第 50 卷 第 1 期：263-274 高电压技术 Vol.50, No.1: 263-274
2024 年 1 月 31 日 High Voltage Engineering January 31, 2024
DOI: 10.13336/j.1003-6520.hve.20221805

基于相关变分模态分解和 CNN-LSTM 的变压器油中溶解气体体积分数预测

范志远[1,2]，杜　江[1,2]
（1. 河北省电磁场与电器可靠性重点实验室(河北工业大学)，天津 300130;
2. 省部共建电工装备可靠性与智能化国家重点实验室(河北工业大学)，天津 300130）

摘　要：为解决变压器油中溶解气体实际监测数据中噪声信号对模型预测性能的影响以及单一长短期记忆神经网络(long short-term memory，LSTM)无法对数据间的深层特征进行有效提取的问题，提出了一种融合了相关变分模态分解(correlation variational mode decomposition，CVMD)、1 维卷积神经网络(one dimensional convolutional neural network，1D-CNN)和 LSTM 的组合预测模型。首先，利用 CVMD 去除原始气体序列中的噪声信号，并将去噪序列分解为 1 组相对平稳的子序列分量；然后，针对各子序列分量分别构建 CNN-LSTM 预测模型，利用 1D-CNN 挖掘数据间的深层特征形成特征向量，并将其输入到 LSTM 中进行预测；最后，对各子序列预测结果叠加重构，得到最终的气体预测值。并通过 4 组对比实验对所提模型进行了全方位、多角度的验证。算例研究结果表明，所提模型单步和多步预测的平均绝对百分比误差分别为 1.53%和 2.09%。相较于现有模型，该文所提模型在单步和多步预测性能上均有明显提升，为变压器在线监测和故障预警提供了重要技术支撑。

关键词：油中溶解气体；相关变分模态分解；1 维卷积神经网络；长短期记忆神经网络；气体体积分数预测

Prediction of Dissolved Gas Volume Fraction in Transformer Oil Based on Correlation Variational Mode Decomposition and CNN-LSTM

FAN Zhiyuan[1,2], DU Jiang[1,2]
(1. Key Laboratory of Electromagnetic Field and Electrical Apparatus Reliability of Hebei Province, Hebei University of Technology, Tianjin 300130, China; 2. State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology, Tianjin 300130, China)

Abstract：To address the issues of noise interference in actual monitoring data of dissolved gases in transformer oil and the ineffective extraction of deep features between data by a single long short-term memory (LSTM) neural network, a combined prediction model combining correlation variational mode decomposition (CVMD), 1D convolutional neural network (1D-CNN) and LSTM is proposed. Firstly, CVMD is used to remove the noise signal from the original gas sequence, and the denoising sequence is decomposed into a group of relatively stable subsequence components. Then, CNN-LSTM prediction models are constructed for each subsequence component, and 1D-CNN is used to mine deep features between data to form feature vectors, which is input into LSTM for prediction. Finally, the prediction results of each sub-sequence is superimposed and reconstructed to obtain the final gas predicted value. Four groups of comparison experiments were carried out to verify the proposed model in a comprehensive and multi-angle way. The results of the example study show that the average absolute percentage errors of single-step and multi-step prediction of the proposed model are 1.53% and 2.09%, respectively. Compared with the existing model, the proposed model in this paper has significantly improved the performance of the single-step and multi-step prediction. The study provides important technical support for transformer on-line monitoring and fault warning.

Key words：dissolved gas in oil; correlation variational mode decomposition; one dimensional convolutional neural network; long short-term memory; gas volume fraction prediction

基金资助项目：国家自然科学基金 (52007047)；天津市自然科学基金重点项目 (19JCZDJC32100)；河北省自然科学基金 (E2018202282)。
Project supported by National Natural Science Foundation of China (52007047), Key Project of Tianjin Natural Science Foundation (19JCZDJC32100), Natural Science Foundation of Hebei Province (E2018202282).

0 引言

作为电网中最关键的枢纽设备之一，变压器的运行状态关系到整个电网的运行可靠性。油中溶解气体是表征变压器运行状态的重要特征量，通过对变压器油中溶解气体的体积分数进行预测，可以提前评估变压器的运行状态，进而引导运行维护人员制定相应检修策略，避免重大事故的发生[1-2]。

对于变压器油中溶解气体体积分数预测，通过对油中溶解气体体积分数历史数据的分析，实现对未来气体体积分数的精准预测。由于变压器工作环境复杂，收集到的数据中通常会掺杂一定程度的噪声信号，使得数据呈现高度非线性及非平稳性[3]。因此，如何有效地进行数据去噪，成为研究人员必须面对的问题。此外，变压器油中溶解气体体积分数预测本质上是一种时间序列预测问题，而传统预测方法往往存在对于非线性时间序列的拟合能力较差，且在大量数据样本下无法有效挖掘数据间的深层特征，具有较大的局限性。

在数据去噪方面，文献[4-5]指出，噪声信号有两个主要特点：一是能量主要集中在高频；二是频谱分布于一个有限的区间。针对以上两个特点，国内外学者进行了一系列探索研究，提出了许多有效的噪声削减方法，主要包括经验模态分解(empirical mode decomposition，EMD)和变分模态分解(variational mode decomposition，VMD)等。EMD 可根据原始数据本身特征将数据分解为多个本征模态分量(intrinsic mode functions，IMF)，实现了数据的自适应分解，因此，EMD 被广泛应用于数据去噪领域，并取得了显著成效。文献[6]利用 EMD 进行数据去噪，准确地呈现出了数据的特征信息，有效抑制了噪声信号并且较完整地保留了有效数据。尽管 EMD 有良好的去噪效果，但由于 EMD 在分解过程中极易出现模态混叠以及端点效应，在一定程度上影响了 EMD 的去噪效果。进一步，为克服上述不足，有学者提出了 VMD 方法。VMD 是一种完全非递归的数据分解方法，采用非递归方法不会传递误差，因此其有效抑制了模态混叠和端点效应。文献[7]首次将 VMD 引入到变压器油中溶解气体体积分数预测领域，其将气体体积分数序列分解为多个周期性和平稳性都较强的 IMF，有效抑制了原始数据中的噪声信号，并同支持向量机(support vector machine，SVM)相结合，实现了变压器油中溶解气体体积分数的预测。尽管 VMD 解决了 EMD 的不足之处，但其仍然存在噪声信号提取不充分、无法判断各 IMF 中噪声信号含量等问题，严重影响了 VMD 的噪声抑制效果。相关变分模态分解(correlation VMD，CVMD)是 VMD 的一种改进方法，其将相关系数与 VMD 相结合，利用 VMD 分解实现噪声信号与有效数据的初步分离，进而利用 IMF 与原始序列之间的相关系数来实现数据去噪。文献[8]将 CVMD 应用于全球定位系统和重力恢复与气候实验的数据去噪中，有效克服了 VMD 的局限性，实现了更好的去噪效果。

在预测方法方面，较常用的方法有 SVM、人工神经网络等。文献[9]提出一种基于时间序列和 SVM 的变压器油中溶解气体体积分数预测模型，实现对变压器故障的早期预警。尽管 SVM 可在小样本、贫信息数据情况下取得较高的预测精度，但随着数据量不断增加，其逐渐出现误差累积、无法准确提取数据时序相关性等问题[10]，大大限制了其应用范围。文献[11]利用长短期记忆神经网络(long short-term memory，LSTM)实现变压器油中溶解气体体积分数预测，解决了 SVM 预测过程中存在的误差积累以及无法准确提取数据时序相关性等问题。然而，单一 LSTM 模型依然存在不能有效提取数据间深层特征的缺陷，限制了其预测精度进一步提升。卷积神经网络(convolutional neural network，CNN)因其自身的结构优势，能够逐层挖掘数据间的抽象规律，实现数据的深度挖掘。文献[12]利用 CNN 挖掘数据间的深层特征，并与 LSTM 相结合，对时序数据进行预测，结果表明，利用 CNN 进行特征提取可加快预测模型的收敛速度，进一步提高预测精度。

