

刘雨佳, 陈洪滨, 金德镇, 等. 2014. 加密自动气象站雨量计资料的质量控制及其相关关系的研究 [J]. 大气科学, 38 (1): 159–170, doi:10.3878/j.issn.1006-9895.2013.13116. Liu Yujia, Chen Hongbin, Jin Dezhen, et al. 2014. Quality control and representativeness of automatic weather station rain gauge data [J]. Chinese Journal of Atmospheric Sciences (in Chinese), 38 (1): 159–170.

加密自动气象站雨量计资料的质量控制 及其相关关系的研究

刘雨佳^{1,2} 陈洪滨¹ 金德镇³ 齐彦斌³ 崔莲³

1 中国科学院大气物理研究所中层大气与全球环境探测重点实验室, 北京 100029

2 中国科学院大学, 北京 100049

3 吉林省人工影响天气办公室, 长春 130062

摘 要 在长春—四平地区 100 km×100 km 的范围内, 分布有平均间隔 10 km 左右的 147 个自动气象站。结合该区域雷达回波强度资料, 对 2007~2011 年 4~10 月的气象站雨量计小时降水数据进行质量控制。多步骤质量控制结果显示, 有 141 个自动站雨量计的数据通过了检查, 删除了 6 个错误站点的数据, 对有疑问时段的数据作了标记。

利用质量控制后的 5 年夏季半年自动站雨量计小时降水数据, 进行相关关系统计分析表明: 距离在 10 km 以内的雨量计测量, 平均相关系数均能达到 0.6 以上; 雨量计距离小于 5 km, 平均相关系数在 0.7 以上; 而站点距离超过 20 km, 相关系数普遍降到 0.4 以下; 随着统计时间的增长 (从分钟到月降水量), 每个雨量计的测量值具有更高的空间代表性。

关键词 雨量计 质量控制 相关系数 空间代表性

文章编号 1006-9895(2014)01-0159-12

中图分类号 P413

文献标识码 A

doi:10.3878/j.issn.1006-9895.2013.13116

Quality Control and Representativeness of Automatic Weather Station Rain Gauge Data

LIU Yujia^{1,2}, CHEN Hongbin¹, JIN Dezhen³, QI Yanbin³, and CUI Lian³

1 Key Laboratory of Middle Atmosphere and Global Environment Observation, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029

2 University of Chinese Academy of Sciences, Beijing 100049

3 Bureau of Jilin Weather Modification, Changchun 130062

Abstract 147 automatic weather stations (AWS) were deployed in 2006 over an area of about 100 km×100 km between Changchun and Siping in Jilin Province. We performed quality control (QC) of the hourly precipitation data from the densely distributed AWS rain gauges for April to October in 2007–2011. After four QC evaluations and a comparison with reflectivity data from C-band weather data in Changchun, data from six gauges were removed and questionable data from several sites were marked but kept for future use. Rain gauge data correlations were calculated after the QC. The results show that the averaged correlation coefficient (CC) is about 0.6 when the distance between gauges is less than 10 km and up to 0.7 or higher when the distance is less than 5 km. CC generally decreases to less than

收稿日期 2013-01-22, 2013-06-03 收修定稿

资助项目 公益性行业 (气象) 科研专项项目GYHY201106046, 国家自然科学基金40975001

作者简介 刘雨佳, 女, 1985年出生, 博士研究生, 主要从事大气探测及应用研究。E-mail: jjpolo@163.com

通讯作者 陈洪滨, E-mail: chb@mail.iap.ac.cn

0.4 when the distance exceeds 20 km. Furthermore, the spatial representativeness of measurements from a single rain gauge increases when the statistical time scale changes from minutes to months.

Keywords Rain gauge, Quality control, Correlation, Spatial representativeness

1 引言

使用雨量计测量降雨可以获得单点的降雨强度和一定时段的累积降水量。自动雨量计应用后,测量降水更加方便快捷,更具有时效性。从 70 年代开始,国内外科学家开始利用联合雨量计的数据来订正雷达估测值,以期得到雷达有效观测范围内更准确的降水数据。随着研究的深入,越来越多的算法和校正因子需要地面建立高密度的地面雨量计站网,雨量计站网密度增加成为了精确测量区域降水的趋势。高密度雨量计组网可以减少降雨测量时间和空间上的不确定性(Steiner et al., 1999; Villarini et al., 2008; 徐亚钦等, 2011)。美国、日本、英国等国家以及我国的许多省市相继建立了相当规模的高密度自动雨量计站网,但雨量计由于测量技术的不同和资料处理方法的差异,造成了时间和空间上不连续等问题(Rodda, 1968; Ciach, 2003; Yang et al., 2009),并存在一定的测量误差。在进行区域降水测量研究时,即使 1 个用来参考的雨量计出现错误,也会对测量结果造成很大影响。因此,在使用加密雨量计资料前,必须进行严格的质量控制。我国对于地面自动气象站观测资料已经制定了三级质量控制业务规范(任芝花和熊安元, 2007),利用常规的格式检查、极值检查、空间一致性检查等方法对自动站的观测资料进行质量控制。张强等(2007)利用人工测量雨量筒周围分布的自动雨量计数据拟合出标准降雨值,对于自动站的降水数据单独进行质量控制。其结果表明,经过与人工雨量筒测量值的对比,此种方法在降水强度较为均匀情况下十分有效,但在对流云降水条件下,有较大误差。丛芳和刘黎平(2011)结合雷达资料,利用相似离度的分析方法对于雨量计资料进行了质量控制,该方法在覆盖面广、强度大的降雨过程中效果较好。

雨量计密度的不断加大,使得其用来订正雷达估测降水的精度随之提高,并在某一校准密度之后雷达对降水的估测精度趋于稳定(东高红和刘黎平, 2012; Šálek et al., 2012)。对于如何找出雨量计在降水测量应用上的最“适宜”密度,可以通过

分析雨量计之间的相关关系来确定其代表性(Böhner, 1994; Marani, 2003)。这样既有利于优化加密雨量计站网,得到更为准确的降水测量值,又能提高雨量计订正雷达估测降水的精度,也为区域气候环境变化监测、工农业生产、天气预报、云雾降水物理研究、人工影响局部天气的效果检验、水文学上估算江河流量及在汛期及时发布洪水警报等提供更加可靠的资料。

Patrions et al. (1979) 利用分布距离在 50 km 左右的雨量计网的月降雨量对于佐治亚州(Georgia)西北部的降水空间相关性进行了研究,结果表明了位置越近的雨量计空间相关关系越好。Kuzuha et al. (2001) 通过间隔 17 km 的自动气象站网,与分布稀疏的地面观测站进行了代表性分析,结果表明相邻的两种雨量计(地面站和自动站)具有一定的相关性,但当距离超过 100 km 时,则无法利用地面站数据代替自动站数据。Ciach et al. (2006) 利用俄克拉荷马州(Oklahoma)一个 9 km²面积内布置的 1 km 间隔的小型雨量计站网,对于降雨空间相关结构进行了研究,认为高密度的雨量计可以用来代替雷达的估测值得到准确的小范围降雨相关关系。

本文利用分布在长春—四平地区面积为 100 km×100 km 的 147 个自动站雨量计(统称为加密站)2007~2011 年 4 月下旬至 10 月上旬的观测资料,通过多重检查以及雷达资料对雨量计进行质量控制,并分析其质量控制效果,以获得适用于该地区自动雨量站质量控制的有效方法和步骤。然后,对自动站雨量计之间的相关关系进行研究,了解其在不同时间尺度下的空间代表性,对开展降水测量对比试验提出建议。

