

基于物理模型驱动的机器学习方法预测超临界二氧化碳管道最大泄漏速率

王一新¹, 陆诗建², 李卫东¹, 滕霖^{1,3*}

1 福州大学石油化工学院, 福州 350108

2 中国矿业大学碳中和研究院, 徐州 221008

3 重庆大学产业技术研究院, 重庆 401329

* 通信作者, tenglin@fzu.edu.cn

收稿日期: 2021-01-15

重庆市自然科学基金(CYY202010102001)和福州大学科研启动基金(GXRC-20041)联合资助

摘要 碳捕集与封存(CCS)项目中涉及的大规模CO₂适合采用超临界管道输送。然而超临界CO₂管道泄漏过程伴随着复杂相变,因此对其最大泄漏速率进行准确预测是目前的研究难点。鉴于传统物理模型方法存在建模复杂、假设过多、计算耗时等缺点,研究提出通过机器学习方法预测超临界CO₂管道最大泄漏速率,分别采用粒子群算法优化的支持向量机(PSO-SVM)和简化处理的卷积神经网络(CNN)对等熵阻塞泄漏模型所生成的泄漏特征数据进行学习,并测试了机器学习模型的预测准确率和泛化能力。研究表明:①物理模型、PSO-SVM、CNN的预测结果与实验数据的平均误差为28.82%;②两种机器学习模型预测精度相差不大,CNN的训练时间远短于PSO-SVM,但PSO-SVM的泛化能力强于CNN,因此,SVM适用于小样本数据精确预测,而CNN更适用于对大数据的学习和预测。本研究成果为超临界CO₂管道最大泄漏速率预测提供了一种高效的新方法。

关键词 机器学习; 超临界二氧化碳; 管道; 泄漏; 卷积神经网络; 支持向量机

A physical model driven machine learning for predicting maximum leakage rate in supercritical CO₂ release

WANG Yixin¹, LU Shijian², LI Weidong¹, TENG Lin^{1,3}

1 College of Chemical Engineering, Fuzhou University, Fuzhou 350116, China

2 Carbon Neutrality Institute, China University of Mining and Technology, Xuzhou 221008, China

3 Chongqing University Industrial Technology Research Institute, Chongqing 401329, China

Abstract Supercritical CO₂ pipelines are suitable to transport the large-scale CO₂ involved in Carbon capture and storage projects. The leakage process of supercritical CO₂ pipelines is accompanied by complex phase changes. Therefore, it is difficult to predict the maximum leakage rate accurately at present. In view of the shortcomings of traditional physical model methods such as complex modeling, too many assumptions and time-consuming calculations, a way of predicting the maximum leakage rate of supercritical CO₂ pipelines by machine learning method was proposed. It used to simply convolutional neural networks (CNN) and support vector machine improved by particle swarm optimization (PSO-SVM) respectively to study the leakage feature data generated by the

引用格式: 王一新, 陆诗建, 李卫东, 滕霖. 基于物理模型驱动的机器学习方法预测超临界二氧化碳管道最大泄漏速率. 石油科学通报, 2023, 01: 102-111

WANG Yixin, LU Shijian, LI Weidong, TENG Lin. A physical model driven machine learning for predicting maximum leakage rate in supercritical CO₂ release. Petroleum Science Bulletin, 2023, 01: 102-111. doi: 10.3969/j.issn.2096-1693.2023.01.007

isentropic choked flow leakage model. The prediction accuracy and generalization ability of the trained machine learning model were tested. The results show that: First, the average error between experimental data and prediction results of physical model, PSO-SVM, CNN is 28.82%. Second, the prediction accuracy of the two machine learning models shows little difference, the training time of CNN is much shorter than that of PSO-SVM, but the generalization ability of PSOSVM is stronger than that of CNN. Therefore, SVM is suitable for accurate prediction of small sample data, while CNN is more suitable for learning and prediction of large sample data. This study provides a new efficient method for predicting the maximum leakage rate of supercritical CO₂ pipelines.

Keywords machine learning; supercritical CO₂; pipeline; leakage; convolutional neural networks; support vector machine

doi: 10.3969/j.issn.2096-1693.2023.01.007

0 引言

近年来,全球极端气候不断显现,以二氧化碳为代表的温室气体减排问题更加引起各国重视^[1]。全球气候变化专门委员会(IPCC)认为碳捕获与封存技术(Carbon Capture and Storage,简称CCS)是降低温室气体排放最有效的措施之一^[2]。CO₂输送作为CCS产业链的关键环节,是连接碳捕集地与封存地的桥梁。研究表明,采用超临界态管道输送CO₂是最为经济的方式^[3]。然而,对于这种非常规气体管道,其安全性有待考量。CO₂的大规模泄漏会造成人员伤亡^[4]。若要评估CO₂泄漏扩散的危害,需知晓其泄漏量。不同于传统油气管道的是,超临界CO₂泄漏过程伴随剧烈相变。研究发现,CO₂在泄漏过程中经历了超临界→气液共存→气态,甚至会生成固态CO₂(干冰)。因此,如何准确预测超临界CO₂管道的泄漏量是进行CO₂泄漏扩散安全评估的关键问题。

传统的研究方法基于流体力学和热力学理论,建立泄漏模型,并采用数值方法求解相应偏微分方程,最终获得超临界CO₂管道的泄漏速率^[5-6]。具有代表性的是,英国伦敦大学学院(UCL)Mahgerefteh团队分别采用均相流模型(HEM)和均相延迟模型(HRM)建立了CO₂管道泄^[7-8]漏模型预测泄漏特性,经与实验结果对比后发现,均相流模型更适用于全尺寸破裂情况,而均相延迟模型则能充分考虑到CO₂相变过程的热力学非平衡特性。除此之外,为了解释超临界CO₂管道泄漏过程的动力学非平衡特性,研究者们采用了漂移流模型(DFM)和双流模型(TFM)^[9-10]。由此可见,针对超临界CO₂管道泄漏这一复杂物理现象,传统的研究方法是对物理模型进行简化,并采用相应的多相流模型进行模拟,该方法具有物理机理明晰的优点,但也存在建模复杂、假设过多、计算耗时等缺点。同时由于存在各类假设,还使得预测结果存在偏差,已难以适应对泄漏扩散规律的快速预测及对风险的实时预警的需要。

