

# 高炉铁水质量鲁棒正则化随机权神经网络建模

李温鹏<sup>1</sup> 周平<sup>1</sup>

**摘要** 高炉炼铁过程运行优化与控制依赖于可靠、稳定的难测铁水质量 (Molten iron quality, MIQ) 指标模型. 针对现有 MIQ 建模方法的不足, 本文提出一种新型的数据驱动鲁棒正则化随机权神经网络 (Random vector functional-link networks, RVFLNs) 算法, 用于实现 MIQ 指标在线估计的鲁棒建模. 首先, 为了提高建模效率和降低计算复杂度, 采用数据驱动典型相关性分析方法从众多变量中提取与 MIQ 相关性最强的变量作为建模输入变量; 其次, 由于传统 RVFLNs 网络的输出权值由最小二乘估计获得, 易受离群数据影响而鲁棒性差, 引入基于 Gaussian 分布加权的 M 估计技术, 提出新型鲁棒 RVFLNs 算法建立多元 MIQ 指标的鲁棒模型; 同时, 在鲁棒加权后的最小二乘损失函数基础上, 进一步引入  $L_1$  和  $L_2$  两个正则化项以构成优化目标函数的 Elastic net, 用于稀疏化 RVFLNs 网络的输出权值矩阵, 解决 RVFLNs 网络多重共线性和过拟合的问题. 最后, 基于某大型高炉工业数据, 进行充分数据实验, 结果表明所提方法具有更高的建模与估计精度以及较强的鲁棒性能.

**关键词** RVFLNs, 鲁棒建模, Gaussian 分布加权 M 估计, 高炉炼铁, 铁水质量

**引用格式** 李温鹏, 周平. 高炉铁水质量鲁棒正则化随机权神经网络建模. 自动化学报, 2020, 46(4): 721–733

**DOI** 10.16383/j.aas.2018.c170670

## Robust Regularized RVFLNs Modeling of Molten Iron Quality in Blast Furnace Ironmaking

LI Wen-Peng<sup>1</sup> ZHOU Ping<sup>1</sup>

**Abstract** Optimal operation and control of a practical blast furnace (BF) ironmaking process depend largely on a reliable model of molten iron quality (MIQ) indices that can not be measured online. Aiming at the shortcomings of the existing MIQ modeling methods, a new data-driven robust regularized random vector functional-link networks (RVFLNs) algorithm is proposed to realize robust modeling of MIQ indices. First, to improve modeling efficiency and reduce computational complexity, the data-driven canonical correlation analysis (CCA) is employed to identify the most influential components from multitudinous factors that affect the MIQ indices to serve as the input variables. Next, since the output weights of traditional RVFLNs are obtained by the least squares approach, the robustness may decrease when the training dataset is contaminated with outliers. To solve this problem, the robust RVFLNs model of MIQ using Gaussian distribution weighted M-estimation is established. Simultaneously, on the basis of the least-square loss function of the robustness, the  $L_1$  regularization and  $L_2$  regularization are introduced to achieve sparse output weight and prevent the overfitting and multicollinearity of the RVFLNs model by forming the Elastic net that optimizes the objective function. Finally, experiments using industrial data from a large balst furnace have demonstrated that the proposed method produces a higher modeling, estimating accuracy and stronger robustness than other modeling methods.

**Key words** Random vector functional-link networks (RVFLNs), robust modeling, Gaussian distribution weighted M-estimator, blast furnace ironmaking, molten iron quality (MIQ)

**Citation** Li Wen-Peng, Zhou Ping. Robust regularized RVFLNs modeling of molten iron quality in blast furnace ironmaking. *Acta Automatica Sinica*, 2020, 46(4): 721–733

收稿日期 2017-11-23 录用日期 2018-02-15  
Manuscript received November 23, 2017; accepted February 15, 2018

国家自然科学基金项目 (61890934, 61790572, 61290323), 中央高校基本科研业务费项目 (N180802003), 辽宁省“兴辽英才计划”项目 (XLYC1907132), 矿冶过程自动控制技术国家 (北京市) 重点实验室开放课题 (BGRIMM-KZSKL-2017-04) 资助

Supported by National Natural Science Foundation of China (61890934, 61790572, 61290323), Fundamental Research Funds for the Central Universities (N180802003), Liaoning Revitalization Talents Program (XLYC1907132), State Key Laboratory of Process Automation in Mining & Metallurgy (Beijing) (BGRIMM-KZSKL-2017-04)

本文责任编辑 贺威

Recommended by Associate Editor HE Wei

1. 东北大学流程工业综合自动化国家重点实验室 沈阳 110819

高炉炼铁是钢铁工业的重要生产环节. 由于工艺相对简单、产量大、劳动生产率高, 因而高炉炼铁仍是现代炼铁的最主要方式, 其产量约占世界生铁总产量的 95% 以上. 如图 1 所示, 现代化的高炉炼铁系统分为高炉本体、上料系统、送风系统、喷吹系统、煤气净化系统、渣铁处理系统和环保除尘系统等几个子系统<sup>[1-2]</sup>. 高炉炼铁时, 铁矿石、焦炭、溶剂按一定比例根据布料制度逐层从高炉顶部装载到

1. State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819

炉喉位置. 同时, 在高炉下部, 将预热的空气、氧气和煤粉通过热风炉鼓入炉缸中. 空气、氧气、煤粉和焦炭在高温作用下发生一系列复杂物理化学反应, 生成大量的高温还原性气体, 这些还原性气体不断向上运动将铁从铁矿石中还原出来. 上行气体最终变为高炉煤气从炉顶回收, 而下行炉料则随着炉缸中焦炭的不断燃烧和铁水的不断滴落逐渐向下运动, 在下降过程中, 炉料经过加热、还原、熔化等一系列复杂的物理化学变化, 最终生成液态的铁水和炉渣从出铁口排出<sup>[1-4]</sup>. 高炉炼铁需要准确判断整个高炉运行态势, 及时调整相关操作制度(如布料制度、热风制度等)及工艺参数, 使炉内煤气分布合理、热量充分利用、渣铁顺利排放, 实现高炉生产的优质、高产、低耗和长寿运行. 为了实现这一目标, 就应对高炉运行状态进行实时监测和有效控制. 然而高炉内部冶炼环境极其严酷, 炉内温度高达 2000 多度, 压强高达标准大气压的 4 倍左右, 且伴随着固、液、气多相共存的状态, 使高炉内部状态的实时监测难以实现, 从而难以对高炉进行控制与操作优化<sup>[1,5]</sup>. 目前, 被广泛用来间接反映高炉运行状态的指标为出铁口的铁水质量指标, 包括铁水温度 (Molten iron temperature, MIT)、硅含量 ([Si])、磷含量 ([P]) 和硫含量 ([S]):

1) MIT 是表征高炉热状态、能量消耗的重要指标. MIT 过低不仅影响高炉顺行, 而且会影响后续

转炉炼钢的运行性能, 增加生产成本. 高炉生产操作时应及时掌握 MIT 及变化趋势, 预见性地采取调控措施, 这对于稳定高炉热制度、减少炉况波动以及提高生铁质量和降低焦比等具有重要意义.

2) 铁水 [Si] 是反映铁水化学热的重要指标. 铁水 [Si] 过高, 渣量增加, 会使生铁变硬变脆, 收得率降低. 另外, 高 [Si] 会使渣中 SiO 含量过高, 影响石灰渣化速度, 延长吹炼时间, 同时也会加剧对炉衬的冲蚀. 因此, 必须对 [Si] 进行严格监视.

3) 硫和磷在钢材中均是有害元素, [S] 过高会使铁硬脆, 产生热脆性和减低铁液流动性, 而 [P] 过高会加重炉衬的冲刷蚀损, 影响炉龄. 因此, 在高炉冶炼过程中应严格监视铁水中的 [S] 和 [P].

采用上述铁水质量指标作为高炉运行状态的评判指标可以较全面地了解高炉内部的运行状态, 为高炉日常操作与调节提供指导. 然而, 现有检测技术难以对铁水质量指标进行直接在线测量, 且离线化验过程滞后时间较长, 通常为 1 小时左右. 另外, 高炉炼铁是一个包含气、固、液三相混合和耦合交错, 物理化学反应极其复杂的过程, 其炉内环境极其恶劣, 导致操作人员难以对其运行状态的变化进行实时监测. 因此有必要建立准确可靠的铁水质量模型来反映高炉当前和预期的铁水质量指标参数变化, 为现场操作人员提供炉况和铁水质量信息<sup>[1,5-8]</sup>.