2 研究区域及资料介绍

吉林省人工影响天气办公室于 2006 年在长春—四平地区布设自动雨量站 96 个,分布在总面积约为 100 km×100 km 的正方形网格上,每个自动气象站相距 10 km 左右。96 个自动气象站(简称人影站,下同)分布在长春市市区、长春市双阳区、公主岭市、梨树县、伊通县以及四平市的东北区域,该地区位于中国东北松辽平原腹地,地势平坦开

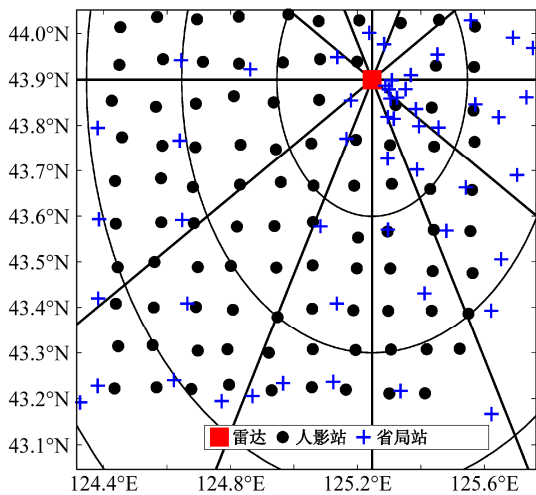


图1 长春—四平地区人影和省局雨量计分布以及雷达所在位置
Fig. 1 Map of rain gauges and radar in the Changchun-Siping region

阔，海拔在 132 m 到 343 m 之间。除 96 个人影站雨量计外，在此区域吉林省气象局布设了 51 个自动气象站（简称省局站，下同），集中分布在长春市市区、市郊和周边乡村，相邻雨量计最大间隔达到 40 km。在长春市市内西南部布设有一部 CINRAD/CC 型多普勒天气雷达，体积扫描时间为 6 min，最大有效探测距离 150 km。全部雨量计与雷达的距离在 3~118 km 之间。图 1 显示两种雨量计分布以及雷达位置情况。

人影站和省局站配备相同型号的温度传感器和自动雨量计。雨量计采用翻斗式雨量感应测定单位时间的降雨量，测量记录的雨强单位是 mm/min，最小分辨率为 0.1 mm，全部雨量计进行统一方法标定和维护（孙海燕等，2006）。除雨量计和温度计外，每个站点均配备一套数据采集器和电源，通过无线通讯方式将数据传送给中心计算机记录与调用。人影站站号排序为 1~100，省局站站号为四位整数（如：1026）。

综合考虑吉林省冬季发生降雪的气候条件、雨量计标定维护部门的记录以及雷达运转情况，本工作选取了 2007 年 4~7 月，2008 年 4~8 月，2009 年 4~9 月，2010 年 4~10 月份和 2011 年 4~10 月份的降水数据，将同一天有测量值的自动站雨量计数据按照站号升序排列，得到每天的分钟降水数据文件。经统计，2007 年 4~7 月共有 24 个降水日；2008 年 4~8 月共有 81 日；2009 年 4~9 月共有 61 日；2010 年 4~10 月共有 77 日；2011 年 4~10 月共有 75 日数据；147 个站点总计有 1685482 个分钟

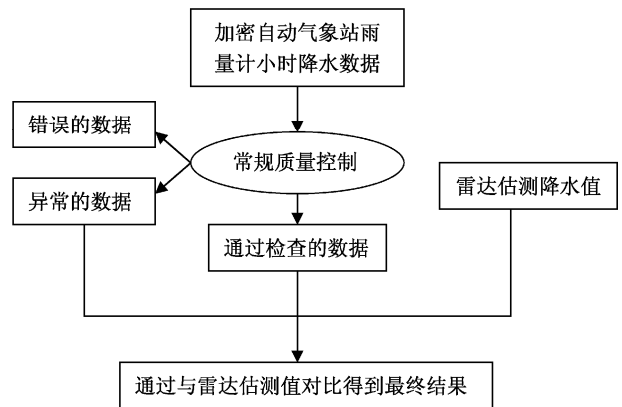


图2 加密站雨量计降水数据质量控制流程
Fig. 2 Flow chart of quality control for the high density rain gauge data

降水数据。考虑到分钟降水数据量较多，时间间隔短，不适宜直接进行质量控制，因此将这些分钟降水数据换算为小时降水数据。

以下对 147 个自动站雨量计的每小时降水数据进行质量控制和分析研究。图 2 为质量控制的主要流程。第 3 节介绍常规检查与结果，第 4 节介绍雨量计与雷达估测降水的比较和结果。

3 利用常规检查对雨量计数据进行质量控制

在长春—四平地区的雨量计日常维护工作中，已经对数据进行了初步的质量控制（孙海燕等，2006），质量控制的主要内容为：（1）通过数据中心的监控与处理，如在晴好天气出现测量值记录则对数据进行删除，并对雨量计进行维护；（2）如在发生明显降雨时雨量计未显示有观测数据，则对该自动站进行维护，在该站维护后至重新标定使用期间的降雨数据进行删除；（3）数据采集系统可以对雨量采集时间进行校对，即中心计算机安装了网络校时软件，如发现雨量计观测时间和中心计算机记录观测时间出入超过雨量采集系统默认范围，则判定雨量计时钟出现问题，要进行校正，出现问题的站点观测数据删除。

根据地面自动站观测资料三级质量控制业务系统中的数据检查方法，结合研究区域自动站雨量计的分布特点，制定了数据连续无变化检查、自动气象站工作情况检查、极值检查和邻近站点比较这 4 个步骤，对 5 年的自动站雨量计降水数据进行质量控制。检查过程中标记为错误的数删除，标记为异常的降水数据保留，并将在下一节通过与雷达

回波强度的对比来验证其可用性。

3.1 数据连续无变化检查

由于测量仪器故障,如雨量计漏斗部分堵塞,承水器收集的降水以匀速渗漏的方式进入翻斗,会造成一段时间内雨量计测量降水值相同(任芝花等, 2010)。根据刘红燕等(2006)的研究表明,即使在层状云降水的情况下,降水强度在1小时内也存在很大的变化,因此可以通过数据连续无变化检查将错误数据挑出。如果某个站点连续出现6个小时或6小时以上降水量相等的情况,则该站点出现连续值的降水数据均记为错误数据(任芝花等, 2010),标记为“W1”。经检查,有4个站点(30号站、80号站、1005号站和4204号站)共40个小时的降水数据未通过检查。我们还对分钟数据进行了同样的检查,如果1小时内分钟降水数据无变化的话,则认为该站该时段的数据错误。检查结果是,没有发现其他站点出现问题。

3.2 自动气象站工作情况检查

对于自动气象站的工作情况检查主要是针对电压和数据采集系统的检查。自动气象站的雨量计和温度计使用同一电源和数据采集系统,温度计在其测量过程中受外界因素的干扰较少,因此可以利用对异常温度数据的检查来判断自动气象站的电源和数据采集系统的工作情况。温度数据的时间和空间代表性都较高,一般而言,相邻两小时的温度数据相差应在 $\pm 8^{\circ}\text{C}$ 之间(杨萍等, 2011)。如果相邻两个小时的温度值差值大于 8°C ,则视为异常数据,根据吉林省4~10月温度变化特点(袭祝香等, 1998),如果出现 -5°C 以下低温或是 38°C 以上高温的测量值时,也视为异常值。经统计,有15个站点共263个温度异常数据,每一年的异常数据占总数据量不超过1%。但异常的温度数据并不代表雨量计出现错误,只是怀疑电压和数据采集系统可能出现问题。因此,只是将温度异常数据相对应的自动站降水数据记为异常数据,标记为“E1”,后面的步骤中会对其进行进一步的检查。

