

DOI: 10.19951/j.cnki.1672-9331.20231201001

文章编号: 1672-9331(2024)01-0088-12

引用格式: 宋超, 赵腾远. 黏土路基回弹模量预测及贝叶斯模型选择研究[J]. 长沙理工大学学报(自然科学版), 2024, 21(1): 88-99.

Citation: SONG Chao, ZHAO Tengyuan. Prediction of the resilient modulus of subgrade clay and Bayesian model class selection[J]. J Changsha Univ Sci Tech (Nat Sci), 2024, 21(1): 88-99.

黏土路基回弹模量预测及贝叶斯模型选择研究

宋超, 赵腾远

(西安交通大学 人居环境与建筑工程学院, 陕西 西安 710049)

摘要:【目的】确定黏土路基回弹模量的最优估计模型, 实现黏土路基回弹模量的准确预测。【方法】采用贝叶斯高斯过程回归方法, 建立了路基土的围压、偏应力、含水率以及干重度与路基回弹模量之间的定量关系, 实现了高斯过程回归参数的准确估计与最优影响因子组合的客观选择, 在模型的复杂度与拟合程度之间达到了自动平衡。【结果】基于所提出的贝叶斯高斯过程回归方法可准确预测路基的回弹模量, 所选最优模型的决定系数(R^2)和平均绝对百分误差(R_{MAPE})分别达到了0.99和1.51%, 与全变量模型的预测性能几乎相同。在100次随机试验中, 最优模型被选择的比率达到了88%。【结论】所提出的贝叶斯高斯过程回归方法不仅可以通过路基土相关物理力学参数准确预测路基的回弹模量, 还可以有效剔除冗余输入变量, 在保证模型拟合程度的情况下, 降低了模型的复杂度, 这对模型的应用与推广具有重要意义。

关键词: 路基工程; 高斯过程回归; 数据驱动; 非参模型; 贝叶斯理论

中图分类号: U416.1、U412.6

文献标志码: A

Prediction of the resilient modulus of subgrade clay and Bayesian model class selection

SONG Chao, ZHAO Tengyuan

(School of Human Settlements and Civil Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: [Purposes] The paper aims to determine the optimal model for predicting the resilient modulus of subgrade soil and predict the resilient modulus of subgrade soil accurately. [Methods] The fully Bayesian Gaussian process regression (fB-GPR) approach is proposed in this study to establish the relationship between the resilient modulus of subgrade soil (M_R) and the confining pressure (σ_3), deviator pressure (σ_d), moisture content (w), dry density (ρ_d). The parameters of the GPR model are estimated accurately and the optimal model for the prediction of the resilient modulus of subgrade soil is determined objectively using the proposed approach in this study. The optimal model reaches a balance between the complexity and fitting degree. [Findings] The findings show that the resilient modulus of subgrade soil can be predicted accurately using the proposed approach in this study. The coefficient of determination (R^2) and the mean absolute percentage error (R_{MAPE}) of the optimal GPR model reach 0.99 and 1.51%, respectively, which is similar to the full model. In 100 experiments, the percentage of the optimal model being selected

收稿日期: 2023-12-01; 修回日期: 2023-12-25; 接受日期: 2024-01-01

基金项目: 国家自然科学基金资助项目(42107204)

通信作者: 赵腾远(1988—)(ORCID: 0000-0002-7007-094X), 男, 教授, 主要从事岩土工程贝叶斯模型选择、可靠度设计与分析、数字岩土工程与机器学习交叉等方面的研究。E-mail: tyzhao@xjtu.edu.cn

投稿网址: <http://cslgxbzk.csust.edu.cn/cslgdxxbzk/home>

reaches 88%. [Conclusions] The proposed fB-GPR approach not only accurately predict the resilient modulus using the easily available indices of subgrade soil but also effectively eliminate the redundant input variables. The proposed approach reduces complexity without compromising the fitting degree of the GPR model, holding significance for its application and promotion in the future.

Key words: subgrade engineering; Gaussian process regression; data-driven; non-parametric modeling; Bayesian framework

Foundation item: Project (42107204) supported by the National Natural Science Foundation of China

Corresponding author: ZHAO Tengyuan (1988—) (ORCID: 0000-0002-7007-094X), male, professor, research interest: Bayesian updating and model class selection in geotechnical engineering, reliability-based design and analysis, machine learning in data-centric geotechnics. E-mail: tyzhao@xjtu.edu.cn

0 引言

路基回弹模量反映了路面结构在外荷载作用下抵抗变形的能力,是路面结构设计的重要参数之一,尤其对于黏土路基,在外力作用下易发生较大变形,因此研究路基回弹模量的主要影响因素和变化规律具有重要的理论意义和应用价值。较高的路基回弹模量是路面结构正常运营的基础,而对其进行准确测试则是路基设计的关键^[1]。

传统试验法可以直接测得路基回弹模量的准确数值,具有快速、高效的特点,但无法考虑路基土性质、外荷载等因素对其演变规律的影响,容易导致运营过程中发生较大变形而影响道路的正常通行^[2]。经验公式法通过路基土的其他物理力学参数间接估计路基的回弹模量,一方面建立了路基土性质与路基回弹模量之间的定量关系,另一方面定量刻画了路基土性质对路基回弹模量影响的大小。例如,陈慧东等^[3]建立了路基动态回弹模量与加州承载比(California bearing ratio, CBR)之间的幂指数关系模型,结果表明未浸水CBR与回弹模量的关系曲线拟合良好;李长生^[4]建立了重载铁路粉质黏土填料回弹模量与路基土的干重度、含水率之间的非线性关系模型,结果表明该模型可准确估计路基在动应力作用下的弹性变形;KIM^[5]建立了路基回弹模量与八面体正应力、八面体剪应力之间的本构关系模型,通过地基土的应力状态实现回弹模量的间接估计,进而达到通过简单的试验测试获得路基土变形性质的目的。然而,由于方程形式的制约,经验公式法的适用范围通常较小,泛化能力较差,导致路基回弹模量估计值与实测值之间存在较大偏差。

近年来,人工智能技术高速发展,机器学习方法已逐步应用于公路工程、隧道工程、基础设施工程等领域。与经验公式法类似,机器学习方法可以通过建立路基回弹模量与其相关的物理力学参数之间的关系模型,进而实现路基回弹模量的预测。进一步讲,机器学习方法是一种数据驱动的方法,不再局限于固有的方程形式,模型更加灵活,适用范围更广,估计结果更加准确^[6]。例如,谭伟^[7]通过人工神经网络算法建立了路基动态回弹模量与含水率、压实度、细粒含量、塑性指数之间的关系模型,结果表明通过该模型估计的动态回弹模量与实测值偏差在5%左右,实现了路基动态回弹模量的准确估计;HEIDARIPANAH等^[8]采用支持向量机方法建立了路基动态回弹模量与围压、偏应力、含水率、干重度之间的定量关系,并探讨了不同核函数对预测结果的影响;KARDANI等^[9]收集了2 813组路基土的弹性模量以及物理力学参数数据,并通过递归增强、决策树、K近邻、随机森林4种机器学习算法建立了路基回弹模量与干容重、加权塑性指数、偏应力、围压、冻融循环次数、含水率之间的相关关系,实现了路基回弹模量的准确估计。然而,传统机器学习方法多为“黑箱模型”,解译性较差,且通常将所有已知物理力学参数作为输入变量。这一方面增加了试验时间、成本,另一方面模型中包含与回弹模量相关性较小的参数,容易导致模型的复杂度增加,泛化能力降低。基于皮尔逊相关系数^[10]、递归特征消除^[11]等方法尽管可以在一定程度上去除冗余输入变量,但皮尔逊相关系数法得到的是输入变量与目标值之间的线性相关系数,对非线性模型适用性较差;递归特征消除法则易将贡献较小的变量包含在最优模型中,导致模型过于复杂。

