

基于非线性最小二乘法的多光谱拟合程序：应用于 $^{12}\text{CH}_4$ 谱线参数分析

马宏亮^{1,2}, 郑健捷^{1,3,4}, 刘 强^{1,3*}, 钱仙妹^{1,3}, 朱文越^{1,3}

1. 中国科学院合肥物质科学研究院安徽光学精密机械研究所中国科学院大气光学重点实验室, 安徽 合肥 230031
2. 安庆师范大学电子工程与智能制造学院, 安徽 安庆 246133
3. 先进激光技术安徽省重点实验室, 安徽 合肥 230037
4. 中国科学技术大学研究生院科学岛分院, 安徽 合肥 230026

摘 要 精确的甲烷分子实验光谱参数在大气科学和天文探测等领域有着广泛的应用,特别是谱线的展宽系数及其温度依赖系数对于甲烷分子浓度廓线的研究尤为重要。精密的实验测量是获得准确谱线参数的重要手段。采用实验测量获取谱线参数时,需要在已知实验条件(浓度,温度,总压力,吸收光程以及气体分子种类的混合比等)的情况下,多次扫描同一波段范围得到多组实验室吸收光谱,然后利用基于非线性最小二乘法的拟合程序处理这些光谱,反演获得所需要的光谱参数。然而,一般常用的单光谱拟合程序处理实验光谱既费时又容易引起拟合过程中的误差传递。针对此问题,采用最小二乘拟合技术和 Levenberg-Marquardt 迭代算法编写了一款适用于处理由可调谐半导体激光吸收光谱技术(TDLAS)所获得的吸收光谱的多光谱拟合程序。该程序可同时处理多张实验光谱,并基于全局拟合方法获得一套光谱参数。详细介绍了该程序的原理、使用方法及数据处理过程。利用多光谱拟合程序中的 Voigt 线型处理了 $2\ 958\sim 2\ 959\ \text{cm}^{-1}$ 波数内甲烷($^{12}\text{CH}_4$)分子6条跃迁谱线的实验光谱,获得了296.0, 251.0, 223.0, 198.0和173.0 K共5组温度下 $^{12}\text{CH}_4$ 分子6条谱线的空气展宽系数。与之前文献报道的该波段内采用单光谱拟合程序得到的相应数据对比结果表明:获得的各温度下的空气展宽系数与参考文献中相应数据差值的百分比处在 $-4.97\%\sim 1.58\%$ 之间,两者数据整体符合较好,并且在30组对比数据中,有4组由单光谱拟合程序得到的空气展宽系数的误差值小于由多光谱拟合程序得到的相应数值,有2组数据显示由两种方法获得的误差值相等,其余24组由多光谱拟合程序获得的数据拟合误差小于由单光谱拟合程序获得的相应数值,表明多光谱拟合程序具有良好的可靠性,适用于气体分子吸收光谱的处理。

关键词 多光谱拟合程序; 甲烷; 吸收光谱; 谱线参数

中图分类号: O433.4 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2021)12-3887-05

引 言

准确测量大气中甲烷气体浓度的时空分布对于研究温室效应以及气候变化等具有十分重要的意义^[1-2]。鉴于此,人们发展了一系列的技术手段来实现这一目的,如基于星载或者地基的甲烷气体浓度廓线探测系统^[3]。由这些探测系统可以直接观测获得多组大气透射率光谱,然后,利用已知的甲烷气体分子谱线参数(谱线的位置、强度、碰撞展宽以及它们与温度的依赖关系等)可以获得所需要的浓度廓线信息。

为了准确获得这些谱线参数,需要在已知实验条件(浓度,温度,压强以及分子种类的混合比等)的情况下,多次扫描同一波段范围得到多组实验测量光谱,然后利用特定的程序拟合这些光谱反演获得。对于一次只能处理一张光谱的拟合程序称为单光谱拟合程序。采用单光谱拟合程序逐张拟合上述光谱,并把得到的数据进行平均或者线性拟合即可获得相应的谱线参数。然而,单光谱拟合程序处理实验光谱最大的缺点是非常耗时,同时,容易引起拟合过程中的误差传递^[4]。为克服以上问题,国际上多个研究小组发展出了多光谱拟合程序^[4-6]。然而,上述多光谱拟合程序都是针对分析

