

利用 SE-GPR 模型对甲醇/柴油混合燃料 柴油机性能的预测

范金宇^{1,2}, 才正¹, 黄朝霞³, 杨晨曦¹, 李品芳^{1,2}, 黄加亮^{1,2}

(1. 集美大学轮机工程学院, 福建 厦门 361021; 2. 福建省船舶与海洋工程重点实验室, 福建 厦门 361021;
3. 集美大学理学院, 福建 厦门 361021)

[摘要] 为了对柴油机的经济性和排放参数进行高效、准确的预测, 根据 4190 型船用柴油机实验数据与边界参数, 建立 AVL-BOOST 甲醇/柴油混合燃料柴油机仿真模型; 利用模型进行仿真实验, 并建立甲醇掺混比、废气再循环 (exhaust gas recirculation, EGR) 率、喷油提前角和进气压力 4 个控制参数对有效油耗率和 NO_x 排放预测数据集; 利用该数据集对 5 种不同核函数的高斯过程回归 (Gaussian process regression, GPR) 模型进行训练; 最后将最优的平方指数高斯过程回归 (squared exponential-Gaussian process regression, SE-GPR) 模型、AVL-BOOST 仿真数据和柴油机实验数据进行对比。结果表明: 在数据量为 180 组时, SE-GPR 模型对有效油耗率和 NO_x 排放均取得拟合关联度 99% 以上, 均方根误差 (root mean square error, RMSE) 分别为 1.859, 0.344 5, 平均绝对误差 (mean absolute error, MAE) 分别为 0.954, 0.248 9; 并且, 相较于 AVL-BOOST 仿真实验, SE-GPR 模型对实验数据具有更好的拟合性。

[关键词] 船用柴油机; 甲醇; 高斯过程回归; 平方指数核函数; 性能预测

[中图分类号] U 664.121, TK 421

Performance Prediction of Methanol/Diesel Blended Diesel Engine Based on SE-GPR Model

FAN Jinyu^{1,2}, CAI Zheng¹, HUANG Zhaoxia³, YANG Chenxi¹, LI Pinfang^{1,2}, HUANG Jialiang^{1,2}

(1. School of Marine Engineering, Jimei University, Xiamen 361021, China;

2. Fujian Provincial Key Laboratory of Naval Architecture and Ocean Engineering, Xiamen 361021, China;

3. School of Science, Jimei University, Xiamen 361021, China)

Abstract: In order to efficiently and accurately predict diesel engine economy and emission parameters, based on the experimental data of the 4190 type marine diesel engine and boundary parameters, an AVL-BOOST simulation model for diesel engines utilizing methanol/diesel blended fuels was established, and a dataset for predicting effective fuel consumption and NO_x emissions was created by using this model, incorporating four operational parameters: methanol blending ratio, exhaust gas recirculation (EGR) rate, injection advance angle, and intake pressure. The dataset was employed to train Gaussian process regression (GPR) models with five different kernel functions. Finally, the best-performing squared exponential Gaussian process regression

[收稿日期] 2023-11-13

[基金项目] 福建省自然科学基金项目 (2022J01812, 2021J01849); 福建省教育厅项目 (JAT210237)

[作者简介] 第一作者: 范金宇, 副教授, 从事船舶柴油机性能优化与故障分析研究。通信作者: 黄加亮, 教授, 博士, 从事现代轮机管理工程研究。E-mail: jluhuang@jmu.edu.cn

http://xuebaobangong.jmu.edu.cn/zkb

(SE-GPR) model was compared with AVL-BOOST simulation data and diesel engine experimental data. The results showed that the SE-GPR model achieves a correlation of over 99% for both effective fuel consumption and NO_x emissions when the dataset contains 180 data sets, with root mean square error (RMSE) values of 1.859, 0.344 5, and mean absolute error (MAE) values of 0.954, 0.248 9. Moreover, compared to AVL-BOOST simulation experiments, the SE-GPR model exhibits a better fit to the experimental data.

Keywords: marine diesel engine; methanol; Gaussian process regression; squared exponential kernel; performance prediction

0 引言

甲醇燃料作为一种可再生能源,近年来引起了人们的广泛关注。甲醇具有清洁燃料的多种优点,包括低碳排放和资源可持续性。与传统的石油燃料相比,甲醇燃料的使用可以显著减少 NO_x 、 CO_2 等有害气体的排放,同时减少细颗粒物和硫化物的产生^[1]。

为了应对排放法规日益严苛的要求,同时提升柴油机的经济性,研究人员一直在寻求新的方法和技术。然而,燃料掺混比、EGR 率、进气压力等船舶柴油机的相关参数,往往具有参数众多、交互作用复杂、非线性相关等特征^[2]。线性回归这类传统的统计学方法只适用于多个参数与单个性能的较为简单的作用关系,而对于复杂作用关系的建模效果有限^[3]。因此,人们引入机器学习作为预测柴油机性能的一种工具。许朵等^[4]首先基于长短期记忆神经网络,建立了发动机 NO_x 排放和燃油消耗率的预测模型,然后结合 NSGA-II 算法,寻找最优控制参数,以达到优化 NO_x 排放、降低油耗率的目的。Liu 等^[5]使用一种改进的支持向量机模型,用于瞬态和稳态时 NO_x 排放质量分数的预测,结果表明,该模型比传统 BP 神经网络、支持向量机模型具有更高的拟合性。但上述研究缺乏在小数据量样本条件下的分析,并且稳态下的 NO_x 排放、油耗率与时间序列之间缺少必然的直接联系^[4],使得长短期记忆预测模型在建立时就受到时间序列限制。Javad 等^[6]在小数据量样本下引入高斯过程回归(Gaussian process regression, GPR)预测模型,对双燃料柴油机的动力性和经济性进行有效预测,结果表明,GPR 预测方法因使用灵活、可描述不确定性及先验性等优势,在柴油机性能参数预测上可以弥补样本的各种缺陷^[7]。

