第 54 卷第 3 期
2026 年 3 月

硅 酸 盐 学 报
JOURNAL OF THE CHINESE CERAMIC SOCIETY

Vol. 54，No. 3
March，2026

http://jccsoc.cbpt.cnki.net

**DOI**：10.14062/j.issn.0454-5648.20250680

# 水泥基材料中钢筋腐蚀智能预测与风险分级

吕余辉[1]，纪豪栋[1]，余倩倩[2]，叶海隆[1]

(1. 香港大学土木工程系，香港 999077；2. 同济大学土木工程学院建筑工程系，上海 200092)

**摘　要**：提出了一种融合物理先验知识与不确定性量化的高斯过程回归(GPR)框架，用于预测水泥基材料中钢筋腐蚀速率，并拓展至腐蚀风险等级划分与工程响应决策。通过在模型中嵌入电化学相关参数并开展不确定性评估，结果显示，该框架在预测精度、泛化能力及边界条件下的稳定性方面均优于传统 GPR，尤其在高腐蚀环境中能够有效降低预测的系统性偏差。进一步分析表明，当水泥基材料孔溶液的氯离子与氢氧根离子浓度比处于 0.2~0.6 的区间时，模型的误差与不确定性显著增加，建议将该区间作为后续实验的重点关注范围。基于预测值及其置信区间，构建了包括常规监测、确认复核、谨慎使用与紧急应对的风险响应策略，以支撑耐久性评估与工程决策。

**关键词**：水泥基材料；钢筋腐蚀；高斯过程回归；工程决策
**中图分类号**：TU511　　**文献标志码**：A　　**文章编号**：0454–5648(2026)03–0947–10
**网络出版时间**：2026–02–13



混凝土结构在服役期内的耐久性退化问题日益受到广泛关注，其中钢筋锈蚀被普遍认为是导致其劣化的关键机制之一。在海洋环境或重工业区等严苛服役条件下，氯离子侵蚀与碳化的协同效应极易诱发钢筋脱钝，显著加速锈蚀进程，进而引发截面削弱、混凝土开裂及保护层剥落等一系列危害，严重威胁结构安全性与服役寿命[1–2]。

腐蚀速率作为表征钢筋劣化程度的关键动力学指标，其准确预测对服役混凝土结构的寿命评估、维修策略制定与风险管控至关重要。为实现对腐蚀行为的量化表征，国内外学者已提出多种预测方法，主要包括经验模型和电化学模型两类。经验模型通常基于实验数据回归分析，建立腐蚀速率与环境及材料因子(如氯离子浓度、pH 值、水泥基材料水胶比等)之间的函数关系，具有形式简洁、计算高效的优势。然而，该类模型在不同材料体系或服役条件下的适应性较差，难以泛化[3–4]。相比之下，电化学模型从腐蚀反应机理出发，通过控制方程描述钢筋表面电化学过程，具备更强的物理基础与解释能力，但其建模过程高度依赖参数，求解复杂，且实际工程应用成本较高[5–7]。因此，亟须发展兼具准确性、可解释性与工程适用性的腐蚀预测方法。

近年来，机器学习(ML)凭借其强大的数据驱动建模能力和对复杂非线性关系的适应性，在腐蚀预测领域已展现出广阔的应用前景。在大气腐蚀、管道腐蚀及多种金属服役环境分析中，ML 模型均展现出优异的预测性能与泛化能力[8–11]。然而，针对钢筋混凝土结构的锈蚀预测，尽管在文献中已有一定的实验数据积累，但实际应用中仍面临数据来源异质性高、特征维度不全及样本分布不均等问题，为腐蚀数据库与 ML 模型的构建带来巨大挑战[12–13]。更重要的是，现有大多数 ML 方法往往侧重于提高点估计精度，忽视了对预测不确定性的量化表达。在结构耐久性评估与管理中，预测不确定性不仅直接影响工程判断的可靠性，也是实现“可信人工智能”决策支持能力的关键要素。因此，开发同时兼顾预测精度与不确定性评估能力的 ML 框架，对提升腐蚀监测的工程实用价值具有重要意义。

基于上述背景，本工作提出一种融合物理先验知识与不确定性量化的高斯过程回归(GPR)框架，用于水泥基材料体系中钢筋腐蚀速率的预测及其风险等级的辅助判断。GPR 具有非参数建模优势，可

收稿日期：2025–09–15。　修订日期：2025–10–02。
基金项目：国家重点研发计划“工程科学与综合交叉”重点专项(2024YFF0510400)。
第一作者：吕余辉(1991—)，女，博士。
通信作者：叶海隆(1988—)，男，博士，副教授。

**Received date:** 2025–09–15.　**Revised date:** 2025–10–02.
**First author:** LYU Yuhui (1991–), female, Ph.D.
**E-mail:** yuhuilyu@hku.hk
**Correspondent author:** YE Hailong (1988–), male, Ph.D., Associate Professor.
**E-mail:** hlye@hku.hk

在不预设函数形式的前提下精确拟合复杂的腐蚀机制，并自然输出预测值的置信区间，为实现“可信结构智能监测”提供理论基础和工具支撑。本文聚焦以下 2 个核心问题展开研究：1)如何提升 GPR 在钢筋腐蚀预测任务中的准确性与鲁棒性？2)GPR 模型输出的不确定性信息能否有效服务于腐蚀等级划分与工程决策响应？

为此，本工作基于典型腐蚀实验数据构建训练集，并引入物理机制相关的先验约束，以增强模型的物理可解释性与泛化能力。在预测阶段，结合腐蚀等级分类机制和基于置信区间宽度的不确定性评判标准，构建“预测值+不确定性”的双维度响应策略框架。该框架通过提供透明的预测置信度、基于物理机制的建模过程以及基于风险的干预建议，体现了“可信人工智能”的核心特征(可解释性、透明性与问责性)，从而有望实现腐蚀风险的精细分级与智能化响应。

# 1 方法与数据

## 1.1 数据集描述

本工作共采集 180 组钢筋腐蚀实验数据，每组数据包含 15 项腐蚀相关特征及其对应的钢筋腐蚀速率。实验数据来源于 30 组完整样本在碳化处理后，分别置于 6 种不同的相对湿度 (RH) 环境下测得的腐蚀数据，RH 条件依次为 56%、68%、75%、84%、92%和 97%。数据的采集与预处理流程详见文献 [14–15] 。