收稿日期: 2020-11-26, 修订日期: 2021-02-11

基金项目: 国家自然科学基金项目(41805014), 中国科学院大气光学重点实验室开放课题基金项目(JJ-19-01), 安徽省高校优秀青年人才支持计划项目(重点项目)(gxyqZD2020032), 先进激光技术安徽省实验室主任基金项目(20191002)资助

作者简介: 马宏亮, 1987年生, 安庆师范大学电子工程与智能制造学院副教授 e-mail: hlgma@foxmail.com

* 通讯作者 e-mail: liuq@aiofm.ac.cn

由傅里叶光谱仪得到的光谱开发的,目前很少有针对分析由 TDLAS 所获得的光谱而开发的此类程序。

本文利用最小二乘法开发了一款适用于处理 TDLAS 所测光谱的多光谱拟合程序,文中对该程序的功能做了详细的介绍;然后基于此程序分析了 $^{12}\text{CH}_4$ 分子在 $2\ 958\sim 2\ 959\ \text{cm}^{-1}$ 波段内多个不同温度下的吸收光谱,获得了 6 条谱线的空气展宽系数,并与之前采用单光谱拟合程序获得的相关数值作了比较,验证了程序的稳定性。

1 多光谱拟合程序

1.1 基于非线性最小二乘法的光谱拟合程序

非线性最小二乘拟合(NLSF)技术广泛用于实验室或实际大气测量光谱的分析工作^[5]。当处理给定的单张实验光谱时,非线性最小二乘方法拟合函数形式为

$$f(x_1, x_2, \dots, x_k) = \sum_{i=1}^N [S_{\nu_i}^m - S_{\nu_i}^c(x_1, x_2, \dots, x_k)]^2 \quad (1)$$

式(1)中, x_1, x_2, \dots, x_k 为需要拟合的参数,例如谱线位置、强度、压力展宽等, $S_{\nu_i}^m$ 为实验测量光谱在频率 ν_i 处的采样数值, $S_{\nu_i}^c(x_1, x_2, \dots, x_k)$ 为光谱在频率 ν_i 处的计算数值, N 为所分析光谱范围内的采样点数。拟合过程中通过迭代算法不断计算 $\frac{\partial f}{\partial x_i}$ 的数值,直到满足所给定的收敛条件为止,此时,会获得一组拟合参数 x_1, x_2, \dots, x_k 。基于此方法编写的单光谱拟合程序一次只能处理一组光谱。

如果同时处理多组给定的实验光谱时,只需对式(1)稍作修改即可获得多光谱非线性最小二乘拟合(NLSF)函数形式为

$$f(x_1, x_2, \dots, x_k) = \sum_{j=1}^M \sum_{i=1}^{N_j} [S_{\nu_i}^m - S_{\nu_i}^c(x_1, x_2, \dots, x_k)]^2 \quad (2)$$

式(2)中, M 为同时处理的光谱张数, N_j 为处理的第 J 张光谱所包含的采样点数, $S_{\nu_i}^m$ 为第 J 张光谱在频率 ν_i 处的实验测量值, $S_{\nu_i}^c(x_1, x_2, \dots, x_k)$ 为第 J 张光谱在频率 ν_i 处的计算值。在单光谱拟合方法的基础上,根据式(2)可以开发多光谱拟合程序。本文的多光谱拟合程序是基于 Fortran 语言编写的,该程序一次可同时完成多张气体分子光谱的拟合(也称为全局拟合),并输出一套光谱参数。有关多光谱拟合程序中涉及的公式推导过程可以参考文献^[5]的附录。

1.2 多光谱拟合程序的结构和功能

多光谱拟合程序的结构如图 1 所示,分为输入、主程序、输出三个功能不同的模块,下面分别就每个模块的功能和用法作相关介绍。

1.2.1 输入模块

如图 1 所示,输入模块一共包括 3 个部分。方框 1 包含需要拟合的实验光谱数据,一般为 txt 格式文件,每个文件对应一张实验光谱,里面包含两列 ASCII 原始实验数据,第一列为采样点的频率,第二列为采样点的信号强度(即待拟合数据),文件的第一行需要包括采样点数。拟合前需要给

