

引用格式:梁秀霞,庞荣荣,郭鹭,等.基于 AMWPSO-LSTM 的多阶段间歇过程故障预测[J].北京化工大学学报(自然科学版),2022,49(2):116-124.

LIANG XiuXia, PANG RongRong, GUO Lu, et al. Multi-stage intermittent process fault prediction based on AMWPSO-LSTM[J]. Journal of Beijing University of Chemical Technology (Natural Science), 2022,49(2):116-124.

# 基于 AMWPSO-LSTM 的多阶段间歇过程故障预测

梁秀霞 庞荣荣 郭鹭 张燕

(河北工业大学 人工智能与数据科学学院,天津 300130)

**摘要:**过程安全对于间歇过程生产具有重要意义,为提高间歇过程生产安全性,提出一种基于改进粒子群算法(AMWPSO)优化长短期记忆网络(LSTM)的间歇过程故障预测模型 AMWPSO-LSTM。针对 LSTM 中的神经元个数、迭代次数、学习率等参数需要人为设置的问题,采用 AMWPSO 对这些参数进行自动寻优。AMWPSO 在原有粒子群优化算法(PSO)中融入了自适应变异和非线性递减惯性权重,提高了 PSO 的参数寻优能力。由于间歇过程具有多阶段性,因此先根据模糊 C 均值聚类(FCM)方法对间歇过程进行阶段划分,再利用 Pearson 相关系数对各阶段实验数据进行相关性分析,以降低系统变量的维数,并建立各阶段  $T^2$  统计量控制限作为系统是否发生故障的指标。实验以青霉素发酵过程数据为例,建立基于 AMWPSO-LSTM 的多阶段故障预测模型,并将该模型的预测结果与基于 LSTM 的多阶段预测模型、基于 PSO-LSTM 的多阶段预测模型的预测结果进行比较,结果表明,基于 AMWPSO-LSTM 的多阶段故障预测模型可取得较高的预测准确度。

**关键词:**粒子群优化;长短期记忆网络;模糊 C 均值聚类;Pearson 相关系数;间歇过程;故障预测

**中图分类号:** TP277 **DOI:** 10.13543/j.bhxbzr.2022.02.014

## 引言

间歇过程作为现代工业生产的一种重要方式,已经广泛应用于染料、食品、制药等众多领域<sup>[1]</sup>。过程安全是间歇过程生产的前提。故障的发生轻则影响生产,重则可能引发事故,威胁人员的生命安全。如果能根据当前的系统工况对系统未来是否会发生故障进行预测,在其还没有对整个系统造成重大危害时,及时发出预警、排除故障,就可以最大程度地减少故障对系统的影响<sup>[2]</sup>。

当前的故障预测技术主要分为两类,即基于物理模型的故障预测方法和基于数据驱动的故障预测方法。基于物理模型的故障预测方法将系统本身的物理特性转换为数学模型,如 Zaidi 等<sup>[3]</sup>使用隐马尔科夫模型对齿轮故障进行预测,具有较高的准确性。但间歇过程物理特性复杂,很难建立可以描述

其本质特性的数学模型。基于数据驱动的故障预测方法包括贝叶斯网络、自回归滑动平均模型(ARMA)等统计学方法和长短期记忆网络(LSTM)、卷积神经网络(CNN)等人工智能方法。高学金等<sup>[4]</sup>采用自回归主元分析(AR-PCA)对间歇过程进行监测,项亚南等<sup>[5]</sup>采用 ARMA 与多向主元分析(MPCA)相结合的方式对间歇过程进行故障预测。但以上传统统计学方法不适用于当下大数据量的训练样本。基于 LSTM、CNN 等人工智能的故障预测方法具有强大的特征提取能力,能够从海量的训练数据中提取关键信息。其中,LSTM 在原有循环神经网络(RNN)结构的基础上,将其隐藏层的神经元替换为具有门控机制的记忆单元,从而使得模型在长序列中有更好的表现。目前,LSTM 已成功应用于股票预测<sup>[6]</sup>、文本处理<sup>[7]</sup>等方面。陈娇娇<sup>[8]</sup>使用 LSTM 模型对间歇过程进行故障预测;李炜等<sup>[9]</sup>采用 LSTM 模型来预测电源车故障,并与其他机器学习方法进行比较,结果表明 LSTM 模型具有更高的预测精度;李京峰等<sup>[10]</sup>将 LSTM 与深度置信网络(DBN)相结合,对航空发动机的剩余寿命进行预测;何群等<sup>[11]</sup>建立 LSTM 模型对齿轮箱的故障状态

收稿日期:2021-06-16

基金项目:天津市企业科技特派员项目(19JCTPJC60300)

第一作者:女,1972年生,教授

E-mail: lxx68@163.com

进行预测,与其他传统方法相比,LSTM 模型的预测性能更好,能够较早地预测故障的发生。

但与其他神经网络类似,LSTM 中的某些参数如神经元个数、迭代次数、学习率等需要人为设置,这些参数通常根据经验获得,大大降低了模型精度<sup>[12]</sup>。因此,本文将融入了自适应变异和非线性递减惯性权重的粒子群优化算法(AMWPSO)与 LSTM 相结合,通过 AMWPSO 对 LSTM 中的参数进行自动寻优,使得基于 LSTM 的故障预测更为准确。并在阶段划分的基础上,提出一种基于 AMWPSO-LSTM 的多阶段间歇过程故障预测模型,之后分别与基于 LSTM 及基于粒子群优化算法(PSO)-LSTM 的多阶段故障预测模型的预测结果相比较,验证了此模型具有更高的预测精度。

