

· 光电测量与检测 ·

基于深度学习的嵌入式 CO₂ 浓度反演算法实现

王彪¹, 杨子腾², 卞广雨², 王冠懿², 赵奕飞³, 薛金波^{1,4}, 程林祥^{1,4}

¹中国科学院长春光学精密机械与物理研究所, 长春 130033;

²吉林大学, 长春 130015;

³吉林农业大学, 长春 130118;

⁴中国科学院大学, 北京 100049

摘要:以 CO₂ 为主的温室气体排放使得全球变暖, 严重影响生态环境, 2021 年习近平主席在二十国集团领导人峰会上提出“中国将力争 2030 年前实现碳达峰、2060 年前实现碳中和”, 因此精确检测 CO₂ 气体浓度具有重要意义。由于 CO₂ 气体吸收谱线的展宽受到气体压力、温度等因素影响, 导致 TDLAS 气体检测系统测量结果误差增大, 因此本文结合 HITRAN 数据库仿真, 提出了基于 BP 神经网络深度学习的 CO₂ 浓度反演算法和嵌入式实现方法, 实现了对气体浓度的补偿, 为嵌入式浓度反演算法设计提供理论依据。该算法可以移植到 STM32F407 中, 经过测试, 气体浓度的检测误差小, 有效提升了气体检测精度, 此方法同样适用于 TDLAS 型的其他气体检测应用场景中。

关键词: CO₂; BP 神经网络; 浓度反演; STM32F407

中图分类号: TN271.5 **文献标识码:** A **doi:** 10.14016/j.cnki.jgzz.2023.05.042

Implementation of embedded CO₂ concentration inversion algorithm based on deep learning

WANG Biao¹, YANG Ziteng², BIAN Guangyu², WANG Guanyi², ZHAO Yifei³,
XUE Jinbo^{1,4}, CHENG Linxiang^{1,4}

¹Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China;

²Jilin University, Changchun 130015, China;

³Jilin Agricultural University, Changchun 130118, China;

⁴University of Chinese Academy of Sciences, Beijing 100049, China

Abstract: The CO₂ Greenhouse gas emissions cause global warming and seriously affect the ecological environment. In 2021, The President Xi proposed at the G20 Summit that "China will strive to achieve carbon peak by 2030 and carbon neutrality by 2060". Therefore, accurate detection of CO₂ concentration has important research significance. Due to the influence of factors such as gas pressure and temperature on the broadening of CO₂ gas absorption spectral lines, the measurement error of TDLAS gas detection system increases. Therefore, this paper proposes a CO₂ concentration inversion algorithm and embedded implementation method based on deep learning of BP neural network, combined with HITRAN database simulation, to achieve compensation for gas concentration and provide a theoretical basis for the design of embedded concentration inversion algorithms. This algorithm can be ported to STM32F407. After testing, the detection error of gas concentration is small, effectively improving gas detection accuracy, and this method is also applicable to other gas detection application scenarios of TDLAS type.

Key words: CO₂; BP neural network; concentration inversion; STM32F407

收稿日期: 2023-02-13

基金项目: 吉林省科技发展计划重点科技研发项目 (No. 20220203016SF)、
中国科学院大学生创新实践训练计划项目 (No. 2022008090)

作者简介: 王彪 (1981-), 男, 博士, 副研究员, 博士生导师, 主要从事嵌入式系统软硬件方面研究。E-mail: wb5996@163.com

<http://www.laserjournal.cn>

1 引言

由于化石燃料的燃烧、煤炭和石油等的生产加工,大气中的二氧化碳含量正在逐年增加,引起温室效应、全球气候变暖、冰川融化等现象。因此,对二氧化碳含量的精确监测具有重要的研究意义。

当 CO₂ 气体检测系统用于大气环境监测时,由于地理位置、海拔的不同,环境压强、温度等因素会影响气体的检测精度。因此在检测二氧化碳浓度时,需要提升气体检测系统的环境适应性,降低环境压强、温度等因素对气体浓度反演结果的影响。

本文基于 TDLAS 技术对大气中的 CO₂ 浓度进行测量。TDLAS 技术是一种新型的大气痕量气体监测技术,近年来在气体检测领域得到了广泛应用。相较于电化学法、质谱法、红外吸收光谱等传统气体浓度检测方法,其具有高选择性、高分辨率、速度快、灵敏度高、通用性强、寿命长等优点^[1-2]。

