

基于变权组合模型的光缆传输损耗预测

张明江, 李红卫, 夏贵进

(国防科技大学 信息通信学院试验训练基地, 西安 710106)

摘要:为使光缆维护管理部门对未来光缆线路传输损耗进行科学预测,以灰色 GM(1,1)模型、三次指数平滑模型和支持向量回归(SVR)模型为基础,构建权重随时间变化的组合预测模型,并进行了实例研究及 Matlab 仿真。仿真结果表明,SVR 模型在三个单项模型中预测效果最优,且变权组合模型无论在拟合期还是检验期,其预测效果都优于其它单项模型。

关键词:光缆传输损耗预测;变权组合模型;支持向量回归;GM(1,1);三次指数平滑

中图分类号:929.11 **文献标识码:**A **文章编号:**1002-5561(2018)05-0045-05

DOI:10.13921/j.cnki.issn1002-5561.2018.05.012

Transmission loss prediction of optical cable based on variable weight combination model

ZHANG Mingjiang, LI Hongwei, XIA Guijin

(Experimental training base of institute of information and communication of national university of defense technology, Xi'an 710106, China)

Abstract: In order to make the department of optical cable maintenance and management make scientific predictions on the transmission loss of the future optical cable, based on the grey GM (1,1) model, the three-exponential smoothing model and the support vector regression (SVR) model, the paper builds a combined forecasting model for the weights change over time, and conducted an example study and Matlab simulation. The simulation results show that the SVR model has the best forecast effect among the three individual models, and the variable weight combination model is superior to other single-term models in both the fitting period and the testing period.

Key words: transmission loss prediction of optical cable; variable weight combination model; support vector regression; GM (1,1); three exponential smoothing

0 引言

光缆网因其具有高速率、大容量和高可靠性等特点,是现代通信网的重要组成部分和主要发展方向,已成为国家重要的战略基础设施。经调查统计,目前我国在用的光缆最长服役时间已接近 30 年,部分早期建设的光缆因线路老化问题,导致传输损耗逐渐变大,造成传输性能不断下降。因此,采取科学方法对光缆线路传输损耗进行预测分析,不仅可以预报排查线路隐患,实现传输故障预警,有效降低光缆阻断的发生率,缩短光缆的故障历时,也可以帮助上级管理部门

全面掌握光缆网的运行状态信息,辅助实现光缆网运维保障力量的全局优化部署,对提高光缆网的通信保障效能至关重要。本文构建基于变权组合模型的光缆传输损耗预测方法,以对光缆线路的传输损耗进行科学预测。

1 光缆线路的传输损耗预测方法分析

损耗是表征光缆线路光纤传输性能的一个重要指标,在排除人为破坏、地质灾害等强外力因素情况下,对于某条具体光缆线路而言,其历年所处外界环境相对比较稳定,线路传输性能的波动劣化是一个缓慢变化过程,主要是由于光缆先天制作工艺、光纤上各接头和熔接点长时间受线路局部应力或外界水分

收稿日期:2018-03-04。

作者简介:张明江(1988-),男,硕士研究生,助教,主要研究方向为通信与指控装备指挥管理。

侵蚀等各种因素累积影响而出现老化现象,在光纤表层或熔接面产生微裂纹所造成的。所以,光缆线路运行时长就成为分析传输损耗变化的主要考虑因素,在实际工作中,光缆维护部门通常定期对所属线路进行监测,全面掌握线路运行状态信息。因此,对某段光缆线路传输损耗进行预测时,考虑到回归预测需掌握各影响因素大量的历史数据,易出现多重共线性和伪回归问题,且信息采集困难,预测难度较大,通常是直接利用监测到的历史损耗数据,通过时间序列或组合模型等定量分析方法建立预测模型,推算出光缆线路传输损耗随时间的变化趋势。