近年来兴起的人工智能技术为预测超临界CO₂管道泄漏量提供了新的手段。机器学习(Machine Learning)作为一种实现人工智能的方法,主要是借助机器的强大运算能力,模拟人类的学习行为,通过分析历史大数据,挖掘事物的潜在规律,实现对事物未来变化趋势的预测^[11-13]。相较于运用传统的理论方法,机器学习方法具有如下优点:(1)机器学习能够挖掘参数之间的非线性和隐藏关系,解决一些传统方法难以解决的非线性问题。(2)传统的物理模型比较复杂,故障诊断效果较差,假设如果太多,建立模型困难,机器学习方法能够抛开复杂的物理模型,只关注输入条件和输出结果,解决了传统方法建模复杂、假设过多的问题。(3)传统物理模型的通用性不强,不同模型误差诊断几率大,而机器学习模型通用性强,节省成本。(4)传统物理模型受专业知识限制,对客观事物的认识深度不够,机器学习方法能发现事物相关特征,专业门槛低。(5)机器的强大运算能力解决了传统方法计算耗时的问题,提高分析预测准确度。Shao等提出了通过电容传感器和机器学习模型测量二氧化碳气液两相流管道中气相体积分数的方法^[14]。他们利用实验数据训练了反向传播神经网络(BPNN)、径向基函数神经网络(RBFNN)、最小二乘法支持向量机(LS-SVM)3种模型,其中RBFNN模型在稳态流动条件下的预测结果优于其他两种模型,在瞬态流动条件下也能很好跟踪气相体积分数的快速变化。马贵阳等针对天然气水合物会给输气管道带来堵塞等严重影响的问题,利用支持向量机结合遗传算法建立了水合物生成预测模型^[15]。夏俏健等在对油气管道腐蚀速率的预测研究中训练了PCA-SVM、PCA-BP、PCA-GRNN、PCA-WNN四种腐蚀速率预测模型,经过对比,PCA-SVM模型的预测效果最好^[16]。李华杰根据管道泄漏负压波检测原理,采用T-S模糊神经网络预测模型实现管道泄漏定位^[17]。Xu等提出了基于粒子群优化(PSO)的自适应变分模态分解(VMD)算法检测液体管道泄漏^[18],该方法展示出强背景噪声去噪能力。张丽娟等针对目

前国内外研究出的多种泄漏检测与定位方法只局限于单条管线的问题,首先采用BP神经网络识别泄漏点在哪条支线上,最后采用线性神经网络准确定位泄漏点^[19]。Lang等针对流体管道中多处泄漏的位置难以定位的问题,提出了一种基于超声速和模拟退火算法改进的灰狼优化算法(SAGWO)的多泄漏位置定位方法。他们建立了SAGWO的目标函数来估计多处泄漏的位置,现场实验表明,该方法能够准确有效地定位流体管道中多个泄漏的位置^[20]。

综上所述,目前机器学习已被应用于油气管道泄漏定位领域,然而针对超临界CO₂管道泄漏这一复杂物理过程的预测方法还有待进一步研究。如图1所示,本研究首先基于等熵阻塞模型建立可预测超临界CO₂最大泄漏速率的物理模型,并基于该物理模型产生大量管道泄漏数据;然后采用粒子群算法优化的支持向量机(PSO-SVM)和简化处理的卷积神经网络(CNN)

分别对物理模型生成的大量数据进行学习;最后得到基于机器学习模型的超临界CO₂管道最大泄漏速率预测方法,并进一步对机器学习模型的适应性进行研究。

1 机器学习方法理论概述

1.1 PSO-SVM算法

支持向量机(SVM)是传统的机器学习算法的一种,一般被用来解决分类问题。SVM基本模型是在多维空间构造一个超平面使正负样本间隔最大。SVM算法不仅能解决分类问题,还可以有效地解决回归问题。由于SVM可以使用非线性的核函数,所以能够解决非线性问题。支持向量机超平面的最优化问题可转化为