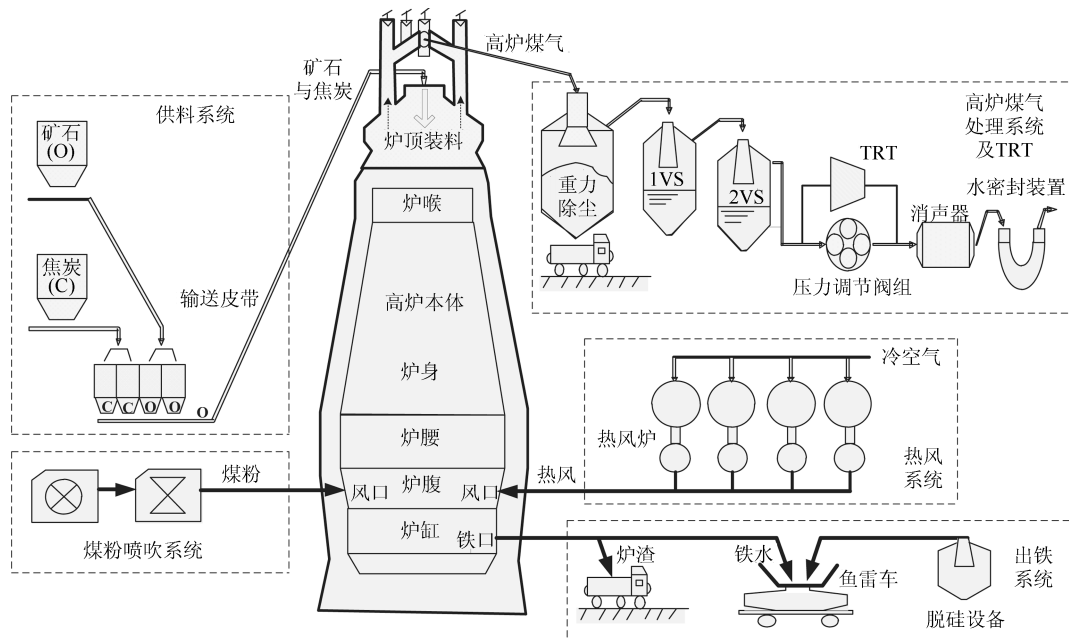


图 1 高炉炼铁工艺示意图

Fig. 1 Diagram of a typical BF ironmaking process

目前常见的铁水质量模型有机理模型、知识推理模型和数据驱动模型三大类<sup>[5]</sup>。机理模型在大量假设条件下, 依赖于高炉炼铁过程的冶炼机理, 从化学反应动力学和流体动力学的角度构建高炉内部反应场, 对反应场内的温度分布、物料运动状态等进行仿真<sup>[9-11]</sup>。高炉炼铁是一个极为复杂、大时滞、高度耦合的非线性动态时变系统, 基于大量假设的机理模型难以用于实际高炉炼铁生产; 知识推理模型完全依赖于系统集成的有限专家知识, 无法应对复杂多样的实际高炉冶炼过程<sup>[12-13]</sup>; 数据驱动铁水质量智能建模不需要了解高炉内部发生的复杂机理, 仅通过数学工具和智能算法对运行数据进行处理而建立需要的铁水质量模型, 因而成为近年高炉炼铁过程建模研究的热点<sup>[1-5, 8]</sup>。

现有数据驱动建模方法主要有多元统计分析方法<sup>[14-15]</sup>、神经网络 (Neural network, NN) 建模方法<sup>[8, 15-18]</sup> 和支持向量回归 (Support vector regression, SVR) 建模方法<sup>[2, 3, 6-7, 19]</sup> 等。其中, 神经网络是一种最常见的数据驱动建模方法, 它是对人脑神经网络进行抽象、简化和模仿所得到的复杂信息处理系统, 能够以任意精度逼近复杂非线性连续函数, 特别适用于处理高炉炼铁这种动态机理复杂、内部状态未知, 仅可获得过程的输入输出数据的复杂工业过程建模<sup>[16]</sup>。常见的神经网络建模算法主要有 BP-NN (BP neural network) 和 RBF-NN (Radial basis function neural network), 但是这些算法普遍存在“过拟合”、易陷入局部极小的问题。另外, 基于批学习的 NN 网络权值和偏差迭代算法容易造成网络训练时间长、收敛速度慢的问题<sup>[16-17, 20-21]</sup>。

近年, 具有学习速度极快且泛化能力好的随机神经网络 (Random vector functional-link networks, RVFLNs) 及其改进算法可有效克服传统神经网络的上述缺点, 被广泛用于高炉炼铁过程建模与控制<sup>[5, 20-21]</sup>。RVFLNs 的最大特点是网络输入权值与偏置在给定范围随机选取, 然后采用最小二乘算法对输出权值进行计算, 因而学习速度极快, 而不像基于梯度的常规 NN 算法需要首先设置很多学习参数, 然后采用批学习策略并花费几分钟甚至几小时时间对众多网络参数进行学习<sup>[5, 21]</sup>。文献 [20] 采用经典 RVFLNs 算法对铁水质量进行建模, 大大加快计算速度; 文献 [21] 进一步采用在线序贯学习型 RVFLNs 实现多元铁水质量的在线软测量建模。实际表明, RVFLNs 比 BP 等常规 NN 具有更高的计算效率, 且 RVFLNs 及其改进算法都具有较高的模型精度。但是, 高炉炼铁生产过程中, 受检测仪表和变送器等装置的故障以及其他异常干扰的影响, 测量数据中经常存在离群点, 即远离数据一般水平的极端大值点或极端小值点。而常规 RVFLNs 的输出

权值是由最小二乘估计得到, 因而鲁棒性不足, 建模时易受离群点干扰, 导致模型准确度下降, 不能为操作人员提供准确的指导。

针对上述问题, 也为了建立稳定、可靠的高炉铁水质量模型, 本文提出一种新型的数据驱动随机神经网络鲁棒建模方法, 用于实现铁水质量指标的鲁棒估计。首先, 为了提高铁水质量建模效率, 采用数据驱动典型相关性分析方法提取与铁水质量指标相关性最强的最主要过程变量作为建模输入变量; 然后, 根据高炉炼铁过程数据特性, 采用基于 Gaussian 分布加权的 M 估计方法建立多元铁水质量指标的鲁棒 RVFLNs 模型; 同时, 在 M 估计鲁棒建模基础上, 进一步引入两个正则化项以防止模型过拟合, 稀疏鲁棒 RVFLNs 的输出权值矩阵。最后, 基于我国华南某大型高炉实际数据建立基于鲁棒正则化 RVFLNs 的多元铁水质量非线性自回归 (Nonlinear autoregressive exogenous, NARX) 模型, 并和其他几类铁水质量建模算法进行对比。结果表明, 相对于对比方法, 本文方法不仅具有更高的模型精度和更快的计算速度, 还解决了常规 RVFLNs 存在的数据建模鲁棒性差的问题。

## 1 基于数据驱动典型相关分析的建模输入变量选择

高炉本体参数较多, 且变量之间存在较强相关性, 如果所有变量全部参与铁水质量建模, 会增加计算复杂度, 影响模型预测的准确性和有效性。现有数据建模研究中, 辅助变量的选取方法主要为成分分析 (Principle component analysis, PCA) 方法<sup>[20-21]</sup>。但是 PCA 技术只针对输入空间进行降维, 未考虑输入输出变量之间的关联性<sup>[1]</sup>。为此, 本文将采用考虑输入输出空间关联关系的典型相关分析 (Canonical correlation analysis, CCA) 方法选择与铁水质量参数相关性最大的几个变量作为输入变量。CCA 是研究多个输入变量与多个输出变量之间相关性的多元统计分析方法, 其基本原理是在两组输入输出关联变量中分别提取有代表性的综合指标, 也就是分别建立两个变量组中各变量的线性组合, 利用这两个综合指标之间的相关关系来反映两组变量之间的整体相关性<sup>[1, 22]</sup>。基于 CCA 的数据降维过程简要描述如下:

基于 CCA 的数据降维需要在两组变量  $X = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p), Y = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_q)$  中分别找出变量的一个线性组合, 即:

$$\begin{cases} U_i = \alpha_1^{(i)} \mathbf{x}_1 + \alpha_2^{(i)} \mathbf{x}_2 + \dots + \alpha_p^{(i)} \mathbf{x}_p = X \alpha^{(i)} \\ V_i = \beta_1^{(i)} \mathbf{y}_1 + \beta_2^{(i)} \mathbf{y}_2 + \dots + \beta_q^{(i)} \mathbf{y}_q = Y \beta^{(i)} \end{cases} \quad (1)$$

式中,  $\boldsymbol{\alpha}^{(i)} = (\alpha_1^{(i)}, \alpha_2^{(i)}, \dots, \alpha_p^{(i)})^T$ ,  $\boldsymbol{\beta}^{(i)} = (\beta_1^{(i)}, \beta_2^{(i)}, \dots, \beta_q^{(i)})^T$  分别为任意非零常系数向量,  $(U_i, V_i)$  为典型变量组.

首先, 寻找通过计算  $\boldsymbol{\alpha}^{(1)}, \boldsymbol{\beta}^{(1)}$ , 分别在每组变量中找到第一对典型变量  $(U_1, V_1)$ , 使其具有最大相关性. 然后, 在每组变量中找到第二对典型变量  $(U_2, V_2)$ , 使其分别与本组内的第一对典型变量  $(U_1, V_1)$  不相关, 并具有次大的相关性. 如此继续下去, 直到两组变量的相关性被提取完为止. 另外, 式(1)中典型变量  $U_i, V_i$  之间的相关系数称为典型相关系数. 可以看出, 典型相关分析研究的问题实际上就是如何选取典型变量的最优线性组合.

在进行变量选择时, 首先要对各对典型变量的典型相关系数进行显著性检验, 若某一对典型变量  $(U_i, V_i)$  的相关程度不显著, 说明这对变量不具有代表性, 舍弃这一对典型变量  $(U_i, V_i)$ . 若相关程度显著, 则保留变量作为最后建模的输入变量. 另外, 在高炉变量中, 各变量均具有一定相关性, 若两个变量相关性达到 80% 以上, 则可近似认为二者呈线性关系, 选择其中一个变量作为建模输入变量即可.