表 1 汇总了用于腐蚀速率预测的 15 项输入特征。为便于分析与建模，根据特征的来源与物理意义，其被划分为以下 4 类：

1) 配合比参数：表征试件的材料组成与设计配比，包括水泥、矿渣、粉煤灰、硅灰占据胶凝材料的比例，水胶比，以及内掺氯盐的质量分数。

2) 材料特性参数：反映硬化砂浆的微观结构和化学状态，包括游离氯离子含量、总孔隙率及孔隙液 pH 值。

3) 环境参数：描述样本暴露时所处的外部环境条件，包括 RH、饱和度和含水率。

4) 电化学参数：与腐蚀反应直接相关，提供关于腐蚀电化学过程的定量信息，包括腐蚀电位、电阻率、氯离子与氢氧根离子浓度比值([$Cl^-$]/[$OH^-$])，以及目标输出变量——腐蚀速率。

本工作数据库虽然覆盖了多种配合比设计及 6 种典型湿度环境，能够在一定程度上反映不同材料组成与环境条件下的腐蚀特征，但总体规模仍为 180 组。在关键腐蚀转折区间(lg([$Cl^-$]/[$OH^-$])∈[0.2, 0.6]) 内，样本分布相对稀疏，这在一定程度上影响了模型在该区间的预测稳定性。

**表 1 腐蚀相关因素的详细信息**

**Table 1 Details of corrosion related factors**

| Category | Corrosion-related factors | Unit | Min. value | Max. value |
|---|---|---|---|---|
| Mixture parameters | Cement proportion | % (in mass) | 20 | 100 |
| | Slag proportion | % (in mass) | 0 | 80 |
| | Fly ash proportion | % (in mass) | 0 | 45 |
| | Silica fume proportion | % (in mass) | 0 | 20 |
| | Water to binder ratio | 1 | 0.30 | 0.55 |
| | Incorporated chloride content | % (in mass) | 0.00 | 2.40 |
| Material properties | Free chloride content | % (in mass) | 0.01 | 1.69 |
| | pH value | 1 | 9.87 | 14.22 |
| | Porosity | 1 | 0.09 | 0.24 |
| Environmental parameters | Relative humidity | % | 56 | 97 |
| | Degree of saturation | 1 | 0.25 | 1.00 |
| | Water content | % | 0.06 | 0.24 |
| Electrochemical parameters | lg [$Cl^-$]/[$OH^-$] | 1 | –1.84 | 4.25 |
| | Electrical resistivity | Ω·m | 12.72 | 12 168.00 |
| | Corrosion potential | mV | –668 | –97 |

## 1.2 模型构建

为实现钢筋腐蚀速率的可预测性，本工作构建了一个以 GPR 为核心的预测框架。考虑到本研究数据库规模有限，且在部分关键腐蚀转折区间样本分布相对稀疏，GPR 被选为建模工具，以在小样本情境下实现可靠预测并量化其不确定性。该模型以配合比参数与电化学参数(尤其是 [$Cl^-$]/[$OH^-$]) 为输入变量，旨在预测不同环境与材料条件下的钢筋腐蚀行为。

选择 GPR 作为建模工具的原因主要包括以下 2 点：

1) 适应不规则采样的建模需求。本工作采集的数据涵盖多种配合比设计与环境暴露条件，各样本在腐蚀龄期、测量指标及特征范围上存在显著差异，呈现出典型的非结构化采样特征。相较于传统回归方法，GPR 能够有效处理这类不规则分布的数据，并通过自适应调整预测方差，反映局部样本的稀疏程度。

2) 具备非线性建模与物理融合能力。GPR 通过核函数灵活刻画输入特征与腐蚀速率之间的复杂非线性关系，无须预设函数形式即可实现良好拟合。

同时，GPR 支持将腐蚀相关的电化学知识融入均值函数或协方差结构中，从而提升模型的可解释性与泛化能力。

1.2.1　模型形式　　本工作将腐蚀速率 $y$ 建模为输入特征空间 $x$ 上的随机函数。输入特征包括[Cl⁻]/[OH⁻]、孔隙率、RH 等变量。该函数受独立高斯噪声 $\varepsilon$ 扰动，其数学形式可表示为：

$$y = f(x) + \varepsilon$$

$$\varepsilon \sim \mathcal{N}(0, \sigma_n^2)$$

式中：$f(x) \sim \mathcal{GP}[m(x), k(x,x')]$ 为 $f(x)$ 服从以 $m(x)$ 为均值函数、$k(x,x')$ 为协方差函数的高斯过程；$\sigma_n^2$ 为观测噪声方差。

为了更好地捕捉腐蚀行为随环境与材料参数变化所呈现的平滑变化趋势，本工作选用高斯径向基函数(RBF)作为协方差函数，其形式如下：

$$k(x,x') = \sigma_f^2 \exp(-\| x - x' \|^2 / 2l^2)$$

式中：$\| x - x' \|$ 为输入向量间的欧几里得距离；$\sigma_f^2$ 为信号方差；$l$ 为长度尺度超参数；用于控制函数变化的平滑程度[16]。

1.2.2　物理先验的引入　　为引入腐蚀机制的领域知识，本工作将高斯过程回归 (GPR) 模型的均值函数设定为基于离子浓度比值 (lg [Cl⁻]/[OH⁻]) 的线性形式，即：

$$m([Cl^-]/[OH^-]) = a \cdot \lg([Cl^-]/[OH^-]) + b$$

此形式来源于腐蚀电化学理论，体现了氯离子浓度相对于氢氧根离子浓度升高时，钢筋腐蚀速率呈上升趋势的基本规律。已有研究在模拟孔隙液中建立了腐蚀速率($i_{corr}$)与离子浓度比值之间的经验关系式[17]，即：

$$i_{corr} = 2.39 + 3.44 \cdot \lg([Cl^-]/[OH^-])$$

结果表明，随着[Cl⁻]/[OH⁻]升高，钢筋腐蚀速率显著增加，这与钝化膜失稳的机理相一致。需要指出的是，[Cl⁻]/[OH⁻]虽已包含在输入特征中，但在均值函数中再次引入并非重复，而是作为电化学规律的先验约束，以确保模型在样本稀疏区间仍保持与机理一致的单调趋势。这种做法既保证了模型在物理层面的合理性，又与输入层面的非线性建模形成互补，从而实现物理约束与数据驱动的双融合。

