

王芬,李腹广,张辉. 风暴单体识别与跟踪(SCIT)算法评估[J]. 气象,2010,36(12):128-133.

风暴单体识别与跟踪(SCIT)算法评估^{* 1}

王 芬 李腹广 张 辉

贵州省黔西南州气象局, 兴义 562400

提 要: 利用兴义新一代多普勒天气雷达复合体扫资料及 WSR-88D 提供的风暴单体识别与跟踪(SCIT)算法对 2007—2008 年发生在贵州省黔西南地区的 40 次天气过程个例进行验证、分析,分别从单体识别、单体跟踪评估、单体位置预报、单体算法评估效果与季节关系分析、单体算法评估效果与距离关系分析、冰雹暴雨评估对比分析 6 个方面进行评估、分析。在充分考虑了当地地理环境、气候背景的前提下对算法评估不太理想的情况进行了误差分析,并进行了算法补偿,提出了解决的办法,一是降低识别阈值,将 SCIT 的 7 个反射率阈值降低一个等级,二是提取反射率因子垂直梯度,用上述两个改进方法对 2007—2008 年的 40 次天气过程重新进行评估,并与算法改进前的评估效果进行对比检验,结果表明,算法改进后评估效果有所提高。

关键词: 多普勒天气雷达, SCIT, 算法评估

Storm Cell Identification and Tracking Algorithm for Assessment of SCIT

WANG Fen LI Fuguang ZHANG Hui

The Meteorological Office of Southwestern Guizhou, Xingyi 562400

Abstract: Using the Xingyi new generation of Doppler weather radar complex scan data and WSR-88D storm cell identification and tracking (SCIT) algorithm, the 40 weather courses occurring in Southwest Guizhou Province in 2007 to 2008 were assessed and analyzed from the six aspects: cell identification, assessment of cell track, cell location prediction, the relationship with the assessing effect, the relation with the assessing effect and the distance, and the comparative analysis between the hail weather and the heavy rainfall. Under the background of taking full account of local geography and climate, the unfavorable situation for algorithm assessment is made analysis and compensation and proposed solutions. First, to lower identification threshold, so that the 7 SCIT reflectivity thresholds are degraded by one level; Second, the vertical reflectivity gradient method is used. With the improved algorithm the weather cases occurring in 2007 and 2008 were re-assessed and re-analyzed, and the comparative results show that the improved algorithm enhances the effect to some extent.

Key words: Doppler weather radar, SCIT (storm cell identification and tracking), algorithm evaluation

引 言

20 世纪 50 年代以来,国内外雷达气象学者在雷达探测强对流天气方面做了大量工作,利用天气雷达对风暴进行自动识别、跟踪和预报的研究已进行了半个世纪,发展了很多理论和方法,提出了一些

风暴识别方法^[1-7]。20 世纪 70 年代 NSSL 的科学家们开发了一系列算法,可以从雷达反射率因子数据中识别出风暴单体,并追踪和外推单体质心的运动^[1]。1982 年 Austin 等提出了三维矩心识别方法,即把每一个风暴单体看作成一个具有三维连续结构的整体,计算相关物理量^[8]。Rosenfeld 对这一方法进行改进,提出通过对风暴形状的几何推理来

* 贵州省气象局青年科技基金项目“风暴单体识别与跟踪 SCIT 算法评估”黔气科合 QN[2009]10 号资助
2009 年 7 月 20 日收稿; 2009 年 11 月 10 日收修定稿
第一作者:王芬,主要从事雷达资料产品二次开发研究。Email:wfen1080@163.com

处理分裂和合并^[9],这就是 1996 年以前在美国天气雷达业务(WSR-88D Build 7.0)中广为应用的风暴系列算法。1997 年 Johnson 提出了利用 7 个反射率因子识别阈值来代替此前唯一的阈值,同时采取特征核抽取技术,对空间相距较近的多个风暴单体进行合并或删除处理,且允许远距离上的二维风暴单体存在^[10]。该算法被 WSR-88D 的 Build 9.0 广泛采用,称为 B9SI 算法^[11]。目前,我国的新一代天气雷达大都使用 WSR-88D 的 Build 9.0 的风暴算法,即风暴单体识别与跟踪算法,简称 SCIT(Storm Cell Identification and Tracking)算法,但是由于 SCIT 算法设计是针对 S 波段的,而我国经济不发达的中西部的多普勒天气雷达是 C 波段,那么 SCIT 算法是否适应 C 波段雷达,应用效果如何,又如何根据本地环境气候进行本地化,这都是亟待解决的问题。