3.3 极值检查

极值是指某个固定测站历史记录中某要素曾出现过的最大值和(或)最小值,气象资料要素值是否超出极值的检查称为极值检查。在雨量计降水资料的统计中,只存在极大值的情况且出现极大值的概率较低。在概率论中我们把概率很接近于0(即在大量重复试验中出现的频率非常低)的事件称为

小概率事件。因此利用小概率事件,找出5年降水数据中的极值标准,以此进行极值检查。将5年数据中相同月份相同站号的自动站大于0的观测值按照从大到小的顺序排列,取其位置在非0观测值数量中排在前1%的降水值(如该站点非0观测值数量小于150,则取其第二位的降水值)作为该站点在该月份的极大值。运用这种方法,求出全部自动站在该月份的极大值,取其中最大值作为5年数据中每个月小时降水极值基准进行检查,如果仍有大于极值基准的降水值存在,则将降水值记为没通过检查的数据。这样可以保证检查出的数据均为概率不大于0.01的小概率事件。在没有通过极值检查的数据中,存在1个小时降水文件中2个或2个以上站点出现极大值的现象,以及一次连续的降水过程中不同站点出现极大值的情况,这都可能是由于强降水云团面积较大及其移动造成的,这两种情况下的极值应视为正常的降水情况。表1为最终极值检查的结果,经统计有28个站点中存在没有通过极值检查的数据,未通过检查的数据共35个,标记为异常数据“E2”。

表1 极值检查结果

Table 1 Results of extremum check

月份	小时降水极值	小时降水数据	未通过极值检查
	(mm/h)	总个数	数据个数
4月	6.4	27371	5
5月	15.1	41753	9
6月	26.3	33075	10
7月	43.8	51247	7
8月	45.2	32444	2
9月	7.4	1631	2

3.4 邻近站点比较(空间一致性检查)

对于区域降水而言,每个参与测量的雨量计之间应存在一定的相关性。邻近站点的比较就是利用这种相关性,找出表现异常的站点。由于降水具有很大的时空不均匀性,因此对降水数据进行空间一致性检查时,应尽量选取距离较近的站点。图3为147个站点与其相邻站点降水值差值比较图。将每个站点2007~2011年可用数据求得总和,然后与其距离最近的4个站点降水值进行比较。图中黑色实心点指示的数值为一个自动站5年降水数据总和分别与其相邻4个站各自降水值总和的差值。由图可见,临近站点之间的差值多数集中在 $\pm 500\text{ mm}$ 之间,但部分站点存在更大的差值。由于研究区域内无明显地形差异,因此可以忽略地形对降水的影

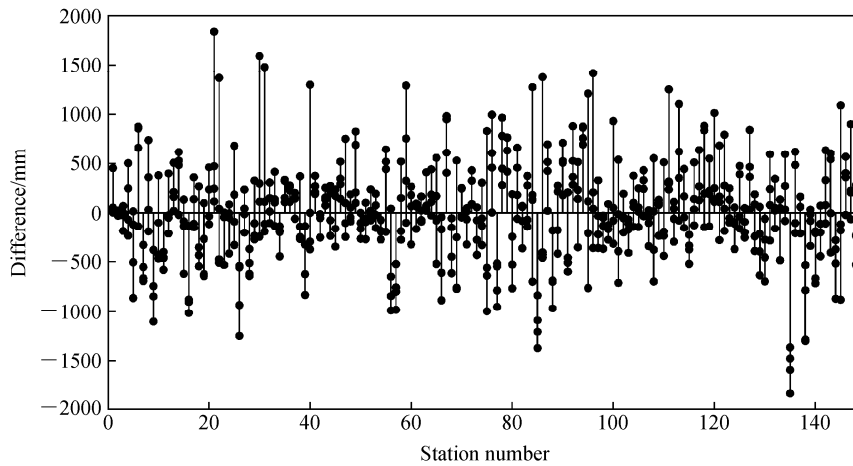


图3 全部站点雨量计与其4个相邻站点雨量计各自5年可用数据降水总和的差值
Fig. 3 Differences of 5 years total rainfall for 147 rain gauges from that at each of their 4 neighbors

响。同样利用小概率事件准则，将全部差值绝对值按照从大到小的顺序排列出，以其总数量前 1% 的位置对应值作为检验值，如果某站点与其相邻站点的差值绝对值超过检验值的个数超过 2 个，则认为这个站点没有通过邻近站点比较。经检查，5 年数据中邻近站点降水总量差值检验值为 1478.5 mm，有 1 个站点（4208 号）未通过此项检验，该站点数据标记为“E3”。

经过以上 4 个步骤的检查，共有 41 个站点中的 298 个数据标记为异常值“E1”和“E2”，1 个站点数据标记为“E3”，而 4 个站点共 40 个小时数据标记为错误数据“W1”。

4 利用雷达数据对雨量计数据进行质量控制

自动站雨量计承水筒内残存雨水或是翻斗误动作（过快或过慢）产生信号，都使得数据采集系统输出错误的降水值。对于这类数据，有时单纯用自身对比检查难以判断，必须使用另一种降水测量方法得到的数据来进行比较。这里，我们使用长春天气雷达的观测数据与雨量计进行对比，并将其作为定性的“检查标准”，如果在常规检查中标记了异常的数据与雷达估测值对比后有很大的差异，则可将异常数据标记为错误数据。

尽管长春站雷达位于 30 m 高度的楼顶，但为了尽量避免地物遮挡，不取最低仰角的观测数据，而是使用 1.5°仰角的雷达反射率（因子）观测数据。又，为了考虑水平风场对降雨数据的影响，雷达回波数据以雨量计为中心并取面积 0.5 km×0.5 km

的正方形格点数据，五个数据进行平均得到与雨量站点相对应的雷达回波反射率因子值 Z （单位 $\text{mm}^6 \text{m}^{-3}$ ）。然后利用关系式 $Z=300R^{1.4}$ （其中 R 为雨强）计算出雷达每次体扫估测的瞬时降水值，求出 1 小时内的降水值均值作为雷达小时估测降水值，计算出的降水值保留一位小数。考虑到雨滴从云底落到地面的时间约为 3~6 mins，因此雷达统计小时降水量时，与雨量计降水数据对应时间提前 6 mins（陈秋萍等，2005）。将全部自动站的雨量计测量值与雷达估测值记录在雨量计对应时间命名的文件中，共得到 3241 个小时的降水数据文件。

