

基于ELM-PNN算法的第24周太阳黑子预测预报

刘志刚^{1†}, 李盼池¹, 许少华²

(1. 东北石油大学 计算机与信息技术学院, 黑龙江 大庆 163318;

2. 山东科技大学 信息科学与工程学院, 山东 青岛 266590)

摘 要: 为了提高太阳黑子预测预报的精度, 提出固定型极限学习过程神经网络(FELM-PNN)和增量型极限学习过程神经网络(IELM-PNN)两种学习算法. FELM-PNN的隐层节点数目固定, 使用SVD求解隐层输出矩阵的Moore-Penrose广义逆, 通过最小二乘法计算隐层输出权值; IELM-PNN逐次增加隐层节点, 根据隐层输出矩阵和网络误差计算增加节点的输出权值. 通过Henon时间序列预测验证了两种方法的有效性, 并实际应用于第24周太阳黑子平滑月均值的中长期预测预报中. 实验结果表明, 两种方法的预测精度均有一定程度的提高, IELM-PNN的训练收敛性优于FELM-PNN.

关键词: 过程神经网络; 极限学习; 网络训练; 广义逆; 太阳黑子数

中图分类号: TP183

文献标志码: A

Prediction and forecast of sunspot numbers in 24 cycle based on ELM-PNN algorithms

LIU Zhi-gang^{1†}, LI Pan-chi¹, XU Shao-hua²

(1. School of Computer and Information Technology, Northeast Petroleum University, Daqing 163318, China;

2. College of Information Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China)

Abstract: In order to improve prediction and forecast accuracy for the sunspot number, two process neural network(PNN) training algorithms of fixed extreme learning machine PNN(FELM-PNN) and incremental extreme learning machine PNN(IELM-PNN) are proposed. The FELM-PNN has fixed numbers of hidden layer nodes, and uses singular value decomposition(SVD) to compute Moore-Penrose generalized inverse of the hidden layer output matrix. The hidden layer output weights are solved by using the least squares method. For the IELM-PNN, the hidden layer nodes are added to the model one by one. The output weights for the added node are computed by according to the hidden layer output matrix and the network output error. The effectiveness of the two proposed methods is verified by Henon time series prediction. The two proposed methods are applied to the 24th cycle sunspot smoothed monthly mean Mid-and-long forecasting problem. The experimental results show that the prediction accuracy of the two methods increased at certain degree, and the training convergence of the IELM-PNN is better than that of the FELM-PNN.

Keywords: process neural network; extreme learning; network training; generalized inverse; sunspot numbers

0 引 言

太阳黑子是太阳活动水平的基本参数, 与地磁扰动、电离层电子浓度有密切关系. 太阳黑子的预测预报是空间预报的一项重要内容, 可为通信、导航、定位提供重要参考信息. 国内外学者对相关预测预报理论进行了广泛研究^[1-2]. 文献[3-7]使用人工神经网络(ANN)进行太阳黑子的预测预报, 取得了一定效果. 但受同步瞬时输入限制, ANN无法体现连续

信号的时间累积效应, 预测精度偏低. 过程神经网络(PNN)^[8]的输入和权值是与时间有关的函数, 通过时间聚合算子可以实现连续输入信号的时间累积, 减小复杂非线性系统逼近时的时间累积误差. 由于PNN计算复杂, 文献[9]提出利用正交基对输入和权函数进行有限项展开, 根据正交特性简化时域聚合运算, 有效降低了时间聚合算子的计算复杂性. 在此基础上, 文献[10]进行了PNN的太阳黑子预测问题研究,

收稿日期: 2016-01-21; 修回日期: 2016-05-25.

基金项目: 国家自然科学基金项目(61170132); 黑龙江省自然科学基金项目(F2015021).

作者简介: 刘志刚(1979—), 男, 副教授, 博士生, 从事神经网络、进化算法的研究; 李盼池(1969—), 男, 教授, 博士生导师, 从事量子优化算法等研究.

[†]通讯作者. E-mail: dqpilzg@163.com

预测精度比普通ANN有一定提升. 但PNN存在以下问题: 1) 隐层过程神经元节点数目不易确定, 目前的“试凑法”降低了模型的逼近精度; 2) 现有的梯度下降训练算法对初值敏感、易陷入局部最小. 极限学习机(ELM)是由Huang等^[11]提出的一种前馈神经网络快速学习算法, 已证明可极大提高学习速度^[12], 已成为国内外研究热点^[13-15].