在对光缆传输损耗的预测研究上,常用的预测方法有指数平滑法^[1]、固定权重组合模型^[2]预测法等。本文在前人研究的基础上,提出将 SVR 模型^[3]用于光缆传输损耗预测研究,构建基于指数平滑、灰色 GM(1,1)和 SVR 预测方法的权重随时间变化的组合预测模型,并进行了实例仿真验证。

2 构建光缆传输损耗变权组合预测模型

2.1 单项模型的选择

2.1.1 灰色 GM(1,1)模型

灰色理论 GM(1,1)模型适合于研究具有“小样本”、“贫信息”的不确定信息,通过对少量已知信息的生成、开发,提取有价值的信息,实现对系统演化规律的正确描述和有效监测^[4]。其预测思路为:通过对原始数列进行一次累加生成运算后,对新生成的数列构建白化形式微分方程,通过最小二乘法求解方程参数,得到方程的时间响应函数 $\hat{x}^{(1)}(k+1)=(x^{(0)}(1)-u/a)e^{-ak}+u/a, k=1,2,\dots,n$,然后通过累减逆还原即可生成原始数列的预测序列,其中 $x^{(0)}$ 为原始数列, $\hat{a}=\begin{bmatrix} a \\ u \end{bmatrix}$ 为白化微分方程的参数, a 为发展系数, u 为灰色作用量。

2.1.2 三次指数平滑模型

指数平滑法是移动平均法的改进和发展,既不需存储很多的历史数据,又考虑了各时期数据的重要性,指数平滑模型可分为一次、二次和三次指数平滑模型^[5]。由于光缆的损耗变化是一个非线性发展过程,因此光缆的损耗预测通常选用更适合非线性时间序列预测的三次指数平滑模型。其预测模型为: $\hat{X}_{t+T}=\alpha_t+b_tT+c_tT^2$, $\hat{X}_{t+T}$ 为 $t+T$ 期的预测值, T 为预测的超前期数,其

中 $\alpha_t=3S_t^{(1)}-3S_t^{(2)}+S_t^{(3)}, b_t=\frac{\alpha}{2(1-\alpha)^2}[(6-5\alpha)S_t^{(1)}-2(5-4\alpha)$

$S_t^{(2)}+(4-3\alpha)S_t^{(3)}], c_t=\frac{\alpha^2}{2(1-\alpha)^2}[S_t^{(1)}-2S_t^{(2)}+S_t^{(3)}], \alpha (0<\alpha<1)$

为平滑系数, $S_t^{(1)}, S_t^{(2)}, S_t^{(3)}$ 分别为时间序列的一次、二次和三次平滑值。

2.1.3 SVR 模型

SVR 模型是一种人工智能机器学习预测方法,和神经网络类似,都是通过对数据样本进行学习和训练构建预测模型,都适用于复杂的非线性预测问题。但是 SVR 模型对数据的样本数量要求远没有神经网络高,更适合于解决小样本、非线性预测问题^[6]。而光缆传输损耗预测通常属于小样本预测问题,故可以选取 SVR 模型对其进行预测。SVR 预测模型主要是基于训练样本集寻找函数 $f(x)=w \times \phi(x)+b$, 使得 $|f(x_i)-y_i| \leq \varepsilon, (i=1,2,\dots,l)$, 其中 $(x_i, y_i) \in X=\{(x_i, y_i) | x_i \in R^N, y_i \in R\}$ 为训练样本点, w 为最优超平面法向量, ε 为不敏感损失系数 ($\varepsilon>0$), b 为阈值。通过引入惩罚参数 C 、松弛变量 ξ_i 和 ξ_i^* , 以及核函数映射方法,将原样本属性空间的非线性回归问题可转化为高维空间的线性凸二次规划问题^[7]。进一步引入拉格朗日乘子和核函数,又可将规划问题转化为其对偶问题求解。因高斯径向基核函数的非线性预测能力较强,本文选取它为核函数,表达式为 $K(x_i, x_j)=\exp(-\|x_i-x_j\|^2/\delta^2)$ 。采用径向基核函数的 SVR 模型的参数有 3 个,即惩罚参数 C 、损失系数 ε 和核函数系数 γ , 由于参数的选取会影响到预测的精度,本文利用遗传算法的全局优化搜索能力对参数进行寻优。

