doi:10.12068/j.issn.1005－3026.2026.20259051

基于物理指导神经网络的中厚板轧制力预测

郭 薇[1,2]，董梓硕[1,2]，李 旭[3]，于 孟[1,2]
（1.首钢集团有限公司 技术研究院，北京 100043；2.绿色可循环钢铁流程北京市重点实验室，北京 100043；
3.东北大学 数字钢铁全国重点实验室，辽宁 沈阳 110819）

摘 要：传统机理模型在复杂耦合条件下难以保持轧制力的精确预测，而纯数据驱动的机器学习模型虽能捕获非线性规律，却缺乏物理知识，易导致过拟合，且其泛化性能与可解释性受到限制．提出一种基于物理指导的神经网络（physics-guided neural network，PGNN）模型，通过在损失函数中融入轧制机理，使模型同时受数据特征与物理规律指导．同时，引入鲸鱼优化算法（whale optimization algorithm,WOA）优化模型超参数以提升模型性能．实验结果表明，所提出的PGNN方法在预测精度、数据依赖性和泛化性能方面均优于传统数据驱动方法，且预测结果与轧制理论高度一致．该方法为轧制力的高精度建模提供了新途径，并展示了在智能制造领域的应用潜力．

关 键 词：中厚板；横向变厚度轧制；轧制力预测；物理指导神经网络；鲸鱼优化算法
中图分类号：TG 335.51 **文献标志码**：A **文章编号**：1005-3026(2026)06-0097-11

Prediction of Rolling Force for Medium and Heavy Plates Based on Physics-Guided Neural Network

GUO Wei[1,2], *DONG Zi-shuo*[1,2], *LI Xu*[3], *YU Meng*[1,2]
(1. Technology Research Institute, Shougang Group Co., Ltd., Beijing 100043, China; 2. Beijing Key Laboratory of Green Recyclable Steel Processes, Beijing 100043, China; 3. State Key Laboratory of Digital Steel, Northeastern University, Shenyang 110819, China. Corresponding author: GUO Wei, E-mail: guowei0514@shougang.com.cn)

Abstract: Accurate prediction of rolling force can hardly be maintained by traditional mechanistic models under complex coupled conditions. Meanwhile, although nonlinear rules can be captured by purely data-driven machine learning models, they lack physical knowledge, which easily leads to overfitting and limits their generalization performance and interpretability. A physics-guided neural network (PGNN) model was proposed. By integrating the rolling mechanism into the loss function, the model was guided simultaneously by data features and physical laws. Meanwhile, a whale optimization algorithm (WOA) was introduced to optimize model hyperparameters to improve model performance. Experimental results indicate that the proposed PGNN method is superior to traditional data-driven methods in terms of prediction accuracy, data dependence, and generalization performance, and the prediction results are highly consistent with rolling theory. This method provides a new pathway for high-precision modeling of rolling force and demonstrates its application potential in the field of intelligent manufacturing.

Key words: medium and heavy plate; transverse variable thickness rolling; rolling force prediction; physics-guided neural network; whale optimization algorithm

钢铁工业在国民经济中占据着至关重要的地位，是衡量一个国家工业化水平与综合国力的重要标志之一[1]．中厚板作为钢铁企业的重要产品，广泛应用于汽车、船舶、建筑等领域．近年来，随着工业领域的快速发展，对板材产品的质量与性能提出了更高要求，钢铁生产已从“以产量提

收稿日期：2025-11-01
作者简介：郭 薇(1982—)，女，北京人，首钢集团有限公司技术研究院正高级工程师，博士．
通信作者：郭 薇，E-mail:guowei0514@shougang.com.cn.

升为主”逐步转向“以质量与成材率提升并重”的发展阶段.因此,构建以信息化技术为核心的钢铁智能制造体系,成为行业发展的必然趋势.在实现轧制过程智能控制的过程中,建立高精度的轧制控制模型,尤其是轧制力预测模型,是其中的关键环节[2].轧制力是板带轧制过程控制的核心之一,也是反映轧制状态的重要参数,其设定精度直接影响厚度板形控制精度及成品质量与产量.同时,轧制力也决定了道次压下量的分配,是制定轧制规程与工艺曲线的依据之一[3].因此,建立高精度的轧制力预测模型,对提升轧制控制水平、保障产品质量具有重要意义.

然而,轧制过程机理复杂,对生产环境、轧制工况高度敏感,涉及变量众多,且各变量之间具有强耦合性与显著的非线性特征,这使得轧制力的精确预测极具挑战性[4].当前轧制力建模方法主要可分为基于机理的建模与基于数据的建模两类.基于机理的轧制力预测模型通常基于金属塑性变形的力学分析,其结果具有明确的物理意义[5-6].但由于轧制过程机理复杂、参数众多,分析中常需做出简化与假设,从而导致计算精度下降[7].此外,传统平板轧制理论普遍假定板材横向厚度一致或相近,而对于诸如MAS(mizushima automatic plan view pattern control system)轧制等特殊工艺,由于在特定道次中宽度方向存在显著厚度变化,采用常规平轧模型计算轧制力将产生较大误差[8].因此,在横向变厚轧制场景下,传统理论模型难以满足实际生产精度要求.另一类是基于有限元分析与数值仿真的方法.该类方法通过建立合理的边界条件与网格划分,可较准确地描述变形区应力应变分布.然而其计算耗时较长,难以满足实时在线控制的需求[9].

近年来,随着机器学习技术的快速发展,数据驱动方法在工业领域得到了广泛关注.机器学习可通过大数据挖掘潜在特征,在处理复杂多变量耦合问题时往往优于传统机理模型[10-11].因此,此类方法已被广泛应用于热轧过程建模与控制中[12-13].已有研究利用随机森林(RF)、深度置信网络(DBN)等模型实现了轧制力预测[14-15].然而,这些方法大多针对平板轧制工况,对于横向厚度不均匀板材的轧制力预测研究仍较少.此外,纯数据驱动模型高度依赖数据质量与数量,而工业生产数据的采集往往耗时且成本高[16].同时,实际数据通常存在噪声与缺失,导致训练样本无法全面覆盖生产场景.更重要的是,纯数据驱动模型缺乏对物理机理的理解,无法有效利用长期积累的领域知识,因而容易出现预测结果与物理规律不符的情况,影响其在实际生产中的可信度[17].

