

时间序列分析方法在郑州市降水量预报中的应用

吕志涛^{1,2}

(1. 河南省地质调查院, 郑州 450001; 2. 河南省地矿局第五地质勘查院, 郑州 450001)

摘要: 根据1971年-2013年郑州市的降水量资料,采用二次多项式拟合提取降水量的趋势分量,采用谐波分析法提取降水量的周期成分,利用自回归模型求解随机成分,最后将三者叠加,构建了郑州市降水量的预报模型。模型计算结果与实测数据对比可知,应用预报模型对降水量进行预报精度较高。因此利用建立的预报模型对2014年-2016三年的降水量进行了预测,为该区水资源的管理提供依据。

关键词: 降水量;时间序列分析;预报;郑州市

中图分类号: TV121.1 **文献标志码:** A **文章编号:** 1672-1683(2014)04-0035-03

Application of Time Series Analysis on the Annual Precipitation of Zhengzhou city

LVU Zhi tao^{1,2}

(1. College of Resources and Environment, North China University of Water Resources and

Electric Power, Zhengzhou 450011, China; 2. Henan Institute of Meteorological Sciences, Zhengzhou 450003, China)

Abstract: Time series analysis method has comprehensive importance and prospects in precipitation forecast. After analyzing precipitation data from 1971 to 2009 in Zhengzhou city, precipitation forecasting model was constructed by time series analysis method in the paper. When establishing forecasting model, First, the trend component of precipitation is picked up by quadratic polynomial calibration, the periodic component is extracted by spectrum analysis and the stochastic component is simulated by using autoregression model. Finally, the forecasting model is established through superposition of these components, and the method has been used to forecast the next three years precipitation in the Zhengzhou city. In comparison with model calculation results and measured data, the precipitation forecast model accuracy meets the requirements, we can use the model to forecast precipitation in the future, this paper gives the precipitation prediction results of three years of 2014~2016. The results would be helpful for water management in this region.

Key words: precipitation; time series analysis; forecast; Zhengzhou city

降水量预报方法很多,如多年平均值法、皮尔逊 ρ 型概率分布曲线等。上述降水量预报方法,虽然在某程度上反映了地区降水量分布特征,但还不能反映降水分布的本质特征。实际上,一个地区降水量的年内分布具有随机性、不确定性,然而其多年分布却存在统计规律性^[1]。人们把某一变量或指标的数值或观测值,按照其出现时间的先后次序,以相同的间隔时间排列一组数值,称为时间序列。目前,分析时间序列的演变规律是掌握系统动力学特性的重要手段。该方法易于掌握,计算工作量小,易于应用推广^[2],因此在国民经济宏观控制、区域综合发展规划、企业经营管理、市场容量预测、气象预报、水文预报、地震前兆预报、农作物病虫害预报、环境污染控制、生态平衡、天文学和海洋学等领域已

有广泛应用。

近年来,时间序列分析法在降水量预报方面取得了十分显著的效果。董晓萌等应用时间序列模型对杨凌张家岗气象站1935年-1955年的四季降水量建立了自回归模型,利用所求得的模型对1956年的四季降水量进行了预报^[3];许晓艳根据浑河流域长系列的降水量资料,运用时间序列分析原理建立了浑河流域降水量预测模型,预测该流域的各站降水量,经检验其预测的中长期结果具有较高的精度^[4]。本文在分析1971年-2013年郑州市降水量(面雨量)的年际变化特征的基础上,应用时间序列分析法对该区的降水量进行了预报,力求为该区水资源的管理提供依据。

收稿日期: 2014-05-16 修回日期: 2014-06-06 网络出版时间: 2014-06-11

网络出版地址: <http://www.cnki.net/kcms/doi/10.13476/j.cnki.nsbdtqk.2014.04.001.html>

基金项目: 河南平原地区地下水污染调查评价(淮河流域); 国土资源大调查项目(1212010634502)

作者简介: 吕志涛(1968-),男,河南禹州人,高级工程师,主要从事气候变化与水文地质工作。E-mail: lztzmd@sina.com

1 时间序列分析方法的基本原理

1.1 预报模型

时间序列分析法也称历史资料延伸预测法,通过对已观测的一系列历史数据进行曲线拟合和参数估计,掌握该序列的发展过程、方向以及趋势,从而对其进行类推或延伸,并来预测下一段时间可能达到的水平。时间序列可分解成趋势成分、周期成分和随机性成分,每一种成分都可以通过一定的数学方法来确定;将三者线性叠加,即可得到总的预报模型;经检验合格后,可用于预报^[5]。预报模型可用下式表示:

$$H(t) = X(t) + P(t) + R(t) \quad (1)$$

式中: $X(t)$ 为趋势项,反映 $H(t)$ 随时间的变化趋势; $P(t)$ 为周期项,反映 $H(t)$ 的周期性变化; $R(t)$ 为随机项,反映随机要素的影响。

