

ZHANG Xiaojian, ZHANG Xianzhong, LI Xinqi, WU Tong, ZHONG Kai, YAN Zhaoai, XU Degang, YAO Jianquan. Uncertainty Evaluation of Atmospheric Temperature Retrieval Using Rayleigh Lidar Based on the Optimal Estimation Method (in Chinese). *Chinese Journal of Space Science*, 2025, 45(4): 995–1006. CSTR:32142.14.cjss.2024-0081. DOI:10.11728/cjss2025.04.2024-0081

基于最优估计法的瑞利激光雷达反演 大气温度不确定度*

张潇剑^{1,2} 张献中^{1,2} 李欣岐^{1,2} 吴同^{1,2}
钟凯^{1,2} 闫召爱^{3,4} 徐德刚^{1,2} 姚建铨^{1,2}

1(天津大学精密仪器与光电子工程学院 天津 300072)

2(天津大学光电信息技术教育部重点实验室 天津 300072)

3(中国科学院国家空间科学中心 北京 100190)

4(中国科学院大学 北京 100049)

摘要 最优估计法 (OEM) 作为一种新型的反演方法, 在激光雷达探测大气环境参数中具有重要应用价值. 为了表征基于 OEM 方法的瑞利激光雷达反演大气温度结果的可靠性, 本文推导了 OEM 不确定度公式, 明确了 OEM 反演结果的不确定度来源, 利用瑞利激光雷达方程模拟回波光子数廓线, 计算了相应的中层大气温度和不确定度, 得到在 OEM 反演过程中的主要不确定度来源, 即参考压强不确定度和噪声不确定度. 利用蒙特卡罗方法 (MCM) 建立 OEM 不确定度验证框架, 对不同不确定度来源产生的不确定度数值进行了验证, 结果表明两种不同方法计算的不确定度数值在海拔 85 km 以下一致性良好, 证明 OEM 不确定度计算数值的准确性. 此外, 基于 OEM 方法对瑞利激光雷达的实测数据进行反演, 并完成了不确定度分析, 为 OEM 方法在激光雷达大气环境探测领域的推广使用提供了依据.

关键词 瑞利激光雷达, 最优估计法, 不确定度, 大气温度, 蒙特卡罗方法

中图分类号 P356

Uncertainty Evaluation of Atmospheric Temperature Retrieval Using Rayleigh Lidar Based on the Optimal Estimation Method

ZHANG Xiaojian^{1,2} ZHANG Xianzhong^{1,2} LI Xinqi^{1,2} WU Tong^{1,2}
ZHONG Kai^{1,2} YAN Zhaoai^{3,4} XU Degang^{1,2} YAO Jianquan^{1,2}

1(School of Precision Instrument and Optoelectronic Engineering, Tianjin University, Tianjin 300072)

* 国家自然科学基金项目资助 (62175184)

2024-06-24 收到原稿, 2024-10-09 收到修定稿

©The Author(s) 2025. This is an open access article under the CC-BY 4.0 License
(<https://creativecommons.org/licenses/by/4.0/>)

2(Key Laboratory of Optoelectronic Information Science and Technology (Ministry of Education),
Tianjin University, Tianjin 300072)

3(National Space Science Center, Chinese Academy of Sciences, Beijing 100190)

4(University of Chinese Academy of Sciences, Beijing 100049)

Abstract As a new retrieval method, the Optimal Estimation Method (OEM) is playing an increasingly important role in detecting the atmospheric environment by lidar. In this study, the necessity of calculating the uncertainty of retrieval results was analyzed first. To characterize the reliability of the atmospheric temperature inversion results by lidar, the OEM uncertainty formula was derived and the sources of uncertainty was clarified. There were four sources of uncertainty, which were smoothing uncertainty, model parameter uncertainty, forward model uncertainty and noise uncertainty. Regarding the problem of using Rayleigh lidar to retrieve the temperature of the middle atmosphere, the forward model was derived based on the lidar equation, ideal gas state equation, and hydrostatic equilibrium as reasonable assumptions. Based on the simulated echo photon profiles of the Rayleigh lidar calculated by forward model, the middle atmospheric temperature and the corresponding uncertainty was calculated, which demonstrated that the main uncertainty sources in the OEM retrieval process are the reference pressure uncertainty and noise uncertainty. Using the Monte Carlo Method (MCM), the OEM uncertainty verification framework was established and the uncertainty values generated by different sources of uncertainty were verified. Afterwards, the total uncertainty value caused by all sources of uncertainty was also verified. Results showed that the uncertainty calculated by two different methods and the total uncertainty were consistent below the altitude of 85 km, proving the accuracy of the OEM uncertainty theories. In addition, temperature retrieval based on measured results by Rayleigh lidar was performed. The characteristics of the measured data were introduced in detail, and the retrieval results and corresponding uncertainties were calculated. The results prove the conclusion that the main uncertainty sources in the OEM inversion process are the reference pressure uncertainty and noise uncertainty. The work of this article paves the way for the applications of lidar in monitoring the atmospheric environment.

Key words Rayleigh lidar, Optimal Estimation Method (OEM), Uncertainty, Atmospheric temperature, Monte Carlo Method (MCM)

0 引言

中层大气为海拔在 10~100 km 范围内的垂直大气区域, 主要包括平流层、中间层以及热层底部. 研究中层大气的压力、密度和温度及其关系, 可以了解大气热力学过程, 对气候研究等具有重要意义. 作为一种重要的中层大气环境探测技术, 瑞利激光雷达具有连续观测、时空分辨率高、信噪比高等优点, 广泛应用于高度 30 km 以上的大气探测^[1]. 瑞利激光雷达工作时, 其发射的激光脉冲与大气分子相互作用生成散射信号, 并被接收端望远镜收集转换为实测光子回波数据^[2]. 实测数据通过反演运算得到大气环境参数. 由于反演过程仅能获得近似解而无法获得唯一解, 需

要计算近似解的不确定度以对其可靠性进行评估^[3]. 因此, 激光雷达的探测精度不仅与激光雷达系统硬件能力相关, 而且还与反演方法和相应不确定度计算方法有关.

目前, 瑞利散射激光雷达多使用 Chanin-Hauchecorne (CH) 方法反演中层大气温度, 该方法由 Hauchecorne 和 Chanin 在 1980 年首次提出^[4]. CH 方法利用雷达方程、理想气体状态方程和静力学平衡方程得到大气反演公式, 从探测上限选取某一温度参考点, 从顶部向下积分获得反演结果. 针对 CH 方法反演结果的不确定度, 已有文献给出相关推导及证明: 2016 年 Leblanc 等^[5]根据不确定度测量指南 (GUM) 建立了大气温度不确定度计算框架, 计算

了不确定度分量及大气温度合成标准不确定度, 对 NDACC 激光雷达温度反演产品进行详细评估; 2023 年 Li 等^[6]使用蒙特卡罗方法验证了 GUM 方法评估不确定度的准确性. 由于 CH 方法的局限性, 顶部参考点的选取会对反演结果产生较大影响, 通常舍弃顶部以下 10~15 km 范围内的反演结果以保证计算结果的准确性, 这会导致探测数据使用率低. 因此, 寻找新方法提升数据利用率十分必要. 2000 年 Roger^[7]提出的最优估计法 (Optimal Estimation Method, OEM) 应用于中层大气反演时能够解决上述问题. OEM 方法以贝叶斯理论为基础, 通过相应的物理假设描述实际探测时的物理过程, 通过优化算法迭代计算出反演结果, 具有充分利用探测数据、无须去噪预处理、反演速度快等优势^[8].