机器学习模型的核心目标是有效利用已有信息对数据规律进行逼近.提高模型性能的关键途径之一,是在学习过程中融合先验知识或领域机理信息[18].只有将领域知识与数据特征相结合,才能使模型既具备数据驱动模型的强拟合能力,又具备专家系统的物理合理性.为弥补纯理论模型与数据驱动模型的不足,学者们开始探索机理与数据融合的双驱动模型,并在理论研究中取得了良好效果[19-20].因此,在钢铁轧制领域,构建融合机理与数据的建模范式,是提高工艺控制精度、提升产品质量的有效途径.近年来,为提升轧制过程控制精度,融合机理与数据的“双驱动”建模已成为研究热点,并在轧制力、板形板厚等预测中取得了一些积极成果.例如,在板形方面,Meng等[21]开发了一种融合机理的深度随机配置网络模型,利用数据驱动方法学习实际板形与机理模型计算结果之间的残差,从而有效补偿了理论模型的固有偏差,提升了最终预报精度.在带钢厚度方面,Wen等[22]提出了一种机理与数据分步协调的预测框架,通过性能评估规则库自适应调度模型,并利用可解释人工智能技术对模型可靠性进行了验证.在轧制力预测方面,Zhang等[23]针对特厚板,通过补偿原则将理论模型与优化后的神经网络相结合,构建了一个融合理论模型与数据驱动的集成模型;Cao等[24]通过建立动态速度场机理模型,并将其输出与过程参数一同输入机器学习模型,从而显著提升了预测精度.这些工作从不同角度验证了理论与数据模型融合在处理热轧这类复杂问题上的巨大潜力.然而,上述研究多停留于机理与数据模型的后集成,这种方式难以将轧制理论中变量关系融入模型的学习过程,易导致模型过拟合.同时,针对中厚板横向变厚度这一非稳态、强非线性的轧制力变化过程,相关研究仍显不足,传统机理模型与纯数据驱动方法均难以兼顾预测精度与稳定性.

基于此,本文提出了一种用于横向厚度变化热轧过程的基于物理指导的神经网络(physics-guided neural network,PGNN).该模型在深度神经网络的损失函数中嵌入了轧制过程中的单调性规律与边界条件,以指导模型学习方向,使其

在数据驱动学习的同时遵循物理规律.同时,引入鲸鱼优化算法(whale optimization algorithm,WOA)优化模型超参数以提升模型性能,从而获得兼具高精度与强泛化性能的轧制力预报模型.

1　基于PGNN的横向变厚度轧制力预测框架

1.1　横向变厚度轧制过程

相较于热连轧,中厚板轧制过程中的展宽比更大.为达到目标尺寸,钢板在轧制过程中通常需在某一轧制道次进行90°旋转,先沿宽度方向轧制,再旋转至长度方向轧制[25].同时,为获得良好平面形状,生产中常采用平面形状控制(plan view pattern control,PVPC)轧制.该方法通过在成型或展宽轧制阶段最后一个道次中动态调整辊缝,使钢板沿宽度方向形成一定的厚度分布差异,从而改善头部和边部形状[26-27].因此,在动态压下后转钢的第一个道次,钢板入口厚度沿宽度方向出现较大变化,如图1所示.

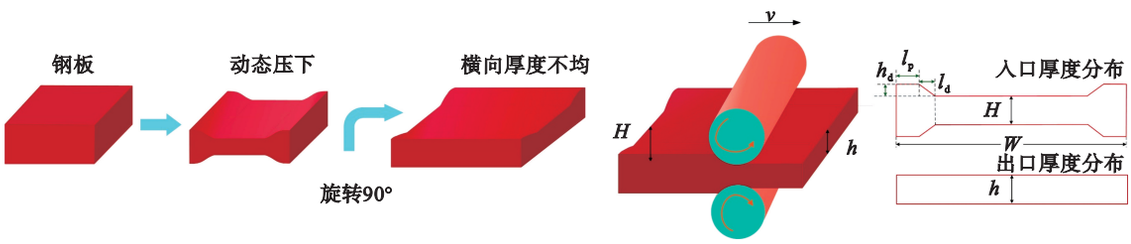


图1　钢板横向变厚度的轧制过程

Fig. 1　Transverse variable thickness rolling process of steel plate

在横向厚度不均的轧制过程中,钢板沿宽度方向的压下量出现变化,这种横向变化可以用3个参数来表示:l_p,l_d和h_d[8].在此情况下,变形区内部形状变得不规则,接触弧投影长度沿宽度出现变化.目前计算横向变厚度轧制力的方法依赖于平板轧制过程的单位轧制力方程.对于横向可变厚度钢板,轧制力P计算式[5]为

$$P_x=\begin{cases}\sigma_s\cdot\left(\dfrac{\pi}{4}\ln\dfrac{h_x}{H}+\dfrac{\pi}{4}+\sqrt{\dfrac{R}{h}}\arctan\left(\sqrt{\dfrac{R}{h}}\alpha\right)-\sqrt{\dfrac{R}{h}}\arctan\left(\sqrt{\dfrac{R}{h}}\varphi\right)\right),\varphi<\gamma;\\ \sigma_s\cdot\left(\dfrac{\pi}{4}\ln\dfrac{h_x}{h}+\dfrac{\pi}{4}+\sqrt{\dfrac{R}{h}}\arctan\left(\sqrt{\dfrac{R}{h}}\varphi\right)\right),\gamma<\varphi<\alpha.\end{cases}\tag{1}$$

$$P=\int_0^W\int_0^{l_w}P_x\mathrm{d}x\mathrm{d}W.\tag{2}$$

式中:σ_s为材料变形抗力,MPa;H和h分别为入口厚度和出口厚度,mm;W表示变形区的平均宽度,mm;h_x表示变形区的厚度,mm;α,φ和γ分别为变形角、接触角和中性角,rad;R为轧辊半径,mm;l_w为接触弧在变形区宽度方向上的投影长度,该长度受下压量沿宽度方向分布影响,如图2所示.

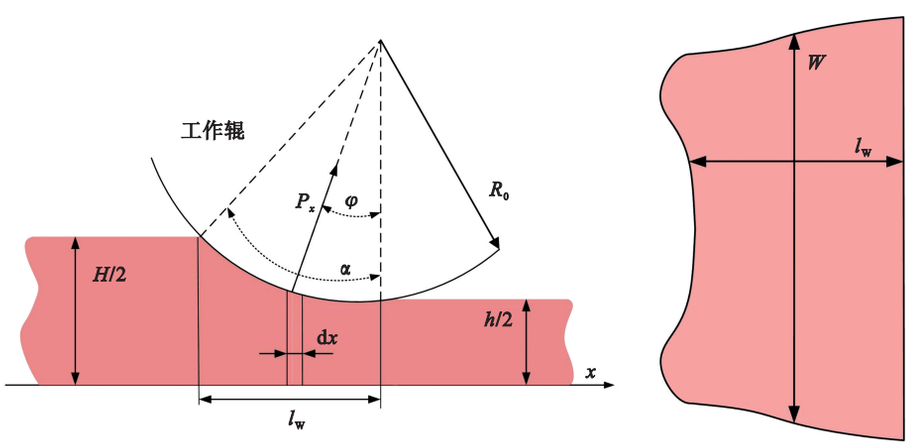


图2　横向变厚度钢板的轧制变形区的示意图

Fig. 2　Schematic diagram of rolling deformation zone for transverse variable thickness steel plate

因此,相对于等厚度轧制,横向变厚度轧制的变形过程所涉及的多变量耦合关系更为复杂.

