

DOI:10.19431/j.cnki.1673-0062.2019.01.007

优化灰色 GM(1,N)-加权 Markov 模型在道路交通噪声 预测中的精度研究

黄朝强,廖基定*,尹邦华

(南华大学 数理学院,湖南 衡阳 421001)

摘要:通过对北京市 2007 年至 2016 年城市道路交通噪声及相关影响因素数据分析,以 GM(1,N)模型为基础,建立了优化灰色-加权 Markov 模型,为有效控制交通噪声污染提供理论依据和决策意见。首先,利用平滑公式对原始数据进行预处理,用数值积分中的 Simpson 公式改变背景值来提高传统多因素 GM(1,N)模型精度。其次,用加权 Markov 模型对得到的模拟值中的异常值进行了修正,将其应用到城市交通噪声的预测上,实证计算表明优化灰色 GM(1,N)模型的模拟值与实际值拟合效果很好,比传统的 GM(1,N)模型精度有较大提高。最后,用该模型对北京市 2017 年和 2018 年城市交通噪声进行预测,基本符合噪声数据实际变化趋势。

关键词:噪声预测;GM(1,N)预测模型;加权 Markov 模型;关联度分析;背景值优化
中图分类号:O211 **文献标志码:**A **文章编号:**1673-0062(2019)01-0035-09

Study on Forecasting Accuracy of Road Traffic Noise by Optimizing Grey GM (1,N)-Weighted Markov Model

HUANG Chaoqiang, LIAO Jiding*, YIN Banghua

(School of Mathematics and Physics, University of South China, Hengyang,
Hunan 421001, China)

Abstract: Based on the analysis of urban road traffic noise and related influencing factors from 2007 to 2016 in Beijing, the study establishes an optimized grey-weighted Markov model based on GM(1,N) model, which provides a theoretical basis and decisive suggestions for effective control of traffic noise pollution. Firstly, it uses the smoothing formula to preprocess the original data, and uses the Simpson formula in numerical integration to

收稿日期:2018-08-11

基金项目:湖南省科技厅软科学课题基金资助项目(2010zk3052)

作者简介:黄朝强(1992-),男,硕士研究生,主要从事概率论与数理统计等方面的研究. E-mail:1196628367@qq.com. * 通信作者:廖基定(1964-),男,教授,博士,主要从事保险风险理论方面的研究. E-mail:liaojiding@163.com

change the background value to improve the accuracy of the traditional multi-factor GM(1, N) model. Secondly, the outliers in the simulated values obtained by the weighted Markov model are modified and applied to the prediction of urban traffic noise. Empirical calculation shows that the simulated value of the Optimized Grey GM(1, N) model fits the actual value very well, and the accuracy of the model is much higher than that of the traditional GM(1, N). Finally, the model is used to predict the urban traffic noise in Beijing in 2017 and 2018, and it basically accords with the actual trend of noise data.

key words: noise prediction; GM(1, N) prediction model; weighted Markov model; correlation analysis; background value optimization

0 引言

当前社会,城市噪声污染日渐严重,不仅对人们的身体健康造成影响,而且对心理也会造成严重的影响,其中最主要来源是交通噪声污染。因此,对于交通噪声污染的控制已经到了迫在眉睫的地步。为了能及时有效的对城市交通噪声污染作出预测,为治理交通噪声污染提供科学的依据,找到一种高精度的预测方法是十分必要的。

当前,对于城市道路交通噪声的预测方法中,通常用灰色系统模型来研究,并且得到的预测结果精度较高。灰色系统模型是用小样本来进行一次累加满足指数分布的一种预测模型^[1],当前对 GM(1, 1) 模型的研究,在理论方面^[2-3]和实际预测方面^[4]都受到了广大专家学者的关注。各学者为了提高该模型的精度,分别就如何提高原始序列的平滑度^[5]、改进序列算子^[6]、改变初始条件^[7]、优化背景值^[8-10]、对最初的数据进行及时更新等方面进行了深入地研究。

对于 GM(1, N) 的研究相对较少,但是在理论^[11-14]和应用^[15-16]方面也有了一定的基础,并且多数是用 Markov 模型^[17-18]来进行修正,本文受沈艳等人^[19-20]利用数值积分对 GM(1, N) 进行背景值优化的启发,用 Simpson 公式对背景值进行优化,并且用加权 Markov 模型对其进行异常值修正,使其精度提高,并用 2007 年至 2016 年北京市城市交通噪声平均声级数据,采用改进的灰色 GM(1, N)-加权 Markov 模型进行了模拟和修正,预测了未来两年的北京市城市交通噪声平均声级的发生趋势。

