

提高最优插值法测量区域降水量 精度的探讨

李建通 杨维生 郭 林 陈泽面

(福建省厦门市气象局, 厦门 361012)

摘 要 首先从理论上研究了要素场相关函数和校准雨量站的分布对最优插值中最优权重系数的影响, 然后用3种相关函数模型对1996~1997年收集的113个时次的雷达和雨量计资料进行计算、检验。结果表明: 采用自适应相关函数模型的最优校准法可以有效提高区域降水量的测量精度。另外, 通过对大误差的产生原因分析, 给出合理设置自动雨量站的建议和解决边界测量精度较低的方法。

关键词: 最优插值; 区域降水量; 精度; 自动雨量站

1 引言

最优插值法测量区域降水量是一种较为简便而又具有一定精度的测量方法, 如何进一步提高最优插值校准法的测量精度是各级气象台站在实际应用中经常要思考的问题, 本文通过多年的应用研究, 提出采用一种自适应的相关函数模型来提高该法的测量精度。

2 基本原理

最优插值法是客观分析的一种方法^[1~3], 每个格点的分析值是由格点的初估值加上订正而求得, 其权重函数不是任选的, 选择的结果使得格点的分析值误差达到最小。

在最优插值法中, 每一格点的分析值可由下式表示:

$$A_{a,k} = A_{g,k} + \sum_{i=1}^N (A_{ob,i} - A_{g,i}) P_i, \quad (1)$$

其中 $A_{a,k}$ 为格点的分析值, $A_{g,k}$ 为格点的初估值, $A_{g,i}$ 为第 i 个测站的初估值, $A_{ob,i}$ 为第 i 个测站的观测值, P_i 为第 i 个测站的权重系数。

假设观测误差之间以及它和偏差场之间误差无关, 并假定相关函数是各向均匀同性, 则最优的权重系数可由下式求得:

$$\sum_{j=1}^N P_i \mu_{ij} + \eta_i P_j = \mu_{ki}, \quad (i = 1, 2, \dots, N) \quad (2)$$

μ_{ij} 表示 i 、 j 两点之间的相关函数, η_i 为第 i 个测站观测值的相对均方差, 在实际计算中

取为零。则 (2) 式可以简化为

$$\sum_{j=1}^N P_i \mu_{ij} = \mu_{ki}, \quad (i = 1, 2, \dots, N) \quad (3)$$

从 (3) 式可知, 只要给定要素场的相关函数, 就可以求解最优的权重系数。

3 最优权重系数的讨论

3.1 相关函数模型

最优插值法定量测量区域降水量中, 要素场的相关函数通常取以下两种形式:

