

基于延拓补偿策略的气体传感器端点效应诊断

朱健松, 邢博轩, 孟凡利, 王浩, 唐坤

(东北大学 信息科学与工程学院, 沈阳 110819)

摘要: 针对经验模态分解(empirical mode decomposition, EMD)处理非平稳信号时因端点效应造成分解结果失真的问题,提出一种基于麻雀搜索算法(sparrow search algorithm, SSA)与长短时记忆(long short-term memory, LSTM)网络的耦合模型,突破传统梯度下降算法易陷入局部最优的局限,显著提升时序预测精度。首先将气体响应信号预处理为周期特征变量;然后采用双向周期延拓策略,通过LSTM-SSA深度训练,生成首尾各延伸一个周期的预测序列;最后利用双向性预测序列构建复合信号,并对其进行EMD分解。以丙酮和甲苯信号为例的实验结果表明,经LSTM-SSA预测后再进行EMD分解时端点效应引起的能量误差分别降低了74.966%和23.368%、正交性系数分别提升了51.444%和34.990%,有效抑制了端点处模态分量的幅值失真,提升了EMD的可靠性,为气体传感信号的特征提取与工业安全监测提供了新思路。

关键词: 经验模态分解;端点效应;麻雀搜索算法;长短时记忆网络;周期延拓

中图分类号: TP212.2 **文献标志码:** A **DOI:** 10.3969/j.issn.1003-1251.2026.01.006

Diagnosis of Endpoint Effects in Gas Sensors Based on Extended Compensation Strategy

ZHU Jiansong, XING Boxuan, MENG Fanli, WANG Hao, TANG Kun

(Northeastern University, Shenyang 110819, China)

Abstract: To address the issue of decomposition result distortion caused by endpoint effects when empirical mode decomposition (EMD) is applied to non-stationary signals, a coupled model based on the sparrow search algorithm (SSA) and long short-term memory (LSTM) network is proposed. This model overcomes the limitation of traditional gradient descent algorithms that are prone to getting stuck in local optima, significantly improving the accuracy of time series prediction. Firstly, the gas response signal is preprocessed into periodic characteristic variables. Then, a bidirectional periodic extension strategy is adopted, and through deep training with LSTM-SSA, a prediction sequence extending one period at both the beginning and the end is generated. Finally, a composite signal is constructed using the bidirectional prediction sequence for EMD. Experimental results using acetone and toluene signals show that when EMD is performed after LSTM-SSA prediction, the energy error caused by endpoint effects is reduced by 74.966% and 23.368%, respectively, and the orthogonality coefficient is increased by 51.444% and 34.990%, respectively. This effectively suppresses the amplitude distortion of mode components at the endpoints and enhances the reliability of EMD, providing a new approach for feature extraction of gas sensing signals and industrial safety monitoring.

收稿日期: 2025-04-09

基金项目: 国家重点研发计划项目(2024YFB4710900);国家自然科学基金项目(62033002);国家级大学生创新创业训练计划项目(03040021301002)

作者简介: 朱健松(2004—),男,本科生;孟凡利(1980—),通信作者,男,教授,博士。

Key words: empirical mode decomposition; endpoint effect; sparrow search algorithm; long short-term memory network; cyclical extension

随着社会经济高速发展和生产能力不断增强,化工、制药等行业的生产过程中产生的挥发性有机化合物(volatile organic compounds, VOCs)也日益增加,不仅对环境造成严重污染,其中的易燃易爆气体还是威胁生产安全的重大隐患^[1]。因此,对大气中 VOCs 浓度进行监测具有重要意义。目前通常采用气体传感器监测 VOCs 浓度,其中将温度调制方法与分解算法相结合进行该类气体的识别是气体传感器研究的重点^[2]。

在信号处理领域,经验模态分解(empirical mode decomposition, EMD)是针对非平稳、非线性信号的一种处理方法^[3],该方法基于迭代筛分原理将信号分解为本征模态函数(intrinsic mode function, IMF)。但是,实际应用中有限长数据序列的边界易产生端点效应,导致 IMF 分量在首尾区间的幅值发生畸变,降低分解可信度并引发模态混叠等问题。因此,开发有效的边界约束策略对提升 EMD 的工程适用性至关重要^[4]。

现有抑制端点效应的策略可分为四类:一是信号处理驱动方法,如镜像延拓、斜率延拓等,该方法可能无法完全包裹信号数据,从而导致数据泄露;二是数学模型驱动方法,如多项式拟合延拓、比例延拓等,该方法适应性差且延拓效果不稳定;三是基于人工智能的驱动方法,如人工神经网络延拓、支持向量机延拓等,该方法对非平稳信号的泛化能力受限;四是混合策略驱动方法,如镜像延拓耦合支持向量机的延拓法,该方法往往具有较高的计算复杂度,且参数调整较为复杂^[5-6]。

基于上述研究背景,本文针对气体传感响应信号的特征提取需求,构建一种基于麻雀搜索算法(sparrow search algorithm, SSA)与长短期记忆网络(long short-term memory, LSTM)的耦合模型,利用该模型对原始信号进行双向延拓,生成与原始信号动态特性高度匹配的复合信号,提升 EMD 分解的可靠性和准确性,进而提高气体传感器的特征提取性能。

