

# 近百年中国降水的测站资料和格点化资料对比\*

贾朋群

(中国气象科学研究院, 北京 100081)

## 提 要

将 Hulme 的全球陆地格点化降水资料在中国陆地区域内同测站值进行了对比分析. 结果表明, 格点化资料能较好地描述降水场的大尺度特征, 40 多个格点值序列代表了全国近 200 个测站记录, 并具有较高的时空覆盖率. 当仅对中国区域内的降水进行分析时, 为使资料的代表性更好, 需对位于国界附近和沿海地区格点的序列值进行必要的修正.

关键词: 中国降水 测站资料 格点化资料 一致性

## 引 言

降水是全球水循环中的重要因子, 其分布和异常可能导致灾害的发生, 因此备受关注. 进入 90 年代, 虽然全球观测资料状况已大为改善, 但是当利用原始测站资料进行区域或全球降水场分析时, 仍面临资料缺测和边远及不发达地区测站稀疏的问题. 如何利用已有测站网记录给出客观真实的地表降水场分布, 一直是气象科学研究的基本问题之一<sup>[1]</sup>.

近几十年来, 许多学者为给出全球和区域性格点化降水场进行了研究<sup>[2]</sup>. 客观的、格点化和高分辨率降水场数据集是建立在原始测站资料的基础上的, 在对原始记录进行某种修正和改进后, 插值到经-纬网格点上, 能够较好地解决测站分布不均和测站记录偶然缺测等问题, 可用于对大气环流模式结果的验证及对全球和区域水循环的研究. 本文针对中国区域的情况, 以 Hulme 给出的全球  $5^{\circ} \times 5^{\circ}$  经纬网格点资料<sup>[3]</sup> (更新到 1900~1995 年) 为例, 对格点化降水资料在中国陆地地区的表现进行了验证, 对其特征及可用性和可能的改进方法提出了看法.

## 1 地表经纬网格点降水资料

许多作者采用不同的方法建立了时间和空间分辨率不尽相同的全球或全球陆地降水经纬网格点降水资料. Legates 总结和比较了其中有代表性的 3 个全球降水和 3 个全球

\* 由国家科委“九五”重中之重项目“我国短期气候预测系统的研究”96-908-01-06 专题和国家人事部“非教育系统留学回国人员科技活动 A 类”资助.  
1997-12-11 收到, 1998-04-22 收到修改稿.

陆地降水资料集<sup>[2]</sup>. 这6个资料集除分辨率和覆盖区域有所不同以外,从得到的方法上可分为两种,即网格点降水资料来自数值化的人工分析的降水图(2个),或是用数学方法对测站降水观测值进行插值得到(4个). 根据降水分析图的方法可以充分利用一切可以利用的资料,而个别测站的缺测和不连续记录的影响也小的多. 用这种方法得到的格点化资料,其空间分辨率受测站密度的影响也减小到最小. 同时,一张好的降水分析图可以将地面状况很好地考虑进去,这一点对于地形起伏较大的山区和降水这种局部变率大的要素场尤为重要. 当然,也正是这一点构成了该方法的主要缺陷,即人为因素的影响,因为不同的人给出的降水分布分析可以存在相当大的差异. 而利用测站资料构建网格资料时,一般都对台站进行十分严格的筛选,特别是要滤掉一些在20~30年的参考期内记录不完整的测站记录,使得原本就较为稀疏的测站网资料无法全部用上. 但后一种方法的客观性和完整性也是前一种方法不可比的,而且该方法还有很大的改进余地. 例如, Hulme 的资料集为了更多地利用测站的不完整记录,针对不同测站群先给出不同的参考期,分别求出格点序列的距平值,再利用不同参考期的重合时段内两组资料的特征,可以将参考期不同的测站资料全部用上.

虽然各种格点化降水资料之间还有较大的差异,但这种差异主要集中在海洋区域,而全球陆地格点降水资料之间具有相当好的一致性. 本文引用的 Hulme 给出的 1900~1995 年全球陆地  $5^{\circ} \times 5^{\circ}$  经纬网格点降水资料是其中较有代表性的一个. 特别是该资料近几年不断更新和改进,资料的使用者也越来越多. 格点资料给出的各格点月平均降水量,是根据格点区域内符合一定条件的测站实测值按站点与格点中心距离倒数加权平均得到的. 全球 780 个格点资料是根据全球陆地和部分海岛 10400 个以上测站的观测记录生成的.

