工程科学学报，第 38 卷，第 9 期：1233–1241，2016 年 9 月
Chinese Journal of Engineering，Vol. 38，No. 9：1233–1241，September 2016
DOI：10.13374/j. issn2095–9389.2016.09.006；http：//journals. ustb. edu. cn

基于稀疏化鲁棒 LS–SVR 与多目标优化的铁水硅含量软测量建模

郭东伟[1,2]，周　平[1,2] ✉

1）东北大学流程工业综合自动化国家重点实验室，沈阳 110819　2）矿冶过程自动控制技术国家重点实验室，北京 102628
✉ 通信作者，E-mail：zhouping@ mail. neu. edu. cn

摘　要　针对高炉炼铁过程的关键工艺指标——铁水硅含量［Si］难以直接在线检测且化验过程滞后的问题，提出一种基于稀疏化鲁棒最小二乘支持向量机（R–S–LS–SVR）与多目标遗传参数优化的铁水［Si］动态软测量建模方法．首先，针对标准最小二乘支持向量机（LS–SVR）的拉格朗日乘子与误差项成正比导致最终解缺少稀疏性的问题，提取样本数据在特征空间映射集的极大无关组来实现训练样本集的稀疏化，降低建模的计算复杂度；其次，标准最小二乘支持向量机的目标函数鲁棒性不足的问题将 IGGIII 加权函数引入稀疏化后的最小二乘支持向量机模型进行鲁棒性改进，得到鲁棒性较强的稀疏化鲁棒最小二乘支持向量机模型；最后，针对常规均方根误差评价模型性能的不足，提出从建模误差与估计趋势评价建模性能的多目标评价指标．在此基础上，利用非支配排序的带有精英策略的多目标遗传算法优化模型参数，从而获得具有最优参数的铁水［Si］在线软测量模型．工业实验及比较分析验证了所提方法的有效性和先进性．

关键词　炼铁；硅含量；建模；最小二乘法；支持向量机；多目标优化

分类号　TP18

Soft-sensor modeling of silicon content in hot metal based on sparse robust LS–SVR and multi-objective optimization

GUO Dong-wei[1,2]，*ZHOU Ping*[1,2] ✉

1）State Key Laboratory of Synthetical Automation for Process Industries，Northeastern University，Shenyang 110819，China
2）State Key Laboratory of Process Automation in Mining & Metallurgy，Beijing 102628，China
✉ Corresponding author，E-mail：zhouping@ mail. neu. edu. cn

ABSTRACT　To solve the problem that the parameter of silicon content（［Si］）in hot mental is difficult to be directly detected and obtained by manual analysis with large time delay，a method of sparse and robust least squares support vector regression（R–S–LS–SVR）was proposed to establish a dynamic model of ［Si］ with the help of the multi-objective genetic optimization of model parameters. First，owing to the issue that the Lagrange multiplier of the standard least squares support vector machine（LS–SVR）is directly proportional to the error term and solves the lack of sparsity，the maximal independent set of sample data in the feature space mapping set was extracted to realize the sparse of the training sample set and reduce the computational complexity of modeling. Next，in view of the problem that the standard least squares support vector machine has no regularization term，a method to improve the modeling robustness was proposed by introducing the IGGIII weighting function into the obtained sparse least squares support vector regression（S–LS–SVR）model. Last，the multi-objective evaluation index that synthesizes the modeling residue and the estimated trend was

收稿日期：2015–11–20
基金项目：国家自然科学基金资助项目（61473064；61290323；61333007）；中央高校基本科研业务费资助项目（N130108001）；辽宁省教育厅科技基金资助项目（L20150186）

presented to compensate for the deficiency of the single root mean square error (RMSE) index. Based on those, an on-line soft sensor model of hot metal [Si] with the optimal parameters was obtained by using the multi-objective genetic algorithm (NSGA-II) with the non-dominated sort and elitist strategy. Industrial verification and analysis show the effectiveness and superiority of the proposed method.
KEY WORDS　ironmaking; silicon content; modeling; least squares methods; support vector machines; multi-objective optimization

高炉炼铁是钢铁生产中的重要单元，其作用就是将固态的铁矿石通过复杂的高温、高压等物理化学变化和多相多场耦合效应，在焦炭、煤气等作用下还原成液态的铁水．一个典型的高炉炼铁系统主要由高炉本体、上料系统、送风系统、高炉烟气净化系统、渣铁处理系统和喷吹燃料系统等组成[1]．由于炼铁高炉内部在高温、高压的条件下进行着复杂的气−固、气−液、固−固、固−液等多相转换，众多变量和参数之间错综复杂和相互耦合，被公认为是最复杂的逆流反应器．一直以来，高炉炼铁过程的建模、控制与优化是维持高炉稳定顺行、高产、优质和低耗的重要手段[1]．目前，由于测量手段的限制，反映铁水质量和高炉热状态的关键指标——硅含量([Si])的测量一般仍采用人工定期抽样化验的方式进行，由于人工检测的滞后性和不精确性，使得基于[Si]质量参数的高炉炉况判断和高炉操作异常困难．因此，实现优质、低耗的高炉运行优化与控制就必须实现铁水[Si]质量参数的在线估计与软测量．

高炉炼铁过程的上述复杂动态特性，使得用于铁水[Si]在线估计的机理模型不易建立．对于难以机理建模的复杂工业过程，数据驱动的智能建模和统计建模受到越来越广泛的关注[2]．Wang 等[3]提出了基于混沌时间序列的铁水[Si]支持向量回归模型(SVR)，在 ±0.1 误差内的命中率大于 88%，有较好的模型预测能力; Saxén 等[4]建立了可实现周期性铁水[Si]预测的离散时间序列模型，将模型的线性部分与非线性部分分别处理，提高了模型的预测能力; 唐贤伦等[5]建立了基于混沌粒子群优化的铁水硅含量 SVR 预报模型，实现预测绝对误差小于 0.03 的样本数占总测试样本数 90% 以上的效果．在众多数据驱动建模方法中，SVR 凭借其在解决小样本、非线性及高维模式识别问题中的优势，得到广泛的应用[6-10]．SVR 有严格的数学基础和稳定的学习机制，能在很大程度上克服“维数灾难”和“过学习”问题[8-10]．现有的针对铁水[Si]的 SVR 建模大多只是静态模型，忽略了高炉炼铁过程的动态特性和输入输出时序和时滞关系[8-9]．另外，由于采集于实际高炉工业过程的样本数据往往包含较大噪音，这样建立的 SVR 模型鲁棒性较差，对噪音很敏感．为此，本文基于多目标参数优化技术，提出一种具有稀疏性和鲁棒性的改进最小二乘支持向量回归(LS-SVR)建模技术，用于对铁水[Si]进行动态软测量建模．首先，以 LS-SVR 为基础，通过提取样本数据在特征空间映射集的极大无关组来实现训练样本集的稀疏化，降低建模的计算复杂度[6,8]；然后，将 IGGIII 权函数引入稀疏化后的 S-LS-SVR 模型，得到鲁棒性较好的 R-S-LS-SVR 模型; 最后，针对常规均方根误差评价建模性能的不足，提出从建模误差与估计趋势综合评价建模性能的多目标评价指标．在此基础上，利用非支配排序的带有精英策略的多目标遗传算法(NSGA-II)优化 R-S-LS-SVR 模型结构参数，从而获得最优参数的铁水[Si]在线估计模型．

