

DOI: 10.19650/j.cnki.cjsi.J2108488

基于三轴融合的漏磁内检测数据缺陷反演方法研究*

卢森骧¹, 神祥凯¹, 张俊楠¹, 刘金海¹, 赵可天²

(1. 东北大学信息科学与工程学院 沈阳 110004; 2. 中海油能源发展装备技术有限公司 天津 300452)

摘要:在管道漏磁检测中,缺陷反演是管道故障诊断的核心部分。考虑漏磁信号的复杂性以及管道环境的多变性,常用的缺陷反演方法多采用传感器单轴信息,从而导致缺陷反演面临缺陷估计尺寸精度低、模型通用性差的问题,难以满足实际应用需求。本文提出基于三轴融合的漏磁内检测数据缺陷反演算法,显著提高漏磁缺陷反演精度。该方法主要由两部分组成,首先,利用提出的加权随机森林算法分别实现单轴信号的缺陷反演;其次,通过本文设计的模糊推理系统实现三轴反演结果决策融合,进而得到精确的缺陷估计尺寸。最后,通过仿真数据与实际管道数据实现该方法的评估。实验结果表明,该方法缺陷反演的长度精度提升23%,宽度精度提升13%,深度精度提升14.7%,具有较好的实验效果。

关键词:漏磁检测;缺陷反演;加权随机森林;三轴决策融合;模糊推理

中图分类号:TE88 TP277 TH878⁺3 文献标识码:A 国家标准学科分类代码:440.55

Research on defect inversion method of magnetic flux leakage internal inspection data based on triaxial fusion

Lu Senxiang¹, Shen Xiangkai¹, Zhang Junnan¹, Liu Jinhai¹, Zhao Ketian²

(1. School of Information Science and Engineering, Northeastern University, Shenyang 110004, China;

2. CNOOC Energy Development Equipment Technology Co., Ltd., Tianjin 300452, China)

Abstract: In the pipeline magnetic flux leakage detection, defect inversion is the core part of pipeline fault diagnosis. Considering the complexity of the magnetic flux leakage signal and the variability of the pipeline environment, the commonly defect inversion methods mostly use sensor uniaxial information, which may cause the defect inversion to bring the problems of low defect estimation size accuracy and poor model versatility. It is difficult to meet the requirement of practical application. This article proposes a three-axis fusion-based defect inversion algorithm for magnetic flux leakage internal inspection data, which significantly improves the inversion accuracy of magnetic flux leakage defect. The method mainly consists of two parts. First, the proposed weighted random forest algorithm is used to realize the defect inversion of single-axis signals. Secondly, the three-axis inversion result decision fusion is achieved through the designed fuzzy inference system. Then, the precise defect size is achieved. Finally, the evaluation of the method is realized through simulation data and practical pipeline data. Experimental results show that the length accuracy of the defect inversion method is increased by 23%, the width accuracy is increased by 13%, and the depth accuracy is increased by 14.7%, which have good experimental results.

Keywords: magnetic flux leakage detection; defect inversion; weighted random forest; three-axis decision fusion; fuzzy reasoning

0 引言

近年来,随着国民经济的不断发展,能源需求急剧增加,油气资源在生产生活中愈发重要。由于管道运输具备可靠性高、运输成本低且不易受气候影响的优势,管道

运输成为油气资源的主要运输方式之一^[1]。

但是随着时间推移,管道往往出现腐蚀、老化和管壁损伤等问题,或因管道内部的潜在缺陷发展成破损引起泄漏事故,这不仅会造成巨大的经济损失,同时也会对生态环境造成一定的影响,并且对人身安全有着极大的威胁。因此,必须定期对油气管道进行检查以及维护,进而

收稿日期:2021-08-28 Received Date: 2021-08-28

* 基金项目:国家自然科学基金(61627809, 61973071, 61703087, 62003080)、辽宁省自然科学基金(2019-KF-03-04)项目资助

保证油气运输过程的安全与高效^[2-3]。

常用的几种检测技术包括超声检测^[4]、涡流检测^[5]以及漏磁检测技术^[6]。其中漏磁内检测技术是目前公认的最有效的铁磁材料无损检测方式^[7-8]，管道漏磁内检测技术通常分为数据预处理、缺陷组件识别、缺陷反演以及管道状态评估4部分，其中缺陷反演是通过漏磁内检测采集的漏磁信号进行有效分析，进而实现缺陷尺寸的估计，是管道故障诊断的核心环节。近年来，国内外众多学者对缺陷反演展开研究。漏磁缺陷反演方法根据工作原理的不同，主要分为基于模型驱动的间接反演方法和基于数据驱动的直接反演方法^[9]。

间接反演法是基于物理机制模拟缺陷信号，通过不断优化参数重建缺陷轮廓。文献[10]提出基于磁通量实现缺陷长度和深度的估计。文献[11]提出一种改进的MHS算法，用于重建缺陷轮廓。间接反演法的求解过程非常复杂，具有计算量大且耗费大量时间的缺点，很难满足工程应用需求。

而直接反演方法往往通过对缺陷数据进行特征提取，建立缺陷信号与缺陷尺寸之间的非线性映射，进而获得准确的缺陷估计尺寸，能够较好地满足实际工程需求。文献[9]提出了一种基于WT-STACK的复杂缺陷反演算法，通过研究特征与缺陷尺寸间的非线性关系，以此构建了动态多轴时频域特征作为输入的迭代堆叠自适应学习模型。文献[12]提出了一种基于强化学习的缺陷深度轮廓重建算法，通过将基于迭代方法的迭代过程嵌入到强化学习过程中以生成学习策略，避免迭代策略设计的问题。