1.2 基于物理指导的神经网络框架

在各类机器学习方法中,神经网络(neural network,NN)具有优异的非线性映射能力.其基本单元为“神经元”,通过输入节点接收数据,经非线性变换后输出响应.深度神经网络(deep neural network,DNN)由输入层、输出层及若干隐含层组成.数据经输入层前向传播至输出层,通过反向传播(backpropagation,BP)算法不断修正参数,实现模型学习.虽然神经网络具有极强的拟合性能,但其搜索空间庞大,易陷入局部最优,同时对样本数据的数量与质量依赖较强.为克服这些问题,可以通过引入合理的约束条件来指导模型收敛方向,从而提高其学习效率.在轧制机理研究中,已有大量专家知识与经验规律可供利用,这些机理知识可作为数据驱动模型的物理约束条件.基于此思想,本文提出一种基于物理指导的神经网络.该模型在传统DNN的损失函数中引入包含物理知识的惩罚项,以此约束模型的解空间.PGNN的总体结构如图3所示.图中,Loss为训练过程总损失函数;MSE_{DNN}表示模型在训练过程中的数值观测误差;$Loss_{theory}$为损失函数的物理损失部分;$f(x_i;\ \theta)$为模型输出,x_i为输入变量,θ为模型参数;v为泊松比;E为轧辊的弹性模量,MPa.

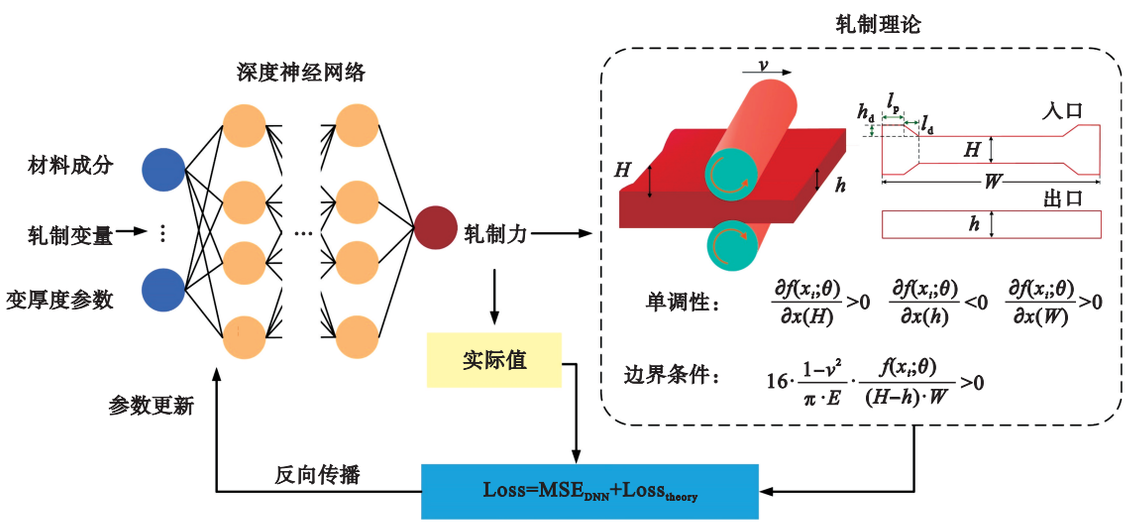


图3 基于PGNN的横向变厚度轧制力预测

Fig. 3 Prediction of transverse variable thickness rolling force based on PGNN

所提出的PGNN是在DNN模型基础上进行拓展,对神经网络的损失函数进行修改,加入基于物理的约束.该约束包括两个主要部分:单调性约束和边界条件约束.

1.2.1 单调性约束

经过研究学者们大量的理论推导和实验,建立了完善的轧制力计算理论体系,其中的基本规律描述了轧制过程中各种变量对轧制力的影响.最直观的是板材入口和出口厚度对轧制力的影响,入口厚度的增加和出口厚度的减小都会使轧制力增大.此外,钢板的入口宽度与轧制力也成正比,虽然宽度的增加不会增加接触弧长度,但会增加变形区面积,因此宽度越大,轧制力越大[28].根据上述描述,轧制过程中影响轧制力的一些单调性规律可表示为式(3).

$$\frac{\partial P}{\partial H}>0;\ \frac{\partial P}{\partial h}<0;\ \frac{\partial P}{\partial W}>0. \tag{3}$$

将式(3)中的单调性规律进行编码并纳入模型的损失函数,作为物理损失的一部分,如式(4)所示.

$$\begin{aligned}\mathrm{Loss_m}=&\lambda_H\cdot\sum_{i=1}^{n}\mathrm{ReLU}\left(-\frac{\partial f(x_i;\theta)}{\partial x(H)}\right)+\\&\lambda_h\cdot\sum_{i=1}^{n}\mathrm{ReLU}\left(\frac{\partial f(x_i;\theta)}{\partial x(h)}\right)+\\&\lambda_W\cdot\sum_{i=1}^{n}\mathrm{ReLU}\left(-\frac{\partial f(x_i;\theta)}{\partial x(W)}\right).\end{aligned} \tag{4}$$

式中:$\mathrm{Loss_m}$为基于机理的单调性损失;$\mathrm{ReLU}(z)=\max(0,z)$为线性整流函数;$n$为输入训练样本的数量;$\lambda_H$,$\lambda_h$和$\lambda_W$分别为理论约束中3个单调性约束的权重.

1.2.2 边界条件约束

在钢板轧制过程中,较大的轧制力会导致轧辊发生弹性变形,这使得轧辊的实际半径比初始半径大,如图4所示.

压扁后的半径可根据Hitchcock公式计算,基于此可以得出模型的边界条件约束,如式(5)所示.

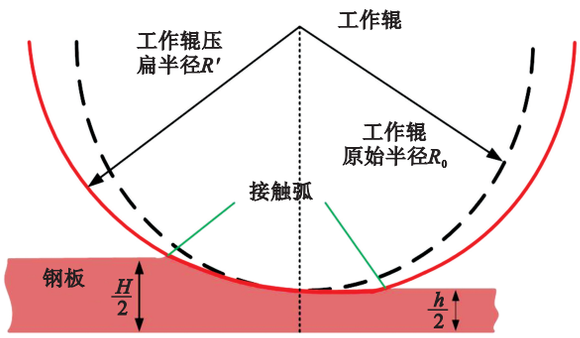


图4　钢板热轧过程的轧辊压扁示意图

Fig. 4　Schematic diagram of roll flattening during hot rolling process of steel plate

$$\left.\begin{aligned} &R'=R_0\cdot\left(1+\frac{C_e\cdot P}{(H-h)\cdot W}\right)>R_0,\\ &C_e=16\cdot\frac{1-\nu^2}{\pi\cdot E}.\end{aligned}\right\}\Rightarrow$$
$$16\cdot\frac{1-\nu^2}{\pi\cdot E}\frac{P}{(H-h)\cdot W}>0. \tag{5}$$

将上述边界条件编码为物理约束的第二部分,如式(6)所示.

$$\mathrm{Loss_b}=\lambda_e\cdot\sum_{i=1}^{n}\mathrm{ReLU}\left(-16\cdot\frac{1-\nu^2}{\pi\cdot E}\frac{f(x_i;\theta)}{(H-h)\cdot W}\right). \tag{6}$$

式中:$\mathrm{Loss_b}$为基于机理的边界条件损失;λ_e为机理约束第二部分的权重.

