

DOI: 10.13718/j.cnki.xdzk.2020.01.018

山地电力架空线路运维质量的智能评估

易校石¹, 刘念²

1. 伊犁师范大学 数学与统计学院, 新疆 伊犁哈萨克自治州 835000; 2. 重庆大学 数学与统计学院, 重庆 401331

摘要: 山地电力架空线路大多沿山架设, 穿越树竹林, 周边的树竹生长缺陷引发短路故障频繁发生, 造成用电客户的大量投诉, 用户体验满意度下降, 抢修抢险费用大幅度攀升, 严重影响供电公司的信誉和售电收益. 为了提升用电客户的满意度和供电公司的售电受益, 采用主成分分析方法提取了山地电力架空线路通道运维质量的 3 个主要影响因素, 它们分别是树竹生长缺陷数、缺陷的消除数和通道运维资金使用合理度. 对树竹的生长规律建立一个预测模型, 从而得到树竹生长缺陷数的一个预测, 根据预测结果划拨架空线路所在供电所的运维资金, 消除树竹生长的缺陷数, 再用 Logistic 回归模型, 构建架空线路通道运维质量智能评估模型, 使供电公司对架空线路运维实现基于数据决策的模式, 尽可能保证供电线路零故障运行.

关键词: 架空线路; 预测模型; 数据决策; 智能评估

中图分类号: TM744

文献标志码: A

文章编号: 1673-9868(2020)01-0124-10

电网运行是否正常影响着国家的发展和人民的生活. 农网电力公司的供电辖区架空线路大多是沿山地架设, 穿越树竹林, 线路因“手拉手”率小, 绝缘化率较低, 抵御外力能力和转移负荷能力均较弱, 特别是在迎峰度夏期间, 出现风雷暴极端恶劣天气的频率较高且持续的时间较长; 高压输电线路通道内的树竹随着时间的增加而不断地生长, 相对于电网线路的树竹缺陷概率增加, 虽有人工蹲守、群众护线以及通道砍伐等应对措施, 但还是引发了大量线路跳闸和 95598 投诉. 现有统计数据显示: 大风雷暴等极端天气产生树竹刮蹭裸导线引发线路故障数占总故障次数的 67.39%, 说明现有的应对措施很难达到理想的效果, 同时还消耗了大量的人力和财力.

树竹是国家不可缺少的资源, 国家退耕还林和封山育林的政策鼓励大面积的树竹种植与再生, 使输电线路在运维通道方面增加了大量的人力、物力和财力. 要解决树障问题, 同时兼顾到国家利益总体达到最优, 可以对树竹生长高度建立一个预测模型, 预测到树竹的生长高度将会对线路通道产生影响时, 再消除树竹缺陷. 有关树竹的生长模型的研究报道较多, 时明芝等人^[1]在 2005 年对 107 杨树阳性冠幅与胸径关系进行了研究, 其线性相关系数达到 93.99%; 张杰伟^[2]在 2010 年对合肥市常见的三类植物(香樟、广玉兰、女贞)分别利用 Logistic 方程、Gompertz 方程、Richards 方程、Weibull 方程以及 S 曲线 5 种模型进行生长模型分析; 刘伟琪等人^[3]在 2013 年考虑影响树木生长规律因素, 通过对园林树木种植密度分析来构建树木的生长模型; 杨楠等人^[4]在 2017 年利用树木的生长模型结合监控设备, 对树木生长高度进行预测, 对输电线路可能造成的威胁进行预警, 从而保证输电线路的正常运行.

由于早期研究时期计算机技术较为落后, 国外学者对运维智能评估和评级的研究仅限于运用聚类分析、因子分析以及相关分析等统计学知识. 比如, 1936 年 Fisher^[5]、1941 年 Durand^[6]根据判别分析方法划

收稿日期: 2019-03-27

基金项目: 国家自然科学基金(11371384); 新疆维吾尔自治区青年博士人才培养项目(2017Q087).

作者简介: 易校石(1991-), 女, 硕士, 助教, 主要从事应用统计的研究.