由于 GPR 是一种依靠核函数进行预测的方法,其预测效果主要由核函数所决定。但目前相关研究只选择了其中一种核函数对柴油机参数预测并进行测试,缺少多种不同核函数在柴油机性能预测之间的对比研究。为解决上述问题,本文分别采用 Mat3/2、Mat5/2、Rational Quadratic、Squared Exponential、Linear 5 种核函数,在柴油机性能参数预测效果上与 BP、SVM 2 种常用方法进行对比研究,为后续在满足排放法规的条件下寻找最优经济性提供了理论依据。

1 甲醇/柴油混合燃料柴油机仿真模型的建立和验证

以济南柴油机厂生产的 4190Z_LC-2 型船用四缸中速柴油机为研究对象。实验在额定工况下展开,根据柴油机结构参数和边界条件,建立 AVL-BOOST 柴油机仿真模型,在仿真模型中调节控制参数,获取仿真实验数据。将得到的仿真数据代入 Matlab 软件建立的 GPR 模型,进一步验证不同核函数的 GPR 模型在预测有效油耗率(制动燃料油耗率)(brake specific fuel consumption, BSFC)、 NO_x 排放的相关性和准确性。

1.1 实验台架搭建

4190 型甲醇/柴油混合燃料柴油机技术参数如表 1 所示。4190 型船用四冲程实验台架采用高压共轨燃油喷射系统,以电控系统(ECU)控制电磁阀启闭,进而控制如喷油提前角等关键参数,从而控制喷油器的喷油规律;甲醇掺混比以体积分数掺混进柴油,再由喷油器喷入气缸。为实现甲醇柴油混合燃料在大体积比例下充分溶合,本实验加入体积分数为 0.3% 的正庚醇,并充分搅拌^[8-9]。

表 1 4190 柴油机技术参数和边界条件
Tab.1 Technical parameters and boundary conditions of 4190 diesel engine

参数	数值	参数	数值	参数	数值
缸数	4	标定转速/($r \cdot \min^{-1}$)	1 000	喷孔数	8
标定功率/kW	220	燃烧室形状	直口 ω 型	排气阀开启角/($^{\circ}$)	58
总排量/L	13.82	压缩比 ε	14:1	排气阀关闭角/($^{\circ}$)	56
气缸直径 \times 行程/ (mm \times mm)	190 \times 210	额定油耗率/ ($g \cdot kW^{-1} \cdot h^{-1}$)	218	进气阀开始角/($^{\circ}$)	66
发火顺序	1-3-4-2	循环喷油量/g	0.394 88	进气阀关闭角/($^{\circ}$)	54

利用 HORIBA MEXA-1600DSEGR 型排气分析仪和 AVL 烟度分析仪测量排气中 NO_x 、CO、THC、Soot 等成分；利用 KISTLER 公司的燃烧分析仪测量缸内压力、温度、放热率等参数；利用 FC2010 型智能油耗仪测量消耗的燃料量^[10]。同时，以 EGR 阀的开度控制 EGR 率的大小。

1.2 控制参数的选择

国际海事组织（IMO）制定的 Tier III 排放标准对船舶 NO_x 排放提出了更加严格的要求，因此，选取 NO_x 和有效油耗率作为柴油机排放参数和柴油机经济性的评价指标。

在 AVL-BOOST 软件中进行仿真实验时，为确保模拟完整的燃烧过程，同时减少排放，且可以兼顾柴油机经济性^[11-13]，本文设置了甲醇掺混比、EGR 率、喷油提前角、进气压力 4 个控制参数，并设置了甲醇掺混比为 0% ~ 20%、EGR 率为 0% ~ 13.5%、喷油提前角为 15.6° ~ 24.6°、进气压力为 0.170 ~ 0.240 MPa。

1.3 AVL-BOOST 仿真模型的建立

AVL-BOOST 仿真模型建立流程如图 1 所示。本文将柴油机看成一个由气缸套、活塞顶面和气缸盖组成的热力学系统。根据柴油机工作过程的物理模型，建立 AVL-BOOST 仿真模型，然后根据柴油机具体结构，输入边界条件和相关参数，最后将仿真数据与柴油机实验数据进行对比验证，直到误差在可接受的范围内。

1.3.1 模型选择

由于本文选用燃料为甲醇/柴油混合燃料，因此选用充分考虑燃料预混合和扩散燃烧的 AVL MCC 燃烧模型；选取 Woschni 1978 作为传热模型；废气涡轮增压器模型采用 AVL-BOOST 系统自带的简化模型，依据压气机与涡轮机在每循环的效率来计算涡轮增压器的能量平衡，这种模型的优点是只需要少量的限制参数就可以表现出涡轮增压器的特性曲线。