无论是直接反演或是间接反演算法，其信息往往依赖单轴传感器采集的漏磁数据，或者无法充分将漏磁信号三轴信息相互融合，从而导致缺陷反演结果具有一定的片面性，进而导致缺陷反演精度低，模型泛化能力差，很难满足实际工程需求。

近年来，随着科技发展以及工程应用需求，信息采集趋向多元化，单一传感器无法满足实际工程需求，多传感器信息融合技术应运而生。文献[13]提出信息不完备条件下漏磁缺陷表征的多传感器融合，文献[14]提出基于多样性特征和多源信息的分子泵故障诊断算法，文献[15]提出基于多源异构信息融合的机械故障诊断方法。上述的融合方法大多基于数据层融合和特征层融合，一定程度上存在灵活性差、容错性弱的问题。

针对缺陷反演精度低、模型通用性差的问题，本文提出基于三轴融合的漏磁内检测数据缺陷反演方法，将漏磁数据三轴信息分别通过加权随机森林实现缺陷的初步反演，进而通过模糊系统实现三轴决策融合得到最终反演结果。试验表明，本文所提算法，显著提高漏磁缺陷的

反演精度。本文创新如下：

1) 提出一种改进的加权随机森林算法。该方法通过不同层次的加权方式，显著提高缺陷反演精度，并具有较好的模型泛化性。

2) 基于漏磁三轴信号之间的特征关系，设计一种基于模糊推理的决策融合算法。该算法充分融合三轴决策信息，提高缺陷反演精度。

1 整体方法介绍

本文提出一种基于三轴融合的漏磁缺陷反演方法，算法的整体框架如图1所示。基于三轴融合的漏磁缺陷方法主要分为两部分：基于加权随机森林的缺陷反演方法和基于模糊推理的三轴决策融合算法。

加权随机森林算法主要包括特征提取、预测网络两部分。首先，三轴信号构建完备的数据集，然后，为充分获取漏磁信号的特征信息，分别提取三轴漏磁信号时域特征与Markov特征，并通过特征选择，去除冗余特征，具体流程如图1中部分2)所示。将获取的特征送入本文提出的改进的随机森林预测网络，通过改进的随机森林预测网络分别生成轴向、径向、周向信号对应的缺陷的长、宽、深，预测网络流程如图1中部分3)所示。

基于模糊推理的三轴决策融合算法。将加权随机森林算法的输出，作为模糊推理决策系统的输入。将单轴反演结果的长、宽、深分别送入图1中的部分4)所示的模糊推理系统，得到融合后的缺陷长、宽、深。缺陷长、宽、深分别对应相应的模糊推理系统，实现缺陷反演的三轴决策融合。将清晰的三轴模糊输入经过模糊化后生成模糊输出，然后通过本文制定的模糊规则实现三轴决策融合，最终通过反模糊化生成融合后的清晰输出，即基于三轴融合信息生成的缺陷反演尺寸。

2 基于加权随机森林的缺陷反演算法

由于随机森林每个基学习器重点学习的样本以及关注的特征不同，使得每个学习器的学习能力不同，预测误差也有所差异，而最终结果为不同学习器的均值，会导致学习能力较差的决策树影响模型回归的精度。

针对上述问题，本文提出一种基于代价敏感学习思想的随机森林算法，即通过引入不同基学习器的预测能力及所选择特征的重要度来对每棵决策树的权重因子进行设计，该方法可以对三轴漏磁数据提取出的特征与缺陷尺寸间的关系进行很好地描述，进而提高缺陷反演的精度。该算法包括特征提取以及预测网络权重设计两部分，算法框架如图1中部分2)、3)所示。

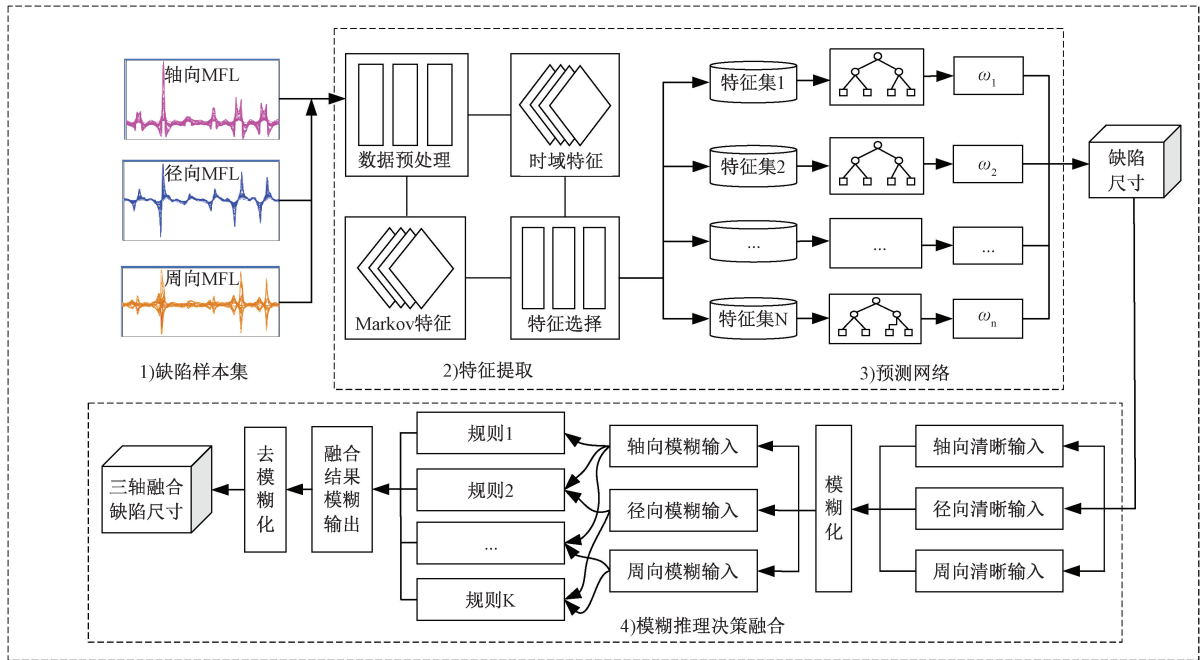


图 1 基于三轴融合漏磁内检测数据的缺陷反演方法框架

Fig. 1 Framework of the defect inversion method based on three-axis fusion magnetic flux leakage internal inspection data