TDLAS 技术测量气体浓度有直接吸收和波长调制法等方法。直接吸收法通过测量气体的吸收谱线直接得到气体的浓度,无需对浓度进行校准,操作简单,容易实施。但由于该方法吸收的光谱信号相比于激光器本身强度可忽略不计,导致其系统检测灵敏度较低,难以达到测量大气中气体浓度的标准;波长调制法测量信号为谐波信号,无法直接反应其中所蕴含的光谱信息,需经过复杂运算才能得到光谱信息。波长调制法不会使被测气体信号与激光器信号相作用,提高了系统的检测灵敏度。与直接吸收法相比,波长调制法复杂度高,抗干扰能力较强。考虑到系统环境的不确定性,本文采用波长调制法检测 CO₂ 气体浓度^[3]。本文采用 VCSEL 激光器作为光源,相比于 DFB 激光器,具有体积更小,效率更高的优点^[4]。

由于 TDLAS 技术测量的气体分子谱线展宽情况会受到环境温度和压强的影响,使测量结果误差增大,因此检测时必须对温度和压强产生的影响进行修正。由于影响是非线性的,传统的拟合方法不能达到很好的拟合效果。因此本文基于 BP 神经网络深度学习算法,结合 HITRAN 数据库,提出了 CO₂ 浓度反演算法:以温度、压强、二次谐波的幅值作为输入,二氧化碳浓度作为输出进行模型训练,实现温度压强补偿。经过测试,该网络模型可以移植到 STM32F407 中,有效提升了系统对环境的适应能力与气体检测精度,并且此方法同样适用于 TDLAS 型的其他气体检测应用场景中。

2 系统整体设计和理论基础

2.1 系统整体设计

本项目研制的基于深度学习 CO₂ 浓度反演算法

系统设计框图如图 1 所示。

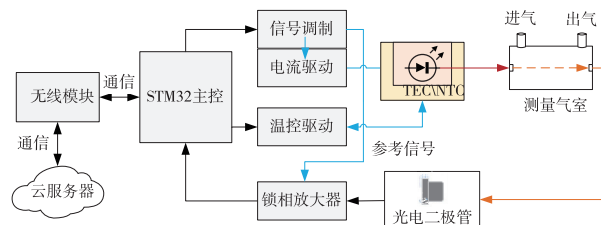


图 1 系统设计框图

系统分为光学部分和电学部分,光学部分实现 CO₂ 对激光的吸收,电学部分主要实现对于激光器的控制,气压温度信号的采集与处理以及无线信号的传输。系统主控通过电流驱动将锯齿波和正弦波叠加产生的调制信号以电流的形式作用在激光器上,产生的调制光束通过气室后被气体吸收^[5]。正交锁相放大技术能够克服环境噪声的干扰,从输出光束中提取特定二次谐波信号,最终计算得到二次谐波幅值。

由于神经网络算法能够移植到单片机中,在传感器采集到温度、压强数据后,结合所提取的二次谐波幅值输入到网络中,便可反演得到 CO₂ 浓度,实现温度压强补偿。

系统采用的主控 ARM 芯片为 STM32F407,该芯片具有丰富的通讯接口,能支撑多种传感器采集外部信息;STM32F407 型单片机主频为 168MHz,响应速度快,能够满足基于原位检测场景对实时性的要求;芯片具有浮点运算单元(Floating Point Unit, FPU),支持浮点数运算,可以对反馈回主控制器的信号进行快速处理,极大地提升了获取与运算数据的效率。

2.2 理论基础

系统基于 TDLAS-WMS 原理实现气体反演算法。TDLAS 技术测量的气体浓度遵从比尔-朗伯定律:

$$\frac{I_t}{I_0} = \exp[-\alpha(v)cl] \quad (1)$$

WMS 技术使用正弦调制方法,采用高频正弦波作为激光器驱动信号中的调制信号^[6],经过调制后信号频率可表示为:

$$v(t) = v_0 + A \cos(\omega t) \quad (2)$$

将(2)带入到(1)中得到:

$$I_t = I_0 \exp[1 - \alpha(v_0 + A \cos(\omega t))cl] \quad (3)$$

此时式(3)表示吸收后的光强信号。对(3)进行傅里叶级数展开,再对中心频率进行泰勒级数展开可以得到下式:

$$H(v_0) = -\frac{1}{4} (I_0 cl A^2) \frac{\partial^2 \alpha}{\partial v^2} \Big|_{v=v_0} \quad (4)$$

