

Improved Entropy Method Establishing Combination Model for Prediction Foundation Settlement

Yanping Gao^{1, a}, Yang Yang^{2, b} and Jie Wang^{1, c}

¹Hebei University of Engineering, Handan, Hebei Province, China

²Hebei Institute of Designing Water Resources and Hydropower, Tianjin, China

^a760761994@qq.com, ^b847200306@qq.com, ^c378205421@qq.com

Keywords: Improved entropy method; Exponential model; Logistic model; Combination model; Settlement prediction

Abstract. On the basis of the exponential model and Logistic model presented at the previous, using improved entropy method to calculate each model weight coefficient to establish a new combination model for prediction foundation settlement, to achieve the goal of comprehensive advantages of both models and smaller prediction error. Respectively using exponential model, Logistic model and a combination of both a model to process power plant foundation settlement data, and with squared error, mean absolute error, mean absolute percentage error, mean square error for each model to evaluate the prediction accuracy. The results show that Prediction error of combination model established by improved entropy method is less than a single forecast model. The combination model has Higher prediction accuracy and effectively predict the power plant foundation settlement.

改进熵值法建立组合模型预测地基沉降

高艳平^{1, a}, 杨洋^{2, b}, 王杰^{1, c}

1. 河北工程大学 水电学院 水利工程, 中国 河北邯郸, 056038

2. 河北省水利水电勘测设计研究院, 中国 天津, 300250

^a760761994@qq.com, ^b847200306@qq.com, ^c378205421@qq.com

摘要: 在前人提出的指数模型和 Logistic 模型基础上, 利用改进熵值法计算各模型权重系数, 建立新的沉降预测组合模型, 以期综合两种模型的优点, 得到较小误差。以某发电厂的地基沉降监测为例, 分别运用指数模型、Logistic 模型以及二者组合模型处理实测沉降数据, 以误差平方和、平均绝对误差、平均绝对百分比误差、均方误差等评价指标对各模型预测精度进行比较。结果表明, 改进熵值法确定的组合模型的预测误差小于单一预测模型, 其预测精度较高, 能够有效的预测发电厂地基基础沉降。

关键字: 改进熵值法; 指数模型; Logistic模型; 组合模型; 沉降预测

1. 引言

沉降预测在工程建筑设计、施工和使用过程中至关重要, 由于土的固结和压缩规律非常复杂, 一些参数很难准确测出, 任何模型的提出均是建立在很多假设基础之上, 则根据实际沉降观测值和观测时间的变化曲线表现出一定的规律, 利用实际沉降观测值建立预测模型的方法被广泛应用在地基沉降中。该类方法的预测模型有很多种, 包括双曲线模型[1]、Weibull 模型[2]、MMF 模型[3]、灰色理论 GM(1, 1)模型[4]、人工神经网络[5]、遗传算法[6]等, 然而, 不同的单个沉降预测模型皆有其优缺点, 不能综合考虑地基沉降的各个因素, 预测沉降时存在一定误差和局限性, 某些情况下无法得到比较满意的预测结果。因此, 需要将多种不同的预测模型进行组合[7], 建立新的预测模型预测地基的沉降量, 以减小预测误差, 提高预测精

度。

自组合模型提出以来,许多专家和学者不断地对不同预测模型进行优化组合研究,通过实际工程分析,得到较好的预测成效。但组合模型的权重系数的确定方法有很多种,且权重系数选取的较合理时,组合模型的预测精度比较高。

本文汇聚指数模型和 Logistic 模型自身优点,进行组合,以达到更好的预测结果。基于 MATLAB 曲线拟合工具箱[8]优化拟合单一模型,运用改进熵值法求出单一模型的权重系数,建立组合模型,并结合某发电厂的工程实例比较预测精度,以验证组合模型的优越性。

