

基于台站降水资料对不同空间 内插方法的比较*

冯锦明^{1,2)} 赵天保^{1,2)} 张英娟^{1,2)}

1) (中国科学院大气物理研究所东亚区域气候-环境重点实验室, 全球变化东亚区域研究中心, 北京 100029)

2) (中国科学院研究生院, 北京 100039)

摘 要 如何把离散的气象台站资料通过合适的空间内插方法转变成规则的网格数据, 对于气候变化分析和模拟研究具有重要的意义。作者利用中国区域 160 个常用台站 10 年降水观测资料为例, 分别采用克里格 (Kriging) 插值、反距离加权、Delaunay 三角剖分线性插值、双谐样条 (Biharmonic Spline) 插值和 Cressman 客观分析等几种常见的空间内插方法, 较为系统地分析和比较了这几种内插方法插值结果之间的异同, 对其优缺点和适用范围进行了适当讨论, 并采用不同的网格分辨率和影响半径对 Cressman 客观分析方法做了进一步分析比较。结果表明: 在台站分布密集的区域不同空间内插方法之间差异较小, 在台站分布稀疏的区域则差异较大; 与更加密集的台站观测资料的比较显示, 自动调节影响半径的 Cressman 客观分析方法与双谐样条插值方法误差相对较小。

关键词: 空间内插; 克里格插值; 三角剖分线性插值; 双谐样条

文章编号 1006-9585 (2004) 02-0261-17 **中图分类号** P459 **文献标识码** A

1 引言

空间数据的精确估计和反演对于地球科学的研究与发展具有重要的意义。气象观测资料作为一种空间数据, 具有一般空间数据所具有的特点, 蕴含着复杂的非线性动力学机制, 在时空分布上具有纷杂多变的时空特征。因而要把这种离散的、不规则的观测资料转换成规则的、可以成为数值模拟的初始场, 或用于模式检验的气候场以及其他诊断分析所用的规则网格点资料, 这对于气候的模拟预测和诊断分析至关重要。观测台站的空间分布很不规则, 要把这种不规则的数据转换成规则的网格数据, 就需要利用特定区域已知的离散观测数据来估计规则格点上的非观测数据, 这就是所谓的“空间内插”。

对于台站稀疏, 而且分布极为不合理的地区来说, 空间内插是研究这些区域空间变量和分布的基本方法。空间内插的一个重要目标就是估计空间变量的变异函数, 变异函数建立的关键是对初始条件的假定和方法运用的理解, 这种主观的假设在某种程度上必然造成空间插值的不确定性。根据空间数据的特征, 不同插值方法其结果之间的精确性存在很大的变化。空间内插方法的选择对于数据采集稀疏和在很小的空间尺

2004-02-16 收到, 2004-04-09 收到修改稿

* 国家重点基础研究发展规划项目 G199043400 资助

度内数据就有变化的多山地区来说尤为重要。目前空间插值方法在大气科学和地理信息系统 (GIS) 研究中已被广泛应用于温度、降水、空气和土壤湿度等空间变量的估计、预测, 以及其空间分布规律、等值线的绘制和各种网格化等的研究中^[1~4]。这些研究一般认为, 有些空间内插方法适用于温度的内插, 有的适用于降水的内插, 即对不同性质的空间变量要使用不同的空间内插方法。

对于空间数据的内插究竟是否存在一种最优的内插法或者说对于某一特定空间变量是否存在某种最优的内插法? 瑞士洛桑矿物大学地球科学研究所的空间插值 97 比较计划 (SIC97), 利用瑞士的降水观测资料系统地分析和比较了各种插值方法之间的差异和不确定性^[5]。Cllins^[6] 用 8 种插值方法在 2 个地区、3 个时间尺度内分别对最高温度和最低温度 2 种温度变量进行了估计, 并对它们进行了比较和分析。研究认为在不同的时空尺度内, 每个方法对不同温度变量估计的误差是不一样的。还有很多研究人员也利用温度、降水以及土壤湿度等气象观测数据在不同的时空尺度内设计了各种数值实验, 对各种常见的插值方法做了不同的比较和分析研究^[7~12]。这些研究的结果一般都认为: 对于空间数据的内插, 一种“包插百量”的最优空间内插法是不存在的; 对于不同的空间变量, 在不同的地域和不同的时空尺度内所谓的“最优”内插法是相对的。

中国大陆位于著名的东亚季风区, 而且地形复杂, 这一地区的降水存在着极其复杂的时空分布特征, 因此如何把中国区域内台站观测的降水资料内插为气候研究所需要的规则网格点数据, 需要进一步探索。本文利用几种常见的空间内插法对中国区域内 160 个常用台站的降水观测资料进行内插, 并利用更加密集的台站观测资料对插值结果进行系统的比较分析和检验, 以期找出在中国区域对降水的较优空间内插方法, 从而为降水以及其他观测资料的诊断分析和同化处理提供参考和依据。

2 几种空间内插方法简介

空间内插法根据研究目的不同有不同的分类方法, 如根据其基本假设可分为几何方法、统计方法、函数方法和物理模型方法等; 根据数学本质又可分为局部函数法和逐点内插法 2 种。本文主要选取了克里格 (Kriging) 插值法、加权反距离法、Delau-nay 三角剖分线性内插、样条函数插值和 Cressman 客观分析等 5 种不同的空间内插方法进行比较研究, 这些方法是 IDL、MATLAB、Surfer 和 Arc/info 等数据分析处理软件以及气象数据的客观分析研究中经常用到的。

2.1 克里格插值法

克里格法被认为是地学统计最主要的方法之一, 是以南非科学家 D. L. Krige 的名字命名的一种空间内插法。从统计意义上说, 克里格法是从变量相关性和变异性出发, 在有限区域内对区域化空间变量的取值进行无偏、最优估计的一种统计方法; 从插值角度上讲, 克里格法以空间结构分析为基础, 充分利用了数据空间场的性质, 在插值过程中对空间数据求线性最优, 可以反映空间场的各向异性, 是一种无偏估计的内插方法。该法的最佳适用条件是空间变量存在着空间相关性。

克里格法内插一般分为 2 步: 第 1 步是对空间场进行结构分析, 也就是说, 在充

分了解空间数据场性质的前提下, 建立空间变量的协方差函数, 提出变异函数 (variogram) 模型; 第 2 步是在该模型的基础上进行克里格计算。克里格空间统计内插的最大优点是以空间统计学作为其坚实的理论基础, 可以克服内插中误差难以分析的问题, 能够对误差做出逐点的理论估计, 而且它也不会产生回归分析的边界效应。缺点是复杂, 计算量大, 另外变异函数需要根据经验人为选定。