2.1 特征提取

漏磁三轴数据主要包括漏磁信号轴向数据、径向数据以及周向数据,构建完备的三轴漏磁数据集后,为充分挖掘缺陷相关特征信息,分别提取数据时域特征以及 Markov 特征。

1) 时域特征提取

(1) 显性特征提取:提取轴向信号的谷谷间距、径向信号的峰谷间距、周向信号的峰谷间距等显性特征用于缺陷反演,如图 2 所示为轴向显性特征示意图。图中, h_1 为左峰谷差, h_2 为右峰谷差, h_3 为谷谷差。

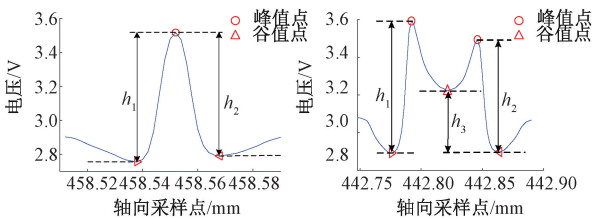


图 2 不同缺陷轴向显性特征示意图

Fig. 2 Schematic diagram of the axial dominance characteristics of different defects

(2) 统计特征提取:除上述信号峰谷差等显性特征外,考虑从漏磁数据的集中趋势、离散程度、分布形态 3 方面,提取统计特征作为其时域特征一部分,主要包括极差、方差以及变异系数 3 部分。

反映数据段样本伸展情况的特征-极差,如式(1)所

示。衡量数据段样本的绝对波动程度的特征-方差,具体如式(2)所示。

$$R = \max(x_i) - \min(x_i) \tag{1}$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n-1} \left(\sum_{i=1}^n X_i^2 - n\bar{X}^2 \right) \tag{2}$$

表示数据段样本的相对波动程度的特征-变异系数,具体如式(3)所示。

$$V = \frac{s}{\bar{X}} \times 100\% \tag{3}$$

2) Markov 特征提取

Markov 特征提取的思路为将三轴漏磁数据的每个通道的信号分为构造成马尔可夫链^[16],然后进一步提取其状态转移次数、状态转移概率等动态特征,这些特征可以反映出管道中出现异常时的变化信息。Markov 特征提取过程分为如下 3 个步骤。

(1) 构造 Markov 链

采用合适的缺陷检测方法可以实现缺陷的精准定位,然后将定位框中的漏磁数据当做一个窗体,设某一里程范围内的窗体为, $X = \{x_1, x_2, \dots, x_L\}$ 通过选择合适的上下边界可以将漏磁数据划分成不同的状态即将窗体范围内的漏磁信号构造成 Markov 链。

(2) 提取 Markov 特征

本文中的 Markov 特征指的是 Markov 链动态变化的状态信息,这里提出两种 Markov 特征,一种是基于各个状态的转移情况,一种是基于状态转移概率矩阵,状态 i 的转移情况定义如式(4)所示。

$$\begin{cases} k_i^+ = \sum_{j=1}^L \mathbf{1}_{(s(j)=i \& s(j+1) > i)} \\ k_i^- = \sum_{j=1}^L \mathbf{1}_{(s(j)=i \& s(j+1) < i)} \\ k_i = \sum_{j=1}^L \mathbf{1}_{(s(j)=i \& s(j+1) = i)} \end{cases} \quad (4)$$

式中: k_i^+ 为状态上移次数, k_i^- 为状态下移次数, k_i 为状态保持次数, $s(j)$ 为 j 时刻的状态, $s(j+1)$ 为 $j+1$ 时刻的状态, L 表示窗体内数据段的长度。

(3) 第二重边界比例系数和中间状态数灵敏度分析

由于选择不同的边界比例系数和划分的状态数会将漏磁信号转换为不同的 Markov 链, 会导致提取的特征各不相同, 从而使得反演的效果不同, 因此需要合理选择第二重边界比例系数 λ_2 和中间状态数 n 的灵敏度。

2.2 基于决策树预测能力的权重设计

1) 指标选择

随机森林是不同决策树的集合, 针对不同决策树预测能力的好坏, 分别通过袋外估计得分 (oobscore)、平均绝对误差得分 (MAEscore) 两种指标, 用来计算不同决策树惩罚权重。

2) 权重设计

根据训练过程中每棵决策树的 oobscore 以及每颗决策树的预测结果的 MAEscore 来对决策树的预测能力进行评估。第 i 棵决策树的权重计算公式如式 (5)、(6) 所示。

$$\omega_i = \frac{\alpha e^{\beta * oobscore_i}}{\sum_{i=1}^N \alpha e^{\beta * oobscore_i}} \quad (5)$$

$$\omega_i = \frac{\alpha e^{-\beta * MAEscore_i}}{\sum_{i=1}^N \alpha e^{-\beta * MAEscore_i}} \quad (6)$$