1　建模问题描述

从提高产品质量和节约能源的角度而言，高炉系统的控制与优化的主要对象是铁水硅含量([Si])，它也是衡量高炉内热状态的重要标志，[Si]过高或过低对于铁水质量、燃料消耗和生产成本有较大的影响[7]．为此，对铁水[Si]进行在线估计或软测量意义重大．

进行[Si]软测量的第 1 步应该分析影响铁水[Si]的主要因素．综合考虑分析高炉送风系统、燃料喷吹系统的现有传感器和炉缸、炉腹的内部可测量参数，进而确定与铁水[Si]密切相关的影响因子有鼓风湿度(g·m^{-3})、热风压力(kPa)、炉腹煤气量(m^3·min^{-1})、设定喷煤量(t·d^{-1})、富氧率(质量分数)、热风温度(℃)、富氧流量(m^3·h^{-1})、炉顶压力(kPa)、气体渗透率(m^3·min^{-1}·kPa)、实际风速(m·s^{-1})、冷风流量(m^3·h^{-1})、理论燃烧温度(℃)、炉腹煤气指数(m·min^{-1})、注煤量(kg·h^{-1})、鼓风动能(kJ·s^{-1})、阻力系数、进料比(质量分数)等．由于影响高炉铁水[Si]的因素众多，并且各变量对于[Si]的影响相差较大，若将其全部引入模型，一方面由于变量维数过高导致训练时间延长和实时性差; 另一方面过多的变量将引入较多的干扰，影响建模精度．因此，降维处理是保证建模效率与精度，提高模型泛化能力的必要过程．本文采用数据分析中常用的主成分分析法(PCA)对变量进行降维，在降低数据样本维度的同时也保持方差贡献最大的特征变量．

通过主成分分析，最终得到影响铁水硅含量[Si](C_{Si})相关性最大的六个变量为热风压力 u_1(kPa)、热风温度 u_2(℃)、富氧率 u_3、设定喷煤量 u_4(t·d^{-1})、鼓风湿度 u_5(g·m^{-3})和炉腹煤气量 u_6(m·min^{-1})．同时，

为了更好地反映高炉的非线性动态特性，将相关输入输出变量的时序关系在建模过程进行考虑. 为此，将之前采样时刻的测量值 $\boldsymbol{U}(t-1)=[u_1(t-1),u_2(t-1),u_3(t-1),u_4(t-1),u_5(t-1),u_6(t-1)]$，$\boldsymbol{U}(t-2),\cdots,\boldsymbol{U}(t-\tau_U)$，$\tau_U\in{}^+$ 以及之前采样时刻铁水[Si]值 $C_{\mathrm{Si}}(t-1),C_{\mathrm{Si}}(t-2),\cdots,C_{\mathrm{Si}}(t-\tau_C)$，$\tau_C\in{}^+$，连同当前采样时刻的检测值 $\boldsymbol{U}(t)$ 作为动态模型的综合输入，即建立的动态软测量模型用于实现如下的非线性动态映射关系：

$$C_{\mathrm{Si}}(t)=f_{\mathrm{R-S-LS-SVR}}\{\boldsymbol{U}(t),\boldsymbol{U}(t-1),\cdots,\boldsymbol{U}(t-\tau_U),C_{\mathrm{Si}}(t-1),\cdots,C_{\mathrm{Si}}(t-\tau_C)\}. \tag{1}$$

式(1)表示需要建立的自回归滑动平均(ARMA)动态软测量模型结构. 本文基于多目标遗传参数优化技术，采用提出的稀疏化鲁棒 LS—SVR 建模算法实现式(1)的自回归滑动平均动态模型，相关算法将在下节中给出.

2 R—S—LS—SVR 建模算法

标准 LS—SVR 通过引入等式约束替代经典 SVR 的不等式约束，虽然使计算的复杂度大大降低，但同时也引入了两个潜在的问题：其一是目标函数中没有正则化项，对于数据中含有噪声、离群点以及误差不服从正态分布的情况导致缺乏鲁棒性；其二是由于支持向量所对应的拉格朗日乘子与误差项成正比导致最终的解缺少稀疏性. 针对 LS—SVR 的上述问题，提出同时兼顾鲁棒性和稀疏性的改进 LS—SVR 算法，即 R—S—LS—SVR.

2.1 标准 LS—SVR 算法

假设给定的训练数据集为 $\{\boldsymbol{x}_i,y_i\}_{i=1}^N$，输入 $\boldsymbol{x}_i\in\mathbf{R}^n$，输出 $y_i\in\mathbf{R}$，其在特征空间中的回归函数为

$$y(\boldsymbol{x})=\boldsymbol{\omega}^{\mathrm{T}}\varphi(\boldsymbol{x})+b. \tag{2}$$

式中：$\varphi(\cdot)$ 为特征空间的非线性映射，$\boldsymbol{\omega}$ 是与特征空间相同维度的权值向量，b 为偏移量.

