

周康辉,郑永光,王婷波,等,2017. 基于模糊逻辑的雷暴大风和非雷暴大风区分方法[J]. 气象,43(7):781-791.

# 基于模糊逻辑的雷暴大风和 非雷暴大风区分方法\*

周康辉<sup>1,2,3</sup> 郑永光<sup>3</sup> 王婷波<sup>4</sup> 蓝 渝<sup>3</sup> 林 建<sup>3</sup>

1 中国气象科学研究院,北京 100081

2 中国科学院大学,北京 100049

3 国家气象中心,北京 100081

4 中国气象局气象干部培训学院,北京 100081

**提 要:** 雷暴大风往往伴随飑线、阵风锋、龙卷等强对流天气而出现,风速大、发展迅速、突发性强,对生命财产安全造成极大威胁,因此对雷暴大风的监测与预报具有重要的意义。然而,雷暴大风监测一直也是强对流监测的难点。本文在地面气象观测站大风记录的基础上,结合多源数据(包括雷达、卫星、闪电、温度、露点等观测数据),利用模糊逻辑算法,实现雷暴大风与非雷暴大风的有效识别,可对雷暴大风进行实时监测。具体算法为:首先,基于历史样本数据的统计得到各变量的概率分布函数,进而得到各参数隶属度函数;然后采用概率重叠面积方法,确定各项质量控制数据的权重;最后通过选取判断概率阈值  $Q$  的方法,区分雷暴大风与非雷暴大风。通过对 2010 年全国 50873 条人工观测大风数据的识别结果检验表明,该算法能有效区分雷暴大风与非雷暴大风,当  $Q$  选取 0.55 时,雷暴大风的识别准确率 POD 约为 0.76,误识别率约为 0.18,雷暴大风 CSI 指数约为 0.67。文中选取了两次大风过程,算法正确地识别了 11 个非雷暴大风记录,5 个雷暴大风记录。本工作能一定程度上提升雷暴大风的监测效果、完善强对流监测业务体系。

**关键词:** 雷暴大风,强对流,模糊逻辑

**中图分类号:** P456.413

**文献标志码:** A

**DOI:** 10.7519/j.issn.1000-0526.2017.07.002

## Fuzzy Logic Algorithm of Thunderstorm Gale Identification Using Multisource Data

ZHOU Kanghui<sup>1,2,3</sup> ZHENG Yongguang<sup>3</sup> WANG Tingbo<sup>4</sup> LAN Yu<sup>3</sup> LIN Jian<sup>3</sup>

1 Chinese Academy of Meteorological Sciences, Beijing 100081

2 University of Chinese Academy of Sciences, Beijing 100049

3 National Meteorological Centre, Beijing 100081

4 China Meteorological Administration Training Centre, Beijing 100081

**Abstract:** Thunderstorm gale (TG) monitoring in the severe convective weather is a complex and important task, in which the difficulty is how to distinguish the TG from non-thunderstorm gale (NTG). Based on the multisource data, including radar, satellite, lightning, temperature and dew-point temperature, this paper proposes a fuzzy logic algorithm to tell them apart, which is proved to be an effective and efficient method. First, get the member functions of the multisource data according to their probability distribution which were extracted from long-term historical data. Second, acquire the weight ratio of each data by calculating the overlap areas of probability distributions. Finally, get the TG probability  $Q$ , and choose a

\* 公益性行业(气象)科研专项(GYHY201406002)、国家重点基础研究发展计划(973 计划)(2013CB430106)和国家自然科学基金项目(41375051)共同资助

2016 年 11 月 7 日收稿; 2017 年 3 月 31 日收修定稿

第一作者:周康辉,主要从事强对流天气研究. Email:zhoukh@cma.gov.cn

threshold of  $Q$  to distinguish TG from NTG. In order to evaluate its performance, the algorithm is used to find TGs in the 50873 gale records of China in 2010. The results show that when  $Q$  is 0.55, the POD of TG is 0.76, and the FAR of TG is 0.18, and the CSI of TG is about 0.67. Two mixing weather processes, caused by cold air and typhoon, are chosen to evaluate its performance, showing 11 NTGs and 5 TGs are correctly identified. The algorithm would enhance the accuracy and effectiveness of the severe weather monitoring significantly.

**Key words:** thunderstorm gale, severe convective weather, fuzzy logic

## 引 言

强对流天气局地性强、发展迅速,天气现象发展剧烈,伴随出现的冰雹、雷暴大风与短时强降水等天气,具有很强的致灾性。中央气象台定义的雷暴大风是指伴随强对流风暴出现的风速超过  $17.2 \text{ m} \cdot \text{s}^{-1}$  的大风天气。雷暴大风往往伴随飏线、阵风锋、龙卷等强对流天气而出现,风速大、发展迅速、突发性强,对生命财产安全造成极大威胁,因此对雷暴大风的监测与预报具有极其重要的意义。对雷暴大风进行实时、有效的监测是对其进行预报与研究的基础。然而,目前我国的气象监测体系中只包含温、压、湿、风等常规气象要素与闪电、降水等天气现象的观测,对于雷暴大风这样一个综合天气现象,并无特定的观测。

雷暴大风往往包括下击暴流大风、阵风锋大风等(郑永光等,2013),其识别的难点在于如何将其与冷空气大风、台风大风、高山站大风、海岛站大风等非强对流大风区分开来。由于缺乏直接的监测数据,往往利用其他数据进行综合判断。2014年之前,将与雷暴、冰雹等强对流天气同时出现的大风视为雷暴大风记录(余蓉等,2012;郑永光等,2013)。国外也采用类似的方法,Lombardo et al(2009)利用极值分析法,结合雷暴人工观测,通过匹配大风与雷暴发生时间,区分雷暴大风与非雷暴大风。2014年之后,由于地面观测调整,取消了雷暴的人工观测(中国气象局综合观测司,2013),使利用雷暴观测数据不再可行。

另一方面,利用天气雷达、卫星云图数据可对雷暴大风等天气现象进行一定识别(吴翠红等,2012;方翀等,2014)。王萍和牛智勇(2014)利用单站雷达数据,发现中层径向辐合(MARC)对雷暴大风、冰雹、短时强降水的识别具有较好的指示性。李国翠等(2013;2014)利用雷达三维组网数据,统计分析了