1 相关理论

1.1 经验模态分解

1.1.1 基本原理

EMD 的核心机制在于通过自适应筛分过程,

将非平稳时序信号解构为具有明确物理意义的多个 IMF 分量,构造的 IMF 分量需严格满足如下两个数学约束^[7]。

1) 单分量性准则

对于任意 IMF 分量,其时域波形的极值点拓扑分布须满足极大值点与极小值点的数目和与过零点的数目相差不超过 1 个。

2) 对称包络约束

在时间域 $t \in [t_0, t_n]$ 内,定义上包络线 $E_{\max}(t)$ 与下包络线 $E_{\min}(t)$ 构成的均值函数 $M(t)$ 为

$$M(t) = \frac{E_{\max}(t) + E_{\min}(t)}{2} \quad (1)$$

$M(t)$ 严格满足

$$\int_{t_0}^{t_n} |M(t)| dt \leq \gamma \quad (2)$$

式中 γ 为预设容差阈值。该条件通过消除基线漂移,保证瞬时频率具有明确的物理意义。

1.1.2 分解过程

EMD 将非平稳信号通过迭代筛选解构为一组 IMF^[8],其实现过程如下。

1) 极值包络提取

设原始信号为 $x(t)$,首先识别 $x(t)$ 的所有局部极大值点与极小值点,然后采用样条插值法分别拟合上、下包络线,通过式(1)计算得到原始信号的包络均值函数 $M_1(t)$ 。

提取初始高频分量 $H_1(t)$,表达式为

$$H_1(t) = x(t) - M_1(t) \quad (3)$$

2) 迭代筛选生成 IMF

将 $H_1(t)$ 作为一个新信号,重复步骤 1),经过 β 次迭代后得到第 β 个高频分量 $H_\beta(t)$ 为

$$H_\beta(t) = H_{\beta-1}(t) - M_\beta(t) \quad (4)$$

式中 $M_\beta(t)$ 为第 β 次迭代时的包络均值函数。

当 $H_\beta(t)$ 满足 IMF 的单分量性和对称包络性时,确定首个 IMF 分量为

$$\text{IMF}_1(t) = H_\beta(t) \quad (5)$$

3) 残差分解与循环终止

从原始信号中分离 $\text{IMF}_1(t)$ 得到残差 $r_1(t)$,表达式为

$$r_1(t) = x(t) - \text{IMF}_1(t) \quad (6)$$

将残差 $r_1(t)$ 作为新的输入信号,循环执行步骤 1) 和步骤 2),直至无法提取新 IMF 分量,过程表示为

$$r_i(t) = r_{i-1}(t) - \text{IMF}_i(t), i=2, \dots, n \quad (7)$$

最终分解结果表示为

$$x(t) = \sum_{i=1}^n \text{IMF}_i(t) + r_n(t) \quad (8)$$

式中 $r_n(t)$ 为单调残余项,是表征信号趋势的分量。

1.1.3 端点效应及其评价

由 EMD 分解过程可知,IMF 分量的提取依赖于对信号局部极值点构建的上、下包络线。然而,由于信号首尾端点通常不具备极值特性,基于样条插值生成的包络线在边界区域易产生过冲或欠冲偏差。这些现象会随着筛分过程逐层向信号内部渗透,从而导致 IMF 分量在边界区域的幅值异常波动、不同频率成分在分解过程中相互干扰。这类由边界包络误差累积导致的信号分解失真现象称为端点效应,该现象的存在严重限制了 EMD 的实际应用效果。

为对端点效应进行诊断与抑制,必须要有适当的评价指标来表征 EMD 过程中端点效应的影响与量化该问题解决方法的有效性。本文采用两个指标进行量化分析^[9-10],具体说明如下。

1) 根据能量守恒定律,原始信号的能量应与分解后所有 IMF 分量(包含残差)的能量总和相同。端点效应的实质在于包络线的过冲或欠冲现象使信号中包含了虚假成分,即 IMF 分量的能量总和会增加,因此可通过比较原始信号的能量与所有分量的能量总和来评价端点效应对分解过程的影响。

在实际计算中,信号的能量有效值(用 G_R 表示)通常采用均方根(RMS)计算,即

$$G_R = \sqrt{\frac{\sum_{t=t_0}^{t_n} [g(t)]^2}{t_n - t_0}} \quad (9)$$

式中 $g(t)$ 表示信号序列(原始信号或 IMF 分量)的能量。

计算所有信号序列的 RMS 有效值后,由下式构造能量误差的评价指标 θ 。

$$\theta = \frac{\left| \sqrt{\sum_{i=1}^n G_{Ri}^2} - G_{R_x} \right|}{G_{R_x}} \quad (10)$$

式中: G_{R_x} 表示原始信号的能量有效值; G_{Ri} 表示第 i 个 IMF 分量的能量有效值。评价指标 θ 越小,端点效应对分解过程的影响越小。

2) EMD 的分解结果在理想情况下应满足正交性,故引入正交性系数(ORT)来评价端点效应的影响,其计算式为

$$\text{ORT} = \sum_{t=t_0}^{t_n} \left[\sum_{j=1}^n \sum_{k=1}^n I_j(t) \cdot I_k(t) / x^2(t) \right] \quad (11)$$

式中 $I_j(t)$ 和 $I_k(t)$ 分别表示第 j 个和第 k 个 IMF 分量,且 $j \neq k$ 。ORT 值越小,端点效应对分解过程的影响越小。

1.2 循环神经网络模型

1.2.1 标准循环神经网络

相较于前馈神经网络的固定连接模式,循环神经网络(recurrent neural network, RNN)通过引入隐状态递归连接,建立时序特征的长程依赖关系,其数学本质可描述为

$$h_t = \sigma(W_{hh} h_{t-1} + W_{xh} x_t + b_h) \quad (12)$$

式中: h_t 表示 t 时刻的隐状态向量; $\sigma(\cdot)$ 表示非线性激活函数; W_{hh} 表示隐状态到隐状态的权重矩阵; W_{xh} 表示输入到隐状态的权重矩阵; x_t 表示在时间步 t 的输入; b_h 表示偏置项向量。这种时间维度上的权值共享机制,赋予网络对时序动态特征的记忆保持能力。