$$\mu(r_{ij}) = \exp(-r_{ij}/a), \quad (4)$$

$$\mu(r_{ij}) = \exp(-r_{ij}^2/a), \quad (5)$$

其中 r_{ij} 是 i, j 两点之间的距离, μ 代表 i, j 两点之间的相关函数。

取不同的 a 值 (a 值从上到下分别为 200、100、40、20、10) 代入上述两个公式, 分别得到图 1 和图 2。

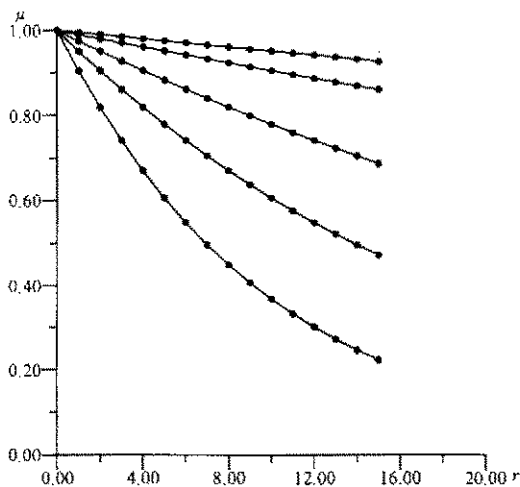


图 1 相关函数 μ 随 r_{ij} 的变化曲线

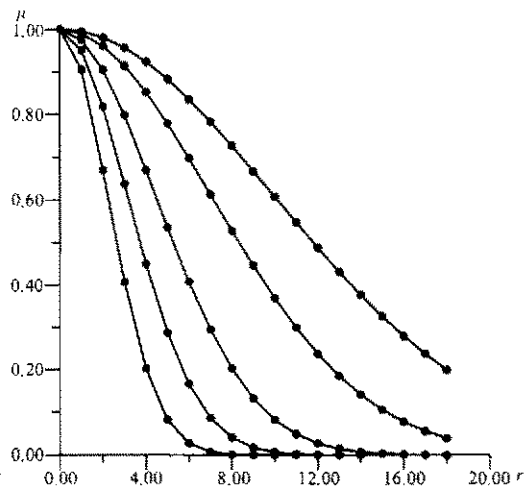


图 2 相关函数 μ 随 r_{ij}^2 的变化曲线

从图 1 和图 2 可以有如下的分析结果:

(1) a 值的选取原则。相关函数趋于零的距离 (称最大影响半径) 由 a 值决定, a 值的大小要根据实际的测站分布和参加站的远近情况确定, 尽量将测站设置在影响半径内的适当位置, 即在测站密集时, a 值要较小, 反之 a 值要取较大, 才能达到最优校准的目的。

(2) 公式 (4) 和公式 (5) 的适用性。在 a 相同时, 比较两张图可知: 图 1 的相关函数随距离衰减较慢, 当测站与格点的距离较远时仍有一定的相关, 这样在测站稀疏时, 格点的校准不致于仅由最近的个别测站决定, 而是所有选定的测站都对之有贡献。若此时采用公式 (5) 则会由于相关函数随距离的迅速衰减, 而造成格点的校准仅由格

点附近的个别测站决定, 远处的测站对格点校准几乎没有贡献, 也就失去求最优权重系数的意义。因此, 图 1 的相关函数适合于测站稀疏时使用, 反之图 2 的相关函数适合测站密集时使用。

从最优插值的物理意义来看: 每个格点校准要由一定半径内的 N 个测站来订正, 且测站离格点的距离 r 越近, 相应的相关函数 μ 也越大, 校准的结果也越好。

我国现有的站网分布基本按行政区划分^[4], 较少考虑合理布局, 现状是区区内有台, 县县有站, 行政区小人口多的地区站点过密, 反之则过疏。同时由于地形千差万别, 各个站的代表范围各不相同, 为了寻求较多的可供校准站而一味扩大搜索半径, 可能导致不合理的校准。因此, 在实际的应用中, 我们要考虑每个格点的校准如何都能根据可供校准站的实际情况而变。

根据上述的分析, 本文提出自适应相关函数模型来达到合理选用相关函数目的, 最终提高最优插值法测量区域降水量的精度。

3.2 最优权重系数和测站分布的关系

最优插值系数不仅与距离有关, 而且与测站的分布有关^[5]。假设测站的分布如图 3a ~ d, P_1 、 P_2 、 P_3 、 P_4 是各个测站相应的权重系数, r 是两点间的距离, μ 是两点间的相关函数, 其中

$$\mu_{ij} = \mu(r) = \mu = \exp\left(-\frac{r}{a}\right). \quad (6)$$

(a) 当格点 K 的周围只有两个等距对称分布的测站时 (图 3a)

$$P_1 = P_2 = \frac{\mu}{1 + \mu^2}; \quad (7)$$

(b) 当两个测站非对称分布, 且站 2 到 K 点的距离是站 1 到 K 点距离的两倍时 (图 3b)

$$P_1 = \frac{\mu(1 + \mu^2)}{(1 + \mu^2 + \mu^4)}, \quad (8)$$

$$P_2 = \frac{\mu^2}{(1 + \mu^2 + \mu^4)}; \quad (9)$$

(c) 当有三个测站, 两边等距各有一个, 第三个和其中的一个相邻很近时 (图 3c)

$$P_2 + P_3 = P_1 = \frac{\mu}{1 + \mu^2}; \quad (10)$$

(d) 当四个测站等距均匀分布在格点的四周时 (图 3d), 则

$$P_1 = P_2 = P_3 = P_4 = \frac{\mu}{1 + 2\mu(\sqrt{2} \times r) + \mu^2}. \quad (11)$$