由此可知,待测气体浓度的表征方式是以二次谐波峰值(即吸收峰中心位置)与标准浓度值进行拟合

表示^[7],二次谐波的幅值正比于通过气体的光强衰减率,因此可根据二次谐波幅值对气体浓度进行推算。

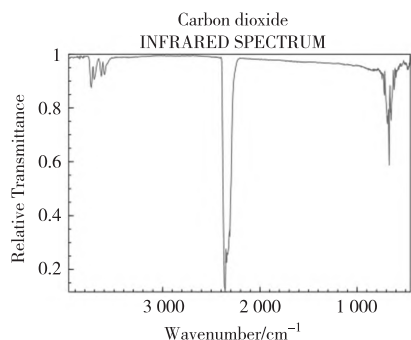


图2 二氧化碳红外吸收谱线

由于温度和压强会影响气体的吸收谱线,所以不能通过简单拟合来计算 CO₂ 浓度。从图 2 的红外光谱图中可以看出,描述气体红外吸收特性的两个重要参量是峰值吸收系数与全线宽,而在不同的气体压强与温度下,两个重要参量的值是不同的^[8],考虑到被测气体吸收谱线的全线宽受到气体压力、温度等因素影响,容易使测量的结果误差增大,在对 CO₂ 气体浓度进行检测时需要对其进行温度压强补偿。

本项目采用 BP 神经网络深度学习算法进行气体浓度反演,提高测量精度。BP 神经网络的全称为误差后向传播网络,具有映射能力强,泛化能力好等优点。通过网络内信号的正向传播及误差的反向传播,不断调整网络各层的权值和偏置,从而获得输入输出向量的映射关系,实现温度压力对气体浓度的补偿。本项目拟采用 3 层拓扑结构的 BP 神经网络模型,其中隐含层包含 16 个结点,网络模型如图 3 所示。

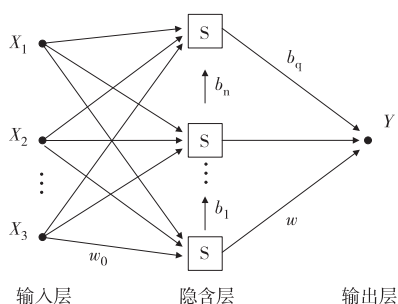


图3 BP神经网络结构

3 算法实现

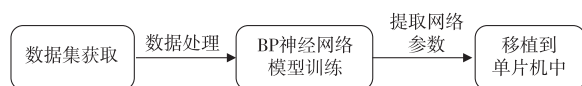


图4 程序整体实现框架图

浓度反演算法的整体设计方案实现流程如图 4 所示。首先从 HITRAN 数据库中提取二氧化碳相关数据,通过仿真程序计算 CO₂ 的二次谐波幅值,将数

据整理得到网络训练的数据集。之后将二氧化碳的温度、压强、二次谐波幅值作为网络输入,二氧化碳浓度作为输出,进行 BP 神经网络模型训练,最后将训练好的网络参数移植到单片机中,在单片机中编写前向传播算法,实现 CO₂ 气体浓度的反演计算。

3.1 数据集获取

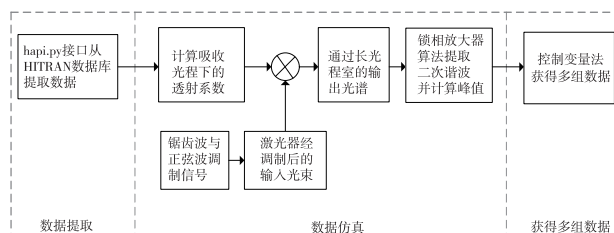


图5 数据集获取流程

数据集获取主要分为三个过程:数据提取、数据仿真及获得多组数据。其流程如图 5 所示,利用 HITRAN 数据库提供的 hapi.py 接口^[9]从数据库中提取需要波段的二氧化碳数据信息。在数据仿真中,根据提取的数据计算二氧化碳在光程气室下的吸收系数与透射系数。由于本项目采用波长调制法来测量气体浓度。首先调制信号为锯齿波与正弦波信号的叠加信号^[10],以电流的形式作用在激光器上产生调制光束。然后通过仿真电流调制信号,并结合光强调制系数得到输入光束,将输入光束与透射系数相乘得到穿过气室后的光谱。最后利用锁相放大算法,从二氧化碳输出光谱中提取出二次谐波并计算二次谐波的幅值。如图 6 与图 7 展示了当二氧化碳浓度为 10 000 g/m³,温度为 296 K,压强为 1 atm 时获得的透射系数与二次谐波曲线。