## 2 基于 Gaussian 分布加权 M 估计的鲁棒正则化 RVFLNs 建模算法

### 2.1 常规 RVFLNs

随机权神经网络 (RVFLNs) 由 Pao 等于 1992 年提出<sup>[23-27]</sup>. 不同于传统基于梯度学习的 NN 需要通过误差的反向传播迭代寻优, 采用随机给定神经元输入权值和偏置, 通过正则化原则计算输出权值建立网络. 实际上, 基于单隐层前馈神经网络的 RVFLNs 以全局逼近理论为基础, 保证隐含层节点参数的随机选取不会影响 NN 的学习能力, 同时网络采用 Moore-Penrose 广义逆矩阵方法一步求得输出权重, 使得算法相对于 BP-NN、SVR 等, 其计算速度提高了数千倍, 为实际工业过程实现在线软测量和在线控制奠定基础<sup>[23-27]</sup>.

给定  $M$  组任意样本集合  $(\mathbf{x}_i, \mathbf{y}_i)$ ,  $\mathbf{y}_i = [y_{i1}, y_{i2}, \dots, y_{im}]^T \in \mathbf{R}^m$ ,  $\mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in \mathbf{R}^n$ , 一个带有  $L$  个隐层节点, 以  $g(x)$  为激活函数的标准单隐层前馈神经网络可以表示为:

$$f_{L,i}(\mathbf{x}_i) = \sum_{j=1}^L \beta_j g(\langle \mathbf{a}_j, \mathbf{x}_i \rangle + b_j), \quad i = 1, 2, \dots, M \quad (2)$$

式中,  $f_{L,i}(\mathbf{x}_i)$  是 RVFLNs 的输出,  $\mathbf{a}_j \in \mathbf{R}^n, j = 1, \dots, L$  是连接输入层与隐层的权值矩阵, 且权值在给定确定范围内随机产生,  $\boldsymbol{\beta}_j =$

$[\beta_{j1}, \beta_{j2}, \dots, \beta_{jm}]^T$  是连接第  $j$  个隐含层节点与输出层神经元的输出权重向量,  $b_j$  是第  $j$  个隐含层节点的偏置,  $\langle \mathbf{a}_j, \mathbf{x}_i \rangle$  表示  $\mathbf{a}_j$  与  $\mathbf{x}_i$  的内积.

当  $f_L$  以 0 误差逼近连续系统的  $M$  个样本时, 即  $\sum_{i=0}^M \|f_L - y_i\| = 0$ , 则式(2)可以转化为寻找  $\boldsymbol{\beta}_j, \mathbf{a}_j$  和  $b_j$ , 使得:

$$\sum_{j=1}^L \beta_j g(\langle \mathbf{a}_j, \mathbf{x}_i \rangle + b_j) = y_i, \quad i = 1, 2, \dots, M \quad (3)$$

用矩阵表示为:

$$H\boldsymbol{\beta} = Y \quad (4)$$

式中:

$$H = \begin{bmatrix} g(\langle \mathbf{a}_1, \mathbf{x}_1 \rangle + b_1) & \cdots & \cdots \\ \vdots & \cdots & \cdots \\ g(\langle \mathbf{a}_1, \mathbf{x}_M \rangle + b_1) & \cdots & \cdots \\ \vdots & \cdots & \cdots \\ g(\langle \mathbf{a}_L, \mathbf{x}_1 \rangle + b_L) & \cdots & \cdots \\ \vdots & \cdots & \cdots \\ g(\langle \mathbf{a}_L, \mathbf{x}_M \rangle + b_L) \end{bmatrix}_{M \times L}, \quad Y = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_M \end{bmatrix}_{M \times m}, \quad \boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_1^T \\ \vdots \\ \boldsymbol{\beta}_L^T \end{bmatrix}_{L \times m}$$

通常, 网络隐层节点数会比训练样本少很多, 从而  $H$  不是方阵, 显然式(4)所示方程组没有解. 为了选取最合适的  $\boldsymbol{\beta}$  尽量使式(4)成立, 那么就需要采用最小二乘法求解上述方程组, 如式(5)所示:

$$\boldsymbol{\beta} = H^\dagger Y \quad (5)$$

式中,  $H^\dagger$  是根据 Moore-Penrose 方法计算得到的  $H$  伪逆矩阵.

Igel'nik 等于 1995 年进一步提出 RVFLNs 中激活函数的约束条件以及随机权值和偏置的理论选取范围<sup>[24]</sup>. Igel'nik 等指出: 对于满足万能逼近性的 RVFLNs, 其激活函数  $g(x)$  应当满足:

$$\int_R g^2(x) dx < \infty \quad \text{or} \quad \int_R [g'(x)]^2 dx < \infty$$

另外, 虽然 Igel'nik 等给出了理论推导的概率空间作为权值  $\mathbf{a}_j$  和偏置  $b_j$  的选取依据, 但是如何确定概率空间上的多个关键参数依然是个比较难的问题. 在文献[28]中, Schmidt 等依据实验结果确定了隐含层随机参数的选择范围并使用了三组数据进行验证, 其明确和简洁的  $[-1, 1]$  区间确定还是为后

续 RVFLNs 理论研究与实际应用提供非常有益的指导.

## 2.2 鲁棒正则化 RVFLNs

上述常规 RVFLNs 采用最小二乘法计算网络输出权值, 当训练数据集中存在离群点时, 最小二乘法将“放大”离群点的影响, 使模型的整体特性发生很大改变. 为了减小数据中离群点对 RVFLNs 数据建模的影响, 拟在 RVFLNs 网络中引入 M 估计技术, 提高模型的鲁棒性. Huber 等于 1964 年提出的 M 估计技术是现在最常用的鲁棒估计方法<sup>[29-31]</sup>. 与使用误差平方和作为目标函数的最小二乘算法不同, M 估计对残差定义了新的目标损失函数, 使其能够对数据中的离群点进行区别对待. 提出的基于 M 估计的 RVFLNs (M-RVFLNs) 算法描述如下.

1) 单输出 M-RVFLNs: 当  $m = 1$  时,  $\beta$  和  $Y$  可以简化为:  $\beta = [\beta_1 \cdots \beta_L]^T_{L \times 1}$ ,  $Y = [\mathbf{y}_1 \cdots \mathbf{y}_L]^T_{M \times 1}$ . 此时, 式 (4) 的最小二乘解为:

$$\hat{\beta} = H^+ Y = (H^T H)^{-1} H^T Y \quad (6)$$

对应的目标函数为:

$$Q = \sum_{i=1}^M r_i^2 = \sum_{i=1}^M (y_i - H_i \beta)^2 \quad (7)$$

引入 M 估计, 式 (7) 关于残差  $r$  的目标函数变为:

$$Q = \sum_{i=1}^M \rho(r_i) = \sum_{i=1}^M \rho(y_i - H_i \beta) \quad (8)$$

式中,  $\rho$  为 M 估计影响函数.

回归系数  $\beta$  (即 RVFLNs 网络输出权值) 的求法如下:

$$\hat{\beta} = \arg \min_{\beta} \sum_{i=1}^M \rho(y_i - H_i \beta) = \arg \min_{\beta} \sum_{i=1}^M \rho(r_i(\beta)) \quad (9)$$

为保证估计结果具有尺度同变性, 引入残差的稳健尺度估计  $\hat{\sigma}$ , 则求解回归系数  $\beta$  形式变为:

$$\hat{\beta} = \arg \min_{\beta} \sum_{i=1}^M \rho \left( \frac{y_i - H_i \beta}{\hat{\sigma}} \right) = \arg \min_{\beta} \sum_{i=1}^M \rho \left( \frac{r_i(\beta)}{\hat{\sigma}} \right) \quad (10)$$

式 (10) 中的残差稳健尺度估计  $\hat{\sigma}$  对 M 估计很重要. 根据鲁棒估计理论,  $\hat{\sigma}$  取绝对离差中位数 (Median absolute deviation, MAD) 除以常数 0.6745. 数值 0.6745 是标准正态分布的 MAD, 当残差服从高斯

分布时, MAD 除以 0.6745 可以保持估计的一致性. 即:

$$\hat{\sigma} = \frac{\text{MAD}}{0.6745} = \frac{\text{median}_i |r_i - \text{median}(r_i)|}{0.6745} \quad (11)$$

令  $\partial \left( \sum_{i=1}^M \rho \left( \frac{r_i(\beta)}{\hat{\sigma}} \right) \right) / \partial \beta = 0$ , 可以得到:

$$\sum_{i=1}^M \varphi \left( \frac{r_i(\beta)}{\hat{\sigma}} \right) H_i^T := \sum_{i=1}^M \rho' \left( \frac{r_i(\beta)}{\hat{\sigma}} \right) H_i^T = 0 \quad (12)$$

定义加权因子

$$w(r_i) := \frac{\varphi(r_i)}{r_i} \quad (13)$$

那么等式 (12) 变为:

$$\sum_{i=1}^M w \left( \frac{r_i(\beta)}{\hat{\sigma}} \right) \times (y_i - H_i \beta) H_i^T = 0 \quad (14)$$

进一步简化为:

$$H^T W H \beta = H^T W Y \quad (15)$$

式中,  $W = \text{diag}\{w_i\}$ , 每一个  $w_i$  是等式 (13) 定义的权值因子, 约去等式两边  $H^T$ , 并进一步写为:

$$H' \beta = Y' \quad (16)$$

式中,  $H' = W H$ ,  $Y' = W Y$ .