2.2 变权组合模型的建立

20 世纪 70 年代, Bates 与 Granger 第一次提出了组合预测模型,与单项预测模型相比,组合预测模型可以弥补单一预测模型的不足,提高预测的精度和可靠性^[8]。组合方法可分为线性和非线性组合,线性组合又分为固定权重和变化权重。由于预测对象所处的环境通常处于变化之中,固定权重组合模型往往不适应这些变化,导致预测精度大为降低,而变权组合模型可以很好地适应环境的变化,故本文在以上三种单项预测模型的基础上,采取变权组合模型对光缆线路的传输损耗进行预测。模型构建过程为:

第一步:分别利用 GM(1,1)模型、三次指数平滑模型和 SVR 模型进行预测,得到对应的预测值 $F_1(t), F_2(t)$ 和 $F_3(t), t=1,2,\dots,n+j$, 其中 n 为建模拟合期数, j 为模型检验期数。由于 SVR 模型需要选定一个最小嵌入

维数 m , 所以组合模型的拟合期通常从第 $m+1$ 期起算。

第二步:求解变权组合模型中单项模型 GM(1,1)、三次指数平滑和 SVR 在第 t 期的权重系数 $w_1(t)$ 、 $w_2(t)$ 和 $w_3(t)$, 求解权重时, 在建模拟合期和模型检验期的方法有所不同, 在拟合期的求解公式为:

$$\begin{cases} \min \left(\sum_{i=1}^3 w_i(t) F_i(t) - y(t) \right)^2 \\ \text{s.t. } \sum_{i=1}^3 w_i(t) = 1, w_i(t) \geq 0, (i=1,2,3) \end{cases} \quad (1)$$

在模型检验期的权重求解公式为:

$$\begin{aligned} w_i(n+1) &= \frac{1}{n} \sum_{t=1}^n w_i(t), w_i(n+2) = \frac{1}{n} \sum_{t=2}^{n+1} w_i(t), \dots, \\ w_i(n+j) &= \frac{1}{n} \sum_{t=j}^{n+j-1} w_i(t) \end{aligned} \quad (2)$$

第三步:求解变权组合模型在建模拟合期和模型检验期的预测值, 公式为:

$$F(t) = \sum_{i=1}^3 w_i(t) F_i(t) \quad (3)$$

其中, $F(t)$ 为变权组合模型在第 t 期的预测值。

第四步:对变权组合预测模型分别在建模拟合期和模型检验期的预测结果进行预测精度分析, 若预测效果较好, 则可以用于预测。

3 实例预测分析

3.1 数据来源

通过调研, 我们得到某单位维护的一段中继段光缆(约 80km)的传输损耗历史监测数据, 每季度对光缆传输损耗监测一次, 统计时间跨度为 2004~2017 年, 部分数据如表 1 所示。前 12 年的数据用于组合模型的建立, 后 2 年的数据用于模型的检验。

表 1 2004~2017 年某中继段损耗测试数值(dB)

年份	1 季度	2 季度	3 季度	4 季度
2004 年	16.20	16.22	16.25	16.29
2005 年	16.26	16.24	16.27	16.32
2006 年	16.32	16.31	16.37	16.39
...	...	...	...	...
2016 年	16.94	16.96	16.99	17.00
2017 年	17.01	17.02	17.05	17.07

3.2 单项模型预测

3.2.1 灰色 GM(1,1)模型预测

我们利用 Matlab 软件编程计算, 将前 12 年的光

缆损耗数据看做初始数列, 通过一次累加, 求解白化微分方程, 得到模型的时间响应函数为: $\hat{x}^{(1)}(t+1) = 17908e^{0.0009t} - 17892$, 取 t 值为 48~55, 将上式作累减还原即可得光缆传输损耗后 2 年的预测值。将观测值和 GM(1,1)在建模拟合期和检验期的预测值绘制比较曲线图, 如图 1 所示。