然而，仅依赖 lg([Cl⁻]/[OH⁻])仍存在一定局限性，例如未能全面表征钢材成分、显微组织、热处理特性以及离子扩散行为等对腐蚀过程的影响。这一设计在当前数据规模下保证了模型结构的简洁性与稳定性，但也表明物理约束的覆盖范围有限。未来研究将进一步结合临界氯离子浓度、扩散系数等指标，以拓展物理约束并提升模型的工程适用性。

此外，为确保模型性能评估的稳健性，本工作采用 10 折交叉验证对数据集进行划分。在每一折中分别训练 GPR 模型，并评估其在对应测试集上的预测效果。该方法能够有效降低由样本划分带来的偶然性影响，从而增强模型的鲁棒性。

## 2　结果与讨论

### 2.1　模型性能对比

为模拟不同工程场景下的信息可获得性，设置 3 组输入特征：

第 1 组：包含全部可用特征变量，涵盖材料、环境与电化学参数，用于刻画模型在完备信息条件下的性能上限；

第 2 组：仅包含配合比参数与[Cl⁻]/[OH⁻]，用于模拟早期设计阶段或仅具备环境侵蚀强度指标时的应用场景，探索在缺乏电化学监测条件下的可用性；

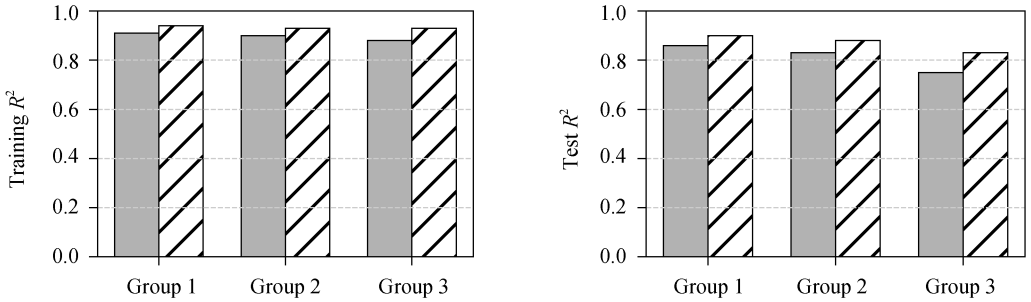
第 3 组：在全特征基础上，结合特征重要性分析与腐蚀机理筛选后的关键变量组合，包括电阻率、[Cl⁻]/[OH⁻]、水泥用量、腐蚀电位和孔隙率等，以构建兼顾精度与简洁性的优化模型，更贴合实际监测与运维需求。通过这种分层设计，可以定量揭示信息减少时模型预测精度的退化规律，并为不同场景下的数据采集与监测策略提供参考。

需要强调的是，所采用的输入特征中包含若干电化学参数，例如腐蚀电位、极化电阻以及[Cl⁻]/[OH⁻]。这些特征是基于腐蚀机理的关键变量选择，其中[Cl⁻]/[OH⁻]是决定钢筋钝化膜稳定性的根本条件，而其他电化学参数则直接反映腐蚀过程的动力学状态。通过引入这类特征，模型不仅能够利用统计规律进行预测，还能保持与电化学腐蚀机理的一致性，从而增强其物理合理性与可解释性。

在工程应用层面，本工作模型具备灵活性。在高价值工程(如跨海大桥、核电厂)中，可以通过定期监测或布设传感器获取[Cl⁻]/[OH⁻]和少量电化学参数，使模型直接服务于耐久性监测与寿命预测；而在一般工程场景中，则可结合氯离子扩散模型、材料与环境条件以及服役时间，近似推算钢筋表面的[Cl⁻]/[OH⁻]，从而减少对现场电化学测试的依赖。由此可见，[Cl⁻]/[OH⁻] 及相关电化学参数在模型中起到了连接宏观条件与微观腐蚀状态的关键作用，使得本工作提出的方法能够兼顾预测精度与应用便

利性。

由图 1 可见，在所有特征组合下，物理先验 GPR 均优于常规 GPR，表现出更高的预测精度和更好的稳定性。特别是在第 3 组仅包含少量关键特征的情况下，物理先验 GPR 仍能保持较高 $R^2$ 的水平，显示出良好的泛化能力。这表明物理先验的引入不仅提升了模型对腐蚀机理知识的利用程度，也显著增强了模型在特征维度受限条件下的鲁棒性和适应性。




(a) Training $R^2$ (b) Test $R^2$

Gray bars represent conventional GPR, and hatched white bars represent GPR with physics.Groups 1–3 represent three feature groups defined in Section 2.1.

图 1 3 组模型的性能对比结果

Fig. 1 Performance comparison of three groups

由表 2 可以进一步量化，物理先验 GPR 在 3 组输入特征配置下均取得了最优或接近最优的预测性能。在组合一和组合二中，其测试集 $R^2$ 分别为 0.90 和 0.88，均为所有模型中最高；在组合三(特征筛选后变量较少)中，其测试集 $R^2$ 仍达到 0.83，明显优于普通 GPR(0.75)、人工神经网络(ANN，0.66)、随机森林(RF，0.69)和极端梯度提升(XGBoost，0.69)。ANN 和 XGBoost 在小样本条件下测试精度下降更为明显。总体而言，物理先验 GPR 在预测精度和稳定性方面均表现突出，尤其在低维特征情形下优势更为明显，同时，GPR 独特的不确定性量化功能能够在预测的同时提供置信区间，为腐蚀风险分级与工程决策提供额外支持，这是其他模型难以直接实现的。