为了防止 RVFLNs 模型过拟合, 稀疏输出权值矩阵, 进一步引入  $L_1$  和  $L_2$  两个正则化项, 此时优化目标函数变为:

$$\hat{\beta} = \arg \min_{\beta} \left\{ \left\| H' \beta - Y' \right\|_2^2 + \lambda \left( \frac{1 - \lambda_0}{2} \right) \|\beta\|_2^2 + \lambda_0 \|\beta\|_1 \right\} \quad (17)$$

式中,  $\lambda_0 \in [0, 1]$ ,  $\lambda \geq 0$ .

令  $\lambda_1 = \lambda \lambda_0$ ,  $\lambda_2 = \lambda(1 - \lambda_0)/2$ , 则式 (17) 变为:

$$\hat{\beta} = \arg \min_{\beta} \left\{ \left\| H' \beta - Y' \right\|_2^2 + \lambda_2 \|\beta\|_2^2 + \lambda_1 \|\beta\|_1 \right\} \quad (18)$$

进一步, 令  $(M + L) \times L$  维矩阵  $H^*_{(M+L) \times L} = (1 + \lambda_2)^{-\frac{1}{2}} \begin{bmatrix} H' \\ \sqrt{\lambda_2} I \end{bmatrix}$ ,  $Y^*_{(M+L)} = \begin{bmatrix} Y' \\ 0 \end{bmatrix}$ , 并使  $\gamma = \frac{\lambda_1}{\sqrt{1 + \lambda_2}}$ ,  $\beta^* = \sqrt{1 + \lambda_2} \beta$ . 则上式变为:

$$\hat{\beta} = \frac{1}{\sqrt{1 + \lambda_2}} \arg \min_{\beta^*} \left\{ \left\| H^* \beta^* - Y^* \right\|_2^2 + \gamma \|\beta^*\|_1 \right\} \quad (19)$$

式(19)乘上一个缩放因子  $1 + \lambda_2$ , 得到下式:

$$\hat{\beta} = \sqrt{1 + \lambda_2} \arg \min_{\beta^*} \left\{ \|H^* \beta^* - Y^*\|_2^2 + \gamma \|\beta^*\|_1 \right\} \quad (20)$$

上述式(20)即可由坐标下降 (Coordinate descent) 算法求解<sup>[32]</sup>.

**评论 1.** 为了解决 RVFLNs 模型固有的多重共线性和过拟合问题, 在鲁棒加权后的最小二乘损失函数基础上, 同时加入输出权值的  $L_2$  范数项和输出权值的  $L_1$  范数项作为惩罚, 构成式(17)所示优化目标函数的弹性网络 (Elastic net)<sup>[33]</sup>, 因而可有效稀疏化输出矩阵, 其效果比岭回归以及 Lasso 等任何单一范数作为正则化项的方法都要好<sup>[33]</sup>.

2) 多输出 M-RVFLNs: 当  $m > 1$  时, M-RVFLNs 为多输出, 此时  $\beta$  和  $Y$  表示如下:

$$\beta = [\beta_{j1}, \dots, \beta_{jh}, \dots, \beta_{jm}] = \begin{bmatrix} \beta_{11} & \cdots & \beta_{1m} \\ \vdots & \ddots & \vdots \\ \beta_{L1} & \cdots & \beta_{Lm} \end{bmatrix}_{L \times m}, \quad \begin{matrix} j = 1, \dots, L \\ h = 1, \dots, m \end{matrix} \quad (21)$$

$$Y = [y_{i1}, \dots, y_{ih}, \dots, y_{im}] = \begin{bmatrix} y_{11} & \cdots & y_{1m} \\ \vdots & \ddots & \vdots \\ y_{M1} & \cdots & y_{Mm} \end{bmatrix}_{M \times m}, \quad \begin{matrix} i = 1, \dots, M \\ h = 1, \dots, m \end{matrix} \quad (22)$$

权值矩阵的维数  $w(r_i(\beta)/\hat{\sigma})$  和  $Y$  一致, 即为如下  $M \times m$  维形式:

$$w = [w_{i1}, \dots, w_{ih}, \dots, w_{im}] = \begin{bmatrix} w_{11} & \cdots & w_{1m} \\ \vdots & \ddots & \vdots \\ w_{M1} & \cdots & w_{Mm} \end{bmatrix}_{M \times m} \quad (23)$$

此时, 多输出权值表示为:

$$\hat{\beta}^{(k+1)} = [\hat{\beta}_{j1}^{(k+1)}, \dots, \hat{\beta}_{jh}^{(k+1)}, \dots, \hat{\beta}_{jm}^{(k+1)}] \quad (24)$$

这里的每一个  $\hat{\beta}_{jh}^{(k+1)}$  由等式(20)求得.

### 2.3 M 估计 Gaussian 分布加权函数及参数选取

基于 M 估计的 RVFLNs 根据残差决定观测值对估计的贡献大小, 所以确定权重非常重要. 通过研究随机变量的分布特性和统计特性, 我们采用统计学中应用较为广泛的高斯 (Gaussian) 分布进行 M 估计权值确定. 设连续随机变量  $X$  服从一个位置参数为  $\mu$ 、尺度 (宽度) 参数为  $\sigma_1$  的概率分布, 且概率

密度函数 (Probability density function, PDF) 为:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{(x-\mu)^2}{2\sigma_1^2}\right)$$

则  $X$  服从高斯分布, 记作  $X \sim N(\mu, \sigma_1^2)$ . 如果参数  $\mu, \sigma_1$  已知, 就可以确定 Gaussian 分布的 PDF 曲线.

根据 Gaussian 分布的 PDF 特性, 在残差处于中间的样本点处取得最大权值, 在残差很小或者很大的样本处取得较小权值, 这样就能削弱离群点对建模带来的不良影响. Gaussian 分布函数参数  $\mu$  和  $\sigma_1$  决定了曲线的特性, 因此如何确定这两个参数是该方法的重要内容:

1)  $\mu$  值的确定: M 估计加权的目的是调和“离群点”和“过拟合”的作用, 因此位置参数  $\mu$  应取标准化残差  $r_i/\hat{\sigma}$  的中间值, 即  $\mu = \text{median}(r_i/\hat{\sigma}_i)$ .

2)  $\sigma_1$  值的确定: 参数  $\sigma_1$  决定了 Gaussian 分布函数曲线的形状, 即样本点越靠近  $\mu$ , 该点的权值也越大, 因此考虑误差的统计特性, 应遵循以下取值原则: 如果误差分布较紧密, 则  $\sigma_1$  取值应偏大; 如果误差分布较分散, 则  $\sigma_1$  取值应偏小. 标准差能反映出预测误差的离散程度, 误差分布越紧密时标准差越小, 误差分布越分散时标准差越大. 为此, 采用标准化残差标准差的倒数来确定宽度参数  $\sigma_1$  的取值, 即:

$$\sigma_1 = \frac{1}{\sqrt{\frac{\sum_{i=1}^M \left\| \frac{r_i/\hat{\sigma}_i - \sum_{i=1}^M (r_i/\hat{\sigma}_i)}{M} \right\|^2}{M}}}$$

### 2.4 鲁棒建模算法实现步骤

所提 M 估计鲁棒 RVFLNs 算法的主要建模过程如下:

- 1) 根据式(6)计算  $\hat{\beta}^{(0)}$ ;
- 2) 利用得到的  $\hat{\beta}^{(0)}$ , 计算残差向量  $\mathbf{r}^{(0)}$  并计算尺度估计值  $\hat{\sigma}$ , 进而求出标准化残差向量  $\mathbf{r}^{(0)}/\hat{\sigma}$ ;
- 3) 根据式(13)计算每个样本对应的权值, 求得权函数方阵  $W^{(0)}$ ;
- 4) 构造  $H'$  和  $Y'$ , 求解  $\beta^{(1)}$ , 并重复迭代计算得到  $\beta^{(2)}, \beta^{(3)}, \dots, \beta^{(k)}, \beta^{(k+1)}$ , 直到每个参数估计值  $|\hat{\beta}_{j,h}^{(k+1)} - \hat{\beta}_{j,h}^{(k)}|/\hat{\beta}_{j,h}^{(k)}, j = 0, 1, \dots, L, h = 1, \dots, m$  都小于给定收敛条件, 如  $10^{-6}$ .

## 3 工业数据测试

### 3.1 模型建立

采用我国华南某大型高炉的炉内数据与铁水质量数据 (采样间隔为 1h) 对所提方法进行数据测试,

根据该高炉炼铁工艺及相关仪器仪表设置, 确定影响 Si 含量、P 含量、S 含量、铁水温度等铁水质量指标的 18 个关键过程变量为: 冷风流量、送风比、热风压力、压差、顶压风量比、透气性、阻力系数、热风温度、富氧流量、富氧率、设定喷煤量、鼓风湿度、理论燃烧温度、标准风速、实际风速、鼓风动能、炉腹煤气量、炉腹煤气指数等. 考虑上述变量间具有很强的相关性, 并且过多的建模变量会加大建模复杂度, 影响模型性能, 因此首先采用典型相关分析 (CCA) 技术在原始高维的建模输入输出变量数据中提取出 4 对典型变量, 其相关系数分别为 0.601、0.551、0.462、0.232. 表 1 表明在 0.05 显著性水平下, 前 3 对典型变量显著相关, 第 4 对典型变量相关性不显著, 因而在后续分析中不予考虑. 表 2 给出了各个高炉本体参数典型变量的标准化系数, 将每一个参数乘以对应的典型相关系数进行加权, 并求其绝对值之和, 可以得到每个参数变量的权值如表 2 最后一列所示. 最后, 确定选取综合权值较大的炉腹煤气量  $u_1$  ( $\text{m}^3$ )、冷风流量  $u_2$  ( $\text{m}^3/\text{min}$ )、富氧流量  $u_3$  ( $\text{m}^3/\text{h}$ )、透气性  $u_4$  ( $\text{m}^3/\text{min} \cdot \text{kPa}$ )、富氧率  $u_5$  (wt%) 和理论燃烧温度  $u_6$  ( $^{\circ}\text{C}$ ) 作为影响 [Si]  $y_1$  (%), [S]  $y_2$  (%), [P]  $y_3$  (%) 和铁水温度 MIT  $y_4$  ( $^{\circ}\text{C}$ ) 等铁水质量指标的最主要因素, 并将这 6 个变

量组成新的样本集作为铁水质量建模输入变量.

