

文章编号: 1006-4354 (2009) 04-0010-04

一种月尺度单站预报方法研究

田 武 文

(陕西省气候中心, 西安 710014)

摘 要: 介绍一种月尺度单站预报方法。通过从前期大气环流的高层、中层环流场及中高层温度场的月平均资料中, 提取一组影响单站气候要素变化的信号源, 组成多元线性回归模型, 同时对同一要素的多个测站进行预报。在对全省 10 个气候区的业务预报中, 经过多年应用, 该方法逐步成为短期气候预测的关键决策依据。利用该方法, 建立全省 14 个测站的预报模型, 并于 2008 年 3 月投入业务使用。利用该模型的预报结果, 3—12 月预测质量大幅提高。

关键词: 月尺度单站预报; 最强影响信号; 预报方法; 预报质量

中图分类号: P456.1

文献标识码: A

我国短期气候预测主要有两类预报方法: 第一类是通过对大气活动中心位置和强度变化的监测诊断、对外强迫热源的周期变化特征及异常发展的监测诊断, 预测未来气候的可能变化; 第二类是通过对具体预报对象与前期环流场进行相关分析, 确定一组影响气候变化的因子源, 建立单站预报模型。两种方法前者注重宏观信号, 后者侧重统计学原理。两类预报方法相互依赖, 互为

参考, 成为预报质量稳定提高的基础。对于省级短期气候预测, 第一类预报方法往往只能对全省降水多少或气温高低给出定性判断, 而对于具体测站的预报, 还需要有逐站预报结果, 为最终的预报结论提供决策依据。

经过多年研究, 总结出一套气候单站预报方法, 一次可对同一要素的多个测站预报。该方法能分析影响测站气候变化的各种因子, 提取影响

收稿日期: 2009-01-15

作者简介: 田武文 (1958—), 男, 陕西户县人, 高工, 从事短期气候预测业务及研究。

果需要读取多个格点 (例如某个划定区域) 的数据, 只要把相应的程序段加入到循环体中即可。

4.2 T639 还提供了用于驱动区域模式的数据产品, 覆盖范围也是 $0^{\circ}\text{E} \sim 180^{\circ}\text{E}$ 、 $0^{\circ}\text{N} \sim 90^{\circ}\text{N}$, 但空间分辨率为 $0.3^{\circ} \times 0.3^{\circ}$, 总格点数为 601×301 , 层次数达 17 层。

参考文献:

- [1] 管成功, 陈起英, 佟华, 等. T639L60 全球中期预报系统预报试验和性能评估 [J]. 气象, 2008, 34 (6): 11-16.
- [2] 沈桐立, 田永详, 葛孝贞, 等. 数值天气预报 [M]. 北京: 气象出版社, 2003: 355-382.
- [3] 谭浩强. Visual C++ 6.0 使用教程 [M]. 北京: 电子工业出版社, 2001: 268-291.

图 1 垂直速度空间剖面图 (10^{-2} hPa/s)

4 讨论

4.1 提供的程序段用以读取 1 个格点的数据, 如

具体测站气候变化的信号源,实现了统计预报模型的动态最优化,能对各种要素未来气候的可能变化做出较准确预测。

1 资料

预报对象为每月向中国气象局上报的 14 个测站的月降水总量和月平均气温预报。其中降水预报的月份为 3—11 月(按 P_s 评分规定, 12—2 月月降水量的气候均值小于 10 mm, 因此不评定预报质量)。因子为中国气象局按月下发的 500、100 hPa 高度的月平均环流资料。

建立模型使用的预报对象为 1961—2007 年, 共 47 个样本。因子使用的资料为 1960—2007 年。本方法利用前期环流资料建立统计预报模型。使用的前期资料为北半球平均环流场, 网格为 10×10 个经纬度, 即每层有 36×8 个格点。前期资料最早是上年 1 月, 由于在月底前发布下月预报, 最晚截止到预报月份的前 2 个月。如预报 5 月降水量, 在 4 月底前发布预报产品, 最晚的资料为 3 月平均环流场。

在短期气候预测使用的资料中, 100 hPa 代表高层引导环流场, 500 hPa 代表中层控制环流。作为因子的拓展, 使用 $H_{100} - H_{500}$ 厚度场资料, 代表高层温度场。上述因子, 保证了每个测站都有丰富的因子源。如预报 6 月降水量, 从上年 1 月到当年 4 月, 共 16 个月、3 层备选因子, 共 $16 \times 3 \times 36 \times 8 = 13\ 824$ 个格点。即对于每一测站的每一要素, 都是从 13 824 个格点上, 找出一组与预报对象相关程度最好的因子源, 建立预报模型。