1 灰色模型和 GM(1, N) 预测模型原理

灰色预测是利用小样本来预测未来的序列,将无规律的原始数据经过一次累加后生成较强的

服从指数分布的数据序列,建立相应的微分方程模型,再用最小二乘法求出相应参数。

设原始数列为

$$X_1^{(0)} = \{x_1^{(0)}(1), x_1^{(0)}(2), \dots, x_1^{(0)}(n)\}$$

另外 $N-1$ 个相关因素序列为

$$X_i^{(0)} = \{x_i^{(0)}(1), x_i^{(0)}(2), \dots, x_i^{(0)}(n)\}, i=2, 3, \dots, N.$$

对数列 $X_i^{(0)}$ 做累加生成,得到 1-AGO 序列

$$X_i^{(1)} = \{x_i^{(0)}(1), \sum_{k=1}^2 x_i^{(0)}(k), \dots, \sum_{k=1}^N x_i^{(0)}(k)\},$$

则得到 GM(1, N) 的影子方程为:

$$\frac{dx_i^{(1)}(t)}{dt} + ax_i^{(1)}(t) = \sum_{i=2}^N b_i x_i^{(1)}(t) \quad (1)$$

于是 GM(1, N) 模型的近似时间响应解为:

$$\hat{x}_1^{(1)}(k+1) = [x_1^{(1)}(1) - \frac{1}{a} \sum_{i=2}^N b_i x_i^{(1)}(k+1)] e^{-ak} + \frac{1}{a} \sum_{i=2}^N b_i x_i^{(1)}(k+1) \quad (2)$$

其预测模拟值为:

$$\hat{x}_1^{(0)}(k+1) = \hat{x}_1^{(1)}(k+1) - \hat{x}_1^{(1)}(k)$$

记参数列 $\hat{a} = [a, b_2, b_3, \dots, b_N]^T$

$$z_1^{(1)}(k) = \frac{1}{2} [x_1^{(1)}(k) + x_1^{(1)}(k-1)] \quad (k=2, 3, \dots, n) \quad (3)$$

由最小二乘解得:

$$\hat{a} = (B^T B)^{-1} B^T Y \quad (4)$$

其中:

$$B = \begin{bmatrix} -z_1^{(1)}(2) & x_2^{(1)}(2) & \cdots & x_N^{(1)}(2) \\ -z_1^{(1)}(3) & x_2^{(1)}(3) & \cdots & x_N^{(1)}(3) \\ \vdots & \vdots & & \vdots \\ -z_1^{(1)}(n) & x_2^{(1)}(n) & \cdots & x_N^{(1)}(n) \end{bmatrix}$$

$$Y = \begin{bmatrix} x_1^{(0)}(2) \\ x_1^{(0)}(3) \\ \vdots \\ x_1^{(0)}(n) \end{bmatrix}$$

从模型中可以看出,模型的精度依赖于背景值的选取,故背景值选取的优劣决定了整个模型的精度,本文对背景值做了如下优化改造。

2 基于数值积分的 GM(1,N) 模型优化

传统的 GM(1,N) 模型背景值通常是用两点之间的中间值来替换(梯形公式),此方法误差较大,并会造成预测值的误差叠加。

对公式(1)两边从 k_i 到 k_i+1 进行积分:

$$\int_{k_i}^{k_i+1} \frac{dx^{(1)}(t)}{dt} dt + a \int_{k_i}^{k_i+1} x^{(1)}(t) dt = \int_{k_i}^{k_i+1} \sum_{i=2}^N b_i x_i^{(1)}(t) dt$$

可得:

$$[x^{(1)}(t)]_{k_i}^{k_i+1} + a \int_{k_i}^{k_i+1} x^{(1)}(t) dt = \int_{k_i}^{k_i+1} \sum_{i=2}^N b_i x_i^{(1)}(t) dt$$

整理得:

$$x^{(0)}(k_i+1) + a \int_{k_i}^{k_i+1} x^{(1)}(t) dt = \sum_{i=2}^N b_i \int_{k_i}^{k_i+1} x_i^{(1)}(t) dt$$

用 $\int_{k_i}^{k_i+1} x^{(1)}(t) dt$ 代替传统的背景值,用机械求积方法近似求解:

$$\int_{k_i}^{k_i+1} x^{(1)}(t) dt \approx \sum_{k=0}^N A_k f(x_k)$$

式中: x_k 称为求积节点; A_k 称为求积系数。

2.1 基于 Simpson 公式的 GM(1,N) 模型

将积分区间 $[a, b]$ 划分为 n 等份,步长 $h = \frac{b-a}{n}$,等距节点为 $x_k = a + kh, k = 1, 2, \dots, n$ 。

