文章编号: 1671- 251X(2010)12- 0024- 04

基于时间序列的矿井瓦斯涌出量预测方法

孟海东， 孙搏， 司子稳， 王睿智， 施兰兰

(内蒙古科技大学矿业工程学院，内蒙古 包头 014010)

摘要: 由于矿井瓦斯浓度的变化受多种因素共同影响，矿井瓦斯涌出量预测经常出现无法获得一部分变量的情况。针对该问题，提出了一种基于时间序列的矿井瓦斯涌出量预测方法，详细介绍了采用时间序列 AR 模型对矿井瓦斯涌出量进行预测的具体实现。实验结果表明，该方法对矿井瓦斯涌出量的预测误差率为 4.3%，预测比较可靠。

关键词: 矿井; 瓦斯涌出量预测; 时间序列; 参数估计; AR 模型

中图分类号: **TD712.5** 文献标识码: **A**

Prediction Method of Mine Gas Emission Based on Time Series

MENG Hai-dong, SUN Bo, SI Zi-wen, WANG Rui-zhi, SHI Lan-lan

(School of Mining Engineering of Inner Mongolia University of Science and Technology, Baotou 014010, China)

Abstract: Because change of mine gas concentration is influenced by various factors, so prediction of mine gas emission can't get some variables. To solve the problem, the paper proposed a prediction method of mine gas emission based on time series. It introduced implementation of using AR model of time series to predict mine gas emission in details. The experiment result showed that error rate of prediction for mine gas emission with the method was 4.3% and the prediction is reliable.

Key words: mine, prediction of gas emission, time series, parameter estimation, AR model

0 引言

我国煤矿瓦斯事故已占到煤矿生产过程所发生事故的 80% 以上，造成的伤亡占特大事故伤亡人数的 90% [1]。因此，必须采取有效的防治措施，而防治工作的关键在于瓦斯涌出量预测。矿井瓦斯涌出量是一个动态过程，瓦斯浓度的变化受多种因素共同影响，矿井瓦斯涌出量预测经常出现无法获得一部分变量的情况。

时间序列分析法是根据客观事物发展的连续规律性，运用过去的历史数据，通过统计分析，进一步推测未来发展趋势的一种方法。由于该时间序列取自某一个随机过程，而该过程的随机特征不随时间的变化而变化，所以又称平稳时间序列分析法。时间序列分析法可通过建立一个描述瓦斯涌出量在一定时间和空间内变化发展的动态模型，反映瓦斯涌出量的变化规律，预测瓦斯涌出量的趋势，对实际

收稿日期: 2010- 08- 23

作者简介: 孟海东(1958-)，男，内蒙古托克托县人，教授，博士，硕士研究生导师，现主要从事矿产资源数据挖掘方面的研究工作。E-mail: 51428554@qq.com

全生产远程管理信息化，促进国内煤炭行业信息化建设的不断发展。

参考文献:

[1] 孙继平. 煤矿安全生产监控系统联网[J]. 工矿自动化，2009(10): 1-4.

[2] 黄富革，周晓芳. 高校信息化数据标准的制定与实施[J]. 企业科技与发展，2009(1): 89-91.

[3] 郝秦霞，赵安新，卢建军. 煤矿安全系统数据资源共享标准的构建[J]. 矿业安全与环保，2008, 35(2): 31-33.

[4] 程杰. 基于 Web 模式的安全信息系统集成标准的研究[D]. 西安: 西安科技大学，2007.

[5] 孙继平. 煤矿安全监控系统联网技术研究[J]. 煤炭学报，2009, 34(11): 108-111.

的瓦斯预测有一定的指导意义。

1 时间序列分析法简介

时间序列分析法就是通过编制和分析时间序列,根据时间序列所反映出来的发展过程、方向和趋势,进行类推或延伸,借以预测下一段时间可能达到的水平。其内容包括收集与整理历史资料;对这些资料进行检查鉴别,排成数列;分析时间数列,从中寻找随时间变化的规律,得出一定的模式;以该模式预测将来的情况。常见的时间序列模型有自回归(AR)模型、滑动平均(MA)模型和自回归滑动平均(ARMA)模型。由于 AR 模型能够更好地反映系统的本质特征,并且 AR 模型是无偏估计。因此,本文采用 AR 模型进行建模,其形式为