本文提出贝叶斯高斯过程回归(fully Bayesian Gaussian process regression, fB-GPR)方法来估计路基回弹模量,该方法不仅具有数据驱动、非参数预测、解译性强的特点,还可实现路基回弹模量的准确预测,剔除模型中的冗余变量,在预测模型的准确度与复杂度之间自动达到平衡,并给出了每个备选模型被选择的概率。另外,本文详细阐述了两种贝叶斯模型证据的估计方法,即基于多维对数正态的方法和Copula方法;首次基于马尔科夫链蒙特卡罗(Markov chain Monte Carlo, MCMC)样本对高斯过程回归模型的复杂度与拟合程度进行量化表征,定量描述了随着模型输入变量的增多,模型复杂度与拟合程度的变化趋势;探讨了当各个备选模型具有不同的先验概率时,本文所提方法模型选择结果的准确性,有力推动了机器学习方法在公路工程中的应用,因而具有重要的现实意义与实践价值。

1 基于高斯过程回归的路基回弹模量预测

高斯过程回归(Gaussian process regression, GPR)可定义为任意有限数量的具有联合高斯分布的随机变量的集合^[12]。高斯过程回归是一种数据驱动的机器学习方法,不但可以对路基回弹模量进行准确预测,而且具有清晰的解译性,预测结果可通过严格的数学推导过程得到,在处理小样本数据过程中具有显著优势。因此,本文采用高斯过程回归方法建立路基回弹模量 M_R 与路基土物理力学参数 $\mathbf{X}=(X_1, X_2, \dots, X_v)$ 之间的关系模型,即 $M_R=f(\mathbf{X})+\varepsilon$,式中, ε 是均值为0、标准差为 σ_ε 的高斯随机变量,表示模型预测残差; v 为输入变量的个数。本文中,以 $\mathbf{X}=(X_1, X_2, \dots, X_v)$ 作为输入变量,路基回弹模量($y=M_R$)作为输出变量。在高斯过程回归中,将模型输出 $y=M_R$ 当成高斯过程(GP),其完全由均值函数 $m(\mathbf{X})$ 和协方差函数 $\mathbf{K}_{\mathbf{x},\mathbf{x}}$ 确定,即 $M_R=GP(m(\mathbf{X}), \mathbf{K}_{\mathbf{x},\mathbf{x}}+\sigma_\varepsilon^2\mathbf{I})$,式中, $\mathbf{x}$ 为与 $\mathbf{X}$ 相对应的输入变量测量值, $\mathbf{I}$ 为与 $\mathbf{K}_{\mathbf{x},\mathbf{x}}$ 大小一致的单位矩阵。通常,假设均值向量 $m(\mathbf{X})$ 为0,协方差函数采用平方指数协方差函数^[13],即:

$$K_{i,j} = \sigma_f^2 \exp\left(-\frac{1}{2} \sum_{t=1}^v \frac{(x_{i,t} - x_{j,t})^2}{l_t^2}\right) + \sigma_n^2 \delta_{i,j} \quad (1)$$

式中: $K_{i,j}$ 为第 i 个预测值 $M_R^{(i)}$ 与第 j 个预测值 $M_R^{(j)}$ 之

间的协方差; σ_f^2 为核函数的波动程度; $\mathbf{x}_i$ 和 $\mathbf{x}_j$ 为 $\mathbf{X}$ 的第 i 组和第 j 组测量值,与 $M_R^{(i)}$ 和 $M_R^{(j)}$ 相对应; $x_{i,t}$ 和 $x_{j,t}$ 为 $\mathbf{x}_i$ 和 $\mathbf{x}_j$ 的第 t 个元素; σ_n^2 为路基回弹模量 M_R 测量误差的方差; $\delta_{i,j}$ 为克罗内克函数,当且仅当 $\mathbf{x}_i=\mathbf{x}_j$ 时, $\delta_{i,j}=1$; l_t 为第 t 个输入变量的相关长度,反映了路基回弹模量 $M_R^{(i)}$ 和 $M_R^{(j)}$ 沿 X_t 变量方向的相关性。通常,当 l_t 值较小时,第 t 个输入变量 X_t 与输出变量 M_R 之间具有较高的相关性,详见文献[12, 14]。因此,给定路基回弹模量($y=M_R$)与相对应的输入变量 $\mathbf{X}=(X_1, X_2, \dots, X_v)$ 的测量值,即可估计高斯过程回归模型参数 $\Theta=[l_t(t=1, 2, \dots, v), \sigma_f, \sigma_\varepsilon, \sigma_n]$,实现路基回弹模量的准确预测。给定新的输入变量 $\mathbf{x}^*$,即可预测对应的路基回弹模量。根据高斯分布的性质,待预测的 M_R 仍然服从高斯分布,其均值和协方差可分别表示为:

$$\begin{cases} \boldsymbol{\mu}^* = \mathbf{K}_{\mathbf{x}^*,\mathbf{x}} [\mathbf{K}_{\mathbf{x},\mathbf{x}} + \sigma_\varepsilon^2 \mathbf{I}]^{-1} \mathbf{y} \\ \boldsymbol{\Sigma}^* = \mathbf{K}_{\mathbf{x}^*,\mathbf{x}^*} - \mathbf{K}_{\mathbf{x}^*,\mathbf{x}} [\mathbf{K}_{\mathbf{x},\mathbf{x}} + \sigma_\varepsilon^2 \mathbf{I}]^{-1} \mathbf{K}_{\mathbf{x},\mathbf{x}}^T \end{cases} \quad (2)$$

式中: $\boldsymbol{\mu}^*$ 为与 $\mathbf{x}^*$ 相对应的路基回弹模量预测值; $\boldsymbol{\Sigma}^*$ 为预测的路基回弹模量的协方差,其对角线元素为预测的路基回弹模量的方差,反映了预测的不确定性的量; $\mathbf{K}_{\mathbf{x}^*,\mathbf{x}}$ 为 M_R 预测值与 M_R 测量值之间的协方差; $\mathbf{K}_{\mathbf{x}^*,\mathbf{x}^*}$ 为 M_R 预测值之间的协方差; $(\cdot)^T$ 表示矩阵转置; $[\mathbf{K}_{\mathbf{x},\mathbf{x}} + \sigma_\varepsilon^2 \mathbf{I}]^{-1}$ 为矩阵 $[\mathbf{K}_{\mathbf{x},\mathbf{x}} + \sigma_\varepsilon^2 \mathbf{I}]$ 的逆矩阵。由式(1)~(2)可知,高斯过程回归模型参数 $\Theta=[l_t(t=1, 2, \dots, v), \sigma_f, \sigma_\varepsilon, \sigma_n]$ 对于预测结果的准确性至关重要。因此,本文将在下一节对高斯过程回归模型的参数估计进行详细论述。

2 模型参数的贝叶斯估计及模型选择

2.1 基于贝叶斯理论参数估计

本节将详细阐述基于贝叶斯理论的高斯过程回归模型参数的估计方法。在贝叶斯理论中,优化后的模型参数通过后验概率密度函数反映,即^[15]:

$$p(\Theta|D) = \frac{p(D|\Theta)p(\Theta)}{\int p(D|\Theta)p(\Theta)d\Theta} \quad (3)$$

式中: D 表示获得的测量数据; $p(\Theta|D)$ 为高斯过程回归模型参数的后验概率密度函数; $p(\Theta)$ 为先验概率密度函数,反映了在未考虑测量数据之前对模型参数的理解与认知情况; $p(D|\Theta)$ 为似然函数,反映了基于高斯过程回归路基回弹模量的预测值

和测量值之间的拟合情况; $\int p(D|\Theta)p(\Theta)d\Theta$ 为常数,与 Θ 无关。

由此可知,合理确定先验概率密度函数 $p(\Theta)$

$$p(\Theta) = \prod_{i=1}^{v+3} p(\Theta_i) = \prod_{i=1}^{v+3} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) \sqrt{\nu\pi\sigma_i^2}} \left\{ \left[1 + \frac{1}{\nu} \frac{(\Theta_i - \mu_i)^2}{\sigma_i^2} \right]^{-\frac{\nu+1}{2}} + \left[1 + \frac{1}{\nu} \frac{(\Theta_i + \mu_i)^2}{\sigma_i^2} \right]^{-\frac{\nu+1}{2}} \right\}, \Theta_i > 0 \quad (4)$$

