

基于在线学习的柴油机起动过程 NO_x 排放预测模型

杜征宇¹, 李华杰², 郭志坤³, 梁永森³, 石磊^{1*}

1. 上海交通大学动力机械及工程教育部重点实验室, 上海 200240;

2. 中国人民解放军 92118 部队, 浙江 舟山 316000; 3. 中国北方发动机研究所, 天津 300400

摘要:为了构建准确的柴油机起动过程 NO_x 排放预测模型, 分析起动过程瞬态运行特点, 基于非线性自回归神经网络, 结合起动试验数据选取模型特征, 对比自注意力机制模型与标准反向传播神经网络模型 NO_x 排放预测效果, 发现自注意力机制模型预测效果良好, 且均方根误差较小; 对比自注意力机制模型在在线梯度下降算法和 FTRL 算法 2 种在线学习的预测效果, 发现采用 FTRL 算法可以使模型在未知工况下获得更好的预测性能; 对比自注意力机制模型学习前、后的预测效果, 发现学习后的模型均方根误差增大, 但预测结果能力良好; 对比在线学习模型与全数据训练集离线学习模型预测效果, 发现两者的决定系数与均方根误差相差不大, 预测效果都较好, 但在线学习所用数据量和训练时间分别减少了 68.7% 和 73.6%, 大幅降低了存储和训练成本。采用 FTRL 算法的自注意力机制模型可减少数据存储和训练资源成本, 并实时反馈预测需求。

关键词: 柴油机; 起动过程; NO_x 预测; 自注意力机制; 在线学习

中图分类号: TK421

文献标志码: A

文章编号: 1673-6397(2024)05-0008-09

引用格式: 杜征宇, 李华杰, 郭志坤, 等. 基于在线学习的柴油机起动过程 NO_x 排放预测模型[J]. 内燃机与动力装置, 2024, 41(5): 8-16.

DU Zhengyu, LI Huajie, GUO Zhikun, et al. Prediction model for NO_x emission of starting process of the diesel engine based on online learning[J]. Internal Combustion Engine & Powerplant, 2024, 41(5): 8-16.

0 引言

柴油机的 NO_x 排放因严重影响空气质量和公共健康而受到人们的日益关注, 目前通常采用便携式排放检测系统(portable emission measurement system, PEMS)进行柴油机 NO_x 排放测试, 但 PEMS 设备具有成本高和测试时间长的缺点, 难以满足广泛、长期及动态测试环境的监测需求^[1], 建立精确的柴油机 NO_x 排放预测模型是当下研究热点之一。

目前关于柴油机 NO_x 排放预测模型的研究多集中于平缓路面、稳定行驶等常规运行条件^[2-4], 但在 PEMS 循环测试中, 柴油机起动过程 NO_x 排放占比超过 40%^[5], 由于柴油机起动过程燃烧不稳定且为瞬态过程, 其 NO_x 排放特性与常规运行状态有较大区别^[6-9], 因此有必要对瞬态起动过程的 NO_x 排放进行深入的建模研究。Seo 等^[10]以柴油机冷起动过程的转矩、转速等参数作为多层感知机的输入, 预测了 NO_x 、 CO_2 等气体的排放; Tang 等^[11]针对混合动力汽车起动过程, 采用遗传算法对神经网络初始权重进行

收稿日期: 2024-05-23

基金项目: 基础产品创新计划(2023027)

第一作者简介: 杜征宇(1998—), 男, 山东东营人, 工学硕士, 主要研究方向为发动机排放与智能控制, E-mail: duzhengyu@sjtu.edu.cn。

* 通信作者简介: 石磊(1977—), 男, 山东潍坊人, 研究员, 工学博士, 博士生导师, 主要研究方向为动力系统及内燃机智能控制, E-mail: shi_lei@sjtu.edu.cn。

优化,得到了较为准确的 NO_x 排放预测结果。但目前关于起动过程中 NO_x 排放的研究普遍未考虑瞬态过程的影响,由于柴油机瞬态过程中各个工况并不独立,每一工况均受之前工况影响,具有高度的时序性,合理选取反映起动过程中瞬态工况的特征参数,并针对复杂的排放规律建立准确的预测模型,提高起动过程中 NO_x 排放预测的准确性^[12]。不同工作条件下,起动过程 NO_x 生成规律较为复杂,在构建排放预测模型时,初始训练集难以覆盖所有工作条件,模型对不同工作环境适应性差;未知工况 NO_x 排放规律与训练集工况差别明显时,会产生较大的预测误差;为了保证模型在不同车型及环境下通用,需适时更新训练模型。若采用历史数据和累积新数据同时重新训练模型,需要较大的存储空间与计算资源,效率较低;若只采用累积新数据重新训练模型也需较长时间,难以迅速反馈当前发动机的预测需求^[13]。因此,本文中以试验数据为训练样本,选取特征参数,对比分析不同模型、不同在线学习方法及同一模型下在线学习前、后的预测结果,并对比不同训练数据集对应的模型预测效果、训练成本,为发动机排放预测模型开发提供参考。

