

基于 SISO 算法的混凝土细观模型压缩行为分析*

白宇飞^{1,2}, 张新宇^{1,2}, 亓晓鹏^{1,2}, 张煜航³, 王志勇^{1,2}

- (1. 太原理工大学 航空航天学院, 太原 030024;
- 2. 山西省材料强度与结构冲击重点实验室, 太原 030024;
- 3. 中国辐射防护研究院 核应急与核安全研究所, 太原 030024)

摘要: 混凝土在外载荷作用下的力学性能受细观组分特性影响, 其典型非均质性使得传统实验或数值方法难以揭示细观结构对宏观力学性能的影响规律. 为有效分析混凝土骨料-砂浆-孔隙三相细观模型在单轴压缩下的峰值应力, 使用 PYTHON 和 ABAQUS 构建混凝土细观模型的二次开发框架, 生成了包含不同骨料体积分数、孔隙率和受压峰值应力的模型数据集. 基于固定描述符压缩筛选 (sure independence screening and sparsifying operator, SISO) 的机器学习算法, 结合 K 折交叉验证筛选最优物理描述符, 给出了不同骨料体积分数与孔隙率对峰值应力的影响公式. 该公式不仅可准确计算目标参数, 还具备一定物理意义, 能够清晰描述峰值应力的变化趋势. 与传统机器学习算法相比, SISO 在保证精度的同时具有计算成本低、可解释性高的明显优势, 克服了经典机器学习的“黑盒”局限性, 为复合材料的多尺度力学分析提供了新方法.

关键词: 混凝土细观模型; 单轴压缩; 多尺度力学研究; 符号回归; 峰值应力

中图分类号: TU37; TP39; O34 **文献标志码:** A **DOI:** 10.21656/1000-0887.450326

Analysis of Compressive Behaviors of Concrete Mesoscale Models Based on the SISO Algorithm

BAI Yufei^{1,2}, ZHANG Xinyu^{1,2}, QI Xiaopeng^{1,2},
ZHANG Yuhang³, WANG Zhiyong^{1,2}

- (1. College of Aeronautics and Astronautics, Taiyuan University of Technology,
Taiyuan 030024, P.R.China;
- 2. Shanxi Key Laboratory of Material Strength and Structural Impact,
Taiyuan 030024, P.R.China;
- 3. Institute for Nuclear Emergency and Safety, China Institute for Radiation Protection,
Taiyuan 030024, P.R.China)

Abstract: The mechanical properties of concrete under external loads are influenced by its mesoscale components. Due to their heterogeneity, experimental and numerical methods struggle to reveal the impacts of mesoscale structures on the macroscopic mechanical behaviors of concrete. To effectively predict the peak stress of a 3-phase (aggregate, mortar and voids) mesoscale model of concrete under uniaxial compression, a frame-

* 收稿日期: 2024-12-09; 修订日期: 2025-04-05

基金项目: 国家自然科学基金 (12272257); 山西省基础研究计划 (202203021211169)

作者简介: 白宇飞 (2000—), 男, 硕士生 (E-mail: f15035694612@163.com);

王志勇 (1982—), 男, 教授, 博士, 博士生导师 (通信作者. E-mail: wangzhiyong@tyut.edu.cn).

引用格式: 白宇飞, 张新宇, 亓晓鹏, 张煜航, 王志勇. 基于 SISO 算法的混凝土细观模型压缩行为分析 [J]. 应用数学和力学, 2026, 47(3): 354-366.

work for mesoscopic concrete was established with PYTHON and ABAQUS, to generate a dataset of models with varying aggregate volume fractions, porosities and peak compressive stresses. The sure independence screening and sparsifying operator (SISSO) machine learning algorithm, combined with the K -fold cross validation for hyperparameter optimization, was employed to derive a formula describing the effects of the aggregate volume fraction and the porosity on the peak stress. The formula accurately describes the peak stress variation trend, thereby achieving precise predictions and offering physical interpretability. Compared to traditional machine learning algorithms, the SISSO demonstrates advantages of maintaining precision while reducing computation costs and improving interpretability. It overcomes the “black box” limitations of conventional methods, offering new insights for multiscale mechanical analyses of composite materials.

Key words: mesoscale concrete model; uniaxial compression; multiscale mechanical analysis; symbolic regression; peak stress

0 引 言

混凝土作为典型的复合材料具有低成本、高性能的优点,被广泛应用于工程建筑领域,其内部结构在不同尺度上的组成成分及物理性质迥异^[1].细观力学中,混凝土主要由砂浆、孔隙、骨料和界面过渡区组成,这些细观组分的物理特性和几何特征呈现明显的非均匀分布,是影响混凝土宏观力学性能的重要因素^[2-3].研究细观结构的材料特征对宏观力学性能的影响是复合材料力学领域中亟待解决的问题,但由于集料分布的随机性较大,难以利用实验和传统数值模拟方法对材料宏细观之间的力学关系进行解析^[4-5].

近年来,数据驱动方法在处理高维非线性问题方面表现出色,为复合材料的多尺度设计优化^[6]及力学性能分析^[7-8]提供有效方法.Zhong 等^[9]将 BP 神经网络和数值模拟结合,有效预测了混凝土经热处理后的力学参数.Ahmed 等^[10]和 Asteris 等^[11]的研究揭示了神经网络可以准确预测混凝土的抗压强度和弹性模量.虽然神经网络在预测混凝土力学性能方面表现出较强的能力,但其“黑盒”特性^[12]导致模型可解释性较差,造成输入到输出的映射存在物理知识的缺失.因此,提升模型可解释性受到更广泛的关注.刘溢凡等^[13]将混凝土的物理参数引入卷积神经网络,以预测应力-应变曲线,从而增强模型的可解释性,为复合材料多尺度力学分析提供了便捷的解决方案.Zhang 等^[14]使用支持向量机准确捕捉到了混凝土界面抗剪强度和土壤参数之间的物理联系.Gkoutakou 等^[15]基于模糊线性回归方法为计算轻量化泡沫混凝土抗压强度提供了可视化的预测模型.此外,随机森林算法^[16]、决策树^[17]及 AdaBoost^[18]等高可解释性模型也被广泛应用于混凝土力学性能的分析.

然而,传统机器学习难以提供明确的表达式,导致无法有效定量描述复杂变量之间的高维非线性关系,这限制了其在力学机理分析中的应用.基于遗传算法^[19]的符号回归方法^[20-21],能对特征和算子模拟自然选择和遗传过程,逐步演化出可拟合数据的数学表达式,从而提升机器学习模型的可解释性.Ouyang 等^[22]提出的固定描述符压缩筛选(SISSO)新算法,可从庞大的特征空间中有效识别和筛选与目标参数相关的最佳低维描述符,进而来构建形式简洁的公式,以实现符号回归.该算法不仅能够高效处理高维数据,还能揭示复杂系统的潜在物理机制^[23].He 等^[24]使用 SISSO 对钢材的疲劳强度进行研究,算法识别到多个具有物理意义的描述符,可用于钢材结构的设计优化.Wei 等^[25]基于 SISSO 对多孔复合材料的力学性能和纳米孔排列方式拟合数学公式,可以精确预测目标参数,并为多孔结构的逆向设计提供改进方案.SISSO 可进一步作为宏细观多尺度分析方法,通过构建显式公式来更直观地分析模型预测能力,探究不同变量间的数学物理联系,从而为复合材料力学研究提供新的思路.