式中: N 为决策树个数, α, β 为调参因子, $oobscore_i$ 为第 i 棵树的袋外估计得分, $MAEscore_i$ 为第 i 棵树的平均绝对误差得分。式 (5)、(6) 表明, 决策树在袋外数据集上的得分越高则在预测中贡献越大, 当决策树的平均绝对误差得分越小, 表示该树的表现越好, 其在预测的贡献越大, 此时应该赋予该颗决策树更大的权重。

2.3 基于样本特征重要度的权重设计

1) 指标选择

根据每棵决策树在建树时随机选择特征集合的不同, 来为每个决策树赋予相应的权重因子。考虑不同特征集之间的关系, 选择皮尔逊相关系数 (pearson)、互信息 (mul)、卡方统计量 (chi) 以及随机森林特征重要度相关系数 (function) 4 个指标, 用来计算不同决策树的权重因子。互信息是衡量随机变量间相互依赖性的

信息度量, 已知随机变量 X 和 Y , 互信息定义如式 (7) 所示。

$$\begin{aligned} I(X; Y) &= H(X) - H(X|Y) = \\ &= H(X) + H(Y) - H(X, Y) = \\ &= \sum_x p(x) \log \frac{1}{p(x)} + \sum_y p(y) \log \frac{1}{p(y)} - \\ &= \sum_{x,y} p(x,y) \log \frac{1}{p(x,y)} = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \end{aligned} \quad (7)$$

式中: $H(X)$ 、 $H(Y)$ 分别为 X 和 Y 的边缘信息熵; $H(X, Y)$ 为 X 和 Y 的联合熵; $H(X|Y)$ 为条件熵; $p(x, y)$ 为随机变量 X 和 Y 的联合概率分布; $p(x)p(y)$ 分别为随机变量 X 和 Y 的边缘概率分布。

2) 权重设计

现有特征集 $F = \{f_1, f_2, \dots, f_m\}$, 通过上述 4 种指标计算得到的特征重要度集合为 $R = \{r_1, r_2, \dots, r_m\}$, 随机抽取的特征子集为 $K = \{f_1, f_2, \dots, f_k\}$ 对应的特征重要度集合为 $RP = \{r_1, r_2, \dots, r_k\}$, 创建的第 i 棵决策树为 T_i , N 棵决策树构成的随机森林集合为 $\Theta = \{T_1, T_2, T_3, \dots, T_N\}$ 对于第 i 棵决策树对应的权重设计公式如式 (8) 所示。

$$\omega_i = \lambda \sum_{i=1}^k \eta \cdot |r_i| \quad (8)$$

式中: $\lambda \eta$ 为调参因子; k 为单棵决策树时的特征子集维度; r_i 为 4 种指标计算出第 i 个特征的相关系数。

3 基于模糊推理的三轴决策融合算法

考虑单轴反演结果具有一定的片面性, 本文提出基于模糊推理的三轴决策融合算法, 将三轴信号经加权随机森林反演初步得到单轴信号的长、宽、深, 然后通过设计的模糊推理系统做更深层次的融合, 提高反演精度。算法整体步骤分为: 模糊化、规则库及推理机的建立以及去模糊化, 其算法框架如图 1 中部分 4) 所示。

3.1 模糊化

模糊化是将精确的变量用模糊变量表示的过程, 该过程可以分为两部分: 1) 模糊语义变量的定义; 2) 隶属度函数的选取。

1) 模糊语义变量的定义

模糊推理系统的输入变量有 3 个, 分别为轴向反演尺寸、径向反演尺寸、周向反演尺寸, 对于长度、宽度、深度进行决策层融合时, 应分别建立相应的模糊推理系统。输入输出的语义变量划分如表 1 所示。

2) 隶属度函数选取

隶属度函数定义了精确的输入变量、输出变量映射到模糊语言变量的隶属值。由于管道缺陷信号的分布近似正态分布, 故本文采用高斯隶属函数法。如图 3 所示为深度决策融合时隶属度函数示意图。

表 1 模糊语义变量划分

Table 1 Classification of fuzzy semantic variables

属性	语义变量
轴向尺寸	极低、低、中、高
径向尺寸	极低、低、中、高
周向尺寸	极低、低、中、高
融合尺寸	极低、低、中、高、极高

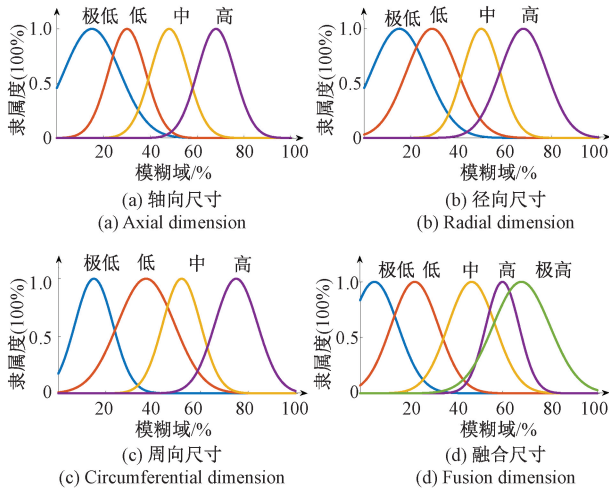


图 3 决策融合隶属度函数

Fig. 3 Decision fusion membership function

3.2 规则库及推理机

1) 规则库

依据专家经验分别对三轴缺陷长度、宽度、深度分别建立相应的模糊规则矩阵。具体构建方法如下:在不同检测条件下多次获取试验场百余个已知缺陷尺寸的三轴漏磁信号;再分别用前文所述方法反演出缺陷的尺寸信息;最后结合已知尺寸,利用最小二乘拟合出误差最小的模糊规则库。以深度决策融合为例,其建立模糊规则库对应的规则矩阵 R 表示如式(9)所示。

$$R = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 4 & 3 & 4 & 5 \\ 4 & 4 & 3 & 5 \\ 4 & 4 & 4 & 5 \end{bmatrix} \quad (9)$$