2.2 反距离加权法

反距离加权法是最常用的空间插值方法之一, 幂次参数控制着权重系数如何随着离开一个网格点距离的增加而下降。对于较大的方次, 较近的数据点被给定一个较高的权重份额, 对于较小的方次, 权重比较均匀地分配给各数据点。可以用下式来表示:

$$Z = \frac{\sum_{i=1}^n \frac{1}{(D_i)^p} Z_i}{\sum_{i=1}^n \frac{1}{(D_i)^p}},$$

式中, Z 是估计值, Z_i 是第 i 个样本; D_i 是距离, p 是幂次, 它显著影响内插的结果, 其选择标准是使平均绝对误差最小。当采样点与网格点重合时, 该网格点被赋予和观测点一致的值, 因此这是一个准确插值。它的特征之一是会在网格区域内产生围绕观测点位置的“靶心”。

2.3 Delaunay 三角化线性内插法

这种方法是通过在数据点之间连线以建立起一个三角形网来进行工作。对原始数据点进行 Delaunay 三角剖分的方法是: 所有三角形的边都不能与另外的三角形相交, 并且任何一个由这种三角形的 3 个顶点所确定的圆不能包含其他任何一个三角形的顶点。每一个三角形定义了一个覆盖该三角形内网格点的面, 三角形的倾斜和标高由定义这个三角形的 3 个原始数据点确定, 给定三角形内的全部网格点都要受到该三角形表面的限制。Delaunay 三角化线性内插就是在这种三角剖分的基础上进一步利用线性插值方法来进行插值计算。

2.4 样条函数插值法

样条函数是使用函数逼近曲面的一种方法。样条内插的本质是利用数学方法产生一条通过一组已知采样点的平滑曲线, 并依据这条曲线来估计每个定点上的属性数值, 在计算过程中采用最小曲率的概念来进行。样条函数易操作, 计算量不大, 它与空间统计方法相比具有以下特点: 不需要对空间方差的结构做预先估计; 不需要做统计假设, 而这些假设往往是难以估计和验证的; 同时, 当表面很平滑时, 也不降低精度。样条函数适合于比较平滑的表面, 一般要求有连续的一阶和二阶导数; 它适合于根据很密的台站内插等值线, 特别是从不规则三角网内插等值线。

样条函数的缺点是难以对误差进行估计, 采样点稀少时效果不好。样条函数的种类很多, 本文将采用双谐样条 (Biharmonic Spline) 插值方法^[13]进行比较。

2.5 Cressman 客观分析方法

Cressman 客观分析采用的是逐步订正的方法, 已被广泛应用于各种气候诊断分析和数值模拟研究中。逐步订正法最主要的根据是 Cressman 客观分析函数, 这种方法是 Cressman 在 1959 年提出的。先给定第一猜测场, 然后用实际观测场逐步修正第一猜测

场,直到订正后的场逼近观测记录为止。

$$\alpha' = \alpha_0 + \Delta \alpha_{ij},$$

其中

$$\Delta \alpha_{ij} = \frac{\sum_{k=1}^K (W_{ijk}^2 \Delta \alpha_k)}{\sum_{k=1}^K W_{ijk}},$$

这里, α 为任一气象要素, α_0 是变量 α 在格点 (i, j) 上的第一猜测值, α' 是变量 α 在格点 (i, j) 上的订正值; $\Delta \alpha_k$ 是观测点 k 上的观测值与第一猜测值之差; W_{ijk} 是权重因子, 在 0.0~1.0 之间变化; K 是影响半径 R 内的台站数。Cressman 客观分析方法最重要的是权重函数 W_{ijk} 的确定, 它的一般形式为

$$W_{ijk} = \begin{cases} \frac{R^2 - d_{ijk}^2}{R^2 + d_{ijk}^2}, & (d_{ijk} < R) \\ 0, & (d_{ijk} \geq R) \end{cases}$$

其中, 影响半径 R 的选取具有一定的人为因素, 一般取一常数。 R 选取的原则是由近及远进行扫描, 常用的几个影响半径是 1、2、4、7 和 10。 d_{ijk} 是格点 (i, j) 到观测点 k 的距离。

Cressman 客观分析方法的缺点是统计平滑功能差, 在采样点稀疏、观测资料贫乏的区域进行空间数据内插时会出现很多“空值”斑点。

3 台站分布及其资料介绍

本文选用的降水资料为 2 种中国区域的台站观测资料和英国的 CRU (Climate Research Unit) 资料。2 种台站资料为中国区域内 160 个常用台站和 730 个密集台站 10 年 (1988 年 12 月至 1998 年 11 月) 的月降水量, 其中, 160 个台站资料主要用于空间内插, 而 730 个台站资料用于检验。CRU 资料为空间分辨率为 0.5° 的网格资料。图 1

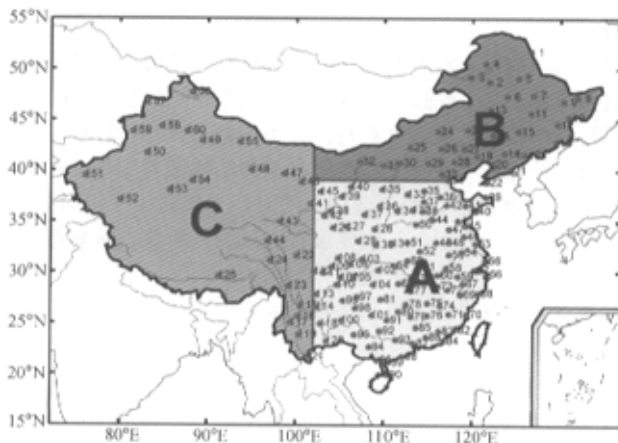


图 1 中国区域 160 个常用观测台站的分布及其分区

和图 2 分别为 160 个常用台站和 730 个密集台站在中国区域的分布；但在图 2 中 10 年观测资料完整的台站为 631 个，且图 1 中所有台站几乎都包含在图 2 中。为了分析和比较数据空间分布对各种插值方法的影响，根据台站的分布密度将整个中国区域分为 A、B、C 3 个区域，其中 A 区台站分布最为密集，B 区次之，C 区台站最为稀疏。同时为了便于统计分析，将 160 个台站序号在图中标出，B 区为小序号台站，A 区为大序号台站。图 3 为根据量值的大小等级所示 160 个台站 10 年平均的年总降水量分布形势。

