

研究论文

DOI: 10.11949/0438-1157.20190716

基于改进回声状态网络的游离氧化钙预测控制

李德健¹, 刘浩然², 刘彬¹, 刘泽仁¹, 王卫涛¹, 闻岩³(¹燕山大学电气工程学院, 河北 秦皇岛 066004; ²燕山大学信息科学与工程学院, 河北 秦皇岛 066004;³燕山大学机械工程学院, 河北 秦皇岛 066004)

摘要: 在非线性时延水泥烧成系统中, 针对传统预测控制方法调节时间长、控制精度不高的问题, 提出一种改进的在线型回声状态网络预测控制模型。首先将带有L1范数约束项的递归最小二乘法与回声状态网络相结合构建在线型预测模型, 解决传统预测控制模型辨识精度较低、无法进行实时预测的问题; 然后基于改进的回声状态网络预测模型, 构建预测控制模型结构, 并采用具有全局优化能力的粒子群算法进行滚动优化, 保证实际输出量快速、准确、平稳地跟随被控量的设定值; 最后利用改进的预测控制模型对水泥烧成系统中的游离氧化钙含量进行预测控制仿真实验, 结果表明改进的预测控制模型具有良好的性能和应用前景。

关键词: 模型预测控制; 神经网络; 回声状态网络; L1正则化; 优化; 烧成系统

中图分类号: TP 183

文献标志码: A

文章编号: 0438-1157 (2019) 12-4749-11

Predictive control of free calcium oxide based on improved echo state network

LI Dejian¹, LIU Haoran², LIU Bin¹, LIU Zeren¹, WANG Weitao¹, WEN Yan³

(¹ College of Electrical Engineering, Yanshan University, Qinhuangdao 066004, Hebei, China; ² College of Information Science and Engineering, Yanshan University, Qinhuangdao 066004, Hebei, China; ³ College of Mechanical Engineering, Yanshan University, Qinhuangdao 066004, Hebei, China)

Abstract: In the nonlinear time-delay cement burning system, an improved on-line echo state network predictive control model is proposed for the problem that the traditional predictive control method has long adjustment time and low control precision. Therefore, we firstly combine the L1-norm constrained recursive least squares method with the echo state network to construct an on-line prediction model to address these issues. Then, the structure of predictive control model is constructed based on the improved prediction model of echo state network. And particle swarm optimization (PSO) algorithm with global optimization capability is utilized for rolling to ensure that the actual output follows the setting value of the controlled variable quickly, accurately and smoothly. Finally, the simulation prediction experiments of the free calcium oxide content in the cement burning system are conducted using the improved predictive control model. The results show that the improved predictive control model has good performance and application prospects.

Key words: model predictive control; neural network; echo state network; L1 regularization; optimization; burning system

收稿日期: 2019-06-25 修回日期: 2019-07-29

通信作者: 刘浩然 (1980—), 男, 硕士, 教授, haoranliuysu@163.com

第一作者: 李德健 (1994—), 男, 硕士研究生, 501807366@qq.com

基金项目: 河北省自然科学基金项目(F2019203320, E2018203398); 国家自然科学基金项目(51641609)

引用本文: 李德健, 刘浩然, 刘彬, 刘泽仁, 王卫涛, 闻岩. 基于改进回声状态网络的游离氧化钙预测控制[J]. 化工学报, 2019, 70(12): 4749–4759

Citation: LI Dejian, LIU Haoran, LIU Bin, LIU Zeren, WANG Weitao, WEN Yan. Predictive control of free calcium oxide based on improved echo state network[J]. CIESC Journal, 2019, 70(12): 4749–4759

引 言

水泥熟料中游离氧化钙(free lime in cement clinker, fCaO)是指水泥熟料中没有以化合状态存在,而是以游离状态存在的氧化钙^[1-2]。在水泥生产过程中,熟料 fCaO 含量的高低直接影响水泥安定性及熟料强度,需要将其含量控制在一定范围内,因此 fCaO 含量值控制方法的研究得到众多学者的关注^[3-4]。目前,国内水泥熟料 fCaO 含量的控制方法主要有动态矩阵控制(dynamic matrix control, DMC)方法和模型预测控制(model predictive control, MPC)方法。DMC 方法虽然算法简单,但控制准确性容易受到现场实际工况的影响,造成实际输出与设定值相差较大^[5-6]。MPC 方法的线性控制理论研究虽然已经比较成熟,但将其应用于实际工业的强非线性对象仍不能得到满意的控制效果。神经网络是一种模仿动物神经网络行为特征的数学模型,能够充分逼近任意复杂的非线性函数,因而神经网络在 MPC 中得到了广泛的应用^[7-8]。但基于经验风险最小化的神经网络在实际应用中仍存在很多问题,如训练速度慢、易陷入局部最优等,一定程度上限制了其在实际工业中的发展^[9]。

回声状态网络(echo state networks, ESN)是一种新型的递归神经网络(recurrent neural networks, RNN),其输入权值和内部权值在初始化阶段随机给定且保持不变,在训练过程中仅需要确定输出权值,这种训练方式实现了预测精度与计算复杂度的折中,使 ESN 在非线性系统辨识方面较传统 RNN 有较大的改进^[10-12]。对于动态系统,由于系统数据不断更新,往往难以全面地采集具有代表性的数据,因此利用基于离线数据训练的 ESN 虽然在短时间内可以描述系统内部状态,但由于其参数不能自适应更新,使得模型难以长时间保持准确。针对以上问题,Jaeger 等^[13]提出了递归最小二乘法回声状态网络(recursive least squares-echo state networks, RLS-ESN),将 RLS 算法用于 ESN 网络的输出权值动态调节,实现 ESN 的在线学习。RLS-ESN 作为在线型预测模型,具有较高的性能指标,但是其网络结构与训练算法仍不够成熟。

网络储备池是 RLS-ESN 的核心,其蕴含系统的运行状态且具有较好的记忆功能。但若储备池规模过小,则无法有效利用储备池强大的映射能力,导致网络出现欠拟合现象,使网络的预测精度降

低;若储备池规模过大,则会引入大量无关特征,导致网络出现过拟合现象,进而影响网络的泛化能力。因此,有效控制储备池规模,建立最优的网络结构,是需要解决的关键问题。为提高网络的泛化能力,众多学者采用正则化方法控制模型规模。汪灵枝等^[14]提出基于贝叶斯正则化的神经网络,有效改善网络过拟合现象;史志伟等^[15]提出 ESN 岭回归学习算法,有效简化模型结构,具有较高的预测精度;黄炳家等^[16]提出带光滑 L1/2 正则化项的神经网络,相比于 L2 范数具有更强的稀疏性和剪枝能力;韩敏等^[17]提出基于 L1 范数正则化的 ESN 网络,利用 L1 范数的特征选择能力,有效控制控制模型的规模,具有良好的泛化能力。

粒子群优化(particle swarm optimization, PSO)算法是近年来发展起来的一种新的进化算法,是一种基于迭代的优化工具^[18-19]。由于 PSO 具有快速收敛、全局寻优和较高的运算精度等优点,被广泛应用于 MPC 的滚动优化环节当中。柴毅等^[20]采用 PSO 算法优化经典 ESN 网络,仿真结果表明其预测控制效果优于经典 PID 控制方法;张天瑜^[21]提出了 ODPSO 与 LS-SVM 相结合的预测控制算法,相比 DMC 控制方法具有更好的控制品质。基于以上分析,本文提出一种基于 L1 范数正则化的递归最小二乘法回声状态网络(L1 recursive least squares-echo state networks, L1RLS-ESN)预测控制模型。首先,为了避免网络出现过拟合现象、有效控制储备池规模,本文通过求得带有 L1 范数约束项的 L1RLS 算法,实现网络输出权值的动态调节,构造具有特征选择能力的在线型 L1RLS-ESN 预测模型;然后基于 L1RLS-ESN 预测模型,构造由粒子群算法进行滚动优化的预测控制结构,使网络能够快速收敛到全局最优解,满足作为控制器的实时性要求;最后通过分析水泥回转窑烧成系统的工艺流程,确定水泥熟料 fCaO 含量的控制变量,建立水泥熟料 fCaO 含量预测控制模型并进行仿真验证。