对流性地面大风的6个雷达识别指标:最大反射率因子、最大垂直积分液态水含量、垂直积分液态水含量随时间变率、风暴最大反射率因子下降高度、风暴体移动速度和垂直积分液态水含量密度等参数,建立隶属函数和权重系数,建立了具有模糊逻辑的对流性地面大风识别方法。以上这些单纯基于雷达数据的雷暴大风识别方法,虽然在一定程度上能够实现雷暴大风的区域性监测,但尚无法实现对雷暴大风的全国监测。

本文尝试在地面气象观测站大风观测记录的基础上,结合多源数据(包括温度、露点等地面气象站观测数据以及闪电、雷达、卫星等遥感数据),利用模糊逻辑算法,建立一套全国范围的雷暴大风自动识别的算法,为我国强对流监测体系实现雷暴大风的实时监测。

## 1 数据与方法

### 1.1 质量控制数据

本算法使用的测风数据来自全国多要素(包括温、压、湿、风)自动气象站,全国约50000个站点。利用国家雷电监测网络的地闪定位数据、全国多普勒雷达基本反射率拼图数据、风云气象卫星的红外亮温等数据对雷暴大风进行识别。

雷暴大风与非雷暴大风训练样本数据使用来自国家气象信息中心提供的826个气象观测站(包括国家基准气候站、国家基本气象站)的“中国气候日值数据集”。该数据集能提供观测站点的人工天气现象观测数据,包括大风、雷暴、冰雹、龙卷等。用冰雹、雷暴、龙卷风等天气现象对大风数据进行识别,得到雷暴大风数据,其作为样本训练数据。具体方法为:对于出现大风的站点,如果大风记录前后1小时有雷暴、冰雹、龙卷等强对流天气现象记录,则认为是雷暴大风天气。

## 1.2 区域划分与季节划分

我国国土面积辽阔,各地气候各异,发生雷暴大风的条件具有较大差异。按照气候特征和便于统计分析原则,将我国大陆地区划分为 7 个区域:东北、西北、华北、华中、华南、西南和西藏(图 1)。

雷暴大风具有较强的季节性,一般而言,春季与夏季为雷暴大风的多发期,秋季发生频次急剧降低,冬季基本无雷暴大风的发生(孙继松等,2013)。为了进一步增加温度与露点对于雷暴大风的区分度,按照我国的气温与露点随月份的变化,本文对季节进行了划分,将 4、5、6、7、8、9 月划分为暖季,10、11、12、1、2、3 月划分为冷季(张世轩等,2011)。本文分别统计暖季和冷季的隶属度函数,进行雷暴大风质量控制(以下简称质控)。

## 1.3 模糊逻辑算法

1965 年,美国数学家 Zadeh(1965;1975)首先提出了模糊集合的概念,标志着模糊数学的诞生。模糊逻辑善于表达界限不清晰的定性知识与经验,借助隶属度函数概念,区分模糊集合,处理模糊关系,模拟人脑实施规则型推理,在气象、工程控制等领域有着广泛的应用(Cho et al, 2006; Gourley et al, 2007)。

由于雷暴大风出现的物理环境仍属于“模型未知”或“不能确定的描述系统”,无法设定精确的阈值判断其是否发生,此时运用“模糊逻辑”,可模仿人脑的方式,实现判断。本文应用模糊逻辑,对大风出现站点对应的温度、露点和周围闪电数量、雷达回波强度与卫星红外亮温等情况,根据大量数据的统计分布规律,确定各项质控物理量参数的质控权重,最终实现对雷暴大风与非雷暴大风的有效识别。

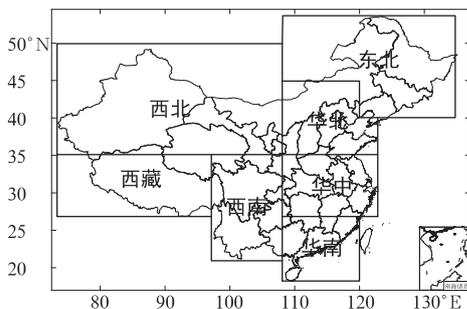


图 1 我国大陆地区分区情况

Fig.1 The area division in Chinese mainland

## 2 雷暴大风识别算法

算法具体流程(图 2)主要包括 4 步:

(1) 利用统计法得到隶属度函数与权重系数: 本文使用 2011—2013 年的全国 826 个观测站 171188 个的大风观测记录提取雷暴大风与非雷暴大风的概率密度分布函数,计算得到隶属度函数与权重函数。

(2) 提取大风记录对应的各项特征参数,包括卫星云顶红外亮温、雷达回波强度、闪电密度、温度、露点等。

(3) 计算得到各项特征参数的隶属度与权重,计算得到概率  $Q$ 。将第二步得到的各项特征参数代入隶属度函数,得到对应隶属度,综合隶属度与权重函数,即可计算雷暴大风概率  $Q$ 。

(4) 通过对历史数据的分析,选取最优的概率阈值 Threshold,用于区分雷暴大风与非雷暴大风。如果  $Q > \text{Threshold}$ ,则判定为雷暴大风;如果  $Q < \text{Threshold}$ ,则判定为非雷暴大风。

### 2.1 隶属度函数

隶属度函数的建立是模糊逻辑算法的基础。隶属度定义为若对论域(研究的范围) $U$ 中的任一元素  $x$ ,都有一个数  $A(x) \in [0, 1]$ 与之对应,则称  $A$  为  $U$  上的模糊集,  $A(x)$  称为  $x$  对  $A$  的隶属度。隶属度函数的确立方法有模糊统计法、专家经验法以及例证法等(王季方和卢正鼎,2000)。本文采用模糊统计法,通过对特征进行大样本概率分布统计,得到隶属度函数。由于样本数量有限,概率密度函数分布不规则,为了得到其更好的概率分布特征,本文利用高斯核密度函数(Silverman, 1986)进行平滑拟合,具体公式如下:

$$\hat{f}(x) = \frac{1}{\sigma \sqrt{2\pi}} \sum_{i=1}^n e^{-\frac{1}{2}(\frac{X_i - x}{\sigma})^2} \quad (1)$$

式中,  $\hat{f}(x)$  为概率密度函数,  $n$  为数据量,  $X_i$  为  $i$ th 观测量,  $x$  为独立变量。  $\sigma$  为平滑因子,利用 Silverman's rule (Silverman, 1986),取  $\sigma = 1.06 SD \cdot n^{-\frac{1}{5}}$ ,其中  $SD$  为样本的标准方差。

得到其雷暴大风的概率分布函数  $F_{TW_i}$  和非雷暴大风的概率分布函数  $F_{NW_i}$  之后,可得到其隶属度函数  $MF_i$ :