**表 2 不同模型在 3 组输入特征下的预测性能对比**

**Table 2 Comparison of ML models in three feature configurations**

| Methods | $R^2$ (Training/Test) | | |
|---|---|---|---|
| | Group 1 | Group 2 | Group 3 |
| ANN | 0.87 / 0.80 | 0.83 / 0.74 | 0.75 / 0.66 |
| RF | 0.94 / 0.83 | 0.87 / 0.71 | 0.83 / 0.69 |
| XGBoost | 0.89 / 0.85 | 0.89 / 0.83 | 0.75 / 0.69 |
| Plain GPR | 0.91 / 0.86 | 0.90 / 0.83 | 0.88 / 0.75 |
| GPR with physics | 0.94 / 0.90 | 0.93 / 0.88 | 0.93 / 0.83 |

基于随机森林特征重要性分析的筛选方法，提取了电阻率、孔隙率、腐蚀电位、$[Cl^-]/[OH^-]$以及水泥用量 5 个关键变量，构成关键特征集。与全特征模型相比，筛选模型在测试集上的精度有所下降，主要原因包括：一是筛选不可避免地舍弃了部分弱相关变量及潜在交互信息，从而削弱了模型的拟合能力；二是在有限的 180 组样本下，一些复杂的交互效应仅能在全特征条件下被捕捉。

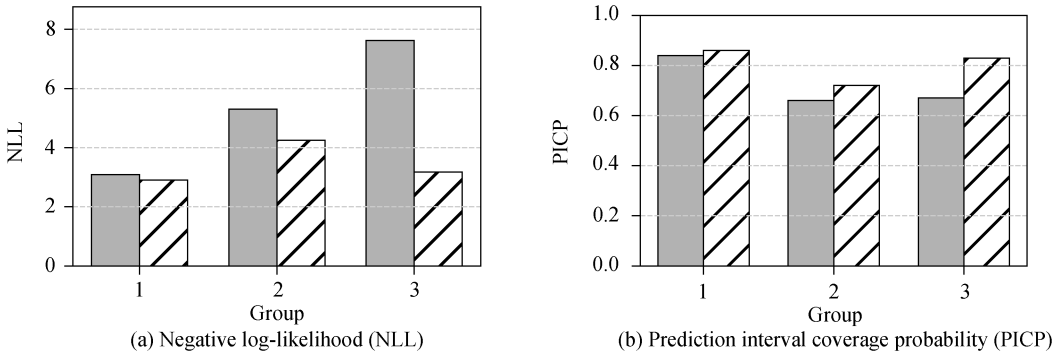
为系统比较模型在不确定性建模与置信度表达方面的性能，本工作评估了常规 GPR 与物理先验 GPR 在 3 组不同输入特征配置下的表现。图 2 为 2 类模型在负对数似然(NLL)和预测区间覆盖率(PICP)2 个指标上的对比结果。

1) NLL 指标分析

在所有特征组合中，物理先验 GPR 的 NLL 值均显著低于普通 GPR，表明其在输出概率分布建模方面具有更强的鲁棒性与可靠性。尤其是在第 3 组仅包含少量关键特征的情形下，普通 GPR 的 NLL 值升高至约 7，显示其在输入信息受限时的不确定性刻画能力显著下降；相比之下，物理先验 GPR 的 NLL 仍维持在约 4，展现出更优的泛化能力与预测稳定性。特别地，NLL 值越小表示模型在概率预测上的拟合越充分。因此，物理先验 GPR 的优势在低维特征条件下更加突出。

2) PICP 指标分析

在所有特征组合中，物理先验 GPR 的 PICP 值均高于普通 GPR，尤其在第二组与第三组中，其预测区间对真实值的覆盖率优势更为显著。该结果表明，物理先验 GPR 所生成的置信区间更具可信度，能够更准确地反映模型预测结果的不确定性。需要




(a) Negative log-likelihood (NLL)　(b) Prediction interval coverage probability (PICP)

Groups 1–3 represent the three feature groups defined in Section 2.1.

图 2　不同特征组合下常规 GPR 与物理先验 GPR 的比较

Fig. 2　Comparison of conventional GPR and GPR with physics under different feature groups

指出的是，PICP 值越接近 1，说明模型给出的预测区间越可靠，因此，物理先验 GPR 在不确定性刻画方面的优势在低维特征条件下表现得尤为突出。

综上所述，物理先验的引入不仅显著提升了模型的预测精度，同时增强了其在不确定性表达方面的能力与整体可信度。在输入特征受限或数据噪声较大的实际工程环境中，物理先验 GPR 展现出更强的适应性与应用潜力。

考虑到工程实践中通常更倾向于使用变量数量较少、获取成本较低的输入特征(如配合比参数与可测量的电化学指标)，本工作最终选取第二组特征配置作为后续分析与模型应用的基础输入方案。该配置在预测性能、数据获取成本与工程实用性之间实现了良好的平衡。

### 2.2　数据规模与模型可靠性检验

为验证 180 组样本是否足以支撑 GPR 模型的可靠性，本工作对不同训练样本规模下的模型性能进行了对比。我们随机抽取 20 至 180 个样本进行训练，并采用 10 折交叉验证计算验证集 $R^2$ (即将数据划分为 10 个子集，轮流以其中 1 个作为验证集，取 10 次结果的平均值) 。结果见表 3。

**表 3　不同训练样本规模下普通 GPR 与物理先验 GPR 的验证性能**

**Table 3　Validation performance of plain GPR and GPR with physics under different sample sizes ($R^2$)**

| Training samples | Plain GPR | GPR with physics |
|---|---|---|
| 20 | 0.50 | 0.54 |
| 40 | 0.69 | 0.76 |
| 80 | 0.75 | 0.81 |
| 120 | 0.80 | 0.86 |
| 160 | 0.83 | 0.88 |
| 180 | 0.83 | 0.88 |

由表可见，随着样本规模的增加，模型性能逐步收敛。当样本数达到 160 至 180 时，验证集 $R^2$ 已趋于收敛并稳定在 0.83~0.88 区间，说明在该低维特征条件下，180 组样本已足以保证 GPR 模型的稳定性与可靠性。此外，物理先验 GPR 在各样本规模下均优于普通 GPR，在小样本阶段提升约 5%~7%，在全样本条件下仍保持约 0.05 的增益，进一步支持所提出方法在现有数据条件下的稳定性与鲁棒性。