1 试验方法与数据处理

1.1 试验方法

试验用发动机为一台自主设计开发的水冷、直喷、四冲程、单缸柴油机,发动机主要参数如表1所示。

表1 发动机主要参数

缸径/mm	行程/mm	压缩比	燃烧室	气门数	额定功率/kW	额定转速/($\text{r} \cdot \text{min}^{-1}$)	最大扭矩/($\text{N} \cdot \text{m}$)	最大扭矩转速/($\text{r} \cdot \text{min}^{-1}$)
126	155	17	ω 型	4	19.5	1 500	132.8	1 200

采用 LabVIEW 编写喷油控制程序,结合 cRIO-9012 嵌入式控制器控制喷油压力、喷油正时、喷油脉宽;结合冷却模块、离心风机、进排气控制阀、稳压罐与真空泵等搭建变海拔环境模拟系统,模拟进、排气压力为 56~101 kPa,海拔高度分别为 0、1、2、3、4 km。在喷油正时、喷油压力等参数相同时,不同海拔高度起动过程的 NO_x 排放如图1所示。

由图1可知:不同海拔高度下 NO_x 排放存在明显差异;海拔高度为 0~2 km 时,随着海拔升高, NO_x 排放降低;海拔高度为 3~4 km 时,随着海拔升高, NO_x 排放升高。主要原因为: NO_x 需要在富氧、高温的条件下生成,海拔高度为 0~2 km 时,缸内氧气质量分数减小,氧气更加稀薄,不利于 NO_x 生成, NO_x 排放降低;海拔高度为 3~4 km 时,滞燃期大幅增加,预混燃烧占比增大,混合气形成更充分,缸内燃烧温度升高,有利于 NO_x 生成, NO_x 排放升高;不同海拔下,由于缸内氧气质量分数和滞燃期的变化, NO_x 排放差异明显。

对该柴油机进行 96 组(66 组平原环境、30 组高海拔环境)起动试验,试验过程中记录相应的边界参数及发动机转速、转矩。以进气压力为 70.1 kPa、喷油正时对应的曲轴转角为上止点前 8° 、喷油压力为 120 MPa 时为例,起动过程中转速、转矩与 NO_x 排放如图2所示。

由图2可知:柴油机起动过程中,转速、转矩的

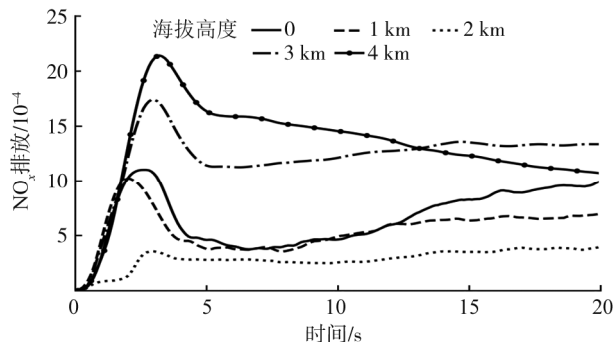


图1 不同海拔高度起动过程 NO_x 排放

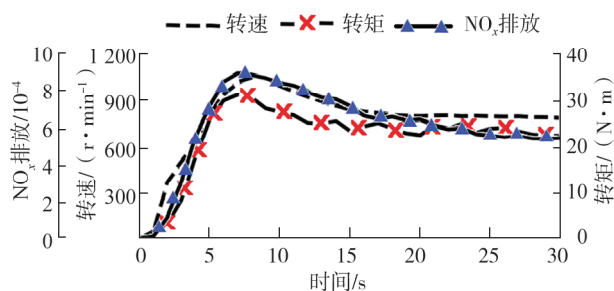


图2 起动过程转速、转矩与 NO_x 排放

变化趋势与 NO_x 排放有一定的一致性。

每组起动试验边界参数(包括喷油压力、喷油正时、进气压力等)保持不变,试验时,发动机的前2个循环由起动机拖动旋转,当转速升高超过怠速转速(800 r/min)时,对循环喷油量进行闭环控制,直至发动机转速稳定在怠速转速。将试验采集数据作为初始离线数据集,在该数据集上进行模型初始训练、验证与评估,并投入模型实际应用中。

1.2 数据处理

为了模拟实时预测未知工况的在线学习环境,将数据集分成离线数据集 L 和在线数据集 O ,采样频率为 0.1 ms,离线数据集包括 66 组平原环境试验数据,共 18 125 个数据样本;在线数据集包括海拔高度分别为 1、2、3、4 km 时的 30 组试验数据,共 8 262 个数据样本点。离线数据集数据样本的 70% 为训练集 L_1 ,用于模型训练;15% 为验证集 L_2 ,用于模型超参数调整;15% 为测试集 L_3 ,用于评估模型预测性能。在线数据集同样划分成训练集 O_1 、验证集 O_2 、测试集 O_3 ,划分比例与离线数据集相同。

Pearson 因数 k 表示参数间的相关性, $k > 0.7$,表明参数间相关性较好^[14]。对试验边界参数与 NO_x 排放进行相关性分析,由于每组试验边界参数保持不变,因此采用 NO_x 累积排放进行相关性分析,计算各边界参数、试验数据与 NO_x 累积排放的相关性分析,结果如表 2 所示,表 2 中 k_p 、 k_x 、 k_s 、 k_n 、 k_m 、 k_j 分别为喷油压力、循环喷油量、喷油正时、转速、转矩、进气压力与 NO_x 累积排放相关的 Pearson 因数。

