

基于聚类分析的双阈值激光雷达云层检测算法

陈思成, 常建华*, 刘振兴, 周妹, 孟园园, 王博业

南京信息工程大学电子与信息工程学院, 江苏 南京 210044

摘要 激光雷达探测精度高、可持续性强, 被广泛应用于云层检测。微分零交叉法是基于激光雷达回波信号进行云参数反演的常用反演方法, 但该方法在信噪比较低的情况下容易造成云层信号的误判, 影响检测精度。因此, 利用云层信号及气溶胶信号的特点, 提出了基于双阈值的云层信号检测算法, 有效排除了气溶胶信号的干扰并精确地筛选出云层信号。同时, 首次将聚类算法应用于云层结构分析, 实现了云分层的精确处理。采用大气辐射测量 (atmospheric radiation measurement, ARM) 数据进行了实验, 实验结果显示, 所提算法实现了 93.62% 的低层云检出正确率、92.78% 的中层云检出正确率和 93.03% 的高层云检出正确率, 相较于传统算法有很大提高。

关键词 激光光学; 激光雷达; 微分零交叉法; 云检测; 大气光学

中图分类号 P413

文献标志码 A

DOI: 10.3788/CJL202249.1110001

1 引言

云覆盖着全球大部分区域, 其通过影响太阳辐射及潜热释放等来参与大尺度环流和全球水循环, 在全球气候变化研究中具有重要作用^[1-3]。目前, 探测云层信息的手段主要有多传感器联合反演^[4-5]、红外测云仪探测^[6-7]、人工观测以及激光测云仪探测^[8]等。多传感器联合反演虽然探测精度高, 但其设备昂贵, 操作复杂。人工观测法在观察云状方面存在明显优势, 然而其主观性强、连续性差, 在观测精度及实时观测方面存在较大缺陷。红外测云仪通过测量红外辐射强度来反演云层信息, 可以直接得到云底高度、云顶高度甚至云量, 但其探测精度受气溶胶、水汽等各种因素的影响。激光测云仪采用激光雷达作为主要设备, 其工作时向大气发射激光脉冲, 通过对接收到的激光回波信号进行分析, 反演出云底高度、云顶高度等云层信息^[9]。

激光雷达在高时空分辨率云结构检测方面具有实时、高效、连续的特点, 已逐渐成为探测大气物理参数及云和气溶胶等时空分布的有效手段, 其中在低空大气气溶胶和云层的空间分布连续监测中获得了广泛的应用^[10]。通常情况下, 云层在对流层大气

中的密度要远高于气溶胶。然而, 当出现与气溶胶层信号相似的薄云层信号或是与云层相似的厚气溶胶层信号干扰时, 难以直接通过激光雷达回波信号提取云层信息^[11-12]。1992年, Pal等^[13]基于激光雷达数据提出微分零交叉法 (differential zero-crossing algorithm) 来反演云参数, 对拟合廓线进行微分处理, 通过微分后的信号反演云底高度、云顶高度。该算法在干扰较小的情况下精度较高, 但当场景中的背景噪声波动较大时, 该方法会产生较大的误差。2010年, 毛飞跃等^[14]在微分零交叉法的基础上结合激光雷达回波信号前后时刻的邻域判断, 有效修正了一些明显的误判。2013年, 王箫鹏等^[15]在微分零交叉法基础上, 利用不同时刻背景噪声的阈值, 实现了精度为 10 m 的云底高度反演。2019年, Cromwell等^[16]通过分割激光雷达图像确定了高时空云位置, 实现了更为高效的云识别。2020年, 陈思颖等^[17]提出了基于改进多尺度法的激光雷达云检测算法, 通过优化尺度个数阈值并添加特征峰, 有效减少了单层多峰情况下云的漏判。以上方法主要通过设定相应的阈值来提取云层的有关信息, 但阈值的设定存在主观性, 在实际应用中容易出现偏差。

为了实现更为有效的云分层, 本文首次将聚类

收稿日期: 2021-09-09; 修回日期: 2021-10-09; 录用日期: 2021-11-12

基金项目: 国家自然科学基金(61875089, 62175114)、江苏省高等学校大学生创新创业训练计划(202010300040)

通信作者: *jianhuachang@nuist.edu.cn

算法与微分零交叉法相结合,提出了一种基于聚类分析的双阈值云检测算法。该算法将改进后的云层峰与层底的回波信号之比即峰底比阈值及背景噪声阈值与微分零交叉法结合,有效避免了噪声信号及气溶胶信号引起的云层信号误判,提高了激光雷达的云层信号筛选精度。同时,根据云的垂直分布特征,结合云层信号微分零点的垂直高度与相对湿度,采用 ISODATA 算法对微分零点进行聚类分析,自适应地将检测到的多段云层信号进行整合并划分,摆脱了传统阈值法在云分层处理上的主观限制,精准地实现了云层的检测与分层。

2 方法与原理

2.1 激光雷达云探测原理

在使用激光雷达进行大气探测时,大气中的分子和粒子会产生瑞利散射和米散射,因此激光雷达方程^[18-19]可写为

$$P(r) = c_0 \frac{1}{r^2} G(r) [\beta_{\text{Ray}}(r) + \beta_{\text{Mie}}(r)] \times \exp\left\{-2 \int_0^r [\alpha_{\text{Ray}}(r') + \alpha_{\text{Mie}}(r')] dr'\right\}, \quad (1)$$

式中: $P(r)$ 为激光雷达接收到的回波信号功率; c_0 为激光雷达常数; $G(r)$ 为几何重叠因子^[20]; r 为检测目标至激光雷达的直线距离(m); $\beta_{\text{Ray}}(r)$ 与 $\beta_{\text{Mie}}(r)$ 分别为大气中粒子与分子的瑞利散射与米散射的后向散射系数; α_{Ray} 与 α_{Mie} 分别为大气中粒子和分子的瑞利散射和米散射的消光系数。由于云与气溶胶的后向散射系数不同,因此当激光在大气中传输遇到云与气溶胶时,其产生的回波信号的起伏会有所不同。图 1 所示为基于大气辐射测量(atmospheric radiation measurement, ARM)网站监测点 2021 年 4 月 19 日 12:00 至 4 月 20 日 08:00 的激光雷达测云

