顺序卡尔曼滤波校准法雷达定量估测降水检验

梁卫芳[1,2] 杨育强[1] 林 行[1] 吴增茂[2] 官 莉[3]

(1. 山东省青岛市气象局, 266003; 2. 中国海洋大学; 3. 南京信息工程大学)

提 要: 利用青岛市气象局多普勒天气雷达基数据和自动雨量计资料,对顺序卡尔曼滤波校准法进行定量估测降水的结果进行了大量样本和一个典型个例检验。5.0mm·h^{-1}以下的弱降水没有参与校准因子的计算,在弱降水、稳定性降水、对流性降水和混合性降水过程中的均方根误差分别为 1.15mm、1.13mm、3.0mm 和 2.18mm, 校准方案对克服弱回波情况下雷达资料的不确定性和自动雨量计数据的不确定性带来降水估测误差有良好的效果。青岛的降水过程中,5.0~20.0mm·h^{-1}降水的样本占有最高的比例,在校准因子计算时拥有较高的权重,该雨量档的误差也较小。稳定性天气过程中,降水回波性质差异不大,校准效果最好。滑动平均法是一个简单适用的雷达资料网格化方案。

关键词: 多普勒天气雷达 顺序卡尔曼滤波校准法 滑动平均法 检验

引 言

及时准确地获取地面某一段时间降水量的区域分布,具有非常重要的意义。例如防汛、水资源管理,以及城市内涝、山体滑坡等由于短时强降水引起的次生灾害的研究、数值预报的改善等。一条可以想到的途径是布设密集的自动天气观测站。但是,目前自动天气观测站在反映中小尺度降水的空间分布问题上仍有很大的困难,而且由于地形等因素的影响,这种雨量信息往往不能代表真实雨量的分布[1]。

新一代多普勒天气雷达反射率因子 Z 具有时空分辨率高,获取及时的优势,其与同一时刻的降水强度 I 有着密切的联系,使得利用多普勒天气雷达资料定量估测区域降水,就成为一条非常重要的途径。但是,多普勒天气雷达的 Z-I 关系具有很大的不确定性,使得提高多普勒天气雷达定量估测降水的精度成为难题。国内外的许多专家学者为提高估测精度而不懈努力,并发展了很多算法,统计方法如 Z-I 关系法(Brands,1985)、概率配对法(Marks,1991)等。随着实时地获取一定密度的地面雨量数据成为可能,多普勒天气雷达-地面自动雨量计联合定量估测降水就成为一条更好的满足实际业务需要的途径。已经有的一些学术研究,如平均校准法(Barnston,1981), 卡尔曼滤波标校法(Ahnert,1986)和变分校准法(Sasaki,1970)等。

1960 年,数学家卡尔曼(R. E Kalman)针对随机过程状态估计问题提出了卡尔曼滤波。之后,卡尔曼滤波在信号处理、最优控制、航天等领域得到广泛应用。Ahnert 等人(1984、1986)用数值模拟的方法首次验证了卡尔曼滤波法测量降水的有效性和稳定性,O'Bannon 等人[2]用 Oklahoma 冬季的实际降

水资料对美国 NEXRAD 降水子系统进行测试,了解卡尔曼滤波某些参数值设置对降水测量精度的影响。尹忠海[3]等人则使用2002年的一个个例,对卡尔曼滤波校准方法估算区域降水量的效果进行了试验,认为该方法能有效地提高雷达定量测量降水的精度,并能较好地体现雷达测量降水场结构精细的特点。

本文以青岛市气象局新一代多普勒天气雷达基数据和自动雨量计数据为研究对象,研究顺序卡尔曼滤波校准法多普勒天气雷达-地面自动雨量计联合定量估测降水在实际业务中的应用,并使用大量样本对应用效果进行评估。

1　顺序卡尔曼滤波校准法定量估测降水的原理

多普勒天气雷达-自动雨量计联合定量估测降水的原理,是引入一个新的量,即校准因子f。用这个校准因子对多普勒天气雷达反射率因子形成的初估场进行调整,得到降水的空间分布。顺序卡尔曼滤波校准法获取校准因子的方法是最优偏差估计。

