

拉曼激光雷达测温的综合多级质量控制技术

李 博^{1,2}, 华灯鑫^{1*}, 周 艳¹, 张 铭¹, 王玉峰¹, 刘 君¹, 狄慧鸽¹

¹西安理工大学机械与精密仪器工程学院, 陕西 西安 710048;

²中国气象科学研究院灾害天气国家重点实验室, 北京 100081

摘要 利用 10 个方程对拉曼激光雷达系统获得的原始数据进行标定, 进而反演温度, 并分别与同时刻、同有效对等高度层的标准温度进行定性对比。选取 50 组样本, 从定量分析(相似性评判、相异性评判, 以及综合评判)角度, 通过权重系数筛选、误差诊断以及标准化处理等手段对这些样本开展特征统计, 筛选出每一级的质量评判参数与分析方法。所获取的各级评判标准均基于较大样本库, 可以较好地代表激光雷达测温的平均性能。实验样本库之外的实时探测样本检验表明, 综合多级质量控制技术可有效地描述雷达测温数据的优劣; 在不同天气背景下激光雷达的测温效果存在差异。

关键词 遥感; 综合多级质量控制技术; 拉曼激光雷达; 温度; 精度

中图分类号 P41 **文献标识码** A

doi: 10.3788/AOS201737.0428003

Synthetical Multilevel Quality Analysis and Control Technique for Raman Lidar Temperature Detection

Li Bo^{1,2}, Hua Dengxin¹, Zhou Yan¹, Zhang Ming¹, Wang Yufeng¹, Liu Jun¹, Di Huige¹

¹*School of Mechanical and Precision Instrument Engineering, Xi'an University of Technology,*

Xi'an, Shaanxi 710048, China;

²*State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing 100081, China*

Abstract The original data acquired by Raman lidar system is calibrated by ten equations. The temperature is inversed, and it is compared with standard temperature at the same moment and the same effective altitude qualitatively. Then 50 groups of data are selected as statistical samples. From the perspective of quantitative analysis (similarity evaluation, difference evaluation, and comprehensive evaluation), the characteristic statistics of the samples is carried out by weight coefficient, error diagnosis and standard processing. Quality evaluation parameters and analysis method of each level are filtered. The acquired evaluation standards of each level are all based on the sample library with lager amount, which can represent average behavior of the temperature detection of lidar better. The test of real-time detection samples beyond the sample library indicate that the comprehensive multilevel control technique can describe the advantages and disadvantages of temperature data detection of lidar effectively. Temperature detection effect of lidar at different weather conditions exists differences.

Key words remote sensing; synthetical multilevel quality analysis and control technique; Raman lidar; temperature; precision

OCIS codes 280.3640; 280.6780

1 引 言

温度、气压、风、湿度是基本大气参数, 尤以温度更为重要, 一方面直接决定了强对流等灾害天气的发生

收稿日期: 2016-11-14; **收到修改稿日期:** 2016-12-12

基金项目: 国家自然科学基金(61675166, 41575027)、中国气象科学研究院灾害天气国家重点实验室基金(2013LASW-B05)、陕西省自然科学基金基础研究计划(2014JQ5176)、陕西省教育厅自然科学基金专项(15JK1506)

作者简介: 李 博(1979—), 男, 博士, 讲师, 主要从事激光雷达大气探测及数据分析等方面的研究。

E-mail: doctorlee@xaut.edu.cn

* **通信联系人.** E-mail: dengxinhua@xaut.edu.cn

发展,另一方面,多数气象参数的诊断计算都以温度为已知项开展^[1]。

与无线探空仪、微波辐射计、电声测温系统、飞机或火箭、卫星遥感等探测方式相比^[2-3],激光雷达测温的优势在于其单色性较好、方向性较强,从而能有效地避免高空漂移造成的探测误差。从理论上讲,瑞利散射激光雷达、差分吸收激光雷达、高光谱分辨率激光雷达、振动拉曼激光雷达、转动拉曼激光雷达等均可实现大气温度探测^[4]。2002年,Behrendt等^[5]搭建了一台振动-转动拉曼激光雷达系统,实现了11 km高度以下大气温度探测。2003年,Hua等^[6]研制了一台气象观测用激光雷达系统,以特殊设计的高精度多通道法布里-珀罗标准具作为核心分光系统,实现了全天时对流层底层大气温度探测。2004年,Balin等^[7]研发了一台多波长米-拉曼激光雷达系统,以双光栅多色仪为分光器件实现大气温度同步探测。2012年,Leblanc等^[8]设计了多望远镜扫描拉曼激光雷达系统,获得了高达平流层底层的大气参数。国内学者在大气参数(温度、水汽、气溶胶等)探测方面也开展了较多的研究工作^[9-14]。

虽然激光雷达探测技术日趋成熟,但探测数据检测技术方面的系统研究比较少见,而这项工作恰恰是不容忽视的,因为激光雷达探测数据在应用于大气科学等领域前,有必要对其可靠性进行评判。

近年来,西安理工大学设计出振动-转动拉曼激光雷达系统,由此实现了大气温度、湿度及气溶胶的实时探测^[15]。本文在这项工作基础上开展进一步研究,根据气象数据常用检验方法^[2,16-21],通过对激光雷达探测温度开展质量控制技术研究,提出一套雷达测温数据的可靠性评价技术,从而实现有效评判激光雷达测温性能的目标。

2 研究资料与方法

2.1 质量控制与标准数据

全国气象基本信息标准化技术委员会在气象行业规范中指出^[22],质量控制的目的是检查(评判、分析)资料(数据)的质量问题,从而保证资料的准确性。可见,质量控制技术的研究,就是有关待测数据精确性的评判参数与评判标准的分析与选定。

本文以激光雷达探测数据为检测对象,以美国国家环境预报中心(NCEP)的Final Operational Global Analysis(FNL)数据为检测标准。FNL数据提供了每6 h更新的、较为丰富的全球三维网格点近百种大气参数数值,已被我国气象学者在研究中广泛使用。张永涛等^[20]、周青等^[21]分别对FNL数据进行了可靠性检验,对标准化处理后的均方根误差(RMSE)等数据的分析表明在我国范围内FNL温度可以较好地代表实况温度。本文也探讨了西安地区FNL温度的精确性,由图1可见,同一高度层上的FNL温度与常规探空温度异常接近,特别是常规探测仅有10个高度层的温度,而FNL资料在常规资料的缺省高度层上同样给出了温度值(共计26个高度层)。FNL温度不仅精度可靠,而且其空间分辨率明显高于常规探空的温度数据。

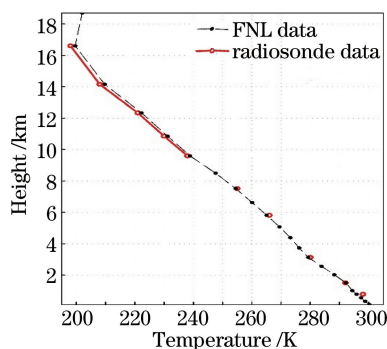


图1 北京时间2013年10月13日20时西安地区FNL温度廓线与常规探空温度廓线

Fig. 1 Atmospheric temperature profile of Xi'an from FNL data and radiosonde data at 20:00 BST on October 13, 2013