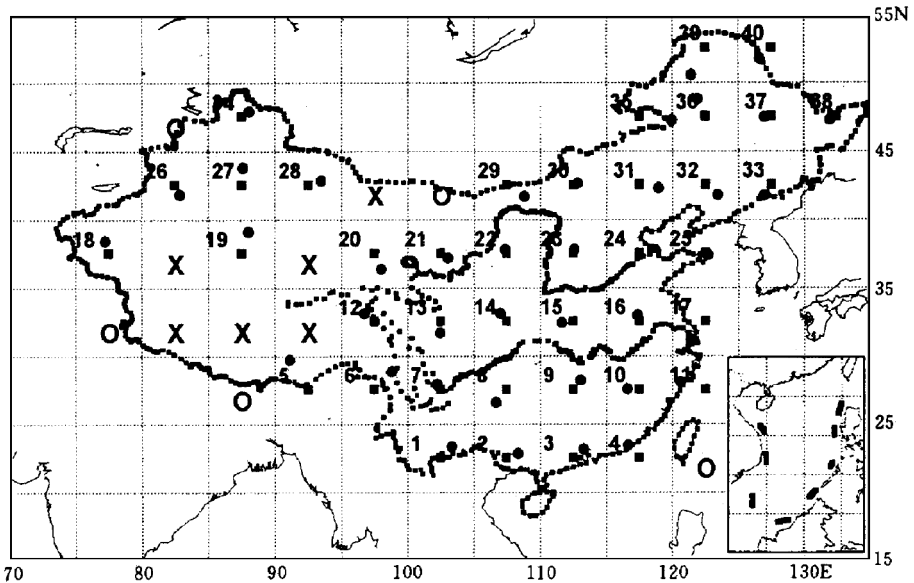


图1 中国陆地区域的格点和对应测站的位置

(每个格点左上角为有记录格点的编号,●测站,■格点,×缺测格点,○周边格点)

中国所在的亚欧大陆东部,位于 Legates 援引的 6 个降水资料集和其他格点资料覆盖率较为完整的北半球中纬度地区. 我国历史和现代丰富的观测资料对全球气象要素场的恢复和建立起着相当重要的作用. 在 Hulme 的资料中,位于中国大陆上的格点应有 46 个,实际有资料格点为 40 个(编号为 1~40,见图 1). 这些格点值利用了国内大约 192 个测站的降水记录. 6 个缺测格点主要位于青藏高原的西部地区. 图 1 还标出 5 个周边格点,这些格点的范围内包含了较多的我国国土,在研究中国降水时,这些格点值也应适当加以考虑. 40 个有资料格点的完整性较好,有 22 个格点资料始于 1900 年 1 月,另有 12 个格点始于 1910 年以前,只有 1 个格点始于 1951 年. 36 个格点序列结束于 1995 年,其余 4 个格点分别结束于 1991 年(1 个)、1993 年(2 个)和 1994 年(1 个).

## 2 中国大陆区域内格点化资料与测站资料的比较

到目前为止,各气象要素资料的全球格点化最为成功的是地表气温资料<sup>[4]</sup>. 这不仅因为气温资料的时空覆盖最全,更为重要的是气温本身分布的均匀性较好. 而其他要素,如降水的时空变率要大得多,而且受地形等因素的影响也更为显著,这就导致单点观测值的代表性相对差一些,在进行格点化处理时需要考虑的因素也较多. 另外降水的零下界也意味着降水序列的非正态分布特征,降水高值区也即是降水方差的高值区. 这些特点均应在降水资料的格点化过程和格点化资料的使用中加以注意.

### 2.1 资料处理

为了验证格点降水资料在我国大陆区域内的表现,在每个有资料格点区域内选择一个最靠近格点中心的测站(图 1),用该测站的观测值与格点值进行比较. 当有多个测站与格点中心的距离相当时,则首先选择记录完整的站. 而当记录的完整性也相差不多时,则计算测站序列与格点序列的相关系数,取相关最好的站作为该格点的比较站. 经这种方法确定的格点和与其进行比较的测站共有 5 对. 在以上序列相关系数的计算和下面格点与测站值的比较中,均限定格点和测站资料的时段为 1961 年 1 月至 1990 年 12 月共 30 年. 因为在这一时段内,格点和测站两种资料均没有缺测. 只是在对两种资料的时间变化进行比较时,我们利用了两种资料尽可能长的记录. 还需指出的是,格点资料给出该格点区域内的平均降水并用格点的中心位置标注,本文仅用最靠近测站中心的一个测站值(也基本上是权重最大的测站)进行比较,从具体降水数值上讲,应该允许二者存在一定差别,但它们的空间相对变化和时间变化趋势应该有较好的一致性,而后者正是本文比较分析的重点.