## 1 基于 LSTM 的故障预测

青霉素发酵过程是典型的间歇过程,其生产过程数据是一组随时间变化的随机变量,尽管组成时间序列的各个序列值不确定,但整个序列的变化具有一定的规律性<sup>[2]</sup>,而 LSTM 擅于处理长时间序列的预测;同时,青霉素发酵过程受多方面因素的影响,具有复杂的非线性,而 LSTM 能很好地提取时间序列数据的非线性特征。因此,本文采用 LSTM 进行故障预测,通过建立 LSTM 故障预测模型,来预测下一批次青霉素发酵过程是否会发生故障。

青霉素发酵过程数据可以写成  $\mathbf{A}(I \times J \times K)$  的三维数据矩阵,其中  $I$  是批次数量, $J$  是过程变量个数, $K$  是采样节点个数。多向主元分析法将  $\mathbf{A}(I \times J \times K)$  沿着批次分解为二维矩阵  $\mathbf{A}(I \times KJ)$ ,然后按照一般的主元分析方法(PCA)进行分析,建立  $T^2$  统计量控制限。PCA 的具体计算方法可参考文献[13]。

$T^2$  统计量控制限  $T_{\text{UCL}}^2$  可以利用式(1)计算得到。

$$T_{\text{UCL}}^2 \sim \frac{c(I^2 - 1)}{I(I - c)} F_{\alpha}(c, I - c) \quad (1)$$

式中, $c$  是所选主元个数, $F_{\alpha}(c, I - c)$  是检验水平为  $\alpha$ 、自由度为  $c$  和  $I - c$  的  $F$  分布的临界值。

青霉素发酵过程由 Pensim 仿真平台模拟实现。用 Pensim 仿真平台模拟产生 30 批采样时间为 400 h、采样间隔为 1 h 的正常数据和一批在采样时间为 220 ~ 250 h 时引入的斜坡为 0.35% 的充气率故障数据。利用 MPCA 方法沿批次分解正常数据,

建立  $T^2$  统计量控制限。

青霉素发酵过程有多种可在线测量变量,这些变量由系统运行累积得到,包含了系统的动态信息,能很好地表征系统的运行状态。本文将这些变量作为 LSTM 的输入,将  $T^2$  统计量作为 LSTM 的输出,以此构建 LSTM 故障预测模型。通过可在线测量变量的值来预测系统未来  $T^2$  统计量的值,若  $T^2$  统计量预测值在控制限范围内,则正常,否则表示发生故障。

假设青霉素发酵过程时间序列集合为  $X$

$$X = \{X_1, X_2, \dots, X_a, \dots, X_L\} \quad (2)$$

式中, $L$  为输入的样本数量。

序列  $X_a$  表示为

$$X_a = \{x_1, x_2, \dots, x_b\} \quad (3)$$

式中, $x_1, x_2, \dots, x_b$  为多种可在线测量变量。

假设  $t$  时刻用于预测的历史数据长度为  $q$ ,则  $t$  时刻的输入为

$$X'(t) = \{X(t - q + 1), X(t - q + 2), \dots, X(t)\} \quad (4)$$

式中, $t \in [q, L)$ 。数据的输入类似于一个在原始序列上的滑动窗口。

输入数据经过 LSTM 神经网络,得到  $t$  时刻的输出,即下一个采样节点  $T^2$  统计量的值。

## 2 青霉素发酵过程分析

### 2.1 阶段划分

与连续生产过程不同的是,间歇过程包含多个阶段,不同阶段具有不同的主导变量和过程特征<sup>[4]</sup>,因此对整个生产过程建立单一的故障预测模型是不准确的,需要先对间歇过程进行阶段划分,然后分阶段建立故障预测模型。Lu 等<sup>[14]</sup>采用 K-means 方法对间歇过程进行阶段划分,但 K-means 是一种硬性划分方法,具有“非彼即此”的性质。模糊 C 均值聚类(FCM)根据计算出的每个样本对于各个类的隶属度大小来进行聚类分析,相比 K-means 更合理,因此本文采取 FCM 方法对青霉素发酵过程进行阶段划分<sup>[15]</sup>。

将得到的 30 批正常数据  $\mathbf{X}(30 \times 17 \times 400)$  先进行标准化处理,然后将按变量展开后的标准化数据沿采样节点切分为 400 个时间片矩阵,计算出每个时间片矩阵  $\mathbf{T}_e$  与其他时间片矩阵  $\mathbf{T}_g$  的相似度  $S_{eg}$ ,并将相似度矩阵  $\mathbf{S}$  作为 FCM 的输入<sup>[16]</sup>。相似度计算公式为

$$S_{eg} = \sum_{i=1}^I \sum_{j=1}^J (T_{e(ij)} - T_{g(ij)})^2 \quad (5)$$

式中,  $I$  为批次数,  $J$  为变量个数, 在本文中  $I = 30$ ,  $J = 17$ 。

阶段划分结果如图 1 所示。本文将生产过程分为 3 个阶段: 第一阶段, 0 ~ 160 h, 菌体生长期; 第二阶段, 161 ~ 280 h, 青霉素合成期; 第三阶段, 281 ~ 400 h, 青霉素自溶期。