当有多个雨量计站时,用各站取得的偏差平均值作为f,则f与空间无关,可以视为"随机游走"问题,可得到离散时间状态下的状态方程和量测方程[4]:

$$f(k+1)=f(k)+W(k) \tag{1}$$

$$Y(k)=f(k)+M(k) \tag{2}$$

由于$W(k)$和$M(k)$均是均值为零的白噪声,且相互独立,所以有:

$$\begin{cases} \mathrm{E}[W(k)]=0 & k\geqslant 0 \\ \mathrm{Cov}[W(k),W(j)]=Q(k)\delta_{kj} & k,j\geqslant 0 \\ \mathrm{E}[M(k)]=0 & k\geqslant 0 \\ \mathrm{Cov}[M(k),M(j)]=F(k)\delta_{kj} & k,j\geqslant 0 \\ \mathrm{Cov}[W(k),M(j)]=Q(k)\delta_{kj} & k,j\geqslant 0 \end{cases} \tag{3}$$

其中,$\delta_{kj}=\begin{cases}0 & k\neq j\\ 1 & k=j\end{cases}$

$Q(k)$是状态噪声方差,$F(k)$是测量噪声方差,参考文献[5]在求解过程中分别由下式给出:

$$Q(k)=a_2(1-a_1^{\,2}) \tag{4}$$

$$F(k)=a_3\eta(k)^{a_4} \tag{5}$$

η为参与校准的自动雨量计的数目。

对方程(1)~(3)求无偏最优解,得顺序卡尔曼滤波校准法的迭代方程组[4]:

$$\hat{f}(k|k)=\hat{f}(k|k-1)+K(k)[Y(k)-\hat{f}(k|k-1)] \tag{6}$$

$$\hat{f}(k|k-1)=\hat{f}(k-1|k-1) \tag{7}$$

$$P(k|k-1)=P(k-1)+Q(k-1) \tag{8}$$

$$P(k|k)=[1-K(k)]P(k|k-1) \tag{9}$$

$$K(k)=P(k|k-1)/[P(k|k-1)+F(k)] \tag{10}$$

其中:$\hat{f}(k|k)$是基于第k小时测量值$Y(k)$、$K(k)$及第$(k-1)$小时估计值$\hat{f}(k|k-1)$而求得的第k小时经滤波后的输出,即所需获得的校准后的偏差估计值,也就是最优的校准因子f;$\hat{f}(k|k-1)$是据第$(k-1)$小时经滤波输出的量$(k-1|k-1)$而作出第k小时的预测估计值;$P(k|k-1)$表示对$\hat{f}(k|k-1)$进行预测时所存在的误差方差;$P(k|k)$为第k小时输出值$f(k)$经滤波后输出值$\hat{f}(k|k)$与状态值$f(k|k)$之间的误差方差,即$P(k)$;$K(k)$表示第k小时的滤波增益;Q为状态方程的误差方差;F为量测方程的误差方差;$Y(k)$是没有经过滤波的校准因子,是滤波器的输入值,获取方法见2.2.2节。

2　顺序卡尔曼滤波校准法的实现

2.1　降水的初估场

2.1.1　反射率因子的选取

在获取多普勒天气雷达的反射率因子之前,首先对其进行质量控制,采用孤立点消除法和倾斜测试法消除杂波[6],对冰水混合物对降水回波的污染,采用55dBz上限阈值限制法。

通常认为,获取降水估测的反射率因子的最佳高度是雷达视屏面以上 1km 高度。本文采用"混合扫描"的方式:在 0 ~20km 使用 3.30°仰角,20 ~40km 使用 2.40°仰角,40 ~60km 使用 1.50° 仰角,60 ~ 150km 使用 0.50°仰角。

2.1.2 雷达资料的网格化

雷达资料网格化方法有最近邻居法、垂直水平线性内插法、三维 Barnes 插值法和变分分析法等[7-8]。本文没有对以上方案进行降维处理,而是重新设计了一套方案。在网格化数据时,认为降水率分布在 X - Y 平面上连续分布,采用参考文献[5]提供的 Z - I 关系式,即 $Z=200I^{1.6}$,首先将获取的二维的反射率因子转换为降水率。然后,设计 400 ×400 个网格距为 1km 的网格,求取每个网格的中心点在雷达站为中心的极坐标中的位置。以该位置为中心,搜索半径 2km 范围内的距离库,对有效数据的降水率进行平均,赋值为网格的降水率数据。以上网格化方案还考虑到在 150km 处,相邻方位的距离约为 2km,2km 半径的空间也可以较好的分辨小尺度降水单体。即使在近距离处,为了补偿风的作用等原因,造成的多普勒天气雷达反射率因子场与自动站位置的偏移,将网格数据视为 2km 半径的空间内降水率的平均,也是合理的,不存在降低数据多普勒天气雷达数据的空间分辨率的问题。由于该方案有滑动平均的意义,因此称之为滑动平均法。该算法对降水精度的影响,将另文阐述。