$$\min \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n \xi_i \quad (1)$$

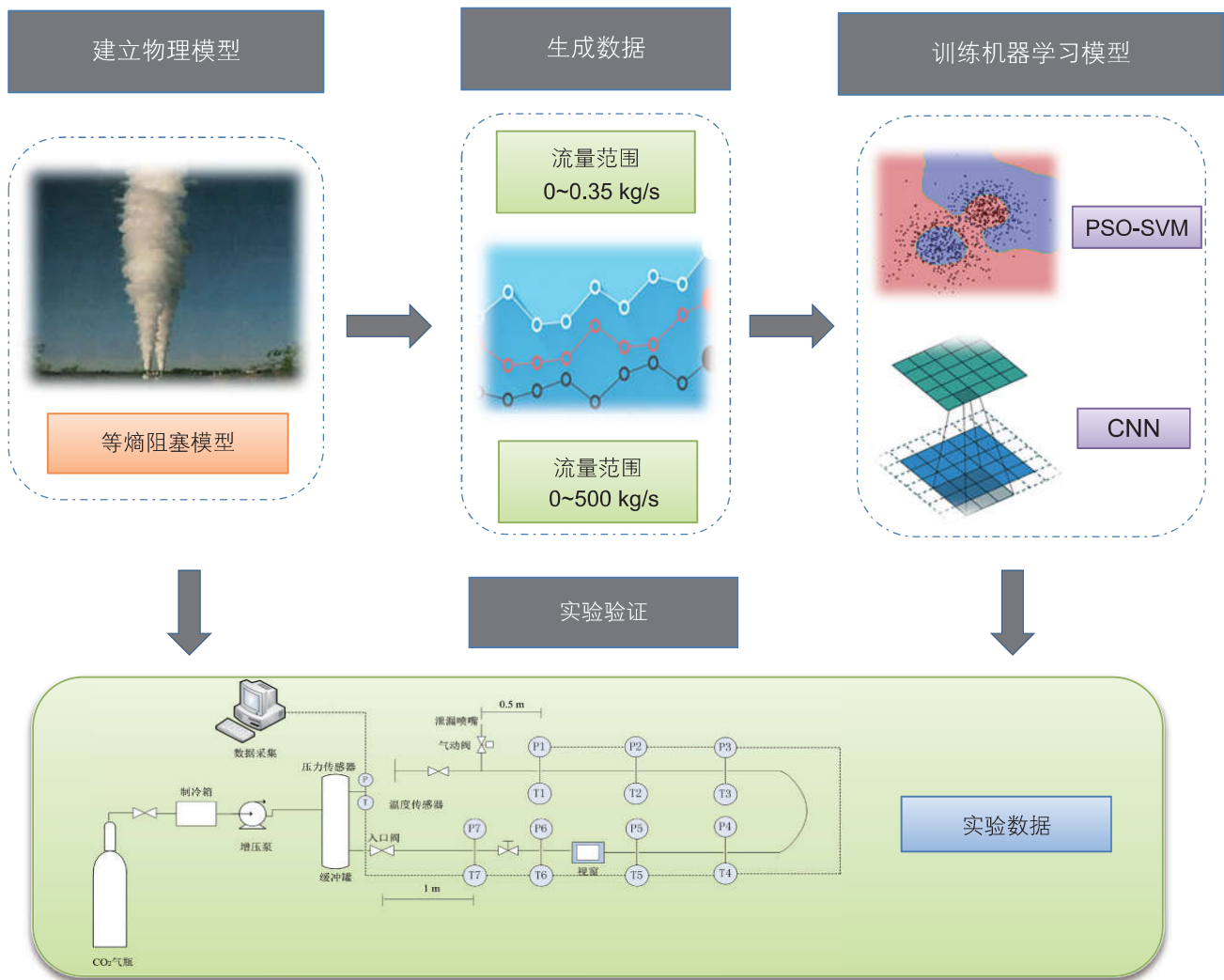


图1 物理模型驱动的机器学习方法框架图

Fig. 1 The frame diagram of physical model driven machine learning approaches

$$\text{s.t.} \begin{cases} y_i [\omega^T \varphi(x_i) + b] \geq 1 - \xi_i \\ \xi_i \geq 0 \end{cases} \quad (2)$$

式中, ω 为超平面向量, C 为惩罚因子, ξ_i 为松弛因子, $\varphi(x)$ 为非线性映射函数, b 为偏置系数。引入拉格朗日算法将最优化问题转化为对偶问题。

$$\begin{cases} \min \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) - \sum_{i=1}^n \alpha_i \\ K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j) \\ \text{s.t.} \sum_{i=1}^n \alpha_i y_i = 0, 0 \leq \alpha \leq C \end{cases} \quad (3)$$

$K(x_i, x_j)$ 为核函数, 本文使用的核函数为高斯核函数, 如式(4)

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (4)$$

式中, σ 为高斯核的带宽。影响 CO_2 最大泄漏速率的变量有很多, 如管径, 泄漏孔径, 温度, 压强等, 如果将这些因素都考虑进去来预测最大泄漏速率则需要大量数据, 所以仅选取影响较大的几个变量对最大泄漏速率进行预测。最大泄漏速率主要取决于管道输送条件(压力、温度)以及泄漏口尺寸, 因此选取泄漏孔径(d)、温度(T)、压力(p)作为自变量, 使用SVM建立预测模型, 将需要预测的数据自变量输入到SVM模型

中, 就能预测出最大泄漏速率。由于SVM在大样本量的情况下计算量过大, 求解非线性问题时难以选择合适的核函数, 所以本文通过粒子群算法(PSO)优化选取合适的核函数参数及惩罚参数, 即 PSO-SVM 算法。PSO 属于进化算法的一种, PSO 先随机初始化每个粒子, 每个粒子分别对应不同惩罚因子(C)和核参数(g)的 SVM 模型, 其属性包括粒子的速度和位置; 在寻找最优解的迭代过程中, 粒子通过跟踪两个极值来更新自己的速度和位置, 这两个极值分别为个体自身的最优解和群体的最优解。本文选取最大泄漏速率的均方误差(MSE)函数作为适应度函数, 通过适应度函数计算每个粒子的适应度来反映 SVM 模型的优劣。如图 2 所示, PSO-SVM 算法的步骤为: 先建立 SVM 模型, 初始化每个粒子的速度和位置; 然后计算粒子的适应度, 评估每个粒子并得到全局最优解; 接着更新每个粒子的速度和位置, 评估每个粒子的函数适应值; 更新每个粒子的历史最优位置, 更新群体的全局最优位置, 判断是否满足结束条件。若满足, 则将得到的最优参数导入到 SVM 模型中, 对模型进行训练; 若不满足, 则继续进行迭代。

1.2 卷积神经网络算法

卷积神经网络(CNN)的理论基础来自 1968 年

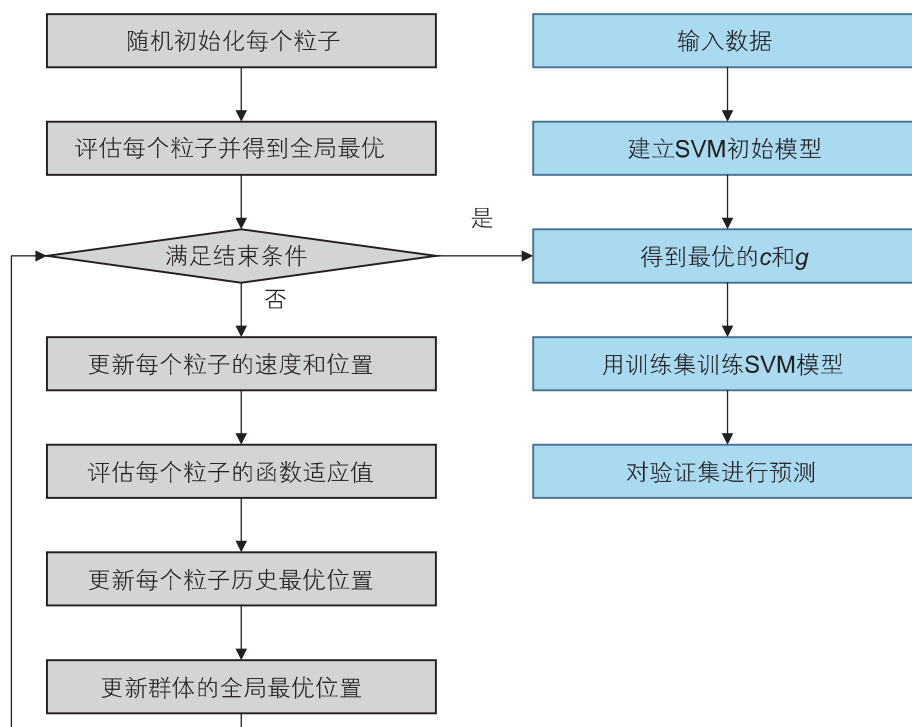


图 2 PSO-SVM 算法流程图

Fig. 2 PSO-SVM algorithm flow chart

Hubel和Wiesel研究猫视觉皮层的论文, Krizhevsky在2012年使用卷积神经网络大幅提高了图像识别的准确率, 此后CNN被大量应用于视觉识别领域^[21]。CNN不需要精确的数学公式, 只需用数据训练就能够实现输入到输出的映射。卷积层权重共享的特性也在很大程度上减少了参数, 降低了卷积网络的复杂程度。由于CNN具有局部感知、权重共享等特点, 所以CNN也可以用来解决回归问题