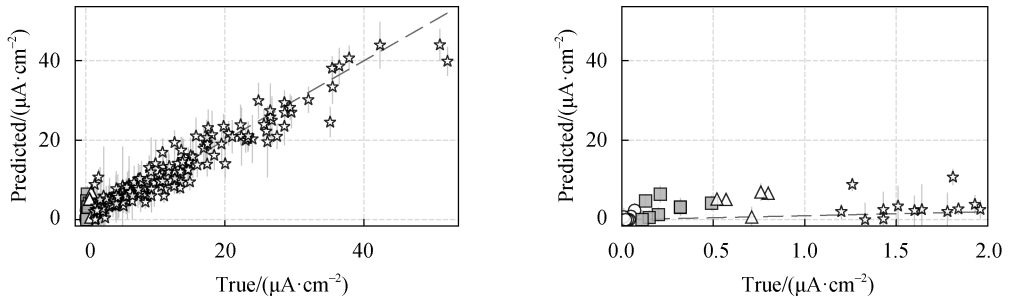
### 2.3　腐蚀等级划分与混淆矩阵分析

2.3.1　腐蚀等级划分　　为更贴近工程实践中对钢筋腐蚀风险的分级管理需求，本工作参考实验室与现场实测研究结果[18–20]，将腐蚀速率进行划分。当钢筋处于钝态时，腐蚀速率通常低于 0.05~0.10 μA/cm$^2$；处于活化状态时，腐蚀速率常见于 0.1~1.0 μA/cm$^2$ 区间；而大于 1.0 μA/cm$^2$ 的值相对少见，往往对应严重腐蚀情形。因此，本工作采用 0.1、0.5 μA/cm$^2$ 和 1.0 μA/cm$^2$ 作为腐蚀速率等级的划分阈值，既与已有研究结果一致，也具有明确的工程意义。低于 0.1 μA/cm$^2$ 表示钢筋处于稳定钝态，可常规巡检；0.1~1.0 μA/cm$^2$ 对应常见的腐蚀发展阶段，需加强监测；大于 1.0 μA/cm$^2$ 则提示高风险，建议采取干预或修复措施。因此，本工作将腐蚀速率(单位：μA/cm$^2$)划分为以下 4 个等级：

- 低腐蚀等级：< 0.1
- 中低腐蚀等级：0.1~0.5
- 中高腐蚀等级：0.5~1.0
- 高腐蚀等级：> 1.0

基于上述等级划分，图 3 展示了模型预测值与真实腐蚀速率等级之间的对应关系。图 3 左图显示了全局范围内预测值与真实值的分布情况，并绘制

了 95%置信区间，用以评估模型在各腐蚀等级区间的整体预测性能。图 3 右图对腐蚀速率在 0~2 μA/cm$^2$ 范围内的低腐蚀区域进行了局部放大，以突出模型在微弱腐蚀状态下的分辨能力。




(a) Full-range results　　(b) Zoomed-in results for the range of 0–2 μA/cm$^2$

Circles, squares, triangles and stars represent low, low–medium, medium–high, and high levels, respectively.

图 3　模型在不同腐蚀等级下的预测结果（真实值与预测值对比）

Fig. 3　Comparison of the model's predicted results at different corrosion levels (True versus Predicted)

由图 3 可以看出，大部分样本点分布接近对角线，说明模型预测结果与真实腐蚀速率等级之间具有较高的一致性，且在各等级区间均表现出良好的拟合能力。置信区间在高腐蚀速率区间呈现显著扩张趋势，反映出模型在该区间内存在较大的不确定性；而在低腐蚀速率区间，大部分样本预测波动较小，输出更为稳定。局部放大图清晰地展示了模型在低腐蚀等级下的预测结果具有较强的聚集性，等级边界明确，表明其在微弱腐蚀状态识别方面具备良好的分辨能力。该特性对于工程实践中关注早期腐蚀检测的应用场景具有重要价值。基于上述分析，本工作进一步通过混淆矩阵分析定量评估模型在各腐蚀等级下的分类性能，以便更全面反映预测结果的可靠性。

2.3.2　混淆矩阵分析　　在上述等级划分的基础上，混淆矩阵(见表 4) 进一步量化了模型在腐蚀等级划分任务中的分类性能，主要观察结果如下：

1) 高腐蚀等级 (>1 μA/cm$^2$) 识别准确率较高。

共计 162 个高腐蚀样本中，绝大多数被模型准确分类，仅有少数出现误判，表明模型在识别严重腐蚀情形方面具有良好的判别能力，具备较强的实用价值。

**表 4　不同腐蚀等级的真实值与预测结果之间的混淆矩阵**

**Table 4　Confusion matrix between the true values and predicted results for different corrosion levels**

| Class | Predicted: low | Predicted: low–medium | Predicted: medium–high | Predicted: high |
|---|---|---|---|---|
| True: low | 7 | 0 | 0 | 3 |
| True: low–medium | 1 | 0 | 0 | 4 |
| True: medium–high | 0 | 0 | 0 | 3 |
| True: high | 1 | 1 | 1 | 159 |

2) 中低与中高等级之间存在一定混淆。

部分中低(0.1~0.5 μA/cm$^2$)与中高(0.5~1.0 μA/cm$^2$)等级的样本被误判为“高腐蚀”，反映出模型在中等腐蚀强度区间的区分能力尚有提升空间。造成该现象的原因为样本分布不均衡，导致模型在训练过程中对高频类别更敏感；中间等级样本的特征边界不显著，增加了判别难度。

3) 低腐蚀等级识别效果较好，但存在一定偏差。

在低腐蚀区间(<0.1 μA/cm$^2$)内，模型成功识别了 7 个样本，但有 3 个被误分类为“高腐蚀”，表明模型在极低腐蚀水平下的置信度可能较低，或存在预测值偏移问题。该误差与测量精度、特征噪声或样本分布有关，需进一步分析其成因以提升模型在微弱腐蚀状态下的稳定性。

