多输出模型确认指标,通过两个算例验证它们的可行性和有效性。

3.1 算例 1

本算例的物理实验数据通过下式获取:

$$y_1^e(x, \theta) = \sin(2x - 0.5\pi)\theta + \varepsilon_1 \quad (17)$$
$$y_2^e(x, \theta) = \cos(0.25\pi x)\theta + 0.2x + \varepsilon_2$$

式中: $x(0 \leq x \leq 6)$ 为可控变量, $\theta(\theta = 1.5)$ 为模型参数, $\varepsilon_1 \sim N(0, 0.2^2)$ 和 $\varepsilon_2 \sim N(0, 0.2^2)$ 为两个实验响应量的测量误差, ε_1 与 ε_2 之间的相关系数 $\rho_{\varepsilon_1, \varepsilon_2} = 0.5$ 。构建的预测模型共 5 个,它们分为两组(见表 1)。设模型响应量的样本量用 n 表示,物理实验观测值的样本量用 m 表示。

表 1 两个测试组预测模型

Tab. 1 Formulas of predictive models in two test cases

测试模型组编号	公 式
第一组	1 $y_1^{m1}(x) = y_1^e(x, \theta = 1.5)$ $y_2^{m1}(x) = y_2^e(x, \theta = 1.5), \rho_{\varepsilon_1, \varepsilon_2} = 0.5$
	2 $y_1^{m2}(x) = y_1^e(x, \theta = 1.2)$ $y_2^{m2}(x) = y_2^e(x, \theta = 1.2), \rho_{\varepsilon_1, \varepsilon_2} = 0.5$
	3 $y_1^{m3}(x) = y_1^e(x, \theta = 1.2)$ $y_2^{m3}(x) = y_2^e(x, \theta = 1.2), \rho_{\varepsilon_1, \varepsilon_2} = -0.6$
第二组	4 $y_1^{m4}(x) = y_1^e[x, \theta \sim N(1.5, 0.2^2)]$ $y_2^{m4}(x) = y_2^e[x, \theta \sim N(1.5, 0.2^2)], \rho_{\varepsilon_1, \varepsilon_2} = 0.5$
	5 $y_1^{m5}(x) = y_1^e[x, \theta \sim N(1.5, 0.4^2)]$ $y_2^{m5}(x) = y_2^e[x, \theta \sim N(1.5, 0.4^2)], \rho_{\varepsilon_1, \varepsilon_2} = 0.5$

3.1.1 测试 1

包括 3 个预测模型。模型 1 与实验过程一致,是一个正确的模型;模型 2 的参数 θ 与实验模型有差异;模型 3 与实验模型相比,不仅模型参数 θ 发生了变化,而且相关系数也发生了变化。显然模型 1 的准确性高于模型 2,模型 2 高于模型 3。这组测试的目的:在模型参数或相关性存在差异的情况下,讨论所建指标的适应性、可行性和有效性。

3.1.1.1 单点位置测试

m 固定, n 变化。在 $x = 2.0$ 处,按照 2.1 部分提出的单点位置指标求解方法和步骤,分别由式(17)生成 $m = 1000$ 组实验观测数据,由表 1 第一测试组模型生成 n 组模型响应量,其中 n 在区间 $[100, 12\ 000]$ 变化,计算模型确认指标值。不确定性条件下,指标值具有一定的随机性,指标均值的置信度为 95% 的置信区间随 n 的变化情况如图 2、图 3 所示。

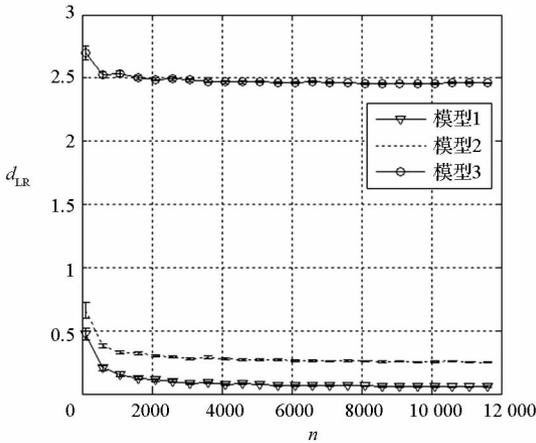


图2 d_{LR} 均值的置信区间在 $x = 2.0$ 处随 n 的变化曲线
 Fig. 2 Confidence interval of the mean of d_{LR} versus the n at validation site $x = 2.0$

