

基于 PSO-LSSVM 的复杂试验不确定度分析

洪谭亮¹, 丁晓红¹, 王海华², 王神龙¹, 徐世鹏¹

(1. 上海理工大学 机械工程学院, 上海 200093; 2. 延锋安道拓座椅有限公司, 上海 201315)

摘要: 提出一种结合粒子群优化 (PSO) 算法和最小二乘支持向量机 (LSSVM) 模型的复杂试验不确定度分析方法。以汽车座椅的安全带拉伸试验为对象, 研究汽车座椅安全带拉伸试验的主要影响因素及其概率密度函数参数, 采用拉丁超立方方法进行试验设计, 并进行安全带拉伸试验有限元仿真。运用 PSO-LSSVM 建立安全带拉伸试验的数学模型, 并与 BP (back propagation) 神经网络建立的数学模型进行对比, 结果显示 PSO-LSSVM 数学模型有更高的预测精度, 满足后续不确定度评定要求。进一步采用蒙特卡罗方法实现安全带拉伸试验的不确定度评定, 并以国家标准规定的方法进行参考, 研究结果表明, 该方法可应用于各种复杂试验不确定度分析中。

关键词: 试验不确定度分析; 粒子群优化; 最小二乘支持向量机; 蒙特卡罗法

中图分类号: TG 707; TB 9 **文献标志码:** A

Uncertainty analysis for complex test based on PSO-LSSVM

HONG Tanliang¹, DING Xiaohong¹, WANG Haihua², WANG Shenlong¹, XU Shipeng¹

(1. School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China;
2. Yanfeng Andaotuo Seat Co., Ltd., Shanghai 201315, China)

Abstract: A complex experimental uncertainty analysis method combining particle swarm optimization (PSO) and least square support vector machine (LSSVM) model was presented. Taking the tensile test of the automobile seat belt as the object, the main influencing factors and probability density function parameters of the tensile test of automobile seat belt were studied. The Latin hypercube method was used to design the test, and the finite element simulation of the tensile test of the seat belt was carried out. PSO-LSSVM was used to establish the mathematical model of the safety belt tension test, and compared with the mathematical model established by back propagation (BP) neural network. The results show that the PSO-LSSVM mathematical model has higher prediction accuracy and meets the requirements of the subsequent uncertainty evaluation. The Monte Carlo method was further used to evaluate the uncertainty of the safety belt tensile test, and the method was compared with the method specified in the national standard. The results show that the proposed method can be applied to the uncertainty analysis of complex experiments.

Keywords: test uncertainty analysis; particle swarm optimization; least square support vector machine; Monte Carlo method

测量不确定度用于表征合理赋予被测量的值的分散性^[1]。目前常用的不确定度评定方法是由《测量不确定度指南》(Guide to the Expression of Uncertainty in Measurement, GUM)^[2]给出的 A 类和 B 类评定。但是当测量系统比较复杂且样本量少的时候, GUM 法评定不确定度效果不是很好。因此, 2008 年 8 个国际权威组织联合发表不确定度补充文件^[3], 用蒙特卡罗法传递分布, 补充了 GUM 法。计量学界广泛运用蒙特卡罗方法, 崔孝海等^[4]采用蒙特卡罗方法对微波功率进行了不确定度分析; 孟令川等^[5]通过蒙特卡罗方法与 GUM 法对圆柱螺纹塞规中径的测量结果进行不确定度评定, 证明了蒙特卡罗方法有更优的评估结果; 朱大业等^[6]采用蒙特卡罗方法对汽车座椅的鞭打试验得分进行不确定度分析, 取得了很好的评价结果。陈怀艳等^[7]通过小功率校准因子说明了蒙特卡罗不确定度评定方法的可靠性。