1 基于 L1 正则化的递归最小二乘法 ESN

1.1 基于 L1 正则化的递归最小二乘法 ESN 结构

ESN 是一种新型的 RNN,其状态方程为^[22]

$$\mathbf{x}(n+1) = f(\mathbf{W}^{\text{in}}\mathbf{u}(n+1) + \mathbf{W}\mathbf{x}(n) + \mathbf{W}^{\text{back}}\mathbf{y}(n)) \quad (1)$$

$$\mathbf{y}(n+1) = f^{\text{out}}(\mathbf{W}^{\text{out}}[\mathbf{u}(n+1), \mathbf{x}(n+1), \mathbf{y}(n)]) \quad (2)$$

式中, $u(n) \in \mathbb{R}^k$, $x(n) \in \mathbb{R}^N$ 和 $y(n) \in \mathbb{R}^L$ 分别表示 n 时刻的输入变量、储备池状态变量和输出变量。 $f(\cdot)$ 为隐藏层神经元激活函数, 通常情况下取双曲正切函数; $f^{\text{out}}(\cdot)$ 为输出层神经元激活函数, 通常情况下取线性函数^[23]。

目前, ESN 采用离线学习作为其训练方式, 即一次性投入所有的训练样本, 训练过程结束后即停止。然而对于实际工业应用中, 随着系统的不断运行, 会有大量新的样本出现, 离线型 ESN 不能自适应调整输出权值, 使模型难以保持准确。

在线学习能够将新样本加入到原有的样本集, 不断对已有模型进行误差修正, 得到动态适应的预测模型, 提高预测准确率。由于 ESN 通过线性回归算法确定输出权值, 基于误差最小化的自适应线性回归算法都可以应用于 ESN 的输出估计。因此, 考虑通过权值的动态调整思路, 将 RLS 引入到 ESN 的训练过程, 实现 ESN 的在线学习。然而, 在网络进行训练过程中, 储备池将会产生大量无关特征参与到输出权值的训练中, 无特征选择能力的 RLS 无法筛选冗余特征, 将导致网络泛化能力变差, 出现过拟合现象。

正则化方法是一种常用的特征选择方法, 其中 L0、L1、L2 三种正则化方法最为常见。然而在实际应用中, 由于 L0 正则化方法不易得到一个便于求解的数学表达式, 因此 L0 范数的求解过程较为困难, 其最优问题通常需要被转化为 L1 的最优化问题; L2 正则化方法通过衰减网络参数, 避免网络出现过拟合现象, 但会影响网络的预测精度; 而 L1 正则化方法可以通过产生系数权值矩阵, 即令网络的无关节点趋近于 0, 达到使网络具有特征选择能力、控制网络的复杂程度的目的^[24-25]。因此, 本文在 RLS 的二次损失函数中加入输出权值的 L1 范数约束项, 构造具有特征选择能力的 L1RLS 自适应模型。然后利用 L1RLS 训练 ESN 网络的输出权值, 使网络具有特征选择能力, 避免网络出现过拟合, 实现 ESN 的在线学习。L1RLS-ESN 网络结构如图 1 所示。图中 n 代表时刻; $u(n)$ 是 n 时刻输入信号, 即 ESN 网络储备池的状态矩阵; $y(n)$ 是数据经过网络训练后的实际输出; $d(n)$ 是期望响应信号; $e(n)$ 是误差信号。

从图 1 所示的 L1RLS-ESN 结构可知, L1RLS-ESN 和 ESN 最主要的区别是输出权值的训练方式。在网络开始训练时, 输入信号通过式(1)、式(2)的状

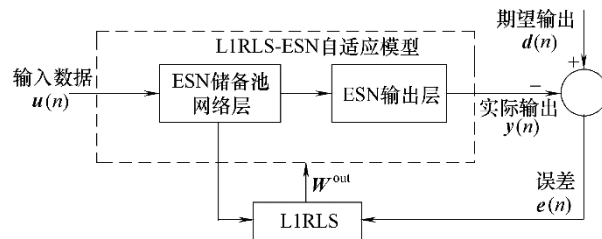


图1 L1RLS-ESN 结构

Fig.1 Structure of L1RLS-ESN

态方程得到网络输出; 然后网络计算出预测输出与实际输出之间的误差, 通过反向传播将误差传递到 L1RLS 中, L1RLS 根据误差值自适应更新输出权值 W^{out} ; 最后将 W^{out} 传递到 L1RLS-ESN 自适应模型进行网络参数的更新。

1.2 基于 L1 正则化的递归最小二乘法 ESN 学习算法

在训练过程中, 考虑在 RLS 的目标函数中加入 L1 范数约束项, 构建 L1RLS-ESN 预测模型。网络通过基于 L1RLS-ESN 目标函数的最小调整, 自适应调整网络输出权值 W^{out} 。其中 L1RLS-ESN 的损失函数为

$$G(n) = \frac{1}{2} \sum_{m=0}^n \lambda^{n-m} |e(m)|^2 + \gamma \|W^{\text{out}}(n)\|_1 \quad (3)$$

式中, $\lambda(n)$ 是 n 时刻遗忘因子, 其作用是离 n 时刻近的误差赋予较大权重, 确保在过去一段时间的观测数据被“遗忘”; $e(m)$ 表示误差信号, $e(m) = d(m) - W^{\text{out}}(n)x(m)$; $\gamma > 0$ 是权衡稀疏性对损失函数贡献大小的正则化参数。

在基于目标函数的最小化求解时, 通常利用梯度下降法对目标函数进行求导。然而, L1RLS-ESN 损失函数中的 L1 范数约束项并不能处处可导, 因此考虑采用次梯度方法对 L1RLS-ESN 的损失函数进行最小化调整。其中次梯度计算公式为^[26]

$$\partial \|h(n)\|_1 = \left\{ d \|d\|_\infty \leq 1, dh(n) = \|h(n)\|_1 \right\} \quad (4)$$

则 $\|W^{\text{out}}(n)\|_1$ 的第 k 个元素表示为

$$\left\{ \partial \|h(n)\|_1 \right\}_k = \begin{cases} \{d \mid |d| \leq 1\} & h_k = 0 \\ \{h_k / |h_k|\} & h_k \neq 0 \end{cases} \quad (5)$$

由式(5)可知, L1 范数的次梯度处处可导。根据上述结果, 求出 L1RLS-ESN 损失函数中 L1 范数约束项 $\|W^{\text{out}}(n)\|_1$ 的次梯度

$$\nabla^s \|W^{\text{out}}(n)\|_1 = \text{sgn}(W^{\text{out}}(n)) \quad (6)$$

式中, ∇^s 表示 n 时刻输出权值 L1 范数的次梯度; $\text{sgn}(\cdot)$ 返回一个整型变量, 若括号中数字大于 0,

则 sgn 返回 1, 数字等于 0, 则返回 0, 数字小于 0, 则返回 -1。将式(6)代入式(3)中, 得到 L1RLS-ESN 损失函数的次梯度

$$\nabla^S G(n) = \frac{1}{2} \nabla J(n) + \gamma \text{sgn}(\hat{\mathbf{W}}^{\text{out}}(n)) \quad (7)$$