1.3.2 混合燃料仿真模型的建立和仿真实验

AVL-BOOST 仿真模型如图 2 所示，在模型建立后要对各种元件边界参数进行设定。主要参数和边界条件如表 1 所示。

在仿真实验中，运用 AVL-BOOST 对甲醇掺混比、EGR 率、喷油提前角、进气压力 4 个控制

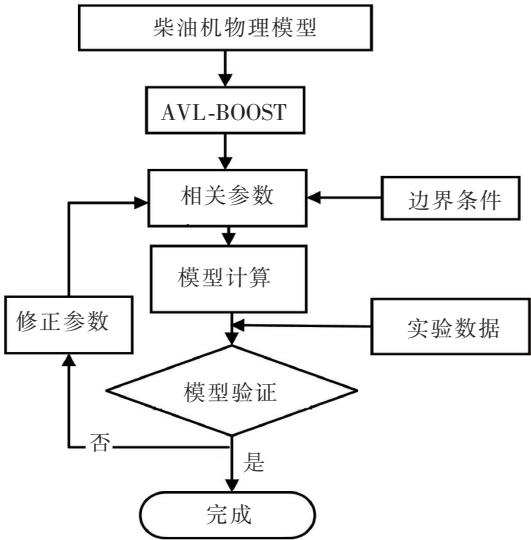


图 1 AVL-BOOST 模型建立流程图
Fig.1 Flowchart of AVL-BOOST model establishment

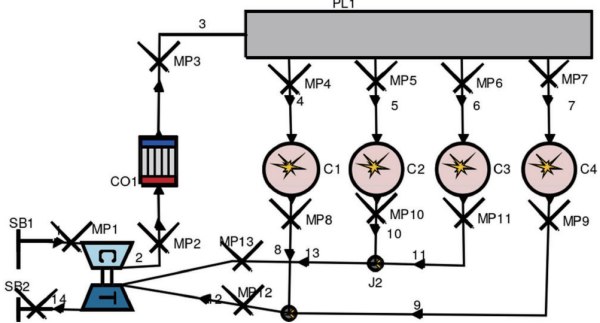


图 2 AVL-BOOST 仿真模型
Fig.2 AVL-BOOST simulation model

参数进行设置。由于甲醇缸内反应极其复杂, 对仿真计算要求较高且计算量较大, 故本文采取 Chang 等^[14]的甲醇/柴油混合燃料的化学反应机理进行研究。在 EGR 率的设置中, 设置方式为返回气体在总混合物 (新鲜空气和进气燃料) 中的比例^[15], 计算公式为

$$\text{EGR}/\% = [m_{\text{EGR}}/(m_{\text{air}} + m_{\text{fuel}} + m_{\text{EGR}})] \times 100。$$

(1)

其中: m_{EGR} 、 m_{air} 、 m_{fuel} 分别为返回气体质量、新鲜空气质量和进气燃料质量。

1.4 AVL-BOOST 仿真实验

由于船用柴油机的缸内平均压力、热效率是有效燃油消耗率和 NO_x 排放的主要影响因素^[16], 因此选取缸内压力和放热率 2 个参数对仿真实验准确性进行验证。

图 3 为额定工况下仿真结果与原机实验结果的对比曲线, 其中: B0 的甲醇掺混比为 0, EGR 率为 0; B10 的甲醇掺混比为 10%, EGR 率为 5%。由图 3 可见, B0、B10 压力曲线和放热率曲线的仿真值均略高于实验值, 这是由于模型未考虑实际中存在的气缸漏气及热损失; 气缸内气体未充分均匀混合, 与实际气体的混合情况存在误差。但由图 3 还可以看出, 误差在 3% 以内, 两条曲线变化趋势基本一致, 因此认为仿真模型可以模拟实机实验。

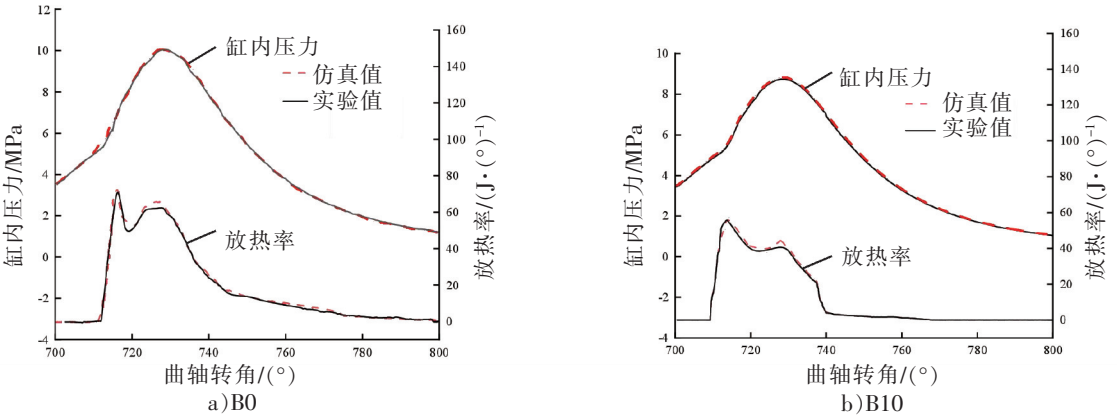


图 3 B0、B10 的缸压和放热率曲线的仿真值与实验值对比
Fig.3 Comparison of simulated and experimental values of cylinder pressure and heat release rate curves of B0 and B10

2 多种核函数的 GPR

GPR 是一种基于统计学原理和贝叶斯理论框架的机器学习方法, 它基于高斯过程的概念, 将目标变量视为联合高斯分布, 对观测数据进行学习, 对未观测数据进行预测, 并提供预测的不确定性估计^[17], 最后给出最大概率出现的目标值。

GPR 流程如图 4 所示, 其中: \mathbf{X} 为甲醇掺混比、EGR 率、喷油提前角、进气压力 (训练输入集); y_1 、 y_2 分别对应有效油耗率和 NO_x 排放质量分数 (训练输出集); \mathbf{X}^* 为预测输入值; \mathbf{Y}^* 为预测输出值; \mathbf{X}' 为已知数据点 (训练集中随机生成)。主要目标就是通过给定的 \mathbf{X} 、 \mathbf{Y} 和不同的核函数, 训练出适配于该模型的映射函数 f , 从而求解出测试输出 \mathbf{Y}^* 。

