

DOI: 10.19650/j.cnki.cjsi.J2514659

基于迁移学习的数控螺旋锥齿轮磨齿机 主轴热误差建模*

薛 芮¹, 韩 江¹, 田晓青¹, 夏 链¹, 王志永²

(1. 合肥工业大学机械工程学院 合肥 230009; 2. 中南林业科技大学机械与智能制造学院 长沙 410004)

摘 要:针对数控 (CNC) 螺旋锥齿轮磨齿机主轴在多工况下热误差建模精度不足、泛化能力较弱以及传统深度学习模型参数依赖经验调节的问题,提出一种结合鲸鱼优化算法 (WOA) 与迁移学习的长短期记忆网络 (LSTM) 主轴热误差建模方法。首先,搭建主轴热误差实验平台,在 1 000、1 500 和 2 000 r/min 这 3 种转速条件下同步采集 10 个温度测点数据与主轴轴向热误差数据,获得主轴从冷启动到热平衡阶段的完整热行为样本。其次,采用 K-means 聚类与灰色关联度分析相结合的方法对温度变量进行筛选,确定 T1、T5 和 T8 为热误差敏感测点,在降低输入冗余和多重共线性影响的同时保留主要热特征信息。进一步地,引入 WOA 对 LSTM 模型的时间步长、学习率、批大小等超参数进行全局优化,建立 WOA-LSTM 热误差预测模型,以提高模型收敛速度、训练稳定性和预测精度。在此基础上,构建基于模型微调的迁移学习策略,利用 1 000 和 1 500 r/min 工况数据完成源域预训练,再迁移至 2 000 r/min 目标工况进行小样本微调,实现跨工况知识复用与快速适应。实验结果表明:所提出的 WOA-LSTM 迁移学习模型在主轴热误差预测中的均方根误差 (RMSE) 和平均绝对误差 (MAE) 分别为 1.364×10^{-3} 和 1.361×10^{-3} mm,决定系数 (R^2) 达到 0.983 9,模型预测性能优于反向传播网络 (BP)、WOA-BP 及常规 LSTM 模型。该方法在复杂热环境下展现出较好的自适应能力和跨工况泛化能力,为机床热误差建模与补偿提供了一种可推广的新思路,具有工程应用价值。

关键词: 螺旋锥齿轮; 主轴热误差; 误差建模; 迁移学习; 长短期记忆网络

中图分类号: TH161 **文献标识码:** A **国家标准学科分类代码:** 460.4

Thermal error modeling of the spindle of CNC spiral bevel gear grinding machine based on transfer learning

Xue Rui¹, Han Jiang¹, Tian Xiaoqing¹, Xia Lian¹, Wang Zhiyong²

(1. School of Mechanical Engineering, Hefei University of Technology, Hefei 230009, China; 2. College of Mechanical and Intelligent Manufacturing, Central South University of Forestry and Technology, Changsha 410004, China)

Abstract: To address the problems of insufficient modeling accuracy, weak generalization under multiple operating conditions, and the heavy reliance of traditional deep learning models on empirical parameter tuning in spindle thermal error prediction for computer numerical control (CNC) spiral bevel gear grinding machines, a spindle thermal error modeling method based on whale optimization algorithm (WOA), transfer learning, and a long short-term memory network (LSTM) network is proposed. First, a spindle thermal error experimental platform is established, and temperature data from 10 measuring points together with the axial thermal error of the spindle are synchronously collected under three rotational speeds of 1 000, 1 500, and 2 000 r/min, thereby obtaining complete thermal behavior samples from cold start to thermal equilibrium. Second, K-means clustering combined with grey relational analysis is employed to screen the temperature variables, and T1, T5, and T8 are identified as thermal error-sensitive measuring points, which preserves the main thermal feature information while reducing input redundancy and multicollinearity. Furthermore, WOA is introduced to globally optimize hyperparameters of the LSTM model, including the time step, learning rate, and batch size, and a WOA-LSTM thermal error prediction model is established to improve convergence speed, training stability, and prediction accuracy. On this basis, a model fine-tuning-based transfer learning strategy is constructed. The source-domain pre-training is completed using data under 1 000 and

收稿日期: 2025-11-07 Received Date: 2025-11-07

* 基金项目: 国家自然科学基金企业创新联合基金重点项目 (U22B2084) 资助

1 500 r/min conditions, and the model is then transferred to the 2 000 r/min target condition for small-sample fine-tuning, thus enabling cross-condition knowledge reuse and rapid adaptation. Experimental results show that the proposed WOA-LSTM transfer learning model achieves a root mean square error (RMSE) of 1.364×10^{-3} mm, an mean absolute error (MAE) of 1.361×10^{-3} mm, and an R-square (R^2) of 0.983 9 in spindle thermal error prediction, outperforming the back propagation network (BP), WOA-BP, and conventional LSTM models. The proposed method exhibits good adaptability and cross-condition generalization ability under complex thermal environments, and provides a new and practical approach for thermal error modeling and compensation of machine tools.

Keywords: spiral bevel gears; spindle thermal error; error modeling; transfer learning; long short-term memory network

0 引 言

螺旋锥齿轮广泛应用于航空航天、汽车、工程机械、轨道交通等现代机械的传动系统中^[1-2]。随着技术的不断发展,对螺旋锥齿轮的制造精度要求也越来越高。数控(computer numerical control, CNC)机床的热误差问题已成为制约加工精度提升的关键因素之一。主轴系统在高速运转过程中,由于轴承摩擦、生热分布不均及冷却条件变化,会导致机床结构产生热变形,从而引起刀具与工件之间的相对位移,进而影响零件的加工精度与表面质量。众多资料表明,精密加工中热误差的占到综合加工误差的40%~70%^[3-5],如何建立一种高精度、可泛化、可迁移的主轴热误差建模方法,已成为机床精度控制与补偿领域的重要研究方向。

近年来,随着智能制造与数据驱动建模技术的快速发展,利用深度学习对热误差进行建模与预测,已逐渐成为解决该问题的有效途径^[6-7]。然而,传统深度学习模型在不同转速、不同环境下的泛化能力有限,模型参数调整依赖经验,难以在多工况条件下实现准确的热误差预测。因此,亟需引入迁移学习与智能优化算法,构建具备自适应与跨工况学习能力的主轴热误差建模框架,以提升机床在复杂热环境下的精度稳定性。

针对机床热误差的建模与补偿问题,国内外众多学者开展大量工作,总体来看,可以分为机理建模^[8-9]、数据驱动建模^[10-11]以及混合建模方法^[12-13]。早期研究多基于热传导与结构变形机理,利用有限元或有限差分方法建立主轴及关键部件的温度场与热变形模型^[14-15]。该类方法能够较好地反映热-结构耦合机理,但模型建立过程复杂,参数获取困难,且对环境边界条件极为敏感,难以实现在线预测。随着传感技术与计算能力的发展,研究者开始采用多元回归、支持向量回归、反向传播网络(back propagation network, BP)等数据驱动方法对主轴热误差进行辨识^[16-17],该类模型具有建模速度快、实时性强的优点,但难以捕捉热误差的非线性与时序相关特性。近年来,深度学习在时间序列预测中的优越性逐渐凸显,循环神经网络、长短期记忆网络(long short-term memory network, LSTM)模型被广泛应用于机床热误差预测,显

著提升了建模精度和鲁棒性^[18]。然而,这些模型普遍存在“数据依赖性强、工况泛化能力弱”的问题,当加工转速、冷却方式或环境温度变化时,模型性能急剧下降。为克服上述局限,部分学者尝试引入迁移学习思想,通过知识迁移实现不同工况间的模型复用^[19-20],在一定程度上缓解了小样本建模难题;同时,群智能优化算法(如遗传算法、粒子群算法、鲸鱼优化算法(whale optimization algorithm, WOA)等)被用于深度网络的超参数寻优,以提升预测精度和收敛效率。但目前迁移策略选择、优化算法协同机制等仍存在不足,尚需进一步研究与完善。