L1RLS-ESN 损失函数的第 i 项次梯度表示为

$$\{\nabla^S G(n)\} = - \sum_{m=0}^n \lambda^{n-m} \mathbf{e}(m) \mathbf{x}^*(m-n+1) + \text{sgn}(\hat{\mathbf{W}}_i^{\text{out}}(n)) \quad (8)$$

为了确定当前时刻的最优输出权值 $\hat{\mathbf{W}}^{\text{out}}(n)$, 设置 L1RLS-ESN 第 i 项的次梯度为 0

$$\sum_{k=0}^{N-1} \hat{\mathbf{W}}_k^{\text{out}}(n) \left\{ \sum_{m=0}^n \lambda^{n-m} \mathbf{x}(m-k) \mathbf{x}^*(m-n+1) \right\} = \sum_{m=0}^n \lambda^{n-m} \gamma(m) \mathbf{x}^*(m-n+1) - \gamma \text{sgn}(\hat{\mathbf{W}}_i^{\text{out}}(n)) \quad (9)$$

将式(9)中, $i=1, \dots, N$ 项整合成一个矩阵的形式, 即可确定修正的正态方程

$$\mathbf{R}(n) \hat{\mathbf{W}}^{\text{out}}(n) = \mathbf{r}(n) - \gamma \text{sgn}(\hat{\mathbf{W}}^{\text{out}}(n)) \quad (10)$$

式中, $\mathbf{R}(n)$ 是指数加权的自相关矩阵; $\mathbf{r}(n)$ 是确定 $\mathbf{x}(n)$ 与 $\gamma(n)$ 之间的互相关矩阵; 这两个变量可由式(11)、式(12)表示

$$\mathbf{R}(n) = \lambda \mathbf{R}(n-1) + \mathbf{x}^*(n) \mathbf{x}^T(n) \quad (11)$$

$$\mathbf{r}(n) = \lambda \mathbf{r}(n-1) + \gamma(n) \mathbf{x}^*(n) \quad (12)$$

将式(10)右侧的方程定义为新变量 $\boldsymbol{\theta}(n)$, 即 $\boldsymbol{\theta}(n) = \mathbf{r}(n) - \gamma \text{sgn}(\hat{\mathbf{W}}^{\text{out}}(n))$ 。则将式(10)简化为

$$\mathbf{R}(n) \hat{\mathbf{W}}^{\text{out}}(n) = \boldsymbol{\theta}(n) \quad (13)$$

$\boldsymbol{\theta}(n)$ 的更新方程可被描述为以下形式的递归方程

$$\boldsymbol{\theta}(n) = \lambda \boldsymbol{\theta}(n-1) + \gamma(n) \mathbf{x}^*(n) - \left\{ \gamma \text{sgn}(\hat{\mathbf{W}}^{\text{out}}(n)) - \lambda \gamma \text{sgn}(\hat{\mathbf{W}}^{\text{out}}(n-1)) \right\} \quad (14)$$

由此, 无须求解正态方程, 只要确定如下形式的迭代方程即可确定当前时刻的最优权值

$$\mathbf{W}^{\text{out}}(n) = \mathbf{W}^{\text{out}}(n-1) + \Delta \mathbf{W}^{\text{out}}(n-1) \quad (15)$$

式中, $\Delta \mathbf{W}^{\text{out}}(n-1)$ 是一个瞬时修正值; 为了得到如式(15)形式的递归方程, 必须将式(14)右侧修改为只存在 $\mathbf{W}^{\text{out}}(n-1)$ 项的形式。因此, 假设权重的 sgn 返回值在一个时间步长不发生变化, 式(14)近似表示为

$$\boldsymbol{\theta}(n) \approx \lambda \boldsymbol{\theta}(n-1) + \gamma(n) \mathbf{x}^*(n) + \gamma(\lambda-1) \text{sgn}(\hat{\mathbf{W}}^{\text{out}}(n-1)) \quad (16)$$

为求取 $\hat{\mathbf{W}}^{\text{out}}(n)$, 将式(13)中除输出权值的其他项都移到方程右侧, 得到以下方程

$$\hat{\mathbf{W}}^{\text{out}}(n) = \mathbf{P}(n) \boldsymbol{\theta}(n) \quad (17)$$

式中, $\mathbf{P}(n)$ 表示的是自相关矩阵的逆矩阵, 即 $\mathbf{P}(n) = \mathbf{R}^{-1}(n)$ 。将式(11)和式(16)代入式(17)当中, 即可

得到当前时刻的最优输出权值

$$\hat{\mathbf{W}}^{\text{out}}(n) = (1 - \mathbf{k}(n) \mathbf{x}^T(n)) \mathbf{P}(n-1) \boldsymbol{\theta}(n-1) + \gamma(n) \mathbf{k}(n) + \gamma \left(\frac{\lambda-1}{\lambda} \right) (1 - \mathbf{k}(n) \mathbf{x}^T(n)) \mathbf{P}(n-1) \text{sgn}(\hat{\mathbf{W}}^{\text{out}}(n-1)) \quad (18)$$

其中, $\mathbf{k}(n)$ 是增益向量, $\mathbf{k}(n) = \frac{\mathbf{P}(n-1) \mathbf{x}^*(n)}{\lambda + \mathbf{x}^T \mathbf{P}(n-1) \mathbf{x}(n)}$ 。

通过上述的求解过程, L1RLS 方法有效地计算出当前时刻的最优输出权值。可以看出, L1RLS 算法相比于传统的前向选择算法, 利用在线型自适应回归算法实现了网络的在线学习; 并且通过在代价函数中加入输出权值的 L1 约束项, 有效解决了大量无关特征参与训练的问题。L1RLS-ESN 工作步骤可归纳如下:

(1) 初始化 ESN 的相关参数 $\mathbf{W}^{\text{in}}, \mathbf{W}$; 设置储备池维数、稀疏度等相关参数, 并使权值矩阵 $\mathbf{W}$ 的谱半径小于 1, 使网络具有回声状态特性;

(2) 初始化 L1RLS 中的相关参数: 令 $\mathbf{W}^{\text{out}}(0)=0$; 增益矩阵 $\mathbf{k}(n)=0$; 设置遗忘因子 λ ; L1 范数正则化参数 γ ; 随机初始化逆矩阵 $\mathbf{P}(n)=\zeta^{-1} \mathbf{I}$, 其中 ζ 表示的是很小的数(一般取 0.00001), $\mathbf{I}$ 表示单位矩阵;

(3) 将样本数据输入到网络, 根据式(1)计算出 ESN 内部状态矩阵 $\mathbf{x}(n)$, 根据式(2)计算网络输出。根据 L1RLS-ESN 的推导公式, 求出增益向量 $\mathbf{k}(n)$ 、自相关矩阵的逆矩阵 $\mathbf{P}(n)$;

(4) 根据式(18)更新 L1RLS-ESN 的输出权值, 即得到当前时刻的预测模型;

(5) 若有新的样本进入则循环执行步骤(4)、步骤(5), 实现模型的自适应调节; 若无新样本进入, 则停止更新。

2 基于 L1 正则化递归最小二乘法 ESN 的 fCaO 预测控制模型

2.1 控制结构

模型预测控制(MPC)是一种特殊的控制方法, 能够有效解决 DMC 方法中控制精度不高的问题, 其模型结构由预测模型、反馈校正、滚动优化和参考轨迹所构成^[27]。但由于传统 MPC 辨识精度低, 调节时间长, 且无法实时预测熟料 fCaO 含量, 因此在水泥烧成系统中的控制效果较差。本文采用在线型 L1RLS-ESN 模型替代传统 MPC 模型中的预测环节, 保证实时准确地预测熟料 fCaO 含量; 并采用计算复