在反向传播过程中,梯度计算涉及参数矩阵的连乘运算,这会导致两种极端情况^[11]。

1) 梯度衰减效应:当激活函数的导数 $\sigma'(\cdot)$ 与权重矩阵 W_{hh} 乘积的绝对值小于 1 时,梯度范数会随着序列长度的增加呈指数衰减,导致长程依赖学习失效。

2) 梯度膨胀现象:当 $\sigma'(\cdot)$ 与权重矩阵 W_{hh} 乘积的绝对值大于 1 时,梯度范数会超线性增长,超过 10^5 时数值溢出,严重影响网络训练稳定性。

只有当 $\sigma'(\cdot)$ 与权重矩阵 W_{hh} 乘积的绝对值等于 1 时,梯度才能保持稳定。

1.2.2 长短时记忆网络

针对传统 RNN 的梯度衰减与梯度膨胀现象,LSTM 通过创新的门控设计系统与细胞状态通道双重机制,在网络架构中实现梯度流的可控传输。LSTM 中一个单元的结构如图 1 所示。

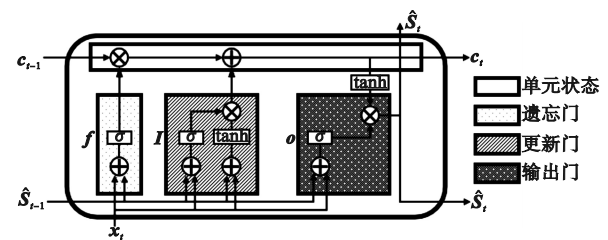


图1 LSTM结构示意图

Fig.1 Diagram of the LSTM architecture

LSTM 通过细胞状态循环机制与门控单元的协同作用,实现对时序信息的长程依赖建模,其核

心运算模块可由以下门控方程表达。

$$\mathbf{f} = \sigma(\mathbf{W}_f \hat{\mathbf{S}}_{t-1} + \mathbf{U}_f \mathbf{x}_t + \mathbf{b}_f) \quad (13)$$

$$\mathbf{I} = \sigma(\mathbf{W}_I \hat{\mathbf{S}}_{t-1} + \mathbf{U}_I \mathbf{x}_t + \mathbf{b}_I) \quad (14)$$

$$\mathbf{o} = \sigma(\mathbf{W}_o \hat{\mathbf{S}}_{t-1} + \mathbf{U}_o \mathbf{x}_t + \mathbf{b}_o) \quad (15)$$

$$\mathbf{c}_t = \mathbf{I} \cdot \tanh(\mathbf{W}_c \hat{\mathbf{S}}_{t-1} + \mathbf{U}_c \mathbf{x}_t + \mathbf{b}_c) + \mathbf{f} \cdot \mathbf{c}_{t-1} \quad (16)$$

$$\hat{\mathbf{S}}_t = \mathbf{o} \cdot \tanh(\mathbf{c}_t) \quad (17)$$

式中: $\mathbf{f}, \mathbf{I}, \mathbf{o}$ 分别表示遗忘门、输入门和输出门的输出; $\hat{\mathbf{S}}_{t-1}$ 为 $t-1$ 时间步时的预测输出; \mathbf{x}_t 为 t 时间步时的输入; $\mathbf{W}_f, \mathbf{W}_I, \mathbf{W}_o$ 分别为遗忘门、输入门、输出门的权重矩阵; $\mathbf{U}_f, \mathbf{U}_I, \mathbf{U}_o$ 分别为遗忘门、输入门、输出门的输入权重矩阵; $\mathbf{b}_f, \mathbf{b}_I, \mathbf{b}_o$ 分别为遗忘门、输入门、输出门的偏置项; \mathbf{c}_t 代表 t 时间步时记忆细胞的存储状态; \mathbf{W}_c 为细胞状态权重矩阵; \mathbf{U}_c 为细胞状态的输入权重矩阵; \mathbf{b}_c 为细胞状态的偏置项。

式(13)中,通过激活函数,利用上一层的输出 $\hat{\mathbf{S}}_{t-1}$ 和当前输入 \mathbf{x}_t 得到遗忘门输出 \mathbf{f} , 其值在 $[0, 1]$ 之间,表示上一个 LSTM 结构中信息保留的概率。输入门执行双阶段操作:式(14)中,通过 Sigmoid 层确定待更新信息的权重,即输入门输出 \mathbf{I} ; 式(16)中,利用 tanh 层生成候选信息内容,并与 \mathbf{I} 结合,用于更新细胞状态 \mathbf{c}_t 。式(15)表示输出门输出 \mathbf{o} , 式(17)通过激活函数 tanh 作用于 LSTM 结构的最新信息,并与式(15)得到的 \mathbf{o} 进行运算,控制最终的输出 $\hat{\mathbf{S}}_t$ 。

在处理信号序列问题时, LSTM 结构能更有效地协调历史记忆单元中的信息分配,比传统 RNN 更具协调和管理记忆能力^[12]。

1.3 麻雀搜索算法

SSA 作为群类智能优化领域的一种新方案,其灵感来源于麻雀群体分层寻找食物的策略。目前该算法已成功应用于神经网络超参数优化、无线传感器网络覆盖优化等工程场景^[13]。

1.3.1 基本原理

SSA 具有动态平衡的仿生优化架构,其核心机制是三个模块的协同运作。

1) 角色分工拓扑。种群被划分为探索者、开发者和警戒者三个功能群体。探索者负责广泛的全局搜索,开发者专注于局部精细搜索,警戒者则实时监测环境变化,确保种群安全。

2) 风险响应机制。当监测到环境威胁因子超出预设安全阈值时,算法自动触发协同预警响应系统。此时全体个体根据威胁等级评估结果,采用莱维飞行策略或高斯扰动模式向安全区域迁移,有效规避局部极值陷阱。