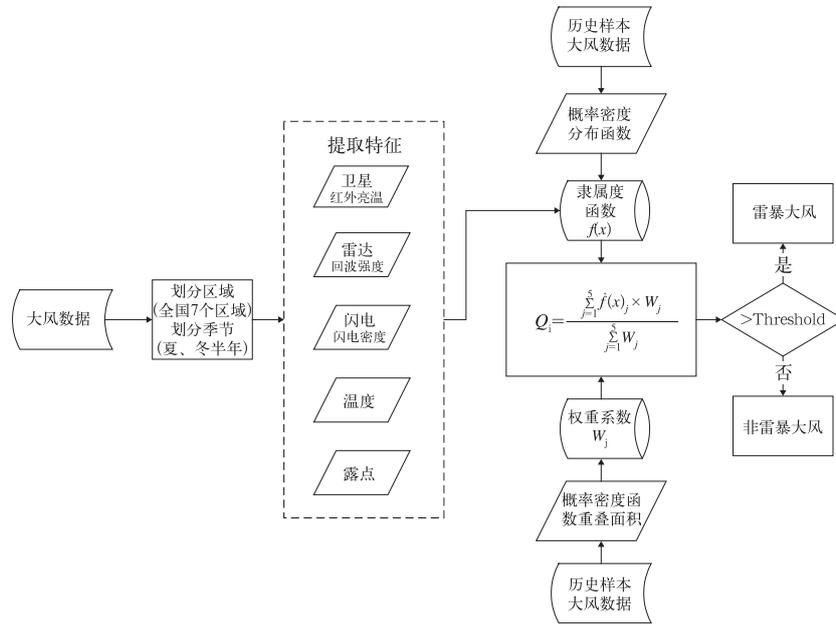


图 2 算法流程

Fig. 2 Algorithm structure

$$MF_i = \frac{F_{TW_i}}{F_{TW_i} + F_{NW_i}} \quad (2)$$

隶属度函数生成之后,根据其分布特征,温度和露点采用了升半岭形分布(式 3)对其进行拟合:

$$MF_i = \begin{cases} 0 & x \leq a \\ \frac{1}{2} + \frac{1}{2} \sin \frac{\pi}{b-a} & a < x < b \\ 1 & x \geq b \end{cases} \quad (3)$$

式中  $a, b$  为拟合参数。

### 2.1.1 雷达回波强度

多普勒天气雷达是监测对流系统的基本手段,强对流系统往往具有反射率大、对流云顶高度高、垂直液态水含量较大的特点(冯桂力等,2010;方翀等,

2015)。大风地区的雷达回波特征,可判断大风是否为对流系统而产生。由于阵风锋等大风速区往往距离强对流主体存在一定的空间距离(徐芬等,2016),因此本文算法中取大风记录点半径 20 km 之内的最大雷达回波参数。由于监测范围较大,因此本文中雷达数据使用的是中国气象局气象探测中心提供的全国雷达基本反射率拼图产品。

由图 3 可见,雷达回波强度对两种大风的区分度较高。非雷暴大风的雷达回波主要分布在 10 dBz 以下,而雷暴大风则主要分布超过 30 dBz,峰值分布于 53 dBz,对应于雷暴大风较强的对流系统。隶属度函数显示,当雷达回波强度超过 30 dBz,其为雷暴大风的概率达到 50%,当雷达回波强度超过

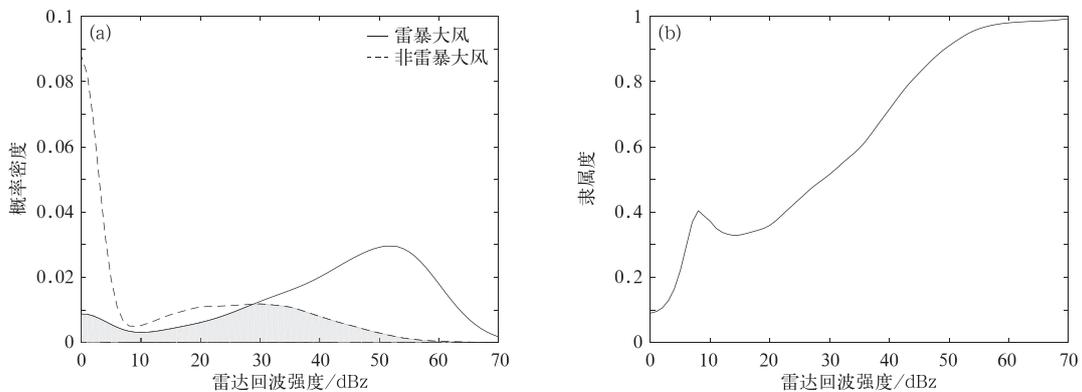


图 3 雷达回波的(a)概率密度函数和(b)隶属度函数

Fig. 3 The probability frequency (a) and function member (b) of radar echo

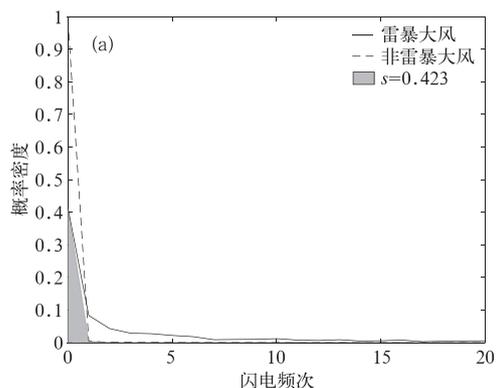
50 dBz,其隶属度超过 90%;隶属度函数在 10 dBz 左右出现一个概率极大值,与对流初生、远离对流主体的阵风锋等环境,雷达回波较弱的情况下的对流性大风相对应。

### 2.1.2 闪电密度

闪电是雷暴中的经常伴随的天气现象(俞小鼎等,2005;郑媛媛等,2004)。雷暴大风必定伴随雷暴等对流系统出现,因此闪电是区分雷暴大风与非雷暴大风的关键因素。本文的统计,取大风记录点半径 21 km(周康辉等,2014)的闪电数量,得到雷暴大风与非雷暴大风的闪电概率密度分布与隶属度函数(图 4)。

闪电对于雷暴大风的具有较好的区分性,其重叠面积  $s=0.423$ 。从概率密度函数来看,超过 60% 的雷暴大风都伴随闪电现象;非雷暴大风的大风天气,也会出现伴随出现闪电的现象,但是其概率极低,累积概率低于 3%。