国际上, 自从 OEM 方法首次被提出并用于作者早期工作^[7], 逐渐被推广应用于遥感领域. 2003 年, Deeter 等^[9]使用 MOPITT 遥感设备对对流层一氧化碳 (CO) 和甲烷气体 (CH₄) 进行探测, 采用 OEM 方法对 CO 浓度进行了计算. 2004 年, Grant 等^[10]使用 OEM 方法通过部分反射式中频雷达脉冲信号计算了中间层电子密度. 随着激光雷达在大气探测领域的广泛应用, OEM 方法在该领域的应用潜力也得以显现. 2013 年, Povey^[11]基于双通道拉曼激光雷达信号, 使用 OEM 方法计算了 0~5 km 海拔范围内气溶胶的后向散射系数和消光系数. 2016 年 Sica 和 Halfele^[12]对位于瑞士的多通道振动拉曼散射激光雷达测量信号进行反演, 使用 OEM 方法计算了在晴天和多云两种不同条件下的水汽混合比. 在中国, 2022 年 Wang^[13]对 OEM 方法的原理进行了说明和分析, 基于瑞利激光雷达探测数据反演中层大气温度的实际过程建立正向模型, 利用仿真数据和实测数据分别对 OEM 方法反演中层大气温度的可行性进行了验证. 上述文献虽然使用 OEM 方法获得了反演结果, 但是对于反演结果的不确定度仅给出计算结果, 缺乏对不确定度系统深入的分析. 由于 OEM 方法与 CH 方法存在原理上的不同, CH 方法通过积分直接计算反演结果, 而 OEM 方法需要迭代求解, 基于 CH 反演方法的不确定度分析在 OEM 方法中并不适用. 因此, 对基于 OEM 方法反演结果的不确定度来源进行分析, 验证其不确定度计算结果的准确性是十分必要的.

本文通过建立反演过程的数学模型, 对 OEM 方法计算结果的不确定度来源进行推导, 并对不同种类的不确定度来源进行分析. 基于仿真数据开展 OEM 反演结果不确定度计算, 确定了最优估计法反演不确定度的主要来源是探测噪声不确定度和参考压强不确定度. 进一步, 使用蒙特卡罗方法 (Monte Carlo Method, MCM) 验证了 OEM 不确定度计算结果的准确性. 采用 OEM 方法对瑞利激光雷达实测数据进行了反演, 通过计算反演结果的不确定度值对不确定度的主要来源进行了验证.

1 原理与方法

1.1 不确定度理论基础

1.1.1 转换方程

对于任意的测量数据 y , 可以通过映射 f 联系未知状态量 x , 未知状态量 x 由测量结果计算得到, 是测量结果的输入数据. 在映射中, x 需要一些参数 b 补充由 x 至 y 的过程, 这些参数与状态量 x 无关. 另外, 由于观测仪器的固有缺陷, 测量结果中存在测量噪声, 还需考虑测量时的噪声项 ϵ . 综上, 可以得到表征测量数据 y 与未知状态量 x 之间的关系, 表示为

$$y = f(x, b) + \epsilon. \quad (1)$$

其中, y 表示测量结果; 映射 f 称为正向公式, 用以描述从状态量到测量结果的整个物理过程, 例如本文所采用的瑞利激光雷达方程; b 称为正向公式参数, 表示所有影响反演过程的参数值, 其中包括激光雷达系统参数, 例如激光脉冲能量、重复频率等; 还包括进行大气热力学计算所使用的值, 例如玻尔兹曼常数 k 、重力加速度 g 等. 某些 b 值的选取会对反演结果造成一定影响. 误差项 ϵ 表示探测器噪声, 通常是一个随机值, 对于激光雷达来说, ϵ 呈泊松分布.

反演是从有限的测量结果中推断未知的输入状态, 利用反演方法通过测量数据计算出正确的近似解. 定义反演方法 R 和近似解 $\hat{x}$, 可以得到

$$\hat{x} = R(y, \hat{b}, x_a, c). \quad (2)$$

其中, 尖角符号表示该值是通过经验获得或计算对某种状态量的估计值; $\hat{b}$ 为正向公式中使用参数的估计值; x_a 为对状态量的先验估计, 即通过已有经验对未

知状态量 $\mathbf{x}$ 的合理估计添加先验信息有利于提高解的可靠性; $\mathbf{c}$ 为反演方法过程中使用到的参数, 例如收敛标准参数等, $\mathbf{c}$ 不存在于正向公式中, 因此将 $\mathbf{c}$ 称为反演方法参数.

将式 (1) 和式 (2) 联立可得

$$\hat{\mathbf{x}} = \mathbf{R} \left[\mathbf{f}(\mathbf{x}, \mathbf{b}) + \epsilon, \hat{\mathbf{b}}, \mathbf{x}_a, \mathbf{c} \right]. \quad (3)$$

式 (3) 称为转换方程, 包含激光雷达系统实际探测过程以及反演方法两部分, 是进行不确定度分析的基础.

1.1.2 不确定度来源

反演问题是一个不适定问题, 其最终计算结果是近似解, 量化该近似解的准确性是必要的^[3]. 不确定度 (Uncertainty) 是指某个估计结果的不确定程度或可信程度^[6]. 在 OEM 方法的计算过程中, 随着迭代次数的增加, 成本函数值不断减小, 当成本函数值最小时定义此时的解为估计解^[14]. 在得到最终解之前, 需要明确估计值偏离真实值的原因, 分析在迭代过程中扰动求解的因素, 也就是 OEM 计算结果的不确定度来源.

真实世界的物理系统非常复杂, 涉及多个因素、多种相互作用, 因此物理模型的建立需要进行取舍. 在 OEM 方法中, 正向方程建立过程的取舍要保证正向方程的真实性和准确性, 这种实际建立的近似正向方程称为正向模型 $\mathbf{F}$, 其与正向公式 $\mathbf{f}$ 的关系为

$$\begin{aligned} \mathbf{F}(\mathbf{x}, \mathbf{b}) &\approx \mathbf{f}(\mathbf{x}, \mathbf{b}, \mathbf{b}'), \\ \Delta \mathbf{f}(\mathbf{f}, \mathbf{b}, \mathbf{b}') &= \mathbf{f}(\mathbf{x}, \mathbf{b}, \mathbf{b}') - \mathbf{F}(\mathbf{x}, \mathbf{b}). \end{aligned} \quad (4)$$

其中, $\mathbf{b}$ 参数被拆分为 $\mathbf{b}$ 和 $\mathbf{b}'$, $\mathbf{b}$ 替换为正向模型参数, $\mathbf{b}'$ 代表正向模型中忽略的某些正向公式参数, $\Delta \mathbf{f}(\mathbf{x}, \mathbf{b}, \mathbf{b}')$ 表示正向模型与正向公式之间的偏差.