尽管现有研究在主轴热误差建模方面取得了诸多进展,但仍存在若干关键问题亟待解决。首先,模型的工况依赖性强。多数深度学习模型需要在单一转速或恒定环境下训练,当工况变化(如转速、负载或冷却条件)时,热源分布与热传递路径随之改变,导致模型失效或预测精度显著下降。其次,模型参数的选择依赖经验。网络结构与超参数(如时间步长、学习率、批大小等)对预测精度影响显著,但传统手动调参或网格搜索方法效率低、易陷入局部最优,难以兼顾精度与泛化性能。然后,迁移学习的策略选择尚不系统。现有文献多侧重直接迁移或微调预训练模型,但缺乏对不同层次冻结策略、学习率衰减机制等的系统研究,迁移效果不稳定。最后,多工况数据的不均衡性与样本不足问题突出。在实际机床运行中,获取全工况温度与误差数据的成本较高,数据分布差异显著,导致模型泛化性能受限。综上所述,构建一种能够在多工况条件下实现知识迁移、参数自适应优化的主轴热误差建模方法,对于提升模型的稳定性、精度与实用性具有重要意义。

针对上述问题,本文提出一种基于迁移学习与WOA的LSTM主轴热误差建模方法。该方法充分融合数据驱动建模与智能优化思想,通过多工况数据的知识迁移与模型参数的全局寻优,实现主轴热误差的高精度预测。具体而言,首先利用低速(如1 000、1 500 r/min)工况下的实验数据预训练LSTM模型,提取机床热行为的时序特征与隐含动态规律;随后,在高速(2 000 r/min)工况下采用迁移学习策略进行微调,实现不同工况间的特征迁移与知识复用,从而显著减少新工况模型的训练样本需求。其次,引入WOA对LSTM的关键超参数(包括

时间步长、学习率、批大小、权重衰减及冻结比例等)进行全局优化,避免人工经验调参的局限性,提高模型收敛速度与稳定性。最后,通过实验验证表明,所提出的 WOA-LSTM 迁移学习模型在主轴热误差预测中较传统 LSTM、BP 网络模型具有更高的拟合精度与更强的跨工况泛化能力,为复杂热环境下机床精度控制提供了一种高效、可扩展的建模新思路。

1 基于 WOA-LSTM 的热误差建模理论

1.1 鲸鱼优化算法

WOA 是由 Seyedali Mirjalili 于 2016 年开发的一种群体智能优化算法,其通过模拟座头鲸捕食的行为,进行搜索代理,解决优化问题^[21]。鲸鱼的捕食行为主要分为 3 类:1)包围猎物;2)发泡网攻击;3)搜索捕食。接下来分别对上述 3 类捕食行为用数学公式进行表示。

1) 包围猎物

假设当前最优个体为猎物位置,其余个体通过更新自身位置不断逼近猎物,形成包围行为,即:

$$\begin{cases} \mathbf{D} = |\mathbf{C} \cdot \mathbf{X}^* - \mathbf{X}(t)| \\ \mathbf{X}(t+1) = \mathbf{X}^* - \mathbf{A} \cdot \mathbf{D} \end{cases} \quad (1)$$

式中: $\mathbf{D}$ 为当前迭代中每个鲸鱼个体的距离; $\mathbf{X}(t+1)$ 鲸鱼在下一时刻的位置; $\mathbf{C}$ 为协同系数向量,用于调整当前位置和当前最优位置之间的差距; $\mathbf{X}(t)$ 为当前迭代次数 t 下每条鲸鱼的位置; $\mathbf{X}^*$ 为当前迭代次数 t 下的全局理想位置; $\mathbf{A}$ 为当前迭代中每条鲸鱼的加速度。

$$\mathbf{A} = 2a \cdot \mathbf{r}_1 - a \cdot \mathbf{I}_{\text{vec}} \quad (2)$$

$$\mathbf{C} = 2 \cdot \mathbf{r}_2 \quad (3)$$

式中: $\mathbf{r}_1, \mathbf{r}_2$ 为 $[0, 1]$ 之间的随机向量; a 为常数,随着迭代次数的增加从 2 线性减小至 0; $\mathbf{I}_{\text{vec}}$ 为单位向量。

2) 发泡网攻击

通过螺旋收缩机制和随机选择策略在局部范围内搜索猎物,平衡开发与探索;

$$\mathbf{X}(t+1) = \begin{cases} \mathbf{X}(t) - \mathbf{A} \cdot \mathbf{D}, & p < 0.5 \\ \mathbf{X}(t) + \mathbf{D} \cdot e^{bl} \cos(2\pi l), & p \geq 0.5 \end{cases} \quad (4)$$

式中: p 为常数,介于 $[0, 1]$ 的随机数; l 为常数,介于 $[0, 1]$ 的任意数; b 为常数,控制着每头鲸鱼向前移动时形成的螺旋线的具体形状, $b = 1$ 时,表示典型的对数螺旋。

3) 搜索猎物

当鲸鱼随机搜索新猎物时,算法通过引入随机权重和距离向量实现全局搜索,防止陷入局部最优。

$$\mathbf{X}(t+1) = \mathbf{X}_{\text{rand}} - \mathbf{A} \cdot |\mathbf{C} \mathbf{X}_{\text{rand}} - \mathbf{X}(t)| \quad (5)$$

式中: $\mathbf{X}_{\text{rand}}$ 为在时间 t 时随机选择的个体位置; $\mathbf{X}(t)$ 当前在时间 t 时个体的位置,当 $A < 1$ 时,该鲸鱼在圈内,

选择螺旋包围搜索;当 $A \geq 1$ 时,该鲸鱼在收缩圈外,应该扩大搜索范围,迫使鲸鱼偏离猎物,并进一步进行全局搜索。

1.2 长短期记忆神经网络

机床主轴的热误差产生于内部热源,随时间累积并与结构耦合的动态过程,其表现形式具有强烈的时序相关性、非线性和迟滞特性。传统静态模型难以刻画温度历史对当前热变形的累积影响。为此,本文引入 LSTM 作为核心建模方法。

LSTM 通过其独特的门控机制(遗忘门、输入门、输出门),能够选择性地记忆和遗忘长期序列信息,有效解决了普通循环神经网络(recurrent neural network, RNN)的梯度消失/爆炸问题^[22],非常适合于捕捉温度序列与热误差序列之间的长期依赖关系^[23]

遗忘门控制是否保留上一时刻的状态信息;输入门决定是否更新当前输入;输出门则负责调节隐藏状态在下一时刻的传递强度。其数学表达式^[24]为:

$$\begin{cases} \mathbf{f}_t = \sigma(\mathbf{W}_f[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f) \\ \mathbf{i}_t = \sigma(\mathbf{W}_i[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_i) \\ \mathbf{o}_t = \sigma(\mathbf{W}_o[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_o) \\ \tilde{\mathbf{C}}_t = \tanh(\mathbf{W}_c[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_c) \\ \mathbf{C}_t = \mathbf{f}_t \odot \mathbf{C}_{t-1} + \mathbf{i}_t \odot \tilde{\mathbf{C}}_t \\ \mathbf{h}_t = \mathbf{o}_t \odot \tanh(\mathbf{C}_t) \end{cases} \quad (6)$$

其中, $\mathbf{x}_t$ 为当前时刻输入特征; $\mathbf{h}_t$ 为隐层输出; $\mathbf{C}_t$ 为记忆单元状态; $\mathbf{W}_f, \mathbf{W}_i, \mathbf{W}_o, \mathbf{W}_c$ 为权重矩阵; $\mathbf{b}_f, \mathbf{b}_i, \mathbf{b}_o, \mathbf{b}_c$ 为偏置向量; $\sigma(\cdot)$ 表示 Sigmoid 函数, $\tanh$ 为双曲正切函数; $\odot$ 为 Hadamard 乘积。