2) 推理机

本文采用的模糊推理方法为 Mamdani 推理法,推理过程可以分为两个阶段:首先对具有多个前件的单一规则进行单独推理,然后将所有规则的结果进行汇总,进而实现三轴反演信号的决策融合。

3.3 去模糊化

将推理得到的三轴决策反演结果利用重心法实现去模糊化过程,从而得到最终的三轴融合后缺陷反演结果,重心法的计算公式如式(10)所示。

$$\tilde{D}^* = \frac{\int_w w\mu(w) dw}{\int_w \mu(w) dw} \quad (10)$$

式中: \tilde{D}^* 表示系统的清晰化输出变量, $w \in W, \mu(w)$ 表示输出的模糊集合对应的隶属度函数。

4 实验结果与分析

4.1 试验场环境描述

本实验的数据来自如图 4 所示的中海油天津管线环路实验平台。该管线回路总长 800 多米,总计 67 根管道,管道内径 200 mm,壁厚 10 mm。管道材料为 X65 碳钢,介质流速为 0.5~3 m/s,操作压力为 3 MPa。本文数据均是通过如图 5 所示漏磁内检测器采集而来。上述实验场的缺陷长度为 9.7~60 mm,缺陷宽度为 9.5~39.6 mm,缺陷深度为 0.335~9.3 mm。



图 4 管道环路实验平台

Fig. 4 The pipeline loop experiment platform

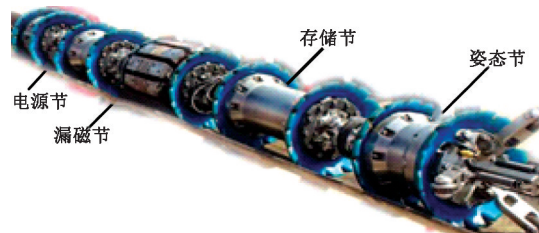


图 5 内检测器

Fig. 5 The internal detector

4.2 实验数据集

本文采用的数据集来自于含有规则缺陷和非规则缺陷的两种牵拉管道。由于数据采集使用的是三轴霍尔传感器,可以采集到 3 个不同方向(轴向、径向、周向)上的漏磁数据,故两种牵拉管道的 3 个方向数据共构成了 6 个互不相同的数据集(DS1~DS6),其中每个数据集的

数据情况如表2所示,其中训练集与测试集比例为8:2,且训练集和测试集无相同或相似样本。

表2 数据集分布及相关特征

陷型	数据类型	数据集	样本数	特征数
规则缺陷	轴向	DS1	176	36
	径向	DS2	176	37
	周向	DS3	176	39
非规则缺陷	轴向	DS4	596	36
	径向	DS5	596	37
	周向	DS6	596	39

4.3 评价指标

本实验采用决定系数(R-square)、平均绝对误差(MAE)、均方根误差(RMSE)以及综合排名(RANK)作为缺陷反演评价指标,各项指标精度均为10折交叉验证的方式得到,评价指标对应公式分别如式(11)~(14)所示。

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (11)$$

$$MAE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (12)$$

$$RMSE(y, \hat{y}) = \sqrt{\frac{1}{n} \sum_{i=1}^n \|y_i - \hat{y}_i\|_2^2} \quad (13)$$

$$RANK = \frac{1}{k} \sum_{i=1}^k S_i \quad (14)$$

式中:其中 n 为样本个数; \bar{y} 为所有样本标签值的平均值; k 为加权随机森林类别数本文为6; S_i 为第 i 种加权方式的排名。

4.4 试验结果分析

本文主要从3方面展开实验分析,首先,利用规则缺陷数据集,对比传统随机森林算法与加权随机森林算法的精度,验证提出改进随机森林算法的有效性;其次,通过模糊融合前后实验分析,验证本文所提出的基于模糊推理决策融合的漏磁数据缺陷反演方法的有效性;最后,通过本文提出的基于三轴融合的漏磁数据缺陷反演方法与其他反演方法在异域数据集中进行实验分析,验证本文所提算法的有效性与泛化性能。

1) 加权随机森林试验结果分析

(1) 有效性分析

利用规则缺陷的三轴漏磁数据 DS1、DS2、DS3 3 种样本集构建随机森林模型,此时通过设置不同的随机种

子(random state)以改变随机森林的构建规则,利用上述6种加权方法与传统随机森林算法分别进行缺陷的长度、宽度、深度反演,各算法对应的长、宽、深预测指标分别如图6~8所示。

