

文章编号: 1673- 9469(2011) 01- 0064- 04

基于熵权法的PHC管桩承载力组合预测

李万庆, 李铮

(河北工程大学 经济管理学院, 河北 邯郸 056038)

摘要: 为克服单项预测方法产生的误差, 利用灰色模型 $GM(1, N)$ 、多元线性回归、BP神经网络等3种单项预测方法建立组合预测模型, 并采用熵值法确定加权系数。通过对PHC管桩承载力进行比较预测, 结果显示 $GM(1, N)$ 法平均绝对百分比误差(MAPE)值为5.4%, 多元线性回归法的MAPE为3.0%, BP神经网络法的MAPE为2.8%, 组合预测法的MAPE为2.3%。因此组合预测法精度较高, 实用性更强。

关键词: PHC管桩; 熵值法; 组合预测; BP神经网络;

中图分类号: TU473.1+1

文献标识码: A

Combination forecasting of bearing capacity of PHC pipe pile based on entropy method

LI Wan-qing, LI Zheng

(School of Economic and Management, Hebei University of Engineering, Hebei Handan 056038, China)

Abstract The combination forecasting model was building to overcome the potential errors generated by single forecast model on the basis of the grey system $GM(1, N)$, multiple linear regression and back-propagation neural network, and the weighting coefficients were determined by the entropy method. The contrast test was conducted to predict the bearing capacity of PHC pile, and the results show that the method means absolute percentage error (MAPE) of $GM(1, N)$ is 5.4%, the MAPE of multiple linear regression is 3.0%, the MAPE of BP neural network method is 2.8%, and the MAPE of the combined forecasting method is 2.3%. Therefore the combined forecasting has high precision and practicability.

Key words: PHC pipe; entropy method; combination forecasting; BP neural network

承载力作为PHC管桩质量的一个重要评价指标, 其预测方法有多种, 较常用的如 $GM(1, N)$ 法、多元线性回归法和BP神经网络法^[1-5]等。由于不同的预测方法是从不同的角度搜集样本信息, 所以预测精度也各有差异。组合预测模型是将不同的预测方法组合起来, 综合利用各种预测方法所提供的信息, 以适当的加权平均形式得出的预测模型。寇毅^[6]通过建立关于时间序列平均移动模型、灰色预测模型、非线性回归模型的组合预测模型, 利用层次分析法得出权重, 在公路货运周转量预测方面取得了良好的效果; 犹勇^[7]以误差平方和最小为目标, 建立关于ARIMA预测模型、卡尔曼滤波预测模型和神经网络预测模型的组合预测

模型, 利用遗传算法求出权重, 在城市交通流量预测中取得了良好效果。

熵权法是一种客观赋权方法, 通过全面研究各个预测方法的误差指标值, 以误差指标值的大小确定各个预测方法在组合预测模型中的权重。本文选取熵权法确定组合模型的权重系数, 建立关于 $GM(1, N)$ 法、多元线性回归法和BP神经网络法的组合预测模型并对PHC管桩承载力进行了预测。

1 组合预测模型的建立

1.1 组合预测理论

设有 m 种单项预测方法, 用 y_{it} 表示第 i 种预

收稿日期: 2010-11-20

作者简介: 李万庆(1954-), 男, 河北临漳人, 教授, 从事工程管理方面的研究。

测方法在 t 时刻的预测值, 则 m 种单项预测方法的组合预测值为

$$y_t^v = \sum_{i=1}^m \omega_i y_{it}, t = 1, 2, \dots, N \quad (1)$$

式中: ω_i 为第 i 种预测方法的加权系数, $i = 1, 2, \dots, m$ 。

1.2 单项预测模型选择

单项预测模型中, GM(1, N) 模型与多元线性回归模型适用于长期预测, 前者不需要分析大量样本但易产生较大误差, 后者预测精度较高但对数据资料要求比较高。BP 神经网络模型所特有的处理复杂非线性问题的功能将弥补因影响因素间关系复杂所导致的建模困难的缺陷, 其缺点在于自身的不稳定性会引起预测结果的不稳定性。本文选取 GM(1, N) 模型、多元线性回归模型、BP 神经网络模型等 3 种单项模型进行组合预测, 利用其互补性, 将短期预测与长期预测有机地结合起来。