数据^[21]的连续时间回波信号图,可以根据垂直高度及信号强度直观地判断出不同时间段的云层信息。

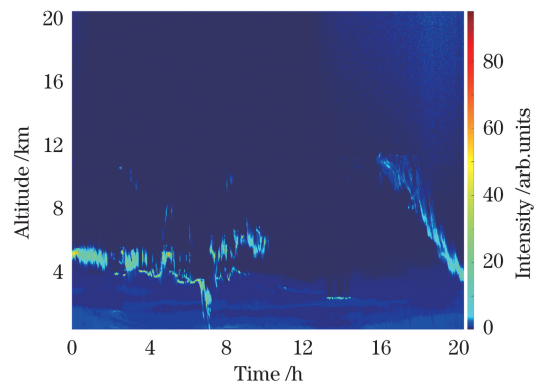


图 1 激光雷达观测的大气回波时空分布图

Fig. 1 Atmospheric echo spatiotemporal distribution pattern observed by lidar

2.2 基于双阈值的云层信号提取算法

当激光照射到云层边界时,由于云的密度远高于气溶胶的密度,大气后向散射回波信号强度 $P(r)$ 迅速增大,当激光穿过云体后, $P(r)$ 迅速减小,产生一个云层回波信号。图 2 所示为微分零交叉法的基本原理,其中, r_b 和 r_t 分别表示云层的云底高度和云顶高度, r_m 表示云层的云峰高度。由图 2 可以看出,一个云层的激光雷达回波信号对应的一阶微分信号表现为一组相邻的波峰、波谷,而云底高度 r_b 、云峰高度 r_m 、云顶高度 r_t 即为其微分信号与高度坐标轴相交的零点。使用微分零交叉法进行云参数反演对激光雷达回波信号的质量要求较高,然而激光雷达在实际使用过程中会不可避免地受到太阳光等各种因素的污染。此外,气溶胶层与云层都会引起激光雷达回波信号的变化,使用微分零交叉法极易将厚气溶胶层误判为云层或是将薄云层误判为气溶胶层。为了减小激光雷达回波信号中噪声信号带来的干扰,需要对原始回波信号进行有

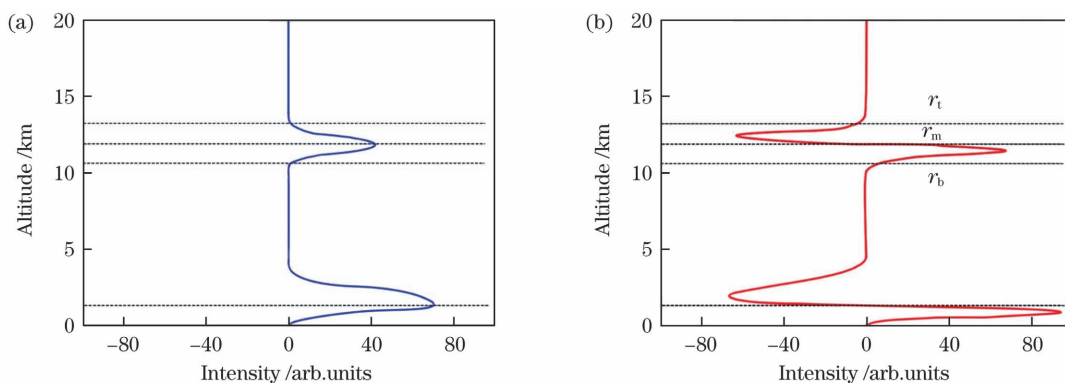


图 2 微分零交叉法基本原理。(a)激光雷达回波信号;(b)一阶微分信号

Fig. 2 Fundamental principle of differential zero-crossing method. (a) Lidar echo signal; (b) first order differential signal

效的预处理,特别是远场处的激光雷达回波信号较弱,容易引起 dP/dr 在零附近波动。Pal 等^[13]对数据进行了 5~11 个点的滑动平均处理,但拟合数据的点数过多会使信号失真,而且从众多的零点中提取出有效的零点仍然具有一定的困难。因此,本文在微分零交叉法的基础上,采用双阈值技术对回波信号进行了处理。

首先是激光雷达信号的预处理。本算法采用 LMD-ITM^[21]算法,对初始信号进行去噪处理,并对去噪后的信号进行 5 个点的滑动平均处理,预处理前与预处理后的激光雷达回波信号如图 3 所示。可以看出,预处理后的激光雷达回波信号的过零点显著减少,有效提高了激光雷达的回波信号质量。

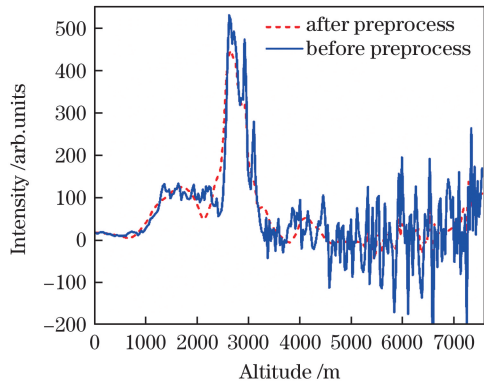


图 3 预处理前后的激光雷达回波信号对比

Fig. 3 Comparison of original and scaled lidar echo signals

然后是双阈值判断。虽然云层及气溶胶层都会引起激光雷达回波信号的变化,但是在可见光波段,云层的激光雷达回波信号强度一般明显强于气溶胶的激光雷达回波信号强度^[22],且仍存在与云层信号极为相似的厚气溶胶信号以及与气溶胶信号相似的薄云层信号。为了从干扰信号中有效选出云层信号,本文选择两种阈值对激光雷达回波信号进行处理。第一个是层峰层底比阈值 X 。经反复实验验证,本文选择由 Morille 等^[23]根据层峰层底修正信号的比值作为初阈值 X ,将云层信号与厚气溶胶信号区分开,其表达式为

$$X = \frac{P(r_m)r_m^2}{P(r_b)r_b^2} = \begin{cases} \geq 4, & \text{cloud} \\ < 4, & \text{aerosol} \end{cases}, \quad (2)$$

式中: $P(r_m)$ 、 $P(r_b)$ 分别为 r_m 、 r_b 处的激光雷达回波信号强度。

第二个是背景噪声阈值 Y 。经初阈值 X 筛选后,信号中可能存在漏判的有效云层信号,需要采用更精确的阈值将薄云层信号与气溶胶信号区分开。为了使云底/云峰的判断更加准确,引入 Pal 等^[13]提

出的背景噪声阈值进行判断。Pal 等^[13]认为,只有激光雷达回波信号中的云峰与其对应云底强度的差值大于 2 倍背景噪声强度时才能有效判断云底/云峰:

$$Y = [P(r_m) - P(r_b)] > 2I, \quad (3)$$

式中: I 为背景噪声。考虑到白天和晚上的背景噪声强度有一定的差异,在这里对 Pal 等人提出的阈值 $2I$ 进行改进。经过多次实验,确定当 $r < 5$ km 时阈值设定为 $3I$,白天当 $r \geq 5$ km 时阈值设定为 $1.5I$,晚上当 $r \geq 5$ km 时阈值取 $25I$ 。