综上所述, 为了提高太阳黑子预测预报精度, 针对PNN现有问题, 本文提出两种极限学习过程神经网络学习算法: 固定型FELM-PNN和增量型IELM-PNN. 1) FELM-PNN利用SVD法和Moore-Penrose广义逆求解隐层输出权值, 经矩阵计算即可完成PNN的训练, 学习速度很快, 但训练收敛性略有不足; 2) IELM-PNN利用增量方式, 逐次增加隐层过程神经元节点, 已有隐层节点的输入权函数系数固定, 根据网络输出误差和隐层输出矩阵计算增加节点的输出权值, 学习速度略慢于FELM-PNN, 但训练收敛性较好, 具有更紧凑的网络结构, PNN模型逼近能力也得到增强. 最后, 以Henon混沌时间序列为例, 对两种方法进行了验证, 并实际应用到的第24周太阳黑子平滑月均值的中长期预测预报中, 预测结果符合实际; 同时使用IELM-PNN做了23个月的预报外推. 通过对学习能力、预测精度的实验对比, FELM-PNN和IELM-PNN均具有较好的时间序列预测精度, 训练速度比梯度下降法的PNN有了很大提高. 训练收敛性方面, IELM-PNN要优于FELM-PNN.

1 PNN模型结构

过程神经元内部包括时空聚合算子, 输入和权值是与时间有关的函数, PNN的结构如图1所示.

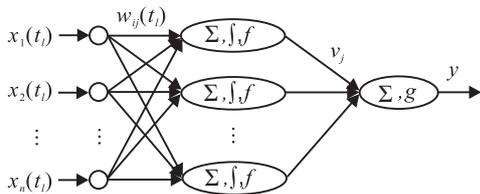


图1 过程神经网络

PNN输入输出映射关系为

$$y = g\left(\sum_{j=1}^m v_j f\left(\sum_{i=1}^n \int_0^T w_{ij}(t) x_i(t) dt - \theta_j\right)\right). \quad (1)$$

选取一组正交基对输入和权函数进行有限项展开, 有

$$x_i(t) = \sum_{l=1}^L a_i^{(l)} b_l(t), \quad w_{ij}(t) = \sum_{l=1}^L w_{ij}^{(l)} b_l(t). \quad (2)$$

根据正交基的正交特性, 式(1)可整理为

$$y = g\left(\sum_{j=1}^m v_j f\left(\sum_{i=1}^n \sum_{l=1}^L w_{ij}^{(l)} a_i^{(l)} - \theta_j\right)\right). \quad (3)$$

2 ELM-PNN学习算法

2.1 FELM-PNN

固定型FELM-PNN是在PNN的输入函数和权函数正交基展开后, 对基函数系数、隐层神经元的阈值进行随机赋值, 将PNN的训练问题转化为最小二乘问题, 即寻找一组合适的隐层输出权值, 使模型以最小的误差逼近非线性系统. 有关ELM理论详见文献^[11]. 下面给出FELM-PNN算法的具体步骤.

Step 1: 通过“试凑法”确定隐层节点数量, 并对权函数基展开系数和隐层神经元阈值进行随机赋值.

Step 2: 计算隐层输出矩阵. 设给定的K个样本为

$$(x_{k1}(t_l), x_{k2}(t_l), \dots, x_{kn}(t_l); d_k), \quad k = 1, 2, \dots, K.$$

为表述方便, 记

$$f(a_{il}^{(k)}, w_{mi}^l, \theta_m) = f\left(\sum_{i=1}^n \sum_{l=1}^L a_{il}^{(k)} w_{mi}^l + \theta_m\right).$$

FELM-PNN隐层输出矩阵为

$$H = \begin{bmatrix} f(a_{i1}^{(1)}, w_{1i}^1, \theta_1) & \cdots & f(a_{in}^{(1)}, w_{ni}^1, \theta_m) \\ \vdots & \ddots & \vdots \\ f(a_{i1}^{(K)}, w_{1i}^K, \theta_1) & \cdots & f(a_{in}^{(K)}, w_{ni}^K, \theta_m) \end{bmatrix}_{K \times m}, \quad (4)$$

FELM-PNN输入输出映射关系可表示为

$$HV = O, \quad (5)$$

其中 $O = [d_1, d_2, \dots, d_K]_{K \times 1}^T$ 和 $V = [v_1, v_2, \dots, v_m]_{m \times 1}^T$ 分别为期望输出向量和隐层输出权值向量.