本文通过有限元模拟细观混凝土单轴受压得到数据集,将骨料体积分数和孔隙率作为特征变量,受压峰值应力作为目标参数,利用 SISSO 算法建立了关于混凝土峰值应力的数学公式.通过引入 10 折交叉验证优

化技术筛选可解释描述符,并提高算法的预测精度及泛化能力,以获得性能最佳的显式表达式.与其他经典机器学习模型相比,SISSO 结果准确且计算成本较低,其高可解释性公式可为细观混凝土模型的反向设计提供可靠的理论依据.

1 细观混凝土有限元模型

在现有研究的基础上^[26-27],利用 PYTHON 程序编写脚本,结合 ABAQUS 组建二次开发框架,建立包含骨料、孔隙和砂浆基质的三相混凝土细观模型.可实现模型自动化构建、赋予材料属性、施加载荷和网格划分等操作,从而组建数据集.需要指出,界面过渡区厚度通常小于 50 μm 且结构复杂,会显著增加建模难度和计算时间^[28].故为提高完整数据集的制作效率,细观模型中未考虑界面过渡区.由于砂浆基质中骨料和孔隙的分布具有随机性和不确定性,使用 Monte-Carlo 法建立试件模型,能够随机均衡地投放骨料和孔隙,并确保骨料、孔隙及其相互之间保持独立分离,不发生重叠或接触.此外,采用 Walraven 公式来确定二维平面内不同粒径骨料的数量分布. Walraven 公式具体如下:

$$P_c(D < D_0) = P_k \left[1.065 \left(\frac{D_0}{D_{\max}} \right)^{0.5} - 0.053 \left(\frac{D_0}{D_{\max}} \right)^4 - 0.012 \left(\frac{D_0}{D_{\max}} \right)^6 - 0.0045 \left(\frac{D_0}{D_{\max}} \right)^8 + 0.0025 \left(\frac{D_0}{D_{\max}} \right)^{10} \right], \quad (1)$$

式中, P_c 表示骨料粒径小于筛孔 D_0 的概率, P_k 为骨料体积百分比, D_{\max} 表示骨料的粒径.为贴合真实混凝土内部形态,细观结构可由多种不同的骨料及孔隙组成,如图 1 所示的椭圆形或多边形骨料,圆形孔隙或圆形与椭圆形夹杂的孔隙等.不同形状的骨料颗粒都会影响混凝土的力学性能^[29],但对其受压峰值应力的影响较小,鉴于圆形几何更有利于直接控制细观参量,可实现快速构建数据集,因此综合平衡精度要求和计算效率,选择使用圆形骨料和孔隙建立混凝土细观模型^[30-33].

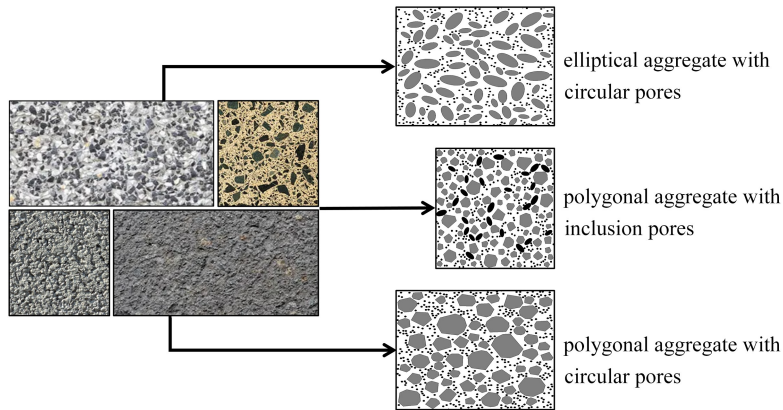


图 1 不同类型的混凝土细观形态

Fig. 1 Different types of meso-scale concrete

注 为了解释图中的颜色,读者可以参考本文的电子网页版本,后同.

混凝土试件样本的尺寸设定为 150 mm×150 mm,骨料采用三级配,其粒径分别设置为 6 mm, 12 mm 和 18 mm,砂浆基质中的孔隙直径均为 2 mm,选取骨料体积分数的范围为 20%~40%,孔隙率范围为 0~10%,可生成 231 组数据.对试件采用三角形平面应变网格进行划分,网格尺寸为 1 mm.

对于材料参数的设置,由于骨料的刚度远高于砂浆,所以在准静态加载模式下可认为骨料不发生破坏,同时砂浆的断裂损伤模式又与混凝土类似,因而分别采用线弹性模型和混凝土塑性损伤模型 (concrete damage plasticity, CDP)^[34] 来模拟骨料和砂浆受载荷的力学行为,二者的材料参数设置见表 1^[35].使用“ABAQUS/Standard”隐式模块进行数值模拟计算,采用准静态单轴压缩试验加载,在试件顶部施加均匀位移

载荷,试件的底部固定约束.生成的混凝土有限元模型如图 2 所示.

表 1 骨料和砂浆的力学参量

Table 1 Mechanical parameters of aggregate and mortar

	compressive strength/MPa	density /($\text{kg}\cdot\text{cm}^{-3}$)	dilatancy angle/($^{\circ}$)	elasticity modulus/GPa	eccentricity/(%)	stress ratio	Poisson's ratio
aggregate	-	2.67	-	43	-	-	0.23
mortar	35	2.40	38	25	0.1	1.16	0.2

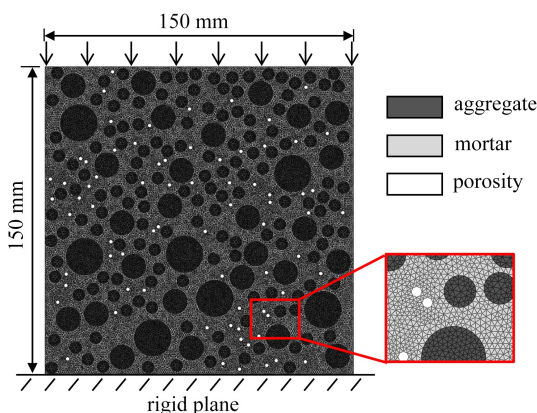


图 2 细观混凝土有限元模型

Fig. 2 The mesoscopic concrete finite element model

对数据集进行批量数值模拟,图 3(a)、3(b)、3(c)展示了部分模拟加载得到的应力-应变曲线.在骨料体积分数和孔隙率分别为 30% 和 1% 下,将数值模拟与相似信息设置下文献中的仿真及试验结果进行对比^[13,36],以验证细观建模的有效性.文献中的模型参数设置与表 1 一致,试验参数见表 2,图 3(d) 显示三者应力-应变曲线趋势相同,峰值应力的误差都小于 10%,吻合较好.由于混凝土内部较强的非均质特性使得其细观模型与实际试件受载荷下的力学行为存在差异,认为应力-应变曲线中下降阶段的误差在合理范围内.基于验证的数值模型,将上述 231 组数据作为实施 SISO 算法的总数据集,并划分为 185 组训练集,46 组测试集.将峰值应力作为表征混凝土开始宏观破坏的临界点,通过研究混凝土的细观组分对峰值应力的影响,来反映其抗压强度的变化规律,对宏观细观多尺度的力学分析具有工程意义.