最后是云顶高度、云底高度的计算。预处理后的信号先经过层峰层底比阈值筛选,提取明显的云层信号。为了防止漏判误判,对处理后的信号进行改进背景噪声阈值处理,选出漏选的薄云层信号。经过两次阈值处理后的信号即为待测候选云层,最后采用微分零交叉法反演计算出云顶高度、云底高度。基于双阈值的云层信号提取算法流程图如图 4 所示。

2.3 基于聚类分析的云分层算法

直观地观察激光雷达回波信号可能会在激光雷达可探测范围内的大气中发现很多层云信号,但相邻云层从相对湿度等物理特性上来看是同一层云,因此,通过聚类算法对各个云层的相对湿度及高度进行聚类分析,可以将同一高度范围内相对湿度大致相同的云层有效地归类在一起。为了在激光雷达可探测范围内实现云分层处理,本文采用 ISODATA 聚类算法对 2.2 节中所提取出的微分零点进行聚类分析。由于同一层云各处的相对湿度基本相同,ISODATA 聚类算法可以将构建的点集进行聚类,将具有相似相对湿度特征的样本聚为一类,即视作同一层云。Wang 等^[24]利用激光雷达数据反演云层信息,认为两层云的垂直距离至少为 0.05 km。因此,为了观测结果的精确度,本文将每两层云之间的距离阈值设定为 0.05 km,若不满足条件则视为同一层云。基于 ISODATA 聚类算法的云分层过程如下。

第一步构建云层零点数据集。根据每个零点处激光雷达所探测到的相对湿度与垂直高度构建出一个二维数据集 $x = \{(r_i, b_i), i = 1, 2, \dots, N\}$,其中 r_i 为垂直高度, b_i 为相对湿度, i 为模式样本的编号, N 为模式样本的总个数。预选 N_c 个零点初始聚类中心 $\{z_1, z_2, \dots, z_{N_c}\}$,随机选取初始位置。预选 $K = 3$ 作为预期的聚类中心数目,设置每一聚类域零点集中最少零点数为 $\theta_N = 2$ 。两个聚类中心零点间的最小距离 $\theta_c = 50$,若两个聚类中心零点间的距离小于 θ_c ,两个聚类需进行合并,在一次迭代运

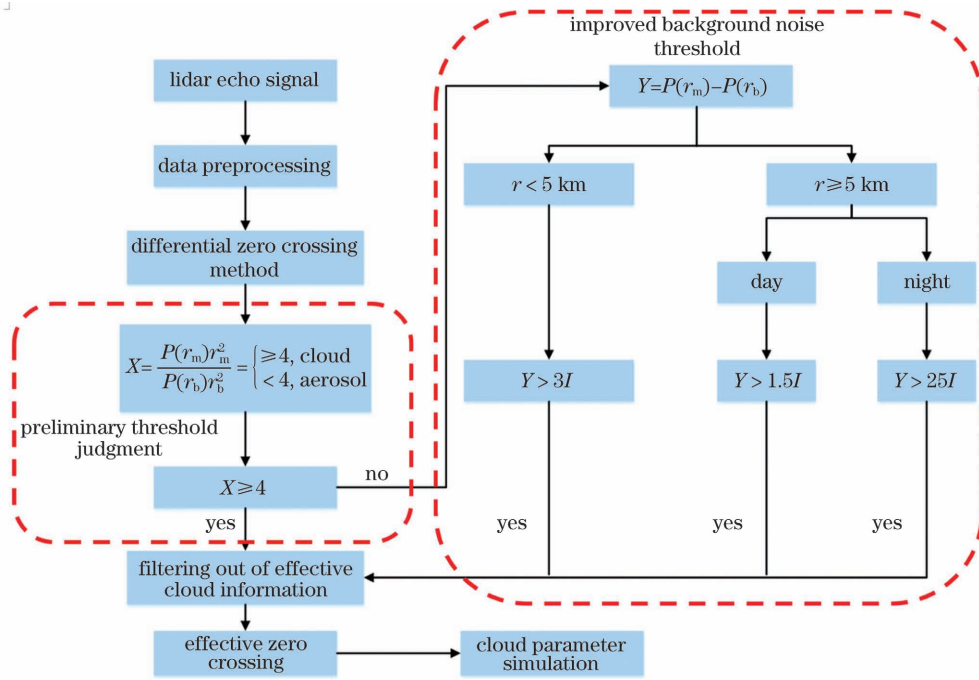


图 4 改进的微分零交叉法流程图

Fig. 4 Flow chart of improved differential zero-crossing algorithm

算中最多可以合并的聚类中心为 2; 设置迭代次数 $M=30$ 。将 N 个模式样本分给最近的聚类 S_j , 若

$$D_j = \min\{|x_i - z_j|, j = 1, 2, \dots, N_c\}, \quad (4)$$

即 $|x_i - z_j|$ 最小 (x_i 为 x 中的元素), 则 $x_i \in S_j$ 。如果 S_j 中的样本数目 $< \theta_N$, 则取消该样本子集, 此时 N_c 减去 1。

修正各聚类中心零点, 有

$$z_j = \frac{1}{N_j} \sum_{x \in S_j} x, j = 1, 2, \dots, N_c, \quad (5)$$

式中: N_j 表示第 j 个类别的样本数

计算各聚类零点域 S_j 中模式样本与各聚类中心间的平均距离:

$$\bar{D}_j = \frac{1}{N_j} \sum_{x \in S_j} |x - z_j|, j = 1, 2, \dots, N_c. \quad (6)$$

计算出全部样本零点和其对应聚类中心零点的总平均距离:

$$\bar{D} = \frac{1}{N_c} \sum_{j=1}^{N_c} \bar{D}_j. \quad (7)$$

第二步是云层零点集的判别分类。该部分主要包括分裂和合并, 在迭代判别的过程中, 如果 $N_c \leq K/2$, 即聚类中心的数目小于或等于规定值的一半, 则对已有聚类进行分裂处理; 如果迭代运算的次数是偶数或 $N_c \geq 2K$, 则进行合并处理, 直至达到最大迭代次数。

在分裂处理中, 首先计算每个聚类零点集中样本距离的标准差:

$$\sigma_j = (\sigma_{1j}, \sigma_{2j}, \dots, \sigma_{nj})^T, \quad (8)$$

式中: 向量的各个分量为

$$\sigma_{aj} = \sqrt{\frac{1}{N_j} \sum_{k=1}^{N_j} (x_{ak} - z_{aj})^2}, \quad (9)$$

