



控制理论与应用  
Control Theory & Applications  
ISSN 1000-8152, CN 44-1240/TP

## 《控制理论与应用》网络首发论文

题目：基于稀疏表示剪枝集成建模的烧结终点位置智能预测  
作者：周平，吴忠卫，张瑞焱，吴永建  
网络首发日期：2022-11-07  
引用格式：周平，吴忠卫，张瑞焱，吴永建. 基于稀疏表示剪枝集成建模的烧结终点位置智能预测[J/OL]. 控制理论与应用.  
<https://kns.cnki.net/kcms/detail/44.1240.TP.20221107.0941.006.html>



**网络首发：**在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

**出版确认：**纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

## 基于稀疏表示剪枝集成建模的烧结终点位置智能预测

周平<sup>1,2†</sup>, 吴忠卫<sup>1</sup>, 张瑞垚<sup>1</sup>, 吴永建<sup>1</sup>

(1. 东北大学流程工业综合自动化国家重点实验室, 辽宁 沈阳 110819;

2. 中国矿业大学煤炭加工与高效洁净利用教育部重点实验室, 江苏 徐州 221116)

**摘要:** 烧结终点位置(Burning through point, BTP)是烧结过程至关重要的参数,直接决定着最终烧结矿的质量.由于BTP难以直接在线检测,因此通过智能学习建模来实现BTP的在线预测并在此基础上进行操作参数调节对提高烧结矿质量具有重要意义.针对这一实际工程问题,首先提出一种基于遗传优化的Wrapper特征选择方法,可选取使后续预测建模性能最优的特征组合;在此基础上,为了解决单一学习器容易过拟合的问题,提出基于随机权神经网络(Random vector functional-link networks, RVFLNs)的稀疏表示剪枝(Sparse representation pruning, SRP)集成建模算法,即SRP-ERVFLNs算法.所提算法采用建模速度快、泛化性能好的RVFLNs作为个体基学习器,采用对基学习器基函数与隐层节点数等参数进行扰动的方式来增加集成学习子模型间的差异性;同时,为了进一步提高集成模型的泛化性能与计算效率,引入稀疏表示剪枝算法,实现对集成模型的高效剪枝;最后,将所提算法用于烧结过程BTP的预测建模.工业数据实验表明所提方法相比于其它方法具有更好的预测精度、泛化性能和计算效率.

**关键词:** 智能预测;特征选择;集成学习;稀疏表示;剪枝;烧结终点位置;随机权神经网络(RVFLNs)

**引用格式:** 周平, 吴忠卫, 张瑞垚, 吴永建. 基于稀疏表示剪枝集成建模的烧结终点位置智能预测. 控制理论与应用.

## Intelligent Prediction of Burning Through Point Based on Sparse Representation Pruning Ensemble Modeling

ZHOU Ping<sup>1,2†</sup>, WU Zhong-wei<sup>1</sup>, ZHANG Rui-yao<sup>1</sup>, WU Yong-jian<sup>1</sup>

(State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang Liaoning 110819, China;

2. Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Science, Xuzhou Jiangsu 221116, China)

**Abstract:** Burning through point (BTP) is a crucial parameter in sintering process, which directly determines the quality of the final sinter. Since the BTP is difficult to directly detect online, it is of great significance to realize the online prediction of BTP through intelligent learning modeling and adjust the operating parameters on this basis to improve the quality of sinter. Aiming at this practical engineering problem, a Wrapper feature selection method based on genetic algorithm is firstly proposed in this paper, which can select the feature combination that optimizes the subsequent predictive modeling performance as much as possible; Secondly, in order to solve the problem of easy overfitting in intelligent modeling of a single learner, a sparse representation pruning (SRP) ensemble modeling algorithm based on random vector functional-link networks (RVFLNs) is proposed, namely SRP-ERVFLNs. The proposed method uses RVFLNs with fast modeling speed and good generalization performance as individual base learners, and perturbs the parameters of the base learner to increase the difference between the ensemble learning sub-models; At the same time, in order to further improve the generalization performance and computational efficiency of the ensemble model, a sparse representation pruning algorithm is introduced to achieve effective pruning of the ensemble model; Finally, the proposed SRP-ERVFLNs algorithm is used for prediction modeling of the BTP in the sintering process. Experiments using industrial data show that the proposed method has better prediction accuracy, generalization performance and computational efficiency than other methods.

**Key words:** Intelligent prediction; feature selection; ensemble learning; sparse representation; pruning; burning through point (BTP); random vector functional-link networks (RVFLNs)

**Citation:** ZHOU Ping, WU Zhong-wei, ZHANG Rui-yao, WU Yong-jian. Intelligent Prediction of Burning Through Point Based on Sparse Representation Pruning Ensemble Modeling. *Control Theory & Applications*.

<sup>†</sup>通信作者. E-mail: zhouping@mail.neu.edu.cn; Tel.: +86 15840310526.

国家自然科学基金资助(61890934, 61890930, 61991400), 兴辽英才项目资助(XLYC1907132).

Supported by the National Natural Science Foundation of China (61890934, 61890930, 61991400) and Liaoning Revitalization Talents Program(XLYC1907132).

## 1 引言

随着钢铁行业的发展,对铁矿石的需求量不断增加,但可用来冶炼的富矿越来越少,为此通常采用烧结法来制造富矿,提高贫矿的利用效率。钢铁制造中,整个钢铁生产的能耗与排放水平受制于最前端的烧结工序,高炉冶炼的铁水质量与前端烧结工序的烧结矿质量也密切相关,因此烧结过程在钢铁行业有着举足轻重的地位。如图1所示,整个烧结系统由烧结台车、布料圆辊、点火炉、保温炉、风箱等部分组成<sup>[1-3]</sup>。烧结生产时,铁矿粉、焦炭、熔剂(石灰石、白云石)以及返矿等根据生产要求按照一定的比例进行配料,然后经过一次混合加水进行充分混合,再经过二次混合加水进行精细制粒。接着布料圆辊将烧结原料平铺到烧结台车上,经过点火炉的前、后两排点火器进行点火,并在烧结机下方依次排列的风箱强力抽风作用下,烧结料层从上到下进行燃烧。燃烧的料层随着烧结台车不断前进,直到料层烧透。料层烧透的位置即为烧结终点位置(Burning through point, BTP)。烧透的烧结矿在机尾被卸下,经过环冷机冷却后被破碎机粉碎,然后进行筛分。筛分后符合工艺要求的烧结矿被运往高炉,而不合格的烧结矿主要作为返矿返回到配料过程返矿仓<sup>[2-4]</sup>。此外,台车下风箱的废气通过大烟道进入脱硫装置脱硫后最终进入发电设备,而台车前端保温炉的作用是保持刚点燃受热后的料层表面温度以防止急冷,其热量一般来源为环冷机的余热回收<sup>[3]</sup>。

烧结过程中,一般需要将BTP控制在倒数第二个风箱附近,即图1中的第23号风箱附近。如果BTP提前,则烧结机的有效烧结面积没有得到充分利用,从而降低了烧结矿的产量;而BTP延后,则烧结料层燃烧不充分,从而导致烧结矿质量下降,返矿量增加。由于BTP难以采用现有仪表进行直接在线检测,因此对BTP进行预测并根据预测结果调节过程参数显得尤为重要。目