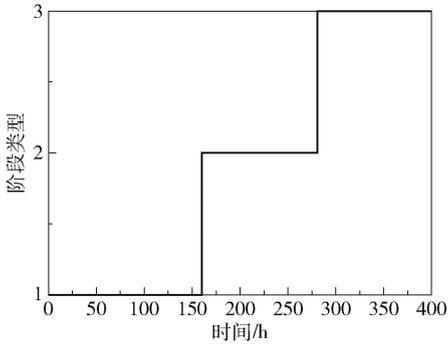


图 1 阶段划分结果图

Fig. 1 Stage division result diagram

按照图 1 的划分结果对样本数据进行阶段划分, 由于菌体生长期各变量不稳定, 因此本文只针对第二阶段和第三阶段进行故障预测。

## 2.2 相关性分析

青霉素发酵过程共有 10 种可在线测量变量。多变量 LSTM 能很好地表征系统的运行状态, 提高系统预测精度。但是如果变量中包含无关变量, 则会大大降低模型的预测精度。因此, 在将这 10 种可在线测量变量作为 LSTM 的输入之前, 先通过 Pearson 相关系数对各阶段的这 10 种变量与  $T^2$  统计量进行相关性分析, 剔除无关变量, 进而提高系统的预测精度。相关性分析结果如表 1 所示, 将  $p < 0.01$  且相关性大于 0.2 的变量作为相应预测模型的输入,  $T^2$  统计量作为输出, 以此来构建第二阶段和第三阶段的 LSTM 故障预测模型。

## 2.3 分阶段建立 $T^2$ 统计量控制限

由于不同阶段的过程特征是变化的, 为了提高系统的准确性, 利用 MPCA 方法沿着批次分解第二阶段样本数据  $A(30 \times 17 \times 120)$ , 建立第二阶段  $T^2$  统计量控制限。同理, 建立第三阶段  $T^2$  统计量控制限。

## 3 AMWPSO-LSTM 多阶段故障预测

与其他神经网络类似, LSTM 中的某些参数需要人为设置, 因此需采用优化算法对 LSTM 中的参

表 1 Pearson 相关性分析结果

Table 1 Pearson correlation analysis results

变量	相关性数值	
	第二阶段 $T^2$	第三阶段 $T^2$
充气率	0.460 **	-0.375 **
搅拌功率	0.209 **	-0.065 **
底物流加速度	0.361 **	0.674 **
底物流温度	0.060 **	-0.426 **
底物浓度	-0.335 **	-0.644 **
溶解氧浓度	-0.348 **	-0.634 **
培养体积	0.261 **	0.814 **
二氧化碳浓度	0.257 **	0.470 **
pH	0.060 **	-0.006
产生的热量	0.363 **	0.805 **

\*\* 表示  $p < 0.01$ 。

数进行自动寻优。通常采用的优化算法有两种: 遗传算法(GA)和粒子群优化算法。有文献将 GA 与 LSTM 相结合对股指进行预测<sup>[17]</sup>。但 GA 的编码过程比较复杂, 因此本文采用 PSO 对 LSTM 中的参数进行自动寻优, 并在原有 PSO 基础上融入自适应变异和非线性递减惯性权重, 本文将其简称为 AMWPSO。

在第 2 节的基础上, 将 AMWPSO 与 LSTM 相结合, 提出一种基于 AMWPSO-LSTM 的多阶段故障预测模型。将 LSTM 的神经元个数、迭代次数、学习率作为 AMWPSO 的优化对象, 其中神经元个数决定模型的拟合能力, 学习率和迭代次数影响模型的训练过程和效果<sup>[18]</sup>。通过 AMWPSO 算法的不断迭代更新, 最终以最优参数构建 LSTM 故障预测模型。

## 3.1 AMWPSO 算法

PSO 的基本原理为在一个  $D$  维空间中,  $m$  个粒子组成一个群体, 每个粒子都有一个速度和位置; PSO 被初始化为一组随机粒子, 通过比较适应度值大小来迭代更新局部最优位置  $p_{best}$  和全局最优位置  $g_{best}$ <sup>[19]</sup>, 进而通过式(6)、(7)来更新粒子的速度和位置, 直至达到最大迭代次数。

$$V_{id}^{k+1} = wV_{id}^k + c_1 rand_1 (p_{best, id}^k - X_{id}^k) + c_2 rand_2 (g_{best, d}^k - X_{id}^k) \quad (6)$$

$$X_{id}^{k+1} = X_{id}^k + V_{id}^{k+1} \quad (7)$$

式中,  $V_{id}^k$ 、 $X_{id}^k$ 、 $p_{best, id}^k$  分别为第  $k$  次迭代时粒子  $i$  的第  $d$  维速度、位置和局部最优位置;  $g_{best, d}^k$  代表第  $k$  次迭代时粒子群的第  $d$  维全局最优位置;  $w$  为惯性权重;

$c_1$  和  $c_2$  为学习因子;  $rand_1$  和  $rand_2$  为  $[0, 1)$  之间的随机值;  $V_{id}^{k+1}$ 、 $X_{id}^{k+1}$  分别为第  $k+1$  次迭代时粒子  $i$  的第  $d$  维速度、位置。

与其他算法相比, PSO 算法结构简单, 便于实现, 但易陷入局部最优解。为了减小粒子陷入局部最优解的风险, 本文在原有 PSO 的基础上融入自适应变异和非线性递减惯性权重, 即 AMWPSO, 以此来提高 PSO 的参数寻优能力, 具体步骤如下。

#### 1) 对 PSO 进行自适应变异

为避免粒子过早收敛, 并提高种群多样性, 对 PSO 进行自适应变异。自适应变异概率公式为

$$rand > \frac{1}{2} \left( 1 + \frac{t}{t_{\max}} \right) \quad (8)$$

式中,  $rand$  函数是一个随机函数, 产生  $[0, 1)$  之间的随机数,  $t$  为迭代次数,  $t_{\max}$  为最大迭代次数。开始时,  $t$  较小,  $rand$  大于不等式右边数值的几率较大; 随着  $t$  的增大, 不等式右边会由 0.5 不断向 1 趋近,  $rand$  大于右边数值的几率则逐渐变小。通过这种方式, 前期变异率大, 后期变异率小, 降低了粒子陷入局部最优解的风险。

