

# 控制与决策

Control and Decision

## 基于注意力LSTM的多阶段发酵过程集成质量预测

高学金, 孟令军, 高慧慧

引用本文:

高学金, 孟令军, 高慧慧. 基于注意力LSTM的多阶段发酵过程集成质量预测[J]. *控制与决策*, 2022, 37(3): 616–624.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2020.1538>

---

## 您可能感兴趣的其他文章

Articles you may be interested in

### 基于批次图像化的卷积自编码故障监测方法

Fault detection of batch image-based convolutional autoencoder

*控制与决策*. 2021, 36(6): 1361–1367 <https://doi.org/10.13195/j.kzyjc.2019.1342>

### 带不相关并行机和有限缓冲MHFS调度的混合启发式算法

Hybrid heuristic algorithm for multi-stage hybrid flow shop scheduling with unrelated parallel machines and finite buffers

*控制与决策*. 2021, 36(3): 565–576 <https://doi.org/10.13195/j.kzyjc.2019.0835>

### 基于局部信息增量与MPLS的质量相关故障检测方法

Quality-related fault detection method based on local information increment and MPLS

*控制与决策*. 2021, 36(7): 1647–1654 <https://doi.org/10.13195/j.kzyjc.2019.1402>

### 基于双权重多邻域保持嵌入的间歇过程故障检测

Fault detection of batch process based on double weight and multiple neighborhoods preserving embedding

*控制与决策*. 2021, 36(12): 3023–3030 <https://doi.org/10.13195/j.kzyjc.2020.0659>

### 基于偏态深度分类的高炉硅含量及波动预测

Prediction of blast furnace silicon content and fluctuation based on skewness depth classification

*控制与决策*. 2021, 36(2): 491–497 <https://doi.org/10.13195/j.kzyjc.2019.1116>

# 基于注意力LSTM的多阶段发酵过程集成质量预测

高学金<sup>†</sup>, 孟令军, 高慧慧

(1. 北京工业大学 信息学部, 北京 100124; 2. 北京工业大学 数字社区教育部工程研究中心, 北京 100124; 3. 北京工业大学 城市轨道交通北京实验室, 北京 100124; 4. 北京工业大学 计算智能与智能系统北京市重点实验室, 北京 100124)

**摘要:** 考虑到发酵过程的动态特征对阶段划分的影响, 为提高模型预测精度, 提出一种基于注意力 LSTM 的多阶段发酵过程质量预测方法. 首先, 将原始三维数据沿批次展开, 对每个时间片矩阵进行偏最小二乘(PLS)分析得到表征过程变量的得分矩阵和表征质量变量的得分矩阵, 采用仿射传播(AP)聚类算法将联合得分矩阵进行聚类, 实现第 1 步划分; 然后, 采用 encoder-decoder 模型将表征过程动态性的动态特征提取出来, 采用 AP 算法对其进行第 2 步划分; 最后, 综合分析两步划分结果, 将生产过程划分为不同的稳定阶段和过渡阶段, 对划分后的各个阶段分别建立注意力长短期记忆(LSTM)集成质量预测模型. 将该方法应用到青霉素发酵仿真数据和大肠杆菌实际生产数据进行验证, 结果表明了所提出方法的可行性和有效性.

**关键词:** 发酵过程; 多阶段; 偏最小二乘; 动态性; 过渡; 质量预测

**中图分类号:** TP277      **文献标志码:** A

**DOI:** 10.13195/j.kzyjc.2020.1538

**引用格式:** 高学金, 孟令军, 高慧慧. 基于注意力 LSTM 的多阶段发酵过程集成质量预测 [J]. 控制与决策, 2022, 37(3): 616-624.

## Integrated quality prediction of multi-stage fermentation process based on attention LSTM

GAO Xue-jin<sup>†</sup>, MENG Ling-jun, GAO Hui-hui

(1. Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China; 2. Engineering Research Center of Digital Community of Ministry of Education, Beijing University of Technology, Beijing 100124, China; 3. Beijing Laboratory for Urban Mass Transit, Beijing University of Technology, Beijing 100124, China; 4. Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing University of Technology, Beijing 100124, China)

**Abstract:** In order to consider the impact of dynamic features of the fermentation process on stage division and improve the prediction accuracy, a quality prediction method based on attention long short-term memory (LSTM) is proposed. Firstly, the original 3D data are unfolded along the batch direction. Partial least square (PLS) analysis is performed on each time slice matrix to obtain the score matrix of process variables and quality variables. The joint score matrices are clustered using the affinity propagation (AP) algorithm. Then the encoder-decoder model is used to extract the dynamic characteristics of the process dynamics, and the AP algorithm is used for the second division. Finally, the production process is divided into different stable phases and transition phases through the comprehensive analysis of the two-step division results. The LSTM integrated quality prediction model is established in each stage after the division. Penicillin fermentation simulation data and *E. coli* production data are tested, and the results demonstrate the feasibility and effectiveness of the proposed method.

**Keywords:** fermentation process; multi-stage; partial least squares; dynamic; transition; quality prediction

## 0 引言

间歇过程作为现代工业过程中极其重要的一种

生产方式, 已经被广泛应用于医药、食品、生物化工、半导体等多品种、高附加值产品的生产中<sup>[1]</sup>. 然而,

收稿日期: 2020-11-17; 录用日期: 2021-01-15.

基金项目: 国家自然科学基金项目(61803005, 61640312, 61763037); 北京市自然科学基金项目(4192011, 4172007); 山东省重点研发计划项目(2018CXGC0608).

责任编辑: 毛志忠.

<sup>†</sup>通讯作者. E-mail: gaoxuejin@bjut.edu.cn.

在某些生产如发酵过程中,质量变量难以实现在线测量. 手工离线测量不仅操作繁琐耗时,给人带来了额外的工作量,而且在取样过程中万一操作不慎,很可能使整个发酵罐染上杂菌,这对生产来说带来了极大浪费. 因此,为提高产品质量和工艺效率,在线质量预测的重要性日益突出<sup>[2-4]</sup>.