前常见的BTP预测模型主要有机理模型<sup>[5-7]</sup>、推理模型<sup>[8-10]</sup>、数据驱动模型<sup>[11-15]</sup>等几类。由于烧结过程物理化学反应极其复杂,数学模型难以建立,基于机理的BTP预测模型因为假设条件过多而难以应用。基于特殊工况或者已知工况的推理模型很难应对复杂多变的工况变化,因而难以长时间准确预测BTP,实用性差。数据驱动BTP预测建模不需要过多了解烧结内部运行的机理,具有建模方法简单、预测准确度高和易工程实现等诸多优点,因此成为近年BTP预测建模的热点。早在2006年,文献[11]就建立了基于支持向量机的BTP预测模型,通过烧结生产过程数据进行了实验验证;文献[12]建立了一种基于改进Takagi-Sugeno模糊模型且具有线性时变参数的BTP预测模型,取得较好的实验验证效果;文献[13]将台车速度和点火温度作为输入,采用神经网络对BTP进行预测,预测结果可一定程度反应实际BTP的变化趋势。上述模型虽然在BTP预测中都取得了一定效果,但是单一学习器建模极易发生过拟合问题,可靠性不足。众多研究表明:具有多学习器的集成学习能够克服单一学习器过拟合的问题,获得比单一学习器更高和更稳定的预测建模性能<sup>[16-17]</sup>。如文献[16]采用AdaBoost.RT集成学习算法实现了烧结BTP的集成预测建模,取得了较好的预测效果。此外,由于烧结过程变量维度较高,如果全部变量都用于数据模型的建立,不仅会加大计算复杂度,而且会由于冗余信息的增加从而降低模型的性能,因此在建模之前进行特征选择非常必要。目前关于BTP数据建模以及其它数据建模文献中通常采用基于皮尔逊相关性分析的特征选择方法或者基于专家经验的特征选择方法。但是,由于这些方法没有将后续模型性能作为特征选择的评价指标,是一种无监督的特征选择方法,难以得到使后续预测模型性能最佳的建模特征变量。显然,这将很大程度限制后续BTP预测建模的最佳性能。

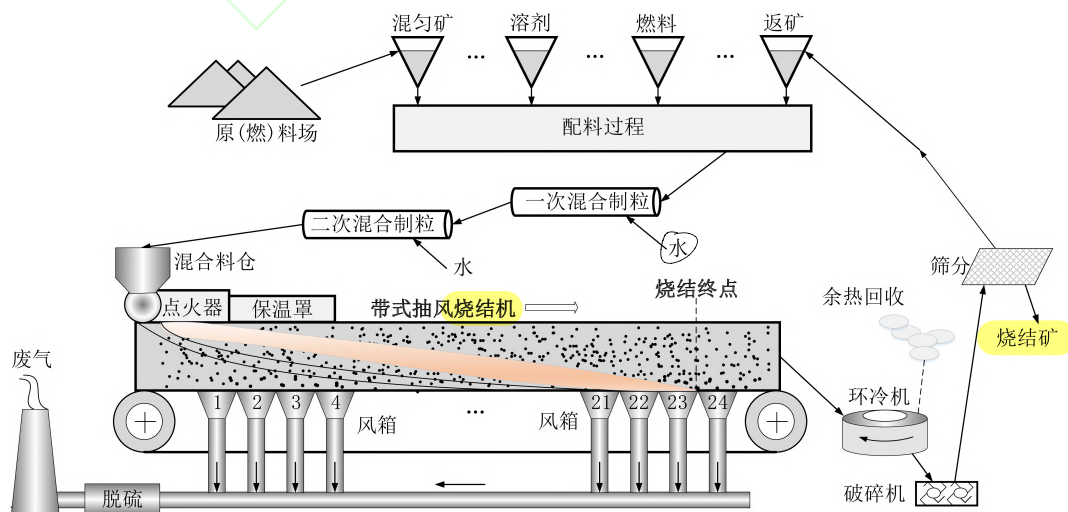


图1 带式抽风烧结工艺流程示意图

Fig. 1 Schematic diagram of the belt-type sintering process



此外,现有大部分集成建模工作,如上文提到的文献[16]的工作,其最终BTP集成模型输出是集成了所有子学习器模型的输出,并没有考虑剔除效果不好和冗余的子学习器模型.实际上,较多数量的子模型不仅会造成预测建模的计算复杂度显著增大,而且最终预测效果也不一定最佳.这意味着集成模型并不是规模越大越好,集成子模型中的一部分通常会比集成所有子模型效果更好<sup>[17]</sup>.基于这一事实,文献[17]提出一种集成神经网络的剪枝算法GASEN,具体如下:首先,先训练所有子模型,然后对所有子模型进行随机权值的分配,并采用遗传算法对各个权值进行改进,最后根据子模型的权重选择有效的子模型集合.文献[18]介绍了一种用于集成模型剪枝的有序剪枝算法.该方法是一种贪婪算法,在每次迭代中选择一个子模型,使已经选择过的子模型集合在训练数据上具有最小的均方误差.虽然这些算法在很大程度上提高了模型的泛化能力,但是相当耗时,满足不了实时预测的要求.针对这一问题,文献[19–20]引入稀疏表示理论来对神经网络集成模型进行剪枝,获得了更高的剪枝效率和更简洁的集成模型.受此启发,本文拟将集成模型的剪枝问题转化为信号的稀疏表示问题,提出稀疏表示剪枝(SRP)算法,实现集成模型的高性能剪枝.

针对上述实际工程问题以及现有方法的不足,本文提出一种新型的基于特征优选与稀疏表示剪枝集成建模的BTP智能预测方法.首先,为了得到使后续预测模型性能最优的建模特征变量,采用基于遗传优化的有监督Wrapper特征选择方法.该方法将后续预测模型的预测精度作为特征选择的评价指标,从而可选取使预测模型性能最佳的特征组合.为了保证预测建模的精度,防止过拟合,提出基于随机权神经网络(RVFLNs)的稀疏表示剪枝(SRP)集成算法,即SRP-ERVFLNs算法.所提算法采用建模速度快、泛化性能好的RVFLNs作为个体基学习器,采用对基学习器参数进行扰动的方式来增加集成学习中子模型间的差异性.为了进一步提高集成学习建模的泛化能力和计算速度,引入稀疏表示算法,实现对智能集成模型的剪枝.基于实际工业数据的烧结过程数据实验表明了所提方法在BTP预测建模的有效性、先进性和实用性.

## 2 基于遗传优化的Wrapper特征优选算法

皮尔逊相关性分析由于计算量小、速度快而成为特征选择的常用方法,但是该方法的评价准则并未将预测模型性能考虑在内,因此选取的特征子集不能完全满足后续模型预测的高精度需

求.而Wrapper特征选择算法可将预测模型的精度作为评价准则,弥补了皮尔逊相关性分析的缺陷.