1 方法分析与资料选取

SCIT 算法由四个子功能组成:风暴单体段、风暴单体质心、风暴单体跟踪和风暴位置预报。SCIT 算法采用了以下几个技术:多反射率因子阈值识别技术、特征核抽取技术、对垂直重叠单体处理、远距离 2D 风暴单体的识别和相近多单体外理。风暴单体段识别反射率因子的径向排列,并输出这些段上的信息到风暴单体质心子功能中。风暴单体质心子功能将段组合成二维分量,并使这些分量垂直相关构成三维单体,再计算单体的属性。单体及它们的属性被输出到风暴单体跟踪及风暴位置预报子功能中。风暴单体跟踪子功能是通过将当前体积扫描发现的单体与前次体积扫描的单体作匹配来监视单体的移动。风暴位置预报子功能是依据风暴移动的历史来预报风暴将来的质心位置^[11]。用 SCIT 算法评估单体可以得到单体以下的几个属性:单体的所属风暴类型识别、单体的移动趋势、单体未来位置预报。

选取 2007—2008 年的 40 次天气过程作为评估对象,利用兴义多普勒天气雷达体扫数据资料及用 WSR-88D 提供的风暴单体识别与跟踪 SCIT 算法对个例天气进行评估。兴义新一代多普勒天气雷达是成都 784 厂生产的 C 波段雷达,雷达中心经纬度为 25.088°N、104.895°E,天线海拔高度为 1329.9 m。观测过程中,雷达参数设置为脉冲重复频率为 900/600

Hz,脉冲宽度为 1,脉冲采样数为 32,距离平均数设为无,多普勒处理设为 FFT,库长设为 250 km,窗口类型为 hamming,观测模式为 VCP21,采用滤波器 2 进行滤波。强度、速度、谱宽 3 个基数据的噪声门限全部设为 SQI & LOG(0.35/5.0 dB)。

本文研究的区域是贵州省黔西南地区(以兴义多普勒天气雷达为圆心,半径 150 km),由于雷达存在静锥区,距离太近探测不到风暴上层,会使探测效果降低,而受地球曲率和雷达自身探测能力的影响,超过 150 km 也会使探测效果降低,因此本文主要研究距雷达站中心 30~150 km 范围内的对流单体。

本文研究所选资料为发生在 2007—2008 年的 40 次冰雹及暴雨天气过程,体扫资料时间间隔为 6 分钟。由于一次天气过程中可能有多个不同类型的单体存在,故选取这 40 次天气过程中每次出现的最具代表性的产生冰雹或暴雨的一个单体分类统计出来作为研究的对象,共选取有单体 40 个。地面降雹时间、最大冰雹直径及降雨量以个乡镇雨量站记录为准。

2 算法评估

选取 2007—2008 年发生的 40 次天气过程作统计分析,在不区分冰雹和暴雨的情况下,用 SCIT 算法对这 40 次过程中出现的最具代表性的 40 个单体分别作识别、跟踪、位置预报分析。

2.1 单体识别评估

将风暴分为如下类型^[11]:

(1)“超级单体风暴”:风暴水平尺度大,可达几十千米,维持时间长,常伴有强风局地冰雹、冰雹、下击暴流等强风暴;

(2)“MCS/Line”:MCS 是中尺度对流系统,由对流单体、多单体风暴和超级单体风暴以各种形式组织而成,包括飚线和中尺度对流辐合体,水平尺度在 20~200 km,包括台风、小低涡、飚线、暴雨团等。MCS/Line 指的是其中单体的风暴构成一个团簇或呈线状排列,它是由一个演变着的单体系列组成的,即它包含着几个处在不同发展阶段的单体,灾害性天气(如冰雹、暴雨、大风等)发生在其中发展成熟的单体下方;