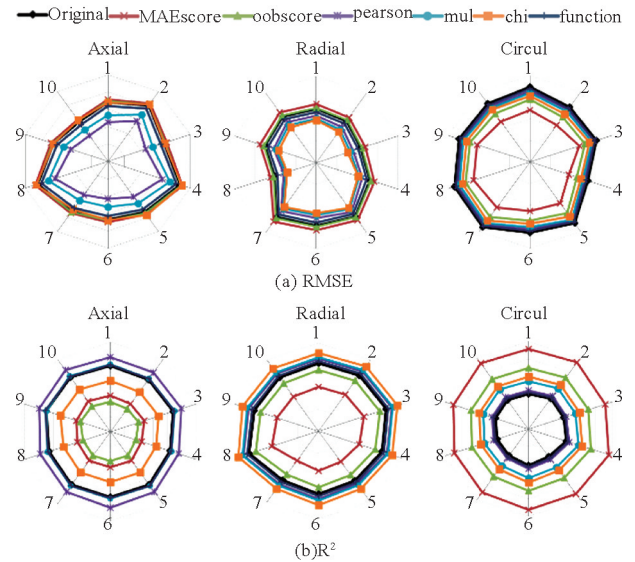


图6 长度预测评估指标对比

Fig. 6 Comparison of length prediction evaluation indicators

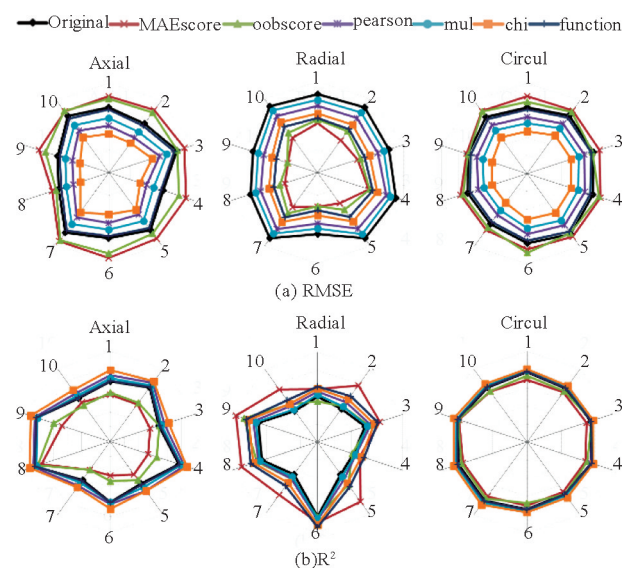


图7 宽度预测评估指标对比

Fig. 7 Comparison of width prediction evaluation indicators

由图8可知,在不同随机种子下进行缺陷尺寸反演时,基于加权随机森林的缺陷反演精度均高于未改进的随机森林算法,证明该算法具有较好的有效性。

(2) 泛化能力验证

为了验证该加权随机森林算法的泛化性能,本试验通过两组不同的数据集进行测试。首先利用非规则缺陷

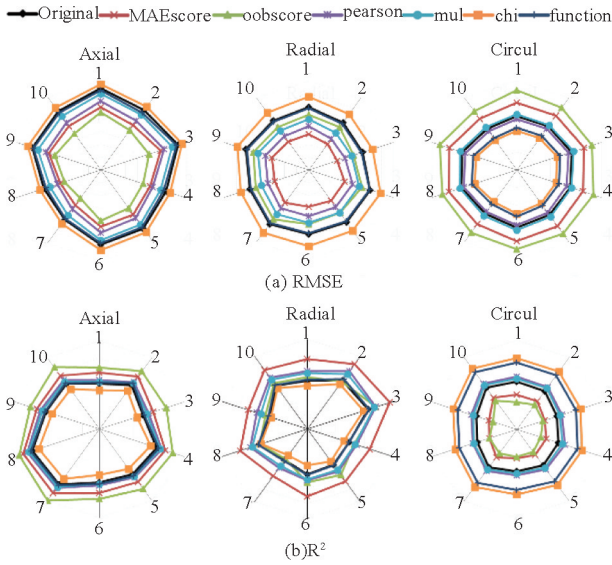


图 8 深度预测评估指标对比

Fig. 8 Comparison of depth prediction evaluation indicators

的三轴漏磁数据即 DS4、DS5、DS6 的训练样本集,建立加权随机森林模型。然后在 DS1~DS6 样本的测试集中进行测试并统计反演结果。各类加权方法综合指标排名如图 9 所示。

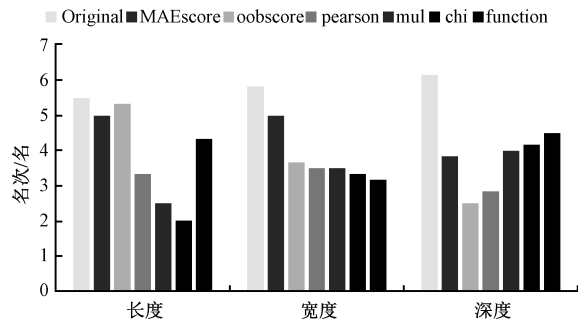


图 9 各类加权方法综合指标对比

Fig. 9 Comparison of comprehensive indicators of various weighting methods

由图 9 可知,在 DS1~DS6 测试集上,对所有算法进行综合排名时,对于缺陷的长度、宽度、深度的估计,本文提出的 6 种加权方式的排名均高于传统随机森林算法。缺陷的长度和宽度会直接反映在漏磁信号的显性特征中,如信号峰谷间距、谷谷间距等。因此,在缺陷的长宽量化时,基于样本特征重要度的 chi 及 function 加权方式能够取得较高的准确度。而缺陷深度对漏磁信号的影响往往耦合了多重信息,难以在信号特征中直接体现。在这种情况下,基于决策树预测能力的 oobscore 方法利用袋外样本来提高模型的泛化能力,在一定程度上解决了缺陷深度对漏磁信号影响的耦合问题,因此可以取得较好的结果。