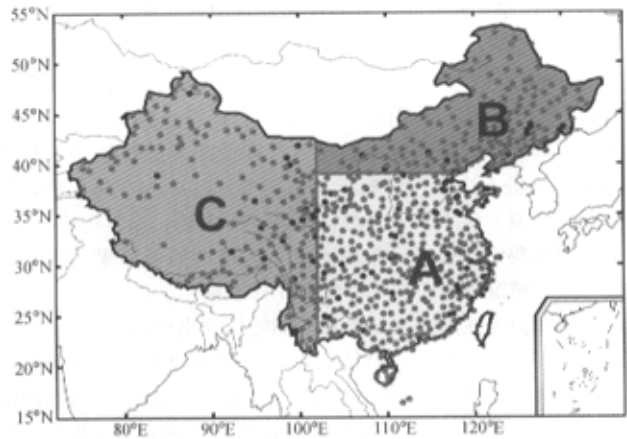


图 2 中国区域 730 个密集台站的分布（深色台站为含有缺失资料的台站）

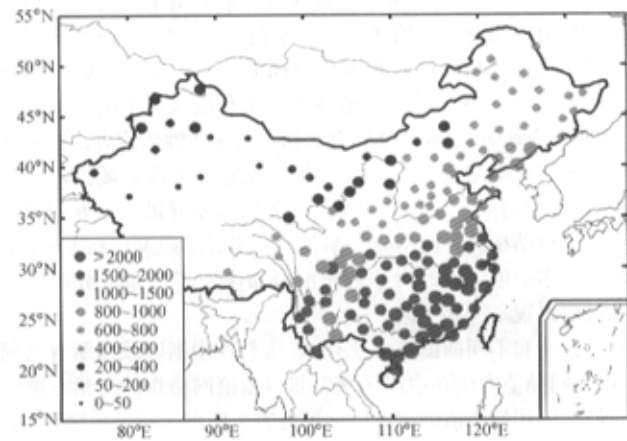


图 3 中国区域 160 个台站 10 年平均年总降水量的分布（单位：mm）

4 不同插值方法 10 年平均年总降水量的比较

本文主要通过以下 2 个步骤来分析比较克里格法、反距离加权法、三角化线性法、

双谐样条法和 Cressman 客观分析法等 5 种内插方法的空间插值能力。首先把离散的 160 个台站资料以一定的空间分辨率内插到规则网格点,与规则网格点的 CRU 资料进行比较;其次再把内插到网格点的值利用双线性插值法反插到原台站和 730 个台站,分别与原始观测资料和更为密集的观测资料进行对比。另外本文还对在不同网格分辨率和影响半径下的 Cressman 客观分析方法进行了比较。

4.1 空间内插后 10 年平均年总降水量的空间分布及其统计特征

把 160 个台站 10 年平均年总降水量的观测资料内插到规则网格点,所取空间分辨率均为 0.5° ,经向范围为 $72^{\circ}\sim 137^{\circ}\text{E}$,网格点数为 131 个,纬向范围为 $15^{\circ}\sim 55^{\circ}\text{N}$,网格点数为 81 个,总网格点数为 10 611 个,其中中国区域内的网格点数为 3 809 个。用来比较的英国 CRU 资料空间分辨率也为 0.5° ,其目前被认为是一种较好的降水同化资料。在克里格方法中,变异函数采用指数形式;在反距离加权法中,距离的幂次取为 3;在 Cressman 客观分析方法中,影响半径采用 1、4、7、10 共 4 种,根据台站密度自动选择。不少空间插值方法不仅可以对空间点进行内插,而且还有一定的外插功能,但外插偏差较大,因此我们只分析比较中国大陆区域的插值结果。图 4 为利用这 5 种方法插值后的降水分布情况,其中图 4f 为 CRU 降水资料。

可以看出,各种插值方法所得结果基本上都能反映出中国 10 年平均年总降水量的空间分布形态;但是不同插值方法所得结果在曲线平滑程度和局部区域空间的分布上存在一定的差异,尤其是在观测台站稀疏的地区,如青藏高原和西北西部地区。克里格方法与双谐样条插值所得的结果与 CRU 比较接近,反距离加权法的空间分布形态较好,但是等值线的平滑程度较差,在西北地区出现较多的“靶心”,这 3 种方法都能将所有要插值的空间格点进行插值处理,即插值结果充满整个空间。三角化线性插值在插值台站所组成的所有三角片之外的区域没有插值结果,Cressman 客观分析方法在台站比较稀疏而网格点又远离台站的区域也没有插值结果,这对于研究某些问题会带来一定不便。在台站最密的 A 区,除了等值线的平滑程度外,不同方法之间没有明显的差异。在 B 区,除了 Cressman 客观分析方法和三角化线性插值方法在东北东北部边缘地区出现空值外,不同方法之间并无太大差异。在整个中国区域内,不同方法的插值结果之间差异最大的是 C 区。在 C 区,50 mm 等值线差异最大,有的呈片状,有的呈点状;由于这一地区台站稀疏,网格分辨率和影响半径的限制使得 Cressman 客观分析方法在该区域出现很多空值,同时台站位置的分布决定了三角剖分插值在该区域的青藏高原西南部地区也出现了一些空值。

对原始 160 个台站和不同插值方法及 CRU 资料在中国区域内所有共同非空插值网格点 2771 个(即不同插值方法在中国区域内非空插值网格点的交集)进行所有点的平均值、标准差、最大值和最小值的统计,结果为表 1 所示。对于平均值和标准差,由于

表 1 原始台站与 5 种内插结果及其 CRU 资料的基本统计情况

统计量	原始台站	克里格法	反距离加权	三角化线性内插	双谐样条	Cressman 客观分析	CRU 资料
平均值	869.0	669.8	686.3	667.5	664.7	665.9	686.2
标准差	530.7	486.2	472.6	483.6	492.1	495.3	497.6
最大值	2576.9	2475.2	2574.3	2481.0	2516.0	2569.5	2378.6
最小值	16.9	16.9	17.5	19.6	-36.6	-97.2	27.2