通过门控结构的协同作用, LSTM 能够动态地“遗忘”无关信息、保留关键特征,从而建立温度变化与热误差之间的非线性映射关系。LSTM 模型可在序列数据中提取长期依赖特征,实现主轴热误差的动态预测,为后续引入鲸鱼优化算法与迁移学习策略奠定模型基础。本文将多点温度作为输入特征向量,热误差作为输出目标,通过时间反向传播算法对网络参数进行迭代更新。

2 磨齿机主轴热误差实验

2.1 磨齿机结构

如图 1 所示为某磨齿机三维结构示意图,其磨齿加工过程中产形轮的转动是通过两直线轴带动刀盘做圆弧插补运动实现。

由图 1 可知机床机构机床有 6 个坐标轴,刀盘位于机床左边由 Y 及 Z 轴带动上下及前后移动,刀盘自身的转动为 C 轴;齿轮位于机床右边,由 X 轴带动纵向移动, B 轴带动绕 Y 轴摆动, A 轴带动绕本身轴线转动。 X 、 Y

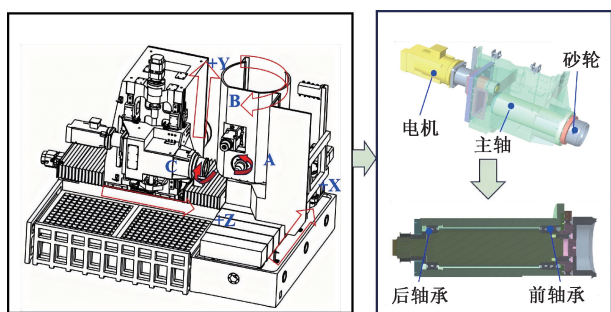


图1 磨齿机三维结构

Fig. 1 Three-dimensional of the gear grinding machine

联动的圆弧插补运动可以视为是虚拟的摇台转动。 Z 轴控制切齿深度, B 轴调整齿轮根锥角。外部电机通过带传动驱动主轴,进而驱动磨齿机砂轮。主轴采用前后轴承进行支承约束,实现稳定回转运动。

2.2 温度传感器布置

针对螺旋锥齿轮磨齿机主轴的热特性分析,需同步获取其温度场与热变形数据。本研究采用 PT100 热电阻传感器(测量范围 $-80^{\circ}\text{C} \sim 200^{\circ}\text{C}$,分辨率 $\pm 0.1^{\circ}\text{C}$)采集主轴系统关键部位的温度信号,并采用激光位移传感器实时测量主轴轴向热误差。所有传感器信号通过同一数据采集系统实现同步采集,以保证温度与热变形数据在时域上的严格对应。

为研究不同转速下主轴的热行为,实验设定 1 000、1 500 和 2 000 r/min 这 3 种工况。每次实验均从冷态开始,主轴以设定转速持续运行 2 h,数据采集间隔为 5 s。该采样时长足以覆盖主轴从启动、温升过渡至准热平衡的完整热过程,确保所获数据包含充分的动态与稳态信息;采样频率则兼顾了热过程的惯性特征与数据有效性,既满足香农采样定理,又避免了数据冗余。

为进一步确保建模与验证的科学性,将每种工况下按时间顺序采集的数据以 7:3 的比例划分为训练集与测试集。该划分方式符合时序建模的验证原则,能客观评估模型在未见数据上的预测性能与泛化能力。

如图 2 所示,为温度传感器的具体布置位置,包括后轴承(T1、T2)、前轴承(T3、T4)、电机(T5、T8)、床身(T6)、立柱(T7)及主轴底座(T9、T10)。该布置方案旨在全面监测主轴系统的温度场分布,为后续热误差建模提供可靠的输入变量。

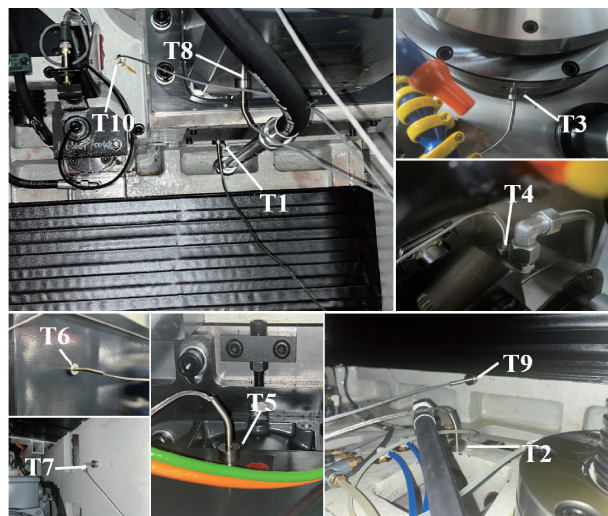


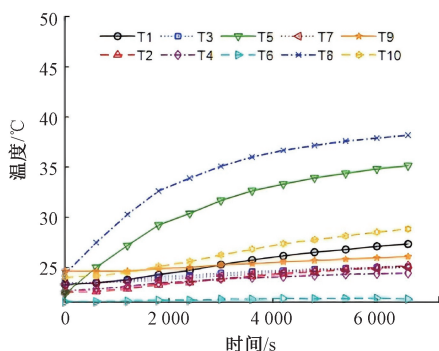
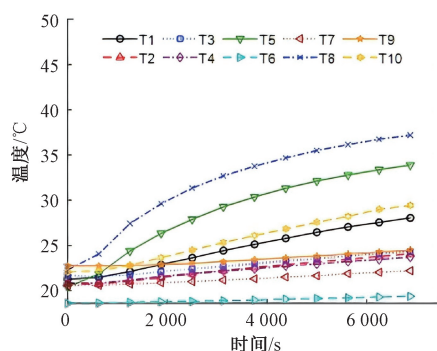
图2 传感器布置

Fig. 2 Sensor location

2.3 实验结果分析

实验过程中采集到的温度和热误差和数据,如图 3 所示。

由图 3 可知,不同转速工况下的温度变化趋势相似,但变化幅度和达到的最高温度不同。特别是在主轴内部轴承处,温度明显升高。初始状态下,各部件温度较低,随着速度的增加,各测点温度上升较快。主轴在预热结束时达到最终稳定转速。在冷却系统作用下达到热平衡。此时,温度波动趋于一致,温升相对较小。前轴承温度明显低于后轴承。因为驱动电机更靠近后轴承,电机产生的热量对后轴承的影响更为显著。由图 3(d)可知,在开始运行工况时,热误差迅速增大,然后达到稳定状

(a) 1 000 r/min 温度曲线
(a) 1 000 r/min temperature curve(b) 1 500 r/min 温度曲线
(b) 1 500 r/min temperature curve

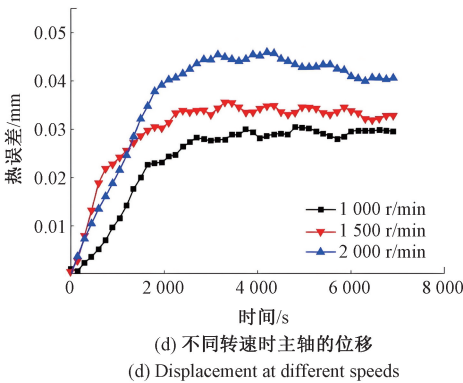
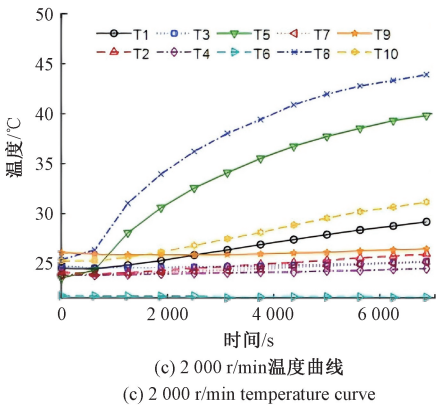


