文章编号:1674-2974(2024)11-0147-11 DOI:10.16339/j.cnki.hdxbzkb.2024116

沥青路面压实密度多指标智能预测模型

阙云[1†],戴伊[2],薛斌[1],章灿林[1],牟宏霖[3],袁燕[1]
(1. 福州大学 土木工程学院,福建 福州 350108;
2. 西南交通大学 土木工程学院,四川 成都 610031;
3. 南平武沙高速公路有限责任公司,福建 南平 353000)

摘 要:为提高现有沥青路面压实密度预测模型预测能力,以厦门翔安机场高速公路的路面段为试验场地,选取分别代表振压过程中谐波比、能量和力学变化的压实控制值CCV、耗散测量值DMV、振动压实值VCV以及温度作为监测指标,采用孤立森林算法进行监测指标异常点识别,基于最小偏二乘回归建立多指标沥青路面压实密度预测模型.结果表明:孤立森林算法可有效识别高维数据异常点,弥补传统方法只能处理一维数据的不足;温度与其他监测指标以及沥青路面密度存在不同程度正相关关系;基于CCV、DMV、VCV的多元回归模型拟合性能优于一元回归模型性能,论证了多指标评价方法可行性;最小偏二乘回归可改善自变量间共线性对模型权重的影响,解决温度权重与实际物理意义相互颠倒的问题;相比普通多元线性回归,最小偏二乘回归能进一步提高模型拟合能力,模型最终在训练集上的决定系数为0.83,在测试集上的决定系数为0.81,具有良好的沥青路面压实密度预测能力.

关键词:沥青路面;智能压实;智能压实测量值;孤立森林;最小偏二乘回归
中图分类号:U416.2 **文献标志码**:A

Intelligent Compaction Density Prediction Model of Asphalt Pavement Based on Multiple Indicators

QUE Yun[1†],DAI Yi[2],XUE Bin[1],ZHANG Chanlin[1],MU Honglin[3],YUAN Yan[1]
(1. College of Civil Engineering, Fuzhou University, Fuzhou 350108,China;
2. School of Civil Engineering, Southwest Jiaotong University, Chengdu 610031,China;
3. Nanping Wusha Expressway Co., Ltd., Nanping 353000,China)

Abstract: To improve the predictive ability of existing models for predicting the compaction density of asphalt pavement, a test site was set up on the upper layer of the Xiang'an Airport Highway Project in Xiamen. The CCV, DMV and VCV, which represent the change of harmonic ratio, energy, and mechanics in the vibration and compression process, respectively, as well as the temperature, were chosen as indicators. The isolation forest algorithm was used to detect outliers in indicators. The density prediction model was established based on the partial least squares regression. The results show that the isolation forest can effectively recognize outliers of high-dimensional data, covering the shortage that traditional methods can only process one-dimensional data. There are

* **收稿日期**:2023-12-11
基金项目:国家自然科学基金资助项目(52008113), National Natural Science Foundation of China(52008113)
作者简介:阙云(1980—),男,江西黎川人,福州大学教授,博士
† 通信联系人,Email:Queyun_2001@fzu.edu.cn

different degrees of positive correlation between temperature, other indicators, and asphalt pavement density. The multiple regression model based on CCV, DMV, and VCV obtains better fitting ability than the unitary regression methods, proving the feasibility of multiple indicators. The partial least squares regression can restrain the adverse impact caused by the approximate collinearity between independent variables, correct the incorrect weight of temperature, and improve the fitting degree compared with the common multiple linear regression methods. The final determination coefficient of the model on the training set is 0.83, and on the test set is 0.81, indicating good predictive ability for asphalt pavement density.

Key words: asphalt pavement; intelligent compaction measurement values; index of continuous compaction monitoring; isolation forest; partial least squares regression

压实是沥青路面工程最关键步骤之一[1],保证压实质量可防止沥青路面过早出现裂缝、坑槽、车辙等病害[2-4]. 为避免不良路面铲除重铺所造成的资源浪费,引入智能压实(Intelligent Compaction, IC)技术尤为重要.

IC原理为采集压路机钢轮振动响应信号,处理信号后获得表征压实度的智能压实测量值(Intelligent Compaction Measurement Values, ICMVs)[5]. ICMVs有3类:一是表征信号畸变的谐波比指标,以压实计值(CMV)[6]、谐振计值(RMV)[7]和压实控制值(CCV)[8]为代表,广泛用于路基与沥青路面压实度预测;二是力学指标,徐光辉等[9]依据路基压实过程中加速度信号与结构抗力线性关系提出抗力指标振动压实值VCV;三是能量指标,有表征土体一个振动周期吸收能量的Omega[10],压路机能量耗散的MDP[11]、振动压路机单位体积压实功E[12]和耗散测量值DMV[13]. 然而,以往多数研究以单一或同一类ICMVs建立压实度回归模型[14],同时使用3类ICMVs预测的尝试较少. 另外,由于沥青混合料为温敏材料,温度的不同会对ICMVs产生影响. 不少研究已论证了温度影响沥青混合料压实特性[15-17]从而影响ICMVs[18-19]. 因此,考虑温度影响下的沥青混合料密度多指标预测模型将更为合理.

由于压路机与沥青混合料接触状态复杂[20],ICMVs往往波动较大甚至出现异常. 异常点会扭曲回归模型中自变量与因变量间相关性. 因此,异常点处理是十分重要的预处理步骤. 传统方法以拉依达准则[21](3σ准则)为代表,适用于正态分布一维数据,应用场景具有局限性. 孤立森林[22](Isolation Forest, iForest)不依赖数据分布而借助疏离程度识别异常点,具有处理高维数据的能力[23]. 因此,对于多个ICMVs组成的高维数据集异常点识别,iForest是一种可尝试的方式.

传统方法以建立简单线性回归模型来预测压实质量[24-25]. 简单线性回归模型需要自变量间满足相互独立假设,但在沥青混合料压实过程中,由于自变量有相似的变化趋势导致多重共线性现象出现,使得线性回归模型出现预测能力差、显著性检验失去意义和系数正负颠倒而违背物理意义等问题[26].

基于以往缺陷,本文以厦门翔安机场高速公路的路面段为试验场地,在考虑沥青混合料温度影响的基础上,选取谐波畸变、力学与能量变化3类智能压实测量值中具有代表性的CCV、DMV和VCV作为监测指标并建立沥青混合料密度预测模型. 为减少异常点对相关性分析以及模型性能的干扰,引入孤立森林算法识别异常点. 使用最小偏二乘回归模型代替传统线性回归模型来改善自变量间共线性情况,为建立符合物理意义且拟合程度好的智能压实回归模型提供参考.