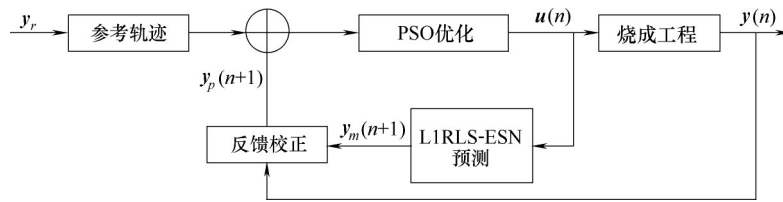


图2 预测控制结构

Fig.2 Structure of predictive control

杂度低、快速搜索到全局最优解的 PSO 算法替代数值解方法进行滚动优化,保证实际被控量输出值在短时间内跟随到设定值,满足控制器的实时性要求。L1RLS-ESN 预测控制算法的流程如图 2 所示。图中 y_r 为被控变量设定值; $y(n)$ 为被控变量实际值; $y_m(n+1)$ 为被控变量预测输出值; $y_p(n+1)$ 为反馈校正输出值; $u(n)$ 为控制变量输入值。

从图 2 可知,控制变量通过 L1RLS-ESN 预测模型和烧成系统,分别得到被控变量预测值和实际值。然后,通过在线的反馈校正环节调整预测输出值,进一步减小预测值和真实值之间的误差。最后,利用预测输出与参考轨迹之间的误差信息,给出带有约束的最优化目标函数,采用 PSO 算法对目标函数进行滚动优化,以给出水泥熟料烧成过程的控制变量修正值,使得系统的实际输出尽可能地接近参考轨迹,达到准确控制 fCaO 含量的目的。

2.2 反馈校正

在水泥烧成过程中,非线性、时变以及各种随机干扰是不可避免的,因此预测模型的输出与实际系统会存在着一定的误差

$$e(n) = y(n) - y_m(n) \quad (19)$$

如果预测输出与实际值偏差较大,控制系统则会出现调节时间长、超调量大甚至发散振荡等现象。为了克服这个误差对系统的影响,利用反馈校正环节对预测输出值进一步进行修正。本文采用的反馈校正公式为

$$y_p(n+1) = y_m(n+1) + \alpha e(n) \quad (20)$$

式中, α 为补偿系数,需根据不同的实际系统进行调节。

2.3 参考轨迹

为降低烧成过程的被控变量设定值 y_{sp} 对预测输出的限定,设定取值为 [0,1] 的柔化系数 C_r 。水泥熟料烧成过程被控变量的参考轨迹可表示为

$$y_r(n+j) = C_r y_r(n) + (1 - C_r) y_{sp}(n) \quad (21)$$

式中, C_r 为柔化系数,能够降低烧成过程的被控变量设定值 y_{sp} 对预测输出的限定,对控制系统的鲁

棒性和动态特性起到重要的作用。如 C_r 的值越大,系统的柔性越高,即系统鲁棒性越高,但是控制的快速性却下降。

2.4 基于 PSO 滚动优化

MPC 是一种基于优化算法的预测控制算法,因此利用 MPC 实现非线性预测控制的另一个关键是要建立有效的滚动优化。模型预估时,在每个采样时刻都要根据当前的预估误差重新计算控制作用的变化量,不断滚动计算,这个过程称为滚动优化^[28]。根据水泥烧成工艺要求,本文以控制变量为寻优目标,使得目标函数最小。其中水泥熟料烧成过程带有约束的最优化目标函数及其约束条件可表示为^[14]

$$\begin{aligned} \min D(n) &= \sum_{i=1}^P [y_r(n+i) - y_p(n+i)] + \\ &\delta \sum_{j=1}^M [u(n+j-1) - u(n+j-2)] \\ \text{s.t. } \Delta u_{\min} &\leq \Delta u(n+i) \leq \Delta u_{\max} \\ u_{\min} &\leq u(n+i) \leq u_{\max}, \quad i = 0, 1, \dots, M-1 \\ y_{\min} &\leq y(n+j) \leq y_{\max}, \quad j = 1, 2, \dots, P \end{aligned} \quad (22)$$

式中, P 为预测时域; M 为控制时域; δ 为控制加权系数,取值范围为 [0,1]; $y_r(n+i)$, $y_p(n+i)$ 分别表示 $k+i$ 时刻参考轨迹的输出和预测模型的输出。

为避免梯度下降法等数值解方法在进行滚动优化时,计算复杂度较高,且不能保证优化结果是全局最优的问题,因此本文考虑采用 PSO 算法进行在线滚动优化,通过求取全局最优解使得式 (22) 目标函数最小,以控制下一时刻的输入量。PSO 是一种非导数优化的群体智能算法,其寻优过程是通过一群粒子在 N 维空间搜索全局最优解的粒子位置而运动。PSO 实现在空间中最优解的搜索,避免数值解方法中复杂的偏微分过程,同时适用于目标函数的一阶导数不存在的情况,不需要多次搜索迭代即可搜索到全局最优解^[29]。假设 N 维空间中 m 个粒子组成一个种群,则每个粒子位置为: $\mathbf{X}_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T$, 粒子相应的速度为 $\mathbf{V}_i = [v_{i1}, v_{i2}, \dots, v_{in}]^T$ 。将这 m 个粒子代入到求解最优化问题的目标函数,求得每个粒

子相对应的适应值,即可求出全局最优解。在第 $n+1$ 次迭代时,种群中粒子 i 的速度和位置分别更新为^[30]

$$\begin{aligned} v_{i,d}^{(n+1)} &= wv_{i,d}^{(n)} + c_1r_1(p_{i,d}^{(n)} - x_{i,d}^{(n)}) + c_2r_2(p_{g,d}^{(n)} - x_{i,d}^{(n)}) \\ x_{i,d}^{(n+1)} &= x_{i,d}^{(n)} + v_{i,d}^{(n+1)} \end{aligned} \quad (23)$$

式中, w 为惯性权重, $v_{i,d}^{(n)}$ 和 $x_{i,d}^{(n)}$ 分别为第 n 次迭代时粒子 i 的第 d 维变量的速度和位置, $p_{i,d}^{(n)}$ 和 $p_{g,d}^{(n)}$ 分别为第 n 次迭代时粒子 i 的第 d 维变量历史最优和全局最优的位置, c_1 和 c_2 分别为自我认知学习因子和社会认知学习因子, r_1 和 r_2 为 $[0,1]$ 区间上均匀分布的随机值。

通过 PSO 寻优, 求出式 (23) 目标函数的最优解, 即为控制变量输入增量值矩阵, 选取控制变量的增量矩阵的第一项作为当前时刻的烧成过程操作变量增量值 $\Delta u(n)$, $\Delta u(n) = [\Delta u_1(n), \Delta u_2(n), \dots, \Delta u_i(n)]^T$ 。

根据该控制变量增量值, L1RLS-ESN 预测控制模型的控制变量输入值可表示为

$$u(n) = u(n-1) + \Delta u(n) \quad (24)$$

根据式 (24) 即可得到当前烧成过程的控制变量输入值, 将该控制变量输入到预测控制中, 循环执

行预测输出值滚动修正过程, 直至被控变量实际值稳定在设定值附近, 即可完成预测控制过程。

本文通过介绍预测模型的选取、滚动优化的寻优过程, 将 L1RLS-ESN 与 PSO 融合到一起构建了水泥烧成过程的预测控制模型。L1RLS-ESN 预测控制模型能够实时预测当前的输出值, 并利用 PSO 滚动优化过程给出控制变量输入值, 保证实际输出量快速、准确、平稳地跟随输出量的设定值改变, 实现 fCaO 含量的实时预测控制。L1RLS-ESN 预测控制模型的工作流程如图 3 所示。