### 2.2 平均降水场

图 2 给出 30 年平均年降水量格点值与测站值的比较. 因为格点的排列顺序是从南向北,由西向东,图中年降水量呈由大到小的锯齿状分布,即大致体现了我国降水东南沿海大于西北内陆的事实. 图中相对误差  $R$  的计算如下式:

$$R = \frac{|P_{stn} - P_{grid}|}{0.5 \times (P_{stn} + P_{grid})} \times 100\% \quad (1)$$

即测站降水量( $P_{stn}$ )与格点降水量( $P_{grid}$ )的差值的绝对值除以二者的均值. 从图 2 可以看

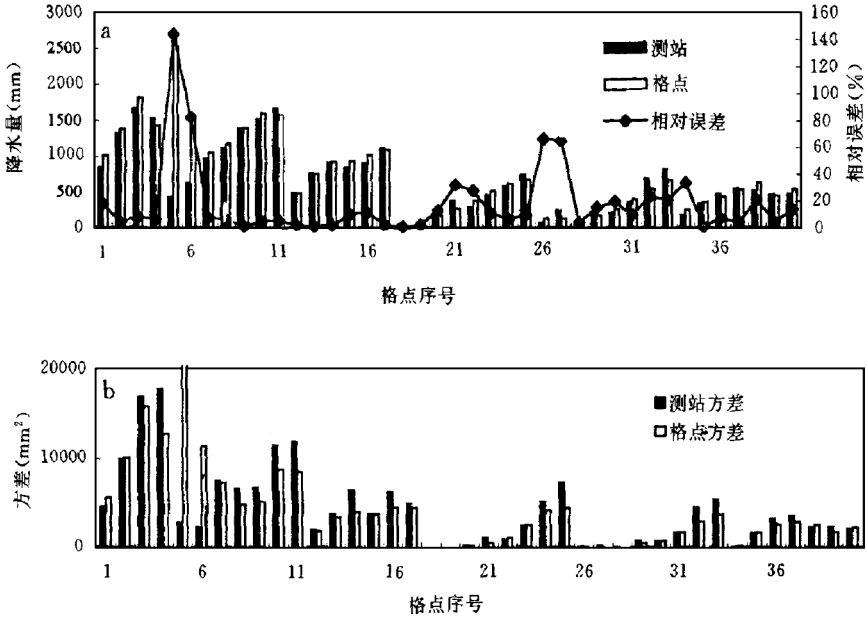


图2 年降水量(a)和降水方差(b)的格点值和测站值

出,两种资料的相对误差基本上都在 20% 以下,降水量较大的点,相对误差更小一些. 需要指出的是,第 5、6 号格点,两种资料的绝对误差和相对误差都很大,如 5 号格点,测站值仅为 432.9 mm,而格点值达 2641.5 mm,为前者 6 倍多. 这是由于该格点中心(27.5°N 92.5°E)位于中国边境线附近,而格点范围内的大部分地区不在中国境内,与格点相对应的测站(29.42°N 91.08°E)位置也偏北,加之该地区地形复杂,造成格点值与测站值存在大的偏差. 6 号格点也有类似的情况,只是误差要小一些.

从两种资料的时间变化的方差分布来看,也基本上同均值分布相似. 但是对均值来讲,两种资料均值的大小比较随机;而方差则大部分是格点资料的方差值小于测站资料的方差值. 这反映了格点值为多个测站值加权平均所得,和单站记录相比,随时间的变化率要小. 总之两种资料在反映中国陆地区域平均降水场方面一致性较好,但位于我国西南边界的两个格点给出的基本上不符合我国区域的降水实况,当只分析国内降水场时,用格点内的测站值代替格点值会更好些.

