

基于随机森林算法的隧(巷)道光面爆破参数选择模型研究^①

侯胜心¹, 马晶晶¹, 唐虹靖¹, 李应方¹, 田军¹, 李麟¹, 李启月²

(1.西藏中金新联爆破工程有限公司, 西藏拉萨 851500; 2.中南大学资源与安全工程学院, 湖南长沙 410083)

摘要:为解决隧(巷)道光面爆破中参数选择困难的问题,构建了基于随机森林算法的隧(巷)道光面爆破参数选择模型。模型以围岩质量分级BQ、隧(巷)道断面面积、炮眼直径为输入参数,以最小抵抗线长度、炮眼间距、炮眼深度为输出参数。使用了24个训练样本、7个测试样本进行模型训练及测试,结果表明,炮眼间距、最小抵抗线长度、炮眼深度3个输出参数值与测试样本对应参数值之间的平均相对误差分别为17%、8.8%、22%,可见模型对于最小抵抗线长度的选择效果较好,对炮眼间距、炮眼深度的选择效果欠佳。

关键词: 随机森林算法; 光面爆破; 参数预测; 隧道掘进; 巷道掘进

中图分类号: TD235

文献标识码: A

doi:10.3969/j.issn.0253-6099.2024.03.007

文章编号: 0253-6099(2024)03-0034-04

Parameter Selection Model Based on Random Forest Algorithm for Smooth Blasting in Tunnel (Roadway) Excavation

HOU Shengxin¹, MA Jingjing¹, TANG Hongjing¹, LI Yingfang¹, TIAN Jun¹, LI Lin¹, LI Qiyue²

(1. Tibet Zhongjinxinlian Blasting Engineer Co Ltd, Lhasa 851500, Xizang, China; 2. School of Resources and Safety Engineering, Central South University, Changsha 410083, Hunan, China)

Abstract: In order to solve the problem of parameter selection for smooth blasting in tunnel/roadway excavation, a parameter selection model based on random forest algorithm was constructed. The input parameters in the model include quality classification BQ of surrounding rock, tunnel (roadway) section area and borehole diameter, and the output parameters include minimum burden, borehole spacing and borehole depth. Then, 24 training samples and 7 test samples were used for training and testing of model, and the results show that the relative errors between the output parameters of borehole spacing, minimum burden, borehole depth and the corresponding parameters of test samples are 17%, 8.8% and 22% respectively on average, indicating that the model is good in selecting minimum burden, but less effective in selecting blasthole spacing and blasthole depth.

Key words: random forest algorithm; smooth blasting; prediction of blasting parameters; tunnel excavation; roadway excavation

当前隧(巷)道开挖方式仍然以钻爆法为主,光面爆破技术是钻爆法开挖的关键技术,合理选择光面爆破参数对隧(巷)道掘进工程围岩的稳定及隧(巷)道的支护有重要作用。然而,隧(巷)道爆破相关理论研究尚不成熟,光面爆破参数主要依照以往工程经验来确定,此过程往往受设计人员主观因素影响,导致光面爆破参数选择不合理、爆破效果不理想。

数理统计方法与计算机技术的发展为解决光面爆破参数选择提供了契机,很多学者将机器学习算法应用于掘进爆破参数选择中^[1-4],其中应用较多的是神经网络算法。但神经网络算法需要大量样本才能保证模型选择的可靠性,因此在光面爆破参数优化的应用上受到了一定限制^[5]。随机森林算法是一种灵活、易于使用的学习算法,它对样本量没有过高的要求,不易出

① 收稿日期: 2023-12-07

基金项目: 新疆维吾尔自治区重大科技专项(2018A03003)

作者简介: 侯胜心(1993—),男,黑龙江佳木斯人,助理工程师,主要从事爆破工程、矿山工程施工等工作。E-mail: houshengxin005@163.com

现过拟合且能处理高维度数据,非常适合作为光面爆破参数选择的机器学习算法。本文在借鉴以往经典光面爆破参数选择案例的基础上,使用以 CART 决策树为基学习器的随机森林算法对隧(巷)道光面爆破参数进行选择,建立随机森林光面爆破参数选择模型,并通过搜索算法确定模型中回归树的数量 k ,优化模型参数选择性能,使模型达到更好的爆破参数选择效果。