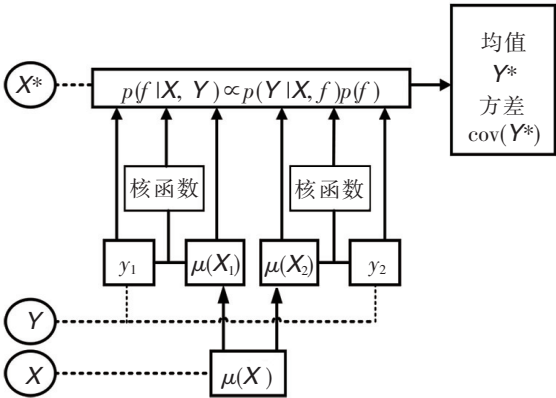


图 4 GPR 流程图
Fig.4 Flow chart of GPR

GPR 模型可以假设为 $\mathbf{Y} = f(\mathbf{X}) + \varepsilon$, 其中: ε 为数据噪声, $\varepsilon \in N(0, \sigma^2)$, σ^2 为噪声方差。GPR 由均值函数和协方差函数 (核函数) 来确定, 计算公式分别为

$$\mu(\mathbf{X}) = E[f(\mathbf{X})], \quad (2)$$

$$k(\mathbf{X}, \mathbf{X}') = E[(f(\mathbf{X}) - \mu(\mathbf{X}))(f(\mathbf{X}') - \mu(\mathbf{X}'))^T]. \quad (3)$$

根据输入集 \mathbf{X} 和 $f(\mathbf{X})$, 得到函数的先验分布, 其计算公式为

$$f(\mathbf{X}) | \mathbf{X} \sim N(\mu(\mathbf{X}), k(\mathbf{X}, \mathbf{X})). \quad (4)$$

以函数集 $f(x)$ 和输出集 \mathbf{Y} 得到函数值域先验分布, 其计算公式为

$$\mathbf{Y} | \mu(\mathbf{X}) \sim N(\mu(\mathbf{X}), \sigma^2). \quad (5)$$

经过训练, 高斯过程的形式可以写成

$$f(x) \sim GP(\mu(\mathbf{X}), k(\mathbf{X}, \mathbf{X}')). \quad (6)$$

此时, 根据贝叶斯定理, 以训练集样本 (\mathbf{X}, \mathbf{Y}) 可以求解出 f 函数的后验分布, 其计算公式为

$$p(f | \mathbf{X}, \mathbf{Y}) \propto p(\mathbf{Y} | \mathbf{X}, f) p(f). \quad (7)$$

一般来说, 均值函数初始设置为 0, 协方差由核函数计算, 由不同核函数来度量输入特征 \mathbf{X} 点和 \mathbf{X}' 点之间的相似性^[18]。本文选用的 5 种核函数分别为 Mat3/2、Mat5/2、Rational Quadratic、Squared Exponential 和 Linear, 它们的计算公式分别为

$$k(\mathbf{X}, \mathbf{X}') = \sigma_f^2 (1 + \sqrt{3}r/l) \exp(-\sqrt{3}r/l), \quad (8)$$

$$k(\mathbf{X}, \mathbf{X}') = \sigma_f^2 (1 + \sqrt{5}r/l + 5r^2/(3l^2)) \exp(-\sqrt{5}r/l), \quad (9)$$

$$k(\mathbf{X}, \mathbf{X}') = \sigma_f^2 (1 + \|\mathbf{X} - \mathbf{X}'\|^2/(2\alpha l^2))^{-\alpha}, \quad (10)$$

$$k(\mathbf{X}, \mathbf{X}') = \sigma_f^2 \exp(-1/2(\mathbf{X} - \mathbf{X}')^T L^{-2}(\mathbf{X} - \mathbf{X}')), \quad (11)$$

$$k(\mathbf{X}, \mathbf{X}') = \sigma_f^2 (\mathbf{X}^T \mathbf{X}'). \quad (12)$$

其中: $r = \sqrt{(\mathbf{X} - \mathbf{X}')^T (\mathbf{X} - \mathbf{X}')}$ 是样本间的欧氏距离; σ_f 、 l 、 α 均为核函数超参数, 一般统称为 θ , 可根据训练样本在 GPR 模型中进行训练再进行确定。

基于该 f 函数后验分布, 新的数据集 (测试集) 中测试集输出值的 \mathbf{Y}^* 和训练集的 \mathbf{Y} 也服从联合高斯分布表达式

$$\begin{bmatrix} \mathbf{Y} \\ \mathbf{Y}^* \end{bmatrix} \sim N \left(0, \begin{bmatrix} \mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma^2 & \cdots & \mathbf{K}(\mathbf{X}, \mathbf{X}^*) \\ \vdots & & \vdots \\ \mathbf{K}(\mathbf{X}, \mathbf{X}^*) & \cdots & \mathbf{K}(\mathbf{X}^*, \mathbf{X}^*) \end{bmatrix} \right). \quad (13)$$

其中: $\mathbf{K}(\mathbf{X}^*, \mathbf{X}) = \mathbf{K}(\mathbf{X}, \mathbf{X}^*)$ 为预测集的协方差函数矩阵; $\mathbf{K}(\mathbf{X}^*, \mathbf{X}^*)$ 为预测集方差阵。则预测值的条件概率满足

$$p(\mathbf{Y}^* | \mathbf{X}^*, \mathbf{X}, \mathbf{Y}) = N(\mathbf{Y}^*, \text{cov}(\mathbf{Y}^*)). \quad (14)$$