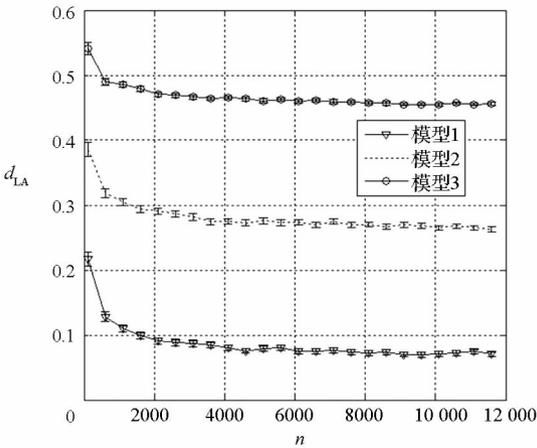
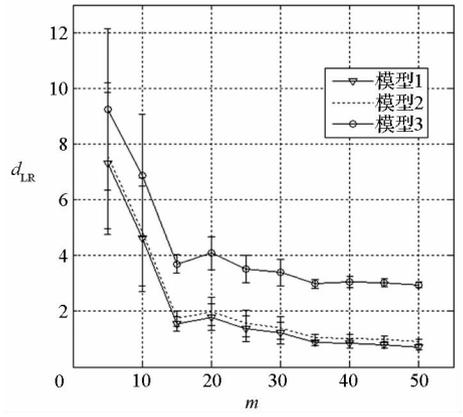


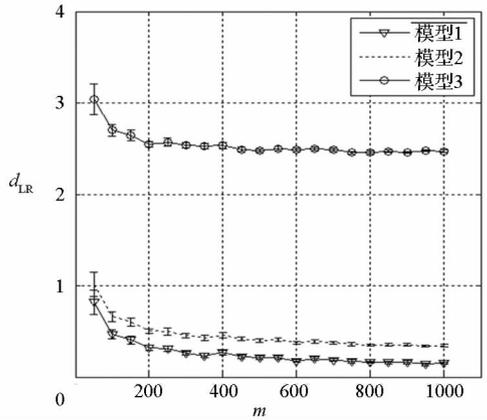
图3 d_{LA} 均值的置信区间在 $x = 2.0$ 处随 n 的变化曲线
 Fig. 3 Confidence interval of the mean of d_{LA} versus the n at validation site $x = 2.0$

由如图2、图3可以看出,在单点位置,随着 n 的增大,模型确认指标 $d_{LA}(Y^m, Y^e)$ 和 $d_{LR}(Y^m, Y^e)$ 的均值及其置信区间都迅速减小,当 $n > 2000$ 时,它们趋于稳定状态,收敛到一个固定值或范围,这表明 $n > 2000$ 时,模型响应量的计算结果趋于它们的解析解,且3个计算模型与物理实验之间的差异程度也处于稳定状态。为了尽可能地提高模型响应量的计算精度,本文在模型确认指标的验证过程中,选取 $n = 10\ 000$ 作为模型响应量样本量。

n 固定, m 变化。在 $x = 2.0$ 处,按2.1部分提出的单点位置指标求解方法和步骤,由表1第一测试组模型生成 $n = 10\ 000$ 组模型响应量,由式(17)生成 m 组实验观测数据,其中 m 在区间 $[5, 1000]$ 变化,计算模型确认指标值。指标均值的置信度为95%的置信区间随 m 的变化情况如图4和图5所示。



(a) m 在区间 $[0, 50]$ 变化图
 (a) m varies on the interval $[0, 50]$



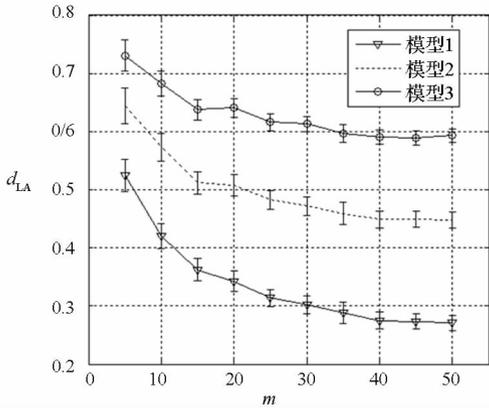
(b) m 在区间 $[50, 1000]$ 变化
 (b) m varies on the interval $[50, 1000]$

图4 d_{LR} 均值的置信区间在 $x = 2.0$ 处随 m 的变化曲线
 Fig. 4 Confidence interval of the mean of d_{LR} versus the m at validation site $x = 2.0$