1.2 趋势项分析

计算趋势成分的方法主要有两种:平滑法(移动平均法、指数平均法)和多项式拟合法。本文采用多项式拟合法进行趋势项分析,表达式为:

$$X_t = C_0 + C_1 t + C_2 t^2 + \dots + C_k t^k = \sum_{k=0}^k C_k t^k \quad (2)$$

式中: X_t 为数据的趋势项, $C_0, C_1, C_2, \dots, C_k$ 为常系数, t 为实测数据。

可采用多元回归方法确定待定系数。一般采用二次多项式 $X_t = b_0 + b_1 t + b_2 t^2$ 拟合趋势成分,应用最小二乘法确定多项式系数,即可确定 X_t 的表达式。为检验拟合结果,需在一定的显著性水平下计算趋势曲线拟合的相关系数 R ,若无最佳拟合函数则认为无趋势项或趋势项不显著。

$$R = \frac{\sqrt{\sum_{i=1}^n (x_i - \hat{x}_i)^2}}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}} \quad (3)$$

式中: n 为实测数据序列的总个数, X_t 为实测序列, \hat{X}_t 为计算序列, \bar{x} 为实测序列的平均值。 R 愈接近 1 时,表明 y 与 $x_i (i=1, 2, \dots, k)$ 的线性关系愈密切,反之愈差。

1.3 周期项分析

本文采用谐波分析法,将剩余序列 $P_i (i=1, 2, \dots, n)$ 看成是由不同周期的规则波叠加而成。因而在分离周期时,逐步分解出一些比较明显的周期波,然后叠加起来作为该时间序列的周期项,把这个周期项进行外推可用于预报。对序列 P_i 可用 L 个波叠加的形式表示其估计值为:

$$P_i = \frac{a_0}{2} + \sum_{k=1}^L [a_k \cos \frac{2\pi kt}{n} + b_k \sin \frac{2\pi kt}{n}] \quad (4)$$

式中: P_i 为序列 P_i 的估计值; L 为谐波个数,取 $n/2$ 的整数部分; K 为谐波序号, $(k=1, 2, \dots, L)$; a_k, b_k 为傅里叶系数,其计算式为:

$$\begin{cases} V_{m-2} = V_{m-1} = 0 \\ V_i = x(t) + V_{i+1} \cdot 2\cos(\omega_j) - V_{i+2} \end{cases} \quad (t = n, n-1, \dots, 2, 1) \quad (5)$$

$$a_0 = \frac{1}{n} \sum_{t=1}^n x(t) \quad \begin{cases} a_k = \frac{2}{n} [V_1 \cos(\omega_j) - V_2] \\ b_{kj} = \frac{2}{n} [V_1 \sin(\omega_j)] \end{cases} \quad (j = 1, 2, \dots, k) \quad (6)$$

通常在 L 个波中选取波动比较显著的几个谐波相加来估计 P_i ,在实际应用中只需选取前 6 个显著谐波即可满足精度要求。若满足式(7),则认为第 k 个波显著,否则不显著。

$$s_k^2 = a_k^2 + b_k^2 > 4s^2 \frac{\ln \frac{k}{n}}{n} \quad (7)$$

式中: α 为检验的显著性水平(一般取 5%); s^2 为系列的方差。

1.4 随机项分析

消除趋势项和近似周期项后的剩余序列为平稳随机序列项 $R(t)$, $R(t) = H(t) - X(t) - P(t)$,可直接对其用自回归模型求解。考虑到 $R(t)$ 与 t 时刻的取值和它前 1 个到 p 个时间间隔 $t-1, t-2, \dots, t-p$ 的取值有关,其自回归模型为:

$$R_t = \Phi_0 + \Phi_1 r_{t-1} + \Phi_2 r_{t-2} + \dots + \Phi_p r_{t-p} \quad (8)$$