近些年来,随着分布式控制系统(DCS)及智能仪表的广泛使用,大量过程数据被采集、记录<sup>[5]</sup>,因此一些基于数据驱动的方法在间歇过程建模、监测和质量预测中得到了广泛应用<sup>[6-11]</sup>. 文献[12]提出的偏最小二乘(partial least square, PLS)是一个广为人知并加以利用的有效工具,但是该方法假设数据是线性的. 为了解决这个问题,文献[13-14]将核技巧与偏最小二乘方法相结合,提出了核偏最小二乘(kernel partial least squares, KPLS)方法. 但是传统多元回归建模方法较为简单,在如今的大数据背景下面对数据量大、价值密度低等问题处理能力远远不足<sup>[15]</sup>. 近年来,神经网络由于强大的非线性信息学习能力,在工业生产中的重要变量预测方面应用较多. 文献[16]利用深度置信网络(DBN)为原油蒸馏装置建立软测量模型. 文献[17]将降噪自编码器与另一个神经网络相结合来改善模型预测性能,提升了模型的鲁棒性. 文献[18]使用分层极限学习机实现半监督软测量的建模. 但是以上方法大多属于静态建模方法,即假设过程处于稳定运行状态,变量之间不存在序列相关性. 而对于大多数实际工业过程而言,过程观测变量呈现出动态特性,并不满足上述假设. 最近,循环神经网络(RNN)在处理时间序列预测时体现出了极大优势<sup>[19]</sup>,传统RNN虽然适用于处理非线性时间序列,但当时间步长较大,网络的误差梯度沿时间反向传播会存在梯度消失与梯度爆炸的问题,难以学习到较为久远的序列信息. LSTM(long short-term memory)网络是由Hochreiter等<sup>[20]</sup>提出的一种改进RNN,其引入内在记忆的线性自循环结构,具有使产生梯度长时间持续流动的路径,其中导数既不消失也不发生爆炸.

多阶段特性是发酵过程中的一个典型特性<sup>[21-22]</sup>,显然对整个生产过程建立单一模型是不完善的,因此许多国内外专家对阶段划分进行了大量研究. 文献[23]提出基于高斯混合模型(Gaussian mixture model, GMM)的阶段划分方法;文献[24]提出了一种线性局部模型逼近方法,实现阶段划分的目的;文献[25]提出了基于K-means的间歇过程子时段划分方法. 但是,上述这些方法属于硬分类方法,忽略了两阶段间的过渡信息. 相比各个稳定阶段的运行状态,阶

段之间的过渡虽然不代表主流的过程操作机理,但却是一种普遍的现象及重要的过程行为,表现为一种动态的渐变趋势. 在此基础上,文献[26]引入模糊隶属度作为两个相邻稳定阶段的权重系数,然后使用K-means进行阶段划分;文献[27]采用模糊C均值聚类(fuzzy C-means clustering, FCM)对时间片进行阶段划分. 但是,上述方法没有考虑到质量变量对阶段划分的影响,而且将每个采样时刻单独看成一个单元进行分析,忽略了不同时刻前后的联系,因此对过程动态特征变化的捕捉不够灵敏. 而发酵过程的多阶段特性很大程度上由表征过程变化的动态性体现出来,在不同阶段发酵过程具有不同的动态特征.

针对以上问题,本文提出一种基于动态特征提取的两步阶段划分算法,并在子阶段建立注意力LSTM集成质量预测模型. 该方法首先将沿批次展开的二维数据矩阵划分为不同的时间片,对每个时间片矩阵进行PLS分析,得到原始过程变量和质量变量的得分矩阵,采用AP(affinity propagation)算法<sup>[28]</sup>将联合得分矩阵进行聚类,实现操作阶段的第1步粗划分. 之后,利用深度学习中的encoder-decoder模型<sup>[29]</sup>提取表征操作过程动态性的动态特征,采用AP算法对其进行第2次阶段划分,综合分析两步划分结果,最终将整个生产操作阶段划分为不同的稳定子时段和过渡子时段. 最后,在各个操作阶段分别建立注意力LSTM集成质量预测模型. 与前人工作相比,本文方法主要有以下优点: 1) 由于时间片PLS建模忽略了时间片前后的联系,且传统的PLS方法属于静态方法,本文提出的方法将原始数据的静态特征和动态特征全部考虑进来,采用encoder-decoder模型提取的动态特征进行2次划分弥补了1次划分的不足. 2) AP聚类算法无需事先指定聚类中心的数目,实际在大多数情况下,对于一个复杂、陌生的工业过程,很难知道其确切的阶段数目. 3) 综合分析两步的划分结果,可以将整个操作阶段划分为稳定阶段和过渡阶段,使得划分结果符合实际生产过程. 4) 在预测时,在传统LSTM网络的基础上加入了注意力机制并引入了集成学习思想,使得预测结果更加精确.

## 1 离线建模策略

### 1.1 阶段划分方法

#### 1.1.1 偏最小二乘(PLS)

PLS方法主要针对因变量和自变量之间建立回归关系. 对于一组生产过程数据,自变量可以认为是一些易测得的过程变量 $X$ ,因变量可以认为是一些难在线测量的质量变量 $Y$ . 传统PLS本质上是一种静

态、线性建模方法. 在进行建模时, 该算法考虑到了  $X$  和  $Y$  两方面的信息, 并且试图最大化它们之间的相关关系, 将它们分解成如下形式:

$$X = \sum_{r=1}^A t_r p_r^T + E = TP^T + E, \quad (1)$$

$$Y = \sum_{r=1}^B u_r q_r^T + F = UQ^T + F. \quad (2)$$

其中:  $A$  和  $B$  分别代表  $X$  和  $Y$  提取的主成分个数, 用交叉验证法来确定其个数;  $T$  和  $U$  分别为对  $X$  和对  $Y$  的得分矩阵,  $P$  和  $Q$  分别为对  $X$  和对  $Y$  的负载矩阵,  $E$  和  $F$  分别为对  $X$  和对  $Y$  的残差矩阵.

PLS 是一种有偏回归方法, 其自变量  $X$  与因变量  $Y$  的最终回归模型可表示为

$$Y = X\beta + E_Y. \quad (3)$$

其中:  $\beta$  是回归系数向量,  $E_Y$  是预测误差.