由于以上原因,闪电的隶属度函数在质控过程中权重较大,可作为识别雷暴大风的重要参量。



### 2.1.3 红外亮温

风云气象卫星能实现全国范围的可见光云图、红外云顶亮温、水汽等气象要素观测。对于监测大范围的中尺度对流系统(mesoscale convective system, MCS)具有独特的优势(郑永光等,2008)。利用大风记录点的云顶红外亮温,同样可有效判断该大风是否伴随强对流系统而出现(方翀等,2014;费海燕等,2016)。与雷达数据相似,算法中同样取大风记录点半径 20 km 之内的云顶红外亮温参数。

由图 5 可见,红外亮温对于雷暴大风与非雷暴大风具有很好的区分性。雷暴大风的云顶红外亮温概率密度峰值出现在  $-50^{\circ}\text{C}$ ,而非雷暴大风的峰值出现在  $0^{\circ}\text{C}$  左右。隶属度函数可见,若红外亮温低于  $-25^{\circ}\text{C}$ ,其为雷暴大风的概率超过 50%;当红外亮温低于  $-50^{\circ}\text{C}$ ,其为雷暴大风概率超过 70%。

### 2.1.4 地面温度

雷暴大风与非雷暴大风地面温度具有一定差异。冷空气大风出现在较强冷空气过程中,因此其地面温度往往较低。因此,温度可作为区分冷空气

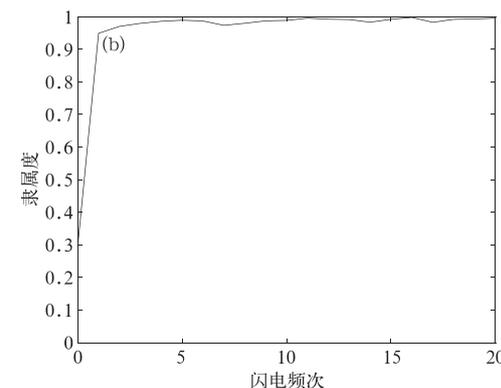


图 4 同图 3,但为闪电频次  
Fig. 4 Same as Fig. 3, but for lightning frequency

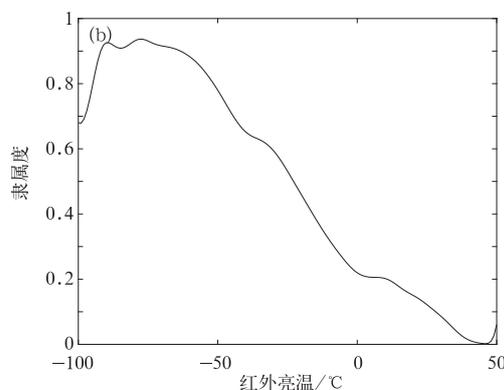
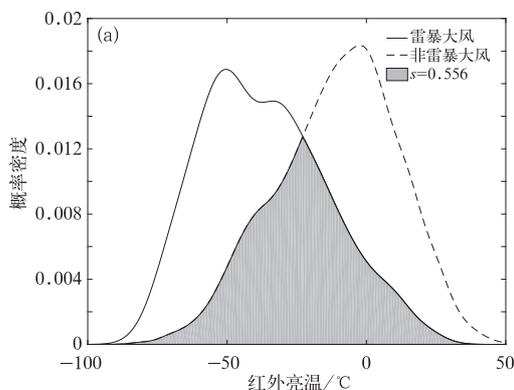


图 5 同图 3,但为红外亮温  
Fig. 5 Same as Fig. 3, but for satellite infrared brightness temperature

大风与雷暴大风的重要参量。

由于篇幅限制,无法给出全国每个区域的概率

分布函数和隶属度函数,这里以西南和西北地区的情况为例(图 6)。由图 6 可见:

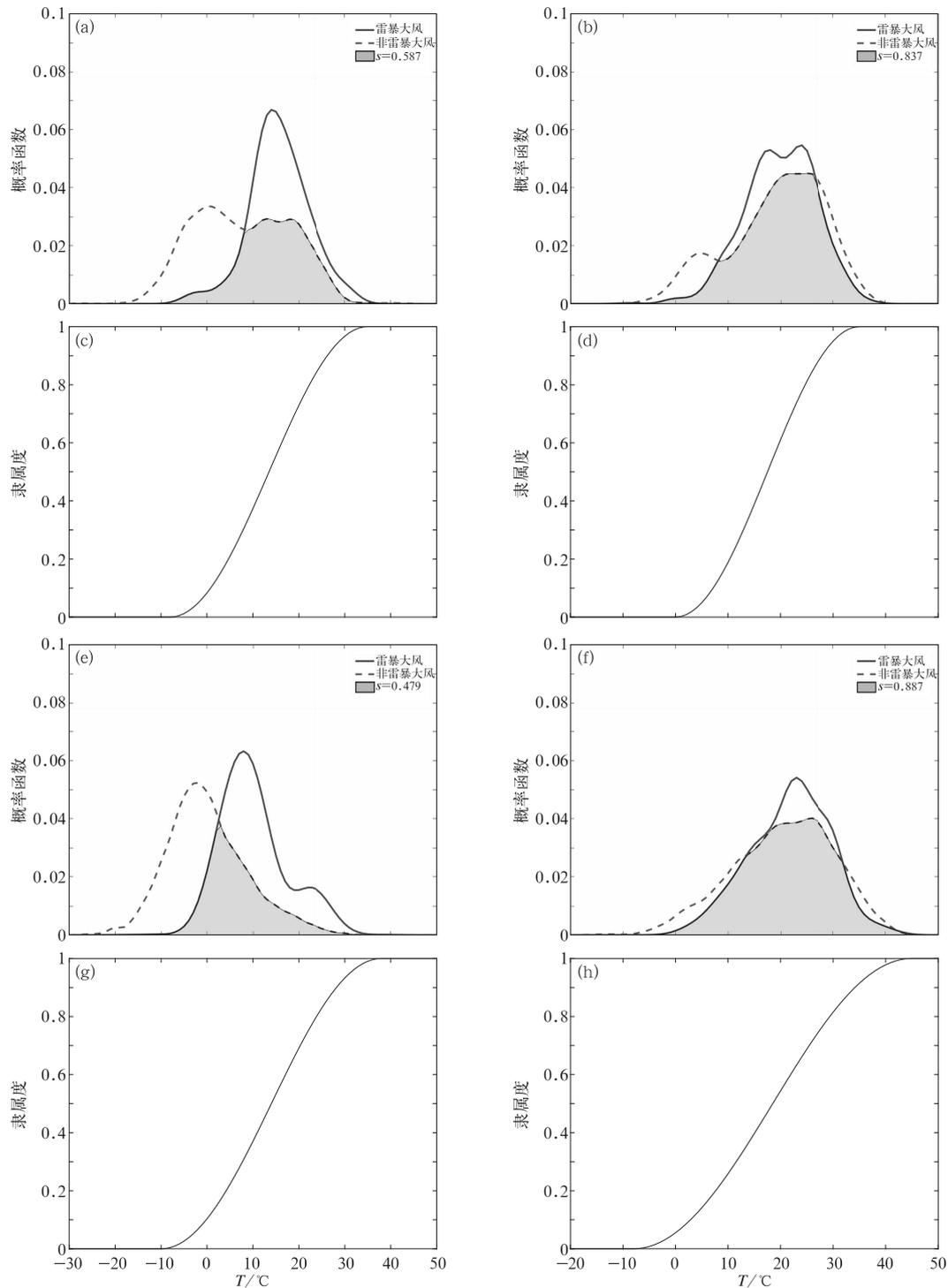


图 6 温度的概率分布函数(a, b, e, f)以及隶属度函数(c, d, g, h)

(a, c)西南地区冷季, (b, d)西南地区暖季, (e, g)西北地区冷季, (f, h)西北地区暖季

Fig. 6 The probability distribution function of temperature (a, b, e, f) and function member (c, d, g, h)

(a, c) southwest area in cold season, (b, d) southwest area in warm season,

(e, g) northwest area in cold season, (f, h) northwest area in warm season

(1) 西南地区冷季,雷暴大风最易出现的温度为 $15^{\circ}\text{C}$ 左右,非雷暴大风出现概率最高的温度为 $0^{\circ}\text{C}$ 左右。冷季,两者重叠面积为 $0.587$ ,具有一定的区分度;暖季,雷暴大风与非雷暴大风出现概率最高的温度均在 $25^{\circ}\text{C}$ 左右,两者重叠面积达到 $0.837$ ,温度对于两者的区分度明显较冬季更低。

(2) 西北地区冷季雷暴大风出现概率最高的温度为 $9^{\circ}\text{C}$ ,非雷暴大风为 $-2^{\circ}\text{C}$ ,两者区分度为 $0.479$ ;暖季,区分度下降至 $0.887$ ,雷暴大风与非雷暴大风最易出现的温度均在 $25^{\circ}\text{C}$ 左右。

由此可见,不同地区、不同季节,温度对于雷暴大风与非雷暴大风的区分度具有较大差异。冬季,雷暴大风发生的温度明显偏高,夏季则与非雷暴大风差异较小;同样的温度,西南地区与西北地区发生雷暴大风的概率也有显著的差异。由此可见,对于全国分区、分季节处理,是很有必要的。

### 2.1.5 地面露点

地面露点是表征地面湿度的重要参量。露点越高,地面水汽含量越大。雷暴大风发生的环境特征需要一定的水汽条件(王秀明等,2013;钟利华等,2011),因此,露点可以作为雷暴大风识别的重要参数。

图7可见,露点温度相对于温度具有更好的区分性:

(1) 冷季,西南地区雷暴大风的概率最大值出现在露点为 $5^{\circ}\text{C}$ ,非雷暴大风则在露点为 $-20^{\circ}\text{C}$ 和 $2^{\circ}\text{C}$ 呈现双峰结构,两者重叠面积仅为 $0.395$ ;暖季,西南地区雷暴大风与非雷暴大风的最大概率均出现在露点为 $8^{\circ}\text{C}$ 左右,重叠面积达到 $0.765$ ,区分性较差。

(2) 西北地区情况与西南地区类似,但是露点呈现出更好地区分布;冷季,雷暴大风最易出现在露点为 $-1^{\circ}\text{C}$ 的情况,而非雷暴大风最易出现在 $-15^{\circ}\text{C}$ ,概率密度曲线重叠面积为 $0.322$ ;暖季,雷暴大风与非雷暴大风的露点区分度变差,重叠面积为 $0.594$ 。

西北地区与西南地区露点温度的差异与气候分布一致。雷暴大风一般发生于强对流环境之中,因此必须具备一定的湿度条件,要求露点温度较高;非雷暴大风,如冷空气大风,则一般发生在较干冷的环境场中。这一差异在冷季显得更为明显。

## 2.2 权重

对于参与统计识别的物理量,包括温度、露点、

闪电密度、雷达回波强度、云顶红外亮温,每个物理量对于雷暴大风的区分度有所差异,比如闪电对雷暴大风具有较高的区分度,当出现闪电的时候,大风记录为雷暴大风的记录超过 $60\%$ 。相对而言,露点和温度的区分度相对较差。因此,需要确定各个物理量在质量控制中的权重 $W$ ,参照Gourley et al (2007)的方法,其计算公式如下:

$$W_j = \frac{1}{A_j} \sum_{j=1}^5 \frac{1}{A_j} \quad (4)$$

式中, $A_j$ 为雷暴大风与非雷暴大风的概率密度分布函数的重叠面积(如图3~7中的阴影部分)。若 $A$ 越小,该参数对与雷暴大风与非雷暴大风的区分度就越佳,其参与质控的比重就越大;反之,则参与质控的比重就越小。

表1为西北地区与西南地区各项物理量参与质控的权重分布情况。可以看到,冷季与暖季具有较大差异。冷季,西北地区露点与温度的质控权重较大,分别达到 $25.69\%$ 与 $17.27\%$ ;暖季,其权重下降,分别为 $17.35\%$ 与 $11.61\%$ 。雷达、闪电、卫星数据的权重系数在暖季高于冷季,其中雷达的权重系数最大,其次为闪电。西南地区情况类似,温度与露点权重相对较低,雷达、卫星、闪电数据权重相对较高。

## 2.3 模糊解码

确定了各个物理量的权重系数之后,则可计算得到总的概率系数 $Q$ :

$$Q_i = \frac{\sum_{j=1}^5 \hat{f}(x)_j \times W_j}{\sum_{j=1}^5 W_j} \quad (5)$$