损失函数量化了模型训练过程中计算值与预期结果之间的偏差程度,影响神经网络的泛化性能和准确性.常规的损失函数主要关注模型输出与实际值之间的差异,忽略了与现有理论认知的一致性.因此,将基于轧制理论编码的式(4)和式(6)纳入损失函数,设计了PGNN的损失函数为

$$\left.\begin{aligned} &\mathrm{Loss_{theory}}=\mathrm{Loss_m}+\mathrm{Loss_b},\\ &\min\mathrm{Loss}=\min\{\mathrm{MSE_{DNN}}+\mathrm{Loss_{theory}}\}.\end{aligned}\right\} \tag{7}$$

训练过程可以看作是最小化损失函数的过程,在这一过程中,物理损失起着约束作用,对不符合物理规律的预测结果进行惩罚,这可以约束模型向与理论认知一致的全局最优解逼近.尽管还存在其他影响因素和边界条件,但在构建PGNN预测模型时,并非所有的理论约束条件都能对提高精度起到积极作用.考虑到过多的约束条件会影响模型的收敛过程,仅将上述影响轧制力的因素作为所提出模型的物理约束.

1.3　模型结构优化

除常规超参数(如隐含层数与节点数)外,PGNN中物理约束项的权重系数同样需要确定.若通过人工试验逐一搜索最优参数组合,计算代价较高且效率低下.为此,本文采用鲸鱼优化算法(WOA)[29]对模型超参数进行自动优化,以获得最优结构.本文利用WOA同时优化网络结构参数(如隐含层数与隐含层节点数)及物理约束权重.其优化流程如图5所示,图中a为线性递减控制参数;$\boldsymbol{A}$为搜索系数向量;C为位置扰动系数;l为螺旋搜索系数;q为[0,1]的随机数.

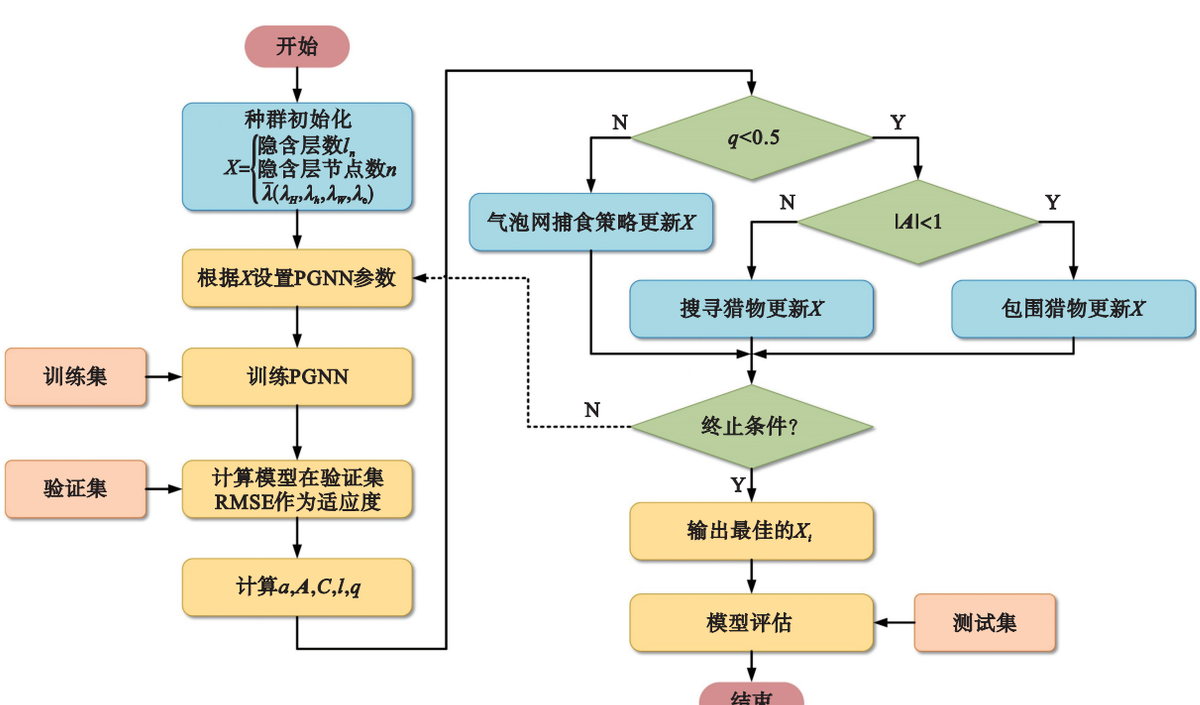


图5　使用WOA优化PGNN超参数的示意图

Fig. 5　Schematic diagram of optimizing PGNN hyperparameters using WOA

2 数据集描述与模型构建

2.1 数据集描述

本研究所采用的数据来源于中厚板产线实际轧制数据.根据轧制理论及生产经验,本文选取对横向变厚度轧制力影响显著的变量作为模型输入,主要包括:材料成分、工艺参数、横向厚度变化特征参数等,各参数见表1.

表1 横向变厚度轧制力数据集的参数

Table 1 Parameters of dataset for transverse variable thickness rolling force

参数符号	参数名称	单位
H	入口厚度	mm
h	出口厚度	mm
l_p	预行程长度(横向厚度参数)	mm
l_d	过渡区长度(横向厚度参数)	mm
h_d	动态压下量(横向厚度参数)	mm
W	入口宽度	mm
v	轧制速度	$m\cdot s^{-1}$
θ	温度	℃
w(C)	碳质量分数	%
w(Mn)	锰质量分数	%
w(Si)	硅质量分数	%
w(Nb+Ti)	微合金质量分数	%
P	轧制力(输出)	kN

在数据预处理阶段,首先采用3σ原则对异常值进行剔除,保留约4 000条有效样本数据作为本研究的主要数据集.随后,将数据集按70%∶15%∶15%的比例划分为训练集、验证集与测试集.数据划分采用随机抽样方式.为消除变量数量级的影响,所有变量在输入模型前均进行了归一化处理.

2.2 基于WOA的模型参数优化

在模型结构设计中,本文各隐含层均采用ReLU函数作为激活函数,输出层采用线性激活函数.学习率设定为0.01,优化算法选用Adam自适应优化器.为进一步提升模型性能,本文使用WOA对PGNN的超参数进行全局搜索.WOA种群规模设为50,最大迭代次数为100.优化目标为最小化验证集的总损失函数值.WOA优化过程中,模型在验证集上的均方根误差(root mean square error,RMSE)变化如图6所示,随着迭代的进行,模型RMSE显著下降并最终收敛,表明WOA有效提升了模型性能.经优化后得到的最优模型结构与参数如表2所示.