表 2 各边界参数与 NO_x 累积排放的相关性分析结果

k_p	k_x	k_s	k_n	k_m	k_j
0.81	0.83	0.79	0.74	0.85	0.76

由表 2 可知,Pearson 因数均大于 0.7,表明各参数与 NO_x 累积排放存在明显相关性,各边界参数可以作为预测模型的输入变量。由于各边界参数的量纲差别较大,需对原数据进行归一化或者标准化处理,本文中选用 Min-Max 归一化方法,归一化后样本参数

$$x' = (x - x_{\min}) / (x_{\max} - x_{\min}), \quad (1)$$

式中: x 为当前样本参数, $x_{\max}$ 为最大样本参数, $x_{\min}$ 为最小样本参数。

2 排放预测模型

2.1 模型特征

时序性神经网络考虑了各特征参数的时序性,可应用于模型构建^[15-16]。起动过程是一个连续且有时序性的过程,本文中采用时序性神经网络并加入转速等参数的历史数据作为模型的特征参数。

非线性自回归(nonlinear autoregressive with external input, NARX)神经网络结合历史时间和外在数据,可用于预测复杂非线性关系的时间序列,尤其适用于非平稳环境的预测^[17]。本文中采用 NARX 神经网络的输入层作为模型特征参数输入,当前工况采用前 n 个时刻的输入、输出延迟及当前时刻的喷油压力、循环喷油量、转速、转矩等参数作为输入特征参数,NARX 神经网络架构如图 3 所示,图中 t 为当前时刻, n 为 NARX 神经网络时间序列步长, p_m 、 n_m 、 M_m 、 E_m 分别为 m 时刻对应的喷油压力、转速、转矩、 NO_x 排放, $m = t, t-1, \dots, t-n$ 。

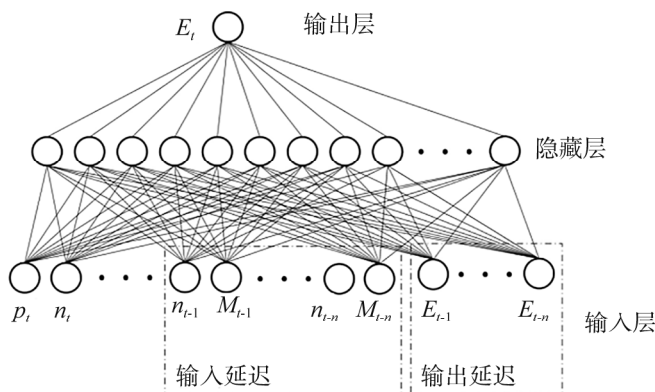


图 3 NARX 神经网络架构图

2.2 自注意力机制

选取模型特征参数后,基于自注意力机制搭建 NO_x 排放预测模型。由于 NO_x 排放与喷油正时、循环喷油量等参数存在明显相关性,且喷油参数变化直接影响转速、转矩,模型特征参数相关性较高,容易造成模型重复学习,降低神经网络识别效率,且携带重复信息的特征参数造成模型预测失准,本文提出了结合 NARX 神经网络特征输入和自注意力机制的 NO_x 排放预测模型。

自注意力机制是近年来机器学习领域中的一种重要思想,它充分利用特征参数的相关性,使模型对不同输入特征参数进行注意力分数加权^[18-19]。自注意力机制原理示意如图4所示。

由图4可知:特征序列中的每个参数通过3个不同的全连接层生成3个向量,分别为 Query 向量 Q 、Key 向量 K 和 Value 向量 V ;将 Q 与 K 进行点积,得到初始注意力分数矩阵 $Q \cdot K^T$;结合 softmax 激活函数 s 对注意力分数矩阵进行标准化,并与 V 相乘后缩放,得到注意力概率分配矩阵;将注意力概率分配矩阵经过线性运算使其维度对应,输出对应的预测矩阵。

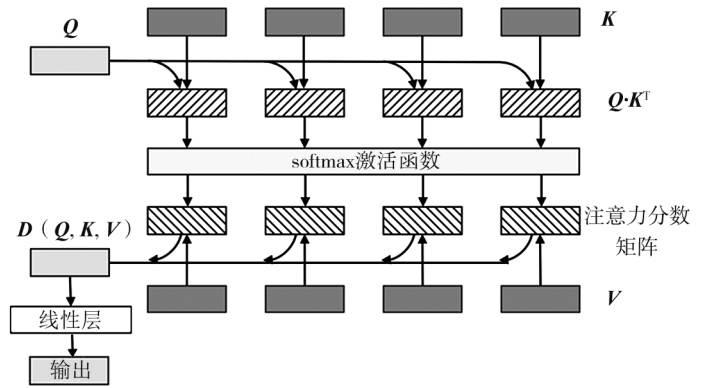


图4 自注意力机制原理示意图

激活函数

$$s(z_i) = e^{z_i} / (\sum_{i=1}^l e^{z_i}), \quad (2)$$

式中: z_i 为第 i 个特征的注意力分数; l 为特征的数量; i 为对应的特征数, $i=1,2,3,\dots,l$ 。

注意力概率分配矩阵

$$D(Q, K, V) = s[(Q \cdot K^T) / \sqrt{d}] V, \quad (3)$$

式中 d 为 K 的维度。

自注意力机制通过注意力分数在一定程度上实现了不同特征参数间的动态权重调整。特征参数变化时,特征参数与预测目标的权重也发生变化,特征参数之间相关性较高时,可以实现更优的拟合效果。结合 NARX 神经网络特征输入和自注意力机制的排放预测模型,主要基于 NARX 神经网络输入层,生成维度大小与输入层特征数量一致的 Q 、 K 和 V 作为自注意力机制模型的输入,计算 NO_x 排放预测结果。