1 随机森林算法

随机森林算法^[6]是一种集成算法,它以 Bagging 算法的思想为基础拓展而来。Bagging 是集成学习方法里的一种基本算法^[7],它通过对样本集随机抽样,将这些抽样的样本作为子学习集,对每个子学习集进行学习得到多个子学习器,将每个子学习器的预测结果进行综合,得到最终的模型选择结果。

随机森林算法的每个子学习器是一个回归树(用作分类时,又称决策树)。通常采用 CART 算法生成回归树。CART 算法由 Bteimain 等人于 1984 年提出,它是一种特殊的决策树算法,其生成基本原理是:回归树在每个节点划分时,将内部节点分为“是”和“否”两个取值,这样回归树就会被分成二叉树的形式^[8]。假设一个二叉树对应着样本集由输入空间 x 到输出变量 $f(x)$ 的函数,输入空间 x 已经被二叉树分为 n 个部分 R_1, R_2, \dots, R_n ,并且在各 R_n 上均有输出值 C_n ,则回归二叉树模型可以表示为:

$$f(x) = \sum_{n=1}^N C_n I \quad (x \in R_n) \quad (1)$$

式中 I 代表决策树的分支映射。

对于每一个样本输入 x_i ,通过该回归二叉树的划分,都能得到一个输出 $f(x_i)$,通过最小化平方误差即可得到 R_1, R_2, \dots, R_n 上的最优输出值 \hat{C}_n ,它是 R_n 上的样本所有输出的均值:

$$\min \sum_{x \in R_n} (y_i - f(x_i))^2 \quad (2)$$

$$\hat{C}_n = \text{ave}(y_i | x_i \in R_n) \quad (3)$$

假设 (j, s) 为样本集第 j 个输入变量和该变量的一个值 s ,可以通过 (j, s) 将样本集输入空间,划分为 $R_1(j, s), R_2(j, s)$ 两个部分:

$$\begin{cases} R_1(j, s) = \{x | x^{(j)} \leq s\} \\ R_2(j, s) = \{x | x^{(j)} > s\} \end{cases} \quad (4)$$

对划分出来的两个部分,通过遍历所有 (j, s) 对,求解:

$$\min_{j, s} \left[\min_{x_i \in R_1(j, s)} (y_i - c_1)^2 + \min_{x_i \in R_2(j, s)} (y_i - c_2)^2 \right] \quad (5)$$

结合两个部分的最优输出 \hat{C}_1 和 \hat{C}_2 ,获得 (j, s) 对,对当前的输入空间进行划分,对划分完成的子节点循环上述操作,直到划分满足条件为止,这样就生成一棵二分回归树^[9]:

$$\begin{cases} \hat{C}_1 = \text{ave}(y_i | x_i \in R_1(j, s)) \\ \hat{C}_2 = \text{ave}(y_i | x_i \in R_2(j, s)) \end{cases} \quad (6)$$

构建随机森林时,通过从样本集有放回的抽取样本,抽取 k 次,每抽取一次便可训练一棵回归树,总共生成 k 棵树。对于不同的 k 值,随机森林模型具有不同的性能,模型训练时用均方误差 MSE 衡量模型的好坏,均方误差 MSE 越小,模型越好,均方误差 MSE 的计算式为:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (7)$$