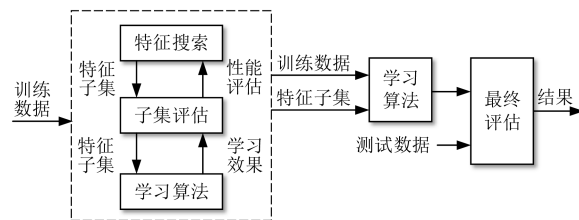


图2 Wrapper特征选择算法流程图

Fig. 2 Flowchart of wrapper feature selection algorithm

Wrapper特征选择算法<sup>[21]</sup>的基本思想是在每一次迭代中,先使用一定的特征搜索策略得到候选特征子集,然后使用后续学习建模算法的预测精度作为评价标准来给候选特征子集打分,满足特定的停止准则时停止迭代,将最优评分的特征子集作为最终结果.因此,对于一个确定的预测模型,Wrapper特征选择算法能够得到使这个模型预测精度达到最优的一个特征子集.Wrapper特征选择算法的实现流程如图2所示.通常,Wrapper特征选择算法选取最优子集的最理想搜索策略是穷举搜索.然而,实际工程问题的特征众多,会使得特征搜索面临很大的计算压力.另一方面,作为启发式随机搜索算法的遗传算法(GA)<sup>[22]</sup>具有良好的全局搜索能力以及较快的求解速度.因而,将GA与Wrapper特征选择算法相结合能够较快的选取使预测模型性能最佳的特征组合.基于该思路,本文提出基于遗传优化的Wrapper特征选择优选算法,如图3所示.遗传优化算法首先初始化一群染色体,这些染色体代表针对不同目标问题的相应解决方案.然后,根据它们的适应度值大小即解决方案的好坏程度确定不同的繁殖机会,给适应度值大的染色体赋予大的繁殖机会.接着通过交叉操作将优秀的染色体基因传给下一代,同时也需要模仿自然界进化过程中存在的变异现象进行变异操作,为新个体的产生提供可能.新的染色体比上几代的染色体更加适应环境,也即针对目标问题的解决方案会越来越好.基于遗传优化的Wrapper特征选择方法的具体步骤如下:

- ① 进行基因编码:“0”表示特征未被选中,“1”表示特征被选中;
- ② 进行种群初始化,生成 $n$ 组不同的特征组合;
- ③ 将后续学习器预测的均方误差的负值作为适应度函数,计算每个个体的适应度值;
- ④ 进行遗传优化的选择、交叉、变异操作;
- ⑤ 未达到设定繁殖代数,转步骤3),否则,转步骤6);
- ⑥ 输出最好的个体即最佳的特征组合.

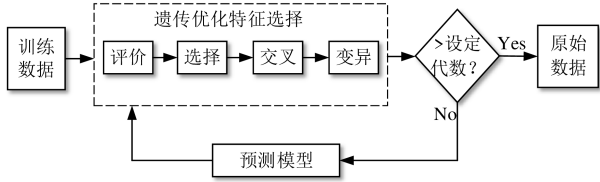


图3 基于遗传优化的Wrapper特征选择

Fig. 3 Wrapper feature selection based on GA

### 3 SRP-ERVFLNs预测建模算法

#### 3.1 算法结构与策略

众多研究表明:集成模型相比于单一学习器模型具有更高的预测精度和泛化能力<sup>[17]</sup>.目前,集成算法主要有Bagging、Boosting、Stacking三种.基于Bagging思想且具有M个基学习器的集成学习首先对原始数据进行M次Bootstrap采样,也即有放回采样得到M个训练数据子集,然后在每个训练子集上训练各个体基学习器,最后将M个基学习器的回归结果通过进行集成,从而得到最终学习结果.该集成学习算法能够有效提高模型的预测精度和防止过拟合.为此,为了实现对BTP的稳定、可靠预测,提出基于Bagging的改进集成学习算法,即SRP-ERVFLNs算法,其整体结构如图4所示.

集成学习要求各基学习器“好而不同”,即基学习器不仅要有一定的精度而且各基学习器个体间要

有差异性.本文采用随机神经网络(RVFLNs)作为基学习器.首先,从单个基学习器的性能来看,RVFLNs具有学习速度快、泛化性能好等优点;其次,在集成学习方法的使用过程中,基学习器的多样性也是影响模型性能的重要因素,而RVFLNs的输入权重和偏置值是在给定范围内随机生成,因此各RVFLNs基学习器具备个体差异性.Bagging中的Bootstrap采样是有放回采样,它是样本数据扰动常用的方法,其目的是增加个体子模型的多样性.在此基础上,所提算法为了进一步增加个体子模型的多样性,对基学习器的参数进行扰动,具体策略为:RVFLNs的激活函数在Sigmoid、Sine、Hardlim和Gaussian中进行多样性选择,隐含层节点数在一定范围内随机生成,最终得到基于Bagging的集成RVFLNs,即ERVFLNs算法.

对集成模型进行剪枝不仅可以删除冗余和性能差的子模型,进而提高预测效果,而且更简洁的集成规模还可以降低存储要求,提高预测计算的速度.因此,进一步提出集成学习性能优化的稀疏表示剪枝(SRP)算法.SRP剪枝算法将剪枝问题转化为稀疏表示问题,然后利用正交匹配追踪(Orthogonal matching pursuit, OMP)算法进行求解.

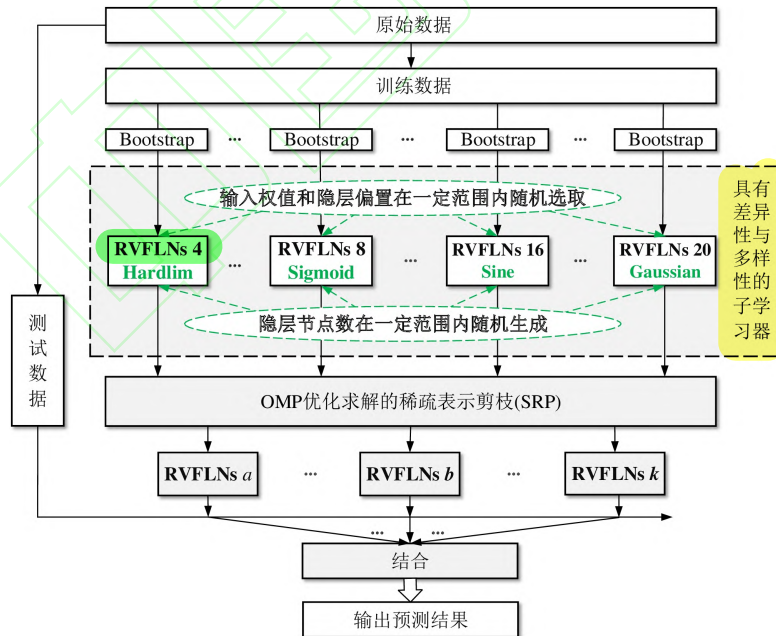


图4 所提SRP-ERVFLNs算法结构图

Fig. 4 SRP-ERVFLNs algorithm structure diagram

#### 3.2 RVFLNs基学习器

RVFLNs最初是Pao和Takefuji等<sup>[23]</sup>提出的一种单隐层前馈型神经网络,其特点就是神经网络的输

入权值和隐含层偏置在一定范围内任意选取,而输出通过式(1)计算:

$$t_j = \sum_{i=1}^L \beta_i J(\alpha_i x_j + b_i), \quad (1)$$

式中:  $L$  是隐含层节点数,  $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{in_y}]$  是连接第  $i$  个隐含层神经元和  $n_y$  个输出神经元的权重向量,  $\alpha_i = [\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{in_x}]$  是第  $i$  个隐含层神经元和  $n_x$  个输入神经元的权重向量,  $b_i$  是第  $i$  个隐含层神经元的偏置, 而  $\mathbf{x}_j$  是输入向量。