式中: R_t 为估计值; Φ_i 为模型自回归系数; p 为模型阶数,用 $AIC(p)$ 准则确定。

当 $AIC(p)$ 值最小时,所对应的 p 值为最佳阶数。

$$AIC(p) = n \ln \sigma_p^2 + 2p \quad (9)$$

式中: n 为序列数据总个数; σ_p^2 为 $AR(p)$ 残差的方差。

1.5 预报模型的建立和检验

将上述趋势分量、周期分量、随机分量线性叠加,即可得到降水量的总预测模型^[6]:

$$H_t = X_t + P_t + R_t \quad (10)$$

建立的模型要经过检验才能判定其是否合理,是否合格。本文采用后验差方法来进行检验模型的精度,即通过后验差比值 c 和小误差频率 P 来检验。设动态样本中前 n 个数据的标准差为 S_2 ,后验数据残差为 e_j ,残差均值为 \bar{e} ,则后验差比值 c 及小误差频率 p 的计算公式为:

$$c = \frac{S_2}{s_1} \quad (11)$$

$$p = \{ |e_j - \bar{e}| < 0.6745s_1 \} \quad (12)$$

表 1 后验预测法预测效果表^[9]

Table 1 Effects of posterior prediction

预测效果	好	合格	勉强	不合格
p	> 0.95	> 0.80	> 0.70	≤ 0.70
c	< 0.35	< 0.50	< 0.65	≥ 0.65

如果 p, c 值都在允许范围内见表 1,则模型可用于计算预报值,否则需要对模型进行检查、分析和重新调参。

2 时间序列分析在降水预报中的应用

本文利用的郑州市 1971 年-2013 年降水量统计资料来源于郑州国家气候观象台。其中 1971 年-2005 年 35 个降水量数据用于建立数学模型,2006 年-2013 年 8 个降水量数据用于检验数学模型的精度。

2.1 趋势项的提取

采用二次多项式拟合法,根据前文的趋势项提取方法,可得趋势项函数为:

$$\hat{x}_t = 726.6591t - 13.2598t^2 + 0.3567t^3 \quad (13)$$

为检验趋势拟合效果,计算趋势曲线拟合的相关系数

R :

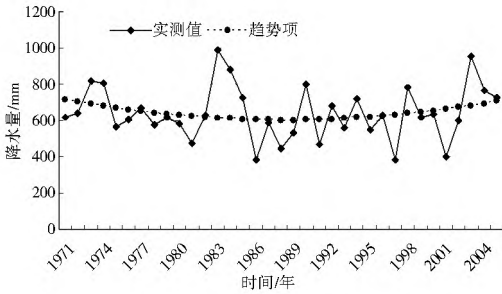


图1 趋势项计算值与实测值比较

Fig.1 Comparison of Trend term calculated and measured values

$$s_2 = \sqrt{\frac{1}{N-5} \sum_{j=1}^5 (e_j - \bar{e})^2} = 18.5718$$

$$c = \frac{s_2}{s_1} = 0.3042$$

$$p = \{ |e_j - \bar{e}| < 0.6745s_1 \} = 1$$

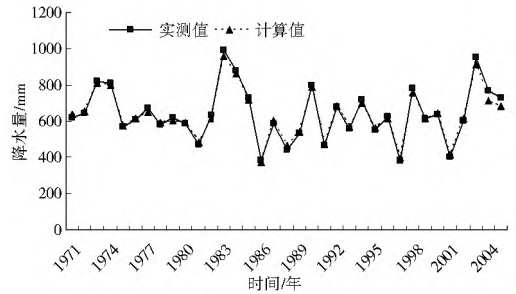


图2 降水量实测值与计算值比较

Fig.2 Comparison of precipitation calculated and measured values

$$R = \frac{\sqrt{S_{\text{回}}}}{\sqrt{S_{\text{总}}}} = \frac{1}{\sqrt{1 + \frac{\sum_{t=1}^n (x_t - \hat{x}_t)^2}{\sum_{t=1}^n (x_t - \bar{x})^2}}} = 0.1975 \quad (14)$$

显然,趋势项并不明显,可以忽略。

2.2 周期项的提取

取 $L = (35 - 1)/2 = 17$ 。计算系数 $a_0, a_k, b_k (k = 1, 2, \dots, 17)$, 其中 $a_0 = -0.014705927$ 。根据周期成分系数计算结果,当 $k = 5, 7, 16$ 时,波显著。则模型为:

$$\hat{p}_t = \frac{a_0}{2} + a_5 \cos\left(\frac{2\pi}{35} \cdot 5t\right) + b_5 \sin\left(\frac{2\pi}{35} \cdot 5t\right) + a_7 \cos\left(\frac{2\pi}{35} \cdot 7t\right) + b_7 \sin\left(\frac{2\pi}{35} \cdot 7t\right) + a_{16} \cos\left(\frac{2\pi}{35} \cdot 16t\right) + b_{16} \sin\left(\frac{2\pi}{35} \cdot 16t\right) \quad (15)$$