LS—SVR 在特征空间的回归建模问题可以描述为求解如下二次规划(QP)问题：

$$\begin{cases}\min\limits_{\boldsymbol{\omega},b}J(\boldsymbol{\omega},b,\boldsymbol{e})=\dfrac{1}{2}\left(\gamma\sum\limits_{i=1}^{N}\|e_i\|^2+\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega}\right),\\ \text{s.t.}\quad y_i=\boldsymbol{\omega}^{\mathrm{T}}\varphi(\boldsymbol{x}_i)+b+e_i,i=1,2,\cdots,N.\end{cases} \tag{3}$$

式中，γ 为权衡结构风险与经验风险的正则化系数，e_i 为误差. 为简化计算，引入拉格朗日乘子

$$L(\boldsymbol{\omega},b,\boldsymbol{e},\boldsymbol{\alpha})=J(\boldsymbol{\omega},b,\boldsymbol{e})-\sum_{i=1}^{N}\alpha_i(\boldsymbol{\omega}^{\mathrm{T}}\varphi(\boldsymbol{x}_i)+b+e_i-y_i). \tag{4}$$

式中，$\alpha_i\in\mathbf{R}$ 为拉格朗日乘子. 令式(4)中的各偏导数为零

$$\begin{cases}\dfrac{\partial L}{\partial\boldsymbol{\omega}}=0\Rightarrow\boldsymbol{\omega}=\sum\limits_{i=1}^{N}\alpha_i\varphi(\boldsymbol{x}_i),\\ \dfrac{\partial L}{\partial b}=0\Rightarrow\sum\limits_{i=1}^{N}\alpha_i=0,\\ \dfrac{\partial L}{\partial e_i}=0\Rightarrow\alpha_i=\gamma e_i,\\ \dfrac{\partial L}{\partial\alpha_i}=0\Rightarrow\boldsymbol{\omega}^{\mathrm{T}}\varphi(\boldsymbol{x}_i)+b+e_i-y_i=0,\end{cases}\quad i=1,2,\cdots,N. \tag{5}$$

消去变量 $\boldsymbol{\omega}$ 和 e_i，得到如下线性方程组：

$$\begin{bmatrix}0 & \boldsymbol{I}^{\mathrm{T}}\\ \boldsymbol{I} & \boldsymbol{\Omega}+\gamma^{-1}\boldsymbol{I}\end{bmatrix}\begin{bmatrix}b\\ \boldsymbol{\alpha}\end{bmatrix}=\begin{bmatrix}0\\ \boldsymbol{y}\end{bmatrix}. \tag{6}$$

式中，$\boldsymbol{y}=(y_1,y_2,\cdots,y_N)^{\mathrm{T}}$，$\boldsymbol{I}^{\mathrm{T}}=(\overbrace{1,1,\cdots,1}^{N})$，$\boldsymbol{\alpha}=(\alpha_1,\alpha_2,\cdots,\alpha_N)^{\mathrm{T}}$，$\boldsymbol{\Omega}$ 是 N 维方阵，$\Omega_{mn}=\varphi(\boldsymbol{x}_m)\varphi(\boldsymbol{x}_n)=K(\boldsymbol{x}_m,\boldsymbol{x}_n)$ 为满足 Mercers 条件的核函数. 本文选用如下高斯核函数：

$$K(\boldsymbol{x}_m,\boldsymbol{x}_n)=\mathrm{e}^{-\frac{\|\boldsymbol{x}_m-\boldsymbol{x}_n\|^2}{\sigma^2}}. \tag{7}$$

2.2 稀疏化改进

由式(5)可知，$\boldsymbol{\omega}$ 为输入向量在特征空间的线性组合，通过寻找输入向量在特征空间的近似基可以一定程度的提高解的稀疏性[8]. 现将训练数据集 $\{\boldsymbol{x}_i,y_i\}_{i=1}^N$ 通过径向基函数 $\varphi(\cdot)$ 映射到高维希尔伯特空间，映射集为 $A=\{\varphi(\boldsymbol{x}_i)\}_{i=1}^N$. 由矩阵分析论可知，若 $\{\varphi(\boldsymbol{x}_i)\}_{i=1}^N$ 线性相关，则至少存在一个 $\varphi(\boldsymbol{x}_q)=\sum\limits_{i=1,i\neq q}^{N}\lambda_i\varphi(\boldsymbol{x}_i)$，其中 $\lambda_i\in\mathbf{R}$. 虽然 $\varphi(\cdot)$ 不能被确切的表达，但 $K(\boldsymbol{x}_q,\boldsymbol{x}_q)=\sum\limits_{i=1,i\neq q}^{N}\sum\limits_{j=1,j\neq q}^{N}\lambda_i\lambda_jK(\boldsymbol{x}_i,\boldsymbol{x}_j)$. 映射集 A 的极大无关组的求解步骤如下.

步骤 1　初始化极大无关组集 $B=\varnothing$，在集合 $S=(1,2,\cdots,N)$ 选取数据 $i=1$ 放到 B 中.

步骤 2　在 S 中依次选取 $i=i+1$，计算 $\min\limits_{\lambda}G(\lambda)=\left(\varphi(\boldsymbol{x}_i)-\sum\limits_{i\in B}\lambda_i\varphi(\boldsymbol{x}_i)\right)^{\mathrm{T}}\left(\varphi(\boldsymbol{x}_i)-\sum\limits_{i\in B}\lambda_i\varphi(\boldsymbol{x}_i)\right)$.

步骤 3　若 $\min\limits_{\lambda}G(\lambda)<\varepsilon$，则说明 $\varphi(\boldsymbol{x}_i)$ 可以由 $\{\varphi(\boldsymbol{x}_i)\mid i\in B\}$ 线性表示，摒弃数据 i；若 $\min\limits_{\lambda}G(\lambda)\geqslant\varepsilon$，则说明 $\varphi(\boldsymbol{x}_i)$ 不可由 $\{\varphi(\boldsymbol{x}_i)\mid i\in B\}$ 线性表示，则 $\{\varphi(\boldsymbol{x}_i),\varphi(\boldsymbol{x}_{i\in B})\}$ 线性无关，将 i 放到集合 B 中.

步骤 4　若迭代次数 $i\leqslant N$，则转到步骤 2；否则终止迭代.