3) 动态角色转换。采用适应度驱动的角色弹性转换策略,每轮迭代中通过分位数排序动态调整个体角色归属。探索者与开发者的群体比例始终保持恒定,个体角色可根据其当前适应度表现进行双向转换。

1.3.2 优化过程

探索者位置更新策略根据环境风险评估动态调整,表达式为

$$\mathbf{X}_{i,j}^{l+1} = \begin{cases} \mathbf{X}_{i,j}^l \cdot \exp\left(\frac{-i}{\alpha \cdot T_{\max}}\right), & R_2 < ST \\ \mathbf{X}_{i,j}^l + Q \cdot \mathbf{L}, & R_2 \geq ST \end{cases} \quad (18)$$

式中: $\mathbf{X}_{i,j}^{l+1}$ 表示第 $l+1$ 次迭代中第 j 维空间的第 i 个麻雀位置; T_{\max} 表示最大迭代次数; α 为 $(0, 1]$ 之间的随机数; \mathbf{L} 为全 1 向量,即 $\mathbf{L} = [1, 1, \dots, 1]$, 其维数为空间维数的最大值; Q 为服从标准正态分布的随机量; R_2 为环境安全系数; ST 表示警戒阈值。当 $R_2 < ST$ 时,探索者在当前位置邻域实施精细开发;当 $R_2 \geq ST$ 时,监测到潜在威胁,个体采用莱维飞行策略向安全区域迁移。

开发者位置更新机制表示为

$$\mathbf{X}_{i,j}^{l+1} = \begin{cases} Q \cdot \exp(\mathbf{X}_{\text{worst}}^l - \mathbf{X}_{i,j}^l / i^2), & i > v/2 \\ \mathbf{X}_p^l + |\mathbf{X}_{i,j}^l - \mathbf{X}_p^l| \cdot \mathbf{A}^+ \cdot \mathbf{L}, & i \leq v/2 \end{cases} \quad (19)$$

式中: \mathbf{X}_p^l 是当前全局探索者所占据的最优位置; $\mathbf{X}_{\text{worst}}^l$ 表示当前全局探索者所占据的最差位置; v 为开发者总数; \mathbf{A}^+ 为伪逆矩阵(由随机正交矩阵 \mathbf{A} 生成)。开发者根据适应度排序位置采取差异化更新:对于排序在前 50% 的个体,引入伪逆矩阵 \mathbf{A}^+ 和全 1 向量 \mathbf{L} 来调整移动方向和步长,从而更有效地向最优位置靠近;当个体的适应度排序处于后 50% 时,说明该个体因资源获取能力不足触发状态迁移协议,将脱离当前区域,执行探索者行为,以寻找新资源。

警戒者动态响应模型为

$$\mathbf{X}_{i,j}^{l+1} = \begin{cases} \mathbf{X}_{\text{best}}^l + \mu \cdot |\mathbf{X}_{i,j}^l - \mathbf{X}_{\text{best}}^l|, & f_i > f_g \\ \mathbf{X}_{i,j}^l + K \cdot \frac{|\mathbf{X}_{i,j}^l - \mathbf{X}_{\text{best}}^l|}{(f_i - f_w) + \varepsilon}, & f_i = f_g \end{cases} \quad (20)$$

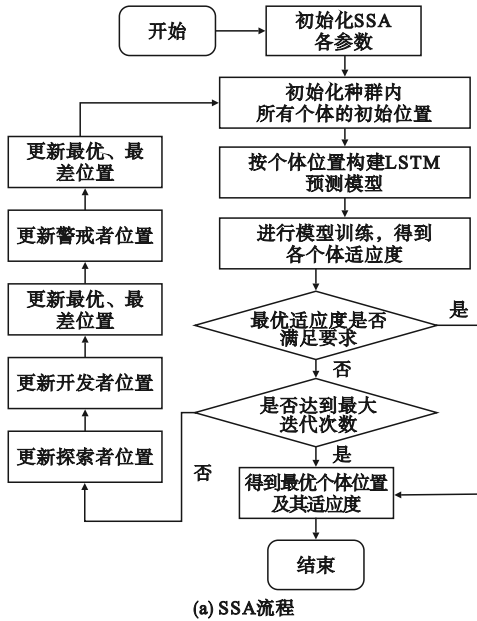
式中: $\mathbf{X}_{\text{best}}^l$ 代表当前全局警戒者所占据的最优位置; $K \sim u(-1, 1)$, 为均匀分布扰动系数; $\mu \sim N(0, 1)$, 为自适应步长因子; f_i 为麻雀个体 i 的适应度值; f_g 和 f_w 分别为当前全局最佳和最差的适应度值; ε 为极小常数。该机制通过双重策略实现威胁响应:当 $f_i > f_g$ 时,采用精英引导策略,向全局最优位置定向逼近;当 $f_i = f_g$ 时,添加极小常数 ε 以避免分母为零,并通过均匀扰动系数 K 的随机