1 研究方法

1.1 智能压实测量值计算方法

CCV[8]、DMV[13]与VCV[9]计算方法如下:

$$\mathrm{CCV}=100\times\frac{A_{0.5\Omega}+A_{1.5\Omega}+A_{2\Omega}+A_{2.5\Omega}+A_{3\Omega}}{A_{0.5\Omega}+A_{\Omega}}\tag{1}$$

式中:A_{Ω}、$A_{0.5\Omega}$、$A_{1.5\Omega}$、$A_{2\Omega}$、$A_{2.5\Omega}$、$A_{3\Omega}$分别为基波幅值、0.5次谐波幅值、1.5次谐波幅值、2次谐波幅值、2.5次谐波幅值、3次谐波幅值.

$$\mathrm{DMV}=\pi\omega_0^3\int_0^{+\infty}\frac{A_{\varpi}^2}{\varpi^3}\mathrm{d}\varpi\tag{2}$$

式中:ω_0为基频,ϖ为圆频率,A_ϖ为频域中圆频率ϖ对应幅值.

$$\mathrm{VCV}=P\sin(\omega t)+mg-\eta mf(\ddot{u},\omega) \tag{3}$$

式中:P为激振力幅值,ω为偏心块角速度,t为时间,m为压路机钢轮质量,g为重力加速度,η为综合修正函数,f为激振频率,$\ddot{u}$为钢轮加速度.

VCV 的单位为 kN/m,DMV 的单位为 $\mathrm{km^{-1}\cdot s^{-1}}$,CCV 无量纲.

1.2　孤立森林算法

由于高效的策略以及优秀的准确性,iForest[22]已广泛用于工业结构化数据异常点检测[27-28]. iForest 原理为:使用超平面一步步切分样本空间,直至每一子空间仅包含一个样本点x_i为止. 由于异常点通常距离正常点族群较远,因此仅需几次切分就可完全孤立出来.图 1 以二维样本点为例,图中正常点x_i需经历 11 次切分才可被完全孤立出来,而异常点x_0仅需 4 次切分,因此切分次数可衡量数据是聚集的还是稀疏的.

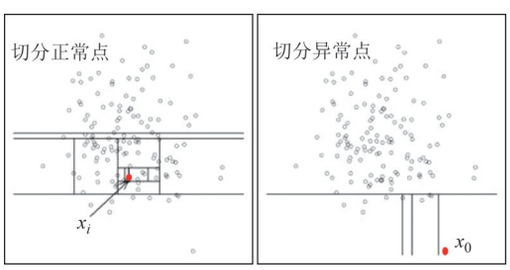


图 1　正常点与异常点的切分[18]

Fig.1　Segmenting normal points and outliers[18]

1.3　最小偏二乘回归

最小偏二乘[29](Partial Least Squares, PLS)回归是集合主成分回归、线性回归和典型相关分析特点的多元回归方法,能一定程度解决自变量间多重共线性问题,使自变量权重尽可能反映其与因变量之间真实的关系.

基本做法为:

先对自变量集$\boldsymbol{X}=(x_1,x_2,\cdots,x_m)^{\mathrm{T}}$提取第一主成分$\boldsymbol{u}_1$,对因变量集$\boldsymbol{Y}=(y_1,y_2,\cdots,y_p)^{\mathrm{T}}$提取第一主成分$\boldsymbol{v}_1$. 如下式所示:

$$\begin{cases}\boldsymbol{u}_1=\boldsymbol{X}\boldsymbol{p}_1\\ \boldsymbol{v}_1=\boldsymbol{Y}\boldsymbol{q}_1\end{cases} \tag{4}$$

式中:$\boldsymbol{p}_1$与$\boldsymbol{q}_1$分别为$\boldsymbol{X}$与$\boldsymbol{Y}$的第一主成分轴向分量.

为使得第一主成分$\boldsymbol{u}_1$与$\boldsymbol{v}_1$相关程度最大的同时携带更多变异信息,有以下优化函数:

$$\max \mathrm{Cov}(\boldsymbol{u}_1,\boldsymbol{v}_1) \tag{5}$$

引入拉格朗日乘子可得:

$$\begin{cases}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{p}_1=\lambda^2\boldsymbol{p}_1\\ \boldsymbol{Y}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{q}_1=\lambda^2\boldsymbol{q}_1\end{cases} \tag{6}$$

因此可将优化函数转化为:

$$\max \lambda \tag{7}$$

所以,第一主成分轴向分量$\boldsymbol{p}_1$与$\boldsymbol{q}_1$实则是$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{X}$和$\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}$最大特征值所对应的单位特征向量. 当得到$\boldsymbol{p}_1$和$\boldsymbol{q}_1$后建立如下回归方程:

$$\begin{cases}\boldsymbol{X}=\boldsymbol{u}_1\boldsymbol{c}_1^{\mathrm{T}}+\boldsymbol{E}\\ \boldsymbol{Y}=\boldsymbol{v}_1\boldsymbol{d}_1^{\mathrm{T}}+\boldsymbol{G}=\boldsymbol{u}_1\boldsymbol{r}_1^{\mathrm{T}}+\boldsymbol{F}\end{cases} \tag{8}$$

式中$\boldsymbol{E}$、$\boldsymbol{G}$分别为回归方程残差矩阵,$\boldsymbol{c}_1$、$\boldsymbol{d}_1$和$\boldsymbol{r}_1$可由下式计算:

$$\begin{cases}\boldsymbol{c}_1=\dfrac{\boldsymbol{X}^{\mathrm{T}}\boldsymbol{u}_1}{\left\|\boldsymbol{u}_1\right\|^2}\\ \boldsymbol{d}_1=\dfrac{\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{v}_1}{\left\|\boldsymbol{v}_1\right\|^2}\\ \boldsymbol{r}_1=\dfrac{\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{u}_1}{\left\|\boldsymbol{u}_1\right\|^2}\end{cases} \tag{9}$$

将残差矩阵$\boldsymbol{E}$视作新的$\boldsymbol{X}$,$\boldsymbol{F}$视作新的$\boldsymbol{Y}$,继续提取第二主成分$\boldsymbol{u}_2$与$\boldsymbol{v}_2$,重复上述步骤直至无残差或达到期望残差为止.

最终可得到$\boldsymbol{X}$、$\boldsymbol{Y}$与各主成分间关系:

$$\begin{cases}\boldsymbol{X}=\boldsymbol{u}_1\boldsymbol{c}_1^{\mathrm{T}}+\boldsymbol{u}_2\boldsymbol{c}_2^{\mathrm{T}}+\cdots+\boldsymbol{u}_r\boldsymbol{c}_r^{\mathrm{T}}+\boldsymbol{E}\\ \boldsymbol{Y}=\boldsymbol{u}_1\boldsymbol{r}_1^{\mathrm{T}}+\boldsymbol{u}_2\boldsymbol{r}_2^{\mathrm{T}}+\cdots+\boldsymbol{u}_r\boldsymbol{r}_r^{\mathrm{T}}+\boldsymbol{F}\end{cases} \tag{10}$$

PLS 回归采用了迭代优化求解方式,通过不断将自变量集和因变量集投影到新的空间建立回归模型,最终避免自变量集共线性对回归模型干扰.