每张光谱文件命名,把所有文件与执行程序放在一个文件夹里。方框 2 为待拟合光谱文件名列表,里面包含方框 1 中需要拟合的文件名称(即方框 1 中命名的文件)。方框 3 为拟合模型的原始输入参数,各参数的名称和意义如表 1 所示。

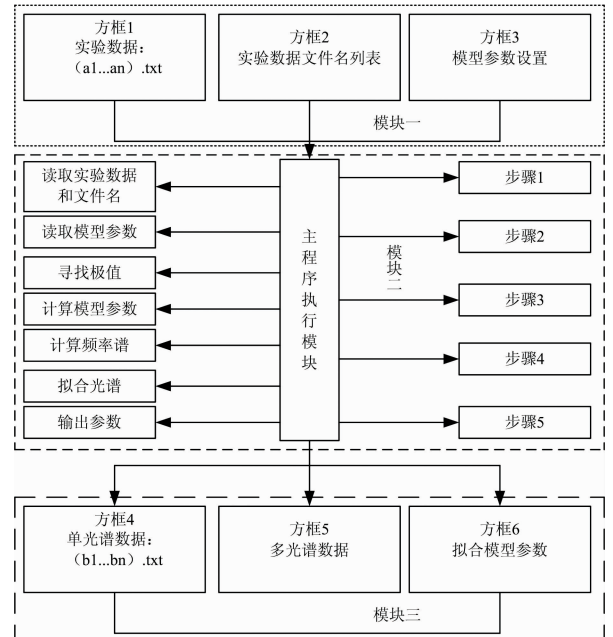


图 1 多光谱拟合程序流程图

Fig. 1 Simplified flow diagram of the multispectrum fitting program

表 1 模型需要输入的原始参数及其意义

Table 1 The initial parameter names and meanings of the model

输入参数	意义
Ω_{296}	296 K 下分子的总配分函数
P	实验压强
T	实验温度
Ω_T	实验温度下的配分函数
$\text{BASE}_0, \text{BASE}_1, \text{BASE}_2$	二次多项式作为光谱基线时,各次项的初始参数
IFLG	判断是否拟合该基线参数,0 表示固定,1 表示拟合
f	实验噪声水平
THE NUMBER OF TRANSITION	拟合的每张光谱中包含的跃迁谱线条数
ν_0	平移后的谱线中心位置初始值
γ_D	跃迁谱线的多普勒展宽半高半宽初始值
γ_P	跃迁谱线的压力展宽半高半宽初始值
A	跃迁谱线的面积初始值
S_H	数据库中相应谱线的强度
E	跃迁谱线的下能级值

1.2.2 主程序模块

如图 1 所示,左边为主程序的数据处理过程,其详细功

能如表 2 所示。

表 2 主程序数据处理过程及意义
Table 2 The data processing and significance of the main program

名称	意义
读取实验数据和文件名	将方框 1 中两列 ASCII 原始实验数据和方框 2 中与实验数据对应的 TXT 文件名称读取到主程序相应的内存中
读取模型参数	将方框 3 之前存储的拟合参数的初始值和实验条件读取到主程序中
寻找极值	搜索实验数据数组的最小值和最大值, 利用最小二乘拟合确定谱线中心位置
计算频率谱	利用当前参数计算频率谱
拟合光谱	对具有给定初始参数的数据执行最小二乘拟合, 并返回一组最佳拟合参数
输出参数	存储拟合之后的光谱数据及相应的谱线参数

如图 1 所示, 右边为主程序的执行过程, 开始拟合时, 程序会依次提示如表 3 所示的 5 步操作。表 3 给出了各操作步骤的名称和意义。

表 3 主程序执行步骤及各步骤意义
Table 3 The execution steps and significance of the main program

操作步骤名称	意义
HOW MANY SPECTRA DO YOU WANT TO SIMULATE	输入需要拟合几张光谱(即方框 1 中文件数量)
PLEASE INPUT THE ZERO FREQUENCY	输入谱线的中心频率平移值
ABSORPTION SPECTRUM: 2; TRANSMISSION SPECTRUM: 1	选择拟合光谱数据是发射谱还是吸收谱
SPECTRUM RECORDED AT THE SAME EXPERIMENTAL CONDITION? YES: 1; NO: 0	选择拟合光谱文件是否在同一实验条件下获得
DO YOU WANT TO PROCESS THE GLOBAL FITTING? YES: 1; NO: 0	是否需要进行全局拟合