由图4、图5可以看出:在单点位置,随着实验观测数据 m 的增大,模型确认指标 $d_{LA}(Y^m, Y^e)$ 和 $d_{LR}(Y^m, Y^e)$ 的均值及其置信区间都迅速减小;当 $m > 200$ 时,它们开始趋于收敛;当 $m = 1000$ 时,它们收敛到一个固定值或范围。

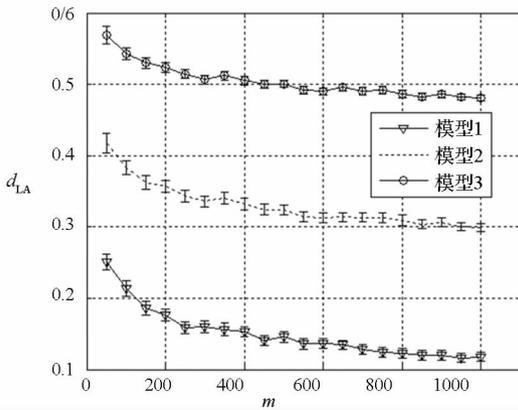
由图4(a)和图5(a)可以看出,当 m 在 $[5, 50]$ 区间取值时,模型确认指标(尤其相对指标 d_{LR})均值的置信区间较大,且相互重叠。这一现象表明当实验观测数据量过少时,运用该指标难以准确区分计算模型之间的优劣,存在判断错误的风险。因此,在实验数据过少,且实验数据误差过大的情况下,提出的模型确认指标具有一定的局限性。

由图4(b)和图5(b)可以看出,当实验观测数据 $m > 50$ 时,局部混合矩模型确认指标都能够很好地度量计算模型与物理实验之间的差异程度,客观地评估计算模型之间的优劣。为了从理



(a) m 在区间 $[0, 50]$ 变化

(a) m varies on the interval $[0, 50]$



(b) m 在区间 $[50, 1000]$ 变化

(b) m varies on the interval $[50, 1000]$

图 5 d_{LA} 均值的置信区间在 $x = 2.0$ 处随 m 的变化曲线

Fig. 5 Confidence interval of the mean of d_{LA} versus the m at validation site $x = 2.0$

论上更好地验证所提指标的可行性和有效性,本文在模型确认指标的验证过程中,选取实验观测数据的样本量 $m = 1000$ 。

n 固定, m 固定。在 $x = 2.0$ 处,按 2.1 部分提出的单点位置指标求解方法和步骤,由式(17)生成 1000 组实验观测数据,由表 1 第一测试组计算模型生成 10 000 组模型响应量,计算模型确认指标,并与文献[11]提出的 PIT 面积法指标进行对比,结果见表 2。

表 2 单点位置第一测试组模型确认指标计算结果

Tab. 2 Metric results of the first test models for single validation sites

指标	模型 1	模型 2	模型 3
PIT 面积法	0.009	0.105	0.184
LA - 3M	0.036	0.238	0.411
LR - 3M	0.037	0.247	2.431

3.1.1.2 多点位置测试

可控制变量 x 在 $(0 \leq x \leq 6)$ 的区间范围内,从 0 开始每间隔 0.01 取一个 x 值,即选取 600 个固定点。同理,对每个 x 值,由式(17)生成 1000 组实验观测数据,由表 1 第一测试组中的 3 个模型分别仿真生成 10 000 组模型响应量,按照 2.2 部分提出的多点位置指标求解方法和步骤,得到的多点位置的模型确认指标值,并与文献[11]提出的 t-pooling 面积指标进行对比,结果见表 3。

表 3 多点位置第一测试组模型确认指标计算结果

Tab. 3 Metric results of the first test models for multiple validation sites

指标	模型 1	模型 2	模型 3
t-pooling	0.012	0.103	0.144
GA - 3M	0.073	0.440	0.594
GR - 3M	0.206	1.694	3.843

本算例测试 1 中,通过表 2 和表 3 可以清晰地看出,无论在固单点位置,还是在多点位置,在模型参数和相关系数存在差异的情况下,本文提出的模型确认指标同 PIT 和 t-pooling 面积指标具有同样的效果,都能够客观地判断出模型 1 优于模型 2,模型 2 优于模型 3,这一结论与定性分析相符,达到第一组测试的目的,验证了所提指标的可行性和有效性。

3.1.2 测试 2

包括 2 个预测模型。模型 4 和模型 5 将确定性参数 θ 误作为不确定性变量,服从均值同为 $\theta = 1.5$,方差不同的正态分布。其中模型 4 中 θ 的方差较小,而模型 5 中方差较大,显然模型 4 与实验数据的一致性高于模型 5。第二组测试的目的:在模型参数的离散程度存在很小差异时,验证所建指标是否能够客观准确地度量计算模型与物理实验之间的差异程度。