基于搭建的自注意力机制模型,设定 n 为 1~10,在完成各时间步长模型训练后,结合数据集 L_3 中的 10 组数据,采用决定系数评估模型性能,决定系数

$$R^2 = 1 - \sum_j (y_j - \hat{y}_j)^2 / [\sum_j (y_j - \bar{y}_j)^2], \quad (4)$$

式中: j 为样本数, $j=1,2,3,\dots,h$; y_j 为真实样本 NO_x 排放; $\hat{y}_j$ 为预测样本 NO_x 排放; $\bar{y}_j$ 为样本平均 NO_x 排放。 R^2 越接近 1,误差越小,模型预测效果越好。

不同时间步长的决定系数如表3所示。由表3可知:为了获得最佳的预测性能,模型时间步长设为5。

表3 不同时间步长的决定系数

时间步长	1	2	3	4	5	6	7	8	9	10
决定系数	0.970 9	0.969 4	0.972 4	0.975 0	0.976 8	0.975 8	0.974 2	0.972 0	0.971 3	0.969 7

2.3 模型对比分析

对比分析自注意力机制模型与标准反向传播(back propagation, BP)神经网络模型,2种模型输入层、输出层的选取相同,为了更充分的反应模型准确度,采用 R^2 和均方根误差作为评价指标,均方根误差越小,模型准确度越高。

均方根误差

$$\delta_R = \sqrt{(\sum (y_j - \hat{y}_j)^2)/h} \quad (5)$$

基于数据集 L_3 , 试验结果和 2 种模型的预测效果对比如图 5 所示。

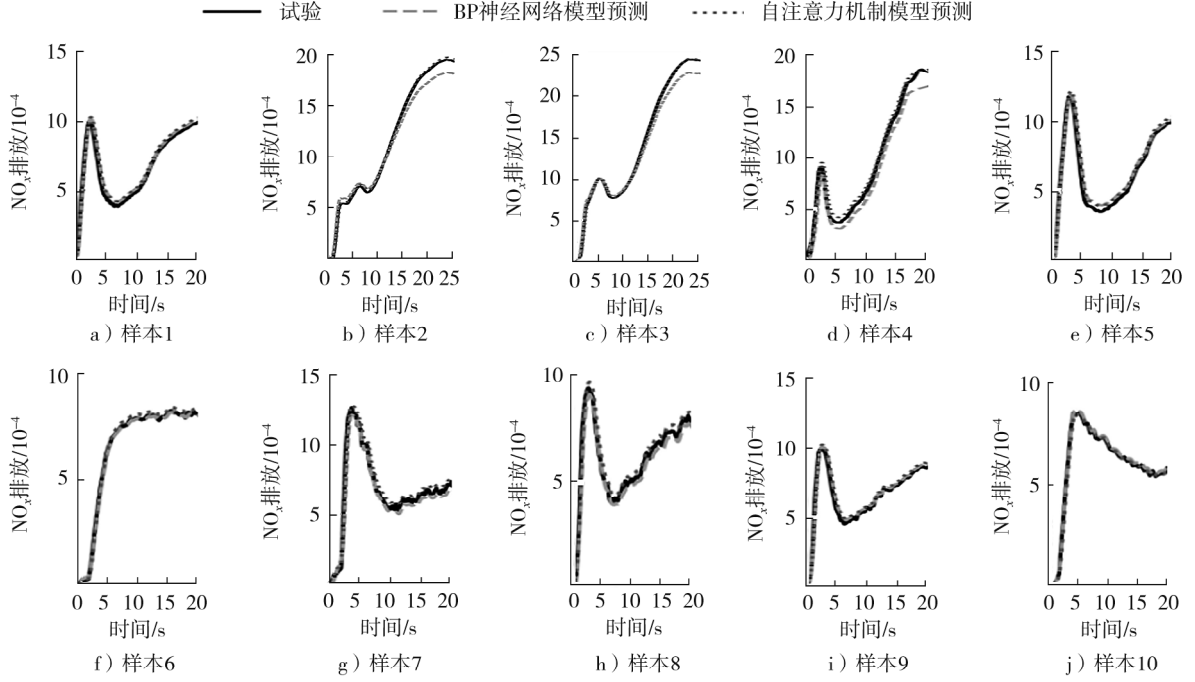


图5 BP神经网络与自注意力机制模型预测效果对比

由图 5 及式(4)(5)可知:BP 神经网络和自注意力机制模型的 R^2 分别为 0.969 15、0.976 88, δ_R 分别为 49.91 与 33.67, 采用自注意力机制模型, 预测样本与真实样本更接近。对比 BP 神经网络模型, 自注意力机制模型结构更为复杂, 表达能力更强, 可以更好的用于起动过程 NO_x 排放预测。