(3)“层状降水”:大块区域的小到中等的反射

率因子,反射率因子一般不超过 40 dBz,强度梯度不大。

表 1 给出 40 次天气过程 SCIT 算法的正确识别率,从表 1 可以看出,小于 40 dBz 的单体的识别率只有 28%,40 dBz 以上的单体有 66%的识别率,大于 50 dBz 的单体有 98%的识别率,整体来讲超级单体风暴的识别率较高,间隔较小的风暴群(线状、簇状)效果不如离散单体识别率高,从反射率因子的角度来说,平均反射率因子越大,相应风暴的识别率越高。

表 1 40 次天气过程 SCIT 算法的正确识别率(POD)

Table 1 The correct recognition rates of the SCIT algorithm (POD) of 40 weather processes

风暴类型	个例数量/个	POD(单体 30~39 dBz)	POD(单体 40~49 dBz)	POD(单体 大于 50 dBz)
超级单体风暴	11	47%	80%	98%
MCS/Line	17	24%	66%	98%
层状降水	12	12%	53%	—
总体	40	28%	66%	98%

2.2 单体跟踪的评估

正确的进行时间关联是跟踪风暴演化的基础,对在两个相继体扫中识别的风暴单体进行时间相关处理以确定第一个被识别的风暴单体的路径。首先利用前一个体扫单体质心的位置确定目前体扫中该单体的初猜位置,然后计算在当前体扫中识别的每个单体与计算的当前每个单体初猜之间的距离,最后计算所有至少已经在相继两个体扫中被识别的目前单体的新的运动向量。单体跟踪算法的评估涉及到以下三个步骤:(1)一个体扫接一个体扫地跟踪所有被识别单体的生命周期;(2)确定不正确的时间关联的个数并记下它们的特性;(3)计算该算法正确跟踪的单体占有被跟踪单体的百分比^[11]。其中时间关联的数量是单体从存在到消亡的全部体扫个数。选取发生在 2007 年至 2008 年发生的 40 次天气过程中出现的最具代表性的 40 个单体作统计分析个例来进行评估,具体评估结果见表 2。

表 2 对于第一个 SCIT 算法正确时间关联的百分比

Table 2 The percentage to the first time association of the SCIT algorithm

个例类型	单体数量	时间关联的数量	正确时间关联的百分比/%
超级单体	11	521	92
MCS/Line	17	1560	89
层状降水	12	1457	81
总体	40	3538	87

从表 2 可以看出,对于第一个 SCIT 算法时间关联的百分比,超级单体的最高为 92%,MCS/Line 次之,为 89%,层状降水最低为 81%。整体来讲超级单体风暴的识别率较高,间隔较小的风暴群(线状、簇状)效果不如离散单体识别率高。

2.3 单体位置预报的评估

选取发生在 2007—2008 年的 40 次天气过程中出现的最具代表性的 40 个单体(每个单体的生命史至少为 2 个体扫),针对各种类型的风暴,对 SCIT 算法的单体位置预报精度进行分析,具体评结果见表 3。

表 3 SCIT 算法对不同预报时效的平均预报误差

Table 3 The average prediction error to different lead time of the SCIT algorithm

预报时效/min	样本中单体的数量/个	平均预报误差/km
5	3538	2.4
15	1816	6.7
30	925	10.4
45	409	17.5
60	210	26.3

从表 3 可以看出,5 分钟的平均预报误差最小仅为 2.4 km,而 1 小时的平均预报误差最大,达到 26.3 km。预报时效与平均预报误差基本成正比,预报时效愈长,则平均预报误差愈大。

2.4 SCIT 算法评估与季节关系分析

用 WSR-88D 提供的风暴单体识别与跟踪(SCIT)算法对发生在贵州省黔西南地区的 2007—2008 年的 40 个天气过程中出现的最具代表性的 40 个单体个例按月分成 2—9 月,并利用 SCIT 算法来分别进行评估、分析,结果见表 4。

表 4 SCIT 算法评估与季节关系分析

Table 4 The analysis of the relationship between the seasons and the assessments of the SCIT algorithm

	2月	3月	4月	5月	6月	7月	8月	9月
两年月天气过程单体的数量	1	2	3	13	8	7	4	2
30 dBz 以上单体识别率/%	70	77	89	68	63	56	55	55
单体跟踪评估(第一个正确时间关联的百分比%)	97	96	96	87	84	80	82	75
15 分钟单体平均预报误差/km	3.6	3.2	3.6	4.9	6.9	9.2	9.3	12.9