2.2 拉曼激光雷达探测技术与10个标定方程温度反演算法

用于温度探测的拉曼激光雷达系统采用Nd:YAG脉冲激光器三倍频输出作为激励光源,激光波长 λ_0 为354.7 nm,重复频率为20 Hz,激光能量为150 mJ。脉冲激光经扩束准直后射向大气。回波信号由直径为600 mm的望远镜接收,并耦合进入多模光纤,经准直透镜后进入由二向色镜和窄带干涉滤光片构成的高

性能拉曼分光系统。该系统可以实现 5 个独立光通道信号的精细分光。其中,第 4 通道和第 5 通道为两路转动拉曼散射信号通道,高低量子通道结合可用于反演大气温度。回波信号经光电倍增管光子计数探测,然后经 Licel 数据采集记录仪输入计算机进行数据反演。该系统的详细介绍可见文献[15]。

温度廓线反演通常采用大气分子反斯托克斯转动拉曼谱线,图 2 为氮气分子反斯托克斯转动拉曼谱线随温度演变趋势。由图 2 可见,随着温度升高(从 200 K 升至 300 K 时),高量子数转动拉曼谱线(远离激光发射波长 354.7 nm 的谱线)强度增强,而低量子数转动拉曼谱线(靠近发射波长的谱线)强度减弱。拉曼激光雷达即利用大气氮气分子散射强度(高、低量子数转动拉曼谱线强度)与温度之间的依赖关系实现大气温度探测与反演。

高、低量子数拉曼散射回波信号强度可用激光雷达方程表示为

$$P_i(J_i, T, z) = C_i \cdot N(z) \cdot \sigma_i(J_i, T) \cdot \exp\left[-2 \int_0^z \alpha(z) dz\right], \quad i = 1, 2, \quad (1)$$

式中 i 表示拉曼通道标号,1、2 分别为高、低量子数转动拉曼通道, $\sigma_i(J_i, T)$ 是温度为 T 、谱线为 J_i 的拉曼散射光谱差分散射截面强度。

实际探测中单条转动拉曼散射谱线强度较弱且不易提取,因此需要同时提取相邻数条具有相同温度依赖性的转动拉曼谱线。这时,利用高、低量子数拉曼散射回波信号强度比值可以确定大气温度,即

$$H(T, z) = \frac{\sum_{J=J_1-\Delta J}^{J_1+\Delta J} P_1(T, z)}{\sum_{J=J_2-\Delta J}^{J_2+\Delta J} P_2(T, z)} \approx \exp\left\{-\left[\frac{A}{T^2(z)} + \frac{B}{T(z)} + C\right]\right\}, \quad (2)$$

式中 z 为探测高度, A 、 B 、 C 为温度标定常数,可以通过拟合求解的方法确定。

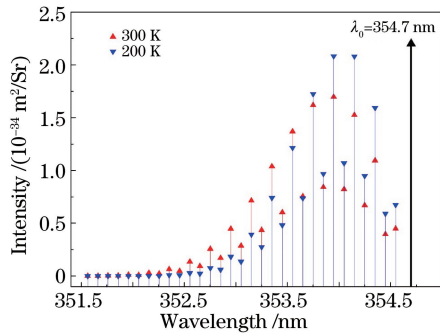


图 2 不同温度背景下氮气转动拉曼散射谱线强度
Fig. 2 Intensity of rotational Raman scattering spectra intensities of N_2 at different temperatures

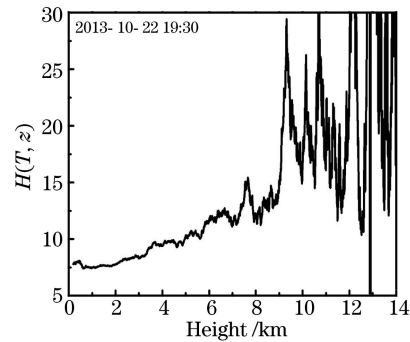


图 3 北京时间 2013 年 10 月 22 日 19 时 30 分拉曼激光雷达 $H(T, z)$ 曲线

Fig. 3 Curve of $H(T, z)$ detected by Raman lidar at 19:30 BST on October 22, 2013

通常温度反演标定采用 3 个方程求解 3 个常数 A 、 B 、 C 。之前的研究表明^[23],依据 $H(T, z)$ 曲线,取值高度层越多,求解的 A 、 B 、 C 常数代表性越好。由此提出利用 10 个方程对标定常数进行标定,进而反演温度。首先,在 $H(T, z)$ 曲线波动较小、线性分布较好的高度范围(以图 3 为例,0.2~10 km)取 10 个不同高度层,得到对应高度处的 $H(T, z)$ 值,再从美国标准大气模型曲线上获得对应高度的温度 T ,代入(2)式的变形中可得

$$\begin{cases} \ln H(T, z_1) = \frac{A}{T^2(z_1)} + \frac{B}{T(z_1)} + C \\ \ln H(T, z_2) = \frac{A}{T^2(z_2)} + \frac{B}{T(z_2)} + C \\ \vdots \\ \ln H(T, z_{10}) = \frac{A}{T^2(z_{10})} + \frac{B}{T(z_{10})} + C \end{cases}, \quad (3)$$

此时为 3 个未知数,10 个标定方程。利用最小二乘法作线性回归,即可得到具有代表性的 3 个常数。

求解得到参数 A 、 B 、 C 之后即可反演温度,反演公式为

$$T(z) = -\frac{B \pm \sqrt{B^2 - 4A [\ln H(z) + C]}}{2 [\ln H(z) + C]} \quad (4)$$

2.3 有效对等高度

以 FNL 数据评判雷达探测数据,在选取同等高度层的温度进行对比时,若选取严格对等高度,则每次探测个例仅能找出 2~3 个可以对比的高度层,由此难以全面评判测温效果。为此,根据大气温度变化特征,引入有效对等高度的概念,即在垂直高度范围内,以 FNL 高度层(含 26 个高度层)为参考高度,查找雷达测温数据在参考高度上下 3 m 范围内的高度作为参考高度的有效对等高度。

2.4 新、旧温度反演算法效果对比

图 4 所示为常规三标定方程与本研究确定的十标定方程算法对 2013 年 10 月 22 日雷达探测个例的反演效果对比。由图 4(a)可见,两种反演数据分布均与 FNL 标准数据分布保持较好的一致性。平均而言,十标定方程反演误差小于三标定方程 2 K 以上,如图 4(b)所示。