式中: μ_i 为位置参数; σ_i 为尺度参数,反映 M_R 的波动范围; ν 为自由度,反映half-t分布的尾部效应; $p(\Theta_i)$ 为第 i 个高斯过程回归模型参数的先验概率密度函数,其中 Θ 的前 v 个元素对应于 v 个相关长度 $l_i(i=1,2,\dots,v)$,后3个元素分别对应于波动程度 σ_r 、预测残差的标准差 σ_e 和测量误差的标准差 σ_n 。

似然函数 $p(D|\Theta)$ 可表示为:

$$p(D|\Theta) = \frac{1}{\sqrt{(2\pi)^{n_m} |K_{x,x} + \sigma_e^2 I|}} \exp\left[-\frac{1}{2} y^T (K_{x,x} + \sigma_e^2 I)^{-1} y\right] \quad (5)$$

式中: n_m 为用于训练的测量数据的组数; $|K_{x,x} + \sigma_e^2 I|$ 为矩阵 $[K_{x,x} + \sigma_e^2 I]$ 的行列式。

先验概率密度函数 $p(\Theta)$ 与似然函数 $p(D|\Theta)$ 非共轭分布,导致后验概率密度函数 $p(\Theta|D)$ 无解析解,故采用MCMC的方法对高斯过程回归模型参数的后验分布进行估计,具体可参考文献[16-17]。

2.2 基于贝叶斯理论的最优模型选择

一般而言,随着输入变量(影响因子)的增多,模型的拟合程度会随之增长,然而,这也增加了模型的复杂度,容易导致模型的泛化能力下降。因此,有必要通过某种方法确定出对路基回弹模量 M_R 真正有作用的输入变量(影响因子)。贝叶斯模型选择方法可以自动平衡模型的复杂度与拟合程度,从而确定出对 M_R 预测真正有贡献的输入变量;并基于MCMC样本对高斯过程回归模型的复杂度与拟合程度进行量化表征,定量描述随着模

与似然函数 $p(D|\Theta)$ 对参数估计至关重要。由于本文中的参数具有物理属性,均为非负数,故先验概率密度函数采用half-t分布,具体表示如下:

型输入变量的增多,模型复杂度与拟合程度的变化趋势。在贝叶斯理论中,每个可能的模型均可计算出其对应的概率,而概率的大小反映了模型的可能性大小。概率最大,即可能性最大的模型即为最优模型。以符号 M_j 表示第 j 个高斯过程回归的备选模型,其对应的模型概率 $P_r(M_j|D)$ 可表示为^[18]:

$$P_r(M_j|D) = \frac{p(D|M_j)P_r(M_j)}{\sum_{j=1}^{n_m} p(D|M_j)P_r(M_j)}, j = 1, 2, \dots, n_m \quad (6)$$

式中: n_m 为备选模型的数量; $P_r(M_j)$ 为先验概率,表示在未考虑测量数据时每个高斯过程回归备选模型被选择的概率,通常 $P_r(M_j) = 1/n_m$; $\sum_{j=1}^{n_m} p(D|M_j)P_r(M_j)$ 为常数。

因此,后验概率 $P_r(M_j|D)$ 与似然函数 $p(D|M_j)$ 成正比,后验概率 $P_r(M_j|D)$ 最大的模型可进一步转化为似然函数 $p(D|M_j)$ 最大的模型,这也是 $p(D|M_j)$ 又被称为“模型证据”的原因。根据文献[14],第 j 个高斯过程回归备选模型 M_j 所对应的模型证据 $p(D|M_j)$ 可表示为:

$$p(D|M_j) = \int \frac{p(D|\Theta_j, M_j)p(\Theta_j|M_j)}{q(\Theta_j)} \cdot q(\Theta_j)d\Theta_j \quad (7)$$

式中: Θ_j 表示模型 M_j 中对应的待估计参数; $q(\Theta_j)$ 是助力计算 $p(D|M_j)$ 值的概率分布函数。在下一小节提出了两种概率分布函数来计算 $p(D|M_j)$ 。

2.3 贝叶斯模型证据估计方法

基于式(7)以及条件概率计算公式,可以推导出模型证据 $p(D|M_j)$ 的对数形式为:

$$\ln[p(D|M_j)] = \overbrace{\frac{1}{N_c} \sum_{q=1}^{N_c} \ln[p(D|\Theta_j^{(q)}, M_j)]}^{\text{拟合程度}} - \overbrace{\frac{1}{N_c} \sum_{q=1}^{N_c} \ln \frac{q(\Theta_j^{(q)})}{p(\Theta_j^{(q)}|M_j)}}^{\text{复杂度(奥卡姆因子)}} \quad (8)$$

式中: $\Theta_j^{(q)}$ 为由概率分布 $q(\Theta_j)$ 生成的第 q 组样本; N_c 为样本数量; $p(D|\Theta_j^{(q)}, M_j)$ 、 $q(\Theta_j^{(q)})$ 和 $p(\Theta_j^{(q)}|M_j)$ 分别为模型 M_j 的第 q 组样本 $\Theta_j^{(q)}$ 所对应的似然函数

值、概率密度函数值和先验概率密度函数值。其中, $p(D|\Theta_j^{(q)}, M_j)$ 可以根据式(5)计算, $p(\Theta_j^{(q)}|M_j)$ 可以根据式(4)计算。

然而,由于 $q(\Theta_j)$ 尚待确定, $q(\Theta_j^{(q)})$ 属于未知。对于如何确定概率分布 $q(\Theta_j)$,本节提出两种 $q(\Theta_j)$ 的估计方法。

2.3.1 基于Copula的概率分布 $q(\Theta_j)$

鉴于2.1小节已经获悉 Θ_j 的后验样本,这里不妨将 $q(\Theta_j)$ 定义为 Θ_j 的后验概率分布,即 $q(\Theta_j)=p(\Theta_j|D, M_j)$ 。然而,由于 $p(\Theta_j|D, M_j)$ 无解析解,因此需要用特定的方法将其构建出来,如Copula方法。

Copula的基本思路是将 $p(\Theta_j|D, M_j)$ 分解为每个参数的边缘概率密度函数 $p(\Theta_{j,t}|D, M_j)$ ($t=1, \dots, v_j+3$)的连乘与Copula函数 $C(\Theta_{j,1}, \Theta_{j,2}, \dots, \Theta_{j,v_j+3})$ 的乘积,即:

$$p(\Theta_j|D, M_j) = C(\Theta_{j,1}, \Theta_{j,2}, \dots, \Theta_{j,v_j+3}) \cdot \prod_{t=1}^{v_j+3} p(\Theta_{j,t}|D, M_j) \quad (9)$$

$$q(\Theta_j) = \frac{\exp\left(-\frac{1}{2}(\ln(\Theta_j) - \mu_{\ln[MCMC]})^T \Sigma_{\ln[MCMC]}^{-1} (\ln(\Theta_j) - \mu_{\ln[MCMC]})\right)}{\sqrt{(2\pi)^{v_j+3} |\Sigma_{\ln[MCMC]}|} \cdot \prod_{t=1}^{v_j+3} \Theta_{j,t}} \quad (10)$$

式中: $\Theta_{j,t}$ 为 Θ_j 的第 t 个参数^[20]。

确定概率分布 $q(\Theta_j)$ 之后,即可计算样本 $\Theta_j^{(q)}$ 对应的概率密度函数值 $q(\Theta_j^{(q)})$ 。然后,根据式(8)可计算得到模型 M_j 证据的对数值,即 $\ln[p(D|M_j)]$ 。最优路基回弹模量预测模型即为 $\ln[p(D|M_j)]$ 的最大值所对应的高斯过程回归模型。由式(8)可知,模型 M_j 证据的对数值由两部分构成:一为反映路基回弹模量预测模型复杂度的因子 $-\frac{1}{N_c} \sum_{q=1}^{N_c} \ln \frac{q(\Theta_j^{(q)})}{p(\Theta_j^{(q)}|M_j)}$,记作 $\ln O$;二为反映模型拟合程度的因子 $\frac{1}{N_c} \sum_{q=1}^{N_c} \ln [p(D|\Theta_j^{(q)}, M_j)]$,记作 $\ln L$ 。前者随着模型输入变量个数的增多,不断减小;后者随着模型输入变量个数的增多,不断增大。当两者之和达到最大值时即表示模型最优^[21]。相比于既有工作^[13-14],该节拓展了贝叶斯模型证据的估计方法和应用场景,定量描述了随着模型输入变量的增多,其复杂度与拟合程度的变化趋势,以可视化的形式展示了贝叶斯模型选择的核心。