综上所述，本文将 CVMD 同 CNN-LSTM 相结合，提出了一种基于 CVMD-CNN-LSTM 的变压器油中溶解气体体积分数预测模型，该模型利用 CVMD 去除变压器油中溶解气体原始序列中的噪声信号，并将其分解为一组平稳的子序列分量，有效克服了传统 VMD 无法判断各 IMF 中噪声信号的含量，进而导致噪声信号提取不充分的问题。此外，该模型还利用 CNN 的结构优势，挖掘油中溶解气体连续数据与非连续数据之间的深层特征，并形成特征向量输入到 LSTM 中，解决了单一 LSTM 模型对数据特征信息提取不充分的缺陷。实例分析表明，本文所提模型能够在去除噪声的同时保留原始序列

中的有用信息，并能对这些信息进行深度挖掘，使得模型能够以更快的收敛速度获得更高的预测精度，大大提高了模型的预测稳定性与可靠性。

1　CVMD-CNN-LSTM 预测模型的基本原理

1.1　相关变分模态分解 CVMD

1.1.1　VMD 的基本原理

VMD 算法是一种基于频域的非递归数据分解方法，其通过求解变分问题来提取各模态分量，主要包括变分问题的构造以及求解，其过程如下[13]：

1）构造变分问题。基于 Hilbert 变换、Wiener 滤波与外差解调构造变分问题，用来寻求 k 个具有中心频率的有限带宽，使每个模态的带宽估计之和最小。带约束的变分问题模型如下：

$$\begin{cases}\min\limits_{\{u_k\},\{\omega_k\}}\left\{\sum\limits_k\left\|\partial_t\left(\left(\delta(t)+\dfrac{\mathrm{j}}{\pi t}\right)u_k(t)\right)\mathrm{e}^{-\mathrm{j}\omega_k t}\right\|_2^2\right\}\\ \sum\limits_k u_k=f\end{cases} \tag{1}$$

式中：u_k 为各模态分量；t 为时间；ω_k 为各模态分量中心频率；$\delta(t)$为冲激函数；∂_t 为对 t 求偏导；f 为原始序列。

2）求解变分问题。通过引入 Lagrange 乘子 $\lambda(t)$ 与二次惩罚因子 α，将约束性变分问题转化为非约束性变分问题：

$$\begin{aligned}L(\{u_k\},\{\omega_k\},\lambda)=\alpha\sum_k\left\|\partial_t\left(\left(\delta(t)+\frac{\mathrm{j}}{\pi t}\right)u_k(t)\right)\mathrm{e}^{-\mathrm{j}\omega_k t}\right\|_2^2+\\ \left\|f(t)-\sum_k u_k(t)\right\|_2^2+\left\langle\lambda(t),f(t)-\sum_k u_k(t)\right\rangle\end{aligned} \tag{2}$$

式中：$\|*\|_2^2$ 为梯度二范数运算；<*>为约束附加项。

利用交替方向乘子法求取式(2)的鞍点，进而得到函数的最优解。

3）经 VMD 分解后，原始序列 f 被分解 k 个模态的叠加，如下：

$$f=M_1+M_2+\ldots+M_k \tag{3}$$

式中：M_i 为第 i 个模态，其中，i=1,2,⋯,k。

1.1.2　CVMD 的基本原理与实现过程

由于传统 VMD 在数据分解过程中仍然会产生较大的精度损失，因此，Y. Shen 和 W. Zheng 等将相关系数与 VMD 相结合，提出一种新的相关变分模态分解(CVMD)。CVMD 通过预分解实现噪声信号与有效数据的初步分离，之后利用 IMF 与原始序列间的相关系数判断各 IMF 中噪声信号的含量，然后通过设置合理的去噪阈值实现数据去噪，最后，将去噪序列作为输入对特征序列进行二次分解。

CVMD 的结构如图 1 所示，其主要过程如下：

1）预分解。对原始序列进行一次变分模态分解，采用中心频率观察法[14]确定 k_1 值，将原始序列分解为 k_1 个 IMF。

2）构造去噪序列。计算各 IMF 与原始序列之间的相关系数 R_i，计算式如下：

$$R_i=\frac{\mathrm{Cov}[f,M_i]}{\sqrt{\mathrm{Var}[f]\times\mathrm{Var}[M_i]}},\ i=1,2,\cdots,k_1 \tag{4}$$

式中：Cov 为协方差运算；Var 为方差运算。

根据 R_i 值及表 1 所描述的相关系数与相关性的对应关系[15]，可得各 IMF 与原始序列间的相关性。

根据相关系数大小判断各 IMF 中噪声信号的含量，并利用各 IMF 的能谱指数确定去噪阈值 c_c，使得噪声信号可以被有效去除，将剩余 IMF 叠加，得到去噪序列 f_{denoise}，其表达式如下：

$$f_{\mathrm{denoise}}=\sum M_i\ ,\ i|c_c<R_i<1 \tag{5}$$

3）二次分解。将去噪序列作为输入，进行二次变分模态分解，利用中心频率观察法确定 k_2 的值，将去噪序列分解成 k_2 个 IMF。

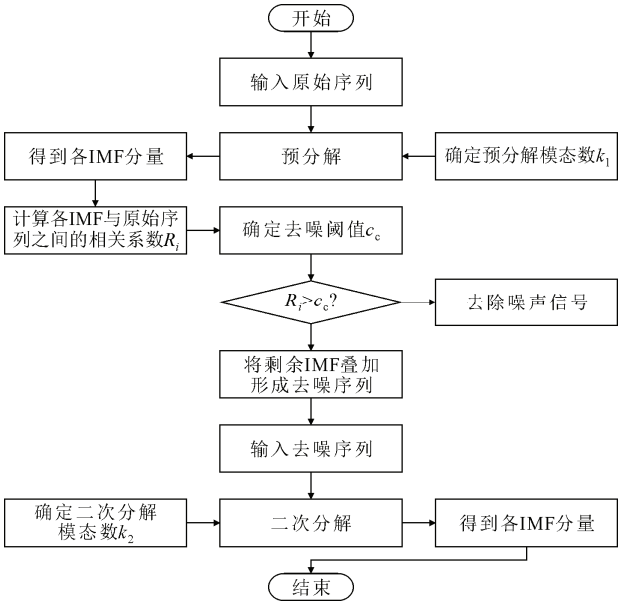


图 1　CVMD 流程图

Fig.1　Flow chart of CVMD

表 1　相关系数与相关性的关系

Table 1　Relationship between the correlation coefficient and correlation

条件	不相关	弱相关	中度相关	强相关
R_i	0~0.1	0.1~0.3	0.3~0.5	0.5~1

1.2 卷积神经网络 CNN

卷积神经网络(convolutional neural network，CNN)是深度学习的代表算法之一，作为具有深度结构的前馈神经网络，其主要由对数据进行卷积计算的卷积层和对网络参数进行下采样的池化层组成，采用局部感知与权值共享的方式深度挖掘数据间的内在联系，进而对数据的深层特征进行有效提取。其中，针对时间序列数据，主要采用 1 维卷积神经网络(one dimensional CNN，1D-CNN)提取特征，其结构图如图 2 所示，数学模型[16-17]描述如下：

$$x_j^l = f\left(\sum_{i=1}^{N} x_j^{l-1} \times k_{ij}^l + b_j^l\right) \tag{6}$$

式中：x_j^l 为第 l 层的第 j 个特征映射；f(*)为激活函数；N 为输入特征映射数量；x_j^{l-1} 为第 l–1 层的第 j 个特征映射；k_{ij}^l 为可训练的卷积核；b_j^l 为偏置项。