为了更好地反映高炉非线性动态特性和输入输出变量的时序和时滞关系, 将当前时刻的输入  $X(t)$ 、上一时刻的输入  $X(t-1)$  和上一时刻的模型输出作为输入, 建立铁水质量的非线性自回归 (NARX) 模型. 同时, 为了消除低幅值高频的随机噪声, 取连续输入变量在一段时间的平均值作为实际模型输入变量. 在隐含层权值  $\mathbf{a}_j$  和偏差  $b_j$  随机范围选取方面, 首先在  $[-1, 1]$  范围内依据高斯分布随机产生  $\mathbf{a}_j$  和  $b_j$ [27]. 在此基础上, 通过增设倍数  $M$  并将  $M$  与权值范围  $[-1, 1]$  相乘以扩大  $\mathbf{a}_j$  与  $b_j$  的选择范围, 即选择范围扩大为  $[-M, M]$ ,  $M \in \mathbb{Z}^+$ . 由图 2 可以看出, 当  $\mathbf{a}_j$  和  $b_j$  的取值范围介于  $[-1, 1]$  时, 铁水质量指标测试均方根误差能够取得最小值, 因而  $\mathbf{a}_j$  与  $b_j$  的最佳选择范围可缩小至  $[-1, 1]$ . 由于 RVFLNs 的输入权值每次都是在一定范围随机产生, 导致每次运算结果具有随机性, 这里取 20 次运算结果的均值作为最终模型输出结果. 为确定铁水质量指标模型的最优正则化参数  $\lambda_0, \lambda$ , 同样取模 20 次平均测试均方根误差的最小值对应的  $\lambda_0, \lambda$  作为最后结果, 如图 3 所示, 结果为  $\lambda_0 = [0.0563, 0.0435, 0.0515, 0.0495]$  和  $\lambda = [3.475\text{E-}04, 1.8\text{E-}04, 2.675\text{E-}04, 2.475\text{E-}04]$ .

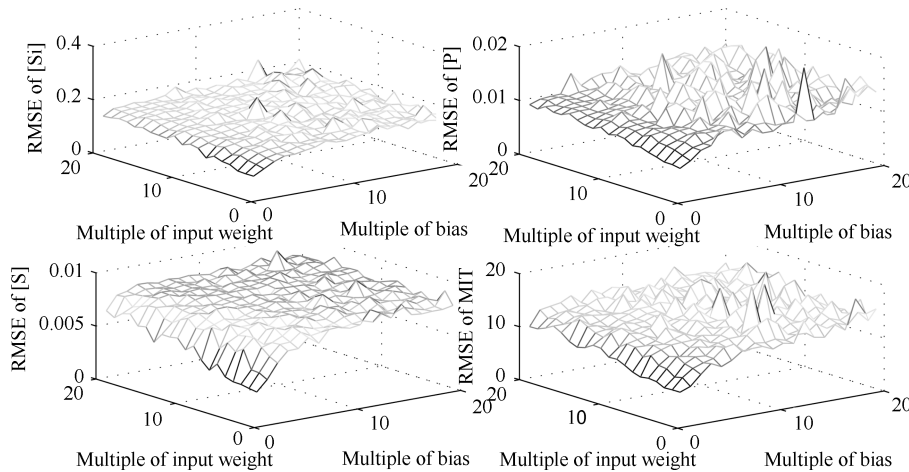


图 2 建模误差 RMSE 与输入权值倍数和输入偏置倍数之间的关系

Fig. 2 The relationship between the modeling RMSE and the input weight multiple and the input bias multiple

表 1 典型相关系数的显著性检验

Table 1 Significance test of canonical correlation coefficient

典型变量	显著性检验指标			
	Wilk's	Chi-SQ	DF	Sig.
1	0.337	299.701	72	0
2	0.527	176.485	51	0
3	0.754	77.786	32	0
4	0.958	11.678	15	0.703

表 2 高炉本体参数典型变量的标准化系数  
Table 2 Standardized canonical coefficients of BF body variables

影响变量	典型变量				变量权值
	1	2	3	4	
冷风流量	14.821	2.609	-3.502	-24.047	11.95769
送风比	-0.669	1.803	0.658	3.633	1.695912
热风压力	0.724	2.664	-2.139	-0.209	2.885878
压差	2.749	1.384	-0.527	2.382	2.655439
顶压富氧率	-0.06	-0.712	0.95	-0.022	0.865848
透气性	5.292	5.441	-6.701	10.201	9.263463
阻力系数	0.801	0.268	-4.063	8.006	2.505639
热风温度	0.587	-0.469	-1.356	-0.268	1.23674
富氧流量	11.697	-4.429	0.07	1.748	9.493758
富氧率	-5.751	3.556	-4.229	-4.556	7.362393
设定喷煤量	-0.931	3.284	4.027	-0.233	4.222921
鼓风湿度	0.533	0.805	1.932	-0.465	1.654862
理论燃烧温度	-3.408	2.774	5.055	1.685	5.906544
标准风速	-2.222	-0.705	2.385	6.454	2.824337
实际风速	-0.224	0.023	0.55	-0.767	0.401351
鼓风动能	-1.85	-1.067	-0.255	-0.002	1.815443
炉腹煤气量	-14.106	-6.874	0.208	15.053	12.34763
炉腹煤气指数	0.292	0.651	-0.006	-0.198	0.535663

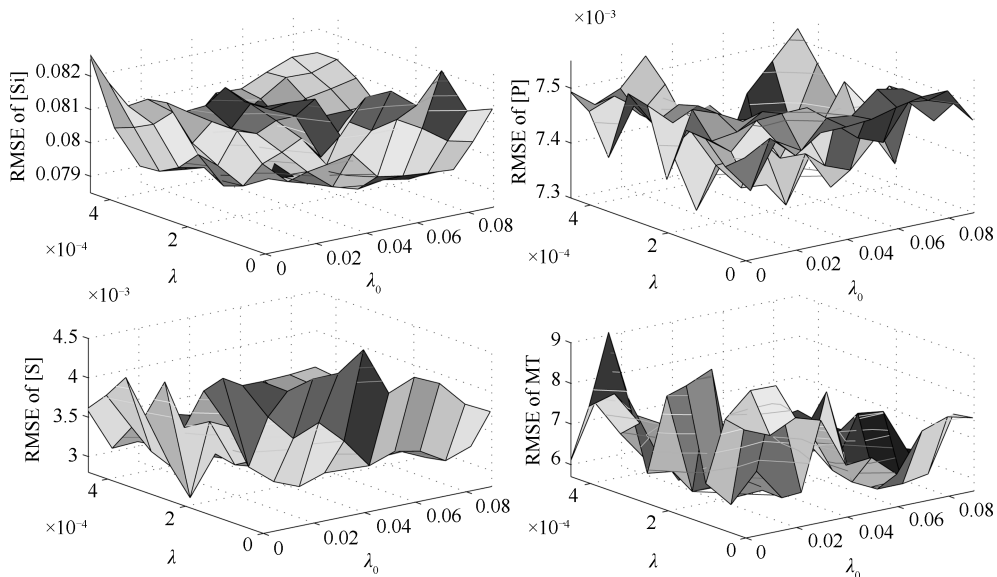


图 3 建模误差 RMSE 与正则化系数之间的关系曲线

Fig. 3 The relation curve between the modeling RMSE and the regularization coefficient

### 3.2 建模及其估计效果

数据中所含离群点比例以及离群点偏离正常取值范围的大小(即离群点幅值)对模型的精度有重要影响. 为了更全面地验证所提方法的鲁棒性, 基于实际高炉炼铁数据进行鲁棒测试的数据设计. 共设计

两组数据, 分别从不同离群点比例时数据建模效果以及不同离群幅值时数据建模效果两个角度来评价模型的建模与鲁棒性能.

第一组数据集用来测试所提方法对不同离群点比例时数据建模的适用性. 首先, 从原始数据中随机选取比例为 0%, 5%, 10%, ...,



50 % 的样本点  $y_{i,Outlier}$  并对其进行处理, 令  $y_{i,Outlier} = y_i + 2 \times sign \times [rand(0, 1) \times y_{max\ min}]$ , 式中  $y_{max\ min} = \max(y_i) - \min(y_i)$ ,  $i = 1, 2, 3, 4$  表示正常炉况下铁水质量参数波动范围内的最大值与最小值之差. 为了使离群点更加不均衡, 在所选择样本点中正向离群点与负向离群点的比例设置为 2:1, 添加正向离群点时令  $sign = 1$ , 添加负向离群点时令  $sign = -1$ .