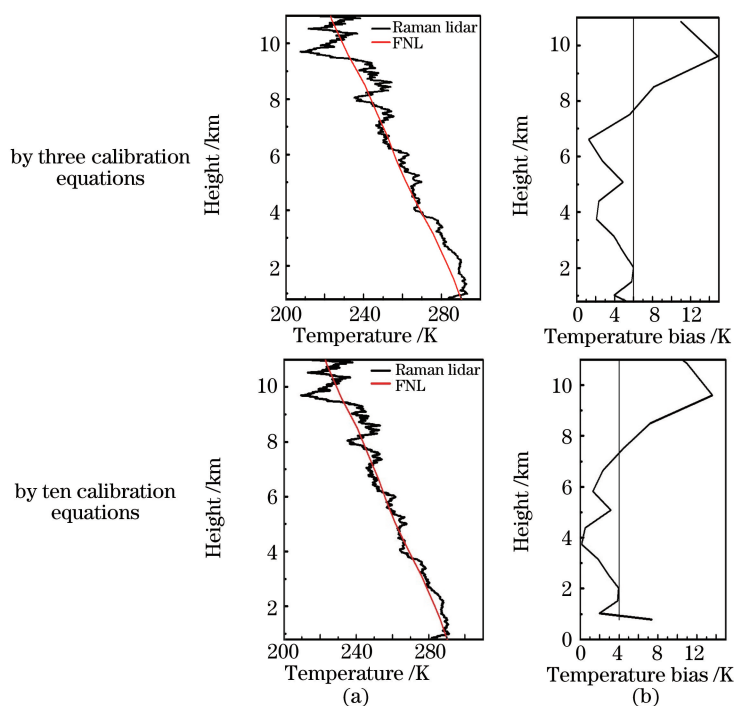


图 4 北京时间 2013 年 10 月 22 日 19 时 30 分不同标定方程反演的(a) 大气温度及(b) 误差廓线
 Fig. 4 (a) Temperature profiles and (b) error profiles inverted from different calibration equations at 19:30 BST on October 22, 2013

为了获取更具代表性的对比结果,随机选取 5 组探测数据,分别用两种算法进行温度反演。由反演效果对比可见(图 5),引入有效对等高度的概念极大地增加了垂直空间内的可对比高度层数,这 5 组探测实例的有效对等高度层数分别为 16,14,16,18,18。

图 5 表明十标定方程反演算法在刻画温度细节方面具有更大优势。在图 5(a)中,有 10 个高度层的新算法反演误差小于传统三标定方程反演误差,所占比率为所有层数的 63%。在图 5(b)中,在 14 个有效高度层中,有 12 个高度层的新算法误差小于传统算法误差,比率为 86%。图 5(c)~(e)可以更明显地揭示出,不仅绝大多数有效对等高度层上的新算法误差小于传统算法,而且新算法的反演精度明显提升,最高可提升 10 K 以上。

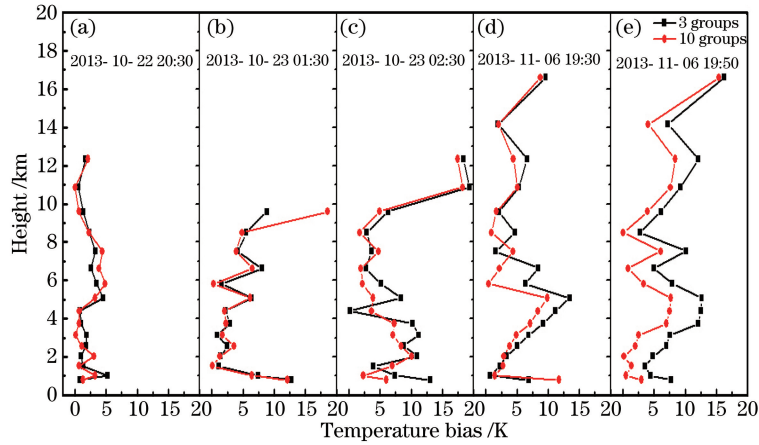


图 5 随机选取 5 组探测数据的两种温度反演算法误差

Fig. 5 Error obtained from the two temperature inversion methods with 5 groups of random selection

3 综合多级质量控制技术的建立

3.1 以定性分析作为质量控制预处理

借鉴当前数据质量检测的常规做法,绘制 FNL 数据温度廓线与雷达探测温度廓线。通过两类廓线直观对比,判断探测数据是否与标准数据趋势一致。如果两者相去甚远,则直接评判为质量不合格。

3.2 定量化质量控制技术研究

除了以常用的图形对比方法定性判断雷达探测数据是否可靠外,本研究更侧重于通过一系列统计研究与分析,筛选出适合评判被检数据的评判参数,并给出其与标准数据之间评判标准的量化值。如果以标准数据评判待测数据的精确程度,有两方面因素需要考虑:其一,两类数据相似/相近(靠近)程度;其二,两类数据相去/相差(离开)程度。统计学认为,当样本数大于 50 时,从中获取的样本规律才能具有代表性,研究所选样本为 2013 年以后的 50 组探测个例。

3.2.1 雷达探测数据与标准数据相近性分析

在海明距离、欧氏距离、相关系数、相似离度(AD)等众多相似判据中,相似离度可以较好地反映两类数据间的相似程度^[17-19]。李博等^[18]还将传统的相似离度表达式扩展为多因子多格点相似离度表达式,即

$$\begin{cases} C_{ij} = \frac{\alpha S_{ij} + \beta D_{ij}}{\alpha + \beta} \\ S_{ij} = \frac{1}{M \times N} \sum_{k=1}^M \sum_{l=1}^N |X_{ijkl} - E_{ij}| \\ D_{ij} = \frac{1}{M \times N} \sum_{k=1}^M \sum_{l=1}^N |X_{ijkl}| \\ X_{ijkl} = X_{ikl} - X_{jkl} \\ E_{ij} = \frac{1}{M \times N} \sum_{k=1}^M \sum_{l=1}^N X_{ijkl} \end{cases}, \quad (5)$$

式中 M 为格点数, N 为因子数, k, l 分别为格点和因子序号。 X 表示因子值, i, j 表示两个不同的样本,本研究中表示待测数据和标准数据,取 $N=1$ 表示单因子。 X_{ijkl} 表示 i 样本与 j 样本对应的逐格点差值, E_{ij} 表示了 i 样本与 j 样本所有格点值的总平均值, $|X_{ijkl}|$ 表示 i, j 两样本的逐个格点值的绝对距离。 D_{ij} 表示两样本所有格点值间的平均距离,反映 i, j 样本的格点值间距离在总平均值上的差异程度,称为值系数。 S_{ij} 表示两样本中各个因子和格点之间的差值 X_{ijkl} 对 E_{ij} 的离散程度,反映了 i, j 两个样本的形相似/趋势相似,称为形系数。 α, β 分别为形系数和值系数对总值的贡献,称为权重系数。

相似离度的形系数、值系数需要通过实验确定。为此,将形系数权重 α 按 0.1 步长从 1 递减至 0、值系数

权重 β 同步长从 0 增加至 1, 随机选取 10 组探测样本, 计算 11 种权重方案下的相似离度 C_{ij} (表 1)。由表 1 可见, 在 10 组随机样本中, 除 2014 年 3 月 31 日个例的相似离度无明显变化外, 其余 9 组相似离度值的变化规律都一致, 即相似离度随着 α 增大、 β 减小而减小。由此表明, 对于激光雷达测温数据, 当形系数为 1、值系数为 0 时, 相似离度普遍取值较小。相似离度越小, 表明样本之间的相似性越好。因此, 在确定雷达测温数据一级质量评判参数的形系数和值系数时, 取能使绝大多数样本相似离度达到最小情况下的权重。