### 1.1.2 encoder-decoder 模型

encoder-decoder 模型, 也称 seq2seq 模型, 它是深度学习中的一个常见的结构. 通过编码部分可以将输入序列编码成一个固定长度的向量, 该向量经过解码部分可以转换成一个输出序列. 用于提取过程动态

特征的 encoder-decoder 模型结构如图 1 所示. 编码部分和解码部分的基本单元可以为 RNN、LSTM 等. 由于 LSTM 网络可以解决传统 RNN 网络中存在的梯度消失和梯度爆炸问题, 选择使用 LSTM 网络作为该模型的基本单元. encoder-decoder 模型的输入要求是序列, 因此将过程变量处理成序列格式作为编码部分的输入. 由于经过编码后得到的向量  $C$  连接了编码和解码部分, 又因为该模型采用 LSTM 作为基本单元, LSTM 存储单元中使用记忆细胞保存信息, 通过输入门、遗忘门和输出门来控制细胞状态 (其中输入门控制信息流对细胞状态的更新, 遗忘门决定从细胞状态中丢弃的信息, 而输出门最终确定存储单元的输出), 该向量  $C$  的值不仅体现了当前时刻  $T$  的动态特征, 而且考虑了时刻  $T$  之前的采样值对  $T$  时刻的影响, 因此向量  $C$  的值体现了与输出序列相关的输入序列的动态变化<sup>[30]</sup>. 此外, 阶段划分完成后在建立质量预测模型时同样可采用 encoder-decoder 模型, 因此在阶段划分时首先将该模型能够提取的信息进行分析, 有助于更好地理解该模型能够处理的不同时刻的数据特点. 综合考虑数据本身的特点以及数据内部的变化特点, 有助于更好地建模, 提升预测精度.

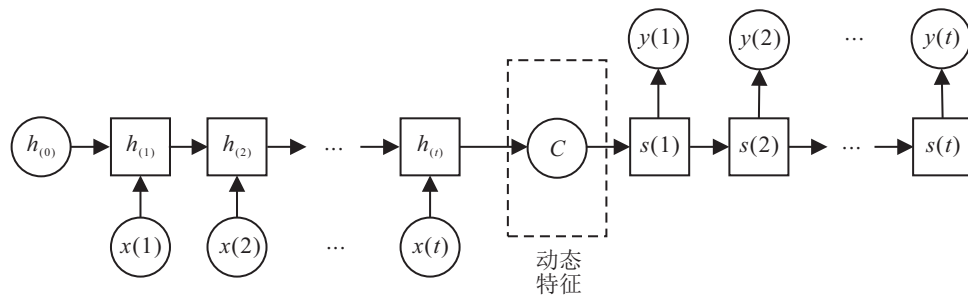


图 1 encoder-decoder 模型结构

### 1.1.3 基于动态特征提取的两步阶段划分方法

发酵过程的历史过程数据是三维矩阵, 由  $X(I \times J_x \times K)$  表示, 其中:  $I$  为批次;  $J_x$  为过程数据变量, 过程数据在发酵过程中是可以在线测量的;  $K$  为每个批次的采样时刻数. 质量数据由  $Y(I \times J_y \times K)$  表示, 在实际生产中质量变量由离线测量获得. 本文假设建模所需的各批次操作时间都是等长的.

为了同时考虑过程变量和质量变量对生产过程阶段划分的影响, 本文将质量变量的得分矩阵  $U$  和过程变量的得分矩阵  $T$  相结合得到联合得分矩阵. AP 聚类算法相较其他聚类算法无需事先指定聚类数目, 因此可以更合理地将过程数据进行划分, 将每个时间片的联合得分矩阵使用 AP 聚类算法进行初步划分. 由于每个联合得分矩阵是对每个独立时间片采用

PLS 算法得到的, 同时传统的 PLS 算法是线性的、静态的, 得到的结果不仅忽略了每个时间片的非线性特征, 也忽略了时间片前后的动态变化信息. 因此, 可以认为使用 AP 聚类算法是对原始数据的静态特征进行的聚类, 得到的聚类结果只能反应出原始数据大致的分布特征. 发酵过程的动态非线性往往与发酵阶段密切相关, 因此有必要将动态非线性特征加以提取并进行细化分.

本文使用训练好的 encoder-decoder 模型进行动态特征提取, 每一个时刻的输入度可以得到对应时刻的动态特征; 其次, 将所有批次得到的动态特征进行组合得到动态特征时间片; 然后, 使用 AP 聚类算法对所有时间片进行划分; 最后, 综合分析两步划分结果, 将发酵过程划分为不同的子阶段, 即:

1) 将三维过程数据  $X(I \times J_x \times K)$  沿批次方向展开, 得到  $K$  个时间片  $X_i(I \times J_x), i = 1, 2, \dots, K$ , 并进行标准化, 即

$$\tilde{x}_{k,j} = \frac{x_{k,j} - \bar{x}_{k,j}}{S_{k,j}} \quad (4)$$

其中:  $j = 1, 2, \dots, J_x, k = 1, 2, \dots, K, \bar{x}_{k,j}$  是第  $k$  采样时刻第  $j$  个过程变量的平均值,  $x_{k,j}$  是第  $k$  采样时刻第  $j$  个过程变量的值,  $S_{k,j}$  是第  $k$  采样时刻第  $j$  个过程变量的标准差. 对质量数据做同样的处理.

2) 对每个时间片做 PLS 分析, 得到  $X$  和  $Y$  的得分矩阵, 将两个得分矩阵左右排成一个矩阵, 得到联合得分矩阵, 对联合得分矩阵使用 AP 聚类算法进行划分, 实现生产过程的初步化分. AP 算法原理如下:

AP 算法根据  $N$  个数据点之间的相似度进行聚类, 其输入是样本的相似度矩阵  $S$ . 所以若要对联合得分矩阵进行聚类, 关键是定义不同时刻矩阵的相似度. 考虑到第 1 步划分是基于数据的静态特征, 因此原始数据存在噪声或奇异值可能会导致存在时间距离较远的两个时刻被划分到同一阶段的现象, 为此在计算两个时刻样本矩阵的相似度时加上时间约束<sup>[31]</sup>, 保证时刻相差较远的采样点不可能被划分到同一阶段. 计算公式如下:

$$S(i, k) = -\|(T, U)_i - (T, U)_k\|A - |k - i|B. \quad (5)$$

其中:  $A$  和  $B$  由交叉验证法确定,  $A$  和  $B$  的取值要保证阶段内的波动尽可能少, 两个相邻阶段间的转换部分允许波动存在, 以此判断错分点大部分处于阶段转换部分. 若无法明显看出两个相邻阶段间转换部分的波动情况, 则进一步采用 Silhouette 准则<sup>[32]</sup> 衡量当前样本点划分到对应阶段的“优劣”程度, 聚类的“优劣”取决于聚类后类内紧致度与类间离散度.