第二组数据集用来测试所提方法对相同离群

点比例但离群点偏离幅值不同时数据建模适用性. 从原始数据中随机选取 20 % 的样本点, 依据  $y_{i,Outlier} = y_i + sign \times a \times [rand(0, 1) \times y_{max\ min}]$  添加离群点, 式中  $y_{max\ min} = \max(y_i) - \min(y_i)$ ,  $a = 0, 0.5, 1, 1.5, \dots, 5$  用来调节幅值大小. 同样为了使离群点更加不均衡, 在所选择样本点中正向离群点与负向离群点的比例设置为 2:1, 添加正向离群点时令  $sign = 1$ , 添加负向离群点时令  $sign = -1$ .

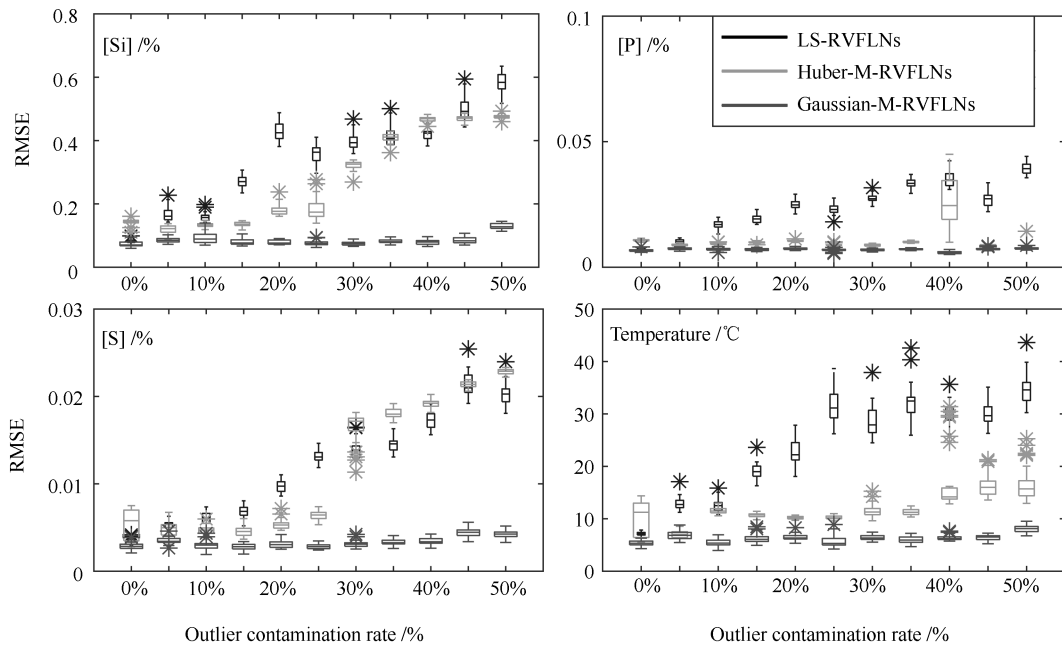


图 4 不同离群点比例时铁水质量估计 RMSE 箱形图

Fig. 4 The box diagram of the Estimation RMSE of MIQ indices with different methods at each of the outlier contamination rates

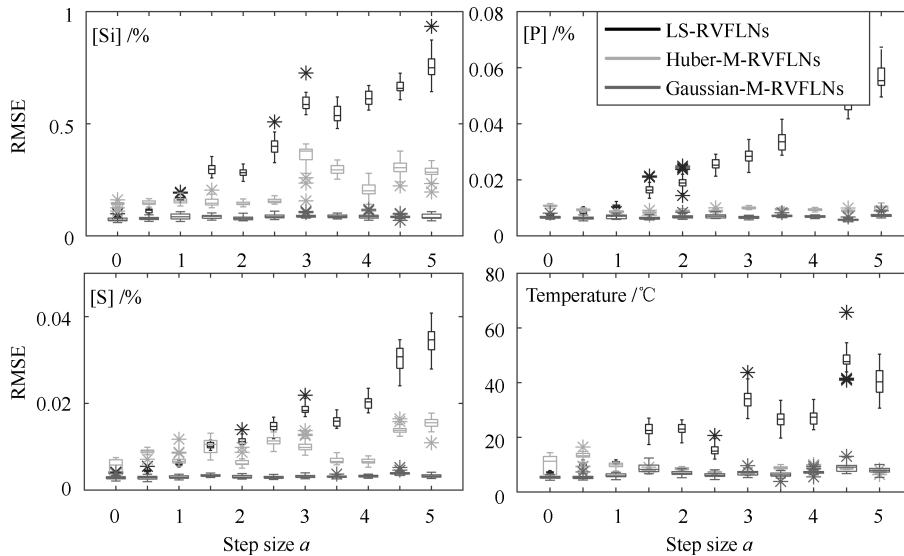


图 5 不同离群幅值时铁水质量估计 RMSE 箱形图

Fig. 5 The box diagram of the Estimation RMSE of MIQ indices with different methods when the amplitudes of the outliers are increased

为了验证所提基于 Gaussian 分布加权 M-估计的鲁棒正则化 RVFLNs (Gaussian-M-RVFLNs) 方法对多元铁水质量指标 (即 [Si]、[P]、[S] 和 MIT) 的建模效果, 将其与基于最小二乘加权的常规 RVFLNs (LS-RVFLNs) 和基于 Huber 加权 M-估计的鲁棒随机神经网络 (Huber-M-RVFLNs) 进行比较, 如图 4 和图 5 所示. 其中图 4 为不同离群点比例时三种方法的多元铁水质量估计均方根误差箱形图, 图 5 为不同离群幅值时三种方法的多元铁水质量估计均方根误差箱形图. 三种方法的激活函数均使用 Sigmoid 函数, 隐层节点数均设置为 30, 权值  $a_j$  和偏差  $b_j$  取值范围均确定为  $[-1, 1]$ . Huber 加权算法中 K 通常取为 1.345, 这样取值的估计方法既是稳健的, 又有较高的估计效率. 另外, 为了避免不同数量级数据间的相互影响, 对所有数据进行归一化处理. 由于随机神经网络的  $a_j$  和  $b_j$  为随机生成, 每次结果不具有唯一性, 为更好地衡量模型的性能, 对每一个数据集分别进行 30 次实验, 通过 30 次数据实验的均方根误差 (Root mean square error, RMSE) 来比较不同方法的鲁棒性能.

从图 4 和图 5 所示的箱型图可以看出, 当没有离群点或者离群点比例和离群幅值较小时, 所提 Gaussian-M-RVFLNs 方法与常规 LS-RVFLNs 的铁水质量估计效果基本相当, 但要好于 Huber-M-RVFLNs 方法. 随着离群点比例以及离群幅值的逐渐增大, 基于传统 LS-RVFLNs 的铁水质量模型估计

精度下降明显, 而所提 Gaussian-M-RVFLNs 建立的铁水质量模型能够始终保持很高的估计精度, 并且远好于对比的其他两种方法. Huber-M-RVFLNs 方法虽然在离群点比例较低和离群幅值较小时可保持较高精度, 具有一定的鲁棒性, 但是当离群点比例较低和离群幅值持续增加时, 其估计性能下降明显, 这在图 4 所示高离群点比例情况时尤其明显.

一个好的鲁棒模型要求在实际模型与理想分布模型差别微小时, 受离群点的影响较小, 接近正确估值, 更重要的是要求实际模型与理想分布模型差别较大时, 估计值也不会受大量离群点的破坏性影响, 依然能够得到接近正常模式下的正确估计. 为此考虑离群比例为 20%、离群幅值步长  $a$  为 2 的多元铁水质量建模与估计效果, 实验结果分别如图 6 和图 7 所示. 为了说明问题, 同样将所提 Gaussian-M-RVFLNs 方法与 LS-RVFLNs 和 Huber-M-RVFLNs 进行比较分析. 可以看出在存在较高离群点比例和较大离群幅值情况下, 所提方法的建模与估计性能最好, 能够根据实时输入数据, 对难测多元铁水质量指标进行准确估计, 且估计趋势基本与实际值一致. 图 8 为不同方法下铁水质量指标测试误差概率密度分布曲线, 可以看出所提 Gaussian-M-RVFLNs 模型和其他两种方法建立的模型相比, 其误差概率密度分布曲线基本为又瘦又高的围绕 0 均值的高斯分布曲线, 即大部分估计误差仅为实际工业数据中所包含的测量噪声.