表 1 11 种权重系数确定的 10 组相似离度 (样本随机选取)

Table 1 10 groups AD determined by 11 groups of weight coefficients (random sample)

α	β	AD / K									
		13/10/22	13/10/22	13/10/23	13/11/06	13/11/06	13/11/15	13/11/16	13/11/16	14/03/31	14/05/06
		19:30	20:30	02:30	19:50	20:30	20:00	19:15	20:07	20:19	20:17
1	0	3.84	1.76	5.32	3.93	3.23	3.83	3.65	4.15	4.01	4.70
0.9	0.1	3.90	1.79	5.45	3.99	3.41	3.88	3.75	4.28	4.00	4.72
0.8	0.2	3.97	1.81	5.59	4.05	3.59	3.93	3.84	4.41	4.00	4.75
0.7	0.3	4.03	1.84	5.72	4.11	3.77	3.99	3.84	4.55	4.00	4.78
0.6	0.4	4.09	1.86	5.86	4.17	3.95	4.04	4.03	4.68	4.00	4.81
0.5	0.5	4.16	1.89	6.00	4.23	4.13	4.09	4.12	4.81	3.99	4.84
0.4	0.6	4.22	1.91	6.13	4.29	4.31	4.14	4.22	4.94	3.99	4.87
0.3	0.7	4.28	1.94	6.27	4.35	4.49	4.19	4.31	5.07	3.99	4.90
0.2	0.8	4.35	1.96	6.41	4.41	4.67	4.24	4.41	5.20	3.98	4.93
0.1	0.9	4.41	1.99	6.54	4.47	4.85	4.30	4.50	5.33	3.98	4.96
0	1	4.47	2.01	6.68	4.53	5.03	4.35	4.60	5.46	3.98	4.99

拉曼雷达测温数据的特性在于对大气温度垂直廓线及空间变化趋势的描述, 而由上段确定的相似离度侧重于两类数据形相似的描述, 这恰符合激光雷达测温数据的特性, 因此可认为这种权重系数的配置是合适的。

表 2 给出了由确定的权重系数计算的 50 组探测样本的相似离度, 有 1 组相似离度在 2 K 以下, 4 组在 2~3 K 间, 21 组在 3~4 K 间, 7 组在 4~5 K 间, 9 组在 5~6 K 间, 8 组在 6~7 K 间。平均相似离度为 4.24 K, 反映了 50 组探测结果与标准数据之间相似性的平均水平, 可以作为第一级质量控制评判标准。

表 2 50 组激光雷达探测样本以及相似离度

Table 2 Lidar detection samples and corresponding AD of 50 groups

Time	13/10/22	13/10/22	13/10/23	13/10/23	13/11/06	13/11/06	13/11/06	13/11/15	13/11/15	13/11/15
AD/K	19:30	20:30	01:30	02:30	19:20	19:50	20:30	19:20	19:40	20:00
AD/K	3.84	1.76	5.40	5.32	3.65	3.93	3.23	2.59	5.04	3.83
Group	13/11/15	13/11/15	13/11/16	13/11/16	13/11/16	13/11/16	13/11/16	13/11/16	13/11/16	13/11/24
AD/K	20:20	20:40	01:35	01:51	02:37	19:00	19:15	20:07	20:22	19:34
AD/K	3.55	2.77	5.87	6.33	2.96	4.04	5.10	4.15	4.67	4.74
Time	13/11/24	13/11/24	13/11/25	13/11/25	13/11/25	13/11/26	13/11/26	13/11/26	13/11/27	13/11/28
AD/K	20:12	20:42	19:20	19:47	20:03	19:33	19:50	20:23	20:00	19:27
AD/K	3.30	3.40	6.15	6.12	6.19	4.05	3.33	3.52	3.22	5.00
Time	13/11/28	13/11/28	13/11/28	13/11/30	13/11/30	13/12/01	13/12/01	13/12/03	13/12/03	13/12/05
AD/K	19:56	20:25	20:43	19:34	20:27	19:33	20:06	19:30	20:46	19:51
AD/K	3.15	3.58	3.07	3.41	3.66	4.31	5.28	3.80	3.02	3.61
Time	13/12/05	13/12/06	13/12/09	13/12/09	13/12/10	13/12/15	13/12/28	14/03/31	14/03/31	14/05/06
AD/K	20:13	20:26	19:39	20:09	20:03	19:57	19:59	20:19	20:42	20:17
AD/K	2.48	6.37	6.48	3.62	5.81	3.00	6.77	4.01	5.01	4.70

3.2.2 两类数据相异性分析

常用的评判待测数据与标准数据相差程度的参数包括: 绝对误差 (ABER)、相对误差 (REER)、平均误差 (AVER)、均方根误差 (RMSE)。前两种参数表示的是逐层误差, 它们的极值是常见的探测精度的度量; 后两种参数表示整层误差。除了较为熟悉的 RMSE 外, ABER 是探测值与标准值之差, AVER 是 ABER 的整层平均, REER 是 ABER 与标准值的比值, 各参数的详细介绍可参见文献[24]。本节通过对这些误差参数的统计分析对比研究, 筛选出能较好地代表两类数据相异特性的误差参数作为本级评判参数。

ABER、REER 的计算结果表明,雷达反演温度极值误差可能会出现在较高层,也可能会出现在 4 km 以下的低层(表 3),但以前者居多(所占比重约 66%)。精度最差的最大误差绝对值为 25.81 K,出现在 2013 年 11 月 16 日 19 时 15 分反演个例中,高度为 12340 m,对应的探测精度为 11.66%。精度最好的最大误差绝对值为 4.87 K,出现在 2013 年 10 月 22 日探测中,高度为 5820 m,对应精度为 1.90%,而这一探测个例的低层精度也最好,数值为-1.13%。总体而言,4 km 以下的探测精度要高于全空间层的探测精度,所选 50 组探测实例的平均精度为 1.86%。