如图3所示, CNN结构包括输入层和隐含层。泄漏孔径(d)、温度(T)、压力(p)由输入层进入隐含层, 其中隐含层通常由卷积层、池化层和全连接层组成。卷积层是CNN的核心, 一般由一组卷积核构成, 卷积核感受野较小, 卷积核能够提取泄漏孔径(d)、温度(T)、压力(p)的更高层次特征。卷积计算公式为

$$Z = W \cdot I + b \quad (5)$$

式中, Z 为中间值, W 为权重系数, I 为输入的泄漏孔径(d)、温度(T)、压力(p), b 为偏置系数。泄漏孔径(d)、温度(T)、压力(p)经过卷积层进行特征提取, 通过神经元的激活函数实现去线性。为了避免CNN反向传播时梯度消失, 采用ReLU激活函数, 其公式为

$$\text{ReLU}(z) = \max(0, z) \quad (6)$$

通常卷积层输出非常多的参数量, 所以常在卷积层后添加池化层, 池化层能从卷积层的输出中提取最显著的特征, 以减少数据和参数量。由于本文中的CNN被用于非线性回归, 所以对CNN进行了简化处理, 不采用池化层。经过上述过程获取的高级特征通过全连接层进行处理, 得到最大泄漏速率。

2 超临界CO₂最大泄漏速率预测物理模型

超临界CO₂管道的泄漏过程是一种高压流体的快

速射流过程, 在泄漏过程的初始阶段, 管内介质的泄漏速度会超过CO₂的当地声速, 在泄漏口处形成壅塞流动^[22]。本文采用等熵阻塞流动模型^[23]计算超临界CO₂管道最大泄漏速率。由于管内初始压力较高, 发生泄漏时泄漏口处流速较大, 忽略泄漏过程管内外热量交换, 可以假设泄漏口处为等熵流动状态。根据能量守恒方程可得:

$$h_0 + \frac{u_0^2}{2} = h_1 + \frac{u_c^2}{2} \quad (7)$$

式中: h_0 为管内介质的初始焓值, kJ/kg; h_1 为管道泄漏口处的介质焓值, kJ/kg; u_c 为泄漏口处介质流速, m/s; u_0 为管内介质流速, 由于管内流体速度较泄漏口处的流速小很多, 因此可以忽略不计。由式(7)可得泄漏速度:

$$u_c = \sqrt{2(h_0 - h_1)} \quad (8)$$

以上二氧化碳的物性参数(压力、温度、焓、熵)都采用GERG2008状态方程计算。泄漏质量流量为:

$$G = \rho_c A u_c \quad (9)$$

式中, G 为管道泄漏质量流量, kg/s; ρ_c 为泄漏口处介质密度, kg/m³; A 为泄漏口面积, m²。

图4为最大泄漏速率计算程序框图。由以上物理模型可见, 最大泄漏速率主要取决于管道输送条件(压力、温度)以及泄漏口尺寸。

3 机器学习的预测结果及讨论

3.1 模型参数选择

3.1.1 PSO-SVM参数选择

本文机器学习模型的编程语言为Python。根据理论研究结果, 本文将泄漏孔径(d)、温度(T)、压力(p)

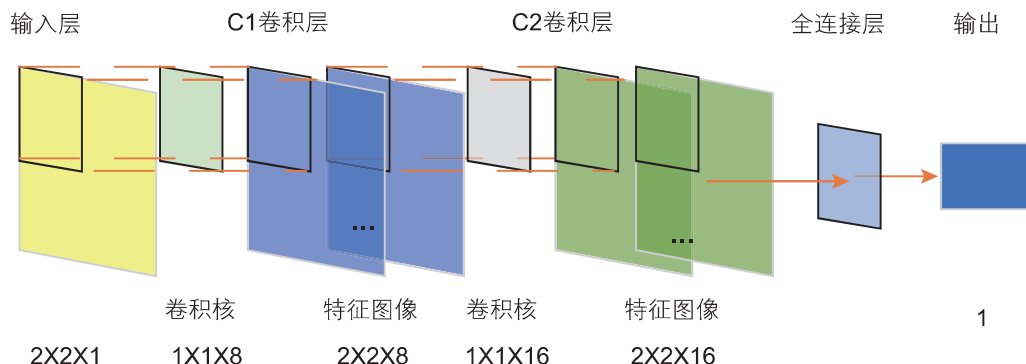


图3 本文所用卷积神经网络算法示意图

Fig. 3 The algorithm schematic diagram of convolutional neural network in this paper