式中 y_i 和 \hat{y}_i 分别为爆破参数实际值和模型预测值。

分别计算不同回归树数量对应的均方误差,取均方误差最小的 k 值,建立随机森林选择模型,流程见图 1。

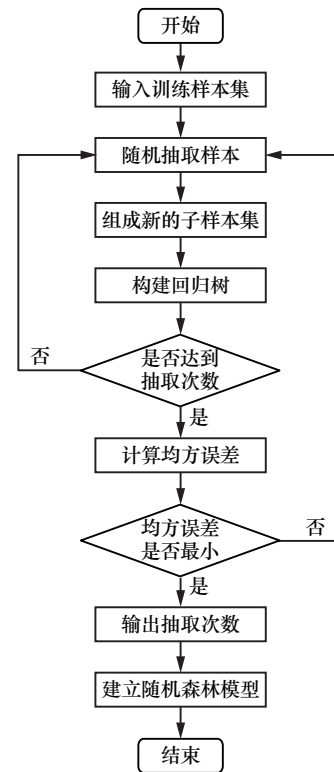


图 1 随机森林算法训练流程

2 光面爆破参数选择模型

2.1 输入输出参数及数据样本

隧(巷)道掘进光面爆破设计所要考虑的参数很多,为保证模型的准确性又尽可能不使模型太过复杂,

选取地质状况(用围岩质量分级 BQ 进行表征)、炮眼直径、隧(巷)道断面大小 3 个参数作为模型的输入参数。最小抵抗线长度、炮眼间距,周边炮眼密集系数等诸多因素都影响着光面爆破效果。最小抵抗线长度控制着岩石能否“爆下来”,炮眼间距、周边眼密集系数(可由炮眼间距和最小抵抗线长度求出)控制着隧(巷)道成型“是否规整”,炮眼深度控制着隧(巷)道掘进爆破工程的掘进效率,所以选择最小抵抗线长度、炮眼间距、炮眼深度作为光面爆破参数选择模型的输出参数。

构建模型所用数据来自文献中经典光面爆破案例^[9],详见表 1,共 31 个样本,其中 X_1 为围岩质量分级 BQ; X_2 为开挖断面, m^2 ; X_3 为炮眼直径, mm ; Y_1 为周边眼间距, cm ; Y_2 为最小抵抗线长度, cm ; Y_3 为炮眼深度, m 。

表 1 原始样本集

序号	X_1	X_2	X_3	Y_1	Y_2	Y_3
1	2	40.21	48	61.5	78.5	2.9
2	2	43.67	40	62.5	75	3
3	3	50	50	58	68	1.8
4	1	6.25	40	45	56	2.3
5	4	13.35	41	33.5	45.5	1
6	4	14.4	40	42	50	1
7	4	32	40	41	55	1
8	1	96.2	48	50	70	5
9	1	60.4	48	50	65.5	5
10	3	50	43	50	60	1.6
11	4	23	42	32.5	50	1.1
12	2	42.0	42	62.5	68	2.25
13	5	45.9	42	80	70	2.5
14	3	48.3	42	57.5	70	1.2
15	3	49.0	42	60	55	1.2
16	3	46.4	42	60	70	2.45
17	1	41.7	42	55	49.5	1.8
18	1	40.9	42	62	75	1.5
19	3	47.6	42	50	55	1.35
20	3	48.2	42	45	45	1.655
21	4	44.7	42	32.5	32.5	1.75
22	5	17.3	42	59	75	1.3
23	2	32.2	42	55	65	2
24	3	26.4	46.5	53	70	2.3
25	2	40	44	70	75	2.5
26	4	15.3	40	37	50	1.1
27	2	46.35	42	62	75	3.5
28	3	42.94	42	47.5	55	2
29	3	18.10	40	55	70	1.3
30	4	17.35	42	59	60	1.4
31	3	32.2	40	65	70	4

2.2 模型参数选择

采用随机森林算法构建光面爆破参数选择模型,其中很重要的模型参数就是回归树数量。模型构建过程就是通过训练光面爆破训练样本,确定适合模型的回归树数量,以达到更好的光面爆破参数选择效果^[10]。

据相关学者的研究^[11],随机森林模型中回归树数量越小,模型性能越差,所以一般选择较大的回归树数量。但模型的复杂度与随机森林模型中回归树数量成正比,构建随机森林模型所花费的时间也会更长,而且随着随机森林模型中回归树的增加,模型性能并不一定会提高。对于隧(巷)道光面爆破参数选择模型而言,合适的回归树数量为 50~1 000,因此采用搜索算法在 50~1 000 范围内搜索出使均方误差 MSE 最小的模型回归树数量。搜索时,先以 10 为步长分别计算各回归树数量对应的均方误差。取到均方误差最小后,在其附近以 1 为步长,分别计算均方误差,取此时均方误差最小的回归树数量作为随机森林模型最终回归树数量。

2.3 模型训练

采用 python 语言作为模型训练编码语言。Scikit-learn 是针对 Python 编程语言的免费软件机器学习库,其内包含了各种机器学习算法,也包括随机森林算法。通过调用 Scikit-learn 库的随机森林算法对隧(巷)道光面爆破样本集进行训练。