表 3 50 组探测精度及最大误差

Table 3 Detecting precisions and maximum errors of 50 groups

Time	13/10/22 19:30	13/10/22 20:30	13/10/23 01:30	13/10/23 02:30	13/11/06 19:20	13/11/06 19:50	13/11/06 20:30	13/11/15 19:20	13/11/15 19:40	13/11/15 20:00
Max error /K	-13.63	4.87	-18.55	18.28	11.76	15.44	16.85	-6.94	-21.51	15.29
Height /m	9600	5820	9600	10860	790	16610	14160	12340	3740	790
Precision /%	-5.88	1.90	-8.04	8.18	4.13	7.39	7.88	-3.13	-8.06	5.37
Low precision /%	-2.55	-1.13	-4.21	3.57	4.13	2.61	3.00	-1.88	-8.06	5.37
Time	13/11/15 20:20	13/11/15 20:40	13/11/16 01:35	13/11/16 01:51	13/11/16 02:37	13/11/16 19:00	13/11/16 19:15	13/11/16 20:07	13/11/16 20:22	13/11/24 19:34
Max error /K	9.24	8.95	31.91	13.95	10.21	15.78	25.81	11.07	25.12	-16.00
Height /m	14160	2570	12340	2030	14160	12340	12340	12340	14160	16610
Precision /%	4.35	3.26	14.35	5.03	4.79	7.13	11.66	5.00	11.81	-7.68
Low precision /%	3.17	3.26	4.61	5.03	-2.33	2.53	-2.98	3.68	2.88	4.17
Time	13/11/24 20:12	13/11/24 20:42	13/11/25 19:20	13/11/25 19:47	13/11/25 20:03	13/11/26 19:33	13/11/26 19:50	13/11/26 20:23	13/11/27 20:00	13/11/28 19:27
Max error /K	11.43	11.08	16.20	-15.79	15.31	20.30	17.79	16.33	14.69	21.08
Height /m	2570	2570	10860	9600	3140	2030	9600	12340	2570	3140
Precision /%	4.24	4.11	7.45	-6.97	5.71	7.51	7.97	7.51	5.56	7.93
Low precision /%	4.24	4.11	5.09	4.75	5.71	7.51	4.11	5.71	5.56	7.93
Time	13/11/28 19:56	13/11/28 20:25	13/11/28 20:43	13/11/30 19:34	13/11/30 20:27	13/12/01 19:33	13/12/01 20:06	13/12/03 19:30	13/12/03 20:46	13/12/05 19:51
Max error /K	14.90	15.70	16.18	12.93	12.94	-12.21	16.79	18.20	16.70	22.61
Height /m	2570	2570	14160	3140	9600	790	8500	10860	14160	16610
Precision /%	5.58	5.88	7.57	4.83	5.79	-4.32	7.24	8.26	7.77	10.71
Low precision /%	5.58	5.88	4.92	4.83	4.41	-4.32	-5.25	4.16	3.40	3.80
Time	13/12/05 20:13	13/12/06 20:26	13/12/09 19:39	13/12/09 20:09	13/12/10 20:03	13/12/15 19:57	13/12/28 19:59	14/03/31 20:19	14/03/31 20:42	14/05/06 20:17
Max error /K	12.09	12.50	-19.12	12.68	-17.78	11.68	15.50	10.47	17.25	-11.21
Height /m	12340	7520	14160	2030	12340	9600	12340	8500	7520	790
Precision /%	5.53	5.22	-8.92	4.68	-8.07	5.09	6.97	4.38	7.01	-3.83
Low precision /%	4.27	-4.00	4.00	4.68	5.23	4.21	-5.23	3.78	2.82	-3.83

Notes: Height is the altitude of the maximum error emerging. Low precision represents the precision below 4 km.

除上段逐层误差分析外,还分析了 50 组个例整层误差参数(AVER 与 RMSE),如表 4 所示。在 50 组探测样本中,AVER 取值范围在 -1.61~8.90 K 之间,RMSE 取值范围在 2.53~10.70 K 之间。由此可见,AVER 侧重于描述探测值与标准值之间离差的平均方向,其值有正有负,正值(负值)表示探测值大于(小于)标准值;而 RMSE 全部为正值,是典型的探测误差大小的量度,其数值大小可以直观反映探测结果的精确性,因此 RMSE 更有利于在本级评判中对两类数据数值大小相差程度进行评判。

求取 50 组 RMSE 的平均值(7.61 K)作为本级评判标准。图 6 所示为 50 组 RMSE 对于平均值的离散情况,红点表示 RMSE,黑色直线表示平均 RMSE,平均值以下(以上)样本个例数为 21(29),而且有较多数量样本的 RMSE 均匀分布在平均值附近。该结果表明可以用该平均值作为描述探测温度与标准温度之间相差程度的整体参考。

表 4 50 组整层误差参数
Table 4 Whole floor error parameters of 50 groups

Time	13/10/22	13/10/22	13/10/23	13/10/23	13/11/06	13/11/06	13/11/06	13/11/15	13/11/15	13/11/15
	19:30	20:30	01:30	02:30	19:20	19:50	20:30	19:20	19:40	20:00
AVER /K	1.55	1.10	-1.61	4.72	3.26	2.92	5.03	-0.01	2.98	2.40
RMSE /K	5.76	2.53	6.96	8.26	5.61	5.93	6.53	3.19	8.64	6.06
Time	13/11/15	13/11/15	13/11/16	13/11/16	13/11/16	13/11/16	13/11/16	13/11/16	13/11/16	13/11/24
	20:20	20:40	01:35	01:51	02:37	19:00	19:15	20:07	20:22	19:34
AVER /K	3.96	4.16	5.96	2.58	2.82	3.85	1.38	4.36	3.08	4.30
RMSE /K	5.75	5.28	10.70	8.02	4.79	6.63	7.87	6.49	7.52	7.81
Time	13/11/24	13/11/24	13/11/25	13/11/25	13/11/25	13/11/26	13/11/26	13/11/26	13/11/27	13/11/28
	20:12	20:42	19:20	19:47	20:03	19:33	19:50	20:23	20:00	19:27
AVER /K	6.82	5.39	4.54	4.10	4.60	8.78	7.47	8.39	8.90	8.07
RMSE /K	7.98	7.03	8.96	8.81	8.77	10.15	8.50	9.59	9.81	10.38
Time	13/11/28	13/11/28	13/11/28	13/11/30	13/11/30	13/12/01	13/12/01	13/12/03	13/12/03	13/12/05
	19:56	20:25	20:43	19:34	20:27	19:33	20:06	19:30	20:46	19:51
AVER /K	7.21	7.99	8.17	8.00	6.57	3.80	4.44	6.49	6.62	6.61
RMSE /K	8.33	9.36	9.40	8.88	7.84	6.71	8.76	8.31	7.67	8.47
Time	13/12/05	13/12/06	13/12/09	13/12/09	13/12/10	13/12/15	13/12/28	14/03/31	14/03/31	14/05/06
	20:13	20:26	19:39	20:09	20:03	19:57	19:59	20:19	20:42	20:17
AVER /K	6.56	4.56	2.83	4.90	5.98	7.94	2.49	3.57	4.82	-1.61
RMSE /K	7.18	8.97	8.51	6.57	10.08	8.74	7.12	5.64	8.07	5.85

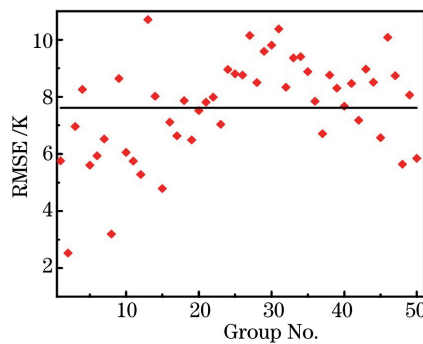


图 6 50 组探测样本 RMSE 和相对平均 RMSE 的离散程度

Fig. 6 Dispersion degree of RMSE and mean RMSE of 50 groups detection samples

需要指出的是,各级质量评判标准的评判参数并非固定值,随着探测样本数增加,其值会发生改变。总体而言,随着统计样本的增多,评判标准的代表性会不断增强。

3.2.3 综合质量评判参数

前文主要从探测数据与标准数据的空间变化趋势以及误差数值大小等不同角度进行分析。实际应用中,可根据不同研究需求有针对性地选取其中一类方法及标准开展评判。但当这两种参数给出的评判结果

