

基于 ARIMA 模型的固原月降水量预测

范小明, 杨长军, 李淑珍, 张成军 (宁夏固原市气象局, 宁夏固原 756000)

摘要 采用 ARIMA 模型来拟合 1957~2013 年 5~10 月的月降水量并给出预测值。结果表明, ARIMA(0, 2, 1) 能较好地拟合 5~10 月的月降水量数据, 准确率较目前业务使用的阴阳历方法有所提高, 可作为月预报的依据。

关键词 自回归积分滑动平均 (ARIMA) 模型; 月降水量; 预测; 固原

中图分类号 S161.6 **文献标识码** A **文章编号** 0517-6611(2013)01-00204-02

Monthly Precipitation Forecasting in Guyuan City Based on ARIMA Model

FAN Xiao-ming et al (Guyuan Meteorological Bureau, Guyuan, Ningxia 756000)

Abstract ARIMA model was adopted to fit monthly precipitation and draw predict value from May to Oct. during 1957-2013. The results showed that ARIMA(0, 2, 1) could be used as monthly forecasting basis for its higher accuracy rate than lunar-solar superposition method.

Key words ARIMA model; Monthly precipitation; Forecasting; Guyuan

研究固原地区月降水量的变化特点并对中长期(旬、月)降水量进行合理预测, 将有利于合理开发利用当地水资源和防灾减灾工作。有研究表明, 该地区降水量变化存在 3、11~13 及 21 年等若干个周期的叠加^[1]。然而, 气象条件复杂多变及难以确定其中长期主要影响因子的制约, 某地的降水量变化趋势存在随机性与不确定性, 目前还难以通过数值模拟来准确预测旬、月等中长期时段的降水量, 同时传统的阴阳历叠加、概率法、外推法等预测模型准确性较低。积分自回归滑动平均模型 (AutoRegressive Integrated Moving Average Model, 简称 ARIMA) 是统计学家博克斯 (Box) 和詹金斯 (Jenkins) 于 70 年代初提出的一著名时间序列预测方法, 所以又称为 box-jenkins 模型、博克斯-詹金斯法^[2]。与平稳时间序列、周期时间序列模型相比, 该建模方法在数学上更为完善, 预测精度更高。因此, 笔者以固原 1957~2012 年 5~10 月降水量为基础, 在对月降水序列进行差分转换的基础上, 通过参数估计和模型检验等步骤, 采用数据处理软件 SPSS19 经过多次拟合试验, 建立固原市 5~10 月月降水量模型。

1 资料与方法

1.1 资料来源 选取固原 1957~2012 年 5~10 月降水量资料, 数据来源于固原基准站。

1.2 分析方法

1.2.1 ARIMA(p, d, q) 模型 ARIMA 模型常用于平稳序列或通过差分而平稳的时间序列分析, ARIMA 是时间序列分析中常用的模型, 也称为带差分的自回归移动平均模型, 可对含有周期成分的时间序列数据进行分析^[3-4]。自回归滑动平均模型 ARMA(p, q) 一般形式为: $X_t - \varphi_1 X_{t-1} - \dots - \varphi_p X_{t-p} = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$ 或 $\varphi(B)X_t = \theta(B)a_t$, ARMA(p, q) 模型由 AR(p) 及滑动平均 MA(q) 混合而成。其中, 等式左边是模型的自回归部分, 非负整数 p 称为自回归阶数, $\varphi(B)$ 称为自回归系数, X_t 是气象要素序列; 等式右边是模型的滑

动平均部分, 非负整数 q 称为滑动平均阶数, $\theta(B)$ 称为滑动平均系数, a_t 是白噪声序列, 是一列相互之间无关, 其均值为 0, 方差为 δ^2 的随机变量序列。p、q 分别是偏自相关函数值和自相关函数值显著不为零的最高阶数。可以看出, 当 p=0 时, 模型是纯滑动平均模型, 记为 ARMA(0, q); 当 q=0 时, 模型是纯自回归模型, 记为 ARMA(p, 0)。ARMA(p, q) 模型可用较少的参数对序列进行较好的拟合, 其自相关和偏自相关函数均呈现拖尾性。ARMA 模型仅适合对平稳序列的分析。实际应用中许多气象要素时间序列并非平稳序列, 不能直接采用 ARMA 模型, 但通常这些序列经过变换处理后变为平稳序列。对它们的分析一般采用加法模型或差分法来处理, 当使用差分法也就是自回归积分滑动平均 ARIMA(p, d, q) 模型, 即 $\varphi(B)X_t = (1-B)^d X_t = \theta(B)a_t$, 式中, d 是差分阶数, B 是后移算子, 即 $BX_t = X_{t-1}$ 。AR(p) 模型的偏自相关函数在 p 阶之后应为零, 称其具有截尾性; AR(p) 模型的自相关函数不能在某一步之后为零 (截尾), 而是按指数衰减 (或呈正弦波形式), 称其具有拖尾性。MA(q) 模型的自相关函数在 q 阶之后应为零, 称其具有截尾性; MA(q) 模型的偏自相关函数不能在某一步之后为零 (截尾), 而是按指数衰减 (或呈正弦波形式), 称其具有拖尾性^[3-4]。利用 ARIMA 模型作月降水量预测, 可在 IBM SPSS Statistics 软件上进行。