### 2.3 降水场的时空变化

图 3 比较了各季节平均降水量的格点和测站值. 季节平均降水量用季节内 3 个月的平均降水之和表示,其中春季是指 3~5 月,夏季 6~8 月,秋季 9~11 月,冬季 12~2 月. 从图 3 可以看出两种资料在各个季节均有较好的一致性. 为了检验格点和测站两种资料的时间变化趋势,首先按年降水量将降水序列分为 4 个级别,即大于 1000 mm(一级),450~1000 mm(二级),250~450 mm(三级)和 250 mm 以下(四级). 如果格点值和测站值属不同级别,则归为较大的一级. 按上述标准,得到一级点有 11 个(1~4,7~11 和 16~17 号),二级点 14 个(12~15,23~25,32~33 和 36~40 号),三级点 5 个(21~22,30~31

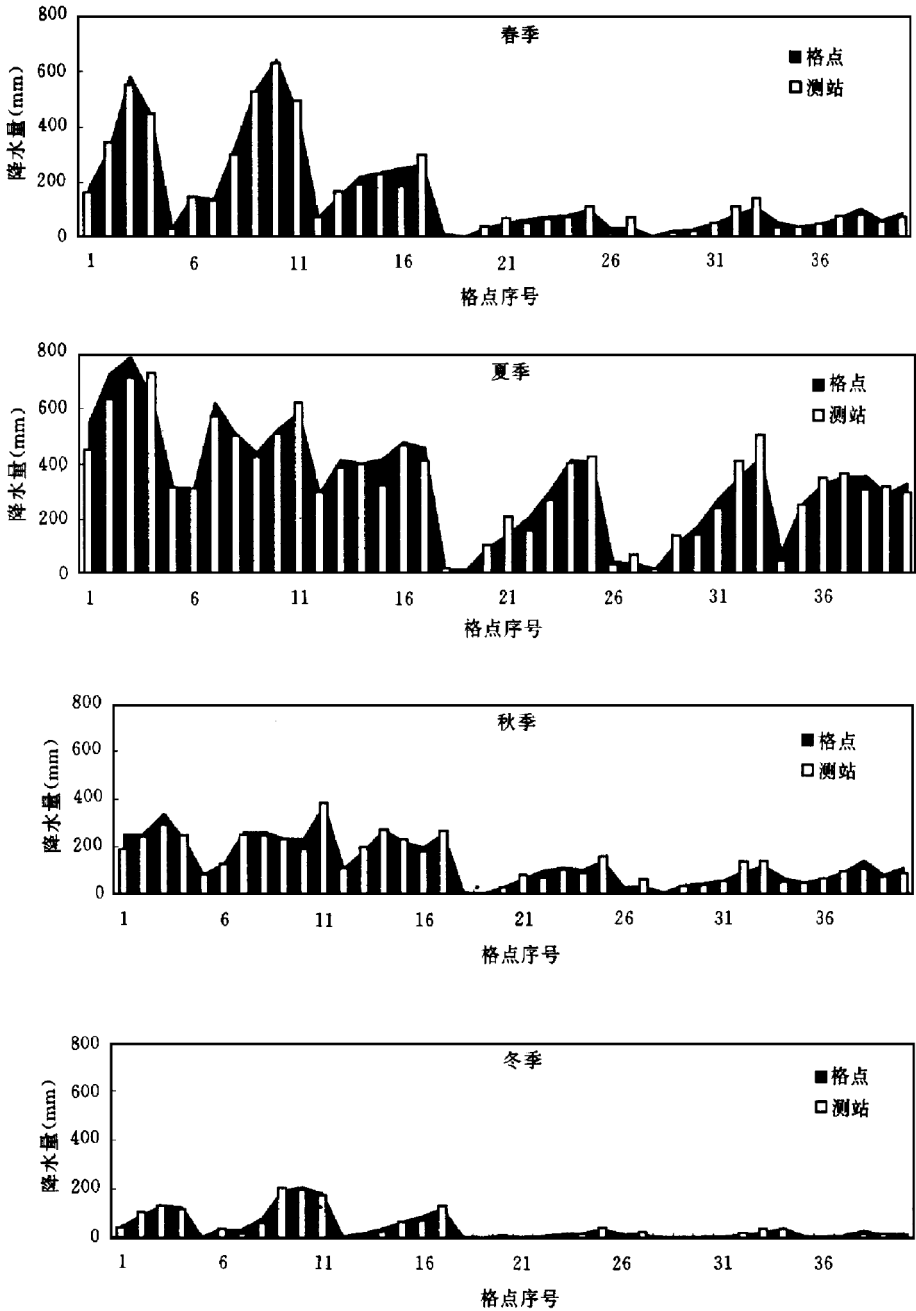


图 3 不同季节平均降水量的格点值和测站值

和 35 号)和四级点 8 个(18~20,26~29 和 34 号)共 38 个点(不包括 5、6 号点)。一级点集中于我国东南地区;二级点包括了青藏高原东部、黄河下游和东北东部;三级点分布于黄河下游、华北西部地区;我国西北地区则为四级少降水区。这一结果与中国降水的区域

划分基本相似<sup>[5]</sup>.