3) 将沿批次展开后得到的二维数据进一步按照变量展开方式进行处理, 然后对每批次数据以窗宽为  $d$  的滑动窗进行连续采样得到 encoder-decoder 模型所需的输入序列. 采用训练好的 encoder-decoder 模型提取动态特征. 将不同批次每个时刻的动态特征组合成为动态特征时间片, 使用 AP 聚类算法进行划分, 实现生产过程的二次划分. 由于提取出来的动态特征和时间有较强的相关性, 并不会出现划分静态特征时间片时两个相邻阶段转换部分存在着较强波动的情况. 由于无法观测两个相邻阶段的波动情况, 采用 Silhouette 准则, 此时定义样本间的相似度为

$$S(i, k) = -\|C_i - C_k\|. \quad (6)$$

其中:  $i, k$  代表不同的时刻;  $C_i, C_k$  代表  $i, k$  时刻对应的动态特征时间片.

4) 综合分析两步划分结果, 将动态特征聚类效果好、静态特征波动较强(聚类效果差)的时刻认为其仍处于合理的稳定阶段内; 将静态特征平稳(聚类效果好)、动态特征聚类效果差的时刻认为其存在过渡趋势, 仍归属为合理的稳定阶段; 将动态特征聚类效果差、静态特征存在较强波动(聚类效果差)的时刻认为其处于过渡阶段.

## 1.2 子阶段质量预测

### 1.2.1 基于注意力机制的 encoder-decoder 模型

前文使用 encoder-decoder 模型进行动态特征提取实现了阶段划分, 对于阶段划分而言, 使用 encoder-decoder 模型提取出的动态特征足够表征过程动态性. 但是在建立预测模型阶段, 为了提高预测精度, 引入注意力机制. 该模型结构如图 2 所示. 使用此种模型可以在预测不同时刻的输出时对应不同的  $C$ . 在时刻  $t$  向量  $C$  使用下式计算:

$$C_{(t)} = \sum_{i=1}^T w_i h_{(i)}. \quad (7)$$

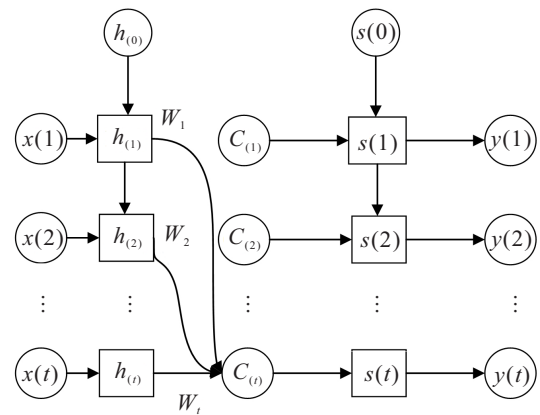


图 2 attention-based encoder-decoder 模型结构

### 1.2.2 注意力 LSTM 集成预测模型

集成学习在很多应用领域有效地提高了回归模型或分类模型的精度<sup>[33-34]</sup>. 集成学习的思想是以多样性为原则建立多个模型, 然后融合它们的输出作为最终结果. 与单一模型相比, 使用集成模型可以明显提高预测精度.

bagging 通过有放回的重采样原始训练集来满足产生多个回归模型的多样性原则, 最后通过将多个预测值取平均来改善预测性能. 考虑到发酵过程中有的变量之间存在较强的相关性, 在训练模型时若存在过多的冗余信息, 则会大大增加过拟合的风险. 同时, 为了满足集成学习要求增强不同模型之间的差异性, 使用 bagging 策略在发酵过程的  $J$  个变量中每次有放回地随机取  $j$  个变量, 共取  $M$  次, 这样便得到了  $M$  个

模型. 最终预测结果由  $M$  个模型的预测结果取平均值得到, 即

$$\hat{y}_t = \frac{1}{M} \sum_{m=1}^M \hat{y}_{m,t}. \tag{8}$$

这样不仅减小了每个模型的训练复杂度, 而且增强了每个模型之间的差异, 符合集成学习的思想. 与传统的软测量模型相比, 本文提出的注意力 LSTM 集成预测模型主要有两个优势: 1) 将过程数据按照过程特点进行阶段划分, 在传统 encoder-decoder 模型的基础上引入注意力机制, 提升模型精度. 2) 使用简单的 bagging 集成策略可以避免使用单一预测模型引起的过拟合问题. 因此, 这两个优势使得集成注意力 LSTM 模型更加适合发酵过程质量变量的在线预测.

### 2 全局质量预测策略

全局质量预测策略如图3所示. 在离线建模过程中, 将正常发酵过程的三维历史数据沿批次方向展开, 对每个时间片做 PLS 分析, 得到  $X$  和  $Y$  的得分矩阵, 将两个得分矩阵左右排成一个矩阵, 得到联合得分矩阵. 采用 AP 聚类算法对时间片的联合得分矩阵进行划分, 实现第 1 步划分; 然后采用 seq2seq 模型提取过程数据的动态特征, 采用 AP 聚类算法实现第 2 步划分; 最后建立子阶段质量预测模型. 生产过程中, 依据采样时刻判断当前所处阶段, 然后根据所属阶段的模型进行预测. 阶段划分后可以使得在同一子阶段内的数据具有高度相似的特征. 将局部模型代替整体模型可以大大提高模型的精度, 减小模型预测误差.