图3 实验过程不同转速时的温度与位移曲线

Fig. 3 Temperature and displacement curves during the experiment at different rotational speeds

态,与温度变化趋势基本一致。这是由于主轴系统内部产生的热量在运行初期对低温部件的冲击作用更强,导致各部件温升较大,从而导致各部件变形较快。在运行后期,系统内部温度逐渐达到热饱和状态,各部分温度波动范围逐渐减小并达到最终稳态,导致热漂移误差增幅减小。随着主轴转速的增加,温度升高,热误差增大。这是因为在较高转速下有更多的产热。主轴转速为 1 000 r/min 时达到热平衡,最高温度为 37.2℃,最大热误差为 33.2×10^{-3} mm。通过对比,主轴在 1 500 和 2 000 r/min 转速下运行时,最大温度分别为 38.3℃ 和 43.0℃,最大热误差分别为 37.5×10^{-3} 和 46.6×10^{-3} mm。

2.4 基于 K-means 的热误差敏感点筛选

1) K-means 聚类算法

过多的温度变量之间存在多重共线性问题,温度点过少可能会导致温度和热误差之间的映射不完整。因此有必要选取合适的温度测点。K-means 聚类算法能够实现样本空间中相近样本的无监督聚类,具有直观、算法时间复杂度低等优点,可以将形态相似的分量进行聚类,减少噪声对整体趋势的影响,降低序列的复杂度,进而提高预测精度,同时也可以减少分量数量,提高模型的训练和预测速度。

如图 4 所示,本文采用 K-means 实现温度变量聚类分簇,通过手肘法确定最佳簇数。由图 4 可以看出,当 $n<3$ 时簇类误差平方和急剧下降,当 $n>3$ 时簇类误差平方和缓慢减小,因此温度测点的最佳聚类数为 $n=3$ 。

温度变量聚类结果,如表 1 所示。10 个温度测点按照相似程度被分为 3 类,再利用灰色关联度分析计算出每个温度变量与热误差之间的关联程度。

2) 灰色关联度分析

聚类分析虽然可以对温度敏感点进行分类处理,但不能具体确定传感器位置和最佳传感器数量。灰色关联

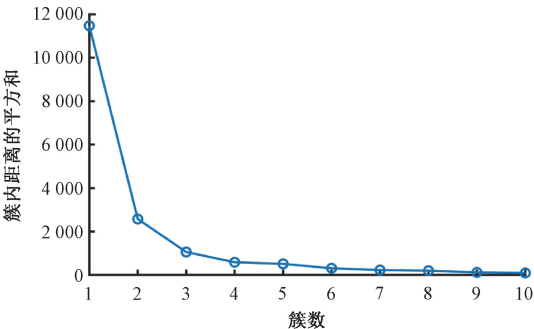


图4 手肘法确定最佳簇数

Fig. 4 Optimal number of clusters by the elbow method

表 1 温度变量聚类结果

Table 1 Temperature variable clustering results

工况	分类	测温点
1 000 r/min	1	T1, T2, T3, T4, T6, T7, T9, T10
	2	T8
	3	T5
1 500 r/min	1	T5, T8,
	2	T2, T3, T4, T6, T7, T9
	3	T1, T10
2 000 r/min	1	T1, T2, T3, T4, T5, T6, T7
	2	T5
	3	T8

度分析(grey relational analysis, GRA)是一种根据系统特征参数序列之间的相关性进行系统分析的数学方法。通过求取灰色关联度,可以去除冗余测量点。GRA 计算如式(7)^[25]所示。

$$\gamma(y, x_i) = \frac{1}{n} \sum_{i=1}^n r(y_k, x_{ik}) \tag{7}$$

式中: y 代表热误差; x_i 代表为第 i 个温度测点; y_k 和 x_{ik} 分别代表热误差和第 i 个温度测点的第 k 个观测值; $\gamma(y, x_i)$ 为热误差和第 i 个温度测点之间的灰色关联度, 由各个观测值的关联度 $\gamma(y, x_{ik})$ 平均而来。 $\gamma(y, x_i)$ 的计算如式(8)^[26] 所示。

$$r(y, x_i) = \frac{\min_i \min_k |y_k - x_{ik}| + \eta \max_i \max_k |y_k - x_{ik}|}{|\min_i \min_k |y_k - x_{ik}| + \eta \max_i \max_k |y_k - x_{ik}|}$$

(8)

式中: η 为分辨率系数。灰色关联度越高, 说明两变量之间的关联度越强。

基于灰色关联度算法, 对主轴热误差与各测温点数据之间的相关性进行了量化分析, 结果如图 5 所示。

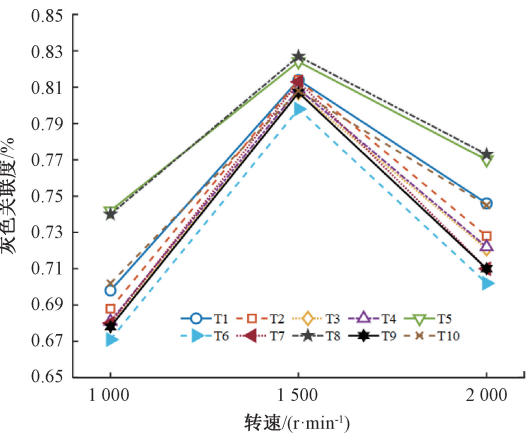


图 5 温度测量点与热误差的灰色关联度分析
Fig. 5 Grey correlation analysis of temperature measurement points and thermal errors

图 5 中关联度值越大, 表明该测点温度与热误差间的相关性越强。由图 5 可见, 位于电机(T5、T8)与后轴承(T1、T2)处的测点关联度普遍高于其他位置。

3) 温度敏感点在不同转速下的变化分析

对 3 种转速工况下的关联度进行排序, 如表 2 所示。

表 2 灰色关联度排序

Table 2 Ranking of grey correlation coefficients

工况转速	筛选结果
1 000 r/min	T5>T8>T10>T1>T2>T4>T7>T3>T9>T6
1 500 r/min	T8>T5>T2>T1>T7>T4>T10>T3>T9>T6
2 000 r/min	T8>T5>T1>T10>T2>T4>T3>T7>T9>T6

上述结果进一步揭示出温度敏感点的动态特性: 电机附近测点 T5 与 T8 在不同转速下始终位列前 2, 表明电机发热是系统最主要且稳定的热源; 后轴承测点 T1 与 T2 的关联度排序随转速升高明显上升(如在 2 000 r/min

时 T1 进入前 3), 说明轴承摩擦热的影响在高速工况下显著增强; 而前轴承及床身结构测点的关联度则普遍较低且相对稳定。这表明, 主轴热误差的温度敏感点集合既具有核心稳定性(电机), 又包含动态变化性(轴承)。

综合 K-means 聚类分析与灰色关联度排序结果, 最终选定 T1、T5、T8 作为模型输入的温度变量。该选择不仅捕捉了最主要的热源影响, 也涵盖了随工况变化的关键动态因素, 为后续采用迁移学习方法提供了依据。

3 基于 WOA-LSTM 的热误差建模

在主轴热误差预测任务中, 温度场与结构变形均呈现显著的时序相关性与非线性耦合特征, 因此本文引入 LSTM 网络模型建立热误差建模, 然而 LSTM 网络性能受到超参数以及滑动窗口大小的设置影响, 为了建立性能准确且理想网络, 必须针对网络训练参数进行深度寻优, 因此引入 WOA 智能优化算法, 构建 WOA-LSTM 模型。