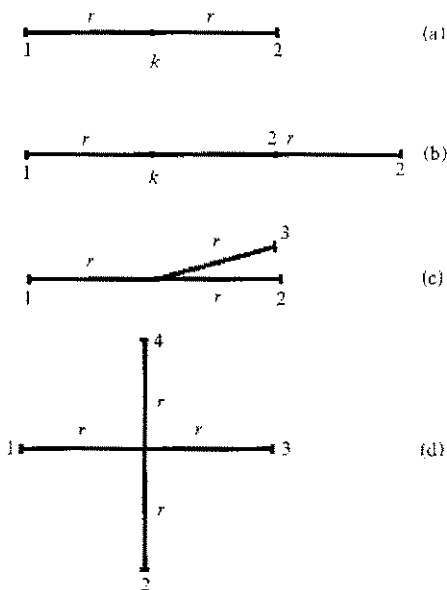


图 3 四种测站分布的示意图

由上面的几个关系可以看出:

(1) 当测站等距均匀对称分布时, 其最优的权重系数相等, 大小由相关函数的型式决定;

(2) 当两个测站非对称分布时, 其权重系数也是不等的;

(3) 当两个测站很近时, 其权重系数的和等于对称等距测站的权重系数。

通过上面的分析可知, 最优插值的精度不仅依赖于气象要素的统计结构, 还依赖于测站点的观测精度及参加校准测站的分布, 而与观测数值的大小无关。

4 实测资料的分析

4.1 厦门四周测站分布特点

厦门雷达站依山面海, 东面为台湾海峡, 西侧 100 km 内除南靖、平和、长泰局部山区外基本为平原地区, 北面为戴云山山脉阻挡影响北部地区低降水回波的测量, 100 km 外西侧由于博平岭山脉的影响, 龙岩地区基本为本雷达的测量盲区, 仅能探测对流发展旺盛的降水回波。漳州地区和安溪一带有较为密集的雨量站网, 其他地区雨量站网较为稀疏, 站网的分布如图 4。