3 模型在线学习

3.1 在线梯度下降与 FTRL 算法

在线学习与传统批处理式机器学习的区别在于在线学习是对在线采集的流式数据的单行样本进行学习, 即在线学习批量为 1。本文中采用在线数据集模拟传感器接收在线数据, 进行样本逐行的训练与预测, 若柴油机实际工况与初始离线数据集采集工况的 NO_x 排放规律不同, 模型将连续接收发动机瞬态过程产生的数据流, 用于模型更新和测试评估。常用的在线学习算法主要有在线梯度下降 (online gradient descent, OGD) 与 FTRL 算法。OGD 是传统梯度下降算法的在线版本, 适用于不断更新的数据流, OGD 算法对学习率的选择非常敏感, 因此采用 Adam 优化器进行学习率的自适应调整, 但 OGD 算法难以解决数据环境非平稳变化时的振荡和收敛问题以及离群噪声对模型准确性影响的问题^[20-21]。

由于正则化限制了权重更新程度, 使模型不会过度学习新样本, 减少了产生灾难性遗忘问题的风险, 因此 FTRL 算法结合了在线梯度下降和正则化, 在每一步梯度计算中都尝试找一个参数, 使历史梯度的加权和与正则化项的和最小, 减少对单个数据点的敏感性和模型的过拟合风险^[22-23]。

OGD 算法第 $t+1$ 时刻的模型参数

$$w_{t+1} = w_t - \eta_t g_t, \quad (6)$$

式中: w_t 为 t 时刻的模型参数, η_t 为 t 时刻的学习率, g_t 为 t 时刻的梯度。

FTRL 算法 t 时刻的学习率

$$\eta_t = \alpha / (\beta + \sqrt{\sum_{s=1}^t g_s^2}), \quad (7)$$

式中: α, β 为超参数,控制模型学习幅度, g_s 为 s 时刻的梯度。

中间变量

$$z_t = g_{1:t} - \sum_{s=1}^t \sigma_s w_s, \quad (8)$$

式中: $g_{1:t}$ 为当前模型对各权重和偏置的计算梯度; σ_s 为中间变量, $\sigma_s = 1/\eta_s - 1/\eta_{s-1}$; w_s 为 s 时刻的模型参数。

FTRL 算法 $t+1$ 时刻的模型参数

$$w_{t+1} = \begin{cases} 0, & |z_t| < \lambda_1 \\ -(z_t - \text{sgn}(z_t) \cdot \lambda_1) / (\lambda_2 + \sum_{s=1}^t \sigma_s), & |z_t| \geq \lambda_1 \end{cases}, \quad (9)$$

式中: λ_1 和 λ_2 为超参数,控制模型学习幅度。选取数据集 O_2 调试超参数后,再进行 NO_x 排放在线预测。

3.2 不同算法预测结果

采用自注意力机制模型,结合数据集 L_3 分析 OGD 算法和 FTRL 算法下 NO_x 排放预测结果如图 6 所示。

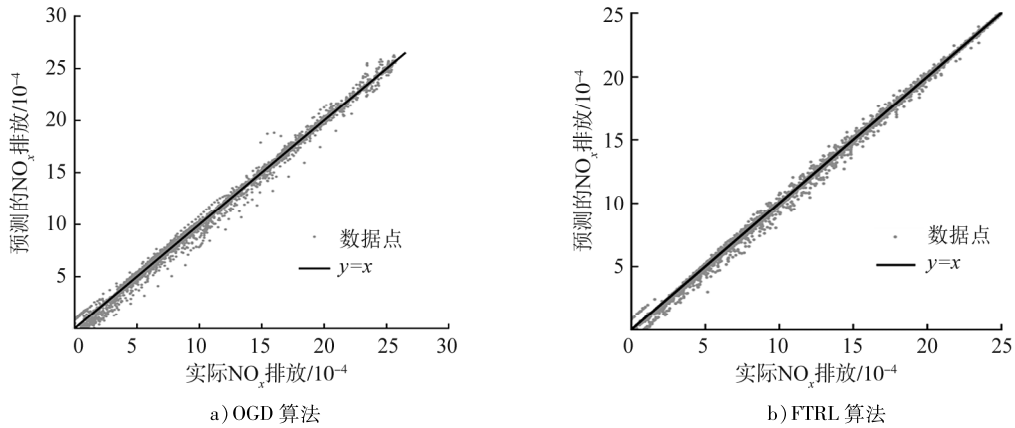


图 6 OGD 和 FTRL 算法的预测效果对比

由图 6 及式(4)(5)可知:采用 OGD、FTRL 算法对应的 R^2 分别为 0.975 12、0.980 39, δ_R 分别为 33.88、24.32; FTRL 算法的 R^2 比 OGD 算法高 0.005 27, δ_R 小 9.56; 2 种算法数据点均集中分布在 $y=x$ 附近。由于起动过程 NO_x 排放随时间的变化趋势不平稳,采用 OGD 算法振荡较大; FTRL 算法由于累积历史梯度、平滑参数和正则化,可以更快地收敛, NO_x 排放预测结果更平稳。