2　现场试验和数据处理

2.1　现场试验

现场试验于福建省厦门市翔安机场高速公路工程的 SMA-13 层进行,路面结构示意图详见图 2,混合料体积特征详见表 1. 试验带长 100 m,宽 2.1 m,划分为 5 个区域后采取不同方案进行振压,如图 3 所示. 区域 1 至区域 5 总振压次数依次为 1 到 5 次,振压前均先静压 1 次. 振压时压路机每完成“单程”和“返程”过程计为 1 次振压. 压路机按照区域 1 到区域 5 的顺序依次完成各区域作业.

试验采用 BW203AD-4 双钢轮振动压路机,其性能参数详见表 2. 压路机与采集装置如图 4 所示,经改造安装,压路机钢轮配备有加速度传感系统及定

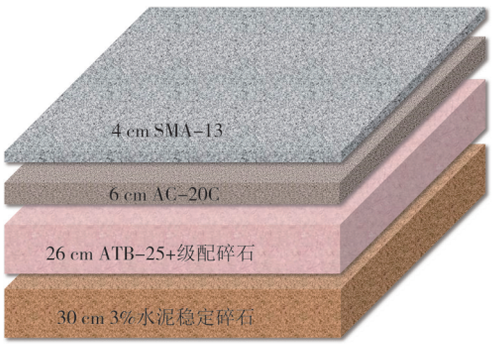


图2 主线主车道、枢纽互通车道路面结构

Fig.2 The pavement structure of principal lines and interconnected lines

表1 SMA-13体积特征参数

Tab.1 The volumetric characteristic indexes of SMA-13

沥青含量/%	实验室标准密度/(g·cm^{-3})	空隙率/%	矿料间隙率/%	沥青饱和度/%	稳定度/kN	流值/mm
5.9	2.576	3.7	17.3	78.4	8.72	30.5

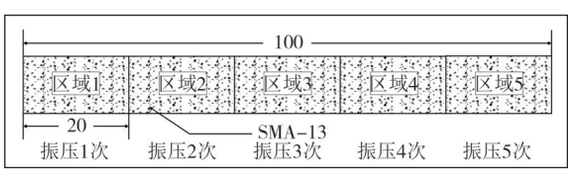


图3 不同振压区域布局(单位:m)

Fig.3 The layout of different areas for vibration compaction test (unit:m)

位系统、底部安装有温度传感器,由驾驶室内的中海达北斗高精度车载定位终端ZDT810接收并处理信号,在压实过程中对沥青混合料表面温度、压路机速度和方位、钢轮加速度响应信号进行实时记录.加速度采集频率为1 000 Hz,定位、速度及温度采集频率为1 Hz.待路面冷却后,使用路面密度仪(Pavement Quality Indicator,PQI)测量密度并采用钻芯取样及表干法计算密度,如图5所示.

表2 BW203AD-4压路机性能参数

Tab.2 The performance parameters of BW203AD-4 roller

轮宽/mm	钢轮载荷/kg	行驶速度/(km·h^{-1})	振动频率/Hz	振动幅度/mm	激振力/kN
2 135	6 600	0~11.0	40/50	0.69/0.29	126/84

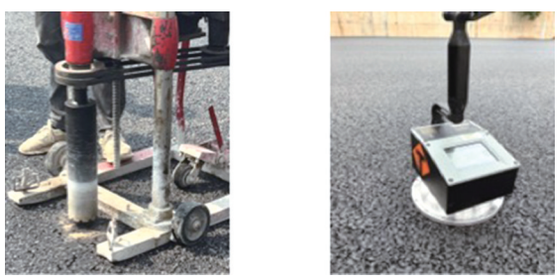
2.2 数据处理

每一区域采集到钢轮加速度信号后,须首先去除压路机加速启动、减速停止及倒车过程中的不稳定部分,保留中间部分用于分析.将处理后的信号切分成包含1 000个采样点的短信号,使用快速傅里叶变换将短信号由时域转换为频域以获得频率幅值信



图4 压路机与采集装置

Fig.4 The roller and sampling devices

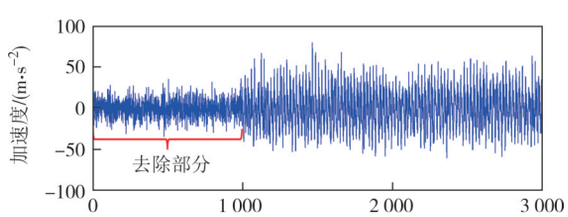


(a)钻芯取样 (b)无核密度仪

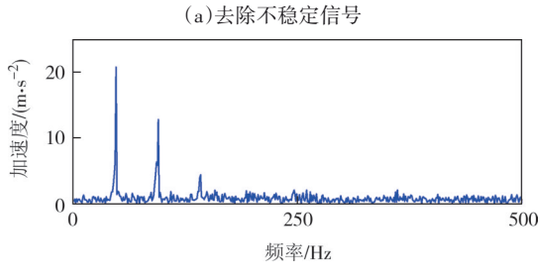
图5 测量路面压实度

Fig.5 Measuring pavement compaction

息,图6所示为加速度信号处理过程.依据式(1)~式(3)分别计算CCV、DMV、VCV.通过定位数据匹配相应温度,最终建立4维数据集.



(a)去除不稳定信号



(b)信号频域图

图6 加速度信号处理过程

Fig.6 The process of acceleration signals

2.3 异常点处理

2.3.1 正态性检验

正态性是检测原始数据缺失点、异常点等预处理手段以及方差分析、回归分析等分析方法的前提.Shapiro-Wilk检验[30]适用于小样本,因此可用于判断试验数据是否符合正态分布.在使用Shapiro-Wilk

检验时,如果p值大于0.05,可认为样本来自的总体分布服从正态分布.

由于不同区域振压时的初始温度不同,为避免温度影响,正态性检验分区进行,结果见表3. 表中各智能压实测量值下角标表示区域编号. 尽管CCV_3、CCV_4以及所有区域DMV检验结果均为非正态分布(即p值小于0.05),但仍可认为试验数据近似服从正态分布,理由见图7,图中y与x含义见式(11)、式(12). 图7以CCV_3和DMV_1的Q-Q图为例,展示了试验数据与理想正态分布之间的偏差. Q-Q图原理为:当数据点基本位于直线附近时,数据服从正态分布. 由于图7中绝大多数点位于直线上,所以可认为数据近似服从正态分布.

$$y_i = \Phi^{-1}\left(\frac{i}{n}\right) \tag{11}$$

式中:Φ为标准正态分布累计概率分布函数,n为样本个数,i为样本升序时样本序号.

$$x = y\sigma + \mu \tag{12}$$

式中:x为分位数,σ为标准差,μ为均值.