2) 模糊推理决策融合缺陷反演方法实验分析

由于缺陷的长度、宽度、深度的决策结果分布规律不同,则将其进行融合时,应建立 3 种不同的模糊推理系统(长度、宽度、深度决策融合建立的模糊推理系统分别记为模糊推理系统 L,模糊推理系统 W,模糊推理系统 D)。将规则缺陷的三轴信号对应反演长、宽、深精度最高的加权随机森林的输出作为模糊推理系统 L,模糊推理系统 W,模糊推理系统 D 的输入,实现漏磁信号三轴决策融合,进而得到精确的反演尺寸。

(1) 融合前后缺陷反演方法对比

为分析模糊推理决策融合缺陷反演方法的有效性,将本文所提算法精度与未融合(利用单轴信号实现缺陷反演)时反演精度作对比,其实验结果如图 10 所示。

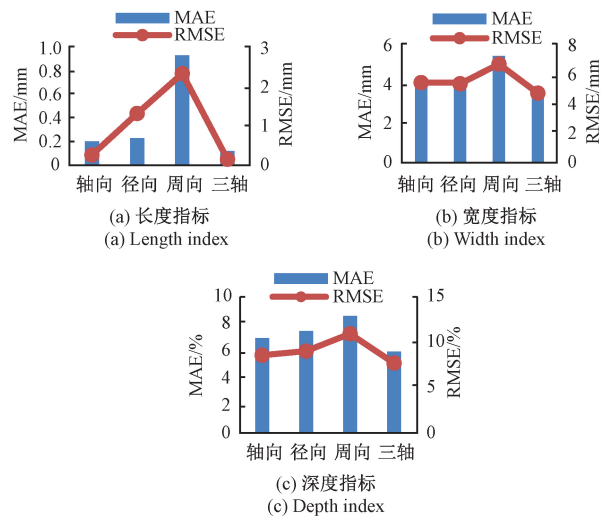


图 10 决策融合前后指标对比

Fig. 10 Comparison of indicators before and after decision fusion

由图 10 可知,融合后反演误差均小于单独利用三轴信号进行独立反演时的误差,证明本文所提算法具有良好的有效性。

(2) 不同决策融合方法对比

为进一步分析本文所提模糊推理决策融合算法的有效性,采用经典的决策融合算法:基于 Bayes 估计的决策融合算法、基于改进的 D-S 证据理论的决策融合算法,与本文所提融合算法进行精度对比,其实验结果如图 11 所示。

由图 11 可知,基于模糊推理的决策融合算法的平方误差以及均方根误差均小于其他经典融合,验证本文所提算法具备良好的有效性。

3) 本文算法整体有效性分析

为验证本文所提方法的有效性,以其他管道非规则缺陷作为测试数据集,将本文所提方法与其他基于机器学习的缺陷反演方法进行对比,实验结果如表 3 所示。

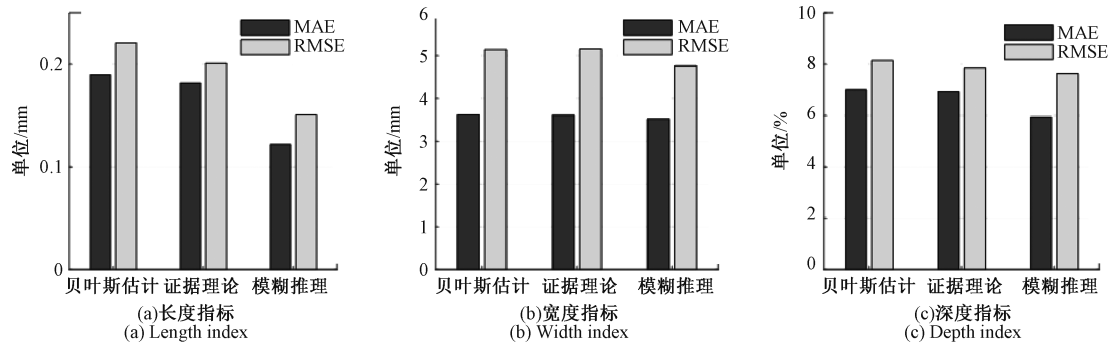


图 11 融合算法指标对比图

Fig. 11 Comparison of fusion algorithm indicators

表 3 不同算法 MAE 对比

Table 3 MAE comparison of different algorithms

测量指标	SVM	随机森林	本文算法
长度/mm	7.782	6.870	5.290
宽度/mm	10.259	8.981	7.788
深度/%	9.30	7.60	6.48

由表 3 可知,在异域数据集中,本文所提方法缺陷反演误差更小,相对于其他方法仍然能保持良好的性能,验证本文所提基于三轴融合的缺陷反演算法有较好的有效性。

5 结 论

针对漏磁数据反演算法精度低的问题,提出一种基于三轴融合的漏磁内检测数据反演算法。首先,构建完备数据集,实现特征提取;其次,通过分析影响随机森林模型泛化误差的因素,引入权重因子,提出基于加权随机森林的缺陷反演算法;最后,设计模糊推理系统实现三轴信号决策融合,得到最终的反演尺寸。实验表明,本文算法相比其他机器学习算法精度明显提升,其中长度精度提升 23%,宽度精度提升 13%,深度精度提升 14.7%。因此本文所提方法具备良好的有效性。

参考文献

- [1] 刘金海,付明芮,唐建华. 基于漏磁内检测的缺陷识别方法[J]. 仪器仪表学报, 2016, 37(11): 2572-2581.
- LIU J H, FU M R, TANG J H. Defect recognition method based on magnetic flux leakage internal inspection[J]. Chinese Journal of Scientific Instrument, 2016, 37(11):2572-2581.
- [2] 杨理践,梁成壮,高松巍,等. 管道漏磁内检测的管壁缺陷漏磁场解析模型[J]. 电子测量与仪器学报,

2021,35(7):106-114.