从表 4 可以看出,2—5 月 4 个月的 30 dBz 以上单体识别率、单体跟踪评估、15 分钟单体位置评估效果都比较好,5 月份以后,SCIT 算法评估效果逐

渐下降,9月效果最差。

2.5 SCIT 算法评估与距雷达距离关系分析

因为雷达波束高度随距离增加而增加,雷达有效照射体积随距离增加可能出现波束未充塞,因此,雷达探测效果随距离增加一般情况下是递减的,那么 SCIT 算法效果也势必受到距离影响。所以,根据黔西南地区气候特征及地理环境,把该区域分为北部(晴隆、普安、兴仁、贞丰)、南部(兴义、望谟、安龙、册亨),从距离上考虑,以雷达中心为零点可划分 3 个距离段,分别为 0~20 km、20~60 km、60~120 km^[12],具体见图 1。

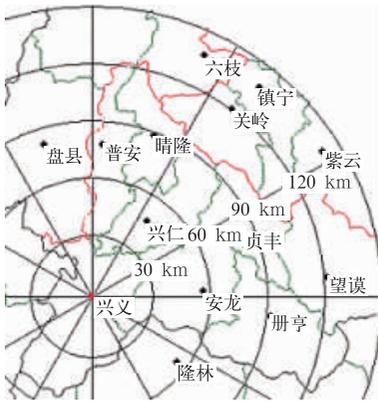


图 1 按气候特点及距雷达中心距离分类

Fig.1 The classification according to climatic features and the distances from the radar center

用 WSR-88D 提供的风暴单体识别与跟踪(SCIT)算法,选取发生在贵州省黔西南地区的 2007—2008 年的 40 个天气过程中出现的 40 个单体个例,把黔西南下属七县一市分别从 30 dBz 以上单体识别率、单体跟踪评估(第一个正确时间关联的平均百分比)、15 分钟单体平均预报误差三个方面进行评估,具体评估效果见表 5。

表 5 SCIT 算法评估与距雷达距离关系分析

Table 5 The analysis of the relationship between the distances and the assessments of the SCIT algorithm

地名	望谟	晴隆	册亨	贞丰	普安	安龙	兴仁	兴义
距雷达中心距离/km	120	90	90	85	82	55	50	0
30 dBz 以上单体平均识别率/%	45	61	62	63	69	72	72	68
第一个正确时间关联的平均百分比/%	80	83	84	86	87	92	94	92
15 分钟单体平均预报误差/km	11.0	7.4	7.2	6.4	5.8	5.5	5.1	5.2

从表 5 可以看出,SCIT 算法应用效果与距雷达

中心距离有关,其中距离雷达最远的望谟应用效果最差,而后随着距雷达距离愈近,应用效果也愈好,但并不是距雷达距离愈近,算法应用效果就愈好,由于雷达扫描存在一个静锥区,即在兴义雷达站上空存在一个雷达探测盲区,导致探测效果降低,而 SCIT 应用效果也势必受到影响。

2.6 冰雹、暴雨评估对比分析

一般冰雹、暴雨的雷达回波特征不同,强度场、速度场、空间结构、天气气候背景也不尽相同,故 SCIT 算法评估这两种不同的天气过程时,效果也不一定相同。对 2007—2008 年 40 个天气过程中出现的 40 个单体,分为冰雹天气过程、暴雨天气过程,用 SCIT 算法分别进行评估,评估结果见表 6。

表 6 冰雹暴雨评估对比分析

Table 6 The comparative analysis between hail and storm assessments

类型	个例数量	POD(单体 30~39 dBz)	POD(单体 40~49 dBz)	POD(单体 ≥50 dBz)	第一个正确时间关联的百分比/%	15 分钟单体平均预报误差/km
冰雹	20	—	70	98	98	4.4
暴雨	20	26	62	—	76.5	9.0

从表 6 可以看出,无论是单体识别,跟踪还是位置预报,SCIT 算法对冰雹的应用效果都更好些,这主要是因为冰雹天气过程多为单体风暴或多单体风暴,属于离散单体,而暴雨天气过程移动不规则,多个风暴相距很近,反射率因水平梯度不明显,因此误差稍大。