2. 单一预测模型

2.1. 指数模型

指数模型[9]于 1959 年曾国熙提出的,根据太沙基一维固结理论,认为土体的平均固结度与时间之间的关系是指数函数。地基沉降预测公式为:

$$S_t = S_{\infty}(1 - e^{-rt}) \quad (1)$$

式中: S_t — t 时对应的地基累计沉降量, mm; t —沉降观测的时间, d; S_{∞} —地基基础最终沉降量, mm, r —待求参数。

指数模型参数少,求解简单,能满足一定的精度要求,其预测中长期沉降数据的精度大,且预测结果一般大于实测值。

2.2. Logistic 模型

Logistic 模型[10]又称皮尔模型、泊松模型,广泛应用于人口学、生态学等领域,是一种生长模型,预测公式为:

$$S_t = \frac{L}{1 + ae^{-bt}} \quad (2)$$

式中: S_t — t 时对应的地基累计沉降量, mm; t —沉降观测的时间, d; L , a , b —待求参数,当 $t \rightarrow \infty$ 时, $S_t \rightarrow L$, 即 L 为最终沉降量。

Logistic 曲线能反应地基沉降的发生、发展、成熟、极限四个阶段,能够预测地基沉降全过程。

3. 加权组合模型

3.1. 组合模型的建立

所谓组合模型,就是用一种新模型将多种单一模型以加权的形式组合起来,综合利用多种单一模型自身反应的有用信息。对于地基基础沉降预测,若应用 n 种预测模型进行沉降预测,每种模型在 t 时刻的沉降预测值为 $S_{it}(i=1, 2, \dots, n; t$ 为实际观测期数, $t=1, 2, \dots, N)$, S_t 为组合模型预测 t 时的沉降值,则组合模型公式为:

$$S_t = \sum_{i=1}^n \mu_i S_{it} \quad (3)$$

式中: μ_i 为第 i 种预测模型的权重系数,且 $\sum_{i=1}^n \mu_i = 1$, $\mu_i \geq 0 (i=1, 2, \dots, n)$ 。

令 Y_t 为 t 时的实测沉降值,则组合模型的预测误差 e_t 为:

$$e_t = S_t - Y_t = \sum_{i=1}^n \mu_i S_{it} - Y_t \quad (4)$$

本文的变权组合模型建立在指数模型和 Logistic 模型的基础上,设在 t 时,这两种单项模型的沉降预测值分别为 S_{1t} 、 S_{2t} , 组合模型的预测值为 S_t , 则组合模型的沉降预测数学表达式为:

$$S_t = \mu_1 S_{1t} + \mu_2 S_{2t} \quad (5)$$

式中： μ_1 、 μ_2 为指数模型、Logistic 模型所占的权重系数，且 $\mu_1 + \mu_2 = 1$ ， $\mu_1 \geq 0$ ， $\mu_2 \geq 0$ 。

为验证组合模型的预测地基基础沉降的合理性和适用性，本文选择误差平方和(SSE)、均方误差(MSE)、平均绝对误差(MAE)、平均绝对百分比误差(MAPE)作为精度的检验标准值。则各指标的表达式如下：

$$SSE = \sum_{t=1}^N (S_t - Y_t)^2 \quad (6)$$

$$MSE = \sqrt{\frac{SSE}{N}} \quad (7)$$

$$MAE = \frac{1}{N} \sum_{t=1}^N |S_t - Y_t| \quad (8)$$

$$MAPE = \frac{1}{N} \sum_{t=1}^N \frac{|S_t - Y_t|}{Y_t} \times 100 \quad (9)$$

3.2. 改进的熵值法计算权重

Bate 和 Granger 首次提出组合模型预测地基沉降，由于其预测精度高，受到国内外学者的重视。为了使组合模型的预测结果更加精确，权重的计算变得更加重要。在申农的信息论中，熵度量系统信息的不确定性，然而，地基基础的实测值与预测模型计算值之间的残差具有不确定性，因此用熵值法确定沉降组合模型的权重是可行的。熵值法是根据各指标反应信息量的大小以及原始数据间的联系确定各指标权重系数，是客观赋值法。当模型信息越多，熵越小，其权重越大；反之亦然。然而传统的熵值法未考虑相对误差序列本身的特点，本文采用文献^[11]提出的改进熵值法计算权重，从相对误差序列服从正态分布的特性出发，按照等概率原则划分区间，并赋予特定区间的误差序列一个修正权重值。其具体计算步骤如下：