将集合 B 中所对应的训练数据集的元素取出组成稀疏后的训练数据集 $\Psi=\{\boldsymbol{x}_k,y_k\}_{k=1}^r$，$r$ 为经稀疏化处理后训练数据集的样本数. Ψ 通过径向基函数映射后为 $\boldsymbol{\Phi}=(\varphi(\boldsymbol{x}_1),\varphi(\boldsymbol{x}_2),\cdots\varphi(\boldsymbol{x}_r))$. 因为 Ψ 是 A 的极大无关组，则

$$\boldsymbol{\omega}=\sum_{k=1}^{r}\beta_k\varphi(\boldsymbol{x}_k)=\boldsymbol{\Phi\beta}. \tag{8}$$

将式(3)中的 $\boldsymbol{\omega}$ 用式(8)替换得

$$\begin{cases}\min\limits_{\boldsymbol{\beta},b}F(\boldsymbol{\beta},b,\boldsymbol{e})=\dfrac{1}{2}\left(\gamma\sum\limits_{i=1}^{r}\|e_i\|^2+(\boldsymbol{\Phi\beta})^{\mathrm{T}}(\boldsymbol{\Phi\beta})\right),\\ \text{s.t. } y_i=(\boldsymbol{\Phi\beta})^{\mathrm{T}}\varphi(\boldsymbol{x}_i)+b+e_i, i=1,2,\cdots,r.\end{cases} \tag{9}$$

2.3　鲁棒改进

为了提高上述稀疏化 LS—SVR，即 S—LS—SVR 算法的鲁棒性能，对式(9)中的误差项 e_i 引入加权因子 ν_i，从而得到如下的优化问题：

$$\begin{cases}\min\limits_{\boldsymbol{\beta},b}F(\boldsymbol{\beta},b,\boldsymbol{e})=\dfrac{1}{2}\left(\gamma\sum\limits_{i=1}^{r}\nu_i\|e_i\|^2+(\boldsymbol{\Phi\beta})^{\mathrm{T}}(\boldsymbol{\Phi\beta})\right),\\ \text{s.t. } y_i=(\boldsymbol{\Phi\beta})^{\mathrm{T}}\varphi(\boldsymbol{x}_i)+b+e_i.\end{cases} \tag{10}$$

引入拉格朗日算子之后可得

$$L(\boldsymbol{\beta},b,\boldsymbol{e},\boldsymbol{\alpha})=F(\boldsymbol{\beta},b,\boldsymbol{e})-\sum_{i=1}^{r}\alpha_i((\boldsymbol{\Phi\beta})^{\mathrm{T}}\varphi(\boldsymbol{x}_i)+b+e_i-y_i). \tag{11}$$

式中：$\boldsymbol{\alpha}\in\mathbf{R}^r$ 为拉格朗日乘子，根据最优条件消去 $\boldsymbol{e}$ 和 $\boldsymbol{\alpha}$ 可得

$$\begin{bmatrix}\boldsymbol{\Phi}^{\mathrm{T}}\sum\limits_{i=1}^{r}\varphi(\boldsymbol{x}_i)\varphi(\boldsymbol{x}_i)^{\mathrm{T}}\boldsymbol{\Phi}-\dfrac{1}{\gamma\boldsymbol{\nu}}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi} & \boldsymbol{\Phi}^{\mathrm{T}}\sum\limits_{i=1}^{r}\varphi(\boldsymbol{x}_i)\\ \sum\limits_{i=1}^{r}\varphi(\boldsymbol{x}_i)\boldsymbol{\Phi}^{\mathrm{T}} & N\end{bmatrix}\cdot\begin{bmatrix}\boldsymbol{\beta}\\ b\end{bmatrix}=\begin{bmatrix}\boldsymbol{\Phi}^{\mathrm{T}}\sum\limits_{i=1}^{r}y_i\varphi(\boldsymbol{x}_i)\\ \sum\limits_{i=1}^{r}y_i\end{bmatrix} \tag{12}$$

式中，$\boldsymbol{\nu}=\mathrm{diag}(\nu_1,\nu_2,\cdots,\nu_r)$．$\boldsymbol{\nu}$ 由 IGGIII 权函数[9]决定，即：

$$\nu_i=\begin{cases}1, & |e_i|<k_1\zeta;\\ k_1\left|\dfrac{\zeta}{e_i}\right|\left(\dfrac{k_2-\left|\dfrac{e_i}{\zeta}\right|}{k_2-k_1}\right)^2, & k_1\zeta\leqslant|e_i|<k_2\zeta;\\ 0, & |e_i|\geqslant k_2\zeta.\end{cases} \tag{13}$$

式中：ζ 为误差的估计标准差；k_1 和 k_2 为相关系数，根据经验值有 $k_1\in[1,3]$，$k_2\in[3.2,6]$．

3　R—S—LS—SVR 参数多目标遗传优化

LS—SVR 经过上述稀疏化和鲁棒性改进后得到 R—S—LS—SVR，它有两个结构参数需要确定，分别是决定离群点惩罚程度的正则项 C 和径向基核函数的伸缩量 σ．常见的模型参数确定方法有网格搜索和交叉验证，但这些算法效率低且易陷入局优．本文将采用基于非支配排序的带有精英策略的多目标优化算法(NSGA—II)对 C 和 σ 进行优化计算，从而得到最终的 NSGAII—R—S—LS—SVR 的模型．

3.1　模型精度多目标评价

参数遗传优化的首要任务是构建性能指标作为适应度函数．常见的建模性能指标大多采用均方根误差(RMSE)，未从整体上考虑模型输出曲线与实际曲线的接近程度和动态趋势．实际上，准确的变化趋势对于动态过程的建模至关重要．为此，提出综合均方根误差和估计曲线与实际曲线相关性的模型精度多目标评价指标．

由数理统计理论可知，两个随机数据矢量 X 和 Y，$E\{(X-E(X))(Y-E(Y))\}$ 称之为 X 与 Y 的协方差或者相关矩，记作 $\mathrm{Cov}(X,Y)$，而两者的相关系数定义为

$$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}. \tag{14}$$

式中，$E(X)$ 和 $E(Y)$ 分别是 X 和 Y 的期望，$\sqrt{D(X)}$ 和 $\sqrt{D(Y)}$ 分别是 X 和 Y 的方差．相关系数 ρ_{XY} 是衡量数据变量 X 与 Y 关系程度的量：$|\rho_{XY}|\to1$ 表示 X 和 Y 之间的线性关系密切，$|\rho_{XY}|\to0$ 表示 X 和 Y 的相关性很差；若 $|\rho_{XY}|=1$ 表示 X 和 Y 依概率 1 存在着线性关系，而 $|\rho_{XY}|=0$ 表示 X 和 Y 不相关．

综上所述，提出模型精度多目标评价指标如下：

$$F_{\mathrm{CEI1}}=\frac{1}{N}\sum_{i=1}^{N}e_i^2, \tag{15}$$

$$F_{\mathrm{CEI2}}=1-\rho_{XY}. \tag{16}$$

该多目标评价指标既可以保证建模过程的平稳性和限制输出曲线的横向偏移量，又可以保证建模过程的准确性和限制输出曲线的纵向偏移量．

3.2　基于 NSGA—II 的模型参数多目标优化

针对于模型精度多目标评价指标，利用 NSGA—II 进行 C，σ 参数寻优．NSGA—II 在工程领域已得到广泛应用[10]，但是存在计算复杂度高、缺少精英策略以及需要人为制定共享参数的问题[11-14]．为此，采用改进的 NAGA—II 算法，以模型精度多目标评价指标为适应度函数，采用实数编码，通过基于进行非支配快速排序和拥挤距离计算的种群进行二进制锦标赛选择，模拟二进制交叉和多项式变异的遗传因子增强种群多样性，主要计算流程如图 1 所示．