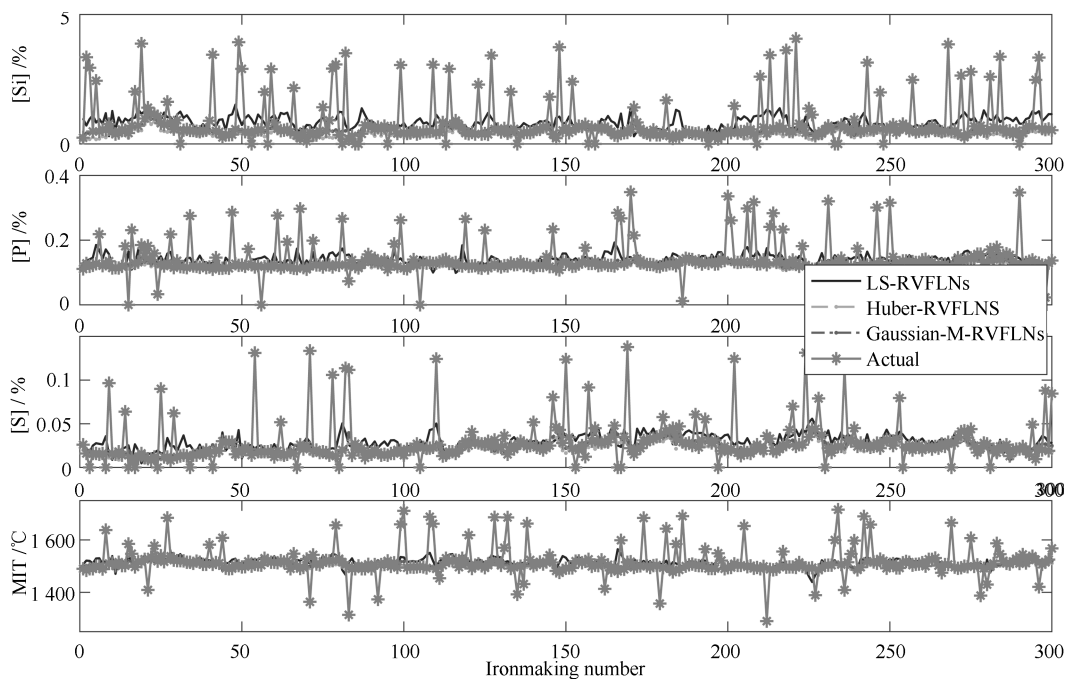


图 6 离群比例为 20% 和离群幅值步长  $a$  为 2 时, 不同 RVFLNs 建模方法铁水质量指标建模效果

Fig. 6 Modeling results of MIQ indices with different methods when outlier contamination rate is twenty percent and amplitude is two

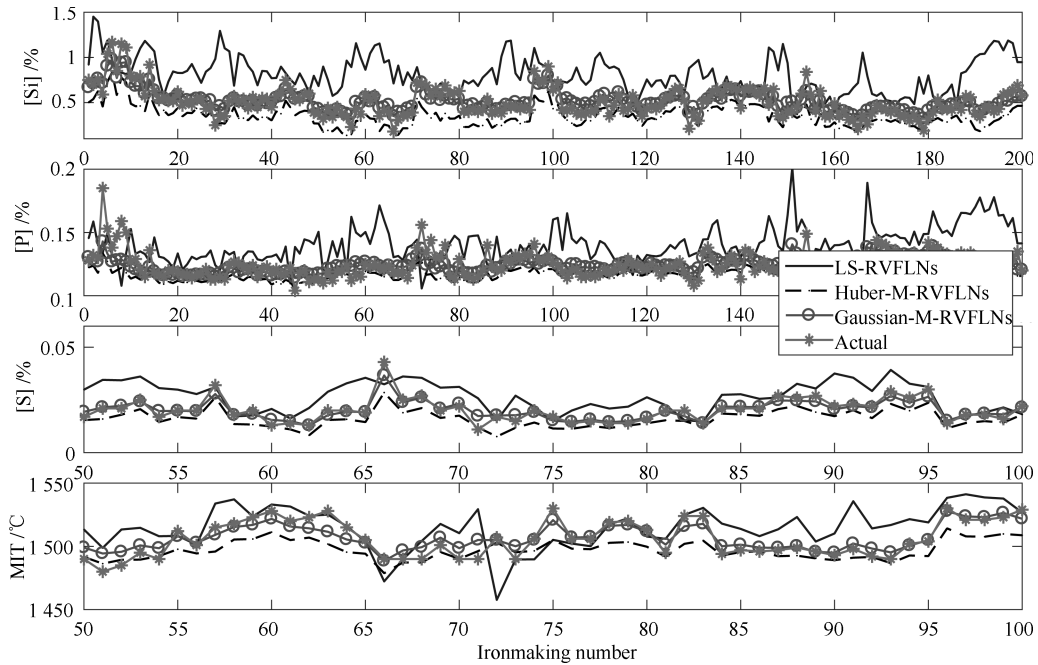


图 7 离群比例为 20% 和离群幅值步长  $a$  为 2 时, 不同 RVFLNs 建模方法铁水质量指标估计效果

Fig. 7 Estimation results of MIQ indices with different methods when outlier contamination rates is twenty percent and amplitude is two

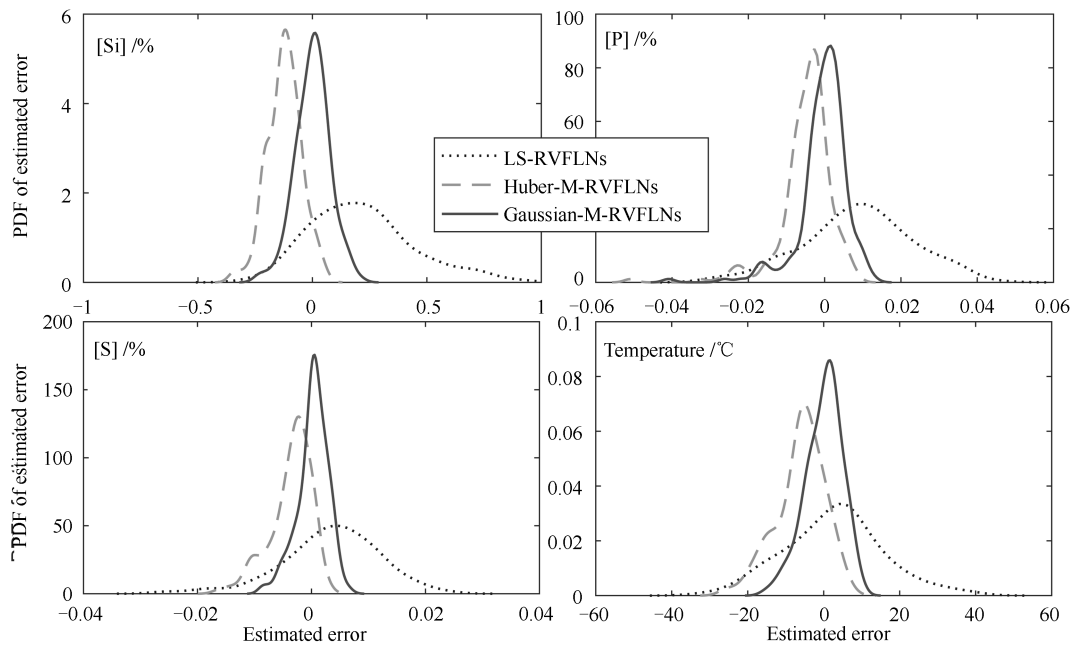


图 8 不同 RVFLNs 建模方法铁水质量估计误差 PDF 曲线

Fig. 8 PDF curve of MIQ estimation error with different methods

图 9 为不同离群点比例时三种不同方法铁水质量建模时 RVFLNs 网络中为 0 的输出权值数量曲线图. 可以看出, 所提 Gaussian-M-RVFLNs 方法建立的铁水质量模型在 4 个铁水质量指标输出权值矩阵中权值为 0 的个数最多, Huber-M-RVFLNs 网

络模型次之, LS-RVFLNs 网络模型最少且输出权值一直不为 0. 这说明, 所提方法由于在鲁棒建模基础上引入  $L_1$  和  $L_2$  两个正则化项范数, 可稀疏化输出权值矩阵, 提高模型泛化性能, 有效避免模型过拟合.

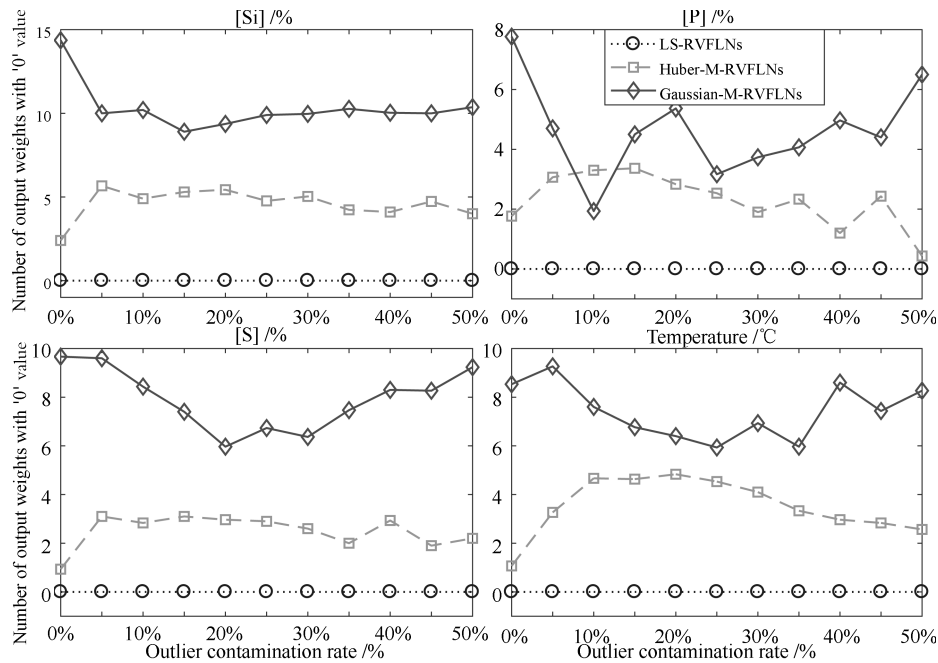


图9 不同RVFLNs建模方法铁水质量估计时网络输出权值为0的数量曲线

Fig. 9 The curve of the number of output weights with '0' value of MIQ estimation error with different methods