给定  $N$  个样本的训练数据集  $\{(\mathbf{x}_j, \mathbf{t}_j)\}_{j=1}^N$ , 其中  $\mathbf{x}_j \in \mathbb{R}^{n_x}$  是输入向量,  $\mathbf{t}_j \in \mathbb{R}^{n_y}$  是对应输出向量。从而等式(1)可以写成如下矩阵形式:

$$T = J\beta, \quad (2)$$

式中:

$$J = \begin{bmatrix} J(\alpha_1, b_1, \mathbf{x}_1) & \cdots & J(\alpha_L, b_L, \mathbf{x}_1) \\ \vdots & \ddots & \vdots \\ J(\alpha_1, b_1, \mathbf{x}_N) & \cdots & J(\alpha_L, b_L, \mathbf{x}_N) \end{bmatrix}_{N \times L}$$

$$= \begin{bmatrix} J_1 \\ \vdots \\ J_N \end{bmatrix}, \quad (3)$$

$$\beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_L \end{bmatrix}_{L \times n_y}, \quad T = \begin{bmatrix} \mathbf{t}_1 \\ \vdots \\ \mathbf{t}_N \end{bmatrix}_{N \times n_y},$$

隐含层和输出层之间的输出权重可以通过求解如下最小二乘问题获得:

$$\min_{\beta \in \mathbb{R}^{L \times n_y}} \psi_{RVFLNs} = \frac{1}{2} \|\beta\|^2 + \frac{\lambda}{2} \|T - J\beta\|^2, \quad (4)$$

式中, 第1项是防止过拟合的正则化项, 第2项是误差向量项,  $\lambda$  是惩罚系数。通过将RVFLNs 相对于  $\beta$  的梯度设为0, 可以得到以下等式:

$$\nabla \psi_{RVFLNs} = \beta + \lambda J^T (T - J\beta) = 0, \quad (5)$$

当矩阵  $J$  列满秩, 即训练样本的数量大于隐含层神经元的数量时, 上式可以求解如下:

$$\beta^* = (J^T J + \frac{I_L}{\lambda})^{-1} J^T T, \quad (6)$$

式中,  $I_L$  是维数为  $L$  的单位矩阵。当矩阵  $J$  行满秩时, 即隐含层神经元的数量大于训练样本的数量, RVFLNs 的输出权重可以求解如下:

$$\beta^* = J^T (J^T J + \frac{I_N}{\lambda})^{-1} T, \quad (7)$$

式中,  $I_N$  是维数为  $N$  的单位矩。

### 3.3 稀疏表示剪枝的集成RVFLNs算法

受文献[19–20]启发, 本文将集成RVFLNs模型的剪枝问题转化为信号稀疏表示问题, 从而提出稀疏表示剪枝(SRP)算法, 用于集成模型的高效剪枝和简化, 具体算法如下。

#### 3.3.1 稀疏表示优化问题及求解的OMP算法

稀疏表示理论由Mallat等<sup>[24]</sup>在1993年提出, 经过不断完善和发展, 目前已经被广泛应用在信号处理、图像处理等领域。稀疏表示的基本原理如下: 信号  $\mathbf{y} \in \mathbb{R}^n$  可以被近似表示为几个信号  $\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_m$  的线性组合, 即  $\mathbf{y} = x_1 \mathbf{d}_1 + x_2 \mathbf{d}_2 + \dots + x_m \mathbf{d}_m$ , 其中  $m > n, \forall i, \mathbf{d}_i \in \mathbb{R}^n, x_i$  是表示系数向量  $\mathbf{x}$  中的元素。信号  $\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_m$  称为原子, 其集合  $D = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_m] \in \mathbb{R}^{n \times m}$  称为字典。稀疏表示试图找到可以近似表示  $\mathbf{y}$  的最小原子个数, 即稀疏表示旨在求解如下问题:

$$\begin{aligned} \min \|\mathbf{x}\|_0 \\ \text{s.t. } \|\mathbf{y} - D\mathbf{x}\|_2^2 < \varepsilon \end{aligned} \quad (8)$$

式中:  $\mathbf{x} \in \mathbb{R}^m, \|\mathbf{x}\|_0$  表示  $\mathbf{x}$  的“ $l_0$  范数”, 即  $\mathbf{x}$  中非零元素个数。对  $\varepsilon \neq 0$  的情况,  $m > n$  的条件不是必须的, 当  $n \geq m$  时, 该问题同样适用。

式(8)所示的稀疏表示优化问题为NP难的组合优化问题, 难以直接求解, 通常只能采用贪婪法或凸松弛法来近似求解。为此, 采用正交匹配追踪(OMP)算法<sup>[25]</sup>来解决式(8)的稀疏表示优化问题。OMP是一种贪婪算法, 其基本思想是在每一步迭代中, 从字典  $D$  中选择一个与残差最相似的原子, 并将已经选择的所有原子进行正交化处理, 直到满足约束条件时, 算法终止迭代。约束条件通常是要选取的原子个数也即稀疏度大小。OMP算法的伪代码如算法1所示。

#### 算法1: 稀疏表示剪枝优化求解的OMP算法

INPUT: 信号  $\mathbf{y}$ , 字典  $D$ , 稀疏度  $K$

BEGIN:

初始残差  $\mathbf{r}_0 = \mathbf{y}$ , 支撑索引集  $\Lambda_0 = \emptyset, k = 1$

for  $k = 1 : K$

寻找支撑索引:  $\lambda_k = \arg \max_{i=1, \dots, m} |\langle \mathbf{r}_{k-1}, \mathbf{d}_i \rangle|$

将最相关字典元素的索引加入索引集:

$\Lambda_k = \Lambda_{k-1} \cup \{\lambda_k\}$

更新残差:  $\mathbf{r}_k = \mathbf{y} - D_{\Lambda_k} (D_{\Lambda_k}^T D_{\Lambda_k})^{-1} D_{\Lambda_k}^T \mathbf{y}$

$k \leftarrow k + 1$

end

OUTPUT: 支撑索引集  $\Lambda_k$  (即所选取原子的索引值的集合)

#### 3.3.2 集成模型的SRP算法

本文将集成模型的剪枝问题转化为3.3.1节所述的信号稀疏表示问题, 从而提出集成模型高效剪枝的SRP算法。SRP算法的基本思想如下: 假设一个集成模型包含  $L$  个子学习器模型, 各子学习器模型在训练数据集  $\{\mathbf{x}^i, \mathbf{y}^i\}, i \in \{1, 2, \dots, N_{\text{train}}\}$  上进行



训练后,对一个训练样本,得到 $L$ 个预测值 $\{y_1^i, y_2^i, \dots, y_L^i\}$ .将此 $L$ 个预测值放入一个向量中,记为 $\mathbf{d}^i \in \mathbb{R}^L$ .然后将所有训练样本对应的这些向量放入一个矩阵,记为 $D \in \mathbb{R}^{N_{\text{train}} \times L}$ .因此字典 $D$ 包含所有子模型对 $N_{\text{train}}$ 个样本的预测值.集成模型最终预测值取 $L$ 个子模型预测值的平均值,即集成模型最终预测值 $\mathbf{y}^* = (y^{*1}, y^{*2}, \dots, y^{*N_{\text{train}}})^T$ 通过式(9)得到:

$$\mathbf{y}^* = D\mathbf{v}, \quad (9)$$

式中,  $\mathbf{v} = \frac{1}{L} \mathbf{1}_{L \times 1}$ ,  $\mathbf{1}_{L \times 1}$  为元素均为1的 $L \times 1$ 维向量.与剪枝前的集成模型在训练集上的均方误差相比,SRP剪枝算法旨在不影响均方误差的前提下,尽可能地移除更多的子学习器模型,使得集成模型尽可能简洁.剪枝前和剪枝后集成模型的输出值分别记为 $\mathbf{y}_{\text{bp}}^*$ 和 $\mathbf{y}_{\text{ap}}^*$ ,因而SRP剪枝算法的目的是选择尽可能少的子模型,使得 $\mathbf{y}_{\text{bp}}^*$ 和 $\mathbf{y}_{\text{ap}}^*$ 几乎相等,即:

$$\begin{aligned} \min \|\mathbf{v}\|_0 \\ \text{s.t. } \|\mathbf{y}_{\text{bp}}^* - D\mathbf{v}\|_2^2 < \varepsilon, \end{aligned} \quad (10)$$