(1) 选择算子．根据非支配排序的结果，选择支配层较低的个体，若同一支配层的个体有多个，选择拥挤距离较大的个体以获得种群的多样性[12]．

(2) 模拟二进制交叉．由于采用实数编码，则交叉后代是父代的线性组合．

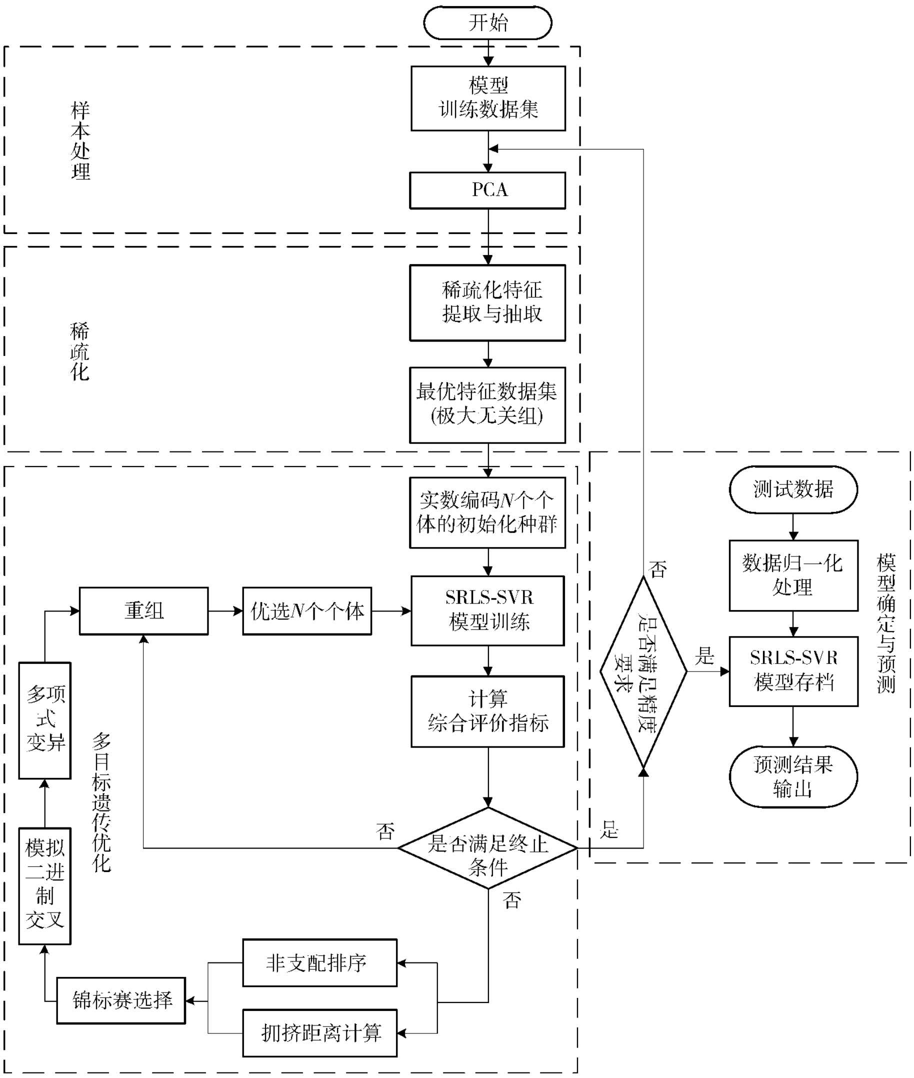

图 1 NSGAII-R-S-LS-SVR 建模流程

Fig. 1 Strategy diagram of NSGAII-R-S-LS-SVR based modeling

$$\begin{cases} G_{a,i}^{t+1} = \dfrac{1}{2}\left[(1-\beta_k(u))G_{1,i}^{t} + (1+\beta_k(u))G_{2,i}^{t}\right], \\ G_{b,i}^{t+1} = \dfrac{1}{2}\left[(1+\beta_k(u))G_{1,i}^{t} + (1-\beta_k(u))G_{2,i}^{t}\right]. \end{cases} \tag{17}$$

式中: $G_{a,i}^{t+1}$ 和 $G_{b,i}^{t+1}$ 分别为第 $t+1$ 代均方根误差与相关系数种群的第 i 个个体; $G_{1,i}^{t}$ 和 $G_{2,i}^{t}$ 分别为由选择算法在第 t 代中选择的两个优良个体; u 为(0,1)均匀分布的随机数; 当 $u>0.5$ 时 $\beta_k = ([2(1-u)]^{(\eta_c+1)})^{-1}$,当 $u\leqslant 0.5$ $\beta_k = (2u)^{(\eta_c+1)^{-1}}$,$\eta_c$ 为交叉分布指数,k 为当代种群的第 k 个个体.

(3) 多项式变异. 变异后的个体如下所示:

$$G_i^{t+1} = G_i^{t+1} + (B^{u} - B^{l})\delta_k. \tag{18}$$

式中: B^{u} 和 B^{l} 分别为优化变量的上界与下界,δ_k 为变异的参数. 当 $r_k>0.5$ 时 $\delta_k = (2r_k)^{(\eta_m+1)^{-1}}$,当 $r_k\leqslant 0.5$ 时 $\delta_k = \{1-[2(1-r_k)]\}^{(\eta_m+1)^{-1}}$,$r_k$ 为来自(0,1)均匀分布的随机数,η_m 为变异分布指数.

4 工业实验与分析

实验使用两组测试数据,即 Data 1 和 Data 2,如图 2 所示. Data 1 是经主成分分析法降维处理后的 270 组实际现场采集的高炉数据,其中训练数据 200 组,测试数据 70 组; Data 2 是在 Data 1 中随机加入 11 个离群点得到的数据. 为了表述方便,我们称 Data 1 为原始数据,Data 2 为离群点数据. 具体实验时,为了更好地说明问题,将提出的 NSGAII-R-S-LS-SVR 算法与标准 LS-SVR 以及极限学习机(ELM)算法[13]进行比较,三者的训练和测试数据完全相同,并且都是在归一化之后导入模型. 另外,采用一些标准指标对模型性能进行评价,这些指标包括建模时间(MT)、均方根误差(RMSE)、平均绝对误差(MAE)、相对误差(RE)、回归系数(RC)和命中率(HR).