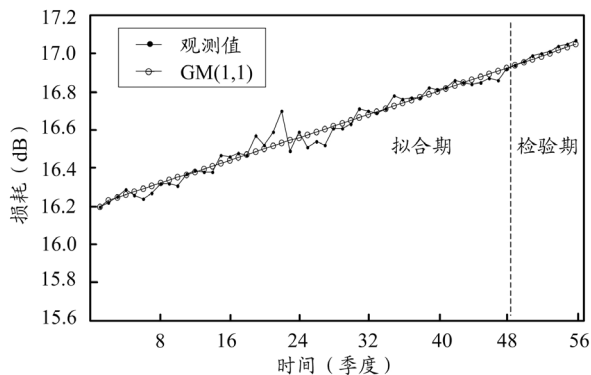


图 1 GM(1,1)模型预测曲线图

3.2.2 三次指数平滑模型预测

在运用三次指数平滑法时, 平滑初始值和平滑系数的选取对预测结果影响较大, 通常当原始数据样本大于 15 项时, 取原始序列第一期的数据作为平滑初始值, 本文取 $x(1)=16.2$ 为平滑初值。平滑系数 α 的选取通常采用经验值法或试算法确定, 但经验值法主观性太强, 试算法计算过程复杂。为提高预测效果, 简化计算过程, 本文采取网格寻优法, 通过 Matlab 编程, 以预测值在检验期平均绝对误差最小为准则, 以 0.001 为最小步长进行循环计算, 最终选取平滑系数 $\alpha=0.082$ 。通过计算, 以第 48 个季度的损耗值为起点, 可得预测公式为 $\hat{x}_{48+T} = 16.9298 + 0.0171T + 0.000075T^2$, 当 T 取值为 1~8 时, 即可求得后 2 年 8 个季度的损耗预测值, 将观测值和三次指数平滑模型的预测值绘制比较曲线图, 如图 2 所示。

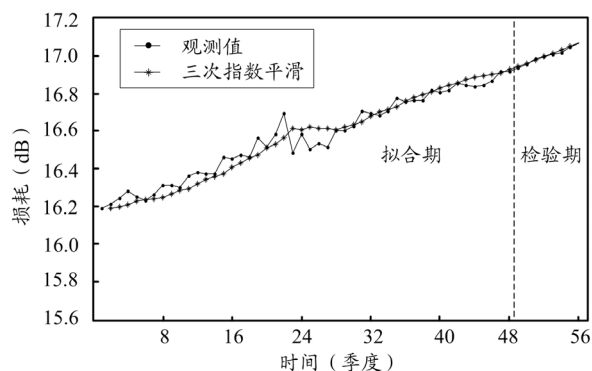


图 2 三次指数平滑模型预测曲线图

3.2.3 SVR 模型预测

本文将前 12 年的损耗时间序列进行空间重构,经过反复计算验证,确定了 SVR 的最小嵌入维 m 取 3,延时系数取 1,表示用前 3 个季度的数据回归预测后一个季度的数据,以此类推,可以得出本模型共建立了 45 个训练样本和 8 个检验样本。通过 Matlab 中的 Libsvm 工具箱编程计算,选取高斯径向基核函数,并采用遗传算法进行 Svmtrain 函数中的最优参数搜索,取最大遗传代数为 100,种群数为 20,交叉验证折数为 5,个体长度为 20。设定核函数系数、惩罚系数和损失系数的取值范围分别为 $[0,100]$ 、 $[0,1000]$ 和 $[0.01,1]$,可求得核函数系数 $\gamma=2.8$,惩罚参数 $C=160$,损失系数 $\varepsilon=0.0103$ 。通过工具箱内的 Svmpredict 函数计算可得各期预测结果,将观测值和 SVR 模型的预测值绘制比较曲线图,如图 3 所示。