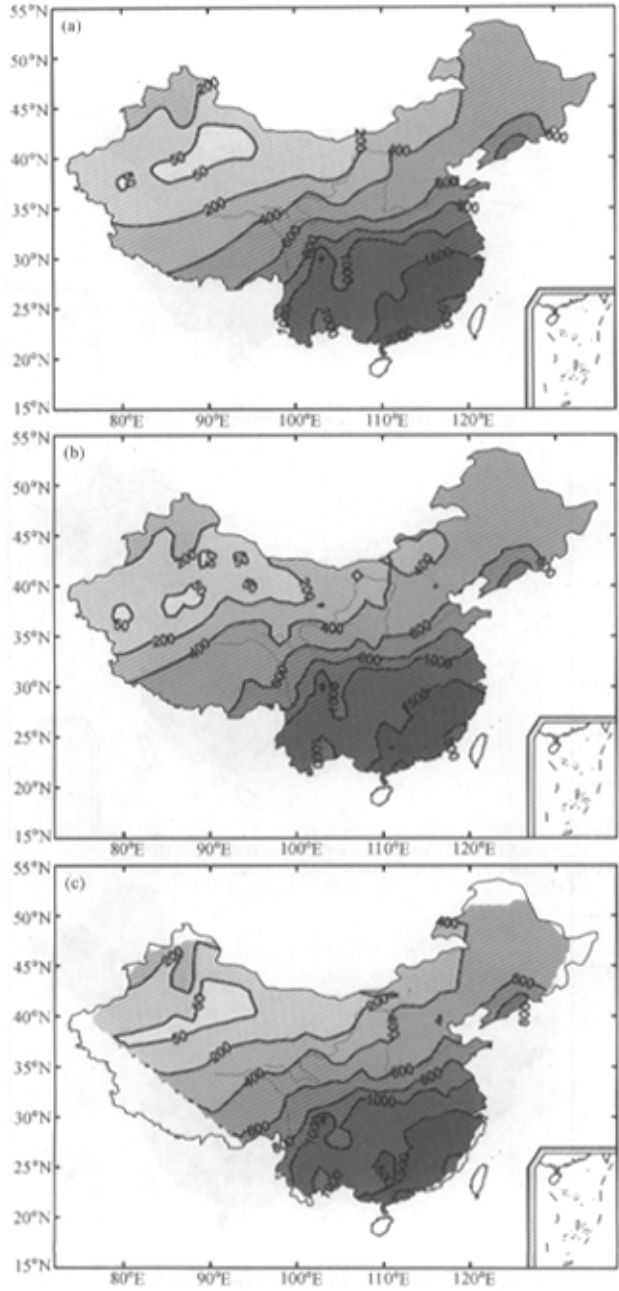


图 4 5 种空间内插结果和 CRU 资料的降水分布 (单位: mm)

(a) 克里格插值法; (b) 反距离加权法; (c) 三角化线性内插; (d) 双谱样条插值; (e) Cressman 客观分析; (f) CRU 资料

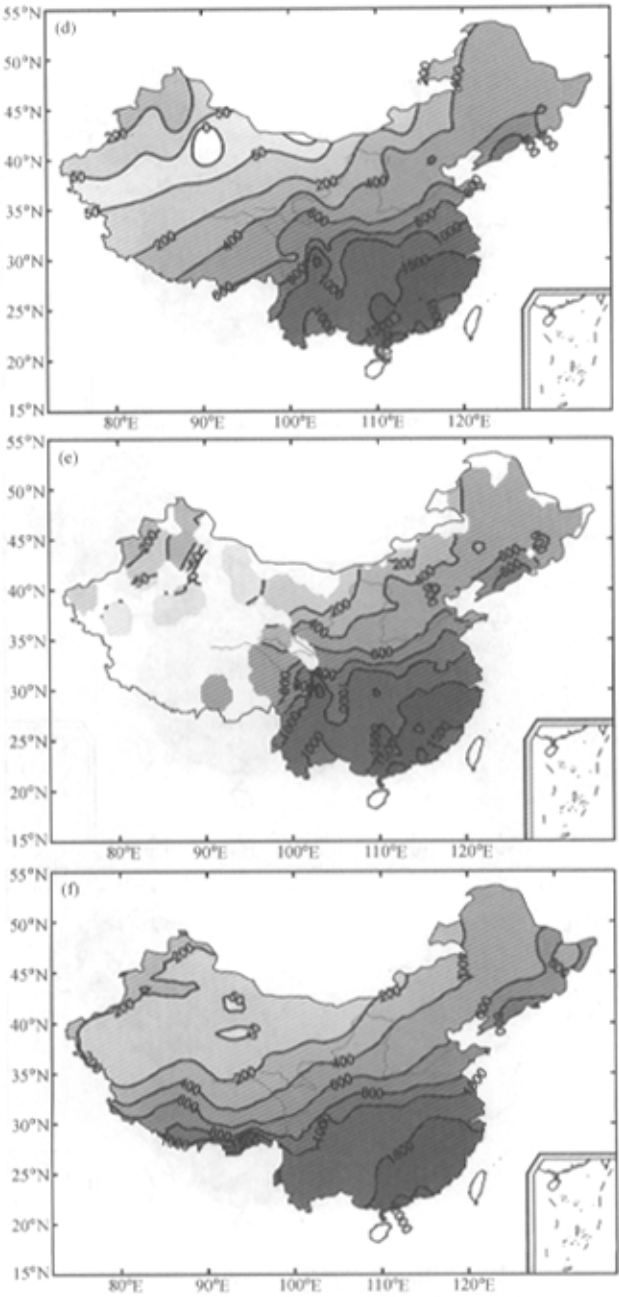


图 4 (续)

台站与插值网格点的样本量和地理位置的差异，二者之间不能直接对比。在 5 种插值方法中，反距离加权法的空间平均降水量最大，比其他 4 种方法的结果大 17~22 mm，它的最大值也最大，与原始台站的最大值最为接近。而克里格法的最大值比其他 4 种方法都要小，这是由于它在插值处理时平滑掉了原始数据的一些峰值信息。对于双谐样条和 Cressman 客观分析 2 种方法，它们的最小值达到负值，这对于降水是不合理的，在实际应用中应予以适当处理。CRU 资料的均值与反距离加权法最为接近，标准差比任一插值方法都大，但最大值比任一插值方法都小，而最小值比任一插值方法都大。

4.2 不同方法的插值结果反插到原始台站后的对比分析

在采用空间内插方法对台站资料进行插值时，如果台站位置正好落在插值网格点上，则对于本文所讨论的 5 种插值方法，若不附加任何平滑处理，此网格点的插值结果应该与台站原始值一致。为了检验插值后的结果与原始台站值有多大偏差，我们将已经插值处理好的网格化数据再利用双线性插值方法反插到原始台站上做进一步分析比较。对于分辨率较高的网格数据，利用双线性方法将其插值到某一点上时所产生的误差较小。

表 2 所示为原始 160 个台站和将 5 种内插方法的结果及 CRU 资料再反插到原始台站后的统计结果，可以看出，Cressman 客观分析方法与原始台站最为接近，反距离加权和双谐样条方法次之，而克里格与三角化线性方法则与原始台站偏离最大。另外，对于平均值、最大值和最小值，CRU 资料比任何一种插值方法的偏差都要大。图 5 所示为不同降水量等级所对应台站数的分布情况，容易看出，三角化线性插值与原始台站偏差最大，尤其是降水量小于 200 mm 的台站数比原始台站数多很多，从而导致其台站平均值偏小最多。图 6 所示为 5 种插值方法和 CRU 资料与原始逐台站降水量偏差百分比，很明显 Cressman 客观分析对于所有台站偏差都很小，其次是反距离加权法、双谐

