

金丽娜, 翟园, 曲静. 西安市各区县气温预测统计方法对比分析 [J]. 陕西气象, 2016 (4): 31-35.

文章编号: 1006-4354 (2016) 04-0031-05

# 西安市各区县气温预测统计方法对比分析

金丽娜, 翟园, 曲静

(西安市气象局, 西安 710016)

**摘要:** 利用西安市 7 个国家一般气象站 1971—2000 年的逐月平均气温, 依次选用最优气候值法、多元线性回归预测法、自回归移动平均法 (ARIMA) 和灰色系统理论法进行回报检验, 模拟出 7 个站点 2001—2013 年的月、季、年平均气温, 再对其进行 PS 评分和异常级分析, 结果显示: 季、年平均气温预测以最优气候值法、灰色系统理论法较为适用, 月平均气温预测中各种方法在时空序列上各有优劣, 多元线性回归预测法应适当避免使用; 从时间尺度来看, 春、夏、秋和年平均气温预测以最优气候值法最优, 冬季平均气温预测以自回归移动平均法最优, 月平均气温预测 3—8 月以最优气候值法最优, 1—2 月、9—12 月以自回归移动平均法最优; 从空间尺度来看, 靠近山区及平原区 (周至、户县、市区、高陵) 月、季、年平均气温预测以灰色系统法最优, 山区与平原区过度区域 (临潼、蓝田、长安) 月、季、年平均气温预测以最优气候值法最优。

**关键词:** 西安区县; 气温预测; PS 评分; 异常级

**中图分类号:** P457.3

**文献标识码:** A

在众多气象要素场中, 统计学方法较多的应用在气温预测上<sup>[1]</sup>, 国内许多地市开展了适用于本地的气温预测研究。如陈杰<sup>[2]</sup>利用最优气候均态模型, 以集宁月平均气温历史资料为依据进行预测检验, 预测准确率较高; 叶晓波<sup>[3]</sup>运用时间序列 ARIMA 模型的建模方法对云南省楚雄市年平均气温进行了分析研究; 唐湘玲<sup>[4]</sup>根据灰色系统理论, 用 GM (1, 1) 方法对阿克苏地区 10 个气象站进行建模和预测; 汪春辉<sup>[5]</sup>采用偏最小二乘回归建立模型, 并进行了改进, 预测能力较

好; 王永弟<sup>[6]</sup>将模糊时间序列模型引入短期气候预测, 并与多种模型进行了精度比较和分析, 结果表明各项精度评定指标优良, 具有一定的实用价值。为了提高西安市区县气温预测准确率, 本研究选取四种较常见的统计学预测方法, 利用历史资料建立模型, 分析研究适用于各区县的月、季、年预测方法, 以期更好的服务于各区县短期气候预测。

## 1 资料

选用的资料为西安 7 个国家一般气象站 (西

**收稿日期:** 2015-09-08

**作者简介:** 金丽娜 (1981—), 女, 汉族, 陕西西安人, 工程师, 硕士, 主要从事气候变化研究。

**基金项目:** 陕西省气象局预报员专项 (2015Y-19)

东省雷暴时空变化与气候变暖的关系 [J]. 广东气象, 2013, 35 (4): 25-28.

[9] 易燕明, 杨兆礼, 万齐林, 等. 近 50 年广东省雷暴、闪电时空变化特征的研究 [J]. 热带气象学报, 2006, 22 (6): 539-546.

[10] 魏凤英. 现代气候统计诊断与预测技术 [M].

北京: 气象出版社, 2007.

[11] 易燕明, 杨兆礼, 万齐林. 广州市闪电密度特征分析 [J]. 资源科学, 2006, 28 (1): 151-156.

[12] 余功梅. 华南地区近 40 年降水的气候特征 [J]. 热带气象学报, 1996, 12 (3): 252-256.