性维持种群多样性。

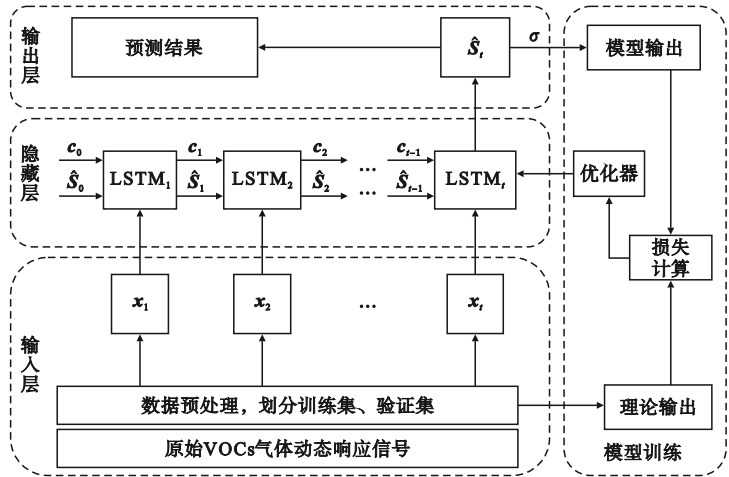
当迭代次数达到预设值 T_{\max} 时,输出全局最优解及适应度。在神经网络参数优化中,该最优解对应于权重矩阵与偏置向量的全局最优配置,可有效规避梯度下降算法的局部收敛风险,提升模型训练效率^[14]。

2 基于 SSA 优化的 LSTM 时间序列双向预测耦合模型

针对 VOCs 气体动态响应信号的时序特性,本文提出一种基于 SSA 优化的 LSTM 耦合预测框架,其结构如图 2 所示,该框架由输出层、隐藏层、输入层、模型训练和 SSA 五个核心模块组成。



(a) SSA流程



(b) LSTM时间序列预测模型

图 2 基于 SSA 的 LSTM 时间序列双向预测耦合模型结构

Fig. 2 Structure of LSTM time series bidirectional forecasting coupled model based on SSA

2.2 误差度量准则与优化策略

本文模型训练架构基于时序信号处理范式构建,具体实施步骤如下。

引入改进型均方根误差 (RMSE) 评估模型性能,用 e_{RMSE} 表示,其计算表达式为

$$e_{\text{RMSE}} = \sqrt{\frac{1}{F} \sum_{a=1}^F (P_a - Y_a)^2} \quad (22)$$

式中: P_a 和 Y_a 分别为第 a 个窗口的预测值和对应的真实值; F 为有效样本总数。

采用 Adam 优化方法进行网络参数迭代优化,使网络损失最小。该优化方法具有自动调整超参数、快速收敛特性及基于梯度一阶矩与二阶矩估计的运算优势^[15-16]。

2.1 数据预处理

原始 VOCs 气体动态响应数据的预处理包括异常值检测与剔除、信号去噪两部分。

为消除量纲差异对模型训练的负面影响,采用最大-最小归一化方法进行线性变换,表达式为

$$x_{\text{norm}}(t) = \frac{x(t) - x_{\min}}{x_{\max} - x_{\min}} \quad (21)$$

式中: $x_{\text{norm}}(t)$ 为归一化后的数据点; x_{\min} 与 x_{\max} 分别为原始信号数据中的最小值和最大值。该线性变换将原始数据投影至 $[0, 1]$ 区间,其优势在于: 维持特征分布相对关系,避免梯度消失或梯度爆炸;降低特征值跨度差异对 Adam 优化器的影响;增强时序波形的形态特征辨识度。将归一化后的数据点划分为训练集和验证集。

为避免模型训练过程中出现过拟合问题,本文在原损失函数基础上添加 L2 范数正则化项,得到的损失函数表示为^[17]

$$J = J_0 + \frac{\lambda}{2m} \sum_{i=1}^m \omega_i^2 = \frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2 + \frac{\lambda}{2m} \sum_{i=1}^m \omega_i^2 \quad (23)$$

式中: J 为改进后的损失函数; J_0 为原损失函数; λ 为正则化系数,用来权衡正则项与原始函数的比重; m 为设定的训练集样本数目; ω_i 为样本 i 的权重; y_i 和 \hat{y}_i 分别为样本 i 的实际输出和模型预测输出。

由此可得权重更新公式为

$$\omega_i \leftarrow \left(1 - \frac{\eta\lambda}{m}\right) \omega_i - \eta \frac{\partial J_0}{\partial \omega_i} \quad (24)$$

式中 η 为学习率,用于控制权重更新的步长。

基于 L2 正则化方法, LSTM 辨识策略的权重更新公式如下。

$$W_f(\tau+1) = W_f(\tau) - \eta \frac{\partial J}{\partial W_f(\tau)} \quad (25)$$

$$U_f(\tau+1) = U_f(\tau) - \eta \frac{\partial J}{\partial U_f(\tau)} \quad (26)$$

式中: W_f 、 U_f 分别为遗忘门的权重矩阵和输入权重矩阵; τ 为训练迭代次数。同理可得到输入门、输出门以及细胞状态的权重矩阵和输入权重矩阵更新公式。

2.3 双向周期延拓与复合信号生成

以训练集数据为初始状态, 通过训练好的 LSTM 模型向前预测 80 个时间步, 生成首部延拓序列; 将原始信号时序反转, 采用相同模型预测尾部延拓序列; 将双向延拓序列分别拼接至原始信号首尾, 形成复合信号, 确保 EMD 分解时边界被完整覆盖。

3 实验与结果分析

实验测试平台主要由密闭检测舱、电信号采集模块、可编程直流电源、气源装置及监控终端组成。密闭检测舱是传感器的测试空间, 测试时密闭, 以有效防止 VOCs 外逸; 电信号采集装置实时记录传感器微电流波动; 可编程直流电源精准调控器件加热参数; 气源系统配置标准混合气体 (N_2 和 O_2 的体积分数分别为 79% 和 21%), 既用于实验前的舱体净化, 也作为检测过程中的基准气源; 监控终端用于实现传感器电流变化的动态可视化监测。