(1) 设 $\{Y_t, t=1, 2, \dots, N\}$ 为地基沉降实测序列，有 n 中拟合模型进行沉降预测，则 S_{it} 为第 i 种沉降预测模型预测 t 时刻的沉降值， e_{it} 为实测值与第 i 中预测模型在 t 时刻的沉降预测值之间的误差序列，可用下式计算：

$$e_{it} = \begin{cases} 1, & \text{当 } \left| \frac{S_t - S_{it}}{S_t} \right| = 1 \text{ 时} \\ \left| \frac{S_t - S_{it}}{S_t} \right|, & \text{当 } 0 = \left| \frac{S_t - S_{it}}{S_t} \right| < 1 \text{ 时} \end{cases} \quad (10)$$

式中： $i=1, 2, \dots, n$ ； $t=1, 2, \dots, N$ 。

(2) 修正相对误差，得新的相对误差 $\{e_{it}\}$ 。把 n 种单项模型预测的相对误差 e_{it} 从小到大排列，且分三种状态：偏小状态、一般状态、偏大状态，这三种状态的区段概率分别为 25%、50%、25%，赋予误差序列的权重分别为 0.3、0.4、0.3。

(3) 对改正的相对误差进行单位化。计算出第 i 种预测模型在 t 时刻沉降预测值与实际观测值的相对误差的比重 P_{it}

$$P_{it} = \frac{e_{it}}{\sum_{t=1}^N e_{it}} \quad (11)$$

(4) 计算第 i 种单项预测模型的熵值 h_i

$$h_i = -k \sum_{t=1}^N P_{it} \ln P_{it} \quad (12)$$

式中: $k = \frac{1}{\ln N}$ 。

(5) 计算第 i 种预测模型的相对误差的变异系数 d_i 。

$$d_i = 1 - h_i \quad (13)$$

(6) 计算组合模型中 n 种预测模型的权重系数 μ_i

$$\mu_i = \frac{1}{n-1} \left(1 - \frac{d_i}{\sum_{i=1}^n d_i} \right) \quad (14)$$

本文组合两种单项预测模型, 则权重系数 $\mu_1 = 1 - \frac{d_1}{d_1 + d_2}$, $\mu_2 = 1 - \frac{d_2}{d_1 + d_2}$

(7) 计算组合模型的预测值

$$S_t = \sum_{i=1}^n \mu_i S_{it} - Y_t \quad (15)$$

4. 工程实例分析

根据某发电厂现场监测资料, 利用指数模型、Logistic 模型及二者组合模型对实测沉降数据进行拟合、预测, 以验证用改进熵值法确定组合模型的优越性。该发电厂位于某海港区内, 所揭露地层主要为第四系全新统 (Q_4) 及上更新统 (Q_3) 海相及海陆交互相地层。从该工程的主厂房监测数据中, 取代表性的观测数据, 即主厂房 B2 点进行沉降预测。其沉降观测时间自 2004 年 8 月 19 日, 到 2008 年 11 月 11 日止, 共观测 21 次。利用 MATLAB 曲线拟合工具箱优化拟合指数模型、Logistic 模型, 得到各模型表达式如下:

指数模型: $S_{1t} = 22.69(1 - e^{-0.005473})$

Logistic 模型: $S_{2t} = \frac{22.21}{1 + 4.148e^{-0.01008}}$

运用改进的熵值法计算各预测模型的权重系数, 得到组合模型的表达式为:

$$S_t = 0.415S_{1t} + 0.585S_{2t}$$

指数模型、Logistic 模型和组合模型的预测结果与实测结果的比较见表 1。

表 1 预测值与实测值的比较

观测次数	实测值 /mm	指数模型		Logistic 模型		组合模型		
		预测值/mm	误差	预测值/mm	误差	预测值/mm	误差	
拟合	1	5.2	3.65	-1.55	5.55	0.35	4.76	-0.44
	2	5.9	7.88	1.98	7.69	1.79	7.77	1.87
	3	10.1	10.47	0.37	9.54	-0.56	9.92	-0.18
	4	12.9	12.54	-0.36	11.43	-1.47	11.89	-1.01
	5	14.6	14.58	-0.02	13.68	-0.92	14.05	-0.55
	6	17.8	17.60	-0.20	17.56	-0.24	17.58	-0.22
	7	17.5	18.37	0.87	18.58	1.08	18.49	0.99
	8	18.4	19.04	0.64	19.43	1.03	19.27	0.87
	9	19.7	19.70	0.00	20.20	0.50	19.99	0.29
	10	19.7	20.06	0.36	20.60	0.90	20.38	0.68
	11	21.6	20.52	-1.08	21.05	-0.55	20.83	-0.77
	12	21.3	20.87	-0.43	21.36	0.06	21.16	-0.14
	13	22.5	21.11	-1.39	21.55	-0.95	21.37	-1.13
	14	22.3	21.31	-0.99	21.69	-0.61	21.54	-0.76
	15	22.6	21.88	-0.72	22.01	-0.59	21.96	-0.64
	16	22.4	22.18	-0.22	22.13	-0.27	22.15	-0.25
预测	17	21.6	22.40	0.80	22.18	0.58	22.27	0.67
	18	22	22.57	0.57	22.20	0.20	22.36	0.36
	19	21.9	22.64	0.74	22.21	0.31	22.39	0.49
	20	22.58	22.68	0.10	22.21	-0.37	22.40	-0.18
	21	22.14	22.69	0.55	22.21	0.07	22.41	0.27

由表 1 可知：组合模型预测值与实际观测值之间的误差在-1.13mm~1.87mm 之间，比单项模型预测的误差范围有所减小。图 1 是指数模型、Logistic 模型、加权组合模型地基沉降预测量—时间曲线与实测沉降量—时间曲线的比较。

由图 1 可知，该观测点的地基沉降已趋于稳定，组合模型预测沉降值介于指数模型、Logistic 模型预测值之间，能够反应地基沉降发展规律。经计算，组合模型预测该观测点的最终沉降为 22.41mm。

指数模型、Logistic 模型、加权组合模型检验标准值见表 2。由表 2 可知：组合模型的误差平方和、平均绝对误差、平均绝对百分比误差、均方误差均小于参加组合的单一预测模型，说明组合模型的预测精度均优于单项模型的预测精度，能够更好的预测地基沉降。

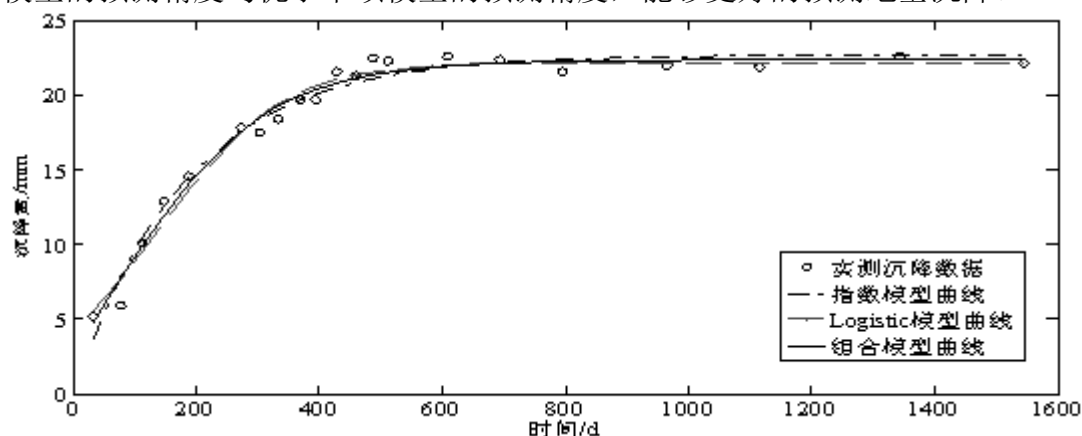


图 1 各模型预测曲线与实测曲线的比较

表 2 各模型评价标准值

模型名称	<i>SSE</i>	<i>MSE</i>	<i>MAE</i>	<i>MAPE</i>
指数模型	14.59	0.83	0.66	5.52
Logistic 模型	12.56	0.77	0.64	4.80
加权组合模型	11.22	0.73	0.61	4.55