式中: 样本特征向量的维数 $i=1, 2, \dots, n$, 其中 n 为最大样本特征维数; k 表示预选的第 k 个聚类中心。

其次是求每一标准差向量 $\{\sigma_j, j=1, 2, \dots, N_c\}$ 中的最大分量, 以 $\{\sigma_{j_{\max}}, j=1, 2, \dots, N_c\}$ 为例, 在任一最大分量集 $\{\sigma_{j_{\max}}, j=1, 2, \dots, N_c\}$ 中, 将一个聚类域中样本距离分布的标准差定义为 θ_s , 若有 $\sigma_{j_{\max}} > \theta_s$, 且满足 $\bar{D}_j > \bar{D}, N_j > \theta_N + 1$ 或 $N_c < \frac{K}{2}$, 则将 z_j 分裂为两个新的零点聚类中心, N_c 加 1。

在合并处理中, 首先计算全部聚类中心的距离:

$$D_{ij} = |z_i - z_j|, i = 1, 2, \dots, N_c - 1 \text{ and } j = i + 1, \dots, N_c. \quad (10)$$

其次是比较 D_{ij} 与 θ_c 的值。将小于 θ_c 的 D_{ij} 值按最小距离次序递增排列, 即 $\{D_{i_1j_1}, D_{i_2j_2}, \dots, D_{i_Lj_L}\}$, 其中 $D_{i_1j_1} < D_{i_2j_2} < \dots < D_{i_Lj_L}$ 。

最后将距离为 $D_{i_kj_k}$ 的两个零点聚类中心 z_{i_k} 和 z_{j_k} 合并, 得新的中心为:

$$z_k^* = \frac{1}{N_{i_k} + N_{j_k}} (N_{i_k} z_{i_k} + N_{j_k} z_{j_k}), \quad k = 1, 2, \dots, L, \quad (11)$$

式中: L 为一次合并中可以合并类别的最多对数。

被合并的两个零点聚类中心分别以其聚类域内的样本数加权,使 z_i^* 成为真正的平均量。

第三步是云层结构判别。根据分类后的微分零点集进行云层结构的判别。在这一步运算中,迭代次数 M 加 1。至此,云层微分零点聚类完成。以 ARM 网站监测站点 Southern Great Plains (SGP) 2021 年 4 月 17 日 11:30 的第 620 根激光

雷达回波信号廓线为例,微分零点聚类结果如图 5 所示,该层云的云底高度和云顶高度为该零点集内改进微分零交叉法计算出的云底高度最小值和云顶高度最大值。可以看出,在 0~9 km 的范围内存在两层云,两层云的云底高度 r_{b1} 、 r_{b2} 分别为 3.24 km、7.78 km,云顶高度 r_{t1} 、 r_{t2} 分别为 6.09 km、8.79 km。

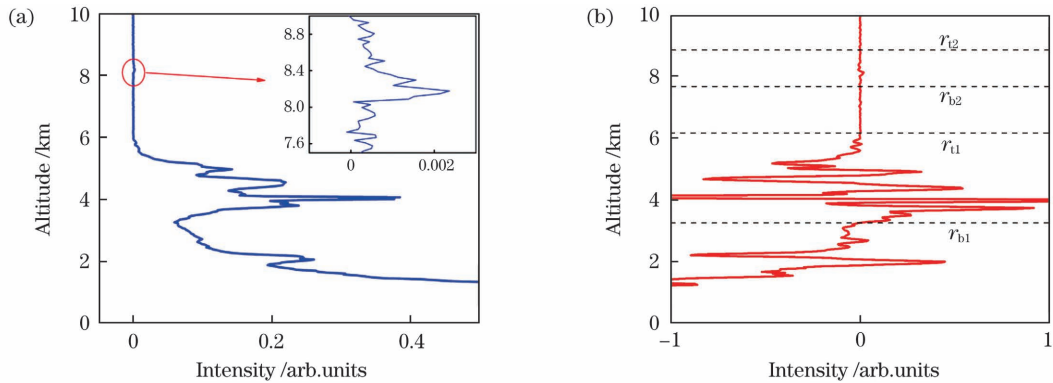


图 5 零点聚类结果。(a)激光雷达回波信号;(b)检测结果

Fig. 5 Zero-point clustering results. (a) Lidar echo signal; (b) cloud detection results