将式 (3) 中的正向公式替换为正向模型, 得到

$$\hat{\mathbf{x}} = \mathbf{R} \left[\mathbf{F}(\mathbf{x}, \mathbf{b}) + \Delta \mathbf{f}(\mathbf{x}, \mathbf{b}, \mathbf{b}') + \epsilon, \hat{\mathbf{b}}, \mathbf{x}_a, \mathbf{c} \right]. \quad (5)$$

当 $\mathbf{x} = \mathbf{x}_a$, $\mathbf{b} = \hat{\mathbf{b}}$ 时, 对正向模型 $\mathbf{F}$ 进行线性处理, 得到

$$\begin{aligned} \mathbf{f}(\mathbf{x}, \mathbf{b}, \mathbf{b}') &= \mathbf{F}(\mathbf{x}_a, \hat{\mathbf{b}}) + \mathbf{K}_x (\mathbf{x} - \mathbf{x}_a) + \\ &\quad \mathbf{K}_b (\mathbf{b} - \hat{\mathbf{b}}) + \Delta \mathbf{f}(\mathbf{x}, \mathbf{b}, \mathbf{b}'). \end{aligned} \quad (6)$$

其中, $\mathbf{K}_x = \partial \mathbf{F} / \partial \mathbf{x}$ 为正向模型对状态量的导数; $\mathbf{K}_b = \partial \mathbf{F} / \partial \mathbf{b}$ 为正向模型对正向模型参数的导数. 将式 (6) 代入式 (2) 可得

$$\begin{aligned} \hat{\mathbf{x}} &= \mathbf{R} \left[\mathbf{F}(\mathbf{x}_a, \hat{\mathbf{b}}) + \mathbf{K}_x (\mathbf{x} - \mathbf{x}_a) + \right. \\ &\quad \left. \mathbf{K}_b (\mathbf{b} - \hat{\mathbf{b}}) + \Delta \mathbf{f}(\mathbf{x}, \mathbf{b}, \mathbf{b}') + \epsilon, \hat{\mathbf{b}}, \mathbf{x}_a, \mathbf{c} \right]. \end{aligned} \quad (7)$$

对式 (2) 中的 $\mathbf{y}$ 进行线性处理, 得到

$$\begin{aligned} \hat{\mathbf{x}} &= \mathbf{R} \left[\mathbf{F}(\mathbf{x}_a, \hat{\mathbf{b}}), \hat{\mathbf{b}}, \mathbf{x}_a, \mathbf{c} \right] + \\ &\quad \mathbf{G}_y \left[\mathbf{K}_x (\mathbf{x} - \mathbf{x}_a) + \mathbf{K}_b (\mathbf{b} - \hat{\mathbf{b}}) + \right. \\ &\quad \left. \Delta \mathbf{f}(\mathbf{x}, \mathbf{b}, \mathbf{b}') + \epsilon \right]. \end{aligned} \quad (8)$$

其中, $\mathbf{G}_y = \partial \mathbf{R} / \partial \mathbf{y}$ 为反演方法对测量结果的导数. 对于一个准确的反演方法, 式 (8) 中 $\mathbf{R}[\mathbf{F}(\mathbf{x}_a, \hat{\mathbf{b}}), \hat{\mathbf{b}}, \mathbf{x}_a, \mathbf{c}]$ 项的值为 $\mathbf{x}_a$. 若该项结果不为 $\mathbf{x}_a$, 证明该反演方法使用不当, 需替换正确的反演方法.

定义矩阵 $\mathbf{A}$ 为平均核矩阵, 其含义为估计解 $\hat{\mathbf{x}}$ 对真实值 $\mathbf{x}$ 的导数, 表示为

$$\mathbf{A} = \mathbf{G}_y \mathbf{K}_x. \quad (9)$$

式 (8) 等号左右两边同时减去 $\mathbf{x}$, 得到反演方法的不确定度来源表达式如下:

$$\begin{aligned} \hat{\mathbf{x}} - \mathbf{x} &= (\mathbf{A} - \mathbf{I})(\mathbf{x} - \mathbf{x}_a) + \\ &\quad \mathbf{G}_y \mathbf{K}_b (\mathbf{b} - \hat{\mathbf{b}}) + \\ &\quad \mathbf{G}_y \Delta \mathbf{f}(\mathbf{x}, \mathbf{b}, \mathbf{b}') + \\ &\quad \mathbf{G}_y \epsilon. \end{aligned} \quad (10)$$

其中, $\mathbf{I}$ 为单位矩阵. 由式 (10) 可知, 估计解 $\hat{\mathbf{x}}$ 的不确定度来源分为 4 项, 具体含义如下.

第 1 项 $(\mathbf{A} - \mathbf{I})(\mathbf{x} - \mathbf{x}_a)$, 定义为平滑不确定度. 平滑不确定度是由于观测系统无法探测大气的精细结构而对反演结果引入平滑处理带来的偏差. 理想状态下, 平均核矩阵的每一行是一个尖峰函数, 其半高全宽定义为该点的距离分辨率. 在底部信噪比较高时, 平均核每一行近似为一个尖峰函数; 随着海拔的提升, 信噪比不断降低, 每一行的半高全宽值不断增大, 该点的距离分辨率降低, 从而引入平滑不确定度. 也可以从另一方面理解平滑不确定度: 理想状态下, $\mathbf{A}$ 为单位矩阵, 此时该项值为 0, 即反演结果中没有先验信息的影响; 在实际使用中, 随着海拔的增高, 矩阵 $\mathbf{A}$ 每一行数值的半高全宽不断增大, 该项数值不为零, 引入先验信息的影响, 使最终的反演结果出现偏差.

第 2 项 $\mathbf{G}_y \mathbf{K}_b (\mathbf{b} - \hat{\mathbf{b}})$, 定义为正向模型参数不确定度. 模型参数不确定度的引入是必然的. 由于测

量仪器的局限性和实际物理过程的复杂性, 一些模型参数自身存在一定的不确定度, 例如重力加速度 g 等, 引入这些值会使最终的反演结果与真实结果出现偏差.

第 3 项 $\mathbf{G}_y \Delta \mathbf{f}(\mathbf{x}, \mathbf{b}, \mathbf{b}')$, 定义为正向模型不确定度, 如式 (4) 所示. 由于正向模型的构建会使用近似和简化处理. 这些处理方式会导致正向模型的局限性, 使正向模型和正向公式之间产生偏差, 并将不确定度引入最终反演结果.

第 4 项 $\mathbf{G}_y \epsilon$, 定义为噪声不确定度. 探测大气环境的激光雷达系统一般使用单光子探测器, 探测器探测的光子数服从泊松分布, 由探测器产生的泊松噪声会对最终的反演结果造成影响.

上述不确定度分析是以线性反演问题为前提进行讨论的. 由于通过瑞利激光雷达探测信号反演中层大气温度为非线性反演问题, 需要通过迭代过程求出最终解, 但不确定度分析结果与上述结果一致, 因此可以沿用至本文讨论的实际问题^[7].

1.2 反演结果不确定度计算与验证框架

1.2.1 反演大气温度结果的正向模型

利用瑞利激光雷达探测数据反演中层大气温度廓线时, 中层大气温度为状态量, 激光雷达回波光子数廓线为测量数据. 中层大气服从理想气体定律以及流体静力平衡关系, 有

$$p = nkT, \quad (11)$$

$$dp = -g(z)\rho(z)dz. \quad (12)$$

式 (11) 表示的理想气体状态方程中, p 为压强, n 为气体分子数密度, $k=1.380649 \times 10^{-23} \text{ J} \cdot \text{K}^{-1}$ 为玻尔兹曼常数. 式 (12) 表示的流体静力平衡表达式中, $g(z)$ 和 $\rho(z)$ 分别为高度在 z 处的重力加速度和大气密度.

瑞利激光雷达方程可表示为

$$N_R(\lambda, z) = \left[\frac{P(\lambda) \Delta t}{hc/\lambda} \right] [\sigma_R(\lambda) n(z) \Delta z] \cdot \left[\frac{A_r}{(z - z_L)^2} \right] [\eta(\lambda) T_r^2(\lambda, z)] + N_B. \quad (13)$$

式中 A_r 表示使用光子计数的瑞利激光雷达接受到不同高度光子数的值.