如图 6 所示为 WOA-LSTM 模型的流程图。首先初始化 WOA 算法的相关参数, 并随机生成个体位置以构建初始种群; 随后, 通过计算个体的适应度函数评估当前超参数配置下 LSTM 模型的性能, 并根据 WOA 算法机制迭代更新个体位置, 直至满足最大迭代次数或预设的误差收敛条件。在获得最优超参数组合后, 进入模型训练阶段, 确定 LSTM 网络结构并配置相应的超参数, 利用训练集完成模型训练, 最终通过测试集对模型性能进行评估, 输出预测结果。该流程实现了 WOA 对 LSTM 关键参数的自动寻优, 提升了模型的泛化能力与预测精度。

WOA 的超参数搜索空间及算法设置, 如表 3 所示。本研究选取时间步长、学习率、批大小、冻结比例等对模型性能影响最显著的超参数作为搜索对象, 并结合预实验中对模型稳定性和计算成本的综合评估, 为每个参数设定了合理的取值范围。其中, 时间步长设置为 20~40, 用于覆盖主轴热误差在 100~200 s 尺度上的主要动态特征; 学习率采用 $1 \times 10^{-5} \sim 1 \times 10^{-2}$ 的对数均匀采样方式, 使不同量级具有相同的探索概率; 批大小与微调轮数则在常用范围内选取, 以兼顾泛化性能与训练效率。冻结比例设为 0.70~1.00, 用于控制源域知识的保留程度, 避免目标域样本不足导致的过拟合。此外, WOA 的种群规模和最大迭代次数分别设为 8 与 10, 能够在保证搜索多样性的同时控制优化过程的计算代价。在本研究中, 被测点的热变形与温升之间的关系由式(9)给出, 即:

$$Y_t = f(Y_{t-1}, T_{t-1}, T_{t-2}, \dots, T_{t-n})$$

(9)

式中: Y_t 和 Y_{t-1} 分别表示热误差的 t 时刻以及 $t-1$ 时刻预测值; $T_{t-1}, T_{t-2}, \dots, T_{t-n}$ 表示过去时刻的温度, n 表示时间窗口。

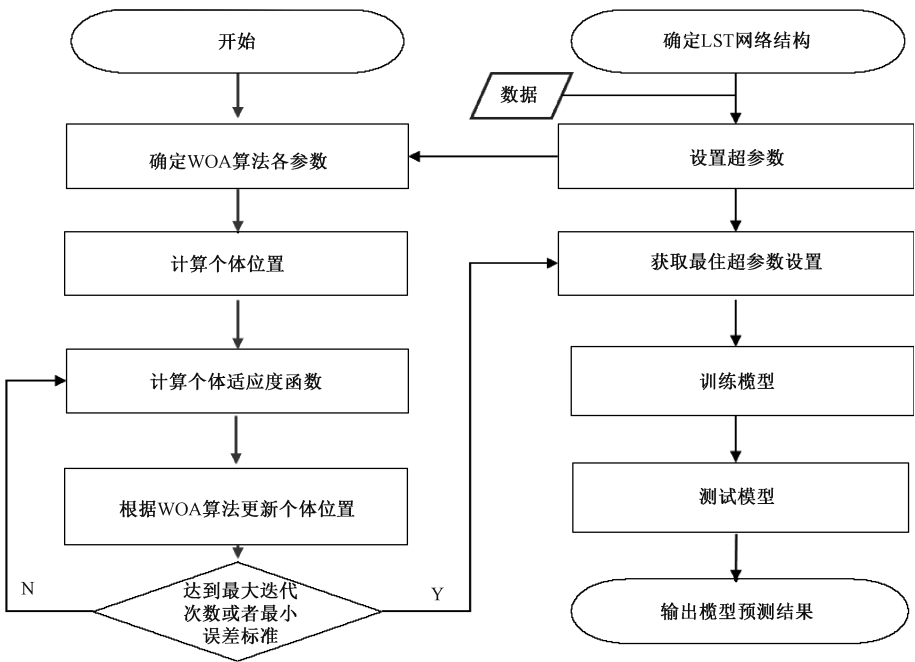


图 6 WOA-LSTM 模型的流程
Fig. 6 Flowchart of the WOA-LSTM model

表 3 WOA 超参数搜索空间与算法设置
Table 3 Search space and algorithm settings of WOA hyperparameters during the transfer learning stage

类别	超参数	取值范围/设置
LSTM 超参数 搜索空间	时间步长	20~40
	微调学习率	$1\times10^{-5}\sim1\times10^{-2}$
	批大小	{16,32,64,128}
	微调轮数	15~80
	冻结比例	0.70~1.00
	权重衰减	$1\times10^{-6}\sim1\times10^{-3}$
WOA 算法设置	种群规模	8
	最大迭代次数	10
	收敛因子 a	2.0~0
	螺旋常数 b	1.0

将温度和热变形数据输入 LSTM 模型中进行训练,然后将 WOA 智能优化算法确定的最优学习率和单元数输入到网络中,以获得模型的最优配置来预测热误差,之后针对预测结果进行进一步性能评价。

3.1 模型性能评价指标

本文依次采用 WOA-LSTM、LSTM、WOA-BP 和 BP 进行预测性能对比。本文采用在 1 000 r/min 数据进行训练,建立热误差模型,确保被测试模型的可比性,同时,以均方根误差 (root mean square error, RMSE)、均方误差

(mean squared error, MSE)、平均绝对误差 (mean absolute error, MAE)、平均绝对百分比误差 (mean absolute percentage error, MAPE) 和决定系数 (R^2) 作为评价模型的指标。RMSE 和 MAE 越小,误差越小,预测性能越高。MAPE 值越小,说明预测值和实际值的偏差越小,模型的预测精度越高。 R^2 越接近 1,表明模型对数据的拟合程度越好。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{10}$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \tag{11}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \tag{12}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \tag{13}$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \tag{14}$$