对网格化以后的格点资料进行时间累计,即形成多普勒天气雷达 - 自动雨量计联合定量估测降水的初估场。

2.2 顺序卡尔曼滤波校准法的实现流程

2.2.1 自动雨量计数据预处理

当一个小时的数据采集结束后,按上节所述方法可以得到降水量的初估场,接着就可以调入相应时段的自动站数据,进行校准。超过 1000.0mm·h^{-1} 的自动雨量计数据看作无效数据[9],小于 5.0mm 的自动雨量计数据存在不确定性[10],这些数据不参与校准因子计算。

2.2.2 校准流程

参考文献[5],校准迭代方程组(6) ~ (11)的初始值赋值为:

$$f_0=1, p_0=0;\ a_1=0.8; a_2=0.1;$$
$$a_3=1.0; a_4=-1.0 \qquad (11)$$

没有经过滤波的偏差校准因子设为:

$$Y(t)=\frac{\sum_{i=1}^{N} I_{gi}(t)}{\sum_{i=1}^{N} I_{ri}(t)} \qquad (12)$$

其中,N 为参加校准的自动雨量计总数。$I_{gi}(t)$为本次参与标校的自动雨量计雨量,$I_{ri}(t)$为与自动雨量计同一位置的初估场的值。这个值是空间的平均值,满足校准迭代方程组的前提条件。将该值和初始值代入迭代方程组(6) ~ (11),得到最优的校准因子 f 和校准误差。校准因子 f 与初估场的乘积即为降水的空间分布。

3 估测结果的检验

检验以青岛市气象局的小时自动雨量计数据为真值。超过 1000.0mm·h^{-1} 的自动雨量计数据看作无效数据,不参与检验。小于 5.0mm·h^{-1} 的自动雨量计数据单独进行检验。

3.1 大量样本的检验

大量样本的检验采用分降水类型,并分雨量档的方法进行。参考强对流天气研究中的短时强降水概念[11],将大于 20.0mm·h^{-1} 的自动雨量计数据视为对流性降水数据,将小于 5.0mm·h^{-1} 的自动雨量计数据视为弱降水数据。然后将检验样本划分为弱降水、稳定性降水、对流性降水、混合性降水四种类型。其中过程中只有弱降水数据的个例,视为弱降水样本。降水持续时间小于 6 小时,

小时雨量数据的对流性降水数据超过50%的个例,视为对流性降水样本。小时雨量5.0~20.0mm·h^{-1}降水数据超过50%的个例,视为稳定性降水样本。其余样本视为混合性降水样本。对2006年以来青岛市的降水过程进行分析,得到34个降水过程。其分型后的样本分布为:弱降水样本4个(37个时次),对流性样本12个(69个时次),稳定性样本6个(113个时次),混合型样本12个(183个时次)。

每个检验样本的小时雨量被分为4档(1):0.1~5.0mm,(2):5.0~20.0mm,(3):20.0~50.0mm,(4):>50.0mm。对每一种降水类型的样本统计了均方根误差、相对误差。

从理论上讲,每一种观测仪器的观测值,都是对真值的一种反映,而非真值本身。而在实际业务应用中,估测结果在一定范围内,对预报结果影响不大,可以认为这个范围内的估测结果是正确的,据此定义了正确率、错误率(包括高估率和低估率)概念,表1为正确估计、高估和低估的定义。

其中:

$$\Delta R = R_g - R_k \tag{13}$$

R_g为自动雨量计数据,R_k为估测的降水。

正确率定义为正确估计样本数占总样本数的比例,高估率定义为高估样本数占总样本数的比例,低估率定义为低估样本数占总样本数的比例。

表1　4个雨量档的正确估计、错误估计(包括高估和低估)的定义(ΔR单位:mm)