3 实验结果与讨论

为了验证本算法的适用性,本文分别进行了模

拟仿真实验和真实数据测试。模拟仿真实验结果如图 6 所示,其中图 6(a)、(c)为模拟的低云、高云信号,图6(b)、(d)为相应的反演结果。模拟的低云云

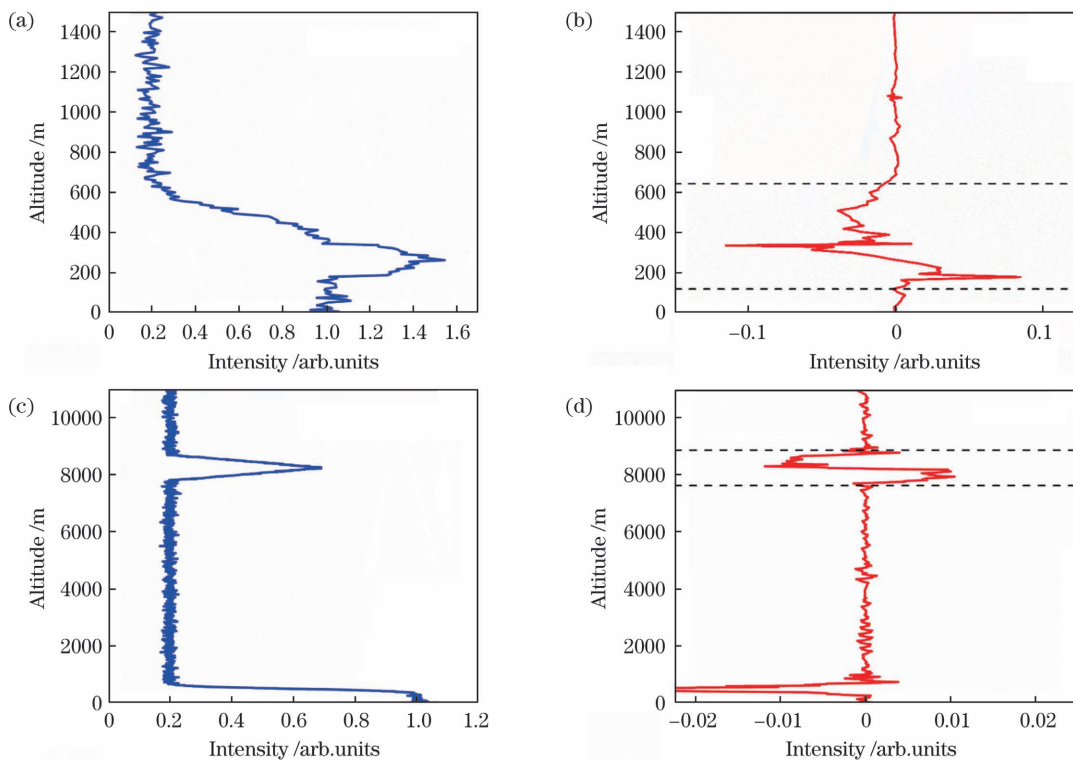


图 6 模拟的激光雷达回波信号及对应的算法处理结果。(a)模拟的低云激光雷达回波信号;(b)低云信号的算法处理结果;(c)模拟的高云激光雷达回波信号;(d)高云信号的算法处理结果

Fig. 6 Simulated lidar echo signal and corresponding algorithm processing results. (a) Simulated low-cloud lidar echo signal; (b) algorithm processing result of low-cloud signal; (c) simulated high cloud lidar echo signal; (d) algorithm processing result of high-cloud signal

底高度、云顶高度分别为 120 m、644 m, 模拟的高云云底高度、云顶高度分别为 7654 m、8823 m。根据实验结果可见, 针对不同高度的云层, 改进后的微分零交叉法都能有效地提取云层信号, 能够很好地确定云底高度及云顶高度。

采用本文所提算法对实测激光雷达回波数据廓线进行了反演。以 ARM 网站监测站点 SGP 的激光雷达实测数据为例, 选取监测时间为 2021 年 4 月 19 日 11:30 的激光雷达回波信号廓线。从图 1 的

回波信号时空分布图中可以看出, 在激光雷达可探测范围内, 2.5~4.5 km 范围以及 4.0~6.5 km 范围内具有明显云层信号; 对 7.0~10.0 km 范围内的信号进行放大后可以发现, 在 8~10 km 区间具有夹杂在噪声信号中的微弱云层信号。激光雷达回波信号图及局部放大信号如图 7 所示, 使用本文提出的改进算法对信号进行处理, 处理后的结果及局部放大如图 8 所示。

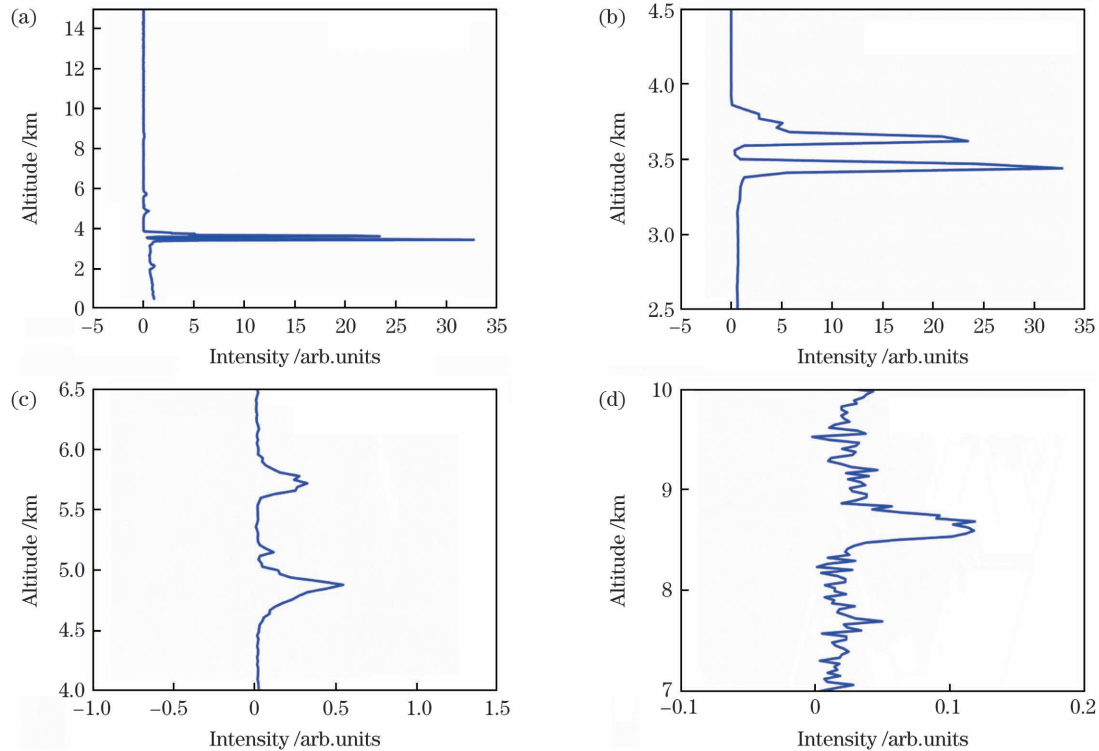


图 7 激光雷达回波信号及云层局部放大信号。(a)激光雷达实测的回波信号;(b) 2.5~4.5 km 区间的放大信号;
(c) 4.0~6.5 km 区间的放大信号;(d) 7.0~10.0 km 区间的放大信号

Fig. 7 Lidar echo signal and partial cloud layer amplified signals. (a) Measured lidar echo signal; (b) amplified signal in 2.5-4.5 km; (c) amplified signal in 4.0-6.5 km; (d) amplified signal in 7.0-10.0 km

图 8 中第一层云的云底高度和云顶高度分别为 3.12 km 和 3.94 km, 第二层云的云底高度和云顶高度分别为 4.47 km 和 6.13 km, 第三层云的云底高度和云顶高度分别为 8.48 km 和 9.50 km。对于

同一站点同一时刻的回波数据, 利用传统微分零交叉法进行处理, 得出了云底高度和云顶高度数据。表 1 将该站点该时刻的标准值与传统微分零交叉法及本文提出的改进算法得出的云层参数数据进行了对比。