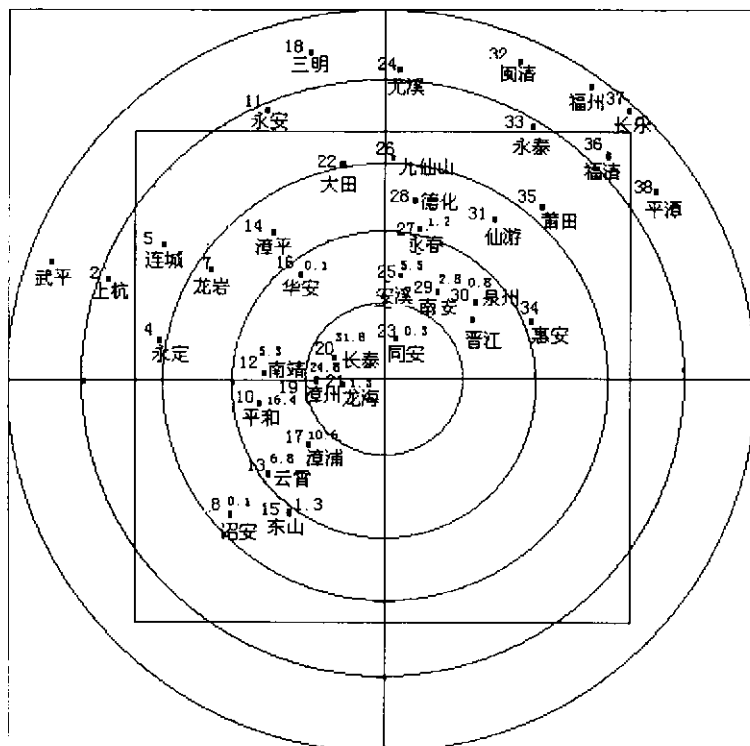


图 4 厦门雷达站周围雨量站分布和部分测站测量的降水量

站点的左上角数字为计算时采用的编号, 右上角的数值为 1997 年 5 月 6 日 23 时~1997 年 5 月 7 日 00 时地面雨量站测得的降水量, 矩形框为最优校准区域, 距离档每圈为 50 km

4.2 自适应的最优权重系数的实现

由上面的分析可知, 雷达站四周雨量站分布不均匀, 漳州较为密集, 地势相对平坦, 测站代表性好, 而其他地区的测站稀疏且代表性差。另外由于实际资料收集和降水分布的不均匀, 更增加了可供校准的雨量站分布的随机性和不均匀性, 若仅采用一种相关函数模型, 是不能达到较好的校准结果。本文根据上述的理论分析, 提出用自适应的相关函数模型求解最优权重系数, 实现方法如下:

给定起始半径 $[R=2$ (格距)], 在此半径内, 满足条件 (雷达回波强度不小于 5 Dbz, 地面的雨量测值大于 0.1 mm/h), 测站的个数不少于 2 个, 则直接采用 (5) 式; 否则, 扩大搜索半径 (增量为 1 个格距), 直至条件满足, 然后判断半径是否大于临界半径 (10 个格距), 小于时用 (5) 式的相关函数模型, 其余用 (4) 式模型。这样每个格点的校准都能根据可供校准雨量站的分布自动选用合适的相关函数模型, 达到更为合理校准的目的。

4.3 资料的计算

首先, 将选定区域雷达低仰角的 PPI 资料转化为 $5 \text{ km} \times 5 \text{ km}$ 分辨率的网格资料, 然后, 将 1 h 内的所有 PPI 资料按时间权重合成 1 h 的逐时雷达平均回波资料, 代表雷达测量的降水初始场。

我们分别用公式 (4)、(5) 和自适应的相关函数 (相应称为方法 I、II、III) 3 种相关函数模型的最优插值法, 对厦门雷达站 1997 年 5 月 6 日 23 时~1997 年 5 月 7 日 00 时的雷达资料 (图 5) 和地面雨量资料 (图 4) 进行实际计算, 结果为表 1 和图 6; 同时我们还对 1996~1997 年采集的 1996 年 3 月 29 日~1996 年 3 月 30 日、1997 年 4 月 13 日~1997 年 4 月 14 日、1997 年 5 月 5 日~1997 年 5 月 7 日、1997 年 8 月 30 日~1997 年 8 月 31 日 4 次降水过程共 113 个时次的雷达资料和相应的地面雨量计资料进