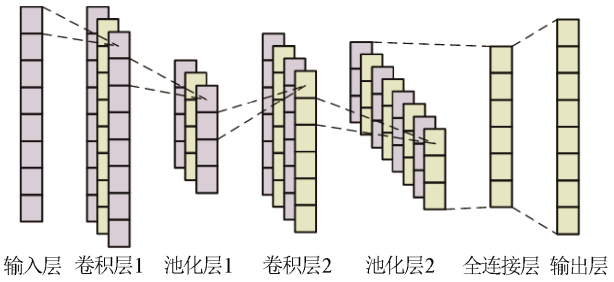


图 2 1D-CNN 结构图

Fig.2 Structure diagram of 1D-CNN

1.3 长短期记忆神经网络 LSTM

1.3.1 LSTM 的基本原理

LSTM 是目前最成熟的循环神经网络(recurrent neural network，RNN)架构，其在 RNN 的基础上增加了 3 个单独的门控单元来控制信息传递，使其对长时间序列数据具有更强的学习能力，有效解决了 RNN 对较长时间跨度序列进行建模时出现的梯度消失和梯度爆炸问题，被广泛应用于时间序列预测领域[18]。

LSTM 由遗忘门、输入门和输出门组成，其网络结构如图 3 所示。遗忘门是将上一时刻神经元的状态选择性遗忘，在保留有用信息的同时避免了无用信息的向后传递；输入门是读取数据并利用激活函数将数据映射至[0,1]之间；输出门是将处理后的数据向下一时刻传递。

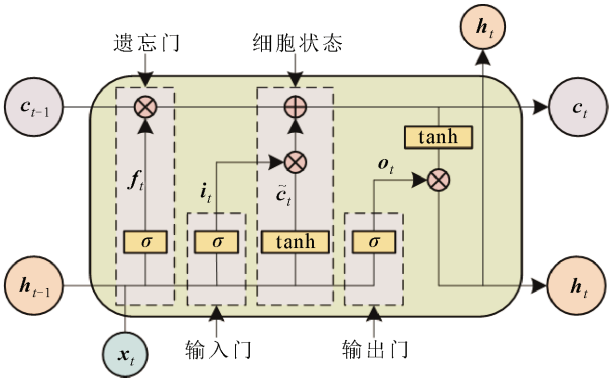


图 3 LSTM 结构图

Fig.3 Structure diagram of LSTM

图 3 中，LSTM 在当前时刻 t 的输入包括输入序列 $\boldsymbol{x}_t$、t–1 时刻的单元状态 $\boldsymbol{c}_{t-1}$ 和 t–1 时刻的中间状态 $\boldsymbol{h}_{t-1}$，输出包括当前时刻 t 的单元状态 $\boldsymbol{c}_t$ 和中间状态 $\boldsymbol{h}_t$，具体计算式如下：

$$\boldsymbol{f}_t = \sigma\left(\boldsymbol{W}_{\mathrm{f}} \cdot \left[\boldsymbol{h}_{t-1}, \boldsymbol{x}_t\right] + \boldsymbol{b}_{\mathrm{f}}\right) \tag{7}$$

$$\boldsymbol{i}_t = \sigma\left(\boldsymbol{W}_{\mathrm{i}} \cdot \left[\boldsymbol{h}_{t-1}, \boldsymbol{x}_t\right] + \boldsymbol{b}_{\mathrm{i}}\right) \tag{8}$$

$$\boldsymbol{o}_t = \sigma\left(\boldsymbol{W}_{\mathrm{o}} \cdot \left[\boldsymbol{h}_{t-1}, \boldsymbol{x}_t\right] + \boldsymbol{b}_{\mathrm{o}}\right) \tag{9}$$

$$\tilde{\boldsymbol{c}}_t = \tanh\left(\boldsymbol{W}_{\mathrm{c}} \cdot \left[\boldsymbol{h}_{t-1}, \boldsymbol{x}_t\right] + \boldsymbol{b}_{\mathrm{c}}\right) \tag{10}$$

$$\boldsymbol{c}_t = \boldsymbol{f}_t \odot \boldsymbol{c}_{t-1} + \boldsymbol{i}_t \odot \tilde{\boldsymbol{c}}_t \tag{11}$$

$$\boldsymbol{h}_t = \boldsymbol{o}_t \odot \tanh\left(\boldsymbol{c}_t\right) \tag{12}$$

式中：$\boldsymbol{f}_t$、$\boldsymbol{i}_t$、$\boldsymbol{o}_t$ 为遗忘门、输入门和输出门在当前时刻 t 的状态计算结果；$\boldsymbol{W}_{\mathrm{f}}$、$\boldsymbol{W}_{\mathrm{i}}$、$\boldsymbol{W}_{\mathrm{o}}$ 为遗忘门、输入门和输出门的权重矩阵；$\boldsymbol{b}_{\mathrm{f}}$、$\boldsymbol{b}_{\mathrm{i}}$、$\boldsymbol{b}_{\mathrm{o}}$ 为遗忘门、输入门和输出门的偏置项；$\tilde{\boldsymbol{c}}_t$ 为当前时刻 t 的单元状态；$\boldsymbol{W}_{\mathrm{c}}$ 为单元状态权重矩阵；$\boldsymbol{b}_{\mathrm{c}}$ 为单元状态偏置项；σ 为 sigmoid 激活函数；tanh 为双曲正切激活函数；$\odot$ 为各元素按位置相乘。

1.3.2 LSTM 关键参数优化

LSTM 存在明显的参数依赖问题，参数选取是否合理直接影响其预测性能。在 LSTM 众多参数中，最大迭代次数、初始学习率以及隐藏层神经元数对模型的训练和拟合效果影响最大，通常情况下这些参数依靠人工经验来确定，主观性和不确定性较强，严重影响了模型的预测准确性。

针对上述问题，文献[19-20]分别采用粒子群优化算法(particle swarm optimization，PSO)和改进果蝇优化算法(improved fruit fly optimization algorithm，IFOA）对 LSTM 网络中相关超参数进行优化，有效克服了依据经验选取参数导致模型预测性能不佳的问题。因此，为了正确选择上述 3 个关键参数的值，本文参照上文思路，采用鹈鹕优化算法(pelican optimization algorithm，POA)对这 3 个关键参数进行寻优，进一步提升模型的预测准确性，同

时使预测结果更加客观。优化过程参照文献[21]，3 个参数的优化范围如表 2 所示。

表 2　LSTM 关键参数优化范围

Table 2　Optimization range of key parameters of LSTM

参数定义	参数符号	参数优化范围
最大迭代次数	MaxEpochs	[10, 200]
初始学习率	InitialLearnRate	[0.001, 0.02]
隐藏层神经元数	numHiddenUnits	[10, 200]

2　CVMD-CNN-LSTM 预测模型的实现过程

本文将数据去噪算法与神经网络模型相结合，提出了一种基于 CVMD-CNN-LSTM 的组合预测模型，模型结构如图 4 所示，具体实现步骤如下：

1）按照文献[22]中的方法对变压器油中溶解气体原始数据进行预处理。

2）对数据进行预分解，采用中心频率观察法确定分解模态数并计算各 IMF 与原始序列之间的相关系数，根据相关系数的大小判断各 IMF 中噪声信号的含量，并同去噪阈值结合实现数据去噪，构建去噪数据集。

3）对去噪数据进行二次分解，同样采用中心频率观察法确定分解模态数。

4）采用增强 Dickey-Fuller 检验(augmented Dickey-Fuller test，ADF)对二次分解后的各模态进行平稳性检验，验证分解的有效性。