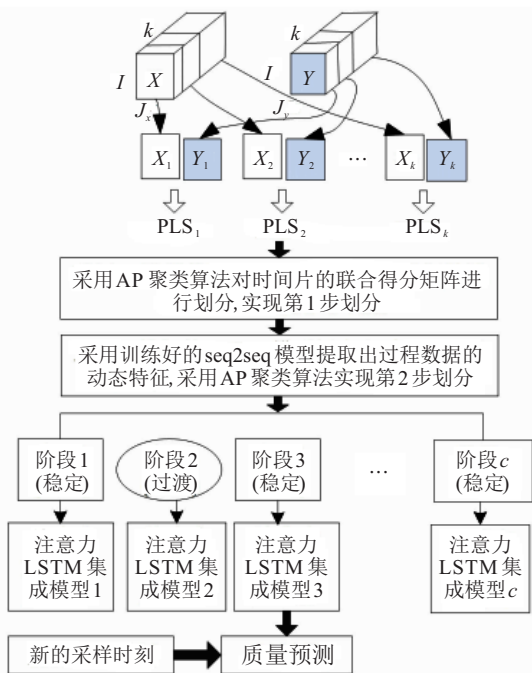


图3 全局质量预测策略

## 3 实验与结果分析

### 3.1 仿真实验

Pensim 仿真平台是国际上较有影响力的青霉素仿真平台, 相关研究已表明其实用性和有效性<sup>[35-36]</sup>.

本文基于该平台对青霉素的发酵过程进行了实验仿真, 设定每批次发酵时间为 400 h, 采样时间间隔为 1 h, 选择 11 个过程变量 (包括通风速率、搅拌功率、底物流加速率、底物温度、溶解氧浓度、排气二氧化碳浓度、酸碱度、温度、反应热、冷水流加速率和底物浓度) 和 1 个质量变量 (产物浓度) 进行监测, 共产生 60 批正常数据, 其中 40 批用来建模, 20 批用来测试.

采用本文所提出方法, 最终划分结果见图 4. 将数据处理完成后, 采用 AP 算法对得到的联合得分矩阵进行聚类. 综合考虑矩阵之间的相似度和时间约束, 采用交叉验证法确定参数  $A$  的值为 2、 $B$  的值为 1 即可, 将 AP 聚类算法的相似度定义为

$$S(i, k) = -\|(T, U)_i - (T, U)_k\|^2 - |k - i|. \tag{9}$$

观察划分结果可明显看出 1~43, 68~142, 197~266, 301~400 时刻不存在波动, 而在 44~67, 143~196, 267~300 时刻波动非常强烈.

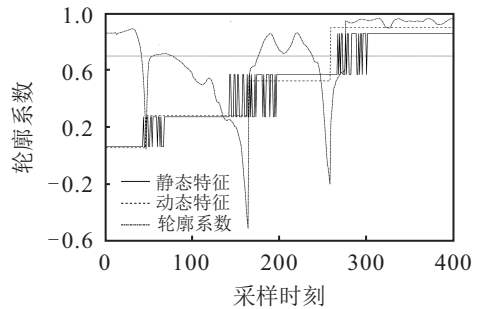


图4 青霉素发酵过程阶段划分结果

采用 encoder-decoder 模型提取动态特征, 该模型采用 LSTM 作为基本单元, 需要设定的主要参数有隐含层节点数目和采样时间窗口大小, 本文采用网格搜索法对两个参数进行交替寻优确定. 最终选择隐含层节点数目为 24, 采样窗宽为 10. 然后采用 AP 聚类算法对发酵过程中的动态特征进行再次划分, 此时相似度的定义如式 (6) 所示.

观察结果可知, 不同于静态特征的相似度函数, 此时不用考虑时间约束即可将动态特征时间片实现较理想的划分, 不存在两个相邻阶段转换时的强烈波动状态. 为此, 引入 Silhouette 准则评价聚类效果的好坏. 对于一系列聚类结果的 Silhouette 指标值, 值越大表示聚类质量越好, 最大值对应的类数是最优的聚类个数, 对应的聚类结果也是最优的. 为了合理判定聚类效果的好坏, 通过多次实验, 将采样点 Silhouette

准则的平均值作为临界值<sup>[32]</sup>,在该仿真实验中将临界值确定为0.7. *Sihouette* 值比0.7低则认为聚类效果很差. 由于采样宽度为10,前9个时刻的动态特征无法提取,观察静态特征划分结果可以发现第1个子阶段的时刻远远比9大,不妨将前9个时刻的划分结果和第10个结果划分到同一阶段. 由 *Sihouette* 值可知,41~54,77~175,238~275时刻的聚类效果不好,而在其他时刻聚类效果较好. 最后综合分析划分1~43,55~142,176~266,276~400时刻为稳定阶段,44~54,143~175,267~275时刻为过渡阶段.

为了验证本文方法的有效性,实验将本文方法与不分阶段的注意力LSTM集成预测模型、使用本文提出的阶段划分方法再分别建立传统LSTM模型、直接对第1步划分结果建立注意力LSTM集成预测模型3种情况进行对比. 采用预测均方根误差RMSE(root mean square error, RMSE)和  $R^2$  作为模型的评价指标,有

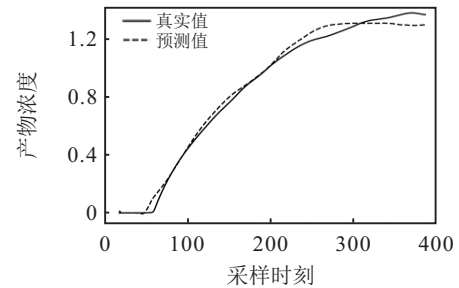
$$RMSE = \sqrt{\frac{1}{K} \sum_{k=1}^K \|y_k - \hat{y}_k\|^2}, \quad (10)$$

$$R^2 = \frac{\sum_{k=1}^K (y_k - \bar{y})^2 - \sum_{k=1}^K (y_k - \hat{y}_k)^2}{\sum_{k=1}^K (y_k - \bar{y})^2}. \quad (11)$$