从样本集中随机选取 7 个样本作为测试集,其他 24 个样本作为训练集。模型训练时,对 24 个样本有放回地随机抽取 k 次,生成 k 个回归树,共同构成随机森林隧(巷)道光面爆破参数选择模型。

这里以对最小抵抗线长度选择的搜索过程为例,先在区间 $[50, 1 000]$ 之间以 10 为步长分别计算不同回归树数量 k 对应的均方误差,结果如图 2 所示。

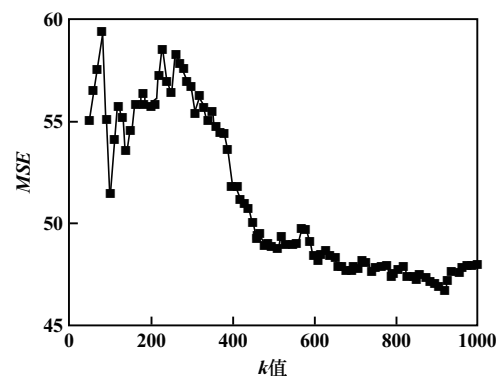


图 2 均方误差随 k 值变化图

从图 2 可以看出, k 值在 920 附近时 MSE 取得最

小值,因此,以1为步长再计算 k 值在 [915, 925] 区间内模型的均方误差,结果如图3所示。从图3可以看出, $k=919$ 时,均方误差最小,因此对最小抵抗线长度进行参数选择时选取 k 为 919。

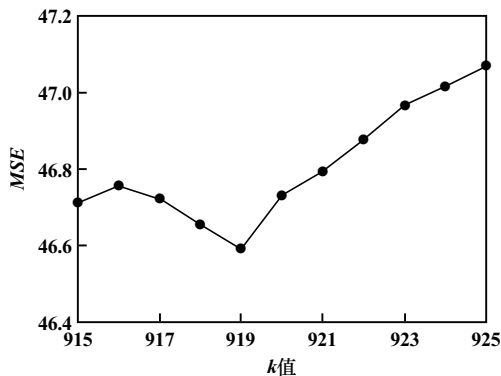


图3 [915, 925] 区间内均方误差随 k 值变化图

按照同样的方法,对周边眼间距、炮眼深度分别进行计算,找出最佳的回归树数量 k ,具体见表2。

表2 各爆破参数对应的最佳回归树数量

爆破参数	最佳回归树数量 k
周边眼间距	948
最小抵抗线长度	919
炮眼深度	87

3 预测结果及分析

将表1中序号26~31共7个样本数据对应的 $X_1 \sim X_3$ 值导入模型算法,对周边眼间距、最小抵抗线长度和炮眼深度3个参数分别进行测试,回归树数量分别为948、919和87。经过随机森林模型运算,最小抵抗线长度、周边眼间距、炮眼深度模型预测值与样本实际值对比见表3。

表3 各参数模型预测值与实际值对比

样本编号	周边眼间距/cm		最小抵抗线长度/cm		炮眼深度/m	
	实际值	预测值	实际值	预测值	实际值	预测值
1	70	57.24	75	67.39	2.5	2.25
2	65	52.24	70	63.09	4	1.55
3	62	61.48	75	69.73	3.5	2.27
4	37	41.12	50	50.37	1.1	1.02
5	59	36.55	60	50.77	1.4	1.16
6	55	49.11	70	60.13	1.3	1.30
7	47.5	58.25	55	57.44	2	1.51

用均方误差 MSE 和相关系数 r 对周边眼间距、最小抵抗线长度和炮眼深度的模型选择性能进行评估,相关系数 r 的计算式为:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (8)$$