分优次等级对信用等级进行了评价. 在判别分析进入了评级领域后, 更多的先进实用的分类评级技术竞相出现, 国外学者开始从大量数据中提取特征来分析行为并作出分类以及预测模型. 随着预测模型的出现和发展, 决策模型渐渐浮出水面, 在 1973 年, Mangasarian^[7] 提出递归分割法, 递归分割法是将总体样本进行分割再预测结果. 在国外学者研究的基础上, 国内也逐渐开始了对数据智能评估评级的研究, 近 10 年当中, 周晖等人^[8] 在 2007 年利用 Logistic 回归模型识别电力客户欠费概率大小, 以以往的客户资料作为数据, 根据 Logistic 回归模型预测欠费的可能性, 即输出结果为有欠费风险与无欠费风险, 从而减少客户欠费风险. 除 Logistic 模型对数据评级之外, 马立新等人^[9] 在 2018 年利用 K 近邻算法对能效指标评级, 其过程为筛选能效指标, 确定分级, 交叉检验选择 k 值, 最后利用模型评估电力用户等级. 本研究利用主成分分析得出影响运维质量的 3 个主要因素, 再 Logistic 回归模型, 根据树竹缺陷预测数、缺陷消缺率和资金的合理使用度 3 个因素评价其运维质量.

本研究的设想是筛选影响架空线路运维质量的树竹缺陷、缺陷消缺数和运维资金使用度 3 个主要因素, 然后根据树竹生长的规律建立一个树竹生长高度预测模型, 预测树竹的生长高度, 根据高度和树竹与架空线路的水平距离给定一个阈值, 预测结果超出这个阈值, 就认为是树竹的一个缺陷, 根据树竹预测的缺陷数, 树竹缺陷的消缺数和资金使用的合理度, 建立一个架空线路运维质量的智能评估系统, 指导各个供电所按需分配运维资金, 避免传统平均划拨运维资金, 树竹缺陷密集的供电所运维资金不足, 树竹缺陷少的供电所获得了大量的剩余运维资金; 节省供电所传统派出人员进行树竹缺陷例行巡查的高成本和抢修抢险的费用, 这种例行巡查可能导致了没有树竹缺陷也在巡查, 有树竹缺陷却没有巡查或没有巡查到, 产生不必要的差旅费用和高额的抢修抢险费用. 根据本研究提出的智能评估系统进行评估, 其结果可以指导各个供电所采用科学的方法消除数竹缺陷, 尽可能地降低线路故障的发生概率, 树立电网公司的优质服务形象.

1 影响架空线路运维的主要因素

影响架空线路运维质量的因素有: 树竹生长高度 x_1 、树障缺陷预测数 x_2 、树竹引发故障抢修数 x_3 、巡视人员数 x_4 、15 天内消除缺陷数 x_5 、30 天内消除缺陷数 x_6 、高峰期投入资金数 x_7 和低峰期投入资金数 x_8 等, 这些因素中有些是相关的, 有些是相互独立的, 为降低模型的识别难度, 需要从这些因素中筛选出相互独立的主要影响因素, 主成分分析能够用较少的变量去解释原来数据中的大部分差异, 通常按照累计贡献率大于 80% 的原则选取主成分个数, 再从主成分中选取权重较大的解释变量作为筛选的主要因素.

设架空线路的影响因素为 x_1, x_2, \dots, x_8 , 记为 $\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_8 \end{pmatrix}$, 将原始影响因素 x_1, x_2, \dots, x_8 综合成 8

个主成分 y_1, y_2, \dots, y_8 , 记为 $\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_8 \end{pmatrix}$, y_i 与 x_1, x_2, \dots, x_8 的关系为:

$$\begin{cases} y_1 = c_{11}x_1 + c_{12}x_2 + \dots + c_{18}x_8 \\ y_2 = c_{21}x_1 + c_{22}x_2 + \dots + c_{28}x_8 \\ \vdots \\ y_8 = c_{81}x_1 + c_{82}x_2 + \dots + c_{88}x_8 \end{cases}$$

$$\mathbf{C} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{18} \\ c_{21} & c_{22} & \cdots & c_{28} \\ \vdots & \vdots & \vdots & \vdots \\ c_{81} & c_{82} & \cdots & c_{88} \end{pmatrix}$$

则主成分写成矩阵形式:

$$\mathbf{Y} = \mathbf{CX}$$

其中 \mathbf{C} 为正交矩阵.

首先获取了 8 个变量 x_1, \dots, x_8 的 270 次观测数据的样本观测矩阵:

$$\mathbf{X}^{(0)} = \begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \\ \vdots \\ x_8^{(0)} \end{pmatrix} = \begin{pmatrix} x_{11}^{(0)} & x_{12}^{(0)} & \cdots & x_{1\ 270}^{(0)} \\ x_{21}^{(0)} & x_{22}^{(0)} & \cdots & x_{2\ 270}^{(0)} \\ \vdots & \vdots & \vdots & \vdots \\ x_{81}^{(0)} & x_{82}^{(0)} & \cdots & x_{8\ 270}^{(0)} \end{pmatrix}$$