2.3 平稳随机项的拟合

平稳随机项可以通过自回归模型 AR(P) 来拟合。用 AIC(p) 准则确定阶数,最终确定 $p = 6$ 。平稳随机项为:

$$r(t) = h(t) - p(t) \quad (16)$$

$$R_t = \Phi_0 + \Phi_1 r_{t-1} + \Phi_2 r_{t-2} + \Phi_3 r_{t-3} + \Phi_4 r_{t-4} + \Phi_5 r_{t-5} + \Phi_6 r_{t-6} \\ = 2536.2305 - 0.0868r_{t-1} - 0.167r_{t-2} - 1.5880r_{t-3} - 0.1890r_{t-4} - 0.463r_{t-5} - 0.9828r_{t-6} \quad (17)$$

因此,把周期项和平稳随机项项叠加起来,得到降水量的预报方程为:

$$h_t = a_0 + a_5 \cos\left(\frac{2\pi}{n} \cdot 5t\right) + b_5 \sin\left(\frac{2\pi}{n} \cdot 5t\right) + a_7 \cos\left(\frac{2\pi}{n} \cdot 7t\right) + b_7 \sin\left(\frac{2\pi}{n} \cdot 7t\right) + a_{16} \cos\left(\frac{2\pi}{n} \cdot 16t\right) + b_{16} \sin\left(\frac{2\pi}{n} \cdot 16t\right) + \Phi_0 + \Phi_1 r_{t-1} + \Phi_2 r_{t-2} + \Phi_3 r_{t-3} + \Phi_4 r_{t-4} + \Phi_5 r_{t-5} + \Phi_6 r_{t-6} \quad (18)$$

从图2中模型计算值和实测值的比较可知,拟合效果较好,说明建立的模型较符合实际。

2.4 模型检验及预报

令表示后验预测的残差,即:

$$e_j = h_{n+1} - h_{n+j} \quad (j = 1, 2, 3, 4, 5)$$

设动态样本中1971年-2005年35个数据的标准差为,2006年-2013年8个数据的后验预测残差的标准差为 s_2 。

根据公式计算得:

$$s_1 = \sqrt{\frac{1}{N-4} \sum_{i=1}^{34} (h_i - \bar{h})^2} = 61.0572$$

表2 后验预测误差表

Table 2 Error of posterior prediction

n	实测值/mm	计算值/mm	e_j	$ e_j - \bar{e} $
36	692	692.23	-0.23	2.29
37	596	607.75	-11.75	13.81
38	658	643.51	14.49	12.43
39	764	738.33	25.67	23.61
40	600.3	589.15	11.15	9.09
41	706.5	691.37	15.13	13.07
42	498.7	524.31	-25.61	27.67
43	390.5	417.29	-26.79	28.85

查表1可知,后验预测效果等级为“好”,可以进行降水量的预报。2014年-2016年的降水量预测值见表3。

表3 降水量预报结果

Table 3 Prediction of precipitation

时间/年	2014	2015	2016
降水量/mm	687.264	705.905	547.880

3 结语

本文运用时间序列分析方法,通过对降水量趋势、周期项和随机项的提取,建立了郑州市降水量和时间之间的数学模型。在模型检验表明,建立的数学模型的预报精度较高,但随着预测周期的增加,预报误差也在逐渐增大。因此为使预测更准确,预报时间长度不宜太长。

参考文献(References):

- [1] 汤成友,官学文,张世明.现代中长期水文预报方法及其应用[M].北京:中国水利水电出版社,2008.(TANG Chengyou, GUANG Xuewen, ZHANG Shiming. Modern Mid to long term Hydrological Forecasting Method and its Application [M]. China Water Power Press. 2008. (in Chinese))
- [2] Martinez J C, Maleski J J, Miller F M. Trends in Precipitation and Temperature in Florida[J]. USA J Hydrol, 2012, 452-453, 259-281.

(下转第56页)