2 单站预报方法

单站预报方法早期也有过大量研究^[1-5], 主要通过回归模型和点聚图等方法预报。但由于计算条件的限制, 在因子更新速度、预报模型因子系数动态变化等环节相对薄弱。本文研究的预报方法充分解决了这一问题。

2.1 相关分析

程序的目的是求出前期环流场的每一层、每一个格点与预报对象的相关系数, 并输出 2 个文件: 一是各月、各站大于给定临界值的因子文件, 内容包括预报站点序号标识、预报月份、对应的因子月份、环流层标识、格点对应的纬度(行)、

经度(列)、按一维排列时的格点位置, 相关系数等; 另一是面相关系数文件, 内容包括站名、预报月份、对应的因子月份、环流层标识、每一个格点上的相关系数等。文件把大量统计信息提供给预报员, 以便进行因子分析和提取因子。

因子分析是, 利用上述输出结果, 分析出与预报对象对应的高相关中心, 利用高相关中心附近的几个格点, 或者一正一负两个高相关中心的若干格点, 作为影响该预报对象气候变化的信号源, 组合成预报因子。通过人工分析, 得出与各站各月最密切的一组因子。针对每一个预报对象, 确定了一组因子之后, 最终形成对同一月同一要素的多个测站的因子库文件。

因子库文件是在建立动态预报模型、业务预报时连接预报对象、因子文件的枢纽文件。例如预报 7 月降水, 对应的因子库文件名为 js_07.txt, 数据内容如下:

```
5
2 3 2 3 2
-6 2 157 193
-7 3 88 89 90
-8 3 158 194
-9 3 193 194 228
-12 1 63 64
6
3 3 3 4 2 1
-2 1 159 160 161
-4 1 55 56 57
-5 3 140 141 175
-12 1 67 68 69 70
-12 3 168 169
5 1 247
5
2 2 2 2 4
-2 1 96 97
-2 3 171 172
-4 2 160 161
-5 2 201 202
-12 2 67 68 -172 -173
...
```

因子库文件内容说明：第1行代表第1个站（榆林）共有5个因子。第2行表示这5个因子分别由2、3、2、3、2个格点组成。第3行第1列数据的绝对值代表因子的月份，月份为负表示为上年的因子，月份为正代表当年的因子；第2列代表环流层标识，标识为1代表500 hPa高度场，2代表100 hPa高度场，3代表100~500 hPa之间的厚度场；第3列以后的各个数据的绝对值代表该因子在月平均环流场中具体格点的位置。

1~7行是第一个预报对象所对应的预报模型的完整描述。通过它指出了预报模型所用的因子个数、每个因子所用的网格点个数、与预报对象对应的因子月份、因子所在的环流层、网格点在月平均环流场中的具体位置以及因子的组成方式等等。预报几个测站，就要对上述结构重复几次。

格点位置的排列方式：把对应的月平均环流场读入2维数组，1维为时间序列，1维为格点位置；格点位置按先列后行的方式排列；如第1行第1、2、3...、36列为第1、2、3...、36个格点，第2行第1列的格点为第37个格点，第3行第1列的格点为第73个格点，余此类推。

因子组合方式为：文件中的格点位置为正，表示在因子组合时，加上该格点的值；格点位置为负，表示在因子组合时，减去该格点（格点位置的绝对值）的值。

如在上述js_07.txt文件中，第1站的第1个因子（第3行）的值，为第157个格点与第193个格点之和，代表了上年6月，在100 hPa月平均高度场上，(30°N~40°N, 130°E)范围内的高度值之和。第1站第5个因子（第7行）的值为第63个格点与第64个格点之和，代表了上年12月，在500 hPa月平均高度场上，(70°N, 80°W~90°W)范围内的高度值之和。上述2个因子是一种因子组合方式，代表了相对于榆林站的2个高相关中心因子。第3站第5个因子（第22行）的值为第67个格点与第68个格点之和减去第172个格点与第173个格点之和，代表了上年12月，在100 hPa月平均高度场上，H(70°N, 40°W~50°W)与(40°N, 110°W~120°W)范围内高度之

差。这是另一种因子的组成方式，它代表了相对于绥德站，由100 hPa的2个正负相关中心形成的一种涛动因子。

2.2 建立变系数的预报模型

因子库文件使预报对象与预报信号形成固定的对应关系，就可以通过标准化制作月预测产品的模式，同时建立起多个测站的动态线性回归预报模型。模型在每次预报前重新运行一次，由于每次运行都使用了上年该月的实况场和环流预报场，因此，模型是一个变系数的预报模型。例如，假设因子库文件不变，利用上述3个数据文件预报2009年7月降水量时，就可利用2008年以前的所有样本，计算出最新的模型统计参数。因此该方法是随着时间变化而不断更新预报模型系数的动态预报模型。