Step 3: 计算隐层输出权值. 由ELM理论, FELM-PNN隐层输出权值为

$$V = (H^T H)^{-1} H^T O. \quad (6)$$

记 $U = H^T H$, 根据SVD法将U分解为

$$U = P \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} Q^T. \quad (7)$$

其中: $\text{rand}(U) = r$, P和Q是正交矩阵, Σ_r 由U的特征值平方根通过对角排列组成. Moore-Penrose广义逆矩阵 U^\dagger 为

$$U^\dagger = Q \begin{bmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{bmatrix} P^T. \quad (8)$$

按式(6), FELM-PNN隐层的输出权值为

$$V = Q \begin{bmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{bmatrix} P^T H^T O. \quad (9)$$

2.2 IELM-PNN

FELM-PNN在本质上是通过固定隐层节点规模, 利用最小二乘原理快速求解隐层输出权值, 计算速

度快,有效解决了PNN训练速度慢的问题.但存在的问题有:1)训练收敛性不高,经实验证明仅为40%左右;2)隐层节点不易确定,需要在多次实验中不断比较误差来确定,较为耗时;3)试凑法确定的网络结构并非最佳,可能存在冗余节点,影响模型的逼近精度. Huang等^[16]给出了以下定理,确定隐层输出矩阵、输出权值和网络误差三者的关系.

定理1 一个具有任意分段连续激励函数的单隐层前馈神经网络,若其隐层参数 w 、 θ 随机选取,则对于任意连续目标函数 y ,一定存在 N 个隐层节点使网络收敛,即

$$\|y - (y_{N-1} + H_{\tilde{N}}(w, \theta, x)v_{\tilde{N}})\| = 0. \quad (10)$$

其中: y_{N-1} 为 $N-1$ 个隐层节点的网络映射函数, $H_{\tilde{N}}$ 和 $v_{\tilde{N}}$ 分别为第 N 个隐层节点的输出矩阵和输出权值,而

$$v_{\tilde{N}} = \frac{\langle y - y_{N-1}, H_{\tilde{N}} \rangle}{\|H_{\tilde{N}}\|^2}. \quad (11)$$

基于定理1,本文给出一种IELM-PNN学习算法,通过逐次增加隐层过程神经元节点,利用已有节点的训练误差和增加节点的输出矩阵,按照定理1计算新增隐层节点的输出权值.此外,PNN基展开后权函数系数较多,ELM随机赋值方式很难得到最优值,因此,IELM-PNN对于新增隐层节点的输入权值系数利用标准粒子群算法进行寻优. IELM-PNN的具体步骤如下.

Step 1: 设隐层过程神经元节点数 $N = 0$,训练误差精度 η ,PNN模型映射的输出误差 $E = O$.

Step 2: 隐层节点 $N = N + 1$.

Step 3: 对当前增加的隐层节点的输入权值函数基展开系数 $w_{i,N}^{(l)}$ 和阈值 θ_N 实施粒子编码.

Step 3.1: 按式(11)和(4)分别计算每个粒子代表的隐层节点的 $v_{\tilde{N}}$ 和该节点的输出矩阵 $H_{\tilde{N}}$;

Step 3.2: 计算粒子适应度,完成粒子进化,并更新全局最优粒子;

Step 3.3: 若达到PSO最大迭代次数,则最优粒子为新增加节点的输入权函数基展开系数,转Step 4,否则转Step 3.1.

Step 4: 计算当前网络的输出误差 $E = E - v_{\tilde{N}}H_{\tilde{N}}$.

Step 5: 若 $\|E\| \leq \eta$,则终止训练,否则转Step 2.

3 应用实例

首先用Henon时间序列预测验证本文方法的有效性,然后实际应用到第24周太阳黑子平滑月均值的中长期预测,并进行23个月的预报外推.实验对比方面,分别从算法的学习能力和预测精度

两个方面进行实验对比.对比算法有:正交基展开PNN(OBPNN)^[9]、FELM-PNN、IELM-PNN.此外,与传统神经网络BP-ANN也进行了对比,以检验本文所提出方法的预测精度是否优于传统神经网络.