3 算例分析

为了对本文所提出的方法进行进一步验证,

式中: v_j 为模型 M_j 所包含的参数的数量; $\Theta_{j,t}$ 为 Θ_j 的第 t 个参数。边缘概率密度函数 $p(\Theta_{j,t}|D, M_j)$ 可通过分析 $\Theta_{j,t}$ 的MCMC样本得到,Copula函数 $C(\cdot)$ 可通过在0-1空间分析MCMC样本之间的相关性得到,具体过程可参考文献[14, 19]。为了计算简便,这里Copula函数 $C(\cdot)$ 采用了高斯形式。当参数数量较少时,上述方法可以准确地对 $p(\Theta_j|D, M_j)$ 进行拟合;但当模型参数较多时,可能会由于Copula函数 $C(\cdot)$ 形式的选取不当导致拟合结果与实际情况偏差较大,进而影响模型证据的计算。

2.3.2 基于多维对数正态的概率分布 $q(\Theta_j)$

鉴于 Θ_j 的每一个元素均为非负数,这里假定 $q(\Theta_j)$ 为多维对数正态分布,其均值为MCMC样本的对数的均值,记作 $\mu_{\ln[MCMC]}$,协方差为MCMC样本的对数的协方差,记作 $\Sigma_{\ln[MCMC]}$,则 $q(\Theta_j)$ 可表达为:

本节通过工程实例详细阐述本文所提方法的计算步骤、预测性能以及模型选择结果。

3.1 实测数据的统计分析

本文选用的数据集来源于文献[22],共包含75组测量数据,土样为取自美国路易斯安那州的路基土,土体类型为粉质黏土。土体物理力学参数包括路基土所承受的围压(σ_3)、路基土体内部由于外部荷载作用而产生的偏应力(σ_d)、路基土的含水率(w)、干重度(ρ_d)以及路基回弹模量(M_R),其基本统计特征(最大值、最小值、平均值以及标准差)见表1。数据分布直方图及物理量之间的线性相关系数 r 如图1所示。由图1可以看出,回弹模量 M_R 与 $X_1=\sigma_3$ 、 $X_3=w$ 具有较高的线性相关性,而与 X_2 、 X_4 的相关性较小,但并不意味着包含 $X_1=\sigma_3$ 和 $X_3=w$ 的模型为最优模型,这一点将在3.2节进行详细阐述。

表1 土体参数的统计特征

Table 1 The statistical characteristics of soil parameters

项目	回弹模量/ MPa	围压/ kPa	偏应力/ kPa	含水率/ %	干重度/ ($\text{kN} \cdot \text{m}^{-3}$)
最大值	108.42	42.00	70.0	22.00	16.78
最小值	28.74	0.00	14.0	12.40	16.20
平均值	58.41	21.00	42.0	16.99	16.40
标准差	22.19	17.15	19.8	3.51	0.21

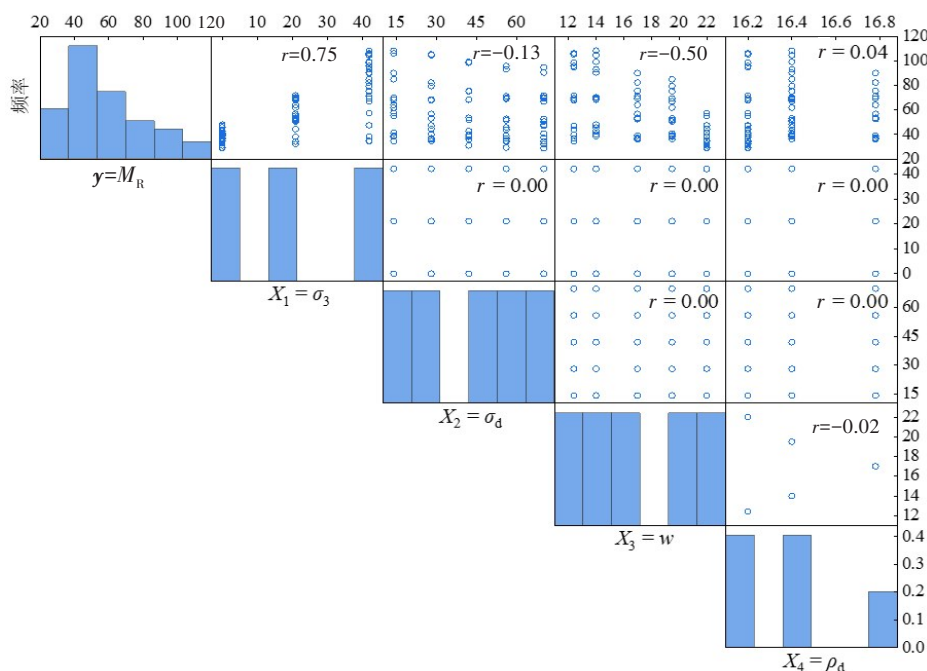


图1 土体参数的分布直方图及相关关系图

Fig. 1 The histograms of the soil parameters and the correlations among them

3.2 基于高斯过程回归的路基回弹模量预测结果分析

本节将基于3.1节中的数据验证本文所提方法的有效性。首先,将数据集随机划分为两部分:训练集(60组)和测试集(15组);其次,基于60组训练数据所对应的3个输入变量($X_1=\sigma_3$, $X_2=\sigma_d$, $X_3=w$)与1个输出变量($y=M_R$)建立高斯过程预测模型,即 M_3 ,见表2。该模型包含6个参数,即 $\Theta=[l_i(t=1,2,3), \sigma_i, \sigma_e, \sigma_n]$,参数的先验分布均为half- t 分布,其中,相关长度 l_i 所对应的half- t 分布的位置参数 μ_i 、尺度参数 σ_i^2 和自由度 ν 分别为0、 10^7 和1;核函数波动程度 σ_i 所对应的 μ_i 、 σ_i^2 和 ν 分别为524.81、262.41和1;预测残差 σ_e 所对应的 μ_i 、 σ_i^2 和 ν 分别为0、5.25和1;目标值的测量误差的标准差 σ_n 所对应的 μ_i 、 σ_i^2 和 ν 分别为5.25、5.25和1。进而可根据式(4)计算得到模型 M_3 的先验分布。基于已有的先验分布与测量数据,采用文献[23]中的切片采样方法即可生成MCMC样本,并得到模型的后验分布。本文中每个参数生成了3 000组MCMC样本,舍弃前500个样本以避免样本不稳定影响预测结果的准确性。

图2给出了模型 M_3 的MCMC样本的先验分布和后验分布。从图2可以看出,后验分布均收敛于某一具体值,说明基于已有的测量数据可对该模型进行有效更新。完成参数估计后,给定新的输