式中 x 和 y 分别代表模型预测值和样本实际值。

相关系数介于0至1之间, r 越大,代表两组数据相关性越好。各爆破参数评估指标值见表4。

表4 各爆破参数模型选择性能评估值

评估指标	周边眼间距/cm	最小抵抗线长度/cm	炮眼深度/m
MSE	142.46	46.02	1.13
r	0.40	0.89	0.67

从表3~4可以看出,在对3种光面爆破参数的选择中,最小抵抗线长度模型预测值与样本实际值之间的相关性最好,炮眼深度次之,周边眼间距的模型预测值与样本实际值之间的相关性最差。7个样本的周边眼间距、最小抵抗线长度和炮眼深度的模型预测值与实际值平均相对误差分别为17%、8.8%和22%,其中,最小抵抗线长度预测值与实际值平均相对误差最小。可以看出,一方面,对于最小抵抗线长度的选择,模型有一定准确性,能够满足工程需要,具有一定的应用价值;另一方面,受当前案例样本数据限制,模型对周边眼间距、炮眼深度的选择效果欠佳,但是随着样本数量增加、样本数据更可靠,模型将会具有更好的性能。

4 结 语

1) 通过随机森林算法训练经典的光面爆破数据样本,确定合适的随机森林回归树数量,构建了隧(巷)道光面爆破参数选择模型。该算法对样本数量要求不高,具有较好的工程应用价值。

2) 使用了24个样本训练模型、7个样本测试模型选择效果,设置炮眼间距、最小抵抗线长度、炮眼深度的回归树数量分别为948、919和87,测试结果表明,3个参数的模型选择参数值与测试样本对应参数值之间的平均相对误差分别为17%、8.8%和22%,为光面爆破参数相关模型研究提供了参考。

3) 基于随机森林的光面爆破参数选择模型对于最小抵抗线长度预测效果较优,在实际应用中,可使用该模型对最小抵抗线长度进行预测,与其他模型配合使用,以求达到更好的爆破参数选择效果。

参考文献:

[1] WU Z, LI Q, KONG D, et al. The ANP-Fuzzy-TOPSIS model for the optimization of the scheme of large-section blasting[J]. Arabian Journal of Geosciences, 2020, 13(2):47.

- [3] 贺建清,林孟源,陈立国,等.考虑围压影响的湖相软黏土 K0 固结经验蠕变模型[J].自然灾害学报,2022,31(1):147-156.
- [4] 魏德永.滇池湖相软土固结渗透试验及参数取值研究[D].昆明:昆明理工大学,2021.
- [5] 李荐华,甘建军,万胜,等.湖相软土固结变形特性试验及微观分析[J].中国农村水利水电,2019(8):120-128.
- [6] 刘嘉正.里下河泻湖相软土土性参数变异性统计分析[J].土工基础,2023,37(6):979-981.
- [7] 甘建军,万胜,李荐华,等.基于沉积环境与工程地质特征的湖相软土分类研究——以鄱阳湖流域四种典型软土为例[J].地质论评,2019,65(4):983-992.
- [8] 邵勇,阎长虹,许宝田,等.湖相软土流变模型识别及其工程应用分析[J].岩土力学,2012,33(8):2383-2387.
- [9] 付贵海,阳标,张根宝,等.湖相软土单轴蠕变特性及其模型研究[J].湖南城市学院学报(自然科学版),2023,32(3):1-7.
- [10] 谷任国,房营光.软土流变的物质基础及流变机制探索[J].岩土力学,2009,30(7):1915-1919.
- [11] 阮永芬,朱尚万,乔文件,等.湖相沉积泥炭质土蠕变特性[J].岩土工程学报,2022,44(增刊1):166-171.
- [12] 邓宗伟,朱志祥,曾向军,等.不同应力水平下湖相软土流变特性研究[J].工业建筑,2014,44(6):67-72.
- [13] 中华人民共和国交通运输部.JTG3430—2020公路土工试验规程[S].北京:人民交通出版社,2020.
- [14] 刘世明,曾国熙.一个考虑次时间效应的一维固结模型[J].浙江大学学报(自然科学版),1990(5):49-60.
- [15] 丁林楠,李国英,魏匡民.描述土体级配分布的级配方程及其适用性[J].岩土力学,2022,43(增刊1):173-183.
- [16] 李广信,张丙印,于玉贞.土力学(第3版)[M].北京:清华大学出版社,2022.
- 引用本文:李和志,任虎程,葛莎,等.含砂量对湘北湖相软土流变特性影响研究[J].矿冶工程,2024,44(3):38-43.