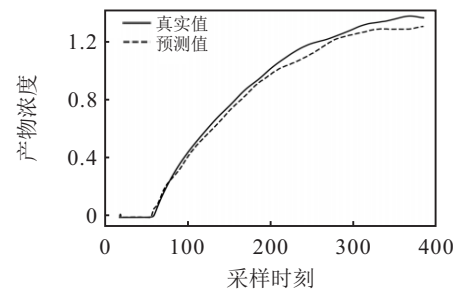
表1为4种不同方法在对20批测试数据进行质量预测后得到的平均RMSE和  $R^2$ . 通过比较可以看出,本文方法的预测精度优于其他3种方法,而且阶段划分后再建立局部模型的效果要比建立一个整体模型效果好. 为进一步说明结果,实验对单一批次的预测曲线进行分析,选择了第10测试批次. 图5为4种方法的产物浓度在线预测结果. 可以看出,图5中(a)和(b)的预测偏差要比(c)和(d)大,原因在于(a)和(b)未考虑到发酵过程的多阶段特征,而不同阶段的数据特征又相差较大,因此对于不同阶段的数据建立全局模型势必会影响预测精度,也说明了建立局部模型的好处. 其次,相比于(c)和(d)的波动幅度更小,原因在于在对发酵过程数据提取动态特征时,采用的也是encoder-decoder模型,而预测模型也同样采用的是这种结构. 因此,该动态特征不仅反映出不同时刻前后变化信息,而且不同时刻动态特征的相似性也能够反映出不同时刻数据对同一个预测模型的适应性. 显然,若两个时刻该特征相差较大,将该两个时刻分配到两个不同的模型则会提升预测精度. 以上实验验证了本文所提出阶段划分方法和预测模型的优良性能.

表1 4种方法对青霉素浓度的预测精度指标

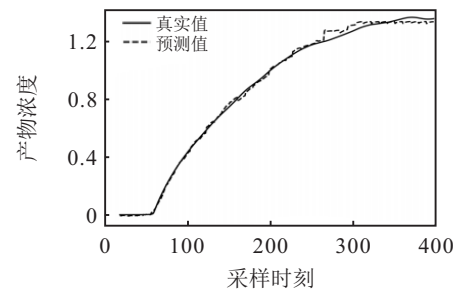
method	LSTM	attention-based LSTM	attention-based LSTM after one-step partition	attention-based LSTM after two-step partition
RMSE	0.0135	0.0129	0.0110	0.0107
$R^2$	0.947	0.956	0.979	0.988



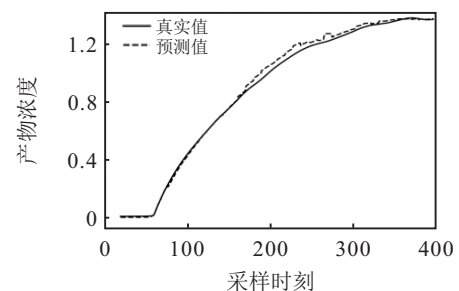
(a) 基于LSTM的预测结果



(b) 基于注意力LSTM的预测结果



(c) 一步划分后基于注意力LSTM的预测结果



(d) 两步划分后基于注意力LSTM的预测结果

图5 青霉素浓度在线预测结果

### 3.2 算法实际验证

生物制药中利用大肠杆菌生产白介素-2是一个典型的多阶段发酵过程,其生产过程的持续时间大约为6~7h. 本文选取北京某生物制药厂的大肠杆菌发酵过程实际生产数据进行算法验证. 选取以下8个过程变量(包括酸碱度、溶解氧、浓度罐压、温度、

搅拌速率、补碳、补氮、通风速率)和1个质量变量(菌体浓度). OD600的值可以间接反应出菌体浓度,工作人员根据其值的变化用来确定下一步操作.生产过程中采用离线测量OD600的值,存在一定的时间滞后性,不能及时反馈并指导生产.因此,对OD600进行准确在线预测有着重要的实际意义.现场的发酵时间为6h,采样间隔为10min.共采集28批正常数据,其中20批用于阶段划分和离线建模,剩余8批用来测试.为使实验结果更具说服力,采用与上节仿真实验相同的对比实验.

图6为采用本文所提出方法得到的划分结果,与青霉素的发酵过程不同的是,大肠杆菌的发酵时间较短,且两个相邻阶段的过渡并不是非常明显,因此首先对大肠杆菌数据的静态特征采用AP算法进行第1步划分,此时相似度函数为

$$S(i, k) = -2\|(T, U)_i - (T, U)_k\| - |k - i|. \quad (12)$$

划分后发现:两个相邻阶段的转换部分不存在明显波动,则进一步采用Silhouette准则确定划分的“优劣”点;第2步划分采用与青霉素发酵过程第2步划分相同的方法即可.最后综合分析两步划分结果,得到最终的划分结果为:1~11, 16~21, 26~36时刻为稳定阶段,12~15, 22~25时刻为过渡阶段.

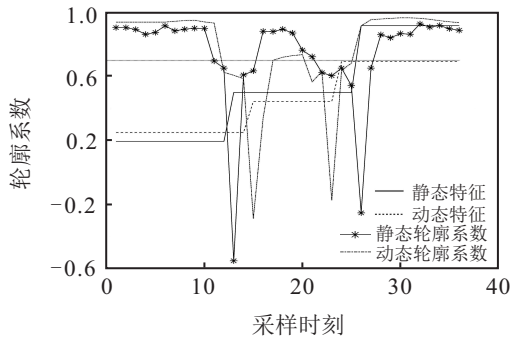
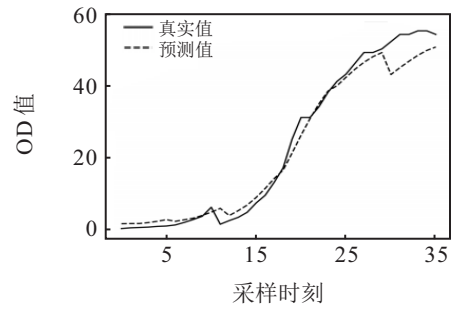