4.1 自动站点降水数据与雷达数据的相似离度

由于雷达估测值本身存在较大误差，而且与雨量计测量值之间存在偏差，因此单纯用相关系数和均值差异等统计方法对两者的一致性进行评判无法说明问题。这里，也采用丛芳和刘黎平（2011）引入的相似离度方法对雷达和雨量计进行比较。相似离度即相似性的差异程度，既考虑到雷达估测降雨量和自动站雨量计实测雨量的形相似情况，又体现了它们之间值的差异，是一个比较全面的相似标准。相似离度 C_{GR} 表示为：

$$C_{GR} = \frac{(S_{GR} + D_{GR})}{2}, \quad (1)$$

$$S_{GR} = \frac{1}{N} \sum_{i=1}^N |x_{GRi} - E_{GR}|, \quad (2)$$

$$D_{GR} = \frac{1}{N} \sum_{i=1}^N |x_{GRi}|, \quad (3)$$

$$x_{GRi} = x_{Gi} - x_{Ri}, \quad (4)$$

$$E_{GR} = \frac{1}{N} \sum_{i=1}^N x_{GRi}, \quad (5)$$

其中, E_{GR} 为雨量计测量值 x_{Gi} 与雷达值 x_{Ri} 之间的平均差值; D_{GR} 是雨量计测值与雷达值之差的绝对值平均, 它能反映出两样本之间在总平均数值上的差异程度, 称为值系数; S_{GR} 为雨量计与雷达差值 x_{GRi} 对 E_{GR} 的离散程度 (李开乐, 1986)。如果雨量计测量值与雷达估测值完全相同, 则相似离度为 0。因此, 相似离度越小, 两种测量值相似性越高。

将小时降水数据按照自动站站号排列, 利用相似离度的方法检查每个站点的整体情况。重点对未通过邻近站点检查 (标记 “E3”) 的站点数据进行检查。表 2 为 147 个自动站相似离度的统计。根据概率统计方法, 将相似离度在 $[0, 0.05)$ 范围内的记为 0, $[0.05, 0.15)$ 内记为 0.1, $[0.15, 0.25)$ 内记为 0.2, $[0.25, 0.35)$ 内记为 0.3, 以此类推。从表中可知, 94.6% 的自动站点相似离度集中在 0.1~0.7 之间, 有 8 个站点相似离度超过 0.7。标记 “E3” 的 4208 号站点, 相似离度为 0.9。该站点雨量计测量值与雷达估测值相似性低, 同时与邻近站点测量值差距大。由于该站点与雷达相距 50 km 左右, 且所处地区无明显地势变化, 因此可以判断该站点雨量计测量出现了错误, 将其标记为错误站点 “W1”。其余的 7 个站点中除 69 号和 87 号站外, 5 个站点与雷达相距不超过 10 km, 所处方位集中在雷达的东南方。我们查看了多个晴天时雷达 1.5° 仰角的平面位置 (PPI) 图像, 发现雷达东南部 5 km 范围内存在回波, 而当仰角抬高到最高的 19.4° 时, 回波消失, 因此确定 1.5° 仰角观测时雷达回波受到地物影响, 出现了与地面雨量计测量更大的偏离, 这并不能说明雨量计出现异常, 应抬高雷达仰角重新对比。

表 2 自动站雨量计降水测量值与雷达估测值相似离度统计
Table 2 The analogy deviation of precipitation between rain gauges and radar measurements

相似离度	站点数量	相似离度大于 0.7 的站点
0.1	1	69 号、87 号、
0.2	1	1008 号、1025 号、
0.3	1	1028 号、1027 号、
0.4	16	1005 号、4208 号
0.5	43	
0.6	56	
0.7	21	
0.8	5	
0.9	3	

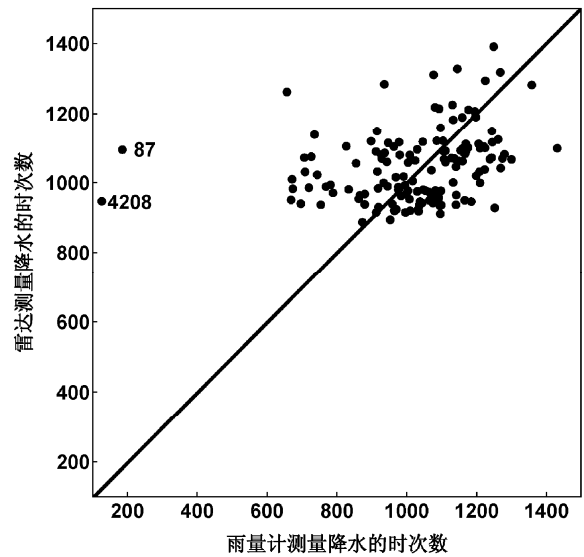


图 4 全部站点 5 年小时降水数据中雨量计与雷达测量到的降水时次数比较

Fig. 4 Comparison of hourly rainfall observation numbers between gauges and radar in five years

4.2 雨量计与雷达测量降水时长的比较

图 4 为 147 个站点 5 年小时降水数据中雨量计测量到降水次数和雷达估测到降水次数的对比。从图中可以看出, 有两个站点 (87 号和 4208 号) 的雨量计测得降水的次数远少于其他站的, 同时较雷达估测降雨的时次数少 909 和 788 次。87 号站点有 44 个小时数据为雨量计测量到降水而雷达未测量到, 有 953 个小时雷达测量到降水而雨量计未测到, 该站与雷达测量时次不匹配数实为 997 个。4208 号站则有 50 个小时数据雨量计测量到而雷达未测到, 838 时次雷达测量到而雨量计未测到。4208 号站点在相似离度检查中已被标记为错误数据, 因此可以进一步证明前面的检查分类是正确的。而 87 号站相似离度为所有站点统计中第二大值 (0.8), 该站点与雷达距离 60 km 左右, 二者之间并没有地物遮挡。在查看了该站点与其周围 4 个邻近站点的 5 年降水总量比较的结果后, 发现该站点较其邻近站点的降水测量差均值为 -1125.5 mm。这说明该站点与相邻站点相比测量值明显偏低, 同时存在雨量计和雷达测量不同步的情况, 因此可以判断这个站点的数据也为错误数据, 记为 “W1”。

4.3 标记异常的雨量计数据与雷达估测值对比

利用雷达估测降水值对标记为 “E1” 和 “E2” 的异常数据进行进一步检查。经过相似离度的比较, 在标记 “E1” 的雨量计数据中, 5 号站数

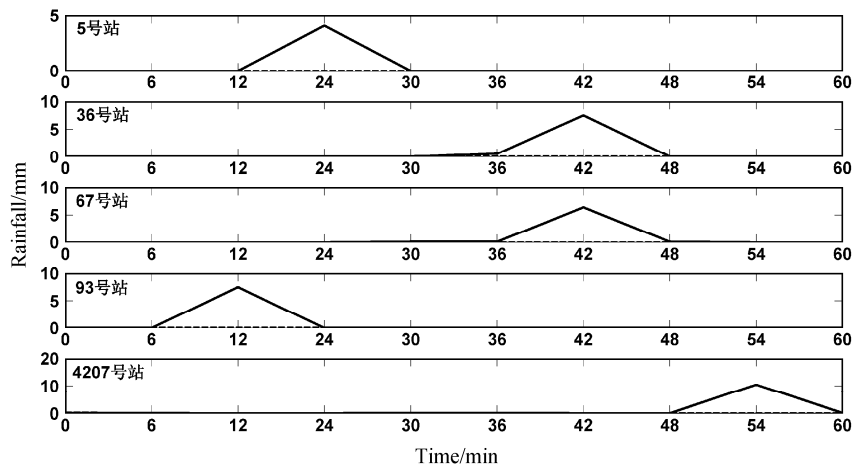


图 5 5 个站点标记 E1 和 E2 数据 1 小时内雨量计与雷达测量值随时间分布 (实线: 雨量计测量值, 虚线: 雷达估测值)

Fig. 5 Time series of rain gauges and radar precipitation marked E1 and E2 at five stations (solid line: gauge precipitation; dotted line: radar precipitation)