3 算法验证

在水泥生产过程中, 生料的组成成分及其煅烧情况决定了水泥熟料 fCaO 含量。而在新型干法水泥生产过程中, 生料是经过原料磨粉、生料成分配比以及均化库进行均化后, 再进入后续烧成系统进行煅烧, 其成分相对稳定。因此, 水泥熟料 fCaO 含量预测控制模型的输入变量不考虑生料成分变化, 只选取反映生料煅烧情况的烟室氮氧化物含量和烟室氧气含量作为模型的输入, 选取水泥熟料 fCaO 含量作为模型输出^[9]。

3.1 L1RLS-ESN 预测模型验证

为验证本文提出的 L1RLS-ESN 算法建立的水泥熟料 fCaO 含量预测模型的性能, 首先采用某水泥生产线的 265 组实际数据对模型进行训练, 其中水泥熟料 fCaO 含量为化验室测量值。并采用 RESN、L1-ESN、RLS-ESN、L1RLS-ESN 方法作为对比验证方法。选取前 200 组作为训练样本, 其余的 65 组作为测试样本, 熟料 fCaO 含量的仿真对比效果如图 4 所示。

通过图 4 中四种方法的仿真效果图可以知, L1-ESN 模型的预测效果要优于通过引入 L2 惩罚项的 RESN 预测模型, 说明在解决 ESN 过拟合问题上, L1 正则化方法要优于 L2 正则化方法, 但是以上两种算法都基于离线学习方法, 在动态系统中可能会存在与实际值曲线偏差较大的问题。RLS-ESN 和 L1RLS-ESN 作为在线学习算法, 基于 L1RLS-ESN 算法的预测模型跟随真实曲线的能力优于 RLS-ESN 模型, 且拟合程度更高, 证明本文提出的 L1 正则化方法对控制 RLS-ESN 模型规模、提高预测精度是有效的。综上所述, 基于 L1RLS-ESN 算法建立的水泥熟料 fCaO 含量仿真预测值更加靠近实际值曲

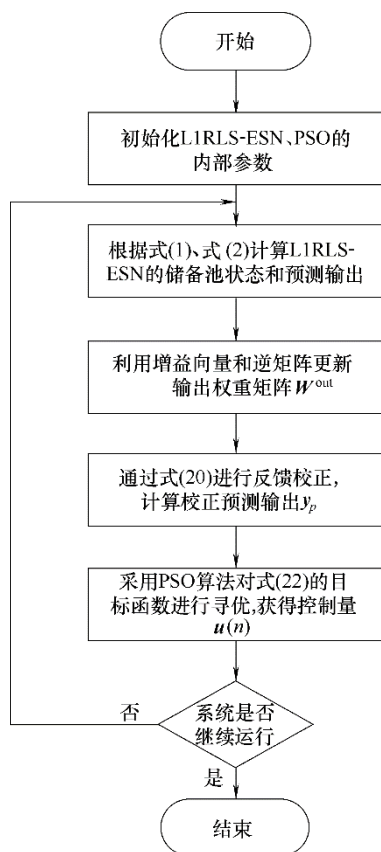


图 3 L1RLS-ESN 预测控制模型工作流程

Fig.3 Workflow chart of L1RLS-ESN predictive control model

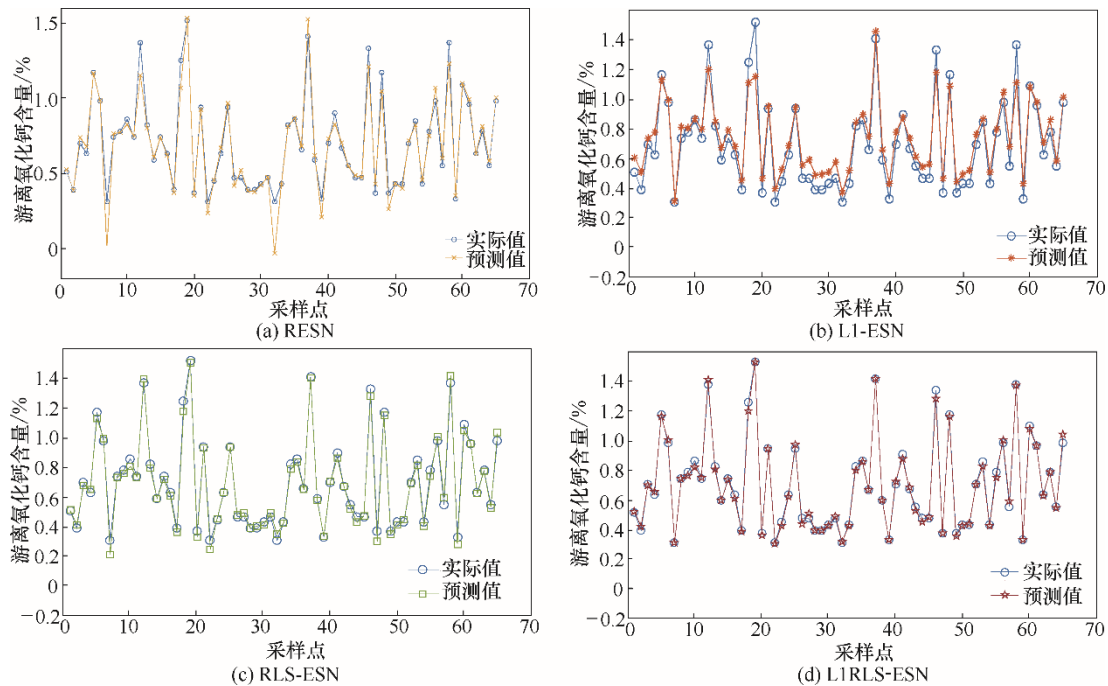


图4 水泥熟料fCaO含量预测仿真结果

Fig.4 Simulation results of fCaO content prediction of cement clinker

线,两者拟合度最高,基于L1RLS-RSN算法建立的水泥熟料fCaO含量预测模型是有效的。

为进一步证明模型的性能,采用均方根误差(root mean square error, RMSE)、平均绝对误差(mean absolute error, MAE)、平均绝对百分误差(mean absolute percentage error, MAPE)、最大绝对误差(maximum absolute error, MAXE)作为性能指标,四种方法建立模型的性能指标如表1所示。

表1 四种方法建立的熟料fCaO预测模型性能统计

Table 1 Performance statistics of clinker fCaO prediction model established by four methods

模型	RMSE	MAPE/%	MAE	MAXE
RESN	0.1852	15.63	0.1032	1.2021
L1-ESN	0.1569	13.21	0.0965	1.1962
RLS-ESN	0.1329	9.54	0.0758	1.0174
L1RLS-ESN	0.1159	8.72	0.0581	0.9834

由表1中模型精度指标数据可以看出,由于离线型L1-ESN、RESN网络不具有记忆功能,两者的误差性能指标值较大;在线型RLS-ESN实现了输出权值的动态调整,提高了算法精度;L1RLS-ESN不仅实现了网络的在线学习,同时控制了储备池规模,剔除了储备池中大量无关特征向量,因此基于L1RLS-ESN算法建立的水泥熟料fCaO含量预测模型的四项误差性能指标值都为最小。

3.2 L1RLS-ESN预测控制模型验证

为测试本文提出的预测控制算法实现水泥熟料烧成过程预测控制的可行性,本文利用日产5000 t水泥熟料生产仿真平台对水泥烧成系统的实际运行状态进行模拟,在线重构实际水泥熟料生产线的各种运行工况^[31]。其中实验仿真平台的实时测量环节是以在线型IQPSO-ELM熟料质量指标软测量仪表为基础^[32]。L1RLS-ESN预测控制模型将软仪表测量值作为实际值,L1RLS-ESN输出值作为预测值,PSO滚动优化环节利用误差信息,实时给出的控制量的修正值,以达到预测控制的目的。其中水泥烧成过程预测控制软件系统如图5所示。