蒙特卡罗模型评定不确定度时需要根据输入量的概率密度函数产生的样本量计算出输出量, 因此建立输入输出模型的精度至关重要。针对复杂试验输入与输出关系的复杂建模, 目前通常采用回归监督机器学习的算法^[8]: BP(back propagation)神经网络、支持向量机(support vector machine, SVM)、迭代决策树(gradient boosting decision tree, GBRT)等回归算法。最小二乘支持向量机、BP 神经网络等算法可用于小样本、高维特征变量模型。BP 神经网络没有给定目标函数的任何假设, 可以实现输入到输出的映射功能, 但是对数据的要求高, 对模型的训练也非常有挑战性^[9]; 而最小二乘支持向量机适合小数量样本, 避免过学习问题, 泛化能力强^[10]。

本文采用具有训练样本少、预测精度高的最小二乘支持向量机算法, 利用 PSO 算法优化 LSSVM 的惩罚系数及核函数参数, 解决最小二乘支持向量机在建模过程中惩罚系数及核函数参数难以确定的问题。以汽车座椅安全带拉伸试验为对象, 采用 PSO-LSSVM 对安全带拉伸试验建模, 考虑安全带拉伸试验主要影响因素的概率密度函数, 对其进行拉丁超立方试验设计, 最后通过蒙特卡罗方法实现对安全带拉伸试验的不确定度分析, 并以 GUM 法对安全带拉伸试验的不确定度分析作为参考。

1 安全带拉伸试验不确定度影响因素

1.1 安全带拉伸试验简介

安全带是汽车发生碰撞过程中保护乘驾人员安全的基本装置, 试验按照 GB14167—2013 程序设置, 座椅通过拉力 F_1 和 F_2 分别对上、下人体模块施加载荷, 如图 1 所示。对于 M1 和 N1 车型, 对上下人体模块沿车辆纵向中心平面与水平方向成向上 $10\pm 5^\circ$ 的方向施加试验载荷, 其中上人体模块施加 $13\,500\pm 200\text{ N}$ 的试验载荷, 下人体模块施加相对座椅质量 20 倍的载荷。试验过程中在 4 s 内加载到国标规定值并至少持续 0.2 s, 且安全带固定点不能失效, 并且不能超过对应的参考面。

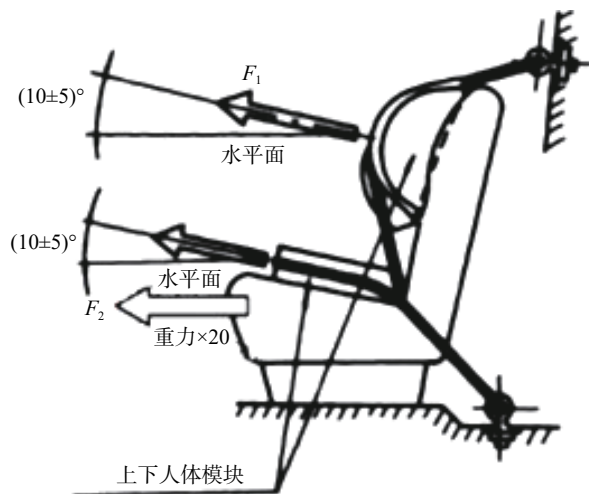


图 1 安全带拉伸试验设置

Fig.1 Safety belt tensile test setup

1.2 主要影响因素及其概率密度分布函数

安全带拉伸试验是最为复杂的汽车座椅静态试验, 试验设置参数有很多, 主要跟假臀和假胸的位置参数有关。根据试验标准和经验, 安全带拉伸试验的主要影响因素有: 假胸中心点与调角器中心点的竖直距离 A 、假胸倾斜的角度 J 、假胸与调角器的水平距离 B 、正视图中假胸中心线到安全带插口的水平距离 O , 如图 2 所示。

根据工程师经验和一般统计规律可以得出 A 为韦布尔分布, J , B 和 O 均为正态分布。

以某型号汽车座椅安全带拉伸试验为例, 以 B 和 O 的尺寸上下极限值作为其各自概率密度分布 95% 置信区间的上下限, 由概率统计区间估计可得出 B , J , O 的正态分布参数。正态分布函数记作 $N(\mu, \sigma^2)$, 其中: μ 决定了其位置; σ 决定了分