入变量 $\mathbf{x}^*$,即可计算得到路基回弹模量。根据式(8)与已有的MCMC样本,通过Copula方法即可进一步计算得到高斯过程回归模型 M_3 的证据对数 $\ln[p(D|M_j)]=-199.22$ 以及反映模型拟合程度的 $\ln L=-174.71$ 和反映模型复杂度的 $\ln O=-24.51$ 。需要说明的是,通过多维对数正态计算得到的 $\ln[p(D|M_j)]$ 、 $\ln L$ 与 $\ln O$ 的值与通过Copula方法得到的值十分接近。由于篇幅有限,本文仅展示了基于Copula方法的计算结果。根据文献[14],高斯过程回归的相关长度体现了输入变量与目标值之间的相关性大小,因此,将相关长度从小到大排序,结果为 $l_3=0.99$, $l_1=2.87$, $l_2=4.87$, $l_4=2\ 688.58$ 。进而,通过逐个增加输入变量,即可得到所有备选模型,见表2。采用与模型 M_3 类似的方法,即可计算得到其他备选模型的 $\ln[p(D|M_j)]$ 、 $\ln L$ 以及 $\ln O$,如表2和图3所示。从表2和图3可以看出,从模型 M_1 到模型 M_3 ,模型的拟合程度($\ln L$)逐渐增加,由-272.13增加到了-174.71,模型的复杂度逐渐增长,导致 $\ln O$ 逐渐下降^[21],由-11.79下降至-24.51,二者之和即为模型证据的对数 $\ln[p(D|M_j)]$,其由-283.92增长至-199.22。继续增加输入变量,即得到 M_4 ,其 $\ln L$ 波动不大,仅增加了0.24,但 $\ln O$ 减小了4.13,导致 $\ln[p(D|M_j)]$ 降低了3.89。因此,模型 M_3 的证据对数值最大,为-199.22,故模型 M_3 为路基回弹模量的最优高斯过程回归预测模型。

进一步计算得到每个模型被选择的概率,见表2。证明了模型 M_3 为路基回弹模量预测的最优模型。可以看出,模型 M_3 被选择的概率达到了0.98,再次 最优模型 M_3 的预测性能将在3.3节进行详细阐述。

表2 路基回弹模量预测备选模型、高斯过程预测结果及模型选择结果

Table 2 The candidate GPR models for the prediction of the resilient modulus of subgrade soil, the prediction results and the model selection results

模型	输入变量				$\ln L$	$\ln O$	$\ln[p(D M_j)]$	概率	数据集	R^2	$R_{MAPE}/\%$
	σ_3	σ_d	w	ρ_d							
M_1			▲		-272.13	-11.79	-283.92	1.60×10^{-37}	训练集	0.28	29.92
									测试集	0.04	32.21
M_2	▲		▲		-219.81	-17.50	-237.31	2.81×10^{-17}	训练集	0.92	9.58
									测试集	0.96	5.48
M_3	▲	▲	▲		-174.71	-24.51	-199.22	0.98	训练集	0.99	1.34
									测试集	0.99	1.51
M_4	▲	▲	▲	▲	-174.47	-28.64	-203.11	0.02	训练集	0.99	1.35
									测试集	0.99	1.50

注: ▲表示模型包含的输入变量; R^2 为决定系数; R_{MAPE} 为平均绝对百分误差。

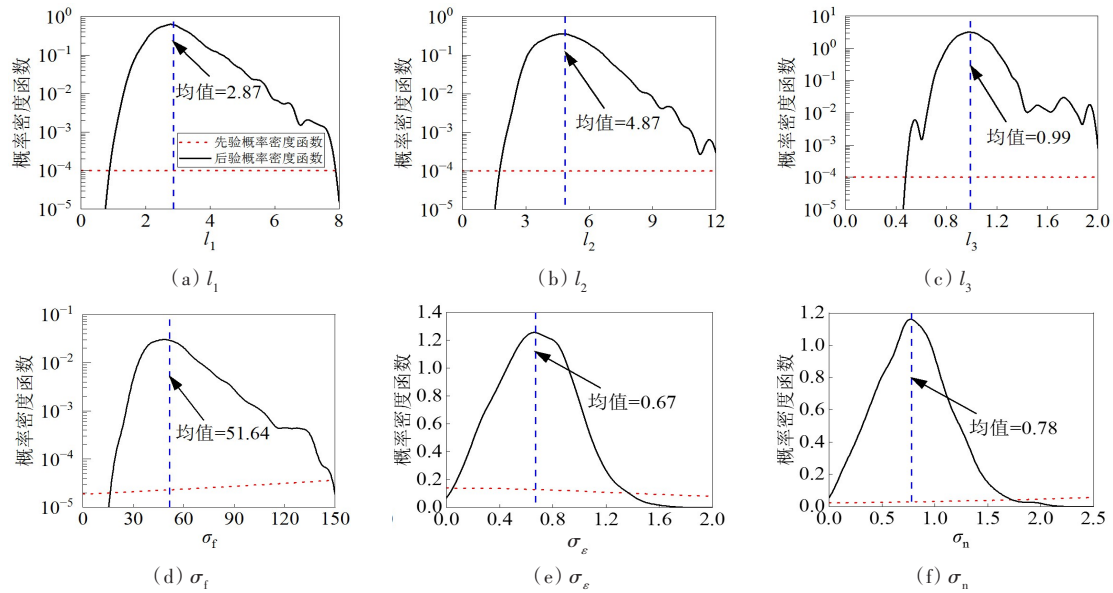


图2 高斯过程回归模型参数的先验和后验概率密度函数

Fig. 2 The prior and posterior probability density functions of GPR parameters

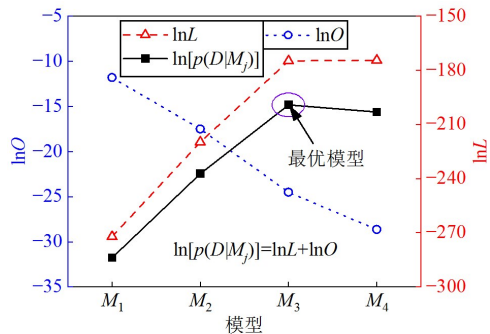


图3 不同GPR模型对应的证据对数 $\ln[p(D|M_j)]$,反映模型拟合程度的 $\ln L$ 以及反映复杂度的 $\ln O$

Fig. 3 The $\ln[p(D|M_j)]$, $\ln L$ and $\ln O$ corresponding to all candidate GPR models for the prediction of the resilient modulus of subgrade soil

3.3 基于贝叶斯理论的模型选择结果分析

本节将详细阐述最优模型 M_3 的预测性能。首先,为对最优模型 M_3 的预测性能进行评估,本文选取了2个评价指标,分别为 R^2 和 R_{MAPE} 。 R^2 是衡量模型预测值与测量值拟合程度的指标, R^2 的范围为0到1, R^2 的值越大,说明模型的预测性能越好; R_{MAPE} 是衡量模型预测值与测量值之间的偏差的指标, R_{MAPE} 的值越小,说明预测值与测量值之间的偏差越小,模型预测性能越好^[14]。 R_{MAPE} 和 R^2 的计算公式如下:

$$R_{MAPE} = \frac{1}{n_m} \sum_{i=1}^{n_m} \left| \frac{\mu_{M_n,i} - y_i}{y_i} \right| \times 100\% \quad (11)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n_m} (y_i - \mu_{M_{n,i}})^2}{\sum_{i=1}^{n_m} (y_i - \mu_y)^2} \quad (12)$$

式中: y_i 和 $\mu_{M_{n,i}}$ 分别为路基回弹模量的测量值和预测值; μ_y 为 n_m 组路基回弹模量测量值的均值。

在本文中,路基回弹模量最优预测模型 M_3 所对应的训练集的 R^2 和 R_{MAPE} 分别为 0.99 和 1.34%, 测试集的 R^2 和 R_{MAPE} 分别为 0.99 和 1.51%, R^2 较大, R_{MAPE} 较小,说明用本文所提出的方法选出的最优模型可以准确预测路基回弹模量。

图 4(a) 给出了最优模型 M_3 的预测值和测量值分布的散点图,同时给出了预测结果的不确定性。从图 4(a) 可以看出,预测值与测量值的散点均分布在 1:1 线附近,且 1:1 线穿过散点 95% 置信区间的比达到了 100%,再次说明了预测结果的准确性。为进行对比,图 4(b) 给出了包含所有输入变量的模型 M_4 的预测值和测量值分布的散点图。从图 4(b) 可以看出,模型 M_4 的预测结果与 M_3 的基本相同,验证了用本文所提出的方法进行模型选择的合理性。