联立式 (11)~(13) 得到正向模型如下:

$$N_R(z) = \frac{C}{(z - z_0)^2} \cdot \frac{p(z)}{kT(z)} + N_B; \quad (14)$$

$$\begin{cases} C = \left[\frac{P(\lambda) \Delta t}{hc/\lambda} \right] [\sigma_R(\lambda) \Delta z A] [\eta(\lambda) T_r^2(z)], \\ p(z) = p_0 \exp \left[\frac{1}{R} \int_z^{z_{\text{top}}} \frac{M(z') g(z')}{T(z')} dz' \right]. \end{cases}$$

其中, $N_R(z)$ 为返回激光雷达接收端的光子数, C 为激光雷达系统参数, $P(\lambda)$ 为激光雷达的发射功率, λ 为发射激光的波长, Δt 为激光雷达正常工作时的有效探测时间, c 为光速, h 为普朗克常数, $\eta(\lambda)$ 为激光雷达系统的总效率, A 为接收望远镜的有效面积, $\sigma_R(\lambda)$ 为瑞利散射截面, Δz 为激光雷达系统探测时的距离分辨率, $p(z)$ 表示随海拔变化的压强曲线, p_0 为参考点压强 (这里为顶部高度的压强值), g 为重力加速度, M 为大气的平均摩尔质量, $T_r(z')$ 为激光光束从发射端输出, 传播至某一高度后返回接收端整个过程的大气透过率, N_B 为噪声.

1.2.2 最优估计法不确定度计算

1.1.2 节中提到的 4 项不确定度值通过计算每一项对应的协方差矩阵得出.

第 1 项, 平滑不确定度, 其计算公式为

$$\mathbf{S}_s = (\mathbf{A} - \mathbf{I}_n) \mathbf{S}_e (\mathbf{A} - \mathbf{I}_n)^T. \quad (15)$$

通过该项的不确定度表达式 $(\mathbf{A} - \mathbf{I})(\mathbf{x} - \mathbf{x}_a)$ 可以看出, 计算该项需要了解真实的状态量 $\mathbf{x}$. 由于测量条件的限制, 真实状态量 $\mathbf{x}$ 的值无法估计, 通常选取平均值状态量 $\bar{\mathbf{x}}$ 代替. $\mathbf{S}_e$ 代表平均状态量的协方差矩阵, 该值由先验协方差矩阵 $\mathbf{S}_a$ 代替计算. 式 (15) 中, 当矩阵 $\mathbf{A}$ 近似为单位矩阵时, 该式等于零, 即当反演结果质量较好时, 该项的值可以忽略. 同时, 由于无法判断先验协方差矩阵 $\mathbf{S}_a$ 代替的准确性, 通过这种方法计算出的不确定度值并不准确. 此外, 在实际应用中反演效果较好的区域内平滑不确定度值与总不确定度值最多相差一个数量级. 综合上述三个原因, 本文忽略平滑不确定度的影响.

第 2 项, 正向模型参数不确定度, 其计算表达式为

$$\mathbf{S}_f = \mathbf{G} \mathbf{K}_b \mathbf{S}_b \mathbf{K}_b^T \mathbf{G}^T. \quad (16)$$

式中, $\mathbf{K}_b = \partial \mathbf{F} / \partial \mathbf{b}$ 为正向模型对正向模型参数的导数. 正向模型参数 $\mathbf{b}$ 不唯一, 在此处的瑞利激光雷达反演中层大气问题中, 正向模型参数包括重力加速度、瑞利散射截面、平均大气摩尔质量、参考压强. 当参数不同时, 其自身不确定度不同, 协方差矩阵 $\mathbf{S}_b$ 也

不同;此外,正向模型对不同正向模型参数的导数也不同,所以矩阵 $\mathbf{K}_b$ 不同.因此,不同正向模型参数需要分别计算.

第3项,正向模型不确定度,其表达式为 $\mathbf{G}_y \Delta \mathbf{f}(\mathbf{x}, \mathbf{b}, \mathbf{b}')$,该项是系统不确定度.在瑞利激光雷达反演中层大气的问题中,正向模型的建立以三个合理模型汇总而成,分别是瑞利激光雷达公式、理想气体状态方程和流体静力平衡.一般认为,中层大气满足理想气体状态方程、流体静力平衡两种合理假设,由此建立正向模型满足准确性要求,且能够满足激光雷达探测过程的实际物理状态.因此,后续计算不考虑该项不确定度.

第4项,噪声不确定度,其计算表达式为

$$\mathbf{S}_\epsilon = \mathbf{G}_y \mathbf{S}_\epsilon \mathbf{G}_y^T. \quad (17)$$

在单光子探测器中,每个光子被探测到被视为一个随机事件,在一段固定时间内光子的到达次数符合泊松分布.因此,噪声的协方差矩阵 $\mathbf{S}_y$ 通常为对角矩阵,表示为

$$\mathbf{S}_{\text{noise}} = \mathbf{G}_y \mathbf{S}_y \mathbf{G}_y^T. \quad (18)$$

综上所述,本文中 OEM 反演结果的不确定度来源主要包括第2项正向模型不确定度和第4项噪声不确定度,总的不确定度计算表达式为

$$\mathbf{S}_{\text{total}} = \sum \mathbf{S}_b + \mathbf{S}_{\text{noise}}. \quad (19)$$

1.2.3 蒙特卡罗方法验证最优估计法不确定度框架

在1.1.1节中建立了转换方程,并通过转换方程进行线性化处理得到反演结果不确定度的来源.该推导过程对使用 OEM 方法解决反演问题具有普适性,但缺乏应用于具体问题的数值计算准确性验证.蒙特卡罗方法(MCM)是一种基于随机抽样和统计模拟的数值计算方法,通过生成大量的随机样本来近似计算复杂问题的数值解或概率分布,特别适用于无法通过解析方法求解的问题.在本文的反演问题中,通常从观测数据中推断出导致这些观测数据的模型参数或系统状态,由于反演问题通常包含复杂的数学模型、大量的未知参数和噪声影响,难以通过解析方法直接求解.因此,本文使用蒙特卡罗方法验证不确定度数值计算的准确性.

通过蒙特卡罗方法计算测量不确定度时,针对每个输入量 X_i ,规定 X_i 根据其数值特征依据相应的概率密度函数进行抽样,通过每一个抽样值计算反演结

果 Y 的一个离散抽样值 Y_i .重复上述过程多次后,获取到大量反演结果,分析大量反演结果可获得 Y 的数值分布情况,以此评估不同来源的不确定度准确性.使用蒙特卡罗方法计算 OEM 不确定度的步骤如下.

(1) 确定需要检验的不确定度来源.

(2) 建立进行不确定度评定所需的验证模型.针对 OEM 方法,验证模型通过更改正向模型获得.

(3) 确定不同不确定度来源的概率密度分布.模型参数不确定度为实际物理量,认为其遵循高斯分布,探测噪声遵循泊松分布.

(4) 明确蒙特卡罗试验的样本量 N .

(5) 在每一次蒙特卡罗试验中,对每个不确定度来源($X_1, X_2, \dots, X_n$),依据不同不确定度来源的不同概率密度分布进行离散抽样.随后准备执行 M 次蒙特卡罗试验,并用符号 $x_{i,r}$ 表示第 i 个不确定度来源的第 r 次离散抽样的值.

(6) 将每个不确定度来源的第 r 次抽样值($x_{1,r}, x_{2,r}, \dots, x_{n,r}$)代入验证模型,计算输出量 Y 的第 r 个对应的反演结果 y_r .在执行了 M 次蒙特卡罗试验后,得到反演结果 Y 的 M 个离散值($y_1, y_2, \dots, y_M$).

(7) 计算 $y_1, y_2, \dots, y_M$ 的平均值作为反演结果 Y 的最佳估计值 y ,通过 y 值计算 M 个离散值的标准差,该标准差即为 y 的标准不确定度.