据的相似离度最大 (4.2), 其余 14 个站点数据的相似离度均小于 1。而在标记“E2”的数据中由于多数站点对应的雨量计测量值只有 1 个, 不适宜使用相似离度的方法来对比。对于出现极值的站点, 如果其对应雷达估测降水量有测量值, 可以证明该站点在这 1 个小时内存在降水现象, 而非由于仪器误差和人为因素造成的降水测量值。利用这种方法, 对于“E1”中相似离度最大的 5 号站和“E2”的数据进行检查, 结果 5 号站有一个小时降水异常值对应的雷达估测值为 0, 极值中有 4 个站点未通过检查。图 5 为未通过检查的站点在降水过程中雨量计测量值与雷达估测值随时间的分布。从图中可以看出, 5 个站点出现异常的测量值均没有对应的雷达估测值, 同时其出现的前 6 分钟和后 6 分钟也均没有雨量计测量值和雷达估测值, 因此判断这 5 个站点的异常值为错误值, 标记为“W2”。错误数据的时段分别为 5 号站: 08071200 (2008 年 7 月 12 日 00:00, 北京时, 下同); 36 号站: 09041916; 67 号站: 08042318; 93 号站: 09042207; 4207 号站: 08042220。

经过以上的质量控制过程, 吉林省长春—四平地区 2007~2011 年 4~10 月的 147 个自动雨量计小时降水数据中有 2 个站点 (87 号和 4208 号) 的数据及 4 个站点 (30 号、80 号、1005 号和 4204 号) 40 个雨量计小时降水量数据记为错误数据“W1”, 而 5 个站点共 5 个小时数据记为错误数据“W2”。在标记错误的数字中, 标记为“W1”的数据由于无法确定雨量计测量的正确值因此要进行删除, 而对于未通过连续检查的 4 个站点来说, 删除数据会

造成雨量计测量次数的缺失, 对于利用雨量计进行相关关系的统计有影响, 因此这 4 个站点的数据也全部删除, 不用于下一步分析研究中。而标记“W2”的数据因为可以证明该时刻并未产生降水, 因此标记“W2”的 5 个小时数据删除, 其余雨量计数据继续保留。标记异常 E1-E3 的数据继续保留。

表 3 为 2007~2011 年全部站点质量控制前后雨量计测量值和雷达估测值的统计值。质量控制后的数据为去除了标记为“W1”和“W2”的站点和数据。从表中可见, 经过质量控制后, 雨量计测量值与雷达估测值的平均绝对差值减小, 相关系数增大, 相似离度减小, 两种数据的一致性更好。

表 3 雨量计测量与雷达估测降水量的对比

Table 3 Comparison of precipitation measurements between rain gauges and radar

对比参数	$ G-R $ 均值 (mm h^{-1})	$ G-R $ 标准差 (mm h^{-1})	相关系数	相似离度
质量控制前	0.57	2.0	0.14	0.62
质量控制后	0.45	1.8	0.21	0.54

注: G 为雨量计测量值, R 为雷达估测值

5 自动气象站雨量计测量相关关系的研究

长春—四平地区加密雨量站网中, 相邻两个站点的距离多为 10 km 左右; 在长春市区内, 雨量计最近距离可达 2~5 km。因此, 通过对加密站网雨量计测量的相关关系研究, 以考查单站雨量计在不同时间尺度上观测的空间代表性。

图6显示质量控制后不同密度雨量计网2007~2011年5年平均降水量分布,三幅小图采用相同的色标。由图可见,随着雨量站密度的增加,年降水的空间不均匀性得到更为明显的展现,很多站直线距离不到50 km,其年平均降水量的差别大于100 mm。可以说,在年月的时间尺度上常规观测站很难代表方圆几十至上百公里范围的降水量。

5.1 自动站雨量计站网的空间相关

本工作采用不同距离之间自动站雨量数据的相关系数来表征单站降水测量的空间代表性。在研究区域东北部,省局站和人影站分布较为密集,雨量计距离小于5 km。这其中1028号站四周的站点最为密集,因此选取该站为“中心站”,计算该站

与区域内人影站和省局站在2007~2011年测得的小时降水量之间的相关系数(correlation coefficient, CC),结果如图7所示。由图7a可以看到,周围有6个站与其距离小于10 km。从图7b中可见,中心站与其相距小于10 km的站点的相关系数达到0.6以上;当距离从10 km增加到20 km左右,相关系数从0.6逐渐下降到0.4;随距离增加,相关系数减小。

再利用5年的小时降水数据,分别以每个自动站点为中心,可以计算与其他站点的平均相关系数CC、平均绝对差值(absolute difference, DA)和平均均方根值(root mean square, RMS),它们随距离远近的分布分别在图8中给出。由图可见,站点

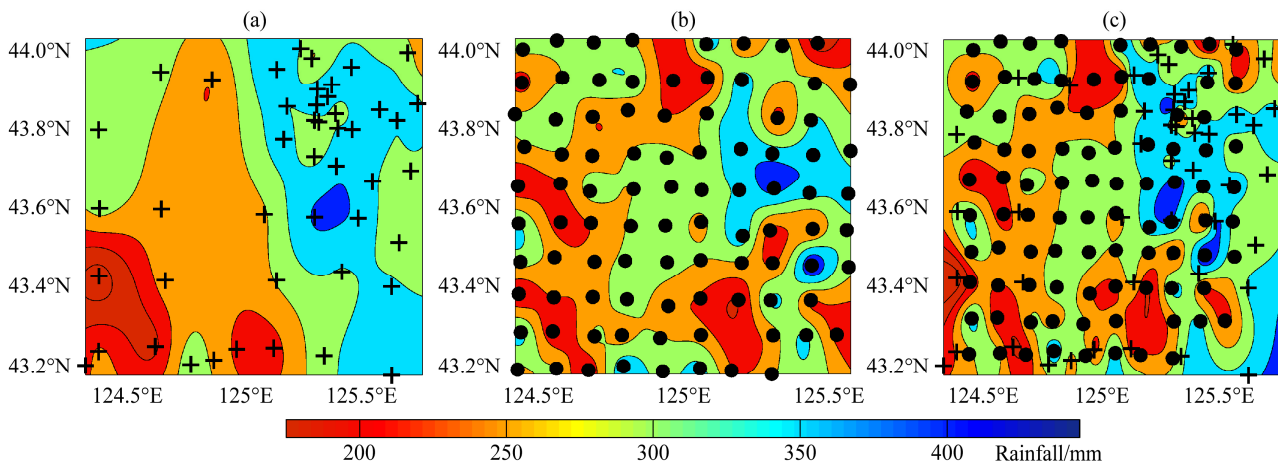


图6 自动站雨量计5年平均降水量分布:(a)省局站雨量计(SJ);(b)人影站雨量计(RY);(c)全部站点雨量计(+为省局站,●为人影站)
Fig. 6 Distribution of the 5-year-average precipitation: (a) Rain gauges, (b) rain gauges, (c) all gauges