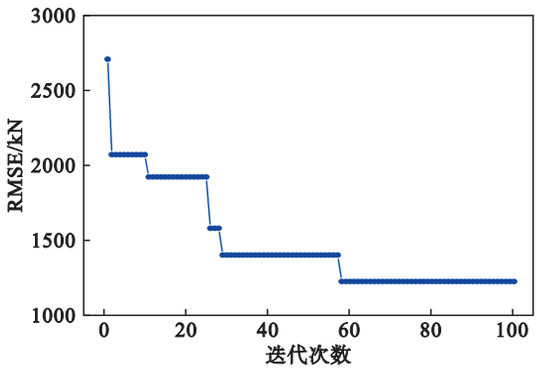


图6 WOA优化模型超参数过程

Fig. 6 Process of optimizing model hyperparameters using WOA

表2 PGNN的超参数优化结果

Table 2 Optimization results of PGNN hyperparameters

参数	优化值
隐含层数量	3
隐含层节点数	20–20–20
λ_H	0.020 4
λ_h	0.000 5
λ_W	0.007 1
λ_e	0.012 9

模型性能评估采用以下4个常用指标:平均绝对误差(mean absolute error,MAE)、均方根误差(RMSE)、平均绝对百分比误差(mean absolute percentage error,MAPE)以及决定系数(coefficient of determination,R^2).其计算式如下:

$$\mathrm{MAE}=\frac{1}{n}\sum_{i=1}^{n}|y_i-\hat{y}_i|, \tag{8}$$

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i-y_i)^2}, \tag{9}$$

$$\mathrm{MAPE}=\frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i-\hat{y}_i}{y_i}\right|\cdot 100\%, \tag{10}$$

$$R^2=1-\frac{\sum_{i=1}^{N}\left(y_i-\hat{y}_i\right)^2}{\sum_{i=1}^{N}\left(y_i-\bar{y}\right)^2}. \tag{11}$$

式中:y_i和$\hat{y}_i$分别为真实值和预测值;$\bar{y}$是实际值的平均值.

3 结果与分析

3.1 预测精度评估

图7展示了模型在测试集上的预测结果比较.图中,对角线表示理想拟合位置,点的颜色表

示预测误差的大小.由图7可见,PGNN模型的预测点更紧密地分布在对角线附近,而传统DNN模型的偏离更明显.同时,红色误差较大的点在PGNN中显著减少,说明其预测误差总体更小.

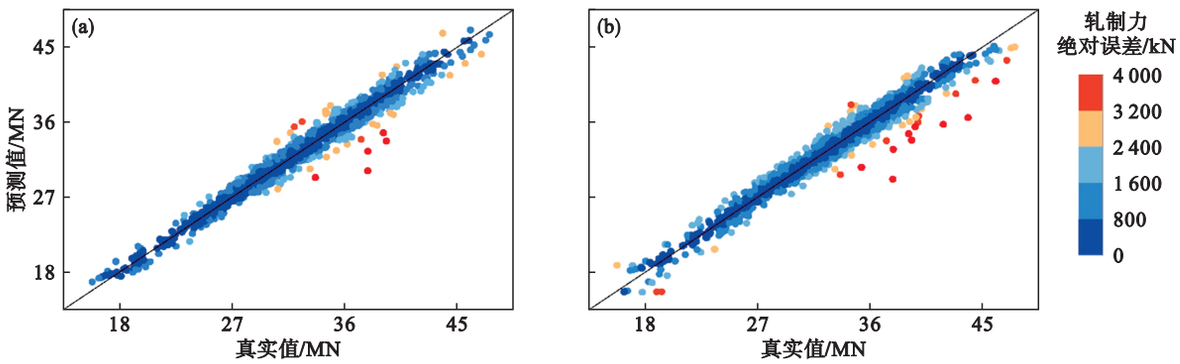


图7　模型轧制力预测结果

Fig. 7　Rolling force prediction results of models

(a)—PGNN; (b)—DNN.

为进一步验证PGNN的有效性,本文选取多种常见机器学习模型:深度神经网络(DNN),支持向量回归(SVR),随机森林(RF),深度置信网络(DBN),极限学习机(ELM)和机理模型作为对比模型,表3为各个模型在测试集上的评价结果.可以看出,所有机器学习模型的预测精度都超过了机理模型.通过简化轧制理论得出的机理模型仅依赖可观测变量,而忽略了其他变量和环境的耦合影响,因此在处理多变量、非线性问题时的准确性较低.相反,机器学习模型可以深度挖掘数据中的耦合规律和潜在影响因素,更适合解决此类问题.

表3　模型对于横向变厚度轧制力预测的评价结果

Table 3　Evaluation results of models for prediction of transverse variable thickness rolling force

评价指标	PGNN	DNN	RF	DBN	SVR	ELM	机理模型
MAE/kN	867. 61	941. 03	1 019. 40	1 076. 62	1 125. 46	1 250. 95	1 683. 28
RMSE/kN	1 204. 13	1 240. 57	1 310. 26	1 355. 80	1 400. 72	1 460. 39	2 255. 71
MAPE/%	2. 63	2. 90	3. 17	3. 44	3. 85	4. 08	9. 15
R^2	0. 972 2	0. 969 3	0. 959 2	0. 955 1	0. 952 2	0. 939 0	0. 900 4

为更直观地展示各模型的预测性能,从测试集中选取3组具有代表性的横向变厚度轧制过程进行实例分析,其关键工艺参数如表4所示.图8展示了3个代表性实例中,不同模型预测值与真实轧制力的对比情况.可以看出,在所有测试实例中,PGNN模型(图中红点)的预测值最接近真实轧制力(图中红色虚线),表现出最优的预测精度.相比之下,传统机理模型在所有实例中均出现明显低估,而DNN,RF等纯数据驱动模型则表现出较大的预测波动性,在不同实例中的预测稳定性不足.这一结果直观地证实了PGNN模型在横向变厚度轧制力预测中的优越性,既能保持高精度,又具备良好的预测稳定性.

表4　轧制案例的部分关键变量

Table 4　Some key variables of rolling case

实例	钢种牌号	入口厚度/mm	出口厚度/mm	入口宽度/mm	动态压下量/mm	实际轧制力/kN
1	Q355ND	229. 95	200. 70	3 522. 30	2. 85	38 408
2	Q355C	242. 04	213. 80	3 975. 70	2. 79	40 138
3	Q355NC	234. 62	199. 12	2 779. 30	1. 63	33 937

3.2　模型对数据量的依赖性

模型对数据量的依赖程度是衡量其实用性的关键指标.在工程实践中,数据采集往往耗时且成本高,尤其对于特殊牌号或新工艺产品,样本构建更加困难.因此,若模型对数据量依赖过高,将限制其在生产中的快速部署.为评估PGNN

的数据依赖性,本文采用不同训练数据量进行对比实验.将训练集随机采样为5种规模:100%,50%,20%,10%,5%,保持验证集与测试集不变.每种情况下重复实验5次取平均值,并计算RMSE与R^2,实验结果如图9所示.随着训练样本数量减少,各模型在训练集与测试集上的性能均呈下降趋势,R^2逐渐降低.这是由于样本减少导致模型学习到的信息不足,同时过拟合现象加剧.然而,在各样本比例下,PGNN的RMSE和R^2始终优于其他模型,且性能下降幅度最小,说明该模型在小样本条件下仍保持良好的泛化性能.这一现象可归因于将领域知识纳入了损失函数,机理约束可以指导模型收敛方向更接近全局最优,从而减少模型的过拟合.上述分析表明,所提出模型有效地降低了模型对数据量的依赖性.