为进一步评估模型在腐蚀等级划分任务中的表现，本工作对不同等级的分类性能进行了分析，采用的指标包括精确率(衡量预测结果的准确性)、召回率(用于衡量高风险样本的覆盖程度)和 $F1$ 值(综合平衡指标)，并计算总体准确率和加权平均 $F1$ 值。结果表明，高腐蚀等级作为工程最关切的风险区间，模型的召回率为 0.98、精确率为 0.94、$F1$ 值为 0.96，说明其对高风险的识别非常可靠。整体来看，模型的总体准确率为 0.92，加权平均 $F1$ 值为 0.91，显示出在全样本层面具有较强的预测稳健性。中间等级样本数量相对有限，其预测性能相对较弱，但这类情况在不均衡分类任务中较为常见，并不影响模型在关键高腐蚀风险区间的有效性与工程应用价值。

综上所述，所建模型在腐蚀等级区分任务中展现出良好的整体趋势识别能力，尤其在高腐蚀等级的识别方面表现出较高的准确性与稳定性。然而，

在中间等级(中低与中高)样本中仍存在一定程度的误分类现象，表明模型在过渡区间的判别能力尚有优化空间。后续工作中可考虑通过样本增强(如合成少数类样本)或引入等级敏感的分类损失函数，以进一步提升模型在腐蚀等级划分任务中的整体性与鲁棒性。

### 2.4　不确定性驱动的工程决策策略

在实际工程管理中，单纯依赖腐蚀速率的单点估计值不足以全面反映潜在风险，尤其在模型预测存在不确定性的情况下，可能导致对风险的过度自信或对关键安全隐患的忽视。为提升模型输出在工程应用中的决策有效性与安全冗余能力，本工作提出一种融合预测不确定性信息的响应策略，用以指导后续管理与维护措施的制定。

该策略基于模型预测的腐蚀等级(高或低)及其对应置信区间的宽度(反映预测不确定性水平)进行分类，并将每个预测样本划分为以下 4 类，对应的工程响应建议详见表 5：

● 高腐蚀等级 + 低不确定性：建议立即采取防护措施或制定维修方案(紧急应对)。

● 高腐蚀等级 + 高不确定性：建议开展复核检测或补充评估，以降低误判风险(谨慎使用)。

● 低腐蚀等级 + 高不确定性：建议继续观测，并结合其他检测手段综合评估风险(确认复核)。

● 低腐蚀等级 + 低不确定性：可接受当前状态，维持常规监测频率(常规监测)。

该分类逻辑实现了从“预测数值”向“工程响应建议”的有效转化，增强了模型在实际管理决策中的可操作性与实用性。

**表 5　基于腐蚀等级与预测不确定性的工程响应决策规则**

**Table 5　Engineering decision strategy based on predicted corrosion level uncertainty**

| Decision category | Condition | Recommendation |
|---|---|---|
| High risk (Direct action) | High corrosion + Low uncertainty | Immediate intervention |
| High risk (Review) | High corrosion + High uncertainty | Manual review |
| Medium–low risk (Monitor) | Low corrosion + High uncertainty | Continuous monitoring |
| Medium–low risk (Accept) | Low corrosion + Low uncertainty | Accept current condition |

为量化模型预测结果的不确定性水平，本工作采用 95%置信区间的半宽度作为衡量指标。当半宽度超过阈值 $T = 4\ \mu A/cm^2$ 时，即判定该预测结果具有较高的不确定性，并在工程应用中建议采取更为保守的响应策略。该阈值的设定基于统计分布与灵敏度分析的综合结果(见表 6)。从分布特征来看，$4\ \mu A/cm^2$ 对应不确定性分布的 P82(即约 82%的样本不确定性低于该值)，处于尾部高不确定性范围；从灵敏度表现来看，阈值从 $3\ \mu A/cm^2$ 提升至 $4\ \mu A/cm^2$ 时，模型性能($R^2$)由 0.831 提升至 0.865，有效样本数增加 30，而继续提高阈值(如 4→5 $\mu A/cm^2$) 仅带来极小的性能提升($R^2$ 增幅 0.006)，边际收益递减。因此，$4\ \mu A/cm^2$ 可视为性能快速提升区间的下限拐点，在排除高不确定性预测与保留足够样本之间取得合理平衡，既具备统计学依据，也符合工程实践中的安全性要求。

**表 6　不同阈值下的不确定性灵敏度分析**

**Table 6　Sensitivity analysis of predictive uncertainty under different thresholds**

| Threshold/($\mu A\cdot cm^{-2}$) | Percentile/% | Valid samples | $R^2$ (Low-uncertainty subset) |
|---|---|---|---|
| 2 | 35.6 | 64 | 0.722 |
| 3 | 65.6 | 118 | 0.831 |
| 4 | 82.2 | 148 | 0.865 |
| 5 | 88.9 | 160 | 0.871 |
| 6 | 93.9 | 169 | 0.876 |
| 8 | 98.3 | 177 | 0.876 |
| 10 | 100.0 | 180 | 0.876 |

基于上述决策逻辑，对全部预测样本进行了响应类别划分与统计分析，结果如图 4 所示。具体表现如下：

1) “高风险(紧急应对)” 类别共包含 103 个样本，即预测结果为高腐蚀等级，且模型置信度较高，可直接用于推动相关防护或维修措施。

2) “高风险(谨慎使用)” 类别包含 62 个样本，虽然预测为高腐蚀等级，但置信区间较宽，存在较大不确定性，建议通过人工复核或补充检测以提高决策可靠性。

3) 中低风险类别中，“接受” 类样本为 13 个，“观察” 类样本为 2 个，数量相对较少。

4) 整体来看，约 36%的样本被归入“复核”或“观察”类别，显示引入不确定性信息在风险控制中具有显著补充价值。

该策略充分体现了“可信人工智能”在工程风险管理中的应用优势：不仅关注预测值本身，还综合考虑模型对自身预测的置信水平，从而实现更稳健、更安全的响应机制，避免因过度依赖单一预测值而引发的误判风险。