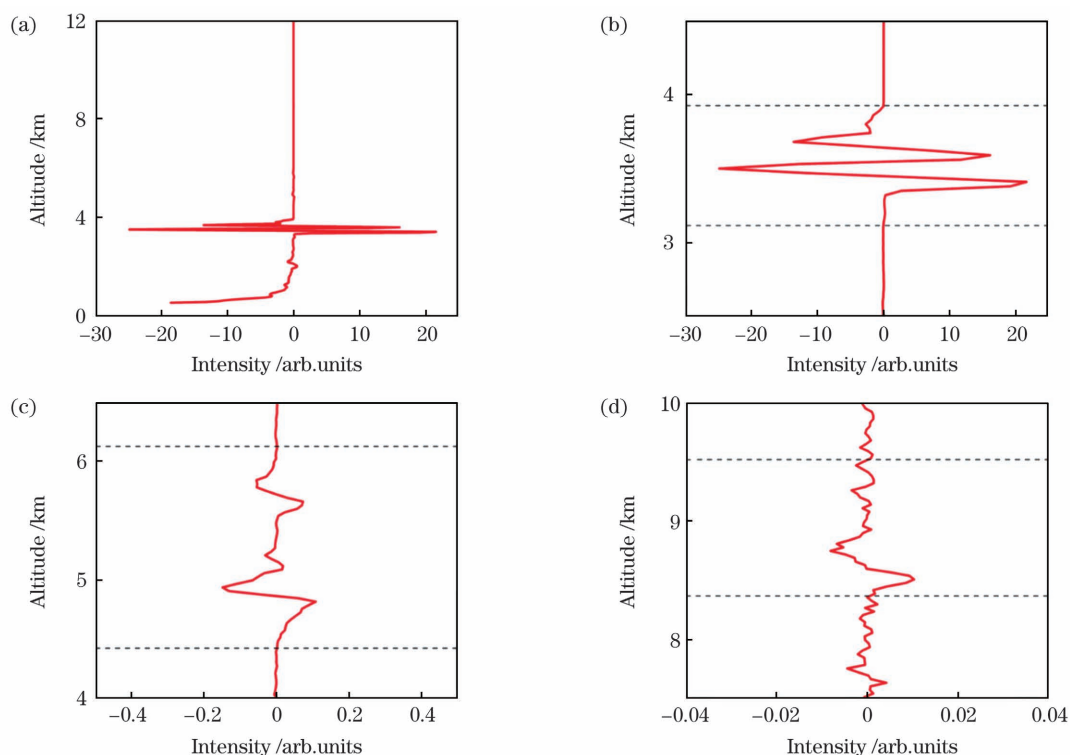


图 8 改进算法的反演结果及对应的云层局部信号放大。(a)激光雷达回波信号的算法处理结果;(b)第一层云层;(c)第二层云层;(d)第三层云层

Fig. 8 Inversion results of improved algorithm and partial cloud layer amplified signals. (a) Algorithm processing result of lidar echo signal; (b) first cloud layer; (c) second cloud layer; (d) third cloud layer

表 1 不同反演算法得出的云层数据及其标准值

Table 1 Cloud layer data obtained by different inversion algorithms and standard values

Method	Cloud layer	Cloud base height /km	Cloud top height /km	Cloud thickness /km
Standard value	Cloud layer 1	3.14	3.98	0.84
	Cloud layer 2	4.51	6.10	1.59
	Cloud layer 3	8.47	9.47	1.0
Improved algorithm	Cloud layer 1	3.12	3.94	0.82
	Cloud layer 2	4.47	6.13	1.66
	Cloud layer 3	8.48	9.50	1.02
Traditional method	Cloud layer 1	3.02	4.03	1.0
	Cloud layer 2	4.38	6.31	1.93
	Cloud layer 3	-	-	-

根据实验反演得到的云层高度数据可以看出,相比传统微分零交叉法,改进后的微分零交叉法在云层的云底高度、云顶高度的测量上具有更好的精确度。相较而言,本文提出的改进微分零交叉法在第一层云的反演中避免了大面积的误判,而传统的微分算法不能很好区分气溶胶和云层,误将部分气溶胶误认为云,因此对第一云层的云底云高的反演存在一定的误差,云层厚度反演值偏高。此外,由于第三层云的信号过于微弱,传统算法难以反演出这一层云的云层信息,而改进后的微分零交叉法通过

双阈值的设定,很好地解决了这一问题,精确地捕捉到了云层信息。

为了进一步验证本算法的优越性,选取 ARM 网站监测站点 SGP 在 2021 年 4 月 19 日至 4 月 20 日的激光雷达实测数据进行分析,激光雷达回波信号包括 2021 年 4 月 19 日 12:00 至 2021 年 4 月 20 日 08:00 共 20 个小时的时间跨度,共测得 667 组信号廓线。微分零交叉法及本算法的检测结果分别如图 9(a)、(b)所示。在图 9(a)中 A 区域,可以发现断断续续的微弱云层信号,然而连续性本就是云层

变化的重要特性之一,通过肉眼对比图 9(a)与图 1,可以发现,A 块区域出现的断点是误判造成的。相比图 9(b)中的 B 区域,图 9(a)中出现了漏选现象,这是因为原始算法的云层信号 dP/dr 大于零的点数达不到阈值的要求,而改进后的算法通过结合层峰层底比阈值及背景噪声阈值,正确地将该区域中的云层信息检测出来。相比图 9(b)中的 C 区域,图 9(a)出现了漏选、误判,这是由于原始算法精度较差,连续的云层信息难以被精确识别,因此云层的检出率低且连续性差。从图 9 可知,在 02:00—06:00 这段时间内,两种算法的云层检测结果保持着较好的一致性。通过对比图 9 与图 1,可以发现,改进算法的结果与图 1 的结果更一致,能够更加可靠地反演激光雷达探测的云层信息。为了验证算法的云层检测性能,本文按照垂直高度将云层分为低层云

($r \leq 2$ km)、中层云 ($2 \text{ km} < r \leq 7$ km) 和高层云 ($r > 7$ km),算法反演得到的云底高度、云顶高度均与 ARM 数据标准值保持在 2 个激光雷达分辨率(本文为 0.06 km)内即视为云层反演正确。以低层云的检测正确率为例,统计出所有含有低层云的廓线数量,若廓线内所有低层云均反演正确则视为有效反演低层云廓线,有效反演低层云廓线数量与低层云廓线总数的比值即为低层云的检测正确率,中层云和高层云的计算方法同理。本文采用 ARM 网站监测站点 2021 年 4 月 17 日至 4 月 30 日共计 33068 组廓线进行实验,并与该站点 ARM 数据的标准值作对比以计算检测正确率。实验结果表明,本文所提改进算法实现了低层云 93.62% 的检测正确率、中层云 92.78% 的检测正确率和高层云 93.03% 的检测正确率。

