

时序预测与因果推理融合的水泵寿命预测

姚国伟

杭州沃德水泵制造有限公司 浙江杭州 311100

摘要：水泵作为工业和民用领域的关键流体输送设备，其寿命预测对于降低运维成本、保障系统稳定运行至关重要。传统的水泵寿命预测方法多依赖单一的时序预测技术，难以深入挖掘设备性能退化背后的因果关系。本文提出将时序预测与因果推理相融合的水泵寿命预测方法，详细阐述融合原理、技术实现流程，通过实际案例验证其有效性。研究表明，该方法能够更精准地捕捉水泵性能变化趋势，识别影响寿命的关键因素，为水泵的预防性维护和全生命周期管理提供科学依据，对提升水泵运行可靠性和经济效益具有重要意义。

关键词：水泵；寿命预测；时序预测

引言

在工业生产、城市供水、农业灌溉等众多领域，水泵的稳定运行直接关系到生产生活的正常开展。随着设备长期运行，水泵部件会逐渐磨损、老化，性能不断退化，若不能及时预测其寿命并采取维护措施，可能导致设备突发故障，引发生产中断、水资源供应不足等严重后果，造成巨大的经济损失和社会影响。传统的水泵寿命预测方法，如基于经验公式、定期检测等方式，存在预测精度低、主观性强、无法适应复杂工况变化等问题。近年来，基于数据驱动的时序预测方法，如神经网络、时间序列分析等，凭借对历史数据的学习能力，在水泵寿命预测中取得了一定进展，但这些方法大多仅关注数据的时间序列特征，忽视了设备性能退化与运行环境、操作条件等因素之间的因果联系，难以从根本上解释设备性能变化的原因，限制了预测的准确性和可靠性。因果推理技术能够挖掘变量之间的因果关系，揭示事件发生的内在机制。将时序预测与因果推理相融合，既可以利用时序预测方法对水泵性能数据的趋势进行分析，又能借助因果推理明确影响水泵寿命的关键因素及其作用机制，为水泵寿命预测提供更全面、深入的分析视角。因此，开展时序预测与因果推理融合的水泵寿命预测研究，对于提高水泵寿命预测精度、实现设备的智能化运维具有重要的理论价值和现实意义。

一、时序预测与因果推理融合的理论基础

（一）时序预测技术概述

时序预测是基于时间序列数据的变化规律，对未来趋势进行预测的技术。常见的时序预测方法包括传统时

间序列分析方法和基于机器学习的方法。传统时间序列分析方法，如自回归滑动平均模型（ARMA）、季节性自回归积分滑动平均模型（SARIMA）等，通过对时间序列数据的平稳性检验、参数估计等步骤，建立数学模型来描述数据的变化规律。这类方法具有原理清晰、计算效率高的优点，但对数据的要求较为严格，在处理非线性、非平稳的复杂数据时，预测精度往往较低。

基于机器学习的时序预测方法，如人工神经网络（ANN）、长短时记忆网络（LSTM）、门控循环单元（GRU）等，具有强大的非线性拟合能力，能够自动学习数据中的复杂特征和模式。以LSTM为例，其独特的门控机制使其能够有效处理长序列数据中的长期依赖问题，在时间序列预测领域得到广泛应用。此外，近年来兴起的Transformer模型，凭借其自注意力机制，在时序预测任务中也展现出优异的性能，能够更好地捕捉序列中的全局信息。

（二）因果推理技术概述

因果推理旨在从数据中发现变量之间的因果关系，揭示事件发生的原因和结果。常见的因果推理方法包括基于因果图的方法和基于因果效应估计的方法。基于因果图的方法，如贝叶斯网络、结构因果模型（SCM）等，通过构建因果图来表示变量之间的因果关系，利用图论和概率推理的方法进行因果分析。结构因果模型可以清晰地描述变量之间的因果机制，通过干预分析和反事实推理，能够回答“如果……会怎样”的问题，为决策提供有力支持。

基于因果效应估计的方法，如倾向得分匹配（PSM）、双重机器学习等，主要用于估计干预变量对结

果变量的因果效应。倾向得分匹配通过对干预组和对照组的样本进行匹配，减少混杂因素的影响，从而更准确地估计因果效应；双重机器学习则结合了机器学习算法和因果推断方法，能够在高维数据环境下有效估计因果效应。