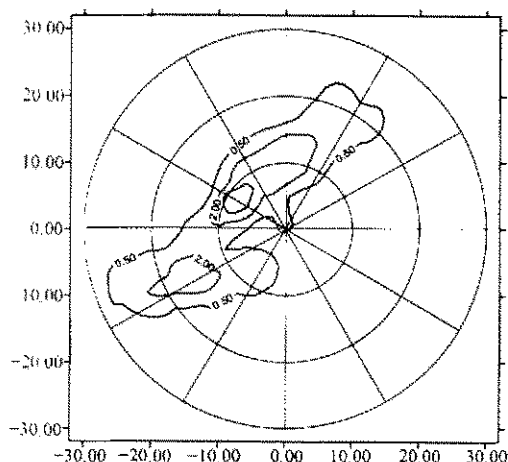


图 5 雷达测量的回波初始雨量分布
等雨量线的数值为 0.5、2.0、5.0 mm

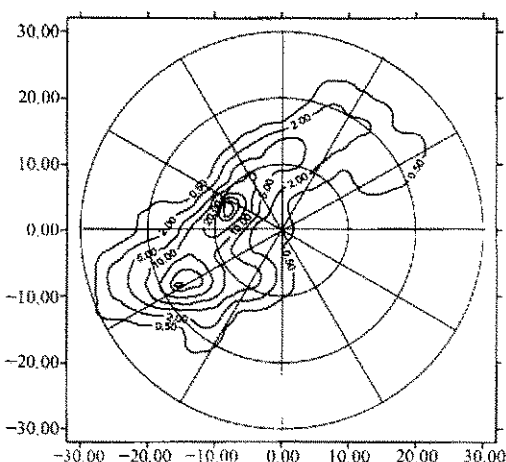


图 6 自适应最优插值校准后的雨量分布
等雨量线数值为 0.5、2、5、10、15、
20、25、30、35 mm

行计算，并读取最优校准后相应的雷达测量值同地面雨量计资料共 1132 个样本作各点的精度比较，结果如表 2 和表 3。精度计算方法如下：

$$\text{误差}(\%) = \frac{|\text{地面雨量计值} - \text{校准后雷达测量值}|}{\text{地面雨量计值}} \times 100\%.$$

表 1 三种校准方法测量精度的比较

站名 (编号)	雨量站测值 (mm)	初始误差 (%)	方法 I 校准后误差 (%)	方法 II 校准后误差 (%)	方法 III 校准后误差 (%)
诏安 (08)	0.1	51	0	0	0
平和 (10)	16.4	89	6	6	6
南靖 (12)	5.3	91	6	6	6
云霄 (13)	6.8	93	5	5	5
东山 (15)	1.3	91	11	2	2
华安 (16)	0.1	44	33	15	15
漳浦 (17)	10.6	87	6	6	6
漳州 (19)	24.8	94	5	5	5
长泰 (20)	31.8	73	1	1	1
龙海 (21)	1.3	76	18	2	2
同安 (23)	0.3	9	5	5	5
安溪 (25)	5.5	57	2	2	2
永春 (27)	1.2	53	4	4	4
南安 (29)	2.8	52	2	2	2
泉州 (30)	0.6	74	6	6	6

表 2 三种校准方法测量精度的比较

	方法 III 优于 方法 I 和 II	方法 III 劣于 方法 I 和 II	方法 II 优于 方法 I	方法 I 优于 方法 II	方法 II 等于 方法 I
次 数	1 033	99	214	46	872
百分比	91.2%	8.8%	18.9%	4.1%	77.0%

注：优于、劣于、等于指的是精度

(1) 从表 1 可以看出：龙海和东山用方法 I 的测量精度低于方法 II 和 III，从雨量站的分布来看，这两个点附近的测站分布是比较密集，这和上面的理论分析的在测站密集时采用 (5) 式的结果是一致的。

(2) 从表 2 可以看出：

(a) 采用自适应相关函数的最优插值法的测量精度在 90% 以上可以不劣于其他两种测量方法；

(b) 在我们使用的雨量站资料中，测站密集地区的雨量站使用的次数较多，因此方法 II 精度优于方法 I 出现的次数多，但总体反映方法 II 和 I 是适用不同的测站分布。

(3) 从表 3 可以看出产生较大误差时主要有 3 种情况：

(a) 测站位于临海的区域，如 8、15、35 号站，产生大误差的原因是可供校准的测站均集中在该点的一侧，最优的权重的校准不能将四周的误差都校准到该点造成；

(b) 测站由于地物的遮挡导致地面有降水, 而相应的雷达上无回波, 在我们校准时由于雷达无回波, 而未作校准, 或有较弱回波而周围无回波用较远的测站进行校准而导致误差较大, 如 4、5、33、36 号站; 另一方面这些测站也位于区域的边界, 也可能产生 (a) 的情况使精度降低;

(c) 测站位于山区, 由于某些特定的回波特点, 也可能出现一些大的误差。但除了 14 号站样本少出现比例高外, 其他如 10、12、16、20、26、27 站出现大误差的比例都低于 15%。