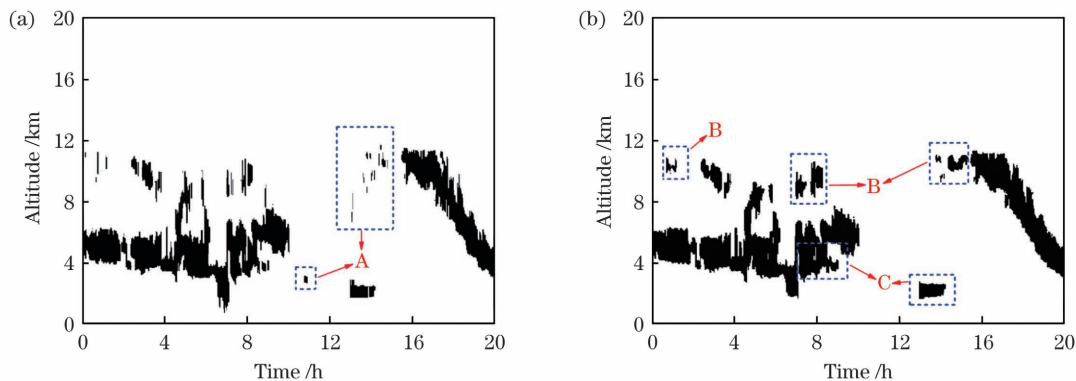


图 9 微分零交叉法与改进算法的检测结果。(a)微分零交叉法;(b)改进算法

Fig. 9 Detection results of differential zero crossing method and improved algorithm; (a) Differential zero-crossing method; (b) improved algorithm

4 结 论

针对传统微分零交叉法存在的缺陷,提出了一种基于聚类分析的双阈值云层检测算法。该算法采用改进后的层峰层底比阈值与背景噪声阈值处理激光雷达信号,精确提取了云层信息。此外,针对云的分层处理,采用 ISODATA 算法对微分零点进行聚类分析,实现了精确的云分层处理。实验结果表明,所提算法相较于传统的微分零交叉法,反演精度有明显提升,可以有效排除气溶胶信号的干扰并提取出云层信息,且对不同高度的云层都有较好的检测效果。所提方法可以进一步探究云层类型和相态的基础,通过结合不同云层内部的物理特性及不同云层结构的回波信号特征,实现对云层类型(如积云、层云和卷云)的区分以及对云相态(如冰云、水云和冰水混合云等)的区分。

参 考 文 献

- [1] Liou K N. An introduction to atmospheric radiation [M]. San Diego: Academic Press, 2002.
- [2] Rossow W B, Schiffer R A. Advances in understanding clouds from ISCCP[J]. Bulletin of the American Meteorological Society, 1999, 80(11): 2261-2287.
- [3] Bourgeois Q, Ekman A M L, Igel M R, et al. Ubiquity and impact of thin mid-level clouds in the tropics[J]. Nature Communications, 2016, 7: 12432.
- [4] Roh W, Satoh M. Extension of a multisensor satellite radiance-based evaluation for cloud system resolving models[J]. Journal of the Meteorological Society of Japan Ser II, 2018, 96(1): 55-63.
- [5] 王健, 崔天翔, 王一, 等. 高分五号可见短波红外高光光谱影像云检测研究[J]. 光学学报, 2021, 41(9): 0928003.
Wang J, Cui T X, Wang Y, et al. Cloud detection

- for GF-5 visible-shortwave infrared advanced hyperspectral image [J]. *Acta Optica Sinica*, 2021, 41(9): 0928003.
- [6] Wang Z E, Sassen K. Cloud type and macrophysical property retrieval using multiple remote sensors [J]. *Journal of Applied Meteorology*, 2001, 40 (10): 1665-1682.
- [7] 莫伟强, 尹淑娴. 激光雷达测云高与人工观测的对比分析 [C]//第 32 届中国气象学会年会 S16 地基遥感观测技术与应用, 北京: 中国气象学会, 2015: 63-70.
Mo W Q, Yin S X. A comparative analysis of cloud height measurement by lidar and artificial observation [C]//The 32nd annual meeting of Chinese Meteorological Society S16 Ground-based Remote Sensing Observation Technology and Application, Beijing: Chinese Meteorological Society, 2015: 63-70.
- [8] Voudouri K A, Giannakaki E, Komppula M, et al. Variability in cirrus cloud properties using a Polly^{XT} Raman lidar over high and tropical latitudes [J]. *Atmospheric Chemistry and Physics*, 2020, 20(7): 4427-4444.
- [9] Zhao C F, Wang Y Z, Wang Q Q, et al. A new cloud and aerosol layer detection method based on micropulse lidar measurements [J]. *Journal of Geophysical Research: Atmospheres*, 2014, 119 (11): 6788-6802.
- [10] Yang Y K, Zhao C F, Wang Q, et al. Aerosol characteristics at the three poles of the Earth as characterized by cloud-aerosol lidar and infrared pathfinder satellite observations [J]. *Atmospheric Chemistry and Physics*, 2021, 21(6): 4849-4868.
- [11] 于思琪, 刘东, 徐继伟, 等. 激光雷达反演大气边界层高度的优化方法 [J]. *光学学报*, 2021, 41(7): 0728002.
Yu S Q, Liu D, Xu J W, et al. Optimization method for planetary boundary layer height retrieval by lidar [J]. *Acta Optica Sinica*, 2021, 41(7): 0728002.
- [12] 魏婷婷, 杨家志, 周国清, 等. 基于改进 EWT 的浅水激光雷达回波信号降噪 [J]. *中国激光*, 2021, 48 (11): 1109001.
Wei T T, Yang J Z, Zhou G Q, et al. Shallow-water LiDAR echo signal denoising based on improved EWT [J]. *Chinese Journal of Lasers*, 2021, 48(11): 1109001.
- [13] Pal S R, Steinbrecht W, Carswell A I. Automated method for lidar determination of cloud-base height and vertical extent [J]. *Applied Optics*, 1992, 31 (10): 1488-1494.
- [14] 毛飞跃, 龚威, 李俊, 等. 基于改进微分零交叉法的米氏散射激光雷达云检测与参数反演 [J]. *光学学报*, 2010, 30(11): 3097-3102.
Mao F Y, Gong W, Li J, et al. Cloud detection and parameter retrieval based on improved differential zero-crossing method for Mie lidar [J]. *Acta Optica Sinica*, 2010, 30(11): 3097-3102.
- [15] 王箫鹏, 宋小全, 闫宝东, 等. 激光雷达观测珠海市云底高度研究 [J]. *光电子·激光*, 2013, 24 (11): 2192-2197.
Wang X P, Song X Q, Yan B D, et al. Research on the observation of cloud-base height for the city of Zhuhai of China with lidar [J]. *Journal of Optoelectronics·Laser*, 2013, 24(11): 2192-2197.
- [16] Cromwell E, Flynn D. Lidar cloud detection with fully convolutional networks [C]//Optical Sensors and Sensing Congress (ES, FTS, HISE, Sensors), June 25-27, 2019, San Jose, California. Washington, D.C.: OSA, 2019: EW6A.4.
- [17] 陈思颖, 王嘉奇, 陈和, 等. 改进简单多尺度法的激光雷达云检测 [J]. *红外与激光工程*, 2020, 49(S2): 20200379.
Chen S Y, Wang J Q, Chen H, et al. Lidar cloud detection based on improved simple multiscale method [J]. *Infrared and Laser Engineering*, 2020, 49(S2): 20200379.
- [18] Gong W, Mao F Y, Song S L. Signal simplification and cloud detection with an improved Douglas-Peucker algorithm for single-channel lidar [J]. *Meteorology and Atmospheric Physics*, 2011, 113(1/2): 89-97.
- [19] Kovalev V A, Eichinger W E. Elastic lidar: theory, practice, and analysis methods [M]. Hoboken: John Wiley & Sons, Inc., 2004.
- [20] Fernald F G. Analysis of atmospheric lidar observations: some comments [J]. *Applied Optics*, 1984, 23(5): 652-653.
- [21] Zhang L Y, Chang J H, Li H X, et al. Noise reduction of LiDAR signal via local mean decomposition combined with improved thresholding method [J]. *IEEE Access*, 2020, 8: 113943-113952.
- [22] Kovalev V A. Solutions in lidar profiling of the atmosphere [M]. Hoboken: John Wiley & Sons, Inc., 2015.
- [23] Morille Y, Haeffelin M, Drobinski P, et al. STRAT: an automated algorithm to retrieve the vertical structure of the atmosphere from single-channel lidar data [J]. *Journal of Atmospheric and Oceanic Technology*, 2007, 24(5): 761-775.
- [24] Wang Z E, Sassen K. Cirrus cloud microphysical property retrieval using lidar and radar measurements. part II: midlatitude cirrus microphysical and radiative properties [J]. *Journal of the Atmospheric Sciences*, 2002, 59(14): 2291-2302.