表 1 给出分级后各级别组中区域平均的测站值和格点值的基本特征. 各降水级别组测站和格点的月平均值非常接近. 这些月平均值对应的年降水量分别大约是 1300 mm (一级)、620 mm (二级)、310 mm (三级)和 120 mm (四级). 各降水级别区域内格点值和测站值变化的方差也相当. 一级和三级降水区的格点方差还大于测站方差. 这一结果表明,对于区域平均而言,格点序列和测站序列具有相似均值和平均变化幅度,这使得两种资料的相互代替具备了基本条件. 无论是格点值还是测站值,都没有在 1961~1990 年的 30 年内显示出明显的变化趋势或振荡特征(图略). 不同级别降水区域 30 年平均变化的线性拟合结果,格点值和测站值也有较好的一致性,或者是微弱的增加(一级和四级降水区),或者是变化几乎为零(二级和三级降水区).

表 1 各降水级别测站和格点序列的统计特征

级别	月平均值(mm)		方差(mm <sup>2</sup> )		线性拟合 30 年变化(mm)	
	测站	格点	测站	格点	测站	格点
一级	106.5	110.3	4499.2	4593.8	0.39	0.31
二级	52.1	51.8	2367.2	2320.6	0.01	-0.06
三级	26.2	27.5	847.3	897.7	-0.04	0.08
四级	10.1	9.8	62.2	46.2	0.3	0.19

## 2.4 降水场的历史变化

在图 4 中给出 1900~1995 年格点和测站有效记录数量的变化. 格点的有效记录数即使是在世纪之初,也保持在半数以上,而且在 1925 年以后,缺测格点保持在 5 个以下. 而与格点相对应的测站记录,则缺测较多. 特别是本世纪前 60 年,仅有格点数量的一半左右. 从 30 年代后半期直到 40 年代末,两种资料的有效记录数的差异最大. 这一时期由于战争频繁,肯定缺测很多. 而基于测站记录的格点资料在这一时期并没有受到很大的

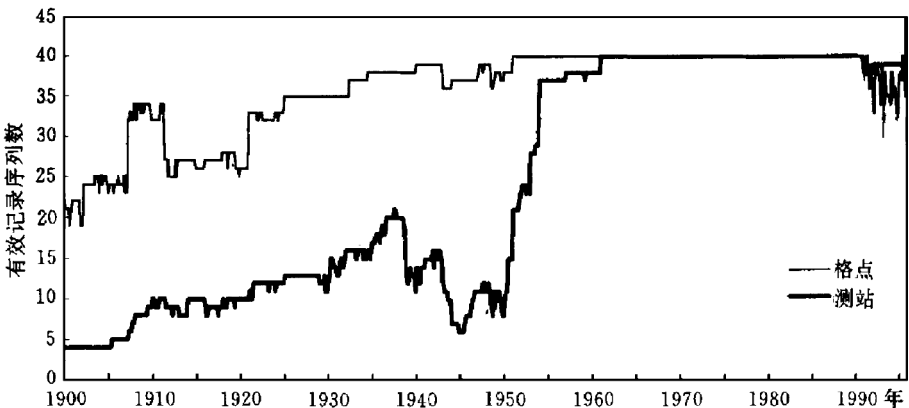


图 4 中国陆地范围内格点和对应的测站有效记录数量的变化

影响,这很可能是运用了某种技术手段对资料进行恢复的结果.从图中看到,只是在1953年后,测站记录数才与格点记录数相当.