可得到

$$I(f) = \int_a^b f(x) dx = (b-a) \sum_{k=0}^n C_k^{(n)} f(x_k),$$

式中 $C_k^{(n)}$ 称为 cores 系数。

$$\begin{cases} u_i^{(0)}(1) = \frac{3x_i^{(0)}(1) + x_i^{(0)}(2)}{4} \\ u_i^{(0)}(k) = \frac{x_i^{(0)}(k-1) + 2x_i^{(0)}(k) + x_i^{(0)}(k+1)}{4}, (k=1, 2, \dots, n-1) \\ u_i^{(0)}(n) = \frac{x_i^{(0)}(n-1) + 3x_i^{(0)}(n)}{4} \end{cases} \quad (6)$$

得到平滑处理后的新序列

$$U_i^{(0)} = \{u_i^{(0)}(1), u_i^{(0)}(2), \dots, u_i^{(0)}(n)\}。$$

2.3 模型的精度检验

对模型进行相对误差(α),小误差概率(p)和后验差比值(C)检验。

设原始序列

$$X_1^{(0)} = \{x_1^{(0)}(1), x_1^{(0)}(2), \dots, x_1^{(0)}(n)\},$$

令 $x = a + th$, 有

$$\begin{aligned} C_k^{(n)} &= \frac{h}{b-a} \int_0^n \prod_{\substack{j=0 \\ j \neq k}}^n \frac{t-j}{k-j} dt \\ &= \frac{(-1)^{n-k}}{nk! (n-k)!} \int_0^n \prod_{\substack{j=0 \\ j \neq k}}^n (t-j) dt \end{aligned}$$

基于模型的稳定性与精度,取 $n=2$,满足三次代数精度,则

$$C_0^{(2)} = \frac{1}{4} \int_0^2 (t-1)(t-2) dt = \frac{1}{6},$$

$$C_1^{(2)} = -\frac{1}{2} \int_0^2 t(t-2) dt = \frac{4}{6},$$

$$C_2^{(2)} = \frac{1}{4} \int_0^2 t(t-2) dt = \frac{1}{6}$$

得到 Simpson 公式:

$$S = \frac{b-a}{6} [f(a) + 4f(\frac{a+b}{2}) + f(b)]$$

则背景值为:

$$\begin{aligned} z^{(1)}(k_i+1) &= \int_{k_i}^{k_i+1} x^{(1)}(t) dt \\ &= \frac{1}{6} [x^{(1)}(k_i) + 4x^{(1)}(k_i + \frac{1}{2}) + x^{(1)}(k_i+1)] \end{aligned} \quad (5)$$

对应的新背景值只需用插值求出 $x^{(1)}(k_i + \frac{1}{2})$ 代入式(5)即可。

2.2 移动平滑法

为了降低预测的误差,减少原始数据中异常值的影响,运用平滑法^[21]进行预处理。

原始序列

$$X_i^{(0)} = \{x_i^{(0)}(1), x_i^{(0)}(2), \dots, x_i^{(0)}(n)\},$$

其中 $x_i^{(0)}(k) \geq 0, k = 1, 2, \dots, n$, 运用以下公式:

模拟序列

$$\hat{x}_1^{(0)} = \{\hat{x}_1^{(0)}(1), \hat{x}_1^{(0)}(2), \dots, \hat{x}_1^{(0)}(n)\}$$

残差序列:

$$\varepsilon(k) = x^{(0)}(k) - \hat{x}^{(0)}(k) (k = 1, 2, \dots, n) \quad (7)$$

第 k 个数的相对误差:

$$\alpha = \frac{1}{n} \sum_{k=1}^n \Delta k,$$

其中 $\Delta k = \left| \frac{\varepsilon(k)}{x^{(0)}(k)} \right|$

原始数列的标准差:

$$S_1 = \sqrt{\frac{1}{n} \sum_{k=1}^n [x^{(0)}(k) - \bar{x}]^2}, \bar{x} = \frac{1}{n} \sum_{k=1}^n x^{(0)}(k);$$

模拟残差的标准差:

$$S_2 = \sqrt{\frac{1}{n} \sum_{k=1}^n [\varepsilon(k) - \bar{\varepsilon}]^2}, \bar{\varepsilon} = \frac{1}{n} \sum_{k=1}^n \varepsilon(k)$$

后验差比值和小误差概率检验公式:

$$C = \frac{S_2}{S_1}, p = p(|\varepsilon(k) - \bar{\varepsilon}| < 0.6745 S_1) \quad (8)$$

预测检验等级通常分为 4 级^[22], 如表 1 所示。

表 1 模拟检验等级表

Table 1 Simulation test grade table

模型的精度等级	p	C
好(一级)	>0.95	<0.35
合格(二级)	>0.80	<0.45
勉强合格(三级)	>0.70	<0.50
不合格(四级)	≤0.70	≥0.65