1.2.3 输出模块

如图 1 所示, 输出模块共有 3 部分。方框 4 为方框 1 中单张光谱经过拟合之后相对应的输出数据。方框 5 为采用全局拟合并且归一化后得到的所有光谱的透过率谱图。方框 6 为对所有光谱拟合后得到的谱线参数, 如基线、谱线位置、多普勒展宽、压力展宽、吸收面积、拟合残差。

2 结果与讨论

本文利用多光谱拟合程序处理了 2 958~2 959 cm⁻¹ 波数内¹²CH₄ 分子 6 条跃迁谱线 5 组不同温度下的空气(Air)展宽实验光谱用以验证其性能。测量这些光谱的有关实验设备细节和详细的压强、温度等实验条件可以参考文献[7]。

图 2(a)为采用多光谱拟合程序同时对 4 张温度为 173 K, 压力为 18.158 kPa, Air+¹²CH₄ 的吸收光谱进行 Voigt 线型拟合所获得的全局拟合输出结果, 拟合过程中多普勒展宽固定在理论值, 图 2(b)为对应的拟合残差。由图 2(b)中的拟合残差可以看出多光谱拟合程序可以较好的完成多张光谱数据同时拟合的任务。

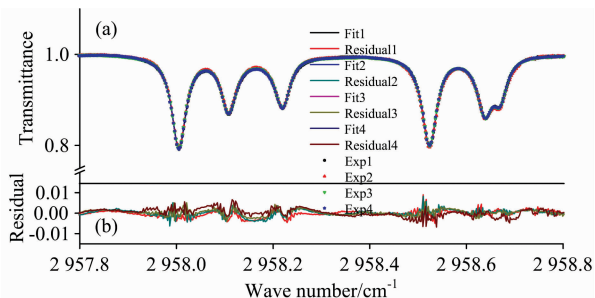


图 2 采用多光谱拟合程序同时拟合 4 张光谱
(a): 测量(散点线)和计算(实线)得到的光谱;
(b): 相应的拟合残差

Fig. 2 Example of multispectrum fitting program operation of with four spectra

(a): The measured (in scatter line) and calculated (in solid line) line profiles with a multi-peak fitting program; (b): The corresponding fitting residuals

采用多光谱拟合程序获得了 5 组温度下¹²CH₄ 分子 6 条谱线的空气展宽系数, 并与文献[7]中采用单光谱拟合程序处理获得的数据进行了对比, 结果如表 4 所示。表中 ν_0 表示跃迁谱线的中心位置, T 表示混合气体的温度, γ_{ref}^{air} 表示取自参考文献[7]的空气展宽系数(第 3 列), γ_{TW}^{air} 表示本文采用多光谱拟合程序得到的空气展宽系数(第 4 列), $\left(\frac{\gamma_{ref}^{air} - \gamma_{TW}^{air}}{\gamma_{TW}^{air}}\right)\%$ 表示本文获得的空气展宽系数与参考文献[7]中相应数据差值的百分比(第 5 列)。由表 4 第 5 列可知, 本文获得的空气展宽系数与参考文献中相应数据差值的百分比处在 -4.97%~1.58% 之间, 这表明采用多光谱拟合程序获得的数据与采用单光谱拟合程序获得的数据整体符合较好, 说明多光谱拟合程序具有良好的稳定性。此外, 需要注意的是, 表中第 3 列和第 4 列括号内的数字代表由数据处理过程所造成的误差, 其对应括号前面相应空气展宽系数的最后一位。由表 4 第 3 列和第 4 列括号内的数值对比可以看出, 在所获得的 30 组空气展宽系数中, 有 4 组由单光谱拟合程序得到的空气展宽系数的误差值小于由多光谱拟合程序得到的相应数值(表中加粗的数值); 有 2 组数据显示由两种方法获得的误差值相等(表中斜体数值); 对于剩余的 24 组数据来说, 由多光谱拟合程序得到的误差值都要小于由单光谱拟合程序得到的相应数值。这主要是由于采用单光谱拟合程序处理数据时, 需要对同一波段、相同实验条件下的多组实验测量光谱逐张处理, 并把得到的数据进行平均或者线性拟合来获得相应的空气展宽系数, 而每处理一张光谱都需要人为的选取一次基线, 并进行一次拟合, 在这一过程中产生的误差会传递给最终的空气展宽系数。然而, 采用多光谱拟合程序处理相应数