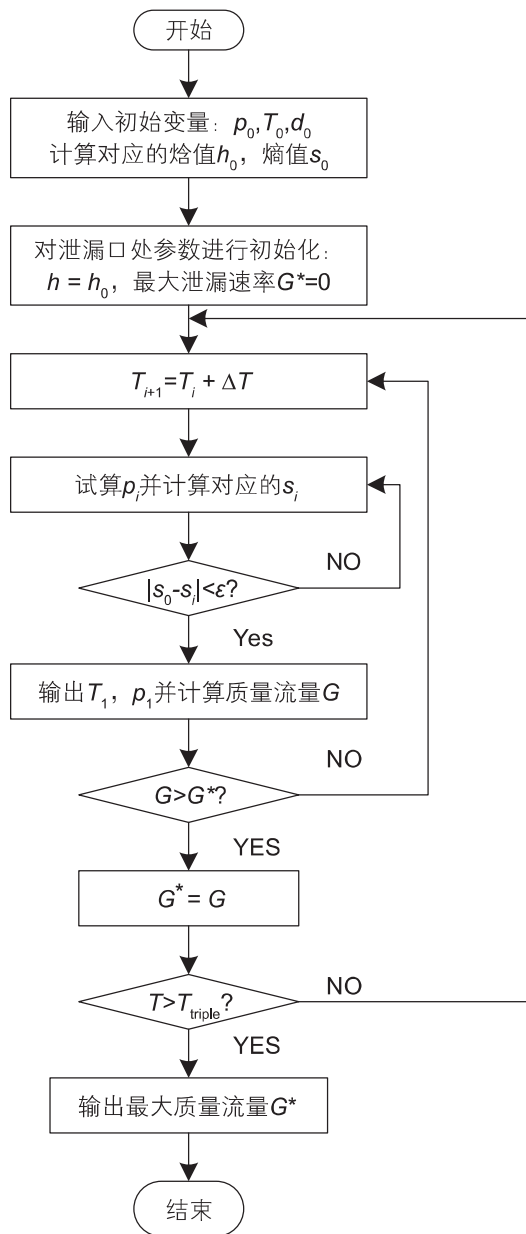


图 4 最大泄漏速率计算框图

Fig. 4 Block diagram for calculation of maximum leakage rate

这 3 个因素作为影响超临界二氧化碳管道最大泄漏速率的特征指标。本文中支持向量机的核函数选用高斯核函数(RBF)，将数据集分为 10 个交叉验证集，在 PSO 中将惩罚因子 C 的取值范围 $[10.0, 500.0]$ ，核函数参数 g 取值范围 $[0.01, 0.1]$ ，适应度函数选择均方误差(MSE)函数，群体搜索的最大迭代次数为 1，搜索远离粒子最优位置的比例因子为 10，搜索终止前群体最优目标值的最小变化值设置为 0.001。种群的粒子数设置为 200。最终经过 PSO 优化之后找到的最优惩罚因子 C 为 476.248，核函数参数 g 为 0.099 048。

3.1.2 CNN 参数选择

由于 CNN 通常被用于图像识别，当将其用于非线性回归时，需要调整 CNN 的结构。本文删除了 CNN 结构中常用的 softmax 层，同时由于本文数据量较小，经过池化后部分局部特征可能无法被识别，导致特征提取效果变差，所以也不采用池化层。

本文的卷积神经网络包括输入层、两个卷积层、一个全连接层以及输出层。先增加一列特征指标，使特征指标变成 4 列，然后对数据进行标准化处理，最后将每行特征指标变为 2×2 的矩阵；经过数据处理，输入层接收的是 $2 \times 2 \times 1$ 的矩阵图像，其中 1 表示单通道。两个卷积层的步长都设置为 1，卷积核尺寸设置为 1×1 ，第一卷积层设置 8 个卷积核，输出的通道数为 8，第二卷积层设置 16 个卷积核，输出的通道数为 16。经过第一层卷积输出 8 个大小为 2×2 的特征图像，经过第二层卷积输出 16 个大小为 2×2 的特征图像。最终全连接层将 16 个 2×2 的图像变成含有 $2 \times 2 \times 16$ 个元素的一维数组，对该行数据进行卷积，卷积核个数为 64，得到含有 64 个元素的一维数组。在训练过程中，为了减少过拟合，采用 dropout 方法，保留数组中 90% 的元素；在预测过程中则全部保留，最后对该数组再进行卷积，卷积核个数为 1，得到长度为 1 的数组。

3.2 模型预测结果分析

3.2.1 机器学习效果对比

本文用来训练模型的数据是通过上述的物理模型生成的。使用该物理模型生成泄漏孔径范围为 1~10 mm，压力范围为 7.5~12 MPa，温度范围为 305~325 K 的数据矩阵，将其中 80% 数据设为训练集，20% 设为测试集。利用训练集训练机器学习模型，通过测试集对训练好的模型进行测试。图 5 对比了 PSO-SVM 和 CNN 的预测结果。x 轴对应物理模型的预测值，y 轴对应机器学习模型的预测值， $y = x$ 这条线代表物理模型预测值和机器学习模型预测值相同，越接近这条线，预测结果越准确。

由图 5 可见，在小孔泄漏工况下两个机器学习模型的预测结果都比较理想。对比 PSO-SVM 模型与 CNN 模型发现，PSO-SVM 模型的决定系数 (R^2) 比 CNN 模型更接近 1，说明 PSO-SVM 模型比 CNN 模型的拟合程度更好。但 CNN 模型的均方误差 (MSE) 和平均绝对误差 (MAE) 都比 PSO-SVM 模型的结果低。进一步对比发现，PSO-SVM 的均方误差约为 CNN 的 30 倍，说明 PSO-SVM 模型的数据变化程度比 CNN 模型

大,但两者的平均绝对误差相差不大,说明两模型预测值误差大致相同。

3.2.2 模型的泛化能力

由于所训练的机器学习模型是针对小孔泄漏工况的数据集的,因此有必要验证模型的泛化能力。由物理模型均匀产生 5938 组用于对比验证的数据集,该数据集中的参数范围为: $d = [10, 100]\text{mm}$, $T = [310, 330]\text{K}$, $p = [8.5, 12]\text{MPa}$ 。可以看出,该压力、温度、泄漏孔径的范围是对用于训练机器学习模型的小孔泄漏数据集范围的扩展,验证数据集的参数范围是所建

立的机器学习模型之前没有接触过的。

如图 6 所示,采用物理模型的预测结果作为 x 轴,机器学习模型的预测结果作为 y 轴。 $y=1.2x$ 、 $y=1.1x$ 、 $y=x$ 、 $y=0.9x$ 、 $y=0.8x$ 这 5 条线分别代表机器学习模型的预测值与物理模型预测值的误差为 20%、10%、0、-10%、-20%。从图中可见,虽然 CNN 和 PSO-SVM 模型只学习了泄漏量在 0~5.33 kg/s 的数据,但两者对泄漏量在 0~500 kg/s 之间的数据的预测结果误差大部分都在 20% 以内。对比两模型发现, SVM 模型预测结果的误差明显小于 CNN 模型预测结果误差,