3 加权 Markov 模型修正

Markov 模型对未来状态的预测方法, 仅取决于当前的状态, 与过去状态是无关的, 即无后效性, 即

$$p(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i_n) = p(X_{n+1} = j | X_n = i_n).$$

在实际研究中, 有的随机变量序列满足“马氏性”, 但通常不满足“齐次性”, 且传统的 Markov 模型并不能有效的利用陈旧数据, 减少了各阶 Markov 模型^[23]的绝对分布在预测中的影响作用。而加权 Markov 模型^[24-25]方法弥补了传统 Markov 模型方法的缺陷。

本文运用加权 Markov 模型的 Markov 性, 对预测值的波动规律进行调整, 修正灰色模型异常预测结果。

3.1 划分状态

运用均值-标准差法^[26], 具体如下: 设随机序列 $X = \{x_1, x_2, \dots, x_n\}$, 原始数列的均值: $\bar{x} =$

$\frac{1}{n} \sum_{i=1}^n x_i$; 原始数列的标准差:

$$S = \sqrt{\frac{1}{n} \sum_{i=1}^n [x_i - \bar{x}]^2} (i \in T)$$

以相对误差 $\bar{x}$ 为标准, 划分为 m 个状态, 得到状态空间 $T = [T_1, T_2, \dots, T_m]$, 若 $T_i \in [Q_i, S_i]$, 其中

$$\begin{cases} Q_i = \inf(T_i) \\ S_i = \sup(T_i) \end{cases}^{\circ}$$

3.2 构造状态转移概率矩阵

用 $p_{ij}^{(k)}$ 表示由 Markov 来状态 T_i 经过 k 步转移到状态 T_j 的概率, 根据原始数据可知, T_i 状态下的数据个数 m_i , 由状态 T_i 经过 k 步转移到状态 T_j 的个数为 $m_{ij}^{(k)}$, 则 k 步转移概率为:

$$p_{ij}^{(k)} = p(T_j | T_i) = p(T_i \rightarrow T_j) = \frac{m_{ij}^{(k)}}{m_i} \quad (i, j = 1, 2, \dots, m) \quad (9)$$

得到 k 步转移概率矩阵:

$$p(k) = \begin{bmatrix} p_{11}^{(k)} & p_{12}^{(k)} & \cdots & p_{1m}^{(k)} \\ p_{21}^{(k)} & p_{22}^{(k)} & \cdots & p_{2m}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1}^{(k)} & p_{m2}^{(k)} & \cdots & p_{mm}^{(k)} \end{bmatrix};$$

$$p(k) = [p(1)]^k \quad (10)$$

因此只需算出一步转移概率矩阵, 就可得出 k 步转移概率矩阵。

3.3 随机序列的马氏检验

用 χ^2 统计量进行马氏性检验, 由转移状态概率矩阵的各行各列之和得到边际概率, 记为 $p_{\cdot j}$, 即为:

$$p_{\cdot j} = \frac{\sum_{i=1}^m m_{ij}}{\sum_{i=1}^m \sum_{j=1}^m m_{ij}} \quad (11)$$

当 n 充分大时, 统计量为:

$$\chi^2 = 2 \sum_{i=1}^m \sum_{j=1}^m m_{ij} \left| \lg \frac{p_{ij}}{p_{\cdot j}} \right| \quad (12)$$

服从自由度为 $(m-1)^2$ 的 χ^2 分布。确定显著水平 α , 查表可得到分位点 $\chi_{\alpha}^2((m-1)^2)$ 的值, 当 $\chi^2 > \chi_{\alpha}^2((m-1)^2)$ 时, 则认为该随机序列满足马氏性, 否则不满足, 也就是该随机序列不能用 Markov 来预测。

3.3.1 计算各阶自相关系数

$$r_k = \frac{\sum_{i=1}^{n-k} (\hat{x}_1^{(0)}(i) - \hat{\bar{x}}_1^{(0)}) (\hat{x}_1^{(0)}(i+k) - \hat{\bar{x}}_1^{(0)})}{\sum_{i=1}^{n-k} (\hat{x}_1^{(0)}(i) - \hat{\bar{x}}_1^{(0)})^2} \quad (13)$$

式中: r_k 为滞后 k ($k \in T$) 阶时的自相关系数;

对各阶自相关系数进行归一化处理, 得

$$W_k = |r_k| / \sum_{k=1}^m |r_k| \quad (14)$$

将 W_k 作为各阶步长的 Markov 模型的权重。

3.3.2 加权 Markov 模型预测

以之前时刻的随机序列为初始状态,结合各阶状态转移概率矩阵即可预测出该时刻随机序列的状态概率 $p_i^{(k)} (i \in T), k=1, 2, \dots, m$ 。将同一状态下的各预测概率加权和作为随机序列在该状态下的预测概率,即:

$$p_i = \sum_{k=1}^m W_k p_i^{(k)} (i \in T) \quad (15)$$

$\max(p_i, i \in T)$ 所对应的 i 为该时刻随机序列的预测状态,然后将预测结果加入初始状态的随机序列即可得到新的随机序列,重复以上过程即可预测较长时间段的状态。