据时,多张光谱是同时处理的,并且它们拥有共同的一套拟合初始参数,因而,可以最大程度上避免这一影响。以上结

果表明,多光谱拟合程序能显著避免单次拟合所引起的误差传递。

表 4 多光谱拟合程序获得的 $^{12}\text{CH}_4$ 空气展宽系数及其与文献[7]相应数值的对比结果
Table 4 The air-broadening coefficients of $^{12}\text{CH}_4$ derived from the multispectrum fitting program and the difference between them and those from Ref. [7]

ν/cm^{-1}	T/K	空气展宽系数/ $(\text{cm}^{-1} \cdot \text{atm}^{-1})$		
		$\gamma_{\text{ref}}^{\text{air}}$	$\gamma_{\text{fit}}^{\text{air}}$	$(\frac{\gamma_{\text{ref}}^{\text{air}} - \gamma_{\text{fit}}^{\text{air}}}{\gamma_{\text{fit}}^{\text{air}}})/\%$
2 958.017 3	296.0±0.3	0.057 6(4)	0.057 7(3)	-0.17
	251.0±0.2	0.065 1(4)	0.065 3(3)	-0.31
	223.0±0.2	0.071 7(12)	0.071 9(3)	-0.28
	198.0±0.2	0.077 8(1)	0.078 2(4)	-0.51
	173.0±0.3	0.086 9(16)	0.088 4(8)	-1.7
2 958.120 0	296.0±0.3	0.064 9(3)	0.065 5(2)	-0.92
	251.0±0.2	0.075 2(3)	0.075 3(1)	-0.13
	223.0±0.2	0.083 6(3)	0.082 3(3)	1.58
	198.0±0.2	0.092 1(1)	0.092 1(2)	0
	173.0±0.3	0.104 2(6)	0.104 4(3)	-0.19
2 958.2329	296.0±0.3	0.064 1(10)	0.066 3(7)	-3.32
	251.0±0.2	0.074 0(3)	0.075 7(2)	-2.24
	223.0±0.2	0.082 1(7)	0.083 5(4)	-1.68
	198.0±0.2	0.089 7(9)	0.092 5(3)	-3.03
	173.0±0.3	0.100 2(16)	0.103 9(10)	-3.56
2 958.5364	296.0±0.3	0.060 1(7)	0.061 0(4)	-1.48
	251.0±0.2	0.069 3(5)	0.070 9(5)	-2.26
	223.0±0.2	0.076 1(10)	0.077 7(7)	-2.06
	198.0±0.2	0.083 2(2)	0.086 1(3)	-3.37
	173.0±0.3	0.092 2(12)	0.095 9(8)	-3.86
2 958.650 8	296.0±0.3	0.061 5(1)	0.063 6(6)	-3.30
	251.0±0.2	0.070 1(6)	0.071 0(5)	-1.27
	223.0±0.2	0.076 5(9)	0.080 5(8)	-4.97
	198.0±0.2	0.083 4(9)	0.085 2(8)	-2.11
	173.0±0.3	0.092 5(15)	0.097 1(14)	-4.74
2 958.682 5	296.0±0.3	0.055 8(18)	0.056 9(10)	-1.93
	251.0±0.2	0.063 9(4)	0.065 8(3)	-2.89
	223.0±0.2	0.070 0(6)	0.071 3(5)	-1.82
	198.0±0.2	0.077 4(8)	0.081 4(3)	-4.91
	173.0±0.3	0.086 4(21)	0.089 9(18)	-3.89

3 结 论

基于最小二乘法和 Levenberg-Marquardt 迭代算法开发了一款可以同时拟合多张光谱并获得一套光谱参数的多光谱拟合程序,该程序适用于处理由 TDLAS 所获得的吸收光谱。利用该程序处理了 2 958~2 959 cm^{-1} 波数内 $^{12}\text{CH}_4$ 分子 6 条跃迁谱线的实验光谱,获得了这些谱线相应的空气展宽

系数。通过比较发现,利用多光谱拟合程序获得的空气展宽系数与之前文献报道的结果整体符合较好,说明该程序具有较好的可靠性。与单光谱拟合程序相比,该程序可以极大的节省实验数据处理时间,并且可以减小数据处理过程中的误差传递。

致谢:作者感谢中国科学技术大学刘安雯副教授在多光谱拟合程序编写以及论文撰写过程中给予的指导。

References

- [1] TIAN Xing, CAO Yuan, WANG Jing-jing, et al (田 兴, 曹 渊, 王静静, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2019, 39(10): 3078.