样本观测矩阵中 $x_{ij}^{(0)}$ 表示第 i 个变量在第 j 次的观测数据. 为克服量纲的影响, 需要将原始数据作标准化处理:

$$x_{ij} = \frac{x_{ij}^{(0)} - \bar{x}_i}{s_i} \quad i = 1, 2, \dots, 8; j = 1, 2, \dots, 270$$

其中

$$\bar{x}_i = \frac{1}{n} \sum_{j=1}^n x_{ij}^{(0)}$$

$$s_i^2 = \frac{1}{n-1} \sum_{j=1}^n (x_{ij}^{(0)} - \bar{x}_i)^2 \quad i = 1, 2, \dots, 8; j = 1, 2, \dots, 270$$

样本观测矩阵的数据经过标准化处理以后, 得到数据矩阵 \mathbf{X} :

$$\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_8 \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1\ 270} \\ x_{21} & x_{22} & \cdots & x_{2\ 270} \\ \vdots & \vdots & \vdots & \vdots \\ x_{81} & x_{82} & \cdots & x_{8\ 270} \end{pmatrix}$$

主成分 y_1, y_2, \dots, y_8 相互正交, 270 个样本点变换后在主成分 y_1 上有最大的方差, 在主成分 y_2 轴上有次大的方差, \dots , 在主成分 y_8 轴上有最小的方差; 同时不同的 y_i 和 y_j 轴的协方差为 0, 即

$$\text{Var}(\mathbf{Y}) = \text{Var}(\mathbf{CX}) = (\mathbf{CX}) \cdot (\mathbf{CX})' = \mathbf{CXX}'\mathbf{C}' = \mathbf{A} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_8 \end{pmatrix}$$

由于 \mathbf{X} 是原始数据经标准化处理后的数据, \mathbf{XX}' 实际上是原始数据的相关系数矩阵 \mathbf{R} , 即

$$\mathbf{R} = \mathbf{XX}'$$

于是得到

$$\mathbf{CRC}' = \mathbf{A} \text{Var}(\mathbf{Y}) = \mathbf{CXX}'\mathbf{C}' = \mathbf{CRC}' = \mathbf{A} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_8 \end{pmatrix}$$

由此得到

$$\mathbf{RC}' = \mathbf{C}'\mathbf{A}$$

即:

$$R \begin{pmatrix} c_{i1} \\ c_{i2} \\ \vdots \\ c_{i8} \end{pmatrix} = \begin{pmatrix} \lambda_i c_{i1} \\ \lambda_i c_{i2} \\ \vdots \\ \lambda_i c_{i8} \end{pmatrix} = \lambda_i \begin{pmatrix} c_{i1} \\ c_{i2} \\ \vdots \\ c_{i8} \end{pmatrix}$$

C' 的第 i 列 $\begin{pmatrix} c_{i1} \\ c_{i2} \\ \vdots \\ c_{i8} \end{pmatrix}$ (或 C 的第 i 行向量的转置) 就是相关系数矩阵 R 的特征值 λ_i 对应的特征向量, 因此求

x_1, \dots, x_8 的主分量转化为求相关系数矩阵 R 的特征值和特征向量, y_1 方差为 λ_1 , y_2 的方差为 λ_2, \dots, y_8 的方差为 λ_8 , 并且满足 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_8$, 将采集样本数据的相关矩阵的特征值计算的结果列于表 1.

表 1 相关系数矩阵 R 的特征值, 贡献率和累积贡献率

主成分	特征值	贡献率/%	累计贡献率/%
主成分 1	8.909 6	74.247	74.25
主成分 2	1.688 1	14.067	88.31
主成分 3	0.655 2	5.460	93.77
主成分 4	0.490 2	4.085	97.86
主成分 5	0.154 4	1.286	99.15

由于相关矩阵 R 的特征值 λ_i 度量了第 i 个主成分 y_i 在 270 次观测中取值变化的大小, 一般按选取累计贡献率大于 90% 的原则选取主成分个数. 由于累积贡献率

$$P = \frac{\sum_{i=1}^3 \lambda_i}{\sum_{i=1}^8 \lambda_i} \geq 93.77\%$$

所以取前 3 个主成分, 而这 3 个主成分中树竹缺陷数、缺陷消缺数和资金合理使用度的权重最大, 因此影响架空线路的主要因素是树竹缺陷数、缺陷消缺数和资金合理使用度.