安市区、周至、户县、长安、蓝田、临潼、高陵) 1971—2000 年的逐月平均气温, 数据提取于地面报表 A 文件, 均经过了质量控制。

## 2 研究方法

依次选用最优气候值法、多线性回归预测法、自回归移动平均法 (ARIMA)、灰色系统理论法进行建模, 模拟出 7 个区县 2001—2013 年的月、季、年平均气温; 再对各模型的试报月进行 PS 评分检验及异常级分析。

### 2.1 最优气候值法

最优气候值法是美国气候中心用于制作地面温度季度预报的一种常用方法, 这种方法认为气候变化具有阶段性、周期性、持续性、转折性等内在规律, 用多年周期的气候平均值作为来年该要素的估计值的方法。其优点在于模型简单便于计算, 且具有低通滤波的功能, 用它可以发现气候变率的时间尺度, 并能识别联合信号, 预报效果较为准确<sup>[7-8]</sup>。

设某地区某要素的年、季或月要素序列为  $x_i$  ( $i=1, 2, \dots, n$ ), 构造序列

$$\bar{x}_i = \frac{1}{k} \sum_{i=1}^k x_i. \quad (1)$$

式中,  $k=1, 2, \dots, 30$ ; 表示计算平均值的时段长度。用  $k$  年的均值作为下一年的预报值, 以预报值与实况值的绝对误差最小为标准, 从统计上得出每年最优的平均年数, 然后取最优平均年数出现频率最多的年数, 作为最优气候值应取的年数, 并依此制作下一年的预报。现以  $n=30$  预测西安市区 2014 年 1 月平均气温举例将模型加以说明。

用西安市区 1971—2000 年 30 a 的资料作为统计样本, 用 2001—2013 年 13 a 的资料作为试验样本。

①分别计算 1, 2,  $\dots$ , 30 a 的 1 月气温平均值: 以  $x_{30}, x_{29}, \dots, x_2, x_1$  表示。 $x_{30}$  为 1971—2000 年共 30 a 的 1 月气温平均值,  $x_{29}$  为 1972—2000 年共 29 a 的 1 月气温平均值,  $x_1$  为 2000 年的 1 月气温值。

②预报 2001 年 1 月气温, 把以上 30 个平均值, 分别看作 2001 年 1 月的预报值, 再以预报值

与实况值的绝对误差最小为标准, 求出试验预报的最优平均年数。

③对 2001—2013 年 1 月的平均气温进行逐年试验预报, 得出每一个试报年 1 月的最优平均年数。

④以最优平均年数出现的频率高和误差小为原则, 确定 1 月平均气温预报的最优平均年数为 2 a, 以此方法确定出西安市 7 个区县的逐月、季、年平均气温预报的最优平均年数。

### 2.2 多元线性回归预测法

多元线性回归分析, 是描述因变量  $y$  与两个或两个以上自变量  $x_j$  ( $j=1, 2, 3, \dots, k$ ) 之间线性关系的一种方法, 应用在预测中时常称为多元线性回归预测方法。假设预测月的气温  $y$  (因变量) 与  $k$  个变量 (自变量)  $x_1, x_2, \dots, x_k$  之间存在线性关系, 那么它们之间的线性回归模型可表述为

$$y = a + b_1 x_1 + b_2 x_2 + \dots + b_k x_k. \quad (2)$$

在预测中, 对多变量进行分析时, 找出影响因变量的主要自变量, 舍弃与因变量关系不大的自变量, 既能提高预测的精度, 又可使预测简化而避免过于繁琐, 因此在预测前先选用相关矩阵法进行自变量的筛选。

本研究建模采用各区县 1971—2000 年的月、季、年平均气温, 模型检验采用 2001—2013 年的月、季、年平均气温, 使用包括预测月前 10 个月的平均气温资料对该温度进行预测。首先利用相关矩阵法进行自变量的选择, 求解各月、季、年的回归方程, 依次预测出 7 个区县 13 年的逐月、季、年平均气温。

### 2.3 自回归移动平均法 (ARIMA)

自回归移动平均模型 (ARIMA 模型, Autoregressive Integrated Moving Average Model), 是由博克思 (Box) 和詹金斯 (Jenkins) 于 20 世纪 70 年代在 “Time Analysis Forecasting and Control” 一书中提出的一种著名的时间序列预报方法。该模型将预报对象的时间序列视为是由某个随机过程产生的, 用一定的数学模型区建立、估计该随机过程。这一模型一旦被识别后可从时间序列的过去值及现在值来预报未来值。