- [2] Ma Hongliang, Sun Mingguo, Zhan Shenlong, et al. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2020, 226: 117605.
- [3] LU Xing-ji, CAO Zhen-song, HUANG Yin-bo, et al (卢兴吉, 曹振松, 黄印博, 等). *Optics and Precision Engineering(光学精密工程)*, 2018, 26(8): 1846.
- [4] Jacquemart D, Mandin J Y, Dana V, et al. *European Physical Journal D*, 2001, 14(1): 55.
- [5] Benner D C, Rinsland C P, Devi V M, et al. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 1995, 53(6): 705.
- [6] Lyulin O M. *Atmospheric and Oceanic Optics*, 2015, 28(6): 487.
- [7] Ma Hongliang, Liu Qiang, Cao Zhensong, et al. *Journal of Quantitative Spectroscopy & Radiative Transfer*, 2016, 171: 50.

A Multispectrum Fitting Program Based on Non-Linear Least-Squares Method for Line Parameters: Application to $^{12}\text{CH}_4$

MA Hong-liang^{1,2}, ZHENG Jian-jie^{1,3,4}, LIU Qiang^{1,3*}, QIAN Xian-mei^{1,3}, ZHU Wen-yue^{1,3}

1. Key Laboratory of Atmospheric Optics, Anhui Institute of Optics and Fine Mechanics, HFIPS, Chinese Academy of Sciences, Hefei 230031, China
2. School of Electrical Engineering and Intelligent Manufacturing, Anqing Normal University, Anqing 246133, China
3. Advanced Laser Technology Laboratory of Anhui Province, Hefei 230037, China
4. Science Island Branch of Graduate School, University of Science and Technology of China, Hefei 230026, China

Abstract The precise knowledge of the experimental spectroscopic line parameters of methane is widely used in atmospheric science and astronomical exploration, especially its collision broadening and temperature-dependent coefficients, which are important to study the methane molecular concentration profile. Precise experimental measurement is the significant method to obtain accurate spectral line parameters. In order to derive unknown parameters through a non-linear least-squares method, it is necessary to the simultaneous treatment of several experimental spectra in the same region were recorded with known experimental conditions (concentration, temperature, total pressure, absorption optical path and mixture ratio of gas molecule species, etc.). Then these spectra were usually fitted separately using spectrum-by-spectrum fitting procedure, but this method was often time-consuming and errors were easily produced. In order to solve this problem, a multispectrum fitting procedure based upon the non-linear least-squares and Levenberg-Marquardt method have been developed, able to treat simultaneously several laboratory spectra obtained by tunable diode laser absorption spectroscopy. This procedure can obtain a set of spectral parameters based on the global fitting method. The theory, data processing and application of the procedure are presented in detail. The laboratory spectra of 6 transition lines of $^{12}\text{CH}_4$ molecules in the $2\,958\sim 2\,959\text{ cm}^{-1}$ region are processed by using the Voigt profile of the multispectrum fitting program, and these lines air-broadening coefficients were obtained at five temperatures (296.0, 251.0, 223.0, 198.0, 173.0 K). The air-broadening coefficients of the present work are compared with the available values reported in the literature, which were obtained by spectrum-by-spectrum fitting procedure, and the difference between them lies within $-4.97\%\sim 1.58\%$. The results show that the broadening coefficients of these transitions in this work agree well with those available values. Among the 30 sets of correlation data, there are 4 sets whose fitting errors obtained by the spectrum-by-spectrum fitting program are smaller than those by the multispectrum fitting program, 2 sets with equal error values, and the remaining 24 sets whose fitting errors obtained by the multispectrum fitting program are smaller. Therefore, the multispectrum fitting program has good reliability and is suitable for processing gas molecular absorption spectrum.

Keywords Multispectrum fitting program; Methane; Absorption spectrum; Line parameter

(Received Nov. 26, 2020; accepted Feb. 11, 2021)

* Corresponding author