为证明以上结论的正确性，分别对该发电厂的锅炉的现场监测数据进行分析。显然，锅炉和主厂房的加载情况不同，地面荷载不一样。图 2 是锅炉的三种模型的预测值与实测值的比较。由图 2 可知，加权组合模型的预测结果和实测值较接近，能够反应地基沉降实际情况。

图 2 锅炉各模型预测曲线与实测曲线的比较

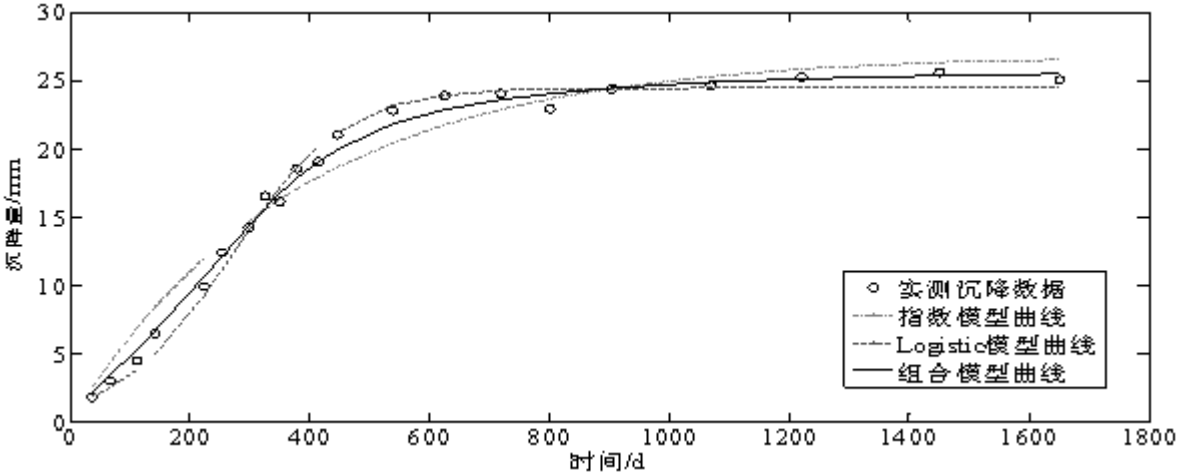


表 3 锅炉各模型评价标准值

模型名称	<i>SSE</i>	<i>MSE</i>	<i>MAE</i>	<i>MAPE</i>
指数模型	42.38	1.42	1.22	12.92
Logistic 模型	12.76	0.78	0.61	5.53
加权组合模型	8.02	0.62	0.51	4.90

表 3 是发电厂锅炉的不同模型的评价指标统计表。由表 3 可知，组合模型的预测精度最高，Logistic 模型的预测精度次之，指数模型最小。综上所述，对于地表的不同加载情况，虽组合模型个别点的观测误差大于单一预测模型，但从整体看，组合模型的预测精度均高于单项预测模型的预测精度。在实际工程中，整体的稳定性比单个点精度高更重要。因此，经改进熵值法计算权重的组合模型预测发电厂地基基础沉降是可行的。

5. 结语

- (1) 单一预测模型均有各自特点，不能全面考虑地基沉降因素，致使地基沉降预测值与实测值之间的误差具有随机性，而改进熵值法能够度量信息的不确定性，且计算简单，避免权重出现负值的情况。因此，用改进熵值法计算组合模型的权重系数具有实用意义。
- (2) 工程实例结果表明：基于指数模型、Logistic 模型建立的组合模型预测地基基础沉降，其误差平方和、均方误差、平均绝对误差、平均绝对百分比误差均小于各单项预测模型，说明运用改进熵值法确定的组合模型的优越性高，能够较好的预测发电厂的地基沉降。