采用前文3.3.1节OMP算法求解上述优化问题.首先确定应该选取的子学习器模型的个数 $k$ ,然后使用OMP算法会找到最重要的 $k$ 个子模型.所提稀疏表示剪枝算法的伪代码如算法2所示.

#### 算法2:集成模型高效剪枝的SRP算法

INPUT: 数据集 $\{\mathbf{x}^i, \mathbf{y}^i\}_{i=1}^{N_{\text{train}}}, k$

BEGIN:

在数据集 $\{\mathbf{x}^i, \mathbf{y}^i\}_{i=1}^{N_{\text{train}}}$ 上训练所有子模型

将所有子模型对所有训练数据的预测值放入字典 $D$ 中

通过式(9)计算 $\mathbf{y}_{\text{bp}}^*$ ,也即 $\mathbf{y}_{\text{bp}}^* = D\mathbf{v}$

采用算法1通过求解式(10)获得 $\mathbf{v}^*$ ,也即

$$\mathbf{v}^* = \text{OMP}(D, \mathbf{y}_{\text{bp}}^*, k)$$

求取所选子模型的索引集 $S$ ,即 $S = \{i | v_i^* \neq 0\}$

OUTPUT:  $S$

最后,所提SRP-ERVFLNs预测建模算法的实现步骤如下:

1) 给定训练集 $P_r$ ,测试集 $P_e$ ,在 $P_r$ 上使用Bootstrap采样得到 $K$ 个训练子集 $\{O^1, O^2, \dots, O^K\}$ ,本文 $K = 20$ ;

2) 在训练子集 $\{O^1, O^2, \dots, O^K\}$ 上分别训练20个RVFLNs子学习器模型,这些子学习器的激活函数选择Sigmoid、Sine、Hardlim和Gaussian各5个,每个RVFLNs的隐含层节点数在 $[a, b]$ 内随机选取;

3) 使用SRP算法对集成模型进行剪枝,得到 $k$ 个RVFLNs子学习器模型,  $k < K = 20$ ;

4) 用 $k$ 个RVFLNs子学习器分别对测试集 $P_e$ 进行预测,得到 $k$ 个预测值;

5) 取 $k$ 个预测值的结合(如取 $k$ 个预测值的平均值)作为最终预测结果.

注1: 算法1是解决式(8)所示稀疏表示优化问题常用的OMP算法,而算法2是本文在算法1基础上提出的稀疏表示剪枝(SRP)算法.所提算法2将稀疏表示优化求解理论应用在集成模型的剪枝问题上,可以实现集成模型的高效剪枝.

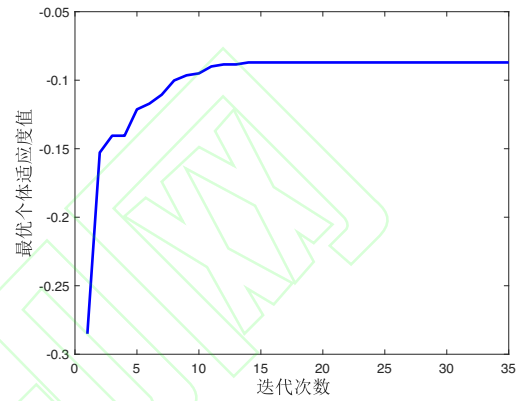


图5 基于遗传优化的Wrapper特征选择过程

Fig. 5 GA-based Wrapper feature selection process

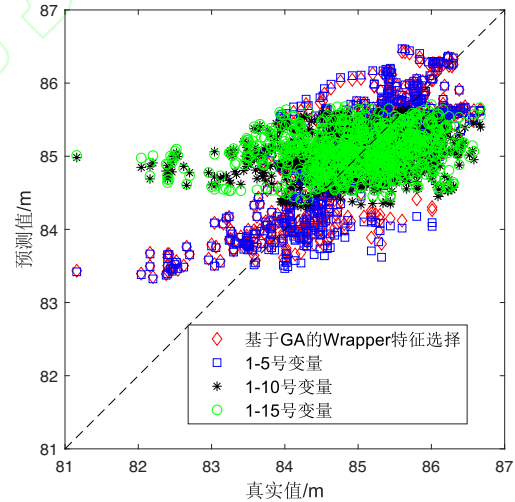


图6 不同特征组合的BTP预测结果比较

Fig. 6 BTP prediction results of different feature combinations

## 4 工业数据实验

### 4.1 基于遗传优化的Wrapper特征选择实验

为了验证所提方法的有效性和先进性,基于柳钢3号360m<sup>2</sup>烧结机的实际生产数据进行工业数据实验.采集到的原始工业数据有26个特征,若从26个特征中选取5个特征,基因串中“1”出现的概率应为5/26.此外,设置初始种群数量为200,采用轮盘赌

选择策略,设置交叉率为0.5,变异率为0.2.由于个体学习器选择RVFLNs,因而将RVFLNs的预测均方误差的负值作为适应度函数,最大迭代次数定为200.为了清晰展示算法迭代过程,在图5中仅绘制算法收敛代数的后25代,可以看出所提算法在15代后即收敛.所提算法输出的最优特征组合为BRP、中间烟道温度、大烟道温度(南)、大烟道温度(北)、大烟道负压(北).为了说明基于遗传优化的Wrapper特征选择算法的优越性,将其与皮尔逊相关性分析方法做比较.为此,采用皮尔逊相关性分析方法分析主要26个特征变量与BTP的相关性,结果如表1所示.然后,使用同样激活函数和隐含层节点数的RVFLNs,将基于遗传优化的Wrapper特征选择算法选取的特征组合,以及皮尔逊相关性分析表1中1-5号、1-10号、1-15号变量分别作为输入变量,对BTP进行预测,预测结果如图6所示.可以看出,使用所提特征选择算法选取的特征所获得的预测结果更接近其真实值,而基于皮尔逊相关性分析得到的1-10号变量、1-15号变量的预测效果都比较差,其原因可能是变量多,引入的噪声干扰较多,从而对预测建模产生不利影响.使用1-5号变量进行预测取得了相对较好的预测效果,但是相比所提算法的预测效果要差.综上分析,相比于传统皮尔逊相关性方法,所提算法能够获得使后续学习器达到更高预测精度的输入特征组合.