5）对二次分解后的各模态进行 Min-Max 归一化处理，计算式如下：

$$y' = \frac{y - y_{\min}}{y_{\max} - y_{\min}} \tag{13}$$

式中：y'为归一化后的数据；y 为原始数据；$y_{\max}$为原始数据集中的最大数据；$y_{\min}$ 为原始数据集中的最小数据。

6）利用 1D-CNN 对各模态归一化后的数据进行充分挖掘，获取序列的深层特征，形成特征向量并输入到 LSTM 中。

7）利用 POA 对影响 LSTM 预测精度的 3 个最关键的超参数——最大迭代次数、初始学习率以及隐藏层神经元数进行寻优，确定上述 3 个关键参数的最优参数组。

8）利用最优参数组下的 LSTM 对各模态进行预测，将各模态的预测值叠加重构，得到最终的预测结果。

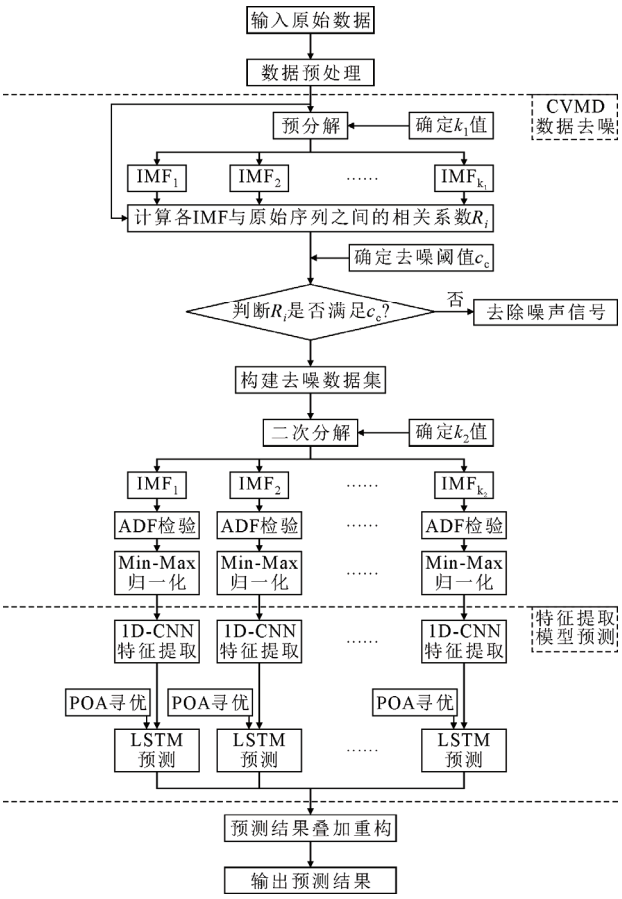


图 4　CVMD-CNN-LSTM 结构图

Fig.4　Structure diagram of CVMD-CNN-LSTM

3　实例分析

3.1　数据选取

本文采用文献[23]提供的变压器油中溶解气体体积分数监测数据，采集时间为 2013-11-07 至 2014-10-25，采集周期为 8 h，共计 1050 组数据，气体体积分数单位为 10^{-6}。其中，将 2013-11-07 至 2014-08-16 的 840 组数据作为训练样本，将 2014-08-17 至 2014-10-25 的 210 组数据作为测试样本。限于文章篇幅，本文以 H_2 为例展开分析。

3.2　模型评价指标

为了对本文所提算法及其对比算法进行准确评估，本算例采用平均绝对误差(mean absolute error，MAE)、均方根误差(root mean square error，RMSE)、平均绝对百分比误差(mean absolute percentage error，MAPE)和拟合优度 R^2 对模型预测效果进行评价。MAE 表征了在不考虑方向的情况下预测值平均误差幅度，RMSE 表征了模型的拟合能力，MAPE 表征了模型预测的准确性，R^2 表征了预测值和真实值之间的相关程度。其中，MAE、RMSE 和 MAPE 为损失型指标，R^2 为增益型指标，即 MAE、

RMSE、MAPE 越小，R^2 越大，模型拟合效果越好，预测结果越准确。各指标计算式如下：

$$y_{\mathrm{MAE}}=\frac{1}{n}\sum_{i=1}^{n}\left|\hat{y}_i-y_i\right| \tag{14}$$

$$y_{\mathrm{RMSE}}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_i-y_i\right)^2} \tag{15}$$

$$y_{\mathrm{MAPE}}=\frac{1}{n}\sum_{i=1}^{n}\frac{\left|\hat{y}_i-y_i\right|}{y_i}\times 100\% \tag{16}$$

$$R^2=1-\frac{\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2} \tag{17}$$

式中：n 为数据样本数量；$\hat{y}_i$ 和 y_i 分别为第 i 时刻油中溶解气体体积分数的预测值与实际值；$\bar{y}$ 为实际数据的平均值。

3.3 数据预处理与 CVMD 分解

因为 H_2 数据序列具备等间隔性，无需对其进行等时间间隔处理，因此本文主要采用插值法对 H_2 数据序列中的异常数据进行修正以及对缺失数据进行填充，即以该数据值前一时刻和后一时刻对应数据均值作为该数据值。

对预处理后的数据进行预分解，利用中心频率观察法确定 k_1 的值，预设 k_1 从小到大取值，当各 IMF 分量的中心频率保持相对稳定，且无模态混叠现象出现时，即认为 k_1 取得最佳值。在 k_1 从小到大取值过程中，各 IMF 分量中心频率值如表 3 所示。

从表 3 可见，在 k_1>6 后，各 IMF 分量的中心频率保持相对稳定，但当 k_1>7 时，高频分量的中心频率非常接近，存在一定的模态混叠现象。此外，对不同 k_1 值下各 IMF 分量进行 ADF 检验，结果显示 k_1=9 时 IMF_1 与 IMF_2 存在单位根，为不平稳序列，其余分量均不存在单位根，为平稳序列。综上所述，当 k_1=8 时，高频分量存在一定的模态混叠现象；当 k_1=9 时，不仅高频分量存在模态混叠现象，低频分量中还存在不平稳序列，故认为 k_1=7 为最佳值。VMD 其他参数的设置参考文献[8]的方法，具体取值如表 4 所示。H_2 序列预分解结果如图 5 所示，从上至下依次为原始序列曲线、IMF_1—IMF_7 曲线。

从 H_2 序列预分解结果看，H_2 序列中明显存在噪声信号，且噪声信号主要存在于高频 IMF 中。为减小噪声信号的影响，后文对 H_2 序列进行去噪处理。首先根据式(4)计算各 IMF 与原始 H_2 序列之间的相关系数，计算结果如表 5 所示。

表 3 不同 k_1 值下各 IMF 分量中心频率

Table 3 Frequency of each IMF component center under different k_1 values

k_1	IMF_1	IMF_2	IMF_3	IMF_4	IMF_5	IMF_6	IMF_7	IMF_8	IMF_9
2	0	0.068							
3	0	0.068	0.329						
4	0	0.017	0.075	0.329					
5	0	0.017	0.075	0.327	0.424				
6	0	0.016	0.072	0.148	0.328	0.424			
7	0	0.016	0.067	0.115	0.217	0.33	0.425		
8	0	0.016	0.067	0.115	0.217	0.329	0.412	0.454	
9	0	0.022	0.055	0.09	0.152	0.228	0.33	0.418	0.473

表 4 VMD 参数设置

Table 4 Parameter settings of VMD

惩罚因子	时间步长	直流效应	初始化频率	收敛容差
2500	0	1	1	10^{-6}

本文参考文献[8]所示方法确定去噪阈值 c_c，结合表 1、表 4 与图 5，确定噪声信号存在于与原始序列不相关的 IMF 分量中，即去噪范围为(0,0.1)。在去噪范围内，以 0.01 为步长，分别令 c_c=0.01n，1≤n≤9 且 n 为正整数，然后在 c_c 的每一个取值下分别构建去噪序列并进行二次分解，最后利用 ADF 检验二次分解后各 IMF 的平稳性，进而判断 c_c 取值是否可行。实验结果表明，当 c_c>0.5 时，无论二次分解 k_2 取何值，利用 ADF 检验二次分解后各 IMF 的平稳性，发现总有 IMF 存在单位根，即为不平稳序列，故 c_c>0.5 的值不可取；当 c_c≤0.05 时，二次分解后的各 IMF 均满足平稳性要求，故 c_c≤0.05 的值均可取。因此，为了在满足分解平稳性要求的同时尽可能去除噪声信号，本文将去噪阈值设定为 0.05。之后利用式(5)构建去噪序列 f_{denoise} 并对去噪序列进行二次分解，同样利用中心频率观察法确定 k_2=4。