表 2 文献中混凝土单轴压缩的试验参数^[36]

Table 2 Experimental parameters for uniaxial compression tests on concrete^[36]

parameter	symbol/unit	value
concrete compressive strength	F_c /MPa	35.46
water-cement ratio	w/c	0.37
coarse aggregate particle size	d /mm	5~20
specimen size	l /mm	150×150×300
cement density	ρ_c /($\text{g}\cdot\text{cm}^{-3}$)	3.1
coarse aggregate density	ρ_m /($\text{g}\cdot\text{cm}^{-3}$)	2.71
mortar elasticity modulus	E_m /GPa	23
concrete elasticity modulus	E_c /GPa	32.4
mortar Poisson's ratio	ν_m	0.2
coarse aggregate Poisson's ratio	ν_c	0.2

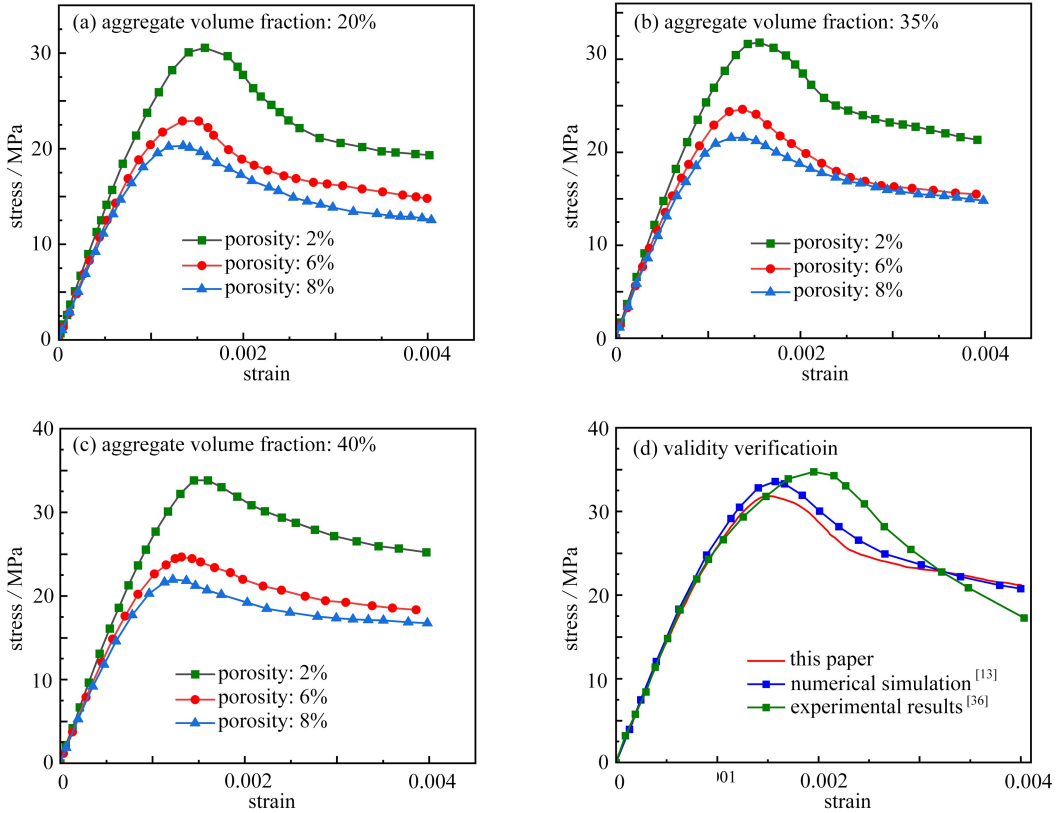


图3 不同细观组分下混凝土细观模型应力-应变曲线及有效性验证^[13,36]

Fig. 3 Stress-strain curves of concrete microscopic models under different microstructural compositions and validation of effectiveness^[13,36]

2 基于 SISSO 分析混凝土单轴压缩下的峰值应力

2.1 SISSO 算法原理

SISSO 是一种基于压缩感知的机器学习方法,用来识别与目标参数相关的描述符,构建特征与目标之间的适当公式,同时具有预测、分类的能力^[22]。相比数值模拟和传统神经网络,该算法无需预设函数形式即可构建公式,从而高效精准地描述变量间的关系。基于 SISSO 研究细观混凝土的骨料体积分数和孔隙率对其单轴受压下峰值应力的影响,能使复杂非均质多尺度问题的分析与预测易于实现。算法拟合的公式具有很高的可移植、可解释特性,从而便于分析力学问题的机理,并判断机器学习模型的预测能力和适用范围,在弥补神经网络“黑盒”缺陷的同时对实际的工程应用展现出极大的便利,算法原理如图 4 所示。

SISSO 始于构造特征空间 Φ 。影响目标参数的初始特征变量构成初始特征空间 Φ_0 , 将一组运算符集合作用于 Φ_0 , 对初始特征进行非线性组合运算产生新特征空间,将该过程表示为 $\hat{H}^{(m)}$:

$$\hat{H}^{(m)} = \{+, -, \times, \div, \dots\} [\phi_1, \phi_2], \quad (2)$$

其中, ϕ_1 和 ϕ_2 表示 Φ 中不同的特征,上标 (m) 表示执行量纲分析,以避免没有物理意义的运算。 Φ_0 由上式转化为 Φ_1 , 初始特征扩充为 Φ_1 中多个不同的新特征。将式(2)迭代 n 次直至特征空间满足预设参数的条件后,得到最终特征空间 Φ_n :

$$\Phi_n \equiv \bigcup_{i=1}^n \hat{H}^{(m)} [\phi_1, \phi_2], \quad \forall \phi_1, \phi_2 \in \Phi_{i-1}, \quad (3)$$

Φ_n 中元素的数量会随 n 的增大呈指数级增长,实现特征扩充,为描述目标参数提供了巨大的特征选择空间。

Φ_n 内部因庞大的特征数量而呈现超高维状态,需要对其进行降维处理以筛选关于目标参数的最优描述符(descriptor)。SISSO 通过联合确定独立筛选(sure independence screening, SIS)和稀疏算子(sparsifying operator, SO)来实现特征降维。SIS 将目标参数与特征之间内积的绝对值作为相关系数对每个特征评分,并选择

Φ_i 中分数最高的前 j 项特征, 构成对应的 i 维特征子空间 S_{iD} , 实现特征空间降维. 在此基础上, 对不同 S_{iD} 做并集运算, SO 再进一步评估该并集中特征组合与目标参数的相关性, 根据 LO 正则化产生最优 i 维描述符 (iD descriptor), 之后对其线性拟合从而生成关于目标参数的公式^[37], 同时使用最小均方根误差 (root mean squared error, RMSE) 来调整公式系数以优化精度, RMSE 的计算公式如下:

$$\delta_{\text{RMSE}} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}, \quad (4)$$

其中, N 表示数据总量, y_i 表示真实值, \hat{y}_i 表示预测值.