多目标遗传参数优化时,C 和 σ 经上述多目标遗传优化的 Pareto 前沿进化过程如图 3 所示. 首先将上

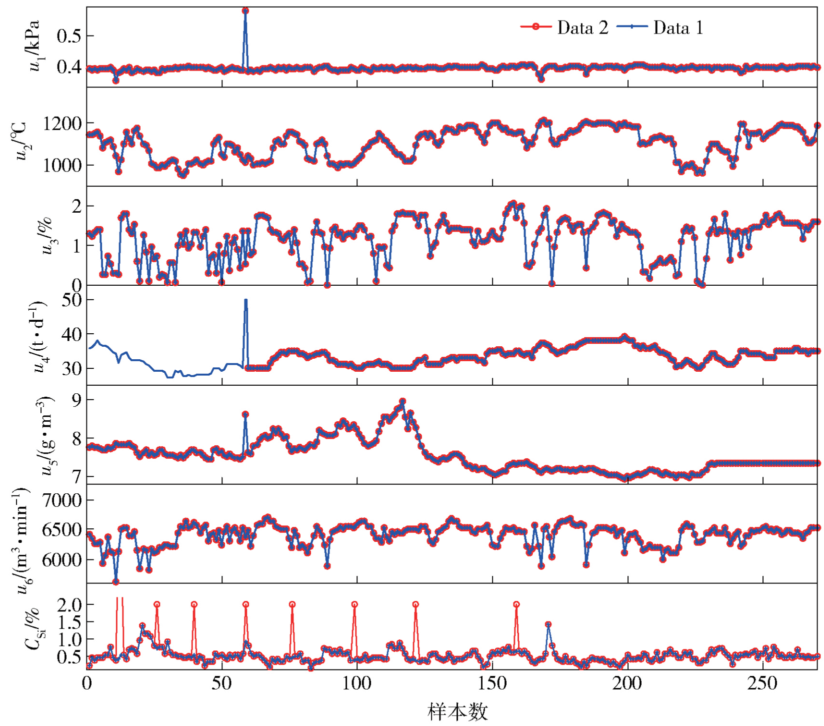

图 2　建模输入输出数据

Fig. 2　Input data and output data of modeling

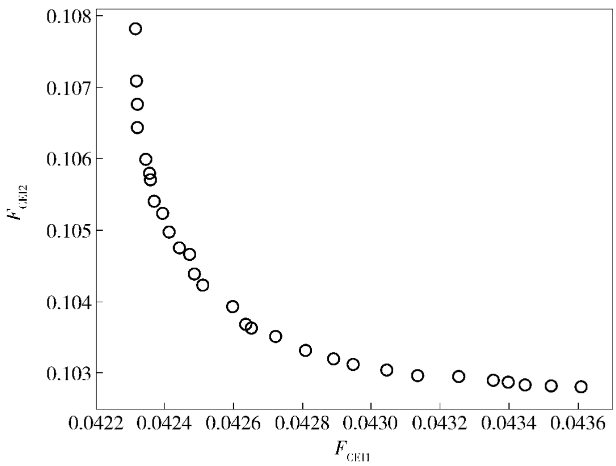

图 3　NSGA—II 参数优化 Pareto 前沿进化过程

Fig. 3　Frontier evolution of Pareto with NSGA—II parameter optimization

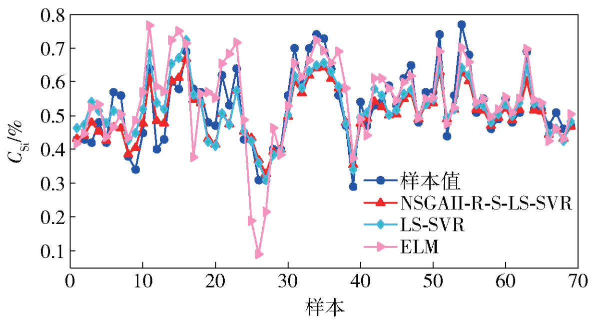

图 4　数据未包含离群点建模时各模型的铁水 [Si] 估计效果

Fig. 4　Modeling results of [Si] with different algorithms using data without outliers

述三种算法针对 Data 1 (即未包含离群点的数据) 进行训练,得到的模型对 [Si] 的估计效果如图 4 和图 5 所示. 从总的估计效果来看,针对没有加入离群点的原始数据建立的 NSGAII—R—S—LS—SVR、LS—SVR、ELM 等铁水 [Si] 软测量模型似乎都可较好地跟踪原始数据的变化,具有较高的估计精度. 但从局部细节来看, LS—SVR、ELM 的数据波动较大,相比之下 NSGAII—R—S—LS—SVR 估计误差概率密度函数(PDF) 曲线对称轴更加逼近于中轴,估计误差自相关曲线更接近白噪声,因而估计效果更好、泛化能力更强. 表 1 给出两种算法建立的模型对铁水 [Si] 估计性能指标的定量统计分析结果,可以看出所提 NSGAII—R—S—LS—SVR 的均方根误差和平均绝对误差最小,训练时间较 LS—SVR 缩短 48.76%, [Si] 估计误差不超过 ±0.1 的命中率达 92.86%. 因此所提方法稀疏化带来的估计精度效果显著.