雨量档	1	2	3	4
正确范围	落在该档	$-5.0<\Delta R<5.0$	$-5.0<\Delta R<5.0$	$-10.0<\Delta R<10.0$
高估	$\Delta R>0.0$	$\Delta R>5.0$	$\Delta R>5.0.0$	$\Delta R>10.0$
低估		$\Delta R<-5.0$	$\Delta R<-5.0.0$	$\Delta R<-10.0$

表2　34个样本的检验结果

样本类型	降水量档	均方根误差/mm	相对误差	正确率	高估率	低估率
弱降水	1	1.15	155%	63%	-	-
稳定性降水	1	1.13	141%	57%	-	-
	2	3.384	43%	78%	8%	14%
对流性降水	1	3	318%	41%	-	-
	2	7.09	79%	36%	15%	49%
	3	12.73	54%	20%	20%	60%
	4	-	-	-	-	-
混合性降水	1	2.18	248%	55%	-	-
	2	5.16	58%	58%	17%	25%
	3	10.87	38%	36%	12%	53%
	4	33.69	49%	8%	6%	87%

表2为检验结果,"-"表示无数据。弱降水本身没有参与校准因子的计算,但是无论是哪种类型的过程,校准结果都有良好的可参考性。使用当地5.0mm·h^{-1}以上的自动雨量计数据计算校准因子,对克服弱回波情况下雷达资料的不确定性[12]带来估测误差也是一条重要的途径。表2显示,稳定性降水的校准效果最好,最差的是对流性降水。校准效果最好的是5.0~20.0mm档的雨量,最差的是50.0mm·h^{-1}以上的雨量。

3.2　典型个例的检验

受0713号超强台风韦帕和西风槽、副热带高压共同影响,2007年9月18日至20日,青岛市出现全区性大暴雨过程,并创下自有气象记录以来,青岛市大暴雨覆盖面积的新

纪录。天气实况、环流形势分析、物理量诊断和雷达资料分析表明，过程的中小尺度系统活跃，对流层的低层扰动强烈[13]。四个雨量档样本数在这个过程中的分布如图1，该过程没有记录到50.0mm·h^{-1}以上的降水数据。

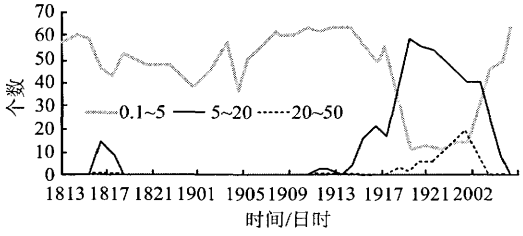


图1 2007年9月18日13时至20日02时(UTC)自动雨量计样本数

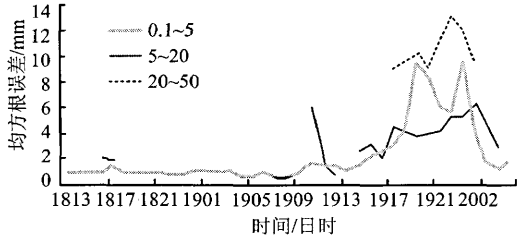


图2 2007年9月18日13时至20日02时(UTC)降水估测的均方根误差

18日15—17时，有26个5.0~10.0mm·h^{-1}的自动雨量计数据参与了校准。18日17时至19日11时没有自动雨量计数据参与，但是弱降水的误差很小，均方根误差约0.9mm(图2)。弱降水过程中使用的校准因子，来自于上一个降水过程，通常都是比较接近5.0mm·h^{-1}的数据，与大量样本检验弱降水校准误差较小一致。19日11时以后，雨势增强，弱降水的误差迅速增大。5.0~20mm·h^{-1}的样本比例最高，其误差也最小，其正确率平均在80%以上。20.0~50mm·h^{-1}数据的样本数少，误差大，多数情况下雨量被低估了。