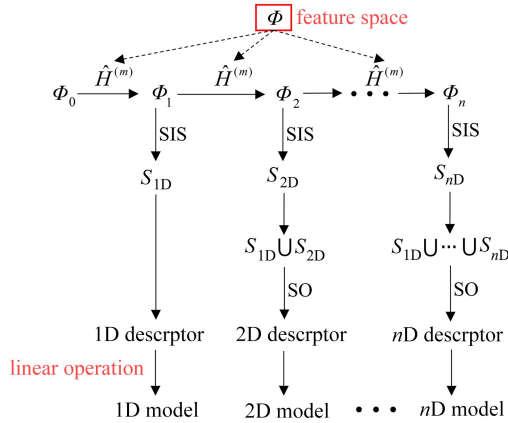


图 4 SISO 算法框架

Fig. 4 Algorithmic framework of SISO

2.2 SISO 算法实施

SISO 算法所需的数据集已在第 1 节中给出, 其中每组样本包含两个独立的特征变量, 即骨料体积分数和孔隙率, 对应的目标参数为峰值应力. 参与构建公式的重要超参数包括运算符、特征复杂度和特征维数. 为使算法具有一定的物理先验知识, 基于混凝土强度经验公式^[38]及 Udrescu 等^[39]对符号回归方法选择算子的归纳, 取运算符集合为

$$\hat{H}^{(m)} = \{ +, -, \times, \div, \exp, \exp(-), \log, ()^{-1}, ()^2, ()^3, \sqrt{\quad} \} [\phi_1, \phi_2]. \quad (5)$$

将上述运算符首先对 Φ_0 中的骨料体积分数和孔隙率展开组合计算, 接着通过 SISO 迭代输出关于峰值应力的最优描述符从而构建公式. 特征复杂度反映了特征中运算符的数量, 复杂度越高, 输出公式中描述符的数学形式越复杂; 特征维数即公式中描述符的数量. 虽然增大特征复杂度和维数能够提高预测精度, 但同时也会带来模型过拟合的风险. 事实上, 为了提高算法的可解释性, 便于其在实际工程中的应用, 描述物理现象的数学公式应尽量保持形式简洁, 并使用较低阶次的多项式. 考虑到算力限制, 故选择特征复杂度和维数均在 1~5 的范围内来训练 SISO 模型从而确定最优超参数组合, 图 5 展示了 1D~5D 模型在不同复杂度下对峰值应力的预测误差.

误差曲线显示, 增加特征维数与提高复杂度都会降低模型的预测误差. 其中, δ_{RMSE} 与 R^2 两项误差均在特征复杂度为 2、特征维数为 3 时趋于收敛. 由于提高超参数值会使得计算成本增加, 同时也会导致过拟合的出现, 所以初步判定特征复杂度为 2 的 3D 模型可构建最佳公式, 对应生成的峰值应力描述符分别为 $P_p + P_a P_p, P_a^{3/2}, e^{-P_p^2}$, 其中, P_a 表示骨料体积分数, P_p 表示孔隙率. 下面将对此模型的稳定性及泛化能力进行评估, 讨论超参数选择的合理性.

2.3 融合 K 折验证构建表征峰值应力的最佳公式

为避免因测试集数据量小而导致的过拟合, 采用 K 折交叉验证提高训练效果. 如图 6 所示, 该方法将训练集随机均匀分为 K 组, 每次使用 $K - 1$ 组作为训练集, 余下一组作为验证集来验证模型预测效果. 上述过程将重复进行 K 次, 每次都选择不同的数据分别作为训练集和验证集. 交叉验证方法能够高效利用有限数据对模型进行相对客观的评估, 从而优化模型预测能力, 避免发生过拟合, 提高其稳定性及泛化能力. 通过模型在

验证集上的平均 R^2 预测误差来评估验证效果, R^2 越接近 1 表示模型的精度越高. 为得到稳定可靠的结果, 取 10 折交叉验证对 2.2 小节中的模型进行评估优化.

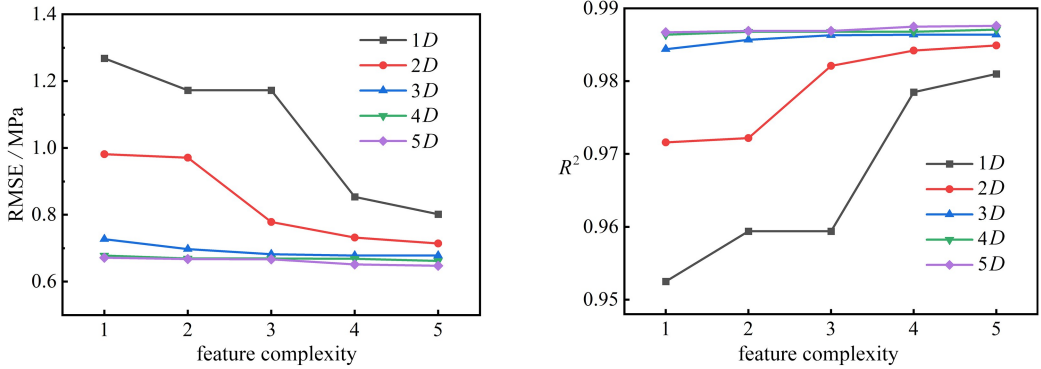


图 5 不同特征复杂度和维数下 SISSO 模型的预测性能

Fig. 5 Performances of the SISSO model with different values of feature complexity and dimension