因此, 柴油机性能参数的预测值 (均值) 和误差区间 (核函数矩阵) 可估计为

$$\mathbf{Y}^* = \mathbf{K}(\mathbf{X}^*, \mathbf{X}) [\mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma I^N]^{-1} \mathbf{Y}, \quad (15)$$

$$\text{cov}(\mathbf{Y}^*) = \mathbf{K}(\mathbf{X}^*, \mathbf{X}^*) - \mathbf{K}(\mathbf{X}^*, \mathbf{X}) [\mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma I^N]^{-1} \mathbf{K}(\mathbf{X}, \mathbf{X}^*)^T. \quad (16)$$

3 模型预测的结果与验证

为了验证 GPR 模型中不同核函数对柴油机排放和经济性预测的准确性, 本文将 AVL-BOOST 仿真实验数据进行整理, 代入 GPR 预测模型, 优化相关参数。此外, 为了证明 GPR 模型相较于 BP 神经网络和 SVM 支持向量机回归预测模型在柴油机性能预测方面的优势, 本文同时做了进一步的对比分析。

3.1 数据集建立

在区间内随机选取数据, 建立具有 200 组数据的数据集。运用 Matlab 将数据集划分为训练集和测试集, 其中: 75% 的数据 (150 组数据) 为训练集, 用于对均值函数、核函数和超参数的设定; 剩下 25% 的数据 (50 组数据) 为测试集, 用于后续验证。为防止实验过拟合, 采用 4 倍交叉验证的方

式^[19]将训练集随机分成 4 份, 轮流将其中的 3 份作训练集, 1 份作测试集, 并分别反复进行 5 次共 20 次模型测试。并在 Matlab 中加装计时模块, 以便对模型的计算时长进行分析。

3.2 评价指标

本文评价模型预测准确性的指标包括拟合关联度 (R)、平均绝对误差 (MAE) 和均方根误差 (RMSE), 它们的计算公式分别为

$$R = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}, \tag{17}$$

$$\text{MAE} = (1/n) \sum_{i=1}^n |X_i - Y_i|, \tag{18}$$

$$\text{RMSE} = \sqrt{(1/n) \sum_{i=1}^n (X_i - Y_i)^2}。 \tag{19}$$

其中: X_i 和 Y_i 分别为测试集的仿真实验值和模型预测值; \bar{X} 和 \bar{Y} 为两组数据的平均值。

3.3 预测结果

表 2 为 GPR 模型中 5 种核函数, 以及 BP、SVM 2 种模型在有效油耗率和 NO_x 排放质量分数预测方面的拟合关联度。为确保实验的可靠性, 采用了 4 倍交叉验证的方法进行了 20 次重复实验, 因此, 表 2 中的拟合关联度以均值 \pm 标准差的形式呈现, 其中, 均值越接近 1, 表示模型对柴油机性能参数预测的拟合性好; 而标准差越接近 0, 表示模型对新区间数据的预测能力越强。表 3 为预测模型的 RMSE 和 MAE 结果。

表 2 预测模型的拟合关联度

表 3 预测模型的 RMSE 和 MAE

Tab.2 Fit correlation degree from model prediction

Tab.3 RMSE and MAE from model prediction

模型	R	
	有效油耗率	NO_x
Mat3/2	0.993 \pm 0.01	0.992 \pm 0.00
Mat5/2	0.989 \pm 0.00	0.994 \pm 0.00
Linear	0.982 \pm 0.02	0.959 \pm 0.03
SE	0.994 \pm 0.00	0.996 \pm 0.00
RQ	0.993 \pm 0.00	0.992 \pm 0.01
BP	0.916 \pm 0.13	0.926 \pm 0.19
SVM	0.984 \pm 0.01	0.983 \pm 0.00

模型	RMSE		MAE	
	有效油耗率	NO_x	有效油耗率	NO_x
Mat3/2	1.979	0.453 0	0.984	0.298 0
Mat5/2	2.321	0.388 5	1.145	0.258 1
Linear	3.137	1.010 7	1.837	0.819 1
SE	1.804	0.344 5	0.919	0.248 9
RQ	2.088	0.431 8	1.164	0.276 2
BP	24.317	0.555 8	6.650	0.422 8
SVM	2.827	0.406 9	1.153	0.243 9

由表 2 可见, 在 GPR 预测模型中, 除 Lin 核函数以外, 其余 4 个核函数的拟合关联度的标准差均大于 0.989。这是因为 GPR 模型具有非参数模型的特性, 没有预先固定的模型结构, 能够自适应地从数据中学习模型的复杂度, 对解决柴油机性能参数预测影响因素多、非线性、数据总量小和噪声大的问题有良好的适配性。由表 2 还可以看出, GPR 模型的 SE 核函数对于柴油机性能参数的拟合关联度最接近 1, 且平均值大于 0.99, 说明 SE-GPR 模型对柴油机性能参数具有 99% 的可信度。

由表 3 可见, SE 核函数对于有效油耗率和 NO_x 排放质量分数的 RMSE 分别为 1.804、0.344 5, MAE 分别为 0.919、0.248 9, 相较于其他函数具有更小的误差。这是因为, 在 SE-GPR 模型中, GPR 在贝叶斯推断框架下进行建模, 通过建立目标变量的先验分布和观测数据之间的条件概率关系, GPR 能够提供对预测结果的全概率估计; SE 核函数的超参数用于控制函数的光滑程度和相似性的范围^[20], 该参数可以根据数据的特征自适应地调整尺度, 从而使得核函数能够更好地适应不

同尺度的数据。因此，SE-GPR 模型在预测不同尺度数据时具有高关联度、误差小的特性，但与其他计算方法相比，SE-GPR 模型的计算时长相差不大（均为 0.3 ~ 0.4 s），因此在计算时间方面不做进一步研究。