由图5可知,通过设定预测控制软件系统相关参数,将仿真平台中被控变量当前值和设定值通过OPC数据通讯模块传入预测控制软件系统中,经过预测控制软件系统中嵌入的预测控制算法进行计算处理,给出当前控制变量的调整量,再通过OPC数据通讯模块将控制变量的调整量写入水泥熟料生产仿真平台,之后循环将被控变量当前值传入预测控制软件系统中,直到被控变量当前值趋近设定值附近,即可完成水泥熟料烧成过程的仿真平台测试实验。

设定预测控制器的参数取为:采样周期为60 s,预测周期 $P=7$,控制周期 $M=5$,补偿系数 $\alpha=1$,柔化

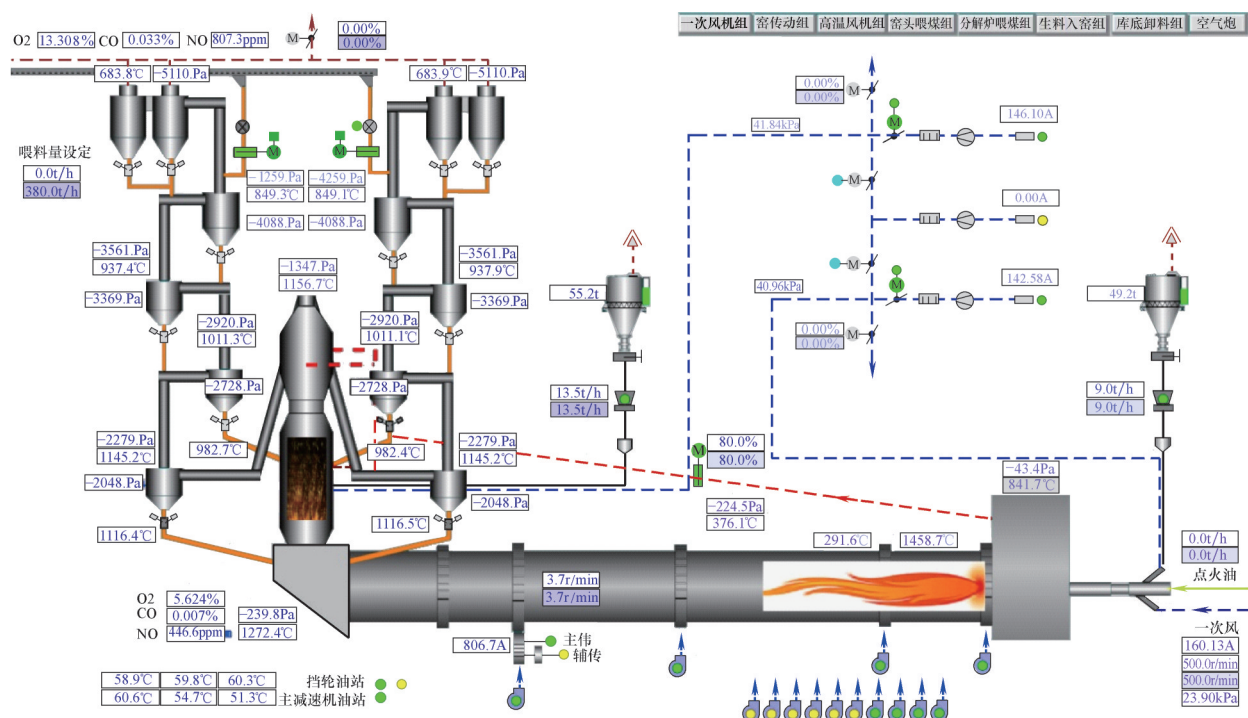


图5 水泥熟料生产仿真平台运行界面

Fig.5 Interface of cement clinker production simulation platform

系数 $C_1=0.5$, 控制加权系 $\delta=0.6$; 设置 PSO 的参数: PSO 算法维度 $N=6$, 粒子群个数为 30, 学习因子 $c_1=c_2=2$, 惯性权重 $w=0.9$, 最大进化代数为 85。为验证 PSO 的收敛性能, 采用训练过程中样本均方差 (MSE) 作为测试指标, 则 PSO 的均方误差曲线如图 6 所示。

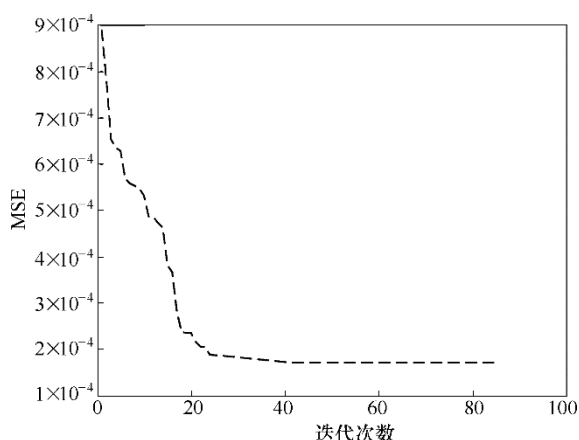


图6 PSO算法均方误差曲线

Fig.6 MSE curve of PSO algorithm

由图 6 可知, PSO 算法经过 41 次迭代, 均方差达到 1.73×10^{-4} 并保持不变。这说明在设定的 85 次迭代次数内, PSO 算法可以达到收敛状态, 且 PSO 的运行时间仅为 3.2 s。因此, PSO 可以在一个采用周期内完成寻优过程, 以确保 L1RLS-ESN 能够快速得到

可靠的预测结果。

在水泥熟料烧成过程中, 实际现场不确定性干扰因素较多, 如燃煤成分、生料成分、预热器塌料、回转窑窑皮脱落等。因此, 通过模拟现场干扰情况对水泥熟料烧成过程进行仿真实验, 以验证控制算法的抗干扰能力。初始时刻被控输出变量设定 $y_{sp}=0$, $k=21, 101, 151, 251$ 时分别阶跃变化到 2%、2.2%、3%、2.5%, 并设定仿真步数 284 时刻之后为干扰条件下的预测控制仿真效果。本文以经典的 DMC 作为对比算法, 预测控制结果和误差值分别如图 7、图 8 所示。在调整输出设定值时, L1RLS-ESN 预测控制模型的烟室氧气含量和氮氧化物含量变化如图 9 所示。

从图 7、图 8 的结果可以看出, 当系统不受外界干扰时, 基于 L1RLS-ESN 神经网络预测控制相比于经典 DMC 预测控制, 能更好地跟踪参考轨迹且输出平稳。而当设定值发生变化时, L1RLS-ESN 模型能够通过调整控制量的值, 快速地跟随被控变量的设定值, 保持了较好的控制品质, 这说明 L1RLS-ESN 具有调节时间短、超调量小、动态特性好的特点。当系统加入外界干扰时, L1RLS-ESN 能够很快恢复稳定, 但 DMC 预测控制模型的跟踪曲线波动较大, 出现了发散现象, 说明 L1RLS-ESN 模型具有一定的抗干扰能力, 鲁棒性较强。从图 9 的结果可以看出,