经过上述步骤后,将该标准不确定度计算结果与 OEM 不确定度计算结果进行对比,能够验证 OEM 方法计算结果的准确性.具体地,通过比较两种方法计算的不确定度值相对偏差来评价 OEM 不确定度数值的准确性,相对偏差计算方法为

$$r = \left| \frac{u_{\text{MCM}} - u_{\text{OEM}}}{u_{\text{MCM}}} \right| \times 100\%. \quad (20)$$

其中, u_{MCM} 表示蒙特卡罗法计算结果(基准值), u_{OEM} 表示最优估计法计算结果(待比较值).当相对偏差为10%时,基本值与比较值的差异相比于计算得出的不确定度数值可忽略,因此可以将10%作为数值准确性判断的标准.

2 结果与分析

2.1 基于仿真数据的最优估计法反演结果不确定度计算

仿真试验以式(14)正向模型为基础,选取某一

海拔范围内的中层大气温度作为温度真值; 设置合适的激光雷达系统参数, 并利用瑞利激光雷达方程生成模拟回波光子数; 依据该回波光子数使用 OEM 方法进行反演, 得到相应海拔范围内的大气温度结果, 计算其不确定度, 并对反演结果的不确定度进行分析. 本文选择 NRLMSISE-00 大气模型在 2022 年 9 月 7 日, 117.169884°E 和 39.113447°N 的输出数据为真值, 模拟激光雷达夜间探测环境, 使用正向模型生成测量结果. 模拟过程中使用的瑞利激光雷达系统参数见表 1.

为避免对流层恶劣天气的影响并有效降低太阳光散射的噪声, 设置瑞利激光雷达所在海拔为 20 km^[15], 瑞利激光雷达探测中层大气的高度范围设置为 30~85 km, 测量原始数据距离分辨率为 100 m, 积分时间为 3600 s, 模拟的回波光子数如图 1 所示.

探测器接收到的激光雷达回波信号中夹杂着背景噪声以及泊松噪声, 随着海拔升高, 回波信号强度减弱, 噪声的影响逐渐明显. 通过计算该条回波廓线的信噪比 (Signal-to-Noise Ratio, SNR) 可以衡量回波信号的质量, 信噪比的计算表达式为

$$R_{\text{SNR}} = \frac{N_{\text{R}}(z) - N_{\text{B}}}{\sqrt{N_{\text{R}}(z)}}. \quad (21)$$

其中, N_{R} 表示探测到的回波信号, N_{B} 表示背景噪声. 针对图 1 所示的光子数廓线, 相应的信噪比如图 2 所示, 其中在 85 km 海拔处信噪比为 7.26.

根据前文所述, 反演结果的不确定度包含模型参数不确定度以及噪声不确定度两部分. 模型参数不确定度包括参考压强 p_0 、大气分子摩尔质量 M 、瑞利散射截面 σ_{R} 以及重力加速度 g 共 4 项, 相关参数的来源以及不确定度对应的数值大小如表 2 所示. 对于模型参数不确定度, 其计算表达式见式 (16), 式中矩阵

表 1 模拟激光雷达系统参数

Table 1 Simulation parameters of the lidar system

系统参数	选值
脉冲能量/mJ	40
重复频率/Hz	50
激光波长/nm	532
望远镜口径/mm	350
系统效率	0.191
噪声光子数	10

K_b 使用数值微分方法计算得出. 根据 1.2.2 节的不确定度计算公式, 计算得到 5 项不同来源的不确定度, 如图 3 所示, 其中的红色廓线表示总不确定度随海拔的变化.

在图 3 中, p_0 代表参考压强、 g 为重力加速度、 M 为大气分子摩尔质量、 N_{B} 为噪声、 σ_{R} 为瑞利散射截面不确定度, u 为总不确定度. 从图 3 可以看出, 噪声不确定度的数值总体上最大, 是影响反演结果的主要原因, 且随着海拔的增加, 其数值持续增大, 在 85 km 处达到最大值 9.24 K. 其次是参考压强的不确定度, 随海拔的增加也呈现不断增大的趋势; 在顶部

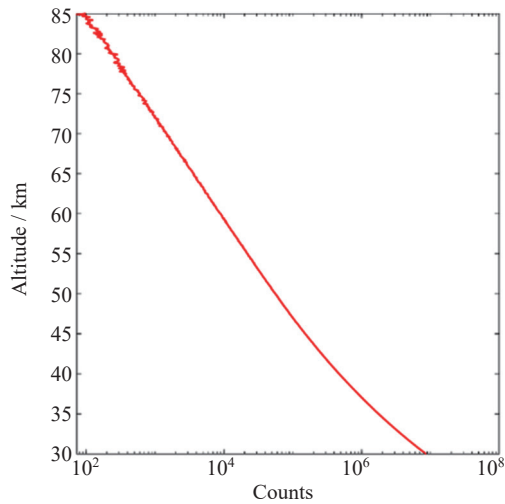


图 1 模拟回波光子数廓线

Fig. 1 Simulated echo photon number profile

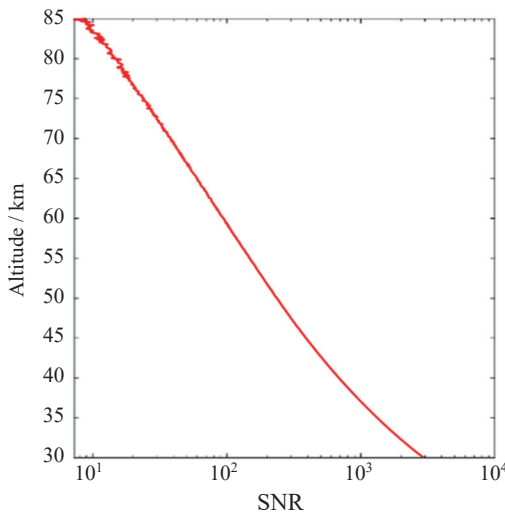


图 2 回波光子廓线信噪比

Fig. 2 SNR of simulated echo photon profile

表 2 不确定度来源汇总
Table 2 Summary of uncertainty sources

系统参数	参数来源	不确定度/(%)
参考压强	Ref. [16]	5
重力加速度	Ref. [17]	0.001
大气分子摩尔质量	Ref. [18]	0.003
瑞利散射截面	Ref. [19]	0.2
噪声	—	服从Poisson分布

时信噪比较低,参考压强不确定度的增长迅速,在 85 km 处达到 9.09 K. 重力加速度和大气分子摩尔质量在 30~80 km 海拔范围内不确定度大小基本保持不变,在 80~85 km 海拔范围内数值降低,该数值变化的原因与实际物理过程无关,仅与反演方法有关. 瑞利激光雷达反演中层大气温度问题属于非线性反演问题,因此应用 OEM 方法求解时,需要通过迭代求出最终的近似解. 在每一次迭代过程中,需要计算上次迭代结果的雅可比矩阵以获得本次的迭代步长和迭代方向. 不同高度的信号质量不同,导致迭代步长和迭代方向也不同,影响这两种因素的来源也不同. 对于顶部的状态量来说,参考压强和噪声对反演结果的影响会大于重力加速度和大气分子摩尔质量的影响,因此后面二者的不确定度值会降低. 综上所述,噪声不确定度与参考压强不确定度是影响反演结果不确定度的两个主要来源,相比之下其他来源要小得多,甚至可以忽略.