(1) GM(1, N) 预测法

GM(1, N) 模型是一阶 N 个变量的微分方程型预测模型, 设 $x_1^{(0)} = (x_1^{(0)}(1), x_1^{(0)}(2), \dots, x_1^{(0)}(n))$ 为所要预测管桩承载力序列; $x_N^{(0)} = (x_N^{(0)}(1), x_1^{(0)}(2), \dots, x_N^{(0)}(n))$ 为影响承载力的因素序列; $x_i^{(1)}(i = 1, 2, \dots, N, N \geq 2)$ 为 $x_i^{(0)}$ 的累加生成数序列(1-AGO 序列); $z_1^{(1)}$ 为 $x_1^{(1)}$ 的紧邻均值生成序列; 则 GM(1, N) 的灰微分方程为: $x^{(0)}(k) = \alpha z_1^{(1)}(k) = \sum_{i=2}^N b_i x_i^{(1)}(k) (k \in K, K = 1, 2, \dots, N)$ 。

(2) 多元线性回归预测法

设承载力 y 为因变量, 自变量 $x_i, (i = 1, 2, 3, \dots, k)$ 为影响因素。则建立多元线性回归方程为:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon$$

式中: $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ —待估的回归系数; ε —不可观测的随机误差项, 具有零均值、同方差且服从正态分布。

(3) BP 神经网络法

选取 3 层 BP 神经网络对 PHC 管桩承载力进行预测。

步骤 1: 应用 $neuff()$ 构造神经网络函数:

$net = neuff(\min \max(P), [s1, s2], \{\tan sig', \log sig'\}, \{traingdx', \learngdm'\})$;

式中, p —PHC 管桩承载力影响因素输入矩阵; $s1, s2$ —隐含层、输出层神经元数目; $\tan sig'$ 、 $\log sig'$ —隐含层/输出层神经元采取的传递函数;

$traingdx'$ —网络采取的训练函数; $learngdm'$ —网络采取的学习函数; net —生成的 BP 网络。

步骤 2: 采用 $iniff()$ 函数进行权值初始化:

$[w1, b1, w2, b2] = iniff(P, sl, \{tansig', s2, \log sig'\})$;

其中: $w1, w2$ —输入层到隐含层、隐含层到输出层的连接权矩阵; $b1, b2$ —隐含层各单元、输出层各单元阈值矩阵。

1.3 熵权法确定加权系数

(1) 熵权法的基本原理

熵的概念源于热力学, 它最先由 Shannon 引入信息论, 现已在工程技术、社会经济等领域得到十分广泛的应用。依据信息论的基本原理, 熵是系统无序程度的一个度量。当系统可能处于几种不同状态, 每种状态出现的概率为 p_i 时, 该系统的熵^[8-9] 定义为

$$E = - \sum_{i=1}^m p_i \ln p_i, i = 1, 2, \dots, m \quad (2)$$

设有 m 种单项预测方法, n 项误差评价指标, 形成的原始指标矩阵 $R = (r_{ij})_{m \times n}$, 对于某个指标 r_j , 有

$$E_j = - k_i \sum_{i=1}^m p_{ij} \ln p_{ij}, k_i \geq 0 \quad (3)$$

式中: $p_{ij} = r_{ij} / \sum_{i=1}^m r_{ij}, r_{ij} \geq 0, i = 1, 2, \dots, m, j = 1, 2, \dots, n$

(2) 熵权法确定组合预测加权系数的步骤

熵权法的基本思想是依据熵的概念和性质, 把各个单项预测方法的各种误差所包含的信息量化, 进而建立基于熵的组合预测权重模型。该方法的具体步骤^[10] 如下所示:

步骤 1: 设 PHC 管桩承载力指标序列 $\{y_t, t = 1, 2, \dots, N\}$ 存在 m 种单项预测方法并对其进行预测, 设第 i 种单项预测方法在 t 时刻的预测值为 y_{it} , 则第 i 种单项预测方法在第 t 时刻的预测相对误差 e_{it} 为

$$e_{it} = \begin{cases} 1, \\ | (y_t - y_{it}) / y_t | \end{cases} \quad (4)$$

其中, 当 $| (y_t - y_{it}) / y_t | \geq 1$ 时, $e_{it} = 1$, 反之, $e_{it} = | (y_t - y_{it}) / y_t |, i = 1, 2, \dots, m, t = 1, 2, \dots, N$ 。