- in Sediments From the North Carolina Continental Slope[J]. *Geochimica et Cosmochimica Acta*, 1999, (63): 427-448.
- [4] Covelli S, Faganeli J, Horvat M, et al. Porewater Distribution and Benthic Flux Measurements of Mercury and Methylmercury in the Gulf of Trieste (Northern Adriatic Sea) [J]. *Estuarine, Coastal and Shelf Science*, 1999, (48): 415-428.
- [5] Choe K Y, Gill G A, Lehman R D, et al. Sediment Water Exchange of Total Mercury and Methylmercury in the San Francisco Bay Delta[J]. *Limnology and Oceanography*, 2004, (49): 1512-1527.
- [6] Azadeh Taghiania Hejazi, Basavarajappa H T, Karbassi A R, et al. Heavy Metal Pollution in Water and Sediments in the Kabini River, Karnataka [J]. *India. Environ Monit Assess.* 2011, (182): 1-13.
- [7] Allen H E, D J Hansen. The Importance of Trace Metal Speciation to Water Quality Criteria [J]. *Water Environment Research*, 1996, 68(1): 42-54.
- [8] Hicks, M B, Scott K J, Hansen, D. J, et al. Toxicity of Cadmium in Sediments: the Role of Acid Volatile Sulfide [J]. *Environmental Toxicology and Chemistry*, 1990, 9(1): 1487-1502.
- [9] Kim Yeongkyoo, Kim Byoung Ki, Kim Kangjoo. Distribution and Speciation of Heavy Metals and Their Sources in Kumho River Sediment, Korea [J]. *Environ Earth Sci*, 2010, 60(5): 943-952.
- [10] Schacht U, Wallmann K, Kutterolf S, et al. Volcanogenic Sediment-Seawater Interactions and the Geochemistry of Pore Waters [J]. *Chemical Geology*, 2008, 249(3-4): 321-338.
- [11] Haese R R, Hensen C, Lange G J. Pore Water Geochemistry of Eastern Mediterranean Mud Volcanoes: Implications for Fluid Transport and Fluid Origin [J]. *Marine Geology*, 2006, 225(1-4): 191-208.
- [12] 薛亚莉. 陕西省渭河流域水资源供需形势分析 [J]. *人民黄河*, 2009, (11): 56-57. (XUE Ya li. Supply and Demand Situation Analysis About Shaanxi Weihe River water [J]. *Yellow River*, 2009, (11): 56-57. (in Chinese))
- [13] 吴立凯. 陕西省渭河流域水质调查及监测措施 [J]. *陕西水利*, 2008, (4): 19-20. (WU Li kai. Water Quality Surveys and Monitoring Measures on Shaanxi Weihe River, Shaanxi Water Resources, 2008, (4): 19-20. (in Chinese))
- [14] 国家环境保护总局. 中华人民共和国地表水环境质量标准 (GB 3838-2002) (SEPA. People's Republic of Surface Water Quality Standards. (GB 3838-2002) (in Chinese))
- [15] Lars Hakanson. An Ecological Risk Index for Aquatic Pollution Control. a Sedimentological Approach. *Water Research*, 1980, (8): 975-1001.

(上接第37页)

- [3] 董晓萌, 罗凤娟, 郭满才, 等. 时间序列的自回归模型在杨凌地区降水量预报中的应用 [J]. *中国农学通报*, 2007, 23(11): 403-407. (DONG Xiaomeng, LUO Fengjuan, GUO Mancai et al. The Autoregressive Model of Time Sequence Apply to Rainfall Forecast in Yangling Area [J]. *Chinese Agricultural Science Bulletin*, 2007, 23(11): 403-407. (in Chinese))
- [4] 许晓艳. 基于时间序列的浑河流域降水量预测模型 [J]. *水土保持应用技术*, 2011(12): 26-27. (XU Xiaoyan. Prediction of Precipitation Model of Hunhe River Based on Time Series [J]. *Technology of Soil and Water Conservation*, 2011, (12): 26-27. (in Chinese))
- [5] 刘杨, 徐淑琴, 董丽丽, 等. 时间序列模型在查哈阳农场降雨量预测中应用 [J]. *黑龙江水利科技*, 2012, 40(2): 5-7. (LIU Yang, XU Shuqin, DONG Lili, et al. Application of Time Series Model in Chahayang Farm in Rainfall Prediction [J]. *Heilongjiang Science and Technology of Water Conservancy*, 2012, 40(2): 5-7. (in Chinese))
- [6] Wang S, Feng J, Liu G. Application of Seasonal Time Series Model in the Precipitation Forecast [J]. *Mathematical and Computer Modelling*, 2013, 58(3-4): 677-683.
- [7] 何书元. 应用时间序列分析 [M]. 北京: 北京大学出版社, 2004. (HE Shuyuan. Application of time series analysis [M]. Beijing. Peking University press, 2004. (in Chinese))
- [8] Sayemuzzaman M, Jha M. Seasonal and Annual Precipitation Time Series Trend Analysis in North Carolina [J]. *United States. Atmospheric Research*, 2014(137): 183-194.
- [9] Vilibic I, Leder N. Long-term Variations in the Mediterranean Sea level Calculated by Spectral Analysis [J]. *Oceanographic Literature Review*, 1997, 44(5): 422.