式中, $W_j$ 为权重系数, $\hat{f}(x)_j$ 为隶属度。

## 2.4 模糊阈值的确定

通过隶属度与权重函数等计算流程后,最终得到概率 $Q$ ,如何通过 $Q$ 判断雷暴大风与非雷暴大风,这里通过选取阈值Threshold进行判断。

首先,引入识别正确率(POD)、误识别率(FAR)、临界成功指数(CSI)等参数,其定义如下:

$$POD = \frac{h}{h+m} \quad (6)$$

$$FAR = \frac{f}{h+f} \quad (7)$$

$$CSI = \frac{h}{h + m + f} \quad (8)$$

式中,  $h$  为正确识别的数量,  $f$  为误识别为雷暴大风数量,  $m$  为雷暴大风漏识别数量。

图 8 给出了 2010 年共 50873 个大风样本在不

同阈值情况下的 POD、FAR、CSI 等参数的分布情况。随着 Threshold 值的增大, POD 值呈现出递减的趋势, 正确识别数量减小, 漏识别数量增多; FAR 值同样呈现降低的趋势, 因为随着 Threshold 的增大, 被错误识别为雷暴大风的大风数量降低; CSI 呈

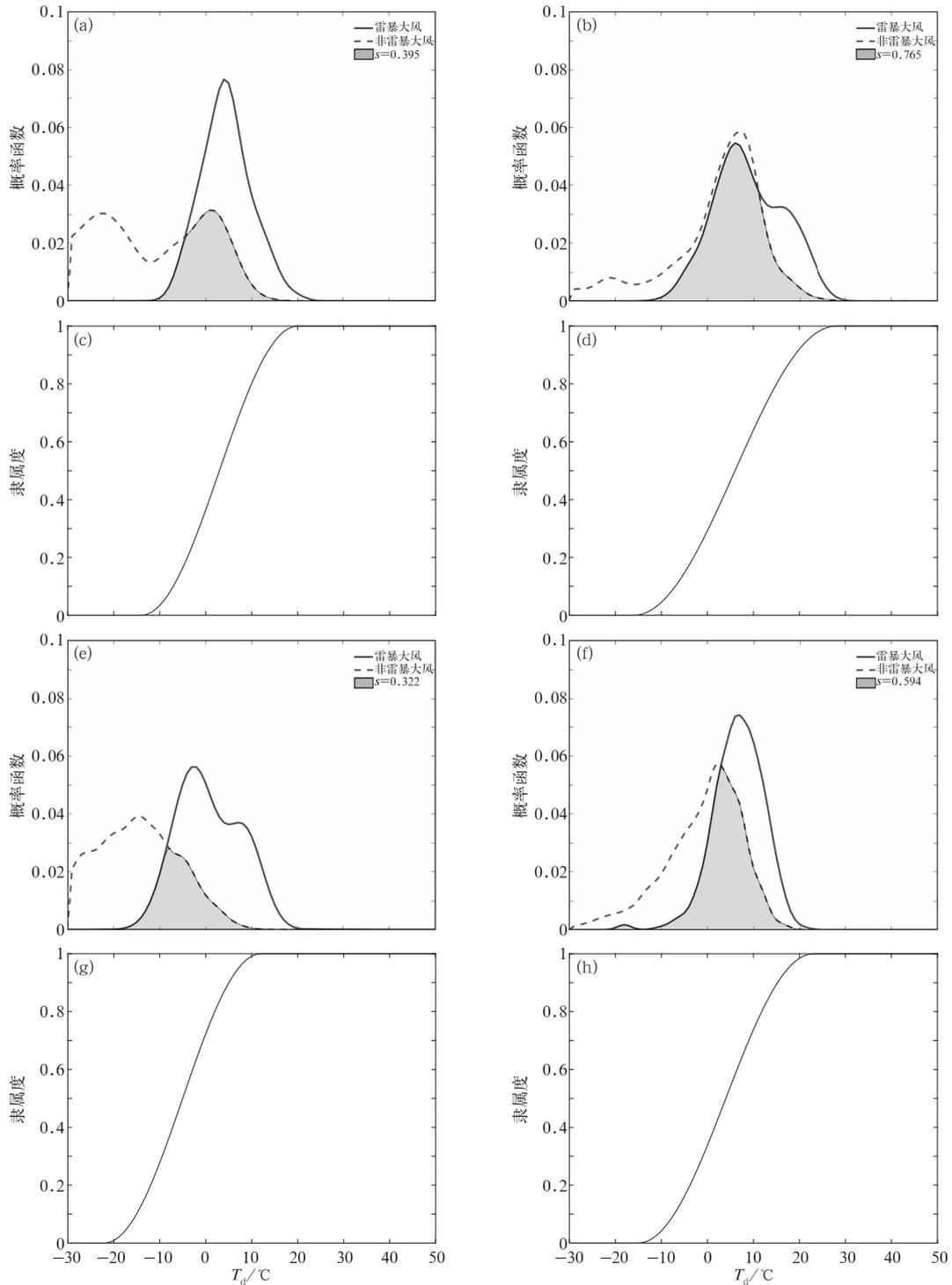


图 7 同图 6, 但为露点温度

Fig. 7 Same as Fig. 6, but for dew point temperature

表 1 两个地区各物理量质控权重比例  
Table 1 Physical quality control weight ratio in Northwest area and Southwest area

地区	参数类型	重叠区域面积		质控权重/%	
		冷季	暖季	冷季	暖季
西北地区	闪电	0.423		19.56	24.35
	卫星亮温	0.556		14.88	18.53
	雷达回波	0.366		22.60	28.15
	温度	0.479	0.887	17.27	11.61
	露点	0.322	0.594	25.69	17.35
西南地区	闪电	0.423		21.23	25.16
	卫星亮温	0.556		16.16	19.14
	雷达回波	0.366		24.55	29.07
	温度	0.587	0.837	15.31	12.72
	露点	0.395	0.765	22.75	13.91

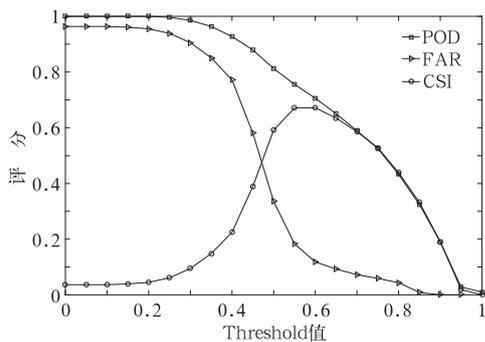


图 8 雷暴大风识别结果检验情况

Fig. 8 Verification of different threshold values of thunderstorm gale identification

现出先增加后降低的趋势,在  $Threshold = 0.55 \sim 0.6$  时达到最大值 0.67,此时雷暴大风识别准确率约为 0.76,误识别率约为 0.18。