3.3 模型对数据质量的依赖性

在工业生产中,受测量误差与传输噪声影响,实际数据往往含有不同程度的噪声.噪声会干扰模型学习过程,降低其泛化性能,特别是纯数据驱动模型容易在高噪声条件下过度拟合噪声特征.因此,评估模型对数据质量的敏感性,对验证其鲁棒性至关重要.

为此,本文在训练集的标签数据中分别加入1%,2%和5%的高斯白噪声.加入噪声后的数据可表示为

$$\left.\begin{aligned} P_{\mathrm{n}} &= P_{\mathrm{or}} + N(0,\sigma_{\mathrm{N}}^2),\\ \sigma_{\mathrm{N}} &= \eta \cdot P_{\mathrm{mean}}. \end{aligned}\right\} \tag{12}$$

其中:P_{or}和P_{n}分别为加噪前后的轧制力;P_{mean}为轧制力平均值;η为噪声系数.

分别向原有训练集的标签中添加1%,2%和5%的噪声,并分别基于该数据建立模型,利用其在测试集上的R^2值来评价模型性能.表5为噪声实验的结果,随着噪声的逐渐增加,所有模型的性能都出现了明显的下降,这种下降可归因于随着噪声水平的上升,模型内的有效特征逐渐模糊,从而阻碍了模型的收敛.但在这一过程中,PGNN的性能始终是最优的,且性能损失最少.这一结果表明,在训练过程中通过轧制理论对模型进行约束,可以有效降低模型对数据质量的依赖性.

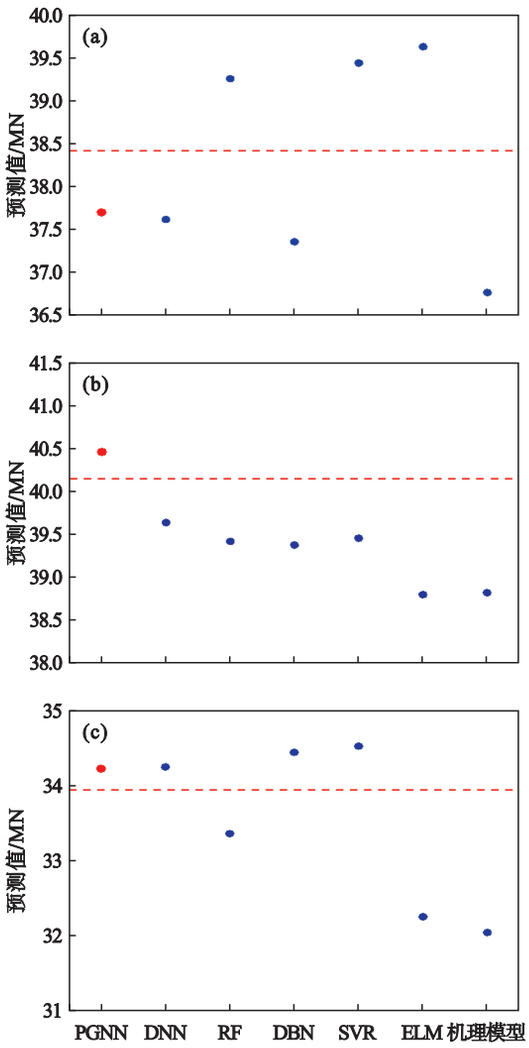
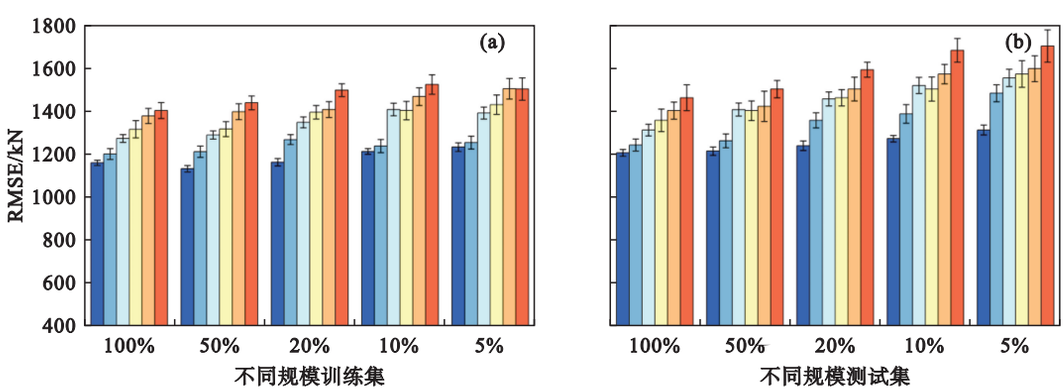


图8 各模型对实例的预测结果对比

Fig. 8 Comparison of prediction results from different models for the case

(a)—实例1;(b)—实例2;(c)—实例3.



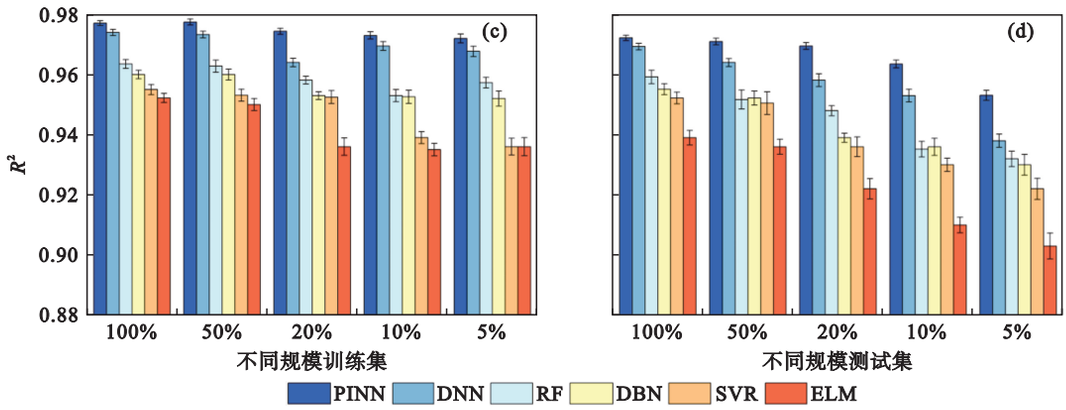


图 9　不同规模训练集下的模型性能

Fig. 9　Model performance under training sets of different scales

(a)—训练集的RMSE;(b)—测试集的RMSE;(c)—训练集的R^2;(d)—测试集的R^2.