2.2 蒙特卡罗方法验证最优估计法不确定度

蒙特卡罗方法设置的样本量 N 对结果验证有重要影响,其大小与试验次数相同. 样本量越大,蒙特卡罗试验次数越多,最终获得的计算结果数量越多,对计算结果的概率分布估计就越准确,但计算时间也会变长;当样本量较小时,可能导致反演结果 Y 的最佳估计值 y 不够准确. 在设置蒙特卡罗试验参数时,需要对样本量和计算时间进行权衡,确保最终计算结果准确可靠的同时,具有较高的计算效率.

以验证噪声不确定度为例,设置样本量大小分别为 100, 1000, 10000, 50000, 获得不同的计算结果,如图 4 所示.

根据图 4 可以看出,样本量越大则蒙特卡罗试验结果越稳定,可靠性越高,但样本值达到一定量级时即可实现收敛,无须再增加蒙特卡罗试验次数. 在噪

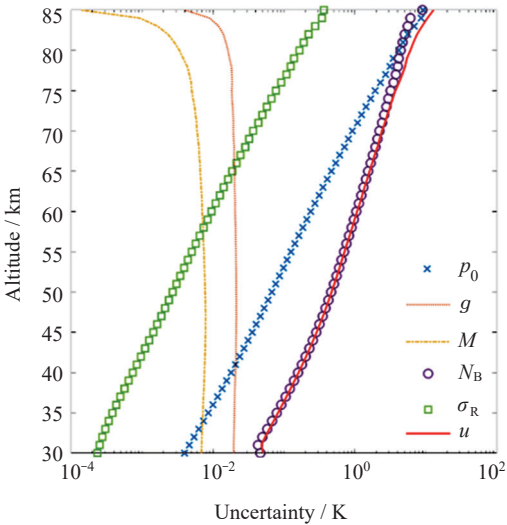


图 3 不确定度结果汇总
Fig. 3 Summary of uncertainty results

声不确定度的样本量确定过程中,50000 次与 100000 次的试验结果相对偏差为 0.96%,远小于 1.2.3 节中规定的 10% 偏差标准,可认为蒙特卡罗试验结果收敛. 通过对 5 种不确定度来源的分析,本文将蒙特卡罗试验次数统一设置为 50000 次,然后计算最佳估计解 y 的标准差. 计算时,首先对不同的不确定度来源分别进行计算,来源类型与表 2 相同,相应的计算结果如图 5 所示.

依据图 5 中的曲线可以看出,蒙特卡罗试验的计算结果与 OEM 方法计算结果在趋势上相符. 模型参数的不确定度中,参考压强 p_0 引入的不确定度、瑞利散射截面的不确定度以及噪声的不确定度值均随着高度不断增加,在海拔上限达到最大值. 模型参数不确定度中的重力加速度不确定度以及平均大气分子摩尔质量两种不确定度通过两种不同计算方法的结果的趋势一致,在顶部达到最小值.

对于数值的准确性判断,依据式 (20) 计算相对偏差的最大值,如表 3 所示. 可以看出,5 种不同来源的不确定度验证结果相对偏差均小于 5%,满足小于 10% 的验证标准,说明 5 项不确定度的 OEM 方法计算结果通过验证.

将 5 项不确定度来源同时代入蒙特卡罗试验计算出总不确定度数值,并与式 (19) 给出的 OEM 方法总不确定度计算公式进行对比,结果如图 6 所示. 两种计算结果趋势一致,总不确定度相对偏差值为 4.11%,同样小于 10% 标准,证明 OEM 总不确定度

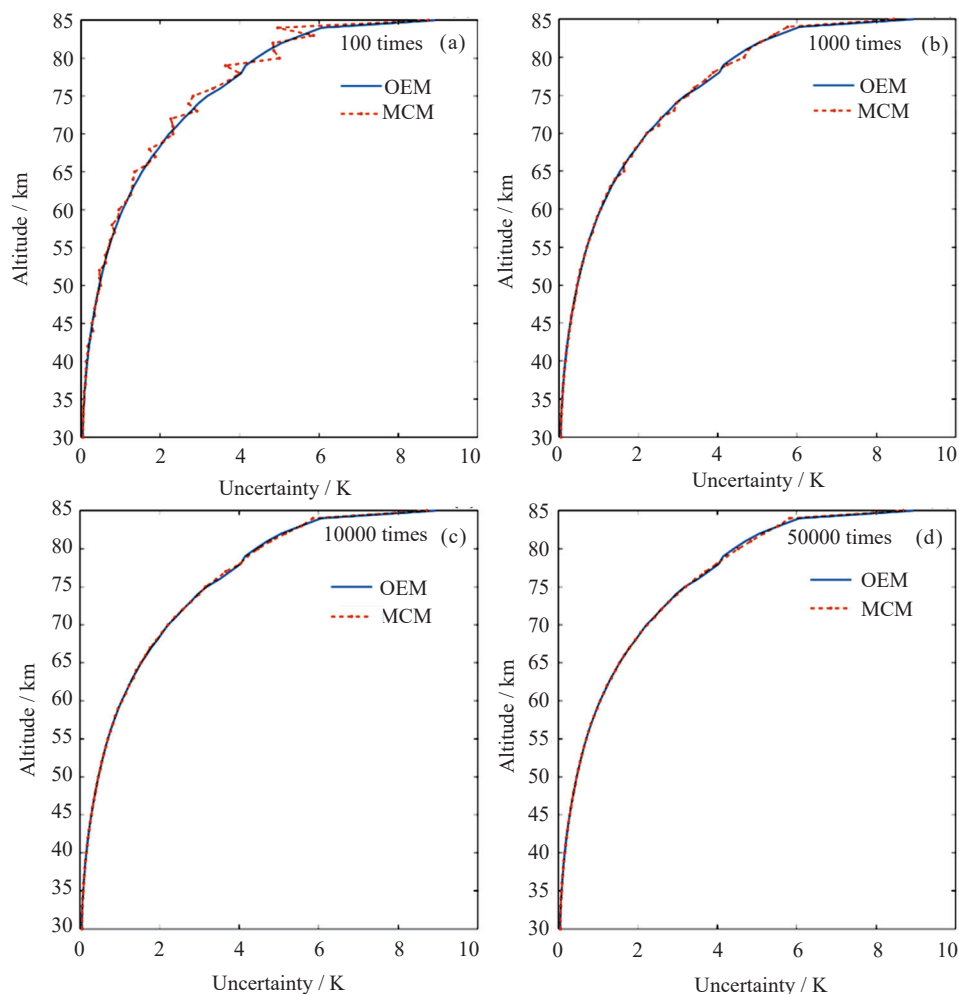


图 4 不同次数的蒙特卡罗试验计算结果

Fig. 4 Calculation results of Monte Carlo experiments with different number of trials

的计算结果是正确的。

综上所述, 最优估计法不确定度计算结果通过蒙特卡罗试验验证, 证明 OEM 不确定度计算过程适用于瑞利激光雷达反演中层大气温度的具体问题, 且计算结果准确可靠。

2.3 基于瑞利激光雷达实测数据不确定度计算

计算实例中, 利用 OEM 方法对中国科学院国家空间科学中心研制的瑞利激光雷达实测回波数据进行反演, 进一步计算反演结果不确定度并进行分析。该激光雷达为地基雷达, 系统参数如表 4 所示, 数据采集地点为 116.68°E, 40.33°N。

所采用的实测数据探测时间为 2018 年 9 月 3 日 17:30 UTC, 原始回波光子的空间分辨率约为 122.8 m, 如图 7(a) 所示。由于机械斩波器的影响, 回波数据在 30 km 高度以下存在异常, 因此选取