该仿真程序能根据不同的二氧化碳浓度、温度与压强计算二次谐波幅值。采用控制变量法分别修改温度、压强和气体浓度值。以上三者作为输入数据进行仿真得到相应的二次谐波信号幅值,进行数据整理,得到 BP 神经网络模型的数据集。

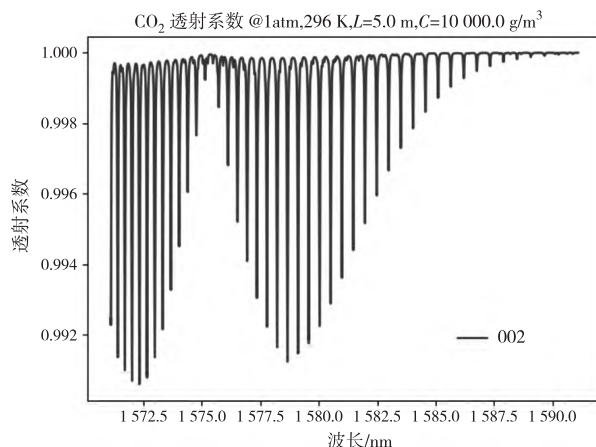


图6 296 K、1 atm 下的 10 000 g/m³ CO₂ 的透射系数

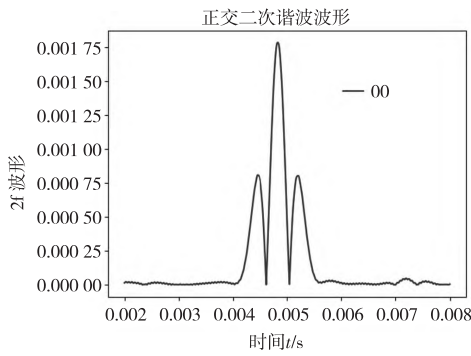


图7 296 K, 1 atm 下的 10 000 g/m³ CO₂ 的二次谐波曲线

3.2 BP 神经网络模型训练

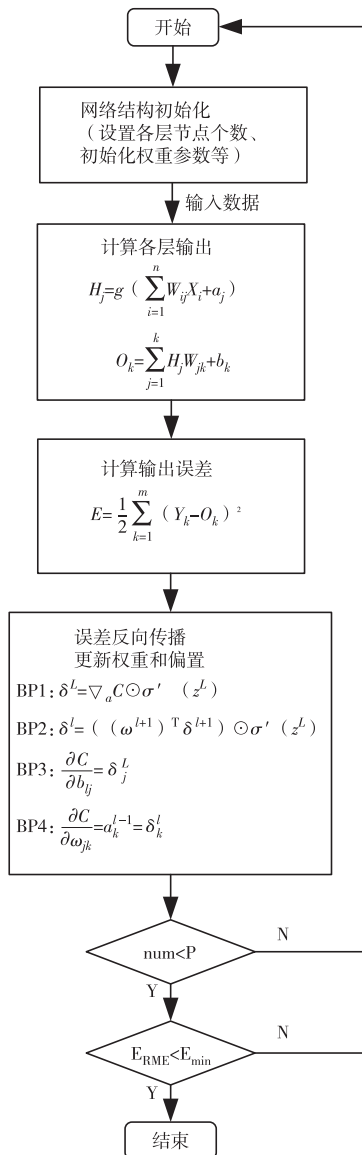


图8 BP 神经网络算法实现框图

BP 神经网络训练主要包括网络结构初始化、前向传播计算每层输出、计算输出误差、反向传播更新权值四个过程,整体流程框架如图 8 所示。在网络初始化中,首先设置各层的节点个数、初始化权值与偏