#### 2) 对 PSO 的惯性权重进行改进

在基本的 PSO 算法中,  $w$  的值是固定的, 固定的  $w$  会限制 PSO 的全局搜索能力和收敛速度, 因此本文将  $w$  改为如下形式

$$w = w_{\max} - \frac{(w_{\max} - w_{\min})}{(1 + e^{-t/t_{\max}})} \quad (9)$$

式中,  $w_{\max}$  和  $w_{\min}$  分别为  $w$  的最大值和最小值。当  $t$  较小时,  $w$  接近于  $w_{\max}$ , 此时算法的全局搜索能力较强; 随着  $t$  的增大,  $w$  逐渐减小, 此时算法的收敛速度较高。这样前期权重大, 后期权重小, 能够灵活地调整算法的全局搜索能力和收敛速度。

### 3.2 AMWPSO 优化 LSTM 的流程

AMWPSO 优化 LSTM 的流程图如图 2 所示, 具体步骤如下。

#### 1) 获取样本数据, 对样本数据进行阶段划分;

2) 划分训练集和测试集, 将前 30 批正常数据作为训练集, 故障数据作为测试集;

3) 对训练集数据做 Pearson 相关性分析, 根据相关性分析结果对所有样本数据进行筛选, 筛选出相关变量;

4) 对经 Pearson 相关性分析和阶段划分后的样本数据进行标准化处理, 得到标准化后的训练集和标准化后的测试集;

5) 确定 PSO 中  $m$ 、 $t_{\max}$ 、 $c_1$ 、 $c_2$ 、 $w_{\max}$  和  $w_{\min}$  的值, 将 LSTM 的神经元个数、迭代次数、学习率作为 PSO 的优化对象, 确定这 3 个参数的取值范围和搜索范围;

6) 随机初始化粒子的速度和位置, 生成一个初始化的粒子位置  $X_i^0 [h, n, l_r]$ , 得到初始化参数的 LSTM 预测模型, 其中,  $h$  代表神经元个数,  $n$  代表迭代次数,  $l_r$  代表学习率;

7) 将标准化后的训练集输入到初始化参数的 LSTM 预测模型中进行预测得到预测值, 将训练集的均方误差 (MSE) 作为适应度函数, 将预测值代入适应度函数, 计算每个粒子的适应度值;

8) 判断是否对粒子群进行自适应变异, 若是, 返回步骤 6); 若否, 继续下一步;

9) 根据初始粒子适应度值确定初始局部最优位置  $p_{\text{best}}$  和全局最优位置  $g_{\text{best}}$ ; 将每个粒子的最佳位置作为其历史最佳位置, 更新  $p_{\text{best}}$  和  $g_{\text{best}}$ , 并按照式 (6) 和式 (7) 迭代更新每个粒子的速度和位置, 进而更新 LSTM 预测模型的参数;

10) 将标准化后的训练集输入到更新参数后的 LSTM 预测模型中进行预测, 将预测值代入适应度函数中计算每个粒子的适应度值; 转至步骤 8), 直至达到  $t_{\max}$ , 以此确定 LSTM 预测模型的参数值;

11) 将标准化后的测试集代入到确定参数的 LSTM 预测模型中进行预测, 并将预测结果与相应控制限比较, 若超出控制限, 则表示发生故障, 发出预警; 否则表示未发生故障。

## 4 仿真结果与分析

### 4.1 基于 LSTM 的全流程故障预测

利用 MPCA 方法沿批次分解正常数据  $X(30 \times 17 \times 400)$ , 建立  $T^2$  统计量控制限。将 10 种可在线测量变量作为 LSTM 的输入,  $T^2$  统计量作为 LSTM 的输出, 同时将正常数据作为训练集, 故障数据作为测试集, 以此构建基于 LSTM 的全流程故障预测模型。

### 4.2 AMWPSO-LSTM 的参数选择及优化结果

设定 AMWPSO 算法参数, 其中  $m = 20$ ,  $t_{\max} = 20$ ,  $c_1$ 、 $c_2 = 1.5$ ,  $w_{\max} = 0.8$ ,  $w_{\min} = 0.6$ ; 设置 LSTM 参数的取值范围和搜索范围, 其中神经元个数的取值范围和搜索范围均为  $[20, 100]$ , 迭代次数的取值范围和搜索范围均为  $[30, 200]$ , 学习率的取值范围和搜索范围均为  $[0.001, 0.1]$ 。