40 km 为有效探测范围的下限。相应的信噪比廓线如图 7(b) 所示, 选取 40~90 km 作为 OEM 反演的高度范围。

使用式 (19) 计算反演结果不确定度, 如图 8 所示。可以看出, 基于实测数据的 OEM 反演结果不确定度与仿真数据趋势基本一致, 主要的不确定度来源为参考压强不确定度和噪声不确定度, 此结论也与仿真结果一致。在 90 km 高度处, 参考压强不确定度值为 7.92 K, 噪声不确定度值为 9.74 K, 总不确定度值为 12.56 K。

3 结语

根据 OEM 反演方法不确定度通用来源类型的推导, 结合瑞利激光雷达反演大气温度的实际问题, 明

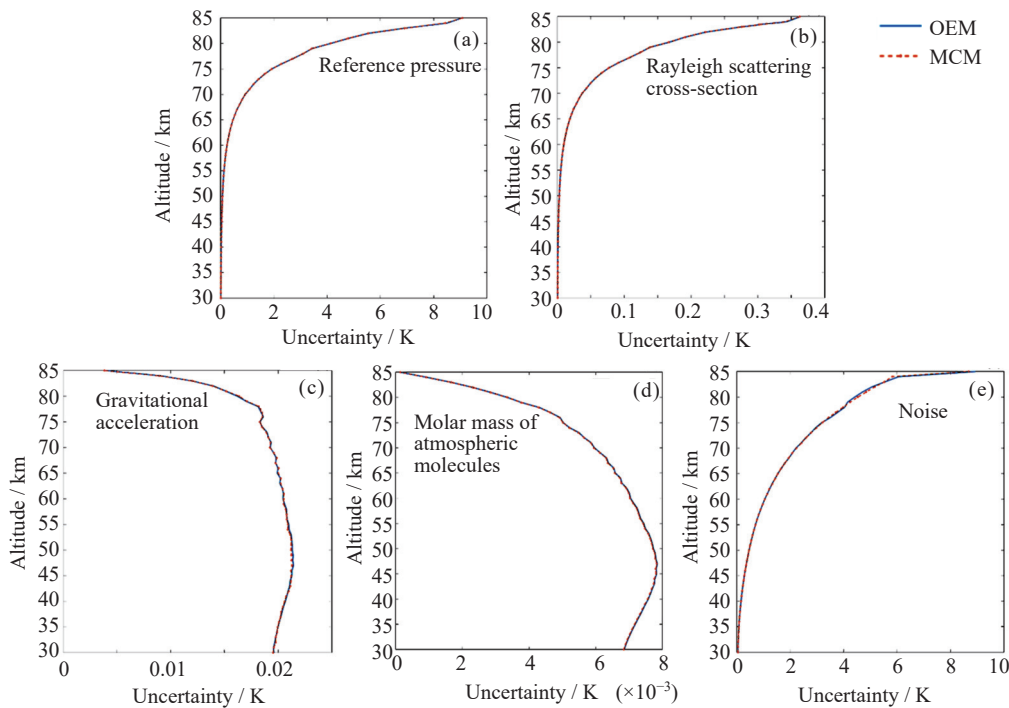


图 5 蒙特卡罗试验验证结果
Fig. 5 Monte Carlo experimental verification results

表 3 相对偏差计算结果汇总表
Table 3 Calculation results of the relative deviation

不确定度来源	相对偏差最大值/(%)
参考压强	0.74
重力加速度	4.00
平均大气分子摩尔质量	1.67
瑞利散射截面	0.59
噪声	4.18

表 4 地基瑞利激光雷达系统参数
Table 4 System parameters of the ground-based Rayleigh lidar

系统参数	参数选值
脉冲能量/mJ	500
脉冲宽度/ns	8
重复频率/Hz	30
激光波长/nm	532
望远镜口径/mm	1000
系统效率	0.02

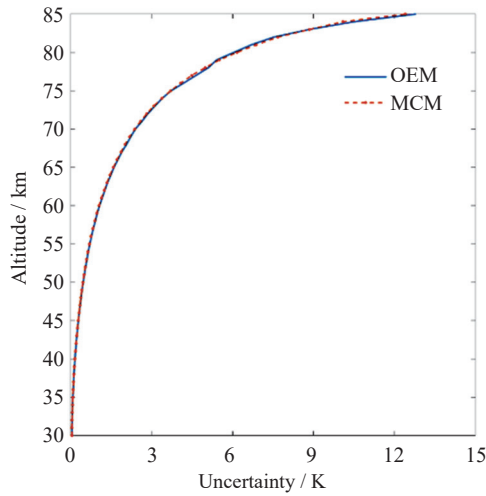


图 6 总不确定度验证结果
Fig. 6 Verification results of total uncertainty

确了其具体不确定度来源; 基于模拟数据通过蒙特卡罗方法证明了 OEM 不确定度计算的准确性, 并在瑞利激光雷达实测中层大气温度中进行了应用与分析. 主要研究结果如下.

基于转换方程对 OEM 反演方法的不确定度来源进行了理论推导, 确定 OEM 反演结果受到平滑不确定度、模型参数不确定度、模型不确定度以及噪声不确定度的影响. 通过分析, 在实际问题处理中可以舍弃平滑不确定度和模型不确定度的影响, 仅需计算模型参数不确定度和噪声不确定度. 依据实际问题建立正向模型, 通过模拟仿真定量分析了瑞利激光雷达反

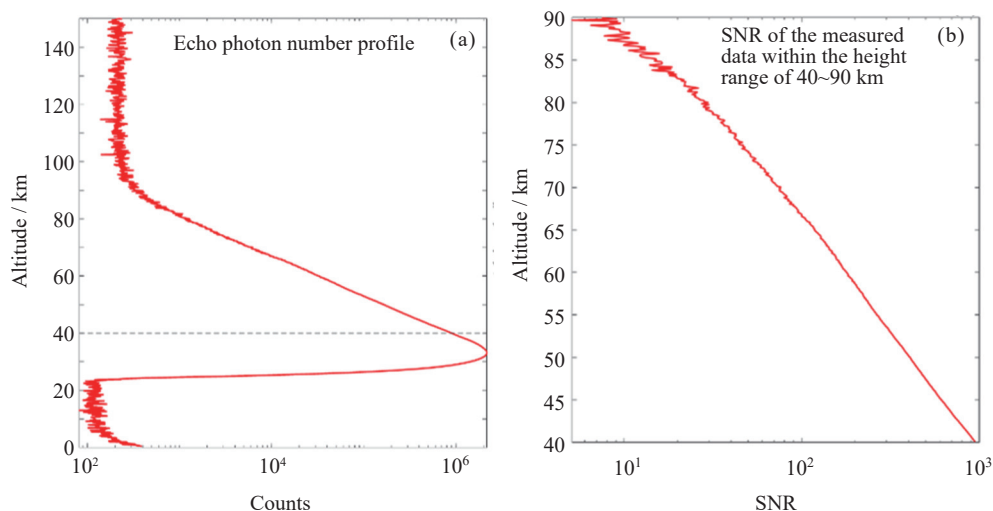


图 7 地基瑞利激光雷达实测数据

Fig. 7 Profiles of the ground-based Rayleigh lidar

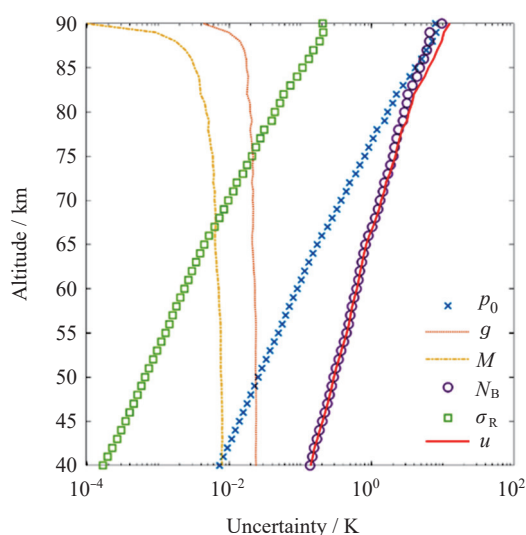


图 8 实测数据反演结果不确定性

Fig. 8 Uncertainty of the retrieval results based on measured data

演中层大气温度的不确定性来源, 确定了 OEM 方法计算不确定性时应主要计算参考压强不确定性以及噪声不确定性, 并对不同不确定性来源的计算结果进行分析和解释。接下来, 利用蒙特卡罗方法验证了 OEM 方法计算不确定性结果的准确性。制定了蒙特卡罗试验验证框架, 选取蒙特卡罗试验样本量为 50000 并确定相对误差最大值小于 10% 的验证标准, 基于此完成了 OEM 计算结果的可靠性验证。对瑞利激光雷达实测数据进行 OEM 反演并计算了反演结果不确定性。根据实测数据计算结果, 在 90 km 高度处