表 2 原始 160 个台站与 5 种内插结果及其 CRU 资料反插到原始台站后的统计情况 mm

统计量	原始台站	克里格法	反距离加权	三角化线性内插	双谐样条	Cressman 客观分析	CRU 资料
平均值	869.0	869.3	868.9	868.3	869.1	868.9	889.7
标准差	530.7	523.6	527.6	521.2	528.1	530.7	524.0
最大值	2576.9	2448.6	2555.7	2402.3	2529.2	2576.9	2236.8
最小值	16.9	30.3	18.2	30.4	22.9	16.9	65.6

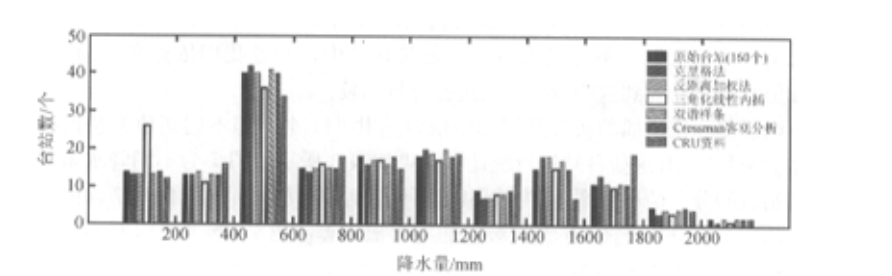


图 5 不同降水量等级所对应台站数的分布（台站总数为 160 个）

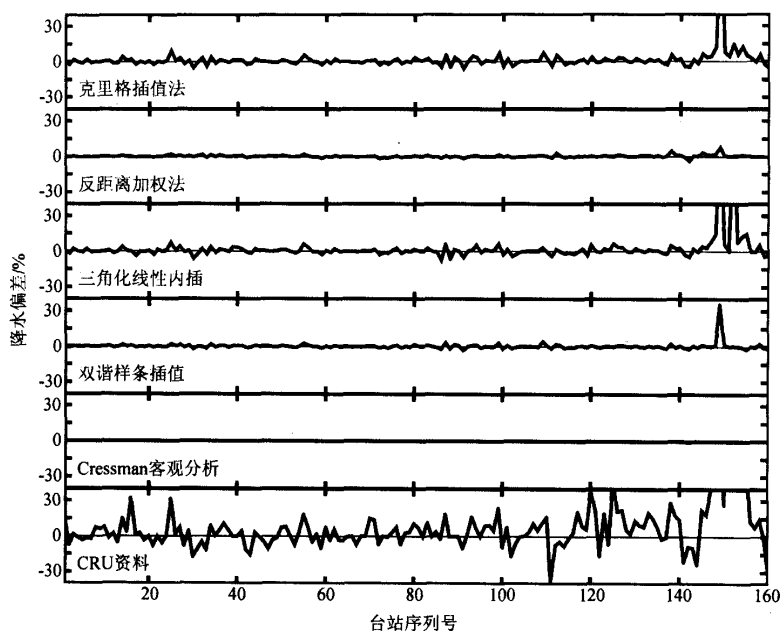


图6 5种内插结果及其CRU资料与原始逐站点降水量偏差百分比

样条、克里格和三角化线性方法。由此可见 Cressman 客观分析方法在数据的客观分析中确实存在着一定的优越性。对于 C 区一些台站, 由于其降水量本身很小, 因此几乎所有方法偏差百分比都很大。而对于 A 区的一些台站, 虽然偏差百分比不是很大, 但降水量的偏差值却很大。

表 1 所示的插值格点平均降水量是整个插值区的区域平均降水量, 而表 2 是由格点反插回去后 160 个台站的平均降水量, 因此, 二者结果不一致但并不矛盾。

4.3 不同内插方法的插值结果与 730 个台站观测资料的比较

在 730 个台站中, 降水资料完整的台站为 631 个。为了将插值结果与这些密集的台站进行比较, 有 2 种方法可以选择, 一是将插值结果再反插到这些台站, 另一是将这些密集的台站利用一种内插方法插值到前面分析所对应的网格点上。后一种方法涉及本文所讨论的主题, 不免陷入循环论证的窘境中, 此处我们依然采用前一种方法, 即将插值结果再反插到这些密集台站进行分析比较。

不同插值方法反插的所有共同非空值台站共 511 个 (即不同插值方法反插的非空值台站的交集), 对这些台站进行统计, 结果如表 3 所示。图 7 为不同降水量等级所对应的台站数的分布情况。图 8a 为按不同区域各种插值方法所得的台站平均降水量与台站观测的偏差百分比, 图 8b 为相对应的均方根误差 (百分比)。可以看出, 所有插值方法所得的台站平均降水量均比密集台站观测的要小, 最大值与标准差也比观测小, 而最小值却比观测大, 这表明由于原始插值台站的限制, 空间内插对于实际降水分布的一些峰值与谷值不能显示出来。虽然台站平均的降水量均偏小, 但某些内插方法所

得到的台站平均百分比偏差却为正值，这是因为在降水量较大的 A 区插值所得的降水偏小，而在降水量较小的 C 区插值所得的降水却又偏大，因而导致了较大的百分比正偏差。很明显，在原始插值台站较密的 A 区和 B 区，各种方法的偏差与均方根误差都比 C 区要小很多。另外，反距离加权法在每个区域都有较大的正偏差。CRU 资料在整个区域的台站平均值比任一插值方法都要大，但与原始台站比较接近，这是因为在降水量较大的 A 区，CRU 资料比原始台站和任一插值方法的结果都偏大较多。但进一步的误差分析表明，CRU 资料除 B 区误差较小外，其他区域的误差都较大（其中 A、B、C 区域划分如图 1 所示）。

进一步的分析表明，在台站密集的 A 区和 B 区，绝大多数台站的偏差百分比都在 10% 以内，在 A 区，其台站大多都为负偏差。在台站稀疏的 C 区，对于克里格、反距离加权与三角化线性方法，大多数台站都有较大的正偏差，少数台站有较大的负偏差。对于双谐波条和 Cressman 客观分析方法，正负偏差分布相当，并且幅度较小一些；CRU 资料在 C 区大多数台站都有较大的正偏差。