图6 大肠杆菌发酵过程阶段划分结果

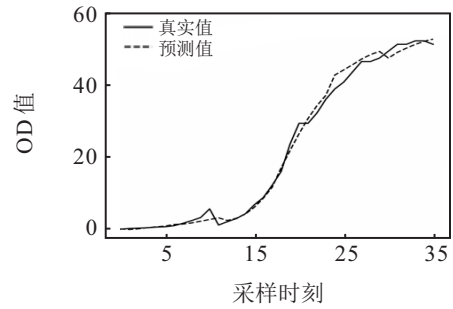
表2为4种不同方法在对8批测试数据进行质量预测后得到的平均RMSE和R<sup>2</sup>.通过比较可以看出,本文所提出方法的预测精度优于其他3种方法.为进一步说明结果,实验对单一批次的预测曲线进行分析,选择了第4测试批次.图7为4种方法的菌体浓度在线预测结果.可以看出,与图7中(a)、(b)、(c)相比,图7(d)的波动幅度更小,说明了本文所提出阶段划分

表2 4种方法对OD值的预测精度指标

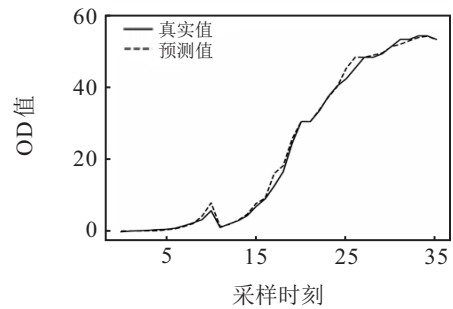
method	LSTM	attention-based LSTM	attention-based LSTM after one-step partition	Attention-based LSTM after two-step partition
RMSE	0.386 5	0.361 5	0.337 2	0.323 1
R <sup>2</sup>	0.936	0.951	0.969	0.978



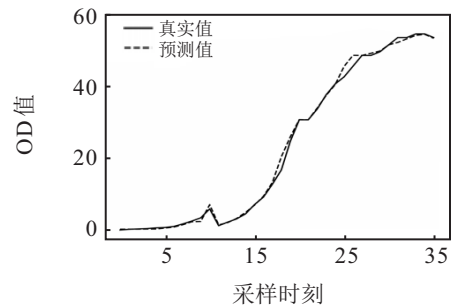
(a) 基于LSTM的预测结果



(b) 基于注意力LSTM的预测结果



(c) 一步划分后基于注意力LSTM的预测结果



(d) 两步划分后基于注意力LSTM的预测结果

图7 OD值在预测结果

方法和预测模型的有效性.

### 4 结论

大多数阶段划分方法将每个时间片单独作为一个单元进行分析,忽略了不同时刻之间的联系.本文考虑发酵过程的动态变化,综合分析每个时刻过程数据本身的特点以及前后时刻的动态变化信息.由于提取动态特征采用的是处理序列问题很好的encoder-decoder模型,在进行预测时也采用的是该结构,划分结果也与最后的预测模型建立了联系,使得每个阶段的数据充分适合该阶段的模型.为了更

好地提升预测精度,在建模阶段引入了注意力机制和集成学习思想。最后,利用青霉素发酵仿真平台和大肠杆菌实例验证了该方法的可行性和有效性,为发酵过程的质量预测提供了一种可行方案。

### 参考文献(References)

- [1] Lv F, Wen C L, Bao Z J, et al. Fault diagnosis based on deep learning[C]. 2016 American Control Conference (ACC). Boston, 2016: 6851-6856.
- [2] 刘毅,陈坤,王海清,等. 选择性递推LSSVR及其在过程建模中的应用[J]. 高校化学工程学报, 2008, 22(6): 1043-1048.  
(Liu Y, Chen K, Wang H Q, et al. Selective recursive LSSVR and its applications in process modeling[J]. Journal of Chemical Engineering of Chinese Universities, 2008, 22(6): 1043-1048.)
- [3] Kourti T, MacGregor J F. Process analysis, monitoring and diagnosis, using multivariate projection methods[J]. Chemometrics and Intelligent Laboratory Systems, 1995, 28(1): 3-21.
- [4] 赵斐,陆宁云,杨毅. 基于工况识别的注塑过程产品质量预测方法[J]. 化工学报, 2013, 64(7): 2526-2534.  
(Zhao F, Lu N Y, Yang Y. Product quality prediction method for injection molding process based on operating mode recognition[J]. CIESC Journal, 2013, 64(7): 2526-2534.)
- [5] Wang X Z, McGreavy C. Automatic classification for mining process operational data[J]. Industrial & Engineering Chemistry Research, 1998, 37(6): 2215-2222.
- [6] 郭小萍,陆宁云,高福荣,等. 间歇过程滑动窗口子时段PCA建模和在线监测[J]. 控制与决策, 2005, 20(9): 1034-1037.  
(Guo X P, Lu N Y, Gao F R, et al. Sub-stage PCA modeling and on-line monitoring with moving windows for batch processes[J]. Control and Decision, 2005, 20(9): 1034-1037.)
- [7] Kadlec P, Gabrys B, Strandt S. Data-driven soft sensors in the process industry[J]. Computers & Chemical Engineering, 2009, 33(4): 795-814.
- [8] 常鹏,王普,高学金. 基于高阶累计统计量的微生物发酵过程监测[J]. 控制与决策, 2017, 32(12): 2273-2278.  
(Chang P, Wang P, Gao X J. Microbial fermentation process monitoring based on higher order cumulant statistics[J]. Control and Decision, 2017, 32(12): 2273-2278.)
- [9] Kadlec P, Grbi R, Gabrys B. Review of adaptation mechanisms for data-driven soft sensors[J]. Computers & Chemical Engineering, 2011, 35(1): 1-24.
- [10] 贾润达,毛志忠,王福利. 基于KPLS模型的间歇过程产品质量控制[J]. 化工学报, 2013, 64(4): 1332-1339.  
(Jia R D, Mao Z Z, Wang F L. KPLS model based product quality control for batch processes[J]. CIESC Journal, 2013, 64(4): 1332-1339.)
- [11] 刘国海,江兴科,梅从立. 基于连续隐Markov模型的发酵过程软测量方法[J]. 控制与决策, 2011, 26(11): 1753-1756.  
(Liu G H, Jiang X K, Mei C L. Research on soft sensing method based on continuous hidden Markov model in fermentation process[J]. Control and Decision, 2011, 26(11): 1753-1756.)
- [12] Wold S, Ruhe A, Wold H, et al. The collinearity problem in linear regression. the partial least squares (PLS) approach to generalized inverses[J]. SIAM Journal on Scientific and Statistical Computing, 1984, 5(3): 735-743.
- [13] Lindgren F, Geladi P, Wold S. The kernel algorithm for PLS[J]. Journal of Chemometrics, 1993, 7(1): 45-59.
- [14] Rosipal R, Tregó L J. Kernel partial least squares regression in reproducing kernel Hilbert space[J]. The Journal of Machine Learning Research, 2002, 2: 97-123.
- [15] 蒋立. 基于自编码器模型的非线性过程监测[D]. 杭州: 浙江大学, 2018.  
(Jiang L. Nonlinear process monitoring based on auto-encoder model[D]. Hangzhou: Zhejiang University, 2018.)
- [16] Shang C, Yang F, Huang D X, et al. Data-driven soft sensor development based on deep learning technique[J]. Journal of Process Control, 2014, 24(3): 223-233.
- [17] Yan W W, Tang D, Lin Y J. A data-driven soft sensor modeling method based on deep learning and its application[J]. IEEE Transactions on Industrial Electronics, 2017, 64(5): 4237-4245.
- [18] Yao L, Ge Z Q. Deep learning of semisupervised process data with hierarchical extreme learning machine and soft sensor application[J]. IEEE Transactions on Industrial Electronics, 2018, 65(2): 1490-1498.
- [19] Ku C C, Lee K Y. Diagonal recurrent neural networks for dynamic systems control[J]. IEEE Transactions on Neural Networks, 1995, 6(1): 144-156.
- [20] Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [21] 赵春晖,王福利,姚远,等. 基于时段的间歇过程统计建模、在线监测及质量预报[J]. 自动化学报, 2010, 36(3): 366-374.  
(Zhao C H, Wang F L, Yao Y, et al. Phase-based statistical modeling, online monitoring and quality prediction for batch processes[J]. Acta Automatica Sinica, 2010, 36(3): 366-374.)
- [22] 邓晓刚,张琛琛,王磊. 基于多阶段多向核熵成分分析