二次分解结果如图 6 所示，从上至下依次为去噪序列曲线、IMF_1—IMF_4 曲线，各 IMF 与原始 H_2 序列之间的相关系数见表 5。由图 6 可见，尽管 H_2 序列经二次分解后依然残留了少量噪声，但噪声的幅度远小于原始序列。对各模态分别进行 ADF 检验，结果显示各模态 Augmented Dickey-Fuller Test Statistic 的值显著小于 3 个置信度(1%，5%，10%)的临界值，同时 p-value 接近于 0，证明各模态均平稳，进一步证明了 CVMD 分解的有效性。

3.4 数据特征提取与预测

将 CVMD 分解后的各模态序列分别输入到 CNN-LSTM 组合预测模型中，CNN-LSTM 组合预

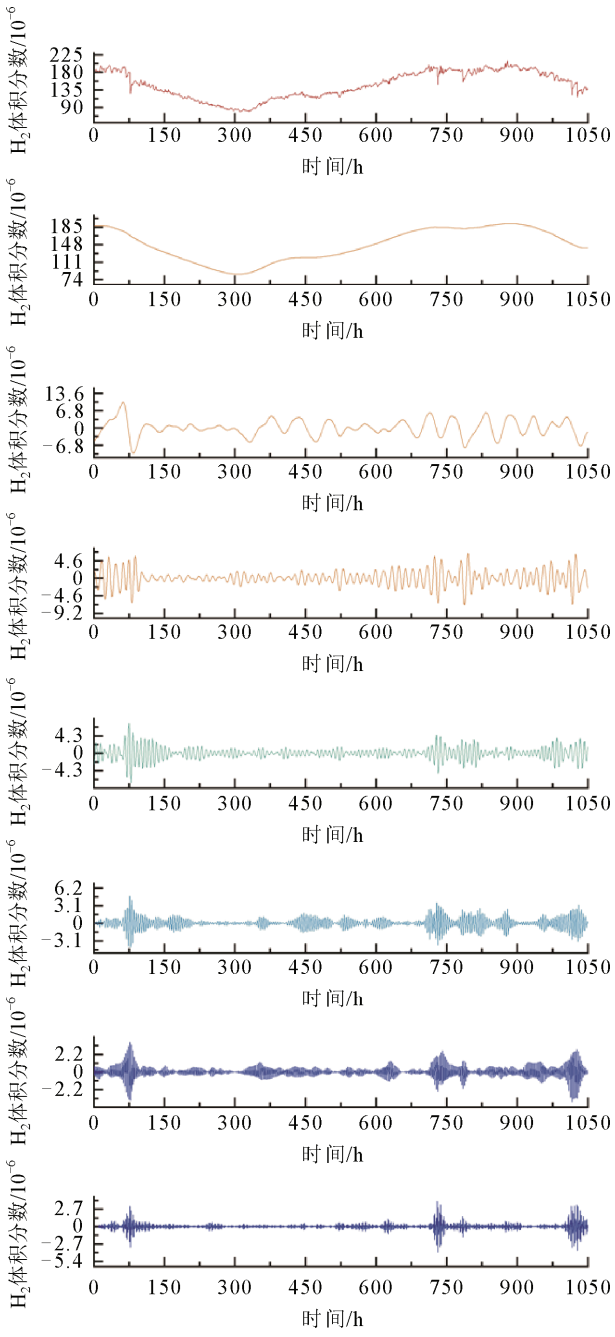


图 5　H_2 序列预分解图

Fig.5　Pre-decomposition diagram of H_2 sequence

表 5　各 IMF 与原始 H_2 序列之间的相关系数

Table 5　The correlation coefficient between each IMF and the original H_2 sequence

各 IMF	IMF_1	IMF_2	IMF_3	IMF_4	IMF_5	IMF_6	IMF_7
相关系数	0.99	0.24	0.1	0.07	0.05	0.04	0.03

测模型由输入层、CNN 网络层、LSTM 网络层和输出层组成。输入层负责将各模态分量输入到 CNN 网络层。因为经过 CVMD 分解后的各分量均为平

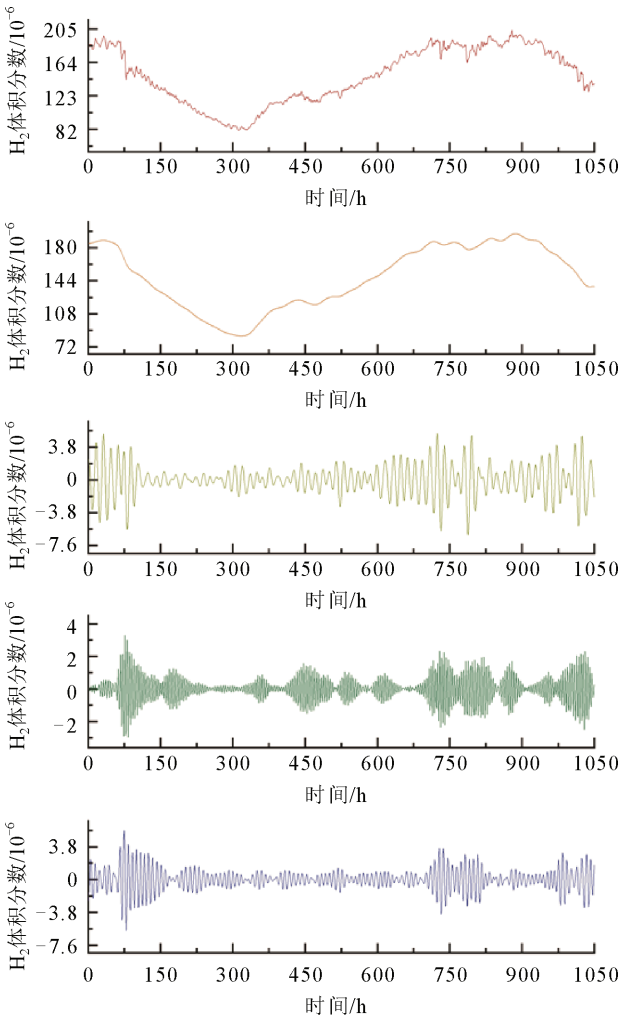


图 6　H_2 序列二次分解图

Fig.6　Secondary decomposition diagram of H_2 sequence

稳分量，为避免网络层数过深造成模型梯度消失或梯度爆炸，CNN 网络层设置 2 层 1 维卷积层和 1 层最大池化层，利用 1 维卷积层充分挖掘数据之间的深层特征，利用池化层对 2 维卷积层所提取的特征进行下采样，去除冗余信息，提高模型的运算效率。为缩短学习周期同时避免模型训练过程中发生过拟合，选择 ReLU 作为激活函数。为在保证模型预测性能的同时缩短训练时间，LSTM 网络层设置 2 层 LSTM 层、1 层 Dropout 层和 1 层 Dense 层。LSTM 层用于对 CNN 网络层提取到的特征进行充分学习，遗忘无效信息，保留有用信息，并对未来的数据进行预测；Dropout 层用于解决模型过拟合问题；Dense 层用于压缩数据，精简网络参数。输出层由一层 Dense 层组成，用于将 LSTM 网络层输出的预测值进行全连接后输出。最后，将各模态分量预测结果叠加重构，得到最终的气体预测值。模型结构如图 7 所示。