不一致时,需进一步构建综合评判参数。

相似高度与 RMSE 分别考虑了不同的评判角度,因此这两类参数间具有相互独立性,如图 7(a)所示。由此构建简捷的计算综合评判参数 Q 的线性关系式

$$Q = \lambda C_{ijm} + \gamma S_{RMSE,m}, \quad m = 1, 2, \dots, 50, \quad (6)$$

式中 m 为样本数; C_{ijm} 为第 m 组样本的相似高度, $S_{RMSE,m}$ 为第 m 组样本的 RMSE; λ 和 γ 分别为相似高度和 RMSE 的权重系数, λ 与 γ 均不小于 0 且和为 1。

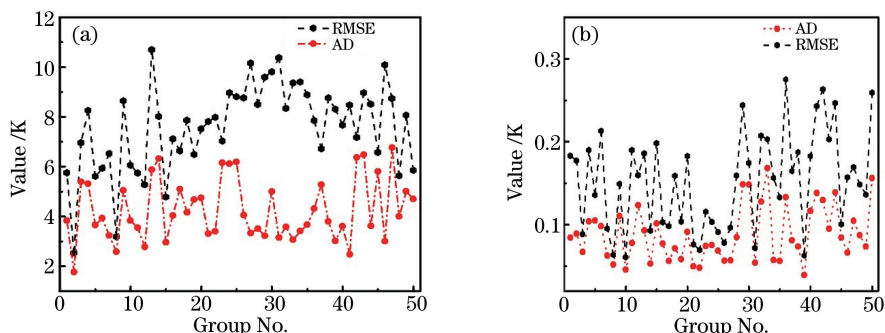


图 7 50 组探测样本标准化处理(a)前、(b)后的相似高度与 RMSE

Fig. 7 AD and RMSE (a) before and (b) after normalizing of 50 groups detection samples

由 50 组相似高度与 RMSE 分布特征[图 7(a)]可见,同组探测样本中 RMSE 值远大于相似高度,可见这两类评判参数之间并不具备可比性。因此,在研究综合参数时需引入数据标准化处理方法,即相似高度与 RMSE 在代入(6)式之前需开展标准化处理,其所用标准化方法可参见文献[18-19]。图 7(b)给出了标准化处理后的相似高度与 RMSE。从理论上讲,标准化处理后的物理参数取值范围在 0~1 之间,且为无量纲数据。由图 7(b)可见,所有样本的相似高度与 RMSE 数值均不超过 0.3,不同种类参数之间具有较好的可比性。

与相似高度权重系数确定方法类似, $\lambda(\gamma)$ 从 0(1)开始取值,以 0.1 为步长逐渐递增(递减),直到 $\lambda(\gamma)$ 取值为 1(0),分别计算 10 组样本对应的综合评判参数值,如表 5 所示。构成综合评判参数 Q 的相似高度与 RMSE 均越小越好,因此 Q 取值越小表明探测数据越可靠。由表 5 可见,随着 $\lambda(\gamma)$ 的增大(减小), Q 值逐渐减小;从细节上进一步分析可发现,当 λ 增高至高值区时, Q 减幅逐渐缓慢,甚至有些样本 Q 值不再减小(如 2013 年 11 月 15 日个例,2013 年 11 月 6 日个例,2013 年 10 月 23 日个例)。根据这些数据分布特征可以选定综合评判参数的权重系数 λ 为 0.9、 γ 为 0.1。

表 5 11 种权重系数确定的 10 组综合评判参数(样本随机选取)

Table 5 10 groups synthetical evaluating parameters chetermined by 11 groups of weight coefficients(random sample)

λ	γ	Q									
		13/10/22		13/10/23		13/11/06		13/11/15		14/03/31	
		19:30	20:30	01:30	02:30	19:50	20:30	20:00	20:07	20:19	20:17
1	0	0.08	0.08	0.07	0.10	0.09	0.06	0.05	0.06	0.08	0.15
0.9	0.1	0.09	0.09	0.07	0.11	0.10	0.07	0.05	0.07	0.09	0.16
0.8	0.2	0.10	0.10	0.07	0.12	0.11	0.07	0.05	0.08	0.10	0.17
0.7	0.3	0.11	0.11	0.07	0.13	0.13	0.07	0.05	0.09	0.10	0.18
0.6	0.4	0.12	0.12	0.08	0.14	0.14	0.08	0.05	0.10	0.11	0.19
0.5	0.5	0.13	0.13	0.08	0.14	0.15	0.08	0.05	0.11	0.12	0.21
0.4	0.6	0.14	0.14	0.08	0.15	0.16	0.08	0.05	0.12	0.12	0.22
0.3	0.7	0.15	0.15	0.08	0.16	0.18	0.09	0.06	0.13	0.13	0.23
0.2	0.8	0.16	0.16	0.08	0.17	0.19	0.09	0.06	0.13	0.14	0.24
0.1	0.9	0.17	0.17	0.09	0.18	0.20	0.09	0.06	0.15	0.14	0.25
0	1	0.18	0.18	0.09	0.19	0.21	0.10	0.06	0.16	0.15	0.26