## 4.2 BTP预测实验

将所提SRP-ERVFLNs算法在样本扰动的基础上,对基学习器的参数进行扰动,具体策略为:生成20个RVFLNs基学习器,激活函数为Sigmoid、Sine、Hardlim和Gaussian的RVFLNs子学习器各5个,每个子学习器隐含层节点数在一定范围内随机选取.为了在随机选取隐含层节点数的同时保证每个子模型的准确性,需要确定单个RVFLNs的最优隐含层节点范围.图7是不同激活函数的RVFLNs在不同隐含层节点下的预测效果.可以看出随着隐含层节点数,在测试集上的表现越来越差,当隐含层节点数在[10, 20]时,不同激活函数RVFLNs学习器在训练集和测试集上均具有较好的拟合效果和稳定的性能.因此,各RVFLNs子学习器的隐含层节点数在[10, 20]中随机选取,以此增加集成模型各子学习器的个体多样性.

将采集的工业烧结数据分为训练集和测试集,训练集取前400组数据,测试集数据取剩下的100组数据.将前文4.1节得到的最优特征组合作为各个学习器的输入变量,集成剪枝后的子模型数量 $k$ 设为10,按照3.3.2节所提算法实现步骤对BTP进行预

测建模.首先,为了验证所提稀疏剪枝算法的有效性,将剪枝前后的建模效果进行对比.图8是稀疏剪枝前ERVFLNs各子模型建模效果以及最终集成模型的建模效果,而图9为所提方法剪枝后的集成模型各子模型建模效果及最终集成模型建模效果,各个子图中的序号为选取的剪枝前对应子学习器模型的序号.此外,图10为采用所提方法剪枝后的各个子学习器模型以及最终集成模型的实际预测效果比较.从中可以看出,由于对表现不佳的子学习器模型进行了合理剪枝和舍去,使得集成模型剪枝后的输出相比剪枝前可更准确地跟踪BTP实际值的变化,预测精度更高.此外,从图10可以看出剪枝后的集成模型的BTP预测误差PDF形状相比剪枝前更窄和更尖,即预测误差更接近白噪声,随机性小,因而模型准确度更高.综上,剪枝后的集成模型不仅结构更为简单,而且相比于剪枝前的复杂模型具有更高的预测精度和更好的泛化性能,进而也验证了所提剪枝算法的有效性和实用性.

为了进一步验证所提算法的有效性和优越性,将所提算法与常规SVM、RVFLNs、BP神经网络等算法以及未剪枝的集成RVFLNs,即ERVFLNs算法进行对比.各算法实现的一些重要参数选取如下:SVM的激活函数选择Sigmoid函数,采用网格搜索法分别确定SVM中的惩罚因子为1、核函数参数为0.01;BP神经网络的隐含层激活函数为Sigmoid函数,隐含层节点数为20;而RVFLNs的隐含层激活函数选为Sigmoid函数,隐含层节点数确定为20;ERVFLNs算法中的相关参数选取与所提算法一致.

表 1 皮尔逊相关性分析结果

Table 1 Result of Pearson correlation analysis

序号	相关性值(绝对值)	变量名	单位
1	0.83	大烟道温度(北)	℃
2	0.78	大烟道温度(南)	℃
3	0.77	废气温度上升点	m
4	0.67	混合烟道温度	℃
5	0.44	15#风箱负压(南)	kPa
6	0.4	点火器前排温度(B北)	℃
7	0.39	料层厚度	mm
8	0.39	大烟道负压(北)	kPa
9	0.38	大烟道负压(南)	kPa
10	0.38	点火炉膛温度	℃
11	0.36	1#主抽风机出口压力	Pa
12	0.35	2#主抽风机入口压力	Pa
13	0.34	15#风箱负压(北)	kPa
14	0.34	1#主抽风机入口压力	Pa
15	0.32	点火器后排温度(A南)	℃



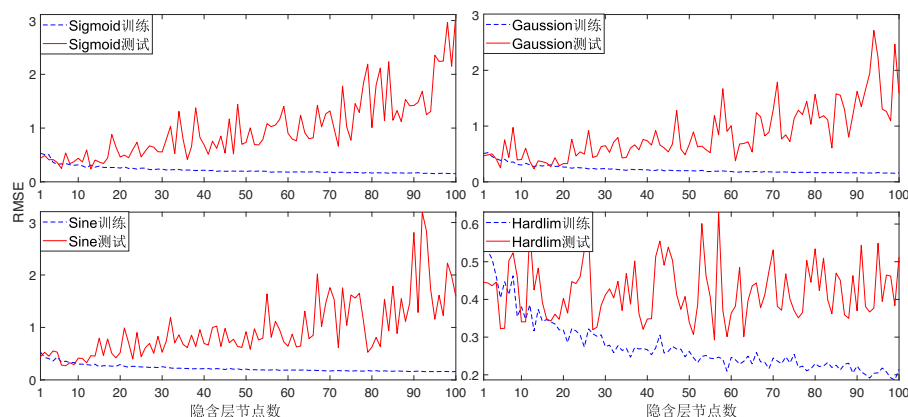


图7 不同激活函数的RVFLNs子学习器预测效果

Fig. 7 The prediction effects of RVFLNs based sub-model with different activation functions

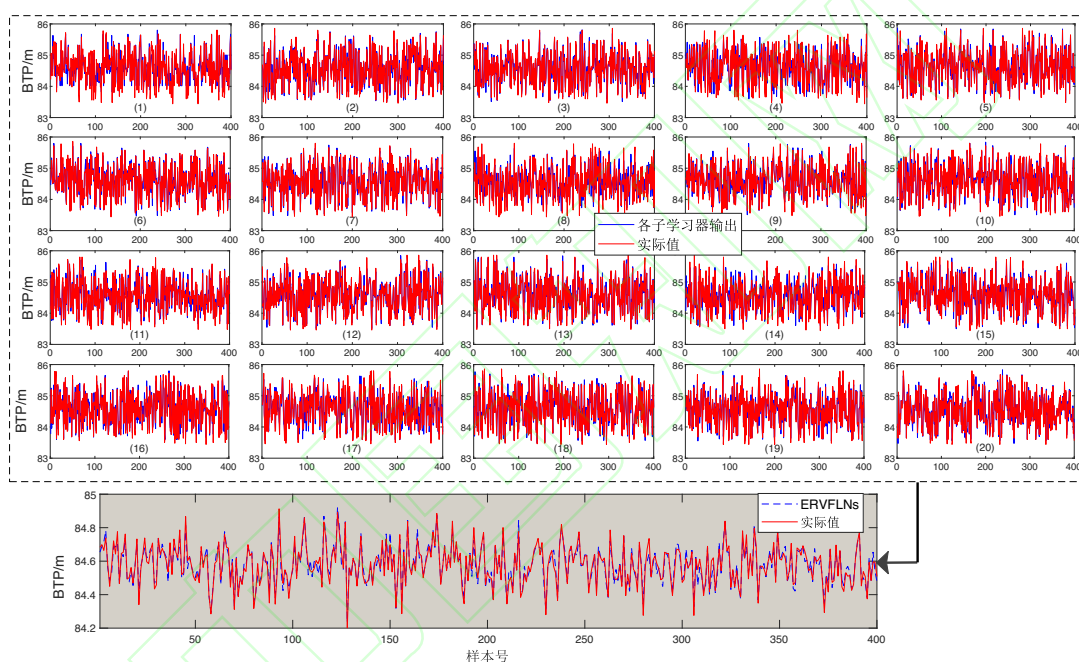


图8 剪枝前的ERVFLNs算法各子学习器预测效果及集成模型最终预测效果

Fig. 8 The prediction effect of each sub-model of ERVFLNs algorithm before pruning and final effect of the integrated model