1.2.2 建立 ARIMA 模型的一般步骤 首先通过差分或其他变换, 使时间序列满足平稳性的要求, 差分转换是最有效的方法; 其次模型识别, 主要是利用自相关函数 ACF、偏自相关函数 PACF 和正态化 BIC 等序列估计模型的大致类型, 并给出几个初步 ARMA(p, q) 模型以待进一步验证和完善; 然后阶数估计和模型诊断, 以预报时残差方差最小为原则, 结合巴叶斯信息准则 (BIC) 对模型的 p、q 阶数作诊断分析, 最终选择合理的模型; 最后预测, 利用最优模型对序列的未来取值或走势进行预测。在以上步骤中, 模型识别、参数估计及模型诊断的过程通常均是不断反馈、逐渐完善的过程。

2 结果与分析

先以 1957~2010 年 5 月降水量 (表 1) 为例, 拟合并计算 2011 年 5 月降水量预测值。最后利用其他各月降水量数据

作者简介 范小明 (1968-), 男, 山西平遥人, 高级工程师, 从事中短期预报工作, E-mail: gyjfxm@163.com。

收稿日期 2012-10-31

给出 2008 ~ 2012 年 5 ~ 10 月各月降水量预测值。

表 1 1957 ~ 2010 年固原基准站 5 月降水量 mm

年份	降水量	年份	降水量	年份	降水量
1957	89.4	1975	47.4	1993	54.3
1958	64.7	1976	12.3	1994	4.1
1959	42.3	1977	16.2	1995	11.7
1960	57.0	1978	70.9	1996	12.0
1961	47.0	1979	18.9	1997	15.7
1962	33.5	1980	32.2	1998	53.7
1963	77.7	1981	2.2	1999	69.4
1964	55.2	1982	40.1	2000	20.2
1965	24.7	1983	72.9	2001	7.8
1966	18.3	1984	86.5	2002	40.5
1967	83.8	1985	86.3	2003	69.3
1968	13.9	1986	36.6	2004	44.6
1969	37.3	1987	40.7	2005	59.9
1970	45.5	1988	76.3	2006	58.2
1971	48.1	1989	20.1	2007	7.0
1972	34.3	1990	93.5	2008	6.4
1973	52.4	1991	88.0	2009	23.3
1974	77.9	1992	20.5	2010	63.3

2.1 最优模型的建立 处理非平稳的时间序列时,可以先建立一个包含趋势成分的模型,对由此初步模型得到的残差项再使用 ARMA (p,q)模型来拟合。ARIMA(p,d,q)模型中的各阶数 p,d,q 不大于 2,经 SPSS 软件试验, $d=2$ 时原序列变为平稳序列。建立 ARIMA(p,d,q)模型,就是对差分后的序列建立 ARMA(p,q)模型。利用 SPSS 软件,当 $p=0,1,2$ 及 $q=0,1,2$ 时得到逐个 ARIMA($p,2,q$)模型的各个参数,最后确定 ARIMA (0,2,1)作为原序列预测的最优模型。从 ARIMA (0,2,1)模型的参数估计结果中 t 统计量的显著性 (Sig. =0.214)可以看出,此时该参数非常显著地接近 0;平稳 R 方为正值且小于 1,表示当前模型优于基本均值模型。

由残差序列的自相关 (ACF) 图形和偏自相关图形 (PACF) (图 1) 可见,2 个图形没有显著的趋势特征 (截尾),均具有拖尾性,可以判断该模型是较为恰当的。

2.2 月降水量的预测分析 利用 ARIMA (0,2,1) 模型对 2011 年 5 月降水量进行预测,结果发现其 5 月降水量为 62.8 mm。ARIMA (0,2,1)模型拟合 1957 ~ 2010 年逐年 5 月降水量 (图 2) 可见,观测线与拟合线总趋势较为接近,但拟合线较观测线略滞后。