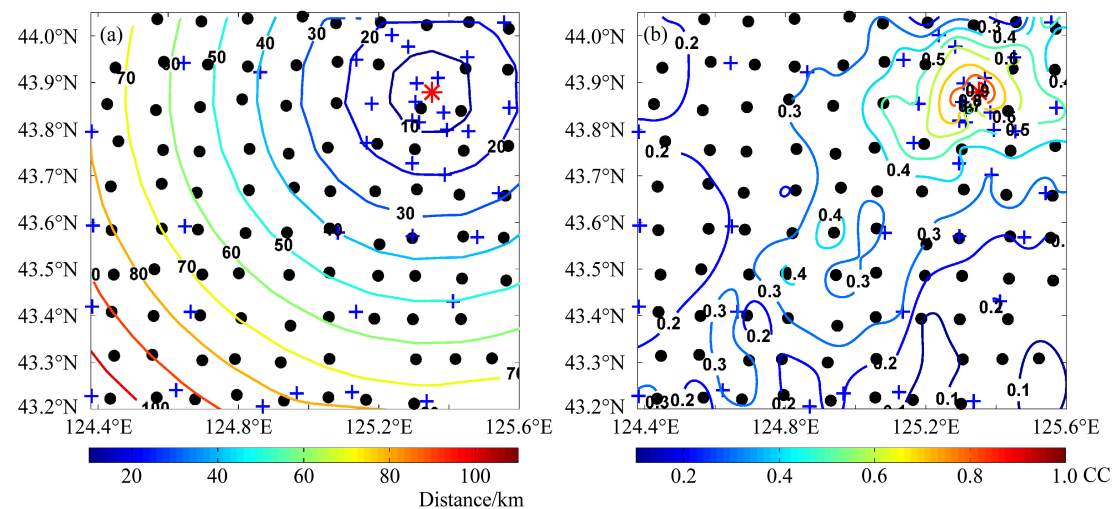


图7 1028号站(a)与其他自动站的距离分布(单位:km)及(b)与其他站测量值相关系数分布(●为人影站,+为省局站,*为1028号站点)
Fig. 7 (a) Distribution of distances from station No.1028 to other 140 stations and (b) correlation coefficients between station No.1028 and other stations (● RY stations, + SJ stations, * station No.1028)

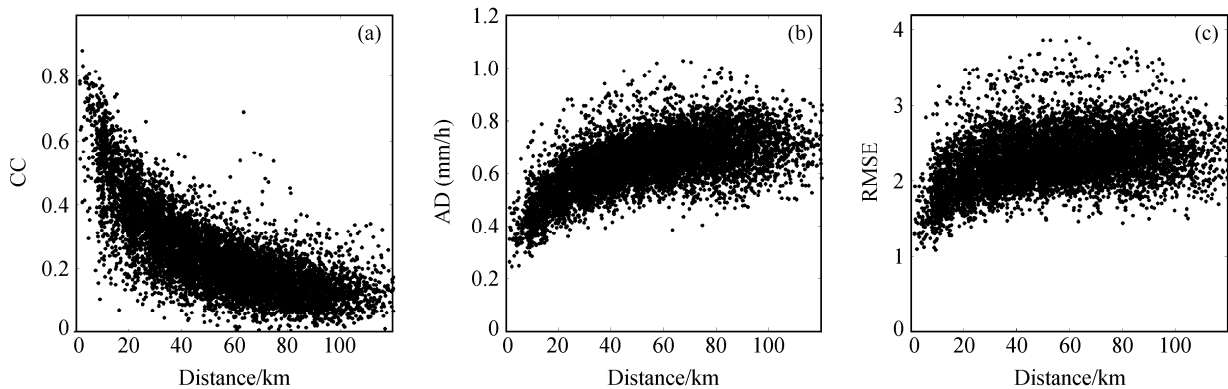


图 8 141 个站点与其他站点 5 年小时降水量的 (a) 平均相关系数、(b) 平均绝对差值以及 (c) 平均均方根随距离的分布

Fig. 8 (a) Distribution of averaged correlation coefficient, (b) absolute difference, and (c) RMS (root mean square) with distance for 5-year hourly precipitation data between every gauge with other 140 gauges

距离越小，相关系数越高。当距离小于 5 km，平均相关系数在 0.7 以上；距离小于 10 km 的站点之间，平均相关系数均能达到 0.6 左右；站点距离超过 20 km 时，降水测值的平均相关系数降到 0.4 以下；当距离大于 60 km 时，平均相关系数小于 0.2。此外，随着站点之间的距离增加，降水测量之间的平均绝对差值和均方根增大。

5.2 不同时间尺度雨量计测量的空间相关性

本节将 2007~2011 年降水数据按照不同时间尺度划分，对雨量计的测量空间相关性进行分析。同样以 1028 号站为基准“中心站”，对其与其他站点在 6 min、30 min、1 h、6 h、12 h 和 24 h 这 6 个时间尺度内降雨数据的相关系数进行统计计算，结果如图 9 所示，其中 1 h 的与图 7b 完全相同。从图 9 可见，随着时间尺度的增长，中心站点与其相邻四周站点测量的相关系数增加。在 6 min 的时间尺度下，中心站点与其相邻小于 10 km 站点的相关系数在 0.4 左右，随着距离的增加，相关系数减小；在 30 min 的时间尺度下，中心站与距离 10 km 以内站的相关系数达到 0.5 左右；在 1 h 的情况下，相关系数增加到 0.6；而在 24 h 的时间尺度下，中心站与周围 20 km 内站点的相关系数都能达到 0.8 以上，同时对于距离超过 40 km 的站点，相关系数的分布也能达到 0.5~0.7。这证明了随着时间尺度的增加，雨量计测量之间的相关性随之提高，空间代表性增大。

在 6 min、30 min、60 min、6 h、12 h、24 h 以及 1 个月 (1 mon) 七个不同时间尺度下，以每个雨量计为中心，对其不同距离的雨量计之间的平均绝对差值、均方根值以及相关系数的均值进行了

表 4 不同时间尺度不同距离雨量计测量比较的统计值

Table 4 Statistics of precipitation measurement comparison between rain gauges at different distances for different time scales

距离/km	不同时间尺度的统计值						
	6 min	30 min	60 min	6 h	12 h	24 h	1 mon
平均绝对差 (mm)							
<10	1.2	1.1	0.9	0.9	0.8	0.6	0.4
10~30	1.4	1.3	1.2	1.0	0.9	0.7	0.5
30~60	1.6	1.5	1.4	1.2	1.1	0.9	0.5
60~90	1.7	1.6	1.5	1.3	1.2	1.0	0.6
>90	1.7	1.7	1.6	1.3	1.2	1.1	0.7
均方根值 (mm)							
<10	6.1	5.8	3.8	2.8	2.2	1.5	0.8
10~30	6.7	6.7	4.5	3.0	2.3	1.7	0.9
30~60	7.1	7.2	5.0	3.4	2.6	2.0	1.0
60~90	7.3	7.5	5.3	3.7	3.0	2.3	1.1
>90	7.4	7.7	5.4	3.8	3.0	2.4	1.2
相关系数							
<10	0.32	0.46	0.56	0.61	0.65	0.81	0.85
10~30	0.17	0.27	0.37	0.55	0.60	0.75	0.83
30~60	0.10	0.16	0.23	0.45	0.51	0.67	0.79
60~90	0.07	0.11	0.16	0.38	0.45	0.61	0.76
>90	0.06	0.10	0.13	0.38	0.44	0.57	0.73

统计，结果在表 4 中给出。从表中可见，在整体情况下，平均绝对差和均方根随着站间距离的增加而增大，相关系数随着距离的增加而减小。随着时间尺度的增加，平均相对误差和均方根值均减小，而相关系数增加。在日 (24 h) 降水量的统计中，10 km 以内距离的雨量计测量的相关系数在 0.8 以上，平均相对误差和均方根也相对较小。在月降水量的统计中，相隔距离在 30 km 以内的站点相关系数都能达到 0.8 以上，同时平均绝对差和均方根均为各自统计中的最小值。表明七个不同时间尺度的降水资料统计中，距离 10 km 以内的雨量计之间日降水量的相关性较好，雨量计单站测量在方圆 10 km 范围