通过上述的比较可以看到观测数据的稀疏程度和各插值方法的数学假设对插值结果有很大的影响。克里格方法作为地学第一定律，是建立在一定的空间假设基础上的，

表 3 更加密集台站与 5 种内插结果和 CRU 资料反插到这些台站后的统计结果 mm

统计量	原始台站	克里格法	反距离加权	三角化线性内插	双谐波条	Cressman 客观分析	CRU 资料
平均值	878.4	853.6	863.4	852.6	851.2	852.5	874.0
标准差	533.1	495.4	481.9	492.9	501.5	503.8	505.5
最大值	2773.9	2448.6	2555.7	2402.3	2529.2	2576.9	2286.2
最小值	15.8	30.3	18.2	30.4	22.9	16.9	52.7

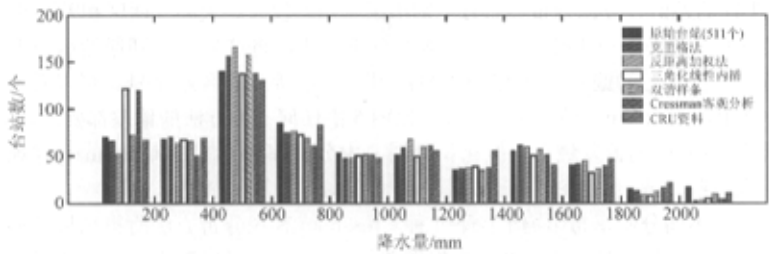


图 7 不同降水量等级所对应台站数的分布（台站总数为 511 个）

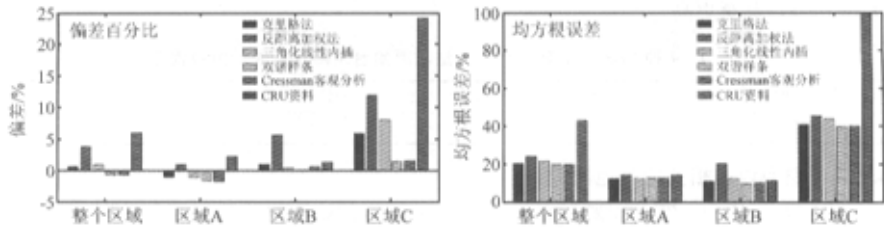


图 8 不同区域台站平均的内插结果及其 CRU 资料与站点观测降水量偏差百分比和均方根误差

在内插过程中平滑掉了很多能够反映局地空间变化的信息；反距离加权法是最简单最易用的空间内插法，其内插的结果不受样本数据空间分布的影响，但是所得到的结果却受到其权重函数幂次的影响；Cressman 客观分析方法作为气象上最常用的客观分析方法，其结果极大地受到数据采样密度和变率的影响；样条插值法适合于非常平滑表面的内插，所以地形梯度对其结果的影响较大；对于三角剖分插值方法来说，采样数据的密度和地形梯度对插值结果有较大的影响。

4.4 Cressman 客观分析方法的进一步比较

对于不同网格分辨率的空间内插，克里格、反距离加权、三角化线性与双谐样条这 4 种方法对某一特定格点的插值不依赖于网格分辨率（如果再将其反插到台站，则依赖于网格分辨率），也就是说所要插值的网格点之间在插值处理时并不相互影响；而 Cressman 客观分析方法在计算时则依赖于网格分辨率，并且影响半径的选取对于插值结果有较大的影响。在此将改变网格分辨率与影响半径来对 Cressman 客观分析方法做进一步的比较研究。

图 9a 是将网格分辨率取为 0.5° ，影响半径取为 4 后 Cressman 客观分析方法所得到的降水分布；图 9b 是影响半径仍为 4，但网格分辨率改为 1° 后的结果，图 9e 为图 9b 与 9a 的差值百分比分布。可以看出，二者之间相差很大，网格分辨率为 1° 的结果虽然缩小了空值区域的范围，但它平滑掉了许多峰值与谷值信息，使得长江中下游地区的降水偏少很多，而华北地区的降水又偏多很多。图 9c 为将网格分辨率改为 1° ，影响半径改为 2 后的结果，它与图 9a 中网格分辨率为 0.5° ，影响半径为 4 的结果极为接近。图 9d 为网格分辨率取 1° ，但影响半径按 $r=1, 4, 7, 10$ 自动选择的降水分布，它与图 4f 中网格分辨率为 0.5° 的结果比较接近，同时又缩小了空值区域的范围。图 9f 为图 9d 与图 4f 的差值百分比分布，只有个别小区域相差较大，大部分地区相差很小。图 10 所示为对应于图 9a~d 4 种 Cressman 客观分析方法，将结果反插到原始插值台站后与原始逐台站降水量的偏差百分比，横坐标为图 1a 中所示的台站序号。可以看出，在这 4 种方法中，图 9a~c 3 种的偏差很大，比图 6 中任何一种方法的偏差都要大，图 9b 的偏差最大；图 9d 的偏差最小，但比相应图 6 中分辨率为 0.5° 的 Cressman 客观分析的偏差要大一些。

表 4 为不同分辨率与影响半径下 4 种 Cressman 客观分析方法所得结果反插到原始台站后相对于原始台站降水量的均方根误差（百分比）。可以看出，自动选择影响半径的客观分析结果均方根误差只有 1%，而固定影响半径的客观分析结果其均方根误差可达到 10%，甚至更高。

表 4 4 种 Cressman 客观分析方法与原始台站降水量的均方根误差 %

分辨率与影响半径	$D_{res}=0.5^\circ, R=4$	$D_{res}=1^\circ, R=4$	$D_{res}=1^\circ, R=2$	$D_{res}=1^\circ, R=1, 4, 7, 10$
均方根误差	9.5	49.9	12.2	1.1

4.5 与空间内插相关的其他问题

在实际工作中，常常需要根据一个时间序列的台站观测资料得到这个时间序列平均的空间场分布，实际上本文所讨论的 10 年平均年总降水量就是这样的问题。这就涉及到是先按台站进行时间平均然后再做空间内插，还是先逐时次做空间内插然后再进