表3　Shapiro-Wilk检验

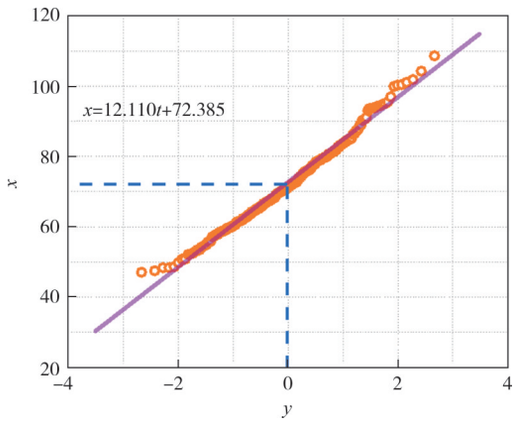
Tab.3　The Shapiro-Wilk test

智能压实测量值	区域	W统计量	p值
DMV_1	区域1	0.963 1	1.069×10^{-10}
DMV_2	区域2	0.975 8	2.043×10^{-4}
DMV_3	区域3	0.972 7	9.788×10^{-8}
DMV_4	区域4	0.975 1	2.277×10^{-12}
DMV_5	区域5	0.966 8	4.978×10^{-14}
CCV_1	区域1	0.996 9	3.527×10^{-1}
CCV_2	区域2	0.997 5	1.187×10^{-1}
CCV_3	区域3	0.985 9	1.138×10^{-2}
CCV_4	区域4	0.992 2	1.397×10^{-2}
CCV_5	区域5	0.998 6	3.732×10^{-1}
VCV_1	区域1	0.986 8	2.966×10^{-1}
VCV_2	区域2	0.991 4	2.324×10^{-1}
VCV_3	区域3	0.987 1	4.690×10^{-1}
VCV_4	区域4	0.974 8	2.891×10^{-1}
VCV_5	区域5	0.987 1	5.978×10^{-2}

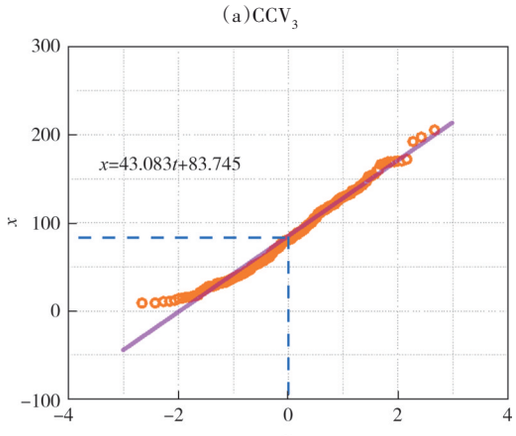
2.3.2　基于iForest的异常点处理

由于3σ准则仅能从一维识别异常点,容易造成数据集中某些样本点从某一维度判定是异常点,而从其他维度来判定又是正常点的矛盾现象出现. iForest可直接从四维样本空间进行异常点识别,综合考虑温度与各ICMVs间相关关系. 对不同区域数据集进行分区检验,iForest算法参数设置见表4.

由于无法直观展示四维空间,使用PCA降维技术将四维样本点投影至三维空间来等效展现,同时在每一特征下标记异常点. 以区域1典型位置数据为例,基于iForest的异常点识别结果见图8、图9.



(a)CCV_3



(b)DMV_1

图7　CCV_3与DMV_1的Q-Q图

Fig.7　The Q-Q plots of CCV_3 and DMV_1

表4　iForest算法参数设置

Tab.4　The parameter setting of the iForest

随机数数量	子树样本数	异常数占比	子数特征数
100	256与样本数间取小值	0.05~0.08	4

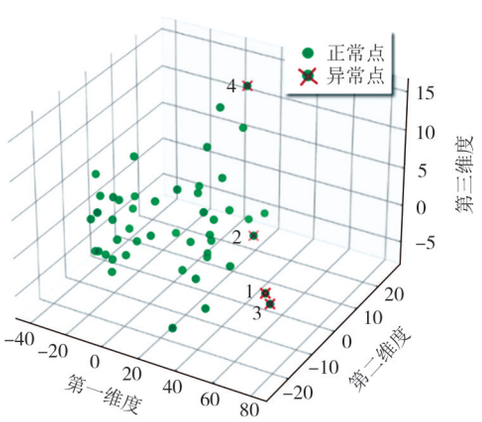
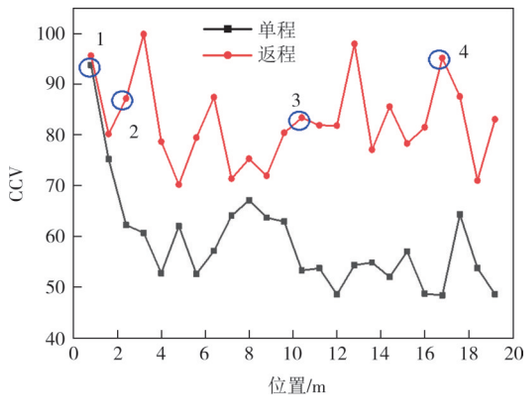
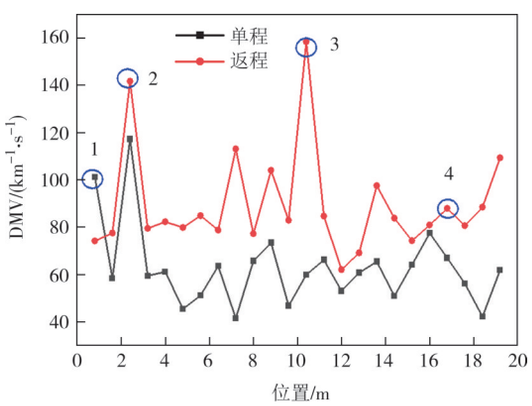


图8　基于iForest的异常点识别

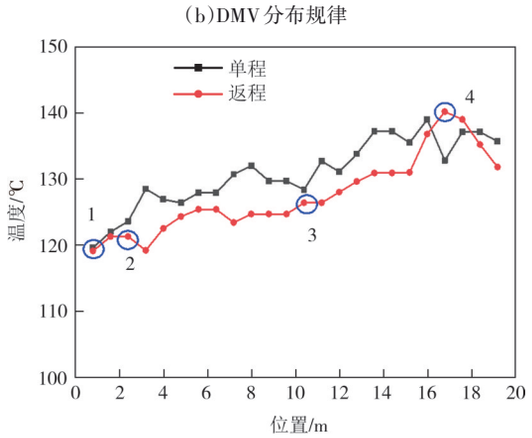
Fig.8　The recognition of outliers based on the iForest