本研究利用 DPS v7.05 平台, 例如: 将西

安 1971—2000 年的 1 月平均气温原始数据资料按时间顺序在编辑器中输入编辑作为分析、建模数据序列, 预测 2001 年 1 月的平均气温; 并以此递归, 依次做出 7 个区县 2001—2013 年逐月、季、年平均气温作为模型检验数据序列。

#### 2.4 灰色系统法

灰色系统是认为区域冷暖年型的变化是由许多因子综合作用的结果, 其中部分特性是不确知的。灰色系统建立的微分方程模型, 是利用无规律的原始数据去以数找数, 通过一定方法处理后, 变成比较有规律的时间序列数据。将时间序列按微分方程拟合, 所得模型称为灰色动态模型 (Grey Model), 简称 GM 模型<sup>[4]</sup>。

本研究利用 DPS v7.05 平台, 验证原始数据量的多少对预测精度的影响, 采用西安各区县年、3 月、6 月、9 月、12 月平均气温作为原始数据, 分别建立 30 a (1971—2000) 预测 2001 年、(1972—2001 预测 2002)、……, 10 a, 5 a 动态模型。结果显示采用 10 a 时对逐月、季、

年平均气温进行预测误差最小。

#### 2.5 PS 检验方法

本研究检验方法使用 2014 年正式实施的用于国家级和省级短期气候预测质量检验的《月、季气候预测检验评分方法》中的 PS 评分方法。评分使用的月、季、年平均气温为多年平均值 (1981—2010 年)。

PS 检验方法的计算公式为

$$P_s = \frac{aN_0 + bN_1 + cN_2}{(N - N_0) + aN_0 + bN_1 + cN_2 + M} \times 100\% \quad (3)$$

其中:  $N_0$  为气候趋势预测正确的站数;  $N_1$  为一级异常预测正确的站数;  $N_2$  为二级异常预测正确的站数;  $M$  为没有预报二级异常而实况出现气温距平  $\geq 3^\circ\text{C}$  或  $\leq -3^\circ\text{C}$  的站数 (漏报站);  $a$ 、 $b$  和  $c$  分别为气候趋势项、一级异常项和二级异常项的权重系数, 本办法分别取  $a=2$ ,  $b=2$ ,  $c=4$ 。气温趋势预测产品各等级划分标准和服务用语见表 1。

表 1 气温趋势预测产品各等级划分标准和服务用语

服务用语	特低	偏低	正常略偏低	正常略偏高	偏高	特高
$\Delta T$	$\Delta T \leq -2.0$	$-2.0 < \Delta T \leq -1.0$	$-1.0 < \Delta T < 0$	$0 \leq \Delta T < 1.0$	$1.0 \leq \Delta T < 2.0$	$\Delta T \geq 2.0$
异常分级	二级异常	一级异常	正常级	正常级	一级异常	二级异常
趋势	负趋势	负趋势	负趋势	正趋势	正趋势	正趋势

注:  $\Delta T$  为平均气温距平 ( $^\circ\text{C}$ )

### 3 试报结果检验与异常分析

#### 3.1 评分结果检验

平均气温评分检验分析: 用上述四种方法对西安市 2001—2013 年各月、季、年平均气温分别进行回报和预测, 采用 PS 评分方法进行评分检验<sup>[9-10]</sup>。表 2 为西安各区县逐月、季、年预测方法评分最高的预测方法一览表。

近 13 年全市以灰色系统法预测年平均气温得分最高, 为 84.72 分。最优气候值法的年评分平均值为 84.59 分, 自回归移动平均法为 78.10 分, 多线性回归法为 77.24 分。市区、长安、蓝田、临潼以最优气候值法预测年平均气温评分最高, 得分在 82.60~86.55 分之间; 周至、户县、高陵以灰色系统法预测年平均气温评分最高, 得