置等参数,然后将数据集输入网络进行前向传播,计算出每层的输出。其中隐含层输出 H_j 需要经过激活函数,输出层 O_k 不需要经过激活函数的计算。最后采用框图中的公式计算均方误差 (MSE),利用反向传播的四个公式调整权重与偏置。四个公式的解释如下^[11]:

BP1: 计算输出层误差 δ^l ;

BP2: 使用下一层误差 δ^{l+1} 表示当前层的误差 δ^l ;

BP3: 代价函数对网络模型偏置的变化率;

BP4: 代价函数对网络模型权重的变化率。

根据以上四个公式编写反向传播算法,经过不断调整网络权重和偏置,使系统误差不断减小。由于采用的是批量梯度下降法,需要判断执行的数据组数 num 是否达到总组数 P,若没有达到则需要继续训练。若系统的误差比期望误差小,则可提前结束训练。

3.3 算法移植

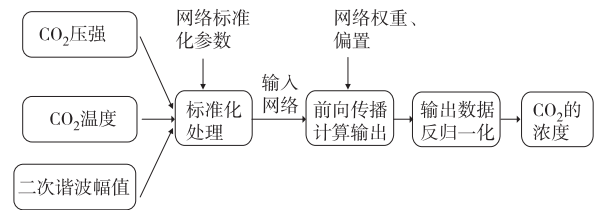


图9 算法移植流程图

算法移植的流程如图 9 所示:训练完成神经网络后,首先提取网络的输入输出标准化参数以及网络的权重与偏置^[12],然后编写 BP 神经网络前向传播过程计算输出数据,最后将数据反归一化后即可得到二氧化碳的浓度。由于程序采用 C 语言编写,并调用 DSP 库实现矩阵运算,所以程序可直接移植到 ARM 芯片中。

移植的重点在于神经网络前向传播算法的编写。根据 BP 神经网络的原理可知:

$$H_j = g\left(\sum_{i=1}^n W_{ij} X_i + a_j\right) \quad (5)$$

$$O_k = \sum_{j=1}^k H_j W_{jk} + b_k \quad (6)$$

其中 H_j 、 O_k 分别为隐含层与输出层的输出; W_{ij} 、 a_j 为输入层与隐含层间的权重; W_{jk} 、 b_k 为输出层与隐含层之间的权重; g 为激活函数。

从公式可以看出,每个隐含层结点的输出为为上一层输入的加权和在经过激活函数后的计算值,每个输出层节点的输出仅为为上一层输入的加权和。由于计算所需的权重和偏置均能从网络中提取,激活函数为我们预先设置的,因此在权重、偏置、激活函数已知的情况下,可通过 C 语言实现以上原理过程,模拟 BP

神经网络的前向传播。

4 测试结果

4.1 仿真结果

本文所研究的二氧化碳吸收峰在 1 581 nm 左右,因此从 HITRAN 数据库中提取波数在 6 285~6 365 之间的二氧化碳数据。通过控制变量法,设置二氧化碳的浓度区间为 100~20 000 g/m³,温度区间为 283~323 K,压强范围为 0.85~1.05 atm,共得到 2 205 组数据,将数据输入到仿真程序中可获得不同环境下的二次谐波幅值。表 1 展示了所获取的部分数据。经过数据处理后,该数据集可用作神经网络模型的训练。

表 1 获取数据部分展示

组号	压强(atm)	温度(K)	二次谐波幅值	浓度(g/m ³)
1	0.85	283	0.001 847 380 386 170	10 050
2	0.9	283	0.001 825 505 905 061	10 050
3	0.95	283	0.001 799 880 327 483	10 050
4	1	283	0.001 771 251 139 416	10 050
5	1.05	283	0.001 739 537 453 210	10 050
6	0.85	283	0.003 655 649 588 540	20 000
7	0.9	283	0.003 611 818 416 209	20 000
8	0.95	283	0.003 560 541 811 296	20 000
9	1	283	0.003 503 340 360 376	20 000
10	1.05	283	0.003 440 401 249 854	20 000

4.2 系统测试结果

该系统采用 3 层网络拓扑结构^[13],隐含层有 16 个结点。网络的输入数据为二氧化碳的温度、压强及二次谐波幅值,输出数据为二氧化碳浓度。将数据按 8:2 的比例划分训练集与验证集,并用验证集检验训练效果。网络的激活函数选择 sigmoid 函数,经过多次训练,发现训练次数为 500,学习率为 0.1 时,网络的拟合效果较好, R^2 可达到 0.999 78。模型测试结果如图 10 所示。