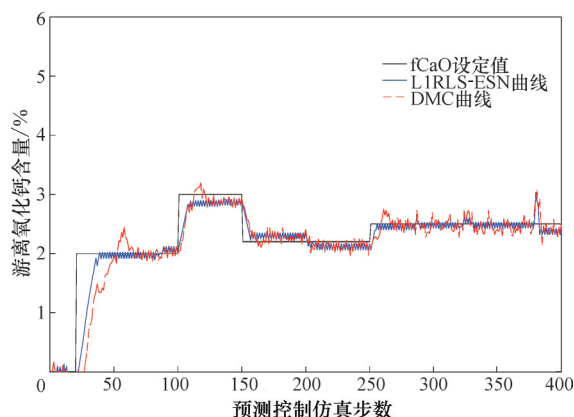


图7 fCaO含量预测控制

Fig.7 Predictive control chart of fCaO content

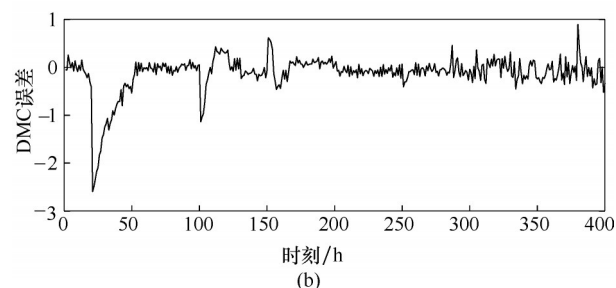
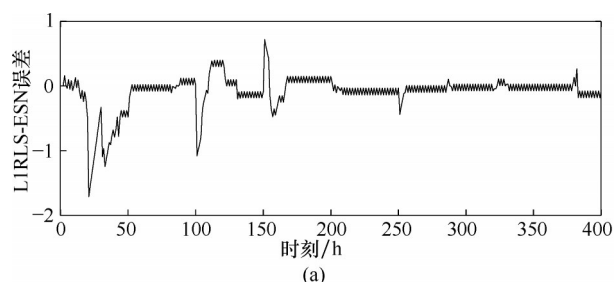


图8 两种预测控制模型误差

Fig.8 Error diagrams of two predictive control models

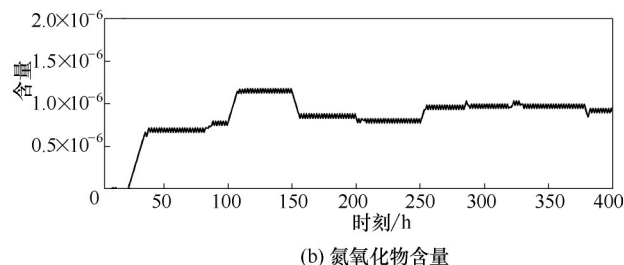
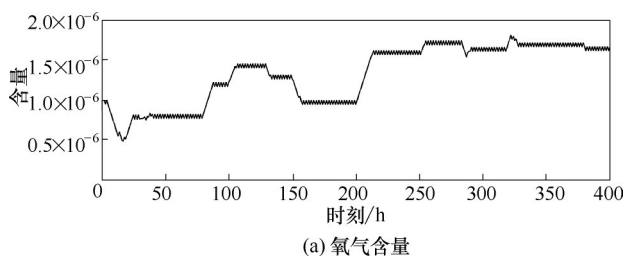


图9 控制量自适应变化

Fig.9 Adaptive variation chart of control volume

在25、100等设定值发生变化的时刻,系统的控制变量自适应发生了变化,说明L1RLS-ESN模型相比于DMC控制方法,具有更好的自适应能力。

综上所述,L1RLS-ESN预测控制算法在对系统实际预测控制的过程中,可以使系统预测输出量快速、准确、平稳地跟随其输出值的设定值;且在受控系统在外界干扰时,预测控制器可以迅速调整输入量,使系统在短时间内恢复到干扰前的平稳状态,说明L1RLS-ESN预测控制器相比于传统预测控制方法,控制性能具有一定程度提高,不仅加快了系统的响应速度,同时还具备更好的动态性能,增强了系统的抗干扰能力,提高了控制系统的鲁棒性。

4 结 论

本文针对非线性时变水泥烧成系统,通过将L1范数与ESN相结合的方法构建在线型L1RLS-ESN预测模型,然后设计基于L1RLS-ESN的预测控制器的网络结构,并采用PSO算法对目标函数进行滚动优化。通过水泥烧成系统的实验验证,得到以下结论。

(1)L1RLS-ESN预测模型实现了实时预测,同时有效控制网络规模,避免网络出现过拟合现象,其预测精度和拟合效果明显优于其他对比算法。

(2)L1RLS-ESN预测控制模型能够保证实际输出量快速、准确、平稳地跟随输出量的设定值,解决了系统超调量高、时间滞后的问题,进而提高了fCaO含量的预测控制精度。

(3)L1RLS-ESN预测控制模型在受到外界干扰的情况下,可以自适应调整输入量,使系统在短时间内恢复到干扰前的平稳状态,证明了算法具有较强的鲁棒性。

符 号 说 明

- C_r —— 柔化系数
- c_1, c_2 —— 分别为自我认知学习因子和社会认知学习因子
- $D(n)$ —— 水泥熟料待优化函数
- $e(n)$ —— 预测值与真实值的误差
- $f(\cdot)$ —— 隐藏层神经元激活函数
- $f^{out}(\cdot)$ —— 输出层神经元激活函数
- $G(n)$ —— L1RLS-ESN的损失函数
- N —— 粒子群维数
- $P(n)$ —— 自相关矩阵的逆矩阵