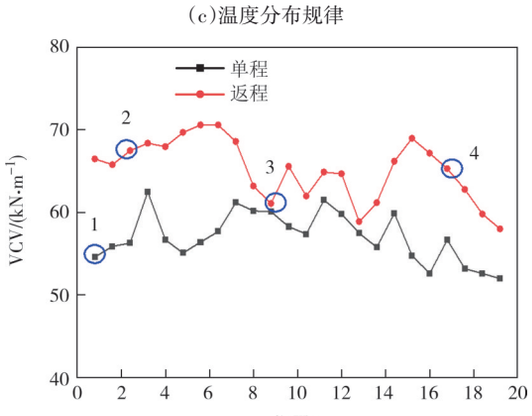
(a)CCV 分布规律



(b)DMV 分布规律



(c)温度分布规律



(d)VCV 分布规律

图 9 每一特征异常点

Fig.9 Outliers for each character

对于区域 1 单程与返程数据,孤立森林算法总共检测出 4 个异常点,其中单程数据有 1 个异常点(编号 1),返程数据有 3 个异常点(编号 2、3、4). 由图 9 可观察到:由于 1 号异常点 CCV 测值以及 2、3 号异常点 DMV 测值均高于总体范围,所以在四维样本空间下能被超平面轻易切分出来. 对于 4 号异常点,其温度测值最高,但 3 个智能压实测量值均没因此高出总体范围而被判定为异常点. 可见,iForest 能基于各特征间存在的正相关关系有效识别异常点.

使用 3σ 准则对区域 1 数据再次检测,每一维度异常点识别结果如表 5 所示. 3σ 准则识别结果为 1 个异常点(对应于图 9 中 3 号异常点),说明在数据波动时该准则对异常点并不敏感. 这是由于当数据标准差 σ 较大时,$[\mu-3\sigma, \mu+3\sigma]$ 也会变大,更多波动较大的点将纳入正常范围. 同时该准则不能考虑多个维度间相关关系,无法识别出 4 号异常点这类多维关系违背整体趋势的异常点.

表 5 基于 3σ 准则的异常点识别

Tab.5 The recognition of outliers based on the 3σ criterion

维度	正常范围	异常点
温度	[112.23, 146.30]	无
CCV	[5.75, 95.61]	无
DMV	[5.52, 146.97]	158.45
VCV	[45.40, 76.86]	无

由图 9 分布规律可知:CCV 与 DMV 在同一振压下产生较大波动,这是由于其在计算公式上考虑了更为细节以及更高频次谐波贡献,不可避免受到高频噪声影响. 此外,由于温度传感器无法安装在压路机的正中间,压路机单程与返程时所测区域有一定偏差,出现 16~18 m 间返程温度高于单程温度的现象. 然而由以上因素引起的数据异常并非由故障或错误引起,不可贸然对数据进行修正. 鉴于异常点占总体数据比例较小,采用直接剔除的方式处理异常点. 此做法将不会对数据原有分布规律造成影响,在保证数据“特色”的同时有效去除异常点.

3 结果分析

3.1 相关性分析

3.1.1 PQI 测量密度与芯样密度相关性

在区域 1~4 各取 4 个芯样后采用表干法测量芯样密度并与相应位置的 PQI 测量密度对比,如图 10 所示. PQI 测量密度与芯样密度之间的决定系数 R^2

为0.86,表明PQI测量密度可以近似代替芯样密度.使用PQI测量密度代替芯样密度的方式可避免试验过程中大量钻芯取样造成路面结构破坏.

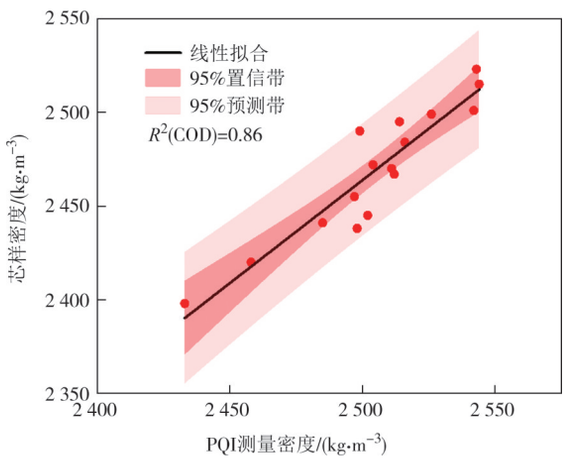


图 10　PQI测量密度与芯样密度相关性

Fig.10　The relationship between the density measured by PQI and core density

3.1.2　不同振压次数对沥青混合料密度影响

图11展示了区域1~5沥青混合料测量密度变化趋势以及每一区域初次振压时平均表面温度.图中密度规定值为沥青混合料试验室标准密度的96%[31],即2 473.0 kg/m^3.当密度超过规定值时表明测点处沥青路面压实度达标.

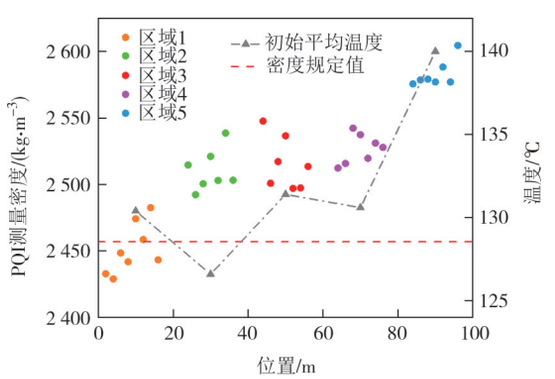


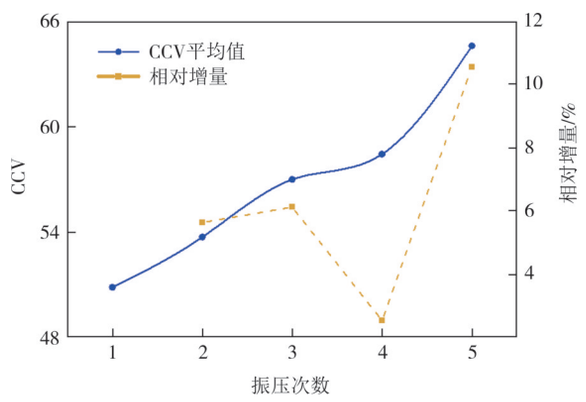
图 11　不同区域密度与振压时初始温度

Fig.11　The initial temperature and density of different areas

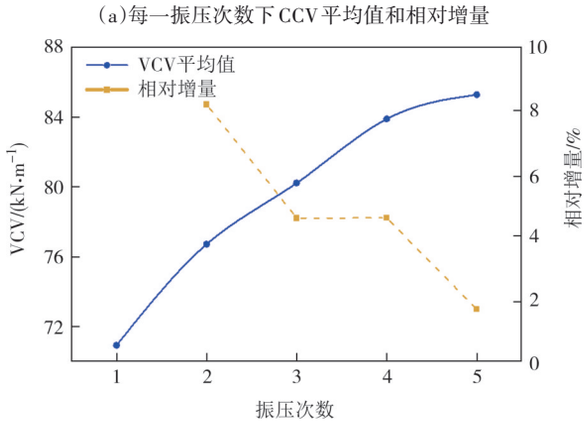
沥青混合料密度随振压次数增加大致呈现增长趋势.区域1与2之间密度增量最大,这是由于初次振压时混合料最为松散,因此能在激振力作用下快速减小空隙并增大密度.区域3与4之间密度增量较小,表明在3次振压下混合料接近密实.为配合现场非试验段施工,一些区域未在混合料摊铺后即刻碾压,因此不同区域初始平均温度存在较大差异.由于区域5初始平均温度明显高于其余区域,且在5次振压下较4次振压时密度有较大增量,因此能一定程度说明高温使得沥青混合料更容易压实.