3.5 实验设计与结果分析

为全方位、多角度验证本文模型的优越性，本文设计 4 组对比实验进行验证。为保证各模型均为最优模型，各模型关键参数均采用 POA 进行优化。

3.5.1 消融实验

为验证 CVMD-CNN-LSTM 各部分的作用，将 LSTM 作为基准模型，分别加入 CVMD、CNN，构建消融模型 CVMD-LSTM、CNN-LSTM。通过消融实验，得到本文模型与消融模型的预测结果对比如表 6 所示。

由表 6 可见，同 LSTM 相比，CNN-LSTM 的 MAE、RMSE、MAPE 分别降低 38.4%、29.6%、32.8%，R^2 提高了 13.8%，验证了采用 CNN 挖掘数据的深层特征可有效提高模型的预测性能。同 LSTM 相比，CVMD-LSTM 的 MAE、RMSE、MAPE 分别降低 40.1%、44.4%、42%，R^2 提高了 16.3%，验证了采用 CVMD 对数据进行去噪以及平稳性分解对模型预测性能的提升效果。同 LSTM 相比，本文所提模型的 MAE、RMSE、MAPE 分别降低 68.6%、71.3%、70.5%，R^2 提高了 20.7%，同时相较于 CVMD-LSTM 和 CNN-LSTM，本文所提模型的 MAE、RMSE、MAPE 更小，R^2 更大，这 2 项结果表明 CVMD 与 CNN 的组合更大限度地发挥了各自的优势，进一步提升了模型的预测性能。表 6 中各模型的测试样本预测曲线如图 8 所示，由图 8 可见本文所提模型的预测值比另外 3 种模型更加贴近实际监测值，证明了本文模型的优越性。

3.5.2 单步预测实验

为进一步验证 CVMD 数据去噪对模型整体预测性能的提升效果，采用 VMD 和 CNN、LSTM 相结合，构建 VMD-CNN-LSTM 和 VMD-LSTM，并与 CVMD-CNN-LSTM 和 CVMD-LSTM 进行比较，构造 2 组单步预测对比实验。模型评价指标对比结果如表 7 所示，预测曲线对比结果如图 9 所示。

从表 7 评价指标的定量分析可见，相较于 VMD-CNN-LSTM，本文所提模型的 MAE、RMSE、MAPE 分别降低 42.2%、41.1%、42.3%，R^2 提高了 1.8%；相较于 VMD-LSTM，CVMD-LSTM 的 MAE、RMSE、MAPE 分别降低 13.8%、19.7%、13.8%，R^2 提高了 3%。从图 9 可见，相较于没有进行数据去噪的模型 VMD-CNN-LSTM 和 VMD-LSTM，采用 CVMD 进行数据去噪的模型 CVMD-CNN-LSTM 和 CVMD-LSTM 的预测曲线与实际监测曲线拟合

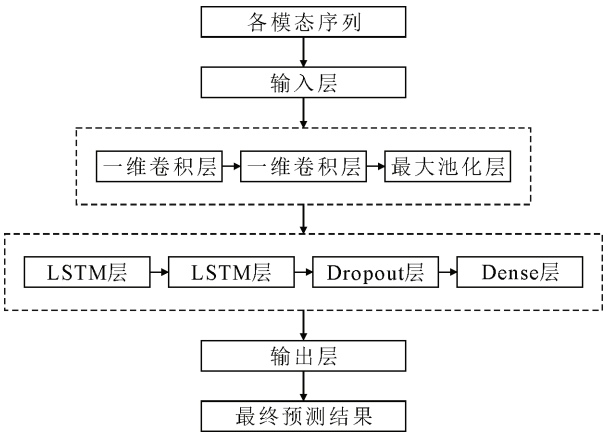


图 7 CNN-LSTM 结构图

Fig.7 Structure diagram of CNN-LSTM

表 6 消融实验结果

Table 6 Ablation experimental results

模型	MAE	RMSE	MAPE/%	R^2
CVMD-CNN-LSTM	2.75	3.28	1.53	0.99
CVMD-LSTM	5.18	6.35	3.01	0.938
CNN-LSTM	5.4	8.04	3.49	0.911
LSTM	8.77	11.42	5.19	0.785

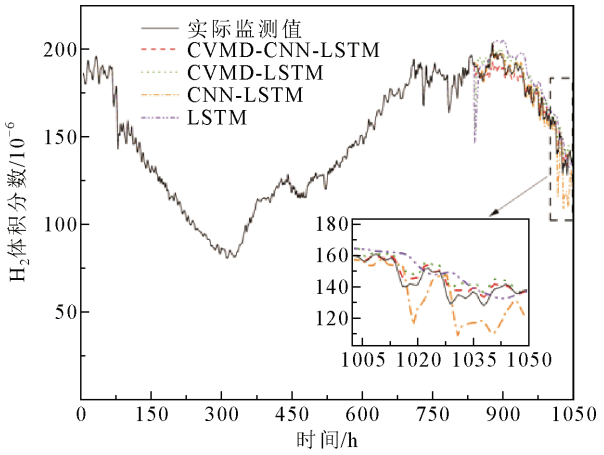


图 8 消融实验各模型预测曲线对比图

Fig.8 Comparison of prediction curves of each model in ablation experiment

表 7 单步预测实验结果

Table 7 Single-step prediction of experimental results

实验序号	模型	MAE	RMSE	MAPE/%	R^2
1	CVMD-CNN-LSTM	2.75	3.28	1.53	0.990
	VMD-CNN-LSTM	4.76	5.57	2.65	0.972
2	CVMD-LSTM	5.18	6.35	3.01	0.938
	VMD-LSTM	6.01	7.91	3.49	0.910

性更好，预测数据的峰值更接近实际监测值，出现拐点的时间也与实际监测值基本相符，这 2 项实验结果也充分表明了 CVMD 数据去噪对模型整体预测性能的提升效果。另外，相较于 VMD-LSTM，VMD-CNN-LSTM 的 MAE、RMSE、MAPE 分别降低 20.8%、29.6%、24.1%，R^2 提高了 6.4%，进一步验证了利用 CNN 挖掘数据的深层特征对模型预测性能的提升效果。

3.5.3　多步预测实验

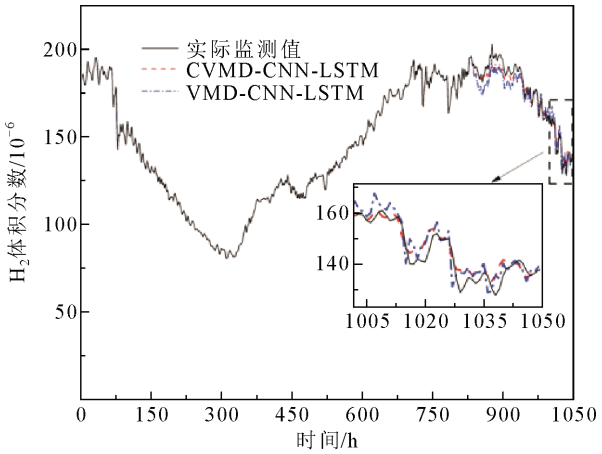
为纵向验证本文模型的预测稳定性与可靠性，对 H_2 体积分数进行了多步预测。本文采用滚动预测方式构建多步预测，即每完成一次单步预测，都将各模态的单步预测结果作为新的特征输入到模型中来完成下一次单步预测，最终得到各模态的多步预测结果。将各模态的多步预测值叠加重构，即可得到 H_2 体积分数的多步预测值。对比模型选择 3.5.1 节与 3.5.2 节中预测精度最高的 2 个模型 CVMD-CNN-LSTM 和 CVMD-LSTM，限于文章篇幅，以 6 步预测为例进行说明，即对未来 2 天的 H_2 体积分数进行延伸预测，表 8 给出了上述 3 种模型多步预测值与实际值的结果对比，表 9 给出了 3 种模型多步预测评价指标的结果对比。