## 4 结论

本文针对高炉炼铁过程多元铁水质量指标难以在线检测,以及采用常规方法难以进行有效建模与在线估计的难题,基于 Gaussian 分布加权 M-估计技术,提出一种新型的数据驱动随机神经网络(RVFLNs)鲁棒建模方法,用于实现多元铁水质量指标的鲁棒估计。基于实际工业数据的研究表明:和常规最小二乘加权 RVFLNs 算法以及经典 Huber 加权 M 估计的 RVFLNs 算法相比,所提 Gaussian 分布加权 M-估计的鲁棒正则化 RVFLNs 算法可有效克服数据中不同离群点比例和不同离群幅值给数据建模带来的不利影响,大大提高 RVFLNs 数据建模的鲁棒性能,使得模型在复杂动态环境中持续有效,并且估计精度较高。同时,在鲁棒建模基础上,在优化性能指标引入两个正则化项,以最大化稀疏化 RVFLNs 网络的输出权值矩阵,这不仅可显著降低模型复杂度,还可有效避免模型过拟合,进一步提高模型的适应性。

## References

- Song He-Da, Zhou Ping, Wang Hong, Chai Tian-You. Non-linear subspace modeling of multivariate molten iron quality in blast furnace ironmaking and its application. *Acta Automatica Sinica*, 2016, **42**(11): 1664–1679  
(宋贺达, 周平, 王宏, 柴天佑. 高炉炼铁过程多元铁水质量非线性子空间建模及应用. *自动化学报*, 2016, **42**(11): 1664–1679)
- Zhou P, Guo D W, Wang H, Chai T Y. Data-driven robust M-LS-SVR-based NARX modeling for estimation and control of molten iron quality indices in blast furnace ironmaking. *IEEE Transactions on Neural Networks and Learning Systems*, 2018, **29**(9): 4007–4021
- Gao C H, Jian L, Luo S H. Modeling of the thermal state change of blast furnace hearth with support vector machines. *IEEE Transactions on Industrial Electronics*, 2012, **59**(2): 1134–1145
- Kuang S B, Li Z Y, Yan D L, Qi Y H, Yu A B. Numerical study of hot charge operation in ironmaking blast furnace. *Minerals Engineering*, 2014, **63**: 45–56
- Zhou Ping, Zhang Li, Li Wen-Peng, Dai Peng, Chai Tian-You. Autoencoder and PCA based RVFLNs modeling for multivariate molten iron quality in blast furnace ironmaking. *Acta Automatica Sinica*, 2018, **44**(10): 1799–1811  
(周平, 张丽, 李温鹏, 戴鹏, 柴天佑. 集成自编码与 PCA 的高炉多元铁水质量随机神经网络建模. *自动化学报*, 2018, **44**(10): 1799–1811)
- Jian L, Gao C H, Xia Z H. Constructing multiple kernel learning framework for blast furnace automation. *IEEE Transactions on Automation Science and Engineering*, 2012, **9**(4): 763–777
- Gao C H, Ge Q H, Jian L. Rule extraction from fuzzy-based blast furnace SVM multiclassifier for decision-making. *IEEE Transactions on Fuzzy Systems*, 2014, **22**(3): 586–596
- Jiang Zhao-Hui, Dong Meng-Lin, Gui Wei-Hua, Yang Chun-Hua, Xie Yong-Fang. Two-dimensional prediction for silicon content of hot metal of blast furnace based on Bootstrap. *Acta Automatica Sinica*, 2016, **42**(5): 715–723  
(蒋朝辉, 董梦林, 桂卫华, 阳春华, 谢永芳. 基于 Bootstrap 的高炉铁水硅含量二维预报. *自动化学报*, 2016, **42**(5): 715–723)
- de Castro J A, Nogami H, Yagi J I. Three-dimensional multiphase mathematical modeling of the blast furnace based on the multifluid model. *ISIJ International*, 2002, **42**(1): 44–52

- 10 de Castro J A, Nogami H, Yagi J I. Transient mathematical model of blast furnace based on multi-fluid concept, with application to high PCI operation. *ISIJ International*, 2000, **40**(7): 637–646
- 11 Chu Man-Sheng, Wang Hong-Tao, Liu Zheng-Gen, Tang Jue. Research progress on mathematical modeling of blast furnace ironmaking process. *Iron and Steel*, 2014, **49**(11): 1–8  
(储满生, 王宏涛, 柳政根, 唐珏. 高炉炼铁过程数学模拟的研究进展. *钢铁*, 2014, **49**(11): 1–8)
- 12 Lvanov E B, Klimovitskii M D, Anisimov E F. Expert system for blast-furnace operators. *Metallurgist*, 2011, **54**(11–12): 730–736
- 13 Liu J K, Wang S Q. Construction of the inference engine of blast furnace expert system. *Journal of Iron and Steel Research, International*, 1998, **5**(2): 22–27
- 14 Shi L, Li Z L, Li J P. Model of hot metal silicon content in blast furnace based on principal component analysis application and partial least square. *Journal of Iron and Steel Research, International*, 2011, **18**(10): 13–16
- 15 Saxén H, Gao C H, Gao Z W. Data-driven time discrete models for dynamic prediction of the hot metal silicon content in the blast furnace — a review. *IEEE Transactions on Industrial Informatics*, 2013, **9**(4): 2213–2225
- 16 Yuan M, Zhou P, Li M L, Li R F, Wang H, Chai T Y. Intelligent multivariable modeling of blast furnace molten iron quality based on dynamic AGA-ANN and PCA. *Journal of Iron and Steel Research, International*, 2015, **22**(6): 487–495
- 17 Wang Wei, Chen Wei-Lin, Ye Yong, Xu Zhi-Hui, Jia Bin. Application of neural network to predict sulphur content in hot metal. *Iron and Steel*, 2006, **41**(10): 19–22  
(王伟, 陈畏林, 叶勇, 徐智慧, 贾斌. 神经网络在高炉铁水硫含量预报中的应用. *钢铁*, 2006, **41**(10): 19–22)
- 18 Yan W W, Tang D, Lin Y J. A data-driven soft sensor modeling method based on deep learning and its application. *IEEE Transactions on Industrial Electronics*, 2017, **64**(5): 4237–4245
- 19 Ma J M, Jiang H C, Huang K Z, Bi Z Q, Man K L. Novel field-support vector regression-based soft sensor for accurate estimation of solar irradiance. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2017, **64**(12): 3183–3191
- 20 Zhou P, Yuan M, Wang H, Chai T Y. Data-driven dynamic modeling for prediction of molten iron silicon content using ELM with self-feedback. *Mathematical Problems in Engineering*, 2015, **2015**: Article ID 326160
- 21 Zhou P, Yuan M, Wang H, Wang Z, Chai T Y. Multivariable dynamic modeling for molten iron quality using online sequential random vector functional-link networks with self-feedback connections. *Information Sciences*, 2015, **325**: 237–255
- 22 Sargin M E, Yemez Y, Erzin E, Tekalp A M. Audiovisual synchronization and fusion using canonical correlation analysis. *IEEE Transactions on Multimedia*, 2007, **9**(7): 1396–1403
- 23 Pao Y H, Takefuji Y. Functional-link net computing: theory, system architecture, and functionalities. *Computer*, 1992, **25**(5): 76–79
- 24 Igel'nik B, Pao Y H. Stochastic choice of basis functions in adaptive function approximation and the functional-link net. *IEEE Transactions on Neural Networks*, 1995, **6**(6): 1320–1329
- 25 Pao Y H, Park G H, Sobajic D J. Learning and generalization characteristics of the random vector functional-link net. *Neurocomputing*, 1994, **6**(2): 163–180
- 26 Scardapane S, Wang D H, Panella M, Uncini A. Distributed learning for random vector functional-link networks. *Information Sciences*, 2015, **301**: 271–284
- 27 Zhang L, Suganthan P N. A comprehensive evaluation of random vector functional link networks. *Information Sciences*, 2016, **367–368**: 1094–1105
- 28 Schmidt W F, Kraaijveld M A, Duin R P W. Feedforward neural networks with random weights. In: *Proceedings of 1992 Pattern Recognition Conference B: Pattern Recognition Methodology and Systems*. The Hague: IEEE, 1992. 1–4
- 29 Huber P J, Ronchetti E M. *Robust Statistics* (Second Edition). Hoboken, New Jersey: John Wiley & Sons, 2009.
- 30 Valdora M, Yohai V J. Robust estimators for generalized linear models. *Journal of Statistical Planning and Inference*, 2014, **146**: 31–48
- 31 Fan J, Yan A L, Xiu N H. Asymptotic properties for M-estimators in linear models with dependent random errors. *Journal of Statistical Planning and Inference*, 2014, **148**: 49–66
- 32 Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 2010, **33**(1): 1–22
- 33 Zou H, Hastie T. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2005, **67**(2): 301–320



李温鹏 东北大学硕士研究生. 2016 年获得烟台大学学士学位. 主要研究方向为数据驱动建模与控制, 机器学习算法. E-mail: weepenli@163.com

(LI Wen-Peng Master student at Northeastern University. He received his bachelor degree from YanTai University in 2016. His research interest covers data-driven modeling and control, and machine learning algorithm.)



周平 东北大学教授. 分别于 2003 年, 2006 年, 2013 年获得东北大学学士学位、硕士学位和博士学位. 主要研究方向为工业过程运行反馈控制, 数据驱动建模与控制. 本文通信作者. E-mail: zhouping@mail.neu.edu.cn

(ZHOU Ping Professor at Northeastern University. He received his bachelor degree, master degree, and Ph.D. degree from Northeastern University in 2003, 2006, and 2013, respectively. His research interest covers operation feedback control of industrial process, data-driven modeling and control. Corresponding author of this paper.)