实验所用检测气体为丙酮与甲苯, 两种气体均为工业生产中常见的易燃易爆气体。通过方波、三角波、矩形波动态加热信号测试进行横向对比实验, 发现方波测试对两种气体选择性较高, 故采用周期为 80 s 的方波动态加热信号进行后续实验。

进行三批次重复实验, 每批次测试气体浓度 (体积浓度, 单位为 10^{-6}) 分别取为 50、100、150、200、250、300。为去除因设备故障、操作失误或环境干扰造成的异常数据或不准确数据, 保留更符合实际测量情况的数据, 进行数据筛选。筛选标准如下:

- 1) 稳态响应电流波动系数小于 5%;
- 2) 基线漂移率小于 0.1% / min;
- 3) 动态响应重复性误差小于 8%。

以丙酮与甲苯气体浓度为 50×10^{-6} 时的动态响应信号为例, 采用基于 SSA 的 LSTM 时间序列

预测, 得到原信号的双向性周期预测信号, 计算所用时间为 0.379 s。图 3 为以丙酮为例的双向周期延拓前后信号序列对比图。

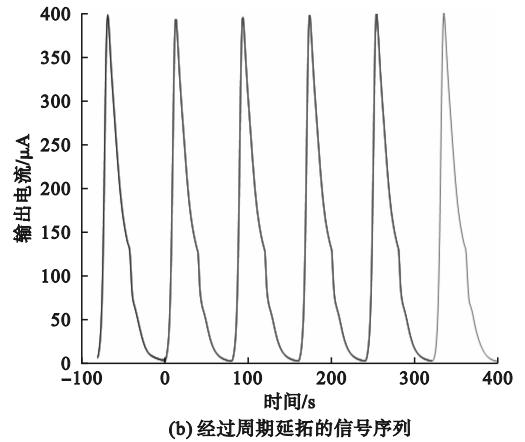
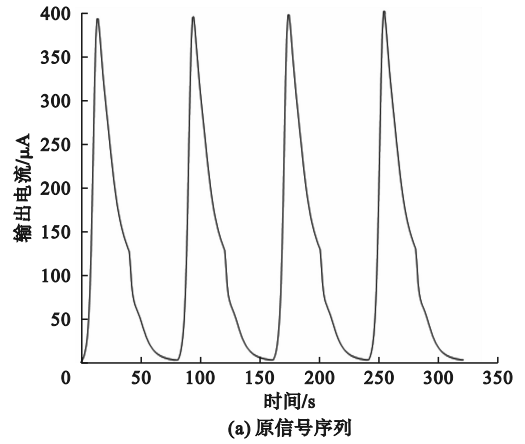


图 3 原信号序列和经过周期延拓的信号序列对比
Fig. 3 Comparison of original signal sequence and periodically extended signal sequences

采用本文端点效应评价指标对预测结果进行量化, 如表 1 所示。

表 1 不同气体信号 EMD 分解的端点效应评价指标
Table 1 Evaluation metrics for endpoint effects in EMD decomposition of different gas signals

分解方法	θ	ORT 值
EMD (丙酮, 下同)	0.216 48	0.217 17
经 LSTM 预测后 EMD	0.081 24	0.155 89
经 LSTM-SSA 预测后 EMD	0.054 19	0.105 45
EMD (甲苯, 下同)	0.216 02	0.205 20
经 LSTM 预测后 EMD	0.166 32	0.168 63
经 LSTM-SSA 预测后 EMD	0.165 54	0.133 40

由表 1 可以看出: 从能量守恒角度, 丙酮与甲苯信号经 LSTM-SSA 预测后进行 EMD 分解的 θ 值比直接进行 EMD 分解降低了 74.966% 和 23.368%, 较经 LSTM 预测后进行 EMD 分解降低

了 33.295% 和 0.469% ;从分解正交性角度,经 LSTM-SSA 预测后进行 EMD 分解的 ORT 值比直接进行 EMD 分解降低了 51.444% 和 34.990% ,较经 LSTM 预测后进行 EMD 分解降低了 32.356% 和 20.891% 。可见,经 LSTM-SSA 预测后进行 EMD 分解显著抑制了 IMF 分量的模态混叠现象,有效解决了端点效应问题导致的序列信号 EMD 分解结果失真现象。

考虑到实际工矿企业中存在复杂的噪声干扰,以丙酮为例,在其信号中加入噪声强度(标准差)分别为 0.1 和 0.5 的高斯噪声,应用本文模型进行双向周期延拓,得到端点效应评价指标如表 2 所示。

表 2 噪声干扰下的端点效应评价指标

Table 2 Evaluation metrics for endpoint effects under noise interference

分解方法	θ	ORT 值
EMD(噪声强度 0.1,下同)	0.180 59	0.228 57
经 LSTM 预测后 EMD	0.061 98	0.120 12
经 LSTM-SSA 预测后 EMD	0.061 90	0.119 98
EMD(噪声强度 0.5,下同)	0.184 02	0.328 28
经 LSTM 预测后 EMD	0.166 15	0.304 70
经 LSTM-SSA 预测后 EMD	0.062 03	0.120 21

由表 2 可见,在信号存在高斯噪声的情况下,无论是 LSTM 模型还是 LSTM-SSA 耦合模型,都可从能量守恒和分解正交性的角度提升 EMD 分解结果的可信度。但明显可以看出:提升噪声强度后,仅经过 LSTM 预测的 EMD 分解评价指标出现了恶化, θ 值从 0.061 98 升至 0.166 15,ORT 值从 0.120 12 升至 0.304 70,表明其在强噪声条件下的抑制端点效应能力有限;LSTM-SSA 模型的 θ 值在提升噪声强度后仍保持在 0.062 00 左右,ORT 值仅略上升至 0.120 21,说明其对噪声鲁棒性更强,能更稳定地抑制端点效应。