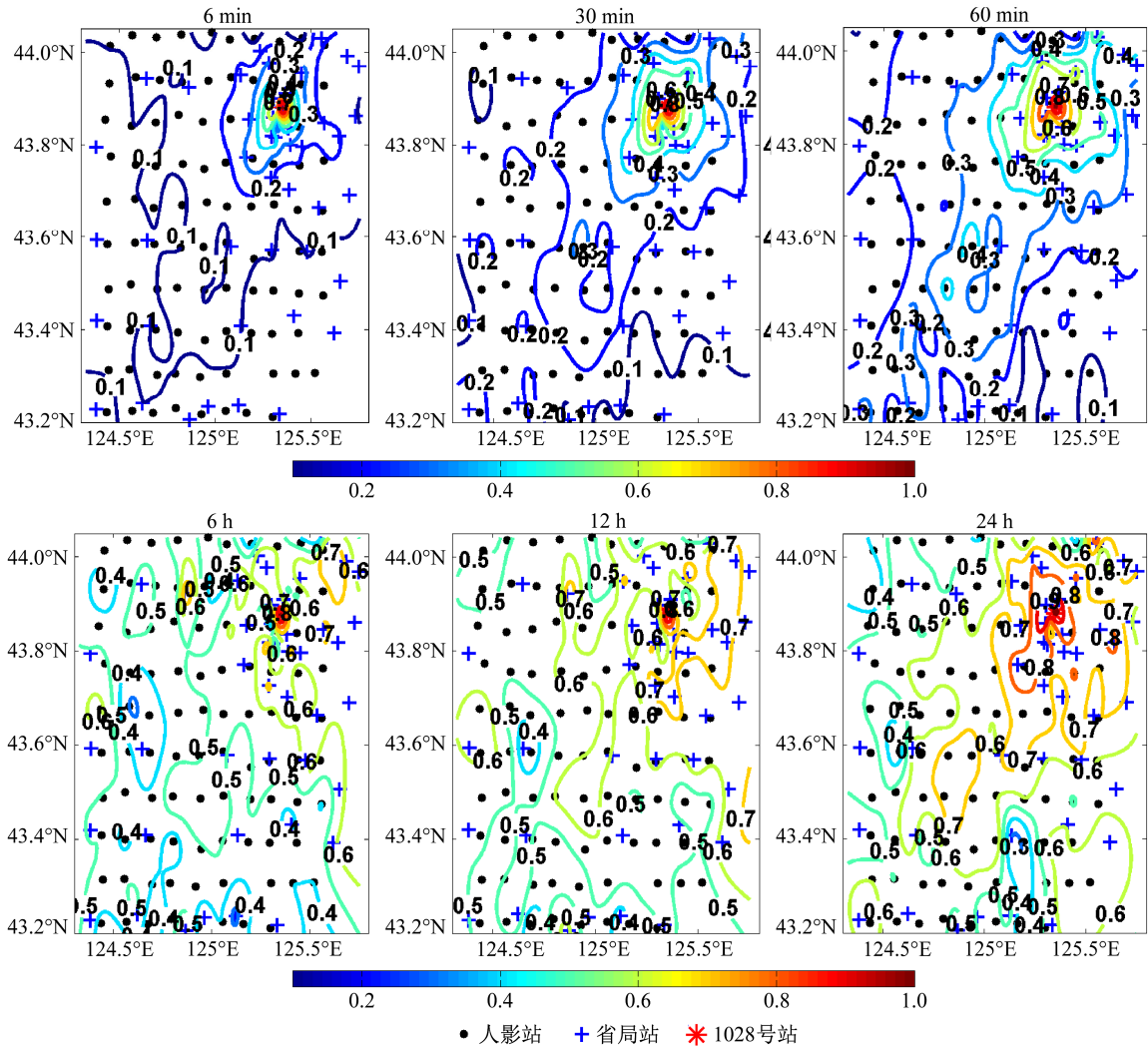


图9 不同时间尺度下雨量计降雨数据的相关系数分布(以1028号站为基准)

Fig. 9 Correlation coefficients distribution of rain gauge rainfall for different time scales (based on station No.1028)

内有较好的代表性。而在月降水量的统计中, 30 km 以内的雨量计之间具有很高的相关性, 雨量计测量值的空间代表性可以方圆 30 km。在其余 5 个短的时间尺度下, 雨量计之间虽有一定的相关性, 但是相邻雨量计测量的相关系数都小于 0.7; 时间尺度越小、距离越大, 相关系数越小, 单站测量的代表性越低。

6 结论

本文利用平均相隔 10 km, 分布在长春—四平地区 100 km×100 km 的 96 个加密自动站雨量计以及 51 个省局站雨量计在 2007~2011 年 4~10 月获得的降水测量数据, 通过数据连续无变化检查、自动气象站工作情况检查、极值检查、邻近站点比较检查以及基于雷达回波强度估测降水的检查, 对

147 个自动站雨量计进行了质量控制, 然后分析研究了雨量计测量之间的相关关系, 得到以下主要结论:

(1) 未通过数据连续无变化检查的数据虽然极少, 但可以确认为错误数据; 未通过自动气象站工作情况检查和极值检查的小时降水数据分别占 0.9%和 0.01%, 未通过邻近站点比较检查的站点有 1 个, 为 4208 号站。这三项检查未通过的数据都记为异常数据 (E1、E2 和 E3), 需要利用其他测量资料做进一步的检查控制。

(2) 利用雷达估测降水值对雨量计数据进行进一步质量控制, 通过相似离度和观测时次对比分析, 可以对标记异常的数据做进一步检查, 确认其中错误的的数据, 判断错误数据类型。位于雷达近距离 (10 km 以内) 的雨量计, 可能由于地物影响而

与雷达相似离度偏大, 可以利用雷达高仰角得到的估测值与雨量计进行对比。经过质量控制后, 有 6 个站点记为错误站点, 其数据不再在分析研究中使用; 其余标记异常的数据继续保留。经此步骤质量控制后的雨量计数据与雷达估测对比显示, 均值差距减小, 相关系数增大。

(3) 利用质量控制后的自动站雨量计小时降水数据进行雨量计相关关系分析显示, 随着站点之间距离的增加, 雨量计之间相关系数减小, 平均绝对差值和均方根值增加。雨量计距离在 10 km 以内, 平均相关系数均能达到 0.6 左右。雨量计距离小于 5 km, 平均相关系数在 0.7 以上。而站点距离超过 20 km, 相关系数普遍降到 0.4 以下。

(4) 通过对 6 min、30 min、1 h、6 h、12 h、24 h 以及 1 mon 等 7 个不同时间尺度的降水量进行相关关系、平均绝对差和均方根值的对比分析, 表明平均绝对差和均方根随着距离的增加而增大, 相关系数随距离的增加而减小。随着时间尺度的增加, 平均相对差和均方根值均减小, 而相关系数增加。在小于 10 km 的距离范围内, 雨量计日降水量和月降水量之间的平均相关系数可以达到 0.8 以上, 说明相距 10 km 以内的雨量计日降水量与月降水量之间具有很高的相关性, 单站降水测量具有一定的空间代表性。