2 树竹缺陷数的预测模型

影响架空线路通道运维质量的主要因素是树竹缺陷数、缺陷消缺数和资金合理使用度 3 个因素, 最主要的因素是树竹缺陷数. 树竹的生长缺陷将会影响到架空线路的运维质量, 一方面, 架空线路正下方的树竹长到一定的高度会影响线路的正常运行; 另一方面, 当架空线路周边的树竹长到一定的高度时, 在大风雷暴等极端天气将产生树竹刮蹭裸导线而引发高架线路故障. 因此需要对架空线路的树竹高度和架空线路周边的树竹生长高度分别给出阈值, 树竹生长高度超过给定的阈值, 即可认为树竹生长产生缺陷. 为避免公司频繁地派出众多的人员进行巡视, 可为树竹生长高度建立一个预测模型, 预测树竹生长的缺陷数. 树竹的生长与很多因素有关, 不同品种树竹的生长规律有一定的差异, 即使是同一个品种的树竹, 在不同的土壤肥沃程度、二氧化碳浓度、光照温度、树竹的间距和行距等生长条件下, 其生长规律也有差异, 应该建立树竹的高度与土壤肥沃程度、二氧化碳浓度、光照温度、树竹的间距和行距的多元线性回归模型^[10], 但要采集到这些数据涉及到很多的部门, 部门之间的数据存在壁垒, 不能共享. 因此在缺失环境数据的情况下, 本研究根据树竹生长的高度与时间的关系规律采集样本数据, 根据采集到树竹生长规律的样本点分布来看, 树竹生长规律与 S 型生物生长曲线相吻合, 也与实际情况吻合. 正常情况下, 春季温度较低, 植被开始生长, 其生长速度较慢, 规模不大, 对架空线路的影响不大; 随着温度的升高, 植被生长达到旺盛期, 其

中以 6 月到 9 月的生长速度最快,达到一定的规模与高度,这时对架空线路的影响较大,在 9 月份因树竹刮蹭裸导线引发的故障报修数达到最大值;进入冬季后,植被生长基本停止,极端恶劣天气减少,树竹不容易刮蹭到电线,且在 10 月之前消除了大部分故障,砍去了对架空线路有较大影响的树竹障碍,因此 10 月之后的故障数呈下降趋势.以月份为横轴,以月度树竹生长高度和故障发生数为纵轴,得到树竹生长与缺陷引发线路故障的联合趋势分布图(图 1).

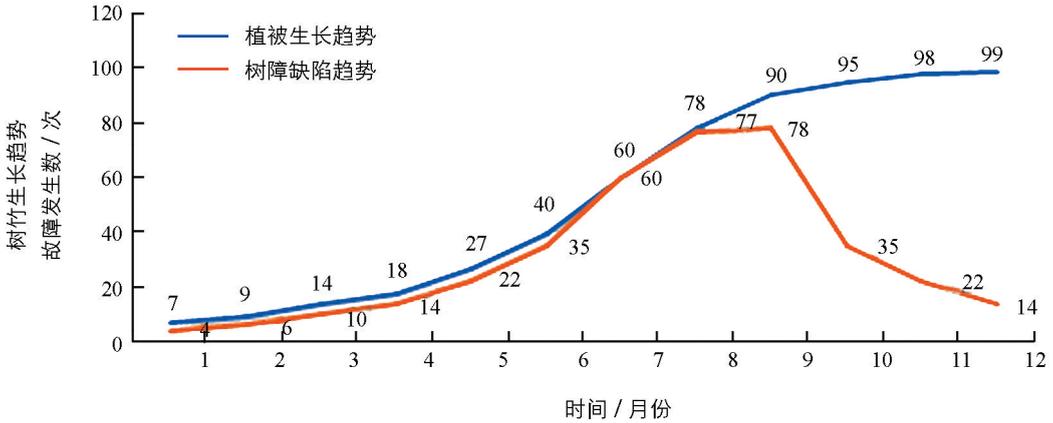


图 1 树竹生长和故障发生数的趋势联合分布图

一年的实际情况如图 2 所示,说明图 1 的一年故障趋势图与实际报修故障数有良好的一致性.