4 总结

(1)5.0mm·h^{-1}以下的自动雨量计数据没有参与校准因子的计算，但是这个雨量档

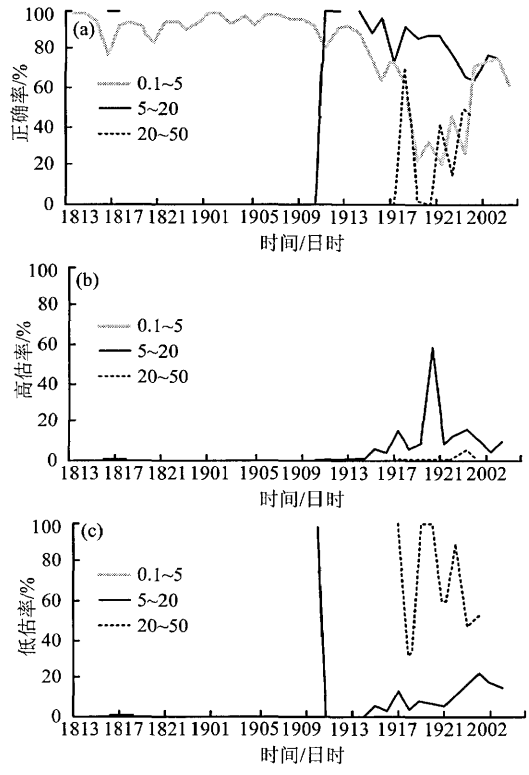


图3 2007年9月18日13时至20日02时(UTC)降水估测正确率(a)、高估率(b)和低估率(c)

的雨量估计误差比较小。本文的校准方案对克服弱回波情况下雷达资料的不确定性和自动雨量计数据的不确定性带来降水估测误差有较好的效果。

(2)在稳定性天气过程中，降水回波性质差异不大，校准效果最好。在混合性天气过程中，5.0~20.0mm·h^{-1}的样本比例最高，其误差也最小。而20.0~50.0mm·h^{-1}数据的样本数比较少，在校准因子计算时，其权重也小，因此误差较大，多数情况下这个雨量档的降水被低估了。校准效果最差的是50.0mm·h^{-1}以上的雨量误差较大，也是这个原因。而对流性天气过程中回波性质复杂权重分散，各个雨量档降水估测的误差都有所增大。因此，在提高顺序kalman滤波校准法天气雷达-自动雨量计联合定量估测降水算法精度的研究中，应重点研究怎样进行降水性质的识别，然后针对不同性质的降水进行

不同的校准。

(3)研究适合天气雷达定量估测降水的网格化方法是必要的。本文设计的滑动平均法是一种简单适用的方法。

参考文献

[1] Beven. k, Wood. E. F, Sivapalan M. On hydrological heterogeneity: catchment morphology and catchment response[J]. Journal of Hydrology, 1988, 100: 129-138.

[2] O'Bannon. T, Ahnert. P. A Study of the Precipitation Processing System on a Winter2type Oklahoma Rainstorm[G]. Preprints of 23rd Radar Meteorology Conf, Boston: AMS, 1986: 99-102.

[3] 尹忠海,张沛源.利用卡尔曼滤波校准方法估测区域降水量[J].应用气象学报,2005,16(2):213-219.

[4] 张培昌,杜秉玉,戴铁丕.雷达气象学[M].北京:气象出版社,2001:508.

[5] Smith. J. A, Krajewski. W. F. Estimation of the mean field bias of radar-rainfall estimates[J]. J A ppl Meteor, 1991, 30: 397-412.

[6] 俞小鼎,姚秀萍,熊廷南,等.多普勒天气雷达原理与业务应用[M].北京:气象出版社,2006:314.

[7] 肖艳姣,刘黎平.新一代天气雷达网资料的三维格点化及拼图方法研究[J].气象学报,2006,64(5):647-657.

[8] 杨毅,邱崇践.多普勒雷达资料格点化方案的比较研究[J].干旱气象,2004,22(2):6-9.

[9] 杨茂水,李树贵,王锡芳.II型自动气象站简介[J].山东气象,2002,22:43-44.

[10] 张霭琛.现代气象观测[M].北京:北京大学出版社.2000:423.

[11] 杨露华,尹红萍,王慧,等.近10年上海地区强对流天气特征统计分析[J].大气科学研究与应用,2007,2:84-91.

[12] 刁秀广,朱君鉴,杨传凤,等.济南CINRAD/SA雷达杂波抑制方案实施[J].山东气象.2005,25:25-29.

[13] 梁卫芳,耿敏,李庆宝.台风"韦帕"全区性大暴雨过程分析[J].气象,待发表.