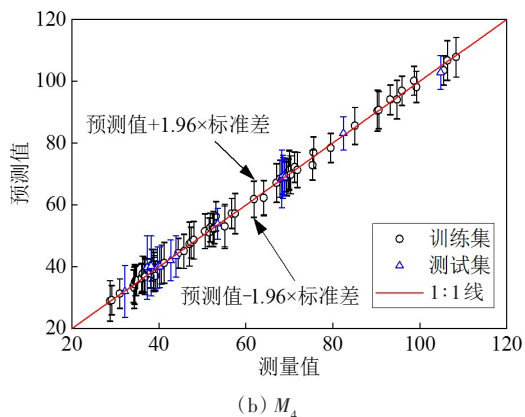
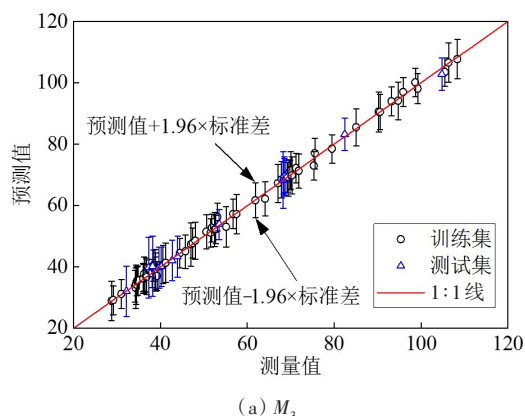
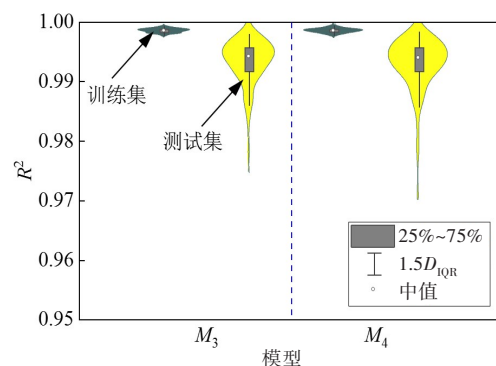


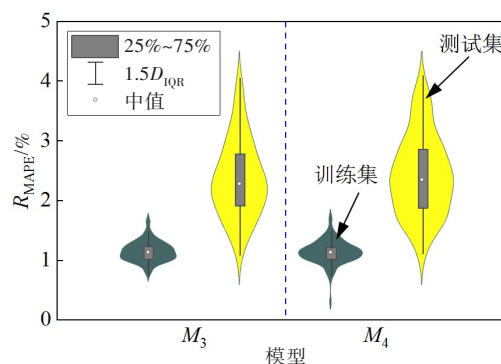
图 4 最优模型 M_3 及全变量模型 M_4 散点图

Fig. 4 The scatter plot of the optimal GPR model M_3 and the full model M_4

以上为一次试验的结果,为对模型的鲁棒性进行验证,将上述过程重复 100 次,得到 100 组对应的 R^2 与 R_{MAPE} ,其中,模型 M_3 被选为最优模型的次数为 88 次。图 5 给出了模型 M_3 所对应的 88 次及模型 M_4 所对应的 100 次预测结果的小提琴图 (图中 D_{IQR} 代表四分位距)。从图 5 可以看出,模型 M_3 的 88 次预测结果的 R^2 的均值在 0.99 以上, R_{MAPE} 的均值小于 3%,且模型 M_4 的预测性能与 M_3 的十分接近,再次说明了 M_3 作为路基回弹模量最优预测模型的合理性,以及用本文所提出的方法预测黏土路基回弹模量结果的准确性与鲁棒性。



(a) R^2



(b) R_{MAPE}

图 5 最优模型 M_3 及全变量模型 M_4 的预测结果统计图

Fig. 5 The violin plot of the prediction performance for the optimal GPR model M_3 and the full model M_4

3.4 不同先验概率下模型选择结果分析

在基于贝叶斯理论的最优模型选择中,先验概率的设定对于贝叶斯模型的结果有一定的影响。上述的先验概率默认所有模型的先验概率相等,这是一种简化假设,计算结果见表 3 中的工况 1。为了验证本文所提方法在不同先验概率下模型选择结果的正确性,进一步对工况 2 和工况 3 的先验概率提出假设。所有工况所用数据完全相同。由表 3 可以看出,工况 2 中模型 M_3 被选择的概率达到了 0.99,工况 3 中模型 M_3 被选择的概率下

降至0.94。换言之,增大 M_3 的先验概率(工况2)确实有利于其被选择出来,然而,这种作用十分有限,相比于工况1(所有模型先验概率均为1/4),其后验概率仅仅增大了0.01;类似地,增大 M_4 的先验概率有效增大了 M_4 被选择的概率,其后验概率由工况1的0.02提高到了工况3的0.06,提高了200%。然而,提高后的后验概率相比于 M_3 而言仍微不足道。由此可知,模型的先验概率对贝叶斯模型选择结果的确有一定的影响,但即使先验概率不准确,基于本文提出的方法依然可以准确地选出最优模型,这进一步证明了本文所提方法的准确性和鲁棒性。

表3 不同先验概率下模型选择结果

Table 3 Model selection results with different prior probabilities

工况	先验概率				模型证据对数值				后验概率			
	M_1	M_2	M_3	M_4	M_1	M_2	M_3	M_4	M_1	M_2	M_3	M_4
工况1	1/4	1/4	1/4	1/4	-283.92	-237.31	-199.22	-203.11	1.60×10^{-37}	2.81×10^{-17}	0.98	0.02
工况2	1/6	1/6	1/2	1/6	-285.71	-239.09	-199.90	-204.90	5.38×10^{-38}	9.50×10^{-18}	0.99	0.01
工况3	1/6	1/6	1/6	1/2	-285.71	-239.09	-201.00	-203.80	1.53×10^{-37}	2.70×10^{-17}	0.94	0.06

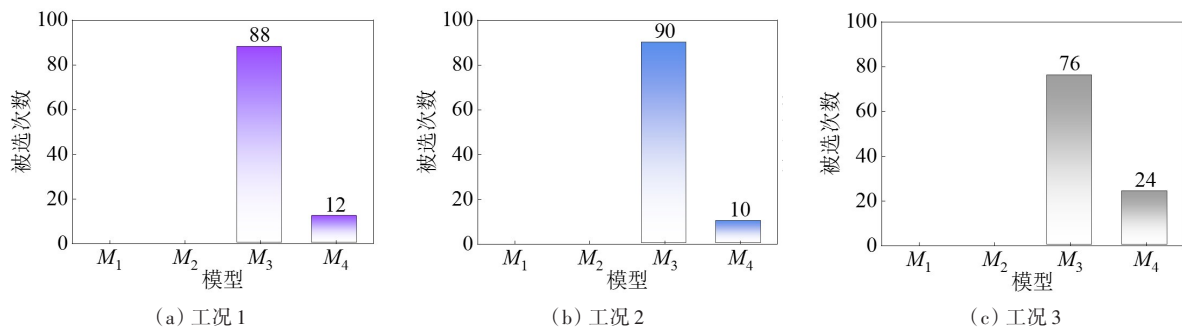


图6 不同先验概率下100次模型选择结果

Fig. 6 Model selection results with different prior probabilities for 100 times

4 讨论

本节将进一步从物理机制的角度对上述模型选择结果进行分析。由表1和图1可以看出,相比于其他输入变量,路基土干重度的变化范围非常小,为16.20~16.78 kN/m³,这意味着本文所用的路基土均来源于同一土层且土质均匀,而本文所提出的贝叶斯高斯过程回归方法为一种数据驱动的算法,较小的波动程度对结果的影响较小,这也导致本文所选出的最优模型未包含干重度。但这并不意味着干重度对黏土路基回弹模量的预测不重要,土体的干重度越大,土体中的孔隙越小,土体颗粒排列越紧密,进而有利于增强土体抵抗竖向