(上接第33页)

参考文献:

- [1] 徐永福.膨胀土的水力作用机理及膨胀变形理论[J].岩土工程学报,2020,42(11):1979-1987.
- [2] 徐永福,程岩,唐宏华.膨胀土边坡失稳特征及其防治技术标准化[J].中南大学学报(自然科学版),2022,53(1):1-20.
- [3] 吴镇清,葛允雷,杨果林,等.膨胀土路堑边坡支挡结构形式试验研究[J].铁道建筑,2017,57(11):79-81.
- [4] 赵乐涛,杨兆林.东边坡不良地质体勘查及安全治理研究[J].矿业工程,2021,19(3):10-14.
- [5] 刘妮娜.普速铁路加筋膨胀土路堤填筑与现场监测研究[J].长沙理工大学学报(自然科学版),2023,20(6):27-38.
- [6] 肖杰,杨和平.膨胀土堑坡格栅加筋柔性支护的数值模拟[J].公路交通科技,2016,33(6):1-8.
- [7] 赵亮,龚壁卫,李青云,等.加筋膨胀土边坡土工格栅的导水作用研究[J].长江科学院院报,2012,29(6):44-47.
- [8] 许英姿,廖旭航,苏超,等.加筋宽度与间距对膨胀土特性影响研究[J].铁道科学与工程学报,2021,18(10):2583-2590.
- [9] Riccardo S, Simon J W, Marti L C. Conceptual hydraulic conductivity model for unsaturated soils at low degree of saturation and its application to the study of capillary barrier systems[J]. Journal of Geotechnical and Geoenvironmental Engineering, 2020,146(10):04020106.
- [10] Kuo-Hsin Y, Joseph N T, Vo D A H, et al. Numerical evaluation of reinforced slopes with various backfill-reinforcement-drainage systems subject to rainfall infiltration [J]. Computers and Geotechnics, 2018,96:25-39.
- [11] 曲永新,张永双,杨俊峰,等.中国膨胀性岩、土一体化工程地质分类的理论与实践[C]//中国地质学会.工程地质学报,2000:140-164.
- [12] 肖杰,杨和平,符浩,等.膨胀土边坡加筋作用的对比模型试验[J].中国公路学报,2014,27(7):24-31.
- 引用本文:李昱,常锦,郭剑雄.考虑格栅水力特性的柔性支护处治膨胀土边坡的数值模拟分析[J].矿冶工程,2024,44(3):28-33.

(上接第37页)

- [2] 单仁亮,汪学清,高文蛟,等.人工神经网络在巷道爆破中的应用研究[J].岩石力学与工程学报,2007,26(增刊1):3322-3328.
- [3] 谢超群,李启月,刘玉丰,等.基于SVR的隧道掘进光面爆破效果预测模型[J].矿冶工程,2022,42(2):28-31.
- [4] 蒋复量,杨文超,郭锦涛,等.基于鲸鱼算法的巷道掘进爆破参数优化与应用[J].南华大学学报(自然科学版),2019,33(2):27-32.
- [5] 李伟,杨仁树.中深孔光面爆破技术在岩石巷道掘进中的应用研究[J].矿冶工程,2013,33(3):21-23.
- [6] Probst P, Wright M, Boulesteix A. Hyperparameters and tuning strategies for random forest[J]. Wiley Interdisciplinary Reviews-Data Mining and Knowledge Discovery, 2019,9(3):1301-1316.
- [7] CHEN C, ZHANG Q, Kashani M H, et al. Forecast of rainfall distribution based on fixed sliding window long short-term memory[J]. Engineering Applications of Computational Fluid Mechanics, 2022,16(1):248-262.
- [8] Janitza S, Celik E, Boulesteix A. A computationally fast variable importance test for random forests for high-dimensional data [J]. Advances in Data Analysis & Classification, 2018,12:885-915.
- [9] 李航.统计学习方法[M].北京:清华大学出版社,2019.
- [10] 刘玉丰.隧(巷)道光面爆破参数选择与效果预测模型研究[D].长沙:中南大学,2022.
- [11] 刘敏,郎荣玲,曹永斌.随机森林中树的数量[J].计算机工程与应用,2015,51(5):126-131.
- 引用本文:侯胜心,马晶晶,唐虹靖,等.基于随机森林算法的隧(巷)道光面爆破参数选择模型研究[J].矿冶工程,2024,44(3):34-37.