(4) 从上述图表可知, 自适应相关函数法能够自动适应可供校准站的分布选用相应的相关函数, 达到合理校准的目的, 校准后的雨量分布和地面测量的雨量分布能够较好地吻合, 且各个测站校准后的误差也明显降低。

(5) 除去上述出现大误差的原因 (占 14.0%), 其余各点校准后的精度基本都控制在 80% 以上, 可以满足日常业务工作的需要。

5 结论、经验和存在的问题

通过对图 1 和图 2 的分析可知, 不同的权重系数适应不同的测站分布, 由于可供校准测站的分布千变万化, 每个格点不可能用固定的相关函数模型和权重系数进行校准。因此, 本文提出的自适应相关函数是一个能根据可供校准测站分布的情况, 自动选择合理相关函数的新方法, 它能有效提高测量区域降水量的精度, 同时对测站分布的要求也较低, 即使在测站不太理想的情况下也能保证一定的精度, 是一种值得在今后的研究和实践中进一步完善的方法。通过几年的实际应用, 有以下的几点体会:

(1) 在我们确定了区域降水量的测量范围后, 应适当扩大校准区域, 将测量区域外围的一些雨量站加入校准, 防止由于可校准测站位于格点一侧而造成大误差, 从而达到提高区域边界的测量精度的目的。

(2) 为了提高测量区域降水量的精度, 应在容易产生大误差的地区如山区、区域边界等地区多设置自动雨量站, 加大测站的密度, 缩小每个测站的代表范围, 从而降低这些地区校准误差。

表 3 出现误差大于 60% 的测站及其所占的百分比

台站编号	出现次数	使用次数	百分比%
4	4	10	40
5	16	19	84.2
7	0	18	0
8	19	54	35.2
10	4	68	5.9
12	5	63	7.9
13	4	48	8.3
14	1	2	50
15	10	47	21.3
16	8	61	13.1
17	2	40	5
19	1	60	1.7
20	3	67	4.5
21	5	71	7.0
22	1	18	5.6
23	4	62	6.5
25	2	52	3.8
26	2	20	10
27	5	54	9.3
28	6	52	1.2
29	1	69	1.4
30	8	61	13.1
31	0	22	0
33	19	23	83.6
34	0	28	0
35	9	23	39.1
36	19	19	100
累计	158	1132	14.0

(3) 当由于地物遮挡造成雷达初始场上无降水回波时, 最优插值法就无法对其进行校准, 从而在估计这些区域降水量时仍然存在较大误差。如何对这种情况的降水进行订正, 是我们进一步提高区域降水量测量精度工作中仍需研究解决的问题。

参 考 文 献

- 1 李建通、张培昌, 最优插值法用于天气雷达测定区域降水量, 台湾海峡, 1996, 15(3), 255~259.
- 2 廖洞贤、王两铭, 数值天气预报原理及其应用, 北京: 气象出版社, 1996, 361~364.
- 3 徐一鸣、丁荣富、李佐凤, 统计插值客观分析方法的试验研究, 气象学报, 1989, 47(2), 237~243.
- 4 王庆安、顾亚进, 气象站网设计探讨, 气象科学, 1988, 3, 72~80.
- 5 张玉玲、吴辉璇, 数值天气预报, 北京: 气象出版社, 1986, 405~408.

A Study of Improving Precision of Measuring Regional Precipitation in Optimum Interpolation Method

Li Jiantong, Yang Weisheng, Guo Lin and Chen Zemian

(Xiamen Department of Meteorology, Xiamen 361012)

Abstract We firstly studied influence of power coefficients in optimum interpolation by correlation function of element field and distribution of gauge. Then three correlation function models were used to calculate 113-hour data of radar and gauge. The result show: the adaptive correlation functions model can far more develop precision of measuring regional precipitation. With analysis of the cause of large errors, we finally gave suggestion of setting automatic gauges and method of solving lower precision on boundary.

Key words: optimum interpolation method; amount of regional precipitation; precision automatic gauge