3 改进

SCIT 算法只考察反射率因子的强度和空间连续性,没有计算纹理结构和梯度,因此,对于识别出的风暴不能给出其内部对流情况,导致反射率因子强度较弱,但对流活跃的单体被遗漏,处理相近多单体时没有优势^[13]。本文在充分考虑了地理环境、气候背景的前提下,针对识别效果不太理想的情况,提出解决的办法,具体如下:

(1) 降低识别阈值,将 SCIT 的 7 个反射率阈值降低一个等级(55、50、45、40、35、30、25 dBz)。

(2) 提取反射率因子垂直梯度因子。风暴处于成熟阶段时,其典型特征之一就是核区内反射率因

子垂直递减率接近于零,因此反射率因子垂直梯度接近零时,表示此时的风暴的对流性最强,相反,则表示风暴的对流性很弱。

在 SCIT 算法的基础上,增加上述提出的两个

改进因子对 2007 年至 2008 年发生的 40 次天气过程重新进行评估,评估结果与改进前的效果进行对比分析,结果见表 7。

表 7 算法改进前后的 SCIT 算法评估对比分析

Table 7 The comparative analysis between the improved algorithm and the SCIT algorithm

30 dBz 以上单体识别率		单体跟踪评估(第一个正确时间关联的百分比)		15 分钟单体位置评估(平均预报误差)	
个例数量/个	28	时间关联的数量/个	3538	样本中单体的数量/个	1816
SCIT 算法评估结果/%	64	SCIT 算法评估结果/%	87.25	SCIT 算法评估结果/km	6.7
改进后的评估结果/%	72	改进后的评估结果/%	92	改进后的评估结果/km	6.3

从表 7 可以看出,改进后的评估效果有所提高,其中,算法改进后 30 dBz 以上单体识别率评估效果较之其他两项稍好,提高了 12.5%,而单体跟踪评

估及 15 分钟单体位置评估为 5.4% 及 6.0%。

用改进后的算法对这 40 个天气个例按照月份重新评估,结果见表 8。

表 8 算法改进前后 SCIT 算法评估与季节关系对比分析

Table 8 The comparative analysis between the improved algorithm and the SCIT algorithm according to the seasons

		2 月	3 月	4 月	5 月	6 月	7 月	8 月	9 月
两年月个例总量/个		1	2	3	13	8	7	4	2
30 dBz 以上单体识别率/%	改进前	70	77	89	68	63	56	55	55
	改进后	73	81	93	72	70	64	64	59
单体跟踪评估(第一个正确时间关联的百分比)/%	改进前	97	96	96	87	84	80	82	75
	改进后	98	97	96	95	92	89	89	80
15 分钟单体平均预报误差/km	改进前	3.6	3.2	3.6	4.9	6.9	9.2	9.3	12.9
	改进后	3.4	2.9	3.3	4.5	6.5	8.5	8.8	12.5

从表 8 可以看出,改进后评估效果整体有所提高,其中,30 dBz 以上单体识别率评估效果中 6、7、8 月的识别率比其他月份改进效果好,这主要因为这 3 个月的回波多为层状回波或间隔较小的风暴群,单体的识别率提高的空间较大。5—8 月的第一个正确时间关联的百分比改进效果比其他月份好。7 月、8 月的 15 分钟单体平均预报误差改进效果比其他月份稍好,大于 0.5 km。

用改进后的算法对这 40 个天气个例按照天气类型(冰雹、暴雨)重新评估,结果见表 9。

从表 9 可以看出,算法改进后冰雹和暴雨的评估效果整体都有所提高,其中冰雹的单体识别率提高了 6.5%,而暴雨提高 17%,冰雹的单体跟踪评估提高了 1%,而暴雨提高了 9%,冰雹的 15 分钟单体平均预报误差减小了 9.1%,而暴雨只减小了 4.4%。整体而言,算法改进后暴雨单体识别率及单体跟踪评估提高较快,而 15 分钟单体平均预报误差提高不明显,冰雹的 15 分钟单体平均预报误差比暴雨的稍好^[14-15]。