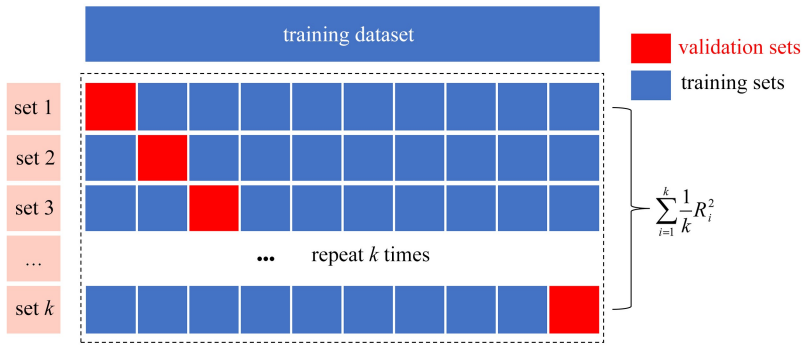


图 6 K 折交叉验证示意图

Fig. 6 Schematic of the K -fold cross-validation test

表 3 展示了 10 折验证中不同描述符组合的出现频次. 2.2 小节中的描述符仅被识别一次, 且与出现频次最高的组合不一致, 同时存在多种其他形式的组合, 反映出此参数条件下, 模型生成公式的随机性较强, 稳定性不足. 进一步分析发现, 10 折训练中, 有 6 折验证集上的模型预测值与数值模拟结果相关性较差, 在小于 30 MPa 的区间内预测值显著低于模拟值, 平均 R^2 为负数, 详见图 7(a). 图 7(c) 显示了出现频次最多的描述符组合在测试集上的表现, 预测结果仍与模拟值存在较大偏差, R^2 仅为 -1.2. 这表明特征复杂度为 2 的 3D 模型性能较差, 难以拟合有效公式. 接下来, 在保证计算效率的同时适当增加一个特征复杂度训练模型, 得到的描述符组合为 $P_a P_p / (P_a + P_p)$, $P_a + P_p^2$ 和 $(P_a + P_p) P_p^2$.

表 3 特征复杂度为 2 时 10 折验证识别的描述符

Table 3 Descriptors identified through 10-fold cross validation with 2-feature complexity

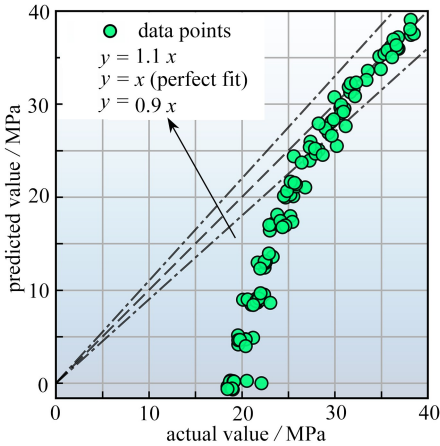
frequency	1st descriptor	2nd descriptor	3rd descriptor
4	$P_p + P_a P_p$	e^{2P_a}	$e^{-P_p^2}$
1	e^{-3P_p}	$P_a + P_p^2$	$\frac{P_p^2}{P_a}$
1	$P_a + P_p^2$	$2P_a + P_p$	$\frac{P_p^2}{P_a}$
1	$P_p + P_a P_p$	$P_a e^{P_a}$	$e^{-P_p^2}$
1	$P_p + P_a P_p$	$P_p^3 - P_a$	$P_p^2 - P_a$
1	$P_p + P_a P_p$	$P_a^{3/2}$	$e^{-P_p^2}$
1	$\frac{P_p}{e^{P_a}}$	$P_a + P_p^2$	$2P_a + P_p$

表 4 显示,经过 10 折交叉验证后该组合可被识别 9 次,稳定性良好.模型在验证集上的平均 R^2 为 0.983 6,测试集上的 R^2 为 0.981 2,图 7(b) 和 7(d) 显示预测值和数值计算结果高度相关,二者均匀分布于 $y = x$ 两侧,表明算法此时可以精确描述骨料体积分数和孔隙率对峰值应力的影响,且具有良好的泛化能力.在该超参数组合下,模型训练时长为 0.04 s,展现了较高的计算效率.

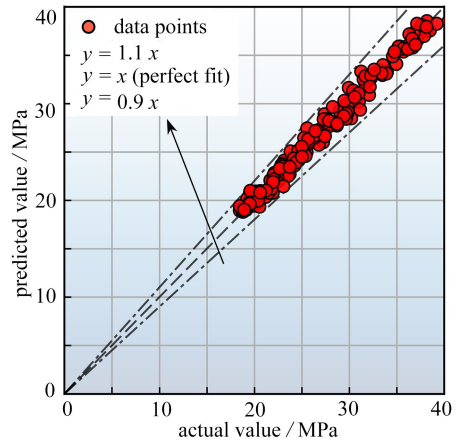
表 4 特征复杂度为 3 时 10 折验证识别的描述符

Table 4 Descriptors identified through 10-fold cross validation with 3-feature complexity

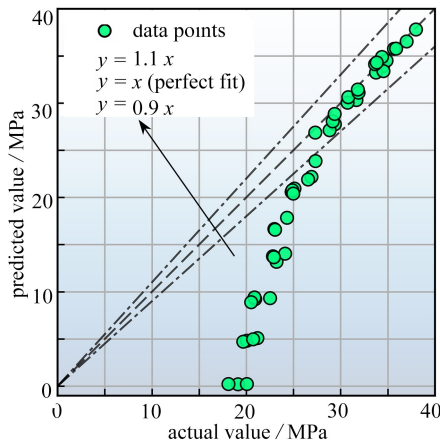
frequency	1st descriptor	2nd descriptor	3rd descriptor
9	$\frac{P_a P_p}{P_a + P_p}$	$P_a + P_p^2$	$(P_a + P_p) P_p^2$
1	$\frac{P_a P_p}{P_a + P_p}$	$P_a + P_p^2$	$P_a P_p^2$



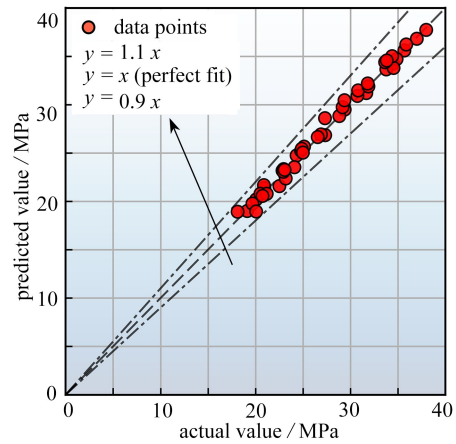
(a) 特征复杂度为 2 的 3D 模型验证集回归图
(a) The regression plot of the 3D model with 2-feature complexity on the validation set



(b) 特征复杂度为 3 的 3D 模型验证集回归图
(b) The regression plot of the 3D model with 3-feature complexity on the validation set



(c) 特征复杂度为 2 的 3D 模型测试集回归图
(c) The regression plot of the 3D model with 2-feature complexity on the test set



(d) 特征复杂度为 3 的 3D 模型测试集回归图
(d) The regression plot of the 3D model with 3-feature complexity on the test set

图 7 不同特征复杂度下 SISSO 模型的预测表现

Fig. 7 SISSO model prediction performances under different-feature complexities

考虑到进一步提高超参数将会导致计算成本增加,因此选择特征复杂度为 3 的 3D 模型作为最优方案.最终得到在第 1 节的材料参数设置下,混凝土受单轴压缩的峰值应力与其骨料体积分数和孔隙率之间的公式为

$$\sigma_{\max} = C_1 \frac{P_a P_p}{P_a + P_p} + C_2 (P_a + P_p^2) + C_3 (P_a + P_p) P_p^2 + b. \quad (6)$$

表5列出了式(6)中的系数项与常数项,其中参数 C_i 作为细观结构的敏感参量,综合考虑了外载荷及混凝土材料参数的影响,用于反映单轴压缩条件下峰值应力对细观结构参数组合的敏感程度,其单位为MPa;常数 b 表示在骨料体积分数和孔隙率均趋近于零时纯砂浆基质的峰值应力.由于所提出的公式适用于骨料和孔隙的含量分别在20%~40%和0%~10%范围内,因此 b 仅作为该工况下的理论参考值.公式中的 $P_a P_p / (P_a + P_p)$, $P_a + P_p^2$ 和 $(P_a + P_p) P_p^2$ 体现了骨料体积分数与孔隙率之间的耦合关系.式(6)满足物理学中一般函数的六项属性^[38],具有一定的物理意义,可对实际工程应用起指导作用.