由此可见,Threshold 值的选取对于雷暴大风的识别至关重要,若太高,则会导致雷暴大风漏识别的情况太多;若太低,则导致误识别率过高。实际工作中,可灵活选取。

### 3 识别结果检验

#### 3.1 检验统计结果

2010 年共 50873 个大风样本,图 9 给出了识别概率  $Q$  的分布情况。由图 9 可见,雷暴大风与非雷暴大风的识别概率具有非常明显差别。超过 93% 的雷暴大风的  $Q$  值超过 0.4,约 78% 的雷暴大风的  $Q$  值超过 0.5;与此相对应,约 91% 的非雷暴大风识别概率低于 0.4,非雷暴大风的  $Q$  超过 0.5 的比例

急剧降低,低于 1%。若  $Threshold = 0.42$ ,算法能识别 92% 的雷暴大风,同时能识别 93% 非雷暴大风;若  $Threshold = 0.55$ ,识别到的雷暴大风比例降至 76%,但是非雷暴大风识别率上升至 99%。

#### 3.2 天气过程识别检验

为了验证上述方法的效果,本节选取了 2013 年 5 月 18 日 16:10 与 8 月 13 日 14:50 两次大风过程。其中,2013 年 5 月 18 日大风过程发生在内蒙古、陕西等地,大风过程以冷空气大风为主,夹杂着强对流天气大风。人工观测记录到大风 10 站次,其中非雷暴大风 8 站次,雷暴大风 2 站次。2013 年 8 月 13 日大风过程属于热带气旋外围的大风。

从雷达拼图(图 10a)与卫星云图(图 10b)可以看到,2013 年 5 月 18 日,冷空气南下,在冷空气前沿,触发东北—西南向狭长的强对流。非雷暴大风主要发生于以红外亮温暗区为特征的冷空气后部;雷暴大风发生于对流系统附近,对流系统周围存在较弱的闪电活动,探测到为数不多的地闪(图 10a)。2013 年 8 月 13 日的大风过程属于台风尤特外围云系触发一系列强对流系统,引发雷暴大风;同时,台风导致的非对流性大风与雷暴大风同时存在。

利用模糊逻辑算法,其各项特征量与识别情况如表 2。由表 2 中可以看到,2013 年 5 月 18 日 16:10 过程温度露点差较大、大气环境较干,属于典型的冷空气南下大风过程。相对而言,非雷暴大风的露点更低,  $T - T_d$  较大,空气层结较干;雷暴大风周围存在对流系统的发生发展,因此其雷达回波强度与红外亮温均显示着对流云的存在,而非雷暴大风对应冷空气大风,底层雷达探测回波反射率为 0,卫

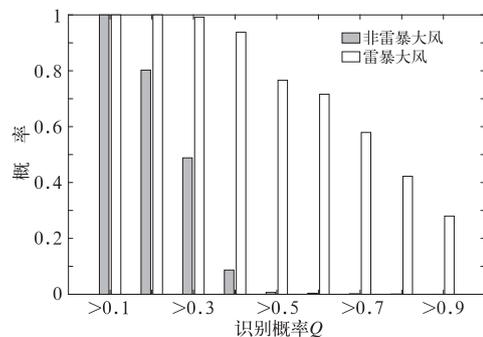


图 9 雷暴大风与非雷暴大风  $Q$  值分布

Fig. 9 Distribution of  $Q$  values of thunderstorm gale (TG) and non-thunderstorm gale (NTG)

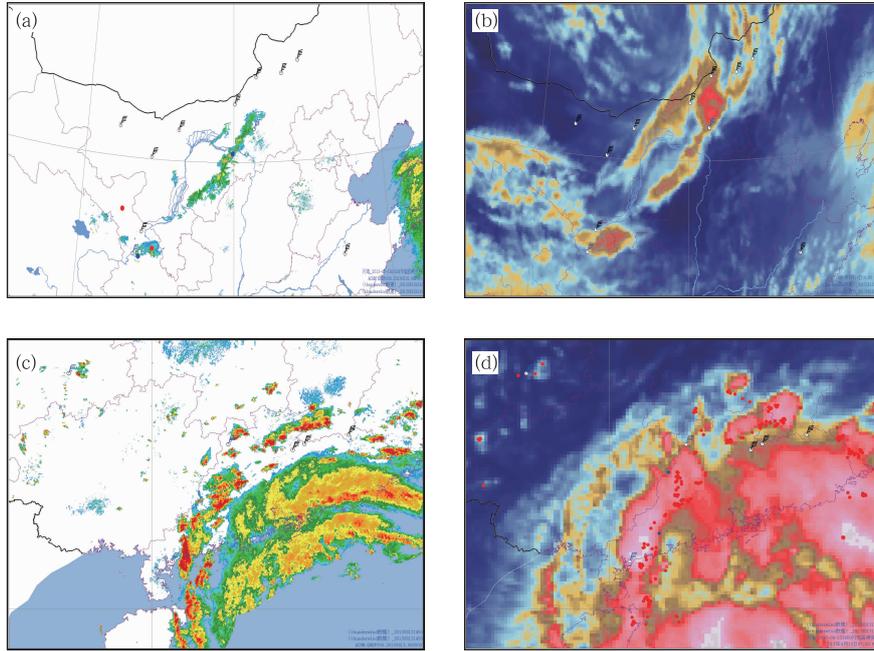


图 10 2013 年 5 月 18 日 16:10(a,b)与 8 月 13 日 14:50(c,d)  
一次大风过程的识别情况

(蓝色为雷暴大风,黑色为冷空气大风;圆点为闪电,红色为正闪,蓝色为负闪)

Fig. 10 Identification result of one high wind case at (a, b) 16:10 BT 18 May  
and (c, d) 14:50 BT 13 August 2013

(Blue stands for TG, black is for NTG; red points represent positive lightning,  
and blue ones are negative lightning)

表 2 两次大风过程的识别情况

Table 2 Identification results of two gale cases

时间	类型	温度/°C	露点温度/°C	雷达回波强度/dBz	红外亮温/°C	闪电数量	Q
2013 年 5 月 18 日 16:10	非雷暴大风	20	-2	0	17	0	0.29
		24	-6	0	-33	0	0.31
		26	-9	0	-47	0	0.39
		20	-1	0	0	0	0.32
		19	-3	0	-45	0	0.34
		27	-8	0	-32	0	0.37
		24	-24	0	24	0	0.24
	27	-6	0	-36	0	0.38	
2013 年 8 月 13 日 14:50	雷暴大风	21	-4	30	-48	0	0.55
		20	3	25	-41	1	0.62
	非雷暴大风	33	10	0	-20	0	0.42
		32	15	0	-21	0	0.43
		33	16	0	-23	0	0.44
雷暴大风	33	14	40	3	14	0.73	
	33	23	50	-74	0	0.74	
33	17	55	-58	5	0.88		