表 9 算法改进前后冰雹暴雨评估对比分析

Table 9 The comparative analysis on hail and storm assessments between the improved algorithm and the SCIT algorithm

类型		冰雹	暴雨
单体识别率	改进前	77	51
	改进后	83.5	68
	改进前后评估提高率(改进后的识别率减去改进前的识别率)	6.5	17
第一个正确时间关联的百分比	改进前	98	76
	改进后	99	85
	改进前后评估提高率(改进后的百分比减去改进前的百分比)	1	9
15 分钟单体平均预报误差百分率/%	改进前	4.4	9.0
	改进后	4.0	8.6
	改进前后评估提高率(改进后得到的平均预报误差千米数减去改进前得到的相关数据再除以改进前的平均预报误差千米数)	-9.1	-4.4

4 小 结

适当降低反射率因子阈值,再把反射率因子垂

直梯度因子考虑进去,探测效果会有一些改善,但也不是阈值越低准确性越高,因为随着阈值的降低也会导致虚报率的提高。究竟反射率阈值为多少才是最合适的,是因地而异的。另外,如何更好地应用雷达资料来开发出本地识别参数也是一个值得研究的方向,本文提到的反射率因子垂直梯度只是其中一个方面,还有更多的要素需要深入探讨,比如:谱宽资料、径向速度、反射率因子水平梯度信息、垂直累积液态水含量等等。

另外,本文评估实验地区为贵州省黔西南州地区,评估对象为 714 型 C 波段雷达,具有独特性,由于各个地区的雷达型号、气候背景、地理环境不同,即使用同样的 SCIT 算法来评估,评估效果也不一定相同,甚至可能有较大出入。

风暴单体识别与跟踪(SCIT)算法是美国 WSR-88D 的经验公式,目前我国尚没有自己的算法。今后,随着探测个例的增加及探测手段的不断提高,要进一步研究、开发适合于本地的雷达识别参数,以便提高本地的雷达探测效果。

参考文献

- [1] 韩雷,王洪庆,谭晓光,等. 基于雷达数据的风暴体识别、追踪及预警的研究进展[J]. 气象,2007,33(1):3-10.
- [2] 张沛源,杨洪平,胡绍萍. 新一代天气雷达在临近预报和灾害性天气警报中的应用[J]. 气象,2008,34(1):3-11.
- [3] 胡明宝,高太长,汤达章. 多普勒天气雷达资料分析与应用[M]. 北京:解放军出版社,2000.
- [4] 张培昌,杜秉玉,戴铁丕. 雷达气象学[M] 北京:气象出版社,2001.
- [5] 王丽荣,胡志群,匡顺四. 应用雷达产品计算风暴相对螺旋度[J]. 气象,2006,32(4):45-51.
- [6] 安英玉,任丽,赵广娜,等. 雷达产品算法及其在预报中的应用[J]. 黑龙江气象,2007,(03):42-44.
- [7] 冯晋勤,汤达章,俞小鼎,等. 新一代天气雷达中气旋识别产品的统计分析[J]. 气象,2010,36(8):47-52.
- [8] Austin G L, Bellon. Very short-range forecast of precipitation by the objective extrapolation of radar and satellite data[M]. In: Broning K, Ed. Nowcasting. Academic Press,1982:177-190.
- [9] Rosenfeld D. Object method for analysis and tracking of convective cells as seen by radar[J]. J Atmos Oceanic Tech, 1987,4:422-434.
- [10] Johnson J T, Mac Keen P L, Witt A, et al. The storm cell identification and tracking algorithm: an enhanced WSR-88D algorithm[J]. Weather and Forecasting,1998,13:263-276.
- [11] 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,2005:187-197.
- [12] 王芬,李腹广. 多普勒天气雷达冰雹探测算法评估及检验改进[J]. 气象科技,2009,37(3):345-348.
- [13] 胡胜,顾松山,庄旭东,等. 风暴的多普勒雷达自动识别[J]. 气象学报,2006,64(6):797-808.
- [14] 王改利,刘黎平. 多普勒雷达资料在暴雨临近预报中的应用[J]. 气象,2005. 31(10):12-15.
- [15] 朱平,李生辰,肖建设,等. 天气雷达回波外推技术应用研究[J]. 气象,2008,34(1):3-9.