此外，为验证 SE-GPR 模型对柴油机性能参数预测的准确性，还必须将 GPR 模型预测数据与 AVL-BOOST 仿真实验数据做进一步对比。图 5 为 SE-GPR 模型对柴油机性能参数预测值与仿真实验值之间的对比图，由图 5 可以看出，GPR 模型对于柴油机性能参数预测值基本位于 95% 预测置信区间以内。另外，GPR 模型的预测数据值都在 45°线（最优目标值： $y = x$ ）附近。基于此结果可知，SE-GPR 模型的预测结果与 AVL-BOOST 的仿真值差别很小。

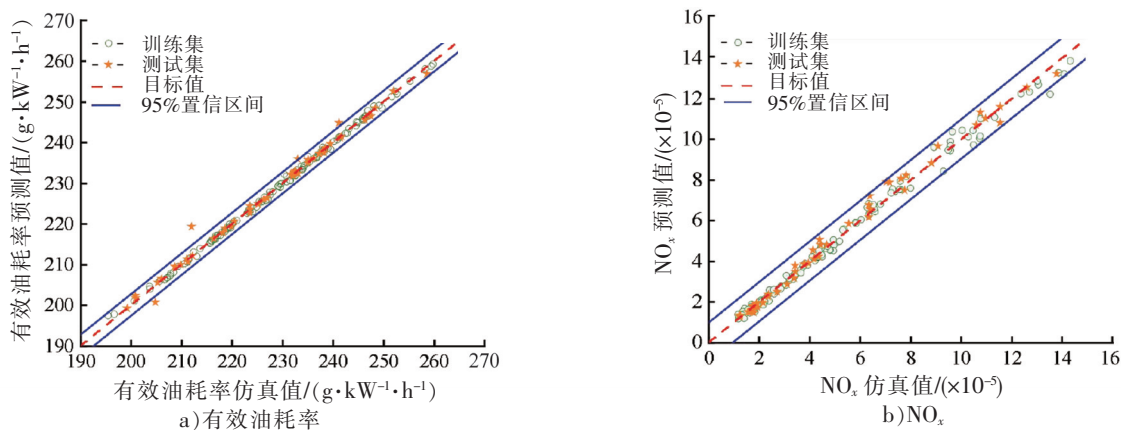


图 5 AVL-BOOST 仿真值与 SE-GPR 预测值对比图(w (甲醇)=0%~20%)
Fig.5 Comparisons of the results from AVL-BOOST simulations and SE-GPR prediction (w (methanol)=0%~20%)

3.4 预测区间拓宽验证和配机验证

为进一步验证 SE-GPR 模型对甲醇掺混比大于 20% 的数据能否进行准确预测，将甲醇掺混比改为 20% ~ 40%，其余不做改变，共进行了 60 组补充实验。AVL-BOOST 仿真实验值与 SE-GPR 模型预测值对比结果如图 6 所示。由图 6 可以看出，虽然 SE-GPR 模型在新的区间内选取输入值，但依然与 AVL-BOOST 的仿真值具有较好的拟合性，并且数据均位于目标值附近。说明 SE-GPR 模型对新数据有较强的适应能力，可以对高甲醇掺混比的新数据进行有效预测。

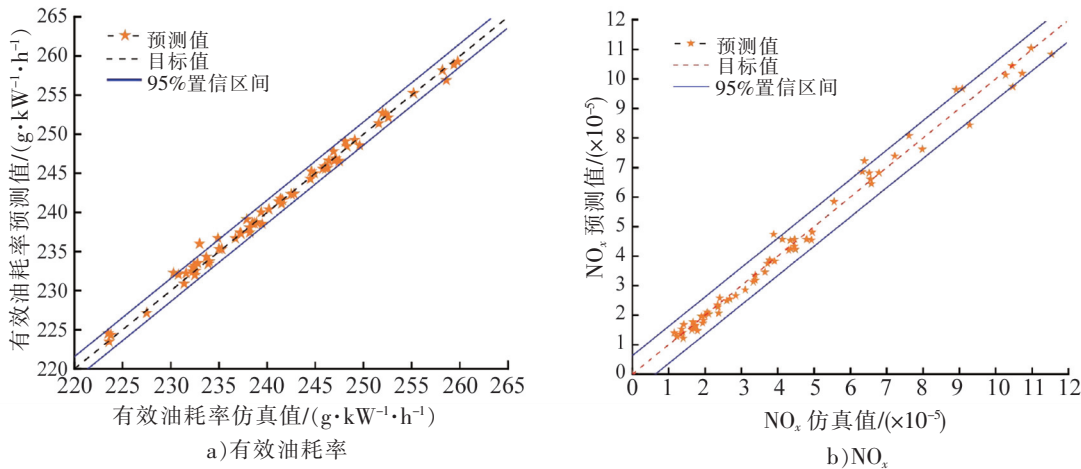
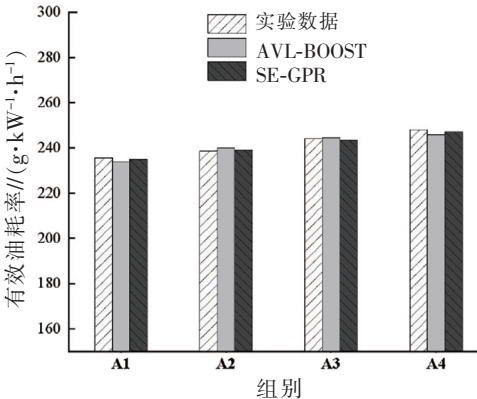


图 6 AVL-BOOST 仿真值与 SE-GPR 预测值对比图(w (甲醇)=20%~40%)
Fig.6 Comparisons of the result from AVL-BOOST simulations and SE-GPR prediction (w (methanol)=20%~40%)