综上所述,对于周期信号,采用基于 SSA 的 LSTM 预测能够减弱信号在 EMD 分解过程中的端点效应,即减小了信号分解过程中的失真程度,并对噪声干扰具有鲁棒性。

4 结论

提出了一种基于 LSTM-SSA 的耦合模型,对时间序列信号进行双向周期延拓,并确定了 EMD 端点效应的诊断方法。以浓度为 50×10^{-6} 的丙酮

和甲苯气体动态响应信号时间序列为样本进行模型训练,得到结论如下。

1) 经 LSTM-SSA 预测后再进行 EMD 分解,生成的延拓序列与原始信号动态特性高度匹配,且分解过程能显著抑制端点处模态分量的幅值失真与模态混叠。丙酮和甲苯信号分解结果的 θ 值分别降低了 74.966% 和 23.368% ,ORT 值分别降低了 51.444% 和 34.990% ,分解结果的可靠性得到了显著提升。

2) 在原始信号中引入不同强度的高斯噪声,对比分析可知 LSTM-SSA 耦合模型较单一 LSTM 模型对噪声干扰的鲁棒性更强。

3) 创新性地构建了 LSTM-SSA 耦合模型,利用 SSA 优化 LSTM 超参数,解决了传统梯度下降算法易陷入局部最优的问题,通过引入 L2 正则化项改进损失函数,进一步提升了预测精度,为周期信号延拓提供了高效解决方案。

4) 本文模型在 VOCs 气体动态信号处理中展现出较强的预测与泛化能力,通过信号分解与 LSTM 神经网络预测的协同设计,有效减弱了 EMD 端点效应,提升了气体传感特征提取的可应用性。

参考文献(References):

- [1] 梁小明. 中国人为源 VOCs 排放清单改进及其特征研究 [D]. 广州:华南理工大学,2022.
- [2] DJEDIDI O, DJEZIRI M A, MORATI N, et al. Accurate detection and discrimination of pollutant gases using a temperature modulated MOX sensor combined with feature extraction and support vector classification [J]. Sensors and Actuators B: Chemical, 2021, 339: 129817.
- [3] GOHARI M, GHORBANI M H. Classification and diagnosis of the rotor unbalance parameters via hybridized EMD and RF [J]. Journal of Vibration Engineering & Technologies, 2025, 13(4): 246.
- [4] 钱欢,蒋沅,王海涛. 基于灰色预测模型的 EMD 端点效应抑制方法 [J]. 南昌航空大学学报(自然科学版), 2021, 35(2): 21-26.
QIAN H, JIANG Y, WANG H T. A method to suppress the end effect of EMD based on GM(1,1) prediction model [J]. Journal of Nanchang Hangkong University (Natural Sciences), 2021, 35(2): 21-26. (in Chinese)
- [5] 张雨萍,张志刚,周庆华,等. EMD 端点效应问题解决方法的研究 [J]. 信息技术, 2024, 48(1): 52-58, 64.
ZHANG Y P, ZHANG Z G, ZHOU Q H, et al. Research on processing method of end effect of empirical mode decomposition [J]. Information Technology, 2024, 48(1): 52-58, 64. (in Chinese)
- [6] 贺智,王强,沈毅,等. 希尔伯特-黄变换端点效应抑制算法综述 [J]. 软件, 2011, 32(10): 1-7.
HE Z, WANG Q, SHEN Y, et al. Survey on end effects mitigation of Hilbert-Huang transform [J]. Software, 2011, 32(10): 1-7. (in Chinese)