（三）融合的必要性及优势

在水泵寿命预测中，单纯的时序预测方法虽然能够对性能数据的变化趋势进行预测，但无法解释为什么会发生这样的变化，难以识别影响水泵寿命的关键因素。而因果推理方法侧重于挖掘因果关系，但在预测未来趋势方面存在不足。将两者融合，时序预测可以为因果推理提供丰富的时间序列数据，帮助发现变量之间随时间变化的因果关系；因果推理则可以为时序预测提供理论依据，明确哪些因素对水泵寿命有重要影响，从而指导时序预测模型的构建和优化^[1]。

通过融合，能够更全面地理解水泵性能退化过程，不仅可以预测水泵性能的未来趋势，还能解释性能变化的原因，提高预测的准确性和可解释性。同时，基于因果关系的分析，还可以为水泵的预防性维护提供更有针对性的建议，如调整运行参数、优化维护策略等，实现设备的精准运维。

二、时序预测与因果推理融合的水泵寿命预测方法

（一）数据采集与预处理

水泵寿命预测需要采集多源数据，包括水泵自身的运行参数，如转速、流量、扬程、电流、电压等，以及运行环境参数，如温度、湿度、水质指标等。这些数据可以通过传感器实时采集，并存储在数据库中。由于采集到的数据可能存在噪声、缺失值等问题，需要进行预处理。对于噪声数据，可采用滤波算法，如中值滤波、卡尔曼滤波等进行去噪处理；对于缺失值，可根据数据特点采用插值法、多重填补法等进行填充^[2]。经过预处理后，将数据划分为训练集和测试集，用于模型的训练和评估。

（二）因果关系分析

利用因果推理技术对预处理后的数据进行因果关系分析。首先，基于领域知识和经验，构建初步的因果图，确定可能影响水泵寿命的因素及其潜在的因果关系^[3]。然后，运用结构因果模型或其他因果推理方法，对因果图进行验证和优化。通过干预分析，模拟改变某些因素对水泵寿命的影响，识别出对水泵寿命有显著影响的关键因素。

（三）时序预测模型构建

根据因果关系分析的结果，选择合适的时序预测方

法构建模型。如果关键因素与水泵性能数据之间存在线性关系，可考虑使用传统时间序列分析方法；若关系呈现非线性特征，则采用基于机器学习的方法，如LSTM、Transformer等。在构建模型时，将因果关系分析得到的关键因素作为模型的输入变量，同时考虑时间序列数据的自相关性和趋势性，优化模型结构和参数。例如，将流量波动、转速变化等关键因素与水泵的历史性能数据一起输入LSTM模型，训练模型学习这些因素与水泵寿命之间的映射关系^[4]。

（四）融合模型训练与优化

将因果推理得到的因果关系知识融入时序预测模型的训练过程。在训练过程中，不仅让模型学习数据的时间序列特征，还引导模型关注关键因素对水泵寿命的因果影响。可以通过设计特殊的损失函数或正则化项，将因果约束引入模型训练，使模型在预测过程中遵循已发现的因果关系。同时，采用交叉验证等方法对融合模型进行优化，调整模型参数，提高模型的泛化能力和预测精度。

（五）寿命预测与结果评估

利用训练好的融合模型对水泵的未来性能进行预测，并根据预设的性能阈值判断水泵的剩余寿命。常见的性能阈值可以根据行业标准、设备制造商的建议或历史经验数据确定。例如，当水泵的扬程下降到初始值的一定比例时，认为水泵达到寿命终点。预测完成后，采用均方误差（MSE）、平均绝对误差（MAE）、决定系数（ R^2 ）等指标对预测结果进行评估，分析模型的预测准确性和可靠性。

三、案例分析

（一）案例背景与数据采集

选取某工业供水系统中的一组离心泵作为研究对象，该组水泵已运行多年，存在不同程度的性能退化问题。在水泵运行过程中，通过安装各类传感器，实时采集水泵的转速、流量、扬程、电流、电压、轴承温度、进水温度、水质pH值等数据，采集频率为每分钟一次，连续采集6个月，共获得约260万条数据记录。

（二）因果关系分析结果

运用结构因果模型对采集的数据进行因果关系分析。经过分析发现，流量波动、轴承温度和水质pH值是影响水泵寿命的关键因素。其中，流量频繁大幅波动会增加叶轮与泵体之间的磨损，加速性能退化；轴承温度过高会导致润滑效果下降，加剧轴承磨损；水质pH值偏离正常范围，会腐蚀水泵部件，缩短设备寿命。