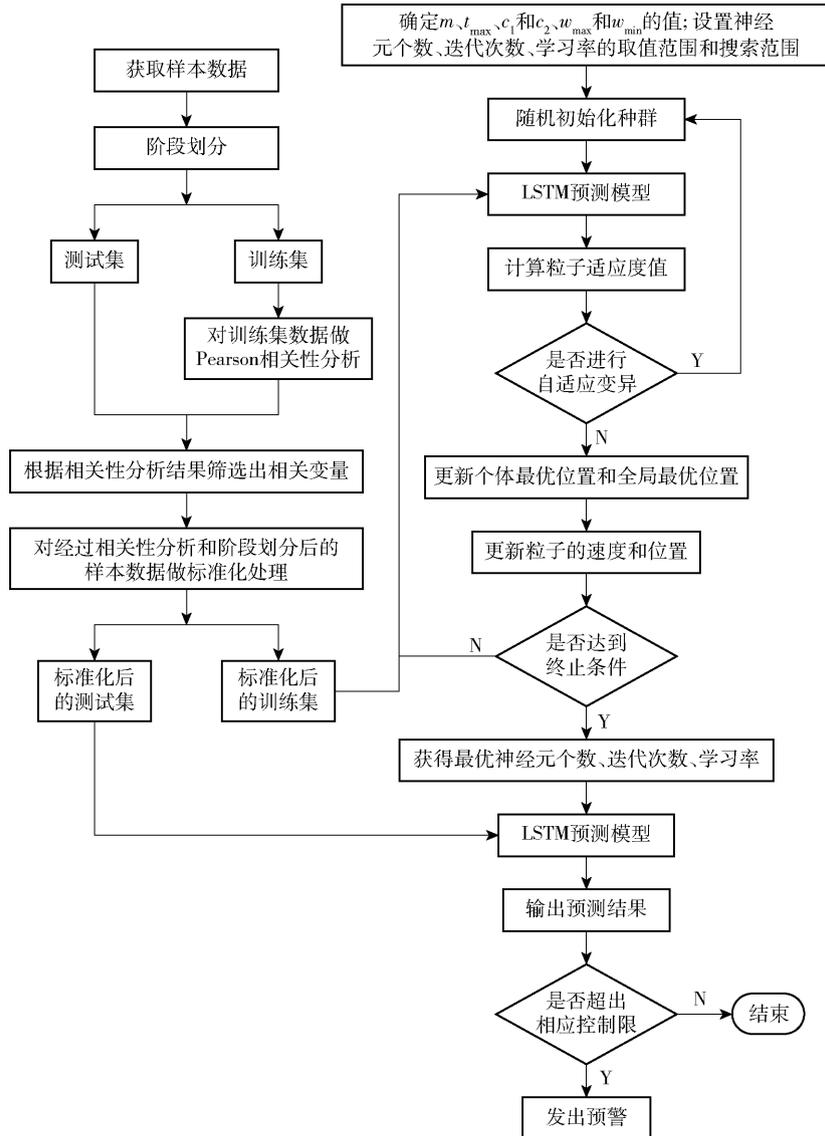


图2 AMWPSO 优化 LSTM 的流程图

Fig. 2 Flow chart of AMWPSO optimizing LSTM

图3(a)为第二阶段 PSO-LSTM 算法的适应度结果,图3(b)为第二阶段 AMWPSO-LSTM 算法的适应度结果。从图中可以看出,PSO 算法约在第3次迭代时就很早地陷入了局部最优解,最终适应度值约为 0.55;AMWPSO 算法增强了寻优能力,在第16次迭代时依然具有寻找更优解的能力,最终适应度值约为 0.51,表明 AMWPSO 的寻优能力更强。

经训练后,确定第二阶段 LSTM 的参数为 [65, 124, 0.024], 第三阶段 LSTM 的参数为 [76, 94, 0.032], 以此分别构建第二阶段和第三阶段基于 AMWPSO-LSTM 的故障预测模型。

### 4.3 仿真分析

图4是基于 LSTM 的全流程故障预测结果,即

未对发酵过程进行阶段划分的故障预测结果。图5(a)、(b)分别为第二阶段、第三阶段基于 LSTM 的故障预测结果。图6(a)、(b)分别为第二阶段、第三阶段基于 PSO-LSTM 的故障预测结果。图7(a)、(b)分别为第二阶段、第三阶段基于 AMWPSO-LSTM 的故障预测结果。

从图4可以看出,基于 LSTM 的全流程故障预测模型的误差较大。在 0~150 h 阶段甚至出现了误报的情况,此时并没有加入扰动,但无论是  $T^2$  的真实值还是预测值均超出控制限,这是因为此阶段为菌体生长期,各个变量是不稳定的。由此说明采用 LSTM 对整个生产过程进行预测是不准确的。从图4、5的对比可以看出,与基于 LSTM 的全流程故

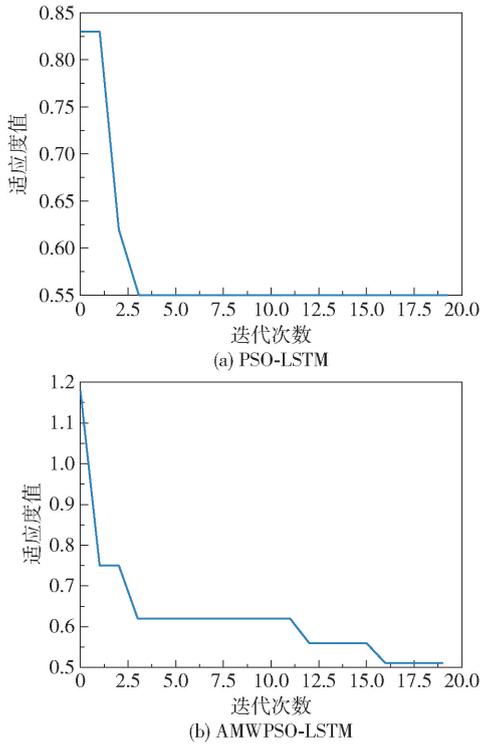


图 3 第二阶段各模型的适应度结果

Fig. 3 The fitness results of each model in the second stage

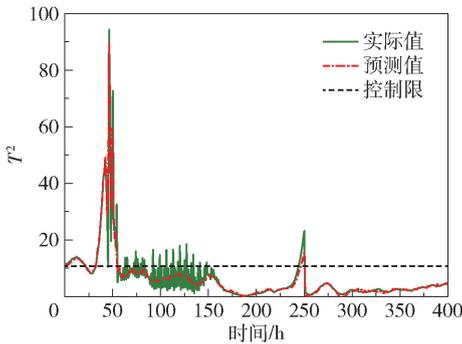


图 4 基于 LSTM 的全流程故障预测结果

Fig. 4 Fault prediction results for the LSTM-based full-process

障预测模型相比,基于 LSTM 的多阶段故障预测模型的拟合效果更好,且滞后时间更短;同时,由于对不同阶段进行划分,避免了出现误报的情况。

从图 5 ~ 7 可以看出,基于 LSTM 的多阶段故障预测模型的预测值曲线在趋势上与真实值非常符合,然而预测值与真实值之间总保持有一定误差,且基于 LSTM 的故障预测模型在第二阶段峰值处的误差最大。这是由于 LSTM 的参数都是通过经验获得的,并不是最优参数。基于 PSO-LSTM 的多阶段故障预测模型相比基于 LSTM 的多阶段故障预测模型,准确度有所提升,但是其在第二阶段峰值处的误