结合数据集 O_3 , 对比不同算法下, NO_x 排放预测结果如图 7 所示。由图 7 可知:对比 FTRL 算法,采用 OGD 算法 NO_x 排放前期更接近试验结果,跟踪误差小,但在排放峰值处出现了较大幅度的振荡。FTRL 与 OGD 算法的 R^2 分别为 0.990 97 和 0.988 15, δ_R 分别为 26.04 和 33.41, 采用 FTRL 算法能够更准确地预测模型在起动过程的 NO_x 排放。

3.3 在线学习前后模型对比

为了研究采用 FTRL 算法在线学习后,自注意力机制模型初始训练工况的预测精度,定义自注意力机制模型采用 FTRL 算法在线学习前的模型为 A 模型,在线学习后的模型为 B 模型,即 A 模型采用数据集 L_1 进行训练, B 模型采用数据集 O_1 进行训练,结合数据集 L_3 , 分析不同模型下 NO_x 预测结果如图 8 所示。

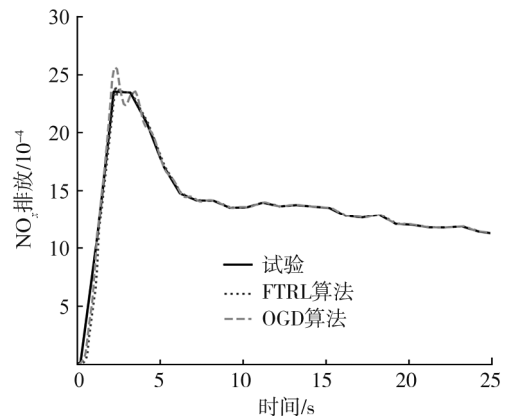


图 7 实时预测效果图

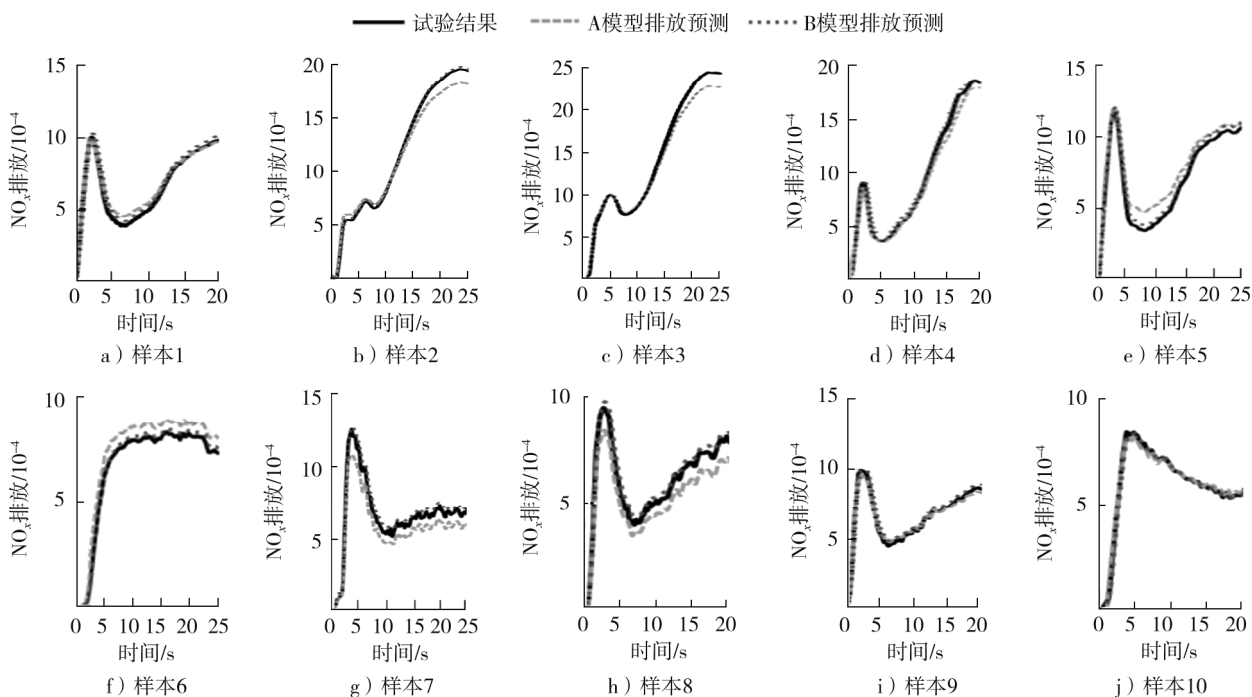


图8 不同模型在离线数据集上的预测效果

由图8及式(4)(5)可知:A、B模型的 R^2 分别为0.976 88和0.962 47, δ_R 分别为33.67和69.26,在线学习前、后模型的预测能力都很好,但在线学习后模型计算精度明显下降。主要原因为B模型采用数据集 O_1 训练,学习了大量高原环境下的 NO_x 排放规律,在预测数据集 L_3 时的精度有所下降,但由于合理的正则化设置,模型没有出现灾难性遗忘问题,对平原工况有较好的预测能力。在训练算力允许的条件下,可以适当对平原工况进行在线学习,使模型充分广泛的适用于各种环境。同时,在线学习在实际应用时可以根据当前误差随时开启与关闭:误差大时开启在线学习,误差小时关闭在线学习,从而提高预测精度,降低计算成本。