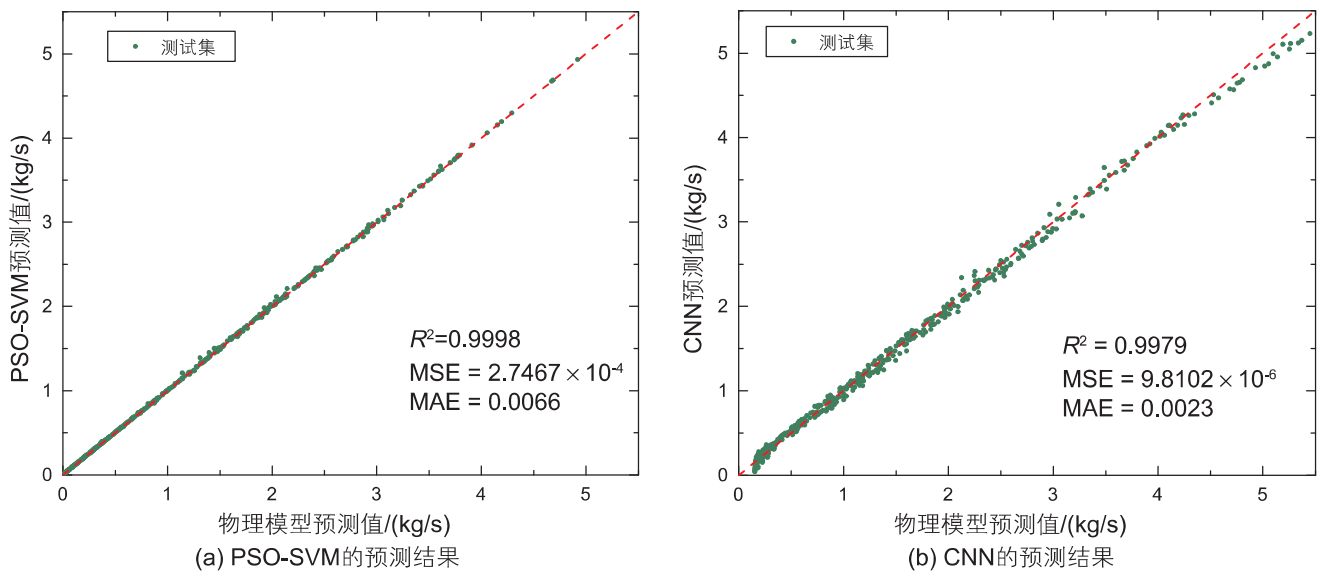


图 5 小孔泄漏工况下最大泄漏速率预测结果

Fig. 5 Forecast results of the maximum leakage rate under the condition of small hole leakage

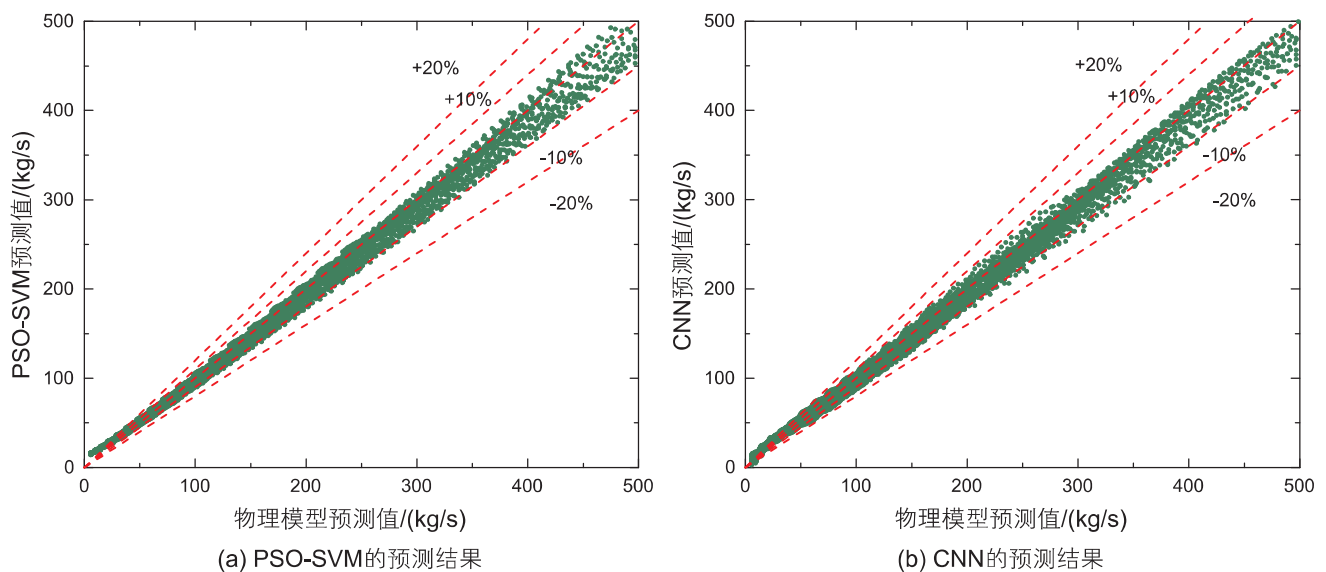


图 6 机器学习模型泛化预测能力对比

Fig. 6 Machine learning models generalize predictive power comparison

SVM模型预测结果的误差基本在 10% 以内，CNN 模型 SVM 模型预测结果的误差基本在 20% 以内。由表 1 可知，PSO-SVM 模型的均方误差(MSE)、平均绝对误差(MAE)均小于 CNN，且决定系数(R^2)大于 CNN，说明 SVM 模型的泛化能力较强。由此可见，PSO-SVM 的泛化能力强于 CNN，但在模型的训练时间方面，CNN 远优于 SVM。针对该训练集 CNN 仅仅花了 6 s，而 SVM 花了 10 min。这还不包括 PSO 为寻找最优超参数所花的时间(超过 4 h)。

3.2.3 实验验证

由于以上的验证数据集都来自于物理模型，是否符合实际情况有待考量，因此本文采用 8 组实验数据^[24-29]分别对 3 个模型(物理模型、PSO-SVM、CNN)进行对比验证。实验组数据见表 2。由图 7 可知，3 个模型的预测值与真实值基本接近，经过计算，误差都在 28.82% 以内。由表 3 可知，3 个模型的 R^2 都非常接近 1，说明在这 9 组数据中 3 个模型的拟合程度较好，但物理模型的 R^2 略高于机器学习模型，这主要是由于训练机器学习模型的数据来自物理模型。