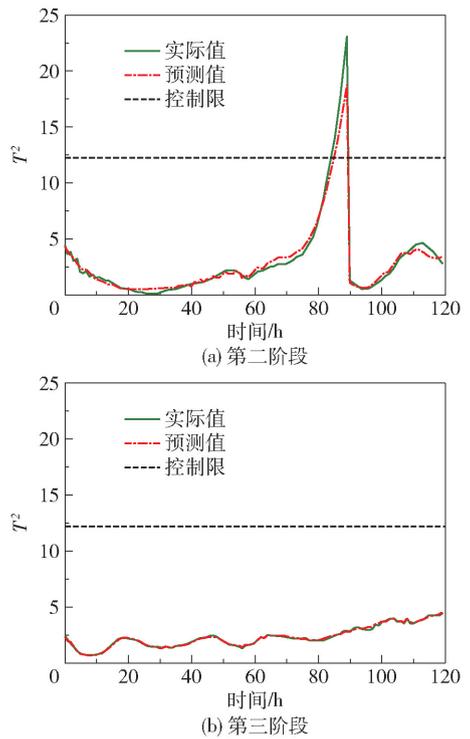


图 5 基于 LSTM 的多阶段故障预测结果

Fig. 5 Results of multi-stage fault prediction based on LSTM

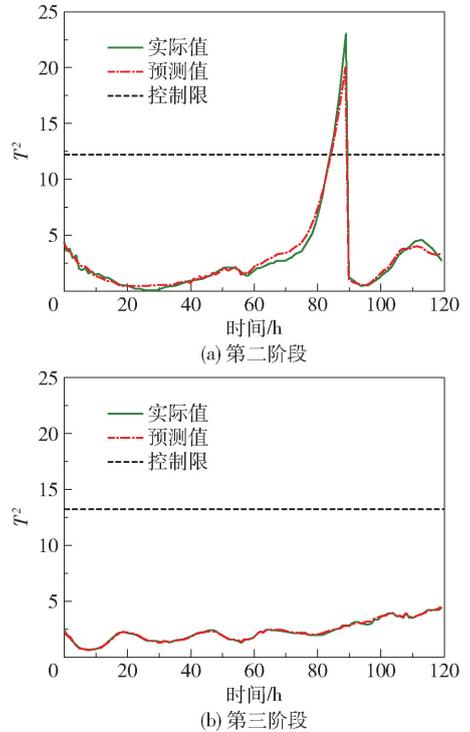


图 6 基于 PSO-LSTM 的多阶段故障预测结果

Fig. 6 Results of multi-stage fault prediction based on PSO-LSTM

差依然较大,并且滞后时间较长。而基于 AMWPSO-LSTM 的多阶段故障预测模型的拟合效

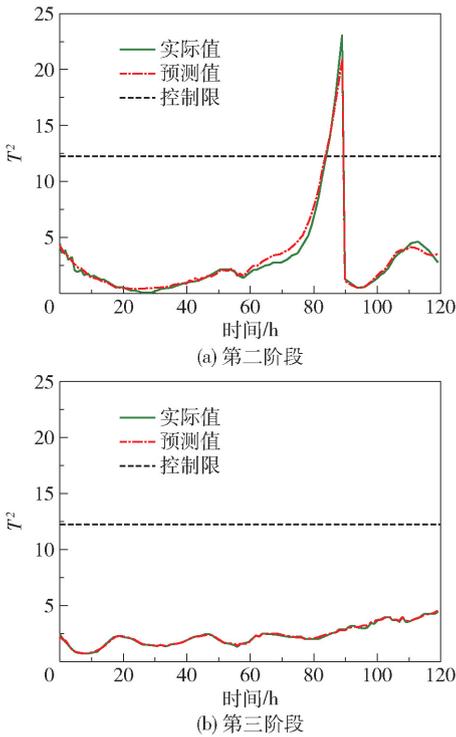


图7 基于 AMWPSO-LSTM 的多阶段故障预测结果

Fig.7 Results of multi-stage fault prediction based on AMWPSO-LSTM

果比前两个模型都要好,同时滞后时间短,能较准确地预测故障的发生时间,并且其在第二阶段峰值处的误差最小。以上结果验证了基于 AMWPSO-LSTM 的多阶段故障预测模型的有效性,同时也验证了利用 AMWPSO 优化 LSTM 的必要性。

本文以平均绝对百分比误差 (MAPE)、均方根误差 (RMSE)、平均绝对值误差 (MAE)、希尔不等系数 (TIC) 和决定系数  $R^2$  作为模型的评价指标,来进一步验证基于 AMWPSO-LSTM 的多阶段故障预测模型的预测性能。表 2、表 3 分别给出了第二阶段、第三阶段各故障预测模型的评价指标计算结果。由表 2 的对比分析可知,与基于 LSTM、PSO-LSTM 的故障预测模型相比,基于 AMWPSO-LSTM 的故障预测模型的 MAPE 分别降低了 59%、33.6%, RMSE 分别降低了 51.8%、32%, MAE 分别降低了 29.3%、15.9%, TIC 分别降低了 54.7%、32.4%,  $R^2$  则分别提升了 3.8%、1.3%。由表 3 的对比分析可知,与基于 LSTM、PSO-LSTM 的故障预测模型相比,基于 AMWPSO-LSTM 的故障预测模型的 MAPE 分别降低了 28.2%、11.3%, RMSE 分别降低了 66.6%、62.1%, MAE 分别降低了 18.3%、8.1%, TIC 分别降低了 18.5%、6.9%,  $R^2$  分别提升了