由表 8 可知，由于误差的累积效应，随着预测步数增加，3 个模型的预测性能均有不同程度下降，但在步数相同的情况下，本文所提模型的预测值与实际值更加接近。由表 9 可知，与 VMD-CNN-LSTM 和 CVMD-LSTM 相比，本文所提模型的 MAE 分别降低 45.5%和 50.1%，RMSE 分别降低 41%和 46.6%，MAPE 分别降低 41.9%和 52.3%，R^2 分别提高 7.7%和 11.5%。结合表 8、表 9 可见，CVMD 数据去噪和 CNN 深层特征提取对于模型单步及多步预测性能均有提升作用，充分展现了本文模型的优越性。

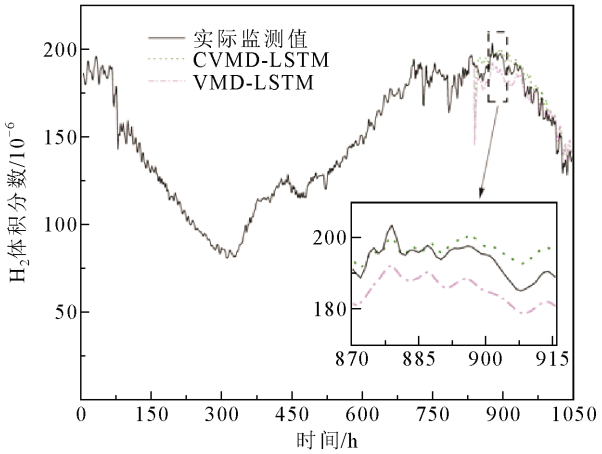
3.5.4　其他特征气体预测实验

为横向验证本文模型的预测稳定性与可靠性，分别对其他 6 种特征气体 2013-11-07 日至 2014-10-25 日的 1050 组数据进行预测，训练样本与测试样本的划分同 H_2，对比模型选择同 3.5.3 节，预测结果如表 10 所示。

从表 10 可以看出，针对其他 6 种特征气体，本文所提模型的 MAPE 稳定在 0.88%~1.72%之间，R^2 稳定在 0.948~0.997 之间，MAE 和 RMSE 也明显优于 VMD-CNN-LSTM 和 CVMD-LSTM，保持着最佳的预测性能，由此可见，本文所提模型在面对



(a) 实验1各模型预测曲线对比图



(b) 实验2各模型预测曲线对比图

图 9　单步预测实验预测曲线对比图

Fig.9　Comparative diagram of prediction curve of single-step prediction experiment

表 8　各模型多步预测值与实际值的结果对比

Table 8　Comparison between the results of multi-step prediction and actual value of each model

实际值	模型预测值		
	CVMD-CNN-LSTM	VMD-CNN-LSTM	CVMD-LSTM
138.7	138.5	142.2	136.2
139.9	138.6	141.9	136.9
141.2	139.3	148.4	136.9
147.6	141.7	150.7	137.8
146.8	143.1	154.7	139.2
148.6	143.8	157.6	139.5

表 9　各模型多步预测评价指标的结果对比

Table 9　Comparison of the results of multi-step prediction and evaluation indexes of each model

模型	MAE	RMSE	MAPE/%	R^2
CVMD-CNN-LSTM	2.97	3.58	2.09	0.923
VMD-CNN-LSTM	5.45	6.07	3.6	0.852
CVMD-LSTM	6.05	6.71	4.38	0.817

表 10 其他特征气体预测结果

Table 10 Prediction results of other characteristic gases

特征气体	模型	MAE	RMSE	MAPE/%	R^2
CH_4	CVMD-CNN-LSTM	0.53	0.65	1.72	0.974
	VMD-CNN-LSTM	0.63	0.77	2.04	0.968
	CVMD-LSTM	0.79	0.97	2.45	0.763
C_2H_2	CVMD-CNN-LSTM	0.008	0.01	1.66	0.973
	VMD-CNN-LSTM	0.014	0.02	2.85	0.970
	CVMD-LSTM	0.02	0.02	3.85	0.950
C_2H_4	CVMD-CNN-LSTM	0.25	0.31	0.88	0.991
	VMD-CNN-LSTM	0.57	0.69	1.99	0.966
	CVMD-LSTM	0.64	0.85	2.18	0.821
C_2H_6	CVMD-CNN-LSTM	0.31	0.36	1.15	0.997
	VMD-CNN-LSTM	0.57	0.73	2.07	0.963
	CVMD-LSTM	0.73	0.85	2.69	0.973
CO	CVMD-CNN-LSTM	4.11	4.25	1.11	0.990
	VMD-CNN-LSTM	9.16	9.39	2.48	0.978
	CVMD-LSTM	10.18	10.72	2.77	0.868
CO_2	CVMD-CNN-LSTM	25.92	27.1	1.06	0.948
	VMD-CNN-LSTM	41.93	44.77	1.72	0.800
	CVMD-LSTM	36.63	49.82	1.51	0.566

不同数据样本时具有更强的自适应能力，进一步证明了本文模型的预测稳定性和可靠性。

4 结论

为了更好地追踪变压器油中溶解气体体积分数的变化规律，本文将数据分解算法与神经网络模型相结合，构建了一种基于 CVMD-CNN-LSTM 的组合预测模型。通过算例仿真得到如下结论：

1）利用 CVMD 去除了原始序列中混入的噪声信号，改善了数据质量，相较于传统 VMD，其 MAPE 降低了 42.3%，有效解决了传统 VMD 算法噪声信号抑制不充分进而导致后续预测模型拟合能力不够的问题，提高了模型整体的预测性能。

2）采用 1D-CNN 充分挖掘数据间的深层特征，减少了人为提取特征所带来的误差，相较于单一 LSTM，其 MAPE 降低了 32.8%，克服了单一 LSTM 模型不能有效提取数据间深层特征的缺陷，提高了模型的预测稳定性与可靠性。

3）本文所提模型综合了 CVMD、1D-CNN 和 LSTM 这 3 者的优势，相较于未采取任何优化的基础 LSTM 模型，MAE、RMSE、MAPE 分别降低了 68.6%、71.3%、70.5%，R^2 提高了 20.7%，具有更高的预测精度与效率。

4）本文所提模型可根据不同的数据样本，自适应调整模型的超参数，实现变压器油中溶解气体体积分数的精准预测，为变压器在线监测和故障预警提供了重要的技术支撑，具有较高的应用价值。

附录见本刊网络版(http://hve.epri.sgcc.com.cn)。

参考文献 References

[1] 刘展程，王 爽，唐 波. 基于 SSA-BiGRU-Attention 模型的变压器油中溶解气体含量预测[J]. 高电压技术，2022，48(8)：2972-2981.
LIU Zhancheng, WANG Shuang, TANG Bo. Prediction of dissolved gas content in transformer oil based on SSA-BiGRU-attention model[J]. High Voltage Engineering, 2022, 48(8): 2972-2981.

[2] DUVAL M, DEPABLA A. Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases[J]. IEEE Electrical Insulation Magazine, 2001, 17(2): 31-41.

[3] 谢 乐，仇 炜，李振伟，等. 基于变分模态分解和门控循环单元神经网络的变压器油中溶解气体预测模型[J]. 高电压技术，2022，48(2)：653-660.
XIE Le, QIU Wei, LI Zhenwei, et al. Prediction model of dissolved gas in transformer oil based on variational modal decompo-sition and recurrent neural network with gated recurrent unit[J]. High Voltage Engineering, 2022, 48(2): 653-660.

[4] DENG L F, ZHAO R Z. Fault feature extraction of a rotor system based on local mean decomposition and Teager energy kurtosis[J]. Journal of Mechanical Science and Technology, 2014, 28(4): 1161-1169.