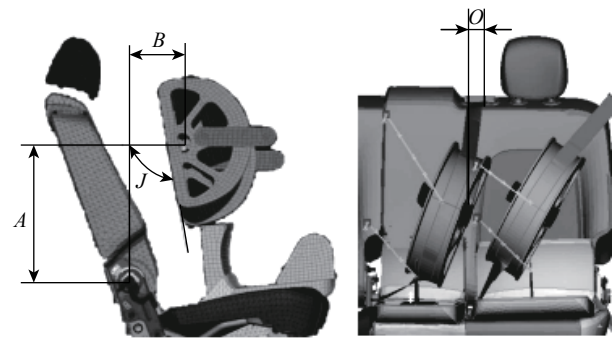


图 2 安全带拉伸试验主要影响因素

Fig. 2 Main influencing factors of safety belt tensile test

布的幅度。同时对 A 设置韦布尔分布参数。韦布尔分布函数记作 $W(\lambda, k)$, 其中: λ 为比例参数; k 为形状参数。 A, B, J, O 的分布函数参数如表 1 所示。

表 1 A, B, J, O 的分布函数参数

Tab.1 Distribution function parameters of A, B, J, O

输入因素	95%置信区间	概率分布类型	分布参数
A	[405, 435]	韦布尔分布	$\lambda=0.75, k=1$
B	[170, 250]	正态分布	$\mu=210, \sigma^2=416.49$
J	[85°, 105°]	正态分布	$\mu=95, \sigma^2=26.03$
O	[80, 100]	正态分布	$\mu=90, \sigma^2=26.03$

2 安全带拉伸试验设计及仿真

2.1 有限元模型

考虑到安全带拉伸试验的成本和一些不可控因素的影响, 汽车座椅安全带拉伸试验有限元模型的建立是本文进行不确定度分析的基础。依据 GB14167—2013 规定的试验方法, 结合某车型利用 ANSA 软件对汽车座椅进行有限元建模。安全带拉伸试验有限元模型如图 3 所示。

2.2 影响因素的试验设计

为了建立安全带拉伸试验输出和输入参数之间的数学模型, 需要在不同输入参数下进行多组安全带拉伸试验, 以便有足够的据来建模^[1]。当试验次数为固定值时, 拉丁超立方采样设计非常适用。拉丁超立方采样设计对每个因子的水平数没有限制, 适用于响应呈现非线性情况。抽样步骤如下:

步骤 1 选定设计变量, 并初始化各个变量的分布区间, 其各个输入参数概率密度函数如表 1 所示。

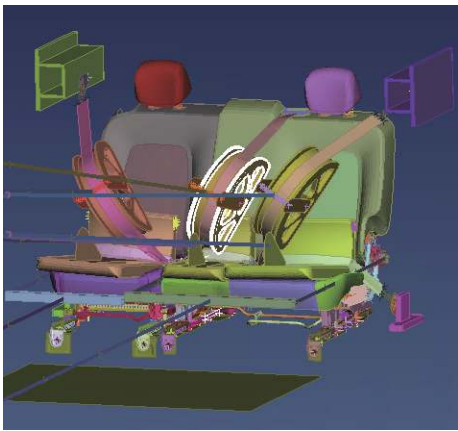


图 3 安全带拉伸试验有限元模型

Fig. 3 Finite element model of safety belt tensile test

步骤 2 确定抽样次数 l , 考虑到设计变量的个数和试验成本, 设定样本数为 30。因为采取了拉丁超立方的方法抽样, 30 组样本能准确地反映输入概率分布中值的分布, 具有足够的代表性, 得到的抽样样本具有足够的建模精度。增加样本数并不能对模型的精度有很高的提升, 同时会增加训练样本的时间。将每个变量 $x_D=(A, J, B, O)$ 按其分布函数将纵坐标轴分成 l 个等分, 每一等分宽度是 $1/l$ 。 x_L^D 为第 D 个变量的第 L 次抽样值, 如图 4 所示。