- 的间歇过程故障检测方法[J]. 化工学报, 2017, 68(5): 1961-1968.  
(Deng X G, Zhang C C, Wang L. Fault detection in batch process by multistage multiway kernel entropy component analysis[J]. CIESC Journal, 2017, 68(5): 1961-1968.)
- [23] Yu J, Qin S J. Multiway Gaussian mixture model based multiphase batch process monitoring[J]. Industrial & Engineering Chemistry Research, 2009, 48(18): 8585-8594.
- [24] Camacho J, Picó J. Online monitoring of batch processes using multi-phase principal component analysis[J]. Journal of Process Control, 2006, 16(10): 1021-1035.
- [25] 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.
- [26] Zhao C H, Wang F L, Lu N Y, et al. Stage-based soft-transition multiple PCA modeling and on-line monitoring strategy for batch processes[J]. Journal of Process Control, 2007, 17(9): 728-741.
- [27] 齐咏生, 王普, 高学金. 基于核主元分析-主元分析的多阶段间歇过程故障监测与诊断[J]. 控制理论与应用, 2012, 29(6): 754-764.  
(Qi Y S, Wang P, Gao X J. Fault detection and diagnosis of multiphase batch process based on kernel principal component analysis-principal component analysis[J]. Control Theory & Applications, 2012, 29(6): 754-764.)
- [28] Frey B J, Dueck D. Clustering by passing messages between data points[J]. Science, 2007, 315(5814): 972-976.
- [29] Cho K, van Merriënboer B, Gulcehre C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[J/OL]. 2014, arXiv: 1406.1078.
- [30] Greff K, Srivastava R K, Koutník J, et al. LSTM: A search space odyssey[J]. IEEE Transactions on Neural Networks and Learning Systems, 2017, 28(10): 2222-2232.
- [31] 高学金, 杨彦霞, 王普, 等. 基于扩展核熵负载矩阵的发酵过程故障监测[J]. 控制理论与应用, 2018, 35(6): 813-821.  
(Gao X J, Yang Y X, Wang P, et al. Fault monitoring of fermentation process based on extended kernel entropy load matrix[J]. Control Theory & Applications, 2018, 35(6): 813-821.)
- [32] 王开军, 张军英, 李丹, 等. 自适应仿射传播聚类[J]. 自动化学报, 2007, 33(12): 1242-1246.  
(Wang K J, Zhang J Y, Li D, et al. Adaptive affinity propagation clustering[J]. Acta Automatica Sinica, 2007, 33(12): 1242-1246.)
- [33] Liu Y, Yang C, Gao Z L, et al. Ensemble deep kernel learning with application to quality prediction in industrial polymerization processes[J]. Chemometrics and Intelligent Laboratory Systems, 2018, 174: 15-21.
- [34] Liu Y, Liang Y, Gao Z L. Industrial polyethylene melt index prediction using ensemble manifold learning-based local model[J]. Journal of Applied Polymer Science, 2017, 134(29): 45094.
- [35] Birol G, Ündey C, Çinar A. A modular simulation package for fed-batch fermentation: Penicillin production[J]. Computers & Chemical Engineering, 2002, 26(11): 1553-1565.
- [36] 刘毅, 王海清. Pensim 仿真平台在青霉素发酵过程的应用研究[J]. 系统仿真学报, 2006, 18(12): 3524-3527.  
(Liu Y, Wang H Q. Pensim simulator and its application in penicillin fermentation process[J]. Journal of System Simulation, 2006, 18(12): 3524-3527.)

### 作者简介

高学金(1973—), 男, 教授, 博士生导师, 从事发酵过程建模、监控与故障诊断等研究, E-mail: gaouxuejin@bjut.edu.cn;

孟令军(1996—), 男, 硕士生, 从事发酵过程建模、质量预测的研究, E-mail: wsyjzmlj@163.com;

高慧慧(1991—), 女, 讲师, 博士, 从事机器学习、深度学习的研究, E-mail: gaohh@bjut.edu.cn.

(责任编辑: 孙艺红)