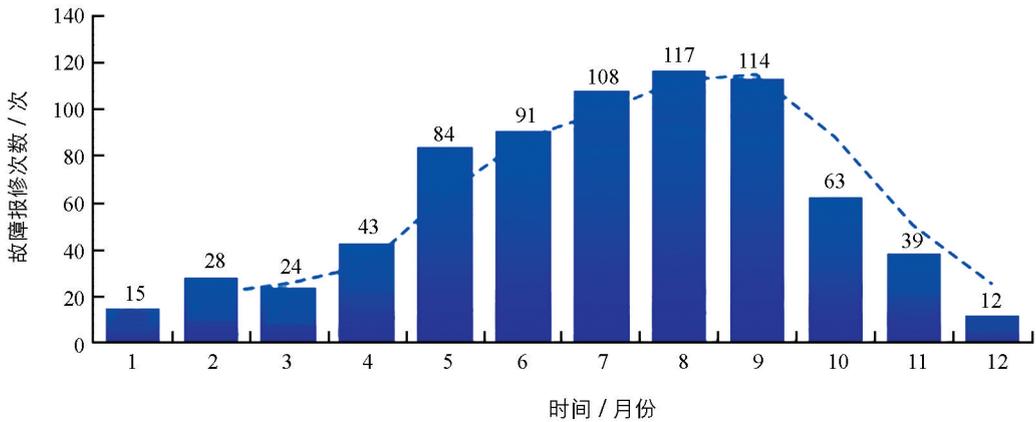


图 2 1 年的报修故障分布图

由图 1 可以看出:树竹生长高度呈现出 S 型曲线,可以用曲线

$$h_i = \frac{1}{\beta_0 + \beta_1 e^{-i}}$$

进行拟合,其中 h_i 表示第 i 个月份树竹的高度.

为了估计模型的参数,将原方程变形得到

$$\frac{1}{h_i} = \beta_0 + \beta_1 e^{-i} \quad (1)$$

作变量代换:

$$y_i = \frac{1}{h_i} \quad i^* = e^{-i}$$

(1)式变为:

$$y_i = \beta_0 + \beta_1 x_i^*$$

对采集的样本训练数据集 $T = \{(e^{-1}, y_1), (e^{-2}, y_2), \dots, (e^{-n}, y_n)\}$,采用最小二乘法得到模型参数的估计式:

$$\hat{\beta}_1 = \frac{l_{i^*y}}{l_{i^*i^*}} = \frac{\sum_{i=1}^n (i^* - \bar{i}^*)(y_i - \bar{y})}{\sum_{i=1}^n (i^* - \bar{i}^*)^2}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{i}^*$$

表 2 中的缺陷预测数是模型预测出的缺陷数, 缺陷消缺率是各个供电所消除缺陷数的比例, 故障抢修数是各个供电所抢修高架线路的故障数, 从表 2 可以看出, 在这一年中仍然有抢修故障存在, 这主要是由于各个供电所的树竹生长缺陷没有消缺完, 也可能是预测产生的误差. 因此还需要对各个供电所的消缺工作进行评估, 督促消缺工作做得差的供电所加强消缺工作, 提高架空线路的运维质量.

表 2 2017 年綦南电网公司树竹生长缺陷数与故障抢修数

供电所编号	供电所 1	供电所 2	供电所 3	供电所 4	供电所 5	供电所 6
缺陷预测数	110	172	181	105	47	114
巡视到位率	0.86	0.88	0.89	0.95	0.9	0.91
缺陷消缺率	0.89	0.92	0.91	1	0.88	0.85
故障抢修数	18	15	23	5	5	11
供电所编号	供电所 7	供电所 8	供电所 9	供电所 10	供电所 11	供电所 12
缺陷预测数	164	108	58	59	65	129
巡视到位率	0.92	0.8	0.84	0.88	0.87	0.9
缺陷消缺率	0.9	0.97	0.86	1	0.89	0.88
故障抢修数	14	27	11	8	10	14
供电所编号	供电所 13	供电所 14	供电所 15	供电所 16	供电所 17	
缺陷预测数	75	151	118	67	42	
巡视到位率	0.81	0.85	0.84	0.69	0.63	
缺陷消缺率	0.90	0.92	0.77	0.86	0.84	
故障抢修数	18	26	23	30	25	

3 运维资金的合理使用度

为降低架空线路在年度迎峰度夏高发期的故障发生次数, 如果在预测出每年 7, 8 月(这 2 个月植被生长速度达到最快)将产生树竹缺陷, 那么一般在每年 4 月~5 月就集中开展树竹消缺工作. 根据树竹缺陷的预测数和抢修故障的分布数, 运维资金在每个月理想使用度见表 3.

表 3 理想运维资金使用度

月份	1	2	3	4	5	6	7	8	9	10	11	12
资金比例/%	7	7	7	13	14	7	7	7	7	10	10	7

之所以选择运维资金理想使用度这个指标, 它对通道的运维质量有较大的影响, 若各个供电所的运维资金使用度与理想资金使用度相似程度越高, 则说明线路通道运维质量比较高; 反之, 相似程度越低, 说明线路通道运维质量比较差. 相似程度用相关系数来衡量:

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

其中:

$$\bar{i} = \frac{\sum_{i=1}^n i}{n}$$

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$$

对采集的綦南电网公司的 17 个供电所的运维资金使用的月度数据, 计算其与理想程度的相似度(表 4).