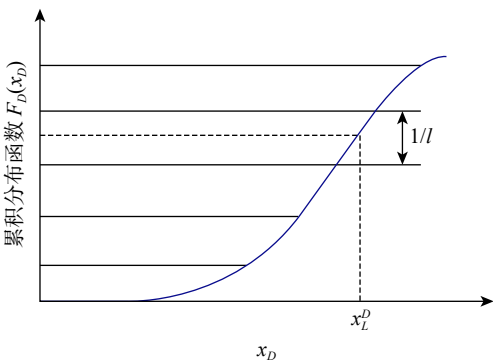


图 4 考虑概率密度分布的拉丁超立方抽样

Fig.4 Latin hypercube sampling considering probability density distribution

步骤 3 产生一个 30×4 的矩阵, 其每一行都是 $(1, l)$ 整数随机序列, a_L^D 是该矩阵的第 L 行 D 列元素。

步骤 4 产生一个 30×4 的矩阵, 其每个元素均服从 $[0,1]$ 均匀分布, p_L^D 是该矩阵的第 L 行 D 列元素。

步骤 5 通过计算得到 30×4 的采样矩阵, x_L^D 为其 L 行 D 列元素。则

$$x_L^D = F_D^{-1} \left(\frac{a_L^D - p_L^D}{l} \right) \tag{1}$$

式中： $L = 1, 2, \cdots, l; D = 1, 2, \cdots, d$ 。

步骤 6 显示拉丁超立方抽样的 30 组数据，如表 2 所示。结合汽车座椅安全带拉伸试验有限元模型，根据试验号调整有限元模型中影响因素 A, B, J, O 的尺寸，可以得到安全带固定点 X 方

表 2 安全带拉伸试验数据样本
Tab.2 Data samples of safety belt tensile test

试验号	A/mm	B/mm	J/mm	O/mm	BUCKLE_X/mm
1	414.80	247.19	96.20	99.34	128.67
2	412.33	193.34	94.22	92.76	131.63
3	415.72	201.35	93.65	82.63	130.17
4	407.52	226.01	90.14	89.57	124.58
5	414.57	230.2	85.81	84.15	120.80
6	408.40	178.66	94.14	94.34	133.80
7	412.32	199.77	97.49	100.33	135.21
8	408.16	190.46	92.72	82.20	130.43
9	405.64	205.12	103.33	92.55	139.13
10	411.59	220.20	91.77	88.37	126.55
11	406.02	216.57	110.35	87.89	145.75
12	420.22	238.13	92.83	86.23	128.08
13	406.36	218.05	95.44	96.11	129.10
14	429.93	212.6	101.2	91.2	137.33
15	412.65	227.19	99.51	88.74	132.69
16	443.2	214.11	92.29	91.59	130.42
17	407.13	184.06	97.1	93.25	136.26
18	422.19	210.87	91.26	86.69	128.54
19	422.99	182.05	88.95	86.88	129.66
20	422.06	233.4	96.67	92.11	129.13
21	408.69	161.32	87.35	96.78	130.02
22	420.43	198.42	84.42	95.64	125.19
23	417.1	209.91	94.59	89.30	131.12
24	407.64	204.67	100.13	84.66	135.40
25	411.91	249.46	98.95	79.60	131.26
26	405.95	207.17	89.68	85.58	125.51
27	433.50	197.22	95.13	93.83	134.30
28	413.28	223.64	98.18	87.73	132.84
29	410.58	188.20	98.29	90.65	136.30
30	410.71	222.49	101.61	90.42	134.93