3.3.3 加权 Markov 优化取值

设 t 时刻灰色预测值为 $\hat{x}^{(0)}(t)$, T_i 为对应指标的预测状态,即状态区间 $[Q_i, S_i]$,对预测值进行优化,优化后的灰色预测值为:

$$\hat{y}(t) = 0.5(Q_i + S_i) \quad (16)$$

4 仿真实例分析

4.1 仿真数据

从《北京市统计年鉴》选取北京市 2007 年至 2016 年道路交通噪声平均值和其相关影响因素,主要有:城市人口数量、城市道路里程、机动车流量等^[27]为例研究改进的灰色 Markov 模型预测精度,表 2 给出了具体数值。

表 2 北京市道路交通平均噪声及其相关影响因素数据表

Table 2 Data table of average road traffic noise and its related influencing factors in Beijing

系统因素	年份									
	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
道路交通噪声 平均声级 $X_1^{(0)}$ /dB	69.9	69.6	69.7	70.0	69.6	69.2	69.1	69.1	69.2	69.3
城市人口数量 $X_2^{(0)}$ /万人	1 676.0	1 771.0	1 860.0	1 961.9	2 018.6	2 069.3	2 114.8	2 151.6	2 170.5	2 172.9
城市道路里程 $X_3^{(0)}$ /km	4 460	6 186	6 247	6 355	6 258	6 271	6 295	6 426	6 423	6 373
机动车拥有量 $X_4^{(0)}$ /万辆	312.80	350.40	401.90	480.90	498.30	520.00	543.70	559.10	561.90	571.70
GDP $X_5^{(0)}$ /亿元	25 669.1	11 392	12 419	14 441	16 627.9	18 350.1	20 330.1	21 944.1	23 685.7	25 669.1
建筑业企业房屋 建筑施工面积 $X_6^{(0)}$ /万 m^2	18 225	19 537	22 721	29 440	36 507	41 660	49 259	56 477	59 777	61 098

4.2 关联度分析主要关联因素

计算 $X_i^{(0)} = \{x_i^{(0)}(1), x_i^{(0)}(2), \dots, x_i^{(0)}(n)\}$

与 $X_i^{(0)} = \{x_i^{(0)}(1), x_i^{(0)}(2), \dots, x_i^{(0)}(n)\} (i=2, 3, \dots, N)$ 的灰色关联系数^[28]:

$$\gamma(x_i^{(0)}(k), x_1^{(0)}(k)) = \frac{\min_i \min_k |x_1^{(0)}(k) - x_i^{(0)}(k)| + \xi \max_i \max_k |x_1^{(0)}(k) - x_i^{(0)}(k)|}{|x_1^{(0)}(k) - x_i^{(0)}(k)| + \xi \max_i \max_k |x_1^{(0)}(k) - x_i^{(0)}(k)|} \quad (17)$$

式中:分辨系数 $\xi \in [0, 1]$,取 $\xi = 0.5$, $X_1^{(0)}$ 与 $X_i^{(0)}$

各影响指标对交通噪声的影响程度大小与 γ_i 成正比, γ_i 越大则说明影响越大,反之越小。

的灰色关联度 $\gamma_i = \frac{1}{n} \sum_{k=1}^n \gamma(x_1^{(0)}(k), x_i^{(0)}(k))$ 。

由公式(17)计算得表 3。

表 3 各影响因素对交通噪声污染的灰色关联度结果

Table 3 Results of grey correlation of factors affecting traffic noise pollution

影响指标	城市人口数量 $x_2^{(0)}$	城市道路里程 $x_3^{(0)}$	机动车拥有量 $x_4^{(0)}$	GDP $x_5^{(0)}$	建筑业企业房屋建筑施工面积 $x_6^{(0)}$
灰色关联度 γ_i	0.780	0.494	0.733	0.390	0.817