0.5%、0.1%。

表 2 第二阶段各模型的预测误差对比

Table 2 Comparison of prediction errors of each model in the second stage

模型	MAPE/%	RMSE	MAE	TIC	$R^2$
LSTM	42.73	0.8219	0.3978	0.0908	0.9534
PSO-LSTM	26.4	0.5822	0.3342	0.0608	0.9766
AMWPSO-LSTM	17.52	0.3958	0.2811	0.0411	0.9892

表 3 第三阶段各模型的预测误差对比

Table 3 Comparison of prediction errors of each model in the third stage

模型	MAPE/%	RMSE	MAE	TIC	$R^2$
LSTM	4.50	0.1061	0.0864	0.0216	0.9864
PSO-LSTM	3.99	0.0934	0.0768	0.0189	0.9895
AMWPSO-LSTM	3.54	0.0354	0.0706	0.0176	0.9909

评价指标中 MAPE、RMSE、MAE、TIC 反映预测误差,它们的值越小越好; $R^2$  反映模型的拟合能力,其值越大越好。由表 2、3 的结果可以看出,基于 PSO-LSTM 的故障预测模型和基于 AMWPSO-LSTM 的故障预测模型的预测误差均比 LSTM 故障预测模型的小,同时拟合效果更好,这体现出对 LSTM 进行参数优化的必要性。而基于 AMWPSO-LSTM 的故障预测模型比基于 PSO-LSTM 的故障预测模型的各项表现都更优,这体现出对 PSO 算法进行改进的必要性。与基于 LSTM 的故障预测模型相比,基于 PSO-LSTM 的故障预测模型的预测误差准确性提升较小,这是因为 PSO-LSTM 模型虽拥有与 AMWPSO-LSTM 模型较为相近的网络结构,而 AMWPSO-LSTM 模型的突出优势是其对 PSO 算法进行了改进,减小了 PSO 算法陷入局部最优解的风险,所以该模型的预测结果比普通的 PSO-LSTM 故障预测模型的预测效果更好。

## 5 结论

(1) 针对 LSTM 的参数需要人为设置的问题,本文通过改进粒子群算法 AMWPSO 对 LSTM 的参数进行自动寻优。AMWPSO 在传统 PSO 的基础上融入了自适应变异和非线性递减惯性权重,以此来减小传统 PSO 算法易陷入局部最优解的风险。针对间歇过程的多阶段性,采取 FCM 对间歇过程进行阶段划分;再在阶段划分的基础上,利用 Pearson 相关

系数法对各阶段数据集进行相关性分析,剔除无关变量,分阶段建立故障预测模型。

(2)将基于 LSTM、PSO-LSTM、AMWPSO-LSTM 的多阶段故障预测模型进行比较,结果表明,基于 AMWPSO-LSTM 的多阶段故障预测模型比其他两类模型的预测误差更小,同时拟合效果也更好。这体现出对 LSTM 进行参数优化和对 PSO 算法进行改进的必要性。

### 参考文献:

[1] 王立敏,杨继胜,于晶贤,等. 基于 T-S 模糊模型的间歇过程的迭代学习容错控制[J]. 化工学报, 2017, 68(3): 1081-1089.  
WANG L M, YANG J S, YU J X, et al. Iterative learning fault-tolerant control for batch processes based on T-S fuzzy model[J]. CIESC Journal, 2017, 68(3): 1081-1089. (in Chinese)

[2] 王姝. 基于数据的间歇过程故障诊断及预测方法研究[D]. 沈阳: 东北大学, 2010.  
WANG S. Data-based fault diagnosis and prediction for batch processes[D]. Shenyang: Northeastern University, 2010. (in Chinese)

[3] ZAIDI S S H, AVIYENTE S, SALMAN M, et al. Prognosis of gear failures in DC starter motors using hidden Markov models[J]. IEEE Transactions on Industrial Electronics, 2011, 58(5): 1695-1706.

[4] 高学金,黄梦丹,齐咏生,等. PDPPO 优化多阶段 AR-PCA 间歇过程监测方法[J]. 化工学报, 2018, 69(9): 3914-3923.  
GAO X J, HUANG M D, QI Y S, et al. Batch process monitoring using multiphase AR-PCA optimized with PDPPO[J]. CIESC Journal, 2018, 69(9): 3914-3923. (in Chinese)

[5] 项亚南,潘丰. 基于多元统计时间序列模型的间歇过程故障预测方法研究[J]. 仪表技术与传感器, 2015(6): 135-137, 142.  
XIANG Y N, PAN F. Batch processes fault method research based on multivariate statistical analysis time series model[J]. Instrument Technique and Sensor, 2015(6): 135-137, 142. (in Chinese)

[6] 裴大卫,朱明. 基于多因子与多变量长短期记忆网络的股票价格预测[J]. 计算机系统应用, 2019, 28(8): 30-38.  
PEI D W, ZHU M. Stock price prediction based on multiple-factor and multi-variable long short-term memory [J]. Computer Systems & Applications, 2019, 28(8):

30-38. (in Chinese)

[7] CAI M, LIU J. Maxout neurons for deep convolutional and LSTM neural networks in speech recognition [J]. Speech Communication, 2016, 77: 53-64.