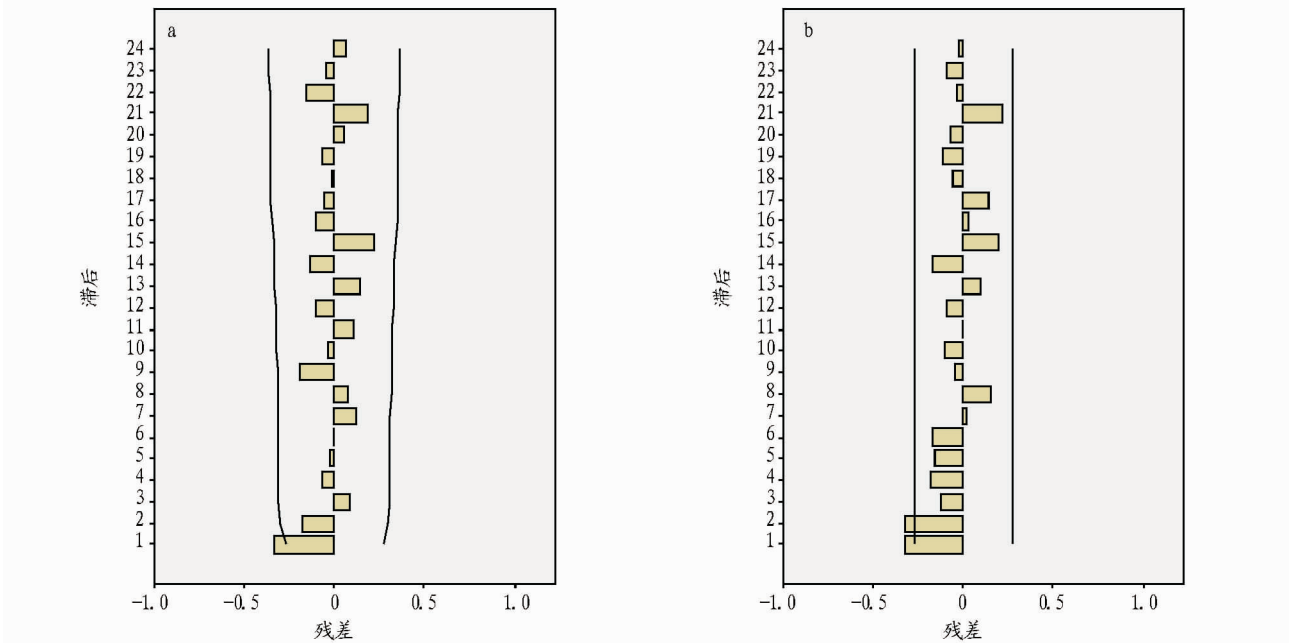


图 1 ARIMA(0,2,1)模型残差的自相关(a)和偏相关(b)

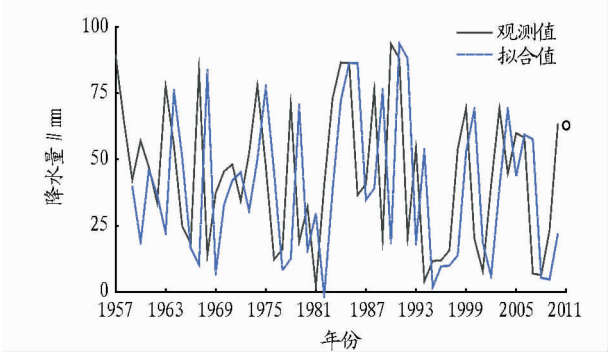


图 2 ARIMA 模型对 1957 ~ 2010 年 5 月降水量的拟合结果

按照上述方法,对 2008 ~ 2013 年逐年 5 ~ 10 月各月降水量进行预测,并给出实际观测值 (表 2)。与目前业务应用的阴阳历方法比较,以距平正、负符号 (预测值大于月平均值为正,预测值小于月平均值为负) 作为评分标准,经统计,表中 29 次月预报正确 11 次,错 18 次,正确率 38%;而 ARIMA 方法正确 16 次,错 13 次,正确率 55%,预报准确率较原方法有所提高。

3 小结与讨论

(1) 采用 ARIMA 模型预测 1957 ~ 2013 年 5 ~ 10 月的月降水量,结果发现 ARIMA (0,2,1) 能较好地拟合月降水量 (下转第 224 页)

-16.9和-13.7℃。2012 年河流开始流冰时间为 1 月 23 日,河流流冰终止时间为 1 月 30 日。

3 物候期观察时应注意事项

3.1 选择物候观测点和物候观测对象

3.1.1 选择物候观测点。在进行物候观测以前,首先应选定观测点。观测点要符合以下 2 项原则:①地点要稳定,可以进行多年观测,不必移动。因为在一个固定地点进行观测的年代越长,记录的物候资料就越宝贵。因此,没有特殊原因,最好不要更换观测地点。②观测地点要有代表性,必须考虑到地形、土壤、植被等情况,尽可能选在平坦或相当开阔的地方。物候观测点选定之后,要将地名、生境、海拔、地形(平地、山地、凹地、坡地等)、位置(在建筑物的哪边,距离建筑物多远)和土壤性质等详细记载下来,作为档案保存。