式中: n 表示样本数; y_i 和 $\hat{y}_i$ 分别表示测量值和预测值。

3.2 建模结果

本文将 1 000 r/min 转速下采集的样本作为训练集训练前述的 4 种模型,并在测试集上验证各个模型性能,4 种模型测试集结果对比,如图 7 所示。

残差的大小表示预测精度的高低。从图 7 可以看出其中,WOA-LSTM 模型的预测曲线与实测曲线基本一致,

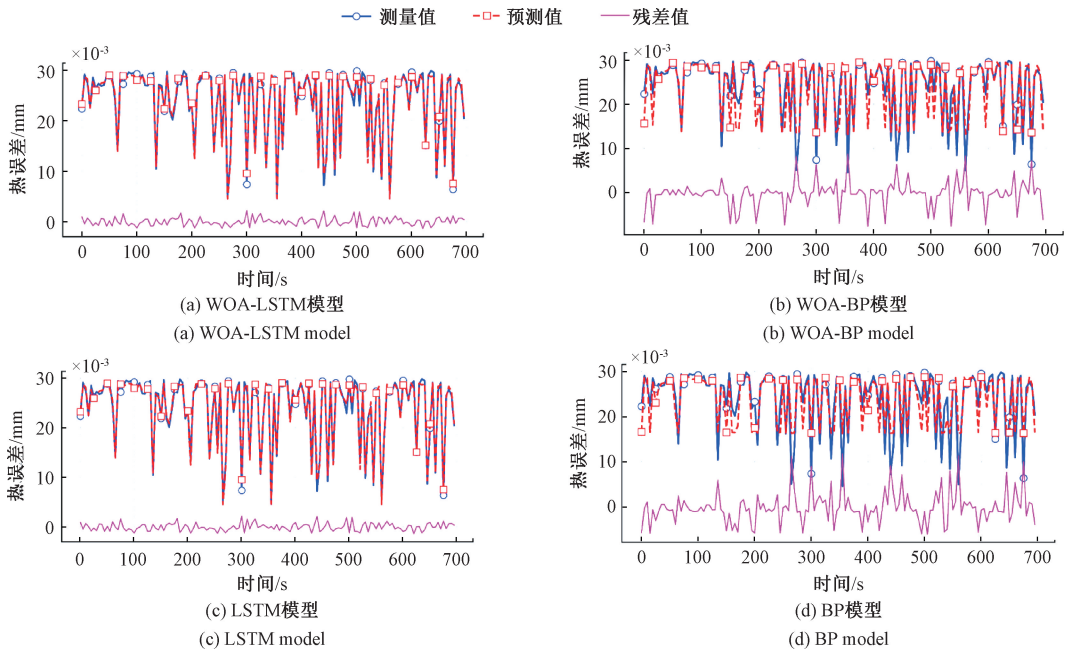


图7 1 000 r/min 时4种模型测试集结果对比

Fig.7 Comparison of test set results of the four models at 1 000 r/min

残余变化大部分在 2×10^{-3} mm 以内,变化较小,预测效果相对更好。而 LSTM 模型的残差波动约为 $\pm3\times10^{-3}$ mm。相比之下,WOA-BP 与 BP 模型的残差峰值分别达到约 $\pm6\times10^{-3}$ 和 $\pm8\times10^{-3}$ mm,误差波动明显增大。虽然 LSTM 和 WOA-BP 的结果对于训练集来说是可以接受的,但是测试集的某些数据误差较大,说明泛化能力不强,无法对训练阶段从未出现过的数据进行较准确预测。相比之下,所提出的 WOA-LSTM 模型能够更准确地预测热误差变化。

WOA-BP 与 WOA-LSTM 超参数优化收敛曲线如

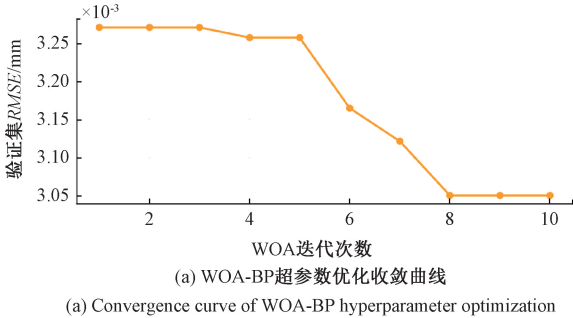


图8所示,为说明超参数搜索的有效性,图8给出了在相同数据划分与训练预算下,WOA-BP 与 WOA-LSTM 两种模型的超参数优化收敛过程:横轴为 WOA 迭代代数,纵轴为迭代到当代为止的最优验证集 RMSE。可以看到,WOA-BP 的验证误差由初始 3.28 随迭代逐步下降,在第8~10代附近收敛至约3.05;WOA-LSTM 在前几代表现为平台期,随后在第9~10代出现显著下降,由3.00快速降至2.41并保持稳定。该结果表明 WOA 能有效发现更优的超参数组合,使模型验证误差单调降低并最终收敛。

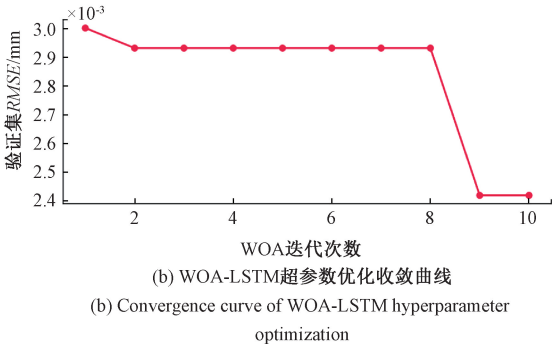


图8 WOA-BP 与 WOA-LSTM 超参数优化收敛曲线

Fig.8 Convergence curves of hyperparameter optimization for WOA-BP and WOA-LSTM

4种模型测试集的统计评价指标,如表4所示。结合表4的结果可知,WOA-LSTM模型的RMSE仅为 0.77×10^{-3} mm、MAE为 0.63×10^{-3} mm,显著优于LSTM(1.92×10^{-3} mm)、WOA-BP(3.03×10^{-3} mm)及BP($3.57\times$

10^{-3} mm)模型。这表明 WOA-LSTM 模型不仅在整体预测精度上最优,而且在瞬态响应中表现出最小的残差抖动与最高的稳定性。

在4种不同的模型中,WOA-LSTM 的误差低于常规

表 4 4 种模型测试集的统计评价指标
Table 4 Comparison of test set performance of four models

模型	$RMSE \times 10^{-3} / \text{mm}$	$MSE \times 10^{-6} / \text{mm}^2$	$MAE \times 10^{-3} / \text{mm}$	$MAPE / \%$	R^2
WOA-LSTM	0.77	0.59	0.63	3.36	0.987 1
LSTM	1.92	3.69	1.24	7.99	0.921 8
WOA-BP	3.03	9.18	1.80	13.50	0.803 6
BP	3.57	12.74	2.43	19.00	0.728 0

BP 和 WOA-BP,说明该模型具有较高的精度。

3.3 基于迁移学习的模型泛化性能提升

由于不同转速工况下主轴热行为存在显著数据分布移位,直接沿用源工况(1 000、1 500 r/min)训练的模型会在升温早期产生尺度性偏差与时滞,难以满足在线补偿需求。为此,本节引入基于模型的迁移学习框架,基于模型的迁移学习流程,如图 9 所示:首先在 1 000 和 1 500 r/min 工况数据上对 WOA-LSTM 模型开展预训练,使其学习到通用化的热行为特征表示;随后通过迁移学习,将源域模型的共享参数传递至目标任务;最终在目标域(2 000 r/min)上对模型进行微调,仅调整与目标工况强相关的非共享参数,同时结合差分目标提升鲁棒性。

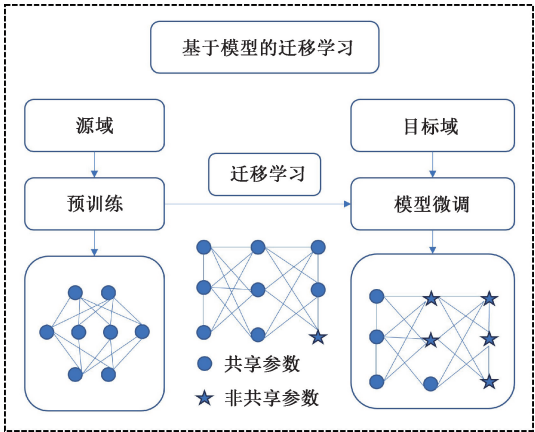
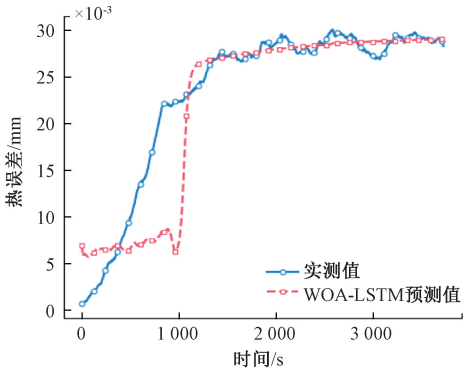