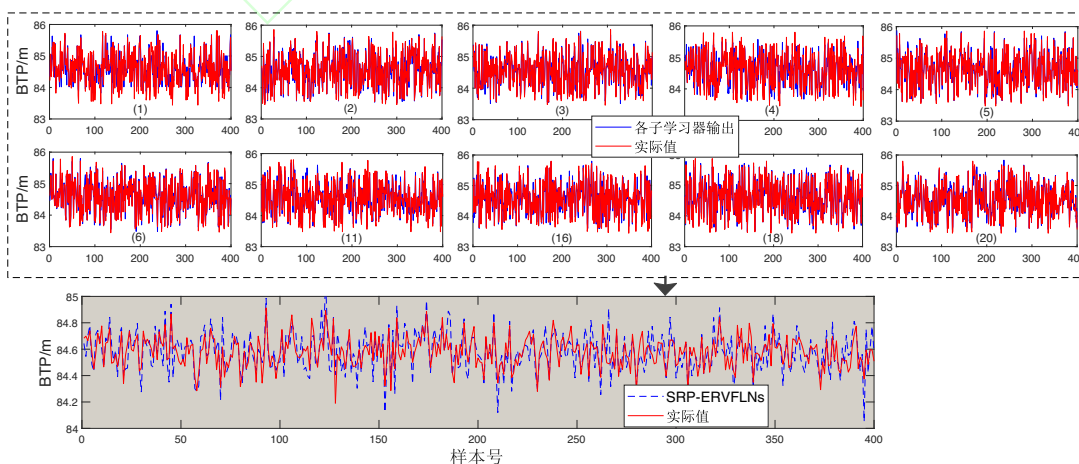


图9 所提SRP-ERVFLNs算法各子学习器预测效果及集成模型最终预测效果

Fig. 9 The prediction effect of each sub-model of the proposed SRP-ERVFLNs algorithm and the final effect of the integrated model

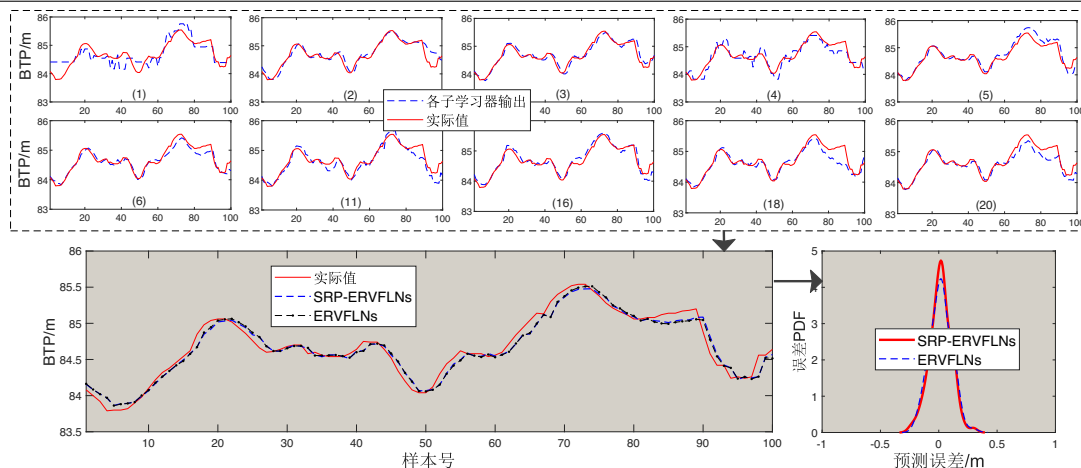


图 10 所提SRP-ERVFLNs算法各子学习器对新数据的预测效果及集成模型最终预测效果

Fig. 10 The prediction effect of each sub-model of the proposed SRP-ERVFLNs algorithm for new data and the final effect of the integrated model

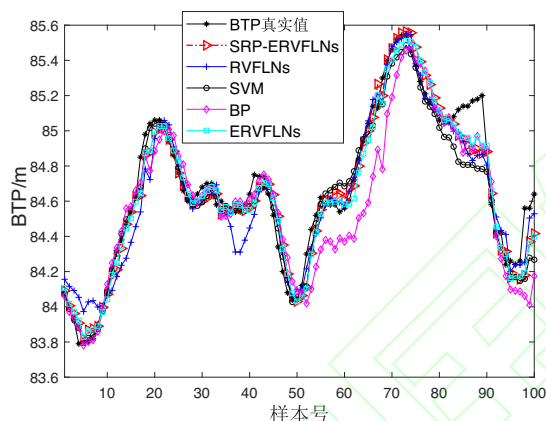


图 11 不同方法的BTP预测比较曲线

Fig. 11 Comparison curve of BTP prediction for different methods

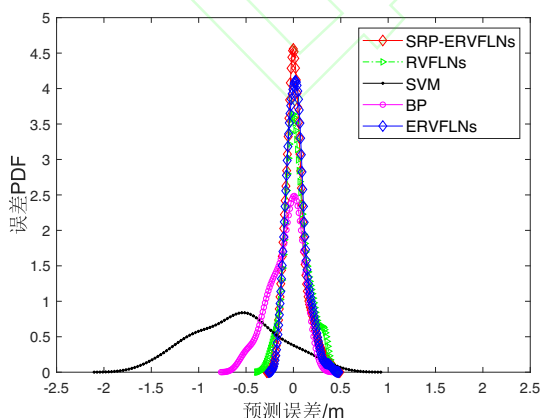


图 12 不同方法的BTP预测误差PDF曲线比较

Fig. 12 BTP prediction error PDF for different methods

图11为所提算法与其他算法的BTP预测效果比较.可以看出,所提算法的BTP预测值能够更好地跟踪BTP实际值的变化趋势.图12为不同方法

的BTP预测误差PDF曲线比较,可以看出所提方法的PDF曲线形状最窄和尖,意味着所提方法建立的预测模型的预测误差更接近随机白噪声,这进一步验证了所提算法的有效性和优越性.表2给出了不同算法的BTP预测均方根误差(RMSE)和计算时间.可以看出,所提算法具有最小的RMSE,也即最高的预测精度.在预测的计算时间上,所提算法所用时间略长于SVM算法所用时间,但是远小于BP神经网络所用时间,而且也小于剪枝前的ERVFLNs算法.以上各个方面充分验证了所提算法能够在保证较高预测精度的基础上实现BTP的快速、稳定预测.

表 2 不同方法BTP预测的RMSE和计算时间比较

Table 2 Comparison of RMSE and prediction time of BTP prediction by different methods

方法	SRP-ERVFLNs	SVM	RVFLNs	BP-NN	ERVFLNs
RMSE/m	0.0829	0.1755	0.1557	0.1761	0.0996
Time/s	0.0094	0.0016	0.0010	0.1066	0.0118

## 5 结论

本文提出了一种新型的基于特征优选与稀疏表示剪枝集成建模的烧结终点位置(BTP)智能预测方法.首先,针对BTP常规预测建模方法难以选取最佳特征组合而影响后续预测建模精度的问题,提出基于遗传优化的新型Wrapper特征选择算法.可选取使后续预测算法性能最优的预测建模特征组合.在此基础上,针对单一学习器容易过拟合而实用性差的问题,提出一种基于随机权神经网络(RVFLNs)的稀疏表示剪枝集成学习算法(SRP-ERVFLNs).该方法采用对基学习器的参数进行扰动的方式来提高RVFLNs子学习器模型的多样性,使用稀疏表示

剪枝算法对集成RVFLNs模型进行剪枝和删减,以进一步提高BTP预测建模的泛化性能与计算速度.实际工业烧结过程数据实验表明:相比剪枝前的ERVFLNs以及常规SVM、RVFLNs和BP神经网络等学习算法,所提SRP-ERVFLNs方法具有更高的预测精度和泛化性能,能够实现BTP的快速准确预测,具有很好实用性.