表5 式(6)中的参数项(单位:MPa)

Table 5 The parameters in formula (6) (unit: MPa)

C_1	C_2	C_3	b
-283.373	15.245	738.123	32.558

为验证公式的有效性,将式(6)的函数曲面与第1节中的数值模拟结果进行对比,如图8所示.结果显示,样本数据点的分布与公式曲面高度吻合,发现混凝土在单轴受压下的峰值应力不是骨料体积分数或孔隙率的单调函数.峰值应力与孔隙率呈反比例关系,同时式(6)中的描述符组合揭示了骨料体积分数和孔隙率对峰值应力的作用是耦合的,在不同孔隙率水平下,骨料体积分数对峰值应力的影响趋势有显著差异.当孔隙率较小时,峰值应力与骨料体积分数呈明显正比关系;随着孔隙率的提高,骨料体积分数对峰值应力的影响逐渐减弱,当孔隙率为10%时,峰值应力对骨料数量变化的敏感性降为最低,即使骨料体积分数超过30%,峰值应力也仅有轻微上升趋势.这是由于外载荷作用下,提高骨料数量可增强混凝土抵抗变形的能力,试件的峰值应力随之增加;而大量孔隙的存在会使混凝土内部承力结构严重缺失,导致试件整体强度迅速减小,峰值应力随即下降,此时骨料数量的变化对结构强度已无明显作用.公式表明,孔隙率是影响混凝土单轴受压下峰值应力的主要因素,这与已有研究结果一致^[40].此外,当退化至 $P_p = 0$ 时,由图8可看出,式(6)在 P_a 边界处依然有精确的计算结果.由于公式中常数项 b 与砂浆的抗压强度相似,数值模拟发现当骨料体积分数和孔隙率均为0时,试件的峰值应力为34.679 MPa,这与式(6)的计算结果接近,二者误差在10%以内.因此,SISSO构建的公式能够较好地反映不含孔隙和骨料的纯砂浆试件在受压下的力学性能,从而进一步证明了其在物理层面上的合理性.

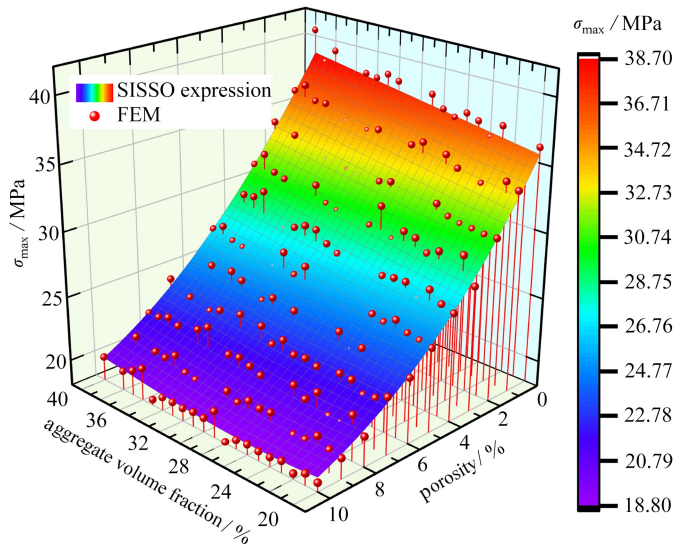


图8 峰值应力公式曲面

Fig. 8 The surface of the peak stress formula

SISSO算法将宏细观间高维非线性关系解耦为可视化的数学表达,量化评估了骨料体积分数和孔隙率对峰值应力的影响机制,为混凝土抗压设计在细观层面提供了结构-性质的准确描述.需要指出的是,式(6)

建立在本文细观组分参数范围的基础上,由于机器学习模型较依赖于训练算法的数据集^[41],所以当输入特征远超样本范围或试件其他力学参数改变时,该公式则不再适用.但算法仍可在不同工况下,对特征和目标参数间的复杂非线性关系进行求解.

3 算法对比验证

为检验 SISSO 的性能,我们将其与 PYTHON 库 scikit-learn 中的其他经典机器学习方法进行了对比,包括支持向量机回归(support vector regression, SVR)^[42]、线性回归(linear regression, LR)^[43]、决策树(decision tree, DT)^[44]、随机森林(random forest, RF)^[45]和梯度增强树(gradient boost decision tree, GBDT)^[46].表 6 给出本文训练不同机器学习模型的计算环境及软硬件参数,图 9 对比展示了五种算法和 SISSO 在测试集上的表现.其中,SISSO 具有最短的计算时间,同时其 R^2 略低于 SVR 与 RF,并高于其他三种模型.表明 SISSO 精度与经典机器学习算法相当,但计算效率卓越.

表 6 计算环境及软硬件参数

Table 6 Computational environment and hardware/software parameters

parameter	value
central processing unit	Intel(R) Core(TM) i7-9750H CPU @ 2.60 GHz
memory	RAM 8 GB
graphics card	NVIDIA GeForce GTX 1650
system	Windows 10
environment	PYTHON 3.8 Scikit-learn 0.23.2 NUMPY 1.22.4

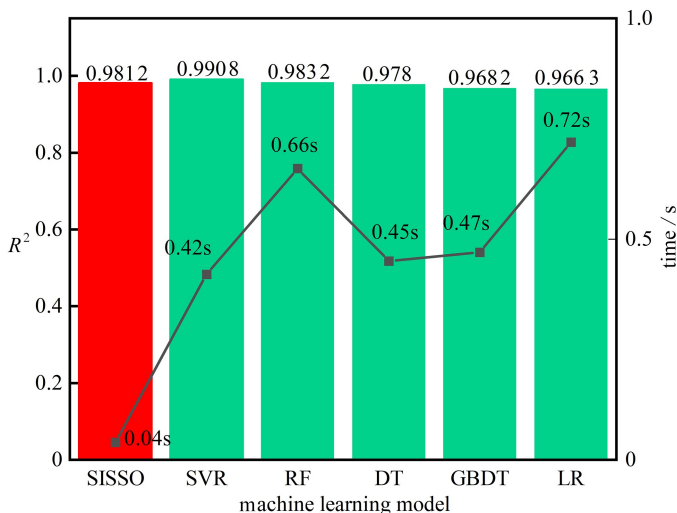


图 9 不同算法在测试集上的 R^2 与计算时间

Fig. 9 R^2 values and computation durations of different algorithms on the test set

SISSO 生成的公式不仅能精确描述目标变量与特征之间的数学关系,还提供了具有物理意义的定量信息,显著提升了模型的可解释性.SISSO 作为一种高效便捷的分析预测工具,其生成的公式能进一步推广到材料的反向设计,即根据目标参数需求反推特征变量,可应用于实际的工程优化.