为进一步验证本文方法的鲁棒性与可行性,这里将数据集按照固定比例(8:2)随机划分为训练集与测试集,并进行了100次独立试验。相比于单次的计算结果,基于100次的统计结果更加具有说服力。图6展示了在不同工况下各模型作为最优模型被选择的统计结果。从图6可以看出,当最优模型 M_3 的先验概率较大时(工况2),最优模型被选出来的次数有所提高,达到90次,相比于工况1,仅增加了2次;反之,全变量模型 M_4 先验概率较大时(工况3),最优模型 M_3 被选出来的次数虽有所下降,但相比于 M_4 而言,仍占据主导地位,比例接近0.8。

变形,并使路基回弹模量增大,这将在后续的研究中进行进一步的论证。

另外,尽管本文所提方法可以准确预测黏土路基回弹模量并进行模型选择,但也存在一定的局限性。当输入变量的数量过多、样本量过大时,该方法的计算效率会显著降低,因而需要融入其他方法加以处理,如稀疏高斯过程回归。对于其他类型的土体也可以基于本文的方法进行推广,但需要重新采集数据、训练模型,这也是数据驱动的典型特点^[24-25]。

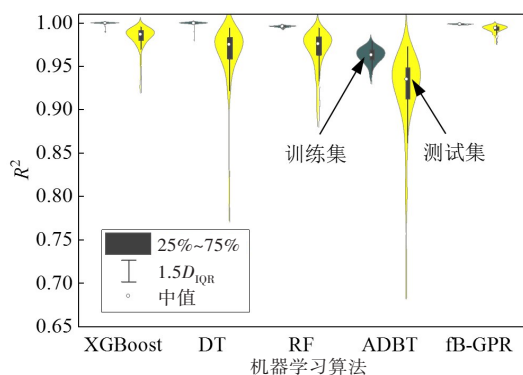
由于本文所提出的方法为数据驱动的方法,故所选出的输入变量的最优组合在采用其他数据驱动的机器学习算法时,应同样具有较好的预测效果。因此,本节讨论了不同机器学习算法对全

变量模型 M_4 和最优模型 M_3 的预测效果,结果见表4。从表4可以看出,极端梯度增强(extreme gradient boosting, XGBoost)、决策树(decision tree, DT)、随机森林(random forest, RF)和自适应增强(adaptive boosting, ADBT)4种机器学习方法对全变量模型 M_4 和最优模型 M_3 均可实现较为准确的路基回弹模量预测,且决定系数 R^2 达到了0.96以上。为对预测结果的可靠性与鲁棒性进行验证,将上述过程重复100次,得到如图7所示的小提琴图。

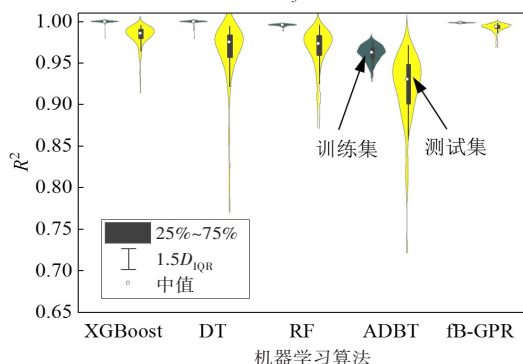
表4 不同机器学习方法路基回弹模量预测结果的 R^2

Table 4 The R^2 for the prediction performance of the resilient modulus of subgrade soil of different machine learning methods

项目	XGBoost	DT	RF	ADBT	fB-GPR
输入变量	X_1, X_2, X_3, X_4				X_1, X_2, X_3
训练集	0.99	1.00	0.99	0.96	0.99
测试集	0.99	0.98	0.99	0.97	0.99
项目	XGBoost	DT	RF	ADBT	fB-GPR
输入变量	X_1, X_2, X_3				
训练集	0.99	1.00	0.99	0.96	0.99
测试集	0.99	0.98	0.99	0.97	0.99



(a) M_3



(b) M_4

图7 不同机器学习方法100次预测结果统计图

Fig. 7 The violin plot for 100 experiments corresponding to different machine learning methods

由图7可以看出,不同机器学习方法100次预测结果的决定系数 R^2 的均值均达到了0.93以上,再次验证了本文所提出的模型选择方法的可靠性。相比于其他机器学习方法,本文提出的贝叶斯高斯过程回归方法不仅可以对路基回弹模量进行较为准确的预测,同时可以对输入变量的最优组合进行客观选择,因而具有重要的现实意义与推广价值。

5 结论

本文提出了贝叶斯高斯过程回归的方法,并将其用于黏土路基回弹模量的预测。所提方法不仅可以通路基土相关的物理力学参数准确预测黏土路基回弹模量,同时实现了输入变量最优组合的客观选择,剔除了冗余输入变量,保证了在路基回弹模量预测模型拟合程度波动不大的情况下,有效降低模型的复杂度,并实现了模型复杂度与拟合程度的定量表征。另外,本文详细阐述了两种贝叶斯模型证据的估计方法,即基于多维对数正态的方法和Copula方法;首次基于马尔科夫链蒙特卡罗(MCMC)样本对高斯过程回归模型的复杂度与拟合程度进行量化表征,定量描述了随着模型输入变量的增多,模型复杂度与拟合程度的变化趋势;探讨了当各个备选模型具有不同的先验概率时,本文所提方法模型选择结果的准确性。主要得出以下结论:

1) 基于本文提出的贝叶斯高斯过程回归方法预测路基回弹模量,最优模型所对应的 R^2 和 R_{MAPE} 分别达到了0.99和1.51%,说明基于本文所提方法选出的最优模型是合理的;在100次试验中,最优模型被选择的比率达到88%,说明本文提出的模型选择方法具有较好的可靠性与鲁棒性。

2) 干重度未包含于最优模型之中,本文分析认为可能是由于本文所采用的黏土数据均来源于同一土层且土质均匀,干重度相比于其他参数波动范围较小,导致其对路基回弹模量的影响较小。

3) 将本文所提方法选出的最优模型应用于其他机器学习方法,均可实现较为准确的路基回弹模量预测,再次验证了本文所提方法的可靠性与有效性,这对机器学习方法在公路工程中的应用与推广具有重要的意义与价值。

[参考文献]