表 1 机器学习模型预测误差

Table 1 Machine learning models predict errors

	R^2	MSE	MAE
PSO-SVM	0.9943	0.0057	0.0523
CNN	0.9897	0.0103	0.0726

表 2 实验组参数

Table 2 Experimental group parameters

编号	泄漏孔径/mm	初始压力/kPa	初始温度/K	最大泄漏量/(kg/s)
Test 1	2	8000	313	0.09
Test 2	2.764	8000	313	0.2
Test 3	3	8000	313	0.227
Test 4	1	9910	315.5	0.033
Test 5	1.01	8140	301	0.04
Test 6	1.01	10130	308	0.04
Test 7	3.175	12500	293	0.9
Test 8	2.5	7970	313	0.15

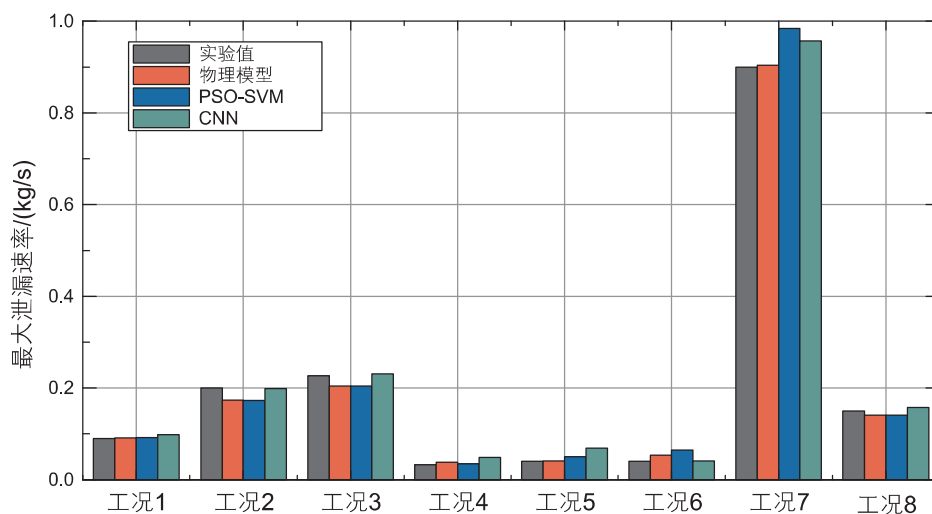


图 7 3 种模型的实验对比验证结果

Fig. 7 The experimental comparison of the three models verified the results

表3 物理模型和机器学习模型的预测误差对比

Table 3 Comparison of prediction errors between physical models and machine learning models

	R^2	MSE	MAE
Model	0.9974	1.903×10^{-4}	0.0103
PSO-SVM	0.9843	1.143×10^{-3}	0.0226
CNN	0.9923	5.574×10^{-4}	0.0154

综上分析,3种模型各有优劣:PSO-SVM模型的预测效果最佳,但训练模型花费的时间较长;相比于PSO-SVM模型,CNN模型的训练时间较短,但机器学习方法不可避免存在黑箱问题;物理模型虽然不存在黑箱问题,但建模过程较为复杂,计算耗时较长。在计算最大泄漏速率的过程中,物理模型计算一组数据的时间大约2s,处理大批量的数据(5938组)时花费时间较长(3 h18 min),无法实现最大泄漏速率的实时预测;而机器学习模型计算同样的大批量数据(5938组)仅需不到1s,计算速度远高于物理模型,这是实现最大泄漏速率的实时预测的关键。

4 结论

本文分别采用两种改进的机器学习算法(PSO-SVM和CNN)预测了超临界CO₂管道最大泄漏速率,测试并分析了机器学习模型的准确率和泛化能力,并通过与实验数据的对比分析,得到以下主要结论。

(1)物理模型、PSO-SVM、CNN的预测结果与实验数据接近,平均误差为28.82%。

(2)PSO-SVM和CNN对小孔泄漏数据的预测结果表明,PSO-SVM模型和CNN模型对小孔泄漏工况的预测精度差别不大,但在对小孔泄漏工况扩展数据集的预测中,PSO-SVM模型的均方误差(MSE),平均绝对误差(MAE),决定系数(R^2)都优于CNN,PSO-SVM模型预测结果的误差基本在10%以内,CNN模型预测结果的误差则在20%以内,PSO-SVM模型的泛化能力强于CNN,但CNN的训练时间远短于PSO-SVM。因此,SVM适用于小样本数据精确预测,而CNN更适用于对大数据的学习和预测。

(3)本研究展示了机器学习在预测超临界CO₂管道最大泄漏速率方面的可行性。研究成果不局限于CO₂管道,对于含相变的多相流管道同样适用。尽管机器学习方法能够抛开复杂的物理模型,只关注输入条件和输出结果,但是仍然存在黑箱问题。基于物理模型驱动的机器学习或许有助于解决该问题。由于缺少现场数据,本文用物理模型生成样本数据,若能够建立管道运行数据库(包括各类失效数据),用实际运行数据作为数据训练集,有望实现对实时最大泄漏速率的预测,为管道风险评估提供参考。

参考文献

- [1] CAI W J, WANG G J, SANTOSO A, et al. Increased frequency of extreme La Niña events under greenhouse warming[J]. Nature Climate Change, 2015, 5(2): 132-137.
- [2] PARSONS W. Strategic analysis of the global status of carbon capture and storage[R]. Global CCS Institute, 2009.
- [3] 吴瑕,李长俊,贾文龙.二氧化碳的管道输送工艺[J].油气田地面工程,2010,29(09):52-53.[WU X, LI C J, JIA W L. Process of pipeline transportation for carbon dioxide [J]. Oil-Gas Field Surface Engineering, 2010, 29(09): 52-53.]
- [4] KLING GW, CLARK MA, WAGNER G, et al. The 1986 lake Nyos gas disaster in cameroon, west Africa[J]. Science, 1987, 236(4798): 169-175.
- [5] MAHGEREFTEH H, DENTON G, RYKOV Y. A hybrid multiphase flow model[J]. AIChE Journal, 2008, 54(9): 2261-2268.
- [6] TENG L, LI Y X, ZHAO Q, et al. Decompression characteristics of CO₂ pipelines following rupture[J]. Journal of Natural Gas Science and Engineering, 2016, 36(Pt. A): 213-223.
- [7] MARTYNOV S, ZHENG W T, MAHGEREFTEH H, et al. Computational and Experimental Study of Solid-Phase Formation during the Decompression of High-Pressure CO₂ Pipelines[J]. Industrial & Engineering Chemistry Research, 2018, 57(20): 7054-7063.
- [8] BROWN S, MARTYNOV S, MAHGEREFTEH H, et al. A homogeneous relaxation flow model for the full bore rupture of dense phase CO₂ pipelines[J]. International Journal of Greenhouse Gas Control, 2013, 17(1): 349-356.
- [9] WOOLLEY R M, FAIRWEATHER M, WAREING C J, et al. An integrated, multi-scale modelling approach for the simulation of multiphase dispersion from accidental CO₂ pipeline releases in realistic terrain[J]. International Journal of Greenhouse Gas Control, 2014, 27(1): 221-238.