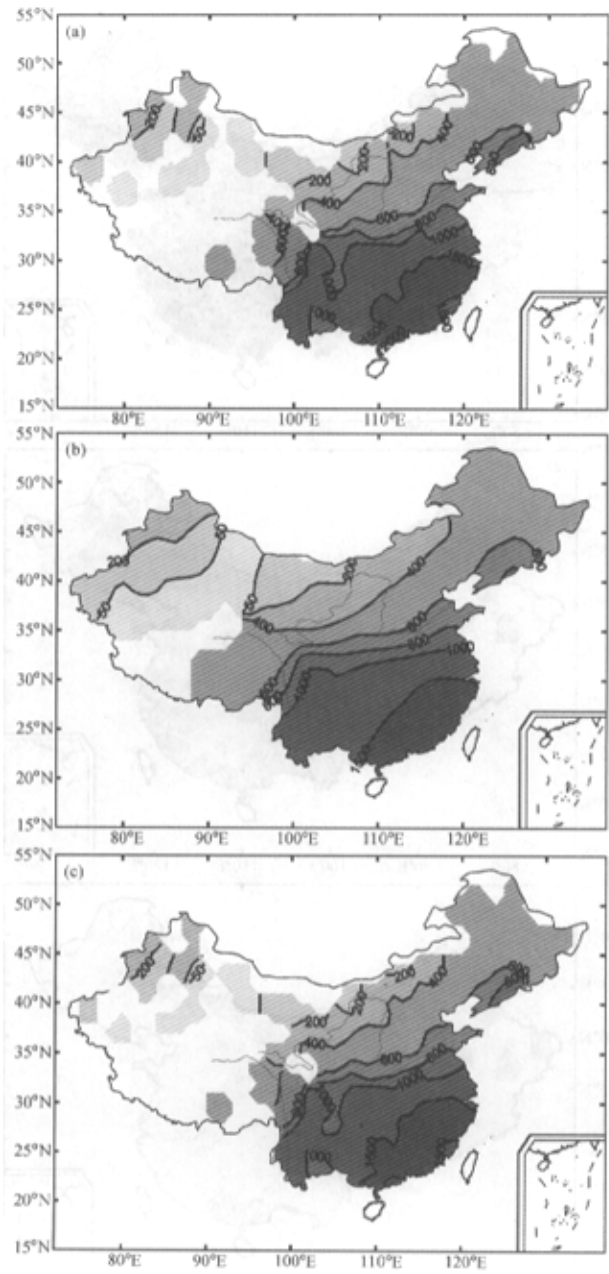


图 9 Cressman 客观分析方法内插结果 (单位: mm)

网格分辨率 D_{res} 与影响半径 R 分别为 (a) $D_{res}=0.5^\circ$, $R=4$; (b) $D_{res}=1^\circ$, $R=4$; (c) $D_{res}=1^\circ$, $R=2$;
(d) $D_{res}=1^\circ$, $R=1, 4, 7, 10$; (e) 图 9b 与 9a 的差值百分比; (f) 图 9d 与图 4f 的差值百分比

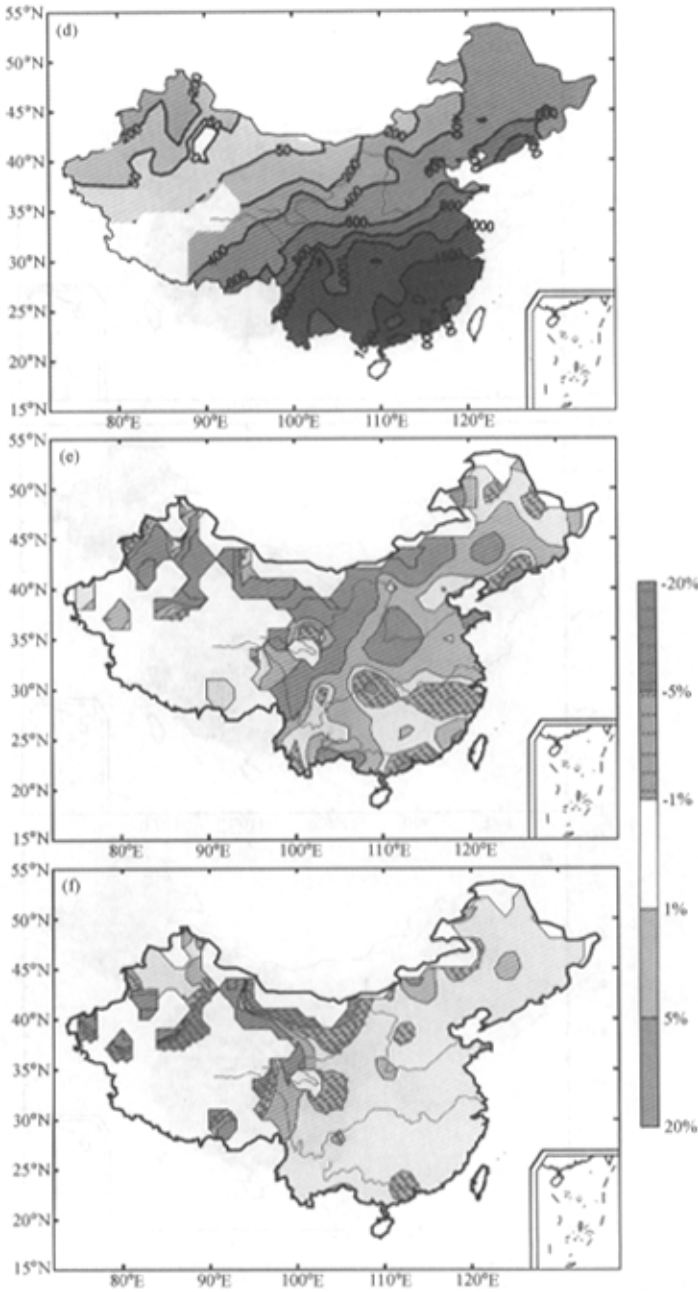


图9 (续)