- [1] 张亚娟. 山西省黄土路基回弹模量研究[D]. 西安: 长安大学, 2015.
ZHANG Yajuan. The research of the loess subgrade modulus of resilience in Shanxi Province [D]. Xi'an: Chang'an University, 2015.
- [2] 张锐, 滕煜晟, 刘闯. 路基回弹弯沉控制方法及试验验证[J]. 长沙理工大学学报(自然科学版), 2019, 16(2): 8-14. DOI: 10.3969/j.issn.1672-9331.2019.02.003.
ZHANG Rui, TENG Yusheng, LIU Chuang. Control method and experimental verification of roadbed rebound deflection [J]. Journal of Changsha University of Science & Technology (Natural Science), 2019, 16(2): 8-14. DOI: 10.3969/j.issn.1672-9331.2019.02.003.
- [3] 陈慧东, 曾重驰, 杨博. 路基填料回弹模量预估方法试验研究[J]. 湖南交通科技, 2023, 49(1): 6-12. DOI: 10.3969/j.issn.1008-844X.2023.01.002.
CHEN Huidong, ZENG Zhongchi, YANG Bo. Experimental study on prediction method of resilience modulus of subgrade filler [J]. Hunan Communication Science and Technology, 2023, 49(1): 6-12. DOI: 10.3969/j.issn.1008-844X.2023.01.002.
- [4] 李长生. 重载铁路粉质黏土填料回弹模量及预估模型研究[J]. 路基工程, 2023(2): 1-6. DOI: 10.13379/j.issn.1003-8825.202205008.
LI Changsheng. Research on resilient modulus and prediction model of silty clay filler for heavy-haul railway [J]. Subgrade Engineering, 2023(2): 1-6. DOI: 10.13379/j.issn.1003-8825.202205008.
- [5] KIM D G. Development of a constitutive model for resilient modulus of cohesive soils [M]. Columbus: The Ohio State University, 2004.
- [6] 赵腾远, 宋超, 何欢. 小样本条件下江苏软土路基回弹模量的贝叶斯估计: 基于静力触探数据与高斯过程回归的建模分析[J]. 岩土工程学报, 2021, 43(增刊2): 137-141.
ZHAO Tengyuan, SONG Chao, HE Huan. Bayesian estimation of resilient modulus of Jiangsu soft soils from sparse data: Gaussian process regression and cone penetration test data-based modelling and analysis [J]. Chinese Journal of Geotechnical Engineering, 2021, 43(sup 2): 137-141.
- [7] 谭伟. 基于物性指标的路基细粒土动态回弹模量预估模型研究[J]. 中外公路, 2023, 43(2): 36-42. DOI: 10.14048/j.issn.1671-2579.2023.02.007.
TAN Wei. Study on dynamic resilience modulus prediction model of subgrade fine-grained soil based on physical property parameters [J]. Journal of China & Foreign Highway, 2023, 43(2): 36-42. DOI: 10.14048/j.issn.1671-2579.2023.02.007.
- [8] HEIDARIPANAH A, NAZEMI M, SOLTANI F. Prediction of resilient modulus of lime-treated subgrade soil using different kernels of support vector machine [J]. International Journal of Geomechanics, 2017, 17(2): 06016020. DOI: 10.1061/(asce) gm. 1943-5622.0000723.
- [9] KARDANI N, AMINPOUR M, NOUMAN A R M, et al. Prediction of the resilient modulus of compacted subgrade soils using ensemble machine learning methods [J]. Transportation Geotechnics, 2022, 36: 100827. DOI: 10.1016/j.trgeo.2022.100827.
- [10] NHU V H, HOANG N D, DUONG V B, et al. A hybrid computational intelligence approach for predicting soil shear strength for urban housing construction: a case study at Vinhomes Imperia Project, Hai Phong city (Vietnam) [J]. Engineering with Computers, 2020, 36(2): 603-616. DOI: 10.1007/s00366-019-00718-z.
- [11] AHMED S M S, GUNEYLI H. Robust multi-output machine learning regression for seismic hazard model using peak crust acceleration case study, Turkey, Iraq and Iran [J]. Journal of Earth Science, 2023, 34(5): 1447-1464. DOI: 10.1007/s12583-022-1616-2.
- [12] RASMUSSEN C E, WILLIAMS C K I. Gaussian processes for machine learning [M]. Cambridge: MIT Press, 2006.
- [13] 宋超, 赵腾远, 许领. 基于贝叶斯高斯过程回归与模型选择的岩石单轴抗压强度估计方法[J]. 岩土工程学报, 2023, 45(8): 1664-1673. DOI: 10.11779/CJGE20220734.
SONG Chao, ZHAO Tengyuan, XU Ling. Estimation of uniaxial compressive strength based on fully Bayesian Gaussian process regression and model selection [J]. Chinese Journal of Geotechnical Engineering, 2023, 45(8): 1664-1673. DOI: 10.11779/CJGE20220734.
- [14] ZHAO T Y, SONG C, LU S F, et al. Prediction of uniaxial compressive strength using fully Bayesian Gaussian process regression (fB-GPR) with model class selection [J]. Rock Mechanics and Rock Engineering, 2022, 55(10): 6301-6319. DOI: 10.1007/s00603-022-02964-y.
- [15] DENG Q X, HE J, CAO Z J, et al. Bayesian learning of Gaussian mixture model for calculating debris flow exceedance probability [J]. Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, 2022, 16(1): 154-177. DOI: 10.1080/17499518.2022.2028849.
- [16] 仇文岗, 顾鑫, 刘汉龙, 等. 基于贝叶斯更新的非饱和土坡参数概率反演及变形预测[J]. 岩土力学, 2022, 43(4): 1112-1122. DOI: 10.16285/j. rsm. 2021.1150.

- ZHANG Wengang, GU Xin, LIU Hanlong, et al. Probabilistic back analysis of soil parameters and displacement prediction of unsaturated slopes using Bayesian updating[J]. *Rock and Soil Mechanics*, 2022, 43(4): 1112-1122. DOI: 10.16285/j.rsm.2021.1150.
- [17] GILKS W R, RICHARDSON S, SPIEGELHALTER D. Markov chain Monte Carlo in practice[M]. Boca Raton: CRC press, 1995.
- [18] SONG C, ZHAO T Y, XU L, et al. Probabilistic prediction of uniaxial compressive strength for rocks from sparse data using Bayesian Gaussian process regression with synthetic minority oversampling technique (SMOTE)[J]. *Computers and Geotechnics*, 2024, 165: 105850. DOI: 10.1016/j.compgeo. 2023. 105850.
- [19] JAWORSKI P, DURANTE F, HARDLE W K, et al. Copula theory and its applications [M]. Berlin: Springer, 2010.
- [20] GRONAU Q F, SARAFOGLOU A, MATZKE D, et al. A tutorial on bridge sampling [J]. *Journal of Mathematical Psychology*, 2017, 81: 80-97. DOI: 10.1016/j.jmp.2017.09.005.
- [21] YAN W M, YUEN K V, YOON G L. Bayesian probabilistic approach for the correlations of compression index for marine clays [J]. *Journal of Geotechnical and Geoenvironmental Engineering*, 2009, 135 (12): 1932-1940. DOI: 10.1061/(asce)gt.1943-5606.0000157.
- [22] PUPPALA A J, MOHAMMAD L N, ALLEN A. Engineering behavior of lime-treated Louisiana subgrade soil [J]. *Transportation Research Record: Journal of the Transportation Research Board*, 1996, 1546(1): 24-31. DOI: 10.1177/0361198196154600103.
- [23] VANHATALO J, RIIHIMÄKI J, HARTIKAINEN J, et al. GPstuff: Bayesian modeling with Gaussian processes [J]. *The Journal of Machine Learning Research*, 2013, 14(1): 1175-1179.
- [24] 陈向阳,王岩.建筑废弃物改良红黏土填料的力学特性研究[J].*交通科学与工程*, 2023, 39(4): 47-53. DOI: 10.16544/j.cnki.cn43-1494/u.2021092623081974.
- CHEN Xiangyang, WANG Yan. Investigation on mechanical properties of red clay filler modified by construction waste[J]. *Journal of Transport Science and Engineering*, 2023, 39 (4): 47-53. DOI: 10.16544/j.cnki.cn43-1494/u.2021092623081974.
- [25] 丁继承.含水率和应力比对路基红砂土回弹模量及累积塑性应变的影响研究[J].*公路与汽运*, 2021 (6): 74-79, 87. DOI: 10.3969/j.issn.1671-2668. 2021. 06.018.
- DING Jicheng. Study on the influence of water content and stress ratio on resilience modulus and cumulative plastic strain of subgrade red sand [J]. *Highways & Automotive Applications*, 2021 (6): 74-79, 87. DOI: 10.3969/j.issn.1671-2668.2021.06.018.

(责任编辑:石月珍;校对:刘平;英文编辑:吴高桥)

(上接第29页)

- Nanjing: Southeast University, 2022.
- [25] 黄晓明,郑彬双.沥青路面抗滑性能研究现状与展望[J].*中国公路学报*, 2019, 32(4): 32-49. DOI: 10.19721/j.cnki.1001-7372.2019.04.003.
- HUANG Xiaoming, ZHENG Binshuang. Research status and progress for skid resistance performance of asphalt pavements [J]. *China Journal of Highway and Transport*, 2019, 32 (4): 32-49. DOI: 10.19721/j.cnki.1001-7372.2019.04.003.
- [26] 孙志林,高超,马帅杰.高抗滑性微表处混合料的制备及性能研究[J].*长沙理工大学学报(自然科学版)*, 2023, 20 (6): 80-90. DOI: 10.19951/j.cnki.1672-9331.20220415001.
- SUN Zhilin, GAO Chao, MA Shuaijie. Study on preparation and performance of micro-surfacing mixture with high skid resistance [J]. *Journal of Changsha University of Science & Technology (Natural Science)*, 2023, 20(6): 80-90. DOI: 10.19951/j.cnki.1672-9331. 20220415001.

(责任编辑:刘平;校对:石月珍;英文编辑:李岳林)