表 4 供电所相似度

供电所编号	供电所 1	供电所 2	供电所 3	供电所 4	供电所 5	供电所 6
相似度	0.441 7	0.991 7	0.951 7	0.951 5	0.965 8	0.971 9
供电所编号	供电所 7	供电所 8	供电所 9	供电所 10	供电所 11	供电所 12
相似度	0.968 8	0.963 7	0.895 3	0.952 9	0.962 4	0.956 8
供电所编号	供电所 13	供电所 14	供电所 15	供电所 16	供电所 17	
相似度	0.867 7	0.942 8	0.843 2	0.795 7	0.346 1	

4 架空线路运维质量的智能评估

运检 PMS2.0 系统内树竹缺陷预测数据、巡视到位的数据及消除缺陷数据, 再采集营销 SG186 系统里架空线路故障报修数据、财务管控系统内的通道运维专项资金数据, 可对各个供电所的架空线路运维质量进行评估.

影响架空线路运维质量的有树竹缺陷预测数、缺陷消缺率和资金的合理使用度 3 个因素, 预测准确度、消缺率和资金使用合理度较高, 架空线路报修的概率就越小, 若 3 个指标与植被生长趋势、故障报修分布趋势、资金合理使用度趋势相同, 架空线路的运维质量就比较好, 这可以用逻辑回归的 Sigmoid 函数进行刻画^[11]. Sigmoid 函数为

$$f(z) = \frac{1}{1 + e^{-z}}$$

其图像见图 3.

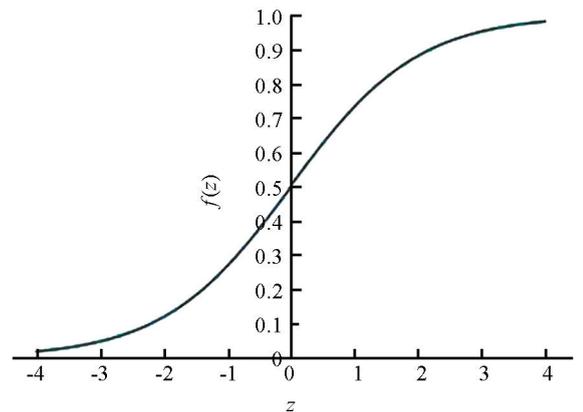


图 3 Sigmoid 函数图像

Sigmoid 函数的函数值在区间(0, 1)内, 图像也呈 S 曲线型, 与树竹的生长规律相当吻合, 用它来评价线路通道的运维质量比较客观^[12]. 首先将树竹生长缺陷预测数 x_1 、消缺率 x_2 、通道运维资金使用合理度 x_3 综合成一个综合指标 z , x_i 的权重记为 ω_i , $\omega^T = (\omega_1, \omega_2, \omega_3)$, $X^T = (x_1, x_2, x_3)$, 得到综合指标的表达式为

$$z = \omega^T X + b = b + \sum_{i=1}^3 \omega_i x_i$$

把采集的綦南电网公司的数据, 以分成两类为例说明评级的方法. 以每年报修 10 次以下作为好, 报修 10 次以上作为差. 权重 ω_i 的估计采用极大似然估计, 记:

$$\varphi(z) = f_{\omega}(z) = \frac{1}{1 + e^{-z}}$$

$$P(Y=1 | X, \omega) = \varphi(\omega^T x + b) = \varphi(z)$$

$$P(Y=0 | X, \omega) = 1 - \varphi(z)$$

$$p(y | \omega, x) = (\varphi(z))^y (1 - \varphi(z))^{1-y}$$

极大似然函数为

$$L(\omega) = P(Y | X, \omega) = \prod_{i=1}^n p(y^{(i)} | x^{(i)}, \omega) = \prod_{i=1}^n (\varphi(z^{(i)}))^{y^{(i)}} (1 - \varphi(z^{(i)}))^{1-y^{(i)}}$$

两边取对数, 得

$$l(\omega) = \log L(\omega) = \sum_{i=1}^n (y^{(i)} \ln \varphi(z^{(i)}) + (1 - y^{(i)}) \ln(1 - \varphi(z^{(i)})))$$

由于 $l(\omega)$ 是负值, 所以引入代价函数

$$J(\omega) = -l(\omega) = -\sum_{i=1}^n (y^{(i)} \ln \varphi(z^{(i)}) + (1 - y^{(i)}) \ln(1 - \varphi(z^{(i)})))$$