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon \qquad (1)$$

式中: ϕ_1、ϕ_2、…、ϕ_p 为模型参数; ε 为白噪声序列,它反映所有其它因素干扰。

式(1)表明, y_t 是自身过去的观察值 y_{t-1}、y_{t-2}、…、y_{t-p} 的线性组合,常记为 AR(p),其中 p 为模型的阶次。若记

$$1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p = 0 \qquad (2)$$

则式(1)可以改成算子形式:

$$\phi_p(B) y_t = \varepsilon_t \qquad (3)$$

式中: B 为移位算子,当 $\phi(B) = 0$ 时为模型的特征方程,特征方程的 p 个根 $\lambda_i (i = 1, 2, 3, \cdots, p)$ 称为 AR 模型的特征根。如果 p 个特征根都在单位圆外,即

$$|\lambda_i| > 1 \qquad (4)$$

则称 AR 模型是稳定的,式(4)又称平稳条件[2]。

2 AR(p) 模型建模过程

AR 模型建模就是对采集到的时间数据序列拟合出适用的 AR(p)模型,其中确定阶数 p 的估计值是关键,建模过程如下。

(1) 模型定阶

模型定阶就是 AR(p)模型中阶数 p 的确定,一般来说 AR 模型中阶数 p 是未知的,若应用模型进行仿真预测,就应该首先确定阶数 p 的值。其确定方法主要有 AIC(Akaike Information Criterion)准则估计法和 BIC 准则估计法等。

AIC 准则是 1971 年日本学者赤池给出的一种适用面非常广泛的统计模型选择准则,称为最小信息准则,运用该准则可以在 AR 模型参数估计的基础上估计阶数 p。首先引入 AIC 准则函数:

$$\mathrm{AIC}(k) = \ln \tilde{\delta}^2(k) + \frac{2k}{m} \quad k = 0, 1, \cdots, P \qquad (5)$$

式中: $\tilde{\delta}^2(k)$ 为 $p = k(1 \leqslant k \leqslant P)$ 时 δ^2 的估计; P 为 p 的估计上界,一般 P 的取值由实际经验而定。取 $\tilde{p}$ 满足:

$$\mathrm{AIC}(\tilde{p}) = \min_{1 \leqslant k \leqslant P} \mathrm{AIC}(k) \qquad (6)$$

则此 $\tilde{p}$ 为阶数 p 的 AIC 准则估计值[3]。

(2) 参数估计

AR(p)模型参数估计的方法有很多,主要包括最小二乘法、解 Yule- Walker 方程法、极大似然估计法和 Burg 算法等。上述方法中,最小二乘法进行参数估计比较简单,参数估计无偏差、精度高,本文选用该方法进行参数估计。已知样本序列 $X = \{x_1, x_2, \cdots, x_n\}$,将式(1)写成矩阵形式,有

$$\boldsymbol{y} = -X\varphi + \boldsymbol{e} \qquad (7)$$

其中 $\boldsymbol{y} = \begin{bmatrix} x_{p+1} \\ x_{p+2} \\ \vdots \\ x_m \end{bmatrix}$, $\varphi = \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_p \end{bmatrix}$, $\boldsymbol{e} = \begin{bmatrix} e_{p+1} \\ e_{p+2} \\ \vdots \\ e_m \end{bmatrix}$。

令

$$S(\varphi) = \sum_{p+1}^{m} (x_n + \varphi_1 x_{n-1} + \varphi_2 x_{n-2} + \cdots + \varphi_p x_{n-p})^2 = \sum_{p+1}^{m} e(n)^2 \qquad (8)$$

求 $\hat{\phi}$,使 $S(\hat{\phi}) = \min\{S(\varphi)\}$,称此时的 $\hat{\phi}$ 为最小二乘估计。由最小二乘估计的运算方法可得 $\hat{\phi}$ 与 δ^2 的最小二乘估计为

$$\hat{\phi} = (X'X)^{-1} X' \boldsymbol{y} \qquad (9)$$

$$\hat{\delta}^2 = \frac{1}{m-p} S(\hat{\phi}) = \frac{1}{m-p} \sum_{p+1}^{m} e(n)^2 \qquad (10)$$

(3) 模型检验

由于 AR(p)模型的识别与估计是在假设随机扰动项是一白噪声的基础上进行的,因此,如果估计的模型确认正确的话,残差应代表白噪声序列;如果残差不代表白噪声序列,则说明模型的识别与估计有误,需重新进行识别与估计。