表 5　噪声对模型 R^2 的影响

Table 5　Effect of noise on R^2 of models

噪声等级	PGNN	DNN	RF	DBN	SVR	ELM
0%	0.972 2	0.969 3	0.958 0	0.953 1	0.952 2	0.939 0
1%	0.971 9	0.963 1	0.951 2	0.944 1	0.947 3	0.926 5
2%	0.968 7	0.958 3	0.940 5	0.927 6	0.939 5	0.912 4
5%	0.962 5	0.940 5	0.927 3	0.909 4	0.912 0	0.882 2

3.4　模型分布外数据泛化性能

作为一个典型的多元耦合过程,轧制过程相关的数据具有广泛分布的特征,更容易受到非常规生产条件的影响,因此需要对模型在具有不同分布数据(包括分布外数据)中的泛化性能进行测试,以保证其实际应用效果.因此,除3.1节提到的数据(记为数据集A)外,实验中还利用了其他生产环境下的横向变厚度轧制过程数据(记为数据集B)来评估模型的性能.两个数据集的轧制力分布如图10所示,可以看出两个数据集在分布范围和分布形式上都有所不同,并且数据集B的数据量较小.在实验中,利用数据集A的训练集建立模型,利用数据集B的测试集评价模型,将受输出范围影响较小的MAPE和R^2作为该部分的评价指标.

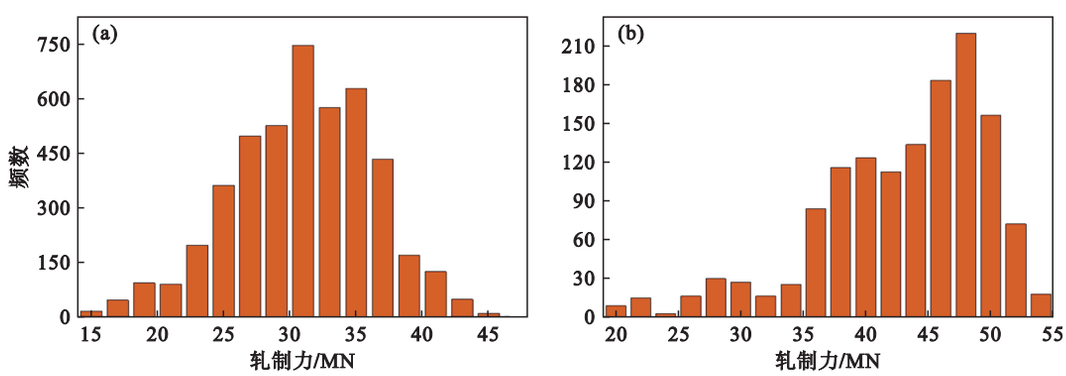


图 10　不同数据集的轧制力分布

Fig. 10　Rolling force distribution of different datasets

(a)—数据集A;(b)—数据集B.

评价结果见表6,表中"*-A"表示使用数据集A作为训练集建立模型、将数据集B作为测试集测试的结果.与表3相比,单纯的机器学习模型在新数据集上的性能下降过多,这表明纯数据驱动模型不具备良好的整体泛化和外推性能.相比之下,PGNN的MAPE和R^2值相对稳定,这种稳定性源于模型在训练过程中受到理论约束,有助于其学习到数据中的理论特征,从而在面对与训练集分布不同的数据时具有更好的鲁棒性.

3.5　模型的理论一致性验证

除了评估模型在各种情况下的精度指标外,还须评估模型与轧制理论之间的一致性,这一结果涉及模型在实际应用中的可信度.理论一致性要求模型在满足精度要求的前提下,输出值与实际测量值之间具有良好的正线性相关,即具有相同的变化趋势.这可以确保模型不会产生与实际

值偏差过大的预测点,从而避免更大的生产损失.通常采用皮尔逊相关系数(r)来衡量模型与理论的全局一致性,如式(13)所示,r越接近1,模型的全局理论一致性越好.图11为各个模型的皮尔逊相关系数计算结果,由于在损失函数中加入了理论约束,与其他模型相比,PGNN的r值更接近于1,这表明PGNN的预测结果更接近实际生产情况.

表6 模型对于不同数据集的泛化性能

Table 6 Generalization performance of models on different datasets

评价指标	PGNN-A	DNN-A	RF-A	DBN-A	SVR-A	ELM-A
MAPE/%	2.98	3.64	4.62	4.41	4.77	5.43
R^2	0.967 3	0.953 3	0.928 4	0.931 4	0.926 5	0.920 4

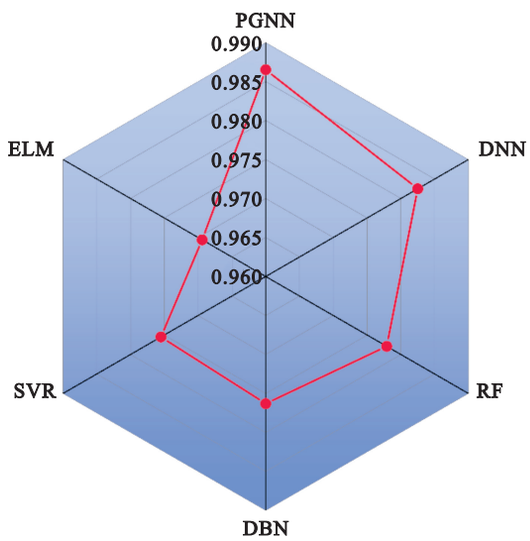


图11 模型的皮尔逊相关系数

Fig. 11 Pearson correlation coefficients of models

$$r=\frac{\sum_{i=1}^{n}(\hat{y}_i-\bar{y})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2\sum_{i=1}^{n}(y_i-\bar{y})^2}}. \tag{13}$$

模型的局部理论一致性在于变量对输出的影响规律.数据驱动模型是从数据中挖掘内部特征,大多数机器学习模型更像一个“黑箱”,虽然采用了统计指标来评估模型的各种性能,但对于模型理解数据的方式仍不清楚,甚至模型所得规律以和实际理论相反的方式影响输出,这会提高模型实际应用的潜在风险.因此,具有局部理论一致性的模型在应用中具有更好的安全性,在实际生产中更受信任.

4 结 论

1)将轧制机理理论与深度神经网络(DNN)模型相结合,提出了一种基于物理指导神经网络(PGNN)的横向变厚板轧制力预测方法.该方法以神经网络为核心,通过在模型损失函数中引入包含轧制机理知识的约束项,指导模型训练过程.模型的最优超参数则通过鲸鱼优化算法(WOA)进行确定.

2)在横向变厚板轧制力预测中,PGNN优于理论模型和纯数据驱动模型,其平均绝对误差(MAE)、均方根误差(RMSE)、平均绝对百分比误差(MAPE)及决定系数(R^2)分别为867.61 kN,1 204.13 kN,2.63%和0.972 2.结果表明,将机理知识融入神经网络能够显著提升模型预测的准确性.

3)PGNN在样本量有限或数据噪声较大的情况下,仍表现出优异的预测性能.随着可用样本数量减少或数据噪声增大,PGNN能够最大限度地降低性能损失.这说明将轧制机理知识引入DNN模型,有效减轻了模型对数据量和数据质量的依赖,从而降低过拟合风险,为模型的快速开发和稳定应用提供了一种新思路.

4)PGNN整体泛化性能良好.在不同数据分布及分布外数据上均表现出稳健的泛化性能.同时,PGNN预测值与实际值的皮尔逊相关系数接近1,表明两者间存在强正线性关系,因此PGNN的预测结果与轧制力变化的理论认知高度一致.