4 结 论

本文基于 SISSO 算法,研究了骨料-砂浆-孔隙三相混凝土细观模型在单轴压缩下的力学性能.将峰值应力作为目标参数,骨料体积分数和孔隙率作为主要特征,经过优化算法超参数,明确了目标和特征之间的显式函数关系,得出了以下结论:

1) 骨料体积分数和孔隙率二者相互耦合,共同决定峰值应力的变化规律.其中,孔隙率占主导作用,当孔隙率为 10%时,骨料体积分数的变化对峰值应力无明显影响.

2) 通过 K 折交叉验证确定了当特征复杂度和维数均为 3 时,SISSO 能够基于最优物理描述符构建可解释的峰值应力计算公式.算法在验证集和测试集上的 R^2 分别为 0.983 6 和 0.981 2,对于不同骨料体积分数和孔隙率均有精确的计算结果,表明公式的泛化能力良好.

3) 相较于经典机器学习算法,SISSO 在保持精度相当的同时计算成本最低,其生成的峰值应力显式公式具有较强的数学物理意义,弥补了“黑盒模型”可解释性低的缺陷.公式形式简洁,可有效简化混凝土的复杂力学分析,并能用于细观模型的逆向设计,进而为复合材料的多尺度力学研究提供新策略.

参考文献(References):

- [1] BERNARD O, ULM F J, LEMARCHAND E. A multiscale micromechanics-hydration model for the early-age elastic properties of cement-based materials[J]. *Cement and Concrete Research*, 2003, **33**(9): 1293-1309.
- [2] BEUSHAUSEN H, DITTMER T. The influence of aggregate type on the strength and elastic modulus of high strength concrete[J]. *Construction and Building Materials*, 2015, **74**: 132-139.
- [3] TIAN Z, YAN Y, LI J, et al. Progressive damage and failure analysis of three-dimensional braided composites subjected to biaxial tension and compression[J]. *Composite Structures*, 2018, **185**: 496-507.
- [4] AHMAD A, FAROOQ F, NIEWIADOMSKI P, et al. Prediction of compressive strength of fly ash based concrete using individual and ensemble algorithm[J]. *Materials*, 2021, **14**(4): 794.
- [5] 李向南, 左晓宝, 周广盼, 等. 混凝土多尺度应力响应方程及其数值模拟[J]. 力学学报, 2022, **54**(11): 3113-3126. (LI Xiangnan, ZUO Xiaobao, ZHOU Guangpan, et al. Equation and numerical simulation on multiscale stress response of concrete[J]. *Chinese Journal of Theoretical and Applied Mechanics*, 2022, **54**(11): 3113-3126. (in Chinese))
- [6] KARAPIPERIS K, STAINIER L, ORTIZ M, et al. Data-driven multiscale modeling in mechanics[J]. *Journal of the Mechanics and Physics of Solids*, 2021, **147**: 104239.
- [7] LV Z, JIANG A, LIANG B. Development of eco-efficiency concrete containing diatomite and iron ore tailings: mechanical properties and strength prediction using deep learning[J]. *Construction and Building Materials*, 2022, **327**: 126930.
- [8] ZHANG X, ZHAO T, LIU Y, et al. A data-driven model for predicting the mixed-mode stress intensity factors of a crack in composites[J]. *Engineering Fracture Mechanics*, 2023, **288**: 109385.
- [9] ZHONG W L, DING H, ZHAO X, et al. Mechanical properties prediction of geopolymers subjected to high temperature by BP neural network[J]. *Construction and Building Materials*, 2023, **409**: 133780.
- [10] AHMED H U, MOHAMMED A S, MOHAMMED A A. Multivariable models including artificial neural network and M5P-tree to forecast the stress at the failure of alkali-activated concrete at ambient curing condition and various mixture proportions[J]. *Neural Computing and Applications*, 2022, **34**(20): 17853-17876.
- [11] ASTERIS P G, MOKOS V G. Concrete compressive strength using artificial neural networks[J]. *Neural Computing and Applications*, 2020, **32**(15): 11807-11826.
- [12] RUDIN C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead[J]. *Nature Machine Intelligence*, 2019, **1**(5): 206-215.
- [13] 刘溢凡, 张杰, 张新宇, 等. 基于“AM-GoogLeNet+BP”联合数据驱动的混凝土细观模型压缩应力-应变曲线预测[J]. 力学学报, 2023, **55**(4): 925-938. (LIU Yifan, ZHANG Jie, ZHANG Xinyu, et al. Prediction of concrete meso-model compression stress-strain curve based on “AM-GoogLeNet+BP” combined data-driven methods [J]. *Chinese Journal of Theoretical and Applied Mechanics*, 2023, **55**(4): 925-938. (in Chinese))
- [14] ZHANG C, JI J, GUI Y, et al. Evaluation of soil-concrete interface shear strength based on LS-SVM[J]. *Geomechanics and Engineering*, 2016, **11**(3): 361-372.
- [15] GKOUNTAKOU F I, PAPADOPOULOS B K. The use of fuzzy linear regression with trapezoidal fuzzy numbers to predict the compressive strength of lightweight foamed concrete[J]. *Mathematical Modelling of Engineering Problems*, 2022, **9**(1): 1-10.
- [16] CHUN P J, UJIKE I, MISHIMA K, et al. Random forest-based evaluation technique for internal damage in reinforced concrete featuring multiple nondestructive testing results[J]. *Construction and Building Materials*,