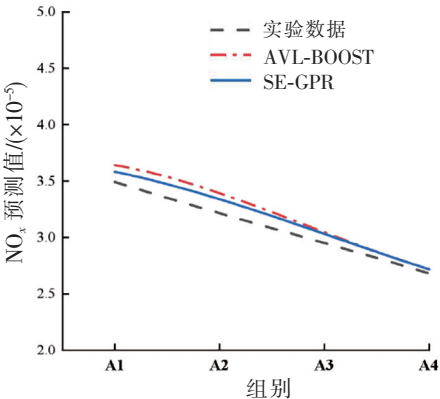
为验证 SE-GPR 模型对实机实验的准确性, 进行如表 4 所示的 8 组对比实验, 结果如图 7 所示。由图 7 可见, AVL-BOOST 仿真试验数据与原机数据的油耗率有误差, 特别在 B1 和 B2 两组试验中, 仿真值小于原机试验值。这是由于较小的喷油提前角度, 雾化质量对燃烧效果产生更大的影响。但由于喷油孔磨损直径变大, 雾化质量降低, 导致油耗率略有升高。但 SE-GPR 模型依靠自身的不确定性估计, 所得预测数据与 AVL-BOOST 仿真实验数据相比, 具有更好的拟合性。

表 4 控制参数设置

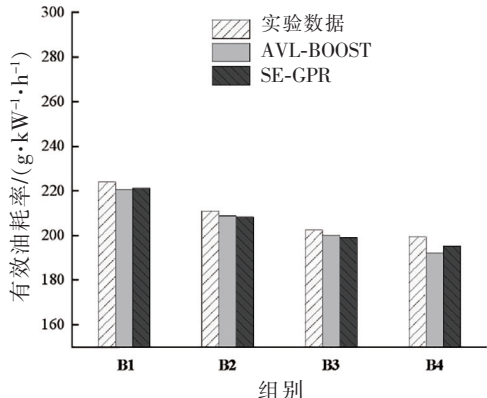
Tab.4 Settings of controlling parameters				
组号	甲醇掺混 比/%	EGR 率 /%	喷油提前 角/(°)	进气压力 /MPa
A1	22	5.0	18.6	0.193
A2	24	7.5	18.6	0.193
A3	26	10.0	18.6	0.193
A4	30	12.5	18.6	0.193
B1	0	0	16.6	0.173
B2	0	0	18.6	0.193
B3	0	0	20.6	0.213
B4	0	0	22.6	0.233



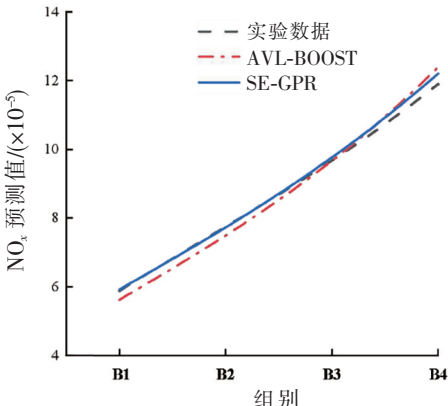
a)有效油耗率(A1~A4)



b)NO_x 预测值(A1~A4)



c)有效油耗率(B1~B4)



d)NO_x 预测值(B1~B4)

图 7 实验数据与 AVL-BOOST、SE-GPR 预测数据对比

Fig.7 Comparisons of the experimental data and those from AVL-BOOST and SE-GPR predictions

3.5 数据集数目确定

为进一步将只适配于额定工况的 SE-GPR 模型推广到全工况, 需减少每个工况点 AVL-BOOST 的仿真次数, 以节约数据集建立的时间。因此, 将之前的数据集进行划分, 分别取其中的 60%、75%、90% (120、150、180) 组数据。如果小样本量的测试结果拟合关联度大于 0.99, 即可认为只进行较少的仿真实验, 就可以达到 SE-GPR 模型对柴油机性能参数的准确预测。

SE-GPR 模型、BP 神经网络和 SVM 支持向量机神经网络分别取数据集 60%、75%、90% 的 R 、RMSE 和 MAE 结果如表 5 所示。由表 5 可知, 在 SE-GPR 模型取 90% 仿真实验数据进行柴油机性能参数有效油耗率和 NO_x 排放质量分数预测时, R 值分别为 0.992, 0.996, 明显高于同样数量的数据集

在 BP 神经网络和 SVM 神经网络训练得出的 R 值；并且，SE-GPR 模型在取 90% 仿真实验数据值时，预测有效油耗率和 NO_x 排放质量分数时的 RMSE 分别为 1.859、0.344 5，MAE 分别为 0.954、0.248 9。由于 SE-GPR 模型采用了概率推断的方法进行建模，可以对每个预测结果提供置信度和方差的信息。在数据量较少的情况下，SE-GPR 模型能够给出预测结果的点估计和不确定性评估，这使得 SE-GPR 模型更适用于处理数据噪声和不确定性较大的情况。因此，当数据集的数量为 90%（180 组数据）时，SE-GPR 模型的预测结果可信用度高，误差小。

表 5 不同数量数据集的拟合度和误差对比
Tab.5 Comparisons of fitting and error of different data sets

模型	数据集	R		RMSE		MAE	
		有效油耗率	NO_x	有效油耗率	NO_x	有效油耗率	NO_x
SE-GPR	60%	0.983	0.993	2.796	0.417 3	1.437	0.287 5
	75%	0.987	0.994	2.522	0.359 7	1.309	0.259 1
	90%	0.992	0.996	1.859	0.344 5	0.954	0.248 9
BP	60%	0.902	0.917	33.571	0.835 1	7.721	1.252 7
	75%	0.910	0.926	24.607	0.557 3	7.295	0.742 5
	90%	0.916	0.926	24.364	0.555 8	6.675	0.422 8
SVM	60%	0.971	0.981	7.188	0.406 9	2.619	0.340 4
	75%	0.979	0.983	5.407	0.434 4	1.762	0.300 0
	90%	0.980	0.983	2.821	0.467 8	1.102	0.243 9