$p_{i,d}^{(n)}, p_{g,d}^{(n)}$ ——分别是PSO算法搜索到的种群最优解和个体最优解

$R(n), r(n)$ ——分别是自相关矩阵和互相关矩阵

v ——PSO算法的粒子飞行速度

$W^{\text{in}}, W, W^{\text{out}}$ ——分别是回声状态网络输入权值矩阵、储备池权值矩阵和输出权值矩阵

w ——惯性权重

y, y_r, y_m, y_p ——分别表示被控变量实际值、设定值、预测值、反馈校正值

α ——补偿系数

γ ——正则化参数

λ ——遗忘因子

下角标

i ——输入、输出样本所属类别标识

参考文献

- [1] Xiao Y, Zhang C, Wang G, et al. Free CaO stabilisation by mixing of BF slag and BOF slag in molten state[J]. Ironmaking & Steelmaking, 2018, **45**(6): 1–9.
- [2] Kaewmanee K, Krammart P, Sumranwanich T, et al. Effect of free lime content on properties of cement-fly ash mixtures[J]. Construction & Building Materials, 2013, **38**(2): 829–836.
- [3] 赵朋程, 刘彬, 高伟, 等. 用于水泥熟料fCaO预测的多核最小二乘支持向量机模型[J]. 化工学报, 2016, **67**(6): 2480–2487.
Zhao P C, Liu B, Gao W, et al. Multiple kernel least square support vector machine model for prediction of cement clinker lime content[J]. CIESC Journal, 2016, **67**(6): 2480–2487.
- [4] 刘彬, 赵朋程, 高伟, 等. 基于粒子群算法与连续型深度信念网络的水泥熟料游离氧化钙预测[J]. 计量学报, 2018, **174**(3): 420–424.
Liu B, Zhao P C, Gao W, et al. Prediction of cement fCaO based on particle swarm optimization and continuous deep belief network [J]. Acta Metrologica Sinica, 2018, **174**(3): 420–424.
- [5] 张嘉英, 王文兰. 基于动态矩阵控制的再热汽温控制系统[J]. 电力自动化设备, 2010, **30**(8): 71–74.
Zhang J Y, Wang W L. Reheated steam temperature control system based on dynamic matrix control[J]. Electric Power Automation Equipment, 2010, **30**(8): 71–74.
- [6] Garcia C E, Morshedi A M. Quadratic programming solution of dynamic matrix control (QDMC) [J]. Chemical Engineering Communications, 1986, **46**(1/2/3): 73–87.
- [7] Kusiak A, Xu G. Modeling and optimization of HVAC systems using a dynamic neural network[J]. Energy, 2012, **42**(1): 241–250.
- [8] 徐湘元, 毛宗源. 时滞系统的神经网络预测控制[J]. 控制理论与应用, 2001, **18**(6): 932–934.
Xu X Y, Mao Z Y. The neural network predictive control of time-delay systems[J]. Control Theory and Applications, 2001, **18**(6): 932–934.
- [9] 郭丹, 李平, 曹江涛. 基于Elman网络的非线性系统神经元自适应预测控制[J]. 计算机仿真, 2003, **20**(8): 55–57.
Guo D, Li P, Cao J T. An adaptive neuron predictive control based on Elman networks for nonlinear system[J]. Computer Simulation, 2003, **20**(8): 55–57.
- [10] Schrauwen B, Wardermann M, Verstraeten D, et al. Improving reservoirs using intrinsic plasticity[J]. Neurocomputing, 2008, **71**(7): 1159–1171.
- [11] Min H, Xu M. Laplacian echo state network for multivariate time series prediction[J]. IEEE Transactions on Neural Networks & Learning Systems, 2018, **29**(1): 238–244.
- [12] Jaeger H, Maass W, Principe J. Introduction to the special issue on echo state networks and liquid state machines[J]. Neural Networks, 2007, **20**(3): 287–289.
- [13] Jaeger H, Haas H. Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication[J]. Science, 2004, **304**(5667): 78–80.
- [14] 汪灵枝, 吴建生, 吴春梅. 偏最小二乘回归的贝叶斯正则化神经网络集成模型在证券分析预测中的应用[J]. 数学的实践与认识, 2007, **37**(14): 197–205.
Wang L Z, Wu J S, Wu C M. Bayesian regularization neural network ensemble model based on partial least squares regression and its application in stock market[J]. Journal of Mathematics in Practice and Theory, 2007, **37**(14): 197–205.
- [15] 史志伟, 韩敏. ESN岭回归学习算法及混沌时间序列预测[J]. 控制与决策, 2007, **22**(3): 258–261.
Shi Z W, Han M. Ridge regression learning in ESN for chaotic time series prediction[J]. Control and Decision, 2007, **22**(3): 258–261.
- [16] 黄炳家, 王健, 温艳青, 等. 带光滑L1/2正则化项的神经网络逆向迭代算法收敛性分析[J]. 中国石油大学学报(自然科学版), 2015, **39**(2): 164–170.
Huang B J, Wang J, Wen Y Q, et al. Convergence analysis of inverse iterative algorithms for neural networks with L1/2 penalty [J]. Journal of China University of Petroleum (Natural Science Edition), 2015, **39**(2): 164–170.
- [17] 韩敏, 任伟杰, 许美玲. 一种基于 L_1 范数正则化的回声状态网络[J]. 自动化学报, 2014, **40**(11): 2428–2435.
Han M, Ren W J, Xu M L. An improved echo state network via L_1 -norm regularization[J]. Acta Automatica Sinica, 2014, **40**(11): 2428–2435.
- [18] Chuang L, Chang H, Tu C, et al. Improved binary PSO for feature selection using gene expression data[J]. Computational Biology & Chemistry, 2008, **32**(1): 29–38.
- [19] Oca M A M D, Stützle T, Birattari M, et al. Frankenstein's PSO: a composite particle swarm optimization algorithm[J]. IEEE Transactions on Evolutionary Computation, 2009, **13**(5): 1120–1132.
- [20] 柴毅, 周海林, 付东莉, 等. 基于ESN和PSO的非线性模型预测控制[J]. 控制工程, 2011, **18**(6): 864–867.
Chai Y, Zhou H L, Fu D L, et al. Nonlinear model predictive control based on ESN and PSO[J]. Control Engineering, 2011, **18**(6): 864–867.
- [21] 张天瑜. QDPSO滚动优化的LS-SVM预测控制研究[J]. 武汉理工大学学报, 2010, (3): 115–119.
Zhang T Y. Research on LS-SVM predictive controlling using QDPSO rolling optimization[J]. Journal of Wuhan University of Technology, 2010, (3): 115–119.
- [22] 徐正阳, 路志英, 刘洪. 基于泄漏积分型回声状态网络的在线学习光伏功率预测[J]. 电力系统及其自动化学报, 2018, **30**(2): 1–7.
Xu Z Y, Lu Z Y, Liu H. Online-learning PV power forecasting

- based on leaky-integrator echo state network[J]. Proceedings of the CSU-EPSC, 2018, 30(2): 1-7.
- [23] 韩敏, 许美玲, 王新迎. 多元时间序列的子空间回声状态网络预测模型[J]. 计算机学报, 2014, 37(11): 2268-2275.
- Han M, Xu M L, Wang X Y. A multivariate time series prediction model based on subspace echo state network[J]. Chinese Journal of Computers, 2014, 37(11): 2268-2275.
- [24] 周燕萍, 业巧林. 基于L1-范数距离的最小二乘对支持向量机[J]. 计算机科学, 2018, 45(4): 100-105.
- Zhou Y P, Ye Q L, Wang X Y. L1-norm distance based least squares twin support vector machine[J]. Computer Science, 2018, 45(4): 100-105.
- [25] 占美全, 邓志良. 基于L1范数的总变分正则化超分辨率图像重建[J]. 科学技术与工程, 2010, 10(28): 6903-6906.
- Zhan M Q, Deng Z L X Y. L1 norm of total variation regularization based super resolution reconstruction for images[J]. Science Technology and Engineering, 2010, 10(28): 6903-6906.
- [26] Eksioglu, E M. Sparsity regularised recursive least squares adaptive filtering[J]. Iet Signal Processing, 2011, 5(5): 480-487.
- [27] 杨平, 彭道刚, 韩璞, 等. 神经网络预测控制算法及其应用[J]. 控制工程, 2003, 10(4): 349-351.
- Yang P, Peng D G, Han P, et al. Neural networks predictive control algorithm and its application study[J]. Control Engineering, 2003, 10(4): 349-351.
- [28] 高昇, 杨延西, 刘军. 模糊遗传滚动优化的LS-SVM预测控制研究[J]. 系统仿真学报, 2007, 19(6): 1277-1280.
- Gao Y, Yang Y X, Liu J. Research on LS-SVM predictive control using fuzzy genetic algorithm rolling optimization[J]. Journal of System Simulation, 2007, 19(6): 1277-1280.
- [29] 席裕庚, 王凡. 非线性系统预测控制的多模型方法[J]. 自动化学报, 1996, 22(4): 456-461.
- Xi Y G, Wang F. Nonlinear multi-model predictive control[J]. Acta Automatica Sinica, 1996, 22(4): 456-461.
- [30] 任佳, 张益波, 高金凤, 等. 热定型过程能耗建模及PSO参数优化[J]. 化工学报, 2011, 62(8): 2206-2211.
- Ren J, Zhang Y B, Gao J F, et al. Modeling and PSO based parameter optimization of heat-setting process[J]. CIESC Journal, 2011, 62(8): 2206-2211.
- [31] 赵朋程. 水泥熟料fCaO含量预测模型及烧成过程优化控制算法研究[D]. 秦皇岛: 燕山大学, 2018.
- Zhao P C. Research on the free lime content prediction model and optimization control algorithm of cement clinker sintering process [D]. Qinhuangdao: Yanshan University, 2018.
- [32] 赵朋程, 刘彬, 孙超, 等. 基于IQPSO优化ELM的熟料质量指标软测量研究[J]. 仪器仪表学报, 2016, 37(10): 2243-2250.
- Zhao P C, Liu B, Sun C, et al. [J] Soft sensor for cement clinker quality indicator based on IQPSO optimize ELM[J]. Chinese Journal of Scientific Instrument, 2016, 37(10): 2243-2250.