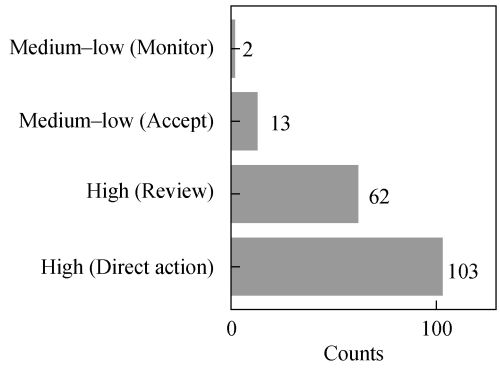



图 4 基于不确定性驱动的工程响应策略下的样本分布统计
Fig. 4 Sample distribution statistics based on uncertainty-driven engineering decision strategy

### 2.5 工程应用示例

为演示本工作提出的风险分级策略在工程情境中的迁移性，本工作构建了一个假想案例。场景设定为沿海隧道衬砌结构，长期处于高氯盐和高湿度环境。考虑到钢筋腐蚀主要发生在钢筋与胶凝材料包裹体界面处，模型输入参数被视作对该关键区域材料与环境特性的描述。基于假设的现场检测与实验室辅助分析(实验室新样本)，获得如下参数：水泥比例 60%(质量分数)、矿渣比例 40%(质量分数)、水胶比 0.45、内掺氯盐含量 1.2%(质量分数)、孔隙率 0.19、环境相对湿度 97%、腐蚀电位−640 mV、电阻率 125 Ω·m、$[Cl^-]/[OH^-]$为 820。采用最少特征组合进行预测，模型输出腐蚀速率为 18.4 μA/cm$^2$，95%置信区间为[13.8, 24.6] μA/cm$^2$，对应高风险等级(>1 μA/cm$^2$)。该案例为假想情境，旨在演示策略的可操作性，而非声称已在现场完成验证。未来工作将进一步结合参数转换与现场实测数据校准，推动该策略在混凝土结构耐久性评估与维护决策中的应用。

## 3 结论

本工作提出了一种融合物理先验与不确定性量化的高斯过程回归(GPR)框架，用于钢筋腐蚀速率建模与耐久性评估。主要结论如下：

1) 构建了物理约束与数据驱动融合的稳健预测模型。通过将电化学参数 $\lg([Cl^-]/[OH^-])$引入 GPR 均值函数，实现了物理一致性与小样本适应性的统一。该模型在有限数据条件下表现出优异的泛化能力，为复杂环境下的钢筋腐蚀预测提供了可靠工具。

2) 识别了腐蚀状态转变的关键临界区并提出实验优化方向。模型结果显示，当 $\lg([Cl^-]/[OH^-])\in[0.2, 0.6]$时预测不确定性显著升高，揭示了钢筋由钝化态向活化态转变的敏感区间。该发现不仅深化了对腐蚀起始机理的理解，也为实验设计提供了依据，提示应优先补充该“高敏感–高不确定”区间的数据，以提升模型精度与可靠性。

3) 形成了融合不确定性量化的风险分级策略并验证工程适用性。提出基于预测值及置信区间的联合判别方法，建立了从常规监测到紧急干预的 4 级响应机制。案例验证表明，该策略能够准确识别需“复核”的边界样本，降低误判风险，为基础设施的耐久性管理与维护决策提供了科学且可操作的支持。

整体而言，该框架为钢筋腐蚀预测与工程耐久性提供了一条新的方法路径。

**参考文献**:

[1] 刘志勇, 缪昌文, 周伟玲, 等. 迁移性阻锈剂对混凝土结构耐久性的保持和提升作用[J]. 硅酸盐学报, 2008, 36(10): 1494–1500.
LIU Zhiyong, MIAO Changwen, ZHOU Weiling, et al. J Chin Ceram Soc, 2008, 36(10): 1494–1500.

[2] 吴智深, 刘加平, 邹德辉, 等. 海洋桥梁工程轻质、高强、耐久性结构材料现状及发展趋势研究[J]. 中国工程科学, 2019, 21(3): 31–40.
WU Zhishen, LIU Jiaping, ZOU Dehui, et al. Strateg Study CAE, 2019, 21(3): 31–40.

[3] OTIENO M B, BEUSHAUSEN H D, ALEXANDER M G. Modelling corrosion propagation in reinforced concrete structures–A critical review[J]. Cem Concr Compos, 2011, 33(2): 240–245.

[4] ZHAO R Q, LI C F, GUAN X M. Advances in modeling surface chloride concentrations in concrete serving in the marine environment: A mini review[J]. Buildings, 2024, 14(6): 1879.

[5] ZHAO Y F, LI H, HU X, et al. Evaluation on steel corrosion risk of seawater sea sand concrete with low water-to-binder ratio by electrochemical measurement[J]. Constr Build Mater, 2024, 417: 135326.

[6] SAMSON G, DEBY F, GARCIAZ J L, et al. An alternative method to measure corrosion rate of reinforced concrete structures[J]. Cem Concr Compos, 2020, 112: 103672.

[7] YU B, YANG L F, WU M, et al. Practical model for predicting corrosion rate of steel reinforcement in concrete structures[J]. Constr Build Mater, 2014, 54: 385–401.

[8] WONG P T, LAI W W, POON C S. Classification of concrete corrosion states by GPR with machine learning[J]. Constr Build Mater, 2023, 402: 132855.

[9] BAMSHAD O, JAMHIRI B, HABIBI A, et al. Predicting corrosion of recycled aggregate concrete under sulfuric acid rain using machine learning and uncertainty analysis[J]. Constr Build Mater, 2024, 438: 137146.

[10] SALAMI B A, RAHMAN S M, OYEHAN T A, et al. Ensemble machine learning model for corrosion initiation time estimation of embedded steel reinforced self-compacting concrete[J]. Measurement, 2020, 165: 108141.

[11] COELHO L B, ZHANG D W, VAN INGELGEM Y, et al. Reviewing machine learning of corrosion prediction in a data-oriented perspective[J].

NPJ Mater Degrad, 2022, 6: 8.

[12] 梁梓豪, 应宗权, 刘梅梅, 等. 基于 Stacking 集成模型融合的钢筋混凝土锈胀开裂预测方法[J]. 中国腐蚀与防护学报, 2024, 44(6): 1601–1609.
LIANG Zihao, YING Zongquan, LIU Meimei, et al. J Chin Soc Corros Prot, 2024, 44(6): 1601–1609.