3.1 Henon混沌时间序列

Henon混沌时间序列是著名的动力学系统,其表达式为

$$\begin{cases} x(t+1) = 1 + y(t) - ax^2(t), \\ y(t+1) = bx(t). \end{cases} \quad (12)$$

当 $a = 1.4, b = 0.3$ 时系统进入混沌状态,此时对Henon进行仿真是动力学系统研究的经典问题,即通过给定的动力学系统轨道演化过程,构造一个非线性系统来模拟原系统.

1)学习样本构造.假定系统初始条件为 $x(0) = y(0) = 0.5$,通过式(12)可产生一个与 t 有关的时间序列数据,取 $t \in [0, 149]$ 的150个数据作为PNN模型的样本数据集,记为 $\{x(r)\}_{r=0}^{149}$.用连续的6个数据拟合成一个函数作为PNN输入,对第7个数据进行预测,以此类推,共计可得到144组样本,随机选择其中的84组作为模型的训练样本,剩余的60组作为测试样本.

2)参数设置.3种方法的正交基展开函数均选用勒让德基函数,正交基函数项数 $L = 21$,误差精度 $\varepsilon = 0.005$.对于FELM-PNN和OBPNN建立1-9-1的网络拓扑结构. OBPNN的最大训练迭代步数为5000次,学习速度 $\alpha = \beta = \gamma = 0.5$,惯性系数 $\eta = 0.1$. IELM-PNN的粒子规模为30,每个粒子编码长度为22位,进化200次.

3)实验对比结果.在算法的学习能力对比方面,针对3种算法和84组训练样本,各独立进行20次训练,训练结果对比如表1所示.在训练时间上,本文所提出的FELM-PNN和IELM-PNN比OBPNN有了很大提高,其中FELM-PNN提高约5~6倍,IELM-PNN提高约3倍,但FELM-PNN训练收敛性不高,仅为40%左右,而IELM-PNN可以达到85%左右.此外,对60组测试样本的预测情况如表2和图2所示,显然,在预测精度上相对于传统的ANN和OBPNN都有了一定的提高,从而表明本文方法是有效的.

表1 3种过程神经网络的训练结果对比

数据集	模型	最少步数	平均步数	平均时间/s	收敛次数
Henon 时间序列	OBPNN	3 896	4 351	76.92	11
	FELM-PNN	134	167	13.74	7
	IELM-PNN	—	—	20.58	18
太阳黑子 平滑月均值	OBPNN	4 263	4 751	78.54	9
	FELM-PNN	153	189	15.59	10
	IELM-PNN	—	—	22.73	17

表2 3种预测模型的预测结果对比

数据集	模型	相关系数 R	确定性系数 E	均方误差 RMSE	平均百分误差 MAPE
Henon 时间序列	BP-ANN	0.929 23	0.866 21	0.112 72	22.407 66
	OBPNN	0.961 73	0.916 42	0.084 32	16.788 54
	FELM-PNN	0.975 46	0.927 31	0.080 19	16.073 46
	IELM-PNN	0.980 12	0.942 78	0.068 21	14.097 73
太阳黑子平滑月均值	BP-ANN	0.896 82	0.886 73	5.740 63	23.202 11
	OBPNN	0.938 31	0.916 23	2.604 27	16.362 79
	FELM-PNN	0.940 97	0.928 58	2.179 18	11.348 93
	IELM-PNN	0.968 54	0.950 07	1.977 21	9.268 46

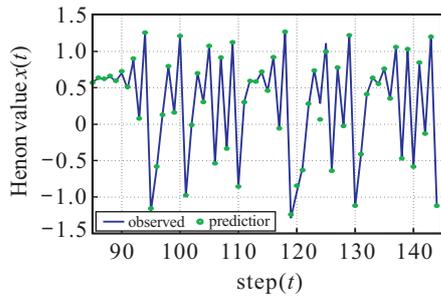


图2 IELM-PNN的Henon时间序列预测

3.2 第24周太阳黑子平滑月均值预测预报

本文选用SIDC目前公布最新的1964年10月~2015年6月的太阳黑子平滑月均值数据,共计609个月。本文利用第20~第23周的太阳黑子数据来预测第24周太阳黑子数,并对SIDC没有公布的2015年7月~2017年6月的太阳黑子数进行外推预报。