最后要说明的是, 本工作研究仅有 5 年的加密雨量站网测量资料, 所以没有在季节和年的时间尺度上进行测量代表性分析。此外, 由于资料时间长度有限, 以及降水时空分布的极大不均匀性, 以现有加密站网的雨量计测量很难直接检验人工增雨的效果。为了获得更为可信和准确的降水资料, 建议在基准和重点台站布置双(多)雨量测量装置, 在更小区域内(如 1 km×1 km)部署更为密集雨量站网, 进行长期的观测对比试验。

致谢 感谢吉林省人工影响天气办公室刘岩和孙海燕, 他们提供了本研究所需的数据资料; 感谢两位审稿专家, 他们对论文修改给出了多条宝贵意见。

参考文献 (References)

Böhner J. 1994. Circulation and representativeness of precipitation and air temperature in the southeast of the Qinghai-Xizang Plateau [J]. *GeoJournal*, 34 (1): 55-66.
陈秋萍, 曾光平, 冯宏芳, 等. 2005. 对流云总降水量和降水效率估测[J]. *应用气象学报*, 16 (2): 260-263. Chen Qiuping, Zeng Guangping, Feng

Hongfang, et al. 2005. The total precipitation and the estimation of precipitation efficiency of convective [J]. *Journal of Applied Meteorological Science (in Chinese)*, 16 (2): 260-263.
Ciach G J. 2003. Local random errors in tipping-bucket rain gauge measurements [J]. *J. Atmos. Oceanic Technol.*, 20: 752-759.
Ciach G J, Krajewski W F. 2006. Analysis and modeling of spatial correlation structure in small-scale rainfall in central Oklahoma [J]. *Advances in Water Resources*, 29 (10): 1450-1463.
丛芳, 刘黎平. 2011. 新一代天气雷达与地面雨量资料的综合分析[J]. *气象*, 37 (5): 532-539. Cong Fang, Liu Liping. 2011. A comprehensive analysis of data from the cinrad and the ground rainfall station [J]. *Meteorological Monthly (in Chinese)*, 37 (5): 532-539.
东高红, 刘黎平. 2012. 雷达与雨量计联合估测降水的相关性分析[J]. *应用气象学报*, 23(1): 30-39. Dong Gaohong, Liu Liping. 2012. Correlation analysis on estimating rainfall using radar-rain gauge calibration [J]. *Journal of Applied Meteorological Science (in Chinese)*, 23 (1): 30-39.
Kuzuha Y, Tomosugi K, Kishii T. 2001. Representativeness of precipitation data [J]. *Journal of Japan Society of Hydrology & Water Resources*, 14 (6): 461-471.
李开乐. 1986. 相似密度及其使用技术 [J]. *气象学报*, 44 (2): 174-177. Li Kaile. 1986. A new similarity parameter and its application [J]. *Acta Meteorological Sinica (in Chinese)*, 44 (2): 174-177.
刘红燕, 雷恒池. 2006. 基于地面雨滴谱资料分析层状云和对流云降水的特征 [J]. *大气科学*, 30 (4): 693-702. Liu Hongyan, Lei Hengchi. 2006. Characteristics of rain from stratiform versus convective cloud based on the surface raindrop data [J]. *Chinese Journal of Atmospheric Sciences (in Chinese)*, 30 (4): 693-702.
Marani M. 2003. On the correlation structure of continuous and discrete point rainfall [J]. *Water Resour. Res.*, 39 (5): 1128, doi:10.1029/2002WR001456.
Patrons A A N, Chen N C J, Miller R L. 1979. Spatial correlations of monthly rainfall: Applications in climatology and weather modification experiments [J]. *Journal of Applied Meteorology*, 18 (6): 719-732.
任芝花, 熊安元. 2007. 地面自动站观测资料三级质量控制业务系统的研制 [J]. *气象*, 33 (1): 19-24. Ren Zhihua, Xiong Anyuan. 2007. Operational system development on three-step quality control of observations from AWS [J]. *Meteorological Monthly (in Chinese)*, 33 (1): 19-24.
任芝花, 赵平, 张强, 等. 2010. 适用于全国自动站小时降水资料的质量控制方法 [J]. *气象*, 36 (7): 123-132. Ren Zhihua, Zhao Ping, Zhang Qiang, et al. 2010. Quality control procedures for hourly precipitation data from automatic weather stations in China [J]. *Meteorological Monthly (in Chinese)*, 36 (7): 123-132.
Rodda J C. 1968. The rainfall measurement problem [M] // *Proceedings of the General Assembly of the International Association of Hydrological Sciences*. Exeter, UK: IAHS Publishing, 78: 215-231.
Šálek M, Štěpánek P, Zahradníček P. 2012. Analysis of rainfall intensities using very dense network measurements and radar information for the brno area during the period 2003-2009 [J]. *Meteorologische Zeitschrift*, 21 (1): 29-35.
Steiner M, Smith J A, Burges S J, et al. 1999. Effect of bias adjustment and

- rain gauge data quality control on radar rainfall estimation [J]. *Water Resour. Res.*, 35: 2487–2503.
- 孙海燕, 崔莲, 齐颖, 等. 2006. 自动雨量计使用情况及问题分析 [J]. *吉林气象*, (2): 26–27. Sun Haiyan, Cui Lian, Qi Ying, et al. 2006. Raingauge usage and problem analysis [J]. *Meteorology of Jilin (in Chinese)*, (2): 26–27.
- Villarini G, Mandapaka P V, Krajewski W F, et al. 2008. Rainfall and sampling uncertainties: A rain gauge perspective [J]. *J. Geophys. Res.*, 113, D11102, doi: 10.1029/2007JD009214.
- 袭祝香, 刘玉英, 张丽, 等. 1998. 近百年来吉林省的气候变化及未来趋势 [J]. *吉林气象*, (2): 30–33. Xi Zhuxiang, Liu Yuying, Zhang Li, et al. 1998. The climate change over the past century and future trends of Jilin Province [J]. *Meteorology of Jilin Province (in Chinese)*, (2): 30–33.
- 徐亚钦, 翟国庆, 黄旋旋, 等. 2011. 利用雷达和自动站资料综合分析风暴移动和发展规律 [J]. *大气科学*, 35 (1): 134–146. Xu Yaqin, Zhai Guoqing, Huang Xuanxuan, et al. 2011. A study of the evolution regularity of storm based on the data of radar and automatic station [J]. *Chinese Journal of Atmospheric Sciences (in Chinese)*, 35 (1): 134–146.
- Yang Daqing, Wang Ninglian, Ye Baisheng, et al. 2009. Recent advances in precipitation-bias correction and application [J]. *Sciences in Cold and Arid Regions*, 1 (3): 193–198.
- 杨萍, 刘伟东, 仲跻芹, 等. 2011. 北京地区自动气象站气温观测资料的质量评估 [J]. *应用气象学报*, 22 (6): 706–715. Yang Ping, Liu Weidong, Zhong Jiqin, et al. 2011. Evaluating the quality of temperature measured at automatic weather stations in Beijing [J]. *Journal of Applied Meteorological Science (in Chinese)*, 22 (6): 706–715.
- 张强, 涂满红, 马舒庆, 等. 2007. 自动雨量站降雨资料质量评估方法研究 [J]. *应用气象学报*, 18 (3): 365–372. Zhang Qiang, Tu Manhong, Ma Shuqing, et al. 2007. Quality assessment of the observational data of automatic precipitation stations in China [J]. *Journal of Applied Meteorological Science (in Chinese)*, 18 (3): 365–372.