参考压强不确定性值为 7.92 K, 噪声不确定性值为 9.74 K, 总不确定性值为 12.56 K, 证明了在瑞利激光雷达反演中层大气问题中计算 OEM 不确定性时主要来源是参考压强不确定性以及噪声不确定性这一结论的正确性。

本文完成了 OEM 反演方法的不确定性计算及验证, 未来可以为各种主动探测设备应用 OEM 方法进行数据反演提供参考。但值得注意的是, 由于缺少完备的大气探测数据库, 在建立正向模型的过程中, 某些参数由于缺少具体数值和不确定性而被舍弃。因此, 某些模型参数的不确定性评估并未在本文体现, 未来希望能通过完备的大气探测数据库建立更加准确的正向模型, 从而获得更加可靠的反演结果不确定性。

致谢 本文使用的瑞利散射激光雷达实测回波光子数廓线数据由中国科学院国家空间科学中心提供。

参考文献

- [1] DONG Zongge, DENG Kai, CAI Wang. Summary of laser radar technology development for atmospheric detection[J]. *Electro-Optic Technology Application*, 2022, **37**(6): 53-57 (董宗戈, 邓凯, 蔡旺. 大气探测激光雷达技术发展综述 [J]. *光电技术应用*, 2022, **37**(6): 53-57)
- [2] ZHANG Xianzhong, ZHONG Kai, WU Tong, et al. Progress of Rayleigh lidar environmental detection in near space[J]. *Aerospace Technology*, 2023(1): 19-42 (张献中, 钟凯, 吴同, 等. 瑞利激光雷达临近空间环境探测研究进展 [J]. *空天技术*, 2023(1): 19-42)

- [3] TARANTOLA A. Inverse Problem Theory and Methods for Model Parameter Estimation[M]. Philadelphia: Society for Industrial and Applied Mathematics, 2005
- [4] HAUCHECORNE A, CHANIN M L. Density and temperature profiles obtained by lidar between 35 and 70 km[J]. *Geophysical Research Letters*, 1980, **7**(8): 565-568
- [5] LEBLANC T, SICA R J, VAN GIJSEL J A E, *et al*. Proposed standardized definitions for vertical resolution and uncertainty in the NDACC lidar ozone and temperature algorithms-Part 3: Temperature uncertainty budget[J]. *Atmospheric Measurement Techniques*, 2016, **9**(8): 4079-4101
- [6] LI X Q, ZHONG K, ZHANG X Z, *et al*. Uncertainty evaluation on temperature detection of middle atmosphere by rayleigh lidar[J]. *Remote Sensing*, 2023, **15**(14): 3688
- [7] RODGERS C D. Inverse Methods for Atmospheric Sounding: Theory and Practice[M]. Singapore: World Scientific Publishing, 2000
- [8] WANG Yu, ZHANG Xianzhong, WU Tong, *et al*. Research on atmospheric temperature retrieval based on Rayleigh lidar using optimal estimation method[J]. *Chinese Journal of Space Science*, 2023, **43**(4): 627-639 (王煜, 张献中, 吴同, 等. 基于最优估计法的瑞利激光雷达反演大气温度研究[J]. *空间科学学报*, 2023, **43**(4): 627-639)
- [9] DEETER M N, EMMONS L K, FRANCIS G L, *et al*. Operational carbon monoxide retrieval algorithm and selected results for the MOPITT instrument[J]. *Journal of Geophysical Research: Atmospheres*, 2003, **108**(D14): 4399
- [10] GRANT J, GRAINGER R G, LAWRENCE B N, *et al*. Retrieval of mesospheric electron densities using an optimal estimation inverse method[J]. *Journal of Atmospheric and Solar-Terrestrial Physics*, 2004, **66**(5): 381-392
- [11] POVEY A. The Application of Optimal Estimation Retrieval to Lidar Observations[D]. Oxford: Oxford University, 2013
- [12] SICA R J, HAEFELE A. Retrieval of water vapor mixing ratio from a multiple channel Raman-scatter lidar using an optimal estimation method[J]. *Applied Optics*, 2016, **55**(4): 763-777
- [13] WANG Y. Research on Atmospheric Temperature Retrieval of Rayleigh Lidar Based on Optimal Estimation Method[D]. Tianjin: Tianjin University, 2022 (王煜. 基于最优估计法的瑞利激光雷达大气温度反演研究[D]. 天津: 天津大学, 2022)
- [14] BEDOYA-VELÁSQUEZ A E, CEOLATO R, LEFEBVRE S. Optimal estimation method applied on ceilometer aerosol retrievals[J]. *Atmospheric Environment*, 2021, **249**: 118243
- [15] LI S J. Research on Key Technologies of Atmospheric Lidar Dual Channel Detection[D]. Tianjin: Tianjin University, 2022(李世杰. 大气激光雷达双通道探测关键技术研究[D]. 天津: 天津大学, 2023)
- [16] JALALI A, SICA R J, HAEFELE A. A middle latitude Rayleigh-scatter lidar temperature climatology determined using an optimal estimation method[J]. *Atmospheric Measurement Techniques Discussions*, 2018, **11**(11): 6043-6058
- [17] MULAIRE W. Department of Defense World Geodetic System 1984: Its Definition and Relationship with Local Geodetic Systems[R]. St. Louis: National Imagery and Mapping Agency, 1984
- [18] JONES F E. The air density equation and the transfer of the mass unit[J]. *Journal of Research of the National Bureau of Standards*, 1978, **83**(5): 419-428
- [19] NICOLET M. On the molecular scattering in the terrestrial atmosphere: an empirical formula for its calculation in the homosphere[J]. *Planetary and Space Science*, 1984, **32**(11): 1467-1468

作者简介



张潇剑 男, 2000年6月出生于河北省邢台市, 现为天津大学精密仪器与光电子工程学院硕士研究生, 主要研究方向为瑞利激光雷达中层大气探测技术。
E-mail: zzxj@tju.edu.cn



钟凯 男, 1984年10月出生于山东省济南市, 天津大学精密仪器与光电子工程学院教授, 博士生导师, 主要研究方向为全固态激光器、非线性光学频率变换技术、太赫兹目标特性、激光雷达探测感知等。
E-mail: zhongkai@tju.edu.cn



徐德刚 (通信作者) 男, 1974年2月出生于山东省青岛市, 天津大学精密仪器与光电子工程学院教授, 博士生导师, 主要研究方向为激光与太赫兹光子学、激光雷达探测感知、太赫兹光谱成像技术以及海洋光纤传感技术等。
E-mail: xudegang@tju.edu.cn