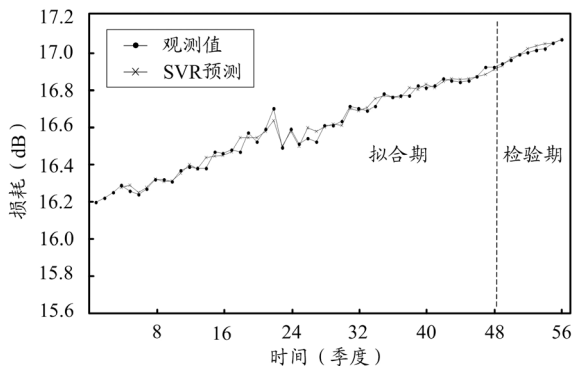


图 3 SVR 模型预测曲线图

3.3 变权组合模型预测及结果分析

在求出各单项模型预测值后,利用 Matlab 编程分别求解式(1)、式(2),即可求得从第 4 个季度开始建模拟合期和模型检验期变权组合模型中各单项模型的权重,如表 2 所示。

利用式(3)可求得变权组合模型在各季度的光缆损耗预测值,将观测值和变权组合模型在建模拟合期和检验期的预测值绘制散点图,如图 4 所示。

表 2 权重表

模型	建模拟合期(季度)					模型检验期(季度)			
	4	5	6	...	48	49	50	...	56
GM(1,1)	0	0.0150	0	...	0	0.2382	0.2435	...	0.2427
三次指数平滑	0	0.5420	1	...	0.4320	0.2319	0.2371	...	0.2177
SVR	1	0.4430	0	...	0.5680	0.5299	0.5194	...	0.5396

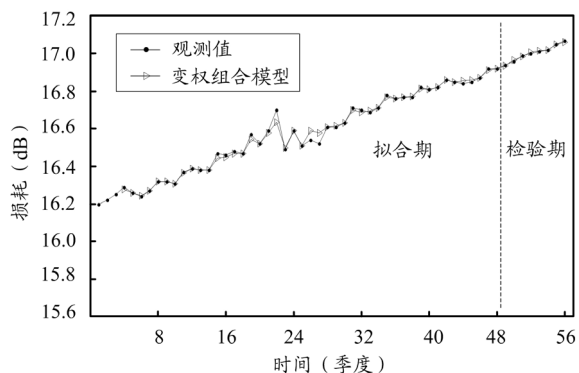


图 4 变权组合模型预测曲线图

通过图 1~图 4 可以较为清晰、直观地看出,在所有的曲线比较图中,变权组合预测模型的预测曲线与观测值曲线最接近,相似度最高,说明变权组合模型的预测精度高于三个单项模型。而在各单项预测模型中,SVR 模型的预测曲线的整体拟合效果要优于 GM(1,1)模型和三次指数平滑模型,这是因为 SVR 这种机器学习算法有更强的非线性预测能力。

为进一步说明变权组合预测模型的优越性,本文引入了两个误差度量指标:平均绝对误差(MAD)和均方根误差(RMSE),分别计算各预测模型在建模拟合期和模型检验期的误差值,以此来定量表征各模型的预测精度,计算结果如表 3 所示。

综合分析表 3,从建模拟合期和模型检验期的 MAD 和 RMSE 误差指标数据可以看出,在拟合期内,本文构建的变权组合模型预测误差最小,SVR 模型预测误差次之;在检验期内,同样是变权组合模型预测误差最小,而三次指数平滑模型预测误差次之。综合考虑拟合期和检验期的预测误差,可以表明变权组合模型