分在 84.68~88.24 分之间。

最优气候值法全市平均评分春季 89.68、夏季 87.33、秋季 82.07、冬季 78.53, 同号率 (预测趋势正确率) 春季 73.26%、夏季 67.77%、秋季 58.61%、冬季 50.55%; 多线性回归法全市平均评分春季 73.36、夏季 79.76、秋季 78.13、冬季 77.70, 同号率春季 37.73%、夏季 46.52%、秋季 47.62%、冬季 45.79%; 自回归移动平均法全市平均评分春季 72.49、夏季 78.35、秋季 78.10、冬季 82.76, 同号率春季 35.90%、夏季 46.15%、秋季 44.69%、冬季 52.75%; 灰色系统法全市平均评分春季 83.48、夏季 84.72、秋季 83.28、冬季 81.43, 同号率春季 72.53%、夏季 69.60%、秋季 62.27%、

表 2 西安各区县月、季、年评分最高的预测方法

时段	周至	户县	临潼	蓝田	长安	市区	高陵
1月	方法4	方法3	方法3	方法3	方法3	方法1	方法3
2月	方法4	方法1	方法1	方法3	方法2	方法2	方法3
3月	方法1	方法1	方法1	方法1	方法1	方法1	方法4
4月	方法4	方法4	方法1	方法1	方法4	方法4	方法1
5月	方法4	方法4	方法1	方法1	方法1	方法4	方法4
6月	方法1	方法1	方法4	方法1	方法1	方法1	方法4
7月	方法4	方法4	方法1	方法3	方法3	方法1	方法4
8月	方法1	方法2	方法2	方法4	方法1	方法4	方法4
9月	方法4	方法4	方法3	方法3	方法1	方法4	方法4
10月	方法2	方法3	方法3	方法3	方法4	方法4	方法4
11月	方法2	方法2	方法4	方法3	方法1	方法1	方法3
12月	方法1	方法1	方法4	方法2	方法3	方法2	方法3
春季	方法4	方法4	方法1	方法1	方法1	方法4	方法1
夏季	方法4	方法4	方法1	方法1	方法1	方法1	方法4
秋季	方法4	方法4	方法1	方法3	方法1	方法4	方法4
冬季	方法4	方法3	方法3	方法3	方法3	方法2	方法3
年	方法4	方法4	方法1	方法1	方法1	方法1	方法4

注：最优气候值法—方法1；多线性回归预测法—方法2；自回归移动平均法—方法3；灰色系统法—方法4。（下同）

冬季 52.01%。

### 3.2 异常气温预测能力对比分析

利用 13 a 中一级异常和二级异常预测准确次数除以预测总次数再乘以 100%，对西安 7 个区县 13 年的月平均气温预测样本进行 4 种客观方法的异常级分析（表 3）。最优气候值法、灰色系统法在准确率及一级异常内表现最优，在二级异常内自回归移动平均法在部分区域表现最优。最优气候值法全市月气温一级异常级概率 28.57%，临潼最低，长安最高；多线性回归法全市月气温一级异常级概率 29.40%，长安最低，户县、蓝田最高；自回归移动平均法一级异常级概率 29.81%，蓝田最低，周至最高；灰色系统法一级异常级概率 29.85%，蓝田、临潼最低，户县最高。最优气候值的一级异常级最少，灰色系统法的二级异常级最少，自回归移动平均法的漏报概率最少。最优气候值的一、二级异常级和漏报概率和最少，为 49.03%，其次是灰色

系统法，为 49.63%。最优气候值法在长安气温预测中异常最大，多线性回归法在市区和户县气温预测中异常最大，自回归移动平均法和灰色系统法均在市区气温预测中异常最大。

## 4 结论

(1) 季、年平均气温预测以最优气候值法、灰色系统理论法较为适用；月平均气温预测中最优气候值法、自回归移动平均法和灰色系统理论法在时空序列上各有优劣；多元线性回归预测法在西安市区县平均气温预测中适当避免使用。