步骤 2: 将各种单项预测方法的预测相对误差序列进行归一化, 即计算第 i 种单项预测方法第 t 时刻的预测相对误差的比重 p_{it} 。

$$p_{it} = \frac{e_{it}}{\sum_{i=1}^N e_{it}}, t = 1, 2, \dots, N \quad (5)$$

显然, $\sum_{i=1}^N p_{it} = 1, i = 1, 2, \dots, m$

步骤 3: 计算第 i 种单项预测方法的预测相对误差的熵值。

$$h_i = -k \sum_{i=1}^N p_{it} \ln p_{it}, i = 1, 2, \dots, m \quad (6)$$

式中, $k = 1/\ln N, 0 \leq h_i \leq 1$ 。

步骤 4: 计算第 i 种单项预测方法预测相对误差序列的变异程度系数。因为 $0 \leq h_i \leq 1$, 根据系统某项指标的熵值的大小与其变异程度相反的原则, 定义第 i 种单项预测方法的预测相对误差序列的变异程度系数 d_i 为

$$d_i = 1 - h_i, i = 1, 2, \dots, m \quad (7)$$

步骤 5: 计算得各种预测方法的加权系数 ω_i 为

$$\omega_i = \frac{1}{m-1} \left[1 - \frac{d_i}{\sum_{i=1}^m d_i} \right], i = 1, 2, \dots, m \quad (8)$$

且 $\sum_{i=1}^m \omega_i = 1$

步骤 6: 计算得组合预测值 \hat{y}_t 为

$$\hat{y}_t = \sum_{i=1}^m \omega_i y_{it}, t = 1, 2, \dots, N \quad (9)$$

2 实例分析

选取文献[11]中前 30 根 PHC 管桩数据作为样本, 分别采用 GM(1, N) 法、多元线性回归法、BP 神经网络法以及基于这 3 种方法的组合预测模型对 PHC 管桩承载力进行对比分析。计算结果如表 1 所示。

表 1 PHC 管桩承载力预测结果

Tab. 1 The prediction of PHC pile bearing capacity

桩号	实际值	GM(1, N) 法	多元线性回归法	BP 神经网络法	组合预测值
1	3 800	3 800.000	3 909.805	3 899.922	3 869.766
2	3 800	3 342.290	3 948.986	3 886.898	3 725.295
3	3 800	4 645.283	3 879.194	3 875.012	4 134.425
4	4 009	4 587.073	3 877.132	3 902.263	4 123.221
5	4 009	4 120.164	3 937.699	3 917.148	3 992.041
6	4 009	3 959.479	3 873.323	3 924.705	3 919.124
7	4 001	4 093.503	3 874.995	3 956.068	3 974.918
8	4 009	3 936.150	3 860.591	3 872.395	3 889.792
9	3 800	4 039.071	3 939.206	3 907.462	3 962.190
10	3 800	3 771.919	3 837.566	3 893.255	3 833.939
11	3 800	3 854.797	3 820.206	3 865.254	3 846.646
12	4 000	3 683.712	4 034.057	4 028.463	3 914.861
13	4 000	3 919.398	4 025.362	4 000.701	3 981.737
14	4 000	4 258.868	4 048.954	3 967.898	4 092.541
15	4 000	3 939.961	3 999.629	3 978.279	3 972.603
16	4 000	4 131.695	4 110.241	4 026.094	4 089.602
17	4 000	4 113.145	4 044.206	4 030.637	4 062.824
18	4 000	4 038.522	3 988.597	3 954.835	3 994.188
19	4 000	4 040.059	4 041.501	3 979.724	4 020.649
20	4 000	3 854.008	3 955.803	4 011.407	3 940.038
21	4 691	3 916.388	4 721.217	4 693.981	4 442.650
22	5 000	6 248.163	5 936.635	5 013.577	5 736.633
23	6 170	6 383.194	6 001.462	6 284.280	6 222.579
24	6 300	5 918.629	5 690.332	6 120.748	5 908.720
25	6 000	5 918.701	5 690.485	6 120.983	5 908.873
26	4 000	4 060.391	3 996.568	3 961.832	4 066.493
27	4 000	3 731.884	4 058.271	4 265.376	4 017.231
28	4 064	4 018.596	4 019.590	4 096.148	4 044.480
29	4 064	3 898.016	4 072.582	4 369.201	4 111.911
30	4 052	4 043.989	3 983.807	3 976.367	4 001.513
MAPE / %		5.4	3.0	2.8	2.3

对表 1 结果分析可以看出, 在 PHC 管桩承载力预测实例中, GM(1, N) 法实际值与预测值平均绝对百分比误差(MAPE)为 5.4%, 大于 5%, 预测精度偏大。多元线性回归法与 BP 神经网络法实际值与 MAPE 分别为 3.0% 和 2.8%, 均小于 5%, 预测精度相近。而组合预测法实际值与 MAPE 为 2.3%, 小于 5%, 比 GM(1, N) 法、多元线性回归法、BP 神经网络法预测效果更好, 精度更高。