(4) 预测及评价

对未来的数据进行预测,计算出相对误差及平均误差率(平均误差率的大小从一定程度上反应预测的精度):

$$E = \frac{q_s - q_r}{q_r} \times 100\% \qquad (11)$$

式中: E 表示相对误差,其平均值即为误差平均

率; q_s 为统计值; q_r 为统计预测值。

3　实例分析

本文以某矿某月 01—15 日实际检测的瓦斯涌出量(见表 1) 为例, 实现对未来 3 天瓦斯涌出量的预测。

表 1　某矿某月实际检测的瓦斯涌出量　　m^3/min

日期	瓦斯涌出量	日期	瓦斯涌出量	日期	瓦斯涌出量
01 日	51.57	07 日	39.73	13 日	40.97
02 日	51.01	08 日	38.74	14 日	41.20
03 日	49.29	09 日	39.13	15 日	39.39
04 日	48.80	10 日	37.71	16 日	39.92
05 日	45.97	11 日	38.95	17 日	39.38
06 日	41.63	12 日	37.99	18 日	39.43

设表 1 所示的样本序列为 $Y_t=\{y_1, y_2, \cdots, y_t\}$, 样本长度 $t=15$, 步长为 1 d。

(1) 将样本序列零均值化

从表 1 可看出, 01—07 日瓦斯涌出量在逐渐减少, 08—15 日趋向平稳, 样本均值显著非零。用 Y_t 表示时间序列, $\overline{Y}$ 表示样本均值, 则

$$\overline{Y}=\frac{1}{15}\sum_{i=1}^{15}Y_t \tag{12}$$

令 $X_t=Y_t-\overline{Y}$, 可得到时间序列零均值$\{X_t\}$, 如表 2 所示。

表 2　时间序列零均值

t	X_t	t	X_t	t	X_t
1	8.764 7	6	−1.175 3	11	−3.855 3
2	8.204 7	7	−3.075 3	12	−4.815 3
3	6.484 7	8	−4.065 3	13	−1.835 3
4	5.994 7	9	−3.675 3	14	−1.605 3
5	3.164 7	10	−5.095 3	15	−3.415 3

(2) 计算自相关函数 ρ_t 和偏自相关函数 $\hat{\Psi}_{kj}$

由于 AR(p) 模型的自相关函数具有托尾特性, 偏自相关函数具有截尾特性, 所以可利用自相关函数曲线和偏自相关函数曲线估计 AR(p) 模型的阶数 p。

AR(p) 模型的自相关函数 $\hat{\rho}_t$ 为

$$\begin{cases}\rho_1=\phi_1+\phi_2\rho_1+\cdots+\phi_p\rho_{p-1}\\ \rho_2=\phi_1\rho_1+\phi_2\rho_2+\cdots+\phi_p\rho_{p-2}\\ \cdots\\ \rho_p=\phi_1\rho_{p-1}+\phi_2\rho_{p-2}+\cdots+\phi_p\end{cases} \tag{13}$$

式(13) 写成矩阵形式为

$$\begin{bmatrix}\rho_1\\ \rho_2\\ \rho_3\\ \vdots\\ \rho_p\end{bmatrix}=\begin{bmatrix}1 & \rho_1 & \rho_2 & \cdots & \rho_{p-1}\\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{p-2}\\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{p-3}\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \cdots & 1\end{bmatrix}\begin{bmatrix}\phi_1\\ \phi_2\\ \phi_3\\ \vdots\\ \phi_p\end{bmatrix} \tag{14}$$

AR(p) 模型的偏自相关函数 $\hat{\Psi}_{kj}$ 为

$$\hat{\Psi}_{kj}=\begin{cases}\phi_j & j=1,2,3,\cdots,p\\ 0 & j=p+1,\cdots,k\end{cases} \tag{15}$$

利用式(14)、式(15) 计算 AR(p) 模型的自相关函数与偏自相关函数并绘制成曲线, 分别如图 1、图 2 所示。

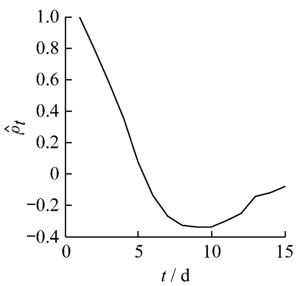

图 1　AR(p) 模型的自相关函数曲线

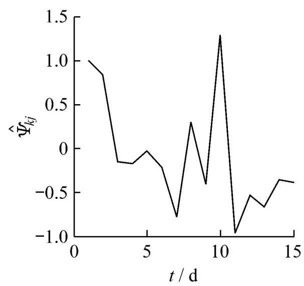

图 2　AR(p) 模型的偏自相关函数曲线

(3) 建立模型

从图 1、图 2 可看出, 自相关函数 ρ_t 随着 t 的增大而衰减, 可认为是拖尾的; 而偏自相关函数 $\hat{\Psi}_{kj}$ 在零附近波动, 可认为是截尾的。由自相关函数与偏自相关函数分别具有拖尾性和截尾性, 可初步认定选用的 AR(p) 模型是合适的, 然后进行参数估计, 即确定 p 的值。