4 结论

利用 AVL-BOOST 软件建立甲醇/柴油混合燃料柴油机仿真模型，完成模型修正和验证，通过设置 4 个柴油机控制参数，对有效油耗率和 NO_x 排放质量分数进行 200 组仿真计算。并对比 5 种不同核函数的 GPR 模型、BP 神经网络和 SVM 支持向量机对柴油机性能参数的预测性能，得到：1）与常用的机器学习方法比较，GPR 模型作为柴油机性能参数预测模型，其拟合关联度标准差均接近于 0，具有良好的泛化性；2）相较于其他核函数模型及其他常用机器学习方法，SE-GPR 模型的拟合关联度均大于 0.99，对于有效油耗率和 NO_x 排放质量分数的 RMSE 分别为 1.859、0.344 5，MAE 分别为 0.954、0.248 9，可对柴油机性能参数进行准确预测，同时，SE-GPR 比 AVL-BOOST 对于实验数据具有更优的拟合性；3）相较 BP 神经网络和 SVM 支持向量机，在总仿真数据集不少于 180 组时，SE-GPR 模型对 4 组性能参数的预测的拟合关联度均高于 0.99，预测有效油耗率和 NO_x 排放质量分数的 RMSE 分别为 1.804、0.344 5，MAE 分别为 0.919、0.248 9，明显小于其他核函数、BP 神经网络和 SVM 支持向量机预测方法，可节约数据集建立时间。

[参考文献]

[1] 黄加亮, 张飞飞, 林航, 等. 双卷流燃烧系统的双燃料发动机燃烧排放特性[J]. 集美大学学报(自然科学版), 2022, 27(6): 531-538.

[2] 杨柏枫, 王玉国, 殷长春, 等. 基于模糊分析的电控柴油机掺烧丁醇性能优化[J]. 大连海事大学学报, 2021, 47(3): 86-95, 119.

[3] 胡登, 邓涛, 黄加亮, 等. 基于正交-主元分析的柴油机燃烧参数匹配优化[J]. 船舶工程, 2020, 42(4): 53-60.

[4] 许朵, 姚崇, 马骏, 等. 船用双燃料发动机综合性能预测及优化[J]. 内燃机学报, 2022, 40(5): 403-411.

[5] LIU B, HU J, YAN F W, et al. A novel optimal support vector machine ensemble model for NO_x emissions prediction of a

- diesel engine[J]. Measurement,2016,92:183-192.
- [6] JAVAD Z, ABBAS R, JOSÉ R N A. The effect of EGR and hydrogen addition to natural gas on performance and exhaust emissions in a diesel engine by AVL fire multi-domain simulation, GPR model, and multi-objective genetic algorithm[J]. International Journal of Hydrogen Energy,2022,47(50):21565-21581.
- [7] 姚小俊, 吴迪. 基于贝叶斯优化高斯过程回归法的再生混凝土力学性能预测[J]. 科学技术与工程,2023,23(7):2968-2975.
- [8] 王晓玲, 沈博, 黄蕾, 等. 醇类助溶剂对甲醇与柴油混合燃料互溶性的影响[J]. 西安石油大学学报(自然科学版), 2017,32(4):98-100,108.
- [9] YUSAF T F, HAMAWAND I, BAKER P, et al. The effect of methanol-diesel blended ratio on CI engine performance[J]. International Journal of Automotive and Mechanical Engineering,2013,8(1):1385-1395.
- [10] 杨捷波, 王彬彬, 叶子泉, 等. 新型相继增压系统对船用柴油机性能影响的研究[J]. 内燃机工程,2021,42(6):52-61.
- [11] 孟雨航, 孙楠楠, 张明泽, 等. EGR 和稀燃对甲醇发动机燃烧及排放特性的影响[J]. 车用发动机,2023(6):10-18, 83.
- [12] 宣熔, 牛梦达, 黄朝霞, 等. EGR 及进气压力对双燃料发动机燃烧和排放的影响[J]. 船舶工程,2021,43(10):76-81, 85.
- [13] 宣熔, 张飞飞, 陈坚, 等. 喷油提前角对双燃料柴油机燃烧和排放的影响[J]. 集美大学学报(自然科学版),2021,26(3):240-245.
- [14] CHANG Y C, JIA M, LI Y P, et al. Application of the optimized decoupling methodology for the construction of a skeletal primary reference fuel mechanism focusing on engine-relevant conditions[J]. Frontiers in Mechanical Engineering,2015,1:00011.
- [15] SOLOIU V, DUGGAN M, HARP S. PFI (port fuel injection) of *N*-butanol and direct injection of biodiesel to attain LTC (low-temperature combustion) for low-emissions idling in a compression engine[J]. Energy,2013,52(1):143-154.
- [16] JI C, HUANG Z, HUANG Y, et al. Performance and emissions characteristics of a gasoline engine with direct injection and spark assisted compression ignition at different engine operating conditions[J]. Energy Conversion and Management,2016,120:289-301.
- [17] 何志昆, 刘光斌, 赵曦晶, 等. 高斯过程回归方法综述[J]. 控制与决策,2013,28(8):1121-1129,1137.
- [18] JÄKEL F, SCHÖLKOPF B, WICHMANN F A. A tutorial on kernel methods for categorization[J]. Journal of Mathematical Psychology,2007,51(6):343-35.
- [19] 崔树银, 汪昕杰. 基于特征工程的集成学习短期光伏功率预测[J]. 科学技术与工程,2022,22(2):532-539.

(责任编辑 马建华 英文审校 郑青榕)