3.4 在线学习与全数据集学习对比分析

在线学习的一个重要目标是在不使用初次训练离线数据集的情况下,仅对新数据进行学习,即可使模型在不同工况下均有良好的预测效果。将在线数据集与离线数据集的训练集、验证集及测试集对应混合,构成全数据集,将全数据集离线训练得到的模型称为C模型。基于全数据集的测试集,B模型与C模型 NO_x 排放预测结果对比如图9所示。

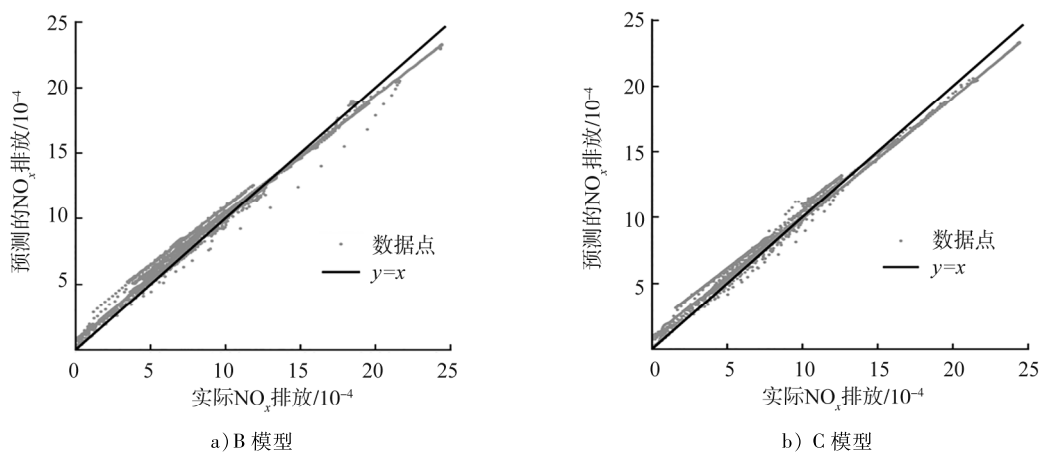


图9 不同模型在全数据集上的预测效果

由图9及式(4)(5)可知:B模型与C模型的 R^2 分别为0.962 81、0.965 63, δ_R 分别为65.04、57.91,2个模型预测精度大致相同。对比C模型,B模型在应用于未知工况时无需重新使用离线数据集所含的66组数据,数据量减少了68.7%。A、B、C 3个模型的训练时间分别为54.26、28.83、109.11 s,相比采用全数据集离线训练的模型C,在线学习后的自注意力机制模型B训练时间减少了73.6%。说明在线学习后的自注意力机制模型比其余模型的训练时间少,可以大幅减少重新训练和存储数据的成本,同时满足了在线预测需求。

4 结论

1)基于自注意力机制构建的模型,预测结果的决定系数和均方根误差分别为0.976 88与33.67,相较于BP神经网络模型,自注意力机制构建模型精度更高,可以更准确地预测 NO_x 排放。

2)FTRL算法预测结果比OGD算法更准确,因此更适用于起动过程 NO_x 排放预测模型。

3)通过在线学习可以使用离线训练集中未覆盖的工况样本进行 NO_x 排放预测模型更新,提高了模型的泛化性。

4)对比采用全数据集离线训练的模型,在线学习后的自注意力机制模型决定系数减小0.002 82,均方根误差增大了7.13,但训练所用数据量和训练时长分别减小了68.7%和73.6%;采用在线学习的自注意力机制模型大幅减少了训练与存储成本,更高效地完成排放预测需求。

参考文献:

- [1] WU C L, PEI Y Q, LIU C T, et al. Insights into the fusion correction algorithm for on-board NO_x sensor measurement results from heavy-duty diesel vehicles[J]. *Energies*, 2023, 16(16): 6082.
- [2] 钱枫,马骋,祝能,等.基于动态改进遗传粒子群-BP的重型车 NO_x 排放预测模型研究[J]. *车用发动机*, 2023(5): 63-71.
- [3] 王志红,董梦龙,张远军,等.基于PSO-SVR的重型柴油车 NO_x 排放预测[J]. *内燃机学报*, 2023, 41(6): 524-531.
- [4] FEBRIANA S B, QUACH N Y, KWANG C O, et al. NO_x emissions prediction in diesel engines: a deep neural network approach[J]. *Environmental Science and Pollution Research I*, 2023(29): 713-722.
- [5] TANG G Z, WANG S B, DU B C, et al. Study on pollutant emission characteristics of different types of diesel vehicles during actual road cold start[J]. *Science of the Total Environment*, 2022, 823: 153598.
- [6] 高松. 汽车发动机冷起动排放控制及应用研究[D]. 沈阳: 东北大学, 2015.
- [7] 苏岩,刘忠长,韩永强,等.进气温度对直喷式柴油机冷起动初始期燃烧和排放的影响[J]. *内燃机工程*, 2007(6): 28-32.
- [8] FANG L, LOU D M, HU Z Y, et al. The emission characteristics of a diesel engine during start-up process at different altitudes[J]. *Energies*, 2019, 12(18): 3556.
- [9] WEILENMANN M, SOLTIC P, SAXER C, et al. Regulated and nonregulated diesel and gasoline cold start emissions at different temperatures[J]. *Atmospheric Environment*, 2005, 39(13): 2433-2441.
- [10] SEO J G, YUN B, KIM J, et al. Development of a cold-start emission model for diesel vehicles using an artificial neural network trained with real-world driving data[J]. *Science of the Total Environment*, 2022, 806: 151347.
- [11] TANG D, ZHANG Z, HUA L, et al. Prediction of cold start emissions for hybrid electric vehicles based on genetic algorithms and neural networks[J]. *Journal of Cleaner Production*, 2023, 420: 138403.
- [12] 杨蓉,杨林,谭盛兰,等.基于遗传算法-优化长短期记忆神经网络的柴油机瞬态 NO_x 排放预测模型研究[J]. *内燃机工程*, 2022, 43(1): 10-17.
- [13] 李志杰. 面向大数据分析的多任务加速在线学习算法研究[D]. 武汉: 武汉大学, 2015.
- [14] 杨飞,卢保玲.基于主成分和贝叶斯正则化的 NO_x 排放量的预测[J]. *热力发电*, 2010, 39(2): 24-27.
- [15] MOLK DARAGH R R, JAFARMADAR S, KHALILARIA S, et al. Prediction of the performance and exhaust emissions of a compression ignition engine using a wavelet neural network with a stochastic gradient algorithm[J]. *Energy*, 2018, 142:

- 1128–1138.
- [16] 戴金池, 庞海龙, 俞妍, 等. 基于 LSTM 神经网络的柴油机 NO_x 排放预测[J]. 内燃机学报, 2020, 38(5): 457–463.
- [17] FEDERICO R, LUCA P, FRANCESCO M, et al. NARX technique to predict torque in internal combustion engines[J]. Information, 2023, 14(7): 417.
- [18] ZHANG Q, LIU Q, YE Q. An attention-based temporal convolutional network method for predicting remaining useful life of aero-engine[J]. Engineering Applications of Artificial Intelligence, 2024, 127: 107241.
- [19] JIANG J J, LI H, MAO Z W, et al. A digital twin auxiliary approach based on adaptive sparse attention network for diesel engine fault diagnosis[J]. Scientific Reports, 2022(12): 675.
- [20] YANG A N, LI D Q, LI G X. A fast adaptive online gradient descent algorithm in over-parameterized neural networks[J]. Neural Processing Letters, 2023(55): 4641–4659.
- [21] 常禧龙, 梁琨, 李文涛. 深度学习优化器进展综述[J]. 计算机工程与应用, 2024, 60(7): 1–12.
- [22] MCMAHAN H B. Follow-the-regularized-leader and mirror descent: equivalence theorems and L1 regularization[C]// Proceedings of 2011 International Conference on Artificial Intelligence and Statistics. Fort Lauderdale, USA: ICAIS, 2011: 525–533.
- [23] 李歌. 面向灾难性遗忘的小样本增量学习方法研究[D]. 北京: 中国矿业大学, 2023.

Prediction model for NO_x emission of starting process of the diesel engine based on online learning

DU Zhengyu¹, LI Huajie², GUO Zhikun³, LIANG Yongsen³, SHI Lei^{1*}

1. Key Laboratory for Power Machinery and Engineering of Ministry of Education, Shanghai Jiao Tong University, Shanghai 200240, China;

2. 92118 Troops of PLA, Zhoushan 316000, China; 3. China North Engine Research Institute, Tianjin 300400, China

Abstract: In order to build an accurate NO_x emission prediction model for diesel engine starting process, the transient operation characteristics of the starting process is analyzed, a nonlinear autoregressive neural network is used to select model features based on starting test data. The prediction performance of the self-attention mechanism model and the standard backpropagation neural network model for NO_x emissions is compared, it is found that the self-attention mechanism model has good prediction performance and small root mean square error. Comparing the predictive performance of self attention mechanism models in two online learning methods, namely online gradient descent algorithm and FTRL algorithm, it is found that adopting FTRL algorithm could improve the predictive performance of the model under unknown working conditions. Comparing the predictive performance of the self attention mechanism model before and after learning, it is found that the root mean square error of the learned model increases, but the predictive ability is good. Comparing the prediction performance of online learning models and full data training set offline learning models, it is found that the coefficient of determination and root mean square error of both models are not significantly different, and the prediction performance is good. However, the amount of data and training time used in online learning have been respectively reduced by 68.7% and 73.6%, significantly reducing storage and training costs. The results indicate that the self-attention mechanism model using FTRL algorithm could reduce data storage and training resource costs, and provide real-time feedback on prediction requirements.

Keywords: diesel engine; starting process; NO_x prediction; self attention mechanism; online learning

(责任编辑: 胡晓燕)