实际生产中,数据总会不同程度受到噪声等外界干扰[14-15],若在常规数据建模时将这些离群点考虑进去会严重的影响建模效果以及模型的泛化能力. 为了验证所提 NSGAII—R—S—LS—SVR 建模算法对数据离群点的鲁棒性能,仍采用 NSGAII—R—S—LS—SVR、LS—SVR 和 ELM 三种算法针对 Data 2 (即包含离群点的数据) 重新进行建模和比较分析,得到的铁水 [Si] 估计

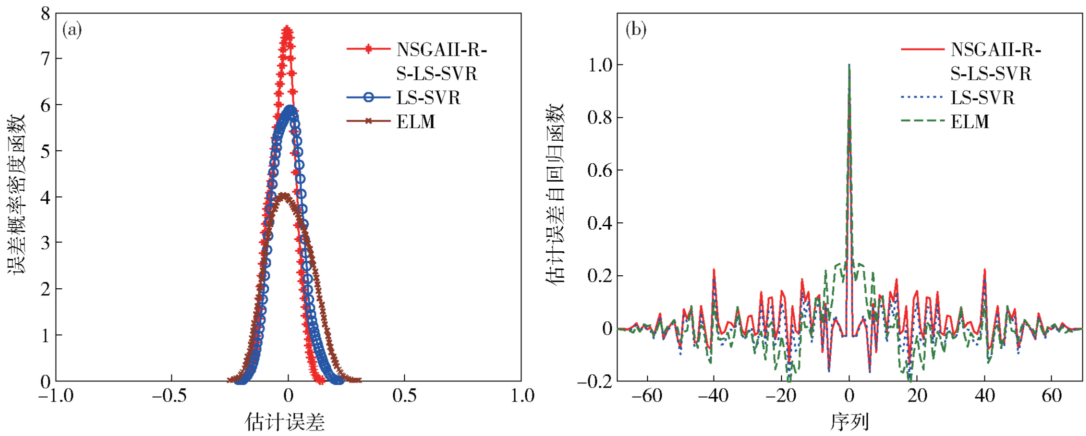

图 5　数据未包含离群点建模时各模型的铁水 [Si] 估计误差概率密度函数(a) 和自相关函数(b)

Fig. 5　PDF (a) and autocorrelation function (b) of [Si] estimation error with different algorithms using data without outliers

表 1　数据未包含离群点建模时各模型性能指标比较

Table 1　Performance comparison of [Si] estimation with different algorithms using data without outliers

算法	MT/s	RC	MAE $\frac{1}{N}\sum_{i=1}^{N}\lvert e_i\rvert$	RMSE $\sqrt{\frac{1}{N}\sum_{i=1}^{N}\lvert e_i\rvert^2}$	HR/% $\frac{1}{N}\sum_{i=1}^{N}H_k\times 100\%$, $H_k=\begin{cases}1, & \lvert e_i\rvert<0.1\\ 0, & \lvert e_i\rvert\geqslant 0.1\end{cases}$	RE $\frac{\sum_{i=1}^{N}\lvert e_i\rvert^2}{\sum_{i=1}^{N}\lvert y_i\rvert^2}$
LS−SVR	0.1472	0.9779	0.0460	0.2144	91.43	0.0110
ELM	0.0470	0.9676	0.0672	0.2593	68.57	0.0356
NSGAII−R−S−LS−SVR	0.0760	0.9458	0.0413	0.2032	92.86	0.0100

效果如图 6 ~ 图 8 所示．可以看出：当训练数据有离群点存在时，LS−SVR 和 ELM 算法所建立的软测量模型的实际估计效果很差，基本不能跟踪实际数据的变化；而所提 NSGAII−R−S−LS−SVR 算法建立的模型对铁水 [Si] 的估计效果仍然很好，能够很好地跟踪其实际值的变化．

由图 7 可以看出：LS−SVR 和 ELM 模型的估计误差概率密度函数(PDF) 曲线延伸范围大，左右很不对称，它们的估计误差自相关曲线虽然形状与白噪声相

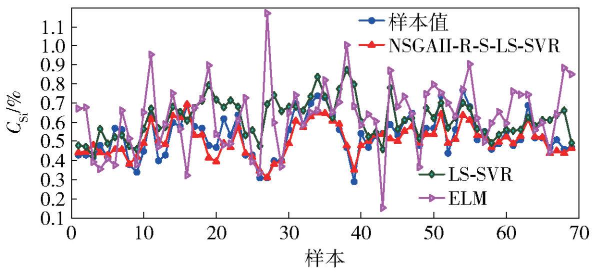

图 6　数据包含离群点建模时各模型的铁水 [Si] 估计效果

Fig. 6　Modeling results of [Si] with different algorithms using data with outliers

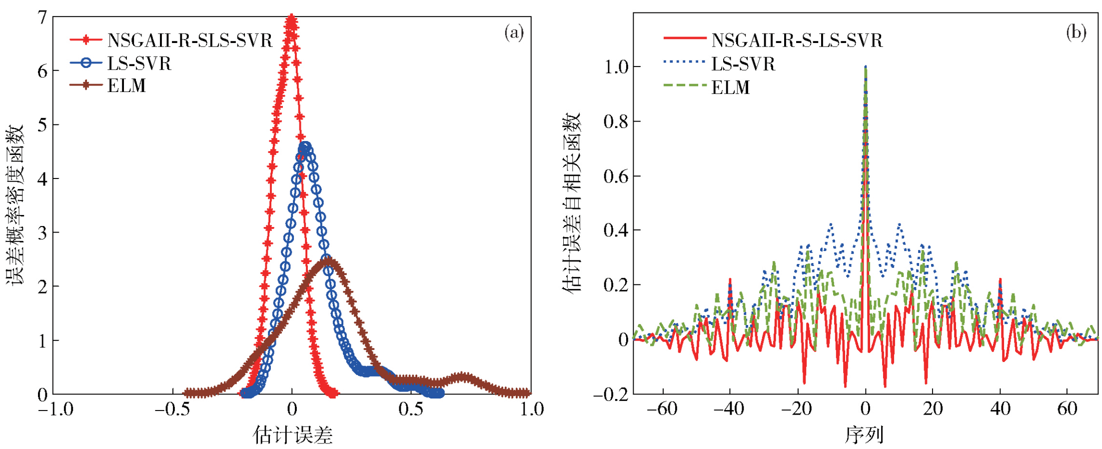

图 7　数据包含离群点建模时各模型的铁水 [Si] 估计误差概率密度函数(a) 和自相关函数(b)

Fig. 7　PDF (a) and autocorrelation function (b) of [Si] estimation error with different algorithms using data with outliers

近，但是振幅较大且偏离零中心线; 而 NSGAII—R—S—LS—SVR 所建立的模型的估计误差概率密度函数曲线延伸范围小，并且是对称、零均值的. 图 8 为针对于离群点数据各算法的回归分析，其中所提算法的实际值与预测值更加接近. 由表 2 可知: 对于离群点数据的建模，所提 NSGAII—R—S—LS—SVR 模型对实际数据估计的均方根误差最小，回归系数更加逼近理想值 1; 命中率达到 91.43%，相比原始数据 92.86% 的命中率，下降波动仅 1.43%，因而相对鲁棒性较强.