[13] 张文生, 曹傅荔, 郅晓, 等. 机器学习方法用于水泥基材料的研究进展[J]. 硅酸盐学报, 2024, 52(11): 3617–3630.
ZHANG Wensheng, CAO Fuli, ZHI Xiao, et al. J Chin Ceram Soc, 2024, 52(11): 3617–3630.

[14] JI H D, YE H L. Machine learning prediction of corrosion rate of steel in carbonated cementitious mortars[J]. Cem Concr Compos, 2023, 143: 105256.

[15] JI H D, LYU Y H, TIAN Z S, et al. Assessment of corrosion probability of steel in mortars using machine learning[J]. Reliab Eng Syst Saf, 2025, 253: 110535.

[16] DERINGER V L, BARTÓK A P, BERNSTEIN N, et al. Gaussian process regression for materials and molecules[J]. Chem Rev, 2021, 121(16): 10073–10141.

[17] TIAN Z S, KANG X J, JI H D, et al. Quantitative relationship between microstructure of steel-concrete interface and chloride-induced corrosion rate of steel in unsaturated cementitious materials[J]. Cem Concr Res, 2025, 188: 107736.

[18] MILLARD S G, LAW D, BUNGEY J H, et al. Environmental influences on linear polarisation corrosion rate measurement in reinforced concrete[J]. NDT E Int, 2001, 34(6): 409–417.

[19] ANDRADE C, MARTÍNEZ I. Techniques for measuring the corrosion rate (polarization resistance) and the corrosion potential of reinforced concrete structures[M]//Non-Destructive Evaluation of Reinforced Concrete Structures. Amsterdam: Elsevier, 2010: 284–316.

[20] ANDRADE C, ALONSO C. On-site measurements of corrosion rate of reinforcements[J]. Constr Build Mater, 2001, 15(2/3): 141–145.

# Intelligent Prediction and Risk Classification of Steel Corrosion in Cementitious Materials

*LYU Yuhui*[1], *JI Haodong*[1], *YU Qianqian*[2], *YE Hailong*[1]
(1. Department of civil engineering, The University of Hong Kong, Hong Kong 999077, China;
2. Department of Building Engineering, College of Civil Engineering, Tongji University, Shanghai 200092, China)

## Extended Abstract

**Introduction** Steel corrosion is a critical deterioration phenomenon in concrete structures, particularly under harsh environmental conditions involving carbonation and chloride ingress. Accurate corrosion rate prediction is essential for structural life assessment and maintenance planning. The existing methods include empirical and electrochemical models; they suffer from a limited adaptability or a high computational complexity. Machine learning (ML) approaches show promise but typically focus on point estimates, while neglecting prediction uncertainty, which is crucial for reliable engineering decision-making. This study was to develop a Gaussian Process Regression (GPR) framework that could integrate physical prior knowledge with uncertainty quantification for steel corrosion rate prediction in cementitious materials, addressing two key questions, *i.e.*, 1) how to enhance GPR accuracy and robustness for corrosion prediction, and 2) whether GPR uncertainty outputs could effectively support corrosion risk classification and engineering decisions.

**Methodology** The study utilized 180 experimental datasets from mortar samples embedded steel exposed to 6 relative humidity (RH) conditions after carbonation treatment. Each dataset contained 15 corrosion-related features categorized as mixture parameters (*i.e.*, cement, slag, fly ash, and silica fume proportions, water-to-binder ratio, and chloride content), material properties (*i.e.*, free chloride content, pH value, and porosity), environmental parameters (*i.e.*, RH, saturation, and water content), and electrochemical parameters (*i.e.*, corrosion potential, electrical resistivity, and $[Cl^-]/[OH^-]$).

The GPR model incorporated physical prior knowledge *via* setting the mean function as a linear relationship with a logarithmic chloride-to-hydroxide ratio, *i.e.*, $m(x) = a \cdot \lg([Cl^-]/[OH^-]) + b$, based on the electrochemical corrosion theory. A Gaussian radial basis function served as a covariance function to capture smooth corrosion behavior variations.

**Results and discussion** The performance evaluation using 10-fold cross-validation demonstrates that the physical prior GPR consistently outperforms conventional GPR across all the feature combinations. The model shows a superior prediction accuracy (*i.e.*, higher $R^2$), a better uncertainty quantification (*i.e.*, lower negative log-likelihood), and more reliable confidence intervals (*i.e.*, higher prediction interval coverage probability).

The corrosion rates are classified into four levels, *i.e.*, low (<0.1 μA/cm$^2$), low-medium (0.1–0.5 μA/cm$^2$), medium-high (0.5–1.0 μA/cm$^2$), and high (>1.0 μA/cm$^2$). The model achieves a high accuracy in identifying severe corrosion conditions, with most of the 159 high corrosion samples correctly classified, though some confusion exists between intermediate levels.

An uncertainty-driven decision strategy is developed using confidence interval width (threshold: 4 μA/cm$^2$) combined with predicted corrosion levels to generate four response categories, *i.e.*, immediate intervention (*i.e.*, high corrosion + low uncertainty), manual review (*i.e.*, high corrosion + high uncertainty), continuous monitoring (*i.e.*, low corrosion + high uncertainty), and accept current condition (*i.e.*, low corrosion + low uncertainty). Approximately 36% of samples are flagged for review or monitoring due to the high uncertainty, demonstrating a value of uncertainty assessment in risk management.

**Conclusions** This work presented a framework combining physical knowledge with probabilistic modeling for trustworthy AI-based corrosion prediction. Key contributions could include 1) effective integration of electrochemical mechanisms into GPR architecture, improving model robustness and interpretability; 2) development of a joint corrosion level and uncertainty classification system; 3) establishment of an uncertainty-based engineering decision framework; and (4) validation that uncertainty quantification effectively could identify potentially problematic predictions requiring additional review.

The framework could advance AI applications in structural health monitoring from simple prediction tools to comprehensive decision support systems, providing a foundation for reliable, interpretable infrastructure maintenance strategies. The approach demonstrated a promising potential for enhancing structural safety through improved corrosion risk assessment and management in challenging service environments.

**Keywords** cementitious materials; steel corrosion; gaussian process regression; engineering decision-making