模型建成后，输出的产品包括模拟质量的统计参数和预报模型参数两个文件：模拟质量参数文件，包括每一个预报对象与各个因子的相关系数，模拟预报结果与实况降水量的复相关系数、回归平方和、残差平方和、各因子在新模型中所占的方差比等，供预报人员监控预报质量；预报模型参数文件，为各测站预报模型的回归系数。

通过上述因子库文件、变系数模型建立的预报模型参数文件，12个月月平均环流场文件，就可以同时预报多个测站的月降水量。

在建立预报模型的过程中，由于预报对象与因子的相关程度会随时间演变，各因子在预报模型中所占的权重也会发生变化，预报员可以根据需要，调整所使用的因子。

2.3 预报订正

利用前期预报模型预报与实况的对比，结合统计学原理，制定订正规则。如围绕大多数气象站都被城市建筑包围或半包围，造成观测温度偏高的现象；当地前期气候实况对后期气候的惯性影响；单站预报结论与周边测站预报截然相反的现象，总结出预报误差出现的规律，制定出预报订正方案。

3 业务应用

3.1 区域预报

以上研究的预报方法，均已制成业务程序，预

报对象为全省 10 个地级市的区域平均降水量和平均气温。组成方式为：以地级市为中心，选取临近的 3~5 个气象站，计算出各站月降水量或月平均气温的平均值。方法自 1998 年开始在业务中应用。期间经过不断研究改善，由于其预测效果较好，逐步成为陕西省短期气候预测的主要决策依据。

3.2 单站预报及效果

2007 年 2 月，中国气象局要求自 3 月 1 日开始，每月向中国气象局发布榆林、定边、绥德、延安、洛川、武功、汉中和安康共 8 站的单站气候预测。而陕西省气候中心还承担着向地市一级的气候预测指导业务，鉴于该原因，气候中心同时制作发布了两种预报，即 10 个区域的向下指导预

报和 8 个测站的向上预报。但一方面由于时间紧迫，当时没有建立单站逐月、逐要素的预报模型；另一方面，目前全国气候预测的主要方法是利用概念模型进行预测^[6]，因此预报员依据自行总结的物理模型，参考 10 个区域的预测结果制作 8 站预报，但与全国相比预报质量较差。

由于所选 8 站对陕西的代表性较差，2008 年 2 月，通过与中国气象局有关部门协商，上报测站增加了 6 个，分别为铜川、宝鸡、咸阳、临潼、渭南、商洛，共 14 个测站。利用本文提出的预报方法，逐月、逐站建立了降水、平均气温预报模型，制作发布了 3—12 月预报。在没有任何订正的情况下， P_s 评分为：降水预测 79.3%、平均气温预测 84.6%（表 1）。

表 1 利用气候单站预报方法预报 2008 年 3—12 月降水、平均气温 P_s 得分

预报对象	3 月	4 月	5 月	6 月	7 月	8 月	9 月	10 月	11 月	12 月	平均
降水	87.5	100	75.0	75.0	50.0	75.0	100	62.5	75.0		79.3
平均气温	75.0	100	87.5	87.5	87.5	87.5	100	75.0	87.5	62.5	84.6

由于该方法的支撑，2008 年陕西省短期气候预测 P_s 评分为：降水预报 79.4%；在温度温度预报中，尽管 1 月评分 12.5%，2 月评分 50%，但由于 3—12 月取得了较好的成绩，全年仍然达到 75.7%，大大提高了陕西省气候预测评分在全国的名次。

4 讨论

短期气候预测没有可视化的预报工具，只有建立在严格遵循统计学原理基础上，针对每个测站、每个要素找出影响其变化的信号源，并在此基础上，通过标准化制作月尺度预报产品的模式，建立多个测站的动态预报模型，才能做出同全国各省比较，具有竞争力的的预报产品。

陕西省短期气候预测使用单站分析预报模型大幅提高预报质量的事实，说明深入研究单站预报模型的必要性。

参考文献：

- [1] 熊敏谔, 陈菊英. 中国 7 月降水分型及其成因 [J]. 气象, 1997, 27 (6): 43-46.
- [2] 陈菊英, 李存强. 黄河中上游地区夏半年旱涝的长期预报因子和工具的剖析 [J]. 高原气象, 1989, 8 (01).
- [3] 陈菊英, 程华琼, 王威. 中国异常增暖来年江淮流域易发生大洪水 [J]. 地球物理学进展, 2007, 22 (4): 1380-1385.
- [4] 陈菊英. 中国旱涝的分析和长期预报研究 [M]. 北京: 中国农业出版社, 1991.
- [5] 陈菊英. 海滦河流域汛期旱涝规律成因和预测研究 [M]. 北京: 气象出版社, 1991.
- [6] 赵振国, 刘海波. 我国短期气候预测的业务技术发展 [J]. 浙江气象, 2003, 24 (3): 1-6.