对于609个连续太阳黑子平滑月均值数据,用连续的10个作为模型输入来预测第11个太阳黑子数,共计可以得到599组样本。选用前521组作为训练样本,后78组作为测试样本,检验模型的预测能力。这78组测试样本为2008年12月~2015年6月隶属于第24周太阳黑子活动。建立输入输出为1-23-1的网络拓扑结构,设定误差精度 $\epsilon = 0.005$,最大迭代步数为5000。针对3种算法各进行20次独立训练,训练对比如表1所示,预测对比如表2所示。此外,对训练收敛性较高的IELM-PNN在训练完成后,对2008年12月~2015年6月的太阳黑子进行中长期预测,对SIDC目前没有公布的2015年7月~2017年6月的太阳黑子进行外推预报,如图3所示。

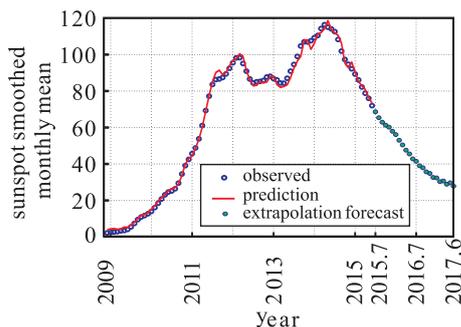


图3 IELM-PNN的太阳黑子平滑月均值中长期预测预报

从表2各种预测指标来看: IELM-PNN对2008年12月~2015年6月的太阳黑子预测是有效的,符合实际数据;对2015年7月~2017年6月的外推预报结果的趋势与文献[17]一致,因此本文所提出的IELM-PNN是有效的。此外,在样本的训练规模上,文献[3]使用1750年1月开始的246年数据来进行预测预报,训练数据规模较大;而本文的数据规模明显减小。

3.3 实验结果分析

1) 预测精度分析。

从表2可以看出,预测精度从高到低依次为IELM-PNN、FELM-PNN、OBPNN和BP-ANN。对时间序列中的 x_{m+1} 预测时,ANN虽考虑了 x_{m-n} , x_{m-n+1}, \dots, x_m 对 x_{m+1} 的影响,但受同步瞬时输入限制,无法同时完成连续离散数据序列的输入,不能体现信号的时间累积效应。IELM-PNN和FELM-PNN都是利用矩阵计算完成模型参数学习,具备ELM良好的泛化特性;而OBPNN受限于梯度下降训练,易陷入局部极值,必然损失精度。此外,IELM-PNN通过逐次增加隐层节点,根据输出误差调整节点数量,网络结构较为紧凑,提高了模型的逼近精度。同时,对新增节点利用PSO进行寻优,相对于随机赋值方式,权函数系数得到了优化,这一点也是FELM-PNN所不具备的。

2) 学习能力分析。

从表1可以看出,学习速度从快到慢依次为FELM-PNN、IELM-PNN和OBPNN,原因在于OBPNN计算过程复杂。此外,权函数基展开后系数较多,对于 $n-m-1$ 的PNN模型结构,若基函数项数为 L ,则模型参数共有 $nmL + m$ 个,利用梯度信息进行调整,迭代次数较多。FELM-PNN只进行向量矩阵计算,通过对隐层参数随机赋值,将输出权值的求解转化为最小二乘问题。IELM-PNN略长于FELM-PNN的原因是,IELM-PNN每次增加节点时,对增加节点的权函数系数不是随机赋值,而是通过一定迭代次数的PSO寻优,计算量稍大。训练收敛性方面,IELM-PNN要高于FELM-PNN,原因在于IELM-PNN逐次增加

隐层节点的方式使得满足训练误差精度的PNN结构更加紧凑,输出误差逐渐逼近训练误差精度.FELM-PNN缺乏这种机制,对于PNN这种内部展开后规模较为庞大的网络模型,随机赋值易导致训练不收敛。

4 结 论

为了提高太阳黑子的预测预报精度,本文提出了FELM-PNN和IELM-PNN两种方法.通过实验验证,两种方法的泛化能力均高于OBPNN,其中FELM-PNN的学习速度最快,训练收敛性略有不足,但对于离线应用问题,FELM-PNN并不影响使用.IELM-PNN通过增量方式逐次增加隐层节点,每次仅计算该节点的输出权值,输入权值利用粒子群进行寻优,达到训练误差精度时的模型结构较为紧凑,因此具有较好的逼近精度,弥补了FELM-PNN的不足.通过对第24周太阳黑子的中长期预测预报,两种方法具有较好的预测精度,与实际数据相符.此外,IELM-PNN对于2015年7月~2017年6月的太阳黑子预报外推,有待于SIDC发布官方数据后方可进行检验。