(三) 时序预测模型构建与融合

根据因果关系分析结果,选择LSTM模型构建时序预测模型。将流量波动幅度、轴承温度、水质pH值以及水泵的历史扬程数据作为模型输入,输出未来一段时间内的扬程预测值。在模型训练过程中,引入因果约束,将流量波动、轴承温度和水质pH值对扬程的因果影响纳入损失函数计算,训练得到融合模型。

(四) 预测结果与分析

使用训练好的融合模型对水泵未来30天的扬程进行预测,并与实际测量值进行对比。结果显示,融合模型的预测均方误差为0.08,平均绝对误差为0.05,决定系数 R^2 达到0.92,预测精度明显高于单独使用LSTM模型(均方误差为0.12,平均绝对误差为0.07, R^2 为0.85)。通过融合模型的预测结果,准确判断出其中一台水泵将在未来15天内达到寿命阈值,为运维人员提前安排维护计划提供了可靠依据,避免了设备突发故障造成的损失。

四、面临的挑战与发展趋势

(一) 面临的挑战

在水泵寿命预测领域,时序预测与因果推理融合技术在实际应用中面临着多重挑战。数据层面,高质量数据是精准预测的根基,但现实中,老旧水泵因传感器配置不完善,难以全面采集运行参数,且采集到的数据易受噪声干扰,存在缺失值问题,严重影响因果推理和时序预测的准确性;此外,多源数据融合时,不同类型数据在时间尺度、精度上的差异,使得预处理和对齐操作复杂繁琐。因果关系挖掘方面,水泵运行环境复杂多变,影响其寿命的因素繁多且彼此关联,因果关系错综复杂,在推理过程中,未被察觉的混杂因素极易导致因果关系误判,而且部分因果关系还会随时间与工况发生动态变化,进一步加大了因果分析难度。模型应用层面,基于机器学习的时序预测模型和复杂的因果推理模型复杂度高,其决策过程晦涩难懂,极大降低了预测结果的可信度,难以满足实际应用需求;同时,融合模型训练和计算对硬件设备性能和算法效率要求极高,不仅耗时久,还消耗大量计算资源,限制了该技术的推广与应用。

(二) 发展趋势

未来,水泵寿命预测技术将围绕多源数据、算法优化和部署模式三个维度实现突破发展。在数据层面,依托传感器技术与物联网的深度融合,实现对水泵运行数据的全面感知,获取涵盖设备参数、环境指标等多维度、

高精度的数据信息,同时引入增强学习算法,使模型能够在与动态运行环境的持续交互中自主学习、优化,精准适应复杂工况变化,提升寿命预测的准确性与适应性;在算法层面,借助人工智能技术的创新发展,推动因果发现向自动化、智能化迈进,开发高效的因果推理算法,从海量数据中自动挖掘复杂因果关系,降低对人工经验的依赖,并结合知识图谱技术,对因果关系进行结构化整合与深度推理,为寿命预测构建坚实的知识基础;在部署模式层面,针对模型计算成本高的痛点,着力研发轻量化的时序预测与因果推理模型,降低模型复杂度与资源消耗,同时融合边缘计算技术,将模型部署于水泵现场边缘设备,实现数据的实时处理与预测,有效减少传输延迟,显著提升系统响应速度与可靠性,全方位推动水泵寿命预测技术的智能化升级。

结论

本文提出的时序预测与因果推理融合的水泵寿命预测方法,通过结合两种技术的优势,为水泵寿命预测提供了一种新的思路和方法。理论分析和案例研究表明,该方法能够更准确地预测水泵寿命,识别影响寿命的关键因素,为水泵的预防性维护和全生命周期管理提供科学依据。尽管目前该方法在实际应用中仍面临数据质量、因果关系复杂性、模型可解释性等挑战,但随着相关技术的不断发展,如多源数据融合、增强学习、因果发现自动化等,时序预测与因果推理融合的水泵寿命预测技术将具有更广阔的应用前景,有望推动水泵运维向智能化、精准化方向发展,为工业生产和社会生活提供更可靠的保障。

参考文献

- [1]皮俊东,肖志怀,任刚,等.水轮机顶盖排水泵动态剩余寿命综合预测方法[J].中国农村水利水电,2023,(07):256-261+267.
- [2]徐仁义.核电厂循环水泵滚柱轴承剩余寿命预测方法研究[D].哈尔滨工程大学,2023.
- [3]贾超红,陈惠英,田慕琴,等.主排水泵寿命预测研究[J].煤矿机械,2022,43(08):196-198.
- [4]李炜,李宗仁,毛海杰.基于反馈控制系统实时寿命预测的延寿策略研究[J].兰州理工大学学报,2021,47(06):74-83.