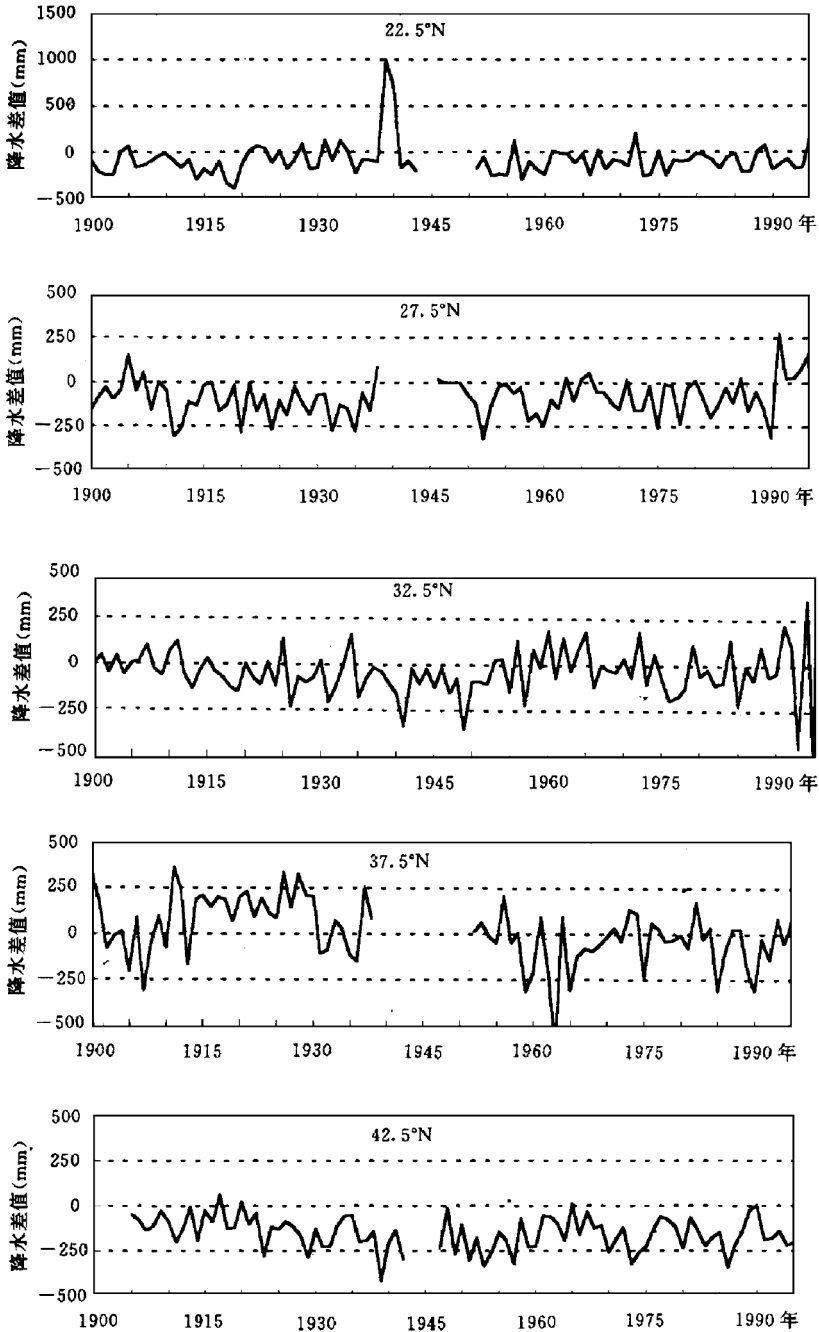


图5 中国东部沿海降水量的格点值和测站值之差的变化

选择降水量较大且有较长记录的我国东部沿海的5个格点,即4、11、17、25和32号格点,它们都位于 $120^{\circ}\text{E}$ 线附近,由南向北从 $22.5^{\circ}\text{N}$ 增加到 $42.5^{\circ}\text{N}$ .在图5中比较了近百年来这些格点值与测站值之差的变化.在沿海地区,一个明显的特征是格点的降水量小于测站记录.在图5给出的个例中,除17号格点和25号格点前期记录以外,均表现的比较明显.这或许是由于在这些格点的范围内,包括了大陆和海洋.格点资料只利用了位于陆地上的测站值,并不反应海洋上的降水状况.而选出的与这些格点进行比较的测站,则大多与海岸线的距离要比格点中心与海岸线的距离更近些,降水量也相对偏多.两种资料的差异在近百年的变化中并没有明显的趋势,这表明它们的时间变化规律是一致的.需引起注意的是,4号和11号格点在前面提到的从30年代后期到40年代末这一时期内,误差值都有一个明显的跳动.这个跳动反映了两种资料都存在的某种不均匀性,它显然同格点化资料因测站资料不全而采用的技术性恢复有关.因为这一时期测站记录太少而无法更加深入地讨论这一问题.