向的位移 $BUCKLE_X$ ，如表 2 所示。

3 建模方法结果分析

选择采样得到前面 20 组样本作为训练集，后面 10 组样本作为预测集。利用粒子群优化算法来寻找最小二乘支持向量机核函数参数和正则化参数的最优值。设置粒子群的种群数为 30，最大迭代次数为 40，学习因子均为 1.5，正则化参数的取值范围为 $[0.1,1000]$ ，核函数参数的取值范围为 $[0.01,1000]$ 。运用 Matlab 软件运行 PSO-LSSVM 模型，将寻找到的最优核函数参数和正则化参数重新更新，最终获得核函数参数为 49.1524，正则化参数为 43.1890。用训练好的 PSO-LSSVM 优化模型对安全带拉伸试验进行预测。

为了验证和说明提出方法的有效性，采用 BP 神经网络模型进行对比。其中 BP 神经网络设置成输入层 4 个节点、隐含层 3 个节点、输出层 1 个节点的网络，隐含层和输出层都是 tansig 激活函数。预测结果如图 5 所示，PSO-LSSVM 的预测结果与真实值更为接近，参数优化有效。因此采用 PSO-LSSVM 建立安全带拉伸试验数学模型能有效地预测 $BUCKLE_X$ 。

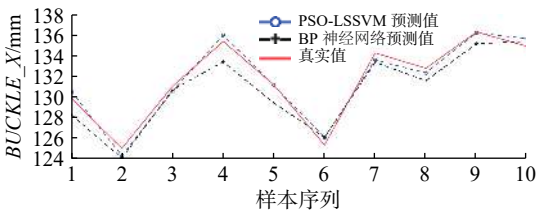


图 5 PSO-LSSVM 与 BP 神经网络建模对比
Fig.5 Comparison of PSO-LSSVM and BP neural network modeling

为了评估回归模型的预测能力，通常采用两个重要指标来验证模型精度，即平均相对误差 e_{mr} 和均方误差 e_{ms} ，其值越小，代表模型性能越好。其中，平均相对误差是指预测值 $\hat{y}_i$ 与真实值 y_i 的绝对误差在真实值中所占比例的平均值，其计算公式为

$$e_{mr} = \frac{1}{s} \sum_{i=1}^s \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \tag{2}$$

均方误差是指真实值与预测值差值的平方的期望值，其计算公式为

$$e_{ms} = \frac{1}{s} \sum_{i=1}^s (y_i - \hat{y}_i)^2 \tag{3}$$

运用 e_{mr} 和 e_{ms} 两个回归预测性能指标, 得到 BP 神经网络和 PSO-LSSVM 两种不同建模方法的 e_{mr} 和 e_{ms} , 其中 PSO-LSSVM 和 BP 神经网络预测结果的 e_{mr} 分别为 0.38% 和 0.63%, e_{ms} 分别为 0.31 和 1.07, 如表 3 所示。

表 3 PSO-LSSVM 与 BP 神经网络的 MRE/MSE 对比图
Tab.3 MRE / MSE comparison between PSO-LSSVM and BP neural network

建模方法	e_{mr}	e_{ms}
PSO-LSSVM	0.38%	0.31
BP神经网络	0.63%	1.07

通过表 3 对比可知, 不论是平均相对误差还是均方误差, PSO-LSSVM 建立的数学模型精度都要高于 BP 神经网络, 其精度满足后续不确定度评定要求。在模型计算效率方面, 采用 4 核 i5-7300、主频 2.5 GHz 的 CPU 计算机运行模型, BP 神经网络运行时间虽然较长, 但在可接受范围之内。

4 安全带拉伸试验不确定度评定

蒙特卡罗法 (Monte Carlo method, MCM)^[14-15], 与 GUM 法不同, 其采用分布传播方式来确定输出量的概率密度函数、估计值、不确定度以及包含概率区间。基于 PSO-LSSVM 安全带拉伸试验的蒙特卡罗不确定度评定方法步骤如下所示:

步骤 1 蒙特卡洛法输入

- a. 定义需要被测量的量 (输出量), 安全带拉伸试验需要测量的量是 $BUCKLE_X$;
- b. 确定与输出量 $BUCKLE_X$ 相关的输入量 $x_1, x_2, \dots, x_d$, 安全带拉伸试验的不确定度来源是 $x_1, x_2, \dots, x_d = (A, J, B, O)$;
- c. 建立输入量 $x_1, x_2, \dots, x_d$ 与输出量 $BUCKLE_X$ 的测量模型; 选 PSO-LSSVM 建立输入输出关系的数学模型;
- d. 为输入量 $x_1, x_2, \dots, x_d$ 设定其概率密度函数; 其中 A 符合韦伯尔分布, B, J, O 符合正态分布;
- e. 选择蒙特卡罗试验次数 M 的大小; 本试验选择试验次数 $M=10^6$, 因为 $M=10^6$ 会为输出量 $BUCKLE_X$ 提供 95% 的置信区间。

步骤 2 蒙特卡洛法传播

- a. 从输入量 $x_1, x_2, \dots, x_d$ 的概率密度分布函数中抽取 10^6 个样本值 $x_R^D (R=1, 2, \dots, MD=1, 2, \dots, d)$,

- x_L^D 为第 D 个变量第 R 个样本值;
- b. 计算输出量 $BUCKLE_X$ 的模型值 $y_R = f(x_R^1, x_R^2, \dots, x_R^d), (R=1, 2, \dots, M)$ 。

步骤 3 蒙特卡罗法输出与结果

- a. 将这 M 个输出量按严格递增的顺序排序, 通过排序模型得到输出量的分布函数如图 6 所示。

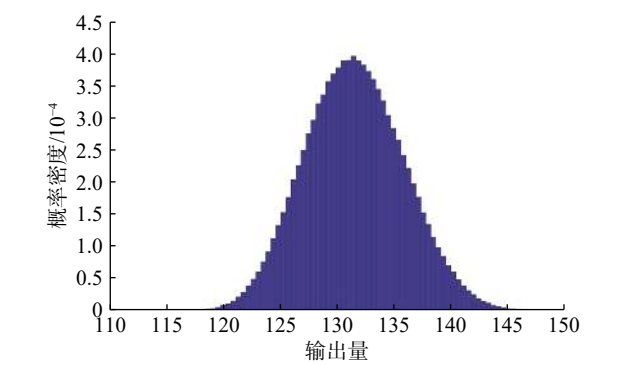


图 6 输出量的概率密度分布函数
Fig.6 Probability density distribution function of output

- b. 确定被测量 $BUCKLE_X$ 的估计值 y 、标准不确定度 $u(y)$, 以及在给定包含概率 95% 时被测量 $BUCKLE_X$ 的包含区间 $[y_l, y_h]$, 如表 4 所示。

表 4 基于 PSO-LSSVM 的安全带拉伸试验蒙特卡罗法不确定度分析结果
Tab.4 Uncertainty analysis results of safety belt tensile test MCM method based on PSO-LSSVM

估计值/mm	标准不确定度	95%概率包含区间/mm	
		左端点	右端点
131.5	4.34	123.28	140.22

目前因为 GUM 法使用普遍、计算简单等优点, 其评定结果也可以用来作为参考。由 GUM 法中的 A 类评定可知, 以 30 组试验数据得到的 $BUCKLE_X$ 结果均值当作估计值, 平均值的标准偏差当作 $BUCKLE_X$ 的标准不确定度, 评定的结果如表 5 所示。

表 5 基于 GUM 的 A 类安全带拉伸试验不确定度分析结果
Tab.5 Uncertainty analysis results of type a seat belt tensile test based on GUM

估计值/mm	标准不确定度	95%概率包含区间/mm	
		左端点	右端点
131.4	4.95	121.74	144.09

由表 5 中 A 类评定结果可知, 除了估计值与蒙特卡罗模型得到的结果较接近外, 标准不确定

度和包含区间误差都比蒙特卡罗评定的误差大,标准不确定度从 4.34 增加到 4.95,增加了 14%;95% 包含的概率区间长度由 16.94 增加到 22.35,增加了 32%。因为 GUM 法规定的 A 类评定在实际使用过程中只是对安全带拉伸试验 *BUCKLE_X* 位移进行了统计分析,没有体现出试验影响因素到试验结果的不确定度传递,且样本数目不多,因此 GUM 法评定安全带拉伸试验不确定度的精度较低,也进一步证实了 GUM 法不适用于复杂试验的不确定度分析。