3.1.2.1 单点位置测试

在 $x = 3.0$ 处,按照 2.1 部分提出的单点位置指标求解方法和步骤,同理,由式(17)生成 1000 组实验观测数据,由表 1 第二测试组计算模型生成 10 000 组模型响应量,得到模型确认指标,并与文献[11]提出的 PIT 面积指标进行对比,结果见表 4。

3.1.2.2 多点位置测试

在可控制变量 $x(0 \leq x \leq 6)$ 的区间范围,从 0 开始每间隔 0.01 取一个 x 值,即选取 600 个固定点。对应于每个 x 值,由式(17)生成 1000 组实验观测数据,由表 1 第二测试组中的 2 个模型生成

10 000组模型响应量,按照 2.2 部分提出的多点位置指标求解方法和步骤,得到模型确认指标值,并与文献[11]提出的 t-pooling 面积指标进行对比,结果见表 5。

表 4 单点位置第二测试组模型确认指标计算结果

Tab.4 Metric results of the second test models for single validation sites

指标	模型 4	模型 5
PIT 面积法	0.048	0.117
LA-3M	0.508	1.013
LR-3M	2.674	10.675

表 5 多点位置第二测试组模型确认指标计算结果

Tab.5 Metric results of the second test models for multiple validation sites

指标	模型 4	模型 5
t-pooling	0.039	0.081
GA-3M	0.381	0.746
GR-3M	1.891	7.299

本算例测试 2 中,由表 4 和表 5 可以清晰地看出,无论在单点位置,还是在多点位置,在确定性参数 θ 被错误地作为不确定性变量的情况下,本文提出的指标同 PIT 和 t-pooling 面积指标具有同样的效果,都能够明确地判断出模型 4 优于模型 5,这一结论与定性分析吻合,达到第二组测试的目的,验证了本文所提指标的可行性和有效性。

对表 2~5 的计算结果进一步分析可以看出:基于混合矩的多输出模型确认相对指标(LR-3M 和 GR-3M)的灵敏度最高,而多输出模型确认绝对指标(LA-3M 和 GA-3M)与 PIT 和 t-pooling 指标的灵敏度处于同一水平,相对较低。按照模型确认指标构建原则,当模型 1 与实验模型完全一致时,其指标值应该为零。从表 2 和表 3 的结果来看,模型 1 的六个指标值都不为零,尤其 GR-3M 的偏差较大(0.206),这一问题的原因不在于指标本身,而在于实验室数据和模型计算存在不确定性。如果进一步提高模型计算精度,增加实验数据量,减少实验误差,模型 1 的指标值最终会收敛于零。

3.2 算例 2

如图 6 所示的矩形截面悬臂梁,自由端承受集中可控力,梁的长度 $L = 2\text{m}$,截面宽度 $b = 50\text{mm}$,截面厚度 $h = 37\text{mm}$,弹性模量 $E \sim N(206.8, 10.5^2)$,输出响应量为梁固定端 A 点的正

应力 σ_A (GPa)、自由端 B 点处截面的转角 θ_B (rad) 和挠度 y_B (mm),实验数据由解析方程式(18)产生,它们的测量误差分别为 $\varepsilon_\theta \sim N(0, 0.005^2)$, $\varepsilon_y \sim N(0, 0.005^2)$, $\varepsilon_\sigma \sim N(0, 0.01^2)$ 。

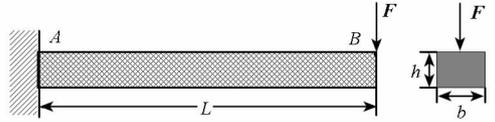


图 6 悬臂梁结构

Fig.6 Diagram of the cantilever beam structure

$$\begin{cases} \sigma_A = \frac{6FL}{bh^2} + \varepsilon_\sigma \\ y_B = \frac{4FL^3}{Ebh^3} + \varepsilon_y \\ \theta_B = \frac{6FL^2}{Ebh^3} + \varepsilon_\theta \end{cases} \quad (18)$$

为节约实验费用,提高效率,工业部门建立了 3 个不确定条件下的悬臂梁模型,这些模型的形式与解析方程相同,但各随机变量的分布函数或分布参数存在差异,具体情况见表 6。