3 结束语

组合预测法的预测值与实际值较为接近, 且组合预测的 MAPE 小于 3 种单项预测模型。基于熵权法的组合预测达到了预期的预测效果, 不仅提高了预测的精度, 还尽可能地保留了 3 种单项预测方法所涵盖的信息。

参考文献:

- [1] 刘思峰, 党耀国, 方志耕. 灰色系统理论及应用[M]. 北京: 科学出版社, 2004.
- [2] 余锦华, 杨维权. 多元统计分析与应用[M]. 广州: 中

山大学出版社, 2005.

- [3] 董长虹. Matlab 神经网络与应用[M]. 北京: 国防工业出版社, 2007.
- [4] 闻新, 周露, 王丹力, 等. Matlab 神经网络应用设计[M]. 北京: 科学出版社, 2000.
- [5] 原建博, 鹿群, 李慧霞. 后注浆提高钻孔灌注桩单桩承载力分析[J]. 河北工程大学学报(自然科学版), 2010, 27(4): 18-22.
- [6] 寇毅. 组合预测模型在公路货运周转量预测中的运用[J]. 中国水运, 2007, 4216-217.
- [7] 犹勇. 组合预测方法在城市交通流量预测中的运用[J]. 重庆工商大学学报(自然科学版), 2007(4): 122-125.
- [8] 王琦, 王花兰, 吴璇, 等. 基于熵值法的城市汽车保有量组合预测[J]. 交通科技与经济, 2009, 6: 53-56.
- [9] 熊崇俊, 宁宣熙, 潘颖莉. 基于熵的组合预测法研究[J]. 科技进步与对策, 2006(11): 68-69.
- [10] 陈华友. 组合预测方法有效性理论及应用[M]. 北京: 科学出版社, 2008.
- [11] 张力. 锤击桩单桩极限承载力的神经网络预测研究[D]. 南京: 东南大学, 2006.

(责任编辑 马立)

(上接第 51 页)

分期完成生态恢复。初期以自然恢复的野生先锋草本植物为主体, 选择适合当地气候特征的浅根类耐性植物加速系统恢复^[9-10]; 在这些先锋植物对环境进行影响改善后, 中后期逐渐引入生态效应及观赏性高的植物类群, 使恢复后的生态系统不断向较理想的顶极群落演替。

4 结语

废弃垃圾填埋场的治理与开辟游览绿地相结合是环境工程建设的积极措施, 对城市的发展有着积极的作用。生态恢复与景观营造是垃圾山改造的两大阶段, 通过人工干预加速受损生态系统的恢复, 待其相对稳定后, 进行景观的营造, 为城市增加公共绿地空间。河南省在废弃垃圾场改造过程中要充分结合本地特点, 以生态恢复为基本, 以景观营造为提升, 从而使垃圾场废弃地发生质的改变。

参考文献:

- [1] 杨锐, 王浩. 景观突围- 垃圾填埋场的生态恢复与景

观重建[J]. 城市发展研究, 2010(8): 81-86.

- [2] 林学端, 廖文波. 垃圾填埋场植被恢复及其环境影响因子的研究[J]. 应用与环境生物学报, 2002, 8(6): 571-577.
- [3] 康汉起, (韩) 吴海泳. 寻找失落的家园- 韩国首尔市兰芝岛世界杯公园生态恢复设计[J]. 中国园林, 2007(8): 55-61.
- [4] 虞蔚君, 丁绍刚. 生命景观- 从垃圾填埋场到清泉公园[J]. 风景园林, 2006(6): 26-31.
- [5] 张金伟, 常江. 城市废弃地景观与生态恢复研究[J]. 现代城市研究, 2007(11): 40-49.
- [6] 计成. 陈植注释, 园冶注释[M]. 北京: 中国建筑工业出版社, 2009.
- [7] 谢东, 敦婉如. 植被生态对城市垃圾的处理利用及改良[J]. 青岛建筑工程学院学报, 2002(3): 41-43.
- [8] 王罗春, 赵由才, 陆雍森, 等. 垃圾填埋场稳定化及其研究现状[J]. 城市环境与城市生态, 2000, 5(13): 36-39.
- [9] 邓毅. 城市景观的生态化设计[J]. 城市问题, 2002(6): 17-20.
- [10] 彼得·拉兹. 废弃场地的质变[J]. 风景园林, 2005(1): 29-36.

(责任编辑 刘存英)