### 3 结果讨论

本文的分析表明,格点化降水资料从整体上看,完全可以代替测站资料对我国陆地降水场进行大尺度的特征分析.由于格点化资料不是针对某一个国家的,所以,当仅对国内降水场进行分析时,应特别对国界附近的格点序列进行分析,以使记录对中国降水更有代表性.文中的比较大多局限于1961~1990年资料覆盖最好的时期,这对于利用格点资料相对于测站资料有效资料时间覆盖长的特点显然是不够的.例如,如上所述,格点资料对30年代后期到40年代末降水量的恢复至少存在着疑点.

格点化资料较好地解决了测站分布不均的问题.由于格点值源于格点内多个测站的观测值,当测站记录偶然发生缺测时,仍能保持格点记录的连续性.从图4中可见,基于近200个测站观测值的40个格点记录,其有效资料的时间覆盖率大大好于用于对比的相同数量的测站记录的时间覆盖率,在60年代以前资料相对匮乏时期尤其如此.

当然,格点资料集的准确性不仅与格点范围内测站数量和记录的质量紧密相关,而且也依赖于格点范围的大小,特别是相对于场变化的空间尺度的大小.对降水场来说, $5^{\circ}\times 5^{\circ}$ 经纬度的格点尺度相对其空间变化还是大了一些,尤其是在低纬山区问题更为突出.文中第5、6号格点值对其范围内北部的中国区域几乎完全没有代表性就是一个明显的例子.本文的结果表明,对5、6号格点值应以相对应的测站值代替才能更好地用于对中国地区的降水场分析.另外,沿海地区格点和测站记录的明显偏差也是在使用格点化资料时应加以注意的.还需指出的是,本文的比较大多基于近30年资料状况较好时期,而格点资料更大的应用价值则是在本世纪前半叶.但因测站资料的不足,很难对二者进行比较.

气象资料的格点化并在气候分析等应用中代替数量庞大且分布不均的测站资料,是一种“浓缩”气象要素场信息的可行方法.中国地域辽阔,降水变化多样,怎样在全球格点化降水场工作的基础上建立中国降水格点资料序列应引起重视.根据中国降水场特征和测站分布的实际情况,在东南沿海等洪涝频繁地区完全可以加密格点,以更充分地利用测站资料反映这些地区降水的多变性.尽管我国降水从整体上讲,变化趋势和振荡特征



均不显著,这一点明显有别于世界许多其他地区<sup>[6]</sup>,但我国与降水有关的灾害不断,说明中国区域内降水的局部变化特征更为重要,这就要求格点化资料在某些特别地区具有更高的空间分辨率。

致谢:对 M. Hulme 博士提供资料和审阅者的细心指正深表谢意。

### 参 考 文 献

- 1 Easterling D R, Peterson T C and Karl T R. On the development and use of homogenized climate datasets. *J. Climate*, 1996, **9**: 1429~1434.
- 2 Legates D R. Global and terrestrial precipitation: A comparative assessment of existing climatologies. *Int. J. Climatol.*, 1995, **15**: 237~258.
- 3 Hulme M. A 1951-80 global land precipitation climatology for evaluation of general circulation models. *Climate Dyn.*, 1992, **7**: 57~72.
- 4 Jones P D and Hulme M. Calculating regional climatic time series for temperature and precipitation: methods and illustrations. *Int. J. Climatol.*, 1996, **16**: 361~377.
- 5 陈菊英. 中国旱涝的分析和长期预报研究. 北京:农业出版社,1991. 1~341.
- 6 Hulme Mike. Recent climate change in the world's drylands. *Geophys. Res. Lett.*, 1996, **23**: 61~64.

## COMPARISON BETWEEN OBSERVATIONAL DATA AND GRID DATA OF PRECIPITATION FOR THE LAST ONE HUNDRED YEARS IN CHINA

Jia Pengqun

(Chinese Academy of Meteorological Sciences, Beijing 100081)

### Abstract

The global grid data of land precipitation are compared with observational data in China. It is shown that the grid data are good to describe the large-scale features of rainfall field. More than 40 grid datasets which have higher spatial and temporal coverage can represent the data about 200 stations. When using grid data to analyse the precipitation in China, it is necessary to modify the grid series near the border and coastal area in China in order to have a better representativeness for data.

**Key words:** Precipitation in China    Observational data    Grid data    Consistency