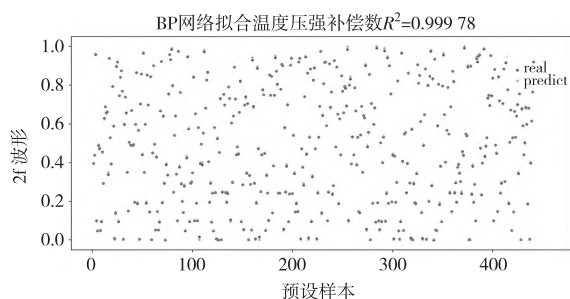


图 10 BP 神经网络拟合效果图

5 结论

本文结合 HITRAN 数据库进行了不同环境条件

下,不同 CO₂ 浓度的二次谐波信号仿真,并将仿真数据用于模型训练。通过采用 BP 神经网络深度学习算法实现了气体浓度的补偿,并将补偿算法移植到 STM32F407 中,相较于传统的拟合方法具有更高的准确性。经过系统测试,在不同环境条件下,气体浓度检测误差小,有效提升了 CO₂ 气体浓度检测精度和系统对不同环境的适应能力,此方法同样适用于 TDLAS 型的其他气体检测应用场景中。

参考文献

- [1] 马砾,范新丽,张硕,王伟峰,魏高明. 基于 TDLAS 技术的 CH₄ 气体检测与温度补偿方法[J]. 光谱学与光谱分析, 2021, 41(11): 3632-3638.
- [2] 杜保鲁,李萌,郭金家等. 基于 TDLAS 技术的海水中溶存 CO₂ 原位测量实验研究[J]. 光谱学与光谱分析, 2022, 42(4): 1264-1269.
- [3] 焦晓峰,孙鹏,管今哥等. 基于 TDLAS 的二氧化碳检测技术综述[J]. 测控技术, 2022, 41(9): 1-8.
- [4] 赵成龙,黄丹飞,刘智颖等. 开放型 TDLAS-WMS 技术 CO₂ 痕量气体检测[J]. 光子学报, 2022, 51(2): 333-342.
- [5] 许江文,王学影,常海涛,蔡静,梅园园,江文松. 基于 TDLAS 的 CO₂ 高温高压谱线参数测量研究[J]. 现代电子技术, 2023, 46(05): 104-108.
- [6] 云玉新,吕天光,韩洪,王泽众,姚金霞,李秀卫,赵笑笑. 气体红外吸收特性受压强与温度的影响分析[J]. 红外与激光工程, 2011, 40(06): 992-996.
- [7] 郁美霞,董刚,胥如迅,喻国丽. 基于 BP 神经网络的精密时基源校准预测模型[J/OL]. 中国测试: 1-7 [2023-04-20].
- [8] 李恒宽. 面向地震监测的 CO₂ 气体检测系统研制[D]. 长春:吉林大学, 2022.
- [9] 贾宇杰,张丕状,戚俊成等. 基于 HITRAN 光谱数据库的合并谱线测温仿真研究[J]. 测试技术学报, 2021, 35(03): 211-215.
- [10] 张可可,刘世萱,陈世哲等. 基于 TDLAS 的二氧化碳浓度检测系统及压强补偿研究[J]. 仪表技术与传感器, 2016, 396(01): 53-55+69.
- [11] Cynthia Eka Pandu, Ismanto Edi, Arifandy M. Imam, Sarbaini S, Nazaruddin N, Manuhutu Melda Agnes, Akbar Muhammad Ali, Abdiyanto. Convolutional Neural Network and Deep Learning Approach for Image Detection and Identification[J]. Journal of Physics: Conference Series, 2022, 2394(1): 291-298.
- [12] 赵振. COx 比例风险模型与 BP 神经网络在肝移植受体预后分析中的应用[D]. 成都:四川大学, 2006.
- [13] 陈远鸣,常建华,沈婉等. 基于改进型 BP 神经网络的 SF₆ 气体传感器[J]. 电子测量与仪器学报, 2017, 31(10): 1582-1588.