表 3 各模型预测精度分析

评价指标	GM(1,1)		三次指数平滑		SVR		变权组合模型	
	拟合期	检验期	拟合期	检验期	拟合期	检验期	拟合期	检验期
MAD	0.0301	0.0105	0.0400	0.0057	0.0193	0.0090	0.0083	0.0044
RMSE	0.0427	0.0129	0.0530	0.0072	0.0264	0.0103	0.0170	0.0051

在本案例光缆线路的传输损耗预测上精度最高预测效果最好,SVR模型在各单项模型中的预测精度最好,所以该变权组合模型可应用到光缆线路传输损耗的预测工作当中。

4 结束语

本文在分析光缆线路传输损耗特点规律的基础上,以具体事例为背景,通过仿真对比GM(1,1)、三次指数平滑和SVR模型对光缆传输损耗的预测效果,最终构建了预测精度更优的变权组合模型,并给出了验证计算过程,可为光缆维护管理部门对光缆线路传输损耗未来的发展趋势预测提供一定的参考。本文分别采用网格寻优法和遗传算法对指数平滑模型的平滑系数 α 和SVR模型的参数进行求解,在一定程度上优化了计算过程。本文构建的基于变权组合模型的光缆传输损耗预测方法绕开了对传输损耗各种复杂影响因素的分析,直接从顶层入手对传输损耗的时间序列数据进行分析预测,后期研究可从传输损耗的各种影响因素和模型的优化两方面考虑,以提高光缆传输损耗预测的可靠性。

参考文献:

- [1] 夏贵进,张曦,张居梅. 基于三次指数平滑法的光纤损耗预测研究[J]. 光通信技术,2014,38(1):35-37.
- [2] 傅蕾,李卫,张曦,等. 基于组合模型的光缆性能指标预测[J]. 光通信技术,2016,40(2):53-55.
- [3] CAI Z J, LU S, ZHANG X B. Tourism demand forecasting by Support vector regression and genetic algorithms [C]//2nd IEEE International Conference on Computer Science and Information Technology, August 8, 2009, Beijing, China. Beijing: IEEE Beijing SECTION, 2009.
- [4] 邓聚龙. 灰色系统理论及应用[M]. 北京:科学出版社,1999.
- [5] 郎茂祥. 预测理论与方法[M]. 北京:清华大学出版社,2011.
- [6] 陈荣,量昌勇,陆文星,等. 基于季节 SVR-PSO 的旅游客流量预测模型研究[J]. 系统工程理论与实践,2014,34(5):1290-1296.
- [7] 孙铁轩,邵春福,计寻,等. 基于 ARIMA 与信息粒化 SVR 组合模型的交通事故时序预测[J]. 清华大学学报(自然科学版),2014,54(3):348-353.
- [8] LEMKE C, GABRYS B. Meta-learning for time series forecasting and forecast combination[J]. Neurocomputing, 2010 (73): 2006-2016.

参考文献著录规则(二)——论文集、专利文献、学位论文和科技报告的著录格式

◆论文集的著录格式:

[序号] 著者. 题名[C]//著者. 专题名:其他题名. 出版地:出版者,出版年:起-止页码.

[1] 白书龙. 植物开花研究[C]//李承森. 植物科学进展. 北京:高等教育出版社,1998:146-163.

◆专利文献著录格式:

[序号] 专利申请者或所有者. 专利名称:专利国别,专利号[文献类型标志]. 公告日期或公开日期[引用日期].

[1] 姜锡洲. 一种温热外敷药制备方案:中国, 88105607.3[P]. 1989-07-26.

◆学位论文著录格式

[序号] 作者. 题名[文献类型标志]. 保存地:保存者,年份.

[1] 李洁. 促进我国国家创新系统中知识流动的政策分析[D]. 北京:中国科学院文献情报中心,2004.

◆科技报告的著录格式:

[序号] 主要责任者. 题名:其他题名信息[文献类型标志]. 其他责任者. 版本项. 出版地:出版者,出版年:起-止页码.

[1] World Health Organization. Factor regulating the immune response:report of WHO Scientific Group[R]. Geneva:WHO,1970.