[5] 屈 雪. 改进的 VMD 算法在天然气管道泄漏信号去噪中的应用[D]. 大庆：东北石油大学，2022.
QU Xue. Application of improved VMD algorithm in gas pipeline leakage signal denoising[D]. Daqing, China: Northeast Petroleum University, 2022.

[6] 米小婷. 基于 MIC 和 EMD 的金融时间序列去噪及其应用研究[D]. 长沙：湖南大学，2021.
MI Xiaoting. Research on denoising and application of financial time series based on MIC and EMD[D]. Changsha, China: Hunan University, 2021.

[7] 肖怀硕，李清泉，施亚林，等. 灰色理论-变分模态分解和 NSGA-Ⅱ优化的支持向量机在变压器油中气体预测中的应用[J]. 中国电机工程学报，2017，37(12)：3643-3653.
XIAO Huaishuo, LI Qingquan, SHI Yalin, et al. Prediction of dissolved gases in oil for transformer based on grey theory-variational mode decomposition and support vector machine improved by NSGA-II[J]. Proceedings of the CSEE, 2017, 37(12): 3643-3653.

[8] SHEN Y F, ZHENG W, YIN W J, et al. Feature extraction algorithm using a correlation coefficient combined with the VMD and its application to the GPS and GRACE[J]. IEEE Access, 2021, 9, 17507-17519.

[9] 黄新波，蒋卫涛，朱永灿，等. 基于时间序列和支持向量机的变压器故障预测[J]. 高电压技术，2020，46(7)：2530-2538.
HUANG Xinbo, JIANG Weitao, ZHU Yongcan, et al. Transformer fault prediction based on time series and support vector machine[J]. High Voltage Engineering, 2020, 46(7): 2530-2538.

[10] 张 鹏，齐 波，张若愚，等. 基于经验小波变换和梯度提升径向基的变压器油中溶解气体预测方法[J]. 电网技术，2021，45(9)：3745-3754.
ZHANG Peng, QI Bo, ZHANG Ruoyu, et al. Dissolved gas prediction in transformer oil based on empirical wavelet transform and gradient boosting radial basis[J]. Power System Technology, 2021, 45(9): 3745-3754.

[11] DAI J J, SONG H, SHENG H H, et al. LSTM networks for the trend

prediction of gases dissolved in power transformer insulation oil[C]// 2018 12th International Conference on the Properties and Applications of Dielectric Materials (ICPADM). Xi'an, China: IEEE, 2018: 666-669.

[12] 邹　智，吴铁洲，张晓星，等. 基于贝叶斯优化 CNN-BiGRU 混合神经网络的短期负荷预测[J]. 高电压技术，2022，48(10)：3935-3945.
ZOU Zhi, WU Tiezhou, ZHANG Xiaoxing, et al. Short-term load forecast based on bayesian optimized CNN-BiGRU hybrid neural networks[J]. High Voltage Engineering, 2022, 48(10): 3935-3945.

[13] 塔　拉. 地震数据去噪中的几种变分模态分解方法[D]. 哈尔滨：哈尔滨工业大学，2020.
TARA P B. Several mode decomposition methods in seismic data denosing[D]. Harbin, China: Harbin Institute of Technology, 2020.

[14] 王　俊，李　霞，周昔东，等. 基于 VMD 和 LSTM 的超短期风速预测[J]. 电力系统保护与控制，2020，48(11)：45-52.
WANG Jun, LI Xia, ZHOU Xidong, et al. Ultra-short-term wind speed prediction based on VMD-LSTM[J]. Power System Protection and Control, 2020, 48(11): 45-52.

[15] LI Y X, LI Y A, CHEN X, et al. Research on ship-radiated noise denoising using secondary variational mode decomposition and correlation coefficient[J]. Sensors, 2017, 18(1): 48.

[16] 杨　威，蒲彩霞，杨　坤，等. 基于 CNN-GRU 组合神经网络的变压器短期故障预测方法[J]. 电力系统保护与控制，2022，50(6)：107-116.
YANG Wei, PU Caixia, YANG Kun, et al. Short-term fault prediction method for a transformer based on a CNN-GRU combined neural network[J]. Power System Protection and Control, 2022, 50(6): 107-116.

[17] KIRANYAZ S, AVCI O, ABDELJABER O, et al. 1D convolutional neural networks and applications: a survey[J]. Mechanical Systems and Signal Processing, 2021, 151: 107398.

[18] MAHRUKH A W, LIAN G X, BIN S S. Prediction of power transformer oil chromatography based on LSTM and RF model[C]// 2020 IEEE International Conference on High Voltage Engineering and Application (ICHVE). Beijing, China: IEEE, 2020: 1-4.

[19] 刘可真，苟家萁，骆　钊，等. 基于粒子群优化–长短期记忆网络模型的变压器油中溶解气体浓度预测方法[J]. 电网技术，2020，44(7)：2778-2784.
LIU Kezhen, GOU Jiaqi, LUO Zhao, et al. Prediction of dissolved gas concentration in transformer oil based on PSO-LSTM model[J]. Power System Technology, 2020, 44(7): 2778-2784.

[20] 李　春，高　飞，王会青. 改进果蝇算法优化 CIAO-LSTM 网络的时序预测模型[J]. 计算机工程与应用，2020，56(11)：129-134.
LI Chun, GAO Fei, WANG Huiqing. Improved fruit fly optimization algorithm for optimizing time series prediction model of CIAO-LSTM network[J]. Computer Engineering and Applications, 2020, 56(11): 129-134.

[21] TROJOVSKÝ P, DEHGHANI M. Pelican optimization algorithm: a novel nature-inspired algorithm for engineering applications[J]. Sensors, 2022, 22(3): 855.

[22] 刘　航，王有元，梁玄鸿，等. 基于多因素的变压器油中溶解气体体积分数预测方法[J]. 高电压技术，2018，44(4)：1114-1121.
LIU Hang, WANG Youyuan, LIANG Xuanhong, et al. Prediction method of the dissolved gas volume fraction in transformer oil based on multi factors[J]. High Voltage Engineering, 2018, 44(4): 1114-1121.

[23] LEWIS Z, WILD J, ALLCOCK M, et al. Assessing the impact of weak and moderate geomagnetic storms on UK power station transformers[J]. Space Weather, 2022, 20(4): 1-12.



FAN Zhiyuan

范志远
1999—，男，硕士生
主要从事变压器及其智能化技术方面的研究工作
E-mail: 1162566160@qq.com



DU Jiang
Ph.D.
Associate professor
Corresponding author

杜　江(通信作者)
1972—，男，博士，副教授，硕导
主要从事变压器及其智能化技术方面的研究工作
E-mail: dj@hebut.edu.cn

收稿日期　2022-10-30　修回日期　2023-01-06　　编辑　卫李静

附录 A

表 A1　变压器油中溶解气体体积分数数据样本示例

Table A1　Sample data of dissolved gas volume fraction data in transformer oil

单位：10^{-6}

采样时间	H_2	CH_4	C_2H_2	C_2H_4	C_2H_6	CO	CO_2
2013–11–07 T03:00	179.4	37.2	0.6	36.3	24.0	294.3	1677
2013–11–07 T11:00	182.7	36.3	0.4	37.0	27.6	294.0	1670
2013–11–07 T19:00	184.6	36.8	0.5	36.5	25.0	293.1	1670
2013–11–08 T03:00	185.7	36.7	0.5	37.1	26.1	293.7	1667
2013–11–08 T11:00	187.0	36.6	0.3	36.2	26.5	293.5	1668
2013–11–08 T19:00	186.8	36.3	0.4	35.8	25.1	292.1	1660
2013–11–09 T03:00	182.6	36.2	0.5	32.5	18.8	293.8	1669
2013–11–09 T11:00	182.7	37.0	0.3	36.5	25.4	290.1	1664
2013–11–09 T19:00	181.8	36.3	0.3	34.4	19.7	293.3	1668