利用梯度下降法, 求使得代价函数 $J(\omega)$ 最小的参数 ω_i ,

$$\begin{aligned} \frac{\partial J(\omega)}{\partial \omega_j} &= -\sum_{i=1}^n (y^{(i)} \frac{1}{\varphi(z^{(i)})} - (1 - y^{(i)}) \frac{1}{1 - \varphi(z^{(i)})}) \varphi(z^{(i)}) (1 - \varphi(z^{(i)})) \frac{\partial z^{(i)}}{\partial \omega_j} = \\ &\sum_{i=1}^n (y^{(i)} (1 - \varphi(z^{(i)})) - (1 - y^{(i)}) \varphi(z^{(i)})) x_j^{(i)} = \\ &\sum_{i=1}^n (y^{(i)} - \varphi(z^{(i)})) x_j^{(i)} \end{aligned}$$

所以得到权重更新的公式为:

$$\omega_j = \omega_j + \eta \sum_{i=1}^n (y^{(i)} - \varphi(z^{(i)})) x_j^{(i)}$$

把綦南电网公司 2017 年的 17 个供电所随机抽取了 13 个供电所(总量的 80%)的数据作为训练集, 利用梯度下降法迭代 500 次, 得到树竹缺陷巡视到位率、消缺率、通道运维资金合理使用度对应的权重系数 $(\omega_1, \omega_2, \omega_3) = (21.491, 1.141, 2.448)$, 由此得到:

$$z = -22.532 + 21.491x_1 + 1.141x_2 + 2.448x_3$$

其 Wald 统计量均通过检验 ($\alpha < 0.05$), 表明该模型有统计学意义。

由此得到对各个供电所的架空线路的运维质量评价类别为:

$$\operatorname{argmax}_p(p(Y=1), p(Y=0)) = \operatorname{argmax}_p(\varphi(z), 1 - \varphi(z)) \quad (2)$$

(2) 式表示把 z 值代入 $p(Y=1)$ 和 $p(Y=0)$ 的表达式, 计算这两个概率值, 若 $p(Y=1)$ 值大, 则判别该供电所对架空线路归属于好, 若 $p(Y=0)$ 的值大, 则判别该供电所对架空线路归属于差^[13]。

为检验本研究给出的分类模型的正确性, 把采集的綦南电网公司剩下的供电所 1, 供电所 7, 供电所 14 和供电所 17 的数据作为测试集, 预测结果和实际评级结果见表 5。

表 5 4 个供电所测试集的真实评级与预测评级

供电所 编号	巡视 到位率	缺陷 消缺率	资金使用 相似度	95598 故障报修数	预测 评级	真实 评级
供电所 1	0.86	0.89	0.441 7	15	0	0
供电所 7	0.92	0.90	0.968 8	8	1	1
供电所 14	0.85	0.92	0.942 8	15	0	0
供电所 17	0.63	0.84	0.346 1	31	0	0

由表 5 可以看出: 只有供电所 7 的架空线路的质量评为好, 其他 3 个供电所的运维质量评为差。

供电所 7 的巡视到位率和资金的使用度与理想程度的相似度较高, 导致故障报修次数较少, 线路正常运转的质量概率较高, 故障报修次数只有 8 次, 这主要是由于巡视到位率和消缺率也没有达到 100% 所引起的, 但相对于其他供电所运维质量还是比较好的;

供电所 1 和供电所 17 的巡视到位率都非常低, 巡视到的缺陷数与预测的缺陷数有较大的差异, 可以理解为巡视人员巡视工作不细致, 敷衍了事, 没有消缺的缺陷数导致线路故障的发生; 消缺资金的使用与理想的消缺资金使用相似程度太低, 运维资金的使用存在较大的漏洞, 缺陷没有及时消除, 故障报修次数较多, 产生高额的抢修抢险费用;

供电所 14 虽然资金使用相似度较高, 主要是巡视到位率较低, 没有巡视到的树竹缺陷导致架空线路故障的发生。

从回归模型 $z = -22.532 + 21.491x_1 + 1.141x_2 + 2.448x_3$ 的权重可以看出: 巡视到位率的权重 (21.491) 最大, 说明巡视到位率对架空线路通道运维质量评估结果的影响程度最大, 其次是消缺资金使用的合理度的权重 (2.448), 最后是消缺率的权重 (1.141), 因为消缺率是由供电公司统一布置, 巡视到位的树竹缺陷基本上消除了, 只是消除缺陷是否及时的问题。