表 6 悬臂梁输入随机变量的分布参数

Tab.6 Distribution parameters of inputs of the cantilever beam

变量	模型 1	模型 2	模型 3
E/GPa	$N(206.8, 16^2)$	$N(180.3, 16^2)$	$N(180.3, 16^2)$
$\varepsilon_\theta/\text{rad}$	$N(0, 0.005^2)$	$N(0, 0.005^2)$	$N(0, 0.005^2)$
ε_y/m	$N(0, 0.005^2)$	$N(0, 0.005^2)$	$N(0, 0.005^2)$
$\varepsilon_\sigma/\text{GPa}$	$N(0, 0.01^2)$	$N(0, 0.01^2)$	$N(0, 0.01^2)$
b/mm	50	50	$N(50, 4.5^2)$
h/mm	37	37	$N(37, 3.5^2)$

同理,本算例 F 分别在单点 1000N 处和 8 个位置 {600, 700, ..., 1300} 处,按照第 2 部分的指标求解方法和步骤,由式(18)生成 1000 组实验观测数据,由表 6 中的 3 个计算模型分别生成 10 000 组模型响应量,得到的模型确认指标计算值,结果见表 7。

表 7 悬臂梁模型确认指标计算结果

Tab.7 Results of validation metric for the cantilever beam models

指标	模型 1	模型 2	模型 3
LA-3M	0.011	0.035	0.095
LR-3M	2.989	7.536	16.671
GA-3M	0.014	0.034	0.091
GR-3M	3.735	5.227	14.013

由表 7 可以看出,无论在单点位置,还是在多点位置,运用本文提出的模型确认指标都能够明确判断出模型 1 优于模型 2,模型 2 优于模型 3,这一结论与定性分析结果吻合,从而再次验证了所提指标的可行性和有效性。

4 结论

基于混合矩的多输出模型确认绝对指标和相对指标及求解方法,能够有效度量模型响应量与实验结果之间差异程度。

1) 由于采用了工程人员比较熟悉且便于应用的各阶矩作为指标的组成要素,使得该指标具有工程应用方便、物理意义明显和分辨率高等优点。

2) 指标既适合单输出模型确认,也适合复杂的具有相关关系的多输出模型确认,可以直接用于工程设计与分析、多模型方案决策、模型预测能力评估和模型校准效果评价等工作之中,对推进工程建模的建设和发展具有重要意义。

3) 指标及其求解方法也存在一定的局限性:一是当实验数据过少且误差过大,或模型计算结果不准确时,指标计算结果会产生一定的偏差。二是指标求解方法是一种比较简单的方法,在提高求解效率和准确性方面,指标的求解方法还有待于进一步深化研究。

参考文献 (References)

[1] Oberkampf W L, Barone M F. Measures of agreement

- between computation and experiment: validation metrics[J]. *Journal of Computational Physics*, 2006, 217(1): 5-36.
- [2] Oberkampf W L, Sindir M, Conlisk A. Guide for the verification and validation of computational fluid dynamics simulations[J]. AIAA, 1998.
- [3] Oberkampf W L, Roy C J. *Verification and validation in scientific computing* [M]. New York, USA: Cambridge University Press, 2010.
- [4] Xiong Y, Chen W, Tsui K L, et al. A better understanding of model updating strategies in validating engineering models[J]. *Computer Methods in Applied Mechanics and Engineering*, 2009, 198(15-16): 1327-1337.
- [5] Messer M, Panchal J H, Krishnamurthy V, et al. Model selection under limited information using a value of information based indicator [J]. *Journal of Mechanical Design*, 2010, 132(12): 1-13.
- [6] Liu Y, Chen W, Arendt P, et al. Toward a better understanding of model validation metrics [J]. *Journal of Mechanical Design*, 2011, 133(7): 1-13.
- [7] Schwer L E. Validation metrics for response histories: perspectives and case studies [J]. *Engineering with Computers*, 2007, 23(4): 295-309.
- [8] Ferson S, Oberkampf W L. Validation of imprecise probability models[J]. *International Journal of Reliability and Safety*, 2009, 3: 3-22.
- [9] Rebba R, Mahadevan S. Validation of models with multivariate output [J]. *Reliability Engineering and System Safety*, 2006, 91(8): 861-871.
- [10] Ferson S, Oberkampf W L, Ginzburg L. Model validation and predictive capability for the thermal challenge problem [J]. *Computer Methods in Applied Mechanics and Engineering*, 2008, 197(29-32): 2408-2430.
- [11] Li W, Chen W, Jiang Z, et al. New validation metrics for models with multiple correlated responses [J]. *Reliability Engineering and System Safety*, 2014, 127: 1-11.
- [12] Jiang X M, Mahadevan S. Bayesian validation assessment of multivariate computational models [J]. *Journal of Applied Statistics*, 2008, 35(1): 49-65.