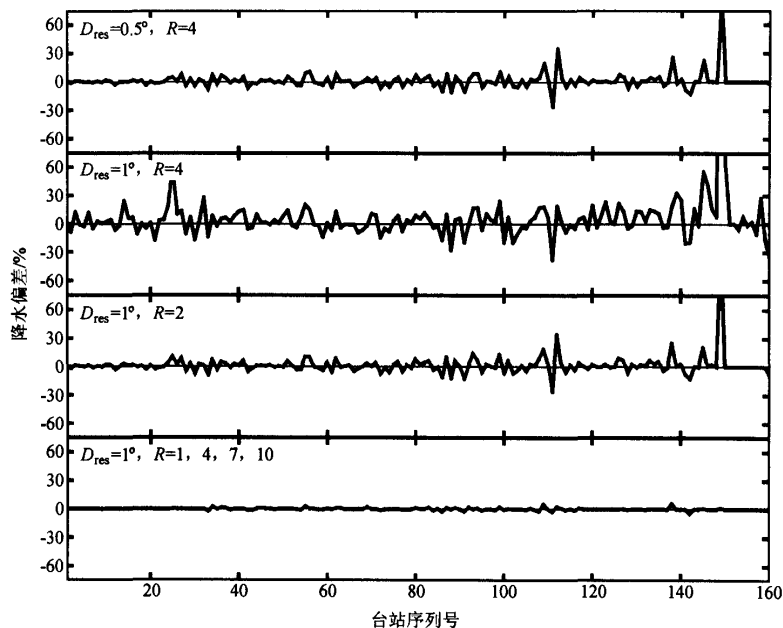


图 10 4 种 Cressman 客观分析方法与原始逐站点降水量偏差百分比

行时间平均，2 种方法之间相差到底有多大？为此，本文选择了 10 年平均的 12 个月逐月 160 个台站的降水观测资料，用它来做 10 年平均年总降水量的空间分布。利用先时间平均后空间内插和先空间内插后时间平均 2 种方法进行对比，从它们的空间分布几乎看不出任何差别，在所有插值的格点中找出这 2 种方法相差最大的格点，统计出 5 种插值方法的二者降水量最大偏差如表 5 所示。可以看出，每种插值方法的二者偏差都非常小，最大也仅为 0.00031 mm，其中三角化线性插值与双谐样条方法这种偏差值更小。因此无论采用什么样的空间内插方法，先时间平均后空间内插和先空间内插后时间平均这 2 种方法偏差都很小，这种偏差已经远远小于观测资料本身的误差，相对于原始台站的量值完全可以忽略不计。

表 5 先时间平均后空间内插和先空间内插后时间平均 2 种方法的最大偏差 mm

克里格插值法	反距离加权法	三角化线性内插	双谐样条插值	Cressman 客观分析
1.3×10^{-4}	1.2×10^{-4}	6.8×10^{-13}	1.6×10^{-9}	3.1×10^{-4}

5 结论

本文利用中国区域 160 个常用台站的降水观测资料，分别采用克里格插值、反距离加权、Delaunay 三角剖分线性插值、双谐样条插值和 Cressman 客观分析等 5 种常见的空间内插方法，较为系统地分析和比较了这几种内插方法插值结果之间的异同，并采用不同的网格分辨率和影响半径对 Cressman 客观分析方法做了进一步研究。主要结

论如下:

(1) 在台站分布密集的区域不同空间内插方法之间差异较小,在台站分布稀疏的区域则差异较大。

(2) 与原始台站的观测资料相比, Cressman 客观分析方法的插值结果与原始台站最为接近,即最“尊重”原始观测数据,反距离加权和双谐波样条方法次之,而克里格与三角化线性方法则与原始台站偏离较大。

(3) 与更加密集的台站观测资料的比较显示根据台站密度自动调节影响半径的 Cressman 客观分析方法与双谐波样条插值方法误差相对较小。

(4) Cressman 客观分析方法会在台站分布稀疏的区域产生很多空值,三角化线性方法也会在所有 Delaunay 三角形外产生空值,这一点给某些问题的研究带来了一定的不便。

(5) 对 Cressman 客观分析方法的进一步比较表明,网格分辨率与影响半径的选取对插值结果有很大的影响,采用自动调节影响半径的 Cressman 客观分析方法比采用单一影响半径的 Cressman 客观分析方法其空间内插结果要好得多。

参 考 文 献

- 1 Ishida, T., and S. Kawashima, Use of Cokriging to Estimate Surface Air Temperature from Elevation, *Theor. Appl. Climatol.*, 1993, **47**, 147~157.
- 2 Dingman, S. L., D. M. Seely-Reynolds, and R. C. Reynolds III, Application of kriging to estimating mean annual precipitation in a region of orographic influence, *Water Resources Bulletin*, 1988, **24**, 329~339.
- 3 Hutchinson, M. F., Interpolating mean rainfall using thin plate smoothing splines, *International Journal of Geographical Information Systems*, 1995, **9**, 385~403.
- 4 冯仲科, 空间数据的最佳内插法 (Kriging 法) 及其在 GIS 中应用的构想, *测绘科技动态*, 1995, (3), 22~26.
- 5 Dubois, G., Spatial interpolation comparison 97: Foreword and introduction, *Journal Geographic Information and Decision Analysis*, 1999, **2** (2), 1~10.
- 6 Collins, F. C., A comparison of Spatial Interpolation Techniques in Temperature Estimation, Proceedings of the Third International Conference/Workshop on Integrating GIS and Environmental Modeling, Santa Barbara, January 21~26, 1996.
- 7 Yukio Sadahiro, Accuracy of areal interpolation: A comparison of alternative methods, *J. Geograph Syst*, 1999, **1** (4), 323~346.
- 8 Zimmerman, D., C. Pavlik, A. Ruggles et al., An Experimental Comparison of Ordinary and Universal Kriging and Inverse Distance Weighting, *Mathematical Geology*, 1999, **31** (4), 375~390.
- 9 Tabios, G. Q., and J. D. Salas, A comparative analysis of techniques for spatial interpolation of precipitation, *Water Resources Bulletin*, 1985, **21** (3), 365~380.
- 10 李新、程国栋、卢玲, 空间内插方法比较, *地球科学进展*, 2000, **15** (3), 260~265.
- 11 苏志侠、程麟生, 两种客观分析方法的比较: 逐步订正和最优内插, *高原气象*, 1994, **13** (2), 194~205.
- 12 姜友谊、黎晓, 数字地面模型内插法的优劣分析, *西安科技学院学报*, 2001, **21** (3), 213~216.
- 13 Sandwell, D. T., Biharmonic Spline Interpolation of GEOS-3 and SEASAT Altimeter Data, *Geophysical Research Letters*, 1987, **14** (2), 139~142.

Intercomparison of Spatial Interpolation Based on Observed Precipitation Data

Feng Jinming^{1,2)}, Zhao Tianbao^{1,2)}, and Zhang Yingjuan^{1,2)}

1) (*Key Laboratory of Regional Climate-Environment for Temperate East Asia*,

Institute of Atmospheric Physics, Chinese Academy of Sciences;

START Regional Center for Temperate East Asia, Beijing 100029

2) (*Graduate School of the Chinese Academy of Sciences, Beijing 100039*)

Abstract How to transfer the scattered observed station data into the regular grid data is vital to the research of climatic change and simulation, which needs spatial interpolation method to achieve. The inter-comparison is made among Kriging, Inverse distance weighting, Delaunay triangulation linear interpolation, Biharmonic spline and Cressman objective analysis spatial interpolation based on observed scattered station precipitation data over China. The advantages, disadvantages and application areas of every spatial interpolation method are discussed. The further research of Cressman objective analysis is also made by altering grid resolution and influencing radius. The results show that biases of different methods are very small in the dense data area and obvious in the sparse data area. The relative errors caused by Cressman objective analysis that influencing radius is adjusted according to the density of stations and Biharmonic spline method are less than other methods.

Key words: spatial interpolation; kriging; delaunay triangulation; biharmonic spline