5 结束语

本研究运用科学的统计分析方法筛选出影响架空线路的主要影响因素, 为树竹的生长建立了一个恰当的预测模型, 从而得到树竹生长的树竹缺陷数^[14]; 利用 Sigmoid 函数为架空线路的运维质量构建了一个评价模型, 对各个供电所的架空线路运维质量进行评价, 指导各个供电所提高运维质量; 将根据树竹生长模型预测出的树竹生长缺陷数作为巡查树竹缺陷数, 然后根据树竹缺陷数, 合理地划拨运维消缺资金, 再根据树竹生长规律, 集中消缺树竹缺陷数, 尽可能保持高架线路的畅通, 减少因线路的运维质量较低所引发的故障数。为了验证本研究模型及算法的有效性, 将本研究建立的模型于 2018 年试用于重庆市綦南电网公司, 公司故障抢修工单比 2017 年下降 18.79%, 95598 停电投诉比 2017 年下降 27.63%, 抢修抢险费用比 2017 年下降 24.76%。本研究提出的智能评估模型为电力公司的架空线路管理提供了基于数据的决策模式, 模型兼具安全性、创新性、经济性以及可推广性, 实现了“用数据管理企业、用信息驱动业务”的目标。

参考文献:

- [1] 时明芝, 张凡东. 107 杨树阳性冠幅与胸径关系的研究 [J]. 福建林业科技, 2005, 32(4): 97-98.
- [2] 张杰伟. 合肥市三种常绿园林树木生长模型的研究 [D]. 合肥: 安徽农业大学, 2010.
- [3] 刘玮琪, 任 慧, 李茵茵, 等. 对园林树木种植疏密度的分析——以北京地区代表性乔木为例 [J]. 中国园艺文摘, 2013, 29(7): 10-13.
- [4] 杨 楠, 刘晓波. 基于树木生长模型的输电线路通道灾害预警方法的研究 [J]. 新型工业化, 2017, 7(12): 11-15.
- [5] FISHER R A. The Use of Multiple Measurements in Taxonomic Problems [J]. Annals of Eugenics, 1936, 7(2): 179-188.
- [6] DURAND D. Risk Elements in Consumer Installment Financing [M]. New York: National Bureau of Economic Research, 1941.
- [7] MANGASARIAN O L. Linear and Nonlinear Separation of Patterns by Linear Programming [J]. Operations Research, 1965, 13(3): 444-452.
- [8] 周 晖, 王 毅, 王 玮, 等. 基于 Logistic 回归模型的电力客户欠费违约概率的预测 [J]. 电网技术, 2007(17): 85-88.
- [9] 马立新, 徐曼如, 吕梦圆. 能耗系统的电力能效测评与分级方法研究 [J]. 电子测量技术, 2018, 41(13): 10-13.
- [10] 吴博文. 泰国洛坤府交通主干道行道树多样性研究及生态效益评价 [D]. 重庆: 西南大学, 2016.
- [11] 王 贤, 张自力. 基于逻辑回归的案件关联 [J]. 西南师范大学学报(自然科学版), 2009, 34(3): 83-87.

- [12] 赵会茹, 霍慧娟, 李春杰. 基于因子分析和 Logistic 回归模型的电力普遍服务受助个体主客观诉求影响因素分析 [J]. 陕西电力, 2014, 42(8): 65-71.
- [13] 张 静. 基于有序 Logistic 模型的发电企业债券信用评级研究 [D]. 天津: 天津财经大学, 2017.
- [14] LUMBRERAS A, JOUVE B, VELCIN J, et al. Role Detection in Online Forums Based on Growth Models for Trees [J]. Social Network Analysis and Mining, 2017, 7: 49-1-49-13.

Intelligent Evaluation of Operation and Maintenance Quality of Mountain Power Overhead Line

YI Xiao-shi¹, LIU Nian²

1. School of Mathematics and Statistics, Yili Normal University, Ili Kazak Autonomous Prefecture Xinjiang 835000, China;

2. School of Mathematics and Statistics, Chongqing University, Chongqing 401331, China

Abstract: Most of the overhead power lines of mountain power systems are erected along the mountain ranges, crossing the tree and bamboo forests, and the surrounding tree and bamboo growth defects cause frequent short-circuit faults, resulting in a large number of complaints from customers, reduced user experience satisfaction and a significant increase in the cost of rush repair and rescue operations and seriously affecting power supply company's reputation and sales revenue. In order to improve the customer satisfaction and the power supply benefit of the power supply company, the paper uses the principal component analysis to extract three main influencing factors of the operation and maintenance quality of the mountain power overhead lines, namely, the number of defects in the growth of trees/bamboo, the number of defects eliminated and the reasonable use of the funds for channel operation and maintenance. A prediction model is established for the growth law of the tree/bamboo, and a prediction of the number of defects in their growth is obtained. The operation and maintenance funds of the power supply station where the overhead line is located are allotted according to the prediction results, Then the logistic regression model is used to construct an intelligent evaluation model for the operation and maintenance quality of the overhead line channel, so that the power supply company can realize the data decision-making mode for the overhead line operation and maintenance, and ensure the zero-fault operation of the power supply line as much as possible.

Key words: overhead line; predictive model; data decision; intelligent assessment

责任编辑 潘春燕