- [10] MUNKEJORD S T, HAMMER M. Depressurization of CO₂-rich mixtures in pipes: Two-phase flow modelling and comparison with experiments[J]. *International Journal of Greenhouse Gas Control*, 2015, 37(1): 398–411.
- [11] KESSLER R C, VAN LOO H M, WARDENAAR K J, et al. Testing a Machine-learning Algorithm to Predict the Persistence and Severity of Major Depressive Disorder From Baseline Self-reports[J]. *Mol Psychiatry*, 2016, 21(10): 1366–1371.
- [12] 冯明刚, 严伟, 葛新民, 等. 利用随机森林回归算法预测总有机碳含量[J]. *矿物岩石地球化学通报*, 2018, 37(3): 475–481. [FENG M G, YAN W, GE X M, et al. Predicting total organic carbon content by random forest regression algorithm [J]. *Bulletin of Mineralogy, Petrology and Geochemistry*, 2018, 37(3): 475–481.]
- [13] 黄天羽, 郭芸莹. 面向动作捕捉的非线性时间序列预测方法研究[J]. *系统仿真学报*, 2018, 30(7): 2808–2815. [HUANG T Y, GUO Y Y. Research of nonlinear time series prediction method for motion capture [J]. *Journal of System Simulation*, 2018, 30(7): 2808–2815.]
- [14] SHAO D, YAN Y, ZHANG W B, et al. Dynamic measurement of gas volume fraction in a CO₂ pipeline through capacitive sensing and data driven modelling[J]. *International Journal of Greenhouse Gas Control*, 2020, 94: 102950.
- [15] 马贵阳, 宫清君, 潘振, 等. 基于支持向量机结合遗传算法的天然气水合物相平衡研究[J]. *天然气工业*, 2017, 37(05): 46–52. [MA G Y, GONG Q J, PAN Z, et al. GA-SVM based study on natural gas hydrate phase equilibrium [J]. *Natural Gas Industry*, 2017, 37(05): 46–52.]
- [16] 夏俏健, 高辉, 高长征, 等. 基于PCA-SVM的油气管道腐蚀速率预测技术研究[J]. *油气田地面工程*, 2020, 39(04): 74–78. [XIA Q J, GAO H, GAO C Z, et al. Prediction technology study of oil and gas pipeline corrosion rate based on PCA-SVM [J]. *Oil-Gas Field Surface Engineering*, 2020, 39(04): 74–78.]
- [17] 李华杰. 基于T-S模糊神经网络预测算法的管道泄漏定位研究与分析[J]. *计算技术与自动化*, 2016, 35(03): 63–67. [LI H J. Study and analysis of gas pipeline leak location based on T-S fuzzy neural network prediction algorithm[J]. *Computing Technology and Automation*, 2016, 35(03): 63–67.]
- [18] XU D, JIANG J C, SHEN G D, et al. An improved variational mode decomposition method based on particle swarm optimization for leak detection of liquid pipelines[J]. *Mechanical Systems and Signal Processing*, 2020, 143(1): 无.
- [19] 张丽娟, 李帆, 王文龙. 两种基于模式识别的枝状燃气管网泄漏定位方法[J]. *天然气工业*, 2007(08): 106–108+143. [ZHANG L J, LI F, WANG Y L. Two leakage positioning methods for dendritic gas pipeline network based on pattern recognition[J]. *Natural Gas Industry*, 2007(08): 106–108+143.]
- [20] LANG X M, LI P, ZHANG B C, CAO J T, et al. Localization of multiple leaks in a fluid pipeline based on ultrasound velocity and improved GWO[J]. *Process Safety and Environmental Protection*, 2020, 137(1): 1–7.
- [21] 管皓, 薛向阳, 安志勇. 深度学习在视频目标跟踪中的应用进展与展望[J]. *自动化学报*, 2016, 42(06): 834–847. [GUAN H, XUE X Y, AN Z Y. Advances on application of deep learning for video Object tracking [J]. *Acta Automatica Sinica*, 2016, 42(06): 834–847.]
- [22] LI K, ZHOU X J, TU R, et al. An experimental investigation of supercritical CO₂ accidental release from a pressurized pipeline[J]. *The Journal of Supercritical Fluids*, 2016, 107(1): 298–306.
- [23] GU S W, LI Y X, TENG L, et al. A new model for predicting the decompression behavior of CO₂ mixtures in various phases[J]. *Process Safety and Environment Protection*, 2018, 120(1): 237–247.
- [24] GU S W, LI Y X, TENG L, et al. An experimental study on the flow characteristics during the leakage of high pressure CO₂ pipelines[J]. *Process Safety and Environment Protection*, 2019, 125(1): 92–101.
- [25] LI K, ZHOU X J, TU R, et al. Investigation of flow characteristics in small-scale highly pressurized leaked CO₂ jet plume from pipeline[J]. *International Journal of Thermal Sciences*, 2019, 141: 160–170.
- [26] FAN X, WANG Y, ZHOU Y, et al. Experimental study of supercritical CO₂ leakage behavior from pressurized vessels[J]. *Energy*, 2018, 150: 342–350.
- [27] WANG Y, ZHOU Y, HUANG Y, et al. Modelling of accidental release process from small rupture of pressure CO₂ vessel[J]. *International Journal of Greenhouse Gas Control*, 2020, 93: 102906.
- [28] AHMAD M, BÖGEMANN-VAN Osch M, BUIT L, et al. Study of the thermohydraulics of CO₂ discharge from a high pressure reservoir[J]. *International journal of greenhouse gas control*, 2013, 19: 63–73.
- [29] ZHOU X, LI K, TU R, et al. A modelling study of the multiphase leakage flow from pressurised CO₂ pipeline[J]. *Journal of hazardous materials*, 2016, 306: 286–294.