由表4中的实际值、传统 GM(1,N)模型的预测值和表6中经过平滑处理的基于 Simpson 公式的 GM(1,N)模型的预测值处理得到图1。

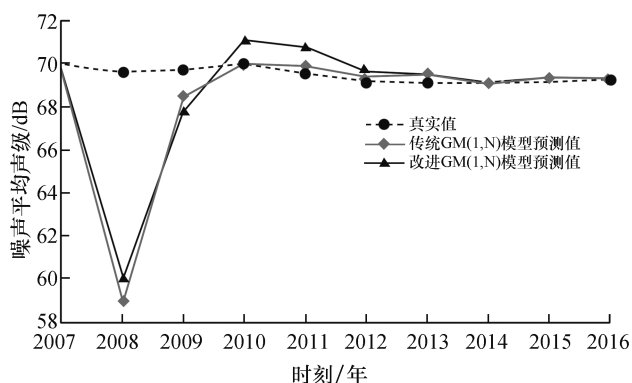


图1 2007年至2016年北京市城市交通噪声实际值与模拟值的拟合图

Fig. 1 Fitting diagram of actual and simulated values of urban traffic noise in Beijing from 2007 to 2016

从图1中可以看出改进的 GM(1,N)模型比传统的 GM(1,N)模型拟合程度更高,但在2008年和2009年的拟合效果很差,因此用加权 Markov 模型对异常值进行反向修正。

4.5 加权 Markov 模型修正

转移概率矩阵确定以后,就可以根据前几年的预测值得到自相关系数归一化之后作为权重,选择最大的权重值来判别属于的状态区间,以此得到加权 Markov 模型修正值。

4.5.1 划分状态区间

根据2010年至2016年统计数据,得到均值 $\bar{x}=69.837$ 标准差 $s=0.783$,用均值标准差法将道路交通噪声平均声级划分为4个状态,即 $T=\{T_1, T_2, T_3, T_4\}$,详见表7。

表7 状态区间划分
Table 7 State interval partition

状态	分级标准	数值区间
T_1	$x \leq \bar{x} - 0.5s$	$x \leq 69.45$
T_2	$\bar{x} - 0.5s < x \leq \bar{x}$	$69.45 < x \leq 69.84$
T_3	$\bar{x} < x \leq \bar{x} + 0.5s$	$69.84 < x \leq 70.23$
T_4	$x > \bar{x} + 0.5s$	$x > 70.23$

4.5.2 构建状态矩阵

根据表6中2010年至2016年的预测值 $\hat{x}_i^{(0)}(t)$ 和表7的状态区间划分标准得到表8。

表8 各年份预测值所在的状态区间

Table 8 The interval status of the predicted values for each year

年份	改进 GM(1,N)模型下的预测值 $\hat{x}_i$	所处状态区间
2016	69.32	T_1
2015	69.39	T_1
2014	69.09	T_1
2013	69.52	T_2
2012	69.65	T_2
2011	70.77	T_3
2010	71.12	T_4

根据表8和公式(9)、式(10),构建状态转移概率矩阵,得到一步到四步状态转移概率矩阵:

$$\begin{aligned}
 p(1) &= \begin{bmatrix} 0.6667 & 0.3333 & 0.0000 & 0.0000 \\ 0.0000 & 0.5000 & 0.5000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \\ 1.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix} \\
 p(2) &= \begin{bmatrix} 0.4445 & 0.3889 & 0.1666 & 0.0000 \\ 0.0000 & 0.2500 & 0.2500 & 0.5000 \\ 1.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.6667 & 0.3333 & 0.0000 & 0.0000 \end{bmatrix} \\
 p(3) &= \begin{bmatrix} 0.2963 & 0.3426 & 0.1944 & 0.1666 \\ 0.5000 & 0.1250 & 0.1250 & 0.2500 \\ 0.6667 & 0.3333 & 0.0000 & 0.0000 \\ 0.4445 & 0.3889 & 0.1666 & 0.0000 \end{bmatrix} \\
 p(4) &= \begin{bmatrix} 0.3642 & 0.2701 & 0.1713 & 0.1744 \\ 0.5834 & 0.2291 & 0.0625 & 0.1250 \\ 0.4445 & 0.3889 & 0.1666 & 0.0000 \\ 0.2963 & 0.3426 & 0.1944 & 0.1666 \end{bmatrix}
 \end{aligned}$$

对随机序列进行马氏性检验。取显著性水平 $\alpha=0.05$,由公式(11)、式(12)计算得统计量 $\chi^2=20.35 > \chi_{\alpha}^2(9)=16.92$,因此该序列通过马氏性检验,接下来进行 Markov 模型修正。

由式(13)计算得到各阶自相关系数为: $r=(0.9943, 0.9759, 0.9738, 0.9667)$ 。

由式(14)进行归一化之后可得到各步长的权重向量: $W=(0.2543, 0.2495, 0.2490, 0.2472)$ 。

再由式(15),根据2010年至2013年的原始数据所属的状态及其状态转移概率矩阵,对北京市城市交通平均噪声声级进行反向修正,结果见表9。

表 9 2009 年北京市城市交通噪声平均声级状态修正
Table 9 Revision of average sound level of urban traffic noise in Beijing in 2009