- [7] HUANG N E, SHEN Z, LONG S R. A new view of nonlinear water waves; the Hilbert spectrum [J]. *Annual Review of Fluid Mechanics*, 1999, 31 : 417 - 457.
- [8] 李超. 改善经验模态分解端点问题的方法 [J]. *石油地质与工程*, 2024, 38 (2) : 72 - 76.
- LI C. Methods to improve the EMD endpoint problems [J]. *Petroleum Geology and Engineering*, 2024, 38 (2) : 72 - 76. (in Chinese)
- [9] 任达千, 吴昭同, 严拱标. EMD 端点效应的评价指标及抑制端点效应的窗函数法 [J]. *制造业自动化*, 2007, 29 (1) : 21 - 24.
- REN D Q, WU Z T, YAN G B. Evaluation of the EMD end effect and its window based method [J]. *Manufacturing Automation*, 2007, 29 (1) : 21 - 24. (in Chinese)
- [10] 郭迪, 沈洋洋, 尹兵. 经验模式分解端点效应抑制方法的研究 [J]. *计算机技术与发展*, 2016, 26 (3) : 89 - 92, 96.
- GUO D, SHEN Y Y, YIN B. Research on method for end effects reduction of empirical mode decomposition [J]. *Computer Technology and Development*, 2016, 26 (3) : 89 - 92, 96. (in Chinese)
- [11] 陈亮, 裴晓辉, 刘韵婷. 基于 LSTM 的大棚环境变量预测 [J]. *沈阳理工大学学报*, 2018, 37 (5) : 13 - 19.
- CHEN L, PEI X H, LIU Y T. Prediction of greenhouse environment variable based on LSTM [J]. *Journal of Shenyang Ligong University*, 2018, 37 (5) : 13 - 19. (in Chinese)
- [12] 王炜, 刘宏伟, 陈永杰, 等. 基于 LSTM 循环神经网络的风力发电预测 [J]. *可再生能源*, 2020, 38 (9) : 1187 - 1191.
- WANG W, LIU H W, CHEN Y J, et al. Wind power forecast based on LSTM cyclic neural network [J]. *Renewable Energy Resources*, 2020, 38 (9) : 1187 - 1191. (in Chinese)
- [13] 金秀章, 张瑾, 陈佳政. 基于麻雀搜索算法优化 TPA-LSTM 的火电厂 NO_x 排放预测 [J/OL]. *控制工程*, 2024; 1 - 9 (2024 - 03 - 18) [2025 - 03 - 15]. <https://link.cnki.net/doi/10.14107/j.cnki.kzgc.20230640>.
- JIN X Z, ZHANG J, CHEN J Z. NO_x emission prediction of thermal power plant based on sparrow search algorithm optimized TPA-LSTM [J/OL]. *Control Engineering of China*, 2024; 1 - 9 (2024 - 03 - 18) [2025 - 03 - 15]. <https://link.cnki.net/doi/10.14107/j.cnki.kzgc.20230640>. (in Chinese)
- [14] 张恒, 张雷, 姚海峰, 等. 基于麻雀搜索算法优化 BP 人工神经网络的短期湍流预报模型研究 [J]. *长春理工大学学报 (自然科学版)*, 2024, 47 (2) : 58 - 65.
- ZHANG H, ZHANG L, YAO H F, et al. Research on short-term turbulence forecasting model based on sparrow search algorithm to optimize BP artificial neural network [J]. *Journal of Changchun University of Science and Technology (Natural Science Edition)*, 2024, 47 (2) : 58 - 65. (in Chinese)
- [15] SHAO Y C, YANG J P, ZHOU W, et al. An improvement of Adam based on a cyclic exponential decay learning rate and gradient norm constraints [J]. *Electronics*, 2024, 13 (9) : 1778.
- [16] 覃聪, 田仲初, 马连峰, 等. Adam 优化神经网络的连续刚构桥施工线形预测 [J]. *交通科学与工程*, 2025, 41 (1) : 98 - 104, 139.
- QIN C, TIAN Z C, MA L F, et al. Construction alignment prediction of continuous rigid frame bridge based on Adam-optimized neural network [J]. *Journal of Transport Science and Engineering*, 2025, 41 (1) : 98 - 104, 139. (in Chinese)
- [17] 徐宝昌, 吕爽, 董秀娟, 等. 基于 L2 正则化 LSTM 的非线性动态系统辨识 [J]. *化工自动化及仪表*, 2021, 48 (1) : 1 - 7.
- XU B C, LV S, DONG X J, et al. Identification of nonlinear dynamic system based on L2 regularized LSTM [J]. *Control and Instruments in Chemical Industry*, 2021, 48 (1) : 1 - 7. (in Chinese)

(责任编辑: 宋颖韬)

(上接第 35 页)

- [12] 孙柱, 吕宪勇, 孙凯信, 等. 基于 Dubins 曲线的轮式机旋耕作业 CCP 算法 [J]. *农业装备与车辆工程*, 2024, 62 (5) : 9 - 13.
- SUN Z, LÜ X Y, SUN K X, et al. Full coverage path planning for wheeled machine rotary tillage operation based on Dubins curve [J]. *Agricultural Equipment & Vehicle Engineering*, 2024, 62 (5) : 9 - 13. (in Chinese)
- [13] 马佳宇, 谭功全, 郑钊, 等. 基于改进 A* 算法的分层平滑安全路径规划 [J/OL]. *重庆工商大学学报 (自然科学版)*, 2024; 1 - 13 [2025 - 03 - 15]. <https://kns.cnki.net/kcms/detail/50.1155.N.20241204.1104.002.html>.
- MA J Y, TAN G Q, ZHENG Z, et al. Hierarchical smooth secure path planning based on improved A* algorithm [J/OL]. *Journal of Chongqing Technology and Business University (Natural Science Edition)*, 2024; 1 - 13 [2025 - 03 - 15]. <https://kns.cnki.net/kcms/detail/50.1155.N.20241204.1104.002.html>. (in Chinese)
- [14] 朱洪波, 殷宏亮. 煤矿救援机器人路径规划研究 [J]. *工矿自动化*, 2024, 50 (12) : 145 - 154.
- ZHU H B, YIN H L. Research on path planning for coal mine rescue robots [J]. *Journal of Mine Automation*, 2024, 50 (12) : 145 - 154. (in Chinese)
- [15] 谢春丽, 陶天艺. 基于混合 A* 算法的机器人路径规划研究 [J/OL]. *南京信息工程大学学报*, 2024; 1 - 13 [2025 - 03 - 15]. <https://link.cnki.net/doi/10.13878/j.cnki.jnuist.20240415001>.
- XIE C L, TAO T Y. Research on path planning of mobile robot based on hybrid A* algorithm [J/OL]. *Journal of Nanjing University of Information Science & Technology*, 2024; 1 - 13 [2025 - 03 - 15]. <https://link.cnki.net/doi/10.13878/j.cnki.jnuist.20240415001>. (in Chinese)
- [16] 林硕, 金恒江, 韩忠华, 等. 随机地图下改进 A* 路径规划算法研究 [J]. *制造技术与机床*, 2025 (2) : 43 - 47.
- LIN S, JIN H J, HAN Z H, et al. Improved A* path planning algorithm based on stochastic map [J]. *Manufacturing Technology & Machine Tool*, 2025 (2) : 43 - 47. (in Chinese)
- [17] 吴庭明. 移动机器人路径规划算法研究 [D]. 西安: 西安电子科技大学, 2023.
- [18] FU H Y, SHAO G Y, ZHAO L. Global path planning of unmanned ship island region based on improved A* algorithm [J]. *Journal of Physics: Conference Series*, 2024, 2858 (1) : 012021.

(责任编辑: 宋颖韬)