根据构建的综合评判参数,分别求取 50 组 Q 值并计算它们的平均值(0.14)作为第三级评判标准。

3.2.4 综合多级质量控制技术

根据上述研究可构建综合多级质量控制(SMQAC)技术,如图8所示。

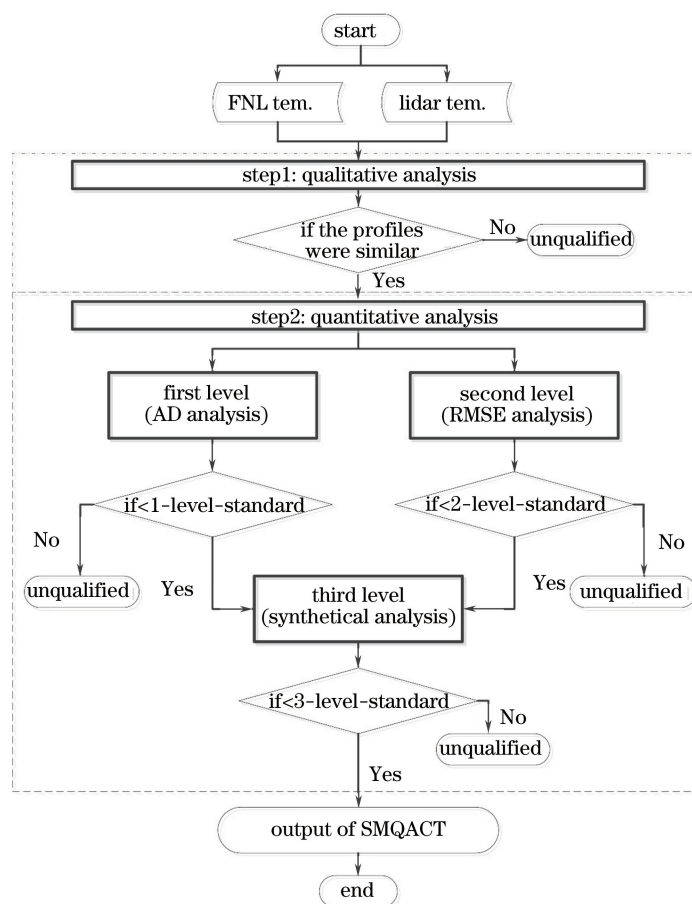


图8 综合多级质量控制技术实现流程

Fig. 8 Flow chart for SMQAC technique

该技术包含两个步骤:第一步,对激光雷达测温数据与 FNL 温度数据进行定性分析,直观比较两类数据廓线分布,若它们分布趋势明显不一致,则评判测温数据不合格;第二步,当两类廓线空间分布趋势大致相似时,评判该次探测可以通过第一步检验,并进入三级定量化检验流程。三级流程分别考虑了雷达探测数据与标准数据的相似高度评判、误差大小评判以及综合评判,每一级均确定了相应的评判标准,各级评判参数小于评判标准时,判定测温数据通过本级质量控制。综合多级质量控制技术各模块之间有明确的内在关系:第一步(依赖图形的定性化评判)相当于雷达测温是否准确的预评判,它与第二步是递进关系,只有通过第一步评判,才能进入第二步评判流程;第二步是三级定量化评判方法,其中一级、二级是并列关系,它们与第三级之间是递进关系,通过前两级的任一级评判,即可进入第三级评判,若两级都不合格,则评判为劣等数据。

选取 50 组实验样本库以外的 3 组实时探测样本对该评判方法进行初步检验。3 组样本按时间顺序分别为:2014 年 6 月 9 日 20 时 07 分(阴天背景下探测),2015 年 7 月 20 日 20 时 02 分(观测点位于对流云南侧边缘),2015 年 7 月 24 日 20 时 56 分(雷阵雨后探测)。第 1 组检验样本全部通过了综合多级质量控制流程;第 2 组检验样本全部未通过控制流程;第 3 组检验通过了定性化检验以及第一级相似高度检验,没有通过第二级、第三级检验。检验结果表明:第 2 组探测质量最差,导致这种问题的原因是降雨过程的影响,实验当日,观测点位于对流云南侧边缘,其北部正有雷阵雨过程发生;第 3 组数据在描述温度垂直趋势变化方面是成功的,但是与标准温度的误差数值太大,这组数据是雷阵雨结束以后探测的,或许是受雨后空气中大量水汽的影响。分析结果表明:一方面,激光雷达在探测无雨状况下的大气温度时,体现出较强优势;另一方面,综合多级质量控制技术可以有效评判雷达测温数据的优劣。

4 结 论

学术界已成功研制了可用于大气温、湿度探测的激光雷达系统,随之而来的问题是如何将这些探测数据有效地推广应用到气象预报等领域中。应用之前的必要环节就是确保探测数据可信。客观地讲,目前还没有一套较为系统的针对激光雷达探测数据的评判方法,这是研究者们共同面临的关键问题,探测数据的准确性无从考证,进而导致雷达探测数据可靠性评判困难。

之前设计了一套振动-转动拉曼激光雷达系统,实现了大气温等参数的实时探测^[15]。以此为基础,通过开展激光雷达常规观测实验,积累了逐日原始探测数据,并通过开展实时激光雷达实验,获取了特定天气过程下的光子/模拟数据。随后研究并设计了激光雷达大气温反演新算法,实现了利用十标定方程进行温度反演参数的标定。与传统的三标定方程反映效果对比表明,十方程标定新算法在刻画温度细节方面具有更大优势。

在这些观测实验与反演算法研究基础上,随机选取 50 组探测实例开展逐一反演,并分别与同时刻、同高度的标准温度进行细致对比。对这些样本数据进行误差诊断与统计分析,相继开展了三级质量评判参数与分析方法的研究。一级定量评判参数选为相似离度,评判标准为 4.24。二级定量评判参数选为 RMSE,评判标准为 7.61。三级为综合评判参数,评判标准为 0.14。所获取的各级质量评判标准均基于较大样本库,因此它们可以较好地反映激光雷达探测温度的平均性能。

三级定量化评判参数分别考虑了待测数据与标准数据之间的相近性特性、相异性特性以及综合特性。在具体应用中,可根据研究需求有针对性地选取某一方面进行评判,如第二级侧重的是误差大小评判,而第一级给出的是温度廓线趋势的评判,第三级给出的是雷达测温效果整体评判。

在第三级综合评判参数中,相似离度权重占到 0.9,而第一级评判所选用的相似离度表示的是两类数据形相似的描述。由此可见,构建的综合评判参数侧重于两类数据垂直空间分布趋势的相似性描述,而激光雷达探测数据的特性恰恰在于对大气参数垂直廓线分布及空间变化趋势的描述。两者的吻合可以表明,综合多级质量控制技术及其构建的综合评判参数较适合激光雷达这类探测数据的评判。

通常认为由于噪声影响,雷达探测数据在高层会产生较大偏差。但研究表明雷达探测最大误差可能会出现在高层,也可能会出现在 4 km 以下的低层,出现在低层的概率约 34%。总体而言,4 km 以下的探测精度(1.86%)高于全空间层精度。

对于未能通过综合多级质量控制技术三级评判的数据,不应急于舍弃,而应考虑是否能从中筛选出相对较好的数据段,并将其应用于后续研究中,以最大限度挖掘雷达探测数据价值。这其中又会涉及到分等级评判数据质量的科学问题,需开展专项研究。

总体而言,综合多级质量控制技术为激光雷达测温数据的评判提供了依据,也为其他类雷达探测数据(如湿度、风等)的评判提供了参考。不同种类探测数据的评判方法有所区别,本研究只是针对测温数据做出尝试,至于激光雷达探测湿度的评判标准,甚至于同时用于温、湿度等参数评判的方法,仍需引入标准化处理方法以后,才能开展更加深入的研究。

参 考 文 献

- [1] Li Bo, Zhao Sixiong. A study of heavy rainfall resulting from transversal trough and shear line-low vortex in the Huaihe River during the Meiyu onset of 2007[J]. Chinese Journal of Atmospheric Sciences, 2009, 33(6): 1148-1164.
李 博, 赵思雄. 2007 年入梅期由横槽与低涡切变引发淮河流域强降雨的诊断研究[J]. 大气科学, 2009, 33(6): 1148-1164.
- [2] Liu Hongyan. The temperature profile comparison between the ground-based microwave radiometer and the other instrument for the recent three years[J]. Acta Meteorologica Sinica, 2011, 69(4): 719-728.
刘红燕. 三年地基微波辐射计观测温度廓线的精度分析[J]. 气象学报, 2011, 69(4): 719-728.
- [3] Li Wanbiao, Wu Longtao, Zhang Chengxiang, *et al.* Retrieval of atmospheric vertical temperature profile over ocean with satellite remote sensing[J]. Acta Scientiarum Naturalium Universitatis Pekinensis, 2003, 39(5): 656-665.
李万彪, 吴龙涛, 张呈祥, 等. 气象卫星遥感探测海面大气温度垂直廓线[J]. 北京大学学报(自然科学版), 2003, 39(5): 656-665.