星探测显示高空存在一定的高层云系;非雷暴大风站点附近均无闪电发生,雷暴大风站点中有 1 个站点周围探测到 1 个闪电。

2013 年 8 月 13 日 14:50,各大风记录的温度和露点较为接近,但是其他 3 个参数具有较大差异。非雷暴大风周围均未探测到闪电,同时距离对流云

存在一定距离,雷达回波为0,红外亮温较高,Q值均在0.5以下。相对而言,雷暴大风记录周围均探测到比较明显的对流活动,Q值均超过0.7。

总体而言,可以看到选取个例中Q具有较好的区分性,若Threshold选取为0.55,能较好地实现雷暴大风与非雷暴大风的分类。

## 4 结论与讨论

本文在地面观测站大风记录的基础上,结合多源数据,包括雷达、卫星、闪电、温度、露点等观测数据,利用模糊逻辑算法,实现了对雷暴大风与非雷暴大风的有效识别。算法利用统计方法得到2011—2013年多源数据的概率分布函数,进而计算得到其隶属度函数。采用概率重叠面积方法,确定各项质控数据的权重,最后通过选取概率阈值的方法,实现对雷暴大风的识别。

通过对2010年50873条人工观测历史数据的识别结果检验表明,该算法能有效区分雷暴大风与非雷暴大风,当判断阈值选取0.55时,对于雷暴大风的识别正确率POD约为0.76,误识别率约为0.18,CSI指数约为0.67,对于非雷暴大风POD约为0.99,对于强对流监测业务应用具有一定的参考意义。文中选取了两次大风过程,算法正确地识别了11个非雷暴大风记录,5个雷暴大风记录,表现出一定的实用性。

在强对流监测中,雷暴大风由于其局地性与强致灾性一直是监测的难点与重点。本文使用模糊逻辑算法,建立了较为科学合理的识别模型,实现了雷暴大风与非雷暴大风的区分,为强对流监测引入了新的思路。算法具有较好的兼容性,后续工作可以尝试将更多的数据加入质控数据源,并将其用于短时强降水与冰雹等强对流监测数据的质控,建立科学合理的强对流监测体系。

## 参考文献

方翀,俞小鼎,朱文剑,等,2015. 2013年3月20日湖南和广东雷暴大风过程的特征分析[J]. 气象,41(11):1305-1314.

方翀,郑永光,林隐静,等,2014. 导致区域性雷暴大风天气的云型分类及统计特征分析[J]. 气象,40(8):905-915.

费海燕,王秀明,周小刚,等,2016. 中国强雷暴大风的气候特征和环境参数分析[J]. 气象,42(12):1513-1521.

冯桂力,王俊,牟容,等,2010. 一次中尺度雷暴大风过程的闪电特征分析[J]. 气象,36(4):68-74.

李国翠,刘黎平,连志鸾,等,2014. 利用雷达回波三维拼图资料识别

雷暴大风统计研究[J]. 气象学报,72(1):168-181.

李国翠,刘黎平,张秉祥,等,2013. 基于雷达三维组网数据的对流性地面大风自动识别[J]. 气象学报,71(6):1160-1171.

孙继松,戴建华,何立富,等,2013. 强对流天气预报的基本原理与技术方法——中国强对流天气预报手册[M]. 北京:气象出版社.

王季方,卢正鼎,2000. 模糊控制中隶属度函数的确定方法[J]. 河南科学,18(4):348-351.

王萍,牛智勇,2014. 基于多普勒天气雷达数据的中层径向辐合自动识别及其与强对流天气的相关性研究[J]. 物理学报,63(1):019201.

王秀明,周小刚,俞小鼎,2013. 雷暴大风环境特征及其对风暴结构影响的对比研究[J]. 气象学报,71(5):839-852.

吴翠红,韦惠红,牛奔,2012. 湖北东部雷暴大风雷达回波特征分析[J]. 大气科学学报,35(1):64-72.

徐芬,杨吉,郑媛媛,等,2016. MIGFA阵风锋识别算法改进与检验[J]. 气象,42(1):44-53.

余蓉,张小玲,李国平,等,2012. 1971—2000年我国东部地区雷暴、冰雹、雷暴大风发生频率的变化[J]. 气象,38(10):1207-1216.

俞小鼎,王迎春,陈明轩,等,2005. 新一代天气雷达与强对流天气预警[J]. 高原气象,24(3):456-464.

张世轩,张璐,孙树鹏,等,2011. 全球变暖情况下中国季节的变化[J]. 高原气象,30(3):659-667.

郑永光,陈炯,朱佩君,2008. 中国及周边地区夏季中尺度对流系统分布及其日变化特征[J]. 科学通报,53(4):471-481.

郑永光,林隐静,朱文剑,等,2013. 强对流天气综合监测业务系统建设[J]. 气象,39(2):234-240.

郑媛媛,俞小鼎,方翀,等,2004. 一次典型超级单体风暴的多普勒天气雷达观测分析[J]. 气象学报,62(3):317-328.

中国气象局综合观测司,2013. 地面气象观测业务调整技术规定[Z].

钟利华,曾鹏,李勇,等,2011. 广西雷暴大风环流特征和物理量诊断分析[J]. 气象,37(1):59-65.

周康辉,杨波,毛冬艳,等,2014. 闪电定位数据与人工观测雷暴数据对比分析[J]. 天气预报,(3):46-51.

Cho Y H, Lee G W, Kim K E, et al, 2006. Identification and removal of ground echoes and anomalous propagation using the characteristics of radar echoes[J]. J Atmos Ocean Technol, 23(9):1206-1222.

Gourley J J, Tabary P, du Chatelet J P, 2007. A fuzzy logic algorithm for the separation of precipitating from nonprecipitating echoes using polarimetric radar observations[J]. J Atmos Ocean Technol, 24(8):1439-1451.

Lombardo F T, Main J A, Simiu E, 2009. Automated extraction and classification of thunderstorm and non-thunderstorm wind data for extreme-value analysis[J]. J Wind Eng Ind Aerodyn, 97(3/4):120-131.

Silverman B W, 1986. Density estimation for statistics and data analysis[M]. New York: CRC Press.

Zadeh L A, 1965. Fuzzy sets[J]. Inf Control, 8(3):338-353.

Zadeh L A, 1975. The concept of a linguistic variable and its application to approximate reasoning—I[J]. Inf Sci, 8(3):199-249.