5 结 论

研究了汽车座椅安全带拉伸试验主要影响因素和概率密度函数参数,采用拉丁超立方方法进行了试验设计。

通过 PSO 算法优化了 LSSVM 的惩罚系数和核函数参数,构建了汽车座椅安全带拉伸试验回归预测模型,并结合 BP 神经网络对预测结果进行了比较。结果表明:采用 PSO-LSSVM 方法能够建立有效的安全带拉伸试验预测模型,该模型的平均相对误差为 0.38%,均方误差为 0.31,具有较高的准确性。

结合 PSO-LSSVM 建模结果,采用蒙特卡罗方法对安全带拉伸试验进行了不确定度分析,并且参考了传统 GUM 法 A 类不确定度评定。结果表明:由于安全带拉伸试验的强非线性,基于 PSO-LSSVM 的蒙特卡罗不确定度分析适用于此复杂试验;GUM 法不适用于非线性复杂试验的不确定度分析。

参考文献:

- [1] 刘建坤,朱家平,郑荣华.测量不确定度评定研究现状及进展[J].现代科学仪器,2013(5): 12-17.
- [2] JJF 1059.1—2012,测量不确定度评定与表示[S].北京:中国标准出版社,2013.
- [3] 刘存成,胡畅.基于 MATLAB 用蒙特卡洛法评定测量不确定度[M].北京:中国质检出版社,2014.
- [4] 崔孝海,曲璐.蒙特卡罗方法在微波功率测量不确定度分析中的应用[J].计量学报,2008,29(1): 77-79.
- [5] 孟令川,朱泽熙.圆柱螺纹塞规中径不确定度评估的蒙特卡洛模拟[J].计量学报,2017,38(S1): 98-103.
- [6] 朱大业,丁晓红,王神龙,等.基于支持向量机模型的复杂非线性系统试验不确定度评定方法[J].机械工程学报,2018,54(8): 177-184.
- [7] 陈怀艳,曹芸,韩洁.基于蒙特卡罗法的测量不确定度评定[J].电子测量与仪器学报,2011,25(4): 301-308.
- [8] 何清,李宁,罗文娟,等.大数据下的机器学习算法综述[J].模式识别与人工智能,2014(4): 327-336.
- [9] WANG J S, HAN S, SHEN N N, et al. Features extraction of flotation froth images and BP neural network soft-sensor model of concentrate grade optimized by shuffled cuckoo searching algorithm[J]. The Scientific World Journal, 2014, 2014: 208094.
- [10] 李帅,王瑛,薛占龙,等.基于 LS-SVM 的接地电阻监测数据回归预测方法[J].国外电子测量技术,2019,38(8): 19-22.
- [11] 韦鹏飞,徐永海,王金浩,等.基于拉丁超立方采样的节点敏感设备暂降免疫水平评估[J].电工技术学报,2018,33(15): 3415-3425.
- [12] 石琴,仇多洋,吴冰,等.基于粒子群优化支持向量机算法的行驶工况识别及应用[J].中国机械工程,2018,29(15): 1875-1883.
- [13] 何邦乐,黄勇,叶颀,等.基于 PSO-LSSVM 的高压电力电缆接头温度预测[J].电力工程技术,2019,38(1): 31-35.
- [14] 刘存成,靳京民.用蒙特卡洛自适应法评定测量不确定度的通用程序[J].工业计量,2016,26(S1): 39-43.
- [15] 沈乐书.大量程液位计远程校准系统研究[D].杭州:中国计量大学,2016: 73-74.

(编辑:董 伟)