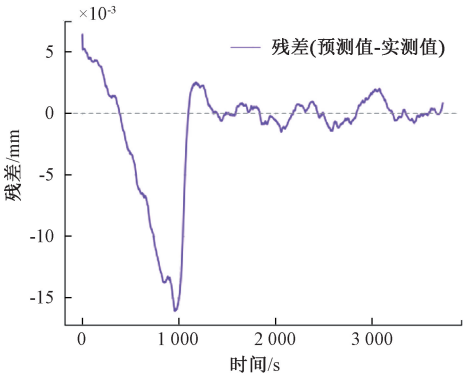
图 9 基于模型的迁移学习流程

Fig. 9 Model-based transfer learning process

如图 10 所示,为主轴 2 000 r/min 时采用迁移学习前的预测结果,图 10(a)为热误差预测结果图,不难看出直接在目标域上使用源域参数的 WOA-LSTM,在升温阶段存在明显的尺度性低估与时间延迟:约 0~15 min 的上升段预测偏低,峰值附近仍呈系统性偏差。这表明模型难以在分布移位(源域到目标域)的早期动态过程中捕捉到真实的热演化轨迹。图 10(b)为预测残差图,从图 10(b)中可以看出,在升温阶段残差波动较大,最大偏差 -15.3×10^{-3} mm,但在稳态阶段残差主要分布于 $\pm 2 \times 10^{-3}$ mm 范围。这说明未经迁移学习的模型在高转速工况下泛化能力有限,难以准确捕捉不同工况下的热特征变化。



(a) 热误差预测结果
(a) Prediction results of thermal errors



(b) 模型预测残差
(b) Model prediction residuals

图 10 迁移学习前的预测结果

Fig. 10 Prediction results before transfer learning

相比之下,迁移学习后预测结果,如图 11 所示。图 11(a)为热误差预测结果,从图 11(a)中可以看出,模型预测值与实测值在整个加热阶段和稳态阶段均高度一致,二者曲线几乎重合,整体趋势吻合良好。预测误差的 RMSE 为 1.364×10^{-3} mm, MAE 为 1.361×10^{-3} mm,决定系数 R^2 达到 0.983 9,说明模型能够较准确地反映主轴热误差的时变规律。显著优于未迁移方案,说明迁移学习有效缓解了由工况差异引起的模型失配,提高了模型的预测精度。图 11(b)为模型预测性能的散点拟合图,可以看到,大部分散点均紧密分布在对角线附近,说明预测结果与实测结果呈现良好的线性相关性,模型总体预

测偏差较小。图 11(c) 给出了预测残差的分布情况。残差点主要集中在 $\pm 0.2 \times 10^{-3}$ mm 范围,且未呈现明显的系统性偏差,表明模型在不同温升阶段均能保持较好的泛化性能与稳定性。

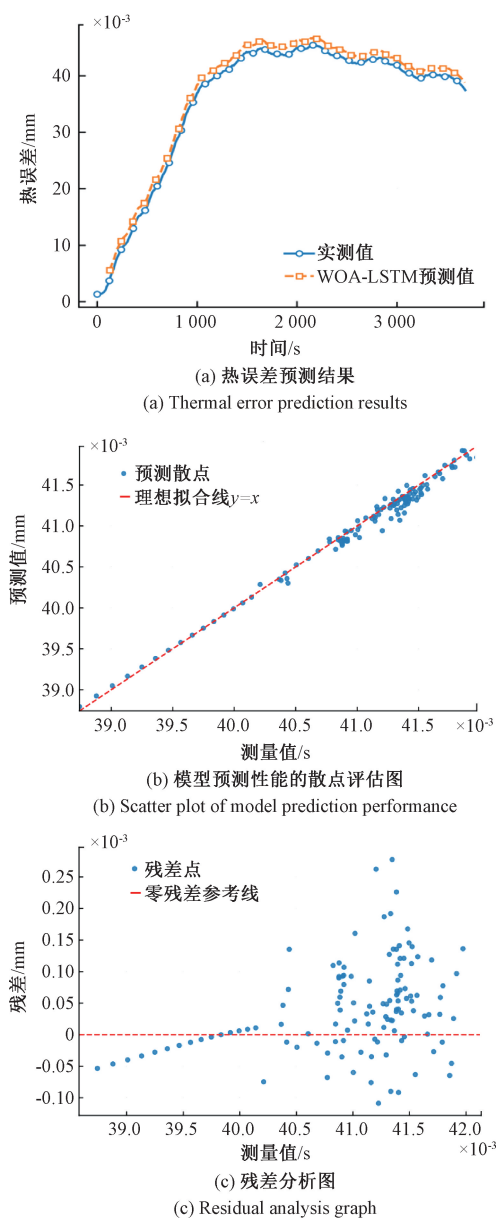


图 11 迁移学习后预测结果

Fig. 11 Prediction results with transfer learning

总体而言,WOA-LSTM 模型在时序预测精度、线性拟合程度及残差分布一致性方面均表现优异。通过引入迁移学习策略,在源域模型的预训练基础上对目标域进行小样本微调,使模型在保持时间序列相关性的同时,能够快速适应 2 000 r/min 工况下的热边界与散热特性变化,从而有效降低了模型在初始阶段的预测误差并缩短了收敛时间。迁移学习后的模型不仅能够为主轴启动后的前

几十分钟内实现对热误差的可靠预测,为机床的实时热误差补偿提供了更稳定的依据,而且在复杂热环境下展现出更强的自适应性与跨工况泛化能力。

本研究未覆盖更极端的负载/环境范围,后续将扩展不同砂轮状态与车间季节性温漂的测试实验,并计划公开更多工况数据集,以进一步推动方法的工程化与普适化。

4 结 论

本文针对数控螺旋锥齿轮磨齿机主轴在多工况下热误差建模精度不足、模型泛化性差的问题,提出了一种结合鲸鱼优化算法与迁移学习策略的长短期记忆神经网络热误差建模方法。通过在低速工况下进行预训练并在高速工况下进行小样本微调,实现了主轴热误差模型的跨工况知识迁移与参数自适应优化。主要结论为:

1) 温度特征优化与输入变量筛选。提出了 K-means 聚类 and GRA 结合的温度测点优化选取温度测量点,去除冗余的温度测量点、避免共线性对模型精度的干扰,同时也降低模型输入温度变量复杂度。

2) 模型优化与迁移策略设计。引入鲸鱼优化算法对 LSTM 模型的时间步长、学习率、批大小及冻结层比例等超参数进行全局寻优,实现了模型的自动化参数调整与快速收敛;通过基于模型的迁移学习策略,实现了低速工况知识向高速工况的有效迁移,显著提升了模型的跨工况泛化性能。

3) 模型性能与对比分析。实验结果表明,所提出的 WOA-LSTM 迁移学习模型在 2 000 r/min 工况下的预测精度达到 $RMSE = 1.364 \times 10^{-3}$ mm、 $MAE = 1.361 \times 10^{-3}$ mm、 $R^2 = 0.9839$,明显优于传统 BP、WOA-BP 及未进行迁移的 LSTM 模型,能够准确捕捉主轴热变形的动态演化过程。

参考文献

- [1] WANG Y N, WANG H, LI K Y, et al. An analytical method to calculate the time-varying mesh stiffness of spiral bevel gears with cracks [J]. Mechanism and Machine Theory, 2023, 188: 105399.
 - [2] ZHANG Y, ZHU L Y, GOU X F. Calculation methods of load distribution ratio for spiral bevel gear [J]. International Journal of Mechanical Sciences, 2023, 257: 108531.
 - [3] 肖捷, 王志永, 于水琴, 等. 基于 RIME-BP 神经网络的磨齿机进给系统热误差预测[J]. 仪器仪表学报, 2024, 45(11): 277-286.
- XIAO J, WANG ZH Y, YU SH Q, et al. Thermal error