[8] 陈娇娇. 多阶段间歇过程的故障检测与预测研究[D]. 天津: 河北工业大学, 2020.  
CHEN J J. Research on fault detection and prediction of multi-stage batch processes [D]. Tianjin: Hebei University of Technology, 2020. (in Chinese)

[9] 李炜,周丙相,蒋栋年,等. 基于多状态时间序列预测学习的电车车故障预测方法[J]. 吉林大学学报(工学版), 2020, 50(4): 1532-1544.  
LI W, ZHOU B X, JIANG D N, et al. Fault prediction of power supply vehicle based on multi-state time series prediction learning [J]. Journal of Jilin University (Engineering and Technology Edition), 2020, 50(4): 1532-1544. (in Chinese)

[10] 李京峰,陈云翔,项华春,等. 基于 LSTM-DBN 的航空发动机剩余寿命预测[J]. 系统工程与电子技术, 2020, 42(7): 1637-1644.  
LI J F, CHEN Y X, XIANG H C, et al. Remaining useful life prediction for aircraft engine based on LSTM-DBN [J]. Systems Engineering and Electronics, 2020, 42(7): 1637-1644. (in Chinese)

[11] 何群,尹飞飞,武鑫,等. 基于长短期记忆网络的风电机组齿轮箱故障预测[J]. 计量学报, 2020, 41(10): 1284-1290.  
HE Q, YIN F F, WU X, et al. Fault prediction of wind turbine gearbox based on long short-term memory network [J]. Acta Metrologica Sinica, 2020, 41(10): 1284-1290. (in Chinese)

[12] ZHANG Y S, YANG S T. Prediction on the highest price of the stock based on PSO-LSTM neural network [C] // 2019 3rd International Conference on Electronic Information Technology and Computer Engineering (EITCE). Xiamen, 2019: 1565-1569.

[13] 项亚南. 基于 MPCA 的工业间歇过程故障检测方法改进研究[D]. 无锡: 江南大学, 2015.  
XIANG Y N. Research on the improvement of fault detection method of industrial batch process based on MPCA [D]. Wuxi: Jiangnan University, 2015. (in Chinese)

[14] LU N Y, GAO F R, WANG F L. Sub-PCA modeling and on-line monitoring strategy for batch processes [J]. AIChE Journal, 2004, 50(1): 255-259.

[15] KARI T, GAO W S. Power transformer fault diagnosis using FCM and improved PCA [J]. The Journal of Engineering, 2017, 2017(14): 2605-2608.

- [16] 梁秀霞, 陈娇娇, 严婷, 等. 基于自适应动态时间规整(DTW)的 GA-FCM 多阶段间歇过程故障诊断[J]. 北京化工大学学报(自然科学版), 2019, 46(5): 87-93.  
LIANG X X, CHEN J J, YAN T, et al. Fault diagnosis of GA-FCM multi-stage batch processes based on adaptive dynamic time warping algorithm[J]. Journal of Beijing University of Chemical Technology (Natural Science), 2019, 46(5): 87-93. (in Chinese)
- [17] 武大硕, 张传雷, 陈佳, 等. 基于遗传算法改进 LSTM 神经网络股指预测分析[J]. 计算机应用研究, 2020, 37(S1): 86-87, 107.  
WU D S, ZHANG C L, CHEN J, et al. Improved LSTM neural network stock index forecast analysis based on genetic algorithm[J]. Application Research of Computers, 2020, 37(S1): 86-87, 107. (in Chinese)
- [18] 李万, 冯芬玲, 蒋琦玮. 改进粒子群算法优化 LSTM 神经网络的铁路客运量预测[J]. 铁道科学与工程学报, 2018, 15(12): 3274-3280.  
LI W, FENG F L, JIANG Q W. Prediction for railway passenger volume based on modified PSO optimized LSTM neural network[J]. Journal of Railway Science and Engineering, 2018, 15(12): 3274-3280. (in Chinese)
- [19] 阳春华, 谷丽珊, 桂卫华. 自适应变异的粒子群优化算法[J]. 计算机工程, 2008, 34(16): 188-190.  
YANG C H, GU L S, GUI W H. Particle swarm optimization algorithm with adaptive mutation[J]. Computer Engineering, 2008, 34(16): 188-190. (in Chinese)

## Multi-stage intermittent process fault prediction based on AMWPSO-LSTM

LIANG XiuXia PANG RongRong GUO Lu ZHANG Yan

(School of Artificial Intelligence and Data Science, Hebei University of Technology, Tianjin 300130, China)

**Abstract:** Process safety is of great significance in batch process production. In order to improve the production safety of batch processing, a fault prediction model AMWPSO-LSTM based on improved particle swarm optimization (AMWPSO) optimized long short-term memory network (LSTM) has been proposed. Given that parameters such as the number of neurons, the number of iterations, and the learning rate in LSTM need to be set manually, AMWPSO is used to optimize these parameters automatically. AMWPSO integrates adaptive mutation and nonlinear decreasing inertia weight into the original particle swarm optimization algorithm (PSO), which improves the parameter optimization ability of PSO. Due to the multi-stage nature of the batch process, it is first divided into stages according to the fuzzy C-means clustering (FCM) method. The Pearson correlation coefficient is then used to carry out correlation analysis on the experimental data for each stage to reduce the dimension of the system variables. The  $T^2$  statistical control limit for each stage is taken as an indicator of whether the system fails. Using experimental data for penicillin fermentation as an example, we have established a multi-stage fault prediction model based on AMWPSO-LSTM and compared the predictions of the model with the predictions of the multi-stage prediction model based on LSTM and the multi-stage test model based on PSO-LSTM. The results show that the multi-stage fault prediction model based on AMWPSO-LSTM improves the prediction accuracy.

**Key words:** particle swarm optimization; long short term memory; fuzzy C-means clustering; Pearson correlation coefficient; batch process; fault prediction

(责任编辑:吴万玲)