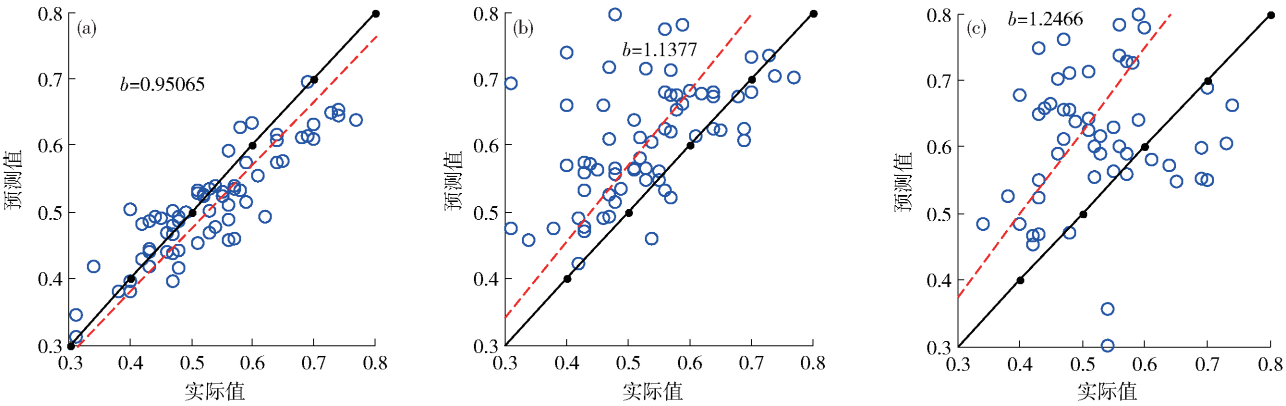

图 8　数据包含离群点建模时 NGGA—II—R—SLS—SVR(a) 、LS—SVR(b) 和 ELM(c) 的回归分析结果

Fig. 8　Regression analysis results of [Si] estimation with NSGA—II—R—SLS—SVR (a) , LS—SVR (b) and ELM (c) using data with outliers

表 2　数据包含离群点建模时各模型性能指标比较

Table 2　Performance comparison of [Si] estimation with different algorithms using data with outliers

算法	MT/s	RMSE	RC	HR/%	RE
LS—SVR	0. 1164	0. 3196	0. 8623	64. 29	0. 0720
ELM	0. 0523	0. 4746	0. 7534	30. 00	0. 2516
NSGAII—R—S—LS—SVR	0. 0934	0. 2056	0. 9506	91. 43	0. 0101

5　结论

(1) 针对机理建模难以准确估计铁水硅含量的难题，利用数据驱动建模的思想，提出一种基于模型精度多目标评价与多目标遗传参数优化的稀疏鲁棒最小二乘支持向量机算法，用于对铁水硅含量进行动态软测量.

(2) 对某大型钢铁厂的实际采集数据进行实验的结果表明所提方法具有良好的估计效果. 与 LS—SVR 和 ELM 相比，所提方法铁水硅含量估计误差小于 ±0. 1 的样本数占测试样本数的 90% 以上(如表 2 所示) ，具有建模时间短且对离群点的鲁棒性强的优势，可用于高炉铁水的实际在线估计.

参　考　文　献

[1] Zhang J L, Wang G W, Shao J G, et al. Comprehensive mathematical model and optimum process parameters of nitrogen free blast furnace. *J Iron Steel Res Int*, 2014, 21(2) : 151

[2] Gao C H, Jian L, Liu X Y, et al. Data-driven modeling based on Volterra series for multidimensional blast furnace system. *IEEE Trans Neural Networks*, 2011, 22(12) : 2272

[3] Wang Y K, Liu X G. Chaotic time series forecasting based on SVM for silicon content in hot metal // *Proceedings of the 33rd Chinese Control Conference*. Nanjing, 2014: 5156

[4] 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 Trans Ind Inf*, 2013, 9(4) : 2213

[5] Tang X L, Zhuang L, Hu X D. The support vector regression based on the chaos particle swarm optimization algorithm for the prediction of silicon content in hot metal. *Control Theory Appl*, 2009, 26(8) : 838
(唐贤伦，庄陵，胡向东. 铁水硅含量的混沌粒子群支持向量机预报方法. 控制理论与应用，2009，26(8) : 838)

[6] Gan L Z, Sun Z H, Sun Y X. Sparse least squares support vector machine. *J Zhejiang Univ Eng Sci*, 2007, 41(2) : 245
(甘良志，孙宗海，孙优贤. 稀疏最小二乘支持向量机. 浙江大学学报(工学版) ，2007，41(2) : 245)

[7] Zhou P, Yuan M, Wang H, et al. Multivariable dynamic modeling for molten iron quality using online sequential random vector functional-link networks with self-feedback connections. *Inf Sci*, 2015, 325: 237

[8] Gan L Z, Liu H K, Sun Y X. Sparse least squares support vector machine for function estimation // *Proceedings of the 3rd International Symposium on Neural Networks*. Chengdu, 2006: 1016

[9] Xu Q F, Zhang J X, Jiang C X, et al. Weighted quantile regression *via* support vector machine. *Expert Syst Appl*, 2015, 42 (13) : 5441

[10] Li Y G, Shen J, Lu J H. Constrained model predictive control of a solid oxide fuel cell based on genetic optimization. *J Power*

Sources，2011，196(14)：5873

[11] Li C P，Zheng Y X，Zhang J. Ore grade interpolation model based on support vector machines optimized by genetic algorithms. *J Univ Sci Technol Beijing*，2013，35(7)：837
(李翠平，郑瑶瑕，张佳. 基于遗传算法优化的支持向量机品位插值模型. 北京科技大学学报，2013，35(7)：837)

[12] Han Y Y，Gong D W，Sun X Y，et al. An improved NSGA-II algorithm for multi-objective lot-streaming flow shop scheduling problem. *Int J Prod Res*，2014，52(8)：2211

[13] Schmidhuber J. Deep learning in neural networks：an overview. *Neural Networks*，2015，61：85

[14] Li C H，Wang Y F，Cai M F，et al. Slope deformation model of metal mines transferred underground mining from open-pit based on support vector machines. *J Univ Sci Technol Beijing*，2009，31(8)：945
(李长洪，王云飞，蔡美峰，等. 基于支持向量机的露天转地下开采边坡变形模型. 北京科技大学学报，2009，31(8)：945)

[15] Agrama F A. Multi-objective genetic optimization for scheduling a multi-storey building. *Autom Constr*，2014，44：119