参考文献:

[1] 王国栋. 创建钢铁企业数字化创新基础设施加速钢铁行业数字化转型[J]. 轧钢, 2022, 39(6): 2-11.
(Wang Guo-dong. Creating a digital innovation infrastructure for iron and steel enterprises to accelerate the digital transformation of the iron and steel industry[J]. *Steel Rolling*, 2022, 39(6): 2-11.)

[2] 彭文, 姬亚锋, 陈小睿, 等. 热轧非稳态过程轧制力自学习模型优化[J]. 东北大学学报(自然科学版), 2019, 40(10): 1408-1412.
(Peng Wen, Ji Ya-feng, Chen Xiao-rui, et al. Optimization of rolling force self-learning model in unsteady process of hot rolling[J]. *Journal of Northeastern University*(*Natural Science*), 2019, 40(10): 1408-1412.)

[3] Cao Y, Cao J G, Wang T, et al. The cold rolling load distribution of the nuclear power zirconium alloy based on the self-adaptive particle swarm optimization algorithm[J]. *International Journal of Advanced Manufacturing*

Technology, 2022, 119(9): 6007-6016.

[4] Yang Y B, Peng Y. Dynamic rolling model based on uniform deformation [J]. *Journal of Manufacturing Processes*, 2020, 58: 1334-1347.

[5] Sims R B . The calculation of roll force and torque in hot rolling mills [J]. *Proceedings of the Institution of Mechanical Engineers*, 1954, 168(1): 191-200.

[6] Liu Y M, Ma G S, Zhao D W, et al. Analysis of hot strip rolling using exponent velocity field and MY criterion[J]. *International Journal of Mechanical Sciences*, 2015, 98: 126-131.

[7] Dong Z S, Li X, Luan F, et al. Prediction and analysis of key parameters of head deformation of hot-rolled plates based on artificial neural networks [J]. *Journal of Manufacturing Processes*, 2022, 77: 282-300.

[8] Ding J G, Wang G Q, He Y H C, et al. Controllable points setting method for plan view pattern control in plate rolling process[J]. *Steel Research International*, 2020, 91(1): 1900345.

[9] 闫松涛, 杨斌, 江海涛. 基于ABAQUS二次开发的6016铝合金热连轧过程有限元分析[J]. 塑性工程学报, 2024, 31(6): 87-94.
(Yan Song-tao, Yang Bin, Jiang Hai-tao. Finite element analysis of hot continuous rolling process of 6016 aluminum alloy based on secondary development of ABAQUS [J]. *Journal of Plasticity Engineering*, 2024, 31(6): 87-94.)

[10] Ji Y F, Song L B, Yuan H, et al. Prediction of strip section shape for hot-rolled based on mechanism fusion data model [J]. *Applied Soft Computing*, 2023, 146: 110670.

[11] Dong Z S, Li X, Luan F, et al. Point and interval prediction of the effective length of hot-rolled plates based on IBES-XGBoost[J]. *Measurement*, 2023, 214: 112857.

[12] 王振华, 龚殿尧, 李广焘, 等. 遗传算法优化神经网络的热轧带钢弯辊力预报模型[J]. 东北大学学报(自然科学版), 2018, 39(12): 1717-1722.
(Wang Zhen-hua, Gong Dian-yao, Li Guang-tao, et al. Bending force prediction model in hot strip rolling based on artificial neural network optimize by genetic algorithm[J]. *Journal of Northeastern University* (*Natural Science*), 2018, 39(12): 1717-1722.)

[13] Ji Y J, Liu S X, Zhou M C, et al. A machine learning and genetic algorithm-based method for predicting width deviation of hot-rolled strip in steel production systems[J]. *Information Sciences*, 2022, 589: 360-375.

[14] Wu Y, Ni H C, Li X, et al. Prediction of bending force in the hot strip rolling process using multilayer extreme learning machine [J]. *Mathematical Problems in Engineering*, 2021, 2021(1): 6682660.

[15] Wang Q N, Song L B, Zhao J W, et al. Application of the gradient boosting decision tree in the online prediction of rolling force in hot rolling[J]. *The International Journal of Advanced Manufacturing Technology*, 2023, 125(1): 387-397.

[16] Huang X L, Guo X J, Li Y L, et al. A novel data enhancement approach to DAG learning with small data samples[J]. *Applied Intelligence*, 2023, 53(22): 27589-27607.

[17] Wang N Z, Zhang D X, Chang H B, et al. Deep learning of subsurface flow via theory-guided neural network [J]. *Journal of Hydrology*, 2020, 584: 124700.

[18] Leiser F, Rank S, Schmidt-Kraepelin M, et al. Medical informed machine learning: a scoping review and future research directions[J]. *Artificial Intelligence in Medicine*, 2023, 145: 102676.

[19] Zhang Z Y, Zhang H, Liu Y, et al. Generalized conditional symmetry enhanced physics-informed neural network and application to the forward and inverse problems of nonlinear diffusion equations[J]. *Chaos Solitons & Fractals*, 2023, 168: 113169

[20] Arzani A, Cassel K W, D'souza R M. Theory-guided physics-informed neural networks for boundary layer problems with singular perturbation [J]. *Journal of Computational Physics*, 2023, 473: 111768.

[21] Meng L M, Ding J G, Li X J, et al. Novel shape control system of hot-rolled strip based on machine learning fused mechanism model [J]. *Expert Systems with Applications*, 2024, 255: 124789.

[22] Wen Y, Ji Y F, Wu B R, et al. A mechanism and data fusion step-wise prediction framework for hot-rolled strip thickness[J]. *Journal of Manufacturing Processes*, 2025, 134: 680-696.

[23] Zhang S H, Deng L, Che L Z, et al. An integrated model of rolling force for extra-thick plate by combining theoretical model and neural network model [J]. *Journal of Manufacturing Processes*, 2022, 75:100-109.

[24] Cao L, Li X, Li X H, et al. Variable speed rolling force prediction with theoretical and data-driven models [J]. *International Journal of Mechanical Sciences*, 2024, 264: 108833.

[25] Rout M, Pal S K, Singh S B. Finite element analysis of cross rolling on AISI 304 stainless steel: prediction of stress and strain fields[J]. *Journal of the Institution of Engineers* (*India*): *Series C*, 2017, 98(1): 27-35.

[26] Zhao Y, Yang Q, He A R, et al. Precision plate plan view pattern predictive model [J]. *Journal of Iron and Steel Research*, *International*, 2011, 18(11): 26-30.

[27] 矫志杰, 何纯玉, 丁敬国, 等. 中厚板轧机平面形状控制技术的工业推广应用[J]. 钢铁, 2019, 54(1): 49-55.
(Jiao Zhi-jie, He Chun-yu, Ding Jing-guo, et al. Industrial popularization and application of plan view pattern control technology for plate mill[J]. *Iron and Steel*, 2019, 54(1): 49-55.)

[28] Wang X L, Zhang W D, Ai Y B. A novel model based on quartic velocity field for hot rolling force of aluminum alloy thick plate [J]. *The International Journal of Advanced Manufacturing Technology*, 2023, 127(11): 5119-5125.

[29] Mirjalili S, Lewis A. The whale optimization algorithm[J]. *Advances in Engineering Software*, 2016, 95: 51-67.