3.1.3　不同振压次数对ICMVs影响

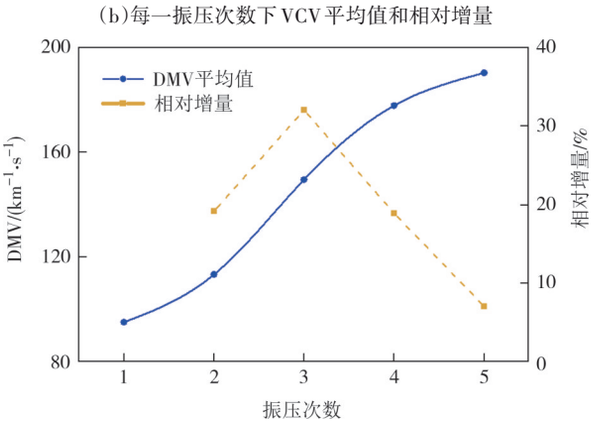
将区域1至5所有ICMVs返程数据按照振压次数重新分组并取平均值,图12为不同振压次数下ICMVs均值以及较前一次均值的相对增量.三者平均值随振压次数增加均呈现增长趋势,能一定程度反映沥青混合料密度变化.相比之下DMV在前4次振压时平均值相对增量最大,这是由于DMV实质上表征畸变信号累积的能量变化,因此比代表畸变信



(a)每一振压次数下CCV平均值和相对增量



(b)每一振压次数下VCV平均值和相对增量



(c)每一振压次数下DMV平均值和相对增量

图 12　不同振压次数下智能压实测量值均值和相对增量

Fig.12　The mean values and relative increment of ICMVs under different vibration passes

号相对变化的CCV以及加速度变化的VCV具有更大的增长范围.

3.1.4 温度对ICMVs影响

将100 m试验带每一区域初次振压时温度与ICMVs以热力图形式展现.由图13可知,DMV与温度之间没有明确的相关性,仅在温差较大的区域1与区域5之间观察到DMV整体增长.而CCV、VCV与温度之间存在较为明显的正相关关系,说明CCV和VCV一定程度上受到温度的影响.理论上,层厚、层间整体刚度、压路机振动频率和温度等因素都能影响ICMVs的测量结果[32],因此在无法控制上述因素相同的情况下,ICMVs与温度的相关性不佳属于正常现象.

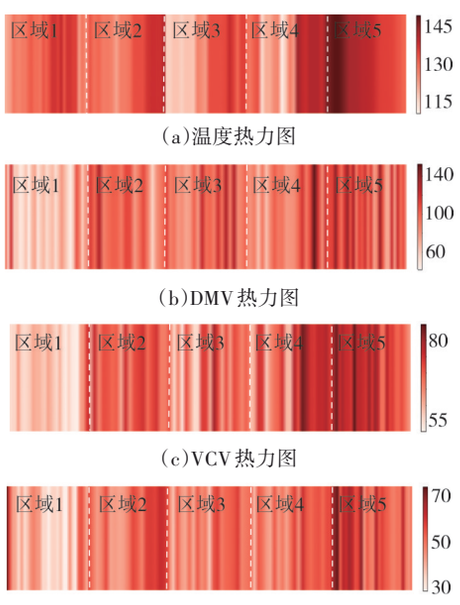


图13 温度与智能压实测量值热力图

Fig.13 Heat maps of temperature and ICMVs

3.2 普通多元线性回归模型

由于VCV、DMV和CCV有着不同的量纲和量级,在建立线性回归模型时,大的量级会使权重偏小,反之亦然.标准化处理可消除量纲和量级对权重的影响,便于权重直接反映自变量与因变量之间的相关关系.Z-score标准化表达式为:

$$x' = \frac{x - \mu}{\sigma} \tag{13}$$

将36处PQI测量密度及其对应的CCV、DMV、VCV和温度数据做标准化处理,以处理后数据组合为数据集,得到的相关系数热力图如图14所示.相关系数在0.3~0.7之间表示两者为中度线性相关,在0.7~1.0之间为强线性相关;相关系数为正数表示正相关,为负数表示负相关.由结果可知,温度与密度、ICMVs之间为中度正相关关系,其中温度与DMV关系最弱,这与图13观察到的现象基本一致.同时,CCV、VCV与密度为强相关,DMV与密度为中度相关,表明三者均具有表征混合料密度的能力.此外,观察CCV、VCV与DMV之间的相关系数可知三者之间存在一定程度的线性相关关系,即共线性关系.这是由于在沥青混合料压实过程中指标均呈现正增长趋势,使得共线性问题出现.

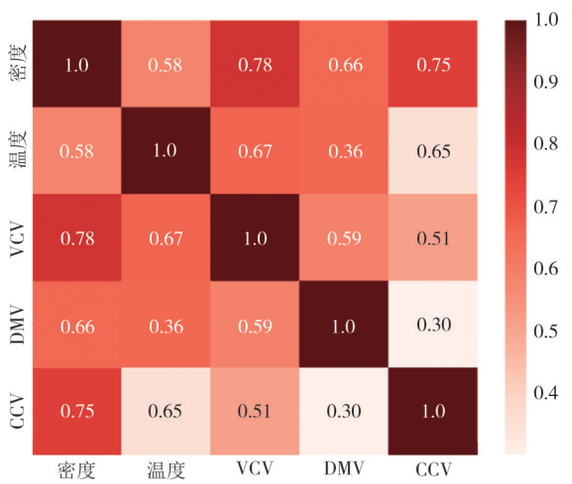


图14 相关系数热力图

Fig.14 The heat map of correlation coefficient

利用数据集建立一元以及多元线性回归模型,结果见表6,表中D为PQI测量密度,T为沥青混合料表面温度.由于自变量组合方式较多,表中仅给出部分结果.