初始年份	初始状态	滞时年	权重	状态 T_1	状态 T_2	状态 T_3	状态 T_4
2013	T_2	4	0.254 3	0.584 3	0.229 1	0.062 5	0.125 0
2012	T_2	3	0.249 5	0.500 0	0.125 0	0.125 0	0.250 0
2011	T_3	2	0.249 0	1.000 0	0.000 0	0.000 0	0.000 0
2010	T_4	1	0.247 2	1.000 0	0.000 0	0.000 0	0.000 0
P_i				0.769 5	0.089 4	0.047 1	0.094 2

由表 9 可以看出,将同一状态的各预测的概率加权求和得到 $\max \{p_i, i \in T\} = 0.769 5$, 则 2009 年道路交通噪声平均声级属于状态 T_1 , 应该落在 $68.67 < x \leq 69.45$ 范围内, 由公式 (16) 取 68.67 和 69.45 的平均值得到 69.09 dB, 而与 2009 年的真实值为 69.7, 修正之后的预测值的相对误差 $PRE = 0.008 8$, 使得模型精度进一步提高。同理可得到 2008 年道路交通噪声平均声级为 70.425 dB。

由表 6 中的实际值、传统 GM(1,N)模型的预测值和加权 Markov 模型修正后的改进 GM(1,N)的预测值处理得到图 2。

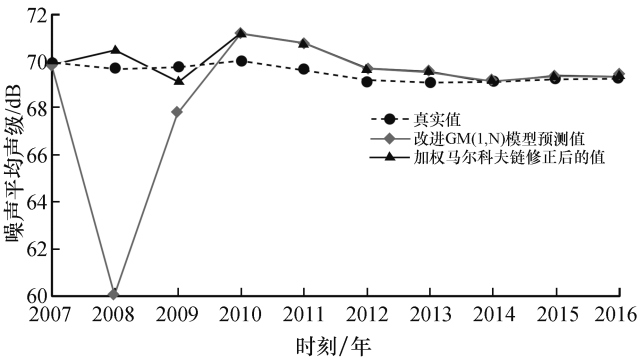


图 2 2007 年至 2016 年北京市城市交通噪声平均声级拟合图

Fig. 2 Fitting map of average sound level of urban traffic noise in Beijing from 2007 to 2016

从图 2 可看出经过加权 Markov 模型修正后的改进 GM(1,5)的预测值与真实值的拟合程度比未修正之前精度更高。

由表 10 可知:应用改进的方法,预测的平均相对误差为 1.679%,经过 Markov 修正后平均相对误差减少为-0.517%。由此可见,本文提出的方法精度较高,能满足实际的需求。

4.6 改进灰色-Markov 模型预测

由 2007 年至 2016 年和 2008 年至 2017 年北

京市城市道路交通噪声声级数据进行优化灰色 GM(1,N)-Markov 模型递推预测,计算得 2017 年和 2018 年北京市城市道路交通噪声声级分别为 69.645 dB 和 70.035 dB。由此可见,随着现在城市交通车辆、人口、建筑施工的不断增加,北京市的交通噪声平均声级接下来两年呈递增趋势。

表 10 改进 GM(1,5)模型与加权 Markov 模型修正后方法的比较

Table 10 Comparison of modified methods between modified and weighted Markov models %	
方法	平均相对误差
改进 GM(1,5)模型	1.679
加权 Markov 模型修正	-0.517

5 结 论

本文通过用基于 Simpson 公式的数值积分的方法改进灰色 GM(1,N)模型中的背景值并且对原始数据进行了平滑处理。结果表明,改进灰色 GM(1,N)模型得到的预测值比传统的灰色 GM(1,N)模型预测值的精度更高。最后,用加权 Markov 模型对改进灰色 GM(1,N)模型的预测值中的异常值进行修正,使得平均相对误差从 1.679% 减少到-0.517%,此结果也说明了该模型的实用性。

本文就小样本条件下的噪声平均声级进行了模拟和预测,在进行数据仿真过程中得出如下结论:

- 1)利用平滑方法对原始序列进行了平滑处理,提高了原始序列的平滑度,降低了异常值对预测结果的影响。
- 2)用 Simpson 公式对背景值进行了改进,优化了传统多因素 GM(1,N)模型的模拟值,将预测噪声声级和实际噪声声级相比较,改进的方法比

传统的方法精度更高。

3)应用加权 Markov 模型对在 Simpson 公式下的 GM(1,N)模型得出的交通噪声声级模拟值中的明显的异常值进行了修正,且整体精度有较大提高。

参考文献:

- [1] 邓聚龙. 灰理论基础[M]. 武汉:华中科技大学出版社,2002:221-233.
- [2] 罗党,刘思峰,党耀国. 灰色模型 GM(1,1)优化[J]. 中国工程科学,2003,5(8):50-53.
- [3] 谢乃明,刘思峰. 离散灰色模型的拓展及其最优化求解[J]. 系统工程理论与实践,2006,26(6):108-112.
- [4] LIU S, FORREST J. The role and position of grey system theory in science development[J]. The journal of grey system,1997,9(4):351-356.
- [5] 芮海田,吴群琪,袁华智,等. 基于指数平滑法和马尔科夫模型的公路客运量预测方法[J]. 交通运输工程学报,2013,13(4):87-93.
- [6] 许秀莉,罗建. GM(1,1)模型的改进方法及其应用[J]. 系统工程与电子技术,2002,24(4):61-63.
- [7] 党耀国,刘思峰,刘斌. 以 $x^{(1)}(n)$ 为初始条件的 GM 模型[J]. 中国管理科学,2005,13(1):132-136.
- [8] 宋中民,同小军,肖新平. 中心逼近式灰色 GM(1,1)模型[J]. 系统工程理论与实践,2001,21(5):110-113.
- [9] 李俊峰,戴文战. 基于插值和 Newton-Cotes 公式的 GM(1,1)模型的背景值构造新方法与应用[J]. 系统工程理论与实践,2004,24(10):122-126.
- [10] 江南,刘小洋. 基于 Gauss 公式的 GM(1,1)模型的背景值构造新方法与应用[J]. 数学的实践与认识,2008,38(7):90-94.
- [11] 吴利丰,刘思峰,姚立根. 基于分数阶累加的离散灰色模型[J]. 系统工程理论与实践,2014,34(7):1822-1827.
- [12] 党耀国,魏龙,丁松. 基于驱动信息控制项的灰色多变量离散时滞模型及其应用[J]. 控制与决策,2017,32(9):1672-1680.
- [13] 毛树华,高明运,肖新平. 基于 Gauss 公式的 GM(1,1)模型的背景值构造新方法与应用[J]. 系统工程理论与实践,2015,35(2):430-436.
- [14] 张可,曲品品,张隐桃. 时滞多变量离散灰色模型及其应用[J]. 系统工程理论与实践,2015,35(8):2092-2103.
- [15] 王正新. 灰色多变量 GM(1,n)模型及其应用[J]. 系统工程理论与实践,2014,34(9):2357-2363.
- [16] 张熙胤,陈兴冲,王常峰. 多变量 GM(1,n)模型在桥梁施工挠度控制中的应用[J]. 世界桥梁,2013,41(5):76-80.
- [17] 樊超,郭亚菲,曹培格,等. 基于灰色-马尔科夫模型的粮食产量预测[J]. 江苏农业科学,2018,46(9):346-349.
- [18] 侯月,廖基定. 基于马尔科夫过程修正的湖南洪涝灾变灰预测研究[J]. 南华大学学报(自然科学版),2017,31(4):95-100.
- [19] 沈艳,孙红影,李丽萍. 优化 GM(1,n)模型在交通噪声预测中的应用和精度分析[J]. 中国安全生产科技技术,2012,8(11):27-31.
- [20] 何满喜,王勤. 基于 Simpson 公式的 GM(1,n)建模的新算法[J]. 系统工程理论与实践,2013,33(1):199-202.
- [21] 杨杰,翁文国. 基于改进无偏灰色模型的燃气供气量的预测[J]. 清华大学学报(自然科学版),2014,54(2):145-148.
- [22] 宁宣熙,刘思峰. 管理预测与决策方法[M]. 2版. 北京:科学出版社,2009:113-145.
- [23] PENG Z, BAO C, ZHAO Y. Weighted markov chains for forecasting and analysis in incidence of infectious diseases in Jiangsu province, China[J]. Journal of biomedical research,2010,24(3):207-214.
- [24] 张祥,陈军,陆雨,等. 降水量的加权型马尔科夫模型预测[J]. 气象水文海洋仪器,2018,35(2):18-20.
- [25] 曹彦,何东进,洪伟,等. 加权马尔科夫链在福建省森林火灾预测中的应用研究[J]. 西南林业大学学报,2014,34(3):62-66.
- [26] 金峻炎,陈进. 加权马氏链在房地产投资回收期预测中的应用[J]. 湖南科技大学学报(自然科学版),2010,25(2):63-66.
- [27] 北京市统计局. 北京市统计年鉴[M]. 北京:中国统计出版社,2007-2016.
- [28] 李辉,彭晓春,钟志强,等. 灰色关联分析法在交通噪声影响因素分析中的应用[J]. 噪声与振动控制,2012,32(1):93-95.

(责任编辑:扶文静)