- [4] Hua Dengxin, Song Xiaoquan. Advances in lidar remote sensing techniques[J]. *Infrared and Laser Engineering*, 2008, 37(s3): 21-27.
华灯鑫, 宋小全. 先进激光雷达探测技术研究进展[J]. *红外与激光工程*, 2008, 37(s3): 21-27.
- [5] Behrendt A, Nakamura T, Onishi M, *et al.* Combined Raman lidar for the measurement of atmospheric temperature, water vapor, particle extinction coefficient, and particle backscatter coefficient[J]. *Applied Optics*, 2002, 41(36): 7657-7666.
- [6] Hua D X, Kobayashi T. UV Rayleigh-Mie Raman lidar for simultaneous measurement of atmospheric temperature and relative humidity profiles in the troposphere[J]. *Japanese Journal of Applied Physics*, 2005, 44(3): 1287-1291.
- [7] Balin I, Serikov I, Bobrovnikov S, *et al.* Simultaneous measurement of atmospheric temperature, humidity, and aerosol extinction and backscatter coefficients by a combined vibrational-pure-rotational Raman lidar [J]. *Applied Physics B*, 2004, 79(6): 775-782.
- [8] Leblanc T, Mc Dermid I S, Walsh T D. Ground-based water vapor Raman lidar measurements up to the upper troposphere and lower stratosphere for long-term monitoring[J]. *Atmospheric Measurement Techniques*, 2012, 5(1): 17-36.
- [9] Zhang Wei, Wu Songhua, Song Xiaoquan, *et al.* Atmospheric boundary layer detected by a Fraunhofer lidar over Qingdao suburb[J]. *Acta Optica Sinica*, 2013, 33(6): 0628002.
张 薇, 吴松华, 宋小全, 等. 夫琅禾费暗线激光雷达探测青岛市郊大气边界层[J]. *光学学报*, 2013, 33(6): 0628002.
- [10] Xie Chenbo, Zhou Jun, Yue Guming, *et al.* New mobile Raman lidar for measurement of tropospheric water vapor[J]. *Acta Optica Sinica*, 2006, 26(9): 1281-1286.
谢晨波, 周 军, 岳古明, 等. 新型车载式拉曼激光雷达测量对流层水汽[J]. *光学学报*, 2006, 26(9): 1281-1286.
- [11] Liu Dong, Tao Zongming, Wu Decheng, *et al.* Development of three-wavelength-Raman- polarization lidar system and case study[J]. *Acta Optica Sinica*, 2013, 33(2): 0228001.
刘 东, 陶宗明, 吴德成, 等. 三波长拉曼偏振激光雷达系统研制及探测个例[J]. *光学学报*, 2013, 33(2): 0228001.
- [12] Tao Z M, Liu D, Zhong Z Q, *et al.* Measurements of cirrus clouds with a three-wavelength lidar[J]. *Chinese Optics Letters*, 2012, 10(5): 050101.
- [13] Ma Xiaomin, Tao Zongming, Ma Mingjun, *et al.* Retrieval method of side-scatter lidar signal based on charge coupled device technique[J]. *Acta Optica Sinica*, 2014, 34(2): 0201001.
麻晓敏, 陶宗明, 马明俊, 等. 基于 CCD 的侧向散射激光雷达信号提取方法[J]. *光学学报*, 2014, 34(2): 0201001.
- [14] Wang Wei, Mao Feiyue, Gong Wei, *et al.* Overlap factor calculation method based on laser intensity distribution and its sensitivity analysis[J]. *Acta Optica Sinica*, 2014, 34(2): 0228005.
王 威, 毛飞跃, 龚 威, 等. 基于激光雷达强度分布的激光雷达重叠因子计算及敏感性分析[J]. *光学学报*, 2014, 34(2): 0228005.
- [15] Wang Yufeng, Gao Fei, Zhu Chengxuan, *et al.* Raman lidar for atmospheric temperature, humidity and aerosols up to troposphere height[J]. *Acta Optica Sinica*, 2015, 35(3): 0328004.
王玉峰, 高 飞, 朱承炫, 等. 对流层高度大气温度、湿度和气溶胶的拉曼激光雷达系统[J]. *光学学报*, 2015, 35(3): 0328004.
- [16] Ren Zhihua, Xiong Anyuan. Operational system development on three-step quality control of observations from AWS [J]. *Meteorological Monthly*, 2007, 33(1): 19-24.
任芝花, 熊安元. 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*, 2007, 33(1): 19-24.
- [17] Li Kaile. A new similarity parameter and its application[J]. *Acta Meteorologica Sinica*, 1986, 44(2): 174-183.
李开乐. 相似离度及其使用技术[J]. *气象学报*, 1986, 44(2): 174-183.
- [18] Li Bo, Zhao Sixiong, Lu Hancheng, *et al.* Test of the synthetical multilevel analog forecast technology in short-term rainstorm prediction[J]. *Journal of Applied Meteorological Science*, 2008, 19(3): 307-314.
李 博, 赵思雄, 陆汉城, 等. 综合多级相似预报技术在暴雨短期预报中的检验[J]. *应用气象学报*, 2008, 19(3): 307-314.
- [19] Zhang Fengqi, Cui Jing, Wang Rensheng. Application of analog deviation to recognition circulation patterns and forecasting of quantifying, timing and locating[J]. *Meteorological Monthly*, 2002, 28(9): 44-48.
张丰启, 崔 晶, 王仁胜. 相似离度在入型判别和定时、定点、定量预报中的应用[J]. *气象*, 2002, 28(9): 44-48.
- [20] Zhang Yongtao, Chen Yaodeng. Error comparison analysis between FNL data and observation data of air temperature, air pressure and ground temperature in Henan province in 2012[J]. *Meteorological and Environmental Sciences*, 2014,

37(1): 93-97.

张永涛, 陈耀登. 2012年FNL气温、气压和地面温度资料与河南省实况观测要素的误差对比分析[J]. 气象与环境科学, 2014, 37(1): 93-97.

- [21] Zhou Qing, Zhao Fengsheng, Gao Wenhua. Comparison and analysis between NCEP/NCAR every-6-hours analysis land surface and air temperature and 753 Chinese stations, observation in 2005[J]. Meteorological Monthly, 2008, 34(2): 83-91.

周 青, 赵凤生, 高文华. NCEP/NCAR 逐时分析与中国实测地表温度和地面气温对比分析[J]. 气象, 2008, 34(2): 83-91.

- [22] 全国气象基本信息标准化技术委员会. 中华人民共和国气象行业标准: 地面气象观测资料质量控制 QX/T118-2010 [S]. 北京: 气象出版社, 2010: 1-8.

- [23] Liu Jun, Hua Dengxin, Li Yan. Rotational Raman lidar for daytime-temperature profiling of the atmospheric boundary layer[J]. Acta Optica Sinica, 2002, 27(5): 755-759.

刘 君, 华灯鑫, 李 言. 大气边界层白天温度测量用转动拉曼激光雷达[J]. 光学学报, 2002, 27(5): 755-759.

- [24] Li B, Liu L P, Zhao S X. The possible mechanism of a type of vortex heavy rainfall during the pre-rainy season in South China[J]. Atmospheric and Oceanic Science Letters, 2011, 4(5): 247-252.