参考文献(References)

- [1] Roshchira E M, Sarychev A P. Approximation of periodicity in sunspot formation of and prediction of the 25th cycle[J]. *Geomagnetism & Aeronomy*, 2015, 55(7): 892-895.
- [2] Gkana A, Zachilas L. Sunspot numbers: Data analysis, predictions and economic impacts[J]. *J of Engineering Science & Technology Review*, 2015, 8(1): 79-85.
- [3] 赵海娟,王家龙,宗位国.用径向基函数神经网络方法预报太阳黑子数平滑月均值[J]. *地球物理学报*, 2008, 51(1): 31-38.
(Zhao H J, Wang J L, Zong W G. Prediction of the smoothed monthly mean sunspot numbers by means of radial basis function neural networks[J]. *Chines J of Geophysics*, 2008, 51(1): 31-38.)
- [4] Taabu S D, Dujanga F M, Seyonga T S. Prediction of ionospheric scintillation using neural network over east african region during ascending phase of cycle 24[J]. *Advances in Space Research*, 2016, 57(1): 1570-1584.
- [5] Tang Jie, Zhang Xiong. Prediction of smoothed monthly mean sunspot number based on chaos theory[J]. *Acta Physica Sinica*, 2012, 61(16):169601-01-169601-07.
- [6] Ding L G, Lan R S, Jiang Y, et al. Prediction of the smoothed monthly mean sunspot area based on neural network[J]. *Trans of Atmospheric Sciences*, 2012, 35(4): 508-512.
- [7] Gkana A. Prediction of ionospheric scintillation using neural network over east african region during ascending phase of sunspot cycle 24[J]. *Advances in Space Research*, 2016, 57(7): 1570-1584.
- [8] He X G. Procedure neural networks[C]. *Proc of Conf on Intelligent Information*. Beijing: Publishing House of Electronic Industry, 2000: 143-146.
- [9] 许少华,何新贵.基于函数正交基展开的过程神经网络学习算法[J]. *计算机学报*, 2004, 27(5): 645-650.
(Xu S H, He X G. Learning algorithms of process neural networks based on orthogonal function basis expansion[J]. *Chinese J of Computers*, 2004, 27(5): 645-650.)
- [10] 丁刚,钟诗胜.基于时变阈值过程神经网络的太阳黑子数预测[J]. *物理学报*, 2007, 56(2): 1224-1227.
(Ding G, Zhong S S. Sunspot number prediction based on process neural network with time-varying threshold functions[J]. *Acta Physica Sinica*, 2007, 56(2): 1224-1227.)
- [11] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: A new learning scheme of feed forward neural networks[C]. *Proc of Int Joint Conf on Neural Networks*. Budapest, 2004: 985-990.
- [12] Huang G B. Extreme learning machine for regression and multiclass classification[J]. *IEEE Trans on Systems*, 2012, 42(2): 513-516.
- [13] Kasun L L C, Zhou H M, Huang G B. Representational learning with extreme learning machine for big data[J]. *IEEE Intelligent System*, 2013, 28(6): 31-34.
- [14] 贺彦林,王晓,朱群雄.基于主成分分析一改进的极限学习机方法的精对苯二甲酸醋酸含量软测量[J]. *控制理论与应用*, 2015, 32(1): 80-85.
(He Y L, Wang X, Zhu Q X. Modeling of acetic acid content in purified terephthalic acid solvent column using principal component analysis based improved extreme learning machine[J]. *Control Theory & Applications*, 2015, 32(1): 80-85.)
- [15] 王超,王建辉.改进式混合增量极限学习机算法[J]. *控制与决策*, 2015, 30(11): 1982-1985.
(Wang C, Wang J H. Improved hybrid incremental extreme learning machine algorithm[J]. *Control and Decision*, 2015, 30(11): 1982-1985.)
- [16] Huang G B, Chei L, Siew C K. Universal approximation using incremental constructive feedforward networks with random hidden nodes[J]. *IEEE Trans on Neural Networks*, 2006, 17(4): 879-892.
- [17] Kane R P. A preliminary estimate of the size of the coming solar cycle 24, based on Ohl's precursor method[J]. *Solar Physics*, 2007, 243(2): 205-217.

(责任编辑:李君玲)