3.1.2 选择物候观测对象。木本植物选择当地常见、分布较广、指示性强、对季节变化反映明显、与农业生产关系密切、群众常用的为主,且是裸地栽培或野生者,生长健壮的达到开花结实 3 年以上的植株 3~5 株进行观测。如山地可选等高梯田内的植株观察;如株间差异大时,应按类选定代表树对全株进行观察,并进行点(代表株)面(生产果园)结合观测,才能确定各物候期。草本植物单株间物候期的差异较大,为避免偶然性,最好选择彼此靠近的数十株作为观测对象。

3.2 观测项目、记载方法的标准和要求 各物候期的观测项目、记载方法要有统一的标准和要求,才能进行比较。对每一物候期的起止日期必须记清。物候观测要随看随记,不要凭记忆在事后补记。

3.3 每个物候期观测的时间,应根据不同时期而定 春夏

(上接第 205 页)

数据。

表 2 2008~2013 年 5~10 月观测、预测值

项目	5 月	6 月	7 月	8 月	9 月	10 月
2008 观测值	6.4	41.9	33.6	64.4	151.5	21.7
2008 预测值	5.4	43.4	59.5	64.7	69.5	72.6
2009 观测值	23.3	22.3	68.7	141.7	23.5	18.6
2009 预测值	4.8	42.0	32.0	64.4	153.9	21.8
2010 观测值	63.3	30.6	116.3	57.9	66.3	60.3
2010 预测值	22.0	22.0	67.8	143.2	23.4	18.6
2011 观测值	25.5	28.6	46.0	84.6	95.6	25.2
2011 预测值	62.8	30.5	116.4	57.8	67.0	61.1
2012 观测值	28.9	82.5	31.8	136.2	95.6	
2012 预测值	28.5	23.7	31.8	66.5	98.0	21.0
2013 预测值	27.8	83.3	30.3	137.5	96.8	25.5

(2)许多气象要素变量属于随机序列,分析其时空分布特点,发现具有一定周期性,但不是很显著。中长期(旬、月)降水量的定量预测是目前预报业务中的难点。笔者经过试

季正是各种植物的萌发、展叶、开花繁殖时期,各种物候现象天天不同,最好每天观测一次,如果时间不允许,也要做到隔日观测。到了秋季,可以隔日或 3 d 观测一次。冬季的初冬和冬末还须进行观测,不可漏掉,到了隆冬季节则必须注意观测河流开始结冰、流冰和流冰终止日期。

3.4 观测人员要固定 观测人员不宜时常变更,更不能几个人轮班观测。因为物候现象时刻变动,由一个人去观测,前后才有联系,如果轮班观测,各人对物候期特征标志的认识未必相同,这样前后的记录就会没有连贯性。

4 结论

在全球气候变暖为主要特征的气候变化背景下,中卫市沙坡头区的区域气候特征也受到比较明显影响。分析宁夏中卫市沙坡头区 1981~2011 年的木本、草本和气象水文观测资料发现,31 年来中卫市沙坡头区的物候期随着气候的变暖,春季萌芽期比 20 世纪 80 年代提前 3~4 d,秋季黄枯期推迟 3~4 d,全年植物生长季延长 6~8 d。同时也对一些观测技术进行探讨,总结了一些气象水文现象的气候指标和物候观测的注意事项。

参考文献

- [1] 姚永刚,王玲,张立,等.北京市八达岭风景区树木物候研究[J].科技园地,2006(1):7-9.
- [2] 徐雨晴,陆佩玲,于强.气候变化对植物物候影响的研究进展[J].资源科学,2004,26(1):129-133.
- [3] 杨国栋,陈效速.论自然景观的季节节奏[J].生态学报,1998,18(3):234-238.
- [4] 国家气象局.农业气象观测规范[M].北京:气象出版社,1993:133-144.

验若干个 ARIMA 模型,认为其对中长期降水量具有可预报性。通过差分等数据转换方法将非平稳序列转换为平稳随机序列,以满足预测的前提,模型使用了残差序列来提高预测的精度。ARIMA 模型假定时间序列未来的变化与过去的一致,因此它适合作中长期(旬、月)降水量的预测。相对于其他方法,如阴阳历叠加法,ARIMA 模型能提高预报旬、月降水量的准确率,为做好中长期预报预测提供重要的量化支持。

参考文献

- [1] 白肇烨,徐国昌.中国西北天气[M].北京:气象出版社,1986:107-112.
- [2] BOX G E P,JENKINS G M. Time series analysis:Forecasting and control[M]. San Francisco:San Francisco Press,1978.
- [3] 黄嘉佑.气象统计分析与预报方法[M].3 版.北京:气象出版社,2004:208-210.
- [4] 刘贤赵,邵金花.烟台地区降水量的 ARIMA 随机模型研究[J].数学的实践与认识,2006,6(8):8-12.