由表6可知,相较于CCV和DMV,VCV与沥青混合料密度相关性更高.同时,观察模型7的决定系数

表6 回归分析表

Tab.6 Regression analysis list

模型	自变量	回归模型	R^2
模型1	CCV	D=0.75CCV	0.56
模型2	VCV	D=0.78VCV	0.61
模型3	DMV	D=0.66DMV	0.44
模型4	CCV、温度	D=0.62CCV+0.36T	0.68
模型5	VCV、温度	D=0.71VCV+0.11T	0.62
模型6	DMV、温度	D=0.49DMV+0.27T	0.48
模型7	CCV、VCV、DMV	D=0.35CCV+0.43VCV+0.23DMV	0.80
模型8	CCV、VCV、DMV、温度	D=0.39CCV+0.57VCV+0.28DMV−0.16T	0.82

R^2可知基于多指标的密度预测模型拟合性能优于基于单一指标的模型. 以各ICMVs建立的一元回归模型(模型1、2、3)在加入温度后(模型4、5、6)决定系数均有提高,说明温度的加入可有效提高模型预测性能,同时,模型7与模型8决定系数之差也证明了此观点. 从拟合优劣的角度来看,由于模型8决定系数(R^2=0.82)最高,因此对沥青混合料密度拟合精度最高. 但与其他模型不同的是,模型8中温度权重为负号,说明温度对预测结果贡献为负,此现象与相关性热力图以及以往研究所得结论不符. 多重共线性是造成温度权重与实际经验相反的原因. 因此,普通线性回归模型并不能在保证精度的同时给出与一般理论相符的预测解释.

3.3　最小偏二乘回归模型

最小偏二乘回归模型中主成分提取个数r通过以下方式确定:以最小化回归模型误差为目标,从提取1个主成分开始依次增加提取主成分个数直至模型误差达到最小值. 本文以均方误差(Mean Square Error,MSE)为优化目标进行提取成分个数选择,均方误差表达式见式(14)所示. 同时,从数据集中随机抽取29组数据(约80%)作为训练集用于模型训练,以余下7组数据作为测试集用于模型性能评估.

$$E_{\mathrm{MSE}}=\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2 \tag{14}$$

式中:y_i为样本真实值,$\hat{y}_i$为样本估计值.

从设定提取主成分为1开始,基于PLS的沥青混合料密度预测模型均方误差和温度权重变化趋势见图15. 随着提取主成分个数增加,均方误差从0.22减少至0.16,且在提取2至4个主成分时保持不变,直至在提取5个主成分时激增到6.39. 温度权重在提取1至2个主成分时为正数,在提取3至4个主成分时为负数,在提取5个主成分时无限接近0. 综上,选择最佳主成分提取数2,在保证温度权重为正数的同时均方误差最小.

最终,在提取2个主成分条件下PLS回归模型在不同数据集下的分析结果见表7.

由表7可知,由PLS所建立的多元线性回归方程对训练集数据具有较好的拟合效果(R^2=0.83),模型在测试集上表现稍差一些(R^2=0.81),但拟合结果仍然较好. 由于模型在训练集与测试集上回归方程显著性检验p值均小于0.05,说明模型的自变量对因变量有显著的共同影响. 同时,模型在训练集和测试集上的均方误差均较小,表明模型具有较高的预测精度. 综合以上结果,PLS回归在解决自变量之间存在共线性问题的同时,具有较好的拟合性能且对沥青混合料密度预测精度较高.

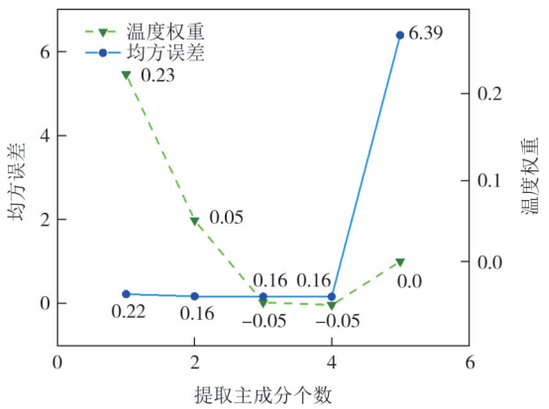


图15　均方误差和温度权重随提取主成分个数变化过程

Fig.15　The processes of mean square error and the weight of temperature with the number of principal components increasing

表7　不同数据集下的PLS回归分析

Tab.7　PLS regression analysis with different data sets

数据集	表达式	R^2	检验p值	均方误差
训练集	D=0.51CCV+0.34VCV+0.24DMV+0.05T	0.83	5.29×10^{-9}	0.16
测试集	D=0.51CCV+0.34VCV+0.24DMV+0.05T	0.81	6.38×10^{-10}	0.18

4　结 论

1) 与传统一维异常点识别方法相比,孤立森林算法可以处理高维数据,因此能在考虑DMV、VCV、CCV以及温度间关系的基础上对异常点进行识别,有效鉴别数据集中的异常点. 由于DMV和CCV考虑了高次谐波贡献,易受到高频噪声干扰使得测值波动大从而被孤立森林判定为异常现象.

2) DMV、VCV和CCV都具有表征沥青混合料密度的能力,其中VCV和CCV与密度相关性最强,DMV最弱. 基于三指标的三元回归模型拟合性能优于单一指标所建立的一元回归模型,表明基于多指标的沥青混合料评估方法可行.

3) 在建立基于简单多元线性回归的压实度预测模型时,自变量间存在共线性问题,导致温度的权重变为负号. 以最小偏二乘回归代替简单多元线性回归后可修正温度权重,提高模型的拟合能力,模型

在训练集与测试集上的决定系数分别为0.83和0.81,具有较高的预测精度.

下一步研究可深入探讨加速信号去噪手段,减小DMV等考虑高次谐波的智能压实测量值受高频噪声干扰程度,使智能压实测量值与其他参数相关关系更为明确.为实现智能压实技术提供更多参考.

参考文献

[1] WANG S Q, SUI X, LENG Z, et al. Asphalt pavement density measurement using non-destructive testing methods: current practices, challenges, and future vision [J]. Construction and Building Materials, 2022, 344: 128154.

[2] KASSEM E, SCULLION T, MASAD E, et al. Comprehensive evaluation of compaction of asphalt pavements and a practical approach for density predictions [J]. Transportation Research Record: Journal of the Transportation Research Board, 2012, 2268(1): 98-107.

[3] MOHAMMAD L N, ELSEIFI M, COOPER S B III, et al. Levels of variability in volumetric and mechanical properties of asphalt mixtures [J]. Journal of Materials in Civil Engineering, 2013, 25(10): 1424-1431.

[4] TRAN N, TURNER P, SHAMBLEY J. Enhanced compaction to improve durability and extend pavement service life: a literature review [R]. Alabama: National Center for Asphalt Technology at Auburn University, 2016: 1-25.

[5] 徐光辉,SHANG G. 智能压实测量值的发展方向[J]. 筑路机械与施工机械化,2018,35(4):19-24.
XU G H, CHANG G. Intelligent compaction measurement values: a road map [J]. Road Machinery & Construction Mechanization, 2018, 35(4): 19-24. (in Chinese)

[6] SANDSTRÖM Â J, PETTERSSON C B Intelligent systems for QA/QC in soil compaction [C]//In Proc., 83rd Annual Transportation Research Board Meeting. Washington D C: Transportation Research Board, 2004, 317(1): 11-14.