## 参考文献:

- [1] BAO Yaping, MA Jinyuan, SONG Qiang. Combination forecasting of sintered ore alkalinity based on grey neural network. *Control Theory & Applications*, 2008, 25(4):791 – 793.  
(鲍雅萍, 马金元, 宋强. 基于灰色神经网络的烧结矿碱度组合预测. 控制理论与应用, 2008, 25(4):791 – 793.)
- [2] WANG Chunsheng, WU Min, CAO Weihua, et al. Intelligent Integrated Modeling and Synthetic Optimization for Blending Process in Lead-Zinc Sintering. *ACTA AUTOMATICA SINICA*, 2009, 35(5): 605 – 612.  
(王春生, 吴敏, 曹卫华等. 铅锌烧结配料过程的智能集成建模与综合优化方法. 自动化学报, 2009, 35(5):605 – 612.)
- [3] Du Sheng, Wu Min, Chen Luefeng, et al. Operating performance assessment based on granular clustering for iron ore sintering process. *Acta Automatica Sinica. ACTA AUTOMATICA SINICA*, 2020, 46.  
(杜胜, 吴敏, 陈略峰等. 基于粒度聚类的铁矿石烧结过程运行性能评价. 自动化学报, 2020, 46.)
- [4] Du S, Wu M, Pedrycz W. An intelligent decision-making strategy based on the forecast of abnormal operating mode for iron ore sintering process. *Journal of Process Control*, 2020, 96(12): 57 – 66.
- [5] Li Y X, Yang C J, Sun Y X. Dynamic time features expanding and extracting method for prediction model of sintering process quality index. *IEEE Transactions on Industrial Informatics*, 2022, 18(3): 1737 – 1745.
- [6] Patisson F, Bellot J P, Ablitzer D. Study of moisture transfer during the strand sintering process. *Metallurgical Transactions B*, 1990, 21(1):37 – 47.
- [7] LI B. Arbehthuber C, Jorgl H P, Raml H. Fuzzy control of an iron ore sintering plant. *Control Engineering Practice*, 1995, 3(12):1669 – 1674.
- [8] WU M, DUAN P, CAO W H, et al. An intelligent control system based on prediction of the burn-through point for the sintering process of an iron and steel plant. *Expert Syst. Appl.*, 2012, 39(5):5971 – 5981.
- [9] LYU Qing, LIU Song, LIU Xiaojie, et al. Intelligent control system based on big data technology for whole production line of sintering quality. *Iron & Steel*, 2018, 53(7):7 – 15.  
(吕庆, 刘颂, 刘小杰等. 基于大数据技术的烧结全产线质量智能控制系统. 钢铁, 2018, 53(7):7 – 15.)
- [10] Du S, Wu M, Chen X, Cao W H. An intelligent control strategy for iron ore sintering ignition process based on the prediction of ignition temperature. *IEEE Transactions on Industrial Electronics*, 2020, 67(2):1223 – 1241.
- [11] Wu X F, Fei M R, Wang H S, et al. Prediction of sinter burn-through point based on support vector machines. *Intelligent Control and Automation*, 2006, 344:722 – 730.
- [12] Chen X, Hu J, WU M, Cao W H. T S fuzzy logic based modeling and robust control for burning-through point in sintering process. *IEEE Transactions on Industrial Electronics*, 2017, 64(12):9378 – 9388.
- [13] GUAN Chunpeng. Optimized Sample Selection and System Modeling of Sinter Burning Through Point Prediction Neural Networks. *Measurement & Control Technology*, 2017, 36(3):86 – 89.  
(汪春鹏. 烧结终点预报神经网络样本优选与系统建模. 测控技术, 2017, 36(3):86 – 89.)
- [14] LI Tao, FENG Qiming, FAN Xiaohui. Adaptive Prediction System of Sintering Through Point Based on Self-organizing Artificial Neural Network. *COMPUTER ENGINEERING AND APPLICATIONS*, 2001,37(6): 127 – 129.  
(李桃, 冯其明, 范晓慧等. 基于自组织神经网络的烧结终点自适应预报系统的开发. 计算机工程与应用, 2001,37(6): 127 – 129.)
- [15] Wu M, Xu C H, She J H, et al. Neural-network-based integrated model for predicting burn through point in lead-zinc sintering process. *Journal of Process Control*, 2012, 22: 925 – 934.
- [16] Wang S H, Li H F, Zhang Y J, et al. A hybrid ensemble model based on ELM and improved AdaBoost.RT algorithm for predicting the iron ore sintering characters. *Computational Intelligence and Neuroscience*, 2019, 1 – 11.
- [17] Zhou Z H, Wu J, Tang W. Ensembling neural networks: many could be better than all. *Artificial Intelligence*, 2002, 137(1-2): 239 – 263.
- [18] Martínez-Munoz G, Hernandez-Lobato D, Suarez A. An analysis of ensemble pruning techniques based on ordered aggregation. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2009, 31(2): 245 – 259.
- [19] Yang J, Bouzerdoum A and Phung S L. A neural network pruning approach based on compressive sampling. *2009 International Joint Conference on Neural Networks*, 2009, 3428 – 3435.
- [20] Yang J, Ma J. Feed-forward neural network training using sparse representation. *Expert Syst. Appl.*, 2019, 116: 255 – 264.
- [21] Idris A, Khan A. Churn prediction system for telecom using filter-wrapper and ensemble classification. *Computer Journal*, 2017, 60(3):410 – 430
- [22] SUN Tiejun, YANG Weidong, CHENG Yanming. Improved genetic algorithm for optimizing prediction model of strip coiling temperature. *Control Theory & Applications*, 2015, 32(8): 1106 – 1113.  
(孙铁军, 杨卫东, 程艳明等. 用改进遗传算法优化的带钢卷取温度预报模型. 控制理论与应用, 2015, 32(8): 1106 – 1113.)
- [23] Pao Y H, Takefuji Y. Functional-link net computing: theory, system architecture, and functionalities. *Computer*, 1992, 25(5):76 – 79.
- [24] Mallat S G, Zhang Z F. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 1993, 41(12): 3397 – 3415.
- [25] Tropp J A, Gilbert A C. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on Information Theory*, 2007, 53(12):4655 – 4666.

## 作者简介:

周平 教授, 目前研究方向为工业过程运行反馈控制、数据驱动建模与控制等, E-mail: zhouping@mail.neu.edu.cn;

吴忠卫 硕士研究生, 目前研究方向为数据驱动建模, E-mail: 394784270@qq.com;

张瑞奎 博士研究生, 目前研究方向为数据驱动过程监测, E-mail: zryao\_neu@163.com;

吴永建 研究员, 目前研究方向为流程工业过程建模、控制与优化, E-mail: wyj\_neu@163.com.