参考文献

- [1] Tan SA. Validation of Hyperbolic Method for Settlement in Clays with Vertical Drain. *Soil and Foundation*. Vol. 35 (1995) No. 1, p. 101-113.
- [2] Wang W, Lu T. Study on prediction model of soft foundation settlement based on Weibull curve [J]. *Yantu Lixue*(Rock and Soil Mechanics), Vol. 28 (2007) No. 4, p. 803-806.
- [3] Weihong L, Junbao W. Study on application of MMF model to prediction of foundation settlement [J]. *Chinese Journal of Underground Space and Engineering*, (2011) No. 4.
- [4] 田晓丽, 杜兴无, 李天斌等. 非等时距GM(1,1)灰色模型在路基沉降预测中的应用[J]. *北方交通*, 2011,02:1-3.
- [5] 李政, 孟德光, 李冰心. 神经网络在软基沉降预测中的应用[J]. *西北农林科技大学学报(自然科学版)*, 2007,07:219-222.
- [6] 夏江, 严平, 庄一舟等. 基于遗传算法的软土地基沉降预测[J]. *岩土力学*, 2004,07:1131-1134.
- [7] 熊春宝, 王海涛. 一种组合模型在地面沉降预测中的应用研究[J]. *建筑技术*, 2013,03:201-204.
- [8] 苏金明, 张莲花, 刘波, 等. *MATLAB工具箱应用*[M]. 北京: 电子工业出版社, 2004: 489~512.
- [9] Xiong C, Li F. Exponential curve method model for predicting foundation settlement in the base pit vicinity [J]. *Geomatics and Spatial Information Technology*, (2011) No. 4, p. 004.
- [10] Wang Z, Zheng M, Wu Y, et al. Study on application of logistic curve model to settlement prediction of roadbed subjected to multilevel loading [J]. *Rock and Soil Mechanics*, (2004) No. 25, p. 901-903.
- [11] 陈启明, 陈华友. 改进的熵值法在确定组合预测权系数中的应用[J]. *统计与决策*, 2011,13:159-160.

References

- [1] Tan SA. Validation of Hyperbolic Method for Settlement in Clays with Vertical Drain. *Soil and Foundation*. Vol. 35 (1995) No. 1, p. 101-113.
- [2] Wang W, Lu T. Study on prediction model of soft foundation settlement based on Weibull curve [J]. *Yantu Lixue*(Rock and Soil Mechanics), Vol. 28 (2007) No. 4, p. 803-806.
- [3] Weihong L, Junbao W. Study on application of MMF model to prediction of foundation settlement [J]. *Chinese Journal of Underground Space and Engineering*, (2011) No. 4.
- [4] Tian Xiaoli, DU Xingwu, Li Tianfu, et al. Application Non-equal Interval GM (1,1) Grey Model in Settlement Prediction of Embankment [J]. *Northern Transportation*, (2008) No. 2, p. 1-3. (In Chinese)
- [5] LI Zheng, Meng Deguang, Li Bingxin. Application of Neural Network in settlement prediction of soft foundation [J]. *Northwest Agriculture and Forestry University (Natural Science)*, (2007) No. 7, p. 219-222. (In Chinese)
- [6] Xia Jiang, Yan Ping, Zhuang Yizhou et al. Prediction of settlement of soft soil foundation based on genetic algorithm [J]. *Rock and Soil Mechanics*, (2004) No. 7, p. 1131-1134. (In Chinese)
- [7] Xiong Chunbao, Wang Haitao. A combination model in application of ground subsidence prediction [J]. *Construction technology*, (2013) No. 3, p. 201-204. (In Chinese)

- [8] Su Jinming, Zhang Lianhua, Paul Liu, et al. Application of MATLAB toolbox [M]. Beijing: Publishing House of electronics industry, (2004), p. 201-204. (In Chinese)
- [9] Xiong C, Li F. Exponential curve method model for predicting foundation settlement in the base pit vicinity [J]. Geomatics and Spatial Information Technology, (2011) No. 4, p. 004.
- [10] Wang Z, Zheng M, Wu Y, et al. Study on application of logistic curve model to settlement prediction of roadbed subjected to multilevel loading [J]. Rock and Soil Mechanics, (2004) No. 25, p. 901-903.
- [11] Chen Qingming, Chen Huayou. Application of improved entropy method determine the weights of forecasting combination [J] Statistics and Decision, (2011) No. 13, p. 159-160.

作者简介: 高艳平, 女, 满族, 河北隆化人, 电话: 13731009657, 电子邮箱: 760761994@qq.com
通讯地址: 河北省邯郸市邯山区河北工程大学水电学院, 邮编: 056038