- 2020, **253**: 119238.
- [17] BEN CHAABENE W, FLAH M, NEHDI M L. Machine learning prediction of mechanical properties of concrete: critical review[J]. *Construction and Building Materials*, 2020, **260**: 119889.
- [18] CUI L, CHEN P, WANG L, et al. Application of extreme gradient boosting based on grey relation analysis for prediction of compressive strength of concrete[J]. *Advances in Civil Engineering*, 2021, **2021**(1): 8878396.
- [19] KOZA J R. *Genetic Programming: on the Programming of Computers by Means of Natural Selection*[M]. MIT Press, 1992.
- [20] SCHMIDT M, LIPSON H. Distilling free-form natural laws from experimental data[J]. *Science*, 2009, **324**(5923): 81-85.
- [21] YI R, GEORGIOU D, LIU X, et al. Mechanics-informed, model-free symbolic regression framework for solving fracture problems[J]. *Journal of the Mechanics and Physics of Solids*, 2025, **194**: 105916.
- [22] OUYANG R, CURTAROLO S, AHMETCIK E, et al. SISSO: a compressed-sensing method for identifying the best low-dimensional descriptor in an immensity of offered candidates[J]. *Physical Review Materials*, 2018, **2**(8): 083802.
- [23] PURCELL T A R, SCHEFFLER M, GHIRINGHELLI L M. Recent advances in the SISSO method and their implementation in the SISSO++ code[J]. *The Journal of Chemical Physics*, 2023, **159**(11): 114110.
- [24] HE N, OUYANG R, QIAN Q. Learning interpretable descriptors for the fatigue strength of steels[J]. *AIP Advances*, 2021, **11**(3): 035018.
- [25] WEI A, YE H, GUO Z, et al. SISSO-assisted prediction and design of mechanical properties of porous graphene with a uniform nanopore array[J]. *Nanoscale Advances*, 2022, **4**(5): 1455-1463.
- [26] ZHANG Y, CHEN Q, WANG Z, et al. 3D mesoscale fracture analysis of concrete under complex loading[J]. *Engineering Fracture Mechanics*, 2019, **220**: 106646.
- [27] 刘溢凡, 马小敏, 王志勇, 等. 基于 ANN 的混凝土均匀化方法解析解[J]. 应用数学和力学, 2024, **45**(5): 554-570. (LIU Yifan, MA Xiaomin, WANG Zhiyong, et al. Analytical solution of the concrete homogenization method based on the ANN[J]. *Applied Mathematics and Mechanics*, 2024, **45**(5): 554-570. (in Chinese))
- [28] BARNES B D, DIAMOND S, DOLCH W L. Micromorphology of the interfacial zone around aggregates in Portland cement mortar[J]. *Journal of the American Ceramic Society*, 1979, **62**(1/2): 21-24.
- [29] ICHINO H, KUWAHARA N, BEPPU M, et al. Effects of the shape, size, and surface roughness of glass coarse aggregate on the mechanical properties of two-stage concrete[J]. *Construction and Building Materials*, 2024, **411**: 134296.
- [30] WANG X F, YANG Z J, YATES J R, et al. Monte Carlo simulations of mesoscale fracture modelling of concrete with random aggregates and pores[J]. *Construction and Building Materials*, 2015, **75**: 35-45.
- [31] 薛刚, 刘毅, 牟一飞. 基于 3D 细观模型的混凝土单轴拉压应力应变关系影响因素研究[J]. 材料导报, 2025, **39**(15): 24060161. (XUE Gang, LIU Yi, MOU Yifei. Analysis of influencing factors of stress-strain curve of three-dimensional mesoscopic concrete under uniaxial tension and compression load[J]. *Materials Reports*, 2025, **39**(15): 24060161. (in Chinese))
- [32] CHEN P, LIU J, CUI X, et al. Mesoscale analysis of concrete under axial compression[J]. *Construction and Building Materials*, 2022, **337**: 127580.
- [33] 肖诗云, 朱梁. 孔隙对混凝土宏观力学性质的影响[J]. 沈阳建筑大学学报(自然科学版), 2016, **32**(4): 608-618. (XIAO Shiyun, ZHU Liang. Study on the effects of voids on macro-mechanical properties of concrete[J]. *Journal of Shenyang Jianzhu University (Natural Science)*, 2016, **32**(4): 608-618. (in Chinese))
- [34] LI B, JIANG J, XIONG H, et al. Improved concrete plastic-damage model for FRP-confined concrete based on true tri-axial experiment[J]. *Composite Structures*, 2021, **269**: 114051.
- [35] 陈青青, 张煜航, 张杰, 等. 含孔隙混凝土二维细观建模方法研究[J]. 应用数学和力学, 2020, **41**(2): 182-194. (CHEN Qingqing, ZHANG Yuhang, ZHANG Jie, et al. Study on a 2D mesoscopic modeling method for concrete with voids[J]. *Applied Mathematics and Mechanics*, 2020, **41**(2): 182-194. (in Chinese))
- [36] 苏捷. 混凝土受压与受拉性能的尺寸效应研究[D]. 长沙: 湖南大学, 2013. (SU Jie. The research on the size effect of concrete behavior in compression and tension[D]. Changsha: Hunan University, 2013. (in Chinese))

- nese))
- [37] 胡红青, 吴邵刚, 郭治廷, 等. 基于 SISSO 和机器学习方法的钙钛矿结构的稳定性预测: 新型容许因子建立与验证[J]. 中国有色金属学报, 2020, **30**(8): 1887-1894. (HU Hongqing, WU Shaogang, GUO Zhiting, et al. New tolerance factor based on SISSO and machine learning for predicting stability of perovskite structure[J]. *The Chinese Journal of Nonferrous Metals*, 2020, **30**(8): 1887-1894. (in Chinese))
- [38] 金浏, 李健, 余文轩, 等. 混凝土动态双轴拉压破坏准则细观数值模拟研究[J]. 力学学报, 2022, **54**(3): 800-809. (JIN Liu, LI Jian, YU Wenxuan, et al. Mesoscopic numerical simulation on dynamic biaxial tension-compression failure criterion of concrete[J]. *Chinese Journal of Theoretical and Applied Mechanics*, 2022, **54**(3): 800-809. (in Chinese))
- [39] UDRESCU S M, TEGMARK M. AI Feynman: a physics-inspired method for symbolic regression[J]. *Science Advances*, 2020, **6**(16): eaay2631.
- [40] 田梦云, 张恩, 曹瑞东, 等. 基于细观尺度的混凝土单轴力学性能仿真计算分析[J]. 应用力学学报, 2020, **37**(3): 975-981. (TIAN Mengyun, ZHANG En, CAO Ruidong, et al. Meso-scale simulation analysis of uniaxial mechanical behavior of concrete[J]. *Chinese Journal of Applied Mechanics*, 2020, **37**(3): 975-981. (in Chinese))
- [41] SHALEV-SHWARTZ S, BEN-DAVID S. *Understanding Machine Learning*[M]. Cambridge, UK: Cambridge University Press, 2014.
- [42] AWAD M, KHANNA R. *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*[M]. Berkeley, CA: Apress, 2015: 67-80.
- [43] MONTGOMERY D C, PECK E A, VINING G G. *Introduction to Linear Regression Analysis*[M]. Wiley, 1982.
- [44] QUINLAN J R. Induction of decision trees[J]. *Machine Learning*, 1986, **1**(1): 81-106.
- [45] SVETNIK V, LIAW A, TONG C, et al. Random forest: a classification and regression tool for compound classification and QSAR modeling[J]. *Journal of Chemical Information and Computer Sciences*, 2003, **43**(6): 1947-1958.
- [46] FRIEDMAN J H. Greedy function approximation: a gradient boosting machine[J]. *The Annals of Statistics*, 2001, **29**(5): 1189-1232.