(2) 从时间尺度来看，近 13 年西安全市各区县春、夏、秋和年平均气温预测以最优气候值法最优，灰色系统法次优；冬季平均气温预测以自回归移动平均法最优；各区县月平均气温预测 3—8 月以最优气候值法最优、灰色系统法次优；1—2 月、9—12 月以自回归移动平均法最优，灰色系统法次优。

(3) 从空间尺度来看，靠近山区及平原区

表 3 4 种预测方法异常概率统计表

%

异常概率等级	预测方法	市区	周至	户县	长安	蓝田	临潼	高陵	全市
准确率	方法 1	53.21	50.00	49.56	46.15	49.36	54.48	54.05	50.97
	方法 2	40.38	42.30	40.38	50.64	48.07	45.51	50.64	45.41
	方法 3	33.32	44.23	48.08	50.00	55.14	48.72	49.04	46.93
	方法 4	47.44	50.64	48.71	50.00	54.49	49.36	51.92	50.37
一级异常内 准确概率	方法 1	78.21	80.13	77.77	82.05	82.05	77.56	79.05	79.54
	方法 2	71.15	72.43	71.79	76.92	79.48	74.36	77.56	74.81
	方法 3	61.53	77.56	78.21	81.41	82.70	78.21	77.56	76.74
	方法 4	80.13	79.49	83.33	79.49	80.13	75.00	83.97	80.22
二级异常内 准确概率	方法 1	93.59	92.95	91.87	94.87	93.59	94.23	92.51	93.37
	方法 2	89.10	89.10	89.10	95.51	95.51	91.67	93.59	91.94
	方法 3	83.97	93.59	92.95	98.72	98.08	93.59	96.15	93.86
	方法 4	91.03	94.23	94.87	92.31	92.95	91.03	94.23	92.95
漏报概率	方法 1	6.41	7.05	8.13	5.13	6.41	5.77	7.49	6.63
	方法 2	10.90	10.90	10.90	4.49	4.49	8.33	6.41	8.06
	方法 3	16.03	6.41	7.05	1.28	1.92	6.41	3.85	6.14
	方法 4	8.97	5.77	5.13	7.69	7.05	8.97	5.77	7.05

(周至、户县、市区、高陵)月、季、年平均气温预测以灰色系统法最优, 最优气候值法次优; 山区与平原区过度区域(临潼、蓝田、长安)月、季、年平均气温预测以最优气候值法最优, 自回归移动平均法次优。

#### 参考文献:

[1] 魏凤英. 我国短期气候预测的物理基础及其预测思路 [J]. 应用气象学报, 2011, 22 (1): 1-11.

[2] 陈杰. 最优气候均态模型在集宁月平均气温预报中的应用 [J]. 资源节约与环保, 2012 (6): 84-85.

[3] 叶晓波, 王松. ARIMA 模型在楚雄市年平均气温预测中的应用研究 [J]. 楚雄师范学院学报, 2012, 27 (6): 16-21.

[4] 唐湘玲, 刘卫平. 灰色预测模型在阿克苏地区气温预测中的应用 [J]. 石河子大学学报: 自然科学版, 2008, 26 (4): 411-414.

[5] 汪春辉, 罗飞, 舒红平. 偏最小二乘回归在气温预测中的研究与应用 [J]. 微计算机信息, 2012, 28 (5): 141-144.

[6] 王永弟. 模糊时间序列模型在短期气候预测中的应用 [J]. 南京信息工程大学学报: 自然科学版, 2012, 4 (4): 316-320.

[7] 乌文奇, 李春云, 郇文河, 等. 应用最优气候值方法进行短期气候预测 [J]. 内蒙古气象, 2008 (1): 12-13.

[8] 赵春雨, 孟莹. 最优气候值方法在短期气候预测中的应用 [J]. 辽宁气象, 1999 (4): 20-21.

[9] 刘长征, 江远安, 毛炜峰, 等. 新疆夏季降水和冬季气温预测方法及效果评估 [J]. 沙漠与绿洲气象, 2015, 9 (2): 1-8.

[10] 田武文, 吴素良, 王娜. 气候预测 PS 评分对业务影响 [J]. 应用气象学报, 2010, 21 (3): 379-384.