Double-Threshold Lidar Cloud Layer Detection Algorithm Based on Cluster Analysis

Chen Sicheng, Chang Jianhua*, Liu Zhenxing, Zhou Mei, Meng Yuanyuan, Wang Boye
*School of Electronic and Information Engineering, Nanjing University of Information Science & Technology,
Nanjing 210044, Jiangsu, China*

Abstract

Objective Clouds cover most of the globe, which plays an important role in the global climate change research by influencing solar radiation and latent heat release and participating in the large-scale circulation and global water cycle. The traditional cloud detection methods mainly include sensor joint inversion, infrared cloud detection, artificial observation, and laser cloud detection, but all of them have some problems, such as strong subjectivity, poor continuity, and too many interference factors. Because of its real-time, efficient and continuous characteristics in cloud structure detection with high spatio-temporal resolution, a lidar has gradually become an effective means to detect the spatio-temporal distributions of atmospheric physical parameters, clouds and aerosols and been widely used in the continuous monitoring of the spatial distributions of atmospheric aerosols and clouds at low and middle altitudes. Normally, clouds are much denser in the tropospheric atmosphere than aerosols. However, it is difficult to extract cloud information directly from lidar echo signals when there is interference from thin cloud signals similar to aerosol signals or thick aerosol signals similar to cloud signals. In this paper, a double-threshold lidar cloud detection algorithm based on cluster analysis is proposed, which can extract cloud information from lidar echo signals and realize layered cloud processing without interference signals.

Methods Based on the differential zero-crossing method combined with threshold and cluster analysis, the cloud signal detection and cloud stratification are realized in this study. First, the IMD-ITM algorithm is used to preprocess the signals to improve the signal-to-noise ratio of the original echo signals. Then, the cloud signal is screened from the preprocessed signal by the ratio of cloud layer peak echo signal to layer bottom echo signal (peak-to-bottom ratio) threshold, and the improved background noise threshold is combined to prevent the thin cloud signal from missing selection. The effective cloud signal obtained by threshold filtering is processed by the differential zero crossing method, and the differential zero of the cloud signal is obtained. Finally, we built up a two-dimensional array by vertical height and relative humidity of cloud signals at differential zeros, and the ISODATA algorithm is used to the cluster analysis of the array. The cloud signals at differential zeros in each array after clustering are in the adjacent height range and have roughly the same relative humidity, which indicates that they affiliate to the same cloud layer. The cloud base and cloud top heights of the stratus cloud are the minimum cloud base height and the maximum cloud top height calculated by the improved differential zero crossing method in the zero set. In order to verify the reliability of the algorithm, the lidar echo signal is simulated to verify the applicability of the proposed algorithm to clouds at different heights, and the continuity and cloud detection rate are compared with those of the traditional differential zero crossing method and the standard values.

Results and Discussions The improved cloud detection algorithm can effectively detect cloud signals and realize cloud structure discrimination. Experiments show that the algorithm can accurately extract cloud signals and determine the heights of cloud base and cloud top for both low cloud and high cloud (Fig. 6). The lidar echo signal profiles are shown in Fig. 7. By comparing with those of the traditional differential zero-crossing method, it can be found that the continuous time signal inversion results of the improved algorithm are significantly better (Fig. 8 and Table 1). It has obvious advantages in eliminating the interference of aerosol signals and extracted thin cloud signals (Fig. 9). Experimental results show that compared with the ARM cloud data, the cloud detection accuracy of the improved algorithm proposed in this paper is 93.62% for the low cloud layer, 92.78% for the middle layer cloud, and 93.03% for the high cloud layer.

Conclusions Aiming at the defects of the traditional differential zero-crossing method, this paper proposes a two-threshold cloud detection algorithm based on cluster analysis. In this algorithm, the lidar signal is processed with the improved thresholds of peak-to-bottom ratio and background noise, and the cloud information is extracted accurately.

In addition, for the layered cloud processing, this paper adopts the ISODATA algorithm for the cluster analysis of differential zeros to achieve accurate layered cloud processing. Experimental results show that compared with the traditional differential zero-crossing method, the proposed algorithm significantly improves the inversion accuracy, effectively eliminates the interference of aerosol signals, extracts cloud information, and has a good detection effect on the clouds with different heights. It is found that on the basis of hierarchical cloud processing, the echo signal characteristics with different types of clouds and different cloud structures can be used to distinguish cloud types such as cumulus, stratus, and cirrus, as well as cloud phase states such as ice cloud, water cloud, and ice-water mixed cloud.

Key words laser optics; lidar; differential zero-crossing method; cloud detection; atmospheric optics