- prediction of gear grinding machine feed system based on RIME-BP neural network [J]. Chinese Journal of Scientific Instrument, 2024, 45(11): 277-286.
- [4] 岳玉霞, 王秀山, 李利军. 五轴数控机床的误差实时补偿控制器[J]. 电子测量技术, 2021, 44(13): 145-149.
- YUE Y X, WANG X SH, LI L J. Real time error compensation controller for five-axis CNC machine tools[J]. Electronic Measurement Technology, 2021, 44(13): 145-149.
- [5] KAFTAN P, MAYR J, PORQUEZ F, et al. Reducing thermal errors with confidence: Uncertainty-based compensation for precision machine tools [J]. CIRP Journal of Manufacturing Science and Technology, 2025, 61: 400-409.
- [6] 魏新园, 王杲, 周京欢, 等. 基于迁移学习的异工况下机床热误差建模方法[J]. 仪器仪表学报, 2023, 44(7): 44-52.
- WEI X Y, WANG G, ZHOU J H, et al. Thermal error modeling method for machine tool under different working conditions based on transfer learning[J]. Chinese Journal of Scientific Instrument, 2023, 44(7): 44-52.
- [7] 马帅, 冷杰武, 陈祝云, 等. 基于数字孪生和深度迁移学习的电主轴热误差建模方法[J]. 机械工程学报, 2025, 61(3): 52-66.
- MA SH, LENG J W, CHEN ZH Y, et al. Thermal error modeling method towards electric spindles based on digital twin and deep transfer learning [J]. Journal of Mechanical Engineering, 2025, 61(3): 52-66.
- [8] LI ZH Y, LI G L, XU K, et al. Thermal error analytical modeling of gear grinding machine full-closed-loop feed system based on equivalent temperature field [J]. Mechanical Systems and Signal Processing, 2025, 224: 112105.
- [9] 朱晚宁, 张浩, 陆旭. 大型液体静压回转工作台静态热误差机理与温度特性研究[J]. 制造技术与机床, 2025, 32(12): 211-219.
- ZHU W N, ZHANG H, LU X. Study on the static thermal error mechanism and temperature characteristics of large-scale hydrostatic rotary table[J]. Manufacturing Technology & Machine Tool, 2025, 32(12): 211-219.
- [10] 杜柳青, 吕发良, 余永维. 数字孪生下基于 DACS-MFAC 的数控机床热误差自适应预测方法[J]. 仪器仪表学报, 2024, 45(4): 248-257.
- DU L Q, LYU F L, YU Y W. DACS-MFAC-based adaptive prediction method for thermal errors of CNC machine tools under digital twin[J]. Chinese Journal of Scientific Instrument, 2024, 45(4): 248-257.
- [11] XUE R, HAN J, XIA L, et al. Thermal error modeling and compensation of CNC spiral bevel gear spindles using a GRU-transformer network [J]. The International Journal of Advanced Manufacturing Technology, 2025, 141(3/4): 1917-1933.
- [12] DAI Y, ZHAN SH Q, HUANG Y J, et al. Thermal management strategy for motorized spindle systems: Graphite coating thermal path reconstruction and hybrid predictive modeling[J]. Applied Thermal Engineering, 2025, 279: 128071.
- [13] WANG M, LU W L, ZHANG K, et al. Thermal error prediction of ball screws in full-time series using working condition data based on a mechanism and data hybrid-driven model[J]. The International Journal of Advanced Manufacturing Technology, 2024, 133(3/4): 1443-1462.
- [14] LIU Y, MA Y X, MENG Q Y, et al. Improved thermal resistance network model of motorized spindle system considering temperature variation of cooling system[J]. Advances in Manufacturing, 2018, 6(4): 384-400.
- [15] 刘阔, 宋磊, 陈虎, 等. 机理驱动的数控机床进给轴时变误差建模和补偿方法[J]. 机械工程学报, 2022, 58(3): 251-258.
- LIU K, SONG L, CHEN H, et al. Mechanism-driven method for time-varying error modeling and compensation of CNC machine tool's feed axes [J]. Journal of Mechanical Engineering, 2022, 58(3): 251-258.
- [16] GUO ZH F, LIU S Y, YAN M Y, et al. Thermal error prediction of electric spindle based on improved whale algorithm optimized neural network[J]. The International Journal of Advanced Manufacturing Technology, 2025, 138(9/10): 4739-4753.
- [17] YAN ZH, WANG Y, XIA ZH J, et al. Dynamic physics-weighted Gaussian process regression for robust thermal error prediction under non-stationary conditions[J]. Engineering Applications of Artificial Intelligence, 2026, 171: 114205.
- [18] SUN CH, LIU ZH Y, QIU CH, et al. An axial attention-BiLSTM-based method for predicting the migration of CNC machine tool spindle thermal error under varying working conditions [J]. The International Journal of Advanced Manufacturing Technology, 2024, 130(3/4):

1405-1419.

- [19] SU H, YIN L, QIU CH CH, et al. A novel transfer learning model based on ED-TCN and RSD domain adaptation for thermal error prediction of multiple machine tools [J]. *Ain Shams Engineering Journal*, 2026, 17(2): 103966.
- [20] 郑悦, 付国强, 雷国强, 等. 变工况下基于迁移学习融合域内对齐的机床主轴热误差模型[J]. *仪器仪表学报*, 2023, 44(5): 33-43.
- ZHENG Y, FU G Q, LEI G Q, et al. Thermal error model of machine tool spindle based on transfer learning and domain-aware alignment under variable operating conditions[J]. *Chinese Journal of Scientific Instrument*, 2023, 44(5): 33-43.
- [21] MAKHADMEH S N, KASSAYMEH S, RJOU B G, et al. Recent advances in multi-objective whale optimization algorithm, its versions and applications [J]. *Journal of King Saud University Computer and Information Sciences*, 2025, 37(7): 200.
- [22] AO X Q, GONG Y J, HE AI X. A review of time series prediction models based on deep learning [J]. *IEEE Access*, 2025, 13: 153696-153712.
- [23] ZENG SH, MA CH, LIU J L, et al. Sequence-to-sequence based LSTM network modeling and its application in thermal error control framework [J]. *Applied Soft Computing*, 2023, 138: 110221.
- [24] LI ZH Y, LI G L, XU K, et al. Data-driven thermal error prediction of spindle with mechanism-reinforced temperature information [J]. *Advanced Engineering Informatics*, 2025, 64: 103005.
- [25] ZHENG K, FANG J, LI J Q, et al. Robust grey relational analysis-based accuracy evaluation method[J]. *Applied Sciences*, 2025, 15(9): 4926.
- [26] HU A Q, XIE N M. Construction and application of a novel grey relational analysis model considering factor coupling relationship [J]. *Grey Systems: Theory and Application*, 2025, 15(1): 1-20.

作者简介



薛芮, 2016 年于安徽芜湖市安徽工程大学获得学士学位, 2021 年于江苏徐州市中国矿业大学获得硕士学位, 现为合肥工业大学博士生, 主要研究方向为数控加工技术、运动控制。

E-mail: xuerui315@163.com

Xue Rui received his B.Sc. degree from Anhui Polytechnic University in 2016, and his M.Sc. degree from China University of Mining and Technology in 2021. He is currently a Ph.D. student at Hefei University of Technology. His main research interests include CNC machining technology and motion control.



韩江(通信作者), 1984 年于合肥工业大学获得学士学位, 2006 年于合肥工业大学获得博士学位, 现为合肥工业大学教授, 主要研究方向为智能数控技术、高档数控机床与数控系统、机器人控制与加工。

E-mail: jianghan@hfut.edu.cn

Han Jiang (Corresponding author) received his B.Sc. and Ph.D. degrees both from Hefei University of Technology in 1984 and 2006, respectively. He is currently a professor at Hefei University of Technology. His main research interests include intelligent numerical control technology, high-end CNC machine tools and numerical control systems, robot control and processing.



王志永, 1996 年于沈阳工业学院获得学士学位, 1999 年于东北大学获得硕士学位, 2006 年于中南大学获得博士学位, 现为中南林业科技大学教授, 主要研究方向为复杂齿轮切齿加工理论及其数字化制造技术。

E-mail: wangzy@csuft.edu.cn

Wang Zhiyong received his B.Sc. degree from Shenyang Institute of Technology in 1996, his M.Sc. degree from Northeastern University in 1999, and his Ph.D. from Central South University in 2006. He is currently an associate professor at Central South University of Forestry and Technology. His main research interests include the theory of complex gear cutting and digital manufacturing technologies.