YANG L J, LIANG CH ZH, GAO S W, et al. Analytical model of pipe wall defect leakage magnetic field for pipeline magnetic flux leakage internal inspection [J]. Journal of Electronic Measurement and Instrumentation, 2021,35(7):106-114.

- [3] 刘金海,赵贺,神祥凯,等. 基于漏磁内检测的自监督缺陷检测方法[J]. 仪器仪表学报, 2020, 41(9): 180-187.
- LIU J H, ZHAO H, SHEN X K, et al. Self-supervised defect detection method based on magnetic flux leakage internal detection [J]. Chinese Journal of Scientific Instrument, 2020, 41(9):180-187.
- [4] 闫蓓,刘孟轲,马宏林. 超声检测系统设计及回波信号特征分析[J]. 电子测量技术, 2019, 42(2):123-129.
- YAN B, LIU M K, MA H L. Ultrasonic testing system design and echo signal characteristic analysis [J]. Electronic Measurement Technology, 2019, 42(2): 123-129.
- [5] 孟繁悦,韩赞东,石承昊. 多通道在线涡流检测系统研制[J]. 电子测量与仪器学报, 2020, 34(5):51-57.
- MENG F Y, HAN Z D, SHI CH H. Development of multi-channel online eddy current detection system [J]. Journal of Electronic Measurement and Instrumentation, 2020,34(5):51-57.
- [6] 王宏安,陈国明. 基于深度学习的漏磁检测缺陷识别方法[J]. 石油机械, 2020, 48(5):127-132.
- WANG H AN, CHEN G M. Flaw recognition method of magnetic flux leakage detection based on deep learning[J]. Petroleum Machinery, 2020, 48(5): 127-132.
- [7] 曹辉,杨理践,刘俊甫,等. 基于数据融合的小波变换漏磁异常边缘检测[J]. 仪器仪表学报, 2019, 40(12): 71-79.

- CAO H, YANG L J, LIU J F, et al. Wavelet transform magnetic flux leakage anomaly edge detection based on data fusion[J]. Chinese Journal of Scientific Instrument, 2019, 40(12):71-79.
- [8] 章卫文, 辛佳兴, 陈金忠, 等. 长输油气管道金属损失漏磁内检测技术研究[J]. 管道技术与设备, 2020(2):25-28.
- ZHANG W W, XIN J X, CHEN J ZH, et al. Research on internal magnetic flux leakage detection technology for metal loss in long-distance oil and gas pipelines [J]. Pipeline Technology and Equipment, 2020(2):25-28.
- [9] YU G, LIU J, ZHANG H, et al. An iterative stacking method for pipeline defect inversion with complex MFL signals[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(6):3780-3788.
- [10] KANDROODI M R, ARAABI B N, BASSIRI M M, et al. Estimation of depth and length of defects from magnetic flux leakage measurements; Verification with simulations, experiments, and pigging data[J]. IEEE Transactions on Magnetics, 2017, 53(3):1-10.
- [11] LI F, FENG J, ZHANG H, et al. Quick reconstruction of arbitrary pipeline defect profiles from MFL measurements employing modified harmony search algorithm[J]. IEEE Transactions on Instrumentation and Measurement, 2018;2200-2213.
- [12] WU Z, DENG Y, LIU J, et al. A reinforcement learning-based reconstruction method for complex defect profiles in MFL inspection [J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70:1-10.
- [13] FU M R, LIU J H, ZHANG H G, et al. Multi-sensor fusion for magnetic flux leakage defect characterization under information incompleteness [J]. IEEE Transactions on Industrial Electronics, 2020, 68(5):4382-4392.
- [14] 林鹏飞, 陶继忠. 基于多样性特征和多源信息的分子泵故障诊断[J]. 真空科学与技术学报, 2020, 40(1):33-39.
- LIN P F, TAO J ZH. Molecular pump fault diagnosis based on diversity characteristics and multi-source information [J]. Journal of Vacuum Science and Technology, 2020, 40(1):33-39.
- [15] 段礼祥, 李涛, 唐瑜, 等. 基于多源异构信息融合的机械故障诊断方法[J]. 石油机械, 2021, 49(2):60-67, 80.
- DUAN L X, LI T, TANG Y, et al. Mechanical fault diagnosis method based on multi-source heterogeneous information fusion [J]. Petroleum Machinery, 2021, 49(2):60-67, 80.
- [16] 刘金海, 臧东, 汪刚. 基于 Markov 特征的油气管道泄漏检测与定位方法[J]. 仪器仪表学报, 2017, 38(4):944-951.
- LIU J H, ZANG D, WANG G. Leak detection and location method for oil and gas pipelines based on Markov feature [J]. Chinese Journal of Scientific Instrument, 2017, 38(4):944-951.

作者简介



卢森骧, 2018 年于东北大学获得博士学位, 现为东北大学讲师, 硕士生导师。主要研究方向是工业人工智能, 故障诊断。

E-mail: lusenxiang@ise.neu.edu.cn

Lu Senxiang received his Ph. D. degree from Northeastern University in 2018. He is currently a lecturer and a master advisor at Northeastern University. His main research interests include industrial artificial intelligence and fault diagnosis.



刘金海 (通信作者), 2009 年于东北大学获得博士学位, 现为东北大学教授, 博士生导师。主要研究方向是工业人工智能, 数据驱动故障诊断理论与技术。

E-mail: jh_lau@126.com

Liu Jinhai (Corresponding author) received his Ph. D. degree from Northeastern University in 2009. He is currently a professor at Northeastern University. His main research interests include industrial artificial intelligence, data-driven fault diagnosis theory and technology.