[7] BRANDL H, ADAM D. Sophisticated continuous compaction control of soils and granular materials [C]//Hamburg: International Conference on Soil Mechanics and Foundation Engineering. 1997.

[8] SCHEROCMAN J A, RAKOWSKI S, UCHIYAMA K. Intelligent compaction, does it exist? [C]//Ontario: Canadian Technical Asphalt Association. 2007.

[9] 徐光辉,高辉,王哲人. 路基压实质量连续动态监控技术[J]. 中国公路学报,2007,20(3):17-22.
XU G H, GAO H, WANG Z R. Continuous dynamic monitor technology on subgrade compaction quality [J]. China Journal of Highway and Transport, 2007, 20(3): 17-22. (in Chinese)

[10] MOONEY M A, RINEHART R V. Field monitoring of roller vibration during compaction of subgrade soil [J]. Journal of Geotechnical and Geoenvironmental Engineering, 2007, 133(3): 257-265.

[11] WHITE D J, THOMPSON M J. Relationships between *in situ* and roller-integrated compaction measurements for granular soils [J]. Journal of Geotechnical and Geoenvironmental Engineering, 2008, 134(12): 1763-1770.

[12] 刘东海, 李子龙, 王爱国. 基于碾压机做功的堆石坝压实质量实时监测与快速评估[J]. 水利学报, 2014, 45(10): 1223-1230.
LIU D H, LI Z L, WANG A G. Roller working-based real-time monitoring and rapid assessment of rock-fill dam compaction quality [J]. Journal of Hydraulic Engineering, 2014, 45(10): 1223-1230. (in Chinese)

[13] 吴龙梁. 基于能量耗散的路基连续压实控制技术研究[D]. 北京:中国铁道科学研究院,2020.
WU L L. Research on continuous compaction control technology of roadbed based on energy dissipation [D]. Beijing: China Academy of Railway Sciences, 2020. (in Chinese)

[14] LIU D H, LIN M, LI S. Real-time quality monitoring and control of highway compaction [J]. Automation in Construction, 2016, 62: 114-123.

[15] 张争奇,边秀奇,杜群乐,等. 沥青混合料压实特性影响因素研究[J]. 武汉理工大学学报, 2012, 34(6): 36-41.
ZHANG Z Q, BIAN X Q, DU Q L, et al. Study on factors effecting on compaction property of asphalt mixture [J]. Journal of Wuhan University of Technology, 2012, 34(6): 36-41. (in Chinese)

[16] CHEN W, WEI K S, WEI J C, et al. Research on the road performance of asphalt mixtures based on infrared thermography [J]. Materials, 2022, 15(12): 4309.

[17] YU S, SHEN S H. Compaction prediction for asphalt mixtures using wireless sensor and machine learning algorithms [J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(1): 778-786.

[18] POLACZYK P, HU W, GONG H R, et al. Improving asphalt pavement intelligent compaction based on differentiated compaction curves [J]. Construction and Building Materials, 2021, 301: 124125.

[19] HU W, JIA X Y, ZHU X Y, et al. Investigating key factors of intelligent compaction for asphalt paving: a comparative case study [J]. Construction and Building Materials, 2019, 229: 116876.

[20] 王永. 铁路路基智能压实控制技术研究[D]. 北京:中国铁道科学研究院,2018.
WANG Y. Study on intelligent compaction control technology of railway subgrade [D]. Beijing: China Academy of Railway Sciences, 2018. (in Chinese)

[21] 魏治文,程琳,来记桃,等. 几种异常值判别准则在安全监测数据处理中的应用[J]. 大坝与安全,2009(1):67-69.
WEI Z W, CHENG L, LAI J T, et al. Application of some statistical criteria in safety monitoring data processing [J]. Dam &

Safety, 2009(1): 67-69.(in Chinese)

[22] LIU F T, TING K M, ZHOU Z H. Isolation forest[C]//2008 Eighth IEEE International Conference on Data Mining. Pisa: IEEE, 2008: 413-422.

[23] 杨晓晖,张圣昌. 基于多粒度级联孤立森林算法的异常检测模型[J]. 通信学报, 2019, 40(8): 133-142.
YANG X H, ZHANG S C. Anomaly detection model based on multi-grained cascade isolation forest algorithm[J]. Journal on Communications, 2019, 40(8): 133-142.(in Chinese)

[24] LI J G, LI P L, SU J F, et al. Coarse aggregate movements during compaction and their relation with the densification properties of asphalt mixture[J]. International Journal of Pavement Engineering, 2021, 22(8): 1052-1063.

[25] CHEN B, YU X, DONG F Q, et al. Compaction quality evaluation of asphalt pavement based on intelligent compaction technology[J]. Journal of Construction Engineering and Management, 2021, 147(9): 04021099.

[26] 马雄威. 线性回归方程中多重共线性诊断方法及其实证分析[J]. 华中农业大学学报(社会科学版), 2008(2): 78-81.
MA X W. Diagnosis and empirical analysis on multicollinearity in linear regression model[J]. Journal of Huazhong Agricultural University (Social Science Edition), 2008(2): 78-81.(in Chinese)

[27] 王月玥,孙琦,钟厚岳. 基于孤立森林算法的轨道交通实时客流告警阈值设定方法研究[J]. 都市快轨交通, 2023, 36(3): 71-76.
WANG Y Y, SUN Q, ZHONG H Y. Setting method of rail transit real-time passenger flow alarm threshold based on isolation forest algorithm[J]. Urban Rapid Rail Transit, 2023, 36(3): 71-76.(in Chinese)

[28] 黄福兴, 周广山, 丁宏, 等. 基于孤立森林算法的电能量异常数据检测[J]. 华东师范大学学报(自然科学版), 2019(5): 123-132.
HUANG F X, ZHOU G S, DING H, et al. Electric energy abnormal data detection based on Isolation Forests[J]. Journal of East China Normal University (Natural Science), 2019(5): 123-132.(in Chinese)

[29] PAGÈS J, TENENHAUS M. Multiple factor analysis combined with PLS path modelling. application to the analysis of relationships between physicochemical variables, sensory profiles and hedonic judgements[J]. Chemometrics and Intelligent Laboratory Systems, 2001, 58(2): 261-273.

[30] SHAPIRO S S, WILK M B. An analysis of variance test for normality (complete samples)[J]. Biometrika, 1965, 52(3/4): 591-611.

[31] 公路工程质量检验评定标准:JTJ 071—1998[S]. 北京:人民交通出版社股份有限公司,2018.
Quality inspection and evaluation standards for highway engineering: JTJ 071—1998[S]. Beijing: China Communications Press Co., Ltd., 2018.(in Chinese)

[32] HU W, JIA X Y, ZHU X Y, et al. Investigating key factors of intelligent compaction for asphalt paving: a comparative case study[J]. Construction and Building Materials, 2019, 229: 116876.