根据式(5)、式(6), 在 AR(p) 模型中选取不同的 p, 所得到的 AIC 值不同, 当使 AIC 值最小时的 p 即为 AR(p) 模型的阶数。本文利用 Matlab 工具箱信号处理中的 Levinson 函数(LPC) 对 AR(p) 模型进行仿真[4], 按 $p=0$ 开始依次递增, 当 AR(p) 模型与原始序列较为一致时, 即满足仿真要求, 可确定 p 的值[5]。图 3 为 3 阶 LPC 估计值与原始信号比较曲线, 从图 3 可看出, 当 $p=3$ 时, 可得到与原始序列比较一致的模拟序列, 即 AR(3) 模型参数比较准确。利用 Matlab 的 result 函数求出最小自回归系数分别为 1.000 0、−0.970 6、0.030 4、0.022 9。

(4) 预测

按 AR(3) 的预测公式:

$$y_t=y_{t-1}+0.970\,6y_{t-2}+0.030\,4y_{t-3}+0.022\,9 \tag{16}$$

文章编号: 1671- 251X(2010)12- 0027- 03

基于一元线性回归的中国煤炭行业年死亡人数与年煤炭产量的预测研究

王文才, 乔旺, 王瑞智, 李刚

(内蒙古科技大学矿业工程学院, 内蒙古 包头 014010)

摘要: 为了预测我国煤炭行业未来安全形势的发展状况, 建立了我国煤炭行业年死亡人数与年煤炭产量的一元线性回归预测模型, 采用该模型对我国未来 3 年煤炭行业年死亡人数与年煤炭产量的发展趋势进行了预测。预测结果表明, 该模型具有一定的可行性, 为预测我国煤炭行业的安全形势提供了可靠的理论依据。

关键词: 煤炭行业; 死亡人数; 煤炭产量; 一元线性回归; 预测

中图分类号: **TD79** 文献标识码: **A**

Prediction Research of Annual Death Numbers and Annual Coal Production in China's Coal Industry Based on Simple Linear Regression

WANG Wen-cai, QIAO Wang, WANG Rui-zhi, LI Gang

(School of Mining Engineering of Inner Mongolia University of Science and Technology, Baotou 014010, China)

收稿日期: 2010- 08- 23

作者简介: 王文才(1963-), 男, 内蒙古伊金霍洛旗人, 教授, 博士, 硕士研究生导师, 现主要从事安全工程及矿业技术经济的教学与研究工作。E-mail: wencai88888@yahoo.com.cn

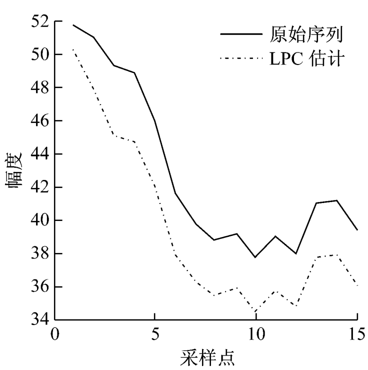

图 3 3 阶 *LPC* 估计值与原始信号比较曲线

得出 $y_{16}=38.89$, $y_{17}=38.25$, $y_{18}=39.54$, 由式(11)得出平均误差率为 4.3%, 说明瓦斯涌出量预测值比较可靠。

4 结语

时间序列预测法是工程统计中的常用方法, 利用它可建立煤矿瓦斯涌出量与深度的函数关系, 从而对矿井未开采的区域进行预测, 具有很强的实际操作性; 为了增强该方法的通用性, 需要开展广泛的实验, 更多地获得煤矿瓦斯涌出量在各种环境条件下的实测数据, 建立一套完整详细的瓦斯涌出量数据库, 进而选定条件由数据库中的实测数据对瓦斯涌出量进行仿真预测; 在建立时间序列时, 要考虑数据的准确性, 同时还要考虑煤层赋存条件及地质构造的影响, 以提高预测的准确性。

参考文献:

[1] 施式亮, 宋译, 何利文, 等. 矿井掘进工作面瓦斯涌出混沌特性判别研究[J]. 煤炭学报, 2006, 31(6): 58-62.

[2] RABINER L R. 语音数字信号处理[M]. 北京: 电子工业出版社, 1993.

[3] QUATITERI T F. 离散时间语音信号处理——原理与应用[M]. 北京: 电子工业出版社, 2004.

[4] 杨位钦, 顾岚. 时间序列分析与动态数据建模[M]. 北京: 北京工业学院出版社, 1986.

[5] 宿晶亮, 栗苹, 刘宁. AR 模型在直升机声学环境模拟中的应用[J]. 探测与控制学报, 2001, 23(2): 45-48.